



А.П. Михайлов, А.П. Петров

**МАТЕМАТИЧЕСКИЕ
МОДЕЛИ СИСТЕМЫ
«ЧЕЛОВЕК-ОБЩЕСТВО»**



**МОСКВА
ФИЗМАТЛИТ®
2022**

УДК 519.6
ББК 22.19
М 69



Издание осуществлено при финансовой поддержке Российского фонда фундаментальных исследований по проекту 21-11-00005, не подлежит продаже

Михайлов А.П., Петров А.П. **Математические модели системы «человек–общество».** — М.: ФИЗМАТЛИТ, 2022. — 456 с. — ISBN 978-5-9221-1909-2.

В монографии изложены подходы к математическому моделированию и изучены конкретные модели процессов из области демографии, психологии, социологии, политологии. В качестве методологической основы принята сформулированная академиком А.А. Самарским триада «модель–алгоритм–программа», описывающая ключевые этапы моделирования. Особое внимание уделяется построению математических моделей на основании гипотез о поведении как индивидов, так и социальных групп, с учетом теоретических представлений и эмпирически полученных закономерностей, установленных в соответствующих отраслях научного знания. Для всех рассматриваемых моделей проведено исследование аналитическими или численными методами, с содержательной трактовкой математических результатов.

Книга предназначена научным работникам, преподавателям, аспирантам и студентам, интересующимся проблемами применения математических методов и моделей к социальным и гуманитарным наукам.

ISBN 978-5-9221-1909-2

© ФИЗМАТЛИТ, 2022

© А. П. Михайлов, А. П. Петров, 2022

ОГЛАВЛЕНИЕ

Введение	10
Литература	16
Глава 1. Некоторые традиционные математические методы соци- альных наук	17
1.1. Применение нелинейной регрессии в социологических исследованиях	17
1.1.1. Основные понятия регрессионного анализа	18
1.1.2. Функциональная спецификация нелинейной регрессии	22
Комментарии и библиографические примечания	29
Литература	29
1.2. Принцип максимального правдоподобия, пробит-анализ и логит-анализ	30
1.2.1. Принцип максимального правдоподобия	30
1.2.2. Пробит-анализ и логит-анализ	33
Комментарии и библиографические примечания	39
Литература	39
Приложение. Некоторые сведения из курса высшей математики и математической статистики	40
1.3. Кластеризация	43
1.3.1. Входные данные для кластерного анализа	43
1.3.2. Иерархическая кластеризация	45
1.3.3. Кластеризация в исследованиях парламентских голосований	48
1.3.4. Кластерный анализ в контексте социологического исследо- вания	49
Упражнения	51
Комментарии и библиографические примечания	51
Литература	52
1.4. Введение в теорию коллективного выбора	53
1.4.1. Вводные примеры	53
1.4.2. Формализация проблемы	55
1.4.3. Правила большинства и парадокс Кондорсе	59
1.4.4. Парадокс Кондорсе при бинарной процедуре	61
1.4.5. Правило Борда	64
1.4.6. Теорема Эрроу	66
Упражнения	68
Комментарии и библиографические примечания	69
Литература	69

1.5. Введение в теорию индексов влияния	69
1.5.1. Индекс Банцафа	70
1.5.2. Взвешенное голосование	73
1.5.3. Парадокс перераспределения депутатов между фракциями	75
1.5.4. Учет коалиционных предпочтений и парадокс вражды	77
Упражнения	80
Комментарии и библиографические примечания	80
Литература	82
Глава 2. Демография	83
2.1. Базовые макромоделли демографии	83
2.1.1. Модель Мальтуса	83
2.1.2. Логистическая модель	85
Упражнения	89
Комментарии и библиографические примечания	90
Литература	95
2.2. Модель Лотки	95
2.2.1. Построение модели	96
2.2.2. Случай мальтузианских населений	99
2.2.3. Анализ устойчивости мальтузианского режима	102
Упражнения	104
Комментарии и библиографические примечания	104
Литература	104
2.3. Модель Лесли	105
2.3.1. Построение модели	105
2.3.2. Уравнение Эйлера–Лотки	108
2.3.3. Возрастно-половая пирамида и демографические волны	111
2.3.4. Учет продолжительного пребывания индивидов в старшей группе	115
2.3.5. Об условиях применимости модели Лесли	116
Упражнения	121
Комментарии и библиографические примечания	121
Литература	124
2.4. Учет половой структуры: бипарентальные модели	124
2.4.1. Модель, не учитывающая брачность	125
2.4.2. Модель, учитывающая брачность	132
Упражнения	140
Комментарии и библиографические примечания	140
Литература	141
2.5. Пример имитационно-математического моделирования в демографии	141
2.5.1. Общее описание имитационной модели	142
2.5.2. Моделирование рождаемости	144

2.5.3. Результаты вычислительных экспериментов и построение демографического прогноза для России	146
2.5.4. Основные требования к имитационной модели	148
Комментарии и библиографические указания	149
Литература	149
Глава 3. Психология	151
3.1. Метод парных сравнений и описание процесса восприятия	151
3.1.1. Закон сравнительных суждений	152
3.1.2. Метод парных сравнений	155
3.1.3. Моделирование когнитивной системы индивида	159
Упражнения	162
Комментарии и библиографические примечания	163
Литература	164
3.2. Простейшая модель подражательного поведения	165
3.2.1. Нейрологическая схема	166
3.2.2. Уравнения модели на микроуровне	168
3.2.3. Учет внешних и внутренних стимулов	173
3.2.4. Базовая математическая модель	174
3.2.5. Стационарные состояния базовой модели	176
3.2.6. Численный эксперимент с нестационарной моделью	179
Упражнения	182
Комментарии и библиографические замечания	182
Литература	183
3.3. Расширенная модель подражательного поведения	183
3.3.1. Понятие социально-политического потенциала (СПП)	184
3.3.2. Основные предположения	186
3.3.3. Построение модели	188
3.3.4. Конкретизация модели	191
3.3.5. Циклический режим эволюции электората	193
Упражнения	196
Комментарии и библиографические примечания	196
Литература	196
3.4. Формирование мнений и установок индивидов, составляющих замкнутую группу	197
3.4.1. Построение и анализ модели формирования мнений	197
3.4.2. Консенсус для некоторых типов групп индивидов	200
3.4.3. Модель формирования установки действия	203
Упражнения	206
Комментарии и библиографические примечания	206
Литература	207
3.5. Социальные взаимодействия в малой группе	208
3.5.1. Традиционная модель	208
3.5.2. Анализ традиционной модели и выявление ее проблемных мест	210

3.5.3. Построение и анализ скорректированной модели	213
3.5.4. Учет структуры группы	216
Упражнения	217
Комментарии и библиографические примечания	217
Литература	218
3.6. Структурный баланс в малой группе	218
3.6.1. ХКХ-подход к структурному балансу	219
3.6.2. Д-подход к структурному балансу	223
3.6.3. Социоматрицы	224
Упражнения	225
Комментарии и библиографические примечания	225
Литература	226
3.7. Модель динамики отношений индивидов в малой группе	226
3.7.1. Основные предположения модели	227
3.7.2. Анализ симметричных стационарных состояний для группы из трех индивидов	229
Упражнения	235
Комментарии и библиографические примечания	235
Литература	236
3.8. Динамическая модель близости позиций пользователей социальных сетей	236
3.8.1. Построение модели	237
3.8.2. Стационарное состояние согласия в случае двух индивидов	242
3.8.3. Стационарное состояние с несимметричными отношениями	243
3.8.4. Вычислительные эксперименты с моделью	245
Упражнения	249
Комментарии и библиографические примечания	249
Литература	250
Глава 4. Социология	251
4.1. Латентно-структурный анализ	251
4.1.1. Основные положения модели латентных классов	251
4.1.2. Отнесение отдельных индивидов к конкретным латентным классам	255
4.1.3. Пример применения модели латентных классов в социологи- ческом исследовании	257
4.1.4. Латентно-структурный анализ в случае непрерывной латент- ной переменной	258
4.1.5. Заключение	262
Упражнения	262
Комментарии и библиографические примечания	262
Литература	263
4.2. Модели марковского типа	263
4.2.1. Динамика численности верующих	264
4.2.2. Элементы общей теории марковских процессов	268
4.2.3. Вертикальная социальная мобильность	271

4.2.4. Модели типа мувер-стейер	275
4.2.5. Расовая динамика	277
Упражнения	281
Комментарии и библиографические примечания	281
Литература	284
4.3. Социальная сегрегация	285
4.3.1. Качественная модель	286
4.3.2. Конкретизация поведенческой гипотезы и построение математической модели	292
4.3.3. Анализ математической модели	294
4.3.4. Заключение	298
Упражнения	299
Комментарии и библиографические примечания	299
Литература	300
4.4. Поведение индивидов при изменении количества степеней свободы	300
4.4.1. Введение: содержательная проблема теста гипотетического монополиста и понятие о степени свободы	301
4.4.2. Основные определения	303
4.4.3. Оценка товарных границ рынка	304
4.4.4. Одновременное определение товарных и географических границ рынка	307
4.4.5. Заключение	312
Упражнения	313
Комментарии и библиографические примечания	313
Литература	313
4.5. Распространение информации в социуме в отсутствие противодействия	314
4.5.1. Базовая модель информационного нападения	314
4.5.2. Учет дополнительных факторов	316
4.5.3. Модель информационного нападения, учитывающая три дополнительных фактора	320
4.5.4. Заключение	324
Упражнения	325
Комментарии и библиографические примечания	325
Литература	329
4.6. Информационное противоборство в социуме	331
4.6.1. Базовая модель информационного противоборства	332
4.6.2. Условие победы при противоборстве произвольного количества партий	334
4.6.3. Учет дополнительных факторов	335
4.6.4. Информационное противоборство при дестабилизирующем воздействии	337
4.6.5. Заключение	342
Упражнения	342
Комментарии и библиографические примечания	343
Литература	343

Глава 5. Политология	344
5.1. Политическое соперничество	344
5.1.1. Постановка задачи и формулировка модели	344
5.1.2. Равновесные уровни власти и их устойчивость	346
5.1.3. Некоторые вычислительные эксперименты с моделью	351
Упражнения	355
Комментарии и библиографические примечания	355
Литература	355
5.2. Моделирование этнополитической самоидентификации	356
5.2.1. Политическая самоидентификация дагестанского этноса в конце 1990-х годов	356
5.2.2. Моделирование динамики взаимодействующих электоратов	361
5.2.3. Динамика взаимодействия электоратов вблизи положения равновесия	367
5.2.4. Интерпретация результатов моделирования взаимодействующих электоратов	371
Упражнения	372
Комментарии и библиографические примечания	372
Литература	372
5.3. Основы моделирования системы «власть–общество»	373
5.3.1. Построение модели	374
5.3.2. Динамика в рамках правовой системы: базовая модель в ли- нейном случае	379
5.3.3. Анархическая и тоталитарная ловушки	382
5.3.4. Заключение	384
Упражнения	385
Комментарии и библиографические примечания	385
Литература	386
5.4. Система «власть–общество» в случае биполярной реакции общества	386
5.4.1. Биполярная реакция общества	387
5.4.2. Формирование структур с внутренними переходными слоями	388
5.4.3. Стационарные контрастные распределения власти	389
5.4.4. Нестационарные контрастные распределения власти	391
5.4.5. Конфедеративные распределения власти	393
Упражнения	394
Комментарии и библиографические примечания	395
Литература	395
5.5. Коррупция и другие несовершенства властных иерархий	396
5.5.1. Учет коррупции в модели «власть–общество»	396
5.5.2. Ущерб от коррупции, стоимость и эффективность ее ограни- чения	400
5.5.3. Сравнительный анализ стратегий ограничения коррупции	402
5.5.4. Учет некомпетентности иерархии	405
5.5.5. Клановые отношения в иерархии	406
Упражнения	408
Комментарии и библиографические примечания	409
Литература	409

5.6. Динамика гонки вооружений	410
5.6.1. Построение базовой модели	411
5.6.2. Анализ базовой модели с постоянными коэффициентами . . .	412
5.6.3. Нелинейная модель	415
5.6.4. Обобщение на случай большего количества стран	420
Упражнения	422
Комментарии и библиографические примечания	422
Литература	422
5.7. Моделирование крупномасштабных исторических процессов	423
5.7.1. Введение	423
5.7.2. Элементы модели мировой динамики	424
5.7.3. Динамика аграрных обществ	425
5.7.4. Моделирование системы древнегреческих полисов	430
Упражнения	432
Комментарии и библиографические примечания	433
Литература	434
Список литературы	436
О авторах	451

Посвящается научной школе
Андрея Николаевича Тихонова —
Александра Андреевича Самарского —
Сергея Павловича Курдюмова

Введение

Использование и развитие математических методов для изучения общественно-политических процессов — актуальное направление прогресса современной науки, позволяющее существенно обогатить традиционные методы анализа, прогноза и принятия решений в данной сфере.

Историю этого направления можно отсчитывать от работы У. Петти [1] (написанной примерно в 1676 г. и опубликованной в 1690 г.), Н. де Кондорсе [2] (опубликовано в 1785 г.) и Т. Мальтуса [3] (опубликовано анонимно в 1798 г.). Позже Д. И. Менделеев в своей обширной монографии [4] анализировал функцию, аппроксимирующую возрастную структуру населения, а с наступлением эпохи социологических опросов математическая статистика превратилась в стандартный инструмент социологов, экономистов и политологов. В целом к концу первой трети XX века математические методы стали широко применяться в науках о человеке и обществе.

В современную эпоху можно говорить уже не только и не столько о математических методах (включающих в себя в первую очередь статистику и задачи оптимизации), сколько о математических моделях в гуманитарных и социальных науках. Это стало возможным благодаря не только прогрессу собственно математики и появлению современных вычислительных средств, но и развитию методологии математического моделирования [5–10].

Ее сущность состоит в замене исходного объекта его «образом», математической моделью и дальнейшем изучении модели с помощью реализуемых на компьютерах вычислительно-логических алгоритмов. Этот «третий метод» познания, конструирования, проектирования сочетает в себе многие достоинства как теории, так и эксперимента. Работа не с самим объектом (явлением, процессом), а с его моделью дает возможность безболезненно, относительно быстро и без существенных затрат исследовать его свойства и поведение в любых мыслимых ситуациях (преимущества теории). В то же время вычислительные (компьютерные, симуляционные, имитационные) эксперименты с моделями объектов позволяют, опираясь на мощь современных вычислительных методов и технических инструментов информатики, подробно и глубоко изучать объекты в достаточной полноте, недоступной чисто теоретическим подходам (преимущества эксперимента). Данный подход кратко выражается триадой «модель–алгоритм–программа». При этом особо подчеркивается роль математического моделирования как инструмента

переработки информации как таковой, т. е. «информационного сырья», в готовый продукт, т. е. в точное знание, на основании которого могут приниматься управленческие решения и осуществляться контроль за их выполнением.

Старт этой методологии и ее дальнейшее бурное развитие пришлось на конец 40-х–начало 50-х годов прошлого столетия, когда СССР был вынужден в кратчайшие сроки лишить США монополии на атомное и термоядерное оружие и средства его доставки, чтобы в дальнейшем достичь паритета между двумя сверхдержавами. Отечественная прикладная математика, коллективы, возглавляемые академиками М. В. Келдышем, А. Н. Тихоновым, А. А. Дородницыным, Г. И. Марчуком, Н. Н. Моисеевым, О. М. Белоцерковским и другими выдающимися учеными, справлялись с возникавшими перед ней проблемами зачастую быстрее и точнее, чем заметно лучше материально обеспеченные заокеанские конкуренты.

Если же говорить о методологии математического моделирования, ее развитии, теоретической и практической институционализации, то наиболее значимый вклад в этот процесс внес академик А. А. Самарский [5–9], сформулировавший представление о математическом моделировании как *методологическом императиве* современной науки, а также его научная школа. При этом характерной чертой этой научной школы всегда было стремление расширить рамки применения методологии математического моделирования, в т. ч. в направлении исследования системы «человек–общество».

Актуальность данного направления определяется в первую очередь тем фактом, что в современном мире возрастает роль гуманитарных взаимодействий, когда люди пытаются больше и осмысленней участвовать в разработке принимаемых государством и обществом решений, основанных на доверии, честности и сотрудничестве. Усложнение социальной реальности выражается в одновременном протекании разнонаправленных процессов, таких как глобализация и борьба за сохранение национального суверенитета, стремление к устойчивому развитию и спонтанное возникновение «турбулентности» в международных отношениях. Эти усложнения создают новые вызовы наукам о человеке и обществе, что требует разработки новых подходов, соответствующих сложности рассматриваемых явлений и процессов. Одним из естественных ответов является развитие и применение методологии математического моделирования к данной сфере.

В настоящее время данная методология прилагается к исследованию широкого круга проблем, в т. ч. к исследованию общества и власти. Первым в нашей стране значительным изданием, посвященным математическим методам и моделированию в гуманитарных науках, стала монография Г. В. Осипова, Э. П. Андреева «Методы измерения в социологии» (издательство «Наука»), появившаяся в 1977 г., когда социология еще рассматривалась в СССР всего лишь как сугубо прикладная научная дисциплина.

С течением времени сформировался ряд направлений, представленных такими социологами, политологами и математиками, как А. И. Антонов, В. М. Карпова, Г. Б. Прончев (Социологический факультет МГУ им. М. В. Ломоносова), В. А. Шведовский (Высшая школа современных социальных наук МГУ), В. Ж. Келле (Институт философии РАН), Г. Г. Татарова (Институт социологии РАН), Ю. Н. Толстова, А. С. Ахременко (Высшая школа экономики), Ю. Н. Гаврилец (Центральный экономико-математический институт РАН), Д. А. Новиков, А. Г. Чхартишвили и Ф. Т. Алескеров (Институт проблем управления РАН), Г. Г. Малинецкий, М. Е. Степанцов (Институт прикладной математики им. М. В. Келдыша РАН), С. Ю. Малков (Академия военных наук), М. Г. Дмитриев, В. К. Финн и М. А. Михеенкова (ФИЦ ИУ РАН), М. А. Александров (РАНХиГС), О. Г. Прончева (МФТИ) и другими исследователями.

Их результаты публикуются, в частности, в ежегодном издании «Математическое моделирование социальных процессов», полностью посвященном данной тематике (к настоящему времени вышло 20 выпусков сборника, изданных в МГУ им. М. В. Ломоносова и ИПМ им. М. В. Келдыша РАН под ред. А. П. Михайлова). Работы, использующие математические модели и методы, регулярно появляются в таких журналах, как «Социология: методология, методы, математическое моделирование» (Институт социологии РАН, выходит с 1991 г. под ред. Г. Г. Татаровой, затем под ред. И. Ф. Девятко) и «Информационные войны» (издается Центром исследований проблем безопасности РАН и Академией военных наук с 2007 г., гл. ред. В. Л. Шульц), а также в ряде других изданий. При этом приходится признать, что российская практика не вполне успевает за мировыми тенденциями, в частности, в области публикационной активности.

Так, за рубежом давно издаются научные журналы, полностью посвященные данной тематике, например с 1971 г. выходит *The Journal of Mathematical Sociology* (издательство Taylor and Francis), с 1981 г. — журнал *Mathematical Social Sciences* (издательство Elsevier). Активно действуют профильные объединения ученых: так, Секция математической социологии Американской социологической ассоциации не только организует научные исследования, но и участвует в образовательной деятельности, в частности выделяет студентам гранты на выполнение дипломных работ в области математической социологии.

В дальнейшем роль математических методов будет только возрастать, причем по той же причине, что и роль всей социологической и политической науки — ведь современное общество не только усложняется, но и пытается осознать свою возрастающую сложность. Все более востребованными становятся, например, научно обоснованные оценки последствий информационных процессов, сопутствующих им политических, экономических и социальных решений. Мировая практика свидетельствует, что исследования в данных областях проводятся с возрастающей ролью математических методов и моделей.

Их актуальность для практической деятельности, в т. ч. по отражению информационных угроз, связана с необходимостью интеллектуальной поддержки принятия решений: в частности, опирающаяся на соответствующие комплексы математических моделей, в перспективе — на информационно-моделирующие системы поддержки принятия решений.

Данной тенденции соответствует позиция как высшего руководства нашей страны, так и научно-педагогической общественности относительно значимости информационно-математических технологий в инновационном развитии глобального информационного общества.

При построении математических моделей в естественных и технических науках используются фундаментальные законы природы, вариационные принципы, различного рода аналогии. Нередко бывает полезным также применение иерархических и феноменологических подходов. Применительно к гуманитарным наукам использование этого богатого исследовательского арсенала сопряжено с принципиальными трудностями. Существенное присутствие «человеческого фактора», сознательное участие в социальных процессах как отдельных людей, так и целых социальных общностей порождает ряд фундаментальных проблем, препятствующих созданию отвечающих реальности математических моделей в этой сфере. К ним относятся следующие.

1. Отсутствие так называемых *фундаментальных законов*, т. е. твердо установленных и допускающих точное количественное выражение универсальных связей между изучаемыми величинами (ср. с различием между правилами и законами в науке, введенным немецким и американским психологом К. Левинем [11]).

2. Плохая «измеримость» социальных показателей. Даже если понятно, что необходимо измерять, то требуются немалые интеллектуальные ухищрения для приближения к истине с приемлемой точностью. Ведь социальные объекты не слишком благосклонны к открытости.

3. Имеет место сильная разномасштабность (от малых социальных групп до этносов и суперэтносов), разнородность и нелинейность изучаемых социальных объектов.

4. Постоянное усложнение социальных объектов, их «рефлексивность», означающая, в частности, то, что новые знания становятся рано или поздно частью объектов, изменяют их свойства и характеристики. Отсюда постоянное развитие математических моделей, их усложнение в «погоне» за изучаемым объектом.

К этим и другим трудностям добавляется специфическая для гуманитарных наук зависимость от власти, от политической конъюнктуры, от тех или иных ожиданий общества. В результате в науке между «двумя культурами» (выражение британского писателя, физикохимика и государственного деятеля Ч. П. Сноу [12]) создается труднопреодолимый барьер, мешающий их взаимному проникновению и пониманию. Проблема взаимоотношений между математическим моделированием в существенно различных областях науки многоаспектна и требует

отдельного обсуждения. В то же время применительно к методологии математического моделирования ввиду ее универсальности можно указать несколько «мостиков» между «двумя культурами», построение которых способно служить первым шагом в преодолении взаимного отчуждения, в их взаимообогащении.

В случае анализа социальных и политических процессов стандартные требования к построению и изучению математических моделей необходимо дополнить (по меньшей мере) следующими принципами, образующими «*Кодекс математика*», исследующего систему «Человек–Общество».

1. Не покушаться на твердо установленные факты и закономерности, избегать соблазна научных «революций», прикрытых математическим «флером».

2. Не подменять собой «предметников» — социологов, демографов, политологов (или, во всяком случае, ясно очерчивать границу между применяемой исследователем математической методологией и его собственными воззрениями на изучаемый предмет).

3. По возможности придерживаться здравого позитивизма, трезво оценивая, какие из социальных и политических процессов можно моделировать с достаточной степенью достоверности, а для каких еще преждевременно говорить об исследовании с необходимой полнотой.

4. Разделять исследования фундаментально-ориентированные (направленные на выяснение общих свойств объектов и в то же время связанные с крупной реальной проблемой) и конкретно-прикладные (опирающиеся на полученные ранее фундаментальные знания и дающие количественные характеристики).

Важнейшую роль при этом играет анализ поведенческих характеристик вовлеченных в процесс акторов. Как правило, при построении конкретных моделей формулируются поведенческие гипотезы, основанные на содержательных теориях и/или эмпирически установленных закономерностях. Вычислительные эксперименты с различными поведенческими гипотезами позволяют установить наиболее состоятельную из них. Данная книга последовательно проводит указанную логику: от поведенческой гипотезы — к построению модели и изучению ее свойств. Помимо вычислительного эксперимента, для анализа моделей используются хорошо разработанные методы теории дифференциальных и интегро-дифференциальных уравнений (в частности, качественные и асимптотические методы), линейной алгебры, теории марковских процессов и т. д. Таким образом, последовательность действий при моделировании включает в себя общую постановку проблемы, построение концептуальной модели, определение участвующих акторов и связей между ними, введение основных переменных и параметров, характеризующих изучаемый объект, вывод соответствующих математических соотношений, их анализ вычислительными и аналитическими методами, содержательную трактовку результатов и разработку рекомендаций на основе проведенного моделирования. В случае необходимости

может быть проведено уточнение всех компонентов моделирования для получения более углубленного и точного знания об изучаемом объекте.

При работе над книгой авторы выбирали модели, отражающие проблематику системы «человек–общество» и основанные на содержательных представлениях о поведении индивидов. При этом в нее сознательно не включена такая крупная область, как моделирование в экономике. Это было сделано по двум причинам, первая из которых является тривиальной — уже существует обширная литература, посвященная этому направлению. Вторая причина состоит в том, что в экономике поведение описывается на основании представления о *Homo Economicus* — рациональном субъекте, принимающем решения на основе принципа максимизации целевой функции. Эта поведенческая гипотеза имеет довольно ограниченную перспективу за пределами экономики, так как она оставляет в стороне принятие решений на основании когнитивного диссонанса, подражания, эмоций и многих других факторов, значимых в социальной и политической жизни.

Что касается стиля книги, то мы стремились избегать громоздких и строгих процедур (которые заинтересовавшийся читатель может найти в специальных изданиях), а обращали основное внимание на описание идей и соответствующих им примеров. Поэтому книга содержит большое количество иллюстраций и упражнений, ее материал без особых затруднений может быть использован для формирования учебных курсов по различным конкретным направлениям математического моделирования в данной сфере.

Часть материала книги апробирована в лекционных курсах, которые читались авторам в течение многих лет на факультете управления и прикладной математики МФТИ и на трех факультетах МГУ — вычислительной математики и кибернетики, социологическом и факультете глобальных процессов.

Материал состоит из пяти глав, в первой из которых представлен ряд предварительных примеров применения традиционных математических методов в социальных науках. Остальные четыре главы посвящены моделированию в демографии, психологии, социологии и политологии: в общей сложности здесь 26 разделов, каждый из которых посвящен одной конкретной модели (иногда — нескольким близким моделям) и содержит краткое введение, а также набор упражнений для самостоятельной работы. Комментарии, библиографические примечания и список литературы приведены в конце каждого раздела единым блоком; к этому списку относятся номера ссылок на литературу в данном пункте. Формулы, таблицы и рисунки нумерованы по пунктам. Таким образом, книга допускает как чтение подряд, с начала до конца, так и по тематическим главам и отдельным разделам.

Подбор материала и манера его изложения отвечают концепциям отечественной школы математического моделирования, всегда находившейся на мировом уровне и широко трактующей данную методологию.

Авторы считают своим приятным долгом поблагодарить академика РАН Б.Н. Четверушкина, академика РАН Ю.С. Осипова, академика РАН В.А. Садовниченко, академика РАН Е.И. Моисеева, д.ф.н. В.И. Добренькова за постоянную поддержку данного направления исследований. Авторы благодарны д.соц.н., к.ф.-м.н. В.А. Шведовскому, д.ф.-м.н. М.Г. Дмитриеву, к.ф.-м.н. Г.Б. Прончеву, к.ф.-м.н. Е.Д. Корнилиной, к.ф.-м.н. О.Г. Прончевой, к.ф.-м.н. А.И. Маслову за многолетнее сотрудничество, а также члену-корреспонденту РАН И.Г. Поспелову, д.ф.-м.н. Г.Г. Малинецкому, к.ф.-м.н. М.Е. Степанцову, к.соц.н. В.М. Карповой, д.ф.-м.н. В.А. Дородницыну, д.ф.-м.н. Н.В. Змитренко, д.ф.-м.н. М.М. Горбунову-Посадову, инженеру С.М. Неретину за полезные обсуждения.

Авторы

*А. П. Михайлов,
А. П. Петров
29 мая 2021 г.*

Литература

1. Petty W. Political Arithmetic. 1690.
2. de Condorcet N. Essai sur l'application de l'analyse a la probabilitй des dйcisions rendues a la pluralitй des voix. 1785.
3. An Essay On The Principle Of Population. 1798 (автор — Т. Мальтус; издание вышло анонимно).
4. Менделеев Д.И. К познанию России. СПб., изд-во А.С. Суворина, 1906 (М., Дрофа, 2002).
5. Самарский А.А. Современная прикладная математика и вычислительный эксперимент // Коммунист, 1983, № 18. С. 31–42.
6. Самарский А.А. Неизбежность новой методологии // Коммунист, 1989, № 1. С. 82–92.
7. Самарский А.А., Михайлов А.П. Компьютеры и жизнь // Советская педагогика, 1987.
8. Самарский А.А., Михайлов А.П. Математическое моделирование. Идеи. Методы. Примеры. М.: Наука, ФИЗМАТЛИТ, 1997, 320 с. (англ. перевод: Samarskii A. A., Mikhailov A. P. Principles of Mathematical Modeling. Ideas, Methods, Examples London and New York. Taylor and Francis, 2002).
9. Самарский А.А., Михайлов А.П. Математическое моделирование в информационную эпоху // Вестник РАН. 2004. Т. 74. № 9.
10. Четверушкин Б.Н., Михайлов А.П. Триада Самарского. К 100-летию со дня рождения академика А.А. Самарского // Вестник Российской академии наук. 2019. Т. 89, № 2. С. 187–193.
11. Левин К. Закон и эксперимент в психологии / К. Левин. Динамическая психология. Сб. статей. С. 23–53.
12. Сноу Ч.П. Две культуры. М.: Прогресс. 1973.

Глава 1

НЕКОТОРЫЕ ТРАДИЦИОННЫЕ МАТЕМАТИЧЕСКИЕ МЕТОДЫ СОЦИАЛЬНЫХ НАУК

1.1. Применение нелинейной регрессии в социологических исследованиях

Исследование связей между переменными, описывающими те или иные характеристики социума, является одной из распространенных математических задач, возникающих при проведении социологических исследований.

В качестве примеров можно привести связь между уровнем безработицы и уровнем преступности либо между количеством голосов, поданных за определенного кандидата на некотором избирательном участке, и средним доходом избирателей этого участка.

В некоторых случаях теоретические соображения или математические модели позволяют выбрать математическую формулу этой связи — так называемую *функциональную спецификацию*, другими словами, указать ее качественный вид — линейная, экспоненциальная, логистическая и т. д. В этом случае основной целью эмпирической обработки данных является нахождение параметров связи, например коэффициентов линейной функции. Однако чаще всего социологические либо математические основания для выбора функциональной спецификации отсутствуют, и тогда этот выбор является неоднозначным. Так, в некоторых обстоятельствах для описания связи можно выбрать как экспоненциальную, так и степенную, либо какую-либо более сложную функциональную зависимость. При этом выбранная функция должна как минимум не противоречить социологическому смыслу исследуемого явления или процесса. Известно, например, что каким бы ни был демографический процесс, численность населения не может возрастать быстрее, чем экспоненциальным образом [1]; соответственно при статистической обработке эмпирических данных, для описания зависимости численности населения от времени, не могут использоваться функции, возрастающие быстрее экспоненты.

Предмет настоящего раздела ограничен данным кругом вопросов. В частности это означает, что здесь не рассматриваются вопросы, связанные с проверкой гипотез, применимостью метода наименьших квадратов и т. д. Первый пункт раздела содержит базовые сведения регрессионного анализа, изучаемые в общем курсе математической статистики для социологов. Во втором пункте рассматривается конкретный пример анализа данных, полученных в ходе эмпирического исследования.

1.1.1. Основные понятия регрессионного анализа

Предположим, что при эмпирическом исследовании изучается связь переменных x , y , определенных в шкале интервалов или отношений. По результатам n наблюдений были получены данные, сведенные в табл. 1.

Таблица 1. Представление эмпирических данных для парного регрессионного анализа

Номер наблюдения	1	2	n
Значения переменной x	x_1	x_2	x_n
Значения переменной y	y_1	y_2	y_n

Отдельные наблюдения чаще всего соответствуют либо различным объектам, либо различным моментам времени сбора данных об одном объекте (в последнем случае в качестве независимой переменной выступает время, а процедуры являются, вообще говоря, более сложными).

Важнейшей предпосылкой для применения регрессионного анализа является предположение о том, что между переменными x , y имеется связь. Его целью является количественное описание этой связи, позволяющее сделать статистические выводы, т. е. обоснованно прогнозировать значения переменной y при известном значении переменной x .

Например, при изучении влияния безработицы на преступность для статистического анализа может быть построена выборка городов из некоторой генеральной совокупности (например, города с населением от 300 тыс. до 1 млн чел.). В этом случае отдельные наблюдения соответствуют различным городам, значения переменной x — уровню безработицы в данном городе (в процентах), значения переменной y — количеству преступлений определенного вида, приходящемуся на 100 тыс. населения. С помощью регрессионного анализа можно получить, например функцию, описывающую «в среднем» количество угонов автомобилей в год при заданном уровне безработицы. Это позволит сделать статистический вывод, имеющий вид прогноза о количестве угонов для города, не попавшего в данную выборку.

Связь между переменными может быть как линейной (рис. 1, *а*), так и нелинейной (рис. 1, *б*). В первом случае скорость изменения одной переменной при изменении другой является постоянной. Например, увеличение безработицы на 1% влечет увеличение количества угонов на 10 единиц (в год, в расчете на 100 тыс. населения). При нелинейной связи увеличение безработицы с 5% до 6% (т. е. на 1%) и увеличение с 9% до 10% приводит, вообще говоря, к различным изменениям в статистике угонов — другими словами, скорость изменения переменной y при изменении переменной x зависит от текущего значения переменной x . На рис. 1, *б* продемонстрирован один из возможных вариантов

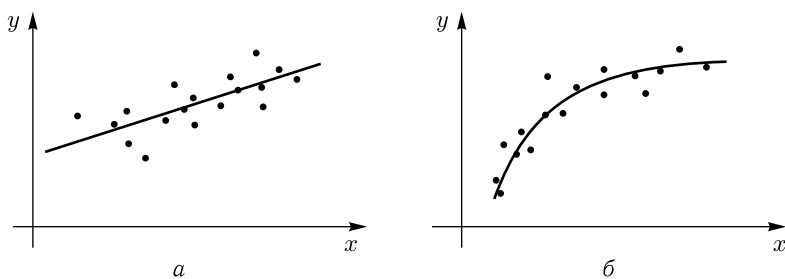


Рис. 1. Линейная зависимость (а) и нелинейная зависимость (б) между переменными x , y

нелинейной зависимости — по мере возрастания x скорость изменения переменной y уменьшается, но остается положительной. В терминах данного примера это означает, что при некотором значении безработицы количество уговоров достигает определенного (достаточно высокого) уровня и при дальнейшем росте безработицы почти не увеличивается.

Другие два примера вида нелинейной зависимости приведены на рис. 2. График на рис. 2, а соответствует случаю, когда увеличение переменной x до определенного значения почти не влияет на y , но после этого значения приводит к резкому уменьшению y .

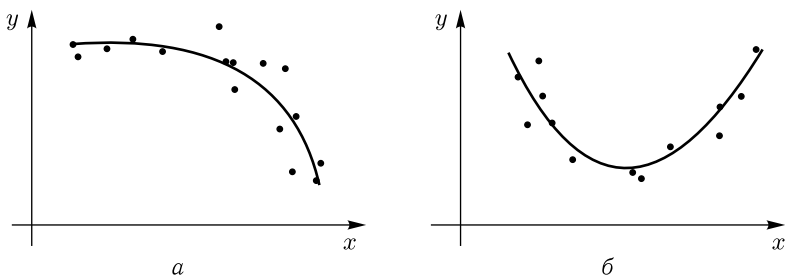


Рис. 2. Некоторые возможные виды нелинейной зависимости между переменными x , y

На рис. 2, б представлена немонотонная зависимость: постепенное увеличение переменной x приводит сначала к уменьшению, затем к увеличению переменной y . В социологических исследованиях такие явления встречаются не очень часто. В качестве примера можно привести зависимость уровня удовлетворенности браком от его продолжительности (см., напр., [2, 3]). Здесь переменная y представляет собой агрегированную величину, полученную в результате обработки ответов респондентов на ряд следующих вопросов.

- Насколько Вы удовлетворены тем, как Ваш супруг Вас понимает (по шкале от «1 — полностью удовлетворен» до «3 — не очень удовлетворен»)?

- Насколько Вы удовлетворены тем, как часто Вы с супругом соглашаетесь по разным поводам (по шкале от «1 — полностью удовлетворен» до «3 — не очень удовлетворен»)?
- Насколько Вы удовлетворены сексуальными отношениями с супругом (по шкале от «1 — полностью удовлетворен» до «3 — не очень удовлетворен»)?
- В сравнении с семьями Ваших знакомых как Вы оцениваете свой брак: лучше, чем большинство других, как большинство других, хуже большинства других?
- Оцените силу своего чувства любви по отношению к супругу (по шкале от «1 — максимально сильно люблю» до «5 — не очень сильно люблю»).

При исследованиях данной тематики анкеты содержат до нескольких десятков подобных вопросов. Ответы на них позволяют построить уровень общей удовлетворенности браком и изучить его зависимость от продолжительности пребывания в этом браке. Большинство исследований такого рода показывает, что с течением времени удовлетворенность браком изменяется немонотонным образом (как на рис. 2, б). Спад приходится на период воспитания детей, и удовлетворенность восстанавливается после уменьшения нагрузки, связанной с выполнением родительскими обязанностями.

Задача регрессионного анализа заключается в том, чтобы математически описать связь между переменными x , y , т.е. построить функцию

$$y = f(x), \quad (1)$$

которая является наиболее точной (в определенном смысле) аппроксимацией эмпирической зависимости. При этом предполагается, что наблюдаемая картина складывается из этой связи и воздействия случайных факторов. Таким образом, наблюдаемое значение y_i представляется в виде суммы,

$$y_i = f(x_i) + \delta_i, \quad (2)$$

где $f(x_i)$ является математическим ожиданием переменной y при условии, что переменная x приняла значение $x = x_i$, δ_i — остаточный член (иногда называемый остатком), т.е. случайная величина, имеющая нулевое математическое ожидание.

Уравнение (1) (а также уравнение вида $y = f(x) + \delta$) называют уравнением регрессии, а график функции $y = f(x)$ — линией регрессии. Переменная x называется регрессором, предиктором или объясняющей переменной, а y — объясняемой переменной.

Если объясняющих переменных несколько, то уравнение регрессии записывается в виде

$$y = f(x^1, x^2, \dots, x^n). \quad (3)$$

В этом случае, в противовес парной регрессии (1), говорят о множественной регрессии.

Линейная регрессионное уравнение (см. рис. 1, а) имеет вид

$$y = kx + b. \quad (4)$$

При выполнении так называемых *условий Гаусса–Маркова* (см., напр., [4]) значения коэффициентов k , b определяются с помощью метода наименьших квадратов. Его применение приводит к системе уравнений

$$k \sum_{i=1}^n x_i^2 + b \sum_{i=1}^n x_i = \sum_{i=1}^n x_i y_i, \quad (5)$$

$$k \sum_{i=1}^n x_i + bn = \sum_{i=1}^n y_i, \quad (6)$$

из которой однозначно определяются k , b .

После нахождения уравнения регрессии (как в линейном, так и нелинейном случаях) следует выяснить, насколько близки к нему эмпирические данные. На рис. 3 показаны два случая, соответствующие одному регрессионному уравнению. Однако на рис. 3, а эмпирические точки расположены значительно ближе к линии регрессии,

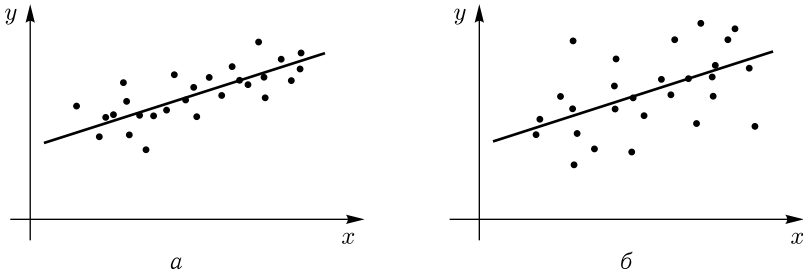


Рис. 3. Для случая, изображенного на левом рисунке, коэффициент детерминации R^2 выше, чем для ситуации, соответствующей правому рисунку

чем на рис. 3, б. Количественной мерой «сконцентрированности» точек около линии регрессии является коэффициент детерминации R^2 , обобщенный лишь для линейной регрессии и вычисляемый по формуле

$$R^2 = 1 - \frac{\sum_{i=1}^n \delta_i^2}{\sum_{i=1}^n (y_i - \bar{y})^2}, \quad (7)$$

где остаточные члены δ_i и среднее значение переменной y даются формулами

$$\delta_i = y_i - f(x_i), \quad (8)$$

$$\bar{y} = \frac{1}{n} \sum_{i=1}^n y_i. \quad (9)$$

Величина R^2 изменяется в пределах от 0 до 1; при этом чем больше R^2 , тем сильнее выражена зависимость переменных (подчеркнем, что это справедливо лишь для линейной регрессии). В частности, значение $R^2 = 1$ соответствует случаю, когда все эмпирические точки лежат на самой линии регрессии, другими словами — когда все остаточные члены равны нулю.

Результат регрессионного анализа считается минимально приемлемым в случае $R^2 > 0,5$ и достаточно убедительным лишь при $R^2 > 0,8$. Если же коэффициент детерминации не превышает величины 0,5, то выбранная гипотеза о регрессионной зависимости признается неверной. При проведении социологических исследований в таком случае следует либо изменить функциональную спецификацию, либо качественно пересмотреть исходную гипотезу о наличии связи между переменными (например, включить в анализ еще одну или несколько независимых переменных).

1.1.2. Функциональная спецификация нелинейной регрессии

Вопрос о функциональной спецификации рассмотрим на примере конкретного исследования, посвященному изучению автобиографической памяти.

Изучаемые социологами мнения, оценки, позиции индивидов зависят (помимо прочего) от автобиографической памяти этих индивидов. Личный опыт и формируемая ими память влияют на принимаемые людьми решения о трудовой деятельности, на политические предпочтения и т. д. Однако многие события или детали со временем забываются. При изучении этого процесса возникает вопрос о том, в какой степени индивид хранит память о событиях, произошедших определенное время назад. В настоящем пункте рассматривается вопрос о выборе функции, описывающей долю сохранившихся воспоминаний как функцию времени.

Эмпирическое исследование [5] проводилось в течение 10 полугодий (т. е. 5 лет), при этом в соответствии с определенной методикой в течение последних двух полугодий проводились припоминания, т. е. попытки вспомнить определенную информацию о событиях, произошедших в течение первых 8 полугодий. Тем самым были получены эмпирические значения y_i , ($i = 1, \dots, 10$) — доли правильных припоминаний о событиях, прошедших i полугодий назад. Задача нелинейного регрессионного анализа заключается в данном случае в том, чтобы

определить функцию $y(t)$, которую можно трактовать как долю информации, оставшуюся в памяти по истечении времени t после события.

Остановимся подробнее на методике эксперимента по изучению автобиографической памяти. Испытуемым, припоминаям события своей жизни, был сам исследователь, В. Вагенаар. На первой стадии эксперимента, длившейся 8 полугодий (т. е. 4 г.), он запротоколировал 1605 событий своей жизни в соответствии со следующей схемой. Относительно каждого события выделялись четыре аспекта, «кто», «что», «где» и «когда», служившие ключами для припоминания. Например, для одного из событий аспекты имели вид

- «кто» — Леонардо да Винчи,
- «что» — я ходил смотреть «Тайную Вечерю»,
- «где» — в Милане, в церкви,
- «когда» — в субботу 19 сентября 1983 г.

Кроме того, протокол содержал дополнительный вопрос о событии, например: «Вместе с кем я туда ходил?».

Вторая стадия эксперимента (9-ое и 10-е полугодия) была полностью посвящена припоминаниям и длилась один год. Таким образом, каждый день припоминалось 4–5 событий, отобранных случайным образом (так что некоторые из них могли быть записаны в начале первой стадии исследования, другие — в конце). Припоминание по каждому из событий проводилось в 4 этапа.

На первом этапе испытуемому сообщался один из аспектов, и задание заключалось в том, чтобы припомнить остальные три. Например, сообщалось

«кто» — Леонардо да Винчи;
требовалось припомнить — «что», «где» и «когда».

На втором этапе сообщался еще один из аспектов, а требовалось припомнить два оставшихся (при этом не обязательно было придерживаться ответов, данных на первом этапе). На третьем этапе становился известным третий аспект, и один последний аспект надо было припомнить. На четвертом этапе, при всех известных аспектах, требовалось ответить на дополнительный вопрос. Таким образом, обработка каждого события сводилась в общей сложности к ответам на 7 вопросов. Всего же в течение года испытуемый ответил на $1605 \times 7 = 11\,235$ вопросов.

По окончании второй стадии эксперимента сопоставлялись даты событий и даты припоминаний, интервал времени между ними округлялся до полугода. Тем самым были получены данные, представленные в табл. 2 и на рис. 4 (см. также упражнение 3).

Нетрудно видеть, что с течением времени, как и следовало ожидать, доля правильных припоминаний уменьшается. Упрощенно говоря, за первые полгода забывается около трети информации, за следующие 4 года — еще около трети. Таким образом, «скорость забывания» не является постоянной, и искомая зависимость $y(t)$ не может быть описана линейной функцией.

Таблица 2. Эмпирическая зависимость доли правильных ответов y от продолжительности интервала времени от события до припоминания

Продолжительность времени t (полугодия)	1	2	3	4	5	6	7	8	9
Доля правильных ответов y (проценты)	70	50	44	46	43	40	35	33	33

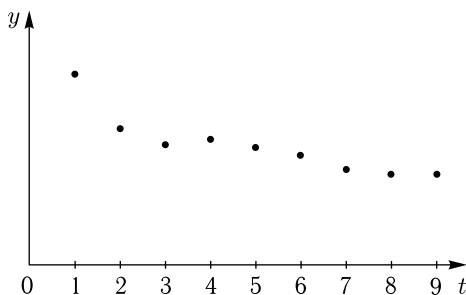


Рис. 4. Графическое отображение данных, представленных в табл. 2

В качестве возможных вариантов спецификации могут быть предложены, например следующие функции:

- экспоненциальная, $y = ae^{kt}$,
 - степенная, $y = at^b$,
 - сумма постоянной и экспоненциальной функций, $y = y_0 + ae^{kt}$.
- Рассмотрим последовательно все эти варианты.

Вариант 1: экспоненциальная функция

Будем искать зависимость между переменными t , y в виде

$$y = ae^{kt}. \quad (10)$$

Чтобы свести задачу к построению линейной регрессии, прологарифмируем уравнение (10),

$$\ln y = \ln a + kt,$$

и сделаем замены

$$\ln y = z, \quad \ln a = b. \quad (11)$$

Таким образом, имеем задачу построения линейной регрессии $z = kt + b$, где исходные данные получаются из табл. 2 в результате пересчета по первой из формул (11). Они представлены во второй и третьей строках табл. 3.

Выполняя вычисления, сведенные в строках 4–5 табл. 3, получаем, что система уравнений для постоянных k , b имеет вид

$$\begin{aligned} 285k + 45b &= 164,01, \\ 45k + 9b &= 33,77. \end{aligned} \quad (12)$$

Таблица 3. К вычислению уравнения регрессии в случае зависимости (10)

i	1	2	3	4	5	6	7	8	9	Сумма
t	1	2	3	4	5	6	7	8	9	45
z	4,25	3,91	3,78	3,83	3,76	3,69	3,56	3,50	3,50	33,77
t^2	1	4	9	16	25	36	49	64	81	285
tz	4,25	7,82	11,35	15,31	18,81	22,13	24,89	27,97	31,47	164,01

Решив ее, получаем значения $k = -0,08$, $b = 4,16$. Проводя обратную замену по второй из формул (11), получаем $a = e^b = 64,07$. Итак, уравнение регрессии принимает вид

$$y = 64,07e^{-0,08t}. \quad (13)$$

Найдем теперь коэффициент детерминации R^2 по формуле (7). Промежуточные вычисления сведены в табл. 4 (при этом учтено, что $\bar{y} = 43,78$), из которой следует, что

$$\sum_{i=1}^n \delta_i^2 = 186,79, \quad \sum_{i=1}^n (y_i - \bar{y})^2 = 1055,56.$$

Окончательно получаем

$$R^2 = 1 - \frac{186,79}{1055,56} = 0,82. \quad (14)$$

Полученное значение коэффициента детерминации является достаточно высоким. Таким образом, процесс забывания автобиографической информации довольно хорошо описывается уравнением (13).

Таблица 4. К вычислению коэффициента детерминации в случае зависимости (10)

t	1	2	3	4	5	6	7	8	9
y	70	50	44	46	43	40	35	33	33
$64,07e^{-0,08t}$	59,14	54,60	50,40	46,52	42,95	39,65	36,60	33,78	31,19
δ	10,86	-4,60	-6,40	-0,52	0,05	0,35	-1,60	-0,78	1,81
δ^2	117,85	21,13	40,95	0,28	0,00	0,13	2,56	0,61	3,29
$y - \bar{y}$	26,22	6,22	0,22	2,22	-0,78	-3,78	-8,78	-10,78	-10,78
$(y - \bar{y})^2$	687,6	38,72	0,05	4,94	0,60	14,27	77,05	116,16	116,16

Вариант 2: степенная функция

Будем искать зависимость между переменными t , y в виде

$$y = at^k. \quad (15)$$

Чтобы свести задачу к построению линейной регрессии, прологарифмируем уравнение (15),

$$\ln y = \ln a + k \ln t,$$

и сделаем замены

$$\ln y = z, \quad \ln a = b, \quad \ln t = x. \quad (16)$$

Таким образом, имеем задачу построения линейной регрессии $z = kx + b$, где исходные данные получаются из табл. 2 в результате пересчета по формулам (16). Они представлены во второй и третьей строках табл. 5.

Таблица 5. К вычислению уравнения регрессии в случае зависимости (15)

i	1	2	3	4	5	6	7	8	9	сумма
t	1	2	3	4	5	6	7	8	9	45
y	70	50	44	46	43	40	35	33	33	394
x	0,00	0,69	1,10	1,39	1,61	1,79	1,95	2,08	2,20	12,80
z	4,25	3,91	3,78	3,83	3,76	3,69	3,56	3,50	3,50	33,77
x^2	0,00	0,48	1,21	1,92	2,59	3,21	3,79	4,32	4,83	22,35
xz	0,00	2,71	4,16	5,31	6,05	6,61	6,92	7,27	7,68	46,71

Из вычислений, сведенных в строках 3–7 таблицы 3, получаем, что система уравнений для постоянных k , b имеет вид

$$\begin{aligned} 22,35k + 12,8b &= 46,71, \\ 12,8k + 9b &= 33,77. \end{aligned} \quad (17)$$

Решив ее, получаем значения $k = -0,32$, $b = 4,20$. Проводя обратную замену по формулам (16), получаем $a = e^b = 66,69$. Итак, уравнение регрессии принимает вид

$$y = 66,69t^{-0,32}. \quad (18)$$

Найдем теперь коэффициент детерминации R^2 по формуле (7). Промежуточные вычисления сведены в табл. 6, из которой следует, что

$$\sum_{i=1}^n \delta_i^2 = 59,50, \quad \sum_{i=1}^n (y_i - \bar{y})^2 = 1055,56.$$

Окончательно получаем

$$R^2 = 1 - \frac{59,50}{1055,56} = 0,94. \quad (19)$$

Полученное значение коэффициента детерминации является заметно более высоким, чем в варианте 1. Таким образом, степенная функция описывает процесс забывания автобиографической информации лучше,

Таблица 6. К вычислению коэффициента детерминации в случае зависимости (15)

t	1	2	3	4	5	6	7	8	9
y	70	50	44	46	43	40	35	33	33
$66,69t^{-0,32}$	66,69	53,42	46,92	42,80	39,85	37,59	35,78	34,28	33,01
δ	3,31	-3,42	-2,92	3,20	3,15	2,41	-0,78	-1,28	-0,01
δ^2	10,96	11,72	8,54	10,27	9,94	5,82	0,61	1,64	0,00
$y - \bar{y}$	26,22	6,22	0,22	2,22	-0,78	-3,78	-8,78	-10,78	-10,78
$(y - \bar{y})^2$	687,60	38,72	0,05	4,94	0,60	14,27	77,05	116,16	116,16

чем экспоненциальная. С точки зрения математической статистики результат регрессионного анализа следует считать очень хорошим.

Вариант 3: сумма постоянной и экспоненциальной функций

Заметим важную особенность обеих рассмотренных выше гипотез (в статистике такие гипотезы называются регрессионными моделями). Они предполагают, что с течением времени может быть забыта *любая* автобиографическая информация. Однако при том, что эмпирические исследования памяти на больших (составляющих десятки лет) временных интервалах отсутствуют, повседневный опыт позволяет высказать гипотезу, что некоторые события остаются в памяти индивида «навечно». Эта гипотеза позволяет сформулировать регрессионную модель, в рамках которой в автобиографической информации можно выделить две составляющих: одна из них забывается по экспоненциальному закону, а другая не забывается вовсе.

В соответствии с этим будем искать зависимость между переменными t, y в виде

$$y = y_0 + ae^{kt}. \tag{20}$$

Здесь постоянная y_0 соответствует «незабываемой» составляющей памяти. Чтобы свести задачу к построению линейной регрессии, прологарифмируем уравнение (10),

$$\ln(y - y_0) = \ln a + kt,$$

и сделаем замены

$$\ln(y - y_0) = z, \quad \ln a = b. \tag{21}$$

Таким образом, получаем задачу построения линейной регрессии $z = kt + b$.

В методических целях упростим задачу, положив $y_0 = 30$ (строго говоря, значение y_0 должно определяться в результате регрессионного анализа наряду с величинами a, k , однако в этом случае анализ будет

существенно более сложным). Опуская промежуточные вычисления, получим систему уравнений

$$\begin{aligned} 285k + 45b &= 85,27, \\ 45k + 9b &= 20,77. \end{aligned} \quad (22)$$

Решив ее, получаем значения $k = -0,31$, $b = 3,86$. Проводя обратную замену по формулам (16), получаем $a = e^b = 47,47$. Следовательно, уравнение регрессии имеет вид

$$y = 30 + 47,47e^{-0,31t}. \quad (23)$$

Найдем теперь коэффициент детерминации R^2 . Снова опуская промежуточные вычисления, получим, что

$$\sum_{i=1}^n \delta_i^2 = 101,51, \quad \sum_{i=1}^n (y_i - \bar{y})^2 = 1055,56.$$

Таким образом, имеем

$$R^2 = 1 - \frac{101,51}{1055,56} = 0,90. \quad (24)$$

Полученное значение коэффициента детерминации говорит о том, что функция вида (20) также может служить для описания процесса забывания автобиографической информации.

Еще один из многочисленных возможных вариантов функциональной спецификации для данных табл. 2 — это зависимость вида $y = ae^{k\sqrt{t}}$ (см. [6]), анализ которой вынесен в упражнение 2 к настоящему разделу.

В конечном счете выбор между перечисленными вариантами зависит от принятия гипотезы относительно существования в памяти индивида такой автобиографической информации, которая не может быть забыта с течением времени. В случае принятия данной гипотезы следует выбрать вариант 3; если же она отклоняется, то выбирать надо между другими вариантами. В свою очередь решение о принятии или отклонении этой гипотезы должно, по возможности, опираться на существующие социологические или психологические теории. Если веские основания в пользу того или иного варианта функциональной спецификации отсутствуют, то в контексте данного социологического исследования результат регрессионного анализа несет весьма ограниченную информацию о связи рассматриваемых переменных, что уменьшает его роль и значение.

Проведенный анализ показывает, что аппроксимация эмпирических данных, полученных в ходе социологического исследования, может быть проведена значительным количеством способов, основанных на различных видах функциональной зависимости между переменными. Наиболее предпочтительной является та функция, которая в большей степени соответствует социологическому содержанию процесса.

Упражнения

1. Построить регрессионное уравнение и найти коэффициент детерминации для данных табл. 2, приняв функциональную спецификацию вида (20), где $y_0 = 25$.

2. Построить регрессионное уравнение и найти коэффициент детерминации для данных табл. 2, приняв функциональную спецификацию вида $y = ae^{k\sqrt{t}}$, предложенную в [6].

3. При проведении эксперимента с автобиографической памятью, данные о количестве успешных припоминаний регистрировались отдельно для 9-го и 10-го полугодий исследования. Приведенная выше табл. 2 представляет результаты, полученные в течение 9-го полугодия. Данные по припоминаниям, сделанным в течение и 10-го полугодия,

Таблица 7. Эмпирическая зависимость доли правильных ответов от продолжительности интервала времени от события до припоминания (10-е полугодие эксперимента)

Продолжительность времени t (полугодия)	2	3	4	5	6	7	8	9	10
Доля правильных ответов y (проценты)	59	48	37	41	39	35	33	33	29

сведены в табл. 7. Задание: по этим данным построить уравнения регрессии для трех вариантов функциональной спецификации, рассмотренных во втором пункте.

Комментарии и библиографические примечания

1. Исследованию удовлетворенности браком посвящены многочисленные работы, в частности [2, 3]. Обзор литературы по этой теме см. в [3].
2. Эмпирическое исследование автобиографической памяти проведено в работе [5].
3. В качестве дополнительного чтения по теме «Регрессия» рекомендуется учебник [4].

Литература

1. Михайлов А.П. Улыбка, мысль и стрела времени, или мой С. П. Курдюмов // Мне нужно быть: Памяти Сергея Павловича Курдюмова. М. УРСС. 2010. С. 338–396.
2. Glenn N.D. The Course of Marital Success and Failure in Five American 10-Year Marriage Cohorts // Journal of Marriage and the Family. 60. P. 569–76.
3. VanLaningham J., Johnson D.R., Amato P. Marital Happiness, Marital Duration, and the U-Shaped Curve: Evidence from a Five-Wave Panel Study // Social Forces, June 2001, 78(4). P. 1313–1341.

4. *Дугерти К.* Введение в эконометрику. М.: ИНФРА-М, 2004. 419 с.
5. *Wagenaar W.* My Memory: A Study of Autobiographical Memory over Six Years // *Cognitive Psychology* 18. P. 225–252. 1986.
6. *Hunt E.* The Mathematics of Behavior. Cambridge University Press, 2007.

1.2. Принцип максимального правдоподобия, пробит-анализ и логит-анализ

Настоящий раздел посвящен изложению основополагающих идей, математических аспектов, а также вопросов содержательной социологической интерпретации результатов пробит-анализа и логит-анализа — широко востребованных в социальных науках методиках установления статистической связи между переменными в случае так называемого *бинарного отклика*.

Область их применения можно описать следующим образом. Предположим, что мы изучаем влияние ряда факторов на переменную, которая может принимать лишь два значения. Например, доучился ли ученик до окончания школы (или бросил ее), разошлась ли молодая семья в течение первых 5 лет после брака (или не разошлась), поступил ли абитуриент в университет (или не поступил). Факторы, т. е. объясняющие переменные, могут быть непрерывными (возраст, доход, трудовой стаж) либо также бинарными (абитуриент — юноша либо девушка), либо заданными в любой другой шкале. Задача, решаемая пробит- или логит-анализом — это обработка эмпирических данных относительно изучаемой выборки и построение статистического вывода о вероятности того или иного значения объясняемой бинарной переменной. Например: по итогам исследования 500 молодых семей требуется сделать вывод о том, что семья с определенными характеристиками (доходом, количеством детей, образованием супругов и т. д.) разведется в течение 5 лет после брака, с вероятностью 30%, и не разведется — с вероятностью 70%. Примеры реальных социологических исследований приведены ниже.

С математической точки зрения пробит- и логит-анализы основаны на принципе максимального правдоподобия, которому посвящен п. 1.2.1.

По сравнению с другими разделами настоящего пособия данный раздел требует большего привлечения материала, изучаемого в курсах высшей математики и математической статистики для социологов. Поэтому для удобства читателей в конце него приведено приложение с краткими необходимыми сведениями.

1.2.1. Принцип максимального правдоподобия

Чтобы продемонстрировать основную идею метода максимального правдоподобия, рассмотрим следующую серию вводных примеров.

Пример 1. Из 100 сотрудников мясокомбината 20 являются мужчинами, а 80 — женщинами. Найти вероятность того, что первый человек,

выходящий с мясокомбината по окончании рабочего дня, окажется женщиной (предполагается, что сотрудники выходят в случайном порядке, т. е. не берутся в учет особенности, связанные с расположением цехов по отношению проходной и т. д.).

Решение. Очевидно, вероятность того, что первый выходящий сотрудник окажется женщиной, составляет $80/100 = 0,8$, а что мужчиной — $20/100 = 0,2$.

В данном примере условия задачи содержали все данные, необходимые для вычисления требуемой вероятности. На основе этой вычисленной вероятности можно высказать прогноз о том, работник какого пола выйдет первым из проходной на следующий день: скорее всего, это будет женщина.

Метод максимального правдоподобия применяется в ситуациях, в некотором смысле обратных по отношению к рассмотренной в Примере 1. Именно, речь идет о случаях, когда известно случившееся событие и требуется высказать обоснованное суждение о параметрах системы, в которой это событие произошло.

Пример 2. В конце рабочего дня из проходной рыбокомбината первой вышла женщина (информация о том, кто вышел вслед за ней, отсутствует). Сотрудников какого пола больше работает на этом комбинате: мужчин или женщин?

Решение. Условия задачи не содержат данных, позволяющих вычислить вероятность того или иного соотношения полов. Однако интуиция подсказывает, что женщин на рыбокомбинате больше, чем мужчин.

Метод максимального правдоподобия позволяет обосновать это интуитивное соображение. Рассмотрим три возможных варианта.

Возможность 1. Среди работников рыбокомбината мужчин больше, чем женщин. В этом случае (пока мы не знали, кто первым выйдет из проходной) мы оценили бы вероятность события «первым работником, вышедшим из проходной, является женщина» как относительно маловероятное — так называемая *априорная вероятность* меньше, чем 50%.

Возможность 2. Если среди работников рыбокомбината мужчин столько же, сколько и женщин, то априорная вероятность данного события равна 50%.

Возможность 3. Если среди работников рыбокомбината женщин больше, чем мужчин, то априорная вероятность того, что первой выйдет женщина, превышает 50%.

После того, как данное событие произошло, мы можем оценить правдоподобность каждой из перечисленных возможностей. В первом случае оказывается, что произошло маловероятное событие, в третьем — весьма вероятное. Поэтому третья возможность является наиболее правдоподобной, первая — наименее правдоподобной. Таким образом, скорее всего, большинство сотрудников рыбокомбината составляют женщины.

Изложенный способ получения оценки о параметрах изучаемой системы (в данном случае — оценки соотношения количества мужчин и женщин на комбинате) называется *принципом максимального правдоподобия*.

Сравнение примеров 1 и 2 показывает, что нахождение вероятности позволяет оценить априорные шансы реализации того или иного случайного события, в то время как принцип максимального правдоподобия применяется, когда это событие уже произошло и требуется оценить параметры изучаемой системы. Вычисление (или оценка) априорных вероятностей является лишь промежуточным шагом в решении задачи.

Следующий пример иллюстрирует более сложную ситуацию, в которой применяется принцип максимального правдоподобия.

Пример 3. Каждый из шести членов экспертного совета должен вынести заключение о некотором проекте — либо поддержать его, либо отклонить. В преддверии решающего заседания был сделан выборочный опрос четырех экспертов: трое из них высказались в поддержку проекта, один — против. Требуется оценить количество экспертов совета, поддерживающих проект.

Решение. В выборку для опроса не попали всего двое членов совета из пяти. Имеются следующие три возможности: ни один из них не поддерживает проект, ровно один или оба поддерживают. Соответственно общее количество сторонников проекта составляет 3, 4 или 5.

Рассмотрим каждый из этих случаев.

Возможность 1. Предположим, что три эксперта из шести поддерживают проект, а трое против. Вычислим априорную вероятность P_3 попадания ровно двух сторонников проекта в выборку. Количество возможных выборок составляет C_6^4 — число сочетаний из 6 по 4. Количество выборок, содержащих трех сторонников и одного противника проекта, равно $C_3^3 C_3^1$. Таким образом, искомая вероятность равна (см. приложения к настоящему разделу)

$$P_3 = \frac{C_3^3 C_3^1}{C_6^4} = \frac{3 \cdot 4! \cdot 2!}{6!} = \frac{3}{15}.$$

Возможность 2. Предположим, что проект поддерживают четыре эксперта. Априорная вероятность P_4 попадания ровно трех из них в выборку составляет

$$P_4 = \frac{C_4^3 C_2^1}{C_6^4} = \frac{4 \cdot 2 \cdot 4! \cdot 2!}{6!} = \frac{8}{15}.$$

Возможность 3. Предположим, что проект поддерживают пять экспертов. Априорная вероятность P_5 попадания ровно двух из них

в выборку составляет

$$P_5 = \frac{C_5^3 C_1^1}{C_6^4} = \frac{5! \cdot 4! \cdot 2!}{3! \cdot 2! \cdot 6!} = \frac{10}{15}.$$

Сравнивая полученные вероятности P_3 , P_4 , P_5 , получаем, что наиболее правдоподобной является последняя возможность. Таким образом, оценка количества сторонников проекта по методу максимального правдоподобия составляет 5 экспертов.

1.2.2. Пробит-анализ и логит-анализ

Рассмотренные выше примеры иллюстрируют применение принципа максимального правдоподобия в простейших ситуациях. Однако в реальных социологических исследованиях этот принцип находит реализацию, как правило, в более сложных с математической точки зрения методах. В первую очередь это относится к пробит-анализу и логит-анализу — двум методикам исследования связей между признаками в случае так называемого *бинарного отклика*.

Другими словами, они применяются в тех случаях, когда изучается влияние непрерывной объясняющей переменной (или переменных) на объясняемую переменную, принимающую лишь два значения (так называемую *бинарную переменную*). Например, влияние дохода индивида на его позицию по отношению к правительству (поддерживаю—не поддерживаю) или влияние текущей зарплаты на решение сотрудника о поиске другой работы (буду искать—не буду искать).

Эмпирические данные в таких случаях могут быть представлены графиком, показанном на рис. 1.

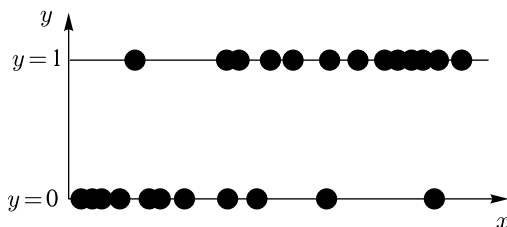


Рис. 1. Эмпирические данные для пробит-анализа или логит-анализа: случай, когда с увеличением значения переменной x возрастает вероятность того, что $y = 1$

Здесь точки отвечают отдельным респондентам, координата x соответствует непрерывной объясняющей переменной (например, доходу), а y может принимать лишь два значения: $y_i = 0$ (i -й индивид не поддерживает правительство) и $y_i = 1$ (поддерживает). Общее количество точек (т.е. респондентов) обозначим через n . Конфигурация, изображенная на рисунке, соответствует ситуации, при которой респонденты с высоким доходом в основном поддерживают правительство, а с низким не поддерживают.

Задача заключается в том, чтобы построить статистический вывод о генеральной совокупности, из которой сделана выборка. Именно, требуется найти вероятность $Y(x)$ того, что индивид с доходом x поддерживает правительство. В общей постановке объясняющих переменных может быть несколько, и задача имеет следующий вид: найти вероятность $Y(x^1, \dots, x^k)$ того, что бинарная переменная имеет значение $y = 1$, если объясняющие переменные имеют значения x^1, \dots, x^k .

В соответствии с методикой пробит-анализа (которую мы изложим для случая одной объясняющей переменной) функция $Y(x)$ ищется в виде интегрального нормального распределения (при необходимости см. в приложении соответствующие сведения из курса высшей математики):

$$Y(x) = \Phi(\alpha + \beta x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\alpha + \beta x} e^{-\frac{z^2}{2}} dz. \quad (1)$$

Здесь α, β — это пока неизвестные параметры, которые определяются на основе принципа максимального правдоподобия из эмпирических данных

$$(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n).$$

Это означает, что требуется найти такие значения α, β , при которых вероятность того, что получили именно такие эмпирические данные, максимальна.

Получим выражение для этой вероятности на основе следующих соображений. Пусть, например, респондент $i = 7$ имеет доход x_7 и поддерживает правительство: $y_7 = 1$. Априорная вероятность этого (т. е. вероятность того, что он поддерживает правительство «до того, как мы это узнали») имеет вид

$$P(y_7 = 1|x_7) = \Phi(\alpha + \beta x_7). \quad (2)$$

Респондент $i = 5$ имеет доход x_5 и не поддерживает правительство: $y_5 = 0$. Априорная вероятность равна

$$P(y_5 = 0|x_5) = 1 - \Phi(\alpha + \beta x_5). \quad (3)$$

Последние две формулы могут быть записаны единообразно:

$$P(y_7 = 1|x_7) = [\Phi(\alpha + \beta x_7)]^{y_7} [1 - \Phi(\alpha + \beta x_7)]^{1-y_7},$$

$$P(y_5 = 0|x_5) = [\Phi(\alpha + \beta x_5)]^{y_5} [1 - \Phi(\alpha + \beta x_5)]^{1-y_5}$$

(появившиеся по сравнению с формулами (2), (3) «новые» множители равны единице, так как содержащиеся в них выражения возводятся в нулевую степень).

Вообще, для любого респондента имеем

$$P(y_i = 0|x_i) = [\Phi(\alpha + \beta x_i)]^{y_i} [1 - \Phi(\alpha + \beta x_i)]^{1-y_i}.$$

По формуле умножения вероятностей получаем, что априорная вероятность $L_{5,7}(\alpha, \beta)$ того, что одновременно $y_5 = 0$ и $y_7 = 1$, равна произведению $P(y_5 = 0|x_5)P(y_7 = 1|x_7)$. Таким образом,

$$L_{5,7}(\alpha, \beta) = [\Phi(\alpha + \beta x_5)]^{y_5} [\Phi(\alpha + \beta x_7)]^{y_7} [1 - \Phi(\alpha + \beta x_5)]^{1-y_5} [1 - \Phi(\alpha + \beta x_7)]^{1-y_7}.$$

Априорная вероятность того, что для всех n респондентов получились именно такие значения y_i , которые были выявлены при опросе, имеет вид

$$L(\alpha, \beta) = \prod_{i=1}^n [\Phi(\alpha + \beta x_i)]^{y_i} [1 - \Phi(\alpha + \beta x_i)]^{1-y_i}. \quad (4)$$

В соответствии с принципом максимального правдоподобия надо найти такие значения α, β , при которых максимизируется априорная вероятность $L(\alpha, \beta)$. Чтобы отыскать максимум функции двух переменных, приравняем к нулю частные производные:

$$\frac{\partial L}{\partial \alpha} = 0, \quad \frac{\partial L}{\partial \beta} = 0. \quad (5)$$

Непосредственное вычисление частных производных от функции вида (4), представляющей собой произведение $2n$ множителей, является довольно громоздкой операцией. Поэтому, воспользовавшись логарифмической производной, сведем уравнения (5) к виду

$$\frac{\partial(\ln L)}{\partial \alpha} = 0, \quad \frac{\partial(\ln L)}{\partial \beta} = 0.$$

Прологарифмируем функцию $L(\alpha, \beta)$, получим

$$\begin{aligned} \ln L(\alpha, \beta) &= y_1 \ln [\Phi(\alpha + \beta x_1)] + (1 - y_1) \ln [1 - \Phi(\alpha + \beta x_1)] + \\ &+ \dots + y_n \ln [\Phi(\alpha + \beta x_n)] + (1 - y_n) \ln [1 - \Phi(\alpha + \beta x_n)] = \\ &= \sum_{i=1}^n \{y_i \ln [\Phi(\alpha + \beta x_i)] + (1 - y_i) \ln [1 - \Phi(\alpha + \beta x_i)]\}. \end{aligned}$$

Теперь вычислим и приравняем к нулю частные производные:

$$\frac{\partial(\ln L)}{\partial \alpha} = \sum_{i=1}^n \left\{ y_i \frac{\Phi'(\alpha + \beta x_i)}{\Phi(\alpha + \beta x_i)} + (1 - y_i) \frac{-\Phi'(\alpha + \beta x_i)}{1 - \Phi(\alpha + \beta x_i)} \right\} = 0, \quad (6)$$

$$\frac{\partial(\ln L)}{\partial \beta} = \sum_{i=1}^n \left\{ x_i y_i \frac{\Phi'(\alpha + \beta x_i)}{\Phi(\alpha + \beta x_i)} + x_i (1 - y_i) \frac{-\Phi'(\alpha + \beta x_i)}{1 - \Phi(\alpha + \beta x_i)} \right\} = 0. \quad (7)$$

Значения параметров α, β находятся из решения системы уравнений (6), (7). На практике эта система решается с помощью программных продуктов, чаще всего, — пакетов статистических программ. Тем самым определяется искомая функция (9) — вероятность того,

что бинарная объясняемая переменная имеет значение $y = 1$, если непрерывная объясняющая переменная имеет значение x .

При интерпретации результата следует иметь в виду два момента.

- Знак параметра β говорит о направлении связи между переменными. Приведенные выше рассуждения строились в предположении, что чем больше значение переменной x , тем выше вероятность того, что $y = 1$ (рис. 1). Такая связь имеет место при $\beta > 0$. Если же пробит-анализ показывает, что $\beta < 0$, то с увеличением значения переменной x возрастает вероятность того, что $y = 0$ (рис. 2).

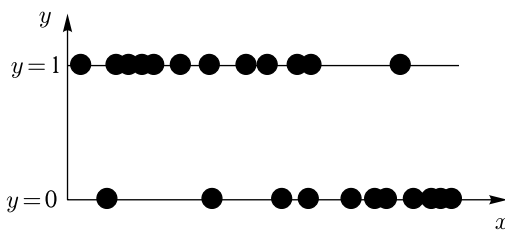


Рис. 2. Если $\beta < 0$, то с увеличением значения переменной x возрастает вероятность того, что $y = 0$

- Чем больше абсолютная величина параметра β , тем сильнее влияет изменение объясняющей переменной на вероятность того, что $y = 1$. Пусть, например, x имеет смысл возраста индивида, $y = 0$ и $y = 1$, отсутствие и наличие у него высшего образования. Тогда, если для мужской части некоторой социальной группы получено значение $\beta = 0,10$, а для женской — $\beta = 0,15$, то это означает, что вероятность наличия высшего образования у индивида тем выше, чем старше этот индивид, причем эта зависимость выражена у женщин сильнее, чем у мужчин.

В социологических исследованиях обычно рассматривается зависимость y не от одной, а от нескольких объясняющих переменных,

$$Y(x_1, x_2, \dots, x_n) = \Phi(\alpha + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_n x_n), \quad (8)$$

при этом некоторые из них могут быть заданы в номинальной или порядковой шкале. Принцип максимального правдоподобия позволяет определить значения всех параметров $\alpha, \beta_1, \beta_2, \dots, \beta_n$ и содержательно интерпретировать результаты.

В качестве примера реального социологического исследования с применением пробит-анализа приведем работу об изучении социальных факторов детской смертности в Кении. При этом мы здесь ограничимся лишь одним из полученных в данной работе результатов, касающегося смертности среди мальчиков (всего 32 216 наблюдений), и факторами, описывающими социальное положение семьи.

Объясняемая переменная пробит-анализа имеет следующие значения: $y = 0$ означает, что мальчик не дожил, и $y = 1$ — дожил до возраста

пять лет. Объясняющие переменные и полученные значения коэффициентов β_i :

x_1 — возраст матери (в годах), $\beta_1 = 0,019$,

x_2 — образование матери (количество лет обучения), $\beta_2 = 0,036$,

x_3 — образование отца (количество лет обучения), $\beta_3 = 0,001$,

x_4 — проживание в сельской местности, $\beta_5 = -0,020$,

x_5 — семья является бедной, $\beta_6 = 0,099$.

Первые три переменные суть непрерывные, в то время как x_4 , x_5 — бинарные. Именно, если семья проживает в сельской местности, то $x_4 = 1$, если в городе, то $x_4 = 0$. Если она является бедной (т.е. удовлетворяет определенным характеристикам), то $x_5 = 1$, если не является, то $x_5 = 0$.

При интерпретации результатов пробит-анализа следует иметь в виду, что главной причиной детской смертности в Кении является малярия.

Полученные значения коэффициентов β_i показывают, что кенийские дети умирают чаще в менее образованных семьях, у более юных матерей и в сельской местности. Парадоксальной выглядит положительная связь между бедностью и выживаемостью детей, выраженная положительным значением пробит-коэффициента $\beta_6 = 0,099$. Этот парадокс связан с тем, что у мальчиков недоедание приводит к возникновению определенного иммунитета против малярии. В то же время у девочек такого иммунитета не возникает, и значение соответствующего пробит-коэффициента отрицательно.

Таким образом, данный пример иллюстрирует возможности пробит-анализа в социологии.

Логит-анализ — это методика, в основных чертах сходная с пробит-анализом, однако вероятность того, что бинарная зависимая переменная имеет значение $y = 1$, ищется не в виде интегрального нормального распределения (1), а в виде функции

$$Y(x_1, \dots, x_n) = \frac{\exp(\alpha + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_n x_n)}{1 + \exp(\alpha + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_n x_n)}. \quad (9)$$

Интерпретация логит-коэффициентов β_i проводится на основе тех же соображений, что и в пробит-анализе.

В качестве примера рассмотрим данные исследования, посвященного изучению факторов, влияющих на то, что британские дети бросают школу, не доучившись до конца. Исследование проводилось в 1978–1984-х годах, размер выборки — 2177 подростков в возрасте 16 лет.

Объясняемая переменная логит-анализа имеет значение $y = 0$, если подросток доучился до конца, и $y = 1$, если он бросил школу.

Объясняющие переменные и полученные значения коэффициентов β_i таковы:

x_1 — уровень дохода (натуральный логарифм от дохода, приходящегося на одного взрослого члена семьи), $\beta_1 = -0,205$;

x_2 — уровень образования главы семьи (количество лет обучения сверх школьного), $\beta_2 = -0,151$;

x_3 — доля прошлого года, в течение которой глава семьи был безработным, $\beta_3 = 0,359$;

x_4 — уровень безработицы в регионе (в процентах), $\beta_4 = 0,007$;

x_5 — натуральный логарифм от количества детей в семье, $\beta_5 = 0,101$;

x_6 — подросток является девочкой ($x_5 = 1$, если да; $x_5 = 0$, если нет), $\beta_6 = -0,256$.

Приведенные значения логит-коэффициентов показывают, что школу чаще бросают дети из бедных и многодетных семей, мальчики, жители регионов с более высокой безработицей и т. д.

Таким образом, применение и социологическая интерпретация результатов логит-анализа технически почти не отличается от пробит-анализа. Выбор между этими методиками представляет собой довольно сложную проблему, которая имеет неоднозначное решение только в тех случаях, когда есть теоретические основания полагать, что существует зависимость (8) или (9). Такая ситуация возникает, например, если известна математическая модель для функции $Y(x_1, \dots, x_n)$, позволяющая определить качественный вид зависимости вероятности Y от объясняющих переменных. В этом случае статистическая обработка эмпирических данных с применением пробит- или логит-анализа используется для того, чтобы определить конкретные значения параметров для этой зависимости. В большинстве ситуаций, однако, такие математические модели отсутствуют, и допустимым является применение как пробит-, так и логит-анализа.

Упражнения

1. Известно, что большинство сотрудников автобазы составляют мужчины, а большинство сотрудников хлебозавода — женщины. Проходные автобазы и хлебозавода расположены рядом, на них нет табличек. Социолог ожидает, пока кто-нибудь из работников этих предприятий выйдет на улицу. Наконец, он видит, что из левой проходной вышел мужчина. Требуется высказать обоснованное суждение о том, какая из проходных (левая или правая) принадлежит автобазе.

Указание. В данном случае нет возможности вычислить (или хотя бы оценить) вероятность того, что автобазе принадлежит, например левая проходная. Поэтому обоснованное суждение может быть высказано о правдоподобии, но не о вероятности. Всего возможны два случая: левая проходная принадлежит либо хлебозаводу, либо автобазе. Их следует рассмотреть поочередно. Предположим, что она принадлежит хлебозаводу. Так как большинство его сотрудников составляют женщины, то событие «мужчина вышел из левой проходной» имеет вероятность ниже, чем 50%, и т. д.

2. Пусть каждый из восьми членов экспертного совета должен вынести заключение о некотором проекте — либо поддержать его, либо отклонить. В преддверии решающего заседания был сделан выборочный опрос пяти экспертов: четверо из них высказались в поддержку

проекта, один — против. Требуется найти наиболее вероятное количество экспертов совета, поддерживающих проект.

Решение. Не попали в выборку для опроса всего трое членов совета из восьми. Возможны следующие четыре варианта: ни один из них не поддерживает проект, ровно один поддерживает проект, ровно двое и ровно трое. Соответственно общее количество сторонников проекта составляет 4, 5, 6 или 7.

Рассмотрим каждый из этих вариантов.

Если проект поддерживают всего четыре эксперта из восьми, то вероятность того, что все они попали в выборку (составляющую пять человек), равна

$$P_4 = \frac{C_4^1}{C_8^5} = \frac{4 \cdot 3!5!}{8!} = \frac{4}{56} = \frac{1}{14}.$$

Если проект поддерживают пятеро экспертов, то вероятность того, что ровно четверо из них попали в выборку, равна

$$P_5 = \frac{C_5^4 C_3^1}{C_8^5} = \frac{5 \cdot 3 \cdot 3!5!}{8!} = \frac{15}{56}.$$

Аналогичным образом

$$P_6 = \frac{C_6^4 C_2^1}{C_8^5} = \frac{6! \cdot 2 \cdot 3!5!}{2!4!8!} = \frac{30}{56},$$

$$P_7 = \frac{C_7^4 C_1^1}{C_8^5} = \frac{7!3!5!}{3!4!8!} = \frac{5}{8} = \frac{35}{56}.$$

Итак, наиболее правдоподобно, что семь экспертов поддерживают проект. Заметим, что в данном примере мы не можем вычислить вероятность того, что проект поддерживают ровно семь экспертов. Вычисленная величина P_7 имеет другой смысл — это вероятность того, что ровно четыре сторонника проекта попали в выборку при условии, что всего этих сторонников семь.

Комментарии и библиографические примечания

1. Анализ социальных факторов детской смертности в Кении проведен в работе [1].
2. Изучение факторов, влияющих на то, что британские дети бросают школу, не доучившись до конца, проведено в [2].

Литература

1. Gor S. Inequality and Child Survival in Kenya: A Probit Model Approach // IARIW 31st General Conference, St-Gallen, Switzerland, August 22-28, 2010. <http://www.iariw.org/papers/2010/poster2Gor.pdf>.
2. Micklewright J. The Analysis of Pooled Cross-Sectional Data // Analyzing social and political change. A Casebook of Methods edited by Angela Dale and Richard B. Davies. p. 92.

Приложение. Некоторые сведения из курса высшей математики и математической статистики

В настоящем приложении в краткой форме приведены некоторые математические сведения, использованные выше.

1. Число сочетаний из n по k

Рассмотрим некоторое множество, состоящее из n объектов. Предположим, что требуется выбрать из него k объектов, причем их порядок в выборке не имеет значения. Такие выборки называются *сочетаниями из n по k* . Например, если исходное множество состоит из $n = 4$ элементов и имеет вид $\{A, B, C, D\}$, то имеются шесть сочетаний по $k = 2$ элемента:

$$AB, AC, AD, BC, BD, CD.$$

Количество сочетаний из n по k обозначается C_n^k , например, $C_4^2 = 6$. Для вычисления величины C_n^k используется следующая формула:

$$C_n^k = \frac{n!}{k!(n-k)!}. \quad (10)$$

Например, C_4^2 можно определить не только путем перебора всевозможных сочетаний (как сделано выше), но и путем вычисления по формуле (10):

$$C_4^2 = \frac{4!}{2!2!} = \frac{1 \cdot 2 \cdot 3 \cdot 4}{1 \cdot 2 \cdot 1 \cdot 2} = 6.$$

В качестве дополнительного примера вычислим величину C_8^3 :

$$C_8^3 = \frac{8!}{3!5!} = \frac{1 \cdot 2 \cdot 3 \cdot 4 \cdot 5 \cdot 6 \cdot 7 \cdot 8}{1 \cdot 2 \cdot 3 \cdot 1 \cdot 2 \cdot 3 \cdot 4 \cdot 5} = 56.$$

Материал данного пункта используется выше (метод максимального правдоподобия).

2. Решение задач по теории вероятностей, требующих подсчета числа сочетаний

Рассмотрим следующую задачу.

Пример 1. Группа состоит из 12 студентов — 7 девушек и 5 юношей. Случайным образом определяют четыре студента, которые будут сидеть во время экзамена на последней парте. Найти вероятность того, что на последнюю парту попадут: (а) четыре девушки, (б) три девушки и один юноша, (в) две девушки и два юноши, (г) одна девушка и три юноши, (д) четыре юноши.

Решение. Общее количество выборок из 12 студентов по 4 равно

$$C_{12}^4 = \frac{12!}{4!8!}.$$

Теперь рассмотрим поочередно каждый из вариантов (а)–(д).

(а) Выборки, при которых на последнюю парту попадут 4 девушки и ни одного юноши (будем называть их выборками типа $4 + 0$), получают путем сочетаний по 4 девушек из 7. Поэтому их количество равно $C_7^4 = \frac{7!}{4!3!}$. Таким образом, вероятность P_{4+0} того, что на последней парте будут 4 девушки (и ни одного юноши), равна

$$P_{4+0} = \frac{C_7^4}{C_{12}^4} = \frac{\frac{7!}{4!3!}}{\frac{12!}{4!8!}} = \frac{7!}{4!3!} \cdot \frac{4!8!}{12!} = \frac{5 \cdot 6 \cdot 7}{1 \cdot 2 \cdot 3} \cdot \frac{1 \cdot 2 \cdot 3 \cdot 4}{9 \cdot 10 \cdot 11 \cdot 12} = \frac{7}{99}.$$

(б) Чтобы получить выборку типа $3 + 1$ (т.е. состоящую из трех девушек и одного юноши), необходимо выбрать 3 девушки из 7, а также одного юношу из 5. Поэтому количество способов, которыми можно сформировать женскую часть выборки, равно C_7^3 , а мужскую — C_5^1 . Таким образом, искомая вероятность равна

$$P_{3+1} = \frac{C_7^3 C_5^1}{C_{12}^4} = \frac{\frac{7!}{3!4!} \frac{5!}{1!4!}}{\frac{12!}{4!8!}} = \frac{7!5}{3!4!1!4!} \cdot \frac{4!8!}{12!} = \frac{35}{99}.$$

(в) Выборка типа $2 + 2$ получается путем выбора 3 девушек из 7, и 2 юношей из 5. Поэтому

$$P_{2+2} = \frac{C_7^2 C_5^2}{C_{12}^4} = \frac{\frac{7!}{2!5!} \frac{5!}{2!3!}}{\frac{12!}{4!8!}} = \frac{7!}{2!2!3!} \cdot \frac{4!8!}{12!} = 5 \cdot 6 \cdot 7 \cdot \frac{2 \cdot 3 \cdot 4}{9 \cdot 10 \cdot 11 \cdot 12} = \frac{42}{99}.$$

(г) Выборка типа $1 + 3$ получается путем выбора 1 девушки из 7, и 3 юношей из 5. Поэтому

$$P_{1+3} = \frac{C_7^1 C_5^3}{C_{12}^4} = \frac{\frac{7!}{1!6!} \frac{5!}{3!2!}}{\frac{12!}{4!8!}} = \frac{7 \cdot 5!}{3!2!} \cdot \frac{4!8!}{12!} = 70 \cdot \frac{2 \cdot 3 \cdot 4}{9 \cdot 10 \cdot 11 \cdot 12} = \frac{14}{99}.$$

(д) Выборка типа $0 + 4$ получается путем выбора 4 юношей из 5. Поэтому

$$P_{0+4} = \frac{C_5^4}{C_{12}^4} = \frac{\frac{5!}{4!1!}}{\frac{12!}{4!8!}} = 5 \cdot \frac{4!8!}{12!} = \frac{1}{99}.$$

Таким образом, все искомые вероятности найдены. Заметим, что варианты (а)–(д) охватывают все возможные типы выборок в данном случае. Поэтому сумма найденных вероятностей равна единице:

$$P_{4+0} + P_{3+1} + P_{2+2} + P_{1+3} + P_{0+4} = \frac{7 + 35 + 42 + 14 + 1}{99} = 1.$$

3. Нормальное распределение

Непрерывная случайная величина называется нормально распределенной с математическим ожиданием μ и дисперсией σ^2 и обозначается $N(\mu, \sigma^2)$, если ее плотность распределения имеет вид

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}.$$

Интегральная функция распределения имеет вид

$$F(x) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{(t-\mu)^2}{2\sigma^2}} dt.$$

Нормальное распределение $N(0, 1)$ называется стандартным, его плотность распределения представляется функцией

$$\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}},$$

а интегральная функция распределения есть

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{t^2}{2}} dt.$$

Таблицы ее значений широко используются как при решении учебных задач, так и проведении исследований.

Любое нормальное распределение может быть приведено к стандартному нормальному путем замены переменной. Так, сделав под знаком интеграла в (1.2.2) замену

$$z = \frac{t-\mu}{\sigma}, \quad t = \sigma z + \mu, \quad dt = \sigma dz,$$

получим

$$F(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\frac{x-\mu}{\sigma}} e^{-\frac{z^2}{2}} dz = \Phi\left(\frac{x-\mu}{\sigma}\right) = \Phi\left(\frac{-\mu}{\sigma} + \frac{1}{\sigma}x\right) = \Phi(\alpha + \beta x).$$

В пробит-анализе параметры

$$\alpha = -\frac{\mu}{\sigma}, \quad \beta = \frac{1}{\sigma}$$

более удобны для социологической интерпретации, чем математическое ожидание μ и дисперсия σ^2 .

1.3. Кластеризация

Кластерный анализ — это общее название весьма многочисленных методов и процедур (не всегда имеющих строгое математическое обоснование), позволяющих проводить классификацию объектов или признаков путем их группировки. При этом количество выделенных групп (кластеров), как правило, также определяется в ходе анализа.

Полученное в результате анализа разбиение на группы, как правило, поддается осмыслению в содержательных социологических терминах и позволяет получить новые знания об обществе.

Например, при определении социальной близости различных этнических групп республики Дагестан с целью поиска и путей урегулирования этнополитических конфликтов в работе [14] был проведен кластерный анализ этносов на основе языкового барьера между ними, религиозной близости и среднего образовательного уровня. Полученный результат использовался при разработке методики прогнозирования развития общественно-политической ситуации в республике.

Заметим, что в прикладных науках (не только социальных) нередко возникают задачи кластеризации большого числа объектов, с необходимостью требующие применения вычислительной техники. Поэтому современное развитие кластерного анализа связано как с разработкой новых методов и их математическим обоснованием, так и с решением задач оптимизации вычислений, возникающих при кластеризации. Однако здесь мы ограничимся описанием наиболее простых, базовых процедур, наиболее употребительных в социальных науках

1.3.1. Входные данные для кластерного анализа

Данные, полученные в результате эмпирических социологических исследований, часто представляются в виде матрицы «объект-признак», строки которой соответствуют единицам измерения (например, человеческим индивидам), а столбцы — измеряемым признакам этих объектов (возраст, пол, уровень дохода и т. д.). При этом, вообще говоря, признаки могут быть измерены в различных типах шкал: например, возраст представляется в шкале отношений, а пол — в номинальной. Далее будем обозначать эту матрицу через X , а значение m -го признака у n -го объекта — через x_{ij} . Здесь $1 \leq n \leq N$, $1 \leq m \leq M$, где N — количество объектов, M — количество признаков.

Например, матрица «объект-признак» может иметь следующий вид (табл. 1). Ее строки соответствуют объектам, т. е. отдельным городам, а столбцы — признакам этих объектов.

Чтобы провести кластеризацию объектов, на основе матрицы «объект-признак» составляют матрицу S сходства объектов, элемент s_{ij} которой описывает, насколько близкими друг другу являются i -й и j -й объекты (см., напр., табл. 3). При этом $0 \leq s_{ij} \leq 1$: $s_{ij} = 1$ соответствует максимально возможному сходству, $s_{ij} = 0$ — минимально возможному. Каждый объект максимально сходен с самим собой, $s_{ii} = 1$,

Таблица 1. Пример матрицы «объект-признак»

Объект	Средний возраст	Средний доход	...	Уровень безработицы
Город 1	41,7	23 200	...	9%
Город 2	32,3	17 500	...	25%
...
Город 50	40,4	36 100	...	3%

имеет место симметрия, $s_{ji} = s_{ij}$, $1 \leq i \leq N$, $1 \leq j \leq N$ (т. е. i -й объект настолько же сходен с j -м, насколько j -й с i -м). В частности, если все признаки измеряются в шкале интервалов или шкале отношений, то мера сходства часто строится на основе коэффициента корреляции Пирсона, изучаемого в рамках стандартного курса математической статистики для социологов). Именно, коэффициент корреляции

$$r_{ij} = \frac{\sum_{m=1}^M (x_{im} - \bar{x}_i)(x_{jm} - \bar{x}_j)}{\sqrt{\sum_{m=1}^M (x_{im} - \bar{x}_i)^2 \sum_{m=1}^M (x_{jm} - \bar{x}_j)^2}}$$

изменяется в пределах от -1 (минимально возможное сходство) до 1 (максимально возможное), поэтому величина

$$s_{ij} = \frac{1 + r_{ij}}{2}$$

может служить мерой сходства i -го и j -го объектов. Иногда в прикладных исследованиях мера сходства нормируется так, что может принимать значения не от 0 до 1 , а от 0 до 100 .

Существует большое количество других методов введения меры сходства, некоторые из них основаны на введении расстояния (например, евклидова) в пространстве признаков. При этом в некоторых случаях может потребоваться предварительное преобразование исходных данных. Вместо матрицы сходства при кластеризации используют также матрицу различий объектов.

Наконец, заметим, что иногда матрица сходства строится без этапа построения матрицы «объект-признак», т. е. непосредственно на основе эмпирических или некоторым иным образом рассчитанных данных.

В качестве примера приведем исследование сходства программ политических партий на выборах в Государственную Думу, проведенное методом латентно-семантического анализа (ЛСА). Этот метод исследования и сопоставления текстов основан на концепции так называемой

синтагматической близости, в соответствии с которой смысл слова связан с контекстом употребления этого слова. Применительно к анализу политических текстов основную идею ЛСА можно пояснить на следующем упрощенном примере. Если в одном тексте фамилия некоторого политика употребляется в окружении таких слов, как «прогресс», «патриотический», «достижения», а в другом тексте — рядом со словами «коррупция», «беспомощность», «имитация», то, скорее всего, первый текст написан сторонником этого политика, а второй — противником. Метод был запатентован группой американских ученых в 1987 г., в настоящее время существует довольно обширная литература, посвященная ему.

В соответствии с методикой ЛСА программа каждой партии разбивается на небольшие фрагменты длиной в несколько десятков или сотен слов, и с помощью специальной процедуры, основанной на понятии синтагматической близости, устанавливается мера сходства между каждыми двумя фрагментами программ двух партий. Затем данные о сходстве отдельных фрагментов агрегируются в один показатель, служащий мерой сходства партийных программ в целом. В табл. 2 представлены полученные таким образом матрицы сходства программ Единой России, КПРФ, ЛДПР и Яблока на выборах 2007 и 2011 гг.

Таблица 2. Матрицы сходства партийных программ 2007 и 2011 гг.

Партия, 2007 г.					Партия, 2011 г.				
	ЕР	КПРФ	ЛДПР	Яблоко		ЕР	КПРФ	ЛДПР	Яблоко
ЕР	—	99	92	87	ЕР	—	87	91	89
КПРФ	99	—	95	91	КПРФ	87	—	96	85
ЛДПР	92	95	—	91	ЛДПР	91	96	—	90
Яблоко	87	91	91	—	Яблоко	89	85	90	—

a

б

Матрица сходства содержит данные, на основании которых (с помощью того или иного метода кластерного анализа) исходные объекты распределяются по кластерам. Даже беглый обзор этих методов выходит далеко за рамки настоящей книги, поэтому мы ограничимся одним примером применения кластерного анализа в реальном исследовании.

1.3.2. Иерархическая кластеризация

Если говорить о применении кластерного анализа в социальных науках, то один из наиболее распространенных методов — это иерархическая кластеризация. Ее главными достоинствами являются простота

и возможность наглядного представления результата, облегчающего его содержательную интерпретацию.

Результат визуализируется путем пошагового построения *дендрограммы*, т. е. связного графа, не содержащего циклов и построенного на основании матрицы сходства. Процедуру иерархической кластеризации построения можно описать следующим образом.

Как указано выше, входные данные для алгоритма имеют вид матрицы сходства (или расстояния) между одиночными объектами.

Шаг 1. Из рассматриваемых N объектов выбираются два наиболее близких (сходных) и объединяются в кластер, что отображается на дендрограмме соединением этих объектов. Например, для данных табл. 2(а), наиболее близкими являются предвыборные программы ЕР и КППРФ — см. рис. 1.

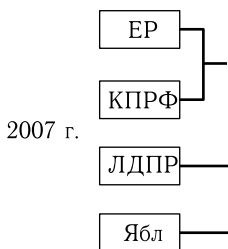


Рис. 1. Первый этап построения дендрограммы

Шаг 2. Рассматриваются получившиеся $N - 1$ объектов (образованный на первом этапе кластер и еще $N - 2$ одиночных объекта), для них вводится новая матрица сходства.

Шаг 3. Два наиболее сходных объекта объединяются в новый кластер.

Шаг 4. Повторяются шаги 2, 3 до тех пор, пока не будет получен единый кластер, объединяющий все N объектов исходной совокупности.

В зависимости от того, какой принцип заложен в построение «новой» матрицы сходства на шаге 2, получаются различные алгоритмы иерархической кластеризации. Рассмотрим наиболее простые из них.

Пусть кластеры C_1, C_2, C_3 содержат n_1, n_2, n_3 объектов соответственно (в частности, одиночный объект тоже понимается как кластер, в этом случае $n_i = 1$). Расстояние между i -м и j -м кластерами обозначим через d_{ij} . Если на некотором шаге кластеры C_2, C_3 объединяются в новый кластер C_4 , то расстояние между C_1 и C_4 принимается равным:

- в методе ближнего соседа $d_{14} = \min(d_{12}, d_{13})$,
- в методе дальнего соседа $d_{14} = \max(d_{12}, d_{13})$,

- во взвешенном попарно-групповом методе со средним

$$d_{14} = \frac{1}{2}(d_{12} + d_{13}),$$

- в невзвешенном попарно-групповом методе со средним

$$d_{14} = \frac{n_2}{n_2 + n_3}d_{12} + \frac{n_3}{n_2 + n_3}d_{13},$$

- в методе Уорда

$$d_{14} = \frac{n_1 + n_2}{n_1 + n_2 + n_3}d_{12} + \frac{n_1 + n_3}{n_1 + n_2 + n_3}d_{13} - \frac{n_1}{n_1 + n_2 + n_3}d_{23}.$$

Можно сказать, что в *методе ближнего соседа* расстояние между двумя кластерами полагается равным расстоянию между двумя наиболее близкими объектами, первый из которых берется из первого кластера, второй — из второго. В противоположность этому, в *методе дальнего соседа* расстояние вычисляется по двум наиболее удаленными объектами. Три других метода основаны на более сложных принципах, и приведенные здесь выражения для d_{14} выводятся как результат довольно громоздких вычислений. Наиболее употребительным в социальных науках является метод Уорда: чаще всего он дает весьма четкие и содержательно осмысляемые кластеры [1].

Как правило, указанные методы приводят к различным результатам — т.е. к различным дендрограммам и разбиениям на кластеры. Однако для данных, представленных в рассматриваемом примере (как для 2007, так и для 2011 годов) полученные дендрограммы для всех методов одинаковы.

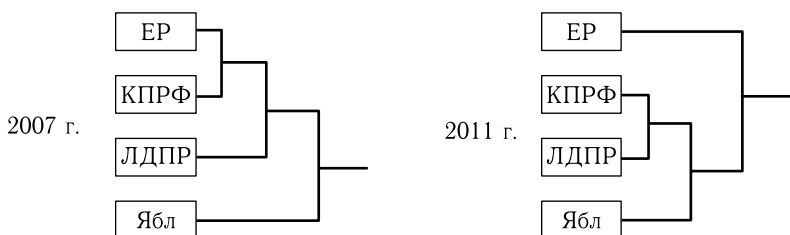


Рис. 2. Иерархическая структура близости программ 2007 и 2011 гг.

Дендрограммы для предвыборных партийных программ, построенные по табл. 2, изображены на рис. 2. Они показывают, что в предвыборную кампанию 2007 г. наиболее близкими были программы EP и KPRF, а наиболее отдаленной от других — программа Яблока. В кампанию 2011 г. произошла определенная перестройка: наиболее близкими были программы KPRF и LDPR, а наиболее отдаленной от других — программа EP.

Разбиение на кластеры проводится следующим образом: на дендрограмме отбрасываются соединительные линии, лежащие правее

некоторой выбранной вертикальной черты (или, что то же самое, заменяются нулями меры сходства, меньшими некоторого порогового значения). В результате граф разбивается на части, и отдельные его части соответствуют кластерам. Например, для данных 2007 г. можно выбрать в качестве порога меру сходства, равную 97, и получить три кластера: КПРФ + ЕР, ЛДПР, Яблоко. Либо можно отсечь меры сходства, меньшие, чем 91,5, и получить два кластера: КПРФ + ЕР + ЛДПР и Яблоко.

Изложенный алгоритм относится к классу агломеративных (или восходящих), т. е. постепенно «собирающих» кластеры из отдельных объектов. Иногда применяются также дивизивные (нисходящие) алгоритмы, в качестве исходного материала принимающие всю совокупность объектов, а кластеры возникают в результате ее деления на части.

1.3.3. Кластеризация в исследованиях парламентских голосований

Кратко разберем еще один пример применения кластерного анализа в социальных науках.

При голосовании в парламенте депутаты принимают решение в поддержку или против некоторого законопроекта, исходя как из «партийных», так и «региональных», «отраслевых» или иных соображений. Поэтому помимо формальной структуры парламента (фракции, комиссии и т. д.) можно выделить и неформальную, основанную на анализе результатов голосований. Определение неформальной структуры и является целью иерархической кластеризации депутатов на основе данных об их голосованиях. В данном примере рассматривается Сенат Палаты представителей США 2003 г.

Для определения близости двух сенаторов рассматриваются все голосования, в которых они оба принимали участие и проголосовали «да» или «нет» (другими словами, исключаются те голосования, при которых хотя один из этих двух сенаторов воздержался или не голосовал). Если один из сенаторов голосует «да», а другой — «нет», то это свидетельствует о различии их позиций, и наоборот — синхронное голосование свидетельствует о сходстве. На основе этих соображений выписывается матрица сходства, и затем проводится иерархическая кластеризация.

Поскольку численность Сената США составляла в 2003 г. 100 человек, то мы не приводим получаемую дендрограмму полностью. Ограничимся лишь ее «свернутым» вариантом, в котором выделены 7 кластеров (см. рис. 3).

Сенаторы, состоящие в Демократической партии, составляют два кластера: «Демократически большой», обозначенный Дб(44) и состоящий из 44 членов, а также имеющий всего трех сенаторов «Демократический малый» — Дм(3). Республиканская партия представлена четырьмя кластерами — большим Рб(38), средним Рс(11), малым Рм(2)

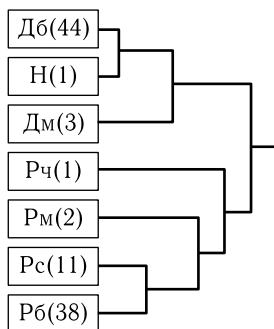


Рис. 3. Дендрограмма Сената США в 2003 г. с выделенными семью кластерами. Пояснения в тексте

и кластером, состоящим из одного сенатора Чефи, — Рч(1). Единственный независимый сенатор составляет отдельный кластер Н(1).

Строго говоря, кластер Рб(38) состоит из 37 республиканцев и одного демократа. За этим исключением, каждый кластер состоит из членов только одной партии. Дендрограмма показывает, что если выделить в Сенате меньшее количество кластеров, то данное свойство остается в силе — демократические кластеры объединяются с демократическими, а республиканские с республиканскими. Таким образом, разделение по партийному признаку оказывается основным.

Каждый кластер можно рассматривать в некотором приближении как отдельную фракцию, голосующую более или менее солидарно. Важный вопрос о том, как численность такой фракции связана с ее влиянием при голосовании, рассматривается ниже.

1.3.4. Кластерный анализ в контексте социологического исследования

Рассмотрим проблему классификации политических убеждений. Вероятно, самым «старым» в современной социологии и политологии является подход, основанный на дихотомии «либералы–консерваторы», задающей одномерную «лево–правую» шкалу. Позднее были развиты и многомерные модели представления политических убеждений. Однако данный, так называемый *пространственный подход*, обладает не только серьезными достоинствами, но и рядом недостатков и ограничений. Альтернативный подход основан на классификации политических убеждений путем кластеризации, не требующей, вообще говоря, построения евклидова (или иного, альтернативного ему) пространства.

Продемонстрируем этот подход на примере исследования, посвященного изучению политических убеждений граждан США в 80-х годах XX века. Для эмпирического исследования были разработаны 12 вопросов по темам, в течение долгих лет находящихся в сфере общественного внимания в США: объем государственных услуг, гарантии

занятости, расходы на оборону, аборт, отношение к СССР, место женщины в обществе, помощь меньшинствам и пр. Для каждого из 12 вопросов респондент должен был ответить, выбрав соответствующее шкальное значение. Например, для вопроса «Должны ли женщины играть равную роль с мужчинами, или же место женщины — дом?» шкала варьировалась от значения 1 (равную роль) до 7 (место женщины — дома) — см. таблицу 3. Для вопроса об отношении к абортам шкала варьировалась от 1 (никогда не запрещать) до 4 (никогда не разрешать). Для сопоставимости различных вопросов все шкалы были перенормированы на единичную длину, и в качестве расстояния между индивидами выбрано евклидово расстояние в 12-мерном пространстве, построенном на этих единичных шкалах. Тем самым была построена матрица расстояний между 483 респондентами, участвовавшими в опросе. Кластеризация проводилась по методу Уорда.

Таблица 3. Средние шкальные значения ответов на некоторые вопросы для кластеров 1 и 5

Вопрос	Минимальное значение шкалы	Среднее по кластеру 1	Среднее по кластеру 5	Максимальное значение шкалы
Должны ли женщины играть равную роль с мужчинами, или место женщины — дом?	1 (равную роль с мужчинами)	1,52	4,93	7 (место женщины — дома)
Следует ли уменьшить количество государственных услуг населению, чтобы уменьшить госрасходы?	1 (не следует уменьшать)	2,45	4,51	7 (следует уменьшить)
Должны ли быть разрешены или запрещены аборты?	1 (никогда не запрещать)	1,65	2,47	4 (никогда не разрешать)

В результате было получено 6 кластеров, основой для их интерпретации послужили средние значения ответов на вопросы. Например (см. табл. 3) для респондентов, входящих в кластер 1, характерны терпимость к аборт, положительное отношение к равноправию женщин с мужчинами и к высокой роли государства в оказании услуг населению. Поэтому данный кластер был осмыслен как либеральный.

В противоположность ему для кластера 5 характерно отрицательное отношение к женскому равноправию, к высокой роли государства в оказании услуг и к аборт — этот кластер был назван консервативным.

Оставшиеся четыре кластера получили имена «квазилибералы», «про-трудовики», «сторонники минимальной роли государства», «экономически умеренные». Были изучены расстояния между кластерами, в частности, показано, что наиболее близкими друг к другу являются кластеры «либералы» и «квазилибералы», а также «консерваторы» и «сторонники минимальной роли государства». Наконец, разбиение группы на кластеры было сопоставлено с иными (не входящими в круг двенадцати указанных вопросов) признаками. Так, в кластерах «консерваторы» и (в особенности) «сторонники минимальной роли государства» зафиксированы наиболее высокий уровень дохода, наибольшая поддержка президента Рейгана и Республиканской партии, наибольшая доля мужчин. Другой пример: поддержка организованных молитв в государственных школах наиболее сильна в кластерах «консерваторы» и «про-трудовики», наибольшую оппозицию этому представляет кластер «либералы». Перечисление закономерностей, установленных в данном исследовании, можно продолжить.

Таким образом, кластерный анализ является важным инструментом социологического исследования, позволяющим изучать структуру общества и обнаруживать нетривиальные связи между переменными, характеризующими индивидов (или иные изучаемые объекты).

Упражнения

1. Для данных табл. 2(а) провести подробно кластеризацию с построением матриц сходства по каждому из пяти методов
2. То же сделать для данных табл. 2(б).

Комментарии и библиографические примечания

1. Более подробно об основных методах кластерного анализа см. [1–3]. Современный анализ некоторых аспектов его применения в социальных науках см. в [4].

2. Методика анализа партийных программ (и прочих текстов, выражающих политические позиции) методом латентно-семантического анализа представлена в работах [5–7], там же получены изложенные здесь результаты.

3. При всем многообразии методов кластеризации и областей их применения иерархическая кластеризация традиционно занимает особое место. Так, уже в 1970-х годах, примерно в 75% опубликованных работ по кластеризации использовались иерархические алгоритмы (см. [8, 9]). Изложенная рекурсивная процедура, в которой для вычисления матрицы расстояний используется только предыдущая матрица, предложена в [10, 11].

4. Исследование голосований в Сенате Палаты представителей США проведено в [12], причем кластеризация является в этой работе

лишь одним из направлений определения неформальной структуры Сената.

5. Изложение примера с кластеризацией индивидов в соответствии с их политическими убеждениями проведено по оригинальной работе [13].

Литература

1. Ким Дж.-О., Мьюллер Ч.У., Клекка У.Р. Факторный, дискриминантный и кластерный анализ. 1989. М.: Финансы и статистика.
2. Мандель И.Д. Кластерный анализ. 1988. М.: Финансы и статистика.
3. Дюран Б., Одед П. Кластерный анализ. М.: «Статистика», 1977. — 128 с.
4. Fonseca J.R.S. Clustering in the field of social sciences: that is your choice // International Journal of Social Research Methodology. V. 16, Is. 5, 2013. P. 403–428.
5. Kornilina E., Petrov A. Research of political blogs by means of LSA // Intellectualization of information processing-2010 conf. proceedings. 2010. P. 508–511.
6. Петров А.П., Корнилина Е.Д. Исследование близости политических позиций методом латентно-семантического анализа // XII Международная научная конференция по проблемам развития экономики и общества. Книга 2. М.: Издательский дом высшей школы экономики. 2012. С. 334–342.
7. Корнилина Е.Д., Петров А.П. Латентно-семантический анализ предвыборных партийных программ на выборах в Государственную Думу 2007 и 2011 годов // М.: Вестник МГУ. Сер. 12: Политические науки, Издательство МГУ. 2013. № 2. С. 90–98.
8. Contreras P., Murtagh F. Methods of hierarchical clustering, arXiv:1105.0121, 2011.
9. Blashfield R.K., Aldenderfer M.S. The literature on cluster analysis Multivariate Behavioral Research 1978, 13. P. 271–295.
10. Lance G.N, Williams W.T. A general theory of classificatory sorting strategies. I. Hierarchical systems. Comput. J., 1967. V. 9, № 4, p. 373–380.
11. Lance G.N, Williams W.T. A general theory of classificatory sorting strategies. II. Clustering systems. Comput. J., 1967. V. 10, № 4, p. 271–276.
12. Jakulin A., Buntine W., La Pira T.M., Brasher H. Analyzing the U.S. Senate in 2003: Similarities, Clusters, and Blocs // Political Analysis (2009) 17(3): 291–310.
13. Fleishman J.A. Types of Political Attitude Structure: Results of a Cluster Analysis // The Public Opinion Quarterly. V. 50, № 3 (Autumn, 1986), pp. 371–386.
14. Лобазова М.В. Результаты применения кластерного анализа к исследованию этнополитического конфликта в Республике Дагестан и на Северном Кавказе // Математическое моделирование социальных процессов. Вып. 3. Под ред. А. П. Михайлова. М. Изд-во МГУ. 2001. С. 171–179.

1.4. Введение в теорию коллективного выбора

Предметом настоящего раздела является теория коллективного выбора, рассматривающая задачу построения определенного коллективного мнения на основе мнений индивидуальных. Такая проблема возникает не только на политических выборах, но также при экспертном оценивании проектов, голосовании в совете директоров крупной компании, на художественном или ученом совете и пр. Фундаментальная проблема состоит в разработке теории, на основе которой можно строить максимально справедливые, оптимальные избирательные системы. Математическая теория коллективного выбора разрабатывается на протяжении более чем полувека, и она оказала значительное влияние на развитие социологической и политической мысли.

Далее здесь после вводных примеров проводится математическая формализация указанной проблемы, рассматриваются правила большинства и парадокс Кондорсе. Материал иллюстрируется подробно разобранным примером, показывающим, как специфика правил голосования повлияла на обсуждаемый в 1956 г. в конгрессе США законопроект о предоставлении школам дополнительной финансовой помощи. Затем рассматривается правило Борда и его практическое применение в политической практике. Наконец, в доступной форме излагается основополагающий результат теории коллективного выбора — теорема Эрроу.

1.4.1. Вводные примеры

В качестве введения в проблематику укажем на разнообразие существующих в сегодняшнем мире избирательных систем, перечислив некоторые из них.

1. Правило, по которому проходят выборы депутатов ГД РФ в одномандатных округах: голосование проводится в один тур, победителем объявляется кандидат, набравший большинство голосов (так называемое *простое* большинство, в отличие от абсолютного большинства, которое составляет 50% голосов плюс один голос).

2. Правило, по которому происходят выборы на должность президента РФ. Голосование происходит в два тура (или в один, если кто-либо из кандидатов наберет в первом туре абсолютное большинство), во второй тур проходят два кандидата, и победителем объявляется кандидат, набравший во втором туре простое большинство.

3. Голосование по доверенности (система Electoral College), по которому проходят выборы президента США. На первой стадии выборов в каждом из штатов избиратели выбирают простым большинством список выборщиков (как правило, каждая политическая партия представляет свой список). На второй стадии выборщики простым большинством выбирают президента.

Заметим, что применение правил 2 и 3 иногда дает разный результат. Например, на президентских выборах в 2001 г. в США большинство избирателей отдали свои голоса за У. Клинтона, однако

применение правила 3 привело к победе Дж. Буша. Этот пример показывает, что политический результат выборов зависит не только от распределения голосов избирателей, но и от конкретного алгоритма, с помощью которого эти многочисленные индивидуальные мнения агрегируются в одно коллективное мнение, отражающее позицию всех избирателей «в целом».

4. Мажоритарная система: на парламентских выборах страна разбивается на одномандатные округа, в каждом из которых выборы депутата проходят по правилу 1 или 2.

5. Пропорциональная система: на парламентских выборах страна представляется как один многомандатный округ. Избиратели голосуют за партийные списки, депутатские мандаты распределяются между партиями пропорционально числу набранных голосов.

Принято считать, что мажоритарная система обеспечивает более тесный контакт между депутатом и его избирателями, в то время как пропорциональная система лучше защищает права политического меньшинства. Компромиссом между ними является система 6.

6. Мажоритарно-пропорциональная система, применяемая в РФ: половина депутатов ГД избирается по мажоритарному принципу, другая половина — по пропорциональному.

Применение правил 4, 5 и 6 может приводить к радикально различным результатам. Например, на выборах (по правилу 6) в ГД РФ в 1993 г. Либерально-демократическая партия победила на выборах по партийным спискам, но не получила ни одного мандата по мажоритарной системе. При проведении выборов по правилу 4 ЛДПР не была бы представлена в Думе вообще, а по правилу 5 — получила бы парламентское большинство.

7. Многомандатные выборы, при которых избиратель может проголосовать за нескольких (а не одного) кандидатов. Победителями объявляются несколько кандидатов, набравших наибольшее количество голосов. По такому правилу проходили выборы в Совет Федерации РФ в 1993 г. — избиратель голосовал за двух кандидатов, и в каждом округе объявлялось по 2 победителя. Однако наибольшее практическое применение эта схема имеет при выборах совета директоров в крупных акционерных обществах: держатель большого пакета акций распределяет свои голоса между несколькими кандидатами в совет директоров так, чтобы иметь в нем максимально возможное влияние.

8. Выборы Папы Римского кардинальским конклавом и определение Олимпийским комитетом города, который будет принимать Олимпиаду, проводятся в несколько туров, пока один из кандидатов не наберет более половины голосов. По итогам каждого тура отсеивается один кандидат, набравший минимальное количество голосов. Например, если в выборах принимают участие пять кандидатов, то победитель может определиться как в первом, так и в четвертом туре.

Мы перечислили далеко не все правила определения результатов коллективного выбора, уделив наибольшее внимание схемам, принятым в политической практике. Особые способы выбора применяются, например, при оценке тех или иных проектов группой экспертов, при выборе города (хозяина Олимпийских игр) и в ряде других практических ситуаций. Каждая из систем имеет свои преимущества и недостатки, а вскрыть их часто помогает анализ математическими средствами.

При описании различных правил голосования мы неявно подразумевали задачу выбора системы, наиболее справедливой в каждом конкретном случае. Другой класс важных задач для математического анализа — это выбор оптимальной стратегии для участника выборов. Простейшая, но стандартная проблема может быть проиллюстрирована следующим примером. Пусть совет директоров предприятия состоит из 10 человек. Очевидно, акционер, имеющий 26% акций, имеет возможность провести в совет двух своих кандидатов. Может ли он провести в совет своего третьего представителя? Какой стратегии ему следует придерживаться при голосовании?

Отметим также, что существует ряд научных и практических проблем, связывающих выборные процессы с другими аспектами социальной жизни. Например, при исследовании распределения власти в иерархии (см. Главу 4) важную роль играет наличие или отсутствие альтернативы между двумя (или более) распределениями власти, а также возможность для общества сделать выбор между ними.

1.4.2. Формализация проблемы

В настоящем пункте проблема коллективного выбора формализуется с математической точки зрения и вводятся основные определения. Теория коллективного выбора имеет применения как в социально-политических науках, так и в экономике, однако, имея в виду общий контекст настоящей книги, мы будем подразумевать политический выбор и иногда пользоваться политической терминологией.

Сама возможность выбора предполагает наличие некоторого множества альтернативных вариантов (далее — альтернатив или, с учетом политического контекста, кандидатов, причем мы будем употреблять эти слова как синонимы), а процедура выбора — это их полное или частичное упорядочивание.

Итак, рассмотрим группу равноправных индивидов (избирателей), каждый из которых имеет свои предпочтения на множестве альтернатив (кандидатов), т. е. для любых двух кандидатов он может указать более предпочтительного или считать их равноценными. Требуется исходя из этих индивидуальных предпочтений построить упорядочение кандидатов, отражающее мнение всей группы в целом (результат выборов).

Общее количество голосующих индивидов (избирателей) обозначим через n , множество этих индивидов — через $Q = \{q_1, \dots, q_n\}$,

возможных кандидатов будем обозначать буквами a, b, c, \dots , а множество этих вариантов — $S = \{a, b, c, \dots\}$.

Тот факт, что i -й избиратель предпочитает кандидата a кандидату b , запишем в виде $aP_i b$ (сильное предпочтение) если он считает их равноценными, то — $aI_i b$ (безразличие). Наконец, если данный индивид предпочитает кандидата a кандидату b либо считает их равноценными, запишем в виде $aR_i b$ (слабое предпочтение).

Сформулируем две основополагающие аксиомы

Аксиома I. Для любого избирателя i и для любых x, y имеет место хотя бы одно из отношений $xR_i y, yR_i x$

Аксиома II. Для любого избирателя i и для любых x, y, z отношения $xR_i y, yR_i z$ влекут за собой $xR_i z$.

Первая из этих аксиом требует, чтобы каждый индивид мог выразить слабое предпочтение для любой пары кандидатов, вторая — чтобы это отношение было транзитивным.

Можно доказать, что выполнении аксиом I, II отношения сильного упорядочения и безразличия также транзитивны.

Слабое предпочтение является обобщением отношения «больше или равно», применяемого к действительным числам. Оно позволяет провести слабое упорядочение (которое мы будем называть просто упорядочением) каждым избирателем всех элементов множества S в том смысле, что если выполнены аксиомы I, II, то для каждого избирателя можно построить его индивидуальную ранжировку всех кандидатов от наиболее до наименее предпочитаемых, при этом в пределах ранжировки некоторые пары могут быть связаны отношением сильного предпочтения, а другие — отношением безразличия. Например, если $S = \{v, x, y, z\}$, то может быть $xP_i y P_i v I_i z$.

Сильное предпочтение обобщает отношения «больше», применяемое к действительным числам. Если выполнены аксиомы I, II и никакие два кандидата не являются для избирателя безразличными, то оно задает сильное предпочтение на множестве S .

Подчеркнем, что термин «отношение предпочтения» применяется к паре кандидатов, а «упорядочение» — ко всему множеству S .

Упорядочения удобно представлять, выписывая элементы множества S в столбец по убыванию предпочтительности сверху вниз.

Пример 1. Для множества $S = \{x, y, z\}$ упорядочение $yP_1 x I_1 z$ может быть представлено в виде

R_1
y
x, z

Поставив рядом упорядочения, сделанные всеми избирателями, получим набор индивидуальных упорядочений (R_1, \dots, R_n) всех членов группы избирателей.

Пример 2. Для системы из трех избирателей и множества кандидатов $A = \{x, y, z, u\}$ один из возможных наборов индивидуальных упорядочений имеет вид

R_1	R_2	R_3
x	u	u
y	y	y, z, x
u, z	z	
	x	

Основной вопрос теории коллективного выбора звучит следующим образом: как наиболее разумным и справедливым образом сформулировать мнение всей группы «в целом» на основе высказанных избирателями индивидуальных предпочтений?

Сформулируем следующее определение.

Функцией общественного благосостояния называется правило, которое каждому набору индивидуальных упорядочений ставит в соответствие определенное коллективное упорядочение.

Название «функция общественного благосостояния» связано с тем, что изначально в этой теории рассматривались упорядочения не кандидатов на избираемые должности, а так называемые *социальные состояния*, связанные с экономическим благосостоянием отдельных индивидов и общества в целом.

Поясним данное выше определение двумя примерами.

Пример 3. Рассмотрим группу из двух избирателей R_1, R_2 , упорядочивающих двух кандидатов x, y . Каждый избиратель имеет три возможности: либо предпочесть кандидата x , либо предпочесть кандидата y , либо оценить их одинаково. Следовательно, всего возможны 9 наборов индивидуальных предпочтений (см. рис. 1). Для коллективного упорядочения (область значений функции общественного благосостояния) есть те же три возможности: на выборах побеждает либо x , либо y , либо ничья. Каждому элементу области определения должен быть сопоставлен ровно один элемент области значений. Различным образом проводя это сопоставление, получаем различные функции общественного благосостояния. Так (см. рис. 1) построим графическое изображение *правила простого большинства*: победителем объявляется тот кандидат, который чаще других занимает первое место в индивидуальных упорядочениях. Применительно к данному случаю это означает: если оба избирателя голосуют за одного кан-

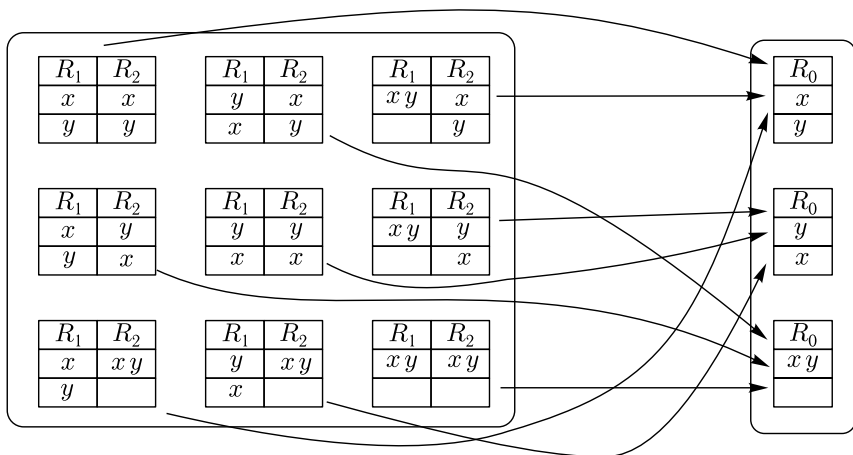


Рис. 1. Правило простого большинства для случая двух избирателей и двух кандидатов

дидата, то он побеждает; если за разных, то выборы заканчиваются вничью.

Рассмотрим, например, три набора индивидуальных упорядочений из верхней части рис. 1. В первом случае оба избирателя голосуют за кандидата x , поэтому он является победителем, соответствующая стрелка направлена к верхнему элементу области значений. Во втором наборе упорядочений первый избиратель голосует за y , второй — за x ; следовательно, результатом выборов является ничья. В третьем наборе из верхней строки один избиратель голосует за x , второй воздержался; следовательно, побеждает x .

Проводя аналогичные рассуждения для оставшихся шести наборов индивидуальных предпочтений, достраиваем все стрелки на рис. 1. Тем самым получено графическое изображение правила простого большинства.

Пример 4 (правило диктатора). Из множества Q всех избирателей выделим некоторого избирателя q_i (будем называть его диктатором) и положим $R_0 = R_i$, т.е. индивидуальное упорядочение, сделанное диктатором, примем в качестве коллективного упорядочения. Например, если для случая двух избирателей и двух кандидатов второй избиратель является диктатором, то функция общественного благосостояния графически изображается на рис. 2.

С чисто формальной, математической точки зрения правило диктатора не содержит никаких противоречий и парадоксов. Однако оно не удовлетворяет понятиям о «разумности и справедливости», предъявляемым обществом к избирательной системе. Поэтому в течение довольно долгого времени в XIX и XX веках усилия теоретиков были направлены на поиск логически непротиворечивого способа построения

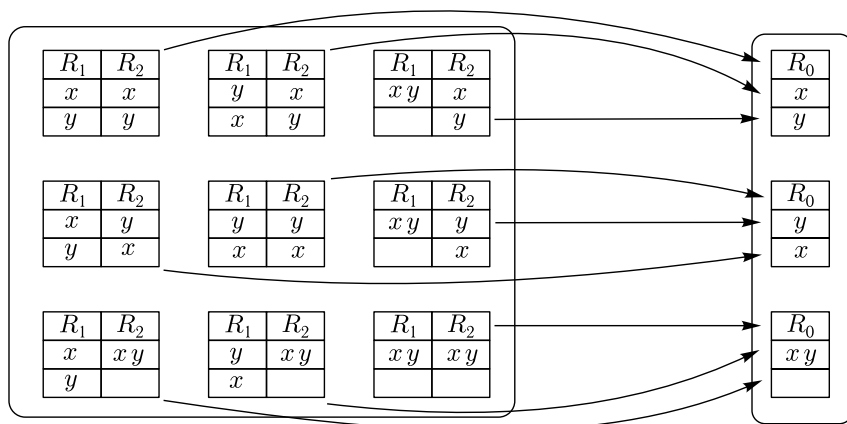


Рис. 2. Правило диктатора для случая двух избирателей и двух кандидатов

максимально разумной и справедливой функции общественного благосостояния.

1.4.3. Правила большинства и парадокс Кондорсе

С античных времен существует консенсус относительно того, что для занятия выборной должности кандидат должен получить большинство голосов избирателей. Но в то же время оказалось, что нет единого подхода к тому, как должно определяться это большинство.

Например, одним из наиболее часто применяемых является *правило простого большинства*, о котором говорилось в примере 3 и которое в общем случае (т. е. для произвольного количества кандидатов) имеет следующий вид. В коллективном упорядочении кандидат x расположен выше кандидата y тогда и только тогда, когда в индивидуальных упорядочениях избирателей он стоит на первом месте большее количество раз, чем кандидат y . К числу достоинств этой функции общественного благосостояния относится простота и легкость реализуемости на массовых выборах: избирателю требуется всего лишь указать в бюллетене одного кандидата, которого он поддерживает; избранным считается кандидат, набравший большинство голосов. Поэтому данное правило применяется во многих странах на парламентских выборах в одномандатных округах. В частности, по принципу простого большинства избраны действующие палата общин Великобритании, обе палаты Конгресса США, половина депутатов Верховной Рады Украины и т. д.; по нему же в различные годы избиралась половина или все депутаты Государственной Думы.

Однако правило простого большинства обладает серьезным недостатком, который мы продемонстрируем на следующем примере.

Пример 5. Рассмотрим систему из 9 избирателей, выбирающих из 3 кандидатов. Пусть набор индивидуальных предпочтений выглядит следующим образом:

$R_1 - R_4$ (4 избирателя)	$R_5 - R_7$ (3 избирателя)	R_8, R_9 (2 избирателя)
x	z	y
y	y	z
z	x	x

По правилу простого большинства победителем является кандидат x , набравший 4 голоса. Менее всего голосов набрал кандидат y . В то же время большинство избирателей (пятеро) в своих индивидуальных упорядочениях ставят кандидата y выше, чем кандидат x . Кроме того, те же пятеро избирателей предпочитают кандидата z кандидату x . Таким образом, победителем оказался кандидат, который для большинства избирателей проигрывает любому другому кандидату при попарном сравнении.

Этот пример показывает, что когда избирателей и кандидатов больше, чем по двое, то понятие «большинства» становится неоднозначным. Проблемы, однако, не возникает, если один из кандидатов набрал более половины голосов, так как в этом случае при попарном сравнении с любым другим кандидатом он является более предпочитаемым для большинства избирателей. Следовательно, проблемными являются лишь ситуации, когда ни один из кандидатов не набрал более половины голосов.

В связи с этим рассмотрим *правило абсолютного большинства*, предполагающее, что в таком случае победитель объявляется лишь после второго тура, в который проходят два кандидата, набравшие наибольшее количество голосов в первом туре. Это правило обладает недостатком того же рода. Чтобы показать это, вернемся к примеру 5. В первом туре кандидаты x, y, z набирают соответственно 4, 3 и 2 голоса. Во второй тур выходят x, z , где побеждает z , набрав 5 голосов избирателей против 4. При этом кандидат y , который в парных сравнениях побеждает всех конкурентов, даже не выходит во второй тур. Таким образом, данное правило обладает следующим недостатком: при наличии кандидата, побеждающего во всех парных сравнениях, этот кандидат может проиграть выборы по системе абсолютного большинства. Тем не менее эта система также используется достаточно широко: например — выборы президента в России и Франции.

Перейдем к следующему варианту правила большинства, которое называется *методом (правилом) Кондорсе*. В соответствии с ним кандидат x получает в коллективном упорядочении более высокое место, чем y , тогда и только тогда, когда большинство избирателей оценивает x выше, чем y .

В этом случае возникает новая трудность, на этот раз связанная с нетранзитивностью. Рассмотрим ее на примере.

Пример 6. Пусть в системе из трех избирателей и множества кандидатов $A = \{x, y, z\}$ набор индивидуальных предпочтений имеет вид

R_1	R_2	R_3
x	y	z
y	z	x
z	x	y

Большинство избирателей ранжируют x выше y , следовательно (по правилу Кондорсе), в групповом упорядочении также x должен стоять выше y . Аналогично: так как большинство избирателей ставят y выше z , а z — выше x , то в групповой ранжировке будет y выше z , а z выше x ; следовательно, y выше, чем x . Таким образом, для данного набора индивидуальных предпочтений одновременно должно быть x выше, чем y , а y выше, чем x . Полученное противоречие доказывает, что *применение правила Кондорсе не всегда позволяет построить в данном примере коллективное упорядочение.*

Этот пример иллюстрирует так называемый *парадокс Кондорсе: применение правила Кондорсе не всегда позволяет построить коллективное упорядочение.*

Другими словами: на основе метода Кондорсе нельзя построить функцию общественного благосостояния, определенную на всем множестве наборов индивидуальных предпочтений. Исключение составляет ситуация, когда множество S состоит всего из двух элементов — в этом случае проблемы нетранзитивности не возникает.

1.4.4. Парадокс Кондорсе при бинарной процедуре

На практике парадокс Кондорсе приводит к наиболее явным проблемам при голосовании по так называемой *бинарной процедуре*, суть которой заключается в следующем. При обсуждении законопроектов в парламенте депутаты имеют право вносить в проект поправки. Рассмотрим самую простую ситуацию, когда к закону предлагается ровно одна поправка. В этом случае голосование может иметь один из трех исходов: принятие закона в его первоначальном виде (этот вариант мы обозначим через *yes*), принятие закона с поправкой (*am* — от *amendment*), отклонение закона (*no*). Этих трех вариантов вполне достаточно для возникновения парадокса Кондорсе.

Бинарная процедура заключается в том, что депутатам предлагается голосовать два раза, выбирая каждый раз всего из двух вариантов. При первом голосовании депутаты решают, включать ли в законопроект поправку. При втором голосовании решается судьба закона — принимать ли его (с поправкой или без нее, в зависимости от результатов

первого тура) или не принимать. Таким образом, множества вариантов для каждого из голосований состоят всего из двух элементов:

$$A_1 = \{yes, am\},$$

$$A_2 = \{yes \text{ или } am, no\}.$$

Бинарная процедура является обычной парламентской практикой, несмотря на то, что парадокс Кондорсе может привести к искажению воли голосующих парламентариев.

Разберем это явление на следующем историческом примере. В 1956 г. в конгрессе США обсуждался законопроект о предоставлении школам дополнительной финансовой помощи. В то время в некоторых школах на юге США еще существовала расовая сегрегация, т.е. раздельное обучение белых и черных учеников. Поэтому в ходе обсуждения законопроекта противниками сегрегации была предложена поправка, суть которой сводилась к тому, чтобы предоставить дополнительное финансирование только школам с совместным обучением белых и черных (так называемая *поправка Пауэлла*). По своему отношению к этому закону конгрессмены разделились на три примерно равные группы. В первую из них входили в основном республиканцы — они были против закона в любом его виде, особенно в первоначальном (так как помощь всем школам означает большие государственные расходы, чем помощь только части школ). Вторая группа конгрессменов (в основном демократы из северных штатов, в т.ч. сам Пауэлл) поддерживала поправку, но в случае ее поражения была готова проголосовать за закон в его первоначальном виде. Демократы из южных штатов (третья группа) поддерживали дополнительное финансирование школ, но поправка Пауэлла была для них неприемлема, так как она фактически предоставляла северным школам преимущество перед южными. Таким образом, набор «индивидуальных» предпочтений по данному вопросу имел следующий вид (мы взяли слово «индивидуальных» в кавычки, так как «индивид» в данном случае — это группа конгрессменов):

Республиканцы	Северные демократы	Южные демократы
<i>no</i>	<i>am</i>	<i>yes</i>
<i>am</i>	<i>yes</i>	<i>no</i>
<i>yes</i>	<i>no</i>	<i>am</i>

Нетрудно видеть, что имеет место парадокс Кондорсе — большинство конгрессменов ранжирует

$$no \text{ выше, чем } am,$$

$$am \text{ выше, чем } yes,$$

$$yes \text{ выше, чем } no.$$

События в Конгрессе развивались следующим образом. В соответствии с бинарной процедурой на первое голосование была выставлена поправка Пауэлла ($A_1 = \{am, yes\}$). Она была принята голосами республиканцев и северных демократов. Соответственно на втором голосовании альтернатива имела вид $A_2 = \{am, no\}$, и поправленный законопроект был провален голосами республиканцев и южных демократов.

Проанализируем этот пример. Большинство конгрессменов поддерживало дополнительное финансирование школ (*yes* выше *no*); если бы поправка Пауэлла не была выдвинута, то закон был бы принят. Однако поправка раздробила голоса демократов, и это привело к отклонению законопроекта (окончательный результат голосования — *no*). Именно свойственный правилу большинства эффект нетранзитивности привел к такому исходу. Интересно, что после того, как демократы уладили внутрипартийные разногласия, в 1957 г. законопроект был вновь представлен конгрессу. На этот раз никаких поправок предложено не было, и при альтернативе $A = \{yes, no\}$ закон был принят (*yes*).

Итак, в данном примере внесение поправки Пауэлла привело к отклонению законопроекта, который был бы принят, не будь она внесена. Такие поправки называются поправками-убийцами. Это довольно редкое явление: например, при исследовании голосований в американском Конгрессе за 1953–2004-е гг. в работе [3] было выявлено всего пять поправок-убийц.

Как следует из рассуждений и выкладок настоящего пункта, ни одно из правил большинства не может служить основой для построения коллективного упорядочения. В этой ситуации могут быть предложены два пути.

Первый из них использует идею, знакомую из курса высшей математики. Как известно, нельзя построить функцию, обратную к функции $y = \sin x$, заданной на всей области определения. Поэтому, чтобы ввести понятие арксинуса, необходимо сначала выделить *часть области определения* функции $y = \sin x$, на которой эта функция монотонна, и лишь после этого удастся построить обратную функцию с помощью известной процедуры. Тем же приемом можно воспользоваться при построении функции общественного благосостояния, а именно: выделить некоторое подмножество множества всех наборов индивидуальных упорядочений и взять это подмножество в качестве области определения функции общественного благосостояния. С помощью такого трюка можно избежать парадокса Кондорсе и построить функцию общественного благосостояния на основе любого из правил большинства. Именно это происходит на обычных политических выборах, например, в большинстве стран (в частности, в России) избирателю предлагается выбирать лишь из тех упорядочений, при которых один кандидат стоит на первом месте, а все остальные кандидаты

равноценны. Политические издержки этого подхода очевидны — ограничение области определения на практике означает ограничение возможности выбора голосующих индивидов, вследствие чего происходит искажение воли избирателей.

Второй путь — это отказ от правил большинства в пользу какого-либо другого принципа определения победителя выборов. В частности, к этому направлению относится правило Борда, которому посвящен следующий раздел.

1.4.5. Правило Борда

Рассмотрим некоторый набор индивидуальных упорядочений. Обозначим через $B_i(x)$ число кандидатов, расположенных ниже кандидата x в ранжировке R_i , и назовем число

$$B(x) = \sum_{i=1}^n B_i(x)$$

числом Борда кандидата x . Правило Борда определяется следующим образом: в коллективном упорядочении кандидат x стоит выше кандидата y тогда и только тогда, когда $B(x) > B(y)$.

Пример 7. Пусть в выборах участвуют четыре кандидата, и набор индивидуальных упорядочений имеет вид

P_1	P_2	P_3
x	y, u	y
y	x, z	x
z		z, u
u		

Построим коллективное упорядочение по правилу Борда. Первый избиратель ранжирует трех кандидатов ниже, чем x ; второй — ноль, третий — двух. Следовательно, число Борда для кандидата x равно

$$B(x) = 3 + 0 + 2 = 5.$$

Для остальных кандидатов имеем

$$B(y) = 2 + 2 + 3 = 7,$$

$$B(z) = 1 + 0 + 0 = 1,$$

$$B(u) = 0 + 2 + 0 = 2.$$

Следовательно, коллективное упорядочение имеет вид $R_{\text{Борда}}$ (см. левую колонку в таблице ниже). Интересно сравнить $R_{\text{Борда}}$ с коллективным упорядочением по Кондорсе для тех же данных. Нетрудно видеть,

что результаты, полученные с помощью того и другого правил, близки, но все же различны:

R_{Borda}	$R_{\text{Condorcet}}$
y	y
x	x
u	u, z
z	

Пример 8. Камнем преткновения для правила Кондорсе стал набор индивидуальных предпочтений из примера 5. Вычислим для этих данных коллективное упорядочение по Борда:

$$B(x) = 2 + 0 + 1 = 3,$$

$$B(y) = 1 + 2 + 0 = 3,$$

$$B(z) = 0 + 1 + 2 = 3.$$

Таким образом, ни один из кандидатов не имеет предпочтения перед другими, т. е. правило Борда дает единственно разумный результат для данных из примера 5.

Иногда правило Борда применяется в политической практике. Так, в островном тихоокеанском государстве Науру, занимающем площадь около 21 кв. км, модифицированное правило Борда применяется на выборах в парламент. Вся страна разбита на 8 многомандатных избирательных округов. От шести из них в парламент избираются по два депутата, от одного — три, еще от одного — четыре депутата. В пределах каждого округа победители определяются следующим образом. Избиратель заполняет бюллетень, ранжируя всех кандидатов. При подсчете голосов за первое место в бюллетене кандидату начисляется 1 балл, за второе — $1/2$, за третье — $1/3$ балла и т. д., затем данные по всем избирателям округа суммируются. Тем самым получается ранжировка всех кандидатов, в соответствии с которой определяются победившие два, три или четыре кандидата. Таким образом, «метод Науру» отличается от классического метода Борда двумя моментами: количеством победителей и методом начисления баллов.

Применяя правило Борда, всегда можно получить мнение группы «в целом» исходя из индивидуальных упорядочений. Другими словами, на основе метода Борда всегда можно построить функцию общественного благосостояния, определенную на всем множестве наборов индивидуальных упорядочений. Однако и оно имеет важный недостаток, который мы покажем на следующем примере.

Пример 9. Рассмотрим следующий набор индивидуальных упорядочений:

R_1	R_2	R_3	R_4	R_5
x	x	x	x	y
y	y	y	y	z
u	u	u	u	u
v	v	v	v	v
w	w	w	w	w
z	z	z	z	x

По правилу простого большинства (или по Кондорсе) побеждает кандидат x , который является наиболее предпочтительным для четверых избирателей. Этот результат выглядит наиболее разумным. Однако правило Борда дает другой исход выборов:

$$B(x) = 5 + 5 + 5 + 5 + 0 = 20,$$

$$B(y) = 4 + 4 + 4 + 4 + 5 = 21.$$

Таким образом, по правилу Борда победителем должен быть объявлен кандидат y .

Данный пример показывает, что метод Борда, как и различные варианты правила большинства, не всегда приводит к результату выборов, соответствующему воле избирателей. Рассмотренное в начале раздела правило диктатора с политической точки зрения еще хуже. Так существует ли вообще «справедливое» правило определения победителя выборов? Ответ на этот вопрос дал в начале 1950-х гг. Кеннет Эрроу, удостоенный за это Нобелевской премии. Необходимо отметить, что появление теоремы Эрроу — этого важнейшего результата политической науки — стало возможным только благодаря анализу политологической проблемы математическими методами.

1.4.6. Теорема Эрроу

Данный (в некотором смысле итоговый) пункт посвящен иному подходу к проблеме. Вместо того, чтобы предлагать все новые функции общественного благосостояния, проведем математическую формализацию минимальных требований к избирательной системе — они называются *аксиомами Эрроу*.

Первое требование называется *аксиомой универсальности*: функция общественного благосостояния должна быть определена на всем множестве наборов индивидуальных упорядочений (а не его части). Из рассмотренных выше функций общественного благосостояния этой аксиоме не удовлетворяет метод Кондорсе.

Перейдем ко второму требованию — *аксиоме независимости от посторонних альтернатив*. Рассмотрим следующую ситуацию.

Пусть некоторого наблюдателя на выборах интересует лишь борьба между несколькими кандидатами, образующими подмножество A_1 множества всех кандидатов A .

Он контролировал предпочтения между этими кандидатами в индивидуальных упорядочениях и в коллективном упорядочении (не обращая внимания на других кандидатов). Затем выяснилось, что счетная комиссия ошиблась с голосами других (не входящих в A_1) кандидатов. Суть аксиомы независимости заключается в том, что тогда коллективное упорядочение кандидатов из A_1 не должно поменяться. Другими словами: функция общественного благосостояния может считаться «справедливой», только если в исправленном коллективном упорядочении кандидаты из A_1 стоят в том же порядке, что первоначально. Это требование может быть сформулировано в виде следующей аксиомы.

Аксиома независимости от посторонних альтернатив. Пусть A_1 — произвольное подмножество множества A . И пусть набор индивидуальных упорядочений изменился так, что индивидуальные предпочтения относительно элементов из A_1 не изменились. Тогда функция общественного благосостояния должна определять, что предпочтения относительно элементов из A_1 в измененном коллективном упорядочении не изменились.

Заметим, что «несправедливость» правила Борда связана именно с тем, что для него не выполняется аксиома независимости. Чтобы это показать, вернемся к примеру 8, и в качестве множества A_1 выберем $A_1 = \{x, y\}$. Для элементов из A_1 индивидуальные предпочтения выглядят следующим образом: $xP_1y, xP_2y, xP_3y, xP_4y, yP_5x$, а в групповом предпочтении имеем yPx . Теперь представим себе, что первый избиратель «передумал» и решил оценить кандидатов w и z выше, чем ранее, поставив их соответственно на второе и третье места, так что теперь индивидуальное упорядочение избирателя P_1 имеет вид x, w, z, y, u, v . Предпочтения относительно элементов из A_1 не изменились — в частности, по-прежнему выполняется xP_1y . Однако групповая ранжировка изменилась: результат применения правила Борда теперь дает $B(x) = 20, B(y) = 19$, и в коллективном упорядочении получится xP_0y . Таким образом, правило Борда не удовлетворяет аксиоме независимости от посторонних альтернатив.

Следующая аксиома называется *принципом единогласия (или принципом Парето)*. В соответствии с ней, если во всех индивидуальных упорядочениях кандидат x стоит выше кандидата y , то так же должно быть и в коллективном упорядочении. Этой аксиоме не удовлетворяют правила простого и абсолютного большинства.

Наконец, последняя аксиома Эрроу говорит о том, что правило голосования должно исключать существование диктатора. Напомним, что если существует избиратель такой, что для любого набора индивидуальных упорядочений коллективное упорядочение совпадает с его индивидуальным упорядочением, то он называется диктатором.

Аксиома отсутствия диктатора: функция общественного благосостояния должна определять, что в группе нет диктатора.

Перечисленные аксиомы позволяют сформулировать основную теорему теории общественного выбора.

Теорема Эрроу. Пусть множество кандидатов A содержит не менее трех элементов, количество голосующих индивидов $n \geq 2$. Тогда функции группового выбора, удовлетворяющей аксиомам универсальности, независимости от посторонних альтернатив, единогласия и отсутствия диктатора, не существует.

Таким образом, не существует идеальной избирательной системы, позволяющей, с одной стороны, точно выразить избирателю свои предпочтения, с другой стороны, справедливым образом составить из этих предпочтений групповую ранжировку. На практике к этим проблемам добавляются дороговизна выборов, некомпетентность избирателей, возможности манипулирования общественным мнением, фальсификации и еще ряд других. Поэтому выбор схемы голосования в каждом конкретном случае является важной общественно-политической задачей.

При этом всегда остается простор для выбора: прямое голосование на президентских выборах (как в РФ) или голосование по доверенности (как в США)? На парламентских выборах: мажоритарная система, или пропорциональная, или мажоритарно-пропорциональная? Вероятно, ответ на этот вопрос должен определяться местными условиями, связанными с политическими традициями и реалиями той или иной страны. Не исключено, что в некоторых случаях предпочтение следует отдавать той избирательной схеме, которая более других устойчива по отношению к фальсификациям.

Отметим, наконец, что схемы, явно неприемлемые для политических выборов, могут оказаться полезными в практических ситуациях с небольшим количеством избирателей — при экспертном оценивании проектов или при голосовании в совете директоров крупной компании.

Упражнения

1. Пусть в выборах участвуют четыре кандидата, а набор индивидуальных упорядочений имеет вид

P_1	P_2	P_3
x	y, u	x
y, u	x, z	z
z		y, u

Построить коллективное упорядочение по правилу Борда.

2. Некоторое государство состоит из трех регионов, в каждом из которых живут по три избирателя. В президентских выборах участвуют два кандидата. Построить для этого государства пример ситуации, в которой прямые выборы и голосование по доверенности (система Electoral College) дают различные результаты.

Комментарии и библиографические примечания

1. Теорема Эрроу была сформулирована и доказана в 1951 г. в основополагающей работе [1].

2. Парадокс Кондорсе был впервые описан его книге [2], вышедшей в 1785 г. и исторически представляющей собой одну из первых работ, в которой математические методы использовались при изучении социально-политических процессов.

3. В качестве дополнительного материала по теме данного раздела можно рекомендовать Главу 4 учебника [4].

Литература

1. Arrow K.J. Social Choice and Individual Values, Yale University Press, 1951.
2. Condorcet N. Essai sur l'application de l'analyse a la probabilité des décisions rendues à la pluralité des voix. 1785.
3. Finocchiaro C.J., Jenkins J.A. In Search of Killer Amendments In the Modern U.S. House // Legislative Studies Quarterly, 2008. 33: 263–294. doi: 10.3162/036298008784310984.
4. Алескеров Ф.Т., Хабина Э.Л., Шварц Д.А. Бинарные отношения, графы и коллективные решения. Издательский дом ГУ ВШЭ. М. 2006. 300 с.

1.5. Введение в теорию индексов влияния

Настоящий раздел посвящен вопросам определения мер влияния консолидированных групп в органах, принимающих решения голосованием. Такими группами являются фракции и иные депутатские объединения в парламентах, неформальные группы в ученых и художественных советах или, например, обладатели пакетов акций в акционерных обществах. Эти группы формируются на основе общих политических взглядов, принадлежности к определенным художественным течениям или бизнес-кланам и т. д.

Для достижения нужного результата при голосовании такие фракции (группы) могут входить в коалиции. Заметим, что в рамках данного научного направления коалицией называется объединение фракций, голосующих одинаково (тем самым не проводится различий между формальными и неформальными союзами фракций).

Коалиции (чаще всего) не образуются произвольным образом, а соответствуют определенному разрезу социальной реальности.

Например, структура парламента в той или иной степени отражает социальную структуру общества. Поэтому парламентские коалиции обычно формируются из идейно близких фракций, представляющих социально близкие общественные слои. Так, для умеренно левой фракции чаще бывает более предпочтительным оставаться в коалиции с крайне левой или центристской, чем с правой фракцией. Тем самым можно говорить о предпочтениях фракций по вступлению в коалиции.

В основе излагаемого ниже подхода к определению влияния фракций лежит идея о том, что власть — это возможность влиять на политические решения. Инструментом власти является здесь решение фракции о вступлении в коалицию с теми, кто «за», или с теми, кто «против». При этом влияние фракции зависит от количества голосов, которое она контролирует. В то же время эта зависимость не носит пропорциональный характер. Например, если в парламенте из 100 депутатов численности фракций составляют 49, 48 и 3 человека и для принятия решения нужен 51 голос, то (при условии, что в пределах каждой фракции парламентарии голосуют солидарно) все они облают равным влиянием: для принятия решения достаточно поддержки любых двух фракций из трех. В общем случае каждой фракции можно приписать определенное значение индекса влияния. Соответствующая теория является предметом настоящего раздела.

В первом пункте проведена необходимая формализация и вводится индекс Банцафа — наиболее простой, но весьма употребительный индекс влияния. Алгоритм иллюстрируется вычислением индексов Банцафа для фракций Верховной Рады Украины VI созыва. Во втором пункте рассмотрено так называемое *взвешенное голосование* — не применяемая ныне (по крайней мере, в парламентах) практика, отмена которой явилась прямым следствием развития теории индексов влияния. В качестве исторического примера рассмотрен Совет графства Нассау (штат Нью-Йорк, США). Третий пункт посвящен парадоксу перераспределения депутатов между фракциями, а четвертый — изучению индекса влияния, учитывающему предпочтения фракций по вступлению в коалиции.

1.5.1. Индекс Банцафа

Рассмотрим парламент, состоящий из N депутатов. Пусть для принятия решения по некоторому вопросу требуется не менее, чем q голосов, это число называется *квотой*. Например, для Государственной Думы РФ $N = 450$, квота составляет 300 голосов (при голосовании по конституционным проектам) и 226 голосов (по всем прочим вопросам).

Все депутаты распределены по n фракциям, численности которых составляют k_1, \dots, k_n человек, при этом некоторые фракции могут состоять из одного депутата. Всюду далее в этом разделе будем полагать, что каждая фракция голосует солидарно (т.е. все выходящие в нее депутаты голосуют одинаково).

Множество фракций, голосующих по некоторому вопросу одинаково, называется *коалицией*. Если коалиция обладает количеством голосов, необходимым для принятия решения, то она называется *выигрывающей*, в противном случае — *проигрывающей*.

В качестве примера рассмотрим Верховную Раду Украины VI созыва (2007–2012 гг.). Она состояла из $N = 450$ депутатов, распределенных по пяти фракциям:

- Партия регионов (ПР) — 175 депутатов,
- Блок Юлии Тимошенко (БЮТ) — 156 депутатов,
- Наша Украина — Народная Самооборона (НУНС) — 72 депутата,
- Коммунистическая партия Украины (КПУ) — 27 депутатов,
- Блок Литвина (БЛ) — 20 депутатов.

Как и в случае российского парламента, для большинства вопросов квота в Верховной Раде составляет $q = 226$ голосов, далее будем рассматривать именно этот случай.

Выигрывающими коалициями являются, например, {ПР, БЮТ} (331 голос) или {ПР, БЮТ, КПУ} (358 голосов). Коалиции {ПР, КПУ} (202 голоса) или {НУНС, БЛ} (92 голоса) являются проигрывающими. Самая большая выигрывающая коалиция включает все пять фракций Верховной Рады.

Подход, лежащий в основе индексе Банцафа, основан на идее о том, что значимость фракции определяется ее возможностями повлиять на исход голосований. В связи с этим введем следующее определение.

Пусть некоторая фракция состоит в выигрывающей коалиции, причем эта коалиция перестанет быть выигрывающей в случае выхода из нее данной фракции. Тогда эта фракция называется ключевой.

Рассмотрим, например, выигрывающую коалицию {ПР, БЮТ, КПУ}. Выясним, какие фракции в ней являются ключевыми. В случае выхода ПР оставшаяся коалиция {БЮТ, КПУ} имеет всего 183 голоса (что меньше квоты $q = 226$), т.е. становится проигрывающей. В случае выхода БЮТ оставшаяся коалиция {ПР, КПУ} также становится проигрывающей (202 голоса). Наконец, в случае выхода КПУ оставшаяся коалиция {ПР, БЮТ} по-прежнему является выигрывающей (331 голос). Таким образом, в исходной коалиции {ПР, БЮТ, КПУ} фракции ПР и БЮТ являются ключевыми, а КПУ не является.

Обозначим через $b(i)$ количество коалиций, в которых i -я фракция является ключевой. *Индексом Банцафа этой фракции называется число*

$$\beta(i) = \frac{b(i)}{\sum_{i=1}^n b(i)}.$$

Таким образом, данный подход предполагает, что влияние фракции пропорционально количеству коалиций, в которых i -я фракция

является ключевой. Нормировка индекса Банцафа такова, что сумма индексов всех фракций равна единице:

$$\beta(1) + \beta(2) + \dots + \beta(n) = 1.$$

Проведем соответствующие вычисления для примера с Верховной Радой. В табл. 1 представлены все выигрывающие коалиции и выписаны фракции, которые являются в них ключевыми.

Таблица 1. Выигрывающие коалиции и ключевые фракции в Верховной Раде Украины VI созыва

№	Выигрывающая коалиция	Ключевые фракции
1	{ПР, БЮТ}	ПР, БЮТ
2	{ПР, НУНС}	ПР, НУНС
3	{БЮТ, НУНС}	БЮТ, НУНС
4	{ПР, БЮТ, НУНС}	—
5	{ПР, БЮТ, КПУ}	ПР, БЮТ
6	{ПР, БЮТ, БЛ}	ПР, БЮТ
7	{ПР, НУНС, КПУ}	ПР, НУНС
8	{ПР, НУНС, БЛ}	ПР, НУНС
9	{БЮТ, НУНС, КПУ}	БЮТ, НУНС
10	{БЮТ, НУНС, БЛ}	БЮТ, НУНС
11	{ПР, БЮТ, НУНС, КПУ}	—
12	{ПР, БЮТ, НУНС, БЛ}	—
13	{ПР, БЮТ, КПУ, БЛ}	ПР, БЮТ
14	{ПР, НУНС, КПУ, БЛ}	ПР, НУНС
15	{БЮТ, НУНС, КПУ, БЛ}	БЮТ, НУНС
16	{ПР, БЮТ, НУНС, КПУ, БЛ}	—

Из табл. 1 следует, что каждая из больших фракций (Партия Регионов, Блок Юлии Тимошенко и НУНС) является ключевой в 8 коалициях, а малые фракции (коммунисты и Блок Литвина) не являются ни в одной. Таким образом,

$$b(ПР) = b(БЮТ) = b(НУНС) = 8, \quad b(КПУ) = b(БЛ) = 0.$$

Следовательно, значения индекса Банцафа равны

$$\beta(ПР) = \beta(БЮТ) = \beta(НУНС) = \frac{8}{24} = \frac{1}{3}, \quad \beta(КПУ) = \beta(БЛ) = 0.$$

В данном примере выигрывающая коалиция БЮТ + НУНС контролирует количество голосов, лишь немного превышающее квоту. Поэтому после ухода из нее всего лишь нескольких депутатов значения индекса Банцафа для всех фракций могут значительно измениться. Подробнее об этом см. в комментариях к настоящему разделу.

1.5.2. Взвешенное голосование

Во многих странах парламенты и избирательные системы формировались на основе исторически сложившихся избирательных округов (графств, деревень и т. д.). При этом количество депутатов не находилось в прямой зависимости от численности избирателей. Например, каждая из двух деревень в 100 и в 520 избирателей могла иметь по одному члену парламента. Тем самым жители большой деревни были в некотором смысле ущемлены: фактически пять их голосов «весили» меньше, чем один голос жителя малой деревни.

Чтобы исправить эту ситуацию и сделать избирателей более равноправными, есть два пути. Первый предполагает, что от большой деревни должно быть избрано впятеро больше депутатов, чем от малой. Этот вариант является в настоящее время общепринятым, несмотря на очевидный недостаток: размеры округов почти никогда не бывают в точности равными и некоторое неравенство избирателей всегда остается. Второй путь заключается в том, чтобы каждая деревня была представлена одним парламентарием, но при этом депутат от большой деревни имел в парламенте 5 голосов или даже 5,2 (а не 1). Таким образом, при голосовании по какому-либо законопроекту голоса депутатов учитываются с весами, пропорциональными количеству избирателей, — это называется *взвешенным голосованием*. Второй вариант привлекателен своей экономичностью, а также тем, что он позволяет (по крайней мере, теоретически) учесть даже самые малые различия в размере округов. Такая практика встречалась сравнительно недавно, и лишь анализ влияния депутатов на основе индекса Банцафа позволил выявить главный недостаток системы взвешенного голосования.

Рассмотрим этот вопрос на конкретном примере. Графство Нассау (штат Нью-Йорк, США) управлялось Советом из шести членов. От самого крупного округа Хэмпстед избирались два советника, от четырех остальных округов — по одному (см. табл. 2). Количество голосов каждого члена Совета зависит от количества избирателей в его округе (для каждого советника от Хэмпстеда — по половине избирателей).

Вычисленные значения индекса Банцафа говорят о том, что (как в 1958 г., так и в 1964 г.) решения зависели от того, каким образом голосовали три наиболее влиятельных советника, а мнения трех наименее влиятельных не оказывали влияния на исход голосования. В конечном счете это означает, что избиратели малого округа (например, Лонг Бич) оказались в неравноправном положении — кого бы

Таблица 2. Количество голосов и значения индекса Банцафа для каждого из членов Совета графства Нассау в 1958 и 1964 гг.

Округ	1958 г.		1964 г.	
	Голосов	Индекс	Голосов	Индекс
Хемпстед (1)	9	1/3	31	1/3
Хэмпстед (2)	9	1/3	31	1/3
Северный Хэмпстед	7	1/3	21	0
Ойстер Бэй	3	0	28	1/3
Глен Коув	1	0	2	0
Лонг Бич	1	0	2	0

они ни избрали в Совет, их представитель был лишен влияния при голосовании.

Любопытная особенность данной ситуации заключается в том, что в период 1958–1964-х гг. изменился состав «влиятельной тройки». За это время население Ойстер Бэй существенно выросло и превысило население Северного Хэмпстеда. В результате представитель одного округа полностью лишился своего влияния, а другого округа обрел влияние.

Кратко скажем о другом примере — он интересен, в частности тем, что представляет один из первых случаев, когда исследование в области социальных наук с необходимостью потребовало применения расчетов на компьютере (публикация об исследовании вышла в 1965 г.). Законодательное собрание штата Нью-Джерси должно было состоять из 21 депутата, имеющих различное количество голосов, пропорциональное количеству избирателей в округе (закон о внедрении взвешенного голосования был принят, но затем отменен Верховным судом штата еще до проведения первых выборов по этому закону). Так, в двух самых маленьких округах было по 22 456 избирателей, соответствующие два депутата имели ровно по одному голосу. Вычисления показали, что индекс Банцафа каждого из них составлял 0,008. Трое наиболее влиятельных законодателей имели соответственно 19,0; 16,1; 12,6 голосов и индексы Банцафа, равные 0,164; 0,133; 0,101. С учетом общего количества депутатов законодательного собрания для вычисления этих индексов потребовалось рассмотреть $2^{21} = 2\,097\,152$ коалиций. Это привело к необходимости разработки специальной компьютерной программы, что было весьма нетипично для начала 1960-х гг. (Вычисления по графству Нассау также проводились с применением компьютера, однако это не было необходимым: индексы Банцафа для парламента из 6 депутатов могут быть легко получены вручную; возможно, это был

один из первых случаев, когда компьютер применялся в социальных исследованиях без особой необходимости.)

Выявленное в обоих случаях неравноправие избирателей послужило причиной изменения местного законодательства, и на основании этих прецедентов было отменено взвешенное голосование в других штатах и графствах США.

1.5.3. Парадокс перераспределения депутатов между фракциями

При переходе депутатов из одной фракции в другую влияние фракций может измениться. Как правило, увеличение численности фракции ведет к увеличению ее влияния (см, напр. ниже, комментарий 1). Однако возможна и обратная ситуация.

Перед тем, как привести пример, введем следующее обозначение.

Ситуацию голосования с квотой q и численностями фракций, равными k_1, \dots, k_n , будем описывать набором чисел $(q; k_1, \dots, k_n)$. Например, ситуация голосования в Верховной Раде Украины VI созыва обозначается через $(226; 175, 156, 72, 27, 20)$.

Рассмотрим теперь ситуацию

$$(10; 4, 6, 1, 1, 1). \quad (1)$$

Ее можно трактовать как голосование в 13-местном парламенте по конституционному вопросу, для принятия которого необходимо 10 голосов. Чтобы вычислить индексы Банцафа, выпишем все выигрывающие коалиции и ключевые фракции в них (табл. 3). При этом фракции будем обозначать буквами А, В, С, D, Е в том же порядке, что и в (1).

Таблица 3. Выигрывающие коалиции и ключевые фракции для голосования в ситуации (1)

№	Выигрывающая коалиция	Ключевые фракции
1	{А, В}	А, В
2	{А, В, С}	А, В
3	{А, В, D}	А, В
4	{А, В, E}	А, В
5	{А, В, С, D}	А, В
6	{А, В, С, E}	А, В
7	{А, В, D, E}	А, В
8	{А, В, С, D, E}	А, В

Очевидно, значения индекса Банцафа равны

$$\beta(A) = \beta(B) = \frac{8}{16} = \frac{1}{2}, \quad \beta(C) = \beta(D) = \beta(E) = 0.$$

Теперь предположим, что один депутат перешел из фракции А во фракцию В. Тем самым имеем голосование

$$(10; 3, 7, 1, 1, 1). \quad (2)$$

Проведем вычисления по той же схеме, что и для голосования в ситуации (1).

Таблица 4. Выигрывающие коалиции и ключевые фракции для голосования в ситуации (2)

№	Выигрывающая коалиция	Ключевые фракции
1	{A, B}	A, B
2	{A, B, C}	A, B
3	{A, B, D}	A, B
4	{A, B, E}	A, B
5	{A, B, C, D}	A, B
6	{A, B, C, E}	A, B
7	{A, B, D, E}	A, B
8	{A, B, C, D, E}	B
9	{B, C, D, E}	B, C, D, E

В соответствии с табл. 4 имеем значения индексов Банцафа:

$$\beta(A) = \frac{7}{19}, \quad \beta(B) = \frac{9}{19}, \quad \beta(C) = \beta(D) = \beta(E) = \frac{1}{19}.$$

Сравнивая влияние фракции В для голосований (1) и (2), получаем, что несмотря на увеличение количества депутатов в ней, значение индекса Банцафа для этой фракции уменьшилось. Это явление называется парадоксом перераспределения или парадоксом дарения (donation paradox).

Причина возникновения парадокса заключается в том, что переход одного депутата из фракции А в В косвенно отразился на трех других фракциях, С, D, E. Если в ситуации (1) они имели нулевые значения индекса Банцафа, то в ситуации (2) их влияние уже не нулевая. Можно сказать, что общая влияние парламента делится теперь не на двух игроков, а на пятерых. Это и привело к уменьшению индекса Банцафа фракции В, несмотря на ее увеличение.

В теории индексов влияния известны также и другие парадоксы. Одно из направлений этой теории — разработка новых мер влияния фракций при голосовании, устойчивых по отношению к этим парадоксам. Некоторые из индексов влияния базируются на несложных соображениях, подобных изложенным выше для индекса Банцафа, другие же основаны на теоретико-игровом подходе.

1.5.4. Учет коалиционных предпочтений и парадокс вражды

Определение влиятельности фракции при голосовании на основе индекса Банцафа использует не высказанное нами явно, но важное предположение. Именно, все коалиции предполагаются здесь равно-возможными. Это далеко не всегда соответствует реальности, даже приблизительно. Например, умеренная левая партия чаще образует коалицию с центристской или радикальной левой, чем с умеренной правой; и почти никогда не блокируется с радикальной правой.

Чтобы продемонстрировать на простом примере, как это соображение может быть учтено при расчете индексов влияния, вернемся к рассмотренному выше примеру с Верховной Радой Украины VI созыва. Предположим, что выигрывающие коалиции с одновременным участием НУНС и КПУ невозможны (это предположение не вполне соответствует действительности, однако оно более реалистично, чем, например, предположение о равновозможности коалиций {НУНС, КПУ} и {НУНС, БЮТ}). В этом случае при расчете индекса Банцафа надо удалить из таблицы 1 строки с выигрывающими коалициями, содержащими одновременно эти фракции. Получим табл. 5.

Таблица 5. Выигрывающие коалиции, не содержащие одновременно НУНС и КПУ

№	Выигрывающая коалиция	Ключевые фракции
1	{ПР, БЮТ}	ПР, БЮТ
2	{ПР, НУНС}	ПР, НУНС
3	{БЮТ, НУНС}	БЮТ, НУНС
4	{ПР, БЮТ, НУНС}	—
5	{ПР, БЮТ, КПУ}	ПР, БЮТ
6	{ПР, БЮТ, БЛ}	ПР, БЮТ
7	{ПР, НУНС, БЛ}	ПР, НУНС
8	{БЮТ, НУНС, БЛ}	БЮТ, НУНС
9	{ПР, БЮТ, НУНС, БЛ}	—
10	{ПР, БЮТ, КПУ, БЛ}	ПР, БЮТ

Нетрудно видеть, что значения индекса Банцафа теперь равны

$$\beta(ПР) = \beta(БЮТ) = \frac{6}{16} = \frac{3}{8}, \quad \beta(НУНС) = \frac{4}{16} = \frac{2}{8}, \quad \beta(КПУ) = \beta(БЛ) = 0.$$

Таким образом, с учетом данного предположения влияние фракции НУНС ниже, чем в случае равновозможности всех коалиций.

Этот пример демонстрирует следующую закономерность: чем ниже способность фракции формировать коалиции, тем ниже ее индекс Банцафа. Однако возможна и обратная ситуация. Пусть голосование имеет вид

$$(51; 50, 25, 25).$$

Обозначим фракции через А (50 депутатов), В и С. Выигрывающими коалициями являются {А, В}, {А, С}, {А, В, С}. Фракция А является ключевой во всех трех, каждая из фракций В, С — в одной. Следовательно, индексы Банцафа равны

$$\beta(A) = \frac{3}{3+1+1} = \frac{3}{5}, \quad \beta(B) = \beta(C) = \frac{1}{5}.$$

Предположим теперь, что фракции В и С не вступают в коалицию друг с другом. В этом случае выигрывающими коалициями являются {А, В}, {А, С}, а значения индекса Банцафа равны

$$\beta(A) = \frac{2}{2+1+1} = \frac{1}{2}, \quad \beta(B) = \beta(C) = \frac{1}{4}.$$

Таким образом, в данном примере влияние фракций В, С увеличилось из-за того, что они не готовы вступать в коалицию друг с другом — это явление называется парадоксом вражды.

Приведенные выше рассуждения о предпочтениях фракций по вступлению в коалиции предполагают, что две фракции либо могут вступать в коалицию, либо не могут. Однако в реальных политических процессах предпочтения имеют более сложный характер: существуют различные градации способности или неспособности образовывать коалицию с тем или иным партнером. В связи с этим в теории индексов влияния при голосованиях с предпочтениями вводится числовая мера предпочтения, f , фракций по вступлению в коалиции. Эта мера изменяется в пределах от 0 до 1 и равна нулю в случае, если фракции никогда не вступают в коалиции. Меры предпочтения определяются на основе так называемых *индексов согласованности фракций*, вычисляемых на основе анализа результатов голосований в парламенте за некоторый прошедший период (чем чаще две фракции голосовали одинаковым образом, тем выше индекс их согласованности). В качестве примера приведем индексы согласованности фракций в Верховной Раде Украины VI созыва (табл. 6).

Для двухпартийной коалиции мера предпочтения f равна индексу согласованности этих двух фракций. Для более широких коалиций существуют различные способы определения величины f . В частности, можно принять f равной минимальному из всех индексов согласованности пар фракций, входящих в коалицию.

Рассмотрим, например, коалицию {БЮТ, НУНС, КПУ}. Согласованность БЮТ и НУНС равна 0,74, БЮТ и КПУ — 0,29, НУНС и КПУ — 0,23. Тогда для данной коалиции имеем $f\{\text{БЮТ, НУНС, КПУ}\} = 0,23$.

Таблица 6. Индексы согласованности фракций в Верховной Раде Украины VI созыва

Партия	ПР	БЮТ	НУНС	КПУ	БЛ
ПР		0,41	0,35	0,69	0,53
БЮТ			0,74	0,29	0,71
НУНС				0,23	0,57
КПУ					0,41
БЛ					

Индекс влияния, обобщающий подход Банцафа на голосования с предпочтениями по формированию коалиций, называется альфа-индексом. Чтобы вычислить его, дополним табл. 1 столбцом, содержащим меры предпочтения (см. табл. 7).

Таблица 7. Выигрывающие коалиции, ключевые фракции и меры предпочтения в Верховной Раде Украины VI созыва

№	Выигрывающая коалиция	Ключевые фракции	Мера предпочтения f
1	{ПР, БЮТ}	ПР, БЮТ	0,41
2	{ПР, НУНС}	ПР, НУНС	0,35
3	{БЮТ, НУНС}	БЮТ, НУНС	0,74
4	{ПР, БЮТ, НУНС}	—	
5	{ПР, БЮТ, КПУ}	ПР, БЮТ	0,29
6	{ПР, БЮТ, БЛ}	ПР, БЮТ	0,41
7	{ПР, НУНС, КПУ}	ПР, НУНС	0,23
8	{ПР, НУНС, БЛ}	ПР, НУНС	0,35
9	{БЮТ, НУНС, КПУ}	БЮТ, НУНС	0,23
10	{БЮТ, НУНС, БЛ}	БЮТ, НУНС	0,57
11	{ПР, БЮТ, НУНС, КПУ}	—	
12	{ПР, БЮТ, НУНС, БЛ}	—	
13	{ПР, БЮТ, КПУ, БЛ}	ПР, БЮТ	0,29
14	{ПР, НУНС, КПУ, БЛ}	ПР, НУНС	0,23
15	{БЮТ, НУНС, КПУ, БЛ}	БЮТ, НУНС	0,23
16	{ПР, БЮТ, НУНС, КПУ, БЛ}	—	

Для каждой фракции вычисляется величина a , равная сумме мер предпочтения всех коалиций, в которых она является ключевой (и представляющая собой аналог величины b для индекса Банцафа). Например, для фракции ПР имеем

$$a(\text{ПР}) = 0,41 + 0,35 + 0,29 + 0,41 + 0,23 + 0,35 + 0,29 + 0,23 = 2,56.$$

После аналогичных вычислений получаем для других фракций:

$$a(\text{БЮТ}) = 3,17, \quad a(\text{НУНС}) = 2,93, \quad a(\text{КПУ}) = a(\text{БЛ}) = 0.$$

Определение вводится аналогично индексу Банцафа. *Альфа-индексом фракции i называется число*

$$\alpha(i) = \frac{a(i)}{\sum_{i=1}^n a(i)}.$$

В соответствии с данным определением имеем для этого примера следующие индексы влияния:

$$\alpha(\text{ПР}) = \frac{2,56}{2,56 + 3,17 + 2,93} = 0,29, \quad \alpha(\text{БЮТ}) = 0,37, \\ \alpha(\text{НУНС}) = 0,34, \quad \alpha(\text{КПУ}) = \alpha(\text{БЛ}) = 0.$$

Таким образом, учет предпочтений фракций по формированию коалиций позволил построить альфа-индексы, более адекватно отражающие влияние фракций при голосовании, чем индексы Банцафа.

Упражнения

1. В комментарии 2 раздела «Комментарии и библиографические примечания» описана ситуация, возникшая в Верховной Раде Украины VI созыва после того, как два депутата изменили свою фракционную принадлежность. Вычислить значения индекса Банцафа, приняв, что один из этих депутатов перешел из БЮТ в БЛ, а другой — из НУНС в ПР.

2. По данным табл. 6 вычислить для каждой коалиции меру предпочтения f , приняв ее равной среднему всех индексов согласованности пар фракций, входящих в коалицию. Вычислить альфа-индексы всех фракций.

Комментарии и библиографические примечания

1. Индекс Банцафа предложен в работе [1]. Лежащие в его основе соображения высказаны в [2] без формулировки четкого алгоритма. Из других индексов влияния наиболее важным является индекс Шепли–Шубика, предложенный в работе [3].

2. В соответствии с украинским законодательством правящая коалиция официально оформляется за подписями всех входящих в нее

депутатов. В Раде VI созыва такую коалицию первоначально составили все члены фракции БЮТ и все, кроме одного, депутаты фракции НУНС. Таким образом, в реальности коалиция {БЮТ, НУНС} имела 227 голосов, что лишь на один голос превышает квоту $q = 226$. После того, как в июне 2008 г. из нее вышли два депутата, коалиция {БЮТ, НУНС} перестала быть выигрывающей. Это привело к политическому кризису и формированию в декабре 2008 г. новой правящей коалиции {БЮТ, НУНС, БЛ}. В марте 2010 г., после перехода еще нескольких депутатов, сформировалась последняя правящая коалиция в Раде VI созыва: {ПР, КПУ, БЛ}. В связи с этими событиями в украинском политическом языке появился термин «тушки», обозначающий депутатов, покинувших фракцию, в составе которой они были избраны. Данный пример ярко иллюстрирует, что влияние фракции не пропорционально количеству контролируемых ею голосов. Выход всего двух депутатов из блока {БЮТ, НУНС} привел к падению влияния этих фракций и в то же время — к резкому росту влияния БЛ и КПУ.

3. Избирательные округа с малочисленным населением и непропорционально большим представительством в парламенте были весьма распространены в Англии в конце XVIII – начале XIX веков. Они получили название «гнилых местечек». Так, в деревнях Олд Сарум и Гаттон жили всего по 7 избирателей, но от каждой из них избиралось по два члена британской Палаты Общин. В то же время такой крупный город, как Манчестер, являлся лишь частью более крупного округа. Всего на выборах 1831 г., 152 депутата были избраны от английских округов численностью менее 100 избирателей (всего депутатов было 658, из них 486 — от Англии).

4. После того как Банцаф получил изложенные в п. 2 результаты относительно взвешенного голосования в графстве Нассау, он подал иск о неконституционности местного избирательного законодательства, что и привело в конечном счете к отмене системы взвешенного голосования в тех штатах, где она существовала или планировалась к применению. Чтобы оценить значение поданного иска, необходимо заметить, что графство Нассау — одно из самых богатых и известных в США. Его северная часть относится к так называемому Золотому Берегу, на котором с конца XIX века располагаются поместья таких семей, как Вандербильты, Рузвельты, Морганы и т. д. Образ жизни довоенного Нассау описан Ф. С. Фитцджеральдом в романе «Великий Гэтсби». Здесь же в 1946–1951-х годах располагалась штаб-квартира Совета Безопасности ООН.

5. Представленные в табл. 6 индексы согласованности фракций относятся к голосованиям по внешнеполитическим вопросам. Строго говоря, именно поэтому выводы четвертого пункта о влиятельности фракций также относятся лишь к голосованиям по внешней политике. Практика исследования деятельности парламентов при голосовании по отдельному кругу вопросов (а не по всему спектру законодательной деятельности) довольно распространена в политической науке.

6. Индексы влияния, учитывающие предпочтения фракций по вступлению в коалиции (в частности, альфа индекс), введены в [4, 5]. Парадокс перераспределения депутатов между фракциями описан, напр., в [6]. Парадокс вражды подробно изучен в работах [7, 8]. Методика определения индекса согласованности фракций при голосовании предложена в [9]. Исследование Верховной Рады Украины VI созыва проведено в [11]. В качестве дополнительного материала по теме данного раздела можно рекомендовать главу 6 известного учебника [10].

Литература

1. *Banzhaf J.F.* Weighted voting doesn't work: A mathematical analysis // *Rutgers Law Review*. 1965. 19(2). P. 317–343.
2. *Penrose L.* The Elementary Statistics of Majority Voting // *Journal of the Royal Statistical Society* (Blackwell Publishing). 1946. 109(1). P. 53–57.
3. *Shapley L.S., Shubik M.* A method for Evaluating the Distribution of Power in a Committee System // *American Political Science Review*. 1954. № 48(3). P. 787–792.
4. *Aleskerov F.* Power indices taking into account agents' preferences. *Mathematics and Democracy*. Berlin, Springer, 2006. P. 1–18.
5. *Алескеров Ф.Т.* Индексы влияния, учитывающие предпочтения участников по созданию коалиций // *ДАН*. 2007. Т. 414. № 5. С. 594–597.
6. *Felsenthal D.S., Machover M.* The Measurement of Voting Power: Theory and Practice, Problems and Paradoxes. Cheltenham: Edward Elgar, 1998. 322 p.
7. *Brams S.J.* Game Theory and Politics. New York: Free Press, 1975. 312 p.
8. *Kilgour D.M.* A formal analysis of the amending formula of Canada's Constitution Act // *Canadian Journal of political science*. 1983. № 16. P. 771–777.
9. *Алескеров Ф.Т., Благовещенский Н.Ю., Сатаров Г.А., Соколова А.В., Якуба В.А.* Оценка влияния групп и фракций в российском парламенте (1994–2003 гг.). Препринт ГУ Высшая Школа Экономики, WP7/2003/01, Москва, 2003.
10. *Алескеров Ф.Т., Хабина Э.Л., Шварц Д.А.* Бинарные отношения, графы и коллективные решения. Издательский дом ГУ ВШЭ. М. 2006. 300 с.
11. *Юрескул Е.А.* Исследование механизма принятия внешнеполитических решений в национальных и наднациональных представительных органах количественными методами (на примере Верховной Рады Украины VI созыва, Бундестага ФРГ 16 законодательного периода, Шестого Европейского Парламента). Факультет глобальных процессов МГУ им. М.В. Ломоносова. Дипломная работа. Москва, 2010.

Глава 2

ДЕМОГРАФИЯ

2.1. Базовые макромоделли демографии

Изучение демографических моделей начнем с математически простых и исторически ранних моделей. Первая из них основана на подходе, предложенном Т. Р. Мальтусом еще в 1798 г., вторая, так называемая *логистическая модель* (П. Ф. Ферхюльст) датируется 1848 годом. Их появление, а также анализ и обсуждение вытекающих из них выводов положили начало математическому моделированию в демографии и вызвали большой интерес не только ученых, но также публицистов и общественных деятелей. Эти модели рассматривают население в наиболее агрегированном виде, не учитывая возрастную или половую структуру. Тем не менее они позволили выделить некоторые общие закономерности демографической динамики, которые были позднее унаследованы более сложными и содержательными подходами.

С математической точки зрения обе модели представляют собой дифференциальные уравнения первого порядка либо их дискретные аналоги. Далее первый пункт, посвящен модели Мальтуса, второй пункт — логистической модели, третий представляет собой краткое заключение. Следующий за ним раздел комментариев и библиографических примечаний содержит, в частности, небольшой обзор исторической дискуссии вокруг модели Мальтуса.

2.1.1. Модель Мальтуса

Подход, предложенный Мальтусом еще в XVIII веке, лежит в основе простейшей демографической модели. Несмотря на то, что он не учитывает социальную структуру населения и множество других социальных факторов, модель Мальтуса описывает основную черту демографической динамики — экспоненциальный (в дискретном времени, геометрический) рост или убывание численности населения. С математической точки зрения причины экспоненциального характера решения — линейность модели и постоянство ее коэффициентов. Данная черта воспроизводится и в более сложных линейных моделях, обсуждаемых ниже. Дискуссии XIX века вокруг модели Мальтуса обозначили многие направления современной демографии (см. комментарии к настоящему разделу).

Идеи, лежащие в основе модели Мальтуса, сводятся к следующему. По данным Мальтуса (относящимся к американским переселенцам

и приятными им в качестве фундаментального закона), с течением времени численность населения растет в геометрической прогрессии, причем в отсутствие сдерживающих факторов период его удвоения составляет 25 лет. В то же время увеличение продуктов земледелия сдерживается плодородием почвы и (по предположению Мальтуса) не может расти быстрее, чем в арифметической прогрессии. Таким образом, если принять значения численности населения и продукта в некоторый фиксированный момент времени за единицу, то с 25-летним шагом «человеческий род размножался бы как 1, 2, 4, 8, 16, 32, 64, 128, 256, в то же время средства существования размножались бы как 1, 2, 3, 4, 5, 6, 7, 8, 9. По истечении двух столетий количество населения относилось бы к средствам существования как 256 к 9».

Обобщим рассуждение Мальтуса, приняв значения временного шага и скорости роста произвольными постоянными.

Рассмотрим численность $N(t)$ населения некоторой территории (страны, города, острова) в момент времени t . Будем предполагать, что ее изменение происходит только за счет рождаемости и смертности, а миграционные потоки отсутствуют. Рассмотрим сначала модель с дискретным временем.

Количество рожденных и количество умерших за единичный интервал времени $(t; t + 1)$ предполагаются пропорциональными численности населения на начало этого периода. При этом коэффициенты пропорциональности считаются постоянными. Таким образом, приращение численности населения за этот интервал времени составит

$$N(t + 1) - N(t) = bN(t) - mN(t),$$

где b и m — количество рожденных и умерших за этот интервал времени, приходящееся в среднем на одного живущего. Эти величины называются соответственно *общими коэффициентами рождаемости и смертности* (чаще всего их выражают в промилле, т.е. вычисляется количество рожденных и умерших в расчете не на одного, а на 1000 чел.).

Из предыдущей формулы следует, что

$$N(t + 1) = (1 + r)N(t),$$

где $r = b - m$. Величина r называется *общим коэффициентом естественного прироста*. Если рождаемость превышает смертность, то $r > 0$, в обратном случае $r < 0$. Численность населения в произвольный момент времени имеет вид

$$N(t) = (1 + r)^t N(0),$$

что и описывает предложенный Мальтусом рост (или, при $r < 0$, убывание) в геометрической прогрессии.

Перейдем теперь к изложению модели Мальтуса в непрерывном времени. По-прежнему предполагается, что миграция отсутствует, а общие коэффициенты рождаемости и смертности являются постоянными.

Поскольку количество рожденных и количество умерших за малый интервал времени длительностью Δt пропорциональны численности населения на начало этого периода, то приращение численности населения за интервал $(t, t + \Delta t)$ составит

$$\Delta N(t) = bN(t)\Delta t - mN(t)\Delta t.$$

Переходя к пределу при $\Delta t \rightarrow 0$, имеем

$$\frac{dN}{dt} = rN.$$

Проинтегрировав данное уравнение, нетрудно получить, что его решение имеет вид

$$N(t) = N_0 e^{rt},$$

где N_0 — численность населения в начальный момент времени $t = 0$.

Таким образом, в случае $r > 0$ модель Мальтуса описывает неограниченный экспоненциальный рост населения. Если $r < 0$ или $r = 0$, то численность соответственно экспоненциально убывает или является постоянной.

В качестве основного вывода относительно модели Мальтуса укажем, что экспоненциальная (или эквивалентная ей геометрическая в дискретном варианте) динамика численности населения обеспечивается двумя факторами: линейностью модели и постоянством общего коэффициента естественного прироста.

2.1.2. Логистическая модель

Модель экспоненциального роста населения, основана на идеях, изложенных Т.Р. Мальтусом в первой главе его книги «Опыт закона о народонаселении». В следующих главах этой книги рассмотрены некоторые факторы, препятствующие экспоненциальному росту, однако не сделаны попытки учесть влияние этих факторов на характер роста. Модель, учитывающая это влияние, была предложена в 1848 г. бельгийским математиком П.Ф. Ферхюльстом и получила название логистической модели. При этом конкретные механизмы влияния ограничивающих факторов на рост численности населения в модели не рассматриваются.

Напомним, что в модели Мальтуса коэффициент общего прироста r полагается не только неизменным во времени, но также не зависящим от текущей численности. Следствием такого предположения является неограниченный рост населения. В противоположность этому в логистическую модель заложено положение, что демографический рост ограничен ввиду ограниченности общего ресурса W (например, пищи), а по мере исчерпания этого ресурса происходит уменьшение коэффициент общего прироста r . В качестве конкретизации этого положения считается, что коэффициент r пропорционален количеству ресурсов, еще не исчерпанных ныне живущим населением. Таким образом, логистическая модель является нелинейной.

Приняв ту или иную гипотезу об изменении общего доступного ресурса с течением времени, можно получить соответствующую модель динамики народонаселения (см. упражнения). В настоящем разделе рассматривается логистическая модель, в рамках которой он предполагается постоянным. Миграция, как и в модели Мальтуса, не учитывается.

Перейдем к построению логистической модели. Обозначим через σ количество ресурса (например, питания), необходимое одному члену популяции. Тогда, если в момент времени t численность населения составляет $N(t)$, то это население использует ресурс в количестве $\sigma N(t)$. Если общее количество ресурса равно постоянной W , то неисчерпанная его часть составляет $W - \sigma N(t)$. Модель основана на гипотезе о том, что общий коэффициент прироста пропорционален неиспользованной части ресурса $W - \sigma N(t)$. Таким образом, прирост населения за временной интервал $(t, t + \Delta t)$ составляет в логистической модели

$$\Delta N(t) = rN(t)\Delta t = a(W - \sigma N(t))N(t)\Delta t.$$

Переходя к пределу при $\Delta t \rightarrow 0$, имеем

$$\frac{dN}{dt} = a(W - \sigma N)N, \quad (1)$$

Обозначим $N_0 = W/\sigma$. Эта величина называется демографической емкостью среды и имеет смысл численности населения, которое способно прокормиться при заданной величине ресурса W и индивидуальном потреблении σ . Введя обозначение $r_0 = a\sigma N_0$, перепишем (1) в виде

$$\frac{dN}{dt} = r_0 \left(1 - \frac{N}{N_0}\right) N. \quad (2)$$

Если $N \ll N_0$, то $dN/dt \approx r_0 N$, т. е. при малой численности рост населения является близким к экспоненциальному. После разделения переменных имеем

$$\frac{dN}{(1 - N/N_0)N} = r_0 dt.$$

Проинтегрировав, получим, что решение уравнения (2) имеет вид

$$N(t) = N(0) \frac{N_0 e^{r_0 t}}{N(0) e^{r_0 t} + (N_0 - N(0))}, \quad (3)$$

где $N(0)$ — численность населения в начальный момент времени $t = 0$.

Рассмотрим некоторые свойства полученного решения.

Из (2) следует, что $dN/dt > 0$, если $N(t) < N_0$, и $dN/dt < 0$, если $N(t) > N_0$. Следовательно, численность населения в логистической модели растет, если его текущее значение меньше емкости среды, и убывает, если оно больше. Кроме того, из (3) получим, вычислив предел, что при любом $N(0)$

$$\lim_{t \rightarrow +\infty} N(t) = N_0.$$

Таким образом, с течением времени численность населения монотонно стремится к емкости среды.

Выясним, имеет ли график решения перегиб. Для этого продифференцируем (2):

$$\frac{d^2 N}{dt^2} = r_0 \left(1 - \frac{2N}{N_0}\right) \frac{dN}{dt}.$$

Следовательно, при $N(t) < N_0/2$ вторая производная решения положительна, а рост является ускоряющимся. При $N(t) > N_0/2$ рост замедляется (рис. 1).

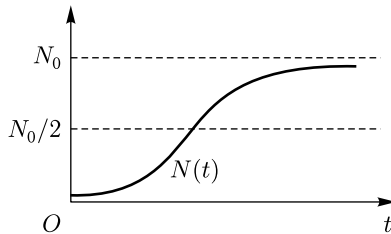


Рис. 1. Решение логистического уравнения (2)

При достаточно больших значениях t численность $N(t)$ приближается к предельному значению N_0 экспоненциальным образом. Действительно, уравнение для разности $z(t) = N_0 - N(t)$ имеет вид $dz/dt = -r_0 z(1 - z/N_0)$, и его линейное приближение $dz/dt = -r_0 z$ имеет решение вида $z(t) = \text{const} \cdot e^{-r_0 t}$. Таким образом, логистическое уравнение (2) описывает экспоненциальный рост в асимптотическом пределе $N(t) \rightarrow 0$ и экспоненциальное затухание при $N(t) \rightarrow N_0$.

Максимально быстро население растет при $N(t) = N_0/2$, что дает возможность предложить метод грубой оценки демографической емкости среды: она равна удвоенной численности населения в период ее максимального роста (в абсолютных показателях). Впрочем, к подобным оценкам следует относиться с большой осторожностью, так как они не учитывают возможные изменения социальных установок.

Рассмотрим теперь ситуацию, когда демографическая емкость среды изменяется с течением времени. Как указано выше, Мальтус допускал, что средства пропитания с течением времени могут возрастать, но не быстрее, чем в арифметической прогрессии (что в непрерывном времени соответствует линейной функции). Мы рассмотрим более общий случай, положив емкость среды произвольной дифференцируемой функцией времени. Уравнение логистической модели принимает тогда вид

$$\frac{dN}{dt} = r_0 \left(1 - \frac{N}{N_0(t)}\right) N. \quad (4)$$

Чтобы исследовать решение этого уравнения, введем переменную $y(t) = N(t)/N_0(t)$, имеющую смысл доли, которую составляет текущее

население от текущего значения емкости среды. Нетрудно получить, что уравнение для нее имеет вид

$$\frac{dy}{dt} = r_0 y \left[(1 - y) - \frac{N'_0(t)}{r_0 N_0(t)} \right].$$

Таким образом, если $N_0(t)$ растет медленнее, чем экспоненциальная функция и $\lim_{t \rightarrow \infty} N'_0(t)/N_0(t) = 0$, то с течением времени численность населения $N(t)$ стремится к $N_0(t)$ (хотя и медленнее, чем в случае постоянной емкости среды N_0). В частности, это относится к случаю линейного роста функции $N_0(t)$. Если же $N_0(t)$ растет экспоненциально $N_0(t) = K_0 e^{\varepsilon t}$, то уравнение примет вид

$$\frac{dy}{dt} = r_0 y \left[1 - \frac{\varepsilon}{r_0} - y \right],$$

т. е. при $t \rightarrow \infty$ численность $N(t)$ будет составлять долю $1 - \varepsilon/r_0$ от емкости $N_0(t)$.

Отметим, что при растущей функции $N_0(t)$ максимальный темп роста населения достигается не при достижении половины емкости, а позже (см. упражнение 5). Вообще, манипулируя функцией $N_0(t)$, можно получить рост численности населения любого вида, в т. ч. удовлетворяющий исторической статистике.

Емкость среды является трудно измеряемой величиной. Необходимо отметить, что если в предшествующие эпохи коэффициент естественного прироста определялся в основном возможностями выживания человека при существовавшем хозяйственном укладе, то в наше время в развитых странах появился фактор, связанный с сознательным решением семьи о рождении определенного, как правило, небольшого количества детей. Поэтому основным фактором, определяющим демографическую динамику в России и ряде других стран, является установка женщин на рождение того или иного количества детей (см. подробнее п. 2.5.). Поэтому модели, не учитывающие социальные особенности среды, не могут претендовать на адекватное описание действительности.

В реальных обществах коэффициент естественного прироста изменяется под влиянием целого ряда факторов. Он может иметь положительный или отрицательный тренд, с наложенными на него колебаниями, возникающими ввиду демографических волн (это понятие будет рассмотрено ниже). Однако при выполнении определенных условий на небольших временных интервалах коэффициент естественного прироста может считаться приблизительно постоянным.

Неограниченный экспоненциальный рост численности населения, с одной стороны, и логистический, протекающий с насыщением, с другой стороны, представляют собой два наиболее значимых сценария демографической динамики. В этом смысле рассмотренные в данном

разделе модели являются наиболее простыми и наиболее фундаментальными представителями линейного и нелинейного подходов.

Нелинейные эффекты существенны лишь на большом масштабе времени, охватывающем много поколений. Но на этом масштабе коэффициенты рождаемости и смертности могут значительно изменяться. Поэтому подавляющее большинство моделей «человеческой» демографии являются линейными; они редко бывают предназначены для анализа и прогноза более чем на 50–70 лет. Главным свойством линейности систем с постоянными коэффициентами является экспоненциальный рост (или спад) численности населения. Такой рост часто называют мальтузианским; он будет продемонстрирован в следующих разделах настоящей главы для более сложных линейных моделей, учитывающих, например, возрастную структуру населения.

Упражнения

1. В рамках модели Мальтуса установить, при каком значении скорости роста количество живущих в любой момент времени $t > 0$ превышает количество умерших.

2. Рассмотреть модель с постоянным миграционным потоком

$$\frac{dN}{dt} = rN + \mu, \quad N(0) = N_0.$$

При каких условиях на коэффициенты уравнения и начальные условия решение описывает рост населения, а при каких — спад?

3. Рассмотреть модель с экспоненциально растущей емкостью среды

$$\frac{dN}{dt} = r_0 \left(1 - \frac{N}{K_0 \exp[Rt]} \right) N, \quad N(0) = N_0.$$

Показать, что независимо от соотношения между r_0 и R население не может возрасти быстрее, чем в модели Мальтуса, т. е. как $\exp(r_0 t)$.

Решение. Решая задачу Коши, получаем

$$N(t) = \frac{N_0 \exp[r_0 t]}{1 + \frac{N_0 r_0}{r_0 - R} (\exp[(r_0 - R)t] - 1)}.$$

Если $r_0 > R$, то при достаточно больших t : $N(t) \sim (1 - R/r_0) \exp(Rt)$, и население растет с той же скоростью, что и емкость, т. е. медленнее, чем по Мальтусу. Если же $R > r_0$, то $\exp[(r_0 - R)t]$ — убывающая функция, $N(t) \sim \frac{N_0 \exp[r_0 t]}{1 + N_0 r_0 / (R - r_0)}$ и рост равен мальтузианскому.

4. Рассмотреть логистическую модель с постоянным миграционным потоком

$$\frac{dN}{dt} = r_0 \left(1 - \frac{N}{K} \right) N + \mu, \quad N(0) = N_0.$$

Показать, что: а) при $\mu > 0$ существует стационарное значение N_{st} такое, что $N(t) \rightarrow N_{st}$ при любом N_0 ; б) если $\mu < 0$ (отрицательное постоянное сальдо миграции), то $N(t) \rightarrow 0$ при достаточно малых N_0 ; в) при фиксированных значениях r_0 и K существуют $\mu < 0$ такие, что $N(t) \rightarrow 0$ при любых N_0 .

5. Для модели (4) с растущей емкостью среды показать, что максимальный темп роста населения достигается при текущем значении населения, более высоком, чем $K(t)/2$.

Решение. Дифференцируя (4), имеем

$$\frac{d^2 N}{dt^2} = r_0 \left(1 - \frac{2N(t)}{K(t)} \right) \frac{dN}{dt} + r_0 \frac{N^2(t)}{K^2(t)} \frac{dK}{dt}.$$

Если t_0 — точка максимального роста, то

$$\left(1 - \frac{2N(t_0)}{K(t_0)} \right) \frac{dN}{dt} \Big|_{t=t_0} + \frac{N^2(t_0)}{K^2(t_0)} \frac{dK}{dt} \Big|_{t=t_0} = 0.$$

Отсюда, поскольку $dK/dt > 0$, следует, что $1 - \frac{2N(t_0)}{K(t_0)} < 0$, т. е. $\frac{N(t_0)}{K(t_0)} < \frac{1}{2}$.

Комментарии и библиографические примечания

1. Идеи Т.Р. Мальтуса, лежащие в основе модели, называемой сегодня его именем, изложены в труде «Опыт закона о народонаселении», первое издание которого [1] вышло анонимно в 1798 г.; см. также русский перевод [2]. Все цитирования Мальтуса приведены нами по этому переводу.

В XIX веке теория Мальтуса обсуждалась весьма широко, при этом возражения критиков относились как к закону геометрической прогрессии роста населения, так и к закону арифметической прогрессии роста продуктов сельского хозяйства. Так, американский экономист и социолог Г. Ч. Кэри (1793–1879) утверждал, что по мере развития человеческого прогресса рождаемость будет падать в связи с увеличением роли интеллектуальной деятельности в жизни человека (причем Кэри связывал это не с психологическими, а с физиологическими факторами), что приведет к прекращению геометрического роста населения. Английский философ, социолог и биолог Г. Спенсер (1820–1903) в труде «Основы биологии», опубликованном в 1864 г., излагал похожий аргумент в эволюционном духе: «*Тип постепенно будет так изменяться, что более развитая нервная система станет притягивать из общего запаса большее количество питания, а вследствие этого <...> она будет неизбежно уменьшать запас, который мог бы пойти на производство новых жизней*» [3]. В терминах нашего времени можно сказать, что Спенсер объясняет предстоящее уменьшение рождаемости с помощью отрицательной обратной связи при

установлении равновесия человека с природой: плодовитость и борьба за выживание будут приводить к интеллектуальному росту, который в свою очередь приведет к уменьшению плодовитости. В результате произойдет сокращение рождаемости до 2–3 детей у каждой родительской пары и стабилизация к постоянной численности населения, при этом увеличится продолжительность жизни.

К. Маркс критиковал Мальтуса с иной позиции: он указывал на ограниченность мальтузианского закона народонаселения, связанную с абстрагированием данного анализа от общего исторического процесса. По Марксу, каждой общественной формации присущ свой специфический закон народонаселения: *«Всякому исторически особенному способу производства в действительности свойственны свои особенные, имеющие исторический характер законы народонаселения. Абстрактный закон населения существует только для растений и животных»* [4]. Данное утверждение Маркса можно трактовать как идею о том, что мальтузианский рост имеет место лишь для некоторых исторических периодов и не может быть экстраполирован на всю временную ось.

Д.И. Менделеев, не одобряя мальтузианства по моральным причинам, все же пользовался геометрической прогрессией для получения грубых оценок. При этом он всегда явно подчеркивал, что это лишь допущение, например: *«Если считать, что прирост сохранится и впредь близким к 1%, т. е. число жителей всей Земли будет удваиваться примерно в 60–70 лет, то через 100 лет, т. е. к 2000 году получится жителей на Земле более 4 млрд., и тогда теснота будет такая же почти повсюду, как теперь в Германии»* [5, С. 228].

Среди возражений по поводу геометрического роста населения следует выделить и несущественные с принципиальной точки зрения, но казавшиеся важными некоторым исследователям XIX века, возражения относительно численного значения периода удвоения, определенного Мальтусом как 25 лет. Критики отмечали, что в наблюдениях, на которые опирался Мальтус (прежде всего, рост населения США в течение 150 лет), прирост населения связан в первую очередь с миграционным приростом, а также с очень высокой долей женщин фертильного возраста в этом миграционном потоке. Удвоение населения за 25 лет (что составляет чуть меньше трех процентов ежегодного прироста) без миграционного пополнения казалось в XIX веке практически невозможным *«по самому устройству физической организации человека»*, под чем имелась в виду, в частности, высокая младенческая и детская смертность. Действительно, по данным того времени (обычно авторы ссылались исследования на Э. Чедвика), в первый год жизни умирало 20% новорожденных, а в первые 5 лет жизни — 30–40% детей (в различных европейских странах). Поэтому возрастание численности населения в 3% автором предисловия к русскому изданию 1895 г. названо *«немыслимым»*. Заметим, что в нашу эпоху детская смертность стала гораздо ниже, и в 2009 г. (по оценкам CIA World Factbook) 16 стран

и территорий имели годовой прирост населения 2,82% и выше, что означает более чем удвоение населения за 25 лет (при этом такие страны, как Нигер, ДР Конго, Буркина Фасо, Руанда совмещают высокий прирост с высокой по сегодняшним меркам детской смертностью — в этих странах 17–19% детей не доживают до 5 лет).

Как мы видим, уже в ранних дискуссиях вокруг модели Мальтуса были подняты важные вопросы, которые не могут быть проигнорированы при математическом моделировании демографических процессов: роль миграции, половозрастная структура населения, границы фертильного возраста, детская смертность и пр. Некоторые из аргументов, высказанных сторонами, кажутся сегодня наивными, другие же поднимают вопросы, и сегодня волнующие демографов.

2. Важными предположениями, лежащими в основе модели Мальтуса, являются постоянство коэффициента естественного прироста и отсутствие (либо, хотя бы, нулевое сальдо) миграции. В реальных обществах коэффициент естественного прироста изменяется под влиянием целого ряда факторов. Он может иметь положительный или отрицательный тренд с наложенными на него колебаниями, возникающими ввиду демографических волн (см. п. 2.3). Однако при выполнении определенных условий, на небольших временных интервалах коэффициент естественного прироста может считаться приблизительно постоянным.

В качестве примера, иллюстрирующего некоторые из возможных ситуаций, рассмотрим данные по четырем различным странам за период 1989–2008-х гг. На рис. 2 приведены графики изменения суммарных коэффициентов рождаемости и смертности для России, Молдовы, Аргентины и Австралии.

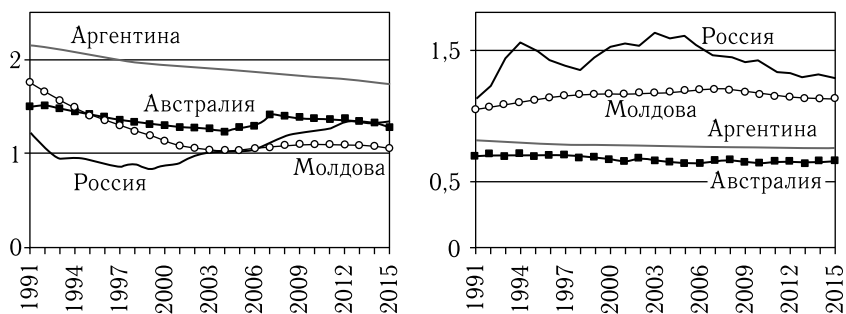


Рис. 2. Суммарные коэффициенты рождаемости (слева) и смертности (справа).
Источник: база данных Всемирного Банка

Эти примеры показывают, что в обществах, испытывающих сильный социальный стресс, изменения демографических показателей происходят наиболее сильно.

Подробный анализ причин изменений этих показателей не является целью настоящей книги. Укажем лишь на ключевую роль такого

социального фактора, как установка женщины на рождение определенного количества детей (см. п. 2.5). Вследствие этого значения коэффициента рождаемости оказываются наиболее близкими в странах, обладающих культурным и историческим родством. Это положение иллюстрируется примером, представленным на рис. 3: такие культурно близкие страны, как Россия и Беларусь, характеризуются близкими значениями коэффициента. То же относится к Саудовской Аравии и Оману. При этом различия между данными парами стран весьма существенны.

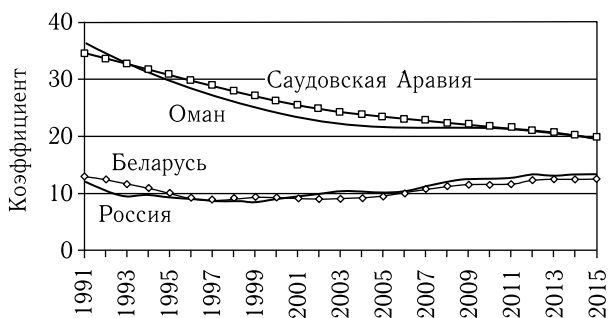


Рис. 3. Суммарные коэффициенты рождаемости по годам. Страны, близкие с культурной и исторической точек зрения, имеют сходную динамику и близкие значения коэффициента. Источник: база данных Всемирного Банка

Как отмечали еще первые критики модели Мальтуса, в ней не учитывается влияние миграции на численность населения. Иногда, действительно, миграция является пренебрежимо малой. В других случаях ее положительное или отрицательное сальдо может внести существенные коррективы в динамику численности.

3. Вопросы ограниченности роста «произведений земли» описаны Мальтусом весьма неубедительно; в отличие от вопроса о численности населения, они являются чисто спекулятивными и не базируются ни на каких статистических данных. Мальтус пишет, что по мере возрастания населения будет возделана вся плодородная земля, и после этого дальнейшее «увеличение количества пищи может быть достигнуто не иначе, как только путем улучшения занятых ранее земель. Эти улучшения, по самим свойствам почвы, не только не могут сопровождаться постоянно возрастающими успехами, но наоборот, последние будут постепенно уменьшаться». Этот тезис получил название закона убывающего плодородия почвы, хотя более точно было бы назвать его законом убывающего прироста плодородия почвы.

Таким образом, возрастание продуктов земледелия в арифметической прогрессии является для Мальтуса не установленной (хотя бы спекулятивно) закономерностью, а оценкой сверху. Один из последователей Мальтуса, известный английский экономист Дж.С. Милль

(1806–1873) уточняет: «Границу, полагаемую производству качествами почвы, не надобно воображать подобною стене, которая стоит неподвижно на известном месте и не задерживает движения, пока совершенно не останавливает его. Скорее, можно сравнить эту границу с очень эластичною и растяжимую повязкою, которая, как бы сильно ни была растянута, все-таки может быть растянута еще шире, но которая стесняет задолго прежде, чем растянется до последнего предела, и стеснение от которой тем сильнее, чем ближе к этому пределу раздвигается она» [6].

Критика этих положений была основана на более оптимистических оценках роста производительности земли. Так, В.И. Ленин писал: «Закон убывающего плодородия почвы вовсе не применим к тем случаям, когда техника прогрессирует, когда способы производства преобразуются; он имеет лишь весьма относительное и условное применение к тем случаям, когда техника остается неизменною» [7].

По вопросу о дискуссиях XIX века относительно мальтузианских идей отметим также курьезную, но получившую известность критику Мальтуса со стороны Н.Г. Чернышевского, первого русского переводчика Милля, доказывавшего в комментариях к переводу, что арифметическая прогрессия лишь незначительно отстает в росте от геометрической [8]. Более конкретно: принятое Мальтусом удвоение население за 25 лет соответствует примерно трем процентам годового прироста. По Чернышевскому, при подсчете за два последовательных года (а не сразу за 25 лет) различие между геометрической и арифметической прогрессиями составит ничтожно малую величину: $\delta = (1 + 0,03)^2 - (1 + 2 \cdot 0,03) = 0,0009$. Таким образом, чтобы улучшение аграрных технологий успевало за ростом населения, оно должно идти лишь с ничтожной добавкой к предполагаемой Мальтусом арифметической прогрессии.

Д.И. Менделеев приводил обширную критику Мальтуса, содержащую, в частности, расчеты, свидетельствующие о том, что даже при фиксированной эффективности земледелия пределы роста аграрного продукта еще очень долго не будут достигнуты. Эти расчеты опирались, с одной стороны, на оценки размера Земной суши, пригодной для земледелия, а с другой, — на текущую площадь обрабатываемой земли [5, с. 221].

Наконец, укажем точку зрения К.А. Тимирязева [9]. Подчеркнуто называя свою точку зрения позицией естествоиспытателя, он писал: «Не подлежит сомнению, что каждое живое существо стремится размножиться в геометрической прогрессии» и соглашался, что ограничения на подобный рост заключаются, прежде всего, в недостатке пищи. Однако в отличие от Мальтуса, Тимирязев считал, что продукты питания могут быть не только плодом сельского хозяйства, но и результатом достижений химии. Но и в этом случае количество производимой пищи имеет верхний предел, наступающий ввиду

ограниченности источников энергии. Синтез органических соединений — химический, либо происходящий в растениях — является реакцией эндотермической, происходящей с поглощением энергии. А поскольку запасы земных энергетических источников ограничены, как и поток (энергия в единицу времени) солнечной энергии, то и количество пищи, которую можно произвести биологическим либо химическим способом, ограничено. Таким образом, взгляды К.А. Тимирязева на ограниченность ресурсов представляли (в некотором смысле) более совершенный вариант риаисуждений Мальтуса.

4. Логистическая модель была предложена в работе [10]. Она допускает также формулировку в дискретном времени, которая широко используется в такой биологической науке, как популяционная динамика (см., напр., [11]).

Литература

1. An Essay On The Principle Of Population. 1798 (издание вышло анонимно, автором является Т. Мальтус).
2. *Мальтус Т.-Р.* Опыт закона о народонаселении. Пер. И.А. Вернера. М., 1895.
3. *Spencer H.* Principles of Biology (1864, 1867; revised and enlarged: 1898), in two volumes.
4. *Маркс К.* Капитал, т. I. — Маркс К., Энгельс Ф. Соч., т. 23, с. 645–646.
5. *Менделеев Д.И.* К познанию России. СПб, изд-во А.С. Суворина, 1906 (переиздание: М., Дрофа, 2002).
6. *Милль Дж.-Ст.* Основы политической экономии. Книга 1. Пер. Н.Г. Чернышевского. 1848.
7. *Ленин В.И.* Аграрный вопрос и «критики Маркса». — Полн. собр. соч., изд. 5., т. 5.
8. *Чернышевский Н.Г.* Замечания на последние четыре главы первой книги Милля (опубликованы Чернышевским в составе издания [7]).
9. *Тимирязев К.А.* Исторический метод в биологии. 1922 (стр. 107 по переизданию 1942 года).
10. *Verhulst P.F.* Notice sur la loi que la population poursuit dans son accroissement. Correspondance mathématique et physique. 1848. 10:113–121.
11. *Ризниченко Г.Ю., Рубин А.Б.* Математические модели биологических продукционных процессов. М.: изд-во МГУ, 1993, 302 с.

2.2. Модель Лотки

В первом разделе настоящей главы были рассмотрены модель Мальтуса, описывающая экспоненциальный демографический рост населения, и логистическая модель, учитывающая его ограничение из-за ограниченности ресурсов. Обе эти модели не принимают во внимание структуру социума, что делает их малоприменимыми для детального

демографического прогноза и анализа — как для глобальных, так и, тем более, для локальных процессов. Настоящий раздел посвящен так называемой *модели Лотки* (или Лотки–Шарпа), учитывающей возрастную структуру населения. Время в данной модели предполагается непрерывным. В связи с этим, она имеет вид линейного интегрального уравнения. Ниже рассматривается модель Лесли, которая является матричным аналогом модели Лотки.

Рассматриваемый ниже вариант модели является, что довольно типично для демографии, «однополым», т. е. рассматривается численность, возрастная структура и скорость роста лишь женского населения. При этом количество мужчин в модели в явном виде не рассматривается; можно условно считать, что в каждой возрастной группе оно равно числу женщин (либо находится в некотором заданном соотношении с ним).

Далее, первый пункт посвящен построению модели, во втором проводится построение решения, описывающего мальтузианские населения, т. е. решения с постоянной возрастной структурой и постоянной скоростью роста. В заключительном, третьем пункте анализируется устойчивость этих решений.

2.2.1. Построение модели

Обозначим через $n(x, t)$ функцию, описывающую распределение женского населения по возрастам, так что $n(x, t)dx$ — это количество женщин, находящихся в возрасте от x до $x + dx$ в момент времени t . Тогда его общая численность в этот момент равна

$$N(t) = \int_0^{+\infty} n(x, t)dx$$

(для реального населения, конечно, функция $n(x, t)$ является финитной, т. е. положительной лишь на конечном отрезке; интегрирование производится фактически по конечному интервалу).

Обозначим через $B(t)$ интенсивность рождений, тогда величина $B(t) dt$ равна количеству девочек, появившихся в течение временного интервала $(t, t + dt)$. Нашей ближайшей целью является получение уравнения для функции $B(t)$.

Введем функцию фертильности $\alpha(x, t)$ так, что количество девочек, рожденных за интервал времени $(t, t + dt)$ женщинами, находящимися в возрасте $(x, x + dx)$, равно $\alpha(x, t)n(x, t) dx dt$. Всюду далее будем считать, что при любом t функция $\alpha(x, t)$ отлична от нуля лишь в пределах конечного отрезка, который обозначим $x_{\min} \leq x \leq x_{\max}$, где $0 \leq x_{\min} < x_{\max}$. При этом интегрирование по x на полуоси $(0, +\infty)$ эквивалентно интегрированию по интервалу (x_{\min}, x_{\max}) .

Тогда количество девочек, рожденных за интервал $(t, t + dt)$ женщинами всех возрастов равно $\int_{x=0}^{+\infty} \alpha(x, t)n(x, t) dx dt$. Таким образом,

получаем

$$B(t) = \int_0^{+\infty} \alpha(x, t)n(x, t) dx.$$

Вероятность для отдельной женщины дожить до возраста x , достигаемого в момент t , обозначим через $l(x, t)$. В соответствии с демографическим содержанием функции $l(x, t)$ всюду далее без дополнительных оговорок будем считать ее убывающей по x достаточно быстро, чтобы соответствующие несобственные интегралы сходились (либо можно считать ее финитной по x с носителем, равномерно ограниченным по t , однако при этом необходимо делать определенные оговорки относительно выражений вида $l(\xi + t, t)/l(\xi, 0)$, см. ниже).

Женщины, находящиеся в возрасте $(x, x + dx)$ в момент времени t , были рождены в интервале $(t - x - dx, t - x)$. С учетом смертности их количество в этот момент времени равно $B(t - x)l(x, t) dx$. С другой стороны (по определению функции $n(x, t)$), их количество равно $n(x, t) dx$. Следовательно,

$$n(x, t) = B(t - x)l(x, t). \quad (1)$$

Таким образом, для интенсивности рождений имеем интегральное уравнение

$$B(t) = \int_0^{+\infty} B(t - x)\alpha(x, t)l(x, t) dx. \quad (2)$$

Оно обладает следующим свойством, затрудняющим анализ: так как переменная интегрирования x изменяется вплоть до бесконечности, то значение функции $B(t)$ в момент времени t выражается через ее значения во *все* предыдущие моменты времени $(t - x)$, в т. ч. соответствующие отрицательным значениям $t - x < 0$.

Между тем соображения демографического характера подсказывают, что задание плотности распределения по возрастам $n(x, 0)$ в начальный момент времени $t = 0$ способно заменить всю информацию об интенсивности рождений при отрицательных значениях t и о вероятности дожить до этого начального момента. Например, информация «50 лет назад в городе N родились 1000 женщин, и для каждой из них вероятность дожить до возраста 50 лет равна 92%» может быть представлена в виде «сегодня в городе N живут 920 женщин 50-летнего возраста». Аналогичные утверждения можно сделать о численности любой возрастной группы в момент времени $t = 0$, т. е. «сегодня».

Соответственно для построения прогноза (т. е. вычисления значений функции $B(t)$ при $t > 0$) необходима информация либо о рождениях и выживаемости в прошлом, либо о распределении по возрастам при $t = 0$. Например, если принять, что продолжительность жизни

индивида не превышает 100 лет, то для построения прогноза на год вперед необходимо иметь либо полную информацию о рождениях за последние 99 лет, либо функцию распределения по возрастам за один (сегодняшний) момент времени.

При этом распределение по возрастам в фиксированный момент времени является, как правило, более доступным для наблюдения и измерения, чем информация о демографических событиях (рождениях), углубленных во времени. Поэтому следует перейти от уравнения (2) к другому уравнению, которое будет ему эквивалентно и в котором будут фигурировать значения функции $B(t-x)$ лишь от неотрицательных аргументов.

Следуя указанным соображениям, разобьем интеграл в правой части уравнения (2), выделив отдельно интенсивности рождений в положительные и в отрицательные моменты времени:

$$B(t) = \int_0^t B(t-x)\alpha(x,t)l(x,t) dx + \int_t^{+\infty} B(t-x)\alpha(x,t)l(x,t) dx.$$

Сделаем во втором интеграле замену $\xi = x - t$, получим

$$B(t) = \int_0^t B(t-x)\alpha(x,t)l(x,t) dx + \int_0^{+\infty} B(-\xi)\alpha(\xi+t,t)l(\xi+t,t) d\xi.$$

Теперь наша ближайшая цель заключается в том, чтобы представить интеграл, содержащий $B(-\xi)$, в виде функции от аргумента t , другими словами — «упаковать» всю предысторию $-\infty < \xi < 0$ в функцию от одного момента времени. Для этого рассмотрим распределение по возрастам в момент $t = 0$. В возрасте от ξ до $\xi + d\xi$ находится $n(\xi, 0)d\xi$ женщин, они родились во временном промежутке от $-\xi - d\xi$ до $-\xi$, следовательно, их количество равно $B(-\xi)l(\xi, 0)d\xi$. Следовательно, $n(\xi, 0)d\xi = B(-\xi)l(\xi, 0)d\xi$, откуда

$$B(-\xi) = \frac{n(\xi, 0)}{l(\xi, 0)}.$$

Таким образом,

$$B(t) = \int_0^t B(t-x)\alpha(x,t)l(x,t) dx + \int_0^{+\infty} n(\xi, 0) \frac{l(\xi+t,t)}{l(\xi, 0)} \alpha(\xi+t,t) d\xi.$$

Обозначив последний интеграл, не содержащий неизвестной функции, через $H(t)$, получим основное уравнение демографической модели

Лотки (или Лотки–Шарпа):

$$B(t) = \int_0^t B(t-x)\alpha(x,t)l(x,t) dx + H(t). \quad (3)$$

Функции $\alpha(x,t)$ и $l(x,t)$ определяются социальными факторами (см. подробнее п.2.5) и в демографических моделях обычно считаются заданными экзогенно. Они определены в области $x \geq 0$, $t \in (-\infty, +\infty)$, являются неотрицательными; без ущерба для демографического смысла задачи их можно считать сколь угодно гладкими. Очевидно, модель является содержательной, если подынтегральная функция не является тождественным нулем, т.е. произведение $\alpha(x,t)l(x,t)$ будет положительным в некоторой области. Наконец, вероятность $l(x,t)$ дожить до возраста x обладает свойствами $l(0,t) = 1$ для любого t и $\partial l/\partial x < 0$ во всей области определения.

Таким образом, уравнение (3) представляет собой интегральное уравнение для интенсивности рождений $B(t)$. В результате его решения определяется функция $B(t)$, после этого не составит труда найти по формуле (1) распределение женского населения по возрастам $n(x,t)$ в любой момент времени.

Ниже рассматриваются некоторые решения уравнения (3) в случае функций $\alpha(x,t)$, $l(x,t)$, не зависящих от времени.

2.2.2. Случай мальтузианских населений

Основной вывод из модели Мальтуса (см. п. 2.1 настоящей главы) состоит в том, что при постоянном коэффициенте прироста численность населения изменяется по экспоненциальному закону $N(t) = N(0) \exp(rt)$. При этом сам коэффициент прироста задается в модели Мальтуса экзогенно (или же экзогенно могут задаваться коэффициенты рождаемости и смертности, разность которых равна r , представляет собой коэффициент прироста).

Модель Лотки, как и модель Мальтуса, является линейной, поэтому следует ожидать, что если функции фертильности и выживаемости не зависят от времени, то она также имеет экспоненциальные режимы. Нахождение таких режимов и анализ их устойчивости являются предметом настоящего и следующего подпунктов соответственно.

Итак, пусть фертильность и смертность не зависят от времени:

$$\alpha = \alpha(x), \quad l = l(x).$$

Тогда уравнение (1) для распределения женщин по возрастам и интегральное уравнение (2) для интенсивности рождений принимают, соответственно, вид

$$n(x,t) = B(t-x)l(x), \quad (4)$$

$$B(t) = \int_0^{+\infty} B(t-x)\alpha(x)l(x) dx, \quad (5)$$

а основное уравнение (3) модели Лотки — вид

$$B(t) = \int_0^t B(t-x)\alpha(x)l(x) dx + H(t). \quad (6)$$

Здесь функция $H(t)$ задается формулой

$$H(t) = \int_0^{+\infty} n(\xi, 0) \frac{l(\xi+t)}{l(\xi)} \alpha(\xi+t) d\xi.$$

В то время, как модель Мальтуса имеет только экспоненциальные решения, более содержательная модель Лотки (5) может иметь также решения другого вида. Поэтому в демографии вводится понятие *мальтузианского населения*, отвечающего решению уравнения (5) вида

$$B(t) = B(0) \exp(rt).$$

Здесь коэффициент прироста r является постоянной величиной, подлежащей определению в результате анализа модели. Подставив в (5), имеем после несложных преобразований

$$\int_0^{+\infty} e^{-rx} \alpha(x) l(x) dx = 1. \quad (7)$$

Равенство (7) представляет собой уравнение для r и называется уравнением Эйлера–Лотки.

Очевидно, левая часть уравнения (7),

$$\Phi(r) = \int_0^{+\infty} e^{-rx} \alpha(x) l(x) dx,$$

монотонно убывает по r . При этом $\Phi(r) < 1$ для достаточно больших (положительных) значений r , и $\Phi(r) > 1$ для достаточно больших по абсолютной величине отрицательных значений r . Следовательно, функция $\Phi(r)$ один раз обращается в единицу, и уравнение Эйлера–Лотки (7) имеет ровно один вещественный корень r_0 . Он называется коэффициентом Лотки.

Все другие корни уравнения (7) являются попарно комплексно-сопряженными, и модуль каждого из них строго меньше, чем r_0 . Таким образом, если интенсивность рождений $B(t)$ имеет несколько гармоник $B(t) = \sum C_i \exp(r_i t)$, то основная из них содержит выражение $\exp(r_0 t)$.

Для описания направления демографического воспроизводства вводится следующее определение.

Определение 1. *Воспроизводство населения называется расширенным, если коэффициент Лотки больше единицы, простым, если он равен единице, и суженным, если он меньше единицы.*

Нередко используется также другое определение расширенного (простого, суженного) воспроизводства, эквивалентное определению 1. Оно основано на понятии нетто-коэффициента воспроизводства населения, непосредственно содержащего возрастные показатели как рождаемости, так и смертности. В рамках модели Лотки нетто-коэффициент имеет вид

$$k_{\text{Netto}} = \int_0^{+\infty} \alpha(x)l(x) dx.$$

Этот демографический показатель имеет смысл количества девочек, которое родит за всю жизнь новорожденная женщина, при условии неизменности возрастных коэффициентов рождаемости и смертности. Таким образом, величина k_{Netto} показывает, в какой мере в стабильном населении поколение «дочерей» замещает поколение «матерей». Это позволяет сформулировать следующее определение.

Определение 2. *Воспроизводство населения называется расширенным, если нетто-коэффициент больше единицы, простым, если он равен единице, и суженным, если он меньше единицы.*

Очевидно, если нетто-коэффициент воспроизводства больше единицы, то $r_0 > 0$, и наоборот. Следовательно, определения 1 и 2 эквивалентны.

Произведение $\alpha(x)l(x)$ называется функцией нетто-фертильности.

Итак, мальтузианскому населению соответствует решение уравнения (5), имеющее вид $B(t) = B(0) \exp(r_0 t)$. Распределение этого населения по возрастам находится из (4):

$$n(x, t) = B(0)l(x)e^{r_0(t-x)} = B(0)l(x)e^{-r_0 x} e^{r_0 t}. \quad (8)$$

Таким образом, возрастная структура определяется функцией $l(x) \exp(-r_0 x)$ и явным образом находится из модели Лотки. Отметим, что в рамках модели Мальтуса общая численность населения также возрастает экспоненциальным образом, но распределение этого населения по возрастам не рассматривается. Связь между решениями моделей Мальтуса и Лотки можно показать следующим образом. Проинтегрируем левую и правую части равенства (8) по x на полуоси $(0, +\infty)$. В левой части получим общую численность населения

$$N(t) = \int_0^{+\infty} n(x, t) dx,$$

а в правой — выражение

$$e^{r_0 t} \int_0^{+\infty} B(0)l(x)e^{-r_0 x} dx.$$

Данный интеграл не зависит от времени; обозначив его через N_0 , получим выражение для общей численности населения в модели Мальтуса

$$N(t) = N_0 e^{r_0 t}.$$

Отметим, впрочем, что экспоненциальный рост (или убывание) общей численности населения в модели Лотки имеет место лишь в мальтузианском режиме, который реализуется начиная с момента $t = 0$ лишь при специально подобранном начальном распределении $n(x, 0)$, а для начальных условий общего вида является аттрактором решения при $t \rightarrow +\infty$. Обоснованию последнего тезиса посвящен следующий подпункт.

2.2.3. Анализ устойчивости мальтузианского режима

В предыдущем пункте была установлена скорость роста мальтузианского населения и его возрастная структура в рамках модели Лотки. Теперь необходимо исследовать его устойчивость. Здесь мы не будем приводить подробное математическое доказательство, а для простоты ограничимся лишь описанием его основной идеи для наиболее естественного случая, когда $n(x, 0)$ непрерывно и в начальный момент времени индивиды присутствуют во всех возрастах вплоть до некоторого возраста X , превышающего возраст выхода из фертильного периода x_{\max} , т. е. $n(x, 0) > 0$ при $0 \leq x \leq X$, где $X > x_{\max}$.

Покажем, что если начальное распределение $n(x, 0)$ (вообще говоря, немальтузианское) удовлетворяет этому условию, то при $t \rightarrow \infty$ на отрезке $0 \leq x \leq X$ устанавливается мальтузианская возрастная структура (8).

В силу непрерывности и положительности функции $n(x, 0)$ существуют ее верхний и нижний барьеры (мажоранта и миноранта) вида (8). Другими словами, существуют постоянные A и B такие, что

$$Al(x) \exp(-r_0 x) \leq n(x, 0) \leq Bl(x) \exp(-r_0 x), \quad 0 \leq x \leq X. \quad (9)$$

Из всех значений постоянной A , при которых выполняется неравенство (9), выберем наибольшее, из всех значений постоянной B — наименьшее. Обозначив их A_0 и B_0 соответственно, имеем

$$A_0 l(x) \exp(-r_0 x) \leq n(x, 0) \leq B_0 l(x) \exp(-r_0 x), \quad 0 \leq x \leq X. \quad (10)$$

На рис. 1 схематически представлены функции, стоящие во всех трех частях неравенства (8).

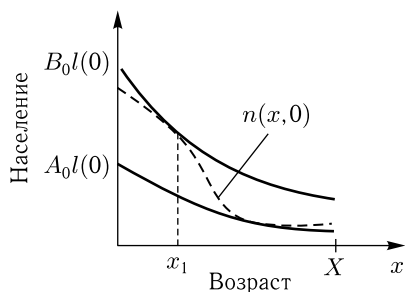


Рис. 1. Население с начальным возрастным распределением $n(x, 0)$ (пунктирная линия), мажорируемое и минорируемое двумя мальтузианскими населениями

Наряду с функцией $n(x, t)$, определяемой из решения модели Лотки с начальной функцией $n(x, 0)$, рассмотрим функции

$$\begin{aligned} n_{\min}(x, t) &= A_0 l(x) \exp[r_0(t - x)], \\ n_{\max}(x, t) &= B_0 l(x) \exp[r_0(t - x)]. \end{aligned} \quad (11)$$

Можно показать, что

$$n_{\min}(x, t) \leq n(x, t) \leq n_{\max}(x, t) \quad (12)$$

при всех $t \geq 0$, $0 \leq x \leq X$.

Другими словами, если два мальтузианских населения являются минорирующим и мажорирующим для данного (немальтузианского, имеющего начальное возрастное распределение $n(x, 0)$) населения в начальный момент времени, то они его минорируют и мажорируют во все последующие моменты.

При этом, поскольку неравенство (12) выполняется, в частности, для фертильного возрастного интервала $x_{\min} \leq x \leq x_{\max}$, то интенсивность рождений данного населения также заключена между интенсивностями рождений этих мальтузианских населений. Поэтому по прошествии некоторого времени $X - x_1$ (см. рис. 1) будут выполняться строгие неравенства (12). Следовательно, будет известен новый, более узкий и точный коридор, ограниченный двумя мальтузианскими населениями, в котором будет находиться данное население. Другими словами, существуют постоянные A_1, B_1 такие, что $A_0 < A_1 < B_1 < B_0$, и выполняются неравенства

$$A_1 l(x) \exp[-r_0 x] \leq n(x, X - x_1) \leq B_1 l(x) \exp(-r_0 x), \quad 0 \leq x \leq X.$$

Далее (по прошествии некоторого следующего интервала времени) для функции $n(x, t)$ можно будет указать еще более узкий коридор двумя мальтузианскими населениями и т. д. Можно доказать, что с каждой новой итерацией ширина коридора асимптотически сужается до нуля, т. е. $\lim_{k \rightarrow \infty} (B_k - A_k) = 0$.

Тем самым произвольное (вообще говоря, немальтузианское в начальный момент $t = 0$) распределение $n(x, t)$ с течением времени неограниченно приближается к мальтузианскому. Это означает, что мальтузианский режим демографического роста (или убывания) в модели Лотки является асимптотически устойчивым.

Упражнения

1. Доказать, что если $a + ib$ является корнем уравнения (7), то $a - ib$ также является корнем и $|a + ib| < r_0$.

2. Получить выражение для интенсивности смертей в случае мальтузианского населения, т. е. найти функцию $M(t)$ такую, что $M(t) dt$ — это количество умерших в интервале времени $(t, t + dt)$.

Комментарии и библиографические примечания

1. Модель Лотки была разработана в работах [1, 2], устойчивость мальтузианского населения по отношению к малым начальным возмущениям доказана в [1], при этом использовались методы теории интегральных уравнений. Приведенное выше доказательство, имеющее ясный демографический смысл и допускающее гораздо более широкий класс начальных функций $n(x, 0)$, получено в [3].

2. Основное ограничение применимости модели Лотки связано с тем, что рождаемость и смертность принимаются зависящими лишь от возраста. В реальных населенных пунктах вероятность того, что в определенном возрастном интервале женщина родит ребенка, может зависеть от значений различных социальных переменных — ее национальности, конфессиональной принадлежности, образовательного уровня, типа населенного пункта, в котором она проживает и т. д. Подробнее о влиянии социальных факторов на рождаемость см. п. 2.5.

3. Проведенный здесь анализ модели Лотки проведен, в основном, для случая функций фертильности и выживания, не зависящих от времени. Вопрос о том, насколько оправданным является подобное предположение, обсуждается в комментариях к следующему пункту, посвященному модели Лесли.

Литература

1. *Sharpe F. R., Lotka A. J.* A Problem of Age-Distribution // *Philos. Mag.* 1911. V. 21. P. 435–438.
2. *Lotka A. J.* A contribution to the theory of self-renewing aggregates, with special reference to industrial replacement // *Ann. Math. Stat.* 1939. V. 10. P. 1–25.
3. *Lotka A. J.* The Stability of the placeNormal Age Distribution // *Proc Natl Acad Sci USA.* 1922 November; 8(11): 339–345.

2.3. Модель Лесли

В настоящем разделе рассматривается демографическая модель с дискретным временем, учитывающая возрастную структуру популяции. Эта модель, разработанная английским биологом и математиком П. Лесли в 40-х годах XX века, является дискретным аналогом модели Лотки, представленной выше. Более конкретно: в модели Лесли распределение женщин по возрастам описывается через численности возрастных групп и тем самым имеет дискретный характер, в отличие от непрерывного распределения в модели Лотки. Таким образом, структура модели Лесли более соответствует характеру статистических данных о населении.

Указанное различие задает иной математический аппарат: если модель Лотки представляла собой интегральное уравнение, то модель Лесли является матричной. В то же время цели анализа и получаемые результаты вполне аналогичны: с течением времени в рамках каждой из моделей формируется мальтузианское население, имеющее постоянную возрастную структуру, определяемую в обеих моделях через старшую гармонику некоторого разложения; в модели Лесли эта структура задается старшим собственным вектором. Скорость роста мальтузианского населения находится из уравнения Эйлера–Лотки, имеющего в данных моделях одинаковый смысл и аналогичную форму. Тем самым модели Лотки и Лесли представляют собой пример ситуации, когда для описания одного и того же процесса могут быть предложены как дискретная, так и непрерывная модели, дающие аналогичные результаты. Между тем модель Лесли зачастую оказывается более удобной и (в определенном смысле) более практичной.

Далее: первый пункт настоящего параграфа раздела построению модели, второй пункт — выводу и анализу уравнения Эйлера–Лотки. В третьем пункте вводится такой широко известный инструмент демографического анализа, как возрастно-половая пирамида, а также кратко описываются демографические волны. Четвертый пункт носит технический характер. Наконец, в пятом пункте рассматриваются условия применимости модели Лесли и обсуждается такой, несколько экзотический для демографии пример, как данные о смертности населения волосовской культуры (II–III тысячелетие до нашей эры).

2.3.1. Построение модели

Данная модель, как и рассмотренная выше модель Лотки, формулируется для закрытого женского населения, в частности это означает, что миграция не учитывается.

Разобьем интервал изменения возраста индивидов на равные отрезки, причем длину этих отрезков выберем равной шагу дискретного времени модели (в практических исследованиях, как правило, этот шаг принимается равным одному или пяти годам). Тем самым, задается разбиение женского населения на возрастные группы.

Пронумеруем группы в порядке возрастания так, что самая младшая из них (например, женщины первого года жизни) будет иметь номер 1, а самая старшая — номер n . Например, для пятилетних возрастных групп обычно принимается $n = 20$. Обозначим через $x_i(t)$ численность i -й возрастной группы в момент времени t . Таким образом, состояние популяции в этот момент времени характеризуется вектором

$$X(t) = \begin{pmatrix} x_1(t) \\ x_2(t) \\ \dots \\ x_n(t) \end{pmatrix}.$$

Обозначим через α_i вероятность рождения девочки у представительницы i -й группы, а через s_i — вероятность для представителя i -й группы дожить до попадания в следующую, $(i + 1)$ -ю группу. Параметры α_i ($i = 1, \dots, n$) и s_i ($i = 1, \dots, n - 1$) — коэффициенты рождаемости и выживания соответственно; в простейшем варианте модели они считаются постоянными во времени. С социологической точки зрения представляет значительный интерес анализ зависимости этих коэффициентов от социальных и экономических параметров общества, однако в собственно демографических моделях значения α_i и s_i предполагаются заданными экзогенно, т. е. извне. Заметим также, что чаще всего в демографии употребление термина «возрастной коэффициент рождаемости» относится к случаю, когда за единицу времени принят год или пять лет, а под «группой» понимается лишь женское население соответствующего возраста.

Перейдем к выводу матричного уравнения модели Лесли.

Поскольку во вторую возрастную группу в момент времени $t + 1$ попадают дожившие до этого момента женщины, которые в момент времени t находились в первой группе, то

$$x_2(t + 1) = s_1 x_1(t). \quad (1)$$

Аналогично получим для следующих возрастных групп

$$x_i(t + 1) = s_{i-1} x_{i-1}(t), \quad i = 3, \dots, n - 1. \quad (2)$$

Численность новорожденных, находящихся в первой возрастной группе в момент времени $t + 1$, и рожденных от матерей из первой возрастной группы, равна $\alpha_1 x_1(t)$, от матерей из второй группы — $\alpha_2 x_2(t)$ и т. д. Таким образом, для общего количества новорожденных имеем

$$x_1(t + 1) = \alpha_1 x_1(t) + \alpha_2 x_2(t) + \dots + \alpha_n x_n(t). \quad (3)$$

Уравнения (1)–(3) агрегируются в одно матричное уравнение

$$X(t + 1) = LX(t), \quad (4)$$

которое является основным уравнением модели. Здесь матрица L , называемая матрицей Лесли, имеет вид

$$L = \begin{pmatrix} \alpha_1 & \alpha_2 & \alpha_3 & \dots & \alpha_{n-1} & \alpha_n \\ s_1 & 0 & 0 & \dots & 0 & 0 \\ 0 & s_2 & 0 & \dots & 0 & 0 \\ 0 & 0 & s_3 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & s_{n-1} & 0 \end{pmatrix}. \quad (5)$$

Все ее элементы неотрицательны, при этом не равны нулю могут быть лишь элементы первой строки (если за единицу времени принимается год, то положительны, как правило, элементы примерно с пятнадцатого по сорок пятый) и поддиагональные элементы (так как коэффициенты s_i имеют смысл вероятностей, то все они лежат в интервале $0 < s_i < 1$).

В рассматриваемом здесь базовом варианте модели все возрастные коэффициенты рождаемости и выживания предполагаются не зависящими от времени, вследствие чего матрица L является постоянной.

Если известно распределение численности населения по возрастным группам в начальный момент $t = 0$, то не составляет труда определить вектор-столбец $X(t)$ для любого момента времени t по формулам

$$\begin{aligned} X(1) &= LX(0), \\ X(2) &= LX(1) = L^2X(0), \\ &\dots \\ X(t) &= LX(t-1) = L^tX(0). \end{aligned}$$

Относительно модели Лесли представляют интерес, в частности следующие вопросы: устанавливается ли с течением времени постоянная возрастная структура, т.е. соотношение между компонентами вектора $X(t)$? Зависит ли эта структура от начальных условий? Могут ли происходить колебания в системе? При каких условиях будет наблюдаться рост, а в каких — убывание общей численности населения?

Изложим нестрого основную идею, лежащую в основе анализа модели Лесли.

Матрица L имеет, вообще говоря, как вещественные, так комплексные собственные значения λ_i , $i = 1, \dots, n$. Пронумеруем их в порядке невозрастания модуля:

$$|\lambda_1| \geq |\lambda_2| \geq \dots \geq |\lambda_n|.$$

Примем упрощающее предположение, что все собственные значения различны, тогда собственные векторы матрицы Лесли $\varphi_1, \varphi_2, \dots, \varphi_n$ образуют базис (при наличии кратных собственных значений анализ модели существенно усложняется). Разложив по этому базису начальный вектор $X(0)$, получим

$$X(0) = C_1\varphi_1 + C_2\varphi_2 + \dots + C_n\varphi_n.$$

Тогда

$$X(1) = C_1 L \varphi_1 + C_2 L \varphi_2 + \dots + C_n L \varphi_n = C_1 \lambda_1 \varphi_1 + C_2 \lambda_2 \varphi_2 + \dots + C_n \lambda_n \varphi_n.$$

Для значений вектора $X(t)$ в следующие моменты времени имеем

$$\begin{aligned} X(t) &= C_1 L^t \varphi_1 + C_2 L^t \varphi_2 + \dots + C_n L^t \varphi_n = \\ &= C_1 \lambda_1^t \varphi_1 + C_2 \lambda_2^t \varphi_2 + \dots + C_n \lambda_n^t \varphi_n. \end{aligned}$$

Отсюда следует, что если $|\lambda_1| > |\lambda_2|$, то при достаточно больших t слагаемое $C_1 \lambda_1^t \varphi_1$ существенно больше, чем сумма всех остальных, и имеет место приближенное равенство $X(t) \approx C_1 \lambda_1^t \varphi_1$.

Таким образом, с течением времени устанавливается режим, при котором:

- соотношение численностей возрастных групп близко к соотношению компонент собственного вектора φ_1 ;
- на каждом временном шаге численность как популяции в целом, так и каждой из возрастных групп изменяется примерно в λ_1 раз.

Тем самым устанавливается мальтузианский режим с коэффициентом естественного прироста, равным $r = \lambda_1 - 1$. Соответствующее население называется также стабильным, что подчеркивает асимптотическую устойчивость данной возрастной структуры.

Чтобы выяснить, имеет ли население тенденцию к росту или к сокращению, надо определить условия, при которых имеет место неравенство $\lambda_1 > 1$. Данному вопросу посвящен следующий пункт.

2.3.2. Уравнение Эйлера–Лотки

Чтобы определить максимальное по модулю собственное значение матрицы L , выпишем ее характеристическое уравнение,

$$L\varphi = \lambda\varphi.$$

Обозначим i -ю компоненту собственного вектора через μ_i и перепишем характеристическое уравнение в покомпонентном виде. В предположении, что $\lambda \neq 0$, получим

$$\begin{aligned} \mu_2 &= \frac{s_1}{\lambda} \mu_1, \\ \mu_3 &= \frac{s_2}{\lambda} \mu_2 = \frac{s_1 s_2}{\lambda^2} \mu_1, \\ &\dots \\ \mu_n &= \frac{s_{n-1}}{\lambda} \mu_{n-1} = \frac{s_1 s_2 \dots s_{n-1}}{\lambda^{n-1}} \mu_1. \end{aligned}$$

Подставив все эти равенства в первую строку векторного уравнения $L\varphi = \lambda\varphi$, получим

$$\alpha_1 \mu_1 + \alpha_2 \frac{s_1}{\lambda} \mu_1 + \dots + \alpha_n \frac{s_1 s_2 \dots s_{n-1}}{\lambda^{n-1}} \mu_1 = \lambda \mu_1$$

или

$$\alpha_1 \frac{1}{\lambda} + \alpha_2 \frac{s_1}{\lambda^2} + \dots + \alpha_n \frac{s_1 s_2 \dots s_{n-1}}{\lambda^n} = 1. \quad (6)$$

Заметим, что входящие в матрицу Лесли вероятности дожития s_i могут быть выражены через доли новорожденных, доживающих до возрастов $i + 1$ и i : $s_i = l_{i+1}/l_i$. Тогда произведения этих вероятностей также могут быть выражены через l_i , например

$$s_1 s_2 \dots s_{n-2} s_{n-1} = l_2 \frac{l_3}{l_2} \dots \frac{l_{n-1}}{l_{n-2}} \frac{l_n}{l_{n-1}} = l_n.$$

Уравнение (6) принимает вид

$$\alpha_1 \frac{1}{\lambda} + \alpha_2 \frac{l_2 b}{\lambda^2} + \dots + \alpha_n \frac{l_n}{\lambda^n} = 1$$

или (с учетом того, что $l_1 = 1$)

$$\sum_{i=1}^n \frac{\alpha_i l_i}{\lambda^i} = 1. \quad (7)$$

Уравнение (7) называется уравнением Эйлера–Лотки. Ему удовлетворяют все ненулевые собственные значения матрицы Лесли.

С помощью уравнения Эйлера–Лотки нетрудно доказать существование и единственность вещественного положительного собственного значения матрицы Лесли. Действительно, левая часть уравнения (7) является положительной монотонно убывающей функцией от λ при $\lambda > 0$, бесконечно возрастающей при $\lambda \rightarrow +0$ и бесконечно убывающей при $\lambda \rightarrow +\infty$. Следовательно, эта функция обращается в единицу ровно при одном значении $\lambda = \lambda_1$.

Отметим, что из проведенного анализа уравнения Эйлера–Лотки не следует, что единственное положительное собственное значение матрицы Лесли является простым и превышает все остальные по модулю. Достаточным условием этого будет положительность хотя бы двух элементов первой строки матрицы Лесли.

Какова роль фундаментальных демографических моделей (таких как обсуждаемая здесь модель Лесли) в изучении динамики реальных населений? Уравнение Эйлера–Лотки позволяет определить скорость роста $r = \lambda_1 - 1$ в мальтузианском режиме. В свою очередь мальтузианский режим устанавливается по прошествии определенного (вообще говоря, не обязательно малого) времени при условии постоянства коэффициентов рождаемости и выживания. Другими словами, скорость роста $r = \lambda_1 - 1$ является характеристикой не реального, а соответствующего ему гипотетического стабильного населения, для которого параметры α_i, s_i постоянны и равны сегодняшним значениям этих параметров для реального населения.

Практическая значимость концепции стабильного населения связана с тем, что (при анализе реальных населений) одна из главных трудностей заключается в необходимости отделить фундаментальные тенденции, свойственные данному населению, от сиюминутного влияния текущей возрастной структуры. Например, увеличение коэффициента

естественного прироста в некоторой стране может свидетельствовать как о тенденции к увеличению рождаемости, так и о том, что в фертильный возраст вступило новое, более многочисленное поколение женщин.

Проведенный выше упрощенный анализ показывает, как в рамках модели Лесли каждому реальному населению (т. е. каждому набору возрастных коэффициентов рождаемости α_i и выживаемости s_i , взятых в совокупности с начальным вектором) сопоставить определенное стабильное мальтузианское население.

Коэффициент естественного прироста стабильного населения (т. е. величина λ_1) называется коэффициентом Лотки или истинным коэффициентом естественного прироста соответствующего реального населения.

Коэффициент Лотки позволяет установить «истинную» динамику численности населения. Действительно, если, например, численность населения при переходе от $t = 0$ к $t = 1$ и затем к $t = 2$ монотонно увеличивается, но при этом имеет место неравенство $\lambda_1 < 1$, то население имеет тенденцию к уменьшению, несмотря на видимый рост численности, возникающий ввиду благоприятных для кратковременного роста начальных условий. Таким образом, тенденция реального населения к увеличению или уменьшению определяется тенденцией соответствующего стабильного населения, именно — соотношением между коэффициентом Лотки и единицей.

Высказанные соображения позволяют сформулировать следующее определение.

Определение 1. *Воспроизводство населения называется расширенным, если коэффициент Лотки больше единицы, простым, если он равен единице, и суженным, если он меньше единицы.*

Широко распространенным в демографии является другое определение расширенного / простого / суженного воспроизводства, эквивалентное определению 1. Оно основано на понятии нетто-коэффициента воспроизводства населения, учитывающего возрастные показатели как рождаемости, так и смертности. В рамках модели Лесли нетто-коэффициент имеет вид

$$k_{\text{Netto}} = \sum_{i=1}^n \alpha_i l_i.$$

Этот демографический показатель имеет смысл количества девочек, которое родит за всю жизнь новорожденная женщина при условии неизменности возрастных коэффициентов рождаемости и смертности. Таким образом, k_{Netto} показывает, в какой мере в стабильном населении поколение «дочерей» замещает поколение «матерей». Это позволяет сформулировать следующее определение.

Определение 2. *Воспроизводство населения называется расширенным, если нетто-коэффициент больше единицы, простым, если он равен единице, и суженным, если он меньше единицы.*

Очевидно, определения 1 и 2 эквивалентны. Действительно, ввиду монотонного убывания левой части уравнения Эйлера–Лотки (7) по λ имеем $\lambda_1 < 1$, если $\sum_{i=1}^n \alpha_i l_i < 1$, и $\lambda_1 > 1$, если $\sum_{i=1}^n \alpha_i l_i > 1$.

Нетрудно видеть, что понятие нетто-коэффициента, а также определения 1, 2 аналогичны соответствующим положениям, изложенным в п. 2.2 для модели Лотки.

2.3.3. Возрастно-половая пирамида и демографические волны

Как показано выше, мальтузианское население имеет возрастную структуру, соответствующую старшему собственному вектору матрицы Лесли. Что касается реальных населений, то их структура может быть весьма разнообразной, что делает актуальной задачу ее визуализации.

Распределение населения по полам и возрастам в данный момент времени обычно представляют с помощью так называемой *возрастно-половой* (популяционной, демографической) *пирамиды*, представляющей собой диаграмму, на которой численность мужского и женского населения каждой возрастной группы отображается с помощью горизонтальной полосы большей или меньшей длины.

Визуализация возрастно-половой структуры населения с помощью пирамиды позволяет, в частности, наглядно представить соотношение между численностью трудоспособного населения, детей и пенсионеров.

В соответствии с общей формой графика выделяют три типа возрастно-половых пирамид: правильная (преобладает молодое население), колокол (за исключением самых старших групп, население равномерно распределено по возрастам) и урна (ниже некоторого возраста численность младших групп ниже, чем старших), см. рис. 1.

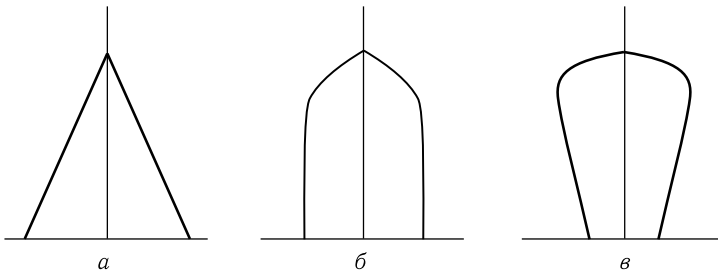


Рис. 1. Типы возрастно-половых пирамид: *a* — правильная, *b* — колокол, *v* — урна

В качестве примеров можно привести возрастно-половые пирамиды 2010 г. для Аргентины (правильный тип, рис. 2), США (пирамида-колокол, рис. 3) и Бельгии (пирамида-урна, рис. 4).

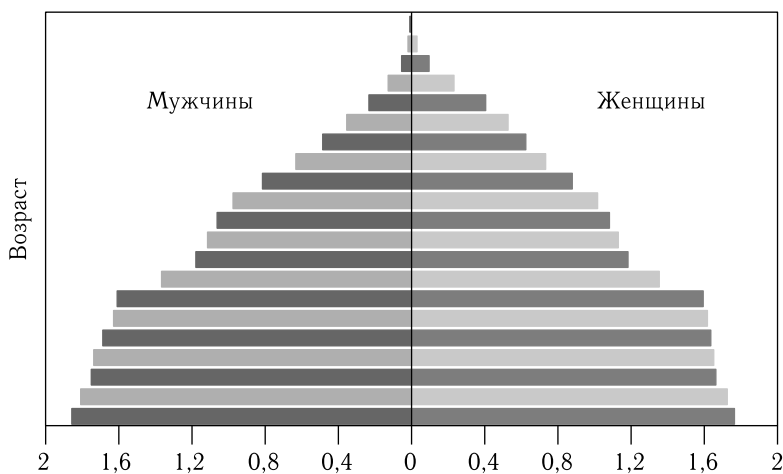


Рис. 2. Возрастно-половая пирамида для Аргентины в 2010 г. По оси абсцисс обозначена численность населения (млн чел.). Источник: US Census Bureau



Рис. 3. Возрастно-половая пирамида для США в 2010 г. По оси абсцисс обозначена численность населения (млн чел.). Источник: US Census Bureau

Тип возрастно-половой пирамиды обычно связывают с характером воспроизводства населения: утверждается, что правильная пирамида соответствует расширенному воспроизводству населения, пирамида-колокол — простому типу воспроизводства, пирамида-урна — суженному типу воспроизводства. Однако, строго говоря, это соответствие является не вполне точным. Действительно, компоненты μ_i старшего собственного вектора матрицы Лесли связаны соотношением

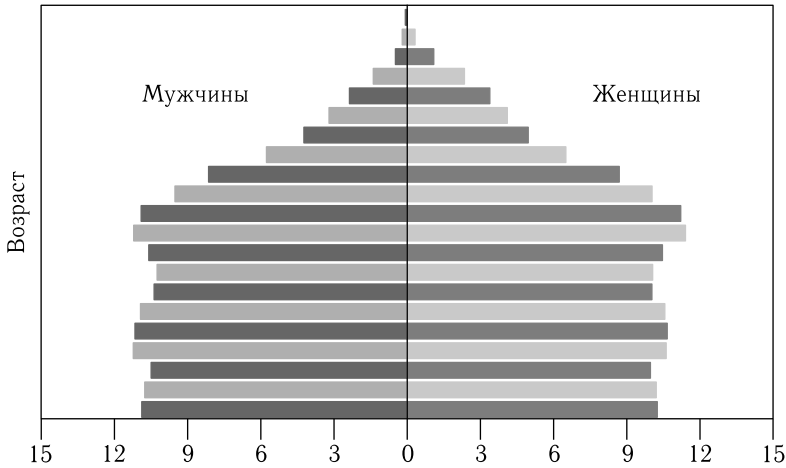


Рис. 4. Возрастно-половая пирамида для Бельгии в 2010 г. По оси абсцисс обозначена численность населения (тыс. чел.). Источник: US Census Bureau

$\mu_{i+1} = (s_i/\lambda_1)\mu_i$. Следовательно, если коэффициент Лотки меньше единицы, но превосходит все коэффициенты выживания ($s_i < \lambda_1$ для любого i), то пирамида имеет правильный тип ($\mu_{i+1} < \mu_i$) даже при установившемся режиме суженного воспроизводства населения.

Однако для развитых государств коэффициенты выживания для младших групп очень близки к единице. Поэтому для реальных населений пирамида-урна возникает вследствие совместного действия двух факторов: не только суженного воспроизводства населения, но и малой смертности в молодых возрастных группах. Действительно, в современных развитых государствах возрастные коэффициенты выживания, как правило, монотонно убывают начиная с группы 5–9 лет (в качестве характерного примера приведем данные для России по состоянию на 2008 г. — рис. 5).

Ввиду убывания коэффициентов выживания s_i почти от единицы до довольно малых значений при $\lambda_1 < 1$ наиболее типичным для развитых стран оказывается случай, при котором $s_i > \lambda_1$ (соответственно $\mu_{i+1} > \mu_i$) для младших возрастных групп и $s_i < \lambda_1$, $\mu_{i+1} < \mu_i$ — для старших. Таким образом, в старших группах возрастно-половая пирамида сужается кверху, а в младших группах сужается книзу. Тем самым возникает пирамида-урна.

Таким образом, пирамида-урна действительно соответствует суженному типу воспроизводства населения, но лишь в случае достаточно низкой детской смертности.

Перейдем к описанию на основании модели Лесли такого явления, как демографические волны.

Пирамиды с относительно ровными краями характерны для стран, население которых не испытывало демографических шоков: прежде

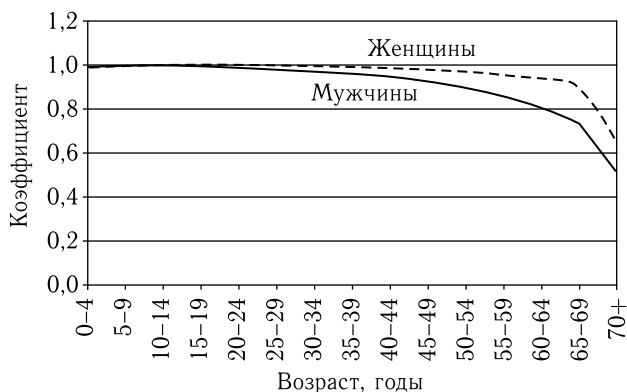


Рис. 5. Коэффициенты выживания по возрастным группам (Россия, 2008)

всего, резкого уменьшения количества рожденных в относительно короткий период времени по сравнению с соседними периодами. Если же подобное малочисленное поколение появилось на свет, то через 20–25 лет оно вступает в фертильный возраст и производит новое поколение, малочисленное по сравнению с рожденными несколько раньше или позже. Тем самым возникает чередование малолюдных и многолюдных поколений, называемое демографической волной.

Актуальность анализа демографических волн связана, в частности, с их влиянием на экономическую жизнь общества: чередование многочисленных и малочисленных поколений приводит к диспропорциям между численностью работающих и численностью пенсионеров, к колебаниям количества школьников и вследствие этого — к колебаниям заполненности школ и т.д.

Проиллюстрируем демографические волны на следующем элементарном примере.

Пример. Пусть матрица Лесли имеет вид

$$L = \begin{pmatrix} 0 & 0 & \alpha \\ s_1 & 0 & 0 \\ 0 & s_2 & 0 \end{pmatrix}, \quad (8)$$

где $\alpha = (s_1 s_2)^{-1}$. Если начальный вектор равен

$$X(0) = \{x_1, x_2, x_3\}^*,$$

то непосредственным вычислением нетрудно получить, что

$$\begin{aligned} X(1) &= \{\alpha x_3, s_1 x_1, s_2 x_2\}^*, \\ X(2) &= \{\alpha s_2 x_2, s_1 \alpha x_3, s_2 s_1 x_1\}^*, \\ X(3) &= \{x_1, x_2, x_3\}^* = X(1) \end{aligned}$$

и т. д.

Чтобы построить в иллюстративных целях возрастную пирамиду при различных значениях t , положим для примера $\alpha = 25/16$, $s_1 = 4/5$, $s_2 = 4/5$ и начальные условия $X(0) = \{6, 2, 1\}^*$. Тогда при $t = 0, 1, 2, 3$ пирамида примет вид (рис. 6).



Рис. 6. Демографическая волна (наиболее многочисленные поколения выделены черным цветом). Пирамида в момент $t = 3$ эквивалентна пирамиде в момент $t = 0$

На рис. 6 видно, что наиболее многочисленное поколение пребывало в младшем возрасте в момент $t = 0$, в среднем — в момент $t = 1$, в старшем — в момент $t = 2$. Оно породило новое многочисленное поколение, которое пребывает в младшем возрасте в момент $t = 3$. Таким образом, через каждую возрастную группу поочередно проходят многочисленные и малочисленные поколения, т. е. имеет место демографическая волна.

В представленном примере лишь один из возрастных коэффициентов рождаемости являлся положительным. Ввиду этого процесс не выходит на режим мальтузианского роста (или убывания) населения. Другими словами, в данном примере демографическая волна не размывается, и население не выходит на режим с постоянным соотношением между возрастными группами. Однако в ситуациях, возникающих в реальной демографии, положительным является более, чем один возрастной коэффициент рождаемости, и (в соответствии с п. 2.3.2) все собственные значения матрицы Лесли имеют модуль, меньший, чем λ_1 , вследствие чего демографические волны размываются: с течением времени различия в численности поколений сглаживаются и устанавливается мальтузианское соотношение между численностями возрастных групп.

2.3.4. Учет продолжительного пребывания индивидов в старшей группе

В модели Лесли предполагается, что за один шаг времени каждый индивид либо переходит в следующую возрастную группу, либо погибает. Однако, так как в старшей группе (например, 100+) индивид может прожить более чем один временной интервал, то модель Лесли может быть видоизменена введением в правый нижний угол матрицы L ненулевого элемента s_n :

$$L = \begin{pmatrix} \alpha_1 & \alpha_2 & \alpha_3 & \dots & \alpha_{n-1} & \alpha_n \\ s_1 & 0 & 0 & \dots & 0 & 0 \\ 0 & s_2 & 0 & \dots & 0 & 0 \\ 0 & 0 & s_3 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & s_{n-1} & s_n \end{pmatrix}.$$

Как указывалось выше, в актуальных (с точки зрения демографии) случаях значительное количество последних коэффициентов рождаемости равны нулю ($n - k > 0$), вследствие чего наличие ненулевого коэффициента выживания s_n приводит лишь к минимальным изменениям в анализе модели. Действительно, в этом случае кратность нулевого собственного значения понижается на единицу, и у матрицы Лесли появляется второе (наряду с числом Фробениуса) положительное собственное значение $\lambda_2 = s_n$. В практически значимых ситуациях оно заведомо меньше числа Фробениуса матрицы L .

Наличие диагональных членов характерно для модели Лefковича, описывающей динамику некоторых биологических популяций и в определенном смысле обобщающей модель Лесли. В модели Лefковича компоненты вектора состояния соответствуют не возрастным группам, а стадиям развития организма, например, у пашенного червя: «кокон» → «ювенил» → «половозрелая стадия 1» → «половозрелая стадия 2» → «половозрелая стадия 3» → «половозрелая стадия 4». С течением времени особи могут не только переходить в следующую стадию, но и задерживаться в текущей. Таким образом, матрица модели Лefковича имеет ненулевые члены в первой строке, на главной диагонали и под ней.

2.3.5. Об условиях применимости модели Лесли

Рассмотрим вопрос о том, при каких условиях модель Лесли может быть применима к описанию реальных демографических процессов. Для этого перечислим предпосылки, при которых она была получена:

- отсутствие миграции (закрытость населения),
- постоянство коэффициентов рождаемости и выживания,
- возможность ограничиться разбиением населения лишь на возрастные (или возрастно-половые) группы, без учета других структурных особенностей.

В отношении первого пункта кратко отметим, что модель может быть очевидным образом обобщена путем введения векторов, описывающих миграционные потоки.

Для иллюстрации динамики возрастных коэффициентов рождаемости приведем их значения для России за 28-летний период (рис. 7), а также данные для Австралии за временной интервал 1921–2006-х гг. (рис. 8).

Данные показывают, что изменения в социуме могут приводить к существенным изменениям коэффициентов рождаемости. При этом имеют место как долговременные тенденции, возникающие ввиду изменений установок женщин на рождение того или иного количества детей, так и краткосрочные изменения, связанные с конкретными социальными ситуациями или даже организованными кампаниями, направленными на повышение рождаемости (напр., некоторый всплеск рождаемости в СССР во время проведения кампании 1980-х гг.).

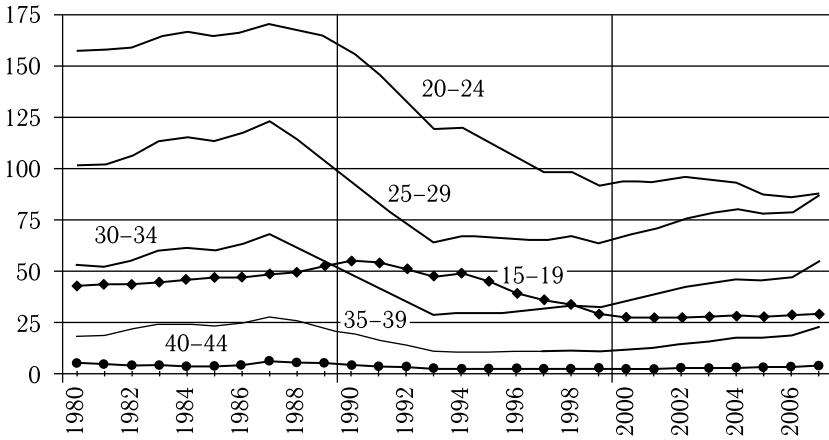


Рис. 7. Количество (по годам) родившихся за год живыми на 1000 женщин соответствующей возрастной группы в России. По данным: Демографический ежегодник России. 2008

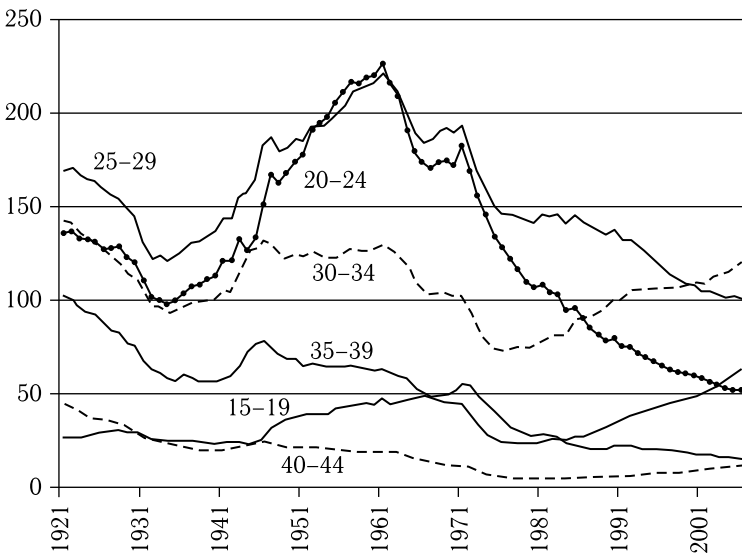


Рис. 8. Количество (по годам) родившихся за год живыми на 1000 женщин соответствующей возрастной группы в Австралии. По данным: Australian Bureau of Statistics

Вероятно, наиболее общая закономерность, которую можно указать в данном аспекте, заключается в том, что чем быстрее изменяется социум, тем быстрее изменяются и возрастные коэффициенты рождаемости и тем короче интервал времени, на котором демографическая

динамика может описываться моделью Лесли (по крайней мере, в плане рождаемости).

Аналогичное утверждение можно высказать и относительно коэффициентов выживания: чем стабильнее социум, тем больше оснований полагать их приблизительно постоянными. Особую роль в изменчивости показателей смертности может играть социальный стресс. Так, для населения России стресс первой половины 90-х годов XX века привел к существенному уменьшению ожидаемой продолжительности жизни (особенно для мужчин), о чем свидетельствует рис. 9.

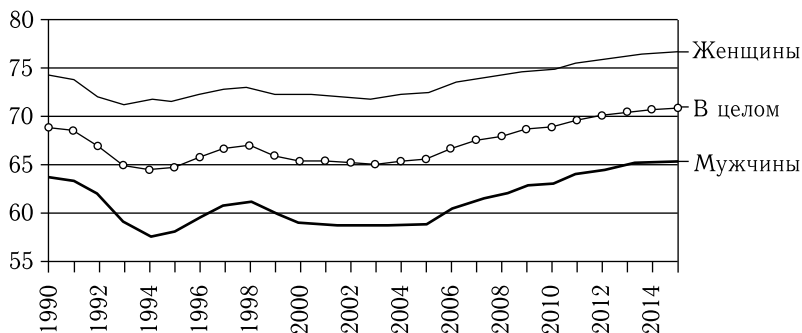


Рис. 9. Ожидаемая (в годах) продолжительность жизни при рождении в России (по данным Всемирного Банка)

Наконец, проиллюстрируем на примере необходимость учета иных (помимо половозрастных) структурных особенностей населения в некоторых случаях.

На рис. 10 представлены данные об ожидаемой продолжительности жизни при рождении для четырех групп населения США за 1970–2007-е гг.: белых женщин, черных женщин, белых мужчин и черных мужчин.

Важным свойством этих данных является различие продолжительности жизни (а следовательно, и коэффициентов выживания) для белого и черного населения США. Это свидетельствует о неоднородности населения и вследствие этого о том, что в некоторых случаях необходимо учитывать не только возрастно-половую структуру, но и иные аспекты, в частности расовый. Показатели, не учитывающие расовую структуру, могут не просто оказаться грубыми. Они способны ввести в заблуждение, так как их динамика может быть связана с изменением соотношения рас в населении.

В качестве примера приведем следующий статистический парадокс. В период с 1900 по 1920 гг. ожидаемая продолжительность жизни белых женщин увеличилась на 20%, а черных — на 40%. Интуитивно может показаться, что ожидаемая продолжительность жизни женщин «в целом» должна увеличиться на некоторую промежуточную между

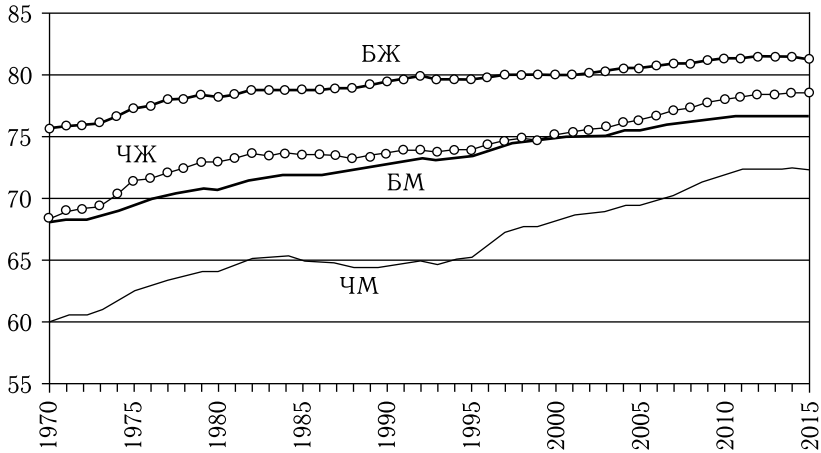


Рис. 10. Ожидаемая (в годах) продолжительность жизни при рождении для белых женщин (БЖ), черных женщин (ЧЖ), белых мужчин (БМ), черных мужчин (ЧМ) в США за 1970–2015 гг. Источник: National Vital Statistics System

20 % и 40 % величину. Однако значение показателя для обеих групп, взятых в совокупности, оказывается меньше, чем для каждой из групп по отдельности: именно, 13 %. Причина парадокса заключается в том, что за 20 лет в женском населении существенно увеличилась доля черных женщин. Описанная ситуация не является исключительной: в частности, ожидаемая продолжительность жизни мужского населения в целом также оказывается меньше, чем для каждой из составляющих его расовых групп. Это свидетельствует о том, что динамика характеристик смертности, не учитывающих расовую структуру населения США, может давать неправильное представление о процессе.

Таким образом, одно из важных ограничений применимости модели Лесли связано с тем, что она учитывает в лучшем случае лишь возрастно-половую структуру населения. Модель не может быть непосредственно применена для исследования демографических процессов, в которых важную роль играет структура населения, основанная на различиях между городским и сельским населением, на расовых, национальных, культурных различиях и т. д. Наиболее широкие возможности для учета структуры населения предоставляют имитационные модели (см. п. 2.5).

Завершая обсуждение вопроса о границах применимости модели Лесли, необходимо отметить, что в XX веке происходили достаточно бурные изменения общества, которые не могли не сказаться на динамике демографических характеристик. В более спокойные исторические периоды элементы матрицы Лесли изменялись, видимо, существенно более медленно.

Так, в палеодемографии укоренилась исследовательская позиция, в соответствии с которой коэффициенты выживания считаются постоянными для интервалов времени, существенно более протяженных, чем в «обычной» демографии. В качестве примера приведем данные относительно смертности населения волосовской культуры (II–III тысячелетие до нашей эры) памятника Сахтыш Па (Ивановская область) см. табл. 1. Заметим, что ввиду специфики археологических исследований, изучить детскую смертность в таких случаях обычно не удается, поэтому в табл. 1 в качестве самой молодой выбрана группа 15–19 лет, а представленные данные имеют вид вероятностей дожития до определенного возраста для индивидов, попавших в эту группу (таким образом, данные не учитывают детскую смертность).

Таблица 1. Вероятность дожития до попадания в ту или иную возрастную группу для индивида, находящегося в возрастной группе 15–19 лет. Источник: [11]

Возраст, гг.	Вероятность дожития		
	Для обоих полов	Для мужчин	Для женщин
15–19	100,0	100,0	100,0
20–24	95,3	95,2	93,75
25–29	93,0	95,2	87,50
30–34	86,0	90,5	81,25
35–39	60,4	66,7	50,0
40–44	39,5	42,8	37,5
45–49	25,6	14,3	37,5
50+	7,0	4,8	12,5

Примечательно, что уже в столь раннюю историческую эпоху характеристики смертности существенно зависели как от биологических, так и от социальных факторов. Так, в молодом возрасте женская выживаемость была ниже мужской ввиду риска смерти при рождении детей, а в старших возрастных группах резко возрастает мужская смертность вместе с вероятностью гибели в опасных ситуациях, связанных с социальной ролью мужчины: на охоте, в военных столкновениях и пр.

Для исследования системы и построения прогноза на временных интервалах, настолько больших, что на них нельзя считать коэффициенты выживания и рождаемости постоянными, можно рассматривать аналог модели Лесли с переменными коэффициентами, причем значения этих коэффициентов также подлежат прогнозированию. Однако в практике построения долгосрочных демографических прогнозов

приходится, как правило, учитывать социальную структуру населения. Поэтому в таких случаях более предпочтительно использовать не модель Лесли, а имитационное моделирование (см. п. 2.5). Однако важнейшие характеристики демографического процесса (такие, как коэффициент Лотки, тип воспроизводства населения и т. д.) могут быть введены лишь на основе таких фундаментальных моделей, как модели Лотки и Лесли.

Упражнения

1. Дана матрица Лесли со 100 возрастными группами (шаг времени 1 г.), т. е. известны $\alpha_1, \dots, \alpha_{100}, s_1, \dots, s_{99}$. Кроме того, известно распределение по возрастам $X = \{x_1, \dots, x_{100}\}^*$. Приняв упрощающее предположение, что в старшей возрастной группе женщина находится в течение ровно одного года, получить выражения для следующих показателей:

- ожидаемую продолжительность жизни при рождении,
- ожидаемую продолжительность жизни для 50-летней (ожидаемая продолжительность жизни — демографический показатель, имеющий смысл среднего количества лет предстоящей жизни человека, достигшего данного возраста),
- общие коэффициенты рождаемости и смертности (отношение количества родившихся живыми и числа умерших в течение календарного года к среднегодовой численности населения, в промилле),
- общий коэффициент естественного прироста (разность общих коэффициентов рождаемости и смертности),
- специальный коэффициент рождаемости (количество рождений в среднем на 1000 женщин в возрасте 15—49 лет).

2. При условии, что все возрастные коэффициенты рождаемости положительны, $\alpha_i > 0$, $i = 1, \dots, n$, доказать примитивность матрицы Лесли.

3. При возведении в какую степень матрица Лесли является положительной, если $n = 5$, $\alpha_1 = \alpha_2 = \alpha_3 = 0$, $\alpha_4 > 0$, $\alpha_5 > 0$? Ответ: матрица L^{16} содержит один нулевой элемент (в правом нижнем углу), $L^{17} > 0$.

Комментарии и библиографические примечания

1. Модель Лесли была предложена в работах [1, 2] для изучения биологических популяций (в первой посвященной на эту тему статье в качестве примера Лесли рассматривал данные для лабораторной популяции коричневых крыс), но позднее стала применяться демографами. Более подробно о создании модели см. в [3, 4].

2. Поскольку в настоящей книге модель Лесли излагается для женского населения, то при использовании статистических данных

надо вводить поправочный коэффициент $\delta = 0,488$ — долю девочек среди новорожденных. Например, если для женщины, попавшей в возрастную группу от 30 до 35 лет, вероятность родить ребенка (находясь в этой группе) равна 51,2%, то вероятность рождения девочки равна $0,488 \cdot 0,512 \approx 0,250$, т. е. 25,0%. В отсутствие поведенческого фактора указанное значение доли девочек среди новорожденных изменяется для различных населений весьма слабо (для сорокалетнего периода СССР и России — от 0,485 до 0,489, [5]) и может считаться биологической константой. О поведенческих причинах вариации значения δ см., напр., [6].

3. Приведенное выше рассуждение о сходимости процесса к мальтузианскому режиму является нестрогим. Чтобы обосновать его, необходимо доказать, что λ_1 превышает по модулю все остальные собственные значения матрицы Лесли, является вещественным, простым и положительным числом, а также, что соответствующий ему собственный вектор может быть выбран вещественным и положительным. Основная идея доказательства заключается в следующем. Для матрицы L существует блочное представление

$$L = \begin{pmatrix} A & 0 \\ D & C \end{pmatrix}. \quad (9)$$

Здесь A и C — квадратные матрицы размерностей k и $n - k$ соответственно, где k — номер такой, что $\alpha_k > 0$, и все $\alpha_i = 0$ при $i > k$. При этом матрица A имеет вид, формально подобный виду исходной матрицы Лесли (но при этом A — невырожденная),

$$A = \begin{pmatrix} \alpha_1 & \alpha_2 & \dots & \alpha_{k-1} & \alpha_k \\ s_1 & 0 & \dots & 0 & 0 \\ 0 & s_2 & \dots & 0 & 0 \\ 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & \dots & s_{k-1} & 0 \end{pmatrix}, \quad (10)$$

а у матрицы C не равны нулю лишь поддиагональные элементы. Через «0» в уравнении (9) обозначена нулевая матрица размерности $k \times n - k$, через D — матрица размерности $n - k \times k$, имеющая лишь один ненулевой элемент (правый верхний, он равен s_k).

Можно доказать, что множество ненулевых собственных значений матрицы L совпадает с множеством собственных значений матрицы A , а также, что достаточным условием примитивности матрицы A является положительность еще хотя бы одного (кроме α_k) элемента первой строки. Таким образом, если хотя бы два возрастных коэффициента рождаемости положительны, то матрица A примитивна. Применяя к ней Теорему Фробениуса–Перрона (см. главу 3), получаем, что λ_1 превышает по модулю все остальные собственные значения матрицы Лесли, является вещественным, простым и положительным, а также, что соответствующий ему собственный вектор может быть выбран вещественным и положительным.

Теорема Фробениуса–Перрона позволяет также получить двустороннюю оценку для максимального по модулю собственного значения матрицы Лесли:

$$1 < \lambda_1 < \sum_{i=1}^k l_i \alpha_i, \text{ если } 1 < \sum_{i=1}^k l_i \alpha_i$$

$$\sum_{i=1}^k l_i \alpha_i < \lambda_1 < 1, \text{ если } \sum_{i=1}^k l_i \alpha_i < 1.$$

4. Об определении расширенного, простого и суженного воспроизводства населения см., например, [5, 7].

5. Помимо нетто-коэффициента, агрегированным образом описывающего рождаемость и смертность, в демографии используется такая характеристика рождаемости, как брутто-коэффициент воспроизводства населения — количество девочек, которое родит за всю жизнь женщина при условии, что она доживет до конца репродуктивного периода.

В терминах модели Лесли брутто-коэффициент имеет вид $k_{\text{Brutto}} = \sum_{i=1}^n \alpha_i$. Он тесно связан с суммарным коэффициентом рождаемости (обозначаемым TFR, total fertility rate) — количеством детей обоих полов, которое родит за всю жизнь женщина при условии, что она доживет до конца репродуктивного периода. Его значение равно

$$\text{TFR} = \frac{k_{\text{Brutto}}}{\delta} = \frac{1}{0,488} \sum_{i=1}^n \alpha_i.$$

Здесь число 0,488 представляет собой долю девочек в общем количестве новорожденных.

Брутто-коэффициент воспроизводства применяется при изучении некоторых демографических процессов, однако его значение ничего не говорит о характере воспроизводстве населения: имеет ли оно тенденцию к росту или к уменьшению.

6. О построении и анализе возрастно-половых пирамид см., например, [5].

7. Утверждение о связи типа возрастно-половой пирамиды с характером воспроизводства населения требует аккуратного определения понятий типов воспроизводства населения. Некоторые авторы определяют расширенный и суженный типы воспроизводства исходя из соотношения между текущим количеством рождений и количеством смертей, например: «расширенное воспроизводство населения — воспроизводство населения, при котором число рождающихся детей превышает число умерших людей» [8] или «расширенное воспроизводство характеризуется превышением числа рождений над числом смертей» [9]. С точки зрения подобных, весьма некорректных определений указанные соотношения между типом половозрастной пирамиды и типом воспроизводства не имеют места. Например, в ряде стран (Финляндия, Вьетнам, Куба и т.д.) рождаемость превышает смертность при

пирамиде-урне. Более корректное определение типа воспроизводства, приведенное выше, основано на нетто-коэффициенте воспроизводства, описывающем замещение поколения «матерей» поколением «дочерей».

8. Более подробно о модели Лефковича см. [10].

9. Фрагмент данного раздела, посвященный палеодемографии, написан по материалам работы [11], из которой взяты, в частности, данные табл. 1, а также комментарий о роли факторов, влияющих на смертность.

Литература

1. *Leslie P.H.* The use of matrices in certain population mathematics. *Biometrika*, 1945. 33(3), 183–212.
2. *Leslie P.H.* «Some further notes on the use of matrices in population mathematics». *Biometrika*, 1948. 35(3–4), 213–245.
3. *Caswell H.* *Matrix Population Models: Construction, analysis, and interpretation.* 2nd Edition, 2001.
4. *Kot M.* *Elements of Mathematical Ecology.* Cambridge University Press, 2001.
5. *Медков В.М.* Демография: Учебное пособие. Серия «Учебники и учебные пособия». — Ростов-на-Дону: «Феникс», 2002. 448 с.
6. *Yadava R.C., Kumar A., Srivastava U.* Sex ratio at birth: A model based approach // *Mathematical Social Sciences*, 2013. 65, p. 36–39.
7. *Борисов В.А.* Демография: Учебник для вузов. С. 242–244.
8. *Яценко Н.Е.* Толковый словарь обществоведческих терминов. СПб.: Изд-во «Лань», 1999. 528 с.
9. *Адамчук В.В., Ромашов О.В., Сорокина М.Е.* Экономика и социология труда: Учебник для вузов. М.: ЮНИТИ, 1999. 407 с.
10. *Логофет Д.О., Клочкова И.Н.* Математика модели Лефковича: репродуктивный потенциал и асимптотические циклы // *Матем. моделирование*, 2002, том 14: 10, 116–126.
11. *Алексеева Т.А., Денисова Р.Я., Козловская М.В., Костылева Е.Л., Крайнов Д.А., Лебединская Г.В., Уткин А.В., Федосова В.Н.* Неолит лесной полосы Восточной Европы (Антропология Сахтышских стоянок). М., «Научный мир», 1997. 191 с.

2.4. Учет половой структуры: бипарентальные модели

В предыдущих разделах настоящей главы рассматривались либо «бесструктурные», т.е. оперирующие лишь с общей численностью населения демографические модели, либо модели, учитывающие возрастную структуру, но не принимающие во внимание возможные различия в численности полов. Однако в некоторых исторических ситуациях, например: в периоды кровопролитных или затяжных войн, социальных катастроф, количество мужчин может оказаться заметно меньше или значительно больше числа женщин. В таких ситуациях учет половой

структуры способен значительно обогатить описание и прогнозирование динамики населения.

Причинами неравенства между численностью мужчин и женщин могут быть также социокультурные и поведенческие факторы. Так, в настоящее время в Китае численность мужчин, особенно в молодом возрасте, существенно превосходит численность женщин. По данным 2006 г., в возрастной группе до 15 лет соотношение численностей мужчин и женщин в этой стране было равно 1,13. Для сравнения в России этот показатель в 2010 г. был равен 1,05 (Росстат). Источник этого дисбаланса в состоит том, что в условиях политики «одна семья — один ребенок» перед китайскими семьями со всей остротой вставал выбор: будет ли единственный ребенок мальчиком или девочкой. В соответствии с культурной спецификой Китая мальчик считается, как правило, более желанным, и родители нередко принимают решение о прерывании беременности в случае нежелательного пола ребенка.

Настоящий раздел посвящен классу моделей, учитывающих половую структуру населения (возрастная структура не рассматривается). Рассматриваемая в первом пункте модель связывает интенсивность рождений непосредственно с численностью мужчин и женщин, а модель второго пункта предполагает более сложный механизм, в соответствии с которым количество мужчин и женщин влияет на количество брачных пар, от которого, в свою очередь, зависит рождаемость.

2.4.1. Модель, не учитывающая брачность

В данной модели рождаемость принимается зависящей от численности живущих в рассматриваемой демографической системе женщин (X) и мужчин (Y). Эта зависимость описывается так называемой *бипарентальной функцией* рождаемости $B(X, Y)$, имеющей смысл количества рождений детей обоих полов в единицу времени. По своим свойствам она похожа на неоклассическую производственную функцию, широко применяемую в макроэкономической теории (где X и Y — объемы факторов производства, чаще всего объемы труда и капитала).

Будем называть *бипарентальной функцией рождаемости* определенную при $X \geq 0$, $Y \geq 0$, дважды дифференцируемую функцию $B(X, Y)$, обладающую следующими свойствами.

1. Для рождения детей необходимы родители обоих полов:

$$B(0, Y) = B(X, 0) = 0.$$

2. При фиксированной численности населения одного пола $B(X, Y)$ — неубывающая функция численности населения другого пола:

$$\frac{\partial B}{\partial X} \geq 0, \quad \frac{\partial B}{\partial Y} \geq 0.$$

Например, если в некотором гипотетическом обществе численность населения составляет 10 тыс. мужчин и столько же женщин,

то появление новой женщины может лишь увеличить рождаемость (либо оставить ее без изменений).

3. При постоянной численности одного из полов приращение каждой следующей единицы численности другого пола дает убывающий прирост рождаемости:

$$\frac{\partial^2 B}{\partial X^2} < 0, \quad \frac{\partial^2 B}{\partial Y^2} < 0.$$

Аналогичное свойство производственных функций называется убывающей предельной производительностью факторов производства, в данном случае можно говорить об убывающей предельной фертильности (демографической производительности) каждого из полов.

4. При одновременном увеличении численности мужчин и женщин в некоторое количество раз рождаемость вырастает во столько же раз:

$$B(kX, kY) = kB(X, Y).$$

Данное равенство называется свойством однородности бипарентальной функции рождаемости; его аналог в макроэкономике имеет смысл постоянной отдачи от масштаба производства. В демографии также можно говорить о постоянной отдаче от масштаба: если количество как мужчин, так и женщин увеличится в некоторое количество раз, то и количество рождений увеличится во столько же раз.

5. При постоянной численности одного из полов и неограниченном увеличении численности другого пола рождаемость возрастает неограниченно:

$$\lim_{X \rightarrow +\infty} B(X, Y) = \lim_{Y \rightarrow +\infty} B(X, Y) = +\infty.$$

Это свойство представляет определенную идеализацию. С учетом свойства 3 смысл его можно упрощенно передать утверждением о том, что при фиксированном значении одного аргумента функция $B(X, Y)$ зависит от другого аргумента примерно как степенная функция с показателем, меньшим единицы.

Из перечисленных свойств, составляющих определение, следует еще одно.

6. Поведение функции $B(X, Y)$ в окрестности координатных осей таково, что

$$\begin{aligned} \frac{\partial B}{\partial X} \Big|_{X=0} &= \lim_{X \rightarrow 0} \frac{B(X, Y)}{X} = \lim_{X \rightarrow 0} \frac{XB(1, Y/X)}{X} = \lim_{X \rightarrow 0} B(1, Y/X) = +\infty, \\ \frac{\partial B}{\partial Y} \Big|_{Y=0} &= \lim_{Y \rightarrow 0} \frac{B(X, Y)}{Y} = +\infty. \end{aligned}$$

Все эти свойства используются ниже при анализе модели.

Представление о поведении функции $B(X, Y)$ дает рис. 1. На левом графике изображен график бипарентальной функции рождаемости в сечении $Y = Y_0 = \text{const} > 0$: функция $B(X, Y_0)$ выходит из начала

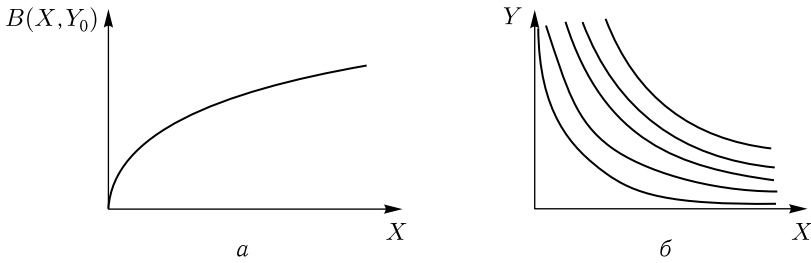


Рис. 1. Бипарентальная функция рождаемости: сечение $Y = Y_0$ (а), линии уровня (б)

координат с бесконечно большой производной и монотонно, неограниченно возрастает, будучи выпуклой вверх. На правом графике представлены линии уровня $B(X, Y) = \text{const}$; из указанных свойств следует, что все они выпуклы вниз.

Примером бипарентальной функции рождаемости служит аналог известной неоклассической производственной функции Кобба–Дугласа $B(X, Y) = AX^\alpha Y^{1-\alpha}$, где $A > 0$, $0 < \alpha < 1$.

В экономической теории с некоторыми оговорками часто применяются также производственная функция Леонтьева и линейная функция ($A_1 > 0$, $A_2 > 0$)

$$B(X, Y) = \min \{A_1 X, A_2 Y\}, \quad B(X, Y) = A_1 X + A_2 Y,$$

удовлетворяющие не всем перечисленным аксиомам. Их аналоги также могут использоваться в качестве бипарентальных функций рождаемости.

Перейдем теперь к построению модели, описывающей динамику численности мужского и женского населения.

Обозначим через $\delta \in (0; 1)$ долю девочек среди новорожденных, а через $\beta_1 > 0$ и $\beta_2 > 0$ — коэффициенты смертности (долю умерших в единицу времени) женщин и мужчин соответственно. Будем предполагать их постоянными. Численность девочек, рожденных в течение интервала времени $(t + \Delta t)$, равна $\delta B(X, Y)\Delta t$, а численность умерших женщин — $\beta_1 X \Delta t$. Следовательно, изменение численности женщин за интервал времени Δt составит

$$\Delta X = X(t + \Delta t) - X(t) = (\delta B(X, Y) - \beta_1 X)\Delta t.$$

Аналогично, изменение численности мужчин равно

$$\Delta Y = Y(t + \Delta t) - Y(t) = ((1 - \delta)B(X, Y) - \beta_2 Y)\Delta t.$$

Проведем, как обычно, в двух предыдущих выражениях предельный переход $\Delta t \rightarrow 0$, тогда получим систему дифференциальных уравнений

$$\begin{aligned} \frac{dX}{dt} &= \delta B(X, Y) - \beta_1 X, \\ \frac{dY}{dt} &= (1 - \delta)B(X, Y) - \beta_2 Y. \end{aligned} \quad (1)$$

Уравнения (1) выражают баланс между рождаемостью и смертностью женщин и мужчин и описывают временную динамику численностей женского и мужского полов.

Все рассмотренные выше в настоящей главе линейные модели имели мальтузианские, т. е. экспоненциальные по времени, решения. Ввиду постоянства коэффициентов и однородности бипарентальной функции рождаемости следует ожидать, что система (1) также имеет мальтузианский режим, реализуемый при определенных начальных условиях. Будем искать его в виде

$$X(t) = X_m e^{rt}, \quad Y(t) = Y_m e^{rt}, \quad (2)$$

где r , X_m , Y_m — постоянные, подлежащие определению в дальнейшем. Подставим (2) в (1), получим

$$(r + \beta_1)X_m = \delta B(X_m, Y_m), \quad (3)$$

$$(r + \beta_2)Y_m = (1 - \delta)B(X_m, Y_m). \quad (4)$$

Система (3), (4) имеет тривиальное решение $X_m = Y_m = 0$ при любом значении r . Докажем, что при некотором (единственном) значении r она имеет также и нетривиальное решение. Из свойства однородности бипарентальной функции рождаемости следует равенство $B(X_m, Y_m) = X_m B(1, Y_m/X_m)$. Обозначим отношение численностей мужчин и женщин Y_m/X_m через S_m , тогда система (3), (4) принимает безразмерный вид

$$r + \beta_1 = \delta B(1, S_m), \quad (5)$$

$$(r + \beta_2)S_m = (1 - \delta)B(1, S_m). \quad (6)$$

Исключив из (5), (6) переменную r , получаем лишь одно уравнение относительно безразмерной величины S_m :

$$\delta B(1, S_m) + \beta_2 - \beta_1 = \frac{(1 - \delta)B(1, S_m)}{S_m}. \quad (7)$$

Докажем существование и единственность его положительного решения, которое обозначим S_{m0} . Введем обозначение $b(S_m) = B(1, S_m)$ и перепишем уравнение (7) в виде

$$\delta S_m b(S_m) + S_m(\beta_2 - \beta_1) = (1 - \delta)b(S_m).$$

Рассмотрим функцию, равную разности левой и правой частей данного равенства:

$$f(S_m) = \delta S_m b(S_m) + S_m(\beta_2 - \beta_1) - (1 - \delta)b(S_m). \quad (8)$$

Ее корни совпадают в области $S_m > 0$ с корнями уравнения (7). Докажем, что эта функция обращается в ноль ровно один раз на полуоси $S_m \in (0, +\infty)$. Для этого заметим, что $f(0) = 0$, первая производная $f'(S_m)$ отрицательна в некоторой правой полуокрестности точки $S_m = 0$ и $f((1 - \delta)/\delta) = (\beta_2 - \beta_1)(1 - \delta)/\delta$.

Если $\beta_2 > \beta_1$, то $f((1-\delta)/\delta) > 0$ (см. рис. 2, а), то функция $f(S_m)$ меняет знак на интервале $0 < S_m < (1-\delta)/\delta$. Единственность корня следует из того, что вторая производная функции (8), равная

$$f''(S_m) = b''(S_m) [\delta S_m - (1-\delta)] + 2\delta b'(S_m),$$

положительна при $0 \leq S_m \leq (1-\delta)/\delta$. В то же время при $S_m > (1-\delta)/\delta$ функция $f(S_m)$ корней не имеет, поскольку в этом интервале ее первая производная

$$f'(S_m) = b'(S_m) [\delta S_m - (1-\delta)] + \delta b(S_m) + \beta_2 - \beta_1$$

положительна.

Аналогично рассматривается случай $\beta_2 < \beta_1$ (см. рис. 2, б). Здесь $f((1-\delta)/\delta) < 0$, и функция $f(S_m)$ знакопостоянна (отрицательна) на интервале $0 < S_m < (1-\delta)/\delta$. Отсюда следует, что корень уравнения (8) лежит правее точки $S_m = (1-\delta)/\delta$.

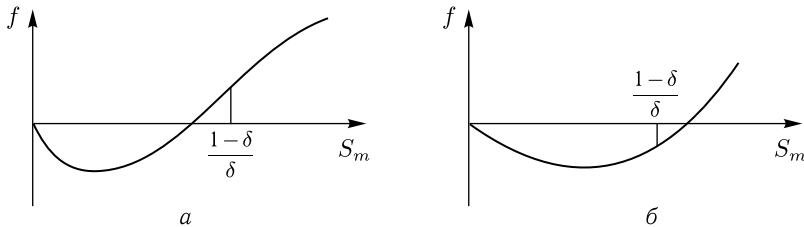


Рис. 2. Геометрическая иллюстрация доказательства существования и единственности решения уравнения (7): случай $\beta_2 > \beta_1$ (а), случай $\beta_2 < \beta_1$ (б)

Из приведенного рассмотрения вытекают следующие выводы.

1) Если $\beta_2 > \beta_1$, то единственный положительный корень уравнения (8) находится в интервале $(0; (1-\delta)/\delta)$. Полученный результат имеет ясную содержательную трактовку: если мужская смертность превышает женскую, то соотношение между численностями мужчин и женщин в мальтузианском режиме ниже, чем соотношение между численностями рождений мальчиков и девочек:

$$\frac{Y_m}{X_m} < \frac{1-\delta}{\delta}.$$

2) Если мужская смертность равна женской, $\beta_2 = \beta_1$, то единственный положительный корень уравнения (8) — величина $S_m = (1-\delta)/\delta$, т.е. в мальтузианском режиме поддерживается то же соотношение между численностями мужчин и женщин, что и между новорожденными мальчиками и девочками.

3) Если мужская смертность ниже женской, $\beta_2 < \beta_1$, то в мальтузианском режиме доля мужчин в общем населении выше, чем доля мальчиков при рождении:

$$\frac{Y_m}{X_m} > \frac{1-\delta}{\delta}.$$

После того как определено значение S_m , из формулы (5) или (6) вычисляется величина r , определяющая характер временной динамики численности населения:

$$r = \delta B(1, S_m) - \beta_1.$$

Если $r > 0$, то имеет место рост численности обоих полов, если $r < 0$, то — ее убывание.

Данные выводы касаются лишь мальтузианского (экспоненциального) режима, т.е. решения системы уравнений (1), имеющего вид (2). Он реализуется непосредственно с момента времени $t = 0$ лишь в случае, когда начальное соотношение численностей полов в точности равно соотношению полов в мальтузианском режиме, $X(0)/Y(0) = X_m/Y_m$. Если же данное равенство не выполняется, то мальтузианский режим является аттрактором процесса в том смысле, что с течением времени отношение $X(t)/Y(t)$ приближается к X_m/Y_m , а скорость роста численности каждого из полов приближается к значению r .

Проиллюстрируем этот процесс, анализируя фазовую плоскость системы уравнений (1) для частного случая, когда бипарентальная функция рождаемости задана формулой

$$B(X, Y) = AX^\alpha Y^{1-\alpha}, \quad (9)$$

где $A > 0$, $0 < \alpha < 1$.

В этом случае задача (1) принимает вид

$$\begin{aligned} \frac{dX}{dt} &= \delta AX^\alpha Y^{1-\alpha} - \beta_1 X, \\ \frac{dY}{dt} &= (1 - \delta)AX^\alpha Y^{1-\alpha} - \beta_2 Y, \end{aligned} \quad (10)$$

а соотношения (3), (4) запишутся в форме

$$\begin{aligned} (r + \beta_1)X_m &= \delta AX_m^\alpha Y_m^{1-\alpha}, \\ (r + \beta_2)Y_m &= (1 - \delta)AX_m^\alpha Y_m^{1-\alpha}. \end{aligned}$$

Первое из этих уравнений дает

$$Y_m = \left[\frac{r + \beta_1}{\delta A} \right]^{\frac{1}{1-\alpha}} X_m, \quad (11)$$

а второе —

$$Y_m = \left[\frac{(1 - \delta)A}{r + \beta_2} \right]^{\frac{1}{\alpha}} X_m. \quad (12)$$

Из (11), (12) получаем уравнение для определения скорости роста r в случае, когда бипарентальная функция рождаемости задана формулой (9):

$$\left[\frac{r + \beta_1}{\delta A} \right]^\alpha = \left[\frac{(1 - \delta)A}{r + \beta_2} \right]^{1-\alpha}. \quad (13)$$

В соответствии с изложенным выше анализом мальтузианского решения (2) задачи (1) с бипарентальной функцией рождаемости общего вида уравнение (13) имеет единственное решение r , которое может быть как положительным, так и отрицательным.

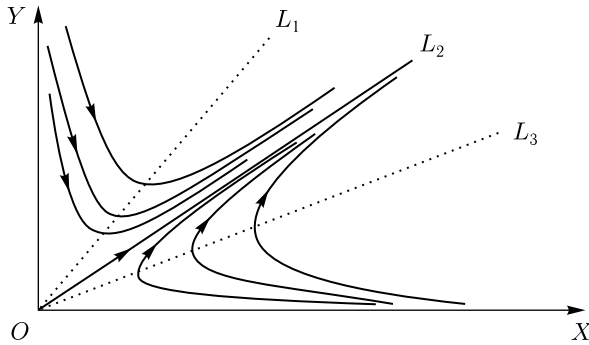


Рис. 3. Фазовый портрет системы (10) в случае $r > 0$

Фазовый портрет системы (1) для случая $r > 0$ представлен на рис. 3. Здесь L_1, L_2, L_3 — прямые, заданные уравнениями

$$\begin{aligned} L_1: \quad Y &= \left[\frac{(1-\delta)A}{\beta_2} \right]^{\frac{1}{\alpha}} X, \\ L_2: \quad Y &= \left[\frac{r+\beta_1}{\delta A} \right]^{\alpha} X = \left[\frac{(1-\delta)A}{r+\beta_2} \right]^{1-\alpha} X, \\ L_3: \quad Y &= \left[\frac{\beta_1}{\delta A} \right]^{\frac{1}{1-\alpha}} X. \end{aligned}$$

Прямые L_1, L_3 состоят из фазовых точек, в которых $dY/dt = 0$ и $dX/dt = 0$ соответственно.

Движение вдоль фазовой траектории L_2 , очевидно, соответствует мальтузианскому режиму, и соотношение полов в нем описывается выражениями (11), (12). Все прочие фазовые кривые описывают асимптотический (при $t \rightarrow +\infty$) выход на мальтузианский режим. При этом, если начальное условие $X(0), Y(0)$ находится в области между прямыми L_1, L_3 , то, начиная с момента времени $t = 0$ возрастает численность как женщин, так и мужчин, поскольку в этой области имеют место неравенства $dX/dt > 0, dY/dt > 0$.

Рассмотрим теперь случай, когда начальная точка $X(0), Y(0)$ находится в области между осью OY и прямой L_1 , в которой имеют место неравенства $dX/dt > 0, dY/dt < 0$. Демографический смысл этих неравенств состоит в том, что если количество мужчин значительно превышает количество женщин, то на первом этапе эволюции системы численность мужчин будет уменьшаться. После того, как соотношение между $X(t)$ и $Y(t)$ достигнет значения, соответствующего прямой L_1 ,

наступает второй этап, на котором приближение к мальтузианскому режиму происходит с одновременным увеличением численности как женщин, так и мужчин.

Аналогичным образом, если начальная точка $X(0), Y(0)$ находится в области между прямой L_3 и осью OX , то на первом этапе происходит увеличение количества мужчин и уменьшение количества женщин. На втором этапе возрастает численность обоих полов.

Во всех трех случаях все фазовые траектории с течением времени неограниченно приближаются к прямой L_2 . Тем самым мальтузианский режим устанавливается при любых начальных значениях численности женского и мужского населения $X(0) > 0, Y(0) > 0$.

Аналогично рассмотрим фазовый портрет системы (10) в случае, когда решение уравнения (13) отрицательно ($r < 0$), а ее аттрактором является мальтузианский режим с убывающим населением (рис. 4). По-прежнему прямая L_1 состоит из фазовых точек, в которых $dY/dt = 0$, а L_3 — из точек, в которых $dX/dt = 0$. Уравнения этих прямых также остаются неизменными, однако меняется их расположение на фазовой плоскости: теперь прямая L_3 лежит выше, чем L_2 , а L_1 — ниже, чем L_2 . Все фазовые траектории с течением времени неограниченно приближаются к прямой L_2 . Тем самым устанавливается режим экспоненциального убывания численности населения обоих полов.

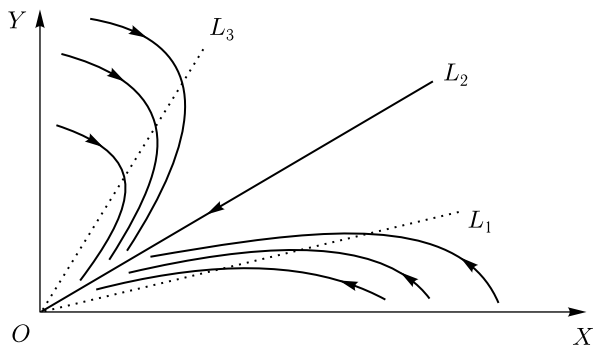


Рис. 4. Фазовый портрет системы (10) в случае $r < 0$

Проведенный анализ показывает, что для демографической системы, описываемой уравнениями (1), мальтузианский режим играет особую роль: он является аттрактором процесса. Каким бы ни было начальное соотношение количества женщин и мужчин, с течением времени устанавливается мальтузианское соотношение между полами и соответствующая скорость роста (положительная, отрицательная или нулевая).

2.4.2. Модель, учитывающая брачность

Данная модель обеспечивает более тонкий учет социально-демографического статуса индивидов: в нее введен постулат о том, что

количество рождений зависит не от численности всех мужчин и женщин, а от количества брачных пар (при этом брак понимается как союз одного мужчины и одной женщины, не состоящих в других браках).

Обозначим количество таких пар в момент времени t через $Z(t)$. Пусть переменные $X(t)$ и $Y(t)$ по-прежнему представляют численность соответственно женского и мужского населения, тогда количество одиноких женщин и мужчин равно $N_1(t) = X(t) - Z(t)$ и $N_2(t) = Y(t) - Z(t)$ соответственно.

Для описания процесса образования пар введем функцию $M(N_1, N_2)$, имеющую смысл скорости образования новых браков, так что $M(N_1, N_2) dt$ есть количество новых браков за время dt . По определению будем называть функцией брачности любую определенную при $N_1 \geq 0, N_2 \geq 0$ непрерывно дифференцируемую функцию, обладающую следующими свойствами.

1. Для образования новых брачных пар необходимо наличие одиноких индивидов обоих полов, т.е.

$$M(0, N_2) = M(N_1, 0) = 0.$$

2. При фиксированной численности одиноких индивидов одного пола $M(N_1, N_2)$ является неубывающей функцией численности одиноких индивидов другого пола:

$$\frac{\partial M}{\partial N_1} \geq 0, \quad \frac{\partial M}{\partial N_2} \geq 0.$$

3. При одновременном увеличении численности одиноких мужчин и женщин в некоторое количество раз количество новообразованных пар вырастает во столько же раз (свойство однородности):

$$M(kN_1, kN_2) = kM(N_1, N_2).$$

Свойства 1–3 являются аналогами свойств 1, 2, 4 бипарентальной функции рождаемости, рассмотренной в первом пункте настоящего раздела.

4. При неограниченном увеличении среднего количества незамужних женщин, приходящегося на одного неженатого мужчину, отношение скорости образования новых браков, заключенных за малое время dt , к количеству мужчин возрастает неограниченно:

$$\lim_{N_1/N_2 \rightarrow +\infty} \frac{M(N_1, N_2)}{N_2} = +\infty.$$

Аналогичное свойство имеет место при неограниченном увеличении отношения количества незамужних мужчин к количеству незамужних женщин:

$$\lim_{N_2/N_1 \rightarrow +\infty} \frac{M(N_1, N_2)}{N_1} = +\infty.$$

С учетом свойства однородности 3 функции брачности свойство 4 можно переписать в форме

$$\lim_{x \rightarrow +\infty} M(x, 1) = +\infty, \quad \lim_{y \rightarrow +\infty} M(1, y) = +\infty,$$

облегчающей его трактовку в виде поведенческой гипотезы.

А именно: *если для вступления в брак мужчина ищет партнершу, удовлетворяющую определенным требованиям, то чем более широким выбором потенциальных невест он обладает, тем быстрее он находит себе пару (аналогично будет для женщины, которая ищет себе пару).*

Более того, условие 4 предполагает, что формально эта скорость может быть сколь угодно большой (впрочем, это требование может быть ослаблено, см. комментарий 5 ниже). Конечно, возможны и альтернативные поведенческие гипотезы, учитывающие, например, что мужчине доступен для «просмотра» лишь ограниченный круг кандидатов в невесты и т. д.

Перейдем теперь к получению уравнений бипарентальной модели, учитывающей брачность.

Семья (брачная пара) образовывается одинокими женщиной и мужчиной; количество таких пар, созданных за интервал времени $(t, t + \Delta t)$, равно, по смыслу функции брачности, величине $M(N_1, N_2)\Delta t$. Введем в рассмотрение следующие демографические параметры: μ_1 — коэффициент смертности женщин, состоящих в браке (так что $\mu_1 dt$ есть доля замужних женщин, умирающих за время dt), μ_2 — коэффициент смертности мужчин, состоящих в браке, μ_3 и μ_4 — коэффициенты разводимости и одновременной смертности обоих супругов. Таким образом, при количестве семей $Z(t)$ число овдовевших мужчин и женщин за тот же интервал времени составит соответственно $\mu_1 Z \Delta t$ и $\mu_2 Z \Delta t$, количество разводов — $\mu_3 Z \Delta t$, а количество случаев одновременной смерти супругов — $\mu_4 Z \Delta t$.

Следовательно, динамика численности брачных пар описывается уравнением

$$\frac{dZ}{dt} = M(N_1, N_2) - \mu_1 Z - \mu_2 Z - \mu_3 Z - \mu_4 Z. \quad (14)$$

Введем в рассмотрение коэффициенты смертности одиноких женщин и мужчин, σ_1 и σ_2 , соответственно. Заметим, что, вообще говоря, $\sigma_1 \neq \mu_1$, $\sigma_2 \neq \mu_2$, т. е. данная модель позволяет учитывать различия в уровне смертности одиноких и семейных индивидов.

Чтобы выписать уравнения для искомой численности одиноких женщин $N_1(t)$ и мужчин $N_2(t)$, введем также следующие величины. Обозначим через α коэффициент фертильности (так что $\alpha \Delta t$ — среднее количество детей, рожденных одной парой за время Δt), а через δ — долю девочек при рождении. Тогда численность девочек, рожденных за интервал времени $(t, t + \Delta t)$, составит $\delta \alpha Z \Delta t$. Количество женщин, ставших одинокими в результате овдовения, равно $\mu_2 Z \Delta t$, в результате

развода — $\mu_3 Z \Delta t$. При численности одиноких женщин, равной N_1 , их смертность за интервал времени $(t, t + \Delta t)$ составит $\sigma_1 N_1 \Delta t$. Наконец, количество женщин, вступивших в брак, равно $M(N_1, N_2) \Delta t$.

Таким образом, динамика численности одиноких женщин описывается «балансным» уравнением

$$\frac{dN_1}{dt} = \delta \alpha Z + \mu_2 Z + \mu_3 Z - \sigma_1 N_1 - M(N_1, N_2). \quad (15)$$

Аналогично, уравнение для численности одиноких мужчин имеет вид

$$\frac{dN_2}{dt} = (1 - \delta) \alpha Z + \mu_1 Z + \mu_3 Z - \sigma_2 N_2 - M(N_1, N_2). \quad (16)$$

Приводя подобные слагаемые в (14)–(16) и обозначая

$$\begin{aligned} \sigma_3 &= \mu_1 + \mu_2 + \mu_3 + \mu_4, \\ k_1 &= \delta \alpha + \mu_2 + \mu_3, \quad k_2 = (1 - \delta) \alpha + \mu_1 + \mu_3, \end{aligned}$$

получим из (14)–(16) систему трех обыкновенных дифференциальных уравнений для трех неизвестных функций $Z(t)$, $N_1(t)$, $N_2(t)$:

$$\frac{dZ}{dt} = -\sigma_3 Z + M(N_1, N_2), \quad (17)$$

$$\frac{dN_1}{dt} = -\sigma_1 N_1 + k_1 Z - M(N_1, N_2), \quad (18)$$

$$\frac{dN_2}{dt} = -\sigma_2 N_2 + k_2 Z - M(N_1, N_2). \quad (19)$$

Каждое из уравнений (17)–(19) содержит нелинейность в правой части, что создает определенные трудности для анализа. Избавляясь от них, перейдем от переменных N_1 , N_2 (количества одиноких женщин и мужчин) к переменным $X = N_1 + Z$, $Y = N_2 + Z$ (численности всех женщин и всех мужчин), складывая соответственно уравнения (17) и (18), затем (17) и (19). Тогда для общего количества женщин, общего количества мужчин и количества семейных пар имеем более простую систему

$$\frac{dX}{dt} = -\sigma_1 X + (\sigma_1 - \sigma_3 + k_1) Z, \quad (20)$$

$$\frac{dY}{dt} = -\sigma_2 Y + (\sigma_2 - \sigma_3 + k_2) Z, \quad (21)$$

$$\frac{dZ}{dt} = -\sigma_3 Z + M(X - Z, Y - Z). \quad (22)$$

Уравнения (20)–(22), дополненные начальными условиями

$$X(0) = X^0, \quad Y(0) = Y^0, \quad Z(0) = Z^0, \quad (23)$$

представляют собой замкнутую модель демографического процесса.

Подобно тому, как было сделано для модели из предыдущего пункта, ввиду постоянства коэффициентов и однородности функции

брачности будем искать для системы (20)–(22) мальтузианское (экспоненциальное) решение, имеющее вид

$$X(t) = X_m e^{rt}, \quad Y(t) = Y_m e^{rt}, \quad Z(t) = Z_m e^{rt}. \quad (24)$$

Поскольку общее количество мужчин не может быть меньше количества женатых мужчин (то же справедливо для женщин), то решение вида (24) должно удовлетворять требованиям

$$X_m \geq Z_m \geq 0, \quad Y_m \geq Z_m. \quad (25)$$

Подставим (24) в (20)–(22), тогда после сокращения обеих частей равенства на множитель e^{rt} получим

$$rX_m = -\sigma_1 X_m + (\sigma_1 - \sigma_3 + k_1)Z_m, \quad (26)$$

$$rY_m = -\sigma_2 Y_m + (\sigma_2 - \sigma_3 + k_2)Z_m, \quad (27)$$

$$rZ_m = -\sigma_3 Z_m + M(X_m - Z_m, Y_m - Z_m). \quad (28)$$

При любом вещественном r эта система имеет тривиальное (нулевое) решение $X_m = Y_m = Z_m = 0$. Чтобы (как в первом пункте) найти нетривиальный мальтузианский режим, необходимо определить значения r , при которых имеется также и ненулевое решение, удовлетворяющее неравенствам (25).

Уравнения (26), (27) позволяют явно выразить величины X_m и Y_m через Z_m :

$$X_m = \frac{\sigma_1 - \sigma_3 + k_1}{r + \sigma_1} Z_m, \quad (29)$$

$$Y_m = \frac{\sigma_2 - \sigma_3 + k_2}{r + \sigma_2} Z_m \quad (30)$$

(здесь мы предполагаем, что $r \neq \sigma_1$, $r \neq \sigma_2$, анализ противоположного случая вынесен в упражнения). Подставив (30) в (28) и воспользовавшись свойством однородности 3 функции брачности, получим уравнение для r — скорости роста численности населения в мальтузианском режиме:

$$r = -\sigma_3 + M\left(\frac{k_1 - \sigma_3 - r}{r + \sigma_1}; \frac{k_2 - \sigma_3 - r}{r + \sigma_2}\right). \quad (31)$$

Его решение ищется в области, определяемой условиями (25), или (что то же самое) при естественном условии неотрицательности аргументов функции брачности:

$$\frac{k_1 - \sigma_3 - r}{r + \sigma_1} \geq 0; \quad \frac{k_2 - \sigma_3 - r}{r + \sigma_2} \geq 0.$$

Каждое из этих неравенств может удовлетворяться либо при положительном знаменателе и неотрицательном числителе, либо при отрицательном знаменателе и неположительном числителе. Таким образом, для полного анализа уравнения (31) необходимо рассмотреть различные комбинации соотношений между параметрами k_1 , k_2 , σ_1 , σ_2 , σ_3 .

Ограничимся здесь разбором лишь одного, наиболее простого случая. Пусть каждая из разностей $\sigma_3 - \sigma_1$, $\sigma_3 - \sigma_2$ меньше каждого из параметров k_1 , k_2 . Тогда решение уравнения (31) ищется в области

$$-\min(\sigma_1, \sigma_2) < r \leq \min(k_1, k_2) - \sigma_3. \quad (32)$$

Чтобы изучить вопрос о его существовании и единственности в области (32), рассмотрим функцию $\varphi(r)$, равную разности левой и правой частей уравнения (31),

$$\varphi(r) = r + \sigma_3 - M\left(\frac{k_1 - \sigma_3 - r}{r + \sigma_1}; \frac{k_2 - \sigma_3 - r}{r + \sigma_2}\right).$$

Корни функции $\varphi(r)$ совпадают с корнями уравнения (31). Докажем, что она меняет знак ровно один раз в полуинтервале (32). На его правом конце функция $\varphi(r)$ положительна: например, если $k_1 < k_2$, то $\varphi(\min(k_1, k_2) - \sigma_3) = \varphi(k_1 - \sigma_3) = k_1$. При приближении к левому концу, т. е. при $r \rightarrow -\min(\sigma_1, \sigma_2)$, имеем $\varphi(r) \rightarrow -\infty$ в силу свойства 4 функции брачности. Следовательно, уравнение (31) в полуинтервале (32) имеет как минимум один корень (см. рис. 5).

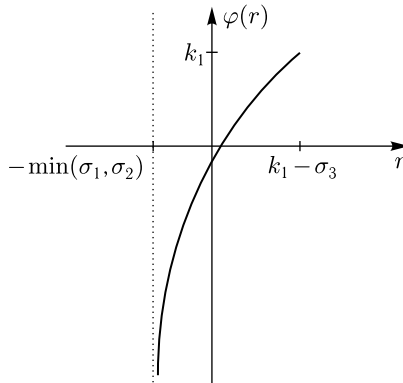


Рис. 5. Геометрическая иллюстрация доказательства существования и единственности корня функции $\varphi(r)$

Докажем его единственность, рассмотрев производную функции $\varphi(r)$,

$$\varphi'(r) = 1 - \frac{\partial M}{\partial N_1} \frac{\sigma_3 - \sigma_1 - k_1}{(r + \sigma_1)^2} - \frac{\partial M}{\partial N_2} \frac{\sigma_3 - \sigma_2 - k_2}{(r + \sigma_2)^2}.$$

Здесь производные $\partial M/\partial N_1$, $\partial M/\partial N_2$ вычисляются в точке $\left(\frac{k_1 - \sigma_3 - r}{r + \sigma_1}; \frac{k_2 - \sigma_3 - r}{r + \sigma_2}\right)$; они неотрицательны ввиду свойства 2. Множители $\partial M/\partial N_1$, $\partial M/\partial N_2$ при этих производных отрицательны при рассматриваемых значениях параметров (см. (32)). Следовательно, $\varphi'(r) > 0$.

Таким образом, существование и единственность решения $r = r_m$ уравнения (31) в полуинтервале (32) доказаны в случае, если каждая из разностей $\sigma_3 - \sigma_1$, $\sigma_3 - \sigma_2$ меньше каждого из параметров k_1 , k_2 .

После того, как определено значение r_m , с помощью формул (30) нетрудно определить всю динамику численности компонент демографической системы (всех женщин, всех мужчин и семейных пар) в мальтузианском режиме. В частности, из установленных свойств функции $\varphi(r)$ следует, что если $\varphi(0) < 0$, то $r_m > 0$ (демографический рост), и если $\varphi(0) > 0$, то $r_m < 0$ (демографический спад).

Итак, в области параметров, задаваемой неравенством $\max(\sigma_3 - \sigma_1, \sigma_3 - \sigma_2) < \min(k_1, k_2)$, у модели (20)–(22) существует мальтузианский режим, при котором численность каждой из компонент населения экспоненциально возрастает, если величина

$$\varphi(0) = \sigma_3 - M \left(\frac{k_1 - \sigma_3}{\sigma_1}; \frac{k_2 - \sigma_3}{\sigma_2} \right)$$

отрицательна, и убывает, если она положительна.

Уравнение (31) при других комбинациях параметров k_1 , k_2 , σ_1 , σ_2 , σ_3 также поддается исследованию, хотя и более громоздкому.

Вопрос об устойчивости мальтузианского режима рассмотрим для частного случая, когда функция брачности имеет вид $M(N_1, N_2) = AN_1^a N_2^{1-a}$, коэффициенты смертности мужчин и женщин равны $\mu_2 = \mu_1$, $\sigma_2 = \sigma_1$, а доля девочек при рождении составляет 50%, т.е. $\delta = 0,5$. Кроме того, как и в приведенном выше доказательстве существования мальтузианского режима, ограничимся случаем, когда параметры модели удовлетворяют неравенству $\sigma_3 - \sigma_1 < k_1$.

Система (20)–(22) при указанных предположениях принимает вид

$$\frac{dX}{dt} = -\sigma_1 X + (\sigma_1 - \sigma_3 + k_1)Z, \quad (33)$$

$$\frac{dY}{dt} = -\sigma_1 Y + (\sigma_1 - \sigma_3 + k_1)Z, \quad (34)$$

$$\frac{dZ}{dt} = -\sigma_3 Z + A(X - Z)^a (Y - Z)^{1-a}. \quad (35)$$

Введем переменные $x(t)$, $y(t)$, имеющие смысл количества соответственно незамужних женщин и неженатых мужчин, приходящихся на одну семью:

$$x(t) = \frac{X(t)}{Z(t)}, \quad y(t) = \frac{Y(t)}{Z(t)}.$$

После несложных, но громоздких преобразований получим из (33)–(35) уравнения для функций $x(t)$, $y(t)$:

$$\frac{dx}{dt} = [\sigma_3 - \sigma_1 - A(x - 1)^a (y - 1)^{1-a}] x + (\sigma_1 - \sigma_3 + k_1), \quad (36)$$

$$\frac{dy}{dt} = [\sigma_3 - \sigma_1 - A(x - 1)^a (y - 1)^{1-a}] y + (\sigma_1 - \sigma_3 + k_1). \quad (37)$$

Нетрудно видеть, что задача (36), (37) имеет ровно одно положительное стационарное решение

$$\begin{aligned} x_s = y_s = \\ = s \equiv \frac{1}{2A} \left[A + \sigma_3 - \sigma_1 + \sqrt{(A + \sigma_3 - \sigma_1)^2 + 4A(\sigma_1 - \sigma_3 + k_1)} \right]. \end{aligned} \quad (38)$$

Наша ближайшая цель заключается в том, чтобы показать асимптотическую устойчивость этого стационарного решения системы (36), (37). Для этого введем в рассмотрение функции

$$\tilde{x}(t) = x(t) - s, \quad \tilde{y}(t) = y(t) - s, \quad (39)$$

характеризующие отклонения переменных $x(t)$, $y(t)$ от стационарных значений. Уравнения для этих отклонений имеют вид

$$\frac{d\tilde{x}}{dt} = [\sigma_3 - \sigma_1 - A(\tilde{x} + s - 1)^a (\tilde{y} + s - 1)^{1-a}] (\tilde{x} + s) + (\sigma_1 - \sigma_3 + k_1), \quad (40)$$

$$\frac{d\tilde{y}}{dt} = [\sigma_3 - \sigma_1 - A(\tilde{x} + s - 1)^a (\tilde{y} + s - 1)^{1-a}] (\tilde{y} + s) + (\sigma_1 - \sigma_3 + k_1). \quad (41)$$

Линеаризовав уравнения (40), (41) в окрестности точки $\tilde{x} = \tilde{y} = 0$, после преобразований получим

$$\frac{d\tilde{x}}{dt} = [\sigma_3 - \sigma_1 - A(s - 1) - Aas] \tilde{x} - A(1 - a)s \tilde{y}, \quad (42)$$

$$\frac{d\tilde{y}}{dt} = [\sigma_3 - \sigma_1 - A(s - 1) - A(1 - a)s] \tilde{y} - Aas \tilde{x}. \quad (43)$$

Нетрудно выписать матрицу линеаризованной системы (42), (43):

$$\begin{pmatrix} \sigma_3 - \sigma_1 - A(s - 1) - Aas & -A(1 - a)s \\ -Aas & \sigma_3 - \sigma_1 - A(s - 1) - A(1 - a)s \end{pmatrix}.$$

Ее характеристическое уравнение имеет вид

$$\begin{aligned} \lambda^2 - \lambda[2(\sigma_3 - \sigma_1) - 2A(s - 1) - As] + \\ + [(\sigma_3 - \sigma_1 - A(s - 1))^2 - As(\sigma_3 - \sigma_1 - A(s - 1))] = 0. \end{aligned} \quad (44)$$

Решая квадратное (44) и учитывая, что значение величины s дается формулой (38), получим

$$\begin{aligned} \lambda_1 = \frac{1}{2} \left[\sigma_3 - \sigma_1 + A - \sqrt{(\sigma_3 - \sigma_1 + A)^2 + 4A(k_1 - \sigma_3 - \sigma_1)} \right], \\ \lambda_2 = \lambda_1 - As. \end{aligned}$$

Следовательно, $\lambda_1 < 0$, $\lambda_2 < 0$ (напомним, что мы здесь рассматриваем случай, когда параметры системы находятся в соотношении $\sigma_3 - \sigma_1 < k_1$) и стационарное решение (38) системы (36), (37) является

асимптотически устойчивым. В терминах исходной задачи (33)–(35) это означает что

$$\frac{X(t)}{Z(t)} \rightarrow x_s, \quad \frac{Y(t)}{Z(t)} \rightarrow y_s \quad \text{при } t \rightarrow \infty,$$

т. е. с течением времени устанавливается постоянное соотношение между численностями компонент населения: незамужних женщин, незамужних мужчин и семейных пар, а именно — мальтузианский режим.

Таким образом, для бипарентальной модели, учитывающей брачность, мальтузианский режим также является аттрактором демографического процесса: с течением времени устанавливаются мальтузианское соотношение между полами и соответствующая скорость роста (положительная, отрицательная или нулевая).

Упражнения

1. Для модели, не учитывающей брачность, определить соотношение полов, наиболее благоприятное для демографического роста, проанализировав величину $\frac{1}{X+Y} \frac{d(X+Y)}{dt}$.

2. Для модели, не учитывающей брачность, найти значение скорости роста r в случае бипарентальной функции рождаемости, имеющей вид $B(X, Y) = A\sqrt{XY}$.

$$\text{Ответ: } r = \frac{-(\beta_1 + \beta_2) + \sqrt{(\beta_1 - \beta_2)^2 + 4\delta(1 - \delta)A^2}}{2}.$$

3. Для модели, учитывающей брачность, рассмотреть случай $r = \sigma_1$ (см. формулу (29)).

Комментарии и библиографические примечания

1. Модель, не учитывающей брачность, была предложена в [1]. Более детальный анализ случая с бипарентальной функцией вида $B(X, Y) = AX^\alpha Y^{1-\alpha}$ проведен авторами.

2. Модель, учитывающей брачность, предложена в [2], доказательство асимптотической устойчивости мальтузианского режима проведено авторами.

3. О влиянии социокультурных факторов на соотношение полов при рождении см. [3].

4. О моделях, учитывающих как возрастную, так и половую структуру, см., напр., [4].

5. Как видно из второго пункта, свойство 4 функции брачности используется при доказательстве существования и единственности решения уравнения $\varphi(r) = 0$ на полуинтервале (32). Отсюда следует возможность ослабления условия, содержащегося в свойстве 4: достаточно потребовать отрицательности предела, $\lim_{r \rightarrow -\min(\sigma_1, \sigma_2)} \varphi(r) < 0$.

К сожалению, формулировка этого требования в терминах исходной задачи является весьма громоздкой.

Литература

1. *Samuelson P.A.*. Generalizing Fisher's "reproductive value": Nonlinear, homogeneous, biparental systems // Proc. Nat. Acad. Sci. USA. V.74, №12. P.5772–5775, December 1977.
2. *Yellin J., Samuelson P.A.* A Dynamical Model for Human Population // Proc. Nat. Acad. Sci. USA. V. 71, № 7. P.2813–2817, July 1974.
3. *Yadava R.C., Kumar A., Srivastava U.* Sex ratio at birth: A model based approach // Mathematical Social Sciences, 2013. 65. P.36–39.
4. *Ianelli M., Martcheva M., Milner F.A.* Gender-Structured Population Modeling / Mathematical Methods, Numerics, and Simulation. country-SIAM. 2005.

2.5. Пример имитационно-математического моделирования в демографии

Рассмотренные в предыдущих разделах настоящей главы математические модели позволяют описать демографическую динамику в общих чертах, ввести важные понятия и обнаружить наиболее значимые закономерности. Другими словами, эти модели принадлежат фундаментальной науке, а для решения практических задач (таких как построение демографических прогнозов) могут быть применимы лишь с рядом серьезных оговорок.

Что касается прикладных исследований, то основным современным направлением моделирования и прогнозирования демографических процессов является построение имитационных моделей, основанных на более конкретном, в сравнении с рассмотренными выше моделями, описании механизмов, лежащих в основе социально-демографической динамики. Применение имитационного моделирования в демографии, как и во многих других дисциплинах, показало эффективность этого подхода, возможности и перспективы которого раскрылись с развитием современных вычислительных средств. В частности, имитационные модели позволяют проводить вычислительные эксперименты по изучению влияния большого количества различных факторов на демографические процессы. Важно также и то, что имитационная модель допускает, как правило, дальнейшее ее расширение за счет учета дополнительных факторов.

Настоящий раздел посвящен краткому описанию модели, имитирующей жизненный цикл населения через цикл отдельного индивида и позволяющей рассчитывать и анализировать различные демографические макропоказатели, касающиеся как рождаемости, так и других демографических процессов.

2.5.1. Общее описание имитационной модели

Имитационные модели формулируются в дискретном времени, причем применительно к демографии временной шаг обычно выбирается равным одному году. В качестве элементарного объекта служит индивид, характеризующийся в каждый момент времени своим статусом, под которым понимается вектор социальных переменных, таких как

- возраст,
- брачное состояние,
- количество детей в семье,
- желаемое количество детей,
- уровень образования,
- вовлеченность в экономическую деятельность,
- уровень материальной обеспеченности.

Для каждого значения статуса определена вероятность того, что индивид доживет до следующего момента времени, а также вероятности перехода за один временной шаг к каждому из других возможных для этого индивида статусов. Например, индивид может родить ребенка (если это женщина) или повысить уровень образования, или и то, и другое; либо ни одно из этих или иных возможных событий не произойдет. В последнем случае изменение статуса сводится к увеличению возраста индивида на 1 год. Вероятности каждого из возможных событий рассчитываются на основе социально-экономической статистики и для самой имитационной модели являются экзогенными, т. е. заданными извне.

Совокупность статусов всех индивидов определяет состояние демографической системы в данный момент времени. Таким образом, можно дать следующее определение.

Определение. *Имитационная демографическая модель представляет собой вероятностный автомат, состоящий из множества состояний популяции, ориентированного графа переходов и распределения вероятностей этих переходов, рассчитанного на основе вероятности смены статуса каждого отдельного индивида из этой популяции.*

В самых общих чертах имитация заключается в том, что изменение модельного статуса данного индивида происходит на основе статистических испытаний (тестов): генерируется случайная величина и ее значение сравнивается с вероятностью того или иного события (например, рождения ребенка). Если случайная величина превышает вероятность события, то считается, что это событие происходит. Такие статистические испытания проводятся по каждой из переменных, после чего индивиду присваивается статус для следующего момента времени.

Порядок действий представлен на блок-схеме (рис. 1). Рассмотрим более подробно порядок расчетов для одного шага по времени (т. е. одного года). Производится перебор всех индивидов; сначала рассматривается индивид № 1. Для него известны все перечисленные компоненты

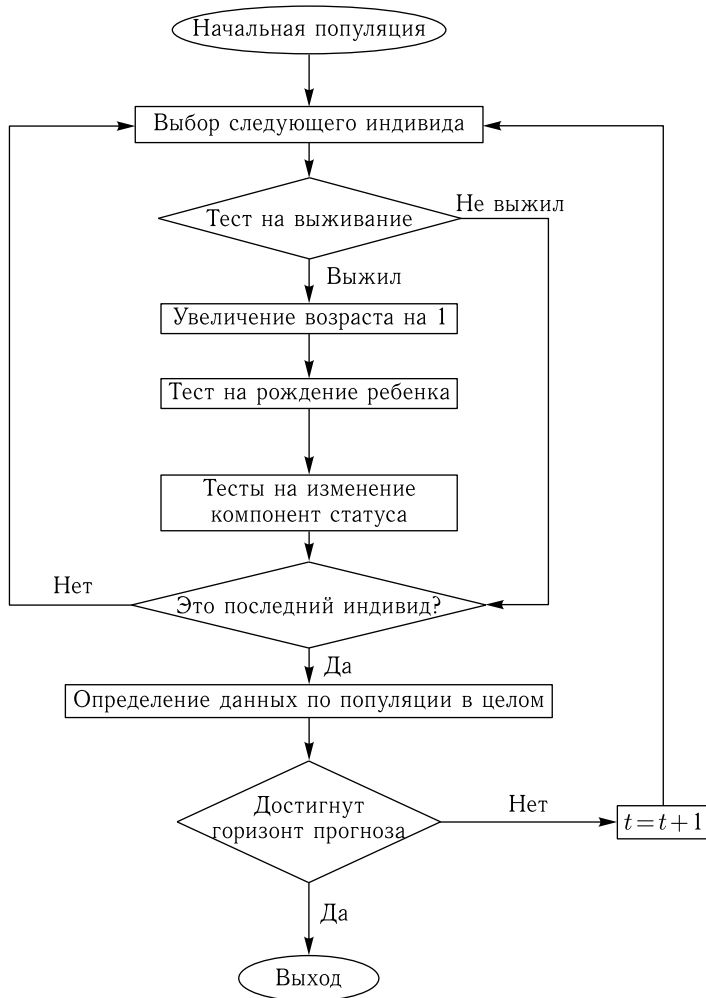


Рис. 1. Блок-схема имитационной демографической модели

статуса, и для данного значения статуса заданы вероятности того, что определенные события произойдут в течение года, а именно:

- вероятность выживания,
- вероятность изменения брачного состояния (вероятность вступления в брак для незамужних и неженатых; вероятность развода для состоящих в браке),
- вероятность рождения ребенка,

и т. д.

Для каждого из этих событий, которые могут произойти с индивидом № 1 в течение года, проводятся статистические испытания. На первом этапе проводится тест на выживание. Если индивид № 1 не проходит его, то он считается выбывшим из популяции, и расчет переходит к индивиду № 2. Если же индивид № 1 успешно проходит тест на выживание, то его возраст (т.е. значение соответствующей компоненты вектора статуса) увеличивается на единицу. В этом случае на втором этапе проводится тест на рождение ребенка у индивида № 1. В случае успешного прохождения теста на следующем временном шаге этот ребенок включается в популяцию. На третьем этапе индивид № 1 проходит тесты на изменение компонент статуса. Тем самым данный индивид получает новое значение вектора статуса, которое будет использовано на следующем временном шаге.

Данная последовательность действий проводится для каждого индивида, после чего либо в популяцию вводятся новые индивиды (в случае положительного сальдо миграции), либо часть индивидов из нее удаляется (в случае отрицательного сальдо).

На основании результатов, полученных для отдельных индивидов, затем вычисляются агрегированные демографические показатели, характеризующие популяцию в целом (см. рис. 1).

На этом шаг имитации заканчивается, и цикл повторяется для обновленной популяции вплоть до последнего расчетного момента времени. Тем самым завершается построение прогноза.

В следующем разделе мы остановимся более подробно на моделировании рождаемости.

2.5.2. Моделирование рождаемости

Ключевую роль в демографическом анализе и, в частности, при построении демографических прогнозов, играет моделирование рождаемости. В конечном счете именно этот процесс определяет направление демографического процесса: является ли воспроизводство населения расширенным, простым или суженным.

Рассмотренные в предыдущих разделах настоящей главы демографические модели предлагают относительно простые способы описания рождаемости. Так, в модели Мальтуса и логистической модели коэффициент рождаемости просто относит количество рождений в единицу времени к общей численности населения. Модели, учитывающие структуру населения, позволяют описать зависимость вероятности рождения ребенка от возраста матери. В бипарентальных моделях рождаемость определяется количеством как женщин, так и мужчин.

Однако для практических целей (например, для построения прогнозов) перечисленные подходы слишком грубы, так как в реальных населенных пунктах рождаемость зависит и от других факторов, основным из которых является желаемое женщиной количество детей в семье, т.е. так называемая *установка на количество детей* (репродуктивная установка). В свою очередь эта установка определяется культур-

ными особенностями социума и социально-экономической ситуацией. Проиллюстрируем данное положение с помощью табл. 1. Представленные в ней показатели демонстрируют, что страны, близкие с точки зрения культуры и социально-экономических параметров, имеют также близкие возрастные коэффициенты рождаемости (ср данные по странам бывшего СССР с другими странами).

Таблица 1. Возрастные коэффициенты рождаемости некоторых стран в 2008 г. Источник: US Census Bureau

Возраст	15–19	20–24	25–29	30–34	35–39	40–44	45–49
Австралия	16,4	53,3	97,3	117,7	59,9	11,4	0,5
Аргентина	54,3	106,8	124,1	109,3	59,5	18,1	2,3
Молдова	27,2	88,8	73,2	42,7	16,2	3,6	0,3
Россия	26,8	94,1	84,1	51,2	19,3	3,8	0,3
Украина	24,9	85,2	74,8	44,7	16,0	4,3	0,5
Беларусь	21,3	87,1	76,8	43,0	15,0	3,2	0,3
Германия	10,8	57,1	91,2	83,4	31,7	5,6	0,3
Сомали	65,1	245,8	301,2	269,7	235,6	123,7	78,9

Материальные факторы также оказывают определенное влияние на рождаемость. В качестве примера приведем данные опроса Левада-Центра (23–26 июля 2010 г.). На вопрос «Что, прежде всего, могло бы повлиять на ваше решение в ближайшие 2–3 года произвести на свет хотя бы одного ребенка?» и было несколько вариантов ответа:

- 28 % респондентов указали «наше решение не будет зависеть ни от каких условий»,
- 18 % респондентов — «если у меня будет уверенность в завтрашнем дне»,
- 18 % — «если будем хорошо зарабатывать»,
- 18 % — «улучшение жилищных условий»,
- 13 % — «если государство будет больше помогать семьям с несовершеннолетними детьми».

Здесь мы ограничиваемся лишь данными по выборке в целом и приводим некоторые ответы, характеризующие влияние материальных факторов на рождаемость.

Имитационное моделирование позволяет учитывать как установку на количество детей, так и уровень материальной обеспеченности семьи, и ряд других факторов, влияющих на принятие индивидами решения о рождении ребенка. Именно, вероятность рождения ребенка в модели принимается зависящей от статуса матери; для определения этой зависимости вводятся градации переменных, представленные в табл. 2.

Таблица 2. Градации переменных, определяющие вероятность рождения ребенка

Переменная	Градации
Возраст матери (8 градаций)	15–19 лет, 20–24 лет, 25–29 лет, 30–34 лет, 35–39 лет, 40–44 лет, 45–49 лет, 50–54 лет
Брачное состояние (2 градации)	В браке, не в браке (подразумевается зарегистрированный брак)
Имеющееся количество детей в семье (5 градаций)	Нет детей, 1 ребенок, 2 ребенка, 3 ребенка, 4 и более детей
Желаемое число детей (6 градаций)	Ни одного, 1 ребенок, 2 ребенка, 3 ребенка, 4 ребенка, 5 и более детей
Уровень образования матери (3 градации)	Ниже среднего, среднее, высшее
Вовлеченность в экономическую деятельность (2 градации)	Занята (работает), Не занята (не работает)
Уровень материальной обеспеченности (3 градации)	Низкий, средний, высокий (по самооценке)

При данном количестве градаций вектор статуса потенциальной матери может принимать $8 \cdot 2 \cdot 5 \cdot 6 \cdot 3 \cdot 2 = 2880$ значений. Для каждого из них в рамках модели экзогенно задается соответствующая вероятность рождения ребенка, оцененная на основании эмпирических данных и позволяющая провести статистическое испытание, как это описано выше.

Аналогичным образом проводятся тесты на изменение всех компонент статуса, что позволяет определить его значение в следующий момент времени.

2.5.3. Результаты вычислительных экспериментов и построение демографического прогноза для России

Для построения прогнозов изменения демографической ситуации на основе построенной имитационной модели определяются четыре основных сценария моделирования, т. е. формулируются предположения относительно изменения параметров модели, соответствующие четырем вариантам демографической политики:

- «высокий» прогнозный сценарий, характеризующийся активной демографической политикой, направленной как на значительное использование ресурса внешней миграции, так и на повышение рождаемости, что находит отражение в увеличении всех соответствующих параметров модели;
- «внешний» сценарий, соответствующий политике, ориентированной лишь на ресурс внешней миграции;

- «внутренний» сценарий, соответствующий политике, направленной на повышение рождаемости (уровень миграционного сальдо фиксируется на показателях 2002 г.);
- «низкий» сценарий соответствует инерционному развитию событий, при котором все параметры модели остаются на уровне показателей 2002 г.

На рис. 2 приведена динамика численности населения России в течение ближайших 100 лет, полученная в результате проведения вычислительных экспериментов с имитационной моделью.

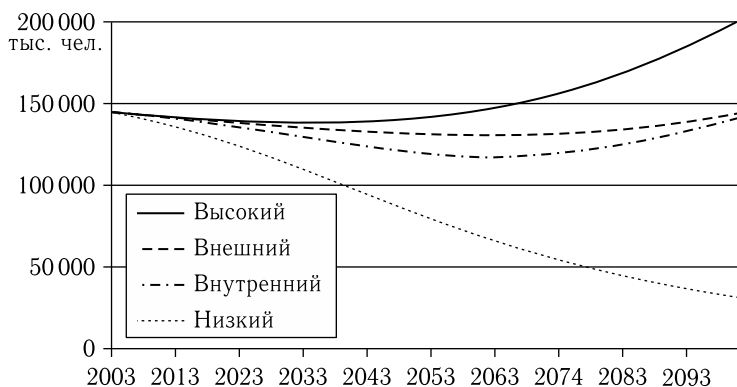


Рис. 2. Четыре варианта прогноза динамики численности населения РФ на 2003–2100-е гг. (прогноз выполнен в 2002 г.)

Сравнительный анализ результатов прогнозирования по четырем сценариям с использованием разработанной имитационной модели позволил сделать ряд выводов о перспективах и последствиях представленных вариантах демографической политики.

Прежде всего, заметим, что три из предлагаемых вариантов с точки зрения динамики численности населения разительно отличаются от «низкого» сценария, ведущего к катастрофическим последствиям — снижению численности населения практически до 30 млн человек. Три прочих варианта демографической политики стабилизируют или даже улучшают демографическую ситуацию в России. Более детальный сравнительный анализ демографических показателей в каждом из сценариев позволяет сделать вывод в пользу варианта стимулирования рождаемости («внутренний» вариант) по сравнению с миграционным вариантом («внешний»). Именно в этом случае не происходит роста некоренного населения, что позволяет избежать необходимости решения миграционных проблем, начинающих проявляться уже в современной России. Общие коэффициенты смертности и рождаемости, соответственно, показатели естественной убыли населения во «внутреннем» варианте имеют гораздо более благоприятный характер, нежели во «внешнем», где практически в течение всего периода моделирования

сохраняется ежегодная естественная убыль на уровне 500–700 тыс. человек.

Имитационная модель позволяет исследовать также динамику половозрастной структуры населения в каждом из сценариев. Показано, что только в случае активной просемейной политики, приводящей к существенному росту желаемого числа детей, возможно преодоление сужения половозрастной пирамиды и впоследствии постепенное торможение процесса старения. При этом формируется возрастная пирамида, весьма благоприятная как с демографической, так и с экономической точки зрения, характеризующаяся постоянным обновлением и пополнением трудовых ресурсов за счет вступления в трудоспособный возраст широких контингентов молодых поколений.

Таким образом, проведенные вычислительные эксперименты с использованием разработанной имитационной модели рождаемости показали, что наилучшими перспективами с точки зрения выхода из демографического кризиса обладает политика, направленная (в первую очередь) на возрождение ценностей семьи, детей, семейного образа жизни и дающая возможность преодоления российского демографического кризиса.

2.5.4. Основные требования к имитационной модели

Изучение опыта применения имитационного моделирования для демографических процессов позволило выделить 6 требований, соответствие которым существенно влияет на адекватность модели и возможность ее применения.

1. Необходимо, чтобы объектом моделирования являлось все население, а не некоторая его часть.
2. Открытость модели, позволяющая на основе простого и не требующего специальной подготовки алгоритма осуществлять дополнение ее новыми параметрами и факторами.
3. Модель должна быть максимально приближена к реальной социально-демографической сути происходящего, отражать наибольшее число возможных взаимосвязей между различными показателями и факторами динамики численности населения (в целях социально-демографической обоснованности и точности).
4. Модель должна обладать свойством факторной полноты, которое отражает необходимость учета максимального количества факторов, влияющих на моделируемый процесс.
5. Важна динамичность имитационной модели, которая заключается в изменении вероятностей перехода из одного состояния в другое с течением времени.
6. Финализируемость модели: она должна иметь законченный вид, позволяющий проводить вычислительные эксперименты с использованием реальных статистических данных.

Описанная выше имитационная модель удовлетворяет всем этим требованиям, что позволило получить хорошо обоснованные прогнозные сценарии, представленные на рис. 2.

Комментарии и библиографические указания

1. Данный раздел подготовлен на основе работ [1–7]. В частности, в этих работах получены прогнозы, представленные на рис. 2.

2. Разработка и применение метода имитационного моделирование начались еще в середине XX века и нашли свое отражение как за рубежом (см. обзор в [8]), так и в работах отечественных демографов [9–14].

3. Важным параметром данной модели является размер исходной совокупности, на основе которой проводится моделирование. Оптимальным с точки зрения вычислительных экспериментов, в первую очередь постоянства их результатов в силу «компенсации» случайностей, являются популяции в 10–20 тыс. человек. Для проведения вычислительных экспериментов на основе предлагаемой имитационной модели рождаемости использовалась популяция, численность которой в 10 000 раз меньше действительной численности населения России. Выбор такой совокупности позволял, во-первых, достичь оптимальной численности с точки зрения обоснованности получаемых результатов, а, во-вторых, существенно упрощал сравнение как с текущими статистическими данными, так и с различными прогнозами. Таким образом, для расчетов, где отправной точкой будет служить 2006 г., удобно использовать выборку, состоящую из 14 275 индивидов (при численности населения 142 млн 753 тыс. человек).

4. При построении прогнозов за основу взяты значения демографических показателей 2002 г., так как на момент проведения исследования [5] это был год последней переписи населения России.

5. Изучению влияния социальных факторов на рождаемость посвящена обширная литература — см., например, [15–22].

6. Более полные результаты опроса Левада-Центра о влиянии материальных факторов на рождаемость см. в [23].

Литература

1. Карнова В.М. Построение и исследование динамической модели рождаемости // Математическое моделирование социальных процессов Выпуск 6, М.: МАКС Пресс, 2004. под ред. А. П. Михайлова.
2. Карнова В.М. Построение и исследование динамической модели рождаемости // Тезисы докладов и выступлений на II Всероссийском социологическом конгрессе «Российское общество и социология в XXI веке: социальные вызовы и альтернативы»: В 3 т. М.: Альфа-М, 2003. Т. 1.
3. Карнова В.М. Миграция или рождаемость — имитационное моделирование как способ выбора ресурса // Социология. № 1–2, 2009.

4. Карпова В. М. Сценарии развития процессов старения населения в зависимости от демографической политики // Социология 4М № 29, 2009.
5. Карпова В. М. Некоторые предложения по изучению путей преодоления демографического кризиса // Демографические исследования № 7 (www.demographia.ru) № 7.
6. Карпова В. М. Проблема депопуляции: разные взгляды на одну катастрофу // Мир психологии. № 3, 2007.
7. Карпова В. М. Имитационная модель рождаемости и возможности ее применения для прогнозирования численности населения России // Мониторинг демографической ситуации в Российской Федерации и тенденций ее изменения. Ежегодный доклад «Демографическая ситуация в РФ», М.: КДУ, 2008.
8. Имитационное моделирование в демографии под. ред. А. Г. Волкова, М.: Статистика, 1980. С. 14–76.
9. Баркалов Н. Б. Микроимитационная модель рождаемости поколения // Городская и сельская семья под. ред. Медкова В. М.. М.: 1987. С. 170–196.
10. Елизаров В. В. Перспективы исследования семьи. М.: Мысль, 1987. С. 127–141.
11. Волков А. Г., Сороко Е. Л. Имитационная модель демографического развития семьи // Демографические процессы и их закономерности. М., 1986.
12. Медков В. М. Микроимитационные модели в демографии: проблемы и возможности использования в прогнозировании динамики населения и демографических процессов // Математическое моделирование социальных процессов. М.: МАКС Пресс, 2002; под ред. А. П. Михайлова.
13. Антонов А. И., Медков В. М., Нетрусов М. А. Демография семьи — проблемы моделирования семейного цикла жизни // Вестник московского университета, сер. 18, Социология и политология. М., 1997. № 2. С. 97–108.
14. Гражданников Е. Д. Прогностические модели социально-демографических процессов. Новосибирск. 1974. С. 65–71.
15. Антонов А. И. Перспективы рождаемости в связи с тенденциями установок детности // Демографические процессы в России XXI века. М., 2002. 59 с.
16. Антонов А. И., Медков В. М. Архангельский ВН Демографические процессы в России XXI века. М., Грааль, 2002, 168 с.
17. Борисов В. А. Желаемое число детей в российских семьях по данным микропереписи населения России 1994 года // Вестник Московского университета. 1997, Серия 18, Социология и политология, 2. 36 с.
18. Медков В. М. Репродуктивное поведение и социальная мобильность: ценностно-ориентационный аспект // Демографические процессы в России в XXI в. М., 2002. 82 с.
19. Гольцова Е. В. Экологические факторы демографического поведения семьи // Социология, 3. 2008. С. 30–33.
20. Шабунова А. А., Калачикова О. Н. Репродуктивный выбор семьи: факторы и условия // Социология, 33, 2008. С. 72–79.
21. Варламова С. Н., Носкова А. В., Седова Н. Н. Семья и дети в жизненных установках россиян // Социс. 2006, 11.
22. Skirbekk V. Fertility trends by social status // Demographic Research. V. 18, pp. 45–180.
23. Сайт Левада-Центра <http://www.levada.ru/press/2010081900.html>.

Глава 3

ПСИХОЛОГИЯ

3.1. Метод парных сравнений и описание процесса восприятия

Одним из широко распространенных в социальных науках методов измерения является метод парных сравнений, цель которого — шкалирование, т. е. измерение некоторого качества объектов. В данном случае речь идет о таких субъективно воспринимаемых качествах, как, например «привлекательность профессий», «уровень коррупции в различных сферах жизни», «привлекательность политических партий для избирателя» и т. д., а не о величинах, для измерения которых существуют объективные методы (длина предметов, сила электрического тока и др.). Процедура данного метода заключается в том, что сначала респонденту предлагается провести попарные сравнения всех объектов по выраженности изучаемого качества, а затем по описанной ниже методике вычисляются итоговые шкальные значения.

Предположим, например, что при построении градостроительного плана проводится социологическое исследование, имеющее своей целью определение привлекательности каждого из 8 районов данного города для жилой застройки. Нетрудно вычислить, что из этих 8 районов можно выбрать $(8 \cdot 7)/2 = 28$ различных пар. Соответственно в ходе опроса каждому респонденту задается 28 вопросов о том, какой из двух районов данной пары он предпочел бы в качестве места для проживания. На этапе финальной обработки результатов данные по всем респондентам сводятся воедино. С помощью специального алгоритма строится шкала, и каждому району присваивается определенное численное значение привлекательности. Описанию этого алгоритма посвящен п. 3.2.

В основе метода парных сравнений лежит известный в психологии закон сравнительных суждений, представляющий собой модель воспринимающего (измеряющего) субъекта; он рассматривается в п. 3.1.1. Алгоритмической стороне метода посвящен п. 3.1.2, который является основным в данном разделе. В комментариях показано, что предположения, лежащие в основе метода парных сравнений, не носят абсолютный характер, а также описан класс ситуаций, в которых эти предположения не выполняются. Наконец, в п. 3.1.3 настоящего раздела

излагается подход к моделированию восприятия субъекта, акцентированный на вопросах эффективности когнитивной системы.

3.1.1. Закон сравнительных суждений

При проведении исследования методом парных сравнений может возникнуть парадоксальная, на первый взгляд, ситуация, когда респондент оценивает объект А как более предпочтительный, чем объект В, а объект В как более предпочтительный, чем С, но при этом объект С более предпочтительный, чем А.

Для объяснения этого парадокса обратимся к модели индивидуального восприятия, основанной на так называемом *законе сравнительных суждений Терстоуна*. Предварительно рассмотрим ситуацию, более простую, чем социологические измерения, а именно: взвешивание человека на напольных весах. Результат взвешивания — показания весов — является случайной величиной. Если весы не дают систематической ошибки, то математическое ожидание этой случайной величины равно значению истинной массы тела. Различия между показаниями весов и истинной массой возникают вследствие множества факторов: усталость пружины после предыдущих взвешиваний, расположение Луны, атмосферное давление и температура воздуха, положение человека на платформе весов и т. д. С приемлемой степенью точности эту случайную величину можно считать нормально распределенной.

Теперь предположим, что необходимо провести поочередное взвешивание двух людей и определить более тяжелого из них. Результат каждого из взвешиваний является нормально распределенной случайной величиной; и математическое ожидание в каждом из двух случаев равняется истинному значению соответствующей массы. При этом с некоторой вероятностью (очевидно, меньшей, чем $1/2$) измеренная масса более тяжелого человека окажется меньше, чем измеренная масса более легкого. Эту вероятность нетрудно определить, если известны истинные массы взвешиваемых людей и характеристика точности весов — дисперсия нормального распределения.

Приведенное рассуждение позволяет по аналогии сформулировать модель измеряющего («проводящего измерения») субъекта, лежащую в основе метода парных сравнений. Для этого обратимся еще раз к примеру с привлекательностью городских районов для проживания.

Исследователь обращается к респонденту с вопросом о том, какой из двух районов тот считает более предпочтительным. Респондент называет один из районов. Рассмотрим модель принятия решения индивидом.

Предполагается, что каждому району соответствует некоторое «истинное значение привлекательности», измеряемое индивидом. Результатом измерения привлекательности конкретного района является случайная величина, причем мы принимаем гипотезу о том, что она является нормально распределенной и ее математическое ожидание равно «истинному значению».

Таким образом, операцию сравнения респондентом двух районов можно представить в виде трех последовательных действий, первыми двумя из которых являются измерение привлекательности первого и второго районов. Результатами этих измерений является нормально распределенные случайные величины X_1 , X_2 , имеющие математические ожидания, равные «истинным значениям привлекательности» соответственно первого и второго районов (напомним, что нахождение этих «истинных значений привлекательности» и составляет цель метода парных сравнений) и стандартное отклонение σ_X .

Третье действие — сравнение результатов этих измерений — респондент проводит путем вычитания и получает случайную величину $Z = X_1 - X_2$, имеющую смысл разности измеренных значений. Она имеет нормальное распределение с математическим ожиданием, равным разности «истинных значений привлекательности». Если, например, в данном эксперименте случайная величина Z реализуется как положительное число, то респондент сообщает исследователю, что он предпочитает первый район. Тем самым привлекательность первого района оказывается измеренной как более высокая, чем второго района (заметим, что это не означает, будто первый район действительно является более привлекательным для жилья: речь идет не об «истинных значениях привлекательности», а лишь о результатах их измерений).

Если исследователь задает вопрос о сравнении данных районов большому количеству респондентов, то он получает статистику, позволяющую ввести меру разности «истинных значений привлекательности». Например, если каждый из районов является более привлекательным примерно для половины респондентов, то «истинные значения» примерно равны. Если же, к примеру, первый район предпочитается подавляющим большинством респондентов, то его «истинное значение привлекательности» существенно выше.

Метод парных сравнений (см. ниже) основан на данной модели измеряющего субъекта. Он оперирует не с двумя измеряемыми (шкалируемыми) объектами, а с большим их количеством. Каждому респонденту предлагается сравнить все эти объекты попарно и в каждой паре выбрать более предпочтительный. Метод парных сравнений позволяет построить шкальные значения объектов на основании всех парных сравнений, взятых в совокупности (например, для случая 8 районов: на основании всех 28 парных сравнений, проведенных каждым из респондентов). Можно сказать, что он соединяет две важные идеи: закон сравнительных суждений Терстоуна (т.е. модель измеряющего субъекта) и алгоритм построения единой шкалы на основе многочисленных парных сравнений.

Перед тем, как перейти к математической формализации модели измеряющего субъекта, рассмотрим вопрос о масштабе искомой шкалы. В каких единицах изменяется привлекательность района?

Можно показать (см. упражнение 3), что наиболее рациональным в данном случае оказывается выбрать единицу измерения привлекательности, равную стандартному отклонению случайной величины X_1 (или, что то же самое, случайной величины X_2). Таким образом, всюду далее будем полагать $\sigma_X = 1$.

Перейдем к формализации изложенных выше рассуждений.

Обозначим через μ_1 и μ_2 искомые истинные значения переменной (например, привлекательности) для первого и второго объектов. Случайные величины X_1 , X_2 , имеющие смысл результатов измерений переменной для первого и второго объектов, имеют нормальные распределения с математическими ожиданиями μ_1 и μ_2 соответственно:

$$\rho_1(x) = \frac{1}{\sqrt{2\pi}} \exp \left[-\frac{(x - \mu_1)^2}{2} \right],$$

$$\rho_2(x) = \frac{1}{\sqrt{2\pi}} \exp \left[-\frac{(x - \mu_2)^2}{2} \right].$$

При отсутствии корреляции между X_1 и X_2 разность $Z = X_1 - X_2$ является, как известно, нормально распределенной случайной величиной с математическим ожиданием $\mu_1 - \mu_2$ и дисперсией $D_Z = 2$. Таким образом, ее плотность распределения имеет вид

$$\rho(x) = \frac{1}{\sqrt{2} \sqrt{2\pi}} \exp \left[-\frac{(x - (\mu_1 - \mu_2))^2}{2(\sqrt{2})^2} \right].$$

Вероятность того, что при измерении величина X_1 будет определена как бóльшая, чем X_2 , или (что то же самое) — вероятность того, что случайная величина Z положительна, вычисляется интегрированием по области $0 < x < \infty$

$$P(X_1 > X_2) = \int_0^{+\infty} \rho(x) dx = \frac{1}{\sqrt{2} \sqrt{2\pi}} \int_0^{+\infty} \exp \left[-\frac{(x - (\mu_1 - \mu_2))^2}{2(\sqrt{2})^2} \right] dx. \quad (1)$$

Выполнив под знаком интеграла замену переменной $t = (x - (\mu_1 - \mu_2))/\sqrt{2}$, получим из (1):

$$P(X_1 > X_2) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{(\mu_1 - \mu_2)/\sqrt{2}} \exp \left[-\frac{t^2}{2} \right] dt = \frac{1}{2} + \Phi_0 \left(\frac{\mu_1 - \mu_2}{\sqrt{2}} \right), \quad (2)$$

где $\Phi_0(x) = \frac{1}{\sqrt{2\pi}} \int_0^x e^{-t^2/2} dt$ — нормированная функция Лапласа.

Из (2) следует, что вероятность $P(X_1 > X_2)$ зависит только от разности $z = \mu_1 - \mu_2$, но не от самих величин μ_1 , μ_2 , взятых по отдельности. Поэтому вероятность $P(X_1 > X_2)$ мы далее будем обозначать $p(z)$.

Следовательно, если известно значение z , показывающее, насколько различаются истинные значения переменной, то с помощью таблицы функции $\Phi_0(x)$ нетрудно определить вероятность $p(z) = P(X_1 > X_2)$ того, что результат измерения будет иметь вид $X_1 > X_2$.

Метод парных сравнений применяется для решения обратной задачи. Именно, известной является эмпирическая вероятность $p(z)$ (определяемая путем опроса респондентов), а величина z подлежит нахождению. Соответствующая таблица (см. табл. 1) может быть легко получена из таблиц функции $\Phi_0(x)$.

Таблица 1. Соотношение между величинами p и z

p	0,1	0,2	0,3	0,4	0,5	0,6	0,7	0,8	0,9
z	-1,28	-0,84	-0,52	-0,25	0	0,25	0,52	0,84	1,28

Таким образом, разность искомых значений μ_1 и μ_2 принимается равной

$$\mu_1 - \mu_2 = z_{12}, \quad (3)$$

где z_{12} — решение уравнения $p_{12} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^z e^{-t^2/2} dt$, а p_{12} — эмпирическая вероятность, определяемая в результате эксперимента. Данное утверждение называется законом сравнительных суждений Терстоуна. Например, если при сравнении двух районов 70 респондентов из 100 ответили, что район А является более предпочтительным для жизни, чем район Б, то $p = 0,7$ и (в соответствии с табл. 1) в качестве разницы искомых «истинных значений привлекательности» следует выбрать $\mu_1 - \mu_2 = 0,52$. Если же район А сочли более предпочтительным 50 респондентов из 100, то $\mu_1 - \mu_2 = 0$.

Рассмотрим теперь, как на основе закона сравнительных суждений проводится шкалирование методом парных сравнений.

3.1.2. Метод парных сравнений

Предположим, что целью исследования является измерение некоторого качества у n объектов, т. е. нахождение значений μ_i , $i = 1, \dots, n$. Каждому из респондентов предъявляются всевозможные пары этих объектов и в каждой паре предлагается выбрать объект, у которого данное качество выражено сильнее. Для простоты далее будем говорить о том, что респонденты выбирают более предпочтительный объект в каждой паре. Количество таких пар, очевидно, равно $n(n-1)/2$.

В соответствии с ответами респондентов исследователь получает эмпирические вероятности (или, что то же самое, относительные частоты) p_{ij} ($i, j = 1, \dots, n$, $i \neq j$), показывающие долю ответов, в которых объект i оценивался как более предпочтительный, чем объект j .

Положив дополнительно $p_{ii} = 0,5$, получим матрицу эмпирических вероятностей P ,

$$P = \begin{pmatrix} 0,5 & p_{12} & p_{13} & \dots & p_{1n} \\ p_{21} & 0,5 & p_{23} & \dots & p_{2n} \\ p_{31} & p_{32} & 0,5 & \dots & p_{3n} \\ \dots & \dots & \dots & \dots & \dots \\ p_{n1} & p_{n2} & p_{n3} & \dots & 0,5 \end{pmatrix}.$$

Эта матрица обладает следующими свойствами:

$$p_{ij} + p_{ji} = 1, \quad i, j = 1, \dots, n, \quad (4)$$

$$0 \leq p_{ij} \leq 1, \quad i, j = 1, \dots, n. \quad (5)$$

Если хотя бы один из элементов равен единице, $p_{ij} = 1$ (и, соответственно, $p_{ji} = 0$), то матрица P (а также матрица Z , см. далее) называется неполной. Это имеет место, если все без исключения респонденты предпочитают i -й объект j -му. Далее мы ограничимся рассмотрением случая полной матрицы, т. е. будем полагать выполненными строгие неравенства

$$0 < p_{ij} < 1, \quad i, j = 1, \dots, n. \quad (6)$$

Неравенство (6) позволяет выполнить так называемое p - z -преобразование, т. е. перейти к переменным z_{ij} ($i, j = 1, \dots, n, i \neq j$) по формулам

$$p_{ij} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z_{ij}} e^{-t^2/2} dt, \quad i, j = 1, \dots, n \quad (7)$$

(на практике пользуются таблицами, подобными табл. 1, либо специализированными программными продуктами). Тем самым из матрицы P получаем матрицу Z -оценок:

$$Z = \begin{pmatrix} 0 & z_{12} & z_{13} & \dots & z_{1n} \\ z_{21} & 0 & z_{23} & \dots & z_{2n} \\ z_{31} & z_{32} & 0 & \dots & z_{3n} \\ \dots & \dots & \dots & \dots & \dots \\ z_{n1} & z_{n2} & z_{n3} & \dots & 0 \end{pmatrix}.$$

Ее элементы удовлетворяют свойству

$$z_{ij} = -z_{ji}, \quad i, j = 1, \dots, n, \quad (8)$$

которое следует из (4), (7).

В соответствии с законом сравнительных суждений (3) имеем для искомым значений μ_i , $i = 1, \dots, n$, систему уравнений

$$\mu_i - \mu_j = z_{ij}, \quad i, j = 1, \dots, n. \quad (9)$$

Если отбросить тождества вида $\mu_i - \mu_i = 0$, $i = 1, \dots, n$, и учесть свойство (8), то получим, что система (9) содержит фактически $(n-1)n/2$

уравнений для n неизвестных. Следовательно, при $n > 3$ количество уравнений превышает количество неизвестных.

При этом уравнения (8) являются, вообще говоря, несогласованными в следующем смысле. Если из уравнения $\mu_1 - \mu_3 = z_{13}$ вычесть уравнение $\mu_1 - \mu_2 = z_{12}$, то получим равенство $\mu_2 - \mu_3 = z_{13} - z_{12}$. С другой стороны, $\mu_2 - \mu_3 = z_{23}$. Таким образом, согласованность уравнений предполагает, что $z_{13} - z_{12} = z_{23}$, что, вообще говоря, не имеет места (поскольку величины z_{ij} получены независимо друг от друга, в результате обработки эмпирических данных).

Таким образом, система уравнений (8) является, как правило, несовместной. Поэтому метод парных сравнений, алгоритм которого излагается ниже, нацелен на получение значений μ_i , $i = 1, \dots, n$, минимизирующих сумму квадратов невязок уравнений (8). Другими словами, цель анализа заключается в том, чтобы определить значения переменных μ_i , при которых функция

$$f(\mu_1, \dots, \mu_n) = \sum_{i,j=1}^n (\mu_i - \mu_j - z_{ij})^2 \quad (10)$$

имеет минимум.

Чтобы найти эти значения, выпишем функцию (10) в развернутом виде,

$$\begin{aligned} f(\mu_1, \dots, \mu_n) = & (\mu_1 - \mu_1 - z_{11})^2 + (\mu_1 - \mu_2 - z_{12})^2 + \dots + (\mu_1 - \mu_n - z_{1n})^2 + \\ & + (\mu_2 - \mu_1 - z_{21})^2 + (\mu_2 - \mu_2 - z_{22})^2 + \dots + (\mu_2 - \mu_n - z_{2n})^2 + \dots + \\ & + (\mu_n - \mu_1 - z_{n1})^2 + (\mu_n - \mu_2 - z_{n2})^2 + \dots + (\mu_n - \mu_n - z_{nn})^2, \end{aligned}$$

и вычислим частные производные. Получим, например для $\partial f / \partial \mu_1$:

$$\begin{aligned} \frac{\partial f}{\partial \mu_1} = & 2[(\mu_1 - \mu_2 - z_{12}) + \dots + (\mu_1 - \mu_n - z_{1n}) - \\ & - (\mu_2 - \mu_1 - z_{21}) - (\mu_n - \mu_1 - z_{n1})] = \\ = & 2 \left[2n\mu_1 - 2 \sum_{i=1}^n \mu_i - \sum_{j=1}^n z_{1j} + \sum_{i=1}^n z_{i1} \right] = 4 \left[n\mu_1 - \sum_{i=1}^n \mu_i + \sum_{i=1}^n z_{i1} \right]. \end{aligned}$$

Аналогичные выражения имеют место для других частных производных функции $f(\mu_1, \dots, \mu_n)$. Приравнявая их к нулю и учитывая соотношение (8), получаем систему n уравнений для n неизвестных μ_i , $i = 1, \dots, n$:

$$\begin{aligned} n\mu_1 - \sum_{i=1}^n \mu_i - \sum_{i=1}^n z_{1i} &= 0, \\ &\dots \\ n\mu_n - \sum_{i=1}^n \mu_i - \sum_{i=1}^n z_{ni} &= 0. \end{aligned} \quad (11)$$

Однако из нее не могут быть однозначно определены значения неизвестных, так как уравнения не являются линейно независимыми. В этом нетрудно убедиться, сложив все уравнения системы (11) и получив (с учетом (8)) тождество.

Действительно, система (11) имеет бесконечное количество решений: если набор чисел $(\mu_1^0, \mu_2^0, \dots, \mu_n^0)$ является решением, то $(\mu_1^0 + C, \mu_2^0 + C, \dots, \mu_n^0 + C)$ также является решением. Это отражает тот факт, что шкала, получаемая с помощью метода парных сравнений, инвариантна относительно сдвига. Другими словами, при оценке привлекательности четырех районов для жилья могут быть получены значения 10, 5, 3, 2 либо 20, 15, 13, 12, либо 5, 0, -2, -3 и т. д. В определенном смысле наиболее удобным является вариант, при котором сумма искомым значений равна нулю, $\mu_1 + \mu_2 + \dots + \mu_n = 0$. Тогда система (11) принимает вид

$$n\mu_1 - \sum_{i=1}^n z_{1i} = 0, \dots, n\mu_n - \sum_{i=1}^n z_{ni} = 0,$$

откуда получаем

$$\mu_1 = \frac{1}{n} \sum_{i=1}^n z_{1i}, \dots, \mu_n = \frac{1}{n} \sum_{i=1}^n z_{ni}. \quad (12)$$

Полученная формула (12) дает искомым результат, т. е. полученные по методу парных сравнений значения выраженности изучаемого качества у рассматриваемых объектов.

Этот результат имеет следующую трактовку. Матрица эмпирических вероятностей P составлена таким образом, что чем более предпочтительным является i -й объект (по сравнению с другими объектами данной совокупности), тем более высокие значения p_{ij} расположены в i -й строке матрицы P . Следовательно, тем более высокие значения z_{ij} находятся в i -й строке матрицы Z и тем выше соответствующее значение μ_i .

Так как для вычисленных по формуле (12) значений μ_i имеет место равенство $\mu_1 + \mu_2 + \dots + \mu_n = 0$, то некоторые из них будут отрицательными. Если, в рамках конкретного исследования, более удобным является расположение всех μ_i в положительной области, то следует произвести сдвиг с учетом того, что общее решение системы (11) имеет вид

$$\mu_1 = \frac{1}{n} \sum_{i=1}^n z_{1i} + C, \dots, \mu_n = \frac{1}{n} \sum_{i=1}^n z_{ni} + C,$$

где C — произвольная постоянная.

Таким образом, метод парных сравнений позволяет проводить шкалирование, т. е. измерение субъективно воспринимаемого качества у множества заданных объектов. В его основе лежит закон сравнитель-

ных суждений, представляющий собой математическую формулировку соответствующей модели воспринимающего (измеряющего) индивида.

3.1.3. Моделирование когнитивной системы индивида

Перейдем к другому аспекту восприятия реальности индивидом. Здесь рассмотрим вопрос о внутренней интеллектуальной организации субъекта, более точно, — о его когнитивной организации. При этом предполагается, что в процессе эволюции человека должна сформироваться наиболее эффективная когнитивная система. Эффективность в данном случае означает высокую способность обрабатывать информацию, получаемую из внешней среды.

Напомним, что в соответствии с определением К. Шеннона, информация — это мера изменения неопределенности, она равна разности энтропий до и после акта получения информации. В частности, информация, полученная при реализации случайного эксперимента равна энтропии этого случайного эксперимента, так как после реализации неопределенность равна нулю (после эксперимента стало известно, какой именно исход реализовался).

Если случайный эксперимент имеет n исходов S_1, \dots, S_n , происходящих с вероятностями p_1, \dots, p_n , то энтропия эксперимента по определению равна

$$E = \sum_{i=1}^n p_i \ln_2 \frac{1}{p_i} \text{ бит.}$$

Соответственно этой же величине равна и информация, полученная в результате эксперимента: $I = E - 0 = E$.

Например, при бросании «честной» монеты возможны два исхода, каждый из них имеет 50-процентную вероятность. Другими словами, $p_1 = p_2 = 1/2$, и для количества информации имеем

$$I = \frac{1}{2} \ln_2 2 + \frac{1}{2} \ln_2 2 = 1.$$

Таким образом, когда нам становится известно, какой стороной упала монета, мы получаем информацию, равную 1 биту.

Используемая в данной теории модель субъекта предполагает, что психологическая ситуация подобна случайному эксперименту в том смысле, что она является для него источником информации, причем количество извлекаемой информации зависит от его когнитивной эффективности. В рассматриваемом здесь базовом варианте модели предполагается, что все исходы равновероятны. Очевидно, чем больше количество исходов возможно в эксперименте, чем больше его энтропия и тем большую информацию может извлечь из психологической ситуации субъект.

Например, если, наблюдая за количеством прохожих на улице, субъект А способен различать лишь два исхода, «много людей» и «мало людей», то по итогам наблюдения он получает 1 бит информации (аналогично тому, как было в эксперименте с бросанием монеты).

Если, в то же время субъект Б способен различить 5 исходов, «очень мало», «мало», «средне», «много», «очень много», то по итогам наблюдения он получает количество информации (в битах), равное

$$I = \frac{1}{5} \ln_2 5 + \frac{1}{5} \ln_2 5 + \frac{1}{5} \ln_2 5 + \frac{1}{5} \ln_2 5 + \frac{1}{5} \ln_2 5 = \ln_2 5 \approx 2,32.$$

Таким образом, из одного и того же наблюдения субъект Б получает 2,32 бита информации, а субъект А — лишь 1 бит. Тем самым субъект Б более когнитивно эффективен.

Что могут дать приведенные рассуждения для понимания «внутреннего устройства» человека? Примем за гипотезу, что сложившаяся когнитивная система субъекта является максимально эффективной, так как только такая система могла сформироваться в процессе эволюции. Другими словами, эволюционная гипотеза дает основания рассматривать задачу максимизации когнитивной эффективности субъекта.

Сформулируем задачу максимизации эффективности. Предполагается, что конкретный субъект в конкретной ситуации способен распознавать следующие аспекты.

а. Определенное количество измерений D этой ситуации. Например, если речь идет об оценке некоторой улицы в плане материальной обеспеченности ее жителей, то в качестве измерений могут выступать: внешний вид жилых домов, принадлежность супермаркетов к сетям экономического или премиального класса, наличие дорогих бутиков одежды или ювелирных изделий, класс запаркованных во дворах автомобилей и т. д.

б. Определенное количество градаций для каждого измерения (различные измерения могут иметь различное количество градаций).

Тем самым образуется «фазовое пространство», состоящее из $n_1 n_2 \dots n_D$ точек (здесь n_i — количество градаций i -го измерения, $i = 1, \dots, D$).

Когнитивное ограничение индивида состоит в том, что суммарное количество градаций по всем измерениям не может превышать некоторого наперед заданного числа R , т. е. $n_1 + n_2 + \dots + n_D \leq R$. Если предположить, что субъект своей волей пытается выстроить собственную когнитивную систему наиболее эффективным образом, то он решает следующую задачу: выбрать количество измерений N и распределить по этим измерениям R градаций так, чтобы фазовое пространство состояло из как можно большего количества точек. Эти точки могут быть отождествлены с различными (и, в рамках базовой модели, равновероятными) исходами; чем их больше, тем выше энтропия «до реализации», и тем большую информацию способен извлечь субъект из ситуации.

Пусть, например, индивид пытается решить задачу о том, насколько оживленной является в данный момент улица, на которой он находится. В рамках этого примера (для определенности) будем считать,

что максимальное суммарное количество градаций по всем измерениям равно $R = 6$. Тогда субъект имеет следующие варианты действий.

А. Выбрать одно измерение (например, «количество людей на улице»): $N = 1$ и шкалу из 6 градаций («очень мало», «мало», «немного меньше среднего» и т. д.). В этом случае он может различить 6 исходов эксперимента, т. е. фазовое пространство имеет 6 точек.

Б. Распределить шесть градаций по двум измерениям (например, «количество людей» и «количество машин») и на каждом из них сформировать шкалу из трех градаций (например, «мало», «средне», «много»). Фазовое пространство имеет в этом случае 9 точек.

В. Распределить шесть градаций по двум измерениям, на одном из них будет шкала из двух градаций, на другом — из четырех градаций. Полученное таким образом фазовое пространство имеет 8 точек.

Г. Распределить шесть градаций по трем измерениям (например, «количество людей», «количество машин», «уровень шума»), на каждом из них будет шкала из двух точек. Фазовое пространство имеет 8 точек.

Очевидно, что наиболее эффективным является вариант Б; в этом случае информация, которую субъект получит из эксперимента, равна $I = \ln_2 9 \approx 3,17$ бит.

Рассмотрим задачу максимизации когнитивной эффективности в более общем виде. В ней можно выделить два вопроса: 1) следует ли равномерно распределять градации по измерениям, 2) какое количество градаций в каждом измерении является оптимальным (или, что то же самое, какое количество измерений является оптимальным)?

Для ответа на первый вопрос обратимся к вариантам Б, В из рассмотренного выше примера. В обоих случаях индивид имеет по два измерения, но в первом варианте, $R = 6$, градаций распределены по ним равномерно (по три на каждое измерение), во втором — неравномерно (две градации на одном измерении, и четыре на другом). Более эффективным в данном примере оказался вариант Б, при котором градации распределены равномерно.

Очевидно, что распределить градации поровну между всеми измерениями не всегда возможно, тогда наиболее равномерным является распределение, при котором количества градаций на шкалах различных измерений отличаются не более, чем на единицу. Обоснование того, что такое распределение является наиболее эффективным, дается следующей леммой (доказательство вынесено в упражнения).

Лемма 1. Если на одной шкале имеется n_1 градаций, а на другой — n_2 , причем $n_1 < n_2$, то перераспределение одной градации со второй шкалы на первую либо не изменит количество точек фазового пространства (если $n_2 - n_1 = 1$), либо увеличит его (если $n_2 - n_1 > 1$).

Таким образом, чем более равномерно распределены градации между измерениями, тем выше когнитивная эффективность индивида.

Перейдем ко второму вопросу: какое количество градаций в каждом измерении является оптимальным? (При этом для простоты будем

производить вычисления так, будто количество градаций быть может нецелым). Распределяя равномерно R градаций по N измерениям, на каждом из измерений получаем шкалу, состоящую из $x = R/N$ градаций. Тогда количество точек фазового пространства равно $n = x^N = x^{R/x}$, а извлекаемая из ситуации информация:

$$I = \ln_2 x^{R/x} = \frac{R}{x} \ln_2 x.$$

Чтобы найти максимум функции $I(x)$, продифференцируем ее. Получим

$$\frac{dI}{dx} = -\frac{R}{x^2} \ln_2 x + \frac{R}{x^2 \ln 2} = \frac{R}{x^2 \ln 2} (1 - \ln x).$$

Отсюда следует, что информация $I(x)$ достигает максимума при $x = e \approx 2,7$.

Округляя, получаем, что оптимальное количество градаций — три, при необходимости — на некоторых измерениях две. Это свойство получило название тернарности.

Итак, в соответствии с изложенной выше гипотезой о том, что эволюционный процесс привел человека к когнитивной организации с наиболее высокой возможной эффективностью, получаем следующую схему когнитивной организации субъекта.

Количество градаций R является индивидуальной характеристикой субъекта: чем выше это значение, тем больше информации он способен извлекать из психологической ситуации. При оценке конкретной ситуации он выделяет в ней $R/3$ измерения (с округлением в большую сторону) и распределяет градации по этим измерениям — на каждом из них формируется шкала с двумя или тремя градациями (при этом, если R кратно трем, то на всех измерениях получается по три градации). Тем самым формируется фазовое пространство, состоящее из максимально возможного при данном значении R количества точек $n(R)$. Сопоставляя психологическую ситуацию с одной из точек этого пространства, индивид извлекает из ситуации определенное количество информации. Если все точки фазового пространства равновероятны, то это количество информации равно $I = \ln_2 n(R)$ бит.

Упражнения

1. Данное упражнение поясняет логику, лежащую в основе закона сравнительных суждений. Каждая монета весит 1 грамм. Весы предназначены для сравнения масс мешочков с монетами и определяют эту массу как целое число граммов (для простоты предполагается, что масса пустого мешочка пренебрежимо мала). В соответствии со своим внутренним алгоритмом, при сравнении двух мешочков, весы на первом этапе измерения взвешивают каждый из них. При этом измеренная масса (результат взвешивания) равна истинной массе с вероятностью 0,4; на 1 грамм больше — с вероятностью 0,2, на 2 грамма

больше — с вероятностью 0,1, и на 1 грамм меньше истинной массы — с вероятностью 0,3. На втором этапе весы сравнивают результаты измерений и делают вывод: масса первого мешочка больше, меньше или равна второй. Требуется выписать ряд распределения случайной величины, имеющей смысл разности измеренных масс, если первый мешочек на 1 грамм тяжелее второго, а также найти вероятность того, что (а) весы правильно определяют первый мешочек как более тяжелый, (б) весы определяют их равными, (в) весы определяют более тяжелый мешочек как более легкий.

2. С помощью весов, описанных в упражнении 1, изучают две кучки монет, одна из которых на 2 грамма тяжелее второй. Найти вероятность того, что весы правильно определяют первую кучку как более тяжелую, что весы определяют их равными, что весы определяют более тяжелую кучку как более легкую.

3. При формулировке закона сравнительных суждений можно не требовать, чтобы единица измерения была равной стандартному отклонению σ случайных величин X_1, X_2 . В этом случае формулу (3) следует заменить на $\mu_1 - \mu_2 = \sigma z_{12}$. Задание: провести формализм метода парных сравнений (включая минимизацию суммы квадратов отклонений) в данном случае.

4. Доказать лемму 1.

Комментарии и библиографические примечания

1. Закон сравнительных суждений был предложен в 1927 г. американским психологом Л. Терстоуном в работах [1, 2] в более общей формулировке, предполагающей, в частности, что дисперсии случайных величин X_1 и X_2 могут быть различными. Там же рассмотрены частные случаи данного закона. Одним из наиболее простых и популярных является вариант, рассмотренный в данном разделе. Иногда его называют также «пятым случаем Терстоуна». Отметим, что существуют аналогичные подходы, опирающиеся на иные (помимо нормального) распределения.

2. Гипотеза о том, что случайные величины X_1, X_2 , имеют нормальное распределение, проверялась Л. Терстоуном экспериментально для воспринимаемых индивидом величин различной природы. Было показано, что она верна в большинстве случаев, однако не всегда. Поэтому, строго говоря, при эмпирическом исследовании ранее не изучавшихся величин требуется соответствующая проверка этой гипотезы.

3. Идеи, лежащие в основе метода парных сравнений, были предложены в [1, 2], алгоритм разработан в [3], см. также статьи [4–6] и учебник [7].

4. При изложении метода парных сравнений в п. 3.2 была сделана оговорка о том, что его алгоритм излагается лишь для случая полной матрицы. Если же имеет место случай неполной матрицы (т. е., хотя бы

одна из величин p_{ij} равна нулю или единице), то этот алгоритм не применим, так как выполнить p - z -преобразование для данного p_{ij} невозможно ввиду того, что уравнение (7) не имеет решений. В этом случае метод парных сравнений принимает более сложную форму — см. [6, 7].

5. Рассмотренный в данном разделе закон сравнительных суждений предполагает, что «измерительный прибор» индивида не имеет систематической ошибки. Однако эмпирически установлено, что при измерениях определенного вида субъект обязательно совершает систематическую ошибку, при этом направление смещения может быть различным для разных субъектов.

Фундаментальным основанием такого подхода является теория психологических конструкторов, предложенная Дж. А. Келли [8]. В соответствии с этой теорией каждый субъект воспринимает окружающий мир с помощью сформированной предыдущим опытом системы личностных конструкторов, т. е. биполярных понятийных шаблонов, таких как «холодный–горячий», «дружеский–враждебный», «умный–глупый» и т. д. При необходимости проанализировать тот или иной объект или процесс происходит активизация тех или иных конструкторов. Например, в определенной ситуации один субъект активизирует конструктор «выгодно-убыточно», а другой субъект — конструктор «честно–нечестно». Эксперименты показали, что при оценке с помощью так называемого *метода категоризации* (когда респондент оценивает промежуточное значение при данных максимальном и минимальном значениях) восприятие несимметрично в том смысле, что при оценке субъект делает систематическую ошибку в сторону позитивного полюса конструктора.

При этом у разных индивидов систематический сдвиг может быть направлен в разные стороны. В качестве примера можно привести следующий эксперимент [9]. Испытуемым предлагали оценить степень светлоты серого листа бумаги, помещенного между двумя образцами, белым и черным. При том, что правильное значение светлоты располагалось ровно посередине шкалы, эксперимент показал, что оценки группировались вокруг двух значений, одно из которых было сдвинуто в «белую» сторону, а другое — в «черную». Тем самым для каждого отдельного индивида можно говорить о том, что он проводит «измерение» светлости тона смещенной оценкой. О других экспериментах, показавших наличие смещения оценки, см. напр., в [10].

Заметим, однако, что подобное явление наблюдается лишь при оценках, сделанных методом категоризации.

6. Изложенная концепция тернарности предложена В. М. Петровым в работе [11].

Литература

1. *Thurstone L.L.* Psychophysical Analysis // Source: The American Journal of Psychology. Jul., 1927. V. 38, № 3. P. 368–389.
2. *Thurstone L.L.* A Law of Comparative Judgment // Psychological Review. 1927. 34. P. 273–286.

3. *Mosteller F.* Remarks on the method of paired comparisons: I. The least-squares solution assuming equal standard deviations and equal correlations // *Psychometrika* 1951. 16. P. 3–9.
4. *Mosteller F.* Remarks on the method of paired comparisons: II. The effect of an aberrant when equal standard deviations and equal correlations are assumed // *Psychometrika*. 1951. 16. P. 203–206.
5. *Mosteller F.* Remarks on the method of paired comparisons: III. A test of significance for paired comparisons when equal standard deviations and equal correlations are assumed // *Psychometrika*. 1951. 16. P. 207–218.
6. *Torgerson N.S.* Theory and Method of scaling. N.Y.: John Wiley and Sons. 1958.
7. *Гусев А.Н., Измайлов Ч.А., Михалевская М.Б.* Измерение в психологии. М. Смысл. 1987.
8. *Келли Дж.* Теория личности. Психология личных конструктов. СПб., Речь, 2000.
9. *Poulton E.C., Simmonds D.C.V.* Subjective zeros, subjectively equal stimulus spacing, and contraction biases in very first judgments of lightness // *Perception & Psychophysics*. 1985. 37. P. 420–428.
10. *Adams-Webber J., Benjafield J.* The Relation between Lexical Marking and Rating Extremity in Interpersonal Judgments // *Canadian Journal of Behavioral Sciences*. 1973. 5. P. 234–241.
11. *Петров В.М.* Тернарность в мышлении, культуре, искусстве: системно-информационные корни бессознательного // Психология. Журнал Высшей школы экономики. 2008. Т. 5. № 4. С. 3–18.

3.2. Простейшая модель подражательного поведения

Поведенческие гипотезы, играющие особую роль в моделировании социальных процессов, формулируются, как правило, для отдельных индивидов (микроуровень), в то время как цель исследования состоит в получении выводов относительно социума в целом (макроуровень). Для построения самосогласованной замкнутой модели необходимо описать взаимную связь между микро- и макроуровнем.

Такую возможность дают, в частности, гипотезы о подражательном поведении, предполагающие, что индивид наблюдает макровеличины, характеризующие действия других членов социума (например, подсчитывает количество сторонников и противников какой-либо партии), и, используя тот или иной психический «механизм», подражает большинству. Тем самым макровеличины влияют на процессы, происходящие на микроуровне. Обратная связь осуществляется путем агрегирования: «голоса» отдельных индивидов складываются, в результате чего получаются такие макровеличины, как численности сторонников различных партий.

Таким образом, рассматриваемая в настоящем разделе представляет подход, в рамках которого «движущая сила» социальных явлений

анализируется на уровне психических процессов, протекающих «внутри» индивида.

Первые три пункта посвящены описанию собственно этих психических процессов, т. е. микроуровню. Именно, рассматривается нейрологическая структура индивида и уравнения для «микропеременных», т. е. таких величин, как величина возбуждения нейроэлемента, интенсивность стимула и т. д. В четвертом пункте формулируется поведенческая гипотеза о том, что информация, получаемая индивидом из социальной среды, становится для него стимулом к тому или иному действию в результате работы подражательного механизма. Эта поведенческая гипотеза, вкуче с описанием микроуровня, позволяют построить базовую математическую модель подражательного поведения на макроуровне. Эта модель имеет вид обыкновенного интегро-дифференциального уравнения. Пятый пункт посвящен аналитическому исследованию стационарных решений, а шестой — численным экспериментам.

3.2.1. Нейрологическая схема

Рассмотрим индивида, стоящего перед выбором — дихотомической (двухвариантной) альтернативой, предполагающей неоднократное, постоянное принятие им решения о предпочтении одного из двух вариантов действий (предпочтения могут меняться с течением времени). При этом действия индивида доступны для наблюдения другим индивидам — членам рассматриваемой социальной группы, а сам он, наблюдая поведение других, подвержен влиянию с их стороны. Это влияние он воспринимает как стимулы для собственного поведения. Например, если биржевой дилер видит, что его коллеги демонстрируют ожидание повышения индексов, то это может быть стимулом к тому, чтобы и он стал ожидать того же. В свою очередь его поведение является стимулом для других игроков.

Поведение индивида, таким образом, обусловлено его реакцией на эти стимулы, определяемой с учетом его «внутренних» свойств. Так как рассматриваемая альтернатива дихотомическая, то возможных реакций всего две: R_1 и R_2 , причем они являются взаимоисключающими. Стимул (а также величину этого стимула), способствующий появлению реакции R_i , обозначим через S_i , $i = 1, 2$. Наиболее содержательными являются ситуации, когда индивид получает одновременно оба стимула (одни игроки финансового рынка торгуют на повышение, другие — на понижение), имеющие, вообще говоря, не равные и переменные во времени интенсивности.

В предположениях данной модели влияние стимулов на формирование реакции описывается с помощью следующей нейрологической схемы (рис. 1).

Звеньями схемы являются большие группы нейронов, т. е. нейроэлементы (возбуждающие L_i и тормозящие B_i ; смысл этих понятий поясняется ниже), связанные между собой и образующие две взаимо-

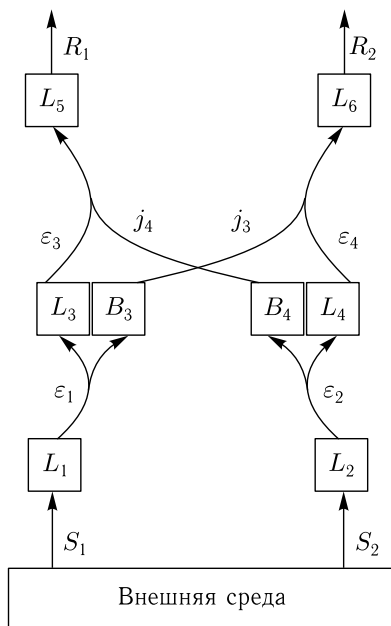


Рис. 1. Нейрологическая схема простейшей модели подражательного поведения

действующих цепи. Одна из них (далее называемая левой) связывает стимул S_1 с реакцией R_1 , другая (правая) — стимул S_2 с реакцией R_2 . Каждая цепь состоит из трех последовательно расположенных возбуждающих нейроэлементов, а также одного тормозящего, спаренного со средним возбуждающим элементом в своей цепи (на рис. 1 спаренные элементы имеют одинаковые номера, например, L_3 и B_3). Тормозящие элементы обеспечивают взаимовлияние цепей: например, начальный стимул S_1 не только способствует появлению конечной реакции R_1 , но и (посредством нейроэлемента B_3) препятствует реакции R_2 . На рис. 1 взаимовлияние цепей соответствует пересечению дуг, исходящих из тормозящих элементов.

Природа нейроэлементов несущественна для наших целей; важен лишь логический принцип их работы, который выглядит следующим образом. На вход каждого нейроэлемента подается определенный стимул (или стимулы). Если он (или разность входящих стимулов) превышает некоторое пороговое значение, то в элементе возникает возбуждение, порождающее на его выходе либо стимул для следующего (в порядке, указанном стрелками на рис. 1) нейроэлемента, либо конечную реакцию (для последнего элемента в цепи). Физическая природа таких величин, как интенсивность стимула или интенсивность возбуждения, в данной модели не существенна. Они могут быть связаны с силой электрического тока между элементами и концентрацией определенных

химических веществ. Для простоты далее будем считать, что все уравнения в данном разделе приведены к безразмерной форме.

Существуют два типа нейроэлементов:

- элемент, возбуждение в котором создает стимул, способствующий возбуждению следующего за ним в нейробиологической схеме элемента, называется *возбуждающим* (величина стимула обозначается через ε_i , а сам возбуждающий элемент — через L_i , $i = 1, \dots, 6$);
- элемент, возбуждение в котором создает стимул, препятствующий возбуждению следующего за ним в нейробиологической схеме элемента, называется *тормозящим* (в этом случае величина стимула обозначается через j_i , а сам тормозящий элемент — через B_i , $i = 3, 4$).

Роль тормозящих нейроэлементов заключается в том, что они обеспечивают взаимодействие левой и правой цепей системы: например, принадлежащий левой цепи элемент B_3 препятствует возбуждению элемента L_6 , принадлежащего правой цепи. В то же время возбуждающие элементы служат распространению возбуждения по «своей» цепи от ее входа до выхода. В целом, нейробиологическая схема описывает конкуренцию левой и правой цепей или, что то же самое, конкуренцию стимулов S_1 и S_2 .

Предваряя подробный анализ (см. следующий пункт) функционирования нейробиологической системы, представленной на рис. 1, дадим здесь краткое описание отношений между ее элементами. Поступающий из внешней среды стимул S_1 действует на низший элемент L_1 левой цепи, вызывая в нем возбуждение, которое производит стимул ε_1 для следующих после L_1 нейроэлементов: возбуждающего элемента L_3 и парного ему тормозящего элемента B_3 . Аналогичным образом стимул S_2 приводит к возбуждению пары элементов L_4 и B_4 .

Далее, элемент L_3 производит стимул ε_3 в следующем за ним элементе L_5 . Однако L_5 получает не только возбуждающий стимул ε_3 , но также тормозящий стимул j_4 , исходящий от элемента B_4 . Таким образом, в нейроэлементе L_5 происходит пересечение левой и правой цепей; при этом, если разность $\varepsilon_3 - j_4$ превышает пороговое значение, то он возбуждается, вызывая реакцию R_1 . Аналогично, в нейроэлементе L_6 при достаточно большой разности $\varepsilon_4 - j_3$ вызывается реакция R_2 . Как будет показано ниже, обе реакции одновременно невозможны, но возможен вариант, при котором не происходит ни одна из них (например, дилер не смог принять решение: готовиться ли к повышению или понижению биржевых индексов).

3.2.2. Уравнения модели на микроуровне

Перейдем к математической формализации процессов, описанных в предыдущем пункте. Рассмотрим сначала принцип функционирования отдельно взятого нейроэлемента цепи. Для определенности будем говорить о возбуждающем элементе L_1 .

Положим, что если он получает стимул интенсивности (которая для простоты пока полагается постоянной) S_1 , превышающей некоторое пороговое значение h , то в нем возникает возбуждение, имеющее величину $E_1 > 0$, пропорциональную величине этого превышения, т. е. разнице между значениями S_1 и h :

$$E_1 = \begin{cases} \alpha(S_1 - h), & S_1 > h, \\ 0, & S_1 \leq h. \end{cases} \quad (1)$$

Здесь α , h — нейрофизиологические постоянные (для простоты будем считать, что они одинаковы для всех нейроэлементов системы). В случае пренебрежимо малого порога h формула (1) переходит в равенство

$$E_1 = \alpha S_1. \quad (2)$$

Реакция нейроэлемента на стимул считается моментальной, так что если интенсивность входящего стимула равна S_1 начиная с момента $t = 0$, то величина возбуждения E_1 дается выражением (1) (или, соответственно, (2)) при $t \geq 0$.

Нейроэлемент L_1 так же мгновенно, без задержек производит возбуждающий стимул интенсивности $\varepsilon_1(t)$ для следующих за ним элементов L_3 , B_3 . Нарастание этой интенсивности является распределенным во времени процессом так, что величина $\varepsilon_1(t)$ изменяется с течением времени от нуля до некоторого максимального значения, пропорционального E_1 (см. рис. 2).

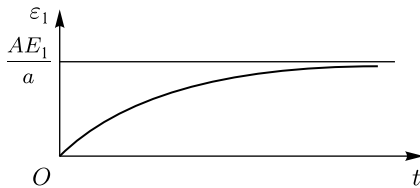


Рис. 2. Процесс нарастания интенсивности $\varepsilon_1(t)$ стимула, производимого нейроэлементом L_1

Наиболее простая модель, описывающая эту стадию процесса, — это дифференциальное уравнение для величины $\varepsilon_1(t)$, имеющее вид

$$\frac{d\varepsilon_1}{dt} = AE_1 - a\varepsilon_1. \quad (3)$$

Здесь постоянные $A > 0$, $a > 0$ отвечают как максимально возможному значению стимула (равному AE_1/a), так и скорости его формирования.

Решая уравнение (3) с начальным условием $\varepsilon_1(0) = 0$, получаем выражение для стимула $\varepsilon_1(t)$, производимого нейроэлементом L_1

и поступающего на вход элементов L_3, B_3 :

$$\varepsilon_1(t) = \frac{AE_1}{a}(1 - e^{-at}).$$

Таким образом (см. рис. 2), функция $\varepsilon_1(t)$ наиболее быстро возрастает при $t = 0$, затем ее рост замедляется с насыщением так, что $\varepsilon_1(t) \rightarrow AE_1/a$ при $t \rightarrow \infty$.

Аналогичным образом действуют и все остальные нейроэлементы, включая и те, что порождают тормозящие стимулы, обозначаемые символом j .

Опираясь на изложенную схему функционирования отдельного элемента системы, рассмотрим теперь формирование конечной реакции на начальные стимулы для простейшей модели подражательного поведения, представленной на рис. 1. При это будем полагать, что оба начальных стимула превосходят пороговое значение: $S_1 > h, S_2 > h$.

Эман 1. На периферический нейроэлемент L_1 поступает стимул S_1 , вызывая в нем возбуждение E_1 :

$$E_1 = \alpha(S_1 - h).$$

Так как сам элемент L_1 является возбуждающим, то на его выходе создается возбуждающий стимул ε_1 , для динамики которого справедливо уравнение (см. (3))

$$\begin{aligned} \frac{d\varepsilon_1}{dt} &= AE_1 - a\varepsilon_1, \quad t > 0, \\ \varepsilon_1(0) &= 0. \end{aligned}$$

Аналогично, в правой цепи имеем для нейроэлемента L_2 такую же динамику возбуждающего стимула ε_2 :

$$\begin{aligned} E_2 &= \alpha(S_2 - h), \\ \frac{d\varepsilon_2}{dt} &= AE_2 - a\varepsilon_2, \quad t > 0, \\ \varepsilon_2(0) &= 0. \end{aligned}$$

Эман 2. Так как ε_1 является возбуждающим стимулом для L_3 и B_3 , то в каждом из них возбуждение возникает лишь, если (и после того, как) величина ε_1 превысит порог h . Таким образом (см. формулу (1)),

$$E_3 = \begin{cases} \alpha(\varepsilon_1 - h), & \varepsilon_1 > h, \\ 0, & \varepsilon_1 \leq h. \end{cases} \quad (4)$$

Следовательно, если максимальное значение стимула ε_1 , равное $AE_1/a = A\alpha S_1/a$, не превышает порог h , то $E_3 = 0$ для любого $t \geq 0$ (т. е. ни в одном из нейроэлементов L_3, B_3 возбуждение не возникает). Если же

$$\frac{A\alpha S_1}{a} > h, \quad (5)$$

то в момент времени $t = t_1$ такой, что $\varepsilon_1(t_1) = h$, в элементах L_3 и B_3 инициируется возбуждение E_3 , стремящееся при $t \rightarrow \infty$ к значению

$$\lim_{t \rightarrow \infty} E_3(t) = \alpha \left(\frac{A\alpha S_1}{a} - h \right).$$

При этом на выходе возбуждающего нейроэлемента L_3 возникает возбуждающий стимул ε_3 , удовлетворяющий уравнению

$$\begin{aligned} \frac{d\varepsilon_3}{dt} &= AE_3 - a\varepsilon_3, & t > t_1, \\ \varepsilon_3(t_1) &= 0, \end{aligned} \quad (6)$$

а на выходе тормозящего элемента B_3 — тормозящий стимул j_3 , имеющий ту же динамику:

$$\begin{aligned} \frac{dj_3}{dt} &= AE_3 - aj_3, & t > t_1, \\ j_3(t_1) &= 0. \end{aligned} \quad (7)$$

При этом, поскольку уравнения (6), (7) содержат одинаковые параметры, то $j_3(t) = \varepsilon_3(t)$ при всех $t \geq 0$.

Аналогично, в правой цепи (см. рис. 1) при выполнении неравенства

$$\frac{A\alpha S_2}{a} > h \quad (8)$$

начиная с момента времени t_2 такого, что $\varepsilon_2(t_2) = h$, возбуждающий стимула ε_2 вызывает в элементах L_4 и B_4 возбуждение E_4 :

$$E_4 = \begin{cases} \alpha(\varepsilon_2 - h), & \varepsilon_2 > h, \\ 0, & \varepsilon_2 \leq h. \end{cases}$$

Вследствие этого на выходе возбуждающего нейроэлемента L_4 возникает стимул ε_4 , имеющий динамику

$$\frac{d\varepsilon_4}{dt} = AE_4 - a\varepsilon_4, \quad t > t_2, \quad \varepsilon_4(t_2) = 0,$$

а на выходе тормозящего нейроэлемента B_4 — стимул j_4 ; для него имеем

$$\frac{dj_4}{dt} = AE_4 - aj_4, \quad t > t_2, \quad j_4(t_2) = 0. \quad (9)$$

При этом $j_4(t) = \varepsilon_4(t)$ при всех $t \geq 0$.

Всюду далее будем считать, что оба неравенства (5), (8) выполняются, т.е. импульс возбуждения идет как по левой, так и по правой цепи, хотя и не обязательно синхронно: если $S_1 > S_2$, то $t_1 < t_2$, а если $S_1 < S_2$, то $t_1 > t_2$ (синхронность имеет место лишь при $S_1 = S_2$).

Непосредственно из уравнений для стимулов, исходящих от элементов L_3, B_3, L_4, B_4 , находим, что их асимптотические при $t \rightarrow \infty$ значения равны

$$\begin{aligned}\lim_{t \rightarrow \infty} \varepsilon_3(t) &= \lim_{t \rightarrow \infty} j_3(t) = \frac{A}{a} \lim_{t \rightarrow \infty} E_3(t) = \frac{A\alpha}{a} \left(\frac{A\alpha S_1}{a} - h \right), \\ \lim_{t \rightarrow \infty} \varepsilon_4(t) &= \lim_{t \rightarrow \infty} j_4(t) = \frac{A}{a} \lim_{t \rightarrow \infty} E_4(t) = \frac{A\alpha}{a} \left(\frac{A\alpha S_2}{a} - h \right),\end{aligned}$$

причем при $S_1 = S_2$ эти конечные (асимптотические) значения равны. Соответственно при импульсе $S_1 > S_2$ по левой проходит более интенсивный импульс, чем по правой, а если $S_1 < S_2$, то — наоборот.

Заман 3. Перейдем теперь к анализу заключительных (выходных) элементов L_5 и L_6 , в которых импульсы левой и правой цепей пересекаются. Суммарным стимулом для нейроэлемента L_5 служит разность $\varepsilon_3 - j_4$ поступающих на его вход возбуждающего и тормозящего стимулов (здесь и далее считается, что если на вход нейроэлемента поступают два стимула, то они действуют аддитивно). Если она превышает пороговое значение h , то в элементе возникает возбуждение

$$E_5 = \begin{cases} \alpha(\varepsilon_3 - j_4 - h), & \varepsilon_3 - j_4 > h, \\ 0, & \varepsilon_3 - j_4 \leq h. \end{cases}$$

Очевидно, что если $\lim_{t \rightarrow \infty} (\varepsilon_3(t) - j_4(t)) \leq h$, то $E_5(t) = 0$, следовательно, $R_1(t) = 0$ при всех $t \geq 0$. Если же $\lim_{t \rightarrow \infty} (\varepsilon_3(t) - j_4(t)) > h$, то начиная с некоторого момента времени t_3 элемент L_5 производит конечную реакцию R_1 , динамика которой описывается задачей Коши

$$\begin{aligned}\frac{dR_1}{dt} &= AE_5 - aR_1, \quad t > t_3, \\ R_1(t_3) &= 0.\end{aligned}$$

Проводя аналогичные рассуждения для нейроэлемента L_6 , нетрудно получить уравнение для реакции R_2 ; для ее возникновения необходимо, чтобы стимул $j_4 - \varepsilon_3$ превысил пороговое значение h :

$$\begin{aligned}\frac{dR_2}{dt} &= AE_6 - aR_2, \quad t > t_4. \\ R_1(t_4) &= 0,\end{aligned}$$

где

$$E_6 = \begin{cases} \alpha(\varepsilon_4 - j_3 - h), & \varepsilon_4 - j_3 > h, \\ 0, & \varepsilon_4 - j_3 \leq h. \end{cases}$$

Так как $j_3(t) = \varepsilon_3(t)$ и $j_4(t) = \varepsilon_4(t)$, то неравенства $\lim_{t \rightarrow \infty} (\varepsilon_3(t) - j_4(t)) > h$, $\lim_{t \rightarrow \infty} (\varepsilon_4(t) - j_3(t)) > h$ не могут выполняться, одновременно а значит, не более чем одна из величин E_5, E_6 может быть отлична от нуля. Таким образом, в соответствии с данным механизмом при любых значениях параметров системы на выходе нейрологической системы реализуется не более чем одна из реакций R_1, R_2 .

Заметим, что в случае пренебрежимо малого порога (т. е. при $h = 0$) ни одна из реакций не будет реализована только при условии $S_1 = S_2$. Если же $S_1 > S_2$, то $\varepsilon_1 > \varepsilon_2$, вслед за этим $\varepsilon_3 > j_4$ и $j_3 > \varepsilon_4$, следовательно, реакция R_1 осуществится, а реакция R_2 нет. И наоборот, если $S_1 < S_2$, то осуществится реакция R_2 .

Таким образом, представленная на рис. 1 нейрологическая схема описывает формирование реакции индивида при поступлении двух противоположно направленных стимулов. Показано, что решающую роль играет соотношение интенсивностей этих стимулов: менее мощный из них подавляется более мощным.

3.2.3. Учет внешних и внутренних стимулов

В предыдущих пунктах природа и структура стимулов S_1 и S_2 не обсуждались, при этом предполагалось, что они являются внешними по отношению к индивиду, поскольку действуют не на внутренние, а на периферические (пограничные) нейроэлементы L_1 и L_2 .

Введем теперь в рассмотрение также и стимулы, порождаемые «внутри индивида». Для наших целей несущественно их происхождение; например, разумно считать, что оно связано с предыдущим опытом индивида и его личностными характеристиками. Действуя внутри левой и правой цепей, эти стимулы складываются со стимулами, создаваемыми периферическими элементами и образуют совместно с ними суммарную картину распространения импульса.

На рис. 3 внутренние стимулы обозначены через w_1 и w_2 , индивид как их источник (действующий не на периферические элементы, а «внутри» сети) — через P . Таким образом, внутренние нейроэлементы L_3 и B_3 получают на вход сумму $w_1 + \varepsilon_1$ стимулов, поступающих соответственно от этого источника и от периферического элемента L_1 . Аналогично, нейроэлементы L_4 и B_4 получают в качестве входящего стимула сумму $w_2 + \varepsilon_2$.

В следующем пункте величины S_1 и S_2 будут увязаны с социальной средой, т. е. с окружающим индивида социумом. При этом величины w_1 и w_2 будут считаться постоянными, а S_1 и S_2 — изменяющимися с течением времени (выше эти величины для простоты предполагались постоянными; данное ограничение не является существенным, оно было принято для того, чтобы иметь возможность получить решение уравнения (3) в явном виде). Таким образом, данная схема позволяет исследовать совместное влияние на формирование реакции со стороны как личностных, внутренних (w_1 , w_2), так и социальных, внешних (S_1 , S_2) по отношению к индивиду факторов.

Введем величину

$$\varphi = w_1 - w_2$$

характеризующую внутреннюю склонность индивида к выбору той или иной реакции. Если $\varphi > 0$, то при отсутствии внешних стимулов индивид чаще (в детерминированном случае — всегда) демонстрирует реакцию R_1 , если $\varphi < 0$, то реакцию R_2 .

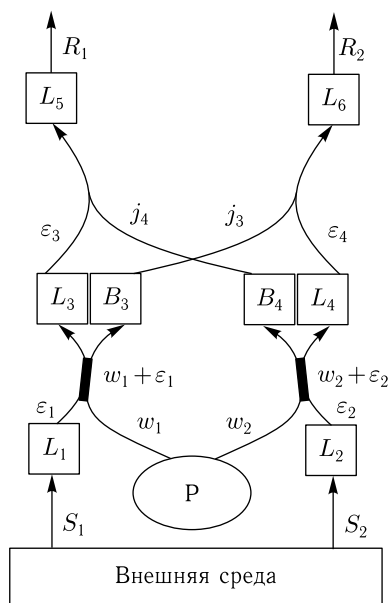


Рис. 3. Нейрологическая схема с внутренними стимулами w_1 , w_2 и внешними стимулами S_1 , S_2

Каждый индивид характеризуется своим значением φ , и считается, что составляющие социум N_0 индивидов распределены с плотностью $N(\varphi)$ так, что $\int_{\varphi_1}^{\varphi_2} N(\varphi) d\varphi$ — количество индивидов, для которых значения величины φ находятся в диапазоне от φ_1 до φ_2 . При этом справедливо очевидное равенство $\int_{-\infty}^{\infty} N(\varphi) d\varphi = N_0$.

После этих предварительных рассмотрений перейдем непосредственно к построению и анализу модели.

3.2.4. Базовая математическая модель

Выше отмечалось, что стимулы S_1 и S_2 являются внешними по отношению к индивиду — они поступают к нему из социальной среды. *Поведенческая гипотеза, принятая в настоящей модели, предполагает, что величина этих стимулов определяется на основе подражательного механизма.* Подражание означает, что чем большее количество членов группы демонстрирует реакцию, например R_1 , тем выше для данного индивида стимул S_1 к этой реакции.

Рассмотрим данное положение более подробно. Каждый член социума наблюдает действия остальных $N_0 - 1$ индивидов, определяя тем самым текущие численности $X(t)$, $Y(t)$ индивидов, демонстрирующих соответственно реакции R_1 и R_2 . Поведенческая гипотеза может быть конкретизирована с помощью простейшей линейной зависимости;

именно, величины стимулов S_1 , S_2 принимаются пропорциональными этим численностям:

$$S_1(t) = CX(t), \quad S_2(t) = CY(t).$$

Введенная таким образом постоянная C характеризует соотношение между восприимчивостью индивида к социальным и личностным факторам. Чем больше величина C , тем более значимы внешние стимулы S_1 , S_2 по сравнению с внутренними w_1 , w_2 .

Для величины возбуждения в нейроэлементе L_1 имеем выражение

$$E_1 = \alpha(CX - h),$$

а с учетом (3) получаем уравнение

$$\frac{d\varepsilon_1}{dt} = AE_1 - a\varepsilon_1 = A\alpha(CX - h) - a\varepsilon_1. \quad (10)$$

Подобным образом для правой цепи справедливо

$$\frac{d\varepsilon_2}{dt} = A\alpha(CY - h) - a\varepsilon_2. \quad (11)$$

Аналогично тому, как выше была введена величина $\varphi = w_1 - w_2$, введем переменную

$$\psi(t) = \varepsilon_1(t) - \varepsilon_2(t).$$

Значение $\psi(t)$ предполагается одинаковым для всех индивидов. Оно имеет смысл определяемого социальной средой сдвига стимулов в сторону реакции R_1 для нейроэлементов L_3 , B_3 , L_4 , B_4 .

Получим уравнение для ψ . Вычитая (11) из (10), имеем

$$\frac{d\psi}{dt} = A\alpha C(X - Y) - a\psi. \quad (12)$$

Далее для простоты ограничимся случаем пренебрежимо малого порога ($h = 0$). Это условие обеспечивает, в частности выполнение равенства $N_0 = X + Y$, т. е. отсутствие индивидов, не демонстрирующих ни одну из реакций.

Наша ближайшая цель заключается в том, чтобы исключить величины X и Y из (12) и тем самым получить окончательное уравнение для функции $\psi(t)$.

Наблюдаемый данным индивидом «другой индивид» демонстрирует первую реакцию, если для него $\varepsilon_3 > j_4$. Поскольку $h = 0$, то (см. (4), (8))

$$E_3 = \alpha\varepsilon_1, \quad E_4 = \alpha\varepsilon_2.$$

Задачи Коши (6), (9) для ε_3 и j_4 принимают вид

$$\begin{aligned} \frac{d\varepsilon_3}{dt} &= AE_3 - a\varepsilon_3 = A\alpha(w_1 + \varepsilon_1) - a\varepsilon_3, & \varepsilon_3(0) &= 0, \\ \frac{dj_4}{dt} &= AE_4 - aj_4 = A\alpha(w_2 + \varepsilon_2) - aj_4, & j_4(0) &= 0. \end{aligned}$$

Из этой пары уравнений следует, что индивид демонстрирует реакцию R_1 , если $w_1 + \varepsilon_1 > w_2 + \varepsilon_2$, т. е. если $\varphi + \psi > 0$ или $-\psi < \varphi < \infty$. Общее количество таких индивидов в момент времени t равно

$$X(t) = \int_{-\psi(t)}^{\infty} N(\varphi) d\varphi; \quad (13)$$

тем самым величина X выражается через переменную ψ .

Проводя аналогичное рассуждение для численности индивидов, демонстрирующих реакцию R_2 , имеем

$$Y(t) = \int_{-\infty}^{-\psi(t)} N(\varphi) d\varphi. \quad (14)$$

С учетом того, что $N_0 = X + Y$, соотношение (12) переписывается в виде

$$\frac{d\psi}{dt} = A\alpha C(2X - N_0) - a\psi. \quad (15)$$

Окончательно из (13) и (15) получаем искомое уравнение для функции $\psi(t)$:

$$\frac{d\psi}{dt} = A\alpha C \left(2 \int_{-\psi}^{\infty} N(\varphi) d\varphi - N_0 \right) - a\psi. \quad (16)$$

Начальное условие $\psi(0)$, необходимое для однозначного определения $\psi(t)$, находится из распределения индивидов между реакциями R_1 и R_2 в момент $t = 0$ и имеет вид

$$X(0) = \int_{-\psi(0)}^{\infty} N(\varphi) d\varphi \quad (17)$$

или, что то же самое, $Y(0) = \int_{-\infty}^{-\psi(0)} N(\varphi) d\varphi$.

Итак, функция $\psi(t)$ находится как решение начальной задачи (16), (17) для интегро-дифференциального уравнения (16). Для получения содержательных выводов относительно исследуемой системы знания динамики $\psi(t)$, как правило, оказывается достаточно, однако при необходимости из выражений (13), (14) могут быть получены также численности $X(t)$ и $Y(t)$ индивидов, демонстрирующих соответственно реакции R_1 и R_2 (а следовательно, и доли таких индивидов $X(t)/[X(t) + Y(t)]$, $Y(t)/[X(t) + Y(t)]$).

3.2.5. Стационарные состояния базовой модели

Чтобы изучить возможные варианты поведения решения задачи (16), (17) при $t \rightarrow \infty$, проанализируем стационарные решения уравнения (16).

Введем обозначение

$$f(\psi) = 2 \int_{-\psi}^{\infty} N(\varphi) d\varphi - N_0 = 2 \int_{-\infty}^{\psi} N(-\varphi) d\varphi - N_0.$$

Эта функция обладает следующими легко проверяемыми свойствами.

1. Предельные значения $f(\psi)$ равны

$$\lim_{\psi \rightarrow -\infty} f(\psi) = -N_0, \quad \lim_{\psi \rightarrow \infty} f(\psi) = N_0,$$

2. Производная $f'(\psi) \geq 0$, причем если $N(\varphi) > 0$ при $-\infty < \varphi < +\infty$, то $f'(\psi) > 0$ при $-\infty < \psi < +\infty$.

3. Если μ — медиана распределения $N(\varphi)$, т. е. значение φ такое, что

$$\int_{-\infty}^{\mu} N(\varphi) d\varphi = \int_{\mu}^{\infty} N(\varphi) d\varphi = N_0/2,$$

то $f(\mu) = 0$.

Введенное обозначение позволяет записать уравнение (16) для функции $\psi(t)$ в виде

$$\frac{d\psi}{dt} = A\alpha C f(\psi) - a\psi. \quad (18)$$

На рис. 4, а, б, в, г проведен геометрический анализ точек равновесия уравнения (18) в наиболее простом случае, когда $N(\varphi)$ имеет ровно одну точку максимума. В каждой из четырех частей рисунка вверху показан график функции $N(-\varphi)$, в середине — взаимное расположение графиков функций $A\alpha C f(\psi)$ и $a\psi$, внизу — знаки производной $d\psi/dt$ в соответствующих интервалах оси ψ . Очевидно, что стационарная точка является асимптотически устойчивой, если в некоторой ее левой полукрестности выполняется неравенство $d\psi/dt > 0$, т. е. $A\alpha C f(\psi) < a\psi$, а в правой полукрестности — $d\psi/dt < 0$, т. е. $A\alpha C f(\psi) > a\psi$.

В случае, изображенном на рис. 4, а, графики функций $A\alpha C f(\psi)$ и $a\psi$ имеют три точки пересечения, разбивающие ось ψ на четыре интервала. Проанализировав знаки производной $d\psi/dt$ на этих интервалах, получаем, что две крайние точки покоя, ψ_1 и ψ_3 , асимптотически устойчивы, а средняя, $\psi_2 = 0$, неустойчива и разбивает ось на две области влияния соответствующих устойчивых стационарных состояний. Таким образом, если начальное значение $\psi(0)$ положительно, то с течением времени решение уравнения (16) стабилизируется к положительному стационарному состоянию. Поскольку (в соответствии с определением функции $\psi(t)$) она представляет собой разность стимулов, $\psi(t) = \varepsilon_1(t) - \varepsilon_2(t)$, то это означает, что с течением времени формируется стационарное состояние системы, в котором для каждого

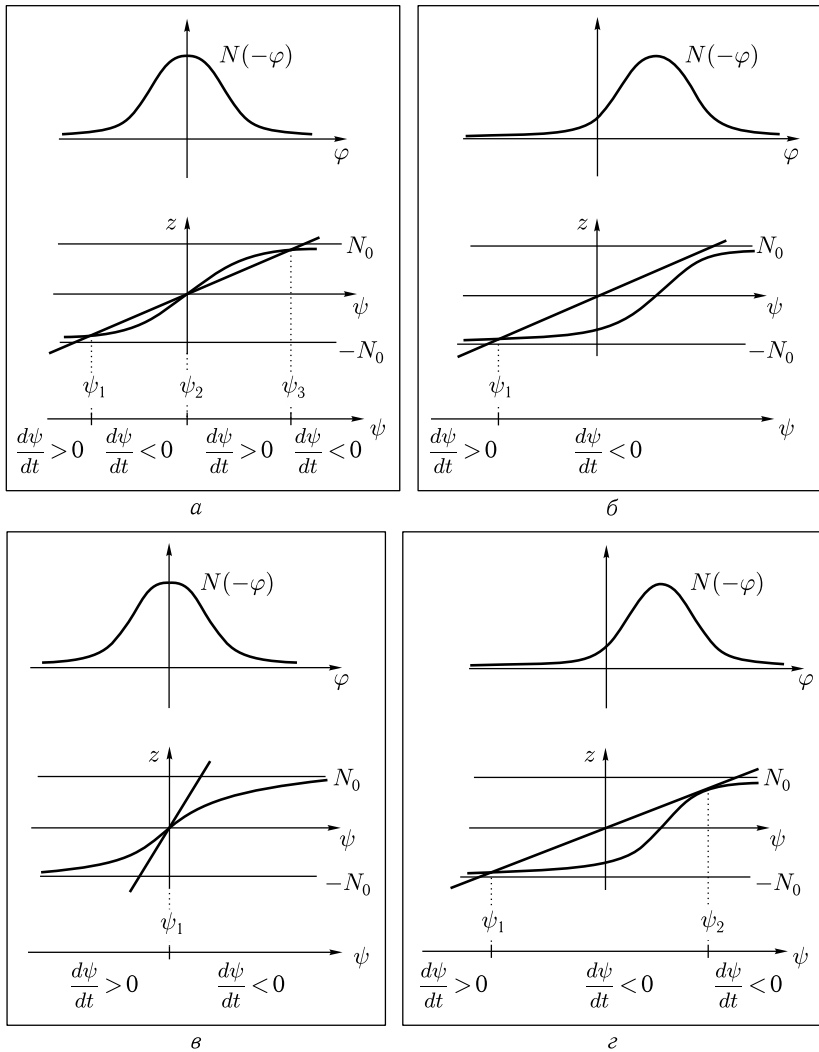


Рис. 4. Анализ стационарных решений уравнения (18). Графики а-г соответствуют различным случаям взаимного расположения прямой $z = a\psi$ и функции $z = A\alpha C f(\psi)$: три точки пересечения (а), одна точка пересечения (б, в), одна точка пересечения и одна точка касания (г)

индивида определяемый социальной средой стимул к реакции R_1 сильнее, чем стимул к реакции R_2 (другими словами, сдвиг внешних стимулов направлен в сторону R_1). Если же начальное значение $\psi(0)$ отрицательно, то реализуется отрицательное стационарное состояние, т. е. сдвиг внешних стимулов в сторону R_2

В случае одной точки пересечения (рис. 4, б и в) единственная точка пересечения ψ_1 является асимптотически устойчивой, стабилизация к ней происходит из любого начального условия $\psi(0)$.

Если же графики функций $A\alpha C f(\psi)$ и $a\psi$ имеют одну точку пересечения (при $\psi = \psi_1$ — см. рис. 4, з) и одну точку касания (при $\psi = \psi_2$), то стационарное состояние $\psi = \psi_1$ является асимптотически устойчивым, а состояние $\psi = \psi_2$ — неустойчивым (причем производная $d\psi/dt$ не меняет знак при переходе через нее, так называемая *точка покоя типа шунт* см. комментарий 4 в конце раздела).

3.2.6. Численный эксперимент с нестационарной моделью

Проведенный выше анализ стационарных состояний позволил установить связь между начальным условием и поведением решения при $t \rightarrow \infty$, другими словами, выделить области начальных значений $\psi(0)$, соответствующие различным конечным значениям функции $\psi(t)$.

Чтобы изучить процесс перехода от начального условия к стационарному состоянию, проведем серию численных экспериментов с моделью. Для этого предварительно приведем уравнение (16) к однопараметрическому виду: введем *системный параметр* $\gamma = \frac{a}{A\alpha C}$ (значением которого, как показано ниже, определяются свойства системы) и обозначим $t' = A\alpha C t$. Получим

$$\frac{d\psi}{dt'} = 2 \int_{-\psi}^{\infty} N(\varphi) d\varphi - N_0 - \gamma\psi. \quad (19)$$

Единицу измерения численности индивидов выберем так, что $N_0 = 1$. Конкретизируя входные данные модели для численного эксперимента, примем в качестве плотности $N(\varphi)$ функцию стандартного нормального распределения

$$N(\varphi) = \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{\varphi^2}{2}\right], \quad (20)$$

а начальное условие возьмем $\psi(0) = 1$. В соответствии с формулой (17) данное начальное условие означает, что

$$X(0) = \int_{-1}^{\infty} N(\varphi) d\varphi = \frac{1}{2} + \int_{-1}^0 N(\varphi) d\varphi > \frac{1}{2}, \quad (21)$$

$$Y(0) = 1 - X(0) < \frac{1}{2},$$

т. е. в момент времени $t = 0$ более половины индивидов демонстрируют реакцию R_1 .

Графики решений для различных значений параметра γ представлены на рис. 5.

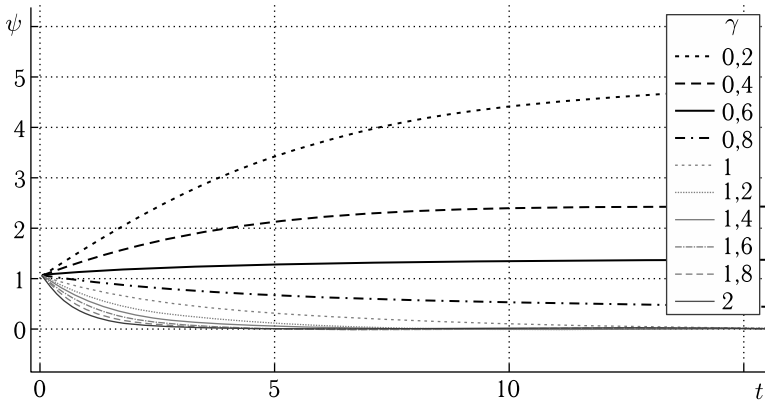


Рис. 5. Результаты численного эксперимента: графики функции $\psi(t')$ при различных значениях γ

Все графики изображают монотонные функции, при достаточно больших значениях γ они убывают до нуля так, что $\lim_{t \rightarrow \infty} \psi(t) = 0$ (что соответствует случаю асимптотической устойчивости стационарного решения $\psi_2 = 0$, изображенному на рис. 4, в), при малых приближаются к положительным стационарным значениям, зависящим от γ (что соответствует случаю асимптотической устойчивости ненулевых стационарных решений ψ_1 и ψ_3 , изображенному на рис. 4, а).

С содержательной точки зрения это означает, что от значения системного параметра γ зависит, какое (с течением времени) установится соотношение между численностями индивидов, демонстрирующих реакции R_1 и R_2 . Пусть, например, имеет место первый случай рис. 4, в: $\lim_{t \rightarrow \infty} \psi(t) = 0$. Тогда в соответствии с формулой (13) для количества $X(t)$ индивидов, выражающих реакцию R_1 , имеем

$$\lim_{t \rightarrow \infty} X(t) = \lim_{t \rightarrow \infty} \int_{-\psi(t)}^{\infty} N(\varphi) d\varphi = \int_0^{\infty} N(\varphi) d\varphi = \frac{1}{2}.$$

Аналогично, из (14) получим, что $\lim_{t \rightarrow \infty} Y(t) = \frac{1}{2}$. Таким образом, если значение γ достаточно велико, то с течением времени установится равенство — реакции R_1 и R_2 будут выражать по половине от общего количества индивидов.

Если же имеет место случай рис. 4, а, т. е. $\lim_{t \rightarrow \infty} \psi(t) = \psi_3 > 0$, то из (13) получаем

$$\lim_{t \rightarrow \infty} X(t) = \lim_{t \rightarrow \infty} \int_{-\psi(t)}^{\infty} N(\varphi) d\varphi = \int_{-\psi_3}^{\infty} N(\varphi) d\varphi = \frac{1}{2} + \int_{-\psi_3}^0 N(\varphi) d\varphi > \frac{1}{2}.$$

Таким образом, если значение γ достаточно мало (и в начальный момент сторонники реакции R_1 находились в большинстве — см. формулу (21)), то с течением времени установится стационарное состояние, при котором более половины членов группы будут демонстрировать реакцию R_1 .

Замечание 1. Если начальное значение функции $\psi(t)$ отрицательно: $\psi(0) < 0$ (что соответствует соотношению $X(0) < Y(0)$), то с течением времени установится стационарное состояние, при котором более половины членов группы будут демонстрировать реакцию R_2 , т. е. $\lim_{t \rightarrow \infty} \psi(t) = \psi_1 < 0$ и $\lim_{t \rightarrow \infty} X(t) < \frac{1}{2}$.

Резюмируем промежуточные результаты настоящего пункта. Если значение параметра γ достаточно велико (мы будем условно называть этот случай режимом рис. 4, в), то из любых начальных условий в группе возникает стационарное состояние, при котором одна половина индивидов демонстрирует реакцию R_1 , а другая половина — реакцию R_2 . Если же значение γ достаточно мало (режим рис. 4, а), то формируется стационарное состояние, в котором одну из реакций выражает более половины членов группы, другую — менее половины. При этом от начальных условий зависит, какая из реакций будет иметь большинство сторонников, а именно та же, которая была более популярной в начальный момент времени.

Для того чтобы найти критическое значение γ_{cr} , разделяющее данные режимы, подставим используемую в данном численном эксперименте конкретную функцию $N(\varphi)$ вида (20) в уравнение (19), получим

$$\frac{d\psi}{dt} = \frac{2}{\sqrt{2\pi}} \int_{-\psi}^{\infty} \exp\left[-\frac{\varphi^2}{2}\right] d\varphi - N_0 - \gamma\psi. \quad (22)$$

Режим (см. рис. 4, а) возникает, если угловой коэффициент прямой $z = \gamma\psi$ меньше, чем у касательной к кривой $z = F(\psi) = 2 \int_{-\psi}^{\infty} N(\varphi) d\varphi - N_0$ в начале координат, т. е. $\gamma < F'(0)$. Непосредственным дифференцированием получаем, что $F'(\psi) = 2N(-\psi)$. Воспользовавшись выражением (20) для функции $N(\varphi)$ получаем, что режим рис. 4, а имеет место, если $\gamma < \gamma_{cr} = 2/\sqrt{2\pi} \approx 0,7979$, а режим рис. 4, в — если $\gamma \geq \gamma_{cr}$.

Итак, если значение системного параметра γ равно или превышает $\gamma_{cr} = 2/\sqrt{2\pi}$, то с течением времени $\psi(t) \rightarrow 0$ и в системе устанавливается равное соотношение между численностями индивидов, выражающих первую и вторую реакции, т. е. $\lim_{t \rightarrow \infty} X(t) = \lim_{t \rightarrow \infty} Y(t)$, несмотря на изначальное доминирование тех, кто выражает первую реакцию. Следовательно, эффект подражания является в таком случае менее значимым, чем собственные предпочтения индивидов, распределение которых описывается функцией $N(\varphi)$.

Обратная ситуация наблюдается при значениях $\gamma < \gamma_{cr}$. Системный параметр γ зависит от нейрологических параметров a , A , α , а также от величины C , описывающей влияние макропеременных X , Y на нейрорепериментальные S_1 , S_2 . Более высокие значения C соответствуют более высокой роли подражания по сравнению с внутренними предпочтениями.

Подобные численные эксперименты могут быть проведены и для других ситуаций.

Таким образом, представленная математическая модель описывает нейрологический механизм подражания индивидов на основе гипотезы о взаимной связи между процессами, происходящими «внутри индивида» и в социальной группе, включающей в себя этого индивида и взаимодействующей с ним. Анализ модели показывает, что в зависимости от параметров системы она имеет одно или несколько устойчивых конечных состояний (аттракторов). Если аттракторов несколько, то от начального условия зависит: какой из них будет реализован в каждом конкретном случае.

Упражнения

1. Решить уравнение (16) для треугольного распределения $N(\varphi) = b - k|\varphi|$. Проанализировать, при каких значениях параметров имеют место случаи, представленные на рис. 4, a и b .
2. Выполнить то же задание для равномерного распределения

$$N(\varphi) = \begin{cases} \frac{1}{b-a}, & a < \varphi < b, \\ 0, & \varphi \leq a, \varphi \geq b. \end{cases}$$

Комментарии и библиографические замечания

1. Базовая модель подражательного поведения, основанная на описании нейроэлементного механизма, была предложена американским ученым, выходцем из России Н. Рашевским ([1, 2], см. также исторический обзор [3]). О современном развитии модели см. [4, 5]. Динамическая модель в виде интегро-дифференциального уравнения сформулирована в работе [5].

2. Принимаемые индивидом решения должны быть не только видны для других, но и восприниматься ими без искажений. Другими словами, модель предполагает, что считывание чужих реакций и подсчет количества индивидов происходят без ошибок.

3. Одно из возможных направлений обобщения модели — описание более сложного механизма подражания. Так, изложенный выше базовый вариант модели предполагал, что подражание основано на наблюдаемых индивидом численностях членов социальной группы, демонстрирующих реакции R_1 и R_2 . Более сложный механизм предлага-

ет концепция социально-психологического потенциала, в соответствии с которой подражательное поведение содержит так называемую *ажитажную составляющую*:

$$S_1 = CX + K \frac{dX}{dt}, \quad S_2 = CY + K \frac{dY}{dt},$$

т. е. стимулом для индивида является не только численность выражающих, например реакцию R_1 , но и скорость роста этой численности.

4. О классификации точек покоя для динамических систем первого порядка см., напр., [6]. Простейший пример точки покоя типа «шунт» дается уравнением $y' = x^2$. Так как в этом примере $y' > 0$ при $x \neq 0$, то точка покоя $x_0 = 0$ является «притягивающей» при $x < 0$ и «отталкивающей» при $x > 0$; это и означает, что она является «шунтом».

Литература

1. *Rashevsky N.* Outline of a Physico-mathematical Theory of Excitation and Inhibition // *Protoplasma*, 1933.
2. *Рашевский Н.* Две модели: подражательное поведение и распределение статуса // Математические методы в современной буржуазной социологии. Сборник статей. Под ред. Г. В. Осипова. М.: Прогресс. 1966. С. 175–197.
3. *Cull P.* The mathematical biophysics of Nicolas Rashevsky // *BioSystems* 2007. Vol. 88. P. 178–184.
4. *Шведовский В.А., Шведовский О.В., Цаплин Н.В.* Математическая модель микродинамики личностных изменений // Математическое моделирование социальных процессов / Под ред. А. П. Михайлова, МАКС Пресс Москва, том 11. С. 98–131.
5. *Петров А.П., Цаплин Н.А.* Вычислительные эксперименты с моделью подражательного поведения // Математическое моделирование социальных процессов. Вып. 14. Под ред. А. П. Михайлова. 2012. С. 81–91.
6. *Эрроусмит Д., Плейс К.* Обыкновенные дифференциальные уравнения. Качественная теория с приложениями. Волгоград, изд-во Платон. 1997.

3.3. Расширенная модель подражательного поведения

Более сложные, чем рассмотренные выше, формы подражательного поведения, позволяет обсуждать концепция социально-политического потенциала (СПП). Как и в модели Рашевского, рассматривается группа индивидов, осуществляющих дихотомический выбор между двумя взаимоисключающими позициями, которые здесь обозначаются через «X» и «Y». Речь идет не о разовом, а о непрерывном выборе: индивиды постоянно совершают поступки в пользу той или иной позиции, причем эти поступки наблюдаются и принимаются

во внимание другими членами группы, являясь для них стимулами к собственным поступкам. При этом один и тот же индивид может почти одновременно совершать поступки, выражающие как поддержку позиции «Х», так и поддержку позиции «У». В этой концепции индивиды рассматриваются как *амбивалентные личности*, в сознании которых сосуществуют противоположные чувства и побуждения и поведение которых может быть внешне противоречивым. Так, например, житель мятежного региона может отвести своего ребенка в федеральную (государственную) школу, оттуда пойти за продуктами к местному торговцу (основные жизненные приоритеты которого лежат не в политической, а экономической сфере) и затем — к общественному или религиозному деятелю, известному оппозиционными настроениями. Первый поступок может быть понят окружающими как признак его лояльности, третий — как признак оппозиционности.

Если в социуме существуют механизмы, синхронизирующие во времени действия составляющих его индивидов или, по крайней мере, значительной их доли, то становится возможным судить о поведении социума с помощью понятия так называемой *типической личности*, понимаемой как *амбивалентная социальная единица, совершающая, вообще говоря, поступки как в пользу позиции «Х», так и в пользу позиции «У» и репрезентирующая осредненное по всем членам этого социума поведение*. Поясним, что для целей данного исследования понятие средней типической личности плодотворно лишь при условии, если в электорате работают механизмы, синхронизирующие во времени действия его членов (или, хотя бы, значительного их числа). Только в этом случае по поведению средней типической личности можно судить, как это и делается ниже, в главе 5, о поведении электората в целом.

3.3.1. Понятие социально-политического потенциала (СПП)

Рассмотрим динамический, т. е. развивающийся во времени процесс выбора некоторой социальной общностью (этнической, конфессиональной, профессиональной и т. д.) позиции (линии поведения, стратегии) по отношению к значимому для нее вопросу (политическому, экономическому, этическому и т. д.). В дальнейшем для определенности и ради краткости наряду со словами «социальная общность» будем употреблять слово «электорат».

Количество дееспособных членов электората $N(t) > 0$ зависит, вообще говоря, от времени в силу как естественных причин (например, вступление новых членов в дееспособный возраст или выбытие из него), так и причин «внешнего» характера (например, благодаря наличию миграционных потоков). Очевидно, что по отношению к выбираемой позиции электорат состоит из трех основных категорий:

$$N(t) = Z(t) + X(t) + Y(t), \quad t \geq t_0, \quad (1)$$

В балансном соотношении (1) через $Z(t) \geq 0$ обозначено число неучастников процесса (нейтралы), а через $X(t) \geq 0$ и $Y(t) \geq 0$ — число сторонников позиции «X» и «Y» соответственно. Другими словами, в момент времени t поступки в поддержку позиции «X» совершают $X(t)$ индивидов, а в поддержку позиции «Y» — $Y(t)$ индивидов. Численность категорий Z, X, Y в общем случае изменяется со временем не только из-за естественных и внешних факторов, но и благодаря «внутренним» причинам — переходам членов социальной общности из одной категории в другую. Более детально структура числа неучастников и сторонников позиций «X» и «Y» описывается соотношениями

$$\begin{aligned} Z(t) &= Z_0(t) + z(t), \\ X(t) &= X_0(t) + x(t), \\ Y(t) &= Y_0(t) + y(t), \end{aligned} \tag{2}$$

где в балансах (2) через $Z_0 \geq 0, X_0 \geq 0, Y_0 \geq 0$ обозначено количество «твердых» нейтралов и сторонников данных позиций, не изменяющих свою позицию ни при каких обстоятельствах (во всяком случае, на рассматриваемых промежутках времени). Таким образом, электорат состоит, вообще говоря, не только из амбивалентных, но и «убежденных» индивидов. Через $z(t) \geq 0, x(t) \geq 0, y(t) \geq 0$ в уравнении (2) обозначены, соответственно, текущие численности членов электората, способных относительно быстро (в сравнении с «твердой» его частью) перейти в другую группу. Величина $n(t) = z(t) + x(t) + y(t)$ представляет собой, таким образом, численность подверженного колебаниям электорального «болота». Очевидно, что если в какой-то момент времени $x(t) = 0$, то $X(t) = X_0(t)$ и текущее число сторонников позиции «X» — минимально возможное (и аналогично для противоположной позиции).

Сторонники и противники той или иной позиции отличаются от нейтралов тем, что совершают определенные действия в ее поддержку — голосуют, агитируют, участвуют в митингах и пикетах, обращаются и выступают в СМИ, оказывают финансовую помощь соответствующим силам, берутся за оружие и т. д.

Социально-политическим потенциалом (СПП) члена электората будем называть его готовность осуществлять поступки в пользу выбранной позиции (соответственно величина СПП представляет собой отнесенную к единице времени меру этой готовности). СПП относится к потенциалам так называемой «гуманитарной» группы, в отличие от, например экономического, научно-технического или военного потенциалов.

Активность («мощность» действий) индивидов может быть выражена через величины $E_X(t)$ и $E_Y(t)$, представляющие собой общее

число производимых в единицу времени равноценных (равных по силе) поступков в пользу той или иной позиции:

$$E_Y(t) = \sum_{i=0}^{Y_0(t)} E_{0,Y}^i(t) + \sum_{i=0}^{y(t)} E_y^i(t) \quad (t \geq t_0), \quad (3)$$

$$E_X(t) = \sum_{i=0}^{X_0(t)} E_{0,X}^i(t) + \sum_{i=0}^{x(t)} E_x^i(t) \quad (t \geq t_0), \quad (4)$$

где $E_{0,X}^i$, $E_x^i(t)$, $E_{0,Y}^i$, $E_y^i(t)$ — индивидуальные интенсивности действий членов электората из категорий $X_0(t)$, $x(t)$, $Y_0(t)$, $y(t)$ соответственно, i — номер индивида.

В настоящем разделе проводится построение и анализ математической модели для определения функций $E_X(t)$, $E_Y(t)$ — интегральных величин (вообще говоря, наблюдаемых и измеряемых), количественно описывающих временную динамику процесса выбора электоратом той или иной линии. Эта динамика, очевидно, зависит от начального состояния социальной общности, ее «внутренних» характеристик (исторического опыта, уровня культуры и т. д.), внешних воздействий на нее (например, со стороны других электоратов) и от многих других факторов и процессов. Поэтому для получения достаточно обозримой модели необходимо упростить первоначальную общую проблему.

3.3.2. Основные предположения

В дальнейшем будем рассматривать ситуацию выбора в относительно краткосрочной перспективе, когда считаются справедливыми следующие предположения.

Предположение 1. Число твердых сторонников позиций «X» и «Y» не изменяется с течением времен, т. е.

$$X_0(t) = X_0 \geq 0, \quad Y_0(t) = Y_0 \geq 0. \quad (5)$$

Предположение 2. Индивидуальные интенсивности поступков твердых сторонников обеих позиций (т. е. членов категорий X_0 , Y_0) постоянны во времени, т. е.

$$\begin{aligned} E_{0,X}^i(t) &= E_{0,X}^i \quad (1 \leq i \leq X_0), \\ E_{0,Y}^i(t) &= E_{0,Y}^i \quad (1 \leq i \leq Y_0). \end{aligned} \quad (6)$$

С учетом предположений 1, 2 получаем, что суммарные активности категорий X_0 , Y_0 (первые члены в правых частях формул (3), (4)) также постоянны и равны величинам

$$E_{0,X} = \sum_{i=0}^{X_0} E_{0,X}^i, \quad E_Y = \sum_{i=0}^{Y_0} E_{0,Y}^i. \quad (7)$$

Они представляют собой неизменные (на рассматриваемом интервале времени) во времени компоненты общей активности противоборствующих сторон. Величина

$$E_{X-Y}^0 = E_{0,X} - E_{0,Y} \quad (8)$$

характеризует постоянную разницу в активности твердых сторонников позиций « X » и « Y ». Она может трактоваться как постоянный «фон», на котором происходят относительно быстрые изменения активности «электорального болота».

В отличие от категорий X_0, Y_0 активности категорий $x(t), y(t)$ (вторые члены в правых частях формул (3), (4)) зависят от времени. Вместе с ними зависят от времени итоговые активности категорий X и Y , их сумма и разность.

В силу предположений 1, 2 исходная задача сводится к изучению динамики поведения членов электорального болота категории $n(t)$ (как уже отмечалось, $n(t) = z(t) + x(t) + y(t)$).

Очевидно, что индивиды не всегда имеют условия для полной реализации своего СПП (например, существующая в какой-то момент времени готовность электората сменить свое политическое руководство «замораживается» до наступления ближайших выборов). Это значит, что интенсивность действий члена электората, вообще говоря, не обязательно равна его индивидуальному потенциалу — она может быть меньше него. Другими словами, потенциал определяет максимально возможную (в данный момент) интенсивность действий индивида. Таким образом, имеют место соотношения

$$E_x^i(t) = \alpha_x^i(t)P_x^i(t), \quad 1 \leq i \leq x(t), \quad (9)$$

$$E_y^i(t) = \alpha_y^i(t)P_y^i(t), \quad 1 \leq i \leq y(t). \quad (10)$$

Здесь $P_x^i(t), P_y^i(t)$ — индивидуальные потенциалы членов электората, принадлежащих соответствующим категориям (i — номер индивида), а коэффициенты $\alpha_x^i(t) \in [0; 1]$, $\alpha_y^i(t) \in [0; 1]$ характеризуют возможность реализации существующего потенциала.

Будем считать справедливым следующее предположение.

Предположение 3. Возможности для полной реализации СПП существуют (т. е. в (9), (10) коэффициенты $\alpha_x^i(t) = 1$, $\alpha_y^i(t) = 1$).

Данное предположение не является принципиальным и носит лишь упрощающий характер. Если коэффициенты $\alpha_x^i(t), \alpha_y^i(t)$ являются экзогенно заданными известными функциями, то общий ход исследования модели остается таким же.

Следующий пункт посвящен построению математической модели для динамики численностей $x(t), y(t)$ и индивидуальных потенциалов $P_x^i(t), P_y^i(t)$ при выполнении предположений 1–3. После определения этих

функций по выписанным выше формулам нетрудно вычислить все оставшиеся искомые величины.

3.3.3. Построение модели

Очевидно, что динамика СПП определяется сложной совокупностью разнообразных факторов и процессов. Будем подразделять их на внутренние и внешние (по отношению к рассматриваемой социальной общности). Первые связаны с «собственными» поведенческими характеристиками электората, т. е. отвечают ситуации, когда электорат изолирован от внешнего мира, а его члены получают стимулы к изменению поведения, исходящие лишь из самого этого электората. Ко вторым относятся влияния окружающей среды: других электоратов, СМИ и прочих внешних по отношению к рассматриваемому электорату влияний — финансовых, экономических, религиозных, культурных и т. д.

В дальнейшем будем считать, что обусловленное внутренними (собственными) причинами поведение членов электората (и формируемая ими часть СПП) основано на *подражании* и описывается следующей *социально-психологической гипотезой*.

1. Член электората осуществляет свой выбор (самоидентификацию) путем подражания, наблюдая и принимая во внимание позитивные и негативные стимулы в ходе общения со своими «соседами» по электорату («сканируя» круг своего общения).

2. Стимулами, определяющими выбор, служат как величина абсолютной разницы между сторонниками и противниками данного курса

$$Q(t) = X(t) - Y(t) = X_0 - Y_0 + x(t) - y(t) = Q_0 + q(t) \quad (11)$$

$$Q_0 = X_0 - Y_0, \quad q(t) = x(t) - y(t), \quad t \geq t_0,$$

так и величина скорости изменения со временем (производной) этой разницы,

$$S(t) = \dot{Q}(t) = \dot{X}(t) - \dot{Y}(t) = \dot{x}(t) - \dot{y}(t) = \dot{q}(t), \quad t \geq t_0 \quad (12)$$

(где точкой обозначено дифференцирование по времени).

В соответствии с данной гипотезой на динамику СПП каждого члена электората влияют «статическая» (определяемая функцией $Q(t)$) и «ажиотажная» (определяемой функцией $S(t)$) характеристики электората как целого. Тем самым осуществляется влияние макроуровня социальной системы, т. е. электората в целом, на ее микроуровень, т. е. на каждого отдельного индивида. Обратная связь, т. е. влияние микроуровня на макроуровень, осуществляется путем агрегирования: «голоса» отдельных членов электората суммируются, формируя при этом численности $X(t)$, $Y(t)$.

Базовая модель подражательного поведения, рассмотренная в предыдущем разделе, также содержит подобную логику взаимодействия

микро- и макроуровней. Однако в базовой модели на формирование стимула влияет лишь «статическая» характеристика электората $Q(t) = X(t) - Y(t)$, но не ее скорость ее изменения.

Итак, в рамках расширенной модели подражательного поведения СПП i -го члена электората зависит от статической и ажиотажной характеристик электората $Q(t)$, $S(t)$, а также от внешних факторов (при этом на разных индивидов действуют, вообще говоря, различные внешние факторы, например разные средства массовой информации). Обозначим через $F_x^i(t)$, $F_y^i(t)$ экзогенно задаваемую «силу» влияния внешних факторов (например, активности СМИ) на формирование СПП. Формализуя приведенные выше рассуждения, запишем следующие выражения для СПП для членов категорий $x(t)$, $y(t)$:

$$\begin{aligned} P_x^i(t) &= P_x^i(Q, S, F_x^i(t)) = \\ &= P_x^i(X_0 - Y_0 + x(t) - y(t), \dot{x}(t) - \dot{y}(t), F_x^i(t)), \quad 1 \leq i \leq x(t), \end{aligned} \quad (13)$$

$$\begin{aligned} P_y^i(t) &= P_y^i(Q, S, F_y^i(t)) = \\ &= P_y^i(X_0 - Y_0 + x(t) - y(t), \dot{x}(t) - \dot{y}(t), F_y^i(t)), \quad 1 \leq i \leq y(t). \end{aligned} \quad (14)$$

Для построения замкнутой модели на основе соотношений (13), (14) необходимо сформулировать предположения и упрощения, дополнительные к тем, что были сделаны ранее.

Прежде всего, заметим, что зависимости в связях (13), (14) предполагают, что каждому члену электората свойственен свой (не такой, как у других) вид зависимости СПП от внешних и внутренних факторов. Так, для одного индивида статическая составляющая $Q(t)$ может быть более значима, чем ажиотажная составляющая $S(t)$, для другого индивида — наоборот. Учитывая практическую невозможность определения индивидуальных функциональных зависимостей в связях (13), (14) для каждого члена электората, используем введенное выше понятие *средней типической личности*, характеризующей рассматриваемую социальную общность. Другими словами, вместо функций P_x^i , P_y^i , F_x^i , F_y^i (характеризующих отдельного, i -го индивида) будем рассматривать лишь функции $P(t)$ и $F(t)$, представляющие, соответственно, СПП типической личности и силу воздействия на нее со стороны внешних факторов.

Понятие типической личности агрегирует разнообразие поведения членов «электорального болота» и позволяет перейти от индивидуальных потенциалов (13), (14) к СПП, характеризующему электорат в целом (или, что то же самое, СПП типической личности):

$$P(t) = P(Q, S, F(t)) = P(X_0 - Y_0 + x(t) - y(t), \dot{x}(t) - \dot{y}(t), F(t)). \quad (15)$$

В следующем пункте данного раздела предложен конкретный вид функции $P(Q, S, F(t))$ в выражении (15). Однако даже это еще не определяет замкнутую модель, так как она содержит две искомые

величины, $P(t)$ и $Q(t)$. Выделим два представляющих практическую значимость случая, в которых данная модель замыкается.

А. Пусть в результате социологических измерений удастся определить величину $Q(t)$ (и, соответственно, $\dot{Q}(t)$). Тогда выражение (15) позволяет определить потенциал $P(t)$ и затем все остальные неизвестные.

Б. Предположим, что на основании теоретико-социологических аргументов либо эмпирически удалось определить временную динамику СПП. В этом случае функция $P(t)$ является известной, и выражение (15) (разумеется, с конкретной функцией трех аргументов $P(Q, S, F(t))$) представляет собой обыкновенное дифференциальное уравнение первого порядка относительно $Q(t)$. Если известно начальное значение $Q(t_0)$, то, решая это уравнение, определяем функцию $Q(t)$, а затем и остальные неизвестные.

Очевидно, что в обоих случаях для нахождения остальных характеризующих процесс величин (E_x, E_y, E_X, E_Y и т. д.) нужна дополнительная связь между численностями $X(t), Y(t)$ (например, знание общей численности колеблющегося электората, т. е. величины $x(t) + y(t)$).

Случай А является довольно простым с математической точки зрения (см. упражнение 1). Поэтому в дальнейшем будем изучать случай Б.

Удобно, продифференцировав (15) по времени,

$$\frac{dP}{dt} = \frac{dP}{dQ} \frac{dQ}{dt} + \frac{dP}{dS} \frac{dS}{dt} + \frac{dP}{dF} \frac{dF}{dt},$$

и переписав последнее выражение в эквивалентном виде

$$\frac{dS}{dt} = \left(\frac{dP}{dt} - \frac{dP}{dQ} \frac{dQ}{dt} - \frac{dP}{dF} \frac{dF}{dt} \right) \left(\frac{dP}{dS} \right)^{-1},$$

перейти (с учетом равенства $dQ/dt = S$) к системе двух обыкновенных дифференциальных уравнений, разрешенных относительно величин Q и S :

$$\begin{cases} \frac{dS}{dt} = \frac{P_t - P_Q S - P_F F_t}{P_S}, \\ \frac{dQ}{dt} = S \end{cases}, \quad t \geq t_0. \quad (16)$$

Динамическая система 2-го порядка (16) с начальными значениями

$$Q(t_0) = Q^0, \quad S(t_0) = S^0 \quad (17)$$

представляет собой замкнутую математическую модель рассматриваемого процесса, полученную в рамках предположений 1–3, социально-психологической гипотезы и представления о средней типичской личности. При заданной функции $P(Q, S, F(t))$ из начальной задачи (16), (17) однозначно (при соответствующих условиях гладкости) находится искомое решение — функции $Q(t), S(t)$.

Помимо задач А и Б модель (16), (17) позволяет ставить вопрос об определении интенсивности внешнего воздействия $F(t)$ на электорат по известным потенциалу $P(t)$ и численности $Q(t)$. Такая постановка может быть названа «обратной задачей социальной психологии».

3.3.4. Конкретизация модели

Для исследования конкретных социальных процессов с помощью расширенной модели подражательного поведения необходимо задать конкретные входные данные — прежде всего, вид функции $P = P(Q, S, F(t))$, определяющей СПП. Для этого сделаем ряд дополнительных упрощений.

Упрощение 1. Внешнее воздействие постоянно во времени:

$$F(t) = \text{const}, \quad F_t = 0, \quad P = P(Q, S, C).$$

Для простоты записи в дальнейшем будем опускать параметрическую зависимость потенциала от постоянной C и записывать его в виде $P = P(Q, S)$.

Упрощение 2. СПП является аддитивной функцией статической и ажиотажной компонент:

$$P(Q, S) = \chi_1 P_1(Q) + \chi_2 P_2(S). \quad (18)$$

Здесь $\chi_1 \geq 0$, $\chi_2 \geq 0$ — веса соответствующих компонент СПП, $\chi_1 + \chi_2 = 1$.

Для определения конкретного вида функций $P_1(Q)$, $P_2(S)$ обратимся к известным из психологии закономерностям восприятия (см. комментарий). Основываясь на них, примем следующий вид компонент СПП:

$$P_1(Q) = \frac{\alpha Q^2}{Q^2 + \beta}, \quad (19)$$

$$P_2(S) = \frac{\mu}{\rho^3 \sqrt{\pi/2}} S^2 \exp \left[-\frac{S^2}{2\rho^2} \right]. \quad (20)$$

Здесь $\alpha > 0$, $\beta > 0$, $\mu > 0$, $\rho > 0$ — параметры, численно характеризующие особенности восприятия и служащие для «привязки» функций к эмпирически установленным конкретным зависимостям. Значения этих параметров являются, вообще говоря, различными для разных электоратов.

Прокомментируем выражения (19), (20).

Четная функция $P_1(Q)$ (рис. 1, а) монотонно растет с ростом $|Q|$ и выходит при $Q \rightarrow \pm\infty$ на асимптотическое значение $P_1(\pm\infty) = \alpha$. При этом ее производная (нечетная функция),

$$\frac{dP_1}{dQ} = \frac{2\alpha\beta Q}{(Q^2 + \beta)^2},$$

положительна при $Q > 0$, отрицательна при $Q < 0$, равна нулю при $Q = 0$, стремится к нулю при $Q \rightarrow \pm\infty$ и достигает своего максимального

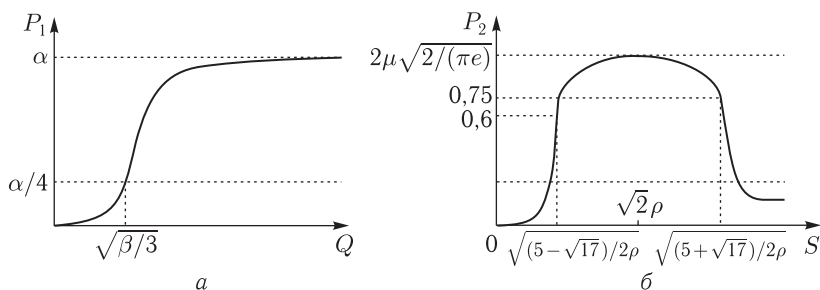


Рис. 1. Зависимость $P_1(Q)$ статической компоненты СПП от разности числа сторонников позиций «X» и «Y» (а), зависимость $P_2(S)$ ажиотажной компоненты СПП от скорости изменения этой разности (б)

по модулю значения, равного $\pm(\sqrt{3}/2)\alpha/\sqrt{\beta}$, в точках $Q = \pm\sqrt{\beta/3}$ (в силу симметрии функции $P_1(Q)$ относительно точки $Q = 0$ ее график приведен лишь при $Q \geq 0$).

Другими словами, с увеличением стимула (аргумента $|Q|$) отклик на стимул (функция $P_1(Q)$) увеличивается с насыщением, изменяясь (при $|Q| > \sqrt{\beta/3}$) все слабее и слабее. Этот эффект можно назвать также *привыканием с утомлением*.

Четная функция $P_2(S)$ (рис. 1, б) монотонно растет с ростом $|S|$ вплоть до точки с координатой $|S| = \sqrt{2}\rho$, где она максимальна и равна $P_2(S)_{\max} = 2\mu\sqrt{2/(\pi e)}$, при $|S| > \sqrt{2}\rho$ она монотонна и уменьшается до нуля при $|S| \rightarrow \pm\infty$ (график функции $P_2(S)$ также приведен лишь в области $S \geq 0$). Ее производная (нечетная функция)

$$\frac{dP_2}{dS} = P_2(S) \frac{2\rho^2 - S^2}{\rho^2 S}$$

равна нулю при $S = 0$ и $|S| = \sqrt{2}\rho$ и стремится к нулю при $|S| \rightarrow \pm\infty$, достигая своих максимальных по модулю значений в точках $S = \pm\rho\sqrt{(5 - \sqrt{17})/2}$ и $S = \pm\rho\sqrt{(5 + \sqrt{17})/2}$.

В отличие от функции $P_1(Q)$ динамическая компонента потенциала $P_2(S)$ увеличивается с увеличением стимула (аргумента $|S|$) лишь до достижения определенного порога восприятия (определяемого значением параметра ρ), а при дальнейшем его нарастании стремится к нулю. В психофизиологии этот эффект интерпретируется как *восприятие растущей по объему информации* (при превышении пропускной способности воспринимающего) *в качестве растущего по амплитуде «шума»* (все более заглушающего полезный сигнал).

Заметим, что определяемые параметрами α , β , μ , ρ значения аргументов функций (19), (20), самих функций и их производных в характерных точках имеют вполне ясную интерпретацию (максимальный отклик на стимул, максимальная скорость роста отклика на стимул, порог восприятия и т. д.).

3.3.5. Циклический режим эволюции электората

При отсутствии на определенном промежутке времени значимых социальных событий, которые могут существенно повысить или понизить готовность членов электората осуществлять поступки в пользу своей позиции, СПП можно приближенно считать постоянным. Поэтому высокой практической значимостью обладает частный случай, когда выполнено предположение.

Предположение 4. СПП является постоянной величиной, т. е.

$$\frac{dP}{dt} = 0. \quad (21)$$

При выполнении данного предположения и упрощений 1, 2 система уравнений (16) принимает вид

$$\begin{cases} \frac{dS}{dt} = -\frac{P_Q S}{P_S}, \\ \frac{dQ}{dt} = S. \end{cases} \quad (22)$$

Будем называть базовой моделью электорального поведения систему (22), в которой потенциал $P(Q, S)$ задан уравнениями (18)–(20). В этом случае из системы (22) имеем

$$\begin{cases} \frac{dS}{dt} = -\frac{2\chi_1\alpha\beta\rho^5 Q \exp(S^2/2\rho^2)}{\sqrt{2/\pi} \chi_2\mu(Q^2 + \beta)^2(2\rho^2 - S^2)}, \\ \frac{dQ}{dt} = S. \end{cases} \quad (23)$$

О подробном анализе базовой модели электорального поведения (23) см. в разделе «Комментарии», здесь же ограничимся рассмотрением случая, когда разность численностей сторонников позиций «X» и «Y», и скорость ее изменения достаточно малы, т. е.

$$Q \ll 1, \quad S \ll 1. \quad (24)$$

С учетом соотношений (24) получаем из (23) систему

$$\begin{cases} \frac{dS}{dt} = -\frac{\chi_1\alpha\rho^3 Q}{\sqrt{2/\pi} \chi_2\mu\beta}, \\ \frac{dQ}{dt} = S. \end{cases} \quad (25)$$

Переходя от (25) к одному уравнению второго порядка, имеем

$$\frac{d^2 Q}{dt^2} = -\frac{\chi_1\alpha\rho^3 Q}{\sqrt{2/\pi} \chi_2\mu\beta}. \quad (26)$$

Общее решение этого уравнения имеет вид

$$Q = Q_0 \sin \frac{2\pi(t - t_*)}{T}, \quad (27)$$

где амплитуда Q_0 и момент времени t_* являются произвольными постоянными, а период колебаний равен

$$T = 2\pi \left(\frac{\chi_2 \mu \beta}{\chi_1 \alpha \rho^3 \sqrt{\pi/2}} \right)^{1/2}, \quad (28)$$

Дифференцируя (27), получаем для скорости изменения разности сторонников различных позиций $S = dQ/dt$ выражение

$$S = \frac{2\pi}{T} Q_0 \cos \frac{2\pi(t - t_*)}{T}. \quad (29)$$

Проанализируем выражения (27), (29). В моменты времени

$$\begin{aligned} t = t_*, \quad t = t_* + \frac{T}{2}, \quad t = t_* + T, \\ t = t_* + \frac{3T}{2}, \quad t = t_* + 2T, \dots \end{aligned}$$

статическая компонента потенциала $Q(t)$ равна нулю, а ажиотажная компонента $S(t)$ достигает своего максимума (по абсолютной величине). И наоборот — в моменты

$$\begin{aligned} t = t_* + \frac{T}{4}, \quad t = t_* + \frac{3T}{4}, \quad t = t_* + \frac{5T}{4}, \\ t = t_* + \frac{7T}{4}, \quad t = t_* + \frac{9T}{4}, \dots \end{aligned}$$

ажитажная компонента $S(t)$ равна нулю, а статическая компонента $Q(t)$ достигает своего максимума (по абсолютной величине). Этот процесс «перекачки потенциала туда и обратно» между ажитажной и статической компонентами аналогичен процессу перехода кинетической энергии механической колебательной системы в потенциальную и обратно. Простейшим примером из области механики является шарик массы m , присоединенный к пружине жесткостью k . Напомним, что здесь величина массы m характеризует инерционные (кинетические) свойства системы, а величина жесткости k — ее силовые (статические) свойства. Период колебаний шарика на пружине дается формулой

$$T_{\text{mech}} = 2\pi(m/k)^{1/2}. \quad (30)$$

Чтобы провести сопоставление между выражением (30) и формулой (28) для периода колебаний электората, выпишем выражение для СПП, определяемое формулами (18)–(20):

$$P(Q, S) = \chi_1 \frac{\alpha Q^2}{Q^2 + \beta} + \chi_2 \frac{\mu}{\rho^3 \sqrt{\pi/2}} S^2 \exp\left(-\frac{S^2}{2\rho^2}\right). \quad (31)$$

Из (31) следует, что ажитажная компонента СПП характеризуется параметрами χ_2 , μ , ρ , более конкретно — их комбинацией $m_{\text{soc}} = \frac{\chi_2 \mu}{\rho^3 \sqrt{\pi/2}}$, которая является аналогом массы шарика. В то же время, аналогом

жесткости пружины является величина $k_{\text{soc}} = \frac{\chi_1 \alpha}{\beta}$, характеризующая статическую компоненту СПП. С учетом этих обозначений формула для периода колебаний электората может быть записана в виде

$$T = 2\pi (m_{\text{soc}}/k_{\text{soc}})^{1/2}, \quad (32)$$

аналогичном выражению (30) для периода колебаний шарика на пружине.

Конечно же, аналогия между колебаниями электората и колебаниями механической системы весьма условна, но (тем не менее) она позволяет дать более ясную интерпретацию зависимости (32).

1. Чем больше удельный вес статической части потенциала в общем постоянном потенциале (т.е. величина k_{soc}), тем сравнительно больше «сила принуждения» личности со стороны большинства и тем быстрее она совершает переходы — колебания (период уменьшается с ростом k_{soc}).

2. Чем больше удельный вес ажиотажной (кинетической) части потенциала, тем сравнительно больше инерционность личности (тем сравнительно меньше, в условиях постоянства потенциала, сила принуждения со стороны большинства) и тем больше времени требуется для изменения ее позиции (период растет с ростом m_{soc}).

Тем самым изучение базовой модели электорального поведения (23) показывает, что существует принципиальная *возможность (до определенной степени) осмысленного управления циклическими режимами эволюции электората* путем изменения параметров системы.

Заметим также, что для электоратов, обладающих разными свойствами (разными значениями величин χ_1 , χ_2 , α , β , μ , ρ), периоды колебаний могут быть близкими, т.е. *может существовать количественное подобие электорального поведения*.

Изученный в п. 3.3.5 режим эволюции со временем числа сторонников позиций «X», «Y» называется *режимом циклической самоидентификации электората*. Он описывает процесс циклического «перемещения» электорального болота то на одну, то на другую позицию. О других режимах самоидентификации см. в комментариях. Период колебаний (28) естественно трактовать как *период самоидентификации электората*.

Таким образом, рассмотренная в данном разделе расширенная модель подражательного поведения представляет собой инструмент изучения гораздо более сложной (по сравнению с базовой моделью, см. предыдущий параграф) социальной структуры, возникающей вследствие подражательного поведения индивидов. Эти возможности возникли благодаря постулированию более сложной поведенческой гипотезы, учитывающей не только статическую, но и ажиотажную (динамическую) компоненту подражательного механизма.

Упражнения

1. Предположим, что в некоторой социальной группе постоянной численностью N_0 отсутствуют как нейтралы, так и твердые сторонники обеих позиций: $Z(t) = 0$, $X_0 = Y_0 = 0$. Таким образом, $x(t) + y(t) = N_0$. В результате эмпирического социологического исследования удалось определить, что количество сторонников позиции «X» имеет вид $x(t) = N_0(0,4 + 0,3 \sin \omega t)$, где $\omega > 0$ — некоторая постоянная. Найти функцию $Q(t)$; затем, используя функции вида (18)–(20), получить выражение $P(t)$, описывающее динамику СПП.

2. Построить фазовый портрет системы (25).

Комментарии и библиографические примечания

1. Концепция социально-политического потенциала сформулирована и разработана в работах [1–4], см. также монографию [5].

2. Формулы (19), (20) соответствуют хорошо известным в психологии эмпирически установленным конкретным зависимостям (законам Вебера–Фехнера, Бугера и Стивенса).

3. Подробный анализ базовой модели электорального поведения (23) проведен в [5]. В частности, для базовой модели электорального поведения (23) без упрощения $Q \ll 1$, $S \ll 1$ изучены также иные (помимо колебательного, рассмотренного в пятом пункте) режимы самоидентификации электората: полный переход электората на позицию «X» или «Y» (однозначная самоидентификация) и переход к стабильному соотношению числа сторонников и противников («поляризованная» самоидентификация).

Литература

1. Шведовский В.А. Динамическая модель электорального поведения // Математическое моделирование. 2000. Т. 12. С. 46–56.
2. Шведовский В.А. Динамическая модель этнополитического конфликта: построение, возможности и результаты применения. // Сб. Математическое моделирование социальных процессов. М.: МГУ. 2000. Вып. 2. С. 31–37.
3. Михайлов А.П., Юхно Л.Ф. Исследование некоторых математических моделей эволюции электората // Дифференциальные уравнения. 2002. Т. 38. С. 1322.
4. Михайлов А.П., Шведовский В.А., Маслов А.И., Ковалев В.Ф. Обобщенная модель электорального поведения и ее применение к изучению этнополитических конфликтов // Математическое моделирование. 2003. Т. 15. С. 39–56.
5. Михайлов А.П. Моделирование системы власть — общество. М.: ФИЗМАТЛИТ, 2006. 144 с.

3.4. Формирование мнений и установок индивидов, составляющих замкнутую группу

Настоящий раздел посвящен построению и анализу двух моделей, описывающих динамику мнений и установок индивидов, составляющих замкнутую группу. Замкнутость в данном случае означает, что индивиды не покидают и не пополняют группу. Обе модели предполагают, что члены группы обмениваются информацией в процессе общения, в результате и возникает данная динамика.

Под установкой (аттитюдом) в социологии и психологии обозначается предрасположенность индивида к действию или реагированию в определенном направлении. Установки, как и мнения индивидов, формируются под влиянием их предшествующего опыта, социальной среды, экономических обстоятельств и т. д. Таким образом, установка действия отличается от более стабильного «внутреннего стандарта». Например, индивид может полагать, что его «норма» составляет 5 сигарет в день, но под влиянием обстоятельств (например, более активного курения товарищей по коллективу) у него может сформироваться установка на выкуривание 15 сигарет в день. В то же время необходимо заметить, что установка не всегда может быть реализована в виде реального действия — например, при физическом отсутствии сигарет.

Первый пункт настоящего раздела посвящен построению модели формирования мнений и ее анализу. Модель имеет вид системы уравнений в дискретном времени. В ее рамках вводится понятие консенсуса, т. е. мнения, разделяемого всеми членами группы. Анализ модели имеет целью, главным образом, изучение вопроса о том, какова должна быть структура отношений между членами группы, чтобы с течением времени в ней сформировался консенсус. Второй пункт посвящен приложению данной теории к анализу нескольких конкретных типов групп. В третьем пункте изучается модель формирования установок, имеющая вид системы обыкновенных дифференциальных уравнений.

3.4.1. Построение и анализ модели формирования мнений

Рассмотрим замкнутую группу k индивидов, находящихся в состоянии неопределенности относительно значения некоторой переменной θ , соответствующей исходу некоторого процесса или явления. В момент времени $t = 0$ у каждого члена группы имеются начальные субъективные представления о распределении вероятностей, с которыми может быть реализован каждый из вариантов.

Например, общество может находиться в неопределенности относительно того, кто победит на следующих президентских выборах. Некоторый конкретный член группы может полагать с вероятностью 50%, что это будет кандидат A_1 , с вероятностью 40% — кандидат A_2 , с вероятностью 10% — кандидат A_3 . Здесь переменная θ принимает значения на множестве $\{A_1, A_2, A_3\}$, а мнение i -го индивида имеет вид вектора $m_i = \{0,5; 0,4; 0,1\}$. Простейший частный случай мнения

возникает, когда индивиды делают точечные оценки переменной θ ; в этом случае мнение можно считать скалярной величиной, равной соответствующему значению этой переменной. Например, первый индивид полагает, что на определенную дату нефть будет стоить $m_1 = 80$, а второй индивид, что $m_2 = 70$ долларов за баррель.

С течением времени индивиды корректируют свои мнения с учетом мнений других членов группы. В рамках данной модели время предполагается дискретным, а переменная θ , относительно значений которой рассматриваются мнения, — скалярной. Мнение индивида i ($i = 1, \dots, k$) в момент времени t выражается распределением вероятности $m_i(t)$ на множестве значений θ . В течение интервала времени $(t, t + 1)$ индивид i корректирует его с учетом мнений других членов группы (а также своего собственного) в момент t : $m_i(t + 1) = f(m_1(t), \dots, m_k(t))$.

Предположение 1. *Влияние мнений других индивидов аддитивно, причем мнения некоторых из этих индивидов более весомы, чем другие (например, i -й член группы может считать кого-то из партнеров по группе более информированным или более опытным).*

Обозначим через w_{ij} вес, приписываемый i -м индивидом мнению j -го. При этом веса предполагаются постоянными, неотрицательными и нормированными на единицу:

$$w_{ij} \geq 0, \quad i, j = 1, \dots, k, \quad \sum_{j=1}^k w_{ij} = 1, \quad i = 1, \dots, k. \quad (1)$$

Чтобы сформулировать предположение о том, как именно влияют мнения, рассмотрим минимальную группу, состоящую всего из двух индивидов.

Предположение 2. *В паре индивидов изменение мнения каждого из них пропорционально разнице мнений и направлено в сторону сближения этих мнений. Коэффициенты пропорциональности равны введенным выше весам.*

В соответствии с данным предположением для мнений индивидов в момент времени $t + 1$ имеем

$$m_1(t + 1) - m_1(t) = w_{12} [m_2(t) - m_1(t)], \quad (2)$$

$$m_2(t + 1) - m_2(t) = w_{21} [m_1(t) - m_2(t)]. \quad (3)$$

Для того чтобы перейти к выводу уравнений для группы, состоящей из произвольного количества индивидов, воспользуемся предположением 1 об аддитивности. Получим для первого индивида уравнение, обобщающее (2) и имеющее следующий вид:

$$m_1(t + 1) - m_1(t) = w_{12} [m_2(t) - m_1(t)] + \dots + w_{1k} [m_k(t) - m_1(t)]. \quad (4)$$

Приводя подобны слагаемые, получим из (4), что

$$m_1(t+1) = [1 - w_{12} - \dots - w_{1k}]m_1(t) + w_{12}m_2(t) + \dots + w_{1k}m_k(t). \quad (5)$$

Но так как $\sum_{j=1}^k w_{1j} = 1$ (см. (1)), то данное уравнение принимает вид

$$m_1(t+1) = w_{11}m_1(t) + w_{12}m_2(t) + \dots + w_{1k}m_k(t) = \sum_{j=1}^k w_{1j}m_j. \quad (6)$$

Тем самым построено уравнение динамики мнения первого индивида. Проводя аналогичные рассуждения для других индивидов, получим, что мнение $m_i(t+1)$ i -го индивида в момент $t+1$ имеет вид

$$m_i(t+1) = \sum_{j=1}^k w_{ij}m_j(t), \quad i = 1, \dots, k. \quad (7)$$

Составим весовую матрицу W , а также вектор-столбец мнений членов группы $M(t)$:

$$W = \begin{pmatrix} w_{11} & w_{12} & \dots & w_{1k} \\ w_{21} & w_{22} & \dots & w_{2k} \\ \dots & \dots & \dots & \dots \\ w_{k1} & w_{k2} & \dots & w_{kk} \end{pmatrix}, \quad M(t) = \begin{pmatrix} m_1(t) \\ m_2(t) \\ \dots \\ m_k(t) \end{pmatrix}.$$

Система уравнений (7) в векторном виде принимает форму

$$M(t+1) = WM(t). \quad (8)$$

Таким образом,

$$M(t) = W^t M(0). \quad (9)$$

Если с течением времени мнения всех членов группы сближаются так, что

$$\lim_{t \rightarrow \infty} m_1(t) = \lim_{t \rightarrow \infty} m_2(t) = \dots = \lim_{t \rightarrow \infty} m_k(t),$$

то говорят, что в группе *достигается консенсус*, а соответствующее предельное мнение $m_{\text{con}} = \lim_{t \rightarrow \infty} m_i(t)$ называется *консенсусом*. Предельный вектор мнений приобретает тогда вид

$$\lim_{t \rightarrow \infty} M(t) = M_{\text{con}} = \begin{pmatrix} m_{\text{con}} \\ m_{\text{con}} \\ \dots \\ m_{\text{con}} \end{pmatrix}. \quad (10)$$

С содержательной точки зрения важный вопрос заключается в том, при каких условиях в группе индивидов гарантированно (т.е. при любых начальных мнениях) возникает консенсус. Чтобы получить эти

условия, сделаем предельный переход $t \rightarrow \infty$ в равенстве (8) и, воспользовавшись выражением (10), получим

$$\begin{pmatrix} m_{\text{con}} \\ m_{\text{con}} \\ \dots \\ m_{\text{con}} \end{pmatrix} = \lim_{t \rightarrow \infty} W^t \begin{pmatrix} m_1(0) \\ m_2(0) \\ \dots \\ m_k(0) \end{pmatrix}. \quad (11)$$

Предположив, что существует $\lim_{t \rightarrow \infty} W^t$, введем в рассмотрение матрицу

$$W^{(\infty)} = \lim_{t \rightarrow \infty} W^t \quad (12)$$

и обозначим ее элементы через $w_{ij}^{(\infty)}$. Из (11) следует, что консенсус достигим из любого начального распределения, если все строки матрицы $W^{(\infty)}$ тождественны, т. е.

$$w_{1j}^{(\infty)} = w_{2j}^{(\infty)} = \dots = w_{kj}^{(\infty)}. \quad (13)$$

Итак, если существует предельная матрица W^∞ , обладающая свойством (12) то в группе достигается консенсус. Достаточные условия даются следующей теоремой.

Теорема 1. Если существует натуральное число n такое, что матрица W^n имеет хотя бы один столбец, состоящий целиком из положительных элементов, то консенсус достигается. Он имеет вид $M_{\text{con}} = W^{(\infty)}M(0)$, где матрица $W^{(\infty)} = \lim_{t \rightarrow \infty} W^t$ имеет вид

$$W^{(\infty)} = \begin{pmatrix} h_1 & h_2 & \dots & h_k \\ h_1 & h_2 & \dots & h_k \\ \dots & \dots & \dots & \dots \\ h_1 & h_2 & \dots & h_k \end{pmatrix}, \quad (14)$$

Здесь h_1, \dots, h_k — компоненты собственного вектора транспонированной матрицы W^* , соответствующего собственному значению $\lambda = 1$, и нормированного так, что $h_1 + h_2 + \dots + h_n = 1$.

Чтобы вычислить консенсус m_{con} с помощью Теоремы 1, подставим выражение (14) в уравнение $M_{\text{con}} = W^{(\infty)}M(0)$. Получим

$$m_{\text{con}} = h_1 m_1(0) + \dots + h_k m_k(0). \quad (15)$$

Таким образом, консенсус представляет собой линейную комбинацию начальных мнений членов группы.

Следующий пункт посвящен применению данных результатов к анализу динамики мнений в нескольких типах группах, различающихся по структуре.

3.4.2. Консенсус для некоторых типов групп индивидов

В рамках данной модели каждая конкретная социальная группа характеризуется матрицей W , отражающей отношения между ее членами. Другими словами, матрица W описывает структуру группы

в аспекте влияния одних ее членов на других. Рассмотрим несколько простых типовых структур.

А. Все члены группы влияют друг на друга

В этом случае все элементы матрицы W положительны, и она удовлетворяет условиям теоремы 1; вследствие этого в группе возникает консенсус. Рассмотрим в качестве примера группу со следующей матрицей W :

$$W = \begin{pmatrix} 0,6 & 0,2 & 0,2 \\ 0,3 & 0,6 & 0,1 \\ 0,3 & 0,1 & 0,6 \end{pmatrix}. \quad (16)$$

Здесь все индивиды в равной степени опираются на свое мнение в предыдущий момент (все диагональные элементы равны 0,6),

Однако первый индивид является наиболее влиятельным: он оказывает на каждого из других членов группы большее влияние, чем они оказывают друг на друга. Поэтому следует ожидать, что достигаемый в группе консенсус определяется в большей степени его начальным мнением, чем мнениями других индивидов. Покажем это, воспользовавшись результатами предыдущего пункта.

Нетрудно вычислить, что собственные значения матрицы W вида (16) равны $\lambda_1 = 1$, $\lambda_2 = \lambda_3 = 0,1$. Собственный вектор транспонированной матрицы W^* , соответствующий собственному значению $\lambda_1 = 1$, имеет вид

$$h = \begin{pmatrix} 3/7 \\ 2/7 \\ 2/7 \end{pmatrix}.$$

Таким образом, получаем выражение консенсуса через начальные мнения:

$$m_{\text{кон}} = \frac{3}{7}m_1(0) + \frac{2}{7}m_2(0) + \frac{2}{7}m_3(0).$$

Полученное выражение показывает, что в данном примере первый член группы оказывает в $3/2$ раза большее влияние на консенсус, чем каждый из двух других индивидов.

Б. Группа, в которой один из членов не влияет на других

«Присоединим» к рассмотренной выше группе еще одного индивида, который испытывает влияние со стороны первых трех, но сам не оказывает влияния на их мнения. Соответствующая матрица W имеет вид

$$W = \begin{pmatrix} 0,6 & 0,2 & 0,2 & 0 \\ 0,3 & 0,6 & 0,1 & 0 \\ 0,3 & 0,1 & 0,6 & 0 \\ 0,1 & 0,1 & 0,1 & 0,7 \end{pmatrix}. \quad (17)$$

Собственные значения данной матрицы равны $\lambda_1 = 1$, $\lambda_2 = \lambda_3 = 0,1$, $\lambda_4 = 0,7$. Собственный вектор транспонированной матрицы W^* ,

соответствующий собственному значению $\lambda_1 = 1$, имеет вид

$$h = \begin{pmatrix} 3/7 \\ 2/7 \\ 2/7 \\ 0 \end{pmatrix}.$$

Отсюда получаем выражение консенсуса через начальные мнения:

$$m_{\text{кон}} = \frac{3}{7}m_1(0) + \frac{2}{7}m_2(0) + \frac{2}{7}m_3(0) + 0 \cdot m_4(0).$$

Таким образом, в данном примере начальное мнение четвертого члена группы не влияет на значение консенсуса. В то же время начальные мнения других индивидов влияют в той же степени, что и в предыдущем примере.

В. Группа, в которой только один индивид влияет на остальных
Рассмотрим группу, описываемую матрицей W вида

$$W = \begin{pmatrix} 1 & 0 & 0 \\ 0,4 & 0,6 & 0 \\ 0,4 & 0 & 0,6 \end{pmatrix}. \quad (18)$$

Здесь первый член группы является единственным, кто оказывает влияние на мнения других индивидов. Собственные значения данной матрицы равны $\lambda_1 = 1$, $\lambda_2 = \lambda_3 = 0,6$. Собственный вектор транспонированной матрицы W^* , соответствующий собственному значению $\lambda_1 = 1$, имеет вид

$$h = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}.$$

Следовательно, консенсус в такой группе равен мнению первого индивида: $m_{\text{кон}} = m_1(0)$.

Г. Составная группа, состоящая из двух подгрупп

В качестве примера ситуации, в которой условия теоремы 1 не выполнены и консенсус не достигается, рассмотрим матрицу

$$W = \begin{pmatrix} 0,6 & 0,4 & 0 & 0 \\ 0,3 & 0,7 & 0 & 0 \\ 0 & 0 & 0,8 & 0,2 \\ 0 & 0 & 0,1 & 0,9 \end{pmatrix}.$$

Данная группа фактически состоит из двух подгрупп: первый и второй индивиды влияют друг на друга; третий и четвертый влияют друг на друга, но влияние между этими парами отсутствует.

Рассмотрим формирование «частичного консенсуса» в первой подгруппе, имеющей матрицу

$$W_1 = \begin{pmatrix} 0,6 & 0,4 \\ 0,3 & 0,7 \end{pmatrix}.$$

Собственные значения данной матрицы равны $\lambda_1 = 1$, $\lambda_2 = 0,3$. Собственный вектор транспонированной матрицы W^* , соответствующий собственному значению $\lambda_1 = 1$, имеет вид

$$h = \begin{pmatrix} 3/7 \\ 4/7 \end{pmatrix}.$$

Отсюда получаем выражение «частичного консенсуса» в подгруппе, состоящей из первого и второго индивидов, через начальные мнения:

$$m_{\text{кон},1,2} = \frac{3}{7}m_1(0) + \frac{4}{7}m_2(0).$$

Аналогично, для подгруппы из третьего и четвертого индивидов имеем

$$m_{\text{кон},3,4} = \frac{1}{3}m_3(0) + \frac{2}{3}m_4(0).$$

При этом, вообще говоря, $m_{\text{кон},1,2} = m_{\text{кон},3,4}$, т.е. консенсус в группе отсутствует.

Таким образом, в зависимости от структуры отношений между индивидами, составляющими замкнутую группу, динамика мнений может приводить как к общегрупповому консенсусу, так и формированию нескольких «частичных консенсусов», отвечающих разбиению данной группы на подгруппы.

Оставшаяся часть раздела посвящена близкой тематике, именно — вопросам формирования установок индивидов в замкнутой группе.

3.4.3. Модель формирования установки действия

Под установкой в психологии и социологии понимают расположенность индивида к определенному направлению действий в конкретной ситуации (см. комментарий 4). Поясним данное понятие на следующем примере. Пусть некоторый индивид считает, что его «норма» курения — b сигарет в день (это представление о «норме» называется установкой намерений индивида на курение). Тогда, если он выкуривает в некоторый день более, чем b сигарет, то при отсутствии внешних факторов его установка действия — это отрицательное число, характеризующее психологическое напряжение данного индивида в направлении уменьшения курения. Если же он выкуривает в какой-либо день менее, чем b сигарет, то (при отсутствии внешних факторов) установка действия — это положительное число, характеризующее психологическое напряжение в направлении увеличения курения. К факторам, влияющим на формирование установки, относятся подражание и следование некоторому «внешнему стандарту».

С течением времени установка изменяется как ввиду действий индивида, направленных на уменьшение психологического напряжения (т.е. ввиду уменьшения или увеличения курения), так и вследствие изменения внешних факторов.

Перейдем к построению математической модели. Рассмотрим замкнутую группу, состоящую из k индивидов, которые в процессе

общения получают информацию об установках друг друга. Замкнутость в данном случае означает, что состав группы с течением времени не изменяется. Обозначим через $x_i(t)$ значение установки i -го индивида в момент времени t . Динамика установки определяется тремя факторами: подражанием другим членам группы, наличием «собственного стандарта» b (т.е. то, что он сам считает для себя нормой) и предъявляемого ему извне «внешнего стандарта» c_i .

Подражание понимается следующим образом. В процессе общения с другими членами группы i -й индивид определяет их текущие значения установок $x_j(t)$, $i = 1, \dots, k$, вычисляет средневзвешенное значение $\sum_{j=1}^k \gamma_{ij} x_j(t)$ и изменяет свою установку в направлении этой величины. Здесь весовые коэффициенты $\gamma_{ij} \geq 0$, $i = 1, \dots, k$, $\sum_{j=1}^k \gamma_{ij} = 1$ описывают относительную значительность членов группы для i -го индивида в качестве «образцов для подражания».

Влияние «собственного стандарта» b_i на динамику установки состоит в том, что он «притягивает» установку к себе, т.е. способствует увеличению величины $x_i(t)$, если $x_i(t) < b_i$, и ее уменьшению, если $x_i(t) > b_i$. Аналогично влияет и «внешний стандарт» c_i .

Предполагается, что данные факторы действуют аддитивно.

Таким образом, модель формирования установок членов замкнутой группы имеет следующий вид:

$$\frac{dx_i}{dt} = A_i \left(\sum_{j=1}^k \gamma_{ij} x_j - x_i \right) + B_i (b_i - x_i) + C_i (c_i - x_i), \quad x_i(0) = x_i^0. \quad (19)$$

Здесь коэффициенты A_i , B_i , C_i описывают склонность i -го индивида к подражанию, следованию «внутреннему стандарту» и «внешнему стандарту» соответственно.

В матричной форме система (19) имеет вид

$$\frac{dX}{dt} = PX + D,$$

где

$$P = \begin{pmatrix} A_1(\gamma_{11}-1) - B_1 - C_1 & A_1\gamma_{12} & \dots & A_1\gamma_{1k} \\ A_2\gamma_{21} & A_2(\gamma_{22}-1) - B_2 - C_2 & \dots & A_2\gamma_{2k} \\ \dots & \dots & \dots & \dots \\ A_k\gamma_{k1} & A_k\gamma_{k2} & \dots & A_k(\gamma_{kk}-1) - B_k - C_k \end{pmatrix},$$

$$X = \begin{pmatrix} x_1 \\ x_2 \\ \dots \\ x_k \end{pmatrix}, \quad D = \begin{pmatrix} B_1 b_1 + C_1 c_1 \\ B_2 b_2 + C_2 c_2 \\ \dots \\ B_k b_k + C_k c_k \end{pmatrix}.$$

Стационарное состояние определяется из системы k линейных алгебраических уравнений

$$A_i \left(\sum_{j=1}^k \gamma_{ij} x_j - x_i \right) + B_i (b_i - x_i) + C_i (c_i - x_i) = 0. \quad (20)$$

В качестве примера рассмотрим группу, состоящую из трех индивидов таких, что $A_i = A$, $B_i = B$, $C_i = C$, $\gamma_{ij} = 1/3$, $i, j = 1, 2, 3$. Модель (19) принимает вид

$$\frac{dx_1}{dt} = A \frac{-2x_1 + x_2 + x_3}{3} + B(b_1 - x_1) + C(c_1 - x_1), \quad x_1(0) = x_1^0, \quad (21)$$

$$\frac{dx_2}{dt} = A \frac{-2x_2 + x_1 + x_3}{3} + B(b_2 - x_2) + C(c_2 - x_2), \quad x_2(0) = x_2^0, \quad (22)$$

$$\frac{dx_3}{dt} = A \frac{-2x_3 + x_1 + x_2}{3} + B(b_3 - x_3) + C(c_3 - x_3), \quad x_3(0) = x_3^0. \quad (23)$$

Матрица системы (21)–(23) имеет вид

$$P = \begin{pmatrix} -\frac{2}{3}A - B - C & \frac{1}{3}A & \frac{1}{3}A \\ \frac{1}{3}A & -\frac{2}{3}A - B - C & \frac{1}{3}A \\ \frac{1}{3}A & \frac{1}{3}A & -\frac{2}{3}A - B - C \end{pmatrix}.$$

Ее собственные значения равны $\lambda_1 = -B - C$, $\lambda_2 = \lambda_3 = -A - B - C$. Поскольку все они отрицательны, то стационарное решение линейной системы (21)–(23) является единственным и асимптотически устойчивым.

Чтобы найти это стационарное решение, положим все производные равными нулю в уравнениях (21)–(23). После приведения подобных слагаемых получим

$$\left(-\frac{2}{3}A - B - C \right) x_1 + \frac{A}{3} x_2 + \frac{A}{3} x_3 = -(Bb_1 + Cc_1), \quad (24)$$

$$\frac{A}{3} x_1 + \left(-\frac{2}{3}A - B - C \right) x_2 + \frac{A}{3} x_3 = -(Bb_2 + Cc_2), \quad (25)$$

$$\frac{A}{3} x_1 + \frac{A}{3} x_2 + \left(-\frac{2}{3}A - B - C \right) x_3 = -(Bb_3 + Cc_3). \quad (26)$$

Сложив уравнения (24)–(26), придем к уравнению

$$(B + C)(x_1 + x_2 + x_3) = B(b_1 + b_2 + b_3) + C(c_1 + c_2 + c_3),$$

из которого следует, что стационарные значения установок удовлетворяют соотношению

$$x_1 + x_2 + x_3 = \frac{B(b_1 + b_2 + b_3) + C(c_1 + c_2 + c_3)}{B + C}. \quad (27)$$

Подставив (27) в уравнение (24), после простых преобразований получаем стационарное значение установки первого индивида:

$$x_1^{st} = \frac{1}{A+B+C} \left[(Bb_1 + Cc_1) + \frac{AB(b_1 + b_2 + b_3) + AC(c_1 + c_2 + c_3)}{3(B+C)} \right]. \quad (28)$$

Аналогично,

$$x_2^{st} = \frac{1}{A+B+C} \left[(Bb_2 + Cc_2) + \frac{AB(b_1 + b_2 + b_3) + AC(c_1 + c_2 + c_3)}{3(B+C)} \right], \quad (29)$$

$$x_3^{st} = \frac{1}{A+B+C} \left[(Bb_3 + Cc_3) + \frac{AB(b_1 + b_2 + b_3) + AC(c_1 + c_2 + c_3)}{3(B+C)} \right]. \quad (30)$$

Итак, у индивидов, составляющих данную группу, с течением времени формируются стационарные установки, задаваемые уравнениями (28)–(30).

Если все индивиды обладают одинаковыми «внутренними стандартами» и «внешними стандартами», т. е. $b_1 = b_2 = b_3$, $c_1 = c_2 = c_3$, то в группе устанавливается «консенсус установок», аналогичный «консенсусу мнений» в модели, рассмотренной выше: $x_1^{st} = x_2^{st} = x_3^{st}$.

Таким образом, рассмотренные в данном разделе модели описывают процессы социальных взаимодействий между членами замкнутой группы, в результате которых при определенных условиях формируется понимаемый тем или иным образом консенсус: разделяемое всеми индивидами мнение по некоторому вопросу или, например, единая для всех членов группы установка на выкуривание определенного количества сигарет в день.

Упражнения

1. Рассмотреть группу с матрицей W , задаваемой формулой (16). Возможны ли начальные условия, при которых сближение мнений индивидов будет немонотонным?

2. Модель формирования установок членов замкнутой группы имеет вид (19) при условии, что все индивиды правильно определяют значения установок друг друга. Предложить различные модификации модели, предполагающие, что индивиды определяют чужие установки с ошибками. Рассмотреть случай, когда индивид $i = 1$ считает текущие установки других членов группы в два раза более близкими к себе, чем на самом деле. Например, если установка второго индивида равна $x_2(t)$, но первый индивид при подражании исходит из того, что она равна $(x_1(t) + x_2(t))/2$.

Комментарии и библиографические примечания

1. Модель динамики мнений излагается в соответствии с работами [1, 2].

2. Вместо сделанного предположения 2 о том, что изменение мнения индивида пропорционально текущей разнице мнений, могут быть высказаны иные гипотезы. Например, вместо формулы (2) можно принять

$$m_1(t+1) - m_1(t) = w_{12} [m_2(t) - m_1(t)]^\alpha, \quad \alpha > 0.$$

В этом случае математическая модель становится нелинейной, что существенно затрудняет ее исследование.

3. Очевидно, что транспонированная матрица W^* — стохастическая, т. е. все ее элементы неотрицательны, а сумма элементов в каждом столбце равна единице (см. разд. 4.2). В связи с этим во избежание недоразумений заметим, что рассматриваемая модель формирования мнений, имеющая вид $M(t+1) = WM(t)$, не представляет собой марковский процесс, она является (в некотором смысле) «транспонированной» к нему. Другими словами, марковским процессом являлась бы модель вида $M(t+1) = W^*M(t)$. С содержательной точки зрения различие между этими двумя типами моделей заключается в следующем. Марковским процессам свойственно «забывание» начальных условий: состояние системы при $t \rightarrow \infty$ зависит от матрицы перехода (в данном случае обозначенной через W^*), но не зависит от начального вектора $M(0)$. В противоположность этому рассмотренная выше модель динамики мнений показывает, что устанавливающийся при $t \rightarrow \infty$ консенсус представляет собой линейную комбинацию от мнений в момент времени $t = 0$. Подробнее о марковских процессах см. в соответствующем разделе главы 4.

4. Модель формирования установок излагается по работе [3], см. также статью [4]. В [3] рассмотрена также нелинейная модель, аналогичная (19). Пример (21)–(23) для группы из трех индивидов предложен авторами, публикуется впервые.

5. В качестве иллюстрации к введенному понятию установки в психологии можно привести следующую цитату, принадлежащую классике психологической науки К.Г. Юнгу: «Хотя общие психологические законы имеют значение для каждого индивида, однако нельзя сказать, что все они характеризуют отдельную личность, поскольку сам способ действия этих законов изменяется в соответствии с его привычной установкой. Привычная установка всегда есть результат всех факторов, способных существенно влиять на психическое, а именно: врожденного предрасположения, влияния среды, жизненного опыта, прозрений и убеждений, приобретенных путем дифференциации, коллективных представлений и др. Без такого безусловно фундаментального значения установки было бы невозможно существование индивидуальной психологии» (см. [5]).

Литература

1. *DeGroot M.H.* Reaching a Consensus // *Journal of the American Statistical Association*, V. 69. Issue 345 (Mar 1974). P. 118–121.
2. *Friedkin N.* A Formal Theory of Social Power // *J. of Mathematical Sociology*. 1986. V. 12(2). P. 103–126.
3. *Ефимов Б.А.* Математическое моделирование формирования установок участников социальной группы // *Математическое моделирование социальных процессов*. Выпуск 10. 2008. Москва. КДУ. С. 67–122.
4. *Гаврилец Ю.Н., Ефимов Б.А.* Изменение предпочтений индивидов в социальной среде // *Экономика и матем. методы*, 1997, т. 33, № 3.
5. *Юнг К.Г.* Психологические типы. М.: Прогресс, 1995.

3.5. Социальные взаимодействия в малой группе

Настоящий раздел посвящен модели взаимодействий в социальной группе, связывающей такие ее характеристики, как уровень дружелюбия, интенсивность деятельности, интенсивность социальных взаимодействий между членами группы.

Главное отличие данной модели от балансовых моделей (см. следующий раздел) заключается в том, что здесь не рассматривается внутренняя топология группы, т. е. не рассматриваются попарные отношения между индивидами (наличие или отсутствие взаимодействий в данной паре). В частности это означает, что уровень дружелюбия, интенсивность деятельности и интенсивность социальных взаимодействий характеризуют группу в целом (а не отдельных индивидов или составленные из них пары).

Первые два пункта настоящего раздела посвящены варианту модели социальных взаимодействий в малой группе, разработанному американским математиком Г. Саймоном. В первом проводится построение этого варианта модели на основе гипотез, отражающих основные положения теории социальных взаимодействий, хорошо известной в психологии. Во втором показана некорректность модели Г. Саймона. Третий пункт посвящен корректировке модели, т. е. построению и исследованию ее «исправленного» варианта, основанного на тех же гипотезах, но конкретизирующего их в иной математической форме. Наконец, в четвертом кратко очерчено одно из возможных направлений развития данной модели.

3.5.1. Традиционная модель

В основе модели лежат следующие представления.

Рассматривается замкнутая группа индивидов, взаимодействующих между собой и участвующих в некоторой общей деятельности. Внешняя, по отношению к группе, среда предъявляет определенные требования по интенсивности этой деятельности. При этом, хотя сама деятельность требует некоторых взаимодействий между индивидами,

одновременно, взаимодействия происходят и помимо нее, из дружеского отношения между членами группы.

Более подробно механизмы, лежащие в основе модели, описаны ниже, в форме соответствующих гипотез. Социальная группа характеризуется следующими четырьмя переменными, каждая из которых является функцией времени:

- интенсивность социальных взаимодействий между членами группы $I(t)$;
- уровень дружелюбия между членами группы $F(t)$,
- интенсивность (трудовой или иной, связанной с социальными взаимодействиями) деятельности, выполняемой членами группы $A(t)$,
- интенсивность деятельности, требуемая от группы внешней средой $E(t)$ — эта переменная рассматривается как экзогенная.

Всюду далее будем считать, что эти переменные принимают значения в неотрицательной области.

Гипотеза 1. Взаимодействия в группе тем более интенсивны, чем выше уровень дружелюбия и интенсивность производимой деятельности. При этом будем считать, что данная зависимость является мгновенной, т. е. временные лаги отсутствуют.

Для математической формализации выберем конкретную зависимость

$$I(t) = \alpha_1 F(t) + \alpha_2 A(t), \quad (1)$$

где $\alpha_1 > 0$, $\alpha_2 > 0$ — постоянные.

Гипотеза 2. Каждому уровню дружелюбия F соответствует определенное значение интенсивности социальных взаимодействий $I^*(F)$, необходимое для его поддержания. Если в течение интервала времени $(t, t + \Delta t)$ интенсивность недостаточна, $I < I^*(F)$, то дружелюбие убывает, $dF/dt < 0$; если же интенсивность превышает этот необходимый уровень ($I > I^*(F)$), то дружелюбие возрастает.

Конкретизируя данную гипотезу, примем $I^*(F) = \beta F$, где $\beta > 0$ — постоянная. Имеем тогда

$$\Delta F = b(I - \beta F)\Delta t.$$

Здесь b — постоянная, характеризующая скорость изменения дружелюбия в группе. Проводя предельный переход $\Delta t \rightarrow 0$, получаем дифференциальное уравнение

$$\frac{dF}{dt} = b(I - \beta F). \quad (2)$$

Гипотеза 3. Факторами, влияющими на изменение интенсивности деятельности группы $A(t)$, являются уровень дружелюбия $F(t)$ и уровень интенсивности, требуемый внешней средой $E(t)$.

Каждому значению дружелюбия F соответствует некоторое значение интенсивности деятельности $A^*(F)$ такое, что ее превышение, $A(t) > A^*(F)$, способствует убыванию интенсивности; недостаточность $A(t) < A^*(F)$ — возрастанию.

Превышение текущей интенсивности деятельности $A(t)$ над интенсивностью $E(t)$, требуемой со стороны внешней среды, способствует ее уменьшению, недостаточность до требуемой интенсивностью — повышению.

Конкретизируя данную гипотезу, примем

$$\Delta A = [c_1(F - \gamma A) + c_2(E(t) - A)] \Delta t,$$

где $\gamma > 0$, $c_1 > 0$, $c_2 > 0$ — постоянные.

Проводя предельный переход $\Delta t \rightarrow 0$, получаем дифференциальное уравнение

$$\frac{dA}{dt} = c_1(F - \gamma A) + c_2(E(t) - A). \quad (3)$$

Уравнения (1)–(3) дополняются начальными условиями

$$F(0) = F_0, \quad A(0) = A_0. \quad (4)$$

Тем самым построена математическая модель социальных взаимодействий в группе индивидов, имеющая вид задачи Коши (1)–(4). Следующий пункт посвящен исследованию этой модели методами теории дифференциальных уравнений.

3.5.2. Анализ традиционной модели и выявление ее проблемных мест

Подставив (1) в (2), после приведения подобных слагаемых получим систему двух обыкновенных дифференциальных уравнений

$$\frac{dF}{dt} = b(\alpha_1 - \beta)F + \alpha_2 bA, \quad (5)$$

$$\frac{dA}{dt} = c_1 F - (\gamma c_1 + c_2)A + c_2 E(t). \quad (6)$$

Проанализируем наиболее простой случай, когда требуемая внешней средой интенсивность деятельности постоянна во времени: $E(t) = E_0$ (см. также упражнение 1).

Соответствующая стационарная задача имеет вид

$$b(\alpha_1 - \beta)F + \alpha_2 bA = 0, \quad (7)$$

$$c_1 F - (\gamma c_1 + c_2)A = -c_2 E_0. \quad (8)$$

Если система не вырождена, т. е.

$$\Delta = b[(\beta - \alpha_1)(\gamma c_1 + c_2) - \alpha_2 c_1] \neq 0, \quad (9)$$

то стационарное решение единственно и дается формулами

$$\begin{aligned} F_{st} &= E_0 \frac{c_2 \alpha_2}{(c_1 \gamma + c_2)(\beta - \alpha_1) - c_1 \alpha_2}, \\ A_{st} &= E_0 \frac{c_2(\beta - \alpha_1)}{(c_1 \gamma + c_2)(\beta - \alpha_1) - c_1 \alpha_2}. \end{aligned} \quad (10)$$

Общее решение системы (5), (6) имеет вид

$$\begin{pmatrix} F(t) \\ A(t) \end{pmatrix} = C_1 \psi_1 e^{\lambda_1 t} + C_2 \psi_2 e^{\lambda_2 t} + \begin{pmatrix} F_{st} \\ A_{st} \end{pmatrix}, \quad (11)$$

где ψ_i , λ_i , $i = 1, 2$, — собственные векторы и собственные значения матрицы системы (5), (6). Чтобы найти их, выпишем ее характеристическое уравнение,

$$\begin{vmatrix} b(\alpha_1 - \beta) - \lambda & \alpha_2 b \\ c_1 & -(\gamma c_1 + c_2) - \lambda \end{vmatrix} = 0,$$

которое после несложных преобразований принимает вид

$$\lambda^2 + \lambda[(\gamma c_1 + c_2) - b(\alpha_1 - \beta)] + \Delta = 0.$$

Решая его, получаем собственные значения

$$\lambda_{1,2} = \frac{1}{2} \left\{ -[(\gamma c_1 + c_2) + b(\beta - \alpha_1)] \pm \sqrt{[(\gamma c_1 + c_2) + b(\beta - \alpha_1)]^2 - 4\Delta} \right\}. \quad (12)$$

Выясним, при каких значениях параметров имеют место неравенства $\text{Re } \lambda_1 < 0$, $\text{Re } \lambda_2 < 0$, обеспечивающие асимптотическую устойчивость стационарного решения (10). Для выполнения этих неравенств необходимо и достаточно выполнения соотношений

$$(\gamma c_1 + c_2) + b(\beta - \alpha_1) > 0, \quad (13)$$

$$\Delta = b[(\beta - \alpha_1)(\gamma c_1 + c_2) - \alpha_2 c_1] > 0. \quad (14)$$

Решая систему неравенств (13), (14), получаем

$$\beta - \alpha_1 > \frac{\alpha_2 c_1}{\gamma c_1 + c_2}. \quad (15)$$

Итак, необходимым и достаточным условием отрицательности собственных значений (12) является выполнение неравенства (15). Отрицательность собственных значений означает, что экспоненциальные слагаемые в (11) убывают с течением времени. Таким образом, если выполняется условие (15), то имеет место асимптотика

$$\begin{pmatrix} F(t) \\ A(t) \end{pmatrix} \rightarrow \begin{pmatrix} F_{st} \\ A_{st} \end{pmatrix} \text{ при } t \rightarrow +\infty, \quad (16)$$

другими словами — стационарное решение (10) является асимптотически устойчивым. Это же условие обеспечивает положительность величин F_{st} , A_{st} .

Перейдем к анализу случая, когда условие (15) не выполняется. Далее будем полагать, что

$$\beta - \alpha_1 < \frac{\alpha_2 c_1}{\gamma c_1 + c_2}. \quad (17)$$

Тогда $\Delta < 0$, и собственные значения λ_1, λ_2 имеют разные знаки. Собственный вектор $\psi_1 = \begin{pmatrix} \psi_{11} \\ \psi_{12} \end{pmatrix}$, соответствующий положительному собственному значению

$$\lambda_1 = \frac{1}{2} \left\{ -[(\gamma c_1 + c_2) + b(\beta - \alpha_1)] + \sqrt{[(\gamma c_1 + c_2) + b(\beta - \alpha_1)]^2 - 4\Delta} \right\}, \quad (18)$$

находится из системы уравнений

$$\begin{pmatrix} b(\alpha_1 - \beta) & \alpha_2 b \\ c_1 & -(\gamma c_1 + c_2) \end{pmatrix} \begin{pmatrix} \psi_{11} \\ \psi_{12} \end{pmatrix} = \lambda_1 \begin{pmatrix} \psi_{11} \\ \psi_{12} \end{pmatrix}$$

и имеет вид

$$\psi_1 = \begin{pmatrix} (\gamma c_1 + c_2) + \lambda_1 \\ c_1 \end{pmatrix}. \quad (19)$$

Следовательно, при достаточно больших значениях t решение системы дифференциальных уравнений (5), (6) имеет асимптотику

$$\begin{pmatrix} F(t) \\ A(t) \end{pmatrix} = C_1 \psi_1 e^{\lambda_1 t} + \begin{pmatrix} F_{st} \\ A_{st} \end{pmatrix}, \quad (20)$$

где постоянная C_1 определяется из начальных условий.

Таким образом, если параметры системы удовлетворяют неравенству (17), то в рамках рассматриваемой модели (1)–(4) уровень дружелюбия в группе $F(t)$ и интенсивность выполняемой ей деятельности, $A(t)$, неограниченно увеличиваются с течением времени. Этот неограниченный рост происходит, невзирая на то, что требуемая внешней средой интенсивность деятельности постоянна во времени: $E(t) = E_0$.

Наличие данного свойства является серьезным недостатком построенной в п. 3.5.1 модели, поскольку для реальных групп индивидов могут реализовываться различные значения параметров, в т. ч. и те, при которых выполняется неравенство (17), вследствие чего в рамках модели происходит неограниченный рост вида (20).

Более внимательный анализ позволяет обнаружить и другие проблемы, позволяющие утверждать, что модель (1)–(4) некорректна. Именно, из формул (10) следует, что $A_{st} = F_{st}(\beta - \alpha_1)/\alpha_2$, т. е. при определенных значениях параметров системы величины F_{st} и A_{st} не могут быть одновременно положительны (и, следовательно, интенсивность взаимодействий I является отрицательной). Аналогично, интенсивность деятельности A_{st} имеет, при некоторых значениях параметров, знак, противоположный знаку требуемой извне интенсивности деятельности E_0 .

Поэтому необходимо внести изменения в модель, чтобы исправить ее логические недостатки и тем самым приблизить ее к реальности. При этом необходимо оставить в силе основные качественные поло-

жения модели, поскольку они базируются на основе хорошо известной в психологии теории социальных взаимодействий.

3.5.3. Построение и анализ скорректированной модели

Чтобы исправить выявленные недостатки модели, сделаем два следующих замечания.

Замечание 1. При конкретизации гипотезы 1 была выбрана формула (1), учитывающая влияние двух факторов (дружелюбия и интенсивности производимой деятельности) на интенсивность взаимодействий аддитивным образом: $I(t) = \alpha_1 F(t) + \alpha_2 A(t)$. Такая форма соответствует независимому действию этих факторов в том смысле, что прирост, например дружелюбия, на величину ΔF приведет к увеличению интенсивности взаимодействий на величину $\Delta I = \alpha_1 \Delta F$, не зависящую от текущего значения $A(t)$. Однако в реальных ситуациях подобная независимость факторов, видимо, не имеет места: величина ΔI должна быть тем больше, чем выше $A(t)$. Другими словами, если интенсивность деятельности в группе низкая, то увеличение дружелюбия ΔF приведет лишь к небольшому увеличению интенсивности взаимодействий ΔI . И наоборот: если интенсивность деятельности $A(t)$ достаточно велика, то и прирост интенсивности взаимодействий будет высоким.

Другая характерная особенность формулы (1) состоит в том, что она предполагает постоянную отдачу от факторов: например, увеличение дружелюбия $\Delta F = 0,1$ с величины $F_0 = 0,3$ до $F_0 + \Delta F = 0,4$ приводит к такому же увеличению интенсивности взаимодействий, что и увеличение с величины $F_0 = 5,3$ до $F_0 + \Delta F = 5,4$. В формальной математической записи это имеет вид $\partial^2 I / \partial F^2 = 0$. Однако более обоснованным выглядит предположение об убывающей отдаче от факторов: $\partial^2 I / \partial F^2 < 0$, $\partial^2 I / \partial A^2 < 0$. Именно, при постоянном значении одного из факторов перманентное увеличение другого фактора приводит к замедляющемуся росту интенсивности взаимодействий.

Отметим, наконец, ценное с содержательной точки зрения свойство формулы (1), которое необходимо сохранить при корректировке модели: при увеличении каждого из факторов в некоторое число раз k интенсивность взаимодействий увеличится в то же число раз, т. е. $I(kF, kA) = kI(F, A)$.

На основе изложенных соображений можно предложить следующую конкретизацию гипотезы 1:

$$I(t) = I_0 F^\alpha(t) A^{1-\alpha}(t), \quad (21)$$

где $I_0 > 0$, $0 < \alpha < 1$ — некоторые постоянные.

Ниже, при анализе корректной модели будет показано, что отказ от аддитивной формулы (1) решает одну из главных проблем модели Г. Саймона: существования лишь таких стационарных решений, в которых значения F_{st} и A_{st} имеют (при некоторых значениях параметров) разные знаки.

Замечание 2. Модель (1)–(4) не учитывает, что для любой группы индивидов существует определенный «физиологический» верхний предел интенсивности деятельности, не зависящий ни от уровня дружелюбия, ни от интенсивности социальных взаимодействий. Обозначим его через A_{\max} . Если с течением времени интенсивность деятельности монотонно увеличивается, $dA/dt > 0$, то с приближением к значению $A = A_{\max}$ должно наступать насыщение так, что $d^2A/dt^2 < 0$ и $\lim_{t \rightarrow +\infty} A(t) = A_{\max}$.

С учетом этого при сохранении психологического содержания гипотезы 3, ее можно конкретизировать в виде

$$\frac{dA}{dt} = [c_1(F - \gamma A) + c_2(E(t) - A)](A_{\max} - A), \quad A(0) = A_0. \quad (22)$$

При этом будем полагать, что как начальное значение, так и требуемая внешней средой интенсивность деятельности E_0 ниже «физиологического» предела:

$$A(0) < A_{\max}, \quad E_0 < A_{\max}. \quad (23)$$

Уравнение, выражающее гипотезу 2 (с соответствующим начальным условием),

$$\frac{dF}{dt} = b(I - \beta F), \quad F(0) = F_0, \quad (24)$$

остается в скорректированной модели без изменений.

Таким образом, полученная модель имеет вид системы уравнений, начальных и дополнительных условий (21)–(24). Проведем ее анализ подобно тому, как это было сделано для исходной модели (1)–(4).

Подставим (21) в (24) и приведем подобные слагаемые в квадратной скобке уравнения (22). Получим систему

$$\frac{dF}{dt} = b(I_0 F^\alpha A^{1-\alpha} - \beta F), \quad (25)$$

$$\frac{dA}{dt} = [c_1 F - (\gamma c_1 + c_2)A + c_2 E(t)](A_{\max} - A). \quad (26)$$

Как и для исходной модели, ограничим анализ случаем, когда требуемая внешней средой интенсивность деятельности постоянна во времени: $E(t) = E_0$. Соответствующая стационарная задача (25), (26) имеет вид

$$b(I_0 F^\alpha A^{1-\alpha} - \beta F) = 0, \quad (27)$$

$$[c_1 F - (\gamma c_1 + c_2)A + c_2 E_0](A_{\max} - A) = 0. \quad (28)$$

Из (27) получаем, что

$$F_{st} = A_{st}(I_0/\beta)^{1/(1-\alpha)}. \quad (29)$$

Подставляя полученное соотношение в (28), после элементарных выкладок получаем, что либо

$$A_{st} = A_{\max}, \quad F_{st} = A_{\max}(I_0/\beta)^{1/(1-\alpha)}, \quad (30)$$

либо

$$\begin{aligned} A_{st} &= \frac{c_2 E_0}{(\gamma c_1 + c_2) - c_1 (I_0/\beta)^{1/(1-\alpha)}}, \\ F_{st} &= \frac{c_2 E_0}{(\gamma c_1 + c_2) - c_1 (I_0/\beta)^{1/(1-\alpha)}} \left(\frac{I_0}{\beta} \right)^{1/(1-\alpha)} \end{aligned} \quad (31)$$

(причем решение (31) существует, лишь если $(\gamma c_1 + c_2) - c_1 (I_0/\beta)^{1/(1-\alpha)} \neq 0$; см. также упражнение 2 в конце раздела).

Изучим устойчивость каждого из полученных стационарных решений. Покажем, что при любых значениях параметров (за исключением особого случая, при котором не применим метод исследования устойчивости по первому приближению) ровно одно из двух стационарных решений (30), (31) асимптотически устойчиво, а второе будет неустойчиво.

В соответствии с методом исследования устойчивости по первому приближению вычислим якобиан системы (25), (26). Получим

$$J = \begin{pmatrix} b[\alpha I_0 F^{\alpha-1} A^{1-\alpha} - \beta] & b(1-\alpha) I_0 F^\alpha A^{-\alpha} \\ c_1 (A_{\max} - A) & -(\gamma c_1 + c_2)(A_{\max} - A) - [c_1 F - (\gamma c_1 + c_2)A + c_2 E_0] \end{pmatrix}.$$

Для стационарного решения (30) задача на собственные значения имеет вид

$$\begin{vmatrix} b\beta(\alpha - 1) - \lambda & b(1-\alpha) I_0 (I_0/\beta)^{\alpha/(1-\alpha)} \\ 0 & [A_{\max}((\gamma c_1 + c_2) - c_1 (I_0/\beta)^{1/(1-\alpha)}) - c_2 E_0] - \lambda \end{vmatrix} = 0.$$

Очевидно, один из корней

$$\lambda_1 = b\beta(\alpha - 1) \quad (32)$$

отрицателен при любых значениях параметров, допустимых в данной модели, в то время как знак второго собственного значения,

$$\lambda_2 = A_{\max}((\gamma c_1 + c_2) - c_1 (I_0/\beta)^{1/(1-\alpha)}) - c_2 E_0, \quad (33)$$

зависит от этих параметров. Если $\lambda_2 < 0$, то стационарное решение (30) асимптотически устойчиво, если $\lambda_2 > 0$, то неустойчиво.

Перейдем к рассмотрению стационарного решения (31). Имеем (собственные значения обозначены здесь буквой μ)

$$\begin{vmatrix} b\beta(\alpha - 1) - \mu & b(1-\alpha) I_0 (I_0/\beta)^{\alpha/(1-\alpha)} \\ c_1 A_{\max} \left(1 - \frac{c_2 E_0}{\lambda_2 + c_2 E_0} \right) & -(\gamma c_1 + c_2) A_{\max} \left(1 - \frac{c_2 E_0}{\lambda_2 + c_2 E_0} \right) - \mu \end{vmatrix} = 0. \quad (34)$$

Здесь учтено, что

$$(\gamma c_1 + c_2) - c_1 (I_0/\beta)^{1/(1-\alpha)} = \frac{\lambda_2 + c_2 E_0}{A_{\max}} \quad (35)$$

(данное равенство следует из (33)). После несложных, но громоздких преобразований получаем из (34)

$$\begin{aligned} \mu^2 + \mu \left\{ (\gamma c_1 + c_2) A_{\max} \left(1 - \frac{c_2 E_0}{\lambda_2 + c_2 E_0} \right) + b\beta(1 - \alpha) \right\} + \\ + A_{\max} \left(1 - \frac{c_2 E_0}{\lambda_2 + c_2 E_0} \right) b(1 - \alpha)\beta \{ (\gamma c_1 + c_2) - c_1 (I_0/\beta)^{1/(1-\alpha)} \} = 0. \end{aligned}$$

Отсюда, еще раз воспользовавшись соотношением (35), получим квадратное уравнение для μ , которое позволит сделать вывод об устойчивости:

$$\mu^2 + \mu \left\{ (\gamma c_1 + c_2) \frac{\lambda_2 A_{\max}}{\lambda_2 + c_2 E_0} + b\beta(1 - \alpha) \right\} + b(1 - \alpha)\beta\lambda_2 = 0. \quad (36)$$

Из теоремы Виета следует, что если корни μ_1 , μ_2 уравнения (36) вещественны, то они удовлетворяют равенствам

$$\mu_1 + \mu_2 = - \left\{ b\beta(1 - \alpha) + \frac{A_{\max}(\gamma c_1 + c_2)}{\lambda_2 + c_2 E_0} \lambda_2 \right\}, \quad (37)$$

$$\mu_1 \mu_2 = b\beta(1 - \alpha)\lambda_2. \quad (38)$$

Таким образом, имеем следующие варианты.

1. Если $\lambda_2 < 0$ (т.е. стационарное решение (30) асимптотически устойчиво), то дискриминант уравнения (36) положителен; в соответствии с теоремой Виета имеем $\mu_1 \mu_2 < 0$. Следовательно, ровно одно из значений μ_1 , μ_2 положительно, откуда следует неустойчивость стационарного решения (31).

2. Если $\lambda_2 > 0$ (т.е. стационарное решение (30) неустойчиво), то дискриминант уравнения (36) может быть как положительным, так и отрицательным. В первом случае из теоремы Виета (37), (38) следует, что $\mu_1 \mu_2 > 0$, $\mu_1 + \mu_2 < 0$; следовательно, $\mu_1 < 0$, $\mu_2 < 0$. Если же дискриминант отрицателен, то $\operatorname{Re} \mu_1 = \operatorname{Re} \mu_2 < 0$ ввиду того, что в квадратном уравнении (36) положителен коэффициент при μ . Таким образом, стационарное решение (31) асимптотически устойчиво.

Итак, приведенное рассуждение доказывает, что если стационарное решение (30) асимптотически устойчиво, то стационарное решение (31) неустойчиво, и наоборот. Первый случай имеет место при $\lambda_2 < 0$, второй — при $\lambda_2 > 0$. Если же $\lambda_2 = 0$, то метод исследования устойчивости по первому приближению не позволяет сделать вывод об устойчивости решений.

Таким образом, скорректированная модель (21)–(24) является корректной: она не содержит тех недостатков, которые свойственны исходной модели (1)–(4).

3.5.4. Учет структуры группы

В качестве одного из направлений развития модели социальных взаимодействий возможно построение модели группы, содержащей внутри себя клан, т.е. «подгруппу» тесно связанных членов группы,

имеющих свои цели, отличные от целей группы в целом. Для осуществления этих целей клан осуществляет некоторую деятельность, отличную от деятельности группы в целом. Обозначим группу через G_1 (ей соответствуют переменные I_1, F_1, A_1, E_1) и клан (ее подгруппу) через G_2 (ему соответствуют переменные I_2, F_2, A_2, E_2). Уравнения соответствующей модели могут иметь вид

$$\frac{dA_1}{dt} = \psi_1(A_1, F_1, A_2, E_1), \quad (39)$$

$$\frac{dF_1}{dt} = \varphi_1(A_1, F_1), \quad (40)$$

$$\frac{dA_2}{dt} = \psi_2(A_2, F_2, A_1, E_2), \quad (41)$$

$$\frac{dF_2}{dt} = \varphi_2(A_2, F_2), \quad (42)$$

где $\psi_i, \varphi_i, i = 1, 2$, — функции, конкретная форма которых должна отражать дополнительные гипотезы модели. Например, если клан оказывает деструктивное влияние на деятельность группы в целом, то $\partial\psi_1/\partial A_2 < 0$. Сформулировав эти дополнительные гипотезы, можно завершить построение модели группы, содержащей клан и исследовать ее методами теории дифференциальных уравнений.

Упражнения

1. Проанализировать модель (21)–(24) в случае, когда интенсивность деятельности, требуемая от группы внешней средой, экспоненциально изменяется с течением времени: $E(t) = E_0 \exp(\nu t)$.

2. Показать, что если система (21)–(24) имеет только одно стационарное решение (см. замечание после формулы (31)), то оно асимптотически устойчиво.

3. Для модели социальных взаимодействий в группе, содержащей внутри себя клан, сформулировать дополнительные гипотезы и конкретизировать их, предложив конкретные функции ψ_i, φ_i . Проанализировать полученную систему уравнений.

Комментарии и библиографические примечания

1. Исходная (изложенная в п. 3.5.1 и п. 3.5.2) модель взаимодействий в социальной группе была предложена Г. Саймоном в работе [1] и является формализацией теории социального обмена Дж. Хоманса [2, 3]. Моделирование группы, содержащей клан, с помощью уравнений вида (39)–(42) также предложено в работе [1].

2. Скорректированная модель (21)–(24) предложена авторами, публикуется впервые.

Литература

1. *Simon H.A.* A Formal Theory of Interaction in Social Groups // *American Sociological Review*. V. 17, №2, Apr., 1952. P. 202–211.
2. *Homans G.C.* The Human Group. N.Y., 1950
3. *Хоманс Дж.* Социальное поведение как обмен // Современная зарубежная социальная психология. М.: Издательство Московского университета, 1984. С. 82–91.

3.6. Структурный баланс в малой группе

Настоящий раздел посвящен моделированию установившихся, статических отношений между индивидами в малой группе. В данном контексте группа считается малой, если каждый ее член имеет определенное мнение (положительное либо отрицательное) о каждом другом члене. Рассматриваемый подход основан на широко известной в психологии теории структурного баланса, в соответствии с которой индивид стремится устранить противоречия между своими оценками (мнениями, отношениями) других индивидов и прочих объектов. В качестве примера можно привести противоречие между оценками «мне не нравится этот человек» и «мне импонируют его политические взгляды».

В качестве другого примера можно привести диссонанс между отношениями «этот человек — мой враг» и «этот человек — друг моего друга». Наличие подобных противоречий приводит к изменению мнений и оценок с течением времени, в частности, к изменению отношений между индивидами. Эта динамика направлена в сторону уменьшения противоречий: в первом из приведенных примеров будет либо улучшаться «мое мнение об этом человеке», либо ухудшаться «мое мнение о данной политической позиции» (либо, в некоторой степени, и то, и другое). В результате возникает так называемое *сбалансированное состояние* индивида, при котором отсутствуют противоречия между отдельными отношениями. Если все члены малой группы адекватно воспринимают отношения между собой, то сбалансированность индивидов возможна лишь при определенной структуре отношений в группе. Группы, имеющие такую структуру, называются сбалансированными.

Стационарная модель, основанная на данной психологической теории, составляет содержание настоящего раздела. В первом пункте излагается исторически более ранний ХКХ-подход, основанный на определенном наборе аксиом относительно условий, при которых конкретный индивид находится в состоянии когнитивного баланса. Во втором пункте рассматривается так называемый *Д-подход*, основанный на несколько ином наборе аксиом. В обоих случаях изучается вопрос о том, какую структуру должна иметь малая группа, чтобы все ее члены находились в состоянии когнитивного баланса. В третьем пункте вводится понятие социоматрицы.

В следующих двух разделах настоящей главы рассматриваются динамические модели, развивающие концепцию структурного баланса.

3.6.1. ХКХ-подход к структурному балансу

Предметом теории структурного баланса являются отношения членов малой группы (индивидов). Эти отношения визуализируются с помощью диаграмм, имеющих вид знаковых графов, вершины которых соответствуют членам группы, а ребра — попарным отношениям между ними. Знаки ребер имеют следующий смысл:

«+» — индивиды относятся друг к другу положительно, «друзья»,

«-» — индивиды относятся друг к другу отрицательно, «враги».

Таким образом, здесь предполагается, что все отношения являются симметричными в том смысле, что индивид А относится к индивиду В так же, как индивид В относится к индивиду А. Исключается ситуация, при которой А считает индивида В другом, а В считает А врагом.

В качестве примера приведем граф, представленный на рис. 1. Здесь индивиды А и В, а также С и D образуют две пары друзей, причем отношения между этими парами враждебные (отрицательные).

В данном разделе рассматриваются малые группы, т. е. такие, в которых каждая пара индивидов связана положительным или отрицательным отношением. Граф малой группы является полным, т. е. каждые две его вершины соединены ребром.

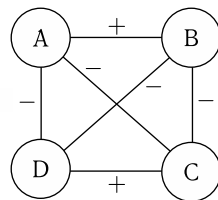


Рис. 1. Пример знакового графа, отображающего отношения в группе из четырех индивидов

Замечание 1. Теория структурного баланса рассматривает также системы, в которых изучаются не только отношения членов группы друг к другу, но и их отношения к некоторым объектам. Данное понятие имеет широкий смысл — объектом может быть какой-либо законопроект, широко обсуждаемое социальное явление, известный политик и т. д. Такие системы также визуализируются в виде знаковых графов. Например, вершина D на рис. 1 может представлять объект, к которому индивиды А и В относятся отрицательно, а С — положительно. Ниже будет рассмотрена динамическая модель, в которой роль объектов играют политические позиции индивидов.

Любых трех членов группы с отношениями между ними (а также соответствующий граф) будем называть триадой.

В соответствии с базовым положением теории структурного баланса *входящий в группу индивид находится в состоянии когнитивного баланса, если для каждой триады, в которую он входит, его отношения с каждым из других членов этой триады, будучи*

выраженными через два других отношения в этой триаде, имеют одну из следующих форм:

(R1) друг моего друга — мой друг,

(R2) враг моего врага — мой друг,

(R3) друг моего врага — мой враг,

(R4) враг моего друга — мой враг.

Для иллюстрации этого положения покажем, что индивид А на рис. 1 находится в состоянии когнитивного баланса. Действительно, он входит в следующие триады: ABC, ABD, ACD. Рассмотрим каждую из них.

- Триада ABC: в ней индивид В — это друг, индивид С — враг, причем В и С — враги. Таким образом, В есть «враг моего врага» и одновременно «мой друг», а С — «враг моего друга», и одновременно «мой враг».
- Триада ABD: в ней индивид В есть «враг моего врага — мой друг», а D — «враг моего друга — мой враг».
- Триада ACD: в ней индивид С есть «друг моего врага — мой враг», и D — «друг моего врага — мой враг».

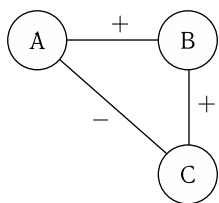


Рис. 2. Каждый из индивидов А, В, С находится в состоянии когнитивного дисбаланса

Еще некоторый член группы входит хотя бы в одну триаду с другой системой отношений, то он находится в состоянии когнитивного дисбаланса; при этом в системе возникают силы и напряжения, направленные на восстановление баланса. Например в ситуации, изображенной на рис. 2, эти силы приведут либо к изменению отношений между А и С на положительные, либо в одной из других пар приведут на отрицательные.

Группа называется структурно сбалансированной, если каждый ее член находится в состоянии когнитивного баланса.

Замечание 2. Строго говоря, реально существующие в группе отношения могут не соответствовать их восприятию индивидами. Например в ситуации, изображенной на рис. 2, индивид А может ошибочно полагать, что между В и С существует неприязнь, т.е. отрицательное отношение. В этом случае он находится в состоянии когнитивного баланса (а каждый из индивидов В и С, правильно воспринимающих все отношения в триаде, — в состоянии дисбаланса). Предметом теории структурного баланса, в наиболее строгом и узком смысле, является восприятие отношений индивидами, а не сами отношения. В настоящем разделе принимается допущение, что все индивиды воспринимают отношения правильно (и, следовательно, одинаково). Это допущение позволяет сделать переход в построении теории с микроуровня на макроуровень, именно — опираясь на понятие сбалансированного индивида, перейти к вопросу о том, какая структура свойственна сбалансированной группе.

Построим индуктивным образом пример структурно сбалансированной группы.

Пусть, например, два члена группы А и В связаны отрицательным отношением (рис. 3).

После добавления в группу третьего члена С каждый из индивидов будет находиться в состоянии когнитивного баланса лишь в том случае, если отношения АС и ВС будут иметь разные знаки.

Действительно, если оба отношения АС и ВС положительны, то индивид С является для индивида А другом и в то же время — «другом моего врага В». Таким образом, индивид А находится в состоянии когнитивного дисбаланса. Нетрудно показать (см. упражнение 1), что В и С также находятся в этом случае в состоянии когнитивного дисбаланса.

Если же оба отношения АС и ВС отрицательны, то индивид С является для А врагом и в то же время — «врагом моего врага В», что приводит к когнитивному дисбалансу индивида А; аналогично — двух других членов группы.

Итак, при добавлении индивида С к группе, изображенной на рис. 3, полученная группа будет сбалансированной лишь если отношения АС и ВС имеют разные знаки. Примем для определенности, что отношение АС отрицательно (рис. 4).

Четвертый член группы D (а также все следующие) будет находиться в состоянии когнитивного баланса лишь в случае, если отношение AD будет иметь знак, противоположный знаку отношений BD, CD. Другими словами, в ситуации, изображенной на рис. 4, можно говорить о том, что сформировались две подгруппы, А и ВС; причем после добавления четвертого члена все индивиды будут находиться в состоянии когнитивного баланса, лишь если D примкнет к одной из подгрупп, т. е. будет иметь положительные отношения со всеми членами этой подгруппы и отрицательные отношения со всеми членами другой. Предположим, для определенности, что отношение AD положительно, а BD и CD отрицательны. Получим ситуацию, представленную на рис. 5.

Далее, при добавлении пятого члена группы Е все индивиды будут находиться в состоянии когнитивного баланса, лишь если Е примкнет к одной из уже существующих групп AD или ВС, т. е. если знак отношений АЕ, DE будет противоположен знаку отношений ВЕ, СЕ. С добавлением каждого следующего члена когнитивный баланс индивидов также будет сохраняться лишь при условии существования двух

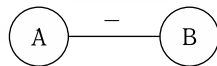


Рис. 3. Два индивида, связанные отрицательным отношением

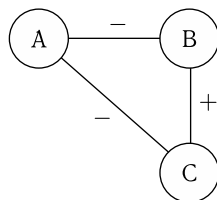


Рис. 4. Группа является сбалансированной, при этом В и С «дружат против» А

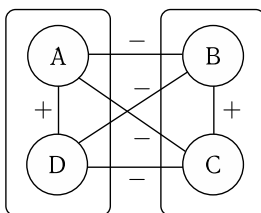


Рис. 5. Группа является сбалансированной и содержит две антагонистические подгруппы AD и BC

подгрупп, члены каждой из которых «дружат против» членов другой подгруппы.

Данные соображения приводят к целесообразности введения следующего определения. *Если малая группа разделена на подгруппы так, что каждый индивид связан положительными отношениями со всеми членами своей подгруппы, а отрицательными — со всеми остальными индивидами, то эти подгруппы называются антагонистическими.*

Таким образом, на примере было продемонстрировано, что малая группа является структурно сбалансированной лишь при наличии в ней двух антагонистических подгрупп (для общности будем считать, что одна из подгрупп может состоять из пустого множества членов, т. е. все отношения в рассматриваемой малой группе положительны).

Граф, соответствующий структурно сбалансированной группе, обладает следующим свойством: множество его вершин можно разбить на два подмножества так, что каждое положительное ребро соединяет вершины, принадлежащие одному подмножеству, а каждое отрицательное ребро — разным. Граф, удовлетворяющий этому свойству, называется сбалансированным по Хайдеру–Картрайту–Харари

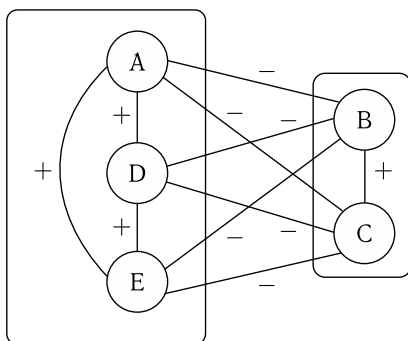


Рис. 6. Структурно сбалансированная группа из пяти членов, отображенная в виде ХКХ-сбалансированного знакового графа

или ХКХ-сбалансированным (см. также комментарии к настоящему разделу), рис. 6.

Если же граф не является ХКХ-сбалансированным, то в нем всегда можно найти три вершины либо попарно связанные отрицательными отношениями (в такой ситуации каждый член группы испытывают когнитивный дисбаланс ввиду того, что для него «враг моего врага» — это «мой враг»), рис. 7, либо отношениями, представленными на рис. 2.

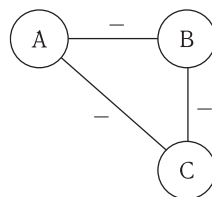


Рис. 7. Три индивида, попарно связанные отрицательными отношениями

3.6.2. Д-подход к структурному балансу

Важным обобщением ХКХ-сбалансированности является сбалансированность по Дэвису (или Д-сбалансированность).

В рамках Д-подхода входящий в группу индивид находится в состоянии когнитивного баланса, если для каждой триады, в которую он входит, его отношения с каждым из других членов этой триады, будучи выраженными через два других отношения в этой триаде, имеют либо одну из форм (R1)–(R4), либо форму

(R5) враг моего врага — мой враг.

На основе данного определения нетрудно построить пример Д-сбалансированной группы аналогично тому, как это сделано выше для ХКХ-сбалансированности. Так, при добавлении индивида С к группе, изображенной на рис. 3, имеются три возможности для построения Д-сбалансированной группы. Две из них — те же, что для примера ХКХ-сбалансированности: либо отношение АС отрицательно, отношение ВС положительно (как на рис. 4), либо наоборот — отношение АС положительно, отношение ВС отрицательно. Отношение (R5) открывает третью возможность: индивид С может быть добавлен так, что оба отношения АС, ВС отрицательны, в результате будет получена Д-сбалансированная группа, изображенная на рис. 7.

В целом, различие между построением примеров ХКХ-сбалансированной и Д-сбалансированной групп можно описать следующим образом. При добавлении каждого следующего индивида (начиная с третьего, см. рис. 3) для ХКХ-сбалансированности необходимо включить его в одну из уже существующих антагонистических подгрупп. Для Д-сбалансированности можно либо включить его в одну из антагонистических подгрупп, либо сделать его основателем новой подгруппы, антагонистической по отношению ко всем предыдущим.

В результате получим, что Д-сбалансированная группа может состоять из произвольного числа антагонистических подгрупп. Поскольку ХКХ-сбалансированная группа состоит не более, чем из двух антагонистических подгрупп, то Д-подход является в определенном смысле обобщением ХКХ-подхода. При этом каждая ХКХ-сбалансированная группа является также Д-сбалансированной.

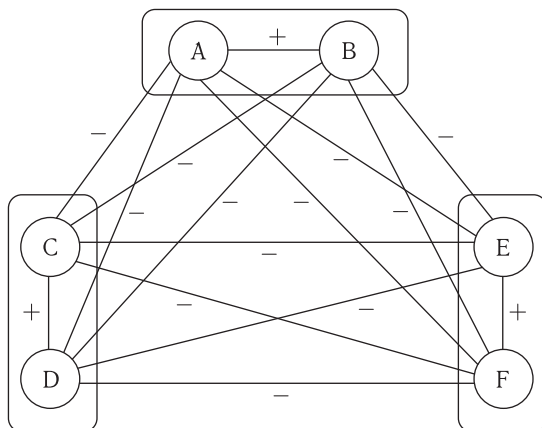


Рис. 8. Д-сбалансированный граф, соответствующий группе из шести индивидов, образующих три антагонистические подгруппы

Соответствующим образом обобщается и понятие сбалансированного графа: если множество вершин графа можно разбить на некоторое количество подмножеств так, что каждое положительное ребро соединяет вершины, принадлежащие одному подмножеству, а каждое отрицательное ребро — разным, то этот граф называется Д-сбалансированным (рис. 8). Если количество этих подмножеств не превышает двух, то граф является также и ХКХ-сбалансированным.

3.6.3. Социоматрицы

Знаковые графы позволяют описать отношения в паре индивидов лишь как «положительные» или «отрицательные», т. е. возможность описания градаций (например, крайне отрицательные, умеренно отрицательные, нейтральные и т. д.) при таком подходе отсутствует. Между тем, существующие методы социометрии (раздела психологии, предметом которого является, в частности, эмпирическое изучение межличностных отношений) позволяют измерять отношения между индивидами, получая на выходе данные в непрерывной шкале.

Именно, отношения между i -м и j -м членами группы могут быть описаны числом x_{ij} так, что $x_{ij} > 0$, если отношение позитивное, и $x_{ij} < 0$, если негативное. В наиболее простом случае отношения считаются симметричными: $x_{ij} = x_{ji}$ ($i, j = 1, \dots, N$, где N — численность малой группы), а нормировка имеет вид $|x_{ij}| \leq 1$. Отношение индивида к самому себе считается максимально положительным: $x_{ii} = 1$, $i = 1, \dots, N$. Составленная из величин x_{ij} квадратная матрица размерности N называется социоматрицей.

На основе социоматрицы может быть построен знаковый граф; при этом учитывается лишь знак отношений, но не их интенсивность. Тем самым социоматрица позволяет определить, является ли конкрет-

ная группа сбалансированной. Пусть, например, социоматрица некоторой группы имеет вид

$$S = \begin{pmatrix} 1 & 0,4 & -0,8 & -0,5 \\ 0,4 & 1 & -0,8 & -0,2 \\ -0,8 & -0,8 & 1 & -0,6 \\ -0,5 & -0,2 & -0,6 & 1 \end{pmatrix}.$$

Знаки элементов S -матрицы показывают, что все отношения в этой группе отрицательны, за исключением отношений между первым и вторым индивидами. Следовательно, знаковый граф системы имеет вид, представленный на рис. 9.

Группа состоит из трех антагонистических подгрупп: первая из них содержит двух индивидов (А и В), другие — по одному индивиду. Таким образом, матрица S описывает группу, которая является Д-сбалансированной, но не является ХКХ-сбалансированной.

Изложенная концепция структурного баланса играет большую роль в исследовании малых групп. В рамках данного направления предполагается, что устойчивыми будут лишь отношения в сообществах, структурно сбалансированных в смысле данного определения или, что то же самое, — группы, графы которых ХКХ-сбалансированы (или, в более общем случае, Д-сбалансированы). Несбалансированность ведет к изменениям отношений, к их качественной перестройке. Следующие два раздела настоящей главы посвящены динамическим моделям, описывающим эти изменения отношений.

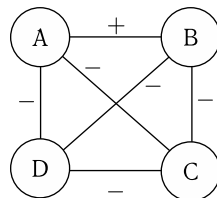


Рис. 9. Группа, соответствующая социоматрице S

Упражнения

1. Показать, что если к группе, представленной на рис. 3, добавить индивида С так, что оба отношения АС и ВС положительны, то индивиды В и С будут находиться в состоянии когнитивного дисбаланса.

2. Для группы, изображенной на рис. 8, показать, что (а) в рамках ХКХ-подхода ни один из индивидов не находится в состоянии когнитивного баланса, (б) в рамках Д-подхода все индивиды находятся в состоянии когнитивного баланса.

Комментарии и библиографические примечания

1. Концепция когнитивного баланса была предложена Ф. Хайдером в статье [1], см. также монографию [2].

2. Основным математическим результатом относительно ХКХ-сбалансированных графов является теорема, сформулированная и доказанная в работе [3]. Она называется теоремой Хайдера–Картрайта–Харари и имеет следующую формулировку.

Теорема. Для полного знакового графа следующие утверждения эквивалентны:

- а) каждый цикл в графе положителен,
- б) каждая замкнутая цепь в графе положительна,
- в) для любых двух вершин все цепи между этими вершинами имеют одинаковый знак,
- г) множество вершин графа можно разбить на два подмножества так, что каждое положительное ребро соединяет вершины, принадлежащие одному подмножеству, а каждое отрицательное ребро — разным.

3. Д-подход к сбалансированности был предложен Д. А. Дэвисом в работе [4].

4. В качестве дальнейшего чтения по тематике данного раздела может быть рекомендована глава 8 учебника [5].

Литература

1. *Heider F.* Attitudes and Cognitive Organization. // *J. of Psych.*, 21, 1946, p. 107–112.
2. *Heider F.* Theory of Interpersonal Relations. N.Y., 1958.
3. *Cartwright D. and Harary F.* Structural Balance: A Generalization of Heider's Theory. — *Psych. Rev.*, 63, 1956. P. 277–293.
4. *Davis J.A.* Clustering and structural balance in graphs // *Human Relations*. 1967. V. 20 P. 181–188.
5. *Kemeny J.G., Snell J.L.* Mathematical Models in the Social Sciences. — New York: Blaisdell Publishing Co., 1962; reprinted by M.I.T. Press, Cambridge, Mass, 1972.

3.7. Модель динамики отношений индивидов в малой группе

Предыдущий раздел настоящей главы был посвящен статическим, установившимся отношениям между индивидами в малой группе. Предметом настоящего и следующего разделов являются две динамические модели, описывающие изменение этих отношений.

В рамках первой модели рассматривается процесс, который может быть кратко описан следующим образом. Каждый индивид из числа составляющих малую группу поочередно встречается с остальными и обменивается с ними мнениями относительно других членов группы, не участвующих в данном разговоре. Этот обмен мнениями приводит к изменению межличностных отношений. Принципы, регулирующие данную динамику, подобны принципам статической модели и являются, в некотором смысле, их динамическими обобщениями. Например, в стационарной модели отношение «друг моего врага — мой враг»

считалось сбалансированным (т. е. устойчивым), а «друг моего врага — мой друг» — неустойчивым. В динамической модели этот принцип принимает форму «если мой враг отзывается о некотором индивиде положительно, то мое мнение об этом индивиде ухудшается». Зависимости такого рода позволяют построить модель, имеющую вид системы обыкновенных дифференциальных уравнений для элементов социоматрицы (см. выше), другими словами — для переменных, описывающих попарные отношения между членами группы.

Важной особенностью данной модели является то, что каждый обмен мнениями происходит в паре индивидов. Это означает, в частности, что каждый член группы делит время своей социальной активности между остальными индивидами. При этом в модели принято предположение, что чем более дружественными являются отношения между двумя членами группы, тем продолжительнее время, в течение которого они обсуждают других индивидов.

Первый пункт данного раздела посвящен построению модели, во втором и третьем пунктах анализируются стационарные отношения для триады (т. е. группы трех индивидов). Показано, в частности, что в рамках данной динамической модели устойчивы те же триады, которые являются ХКХ-сбалансированными в рассмотренной выше статической модели.

3.7.1. Основные предположения модели

Как указано в преамбуле в самых общих чертах заложенный в данную модель социально-психологический механизм выглядит следующим образом. Члены группы обмениваются мнениями относительно других (не участвующих в данном, конкретном разговоре) членов группы. Под влиянием высказанных мнений отношения между индивидами с течением времени изменяются. Для описания данной динамики и служит настоящая модель.

Рассмотрим ее основные предпосылки.

Предположение 1. Рассматривается закрытая группа, состоящая из N индивидов. С течением времени отношения в группе изменяются только под влиянием самих этих отношений. Другими словами, отсутствуют внешние влияния; в частности, индивиды не покидают и не пополняют группу.

Предположение 2. Между членами группы имеются два вида отношений: аффективные (эмоциональные) и отношения взаимовлияния.

Предположение 3. Аффективное отношение индивида i по отношению к индивиду j может быть описано скалярной непрерывной величиной X_{ij} , нормированной так, что $|X_{ij}| \leq 1$. Положительные значения соответствуют позитивным (дружеским) отношениям, отрицательные значения — негативным (враждебным) отношениям. При этом, вообще говоря, $X_{ij} \neq X_{ji}$.

Предположение 4. Отношения взаимовлияния предполагаются диадическими (бинарными, попарными) и могут быть описаны скалярной непрерывной величиной Y_{ij} , имеющей смысл времени, в течение которого индивиды i и j обсуждают других членов группы. Чем больше величина Y_{ij} , тем более продолжительным является обмен мнениями между i -м и j -м членами группы и тем сильнее эти два индивида влияют друг на друга. Непосредственно из данного определения следует, что $Y_{ij} = Y_{ji}$.

Предположение 5 (правдивость). Коммуницируя друг с другом, члены группы высказывают свои истинные мнения.

Предположение 6. Динамика аффективных отношений в группе определяется следующими зависимостями:

- «мое отношение к другому члену группы улучшается, если я слышу о нем хорошее от моего друга или плохое от моего врага; мое отношение к нему ухудшается, если я слышу о нем плохое от моего друга или хорошее от моего врага»
- «сила влияния моего собеседника на мое отношение к другим индивидам зависит от продолжительности разговора»

Предположение 7. Индивиды более склонны к общению с теми, кто им симпатичен.

Перейдем к выводу уравнений модели. При этом высказанным выше предположения будет придана более конкретная форма.

Пусть i -й и k -й члены группы обмениваются мнениями, в частности — индивид k высказывает индивиду i свое мнение X_{kj} о члене группы j . Вследствие этого изменяется отношение i к j . Именно, в соответствии с предположением 6 отношение X_{ij} i -го индивида к j -му улучшается, если X_{ik} и X_{kj} имеют один знак, и ухудшается если они имеют разные знаки. В то же время в соответствии с предположением 6 сила этого влияния тем выше, чем больше продолжительность Y_{ik} времени, в течение которого обмениваются мнениями i -й и k -й члены группы. Примем, что при этом время Y_{ik} распределяется поровну между всеми предметами обсуждения (т. е. остальными членами группы), так что j -й индивид обсуждается в течение времени $Y_{ik}/(N-2)$.

Простейший конкретный способ формализации данного рассуждения заключается в том, чтобы принять изменение величины X_{ij} зависящим от комбинации величин $X_{ik}X_{kj}Y_{ik}/(N-2)$. С учетом того, что i -й индивид обсуждает j -го с каждым другим членом группы, получаем

$$\frac{dX_{ij}}{dt} = Q_i(1-X_{ij})(X_{ij}+1) \sum_{k \neq i, j} \frac{X_{ik}X_{kj}Y_{ik}}{N-2}, \quad i, j = 1, \dots, n, \quad i \neq j. \quad (1)$$

Здесь Q_i — коэффициент, описывающий скорость восприятия i -го индивида, а множители $(1-X_{ij})$ и $(X_{ij}+1)$ описывают эффект насы-

щения при приближении к максимально возможному значению $X_{ij} = 1$ и минимально возможному значению $X_{ij} = -1$.

Теперь рассмотрим величину Y_{ij} , имеющую смысл времени, в течение которого индивиды i и j обсуждают других членов группы. Обозначим через T общую продолжительность социальной активности в группе (например, производственной деятельности группы, подразумевающей контакты между ее членами) в течение некоторого периода например месяца. Положим, что в ходе этой социальной активности каждый $\frac{T}{N-1}$ индивид проводит с каждым другим одинаковое время, равное $\frac{T}{N-1}$, причем часть его может быть посвящена обмену мнениями относительно других членов группы. В соответствии с предположением 7 доля времени, выделяемая на разговоры, зависит от отношений между этими индивидами: индивиды хотят больше разговаривать с людьми, которые им нравятся, чем с людьми, которые не нравятся. Конкретизируя данное предположение, примем

$$Y_{ij} = \frac{T}{N-1} \exp [\alpha(X_{ij} - 1)]. \tag{2}$$

Таким образом, если отношения между i -м и j -м членами группы максимально дружелюбные ($X_{ij} = 1$), то на обмен мнениями уходит максимально возможное время: $Y_{ij} = \frac{T}{N-1}$; при уменьшении дружелюбия продолжительность обмена мнениями убывает экспоненциально. Степень этого убывания отражается коэффициентом $\alpha > 0$.

Подставляя (2) в (1), получаем

$$\frac{dX_{ij}}{dt} = \frac{Q_i T}{(N-1)(N-2)} (1 - X_{ij}^2) \sum_{k \neq i, j} X_{ik} X_{kj} \exp [\alpha(X_{ik} - 1)], \tag{3}$$

$i, j = 1, \dots, n, \quad i \neq j.$

Обозначив

$$q_i = \frac{Q_i T}{(N-1)(N-2)},$$

получим систему уравнений данной динамической модели отношений в малой группе:

$$\frac{dX_{ij}}{dt} = q_i (1 - X_{ij}^2) \sum_{k \neq i, j} X_{ik} X_{kj} \exp [\alpha(X_{ik} - 1)], \tag{4}$$

$i, j = 1, \dots, n, \quad i \neq j.$

С учетом того, что аффективные отношения между индивидами вообще говоря несимметричны, $X_{ij} \neq X_{ji}$, то для группы из N индивидов система (4) содержит $N^2 - N$ уравнений.

3.7.2. Анализ симметричных стационарных состояний для группы из трех индивидов

В настоящем пункте исследуется построенная модель (4) для случая триады, т. е. группы, состоящей из трех индивидов. При этом одной из основных целей данного анализа является изучение соответствия

между динамической моделью (4) и рассмотренной выше статической моделью, в рамках которой понятие сбалансированной (т. е. устойчивой) малой группы вводилось на основе когнитивного баланса отдельных индивидов.

Для случая группы из трех индивидов система (4) содержит шесть уравнений:

$$\frac{dX_{12}}{dt} = q_1(1 - X_{12}^2)X_{13}X_{32} \exp[\alpha(X_{13} - 1)], \quad (5)$$

$$\frac{dX_{13}}{dt} = q_1(1 - X_{13}^2)X_{12}X_{23} \exp[\alpha(X_{12} - 1)], \quad (6)$$

$$\frac{dX_{21}}{dt} = q_2(1 - X_{21}^2)X_{23}X_{31} \exp[\alpha(X_{23} - 1)], \quad (7)$$

$$\frac{dX_{23}}{dt} = q_2(1 - X_{23}^2)X_{21}X_{13} \exp[\alpha(X_{21} - 1)], \quad (8)$$

$$\frac{dX_{31}}{dt} = q_3(1 - X_{31}^2)X_{32}X_{21} \exp[\alpha(X_{32} - 1)], \quad (9)$$

$$\frac{dX_{32}}{dt} = q_3(1 - X_{32}^2)X_{31}X_{12} \exp[\alpha(X_{31} - 1)]. \quad (10)$$

Она имеет значительное количество стационарных решений, при которых значения переменных X_{ij} , $i, j = 1, 2, 3$, $i \neq j$, равны нулю либо ± 1 . Поскольку значение отношения, равное $X_{ij} = X_{ji} = 1$, соответствует максимально возможному дружелюбию i -го и j -го индивидов, то оно аналогично положительному отношению в статической модели, показанному на рис. 1 слева. Аналогично, значение $X_{ij} = X_{ji} = -1$ соответствует отрицательному отношению в статической модели (рис. 1, справа).



Рис. 1. Положительное и отрицательное отношения в статической модели, соответствующие значениям $X_{ij} = X_{ji} = 1$ и $X_{ij} = X_{ji} = -1$ в динамической модели

Соответственно для каждой из возможных структур отношений в триаде (рис. 2) в статической модели соответствует определенное стационарное решение в динамической модели.

Для исследования стационарных решений системы (5)–(10) на устойчивость вычислим ее матрицу Якоби

$$G = \begin{pmatrix} \partial F_{12}/\partial X_{12} & \dots & \partial F_{12}/\partial X_{32} \\ \dots & \dots & \dots \\ \partial F_{32}/\partial X_{12} & \dots & \partial F_{32}/\partial X_{32} \end{pmatrix}. \quad (11)$$

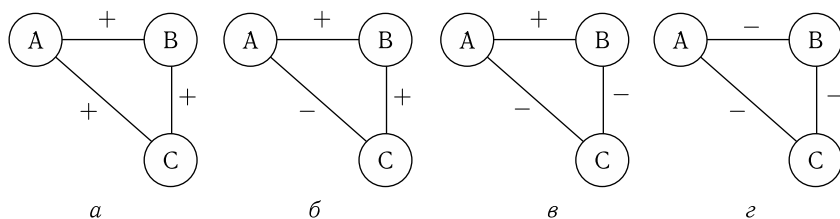


Рис. 2. Возможные варианты структуры отношений в триаде (в соответствии со статической моделью)

Здесь через F_{ij} , $i, j = 1, 2, 3$, $i \neq j$, обозначены правые части соответствующих уравнений из системы (5)–(10). В данном случае она является квадратной матрицей, имеющей по 6 строк и столбцов; ее полный вид приведен в комментариях к настоящему разделу.

Рассмотрим последовательно каждый из вариантов структуры отношений в триаде, представленных на рис. 2.

а. Отношения между всеми членами триады положительны (рис. 2, а):

$$X_{ij} = 1, \quad i, j = 1, 2, 3, \quad i \neq j. \quad (12)$$

Подставив данные значения X_{ij} в формулу (11) для матрицы Якоби (см. комментарий 2), получим для рассматриваемого стационарного решения (12):

$$G = \begin{pmatrix} -2q_1 & 0 & 0 & 0 & 0 & 0 \\ 0 & -2q_1 & 0 & 0 & 0 & 0 \\ 0 & 0 & -2q_2 & 0 & 0 & 0 \\ 0 & 0 & 0 & -2q_2 & 0 & 0 \\ 0 & 0 & 0 & 0 & -2q_3 & 0 \\ 0 & 0 & 0 & 0 & 0 & -2q_3 \end{pmatrix}. \quad (13)$$

Собственные значения λ_k , $k = 1, \dots, 6$, матрицы (13) равны ее диагональным членам. Очевидно, все они отрицательны. Тем самым стационарное решение (12) системы (5)–(10) является асимптотически устойчивым. Это соответствует сбалансированности данной триады в статической модели (как в рамках ХКХ-подхода к когнитивному балансу, так и в рамках Д-подхода).

б. Отношения в двух парах индивидов положительны, в третьей паре отношения отрицательны (рис. 2, б):

$$X_{12} = X_{21} = X_{13} = X_{31} = 1, \quad X_{23} = X_{32} = -1. \quad (14)$$

Матрица Якоби (11) в этом случае принимает вид

$$G = \begin{pmatrix} 2q_1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 2q_1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 2q_2 \exp(-2\alpha) & 0 & 0 & 0 \\ 0 & 0 & 0 & 2q_2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 2q_3 \exp(-2\alpha) & 0 \\ 0 & 0 & 0 & 0 & 0 & 2q_3 \end{pmatrix}.$$

Как и в предыдущем случае, собственные значения матрицы равны ее диагональным членам. Однако в данном случае они положительны. Следовательно, стационарное решение (14) неустойчиво, что соответствует несбалансированности (как ХКХ-, так и Д-несбалансированности) данной триады в статической модели.

в. Отношения в одной паре индивидов положительны, в двух других они отрицательны, т.е. «двое дружат против третьего» (рис. 2, в):

$$X_{12} = X_{21} = 1, \quad X_{13} = X_{31} = X_{23} = X_{32} = -1. \quad (15)$$

Для данной триады матрица Якоби (11) равна

$$A = \begin{pmatrix} -2q_1 e^{-2\alpha} & 0 & 0 & 0 & 0 & 0 \\ 0 & -2q_1 & 0 & 0 & 0 & 0 \\ 0 & 0 & -2q_2 e^{-2\alpha} & 0 & 0 & 0 \\ 0 & 0 & 0 & -2q_2 & 0 & 0 \\ 0 & 0 & 0 & 0 & -2q_3 e^{-2\alpha} & 0 \\ 0 & 0 & 0 & 0 & 0 & -2q_3 e^{-2\alpha} \end{pmatrix},$$

ее собственные значения отрицательны. Таким образом, стационарное решение (15) системы (5)–(10) является асимптотически устойчивым, что соответствует как ХКХ-, так и Д-сбалансированности данной триады в статической модели.

г. Отношения во всех парах индивидов отрицательны (рис. 2, г):

$$X_{12} = X_{21} = X_{13} = X_{31} = X_{23} = X_{32} = -1. \quad (16)$$

Данный тип триады — единственный, для которого в статической модели проявляется различие между ХКХ- и Д-подходами. Именно, триада, представленная на рис. 2, г, является Д-сбалансированной, но не ХКХ-сбалансированной. Чтобы определить, какому из этих подходов соответствует динамическая модель, рассмотрим матрицу Якоби (11),

которая для стационарного решения (16) принимает вид

$$G = \begin{pmatrix} 2q_1 e^{-2\alpha} & 0 & 0 & 0 & 0 & 0 \\ 0 & 2q_1 e^{-2\alpha} & 0 & 0 & 0 & 0 \\ 0 & 0 & 2q_2 e^{-2\alpha} & 0 & 0 & 0 \\ 0 & 0 & 0 & 2q_2 e^{-2\alpha} & 0 & 0 \\ 0 & 0 & 0 & 0 & 2q_3 e^{-2\alpha} & 0 \\ 0 & 0 & 0 & 0 & 0 & 2q_3 e^{-2\alpha} \end{pmatrix}.$$

Ее собственные значения положительны. Следовательно, стационарное решение (16) является неустойчивым. Данный вывод соответствует тому, что в статической модели триада, изображенная на рис. 2, *г*, является ХКХ-несбалансированной. Таким образом, динамическая модель, рассматриваемая в настоящем разделе, соответствует ХКХ-подходу в статической модели.

Помимо рассмотренных выше, у данной динамической модели существует значительное количество других стационарных решений. Эти решения не имеют аналогов в статической модели, рассмотренной в предыдущем разделе, и связаны со следующими двумя свойствами динамической модели.

- Отношения между индивидами могут быть несимметричными, т. е., вообще говоря, $X_{ij}(t) \neq X_{ji}(t)$. С психологической точки зрения это означает, что индивид А относится к индивиду В максимально положительно, а В к А — максимально отрицательно.
- Система (5)–(10) имеет стационарные решения, при которых некоторые (или даже все) значения величин X_{ij} равны нулю. Содержательно равенство $X_{ij} = 0$ означает, что отношения между i -м и j -м индивидами являются нейтральными.

Ввиду второго из этих свойств в динамической модели (5)–(10) имеется бесконечное количество стационарных решений, в т. ч. неизолированных. Например, $X_{12} = C$, $X_{13} = X_{21} = X_{23} = X_{31} = X_{32} = 0$ является стационарным решением при любом значении $C \in [0; 1]$. Для исследования устойчивости неизолированных решений нельзя пользоваться методом первого приближения (который был использован в предыдущем пункте). Поэтому мы не проводим здесь исчерпывающее исследование стационарных решений модели (5)–(10), ограничиваясь лишь двумя типами решений в дополнение к рассмотренным выше.

д. В триаде присутствует хотя бы одна пара индивидов с несимметричными стационарными отношениями, причем все $X_{ij} \neq 0$, $i, j = 1, 2, 3$, $i \neq j$.

В данном случае речь идет о стационарных решениях системы уравнений (5)–(10), при которых для одной или нескольких пар индивидов имеют место равенства $X_{ij} = -X_{ji} = 1$, при этом $|X_{ij}| = 1$ для всех $i, j = 1, 2, 3$, $i \neq j$. Предположим, для определенности, что

$$X_{12} = 1, X_{21} = -1, |X_{13}| = |X_{31}| = |X_{23}| = |X_{32}| = 1. \quad (17)$$

Подставив данные значения X_{ij} в выражение для матрицы Якоби (11), получим

$$G = \begin{pmatrix} -2q_1 X_{13} X_{32} \times & 0 & 0 & 0 & 0 & 0 \\ \times e^{\alpha(X_{13}-1)} & & & & & \\ 0 & -2q_1 X_{13} X_{23} & 0 & 0 & 0 & 0 \\ 0 & 0 & 2q_2 X_{23} X_{31} \times & 0 & 0 & 0 \\ & & \times e^{\alpha(X_{23}-1)} & & & \\ 0 & 0 & 0 & 2q_2 X_{23} \times & 0 & 0 \\ & & & \times X_{13} e^{-2\alpha} & & \\ 0 & 0 & 0 & 0 & 2q_3 X_{31} X_{32} \times & 0 \\ & & & & \times e^{\alpha(X_{32}-1)} & \\ 0 & 0 & 0 & 0 & 0 & -2q_3 X_{32} X_{31} \times \\ & & & & & \times e^{\alpha(X_{31}-1)} \end{pmatrix}.$$

Собственные значения данной матрицы равны ее диагональным элементам. При этом два из них, именно

$$\lambda_2 = -2q_1 X_{13} X_{23}, \quad \lambda_4 = 2q_2 X_{23} X_{13} e^{-2\alpha},$$

имеют разные знаки. Следовательно, одно из них положительно. Таким образом, стационарное решение вида (17) является неустойчивым. Поскольку данный вывод не зависит от знаков величин X_{13} , X_{23} , X_{31} , X_{32} , то он имеет место независимо от того, присутствуют ли в системе другие несимметричные отношения (помимо того, что $X_{12} = -X_{21}$).

e. Все стационарные отношения в группе являются нейтральными. Соответствующее стационарное решение системы (5)–(10) имеет вид

$$X_{ij} \neq 0, \quad i, j = 1, 2, 3, \quad i \neq j, \quad (18)$$

и является неустойчивым. Строгое доказательство неустойчивости требует применения довольно сложного аппарата функций Ляпунова, поэтому здесь мы ограничимся следующим простым, но нестрогим рассуждением.

Рассмотрим систему обыкновенных дифференциальных уравнений (5)–(10) с начальными условиями

$$X_{ij}(0) = \varepsilon > 0, \quad i, j = 1, 2, 3, \quad i \neq j, \quad (19)$$

где ε может быть выбрано сколь угодно малым. Тогда при $t = 0$ правые части всех уравнений (5)–(10) положительны, вследствие чего все переменные $X_{ij}(t)$ возрастают. Далее, с увеличением t правые части остаются положительными, вследствие чего все переменные $X_{ij}(t)$ монотонно возрастают, асимптотически приближаясь при $t \rightarrow \infty$ к стационарным значениям $X_{ij} = 1$, $i, j = 1, 2, 3$, $i \neq j$. Тем самым показано, что в сколь угодно малой окрестности стационарного решения (18) существует начальное условие вида (19), из которого с течением времени

решение системы дифференциальных уравнений (5)–(10) стабилизируется к другому стационарному u решению. Следовательно, стационарное решение (18) неустойчиво.

Таким образом, построенная и изученная в настоящем разделе модель (4) является динамическим обобщением статической модели, рассмотренной выше. В случае группы, состоящей из трех индивидов, она имеет асимптотически устойчивые стационарные решения, соответствующие ХКХ-сбалансированным триадам статической модели, а также неустойчивые стационарные решения, соответствующие ХКХ-несбалансированным триадам статической модели. Кроме того, динамическая модель имеет ряд других стационарных решений. В частности, показана неустойчивость решения, соответствующего триаде, в которой все отношения являются нейтральными.

Упражнения

1. Найти общее количество изолированных стационарных решений системы (5)–(10).

2. Для группы из четырех индивидов доказать, что стационарное решение, соответствующее сбалансированной группе, состоящей из двух антагонистических подгрупп, в каждой из которых состоят по два индивида (см. рис. 5 из предыдущего раздела), является устойчивым.

Комментарии и библиографические примечания

1. Изложенная в настоящем разделе модель представляет собой значительно измененный и упрощенный вариант модели, предложенной в работе [1].

2. Матрица Якоби (11) системы (5)–(10) является довольно громоздкой. Поэтому мы выпишем ее не в виде таблицы, а в виде отдельных столбцов. Ниже через G_k , $k = 1, \dots, 6$, обозначен k -й столбец ее матрицы:

$$G_1 = \begin{pmatrix} -2q_1 X_{12} X_{13} X_{32} \exp[\alpha(X_{13} - 1)] \\ q_1(1 - X_{13}^2) X_{23} \exp[\alpha(X_{12} - 1)](1 + \alpha X_{12}) \\ 0 \\ 0 \\ 0 \\ q_3(1 - X_{32}^2) X_{31} \exp[\alpha(X_{31} - 1)] \end{pmatrix},$$

$$G_2 = \begin{pmatrix} q_1(1 - X_{12}^2) X_{32} \exp[\alpha(X_{13} - 1)](1 + \alpha X_{13}) \\ -2q_1 X_{13} X_{12} X_{23} \exp[\alpha(X_{12} - 1)] \\ 0 \\ q_2(1 - X_{23}^2) X_{21} \exp[\alpha(X_{21} - 1)] \\ 0 \\ 0 \end{pmatrix},$$

$$\begin{aligned}
 G_3 &= \begin{pmatrix} 0 \\ 0 \\ -2q_2 X_{21} X_{23} X_{31} \exp[\alpha(X_{23} - 1)] \\ q_2(1 - X_{23}^2) X_{13} \exp[\alpha(X_{21} - 1)] (1 + \alpha X_{21}) \\ q_3(1 - X_{31}^2) X_{32} \exp[\alpha(X_{32} - 1)] \\ 0 \end{pmatrix}, \\
 G_4 &= \begin{pmatrix} 0 \\ q_1(1 - X_{13}^2) X_{12} \exp[\alpha(X_{12} - 1)] \\ q_2(1 - X_{21}^2) X_{31} \exp[\alpha(X_{23} - 1)] (1 + \alpha X_{23}) \\ -2q_2 X_{23} X_{21} X_{13} \exp[\alpha(X_{21} - 1)] \\ 0 \\ 0 \end{pmatrix}, \\
 G_5 &= \begin{pmatrix} 0 \\ 0 \\ q_2(1 - X_{21}^2) X_{23} \exp[\alpha(X_{23} - 1)] \\ 0 \\ -2q_3 X_{31} X_{32} X_{21} \exp[\alpha(X_{32} - 1)] \\ q_3(1 - X_{32}^2) X_{12} \exp[\alpha(X_{31} - 1)] (1 + \alpha X_{31}) \end{pmatrix}, \\
 G_6 &= \begin{pmatrix} q_1(1 - X_{12}^2) X_{13} \exp[\alpha(X_{13} - 1)] \\ 0 \\ 0 \\ 0 \\ q_3(1 - X_{31}^2) X_{21} \exp[\alpha(X_{32} - 1)] (1 + \alpha X_{32}) \\ -2q_3 X_{32} X_{31} X_{12} \exp[\alpha(X_{31} - 1)] \end{pmatrix}.
 \end{aligned}$$

Литература

1. Killworth P.D., Bernard H.R. A Model Of Human Group Dynamics // Social Science Research. 1976. 5. P. 173–224.

3.8. Динамическая модель близости позиций пользователей социальных сетей

Настоящий раздел посвящен динамической модели, описывающей взаимосвязь межличностных отношений и политических позиций взаимодействующих индивидов, образующих замкнутую группу. Таким образом, одно из основных отличий рассматриваемой ниже модели от изученной выше модели отношений между членами малой группы состоит в том, что здесь процесс взаимодействия индивидов заключается в обсуждении некоторых вопросов, представляющих общий интерес, а не других членов группы. Для определенности, далее будем говорить о политических вопросах и, соответственно, о политических позициях индивидов.

Другое важное отличие связано с тем, что рассматриваемые здесь взаимодействия носят черты массовой (а не попарной) коммуникации: мнение каждого индивида в равной степени доступно всем остальным. Напомним, что при попарной коммуникации (см построение модели в предыдущем разделе) время социальной активности каждого индивида делится между потенциальными собеседниками, причем лишь часть этого времени посвящена обсуждениям, которые влияют на изменение отношений между членами группы. В рамках рассмотренной в выше модели два индивида влияют друг на друга тем сильнее, чем продолжительнее время их общения, которое в свою очередь зависит от симпатии между ними. В противоположность этому, при массовой коммуникации индивид публикует свое мнение, делая его доступным одновременно (и в равной мере) для всех членов группы. Такая ситуация имеет место, например в сообществах пользователей интернет-форумов или блог-платформ. В то же время аффективные, эмоциональные отношения (т. е. отношения симпатии или антипатии) между индивидами в данной модели являются попарными.

С учетом указанных отличий данная модель родственна динамической модели, изложенной выше, поскольку обе они развивают концепцию структурного баланса.

Далее, вводятся основные понятия, формулируются предпосылки, и на их основе вводятся уравнения модели. Второй и третий пункты посвящены ее аналитическому исследованию в базовых случаях а четвертый — численным экспериментам с моделью.

3.8.1. Построение модели

В преамбуле к настоящему разделу были описаны некоторые черты моделируемого процесса. Рассмотрим его более подробно. Предметом обсуждения в группе являются некоторые (например, политические) темы. Все индивиды имеют в каждый момент времени $t \geq 0$ определенные позиции по каждой теме. Кроме того, между индивидами имеются попарные отношения, именно: у каждого члена группы есть представления о том, насколько близким к нему является каждый из других индивидов. С течением времени отношения изменяются в соответствии с принципом *«если индивид высказывает близкие к моим политические позиции, то я считаю его более близким ко мне, чем ранее, и наоборот»*. С другой стороны, отношения между членами группы влияют на изменение их политических позиций, что можно условно описать словами *«позиция, высказанная моим другом, притягивает, а высказанная моим врагом, отталкивает мою позицию»*. Тем самым отношения между индивидами и их политические позиции влияют друг на друга. Для описания данной динамики и служит настоящая модель.

Перейдем к формулировке конкретных предположений, на основе которых будут получены уравнения модели.

Предположение 1. Рассматривается замкнутая малая группа, состоящая из $K \geq 2$ индивидов, обсуждающих $N \geq 1$ тем (которые мы также будем называть вопросами). Замкнутость означает, что в течение рассматриваемого интервала времени индивиды не покидают и не пополняют группу.

Это предположение является стандартным для традиционных моделей групповой динамики. Напомним, что группа индивидов считается малой, если каждый входящий в нее индивид имеет определенное отношение (в данном случае представление о близости, выражаемое в виде субъективной дистанции — см. ниже) о каждом другом члене группы.

Предположение 2. Позиция i -го индивида по n -му вопросу в момент времени t может быть охарактеризована скалярной переменной $z_{in}(t) \in (-\infty, +\infty)$, где $i = 1, \dots, K$, $n = 1, \dots, N$, $t \geq 0$.

Представление политической позиции в виде точки на одномерном континууме соответствует наиболее распространенному в политологии подходу.

Значения этой функции могут трактоваться, например следующим образом: если $z_{in}(t) < 0$, то i -й индивид относится к обсуждаемому политическому проекту (или политическому деятелю) отрицательно, и если $z_{in}(t) > 0$, то положительно.

Будем полагать, что индивиды *не обманывают друг друга*, т.е. каждый из них высказывает по каждому вопросу именно ту позицию, которую он действительно занимает. Высказываемые позиции индивида известны всем остальным.

Следующим предположением вводится величина, характеризующая аффективные (эмоциональные) отношения между индивидами.

Предположение 3. Аффективные отношения между двумя индивидами могут быть описаны с помощью *субъективной дистанции*. Именно, у i -го индивида есть представления о том, насколько близок к нему в данный момент времени каждый из прочих индивидов группы; эти представления описываются неотрицательной переменной величиной $x_{ij}(t)$, которая называется субъективной дистанцией от i -го индивида до j -го (с точки зрения индивида, указанного первым). При этом, вообще говоря, два индивида могут по-разному воспринимать отношения между ними, т.е. $x_{ij}(t) \neq x_{ji}(t)$. Здесь и далее $i, j = 1, \dots, K$, $i \neq j$.

Введенная таким образом субъективная дистанция аналогична величине $X_{ij}(t)$ из динамической модели, рассмотренной выше, но отличается от нее калибровкой. Именно, субъективная дистанция равна нулю ($x_{ij} = 0$) при максимально положительном аффективном отношении ($X_{ij} = 1$). Максимально отрицательному отношению $X_{ij} = -1$ соответствует бесконечная субъективная дистанция, $x_{ij} = +\infty$.

Перейдем к предположениям модели, описывающим динамику введенных выше переменных $x_{ij}(t)$, $z_{in}(t)$, $i, j = 1, \dots, K$, $i \neq j$, $n = 1, \dots, N$.

Предположение 4. Отношение i -го члена группы к j -му с течением времени изменяется из-за высказываний j -го индивида: именно, субъективная дистанция уменьшается, если высказанные j -м индивидом позиции достаточно близки к позициям i -го индивида, и увеличивается, если они далеки от нее. В данном контексте считается, что позиции по n -ному вопросу в момент времени t достаточно близки, если $|z_{in}(t) - z_{jn}(t)| < a$, где a — некоторый заданный пороговый параметр (для простоты будем полагать, что значение этого параметра является одинаковым для всех индивидов и всех вопросов).

Различия в позициях по обсуждаемым вопросам учитываются с весовыми коэффициентами α_{in} , описывающими субъективную важность каждого из вопросов с точки зрения i -го индивида, нормированными так, что $\sum_{n=1}^N \alpha_{in} = 1$. Полученная в результате этого величина

$$S_{ij} = \sum_{n=1}^N \alpha_{in} [(z_{in} - z_{jn})^2 - a^2] dt, \quad i, j = 1, \dots, K, i \neq j,$$

может рассматриваться как действующий в течение интервала времени dt стимул для такой реакции, как изменение i -м индивидом отношения к j -му. Таким образом,

$$\frac{dx_{ij}}{dt} = f(x_{ij}, S_{ij}), \quad i, j = 1, \dots, K, i \neq j. \quad (1)$$

Здесь функция $f(x_{ij}, S_{ij})$ определена при $x_{ij} \geq 0$, $-\infty < S_{ij} < +\infty$ и обладает следующими свойствами.

1. Если $S_{ij} < 0$, то $f(x_{ij}, S_{ij}) < 0$ при всех $x_{ij} > 0$. Например, если позиции i -го и j -го членов группы достаточно близки по всем вопросам (т. е. $|z_{in} - z_{jn}| < a$ для всех $n = 1, \dots, N$), то $S_{ij} < 0$ и стимул действует в направлении уменьшения субъективной дистанции x_{ij} .
2. Если $S_{ij} = 0$, то $f(x_{ij}, S_{ij}) = 0$ при всех $x_{ij} > 0$.
3. Если $S_{ij} > 0$, то $f(x_{ij}, S_{ij}) > 0$ при всех $x_{ij} > 0$.
4. Поскольку субъективная дистанция x_{ij} по своему содержательному смыслу является неотрицательной величиной, то вид функции $f(x_{ij}, S_{ij})$ должен обеспечивать, чтобы при положительном начальном условии $x_{ij}(0) > 0$ функция оставалась неотрицательной при всех $t > 0$.

Для конкретизации предположения 4 может быть предложена функция $f(x_{ij}, S_{ij}) = \gamma_i x_{ij} S_{ij}$, удовлетворяющая всем перечисленным условиям. Тогда уравнение (1) принимает вид

$$\frac{dx_{ij}}{dt} = \gamma_i x_{ij} \sum_{n=1}^N \alpha_{in} [(z_{in} - z_{jn})^2 - a^2], \quad i, j = 1, \dots, K, i \neq j. \quad (2)$$

Данное уравнение показывает содержательный смысл введенного выше параметра $\gamma_i > 0$: он характеризует i -го индивида с точки зрения того, насколько быстро или медленно он меняет свое отношение к другим членам группы. Другими словами, параметр $\gamma_i > 0$ описывает скорость реакции i -го члена группы в части изменения межличностных отношений.

Заметим, что в соответствии с уравнением (2) i -й и j -й индивиды могут иметь (по крайней мере, в течение некоторого интервала времени) противоположные взгляды на отношения между ними. Действительно, пусть $N = 2$, т. е. в группе обсуждаются всего две темы. По первой из них позиции i -го и j -го индивидов близки: $|z_{i1} - z_{j1}| < a$, а по второй далеки, $|z_{i2} - z_{j2}| > a$. И пусть для i -го члена группы первая тема является более важной, чем вторая, а для j -го члена, наоборот, вторая тема является более важной, чем первая: $\alpha_{i1} > \alpha_{i2}$, $\alpha_{j1} < \alpha_{j2}$. В этом случае, i -й индивид считает j -го своим единомышленником, однако j -й считает i -го своим оппонентом. Из уравнения (2) в таком случае получаем, что $dx_{ij}/dt < 0$, но $dx_{ji}/dt > 0$.

Предположение 5. С течением времени позиции индивидов по каждому из вопросов изменяются под влиянием высказываний членов группы. При этом высказанная достаточно близким (с точки зрения субъективной дистанции слушателя) членом группы позиция способствует приближению позиции слушателя к ней, а высказанная далеким членом группы позиция — отдалению от нее. В данном контексте j -й член группы считается достаточно близким к i -му (с точки зрения i -го), если $x_{ij} < R_i$, где параметр $R_i > 0$ характеризует i -го индивида с точки зрения того, насколько близких с точки зрения субъективной дистанции пользователей он воспринимает как «собеседников, к мнению которых стоит прислушаться».

Таким образом, влияние позиции j -го члена группы по n -ной теме вопросу на позицию i -го члена группы по этой теме описывается функцией $g_{ij}(z_{in}, z_{jn}, x_{ij})$, обладающей следующими свойствами.

5. Если $z_{jn} > z_{in}$ и $x_{ij} < R$, то $g_{ij}(z_{in}, z_{jn}, x_{ij}) > 0$.
6. Если $z_{jn} > z_{in}$ и $x_{ij} > R$, то $g_{ij}(z_{in}, z_{jn}, x_{ij}) < 0$.
7. Если $z_{jn} < z_{in}$ и $x_{ij} < R$, то $g_{ij}(z_{in}, z_{jn}, x_{ij}) < 0$.
8. Если $z_{jn} < z_{in}$ и $x_{ij} > R$, то $g_{ij}(z_{in}, z_{jn}, x_{ij}) > 0$.

Поскольку на изменение позиции i -го индивида по вопросу n влияют позиции всех остальных членов группы, то они учитываются с весовыми коэффициентами $\beta_{ijn} \geq 0$, $j \neq i$, нормированными так,

что $\sum_{\substack{j=1 \\ j \neq i}}^K \beta_{ijn} = 1$. Для удобства положим по определению $\beta_{iin} = 0$,

что позволяет не делать исключение $j \neq i$ под знаком суммирования:

$$\sum_{j=1}^K \beta_{ijn} = 1.$$

Таким образом (в соответствии с предположением 5), имеем уравнение

$$\frac{dz_{in}}{dt} = \sigma_i \sum_{j=1}^K \beta_{ijn} g_{ijn}(z_{in}, z_{jn}, x_{ij}), \quad i = 1, \dots, K, \quad n = 1, \dots, N. \quad (3)$$

Здесь параметр $\sigma_i > 0$ характеризует i -го индивида с точки зрения того, насколько быстро или медленно он меняет свои позиции по обсуждаемому тема. Другими словами, параметр σ_i описывает скорость реакции i -го члена группы в части изменения его позиций (например, политических).

Простейшая функция, обладающая перечисленными выше свойствами, имеет вид

$$g_{ij}(z_{in}, z_{jn}, x_{ij}) = (z_{jn} - z_{in})(R_i - x_{ij}). \quad (4)$$

Подставляя (4) в (3), получаем в окончательном виде уравнение, описывающее динамику позиций индивидов по обсуждаемым темам:

$$\frac{dz_{in}}{dt} = \sigma_i \sum_{j=1}^K \beta_{ijn} (z_{jn} - z_{in})(R_i - x_{ij}), \quad i = 1, \dots, K, \quad n = 1, \dots, N. \quad (5)$$

Система уравнений вида (2), (5) образует замкнутую модель.

В соответствии с предположением 5 изменения мнений членов группы пренебрежимо мало зависят от других причин: мнений, внешних по отношению к данной группе, текущих политических событий и т. д. С практической точки зрения это ограничивает область применимости модели «фундаментальными» вопросами, изменение позиции по которым происходит не в результате появления новой информации, а вследствие осмысления и самопозиционирования в дискуссии, например в ходе предвыборной политической борьбы.

Итак, система уравнений модели динамики политических позиций пользователей имеет вид (2), (5). Она дополняется начальными условиями вида

$$x_{ij}(0) = x_{ij}^0, \quad z_{in}(0) = z_{in}^0. \quad (6)$$

Поскольку $i, j = 1, \dots, K, j \neq i, n = 1, \dots, N$, система содержит $N^2 - N$ количество уравнений вида (2) и KN уравнений вида (5). Например, в случае трех индивидов и трех тем система имеет 6 уравнений, описывающих динамику переменных типа $x_{ij}(t)$, и 9 уравнений динамики переменных типа $z_{in}(t)$.

Перейдем к исследованию построенной модели (2), (5), (6).

3.8.2. Стационарное состояние согласия в случае двух индивидов

Рассмотрим случай, когда группа состоит всего из двух индивидов.

В этом случае уравнения вида (5) имеют в правой части лишь по одному слагаемому:

$$\frac{dz_{1n}}{dt} = \sigma_1(z_{2n} - z_{1n})(R_1 - x_{12}), \quad n = 1, \dots, N, \quad (7)$$

$$\frac{dz_{2n}}{dt} = \sigma_2(z_{1n} - z_{2n})(R_2 - x_{21}), \quad n = 1, \dots, N. \quad (8)$$

Модель содержит в данном случае $2N + 2$ обыкновенных дифференциальных уравнения для такого же количества переменных. Чтобы упростить ее, вычтем уравнение (8) из (7), получим

$$\frac{d(z_{1n} - z_{2n})}{dt} = -(z_{1n} - z_{2n})[\sigma_1(R_1 - x_{12}) + \sigma_2(R_2 - x_{21})].$$

Введем переменную, имеющую смысл разности позиций двух индивидов по вопросу n : $y_n = z_{1n} - z_{2n}$. Таким образом, получим систему $N + 2$ уравнений

$$\frac{dx_{12}}{dt} = \gamma_1 x_{12} \sum_{n=1}^N \alpha_{1n} [y_n^2 - a^2], \quad (9)$$

$$\frac{dx_{21}}{dt} = \gamma_2 x_{21} \sum_{n=1}^N \alpha_{2n} [y_n^2 - a^2], \quad (10)$$

$$\frac{dy_n}{dt} = -y_n [\sigma_1(R_1 - x_{12}) + \sigma_2(R_2 - x_{21})], \quad n = 1, \dots, N. \quad (11)$$

Уравнения для стационарных состояний системы (9)–(11) имеют вид

$$x_{12} \sum_{n=1}^N \alpha_{1n} [y_n^2 - a^2] = 0, \quad (12)$$

$$x_{21} \sum_{n=1}^N \alpha_{2n} [y_n^2 - a^2] = 0, \quad (13)$$

$$y_n [\sigma_1(R_1 - x_{12}) + \sigma_2(R_2 - x_{21})] = 0, \quad n = 1, \dots, N. \quad (14)$$

Очевидно, система (12)–(14) имеет решение $x_{12} = 0$, $x_{21} = 0$, $y_n = 0$, $n = 1, \dots, N$. Назовем это стационарное решение *состоянием согласия*, так как оно соответствует ситуации, при которой каждый из индивидов считает другого единомышленником и они занимают одну и ту же позицию по всем вопросам: $z_{1n} = z_{2n}$ (т. е. разность позиций $y_n = 0$) для всех $n = 1, \dots, N$.

Чтобы исследовать состояние согласия на устойчивость, линеаризуем систему (9)–(11) в его окрестности. Получим

$$\frac{dx_{12}}{dt} = -a^2\gamma_1 x_{12} \sum_{n=1}^N \alpha_{1n}, \quad (15)$$

$$\frac{dx_{21}}{dt} = -a^2\gamma_2 x_{21} \sum_{n=1}^N \alpha_{2n}, \quad (16)$$

$$\frac{dy_n}{dt} = -y_n [\sigma_1 R_1 + \sigma_2 R_2], \quad n = 1, \dots, N. \quad (17)$$

Матрица линеаризованной системы (15)–(17) имеет вид

$$A = \begin{pmatrix} -a^2\gamma_1 \sum_{n=1}^N \alpha_{1n} & 0 & 0 & \dots & 0 \\ 0 & -a^2\gamma_2 \sum_{n=1}^N \alpha_{2n} & 0 & \dots & 0 \\ 0 & 0 & -\sigma_1 R_1 - \sigma_2 R_2 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & -\sigma_1 R_1 - \sigma_2 R_2 \end{pmatrix}.$$

Ее собственные значения равны диагональным элементам

$$\lambda_n = -[\sigma_1 R_1 + \sigma_2 R_2] < 0, \quad n = 1, \dots, N,$$

$$\lambda_{N+1} = -a^2\gamma_1 \sum_{n=1}^N \alpha_{1n} < 0, \quad \lambda_{N+2} = -a^2\gamma_2 \sum_{n=1}^N \alpha_{2n} < 0.$$

Поскольку все собственные значения отрицательны, то исследуемое стационарное решение (состояние согласия) является асимптотически устойчивым.

3.8.3. Стационарное состояние с несимметричными отношениями

При изучении модели динамики отношений в группе, рассмотренной выше, несимметричными назывались отношения в паре индивидов, при которых один член группы относился к другому положительно, а тот к нему — отрицательно. В данной же модели аффективные отношения характеризуются субъективной дистанцией x_{ij} , причем i -й индивид воспринимает j -го как «собеседника, к мнению которых стоит прислушаться», если $x_{ij} < R_i$. Поэтому в рамках этой модели *будем называть отношения i -го и j -го индивидов несимметричными, если $x_{ij} < R_i$, $x_{ji} > R_i$ или наоборот: $x_{ij} > R_i$, $x_{ji} < R_i$* . Настоящий пункт посвящен анализу этих решений, в первую очередь — вопросу о том, являются ли они устойчивыми. При этом ограничимся случаем двух индивидов, обсуждающих одну тему, причем эти индивиды имеют

одинаковую скорость реакции в части изменения межличностных отношений: $\gamma_1 = \gamma_2 = \gamma$.

В этом случае система уравнений модели (9)–(11) принимает вид

$$\frac{dx_{12}}{dt} = \gamma x_{12}(y^2 - a^2), \quad (18)$$

$$\frac{dx_{21}}{dt} = \gamma x_{21}(y^2 - a^2), \quad (19)$$

$$\frac{dy}{dt} = -y[\sigma_1(R_1 - x_{12}) + \sigma_2(R_2 - x_{21})]. \quad (20)$$

Соответствующая стационарная система

$$\gamma x_{12}(y^2 - a^2) = 0, \quad (21)$$

$$\gamma_2 x_{21}(y^2 - a^2) = 0, \quad (22)$$

$$-y[\sigma_1(R_1 - x_{12}) + \sigma_2(R_2 - x_{21})] = 0 \quad (23)$$

имеет множество неизолированных решений, даваемых формулами

$$y = a, \quad (24)$$

$$\sigma_1 x_{12} + \sigma_2 x_{21} = \sigma_1 R_1 + \sigma_2 R_2. \quad (25)$$

Важное свойство равенства (26) заключается в том, что если $x_{12} > R_1$, то $x_{21} < R_2$; и наоборот, если $x_{12} < R_1$, то $x_{21} > R_2$. Таким образом, стационарное решение (24), (26) соответствует несимметричным отношениям между индивидами.

Поскольку эти решения не являются изолированными (они составляют прямую в трехмерном пространстве переменных x_{12}, x_{21}, y), то для анализа их устойчивости непригодны базовые методы, такие как метод исследования устойчивости по первому приближению. В связи с этим сделаем замену переменных, которая позволит уменьшить размерность системы (это возможно благодаря сделанному выше упрощающему предположению $\gamma_1 = \gamma_2 = \gamma$). Замена имеет вид

$$p = \sigma_1 x_{12} + \sigma_2 x_{21}. \quad (26)$$

Систему дифференциальных уравнений для функций $p(t)$, $y(t)$ получим, умножив уравнения (18), (19) на σ_1 и σ_2 , соответственно, и сложив их. Получим

$$\frac{dp}{dt} = \gamma p(y^2 - a^2), \quad (27)$$

$$\frac{dy}{dt} = -y[\sigma_1 R_1 + \sigma_2 R_2 - p]. \quad (28)$$

Эта система имеет следующие стационарные решения:

$$y = 0, \quad p = 0, \quad (29)$$

$$y = a, \quad p = \sigma_1 R_1 + \sigma_2 R_2, \quad (30)$$

$$y = -a, \quad p = \sigma_1 R_1 + \sigma_2 R_2. \quad (31)$$

Первое из них соответствует состоянию согласия, а решения (30), (31) — несимметричным отношениям. Поскольку эти решения суть изолированные, то для исследования их на устойчивость применим метод первого приближения.

В соответствии с ним составим матрицу Якоби системы (27), (28). Получим

$$G = \begin{pmatrix} \gamma(y^2 - a^2) & 2y\gamma p \\ y & -[\sigma_1 R_1 + \sigma_2 R_2 - p] \end{pmatrix}. \quad (32)$$

Подставив значения y и p из равенств (30), получим

$$G = \begin{pmatrix} 0 & 2a\gamma(\sigma_1 R_1 + \sigma_2 R_2) \\ a & 0 \end{pmatrix}. \quad (33)$$

Для собственных значений матрицы вида (33) имеем уравнение

$$\lambda^2 - 2a^2\gamma(\sigma_1 R_1 + \sigma_2 R_2) = 0,$$

откуда

$$\lambda_1 = a\sqrt{2\gamma(\sigma_1 R_1 + \sigma_2 R_2)}, \quad \lambda_2 = -a\sqrt{2\gamma(\sigma_1 R_1 + \sigma_2 R_2)}.$$

Поскольку одно из собственных значений положительно, то стационарное решение (30) неустойчиво. Аналогичным образом нетрудно показать, что стационарное решение (31) также неустойчиво.

Таким образом, в рамках данной модели несимметричные отношения индивидов неустойчивы, как и в модели динамики отношений индивидов в малой группе.

3.8.4. Вычислительные эксперименты с моделью

Настоящий раздел посвящен численному исследованию нестационарных решений построенной математической модели, главным образом, для случая двух тем ($N = 2$) и двух индивидов ($K = 2$). Количество уравнений модели (2), (5), (6) в этом случае равно шести. Приведено также краткое описание результатов для случая одной темы и трех индивидов ($N = 1, K = 3$).

Система уравнений (2), (5) в случае $N = 2, K = 2$ исследовалась при различных значениях параметров и различных начальных условиях (6). Было показано, в частности, что все стационарные решения, кроме состояния согласия, являются неустойчивыми.

Приведем пример численного эксперимента, при котором происходит стабилизация решения к состоянию согласия. Конкретные начальные условия и значения параметров имеют вид

$$\text{субъективные дистанции: } x_{12}(0) = 1, \quad x_{21}(0) = 2, \quad (34)$$

$$\text{позиции индивидов по первой теме: } z_{11} = 1, \quad z_{21} = 5, \quad (35)$$

$$\text{позиции индивидов по второй теме: } z_{12} = 5, \quad z_{22} = 1, \quad (36)$$

$$\text{параметры уравнений: } a = 5, \quad R = 5, \quad \gamma_1 = \gamma_2 = 1, \quad \sigma_1 = \sigma_2 = 0,5. \quad (37)$$

При этих данных модель соответствует следующей социальной ситуации. Соотношения

$$x_{12}(0) = 1 < R_1 = 5,$$

$$x_{21}(0) = 2 < R_2 = 5$$

означают, что в начальный момент времени $t = 0$ каждый из членов группы воспринимает другого как достаточно близкого. Разность начальных позиций индивидов по первой теме равна $y_1(0) = z_{11}(0) - z_{21}(0) = -4$, по второй теме — соответственно $y_2(0) = z_{12}(0) - z_{22}(0) = 4$. Поскольку

$$|y_1(0)| = |y_2(0)| = 4 < a = 5,$$

то начальные позиции индивидов достаточно близки, чтобы субъективные дистанции $x_{12}(t)$ и $x_{21}(t)$ уменьшались. Таким образом, сочетание параметров и начальных условий в данном случае таково, что на основании общих соображений следует ожидать дальнейшего улучшения аффективных отношений и сближения позиций по обсуждаемым темам (хорошие отношения ведут к сближению позиций, а близкие позиции — к улучшению отношений). Результаты численного эксперимента, представленные на рис. 1, 2, соответствуют этим представлениям. Именно, графики на рис. 1 описывают уменьшение величин $x_{12}(t)$ и $x_{21}(t)$ вплоть до нулевых значений. В то же время рис. 2

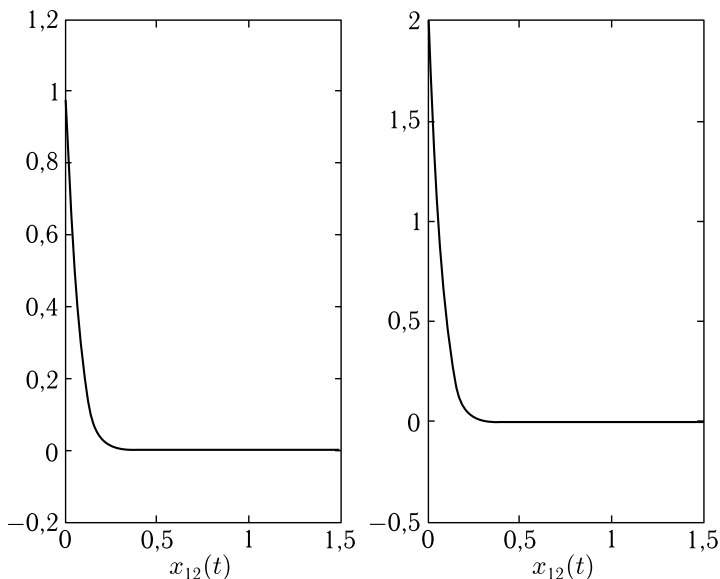


Рис. 1. Динамика субъективных дистанций (численный расчет для случая двух индивидов и двух тем)

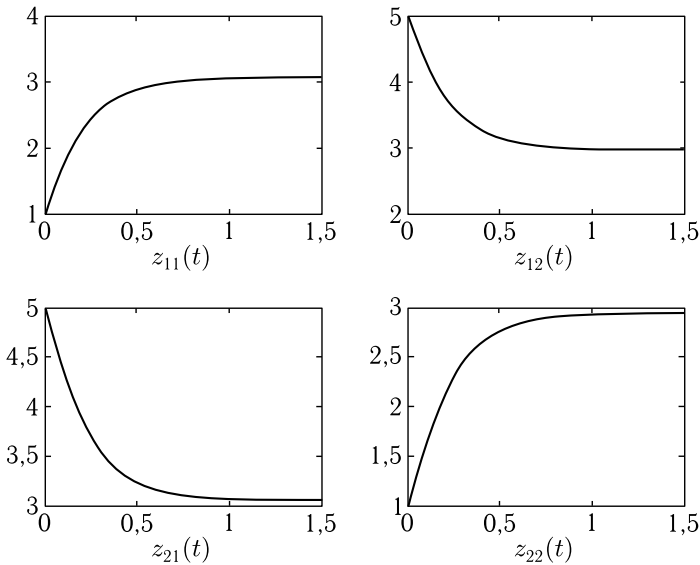


Рис. 2. Динамика позиций участников (численный расчет для случая двух индивидов и двух тем)

показывает, что позиции индивидов по каждому из вопросов сближаются при $t \rightarrow \infty$, причем возникающий консенсус равен среднему арифметическому от начальных позиций. Другими словами,

$$\lim_{t \rightarrow \infty} z_{11}(t) = \lim_{t \rightarrow \infty} z_{21}(t) = \frac{z_{11}(0) + z_{21}(0)}{2} = \frac{1 + 5}{2} = 3$$

(и аналогично для позиций по второй теме). Такое «срединное» положение консенсуса связано с тем, что в данном примере индивиды обладают одинаковой скоростью реакции i -го члена группы в части изменения его позиций: $\sigma_1 = \sigma_2$.

Как показали другие эксперименты из проведенной серии, если первый индивид меняет позиции по обсуждаемым вопросам медленнее второго, т.е. $\sigma_1 < \sigma_2$, то консенсусная позиция оказывается ближе к начальной позиции первого индивида, чем второго.

Перейдем к краткому описанию результатов некоторых вычислительных экспериментов для случая группы из трех индивидов, обсуждающих одну тему.

На рис. 3 представлена динамика субъективных дистанций между индивидами при формировании состояния согласия. Графики (визуально идентичные) показывают, что с течением времени все субъективные дистанции убывают так, что

$$\lim_{t \rightarrow \infty} x_{ij} = 0, \quad i, j = 1, 2, 3, \quad i \neq j.$$

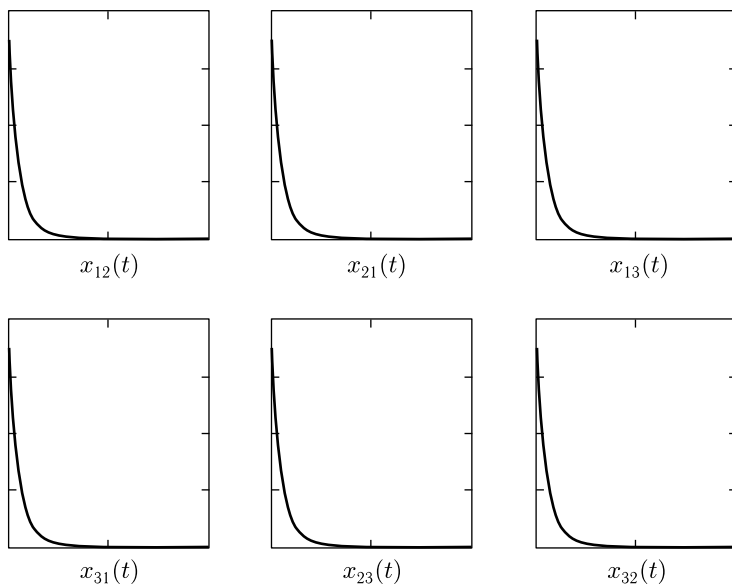


Рис. 3. Динамика субъективных дистанций при формировании состояния согласия (численный расчет для случая трех индивидов и одной темы)

Рис. 4 представляет процесс формирования состояния, при котором «первый и второй индивиды дружат против третьего». Обе субъективные дистанции (одна из которых характеризует отношение первого индивида ко второму, а другая — отношение второго к первому) между первым и вторым членами группы убывают:

$$\lim_{t \rightarrow \infty} x_{12} = \lim_{t \rightarrow \infty} x_{21} = 0.$$

В то же время третий индивид с течением времени отдаляется от них обоих: неограниченно увеличиваются как величины $x_{31}(t)$, $x_{32}(t)$, характеризующие отношение третьего члена группы к первому и второму, так и функции $x_{13}(t)$, $x_{23}(t)$, описывающие отношение первого и второго к третьему.

Таким образом, в данной модели динамики позиций пользователей социальных сетей устойчивыми являются те же групповые структуры, что и в модели динамики отношений в группе, изложенной в начале настоящей главы. В то же время здесь моделируются несколько иные ситуации; прежде всего, разница заключается в том, что рассматриваются не только аффективные отношения между индивидами, но и их позиции по обсуждаемым вопросам (подробнее о различиях между моделями см. в преамбуле к настоящему разделу). Будучи основанными, в определенном смысле, на общем базисе, эти модели дополняют друг друга.

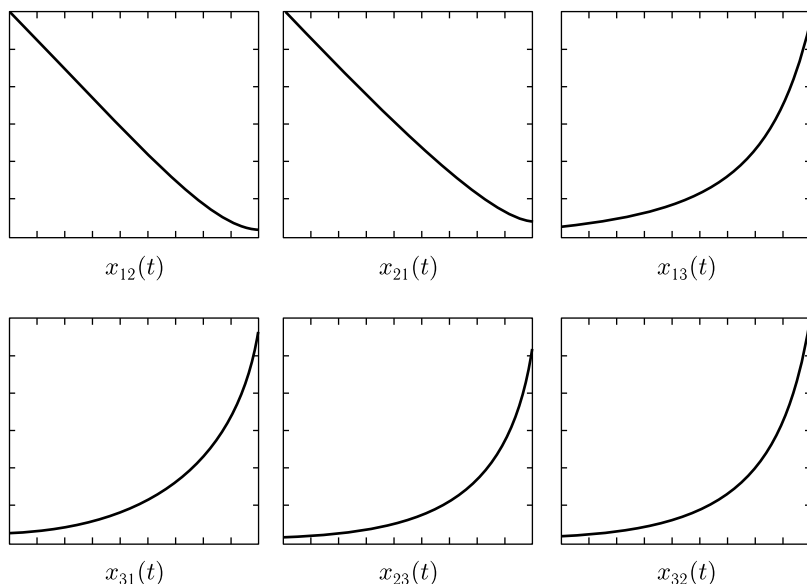


Рис. 4. Динамика субъективных дистанций при формировании состояния «двое друзей против третьего» (численный расчет для случая трех индивидов и одной темы)

Упражнения

1. Рассмотреть систему (18)–(20) в случае $\gamma \ll 1$. Построить асимптотику решения по малому параметру. Предложить содержательную трактовку результатов.

2. Рассмотреть систему (18)–(20) в случае $\sigma_1 = \sigma_2 \ll 1$. Построить асимптотику решения по малому параметру. Предложить содержательную трактовку результатов.

Комментарии и библиографические примечания

1. Изложенная в настоящем разделе модель предложена в работе [1], см. также [2, 3].

2. Представление политической позиции в виде точки на одномерной шкале широко распространено в политологии. Данный подход был предложен в [3], обзор современного состояния теории см. в [5].

3. Эмпирические методы измерения отношений между индивидами (в терминах данной модели — субъективных дистанций) хорошо известны в психологии. В то же время разработка методов измерения близости политических позиций является актуальной задачей современной политологии. Обзор существующих методов лежит вне рамок настоящей работы; отметим лишь, что все они основаны

на анализе текстов, выражающих эти позиции. Наиболее известная методика такого рода представлена в проекте «Манифесто» [6], ее основным элементом является специального вида вопросник. Альтернативный подход предложен в работах [7–10], он основан на процедуре латентно-семантического анализа [11], позволяющего с помощью алгоритмической (реализованной программным образом) процедуры измерить степень близости между высказанными в текстах позициями на основе анализа контекста, в котором употребляются фрагменты текста и отдельные слова.

Литература

1. Корнилина Е.Д., Петров А.П. Динамическая модель близости позиций пользователей социальных сетей // Математическое моделирование. 2012. Т. 24. № 10. С. 89–97.
2. Gorbatiukov E., Kornilina E., Mikhailov A., Petrov A. Mathematical Model of Opinion Dynamics in Social Groups // Mediterranean Journal of Social Sciences. V. 4. № 10. October 2013. p. 380–387. Doi:10.5901/mjss.2013.v4n10p380.
3. Корнилина Е.Д., Петров А.П. Модель динамики близости политических позиций // Теория активных систем — 2011 (ТАС-2011). Труды международной научно-практической конференции. ИПУ РАН, Москва, 14–16 ноября 2011 г. Том 2. С. 270–273.
4. Downes A. An economic Theory of Democracy. N.Y., Harper&Row, 1957.
5. Ахременко А.С. Структуры электорального пространства. М.: Социально-политическая мысль, 2007. 320 с.
6. Volkens A. Strengths and weaknesses of approaches to measuring policy positions of parties // Electoral Studies. 2007. Т. 26. № 1. С. 108–120.
7. Корнилина Е.Д. Применение латентно-семантического анализа к определению позиции политических блогов // Социология. 2011 г. № 1 С. 162–164.
8. Петров А.П., Корнилина Е.Д. Исследование близости политических позиций методом латентно-семантического анализа // XII Международная научная конференция по проблемам развития экономики и общества. Книга 2. М.: Издательский дом высшей школы экономики. 2012. С. 334–342.
9. Kornilina E., Petrov A. Research of political blogs by means of LSA // Intellectualization of information processing-2010 conf. proceedings — 2010 P. 508–511.
10. Корнилина Е.Д., Петров А.П. Латентно-семантический анализ предвыборных партийных программ на выборах в Государственную Думу 2007 и 2011 годов // М.: Вестник МГУ. Сер. 12: Политические науки, Издательство МГУ. 2013. № 2. С. 90–98.
11. Landauer T., Foltz P.W., Laham D. Introduction to Latent Semantic Analysis. Discourse Processes 1998. V. 25. P. 259–284.

Глава 4

СОЦИОЛОГИЯ

4.1. Латентно-структурный анализ

При проведении социологических опросов нередко возникают ситуации, когда ответы респондентов на различные вопросы коррелируют в том смысле, что, например, почти все индивиды, ответившие положительно на первый вопрос социологической анкеты, ответили отрицательно на третий и четвертый вопросы. Такие закономерности свидетельствуют о том, что в отношении изучаемой в данном опросе проблематики социум обладает латентной (скрытой) структурой. Например, он может состоять из нескольких классов, так что представители одного и того же класса отвечают на вопросы почти одинаковым образом.

Математическим аппаратом для выявления скрытой структуры социума служит латентно-структурный анализ, которому посвящен настоящий раздел. Он является латентным в том смысле, что указанные классы задают размежевание социума, которое не является априорно явным, как, например разбиение по возрастному или национальному признаку. Если социолога интересуют явные признаки, то он уже при планировании исследования может концентрировать свое внимание, например, на проблеме того, как люди разных национальностей относятся к разводам; или как относятся к разводам люди разных возрастов. В противоположность этому латентно-структурный анализ не использует априорных предположений о том, какие именно факторы определяют отношение индивидов к данному кругу вопросов. Более того, может оказаться, что латентная структура классов не соответствует никакому из «явных» разбиений социума.

Первые три пункта данного раздела посвящены так называемой *модели латентных классов* — варианту латентно-структурного анализа, при котором латентная переменная принимает значения на конечном множестве. В четвертом пункте кратко рассматривается более сложный случай, когда областью значений латентной переменной является непрерывное множество.

4.1.1. Основные положения модели латентных классов

В качестве вводного примера рассмотрим следующую ситуацию. Предположим, что в ходе некоторого социологического опроса респондентам задавалось 5 дихотомических (т. е. предполагающих лишь два

варианта ответов, например ДА и НЕТ) вопросов. В ходе обработки результатов опроса было замечено, что большинство респондентов, ответивших ДА на первый вопрос, ответило НЕТ все остальные вопросы. Что касается тех респондентов, которые ответили НЕТ на первый вопрос, то большинство из них либо отвечали ДА на третий и четвертый, и НЕТ — на второй и пятый, либо наоборот. Таким образом, множество всех респондентов можно описать как состоящее в основном из трех групп, называемых латентными классами. Каждый класс представлен своим профилем ответов на вопросы, см. табл. 1, в которой ответ ДА обозначен знаком «+», а ответ НЕТ — знаком «-».

Таблица 1. Пример латентных классов, возникающих при ответах респондентов на 5 вопросов

Латентный класс	Вопрос 1	Вопрос 2	Вопрос 3	Вопрос 4	Вопрос 5
Первый	+	-	-	-	-
Второй	-	-	+	+	-
Третий	-	+	-	-	+

Полученное разбиение на классы говорит о том, что позиция конкретного индивида по данному кругу вопросов определяется (в основном) некоторой латентной переменной, связанной с этим разбиением. Эту переменную можно понимать как номер класса. Таким образом, в данном примере она принимает значения на множестве $\{1; 2; 3\}$.

Вариант латентно-структурного анализа, при котором область изменения латентной переменной является конечным множеством, называется *моделью латентных классов*.

В реальном социологическом исследовании закономерности, подобные представленной в табл. 1, как правило, не бывают настолько ярко выраженными, чтобы их удалось установить при «визуальном осмотре» данных. Поэтому при выявлении латентной структуры социума возникают следующие вопросы.

— Как выявить латентные классы?

— К какому классу отнести «нетипичных» индивидов, ответы которых не вписываются напрямую ни в один из классов?

Анализ этих проблем составляет основное содержание математического анализа модели латентных классов (отметим, что в реальных социологических исследованиях важную роль играют также вопросы статистической обработки данных).

Рассмотрим данную методику для простейшего случая, когда респондентам задаются три вопроса, а количество латентных классов равно двум. Схема представлена на рис. 1. Через s^1 и s^2 обозначены доли классов в общей численности респондентов, при этом $s^1 + s^2 = 1$. Вероятности положительного ответа респондента из i -го класса на j -й вопрос обозначены через p_i^j ($i = 1, 2, 3, j = 1, 2$).

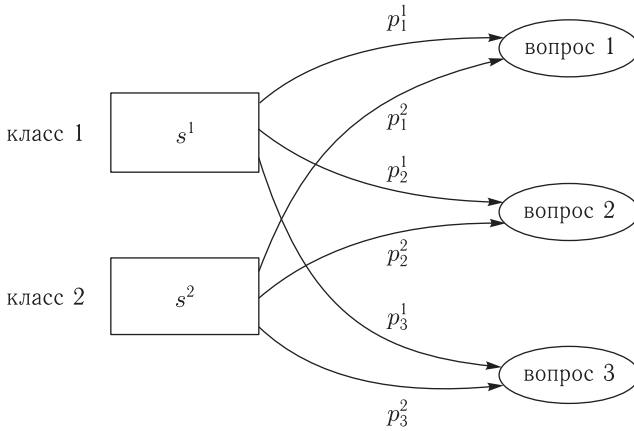


Рис. 1. Схема модели латентных классов для случая двух классов и трех вопросов

Введем также эмпирически определяемые доли респондентов q_i , q_{ik} , q_{123} ($i, k = 1, 2, 3$; $i \neq k$), положительно ответивших на вопросы, номера которых присутствуют в нижних индексах. Например, q_2 — это доля индивидов, положительно ответивших на второй вопрос (и произвольным образом — на первый и третий вопросы); q_{12} — доля индивидов, положительно ответивших на первый и второй вопросы (и произвольным образом — на третий). Эти величины составляют входную информацию и для математической модели являются экзогенно заданными; в социологии они называются наблюдаемыми переменными. Величины вида p_i^j , s^j подлежат определению. Пример входных данных для анализа дается табл. 2.

Перейдем к составлению системы уравнений для случая двух классов и трех вопросов.

Доля (от общей численности респондентов N_0) членов первого латентного класса, положительно ответивших на первый вопрос (и произвольным образом — на остальные вопросы), составляет $s^1 p_1^1$, членов второго класса — $s^2 p_1^2$. С другой стороны (см. табл. 2), доля положительно ответивших на первый вопрос составляет $q_{123} + q_{12} + q_{13} + q_1$. Таким образом, получаем уравнение

$$s^1 p_1^1 + s^2 p_1^2 = q_{123} + q_{12} + q_{13} + q_1. \quad (1)$$

Аналогично, рассматривая доли респондентов, положительно ответивших, соответственно, на второй и третий вопросы, получаем уравнения

$$s^1 p_2^1 + s^2 p_2^2 = q_{123} + q_{12} + q_{23} + q_2, \quad (2)$$

$$s^1 p_3^1 + s^2 p_3^2 = q_{123} + q_{13} + q_{23} + q_3. \quad (3)$$

Таблица 2. Пример входных данных для модели латентных классов

Вопрос 1	Вопрос 2	Вопрос 3	Кол-во респондентов	Доля респондентов
+	+	+	126	$q_{123} = 0,0252$
+	+	-	294	$q_{12} = 0,0588$
+	-	+	504	$q_{13} = 0,1008$
-	+	+	494	$q_{23} = 0,0988$
+	-	-	1176	$q_1 = 0,2352$
-	+	-	86	$q_2 = 0,0172$
-	-	+	1976	$q_3 = 0,3952$
-	-	-	344	0,0688
Всего респондентов			$N_0 = 5000$	

Доля (от общей численности респондентов) членов первого латентного класса, положительно ответивших на первый и второй вопросы (и произвольным образом — на третий вопрос), составляет $s^1 p_1^1 p_2^1$, членов второго класса — $s^2 p_1^2 p_2^2$. С другой стороны (см. табл. 2), доля положительно ответивших на первый и второй вопросы составляет $q_{123} + q_{12}$. Таким образом, получаем уравнение

$$s^1 p_1^1 p_2^1 + s^2 p_1^2 p_2^2 = q_{123} + q_{12}. \quad (4)$$

Аналогично, рассматривая доли респондентов, положительно ответивших на первый и третий, а затем — на второй и третий вопросы, получим уравнения

$$s^1 p_1^1 p_3^1 + s^2 p_1^2 p_3^2 = q_{123} + q_{13}, \quad (5)$$

$$s^1 p_2^1 p_3^1 + s^2 p_2^2 p_3^2 = q_{123} + q_{23}. \quad (6)$$

Наконец, доля (от общей численности респондентов) членов первого латентного класса, положительно ответивших на все три вопроса, составляет $s^1 p_1^1 p_2^1 p_3^1$, членов второго класса через $s^2 p_1^2 p_2^2 p_3^2$. С другой стороны (см. табл. 2), доля положительно ответивших на первый вопрос составляет q_{123} . Таким образом, получаем уравнение

$$s^1 p_1^1 p_2^1 p_3^1 + s^2 p_1^2 p_2^2 p_3^2 = q_{123}. \quad (7)$$

Еще одно уравнение следует из того, что два латентных класса полностью исчерпывают множество респондентов:

$$s^1 + s^2 = 1. \quad (8)$$

Уравнения (1)–(8) представляют модель латентных классов для случая двух классов и трех вопросов. Эта система содержит 8 уравнений для 8 неизвестных, но является нелинейной. Поэтому существование и единственность решения для нее, вообще говоря, не гарантированы. Выясним, сколько уравнений и сколько неизвестных содержит модель латентных классов в более общем случае, когда респонденты отвечают на n вопросов, а количество классов равно k . Неизвестными являются доли классов s^1, \dots, s^k , а также вероятности положительных ответов на вопросы, p_i^j , $i = 1, \dots, n$; $j = 1, \dots, k$; таким образом, система имеет $k + kn$ неизвестных. Нетрудно показать, что количество уравнений равно 2^n (см. упражнение 1). Таким образом, с формальной точки зрения при различных значениях k и n возможны как случаи, при которых количество уравнений больше, чем количество неизвестных, так и наоборот (они равны лишь в разобранном выше случае $k = 2$, $n = 3$). Если количество вопросов n (или, что то же самое, количество уравнений 2^n) недостаточно велико, то провести разбиение множества респондентов на заданное количество латентных классов k единственным образом невозможно. С точки зрения применимости данного метода в социологических исследованиях это означает, что при фиксированном количестве вопросов можно провести разбиение на количество классов, достаточно малое, чтобы удовлетворялось неравенство $2^n \geq k + kn$, т. е. $k \leq 2^n / (n + 1)$. На практике, однако, это условие оказывается не слишком ограничительным, например: если респондентам задается $n = 8$ вопросов, то максимально возможное количество латентных классов равно 28 при том, что социологическому осмыслению поддается лишь существенно меньшее количество латентных классов. Таким образом, количество уравнений обычно превышает количество неизвестных (за исключением случая $k = 2$, $n = 3$). В этом случае система почти всегда является несовместной, и в качестве искомых величин принимаются значения, удовлетворяющие уравнениям наилучшим образом (например, минимизирующие сумму квадратов невязок уравнений). То же относится и к случаю, когда система (1)–(8) несовместна. Заметим также, что если количество вопросов равно трем, то невозможно разбить респондентов более чем на два класса.

4.1.2. Отнесение отдельных индивидов к конкретным латентным классам

После того, как найдены значения s^j , p_i^j , $i = 1, \dots, n$; $j = 1, \dots, k$, возникает вопрос, как распределить по классам конкретных индивидов, демонстрирующих «нетипичное» поведение.

Рассмотрим, для примера, значения

$$\begin{aligned} s^1 &= 0,4, & p_1^1 &= 0,9, & p_2^1 &= 0,2, & p_3^1 &= 0,2, \\ s^2 &= 0,6, & p_1^2 &= 0, & p_2^2 &= 0,2, & p_3^2 &= 0,9 \end{aligned} \quad (9)$$

(с помощью прямой подстановки нетрудно убедиться, что они соответствуют входным данным, представленным в табл. 2). Содержательно

эти данные означают, что первый класс состоит в основном из индивидов, отвечающих положительно на первый вопрос и отрицательно на остальные, а второй класс — из индивидов, отвечающих положительно на третий вопрос и отрицательно на остальные вопросы. Поскольку эти два класса полностью исчерпывают множество респондентов, то в них входят и «нетипичные» индивиды.

Рассмотрим, например 504 респондента, которые положительно ответили на первый и третий вопросы, и отрицательно — на второй (см. соответствующую строку в табл. 2). Вычислим, сколько из них относятся к первому классу, а сколько — ко второму. Для представителя первого латентного класса вероятность того, что он ответит на вопросы подобным образом, составляет

$$p_1^1(1 - p_2^1)p_3^1 = 0,9 \cdot (1 - 0,2) \cdot 0,2 = 0,144.$$

Таким образом, поскольку численность первого класса составляет $N_0 s^1 = 2000$ респондентов, то $2000 \cdot 0,144 = 288$ ответили положительно на первый и третий вопросы и отрицательно — на второй. Аналогично, получаем численность ответивших подобным образом членов второго класса:

$$N_0 s^2 p_1^2(1 - p_2^2)p_3^2 = 3000 \cdot 0,1 \cdot (1 - 0,2) \cdot 0,9 = 216.$$

Итак, из 504 индивидов, ответивших подобным образом, 288 относятся к первому классу, а 216 — ко второму. Следовательно, конкретный индивид относится к первому классу с вероятностью, равной $288/504$, и ко второму классу — с вероятностью, равной $216/504$.

Аналогичным образом могут быть получены вероятности принадлежности к тому или иному классу для остальных респондентов.

Задача, близкая к только что рассмотренной, возникает в случае неполной информации. Предположим, например, что для некоторого индивида известно, что он положительно ответил на первый и второй вопросы, но неизвестно, как он ответил на третий вопрос. С какой вероятностью данный индивид относится к тому или иному классу? Такая постановка задачи предполагает, что индивид является представителем социума, для которого рассмотренная выборка является репрезентативной. Другими словами, выявленные латентные классы характеризуют весь социум численности N_{soc} (а не только участников социологического опроса).

Аналогично тому, как это сделано выше для случая, когда известна информация об ответах на все три вопроса, вычислим численности членов первого и второго латентных классов, положительно ответивших на первые два вопроса. Она равны: $N_{\text{soc}} s^1 p_1^1 p_2^1 = 0,072 N_{\text{soc}}$ и $N_{\text{soc}} s^2 p_1^2 p_2^2 = 0,016 N_{\text{soc}}$. Следовательно, вероятность того, что индивид принадлежит к первому латентному классу, составляет $0,072 N_{\text{soc}} / (0,072 N_{\text{soc}} + 0,016 N_{\text{soc}}) = 72/88 \approx 0,82$.

4.1.3. Пример применения модели латентных классов в социологическом исследовании

Первое исследование, в котором был применен латентно-структурный анализ, проводилось в рамках большого проекта «Американский солдат: социальная психология во время Второй мировой войны». В его рамках респондентам (американским военнослужащим) задавались, в частности, следующие четыре вопроса.

Вопрос 1: в целом, считаете ли Вы, что армией управляют хорошо?

Вопрос 2: как Вы думаете, останется ли у Вас хорошее отношение к армии после окончания службы?

Вопрос 3: в целом, думаете ли Вы, что армия относится к Вам справедливо?

Вопрос 4: как Вы думаете, пытается ли армия делать все возможное для блага тех, кто в ней служит?

Результаты опроса представлены в Табл.3: плюсы соответствуют ответам, показывающим положительное отношение к армии.

Таблица 3. Данные опроса, проведенного в рамках проекта «Американский солдат»

Вопрос 1	Вопрос 2	Вопрос 3	Вопрос 4	Численность респондентов
+	+	+	+	75
+	+	+	-	69
+	+	-	+	55
+	-	+	+	42
-	+	+	+	3
+	+	-	-	96
+	-	+	-	60
+	-	-	+	45
-	+	+	-	16
-	+	-	+	8
-	-	+	+	10
+	-	-	-	199
-	+	-	-	52
-	-	+	-	25
-	-	-	+	16
-	-	-	-	229

Полученные результаты имеют вид

$$s^1 = 0,4243, \quad p_1^1 = 0,9240, \quad p_2^1 = 0,6276, \quad p_3^1 = 0,5704, \quad p_4^1 = 0,5125, \\ s^2 = 0,5757, \quad p_1^2 = 0,4324, \quad p_2^2 = 0,1871, \quad p_3^2 = 0,1008, \quad p_4^2 = 0,0635.$$

Данные значения величин p_i^j показывают, что по каждому из четырех вопросов представители первого класса имеют большую вероятность положительного ответа, чем представители второго класса. Таким образом, результат имеет ясную социологическую трактовку: разбиение на латентный классы произошло в соответствии с отношением к армии: первый характеризуется (в целом) положительным отношением, второй класс — отрицательным. При этом, однако, даже среди тех, кто положительно относится к армии, лишь немногим более половины опрошенных считают, что армия относится к ним справедливо и делает все возможное для блага военнослужащих.

4.1.4. Латентно-структурный анализ в случае непрерывной латентной переменной

Как указано выше, номер класса можно понимать как латентную переменную x , заданную на множестве индивидов. В приведенном примере об американской армии эта переменная принимала значения на множестве $\{1; 2\}$, причем $x = 1$ соответствовало положительному отношению респондента к армии, $x = 2$ — отрицательному.

Однако не все, кто относится к армии хорошо, относятся к ней одинаково хорошо (и не все, кто относится к ней плохо, относятся к ней одинаково плохо). Поэтому для латентной переменной можно было бы ввести больше дискретных градаций (т. е. больше латентных классов) или считать ее непрерывной. Настоящий пункт посвящен варианту латентно-структурного анализа в случае непрерывной латентной переменной.

Без ограничения общности будем считать, что эта переменная принимает значения на отрезке $[0; 1]$. Для примера, далее будем считать, что речь идет об отношении респондентов к армии, причем «направление» латентной переменной выбрано так, что $x = 0$ соответствует *максимально* положительному отношению, а $x = 1$ — *максимально* отрицательному. Если в модели латентных классов предполагалось, что распределение индивидов характеризуется долями классов s^1, \dots, s^k (при этом $\sum_{j=1}^k s^j = 1$), то теперь оно описывается плотностью $p(x)$ так, что $p(x) dx$ есть доля респондентов, для которых значения переменной находится в интервале $(x, x + dx)$, причем

$$\int_0^1 p(x) dx = 1. \quad (10)$$

Как и в случае модели латентных классов, предполагается, что каждый из n задаваемых респондентам вопросов является дихотомии-

ческим, т. е. допускающим лишь ответы ДА и НЕТ. Вопрос с номером i характеризуется так называемой *функцией вопроса* $f_i(x)$, имеющей смысл вероятности того, что индивид со значением латентной переменной, равным x , даст на него положительный ответ. Таким образом, функция вопроса $f_i(x)$ является аналогом набора вероятностей p_i^1, \dots, p_i^k в модели латентных классов. Очевидно, что

$$0 \leq f_i(x) \leq 1 \text{ при } 0 \leq x \leq 1 \text{ для любого } i = 1, \dots, n.$$

Качественные характеристики функции $f_i(x)$ несут определенную информацию о «направлении» i -го вопроса. Так, монотонно возрастающая функция $f_i(x)$ соответствует вопросу, для которого ответ ДА свидетельствует об отрицательном отношении респондента к армии (рис. 2, а): действительно, чем больше значение x , тем хуже он относится к армии и тем более вероятно, что он ответит на данный вопрос положительно. Аналогично, монотонно убывающая функция $f_i(x)$ соответствует вопросу, для которого ответ ДА свидетельствует о положительном отношении (рис. 2, б). В частности, подобный вид имеют все функции вопросов исследования «Американский солдат», рассмотренные в третьем пункте. Возможны также вопросы, для которых вероятность положительного ответа достигает максимума (рис. 2, в) или минимума (рис. 2, г) в одной из внутренних точек отрезка $0 \leq x \leq 1$.

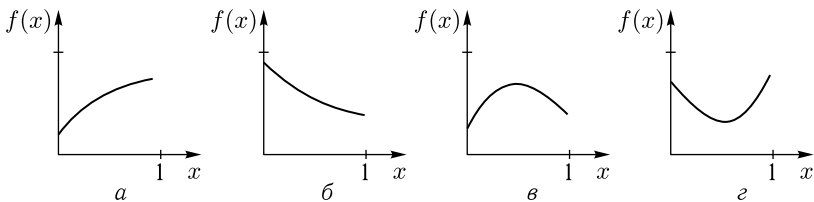


Рис. 2. Возможные формы функции вопроса $f_i(x)$

В соответствии со введенными определениями доля индивидов, положительно ответивших на i -й вопрос, равна

$$\int_0^1 p(x) f_i(x) dx, \quad i = 1, \dots, n.$$

Входные данные для латентно-структурного анализа с непрерывной латентной переменной имеют такой же вид, как и рассмотренный выше, т. е. вид, подобный табл. 2 и табл. 3. Поэтому измеренная эмпирически доля ответивших на каждый вопрос (или на комбинацию вопросов) может быть представлена так же, как и в модели латентных классов.

Таким образом, имеем (для простоты далее ограничимся рассмотрением случая трех вопросов)

$$\int_0^1 p(x)f_1(x) dx = q_{123} + q_{12} + q_{13} + q_1, \quad (11)$$

$$\int_0^1 p(x)f_2(x) dx = q_{123} + q_{12} + q_{23} + q_2, \quad (12)$$

$$\int_0^1 p(x)f_3(x) dx = q_{123} + q_{13} + q_{23} + q_3. \quad (13)$$

Полученные формулы (11)–(13) представляют собой аналоги формул (1)–(3). Далее, рассмотрев долю респондентов, положительно ответивших на первый и второй вопросы, получаем уравнение

$$\int_0^1 p(x)f_1(x)f_2(x) dx = q_{123} + q_{12}, \quad (14)$$

которое является аналогом уравнения (4). Аналогично, для долей респондентов, положительно ответивших на первый и третий, а затем — на второй и третий вопросы, имеем

$$\int_0^1 p(x)f_1(x)f_3(x) dx = q_{123} + q_{13}, \quad (15)$$

$$\int_0^1 p(x)f_2(x)f_3(x) dx = q_{123} + q_{23}. \quad (16)$$

Наконец, рассмотрев численность индивидов, положительно ответивших на все три вопроса, получаем аналог уравнения (7), который имеет вид

$$\int_0^1 p(x)f_1(x)f_2(x)f_3(x) dx = q_{123}. \quad (17)$$

В построенной системе уравнений (10)–(17) неизвестными являются функции $p(x)$, $f_1(x)$, $f_2(x)$, $f_3(x)$. Очевидно, такая система (в которой неизвестны функции, а не числа) имеет бесконечное количество решений. Поэтому решение данной системы можно искать в том или ином классе функций, если это не противоречит социологическому смыслу задачи.

В частности, если содержательный смысл задаваемых респондентам вопросов предполагает, что все функции $f_i(x)$ монотонны (см. рис. 2, а, б и комментарии к ним в тексте), то в некотором приближении эти функции можно считать линейными:

$$f_i(x) = a_i x + b_i, \quad i = 1, 2, 3. \quad (18)$$

Таким образом, каждая функция вопроса $f_i(x)$ характеризуется двумя параметрами. С учетом того, что система (10)–(17) состоит из 8 уравнений, а вследствие этого количество неизвестных параметров в данном случае также должно быть равно восьми, то получаем, что для функции плотности $p(x)$ также остается два параметра

Заметим, что с социологической точки зрения существенно более ценной была бы параметризация, допускающая немонотонную плотность $p(x)$. Действительно, распространены ситуации, когда наибольшее количество индивидов характеризуется промежуточным (а не крайним) значением переменной: например, если бы респондентов попросили оценить свое отношение к военной службе по пятибалльной шкале, то, вероятно, максимум распределения достигался бы на оценке 2, 3 или 4, но не 1 или 5. Однако для того, чтобы методом латентно-структурного анализа можно было построить немонотонную функцию плотности $p(x)$, недостаточно задать респондентам всего лишь по три вопроса (и тем самым располагать лишь двумя параметрами для описания этой функции). Напомним, что если количество вопросов равно трем, то модель латентных классов не позволяет разбить респондентов более чем на два класса. В данном случае ситуация аналогична.

Итак, примем для функции $p(x)$ представление

$$p(x) = a_0 x + b_0. \quad (19)$$

Таким образом, имеем систему восьми уравнений (10)–(17) для восьми неизвестных $a_i, b_i, i = 0, \dots, 3$. Подставляя формулы (18), (19) в (10)–(17), после ряда элементарных преобразований имеем

$$\frac{a_0}{2} + b_0 = 1, \quad (20)$$

$$\frac{a_0 a_1}{3} + \frac{a_0 b_1 + b_0 a_1}{2} + b_0 b_1 = q_{123} + q_{12} + q_{13} + q_1, \quad (21)$$

$$\frac{a_0 a_2}{3} + \frac{a_0 b_2 + b_0 a_2}{2} + b_0 b_2 = q_{123} + q_{12} + q_{23} + q_2, \quad (22)$$

$$\frac{a_0 a_3}{3} + \frac{a_0 b_3 + b_0 a_3}{2} + b_0 b_3 = q_{123} + q_{13} + q_{23} + q_3, \quad (23)$$

$$\frac{a_0 a_1 a_2}{4} + \frac{b_0 a_1 a_2 + a_0 a_1 b_2 + a_0 b_1 a_2}{3} + \frac{a_0 b_1 b_2 + b_0 a_1 b_2 + b_0 b_1 a_2}{2} + b_0 b_1 b_2 = q_{123} + q_{12}, \quad (24)$$

$$\frac{a_0 a_1 a_3}{4} + \frac{b_0 a_1 a_3 + a_0 a_1 b_3 + a_0 b_1 a_3}{3} + \frac{a_0 b_1 b_3 + b_0 a_1 b_3 + b_0 b_1 a_3}{2} + b_0 b_1 b_3 = q_{123} + q_{13}, \quad (25)$$

$$\frac{a_0 a_2 a_3}{4} + \frac{b_0 a_2 a_3 + a_0 a_2 b_3 + a_0 b_2 a_3}{3} + \frac{a_0 b_2 b_3 + b_0 a_2 b_3 + b_0 b_2 a_3}{2} + b_0 b_2 b_3 = q_{123} + q_{23}, \quad (26)$$

$$\frac{a_0 a_1 a_2 a_3}{5} + \frac{a_0 a_1 (a_2 b_3 + b_2 a_3) + (a_0 b_1 + b_0 a_1) a_2 a_3}{4} + \frac{a_0 a_1 b_2 b_3 + (a_0 b_1 + b_0 a_1) (a_2 b_3 + b_2 a_3) + b_0 b_1 a_2 a_3}{3} + \frac{(a_0 b_1 + b_0 a_1) b_2 b_3 + b_0 b_1 (a_2 b_3 + b_2 a_3)}{2} + b_0 b_1 b_2 b_3 = q_{123}. \quad (27)$$

Решение этой системы позволяет определить значения коэффициентов a_i , b_i ($i = 0, 1, 2, 3$) и тем самым получить искомые функции вопросов $f_i(x)$, $i = 1, 2, 3$, и плотность распределения $p_0(x)$. Тем самым латентная структура социума определена.

4.1.5. Заключение

Латентно-структурный анализ является эффективным инструментом в исследовании общества, к сожалению, практически не применяемым российскими социологами. До какого-то времени его распространение сдерживалось трудоемкостью (а иногда — и содержательной сложностью) математических вычислений. Однако современные программные средства позволяют использовать модель с практически любым количеством вопросов и латентных классов.

Упражнения

1. Рассмотреть модель латентных классов, когда количество вопросов, задаваемых респондентам, равно произвольному числу n . Найти количество уравнений, описывающих долю индивидов, положительно ответивших на один вопрос (и произвольным образом — на все остальные), положительно ответивших на два вопроса (и произвольным образом — на все остальные), и т. д. Доказать, что вся система содержит 2^n уравнений.

2. Показать, что если социологическая анкета состоит из n дихотомических вопросов, а функции вопросов $f_i(x)$, $i = 1, \dots, n$, имеют представление (18), то для функции плотности $p(x)$ можно выбрать представление с числом параметров, равным $2^{n-1}/n$.

Комментарии и библиографические примечания

1. Латентно-структурный анализ был разработан и впервые применен американским социологом П. Лазарсфельдом во второй половине

40-х годов XX века. Теоретическим основам метода посвящена работа [1] (см. также русский перевод [2]), а его применению — вышедшая в том же издании работа [3].

2. Проведенное в данном разделе изложение несколько отличается от оригинального (см. [1, 2]). Так, непрерывная латентная переменная в [1] задана на множестве $(-\infty; \infty)$, и функции $p(x)$, $f_i(x)$ имеют представления в виде полиномов. Это приводит к тому, что функции, которые, по замыслу автора, имеют смысл вероятности, могут принимать отрицательные значения и/или превышать единицу. Чтобы избежать данного логического недочета, в данной работе мы определяем латентную переменную на конечном множестве и предлагаем линейные представления для функций $p(x)$, $f_i(x)$.

3. Простое и весьма краткое изложение основных латентно-структурного анализа дано в [4]. Работы на русском языке крайне малочисленны: см., например, [5, 6].

Литература

1. *Lazarsfeld P. F.* The Logical and Mathematical Foundations of Latent Structure Analysis // Chapter 10 in Stouffer S. A., et al. Measurement and Prediction, Volume IV of The American Soldier: Studies in Social Psychology in World War II. Princeton University Press. 1950.
2. *Лазарсфельд П.* Логические и математические основания латентно-структурного анализа // Математические методы в современной буржуазной социологии. Под ред. Г. В. Осипова. М. Прогресс. 1966. С. 344–401 (перевод предыдущей работы в данном списке).
3. *Lazarsfeld P. F.* Some Latent Structures // Chapter 11 in Stouffer S. A., et al. Measurement and Prediction, Volume IV of The American Soldier: Studies in Social Psychology in World War II. Princeton University Press. 1950.
4. *Henry N. W.* Latent Structure Analysis at Fifty. Paper presented at the 1999 Joint Statistical Meetings, Baltimore MD, August 11, 1999. URL <http://www.people.vcu.edu/~nhenry/LSA50.htm>.
5. *Осипов Г. В., Андреев Э. П.* Методы измерения в социологии. Андреев, Осипов. Москва: Наука, 1977.
6. *Толстова Ю. Н.* Измерение в социологии. М.: Инфра-М, 2003.

4.2. Модели марковского типа

Как в естественных, так и в социальных науках марковскими называются процессы, при протекании которых *отсутствуют эффекты последствия*. Другими словами, вероятность изменения состояния элементов (частиц, индивидов) рассматриваемой системы в любой момент времени не зависит от их состояния в любые предыдущие моменты.

Применительно к задачам социологии под системой, в моделях марковского типа, обычно понимается некоторая, рассматриваемая

в дискретные моменты времени социальная общность (социум), каждый член которой может находиться в одной из нескольких групп (состояний, классов и т.д.). С течением времени члены общности могут переходить из одной группы в другую. Причем, если индивид остается в своем классе, то этот акт также называется переходом (тождественным). В каждый из моментов времени состояние социума характеризуется конечномерным вектором-столбцом. Его компоненты, расположенные в определенном порядке, имеют смысл численностей (или долей) рассматриваемых групп. Конечная цель математического исследования — нахождение этих компонент как функций времени, для чего используется хорошо разработанный и относительно простой аппарат марковских процессов.

Ниже изучаются модели, в которых для всех индивидов социума, находящихся в некотором классе, вероятность перехода в другой класс одна и та же. С чисто математической точки зрения эти модели отвечают «стандартным» марковским процессам. Также приводятся некоторые сведения соответствующей математической теории. Далее, рассматриваются модели социума, имеющего несколько латентных (скрытых) сегментов таких, что вероятности перехода различны для представителей различных сегментов. В этих более сложных ситуациях происходят несколько одновременных марковских процессов, причем непосредственному наблюдению доступна лишь их *смесь*.

В комментариях более подробно обсуждается понятие марковских процессов, связанные с ними измерения в рамках социологических исследований, а также приводятся некоторые результаты, дополняющие основное содержание данного раздела и библиографические ссылки.

Напомним о главном принципиальном ограничении, свойственном всем моделям марковского типа. Они опираются на основное допущение марковских процессов: *вероятности перехода индивида из одного класса в другой не зависят от предыстории этого индивида*. Обоснование этого допущения — отдельная непростая задача, без решения которой адекватность модели исследуемому явлению не гарантирована.

4.2.1. Динамика численности верующих

Последовательность математических процедур, применяемых в теории марковских процессов, проиллюстрируем на примере социологического исследования, проведенного в 2005 г., в ходе которого были опрошены около 6000 респондентов из 78 субъектов РФ. Респондентам задавались, в частности, вопросы об их отношении к вере в момент опроса и в предшествующие годы. Их ответы сведены в табл. 1.

Заметим, что в рамках настоящей модели не учитываются явления рождения, миграции и умирания в исследуемом социуме. Единственный рассматриваемый вид движения — это переход индивидов из одной группы в другую. При построении модели будем рассматривать динамику процесса в дискретном времени, приняв «нынешний»

Таблица 1. Отношение респондентов к вере (процент от количества опрошенных)

Отношения: стал	Был верующим	Был неверующим	Колесался между верой и неверием	Затруднился ответить
Стал/остался верующим	47,7	4,6	3,2	1,1
Стал/остался неверующим	1,4	16,4	0,3	0,3
Стал/остался колеблющимся между верой и неверием	1,0	3,2	8,6	0,4
Затруднился ответить	1,2	0,7	0,4	9,5

момент за $t = 1$, а «предыдущий», за $t = 0$. Всюду далее группы верующих, неверующих, колеблющихся и затрудняющихся ответить нумеруются соответственно индексами 1, 2, 3, 4. Обозначая через m_{ij} ($i, j = 1, \dots, 4$) долю от общего количества опрошенных, перешедших из i -го состояния в j -е, и используя данные табл. 1, получим составленную из чисел m_{ij} матрицу, элементы которой равны этим долям:

$$M = \begin{pmatrix} 0,477 & 0,014 & 0,010 & 0,012 \\ 0,046 & 0,164 & 0,032 & 0,007 \\ 0,032 & 0,003 & 0,086 & 0,004 \\ 0,011 & 0,003 & 0,004 & 0,095 \end{pmatrix}.$$

Состояние системы в момент времени t описывается вектором

$$X(t) = \begin{pmatrix} x_1(t) \\ x_2(t) \\ x_3(t) \\ x_4(t) \end{pmatrix},$$

где $x_i(t)$ — доля i -й группы в общем количестве опрошенных.

Суммированием элементов строк матрицы M нетрудно получить вектор $X(0)$, характеризующий «предыдущий» момент, а суммированием элементов столбцов — вектор «нынешнего» момента $X(1)$:

$$X(0) = \begin{pmatrix} 0,513 \\ 0,249 \\ 0,125 \\ 0,113 \end{pmatrix}, \quad X(1) = \begin{pmatrix} 0,566 \\ 0,184 \\ 0,132 \\ 0,118 \end{pmatrix}.$$

Матрица M отвечает исходным данным, полученным в результате опроса. На их основе необходимо составить матрицу P перехода системы из состояния при $t = 0$ в состояние при $t = 1$. Ее элементами служат вероятности p_{ij} переходов индивидов из j -й группы в i -ю:

$$P = \begin{pmatrix} p_{11} & \cdots & p_{14} \\ \cdots & \cdots & \cdots \\ p_{41} & \cdots & p_{44} \end{pmatrix}.$$

Во всех моделях настоящего раздела рассматривается лишь простейший случай, когда матрица P является постоянной.

Важным свойством матрицы P является то, что сумма элементов каждого ее столбца равна единице. Это следует из того, что элементы j -го столбца $p_{1j}, p_{2j}, \dots, p_{nj}$ представляют собой вероятности всевозможных переходов из j -й группы, а значит — вероятности событий, образующих полную группу событий (т.е. вероятности непересекающихся событий, одно из которых точно произойдет).

Вычислим теперь для рассматриваемых данных значения p_{i1} , т.е. вероятности для человека, находящегося при $t = 0$ в первой группе, попасть при $t = 1$ в i -ю группу ($i = 1, \dots, 4$). Как установлено выше, из каждой тысячи опрошенных в момент времени $t = 0$ к первой группе в среднем принадлежали 513. Из них в момент времени $t = 1$ в первой группе оставались 477, во вторую группу перешли 14, в третью — 10, в четвертую — 12. Следовательно, для представителя первой группы вероятность остаться в ней равна $p_{11} = \frac{477}{513}$, вероятность попасть во вторую, третью, четвертую —

$$p_{21} = \frac{14}{513}, \quad p_{31} = \frac{10}{513}, \quad p_{41} = \frac{12}{513}.$$

Аналогично найдем все остальные вероятности p_{ij} переходов из j -й группы в i -ю и получим матрицу перехода $P = \|\|p_{ij}\|\|$ для данного случая:

$$P = \begin{pmatrix} \frac{477}{513} & \frac{46}{249} & \frac{32}{125} & \frac{11}{113} \\ \frac{14}{513} & \frac{164}{249} & \frac{3}{125} & \frac{3}{113} \\ \frac{10}{513} & \frac{32}{249} & \frac{86}{125} & \frac{4}{113} \\ \frac{12}{513} & \frac{7}{249} & \frac{4}{125} & \frac{95}{113} \end{pmatrix}.$$

Прямая проверка показывает, что, как и следовало ждать, $X(1) = PX(0)$.

В предположении, что матрица P с течением времени остается неизменной, вектор $X(t)$ для любого значения t вычисляется через

вектор $X(t-1)$ по простой рекуррентной формуле

$$X(t) = PX(t-1) = P^2X(t-2) = \dots = P^tX(0).$$

В частности, нетрудно вычислить вектор $X(2)$,

$$X(2) = \begin{pmatrix} \frac{477}{513} & \frac{46}{249} & \frac{32}{125} & \frac{11}{113} \\ \frac{14}{513} & \frac{164}{249} & \frac{3}{125} & \frac{3}{113} \\ \frac{10}{513} & \frac{32}{249} & \frac{86}{125} & \frac{4}{113} \\ \frac{12}{513} & \frac{7}{249} & \frac{4}{125} & \frac{95}{113} \end{pmatrix} \begin{pmatrix} 0, 566 \\ 0, 184 \\ 0, 132 \\ 0, 118 \end{pmatrix} \approx \begin{pmatrix} 0, 606 \\ 0, 143 \\ 0, 130 \\ 0, 122 \end{pmatrix},$$

и сделать краткосрочный «прогноз» относительно численности каждой из групп.

В связи с вопросом о прогнозе дадим следующее определение.

Определение. Неотрицательная матрица, обладающая свойством

$$p_{1j} + p_{2j} + \dots + p_{nj} = 1$$

для любого $j = 1, \dots, n$, называется стохастической (или марковской, или вероятностной).

Она имеет следующие основные свойства (см. также упражнения 1, 2).

1. Все элементы стохастической матрицы неотрицательны.
2. Стохастическая матрица имеет собственное значение, равное единице.
3. Любая натуральная степень стохастической матрицы также является стохастической матрицей.

Отметим, что полученная выше матрица P является стохастической.

При последовательном вычислении с помощью матрицы P векторов $X(3)$, $X(4)$ и т. д. оказывается, что с ростом t вектор $X(t)$ приближается к некоторому предельному вектору $X_s = \lim_{t \rightarrow \infty} X(t)$, который, по построению, очевидно является решением уравнения

$$X_s = PX_s,$$

т. е. собственным вектором матрицы P , соответствующим собственному значению, равному единице.

Полученное значение вектора $X(2)$ может быть использовано для построения прогноза численности верующих россиян на ближайшее будущее. Однако при изучении ее динамики на достаточно большом временном интервале необходимо учитывать как демографические процессы, так и неизбежное изменение с течением времени матрицы перехода P . Поэтому вектор X_s нельзя трактовать как долгосрочный прогноз состояния системы. Практическая ценность вектора X_s

заключается в том, что он позволяет определить общую тенденцию, т. е. направление процесса.

Действительно, на основе одной только краткосрочной динамики, т. е. путем сравнения векторов $X(0)$, $X(1)$, невозможно определить, имеет ли та или иная конфессиональная группа тенденцию к увеличению или уменьшению численности. Например, см. упражнение 4, в котором имеет место случай, когда для некоторой группы выполняется неравенство $x_i(0) < x_i(1)$, однако, тем не менее, значение i -й компоненты вектора X_s меньше, чем $x_i(0)$. Такая ситуация возникает в случае, когда в нулевой момент времени достаточно большая часть индивидов находится в группе, из которой переход в i -ю группу происходит наиболее активно.

Для данных указанного социологического исследования [11] собственный вектор имеет вид

$$X_s = \begin{pmatrix} 0,704 \\ 0,73 \\ 0,89 \\ 0,134 \end{pmatrix}.$$

Замечание. В рассматриваемом случае все элементы матрицы P положительны. Но при моделировании социальных процессов нередко встречаются ситуации, когда один или несколько элементов матрицы перехода равны нулю. Например, если находящиеся в некоторой группе индивиды не могут перейти в другие группы, то в соответствующем столбце диагональный элемент равен единице, а остальные элементы — нулю. Такая группа называется поглощающей. В частности, поглощающие группы возникают при моделировании социальных систем, в которых разделение на группы проводится по возрасту: поглощающей является старшая возрастная группа (см., напр., разд. 3 в главе 1). Другой пример поглощающих групп есть в разделе комментарии. Заметим, что если имеется поглощающая группа j , то у матрицы P^n ($n \in N$) все элементы j -го столбца, кроме диагонального, также равны нулю (см. упражнение 5).

Таким образом, представленная модель позволяет как построить краткосрочный прогноз количества верующих, неверующих и колеблющихся, так и определить общую тенденцию процесса.

4.2.2. Элементы общей теории марковских процессов

Рассмотрим более подробно основы математического аппарата марковских процессов.

В представленной выше модели динамики конфессиональной принадлежности россиян состояние системы, т. е. распределение индивидов по группам, с течением времени приближается к собственному вектору X_s (стационарному состоянию). Является ли это свойство существования «социального аттрактора» общим для всех марковских процессов?

Дадим ответ с помощью простого примера.

Пример 1. Пусть матрица перехода имеет вид

$$P = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

Тогда каждый индивид на каждом временном шаге меняет свою групповую принадлежность:

$$\begin{aligned} x_1(0) &= x_2(1) = x_1(2) = x_2(3) = \dots, \\ x_2(0) &= x_1(1) = x_2(2) = x_1(3) = \dots. \end{aligned}$$

Данная социальная система, несмотря на наличие единичного собственного значения и соответствующего ему собственного вектора $X_s = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$, не стремится к стационару, а испытывает периодические колебания (кроме особого случая, когда начальные условия выбраны так, что $x_1(0) = x_2(0)$).

Итак, не во всех случаях марковский процесс приводит к реализации стационарного состояния. Поэтому опишем достаточные условия стабилизации системы к стационару, для чего введем сначала следующие определения.

Определение. Стохастическая матрица (п.4.2.1) называется примитивной (регулярной), если некоторая ее натуральная степень положительна.

Определение. Марковский процесс с примитивной матрицей называется регулярным.

Для примитивных матриц имеет место следующая теорема Фробениуса–Перрона.

Теорема 1 (Фробениуса–Перрона). Пусть A — примитивная (регулярная) матрица. Тогда она имеет собственное значение λ_F такое, что

- 1) λ_F является вещественным, положительным, простым (т. е. единичной кратности),
- 2) λ_F строго превосходит модули всех других собственных значений,
- 3) имеют место неравенства для сумм по столбцам

$$\min_j \sum_i a_{ij} \leq \lambda_F \leq \max_j \sum_i a_{ij}$$

(причем оба неравенства являются строгими, если $\min_j \sum_i a_{ij} < \max_j \sum_i a_{ij}$), а также аналогичные неравенства для соответствующих сумм по строкам,

- 4) все координаты собственного вектора h_F , соответствующего λ_F , положительны.

При этом λ_F и h_F называются числом и вектором Фробениуса матрицы A . Далее будем считать, что h_F нормирован так, что сумма его компонент равна единице:

$$h_{F1} + h_{F2} + \dots + h_{Fn} = 1.$$

Теорема Фробениуса–Перрона дает возможность сформулировать достаточное условие стабилизации системы к стационарному состоянию. Заметим, что эти условия легко проверяются для матриц не очень большой размерности.

Пусть матрица перехода P — примитивная матрица. Применим к ней теорему Фробениуса–Перрона. Поскольку она является стохастической (см. свойство 3 стохастических матриц), сумма элементов каждого ее столбца равна единице, то непосредственно из третьего утверждения теоремы следует, что $\lambda_F = 1$. Существенным является не только само существование собственного значения $\lambda = 1$ (вытекающего из свойства 2), но и то, что оно является числом Фробениуса, так как из первого и второго утверждений теоремы следует, что все другие собственные значения матрицы P имеют модули, меньшие единицы.

Сделать вывод о поведении вектора $X(t)$ при $t \rightarrow \infty$ проще всего, если все собственные значения матрицы P различны. В этом случае начальный вектор $X(0)$ может быть разложен по ее собственным векторам $h_F = h_1, h_2, \dots, h_n$ как по базису:

$$X(0) = C_1 h_F + C_2 h_2 + \dots + C_n h_n.$$

Тогда после k -кратного умножения слева на матрицу P имеем

$$\begin{aligned} P^k X(0) &= C_1 P^k h_F + C_2 P^k h_2 + \dots + C_n P^k h_n = \\ &= C_1 \lambda_F^k h_F + C_2 \lambda_2^k h_2 + \dots + C_n \lambda_n^k h_n, \end{aligned}$$

а так как $\lambda_F = 1$ есть максимальное по модулю собственное значение, то все другие собственные значения по модулю строго меньше единицы, а значит, $\lambda_i^k \rightarrow 0$ при $k \rightarrow \infty$ для всех номеров $i = 2, \dots, n$. Тем самым

$$\lim_{k \rightarrow \infty} P^k X(0) = C_1 h_F,$$

т. е. стационарный вектор $X_s = C_1 h_F$ есть вектор Фробениуса, нормированный на общее количество индивидов.

В более общем случае, т. е. при наличии, вообще говоря, кратных собственных значений у матрицы перехода, механизм стабилизации к стационарному состоянию является менее наглядным. В этом случае полезной оказывается следующая теорема.

Теорема 2. Если P — примитивная матрица, то

$$\lim_{k \rightarrow \infty} P^k = \begin{pmatrix} h_{F1} & h_{F1} & \dots & h_{F1} \\ h_{F2} & h_{F2} & \dots & h_{F2} \\ \dots & \dots & \dots & \dots \\ h_{Fn} & h_{Fn} & \dots & h_{Fn} \end{pmatrix}.$$

Матрица $\lim_{k \rightarrow \infty} P^k$ обладает следующим свойством: при умножении ее на любой вектор X получается вектор $h_F \sum_{i=1}^n x_i$. Действительно,

$$\begin{aligned} \lim_{k \rightarrow \infty} P^k \begin{pmatrix} x_1 \\ x_2 \\ \dots \\ x_n \end{pmatrix} &= \begin{pmatrix} h_{F1} & h_{F1} & \dots & h_{F1} \\ h_{F2} & h_{F2} & \dots & h_{F2} \\ \dots & \dots & \dots & \dots \\ h_{Fn} & h_{Fn} & \dots & h_{Fn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \dots \\ x_n \end{pmatrix} = \\ &= \begin{pmatrix} h_{F1}(x_1 + x_2 + \dots + x_n) \\ h_{F2}(x_1 + x_2 + \dots + x_n) \\ \dots \\ h_{Fn}(x_1 + x_2 + \dots + x_n) \end{pmatrix} = (x_1 + x_2 + \dots + x_n) \begin{pmatrix} h_{F1} \\ h_{F2} \\ \dots \\ h_{Fn} \end{pmatrix}. \end{aligned}$$

Итак,

$$\lim_{k \rightarrow \infty} X(k) = \lim_{k \rightarrow \infty} P^k X(0) = (x_1 + x_2 + \dots + x_n) h_F.$$

В частности, если x_i ($i = 1, \dots, n$) имеет смысл доли социума, находящейся в i -й группе, то $x_1 + x_2 + \dots + x_n = 1$ и $\lim_{k \rightarrow \infty} X(k) = h_F$, т. е. финальное соотношение между численностями групп определяется вектором Фробениуса.

Из теоремы 2 следует также устойчивость стационарного состояния: малые возмущения начального вектора $X(0)$ приводят к малым изменениям финального вектора $\lim_{k \rightarrow \infty} X(k) = (x_1 + x_2 + \dots + x_n) h_F$.

Таким образом, примитивность стохастической матрицы — это нетрудно проверяемое и достаточно общее достаточное условие стабилизации марковского процесса к стационарному состоянию (которое является единственным). Оно находится либо путем определения вектора Фробениуса и его перенормировки на общее количество индивидов, либо непосредственным решением уравнения $X_s = P X_s$ с помощью хорошо известных вычислительных методов.

Применительно к прикладным задачам (в т. ч. из области социологии) существование стационарного состояния означает, что система имеет общую тенденцию к стабилизации, а его отсутствие указывает на циклический характер изменений.

4.2.3. Вертикальная социальная мобильность

Пусть в обществе можно выделить несколько непересекающихся классов, упорядоченных по принципу «сверху–вниз». Изучаемый феномен — *вертикальная социальная мобильность*, т. е. перемещение индивидов между классами с течением времени. Индивидами в данном случае считаются только мужчины — подразумевается, что групповая принадлежность всех членов семьи определяется положением ее главы. Тем самым при рождении мальчик попадает в тот же класс, что и его

отец. С течением времени он может в силу разных причин оказаться, вообще говоря, в каком-либо другом классе.

Описываемая ниже модель, отвечающая социальным реалиям Британии первой половины XX века, была предложена для трехклассовой и семиклассовой систем, она содержит следующие предположения:

- в каждой семье рождается в среднем один мальчик, семьи не распадаются, все дети рождаются в браке;
- миграция пренебрежимо мала.

В результате социологического исследования были выделены семь групп, описание которых и их доли в мужском населении сведены в табл. 2. Также в этой таблице приведены результаты агрегирования,

Таблица 2. Данные для семиклассового и трехклассового вариантов модели социальной мобильности

Описание группы	%	Класс	%
1. Professional and high administrative (топ-менеджеры)	2,9	Высший	7,5
2. Managerial and executive (руководители высокого уровня)	4,6		
3. Higher grade supervisory and non-manual (руководители среднего уровня)	9,4	Средний	63,4
4. Lower grade supervisory and non-manual (руководители низшего уровня)	13,1		
5. Skilled manual and non-manual (квалифицированный ручной и умственный труд)	40,9		
6. Semi-skilled manual (полуквалифицированный ручной труд)	17,0	Низший	29,1
7. Non-skilled manual (неквалифицированный ручной труд)	12,1		

необходимые для трехклассового варианта модели (см. два последних столбца). Заметим, что основой для классовой стратификации в данном исследовании является не характер потребления или размер располагаемых ресурсов (что более типично для сегодняшней социологии, см. раздел комментарии), а характер выполняемой работы и позиция, занимаемая в служебной иерархии. Описания групп даны на языке оригинала ввиду того, что в для них не всегда можно найти точный русский эквивалент (указанные в таблице эквиваленты следует считать приблизительными).

Полученные на основе данных исследования матрицы эмпирических вероятностей перехода для семиклассового и трехклассового вариантов

модели имеют, соответственно, вид

$$P_7 = \begin{pmatrix} 0,338 & 0,107 & 0,035 & 0,021 & 0,009 & 0 & 0 \\ 0,146 & 0,267 & 0,101 & 0,039 & 0,024 & 0,013 & 0,008 \\ 0,202 & 0,227 & 0,188 & 0,112 & 0,075 & 0,041 & 0,036 \\ 0,062 & 0,120 & 0,191 & 0,212 & 0,123 & 0,088 & 0,083 \\ 0,140 & 0,206 & 0,357 & 0,430 & 0,473 & 0,391 & 0,364 \\ 0,047 & 0,053 & 0,067 & 0,124 & 0,171 & 0,312 & 0,235 \\ 0,015 & 0,020 & 0,061 & 0,062 & 0,125 & 0,155 & 0,274 \end{pmatrix},$$

$$P_3 = \begin{pmatrix} 0,45 & 0,05 & 0,01 \\ 0,48 & 0,70 & 0,50 \\ 0,07 & 0,25 & 0,49 \end{pmatrix}.$$

Напомним, что элемент матрицы, стоящий на пересечении i -й строки и j -го столбца, имеет смысл эмпирической вероятности того, что индивид, находящийся в момент времени t в j -м классе, окажется в момент $t + 1$ в j -м классе.

Поясним, что матрица P_3 не может быть получена непосредственным агрегированием элементов матрицы P_7 , поскольку для ее вычисления необходимо использовать также данные о численности каждой из семи групп (см. упражнение 6).

Очевидно, что матрицы P_7 и P_3 являются примитивными, поэтому моделируемая социальная система с течением времени стабилизируется к стационарному состоянию, которое может быть найдено из уравнений $X_s = P_7 X_s$ и $X_s = P_3 X_s$ для семи- и трехгрупповой стратификации, соответственно.

Анализ матриц P_7 и P_3 показывает, в частности, что в рассматриваемой ситуации индивиды из высшего и низшего классов с примерно равной вероятностью остаются в своем классе или переходят в средний. В то же время наименее мобильными являются представители «низшего среднего класса»: максимальный элемент матрицы P_7 — это $p_{55} = 0,473$.

Для социологии одной из наиболее интересных представляется задача определения на основе описанной модели среднего количества поколений, которые семья проводит в своем классе (см. упражнение 7). Чем большее значение имеет эта величина, тем менее мобильным является общество.

Для предварительного количественного сравнения различных социумов в отношении их социальной мобильности полезно ввести и использовать следующие «полярные» идеализации.

Совершенно мобильным называется общество, в котором вероятность попасть в данную группу не зависит от группы, в которой

индивид был рожден. В этом случае матрица перехода дается выражением (для краткости ограничимся трехклассовым вариантом модели)

$$P_m = \begin{pmatrix} a & a & a \\ b & b & b \\ 1 - a - b & 1 - a - b & 1 - a - b \end{pmatrix},$$

где $0 < a + b < 1$. Его можно назвать также обществом равных возможностей.

Абсолютно ригидным называется общество, в котором вертикальная мобильность полностью отсутствует. Тогда матрица перехода имеет вид

$$P_r = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Естественно считать, что чем ближе некоторый социум к совершенно мобильному (матрица P_m), чем менее он сходен с абсолютно ригидным (матрица P_r), тем более мобильным он является. Заметим, однако, что теория (как и систематические эмпирические исследования по определению вероятностей перехода), сопоставляющая каждой стохастической матрице скалярную величину, которую можно было трактовать как общий, агрегированный уровень мобильности общества, пока отсутствует.

Модель также позволяет исследовать вертикальную мобильность в некоторых социумах, имеющих более сложную структуру. Рассмотрим, как пример, общество, состоящее из двух каст, в каждой из которых выделяются два класса — высший и низший. В течение жизни индивид остается в одной и той же касте, но может перемещаться между классами. Матрица перехода имеет в этом случае вид

$$P = \begin{pmatrix} a & b & 0 & 0 \\ 1 - a & 1 - b & 0 & 0 \\ 0 & 0 & c & d \\ 0 & 0 & 1 - c & 1 - d \end{pmatrix},$$

где первые две строки и столбца соответствуют классам первой касты, последние две строки и столбца — классам второй касты.

Модель допускает и другие обобщения. Например, в нее может быть введена дифференцированная фертильность — различное среднее количество мальчиков, рождающихся в семьях из разных классов. Однако в таком случае матрица перехода не является стохастической, что усложняет исследование (см. раздел Комментарии).

Приведенные выше примеры свидетельствуют о том, что марковские процессы могут быть эффективным средством исследования вертикальной мобильности в социуме.

4.2.4. Модели типа мувер-стейер

В рамках рассмотренных выше моделей марковского типа индивиды предполагались полностью идентичными — для всех них вероятность переходов одинаковы. Это предположение оправдано не всегда. Например, в модели динамики количества верующих (см. первый пункт) тот эмпирический факт, что из 513 верующих 14 стали неверующими, трактовался так, что для каждого из верующих вероятность стать неверующим в течение модельной единицы времени равна $14/513$. Более реалистично предположить, что некоторые из этих 513 верующих настолько тверды, что достоверно остаются в своем классе. Такие индивиды называются «оседлыми», *стейерами* (stayers). Обозначим численность верующих стейеров через s_1 . Тогда для каждого «кочующего», *мувера* (mover), т. е. индивида, способного к переходам между классами, вероятность из верующего стать неверующим увеличивается по сравнению с моделью из первого пункта и равна $\frac{14}{513 - s_1}$.

В используемых для анализа подобных ситуаций моделях типа «мувер-стейер» предполагается наличие двух типов населения, причем представители только одного из них (муверы) могут переходить из одного класса в другой. Каждый представитель другого типа населения (стейеры) пребывает в одном и том же классе во все моменты времени. Наблюдаемая при социологическом исследовании динамика складывается из двух указанных процессов; при этом, строго говоря, нет возможности эмпирически отделить стейеров от муверов, остающихся в одном классе на протяжении всего времени наблюдения (см. упражнение 8).

Рассмотрим модель «мувер-стейер» для простейшего случая двух классов.

Обозначим через s_1 и s_2 численность стейеров первого и второго классов соответственно через $m_1(t)$, а через $m_2(t)$ — численность муверов, находящихся в первом и втором классах в начальный момент времени $t = 0$. Матрица перехода для муверов имеет вид

$$L = \begin{pmatrix} \alpha_1 & 1 - \alpha_2 \\ 1 - \alpha_1 & \alpha_2 \end{pmatrix},$$

где $0 \leq \alpha_1 \leq 1$, $0 \leq \alpha_2 \leq 1$ — вероятности для муверов, находящихся в первом и втором классе соответственно, остаться в нем в следующий момент времени.

Наблюдаемыми при социологическом исследовании является общая численность $n_i(t)$ i -го класса в момент времени t при всех $i = 1, 2$, $t = 0, 1, \dots, T$ (здесь T — продолжительность временного интервала), а также численности $b_{ij}(t)$ перешедших из i -го класса в j -й между моментами t и $t + 1$ ($i = 1, 2$, $j = 1, 2$). Соотношения между ними удобно отображать в виде схемы (рис. 1).

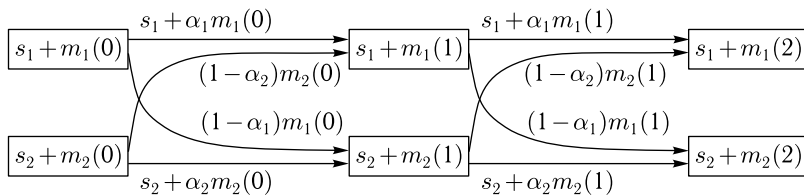


Рис. 1. Схема динамики численности муверов и стейеров

Задача заключается в том, чтобы определить в интересующий момент времени численность стейеров и муверов в каждом классе, а также матрицу перехода для муверов (при стандартном предположении, что эта матрица постоянна).

Имеют место следующие связи между введенными выше величинами (см. рис. 1):

$$s_1 + m_1(t) = n_1(t), \quad (1)$$

$$s_2 + m_2(t) = n_2(t), \quad (2)$$

$$s_1 + \alpha_1 m_1(t) = b_{11}(t), \quad (3)$$

$$s_2 + \alpha_2 m_2(t) = b_{22}(t), \quad (4)$$

$$(1 - \alpha_1)m_1(t) = b_{12}(t), \quad (5)$$

$$(1 - \alpha_2)m_2(t) = b_{21}(t), \quad (6)$$

причем при всех t для правых частей выполняются балансные условия

$$n_1(t) = b_{11}(t) + b_{12}(t), \quad n_2(t) = b_{22}(t) + b_{21}(t), \quad (7)$$

$$n_1(t+1) = b_{11}(t) + b_{21}(t), \quad n_2(t+1) = b_{22}(t) + b_{12}(t). \quad (8)$$

Заметим, что (5), (6) могут быть исключены как следствия предыдущих: например, равенство (5) следует из (1) и (3) с учетом (7).

Для определения шести неизвестных s_1 , s_2 , $m_1(0)$, $m_2(0)$, α_1 , α_2 имеем систему шести уравнений

$$s_1 + m_1(0) = n_1(0), \quad (9)$$

$$s_2 + m_2(0) = n_2(0), \quad (10)$$

$$s_1 + \alpha_1 m_1(0) = b_{11}(0), \quad (11)$$

$$s_2 + \alpha_2 m_2(0) = b_{22}(0), \quad (12)$$

$$s_1 + \alpha_1(a_1(1) - s_1) = b_{11}(1), \quad (13)$$

$$s_2 + \alpha_2(a_2(1) - s_2) = b_{22}(1), \quad (14)$$

связывающих эти величины с измеренными при социологическом исследовании значениями $a_1(0)$, $b_{11}(0)$, ..., $a_1(1)$, $b_{11}(1)$. Система (9)–(14) распадается на две системы: уравнения (9), (11), (13) содержат только

$s_1, m_1(0), \alpha_1$, а уравнения (10), (12), (14) — только $s_2, m_2(0), \alpha_2$.
Имеем, в частности (вторая тройка неизвестных находится аналогично)

$$\alpha_1 = \frac{b_{11}(1) - b_{11}(0)}{n_1(1) - n_1(0)}, \quad s_1 = \frac{b_{11}(0)n_1(1) - b_{11}(1)n_1(0)}{[n_1(1) - n_1(0)] - [b_{11}(1) - b_{11}(0)]}, \quad m_1(0) = n_1(0) - s_1.$$

После этого нетрудно рекуррентно вычислить значения $m_1(t), m_2(t)$ для всех последующих моментов t и определить всю динамику искомых величин.

Если вероятности α_1, α_2 не равны нулю или единице (что на практике бывает лишь в исключительных случаях), то с течением времени распределение муверов по классам стабилизируется к стационарному состоянию, которое может быть найдено стандартным способом. В случае, если система содержит более двух классов, для вычисления всех переменных модели типа мувер-стейер требуются данные панельного наблюдения (см. комментарий 1 в конце главы) для более чем трех моментов времени.

Модели типа мувер-стейер — это пример совокупности процессов, в которой чисто марковский процесс сочетается с каким-либо другим. В данном случае указанный другой процесс заключается в тождественных (см. первый пункт) переходах стейеров, т. е. их приверженности своему классу.

4.2.5. Расовая динамика

Рассмотренные выше марковские процессы, как и расширяющие их модели типа «мувер-стейер», предполагают, что в каждый момент времени система характеризуется вектором, размерность которого равна количеству классов в социальной системе; за один шаг времени происходит переход от одного вектора к другому вектору той же размерности.

В настоящем разделе рассматривается модель марковского типа, предполагающая, однако, что в течение одного временного цикла система поочередно описывается векторами, имеющими различную размерность.

Конкретный изучаемый социальный процесс состоит в следующем (рис. 2).

Рассматривается население, содержащее три расы: белую, черную и смешанную (мулаты). Отметим, что в данном случае речь идет не о собственно биологической расе, а о «гражданском состоянии», т. е. той расе, которая указана в документах индивида и не всегда совпадает с его биологической расой (принадлежность к той или иной расе зачастую увязывается с социальным положением и престижем). Миграция считается пренебрежимо малой. Единица времени увязывается с демографическим циклом, так что номер поколения служит номером момента модельного времени.

Анализ цикла начинается с рассмотрения поколения детей, часть которых принадлежит к белой расе, другая — к смешанной,



Рис. 2

третья — к черной (см. верхний прямоугольник на рис. 2). Социальная система в этой фазе описывается трехмерным вектором, при этом значение каждой компоненты равно численности соответствующей расы в данном поколении. Дети подрастают и образуют пары (этап 1), которые классифицируются по 9 группам. На этой фазе система описывается девятимерным вектором.

Став членом новообразованной семьи, индивид может получить дополнительные социальные возможности, ввиду чего становится возможным расовое «переключение» (этап 2), т. е. незаконная формальная (не биологическая) смена расы в плане официального учета населения. Например женщина, рожденная мулаткой и ставшая женой белого, может получить возможность заявлять себя как белую. На следующей фазе цикла семейные пары производят детей (причем каждый из 9 типов пар характеризуется, вообще говоря, своей фертильностью, т. е. средним количеством детей на семью). Каждому из новорожденных приписывается определенная раса (например, ребенок белого и мулатки должен быть зарегистрирован как мулат, но в нарушение установленного порядка может быть зарегистрирован как белый). Затем начинается первый этап следующего цикла модели.

Рассмотрим описанный процесс более подробно. При этом порядковый номер поколения будем отождествлять с моментом времени. Будем полагать, что в каждой группе мальчиков и девочек рождается поровну.

Этап 1. Обозначим через $x_1(t)$, $x_2(t)$, $x_3(t)$ численности белого, мулатского и черного населения соответственно в поколении t . Введем трехмерный вектор $X(t)$, компонентами которого являются указанные

численности, а также девятимерный вектор $Y(t)$, компонентами которого являются численности пар в поколении t . Соответственно имеем девять типов семей:

- белая женщина — белый мужчина ($y_1(t)$),
- белая женщина — мулат мужчина ($y_2(t)$),
- белая женщина — черный мужчина ($y_3(t)$),
- мулатская женщина — белый мужчина ($y_4(t)$),
- мулатская женщина — мулат мужчина ($y_5(t)$),
- мулатская женщина — черный мужчина ($y_6(t)$),
- черная женщина — белый мужчина ($y_7(t)$),
- черная женщина — мулат мужчина ($y_8(t)$),
- черная женщина — черный мужчина ($y_9(t)$).

Связь между $Y(t)$ и $X(t)$ дается равенством $Y(t) = MX(t)$, где матрица M имеет вид

$$M = \begin{pmatrix} m_{11} & 0 & 0 \\ m_{21} & 0 & 0 \\ m_{31} & 0 & 0 \\ 0 & m_{42} & 0 \\ 0 & m_{52} & 0 \\ 0 & m_{62} & 0 \\ 0 & 0 & m_{73} \\ 0 & 0 & m_{83} \\ 0 & 0 & m_{93} \end{pmatrix}.$$

Здесь m_{11} , m_{21} , m_{31} — вероятности для белой женщины выйти замуж за белого, мулатского и черного мужчину соответственно. Например, численность семей первой группы равна доле белых женщин, вышедших замуж за белого мужчину: $y_1 = m_{11}x_1$. Аналогичный смысл имеют и другие ненулевые элементы матрицы M . При этом сумма элементов каждого столбца матрицы равна единице.

Заман 2. В соответствии с описанной выше схемой модели (рис. 2), рассмотрим изменение расового состава семей. Обозначим через s_{ij} ($1 \leq i, j \leq 9$) вероятность для семьи из j -й группы перейти между браком и рождением ребенка в i -ю группу. Тогда распределение семей поколения t к моменту рождения ребенка будет описываться девятимерным вектором $Z(t)$ таким, что $Z(t) = SY(t)$. Здесь S — стохастическая матрица девятого порядка, составленная из элементов s_{ij} , она аналогична матрице перехода в модели вертикальной социальной мобильности.

Заман 3. Для описания дифференцированной по типам родительских пар фертильности вводится диагональная матрица $R = \text{diag}\{r_1, r_2, \dots, r_9\}$, где r_i — среднее количество детей, рождаемый семьей

i -й группы. Тогда распределение детей (т. е. членов $t + 1$ -го поколения) по эти группам описывается девятимерным вектором $B(t + 1) = RZ(t)$.

Заман 4. После этого проводится классификация новорожденных по расам, описываемая матрицей классификации C , элемент c_{ij} которой имеет смысл вероятности для ребенка, рожденного от родительской пары из j -й группы, быть зарегистрированным принадлежащим к i -й расе. Матрица C состоит из трех строк и девяти столбцов, и по данным пуэрториканской переписи 1920 г. имеет вид

$$C = \begin{pmatrix} 0,997 & 0,259 & 0,092 & 0,259 & 0,013 & 0,010 & 0,092 & 0,10 & 0 \\ 0,003 & 0,741 & 0,827 & 0,741 & 0,985 & 0,758 & 0,827 & 0,758 & 0,013 \\ 0 & 0 & 0,081 & 0 & 0,002 & 0,232 & 0,081 & 0,232 & 0,987 \end{pmatrix}.$$

Заметим, что в соответствии с существовавшими в то время правовыми нормами белым должен был считаться ребенок белых родителей, черным — ребенок черных родителей, во все остальных случаях ребенок должен был считаться мулатом. Таким образом, «идеальная» матрица расовой классификации имеет вид

$$C_{\text{ideal}} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix},$$

т. е. некоторые элементы матрицы C отличаются от «законных» значений на 25%.

Итак, с учетом всех стадий цикла основное уравнение модели имеет итоговый вид

$$X(t + 1) = CRSMX(t).$$

Формально модель выглядит как марковский процесс с матрицей перехода $P = CRSM$. Однако на самом деле элементы матрицы M зависят от текущего распределения $X(t)$, поскольку вероятность выйти замуж, например за белого мужчину, зависит от доли белых в мужском населении: $M = M(X(t))$. Модель является линейной ($X(t + 1) = PX(t)$), где P — постоянная матрица) в случае, если принять, что пары могут быть образованы лишь мужчиной и женщиной одной расы, т. е. $m_{11} = m_{52} = m_{93} = 1$, все остальные $m_{ij} = 0$. При этом, если все коэффициенты фертильности равны двум, $r_1 = \dots = r_9 = 2$, то численность населения является постоянной, а процесс — марковским. В противном случае динамика социальной системы складывается из изменений расового состава, описываемого чисто марковским процессом, и демографического роста либо спада. В более общем случае, т. е. при дифференцированной фертильности, будет изменяться как расовый состав, так и общая численность населения.

Подобного рода исследования и результаты могут быть весьма важными для анализа и регулирования социально-демографической ситуации в обществе.

Упражнения

1. Не используя теорему Фробениуса–Перрона, доказать, что стохастическая матрица имеет собственное значение, равное единице.

2. Доказать, что любая стохастическая матрица в натуральной степени также является стохастической матрицей.

3. Рассмотреть модель динамики численности верующих (первый пункт). Для индивида, находящегося при $t = 1$ в i -й группе, найти вероятность q_{ij} того, что при $t = 0$ он находился в j -й группе.

4. Предложить пример, при котором для некоторой группы имеет место неравенство $x_i(0) < x_i(1)$, однако значение i -й компоненты вектора X_s меньше, чем $x_i(0)$ (т. е. при общей тенденции к уменьшению численности группы на первом временном шаге наблюдалось ее увеличение).

5. Доказать, что если имеется поглощающая группа j , то не только у матрицы P все, кроме одного, элементы j -го столбца равны нулю, но и любой натуральной степени этой матрицы. Каков социологический смысл этого утверждения?

6. Для модели вертикальной социальной мобильности (второй пункт) на основе данных табл. 2 и матрицы P_7 получить матрицу P_3 .

7. Для трехклассовой модели (второй пункт) определить, сколько в среднем поколений проводит семья в высшем классе? То же сделать для среднего и низшего классов.

8. В рамках модели типа мувер-стейер определить вероятность того, что индивид является стейером, если он находился в первом классе при $t = 0$, $t = 1$ и $t = 2$.

Комментарии и библиографические примечания

1. Строго говоря, марковский процесс — это семейство случайных величин, индексированных некоторым параметром (как правило, не обязательно имеющим смысл времени). Другими словами, марковский процесс — это математический объект. Однако в приложениях этот термин применяют и самим моделируемым процессам — социальным, физическим и пр. В данной книге мы также придерживаемся этого широкого понимания термина «марковский процесс».

Марковским процессам противопоставляются немарковские процессы, в которых «будущее» (момент времени $t + 1$) зависит не только от «настоящего» (момент t), но и «прошлого» (момент $t - 1$). Типичным немарковским процессом является колебание курса акций. Процессы социальной динамики являются марковскими лишь в некотором приближении. Например, изложенная в третьем пункте модель социальной мобильности предполагает, что для рожденного, например в среднем классе ребенка, вероятность попасть в высший класс не зависит от того, к какому классу принадлежит его дед. Чем лучше выполняется в некотором социуме это предположение, тем больше оснований

считать процесс приближенно марковским, тем более адекватной является модель.

Марковские процессы с дискретным временем и дискретными состояниями называются марковскими цепями. Таким образом, рассматриваемые здесь модели являются марковскими цепями, однако мы не употребляем этот термин, поскольку он не прижился в социальных науках.

2. Заметим, что одним из существенных аспектов проведения конкретных социологических исследований в данном случае является обработка и анализ данных: статистическое обоснование утверждений о количестве групп (классов) и построение матриц перехода зачастую оказывается нетривиальной задачей. В настоящей книге вопросы статистического анализа не рассматриваются, при обсуждении соответствующих моделей даются библиографические ссылки. Поскольку здесь не рассматриваются вопросы статистической обработки данных, мы, несколько пренебрегая строгостью, не проводим различий между наблюдаемыми в результате социологического исследования частотами, и вероятностями, приближениями к которым являются эти частоты. Другими словами, в данном разделе, все опросы предполагаются «идеально репрезентативным» в том смысле, что частоты (доли по рассматриваемой выборке) равны соответствующим значениям вероятностей (долям по населению в целом).

3. Социологические данные, необходимые для оснащения моделей марковского типа реальными данными, собираются (чаще всего) путем проведения так называемых *панельных исследований*, предполагающих изучение одной и той же выборки в различные моменты времени (другим вариантом получения информации, необходимой для оснащения конкретными данными моделей марковского типа, является опрос, при котором респондента опрашивают о его принадлежности к той или иной группе в настоящий момент и в некоторый момент в прошлом).

Заметим, что, в случае панельного наблюдения исследователь обладает более полной информацией, а именно так называемой «историей» каждого члена выборки. «История» содержит данные о его принадлежности к тому или иному классу во все рассматриваемые дискретные моменты времени и, таким образом, позволяет связывать распределение по классам в момент t с распределением не только в момент $t - 1$, но и в другие предшествующие моменты времени. Таким образом, при построении математической модели, основанной на марковских процессах, происходит сознательный отказ от части информации, собранной при панельном исследовании.

4. Модель динамики количества верующих была предложена в работе [1]. При проведении социологического опроса помимо общих вопросов об отношении к религии (был ли и стал ли верующим и т. д.) респондентам задавались также вопросы о конфессиональной принадлежности. Построенные по ответам вероятности перехода (умноженные на 1000) представлены в табл. 3.

Таблица 3. Вероятности перехода, умноженные на 1000. Используются сокращения: вбв — верующие, не относящие себя ни к какому вероисповеданию, др. веры — придерживающиеся другой веры

		Был									
		Правосл.	Старобр.	Католик	Мусульм.	Буддист	Иудаист	Протест	Др. веры	Вбв	Невер.
Стал	Правосл.	962	88	143	16	0	0	0	94	172	146
	Старобр.	1	877	0	0	0	0	0	31	0	2
	Католик	1	0	810	3	0	0	0	0	6	1
	Му- сульм.	3	0	0	959	0	0	0	0	19	4
	Буддист	0	0	0	10	1000	0	0	0	3	2
	Иудаист	1	0	0	3	0	1000	0	0	3	0
	Протест	2	0	0	6	0	0	1000	0	0	0
	Др. веры	3	0	0	0	0	0	0	875	17	7
	Вбв	19	0	0	3	0	0	0	0	714	58
	Невер.	8	35	47	0	0	0	0	0	67	781

5. Современное и весьма полное изложение теории марковских цепей, в частности доказательство теоремы 2, можно найти, напр. в [2, 3].

6. Модель вертикальной социальной мобильности была построена на основе крупномасштабного социологического исследования «Социальная мобильность в Британии», проведенного в конце 40-х годов XX века, и предложена в работах [4, 5]. О современной теории социальной стратификации см., напр., обзоры в [6, 7]; одним из основных ее направлений является так называемый *ресурсный подход*, основанный на предложенной французским социологом П. Бурдьё [8] теории различных видов капитала: экономический капитал, социальный, культурный и пр. Заметим, что на каких бы социологических основаниях ни строилась стратификация (уровень потребления индивида, объем располагаемых им ресурсов и пр.), вертикальная мобильность индивидов может быть описана моделями того же типа, что и рассмотренная в данном разделе.

7. Модель социальной мобильности допускает ряд обобщений. В частности, в нее может быть введено предположение о дифференцированной фертильности: в разных классах на одного отца в среднем приходится различное количество детей. В этом случае матрица

перехода не является стохастической, и стационарное состояние, вообще говоря, отсутствует.

8. Модель типа мувер-стейер (mover-stayer) была предложена в [9]). В качестве примера ее приложения укажем работу [10], посвященную исследованию безработицы и рынка труда во Франции. Содержательная сторона в данном случае состоит в том, что некоторые индивиды могут переходить из категории безработных в занятые и/или обратно; тем самым они являются муверами. В то же время некоторые другие имеют твердую работу либо являются безнадежными безработными; они составляют множество стейеров. Таким образом, деление всех индивидов на муверов и стейеров продиктовано содержанием рассматриваемого явления.

9. Модель расовой динамики предложена и изучена в [11].

Литература

1. *Ефремов В.А., Леонтьев В.А.* Использование аппарата Марковских процессов в социологических исследованиях межконфессиональных отношений в России // Математическое моделирование социальных процессов. Вып. 8. М.: 2006, с. 15–29.
2. *Seneta E.* Non-negative Matrices and Markov Chains. Springer, 2006.
3. *Kemeny J., Snell J.* Finite Markov Chains. New York: Springer-Verlag, 1976.
4. *Prais S.J.* Measuring social mobility. Journal of the Royal Statistical Society. Series A (General), 118(1), p. 56–66. 1955.
5. *Prais S.J.* The formal theory of social mobility. Population Studies, 9(1), p. 72–81. 1955.
6. *Тихонова Н.Е.* Модель социальной стратификации российского общества: эвристические возможности различных теоретических подходов // Россия реформирующаяся. Ежегодник / Отв. Ред. М.К. Горшков. Вып. 6. М.: Институт социологии РАН, 2007. С. 112–146.
7. *Тихонова Н.Е.* Низший класс в социальной структуре российского общества с. 24 // Социологические исследования. 2011. №5.
8. *Бурдые П.* Формы капитала // Экономическая социология. Электронный журнал. Т.3 №5. ноябрь 2002. С. 60–74
9. *Blumen I., Kogan M., McCarthy P.J.* The Industrial Mobility of Labor as a Probability Process. Cornell University Press. Ithaca. New York, 1955.).
10. *Beffy M., Coudin E., Rathelot R.* Who is confronted to insecure labor market histories? Some evidence based on the French labor market transitions. Institut National de la Statistique et des Études Économiques. Série des documents de travail de la Direction des Études et Synthèses Économiques. G2008/10.
11. *Montgomery J.D.* The Population Dynamics of Black-White-Mulatto Racial Systems // American Journal of Sociology V. 117, №1 (July 2011), p. 46–89.

4.3. Социальная сегрегация

Настоящий раздел посвящен модели, описывающей социальную сегрегацию, под которой понимается процесс разделения различных групп населения по месту жительства в соответствии с расовой, этнической, религиозной принадлежностью. Далее будем называть эти группы общинами.

Более конкретно, рассматривается сегрегация, происходящая вследствие добровольных действий индивидов: каждый из них выбирает место жительства, стараясь оказаться по возможности «среди своих».

Таким образом, из рассмотрения исключаются (по крайней мере) два вида сегрегации. Во-первых, не рассматривается принудительная сегрегация, при которой разделение возникает в результате официального или неофициального запрета для одной из общин на проживание в некоторой местности. Во-вторых, здесь не рассматривается сегрегация, основанная на экономических механизмах. Имеются в виду процессы, при которых жители меняют место жительства из экономических соображений. Если, например, в местности А жизнь стала более дорогой, чем в местности Б, то наиболее бедные жители местности А будут переезжать в Б, а наиболее богатые жители Б будут переезжать в более «роскошную» местность А. Таким образом, экономический механизм приведет к раздельному проживанию богатых и бедных. При этом экономически вызванная сегрегация может коррелировать с расовой. Например, если белое население является более обеспеченным, чем черное, то тем самым возникнет раздельное проживание белых и черных, но эта сегрегация будет лишь внешним проявлением сегрегации экономической. Это явление в настоящей работе не рассматривается.

Собственно социальная сегрегация основана на следующем механизме. Процесс начинается с того, что в результате миграции возникает ситуация совместного проживания двух общин в пределах одного квартала. Если в отношениях между общинами присутствуют конфликты, то некоторые представители одной из них выезжают из квартала. На освободившиеся места въезжают новые представители другой общины. Тем самым пропорция изменяется в пользу второй общины, в результате чего еще некоторые представители первой общины покидают квартал, освобождая места для второй общины. Таким образом, с течением времени доля второй общины в населении постоянно возрастает. Этот социальный механизм описывается математической моделью, имеющей вид системы двух обыкновенных дифференциальных уравнений.

Первый пункт настоящего раздела посвящен формулировке поведенческой гипотезы, а также разработке и анализу качественной модели, описывающей направление динамики численностей членов общин. Во втором пункте проведена конкретизация поведенческой гипотезы, что позволило построить математическую модель, имеющую вид

системы двух обыкновенных дифференциальных уравнений; третий пункт посвящен анализу этой модели.

4.3.1. Качественная модель

Рассмотрим население, состоящее из двух общин, которые будем обозначать X и Y . Численность членов общин, проживающих в данном квартале в момент времени t , обозначим через $x(t)$ и $y(t)$.

Модель базируется на следующей поведенческой гипотезе. Предполагается, что индивид удовлетворен соседством в данном квартале, лишь если отношение численности «чужой» общины к численности его собственной общины не превышает некоторой величины, называемой порогом толерантности этого индивида. Значение порога толерантности не меняется с течением времени и варьируется между индивидами: например, для кого-либо может быть достаточно не находиться в меньшинстве (т. е. порог равен единице), другой индивид может быть удовлетворен, если его община составляет не менее 30% населения квартала (в этом случае порог равен $7/3$). Если текущее соотношение численностей ниже порога толерантности, то индивид удовлетворен соседством в квартале и он остается в нем (или въезжает в квартал, если не жил в нем ранее); если же порог превышен, то индивид покидает квартал (или не въезжает в него, если не живет в нем). При этом для индивида, покинувшего квартал, сохраняется возможность с течением времени вернуться в него (и наоборот, новоприбывший житель может его впоследствии покинуть). Индивиды обладают полной информацией о количестве жителей квартала, принадлежащих к каждой из общин, но они не согласовывают свои действия и не строят предположений о поведении других индивидов.

Отметим: что в отличие ряда других моделей, рассматриваемых в настоящей книге, в данной модели отсутствует понятие типической личности.

Сначала рассмотрим случай, когда численность кандидатов на проживание в квартале (т. е. тех, кто либо живет в нем, либо может въехать) фиксирована, причем емкость квартала позволяет проживать в нем всем этим кандидатам.

В соответствии со сформулированной поведенческой гипотезой каждая община может быть охарактеризована распределением порогов толерантности среди ее членов. Для того чтобы формализовать данную характеристику общины X , пронумеруем ее членов в порядке невозрастания порога толерантности и каждому из этих членов (вернее, соответствующему номеру) поставим в соответствие его значение порога толерантности; получим гистограмму, изображенную на рис. 1 (слева). При достаточно большой численности N проведем предельный переход, получив непрерывную, невозрастающую функцию $f(n)$, которую назовем распределением толерантности в данной общине.

Пример 1. На рис. 2 представлен график функции распределения толерантности общины, состоящей из $N = 100$ членов. Все они удо-

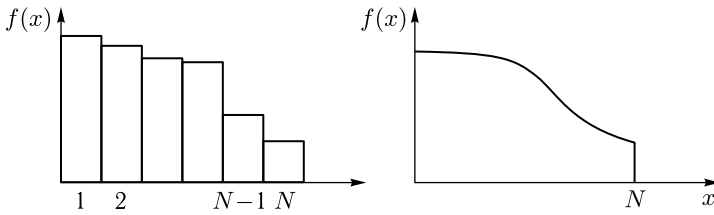


Рис. 1. Распределение толерантности в общине

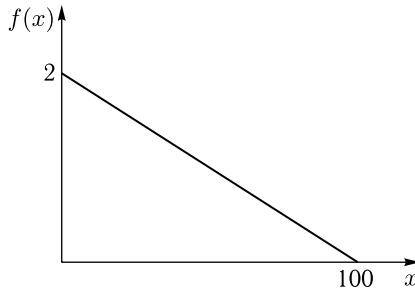


Рис. 2. Пример функции распределения толерантности

влетворены кварталом, когда отношение численности «чужой» общины к своей равно нулю (т.е. «чужая» община отсутствует). В то же время самый толерантный член удовлетворен кварталом, когда «чужая» община в два раза (или меньше) превышает его собственную. Между этими крайними точками толерантность плюсов распределена линейно: $f(x) = 2 - 0,02x$.

Функция $y = xf(x)$ имеет смысл максимальной численности общины Y , при которой x наиболее толерантных членов общины X удовлетворены соседством в квартале. Таким образом, если $y < xf(x)$, то численность общины X возрастает, а если $y > xf(x)$, то убывает. Аналогично, если распределение толерантности в общине Y дается функцией $g(y)$, то $x = yg(y)$ — это максимальная численность общины X , при которой y наиболее толерантных членов общины Y удовлетворены соседством в квартале. Если $x < yg(y)$, то численность общины Y возрастает, а если $x > yg(y)$, то убывает. Изобразим функции $y = xf(x)$ и $x = yg(y)$ на одном графике. Один из возможных вариантов получаемой при этом конфигурации представлен на рис. 3.

Состоянию системы, при котором в квартале проживают x представителей первой общины и y представителей второй, соответствует точка (x, y) прямоугольника $0 \leq x \leq N_x$, $0 \leq y \leq N_y$, где N_x , N_y — максимально возможные численности общин. Кривые $y = xf(x)$ и $x = yg(y)$ разбивают данный прямоугольник на четыре области, каждая из которых характеризуется определенным направлением динамики (возрастанием или убыванием) численности членов общин.

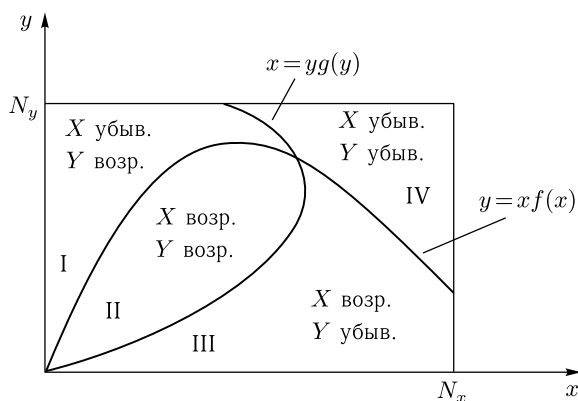


Рис. 3. Области значений переменных, в которых динамика численности членов общин X, Y, проживающих в квартале, имеет постоянное направление (конфигурация с четырьмя областями)

Так, если точка (x_0, y_0) , изображающая начальное количество живущих в квартале членов общин X, Y, находится в области III, то с течением времени количество представителей общины X будет возрастать, а Y — убывать. Аналогично, если точка (x_0, y_0) расположена в области IV, то будет убывать численность жителей обеих общин.

Чтобы качественно описать динамику системы, составим ее фазовый портрет (рис. 4), аналогично тому, как это делается для систем обыкновенных дифференциальных уравнений. В данном случае фазовый портрет удастся построить, несмотря на отсутствие уравнений, благодаря тому, что фазовые кривые горизонтальны при пересечении кривой $x = yg(y)$ и вертикальны при пересечении кривой $y = xf(x)$.

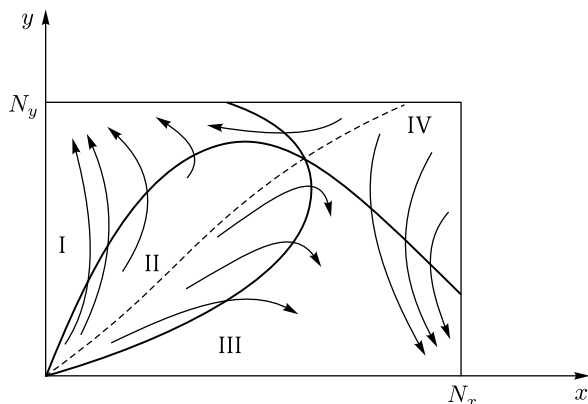


Рис. 4. Фазовый портрет, соответствующий конфигурации, представленной на рис. 3. Сепаратриса изображена пунктирной линией

В области II фазовые кривые направлены вверх и вправо (так как численность жителей квартала, принадлежащих к каждой общине, возрастает). Если начальная точка (x_0, y_0) расположена в этой области, то с течением времени система перейдет из области II в область I либо в область III в зависимости от того, по какую сторону от сепаратрисы расположено начальное условие и от того, какая община более быстро заполняет квартал. Если, например, система перейдет в область I, то начиная с некоторого момента времени община X начнет убывать и с течением времени сформируется устойчивое стационарное состояние $x = 0, y = N_y$, при котором все население квартала принадлежит к общине Y. И наоборот, если система перейдет в область III, то сформируется устойчивое стационарное состояние $x = N_x, y = 0$.

Аналогично, если начальная точка (x_0, y_0) расположена в области IV, то с течением времени система перейдет либо в область I, либо в область III с последующим формированием одного из указанных состояний. Если же система находится в области I либо III уже в начальный момент времени, то она остается в ней в любой момент времени.

Таким образом, в данном случае система имеет лишь два устойчивых стационарных состояния, $x = 0, y = N_y$ и $x = N_x, y = 0$, т. е. одна из общин полностью вытесняет другую из квартала.

Представленное на рис. 3 и рис. 4 взаимное расположение графиков функций $y = xf(x)$ и $x = yg(y)$ не является единственно возможным. Один из других вариантов изображен на рис. 5.

Отличие от рассмотренного выше случая заключается в том, что кривые $y = xf(x)$ и $x = yg(y)$ на рис. 5 имеют четыре точки пересечения (считая начало координат), а не две, как на рис. 3 и на рис. 4.

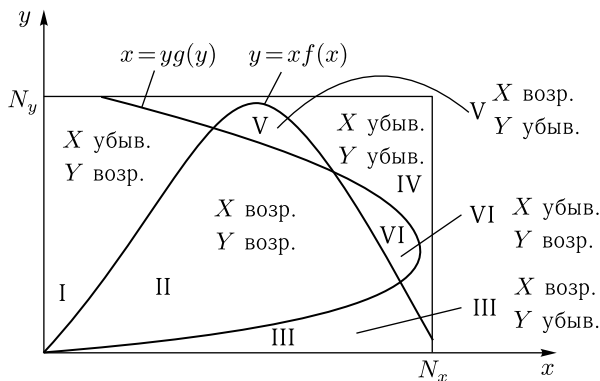


Рис. 5. Области значений переменных, в которых динамика численности членов общин X, Y, проживающих в квартале, имеет постоянное направление (конфигурация с трехкратным пересечением кривых $y = xf(x)$ и $x = yg(y)$ в области $0 < x < N_x, 0 < y < N_y$)

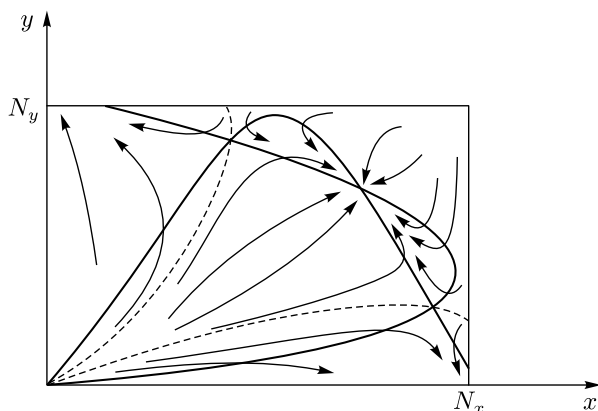


Рис. 6. Фазовый портрет, соответствующий конфигурации, представленной на рис. 5. Сепаратрисы изображены пунктирными линиями

Соответствующий фазовый портрет представлен на рис. 6. Как и в случае рис. 4, фазовые кривые горизонтальны при пересечении кривой $x = yg(y)$ и вертикальны при пересечении кривой $y = xf(x)$.

Как показывает рис. 6, система имеет в данном случае три асимптотически устойчивых стационарных состояния: помимо аттракторов $(0; N_y)$, $(N_x; 0)$ (соответствующих ситуации, когда в квартале с течением времени остается лишь одна из общин) присутствует также аттрактор, при котором общины сосуществуют. К какому из них произойдет стабилизация — зависит от начальных условий.

Помимо конфигураций, представленных на рис. 3–6, могут иметь место и другие. В частности возможен случай, когда в окрестности начала координат кривая $x = yg(y)$ расположена выше, чем кривая $y = xf(x)$.

В проведенном выше анализе предполагалось, что емкость квартала N_0 достаточно велика, чтобы вместить всех кандидатов на проживание в нем, т. е. $N_0 \geq N_x + N_y$. Это предположение является реалистичным лишь в некоторых, весьма особенных случаях: например, когда как X , так и Y представляют собой малочисленные общины мигрантов в чужой стране. Значительно более распространенной является ситуация, когда емкость квартала N_0 существенно меньше численностей общин N_x , N_y , например, коренное население европейского города и крупная мигрантская община. В этом случае областью изменения переменных x , y является не прямоугольник $0 \leq x \leq N_x$, $0 \leq y \leq N_y$, а треугольник $x \geq 0$, $y \geq 0$, $x + y \leq N_0$, где N_0 — емкость квартала (см. рис. 7, а). Возможны также случаи, когда N_0 меньше численности лишь одной из общин (см. рис. 7, б) либо больше каждой из них, но меньше их суммы: $\max(N_x, N_y) < N_0 < N_x + N_y$ (рис. 7, в).

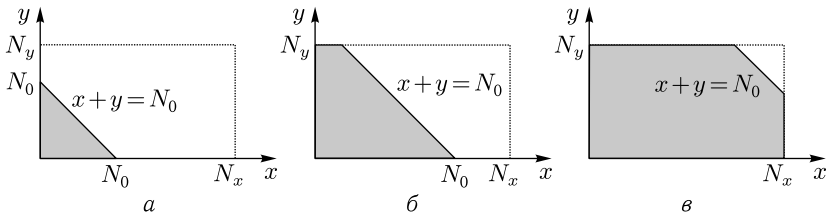


Рис. 7. Возможные варианты области изменения переменных x, y (помимо случая $N_0 \geq N_x + N_y$)

Рассмотрим случай

$$N_0 \ll \min(N_x, N_y), \quad (1)$$

которому соответствует рис. 7, а. В этой ситуации возможны две конфигурации, отличающиеся тем, какая из кривых $y = xf(x)$, $x = yg(y)$ расположена выше (случай пересечения кривых исключается неравенством (1), так как каковы бы ни были координаты точки пересечения, всегда можно выбрать N_0 настолько малым, чтобы она осталась за пределами треугольника $x \geq 0, y \geq 0, x + y \leq N_0$). Для каждой из этих конфигураций имеем три области, отличающиеся друг от друга направлением динамики (возрастанием или убыванием) численности членов общин (см. рис. 8).

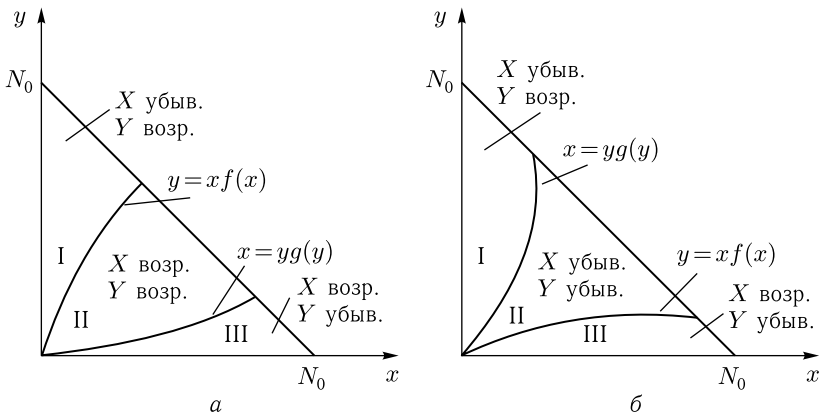


Рис. 8. Области значений переменных, в которых динамика численности членов общин X, Y имеет постоянное направление (при выполнении условия (1))

Соответствующие фазовые портреты представлены на рис. 9.

Если кривая $y = xf(x)$ расположена выше кривой $x = yg(y)$ (рис. 8, а, рис. 9, а), то помимо аттракторов $x = 0, y = N_0$, $x = N_0, y = 0$ (соответствующих ситуации, когда в квартале с течением времени остается лишь одна из общин) имеется также континуум

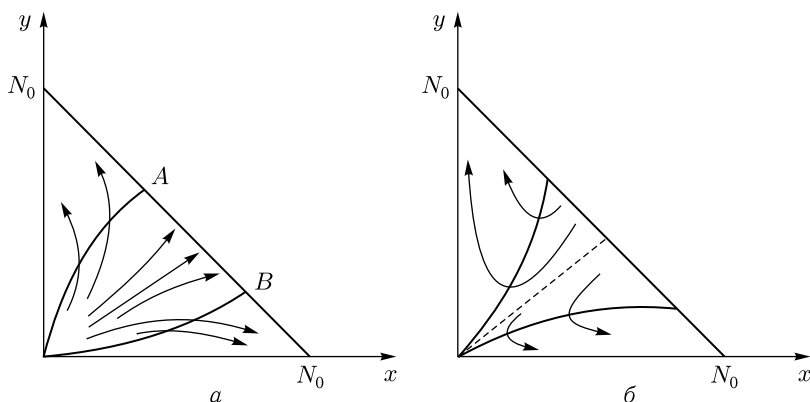


Рис. 9. Фазовые портреты, соответствующие конфигурациям, представленным на рис. 8

устойчивых состояний, соответствующих внутренним точкам отрезка AB на рис. 9, a . Таким образом, в зависимости от начальных условий в данном случае могут сформироваться стационарные состояния с различным соотношением численностей общин.

Если же $y = xf(x)$ расположена ниже кривой $x = yg(y)$ (рис. 8, b , рис. 9, b), то может сформироваться лишь одно из двух асимптотически устойчивых стационарных решений $x = 0, y = N_0, x = N_0, y = 0$, т.е. в квартале с течением времени остается лишь одна из общин. Различие по отношению со случаями, изображенными на рис. 3–рис. 6, заключается в том, что здесь (т.е. при $N_0 \ll \min(N_x, N_y)$) в стационарном состоянии квартал полностью заселен одной из общин, причем за его пределами остаются другие члены этой общины, которые тоже были бы готовы в него въехать при наличии свободных мест. В то же время в ситуациях, соответствующих рис. 3–рис. 6 (т.е. при $N_0 > \max(N_x, N_y)$), в квартале остаются свободные места.

4.3.2. Конкретизация поведенческой гипотезы и построение математической модели

Изложенный выше анализ опирался лишь на качественную зависимость между текущим соотношением численностей общин и направлением динамики, другими словами — на зависимость между величинами $x(t), y(t)$, с одной стороны, и знаками величин $dx/dt, dy/dt$ с другой. Для того чтобы получить дифференциальные уравнения, определяющие значения самих производных $dx/dt, dy/dt$ (а не только их знаков), необходимо конкретизировать сформулированную выше поведенческую гипотезу.

Как указано выше, в соответствии с этой гипотезой, если $x < yg(y)$, то численность общины Y возрастает, а если $x > yg(y)$, то она убывает. Изобразим график функции $x = yg(y)$, обозначим через x_1 ее макси-

мальное значение, y_1 — значение аргумента, при котором достигается максимум, а также укажем стрелками направление изменения численности членов общины Y , рис. 10, а (данный рисунок соответствует случаю, когда функция $x = yg(y)$ имеет ровно один локальный максимум $x_1 < N_x$; все рассуждения легко обобщаются на более сложные случаи).

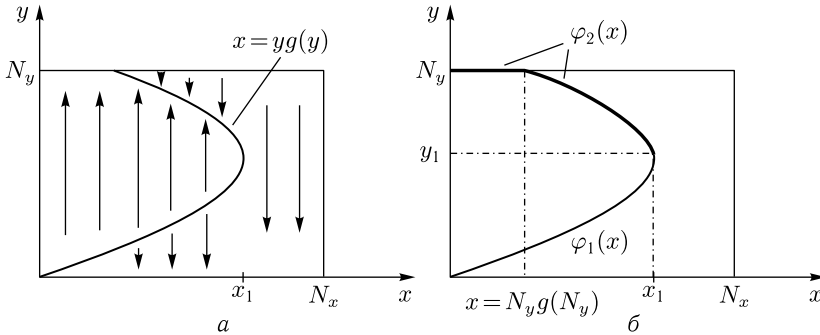


Рис. 10. Направление изменения численности членов общины Y и определение функций $y = \varphi_1(x)$, $y = \varphi_2(x)$

Обозначим через $y = \varphi_1(x)$ функцию, обратную по отношению к $x = yg(y)$ на отрезке $0 \leq y \leq y_1$, а через $y = \varphi_2(x)$ — функцию, обратную по отношению к $x = yg(y)$ на $y_1 \leq y \leq N_y$ и равную N_y при $0 \leq x \leq N_y g(N_y)$, см. рис. 10, б.

Для изменения численности общины Y имеем следующее.

При $x \leq x_1$: $\frac{dy}{dt} < 0$, если $y < \varphi_1(x)$,

$$\frac{dy}{dt} > 0, \quad \text{если } \varphi_1(x) < y < \varphi_2(x),$$

$$\frac{dy}{dt} < 0, \quad \text{если } y > \varphi_2(x).$$

При $x > x_1$: $\frac{dy}{dt} < 0$.

Простейшее дифференциальное уравнение, отражающее данную закономерность, имеет вид

$$\frac{dy}{dt} = \begin{cases} -\alpha y(y - \varphi_1(x))(y - \varphi_2(x)), & 0 \leq x \leq x_1, \\ -\alpha y, & x_1 \leq x \leq N_x. \end{cases} \quad (2)$$

Аналогично получим уравнение, описывающее динамику численности проживающих в квартале членов общины X .

Изобразим график функции $y = xf(x)$, обозначим через y_2 ее максимальное значение, x_2 — значение аргумента, при котором достигается максимум, рис. 11 (при этом ограничимся случаем, когда $y_2 < N_y$ и функция $y = xf(x)$ имеет ровно один локальный максимум). Обозначим через $x = \psi_1(y)$ функцию, обратную по отношению к $y = xf(x)$

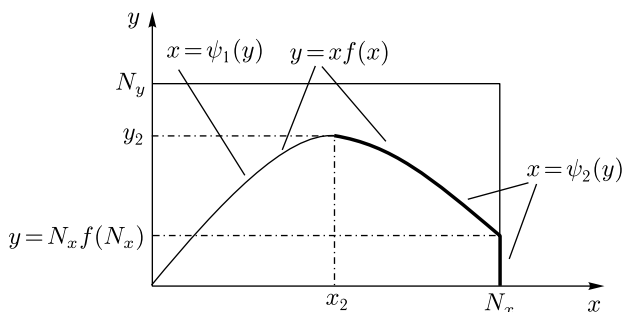


Рис. 11. Направление изменения численности членов общины X и определение функций $x = \psi_1(y)$, $x = \psi_2(y)$

на отрезке $0 \leq x \leq x_2$, а через $x = \psi_2(y)$ — функцию, обратную по отношению к $y = xf(x)$ на отрезке $x_2 \leq x \leq N_x$ и равную N_x при $0 \leq y \leq N_x f(N_x)$.

Простейшее дифференциальное уравнение, качественно отражающее динамику численности проживающих в квартале членов общины X, имеет вид

$$\frac{dx}{dt} = \begin{cases} -\beta x(x - \psi_1(y))(x - \psi_2(y)), & 0 \leq y \leq y_2, \\ -\beta x, & y_2 \leq y \leq N_y. \end{cases} \quad (3)$$

Уравнения (2), (3) являются конкретизацией поведенческой гипотезы, сформулированной в первом пункте. Они дополняются начальными условиями

$$x(0) = x^0, \quad y(0) = y^0, \quad (4)$$

где $0 < x^0 \leq N_x$, $0 < y^0 \leq N_y$.

4.3.3. Анализ математической модели

Проведенный выше анализ качественной модели показал, что стационарные решения $x = 0$, $y = N_y$, $x = N_x$, $y = 0$ являются асимптотически устойчивыми, в то время как стационарные состояния, при которых общины сосуществуют в квартале, могут быть как устойчивыми, так и неустойчивыми.

Изучим теперь вопрос об устойчивости стационарных решений модели (2)–(4).

Рассмотрим сначала решение

$$x = 0, \quad y = N_y. \quad (5)$$

В его окрестности уравнения (2), (3) принимают вид

$$\frac{dx}{dt} = -\beta x, \quad (6)$$

$$\frac{dy}{dt} = -\alpha y(y - \varphi_1(x))(y - N_y). \quad (7)$$

Очевидно, что при достаточно малых δ в области вида $0 < x < \delta$, $N_y - \delta < y < N_y$ имеем $dx/dt < 0$, $dy/dt > 0$. Отсюда следует, что если начальные условия (4) лежат в этой области, то $x(t) \rightarrow 0$, $N(t) \rightarrow N_y$ при $t \rightarrow \infty$, что согласуется с результатом анализа качественной модели.

Аналогично доказывается асимптотическая устойчивость решения

$$x = N_x, \quad y = 0. \quad (8)$$

Перейдем к исследованию стационарных решений, описывающих состояния, при которых общины сосуществуют в квартале. В окрестности каждого из них уравнения (3), (2) принимают вид

$$\frac{dx}{dt} = -\beta x(x - \psi_1(y))(x - \psi_2(y)), \quad (9)$$

$$\frac{dy}{dt} = -\alpha y(y - \varphi_1(x))(y - \varphi_2(x)). \quad (10)$$

Уравнения для рассматриваемых стационарных решений имеют вид

$$(x - \psi_1(y))(x - \psi_2(y)) = 0, \quad (11)$$

$$(y - \varphi_1(x))(y - \varphi_2(x)) = 0. \quad (12)$$

Количество стационарных решений зависит от вида функций распределения толерантности $y = f(x)$, $x = g(y)$. Каждое из них может быть исследовано на устойчивость, например с помощью метода исследования устойчивости по первому приближению.

Продемонстрируем это для случая линейных функций распределения толерантности:

$$f(x) = a - bx, \quad g(y) = c - ky. \quad (13)$$

Тогда максимальная численность общины Y, при которой x наиболее толерантных членов общины X удовлетворены соседством в квартале, дается функцией

$$y = xf(x) = ax - bx^2. \quad (14)$$

Максимальное значение функции (14) достигается при $x_2 = a/(2b)$, оно равно $y_2 = a^2/(4b)$. Чтобы найти обратные функции $x = \psi_1(y)$, $x = \psi_2(y)$, выразим x через y из равенства (14), получим

$$\psi_1(y) = \frac{a - \sqrt{a^2 - 4by}}{2b}, \quad \psi_2(y) = \frac{a + \sqrt{a^2 - 4by}}{2b}. \quad (15)$$

Аналогично, максимальная численность общины X, при которой y наиболее толерантных членов общины Y удовлетворены соседством в квартале, дается функцией

$$x = yg(y) = cy - ky^2. \quad (16)$$

Максимальное значение функции (16) достигается при $y_1 = c/(2d)$, оно равно $x_1 = c^2/(4d)$. Для обратных функций имеем

$$\varphi_1(x) = \frac{c - \sqrt{c^2 - 4kx}}{2k}, \quad \varphi_2(x) = \frac{c + \sqrt{c^2 - 4kx}}{2k}. \quad (17)$$

Система (9), (10) принимает в данном случае вид

$$\frac{dx}{dt} = -\beta x \left(x - \frac{a - \sqrt{a^2 - 4by}}{2b} \right) \left(x - \frac{a + \sqrt{a^2 - 4by}}{2b} \right), \quad (18)$$

$$\frac{dy}{dt} = -\alpha y \left(y - \frac{c - \sqrt{c^2 - 4kx}}{2k} \right) \left(y - \frac{c + \sqrt{c^2 - 4kx}}{2k} \right). \quad (19)$$

Уравнение (18) можно записать в виде

$$\frac{dx}{dt} = -\beta x \left(x - \frac{a}{2b} + \frac{\sqrt{a^2 - 4by}}{2b} \right) \left(x - \frac{a}{2b} - \frac{\sqrt{a^2 - 4by}}{2b} \right), \quad (20)$$

откуда

$$\frac{dx}{dt} = -\beta x \left(\left(x - \frac{a}{2b} \right)^2 - \frac{a^2 - 4by}{4b^2} \right). \quad (21)$$

Аналогично, из (19) имеем

$$\frac{dy}{dt} = -\alpha y \left(\left(y - \frac{c}{2k} \right)^2 - \frac{c^2 - 4kx}{4k^2} \right). \quad (22)$$

Рассмотрим вопрос об устойчивости стационарного решения (x_s, y_s) такого, что

$$x_s > 0, \quad y_s > 0, \quad (23)$$

описывающего ситуацию, при которой две общины сосуществуют в квартале. Приравнявая к нулю производные в (21), (22) и учитывая (23), получаем, что x_s, y_s удовлетворяют системе

$$\left(x_s - \frac{a}{2b} \right)^2 - \frac{a^2 - 4by_s}{4b^2} = 0, \quad (24)$$

$$\left(y_s - \frac{c}{2k} \right)^2 - \frac{c^2 - 4kx_s}{4k^2} = 0. \quad (25)$$

Матрица Якоби, взятая в стационарных состояниях, удовлетворяющих уравнениям (24), (25), имеет вид

$$J(x_s, y_s) = \begin{pmatrix} -2\beta x_s \left(x_s - \frac{a}{2b} \right) & -\frac{\beta x_s}{b} \\ -\frac{\alpha y_s}{k} & -2\alpha y_s \left(y_s - \frac{c}{2k} \right) \end{pmatrix}. \quad (26)$$

Если оба собственных значения матрицы вида (26) имеют отрицательные действительные части, то стационарное решение (x_s, y_s) асимптотически устойчиво, если же хотя бы одно из собственных значений имеет положительную действительную часть, то стационарное решение неустойчиво.

Заметим, что помимо решений, рассмотренных выше, стационарных решений (5), (8), (23) система (18), (19) имеет также стационарное решение $x_s = y_s = 0$. Оно не может быть изучено методом исследования устойчивости по первому приближению, так как матрица Якоби в точке $x_s = y_s = 0$ является нулевой матрицей и ее собственные значения равны нулю, что не позволяет сделать вывод относительно устойчивости. Однако неустойчивость этого стационарного решения может быть показана путем анализа знаков производных dx/dt , dy/dt в окрестности этой точки (см. упражнение 2).

Пример 2. Пусть численность возможных кандидатов общины X на проживание в данной местности составляет $N_x = 105$ тыс. человек, $a = 13/3$, $b = 0,04$. Община Y имеет аналогичные характеристики: $N_y = 105$ тыс. человек, $c = 13/3$, $k = 0,04$. Таким образом,

$$f(x) = \frac{13}{3} - 0,04x, \quad g(y) = \frac{13}{3} - 0,04y. \quad (27)$$

С содержательной точки зрения это означает, что наиболее толерантный член общины X считает соседство удовлетворительным, если на каждого «своего» приходится не более чем $13/3$ членов общины Y; другими словами, если община X составляет не менее чем примерно 19% населения. В то же время наименее толерантный член общины X удовлетворен соседством, если на каждого «своего» приходится не более чем $f(105) = 13/3 - 4,2 \approx 0,13$ членов общины Y; другими словами, если община X составляет не менее чем примерно 88,5% населения. Аналогичные утверждения справедливы для наиболее и наименее толерантных членов общины Y.

Уравнения (21), (22) для численностей членов общин, проживающих в квартале, принимаю в данном случае вид

$$\frac{dx}{dt} = -\beta x \left(\left(x - \frac{325}{6} \right)^2 - \frac{625(169 - 1,44y)}{36} \right), \quad (28)$$

$$\frac{dy}{dt} = -\alpha y \left(\left(y - \frac{325}{6} \right)^2 - \frac{625(169 - 1,44x)}{36} \right). \quad (29)$$

Эта система дифференциальных уравнений имеет четыре стационарных решения:

$$\begin{aligned} x_{s1} = y_{s1} = \frac{250}{3}, \quad x_{s2} = 100, \quad y_{s2} = \frac{100}{3}, \\ x_{s3} = \frac{100}{3}, \quad y_{s3} = 100, \quad x_{s4} = y_{s4} = 0. \end{aligned}$$

Чтобы изучить устойчивость первого из них, воспользуемся методом исследования устойчивости по первому приближению. Матрица

Якоби (26), взятая в точке (x_{s1}, y_{s1}) , равна

$$J(x_{s1}, y_{s1}) = J\left(\frac{250}{3}, \frac{250}{3}\right) = \frac{6250}{3} \begin{pmatrix} -\frac{7\beta}{3} & -\beta \\ -\alpha & -\frac{7\alpha}{3} \end{pmatrix}. \quad (30)$$

Характеристическое уравнение имеет вид $\lambda^2 + \frac{7}{3}\lambda(\alpha + \beta) + \frac{40\alpha\beta}{9} = 0$. Оба его корня вещественны; из теоремы Виета следует, что они отрицательны. Следовательно, стационарное решение (x_{s1}, y_{s1}) асимптотически устойчиво.

Для стационарного решения (x_{s2}, y_{s2}) имеем

$$J(x_{s2}, y_{s2}) = J\left(100, \frac{100}{3}\right) = 2500 \begin{pmatrix} -\frac{11\beta}{3} & -\beta \\ -\frac{\alpha}{3} & \frac{5\alpha}{9} \end{pmatrix}.$$

Характеристическое уравнение $\lambda^2 + \frac{33\beta - 5\alpha}{9}\lambda - \frac{64\alpha\beta}{27} = 0$ имеет два вещественных корня; из теоремы Виета следует, что они имеют противоположные знаки. Следовательно, стационарное решение (x_{s2}, y_{s2}) неустойчиво.

Аналогично показывается, что неустойчиво стационарное решение (x_{s3}, y_{s3}) (симметричное по отношению к (x_{s2}, y_{s2})). Можно доказать (см. упражнение 2), что стационарное решение (x_{s4}, y_{s4}) также является неустойчивым.

Таким образом, единственным асимптотически устойчивым стационарным решением в данном примере является решение $x_{s1} = y_{s1} = \frac{250}{3}$. Это случай качественно представлен на рис. 5, рис. 6.

4.3.4. Заключение

Представленная в настоящем разделе модель описывает механизм социальной сегрегации, основанный на неудовлетворенности индивидов проживанием «среди чужих», сформулированном здесь в виде поведенческой гипотезы. Один из возможных социологических сценариев описывает полный исход одной из общин из квартала. Пусть, например, изначально община X находилась в меньшинстве, но некоторые из ее членов были удовлетворены текущим соотношением. Однако по мере того, как другие (менее толерантные) члены общины X покидают квартал, соотношение меняется еще более в пользу общины Y. Ввиду этого некоторые из ранее удовлетворенных членов общины X становятся неудовлетворенными и также покидают квартал. Этот процесс продолжается до тех пор, пока вся община X не покинет квартал.

Возможны и другие сценарии: в частности, в случае достаточно крупных общин существует континуум устойчивых состояний, характеризующихся различным соотношением численностей общин.

С методологической точки зрения отметим, что на качественном уровне ряд выводов может быть получен без конкретизации поведенческой гипотезы. Это позволяет сделать вывод об устойчивости основных свойств модели по отношению к выбору конкретных функций, описывающих динамику процесса.

Упражнения

1. Для случая $x_1 > N_x$, $y_2 > N_y$ изобразить области значений переменных, в которых динамика численности членов общин X, Y, проживающих в квартале, имеет постоянное направление (аналогично рис. 3) и построить фазовый портрет (аналогично рис. 4). Определить функции $y = \varphi_1(x)$, $y = \varphi_2(x)$, $x = \psi_1(y)$, $x = \psi_2(y)$. Получить аналоги уравнений (2), (3).

2. Для примера 2 доказать неустойчивость стационарного решения $x_{s4} = y_{s4} = 0$. Для этого проанализировать знаки производных dx/dt , dy/dt в окрестности этой точки.

3. Рассмотреть пример 2 при ограничении $x + y \leq 100$. Построить множество устойчивых стационарных решений.

Комментарии и библиографические примечания

1. Некоторые международные документы ограничивают понятие сегрегации насильственными процессами. Так, в документе [1] Европейской комиссии против расизма и нетолерантности отмечается, что «добровольное действие отделения себя от других индивидов <...> не образует сегрегации». В соответствии с данными представлениями изучаемый в настоящем разделе процесс не рассматривается как сегрегация. Тем не менее вслед за работами [2, 3] в науке укоренилось более широкое понятие сегрегации, включающее в себя данный процесс.

2. Основные положения были введены нобелевским лауреатом по экономике 2005 г. Т. Шеллингом в работах [2, 3], там же предложена и изучена качественная модель. Конкретизация поведенческой гипотезы и математическая модель, имеющая вид системы обыкновенных дифференциальных уравнений, разработаны авторами и публикуются впервые.

3. Сегрегация довольно часто наблюдается в современных европейских городах. Так, в лондонском округе Излингтон доля белого населения составляет 68%, а в округе Ньюхэм — лишь 20% (по переписи 2011 года). Однако такого рода неоднородность населения возникает, по-видимому, в результате совместного действия двух явлений: механизма сегрегации, рассмотренного в данном разделе, а также экономической сегрегации.

Литература

1. European Commission against Racism and Intolerance General Policy Recommendation № 7: On national legislation to combat racism and racial discrimination. Adopted by ECRI on 13 December 2002.
2. *Schelling T.C.* Models of Segregation // *The American Economic review*. V. 59, № 2, (May 1969). P. 488-493.
3. *Schelling T.C.* Dynamic Models of Segregation // *Journal of Mathematical Sociology*, 1971. V. 1. P. 143–186.

4.4. Поведение индивидов при изменении количества степеней свободы

В некоторых социальных ситуациях поведение индивидов зависит от определенного фактора, который можно трактовать как степень свободы. Приведем простейший пример: пусть в некотором районе планируется открытие детских спортивных секций. Количество этих секций и их направленность могут быть различными, при этом потребность в каждой отдельной секции зависит от наличия или отсутствия альтернатив. Например, гандболом готовы заниматься 35 детей, если это будет единственный доступный вид спорта; только 22 ребенка, если ему будет альтернатива в виде волейбола, и лишь 5 детей, если будут открыты также футбольная секция. Некоторые дети предпочитают гандболу любой из двух альтернативных вариантов, другие — только футбол или, наоборот, только волейбол. В то же время планирующие органы исходят из двух критериев эффективности: с одной стороны, следует открывать только те секции, которые будут в достаточной степени востребованы, с другой стороны, желательно охватить как можно большее количество детей. Требуется построить прогноз востребованности секций при самых различных их сочетаниях. Например, если максимально возможный план заключается в открытии пяти секций, то необходим прогноз количества вовлеченных детей в следующих ситуациях: если будет открыта ровно одна секция (5 возможных случаев), ровно две секции (10 возможных случаев) и т. д. — всего 31 случай. Очевидно, что даже при столь небольшом количестве возможных секций построение прогноза не может быть проведено на основе исключительно социологического опроса, а требуется привлечение методов математического моделирования.

В самых общих чертах суть излагаемого в настоящем разделе подхода к построению прогноза востребованности, например гандбольной секции, заключается в том, что каждая альтернатива гандболу рассматривается как степень свободы индивидов; и количество членов данной секции рассматривается как функция от количества степеней свободы.

Данный подход является универсальным в том плане, что он применим ко всем ситуациям, в которых поведение индивидов зависит

от количества степеней свободы. В частности, изложение ведется на содержательном материале той проблемы, для которой был разработан этот подход, именно проблемы оценки границ рынка с помощью так называемого *теста гипотетического монополиста*.

Далее, первый пункт представляет собой введение в данную проблематику, во втором вводятся основные понятия, используемые в изложении. Третий и четвертый пункты являются основными: в них излагается концепция степени свободы индивида, формулируются соответствующие поведенческие гипотезы и на их основе разрабатывается алгоритм построения прогноза поведения индивида. Пятый пункт посвящен заключительным выводам.

4.4.1. Введение: содержательная проблема теста гипотетического монополиста и понятие о степени свободы

Хорошо известно, что антимонопольное регулирование рыночной экономики является одной из важнейших задач государства. Решение вопроса о необходимости проведения соответствующих мероприятий в каком-либо конкретном случае разбивается на три этапа.

1. Определение границ соответствующего рынка.
2. Оценка степени доминирования (рыночной власти) того или иного участника (или группы участников) в пределах этого рынка.
3. В случае, если доминирование действительно наблюдается, рассматривается вопрос о применении конкретных антимонопольных мер.

Одним из основных инструментов, применяемых при решении первого пункта данного плана, является так называемый «тест гипотетического монополиста» (ТГМ). В самых общих чертах его идея выглядит следующим образом. Продуктовые границы рынка определяют относительно обособленную группу товаров, так что при повышении цены на часть товаров этой группы покупатели переключают свое потребление в основном на другие товары той же группы. Таким образом, при повышении цены на часть товаров рынка будет наблюдаться значительное падение спроса на эти товары, а при повышении цены на все товары, относящиеся к рынку, падение спроса будет небольшим. Следовательно, фирма, желающая установить монопольные цены, должна стремиться к контролю над продажами всех товаров, составляющих рынок. В свою очередь антимонопольный орган должен определить границы рынка, не дожидаясь его реального захвата какой-либо фирмой и установления монопольных цен.

В утрированном виде проблему можно пояснить следующим примером. Предположим, что в некотором городе определенная фирма сконцентрировала в своих руках всю торговлю клубникой. Является ли эта фирма монополистом? Контролирует ли она цену данного товара? Ответ зависит от того, образует ли клубника отдельный рынок или представляет часть лишь более обширного рынка (например, рынка фруктов или фруктов и овощей). В первом случае данная фирма имеет

возможность увеличить прибыль, воспользовавшись своим монопольным положением и подняв цену клубники. Во втором случае увеличение цены приведет к переключению покупателей на другие фрукты и уменьшению прибыли фирмы (количественные аспекты данного рассуждения рассматриваются ниже).

В подобных случаях невозможны ни эксперимент, ни наблюдение, ни имитация, ни рассуждение по аналогии. Поэтому математическое моделирование выступает как необходимый инструмент для принятия управленческих решений, имеющего практические последствия. Фактически упомянутый выше Приказ ФАС № 108 предусматривает обязательное применение моделирования в целях определения границ рынков.

Реализация ТГМ в обязательном порядке предусматривает построение прогноза падения спроса на релевантные товары при повышении цен на них. Основная идея данного метода была высказана достаточно давно, однако до последнего времени она не была разработана до уровня, допускающего алгоритмизацию с программной реализацией на ЭВМ и эффективное применение на практике. Относительно активно ТГМ применяется Еврокомиссией в случаях минимального количества гипотез о границах рынка. Например, при принятии в 2001 г. решения по слиянию «CVC/Lenzing» антимонопольный орган решал вопрос о том, образуют ли штапельное вискозное волокно и древесное волокно один общий рынок или два разных рынка. С технической точки зрения эта задача является существенно более простой, чем если бы релевантными являлись не два, а три или более товаров. Разработка методов ТГМ приводила лишь к построению теоретических алгоритмов, не пригодных к практическому применению ввиду необходимости сбора нереально больших объемов информации. В частности, в существующих работах, посвященных алгоритмизации ТГМ (см. библиографические примечания в конце раздела), предполагаются известными все эластичности спроса (включая определяющие близость товаров перекрестные эластичности) и необходимые экзогенные параметры.

Однако в практической деятельности задачи, возникающие при применении ТГМ (за исключением случаев минимального количества гипотез о границах рынка), являются задачами с неполной информацией, поскольку собрать «сколь угодно полную» информацию оказывается возможным лишь в исключительных случаях.

Ввиду неполноты информации результатов опроса покупателей оказывается недостаточно для построения прогноза спроса при изменении цен на те или иные товары, а следовательно, и для определения границ рынка. Поэтому применение ТГМ требует выдвижения определенных гипотез о поведении покупателей.

Основная поведенческая гипотеза, применяемая в наших алгоритмах, связана с построением верхней и нижней границ изменения спроса с последующей интерполяцией. При этом вводится переменная,

называемая «степенью свободы» переключения покупателя, а спрос рассматривается как функция степени свободы.

4.4.2. Основные определения

Понятие гипотетического монополиста основано на идее об определении границ соответствующего рынка, представляющего собой относительно обособленную группу товаров, испытывающих конкуренцию, в основном, только со стороны других продуктов этого же рынка. Очевидно, что эти продукты должны в значительной степени замещать друг друга. В то же время возможность замещения продуктов, составляющих рынок, продуктами, не входящими в этот рынок, должна быть незначительной. Если бы этот рынок контролировался неким гипотетическим монополистом (стремящимся к увеличению прибыли и не находящимся под воздействием мер антимонопольного регулирования), то проведенное этим монополистом небольшое, но существенное и долгосрочное повышение цены на *все* товары этого рынка привело бы к увеличению его прибыли. Причина заключается в том, что покупатели не смогли бы в достаточной мере отказаться от потребления продуктов этого рынка ввиду недостаточной возможности заместить эти продукты другими продуктами, в него не входящими. В то же время повышение цен лишь на часть продуктов данного рынка приведет к уменьшению прибыли, полученной от торговли этими продуктами, так как покупатели смогут заместить эти товары другими продуктами, входящими в рынок, но не испытывавшими повышения цен.

В соответствии с изложенным, дается следующее определение.

Определение. Рынком называется группа товаров (в частности, группа может состоять из одного товара) с географической областью, на которой эта группа продуктов продается, такая что гипотетическая фирма (стремящаяся к увеличению прибыли и не являющаяся объектом антимонопольного регулирования), являющаяся единственным продавцом этих товаров:

1) может повысить цену на эту группу продуктов, увеличив свою прибыль,

2) не может увеличить свою прибыль, повысив цену только на часть товаров этой группы и оставив цены на другие товары этой группы без изменений.

Задача определения границ рынка заключается в выделении товаров, составляющих рынок, из некоторого более широкого множества товаров. Это множество мы будем называть *максимальной группой* или *группой*. Перед проведением ТГМ максимальная группа товаров должна быть предварительно отобрана экспертами антимонопольного органа. Подмножества максимальной группы называются *подгруппами*. Они являются кандидатами на то, чтобы быть рынком. Максимальной подгруппой является сама максимальная группа. Минимальными подгруппами являются подгруппы из одного товара. Если максимальная подгруппа содержит K товаров, то в ней можно выделить $2^K - 1$

различных подгрупп (для трех товаров имеется уже семь подгрупп — кандидатов в рынки).

Покупатель рассматривается как в целом рациональный субъект и как источник информации для принятия решений об определении границ рынка. Каждый покупатель может приобретать один или несколько товаров максимальной группы. При повышении цены на покупаемый им товар покупатель может либо переключиться на другой товар, либо уйти с рынка или сократить закупки подорожавшего товара, либо не отреагировать на повышение цен, либо применить смешанную стратегию. Рациональность поведения покупателя состоит, например в том, что если в начальный момент времени покупатель предпочитал товар G_1 товару G_2 , то после повышения цены на товар G_2 это предпочтение сохранится. Покупатель является источником информации, поскольку при проведении опроса он выступает в качестве респондента, описывающего свое поведение в случае повышения цены на ту или иную подгруппу товаров.

В следующих двух разделах идея интерполяции спроса рассматривается сначала в приложении к задаче определения товарных границ рынка, а затем — к более сложной задаче одновременной оценки товарных и географических границ рынка.

4.4.3. Оценка товарных границ рынка

В рамках предлагаемого подхода покупателям предъявляются список релевантных товаров и предлагают дать ответ на следующие вопросы.

1. *Какие товары из предложенного списка Вы закупаете?*
2. *В каком объеме Вы закупаете каждый из предложенных товаров?*
3. *На какую величину Вы сократите закупки этого товара при повышении цены только на этот товар на 10% (укажите, на сколько процентов от 0% до 100% Вы сократите закупки).*
4. *На какую величину Вы сократите закупки этого товара при повышении цены на все товары предложенного списка на 10% (укажите, на сколько процентов от 0% до 100% Вы сократите закупки).*

Минимальные подгруппы, подлежащие исследованию, — это подгруппы из одного товара. Для вычисления изменения прибыли на таких подгруппах не требуется никаких дополнительных гипотез или предположений, так как ответами на первый вопрос анкеты покупатели обеспечивают информацию, необходимую для анализа этих подгрупп. Максимальная подгруппа, подлежащая исследованию, — это подгруппа из всех релевантных товаров, т. е. совпадающая с данной, максимальной товарной группой. Для вычисления изменения прибыли на этой подгруппе также не требуется никаких дополнительных гипотез или предположений, так как ответами на второй вопрос анкеты покупатели обеспечивают информацию, необходимую для анализа этой подгруппы.

Чтобы дать обоснованную оценку приращения прибыли гипотетического монополиста в случае повышения цен на различные подгруппы товаров (кроме подгрупп из одного товара и максимальной подгруппы), необходимо ввести *основную поведенческую гипотезу*, основанную на следующих рассуждениях.

Пусть цена повышается только на товар G_1 . Тогда покупатели перераспределят часть своего спроса на другие товары данной группы и частично уменьшат закупки этого товара.

Пусть теперь цена повышается в равной мере уже на два товара: G_1 и G_2 . В этом случае перераспределение спроса с G_1 на G_2 будет происходить в меньшей степени, чем при повышении цены только на G_1 (возможно, перераспределения между G_1 и G_2 вообще не будет, а возможен и обратный переток спроса — от G_2 к G_1). В то же время перераспределение спроса с G_1 на другие товары данной группы, а также доля покупателей товара G_1 , отказавшихся от покупок товаров группы в результате повышения цен, не зависит от того, повышалась ли цена на товар G_2 .

Таким образом, падение спроса на товар G_1 в случае повышения цены только на G_1 будет не меньшим, чем в случае повышения цены на G_1 и G_2 . Вообще говоря, чем больше подгруппа товаров (содержащая G_1), на которые повышается цена, тем меньше падение спроса на товар G_1 . Другими словами, возможности покупателя переключиться на потребление другого товара определяются «свободой выбора», определяемой в свою очередь рыночной долей товаров, на которые не происходит повышения цен.

Это утверждение и составляет смысл принятой в рамках настоящего подхода поведенческой гипотезы [4, 5]:

Пусть Ω_1 и Ω_2 — некоторые подгруппы товаров, содержащие товар G_k . Обозначим через s_1 (соответственно, s_2) долю в общем объеме закупок товаров группы, приходящуюся на товары подгруппы Ω_1 , (соответственно, Ω_2), т. е.

$$s_1 = \frac{\sum_{\Omega_1} q_i^0}{\sum_{i=1}^N q_i^0}, \quad s_2 = \frac{\sum_{\Omega_2} q_i^0}{\sum_{i=1}^N q_i^0},$$

где в числителе каждого из выражений суммирование производится по всем товарам соответствующей подгруппы.

Основная поведенческая гипотеза формулируется следующим образом: если $s_1 \geq s_2$, то падение спроса на товар G_k при повышении цены на все товары подгруппы Ω_1 будет не большим, чем при повышении цены на все товары подгруппы Ω_2 .

В частности, падение спроса на некоторый товар будет минимальным при повышении цены на все товары рассматриваемой группы и максимальным при повышении цены только на этот товар. Таким образом, если ν — падение спроса на некоторый товар G_i при повышении цены на содержащую этот товар подгруппу Ω , то $0 < \mu_i < \nu < \nu_i$.

Опираясь на сформулированную поведенческую гипотезу, дадим оценку приращения прибыли гипотетического монополиста в случае повышения цен на различные подгруппы товаров (кроме подгрупп из одного товара и максимальной подгруппы, совпадающей с исходной товарной группой). Для того чтобы вычислить изменение прибыли на некоторой подгруппе товаров при повышении цен на них, необходимо вычислить изменение спроса на каждый из товаров этой подгруппы.

Конкретизируя поведенческую гипотезу, введем количественную меру зависимости падения спроса от доли s . Именно, *при повышении цен на некоторую подгруппу Ω положим падение спроса на входящий в нее товар линейно зависящим от суммарной доли продаж товаров, не входящих в эту подгруппу.*

В этом случае будем говорить о линейной интерполяции спроса.

Обозначим через Δq_k изменение спроса на k -й товар при повышении цен на некоторую включающего его подгруппу товаров Ω , а через s — долю в общем объеме закупок товаров группы, приходящуюся на товары подгруппы Ω :

$$s = \frac{\sum_{\Omega} q_i^0}{\sum_{i=1}^N q_i^0}.$$

Основную поведенческую гипотезу запишем в виде

$$\Delta q_k(s) = q_k^0 [-a(1-s) + b],$$

где $a > 0$ и b — некоторые, пока неизвестные коэффициенты. Знак «минус» перед коэффициентом a означает, что изменение спроса отрицательно (в нашей терминологии, изменение спроса равно его падению, взятому с противоположным знаком). Численные значения коэффициентов a , b можно определить на основе анализа сводной анкеты, суммирующей полученные от покупателей ответы на указанные выше вопросы.

При повышении цены только на некоторый товар G_k (т. е. при $s = q_i^0 / \sum_{i=1}^K q_i^0$) изменение спроса равно $-\nu_i q_k^0$. В то же время при повышении на все товары исследуемой большой группы (т. е. при $s = 1$) изменение спроса равно $-\mu_i q_k^0$. Из этих двух условий нетрудно найти значения коэффициентов a , b :

$$a = \frac{\nu_i - \mu_i}{1 - q_k^0 / \sum_{i=1}^K q_i^0}, \quad b = -\mu_i.$$

Таким образом, выражение для изменения спроса на товар G_k , произошедшее вследствие повышения цен на товары подгруппы Ω , занимаю-

щей долю s в общем объеме закупок товаров, имеет вид

$$\Delta q_k(s) = q_k^0 \left[\frac{\nu_i - \mu_i}{1 - q_k^0 / \sum_{i=1}^K q_i^0} (s - 1) - \mu_i \right]. \quad (1)$$

Вычислим изменение прибыли продавца на товаре G_k , произошедшее ввиду указанного повышения цены. Обозначим через p_k^0 цену на товар G_k до повышения цен, а через c_k — себестоимость единицы этого товара. Тогда прибыль от торговли товаром G_k до повышения цен имеет вид

$$\Pi_k^0 = q_k^0 (p_k^0 - c_k),$$

а после z -процентного повышения цены на товары подгруппы Ω

$$\Pi_k = (q_k^0 + \Delta q_k) ((1 + z)p_k^0 - c_k).$$

Приращение прибыли на товаре G_k имеет вид

$$\Delta \Pi_k = \Pi_k - \Pi_k^0.$$

Общее приращение прибыли гипотетического монополиста на подгруппе товаров Ω вычисляется суммированием приращений прибыли по каждому из товаров этой подгруппы:

$$\Delta \Pi = \sum_{k \in \Omega} [(q_k^0 + \Delta q_k) ((1 + z)p_k^0 - c_k) - q_k^0 (p_k^0 - c_k)]. \quad (2)$$

Алгоритм применения ТГМ предусматривает, что на исследуемой подгруппе товаров Ω по указанной формуле вычисляется приращение прибыли. Если оно отрицательно, то Ω не является рынком. Если же приращение прибыли положительно, то вычисляется приращение прибыли на каждом из подмножеств Ω , содержащем на один товар меньше, чем Ω . Если все эти приращения прибыли отрицательны, то Ω является рынком, в противном случае Ω — это подгруппа, бóльшая, чем рынок.

4.4.4. Одновременное определение товарных и географических границ рынка

В настоящем подразделе идея оценки границ рынка методом интерполяции спроса прилагается к решению следующей задачи. Имеются несколько точек продаж, в которых продаются несколько товаров. Требуется определить одновременно товарные и географические границы рынка, т.е. множество товаров и точек продажи, контроль над которыми позволяет гипотетическому монополисту повысить цены с увеличением прибыли. При этом ответ может выглядеть, например, следующим образом: увеличение прибыли приносит как контроль над

продажами товара G_1 в пяти торговых точках A_1, A_2, A_3, A_4, A_5 , так и контроль над продажами товаров G_1, G_2 в трех торговых точках A_1, A_2, A_3 , так и контроль над продажами товаров G_1, G_2, G_3 в одной торговой точке A_1 и т. д.

Мы опишем кратко идею одновременной оценки товарных и географических границ методом интерполяции спроса, а конкретный расчет приведем для примера.

Необходимая для применения метода информация должна быть получена из опроса покупателей. В анкете должны быть описаны множество товаров и множество точек продаж, а также содержаться следующие вопросы.

1. *Какие товары, в каких точках продаж и в каком количестве Вы покупаете?*

Далее респондент отвечает по каждому товару и каждой точке продаж.

2. *Если цена на этот товар в этой точке увеличится а 10%, то на сколько Вы сократите закупки этого товара в этой точке?*

3. *Если цена на все товары в этой точке увеличится а 10%, то на сколько Вы сократите закупки этого товара в этой точке?*

4. *Если цена на этот товар во всех точках увеличится а 10%, то на сколько Вы сократите закупки этого товара в этой точке?*

Предположим, что требуется оценить изменение спроса на некоторый товар в некоторой точке продаж при 10-процентном повышении цены на некоторую группу товаров в некоторых точках продаж (в т. ч. на товар, изменение спроса на который оценивается). Воспользуемся идеей оценки на основе линейной интерполяции спроса.

При оценке только товарных границ рынка изменение спроса на товар рассматривалось как линейная функция одной переменной, а именно доли группы (на которую повышается цена) в общем объеме продаж. При одновременной оценке товарных и географических границ рынка мы рассматриваем изменение спроса на товар G_j в точке A_i как линейную функцию двух переменных, а именно: (*) доли группы товаров (содержащей товар G_j) в общем объеме продаж в точке A_i и (**) доли группы точек (содержащей точку A_i) в общем объеме продаж товара G_j .

Таким образом, изменение спроса (а значит, и сам спрос) геометрически изображается плоскостью в пространстве трех переменных. Вертикальная ось имеет смысл самого изменения спроса, а горизонтальные оси — соответствующие доли. Ответы на вопросы 2–4 (суммированные по всем покупателям) позволяют построить три точки, задающие плоскость, которая является двумерной линейной интерполяцией спроса.

С ее помощью оценивается падение спроса на данный товар в данной точке продаж.

Итак, предположим, что требуется оценить изменение прибыли продавца при повышении цен на некоторые товары в некоторых точках продаж. Для каждого товара, входящего в ту группу, и для каждой точки его продаж оценивается падение спроса указанным методом. После этого стандартным способом вычисляется изменение прибыли на этом товаре в этой точке и т. д.

Проведем конкретный расчет для следующего примера.

Пример. В каждой из трех точек продаж A_1 , A_2 , A_3 продаются товары G_1 , G_2 . Одна единица товара G_1 , с потребительской точки зрения, эквивалентна одной единице G_2 (т. е., если для некоторого потребителя товары взаимозаменяемы, то он замещает одну единицу одного товара одной единицей другого товара). Себестоимость единицы каждого товара для продавца составляет $c = 4$ руб, а продажная цена — $p^0 = 10$ руб.

В точке продаж A_1 продается в месяц 10 единиц товара G_1 (для удобства будем называть соответствующих покупателей B_1) и 10 единиц товара G_2 (для удобства будем называть соответствующих покупателей B_4) и т. д. Соответствующие данные сведены в таблицу, каждая ячейка которой содержит название покупателей и объем продаж.

Точки продаж	G_1	G_2
A_1	$B_1, 10$	$B_4, 10$
A_2	$B_2, 20$	$B_5, 20$
A_3	$B_3, 20$	$B_6, 20$

Был проведен опрос покупателей, который показал следующие результаты.

Категория покупателей	Снижение спроса при повышении цены (в %)		
	Только на данный товар в данной точке	На оба товара в данной точке	На данный товар во всех трех точках
B_1	40	20	10
B_2	40	20	20
B_3	40	10	20
B_4	40	20	10
B_5	40	10	20
B_6	40	20	20

Построим линейную функцию падения спроса для покупателей B_1 . Обозначим долю падения спроса через z ($0 < z \leq 1$), долю группы товаров в точке продаж A_1 — через x ($0 < x \leq 1$), долю группы товаров во всем объеме продаж товара G_1 — через y ($0 < y \leq 1$). Линейная функция падения спроса определяется как плоскость, проходящая через три точки:

$$x = 0,5; y = 0,2; z = 0,4$$

(при повышении цены только на товар G_1 только в точке продаж A_1 доля товара в точке составляет 0,5 (10 из 20), доля в общем объеме продаж товара G_1 — 0,2 (10 из 50), падение спроса составляет 40%);

$$x = 0,5; y = 1; z = 0,1$$

(при повышении цены только на товар G_1 во всех точках продаж доля товара в точке A_1 составляет 0,5 (10 из 20), доля в общем объеме продаж товара G_1 — 1 (50 из 50), падение спроса составляет 10%);

$$x = 1; y = 0,2; z = 0,2$$

(при повышении цены на оба товара только в точке продаж A_1 доля товара в точке составляет 1 (20 из 20), доля в общем объеме продаж товара G_1 — 0,2 (10 из 50), падение спроса составляет 20%).

По известной формуле строится уравнение плоскости, проходящей через три указанные точки:

$$\begin{vmatrix} x - 0,5 & y - 0,2 & z - 0,4 \\ 0,5 - 0,5 & 1 - 0,2 & 0,1 - 0,4 \\ 1 - 0,5 & 0,2 - 0,2 & 0,2 - 0,4 \end{vmatrix} = 0.$$

После несложных преобразований получаем уравнение для линейной функции падения спроса покупателей B_1 :

$$0,16x + 0,15y + 0,4z - 0,27 = 0.$$

Аналогичным образом получаем уравнения для линейных функций падения спроса других покупателей:

$$B_2: 0,12x + 0,1y + 0,3z - 0,22 = 0,$$

$$B_3: 0,18x + 0,1y + 0,3z - 0,25 = 0,$$

$$B_4: 0,16x + 0,15y + 0,4z - 0,27 = 0,$$

$$B_5: 0,18x + 0,1y + 0,3z - 0,25 = 0,$$

$$B_6: 0,12x + 0,1y + 0,3z - 0,22 = 0.$$

С помощью найденных линейных функций протестируем различные группы товаров и точек продаж.

Рассмотрим, например, товар A_1 и точку продаж G_1 . Выясним, может ли гипотетическая фирма, являющаяся единственным продавцом данного товара в данной точке, повысить цену на 10% с выгодой для себя. Первоначальный спрос составляет 10 единиц, цена — 10 руб., а себестоимость — 4 руб. Следовательно, прибыль продавца составляет $\Pi = 10 \cdot (10 - 4) = 60$ руб. После гипотетического повышения цены

на товар G_1 до 11 руб. спрос упадет на 40% и составит 6 единиц. Прибыль окажется равной $\Pi = 6 \cdot (11 - 4) = 42$ руб. Падение прибыли свидетельствует о том, что товар G_1 в точке продаж A_1 не является рынком.

Рассмотрим теперь множество, состоящее из обоих товаров, продаваемых в точке A_1 . Первоначально прибыль продавца составляет $\Pi = 20 \cdot (10 - 4) = 120$ руб. После гипотетического повышения цены на оба товара до 11 руб. спрос покупателей B_1 на первый товар упадет на 20% и составит 8 единиц, а спрос покупателей B_4 на товар G_1 упадет также на 20% и составит 8 единиц. Прибыль окажется равной $\Pi = 16 \cdot (11 - 4) = 112$ руб. Падение прибыли свидетельствует о том, что точка продаж A_1 не является рынком.

Рассмотрим товар G_1 , продаваемый во всех трех точках. Первоначально прибыль продавца составляет $\Pi = 50 \cdot (10 - 4) = 300$ руб. После гипотетического повышения цены на товар G_1 во всех точках до 11 руб. спрос покупателей B_1 в первой точке упадет на 10% и составит 9 единиц, спрос покупателей B_2 в точке A_2 упадет на 20% и составит 16 единиц, а спрос покупателей B_3 в точке A_3 упадет на 20% и составит также 16 единиц. Прибыль окажется равной $\Pi = 41 \cdot (11 - 4) = 287$ руб. Падение прибыли свидетельствует о том, что товар G_1 не образует рынок.

Для исследования перечисленных выше кандидатов на рынок интерполяция не требовалась. Продемонстрируем, каким образом она проводится, на двух следующих кандидатов рынки.

Рассмотрим множество, состоящее из обоих товаров, продаваемых в точке A_2 , и товара G_1 , продаваемого в точке A_3 . Первоначально прибыль продавца составляет $\Pi = 60 \cdot (10 - 4) = 360$ руб. После гипотетического повышения цены до 11 руб. спрос покупателей B_2 на первый товар определится из уравнения

$$B_2: 0,12x + 0,1y + 0,3z - 0,22 = 0.$$

Решив уравнение при $x = 1$ (40 единиц из 40), $y = 0,8$ (40 единиц из 50), найдем $z = 0,07$, т.е. спрос покупателей B_2 упадет на 7% и составит 18,7 единиц. Спрос покупателей B_3 на первый товар определится из уравнения

$$B_3: 0,18x + 0,1y + 0,3z - 0,25 = 0.$$

Решив уравнение при $x = 0,5$ (20 единиц из 40), $y = 0,8$ (40 единиц из 50), найдем $z = 0,27$, т.е. спрос покупателей B_3 упадет на 27% и составит 14,7 единиц.

Спрос покупателей B_5 на товар G_2 определится из уравнения

$$B_5: 0,18x + 0,1y + 0,3z - 0,25 = 0,$$

где $x = 1$ (40 единиц из 40), $y = 0,4$ (20 единиц из 50). Решив уравнение, найдем $z = 0,1$, т.е. спрос покупателей B_5 упадет на 10% и составит 18 единиц. Общий спрос в рассматриваемой

группе составит $18,7 + 14,7 + 18 = 51,4$ ед. Прибыль окажется равной $\Pi = 51,4 \cdot (11 - 4) = 359,8$ руб. Падение прибыли свидетельствует о том, что рассматриваемая группа не является рынком.

Рассмотрим теперь множество, состоящее из обоих товаров, продаваемых в точке A_2 , и товара G_2 , продаваемого в точке A_3 . Первоначально прибыль продавца составляет $\Pi = 60 \cdot (10 - 4) = 360$ руб. После гипотетического повышения цены до 11 руб. спрос покупателей B_2 на первый товар определится из уравнения

$$B_2: 0,12x + 0,1y + 0,3z - 0,22 = 0.$$

Решив уравнение при $x = 1$ (40 единиц из 40), $y = 0,4$ (20 единиц из 50), найдем $z = 0,2$, т.е. спрос покупателей B_2 упадет на 20% и составит 16 единиц.

Спрос покупателей B_5 на товар G_2 определится из уравнения

$$B_5: 0,18x + 0,1y + 0,3z - 0,25 = 0.$$

Решив уравнение при $x = 1$ (40 единиц из 40), $y = 0,8$ (40 единиц из 50), найдем $z = 0$, т.е. спрос покупателей B_5 не упадет (в рассматриваемом приближении) и составит 20 единиц. Спрос покупателей B_6 на товар G_2 определится из уравнения

$$B_6: 0,12x + 0,1y + 0,3z - 0,22 = 0.$$

Решив уравнение при $x = 0,5$ (20 единиц из 40), $y = 0,8$ (40 единиц из 50), найдем $z = 0,27$, т.е. спрос покупателей B_6 упадет на 27% и составит 14,7 единиц.

Общий спрос в рассматриваемой группе составит $18 + 20 + 14,7 = 52,7$ ед. Прибыль окажется равной $\Pi = 52,7 \cdot (11 - 4) = 368,9$ руб. Увеличение прибыли свидетельствует о том, что рассматриваемая группа является рынком (нетрудно показать, что увеличение цены на любом ее меньшем подмножестве приводит к уменьшению прибыли).

Итак, фирма, являющаяся единственным продавцом обоих товаров в точке A_2 и товара G_2 в точке A_3 , может повысить на 10% цену на эти товары, получив при этом увеличение прибыли.

4.4.5. Заключение

Концепция степени свободы индивида и возможность ее применения в математическом моделировании проиллюстрированы в данном разделе на материале теста гипотетического монополиста. Однако потенциальная сфера применения данной концепции гораздо шире. Эта сфера распространяется в области, где существует необходимость построения прогнозов поведения индивидов при непостоянном количестве стоящих перед ними альтернативных вариантов.

В качестве примера подобной ситуации можно привести выборы депутата в одномандатном округе, проходящие в один тур. Количество голосов, поданных за конкретного кандидата, зависит не только от общего количества кандидатов, но и от сложной системы предпочтений

избирателей. Например, некоторые избиратели предпочитают кандидату А лишь кандидата В, другие — лишь кандидата С или любого из этих двух и т. д. Поскольку подробное исследование данной системы предпочтений социологическими методами не всегда возможно, то возникает задача построения оценок на основе неполной информации. Такие оценки могут быть построены с помощью методики, разработанной в данном разделе.

Упражнения

1. Предложить несколько примеров социальных ситуаций, в которых построение прогноза поведения индивидов может быть проведено на основе концепции степени свободы, изложенной в настоящем разделе.
2. В заключении в качестве примера указана ситуация с выборами депутата. Разработайте конкретную модель для данного случая по методике, изложено в разделе.

Комментарии и библиографические примечания

1. Подробный обзор работ по тематике ТГМ см. в [1]. В качестве известных работ, посвященных алгоритмизации ТГМ, в которых предполагаются известными все эластичности спроса (включая определяющие близость товаров перекрестные эластичности, и необходимые экзогенные параметры) укажем [2, 3].
2. Настоящий раздел написан по материалам работ [4, 5].
3. Относительно реальной практики применения ТГМ антимонопольными органами отметим, что при определении продуктовых границ рынка в 1999–2001 годах Европейская комиссия принимала решение на основе ТГМ в 11 % случаев [1]. В Российской Федерации применение ТГМ предусмотрено Приказом Федеральной антимонопольной службы от 25 марта 2006 г. № 108.

Литература

1. *Алешин Д.А.* Новые подходы ФАС России к анализу товарных рынков: применение теста гипотетического монополиста // Анализ товарных рынков в антимонопольном регулировании. Технологии и алгоритмы. Под ред. Д.А. Алешина. М.: ФАС России, Маркет ДС Корпорейшн, 2007. С. 5–19.
2. *Dobbs I.M.* Demand Cost Elasticities and Pricing Benchmarks in the Hypothetical Monopoly Test: The Consequences of a Simple SSNIP.
3. *Dobbs I.M.* Defining Markets for ex ante Regulation Using the Hypothetical Monopoly Test // International Journal of the Economics of Business 2006, 13(1), 83–109. Applied Economics Letters 2003, 10, 545–548.
4. *Михайлов А.П., Петров А.П., Алешин Д.А.* Математические модели и алгоритмы определения границ рынка с помощью теста гипотетического

монополиста // Анализ товарных рынков в антимонопольном регулировании. Технологии и алгоритмы. Под ред. Д. А. Алешина. М.: ФАС России, Маркет ДС Корпорейшн, 2007. С. 20–54.

5. Михайлов А.П., Петров А.П., Алешин Д.А. Антимонопольная деятельность и математическое моделирование // Информационные технологии и вычислительные системы, 2007, № 4. С. 12–21.

4.5. Распространение информации в социуме в отсутствие противодействия

Современные подходы к изучению гуманитарных проблем информационной безопасности (т. е. проблем, связанных с наличием «человеческого фактора») могут быть существенно обогащены математическими методами. К ним относится (прежде всего) использование соответствующих математических моделей. Данный раздел посвящен построению и исследованию математических моделей распространения информации в социуме в отсутствие противодействия. Далее будут изучены модели информационного противоборства.

Эти два класса моделей описывают два класса процессов распространения информации. Под противоборством мы понимаем процесс распространения в социуме нескольких конкурирующих информационных потоков. Термин «конкурирующие» означает в данном случае, что каждый индивид участвует в распространении не более чем одного потока. Например, если две политические партии пропагандируют два противоположных взгляда на одно и то же событие, то каждый член общества поддерживает не более чем один из них. В таком случае каждая из партий, рекрутируя сторонников, делает их недоступными для другой партии и тем самым противодействует ей. Аналогичный пример распространения информации в отсутствие противодействия — это распространение лишь одной точки зрения на событие. Такой процесс является более простым как в математическом, так и в содержательном аспекте. В определенном смысле настоящий раздел можно рассматривать как прелюдию к анализу более сложного и содержательного процесса информационного противоборства.

С математической точки зрения модели настоящего раздела представляют собой уравнения и системы обыкновенных дифференциальных уравнений. Общий план представления моделей соответствует иерархическому принципу: в первом пункте рассматривается базовая модель, во втором пункте вводится ряд дополнительных факторов распространения информации, а также проводится отдельное изучение каждого из этих факторов. Модель, учитывающая все эти факторы в совокупности, представлена в третьем пункте.

4.5.1. Базовая модель информационного нападения

Рассмотрим социальную общность, состоящую из N индивидов. В ней распространяется некоторая информация; индивиды, облада-

ющие ей (и распространяющие ее), называются адептами. Их численность в момент времени t обозначим через $X(t)$, численность игнорантов (неохваченных информацией индивидов) равна $N - X(t)$. В начальный момент времени $t = 0$ никто из членов этого социума не обладает данной информацией. При $t > 0$ информация распространяется двумя способами: через СМИ и путем межличностной коммуникации с людьми, охваченными ранее. Интенсивность распространения информации этими способами описывается параметрами α и β соответственно. Заметим, что интенсивность распространения информации через межличностную коммуникацию при этом пропорциональна также числу адептов. Предполагается, что скорость распространения информации (т.е. скорость увеличения количества адептов) складывается из скоростей распространения информации через СМИ и коммуникацию. Эта скорость пропорциональна количеству игнорантов, т.е. $N - X(t)$.

Введем следующие предположения.

Предпосылка А. Однажды получив информацию, игнорант становится адептом и немедленно начинает распространять ее.

Предпосылка В. Обратного перехода из адептов в игноранты не происходит.

Предпосылка С. Социум является однородным в том смысле, что: (а) все индивиды в равной мере получают информацию от СМИ, (б) при межличностной коммуникации каждый индивид в равной мере получает информацию от всех других индивидов.

Предпосылка Д. Параметры α и β , описывающие интенсивность распространения информации через СМИ и межличностную коммуникацию, суть постоянные величины.

При этих предположениях получаем базовую модель информационного нападения, имеющую вид

$$\frac{dX}{dt} = (\alpha + \beta X)(N - X), \quad (1)$$

$$X(0) = 0. \quad (2)$$

Решение этой задачи Коши имеет вид (рис. 1)

$$X(t) = N \frac{\alpha \exp [(\alpha + \beta N)t] - \alpha}{\alpha \exp [(\alpha + \beta N)t] + \beta N}. \quad (3)$$

Из несложного анализа следует, что оно положительно (кроме момента $t = 0$), монотонно возрастает с течением времени и стремится снизу к величине N . Таким образом, в итоге (формально, при $t \rightarrow \infty$) все члены группы переходят в категорию адептов.

Более тонкую градацию режимов распространения информации дает анализ скорости нарастания информационной угрозы в зависимости

от параметров модели. Нетрудно показать (см. упражнение 1, ниже), что при выполнении условия

$$N_{cr} = \frac{1}{2} \left(N - \frac{\alpha}{\beta} \right) > 0 \quad (4)$$

существует момент времени $t_{cr} > 0$, когда достигается максимальная скорость вербовки (доказательство данного утверждения вынесено в упражнения). Здесь $N_{cr} = N(t_{cr})$ — количество индивидов, завербованных в момент t_{cr} . Назовем ее точкой максимума «ажитожа».

Из анализа неравенства (4) следует, что эта точка существует не для любых значений параметров модели. В зависимости от соотношений между параметрами модели возможны три режима распространения информации.

1. $\alpha > \beta N$: в этом случае точка максимума ажиотажа отсутствует, а скорость перехода игнорантов в адепты максимальна при $t = 0$ и с течением времени уменьшается, рис. 1, а.

2. $\alpha = \beta N$: тогда $N_{cr} = 0$, т. е. максимум ажиотажа имеет место в начальный момент времени $t = 0$ и с течением времени скорость перехода в адепты уменьшается, рис. 1, б.

3. $\alpha < \beta N$: скорость перехода в адепты при $t < t_{cr}$ увеличивается, при $t = t_{cr}$ она максимальна, а при $t > t_{cr}$ уменьшается. При этом максимум ажиотажа достигается при численности адептов, равной N_{cr} , рис. 1, в.

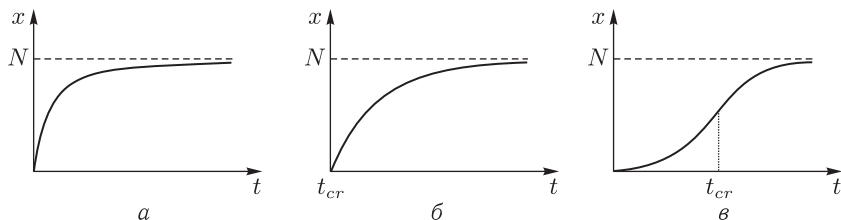


Рис. 1. Базовая модель: три режима (пояснения в тексте)

4.5.2. Учет дополнительных факторов

Перейдем от рассмотрению серии моделей, учитывающих ряд дополнительных (по сравнению с базовой моделью) факторов, влияющих на распространение информации в социуме, именно: усвоение информации индивидами лишь после двухкратного ее получения, забывание (утра интереса) информации адептами, а также неполный охват социума средствами массовой информации.

Двухшаговое усвоение информации

Практика показывает, что информационное воздействие является результативным, если оно является не одноразовым, а проводится в виде кампании, обеспечивающей многократное (в самом простом случае — двукратное) получение индивидом одной и той же инфор-

мации. Соответственно в данном разделе рассматривается вариант модели распространения информации, предполагающий, что игнорант начинает распространять информацию лишь после того, как получит ее дважды. Это предположение изменяет предпосылку А базовой модели, рассмотренной в первом пункте.

Индивидов, получивших информацию ровно один раз, будем называть предадептами. Таким образом, предадепты, хотя и знакомы с информацией, но не распространяют ее далее. Обозначим их численность через $x(t)$. Адепты рекрутируются из предадептов, поэтому

$$\frac{dX}{dt} = x(\alpha + \beta X). \quad (5)$$

Поскольку адепты рекрутируются из игнорантов, численность которых в момент времени t составляет $N_0 - X(t) - x(t)$, то уравнение динамики численности предадептов имеет вид

$$\frac{dx}{dt} = (N - X - 2x)(\alpha + \beta X). \quad (6)$$

Система уравнений (5), (6) представляет собой модель информационного нападения, учитывающую двухшаговое усвоение информации индивидами. Нетрудно видеть, что стационарное решение этой системы (5), (6) единственно, асимптотически устойчиво и имеет вид $x_{st} = 0$, $X_{st} = N$. Таким образом, процесс распространения информации завершается тем же, что и в базовой модели: все индивиды переходят в категорию адептов.

Распространение информации с учетом забывания

Положим теперь, что не выполняется предпосылка В базовой модели, т. е. возможен обратный переход из адептов в игноранты в результате забывания информации (или потери интереса к ней). Тогда уравнение модели принимает вид

$$\frac{dX}{dt} = (\alpha + \beta X)(N - X) - \gamma X. \quad (7)$$

Здесь параметр $\gamma > 0$ характеризует интенсивность забывания информации адептами. Найдем стационарное решение уравнения (7). Положив $dX/dt = 0$, получим

$$(\alpha + \beta X)(N - X) - \gamma X = 0.$$

При любых положительных значениях α , β , γ и N это уравнение имеет два корня, один из которых положителен. Таким образом, стационарное решение всегда существует, единственно и равно

$$X_{st} = \frac{1}{2\beta} \left(-\gamma - \alpha + \beta N \sqrt{(\gamma + \alpha - \beta)N^2 + 4\alpha\beta N} \right). \quad (8)$$

Нетрудно видеть, что $dX/dt > 0$ при $X < X_{st}$ и $dX/dt < 0$ при $X > X_{st}$; следовательно, данное решение является асимптотически устойчивым.

Перейдем от анализа стационарного решения к вопросу о максимуме ажиотажа аналогично тому, как это было рассмотрено для базовой

модели выше. В этом случае анализ второй производной функции $X(t)$ показывает, что возможны три режима распространения информации, соответствующие трем случаям, представленным на рис. 1 для базовой модели.

1. $\alpha + \gamma > \beta N$: точка максимума ажиотажа отсутствует, а скорость перехода игнорантов в адепты максимальна при $t = 0$ и с течением времени уменьшается.
2. $\alpha + \gamma = \beta N$: максимум ажиотажа имеет место в начальный момент времени $t = 0$ и с течением времени скорость перехода в адепты уменьшается.
3. $\alpha + \gamma < \beta N$: скорость перехода игнорантов в адепты при $t < t_{cr}$ увеличивается, при $t = t_{cr}$ она максимальна, а при $t > t_{cr}$ она уменьшается.

Рассмотрим более подробно модель распространения информации с учетом забывания в важном с прикладной точки зрения случае сильной пропаганды (α много больше, чем βN). Уравнение (7) может быть решено аналитически, однако ввиду громоздкости точного решения, затрудняющего содержательную трактовку результатов, более эффективным является применение метод пограничных функций теории сингулярных возмущений. В соответствии с этим методом неизвестная функция $X(t)$ представляется в виде суммы регулярного $\bar{X}_0(t)$, $\bar{X}_1(t)$ и погранслоного рядов по малому параметру $1/\alpha$:

$$X(t) = \bar{X}_0(t) + \Pi_0 X(\tau) + \frac{1}{\alpha} [\bar{X}_1(t) + \Pi_1 X(\tau)] + O\left(\frac{1}{\alpha^2}\right). \quad (9)$$

Здесь $\tau = \alpha t$, $\Pi_0 X(\tau)$, $\Pi_1 X(\tau)$ — погранслойные функции такие, что $\lim_{t \rightarrow \infty} \Pi_i X(\tau) = 0$, $i = 0, 1$. Подставляя разложение (9) в уравнение (7) и начальное условие (2), затем приравнивая коэффициенты при соответствующих степенях параметра $1/\alpha$, получим уравнения для нулевого и первого приближения функции $X(t)$. Решая их и подставляя найденные выражения в (11), получим первое приближение решения этого уравнения:

$$X(t) = (1 - \exp(-\alpha t)) X_{st} + \frac{1}{\alpha} [\beta N - \gamma + \exp(-2\alpha t) + \exp(-\alpha t)(-2\beta N + \beta N \alpha t + \gamma - \gamma \alpha t)] + O\left(\left(\frac{1}{\alpha}\right)^2\right).$$

Таким образом, в случае сильной пропаганды в главном приближении имеет место экспоненциальное приближение решения к стационарному значению количества адептов, даваемому формулой (8).

Распространение информации в социуме, состоящем из двух групп

Базовая модель (1), (2) предполагает однородность социума в том смысле, что все индивиды в равной мере получают информацию как от СМИ (а), при межличностной коммуникации каждый индивид в рав-

ной мере получает информацию от всех других индивидов (предпосылка С) (б).

В настоящем разделе рассматривается более сложная ситуация. Именно, заменим предпосылку С на следующие положения:

- социум состоит из двух групп, в различной мере пользующихся средствами массовой информации (в частности, одна из групп может вообще не получать информацию из СМИ),
- распространение информации путем межличностной коммуникации может проходить внутри каждой из групп, вообще говоря, более (или, наоборот, менее) интенсивно, чем между группами,
- переходы индивидов из одной группы в другую отсутствуют.

Обозначим количество adeptов в i -й группе в момент времени t через $X_i(t)$. Модель имеет вид следующей системы уравнений:

$$\frac{dX_1}{dt} = (\alpha_1 + \beta_g X_1 + \beta_m X_2)(N_1 - X_1), \quad (10)$$

$$\frac{dX_2}{dt} = (\alpha_2 + \beta_m X_1 + \beta_g X_2)(N_2 - X_2). \quad (11)$$

Здесь параметры α_1 , α_2 характеризуют интенсивность получения информации от СМИ индивидами из первой и второй группы соответственно, β_g , β_m — интенсивности распространения информации при межличностной коммуникации внутри групп и между группами.

Частный случай для модели настоящего пункта представляет ситуация, когда одна из групп не пользуется средствами массовой информации. Тогда игноранты первой группы получают информацию как от СМИ, так и от adeptов обеих групп, игноранты второй подгруппы — только от adeptов. В соответствии с этим в уравнении (11) имеем $\alpha_2 = 0$.

Анализ модели проведем в общем случае $\alpha_2 \geq 0$. Система (10), (11) имеет три положения равновесия, одно из которых (N_1, N_2) находится в первой четверти. Чтобы исследовать его на устойчивость, линеаризуем систему в окрестности данного положения равновесия. После замены переменных $\xi_1 = X_1 - N_1$, $\xi_2 = X_2 - N_2$ получим

$$\frac{d\xi_1}{dt} = -\xi_1(\alpha_1 + \beta_g N_1 + \beta_m N_2), \quad (12)$$

$$\frac{d\xi_2}{dt} = -\xi_2(\alpha_2 + \beta_m N_1 + \beta_g N_2). \quad (13)$$

Матрица этой системы,

$$\begin{pmatrix} -(\alpha_1 + \beta_g N_1 + \beta_m N_2) & 0 \\ 0 & -(\alpha_2 + \beta_m N_1 + \beta_g N_2) \end{pmatrix},$$

имеет собственные значения $\lambda_1 = -(\alpha_1 + \beta_g N_1 + \beta_m N_2)$ и $\lambda_2 = -(\alpha_2 + \beta_m N_1 + \beta_g N_2)$. Оба они отрицательны, следовательно, положение равновесия является устойчивым узлом.

Таким образом, с течением времени все индивиды обеих групп становятся адептами.

Вид фазового портрета зависит от соотношения между величинами λ_1 и λ_2 : если $|\lambda_1| > |\lambda_2|$, то фазовый портрет имеет вид, представленный на рис. 2, а, если же $|\lambda_1| < |\lambda_2|$, тогда — вид, представленный на рис. 2, б.

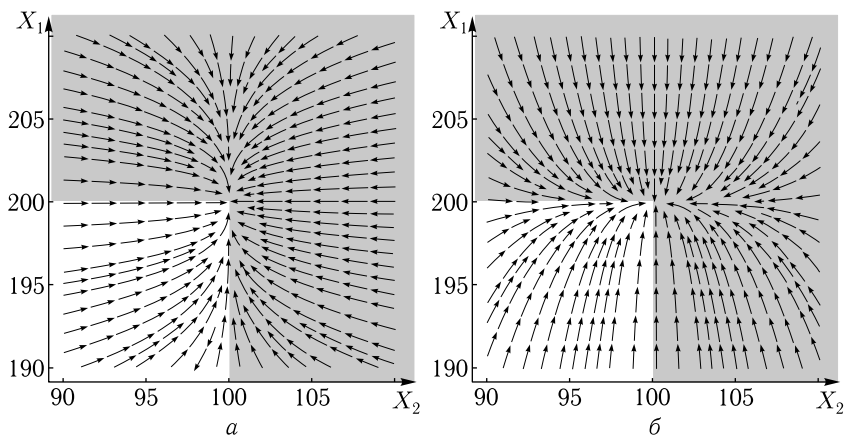


Рис. 2. Фазовый портрет системы (12)–(13). Область изменения переменных выделена белым цветом. Значения параметров на а: $N_1 = 100$; $N_2 = 200$; $\alpha_1 = 0,1$; $\alpha_2 = 0,5$; $\beta_1 = 0,9$; $\beta_2 = 0,3$. Значения параметров на б: $N_1 = 1000$; $N_2 = 200$; $\alpha_1 = 0,5$; $\alpha_2 = 0,1$; $\beta_1 = 0,3$; $\beta_2 = 0,9$

4.5.3. Модель информационного нападения, учитывающая три дополнительных фактора

Данный пункт посвящен построению и анализу модели, учитывающей все три рассмотренных выше фактора. Мы будем называть ее расширенной моделью.

Построение модели

Примем вместо предпосылок А–С следующие положения.

- Социум неоднороден и состоит из двух групп индивидов: одна из них может получать информацию как от СМИ, так и при межличностной коммуникации; другая группа не пользуется СМИ и может получать информацию только при коммуникации с другими индивидами.
- Игнорант становится адептом за два шага. После первого получения информации он становится предадептом (не распространяет информацию далее), после второго охватывания присоединяется к категории адептов. Преадепт получает информацию точно так же, как и неохваченный индивид.

- Индивиды могут забывать информацию (терять к ней интерес), т. е. переходить из адептов в предадепты, а из предадептов — в игноранты.

Таким образом, модель учитывает в совокупности все три фактора, рассмотренных по отдельности в предыдущем пункте.

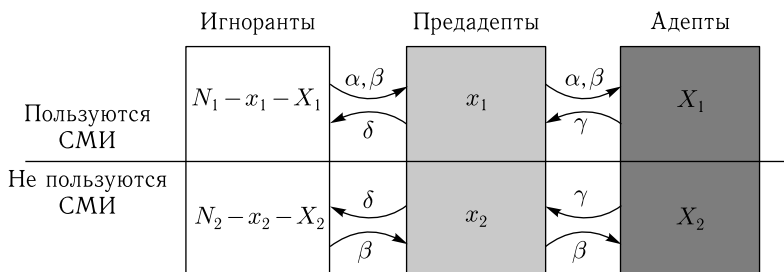


Рис. 3. Схема распространения информации в модели информационного нападения

Процесс распространения информации в социуме (с учетом данных предпосылок) схематично представлен на рис. 3. Индивиды переходят из игнорантов в предадепты, а затем из предадептов — в адепты со скоростью, зависящей от параметров α и β (в первой группе) или только от β (во второй группе). Преадепты забывают информацию и вновь переходят в группу игнорантов со скоростью, характеризующейся параметром δ . Адепты возвращаются в состояние предадептов со скоростью, описываемой параметром γ .

Основанная на этих предпосылках модель имеет вид

$$\frac{dX_1}{dt} = x_1(\alpha + \beta[X_1 + X_2]) - \gamma X_1, \quad (14)$$

$$\frac{dX_2}{dt} = \beta x_2[X_1 + X_2] - \gamma X_2, \quad (15)$$

$$\frac{dx_1}{dt} = (\alpha + \beta[X_1 + X_2])[N_1 - 2x_1 - X_1] - \delta x_1 + \gamma X_1, \quad (16)$$

$$\frac{dx_2}{dt} = \beta[X_1 + X_2][N_2 - 2x_2 - X_2] - \delta x_2 + \gamma X_2, \quad (17)$$

$$x_1(0) = x_2(0) = X_1(0) = X_2(0) = 0. \quad (18)$$

Здесь X_i — число адептов из i -й группы (здесь и далее $i = 1, 2$); x_i — число предадептов из i -й группы, $i = 1, 2$; N_i — численность i -й группы; α, β — параметры, характеризующие интенсивность распространения информации через СМИ и путем межличностной коммуникации соответственно; δ, γ — параметры, характеризующие интенсивность забывания информации предадептами и адептами соответственно.

Для анализа модели (14) рассмотрим два случая, в которых соотношения между параметрами допускают наиболее простую содержательную трактовку. Сначала будет изучается случай, когда забывание

информации адептами происходит со слабой интенсивностью ($\gamma \ll 1$), а затем будет рассматриваться случай сильной агитации ($\alpha \gg 1$). В обоих случаях будет проводится построение асимптотики решения по малому параметру.

Асимптотика стационарного решения системы уравнений расширенной модели при слабом забывании информации адептами

Рассмотрим стационарную систему, соответствующую системе (14)–(17):

$$x_1 (\alpha + \beta[X_1 + X_2]) - \gamma X_1 = 0, \quad (19)$$

$$\beta x_2 [X_1 + X_2] - \gamma X_2 = 0, \quad (20)$$

$$(\alpha + \beta[X_1 + X_2]) [N_1 - 2x_1 - X_1] - \delta x_1 + \gamma X_1 = 0, \quad (21)$$

$$\beta [X_1 + X_2] [N_2 - 2x_2 - X_2] - \delta x_2 + \gamma X_2 = 0. \quad (22)$$

Слабому забыванию информации адептами соответствует соотношение $\gamma \ll 1$. В этом случае стационарная система (19)–(22) является регулярно возмущенной, и неизвестные величины раскладываются в степенные ряды по малому параметру:

$$x_{1,\text{st}} = x_1^0 + \gamma x_1^1 + o(\gamma), \quad (23)$$

$$x_{2,\text{st}} = x_2^0 + \gamma x_2^1 + o(\gamma), \quad (24)$$

$$X_{1,\text{st}} = X_1^0 + \gamma X_1^1 + o(\gamma), \quad (25)$$

$$X_{2,\text{st}} = X_2^0 + \gamma X_2^1 + o(\gamma). \quad (26)$$

Подставляя разложения (23)–(26) в уравнения (19)–(22) и приравнивая коэффициенты при соответствующих степенях γ , получим уравнения для нулевого и первого приближения каждой из функций. Решая их и подставляя найденные выражения в (23)–(26), получим первое приближение стационарного решения:

$$x_1 = \frac{\gamma N_1}{\alpha + \beta [N_1 + N_2]} + o(\gamma), \quad (27)$$

$$x_2 = \frac{\gamma N_2}{\alpha + \beta [N_1 + N_2]} + o(\gamma), \quad (28)$$

$$X_1 = N_1 \left[1 - \frac{\gamma (\alpha + \delta + \beta [N_1 + N_2])}{(\alpha + \beta [N_1 + N_2])^2} \right] + o(\gamma), \quad (29)$$

$$X_2 = N_2 \left[1 - \frac{\gamma (\delta + \beta [N_1 + N_2])}{(\beta [N_1 + N_2])^2} \right] + o(\gamma). \quad (30)$$

Выписав матрицу Якоби для системы (15)–(18) и найдя ее собственные значения, нетрудно показать, что полученное стационарное решение является асимптотически устойчивым при любых положительных значениях параметров α , β , и δ . (см. упражнение 3).

Асимптотика нестационарного решения системы уравнений расширенной модели информационного нападения при сильной агитации

Рассмотрим теперь нестационарные решения системы (10)–(14) для случая сильной агитации, т. е. параметр α положим достаточно большим. С учетом уравнений (10), (12) «достаточно большое α » означает, что α должно быть много больше, чем $N\beta$, где N — порядок суммарной численности первой и второй группы. При данном предположении становится возможным применение методов теории сингулярных возмущений. В соответствии с методом пограничных функций представим неизвестные функции в виде суммы регулярного и погранслоного рядов по малому параметру $\varepsilon = 1/\alpha$:

$$X_1(t) = \overline{X}_1^0(t) + \Pi_0 X_1(\tau) + \varepsilon \left[\overline{X}_1^1(t) + \Pi_1 X_1(\tau) \right] + o(\varepsilon), \quad (31)$$

$$X_2(t) = \overline{X}_2^0(t) + \Pi_0 X_2(\tau) + \varepsilon \left[\overline{X}_2^1(t) + \Pi_1 X_2(\tau) \right] + o(\varepsilon), \quad (32)$$

$$x_1(t) = \overline{x}_1^0(t) + \Pi_0 x_1(\tau) + \varepsilon \left[\overline{x}_1^1(t) + \Pi_1 x_1(\tau) \right] + o(\varepsilon), \quad (33)$$

$$x_2(t) = \overline{x}_2^0(t) + \Pi_0 x_2(\tau) + \varepsilon \left[\overline{x}_2^1(t) + \Pi_1 x_2(\tau) \right] + o(\varepsilon). \quad (34)$$

Здесь $\tau = t/\varepsilon$, $\overline{X}_i^j(t)$, $\overline{x}_i^j(t)$ — регулярные члены асимптотики, $\Pi_j X_i(\tau)$, $\Pi_j x_i(\tau)$ — погранслоные функции, экспоненциально убывающие при $\tau \rightarrow \infty$, $i = 0, 1$.

Подставляя разложения (28)–(31) в уравнения (10)–(14) и приравнявая коэффициенты при соответствующих степенях ε , получим уравнения относительно нулевого и первого приближения функций $X_1(t)$ и $x_1(t)$. Решая их и подставляя найденные выражения в (28)–(31), получим первое приближение этой системы:

$$X_1(t) = N_1 - N_1 e^{-t/\varepsilon} \left(1 + \frac{t}{\varepsilon} \right) + \varepsilon N_1 \left[-\gamma + \frac{1}{6} e^{-t/\varepsilon} \left(-\delta \left(\frac{t}{\varepsilon} \right)^3 + \left(\frac{t}{\varepsilon} \right)^2 (6\beta N_1 (e^{-t/\varepsilon} + 1) + 3\gamma) + \frac{2t}{\varepsilon} (6\beta N_1 (e^{-t/\varepsilon} - 1) + 3\gamma) + 6\gamma \right) \right], \quad (35)$$

$$x_1(t) = N_1 e^{-t/\varepsilon} \frac{t}{\varepsilon} + \varepsilon N_1 \left[\gamma + \frac{1}{6} e^{-t/\varepsilon} \left(\delta \left(\frac{t}{\varepsilon} \right)^3 - \left(\frac{t}{\varepsilon} \right)^2 (3\delta 6\beta N_1 (e^{-t/\varepsilon} + 1) + 3\gamma) - \frac{6t}{\varepsilon} (\beta N_1 (e^{-t/\varepsilon} - 3) + \gamma) - 6\gamma - 12\beta N_1 (1 - e^{t/\varepsilon}) \right) \right]. \quad (36)$$

При больших значениях α данное приближение достаточно точно описывает динамику численности адептов и предадептов. На рис. 4

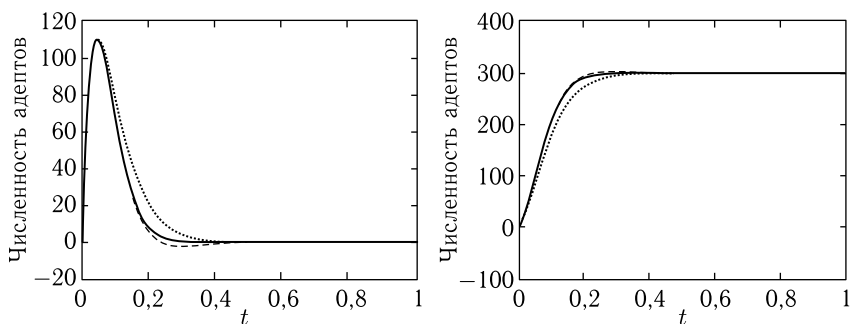


Рис. 4. Сравнение численного и аналитического решения системы (14)–(17) с параметрами (37): толстая сплошная линия — решение, полученное численно, пунктирная — нулевое приближение, тонкая сплошная — первое приближение (для численности адептов первая и третья кривые визуально сливаются)

представлено сравнение результатов, полученных при численном и аналитическом решении, для следующих значений параметров:

$$N_1 = 300, N_2 = 100, \gamma = 0,02, \delta = 0,01, \alpha = 20, \beta = 0,03. \quad (37)$$

Максимальное значение модуля разности между численным решением и асимптотикой первого приближения в данном случае составило 3,2 для функции $X_1(t)$ и 4,3 для функции $x_1(t)$.

Сравнение левого и правого графиков на рис. 4 показывает, что этап быстрого возрастания численности адептов более продолжителен, чем у предадептов. Это объясняется тем, что индивиды переходят в адепты из предадептов. При этом максимальная скорость роста адептов наблюдается в момент достижения максимума численности предадептов.

Построить аналогичные асимптотики для переменных $x_2(t)$, $X_2(t)$ не удастся ввиду сложности системы уравнений, поэтому приведем лишь графики численного решения (рис. 5). Они показывают, что информация охватывает первую группу быстрее, чем вторую. Очевидно, это связано с тем, что первая группа пользуется СМИ.

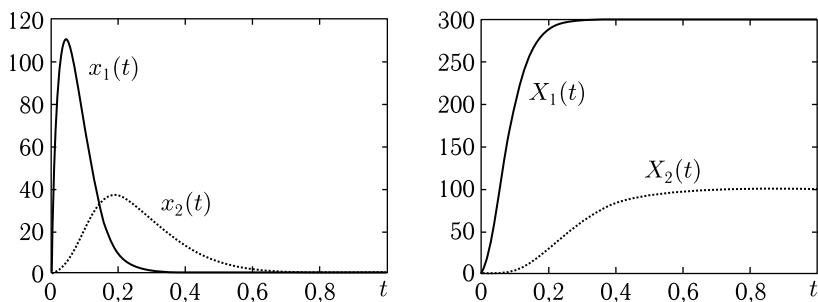


Рис. 5. Численное решение системы (14)–(17) с параметрами (37)

4.5.4. Заключение

В настоящем разделе рассмотрен ряд моделей распространения информации в отсутствие противодействия, учитывающих как передачу информации при межличностной коммуникации, так и пропаганду через СМИ. Разработаны способы математического описания таких факторов распространения информации, как неполный охват социума СМИ массовой информации, усвоение информации лишь после неоднократного ее получения, забывания информации индивидами. На основе базовой модели построена математическая модель информационного противоборства в социуме, учитывающая перечисленные факторы. Результатам математического исследования дана содержательная трактовка.

Упражнения

1. Доказать, что выполнение неравенства (2) является необходимым и достаточным условием существования момента времени $t_{cr} > 0$, при котором имеет место максимум ажиотажа.

2. Для уравнения (9) из раздела «Распространение информации с учетом забывания» провести анализ существования точки максимума ажиотажа, аналогично тому, как это сделано для базовой модели в первом пункте.

3. Построить асимптотику первого приближения (27)–(30) стационарного решения системы (14)–(17) в случае $\gamma \ll 1$. Доказать, что это стационарное решение является асимптотически устойчивым.

Комментарии и библиографические примечания

1. Базовая модель информационного нападения предложена в работах [1, 2]. Дополнительные факторы введены в [3] и исследованы в [3, 4].

2. Метод пограничных функций А. Б. Васильевой, использованный при анализе моделей данного раздела, основан на ранних результатах А. Н. Тихонова [5–7], изложен в ряде монографий, из которых для первого чтения может быть рекомендована книга [8].

3. Обзор социологической литературы по вопросам влияния информации на общество, распространения информации и т. д. мог бы стать самостоятельным многотомным изданием, здесь мы отметим лишь работы [9–11], ни в какой мере не претендуя на полноту либо репрезентативность обзора.

Первые математические модели распространения одного слуха были предложены довольно давно [12, 13]. В самых общих чертах в этих моделях предполагается, что в каждый момент времени некоторые индивиды из числа образующих социальную группу обладают определенной информацией и передают ее другим индивидам. Тем самым происхо-

дит распространение этой информации. Различают детерминированные и стохастические модели.

Механика модели работы [12] выглядит следующим образом. В каждый момент времени, каждый член социума относится к одному из трех категорий: игноранты, спредеры (распространители, аналог понятия «адепты» в моделях настоящего раздела), стифлеры (дословно, *stiflers* — это «подавители»). Игноранты еще не знакомы со слухом, спредеры знают слух и распространяют его, стифлеры знают, но не распространяют. Изначально один член социума является спредером, все остальные — игнорантами. Частота встреч игнорантов и спредеров пропорциональна произведению текущих численностей этих категорий (это предполагает однородность социума). Переходы индивидов из одной категории в другую происходят в трех случаях: если игнорант встречается со спредером, то он тоже становится спредером (i), если встречаются два спредера, то оба они становятся стифлерами (ii), если спредер встретил стифлера, то он тоже становится стифлером (iii). Далее мы будем называть переход спредеров в категорию стифлеров стифлинг-эффектом.

Отличие предпосылок модели [13] состоит в том, что при взаимодействии двух спредеров только один из них превращается в стифлера (а второй остается спредером), т. е. стифлинг-эффект носит менее интенсивный характер.

Эти ранние работы развивались под влиянием моделей эпидемиологии (впрочем, это влияние остается весьма сильным до сих пор). Уже когда в работе [12] вводились понятия игнорантов, спредеров и стифлеров, то ее авторы указывали, что эти категории аналогичны категориям «подверженные», «инфицированные» и «отболевшие» в эпидемиологии. В то же время речь идет не только о перенесении идей из одной научной отрасли в другую, но также о поиске различий. Так, те же авторы видят различие в том, что, моделях эпидемиологии переход из категории «инфицированные» в категорию «отболевшие» (т. е. выздоровление) происходит просто с течением времени, без влияния соседей по социуму; в то время как в разработанной ими модели слуха стифлинг-эффект происходит вследствие встречи с другим спредером или стифлером.

Основной вопрос этих двух моделей состоит в том, чтобы определить количество игнорантов в конце процесса. Ответ имеет следующий вид: если начальное количество спредеров равно одному, то конечное (формально: при бесконечно большом времени) количество игнорантов в обеих моделях равно примерно 0,203 от общей численности. В работах [14, 15] для модели [12] изучался также вопрос о финальном количестве игнорантов в случае, если начальное количество спредеров больше единицы. Показано, в частности, что при достаточно большом количестве спредеров в начальный момент категория игнорантов в конце процесса составляет примерно $1/e$ от своей начальной численности.

Социологический смысл данного результата можно пояснить следующим образом. Повседневный опыт говорит, что когда циркуляция слуха в обществе заканчивается, то в нем, как правило, остаются довольно многочисленные игноранты. Почему так происходит? Почему слух не охватывает все общество? Модели работ [12, 13] объясняют это тем, что все спредеры (ввиду коммуникаций между собой и со стифлерами, сами становятся стифлерами) не успев передать слух оставшимся игнорантам.

На наш взгляд, данное объяснение является неудовлетворительным, поскольку оно объясняет прекращение циркуляции слуха лишь системой коммуникаций между индивидами, полностью игнорируя характер информационного сообщения, которое передается посредством слуха. Другими словами, каждый из множества слухов должен охватить одну и ту же долю населения! На самом деле даже повседневный опыт свидетельствует о том, что некоторые слухи захватывают более широкую публику (и распространяются гораздо быстрее), чем другие. Более того: некоторые слухи могут охватывать практически все взрослое население какой-нибудь страны, тогда как многие информационные сообщения прекращают циркуляцию уже после второго или третьего спредера, т. е. фактически даже не становятся слухами в обыденном смысле этого слова. Это говорит о том, что влияние стифлинг-эффекта на распространение информации в социуме (на самом деле) довольно мало и им можно пренебречь. По этой причине рассматриваемые в данном разделе модели не содержат стифлинг-эффекта.

При этом нельзя не признать, что работы [12, 13] породили довольно крупное научное направление. Оно включает сегодня обширную библиографию и развивается, в основном, как раздел математики, в определенном отрыве от социальных наук.

Укажем совсем кратко некоторые другие направления в данной области. Изучению процессов распространения информации в социальных сетях посвящены многочисленные работы — в качестве примера можно указать [16, 17]. Довольно редкий пример моделирования распространения слухов с опорой на конкретный социальный механизм представляет статья [18]. В ней построена модель распространения информации, акцентированная на механизме «узнал на работе — рассказал в семье, узнал в семье — рассказал на работе». Модели конкурирующих слухов известны гораздо меньше, хотя появились также довольно давно, по крайней мере в работе 1977 г. [19]. Среди современных работ можно указать, например [20, 21].

Перечисленные выше работы посвящены распространению слуха (в данной области науки это слово не имеет оттенка «недостоверная информация»). Другими словами, в этих моделях информационное сообщение передается только от человека к человеку, т. е. исключительно путем межличностных коммуникаций.

В реальной жизни, во многих случаях существенную роль в распространении информационных потоков играют также масс-медиа.

В качестве эмпирического примера, демонстрирующего сочетание распространения информации через масс-медиа и через слухи, можно указать работу [22], посвященную исследованию влияния радиостанции RTLM, пропагандирующей насилие, на геноцид в Руанде в 1994 г. В ней было показано, что наиболее интенсивные этнические чистки происходили в деревнях (принимавших призывы к ненависти от радиостанции), а также в соседних с ними деревнях (получавших те же призывы в пересказе своих соседей). Отсюда был сделан вывод о том, что при изучении распространения информации в социуме нельзя пренебречь ни вещанием СМИ, ни межличностной коммуникацией.

4. Приведем обоснование уравнения, лежащего в основе модели информационного нападения, через описание процессов на уровне отдельных индивидов [4]. Для этого рассмотрим некоторого игноранта в некоторый момент времени t_0 и предположим, что он может получить информацию лишь от одного источника. Если этот источник действует с постоянной интенсивностью, то для этого индивида вероятность остаться игнорантом в момент времени $t > t_0$ составляет

$$q(t) = \exp[-\lambda(t - t_0)], \quad (4)$$

где $\lambda > 0$ — некоторый параметр. Определим его содержательный смысл, имея в виду необходимость связать его с параметрами α и β . Из формулы (6) получаем: $q(t_0 + 1) = \exp(-\lambda)$, $-\lambda = \ln q(t_0 + 1)$. Таким образом, $(-\lambda)$ — это логарифм вероятности того, что индивид, являющийся игнорантом «сейчас» (т. е. в момент времени t_0), останется игнорантом в ближайшую единицу времени, т. е.

$$\lambda = \ln [1/q(t_0 + 1)]. \quad (5)$$

Обозначим через P_α вероятность того, что СМИ охватят данного игноранта в течение одной единицы времени (эта величина является характеристикой интенсивности вещания СМИ, и наша цель состоит в том, чтобы связать ее с параметром α). Аналогично, обозначим через P_β вероятность того, что некоторый конкретный адепт сообщит информацию данному игноранту в течение одной единицы времени. Тогда вероятность остаться неохваченным информацией в течение единицы времени равна

$$q(t_0 + 1) = (1 - P_\alpha)(1 - P_\beta)^{X(t_0)} \quad (6)$$

(здесь учтено, что взаимодействия в социуме однородны и количество адептов в момент t_0 равно $X(t_0)$). Подставив (6) в (5), получаем

$$\lambda = \ln \frac{1}{(1 - P_\alpha)} + X(t_0) \ln \frac{1}{(1 - P_\beta)}.$$

Теперь перейдем непосредственно к выводу уравнения (1) базовой модели информационного нападения. Рассмотрим опять индивида, который в момент t_0 является игнорантом, и обозначим через $p(t)$ вероятность того, что в момент времени $t > t_0$ он будет адептом:

$$p(t) = 1 - q(t) = 1 - \exp[-\lambda(t - t_0)]. \quad (7)$$

Положим $t = t_0 + dt$, получим

$$p(t_0 + dt) = 1 - \exp(-\lambda dt) = \lambda dt + O((\lambda dt)^2), \quad (8)$$

т. е.

$$p(t_0 + dt) = \left(\ln \frac{1}{1 - P_\alpha} + X(t_0) \ln \frac{1}{1 - P_\beta} \right) dt + O((\lambda dt)^2).$$

Следовательно, математическое ожидание прироста количества адептов за время dt равно

$$\begin{aligned} X(t_0 + dt) - X(t_0) &= \\ &= \left[\left(\ln \frac{1}{1 - P_\alpha} + X(t_0) \ln \frac{1}{1 - P_\beta} \right) dt + O((\lambda dt)^2) \right] (N - X(t_0)). \end{aligned}$$

Поделив левую и правую части уравнения на dt и совершив предельный переход $dt \rightarrow 0$, получим

$$\frac{dX}{dt} = \left(\ln \frac{1}{1 - P_\alpha} + X \ln \frac{1}{1 - P_\beta} \right) (N - X).$$

Вводя обозначения

$$\ln \frac{1}{1 - P_\alpha} = \alpha, \quad \ln \frac{1}{1 - P_\beta} = \beta,$$

получаем уравнение (1).

Заметим, что при малых P_α , P_β имеют место приближенные равенства $P_\alpha \approx \alpha$, $P_\beta \approx \beta$. Таким образом, если интенсивность вещания достаточно слаба (так, что величина P_α является малой), то параметр α имеет смысл вероятности того, что игнорант будет охвачен информацией от СМИ за единицу времени. Аналогично, если интенсивность передачи информации путем межличностной коммуникации достаточно мала, то параметр β имеет смысл вероятности того, что игнорант будет охвачен информацией от некоторого адепта (любого конкретного адепта) за единицу времени.

Тем самым проведено обоснование базовой модели информационного нападения через описание процессов на уровне отдельных индивидов, а также показан содержательный смысл коэффициентов α , β .

Литература

1. Самарский А.А., Михайлов А.П. Математическое моделирование. Идеи, методы, примеры. М.: Наука, 2001. 320 с.
2. Михайлов А.П., Ключов Н.В. О свойствах простейшей математической модели распространения информационной угрозы, Математическое моделирование социальных процессов, вып. 4. М.: МАКС Пресс, 2002. С. 115–123.
3. Михайлов А.П., Петров А.П., Маревцева Н.А., Третьякова И.В. Развитие модели распространения информации в социуме // Математическое моделирование. 2014. Т. 26, № 3. С. 65–74. (англ. перевод: *Mikhailov A.P., Petrov A.P., Marevtseva N.A., Tretiakova I.V.* Development of a Model of Information Dissemination in Society // *Mathematical Models and Computer Simulations*, 2014. V. 6, № 5, pp. 535–541, 2014.).
4. Петров А.П., Прончева О.Г. Исследование моделей информационного нападения и информационного противоборства в структурированном социуме // Математическое моделирование социальных процессов. Вып. 17. Под ред. А.П. Михайлова — М.: Эконинформ, 2015. С. 136–149.
5. Тихонов А.Н. О зависимости решений дифференциальных уравнений от малого параметра. Матем. сб., 1948, 22(64).
6. Тихонов А.Н. О системах дифференциальных уравнений, содержащих параметры. — Матем. сб., 1948, 27(69).
7. Тихонов А.Н. Системы дифференциальных уравнений, содержащих малые параметры при производных. Матем. сб., 1952, 31(73).
8. Васильева А., Бутузов В. Асимптотические методы в теории сингулярных возмущений. — Высшая школа Москва, 1990. 208 с.
9. Herman E., Chomsky N. A propaganda model. *Media and Cultural Studies*. 2006.
10. DiFonzo N., Bordia P. Rumor, gossip and urban legends. *Diogenes*. 2007. 54(1), 19–35.
11. Pronchev G.B., Muraviev V.I. Социальные сети как фактор перехода России к инновационному развитию // *Социология*. № 3. 2011. С. 36–56.
12. Daley D.J., Kendall D.G. Stochastic Rumors // *Journal of the Institute of Mathematics and its Applications*. 1964. V. 1. P. 42–55.
13. Maki D.P., Thompson M. *Mathematical Models and Applications*. Prentice-Hall. Englewood Cliffs. 1973.
14. Belen S. The behaviour of stochastic rumours. PhD Thesis, The University of Adelaide. 2008.
15. Belen S., Pearce C.E.M. Rumours with general initial conditions. *ANZIAM J.*, 2004. 4, 393–400.
16. Nekovee M., Moreno Y., Bianconi G., Marsili M. Theory of Rumor Spreading in Complex Social Networks. *Physica A*. 2007. 374. P. 457–470.
17. Губанов Д.А., Новиков Д.А., Чхартишвили А.Г. Социальные сети: модели информационного влияния, управления и противоборства. — М.: ФИЗМАТЛИТ, 2010, 228 с.
18. Шведовский В.А. Моделирование распространения информации в смежных социальных группах // *Математические методы в социологическом исследовании*, Наука, М., 1981. С. 207–214.
19. Osei G.K., Thompson J.W. The supersession of one rumour by another. *J. of Applied Probability*. 1977. 14(01). P. 127–134.

20. Kaligotla C., Yücesan E., Chick S.E. An agent based model of spread of competing rumors through online interactions on social media. In Proceedings of the 2015 Winter Simulation Conference (3985-3996). IEEE Press. 2015.
21. Escalante R., Odehnal M. A deterministic mathematical model for the spread of two rumors. arXiv:1709.01726 [physics.soc-ph].
22. Yanagizawa-Drott D. Propaganda and Conflict: Evidence from the Rwandan Genocide. The Quarterly Journal of Economics, 129(4). P. 1947–1994. 2014. doi: 10.1093/qje/qju020.

4.6. Информационное противоборство в социуме

В предыдущем разделе рассмотрен ряд моделей распространения информации в социуме в отсутствие противодействия. Здесь рассматриваются построенные на их основе модели информационного противоборства.

Изучаемый процесс состоит в следующем. Противоборствующие стороны (партии) распространяют в социуме информационные потоки с помощью аффилированных СМИ. Если индивид получает информацию одной из партий, то он становится ее сторонником (адептом) и начинает распространять эту информацию дальше при межличностной коммуникации с другими членами социума; при этом он становится невосприимчивым к пропаганде другой стороны. Тем самым изначально нейтральные индивиды становятся сторонниками противоборствующих партий. Информационные потоки являются антагонистическими в том смысле, что каждый индивид может быть сторонником, т. е. распространителем информации не более, чем одной из партий. Основной вопрос состоит в том, какая из сторон одержит победу в противоборстве, получив большинство сторонников к определенному моменту времени (например, при t , стремящемся к бесконечности).

С математической точки зрения модели настоящего раздела представляют собой уравнения и системы обыкновенных дифференциальных уравнений. Как и выше, рассматривается ряд моделей: от более простых к более сложным. В первом пункте рассматривается базовая модель, описывающая противоборство двух сторон в неструктурированном социуме. Выводится так называемое *условие победы*, т. е. неравенство для параметров модели, выполнение которого необходимо и достаточно для победы одной из партий; а выполнение обратного неравенства — для победы другой партии. Это условие обобщается во втором пункте на случай противоборства произвольного количества сторон. Далее, рассматривается модель, учитывающая дополнительные факторы, введенные в предыдущем разделе для описания распространения информации. Данная модель имеет вид задачи Коши для нелинейной системы из восьми обыкновенных дифференциальных уравнений. Проведены серии вычислительных экспериментов, при этом наибольшее внимание уделялось случаю, когда одна из сторон имеет более интенсивно работающее СМИ, но информация другой стороны

более интенсивно распространяется при межличностной коммуникации. В четвертом пункте рассмотрена модель с дестабилизирующим воздействием, под которым понимается краткосрочное периодическое увеличение интенсивности медийного вещания одной из партий. Краткое заключение приведено в пятом пункте.

4.6.1. Базовая модель информационного противоборства

Рассмотрим социальную общность, состоящую из N индивидов. Две стороны противоборства распространяют через аффилированные СМИ антагонистические потоки информации. Индивиды, принявшие информацию определенной партии, называются ее адептами или сторонниками. Адепт распространяет информацию своей партии среди других индивидов и является невосприимчивым к информации противоположной стороны. Победителем противоборства считается сторона, имеющая большее количество адептов к определенному моменту времени (например, при t , стремящемся к бесконечности).

Численности сторонников двух противоборствующих партий в момент времени t обозначим через $X(t)$, $Y(t)$, численность нейтралов, т.е. неохваченных партиями индивидов тогда составит $N - X(t) - Y(t)$. Для определенности будем полагать, что в начальный момент времени $t = 0$ никто из членов социума не принадлежит ни к одной из сторон; другие начальные условия могут быть рассмотрены аналогичным образом. Информация распространяется двумя способами: через СМИ и путем межличностной коммуникации с людьми, охваченными ею ранее. Интенсивность распространения информации этими способами описывается параметрами α_x, β_x (для партии X), α_y, β_y (для партии Y). Заметим, что интенсивность распространения информации через межличностную коммуникацию при этом пропорциональна также числу сторонников. Далее, предполагается, что скорость распространения информации (т.е., скорость увеличения количества адептов некоторой конкретной партии) складывается из скоростей распространения информации через СМИ и коммуникацию. Эта скорость пропорциональна численности нейтралов, т.е. $N - X(t) - Y(t)$. Приведенные рассуждения позволяют выписать базовую модель информационного противоборства, имеющую следующий вид:

$$\frac{dX}{dt} = (\alpha_x + \beta_x X)(N - X - Y), \quad (1)$$

$$\frac{dY}{dt} = (\alpha_y + \beta_y Y)(N - X - Y), \quad (2)$$

$$X(0) = X^0, \quad Y(0) = Y^0. \quad (3)$$

Перечислим предположения, учтенные при получении данной модели.

Предпосылка А. Однажды получив информацию некоторой партии, нейтральный индивид становится ее сторонником и немедленно начинает распространять эту информацию.

Предпосылка В. Обратного перехода из сторонников партий в нейтралы не происходит, как и перехода индивидов из одной партии в другую.

Предпосылка С. Социум является однородным в том смысле, что все индивиды в равной мере получают информацию от СМИ (а), каждый индивид в равной мере получает информацию от всех других индивидов (б).

Предпосылка Д. Параметры $\alpha_x, \alpha_y, \beta_x, \beta_y$, описывающие интенсивность распространения информации партий через СМИ и межличностную коммуникацию, суть постоянные величины.

Основной целью анализа модели (1)–(3) является получение условия победы, т. е. соотношения между параметрами определяющего партию-победителя противоборства (при этом будем полагать, что противоборство длится достаточно долго, т. е. определяется победитель при $t \rightarrow \infty$). Чтобы получить его, поделим (1) на (2):

$$\frac{dX}{dY} = \frac{\alpha_x + \beta_x X}{\alpha_y + \beta_y Y}. \quad (4)$$

Отсюда

$$\beta_y d(\ln(\alpha_x + \beta_x X)) - \beta_x d(\ln(\alpha_y + \beta_y Y)) = 0,$$

т. е.

$$\left(\frac{\alpha_x + \beta_x X(t)}{\alpha_x + \beta_x X(0)} \right)^{\beta_y} = \left(\frac{\alpha_y + \beta_y Y(t)}{\alpha_y + \beta_y Y(0)} \right)^{\beta_x}. \quad (5)$$

Противоборство заканчивается вничью, если $X(t) \rightarrow N/2, Y(t) \rightarrow N/2$ при $t \rightarrow \infty$. Партия X одерживает победу, если при достаточно больших t $X(t) > N/2, Y(t) < N/2$, партия Y — в обратном случае. Следовательно, соотношение между параметрами, соответствующее победе партии X , имеет вид

$$\frac{\beta_x}{\ln \frac{1 + \beta_x N / (2\alpha_x)}{1 + \beta_x X(0) / (2\alpha_x)}} > \frac{\beta_y}{\ln \frac{1 + \beta_y N / (2\alpha_y)}{1 + \beta_y Y(0) / (2\alpha_y)}}, \quad (6)$$

а соотношение с противоположным знаком означает победу партии Y . В частности, если начальные условия имеют вид

$$X(0) = 0, \quad Y(0) = 0, \quad (7)$$

то неравенство (6) принимает вид

$$\frac{\beta_x}{\ln [1 + \beta_x N / (2\alpha_x)]} > \frac{\beta_y}{\ln [1 + \beta_y N / (2\alpha_y)]}. \quad (8)$$

Неравенство (6) (а также его частный случай (8)) называется условием победы.

Ограничимся далее анализом случая (7), т. е. начальными условиями, соответствующими ситуации, в которой начальные численности обеих партий равны нулю.

Прежде всего заметим, что при одновременном выполнении условия (8) и условия $\alpha_x > \alpha_y$ победа партии X достигается в режиме постоянного доминирования, т. е. $X(t) > Y(t)$ при всех $t > 0$. Если же неравенство (8) выполнено, но $\alpha_x < \alpha_y$, то противоборство проходит со «сменой лидера», т. е. $X(t) < Y(t)$ при $0 < t < t_c$ и $X(t) > Y(t)$ при $t > t_c$ (задача нахождения величины t_c вынесена в упражнения).

Если $\beta_x = \beta_y$, то неравенство (8) эквивалентно неравенству $\alpha_x > \alpha_y$. Это имеет очевидный смысл: если интенсивности передачи конкурирующих сообщений при межличностной коммуникации равны, то побеждает партия, имеющая более сильное вещание СМИ. Аналогично, если $\alpha_x = \alpha_y$, то неравенство (8) эквивалентно неравенству $\beta_x > \beta_y$.

Чтобы рассмотреть нетривиальную ситуацию, в которой имеется относительно простое явное решение, рассмотрим случай, когда информация партии Y распространяется через межличностные коммуникации в два раза более интенсивно, т. е. $\beta_y = 2\beta_x$. Условие (8) для победы партии X принимает в данной ситуации вид

$$\alpha_y < 4\alpha_x^2 / (4\alpha_x + \beta_x N). \quad (9)$$

В частности, если $\beta_x N \ll \alpha_x$, то оно принимает еще более простой вид $\alpha_x > \alpha_y$. Таким образом, если партия X имеет более интенсивное вещание СМИ, чем Y , и это вещание существенно превосходит по интенсивности передачу информации при межличностной коммуникации, то она побеждает, даже если информация партии Y передается при межличностной коммуникации более интенсивно.

4.6.2. Условие победы при противоборстве произвольного количества партий

Рассмотрим теперь противоборство произвольного количества партий, которое обозначим через M . Базовая модель очевидным образом обобщается на этот случай; она представляет собой систему M нелинейных обыкновенных дифференциальных уравнений

$$\frac{dX_i}{dt} = (\alpha_i + \beta_i X_i) \left(N - \sum_{i=1}^M X_i \right), \quad X_i(0) = X_i^0, \quad 1 \leq i \leq M, \quad t > 0. \quad (10)$$

Здесь X_i — количество сторонников i -й партии, начальные численно-сти X_i^0 удовлетворяют неравенствам $X_i^0 \geq 0$, $\sum_{i=1}^M X_i^0 < N$.

Система (10) имеет решение в рекуррентном виде

$$\beta_{i+1} X_{i+1}(t) = C_i (\alpha_i + \beta_i X_i(t))^{\beta_{i+1}/\beta_i} - \alpha_{i+1}, \quad 1 \leq i \leq M-1, \quad (11)$$

где константы C_i даются формулами

$$C_i = [\alpha_{i+1} + \beta_{i+1} X_i(0)] / [\alpha_i + \beta_i X_i(0)], \quad 1 \leq i \leq M-1.$$

Отсюда находится достаточное условие победы одной из партий. Введем «функцию победы» i -й партии $V_i(\alpha_i, \beta_i, X_i(0), N)$, $1 \leq i \leq M$, следующим образом:

$$V_i(\alpha_i, \beta_i, X_i(0), N) = \beta_i \ln \left[\left(1 + \frac{\beta_i N}{2\alpha_i} \right) / \left(1 + \frac{\beta_i X_i(0)}{\alpha_i} \right) \right]^{-1}. \quad (12)$$

Тогда победителем является партия, имеющая наибольшее значение функции победы (обоснование данного утверждения вынесено в упражнении).

4.6.3. Учет дополнительных факторов

В предыдущем разделе, посвященном распространению информации в отсутствие противоборства, был введен ряд дополнительных факторов, влияющих на распространение информации в социуме, именно — усвоение информации индивидами лишь после двукратного ее получения, забывание (утеря интереса) информации адептами, а также неполный охват социума средствами массовой информации. Перейдем к модели информационного противоборства, учитывающей все эти факторы в совокупности. Она имеет вид

$$\frac{dX_1}{dt} = x_1(\alpha_x + \beta_x(X_1 + X_2)) - \gamma_x X_1, \quad (13)$$

$$\frac{dX_2}{dt} = \beta_x x_2(X_1 + X_2) - \gamma_x X_1, \quad (14)$$

$$\frac{dY_1}{dt} = y_1(\alpha_y + \beta_y(Y_1 + Y_2)) - \gamma_y Y_1, \quad (15)$$

$$\frac{dY_2}{dt} = \beta_y y_2(Y_1 + Y_2) - \gamma_y Y_2, \quad (16)$$

$$\frac{dx_1}{dt} = (\alpha_x + \beta_x(X_1 + X_2))(N_1 - X_1 - Y_1 - 2x_1 - y_1) + \gamma_x X_1 - \delta_x x_1, \quad (17)$$

$$\frac{dx_2}{dt} = \beta_x(X_1 + X_2)(N_2 - X_2 - Y_2 - 2x_2 - y_2) + \gamma_x X_2 - \delta_x x_2, \quad (18)$$

$$\frac{dy_1}{dt} = (\alpha_y + \beta_y(Y_1 + Y_2))(N_1 - X_1 - Y_1 - x_1 - 2y_1) + \gamma_y Y_1 - \delta_y y_1, \quad (19)$$

$$\frac{dy_2}{dt} = \beta_y(Y_1 + Y_2)(N_2 - X_2 - Y_2 - x_2 - 2y_2) + \gamma_y Y_2 - \delta_y y_2, \quad (20)$$

$$X_1(0) = X_2(0) = Y_1(0) = Y_2(0) = x_1(0) = x_2(0) = y_1(0) = y_2(0) = 0. \quad (21)$$

Здесь первая группа состоит из членов социума, пользующихся средствами массовой информации, и имеет численность N_1 , вторая группа не пользуется СМИ и имеет численность N_2 ; переходы индивидов между группами отсутствуют. Нейтральный член социума, получив информацию одной из партий, становится ее предадептом: он еще не распространяет ее информацию, но уже является недоступным для агитации противоположной партии. Если предадепт повторно получает информацию этой партии, то он становится ее адептом (сторонником). Численность адептов партии X в i -й группе ($i = 1, 2$) обозначена через X_i , численность предадептов — через x_i ; аналогично — для

партии Y . При забывании (утере интереса) адепты переходят в состояние предадептов с интенсивностями γ_x, γ_y , а предадепты — в нейтралов с интенсивностями δ_x, δ_y . Таким образом, в данной модели не предполагается выполнение предпосылок А, В, С, сформулированных в первом пункте: каждая из них заменена более сложным положением о процессе информационного противоборства.

Модель (13)–(21) слишком сложна для полноценного аналитического исследования, поэтому для ее изучения более подходит численный эксперимент. Подбор параметров для численного эксперимента следует осуществлять так, чтобы его условия соответствовали осмысленной и представляющей содержательный интерес социальной ситуации. Пример такой социальной ситуации представляет случай, когда одна из сторон имеет более активно пропаганду через СМИ, а информация другой стороны более активно распространяется при межличностной коммуникации: $\alpha_1 > \alpha_2$, $\beta_2 > \beta_1$. Результаты проведенных численных экспериментов показывают, что для каждой из групп динамика имеет, в общих чертах, следующий вид: сначала увеличивается и достигает максимума численность предадептов, затем она уменьшается, при этом происходит рост численности адептов. На рис. 1 и рис. 2 изображена динамика численности адептов и предадептов из обеих групп для следующих значений параметров:

$$N_1 = 100, N_2 = 200, \gamma_x = 0,01, \gamma_y = 0,3, \delta_x = 0,1, \delta_y = 0,02, \\ \alpha_1 = 0,13, \alpha_2 = 0,06, \beta_1 = 0,04, \beta_2 = 0,1,$$

Рис. 1, рис. 2 показывают, что большинство членов первой группы (т. е. индивидов, пользующихся СМИ) с течением времени становятся сторонниками партии X (т. е. $X_1(t) > Y_1(t)$ при достаточно больших t), а большинство членов второй группы (т. е. не пользующихся СМИ) — партии Y (т. е. $X_2(t) < Y_2(t)$ при достаточно больших t). Но если в первой группе преимущество партии X незначительно, то во второй группе разрыв в пользу партии Y намного больше. Это происходит

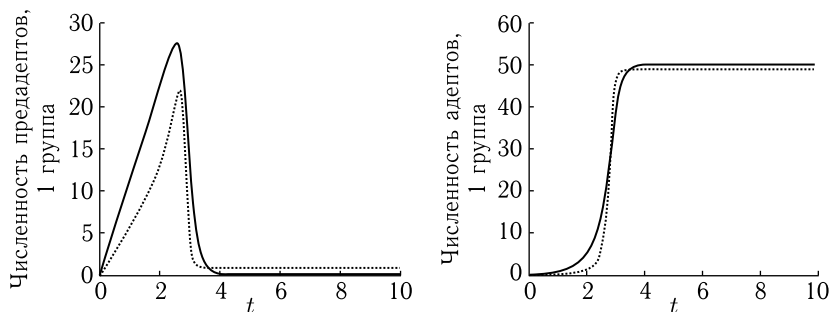


Рис. 1. Динамика численности адептов и предадептов из первой группы. Сплошная линия — $x(t)$ и $X(t)$ соответственно, пунктирная линия — $y(t)$ и $Y(t)$

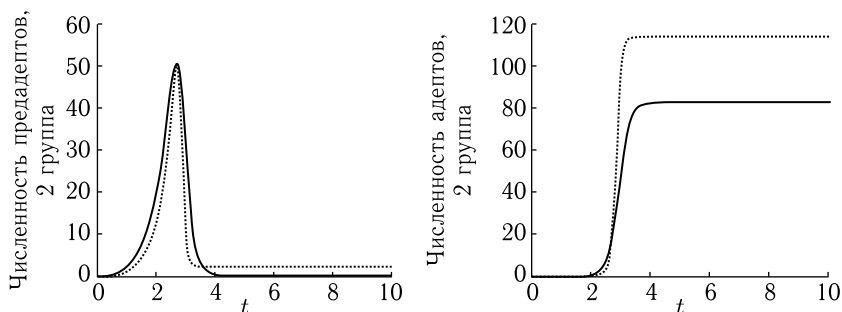


Рис. 2. Динамика численности адептов и предадептов из второй группы. Сплошная линия — $x(t)$ и $X(t)$ соответственно, пунктирная линия — $y(t)$ и $Y(t)$

из-за того, что «вирусность» информации партии Y относится к передаче информации как внутри каждой из групп, так и при коммуникации индивидов из обеих групп, в то время как более интенсивная медийная пропаганда партии X имеет место лишь в пределах первой группы. При увеличении значений параметров α_1 и β_2 имеет место более полный «захват» второй группы партией Y , более сильное преобладание партии X в первой группе. Увеличение коэффициентов забывания информации приводит к сдвигу стационарного распределения в пользу партии Y .

4.6.4. Информационное противоборство при дестабилизирующем воздействии

В изложенных выше моделях интенсивность распространения информации через СМИ предполагалась постоянной, что устанавливается предпосылкой D, сформулированной в первом пункте. Соответственно с течением времени в них формируются стационарные решения, соответствующие постоянным (установившимся) численностям сторонников той и другой партий. В то же время многие реальные информационные кампании характеризуются непостоянством интенсивности СМИ. Усилившаяся пропаганда одной из партий способна дестабилизировать сложившееся соотношение противоборствующих сторон. В настоящем разделе рассматривается случай периодического дестабилизирующего воздействия. Именно, предполагается, что к «фоновой» постоянной интенсивности пропаганды одной из сторон добавляется еще одна, периодическая, компонента. Таким образом, параметр, характеризующий интенсивность распространения информации партии X через СМИ, имеет вид кусочно-постоянной периодической функции времени:

$$\alpha_1(t) = \begin{cases} \alpha_1^*, & t \in [iT; iT + t_{sw}), \\ \alpha_1^* + h, & t \in [iT + t_{sw}; iT + T), \end{cases} \quad i = 0, 1, 2, \dots \quad (22)$$

Здесь $T > 0$ — период, $t_{sw} \in (0; T)$ — продолжительность дестабилизирующего воздействия в течение каждого периода. Наиболее интересным

представляется случай

$$\alpha_1^* < \alpha_2 < \alpha_1^* + h, \quad (23)$$

при котором в «фоновом» режиме более интенсивно через СМИ пропагандируется информация партии Y , а в течение интервалов дестабилизации — партии X . Кроме того, учтем забывание информации индивидами и положим для простоты, что скорость этого забывания, а также интенсивность передачи при межличностной коммуникации равны для информации партий X , Y . Таким образом, модель имеет вид

$$\frac{dX}{dt} = (\alpha_1(t) + \beta X)(N - X - Y) - \gamma X, \quad (24)$$

$$\frac{dY}{dt} = (\alpha_2 + \beta Y)(N - X - Y) - \gamma Y, \quad (25)$$

$$X(0) = 0, \quad Y(0) = 0. \quad (26)$$

Здесь α_2 — постоянная, а функция $\alpha_1(t)$ дается выражением (22).

Нетрудно убедиться, что система (24), (25) не имеет стационарного решения при $\gamma > 0$ (обоснование данного утверждения вынесено в упражнения); при $t \rightarrow \infty$ решение выходит на периодический режим. Аналитическое исследование этого режима проведем в случае, когда интенсивность распространения информации путем межличностной коммуникации мала, т. е. $\beta \ll 1$. Построим асимптотические разложения периодического решения $\tilde{X}(t)$, $\tilde{Y}(t)$ системы уравнений (24), (25) по данному малому параметру. Подчеркнем, что будет построена асимптотика периодического решения, а не задачи Коши (24)–(26).

Поскольку система является регулярно возмущенной, то функции $\tilde{X}(t)$ и $\tilde{Y}(t)$ раскладываются в степенные ряды по параметру β :

$$\begin{aligned} \tilde{X}(t) &= X_0(t) + \beta X_1(t) + \dots + \beta^k X_k(t) + o(\beta^k), \\ \tilde{Y}(t) &= Y_0(t) + \beta Y_1(t) + \dots + \beta^k Y_k(t) + o(\beta^k). \end{aligned} \quad (27)$$

Каждый член асимптотики (27) строится как периодическая функция. Поэтому дифференциальные уравнения для членов асимптотики рассматриваются на одном периоде $[0, T]$ с условиями

$$X_i(0) = X_i(T), \quad Y_i(0) = Y_i(T), \quad i = 0, 1, 2, \dots \quad (28)$$

Далее будем обозначать

$$\tilde{X}(0) = X^0, \quad \tilde{Y}(0) = Y^0, \quad X_i(0) = X_i^0, \quad Y_i(0) = Y_i^0, \quad i = 0, 1, 2, \dots$$

При этом, поскольку функция $\alpha_1(t)$ является кусочно-постоянной (см. формулу (22)), то отдельно рассматриваются первая часть $[0; t_{\text{sw}}]$ периода $[0; T]$ и вторая его часть. На каждой из этих частей функция $\alpha_1(t)$ является постоянной, а состыкование решений происходит по непрерывности:

$$X_i(t_{\text{sw}} - 0) = X_i(t_{\text{sw}} + 0), \quad Y_i(t_{\text{sw}} - 0) = Y_i(t_{\text{sw}} + 0), \quad i = 0, 1, 2, \dots \quad (29)$$

Соотношения вида (28), (29) дополняют уравнения для членов асимптотического разложения (27).

Перейдем к построению нулевого члена асимптотики. Суть применяемого здесь подхода состоит в следующем. На отрезке $[0; t_{sw}]$ решается задача Коши с начальным условием $X_1(0) = X_1^0$, $Y_1(0) = Y_1^0$, где X_1^0 , Y_1^0 — пока неизвестные величины. В результате получаем функции $X_1(t)$, $Y_1(t)$ на данном отрезке и определяем величины $X_1(t_{sw})$, $Y_1(t_{sw})$, зависящие от X_1^0 , Y_1^0 . Далее решается задача Коши на отрезке $[t_{sw}; T]$, и определяются значения $X_1(T)$, $Y_1(T)$, зависящие от $X_1(t_{sw})$, $Y_1(t_{sw})$, тем самым — от X_1^0 , Y_1^0 . Наконец, получаем уравнения для X_1^0 , Y_1^0 из условий периодичности (28). Решив эти уравнения, определяем значения X_1^0 , Y_1^0 . Тем самым строится асимптотика нулевого порядка. Тот же подход может быть применен к построению членов асимптотики (27) более высоких порядков.

Перейдем к реализации изложенного подхода. Подставив разложение (27) в уравнения (24), (25) и рассмотрев слагаемые, не содержащие параметра β , получим уравнения для нулевого приближения. Они имеют вид

$$\begin{cases} \frac{dX_0}{dt} = \alpha_1(N - X_0 - Y_0) - \gamma X_0, \\ \frac{dY_0}{dt} = \alpha_2(N - X_0 - Y_0) - \gamma Y_0. \end{cases} \quad (30)$$

В соответствии с формулой (22) на промежутке $[0; t_{sw})$ имеем $\alpha_1 = \alpha_1^*$. Решив на нем систему (30) с учетом данного соотношения, получаем

$$\begin{pmatrix} X_0(t) \\ Y_0(t) \end{pmatrix} = C_1 \begin{pmatrix} \alpha_1^* \\ \alpha_2 \end{pmatrix} e^{-(\alpha_1^* + \alpha_2 + \gamma)t} + C_2 \begin{pmatrix} 1 \\ -1 \end{pmatrix} e^{-\gamma t} + \frac{N}{\alpha_1^* + \alpha_2 + \gamma} \begin{pmatrix} \alpha_1^* \\ \alpha_2 \end{pmatrix}. \quad (31)$$

Подставив $t = 0$ в (31) и воспользовавшись начальными условиями $X_0(0) = X_0^0$, $Y_0(0) = Y_0^0$, получаем значения постоянных C_1 и C_2 :

$$C_1 = -\frac{N}{\gamma + \alpha_1^* + \alpha_2} + \frac{X_0^0 + Y_0^0}{\alpha_1^* + \alpha_2}, \quad C_2 = \frac{\alpha_2 X_0^0 - \alpha_1^* Y_0^0}{\alpha_1^* + \alpha_2}. \quad (32)$$

Таким образом, на отрезке $[0; t_{sw}]$ асимптотика нулевого приближения имеет вид

$$\begin{pmatrix} X_0(t) \\ Y_0(t) \end{pmatrix} = \left[\frac{N}{\gamma + \alpha_1^* + \alpha_2} + \frac{X_0^0 + Y_0^0}{\alpha_1^* + \alpha_2} \right] \begin{pmatrix} \alpha_1^* \\ \alpha_2 \end{pmatrix} \exp[-(\alpha_1^* + \alpha_2 + \gamma)t] + \frac{\alpha_2 X_0^0 - \alpha_1^* Y_0^0}{\alpha_1^* + \alpha_2} \begin{pmatrix} 1 \\ -1 \end{pmatrix} \exp(-\gamma t) + \frac{N}{\alpha_1^* + \alpha_2 + \gamma} \begin{pmatrix} \alpha_1^* \\ \alpha_2 \end{pmatrix}. \quad (33)$$

На полуинтервале $[t_{sw}; T)$ уравнения для функций $X_0(t)$, $Y_0(t)$ получаются путем подстановки $\alpha_1 = \alpha_1^* + h$ в уравнения (30). Эти уравнения дополняются (ввиду условий непрерывности (29)) начальным

условием, получаемым путем подстановки $t = t_{\text{sw}}$ в (33). Решение получаемой таким образом задачи Коши имеет вид

$$\begin{pmatrix} X_0(t) \\ Y_0(t) \end{pmatrix} = C_1^* \begin{pmatrix} \alpha_1^* + h \\ \alpha_2 \end{pmatrix} \exp[-(\alpha_1^* + h + \alpha_2 + \gamma)t] + \\ + C_2^* \begin{pmatrix} 1 \\ -1 \end{pmatrix} \exp[-\gamma t] + \frac{N}{\alpha_1^* + h + \alpha_2 + \gamma} \begin{pmatrix} \alpha_1^* + h \\ \alpha_2 \end{pmatrix}, \quad (34)$$

где

$$\begin{aligned} C_1^* = & \frac{2\alpha_2 e^{(\alpha_1^* + \alpha_2 + h)t_{\text{sw}}} + (\alpha_1^* - \alpha_2) e^{ht_{\text{sw}}}}{(\alpha_1^* - \alpha_2 + h)(\alpha_1^* + \alpha_2)} X^0 + \\ & + \frac{(-2\alpha_1^* e^{(\alpha_1^* + \alpha_2 + h)t_{\text{sw}}} + (\alpha_1^* - \alpha_2) e^{ht_{\text{sw}}})}{(\alpha_1^* - \alpha_2 + h)(\alpha_1^* + \alpha_2)} Y^0 + \\ & + \frac{(\alpha_1^* - \alpha_2) N e^{(\alpha_1^* + \alpha_2 + h + \gamma)t_{\text{sw}}} - (\alpha_1^* - \alpha_2) e^{ht_{\text{sw}}}}{(\alpha_1^* - \alpha_2 + h)(\alpha_1^* + \alpha_2 + \gamma)} + \\ & + \frac{(\alpha_2 - \alpha_1^* - h) N e^{(\alpha_1^* + \alpha_2 + h + \gamma)t_{\text{sw}}}}{(\alpha_1^* - \alpha_2 + h)(\alpha_1^* + h + \alpha_2 + \gamma)}, \quad (35) \end{aligned}$$

$$\begin{aligned} C_2^* = & \frac{1}{(\alpha_1^* + \alpha_2)(\alpha_1^* - \alpha_2 + h)(\alpha_1^* + \alpha_2 + \gamma)} \times \\ & \times \{ [\alpha_1^* \alpha_2 (\alpha_1^* + 2\alpha_2 + h + \gamma) \alpha_1^* \alpha_2 - (\alpha_1^* \alpha_2 + \alpha_2^2 + \alpha_2 \gamma) e^{-(\alpha_1^* + \alpha_2)t_{\text{sw}}} h] X^0 + \\ & + [-\alpha_1^* (\alpha_1^{*2} + \alpha_1^* (h + \gamma) - h\gamma) - (\alpha_2 + h + \gamma + 2\alpha_1^*) \alpha_1^* \alpha_2 - \\ & - (\alpha_1^* + \alpha_2 + \gamma) \alpha_2 h e^{-(\alpha_1^* + \alpha_2)t_{\text{sw}}}] Y^0 + \\ & + (\alpha_1^* \gamma + \alpha_2 N) \alpha_2 h e^{-(\alpha_1^* + \alpha_2)t_{\text{sw}}} - (\alpha_1^* + \alpha_2) \alpha_2 h N e^{\gamma t_{\text{sw}}} \}. \quad (36) \end{aligned}$$

Выражение (34) позволяет определить значения функций $X_0(t)$, $Y_0(t)$ при $t = T$ и получить из условий периодичности (28) систему уравнений для X_0^0 , Y_0^0 :

$$\begin{cases} X_0^0 = C_1^* (\alpha_1^* + h) e^{-(\alpha_1^* + h + \alpha_2 + \gamma)T} + C_2^* e^{-\gamma T} + \frac{(\alpha_1^* + h)N}{\alpha_1^* + h + \alpha_2 + \gamma}, \\ Y_0^0 = C_1^* \alpha_2 e^{-(\alpha_1^* + h + \alpha_2 + \gamma)T} - C_2^* e^{-\gamma T} + \frac{\alpha_2 N}{\alpha_1^* + h + \alpha_2 + \gamma}. \end{cases} \quad (37)$$

Здесь постоянные C_1^* , C_2^* линейно зависят от X_0^0 , Y_0^0 . Таким образом, соотношения (37) представляют собой систему двух линейных алгебраических уравнений для X_0^0 , Y_0^0 . Нетрудно показать, что при выполнении условия (23) определитель этой системы отличен от нуля. При выполнении этого условия значения X_0^0 , Y_0^0 определяются однозначно. Тем самым асимптотика нулевого порядка на отрезке $[0; T]$

построена; она дается формулами (33)–(36). При $t > T$ функции $X_0(t)$, $Y_0(t)$ продолжаются периодически с периодом T .

Уравнения для следующих (после нулевого) членов асимптотического разложения (27) имеют вид ($i = 1, 2, \dots$)

$$\begin{cases} \frac{dX_i}{dt} = -(\alpha_1 + \gamma)X_i - \alpha_1 Y_i + f_i(t), \\ \frac{dY_i}{dt} = -\alpha_2 X_i - (\alpha_2 + \gamma)Y_i + g_i(t). \end{cases} \quad (38)$$

Здесь функции $f_i(t)$, g_i выражаются через предыдущие члены асимптотики. В частности

$$\begin{aligned} X_1(t) &= X_0(t)N - (X_0(t))^2 - X_0(t)Y_0(t), \\ Y_1(t) &= Y_0(t)N - (Y_0(t))^2 - X_0(t)Y_0(t), \end{aligned}$$

где функции $X_0(t)$, $Y_0(t)$ уже известны к моменту определения функций $X_1(t)$, $Y_1(t)$.

Уравнения (38) дополняются условиями периодичности (28) и условиями непрерывности (29). Тем самым построен метод определения всех членов асимптотического разложения (27).

Проведенное аналитическое исследование дополняется численным экспериментом. В качестве примера приведем расчет со следующими параметрами:

$$\begin{aligned} \alpha_1^* &= 0,6; \quad h = 2; \quad \alpha_2 = 1,4; \quad \beta = 0,1; \\ \gamma &= 0,3; \quad N = 10; \quad T = 4; \quad t_{sw} = 2. \end{aligned} \quad (39)$$

Поведение системы демонстрируется графиками, приведенными на рис. 3 и рис. 4. Сплошные кривые представляют решение системы (24)–(26), полученное численно. Пунктирные кривые соответствуют нулевому порядку (33)–(36) асимптотики (27), т.е. (с определенной точностью) решению при $t \rightarrow \infty$. В целом, поведение системы при дебалансирующем периодическом воздействии можно охарактеризовать

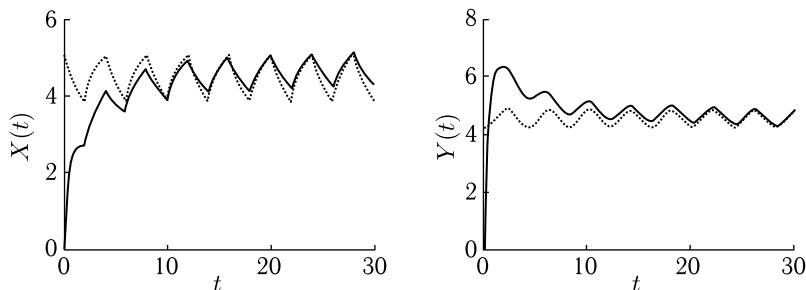


Рис. 3. Решение системы (24)–(25) с параметрами (39), полученное численно (сплошная линия), и асимптотика нулевого приближения (33)–(36) (пунктир).

На левом рисунке — график $X(t)$, на правом — $Y(t)$

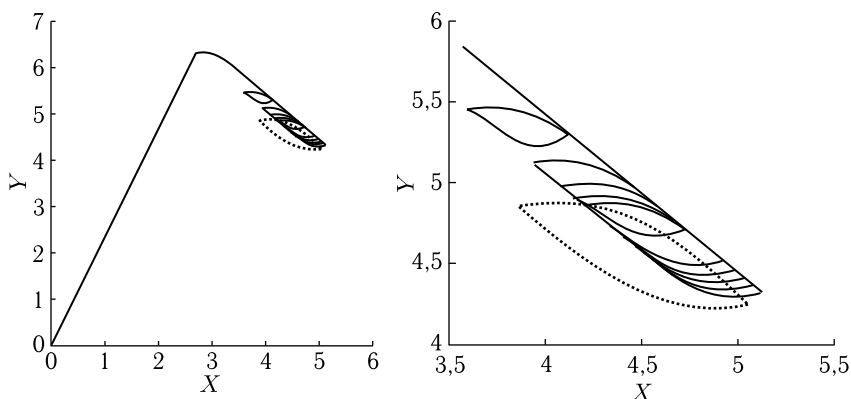


Рис. 4. Фазовая траектория $\{X(t); Y(t)\}$ системы (24)–(25) с параметрами (39), изображенная в мелком и крупном масштабах. Сплошная линия — численное решение, пунктир — асимптотика нулевого приближения (33)–(36)

как выход на периодический режим после переходного периода. В этом периодическом режиме чередуются интервалы времени, в течение которых то одна, то другая сторона имеет преимущество в численности adeptов.

4.6.5. Заключение

В настоящем разделе рассмотрен ряд моделей информационного противоборства, учитывающих как передачу информации при межличностной коммуникации, так и пропаганду через СМИ. Базовая модель учитывает основные факторы противоборства, следующие модели учитывают неполный охват социума СМИ массовой информации, усвоение информации лишь после неоднократного ее получения, забывания информации индивидами, а также рассматривают процесс противоборства при периодическом дестабилизирующем воздействии. Результатам математического исследования дана содержательная трактовка.

Упражнения

1. Если выполнено условие (8) для победы первой партии в базовой модели информационного противоборства, но $\alpha_1 < \alpha_2$, то имеет место «смена лидера», т. е. $X_1(t) < X_2(t)$ при $0 < t < t_c$ и $X_1(t) > X_2(t)$ при $t > t_c$. Определите величину t_c .
2. Получить решение (11) системы (10) модели противоборства произвольного количества партий.
3. Доказать, что победу в противоборстве произвольного количества партий одерживает партия, имеющая наибольшее значение функции победы, даваемой формулой (12).
4. Показать, что система (24), (25) не имеет стационарного решения при $\gamma > 0$.

Комментарии и библиографические примечания

1. Базовая модель информационного противоборства предложена в работах [1]. Описание дополнительных факторов разработано в [2] для модели распространения информации без противодействия; эти факторы включены в модель противоборства в [3]. Изложенная в четвертом пункте модель противоборства при периодической дестабилизации предложена в [4].

2. Краткий обзор других подходов к моделированию влияния информации на общество, информационного противоборства приведен в комментариях к предыдущему разделу.

3. Если в модели с дестабилизирующим воздействием (24)–(26) положить $\gamma = 0$ (отсутствует забывание информации индивидами), то задача имеет стационарное решение (а не периодическое, как при $\gamma > 0$). Этот случай рассмотрен в работе [5].

Литература

1. Михайлов А. П., Маревцева Н. А. Модели информационной борьбы // Математическое моделирование. Т. 23. 2011. № 10. С. 19–32.
2. Михайлов А. П., Петров А. П., Маревцева Н. А., Третьякова И. В. Развитие модели распространения информации в социуме // Математическое моделирование. 2014. Т. 26, № 3. С. 65–74 (англ. перевод: *Mikhailov A. P., Petrov A. P., Marevtseva N. A., Tretyakova I. V. Development of a Model of Information Dissemination in Society // Mathematical Models and Computer Simulations. 2014. V. 6, № 5. P. 535–541.*)
3. Петров А. П., Прончева О. Г. Исследование моделей информационного нападения и информационного противоборства в структурированном социуме // Математическое моделирование социальных процессов. Вып. 17. Под ред. А. П. Михайлова — М.: Эконинформ, 2015. С. 136–149.
4. Михайлов А. П., Петров А. П., Прончева О. Г., Маревцева Н. А. Модель информационного противоборства в социуме при периодическом дестабилизирующем воздействии // Математическое моделирование. 2017. Т. 29, № 2 С. 23–32 (англ. перевод: *Mikhailov A. P., Petrov A. P., Proncheva O. G., Marevtseva N. A. A Model of Information Warfare in a Society Under a Periodic Destabilizing Effect // Mathematical Models and Computer Simulations. 2017. V. 9, № 5. P. 580–586.*)
5. Михайлов А. П., Петров А. П., Прончева О. Г. Модель информационного противоборства в социуме с кусочно-постоянной функцией дестабилизирующего воздействия // Математическое моделирование. 2018. Т. 30, № 7. С. 47–60.

Глава 5

ПОЛИТОЛОГИЯ

5.1. Политическое соперничество

Рассматривается модель, описывающая временную динамику установления равновесия в системе двух конкурирующих в борьбе за власть политических сил (например, ветвей власти). Изучена зависимость равновесных значений уровней власти, времени прихода к равновесию и ряда других величин от параметров системы — поведенческих характеристик партнеров (в т. ч., от степени их «компетентного властолюбия»). Каждая из политических сил представляется в данной модели как единое целое (в отличие от рассмотренных в разд. 5.3–5.5 моделей системы «власть–общество», уделяющих основное внимание внутренней структуре политической иерархии). Вследствие этого модель имеет вид системы нелинейных обыкновенных дифференциальных уравнений. Первый пункт раздела посвящен построению модели, второй — аналитическому исследованию стационарных решений методами качественной теории дифференциальных уравнений, третий — численным экспериментам и содержательным выводам.

5.1.1. Постановка задачи и формулировка модели

Основными искомыми величинами в описываемой модели являются текущие уровни власти партнеров (данное понятие вполне аналогично тому, что используется в модели «власть–общество», см. далее в настоящей главе). Они описываются безразмерными функциями времени $p_1(t) > 0$, $p_2(t) > 0$, $t \geq t_0$ (t_0 — время, начиная с которого изучается эволюция системы).

Считается, что у каждой ветви имеется свое представление о том, какой уровень власти она должна была бы иметь в данный момент времени и каким должен был бы быть уровень власти у ее конкурента. Эти представления описываются следующими заданными функциями времени:

$p_{11}(t) > 0$ — уровень власти, который первый партнер хотел бы иметь для себя;

$p_{12}(t) > 0$ — уровень власти, который первый партнер хотел бы иметь для второго;

$p_{22}(t) > 0$ — уровень власти, который второй партнер хотел бы иметь для себя;

$p_{21}(t) > 0$ — уровень власти, который второй партнер хотел бы иметь для первого.

Считается также, что величина $p_1 + p_2$, как и величины $p_{11} + p_{12}$ и $p_{22} + p_{21}$ находятся в конституционных рамках, т. е. ограничены сверху и снизу максимальными и минимальными полномочиями. Строго говоря, для гарантированного выполнения этих ограничений в модель необходимо вводить соответствующий «регулятор» (например, активность конституционного суда или гражданского общества), но для главной цели настоящей работы, изучения ситуаций, близких к равновесным, такое дополнение несущественно.

Относительно введенных величин возможны три следующие ситуации. Если одновременно выполняются соотношения

$$p_{11}(t) = p_{21}(t), \quad p_{12}(t) = p_{22}(t), \quad t \geq t_0, \quad (1)$$

$$p_1(t) = p_{11}(t), \quad p_2(t) = p_{22}(t), \quad t \geq t_0, \quad (2)$$

то, очевидно, стороны согласны друг с другом и не предпринимают никаких действий к изменению величин $p_1(t)$ и $p_2(t)$. В этом случае имеется «согласованное равновесие».

Если равенства (1) выполнены, а равенства (2) нет, то партнеры эволюционируют к согласованному равновесию (без борьбы и конкуренции).

Наконец, если хотя бы одно из равенств (1) не выполняется, то стороны предпринимают действия в защиту своей позиции, в результате чего приходят к «вынужденному равновесию» — компромиссу (в дальнейшем будем называть его просто равновесием). Именно этот случай изучается в данном разделе.

При формулировке математической модели делаются следующие предположения.

1. В общем случае каждый из партнеров ведет себя таким образом, чтобы реализовать свои представления об уровне власти как для себя, так и для конкурента.

В этом варианте партнеры действуют в некотором смысле симметрично. Возможны также и другие варианты поведения. Например, партнеры (один или оба) безразлично относятся к факту превышения своей власти над равновесным уровнем или к факту понижения уровня власти партнера по сравнению с равновесным и т. п. Переход к этим и другим вариантам поведения не составляет особого труда, поскольку структура математической модели в принципе не меняется. В дальнейшем будем рассматривать «симметричную» модель поведения.

2. Интенсивность действий партнеров (число «единичных» актов в поддержку своей позиции за единицу времени) определяется пропорционально трем факторам:

- разности между желаемым и текущим уровнями власти;
- самими текущими значениями уровней власти;
- компетентным властолюбием партнеров, т. е. желанием, подкрепленным соответствующей компетентностью (интеллектом, амбициями, информационными и финансовыми ресурсами, опытом и т. д.).

Основываясь на предположениях (1, 2), приходим к следующей математической модели:

$$\begin{aligned} \frac{dp_1}{dt} &= f_{11}p_1(p_{11} - p_1) + f_{21}p_2(p_{21} - p_1) \\ \frac{dp_2}{dt} &= f_{22}p_2(p_{22} - p_2) + f_{12}p_1(p_{12} - p_2), \end{aligned} \quad (3)$$

где

$$f_{ij} = f_{ij}(t, p_1, p_2, p_{11}, p_{22}, p_{12}, p_{21}) \geq 0, \quad i, j = 1, 2,$$

— степени компетентного властолюбия партнеров, заданные функции времени, искомым величин и желаемых партнерами уровней власти.

Система (3) — нелинейная система двух обыкновенных дифференциальных уравнений первого порядка, коэффициенты которой зависят в общем случае от времени, параметров модели и решения. Если известны начальные условия при $t = t_0$,

$$p_1(t_0) = p_1^0 > 0, \quad p_2(t_0) = p_2^0 > 0. \quad (4)$$

то из (3), (4) однозначно находится решение — пара функций $p_1(t)$, $p_2(t)$ для всех $t > t_0$.

Итак, при сделанных предположениях динамика изменения уровней власти противоборствующих сторон описывается моделью (3), (4), включающей параметры p_{ij} , f_{ij} , p_i^0 , $i, j = 1, 2$. Зная их, можно получить временную зависимость искомым функций $p_1(t)$, $p_2(t)$ (в частных случаях — аналитически, в общем случае, ввиду нелинейности задачи, — численно).

Построенная модель допускает ряд естественных обобщений (борьба между несколькими партнерами, вмешательство внешних сил и т. д.). Отметим также, что известные модели соперничества (хищник–жертва, гонка вооружений, боевые действия армий) описывают системы с принципиально иными свойствами, целями и поведенческими характеристиками.

5.1.2. Равновесные уровни власти и их устойчивость

Важным частным случаем модели, позволяющим применить аналитические методы исследования, является ситуация, когда поведенческие характеристики партнеров f_{ij} и «желаемые» уровни власти p_{ij} не зависят от времени t . Тогда система (3) представляет собой нелинейную автономную систему дифференциальных уравнений. Такие системы удобно изучать в плоскости переменных p_1 , p_2 (фазовой плоскости).

Для этого случая будем находить положения ее равновесия, т. е. те точки на фазовой плоскости, в которых правые части уравнений (3) обращаются в нуль одновременно, а также выяснять тип этих положений равновесия.

Далее будем изучать модель, в которой коэффициенты f_{ij} , p_{ij} постоянны.

Итак, положения равновесия автономной системы уравнений (3) определяются из условий

$$\frac{dp_1}{dt} = \frac{dp_2}{dt} = 0,$$

что равносильно выполнению соотношений

$$p_2 = -\frac{f_{11}p_1(p_{11} - p_1)}{f_{21}(p_{21} - p_1)}, \quad (5)$$

$$p_1 = -\frac{f_{22}p_2(p_{22} - p_2)}{f_{12}(p_{12} - p_2)}. \quad (6)$$

Очевидно, $p_1 = 0$, $p_2 = 0$ — одна из таких точек равновесия (тривиальная). Будем считать далее, что $p_{11} \neq p_{21}$ и $p_{22} \neq p_{12}$ (см. (1)). Тогда из (5) следует $p_1 \neq p_{21}$, а из (6) следует $p_2 \neq p_{12}$. Кроме того, из (5), (6) вытекают равенства $p_1 \neq p_{11}$, $p_2 \neq p_{22}$.

Подставляя (5) в (6) и сокращая обе части полученного соотношения на $p_{21} - p_1 \neq 0$, получим уравнение относительно p_1 ,

$$\begin{aligned} f_{12}f_{21}(p_{21} - p_1)(f_{21}p_{12}p_{21} + (f_{11}p_{11} - f_{21}p_{12})p_1 - f_{11}p_1^2) = \\ = f_{11}f_{22}(p_{11} - p_1)(f_{21}p_{22}p_{21} + (f_{11}p_{11} - f_{21}p_{22})p_1 - f_{11}p_1^2), \quad (7) \end{aligned}$$

которое определяет еще три решения системы (5), (6) (в дополнение к тривиальному).

Нас интересуют лишь положительные решения системы уравнений (5), (6) (соответственно уравнения (7)).

Пусть исходные данные f_{ij} , p_{ij} — любые положительные числа и такие, что $p_{11} \neq p_{21}$ и $p_{22} \neq p_{12}$. Тогда из (5) следует, что величины $p_{11} - p_1$ и $p_{21} - p_1$ имеют разные знаки. Это означает, что значение p_1 лежит в интервале с концами p_{11} и p_{21} . Аналогично, из (6) следует, что значение p_2 лежит в интервале с концами p_{22} и p_{12} . На одном из концов функция $p_2 = p_2(p_1)$ из (5) обращается в ноль, на другом — в ∞ , в промежутке монотонна. Аналогично ведет себя функция $p_1 = p_1(p_2)$ из (6) на соответствующем интервале.

Таким образом, система (5), (6) имеет по крайней мере одно положительное решение, причем все ее существующие положительные решения (не больше трех) лежат в области, где p_1 меняется между p_{11} и p_{21} , а p_2 — между p_{22} и p_{12} . Будем называть эту область в плоскости p_1, p_2 «властным прямоугольником» и обозначать P (см. рис. 1).

Нахождение точек равновесия внутри прямоугольника P , аналитически или численно, не представляет трудностей.

Исследуем устойчивость положительных точек равновесия системы (3). Положение равновесия считается асимптотически устойчивым, если

$$\left\| \begin{pmatrix} p_1(t) \\ p_2(t) \end{pmatrix} \right\| \rightarrow \left\| \begin{pmatrix} p_1^* \\ p_2^* \end{pmatrix} \right\|$$

при $t \rightarrow \infty$ и начальных данных, достаточно близких к $\|p_1^*, p_2^*\|^T$.

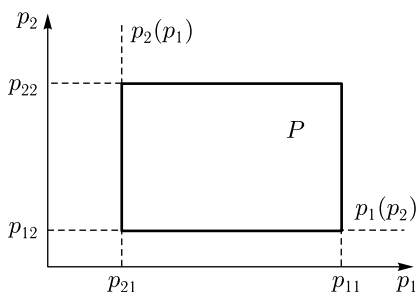


Рис. 1. Властный прямоугольник

Представим систему (3) в виде

$$\frac{dp_1}{dt} = g_1(p_1, p_2), \quad \frac{dp_2}{dt} = g_2(p_1, p_2).$$

Пусть λ_1, λ_2 — собственные значения матрицы

$$A = \begin{pmatrix} \frac{\partial g_1}{\partial p_1} & \frac{\partial g_1}{\partial p_2} \\ \frac{\partial g_2}{\partial p_1} & \frac{\partial g_2}{\partial p_2} \end{pmatrix}.$$

Согласно классификации положений равновесия линейных автономных систем положение равновесия асимптотически устойчиво, если $\text{Re } \lambda_1 < 0, \text{Re } \lambda_2 < 0$. Как известно,

$$\text{tr } A = \lambda_1 + \lambda_2, \quad \det A = \lambda_1 \lambda_2.$$

Тогда условие асимптотической устойчивости эквивалентно условиям

$$\text{tr } A < 0, \quad \det A > 0. \quad (8)$$

Матрица A для системы (3) имеет вид

$$A = \begin{pmatrix} f_{11}(p_{11} - 2p_1) - f_{21}p_2 & f_{21}(p_{21} - p_1) \\ f_{12}(p_{12} - p_2) & f_{22}(p_{22} - 2p_2) - f_{12}p_1 \end{pmatrix}.$$

Отсюда следует

$$\begin{aligned} \text{tr } A &= f_{11}p_{11} + f_{22}p_{22} - (2f_{11} + f_{12})p_1 - (2f_{22} + f_{21})p_2, \\ \det A &= [f_{11}(p_{11} - 2p_1) - f_{21}p_2][f_{22}(p_{22} - 2p_2) - \\ &\quad - f_{12}p_1] - f_{12}f_{21}(p_{21} - p_1)(p_{12} - p_2). \end{aligned} \quad (9)$$

Рассмотрим, какого типа могут быть положения равновесия системы (3). Собственные значения матрицы A выражаются следующим образом:

$$\lambda_{1,2} = \frac{\text{tr } A \pm \sqrt{(\text{tr } A)^2 - 4 \det A}}{2}.$$

Обозначим $D = (\text{tr } A)^2 - 4 \det A$ и исследуем его знак.

Напомним, что если выполнены условия (8), то существует два основных типа устойчивых положений равновесия. При $D > 0$ собственные значения вещественны и отрицательны; положение равновесия — устойчивый узел. При $D < 0$ собственные значения комплексносопряженные с отрицательной действительной частью; положение равновесия — устойчивый фокус.

Обозначим $\partial g_1/\partial p_1 = q_1$, $\partial g_2/\partial p_2 = q_2$. В этих обозначениях

$$\begin{aligned}\operatorname{tr} A &= q_1 + q_2, \\ \det A &= q_1 q_2 - f_{12} f_{21} (p_{21} - p_1)(p_{12} - p_2), \\ D &= (q_1 - q_2)^2 + 4f_{12} f_{21} (p_{21} - p_1)(p_{12} - p_2).\end{aligned}$$

Второе слагаемое в формуле для D , следовательно и величина D положительны, если величины $p_{21} - p_1$ и $p_{12} - p_2$ одного знака.

Таким образом, если у системы (3) существует асимптотически устойчивое положение равновесия, то при указанных соотношениях параметров это положение равновесия есть устойчивый узел.

Будем называть властный прямоугольник «типичным», если его параметры удовлетворяют условиям $p_{21} < p_1 < p_{11}$ и $p_{12} < p_2 < p_{22}$. Из вышесказанного следует, что в типичном властном прямоугольнике положение равновесия может быть только узлом, устойчивым или неустойчивым.

Исследуем вопрос о том, при каких параметрах модели f_{ij} , p_{ij} у нее имеется устойчивое положение равновесия.

В общем случае поле направлений системы (3) при любых положительных параметрах модели таково, что на границах прямоугольника P все векторы направлены внутрь этой области. Будем предполагать, что предельных циклов в задаче нет. Тогда это означает, что внутри властного прямоугольника всегда имеются асимптотически устойчивые точки равновесия системы (точки притяжения решения).

Рассмотрим теперь некоторые частные случаи. Остановимся сначала на одном из вариантов модели, когда решение системы (3) находится аналитически. Предположим, что

$$\begin{aligned}p_{11} &= p_{22}, & p_{21} &= p_{12}, \\ f_{11} &= f_{22}, & f_{21} &= f_{12}.\end{aligned}$$

В этом случае задача (5), (6) имеет решение на диагонали властного прямоугольника:

$$p_1 = p_2 = \frac{f_{11}p_{11} + f_{12}p_{12}}{f_{11} + f_{12}}. \quad (10)$$

Для него матрица A симметрическая, спектр ее веществен. Подставляя выражения (10) в (9), получим

$$\begin{aligned}\frac{1}{2}\operatorname{tr} A &= -(f_{12}p_{12} + f_{11}\frac{f_{11}p_{11} + f_{12}p_{12}}{f_{11} + f_{12}}), \\ \det A &= (f_{11}p_{11} + f_{12}p_{12})[f_{11}(f_{11} - f_{12})p_{11} + f_{12}(f_{12} + 3f_{11})p_{12}].\end{aligned}$$

Таким образом, $\text{tr} A < 0$ при любых положительных параметрах, а $\det A > 0$, если

$$f_{11}(f_{11} - f_{12})p_{11} + f_{12}(f_{12} + 3f_{11})p_{12} > 0.$$

Обозначая $\Delta p = p_{11} - p_{12}$, перепишем это условие в виде

$$p_{12} > \frac{f_{11}(-f_{11} + f_{12})}{(f_{11} + f_{12})^2} \Delta p. \quad (11)$$

Условие (11) есть необходимое и достаточное. Если оно нарушается, то положение равновесия (10) перестает быть асимптотически устойчивым.

Если $f_{11} > f_{12}$, то условие (11) выполняется, т. е. это неравенство — достаточное условие устойчивости. Смысл его в том, что равным партнерам для достижения устойчивого равновесия достаточно прикладывать усилия в свою пользу не меньше, чем в пользу противника. Кроме того, поскольку множитель перед Δp в правой части (11) меньше единицы, то условие $p_{12} > \Delta p$ также является достаточным для устойчивости.

Если условие (11) нарушается, то точка равновесия (10) является неустойчивой. Аналитические исследования, подтвержденные расчетами, показывают, что в этом случае помимо точки (10) возникают два симметрично расположенных относительно центра прямоугольника P устойчивых узла. Причем, когда условие (11) становится равенством, точка (10) становится трехкратным решением системы (5), (6), а затем от нее отделяются две указанные точки.

Смысл такой фазовой картины состоит в том, что в этом случае при равных начальных данных небольшое отклонение уровня власти в пользу одного из партнеров приводит к устойчивому равновесию, преимущественному для этого партнера.

Заметим, что при дополнительном условии $f_{ij} = f$ система (3) имеет решение

$$p_1 + p_2 = \frac{p_{11} + p_{12}}{1 + C(p_{11} + p_{12}) \exp(-(p_{11} + p_{12})ft)},$$

где C — произвольная постоянная. Отсюда следует, что при $t \rightarrow \infty$ $p_1 + p_2 \rightarrow p_{11} + p_{12}$, что наглядно иллюстрирует симметричность положений равновесия системы (3) в этом случае относительно диагонали властного прямоугольника.

Рассмотрим теперь названный нами типичным вариант борьбы, когда $p_{21} < p_1 < p_{11}$ и $p_{12} < p_2 < p_{22}$ (основной случай). Обозначая $\Delta p_1 = p_{11} - p_{21}$, $\Delta p_2 = p_{22} - p_{12}$, из условия $\text{tr} A < 0$ получим следующее достаточное условие устойчивости:

$$\text{tr} A < -f_{11}(p_1 - \Delta p_1) - f_{22}(p_2 - \Delta p_2) - f_{12}p_1 - f_{21}p_2 < 0.$$

Это неравенство выполняется, в частности, если

$$p_{21} \geq \Delta p_1, \quad p_{12} \geq \Delta p_2. \quad (12)$$

Не выписывая полностью соответствующих выражений, отметим, что условия (12) являются достаточными также и для выполнения неравенства $\det A > 0$ (см. (9)).

Таким образом, для типичного случая мы получили условия (12), которые являются достаточными для устойчивости существующих внутри P точек равновесия.

Смысл их состоит в том, что разброс во мнениях партнеров об уровнях власти каждого из них не должен превышать нижней оценки этого уровня.

При этом точки равновесия, как было показано, являются устойчивыми узлами. При нарушении указанных условий во властном прямоугольнике могут быть как устойчивые, так и неустойчивые точки равновесия (также узлы) аналогично рассмотренному случаю равных партнеров.

Итак, при типичном соотношении параметров задачи положение равновесия внутри соответствующего властного прямоугольника P всегда является узлом: устойчивым или неустойчивым.

Рассмотрим теперь вариант борьбы, когда типичное соотношение параметров для одного из партнеров, например второго, нарушено, а именно: $p_{21} < p_1 < p_{11}$, $p_{22} < p_2 < p_{12}$ (нетипичный случай).

В этом случае дискриминант $D = (\operatorname{tr} A)^2 - 4 \det A$, определяющий тип положения равновесия, может быть отрицательным, т. е. положение равновесия может быть фокусом, устойчивым или неустойчивым. Необходимым и достаточным для устойчивости фокуса является условие $\operatorname{tr} A < 0$. Из него аналогично предыдущему получается достаточное условие

$$\operatorname{tr} A < -f_{11}(p_1 - \Delta p_1) - f_{12}p_1 - (f_{22} + f_{21})p_2 < 0,$$

откуда следует также достаточное условие (ср. с (12))

$$p_{21} \geq \Delta p_1. \quad (13)$$

Отметим, что условие (13) является достаточным также для выполнения неравенства $\det A > 0$, т. е. является достаточным условием устойчивости как для узла, так и для фокуса в этом варианте.

5.1.3. Некоторые вычислительные эксперименты с моделью

Приведем характеристики положений равновесия системы для случая равных партнеров при различных входных параметрах. Пусть

$$f_{11} = f_{22} = 1, \quad p_{11} = p_{22} = 1, \quad f_{12} = f_{21} = f, \quad p_{12} = p_{21} = p < 1,$$

где p, f — параметры. Тогда властный прямоугольник — это квадрат $p < p_1, p_2 < 1$. Необходимое и достаточное условие устойчивости (11) имеет вид

$$p > \frac{f-1}{f-1+(f+1)^2}. \quad (14)$$

При $f \leq 1$ это условие выполняется для любых значений $p > 0$. При этом имеется единственная устойчивая точка равновесия (узел).

Возьмем $f = 2$ и будем варьировать значение p . При этом условие (14) имеет вид $p > 0,1$. Приведем результаты вычислений точек равновесия задачи (исключая нулевую), иллюстрирующие изменение их взаимного расположения:

$$p = 0,40: \quad p_1 = p_2 = 0,60; \quad p_1, p_2 = 0,10 \pm 0,38i; \quad p_1, p_2 = 0,10 \mp 0,38i;$$

$$p = 0,11: \quad p_1 = p_2 = 0,41; \quad p_1, p_2 = 0,39 \pm 0,14i; \quad p_1, p_2 = 0,39 \mp 0,14i;$$

$$p = 0,10: \quad p_1 = p_2 = 0,4 - \text{тройной узел};$$

$$p = 0,09: \quad p_1 = p_2 = 0,39; \quad p_1, p_2 = 0,55, 0,27; \quad p_1, p_2 = 0,27, 0,55;$$

$$p = 0,0: \quad p_1 = p_2 = 0,37; \quad p_1, p_2 = 0,78, 0,11; \quad p_1, p_2 = 0,11, 0,78.$$

В последних двух случаях внедиагональные точки равновесия устойчивы и отстоят от диагонали тем дальше, чем меньше значение p .

Такое смещение устойчивого равновесия с диагонали властного прямоугольника и возникающая в связи с этим возможность бифуркации вызваны здесь взаимным стремлением партнеров уменьшить желаемый уровень власти для конкурента.

Рассмотрим теперь некоторые характеристики положений равновесия для нетипичного случая. Пусть

$$f_{11} = f_{22} = 1, \quad p_{11} = p_{12} = 1, \quad f_{12} = f_{21} = f, \quad p_{22} = p_{21} = p < 1.$$

Здесь также $p < p_1, p_2 < 1$ — властный прямоугольник. Условие $D < 0$ (условие того, что положение равновесия есть фокус) имеет вид

$$[(1-p) - (2-f)(p_1-p_2)]^2 < 4f^2(p_1-p)(1-p_2).$$

Достаточным условием устойчивости фокуса в этом случае является соотношение

$$p_1 + p_2 > \frac{1+p}{2+f},$$

а из (13) следует условие $p \geq 0,5$. Возьмем $p = 0,5$, что гарантирует устойчивость полученных положений равновесия, и будем варьировать значение f , см. табл. 1.

Из приведенных данных видно, как с увеличением значения f меняется фазовый портрет задачи: точка устойчивого равновесия системы из узла превращается в фокус.

Таким образом, при усилении борьбы партнеров друг против друга траектории приближения к положению равновесия приобретают немонотонный характер.

Таблица 1

f	p_1	p_2	Тип
0,05	0,9866	0,5417	Узел
0,1	0,9722	0,5726	Узел
0,15	0,9572	0,5970	Узел
0,2	0,9421	0,6170	Фокус
0,5	0,8562	0,6912	Фокус
1,0	0,7500	0,7500	Вырожд. узел ($D=0$)
2,0	0,6423	0,8071	Фокус
5,0	0,5561	0,8798	Фокус

Применим описанную модель для анализа следующей ситуации. Предположим, для некоторых параметров модели (допустимых границ уровней власти и степеней компетентного властолюбия партнеров) известно устойчивое положение равновесия этой системы, т. е. значения p_1 , p_2 . Предположим, что эти значения не устраивают одного из партнеров, например первого, и он хотел бы изменить ситуацию в более выгодную для себя сторону (увеличить значение p_1). Какие усилия он должен приложить для этого, а именно, как он должен изменить степень своего властолюбия (f_{11})? Ниже приводятся соответствующие результаты вычислений.

Вариант 1. Для значений

$$0,6 \leq p_1 \leq 1,0, \quad 0,8 \leq p_2 \leq 1,2, \\ f_{21} = f_{12} = 1, \quad f_{11} = 1,25, \quad f_{22} = 0,8$$

точкой устойчивого равновесия системы является точка $p_1 = 0,8$, $p_2 = 1,0$. Покажем, как меняется значение f_{11} в зависимости от увеличения равновесного значения p_1 на несколько процентов, см. табл. 2.

Таблица 2

p_1	0,8 (0%)	0,816 (2%)	0,840(5%)	0,864(8%)	0,872(9%)
p_2	1,000	0,998	0,995	0,991	0,990
f_{11}	1,250	1,435	1,776	2,228	2,414

Во второй строке приведены соответствующие равновесные значения p_2 . Отсюда видно, что увеличение своего властного уровня всего на 9% при этих параметрах требует от первого партнера почти двукратного увеличения степени своего властолюбия (при этом уровень власти второго партнера практически не меняется). Ситуация еще более

усугубляется, если отправная точка устойчивого равновесия расположена близко к его верхней границе — уровню власти p_{11} . Рассмотрим в связи с этим другой вариант.

Вариант 2. Для значений

$$0,6 \leq p_1 \leq 1,1, \quad 0,8 \leq p_2 \leq 1,3,$$

$$f_{21} = f_{12} = 1, \quad f_{11} = 4,8, \quad f_{22} = 3,3$$

точкой устойчивого равновесия системы является точка $p_1 = 1,0$, $p_2 = 1,2$.

Соответствующая зависимость изменения величины f_{11} от увеличения p_1 имеет вид, см. табл. 3.

Таблица 3

p_1	1,0 (0%)	1,02 (2%)	1,05 (5%)	1,08 (8%)	1,09 (9%)
p_2	1,200	1,197	1,195	1,192	1,191
f_{11}	4,800	6,163	10,21	26,50	53,56

В этом случае увеличение своего уровня власти на ту же величину, что и в предыдущем варианте, требует на порядок больше усилий (эффект насыщения).

В заключение сделаем качественную оценку времени прихода системы к устойчивому положению равновесия a . Пусть у матрицы A в силу устойчивости равновесия $\operatorname{Re} \lambda_1 \leq \operatorname{Re} \lambda_2 < 0$. Тогда

$$|y(t) - a| \sim |y(t_0) - a| e^{t \operatorname{Re} \lambda_2}.$$

Если задать малое $\varepsilon > 0$, то условие

$$\frac{|y(t) - a|}{|y(t_0) - a|} \leq \varepsilon$$

эквивалентно условию

$$t \geq \frac{\ln \varepsilon}{\operatorname{Re} \lambda_2}.$$

Проиллюстрируем эту оценку на примере вариантов 1, 2 последнего из рассмотренных случаев.

В варианте 1 для первых трех наборов параметров (колонок таблицы) имеем следующие оценки времени прихода к равновесию (a именно, при уменьшении начального отклонения от положения равновесия в 100 и 1000 раз):

$$t \sim \begin{array}{ll} \varepsilon = 0,01: & 3,432 \ 3,355 \ 3,260, \\ \varepsilon = 0,001: & 5,142 \ 5,033 \ 4,890. \end{array}$$

В варианте 2 также для первых трех колонок таблицы имеем

$$t \sim \begin{array}{ll} \varepsilon = 0,01: & 1,020 \ 1,0100,997, \\ \varepsilon = 0,001: & 1,543 \ 1,5141,494). \end{array}$$

В обоих случаях время установления равновесия, естественно, уменьшается с ростом усилия f_{11} .

Сформулируем основные выводы, следующие из анализа построенной и изученной в данном разделе модели.

1. Система (3) всегда имеет положительные точки равновесия. Все они (их не больше трех) лежат внутри властного прямоугольника P . Другими словами, равновесные уровни власти располагаются между уровнями, которые партнеры желали бы иметь для себя и для соперника (компромисс властей).

2. Во всех рассмотренных (частных) случаях соперничества (см. условия (12), (13)) гарантией устойчивости положения равновесия является императив: «Не желай для себя слишком большой власти, а для партнера — слишком малой власти».

3. Стремление партнеров изменить устойчивое равновесие в свою пользу должно сопровождаться мобилизацией соответствующих ресурсов, причем относительная «плата» за изменение может быть несоразмерно больше относительной величины желаемого изменения.

Эти и некоторые другие выводы имеют вполне ясную политологическую интерпретацию и формулируются в виде количественных соотношений. Они могут быть конкретизированы применительно к соответствующей ситуации, описываемой моделью. Данная «точечная» модель в сопряжении с «пространственными» моделями может использоваться также и для более детального анализа важных проблем функционирования властных структур (реализация принципа разделения властей, кламовость и т. п.).

Упражнения

1. Рассмотреть случай, когда компетентное властолюбие одной из сторон в отношении собственной власти является возрастающей функцией ее текущего количества власти: $f_{11}(p_1)$ — возрастающая функция.

2. Рассмотреть случай, когда $f_{11}(p_1)$ и $f_{12}(p_1)$ — возрастающие функции.

Комментарии и библиографические примечания

1. Модель политического соперничества предложена в работах [1, 2].

Литература

1. Михайлов А. П., Юхно Л. Ф. Простейшая модель установления равновесия между двумя ветвями власти // Матем. моделирование. 2001. Т. 13, № 1. С. 65–76.
2. Михайлов А. П., Юхно Л. Ф., Маслов А. И. Динамическая модель конкуренции между политическими силами // Доклады Академии наук. 2000. Т. 374, № 4. С. 469–474.

5.2. Моделирование этнополитической самоидентификации

В третьем разделе главы 3 была рассмотрена концепция социально-политического потенциала (СПП), рассматривающая индивида, осуществляющего дихотомический выбор между двумя взаимоисключающими позициями «Х» и «У» и постоянно осуществляющего поступки в пользу той или иной позиции. При этом, если в социуме работают механизмы, синхронизирующие во времени действия его членов (или, хотя бы, значительного их числа), то о поведении этого социума в целом можно судить по поведению так называемой *типической личности*, применительно к которой и рассматривается динамика СПП.

Настоящий раздел посвящен применению теории СПП к изучению конкретного класса процессов, именно, процессов политической самоидентификации. Социальная группа, находящаяся в процессе самоидентификации путем осуществления дихотомического выбора, в данном контексте называется электоратом. В первом пункте излагаются результаты прикладного исследования по самоидентификации дагестанского этноса, проведенного в конце 1990-х годов. Под позициями «Х» и «У» в данном случае понимаются позиции «за сепаратизм» и «против сепаратизма» соответственно. Наибольший практический интерес представляет так называемая *социально-политическая напряженность* (СПН) — динамическая величина, имеющая смысл интенсивности поступков в пользу позиции «Х». В ходе моделирования построен прогноз изменения СПН и проведено сопоставление с реальными данными.

Во втором изучается более сложная (с математической точки зрения) задача о взаимодействии двух электоратов, каждый из которых характеризуется своей «типической личностью».

5.2.1. Политическая самоидентификация дагестанского этноса в конце 1990-х годов

Общеизвестно, что северокавказские субъекты РФ еще в советское время относились к числу наиболее неблагополучных в экономическом, социальном и политическом отношениях. Трудоизбыточность и особое геополитическое положение, сосуществование разных конфессий и традиций, многоэтничность и драматическая история — вот далеко не самый полный перечень причин, постоянно порождающих повышенную, по сравнению со среднероссийским уровнем, социально-политическую напряженность в этом регионе. После распада СССР, начала экономических и политических реформ, проблемы, таившиеся ранее «под спудом», обнаружили с особой остротой и силой.

В заметной части северокавказских этносов явно или завуалированно обозначилась тенденция решать проблемы своего региона сепаратно от России. При этом выдвигались различные обоснования, прежде всего, этнического, а также религиозного и исторического характера,

рассматривались различные степени политической независимости от российской государственности (вплоть до создания новых самостоятельных государств) и применялись различные методы борьбы, вплоть до жесткого вооруженного противостояния. Тем самым в ряде этих регионов возникли этнополитические конфликты между сторонниками сепаратизма и противниками этой линии, считающими, что решение основных проблем этноса следует искать, оставаясь в составе единого государства.

Одним из весьма проблемных в этом отношении регионов являлся Дагестан, причем сепаратистские настроения (согласно социологическим данным затрагивавшие значительную часть населения) зачастую проявлялись скрытно, принимая форму тех или иных действий, дестабилизирующих социально-политическую ситуацию в регионе. Поэтому выраженная количественно интенсивность этих действий служит (опосредовано) текущей наблюдаемой мерой сепаратистских проявлений (или шире — социально-политической напряженности (СПН)) в регионе. В терминах рассмотренной в разд. 3 главы 3 модели речь идет о наблюдаемой величине E_x , где позиция «X» — это позиция в пользу сепаратизма.

Применительно к Дагестану (и к большей части северокавказских этносов) электорат разбивается на четыре группы:

1) твердые сторонники (противники) позиции «X» или «Y», или нейтралы (по данным социальной психологии их число не превышает нескольких процентов);

2) действующие в пользу позиции «X» («Y», «Z») в соответствии с абсолютным послушанием (например, так называемые *мюриды* или берущие на себя обязательства подчинения за оплату как «рекрутированные ваххабиты»);

3) находящиеся под действием пропаганды или обоснованных социальных чувств, например родственники и друзья, пострадавшие от происходящих в регионе боевых действий (они составляют от 20% до 40% от численности этноса);

4) выбирающие свою позицию в результате подражания, «заряженности» настроением, преобладающие в их социальной группе.

Из указанных групп и формируются категории $X(t)$, $Y(t)$, $Z(t)$, уравнения для динамики численности которых построены и изучены в разд. 3 главы 3.

Напомним, что модели предусматривают синхронизацию действий категорий «X» и «Y» (в данном случае такими «маркирующими датами» могут служить день рождения имама Шамиля, выборы в органы власти и т. д.).

Заметим также, что согласно этнографическим представлениям и данным, рассмотренные в разд. 3 главы 3 модели хорошо соответствуют поведению жителей крупных сел и небольших городов республики, в которых проживает основная часть населения (причем в условиях жесткого противостояния). Дальнейшие обобщения этих моделей

учитывают поведение горожан, межрегиональные отношения и другие значимые процессы. После этих предварительных замечаний опишем некоторые результаты моделирования.

Рассматривалась временная динамика СПН в Дагестане с начала 1998 г. Использовалась модифицированная базовая модель (с учетом дополнительных, за счет миграции, «перетоков» между категориями «X», «Y», «Z»). Начальное соотношение сторонников и противников позиции «X» (сепаратизма) определялась из данных социологических опросов, в качестве «маркирующего» события были взяты февральские выборы мэра Махачкалы, начальное значение скорости изменения разницы между числом сторонников и противников (наиболее трудно определяемая «стартовая» величина) подбиралась полуэмпирически.

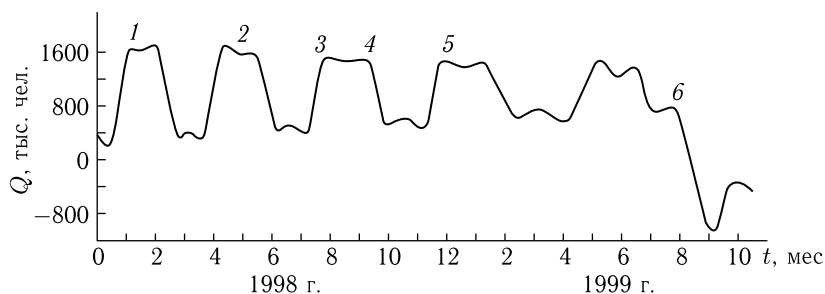


Рис. 1. Динамика социально-политической напряженности в Дагестане. По оси абсцисс отложено время в месяцах, по оси ординат — разница в числе сторонников и противников сепаратизма в тыс. человек

На рис. 1 показана полученная из вычислительных экспериментов временная динамика СПН. Номерам на кривой СПН отвечают хорошо известные события, приходящиеся на очередной период обострения:

«1» — борьба местных элит за кресло мэра Махачкалы (конец января—первая декада февраля),

«2» — захват здания Госсовета братьями Хачилаевыми (21 мая),

«3» — убийство муфтия (21 августа),

«4» — поход на Махачкалу (конец сентября),

«5» — убийство директора Центрального рынка Махачкалы (21 декабря),

«6» — начало боевых действий.

Первый же прогноз обострения (начало февраля) оказался весьма точным, что позволило не только убедиться в адекватности модели, но и оснастить ее дополнительными реальными данными. В дальнейшем прогнозы по модели также хорошо совпадали с реальными событиями. Динамика СПН в течение почти полутора лет носит циклический характер, повторяющиеся циклы политической самоидентификации имеют стабильный «основной» период, равный 120–130 дням. После проведения боевых действий в августе–сентябре 1999 г. динамика СПН

заметно изменяется — сторонников сепаратизма становится значительно меньше числа его противников, и исчезает ярко выраженная циклическая динамика СПН (амплитуда колебаний уменьшается, периодическая структура становится гораздо менее ощутимой). Другими словами, наступает относительная стабилизация политической обстановки в регионе (сильное внешнее вмешательство перевело систему на другую траекторию). Данный пример хорошо демонстрирует описываемые моделью режимы самоидентификации применительно к конкретному периоду эволюции дагестанского этноса.

С целью дополнительной проверки и подтверждения гипотез модели динамики группового противоборства сторонников и противников определенного курса действий и получаемых из нее выводов и прогнозов изучалась хронологическая последовательность полученных из различных источников данных о социально-политических конфликтах в регионе. Предполагалось, что такой анализ динамики конфликтов позволил бы выявить временные закономерности поведения социума (в т.ч. циклы самоидентификации), которые затем могли бы быть сопоставлены с модельными прогнозами. Следует отметить и самостоятельную ценность математической обработки временного событийного ряда для нахождения тенденции в поведении изучаемого объекта (и получения на этой основе, хотя бы, краткосрочного дополнительного прогноза).

Для решения указанной задачи необходимо выбрать некую единую меру оценки конфликтного поведения социума, т.е. решать задачу в рамках общей проблемы снижения размерности многомерного признака. Действительно, типичны ситуации, когда приходится сравнивать между собой и упорядочивать по некоторому не поддающемуся непосредственному измерению свойству ряд явлений (в т.ч. социальных). При этом общее представление о степени проявления анализируемого латентного (т.е. не поддающегося непосредственному измерению) свойства складывается как результат определенного суммирования ряда частных (и поддающихся измерению) характеристик, от которых зависит в конечном счете это свойство.

В качестве такой меры был принят уровень СПН, понимаемый как интегральный показатель опасности конфликтов произошедших в регионе в течение дня. Более опасному конфликту присваивалось по методике, изложенной ниже, более высокое численное значение в баллах, а затем эти значения суммировались по всем конфликтам, произошедшим за день (шкалирование).

Методика шкалирования заключается в следующем. Информация о каждом событии (конflikте) структурируется по ряду характеристик. В свою очередь каждая характеристика разбивается по уровням значимости, каждому из которых присваивается свое значение в баллах. Некоторые характеристики конфликта имеют прямую количественную меру (например, число людей, участвующих в конфликте, или продолжительность конфликта в днях), и шкалирование этих

характеристик проводится интервальным образом, т. е. диапазон возможных числовых значений характеристики разбивался на интервалы, и каждому интервалу чисел присваивалась своя балльная оценка, тем большая, чем выше границы интервала. Характеристики конфликта, не имеющие прямой количественной меры, подразделялись по показателям значимости методом экспертной оценки. Например, если некоторый конфликт продолжался 3 дня и в ходе его применялось легкое стрелковое оружие, то ему присваивается 3 балла по длительности и 2 балла по вооруженности (по шкале от 1 до 4). После учета остальных характеристик получаем интегральный показатель для данного конфликта. Суммированием этих показателей для всех конфликтов некоторого дня получаем итоговое значение уровня СПН за день.

Основным источником информации служили ежемесячные оперативные сводки региональных управлений МВД. Для целей данного исследования этот источник оказался наиболее значимым (сводки содержат подробное ежедневное описание событий и относительно хорошо структурированы). После их обработки по вышеуказанной методике времени получается временной ряд уровня СПН за определенный период. Графическое представление такого ряда в виде гистограммы приведено на рис. 2 на примере Дагестана. Время наблюдения — с июля 1998 г. по 6 августа 2000 г. По оси абсцисс отложены дни, отсчитанные от 1.07.98 г. (цифра 1 соответствует 1.07.1998 г.). По оси ординат отложено значение уровня СПН в баллах.

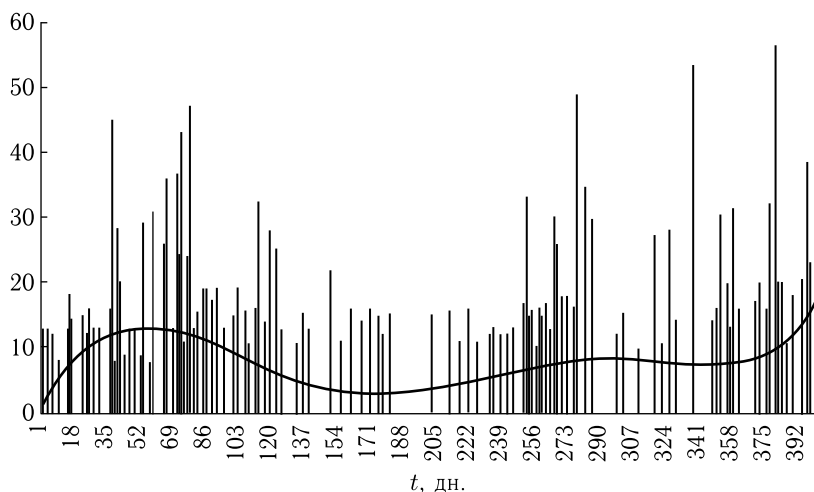


Рис. 2. Временной ряд (гистограмма) уровня СПН в Дагестане с 1 июля 1998 г. по 6 августа 2000 г.

Сопоставим полученные результаты с событийным рядом сводок. Зачерненный участок на гистограмме над числом 77, показывающий высокую временную «плотность» уровня СПН, соответствует многодневному маршу-походу в сентябре 1998 г. группы в несколько сот человек в защиту местного лидера Хачилаева. Одновременно в это же время происходили массовые многодневные выступления шоферов-«дальнобойщиков» с перекрытием автомобильной дороги федерального значения в знак протеста против поборов со стороны таможенников. Участок сгущения (увеличения интегрального уровня СПН) над числом 267 отвечает массовым выступлениям протеста в марте 1999 г. против результатов выборов в Народное собрание республики. Наконец, резкое увеличение уровня СПН в конце рассматриваемого периода отвечает ситуации, складывавшейся в Дагестане перед началом вооруженных действий в августе–сентябре 1999 г.

Роль шкалирования, построения и анализа временных рядов не ограничивается описанием (постфактум) эволюции конфликтов во времени и частичном прогнозировании нарастания их остроты (хотя сопоставление этих результатов с прогнозами по динамической модели уже весьма важно для их обоснования и повышения их точности). Соответствующие математические процедуры обработки временных рядов более подробно проясняют их внутреннюю структуру — временные тренды (например, сплошной жирной линией на рис. 2 представлен тренд, основанный на аппроксимации ряда полиномом 5-й степени и характеризующий относительно медленные изменения уровня СПН) и их различные спектральные характеристики (например, с помощью соответствующих пакетов определяются частота основных гармоник процесса и их фаза, что позволяет оперативно перенастраивать динамическую модель в соответствии с изменениями ситуации).

Тем самым комплексное использование адекватных математических моделей, соответствующих программных средств и информационных ресурсов дает возможность моделировать и прогнозировать, фактически в режиме реального времени, протекание реальных политических процессов (в данном случае, этнополитических конфликтов).

5.2.2. Моделирование динамики взаимодействующих электронов

Настоящий пункт посвящен обобщению расширенной модели подражательного поведения (см. разд. 3 главы 3) на случай двух различных по своим свойствам социальных общностей «1» и «2», участвующих в протекающем во времени процессе выбора между двумя различными позициями «X» и «Y». Поскольку речь идет о выборе, в дальнейшем, ради краткости, употребляется также слово «электорат» как синоним понятия «социальная общность».

Хотя рассматриваемые электораты как таковые, не будут идентичными, но образуют единую систему, поскольку их связывают два следующих фактора.

1. Необходимость выбора между позициями «X» и «Y» (имеется в виду, что оба электората стоят перед одной и той же дилеммой, например «за сепаратизм» и «против сепаратизма»).

2. В динамическом процессе выбора электораты определенным образом взаимодействуют между собой; тем самым процесс отличен от ситуации с двумя изолированными друг от друга общностями.

При построении модели будем следовать идеологии и порядку рассуждений, принятых в разд. 3 главы 3, выделяя при этом новые понятия, постановки, предположения и результаты, возникающие в связи с изучением поведения не одной, а двух социальных общностей.

Обозначим через $N_1(t)$, $N_2(t)$ численности способных к действию (выбору позиции) членов первого и второго электоратов соответственно.

В общем случае величины $N_1(t)$, $N_2(t)$ изменяются с течением времени в результате естественного или вызванного внешними причинами (миграция) увеличения или уменьшения за счет того, что часть электората перестает идентифицировать себя с ним либо, наоборот, появляются новые члены, готовые участвовать в решении его проблем и т. д. В отношении выбора той или иной позиции каждая из численностей $N_1(t)$, $N_2(t)$ разбивается на три категории:

$$\begin{aligned} N_1(t) &= Z_1(t) + X_1(t) + Y_1(t), \\ N_2(t) &= Z_2(t) + X_2(t) + Y_2(t), \quad t \geq t_0, \end{aligned} \quad (1)$$

где через X_i , Y_i , Z_i ($i = 1, 2$) обозначены численности сторонников позиций «X», «Y» и нейтралов соответственно.

Более детально структура численности сторонников данных позиций и нейтралов описывается соотношениями

$$\begin{aligned} X_i(t) &= X_{i,0}(t) + x_i(t); \quad Y_i(t) = Y_{i,0}(t) + y_i(t), \\ Z_i(t) &= Z_{i,0}(t) + z_i(t), \quad i = 1, 2. \end{aligned} \quad (2)$$

Здесь величины $X_{i,0} \geq 0$, $Y_{i,0} \geq 0$, $Z_{i,0} \geq 0$ — это численности «твердых» сторонников, противников и нейтралов, либо вовсе не меняющих свою позицию, либо способных изменить ее на временных масштабах значительно больших, чем рассматриваемые промежутки времени. Величины $x_i(t)$, $y_i(t)$, $z_i(t)$ представляют колеблющуюся часть общностей 1, 2, их электоральное «болото». Суммарную активность («энергию») сторонников позиций «X», «Y» в i -м электорате (число «равноценных» акций в единицу времени) будем обозначать

через $E_{i,X}(t)$, $E_{i,Y}(t)$ соответственно. Тогда

$$E_{i,X}(t) = \sum_{j=1}^{X_{i,0}(t)} E_{i,0,X}^j(t) + \sum_{j=1}^{x_i(t)} E_{i,x}^j(t), \quad (3)$$

$$E_{i,Y}(t) = \sum_{j=1}^{Y_{i,0}(t)} E_{i,0,Y}^j(t) + \sum_{j=1}^{y_i(t)} E_{i,y}^j(t), \quad (4)$$

где $E_{i,0,X}^j$, $E_{i,x}^j(t)$, $E_{i,0,Y}^j$, $E_{i,y}^j(t)$ — индивидуальные интенсивности действий членов соответствующих категорий электоратов, j — номер индивида.

Знание величин из равенств (3), (4) позволяет определить полную временную динамику рассматриваемого процесса, все его релевантные характеристики.

Например, величина

$$E_{X-Y} = \sum_{i=1}^2 [E_{i,X}(t) - E_{i,Y}(t)]$$

описывает абсолютную степень превосходства одной позиции над другой, а величина

$$\frac{E_{X-Y}}{E_{X+Y}} = \frac{\sum_{i=1}^2 [E_{i,X}(t) - E_{i,Y}(t)]}{\sum_{i=1}^2 [E_{i,X}(t) + E_{i,Y}(t)]}$$

описывает ту же степень превосходства в отношении к уровню общей активности.

Очевидно, динамика указанных (и других, относящихся к делу) величин определяется многими факторами — «внутренними» характеристиками электоратов (исторической памятью, уровнем и особенностями культуры, и т. д.), внешними воздействиями, состоянием социальной общности в начале процесса выбора и т. д. Описать и учесть все эти факторы не представляется возможным. Поэтому для построения достаточно компактной и работоспособной модели следует провести соответствующие упрощения первоначальной общей модели выбора.

В дальнейшем будем рассматривать ситуацию на сравнительно небольших временных масштабах по отношению к некоторому характерному времени. Данное сужение исходной проблемы означает, что справедливы следующие упрощающие допущения.

Предположение 1. Число твердых сторонников и противников позиций «X» и «Y» не изменяется со временем, т. е. $X_{i,0}$, $Y_{i,0}$ ($i = 1, 2$) — неотрицательные постоянные.

Предположение 2. Индивидуальные активности (интенсивности поступков) твердой части электората не изменяются со временем, т. е. $E_{i,0,X}^j$ ($i = 1, 2, 0 \leq j \leq X_{i,0}$), $E_{i,0,Y}^j$ ($i = 1, 2, 0 \leq j \leq Y_{i,0}$) — неотрицательные постоянные.

С учетом данных предположений исходная проблема сводится к изучению динамики поведения электорального болота, т. е. категорий

$$n_i(t) = x_i(t) + y_i(t) + z_i(t), \quad i = 1, 2.$$

Обозначим через $P_{i,x}^j(t)$, $P_{i,y}^j(t)$ индивидуальные потенциалы членов электората, принадлежащих соответствующим категориям (j — номер индивида), и будем считать, что у всех индивидов имеются возможности для полной реализации СПП, т. е. потенциал определяет максимально возможную (в данный момент) интенсивность действий индивида:

$$\begin{aligned} P_{i,x}^j(t) &= E_{i,x}^j(t), \quad 1 \leq j \leq x(t), \\ P_{i,y}^j(t) &= E_{i,y}^j(t), \quad 1 \leq j \leq y(t). \end{aligned}$$

Динамика СПП членов обоих рассматриваемых электоратов основана на расширенной модели подражательного поведения (см. разд. 3 главы 3). Поэтому стимулами, определяющими выбор, являются как величины разницы между сторонниками и противниками данной позиции в электоратах «1» и «2»,

$$Q_i(t) = X_i(t) - Y_i(t) = X_{i,0} - Y_{i,0} + x_i(t) - y_i(t) = Q_{i,0} + q_i(t), \quad (5)$$

$$Q_{i,0} = X_{i,0} - Y_{i,0}, \quad q_i(t) = x_i(t) - y_i(t), \quad i = 1, 2,$$

так и величина скорости изменения со временем (производной) этой разницы,

$$S_i(t) = \dot{Q}_i(t), \quad i = 1, 2, \quad (6)$$

где точкой обозначено дифференцирование по времени.

Тем самым СПП членов каждой из двух общностей формируется, условно говоря, «статической» и «ажиотажной» компонентами. Принципиальным отличием от ситуации с одним электоратом является зависимость СПП не только от характеристик «своего» электората, но также и от аналогичных характеристик электората-партнера.

Таким образом,

$$P_{i,x}^j(t) = f_{i,x}^j(Q_1(t), S_1(t), Q_2(t), S_2(t)), \quad i = 1, 2, \quad 1 \leq j \leq x(t)$$

$$P_{i,y}^j(t) = f_{i,y}^j(Q_1(t), S_1(t), Q_2(t), S_2(t)), \quad i = 1, 2, \quad 1 \leq j \leq y(t).$$

Воспользовавшись понятием типической личности для каждого из электоратов, перейдем к более простой формулировке:

$$P_1(t) = f_1(Q_1(t), S_1(t), Q_2(t), S_2(t)); \quad (7)$$

$$P_2(t) = f_2(Q_1(t), S_1(t), Q_2(t), S_2(t)). \quad (8)$$

Соотношения (7), (8) содержат четыре неизвестные величины P_1 , P_2 , Q_1 , Q_2 . При том что, вообще говоря, возможны разные постановки исходной проблемы, наиболее перспективной является следующая. На рассматриваемом промежутке времени известны, из каких-то данных или соображений, потенциалы $P_1(t)$, $P_2(t)$; требуется найти из уравнений (7), (8) две неизвестные функции $Q_1(t)$, $Q_2(t)$, а затем и все остальные искомые величины.

Для построения замкнутой модели на основе соотношений (7), (8) необходимо конкретизировать данные функции в соответствии с психофизиологическими закономерностями восприятия.

В дальнейшем будем считать справедливым следующее упрощение.

Упрощение 1. Вклады стимулов $Q_1(t)$, $S_1(t)$, $Q_2(t)$, $S_2(t)$ потенциалы $P_1(t)$, $P_2(t)$ «нелинейно аддитивны», т. е.

$$P_1(t) = \tilde{P}_1^1(Q_1) + \tilde{P}_1^2(S_1) + \tilde{P}_1^3(Q_2) + \tilde{P}_1^4(S_2), \quad (9)$$

$$P_2(t) = \tilde{P}_2^1(Q_2) + \tilde{P}_2^2(S_2) + \tilde{P}_2^3(Q_1) + \tilde{P}_2^4(S_1). \quad (10)$$

Здесь неотрицательные функции в правых частях (напомним, что потенциалы $P_1(t)$, $P_2(t)$ по своему смыслу — неотрицательные функции) суть вклады в потенциал соответствующего электората стимулов, характеризующих «собственный» вклад или вклад другого электората.

При этом функции $\tilde{P}_1^1(Q_1)$, $\tilde{P}_1^3(Q_2)$, $\tilde{P}_2^1(Q_1)$, $\tilde{P}_2^3(Q_2)$, описывающие вклады «статических» компонент в правых частях равенств (9), (10) монотонно возрастают от нуля при $Q_1 = Q_2 = 0$ до некоторого асимптотического значения при $Q_1 \rightarrow \infty$, $Q_2 \rightarrow \infty$, т. е. с увеличением стимула отклик на него увеличивается с «насыщением», изменяясь слабее и слабее. Данный эффект называется привыканием с утомлением.

В отличие от «статических» компонент потенциалов $P_1(t)$, $P_2(t)$, функции $\tilde{P}_1^2(S_1)$, $\tilde{P}_1^4(S_2)$, $\tilde{P}_2^2(S_2)$, $\tilde{P}_2^4(S_1)$, т. е. «динамические» (ажиотажные) компоненты потенциалов, растут от нуля при $S_1 = S_2 = 0$ до некоторой максимальной величины при конечных значениях S_1 , S_2 (при достижении некоторого порога восприятия), а затем монотонно стремятся к нулю при $S_1 \rightarrow \infty$, $S_2 \rightarrow \infty$. Данный эффект означает, что растущая по объему информация при превышении пропускной способности воспринимающего становится растущим по амплитуде «шумом», заглушающим основной полезный сигнал. Напомним, что в соответствии со смыслом потенциала и его компонент все введенные выше функции — четные, т. е. не меняют своего значения при смене знаков аргументов.

Математическая запись функций, обладающих описанными свойствами (эмпирически полученными зависимостями), выглядит следующим образом (один из вариантов):

$$\tilde{P}_1^1(Q_1) = \frac{a_1 Q_1^2}{Q_1^2 + b_1}, \quad \tilde{P}_1^2(S_1) = \frac{c_1 S_1^2}{S_1^4 + d_1}, \quad (11)$$

$$\tilde{P}_1^3(Q_2) = \frac{a_2 Q_2^2}{Q_2^2 + b_2}, \quad \tilde{P}_1^4(S_2) = \frac{c_2 S_2^2}{S_2^4 + d_1},$$

$$\tilde{P}_2^1(Q_2) = \frac{a_3 Q_2^2}{Q_2^2 + b_3}, \quad \tilde{P}_2^2(S_2) = \frac{c_3 S_2^2}{S_2^4 + d_3},$$

$$\tilde{P}_2^3(Q_1) = \frac{a_4 Q_1^2}{Q_1^2 + b_4}, \quad \tilde{P}_2^4(S_1) = \frac{c_4 S_1^2}{S_1^4 + d_4}. \quad (12)$$

Здесь все параметры a_i , c_i неотрицательны, а b_i , d_i строго положительны. Заметим, что все они имеют ясную содержательную интерпретацию.

Так, параметры a_i равны предельным значениям (при $Q_1 \rightarrow \pm\infty$, $Q_2 \rightarrow \pm\infty$) соответствующих «статических» (монотонно растущих с ростом $|Q_1|$, $|Q_2|$) компонент потенциалов P_1 , P_2 , величины $2a_i$ равны значениям их вторых производных в нуле, а величины $\pm\sqrt{b_i/3}$ и $\pm 3\sqrt{3b_i}/8a_i$ равны координатам точек перегиба потенциалов и значениям их первых производных по аргументам S_1 , S_2 в этих точках, соответственно (максимальным по модулю значениям первых производных во всей области определения потенциалов). В свою очередь параметры $\pm d_i$ в соответствующих «ажитажных» компонентах потенциалов P_1 , P_2 определяют координаты максимумов этих компонент (причем $S_{i,\max} = d_i^{1/4}$, а комбинации $c_i/(2\sqrt{d_i})$ равны максимальным значениям компонент).

Дальнейшая конкретизация общей модели связана со следующим

Упрощением 2. Потенциалы $P_1(t)$, $P_2(t)$ не изменяются со временем, то есть $\dot{P}_1(t) = \dot{P}_2(t) = 0$.

Данное упрощение означает, что рассматривается не общий, но достаточно широкий и важный класс процессов, имеющих, образно говоря, мобилизационный военный характер. В этой ситуации потенциалы электоратов «1», «2» в течение некоторого времени «возбуждены» соответствующими событиями, и уровень этого возбуждения поддерживается постоянным. Очевидно, что система уравнений (11) в этом случае становится автономной.

Наконец, введем заключительное упрощение, означающее, что влияние каждого из электоратов «1», «2» на своего партнера осуществляется в основном по каналу, «реагирующему» на величины самих численностей $Q_1(t)$, $Q_2(t)$, в то время как влиянием скоростей $\dot{Q}_1(t)$, $\dot{Q}_2(t)$ на их изменение можно пренебречь. Иными словами, для электоратов, осуществляющих выбор, важна, прежде всего, «статическая» компонента потенциала электората-партнера. Итак, упростим.

Упрощение 3. Вклад ажитажной (динамической) компоненты потенциала некоторого электората в СПП типической личности другого электората незначителен: $c_2 = c_4 = 0$.

С учетом данного упрощения получаем из равенств (9)–(12) следующие выражения для потенциалов:

$$P_1 = \frac{a_1 Q_1^2}{Q_1^2 + b_1} + \frac{a_2 Q_2^2}{Q_2^2 + b_2} + \frac{c_1 (dQ_1/dt)^2}{(dQ_1/dt)^4 + d_1}, \quad (13)$$

$$P_2 = \frac{a_3 Q_2^2}{Q_2^2 + b_3} + \frac{a_4 Q_1^2}{Q_1^2 + b_4} + \frac{c_3 (dQ_2/dt)^2}{(dQ_2/dt)^4 + d_3}. \quad (14)$$

Динамическая система 2-го порядка (13), (14) с начальными значениями

$$Q_1(t_0) = Q_1^0, \quad Q_2(t_0) = Q_2^0 \quad (15)$$

представляет собой замкнутую математическую модель рассматриваемого процесса, полученную в рамках сформулированных предположений и упрощений. При заданных значениях постоянных потенциалов P_1 , P_2 и прочих входящих в уравнения параметров из начальной задачи (13)–(15) однозначно (при соответствующих условиях гладкости) находится искомое решение — функции $Q_1(t)$, $Q_2(t)$.

5.2.3. Динамика взаимодействия электоратов вблизи положения равновесия

Рассмотрим систему уравнений (13), (14) применительно к важному классу ситуаций, когда выполняются неравенства

$$|Q_1| \ll \min(\sqrt{b_1}, \sqrt{b_4}), \quad |Q_2| \ll \min(\sqrt{b_2}, \sqrt{b_3}),$$

$$|dQ_1/dt| \ll \min(\sqrt{d_1}, \sqrt{d_4}), \quad |dQ_2/dt| \ll \min(\sqrt{d_2}, \sqrt{d_3}).$$

Тогда указанные уравнения принимают вид

$$P_1 = \frac{a_1}{b_1} Q_1^2 + \frac{a_2}{b_2} Q_2^2 + \frac{c_1}{d_1} (dQ_1/dt)^2, \quad (16)$$

$$P_2 = \frac{a_3}{b_3} Q_2^2 + \frac{a_4}{b_4} Q_1^2 + \frac{c_3}{d_3} (dQ_2/dt)^2. \quad (17)$$

Уравнения (2) описывают динамику относительно малых отклонений численностей $Q_1(t)$, $Q_2(t)$ от положения $Q_1 = Q_2 = 0$, $S_1 = S_2 = 0$, т. е. от состояния равновесия. В этом состоянии (его можно называть также состоянием покоя) очевидно, потенциалы $P_1 = P_2 = 0$, а при малых отклонениях от него потенциалы относительно малы.

С целью нахождения «безразмерных» («комбинированных», «системных») параметров, от которых зависит решение (16), (17), преобразуем ее нормировкой на P_1 , P_2 к виду

$$1 = \bar{Q}_1^2 + \bar{Q}_2^2 + (d\bar{Q}_1/d\bar{t})^2, \quad (18)$$

$$1 = \bar{\alpha}_1 \bar{Q}_2^2 + \bar{\alpha}_2 \bar{Q}_1^2 + \bar{\alpha}_3 (d\bar{Q}_2/d\bar{t})^2. \quad (19)$$

Здесь \bar{Q}_1 , \bar{Q}_2 — перенормированные численности электоратов, \bar{t} — перемасштабированное время, даваемые формулами

$$\bar{Q}_1 = Q_1 \sqrt{a_1/(P_1 b_1)}, \quad \bar{Q}_2 = Q_2 \sqrt{a_2/(P_1 b_2)}, \quad \bar{t} = t \sqrt{a_1 d_1/(b_1 c_1)}. \quad (20)$$

Безразмерные системные параметры, входящие в уравнение (19), равны, соответственно,

$$\bar{\alpha}_1 = \frac{P_1 a_3 b_2}{P_2 a_2 b_3}, \quad \bar{\alpha}_2 = \frac{P_1 a_4 b_1}{P_2 a_1 b_4}, \quad \bar{\alpha}_3 = \frac{P_1 a_1 b_2 c_3 d_1}{P_2 a_2 b_1 c_1 d_3}. \quad (21)$$

«Размерности» основных входящих в модели (13)–(21) величин естественно, согласуясь с их смыслом, определять из следующих

соображений: искомые величины Q_1 , Q_2 измеряются числом людей (разницей между числом сторонников и противников выбираемой позиции), потенциалы P_1 , P_2 измеряются числом поступков, совершаемых типической личностью электората в единицу времени, время понимается в обычном смысле. Тем самым в рассматриваемых моделях естественно выбрать следующие три независимые основные размерности — время, число людей и число поступков. Через них легко определяются размерности всех фигурирующих величин.

В дальнейшем черточки над величинами в (18), (19) будем опускать, записав эту систему в виде

$$1 = Q_1^2 + Q_2^2 + (dQ_1/dt)^2, \quad (22)$$

$$1 = \alpha_1 Q_2^2 + \alpha_2 Q_1^2 + \alpha_3 (dQ_2/dt)^2. \quad (23)$$

Разрешив уравнения (22), (23) относительно производных, получим для искомым функций $Q_1(t)$, $Q_2(t)$ нелинейную автономную динамическую систему второго порядка

$$\begin{aligned} \frac{dQ_1}{dt} &= \pm \sqrt{1 - Q_1^2 - Q_2^2}, \\ \frac{dQ_2}{dt} &= \pm \frac{1}{\sqrt{\alpha_3}} \sqrt{1 - \alpha_1 Q_2^2 - \alpha_2 Q_1^2}, \quad t > 0. \end{aligned} \quad (24)$$

Система (24) содержит в себе четыре варианта (т.е. четыре варианта модели (22), (23)), отвечающие четырем разным комбинациям знаков «+» и «-» в правых частях ее уравнений (знаки, в свою очередь, отвечают возрастанию или убыванию решений с течением времени). Очевидно, что в силу симметрии достаточно изучить лишь один случай из четырех (например, комбинацию знаков «+», «+»), а все остальные «ветви» решения получаются из него соответствующей заменой знаков перед функциями $Q_1(t)$, $Q_2(t)$.

Разделив второе уравнение (24) на первое, получим вспомогательное уравнение первого порядка

$$\frac{dQ_2}{dQ_1} = \pm \frac{1}{\sqrt{\alpha_3}} \frac{\sqrt{1 - \alpha_1 Q_2^2 - \alpha_2 Q_1^2}}{\sqrt{1 - Q_1^2 - Q_2^2}}, \quad (25)$$

описывающее поведение решения в плоскости состояний — фазовой плоскости (Q_1, Q_2) . Оно содержит два варианта, отвечающих знакам «+» и «-», т.е. имеет две ветви решения, достаточно изучить лишь одну из ветвей, например соответствующую знаку «+».

Для получения подробной временной динамики процесса выбора проведем совместное исследование (24) и (25), изучив прежде всего поведение решения на фазовой плоскости (Q_1, Q_2) .

Как легко видно из (25), область допустимых значений (ОДЗ) определяется из условия неотрицательности подкоренных выражений

в правой части. В терминах плоскости (Q_1, Q_2) ОДЗ представляет собой область пересечения единичного круга (с границей), задаваемой неравенством

$$Q_1^2 + Q_2^2 \leq 1, \quad (26)$$

и площадь внутри эллипса (с границей), задаваемого уравнением

$$\alpha_1 Q_2^2 + \alpha_2 Q_1^2 \leq 1. \quad (27)$$

В зависимости от значений параметров α_1, α_2 возможны различные варианты взаимного расположения фигур, ограниченных указанными линиями (см. упражнение 2). Мы рассмотрим лишь некоторые, наиболее содержательные из них.

Однако предварительно необходимо сделать важное разъяснение по вопросу о том, как следует понимать решение $Q_1(t), Q_2(t)$ системы (25) (и уравнения (25)). По своему смыслу функции $Q_1(t), Q_2(t)$ ограничены и непрерывны. Что же касается производных $dQ_1/dt, dQ_2/dt$, то они, вообще говоря, могут претерпевать разрыв в какие-то моменты времени в ситуациях, когда продолжить решение необходимо (процесс выбора рассматривается при всех $t > 0$), но невозможно при сохранении непрерывности этих производных (другими словами, допускается решение в слабом смысле). При этом по обе стороны точек разрыва должны выполняться уравнения (24). Таким образом, на разрывах производные $dQ_1/dt, dQ_2/dt$ меняют лишь свой знак, но не абсолютную величину (симметричное «отражение» функций $Q_1(t), Q_2(t)$ относительно точек разрыва).

Итак, рассмотрим случай, когда параметры уравнений (24), (25) удовлетворяют неравенствам $\alpha_1 < 1, \alpha_2 < 1$ (при этом будем считать, что в правой части уравнения (25) стоит знак «+», т.е. производная dQ_2/dQ_1 неотрицательна). Геометрически это означает, что круг (26) полностью находится внутри эллипса (27), см. рис. 3.

Область допустимых значений — это круг (26), сплошные линии внутри круга соответствуют интегральным кривым. Вдоль них функция $Q_2(Q_1)$ монотонно возрастает во всей ОДЗ. Ее производная бесконечна на окружности $Q_1^2 + Q_2^2 = 1$, на оси Q_1 она монотонно изменяется от $+\infty$ при $Q_1 = \pm 1$ до $1/\sqrt{\alpha_3}$ при $Q_1 = 0$. На оси Q_2 производная dQ_2/dQ_1 также монотонно изменяется в диапазоне от $+\infty$ при $Q_2 = \pm 1$ до $1/\sqrt{\alpha_3}$ при $Q_2 = 0$.

Стрелки на рис. 3 обозначают движение по интегральным кривым с течением времени. При этом по достижении максимальных или минимальных значений функций $Q_1(t), Q_2(t)$ происходит «отражение» от окружности так, что $Q_1^2 + Q_2^2 = 1$. Если обозначить момент отражения через t^* , то производная dQ_1/dt , равная нулю в этой точке, непрерывна, а производная dQ_2/dt имеет разрыв такой, что $(dQ_2/dt)|_{t=t^*+0} = -(dQ_2/dt)|_{t=t^*-0}$. Тем самым знаки производных по времени меняются на противоположные, и та же фазовая кривая проходится при $t > t^*$ в обратном направлении вплоть до следующего

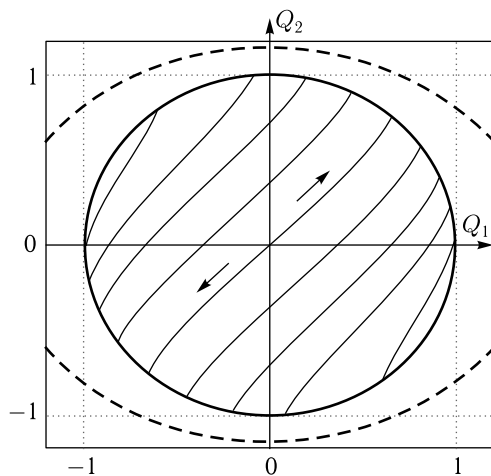


Рис. 3. Фазовый портрет системы (24) при $\alpha_1 < 1$, $\alpha_2 < 1$

отражения, которое происходит на противоположной стороне окружности. В целом, имеет место незатухающий колебательный процесс с отражениями. Его механической аналогией является процесс колебания маятника, зажатого между двумя стенками (симметричными относительно точки подвеса) и поочередно отражающегося от каждой из этих стенок.

Другой важный случай представлен на рис. 4: здесь $\alpha_1 < 1$, $\alpha_2 > 1$.

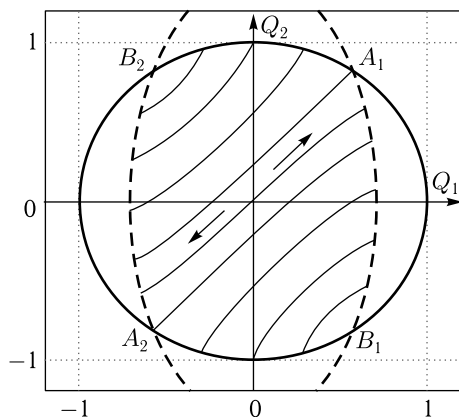


Рис. 4. Фазовый портрет системы (24) при $\alpha_1 < 1$, $\alpha_2 > 1$

Границей ОДЗ являются части границы круга (жирные сплошные линии — верхняя и нижняя границы) и эллипса (жирные штриховые линии — правая и левая границы). Координаты Q_{1I} , Q_{2I} точек пе-

ресекаются A_1, A_2, B_1, B_2 единичной окружности и эллипса равны, соответственно,

$$Q_{1I} = \pm \sqrt{\frac{1 - \alpha_1}{\alpha_2 - \alpha_1}} < 1, \quad Q_{2I} = \pm \sqrt{\frac{\alpha_2 - 1}{\alpha_2 - \alpha_1}} < 1.$$

При $|Q_1| < |Q_{1I}|$ граница ОДЗ — часть окружности, при $|Q_1| < |Q_{1I}|$ — часть эллипса.

Решение системы (24) имеет (как и в случае рис. 3) колебательный характер: каждая фазовая кривая проходит поочередно то в одном, то в другом направлении. Отличие от случая рис. 3 заключается в том, что здесь на одном конце каждой траектории происходит отражение от окружности, на другом конце — от эллипса. Так, если обозначить через t_1^* момент отражения решения от окружности, а через t_2^* — момент отражения от эллипса, то при $t = t_1^*$ производная dQ_1/dt , равная нулю в этой точке, непрерывна, а производная dQ_2/dt имеет разрыв такой, что $(dQ_2/dt)|_{t=t_1^*+0} = - (dQ_2/dt)|_{t=t_1^*-0}$. И наоборот, при $t = t_2^*$ производная dQ_2/dt , равная нулю в этой точке, непрерывна, а производная dQ_1/dt имеет разрыв такой, что $(dQ_1/dt)|_{t=t_2^*+0} = - (dQ_1/dt)|_{t=t_2^*-0}$.

5.2.4. Интерпретация результатов моделирования взаимодействующих электоров

Просуммируем вытекающие из результатов второго и третьего пунктов выводы и дадим их интерпретацию.

1. Динамика численностей $Q_1(t), Q_2(t)$ взаимодействующих электоров в окрестности положения равновесия при любых допустимых входных данных модели (24) имеет колебательный, циклический характер.

2. Колебания величин $Q_1(t), Q_2(t)$ всегда происходят «синхронно», в том смысле, что их экстремумы достигаются одновременно.

3. Реализуются либо «синфазные» (функции $Q_1(t), Q_2(t)$ возрастают и убывают одновременно, их максимумы и минимумы совпадают по времени), либо противофазные (когда $Q_1(t)$ возрастает, то $Q_2(t)$ убывает, и наоборот) режимы колебаний. Принадлежность к одному из этих двух режимов определяется лишь начальным состоянием системы (точнее, знаком производной dQ_2/dQ_1 в начальный момент) и не зависит от параметров $\alpha_1, \alpha_2, \alpha_3$. Заметим, что в третьем пункте изучался случай синфазных режимов, для противофазных режимов все результаты полностью идентичны (при этом вместо функции $Q_2(t)$ следует рассматривать функцию с противоположным знаком, т.е. величину $Q_2(t)$).

4. Амплитуда и период колебаний определяются как начальными значениями $Q_1(0), Q_2(0)$, так и значениями параметров $\alpha_1, \alpha_2, \alpha_3$, причем для заданных параметров наибольшую амплитуду имеют колебания вокруг состояния $Q_1 = Q_2 = 0$ (амплитуда уменьшается до нуля при удалении от данного состояния).

5. При $\alpha_1 < 1$, $\alpha_2 < 1$ (см. рис. 3) электорат «2» более слабо, чем «1», реагирует на значения величин Q_1 , Q_2 . После достижения величиной Q_1 своего максимального по модулю значения (на границе круга) решение можно продолжить лишь симметрично, «отразив» его от окружности в сторону уменьшающихся значений Q_1 (при этом производная dQ_2/dt разрывна). Очевидно, что максимальное (формально) значение величины $|Q_2|$ (на границе эллипса) не может быть достигнуто — под действием более «сильного» первого электората разница между числом сторонников и противников выбираемой позиции для электората «2» начинает уменьшаться (или увеличиваться) «досрочно».

Из этого результата вытекает следующая важная интерпретация: динамику колебаний (их амплитуду и частоту) в основном диктует «сильный», т. е. более восприимчивый к величине стимулов Q_1 , Q_2 электорат. Тем самым возникает конструктивно определяемое понятие ведущего и ведомого электората в процессе выбора ими позиции.

При $\alpha_1 < 1$, $\alpha_2 < 1$ (рис. 3) ведущим является первый электорат, при $\alpha_1 > 1$, $\alpha_2 > 1$ — второй (см. упражнение 4), а в случаях $\alpha_1 > 1$, $\alpha_2 < 1$, $\alpha_1 < 1$, $\alpha_2 > 1$ (рис. 4), электораты занимают ведущее место попеременно.

Упражнения

1. Какой содержательный смысл имеет система (13), (14) в случае $a_2 \ll 1$, $a_4 \ll 1$? То же — в случае $c_1 \ll 1$, $c_3 \ll 1$.
2. Перечислить все возможные способы взаимного расположения круга (26) и площади внутри эллипса (27).
3. Найти решение системы (24) в случае $\alpha_1 = \alpha_2 = 1$.
4. Построить фазовый портрет системы (24) в случае $\alpha_1 > 1$, $\alpha_2 > 1$. Дать результатам содержательную интерпретацию.

Комментарии и библиографические примечания

1. Политическая самоидентификация дагестанского этноса в конце 1990-х гг. исследована в работе [1], см. также главу VII монографии [2] и статьи [3, 4].

2. Модель динамики взаимодействующих электоратов построена в работе [5], ее исследование проведено в [6].

Литература

1. Михайлов А.П., Шведовский В.А. Математическое моделирование региональных конфликтов в контексте глобализации (на примере Северного Кавказа) // Вестник Московского университета. Серия 18: Социология и политология. 2005. Т. 18, № 1. С. 186–197.
2. Михайлов А.П. Моделирование системы «власть–общество». М. ФИЗМАТЛИТ, 2006. 144 с.

3. Михайлов А. П., Шведовский В. А., Маслов А. И., Ковалев В. Ф. Обобщенная модель электорального поведения и ее применение к изучению этнополитических конфликтов // Математическое моделирование. 2003. Т. 15. С. 39–56.
4. Шведовский В. А. Динамическая модель этнополитического конфликта: построение, возможности и результаты применения // Сб. Математическое моделирование социальных процессов. М.: МГУ. 2000. Вып. 2. С. 31–37.
5. Шведовский В. А., Михайлова П. А. Построение модели взаимодействия электоратов // Математическое моделирование. 2008. Т. 20, № 7. С. 107–118.
6. Шведовский В. А., Михайлова П. А. Динамика взаимодействующих электоратов в окрестности положения равновесия // Математическое моделирование. 2009. Т. 21, № 7. С. 43–54.

5.3. Основы моделирования системы «власть–общество»

В современном мире возрастает роль гуманитарных взаимодействий, когда люди все больше и осмысленней участвуют в разработке принимаемых государством и обществом «правил игры», основанных на доверии, честности и справедливости. Одно из проявлений этого процесса — непосредственное влияние на власть не только общества в целом, но и каждого из его членов. Это делает крайне необходимым изучение моделей режимов функционирования власти, находящейся в диалоге и взаимодействии как с обществом, так и с отдельными гражданами.

Настоящий раздел, как и два следующих, посвящен исследованию системы «власть–общество». Модель имеет вид дифференциального уравнения в частных производных параболического типа. При этом роль «пространственной» координаты играет переменная, отвечающая положению инстанции в иерархии, а сама неизвестная функция описывает объем власти, реализуемой инстанциями. Изменение этой функции с течением времени происходит вследствие как процессов внутреннего перераспределения власти между инстанциями, так и реакции общества на текущий уровень власти. Тем самым модель описывает самоорганизацию системы «власть–общество» с возникновением неоднородной пространственной структуры.

В первом пункте введены основные понятия и предпосылки и представлена базовая модель. Важнейший частный случай — модель правовой системы — анализируется во втором пункте уравнение модели в этом случае является линейным. Показано, что в условиях определенного согласования параметров, система не выходит из правового поля, если она в нем находится ввиду начального условия или своей предыстории. При нарушении этих условий распределение власти может выйти за границы правового поля. Дается интерпретация этого феномена, обсуждаются способы его парирования. Далее, в третьем пункте, модель используется для изучения так называемых *анархических*

и тоталитарных ловушек. Рассматриваются политологические сценарии, при которых система выходит из правового поля под действием собственных, внутренних сил, вырабатываемых в специфической структуре общественных настроений.

5.3.1. Построение модели

Изложение модели начнем с введения используемых в ней терминов.

Иерархия — упорядоченная по старшинству, по подчиненности совокупность институтов (инстанций, постов, должностей), наделенных властными полномочиями от имени государства (т. е. по конституции, законам, уставам, нормам, инструкциям, правилам).

Под властными полномочиями какой-либо инстанции подразумевается возможный уровень (сила, степень) ее влияния на жизнь и поведение остальных партнеров, составляющих систему «власть–общество». Считается, что объем властных полномочий любого института можно, в принципе, оценить с помощью некоторого совокупного критерия, включающего в себя формальное положение инстанции в иерархии, величину находящихся в ее распоряжении людских, финансовых, материальных, информационных, интеллектуальных, законодательных, распорядительных и иных видов ресурсов, размер и местоположение контролируемой ею территории, престижность в глазах общественного мнения и экспертов и т. д.

Для целей макроmodellирования, изучения принципиальных свойств властных структур и сравнительного анализа важно, что задача такого рода может быть, вообще говоря, решена, причем на количественной основе. При этом нет необходимости вводить какие-либо абсолютные единицы измерения властных полномочий — достаточно принять полномочия самого старшего звена за единицу (или 100%), тогда полномочия остальных звеньев будут выражаться в долях (или процентах) по отношению к полномочиям высшего иерарха. Известно, что в практической деятельности конкретных иерархий чиновники весьма неплохо оценивают относительную весомость тех или иных должностей. Всюду далее подразумевается, что существует числовая характеристика, характеризующая количество властных полномочий того или иного органа власти.

Дальнейшее уточнение понятия власти связано с максимальными и минимальными властными полномочиями. Первое описывает то, что именно орган власти максимально может предпринять в какой-либо ситуации в соответствии с законодательством. Минимальные полномочия фиксируют действия, которые инстанция обязана, должна осуществлять всегда. Хорошей количественной иллюстрацией этих понятий служат статьи Уголовного кодекса, определяющие за один и тот же вид преступления максимальные и минимальные сроки наказания (скажем, от 3 до 5 лет). Другими словами, максимальные и минимальные полномочия дают законные верхние и нижние границы власти каждой

инстанции, определяемые конституцией, системой прецедентов, традициями и т. д. (границы правового поля).

Наконец, еще одно используемое в модели понимание власти относится к фактически осуществляемому в данный момент данной инстанцией уровню властного влияния (величине власти). Очевидно, что верхние и нижние границы власти достигаются не всегда и не всюду. Пусть, например, суды по некоторой статье Уголовного кодекса, предусматривающей лишение свободы сроком от 3 до 5 лет, вынесли за достаточно длительное время много приговоров со средним сроком 4 года. Тогда реализованная ими по данной статье власть составила 80 % от их максимальных и 133 % от их минимальных полномочий. Для нахождения именно этой заранее неизвестной (искомой) величины (текущего распределения власти, или профиля власти в иерархии) и строится математическая модель.

Как же найти распределение власти в иерархии, какова его динамика, и от чего она зависит? Для ответа на эти вопросы рассмотрим главного партнера властной структуры. Гражданское общество — «негосударственная» часть системы «власть–общество», не обладающая непосредственно официальной властью. Сюда относятся отдельные граждане и их профессиональные, культурные и прочие объединения, семьи и частные предприятия. Принципиальное отличие гражданского общества от структур власти состоит в том, что его элементы не могут напрямую приказывать от имени государства ни друг другу, ни (тем более) звеньям иерархии.

Реакция гражданского общества — ответ (положительный, отрицательный или безразличный) гражданского общества на действия того или иного института власти в тот или иной момент времени (путем выборов, референдумов, через СМИ, с помощью митингов, забастовок). Под ответом общества подразумевается его сопротивление (протест) или его одобрение (поддержка). Под действиями иерархии понимается активность или неактивность ее звеньев, т. е. уровень наличной власти в инстанциях (о нем уже шла речь выше).

Основные модельные предположения формулируются следующим образом.

1. Все партнеры в системе «власть–общество» законопослушны (законы соблюдаются, налоги платятся, приказы выполняются).

2. Знак и величина (интенсивность) обмена властью между иерархической структурой и конституцией определяются реакцией системы (суммой реакций «граждан» и «чиновников»).

Имеется в виду, что если в данный момент времени выражена отрицательная (суммарная) реакция на действия данного звена иерархии, то это побуждает инстанцию уменьшить осуществляемую ею власть (причем тем больше, чем сильнее протест), как бы отложив про запас некоторое ее количество в конституционный резервуар. Положительная же реакция стимулирует инстанцию увеличить, в меру продемонстрированной поддержки, уровень реализуемой ею власти,

взяв необходимые ресурсы из законодательства. Такое понимание подразумевает, конечно же, что реакция выражена в нужный момент и правильно адресована.

3. *Основной поведенческий постулат.* В иерархии власть может передаваться только от инстанций с большей текущей властью к инстанциям с меньшей текущей властью (причем скорость передачи тем больше, чем больше разница между значениями текущей власти в инстанциях).

Основной поведенческий постулат касается взаимоотношений уже внутри самой властной структуры. Здесь любая инстанция принимает к исполнению какие-то властные распоряжения, идущие от старших звеньев и в свою очередь передает какие-то приказы младшим звеньям. При этом происходит некоторое перераспределение власти между ступенями иерархии (напомним, что речь идет не об изменении властных полномочий, а об изменении реально осуществляемого, текущего уровня власти).

Совокупность передаваемых распоряжений формирует своего рода поток власти между инстанциями — количество власти, получаемое в единицу времени младшей ступенью от старшей ступени. Описанный механизм передачи властных распоряжений отвечает иерархической процедуре приказов по команде. В модели предусмотрен также хорошо известный и часто применяемый механизм команд «через голову» (и введена соответствующая поведенческая характеристика), когда старший приказывает не только непосредственному подчиненному, но части более младших звеньев или всем им сразу.

4. *Локальный «закон сохранения власти»:* скорость изменения величины власти в инстанции определяется разностью протекающих через нее потоков власти и интенсивностью реакции системы

Рассмотрим теперь принятую в модели схему взаимодействий партнеров в системе «власть–общество» (см. рис. 1). Она основана на следующих положениях.

1. Гражданское общество непосредственно или через своих представителей принимает (дополняет, изменяет) законодательство или воздействует на властные структуры иными способами. Оно, таким образом, выступает как источник («заказчик», «хозяин») власти для государственной иерархической структуры (влияя на ее поведение выражением своей реакции).

2. Иерархия рассматривается не сама по себе, а как открытая система, взаимодействующая с гражданским обществом. Законодательство (в широком понимании, включающем конституцию, законы, правила, уставы и т. д.) служит для иерархии своеобразным резервуаром, из которого ее звенья могут по мере необходимости либо черпать дополнительные порции власти, либо возвращать обратно ее избыточные доли. Тем самым между конституцией и иерархией осуществляется как бы обмен властью с гражданским обществом — учредителем конституции.

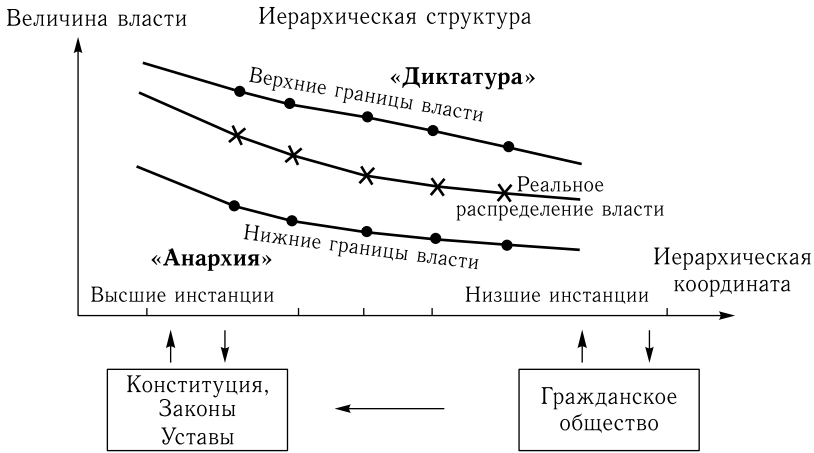


Рис. 1. Схема взаимодействия партнеров в системе «власть–общество»

3. Внутри самой иерархии происходит перераспределение текущей власти между составляющими ее инстанциями в соответствии с принятыми в иерархии взаимодействиями.

4. По отношению к гражданскому обществу властные структуры выступают как принуждающие или освобождающие институты, вводящие и отменяющие те или иные ужесточения или послабления.

Уравнение модели «власть–общество» получено на основе вышеизложенных соображений и имеет вид

$$\frac{\partial p}{\partial t} = \frac{\partial p}{\partial x} \left[\kappa \left(p, \frac{\partial p}{\partial x}, x, t \right) \frac{\partial p}{\partial x} \right] + F(p, x, t) + \int_0^1 \chi(p(x', t), p(x, t), x', x) [p(x', t) - p(x, t)] dx' \quad (1)$$

Здесь уровень инстанции в иерархии обозначен через x , при этом $x = 0$ соответствует высшему уровню иерархии, $x = 1$ — низшему. Количество власти, реализуемой инстанцией x в момент времени t , обозначено через $p(x, t)$. Коэффициент $\kappa(p \partial p / \partial x, x, t)$ имеет смысл величины, обратной «ответственности» инстанций иерархии, функция $F(p, x, t)$ описывает реакцию общества на текущее распределение власти. Интегральный член соответствует так называемому механизму команд «через голову», т. е. передаче власти не от начальника к непосредственному подчиненному, а к некоторым или сразу всем нижестоящим инстанциям.

Уравнение (1) рассматривается с краевыми условиями

$$\frac{\partial p}{\partial x} \Big|_{x=0} = \frac{\partial p}{\partial x} \Big|_{x=1} = 0, \quad (2)$$

имеющими смысл отсутствия потоков власти через границы иерархии, а начальное условие

$$p(x, 0) = p^0(x) \quad (3)$$

описывает исходное распределение власти в иерархии.

Условия (2) не выполняются, если над высшей инстанцией иерархии есть некий «внешний» орган, от которого (или к которому) власть может перетекать от этой высшей инстанции. Например, марионеточные правители могут получать весомую поддержку от внешних сил, а руководители некоторых других стран — наоборот, тратить часть своего властного ресурса на борьбу с этими силами.

Задача (1)–(3) — это и есть модель «власть–общество». Ее входными данными являются: а) начальное распределение власти и б) конкретный вид реакции гражданского общества. Модель будет полностью определена, если эти входные данные будут измерены экспертными или социологическими методами.

Модель математически реализует описанную выше схему взаимодействий в системе «власть–общество», которая предстает как самоорганизующаяся, неоднородная, с системой прямых и обратных связей. Объем власти, осуществляемый любым институтом в любой момент времени, отнюдь не произволен, а является результатом совместного действия всех компонент системы (идущих по иерархической лестнице властных распоряжений, реакции общества, действующего законодательства, начального состояния системы и т. д.) и изменяется при их изменении. В математическом плане модель представляет собой начально-краевую задачу для интегро-дифференциального уравнения в частных производных параболического типа.

Построенная таким образом модель замкнута и корректна, т. е. при известных входных данных однозначно определяет решение. Под входными данными имеют в виду число инстанций в иерархии и ее топология, известные властные полномочия инстанций и поведенческие характеристики системы. Под решением понимается распределение (профиль) власти в иерархии, знание которой позволяет вычислить потоки власти и иные, относящиеся к делу, характеристики системы.

Всюду далее будем рассматривать модель при следующих упрощающих предположениях: пренебрежем механизмом команд «через голову» и будем считать коэффициент $\kappa(p, \partial p / \partial x, x, t)$ константой (далее будем обозначать $\varepsilon = \sqrt{\kappa}$), а реакцию общества — не зависящей от времени (последнее предположение можно считать справедливым, если динамика власти рассматривается на протяжении временного отрезка, меньшего, чем характерное время эволюции общественного сознания). Уравнение (1) принимает тогда вид

$$\frac{\partial p}{\partial t} = \varepsilon^2 \frac{\partial^2 p}{\partial x^2} + F(p, x). \quad (4)$$

Одна из наиболее естественно формализуемых систем «власть–общество» — правовая система. Поэтому в следующем пункте рассматривается простейшая конкретизация данной модели, содержательно отвечающая именно этому случаю.

5.3.2. Динамика в рамках правовой системы: базовая модель в линейном случае

Система «власть–общество» называется правовой, если ее реакция на действия любой инстанции иерархии всегда направлена на удержание распределения власти в рамках предписанных ей полномочий (подразумевается, что максимальные и минимальные полномочия иерархической структуры соответствуют праву в общепринятом смысле этого понятия). Подобный тип реакции отвечает правовому общественному сознанию.

Другими словами, правовая система отрицательно реагирует на превышение уровня власти относительно разумного, с точки зрения общества, интервала значений, усиливая реакцию с увеличением этого отклонения (при отклонениях, уменьшающих средний уровень власти, реакция обратная, т. е. система требует увеличения объема власти, реагируемой иерархией).

В самом простом случае этот «разумный с точки зрения общества интервал значений» сводится просто к кривой на плоскости (x, p) . Это значит, что для каждой инстанции x существует некоторое желательное, с точки зрения общества, количество власти $p_0(x)$. Далее будем называть эту функцию идеальным распределением власти. Если в какой-то момент времени текущее количество власти $p(x, t)$ больше, чем $p_0(x)$, то реакция общества воздействует на иерархию таким образом, чтобы уменьшить $p(x, t)$, и наоборот. Таким образом, $p_0(x)$ является *единственным* желательным, с точки зрения общества, и потому устойчивым распределением власти. Это имеет место, например, в случае линейной (по p) функции реакции общества $F(p, x)$.

Отметим, что при линейной реакции общества устойчивость желательного, с точки зрения общества, распределения носит глобальный характер (устойчивость по отношению к математически сколь угодно большим возмущениям). В противоположность этому при некоторых деформациях общественно-политического сознания (см. ниже об анархической и авторитарной ловушках) устойчивость является лишь локальной.

В модели (2)–(4) сформулированные выше представления реализуются заданием соответствующей функции $F(p, x)$. На рис. 2 качественно изображены минимальные и максимальные (соответственно, $p_1(x, t)$ и $p_2(x, t)$) властные полномочия инстанций иерархии в некоторый момент времени. По своему смыслу $p_1(x, t)$ и $p_2(x, t)$ — монотонно убывающие по x положительные функции, причем $p_2(x, t) > p_1(x, t)$, $0 \leq x \leq l$, $t \geq 0$. Область, лежащая между кривыми $p_1(x, t)$, $p_2(x, t)$, — правовое поле, область сверху от $p_2(x, t)$ условно называется «диктатурой»,

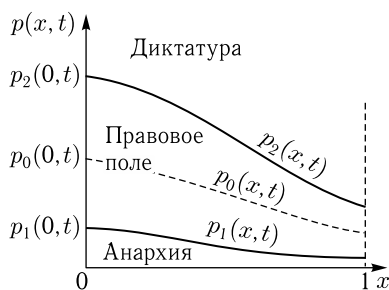


Рис. 2. Взаимное расположение минимальных полномочий $p_1(x, t)$, идеального распределения власти $p_0(x, t)$ и максимальных полномочий $p_2(x, t)$

область снизу от кривой $p_1(x, t)$ — «анархией». Искомое решение, распределение власти $p(x, t)$, может, вообще говоря, в разные моменты времени либо полностью находиться в правой области, либо частично или полностью быть вне нее.

Итак, правовая система реагирует на отклонение власти от идеального значения. В количественном плане эта реакция зависит не только от разности $p(x, t) - p_0(x)$, но, вообще говоря, и от времени t , может быть различной для разных звеньев иерархии (координата x), и может включать более тонкую зависимость от величины $p(x, t)$.

Перейдем к изучению стационарных (не зависящих от времени) распределений власти в правовой системе. При этом ограничимся линейной функцией реакции общества

$$F(p, x) = k_1(p_0(x) - p). \quad (5)$$

Здесь коэффициент $k_1 > 0$ характеризует реакцию общества в аспекте того, насколько сильно общество реагирует на отклонение распределения власти от «идеального» распределения. Желательное, с точки зрения общества, распределение власти $p_0(x)$ и властные полномочия $p_1(x)$, $p_2(x)$ не меняются со временем; примем для простоты, что они линейно убывают с ростом координаты x , т. е.

$$p_0(x) = H - kx, \quad H > 0, \quad p_1(x) = (1 - \alpha)p_0(x), \quad p_2(x) = (1 + \alpha)p_0(x). \quad (6)$$

Здесь параметр $k > 0$ описывает убывание функции $p_0(x, t)$ с ростом координаты x . Значение $k = 0$ соответствовало бы ситуации, при которой все властные инстанции (президент, председатель правительства и т. д. вплоть до самого мелкого чиновника) должны, с точки зрения общества, реализовывать одинаковое количество власти. Чем больше значение k , тем больше разность между объемами власти, реализуемыми инстанциями различного уровня.

При этом линейность функции $p_0(x, t)$ соответствует равномерному убыванию «идеального» распределения власти по иерархии: самая старшая инстанция реализует больше власти, чем вторая инстанция,

на столько же, на сколько вторая инстанция больше, чем третья. Величина

$$\frac{H - k}{H} = \frac{p^0(1)}{p^0(0)} = \frac{p_1(1)}{p_1(0)} = \frac{p_2(1)}{p_2(0)}$$

представляет собой «перепад» полномочий, т. е. отношение полномочий низших и высших инстанций иерархии.

Продолжим описание содержательного смысла коэффициентов в выражениях (6). Параметр α ($0 < \alpha < 1$) характеризует относительную разницу между $p_0(x, t)$ и функциями $p_1(x)$, $p_2(x)$ (таким образом, здесь мы полагаем, что $p_0(x) = (p_1(x) + p_2(x))/2$, т. е. минимальные и максимальные полномочия в равной степени удалены от «идеального» распределения власти).

Итак, базовая модель правовой системы имеет вид

$$\frac{\partial p}{\partial t} = \varepsilon^2 \frac{\partial^2 p}{\partial x^2} + k_1(H - kx - p), \quad 0 < x < 1, \quad t > 0, \quad (7)$$

$$\left. \frac{\partial p}{\partial x} \right|_{x=0} = \left. \frac{\partial p}{\partial x} \right|_{x=1} = 0, \quad p(x, 0) = p^0(x). \quad (8)$$

Заметим, что модель (7), (8) линейна; для нее относительно нетрудно выписать общее решение (см. упражнения).

Проведем анализ стационарного решения. Оно находится из краевой задачи

$$\varepsilon^2 \frac{\partial^2 p}{\partial x^2} + k_1(H - kx - p) = 0, \quad 0 < x < 1, \quad \left. \frac{\partial p}{\partial x} \right|_{x=0} = \left. \frac{\partial p}{\partial x} \right|_{x=1} = 0. \quad (9)$$

Задача (9) имеет хорошо известное единственное (неотрицательное) решение, даваемое монотонно убывающей по x функцией

$$p(x) = \frac{k\varepsilon/\sqrt{k_1}}{\exp(\sqrt{k_1}/\varepsilon) - \exp(-\sqrt{k_1}/\varepsilon)} \left[(1 - \exp(-\sqrt{k_1}/\varepsilon)) \exp(x\sqrt{k_1}/\varepsilon) + (1 - \exp(\sqrt{k_1}/\varepsilon)) \exp(-x\sqrt{k_1}/\varepsilon) \right] + (H - kx). \quad (10)$$

Качественный вид этого решения показан на рис. 3.

Количество власти, реализуемой верхними инстанциями иерархии, несколько ниже, чем «идеальное». Таким образом, для них имеет место явление принижения власти. Аналогично, для низших инстанций имеет место превышение власти. Тем самым идеальное распределение власти в системе не может быть реализовано.

Рассмотрим вопрос о том, каким образом эффект принижения (и эффект превышения) власти зависит от параметров системы. Вычислим величину принижения при $x = 0$:

$$H - p(0) = - \frac{k\varepsilon/\sqrt{k_1}}{\exp(\sqrt{k_1}/\varepsilon) - \exp(-\sqrt{k_1}/\varepsilon)} \times \left[(1 - \exp(-\sqrt{k_1}/\varepsilon)) + (1 - \exp(\sqrt{k_1}/\varepsilon)) \right]. \quad (11)$$

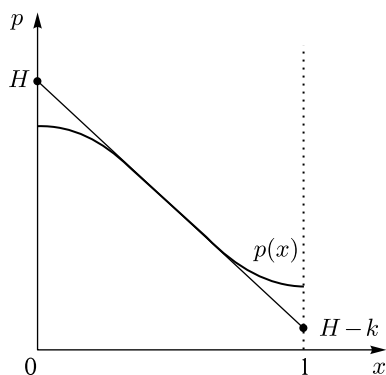


Рис. 3. Решение (10) краевой задачи (9)

Очевидно, $H - p(0) \rightarrow 0$ при $k \rightarrow 0$. Таким образом, чем меньше перепад власти между инстанциями иерархии, тем меньше принижение. Другими словами, отклонение стационарного распределения власти от «идеального» является более значительным в тех системах, в которых верхние инстанции имеют слишком много власти по сравнению с низшими.

Рассмотрим теперь, как величина $H - p(0)$ зависит от комбинации параметров $\varepsilon/\sqrt{k_1}$. Рассмотрим поведение этой величины при $\varepsilon/\sqrt{k_1} \rightarrow 0$. Имеем

$$\begin{aligned} \lim_{\varepsilon/\sqrt{k_1} \rightarrow 0} \frac{H - p(0)}{\varepsilon/\sqrt{k_1}} &= -k \lim_{\varepsilon/\sqrt{k_1} \rightarrow 0} \frac{(1 - \exp(-\sqrt{k_1}/\varepsilon)) + (1 - \exp(\sqrt{k_1}/\varepsilon))}{\exp(\sqrt{k_1}/\varepsilon) - \exp(-\sqrt{k_1}/\varepsilon)} = \\ &= -k \lim_{\varepsilon/\sqrt{k_1} \rightarrow 0} \frac{2 - \exp(-\sqrt{k_1}/\varepsilon) - \exp(\sqrt{k_1}/\varepsilon)}{\exp(\sqrt{k_1}/\varepsilon) - \exp(-\sqrt{k_1}/\varepsilon)} = k. \end{aligned}$$

Таким образом, при сильной реакции общества (т. е. при большом значении параметра k_1) эффект принижения власти уменьшается (то же относится к эффекту превышения на низшем конце иерархии).

Резюмируя результаты анализа эффектов превышения и принижения власти, получаем, что распределение власти наиболее близко к идеальному в тех системах, где высшие инстанции имеют не слишком много власти по сравнению с низшими (а), реакция общества достаточно сильна (б).

5.3.3. Анархическая и тоталитарная ловушки

Еще более сложные, причем неосознанные эволюции распределения власти в иерархиях возможны при внешне, быть может, незначительных, деформациях «сознания» системы по отношению к правовому. Они носят характер катастрофических сценариев: либо неограниченного роста властного влияния, либо, наоборот, его полного исчезновения (авторитарные и анархические ловушки). Слово «ловушка» здесь озна-

чает, что первоначально за рамки законных полномочий выходит лишь малая часть инстанций иерархии, которые со временем звено за звеном вытягивают всю властную структуру из правового поля.

Опишем основные стадии развития анархической ловушки. Она может реализовываться, например тогда, когда тип сознания граждан в основном правовой, но при значениях уровня власти (активности) инстанций ниже некоторого критического уровня становится неправовым — общество, разочарованное иерархией, стремится избавиться от ее власти. Если в некоторый (например, начальный) момент часть старших инстанций реализует меньшее количество власти, чем критическое, то этого может оказаться достаточно, чтобы с течением времени в «подкритическую» область попали все остальные звенья иерархии и произошло необратимое уменьшение уровня власти до нуля (полная анархия). Авторитарная ловушка развивается противоположным образом — волна превышения власти распространяется от младших к старшим звеньям. Заметим, что в обоих случаях выход из правового поля происходит вследствие действия социальных сил и политических механизмов, а не ввиду того, что, например кто-то, «хочет установить диктатуру». Главной причиной затягивания системы в ловушку является специфический тип сознания, при котором общество отказывается от поддержки правой системы.

Рассмотрим, как моделируются анархическая и тоталитарная ловушки в модели (2)–(4). Пусть функция $F(p, x)$ имеет вид как на рис. 4. Значения p , находящиеся в промежутке $p_{\text{crit},1} < p < p_{\text{crit},2}$, соответствуют правовому сознанию общества: его реакция направлена на то, чтобы приблизить распределение власти к «идеальному». При этом лишь в интервале $\tilde{p}_1 < p < \tilde{p}_2$ общество реагирует тем сильнее, чем больше отклонение текущего распределения власти от «идеального». Интервалы $p_{\text{crit},1} < p < \tilde{p}_1$, $\tilde{p}_2 < p < p_{\text{crit},2}$ можно условно назвать «слабоправовой областью» общественного сознания в том смысле, что чем сильнее отклонено распределение власти, тем слабее реагирует общество. Если же $p < p_{\text{crit},1}$, то разочарованное общество действует в направлении уменьшения власти, реализуемого иерархией, и имеет место

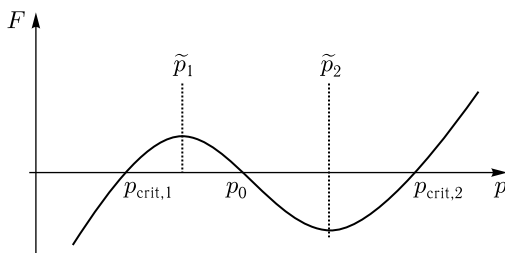


Рис. 4. Качественный вид функции $F(p, x)$, при котором возникают анархическая и тоталитарная ловушки

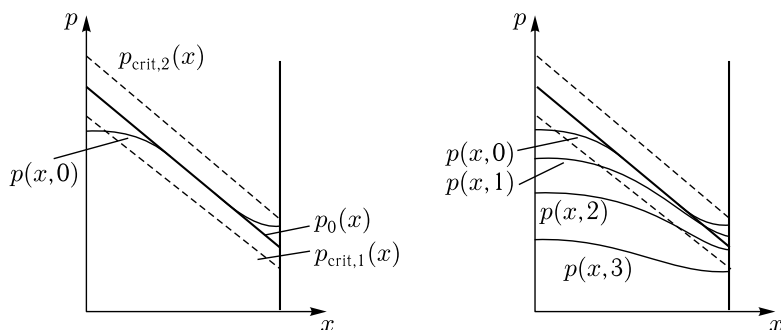


Рис. 5. Динамика анархического сценария

переход к анархии. Аналогично, при $p > p_{crit,2}$ общество поддерживает дальнейшее усиление власти — происходит переход к диктатуре.

Рассмотрим динамику анархического сценария. На рис. 5 изображены значения функции $p(x, t)$ в различные моменты времени. Левый рисунок показывает взаимное расположение функций $p_{crit,1}(x)$, $p_0(x)$, $p_{crit,2}(x)$, а также начального распределения власти $p(x, 0)$. В области высших инстанций (при малых значениях x) это начальное распределение расположено в области $p(x, 0) < p_{crit,1}(x)$. Относительно этих инстанций реакция общества направлена на их дальнейшее ослабление, т. е. она «тянет» распределение власти $p(x, t)$ вниз. Поэтому с течением времени функция $p(x, t)$ опускается вниз, вследствие чего возрастает диапазон значений x , при котором $p(x, t) < p_{crit,1}(x)$, он постепенно увеличивается. Так, на правом рисунке видно, что функция $p(x, 1)$ расположена ниже, чем $p(x, 0)$, функция $p(x, 2)$ — еще ниже. Соответственно область значений переменной x , в которой $p(x, t) < p_{crit,1}(x)$, существенно больше, чем аналогичная область для функции $p(x, 0)$. Наконец, функция $p(x, 3)$ полностью расположена ниже, чем пунктирная прямая $p_{crit,1}(x)$. Тем самым дальнейшая динамика сводится к уменьшению количества власти каждой из инстанций иерархии, что является кульминацией анархического сценария.

Аналогично развивается процесс формирования тоталитарной ловушки. Его предварительным условием является попадание инстанций низшей части иерархии в область $p(x, 0) > p_{crit,2}(x)$. Относительно этих инстанций общество влияет в направлении увеличения реализуемой ими власти и т. д. Результатом является выполнение неравенства $p(x, t) > p_{crit,2}(x)$ для всех x при достаточно больших значениях времени, что соответствует сценарию установления диктатуры.

5.3.4. Заключение

Рассмотренные в данном (и следующем) разделе сценарии весьма условны (хотя бы потому, что при заметном превышении или принижении власти модель, строго говоря, не применима, так как нарушается

предположение о законопослушности системы). Однако они демонстрируют потенциальное богатство возможных поведений распределения власти в иерархии, заложенных в модель, которая содержит различные прямые и обратные связи, нелинейности и пространственно-временные вариации характеристик объекта, а также указывают на возможность осмысленного толкования эволюции изучаемой системы.

Упражнения

1. Получить общее решение начально-краевой задачи (7), (8).
2. Исследовать эффекты принижения и превышения власти для случая слабой реакции общества. Для этого рассмотреть величину $H - p(0)$ в предельном переходе $\sqrt{k_1}/\varepsilon \rightarrow 0$.

Комментарии и библиографические примечания

1. Модель «власть–общество» была предложена в работах [1, 2]. Наиболее полно результаты отражены монографии [3].

2. Некоторым направлениям развития модели «власть–общество» посвящены следующие два раздела настоящей главы. Среди других направлений отметим исследование древовидных (пирамидальных) иерархий. Напомним, что модели настоящего раздела рассматривали иерархию как цепочку инстанций: у каждого начальника есть ровно один подчиненный. Модели пирамидальных иерархий учитывают, что начальник может иметь несколько подчиненных, так что общее количество чиновников является непостоянной (и, как правило, монотонно возрастающей) функцией иерархического уровня. В базовом случае (отсутствие механизма передачи команд «через голову», постоянство коэффициента k , отсутствие явной зависимости реакции общества от времени) уравнение модели с пирамидальной иерархией имеет вид

$$n(x) \frac{\partial p}{\partial t} = \varepsilon^2 \frac{\partial}{\partial x} \left(n(x) \frac{\partial p}{\partial x} \right) + n(x) F(p, x).$$

Здесь $n(x)$ — количество инстанций, находящихся на иерархическом уровне x . При $n(x) \equiv 1$ данное уравнение принимает форму уравнения (4). Построению и анализу модели с пирамидальной иерархией посвящены работы (4), (4) (см. также монографию [3]).

3. В тех случаях, когда количество иерархических уровней невелико, модель «власть–общество» применяется в дискретной форме. При этом вторая производная d^2p/dx^2 заменяется на разностную производную; для верхней и нижней инстанций учитываются краевые условия. Например для случая трех иерархических уровней (которые могут пониматься как федеральный, региональный и местный

уровни), система уравнений базовой модели «власть–общество» имеет следующий вид:

$$\begin{cases} \frac{dp_1}{dt} = F_1(p_1) + \kappa(p_2 - p_1), \\ \frac{dp_2}{dt} = F_2(p_2) + \kappa(p_1 - 2p_2 + p_3), \\ \frac{dp_3}{dt} = F_3(p_3) + \kappa(p_2 - p_3). \end{cases}$$

Здесь $F_i(p_i)$ — реакция общества на количество власти, реализуемое i -й инстанцией ($i = 1, 2, 3$). Дискретная форма позволяет рассматривать, в частности, иерархии сложной структуры — например, такие, которые возглавляются не единоличным первоиерархом, а дуумвиратом [6].

Литература

1. Михайлов А.П. Математическое моделирование власти в иерархических структурах // Математическое моделирование. 1994. Т.6, №6. С. 108–138.
2. Mikhailov A.P. Mathematical Modeling of Power Distribution in State Hierarchical Structures Interacting with Civil Society // Proceedings of 14th IMACS World Congress, Atlanta, USA. 1994. V.2. P. 828–830.
3. Михайлов А.П. Моделирование системы «власть–общество». М.: Наука, ФИЗМАТЛИТ, 2006. 250 с.
4. Михайлов А.П., Савельев А.В. Обоснование моделей властных иерархий через их микроописание // Математическое моделирование. 2001. Т.13., №4. С. 19–34.
5. Михайлов А.П., Ланкин Д.Ф. О конструкциях властных иерархий // Математическое моделирование. 2009. Т.21, №8. С. 108–120.
6. Михайлов А.П., Горбатилов Е.А. Базовая модель дуумвирата в системе «власть–общество» // Математическое моделирование. 2012. Т.24, №1. С. 33–45.

5.4. Система «власть–общество» в случае биполярной реакции общества

Модель «власть–общество», рассмотренная в предыдущем разделе, прилагается к анализу некоторых конкретных политологических сценариев. При этом реакция общества предполагается биполярной (амбивалентной). Другими словами, предполагается, что существует два распределения власти, каждое из которых является желательным с точки зрения общества. В силу этого оба этих распределения являются также локально устойчивыми. Одно из них называется партиципаторным распределением и соответствует сравнительно малому количеству власти, реализуемой иерархией, другое называется *распределением сильной руки* и соответствует сравнительно большому количеству власти. Показано, что при такой реакции общества суще-

ствуют также стационарные и нестационарные решения с внутренним переходным слоем: часть инстанций реализует большое количество власти (как в случае сильной руки), другая часть — малое (как в случае партиципаторного распределения). С математической точки зрения такие решения относятся к классу контрастных структур; их анализ проводится методами теории сингулярных возмущений. Показано, что результат эволюции системы (т. е. распределение власти при больших значениях t) завит не только от предпочтений общества, но и от начальных условий. Полученным решениям дана содержательная политологическая трактовка: в частности, рассмотрены процессы установления властной вертикали и формирования конфедеративного распределения, при котором формально высшие инстанции обладают сравнительно небольшой властью (в противоположность инстанциям некоторого среднего иерархического уровня).

Далее вводится понятие амбивалентной реакции общества, которое является ключевым для сценариев данного раздела. Второй пункт посвящен описанию динамического процесса формирования нестационарной контрастной структуры из начальных условий. С течением времени она движется и преобразуется (при определенных условиях) в стационарную контрастную структуру, что является предметом третьего и четвертого пунктов.

5.4.1. Биполярная реакция общества

Рассмотрим уравнение базовой модели «власть–общество» (см. уравнение (5) в разд. 5.3 настоящей главы):

$$\frac{\partial p}{\partial t} = \varepsilon^2 \frac{\partial^2 p}{\partial x^2} + F(p, x), \quad (1)$$

$$\left. \frac{\partial p}{\partial x} \right|_{x=0} = \left. \frac{\partial p}{\partial x} \right|_{x=1} = 0, \quad p(x, 0) = p^0(x). \quad (2)$$

Далее будем полагать, что существуют два устойчивых распределения власти, каждое из которых является, в некотором смысле, желательным с точки зрения общества. Обозначим их через $p_1(x)$ и $p_3(x)$, причем будем полагать эти функции достаточно гладкими и удовлетворяющими неравенству $p_1(x) < p_3(x)$ при $0 \leq x \leq 1$. В качестве конкретной функции $F(p, x)$, обладающей требуемым свойством, возьмем функцию с кубической нелинейностью. Таким образом, рассматривается уравнение

$$\frac{\partial p}{\partial t} = \varepsilon^2 \frac{\partial^2 p}{\partial x^2} - k_1(x)(p - p_1(x))(p - p_2(x))(p - p_3(x)), \quad (3)$$

где функции $k_1(x)$, $p_i(x)$, $i = 1, 2, 3$, имеют непрерывные производные, $k_1(x) > 0$, $p_1(x) < p_2(x) < p_3(x)$ при $0 \leq x \leq 1$.

Далее будем называть $p_3(x)$ распределением сильной руки, а $p_1(x)$ — партиципаторным распределением власти. (В социологии партиципаторной демократией принято называть систему, для которой

характерно более активное прямое участие граждан в политическом управлении (в противоположность системам, где они осуществляют управление лишь через выбранных представителей — депутатов и т. д.). Типичный пример партципации — распределение части местного бюджета на собрании жителей. Таким образом, реакцию общества в данном случае формирует политическая культура, основанная на сосуществовании двух взглядов на желательное распределение власти. Каждый из этих взглядов может быть представлен определенной политической партией. Мы полагаем, что избиратель в своих электоральных предпочтениях ориентируется не только на экономическую программу партии, личности ее лидеров и т. д., но и на рассматриваемый нами аспект, взгляды партии на вопрос о полномочиях властной иерархии. Более того, опыт показывает, что в предвыборной борьбе лозунги типа «укрепить властную вертикаль», «укрепить структуру власти» или «власть должна чаще советоваться с народом» и т. д. зачастую играют более заметную роль, чем соображения чисто экономического характера.

Относительно параметров модели будем также полагать, что $\varepsilon \ll \ll 1$; это условие выполняется, в частности, если властная иерархия достаточно протяженная. В силу малости ε уравнение (1) является сингулярно возмущенным.

Так как $\frac{\partial F}{\partial p}(p_1(x), x) < 0$, $\frac{\partial F}{\partial p}(p_2(x), x) > 0$, $\frac{\partial F}{\partial p}(p_3(x), x) < 0$, то распределения $p_1(x)$, $p_3(x)$ устойчивы, а $p_2(x)$ неустойчиво. Назовем область $p < p_2(x)$ партиципаторным полем, а область $p > p_2(x)$ — полем сильной руки. Таким образом, функция $p_2(x)$ имеет смысл линии раздела полей. Назовем функции

$$h_1(x) = p_2(x) - p_1(x), \quad h_3(x) = p_3(x) - p_2(x)$$

шириной партиципаторного поля и шириной поля сильной руки соответственно (рис. 1). Выполнение неравенства $h_1(x) > h_3(x)$ для любого x означает, что идея партиципаторной демократии имеет более широкую общественную поддержку, чем идея демократии элит, и на-

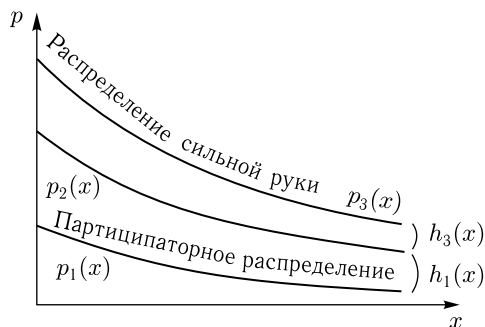


Рис. 1. Поле сильной руки и партиципаторное поле

оборот. Наиболее содержательными являются случаи, при которых функция $h_1(x) - h_3(x)$ меняет знак на интервале $(0, 1)$.

Ниже будут рассмотрены некоторые решения уравнения (3) и выяснен их содержательный смысл.

5.4.2. Формирование структур с внутренними переходными слоями

Если $\partial^2 p / \partial x^2 = O(\varepsilon^0)$, то решение уравнения (3) близко к решению вырожденного уравнения

$$\frac{\partial p}{\partial t} = -k_1(x)(p - p_1(x))(p - p_2(x))(p - p_3(x)), \quad (4)$$

в котором переменная x является параметром. Нетрудно убедиться (см. упражнение 1), что если при всех значениях x в момент $t = 0$ функция $p(x, 0)$ принимает значения в партиципаторном поле, т.е. $p(x, 0) < p_2(x)$, то при достаточно больших значениях t распределение власти будет близким к партиципаторному. Если же функция $p(x, 0) > p_2(x)$ для всех x , то при достаточно больших t распределение власти будет близким к распределению сильной руки.

Рассмотрим теперь случай, когда начальное распределение власти $p(x, 0)$ принимает значения в поле сильной руки при $x < x^*$, а в партиципаторном поле — при $x > x^*$, где $0 < x^* < 1$ (рис. 2, слева). Тогда в соответствии с уравнением (4) распределение $p(x, t)$ притянется к распределению сильной руки при $x < x^*$, а к партиципаторному распределению — при $x > x^*$ (рис. 2, справа). В окрестности точки x^* образуется переходный слой, в котором будет $\partial p / \partial x \gg 1$. Сформировавшись, этот переходный слой движется налево или направо; данному нестационарному процессу посвящен четвертый пункт.

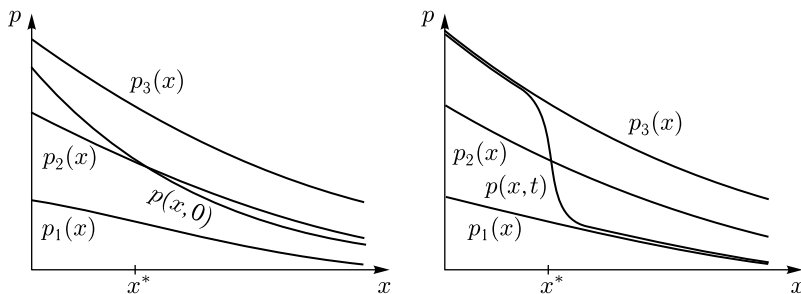


Рис. 2. Формирование внутреннего переходного слоя

Решения с внутренними переходными слоями называются контрастными структурами (КС). Следующие пункты посвящен анализу эволюции распределения власти методами теории КС.

5.4.3. Стационарные контрастные распределения власти

Метод пограничных функций позволяет построить асимптотику по малому параметру ε стационарного решения задачи (2), (3),

принадлежащего к классу КС. В нулевом приближении эта асимптотика имеет вид

$$p_{st}(x, \varepsilon) = \frac{1 - \exp[\sqrt{2\gamma(x_0)}(x - x_0)/\varepsilon]}{1 + \exp[\sqrt{2\gamma(x_0)}(x - x_0)/\varepsilon]} \frac{p_3(x) - p_1(x)}{2} + \frac{p_3(x) + p_1(x)}{2}. \quad (5)$$

Здесь x_0 — главный член асимптотики точки перехода, который находится из уравнения

$$h_1(x_0) = h_3(x_0), \quad (6)$$

или, что то же самое, $2p_2(x_0) = p_3(x_0) + p_1(x_0)$ (для простоты мы здесь полагаем, что уравнение (6) имеет единственное решение на интервале $0 < x < 1$). Таким образом, в точке перехода стационарного решения ширина партиципаторного поля равна ширине поля сильной руки: $h_1(x_0) = h_3(x_0)$.

Из формулы (5) следует, что распределение власти $p_{st}(x)$ близко к распределению сильной руки $p_3(x)$ при $0 \leq x < x_0$, и к партиципаторному распределению $p_1(x)$ при $x_0 < x \leq 1$ (рис. 3). При этом в окрестности точки x_0 имеем $\partial p / \partial x \cong \varepsilon^{-1}$. Назовем такие распределения власти контрастными.

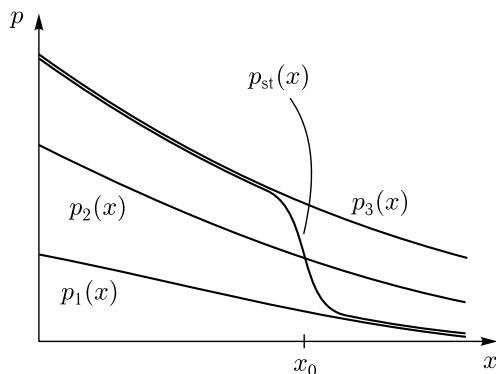


Рис. 3. Стационарное контрастное распределение власти

Исследование устойчивости (см. комментарии в конце раздела) показало, что стационарные контрастные распределения власти, близкие к распределению сильной руки на верхних уровнях власти ($p_{st}(x) \approx p_3(x)$ при $0 \leq x < x_0$) и к партиципаторному распределению на нижних уровнях власти ($p_{st}(x) \approx p_1(x)$ при $x_0 < x \leq 1$), устойчивы, если ширина поля сильной руки больше ширины партиципаторного поля на верхних уровнях власти ($h_3(x) > h_1(x)$ при $0 \leq x < x_0$) и меньше него на нижних ($h_3(x) < h_1(x)$ при $x_0 < x \leq 1$). И наоборот, если $h_3(x) < h_1(x)$ при $0 \leq x < x_0$ и $h_3(x) > h_1(x)$ при $x_0 < x \leq 1$, то решение неустойчиво.

5.4.4. Нестационарные контрастные распределения власти

Вернемся к нестационарному уравнению (3). Пусть в соответствии со сценарием, описанным во втором пункте, к некоторому моменту времени $t = t_0$ сформировалась КС с переходным слоем в окрестности точки $x = x^*$ (рис. 2). Тогда при $t > t_0$ решение будет принадлежать к классу нестационарных КС: $p(x, t) \approx p_3(x)$ при $x < R(t)$ и $p(x, t) \approx p_1(x)$ при $x > R(t)$, где точка перехода $R(t)$ является некоторой функцией времени. Соответствующее распределение власти будем называть нестационарным контрастным распределением власти (НКРВ).

Главный член асимптотики нестационарной КС имеет вид, аналогичный виду главного члена (5) асимптотики стационарной КС, а именно:

$$p(x, t, \varepsilon) = \frac{1 - \exp[\sqrt{2\gamma(R)}(x - R)/\varepsilon]}{1 + \exp[\sqrt{2\gamma(R)}(x - R)/\varepsilon]} \frac{p_3(x) - p_1(x)}{2} + \frac{p_3(x) + p_1(x)}{2}, \quad (7)$$

где функция $R = R(t)$ удовлетворяет уравнению

$$\frac{dR}{dt} = -\varepsilon [2p_2(R) - p_3(R) - p_1(R)] \sqrt{k_1(R)/2} \quad (8)$$

или, что то же самое,

$$\frac{dR}{dt} = \varepsilon [h_3(R) - h_1(R)] \sqrt{k_1(R)/2}. \quad (9)$$

Распределение власти $p(x, t)$ близко к распределению сильной руки $p_3(x)$ при $0 \leq x < R$ и к партиципаторному распределению $p_1(x)$ при $R < x \leq 1$. Величина dR/dt представляет собой скорость движения переходного слоя.

Рассмотрим теперь некоторые важнейшие случаи применения формулы (9).

Построение властной вертикали. Пусть поле сильной руки шире партиципаторного поля, $h_3(x) > h_1(x)$, для любого $x \in [0, 1]$, что означает склонность общества к варианту сильной руки. В этом случае стационарной контрастной структуры не существует, так как уравнение (6) не имеет решений. После формирования КС она движется (рис. 4, слева) в соответствии с формулой (9). При этом $dR/dt > 0$, и за время порядка ε^{-1} точка перехода $R(t)$ доходит до правого конца отрезка $[0, 1]$. Тем самым устанавливается распределение власти, близкое при $x \in [0, 1]$ к распределению сильной руки (рис. 4, справа). Этот процесс мы называем построением властной вертикали. Заметим, что в соответствии со сказанным в разд. 5.2 после формулы (4): если в начальный момент времени $t = 0$ функция $p(x, 0)$ принимает значения только в партиципаторном поле, то при достаточно больших значениях t распределение власти будет близким к партиципаторному даже при выполнении условия $h_3(x) > h_1(x)$. Для того чтобы при больших значениях t установилось распределение власти, близкое к распределению сильной руки, необходимо, чтобы в момент времени $t = 0$

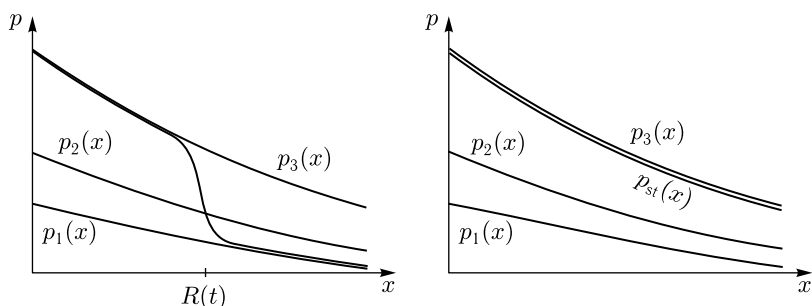


Рис. 4. Построение властной вертикали. На левом рисунке $dR/dt > 0$, так как $h_3(x) > h_1(x)$

функция $p(x, 0)$ была гладкой и хотя бы в одной точке интервала $(0, 1)$ принимала значение в поле сильной руки.

Установление партиципаторного распределения власти. Пусть $h_3(x) < h_1(x)$ для любого $x \in [0, 1]$, что означает склонность общества к партиципаторному распределению власти. В этом случае также СКРВ не существует, так как уравнение (6) не имеет решений. После формирования КС она движется (рис. 5, слева) в соответствии с формулой (9). При этом $dR/dt < 0$, и за время порядка ε^{-1} точка перехода $R(t, \varepsilon)$ доходит до левого конца отрезка $[0, 1]$. Тем самым устанавливается распределение власти, близкое при $x \in [0, 1]$ к партиципаторному распределению (рис. 5, справа). Заметим, что в соответствии со сказанным в разд. 5.2, если в начальный момент времени $t = 0$ функция $p(x, 0)$ принимает значения только в поле сильной руки, то при достаточно больших значениях t распределение власти будет близким к распределению сильной руки даже при выполнении условия $h_3(x) < h_1(x)$. Для того чтобы при больших значениях t установилось распределение власти, близкое к партиципаторному, необходимо, чтобы

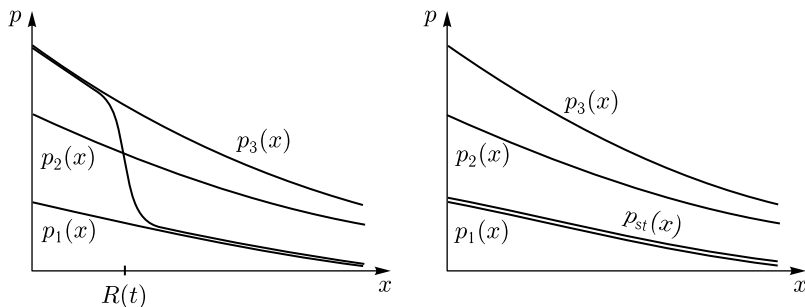


Рис. 5. Установление партиципаторного распределения. На левом рисунке $dR/dt < 0$, так как $h_1(x) > h_3(x)$

в момент времени $t = 0$ функция $p(x, 0)$ была гладкой и хотя бы в одной точке интервала $(0, 1)$ принимала значение в партиципаторном поле.

Установление СКРВ. Пусть существует точка $x_0 \in (0, 1)$ такая, что $h_1(x_0) = h_3(x_0)$, $h_1(x) < h_3(x)$ при $x < x_0$ и $h_1(x) > h_3(x)$ при $x > x_0$. Тогда в соответствии с третьим пунктом (см. формулу (6)) существует устойчивое СКРВ с переходом в окрестности точки x_0 . Пусть, кроме того, в момент $t = 0$ функция $p(x, 0)$ принимает значения в поле сильной руки при $x < x^*$ и в партиципаторном поле при $x > x^*$, где $x_0 < x^* < 1$. Тогда распределение $p(x, t)$ притянется к распределению сильной руки при $x < x^*$ и к партиципаторному распределению при $x > x^*$. После формирования КС с переходом в точке $x = x^*$ она движется в соответствии с уравнением (9). Тогда, если $x^* < x_0$, то будет $dR/dt > 0$. Если же $x^* > x_0$, то $dR/dt < 0$. Таким образом, при достаточно больших значениях t точка перехода $R(t)$ будет находиться в малой по ε окрестности точки x_0 . Тем самым устанавливается СКРВ с переходом в окрестности точки x_0 (рис. 6).

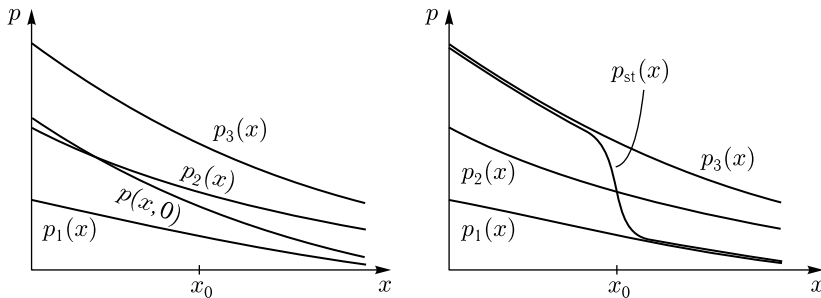


Рис. 6. Установление стационарного контрастного распределения власти. Слева показано начальное распределение, справа — стационарное решение $p_{st}(x)$. Точка x_0 определяется из условия (6)

5.4.5. Конфедеративные распределения власти

Всюду выше мы рассматривали СКРВ, близкие к распределению сильной руки при $x < x_0$ и близкие к партиципаторному распределению при $x > x_0$. Только такие СКРВ могут соответствовать последовательному убыванию количества полномочий с понижением уровня инстанции в иерархии. Однако на практике иногда реализуются распределения другого типа — такие, что некоторая (не высшая) инстанция обладает максимальным по всей иерархии количеством полномочий. Например, в конфедеративных государствах формальный глава конфедерации обладает количеством полномочий меньшим, чем главы субъектов конфедерации. Такие ситуации могут описываться конфедеративными СКРВ — решениями, близкими к партиципаторному распределению при $x < x_0$ и близкими к распределению сильной руки при $x > x_0$ (см. ниже: рис. 7).

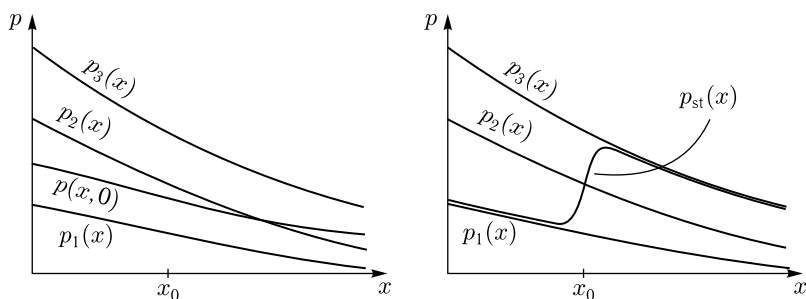


Рис. 7. Установление конфедеративного стационарного контрастного распределения власти. Слева показано начальное распределение, справа — стационарное решение $p_{st}(x)$. Точка x_0 определяется из условия (6)

Изложенные выше критерий устойчивости и формула (9) для скорости движения переходного слоя модифицируются в случае конфедеративных СКРВ. Формула (6) для главного члена асимптотики точки перехода x_0 остается справедливой.

Критерий устойчивости имеет следующий вид: конфедеративные СКРВ, близкие к партиципаторному распределению на верхних уровнях власти ($p \approx p_1(x)$ при $0 \leq x < x_0$) и к распределению сильной руки на нижних уровнях власти ($p \approx p_3$ при $x_0 < x \leq 1$), устойчивы, если ширина поля сильной руки меньше ширины партиципаторного поля на верхних уровнях власти ($h_3(x) < h_1(x)$ при $0 \leq x < x_0$) и больше на нижних ($h_3(x) > h_1(x)$ при $x_0 < x \leq 1$). Наоборот, если $h_3(x) > h_1(x)$ при $0 \leq x < x_0$ и $h_3(x) < h_1(x)$ при $x_0 < x \leq 1$, то конфедеративное СКРВ неустойчиво.

В формулах (8), (9) для скорости движения переходного слоя правая часть в случае конфедеративных СКРВ меняет знак. Так, вместо формулы (9) имеем

$$\frac{dR}{dt} = -\varepsilon [h_3(R) - h_1(R)] \sqrt{k_1(R)/2}. \quad (10)$$

Конфедеративные СКРВ формируются по следующему сценарию. Пусть существует точка $x_0 \in (0, 1)$ такая, что $h_1(x_0) = h_3(x_0)$, $h_1(x) > h_3(x)$ при $x < x_0$ и $h_1(x) < h_3(x)$ при $x > x_0$. Тогда существует устойчивое конфедеративное СКРВ с переходом в окрестности точки x_0 .

Пусть также для начального распределения власти $p^0(x)$ выполнены условия: $p^0(x)$ находится частично в партиципаторном поле, частично — в поле сильной руки так, что существует хотя бы одно $x_1 \in (0, x_0)$ такое, что $p^0(x_1) < p_2(x_1)$, и хотя бы одно $x_2 \in (x_0, 1)$ такое, что $p^0(x_2) > p_2(x_2)$, см. рис. 7, слева. После формирования КС она движется в соответствии с уравнением (10). Тогда, если $R(t) < x_0$, то $dR/dt > 0$, если же $R(t) > x_0$, то $dR/dt < 0$. Таким образом, при достаточно больших значениях t точка перехода $R(t)$ будет находиться

в малой по ε окрестности точки x_0 . Тем самым устанавливается конфедеративное СКРВ с переходом в окрестности точки x_0 (рис. 7, справа).

Кратко подведем итог данного раздела. Модель «власть–общество» с биполярной (амбивалентной) реакцией общества позволяет рассматривать целый ряд сценариев, имеющих содержательную политологическую трактовку. Мы рассмотрели четыре сценария: построение властной вертикали, установление партиципаторного распределения власти, контрастного и конфедеративного контрастного распределений.

Упражнения

1. Для уравнения (4) показать, что если при всех значениях x в момент $t = 0$ функция $p(x, 0)$ принимает значения в партиципаторном поле, т.е. $p(x, 0) < p_2(x)$, то при достаточно больших значениях t распределение власти будет близким к партиципаторному. Если же при $t = 0$ функция $p(x, 0)$ принимает значения в поле сильной руки, то при достаточно больших значениях t распределение власти будет близким к распределению сильной руки.

2. Как выглядит (на качественном уровне) эволюция распределения власти, если $h_1(x) < h_3(x)$ при всех $0 \leq x \leq 1$, $p_2(x) < p(x, 0) < p_3(x)$ при $0 \leq x < x_1$, $x_2 < x \leq 1$ и $p_1(x) < p(x, 0) < p_2(x)$ при $x_1 < x < x_2$, где x_1, x_2 — некоторые точки такие, что $0 < x_1 < x_2 < 1$.

Комментарии и библиографические примечания

1. Данный раздел написан на материале работы [1]. Устойчивость по Ляпунову контрастных структур задачи (2), (3) была исследована в работах [2, 3]. Основным математическим инструментом данных работ является метод пограничных функций (см., напр., [4]).

2. Математически строгим обоснованием сценариев установления СКРВ и конфедеративного стационарного распределения является теорема В. Ф. Бутузова и И. В. Неделько, доказанная в работе [5]; выполнение ее условий следует из сделанных нами предположений относительно начального распределения власти.

Литература

1. Дмитриев М. Г., Жукова Г. С., Петров А. П. Асимптотический анализ модели «власть–общество» для случая двух устойчивых распределений власти // Математическое моделирование. 2004. Т. 16, № 5. С. 23–34.
2. Васильева А. Б. Об устойчивости контрастных структур // Математическое моделирование. 1991. Т. 3, № 4. С. 114–123.
3. Васильева А. Б., Никитин А. Г., Петров А. П. Асимптотический метод исследования контрастных структур и его приложение к теории гидромагнитного динамо // Математическое моделирование. 1995. Т. 7. № 2. С. 61–71.

4. Васильева А.Б., Бутузов В.Ф. Асимптотические методы в теории сингулярных возмущений. М.: Высшая школа. 1990. 208 с.
5. Бутузов В.Ф., Неделько И.В. О глобальной области влияния устойчивых решений с внутренними слоями // Математическое моделирование. 2001. Т. 192, №5. С. 13–52.

5.5. Коррупция и другие несовершенства властных иерархий

В двух предыдущих разделах система «власть–общество» рассматривалась в предположении законопослушной и (в некотором смысле) совершенной иерархии: предполагалось, что на динамику распределения власти не оказывают влияние такие явления, как коррупция, клановость, некомпетентность. Здесь мы переходим к изучению такого известного отклонения от законопослушности, как коррупция. Анализ проводится на основе системно-социального подхода, предполагающего, что объектом моделирования является не отдельная инстанция, а иерархия в целом. Другими словами, предпосылки базовой модели «власть–общество» изменяются таким образом, чтобы учесть поведенческие характеристики коррумпированной инстанции. Тем самым на основе базовой модели системы построена математическая модель системы «власть–общество» с коррумпированной властной иерархией. Введены и математически формализованы такие понятия, как коррупция, степень коррумпированности, ущерб от коррупции, стоимость и эффективность ее подавления. Приведены результаты вычислительных экспериментов с моделью, показывающие, что в рамках фиксированного антикоррупционного ресурса существуют как более, так и менее эффективные стратегии подавления коррупции. Эффективность сильно зависит от строения властного «древа», степени коррумпированности в различных частях иерархии, поведенческих установок партнеров и других характеристик системы.

Далее первый и второй пункты посвящены, соответственно, построению модели «власть–общество» с коррумпированной иерархией и формализации понятий, связанных с ущербом от коррупции, а также стоимостью и эффективностью ее ограничения. В третьем пункте формализуется и рассматривается задача ограничения коррупции, приведен иллюстративный пример. В четвертом и пятом пунктах кратко рассматриваются иные (помимо коррупции) несовершенства иерархии: некомпетентность и клановость.

5.5.1. Учет коррупции в модели «власть–общество»

Рассмотренные в предыдущих разделах варианты модели «власть–общество» предполагают «идеальность» системы в том смысле, что властные инстанции считаются законопослушными и компетентными. Это означает, в частности, что властные распоряжения отдаются и реализуются в том объеме, который соответствует интересам общества. Между тем можно условно выделить по крайней мере

три класса явлений, приводящих к отклонению действий властных иерархий от надлежащего. К первому из них относится «искренняя» некомпетентность власти, порождаемая плохой информированностью, чрезмерным объемом работы, недостаточным профессиональным уровнем, отсутствием нужных ресурсов и т. д. Ко второму классу относится то, что мы называем клановостью: возникновение групп чиновников, практикующих «перетягивание власти» в пользу членов этих групп. Наконец, третий класс явлений — это сознательно совершаемые чиновниками и скрываемые ими неправильные действия. Не претендуя на юридическую точность, дадим следующее определение:

коррупция — тайное, корыстное использование государственной власти в интересах частных групп.

Выделим три следующих вида коррупционной активности (отсутствующей в законопослушных иерархиях).

«Обычная» коррупция — акты, которые осуществил бы законопослушный институт власти в соответствии со своей служебной инструкцией, но совершаемые за дополнительное вознаграждение (взятку). Типичный пример: поборы при различного рода регистрациях и оформлении законных документов.

Коррупция бездействия — акты, не совершенные институтом власти из-за взятки (законопослушный чиновник в данной ситуации употребил бы свои полномочия). К ним относится, например, неосновательный отказ от возбуждения уголовных дел.

Коррупция сверхдействия — акты, осуществляемые подкупленным институтом власти сверх надлежащих, т. е. такие, которые законопослушный чиновник не совершил бы (например, выдача водительских прав без сдачи экзамена).

Перейдем к математической формализации понятий, связанных с коррупцией. Для этого сначала сравним действия законопослушной и коррумпированной инстанций, находящихся на i -м уровне иерархии (для простоты будем полагать, что этот уровень не есть самый нижний, т. е. $0 \leq i < N$) — см. рис. 1).



Рис. 1. Сравнение законопослушной и коррумпированной инстанций

Пусть в единицу времени идеальная инстанция отдает инстанции из слоя с номером $i + 1$ (т. е. своему подчиненному чиновнику) $M^0 > 0$ равноценных по своей «стоимости» властных распоряжений, например шесть приказов об оформлении разрешений для шести различных фирм на проведение каждой из них одинаковой экспортной операции (сплошные стрелки на рис. 1, $M^0 = 6$).

Коррупцированная инстанция (рис. 1) может одновременно практиковать, вообще говоря, все три перечисленных выше вида коррупции:

- $M > 0$ приказов отдается, как и в законопослушном случае, т. е. без взятки (сплошные стрелки);
- $M_0 \geq 0$ отвечает обычной коррупции — правильные действия, но за взятку b_0 (стрелки с кружочком);
- приказы числом $M_- \geq 0$ не отдаются в результате коррупции бездействия (стрелки с крестиком);
- отдаются $M_+ \geq 0$ лишних приказов, которых не было бы в действиях идеальной инстанции (коррупция сверхдействия, пунктирные стрелки).

В примере на рис. 1 две фирмы получают нужное им разрешение без дополнительных проблем ($M = 2$), двум фирмам приходится потратиться на взятку ($M_0 = 2$), две фирмы не получают причитающееся им разрешение ($M_- = 2$) и, наконец, появляются две фирмы, не получившие бы разрешение у законопослушного чиновника ($M_+ = 2$).

Введем понятие степени коррупцированности инстанции с номером i в данный момент времени t . Отнормируем отвечающие коррупционным актам числа $M_0(i, t)$, $M_-(i, t)$, $M_+(i, t)$ на число правильных действий $M(i, t)$ (считаем, что оно не равно нулю) и обозначим через

$$c_0(i, t) = \frac{M_0}{M}, \quad c_-(i, t) = \frac{M_-}{M}, \quad c_+(i, t) = \frac{M_+}{M}$$

парциальные степени коррупцированности, соответствующие вышеуказанным типам коррупции. Из них складывается общая коррупцированность i -го института:

$$c(i, t) = c_0 + c_- + c_+.$$

Так, для примера, изображенного на рис. 1, имеем с очевидностью, что $c_0 = c_- = c_+ = 1$, $c = 3$.

Степень коррупцированности иерархии в целом вводится как средневзвешенная величина с учетом уровня текущей власти $p(x, t)$ в разных инстанциях иерархии:

$$C(t) = \frac{\int_0^l c(x, t)p(x, t) dx}{\int_0^l p(x, t) dx}. \quad (1)$$

В частности, если $c(x, t) = c(t)$, т. е. коррумпированность одинакова во всех звеньях властной структуры (это возможно, например, при $c_0(x, t) = c_0(t)$, $c_-(x, t) = c_-(t)$, $c_+(x, t) = c_+(t)$), то $C(t) = c(t)$ и средневзвешенная характеристика иерархии совпадает с локальной.

Уравнение модели «власть–общество», учитывающее коррумпированность иерархии, имеет вид

$$\frac{\partial p}{\partial t} = \frac{\partial}{\partial x} \left(k_f \frac{\partial p}{\partial x} \right) + F(x, t, p), \quad 0 < x < 1, \quad t > t_0, \quad (2)$$

где

$$k_f = \left[1 + \frac{c_+(x, t) - c_-(x, t)}{1 + c_0(x, t) + c_-(x, t)} \right] k(x, t, p). \quad (3)$$

Вместе с такими же, как и в базовой модели, граничными условиями

$$\frac{\partial p}{\partial x} \Big|_{x=0} = \frac{\partial p}{\partial x} \Big|_{x=1} = 0, \quad (4)$$

имеющими смысл отсутствия потоков власти через границы иерархии, а также начальным условием

$$p(x, 0) = p^0(x) \quad (5)$$

и заданными поведенческими характеристиками k , F , c_0 , c_- , c_+ , уравнения (2), (3) образуют замкнутую математическую модель «власть–общество» с коррумпированной иерархией, из которой для всех $0 \leq x \leq 1$, $t \geq t_0$ однозначно находится решение — функция $p(x, t)$.

Нетрудно видеть, что при нулевой коррумпированности ($c_0 = c_- = c_+ = 0$) модель (2)–(5) принимает вид базовой модели, рассмотренной в разд. 5.3 настоящей главы. Отметим, что это происходит также в случае $c_-(x, t) \equiv c_+(x, t)$, $c_0(x, t) \neq 0$, т. е. обычная коррупция не меняет распределение власти (и потоки власти) в иерархии, а коррупция бездействия и коррупция сверхдействия «уравновешивают» в данном смысле друг друга.

Перечислим допущения, сделанные при построении модели (2)–(5), охарактеризовав тем самым границы ее применимости (эти допущения использовались при выводе модели, см. комментарий 1 в конце раздела).

1. Рассматривается цепочечная властная структура с одним чиновником на каждом иерархическом уровне.

2. Команды «через голову» в иерархии отсутствуют (о механизме «команд через голову» см. в разд. 5.3).

3. Степени коррумпированности $c_0(x, t)$, $c_-(x, t)$, $c_+(x, t)$ — непрерывные функции иерархической переменной x .

4. Степень ответственности инстанций не зависит от уровня их коррумпированности (функция k не зависит явно от c_0 , c_- , c_+).

5. Реакция системы F имеет тот же функциональный вид, что для законопослушной иерархии (т. е. F не зависит явно от степеней коррумпированности).

5.5.2. Ущерб от коррупции, стоимость и эффективность ее ограничения

Положим по определению, что ущерб от действий коррумпированной инстанции x равен

$$d(x, t) = -\frac{c_0 b_0 + c_- + c_+}{1 + c_0 + c_-} k \frac{\partial p}{\partial x}, \quad 0 < x < 1, \quad t > t_0. \quad (6)$$

Это определение опирается на следующие содержательные аргументы. Наносимый обществу коррупцией в единицу времени локальный абсолютный ущерб $d(x, t)$ складывается из трех парциальных ущербов, отвечающих коррупционным актам соответствующих типов. Для обычной коррупции ущерб пропорционален средней величине взятки b_0 , умноженной на количество распоряжений «за взятку», поскольку потратившая на взятку «фирма» так или иначе «взыщет» ее стоимость с общества (например, повысив цену на предлагаемые ей товары и услуги). В отношении двух других типов коррупции ущерб равен либо не потраченному на благо общества властному ресурсу (бездействие), либо ресурсу, потраченному в пользу частных интересов (сверхдействие), а величина самой взятки считается пренебрежимо малой по сравнению с ущербом от коррупционных действий. Во всех трех компонентах ущерб пропорционален властному ресурсу данной инстанции, т. е. проходящему через нее потоку власти, т. е. пропорционален $k \partial p / \partial x$.

Локальная абсолютная польза $u(x, t)$ от действий инстанции складывается из пользы от действий, произведенных без взятки, и действий, отвечающих обычной коррупции. Применяя такие же рассуждения, как при формализации ущерба, приходим к выражению

$$u(x, t) = -\frac{1 + c_0}{1 + c_0 + c_-} k \frac{\partial p}{\partial x}, \quad 0 < x < 1, \quad t > t_0. \quad (7)$$

Интегрируя (6) и (7) по всей протяженности иерархии от $x = 0$ до $x = 1$ и деля первое выражение на второе, получим *общий относительный ущерб*, наносимый всеми звеньями коррумпированной иерархии:

$$D(t) = \left[\int_0^1 \frac{c_0 b_0 + c_- + c_+}{1 + c_0 + c_-} k \frac{\partial p}{\partial x} dx \right] / \left[\int_0^1 \frac{1 + c_0}{1 + c_0 + c_-} k \frac{\partial p}{\partial x} dx \right]. \quad (8)$$

В случае однородной коррумпированности инстанций, т. е. когда c_0 , c_+ , c_- , не зависят от x , формула (8) приобретает вид

$$D(t) = \frac{c_0(t) b_0 + c_-(t) + c_+(t)}{1 + c_0(t)}. \quad (9)$$

Если, например $c_- = c_+ = 0$, $c_0 = 1$, т. е. коррупция бездействия и коррупция сверхдействия отсутствуют и половина действий иерархии осуществляется в режиме обычной коррупции, то $D = b_0/2$, т. е. общий относительный ущерб равен половине средней взятки b_0 .

В отношении мер ограничения коррупции оговоримся, что здесь мы рассматриваем только акции, непосредственно направленные на устранению из структур власти коррупции чиновников: сбор оперативной информации, уголовное и судебное преследования и т. д. Разумеется, проведение этих мероприятий требует соответствующих затрат. Их «стоимость» (в единицу времени) будем определять, исходя из следующих соображений. Прежде всего, очевидно, что она напрямую зависит от текущего уровня власти $p(x, t)$ коррупционера и тем больше, чем больше этот уровень. Добиться осуждения нечестного чиновника из федерального эшелона власти гораздо труднее, чем в случае с фигурой местного масштаба. Примем линейную, пропорциональную зависимость стоимости от величины $p(x, t)$. Кроме того, эти затраты, естественно, тем выше, чем существеннее текущий ущерб $d(x, t)$ от действий коррупционера, уже хотя бы потому, что у последнего с ростом наносимого им ущерба увеличиваются возможности для своей «защиты» (примем поэтому затраты пропорциональными также величине $d(x, t)$). Тогда *локальная абсолютная стоимость* подавления коррупции дается выражением

$$v(t) = \varepsilon_0 p(x, t) d(x, t) = -\varepsilon_0 \frac{c_0 b_0 + c_- + c_+}{1 + c_0 + c_-} p k \frac{\partial p}{\partial x}, \quad 0 < x < l, \quad t > t_0, \quad (10)$$

где ε_0 — коэффициент, который будет определен ниже.

Проинтегрировав (10) по x от $x = 0$ до $x = 1$, получим общую абсолютную цену затрат на полное искоренение коррупции в иерархии; деля ее на общую абсолютную пользу (интеграл от (7) по x от $x = 0$ до $x = 1$), придем к выражению

$$V(t) = \left[\varepsilon_0 \int_0^l \frac{c_0 b_0 + c_- + c_+}{1 + c_0 + c_-} p k \frac{\partial p}{\partial x} dx \right] / \left[\int_0^l \frac{1 + c_0}{1 + c_0 + c_-} k \frac{\partial p}{\partial x} dx \right] \quad (11)$$

для *общей относительной стоимости* подавления коррупции в иерархии. Нормирующий множитель $\varepsilon_0(t) > 0$ может трактоваться как характеристика «единичной» затраты. Поскольку речь идет об относительных величинах, то он имеет вспомогательное значение и его можно выбирать разными способами, например, зная решение $p(x, t)$ по формуле

$$\varepsilon_0 = \left[\int_0^l k \frac{\partial p}{\partial x} dx \right] / \left[\int_0^l k p \frac{\partial p}{\partial x} dx \right]. \quad (12)$$

Нормировка (12) такова, что для однородной ($c_0 = c_0(t)$, $c_- = c_-(t)$, $c_+ = c_+(t)$) коррумпированности имеем $D(t) = V(t)$, т. е. ущерб и затраты равны.

Одно из возможных направлений применения модели «власть–общество» с коррумпированной иерархией — это анализ антикоррупционных стратегий. Проблема ставится следующим образом. Пусть в единицу времени выделяются некоторые средства (часть госбюджета)

на «полицейские меры» по уменьшению коррумпированности иерархии. Как распорядиться ими наиболее эффективно, т. е. достичь в рамках фиксированного антикоррупционного ресурса максимального эффекта по критерию «затраты–прибыль»? Под «прибылью» здесь понимается «предотвращенный ущерб», т. е. ущерб, который был бы нанесен обществу, если бы мероприятия против коррупции не проводились. Введем понятие *эффективности подавления коррупции* как отношения предотвращенного ущерба к стоимости соответствующих мер:

$$E(t) = \frac{\Delta D}{\Delta V} = \frac{D(C + \Delta C) - D(C)}{V(C + \Delta C) - V(C)}, \quad (13)$$

где ΔC — уменьшение степени коррумпированности, ΔV — стоимость этого изменения, ΔD — уменьшение наносимого коррупцией ущерба, обусловленное проведенными мероприятиями.

Эффективность $E(t)$ может сложно (и сильно) зависеть от применяемых стратегий подавления коррупции. Поэтому первостепенное значение имеет сравнительный анализ различных стратегий подавления коррупции. Понятие стратегии может пониматься различным образом. Мы конкретизируем его, понимая под стратегией направленность антикоррупционных мер преимущественно против определенных уровней иерархии. Именно, будем говорить о стратегии преимущественного

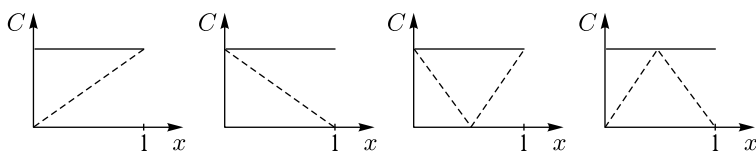


Рис. 2. Стратегии подавления коррупции: подавление старших инстанций (а), младших (б), средних (в), крайних (г). Коррумпированность «до применения мер» обозначена сплошной линией, «после применения» — пунктиром

подавления коррупции среди старших инстанций, если коррумпированность среди этих инстанций уменьшилась в результате применения антикоррупционных мер сильнее, чем в средней и нижней частях иерархии (рис. 2). Аналогично вводятся стратегии преимущественного подавления коррупции среди младших, средних и крайних инстанций.

5.5.3. Сравнительный анализ стратегий ограничения коррупции

Перейдем к сравнительному анализу эффективности антикоррупционных стратегий на основе вычислительных экспериментов со стационарной моделью (2)–(5). Для всех расчетов, проводившихся на отрезке $0 \leq x \leq 1$ с нулевыми начальными данными от момента $t = 0$ до момента установления стационарного решения $p(x)$, в модели бралось $k(x, t, p) \equiv 0,005$, реакция общества выбиралась в виде $F = p^0(x) - p(x, t)$.

Промежуточная эффективность вычисляется по значениям ущерба и затрат для двух «соседних» (см. таблицы ниже) состояний коррумпированности. В качестве «итогового» значения эффективности берется значение, получающееся при условии полного подавления коррупции для чиновников, занимающих соответствующее место в иерархии (исходному профилю коррупции отвечает ее начальное состояние). При этом основное внимание уделяется сравнительному анализу эффективности в зависимости от вида коррупции и от *степени централизации* властных полномочий в иерархии. Степень централизации (децентрализации) задается видом функции $p^0(x)$, ее характеристикой служит значение $p^0(x=0,5)$ — чем меньше эта величина по сравнению с $p^0(0) = 1$, тем более централизована иерархия (рис. 3).

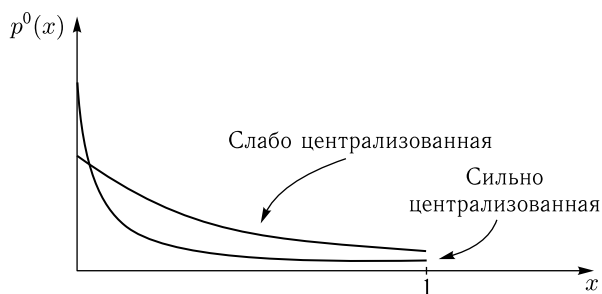


Рис. 3. Распределение власти в слабо централизованной и в сильно централизованной иерархиях

Для определенности возьмем функцию $p^0(x)$, имеющую вид

$$p^0(x) = \frac{1}{[1 + (100^{1/n} - 1)x]^n}.$$

При различных значениях параметра n получаем различные степени централизации: чем выше этот параметр, тем более пологим является график функции $p^0(x)$, тем ниже степень централизации. Для численных экспериментов приняты следующие значения:

- сильная централизация: $n = 1$, при этом $p^0(0,5) = 0,02$,
- средняя централизация: $n = 2$, при этом $p^0(0,5) = 0,033$,
- слабая централизация: $n = 8$, при этом $p^0(0,5) = 0,072$.

Рассмотрим коррупцию сверхдействия (исходное значение коррумпированности $c_+(x) \equiv 2$). Результаты численного расчета для общего относительного ущерба, общей относительной стоимости и эффективности подавления коррупции при стратегиях подавления «младших» и «старших» чиновников представлены в табл. 1, 2 (в левом столбце даны последовательные профили степени коррумпированности). «Итоговое» значение эффективности для каждой из стратегий — это число в нижней строке и правом столбце соответствующей таблицы.

Т а б л и ц а 1. Результаты численного эксперимента (коррупция сверхдействия, стратегия: подавление «младших» чиновников)

$c_+(x)$	D	V	E
2	2	2	—
$2 - 0,1x$	1,974	1,987	2,009
$2 - 0,2x$	1,948	1,974	2,011
$2 - 0,3x$	1,922	1,961	2,013
$2 - 0,4x$	1,896	1,948	2,014
$2 - 0,5x$	1,87	1,935	2,017
$2 - 2x$	1,487	1,747	2,024

Т а б л и ц а 2. Результаты численного эксперимента (коррупция сверхдействия, стратегия: подавление «старших» чиновников)

$c_+(x)$	D	V	E
2	2	2	—
$1,9 + 0,1x$	1,926	1,913	0,848
$1,8 + 0,2x$	1,852	1,826	0,848
$1,7 + 0,3x$	1,778	1,738	0,848
$1,6 + 0,4x$	1,704	1,65	0,847
$1,5 + 0,5x$	1,629	1,562	0,847
$2x$	1,484	1,201	0,843

Аналогичные таблицы можно привести также для стратегий подавления «средних» и «крайних» чиновников.

Аналогичные таблицы можно привести также для случаев коррупции бездействия и «обычной коррупции». Наиболее важная величина — это «итоговое» значение эффективности, результаты для которой по различным типам коррупции и различным стратегиям ее подавления сведены в табл. 3. Подобным образом, табл. 4, 5 содержат значения эффективности, рассчитанные для средне- и слабоцентрализованной иерархий.

Т а б л и ц а 3. Эффективность четырех стратегий при подавлении коррупции различных типов. Случай сильно централизованной иерархии

Тип коррупции	Стратегия: подавление			
	Младших	Старших	Средних	Крайних
c_0	2,89	0,8	2,21	0,75
c_+	2,02	0,84	1,62	0,8
c_-	1,33	0,91	1,28	0,89

Сопоставление значений эффективности подавления коррупции для иерархий с различной степенью централизации (от сильной к слабой) позволяет сделать следующие выводы.

1. Эффективность подавления младших чиновников для всех типов коррупции заметно падает.

2. При подавлении старших чиновников эффективность для всех типов коррупции незначительно возрастает.

Таблица 4. Эффективность четырех стратегий при подавлении коррупции различных типов. Случай среднецентрализованной иерархии

Тип коррупции	Стратегия: подавление			
	Младших	Старших	Средних	Крайних
c_0	2,45	0,84	2,01	0,78
c_+	1,85	0,87	1,55	0,81
c_-	1,29	0,93	1,25	0,9

Таблица 5. Эффективность четырех стратегий при подавлении коррупции различных типов. Случай слабоцентрализованной иерархии

Тип коррупции	Стратегия: подавление			
	Младших	Старших	Средних	Крайних
c_0	2	0,85	1,73	0,79
c_+	1,64	0,88	1,42	0,82
c_-	1,23	0,94	1,2	0,91

3. Эффективность подавления средних чиновников уменьшается для всех типов коррупции.

4. При подавлении крайних чиновников эффективность ненамного возрастает для всех типов коррупции.

5. Во всех случаях наибольшее значение эффективности достигается при давлении на младших чиновников, наименьшее — при давлении на крайних чиновников.

6. Отношение наибольшей к наименьшей эффективности подавления коррупции максимально в случае обычной коррупции (эффективность ее подавления сильно зависит от стратегии).

7. Отношение наибольшей к наименьшей эффективности подавления коррупции минимально в случае коррупции бездействия (эффективность ее подавления слабо зависит от стратегии).

5.5.4. Учет некомпетентности иерархии

Некомпетентность определяется здесь как непреднамеренное и бескорыстное принятие и реализация властными структурами ошибочных решений. При этом (в отличие, скажем, от коррупции) власть действует честно и открыто. Некомпетентность может быть вызвана разнообразными причинами: недостаточный уровень профессионального образования и опыта работы, плохая информированность, чрезмерный объем работы и отсутствие ресурсов, необходимых для выполнения своих обязанностей, неоптимальная структура иерархии,

личные качества чиновников (ангажированность, халатность, безразличие) и т. д.

Логика построения и исследования модели с некомпетентной иерархией близка к логике для модели коррумпированных иерархий. Первый шаг — классификация видов некомпетентности. Их существует, по меньшей мере, два.

1. Некомпетентность бездействия. Отсутствие в деятельности инстанций актов, которые компетентный институт власти обязательно осуществил бы в соответствии со своими служебными обязанностями (качественная профилактика техногенных катастроф, своевременное парирование эпидемиологических угроз и т. п.).

2. Некомпетентность сверхдействия — наличие в деятельности инстанций актов, которые компетентный институт власти никогда бы не осуществил (скоропалительные и непродуманные социальные, политические и иные реформы, чрезмерная регламентация жизни гражданского общества и т. п.).

Уравнение модели в случае линейной реакции общества имеет вид

$$\frac{\partial p}{\partial t} = \frac{\partial}{\partial x} \left(\kappa_f \frac{\partial p}{\partial x} \right) + F(x, t, p), \quad 0 < x < 1, \quad t > t_0, \quad (14)$$

где $\kappa_f = (1 + b_+ - b_-) \kappa(x)$. Здесь величины b_+ , b_- аналогичны рассмотренным ранее характеристикам c_+ , c_- . Таким образом, формально уравнение (14) для некомпетентной иерархии совпадает с уравнением (2) для коррумпированной иерархии, но коэффициент κ_f определяется иначе.

В соответствии с указанной логикой вводятся общий относительный ущерб от некомпетентности и общая относительная стоимость ее подавления, а также эффективность подавления. При этом стоимость мероприятий по ограничению некомпетентности вводится исходя из следующих предположений:

1) стоимость ограничения тем выше, чем выше положение инстанции в иерархии, а соответственно, ее текущая власть;

2) стоимость ограничения тем меньше, чем больше (и тем самым очевиднее) текущий ущерб, приносимый неправильными действиями инстанции.

Первое из этих предположений полностью аналогично соответствующему предположению для коррумпированной иерархии, тогда как второе по смыслу совершенно противоположно своему аналогу. Это приводит к существенно другим стратегиям ограничения дефектов несовершенств властных структур. В частности, наиболее предпочтительной стратегией является ограничение некомпетентности не младших, как в случае с коррупцией, а старших инстанций.

5.5.5. Клановые отношения в иерархии

В данном разделе мы кратко рассмотрим еще один тип несовершенств властных иерархий, «кланов», т. е. частей («поддеревьев») иерархии, имеющих, как правило, вертикальное строение и состоящих

из инстанций, преследующих общие цели, не совпадающие, вообще говоря, с целями общества. Для описания феномена «клановости» используем дискретный вариант модели древовидной иерархии, поскольку в нем предусмотрена возможность анализа топологии властного дерева и задания характеристик для каждой конкретной инстанции.

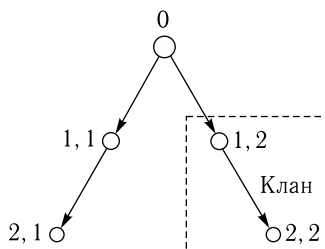


Рис. 4. Пятиточечная иерархия с кланом

Рассмотрим простейшее пятиточечное дерево, представляющее собой две трехточечные ветви с общей корневой вершиной (рис. 4). Значение власти в высшей вершине обозначим через p_0 , в вершинах первого (среднего) слоя — p_{11} , p_{12} , в вершинах второго (нижнего) слоя — через p_{21} , p_{22} . Будем считать, что правое поддерево (вершины p_{12} , p_{22}) составляет клан. Стратегия клана (один из возможных сценариев) состоит в том, что он пытается доступными ему способами максимизировать отношение своей суммарной власти к суммарной власти левого поддерева (т. е. отношение $\beta = (p_{12} + p_{22}) / (p_{11} + p_{21})$) или к власти иерарха (т. е. отношение $\gamma = (p_{12} + p_{22}) / p_0$). Высший иерарх пытается (также один из возможных сценариев) доступными ему способами парировать эту тенденцию. Левое поддерево (вершины p_{11} и p_{21}) считается пассивным (не является кланом).

Эффективный способ достижения своей цели кланом заключается в увеличении его членами степени своего властолюбия. Суммарная реакция системы берется в виде

$$F_{ij}(t) = k_1(p_{ij}(t) - p_{ij}^0) + F_{H,ij},$$

где p_{ij} — текущие уровни власти, p_{ij}^0 — «идеальные» уровни власти: $p_0^0 = H$, $p_{11}^0 = p_{12}^0 = H - k/2$, $p_{21}^0 = p_{22}^0 = H - k$. Здесь k_1 — интенсивность реакции гражданского общества, пропорциональной отклонению уровня власти от «идеального»; параметр k характеризует «перепад» полномочий между инстанциями — линейная функция от уровня иерархии). Под F_H понимается степень властолюбия рассматриваемой инстанции. При этом $F_H > 0$ для клана и $F_H = 0$ для не-клана и для высшего иерарха.

Система уравнений, определяющая стационарное распределение власти в иерархии, выглядит следующим образом:

$$\begin{aligned} -k_0(p_0 - p_{11}) - k_0(p_0 - p_{12}) - k_1(p_0 - p_0^0) &= 0, \\ k_0(p_0 - p_{11}) - k_0(p_{11} - p_{21}) - k_1(p_{11} - p_{11}^0) &= 0, \\ k_0(p_0 - p_{12}) - k_0(p_{12} - p_{22}) - k_1(p_{12} - p_{12}^0) + F_H &= 0, \\ k_0(p_{11} - p_{21}) - k_1(p_{21} - p_{21}^0) &= 0, \\ k_0(p_{12} - p_{22}) - k_1(p_{22} - p_{22}^0) + F_H &= 0, \end{aligned} \quad (15)$$

где k_0 — характеристика ответственности инстанций. Данная система получается из подсчета баланса власти в инстанциях.

Численное исследование системы приводит к следующим выводам.

1. При увеличении властолюбия клана уровень власти во всем дереве (а не только внутри клана) увеличивается (этот результат согласуется с принципом максимума для уравнений параболического типа).

2. Несмотря на то, что увеличение власти идет во всем дереве, отношение власти клана к власти левого поддерева и к власти иерарха увеличивается.

Заметим, что величина F_H не может быть сколь угодно большой, так как при достаточно большом властолюбии клана стационарного распределения власти в иерархии не существует. В данном примере «предельное» значение $F_H \simeq 0,1-0,2$. Перебор различных комбинаций параметров показал в случае базовой модели, что максимальная величина $\beta = (p_{12} + p_{22})/(p_{11} + p_{21}) \approx 1,5$.

Иерарх может парировать «притязания» клана, используя свое собственное властолюбие, чтобы вернуть величину β к первоначальному значению. Оказывается, что таким образом можно полностью парировать лишь увеличение отношения власти клана к своей собственной власти, т.е. сохранить величину $\gamma = (p_{12} + p_{22})/p_0$. Для этого необходимо увеличить властолюбие иерарха на величину $F_H^0 \geq 2 \cdot F_H$. Парировать же увеличение отношения власти клана к власти левого поддерева (т.е. уменьшить величину β) таким способом полностью нельзя. Можно лишь несколько уменьшить это отношение, но ценой очень больших усилий (увеличив свое властолюбие намного больше, чем клан в целом).

Резюмируя данный раздел, отметим, что рассмотренные в нем модели предоставляют возможность (при соответствующем «оснащении» конкретными данными) для осмысленного и научно обоснованного планирования и анализа мероприятий по уменьшению коррупции, некомпетентности и клановости в конкретных системах «власть–общество».

Упражнения

1. Пояснить, каким образом выводы 6, 7 (п. 4, после табл. 5) следуют из табл. 3–5.

2. Пусть в системе (15) параметр k_1 варьируется от уравнения к уравнению. Как это скажется на величине β ? Дать ответу содержательную трактовку в следующих терминах: «Чтобы повысить свою власть, клан может поступить следующим образом...».

Комментарии и библиографические примечания

1. Данный раздел написан на материале работ [1, 2]; см. также главу V в монографии [3]. В частности, в этих источниках содержится полный вывод уравнения (2) модели с коррумпированной иерархией.

2. Введенное определение степени коррумпированности не является единственно возможным. Так, величины M_0 , M_- , M_+ можно нормировать на сумму $M_0 + M_- + M_+$, т.е. на общее число действий чиновника или каким-то иным образом. При построении моделей коррумпированных властных структур и интерпретации вытекающих из них результатов важно лишь придерживаться единообразного толкования понятия степени коррумпированности.

3. Из различных видов несовершенств властных иерархий, рассматриваемых в данном разделе (коррупция, некомпетентность, клановость), наиболее широко изучается коррупция. При этом большая часть работ базируется на подходе Ното *Economicus*, и предметом моделирования в них является не иерархия в целом, а отдельная инстанция (и ее взаимодействие со своим окружением). Первыми работами, в которых привлекались математические модели для изучения коррупции, были [4, 5]. Позднее феномен коррупции исследовался с помощью теоретико-игровых методов [6, 7]. Рассматривалось возникновение коррупции в системе принципал–агент–супервайзер, коррупция в иерархических структурах различных простейших топологий, исследовался вопрос об эффективности предоставляемых услуг при наличии возможности коррупции и так далее. Также существует множество эмпирических работ о коррупции на уровне стран (например, [8]) и на уровне отдельных организаций [9]. Если говорить об отечественных публикациях, то можно выделить работу [10], посвященную вопросу об устройстве контролирующих коррупцию организаций. а также статью [11], в которой предложена имитационная модель коррупционного поведения в иерархических системах, учитывающая индивидуальные стратегии отдельных элементов и позволяющая описывать коллективное поведение достаточно больших групп.

Литература

1. *Mikhailov A. P.* Efficient Strategies of Corruption Suppression in State Power Hierarchies // Proceedings of 15th IMACS World Congress, Berlin, Germany. 1997. V. III. P. 727–733.
2. *Михайлов А. П.* Модель коррумпированных властных иерархий // Математическое моделирование. 1999. Т. 11, № 1. С. 3–19.

3. Михайлов А.П. Моделирование системы «власть–общество». М.: ФИЗМАТЛИТ, 2006. 144 с.
4. Rose-Ackerman S. The political economy of corruption // *Corruption and the global economy*. 1997. 31, 60.
5. Ackerman S.R. *Corruption: A study in political economy*. New York: Academic Press, 1978.
6. Tirole J. Hierarchies and bureaucracies: On the role of collusion in organizations // *JL Econ. & Org.*, 1986. 2. P. 181.
7. Bac M. Corruption, supervision, and the structure of hierarchies // *The Journal of Law, Economics, and Organization*. 1996. V. 12, №2. P. 277–298.
8. Mauro P. Corruption and growth // *The quarterly journal of economics*. 1995. V. 110, №3. P. 681–712.
9. Bertrand M., Djankov S., Hanna R., Mullainathan S. Does corruption produce unsafe drivers? (№w12274) // *National Bureau of Economic Research*. 2006.
10. Васин А. А., Картунова П. А., Уразов П. А. Модели организации государственных инспекций и борьбы с коррупцией // *Математическое моделирование*, 2010. Т. 22, № 4. С. 67–89.
11. Зенюк Д.А., Малинецкий Г.Г., Фаллер Д.С. Имитационная модель коррупции в иерархических системах // *Компьютерные исследования и моделирование*, 2014. V. 6, № 2. P. 321–329.

5.6. Динамика гонки вооружений

Рассматриваются примеры наиболее простых математических моделей, описывающих процесс гонки вооружений между двумя или несколькими странами или блоками стран. Под гонкой вооружений здесь понимается изменение (увеличение или уменьшение) со временем военных потенциалов стран или блоков, стремящихся обеспечить свою безопасность в случае возможного возникновения вооруженного конфликта между ними. Цели соперников могут быть весьма разнообразными — достичь паритета (равенства потенциалов), получить абсолютное превосходство над противником (фактически лишаящее его суверенитета), обеспечить себе пусть относительно небольшой, но грозный оппоненту неприемлемым для него ущербом уровень вооружений, а также одновременное стремление сторон к взаимному гармоничному, т. е. не грозящему их безопасности сокращению вооружений и т. д.

Поведенческие гипотезы описываемых моделей формулируются, естественно, не для отдельных индивидов, а для целых стран, каждая из которых понимается как единый актер, не имеющий внутренней структуры и внутренних противоречий. Тем самым странам приписываются определенные элементы человеческой психологии — такие, например, как неприязнь и недоверие. В этом смысле моделирование гонки вооружений между странами в какой-то степени подобно моделирова-

нию взаимоотношений между индивидами: как обычно, поведенческие гипотезы позволяют построить концептуально обоснованную модель, которая затем изучается математическими средствами.

В первом пункте данного раздела, выводятся уравнения простейшей (базовой) модели. Ее анализу посвящен второй пункт в частности, рассматривается вопрос о том, при каких условиях гонка развивается с неограниченным ростом вооружений, а при каких она переходит в стационарное состояние, при котором количество вооружений каждой из сторон становится постоянным во времени. Далее изучаются два обобщения модели — ее нелинейный вариант, учитывающий ограниченность экономических возможностей стран, а также обобщение на случай произвольного количества стран, пятый пункт посвящен заключительным выводам и комментариям.

5.6.1. Построение базовой модели

Рассмотрим гонку вооружений между двумя странами. Общее количество вооружений, имеющихся у i -й страны в момент времени t , обозначим через $m_i(t)$ (всюду далее будем предполагать без дополнительного упоминания, что индекс i может принимать значения 1 и 2). Для простоты будем полагать, что объемы вооружений измеряются в денежных единицах.

Положим, что скорость изменения количества вооружений dm_i/dt определяется следующими тремя факторами.

1. Количество вооружений у противника. Поведенческая гипотеза состоит в том, что чем лучше вооружен противник, тем быстрее страна наращивает собственные вооружения. Простейшая (линейная) конкретизация данной гипотезы предполагает, что выражение для скорости dm_1/dt содержит слагаемое, пропорциональное величине m_2 (аналогично — для скорости dm_2/dt). Коэффициент пропорциональности обозначим через $\alpha_1(t)$.

2. Износ имеющихся вооружений. Соответствующее слагаемое в формуле для dm_i/dt отрицательно и описывается членом $-\beta_i(t)m_i$. Чем больше коэффициент $\beta_i(t)$, тем быстрее выбывают вооружения i -й страны.

3. Уровень общего недоверия к противнику, не зависящий от количества вооружений. Соответствующая поведенческая гипотеза состоит в том, что выражение для dm_i/dt содержит не зависящее от m_1 , m_2 слагаемое $\gamma_i(t)$. Заметим, что страны, участницы гонки вооружений, могут не доверять друг другу в разной степени. Если, например, первая страна не доверяет второй в большей степени, чем вторая страна не доверяет первой, то имеет место неравенство $\gamma_1(t) > \gamma_2(t)$.

Заметим, что перечисленные факторы неявно предполагают отсутствие «третьей стороны», которую участникам гонки вооружений приходилось бы также учитывать в своей политике.

Перечисленные положения позволяют сформулировать уравнения модели гонки вооружений в следующем виде:

$$\frac{dm_1}{dt} = \alpha_1(t)m_2 - \beta_1(t)m_1 + \gamma_1(t), \quad (1)$$

$$\frac{dm_2}{dt} = \alpha_2(t)m_1 - \beta_2(t)m_2 + \gamma_2(t). \quad (2)$$

Здесь $m_i(t) \geq 0$, $\alpha_i(t) > 0$, $\beta_i(t) > 0$ по содержательному смыслу этих величин. Функции $\gamma_i(t)$, вообще говоря, могут принимать отрицательные значения, что соответствует доверительным отношениям между странами («отрицательное недоверие»). Однако в этом случае политика первой и второй стран в области вооружений определяется, скорее, их взаимоотношениями с другими странами, а не между собой. Вследствие этого рассматриваемая нами модель гонки вооружений между двумя странами становится неприменимой. Поэтому всюду далее мы будем полагать, что функции $\gamma_i(t)$ неотрицательны.

Одна из возможных трактовок уравнения имеет следующий вид. Первая страна увеличивает количество своих вооружений ($dm_1/dt > 0$), если их текущее количество $m_1(t)$ меньше, чем величина $\bar{m}_1(t)$, даваемая выражением

$$\bar{m}_1(t) = \frac{1}{\beta_1(t)}(\alpha_1(t)m_2(t) + \gamma_1(t)),$$

и позволяет этому количеству уменьшиться в противоположном случае (и аналогично — для второй страны). Данное утверждение может трактоваться как поведенческая гипотеза, лежащего в основе модели гонки вооружений.

Следующий пункт посвящен анализу модели (1), (2) в случае постоянных коэффициентов.

5.6.2. Анализ базовой модели с постоянными коэффициентами

Итак, пусть коэффициенты $\alpha_i(t), \beta_i(t), \gamma_i(t)$ постоянны. Тогда система (1), (2) принимает вид

$$\frac{dm_1}{dt} = \alpha_1 m_2 - \beta_1 m_1 + \gamma_1, \quad (3)$$

$$\frac{dm_2}{dt} = \alpha_2 m_1 - \beta_2 m_2 + \gamma_2. \quad (4)$$

Исследуем систему (3), (4) с целью определить качественное поведение функций $m_1(t), m_2(t)$ во времени. Соответствующая стационарная задача имеет вид

$$\alpha_1 m_2 - \beta_1 m_1 + \gamma_1 = 0, \quad (5)$$

$$\alpha_2 m_1 - \beta_2 m_2 + \gamma_2 = 0. \quad (6)$$

Очевидно, что если $\beta_1\beta_2 - \alpha_1\alpha_2 \neq 0$, то положение равновесия является единственным и дается выражениями

$$m_1^* = \frac{\alpha_1\gamma_2 + \beta_2\gamma_1}{\beta_1\beta_2 - \alpha_1\alpha_2}, \quad m_2^* = \frac{\alpha_2\gamma_1 + \beta_1\gamma_2}{\beta_1\beta_2 - \alpha_1\alpha_2}. \quad (7)$$

Стационарное решение (7) имеет содержательный смысл лишь при выполнении неравенства

$$\beta_1\beta_2 - \alpha_1\alpha_2 > 0, \quad (8)$$

обеспечивающего положительность значений m_1^* , m_2^* .

Чтобы изучить устойчивость стационарного решения (7), перепишем систему (3), (4) в матричном виде:

$$\frac{d}{dt} \begin{pmatrix} m_1 \\ m_2 \end{pmatrix} = \begin{pmatrix} -\beta_1 & \alpha_1 \\ \alpha_2 & -\beta_2 \end{pmatrix} \begin{pmatrix} m_1 \\ m_2 \end{pmatrix} + \begin{pmatrix} \gamma_1 \\ \gamma_2 \end{pmatrix}. \quad (9)$$

Собственные значения матрицы системы (9) находятся из характеристического уравнения

$$\begin{vmatrix} -\beta_1 - \lambda & \alpha_1 \\ \alpha_2 & -\beta_2 - \lambda \end{vmatrix} = \lambda^2 + \lambda(\beta_1 + \beta_2) + \beta_1\beta_2 - \alpha_1\alpha_2 = 0, \quad (10)$$

имеющего дискриминант

$$\text{Discr} = (\beta_1 + \beta_2)^2 - 4(\beta_1\beta_2 - \alpha_1\alpha_2) = (\beta_1 - \beta_2)^2 + 4\alpha_1\alpha_2,$$

который положителен ввиду положительности коэффициентов α_1 , α_2 . Отсюда следует, что уравнение (10) имеет два различных вещественных корня. В соответствии с теоремой Виета сумма этих корней равна $\lambda_1 + \lambda_2 = -(\beta_1 + \beta_2) < 0$, а произведение: $\lambda_1\lambda_2 = \beta_1\beta_2 - \alpha_1\alpha_2$. Следовательно, если выполнено условие (8), то оба собственных значения λ_1 , λ_2 отрицательны и точка покоя (m_1^*, m_2^*) асимптотически устойчива. Гонка вооружений в этом случае называется *стабильной*. Если же $\beta_1\beta_2 - \alpha_1\alpha_2 < 0$, то один из корней λ_1 , λ_2 положителен и точка покоя (m_1^*, m_2^*) неустойчива.

Итак, неравенство (8) обуславливает не только положительность величин m_1^* , m_2^* , но и асимптотическую устойчивость стационарного решения (7). Содержательный смысл этого условия поясняется следующим рассуждением. Путь, например, параметры α_1 , β_1 и β_2 неизменны, а параметр α_2 увеличивается. Это означает, что первая страна не меняет свою стратегию в области вооружений, а вторая наращивает вооружения при неизменном темпе износа оружия (параметр β_2). Тогда при достаточно большой величине α_2 равновесие станет заведомо невозможным, а неравенство (8) обязательно нарушится. Заметим, что если оба параметра γ_1 , γ_2 , характеризующие взаимное недоверие, равны нулю, то положению равновесия отвечает отсутствие вооружений у обеих сторон.

Фазовый портрет системы (3), (4) в случае выполнения условия (8) представлен на рис. 1. Пунктирные прямые отвечают изоклинам нуля

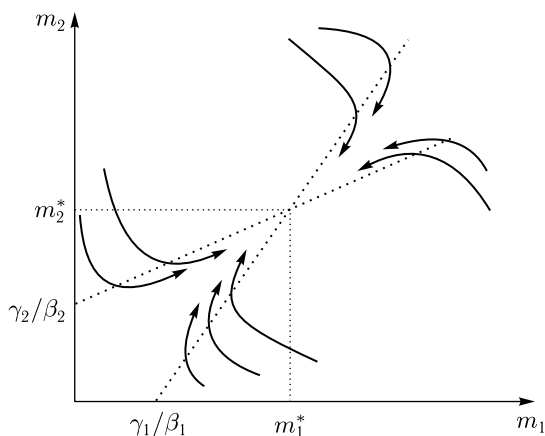


Рис. 1. Фазовый портрет системы (3), (4) в случае выполнения условия (8)

($m_2 = (\alpha_2/\beta_2)m_1 + \gamma_2/\beta_2$) и бесконечности ($m_2 = (\beta_1/\alpha_1)m_1 - \gamma_1/\alpha_1$). Изоклина нуля имеет наклон, меньший, чем изоклина бесконечности (это следует из неравенства (8)). Сплошными линиями соответствуют фазовые траектории. Стрелками показано направление движения по фазовым траекториям с течением времени. Функции $m_1(t)$, $m_2(t)$ при возрастании t стремятся к стационарным значениям.

Чтобы проанализировать развитие гонки вооружений в случае

$$\beta_1\beta_2 - \alpha_1\alpha_2 < 0, \quad (11)$$

вернемся к модели в матричной форме, именно, к уравнению (9). Нетрудно получить, что собственные значения λ_1 , λ_2 , определяемые уравнением (10), и соответствующие им собственные векторы имеют вид

$$\lambda_1 = -\frac{\beta_1 + \beta_2}{2} + \sqrt{\frac{(\beta_1 - \beta_2)^2}{4} + \alpha_1\alpha_2}, \quad (12)$$

$$h_1 = \begin{pmatrix} \alpha_1 \\ \frac{\beta_1 - \beta_2}{2} + \sqrt{(\beta_1 - \beta_2)^2/4 + \alpha_1\alpha_2} \end{pmatrix},$$

$$\lambda_2 = -\frac{\beta_1 + \beta_2}{2} - \sqrt{\frac{(\beta_1 - \beta_2)^2}{4} + \alpha_1\alpha_2}, \quad (13)$$

$$h_2 = \begin{pmatrix} \alpha_1 \\ \frac{\beta_1 - \beta_2}{2} - \sqrt{(\beta_1 - \beta_2)^2/4 + \alpha_1\alpha_2} \end{pmatrix}.$$

Нетрудно видеть, что $\lambda_2 < 0$ при любых положительных значениях параметров α_i, β_i , $i = 1, 2$, а выполнение неравенства (11) влечет $\lambda_1 > 0$.

Общее решение системы (9) имеет вид

$$\begin{pmatrix} m_1(t) \\ m_2(t) \end{pmatrix} = C_1 h_1 e^{\lambda_1 t} + C_2 h_2 e^{\lambda_2 t} + \begin{pmatrix} m_1^* \\ m_2^* \end{pmatrix}. \quad (14)$$

Следовательно, при $t \rightarrow +\infty$ численности вооружений первой и второй сторон неограниченно возрастают (гонка вооружений в этом случае называется *нестабильной*), и при этом относятся друг к другу как компоненты вектора h_1 :

$$\lim_{t \rightarrow \infty} \frac{m_1(t)}{m_2(t)} = \frac{\alpha_1}{(\beta_1 - \beta_2)/2 + \sqrt{(\beta_1 - \beta_2)^2/4 + \alpha_1 \alpha_2}}.$$

В частности, если вооружения изнашиваются с одинаковой скоростью, $\beta_1 = \beta_2$, то

$$\lim_{t \rightarrow \infty} \frac{m_1(t)}{m_2(t)} = \sqrt{\frac{\alpha_1}{\alpha_2}}.$$

Таким образом, если первая страна наращивает собственные арсеналы в ответ на имеющееся количество вооружений второй страны, например в три раза быстрее нее (см. соотв. поведенческую гипотезу выше), то в итоге она будет иметь преимущество в $\sqrt{3}$ раз.

5.6.3. Нелинейная модель

В реальной международной ситуации ни одна страна не имеет возможности интенсивно наращивать вооружения в течение значительного времени (не стоит путать рост вооружений с военными расходами: на практике большая часть расходов лишь компенсирует выбытие старого или выбывшего оружия). Основное ограничение связано с возможностями экономики.

В соответствии с этим в модель (3), (4) вносится следующее изменение. Положим, что количество вооружений i -й страны ($i = 1, 2$) не может превышать некоторой величины $M_i > 0$. Тогда вместо (3), (4) имеем

$$\frac{dm_1}{dt} = \left(1 - \frac{m_1}{M_1}\right) (\alpha_1 m_2 - \beta_1 m_1 + \gamma_1), \quad (15)$$

$$\frac{dm_2}{dt} = \left(1 - \frac{m_2}{M_2}\right) (\alpha_2 m_1 - \beta_2 m_2 + \gamma_2). \quad (16)$$

Решив соответствующую стационарную задачу, получаем, что система (15), (16) имеет четыре точки покоя:

$$P_1(m_1^*, m_2^*), P_2(M_1, M_2), P_3\left(\frac{\alpha_1 M_2 + \gamma_1}{\beta_1}, M_2\right), P_4\left(M_1, \frac{\alpha_2 M_1 + \gamma_2}{\beta_2}\right).$$

Чтобы исследовать их на устойчивость, вычислим матрицу Якоби системы (15), (16). Она имеет вид

$$J = \begin{pmatrix} -\frac{\alpha_1 m_2 - \beta_1 m_1 + \gamma_1}{M_1} - \beta_1 \left(1 - \frac{m_1}{M_1}\right) & \alpha_1 \left(1 - \frac{m_1}{M_1}\right) \\ \alpha_2 \left(1 - \frac{m_2}{M_2}\right) & -\frac{\alpha_2 m_1 - \beta_2 m_2 + \gamma_2}{M_2} - \beta_2 \left(1 - \frac{m_2}{M_2}\right) \end{pmatrix}. \quad (17)$$

Как показано ниже, знаки корней (либо их действительных частей, если корни комплексные) характеристического уравнения и, следовательно, поведение решений системы (15), (16) различаются в следующих четырех случаях:

$$(a) \quad m_1^* < M_1, \quad m_2^* < M_2, \quad (18)$$

$$(б) \quad m_1^* > M_1, \quad m_2^* > M_2, \quad (19)$$

$$(в) \quad m_1^* > M_1, \quad m_2^* < M_2, \quad (20)$$

$$(г) \quad m_1^* < M_1, \quad m_2^* > M_2, \quad (21)$$

(возможны также ситуации, при которых $m_i^* = M_i$ хотя бы для одной из стран; соответствующий анализ вынесен в упражнения).

Рассмотрим каждый из них поочередно.

Случай (а): $m_1^* < M_1, m_2^* < M_2$.

Вычислим якобиан (17) в точке покоя $P_1(m_1^*, m_2^*)$. Получим

$$J_1 = \begin{pmatrix} -\beta_1 \left(1 - \frac{m_1^*}{M_1}\right) & \alpha_1 \left(1 - \frac{m_1^*}{M_1}\right) \\ \alpha_2 \left(1 - \frac{m_2^*}{M_2}\right) & -\beta_2 \left(1 - \frac{m_2^*}{M_2}\right) \end{pmatrix}. \quad (22)$$

Характеристическое уравнение $\det(J_1 - \lambda I) = 0$ (где I — единичная матрица) имеет вид

$$\lambda^2 + \lambda \left[\beta_1 \left(1 - \frac{m_1^*}{M_1}\right) + \beta_2 \left(1 - \frac{m_2^*}{M_2}\right) \right] + [\beta_1 \beta_2 - \alpha_1 \alpha_2] \left(1 - \frac{m_1^*}{M_1}\right) \left(1 - \frac{m_2^*}{M_2}\right) = 0. \quad (23)$$

Дискриминант

$$\text{Discr}_1 = \left[\beta_1 \left(1 - \frac{m_1^*}{M_1}\right) - \beta_2 \left(1 - \frac{m_2^*}{M_2}\right) \right]^2 + 4\alpha_1 \alpha_2 \left(1 - \frac{m_1^*}{M_1}\right) \left(1 - \frac{m_2^*}{M_2}\right) \quad (24)$$

положителен ввиду неравенств (18), следовательно, корни $\lambda_{11}, \lambda_{12}$ (здесь и далее первая цифра в нижнем индексе соответствует номеру соответствующей точки покоя; вторая цифра — номеру корня) характеристического уравнения (23) вещественны. Их сумма равна

$$\lambda_{11} + \lambda_{12} = - \left[\beta_1 \left(1 - \frac{m_1^*}{M_1}\right) + \beta_2 \left(1 - \frac{m_2^*}{M_2}\right) \right], \quad (25)$$

а произведение —

$$\lambda_{11}\lambda_{12} = [\beta_1\beta_2 - \alpha_1\alpha_2] \left(1 - \frac{m_1^*}{M_1}\right) \left(1 - \frac{m_2^*}{M_2}\right). \quad (26)$$

Из неравенств (18), (25), (26) следует, что

- если $\beta_1\beta_2 > \alpha_1\alpha_2$, то $\lambda_{11} < 0$, $\lambda_{12} < 0$, точка покоя P_1 — устойчивый узел,
- если $\beta_1\beta_2 < \alpha_1\alpha_2$, то λ_{11} , λ_{12} имеют разные знаки, точка покоя P_1 — седло.

Перейдем к точке покоя $P_2(M_1, M_2)$. Якобиан (17) принимает в ней вид

$$J_2 = \begin{pmatrix} -\frac{\alpha_1 M_2 - \beta_1 M_1 + \gamma_1}{M_1} & 0 \\ 0 & -\frac{\alpha_2 M_1 - \beta_2 M_2 + \gamma_2}{M_2} \end{pmatrix}. \quad (27)$$

Корни характеристического уравнение $\det(J_2 - \lambda I) = 0$ равны

$$\lambda_{21} = -\frac{\alpha_1 M_2 - \beta_1 M_1 + \gamma_1}{M_1}, \quad (28)$$

$$\lambda_{22} = -\frac{\alpha_2 M_1 - \beta_2 M_2 + \gamma_2}{M_2}. \quad (29)$$

Можно показать (см. упражнение 1 к настоящему разделу), что если $\beta_2\beta_1 - \alpha_1\alpha_2 > 0$ и выполняются условия (18), то хотя бы одно их значений λ_{21} , λ_{22} положительно. Таким образом, точка покоя P_2 неустойчива.

Для точки $P_3((\alpha_1 M_2 + \gamma_1)/\beta_1, M_2)$ якобиан (17) имеет вид

$$J_3 = \begin{pmatrix} \frac{\alpha_1 M_2 - \beta_1 M_1 + \gamma_1}{M_1} & \alpha_1 \left(1 - \frac{\alpha_1 M_2 + \gamma_1}{\beta_1 M_1}\right) \\ 0 & \frac{(\beta_1\beta_2 - \alpha_1\alpha_2)(M_2 - m_2^*)}{\beta_1 M_2} \end{pmatrix}. \quad (30)$$

Таким образом,

$$\lambda_{31} = \frac{\alpha_1 M_2 - \beta_1 M_1 + \gamma_1}{M_1} = -\lambda_{21}, \quad (31)$$

$$\lambda_{32} = \frac{(\beta_1\beta_2 - \alpha_1\alpha_2)(M_2 - m_2^*)}{\beta_1 M_2}. \quad (32)$$

Аналогично, рассмотрев точку покоя $P_4(M_1, (\alpha_2 M_1 + \gamma_2)/\beta_2)$, получаем

$$\lambda_{41} = \frac{(\beta_1\beta_2 - \alpha_1\alpha_2)(M_1 - m_1^*)}{\beta_2 M_1}, \quad (33)$$

$$\lambda_{42} = \frac{\beta_2(\alpha_2 M_1 - \beta_2 M_2 + \gamma_2)}{M_2} = -\lambda_{22}. \quad (34)$$

Из (31) и (34) следует, что если точка покоя P_2 асимптотически устойчива (и имеет тип «узел», $\lambda_{21} < 0$, $\lambda_{22} < 0$), то P_3 , P_4 — неустойчивы. Если же точка P_2 неустойчива (хотя бы одна из величин λ_{21} , λ_{22} положительна), то каждая из точек P_3 , P_4 может быть как устойчивой, так и неустойчивой (в зависимости от параметров, в частности от значения выражения $\beta_1\beta_2 - \alpha_1\alpha_2$).

Тем самым оказываются возможными различные ситуации, некоторые из которых проиллюстрированы рис. 2–4.

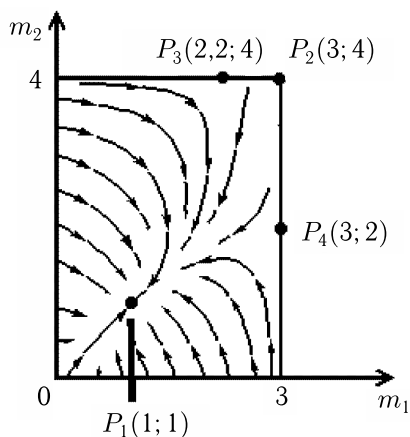


Рис. 2. Фазовый портрет системы (15), (16) при $\alpha_1 = \alpha_2 = 0,2$, $\beta_1 = \beta_2 = 0,4$, $\gamma_1 = \gamma_2 = 0,2$, $M_1 = 3$, $M_2 = 4$. Точка P_1 — устойчивый узел, P_2 — неустойчивый узел, P_3 , P_4 — седла

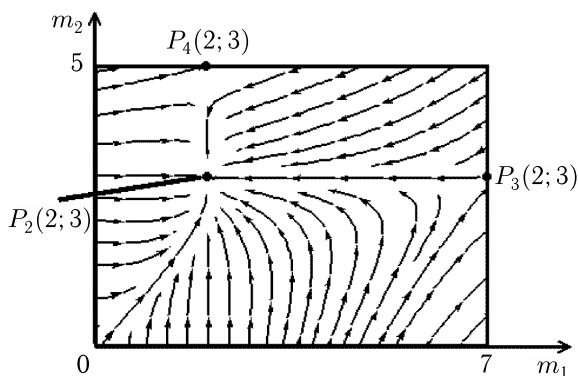


Рис. 3. Фазовый портрет системы (15), (16) при $\alpha_1 = \alpha_2 = 0,4$, $\beta_1 = \beta_2 = 0,2$, $\gamma_1 = \gamma_2 = 0,2$, $M_1 = 2$, $M_2 = 3$. Точка $P_1(-1; -1)$ (седло) находится в отрицательной области значений переменных, P_2 — устойчивый узел, P_3 , P_4 — седла

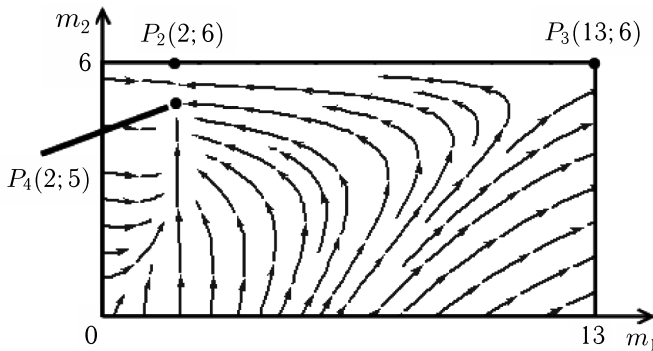


Рис. 4. Фазовый портрет системы (15), (16) при $\alpha_1 = \alpha_2 = 0,4$, $\beta_1 = \beta_2 = 0,2$, $\gamma_1 = \gamma_2 = 0,2$, $M_1 = 2$, $M_2 = 6$. Точка $P_1(-1; -1)$ (седло) находится в отрицательной области значений переменных, P_2, P_3 — седла, P_4 — устойчивый узел

С содержательной точки зрения неравенство (18), определяющее случай (а), означает, что обе страны обладают значительным экономическим потенциалом, позволяющим довести объемы вооружений до более высоких значений, чем равновесное состояние $P_1(m_1^*, m_2^*)$. Однако в случае стабильной гонки вооружений ($\beta_1\beta_2 - \alpha_1\alpha_2 > 0$) точка покоя $P_1(m_1^*, m_2^*)$ является асимптотически устойчивой и определяет количество вооружений при $t \rightarrow \infty$. Если же гонка нестабильна ($\beta_1\beta_2 - \alpha_1\alpha_2 < 0$), то хотя бы одна из сторон достигает максимально возможного (с точки зрения возможностей экономики) уровня.

Случай (б): $m_1^ > M_1$, $m_2^* > M_2$.*

Предварительно к анализу точек покоя заметим, что в данном случае (в отличие от рассмотренного выше) величины m_1^* , m_2^* могут быть только положительными. Следовательно, $\beta_1\beta_2 > \alpha_1\alpha_2$.

Рассмотрев точку покоя $P_1(m_1^*, m_2^*)$ и корни характеристического уравнения (23) аналогично тому, как это сделано для случая (а), получим неравенство $\lambda_{11} + \lambda_{12} > 0$. Следовательно, точка P_1 неустойчива.

Исходя из выражений (28), (29) для величин λ_{21} , λ_{22} нетрудно показать, что точка покоя P_2 асимптотически устойчива ($\lambda_{21} < 0$, $\lambda_{22} < 0$) тогда и только тогда, когда выполняются неравенства

$$\frac{\beta_1}{\alpha_1} M_1 - \frac{\gamma_1}{\alpha_1} < M_2, \quad (35)$$

$$M_2 < \frac{\alpha_2}{\beta_2} M_1 + \frac{\gamma_2}{\beta_2}. \quad (36)$$

Из этих неравенств также следует, что $\lambda_{31} = -\lambda_{21} > 0$, $\lambda_{42} = -\lambda_{22} > 0$ и точки P_3, P_4 неустойчивы. Нетрудно убедиться, что при условии (19) хотя бы одно из неравенств (35), (36) обязательно выполняется (мы оставляем вне рассмотрения случаи, при которых

хотя бы один из корней характеристического уравнения λ_{ij} равен нулю и метод исследования устойчивости по первому приближению не применим).

Если обратный знак имеет неравенство (35), то $\lambda_{21} > 0$. Тогда $\lambda_{31} = -\lambda_{21} < 0$, точка покоя P_3 асимптотически устойчива, точки P_2, P_4 неустойчивы. Аналогично, если обратный знак имеет неравенство (36), то точка покоя P_2 асимптотически устойчива, точки P_3, P_4 неустойчивы, рис. 5.

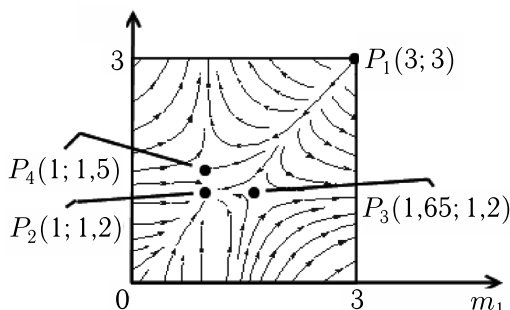


Рис. 5. Фазовый портрет системы (15), (16) при $\alpha_1 = \alpha_2 = 0,3$, $\beta_1 = \beta_2 = 0,4$, $\gamma_1 = \gamma_2 = 0,3$. Точка $P_1(3; 3)$ — неустойчивый узел, P_2 — устойчивый узел, P_3, P_4 — седла

Таким образом, в случае (б), определяемом неравенством (19), асимптотически устойчивой является ровно одна из точек покоя P_2, P_3, P_4 . Две другие из них, а также точка P_1 неустойчивы.

С содержательной точки зрения неравенство (19), определяющее случай (б), означает, что экономический потенциал каждой из стран достаточно мал — так, что участники гонки неспособны довести объемы вооружений до равновесного состояния $P_1(m_1^*, m_2^*)$, в результате чего объемы вооружений сторон остаются сравнительно низкими.

Анализ случаев (в) $m_1^* > M_1$, $m_2^* < M_2$ и (г) $m_1^* < M_1$, $m_2^* > M_2$ вынесен в упражнение 2.

5.6.4. Обобщение на случай большего количества стран

Модель гонки вооружений (9), обобщенная на случай произвольного количества стран, имеет вид

$$\frac{d}{dt} \begin{pmatrix} m_1 \\ m_2 \\ \dots \\ m_n \end{pmatrix} = \begin{pmatrix} -\beta_1 & \alpha_{12} & \dots & \alpha_{1n} \\ \alpha_{21} & -\beta_2 & \dots & \alpha_{2n} \\ \dots & \dots & \dots & \dots \\ \alpha_{n1} & \alpha_{n2} & \dots & -\beta_n \end{pmatrix} \begin{pmatrix} cm_1 \\ m_2 \\ \dots \\ m_n \end{pmatrix} + \begin{pmatrix} c\gamma_1 \\ \gamma_2 \\ \dots \\ \gamma_n \end{pmatrix}. \quad (37)$$

При этом становится возможным описание более широкого класса отношений между странами, чем в модели (9). Именно, гонка вооружений между двумя странами подразумевает недоверие между ними, что выражается положительностью коэффициентов α_1, α_2 в уравнении (9): чем более вооружена одна страна, тем более испытывает необходимость вооружаться другая. Другими словами, поведенческая гипотеза, лежащая в основе гонки вооружений двух стран, предполагает негативное отношение к арсеналам оппонента. Если же рассматривается система, состоящая из большего количества стран, то некоторые из могут находиться в союзнических отношениях.

Один из возможных типов поведения состоит в том, что «младший партнер» в военном блоке экономит на собственной армии, поскольку находится под защитой «старшего партнера». Другими словами, чем большим количеством вооружений обладает «старший партнер» (j -я страна), тем меньше вооружается «младший партнер» (i -я страна). Тем самым $\alpha_{ij} < 0$. В то же время количество вооружений «младшего партнера» не оказывает влияния на скорость роста вооружений «старшего»: $\alpha_{ji} = 0$. В качестве примера подобной ситуации можно привести отношения США и Японии. Если блок состоит из одного «старшего партнера» и нескольких «младших партнеров», то возможны различные варианты: например, каждый из «младших» может считать, что объем вооружений других «младших» также способствует его безопасности и тем самым создает возможность для уменьшения собственных арсеналов.

Возможны также следующие типы отношений между странами (и соответствующие им типы поведения):

- $\alpha_{ij} < 0, \alpha_{ji} < 0$ — военный альянс, в котором каждый из членов вносит вклад в обороноспособность другого (США, Великобритания);
- $\alpha_{ij} = 0, \alpha_{ji} = 0$ — нейтрально относящиеся друг к другу страны (Непал, Ботсвана);
- $\alpha_{ij} > 0, \alpha_{ji} = 0$ — i -я страна воспринимает j -ю как угрозу, но при этом j -я страна нейтрально относится к объему вооружений i -й (Монголия, США);
- $\alpha_{ij} > 0, \alpha_{ji} > 0$ — антагонистические отношения (СССР, США).

Эти соотношения позволяют описывать различные внешнеполитические конфигурации. Путь, например, в гонке вооружений участвуют два блока:

- А) страна 1 является «старшим партнером» в блоке, страны 2, 3 — «младшие партнеры», которые воспринимаются как угроза странами блока Б;
- Б) страны 4, 5 образуют военный альянс, в котором каждый из членов вносит вклад в обороноспособность другого и воспринимается как угроза странами блока А.

Такая внешнеполитическая конфигурация описывается следующей матрицей:

$$\begin{pmatrix} - & 0 & 0 & + & + \\ - & - & - & + & + \\ - & - & - & + & + \\ + & + & + & - & - \\ + & + & + & - & - \end{pmatrix}$$

(на месте коэффициентов α_{ij} , $-\beta_i$ в формуле (37) стоят их знаки).

Аналогичным образом могут быть описаны другие конфигурации. Соответствующие модели анализируются теми же методами, что и модель гонки вооружений двух стран.

Упражнения

1. Рассмотреть случай (а) $m_1^* < M_1$, $m_2^* < M_2$, точку покоя $P_2(M_1, M_2)$: доказать, что если $\beta_2\beta_1 - \alpha_1\alpha_2 > 0$ и выполняются условия (18), то хотя бы одно из значений λ_{21} , λ_{22} , даваемых выражениями (28), (29), положительно. Таким образом, точка покоя P_2 неустойчива.

2. Проанализировать случаи (в) $m_1^* > M_1$, $m_2^* < M_2$ и (г) $m_1^* < M_1$, $m_2^* > M_2$. В частности: доказать, что в каждом из них имеет место неравенство $\beta_1\beta_2 > \alpha_1\alpha_2$, затем доказать, что дискриминант характеристического уравнения положителен (при этом рекомендуется сделать замены $|\beta_1(1 - m_1^*/M_1)| = p$, $\beta_2(1 - m_2^*/M_2) = q$).

Комментарии и библиографические примечания

1. Базовая модель гонки вооружений была разработана английским физиком и метеорологом Л. Ф. Ричардсоном и опубликована в книге [1]. После этого в 60-х годах XX века появилось значительное количество работ, посвященных моделированию гонки вооружений. Так, библиография в обзоре [2], сделанном в 1984 году, содержит 224 публикации по данной тематике (см. также более поздний обзор [3]).

2. Нелинейная модель (15), (16) предложена в работе [4]. Нетрудно видеть, что при переходе от линейной модели к нелинейной использован тот же подход, что в демографии (см. главу 1): от модели Мальтуса к логистической модели.

3. Обобщение на случай произвольного количества стран проведено в работе [5].

Литература

1. Richardson L.F. Arms and Insecurity. Pittsburgh. Boxwood Press. 1960.
2. Anderton C.H. Arms Race Modeling: Categorization and Systematic Analysis. Paper presented at the International Studies Association, Washington. 1985.

3. Schrod P. A. Patterns, Rules and Learning: Computational Models of International Behavior. 2004 (2nd Edition). Parus Analytical Systems, Vinland, Kansas, USA.
4. Lehmann B., McEwen J., Lane B. Modifying the Richardson Arms Race Model With a Carrying Capacity.
5. Schrod P. A. Richardson's N-nation model and the balance of power // American Journal of Political Science. 1978. V. 5 № 1. P. 364–390.

5.7. Моделирование крупномасштабных исторических процессов

Данный раздел посвящен моделям протяженных (как в пространстве, так и во времени) процессов. С математической точки зрения такие модели обычно имеют вид нелинейных систем обыкновенных дифференциальных уравнений либо уравнений с дискретным временем. При этом модели, нацеленные на выявление основных качественных закономерностей процесса, могут быть относительно компактными, в то время как задача прогнозирования обычно приводит к системе большого количества уравнений, для исследования которой требуется анализ численными методами.

Краткий обзор методов и подходов к моделированию крупномасштабных процессов приводится в первом пункте. Второй пункт посвящен изучению наиболее крупномасштабных, с точки зрения географического охвата, процессов. Другими словами, рассматривается одна из так называемых *моделей мировой динамики*. Общим для всех моделей этого класса является подход, связанный с опорой на высокоагрегированные переменные — например, на объем экономики всего мира.

Следующие два пункта посвящены процессам, протяженным во времени, т. е. моделированию исторических процессов. В третьем изучается конкретная социально-историческая формация, именно, аграрное общество. Четвертый пункт посвящен еще более конкретному предмету: рассматривается модель системы древнегреческих полисов в период Пелопоннесской войны (431–404 гг. до н.э.).

5.7.1. Введение

Во множестве разнообразных методов моделирования крупномасштабных процессов можно условно выделить несколько взаимодействующих между собой подходов. Экстраполяционные методы основаны на аппарате математической статистики. Методы экспертных оценок имеют особую роль в данной области, т. к. некоторые характеристики, относящиеся к будущему периоду (например, темпы технологического развития в предстоящие десятилетия) не могут быть оценены иначе как экспертно. Сценарный подход предполагает построение одного или, чаще, нескольких сценариев, например верхнего (иногда понимаемого как оптимистический), среднего и нижнего (пессимистического).

Математическое моделирование позволяет интегрировать все перечисленные подходы. Например, модель может иметь вид системы обыкновенных дифференциальных уравнений и иметь в качестве входных данных оценки некоторых величин (таких, как коэффициенты рождаемости), полученные путем экстраполяции, других величин (темпов роста технологий) полученных экспертно. В свою очередь результат математического моделирования т. е. прогноз, может иметь вид набора сценариев.

При этом роль математического моделирования не исчерпывается построением прогнозов. Оно предлагает исследовательский аппарат, позволяющий выявлять и анализировать как количественные, так и качественные закономерности крупномасштабных процессов.

Эти положения иллюстрируются примерами, представленными в следующих пунктах настоящего раздела.

5.7.2. Элементы модели мировой динамики

Известны несколько различных моделей мировой динамики, (см. примечание 1 в конце раздела) и всем им присущи определенные характерные черты, в первую очередь сравнительно большое количество нелинейных уравнений, сгруппированных в несколько блоков, высокая степень агрегирования переменных, а также опора на численный анализ при исследовании системы. В качестве примера рассмотрим наиболее известную и раннюю модель такого рода, известную как модель Форрестера (далее будет называть ее просто моделью мировой динамики).

Предметом исследования является мировая система, под которой понимается совокупность взаимодействующих друг с другом социальных систем, технологий и окружающей среды, включающей в себя ресурсные запасы. Модель состоит из 22 соотношений, включающих в себя как нелинейные динамические уравнения в дискретном времени, так и связи между значениями различных функций, взятыми в один и тот же момент времени. Не имея целью полное описание модели, рассмотрим лишь некоторые ее уравнения из числа наиболее важных.

Динамика численности населения $P(t)$ описывается уравнением

$$P(t+1) = P(t)[1 + b(c, f, d, z) - m(c, f, d, z)], \quad (1)$$

где b , m — коэффициенты рождаемости и смертности соответственно, зависящие от удельного потребления c , удельного выпуска продуктов питания f , плотности населения d , объема экологических загрязнений z .

Потребление непродовольственных благ $c(t)$ принимается пропорциональным такой величине, как объем капитала, непосредственно участвующего в создании этих благ, приходящегося в среднем на одного индивида:

$$c(t) = ax(t). \quad (2)$$

Выражение для только что введенной величины имеет вид

$$x(t) = \frac{AR(t)K(t)}{P(t)}, \quad (3)$$

где A — некоторая постоянная, а $R(t)$, $K(t)$ — объемы природных ресурсов и капитала соответственно. Динамика этих переменных описывается уравнениями

$$R(t) = R(t-1) - \theta(c(t))P(t), \quad (4)$$

$$K(t) = K(t-1) - \delta K(t-1) + s(c(t))P(t). \quad (5)$$

Здесь функция $\theta(c(t))$ описывает истощение природных ресурсов в зависимости от уровня потребления, δ — норма выбытия капитала, $s(c(t))$ — норма сбережения. Функции $\theta(c)$, $s(c)$ являются монотонно возрастающими и задаются экзогенно.

Совокупность уравнений (2)–(5) представляет собой блок модели мировой динамики, описывающий производство и потребление непродовольственных благ. Если в какой-то конкретной ситуации можно было бы считать численность населения $P(t)$ заданной (например, постоянной), то данные уравнения представляли бы замкнутую систему

Другие уравнения модели включают описание динамики производства продовольствия, количества свободной земли и сельхозугодий, неустранимых экологических загрязнений и т. д. В целом, модель представляет собой замкнутую (и не распадающуюся на независимые подсистемы) систему динамических уравнений. Ее исследование проводится численными методами.

Критика модели мировой динамики обсуждается в комментариях к данному разделу.

5.7.3. Динамика аграрных обществ

В качестве примера моделирования крупномасштабных исторических процессов рассмотрим модель аграрных обществ. В данной модели отсутствует страновая спецификация; другими словами, моделируется тип общественно-экономической формации, а не конкретная страна.

В отличие от моделей мировой динамики, данный подход не содержит внутри себя проблему излишне высокой степени агрегирования. Это связано с тем, что составляющие элементы конкретного типа общества (в частности, аграрного) существенно более однородны, чем элементы современного мира, в котором присутствуют страны и даже континенты, находящиеся на разных ступенях развития. Так, абсолютное большинство населения аграрной страны — это крестьяне, находящиеся в сходных социально-экономических условиях. В частности, это позволяет принять факторы рождаемости и смертности общими для всего населения. Экономические субъекты аграрного общества — это, в основном, сельскохозяйственные предприятия, использующие одни и те же ресурсы и находящиеся на сопоставимых уровнях

технологического развития. Поэтому применение таких переменных, как совокупная численность индивидов, совокупный выпуск и т. д. для описания базовых черт демографической и экономической динамики в данном случае выглядит более обоснованным, чем в моделях типа Форрестера, рассматривающих мир как единую систему.

В целом, социально-экономическая динамика описывается в модели аграрного общества следующим образом. Основными переменными являются численность населения и объем накопленных обществом запасов, причем демографическая динамика зависит от количества запасов. Сами запасы формируются как остаток произведенного продукта после вычета потребления. Производство продукта описывается функцией, зависящей, в частности от численности населения. Тем самым модель имеет вид замкнутой системы уравнений.

Специфика аграрного общества учитывается особым характером производственной функции: именно, если количество крестьян достаточно велико, то их наделы уменьшаются ввиду конечности земельного ресурса. При этом производительность (максимально количество аграрного продукта, которое может быть получено с одной единицы земли) также полагается ограниченной функцией. Сочетание таких факторов, как ограниченность количества обрабатываемой земли и ограниченность производительности, приводит к ограниченности выпуска аграрной экономики. Вследствие этого модель не имеет неограниченно возрастающих решений: в системе возникают либо устойчивые стационарные состояния, либо циклы. Это обстоятельство отличает модель аграрного общества от моделей экономического роста, широко распространенных в математической экономике, а также от модели мировой динамики, рассмотренной в предыдущем пункте.

Перейдем к построению модели. Демографическая динамика описывается уравнением

$$\frac{dN}{dt} = ND(x).$$

Здесь $N(t)$ — численность населения, x — средний объем материальных накоплений, приходящийся на одного индивида, который можно с определенной степенью условности ассоциировать с уровнем жизни. Очевидно, общий объем накоплений общества равен $X = xN$.

Темп демографического роста $D(x)$ обладает следующими свойствами:

- 1) функция $D(x)$ — монотонно возрастающая, определена при $x > 0$,
- 2) $D(x) > 0$ при значениях x , превышающих некоторое значение x_d , а $D(x) < 0$ при $x < x_d$, т. е. x_d — критическое значение выпуска, при котором рождаемость равна смертности,
- 3) $D(x)$ ограничена при $x \rightarrow \infty$,
- 4) $D(x) \rightarrow -\infty$ при $x > 0$, поскольку в голод численность может убывать сколь угодно быстро ввиду высокой смертности.

Простейшей функцией, удовлетворяющей данным свойствам, является

$$D(x) = r \left(1 - \frac{x_d}{x} \right).$$

Выпуск экономики аграрного общества описывается производственной функцией вида

$$Y = a(x)NR(N),$$

где a — производительность одного крестьянина с единицы площади обрабатываемой земли, $R(N)$ — площадь, обрабатываемая одним крестьянином. При этом принимается, что

$$R(N) = \begin{cases} s, & N \leq S_0/s, \\ \frac{S_0}{N}, & N \geq S_0/s, \end{cases} \quad (6)$$

где s — максимальная площадь земли, обрабатываемая одним крестьянином, S_0 — доступное количество земли. Формулу (6) можно переписать в виде

$$R(N) = \min \left(s, \frac{S_0}{N} \right).$$

Производительность $a(x)$ обладает следующими свойствами.

5. Функция $a(x)$ определена при всех $x \geq 0$ и ограничена сверху и снизу положительными константами.

6. Функция $a(x)$ монотонно возрастающая, так как при низком уровне жизни (т.е. при малых значениях x) возрастает социальная нестабильность, препятствующая нормальному ведению хозяйства и приводящая к снижению производительности труда.

Простейшей функцией, удовлетворяющей данным свойствам, является

$$a(x) = a_0 + a_1 \frac{x^2}{x^2 + x_a^2}.$$

Рост накоплений общества происходит за счет разницы между выпуском Y и потреблением $C = cN$, причем удельное потребление c описывается линейной функцией

$$c(x) = c_0 + c_1x.$$

С учетом сделанных допущений базовая модель аграрного общества имеет следующий вид:

$$\frac{dN}{dt} = Nr \left(1 - \frac{x_d}{x} \right), \quad (7)$$

$$\frac{d(xN)}{dt} = N \left[\left(a_0 + a_1 \frac{x^2}{x^2 + x_a^2} \right) \min \left(s, \frac{S_0}{N} \right) - (c_0 + c_1x) \right]. \quad (8)$$

Переходя от удельных материальных накоплений к суммарным по формуле $X = xN$, перепишем модель в виде

$$\frac{dN}{dt} = Nr \left(1 - \frac{Nx_d}{X} \right), \quad (9)$$

$$\frac{dX}{dt} = \left(a_0 + a_1 \frac{X^2}{X^2 + N^2 x_a^2} \right) \min(Ns, S_0) - (c_0 N + c_1 X). \quad (10)$$

Проведем анализ модели (9), (10) при $a_1 = 0$ (в случае $a_1 > 0$ см. примечание в конце раздела). Производительность труда в этом случае является постоянной: $a = a_0$. Система (9), (10) принимает вид

$$\frac{dN}{dt} = Nr \left(1 - \frac{Nx_d}{X} \right), \quad (11)$$

$$\frac{dX}{dt} = a_0 \min(Ns, S_0) - (c_0 N + c_1 X). \quad (12)$$

Рассмотрим вопрос о существовании стационарного решения данной системы. Положив производные нулю,

$$\begin{aligned} Nr \left(1 - \frac{Nx_d}{X} \right) &= 0, \\ a_0 \min(Ns, S_0) - (c_0 N + c_1 X) &= 0, \end{aligned}$$

выразив из первого уравнения $X = Nx_d$ и подставив во второе, получим, что

$$a_0 \min(Ns, S_0) - N(c_0 + c_1 x_d) = 0.$$

Так как $N = X = 0$ не принадлежит области определения правой части уравнения (11) (хотя и является аттрактором — см. упражнение 1), то стационарное решение существует, лишь если

$$\min(N_{st} s, S_0) = S_0, \quad (13)$$

где N_{st} — искомое стационарное состояние. В этом случае имеем

$$N_{st} = \frac{a_0 S_0}{c_0 + c_1 x_d}, \quad X_{st} = \frac{x_d a_0 S_0}{c_0 + c_1 x_d}. \quad (14)$$

С учетом (13) получаем, что система (11), (12) имеет стационарное решение при выполнении условия $s > s_{cr} = (c_0 + c_1 x_d) / a_0$, и это решение имеет вид (14).

С содержательной точки зрения данный результат означает, что стационарное решение существует, лишь если средний крестьянин способен обработать достаточно большое количество земли: $s > s_{cr}$. При этом критическое значение s_{cr} тем выше, чем выше практикуемый в данном обществе уровень потребления $c_0 + c_1 x_d$ и чем ниже параметр a_0 , характеризующий потребление.

Если $s \leq s_{cr}$ (см. упражнение 1), то $N(t) \rightarrow 0$, $X(t) \rightarrow 0$ при $t \rightarrow \infty$, хотя точка $X = N = 0$, строго говоря, не является стационарным решением, т. к. не принадлежит области определения правой

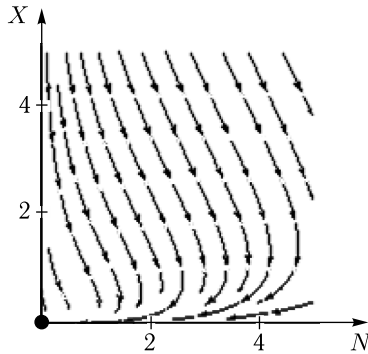


Рис. 1. Фазовая плоскость системы (11), (12) при $s \leq s_{cr}$. Параметры расчета: $r = 0,5$; $x_d = 0,3$; $a_0 = 0,4$; $s = 0,9$; $c_0 = 0,8$; $c_1 = 0,3$

части уравнения (11). Пример такой ситуации приведен на рис. 1: все фазовые траектории ведут в начало координат.

Если же $s > s_{cr}$, то стационарное решение имеет вид

$$N_{st} = \frac{a_0 S_0}{c_0 + c_1 x_d}, \quad X_{st} = \frac{x_d a_0 S_0}{c_0 + c_1 x_d}.$$

Исследуем его на устойчивость. Матрица Якоби дается выражением

$$J = \begin{pmatrix} r(1 - 2Nx_d/X) & N^2 r x_d / X^2 \\ -c_0 & -c_1 \end{pmatrix}.$$

При $N = N_{st}$, $X = X_{st}$ получаем характеристическое уравнение

$$\lambda^2 + \lambda(c_1 + r) + c_1 r + c_0 r / x_d = 0.$$

Его корни равны

$$\lambda_{1,2} = 0,5 \left[-(c_1 + r) \pm \sqrt{(c_1 - r)^2 - 4c_0 r / x_d} \right]. \quad (15)$$

Отсюда следует, что в зависимости от знака подкоренного выражения стационарное решение является либо устойчивым узлом (рис. 2, а), либо устойчивым фокусом (рис. 2, б).

Таким образом, если выполнено неравенство $s > s_{cr}$, то система (9), (10) стабилизируется к стационарному состоянию, в котором как численность населения, так и объем материальных запасов общества (а следовательно, и объем запасов на одного крестьянина) являются постоянными и положительными величинами. Данное неравенство можно переписать в виде $a_0 s_{cr} > c_0 + c_1 x_d$. В его левой части стоят параметры производства (производительность земли, и количество земли, обрабатываемой одним крестьянином), в правой — параметры потребления. Поэтому полученный результат можно трактовать следующим образом:

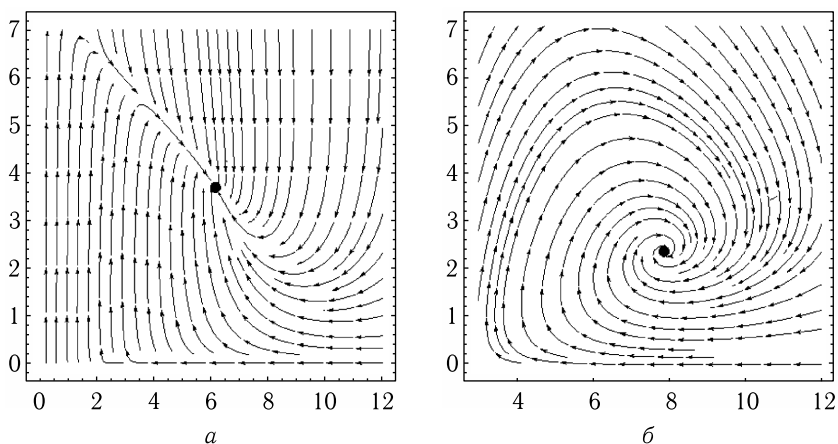


Рис. 2. Фазовая плоскость системы (11), (12) при $s > s_{cr}$. Параметры расчета:
 а) $r = 0,1$; $x_d = 0,6$; $a_0 = 0,7$; $s = 1,5$; $c_0 = 0,6$; $c_1 = 0,9$; $S_0 = 10$; б) $r = 0,5$;
 $x_d = 0,3$; $a_0 = 0,7$; $s = 1,5$; $c_0 = 0,8$; $c_1 = 0,3$; $S_0 = 10$

если производство способно покрыть потребление, то система (11), (12) имеет положительное стационарное решение.

Если же $a_0 s_{cr} < c_0 + c_1 x_d$, то с течением времени численность населения и материальных запасов убывают до нуля (упражнение 1).

Наконец, если рассматривается более полная модель (9), (10), то численные эксперименты показывают, что в ней возможны также периодические колебания, имеющие смысл демографических циклов. Именно, фаза демографического роста приводит к недостаточности ресурсов и (вследствие этого) к демографическому спаду. В свою очередь, это приводит к избылию ресурсов и демографическому росту, который затем сменяется новым спадом (см. примечание 3 в конце раздела).

5.7.4. Моделирование системы древнегреческих полисов

Настоящий пункт посвящен краткому описанию модели процессов производства, распределения, обмена и потребления в древнегреческих полисах в период Пелопоннесской войны (431–404 г. до н.э.).

Полисы представляли собой рабовладельческие города-государства, большая часть из которых была объединена в две коалиции: Пелопоннесский союз (во главе со Спартой) и Афинский союз. Многолетние политические и экономические трения между этими коалициями переросли в 431 г. в затяжную и крупномасштабную войну.

При построении модели в каждом полисе выделяется 14 социальных слоев: крупные землевладельцы, средние землевладельцы, крестьяне-арендаторы, богатые, средние и мелкие ремесленники, рабы и т. д. Учитывается 11 видов товаров: зерно, вино, масло, сельскохозяйственные орудия труда, рабы и т. д. Производство каждого товара каждым слоем учитывается отдельно, при этом принимаются во вни-

мание исторические данные о площади земель и численности рабов у каждого слоя, а также об урожайности. То же относится к ремесленной деятельности. Часть произведенного продукта потребляется самим производителем и его рабами (согласно нормам потребления), остаток выставляется на рынок.

Сфера обмена учитывает, что каждый торгующий слой выставляет на продажу свою продукцию, приобретая продукцию других слоев. При этом принимается во внимание внешняя торговля каждого полиса (с другими полисами); модель учитывает три вида пошлин: на ввоз товара, на вывоз и на торговлю на рынке. Механизм обмена имеет следующий вид. По каждому товару подсчитывается общее предложение и общий спрос, и в случае излишка этого товара спрос считается удовлетворенным полностью, а излишек распределяется пропорционально между продавцами. Аналогично, в случае дефицита он распределяется между покупателями.

Денежные потоки на рынке соответствуют товарным, при этом принимаются во внимание имеющиеся исторические данные об уровне цен. Денежные потоки, связанные с государством, включают поступления от пошлин и налогов, а также расходы, в т.ч. на ведение войны: жалование воинов, постройка и ремонт кораблей.

Специфика военного времени учитывается также в описании трудового ресурса: в годы активных боевых действий его величина уменьшается на количество мобилизованных крестьян и ремесленников, которые (за вычетом потерь) возвращаются к хозяйственной деятельности окончании боев. Наконец, учитывается уменьшение количества сельскохозяйственных земель в результате опустошительных походов вражеской армии. При восстановлении земель считается, например, что виноградники начинают давать продукцию через три года после разорения, оливковые деревья — через 12–15 лет. Военная добыча распределяется между воинами и учитывается в объеме запасов соответствующих слоев населения.

В целом, модель представляет собой довольно детальное описание исследуемой системы. С математической точки зрения она представляет собой систему динамических уравнений с дискретным временем. Ввиду ее громоздкости ограничимся лишь одним уравнением, иллюстрирующим модель.

Так, военные расходы i -го полиса в год t равны

$$VR_i(t) = C1_i(t)VK_i(t) + C2_i(t)HK_i(t) + \sum_{j=1}^4 \sum_{k=2}^{14} C3_{ij}(t)W'_{ikj}(t) + \sum_{k=1}^{14} C3_{i4}(t)WR'_{ik}(t).$$

Здесь $C1_i$ — стоимость ремонта одного корабля в i -м полисе, $C2_i$ — стоимость постройки одного нового корабля в i -м полисе, $C3_i$ — годовая плата одному воину в i -м полисе в войске рода j (всадники: $j = 1$,

гоплиты, т. е. тяжеловооруженные воины: $j = 2$, легковооруженные воины: $j = 3$, служащие во флоте: $j = 4$), VK_i — количество кораблей в i -м полисе, построенных ранее данного года, HK_i — количество кораблей, построенных в данный год в i -м полисе, W'_{ikj} — количество людей, взятых в i -м полисе из социального слоя k в войско рода j , WR'_{ik} — количество рабов, взятых во флот в i -м полисе у социального слоя k .

Для данных переменных, разумеется, в модели содержатся свои уравнения. В системе уравнений содержатся также параметры; значения многих из них оценены на основании исторической информации. Например, при определении потребности в одежде учитывалось, что срок службы хитона и плаща составляет два года, наемный работник зарабатывал в год 360 драхм, или 2160 оболмов, нормы потребления зерна и мяса составляли, соответственно, 66 оболмов и 30 оболмов (данное слово обозначает как меру веса, так и денежную единицу) и т. д. Некоторые параметры квантифицировались на основе общих представлений: например, норма в модели положена норма потребления вина 180 литров в год.

Для оставшихся параметров проводилась калибровка модели. Этот метод характерен для изучения систем с недостаточным количеством информации о численных значениях входящих в нее величин. Проводится серия вычислительных экспериментов, в которой варьируются неизвестные параметры и изучаются полученные таким образом решения. Как правило, для части диапазона варьируемых параметров получаются очевидно непригодные результаты: либо система демонстрирует качественно несодержательное решение (отсутствие осмысленных стационарных решений, стремление к нулю или отрицательным значениям вместо положительных и т. д.), либо для каких-то переменных получатся значения, несовместимые с существующими представлениями (например, неправдоподобно высокие для данной социально-экономической системы значения продуктивности). Такие значения варьируемых параметров отвергаются, и оставшаяся область их изменения представляет собой конечную оценку данных параметров. Такого рода оценки являются иногда основным результатом исследования.

Упражнения

1. Рассмотреть систему (11), (12) в случае, когда $s < s_{cr}$. Показать, что $N(t) \rightarrow 0$, $X(t) \rightarrow 0$ при $t \rightarrow \infty$, хотя точка $X = N = 0$, строго говоря, не является стационарным решением, т. к. не принадлежит области определения правой части уравнения (11). Указание: принять во внимание, что решение притягивается к прямой $X = Nx_d$.

2. Положив $a_1 \ll 1$ в уравнении (8), построить первый член асимптотического разложения решения системы (7), (8) по малому параметру a_1 в случае $s > s_{cr}$.

Комментарии и библиографические примечания

1. Широкий обзор моделей мировой динамики, а также методов и подходов их исследования приведен в [1]. В этой же монографии обсуждаются, в частности, вопросы применения моделирования и прогнозирования в системах поддержки принятия решений. Анализ места моделей исторических процессов в общем контексте исторической науки проводится в [2].

2. Исторически первые модели мировой динамики были построены Дж. Форрестером [3], Д. Медоузом и др. [4]. Более современные модели, описывающие, в частности неравновесные процессы и кризисы, предложены в [1].

3. Критики указывают [5], что описание производства в модели мировой динамики (см. формулы (2)–(5)) — недостаточно прозрачное, поскольку не содержит в явном виде понятие экономического выпуска; при этом сведение данных положений модели мировой динамики к общепринятому и более прозрачному методу производственных функций выявляет несостоятельность такой модели производственного процесса.

Более конкретно: поскольку потребление — это часть выпуска (за исключением редких и коротких периодов, когда имеется возможность «продать» выпуски прошлых лет), то для каждого конкретного года выпуск $y(t)$ превышает потребление:

$$y(t) > c(t) = ax = \frac{aAR(t)K(t)}{P(t)}.$$

Выражение в правой части содержит произведение факторов производства, которыми в данной модели являются объемы ресурсов и капитала. Таким образом, например при двукратном увеличении количества доменных печей (т. е. капитала) и доступной руды (ресурса) выпуск увеличивается в четыре раза. В противоположность этому как экономическая теория, так и практика свидетельствуют, что при увеличении ресурса и капитала в некоторое количество раз выпуск возрастает примерно во столько же раз. Другими словами, производственная функция должна обладать свойством однородности: $y(\lambda R, \lambda K) = \lambda y(R, K)$, $\lambda > 0$.

Данная проблема носит частный характер, она присутствует в конкретной модели Форрестера. Однако указывается также на ряд других проблем, свойственных всему классу моделей мировой динамики, в частности: чрезвычайно высокий уровень агрегирования переменных. Например, численность населения учитывается совокупно для всего мира, вследствие чего, к примеру, факторы рождаемости полагаются одинаковыми для всех стран мира; поэтому соответствующие уравнения не имеют обоснований в виде гипотез о поведении индивидов. Действительно, для такого уровня агрегирования просто не существует единой гипотезы о репродуктивном поведении. Таким образом, уравнения модели Форрестера и других моделей мировой динамики не

основаны на поведенческих гипотезах. Аналогичным образом они не основаны на эмпирических закономерностях, экономических и социальных законах, а получены лишь на основании общих соображений.

Другая линия критики основана на том, что модели мировой динамики игнорируют такие существенные экономические факторы, как технический прогресс и связанное с ним замещение ресурсов (например, по мере развития технологий, угля, в качестве источника энергии, замещается нефтью, затем газом, возобновляемыми источниками и т. д.).

Наиболее резкое отношение ряда экономистов как к заложенным в модель Форрестера положениям, так и полученным на ее основе прогнозам высказал нобелевский лауреат П. Кругман: «Мусор заложили — мусор получили» [6].

С другой стороны, сторонники данного подхода указывают, что в отношении некоторых переменных прогнозы различных моделей мировой динамики оказываются довольно точными (см., напр., [7, 8]). Одно из возражений против этого аргумента состоит в том, что для этих переменных прогнозы либо лишь продолжают устоявшиеся тренды, доступные для простейших регрессионных прогнозов, либо имеют вид довольно широкой «вилки» между различными сценариями.

В целом, можно сказать, что дискуссии вокруг моделей мировой динамики продолжаются более сорока лет и пока не привели к консенсусу.

4. Построение модели аграрного общества излагается по работе [9], там же представлены численные эксперименты в случае $a_1 > 0$. Анализ случая $a_1 = 0$ проводится впервые.

5. Если в системе (9), (10) производительность крестьянина зависит от объема материальных накоплений, т. е. $a_1 \neq 0$ (а также, конечно, $a_0 > 0$, $a_x > 0$), то при некоторых значениях параметров в ней возникает предельный цикл. Другими словами, как численность населения $N(t)$, так и общий объем $X(t)$ испытывают нелинейные периодические колебания. Этот случай разобран в работе [5].

6. Подробному изложению модели системы древнегреческих полисах в период Пелопоннесской войны посвящена книга [10]. Более краткое описание в рамках общего контекста исследований данной тематике дается в монографии [11].

Литература

1. Садовничий В. А., Акаев А. А., Коротаев А. В., Малков А. В. Моделирование и прогнозирование мировой динамики // М.: ИСПИ РАН, 2012. 359 с.
2. Малинецкий Г. Г. Нелинейная динамика — ключ к теоретической истории? // Общественные науки и современность. 1996. Т. 4. Р. 98–111.
3. Forrester J. W. World Dynamics. Cambridge, MA: Wright-Allen Press. 1971.

4. Meadows D. H., Meadows D. L., Randers J., Behrens W. W. III. The Limits to Growth: A Report for the Club of Rome's Project on the Predicament of Mankind. New York, NY: Universe Books. 1972.
5. Nordhaus W.D. World Dynamics: Measurement Without Data // The Economic Journal. V. 83, № 332 (Dec., 1973). P. 1156–1183.
6. Krugman P. Limits to growth and related stuff. April 22, 2008. URL: <https://krugman.blogs.nytimes.com/2008/04/22/limits-to-growth-and-related-stuff>.
7. Turner G.M.. A comparison of The Limits to Growth with 30 years of reality // Global environmental change. 2008. V. 18, № 3. P. 397–411.
8. van Vuuren D.P., Faber A. Growing within Limits — A Report to the Global Assembly 2009 of the Club of Rome. Netherlands Environmental Assessment Agency. 2009.
9. Малков С. Ю. Математическое моделирование исторической динамики: подходы и модели // Дмитриев М. Г., Дубовский С. В., Жукова Г. С., Малков С. Ю., Михайлов А. П., Петров А. П., Степанцов М. Е. Моделирование социально-политической и экономической динамики. М.: РГСУ. 2004. С. 76–188.
10. Гусейнова А. С., Павловский Ю. Н., Устинов В. А. Опыт имитационного моделирования исторического процесса. М.: Наука, 1984.
11. Белотелов Н. В., Бродский Н. В., Павловский Ю. Н. Сложность. Математическое моделирование. Гуманитарный анализ. Исследование исторических, военных, социально-экономических и политических процессов. — Либроком. 2009.

Список литературы

Литература к Введению

1. *Petty W.* Political Arithmetic. 1690.
2. *de Condorcet N.* Essai sur l'application de l'analyse a la probabilitй des dйcisions rendues a la pluralitй des voix. 1785.
3. An Essay On The Principle Of Population. 1798 (автор — Т. Мальтус; издание вышло анонимно).
4. *Менделеев Д.И.* К познанию России. СПб, изд-во А.С. Суворина, 1906 (переиздание: М., Дрофа, 2002).
5. *Самарский А.А.* Современная прикладная математика и вычислительный эксперимент. Коммунист, 1983, № 18. С. 31–42.
6. *Самарский А.А.* Неизбежность новой методологии. Коммунист, 1989, № 1. С. 82–92.
7. *Самарский А.А., Михайлов А.П.* Компьютеры и жизнь. Педагогика, 1987.
8. *Самарский А.А., Михайлов А.П.* Математическое моделирование. Идеи. Методы. Примеры. М., Наука, ФИЗМАТЛИТ, 1997, 320 с. (англ. перевод: Samarskii A. A., Mikhailov A. P. Principles of Mathematical Modeling. Ideas, Methods, Examples London and New York. Taylor and Francis, 2002).
9. *Самарский А.А., Михайлов А.П.* Математическое моделирование в информационную эпоху // Вестник РАН. 2004. Т. 74. № 9.
10. *Четверушкин Б.Н., Михайлов А.П.* Триада Самарского. К 100-летию со дня рождения академика А. А. Самарского // Вестник Российской Академии Наук. 2019. Т. 89, № 2. С. 187–193.
11. *Левин К.* Закон и эксперимент в психологии // К. Левин. Динамическая психология. Сб. статей. С. 23–53.
12. *Сноу Ч.П.* Две культуры. М.: Прогресс. 1973.

Литература к Главе 1

1.1

1. *Михайлов А.П.* Улыбка, мысль и стрела времени, или мой С.П. Курдюмов // Мне нужно быть: Памяти Сергея Павловича Курдюмова. М. УРСС. 2010. С. 338–396.
2. *Glenn N.D.* The Course of Marital Success and Failure in Five American 10-Year Marriage Cohorts // Journal of Marriage and the Family. 60 p. 569–76.
3. *VanLaningham J., Johnson D.R., Amato P.* Marital Happiness, Marital Duration, and the U-Shaped Curve: Evidence from a Five-Wave Panel Study // Social Forces, June 2001, 78(4). P. 1313–1341.

4. *Дуггерти К.* Введение в эконометрику. М.: ИНФРА-М, 2004. 419 с.
5. *Wagenaar W.* My Memory: A Study of Autobiographical Memory over Six Years // *Cognitive Psychology* 18. P. 225–252. 1986.
6. *Hunt E.* *The Mathematics of Behavior.* Cambridge University Press, 2007.

1.2

1. *Gor S.* Inequality and Child Survival in Kenya: A Probit Model Approach // IARIW 31st General Conference, St-Gallen, Switzerland, August 22-28, 2010. <http://www.iariw.org/papers/2010/poster2Gor.pdf>.
2. *Micklewright J.* The Analysis of Pooled Cross-Sectional Data // *Analyzing social and political change. A Casebook of Methods* edited by Angela Dale and Richard B. Davies. p. 92.

1.3

1. *Ким Дж.-О., Мьюллер Ч.У., Клекка У.Р.* Факторный, дискриминантный и кластерный анализ. 1989. М.: Финансы и статистика.
2. *Мандель И.Д.* Кластерный анализ. 1988. М.: Финансы и статистика.
3. *Дюран Б., Оделл П.* Кластерный анализ. М.: «Статистика», 1977. 128 с.
4. *Fonseca J.R.S.* Clustering in the field of social sciences: that is your choice // *International Journal of Social Research Methodology.* Volume 16, Issue 5, 2013. P. 403–428.
5. *Kornilina E., Petrov A.* Research of political blogs by means of LSA // *Intellectualization of information processing-2010 conf. proceedings* — 2010 P. 508–511.
6. *Петров А.П., Корнилина Е.Д.* Исследование близости политических позиций методом латентно-семантического анализа // XII Международная научная конференция по проблемам развития экономики и общества. Книга 2. М.: Издательский дом высшей школы экономики. 2012. С. 334–342.
7. *Корнилина Е.Д., Петров А.П.* Латентно-семантический анализ предвыборных партийных программ на выборах в Государственную Думу 2007 и 2011 годов // М.: Вестник МГУ. Сер. 12: Политические науки, Издательство МГУ. 2013. № 2. С. 90–98.
8. *Contreras P., Murtagh F.* Methods of hierarchical clustering, arXiv:1105.0121, 2011.
9. *Blashfield R.K., Aldenderfer M.S.* The literature on cluster analysis *Multivariate Behavioral Research* 1978, 13: 271–295.
10. *Lance G.N, Williams W.T.* A general theory of classificatory sorting strategies. 1. Hierarchical systems. *Comput. J.*, 1967. V. 9, № 4, p. 373–380.
11. *Lance G.N, Williams W.T.* A general theory of classificatory sorting strategies. II. Clustering systems. *Comput. J.*, 1967. V. 10, № 4, p. 271–276.
12. *Jakulin A., Buntine W., La Pira T.M., Brasher H.* Analyzing the U.S. Senate in 2003: Similarities, Clusters, and Blocs // *Political Analysis* (2009) 17(3): 291–310.

13. *Fleishman J.A.* Types of Political Attitude Structure: Results of a Cluster Analysis // *The Public Opinion Quarterly*. V. 50, № 3 (Autumn, 1986), pp. 371–386.
14. *Лобазова М.В.* Результаты применения кластерного анализа к исследованию этнополитического конфликта в Республике Дагестан и на Северном Кавказе // *Математическое моделирование социальных процессов*. Вып. 3. Под ред. А. П. Михайлова. М. Изд-во МГУ. 2001. С. 171–179.

1.4

1. *Arrow K.J.* Social Choice and Individual Values, Yale University Press, 1951.
2. *Condorcet N.* Essai sur l'application de l'analyse a la probabilité des décisions rendues à la pluralité des voix. 1785.
3. *Finocchiaro C.J., Jenkins J.A.* In Search of Killer Amendments In the Modern U.S. House // *Legislative Studies Quarterly*, 2008. 33: 263–294. doi: 10.3162/036298008784310984.
4. *Алескеров Ф.Т., Хабина Э.Л., Шварц Д.А.* Бинарные отношения, графы и коллективные решения. Издательский дом ГУ ВШЭ. М. 2006. 300 с.

1.5

1. *Banzhaf J.F.* Weighted voting doesn't work: A mathematical analysis // *Rutgers Law Review*. 1965. 19(2). P. 317–343.
2. *Penrose L.* The Elementary Statistics of Majority Voting // *Journal of the Royal Statistical Society* (Blackwell Publishing). 1946. 109(1). P. 53–57.
3. *Shapley L.S., Shubik M.* A method for Evaluating the Distribution of Power in a Committee System // *American Political Science Review*. 1954. № 48(3). P. 787–792.
4. *Aleskerov F.* Power indices taking into account agents' preferences. *Mathematics and Democracy*. Berlin, Springer, 2006. P. 1–18.
5. *Алескеров Ф.Т.* Индексы влияния, учитывающие предпочтения участников по созданию коалиций // *ДАН*. 2007. Т. 414. № 5. С. 594–597.
6. *Felsenthal D.S., Machover M.* The Measurement of Voting Power: Theory and Practice, Problems and Paradoxes. Cheltenham: Edward Elgar, 1998. 322 p.
7. *Brams S.J.* Game Theory and Politics. New York: Free Press, 1975. 312 p.
8. *Kilgour D.M.* A formal analysis of the amending formula of Canada's Constitution Act // *Canadian Journal of political science*. 1983. № 16. P. 771–777.
9. *Алескеров Ф.Т., Благовещенский Н.Ю., Сатаров Г.А., Соколова А.В., Якуба В.А.* Оценка влияния групп и фракций в российском парламенте (1994–2003 гг.). Препринт ГУ Высшая Школа Экономики, WP7/2003/01, Москва, 2003.
10. *Алескеров Ф.Т., Хабина Э.Л., Шварц Д.А.* Бинарные отношения, графы и коллективные решения. Издательский дом ГУ ВШЭ. М. 2006. 300 с.
11. *Юрескул Е.А.* Исследование механизма принятия внешнеполитических решений в национальных и наднациональных представительных органах количественными методами (на примере Верховной Рады Украины VI созыва, Бундестага ФРГ 16 законодательного периода, Шестого Европейского Парламента). Факультет глобальных процессов МГУ им. М.В. Ломоносова. Дипломная работа. Москва, 2010.

Литература к Главе 2

2.1

1. An Essay On The Principle Of Population. 1798 (издание вышло анонимно, автором является Т. Мальтус).
2. Мальтус Т.-Р. Опыт закона о народонаселении. Пер. И. А. Вернера. М., 1895.
3. Spencer H. Principles of Biology (1864, 1867; revised and enlarged: 1898), in two volumes.
4. Маркс К. Капитал, т. I. — Маркс К., Энгельс Ф. Соч., т. 23, с. 645–646.
5. Менделеев Д.И. К познанию России. СПб, изд-во А.С. Суворина, 1906 (переиздание: М., Дрофа, 2002).
6. Милль Дж.-Ст. Основы политической экономии. Книга I. Пер. Н.Г. Чернышевского. 1848.
7. Ленин В.И. Аграрный вопрос и «критики Маркса». — Полн. собр. соч., изд. 5., т. 5.
8. Чернышевский Н.Г. Замечания на последние четыре главы первой книги Милля (опубликованы Чернышевским в составе издания [7]).
9. Тимирязев К.А. Исторический метод в биологии. 1922 (стр. 107 по переизданию 1942 года).
10. Verhulst P.F. Notice sur la loi que la population poursuit dans son accroissement. Correspondance mathématique et physique. 1848. 10:113–121.
11. Ризниченко Г.Ю., Рубин А.Б. Математические модели биологических продукционных процессов. М.: изд-во МГУ, 1993, 302 с.

2.2

1. Sharpe F.R., Lotka A.J. A Problem of Age-Distribution // Philos. Mag. 1911. V. 21. P. 435–438.
2. Lotka A.J. A contribution to the theory of self-renewing aggregates, with special reference to industrial replacement // Ann. Math. Stat. 1939. V. 10. P. 1–25.
3. Lotka A.J. The Stability of the placeNormal Age Distribution // Proc Natl Acad Sci USA. 1922 November; 8(11): 339–345.

2.3

1. Leslie P.H. The use of matrices in certain population mathematics. Biometrika, 1945. 33(3), 183–212.
2. Leslie P.H. «Some further notes on the use of matrices in population mathematics». Biometrika, 1948. 35(3–4), 213–245.
3. Caswell H. Matrix Population Models: Construction, analysis, and interpretation. 2nd Edition, 2001.
4. Kot M. Elements of Mathematical Ecology. Cambridge University Press, 2001.
5. Медков В.М. Демография: Учебное пособие. Серия «Учебники и учебные пособия». Ростов-на-Дону: «Феникс», 2002. 448 с.
6. Yadava R.C., Kumar A., Srivastava U. Sex ratio at birth: A model based approach // Mathematical Social Sciences, 2013. 65, p. 36–39.

7. Борисов В. А. Демография: Учебник для вузов. С. 242–244.
8. Яценко Н. Е. Толковый словарь обществоведческих терминов. СПб.: Изд-во «Лань», 1999. 528 с.
9. Адамчук В. В., Ромашов О. В., Сорокина М. Е. Экономика и социология труда: Учебник для вузов. М.: ЮНИТИ, 1999. 407 с.
10. Логофет Д. О., Клочкова И. Н. Математика модели Лефковича: репродуктивный потенциал и асимптотические циклы // Матем. моделирование, 2002, том 14: 10, 116–126.
11. Алексеева Т. А., Денисова Р. Я., Козловская М. В., Костылева Е. Л., Крайнов Д. А., Лебединская Г. В., Уткин А. В., Федосова В. Н. Неолит лесной полосы Восточной Европы (Антропология Сахтышских стоянок). М., «Научный мир», 1997. 191 с.

2.4

1. Samuelson P. A.. Generalizing Fisher's "reproductive value": Nonlinear, homogeneous, biparental systems // Proc. Nat. Acad. Sci. USA. V. 74, № 12. P. 5772–5775, December 1977.
2. Yellin J., Samuelson P. A. A Dynamical Model for Human Population // Proc. Nat. Acad. Sci. USA. Vol. 71, No. 7. P. 2813–2817, July 1974.
3. Yadava R. C., Kumar A., Srivastava U. Sex ratio at birth: A model based approach // Mathematical Social Sciences, 2013. 65. P. 36–39.
4. Ianelli M., Martcheva M., Milner F. A. Gender-Structured Population Modeling / Mathematical Methods, Numerics, and Simulation. country-SIAM. 2005.

2.5

1. Карпова В. М. Построение и исследование динамической модели рождаемости // Математическое моделирование социальных процессов Выпуск 6, М.: МАКС Пресс, 2004. под ред. А. П. Михайлова.
2. Карпова В. М. Построение и исследование динамической модели рождаемости // Тезисы докладов и выступлений на II Всероссийском социологическом конгрессе «Российское общество и социология в XXI веке: социальные вызовы и альтернативы»: В 3 т. М.: Альфа-М, 2003. Т. 1.
3. Карпова В. М. Миграция или рождаемость — имитационное моделирование как способ выбора ресурса // Социология. № 1–2, 2009.
4. Карпова В. М. Сценарии развития процессов старения населения в зависимости от демографической политики // Социология 4М № 29, 2009.
5. Карпова В. М. Некоторые предложения по изучению путей преодоления демографического кризиса // Демографические исследования № 7 (www.demographia.ru) № 7.
6. Карпова В. М. Проблема депопуляции: разные взгляды на одну катастрофу // Мир психологии. № 3, 2007.
7. Карпова В. М. Имитационная модель рождаемости и возможности ее применения для прогнозирования численности населения России // Мониторинг демографической ситуации в Российской Федерации и тенденций ее изменения. Ежегодный доклад «Демографическая ситуация в РФ», М.: КДУ, 2008.

8. Имитационное моделирование в демографии под. ред. А.Г. Волкова, М.: Статистика, 1980. С. 14–76.
9. Баркалов Н.Б. Микроимитационная модель рождаемости поколения // Городская и сельская семья под. ред. Медкова В.М. М.: 1987. С. 170–196.
10. Елизаров В.В. Перспективы исследования семьи. М.: Мысль, 1987. С. 127–141.
11. Волков А.Г., Сороко Е.Л. Имитационная модель демографического развития семьи // Демографические процессы и их закономерности. М., 1986.
12. Медков В.М. Микроимитационные модели в демографии: проблемы и возможности использования в прогнозировании динамики населения и демографических процессов // Математическое моделирование социальных процессов. М.: МАКС Пресс, 2002; под ред. А.П. Михайлова.
13. Антонов А.И., Медков В.М., Нетрусов М.А. Демография семьи — проблемы моделирования семейного цикла жизни // Вестник московского университета, сер. 18, Социология и политология. М., 1997. №2. С. 97–108.
14. Гражданников Е.Д. Прогностические модели социально-демографических процессов. Новосибирск. 1974. С. 65-71.
15. Антонов А.И. Перспективы рождаемости в связи с тенденциями установок детности // Демографические процессы в России XXI века. М., 2002. 59 с.
16. Антонов А.И., Медков В.М. Архангельский ВН Демографические процессы в России XXI века. М., Грааль, 2002, 168 с.
17. Борисов В.А. Желаемое число детей в российских семьях по данным микропереписи населения России 1994 года // Вестник Московского университета. 1997, Серия 18, Социология и политология, 2. 36 с.
18. Медков В.М. Репродуктивное поведение и социальная мобильность: ценностно-ориентационный аспект // Демографические процессы в России в XXI в. М., 2002. 82 с.
19. Гольцова Е.В. Экологические факторы демографического поведения семьи // Социология, 3. 2008. С. 30–33.
20. Шабунова А.А., Калачикова О.Н. Репродуктивный выбор семьи: факторы и условия // Социология, 33, 2008. С. 72–79.
21. Варламова С.Н., Носкова А.В., Седова Н.Н. Семья и дети в жизненных установках россиян // Социс. 2006, 11.
22. Skirbekk V. Fertility trends by social status // Demographic Research. Vol. 18, pp. 45–180.
23. Сайт Левада-Центра <http://www.levada.ru/press/2010081900.html>.

Литература к Главе 3

3.1

1. Thurstone L.L. Psychophysical Analysis // Source: The American Journal of Psychology. Jul., 1927. Vol. 38, № 3. P. 368–389.
2. Thurstone L.L. A Law of Comparative Judgment // Psychological Review. 1927. 34. P. 273–286.

3. *Mosteller F.* Remarks on the method of paired comparisons: I. The least-squares solution assuming equal standard deviations and equal correlations // *Psychometrika* 1951. 16. P. 3–9.
4. *Mosteller F.* Remarks on the method of paired comparisons: II. The effect of an aberrant when equal standard deviations and equal correlations are assumed // *Psychometrika*. 1951. 16. P. 203–206.
5. *Mosteller F.* Remarks on the method of paired comparisons: III. A test of significance for paired comparisons when equal standard deviations and equal correlations are assumed // *Psychometrika*. 1951. 16. P. 207–218.
6. *Torgerson N.S.* Theory and Method of scaling. N.Y.: John Wiley and Sons. 1958.
7. *Гусев А.Н., Измайлов Ч.А., Михалевская М.Б.* Измерение в психологии. М. Смысл. 1987.
8. *Келли Дж.* Теория личности. Психология личных конструктов. СПб., Речь, 2000.
9. *Poulton E.C., Simmonds D.C.V.* Subjective zeros, subjectively equal stimulus spacing, and contraction biases in very first judgments of lightness // *Perception & Psychophysics*. 1985. 37. P. 420–428.
10. *Adams-Webber J., Benjafield J.* The Relation between Lexical Marking and Rating Extremity in Interpersonal Judgments // *Canadian Journal of Behavioral Sciences*. 1973. 5. P. 234–241.
11. *Петров В.М.* Тернарность в мышлении, культуре, искусстве: системно-информационные корни бессознательного // Психология. Журнал Высшей школы экономики. 2008. Т. 5. № 4. С. 3–18.

3.2

1. *Rashevsky N.* Outline of a Physico-mathematical Theory of Excitation and Inhibition // *Protoplasma*, 1933.
2. *Рашевский Н.* Две модели: подражательное поведение и распределение статуса // Математические методы в современной буржуазной социологии. Сборник статей. Под ред. Г.В. Осипова. М.: Прогресс. 1966. С. 175–197.
3. *Cull P.* The mathematical biophysics of Nicolas Rashevsky // *BioSystems* 2007. Vol. 88. P. 178–184.
4. *Шведовский В.А., Шведовский О.В., Цаплин Н.В.* Математическая модель микродинамики личностных изменений // Математическое моделирование социальных процессов / Под ред. А.П. Михайлова, МАКС Пресс Москва, том 11, с. 98–131.
5. *Петров А.П., Цаплин Н.А.* Вычислительные эксперименты с моделью подражательного поведения // Математическое моделирование социальных процессов. Вып. 14. Под ред. А.П. Михайлова. 2012. С. 81–91.
6. *Эрроусмит Д., Плейс К.* Обыкновенные дифференциальные уравнения. Качественная теория с приложениями. Волгоград, изд-во Платон. 1997.

3.3

1. *Шведовский В.А.* Динамическая модель электорального поведения // Математическое моделирование. 2000. Т. 12. С. 46–56.

2. Шведовский В.А. Динамическая модель этнополитического конфликта: построение, возможности и результаты применения. // Сб. Математическое моделирование социальных процессов. М.: МГУ. 2000. — Вып. 2. С. 31–37.
3. Михайлов А.П., Южно Л.Ф. Исследование некоторых математических моделей эволюции электората // Дифференциальные уравнения. 2002. Т. 38. С. 1322.
4. Михайлов А.П., Шведовский В.А., Маслов А.И., Ковалев В.Ф. Обобщенная модель электорального поведения и ее применение к изучению этнополитических конфликтов // Математическое моделирование. 2003. Т. 15. С. 39–56.
5. Михайлов А.П. Моделирование системы власть–общество. М.: ФИЗМАТЛИТ, 2006. 144 с.

3.4

1. DeGroot M.H. Reaching a Consensus // Journal of the American Statistical Association, Vol. 69. Issue 345 (Mar 1974). P. 118–121.
2. Friedkin N. A Formal Theory of Social Power // J. of Mathematical Sociology. 1986. Vol. 12(2). P. 103–126.
3. Ефимов Б.А. Математическое моделирование формирования установок участников социальной группы // Математическое моделирование социальных процессов. Выпуск 10. 2008. Москва. КДУ. С. 67–122.
4. Гаврилец Ю.Н., Ефимов Б.А. Изменение предпочтений индивидов в социальной среде // Экономика и матем. методы, 1997, т. 33, №3.
5. Юнг К.Г. Психологические типы. М.: Прогресс, 1995.

3.5

1. Simon H.A. A Formal Theory of Interaction in Social Groups // American Sociological Review. V. 17, №2, Apr., 1952. P. 202–211.
2. Homans G.C. The Human Group. N.Y., 1950
3. Хоманс Дж. Социальное поведение как обмен // Современная зарубежная социальная психология. М.: Издательство Московского университета, 1984. С. 82–91.

3.6

1. Heider F. Attitudes and Cognitive Organization. // J. of Psych., 21, 1946, p. 107–112.
2. Heider F. Theory of Interpersonal Relations. N.Y., 1958
3. Cartwright D. and Harary F. Structural Balance: A Generalization of Heider's Theory. Psych. Rev., 63, 1956. P. 277–293.
4. Davis J.A. Clustering and structural balance in graphs // Human Relations. 1967. V. 20. P. 181–188.
5. Kemeny J.G., Snell J.L. Mathematical Models in the Social Sciences. — New York: Blaisdell Publishing Co., 1962; reprinted by M.I.T. Press, Cambridge, Mass, 1972.

3.7

1. Killworth P.D., Bernard H.R. A Model Of Human Group Dynamics // Social Science Research. 1976. 5. P. 173–224.

3.8

1. Корнилина Е.Д., Петров А.П. Динамическая модель близости позиций пользователей социальных сетей // Математическое моделирование. 2012. Т. 24. № 10. С. 89–97.
2. Gorbatikov E., Kornilina E., Mikhailov A., Petrov A. Mathematical Model of Opinion Dynamics in Social Groups // Mediterranean Journal of Social Sciences. V. 4. № 10. October 2013. P. 380–387. Doi:10.5901/mjss.2013.v4n10p380.
3. Корнилина Е.Д., Петров А.П. Модель динамики близости политических позиций // Теория активных систем — 2011 (ТАС-2011). Труды международной научно-практической конференции. ИПУ РАН, Москва, 14–16 ноября 2011 г. Том 2. С. 270–273.
4. Downes A. An economic Theory of Democracy. N.Y., Harper&Row, 1957.
5. Ахременко А.С. Структуры электорального пространства. М.: Социально-политическая мысль, 2007. 320 с.
6. Volkens A. Strengths and weaknesses of approaches to measuring policy positions of parties // Electoral Studies. 2007. Т. 26. № 1. С. 108–120.
7. Корнилина Е.Д. Применение латентно-семантического анализа к определению позиции политических блогов // Социология. 2011 г. № 1 С. 162–164.
8. Петров А.П., Корнилина Е.Д. Исследование близости политических позиций методом латентно-семантического анализа // XII Международная научная конференция по проблемам развития экономики и общества. Книга 2. М.: Издательский дом высшей школы экономики. 2012. С. 334–342.
9. Kornilina E., Petrov A. Research of political blogs by means of LSA // Intellectualization of information processing-2010 conf. proceedings 2010. P. 508–511.
10. Корнилина Е.Д., Петров А.П. Латентно-семантический анализ предвыборных партийных программ на выборах в Государственную Думу 2007 и 2011 гг. // М.: Вестник МГУ. Сер. 12: Политические науки, Издательство МГУ. 2013. № 2. С. 90–98.
11. Landauer T., Foltz P.W., Laham D. Introduction to Latent Semantic Analysis. Discourse Processes 1998. V. 25. P. 259–284.

Литература к Главе 4

4.1

1. Lazarsfeld P.F. The Logical and Mathematical Foundations of Latent Structure Analysis // Chapter 10 in Stouffer S.A., et al. Measurement and Prediction, Volume IV of The American Soldier: Studies in Social Psychology in World War II. Princeton University Press. 1950.
2. Лазарсфельд П. Логические и математические основания латентно-структурного анализа // Математические методы в современной буржуазной социологии. Под ред. Г.В. Осипова. — М. Прогресс. 1966. С. 344–401 (перевод предыдущей работы в данном списке).
3. Lazarsfeld P.F. Some Latent Structures // Chapter 11 in Stouffer S.A., et al. Measurement and Prediction, Volume IV of The American Soldier: Studies in Social Psychology in World War II. Princeton University Press. 1950.

4. *Henry N.W.* Latent Structure Analysis at Fifty. Paper presented at the 1999 Joint Statistical Meetings, Baltimore MD, August 11, 1999. URL <http://www.people.vcu.edu/~nhenry/LSA50.htm>.
5. *Осипов Г.В., Андреев Э.П.* Методы измерения в социологии. Андреев, Осипов. Москва: Наука, 1977.
6. *Толстова Ю.Н.* Измерение в социологии. М.: Инфра-М, 2003.

4.2

1. *Ефремов В.А., Леонтьев В.А.* Использование аппарата Марковских процессов в социологических исследованиях межконфессиональных отношений в России // Математическое моделирование социальных процессов. Вып. 8. М.: 2006. С. 15–29.
2. *Seneta E.* Non-negative Matrices and Markov Chains. Springer, 2006.
3. *Kemeny J., Snell J.* Finite Markov Chains. New York: Springer-Verlag, 1976.
4. *Prais S.J.* Measuring social mobility. Journal of the Royal Statistical Society. Series A (General), 118(1). P. 56–66. 1955.
5. *Prais S.J.* The formal theory of social mobility. Population Studies, 9(1). P. 72–81. 1955.
6. *Тихонова Н.Е.* Модель социальной стратификации российского общества: эвристические возможности различных теоретических подходов // Россия реформирующаяся. Ежегодник / Отв. Ред. М.К. Горшков. Вып. 6. М.: Институт социологии РАН, 2007. С. 112–146.
7. *Тихонова Н.Е.* Низший класс в социальной структуре российского общества с. 24 // Социологические исследования. 2011. №5.
8. *Бурдые П.* Формы капитала // Экономическая социология. Электронный журнал. Т.3 №5. ноябрь 2002. С. 60–74.
9. *Blumen I., Kogan M., McCarthy P.J.* The Industrial Mobility of Labor as a Probability Process. Cornell University Press. Ithaca. New York, 1955.).
10. *Beffy M., Coudin E., Rathelot R.* Who is confronted to insecure labor market histories? Some evidence based on the French labor market transitions. Institut National de la Statistique et des Études Économiques. Série des documents de travail de la Direction des Études et Synthèses Économiques. G2008/10.
11. *Montgomery J.D.* The Population Dynamics of Black-White-Mulatto Racial Systems // American Journal of Sociology. V. 117, № 1 (July 2011). P. 46–89.

4.3

1. European Commission against Racism and Intolerance General Policy Recommendation №7: On national legislation to combat racism and racial discrimination. Adopted by ECRI on 13 December 2002.
2. *Schelling T.C.* Models of Segregation // The American Economic review. V. 59, № 2, (May 1969). P. 488-493.
3. *Schelling T.C.* Dynamic Models of Segregation // Journal of Mathematical Sociology, 1971. V. 1. P. 143–186.

4.4

1. *Алешин Д.А.* Новые подходы ФАС России к анализу товарных рынков: применение теста гипотетического монополиста // Анализ товарных рынков в антимонопольном регулировании. Технологии и алгоритмы. Под ред. Д.А. Алешина. М.: ФАС России, Маркет ДС Корпорейшн, 2007. С. 5–19.
2. *Dobbs I.M.* Demand Cost Elasticities and Pricing Benchmarks in the Hypothetical Monopoly Test: The Consequences of a Simple SSNIP
3. *Dobbs I.M.* Defining Markets for ex ante Regulation Using the Hypothetical Monopoly Test // International Journal of the Economics of Business 2006, 13(1), 83–109. Applied Economics Letters 2003, 10, 545–548.
4. *Михайлов А.П., Петров А.П., Алешин Д.А.* Математические модели и алгоритмы определения границ рынка с помощью теста гипотетического монополиста // Анализ товарных рынков в антимонопольном регулировании. Технологии и алгоритмы. Под ред. Д. А. Алешина. М.: ФАС России, Маркет ДС Корпорейшн, 2007. С. 20–54.
5. *Михайлов А.П., Петров А.П., Алешин Д.А.* Антимонопольная деятельность и математическое моделирование // Информационные технологии и вычислительные системы, 2007, № 4. С. 12–21.

4.5

1. *Самарский А.А., Михайлов А.П.* Математическое моделирование. Идеи, методы, примеры. Москва, Наука, 2001 г., 320 с.
2. *Михайлов А.П., Ключосов Н.В.* О свойствах простейшей математической модели распространения информационной угрозы, Математическое моделирование социальных процессов, вып. 4. М.: МАКС Пресс, 2002. С. 115–123.
3. *Михайлов А.П., Петров А.П., Маревцева Н.А., Третьякова И.В.* Развитие модели распространения информации в социуме // Математическое моделирование. 2014. Т. 26, № 3. С. 65–74. (англ. перевод: *Mikhailov A.P., Petrov A.P., Marevtseva N.A., Tretiakova I.V.* Development of a Model of Information Dissemination in Society // Mathematical Models and Computer Simulations, 2014. Vol. 6, № 5, pp. 535–541, 2014.).
4. *Петров А.П., Прончева О.Г.* Исследование моделей информационного нападения и информационного противоборства в структурированном социуме // Математическое моделирование социальных процессов. Вып. 17. Под ред. А. П. Михайлова М.: Эконинформ, 2015. С. 136–149.
5. *Тихонов А.Н.* О зависимости решений дифференциальных уравнений от малого параметра. Матем. сб., 1948, 22(64).
6. *Тихонов А.Н.* О системах дифференциальных уравнений, содержащих параметры // Матем. сб., 1948, 27(69).
7. *Тихонов А.Н.* Системы дифференциальных уравнений, содержащих малые параметры при производных. Матем. сб., 1952, 31(73).
8. *Васильева А., Бутузов В.* Асимптотические методы в теории сингулярных возмущений. — Высшая школа Москва, 1990. 208 с.
9. *Herman E., Chomsky N.* A propaganda model. Media and Cultural Studies. 2006.

10. DiFonzo N., Bordia P. Rumor, gossip and urban legends. *Diogenes*. 2007. 54(1), 19–35.
11. Pronchev G.B., Muraviev V.I. Социальные сети как фактор перехода России к инновационному развитию // *Социология*. № 3. 2011. С. 36–56.
12. Daley D.J., Kendall D.G. Stochastic Rumors // *Journal of the Institute of Mathematics and its Applications*. 1964. V. 1. P. 42–55.
13. Maki D.P., Thompson M. *Mathematical Models and Applications*. Prentice-Hall. Englewood Cliffs. 1973.
14. Belen S. The behaviour of stochastic rumours. PhD Thesis, The University of Adelaide. 2008.
15. Belen S., Pearce C.E.M. Rumours with general initial conditions. *ANZIAM J.*, 2004. 4, 393–400.
16. Nekovee M., Moreno Y., Bianconi G., Marsili M. Theory of Rumor Spreading in Complex Social Networks. *Physica A*. 2007. 374. P. 457–470.
17. Губанов Д.А., Новиков Д.А., Чхартишвили А.Г. Социальные сети: модели информационного влияния, управления и противоборства. — М.: ФИЗМАТЛИТ, 2010, 228 с.
18. Шведовский В.А. Моделирование распространения информации в смежных социальных группах // *Математические методы в социологическом исследовании*, Наука, М., 1981. С. 207–214.
19. Osei G.K., Thompson J.W. The supersession of one rumour by another. *J. of Applied Probability*. 1977. 14(01). P. 127–134.
20. Kaligotla C., Yücesan E., Chick S.E. An agent based model of spread of competing rumors through online interactions on social media. In *Proceedings of the 2015 Winter Simulation Conference (3985-3996)*. IEEE Press. 2015.
21. Escalante R., Odehnal M. A deterministic mathematical model for the spread of two rumors. arXiv:1709.01726 [physics.soc-ph].
22. Yanagizawa-Drott D. Propaganda and Conflict: Evidence from the Rwandan Genocide. *The Quarterly Journal of Economics*. 2014. 129(4). P. 1947–1994. doi: 10.1093/qje/qju020.

4.6

1. Михайлов А.П., Маревцева Н.А. Модели информационной борьбы // *Математическое моделирование*. Т. 23. 2011. № 10. С. 19–32.
2. Михайлов А.П., Петров А.П., Маревцева Н.А., Третьякова И.В. Развитие модели распространения информации в социуме // *Математическое моделирование*. 2014. Т.26, №3. С. 65–74 (англ. перевод: *Mikhailov A.P., Petrov A.P., Marevtseva N.A., Tretiakova I.V. Development of a Model of Information Dissemination in Society // Mathematical Models and Computer Simulations*. 2014. V. 6, №5. P. 535–541.).
3. Петров А.П., Прончева О.Г. Исследование моделей информационного нападения и информационного противоборства в структурированном социуме // *Математическое моделирование социальных процессов*. Вып. 17. Под ред. А.П. Михайлова — М.: Эконинформ, 2015. С. 136–149.
4. Михайлов А.П., Петров А.П., Прончева О.Г., Маревцева Н.А. Модель информационного противоборства в социуме при периодическом дестабилизирующем воздействии // *Математическое моделирование*. 2017. Т.29,

- № 2. С. 23–32 (англ. перевод: *Mikhailov A. P., Petrov A. P., Proncheva O. G., Marevtseva N. A.* A Model of Information Warfare in a Society Under a Periodic Destabilizing Effect // *Mathematical Models and Computer Simulations*. 2017. V. 9, № 5. P. 580–586.
5. *Михайлов А. П., Петров А. П., Прончева О. Г.* Модель информационного противоборства в социуме с кусочно-постоянной функцией дестабилизирующего воздействия // *Математическое моделирование*. 2018. Т. 30, № 7. С. 47–60.

Литература к Главе 5

5.1

1. *Михайлов А. П., Южно Л. Ф.* Простейшая модель установления равновесия между двумя ветвями власти // *Матем. моделирование*. 2001. Т. 13, № 1. С. 65–76.
2. *Михайлов А. П., Южно Л. Ф., Маслов А. И.* Динамическая модель конкуренции между политическими силами // *Доклады Академии наук*. 2000. Т. 374, № 4. С. 469–474.

5.2

1. *Михайлов А. П., Шведовский В. А.* Математическое моделирование региональных конфликтов в контексте глобализации (на примере Северного Кавказа) // *Вестник Московского университета. Серия 18: Социология и политология*. 2005. Т. 18, № 1. С. 186–197.
2. *Михайлов А. П.* Моделирование системы «власть–общество». М. ФИЗМАТЛИТ, 2006. 144 с.
3. *Михайлов А. П., Шведовский В. А., Маслов А. И., Ковалев В. Ф.* Обобщенная модель электорального поведения и ее применение к изучению этнополитических конфликтов // *Математическое моделирование*. 2003. Т. 15. С. 39–56.
4. *Шведовский В. А.* Динамическая модель этнополитического конфликта: построение, возможности и результаты применения // *Сб. Математическое моделирование социальных процессов*. М.: МГУ. 2000. Вып. 2. С. 31–37.
5. *Шведовский В. А., Михайлова П. А.* Построение модели взаимодействия электоратов // *Математическое моделирование*. 2008. Т. 20, № 7. С. 107–118.
6. *Шведовский В. А., Михайлова А. П.* Динамика взаимодействующих электоратов в окрестности положения равновесия // *Математическое моделирование*. 2009. Т. 21, № 7. С. 43–54.

5.3

1. *Михайлов А. П.* Математическое моделирование власти в иерархических структурах // *Математическое моделирование*. 1994. Т. 6, № 6. С. 108–138.
2. *Mikhailov A. P.* Mathematical Modeling of Power Distribution in State Hierarchical Structures Interacting with Civil Society // *Proceedings of 14th IMACS World Congress, Atlanta, USA*. 1994. V. 2. P. 828–830.
3. *Михайлов А. П.* Моделирование системы «власть–общество». М.: Наука, ФИЗМАТЛИТ, 2006. 250 с.
4. *Михайлов А. П., Савельев А. В.* Обоснование моделей властных иерархий через их микроописание // *Математическое моделирование*. 2001. Т. 13., № 4. С. 19–34.

5. Михайлов А. П., Ланкин Д. Ф. О конструкциях властных иерархий // Математическое моделирование. 2009. Т. 21, № 8. С. 108–120.
6. Михайлов А. П., Горбатииков Е. А. Базовая модель дуумвирата в системе «власть–общество» // Математическое моделирование. 2012. Т. 24, № 1. С. 33–45.

5.4

1. Дмитриев М. Г., Жукова Г. С., Петров А. П. Асимптотический анализ модели «власть–общество» для случая двух устойчивых распределений власти // Математическое моделирование. 2004. Т. 16, № 5. С. 23–34.
2. Васильева А. Б. Об устойчивости контрастных структур // Математическое моделирование. 1991. Т. 3, № 4. С. 114–123.
3. Васильева А. Б., Никитин А. Г., Петров А. П. Асимптотический метод исследования контрастных структур и его приложение к теории гидромагнитного динамо // Математическое моделирование. 1995. Т. 7, № 2. С. 61–71.
4. Васильева А. Б., Бутузов В. Ф. Асимптотические методы в теории сингулярных возмущений. М.: Высшая школа. 1990. 208 с.
5. Бутузов В. Ф., Неделько И. В. О глобальной области влияния устойчивых решений с внутренними слоями // Математическое моделирование. 2001. Т. 192, № 5. С. 13–52.

5.5

1. Mikhailov A. P. Efficient Strategies of Corruption Suppression in State Power Hierarchies // Proceedings of 15th IMACS World Congress, Berlin, Germany. 1997. V. III. P. 727–733.
2. Михайлов А. П. Модель коррумпированных властных иерархий // Математическое моделирование. 1999. Т. 11, № 1. С. 3–19.
3. Михайлов А. П. Моделирование системы «власть–общество». М.: ФИЗМАТЛИТ, 2006. 144 с.
4. Rose-Ackerman S. The political economy of corruption // Corruption and the global economy. 1997. 31, 60.
5. Ackerman S. R. Corruption: A study in political economy. New York: Academic Press, 1978.
6. Tirole J. Hierarchies and bureaucracies: On the role of collusion in organizations // J. Econ. & Org., 1986. 2. P. 181.
7. Bac M. Corruption, supervision, and the structure of hierarchies // The Journal of Law, Economics, and Organization. 1996. V. 12, № 2. P. 277–298.
8. Mauro P. Corruption and growth // The quarterly journal of economics. 1995. V. 110, № 3. P. 681–712.
9. Bertrand M., Djankov S., Hanna R., Mullainathan S. Does corruption produce unsafe drivers? (№ w12274) // National Bureau of Economic Research. 2006.
10. Васин А. А., Картунова П. А., Уразов П. А. Модели организации государственных инспекций и борьбы с коррупцией // Математическое моделирование, 2010. Т. 22, № 4. С. 67–89.

11. *Зенюк Д.А., Малинецкий Г.Г., Фаллер Д.С.* Имитационная модель коррумпции в иерархических системах // Компьютерные исследования и моделирование, 2014. V. 6, № 2. P. 321–329.

5.6

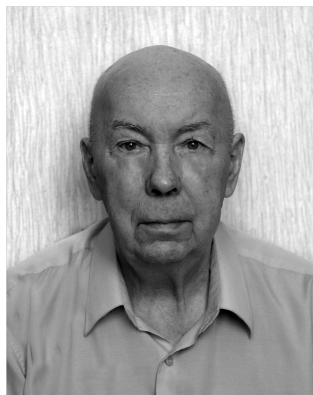
1. *Richardson L.F.* Arms and Insecurity. Pittsburgh. Boxwood Press. 1960.
2. *Anderton C.H.* Arms Race Modeling: Categorization and Systematic Analysis. Paper presented at the International Studies Association, Washington. 1985.
3. *Schrodt P.A.* Patterns, Rules and Learning: Computational Models of International Behavior. 2004 (2nd Edition). Parus Analytical Systems, Vinland, Kansas, USA.
4. *Lehmann B., McEwen J., Lane B.* Modifying the Richardson Arms Race Model With a Carrying Capacity.
5. *Schrodt P.A.* Richardson's N-nation model and the balance of power // American Journal of Political Science. 1978. V. 5 № 1. P. 364–390.

5.7

1. *Садовничий В.А., Акаев А.А., Коротаев А.В., Малков А.В.* Моделирование и прогнозирование мировой динамики // М.: ИСПИ РАН, 2012. 359 с.
2. *Малинецкий Г.Г.* Нелинейная динамика — ключ к теоретической истории? // Общественные науки и современность. 1996. Т. 4. P. 98–111.
3. *Forrester J.W.* World Dynamics. Cambridge, MA: Wright-Allen Press. 1971
4. *Meadows D.H., Meadows D.L., Randers J., Behrens W.W.* III. The Limits to Growth: A Report for the Club of Rome's Project on the Predicament of Mankind. New York, NY: Universe Books. 1972.
5. *Nordhaus W.D.* World Dynamics: Measurement Without Data // The Economic Journal. V. 83, № 332 (Dec., 1973). P. 1156–1183.
6. *Krugman P.* Limits to growth and related stuff. April 22, 2008. URL: <https://krugman.blogs.nytimes.com/2008/04/22/limits-to-growth-and-related-stuff>.
7. *Turner G.M.* A comparison of The Limits to Growth with 30 years of reality // Global environmental change. 2008. V. 18, No 3. P. 397–411.
8. *van Vuuren D.P., Faber A.* Growing within Limits — A Report to the Global Assembly 2009 of the Club of Rome. Netherlands Environmental Assessment Agency. 2009.
9. *Малков С.Ю.* Математическое моделирование исторической динамики: подходы и модели // *Дмитриев М.Г., Дубовский С.В., Жукова Г.С., Малков С.Ю., Михайлов А.П., Петров А.П., Степанцов М.Е.* Моделирование социально-политической и экономической динамики. М.: РГСУ, 2004. С. 76–188.
10. *Гусейнова А.С., Павловский Ю.Н., Устинов В.А.* Опыт имитационного моделирования исторического процесса. М.: Наука, 1984.
11. *Белотелов Н.В., Бродский Н.В., Павловский Ю.Н.* Сложность. Математическое моделирование. Гуманитарный анализ. Исследование исторических, военных, социально-экономических и политических процессов. Либликом. 2009.

О авторах

Михайлов Александр Петрович, доктор ф.-мат. наук, профессор, главный научный сотрудник ИПМ им. М. В. Келдыша РАН. Выпускник МФТИ, ученик С. П. Курдюмова и А. А. Самарского. Основные научные результаты относятся к теории самоорганизации в нелинейных средах и математическому моделированию в социологии и политологии. Более 25 лет занимается исследованиями системы «человек–общество», руководитель и исполнитель многочисленных проектов, поддержанных РФФИ, РГНФ, РНФ. Руководил лабораторией и кафедрой на социологическом факультете и специализированным семинаром на факультете вычислительной математики и кибернетики МГУ им. М. В. Ломоносова. Член редколлегий журналов «Математическое моделирование», «Социология 4М», «Computational Mathematics and Informational Technologies», главный редактор ежегодника «Математическое моделирование социальных процессов», рецензент РФФИ. Автор ряда оригинальных моделей, в частности модели системы «власть–общество». Ряд математических моделей реализован в виде программных комплексов и внедрен в работу органов государственной власти. Текущее значение индекса Хирша: 36. Более подробную информацию см. в энциклопедическом издании «Социология в лицах. История и современность».



Петров Александр Пхоун Чжо, доктор ф.-мат. наук, ведущий научный сотрудник ИПМ им. М. В. Келдыша РАН. Выпускник кафедры математики физического факультета МГУ им. М. В. Ломоносова, ученик А. Б. Васильевой. Основные научные результаты относятся к теории сингулярных возмущений, математическому моделированию в астрофизике и в социальных науках. Более 15 лет занимается исследованиями системы «человек–общество», руководитель и исполнитель многочисленных проектов, поддержанных РФФИ, РГНФ, РНФ. Читал лекции по математическому моделированию социально-экономических процессов в Российском университете дружбы народов, Российском государственном социальном университете, на социологическом факультете и факультете глобальных процессов МГУ. Ученый секретарь журнала «Computational Mathematics and Informational Technologies», рецензент журнала «Математическое моделирование» и РФФИ. Результаты в области математического моделирования информационного противоборства и распространения информации в социуме получили широкое освещение в средствах массовой информации. Текущее значение индекса Хирша: 14.



Научное издание

МИХАЙЛОВ Александр Петрович
ПЕТРОВ Александр Пхоун Чжо

МАТЕМАТИЧЕСКИЕ МОДЕЛИ СИСТЕМЫ
«ЧЕЛОВЕК–ОБЩЕСТВО»

Редактор *Е.И. Ворошилова*
Оригинал-макет: *Д.П. Вакуленко*
Оформление переплета: *А.В. Андросов*

Подписано в печать 24.02.2021. Формат 60×90/16. Бумага офсетная.
Печать офсетная. Усл. печ. л. 28,5. Уч.-изд. л. 31,35. Тираж 300 экз.
Заказ №

Издательская фирма «Физико-математическая литература»
МАИК «Наука/Интерпериодика»
117342, г. Москва, ул. Бултерова, д. 17Б
E-mail: porsova@fml.ru, sale@fml.ru
Сайт: <http://www.fml.ru>
Интернет-магазин: <http://www.fmlib.ru>

Отпечатано с электронных носителей издательства
в АО «Первая Образцовая типография»
Филиал «Чеховский Печатный Двор»
142300, Московская область, г. Чехов, ул. Полиграфистов, д. 1
Сайт: www.chpd.ru. E-mail: sales@chpd.ru, тел.: 8 (499) 270-73-59

