

**Л. А. Петросян
Н. А. Зенкевич
Е. В. Шевкопляс**

Теория игр

2-е издание

Рекомендовано УМО в области инновационных междисциплинарных образовательных программ в качестве учебника по направлению 010500 «Математическое обеспечение и администрирование информационных систем»

Санкт-Петербург
«БХВ-Петербург»
2012

УДК 512.8(075.8)
ББК 22.14+22.19я73
ПЗ0

Петросян, Л. А.

ПЗ0 Теория игр: учебник / Л. А. Петросян, Н. А. Зенкевич,
Е. В. Шевкопляс. — 2-е изд., перераб. и доп. — СПб.: БХВ-Петербург,
2012 — 432 с.: ил. — (Учебная литература для вузов)

ISBN 978-5-9775-0484-3

Учебник предназначен как для первоначального, так и для углубленного изучения теории игр. Проведено систематическое исследование математических моделей принятия решений несколькими сторонами в условиях конфликта. Представлено последовательное изложение единой теории статических и динамических игр. Рассмотрены все основные классы игр: конечные и бесконечные антагонистические игры, бескоалиционные и кооперативные игры, многошаговые и дифференциальные игры. Для закрепления материала в каждой главе содержатся задачи и упражнения разной степени сложности.

Во втором издании расширены разделы, касающиеся статической теории кооперативных решений и динамических кооперативных игр, а также игр с неполной информацией. Уточнены и изменены доказательства отдельных утверждений. Применен новый единый подход к исследованию оптимального поведения игроков в позиционных и дифференциальных играх.

*Для студентов и аспирантов математических, экономических,
управленческих и технических направлений и специальностей*

УДК 512.8(075.8)
ББК 22.14+22.19я73

Группа подготовки издания:

Главный редактор	<i>Екатерина Кондукова</i>
Зам. главного редактора	<i>Евгений Рыбаков</i>
Зав. редакцией	<i>Григорий Добин</i>
Компьютерная верстка	<i>Екатерины Шевкопляс</i>
Корректор	<i>Наталья Першакова</i>
Дизайн серии	<i>Инны Тачиной</i>
Оформление обложки	<i>Елены Беляевой</i>
Фото	<i>Кирилла Сергеева</i>
Зав. производством	<i>Николай Тверских</i>

Рецензенты:

Н. Н. Петров, д-р физ.-мат. наук, проф., завкафедрой исследования операций СПбГУ;
А. Ю. Гарнаев, д-р физ.-мат. наук, проф. кафедры компьютерного моделирования
и микропроцессорных систем СПбГУ.

Подписано в печать 30.09.11.

Формат 70×100^{1/16}. Печать офсетная. Усл. печ. л. 34,83.

Тираж 1200 экз. Заказ №

"БХВ-Петербург", 190005, Санкт-Петербург, Измайловский пр., 29.

Санитарно-эпидемиологическое заключение на продукцию
№ 77.99.60.953.Д.005770.05.09 от 26.05.2009 г. выдано Федеральной службой
по надзору в сфере защиты прав потребителей и благополучия человека.

Отпечатано с готовых диапозитивов
в ГУП "Типография "Наука"
199034, Санкт-Петербург, 9 линия, 12

ISBN 978-5-9775-0484-3

© Петросян Л. А., Зенкевич Н. А., Шевкопляс Е. В., 2011
© Оформление, издательство "БХВ-Петербург", 2011

Оглавление

Предисловие	6
Введение	8
1 Матричные игры	12
§ 1.1. Определение антагонистической игры в нормальной форме	12
§ 1.2. Максиминные и минимаксные стратегии	16
§ 1.3. Ситуации равновесия	18
§ 1.4. Смешанное расширение игры	22
§ 1.5. Некоторые сведения из теории выпуклых множеств	25
§ 1.6. Существование решения в классе смешанных стратегий	28
§ 1.7. Свойства оптимальных стратегий и значения игры	30
§ 1.8. Доминирование стратегий	38
§ 1.9. Вполне смешанные и симметричные игры	43
§ 1.10. Итеративные методы решения матричных игр	48
§ 1.11. Упражнения и задачи	52
2 Бесконечные антагонистические игры	55
§ 2.1. Бесконечные игры	55
§ 2.2. Ситуация ε -равновесия	58
§ 2.3. Смешанные стратегии	63
§ 2.4. Игры с непрерывной функцией выигрыша	70
§ 2.5. Игры с выпуклой функцией выигрыша	76
§ 2.6. Одновременные игры преследования	85
§ 2.7. Один класс игр с разрывной функцией выигрыша	90
§ 2.8. Бесконечные игры поиска	93
§ 2.9. Покер	98
§ 2.10. Упражнения и задачи	116
3 Неантагонистические игры	119
§ 3.1. Определение бескоалиционной игры в нормальной форме	119
§ 3.2. Принципы оптимальности в бескоалиционных играх	123
§ 3.3. Смешанное расширение бескоалиционной игры	129
§ 3.4. Существование ситуации равновесия по Нэшу	133
§ 3.5. Существование ситуации равновесия в конечной игре n лиц	134
§ 3.6. Модификации концепции равновесия по Нэшу	137

§ 3.7. Свойства оптимальных решений	141
§ 3.8. Эволюционно устойчивые стратегии	145
§ 3.9. Равновесие в совместных смешанных стратегиях	149
§ 3.10. Задача о переговорах	152
§ 3.11. Игры в форме характеристической функции	159
§ 3.12. C -ядро и NM -решение	165
§ 3.13. Вектор Шепли	173
§ 3.14. Вектор Шепли и потенциал	179
§ 3.15. Упражнения и задачи	182
4 Многошаговые игры	187
§ 4.1. Определение динамической игры с полной информацией	187
§ 4.2. Равновесие по Нэшу	190
§ 4.3. Основные функциональные уравнения	194
§ 4.4. Иерархические игры	196
§ 4.5. Иерархические игры (кооперативный вариант)	198
§ 4.6. Многошаговые игры с неполной информацией	204
§ 4.7. Стратегия поведения	210
§ 4.8. Функциональные уравнения для одновременных многошаговых игр	216
§ 4.9. Построение единственного равновесия по Нэшу	223
§ 4.10. Структура множества абсолютных равновесий по Нэшу	227
§ 4.11. Индифферентное равновесие в позиционных играх	234
§ 4.12. Стратегии наказания и «народные теоремы»	237
§ 4.13. Кооперация в многошаговых играх	241
§ 4.14. Кооперативные стохастические игры	250
§ 4.15. Марковские игры	261
§ 4.16. Упражнения и задачи	277
5 Антагонистические дифференциальные игры	284
§ 5.1. Антагонистические дифференциальные игры	284
§ 5.2. Многошаговые игры с полной информацией	292
§ 5.3. Существование ситуаций ε -равновесия	296
§ 5.4. Дифференциальные игры преследования на быстродействие	301
§ 5.5. Существование оптимальной программной стратегии убегающего	307
§ 5.6. Основное уравнение	310
§ 5.7. Методы последовательных приближений	316
§ 5.8. Примеры решения дифференциальных игр преследования	320
§ 5.9. Игры преследования с задержкой информации у преследователя	323
§ 5.10. Упражнения и задачи	329
6 Неантагонистические дифференциальные игры	333
§ 6.1. Принцип динамического программирования	333
§ 6.2. Принцип максимума Понтрягина	338
§ 6.3. Равновесие по Нэшу в программных стратегиях	341
§ 6.4. Равновесие по Нэшу в позиционных стратегиях	345
§ 6.5. Конкурентная реклама с двумя участниками	347
§ 6.6. Игры с бесконечной продолжительностью	350
§ 6.7. Модель конкуренции с бесконечной продолжительностью	352

§ 6.8. Упражнения и задачи 354

**7 Кооперативные дифференциальные игры в форме
характеристической функции 356**

§ 7.1. Определение кооперативной игры 356

§ 7.2. Дележи 357

§ 7.3. Дележи в динамике 359

§ 7.4. Принцип динамической устойчивости 361

§ 7.5. Динамически устойчивые решения 362

§ 7.6. Процедура распределения дележа 363

§ 7.7. Управление загрязнением окружающей среды 365

§ 7.8. Упражнения и задачи 374

**8 Кооперативные дифференциальные игры двух лиц
с дисконтированием 377**

§ 8.1. Постановка задачи 377

§ 8.2. Кооперативные игры с бесконечной продолжительностью 391

§ 8.3. Игры с нетрансферабельными выигрышами 397

§ 8.4. Упражнения и задачи 409

Литература 410

Предметный указатель 422

Предисловие

Теория игр — это раздел математики, в котором исследуются математические модели принятия решений в условиях конфликта, т. е. в условиях столкновения сторон, каждая из которых стремится воздействовать на развитие конфликта в своих собственных интересах. Теорию математических моделей принятия оптимальных решений принято называть исследованием операций, поэтому теорию игр следует рассматривать как прикладную математическую теорию — составную часть исследования операций.

Несмотря на наличие богатой монографической и специальной литературы по теории игр, учебников, покрывающих этот раздел математики, сравнительно немного и в них рассматриваются в основном отдельные разделы теории игр. Настоящий учебник восполняет этот пробел и является существенным развитием книги Петросян Л. А., Зенкевич Н. А., Семина Е. А. «Теория игр», М.: Высшая школа, Книжный дом «Университет», 1998, в которой впервые в отечественной литературе было дано систематическое изложение единой теории статических и динамических игр. В новой книге изложение динамических игр распространено на случай кооперативных дифференциальных игр, существенно расширены разделы, касающиеся статической теории кооперативных решений и динамических кооперативных игр, а также игр с неполной информацией. Уточнены и изменены доказательства отдельных утверждений. Здесь впервые в отечественной учебной литературе используется единый подход к исследованию оптимального поведения игроков в позиционных играх, основанный на концепции динамической устойчивости (состоятельности во времени) решений неантагонистических игр, а также свойствах сильной динамической устойчивости, динамической совместимости, согласованности и других динамических характеристиках состоятельности различных оптимальных решений. На базе изложенного подхода приведено систематическое изложение бескоалиционных, коалиционных и кооперативных принципов оптимальности в различных классах позиционных игр.

В учебнике отражено большинство актуальных направлений современной теории игр. Он методически построен так, что понятие модели конфликта (игры) развивается от простой (матричные игры) до наиболее сложной (дифференциальные игры). Большинство университетских учебных программ предполагает чтение отдельных разделов или специальных курсов по теории игр. Данный учебник построен таким образом, чтобы каждая глава могла служить основой такого курса. Для предварительного ознакомления с теорией игр достаточно изучить материал гл. 1. Типовой курс по теории игр может быть построен на основе гл. 1, 3 и 4. В учебнике полно изложены теории антагонистических игр (гл. 1, 2, 4, 5), неантагонистических игр (гл. 3, 4, 6) и кооперативных игр (3, 4, 7, 8). В учебных дисциплинах «Системный анализ» и «Модели принятия решений» целесообразно использовать гл. 3 и 4. При построении курса лекций полезно также воспользоваться приведенным списком специальной литературы.

Во всех главах приводятся многочисленные примеры, иллюстрирующие основные положения теории. Некоторые из них представляют самостоятельный интерес. В конце каждой главы приведены упражнения для индивидуальной работы, расположенные в порядке изложения материала и возрастания сложности. В ряде случаев они существенно дополняют содержание главы. Систематическое решение этих упражнений является важной формой изучения теории игр. Для усвоения основных понятий и результатов, приведенных в учебнике, достаточно знания курса математики в объеме университетской программы. Наиболее сложными в математическом отношении являются главы 2 и 5, которые предназначены для студентов математических специальностей.

В списке рекомендованной литературы приведены основная (основные учебники и задачники), дополнительная (монографии и учебные пособия) и специальная (статьи, справочники, обзоры, сборники статей) литература. В список дополнительной литературы включены также статьи, которые цитируются в основном тексте. Вместе с тем библиография не претендует на полноту.

Учебник может быть использован как для первоначального, так и для углубленного изучения теории игр. Он предназначен для студентов и аспирантов, обучающихся по направлению «Прикладная математика и информатика» и специализирующихся в области исследования операций, системного анализа, методов оптимизации, математической кибернетики, математического моделирования, будет также полезен студентам и аспирантам экономических, управленческих и технических направлений, изучающим математические методы принятия решений в сложных системах управления. Книга заинтересует специалистов, развивающих теорию игр, исследование операций, теорию управления, математическую экономику, теорию менеджмента и их приложения.

Учебник написан на основе курсов «Теория игр и исследование операций», «Системный анализ», «Математические модели принятия решений в экономике и управлении», «Исследование систем управления», а также ряда специальных курсов по разделам и приложениям теории игр, прочитанных Л. А. Петросяном, Н. А. Зенкевичем и Е. В. Шевкопляс студентам старших курсов и аспирантам на факультете прикладной математики — процессов управления (ПМ-ПУ) и в Высшей школе менеджмента Санкт-Петербургского государственного университета.

Благодарности. Параграфы 7, 9 гл. 1, § 5, 10 гл. 3, § 4–6, 8 и 9 гл. 4, § 2–6, 8 гл. 5 написаны совместно с Е. А. Семиной, за что авторы выражают Елене Александровне искреннюю признательность.

Мы благодарим Е. М. Парилину, А. А. Седакова, С. Ю. Костюнина и В. А. Клеменова, а также студентов и аспирантов кафедры математической теории игр и статистических решений факультета прикладной математики — процессов управления Санкт-Петербургского государственного университета за помощь при подготовке рукописи.

Выражаем особую благодарность Н. Н. Петрову и Н. И. Наумовой за ценные замечания и предложения.

Авторы

Введение

В.1. Математическая теория игр является составной частью исследования операций. Она находит широкое применение в различных областях человеческой деятельности, таких, как экономика и менеджмент, промышленность и сельское хозяйство, военное дело и строительство, торговля и транспорт, связь и т. д. В настоящем учебнике изложены основные понятия и результаты теории игр.

В.2. Задачи исследования операций можно классифицировать по уровню информации о ситуации, которой располагает субъект, принимающий решение. Наиболее простыми уровнями информации о ситуации являются детерминированный (когда условия, в которых принимаются решения, известны полностью) и стохастический (когда известно множество возможных вариантов условий и их вероятностное распределение). В этих случаях задача сводится к нахождению экстремума функции (или ее математического ожидания) при заданных ограничениях. Методы решения таких задач изучаются в курсах математического программирования или методов оптимизации.

Наконец, третий уровень — неопределенный, когда известно множество возможных вариантов, но без какой-либо информации об их вероятностях. Такой уровень информации о ситуации является наиболее сложным. Эта сложность оказывается принципиальной, так как могут быть не ясны сами принципы оптимального поведения. Следуя определению Н.Н. Воробьева, теория игр — это теория математических моделей принятия решений в условиях неопределенности, когда принимающий решение субъект («игрок») располагает информацией лишь о множестве возможных ситуаций, в одной из которых он в действительности находится, о множестве решений («стратегий»), которые он может принять, и о количественной мере того «выигрыша», который он мог бы получить, выбрав в данной ситуации данную стратегию. Установление принципов оптимального поведения в условиях неопределенности, доказательство существования решений, удовлетворяющих этим принципам, указание алгоритмов нахождения решений и составляют содержание теории игр.

В.3. Неопределенность, с которой мы встречаемся в теории игр, может иметь различное происхождение. Однако, как правило, она является следствием сознательной деятельности другого лица (лиц), отстаивающего свои интересы. В связи с этим под теорией игр часто понимают теорию математических моделей принятия оптимальных решений в условиях конфликта. Таким образом, моделями теории игр можно в принципе содержательно описывать весьма разнообразные явления: экономические, правовые и классовые конфликты, взаимодействие человека с природой, биологическую борьбу за существование и т. д. Все такие модели в теории игр принято называть играми.

Математическое описание игры сводится к перечислению всех действующих в ней игроков, указанию для каждого игрока всех его стратегий, а также численного выигрыша

ша, который он получит после того, как игроки выберут свои стратегии. В результате игра становится формальным объектом, который поддается математическому анализу.

В.4. Игры можно классифицировать по различным признакам. Во-первых, бескоалиционные игры, в которых каждая коалиция (множество игроков, действующих совместно) состоит лишь из одного игрока. Так называемая кооперативная теория бескоалиционных игр допускает временные объединения игроков в коалиции в процессе игры с последующим разделением полученного выигрыша или принятия совместных решений. Во-вторых, коалиционные игры, в которых принимающие решения игроки согласно правилам игры объединены в фиксированные коалиции. Члены одной коалиции могут свободно обмениваться информацией и принимать полностью согласованные решения.

По выигрышу игры можно разделить на антагонистические и игры с ненулевой суммой.

По характеру получения информации — на игры в нормальной форме (игроки получают всю предназначенную им информацию до начала игры) и динамические игры (информация поступает игрокам в процессе развития игры).

По количеству стратегий — на конечные и бесконечные игры.

В.5. Учебник состоит из восьми глав. Первая глава содержит основные сведения из теории конечных антагонистических (матричных) игр. Здесь доказывается теорема существования ситуации равновесия в классе смешанных стратегий, выводятся свойства оптимальных смешанных стратегий, приведены методы решения матричных игр. Хотя антагонистический конфликт является очень специальным случаем конфликта, возникающего в конкретных сферах приложений, тем не менее, в первой главе приводятся многочисленные примеры задач поиска и преследования, которые моделируются матричными играми.

Во второй главе рассматриваются бесконечные антагонистические игры или игры с бесконечным числом стратегий у каждого из игроков. Здесь теоремы существования ситуаций равновесия справедливы далеко не во всех случаях. К важным условиям существования относятся свойства функции выигрыша. В главе приводится доказательство существования ситуации равновесия в смешанных стратегиях, когда функция выигрыша является непрерывной. Однако существует достаточно большое число классов игр, для которых это обстоятельство не имеет места, но равновесие, тем не менее, существует. К играм такого типа относятся дуэли и покер. Рассмотрение игр типа покера интересно еще и тем, что позволяет обосновать стратегию «блефа», часто встречающуюся на практике. В главе рассмотрены также приложения к задачам поиска и преследования.

Третья глава посвящена статическим неантагонистическим играм. В качестве принципа оптимальности в таких моделях обычно используется равновесие по Нэшу. Приводится доказательство существования равновесия по Нэшу в смешанных стратегиях в играх с конечным числом стратегий, исследуются свойства равновесий и приводятся некоторые модификации равновесия по Нэшу. Исследуются и другие принципы оптимальности, такие, как решения оптимальные по Парето и арбитражные схемы. В связи с серьезными приложениями в биологии и экономике, даются определение и свойства эволюционно-устойчивых стратегий. В последнее время возродился интерес к использованию коррелированных действий в неантагонистических играх. В этой связи в главу включен материал о равновесиях в совместных смешанных стратегиях. Вторая половина главы посвящена кооперативной теории игр. Здесь мы ограничиваемся изложением случая, когда выигрыши игроков трансферабельны и игра задается характеристиче-

ской функцией. Предполагается, что при кооперации игроки максимизируют свой суммарный выигрыш. Поэтому задача заключается в дележе полученного максимального выигрыша, который устраивал бы всех игроков. Исходя из такого понимания кооперации, в главе приводится классическое понятие характеристической функции и основные принципы оптимальности (C -ядро, NM -решение и вектор Шепли). Доказаны теоремы существования непустого ядра. Отдельно исследованы выпуклые и простые игры, доказано существование C -ядра. Теоретические результаты иллюстрируются примерами из социально-экономической сферы.

В четвертой главе исследуются позиционные многошаговые игры (игры в развернутой форме). Эта глава имеет особое значение, поскольку она служит базой для понимания следующих глав, в которых описываются дифференциальные игры. Наиболее изученным классом игр являются игры с полной информацией. Для них доказывается существование абсолютного равновесия по Нэшу, т. е. такого равновесия, сужение которого в каждой подыгре, является равновесием в этой подыгре. Однако, кроме абсолютных равновесий, существует достаточно представительный класс равновесий в стратегиях наказания. Приводится характеристика этого класса равновесий и приводятся теоремы, характеризующие этот класс равновесий. Особенностью равновесий по Нэшу является их неединственность и неэквивалентность в том смысле, что выигрыши игроков в разных ситуациях равновесия могут серьезно различаться. Поэтому нетривиальным вопросом является выбор конкретного равновесия по Нэшу. В главе предлагается в качестве представителя равновесий по Нэшу выбрать индифферентное равновесие, существование и единственность которого доказывается. Особое место занимают иерархические игры. В главе приводится решение иерархических игр с древовидной и ромбовидной структурой. Отдельно исследованы кооперативные позиционные игры. Здесь возникает новая проблема — построение динамически устойчивых (состоятельных во времени) принципов оптимальности. Эта проблема решается с помощью введения процедур распределения дележа и основанной на них регуляризации игры. Отдельный параграф посвящен построению принципов оптимальности в играх с переменным коалиционным разбиением.

В пятой главе рассматриваются антагонистические дифференциальные игры. Изложение материала ведется на примере дифференциальных игр преследования. Однако полученные результаты могут быть легко использованы и в более общем случае. Доказывается основополагающая теорема о существовании ситуации ε -равновесия в классе кусочно-программных стратегий, выводится уравнение Айзекса–Беллмана для функции значения игры, приводятся итеративные методы решения этого уравнения. Результаты теории иллюстрируются на модельных примерах простого преследования и преследования при наличии сил трения. В указанных случаях находится решение уравнения Айзекса–Беллмана в явной форме. Исследуется задача преследования на быстроедействие, и приводятся теоремы, указывающие на связь между задачами преследования на быстроедействие и с предписанной продолжительностью. Отдельный параграф посвящен задаче преследования с задержкой информации у преследователя специального вида. Обоснован вид оптимальной стратегии убегающего, которая в этом случае включает в себя случайный выбор управляющего воздействия.

Здесь следует заметить, что уровень строгости решений дифференциальных игр и, в частности, игр преследования, базирующихся на решении уравнения Айзекса–Беллмана, ограничивается областью фазовых переменных, для которых указанное уравнение имеет смысл. Строгое обоснование решений может быть получено с использованием фундаментальных результатов Н. Н. Красовского и его учеников. Именно

на основе формализации дифференциальной игры, предложенной Н. Н. Красовским, оказывается возможным связать дескриптивную теорему о значении игры и ситуации равновесия с обобщенным минимаксным решением уравнения Айзекса–Беллмана (см. [Красовский, Котельникова, 2010]). Это же замечание относится и к неантагонистическим дифференциальным играм, рассмотренным в следующей главе, для которых подобные результаты еще не получены, но их получение является лишь делом времени.

Неантагонистическим некооперативным дифференциальным играм посвящена шестая глава. В некооперативном случае в качестве принципа оптимальности берется равновесие по Нэшу в позиционных стратегиях. Для определения абсолютных равновесий в регулярном случае обосновывается техника использования систем дифференциальных уравнений в частных производных Гамильтона–Якоби–Беллмана. И хотя эта техника трудно применима для решения сложных задач, в ряде конкретных случаев удается найти явное решение. Это касается, прежде всего, приведенных в главе задач управления совместным предприятием, ограничения вредных выбросов в атмосферу и совместной добычи ограниченного природного ресурса. Во всех указанных задачах равновесия находятся в явной форме.

В седьмой и восьмой главах рассматриваются дифференциальные кооперативные игры. При формировании кооперативного соглашения используется принцип оптимальности классической кооперативной теории. Однако в динамике, так же как и в дискретных игровых задачах, здесь имеет место нарушение динамической устойчивости (состоятельности во времени) основных принципов оптимальности. Поэтому проводится регуляризация игры путем введения нового управляющего воздействия — процедуры распределения дележа, что обеспечивает динамическую устойчивость принципа оптимальности. Кооперативное решение находится для прикладных задач, рассмотренных также в некооперативном случае.

В учебнике принята тройная нумерация подразделов (пунктов) и формул: номер главы, номер параграфа, номер подраздела. Для рисунков и таблиц в рамках главы используется сквозная (двойная) нумерация. Основным структурным элементом учебника является подраздел (пункт), на который и делаются ссылки в тексте. Например, пример 1 п. 2.1.3 или теорема п. 4.2.5.

Глава 1

Матричные игры

§ 1.1. Определение антагонистической игры в нормальной форме

1.1.1. Начнем изучение теории игр с простейшей статической модели — матричной игры, в которой участвуют два игрока, множество стратегий каждого из игроков конечно, а выигрыш одного игрока равен проигрышу другого.

Определение. Система

$$\Gamma = (X, Y, K), \quad (1.1.1)$$

где X и Y — непустые множества, и функция $K : X \times Y \rightarrow R^1$, называется антагонистической игрой в нормальной форме.

Элементы $x \in X$ и $y \in Y$ называются стратегиями игроков 1 и 2 соответственно в игре Γ , элементы декартового произведения $X \times Y$ (т. е. пары стратегий (x, y) , где $x \in X$ и $y \in Y$) — ситуациями, а функция K — функцией выигрыша игрока 1. Выигрыш игрока 2 в ситуации (x, y) полагается равным $[-K(x, y)]$; поэтому функция K также называется функцией выигрыша самой игры Γ , а игра Γ — игрой с нулевой суммой. Таким образом, используя принятую терминологию, для задания игры Γ необходимо определить множества стратегий X, Y игроков 1 и 2, а также функцию выигрыша K , заданную на множестве всех ситуаций $X \times Y$.

Игра Γ интерпретируется следующим образом. Игроки одновременно и независимо выбирают стратегии $x \in X, y \in Y$. После этого игрок 1 получает выигрыш, равный $K(x, y)$, а игрок 2 — $(-K(x, y))$.

1.1.2. Определение. Игра $\Gamma' = (X', Y', K')$ называется подыгрой игры $\Gamma = (X, Y, K)$, где $X' \subset X, Y' \subset Y$, а функция $K' : X' \times Y' \rightarrow R^1$ является сужением функции K на $X' \times Y'$.

В данной главе будут рассматриваться главным образом антагонистические игры, в которых множества стратегий игроков конечны.

Определение. Антагонистические игры, в которых оба игрока имеют конечные множества стратегий, называются матричными.

Пусть игрок 1 в матричной игре (1.1.1) имеет всего m стратегий. Упорядочим множество X стратегий первого игрока, т. е. установим взаимно однозначное соответствие между множествами $M = \{1, 2, \dots, m\}$ и X . Аналогично, если игрок 2 имеет n стратегий, то можно установить взаимно однозначное соответствие между множе-

ствами $N = \{1, 2, \dots, n\}$ и Y . Тогда игра Γ полностью определяется заданием матрицы $A = \{a_{ij}\}$, где $a_{ij} = K(x_i, y_j)$, $(i, j) \in M \times N$, $(x_i, y_j) \in X \times Y$, $i \in M, j \in N$ (отсюда и название игры — матричная). При этом игра Γ реализуется следующим образом. Игрок 1 выбирает строку $i \in M$, а игрок 2 (одновременно с игроком 1 и независимо от него) выбирает столбец $j \in N$. После этого игрок 1 получает выигрыш (a_{ij}) , а игрок 2 получает $(-a_{ij})$. Если выигрыш равен отрицательному числу, то речь идет о фактическом проигрыше игрока.

Игру Γ с матрицей выигрышей A обозначим Γ_A и назовем $(m \times n)$ -игрой согласно размерности матрицы A . Если из изложения понятно, об игре с какой матрицей идет речь, то индекс A будем опускать.

Нумерация стратегий в матричной игре может производиться различными способами, поэтому каждому отношению порядка, строго говоря, соответствует своя матрица. Таким образом, конечная антагонистическая игра может быть описана различными матрицами, отличающимися друг от друга лишь порядком строк и столбцов.

1.1.3. Пример 1 (Оборона города). Этот пример известен в литературе под названием «игра полковника Блотто» [Дрешер, 1964]. Полковник Блотто имеет m полков, а его противник — n полков. Противник защищает две позиции. Позиция будет занята полковником Блотто, если на ней наступающие полки окажутся в численном превосходстве. Противоборствующим сторонам требуется распределить полки между двумя позициями.

Определим выигрыш полковника Блотто (игрока 1) на каждой позиции. Если у него на позиции полков больше, чем у противника (игрока 2), то его выигрыш на этой позиции равен числу полков противника плюс один (занятие позиции равносильно захвату одного полка). Если у игрока 2 полков на позиции больше, чем у игрока 1, то игрок 1 теряет все свои полки на этой позиции и еще единицу (за потерю позиции). Если обе стороны имеют одинаковое число полков на позиции, то имеет место ничья и каждая из сторон ничего не получит. Общий выигрыш игрока 1 равен сумме выигрышей на обеих позициях.

Игра, очевидно, антагонистическая. Опишем стратегии игроков. Пусть, для определенности, $m > n$. Игрок 1 имеет следующие стратегии: $x_0 = (m, 0)$ — послать все полки на первую позицию; $x_1 = (m - 1, 1)$ — послать $(m - 1)$ полков на первую позицию, а один — на вторую; $x_2 = (m - 2, 2), \dots, x_{m-1} = (1, m - 1), x_m = (0, m)$. Противник (игрок 2) имеет следующие стратегии: $y_0 = (n, 0), y_1 = (n - 1, 1), \dots, y_n = (0, n)$.

Пусть игрок 1 выбрал стратегию x_0 , а игрок 2 — стратегию y_0 . Вычислим выигрыш a_{00} игрока 1 в этой ситуации. Поскольку $m > n$, на первой позиции выигрывает игрок 1. Его выигрыш равен $n + 1$. Следовательно, $a_{00} = n + 1$.

Теперь вычислим a_{01} . Поскольку $m > n - 1$, то на первой позиции выигрыш игрока 1 равен $n - 1 + 1 = n$. На второй позиции выигрывает игрок 2. Следовательно, проигрыш игрока 1 на этой позиции равен единице. Таким образом, $a_{01} = n - 1$. Рассуждая аналогично, получаем $a_{0j} = n - j + 1 - 1 = n - j$, $1 \leq j \leq n$. Далее, если $m - 1 > n$, то $a_{10} = n + 1 + 1 = n + 2$, $a_{11} = n - 1 + 1 = n$, $a_{1j} = n - j + 1 - 1 - 1 = n - j - 1$, $2 \leq j \leq n$.

В общем случае (для любых m и n) элементы a_{ij} , $i = \overline{0, m}$, $j = \overline{0, n}$, матрицы выигрышей вычисляются следующим образом:

$$a_{ij} = K(x_i, y_j) = \begin{cases} n+2, & \text{если } m-i > n-j, \quad i > j, \\ n-j+1, & \text{если } m-i > n-j, \quad i = j, \\ n-j-i, & \text{если } m-i > n-j, \quad i < j, \\ -m+i+j, & \text{если } m-i < n-j, \quad i > j, \\ j+1, & \text{если } m-i = n-j, \quad i > j, \\ -m-2, & \text{если } m-i < n-j, \quad i < j, \\ -i-1, & \text{если } m-i = n-j, \quad i < j, \\ -m+i-1, & \text{если } m-i < n-j, \quad i = j, \\ 0, & \text{если } m-i = n-j, \quad i = j. \end{cases}$$

Таким образом, при $m = 4, n = 3$, рассмотрев всевозможные ситуации, получим матрицу выигрышей A этой игры:

$$A = \begin{matrix} & y_0 & y_1 & y_2 & y_3 \\ \begin{matrix} x_0 \\ x_1 \\ x_2 \\ x_3 \\ x_4 \end{matrix} & \begin{bmatrix} 4 & 2 & 1 & 0 \\ 1 & 3 & 0 & -1 \\ -2 & 2 & 2 & -2 \\ -1 & 0 & 3 & 1 \\ 0 & 1 & 2 & 4 \end{bmatrix} \end{matrix}.$$

Пример 2 (Игра на уклонение) [Гейл, 1960]. Игроки 1 и 2 выбирают целые числа i и j из множества $\{1, \dots, n\}$, при этом игрок 1 выигрывает величину $|i - j|$. Игра антагонистическая. Матрица выигрышей этой игры квадратная, размера $(n \times n)$, где $a_{ij} = |i - j|$. Так, если $n = 4$, матрица A игры принимает вид

$$A = \begin{matrix} & 1 & 2 & 3 & 4 \\ \begin{matrix} 1 \\ 2 \\ 3 \\ 4 \end{matrix} & \begin{bmatrix} 0 & 1 & 2 & 3 \\ 1 & 0 & 1 & 2 \\ 2 & 1 & 0 & 1 \\ 3 & 2 & 1 & 0 \end{bmatrix} \end{matrix}.$$

Пример 3 (Дискретная игра типа дуэли) [Гейл, 1960]. Игроки продвигаются навстречу друг другу на n шагов. После каждого сделанного шага игрок может выстрелить или нет, но во время игры он может выстрелить только один раз. Считается, что вероятность того, что игрок попадает в своего противника, если выстрелит, продвинувшись на k шагов, равна k/n ($k \leq n$).

Стратегия игрока 1(2) заключается в принятии решения стрелять на i -м (j -м) шаге. Пусть $i < j$ и игрок 1 принимает решение стрелять на i -м шаге, а игрок 2 — на j -м шаге. Тогда выигрыш a_{ij} игрока 1 определяется формулой

$$a_{ij} = \frac{i}{n} - \left(1 - \frac{i}{n}\right) \frac{j}{n} = \frac{n(i-j) + ij}{n^2}.$$

Таким образом, выигрыш a_{ij} — это разность вероятностей поражения противника и собственной гибели в дуэли.

В случае $i > j$ первым стреляет игрок 2 и $a_{ij} = -a_{ji}$. Если же $i = j$, то полагаем $a_{ij} = 0$. Так, если положить $n = 5$, то матрица этой игры, умноженная на 25, имеет вид

$$A = \begin{bmatrix} 0 & -3 & -7 & -11 & -15 \\ 3 & 0 & 1 & -2 & -5 \\ 7 & -1 & 0 & 7 & 5 \\ 11 & 2 & -7 & 0 & 15 \\ 15 & 5 & -5 & -15 & 0 \end{bmatrix}.$$

Пример 4 (Игра «нападение — защита»). Пусть игрок 1 намерен атаковать один из объектов c_1, \dots, c_n , которые имеют положительные ценности $\tau_1 > 0, \dots, \tau_n > 0$. Игрок 2 защищает один из этих объектов. Будем считать, что если атакован незащищенный объект c_i , то он с достоверностью уничтожается (игрок 1 выигрывает τ_i), а защищенный — поражается с вероятностью $1 > \beta_i > 0$ (объект c_i выдерживает нападение с вероятностью $1 - \beta_i > 0$), т. е. игрок 1 выигрывает (в среднем) $\beta_i \tau_i$, $i = 1, 2, \dots, n$.

Тогда задача выбора объекта нападения (для игрока 1) и объекта защиты (для игрока 2) сводится к матричной игре с матрицей выигрышей

$$A = \begin{bmatrix} \beta_1 \tau_1 & \tau_1 & \dots & \tau_1 \\ \tau_2 & \beta_2 \tau_2 & \dots & \tau_2 \\ \dots & \dots & \dots & \dots \\ \tau_n & \tau_n & \dots & \beta_n \tau_n \end{bmatrix}.$$

Пример 5 (Игра дискретного поиска). Имеется n ячеек. Игрок 2 прячет предмет в одной из n ячеек, а игрок 1 хочет его найти. При проверке i -й ячейки игрок 1 тратит $\tau_i > 0$ усилий, при этом вероятность найти предмет в i -й ячейке (если там он спрятан) равна $0 < \beta_i \leq 1$, $i = 1, 2, \dots, n$.

Если предмет найден, то игрок 1 получает доход α . Стратегиями игроков являются номера ячеек, в которых игроки соответственно прячут и ищут предмет. Выигрыш игрока 1 равен разности между ожидаемым доходом и усилиями, затраченными на поиск предмета. Таким образом, задача поиска и прятания предмета сводится к матричной игре с матрицей выигрышей

$$A = \begin{bmatrix} \alpha\beta_1 - \tau_1 & -\tau_1 & -\tau_1 & \dots & -\tau_1 \\ -\tau_2 & \alpha\beta_2 - \tau_2 & -\tau_2 & \dots & -\tau_2 \\ \dots & \dots & \dots & \dots & \dots \\ -\tau_n & -\tau_n & -\tau_n & \dots & \alpha\beta_n - \tau_n \end{bmatrix}.$$

Пример 6 (Поиск «шумного» объекта.) Предположим, что игрок 1 ведет поиск подвижного объекта (игрок 2) с целью его обнаружения. Игрок 2 преследует противоположную цель (т. е. стремится уклониться от обнаружения). Игрок 1 может двигаться со скоростями $\alpha_1 = 1$, $\alpha_2 = 2$, $\alpha_3 = 3$, а игрок 2 — соответственно со скоростями $\beta_1 = 1$, $\beta_2 = 2$, $\beta_3 = 3$. Дальность действия средства обнаружения игрока 1 в зависимости от скоростей движения участников игры приведена в матрице

$$D = \begin{matrix} & \beta_1 & \beta_2 & \beta_3 \\ \alpha_1 & \begin{bmatrix} 4 & 5 & 6 \end{bmatrix} \\ \alpha_2 & \begin{bmatrix} 3 & 4 & 5 \end{bmatrix} \\ \alpha_3 & \begin{bmatrix} 1 & 2 & 3 \end{bmatrix} \end{matrix}.$$

Стратегиями игроков являются скорости движения, а в качестве выигрыша игрока 1 в ситуации (α_i, β_j) примем производительность поиска $a_{ij} = \alpha_i \delta_{ij}$, $i = \overline{1, 3}, j = \overline{1, 3}$, где δ_{ij} — элемент матрицы D . Тогда задача выбора скоростей игроков при поиске — укловнении может быть представлена матричной игрой с матрицей

$$A = \begin{matrix} & \begin{matrix} \beta_1 & \beta_2 & \beta_3 \end{matrix} \\ \begin{matrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{matrix} & \begin{bmatrix} 4 & 5 & 6 \\ 6 & 8 & 10 \\ 3 & 6 & 9 \end{bmatrix} \end{matrix}.$$

§ 1.2. Максиминные и минимаксные стратегии

1.2.1. Рассмотрим антагонистическую игру $\Gamma = (X, Y, K)$. Здесь каждый из игроков выбором стратегии стремится максимизировать свой выигрыш. Но для игрока 1 он определяется функцией $K(x, y)$, а для игрока 2 как $(-K(x, y))$, т. е. цели игроков прямо противоположны. При этом заметим, что выигрыш игрока 1(2) определен на ситуациях $(x, y) \in X \times Y$, складывающихся в процессе игры. Но каждая ситуация, а следовательно, и выигрыш игрока зависят не только от его выбора, но и от того, какая стратегия будет выбрана противником. Поэтому, стремясь получить возможно больший выигрыш, каждый игрок должен учитывать поведение противника.

Поясним сказанное на примере игры «оборона города». Если игрок 1 хочет получить максимальный выигрыш, то он должен принять стратегию x_0 (или x_4). В этом случае, если игрок 2 применит стратегию $y_0(y_3)$, то первый получит выигрыш, равный 4 единицам. Но если игрок 2 применит стратегию y_3 (соответственно y_0), то игрок 1 получит выигрыш, равный 0, т. е. потеряет 4 единицы. Аналогичные рассуждения можно провести и для игрока 2.

В теории игр предполагается, что оба игрока действуют разумно, т. е. стремятся к получению максимального выигрыша, считая, что соперник действует наилучшим (для себя) образом. Что может себе гарантировать игрок 1? Пусть игрок 1 выбрал стратегию x . Тогда в худшем случае он выиграет $\min_y K(x, y)$. Поэтому игрок 1 всегда может гарантировать себе выигрыш $\max_x \min_y K(x, y)$. Если отказаться от предположения достижимости экстремума, то игрок 1 может всегда получить выигрыш, сколь угодно близкий к величине

$$\underline{v} = \sup_{x \in X} \inf_{y \in Y} K(x, y), \quad (1.2.1)$$

которую будем называть *нижним значением игры*.

Если же внешний экстремум в (1.2.1) достигается, то величина \underline{v} называется также *максимумом*; принцип построения стратегии x , основанный на максимизации минимального выигрыша, — *принципом максимина*, а выбираемая в соответствии с этим принципом стратегия x — *максиминной стратегией* игрока 1.

Для игрока 2 можно провести аналогичные рассуждения. Пусть он выбрал стратегию y . Тогда в худшем случае он проиграет $\max_x K(x, y)$. Поэтому второй игрок всегда может себе гарантировать проигрыш не более, чем $\min_y \max_x K(x, y)$. Число

$$\bar{v} = \inf_{y \in Y} \sup_{x \in X} K(x, y) \quad (1.2.2)$$

называется *верхним значением* игры Γ , а в случае достижения внешнего экстремума в (1.2.2) и минимаксом. При этом принцип построения стратегии y , основанный на минимизации максимальных потерь, называется *принципом минимакса*, а выбираемая в соответствии с этим принципом стратегия y — *минимаксной стратегией* игрока 2. Подчеркнем, что существование минимаксной (максиминной) стратегии определяется достижимостью внешнего экстремума в (1.2.2) ((1.2.1)).

Пусть задана матричная $(m \times n)$ - игра Γ_A . Тогда экстремумы в (1.2.1) и (1.2.2) достигаются, а нижнее и верхнее значения игры соответственно равны

$$\underline{v} = \max_{1 \leq i \leq m} \min_{1 \leq j \leq n} a_{ij}, \quad (1.2.3)$$

$$\bar{v} = \min_{1 \leq j \leq n} \max_{1 \leq i \leq m} a_{ij}. \quad (1.2.4)$$

Минимакс и максимин для игры Γ_A могут быть найдены по следующей схеме:

$$\left[\begin{array}{cccc} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{array} \right] \left. \begin{array}{l} \min_j a_{1j} \\ \min_j a_{2j} \\ \dots \\ \min_j a_{mj} \end{array} \right\} \max_i \min_j a_{ij}$$

$$\underbrace{\begin{array}{cccc} \max_i a_{i1} & \max_i a_{i2} & \dots & \max_i a_{in} \end{array}}_{\min_j \max_i a_{ij}}$$

Так, в игре Γ_A с матрицей

$$A = \begin{bmatrix} 1 & 0 & 4 \\ 5 & 3 & 8 \\ 6 & 0 & 1 \end{bmatrix}$$

нижнее значение (максимин) \underline{v} и максиминная стратегия i_0 первого игрока равны $\underline{v} = 3$, $i_0 = 2$, а верхнее значение (минимакс) \bar{v} и минимаксная стратегия j_0 второго игрока $\bar{v} = 3$, $j_0 = 2$.

1.2.2. Для любой игры $\Gamma = (X, Y, K)$ справедливо следующее утверждение.

Лемма. В антагонистической игре Γ

$$\underline{v} \leq \bar{v} \quad (1.2.5)$$

или

$$\sup_{x \in X} \inf_{y \in Y} K(x, y) \leq \inf_{y \in Y} \sup_{x \in X} K(x, y). \quad (1.2.6)$$

Доказательство. Пусть $x \in X$ произвольная стратегия игрока 1. Тогда имеем

$$K(x, y) \leq \sup_{x \in X} K(x, y).$$

Отсюда получаем

$$\inf_{y \in Y} K(x, y) \leq \inf_{y \in Y} \sup_{x \in X} K(x, y).$$

Теперь заметим, что в правой части последнего неравенства стоит константа, а значение $x \in X$ выбиралось произвольно. Поэтому выполняется неравенство

$$\sup_{x \in X} \inf_{y \in Y} K(x, y) \leq \inf_{y \in Y} \sup_{x \in X} K(x, y).$$

§ 1.3. Ситуации равновесия

1.3.1. Рассмотрим вопрос об оптимальном поведении игроков в антагонистической игре. Естественно считать оптимальной в игре $\Gamma = (X, Y, K)$ такую ситуацию $(x^*, y^*) \in X \times Y$, отклоняться от которой невыгодно ни одному из игроков. Такая ситуация (x^*, y^*) называется *равновесной*, а принцип оптимальности, основанный на построении равновесной ситуации, — *принципом равновесия*. Конечно, для этого необходимо существование равновесия (т. е. чтобы принцип оптимальности был реализуем).

Определение. В антагонистической игре $\Gamma = (X, Y, K)$ ситуация (x^*, y^*) называется *ситуацией равновесия или седловой точкой*, если

$$K(x, y^*) \leq K(x^*, y^*), \quad (1.3.1)$$

$$K(x^*, y) \geq K(x^*, y^*) \quad (1.3.2)$$

для всех $x \in X$ и $y \in Y$.

Множество всех ситуаций равновесия в игре Γ обозначим через

$$Z(\Gamma) \subset X \times Y.$$

Для матричной игры Γ_A речь идет о *седловых точках* матрицы выигрышей A , т. е. таких точках (i^*, j^*) , что для всех $i \in M$ и $j \in N$ выполняются неравенства

$$a_{ij^*} \leq a_{i^*j^*} \leq a_{i^*j}.$$

В седловой точке элемент матрицы $a_{i^*j^*}$ является одновременно минимумом в своей строке и максимумом в своем столбце. Например, в игре с матрицей

$$\begin{bmatrix} 1 & 0 & 4 \\ 5 & 3 & 8 \\ 6 & 0 & 1 \end{bmatrix} \text{ ситуация } (2,2) \text{ является равновесной.}$$

1.3.2. Множество ситуаций равновесия в антагонистической игре Γ обладает свойствами, которые позволяют говорить об оптимальности ситуации равновесия и входящих в нее стратегий.

Теорема. Пусть (x_1^*, y_1^*) , (x_2^*, y_2^*) — две произвольные ситуации равновесия в антагонистической игре Γ . Тогда

$$1) K(x_1^*, y_1^*) = K(x_2^*, y_2^*);$$

$$2) (x_1^*, y_2^*) \in Z(\Gamma), (x_2^*, y_1^*) \in Z(\Gamma).$$

Доказательство. Из определения ситуации равновесия для всех $x \in X$ и $y \in Y$ имеем

$$K(x, y_1^*) \leq K(x_1^*, y_1^*) \leq K(x_1^*, y); \quad (1.3.3)$$

$$K(x, y_2^*) \leq K(x_2^*, y_2^*) \leq K(x_2^*, y). \quad (1.3.4)$$

Подставим в левую часть неравенства (1.3.3) x_2^* , в правую — y_2^* , в левую часть неравенства (1.3.4) — x_1^* и в правую y_1^* . Тогда получим

$$K(x_2^*, y_1^*) \leq K(x_1^*, y_1^*) \leq K(x_1^*, y_2^*) \leq K(x_2^*, y_2^*) \leq K(x_2^*, y_1^*).$$

Откуда следует равенство

$$K(x_1^*, y_1^*) = K(x_2^*, y_2^*) = K(x_2^*, y_1^*) = K(x_1^*, y_2^*). \quad (1.3.5)$$

Покажем справедливость второго из утверждений. Рассмотрим ситуацию (x_2^*, y_1^*) . Тогда из (1.3.3)–(1.3.5) имеем

$$K(x, y_1^*) \leq K(x_1^*, y_1^*) = K(x_2^*, y_1^*) = K(x_2^*, y_2^*) \leq K(x_2^*, y) \quad (1.3.6)$$

для всех $x \in X, y \in Y$. Доказательство равновесности ситуации $(x_1^*, y_2^*) \in Z(\Gamma)$ проводится аналогично.

Из теоремы следует, что функция выигрыша принимает одно и то же значение во всех ситуациях равновесия. Поэтому разумно ввести следующее определение.

Определение. Пусть (x^*, y^*) — ситуация равновесия в игре Γ . Тогда число

$$v = K(x^*, y^*) \quad (1.3.7)$$

называется значением игры Γ .

Из второго утверждения теоремы следует, в частности, такой факт. Обозначим X^* и Y^* проекции множества $Z(\Gamma)$ на X и Y соответственно, т. е.

$$X^* = \{x^* \mid x^* \in X, \exists y^* \in Y, (x^*, y^*) \in Z(\Gamma)\},$$

$$Y^* = \{y^* \mid y^* \in Y, \exists x^* \in X, (x^*, y^*) \in Z(\Gamma)\}.$$

Тогда множество $Z(\Gamma)$ можно представить в виде декартового произведения

$$Z(\Gamma) = X^* \times Y^*. \quad (1.3.8)$$

Доказательство (1.3.8), как следствие второго утверждения теоремы, предоставим читателю.

Определение. Множество $X^*(Y^*)$ называется множеством оптимальных стратегий игрока 1(2) в игре Γ , а его элементы — оптимальными стратегиями игрока 1(2).

Заметим, что равенство (1.3.5) указывает на взаимозаменяемость оптимальных стратегий, т. е. любая пара оптимальных стратегий образует ситуацию равновесия, а выигрыш в ней равен значению игры.

1.3.3. Оптимальность поведения игроков не изменится, если в игре множества стратегий остаются прежними, а функция выигрыша умножается на положительную константу (или к ней прибавляется постоянное число).

Лемма о масштабе. Пусть $\Gamma = (X, Y, K)$ и $\Gamma' = (X, Y, K')$ — две антагонистические игры, причем

$$K' = \beta K + \alpha, \quad \beta > 0, \alpha = \text{const}, \beta = \text{const}. \quad (1.3.9)$$

Тогда

$$Z(\Gamma') = Z(\Gamma), \quad v_{\Gamma'} = \beta v_{\Gamma} + \alpha. \quad (1.3.10)$$

Доказательство. Пусть (x^*, y^*) — ситуация равновесия в игре Γ . Тогда имеем

$$K'(x^*, y^*) = \beta K(x^*, y^*) + \alpha \leq \beta K(x^*, y) + \alpha = K'(x^*, y),$$

$$K'(x, y^*) = \beta K(x, y^*) + \alpha \leq \beta K(x^*, y^*) + \alpha = K'(x^*, y^*),$$

для всех $x \in X$ и $y \in Y$. Следовательно, $(x^*, y^*) \in Z(\Gamma')$, $Z(\Gamma) \subset Z(\Gamma')$.

Обратно, пусть $(x, y) \in Z(\Gamma')$. Тогда

$$K(x, y) = (1/\beta)K'(x, y) - \alpha/\beta$$

и, рассуждая аналогично, получаем, что $(x, y) \in Z(\Gamma)$. Следовательно, $Z(\Gamma) = Z(\Gamma')$, при этом выполняется равенство

$$v_{\Gamma'} = K'(x^*, y^*) = \beta K(x^*, y^*) + \alpha = \beta v_{\Gamma} + \alpha.$$

Содержательно данная лемма говорит о стратегической эквивалентности двух игр, отличающихся лишь началом отсчета выигрышей, а также масштабом их измерения.

1.3.4. Теперь установим связь между принципом равновесия и принципами минимакса и максимина в антагонистической игре.

Теорема. *Для того чтобы в игре $\Gamma = (X, Y, K)$ существовала ситуация равновесия, необходимо и достаточно, чтобы существовали минимакс и максимин*

$$\min_y \sup_x K(x, y), \quad \max_x \inf_y K(x, y) \quad (1.3.11)$$

и выполнялось равенство:

$$\underline{v} = \max_x \inf_y K(x, y) = \min_y \sup_x K(x, y) = \bar{v}. \quad (1.3.12)$$

Доказательство. Необходимость. Пусть $(x^*, y^*) \in Z(\Gamma)$. Тогда для всех $x \in X$ и $y \in Y$ выполняются следующие неравенства:

$$K(x, y^*) \leq K(x^*, y^*) \leq K(x^*, y) \quad (1.3.13)$$

отсюда

$$\sup_x K(x, y^*) \leq K(x^*, y^*). \quad (1.3.14)$$

Вместе с тем имеем

$$\inf_y \sup_x K(x, y) \leq \sup_x K(x, y^*). \quad (1.3.15)$$

Сравнивая (1.3.14) и (1.3.15), получаем

$$\inf_y \sup_x K(x, y) \leq \sup_x K(x, y^*) \leq K(x^*, y^*). \quad (1.3.16)$$

Рассуждая аналогично, приходим к неравенствам

$$K(x^*, y^*) \leq \inf_y K(x^*, y) \leq \sup_x \inf_y K(x, y). \quad (1.3.17)$$

Таким образом,

$$\inf_y \sup_x K(x, y) \leq \sup_x K(x, y^*).$$

С другой стороны, всегда выполняется обратное неравенство (1.2.6). Итак, получаем

$$\sup_x \inf_y K(x, y) = \inf_y \sup_x K(x, y), \quad (1.3.18)$$

при этом неравенства (1.3.16), (1.3.17) выполняются как равенства

$$\min_y \sup_x K(x, y) = \sup_x K(x, y^*) = K(x^*, y^*),$$

$$\max_x \inf_y K(x, y) = \inf_y K(x^*, y) = K(x^*, y^*),$$

т. е. внешние экстремумы $\min \sup$ и $\max \inf$ достигаются в точках y^* и x^* соответственно.

Достаточность. Пусть существуют $\min \sup$ и $\max \inf$,

$$\max_x \inf_y K(x, y) = \inf_y K(x^*, y); \quad (1.3.19)$$

$$\min_y \sup_x K(x, y) = \sup_x K(x, y^*) \quad (1.3.20)$$

и выполняется равенство (1.3.12). Покажем, что ситуация (x^*, y^*) является равновесной. Действительно,

$$K(x^*, y^*) \geq \inf_y K(x^*, y) = \max_x \inf_y K(x, y); \quad (1.3.21)$$

$$K(x^*, y^*) \leq \sup_x K(x, y^*) = \min_y \sup_x K(x, y). \quad (1.3.22)$$

Согласно равенству (1.3.12), $\min \sup$ равен $\max \inf$, а из (1.3.21), (1.3.22) следует, что он равен также и величине $K(x^*, y^*)$, т. е. неравенства в (3.21), (3.22) выполняются как равенства. Теперь имеем

$$K(x^*, y^*) = \inf_y K(x^*, y) \leq K(x^*, y),$$

$$K(x^*, y^*) = \sup_x K(x, y^*) \geq K(x, y^*)$$

для всех $x \in X$ и $y \in Y$, т. е. $(x^*, y^*) \in Z(\Gamma)$.

Заметим, что в ходе доказательства показано, что общее значение $\min \sup$ и $\max \inf$ равно $K(x^*, y^*) = v$ — значению игры, при этом любая $\min \sup$ ($\max \inf$) стратегия $y^*(x^*)$ в условиях теоремы является оптимальной, т. е. ситуация (x^*, y^*) является равновесной.

Из доказательства теоремы получаем следующее утверждение.

Следствие 1. Если $\min \sup$ и $\max \inf$ в (1.3.11) существуют и достигаются на \bar{y} и \bar{x} соответственно, то

$$\max_x \inf_y K(x, y) \leq K(\bar{x}, \bar{y}) \leq \min_y \sup_x K(x, y). \quad (1.3.23)$$

Игры, в которых существуют ситуации равновесия, называются *вполне определенными*. Поэтому данная теорема устанавливает критерий вполне определенной игры и может быть переформулирована следующим образом. Для того чтобы игра была вполне определена, необходимо и достаточно, чтобы существовали $\min \sup$ и $\max \inf$ в (1.3.11) и выполнялось равенство (1.3.12).

Заметим, что в матричной игре Γ_A экстремумы в (1.3.11) всегда достигаются, поэтому теорема принимает следующий вид.

Следствие 2. Для того чтобы матричная $(m \times n)$ -игра Γ_A была вполне определена, необходимо и достаточно выполнение равенства

$$\min_{j=1,2,\dots,n} \max_{i=1,2,\dots,m} \alpha_{ij} = \max_{i=1,2,\dots,m} \min_{j=1,2,\dots,n} \alpha_{ij}. \quad (1.3.24)$$

Например, в игре с матрицей $\begin{bmatrix} 1 & 4 & 1 \\ 2 & 3 & 4 \\ 0 & -2 & 7 \end{bmatrix}$ ситуация (2,1) является равновесной.

При этом

$$\max_i \min_j a_{ij} = \min_j \max_i a_{ij} = 2.$$

С другой стороны, игра с матрицей $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ не имеет ситуации равновесия, поскольку

$$\min_j \max_i a_{ij} = 1 > \max_i \min_j a_{ij} = 0.$$

Заметим, что игры, сформулированные в примерах 1 – 3 (п. 1.1.3), не являются вполне определенными, а игра в примере 6 вполне определена и ее значение $v = 6$.

§ 1.4. Смешанное расширение игры

1.4.1. Рассмотрим матричную игру Γ_A . Если в ней существует ситуация равновесия, то минимакс равен максимину, причем согласно определению ситуации равновесия каждый из игроков может сообщить свою оптимальную (максиминную) стратегию противнику и от этого ни один из игроков не может получить дополнительную выгоду. Теперь предположим, что в игре Γ_A не существует ситуации равновесия. Тогда согласно теореме п. 1.3.4 и лемме п. 1.2.2 имеем

$$\min_j \max_i \alpha_{ij} - \max_i \min_j a_{ij} > 0. \quad (1.4.1)$$

В этом случае максиминная и минимаксная стратегии не являются оптимальными. Более того, игрокам бывает невыгодно их придерживаться, так как они могут получить больший выигрыш. Однако сообщение о выборе стратегии противнику может привести к еще большим потерям, чем в случае максиминной или минимаксной стратегии.

Действительно, пусть матрица A имеет вид

$$A = \begin{bmatrix} 7 & 3 \\ 2 & 5 \end{bmatrix}.$$

Для такой матрицы $\min_j \max_i \alpha_{ij} = 5$, $\max_i \min_j a_{ij} = 3$, т. е. ситуации равновесия не существует.

Обозначим через i^* максиминную стратегию игрока 1 ($i^* = 1$), а минимаксную стратегию игрока 2 через j^* ($j^* = 2$). Пусть игрок 2 придерживается стратегии $j^* = 2$, а игрок 1 выберет стратегию $i = 2$. Тогда последний получит выигрыш 5, т. е. на 2 единицы больше, чем максимин. Однако если игрок 2 догадается о выборе игрока 1, то он изменит стратегию на $j = 1$, и тогда первый получит выигрыш лишь 2 единицы, т. е. на единицу меньше, чем в случае максимина. Аналогичные рассуждения можно провести и для второго игрока. По существу вопрос стоит о том, как разделить между игроками величину (1.4.1)?

Оказывается, что в этом случае игрокам разумно действовать случайно, что обеспечивает наибольшую скрытность выбора стратегии. Результат выбора не может стать известным противнику, поскольку до реализации случайного механизма не известен самому игроку.

1.4.2. Определение. *Случайная величина, значениями которой являются стратегии игрока, называется его смешанной стратегией.*

Так, для матричной игры Γ_A смешанной стратегией игрока 1 является случайная величина, значениями которой являются номера строк $i \in M$, $M = \{1, 2, \dots, m\}$ матрицы A . Аналогично определяется смешанная стратегия игрока 2, значениями которой являются номера $j \in N$ столбцов матрицы A .

Учитывая только что введенное определение смешанных стратегий, прежние стратегии будем называть «чистыми». Так как случайная величина характеризуется своим распределением, то будем отождествлять в дальнейшем смешанную стратегию с вероятностным распределением на множестве чистых стратегий. Таким образом, смешанная стратегия x игрока 1 в игре есть m -мерный вектор

$$x = (\xi_1, \dots, \xi_m), \sum_{i=1}^m \xi_i = 1, \xi_i \geq 0, i = 1, \dots, m. \quad (1.4.2)$$

Аналогично, смешанная стратегия y игрока 2 есть n -мерный вектор

$$y = (\eta_1, \dots, \eta_n), \sum_{j=1}^n \eta_j = 1, \eta_j \geq 0, j = 1, \dots, n. \quad (1.4.3)$$

При этом $\xi_i \geq 0$ и $\eta_j \geq 0$ — вероятности выбора чистых стратегий $i \in M$ и $j \in N$ соответственно при использовании игроками смешанных стратегий x и y .

Обозначим через X и Y соответственно множества смешанных стратегий первого и второго игроков. Нетрудно заметить, что множество смешанных стратегий каждого игрока — компакт в соответствующем конечномерном евклидовом пространстве (замкнутое, ограниченное множество).

Определение. Пусть $x = (\xi_1, \dots, \xi_m) \in X$ — смешанная стратегия игрока 1. Тогда множество индексов

$$M_x = \{i | i \in M, \xi_i > 0\}, \quad (1.4.4)$$

где $M = \{1, 2, \dots, m\}$, назовем спектром стратегии x .

Аналогично для смешанной стратегии $y = (\eta_1, \dots, \eta_n) \in Y$ игрока 2 спектр N_y определяется следующим образом:

$$N_y = \{j | j \in N, \eta_j > 0\}, \quad (1.4.5)$$

где $N = \{1, 2, \dots, n\}$. Таким образом, спектр смешанной стратегии состоит из таких чистых стратегий, которые выбираются с положительными вероятностями.

Для любой смешанной стратегии x спектр $M_x \neq \emptyset$, поскольку вектор x имеет неотрицательные компоненты, сумма которых равна 1 [см. (1.4.2)].

Рассмотрим смешанную стратегию $u_i = (\xi_1, \dots, \xi_1, \dots, \xi_m) \in X$, где $\xi_i = 1$, $\xi_j = 0$, $j \neq i$, $i = 1, 2, \dots, m$. Такая стратегия предписывает выбор i -ой строки матрицы A с вероятностью 1. Естественно отождествлять смешанную стратегию $u_i \in X$ с выбором i -й строки, т. е. с чистой стратегией $i \in M$ игрока 1. Аналогично отождествим смешанную стратегию $w_j = (\eta_1, \dots, \eta_j, \dots, \eta_n) \in Y$, где $\eta_j = 1$, $\eta_i = 0$, $i \neq j$, $j = 1, \dots, n$ с чистой стратегией $j \in N$ игрока 2. Тем самым мы получили, что множество смешанных стратегий игрока есть расширение его пространства чистых стратегий.

Определение. Пара (x, y) смешанных стратегий игроков в матричной игре Γ_A называется ситуацией в смешанных стратегиях.

Определим выигрыш игрока 1 в ситуации (x, y) в смешанных стратегиях для матричной $(m \times n)$ -игры Γ_A как математическое ожидание его выигрыша при условии, что игроки используют смешанные стратегии соответственно x и y . Выбор стратегий игроками осуществляется независимо друг от друга, поэтому математическое ожидание выигрыша $K(x, y)$ в ситуации (x, y) в смешанных стратегиях $x = (\xi_1, \dots, \xi_m)$, $y = (\eta_1, \dots, \eta_n)$ равно

$$K(x, y) = \sum_{i=1}^m \sum_{j=1}^n a_{ij} \xi_i \eta_j = (xA)y = x(Ay). \quad (1.4.6)$$

При этом функция $K(x, y)$ является непрерывной по $x \in X$ и $y \in Y$. Заметим, что выигрыши $K(i, y)$, $K(x, j)$ при применении одним из игроков чистой стратегии (i или j соответственно), а другим — смешанной стратегии (y или x) имеют вид

$$K(i, y) = K(u_i, y) = \sum_{j=1}^n a_{ij} \eta_j = a_i y, \quad i = 1, \dots, m,$$

$$K(x, j) = K(x, w_j) = \sum_{i=1}^m a_{ij} \xi_i = x a^j, \quad j = 1, \dots, n,$$

где a_i — i -я строка, а a^j — j -й столбец $(m \times n)$ -матрицы A .

Таким образом, от матричной игры $\Gamma_A = (M, N, A)$ мы перешли к новой игре $\bar{\Gamma}_A = (X, Y, K)$, где X и Y — множества смешанных стратегий в игре Γ_A и K — функция выигрыша в смешанных стратегиях (математическое ожидание выигрыша). Игру $\bar{\Gamma}_A$ будем называть *смешанным расширением* игры Γ_A . Игра Γ_A является подыгрой для $\bar{\Gamma}_A$, т. е. $\Gamma_A \subset \bar{\Gamma}_A$.

1.4.3. Определение. Ситуация (x^*, y^*) в игре $\bar{\Gamma}_A$ образует ситуацию равновесия, а число $v = K(x^*, y^*)$ является значением игры $\bar{\Gamma}_A$, если для всех $x \in X$ и $y \in Y$

$$K(x, y^*) \leq K(x^*, y^*) \leq K(x^*, y). \quad (1.4.7)$$

Теорема п. 1.3.2 очевидным образом справедлива и для ситуаций равновесия в смешанном расширении $\bar{\Gamma}_A$ игры Γ_A . Более того, согласно теореме п. 1.3.4 стратегии x^* и y^* являются соответственно максиминной и минимаксной, поскольку внешние экстремумы в (1.3.11) достигаются (функция $K(x, y)$ непрерывна на компактных множествах X и Y).

В лемме п. 1.3.3 была показана стратегическая эквивалентность двух игр, отличающихся лишь началом отсчета выигрышей, а также масштабом их измерения (лемма о масштабе). Оказывается, что если две матричные игры Γ_A и $\Gamma_{A'}$ находятся в условиях этой леммы, то их смешанные расширения стратегически эквивалентны. Формально этот факт устанавливается следующим утверждением.

Лемма. Пусть Γ_A и $\Gamma_{A'}$ — две матричные $(m \times n)$ -игры, где

$$A' = \alpha A + B, \quad \alpha > 0, \alpha = \text{const},$$

а B — матрица с одинаковыми элементами β , т. е. $\beta_{ij} = \beta$ для всех i и j . Тогда $Z(\bar{\Gamma}_{A'}) = Z(\bar{\Gamma}_A)$, $\bar{v}_{A'} = \alpha \bar{v}_A + \beta$, где $\bar{\Gamma}_{A'}$ и $\bar{\Gamma}_A$ — смешанные расширения игр $\Gamma_{A'}$ и Γ_A соответственно, а $\bar{v}_{A'}$, \bar{v}_A — значения игр $\bar{\Gamma}_{A'}$ и $\bar{\Gamma}_A$.

Доказательство. Обе матрицы A и A' одинаковой размерности $m \times n$; поэтому множества смешанных стратегий в играх $\Gamma_{A'}$ и Γ_A совпадают. Покажем, что для любой ситуации (x, y) в смешанных стратегиях выполняется равенство

$$K'(x, y) = \alpha K(x, y) + \beta, \quad (1.4.8)$$

где K' и K — выигрыши игрока 1 в играх $\bar{\Gamma}_{A'}$ и $\bar{\Gamma}_A$ соответственно.

Действительно, для всех $x \in X$ и $y \in Y$ имеем

$$K'(x, y) = xA'y = \alpha(xAy) + xBy = \alpha K(x, y) + \beta.$$

Из леммы о масштабе следует, что $Z(\bar{\Gamma}_{A'}) = Z(\bar{\Gamma}_A)$, $\bar{v}_{A'} = \alpha \bar{v}_A + \beta$.

Пример 7. Проверим, что стратегии $y^* = (1/2, 1/4, 1/4)$, $x^* = (1/2, 1/4, 1/4)$ оптимальны, а $v = 0$ — значение игры $\bar{\Gamma}_A$ с матрицей

$$A = \begin{bmatrix} 1 & -1 & -1 \\ -1 & -1 & 3 \\ -1 & 3 & -1 \end{bmatrix}.$$

Упростим матрицу A (с целью получения максимального числа нулей). Прибавляя ко всем элементам матрицы A единицу, получим матрицу

$$A' = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 0 & 4 \\ 0 & 4 & 0 \end{bmatrix}.$$

Каждый элемент матрицы A разделим на 2. Новая матрица принимает вид

$$A'' = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 2 \\ 0 & 2 & 0 \end{bmatrix}.$$

По лемме значение игр связано равенством $v_{A''} = \frac{1}{2}v_{A'} = \frac{1}{2}(v_A + 1)$. Таким образом, требуется проверить, что значение игры $\Gamma_{A''}$ равно $1/2$. Действительно, $K''(x^*, y^*) = x^{*T} A'' y^* = 1/2$. С другой стороны, для каждой стратегии $y \in Y$, $y = (\eta_1, \eta_2, \eta_3)$ имеем $K''(x^*, y) = \frac{1}{2}\eta_1 + \frac{1}{2}\eta_2 + \frac{1}{2}\eta_3 = \frac{1}{2} \cdot 1 = \frac{1}{2}$, а для всех $x = (\xi_1, \xi_2, \xi_3)$, $x \in X$, $K''(x, y^*) = \frac{1}{2}\xi_1 + \frac{1}{2}\xi_2 + \frac{1}{2}\xi_3 = \frac{1}{2}$. Следовательно, указанные стратегии x^*, y^* являются оптимальными, а $v_A = 0$.

В дальнейшем, говоря о матричной игре Γ_A , будем предполагать, что речь идет о ее смешанном расширении $\bar{\Gamma}_A$.

§ 1.5. Некоторые сведения из теории выпуклых множеств и систем линейных неравенств

Этот параграф носит вспомогательный характер и при первом чтении может быть опущен. Однако для понимания доказательств последующих утверждений полезно напомнить широко распространенные понятия и результаты. Большинство из них будет приведено без доказательств, в необходимых случаях даны ссылки на специальную литературу.

1.5.1. Множество $M \subset R^m$ называется *выпуклым*, если вместе с любыми двумя точками этого множества $x_1, x_2 \in M$ в нем содержатся все точки отрезка $\lambda x_1 + (1 - \lambda)x_2, 0 \leq \lambda \leq 1$. Понятие выпуклого множества можно сформулировать и в более общем, но эквивалентном виде.

Множество $M \subset R^m$ называется выпуклым, если вместе с точками x_1, \dots, x_k из M , оно содержит и все точки вида

$$x = \sum_{i=1}^k \lambda_i x_i, \lambda_i \geq 0, \sum_{i=1}^k \lambda_i = 1,$$

называемые выпуклыми линейными комбинациями точек x_1, \dots, x_k .

Пересечение выпуклых множеств всегда выпукло.

Рассмотрим систему линейных неравенств

$$xA \leq b$$

или

$$xa^j \leq \beta_j, j \in N, N = \{1, \dots, n\}, \quad (1.5.1)$$

где $A = [a^j, j \in N]$ — $(m \times n)$ -матрица, $x \in R^m$, $b = (\beta_1, \dots, \beta_n) \in R^n$.

Обозначим как $\tilde{X} = \{x | xA \leq b\}$ множество решений системы (1.5.1). Непосредственно из определения следует, что \tilde{X} — выпуклое множество. Множество \tilde{X} называется *выпуклым многогранным множеством*, заданным системой ограничений (1.5.1).

1.5.2. Точка $X \in M$, где M — выпуклое множество, называется *крайней точкой*, если из условия $x = \lambda x_1 + (1 - \lambda)x_2$, $x_1 \in M, x_2 \in M$ и $0 < \lambda < 1$ следует, что $x_1 = x_2 = x$. Содержательно определение означает, что $x \in M$ — крайняя точка, если не существует отрезка, содержащего две точки из M , для которого x является внутренней. Заметим, что крайняя точка выпуклого множества всегда является граничной, обратное неверно.

Пусть X — выпуклое многогранное множество, заданное системой ограничений (1.5.1). Тогда справедливы следующие утверждения.

Теорема. Множество \tilde{X} имеет крайние точки тогда и только тогда, когда $\text{rank} A = \text{rank}[a^j, j \in N] = m$ [Ашманов, 1981].

Теорема. Для того чтобы точка $x_0 \in X$ была крайней, необходимо и достаточно, чтобы она была решением системы

$$x_0 a^j = \beta_j, j \in N_1; \quad (1.5.2)$$

$$x_0 a^j \leq \beta_j, i \in N \setminus N_1, \quad (1.5.3)$$

где $N_1 \subset N$, $\text{rank}[a^i, i \in N_1] = m$ [Ашманов, 1981].

Последняя теорема дает алгоритм нахождения крайних точек множества \tilde{X} . Для этого необходимо рассмотреть столбцовые базисы матрицы A , решить систему линейных уравнений (1.5.2) и проверить выполнение неравенств (1.5.3). Однако такой способ поиска крайних точек многогранного множества мало пригоден для практики, поскольку он связан с полным перебором всевозможных столбцовых базисов матрицы A .

1.5.3. *Выпуклой оболочкой множества P* будем называть пересечение всех выпуклых множеств, содержащих P , и обозначать $\text{conv}(P)$. Данное определение эквивалентно следующему. Выпуклая оболочка множества P состоит из всех выпуклых линейных комбинаций всевозможных точек из P , т. е.

$$\text{conv}(P) = \{x | x = \sum_{i=1}^n \lambda_i x_i, \sum_{i=1}^n \lambda_i = 1, \lambda_i \geq 0, x_i \in P\}.$$

Выпуклая оболочка конечного числа точек называется *выпуклым многогранником*, порожденным этими точками. Выпуклый многогранник порожден своими крайними точками. Так, если рассмотреть множество X смешанных стратегий игрока 1 в $(m \times n)$ -игре, то $X = \text{conv}\{u_1, \dots, u_m\}$,

где $u_i = (0, \dots, 0, 1, 0, \dots, 0)$ — орты пространства R^m или чистые стратегии игрока 1. Множество X является выпуклым многогранником размерности $(m-1)$ и называется также $(m-1)$ -мерным симплексом или фундаментальным симплексом). При этом все векторы u_i (чистые стратегии) являются крайними точками многогранника X . Аналогичные утверждения справедливы для множества Y смешанных стратегий игрока 2.

Конусом C называется множество таких точек, что если $x \in C, \lambda \geq 0$, то $\lambda x \in C$.

Содержательно, конус C — это такое подмножество R^m , которое вместе с точкой x содержит и всю полупрямую (x) , где

$$(x) = \{y \mid y = \lambda x, \lambda \geq 0\}.$$

Конус C называется *выпуклым конусом*, если выполняется условие: для всех $x, y \in C$ справедливо $x + y \in C$. Другими словами, конус C — выпуклый, если он замкнут относительно операции сложения. Можно дать и другое эквивалентное определение. Конус называется выпуклым, если он является выпуклым множеством. Сумма выпуклых конусов $C_1 + C_2 = \{c \mid c = c_1 + c_2, c_1 \in C_1, c_2 \in C_2\}$ и их пересечение $C_1 \cap C_2$ также являются выпуклыми конусами.

Непосредственной проверкой определения можно показать, что множество $C = \{x \mid xA \leq 0\}$ решений однородной системы линейных неравенств, соответствующей (1.5.1), является выпуклым конусом.

Пусть \tilde{X} — выпуклое многогранное множество, заданное системой ограничений (1.5.1), записанной в эквивалентной форме

$$\sum_{i=1}^m \xi_i a_i \leq b, \quad (1.5.4)$$

где $x = (\xi_1, \dots, \xi_m) \in R^m$, a_i — i -я строка матрицы A , $i = 1, \dots, m$. Предположим, что $\text{rank} A = r, r \leq m$, и векторы a_1, \dots, a_r образуют строчечный базис матрицы A . Разложим остальные строки по базису

$$a_j = \sum_{i=1}^r \delta_{ij} a_i, \quad j = r+1, \dots, m. \quad (1.5.5)$$

Подставляя (1.5.5) в (1.5.4), получим эквивалентную (1.5.4) систему неравенств

$$\sum_{i=1}^r (\xi_i + \sum_{j=r+1}^m \xi_j \delta_{ij}) a_i \leq b. \quad (1.5.6)$$

Обозначим через X_0 множество векторов $x = (\xi_1, \dots, \xi_m) \in R^m$, удовлетворяющих неравенствам (1.5.6) и условию $\xi_j = 0, j = r+1, \dots, m$. По теореме п. 1.5.2, множество X_0 имеет крайние точки. Справедлива следующая теорема [Ашманов, 1981].

Теорема о представлении многогранного множества. Пусть \tilde{X} — многогранное множество, заданное системой ограничений (1.5.4). Тогда

$$\tilde{X} = M + C,$$

где $M + C = \{x \mid x = y + z, y \in M, z \in C\}$, M — выпуклый многогранник, порожденный крайними точками многогранного множества X_0 , заданного (1.5.6), а $C = \{x \mid xA \leq 0\}$ — выпуклый конус.

Из теоремы, в частности, следует, что если множество \tilde{X} решений системы (1.5.4) ограничено, то \tilde{X} — выпуклый многогранник.

1.5.4. Напомним, что задача нахождения $\min cx$ при ограничениях

$$xA \geq b, \quad x \geq 0, \quad (1.5.7)$$

где A — $(m \times n)$ -матрица, $c \in R^m$, $x \in R^m$, $b \in R^n$ называется прямой стандартной задачей линейного программирования, а задача, заключающаяся в определении $\max by$ при ограничениях

$$Ay \leq c, \quad y \geq 0, \quad (1.5.8)$$

где $y \in R^n$, двойственной задачей линейного программирования для (1.5.7). Вектор $x \in R^n$, удовлетворяющий системе (1.5.7), называется допустимым решением задачи (1.5.7). Аналогично вводится понятие допустимого решения $y \in R^n$ задачи (1.5.8). Допустимое решение $\bar{x}(\bar{y})$ называется оптимальным решением задачи (1.5.7) [(1.5.8)], если на нем достигается минимум (максимум) функции $sx(by)$ на множестве всех допустимых решений.

Справедливо следующее утверждение [Ашманов, 1981].

Теорема двойственности. Если обе задачи (1.5.7), (1.5.8) имеют допустимые решения, то они обе имеют оптимальные решения \bar{x}, \bar{y} , соответственно, при этом

$$c\bar{x} = b\bar{y}.$$

1.5.5. В заключение параграфа приведем одно свойство выпуклых функций. Сначала напомним, что функция $\varphi : M \rightarrow R^1$, где $M \subset R^m$ — выпуклое множество, называется выпуклой, если

$$\varphi(\lambda x_1 + (1 - \lambda)x_2) \leq \lambda\varphi(x_1) + (1 - \lambda)\varphi(x_2) \quad (1.5.9)$$

для всех $x_1, x_2 \in M$ и $\lambda \in [0, 1]$. Если же в (1.5.9) выполняется обратное неравенство, то функция φ называется вогнутой. Пусть $\varphi_i(x)$ — выпуклые на M функции, $i = 1, \dots, n$. Тогда верхняя огибающая $\psi(x)$ этого семейства функций

$$\psi(x) = \max_{i=1, \dots, n} \varphi_i(x) \quad (1.5.10)$$

является выпуклой на M .

Действительно, по определению выпуклой функции для $x_1, x_2 \in M$ и $\alpha \in [0, 1]$ имеем

$$\begin{aligned} \varphi_i(\alpha x_1 + (1 - \alpha)x_2) &\leq \alpha\varphi_i(x_1) + (1 - \alpha)\varphi_i(x_2) \leq \\ &\leq \alpha \max_i \varphi_i(x_1) + (1 - \alpha) \max_i \varphi_i(x_2). \end{aligned}$$

Отсюда получаем

$$\psi(\alpha x_1 + (1 - \alpha)x_2) = \max_i \varphi_i(\alpha x_1 + (1 - \alpha)x_2) \leq \alpha\psi(x_1) + (1 - \alpha)\psi(x_2),$$

что и требовалось доказать.

Аналогично можно показать вогнутость нижней огибающей (в (1.5.10) берется минимум по i) семейства вогнутых функций.

§ 1.6. Существование решения матричной игры в классе смешанных стратегий

Докажем, что произвольная матричная игра вполне определена в классе смешанных стратегий.

1.6.1. Основная теорема матричных игр. Всякая матричная игра имеет ситуацию равновесия в смешанных стратегиях [Фон Нейман, 1928].

Доказательство. Пусть Γ_A — произвольная $(m \times n)$ -игра со строго положительной матрицей $A = \{a_{ij}\}$, т. е. $a_{ij} > 0$ для всех $i = \overline{1, m}$ и $j = \overline{1, n}$. Покажем, что в этом случае теорема справедлива. Для этого рассмотрим вспомогательную задачу линейного программирования

$$\min xu, \quad xA \geq w, \quad x \geq 0 \quad (1.6.1)$$

и двойственную ей задачу п. (1.5.4)

$$\max yw, \quad Ay \leq u, \quad y \geq 0, \quad (1.6.2)$$

где $u = (1, \dots, 1) \in R^m$, $w = (1, \dots, 1) \in R^n$. Из строгой положительности матрицы A следует, что существует такой вектор $x > 0$, для которого $xA > w$, т. е. задача (1.6.1) имеет допустимое решение. С другой стороны, вектор $y = 0$ является допустимым решением задачи (1.6.2). Поэтому по теореме о двойственности линейного программирования (см. п. 1.5.4) обе задачи (1.6.1) и (1.6.2) имеют оптимальные решения \bar{x}, \bar{y} соответственно, при этом

$$\bar{x}u = \bar{y}w = \Theta > 0. \quad (1.6.3)$$

Рассмотрим векторы $x^* = \bar{x}/\Theta$ и $y^* = \bar{y}/\Theta$ и покажем, что они являются оптимальными стратегиями игроков 1 и 2 соответственно в игре $\bar{\Gamma}_A$, при этом значение игры равно $1/\Theta$.

Действительно, из (1.6.3) имеем

$$x^*u = (\bar{x}u)/\Theta = (\bar{y}w)/\Theta = y^*w = 1,$$

а из допустимости \bar{x} и \bar{y} для задач (1.6.1), (1.6.2), следует, что $x^* = \bar{x}/\Theta \geq 0$ и $y^* = \bar{y}/\Theta \geq 0$, т. е. x^* и y^* — смешанные стратегии игроков 1 и 2 в игре Γ_A .

Вычислим выигрыш игрока 1 в ситуации (x^*, y^*) :

$$K(x^*, y^*) = x^*Ay^* = (\bar{x}A\bar{y})/\Theta^2. \quad (1.6.4)$$

С другой стороны, из допустимости векторов \bar{x} и \bar{y} для задач (1.6.1), (1.6.2) и равенства (1.6.3) имеем

$$\Theta = w\bar{y} \leq (\bar{x}A)\bar{y} = \bar{x}(A\bar{y}) \leq \bar{x}u = \Theta. \quad (1.6.5)$$

Таким образом, $\bar{x}A\bar{y} = \Theta$, из (1.6.4) получаем, что

$$K(x^*, y^*) = 1/\Theta. \quad (1.6.6)$$

Пусть $x \in X$ и $y \in Y$ — произвольные смешанные стратегии игроков 1 и 2. Тогда выполняются неравенства

$$K(x^*, y) = (x^*A)y = (\bar{x}A)y/\Theta \geq (wy)/\Theta = 1/\Theta, \quad (1.6.7)$$

$$K(x, y^*) = x(Ay^*) = x(A\bar{y})/\Theta \leq (xu)/\Theta = 1/\Theta. \quad (1.6.8)$$

Сравнивая (1.6.6)–(1.6.8), получаем, что (x^*, y^*) ситуация равновесия, а $1/\Theta$ — значение игры Γ_A со строго положительной матрицей A .

Теперь рассмотрим $(m \times n)$ -игру $\Gamma_{A'}$ с произвольной матрицей $A' = \{a'_{ij}\}$. Тогда существует такая константа $\beta > 0$, что матрица $A = A' + B$ строго положительна, где $B = \{\beta_{ij}\}$ — $(m \times n)$ -матрица, $\beta_{ij} = \beta$, $i = \overline{1, m}$, $j = \overline{1, n}$. В игре Γ_A существует ситуация равновесия (x^*, y^*) в смешанных стратегиях, а значение игры равно $v_A = 1/\Theta$, где Θ определяется как в (1.6.3).

Из леммы п. 1.4.3 следует, что $(x^*, y^*) \in Z(\bar{\Gamma}_{A'})$ — ситуация равновесия в игре $\Gamma_{A'}$ в смешанных стратегиях, а значение игры равно $v_{A'} = v_A - \beta = 1/\Theta - \beta$. Теорема доказана.

Неформально факт существования решения в классе смешанных стратегий означает, что игроки всегда могут снять неопределенность выбора стратегии, с которой они столкнулись перед началом игры, рандомизируя множество чистых стратегий. Следует отметить, что не всегда в антагонистических играх существует решение в смешанных стратегиях. Примеры таких игр с бесконечным числом стратегий приведены в 2.3, 2.4.

Заметим также, что доказательство теоремы конструктивно, поскольку сводит решение матричной игры к задаче линейного программирования, при этом алгоритм решения игры $\Gamma_{A'}$ следующий.

- 1) По матрице A' строится строго положительная матрица $A = A' + B$, где $B = \{\beta_{ij}\}, \beta_{ij} = \beta > 0$.
- 2) Решаются задачи линейного программирования (1.6.1), (1.6.2). Находятся векторы \bar{x}, \bar{y} и число Θ [см. (1.6.3)].
- 3) Строятся оптимальные стратегии игроков 1 и 2 соответственно,

$$x^* = \bar{x}/\Theta, \quad y^* = \bar{y}/\Theta.$$

- 4) Вычисляется значение игры $\Gamma_{A'}$

$$v_{A'} = 1/\Theta - \beta.$$

Пример 8. Рассмотрим матричную игру Γ_A , определенную матрицей

$$A = \begin{bmatrix} 4 & 0 \\ 2 & 3 \end{bmatrix}.$$

Соответствующие ей задачи линейного программирования имеют следующий вид:

$$\begin{array}{ll} \min (\xi_1 + \xi_2), & \max (\eta_1 + \eta_2), \\ 4\xi_1 + 2\xi_2 \geq 1, & 4\eta_1 \leq 1 \\ 3\xi_2 \geq 1, & 2\eta_1 + 3\eta_2 \leq 1, \\ \xi_1 \geq 0, \xi_2 \geq 0, & \eta_1 \geq 0, \eta_2 \geq 0. \end{array}$$

Заметим, что эти задачи в эквивалентной форме могут быть записаны для ограничений типа равенств:

$$\begin{array}{ll} \min (\xi_1 + \xi_2), & \max (\eta_1 + \eta_2), \\ 4\xi_1 + 2\xi_2 - \xi_3 = 1, & 4\eta_1 + \eta_3 = 1 \\ 3\xi_2 - \xi_4 = 1, & 2\eta_1 + 3\eta_2 + \eta_4 = 1, \\ \xi_1 \geq 0, \xi_2 \geq 0, \xi_3 \geq 0, \xi_4 \geq 0, & \eta_1 \geq 0, \eta_2 \geq 0, \eta_3 \geq 0, \eta_4 \geq 0. \end{array}$$

Таким образом, любой метод решения задач линейного программирования может быть приспособлен для решения матричных игр. Наиболее распространенным методом решения таких задач является симплекс-метод, систематическое изложение которого можно найти в [Ашманов, 1981, Гейл, 1960, Ху, 1974].

§ 1.7. Свойства оптимальных стратегий и значения игры

Рассмотрим свойства оптимальных стратегий, которые в ряде случаев помогают находить значение игры и ситуацию равновесия.

1.7.1. Пусть $(x^*, y^*) \in X \times Y$ — ситуация в смешанных стратегиях в игре Γ_A . Оказывается, что для проверки ситуации (x^*, y^*) на равновесность, неравенства (1.4.7) достаточно проверять не для всех $x \in X$ и $y \in Y$, а лишь для $i \in M$ и $j \in N$, поскольку справедливо следующее утверждение.

Теорема. Для того, чтобы ситуация (x^*, y^*) была равновесной в игре Γ_A , а число $v = K(x^*, y^*)$ — значением игры Γ_A , необходимо и достаточно выполнение следующих неравенств для всех $i \in M$ и $j \in N$:

$$K(i, y^*) \leq K(x^*, y^*) \leq K(x^*, j). \quad (1.7.1)$$

Доказательство. Необходимость. Пусть (x^*, y^*) — ситуация равновесия в игре Γ_A . Тогда

$$K(x, y^*) \leq K(x^*, y^*) \leq K(x^*, y)$$

для всех $x \in X, y \in Y$. Поэтому, в частности, для $u_i \in X$ и $w_j \in Y$ имеем

$$K(i, y^*) = K(u_i, y^*) \leq K(x^*, y^*) \leq K(x^*, w_j) = K(x^*, j)$$

для всех $i \in M$ и $j \in N$.

Достаточность. Пусть (x^*, y^*) — пара смешанных стратегий, для которой выполняются неравенства (1.7.1). Пусть также $x = (\xi_1, \dots, \xi_m) \in X$ и $y = (\eta_1, \dots, \eta_n) \in Y$ — произвольные смешанные стратегии игроков 1 и 2 соответственно. Умножая первое и второе неравенства (1.7.1) на ξ_i и η_j соответственно и суммируя, получаем

$$\sum_{i=1}^m \xi_i K(i, y^*) \leq K(x^*, y^*) \sum_{i=1}^m \xi_i = K(x^*, y^*), \quad (1.7.2)$$

$$\sum_{j=1}^n \eta_j K(x^*, j) \geq K(x^*, y^*) \sum_{j=1}^n \eta_j = K(x^*, y^*). \quad (1.7.3)$$

При этом имеем

$$\sum_{i=1}^m \xi_i K(i, y^*) = K(x, y^*), \quad (1.7.4)$$

$$\sum_{j=1}^n \eta_j K(x^*, j) = K(x^*, y). \quad (1.7.5)$$

Подставляя (1.7.4), (1.7.5) в (1.7.2) и (1.7.3) соответственно и учитывая произвольность стратегий $x \in X$ и $y \in Y$, получаем равновесность ситуации (x^*, y^*) .

Следствие. Пусть (i^*, j^*) — ситуация равновесия в игре Γ_A . Тогда ситуация (i^*, j^*) равновесна и в игре $\bar{\Gamma}_A$.

Пример 10 (Решение игры на уклонение). Предполагается, что игроки выбирают целые числа i и j между 1 и n , а игрок 1 выигрывает величину $a_{ij} = |i - j|$, т. е. расстояние между числами i и j .

Пусть первый игрок придерживается стратегии $x^* = (1/2, 0, \dots, 0, 1/2)$. Тогда

$$K(x^*, j) = 1/2|1 - j| + 1/2|n - j| = 1/2(j - 1) + 1/2(n - j) = (n - 1)/2$$

для всех $1 \leq j \leq n$.

а) Пусть $n = 2k + 1$ — нечетно. Тогда игрок 2 имеет такую чистую стратегию $j^* = (n + 1)/2 = k + 1$, что

$$a_{ij^*} = |i - (n + 1)/2| = |i - k - 1| \leq k = (n - 1)/2$$

для всех $i = 1, 2, \dots, n$.

б) Предположим, что $n = 2k$ — четно. Тогда игрок 2 имеет такую стратегию $y^* = (0, 0, \dots, 1/2, 1/2, 0, \dots, 0)$, где $\eta_k^* = 1/2, \eta_{k+1}^* = 1/2, \eta_j^* = 0, j \neq k + 1, j \neq k$, и

$$\begin{aligned} K(j, y^*) &= 1/2|i - k| + 1/2|i - k - 1| \leq \\ &\leq 1/2k + 1/2(k - 1) = (n - 1)/2, \quad \text{для всех } 1 \leq i \leq n. \end{aligned}$$

Теперь, используя теорему, нетрудно убедиться, что значение игры $v = (n - 1)/2$, игрок 1 имеет оптимальную стратегию x^* , а оптимальная стратегия игрока 2 равна j^* , если $n = 2k + 1$, и y^* , если $n = 2k$.

1.7.2. Приведем результаты, являющиеся непосредственным следствием теоремы п. 1.7.1.

Теорема. Пусть Γ_A — $(m \times n)$ -игра. Для того, чтобы ситуация в смешанных стратегиях (x^*, y^*) была равновесной в игре $\bar{\Gamma}_A$, необходимо и достаточно выполнение равенства

$$\max_{1 \leq i \leq m} K(i, y^*) = \min_{1 \leq j \leq n} K(x^*, j). \quad (1.7.6)$$

Доказательство. Необходимость. Если (x^*, y^*) — ситуация равновесия, то согласно теореме п. 1.7.1 имеем

$$K(i, y^*) \leq K(x^*, y^*) \leq K(x^*, j)$$

для всех $i \in \{1, \dots, m\}$, $j \in \{1, \dots, n\}$. Тогда

$$K(i, y^*) \leq K(x^*, j)$$

для всех i и j . Предположим противное, т. е. (1.7.6) не выполнено. Тогда

$$\max_{1 \leq i \leq m} K(i, y^*) < \min_{1 \leq j \leq n} K(x^*, j).$$

Следовательно, имеют место неравенства

$$\begin{aligned} K(x^*, y^*) &= \sum_{i=1}^m \xi_i^* K(i, y^*) \leq \max_{1 \leq i \leq m} K(i, y^*) < \min_{1 \leq j \leq n} K(x^*, j) \leq \\ &\leq \sum_{j=1}^n \eta_j^* K(x^*, j) = K(x^*, y^*). \end{aligned}$$

Полученное противоречие и доказывает необходимость условия теоремы.

Достаточность. Пусть пара смешанных стратегий (\tilde{x}, \tilde{y}) такова, что $\max_i K(i, \tilde{y}) = \min_j K(\tilde{x}, j)$. Покажем, что в этом случае (\tilde{x}, \tilde{y}) является ситуацией равновесия в игре $\bar{\Gamma}_A$.

Справедливы соотношения

$$\begin{aligned} \min_{1 \leq j \leq n} K(\tilde{x}, j) &\leq \sum_{j=1}^n \tilde{\eta}_j K(\tilde{x}, j) = K(\tilde{x}, \tilde{y}) = \\ &= \sum_{i=1}^m \tilde{\xi}_i K(i, \tilde{y}) \leq \max_{1 \leq i \leq m} K(i, \tilde{y}). \end{aligned}$$

Следовательно, мы имеем

$$K(i, \tilde{y}) \leq \max_{1 \leq i \leq m} K(i, \tilde{y}) = K(\tilde{x}, \tilde{y}) = \min_{1 \leq j \leq n} K(\tilde{x}, j) \leq K(\tilde{x}, j)$$

для всех $1 \leq i \leq m$ и $1 \leq j \leq n$. Тогда по теореме п. 1.7.1, ситуация (\tilde{x}, \tilde{y}) образует равновесие в игре $\bar{\Gamma}_A$.

Из доказательства следует, что любое из чисел в (1.7.6) равно значению игры.

1.7.3. Теорема. Для матричной игры Γ_A справедливо следующее соотношение:

$$\max_x \min_j K(x, j) = v_A = \min_y \max_i K(i, y), \quad (1.7.7)$$

причем экстремумы по смешанным стратегиям x и y в (1.7.7) достигаются на оптимальных стратегиях игроков.

Теорема является следствием теорем п. 1.3.4, 1.7.2 и ее доказательство предоставляем читателю.

1.7.4. Теорема. В матричной игре Γ_A множества оптимальных смешанных стратегий X^* и Y^* игроков являются выпуклыми многогранниками.

Доказательство. По теореме п. 1.7.1, множество X^* является множеством всех решений системы неравенств

$$\begin{aligned} xa^j &\geq v_A, \quad j \in N, \\ xu &= 1, \\ x &\geq 0, \end{aligned}$$

где $u = (1, \dots, 1) \in R^m$, v_A — значение игры. Таким образом, X^* — выпуклое многогранное множество (1.5.1). С другой стороны, $X^* \subset X$, где X — выпуклый многогранник (1.5.3). Следовательно, X^* — ограничено. Следовательно, по теореме п. 1.5.3 множество X^* — выпуклый многогранник.

Аналогично доказывается, что Y^* — выпуклый многогранник.

1.7.5. В качестве примера использования теоремы п. 1.7.3 приведем геометрическое решение игр с двумя стратегиями у одного из игроков $(2 \times n)$ и $(m \times 2)$ -игры. Такой подход в литературе также называется *графоаналитическим методом решения игр*. В основе графоаналитических методов лежит свойство оптимальных стратегий x^* и y^* доставлять внешние экстремумы в равенстве

$$v_A = \max_x \min_j K(x, j) = \min_y \max_i K(i, y).$$

Пример 11. $((2 \times n)$ -игра.) Рассмотрим игру, в которой игрок 1 имеет две стратегии, а игрок 2 — n стратегий. Матрица имеет вид

$$A = \begin{bmatrix} \alpha_{11} & \alpha_{12} & \dots & \alpha_{1n} \\ \alpha_{21} & \alpha_{22} & \dots & \alpha_{2n} \end{bmatrix}.$$

Пусть игрок 1 выбрал смешанную стратегию $x = (\xi, 1 - \xi)$, а игрок 2 чистую стратегию $j \in N$. Тогда выигрыш игрока 1 в ситуации (x, j) равен

$$K(x, j) = \xi \alpha_{1j} + (1 - \xi) \alpha_{2j}. \quad (1.7.8)$$

Геометрически он представляет собой прямую в координатах (ξ, K) . Таким образом, каждой чистой стратегии j соответствует своя прямая. Графиком функции

$$H(\xi) = \min_j K(x, j)$$

является нижняя огибающая семейства прямых (1.7.8). Эта функция вогнута как нижняя огибающая семейства вогнутых (в данном случае линейных) функций (п. 1.5.5).

Точка ξ^* , в которой достигается максимум функции $H(\xi)$ по $\xi \in [0, 1]$, и дает требуемое оптимальное решение $x^* = (\xi^*, 1 - \xi^*)$ и значение игры $v_A = H(\xi^*)$.

Для определенности рассмотрим игру с матрицей

$$A = \begin{bmatrix} 1 & 3 & 1 & 4 \\ 2 & 1 & 4 & 0 \end{bmatrix}.$$

Для каждого $j = 1, 2, 3, 4$ имеем: $K(x, 1) = -\xi + 2$, $K(x, 2) = 2\xi + 1$, $K(x, 3) = -3\xi + 4$, $K(x, 4) = 4\xi$. Нижняя огибающая $H(\xi)$ семейства прямых $\{K(x, j)\}$ и сами прямые $K(x, j)$, $j = 1, 2, 3, 4$ изображены на рис. 1.1. Максимум $H(\xi^*)$ функции $H(\xi)$ находится на пересечении первой и четвертой прямых. Таким образом, ξ^* — решение уравнения

$$4\xi^* = -\xi^* + 2 = v_A.$$

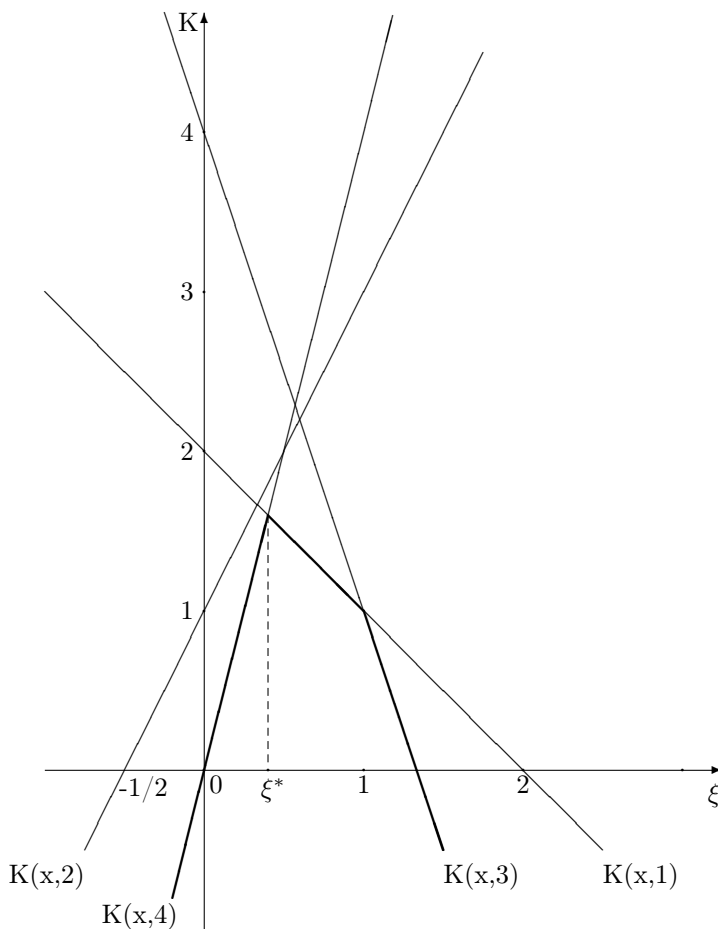


Рис. 1.1. Геометрическое решение $(2 \times n)$ -игры

Откуда получаем оптимальную стратегию $x^* = (2/5, 3/5)$ игрока 1 и значение игры $v_A = 8/5$. Оптимальную стратегию игрока 2 найдем из следующих соображений. Заметим, что в рассматриваемом случае $K(x^*, 1) = K(x^*, 4) = v_A = 8/5$.

Для оптимальной стратегии $y^* = (\eta_1^*, \eta_2^*, \eta_3^*, \eta_4^*)$ должно выполняться равенство

$$v_A = K(x^*, y^*) = \eta_1^* K(x^*, 1) + \eta_2^* K(x^*, 2) + \eta_3^* K(x^*, 3) + \eta_4^* K(x^*, 4).$$

При этом $K(x^*, 2) > 8/5$, $K(x^*, 3) > 8/5$; следовательно, $\eta_2^* = \eta_3^* = 0$, а η_1^*, η_4^* можно найти из условия (1.7.1):

$$\eta_1^* + 4\eta_4^* = 8/5,$$

$$2\eta_1^* = 8/5.$$

Таким образом, $\eta_1^* = 4/5$, $\eta_4^* = 1/5$ и оптимальная стратегия игрока 2 равна $y^* = (4/5, 0, 0, 1/5)$.

Пример 12 (($m \times 2$)-игра). В этом примере две стратегии имеет игрок 2, а игрок 1 имеет m стратегий. Тогда матрица A имеет вид

$$A = \begin{bmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \\ \dots & \dots \\ \alpha_{m1} & \alpha_{m2} \end{bmatrix}.$$

Анализ этой игры проводится аналогично. Действительно, пусть $y = (\eta, 1 - \eta)$ — произвольная смешанная стратегия игрока 2. Тогда выигрыш игрока 1 в ситуации (i, y) равен

$$K(i, y) = \alpha_{i1}\eta + \alpha_{i2}(1 - \eta) = (\alpha_{i1} - \alpha_{i2})\eta + \alpha_{i2}.$$

График функции $K(i, y)$ — прямая. Рассмотрим верхнюю огибающую этих прямых, т. е. функцию

$$H(\eta) = \max_i [(\alpha_{i1} - \alpha_{i2})\eta + \alpha_{i2}].$$

Функция $H(\eta)$ — выпуклая (как верхняя огибающая семейства выпуклых функций).

Точка минимума η^* функции $H(\eta)$ дает оптимальную стратегию $y^* = (\eta^*, 1 - \eta^*)$ и значение игры $v_A = H(\eta^*) = \min_{\eta \in [0,1]} H(\eta)$.

1.7.6. Приведем результат, полезный при отыскании решения игры.

Теорема. Пусть $x^* = (\xi_1^*, \dots, \xi_m^*)$ и $y^* = (\eta_1^*, \dots, \eta_n^*)$ — оптимальные стратегии в игре Γ_A и v_A — значение игры. Тогда для любого i , при котором $K(i, y^*) < v_A$, имеет место равенство $\xi_i^* = 0$, а для любого j такого, что $v_A < K(x^*, j)$ имеет место равенство $\eta_j^* = 0$. Обратно, если $\xi_i^* > 0$, то $K(i, y^*) = v_A$, а если $\eta_j^* > 0$, то $K(x^*, j) = v_A$.

Доказательство. Допустим, что для некоторого $i_0 \in M$, $K(i_0, y^*) < v_A$ и $\xi_{i_0}^* \neq 0$. Тогда получаем, что

$$K(i_0, y^*)\xi_{i_0}^* < v_A\xi_{i_0}^*.$$

Для всех $i \in M$, $K(i, y^*) \leq v_A$, поэтому

$$K(i, y^*)\xi_i^* \leq v_A\xi_i^*.$$

Следовательно, $K(x^*, y^*) < v_A$, что противоречит тому, что v_A — значение игры. Вторая часть теоремы доказывается аналогично.

Этот результат является аналогом теоремы о дополняющей нежесткости [Ху, 1974] или, как ее еще называют, канонической теоремой равновесия для задачи линейного программирования [Гейл, 1960].

Определение. Чистая стратегия $i \in M$ ($j \in N$) игрока 1 (2) называется существенной или активной стратегией, если существует оптимальная стратегия $x^* = (\xi_1^*, \dots, \xi_m^*)$ ($y^* = (\eta_1^*, \dots, \eta_n^*)$) этого игрока, для которой $\xi_i^* > 0$ ($\eta_j^* > 0$).

Из определения и последней теоремы следует, что для каждой существенной стратегии i игрока 1 и любой оптимальной стратегии $y^* \in Y^*$ игрока 2 в игре Γ_A выполняется равенство

$$K(i, y^*) = a_i y^* = v_A.$$

Аналогичное равенство имеет место для любой существенной стратегии $j \in N$ игрока 2 и оптимальной стратегии $x^* \in X^*$ игрока 1

$$K(x^*, j) = a^j x^* = v_A.$$

Если для чистой стратегии $i \in M$ и смешанной стратегии $y \in Y$ выполняется равенство $a_i y = v_A$ то говорят, что стратегия i уравнивает смешанную стратегию y в игре Γ_A .

Таким образом, в данной терминологии теорему можно переформулировать следующим образом. Если чистая стратегия игрока существенна, то она уравнивает любую оптимальную стратегию противника.

Знание спектра оптимальной стратегии упрощает нахождение решения игры. Действительно, пусть M_{x^*} — спектр оптимальной стратегии x^* игрока 1. Тогда каждая оптимальная стратегия $y^* = (\eta_1^*, \dots, \eta_n^*)$ игрока 2 и значение игры v удовлетворяют системе неравенств

$$\begin{aligned} a_i y^* &= v, \quad i \in M_{x^*}, \\ a_i y^* &\leq v, \quad i \in M \setminus M_{x^*}, \\ \sum_{j=1}^n \eta_j^* &= 1, \quad \eta_j^* \geq 0, \quad j \in N. \end{aligned}$$

При этом в спектр M_{x^*} любой оптимальной стратегии x^* могут входить лишь существенные стратегии..

1.7.7. В заключение параграфа приведем аналитическое решение игры «нападение-защита» (см. пример 4 п. 1.1.3)

Пример 13 [Sakaguchi, 1973]. Рассмотрим игру с $(n \times n)$ матрицей A .

$$A = \begin{bmatrix} \beta_1 \tau_1 & \tau_1 & \dots & \tau_1 \\ \tau_2 & \beta_2 \tau_2 & \dots & \tau_2 \\ \dots & \dots & \dots & \dots \\ \tau_n & \tau_n & \dots & \beta_n \tau_n \end{bmatrix}.$$

Здесь $\tau_i > 0$ — ценность, а $0 < \beta_i < 1$ — вероятность поражения объекта C_i , $i = 1, 2, \dots, n$ при условии, что он защищен. Пусть $\tau_1 \leq \tau_2 \leq \dots \leq \tau_n$. Определим функцию φ от целых чисел $1, 2, \dots, n$ следующим образом:

$$\varphi(k) = \left\{ \sum_{i=k}^n (1 - \beta_i)^{-1} - 1 \right\} / \sum_{i=k}^n (\tau_i (1 - \beta_i))^{-1} \quad (1.7.9)$$

и пусть $l \in \{1, 2, \dots, n\}$ — целое число, доставляющее максимум функции $\varphi(k)$, т. е.

$$\varphi(l) = \max_{k=1,2,\dots,n} \varphi(k). \quad (1.7.10)$$

Установим свойства функции $\varphi(k)$. Обозначим символом R один из знаков отношения порядка $\{>, =, <\}$. В этом случае

$$\varphi(k)R\varphi(k+1) \quad (1.7.11)$$

тогда и только тогда, когда

$$\tau_k R \varphi(k), \quad k = 1, 2, \dots, n-1, \quad \tau_0 \equiv 0. \quad (1.7.12)$$

Действительно, из (1.7.9) получаем

$$\frac{\varphi(k)}{\tau_k} \frac{(1-\beta_k)^{-1}}{\sum_{i=k+1}^n (\tau_i(1-\beta_i))^{-1}} + \varphi(k) = \varphi(k+1) + \frac{(1-\beta_k)^{-1}}{\sum_{i=k+1}^n (\tau_i(1-\beta_i))^{-1}}.$$

Тогда имеем

$$\left[\frac{\varphi(k)}{\tau_k} - 1 \right] \frac{(1-\beta_k)^{-1}}{\sum_{i=k+1}^n (\tau_i(1-\beta_i))^{-1}} + \varphi(k) = \varphi(k+1). \quad (1.7.13)$$

Заметим, что коэффициент в (1.7.13), стоящий после квадратных скобок, положительный. Поэтому из (1.7.13) получаем эквивалентность соотношений (1.7.11) и (1.7.12).

Теперь так как $\varphi(l) \geq \varphi(l-1)$ или $\varphi(l) \geq \varphi(l+1)$, (в этом случае $\tau_{l-1} \leq \varphi(l-1)$ или $\tau_l \geq \varphi(l)$), то из соотношений (1.7.10), (1.7.11) имеем неравенство

$$\tau_{l-1} \leq \varphi(l) \leq \tau_l. \quad (1.7.14)$$

Найдем оптимальные стратегии в игре Γ_A . Напомним, что мы предполагаем выполненными неравенства $\tau_1 \leq \tau_2 \leq \dots \leq \tau_n$. Тогда оптимальными смешанными стратегиями $x^* = (\xi_1^*, \dots, \xi_n^*)$ и $y^* = (\eta_1^*, \dots, \eta_n^*)$ игроков 1 и 2 соответственно являются следующие:

$$\xi_i^* = \begin{cases} 0, & i = 1, \dots, l-1, \\ (\tau_i(1-\beta_i))^{-1} / \sum_{j=l}^n (\tau_j(1-\beta_j))^{-1}, & i = l, \dots, n, \end{cases} \quad (1.7.15)$$

$$\eta_j^* = \begin{cases} 0, & j = 1, \dots, l-1, \\ (\tau_j - \varphi(l)) / (\tau_j(1-\beta_j)), & j = l, \dots, n, \end{cases} \quad (1.7.16)$$

а значение игры равно

$$v_A = \varphi(l).$$

Действительно, $\xi_i^* \geq 0, i = 1, 2, \dots, n$ и $\sum_{i=1}^n \xi_i^* = 1$. Из определения $\varphi(l)$ и (1.7.14) получаем, что $\eta_j^* \geq 0, j = 1, 2, \dots, n$ и $\sum_{j=1}^n \eta_j^* = 1$.

Пусть $K(x^*, j)$ — выигрыш игрока 1 в ситуации (x^*, j) , аналогично $K(i, y^*)$ — выигрыш в ситуации (i, y^*) . Подставляя (1.7.15), (1.7.16) в функцию выигрыша и используя предположение о неубывании ценностей объектов, а также (1.7.14), получаем

$$K(x^*, j) = \begin{cases} \sum_{i=l}^n \tau_i \xi_i^* = \varphi(l) + \left(\sum_{j=l}^n (\tau_j(1-\beta_j))^{-1} \right)^{-1} > \varphi(l), & j = \overline{1, l-1}, \\ \sum_{i=l}^n \tau_i \xi_i^* - (1-\beta_j) \tau_j \xi_j^* = \varphi(l), & j = \overline{l, n}, \end{cases}$$

$$K(i, y^*) = \begin{cases} \tau_i \leq \varphi(l), & i = \overline{1, l-1}, \\ \tau_i - \tau_i(1 - \beta_i)\eta_i^* = \varphi(l), & i = \overline{l, n}. \end{cases}$$

Таким образом, для всех $i, j = 1, \dots, n$ выполняются неравенства

$$K(i, y^*) \leq \varphi(l) \leq K(x^*, j).$$

Тогда по теореме п. 1.7.1, x^* и y^* — оптимальные стратегии игроков и $v_A = \varphi(l)$ — значение игры. Игра решена.

§ 1.8. Доминирование стратегий

Сложность решения матричной игры возрастает с увеличением размеров матрицы A . Вместе с тем, в ряде случаев анализ матрицы выигрышей позволяет сделать вывод, что некоторые чистые стратегии не входят в спектр оптимальной стратегии. Это приводит к замене первоначальной матрицы на матрицу выигрышей меньшей размерности.

1.8.1. Определение. Говорят, что стратегия x' игрока 1 доминирует стратегию x'' в $(m \times n)$ -игре Γ_A , если для всех чистых стратегий $j \in \{1, \dots, n\}$ игрока 2 выполняются неравенства

$$x' a^j \geq x'' a^j. \quad (1.8.1)$$

Аналогично, стратегия y' игрока 2 доминирует его стратегию y'' , если для всех чистых стратегий $i \in \{1, \dots, m\}$ игрока 1

$$a_i y' \leq a_i y''. \quad (1.8.2)$$

Если неравенства (1.8.1), (1.8.2) выполняются как строгие, то говорят о *строгом доминировании*. Частным случаем доминирования стратегий является их эквивалентность.

Определение. Будем называть стратегии x' и x'' игрока 1 эквивалентными в игре Γ_A , если для всех $j \in \{1, \dots, n\}$

$$x' a^j = x'' a^j,$$

и обозначать $x' \sim x''$.

Для двух эквивалентных стратегий x' и x'' выполняется (для каждого $y \in Y$) равенство

$$K(x', y) = K(x'', y).$$

Аналогично, стратегии y' и y'' игрока 2 эквивалентны ($y' \sim y''$) в игре Γ_A , если для всех $i \in \{1, \dots, m\}$

$$a_i y' = a_i y''.$$

Отсюда имеем, что для любой смешанной стратегии $x \in X$ игрока 1 выполняется равенство

$$K(x, y') = K(x, y'').$$

Для чистых стратегий введенные определения трансформируются следующим образом. Если чистая стратегия i' игрока 1 доминирует стратегию i'' , а чистая стратегия j' игрока 2 — стратегию j'' того же игрока, то для всех $i = 1, \dots, m$; $j = 1, \dots, n$ выполняются неравенства

$$a_{i'j} \geq a_{i''j}, \quad a_{ij'} \leq a_{ij''}.$$

Это можно записать в векторной форме следующим образом:

$$a_{i'} \geq a_{i''}, \quad a^{j'} \leq a^{j''}.$$

Эквивалентность пар стратегий $i', i'' (i' \sim i'')$ и $j', j'' (j' \sim j'')$ означает выполнение равенства $a_{i'} = a_{i''} (a^{j'} = a^{j''})$.

Определение. Будем говорить, что стратегия $x''(y'')$ игрока 1(2) доминируема, если существует стратегия $x' \neq x'' (y' \neq y'')$ этого игрока, которая доминирует $x''(y'')$. В противном случае стратегия $x''(y'')$ недоминируема.

Аналогично стратегия $x''(y'')$ игрока 1 (2) называется *строго доминируемой*, если существует стратегия $x'(y')$ этого игрока, которая строго доминирует $x''(y'')$, т. е. для всех $j = \overline{1, n} (i = \overline{1, m})$ выполняются неравенства

$$x'a^j > x''a^j, \quad a_i y' < a_i y''.$$

В противном случае говорят, что стратегия $x''(y'')$ игрока 1 (2) недоминируема строго.

1.8.2. Покажем, что игроки могут не использовать доминируемые стратегии. Этот факт устанавливает следующее утверждение.

Теорема. Если в игре $\bar{\Gamma}_A$ стратегия x' одного из игроков доминирует оптимальную стратегию x^* , то стратегия x' также оптимальна.

Доказательство. Пусть, для определенности, x' и x^* — стратегии игрока 1. Тогда в силу доминирования

$$x'a^j \geq x^*a^j$$

для всех $j = \overline{1, n}$. Следовательно, в силу оптимальности стратегии x^* (см. 1.7.3), получаем

$$v_A = \min_j x^*a^j \geq \min_j x'a^j \geq \min_j x^*a^j = v_A$$

для всех $j = \overline{1, n}$. Следовательно, согласно теореме п. 1.7.3, стратегия x' также оптимальна.

Итак, оптимальная стратегия может быть доминируема лишь оптимальной стратегией. С другой стороны, никакая оптимальная стратегия не является строго доминируемой, поэтому игроки не должны использовать строго доминируемые стратегии.

Теорема. Если в игре $\bar{\Gamma}_A$ стратегия x^* одного из игроков оптимальна, то x^* — недоминируема строго.

Доказательство. Пусть, для определенности, x^* — оптимальная стратегия игрока 1. Предположим, что x^* — строго доминируема, т. е. существует такая стратегия $x' \in X$, что

$$x'a^j > x^*a^j, \quad j = 1, 2, \dots, n.$$

Следовательно,

$$\min_j x'a^j > \min_j x^*a^j.$$

Но в силу оптимальности $x^* \in X$ выполняется равенство $\min_j x^*a^j = v_A$. Тогда справедливо строгое неравенство

$$\max_x \min_j xa^j > v_A,$$

что противоречит тому, что v_A — значение игры (1.7.3). Полученное противоречие доказывает теорему.

Понятно, что обратное утверждение, вообще говоря, неверно. Так, в игре с матрицей $\begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}$ 1-я и 2-я чистые стратегии игрока 1 недоминируемы строго, но они неоптимальны.

С другой стороны, интуитивно понятно, что если i -я строка матрицы A (j -й столбец) доминируема, то нет необходимости приписывать ей (ему) положительную вероятность. Таким образом, для нахождения оптимальных стратегий вместо игры Γ_A , достаточно решить подыгру $\Gamma_{A'}$, где A' — матрица, получаемая из матрицы A вычеркиванием доминируемых строк и столбцов.

Прежде чем перейти к точной формулировке и доказательству этого результата, введем понятие *расширения смешанной стратегии* x на i -м месте. Если $x = (\xi_1, \dots, \xi_m) \in X$ и $1 \leq i \leq m+1$, то расширением стратегии x на i -м месте будем называть вектор $\bar{x}_i = (\xi_1, \dots, \xi_{i-1}, 0, \xi_i, \dots, \xi_m) \in R^{m+1}$. Так, расширением вектора $(1/3, 2/3, 1/3)$ на 2-м месте является вектор $(1/3, 0, 2/3, 1/3)$; расширением на 4-м месте — вектор $(1/3, 2/3, 1/3, 0)$; расширением на 1-м месте — вектор $(0, 1/3, 2/3, 1/3)$.

1.8.3. Теорема. Пусть Γ_A — $(m \times n)$ -игра. Предположим, что i -я строка матрицы A доминируема (т. е. доминируема чистая стратегия i первого игрока) и пусть $\Gamma_{A'}$ — игра с матрицей A' , получаемой из A вычеркиванием i -й строки. Тогда справедливы следующие утверждения.

- 1) $v_A = v_{A'}$.
- 2) Всякая оптимальная стратегия y^* игрока 2 в игре $\Gamma_{A'}$ является оптимальной и в игре Γ_A .
- 3) Если x^* — произвольная оптимальная стратегия игрока 1 в игре $\Gamma_{A'}$ и \bar{x}_i^* — расширение стратегии x^* на i -м месте, то \bar{x}_i^* — оптимальная стратегия этого игрока в игре Γ_A .
- 4) Если i -я строка матрицы A строго доминируема, то произвольная оптимальная стратегия \bar{x}^* игрока 1 в игре Γ_A может быть получена из некоторой оптимальной стратегии x^* в игре $\Gamma_{A'}$ расширением на i -м месте.

Доказательство. Не нарушая общности, можно предположить, что доминируемой является последняя m -я строка. Пусть $x = (\xi_1, \dots, \xi_m)$ — смешанная стратегия, которая доминирует строку m . Если $\xi_m = 0$, то из условия доминирования для всех $j = 1, 2, \dots, n$ получаем

$$\sum_{i=1}^m \xi_i \alpha_{ij} = \sum_{i=1}^{m-1} \xi_i \alpha_{ij} \geq \alpha_{mj},$$

$$\sum_{i=1}^{m-1} \xi_i = 1, \quad \xi_i \geq 0, \quad i = 1, \dots, m-1. \quad (1.8.3)$$

В противном случае ($\xi_m > 0$), рассмотрим вектор $x' = (\xi'_1, \dots, \xi'_m)$, где

$$\xi'_i = \begin{cases} \xi_i / (1 - \xi_m), & i \neq m, \\ 0, & i = m. \end{cases} \quad (1.8.4)$$

Компоненты вектора x' неотрицательны, $(\xi'_i \geq 0, i = 1, \dots, m)$ и $\sum_{i=1}^m \xi'_i = 1$. С другой стороны, для всех $j = 1, \dots, n$ имеем

$$\frac{1}{1 - \xi_m} \sum_{i=1}^m \xi_i \alpha_{ij} \geq \alpha_{mj} \frac{1}{1 - \xi_m} \sum_{i=1}^m \xi_i$$

или

$$\frac{1}{1 - \xi_m} \sum_{i=1}^{m-1} \xi_i \alpha_{ij} \geq \alpha_{mj} \frac{1}{1 - \xi_m} \sum_{i=1}^{m-1} \xi_i.$$

Учитывая (1.8.4), получаем

$$\begin{aligned} \sum_{i=1}^{m-1} \xi'_i \alpha_{ij} &\geq \alpha_{mj} \sum_{i=1}^{m-1} \xi'_i = \alpha_{mj}, \quad j = 1, \dots, n, \\ \sum_{i=1}^{m-1} \xi'_i &= 1, \quad \xi'_i \geq 0, \quad i = 1, \dots, m-1. \end{aligned} \quad (1.8.5)$$

Таким образом, всегда из доминирования m -й строки следует, что она не превосходит выпуклую линейную комбинацию остальных $m-1$ строк [(1.8.5)].

Пусть $(x^*, y^*) \in Z(\Gamma_{A'})$ — ситуация равновесия в игре $\Gamma_{A'}$, $x^* = (\xi_1^*, \dots, \xi_{m-1}^*)$, $y^* = (\eta_1^*, \dots, \eta_n^*)$. Для доказательства утверждений 1, 2, 3 теоремы достаточно показать, что $K(\bar{x}_m^*, y^*) = v_{A'}$ и

$$\sum_{j=1}^n \alpha_{ij} \eta_j^* \leq v_{A'} \leq \sum_{i=1}^{m-1} \alpha_{ij} \xi_i^* + 0 \cdot \alpha_{mj} \quad (1.8.6)$$

для всех $i = 1, \dots, m$, $j = 1, \dots, n$.

Первое равенство очевидно, а из оптимальности стратегий (x^*, y^*) в игре $\Gamma_{A'}$ следует выполнение неравенств

$$\sum_{j=1}^n \alpha_{ij} \eta_j^* \leq v_{A'} \leq \sum_{i=1}^{m-1} \alpha_{ij} \xi_i^*, \quad i = \overline{1, m-1}, \quad j = \overline{1, n}. \quad (1.8.7)$$

Из (1.8.7) очевидным образом следует правое из неравенств (1.8.6). Докажем левое неравенство. Для этого достаточно показать, что

$$\sum_{j=1}^n \alpha_{mj} \eta_j^* \leq v_{A'}.$$

Из неравенств (1.8.3), (1.8.5) получаем

$$\sum_{j=1}^n \alpha_{mj} \eta_j^* \leq \sum_{j=1}^n \sum_{i=1}^{m-1} \alpha_{ij} \xi'_i \eta_j^* \leq \sum_{i=1}^{m-1} v_{A'} \xi'_i = v_{A'},$$

что и доказывает первую часть теоремы.

Для доказательства второй части теоремы (утверждение 4) достаточно заметить, что в случае строгого доминирования m -й строки, неравенства (1.8.3), (1.8.5) выполняются как строгие для всех $j = \overline{1, n}$. Поэтому

$$\sum_{j=1}^n \alpha_{mj} \eta_j^* < \sum_{j=1}^n \sum_{i=1}^{m-1} \alpha_{ij} \xi_i' \eta_j^* \leq v_{A'}.$$

Тогда из теоремы п. 1.7.6 получаем, что у любой оптимальной стратегии игрока 1 в игре Γ_A m -я компонента равна нулю. Теорема доказана.

Сформулируем теорему о доминировании для второго игрока, доказательство которой опустим.

Теорема. Пусть Γ_A — $(m \times n)$ - игра. Предположим, что j -й столбец матрицы A доминируем и пусть $\Gamma_{A'}$ — игра с матрицей A' , полученной из матрицы A вычеркиванием j -го столбца. Тогда справедливы следующие утверждения:

- 1) $v_A = v_{A'}$.
- 2) Всякая оптимальная стратегия x^* игрока 1 в игре $\Gamma_{A'}$ является оптимальной и в игре Γ_A .
- 3) Если y^* — произвольная оптимальная стратегия игрока 2 в игре $\Gamma_{A'}$ и \bar{y}_j^* — расширение стратегии y^* на j -м месте, то \bar{y}_j^* — оптимальная стратегия игрока 2 в игре Γ_A .
- 4) Далее, если j -й столбец матрицы A строго доминируем, то произвольная оптимальная стратегия \bar{y}^* игрока 2 в игре Γ_A может быть получена из некоторой оптимальной стратегии y^* в игре $\Gamma_{A'}$ расширением на j -м месте.

1.8.4. Обобщим полученные результаты и подведем итоги. Теоремы п. 1.8.3 дают алгоритм понижения размерности матрицы игры. Так, если строка (столбец) матрицы не больше (не меньше) некоторой выпуклой линейной комбинации остальных строк (столбцов) этой матрицы, то для нахождения решения игры можно эту строку (столбец) вычеркнуть. При этом расширение оптимальных стратегий в игре с усеченной матрицей даст оптимальное решение исходной игры. Если неравенства выполнялись как строгие, то множество оптимальных стратегий в первоначальной игре можно получить расширением множества оптимальных стратегий усеченной игры, в противном случае при такой процедуре оптимальные стратегии можно потерять. Поясним применение данных теорем на примере.

Пример 14. Рассматривается игра с матрицей

$$A = \begin{bmatrix} 2 & 1 & 1 & 0 \\ 2 & 3 & 1 & 3 \\ 3 & 1 & 2 & 0 \\ 0 & 3 & 0 & 6 \end{bmatrix}.$$

Так как 3-я строка a_3 превосходит первую ($a_3 \geq a_1$), то, вычеркивая первую строку, получаем

$$A_1 = \begin{bmatrix} 2 & 3 & 1 & 3 \\ 3 & 1 & 2 & 0 \\ 0 & 3 & 0 & 6 \end{bmatrix}.$$

В этой матрице 1-й столбец a^1 превосходит 3-й столбец a^3 . Поэтому получаем

$$A_2 = \begin{bmatrix} 3 & 1 & 3 \\ 1 & 2 & 0 \\ 3 & 0 & 6 \end{bmatrix}.$$

В последней матрице никакая строка (столбец) не доминируется другой строкой (столбцом). Вместе с тем, 1-й столбец a^1 превосходит выпуклую линейную комбинацию столбцов a^2 и a^3 , так как $a^1 \geq 1/2a^2 + 1/2a^3$, поскольку $3 > 1/2 + 1/2 \cdot 3$, $1 = 1/2 \cdot 2 + 1/2 \cdot 0$, $3 = 0 \cdot 1/2 + 1/2 \cdot 6$. Исключая 1-й столбец, получаем

$$A_3 = \begin{bmatrix} 1 & 3 \\ 2 & 0 \\ 0 & 6 \end{bmatrix}.$$

В этой матрице 1-я строка эквивалентна смешанной стратегии $x = (0, 1/2, 1/2)$, поскольку $1 = 1/2 \cdot 2 + 0 \cdot 1/2$, $3 = 0 \cdot 1/2 + 6 \cdot 1/2$. Таким образом, исключая 1-ю строку, получаем матрицу

$$A_4 = \begin{bmatrix} 2 & 0 \\ 0 & 6 \end{bmatrix}.$$

Оптимальные стратегии x^* и y^* игроков в игре с этой матрицей равны $x^* = y^* = (3/4, 1/4)$, при этом значение v игры равно $3/2$.

Последняя матрица получена вычеркиванием первых двух строк и столбцов, поэтому оптимальными стратегиями игроков в исходной игре являются расширения указанных стратегий на 1-м и 2-м местах, т. е. $\bar{x}_{12}^* = \bar{y}_{12}^* = (0, 0, 3/4, 1/4)$.

§ 1.9. Вполне смешанные и симметричные игры

Знание спектра оптимальной стратегии упрощает нахождение решения игры. В спектр оптимальной стратегии могут входить лишь существенные чистые стратегии игрока. При этом никакая существенная стратегия не является строго доминируемой, что непосредственно следует из теорем §1.8.

1.9.1. Рассмотрим класс игр, в котором знание спектра достаточно для нахождения решения игры.

Определение. Стратегия $x(y)$ игрока 1 (2) называется вполне смешанной, если ее спектр состоит из множества всех стратегий игрока, т. е. $M_x = M(N_y = N)$. Ситуация равновесия (x^*, y^*) называется вполне смешанной, если стратегии x^* и y^* — вполне смешанные. Игра Γ_A называется вполне смешанной, если каждая ситуация равновесия в ней является вполне смешанной.

Следующая теорема утверждает, что вполне смешанная игра имеет единственное решение.

Теорема. Вполне смешанная $(m \times n)$ -игра Γ_A имеет единственную ситуацию равновесия (x^*, y^*) и квадратную матрицу $(m = n)$. Если $v_A \neq 0$, то матрица A невырожденная и

$$x^* = \frac{uA^{-1}}{uA^{-1}u}, \quad (1.9.1)$$

$$y^* = \frac{A^{-1}u}{uA^{-1}u}, \quad (1.9.2)$$

$$v_A = \frac{1}{uA^{-1}u}. \quad (1.9.3)$$

Доказательство. Пусть $x^* = (\xi_1^*, \dots, \xi_m^*) \in X^*$ и $y^* = (\eta_1^*, \dots, \eta_n^*) \in Y^*$ — произвольные оптимальные стратегии игроков, а v_A — значение игры Γ_A . Поскольку Γ_A — вполне смешанная игра, x^* и y^* — вполне смешанные стратегии, которые (и только они) являются решениями систем линейных неравенств п. 1.7.6:

$$xa^j = v_A, \quad xu = 1, \quad x > 0, \quad j = 1, \dots, n, \quad (1.9.4)$$

$$ya_i = v_A, \quad yw = 1, \quad y > 0, \quad i = 1, \dots, m, \quad (1.9.5)$$

где $u = (1, \dots, 1) \in R^m$, $w = (1, \dots, 1) \in R^n$.

Покажем, что решение вполне смешанной игры (x^*, y^*) единственно. Множества X^*, Y^* , заданные (1.9.4) и (1.9.5) являются непустыми выпуклыми многогранниками и, следовательно, имеют крайние точки. Согласно второй из теорем п. 1.5.2 имеем

$$m \leq \text{rank}[a^1, \dots, a^n, u] = \text{rank}[A, u] \leq m, \quad (1.9.6)$$

$$n \leq \text{rank}[a_1, \dots, a_m, w] = \text{rank}[A, w] \leq n. \quad (1.9.7)$$

Теперь из этой же теоремы следует, что множества X^*, Y^* имеют по одной крайней точке и, следовательно, состоят только из них (как выпуклые многогранники, содержащие единственную крайнюю точку). Единственность решения (x^*, y^*) доказана.

Пусть $v_A = 0$. Тогда однородная система

$$xa^j = v_A, \quad j = \overline{1, n}$$

имеет ненулевое решение, откуда $\text{rank}(A) < m$. Поскольку $\text{rank}[A, u] = m$, имеем: $\text{rank}(A) = m - 1$. Аналогично, из (1.9.5) и (1.9.7) следует, что $\text{rank}(A) = n - 1$. Следовательно, $n = m$.

Пусть $v_A \neq 0$. Тогда

$$\text{rank}(A) = \text{rank}[A, v_A u] = \text{rank}[A, u] = m,$$

$$\text{rank}(A) = \text{rank}[A, v_A w] = \text{rank}[A, w] = n.$$

Отсюда получаем $n = m = \text{rank}(A)$, т. е. A — невырожденная матрица. Система уравнений $x^* A = v_A u$ имеет решение

$$x^* = v_A u A^{-1}.$$

Запишем решение системы $A y^* = v_A u$:

$$y^* = A^{-1} v_A u.$$

Тогда

$$v_A = \frac{1}{uA^{-1}u}.$$

Теорема доказана.

Справедливо и обратное утверждение, доказательство которого предоставляем читателю.

Теорема. Пусть в $(m \times n)$ -игре Γ_A матрица A является невырожденной. Тогда, если игрок 2 имеет в Γ_A вполне смешанную оптимальную стратегию, то игрок 1 имеет единственную оптимальную стратегию x^* (1.9.1). Если в игре Γ_A вполне смешанную оптимальную стратегию имеет игрок 1, то игрок 2 имеет единственную оптимальную стратегию y^* (1.9.2), при этом значение игры v_A определено в (1.9.3).

Пример 15 $((2 \times 2)$ – игры). Пусть задана (2×2) -игра с матрицей

$$A = \begin{bmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \end{bmatrix}.$$

Произвольная смешанная стратегия x игрока 1 может быть записана в виде $x = (\xi, 1 - \xi)$, где $0 \leq \xi \leq 1$. Аналогично, смешанная стратегия игрока 2 имеет вид $y = (\eta, 1 - \eta)$, где $0 \leq \eta \leq 1$. Выигрыш в ситуации (x, y) равен

$$K(x, y) = \xi[\alpha_{11}\eta + \alpha_{12}(1 - \eta)] + (1 - \xi)[\alpha_{21}\eta + \alpha_{22}(1 - \eta)].$$

Предположим теперь, что в игре Γ_A нет ситуации равновесия в чистых стратегиях (в противном случае решение просто найти из равенства минимакса и максимина) и пусть $x^* = (\xi^*, 1 - \xi^*)$, $y^* = (\eta^*, 1 - \eta^*)$ — произвольные оптимальные стратегии соответственно первого и второго игроков. Ситуация (x^*, y^*) и игра Γ_A являются вполне смешанными ($\xi^* > 0$ и $\eta^* > 0$). Поэтому по теореме п.1.9.1 в игре существует единственная пара оптимальных смешанных стратегий, которые являются решением системы уравнений

$$\alpha_{11}\eta^* + (1 - \eta^*)\alpha_{12} = v_A,$$

$$\alpha_{21}\eta^* + (1 - \eta^*)\alpha_{22} = v_A,$$

$$\alpha_{11}\xi^* + (1 - \xi^*)\alpha_{21} = v_A,$$

$$\alpha_{12}\xi^* + (1 - \xi^*)\alpha_{22} = v_A.$$

Если добиваться того, чтобы $v_A \neq 0$ (например, если все элементы матрицы A положительны, то это неравенство выполняется), то решение игры

$$v_A = \frac{1}{uA^{-1}u}, \quad x^* = v_A u A^{-1}, \quad y^* = v_A A^{-1} u,$$

где $u = (1, 1)$. Так, легко проверить, что у матрицы

$A = \begin{bmatrix} 1 & 0 \\ -1 & 1 \end{bmatrix}$ не существует седловой точки. Обратная матрица A^{-1} равна

$$A^{-1} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}.$$

Тогда $v_A = 1/3$, $x^* = (2/3, 1/3)$, $y^* = (1/3, 2/3)$.

1.9.2. Исследуем частный класс игр с матрицами специального вида.

Определение. Игра Γ_A с квадратной матрицей A называется симметричной, если матрица A — кососимметричная, т. е. если $a_{ij} = -a_{ji}$ для всех i и j .

В этом случае все диагональные элементы матрицы A равны 0, т. е. $a_{ii} = 0$ при всех i . Для кососимметричной матрицы A всегда выполняется условие $A^T = -A$. Поскольку матрица A квадратная, множества смешанных стратегий игроков совпадают, т. е. $X = Y$.

Докажем теорему о свойствах решения симметричной игры Γ_A , которая полезна при отыскании ситуации равновесия.

Теорема. Пусть Γ_A — симметричная игра. Тогда

$$v_A = 0$$

и множества оптимальных стратегий игроков совпадают, т. е.

$$X^* = Y^*.$$

Доказательство. Пусть A — матрица игры и $x \in X$ — произвольная стратегия. Тогда $xAx = xA^T x = -xAx$. Следовательно, $xAx = 0$.

Пусть $(x^*, y^*) \in Z(A)$ — ситуация равновесия, а v_A — значение игры. Тогда

$$v_A = x^* A y^* \leq x^* A y,$$

$$v_A = x^* A y^* \geq x A y^*$$

для всех $x \in X, y \in Y$. Следовательно,

$$v_A \leq x^* A x^* = 0, \quad v_A \geq y^* A y^* = 0.$$

Откуда получаем $v_A = 0$.

Пусть стратегия x^* оптимальна в игре Γ_A . Тогда (см. теорему п. 1.7.1)

$$x^* A \geq 0.$$

Однако отсюда следует, что $x^*(-A^T) \geq 0$, поэтому $x^* A^T \leq 0$. Таким образом, получаем

$$Ax^* \leq 0.$$

Значит, по той же теореме п. 1.7.1, x^* — оптимальная стратегия игрока 2. Таким образом, доказано, что $X^* \subset Y^*$. Обратное включение доказывается аналогично.

В дальнейшем на основании равенства $X^* = Y^*$, говоря об оптимальной стратегии игрока в симметричной игре, мы не будем указывать, о каком именно игроке идет речь.

Пример 16. Решим игру с матрицей

$$A = \begin{bmatrix} 0 & -1 & 1 \\ 1 & 0 & -1 \\ -1 & 1 & 0 \end{bmatrix}.$$

Пусть $x^* = (\xi_1^*, \xi_2^*, \xi_3^*)$ — оптимальная стратегия в игре Γ_A . Тогда должны выполняться неравенства

$$\begin{aligned} \xi_2^* - \xi_3^* &\geq 0, & -\xi_1^* + \xi_3^* &\geq 0, & \xi_1^* - \xi_2^* &\geq 0, \\ \xi_1^* + \xi_2^* + \xi_3^* &= 1, & \xi_1^* &\geq 0, & \xi_2^* &\geq 0, & \xi_3^* &\geq 0. \end{aligned} \quad (1.9.8)$$

Покажем, что эта игра вполне смешанная. Действительно, пусть $\xi_1^* = 0$. Тогда из системы неравенств (1.9.8) получаем систему

$$\xi_2^* - \xi_3^* \geq 0, \quad \xi_3^* \geq 0, \quad -\xi_2^* \geq 0,$$

$$\xi_1^* + \xi_2^* + \xi_3^* = 1,$$

которая не имеет неотрицательного решения. Аналогичные рассуждения показывают невозможность случаев $\xi_2^* = 0$ и $\xi_3^* = 0$. Поэтому игра Γ_A — вполне смешанная. Следовательно, компоненты $\xi_1^*, \xi_2^*, \xi_3^*$ являются решением системы

$$\xi_2^* - \xi_3^* = 0, \quad -\xi_1^* + \xi_3^* = 0, \quad \xi_1^* - \xi_2^* = 0,$$

$$\xi_1^* + \xi_2^* + \xi_3^* = 1, \quad \xi_i > 0, \quad i = 1, 2, 3.$$

Эта система имеет единственное решение. Оптимальной стратегией является вектор $x^* = (1/3, 1/3, 1/3)$.

Пример 17. Решим дискретную игру типа дуэли с пяти шагов и одним выстрелом у каждого игрока, сформулированную в п. 1.1.4 (см. пример 3). Матрица A выигрышей игрока 1 является симметричной и имеет вид

$$A = \begin{bmatrix} 0 & -3 & -7 & -11 & -15 \\ 3 & 0 & 1 & -2 & -5 \\ 7 & -1 & 0 & 7 & 5 \\ 11 & 2 & -7 & 0 & 15 \\ 15 & 5 & -5 & -15 & 0 \end{bmatrix}.$$

Заметим, что 1-я стратегия каждого игрока (1-я строка и 1-й столбец матрицы) строго доминируема, поэтому она не может быть существенной и ее можно вычеркнуть. В полученной усеченной матрице

$$A' = \begin{bmatrix} 0 & 1 & -2 & -5 \\ -1 & 0 & 7 & 5 \\ 2 & -7 & 0 & 15 \\ 5 & -5 & -15 & 0 \end{bmatrix}$$

не все стратегии являются существенными.

Действительно, из симметричности игры $\Gamma_{A'}$ следует, что $v_{A'} = 0$. Если бы все стратегии были существенными, то оптимальная стратегия x^* была бы решением системы уравнений

$$x^* a'^j = 0, \quad j = 2, 3, 4, 5,$$

$$\sum_{i=2}^5 \xi_i^* = 1,$$

которая решения не имеет.

Перебирая варианты, остановимся на существенной подматрице A'' , составленной из строк и столбцов матрицы A с номерами 2, 3 и 5:

$$A'' = \begin{bmatrix} 0 & 1 & -5 \\ -1 & 0 & 5 \\ 5 & -5 & 0 \end{bmatrix}.$$

Игра с матрицей A'' является вполне смешанной и имеет единственное решение $\bar{y} = \bar{x} = (5/11, 5/11, 1/11)$.

Теперь в исходной игре рассмотрим стратегии $x^* = y^* = (0, 5/11, 5/11, 0, 1/11)$, которые и являются оптимальными.

Таким образом, окончательно имеем: $v_A = 0$, ситуация равновесия (x^*, y^*) — единственная. С точки зрения правил игры получаем, что дуэлянту не следует стрелять на 1-м шаге, он должен стрелять с равной вероятностью после 2-го и 3-го шагов, никогда после 4-го шага и лишь с малой вероятностью стрелять в упор.

§ 1.10. Итеративные методы решения матричных игр

Распространенный способ решения матричной игры путем сведения ее к задаче линейного программирования обладает тем недостатком, что процесс решения задачи линейного программирования существенно усложняется для матриц большой размерности. В таких случаях обычно используют методы декомпозиции задачи линейного программирования, когда вместо решения задачи с исходной матрицей строится координирующая задача с матрицей, у которой мало строк, но много столбцов. На каждой итерации координирующей задачи решается некоторая совокупность вспомогательных задач линейного программирования с матрицами меньших размерностей. К сожалению, декомпозиционные методы эффективны лишь для матриц специального вида (например, блочно-диагональных).

1.10.1. Итеративный метод Брауна — Робинсона (метод фиктивного разыгрывания [Robinson, 1950]). Идея метода — многократное фиктивное разыгрывание игры с заданной матрицей выигрыша. Одно повторение игры будем называть партией. Пусть разыгрывается игра с $(m \times n)$ матрицей $A = \{a_{ij}\}$. В 1-й партии оба игрока выбирают совершенно произвольные чистые стратегии. В k -й партии каждый игрок выбирает ту чистую стратегию, которая максимизирует его ожидаемый выигрыш против наблюдаемого эмпирического вероятностного распределения противника за $(k - 1)$ партий.

Итак, предположим, что за первые k разыгрываний игрок 1 использовал i -ю стратегию ξ_i^k раз ($i = 1, \dots, m$), а игрок 2 использовал j -ю стратегию η_j^k раз ($j = 1, \dots, n$). Тогда в $(k + 1)$ -й партии игрок 1 будет использовать i_{k+1} -ю стратегию, а игрок 2 — свою j_{k+1} -ю стратегию, где

$$\bar{v}^k = \max_i \sum_j a_{ij} \eta_j^k = \sum_j a_{i_{k+1}j} \eta_j^k$$

и

$$\underline{v}^k = \min_j \sum_i a_{ij} \xi_i^k = \sum_i a_{ij_{k+1}} \xi_i^k.$$

Пусть v — значение матричной игры Γ_A . Рассмотрим отношения

$$\bar{v}^k/k = \max_i \sum_j \alpha_{ij} \eta_j^k/k = \sum_j \alpha_{i_{k+1}j} \eta_j^k/k,$$

$$\underline{v}^k/k = \min_j \sum_i \alpha_{ij} \xi_i^k/k = \sum_i \alpha_{ij_{k+1}} \xi_i^k/k.$$

Векторы $x^k = (\xi_1^k/k, \dots, \xi_m^k/k)$ и $y^k = (\eta_1^k/k, \dots, \eta_n^k/k)$ являются смешанными стратегиями игроков 1 и 2 соответственно, поэтому по определению значения игры имеем

$$\max_k \underline{v}^k/k \leq v \leq \min_k \bar{v}^k/k.$$

Таким образом, получен некоторый итеративный процесс, позволяющий находить приближенное решение матричной игры, при этом степень близости приближения к истинному значению игры определяется длиной интервала $[\max_k \underline{v}^k/k, \min_k \bar{v}^k/k]$. Сходимость алгоритма гарантируется теоремой [Robinson, 1950].

Теорема.

$$\lim_{k \rightarrow \infty} (\min_k \bar{v}^k/k) = \lim_{k \rightarrow \infty} (\max_k \underline{v}^k/k) = v.$$

Пример 18. Найти приближенное решение игры с матрицей

$$A = \begin{matrix} & \begin{matrix} a & b & c \end{matrix} \\ \begin{matrix} \alpha \\ \beta \\ \gamma \end{matrix} & \begin{bmatrix} 2 & 1 & 3 \\ 3 & 0 & 1 \\ 1 & 2 & 1 \end{bmatrix} \end{matrix}.$$

Обозначим как α, β, γ стратегии игрока 1 и a, b, c — стратегии игрока 2. Пусть сначала игроки выбрали стратегии α и a , соответственно. Если игрок 1 выбрал стратегию α , то игрок 2 может получить один из выигрышей (2, 1, 3). Если игрок 2 выбрал стратегию a , то игрок 1 может получить один из выигрышей (2, 3, 1). Во 2-й и 3-й партиях игрок 1 выбирает стратегию β , а игрок 2 выбирает b , поскольку эти стратегии обеспечивают наилучший результат и т. д.

В табл. 1.1 приведены результаты разыгрываний, в этой таблице указаны стратегия игрока, накопленный выигрыш и средний выигрыш.

Таким образом, за 12 партий мы получили приближение решения

$$x^{12} = (1/4, 1/6, 7/12), \quad y^{12} = (1/12, 7/12, 1/3),$$

а точность может быть оценена числом $4/12 = 1/3$. Основным недостатком рассмотренного метода является его малая скорость сходимости, которая уменьшается с ростом размерности матрицы. Это является также следствием немонотонности последовательностей \bar{v}^k/k и \underline{v}^k/k .

Рассмотрим другой итеративный алгоритм, который избавлен от указанного недостатка.

1.10.2. Монотонный итеративный алгоритм решения матричных игр [Садовский, 1978]. Рассмотрим смешанное расширение $\Gamma_A = (X, Y, K)$ матричной игры с $(m \times n)$ -матрицей A .

Обозначим как $x^N = (\xi_1^N, \dots, \xi_m^N) \in X$ приближение оптимальной стратегии первого игрока на N -й итерации и $c^N \in R^N$, $c^N = (\gamma_1^N, \dots, \gamma_n^N)$ — вспомогательный вектор. Алгоритм позволяет находить (точно и приближенно) оптимальную стратегию игрока 1 и значение игры v .

В начале процесса игрок 1 выбирает произвольную чистую стратегию i_0 , т. е. $x^0 = (0, \dots, 1, \dots, 0) = u_{i_0}$ и вектор $c^0 = a_{i_0}$, где a_{i_0} — строка матрицы A , имеющая номер i_0 .

Итеративный процесс строится следующим образом. Пусть выполнена $N-1$ -я итерация и получены векторы x^{N-1}, c^{N-1} . Тогда x^N и c^N вычисляются по следующим итеративным формулам:

$$x^N = (1 - \alpha_N)x^{N-1} + \alpha_N \tilde{x}^N, \quad (1.10.1)$$

$$c^N = (1 - \alpha_N)c^{N-1} + \alpha_N \tilde{c}^N, \quad (1.10.2)$$

где параметр $0 \leq \alpha_N \leq 1$. Векторы \tilde{x}^N и \tilde{c}^N будут получены ниже.

Таблица 1.1. Выигрыши игроков

Номер партии	Выбор игрока	Выбор игрока	Выигрыш			Проигрыш			$\frac{\overline{v}^k}{k}$	$\frac{\underline{v}^k}{k}$
			игрока 1			игрока 2				
			α	β	γ	a	b	c		
1	α	a	2	3	1	2	1	3	3	1
2	β	b	3	3	3	5	1	4	3/2	1/2
3	β	b	4	3	5	8	1	5	5/3	1/3
4	γ	b	5	3	7	9	3	6	7/4	3/4
5	γ	b	6	3	9	10	5	7	9/5	5/5
6	γ	b	7	3	11	11	7	8	11/6	7/6
7	γ	b	8	3	13	12	9	9	13/7	9/7
8	γ	c	11	4	14	13	11	10	14/8	10/8
9	γ	c	14	5	15	14	13	11	15/9	11/9
10	γ	c	17	6	16	15	15	12	17/10	12/10
11	α	c	20	7	17	17	16	15	20/11	15/11
12	α	b	21	7	19	19	17	18	21/12	17/12

Рассмотрим вектор $c^{N-1} = (\gamma_1^{N-1}, \dots, \gamma_n^{N-1})$ и выберем такие индексы j_k , на которых достигается минимум

$$\min_{j=1, \dots, n} \gamma_j^{N-1} = \gamma_{j_1}^{N-1} = \gamma_{j_2}^{N-1} = \dots = \gamma_{j_k}^{N-1}.$$

Обозначим через

$$\underline{v}^{N-1} = \min_{j=1, \dots, n} \gamma_j^{N-1} \tag{1.10.3}$$

и $J^{N-1} = \{j_1, \dots, j_k\}$ множество индексов, на которых (1.10.3) достигается.

Пусть $\Gamma^N \subset \Gamma_A$ — подыгра игры Γ_A с матрицей $A^N = \{a_{ij}^{N-1}\}$, $i = 1, \dots, m$, а индекс $j^{N-1} \in J^{N-1}$. Решаем подыгру и находим оптимальную стратегию $\tilde{x}^N \in X$ игрока 1. Пусть $\tilde{x}^N = (\tilde{\xi}_1^N, \dots, \tilde{\xi}_m^N)$.

Вычислим вектор $\tilde{c}^N = \sum_{i=1}^m \tilde{\xi}_i^N a_i$. Пусть вектор \tilde{c}^N имеет компоненты $\tilde{c}^N = (\tilde{\gamma}_1^N, \dots, \tilde{\gamma}_n^N)$. Рассмотрим $(2 \times n)$ -игру с матрицей

$$\begin{bmatrix} \gamma_1^{N-1} & \dots & \gamma_n^{N-1} \\ \tilde{\gamma}_1^N & \dots & \tilde{\gamma}_n^N \end{bmatrix}.$$

Найдем оптимальную стратегию $(\alpha_N, 1 - \alpha_N)$, $0 \leq \alpha_N \leq 1$ в этой подыгре.

Подставляя найденные значения $\tilde{x}^N, \tilde{c}^N, \alpha_N$ в (1.10.1), (1.10.2), находим x^N и c^N . Процесс продолжаем до тех пор, пока не выполнится равенство $\alpha_N = 0$ или не будет достигнута требуемая точность вычислений. Сходимость алгоритма гарантируется следующей теоремой [Садовский, 1978].

Теорема. Пусть $\{x^N\}, \{\underline{v}^N\}$ — итеративные последовательности, определяемые (1.10.1), (1.10.3). Тогда справедливы следующие утверждения:

1) $\underline{v}^N > \underline{v}^{N-1}$, т. е. последовательность $\{\underline{v}^{N-1}\}$ строго монотонно возрастает;

2)

$$\lim_{N \rightarrow \infty} \underline{v}^N = \underline{v} = v; \quad (1.10.4)$$

3) $\lim_{N \rightarrow \infty} x^N = x^*$, где $x^* \in X^*$ — оптимальная стратегия игрока 1.

Пример 19. Решим, используя монотонный алгоритм, игру с матрицей

$$A = \begin{bmatrix} 2 & 1 & 3 \\ 3 & 0 & 1 \\ 1 & 2 & 1 \end{bmatrix}.$$

Итерация 0. Пусть игрок 1 выбрал 1-ю строку матрицы A , т. е. $x^* = (1, 0, 0)$ и $c^0 = a_1 = (2, 1, 3)$. Вычислим $\underline{v}^0 = \min_j \gamma_j^0 = \gamma_2^0 = 1$, $J^0 = 2$.

Итерация 1. Рассмотрим подыгру $\Gamma^1 \subset \Gamma$ с матрицей

$$A^1 = \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix}.$$

Оптимальной стратегией \tilde{x}^1 игрока 1 является вектор $\tilde{x}^1 = (0, 0, 1)$. Тогда $\tilde{c}^1 = a_3 = (1, 2, 1)$. Решаем (2×3) -игру с матрицей $\begin{bmatrix} 2 & 1 & 3 \\ 1 & 2 & 1 \end{bmatrix}$. Заметим, что 3-й столбец матрицы доминируем, поэтому рассмотрим матрицу $\begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$. В силу симметрии оптимальной стратегией игрока 1 в этой игре является вектор $(\alpha_N, 1 - \alpha_N) = (1/2, 1/2)$.

Вычисляем x^1 и c^1 по формулам (1.10.1), (1.10.2). Имеем

$$\begin{aligned} x^1 &= 1/2 x^0 + 1/2 \tilde{x}^1 = (1/2, 0, 1/2), \\ c^1 &= 1/2 c^0 + 1/2 \tilde{c}^1 = (3/2, 3/2, 2), \\ \underline{v}^1 &= \min_j \gamma_j^1 = \gamma_1^1 = \gamma_2^1 = 3/2 > \underline{v}^0 = 1. \end{aligned}$$

Множество индексов имеет вид $J^1 = \{1, 2\}$.

Итерация 2. Рассмотрим подыгру $\Gamma^2 \subset \Gamma$ с матрицей

$$A = \begin{bmatrix} 2 & 1 \\ 3 & 0 \\ 1 & 2 \end{bmatrix}.$$

Первая строка в этой матрице доминируема, поэтому достаточно рассмотреть подматрицу

$$\begin{bmatrix} 3 & 0 \\ 1 & 2 \end{bmatrix}.$$

Оптимальной стратегией игрока 1 в этой игре является вектор $(1/4, 3/4)$, поэтому $\tilde{x}^2 = (0, 1/4, 3/4)$.

Вычислим $\tilde{c}^2 = 1/4a_2 + 3/4a_3 = (3/2, 3/2, 1)$ и рассмотрим (2×3) -игру с матрицей

$$\begin{bmatrix} 3/2 & 3/2 & 1 \\ 3/2 & 3/2 & 2 \end{bmatrix}.$$

Вторая стратегия игрока 1 доминирует первую, поэтому $\alpha_2 = 0$. Таким образом, вычисления закончены $x^* = x^1 = (1/2, 0, 1/2)$; значение игры равно $v = v^1 = 3/2$, а оптимальная стратегия игрока 2 имеет вид $y^* = (1/2, 1/2, 0)$ (см. пример 18).

§ 1.11. Упражнения и задачи

1. Каждый из двух игроков показывает другому m пальцев на руке, $(1 \leq m \leq n, n \leq 5)$ и одновременно называет число пальцев, которое, по его мнению, может показать противник. Если один игрок угадывает правильно, а другой неправильно, то тот, который угадал, выигрывает сумму, равную числу пальцев, показанных обоими игроками. Во всех остальных случаях выигрыши обоих игроков считаются нулевыми.

(а) Сколько стратегий имеет каждый игрок при $n = 3$?

(б) Построить матрицу игры для $n = 2$.

2. *Распределение поисковых усилий.* В одной из n ячеек игрок 2 прячет предмет. Игрок 1 имеет в распоряжении r ищущих, которые должны быть распределены по ячейкам для поиска предмета. Например, в первую ячейку могут быть направлены $(r - 1)$ ищущих, один — во вторую ячейку, а в остальные ячейки — ни одного и т. п.

Предполагается, что известна вероятность обнаружения предмета в i -й ячейке (если он там находится) при поиске одним ищущим. Обнаружение предмета каждым из ищущих — независимые события.

Выигрыш игрока 1 — вероятность обнаружения предмета при заданном распределении ищущих.

(а) Вычислить число m чистых стратегий игрока 1.

(б) Построить матрицу игры.

3. *Поиск многих предметов.* Игрок 2 прячет m черных шаров в n урнах. Общее количество шаров (черных и белых), находящихся в j -й урне, равно $l_j, j = 1, \dots, n, l_j > m$.

Противник (игрок 1) старается обнаружить максимальное число черных шаров, имея возможность проверить одну из урн. При проверке i -й урны игрок 1 наугад (равновероятно) выбирает m шаров из l_i и его выигрыш равен математическому ожиданию количества черных шаров в выборке из m шаров.

(а) Пусть в i -й урне спрятаны p_i черных шаров. Вычислить вероятность β_{ij} того, что выбранная из i -й урны группа r шаров содержит ровно j черных.

(б) Построить матрицу игры.

4. *Противовоздушная оборона.* В системе ПВО объекта могут применяться три типа средств поражения воздушной цели (1, 2, 3), которые должны быть распределены между двумя стартовыми установками. У противника (игрока 2) имеется два типа самолетов (тип 1 и тип 2). Вероятности поражения самолетов одним средством сведены в матрицу

$$\begin{array}{cc} & \begin{matrix} 1 & 2 \end{matrix} \\ \begin{matrix} 1 \\ 2 \\ 3 \end{matrix} & \begin{bmatrix} 0.3 & 0.5 \\ 0.5 & 0.3 \\ 0.1 & 0.6 \end{bmatrix} \end{array}.$$

Предполагается, что возможно нападение только одним из самолетов.

Выигрыш игрока 1 — вероятность поражения самолета системой ПВО.

(а) Построить матрицу игры.

(б) Выяснить, имеется ли решение в чистых стратегиях.

5. Найти ситуации равновесия и значения следующих игр:

$$(a) \begin{bmatrix} 3 & 5 \\ 3 & 2 \end{bmatrix}; \quad (b) \begin{bmatrix} 1/2 & 0 & 1/2 \\ 1 & 3/2 & 2 \\ 0 & -1 & 7/4 \end{bmatrix}.$$

6. Проверить, что $v = 2$ и пара (x^*, y^*) , где $x^* = (0, 0, 1)$, $y^* = (2/5, 3/5, 0)$ — соответственно значение и ситуация равновесия в игре с матрицей $\begin{bmatrix} 3 & -2 & 4 \\ -1 & 4 & 2 \\ 2 & 2 & 6 \end{bmatrix}$.

7. Пусть $A'(A'')$ — подматрица матрицы A , получающаяся вычеркиванием ряда строк (столбцов) A . Показать, что выполняются неравенства $v_{A'} \leq v_A \leq v_{A''}$, где $v_{A'}, v_{A''}$ — значения игр $\Gamma_{A'}, \Gamma_{A''}$ соответственно.

8. Рассматривается игра $\Gamma_{A'}$ с матрицей $\begin{bmatrix} -1 & 3 & -3 \\ 2 & 0 & 3 \\ 2 & 1 & 0 \end{bmatrix}$. Значение игры $v_A = 1$ и оптимальная стратегия игрока 1 есть $x^* = (1/3, 2/3, 0)$. Найти оптимальную стратегию y^* игрока 2.

9. Решить графически игру с матрицей $\begin{bmatrix} -4 & 0 \\ 3 & -2 \\ 5 & -3 \\ -1 & -1 \end{bmatrix}$.

10. Показать, что строго доминируемая стратегия не может быть существенной.

11. Показать, что 3-я строка матрицы A доминируема, где $A = \begin{bmatrix} 20 & 0 \\ 0 & 8 \\ 4 & 5 \end{bmatrix}$.

12. Показать, что выбор 1-го столбца эквивалентен смешанной стратегии $y = (0, 1/3, 2/3)$, где матрица игры имеет вид $\begin{bmatrix} 1 & 3 & 0 \\ 2 & 0 & 3 \end{bmatrix}$.

13. Используя понятие доминирования, найти решение игры с матрицей $\begin{bmatrix} 1 & 7 & 2 \\ 6 & 2 & 7 \\ 5 & 1 & 6 \end{bmatrix}$.

14. Доказать теорему п. 1.7.3.

15. Решить игру поиска с одной попыткой. Игрок 2 прячет предмет в одну из n ячеек. Игрок 1 ищет его в одной из этих ячеек, при этом вероятность обнаружения предмета в i -й ячейке равна $\beta_i > 0$, $i = 1, \dots, n$ (при условии, что он там находится). Показать, что рассматриваемая игра вполне смешанная. Найти решение игры.

16. Решить игру дискретного поиска (пример 5, п. 1.1.3) в предположении $\alpha\beta_i - \tau_i \neq 0$, $i = 1, \dots, n$.

Указание. Воспользоваться результатом п. 1.7.7.

17. *Игра поиска двух предметов.* Игрок 2 прячет два предмета в n ячейках (можно оба в одной ячейке). Цель игрока 1 — обнаружить хотя бы один предмет, при этом он имеет возможность проверить одну ячейку ($\beta_i > 0$ — вероятность обнаружения одного предмета в i -й ячейке при условии, что он там находится). Если в i -й ячейке находятся одновременно два предмета, то вероятность их одновременного обнаружения равна β_i^2 . Таким образом, матрица $A = \{\alpha_{ka}\}$, $a = (i, j)$, $j = 1, \dots, n$, имеет вид

$$\begin{aligned} \alpha_{ka} &= 0, & i &= j, \quad i \neq k, \\ \alpha_{ka} &= \beta_i, & i &= k, \quad i \neq j, \\ \alpha_{ka} &= \beta_j, & j &= k, \quad i \neq j, \\ \alpha_{ka} &= \beta_i(2 - \beta_i), & i &= j = k. \end{aligned}$$

Решить игру.

18. Решить игру поиска многих предметов (см. упр. 3).

19. *Игра поиска нескольких множеств на плоскости.* Заданы набор n фиксированных компактных выпуклых множеств $K_1, K_2, \dots, K_n \subset R^2$ и система m конгруэнтных между собой

компактных выпуклых множеств $T_1, \dots, T_m \subset R^2$. Дискретная одновременная игра поиска заключается в следующем. Игрок 2 прячет m множеств T_j ($j = 1, \dots, m$) в n множествах K_i ($i = 1, \dots, n$) таким образом, что они пересекают K_i . Тот факт, что p_i множеств спрятаны в K_i , означает, что совокупность множеств $\{T_j\}$ в количестве p_i единиц бросается на плоскость случайно. Чистая стратегия a игрока 2 имеет вид

$$a = (p_1, p_2, \dots, p_n) \in R^n, \sum_{i=1}^n p_i = m,$$

где p_i — количество множеств T_j , спрятанных в множестве K_i .

Игрок 1 может проверить одно из множеств K_i , бросая случайно в K_i точку x . Выигрыш игрока 1 — математическое ожидание числа множеств $\{T_j\}$, которым принадлежит x .

Найти решение игры.

20. *Игра поиска с двумя попытками у ищущего.* Игрок 2 прячет предмет в одной из n ячеек, а игрок 1 (ищущий) производит поиск в одной из этих ячеек, имея возможность просмотреть две ячейки (повторный просмотр ячеек не допускается). Множество чистых стратегий игрока 1 состоит из несовпадающих пар (i, j) , $i = 1, \dots, n$, $j = 1, \dots, n$, $i \neq j$ и содержит C_n^2 элементов. Множество чистых стратегий игрока 2 определяется индексом k , $k = 1, \dots, n$ и содержит n элементов. Матрица выигрышей имеет вид $B = \{\beta_{(i,j)k}\}$, где

$$\beta_{(i,j)k} = \begin{cases} \delta_k, & \text{если } i = k \text{ или } j = k, (\delta_k > 0), \\ 0, & \text{в противном случае.} \end{cases}$$

Решить игру.

21. В игре поиска с двумя попытками у ищущего рассмотреть случай, когда множество чистых стратегий игрока 1 состоит из всевозможных пар (i, j) и содержит n^2 элементов. Решить игру в предположении

$$\sum_{k=1}^{n-1} \delta_n / \delta_k < 1.$$

22. В игре на уклонение (п. 1.7.1) показать, что игрок 1 всегда имеет единственную оптимальную стратегию.

23. *Простейшая игра поиска.* В простейшей игре поиска участвуют два игрока. Игрок 2 прячет предмет в одном из n ящиков. Противник (игрок 1) старается обнаружить спрятанный предмет. Если предмет спрятан в i -м ящике, а игрок 1 (ищущий) ищет его в i -м ящике, то вероятность обнаружения объекта (выигрыш игрока 1) равна σ_i , причем $0 < \sigma_1 < \sigma_2 < \dots < \sigma_n < 1$. Если же предмет спрятан в i -м ящике, а игрок 1 ищет его в j -м ящике, то вероятность обнаружения объекта равна 0.

(а) Построить матрицу игры.

(б) Найти решение игры в классе вполне смешанных стратегий.

Глава 2

Бесконечные антагонистические игры

§ 2.1. Бесконечные игры

2.1.1. В этой главе рассматриваются антагонистические игры, которые отличаются от матричных тем, что в них один или оба игрока имеют бесконечное (счетное или континуум) множество стратегий. С теоретико-игровой точки зрения это отличие малозначительно, поскольку игра остается антагонистической и проблема состоит в использовании более сложного аналитического аппарата исследования.

Таким образом, будем исследовать общие антагонистические игры, т. е. системы вида

$$\Gamma = (X, Y, H), \quad (2.1.1)$$

где X и Y — произвольные бесконечные множества, элементы которых являются стратегиями игроков 1 и 2 соответственно, а $H : X \times Y \rightarrow R^1$ — функция выигрыша игрока 1. Напомним, что правила антагонистической игры изложены в п. 1.1.1. Выигрыш игрока 2 в ситуации (x, y) равен $[-H(x, y)]$, $x \in X, y \in Y$ (игра антагонистическая). В этой главе будем рассматривать такие игры, у которых функция H ограничена.

2.1.2. Пример 1 (Одновременная игра преследования на плоскости). Пусть S_1 и S_2 — множества на плоскости. Игра Γ заключается в следующем. Игрок 1 выбирает некоторую точку $x \in S_1$, а игрок 2 выбирает точку $y \in S_2$. При совершении выбора игроки 1 и 2 не имеют информации о действиях противника, поэтому подобный выбор удобно интерпретировать как одновременный. В этом случае точки $x \in S_1, y \in S_2$ являются стратегиями игроков 1 и 2 соответственно. Таким образом, множества стратегий игроков совпадают с множествами S_1 и S_2 на плоскости.

Целью игрока 2 является минимизация расстояния между ним и игроком 1 (игрок 1 преследует противоположную цель). Поэтому под выигрышем $H(x, y)$ игрока 1 в этой игре будем понимать евклидово расстояние $\rho(x, y)$ между точками $x \in S_1$ и $y \in S_2$, т. е. $H(x, y) = \rho(x, y)$, $x \in S_1, y \in S_2$. Выигрыш игрока 2 полагаем равным выигрышу игрока 1, взятому с обратным знаком, а именно $[-\rho(x, y)]$ (игра антагонистическая).

Пример 2 (Поиск на отрезке) [Дюбин, Суздаль, 1981]. Простейшей игрой поиска с бесконечным числом стратегий является следующая игра.

Игрок 2 (прячущийся) выбирает точку $y \in [0, 1]$, а игрок 1 (ищущий) выбирает одновременно и независимо точку $x \in [0, 1]$. Точка y считается «обнаруженной», если $|x - y| \leq l$, где $0 < l < 1$. В этом случае игрок 1 выигрывает величину $+1$; во всех остальных случаях его выигрыш полагается равным 0. Игра антагонистическая.

Таким образом, функция выигрыша имеет вид

$$H(x, y) = \begin{cases} 1, & \text{если } |x - y| \leq l, \\ 0, & \text{если } |x - y| > l. \end{cases}$$

Выигрыш игрока 2 полагается равным $[-H(x, y)]$.

Пример 3 (Поиск на сфере.) Пусть в R^3 задана сфера C радиуса R . Игрок 1 (ищущий) выбирает систему из точек $x_1, x_2, \dots, x_s \in C$, а игрок 2 — одну точку $y \in C$. Выборы точек осуществляются игроками одновременно и независимо друг от друга. Игрок 2 считается обнаруженным, если точка $y \in C$ оказывается в r -окрестности одной из точек x_j , $j = 1, \dots, s$. Здесь под r -окрестностью точки x_j будем понимать сферический сегмент с вершиной в точке x_j и радиусом основания r (рис. 2). В дальнейшем r -окрестность точки x_j будем обозначать через $S(x_j, r)$.

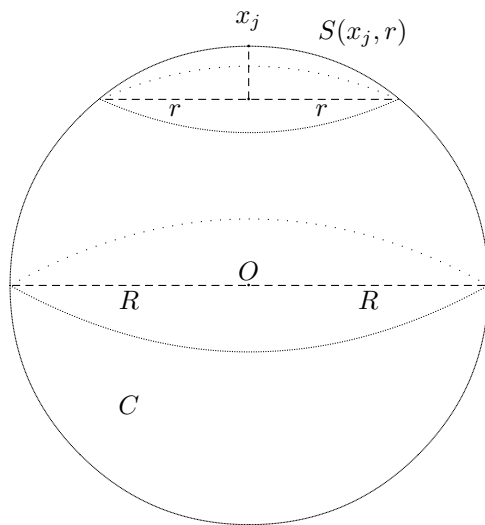


Рис. 2.1. Поиск на сфере

Целью игрока 1 является обнаружение игрока 2. Игрок 2 преследует противоположную цель. В соответствии с этим положим выигрыш игрока 1 равным

$$H(x, y) = \begin{cases} 1, & \text{если } y \in M_x, \\ 0, & \text{в противном случае,} \end{cases}$$

где $x = (x_1, \dots, x_s)$ и $M_x = \cup_{j=1}^s S(x_j, r)$. Выигрыш игрока 2 полагается равным $[-H(x, y)]$.

Пример 4 (Шумная дуэль) [Карлин, 1964]. Каждому из двух дуэлянтов разрешается выстрелить только один раз. Предполагается, что оба они имеют «шумные»

пистолеты, так что каждый знает, когда выстрелил его противник. Предполагается также, что функция меткости $p_1(x)$ (вероятность попадания при стрельбе в момент времени x) игрока 1 определена на $[0, 1]$, непрерывна, монотонно возрастает по x и $p_1(0) = 0$, $p_1(1) = 1$. Аналогично, точность выстрела игрока 2 описывается функцией $p_2(y)$ на $[0, 1]$, где $p_2(0) = 0$, $p_2(1) = 1$. Если игрок 1 поражает игрока 2, то первый получает выигрыш $+1$. Если игрок 2 поражает игрока 1, то игрок 1 получает -1 , если оба игрока стреляют одновременно и с одинаковым результатом (успешным или нет), то выигрыш игрока 1 равен 0.

Структура информации в этой игре (тот факт, что оружие шумное) принимается во внимание при составлении функции выигрыша $H(x, y)$. Если $x < y$, то вероятность того, что игрок 1 поразит противника, равна $p_1(x)$ и вероятность того, что игрок 1 промахнется, равна $1 - p_1(x)$. Если игрок 2 еще не стрелял и знает, что игрок 1 больше не может выстрелить, то игрок 2 будет увеличивать свои шансы на успех, ожидая, пока y не станет равным 1. Таким образом, если игрок 1 промахнется в момент x , то он наверняка будет поражен игроком 2, если $x < y$; следовательно,

$$H(x, y) = p_1(x) + (-1)[1 - p_1(x)], \quad x < y.$$

Аналогично имеем

$$H(x, y) = p_2(y)(-1) + [1 - p_2(y)] \cdot 1, \quad x > y$$

и

$$H(x, y) = p_1(x)[1 - p_2(y)] + p_2(y)[1 - p_1(x)](-1), \quad x = y.$$

Таким образом, функция выигрыша $H(x, y)$ в игре равна

$$H(x, y) = \begin{cases} 2p_1(x) - 1, & x < y, \\ p_1(x) - p_2(y), & x = y, \\ 1 - 2p_2(y), & x > y, \end{cases}$$

где $x \in [0, 1]$, $y \in [0, 1]$.

Пример 5 (Бесшумная дуэль) [Карлин, 1964]. Снова каждому из дуэлянтов разрешается выстрелить только один раз, но в этом случае ни один из дуэлянтов не может определить, выстрелил его противник или нет.

Предположим для простоты, что функции меткости заданы следующим образом: $p_1(x) = p_2(x) = x$. Тогда функция выигрыша, описывающая игру, имеет вид

$$H(x, y) = \begin{cases} x - (1 - x)y, & x < y, \\ 0, & x = y, \\ -y + (1 - y)x, & x > y, \end{cases}$$

$x \in [0, 1]$, $y \in [0, 1]$. Построение функции выигрыша $H(x, y)$ в этой игре производится так же, как и в примере 4, за исключением того, что в данном случае ни один из игроков не может определить момента выстрела противника, если только этот выстрел не оказался успешным.

Пример 6 (Поиск «шумного» объекта). Рассматривается задача поиска «шумного» объекта (игрок 2) подвижным средством обнаружения (игрок 1). Дальность действия

$l(x, y)$, средства обнаружения в зависимости от скоростей $x \in [x_0, x_1]$ и $y \in [y_0, y_1]$ игроков 1 и 2 соответственно имеет вид

$$l(x, y) = \bar{l}(y) \frac{(x_1 - x)}{(x_1 - x_0)},$$

где $\bar{l}(y) = l_0 + \beta(y - y_0)$,

$$\beta = \frac{(l_1 - l_0)}{(y_1 - y_0)},$$

$l_1 = \bar{l}(y_1), l_0 = \bar{l}(y_0)$. Положительные числа l_1, l_0 считаются заданными. Таким образом,

$$l(x, y) = \frac{l_0(y_1 - y) + l_1(y - y_0)}{(y_1 - y_0)} \frac{(x_1 - x)}{(x_1 - x_0)}.$$

В качестве функции выигрыша $H(x, y)$ игрока 1 понимается производительность поиска, т. е. просмотренная площадь в единицу времени $H(x, y) = 2x \cdot l(x, y)$. Выигрыш игрока 2 полагаем равным $[-H(x, y)]$. Таким образом, получаем игру с функцией выигрыша

$$H(x, y) = 2x \frac{l_0(y_1 - y) + l_1(y - y_0)}{(y_1 - y_0)} \frac{(x_1 - x)}{(x_1 - x_0)},$$

где $x \in [x_0, x_1], y \in [y_0, y_1]$.

2.1.3. В заключение отметим специальный класс антагонистических игр, в которых $X = Y = [0, 1]$. В этих играх ситуации — суть пары чисел (x, y) , где $x, y \in [0, 1]$. Эти пары задают точки единичного квадрата, поэтому такие игры называются играми на *единичном квадрате*. Класс игр на единичном квадрате во многом характеризует бесконечные антагонистические игры и поэтому является базовым при исследовании бесконечных игр. В частности, примеры 2, 4, 5 — примеры игр на единичном квадрате. Пример 6 также является игрой на единичном квадрате, если положить $x_0 = y_0 = 0, x_1 = y_1 = 1$.

§ 2.2. Ситуация ε —равновесия, ε —седловые точки и ε —оптимальные стратегии

2.2.1. Как и во всякой антагонистической игре $\Gamma(X, Y, H)$, в бесконечной игре принципом оптимального поведения игроков является принцип равновесия. Оптимальной (равновесной) является такая ситуация (x^*, y^*) , для которой выполняются неравенства

$$H(x, y^*) \leq H(x^*, y^*) \leq H(x^*, y) \quad (2.2.1)$$

при всех $x \in X, y \in Y$. Этот принцип реализуется в игре Γ в том и только в том случае, когда

$$v = \bar{v} = \underline{v} = H(x^*, y^*),$$

где

$$\underline{v} = \max_x \inf_y H(x, y), \quad \bar{v} = \min_y \sup_x H(x, y), \quad (2.2.2)$$

т. е. внешние экстремумы максимина и минимакса достигаются и нижнее значение игры \underline{v} равно верхнему значению \bar{v} . Такая антагонистическая игра Γ называется *вполне определенной*, а число v — значением игры (см. п. 1.3.4 гл. I).

Для матричных игр существование и равенство максимина минимаксу было доказано в классе смешанных стратегий (см. §6 гл. I), поэтому решение игры заключалось в нахождении их общего значения v и тех стратегий x^*, y^* , на которых достигаются внешние экстремумы в (2.2.2).

Для бесконечных игр существование внешних экстремумов в (2.2.2), вообще говоря, не обязательно.

2.2.2. Пример 7. Пусть каждый из игроков 1 и 2 выбирает число из открытого интервала $(0, 1)$, после чего игрок 1 получает выигрыш, равный сумме выбранных чисел. Таким образом, получаем игру на открытом единичном квадрате с функцией выигрыша $H(x, y)$ игрока 1:

$$H(x, y) = x + y, \quad x \in (0, 1), \quad y \in (0, 1). \quad (2.2.3)$$

Здесь ситуация $(1, 0)$ была бы равновесной, если бы 1 и 0 входили в число стратегий игроков, а значением игры v являлось $v = 1$. В действительности внешние экстремумы в (2.2.2) не достигаются, а верхнее и нижнее значения игры равны между собой. Поэтому $v = 1$ и игрок 1, выбирая число $1 - \varepsilon$, $\varepsilon > 0$, достаточно близкое к 1, всегда может получить выигрыш, достаточно близкий к значению игры. С другой стороны, игрок 2, выбирая число $\varepsilon > 0$ достаточно малым (близким к 0), может гарантировать, что его проигрыш будет сколь угодно близким к значению игры.

2.2.3. Определение. Ситуация $(x_\varepsilon, y_\varepsilon)$ в антагонистической игре $\Gamma = (X, Y, H)$ называется ситуацией ε -равновесия, если для любых стратегий $x \in X$ и $y \in Y$ игроков 1 и 2 соответственно, выполняется неравенство:

$$H(x, y_\varepsilon) - \varepsilon \leq H(x_\varepsilon, y_\varepsilon) \leq H(x_\varepsilon, y) + \varepsilon. \quad (2.2.4)$$

Точка $(x_\varepsilon, y_\varepsilon)$, для которой имеет место (2.2.4), называется ε -седловой точкой, а стратегии x_ε и y_ε — ε -оптимальными стратегиями игроков 1 и 2 соответственно.

Полезно сравнить определения ситуации равновесия (2.2.1) и ε -равновесия (2.2.4). Если отклонение от оптимальной стратегии приводит лишь к уменьшению выигрыша этого игрока, то отклонение от ε -оптимальной стратегии может привести к его увеличению, но не более чем на ε .

Так, ситуация $(1 - \varepsilon, \varepsilon)$, $0 < \varepsilon < 1$ является ε -равновесной в примере 7, а стратегии $x_\varepsilon = 1 - \varepsilon$, $y_\varepsilon = \varepsilon$ — ε -оптимальными стратегиями игроков 1 и 2 соответственно.

2.2.4. Заметим, что для двух стратегически эквивалентных игр $\Gamma = (X, Y, H)$ и $\Gamma' = (X, Y, H')$, где $H' = \beta H + \alpha$, $\beta > 0$, справедливы следующие результаты. Если $(x_\varepsilon, y_\varepsilon)$ — ситуация ε -равновесия в игре Γ , то она является ситуацией $(\beta\varepsilon)$ -равновесия в игре Γ' (сравните с леммой о масштабе §1.3).

2.2.5. Основное свойство ε -оптимальных стратегий дает следующая теорема.

Теорема. Для того, чтобы существовало значение $v = \inf_y \sup_x H(x, y) = \sup_x \inf_y H(x, y) < +\infty$ антагонистической игры $\Gamma = (X, Y, H)$, необходимо и достаточно, чтобы для любого $\varepsilon > 0$, существовали ε -оптимальные стратегии $x_\varepsilon, y_\varepsilon$ игроков 1 и 2, при этом

$$\lim_{\varepsilon \rightarrow 0} H(x_\varepsilon, y_\varepsilon) = v. \quad (2.2.5)$$

Доказательство. Необходимость. Пусть игра Γ имеет конечное значение v . Для любого $\varepsilon > 0$ выберем стратегию y_ε из условия

$$\sup_{x \in X} H(x, y_\varepsilon) - \frac{\varepsilon}{2} \leq v \quad (2.2.6)$$

и стратегию x_ε из условия

$$\inf_{y \in Y} H(x_\varepsilon, y) + \frac{\varepsilon}{2} \geq v. \quad (2.2.7)$$

Из (2.2.2), (2.2.6), (2.2.7) получаем неравенство

$$H(x, y_\varepsilon) - \frac{\varepsilon}{2} \leq v \leq H(x_\varepsilon, y) + \frac{\varepsilon}{2} \quad (2.2.8)$$

для всех стратегий x, y . Следовательно,

$$|H(x_\varepsilon, y_\varepsilon) - v| \leq \frac{\varepsilon}{2}. \quad (2.2.9)$$

Из неравенств (2.2.8), (2.2.9) следуют соотношения (2.2.4), (2.2.5).

Достаточность. Если для любого числа $\varepsilon > 0$ выполняются неравенства (2.2.4), то

$$\begin{aligned} \bar{v} &= \inf_y \sup_x H(x, y) \leq \sup_x H(x, y_\varepsilon) \leq H(x_\varepsilon, y_\varepsilon) + \varepsilon \leq \\ &\leq \inf_y H(x_\varepsilon, y) + 2\varepsilon \leq \sup_x \inf_y H(x, y) + 2\varepsilon = \underline{v} + 2\varepsilon. \end{aligned} \quad (2.2.10)$$

Отсюда заключаем, что $\bar{v} \leq \underline{v}$, но согласно лемме п. 1.2.2 гл. I справедливо противоположное неравенство. Таким образом, остается доказать, что значение игры Γ конечно. Возьмем такую последовательность $\{\varepsilon_n\}$, что $\lim_{n \rightarrow \infty} \varepsilon_n = 0$. Пусть $\varepsilon_k \in \{\varepsilon_n\}$, $\varepsilon_{k+m} \in \{\varepsilon_n\}$, где m — любое фиксированное натуральное число. Имеем

$$\begin{aligned} H(x_{\varepsilon_{k+m}}, y_{\varepsilon_k}) + \varepsilon_{k+m} &\geq H(x_{\varepsilon_{k+m}}, y_{\varepsilon_{k+m}}) \geq H(x_{\varepsilon_k}, y_{\varepsilon_{k+m}}) - \varepsilon_{k+m}, \\ H(x_{\varepsilon_k}, y_{\varepsilon_{k+m}}) + \varepsilon_k &\geq H(x_{\varepsilon_k}, y_{\varepsilon_k}) \geq H(x_{\varepsilon_{k+m}}, y_{\varepsilon_k}) - \varepsilon_k. \end{aligned}$$

Таким образом,

$$|H(x_{\varepsilon_k}, y_{\varepsilon_k}) - H(x_{\varepsilon_{k+m}}, y_{\varepsilon_{k+m}})| \leq \varepsilon_k + \varepsilon_{k+m} = \delta_{km}.$$

Так как $\lim_{k \rightarrow \infty} \delta_{km} = 0$ при любом фиксированном значении m , то существует конечный предел $\lim_{\varepsilon \rightarrow 0} H(x_\varepsilon, y_\varepsilon)$. Из соотношения (2.2.9) получаем неравенство $|H(x_\varepsilon, y_\varepsilon) - v| \leq \varepsilon$; следовательно, $v = \lim_{\varepsilon \rightarrow 0} H(x_\varepsilon, y_\varepsilon)$. Теорема доказана.

2.2.6. Для иллюстрации приведенных в этом параграфе определений рассмотрим подробно пример 1 п. 1.1.2.

Пример 8. Предположим, что множества S_1 и S_2 представляют собой замкнутые круги с радиусами R_1 и R_2 ($R_1 < R_2$). Найдем нижнее значение игры

$$\underline{v} = \max_{x \in S_1} \min_{y \in S_2} \rho(x, y).$$

Пусть $x_0 \in S_1$. Тогда $\min_y \rho(x_0, y)$ достигается в точке y_0 пересечения прямой, проходящей через центр O_1 круга S_2 и точку x_0 , с границей круга S_2 . Очевидно, что величина $\min_{y \in S} \rho(x_0, y)$ достигает максимального значения в точке $M \in S_1$, являющейся точкой пересечения линий центров OO_1 (рис. 2.2) с границей круга S_1 , наиболее удаленной от точки O_1 .

Таким образом, $\underline{v} = |O_1 M| - R_2$.

Для вычисления верхнего значения игры

$$\bar{v} = \min_{y \in S_2} \max_{x \in S_1} \rho(x, y)$$

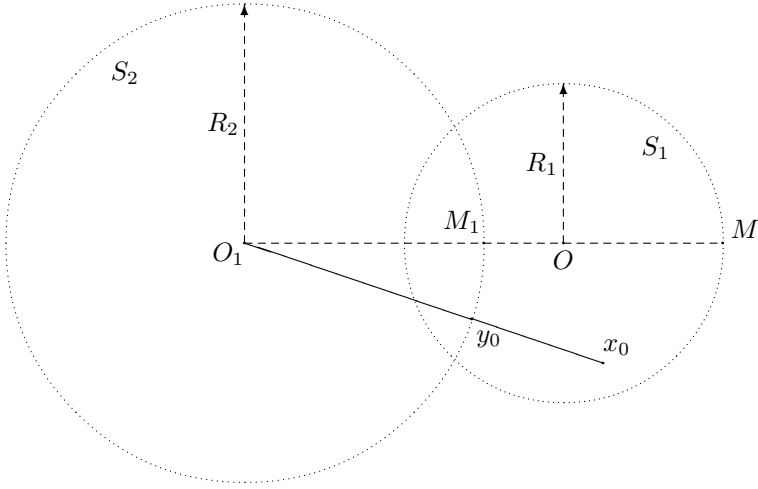


Рис. 2.2. Нижнее значение игры

рассмотрим два случая.

Случай 1. Центр O круга S_1 принадлежит множеству S_2 (рис. 2.3). Для каждого $y_0 \in S_2$ точка x_0 , доставляющая $\max_{x \in S_1} \rho(x, y_0)$, строится следующим образом.

Пусть x_0^1 и x_0^2 — точки пересечения прямой $O_1 y_0$ с границей круга S_1 , а x_0^3 — точка пересечения прямой $O y_0$ с границей круга S_1 , наиболее удаленная от точки y_0 . Тогда x_0 определяется из условия

$$\rho(x_0, y_0) = \max_{i=1,2,3} \rho(x_0^i, y_0).$$

По построению, для всех $y_0 \in S_2$

$$\max_{x \in S_1} \rho(x, y_0) = \rho(x_0, y_0) \geq R_1.$$

Однако при $y_0 = O$ получаем

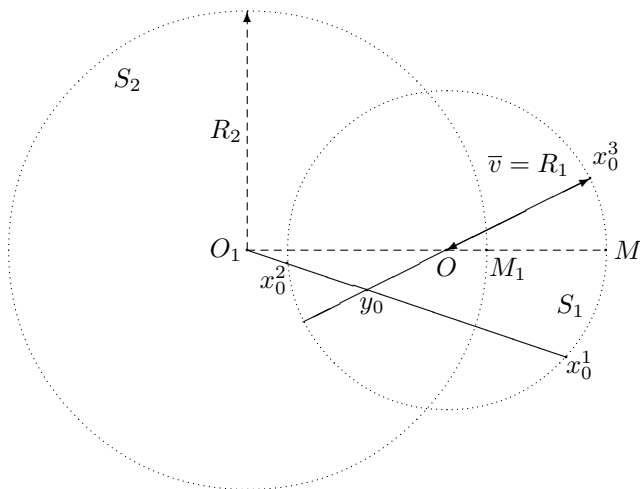
$$\max_{x \in S_1} \rho(x, O) = R_1,$$

следовательно,

$$\min_{y \in S_2} \max_{x \in S_1} \rho(x, y) = \bar{v} = R_1.$$

Непосредственно видно, что точка $O \in S_2$ и поэтому $\bar{v} = R_1 \geq |\overline{O_1 M}| - R_2 = \underline{v}$. При этом равенство возможно лишь при условии, что O принадлежит границе множества S_2 .

Таким образом, если в случае 1 точка O не принадлежит границе множества S_2 , то значения игры и ситуации равновесия не существует. Если же точка O принадлежит границе множества S_2 , то существует ситуация равновесия, при этом оптимальная стратегия игрока 1 заключается в выборе точки M , лежащей на пересечении линии центров OO_1 с границей множества S_1 и наиболее удаленной от точки O_1 . Оптимальная стратегия игрока 2 заключается в выборе точки $y \in S_2$, совпадающей с центром O круга S_1 . Значение игры при этом равно $v = \underline{v} = \bar{v} = R_1 + R_2 - R_2 = R_1$.

Рис. 2.3. Случай 1, когда $O \in S_2$

Случай 2. Центр круга $O \notin S_2$. Этот случай рассматривается как вариант случая 1, когда центр круга S_1 принадлежит границе множества S_2 . Вычислим величину \bar{v} (рис. 2.4).

Пусть $y_0 \in S_2$. Тогда точка x_0 , доставляющая $\max_{x \in S_1} \rho(x, y_0)$, совпадает с точкой пересечения x_0 прямой, проходящей через y_0 и центр O круга S_1 с границей круга S_1 , наиболее удаленной от точки y_0 .

Действительно, круг радиусом $\overline{x_0 y_0}$ с центром в точке y_0 содержит S_1 и его граница касается границы круга S_1 в единственной точке x_0 . Очевидно, что величина $\max_{x \in S_1} \rho(x, y) = \rho(x_0, y)$ достигает минимума в точке M_1 пересечения отрезка OO_1 с границей круга S_2 . Таким образом, в рассматриваемом случае

$$\bar{v} = \min_{y \in S_2} \max_{x \in S_1} \rho(x, y) = |\overline{O_1 M}| - R_2 = \underline{v}.$$

Оптимальные стратегии заключаются в выборе $M \in S_1$ и $M_1 \in S_2$ игроками 1 и 2 соответственно.

Если в качестве множеств стратегий в примере 1 п. 1.1.2 рассматривать открытые круги S_1 и S_2 , то в случае 2 значение игры существует и равно

$$\underline{v} = \sup_{x \in S_1} \inf_{y \in S_2} \rho(x, y) = \inf_{y \in S_2} \sup_{x \in S_1} \rho(x, y) = \bar{v} = |\overline{O_1 M}| - R_2 = \underline{v}.$$

Однако оптимальных стратегий не существует, поскольку $M \notin S_1$, $M_1 \notin S_2$. Тем не менее, для любого $\varepsilon > 0$ существуют ε -оптимальные стратегии — это точки из ε -окрестностей точек M и M_1 , принадлежащие соответственно множествам S_1 и S_2 .

В заключение отметим, что игра в примере 6 имеет ситуацию равновесия в чистых стратегиях (см. упр. 7), а игры в примерах 1 – 5, вообще говоря, не имеют ситуации равновесия и значения игры. Так, в примере 2 лишь при $x^* = 1/2$ у игрока 1 есть оптимальная стратегия $l \geq 1/2$, а значение игры равно единице (у игрока 2 оптимальной является любая стратегия).

В соответствии с этими мерами они реализуют (например, с помощью таблицы случайных чисел) случайный выбор стратегий $x \in X$ и $y \in Y$. После этого игрок 1 получает выигрыш $H(x, y)$. Стратегии $\mu \in \bar{X}$, $\nu \in \bar{Y}$ называются смешанными, а $x \in X$, $y \in Y$ — чистыми стратегиями в игре Γ .

Введение смешанного расширения бесконечной игры требует определенных пояснений. Множества \bar{X} и \bar{Y} зависят от того, на каких σ -алгебрах \mathcal{X} и \mathcal{Y} рассматриваются вероятностные меры. В случае матричных игр (множества X и Y конечны) в смешанном расширении игроки выбирали свои стратегии согласно вероятностным распределениям на множествах X и Y . Если X — бесконечное множество и мы будем поступать так же, как в конечном случае, то необходимо рассматривать меры, для которых измеримы все подмножества бесконечного множества X . Однако таких мер сравнительно мало: это меры, сосредоточенные на не более чем счетных множествах точек. Используя только такие меры, игроки обедняют свои возможности (и далеко не всегда могут гарантировать существование ситуации равновесия в смешанных стратегиях). Поэтому используют менее обширные σ -алгебры, на которых определяют вероятностные меры. Тогда возможных вероятностных мер существенно больше (и, как правило, гарантируется существование ситуации равновесия в смешанных стратегиях). Однако в этом случае не всякая функция H на $X \times Y$ окажется измеримой, поэтому нельзя определить математическое ожидание выигрыша и тем самым понятие равновесия, значения игры и оптимальных стратегий. Таким образом, здесь необходим известный компромисс. С точки зрения проблемы нахождения решения желательно, чтобы смешанные стратегии имели наиболее простой вид и в то же время в этом расширении существовало, по крайней мере, значение игры.

Строго говоря, интеграл в (2.3.1) должен браться по мере $\mu \times \nu$ на декартовом произведении $X \times Y$. Однако согласно правилам антагонистической игры смешанные стратегии (меры) μ и ν игроками выбираются одновременно и независимо друг от друга, т. е. вероятностные меры μ и ν стохастически независимы.

Определение. Ситуацией (μ, ν) в смешанных стратегиях называется пара вероятностных мер $\mu \in \bar{X}$, $\nu \in \bar{Y}$, которые стохастически независимы.

Таким образом, в ситуации (μ, ν) в смешанных стратегиях выигрыш $K(\mu, \nu)$ равен повторному интегралу (2.3.1). Одноточечные множества принадлежат σ -алгебре подмножеств множества стратегий, на которой определяются вероятностные меры, поэтому каждой чистой стратегии $x(y)$ можно поставить в соответствие вероятностную меру $\mu_x \in \bar{X}$ ($\nu_y \in \bar{Y}$), сосредоточенную в точке $x \in X$ ($y \in Y$). отождествляя стратегии x и μ_x , y и ν_y , видим, что чистые стратегии являются частным случаем смешанных, т. е. справедливы включения $X \subset \bar{X}$, $Y \subset \bar{Y}$. Тогда выигрыши игрока 1 в ситуациях (x, ν) и (μ, y) равны, соответственно, математическим ожиданиям:

$$K(x, \nu) = K(\mu_x, \nu) = \int_Y H(x, y) d\nu(y), \quad (2.3.2)$$

$$K(\mu, y) = K(\mu, \nu_y) = \int_X H(x, y) d\mu(x), \quad (2.3.3)$$

где интегралы в (2.3.1), (2.3.2), (2.3.3) понимаются в смысле Лебега — Стильеса. Если же распределения $\mu(x)$, $\nu(y)$ имеют плотности $f(x)$ и $g(y)$, т. е. $d\mu(x) = f(x)dx$ и $d\nu(y) = g(y)dy$, интегралы в (2.3.1), (2.3.2), (2.3.3) понимаются в смысле Римана — Стильеса. Таким образом, $\Gamma \subset \bar{\Gamma}$ — подыгра своего смешанного расширения $\bar{\Gamma}$. Будем считать, что

все интегралы в (2.3.1), (2.3.2), (2.3.3) существуют, каковы бы ни были вероятностные меры μ и ν .

Определение. Пусть $\Gamma = (X, Y, H)$ — антагонистическая игра, а $\bar{\Gamma} = (\bar{X}, \bar{Y}, K)$ — ее смешанное расширение. Тогда ситуация $(\mu^*, \nu^*) \in \bar{X} \times \bar{Y}$ называется ситуацией равновесия в игре Γ в смешанных стратегиях, если для всех $\mu \in \bar{X}$ и $\nu \in \bar{Y}$ выполняются неравенства:

$$K(\mu, \nu^*) \leq K(\mu^*, \nu^*) \leq K(\mu^*, \nu), \quad (2.3.4)$$

т. е. (μ^*, ν^*) — ситуация равновесия в смешанном расширении игры Γ , а $\mu^*(\nu^*)$ — оптимальная стратегия игрока 1 (2) в $\bar{\Gamma}$.

Аналогично, ситуация $(\mu_\varepsilon^*, \nu_\varepsilon^*) \in \bar{X} \times \bar{Y}$ называется ситуацией ε -равновесия в игре Γ в смешанных стратегиях, если для всех $\mu \in \bar{X}$ и $\nu \in \bar{Y}$ выполняются неравенства

$$K(\mu, \nu_\varepsilon^*) - \varepsilon \leq K(\mu_\varepsilon^*, \nu_\varepsilon^*) \leq K(\mu_\varepsilon^*, \nu) + \varepsilon, \quad (2.3.5)$$

т. е. $\mu_\varepsilon^*(\nu_\varepsilon^*)$ — ε -оптимальная стратегия игрока 1 (2) в $\bar{\Gamma}$.

2.3.3. Подобно тому, как это доказывалось для матричных игр, можно показать, что если функции выигрыша игр $\Gamma = (X, Y, H)$ и $\Gamma' = (X, Y, H')$ связаны равенством $H'(x, y) = \alpha H(x, y) + \beta$, $\alpha > 0$, то множества ситуаций равновесия в играх Γ и Γ' совпадают, а значения игр связаны соотношением $v(\Gamma') = \alpha v(\Gamma) + \beta$ (см. § 1.4).

2.3.4. Ситуации равновесия в смешанных стратегиях обладают такими же свойствами, как и в случае матричных игр, что следует из приведенных ниже теорем.

Теорема. Для того, чтобы пара (μ^*, ν^*) , $\mu^* \in \bar{X}$, $\nu^* \in \bar{Y}$ была ситуацией равновесия (ε -равновесия) в смешанных стратегиях в игре Γ , необходимо и достаточно для всех $x \in X, y \in Y$ выполнение неравенств:

$$K(x, \nu^*) \leq K(\mu^*, \nu^*) \leq K(\mu^*, y), \quad (2.3.6)$$

$$(K(x, \nu^*) - \varepsilon \leq K(\mu^*, \nu^*) \leq K(\mu^*, y) + \varepsilon). \quad (2.3.7)$$

Доказательство. Необходимость теоремы очевидна, поскольку чистые стратегии являются частным случаем смешанных. Докажем достаточность для (2.3.6) (для (2.3.7) это доказывается аналогично). Пусть μ и ν — произвольные смешанные стратегии игроков 1 и 2 соответственно. Тогда из (2.3.1), (2.3.2) и (2.3.6) получаем

$$K(\mu, \nu^*) = \int_X K(x, \nu^*) d\mu(x) \leq K(\mu^*, \nu^*),$$

$$K(\mu^*, \nu) = \int_Y K(\mu^*, y) d\nu(y) \geq K(\mu^*, \nu^*).$$

Отсюда вытекают неравенства (2.3.4), что и требовалось доказать.

Из теоремы, в частности, следует, что если (x^*, y^*) — ситуация равновесия (ε -равновесия) в чистых стратегиях в игре Γ , то она является и ситуацией равновесия (ε -равновесия) в смешанном расширении $\bar{\Gamma}$, при этом значение игры v сохраняется.

Заметим, что смешанное расширение $\bar{\Gamma}$ является антагонистической игрой, поэтому относительно $\bar{\Gamma}$ справедливо понятие вполне определенной игры (п. 2.2.1), а также теорема п. 2.2.5, только речь теперь идет о ситуации равновесия и значении игры в смешанных стратегиях.

2.3.5. Теорема. Для того, чтобы игра $\Gamma = (X, Y, H)$ имела значение v в смешанных стратегиях, т. е. $\sup_{\mu} \inf_{\nu} K(\mu, \nu) = \inf_{\nu} \sup_{\mu} K(\mu, \nu) = v$, необходимо и достаточно выполнение равенства

$$\sup_{\mu} \inf_y K(\mu, y) = \inf_{\nu} \sup_x K(x, \nu) = v. \quad (2.3.8)$$

Если при этом игроки имеют оптимальные стратегии, то внешние экстремумы в (2.3.8) достигаются и равенства

$$\inf_y K(\mu^*, y) = v, \quad (2.3.9)$$

$$\sup_x K(x, \nu^*) = v \quad (2.3.10)$$

являются необходимыми и достаточными условиями оптимальности смешанных стратегий $\mu^* \in \bar{X}$ и $\nu^* \in \bar{Y}$.

Доказательство. Пусть v — значение игры. Тогда по определению

$$v = \sup_{\mu} \inf_{\nu} K(\mu, \nu). \quad (2.3.11)$$

Для фиксированной стратегии μ , множество $\{K(\mu, \nu) | \nu \in \bar{Y}\}$ — выпуклая оболочка чисел $K(\mu, y)$, $y \in \bar{Y}$. Так как точная нижняя граница любого множества действительных чисел совпадает с точной нижней границей выпуклой оболочки этих чисел, то

$$\inf_{\nu \in \bar{Y}} K(\mu, \nu) = \inf_{y \in Y} K(\mu, y). \quad (2.3.12)$$

Равенство (2.3.12) можно получить также из следующих соображений. Поскольку $Y \subset \bar{Y}$, имеем

$$\inf_{\nu \in \bar{Y}} K(\mu, \nu) \leq \inf_{y \in Y} K(\mu, y).$$

Предположим, что неравенство строгое, т. е.

$$\inf_{\nu} K(\mu, \nu) < \inf_y K(\mu, y).$$

Это значит, что при некотором достаточно малом $\varepsilon > 0$ выполняется неравенство

$$\inf_{\nu} K(\mu, \nu) + \varepsilon < \inf_y K(\mu, y).$$

Таким образом, при всех $y \in Y$

$$K(\mu, y) > \inf_{\nu} K(\mu, \nu) + \varepsilon. \quad (2.3.13)$$

Теперь, переходя к смешанным стратегиям в (2.3.13), получаем

$$\inf_{\nu} K(\mu, \nu) \geq \inf_{\nu} K(\mu, \nu) + \varepsilon.$$

Полученное противоречие и доказывает (2.3.12). Возьмем супремум по μ в (2.3.12). Тогда

$$v = \sup_{\mu} \inf_y K(\mu, y).$$

Аналогично доказывается правое из равенств в (2.3.8). Обратно, если (2.3.8) выполнено, то из (2.3.12) следует, что v является значением игры.

Пусть теперь μ^*, ν^* — оптимальные стратегии игроков 1 и 2 соответственно. По теореме п. 1.3.4 внешние экстремумы в (2.3.8) достигаются, а (2.3.9), (2.3.10) являются необходимыми и достаточными условиями оптимальности смешанных стратегий μ^* и ν^* .

В п. 2.3.2 отмечалось, что введение смешанных стратегий в бесконечной антагонистической игре зависит от способа рандомизации множества чистых стратегий. Однако из (2.3.8) следует, что значение v игры не зависит от способа рандомизации. Так, для доказательства его существования достаточно найти хотя бы одно смешанное расширение игры, для которого выполнялось бы равенство (2.3.8).

Следствие. Для любой антагонистической игры $\Gamma = (X, Y, H)$, имеющей значение v в смешанных стратегиях, справедливо неравенство:

$$\sup_x \inf_y H(x, y) \leq v \leq \inf_y \sup_x H(x, y). \quad (2.3.14)$$

Доказательство. Из теоремы п. 2.3.5 следует:

$$\sup_x \inf_y H(x, y) \leq \sup_\mu \inf_y K(\mu, y) = v = \inf_\nu \sup_x K(x, \nu) \leq \inf_y \sup_x H(x, y).$$

2.3.6. Из (2.3.14) следует один из способов приближенного решения антагонистической игры. Действительно, пусть внешние экстремумы в (2.3.14) достигаются, т. е.

$$v^- = \max_x \inf_y H(x, y) = \inf_y H(x^0, y), \quad (2.3.15)$$

$$v^+ = \min_y \sup_x H(x, y) = \sup_x H(x, y^0), \quad (2.3.16)$$

и пусть $\alpha = v^+ - v^-$. Тогда максиминная стратегия x^0 игрока 1 и минимаксная стратегия y^0 игрока 2 с точностью до α описывают оптимальное поведение игроков и могут быть взяты в качестве приближенного решения игры Γ . Таким образом, в этом случае задача сводится к нахождению максиминных и минимаксных стратегий игроков 1 и 2 соответственно, а точность приближенного решения определяется величиной $\alpha = v^+ - v^-$, при этом значение игры v согласно (2.3.14) лежит в интервале $v \in [v^-, v^+]$. Способам нахождения решения задач (2.3.15), (2.3.16) посвящена теория минимакса [Демьянов, Малоземов, 1972], [Данскин, 1970].

2.3.7. Как и в случае матричных игр, для бесконечных игр важную роль играет понятие спектра смешанной стратегии.

Определение. Пусть $\Gamma = (X, Y, H)$ — антагонистическая игра. Тогда чистую стратегию $x_0 \in X$ ($y_0 \in Y$) игрока 1 (2) называют точкой концентрации его смешанной стратегии $\mu(\nu)$, если $\mu(x_0) > 0$ ($\nu(y_0) > 0$).

Определение. Чистая стратегия $x_0 \in X$ ($y_0 \in Y$), где X (соответственно, Y) — топологическое пространство, называется точкой спектра смешанной стратегии $\mu(\nu)$, заданной на борелевской σ -алгебре подмножеств множества $X(Y)$, если для любой измеримой окрестности ω точки $x_0(y_0)$ имеет место неравенство:

$$\mu(\omega) = \int_{\omega} d\mu(x) > 0 \quad (\nu(\omega) = \int_{\omega} d\nu(y) > 0).$$

Спектром смешанной стратегии $\mu(\nu)$ назовем наименьшее замкнутое множество, μ -мера (ν -мера) которого равно единице.

Точки концентрации смешанной стратегии являются точками спектра; обратное, вообще говоря, неверно. Так, чистые стратегии, в которых смешанная стратегия имеет положительную плотность, являются точками спектра, но они не являются точками концентрации.

Спектр смешанной стратегии μ (соответственно, ν) будем обозначать $X_\mu(Y_\nu)$. Докажем аналог теоремы п. 1.7.6 гл. I о дополняющей нежесткости для бесконечных игр.

Теорема. Пусть $\Gamma = (X, Y, H)$ — антагонистическая игра, имеющая значение v . Тогда, если $x_0 \in X$, а ν^* — оптимальная смешанная стратегия игрока 2 и

$$K(x_0, \nu^*) < v, \quad (2.3.17)$$

то x_0 не может быть точкой концентрации какой-либо оптимальной стратегии игрока 1.

Аналогичный результат справедлив и для точек концентрации оптимальных стратегий игрока 2.

Доказательство. Из оптимальности смешанной стратегии $\nu^* \in \bar{Y}$ следует, что для всех $x \in X$ выполняется неравенство:

$$K(x, \nu^*) \leq v.$$

Интегрируя его по оптимальной смешанной стратегии (мере) μ^* игрока 1 на множестве $X \setminus \{x_0\}$, получаем

$$\int_{X \setminus \{x_0\}} K(x, \nu^*) d\mu^*(x) \leq v \int_{X \setminus \{x_0\}} d\mu^*(x).$$

Пусть $\mu^*(x_0) > 0$, т. е. x_0 — точка концентрации оптимальной смешанной стратегии μ^* игрока 1. Тогда из (2.3.17) имеем

$$K(x_0, \nu^*) \mu^*(x_0) < v \mu^*(x_0).$$

Складывая два последних неравенства, получаем противоречие

$$v = \int_X K(x, \nu^*) d\mu^*(x) = K(\mu^*, \nu^*) < v.$$

Следовательно, $\mu^*(x_0) = 0$ для всех оптимальных стратегий $\mu^* \in \bar{X}$.

2.3.8. Для бесконечных антагонистических игр можно ввести понятие доминирования стратегий аналогично тому, как это делалось в §1.8.

Определение. Стратегия $\mu_1 \in \bar{X}$ игрока 1 строго доминирует стратегию $\mu_2 \in \bar{X}$ ($\mu_1 \succ \mu_2$), если

$$H(\mu_1, y) > H(\mu_2, y),$$

для всех $y \in Y$. Аналогично, стратегия $\nu_1 \in \bar{Y}$ игрока 2 строго доминирует стратегию $\nu_2 \in \bar{Y}$ ($\nu_1 \succ \nu_2$), если

$$H(x, \nu_1) < H(x, \nu_2),$$

для всех $x \in X$. Стратегии μ_2 и ν_2 называются строго доминируемыми, если существуют $\mu_1 \succ \mu_2$ и $\nu_1 \succ \nu_2$.

Если последние неравенства выполняются как нестрогие, то говорят, что μ_1 доминирует μ_2 ($\mu_1 \succeq \mu_2$) и ν_1 доминирует ν_2 ($\nu_1 \succeq \nu_2$).

Приведем без доказательства теоремы о доминировании, аналогичные теоремам п.1.8.3.

Теорема. Для бесконечной антагонистической игры, имеющей решение, ни одна строго доминируемая чистая стратегия игрока не содержится в спектрах его оптимальных смешанных стратегий.

Теорема. Пусть $\Gamma = (X, Y, H)$ — бесконечная антагонистическая игра, имеющая решение (X и Y — топологические пространства), и каждый элемент открытого множества $X^0 \subset X$ доминируется некоторой стратегией μ^0 , спектр которой не пересекается с X^0 . Тогда всякое решение игры $\Gamma' = (X \setminus X^0, Y, H)$ является решением игры Γ .

Аналогичная теорема верна и для стратегий игрока 2.

2.3.9. В этом параграфе рассмотрены свойства оптимальных (ε -оптимальных) смешанных стратегий в предположении существования решения игры. Матричная игра вполне определена в смешанных стратегиях, т. е. всегда существуют значение и ситуация равновесия, что следует из теоремы п.1.6.1 гл. I. Возможности решения бесконечных антагонистических игр в смешанных стратегиях ограничены, что показывает следующий пример.

Пример 9 (Игра, не имеющая значения в смешанных стратегиях) [Воробьев, 1984]. Рассмотрим игру $\Gamma = (X, Y, H)$, где $X = Y = \{1, 2, \dots\}$ — множество натуральных чисел, а функция выигрыша имеет вид

$$H(x, y) = \begin{cases} 1, & \text{если } x > y, \\ 0, & \text{если } x = y, \\ -1, & \text{если } x < y. \end{cases}$$

Эта игра не имеет значения в чистых стратегиях. Покажем, что она не имеет значения и в смешанных стратегиях.

Пусть μ — произвольная смешанная стратегия игрока $d\mu(x) = \delta_x$, где $\delta_x \geq 0$ и $\sum_{x=1}^{\infty} \delta_x = 1$. Возьмем $\varepsilon > 0$ и найдем y_ε такое, что

$$\sum_{x \leq y_\varepsilon} \delta_x > 1 - \varepsilon.$$

Тогда

$$\begin{aligned} K(\mu, y_\varepsilon) &= \sum_{x=1}^{\infty} \delta_x H(x, y_\varepsilon) = \sum_{x \leq y_\varepsilon} \delta_x K(x, y_\varepsilon) + \sum_{x > y_\varepsilon} \delta_x K(x, y_\varepsilon) = \\ &= - \sum_{x \leq y_\varepsilon} \delta_x + \sum_{x > y_\varepsilon} \delta_x < -1 + 2\varepsilon. \end{aligned}$$

В силу произвольности $\varepsilon > 0$ и так как $H(x, y)$ не принимает значений, меньших -1 , имеем

$$\inf_y K(\mu, y) = -1.$$

Следовательно, поскольку стратегия μ произвольна,

$$\underline{v} = \sup_{\mu} \inf_y K(\mu, y) = -1.$$

Рассуждая аналогично, получаем

$$\bar{v} = \inf_{\nu} \sup_x K(x, \nu) = 1.$$

Поскольку $\bar{v} > \underline{v}$, то игра Γ не имеет значения в смешанных стратегиях. Как будет показано в следующем параграфе, непрерывности функции выигрыша и компактности пространства стратегий достаточно для того, чтобы игра имела решение (значение и оптимальные стратегии) в смешанном расширении.

§ 2.4. Игры с непрерывной функцией выигрыша

2.4.1. В данном параграфе рассмотрим антагонистические игры $\Gamma = (X, Y, H)$ в предположении, что пространства стратегий X и Y — метрические компакты (чаще всего они будут подмножествами евклидовых пространств), а функция H непрерывна по обоим переменным. Под множествами \bar{X} , \bar{Y} смешанных стратегий игроков 1 и 2 будем понимать множества вероятностных мер, заданных на σ -алгебрах \mathcal{X} и \mathcal{Y} борелевских множеств пространств X и Y соответственно. Тогда выигрыш $K(\mu, \nu)$ игрока 1 в ситуации $(\mu, \nu) \in \bar{X} \times \bar{Y}$ в смешанных стратегиях — измеримая функция относительно борелевской σ -алгебры $\mathcal{X} \times \mathcal{Y}$, которая определяется интегралом (2.3.1) и представляет собой математическое ожидание выигрыша по вероятностной мере $\mu \times \nu$.

Игру $\Gamma = (X, Y, H)$, определенную указанным выше способом, будем называть непрерывной игрой.

2.4.2. Теорема. Если $\Gamma = (X, Y, H)$ — бесконечная антагонистическая игра, имеющая значение v и ситуацию равновесия (μ^*, ν^*) , а функции $K(\mu^*, y)$, $K(x, \nu^*)$ — непрерывны соответственно по y и x , то справедливы равенства

$$K(\mu^*, y) = v, \quad y \in Y_{\nu^*}, \quad (2.4.1)$$

$$K(x, \nu^*) = v, \quad x \in X_{\mu^*}, \quad (2.4.2)$$

где Y_{ν^*} , X_{μ^*} — спектры смешанных стратегий ν^* и μ^* соответственно.

Доказательство. Из теоремы п. 2.3.4 следует, что неравенство

$$K(\mu^*, y) \geq v \quad (2.4.3)$$

выполняется для всех точек $y \in Y$. Если (2.4.1) не выполнено, то существует такая точка $y_0 \in Y_{\nu^*}$, что $K(\mu^*, y_0) > v$. В силу непрерывности функции $K(\mu^*, y)$ неравенство (2.4.3) в некоторой окрестности ω точки y_0 выполняется как строгое. Из того, что $y_0 \in Y_{\nu^*}$ — точка спектра смешанной стратегии ν^* , следует, что $\nu^*(\omega) > 0$. Отсюда и из неравенства (2.4.3) получаем

$$v = K(\mu^*, \nu^*) = \int_Y K(\mu^*, y) d\nu^*(y) > v.$$

Противоречие доказывает справедливость (2.4.1). Равенство (2.4.2) доказывается аналогично.

Данный результат является аналогом теоремы о дополняющей нежесткости п. 1.7.6. Напомним, что чистая стратегия x , входящая в спектр оптимальной стратегии, называется существенной. Таким образом, теорема утверждает, что для существенных стратегий должны быть выполнены равенства (2.4.1), (2.4.2).

Теорема п. 2.4.2 справедлива для любой непрерывной игры, поскольку справедливо следующее утверждение.

2.4.3. Лемма. *Если функция $H : X \times Y \rightarrow R^1$ непрерывна на $X \times Y$, то интегралы $K(\mu, y)$ и $K(x, \nu)$ являются соответственно непрерывными функциями от y и x для любых фиксированных смешанных стратегий $\mu \in \bar{X}$ и $\nu \in \bar{Y}$.*

Доказательство. Функция $H(x, y)$ непрерывна на компакте $X \times Y$, поэтому она равномерно непрерывна. Возьмем произвольное $\varepsilon > 0$ и найдем такое $\delta > 0$, что как только $\rho(y_1, y_2) < \delta$, то для любого x выполняется неравенство:

$$|H(x, y_1) - H(x, y_2)| < \varepsilon, \quad (2.4.4)$$

где $\rho(\cdot)$ — метрика в пространстве Y .

Тогда

$$\begin{aligned} |K(\mu, y_1) - K(\mu, y_2)| &= \left| \int_X H(x, y_1) d\mu(x) - \int_X H(x, y_2) d\mu(x) \right| = \\ &= \left| \int_X [H(x, y_1) - H(x, y_2)] d\mu(x) \right| \leq \int_X |H(x, y_1) - H(x, y_2)| d\mu(x) < \varepsilon \int_X d\mu(x) = \varepsilon. \end{aligned} \quad (2.4.5)$$

Следовательно, функция $K(\mu, y)$ непрерывна по y .

Аналогично доказывается непрерывность функции $K(x, \nu)$ по x .

2.4.4. Сформулируем основную теорему данного параграфа.

Теорема. *Бесконечная антагонистическая игра $\Gamma = (X, Y, H)$, где X, Y — метрические компакты, а H — непрерывная функция на их произведении, имеет решение в смешанных стратегиях (значение и оптимальные стратегии).*

Доказательство теоремы основано на аналитических свойствах смешанного расширения игры $\bar{\Gamma} = (\bar{X}, \bar{Y}, K)$ и некоторых вспомогательных результатах.

2.4.5. Напомним, что последовательность борелевских мер μ_n , $n = 1, 2, \dots$, заданных на борелевской σ -алгебре \mathcal{X} компактного метрического пространства X , называется слабо сходящейся, если

$$\lim_{n \rightarrow \infty} \int_X \varphi(x) d\mu_n(x) = \int_X \varphi(x) d\mu(x) \quad (2.4.6)$$

для любой непрерывной функции $\varphi(x)$, $x \in X$.

Лемма. *В условиях теоремы п. 2.4.4 множества смешанных стратегий \bar{X} и \bar{Y} (множества борелевских вероятностных мер) — метрические компакты в топологии слабой сходимости.*

Приведем схему доказательства для множества смешанных стратегий \bar{X} (для \bar{Y} рассуждения аналогичны). Пространство борелевских мер \bar{X} , заданных на борелевской σ -алгебре \mathcal{X} компактного метрического пространства X , метризуемо, поскольку в \bar{X} можно ввести метрику

$$\rho(\mu', \mu'') = \max(\rho', \rho''),$$

где ρ' и ρ'' — нижние границы таких чисел r' и r'' соответственно, что для любого замкнутого множества $F \subseteq X$

$$\mu'(F) < \mu''(V_{r'}(F)) + r', \quad \mu''(F) < \mu'(V_{r''}(F)) + r'',$$

где $V_r(F) = \{x \in X : \min_{z \in F} \rho_1(x, z) < r\}$, $r > 0$, и $\rho_1(\cdot)$ — метрика в пространстве X .

Известно [Brans, 1994], что сходимость в этом метрическом пространстве равносильна слабой сходимости, а семейство мер μ на борелевской σ -алгебре пространства X слабо компактно (т. е. компактно в описанном выше метрическом пространстве всех борелевских мер) тогда и только тогда, когда это семейство равномерно ограничено

$$\mu(X) \leq c \quad (2.4.7)$$

и равномерно плотно, т. е. для любого $\varepsilon > 0$ существует такой компакт $A \subseteq X$, что

$$\mu(X \setminus A) \leq \varepsilon. \quad (2.4.8)$$

Условие (2.4.8) следует из компактности X , а (2.4.7) — из того, что меры $\mu \in \overline{X}$ нормированы ($\mu(X) = 1$).

2.4.6. Заметим, что в условиях теоремы п. 2.4.4 множество смешанных стратегий $\overline{X}(\overline{Y})$ игрока 1 (2) является компактом и в обычном смысле, поскольку в данном случае слабая сходимость последовательности мер $\{\mu_n\}$, $n = 1, 2, \dots$ равносильна сходимости в обычном смысле:

$$\lim_{n \rightarrow \infty} \mu_n(A) = \mu(A)$$

для любого борелевского множества $A \subseteq X$, такого, что его граница A' имеет меру нуль: $\mu(A') = 0$.

Доказательство этого результата представляет определенные технические сложности. Его можно найти, например, в [Дрешер, 1964].

2.4.7. Обозначим через \underline{v} и \overline{v} соответственно нижнее и верхнее значения игры $\Gamma = (X, Y, H)$.

$$\underline{v} = \sup_{\mu} \inf_y K(\mu, y), \quad \overline{v} = \inf_{\nu} \sup_x K(x, \nu). \quad (2.4.9)$$

Лемма. В условиях теоремы п. 2.4.4 экстремумы в (2.4.9) достигаются, поэтому

$$\underline{v} = \max_{\mu \in \overline{X}} \min_{y \in Y} K(\mu, y), \quad \overline{v} = \min_{\nu \in \overline{Y}} \max_{x \in X} K(x, \nu). \quad (2.4.10)$$

Доказательство. Поскольку $H(x, y)$ непрерывна, то по лемме п. 2.4.3 для любой меры $\mu \in \overline{X}$ функция

$$K(\mu, y) = \int_X H(x, y) d\mu(x)$$

непрерывна по y . Так как Y — компакт, то $K(\mu, y)$ в некоторой его точке будет достигать минимума.

По определению \underline{v} , для любого n существует такая мера $\mu_n \in \overline{X}$, что

$$\min_y K(\mu_n, y) \geq \underline{v} - 1/n.$$

Так как \overline{X} — компакт в топологии слабой сходимости (лемма п. 2.4.5), то из последовательности $\{\mu_n\}_{n=1}^{\infty}$, $\mu_n \in \overline{X}$ можно выбрать слабо сходящуюся подпоследовательность. Пусть сама последовательность $\{\mu_n\}_{n=1}^{\infty}$ слабо сходится к некоторой мере $\mu_0 \in \overline{X}$. Тогда

$$\begin{aligned} \lim_{n \rightarrow \infty} K(\mu_n, y) &= \lim_{n \rightarrow \infty} \int_X H(x, y) d\mu_n(x) = \\ &= \int_X H(x, y) d\mu_0(x) = K(\mu_0, y), \quad y \in Y. \end{aligned}$$

Но $K(\mu_0, y)$ не меньше \underline{v} для каждого $y \in Y$. Следовательно, $\min_y K(\mu_0, y) \geq \underline{v}$ и на $\mu_0 \in \bar{X}$ достигается требуемый максимум.

Аналогично доказывается, что $\inf \sup$ в (2.4.9) можно заменить на $\min \max$.

2.4.8. Перейдем непосредственно к доказательству теоремы п. 2.4.4.

Доказательство. Так как X и Y — метрические компакты, то для любого целого n существуют конечные $(1/n)$ -сети

$$X_n = \{x_1^n, \dots, x_{r_n}^n\}, \quad X_n \subset X, \quad Y_n = \{y_1^n, \dots, y_{s_n}^n\}, \quad Y_n \subset Y$$

соответственно множеств X и Y . Это означает, что для любых точек $x \in X$ и $y \in Y$ найдутся такие точки $x_i^n \in X_n$ и $y_j^n \in Y_n$, что

$$\rho_1(x, x_i^n) < 1/n, \quad \rho_2(y, y_j^n) < 1/n, \quad (2.4.11)$$

где $\rho_1(\cdot), \rho_2(\cdot)$ — метрики пространств X и Y соответственно.

Для произвольного целого n построим матричную игру с матрицей $A_n = \{\alpha_{ij}^n\}$, где

$$\alpha_{ij}^n = H(x_i^n, y_j^n), \quad x_i^n \in X_n, \quad y_j^n \in Y_n. \quad (2.4.12)$$

Игра с матрицей A_n имеет значение θ_n и оптимальные смешанные стратегии $p_n = (\pi_1^n, \dots, \pi_{r_n}^n)$, $t_n = (\tau_1^n, \dots, \tau_{s_n}^n)$ игроков 1 и 2 соответственно (см. теорему п. 1.6.1 гл. I).

Функция $H(x, y)$ непрерывна на декартовом произведении $X \times Y$ метрических компактов, поэтому она равномерно непрерывна, т. е. для заданного $\varepsilon > 0$ можно найти такое $\delta > 0$, что как только

$$\rho_1(x, x') < \delta, \quad \rho_2(y, y') < \delta,$$

то

$$|H(x, y) - H(x', y')| < \varepsilon. \quad (2.4.13)$$

Выберем n настолько большим, чтобы $1/n < \delta$, и определим стратегию $\mu_n \in \bar{X}$ по правилу

$$\mu_n(F) = \sum_{\{i | x_i^n \in F, x_i^n \in X_n\}} \pi_i^n \quad (2.4.14)$$

для каждого борелевского множества F пространства X . Таким образом, имеем

$$K(\mu_n, y_j^n) = \sum_{i=1}^{r_n} \alpha_{ij}^n \pi_i^n \geq \theta_n. \quad (2.4.15)$$

Если $\rho_2(y, y_j^n) < \delta$, то, согласно (2.4.4), (2.4.5) и (2.4.13), получаем

$$|H(x, y) - H(x, y_j^n)| < \varepsilon,$$

$$|K(\mu_n, y) - K(\mu_n, y_j^n)| < \varepsilon.$$

Следовательно, для любого $y \in Y$ (Y_n — $(1/n)$ -сеть множества Y) имеем

$$K(\mu_n, y) > \theta_n - \varepsilon. \quad (2.4.16)$$

Так как $\min_y K(\mu_n, y)$ достигается (лемма п. 2.4.7), то

$$\underline{v} > \theta_n - \varepsilon. \quad (2.4.17)$$

Аналогично можно показать, что

$$\bar{v} < \theta_n + \varepsilon. \quad (2.4.18)$$

Из (2.4.17) и (2.4.18) получаем

$$\underline{v} > \bar{v} - 2\varepsilon.$$

Но по лемме п. 1.2.2 гл. I неравенство $\underline{v} \leq \bar{v}$ выполняется всегда. Учитывая произвольность $\varepsilon > 0$, получаем

$$\underline{v} = \bar{v}, \quad (2.4.19)$$

тогда из леммы п. 2.4.7 и (2.4.19) следует утверждение теоремы (см. п. 2.2.1).

2.4.9. Следствие. *Имеет место равенство:*

$$v = \lim_{n \rightarrow \infty} \theta_n, \quad (2.4.20)$$

где $\theta_n = v(A_n)$ — значение матричной игры с матрицей A_n (2.4.12).

2.4.10. Из доказательства теоремы п. 2.4.4 следует, что непрерывную игру можно с любой степенью точности аппроксимировать конечными играми. Более того, справедлив следующий результат.

Теорема. *Бесконечная антагонистическая игра $\Gamma = (X, Y, H)$, где X, Y — метрические компакты, а H — непрерывная функция на их произведении, при любом $\varepsilon > 0$ имеет ε -оптимальные смешанные стратегии с конечным спектром.*

Доказательство теоремы следует из доказательства (п. 2.4.8) теоремы п. 2.4.4. Действительно, по игре Γ построим матричные игры с матрицами A_n и смешанные стратегии $\mu_n \in \bar{X}$, определяемые соответственно (2.4.12), (2.4.14) для произвольного целого n . Стратегии $\nu_n \in \bar{Y}$ игрока 2 определяются аналогичным образом:

$$\nu_n(G) = \sum_{\{j | y_j^n \in G, y_j^n \in Y_n\}} \nu_j^n, \quad (2.4.21)$$

где $t^n = (\tau_1^n, \dots, \tau_{s_n}^n)$ — оптимальная смешанная стратегия игрока 2 в игре с матрицей A_n и значением θ_n .

По построению имеем

$$\theta_n = \sum_{i=1}^{r_n} \sum_{j=1}^{s_n} \alpha_{ij}^n \pi_i^n \tau_j^n = K(\mu_n, \nu_n), \quad (2.4.22)$$

где $K(\mu, \nu)$ — выигрыш в смешанных стратегиях (μ, ν) в игре Γ . Из (2.4.16) и аналогичного неравенства для стратегии ν_n получаем, что для произвольного $\varepsilon > 0$ найдется номер n , такой, что

$$K(x, \nu_n) - \varepsilon < \theta_n < K(\mu_n, y) + \varepsilon \quad (2.4.23)$$

для всех $x \in X$ и $y \in Y$. Учитывая, что стратегии μ_n и ν_n имеют конечный спектр X_n и Y_n соответственно (X_n, Y_n — конечные $1/n$ -сети множеств X и Y соответственно), получаем утверждение теоремы (см. п. 2.3.4).

2.4.11. Объединяя результаты теорем п. 2.4.4 и 2.4.10, можно сделать вывод, что бесконечная антагонистическая игра с непрерывной функцией выигрыша и компактными множествами стратегий для любого $\varepsilon > 0$ имеет ε -оптимальные стратегии игроков, являющиеся смесями конечного числа чистых, а также смешанные оптимальные стратегии в классе борелевских вероятностных мер. В частности, эти результаты справедливы для игр на квадрате (п. 2.1.3) с непрерывной функцией выигрыша.

2.4.12. Имеется большое число работ, в которых доказываются существование значения бесконечных антагонистических игр. Наиболее общий результат в этом направлении принадлежит Сайону. Для игр с компактными пространствами стратегий и полунепрерывными функциями выигрыша известны результаты [Пек, Далмидж], [Яновская, 1973], [Petrosyan, 1993]. Покажем, что в некоторых направлениях они не поддаются обобщению.

Пример 10 (Игра на квадрате, не имеющая значения в смешанных стратегиях) [Сайон, Вульф]. Рассматривается антагонистическая игра $\Gamma = (X, Y, H)$, где $X = Y = [0, 1]$, а функция выигрыша H имеет вид

$$H(x, y) = \begin{cases} -1, & \text{если } x < y < x + 1/2, \\ 0, & \text{если } x = y \text{ или } y = x + 1/2, \\ 1, & \text{если } y < x \text{ или } x + 1/2 < y. \end{cases}$$

Эта функция имеет разрывы на прямых $y = x$ и $y = x + 1/2$. Покажем, что

$$\sup_{\mu} \inf_{\nu} K(\mu, \nu) = 1/3, \quad \inf_{\mu} \sup_{\nu} K(\mu, \nu) = 3/7. \quad (2.4.24)$$

Пусть μ — вероятностная мера на $[0, 1]$. Если $\mu([0, 1/2)) \leq 1/3$, то положим $y_{\mu} = 1$. Если же $\mu([0, 1/2)) > 1/3$, то выберем $\delta > 0$, такое, что $\mu([0, 1/2 - \delta]) > 1/3$, и положим $y_{\mu} = 1/2 - \delta$. В каждом из этих случаев получаем неравенства

$$\inf_{\nu} K(\mu, \nu) \leq K(\mu, y_{\mu}) \leq 1/3,$$

которые доказываются непосредственной проверкой.

С другой стороны, если μ выбрано так, что $\mu(\{0\}) = \mu(\{1/2\}) = \mu(\{1\}) = 1/3$, то для всех $y \in [0, 1]$ имеем

$$\int_0^1 H(x, y) d\mu(x) = 1/3 [H(0, y) + H(1/2, y) + H(1, y)] \geq 1/3.$$

Следовательно, доказано первое из равенств (2.4.24).

Теперь пусть ν — какая-либо вероятностная мера на $[0, 1]$. Если $\nu([0, 1)) \geq 3/7$, то положим $x_{\nu} = 1$. Если $\nu([0, 1)) < 3/7$, то $\nu(\{1\}) > 4/7$, и в этом случае положим $x_{\nu} = 0$, если $\nu([0, 1/2)) \leq 1/7$; если же $\nu([0, 1/2)) > 1/7$, то выберем $\delta > 0$ так, чтобы $\nu([0, 1/2 - \delta]) > 1/7$, и положим $x_{\nu} = 1/2 - \delta$. В каждом из указанных случаев убеждаемся, что

$$\sup_{\mu} K(\mu, \nu) > K(x_{\nu}, \nu) > 3/7.$$

С другой стороны, если ν выбрано так, что

$$\nu(\{1/4\}) = 1/7, \quad \nu(\{1/2\}) = 2/7, \quad \nu(\{1\}) = 4/7,$$

то для любого $x \in [0, 1]$ имеем

$$\int_0^1 H(x, y) d\nu(y) = \frac{1}{7} [H(x, \frac{1}{4}) + 2H(x, \frac{1}{2}) + 4H(x, 1)] \leq \frac{3}{7}.$$

Таким образом, доказано второе из равенств (2.4.24).

§ 2.5. Игры с выпуклой функцией выигрыша

В § 2.4 при достаточно общих предположениях было доказано существование решения в бесконечных антагонистических играх с непрерывной функцией выигрыша и компактными множествами стратегий. Вместе с тем представляет теоретический и практический интерес выделение таких классов игр, когда один или оба игрока имеют оптимальные чистые стратегии. Такие игры рассматриваются в данном параграфе.

2.5.1. Определение. Пусть $X \subset R^m$, $Y \subset R^n$ — компакты, множество Y — выпукло, функция $H : X \times Y \rightarrow R^1$ непрерывна по совокупности аргументов и выпукла по $y \in Y$ при любом фиксированном значении $x \in X$. Тогда игра $\Gamma(X, Y, H)$ называется игрой с выпуклой функцией выигрыша (выпуклая игра).

Приведем симметричное определение относительно игрока 1.

Определение. Если $X \subset R^m$, $Y \subset R^n$ — компакты, множество X выпукло, функция выигрыша H непрерывна по совокупности аргументов и вогнута по $x \in X$ при любом фиксированном $y \in Y$, то игра $\Gamma = (X, Y, H)$ называется игрой с вогнутой функцией выигрыша (вогнутая игра).

Если же $X \subset R^m$, $Y \subset R^n$ — выпуклые компакты, а непрерывная по совокупности аргументов функция выигрыша $H(x, y)$ вогнута по x при любом фиксированном y и выпукла по y при каждом x , то игра $\Gamma(X, Y, H)$ называется игрой с вогнуто-выпуклой функцией выигрыша (вогнуто-выпуклая игра).

Рассмотрим игры с выпуклой функцией выигрыша. Аналогичные результаты справедливы и для вогнутых игр.

Теорема. Пусть $\Gamma = (X, Y, H)$ — выпуклая игра. Тогда игрок 2 имеет оптимальную чистую стратегию, при этом значение игры равно

$$v = \min_{y \in Y} \max_{x \in X} H(x, y). \quad (2.5.1)$$

Доказательство. Так как X и Y — метрические компакты (в метрике евклидовых пространств R^m и R^n), а функция H непрерывна на произведении $X \times Y$, то, согласно теореме п. 2.4.4, в игре Γ существует значение v и оптимальные смешанные стратегии μ^*, ν^* . Известно, что множество вероятностных мер с конечным носителем всюду плотно в множестве всех вероятностных мер на Y [Прохоров, Рязанов, 1967]. Поэтому существует последовательность смешанных стратегий ν^n с конечным спектром, слабо сходящаяся к ν^* . Пусть спектр стратегии ν^n состоит из точек $y_n^1, \dots, y_n^{k_n}$ и они выбираются с вероятностями $\eta_n^1, \dots, \eta_n^{k_n}$. Тогда в силу выпуклости функции H имеем

$$K(x, \nu^n) = \sum_{j=1}^{k_n} \eta_j^n H(x, y_n^j) \geq H(x, \bar{y}^n), \quad (2.5.2)$$

где $\bar{y}^n = \sum_{j=1}^{k_n} \eta_j^n y_n^j$. Переходя к пределу при $n \rightarrow \infty$ в неравенстве (2.5.2) (если необходимо, то следует рассмотреть подпоследовательность $\{\bar{y}^n\}$), получаем

$$K(x, \nu^*) \geq H(x, \bar{y}), \quad x \in X \quad (2.5.3)$$

где \bar{y} — предельная точка последовательности $\{\bar{y}^n\}$. Из (2.5.3) и леммы п. 2.4.3 имеем

$$\max_x K(x, \nu^*) \geq \max_x H(x, \bar{y}). \quad (2.5.4)$$

Пусть неравенство (2.5.4) строгое. Тогда

$$v = \max_x K(x, \nu^*) > \max_x H(x, \bar{y}) \geq \min_{\nu} \max_x H(x, \nu) = v,$$

что невозможно. Таким образом, $\max_x H(x, \bar{y}) = \max_x K(x, \nu^*) = v$ и из теоремы п. 2.3.5 получаем, что \bar{y} — оптимальная стратегия игрока 2.

Установим справедливость равенства (2.5.1). Так как $\bar{y} \in Y$ — оптимальная стратегия игрока 2, то

$$v = \max_x H(x, \bar{y}) \geq \min_y \max_x H(x, y).$$

С другой стороны, выполняется неравенство

$$v = \min_{\nu} \max_x K(x, \nu) \leq \min_y \max_x H(x, y).$$

Сравнивая последние неравенства, получаем (2.5.1).

2.5.2. Напомним, что функция $\varphi : Y \rightarrow R^1$, где $Y \subset R_n$, Y — выпуклое множество, *строго выпукла*, если для всех $\lambda \in (0, 1)$ выполняется строгое неравенство:

$$\varphi(\lambda y_1 + (1 - \lambda)y_2) < \lambda\varphi(y_1) + (1 - \lambda)\varphi(y_2), \quad y_1, y_2 \in Y, \quad y_1 \neq y_2.$$

Теорема. Пусть $\Gamma = (X, Y, H)$ — выпуклая игра со строго выпуклой функцией выигрыша. Тогда игрок 2 имеет единственную оптимальную стратегию, которая является чистой.

Доказательство. Пусть μ^* — оптимальная стратегия игрока 1, $\varphi(y) = K(\mu^*, y)$ и v — значение игры. Если \bar{y} — точка спектра оптимальной стратегии игрока 2, то выполняется равенство (2.4.2):

$$K(\mu^*, \bar{y}) = v.$$

Однако для всех $y \in Y$ имеем неравенство $K(\mu^*, y) \geq v$, поэтому

$$\varphi(\bar{y}) = \min_{y \in Y} \varphi(y) = v.$$

Функция $\varphi(y)$ является строго выпуклой, поскольку для $\lambda \in (0, 1)$ имеет место неравенство:

$$\begin{aligned} \varphi(\lambda y_1 + (1 - \lambda)y_2) &= \int_X H(x, \lambda y_1 + (1 - \lambda)y_2) d\mu^*(x) < \\ &< \lambda \int_X H(x, y_1) d\mu^*(x) + (1 - \lambda) \int_X H(x, y_2) d\mu^*(x) = \lambda\varphi(y_1) + (1 - \lambda)\varphi(y_2). \end{aligned} \quad (2.5.5)$$

Из (2.5.5) следует, что функция $\varphi(y)$ не может достигать минимума в двух различных точках. С другой стороны, существование точки минимума \bar{y} функции $\varphi(y)$ гарантируется теоремой п. 2.5.1, что завершает доказательство.

2.5.3. Приведем без доказательства результаты, симметричные теоремам по п. 2.5.1 и 2.5.2 для вогнутых и вогнуто-выпуклых игр. **Теорема.** Пусть $\Gamma = (X, Y, H)$, $X \subset R^m$, $Y \subset R^n$ — вогнутая игра. Тогда значение игры v вычисляется по формуле

$$v = \max_x \min_y H(x, y) \quad (2.5.6)$$

а каждая чистая стратегия x^* , на которой достигается $\max \min$ (2.5.6), является оптимальной для игрока 1. Если, кроме того, функция $H(x, y)$ строго вогнута по x при каждом фиксированном $y \in Y$, то оптимальная стратегия игрока 1 единственна.

Теорема. Пусть $\Gamma = (X, Y, H)$, $X \subset R^m$, $Y \subset R^n$ — *вогнута-выпуклая игра*. Тогда значение игры v равно

$$v = \min_y \max_x H(x, y) = \max_x \min_y H(x, y). \quad (2.5.7)$$

В игре Γ всегда существует ситуация равновесия (x^*, y^*) , где $x^* \in X$, $y^* \in Y$ — чистые стратегии игроков 1 и 2, на которых достигаются внешние экстремумы в (2.5.7). Если при этом функция $H(x, y)$ строго вогнута (выпукла) по переменной $x(y)$ при любом фиксированном $y \in Y$ ($x \in X$), то игрок 1 (2) имеет единственную оптимальную стратегию, которая является чистой.

2.5.4. Выясним структуру оптимальной стратегии игрока 1 в *выпуклой* игре $\Gamma = (X, Y, H)$.

Теорема. В *выпуклой* игре $\Gamma = (X, Y, H)$, $Y \subset R^n$, игрок 1 имеет оптимальную смешанную стратегию μ^* с конечным спектром, состоящим не более чем из $(n+1)$ -й точки множества X .

Доказательство этого результата основано на известной теореме Хелли о выпуклых множествах, которую мы приведем без доказательства [Рокафеллар, 1973, с. 210; Давыдов, 1978, с. 107].

Теорема Хелли. Пусть K — семейство из не менее чем $n+1$ выпуклого множества в R^n , причем каждое множество из K компактно. Тогда, если каждые $n+1$ из множества семейства K имеют общую точку, то существует точка, общая всем множествам семейства K .

Прежде чем перейти непосредственно к доказательству теоремы (2.5.5), докажем ряд вспомогательных утверждений.

Пусть функция $H(x, y)$ непрерывна на произведении $X \times Y$ компактных множеств $X \subset R^m$, $Y \subset R^n$. Обозначим как $X^r = X \times \dots \times X$ декартово произведение r множеств X .

Рассмотрим функцию $\varphi : X^r \times Y \rightarrow R^1$:

$$\varphi(x_1, \dots, x_r, y) = \max_{1 \leq i \leq r} H(x_i, y).$$

Лемма. Функция $\varphi(x_1, \dots, x_r, y)$ непрерывна на $X^r \times Y$.

Доказательство. Функция $H(x, y)$ непрерывна на компактном множестве $X \times Y$, поэтому и равномерно непрерывна на нем. Тогда для любого $\varepsilon > 0$ найдется $\delta > 0$, такое, что из неравенств $\rho_1(\bar{x}, \bar{\bar{x}}) < \delta$, $\rho_2(y_1, y_2) < \delta$ следует неравенство $|H(\bar{x}, y_1) - H(\bar{\bar{x}}, y_2)| < \varepsilon$, где $\rho_1(\cdot)$, $\rho_2(\cdot)$ — расстояния в R^m и R^n соответственно. Имеем

$$\begin{aligned} & |\varphi(\bar{x}_1, \dots, \bar{x}_r, y_1) - \varphi(\bar{\bar{x}}_1, \dots, \bar{\bar{x}}_r, y_2)| = \\ & = \left| \max_{1 \leq i \leq r} H(\bar{x}_i, y_1) - \max_{1 \leq i \leq r} H(\bar{\bar{x}}_i, y_2) \right| = |H(\bar{x}_{i_1}, y_1) - H(\bar{\bar{x}}_{i_2}, y_2)|, \end{aligned}$$

где

$$H(\bar{x}_{i_1}, y_1) = \max_{1 \leq i \leq r} H(\bar{x}_i, y_1), \quad H(\bar{\bar{x}}_{i_2}, y_2) = \max_{1 \leq i \leq r} H(\bar{\bar{x}}_i, y_2).$$

Если $\rho_1(\bar{x}_i, x_i) < \delta$ для $i = 1, \dots, r$, $\rho_2(y_1, y_2) < \delta$ и если $H(\bar{x}_{i_1}, y_1) \geq H(\bar{\bar{x}}_{i_2}, y_2)$, то

$$0 \leq H(\bar{x}_{i_1}, y_1) - H(\bar{\bar{x}}_{i_2}, y_2) \leq H(\bar{x}_{i_1}, y_1) - H(\bar{\bar{x}}_{i_1}, y_2) < \varepsilon.$$

Аналогичные неравенства имеют место в случае $H(\bar{x}_{i_1}, y_1) \leq H(\bar{\bar{x}}_{i_2}, y_2)$.

Лемма. В выпуклой игре $\Gamma = (X, Y, H)$, $Y \subset R^n$ значение игры v равно

$$v = \min_y \max_x H(x, y) = \max_{x_1, \dots, x_{n+1}} \min_y \max_{1 \leq i \leq n+1} H(x_i, y), \quad (2.5.8)$$

где $y \in Y$, $x_i \in X$, $i = 1, \dots, n+1$.

Доказательство. Введем обозначение:

$$\Theta = \max_{x_1, \dots, x_{n+1}} \min_y \max_{1 \leq i \leq n+1} H(x_i, y).$$

Так как $\min_y \max_{1 \leq i \leq n+1} H(x_i, y) \leq \min_y \max_x H(x, y) = v$ для каждой системы точек $(x_1, \dots, x_{n+1}) \in X^{n+1}$, то

$$\Theta \leq v. \quad (2.5.9)$$

Для произвольного фиксированного набора стратегий $x_i \in X$, $i = 1, \dots, n+1$, рассмотрим систему неравенств относительно y :

$$H(x_i, y) \leq \Theta, \quad y \in Y, \quad i = 1, \dots, n+1. \quad (2.5.10)$$

Покажем, что система (2.5.10) имеет решение.

Действительно,

$$\Theta \geq \min_y \max_{1 \leq i \leq n+1} H(x_i, y) = \max_{1 \leq i \leq n+1} H(x_i, \bar{y}) \geq H(x_i, \bar{y}), \quad i = 1, \dots, n+1.$$

Таким образом, \bar{y} удовлетворяет системе (2.5.10).

Следовательно, система (2.5.10) имеет решение для любых $x_i \in X$, $i = 1, 2, \dots, n+1$.

Зафиксируем x и рассмотрим множество

$$D_x = \{y \mid H(x, y) \leq \Theta\}.$$

Функция $H(x, y)$ выпукла и непрерывна по y , поэтому множество D_x выпукло и замкнуто при каждом x . Множества $\{D_x\}$ образуют систему выпуклых компактных множеств в R^n , причем в силу того, что неравенства (2.5.10) всегда имеют решение, любой набор по $(n+1)$ -му множеству системы $\{D_x\}$ имеет непустое пересечение. Поэтому по теореме Хелли существует точка $y_0 \in Y$, общая для всех множеств D_x , т. е. такая, что

$$H(x, y_0) \leq \Theta \quad (2.5.11)$$

при любых $x \in X$. Предположим, что $\Theta \neq v$. Тогда из (2.5.9) и (2.5.11) имеем

$$\Theta < v = \min_y \max_x H(x, y) \leq \max_x H(x, y_0) \leq \Theta,$$

т. е. $\Theta < \Theta$. Полученное противоречие и доказывает (2.5.8).

Перейдем к доказательству теоремы.

Доказательство. Из предыдущей леммы имеем

$$\begin{aligned} v &= \max_{x_1, \dots, x_{n+1}} \min_y \max_{1 \leq i \leq n+1} H(x_i, y) = \min_y \max_{1 \leq i \leq n+1} H(\bar{x}_i, y) = \\ &= \min_y \max_p \sum_{i=1}^{n+1} H(\bar{x}_i, y) \pi_i, \end{aligned} \quad (2.5.12)$$

где $\bar{x}_1, \dots, \bar{x}_{n+1}$ — векторы, на которых достигается внешний максимум в (2.5.8),

$$p = (\pi_1, \dots, \pi_{n+1}) \in R^{n+1}, \quad \pi_i \geq 0, \quad \sum_{i=1}^{n+1} \pi_i = 1. \quad (2.5.13)$$

Рассмотрим функцию

$$K(p, y) = \sum_{i=1}^{n+1} H(x_i, y) \pi_i, \quad y \in Y, \quad p \in P,$$

где P состоит из векторов, удовлетворяющих (2.5.13). Функция $K(p, y)$ непрерывна по p и y , выпукла по y и вогнута по p , а множества $Y \subset R^n$, $P \subset R^{n+1}$ — компакты в соответствующих евклидовых пространствах. Поэтому по теореме п. 2.5.3 и из (2.5.12) имеем

$$v = \min_y \max_p \sum_{i=1}^{n+1} H(\bar{x}_i, y) \pi_i = \max_p \min_y \sum_{i=1}^{n+1} H(\bar{x}_i, y) \pi_i. \quad (2.5.14)$$

Из (2.5.8) и (2.5.14) следует существование таких $p^* \in P$ и $y^* \in Y$, что для всех $x \in X$ и $y \in Y$ выполняется неравенство:

$$H(x, y^*) \leq v \leq \sum_{i=1}^{n+1} H(\bar{x}_i, y) \pi_i^*.$$

Теорема доказана.

Сформулируем теорему о структуре оптимальной стратегии игрока 2 в вогнутой игре $\Gamma = (X, Y, H)$.

Теорема. В вогнутой игре $\Gamma = (X, Y, H)$, $X \subset R^m$, игрок 2 имеет оптимальную смешанную стратегию v^* с конечным спектром, состоящим не более чем из $(m+1)$ -й точки множества Y .

Доказательство теоремы аналогично доказательству предыдущей теоремы.

2.5.5. Суммируем результаты теорем для выпуклых игр, доказанные в этом параграфе.

Теорема. Пусть $\Gamma = (X, Y, H)$, $X \subset R^m$, $Y \subset R^n$ — выпуклая игра. Тогда значение v игры Γ определяется по формуле

$$v = \min_y \max_x H(x, y).$$

Игрок 1 обладает оптимальной смешанной стратегией μ_0 с конечным спектром, состоящим не более чем из $(n+1)$ -й точки множества X . В то же время все чистые стратегии y_0 , на которых достигается $\min_y \max_x H(x, y)$, являются оптимальными

для игрока 2. Если, кроме того, функция $H(x, y)$ при каждом фиксированном $x \in X$ строго выпукла по y , то оптимальная стратегия игрока 2 единственна.

Проиллюстрируем эти результаты на примере.

Пример 11. Рассмотрим частный случай примера 1 (см. п. 2.1.2). Пусть $S_1 = S_2 = S$, а множество S представляет собой замкнутый круг на плоскости с центром в точке O и радиусом R .

Функция выигрыша $H(x, y) = \rho(x, y)$, $x \in S$, $y \in S$, где $\rho(\cdot)$ — функция расстояния в R^2 , является строго выпуклой по y при любом фиксированном x , а S — выпуклое множество. Поэтому согласно теореме п. 2.5.5 значение игры v равно

$$v = \min_{y \in S} \max_{x \in S} \rho(x, y). \quad (2.5.15)$$

Вычисляя $\min \max$ в (2.5.15), получаем, что $v = R$ (см. пример 8 п. 2.2.6). При этом точка $y_0 \in S$, на которой достигается минимум выражения $\max_{x \in S} \rho(x, y)$, единственная и совпадает с центром круга S (т. е. точкой O). Эта точка и является оптимальной стратегией игрока 2 (минимизирующего). Теорема утверждает, что у игрока 1 (максимизирующего) существует оптимальная смешанная стратегия, предписывающая положительную вероятность не более чем трем точкам множества S . Однако вследствие симметрии множества S в действительности оптимальная смешанная стратегия μ_0 игрока 1 предписывает с вероятностью $1/2$ выбирать любые две диаметрально противоположные точки на границе множества S . Для доказательства оптимальности стратегий μ_0, y_0 достаточно установить, что $K(x, y_0) \leq K(\mu_0, y_0) \leq K(\mu_0, y)$ для всех $x, y \in S$, где K — математическое ожидание выигрыша, $K(\mu_0, y_0) = R/2 + R/2 = R$. Действительно, $K(x, y_0) = \rho(0, x) \leq R$ и $K(\mu_0, y) = \rho(x_1, y)/2 + \rho(x_2, y)/2 \geq R$, где x_1 и x_2 — произвольные диаметрально противоположные точки на границе круга S . Оптимальность стратегий μ_0 и y_0 доказана.

2.5.6. Рассмотрим частный случай выпуклой игры $\Gamma = (X, Y, H)$ когда $X = Y = [0, 1]$, т. е. выпуклую игру на единичном квадрате. Из теоремы п. 2.5.5 следует, что игрок 2 всегда имеет оптимальную чистую стратегию $y_0 \in [0, 1]$, а игрок 1 — смешанную, сосредоточенную не более чем на двух точках, при этом значение игры равно

$$v = \min_{y \in [0, 1]} \max_{x \in [0, 1]} H(x, y). \quad (2.5.16)$$

Множество всех существенных стратегий $\{x\} \subset [0, 1]$ игрока 1 является подмножеством решений уравнений (п. 2.4.2)

$$H(x, y_0) = v, \quad x \in [0, 1], \quad (2.5.17)$$

где y_0 — оптимальная стратегия игрока 2. Чистые стратегии (2.5.17) игрока 1, удовлетворяющие равенству (2.5.17), иногда называются *уравновешивающими*. Множество всех уравновешивающих стратегий игрока 1 замкнуто и ограничено, т. е. компактно. Оптимальной чистой стратегией игрока 2 является любая точка $y_0 \in [0, 1]$, на которой достигается (2.5.16).

Обозначим через $H'_y(x, y)$ частную производную функции H по y (при $y = 0$ и $y = 1$ понимается соответственно правая и левая производные).

Лемма. Если y_0 — оптимальная стратегия игрока 2 в выпуклой игре на единичном квадрате с функцией выигрыша H , дифференцируемой по y и $y_0 > 0$, то найдется уравновешивающая стратегия x' игрока 1, для которой

$$H'_y(x', y_0) \leq 0. \quad (2.5.18)$$

Если же $y_0 < 1$, то существует такая уравнивающая стратегия x'' игрока 1, что

$$H'_y(x'', y_0) \geq 0. \quad (2.5.19)$$

Доказательство. Докажем (2.5.18). (Вторая часть леммы доказывается аналогично.) Предположим противное, а именно: для каждой уравнивающей стратегии \bar{x} игрока 1 выполняется неравенство $H'_y(\bar{x}, y_0) > 0$, т. е. функция $H(\bar{x}, \cdot)$ в точке y_0 строго возрастает. Это означает, что найдутся такие $\varepsilon(\bar{x}) > 0$ и $\Theta(\bar{x}) > 0$, что для $y \in [0, 1]$ удовлетворяющих неравенству $\Theta(\bar{x}) > y_0 - y > 0$, выполняется неравенство:

$$H(\bar{x}, y) < H(\bar{x}, y_0) - \varepsilon(\bar{x}).$$

В силу непрерывности функции H имеем, что для каждой уравнивающей стратегии \bar{x} и $\varepsilon(\bar{x})/2$ найдется такое $\delta(\bar{x}) > 0$, что при $\Theta(\bar{x}) > y_0 - y > 0$ выполняется неравенство

$$H(x, y) < H(\bar{x}, y) - \frac{\varepsilon(\bar{x})}{2} < H(\bar{x}, y_0) - \frac{\varepsilon(\bar{x})}{2} = H(x, y_0) - \frac{\varepsilon(\bar{x})}{2}$$

для всех, уравнивающих стратегий x , для которых $|x - \bar{x}| < \delta(\bar{x})$. Множество уравнивающих стратегий компактно, поэтому его можно покрыть конечным числом таких $\delta(\bar{x})$ -окрестностей. Пусть ε — наименьшее из всех соответствующих чисел $\varepsilon(\bar{x})$. Тогда имеем неравенство, справедливое для всех уравнивающих стратегий x (в том числе и для всех существенных стратегий)

$$H(x, y) \leq H(x, y_0) - \frac{\varepsilon}{2},$$

где

$$y_0 - \min \Theta(\bar{x}) < y < y_0.$$

Пусть μ_0 — оптимальная смешанная стратегия игрока 1. Последнее неравенство справедливо для всех точек спектра стратегии μ_0 , поэтому, интегрируя, получаем

$$K(\mu_0, y) \leq K(\mu_0, y_0) - \frac{\varepsilon}{2} = v - \frac{\varepsilon}{2},$$

что противоречит оптимальности стратегии μ_0 .

Теорема. Пусть Γ — выпуклая игра на единичном квадрате с функцией выигрыша H , дифференцируемой по y при любом x , y_0 — чистая оптимальная стратегия игрока 2, а v — значение игры. Тогда:

- 1) если $y_0 = 1$, то среди оптимальных стратегий игрока 1 имеется чистая стратегия x' , для которой выполняется (2.5.18);
- 2) если $y_0 = 0$, то среди оптимальных стратегий игрока 1 имеется чистая стратегия x'' , для которой выполняется (2.5.19);
- 3) если $0 < y_0 < 1$, то среди оптимальных стратегий игрока 1 найдется такая, которая является смесью двух существенных стратегий x' и x'' , удовлетворяющих (2.5.18), (2.5.19) с вероятностями α и $1 - \alpha$, $\alpha \in [0, 1]$.

При этом α является решением уравнения

$$\alpha H'_y(x', y_0) + (1 - \alpha) H'_y(x'', y_0) = 0. \quad (2.5.20)$$

Доказательство. Пусть $y_0 = 1$. Тогда найдется уравнивающая стратегия x' игрока 1, для которой выполняется (2.5.18). Тогда из выпуклости функции $H(x', y)$ следует, что она не возрастает по y на всем промежутке $[0, 1]$, достигая при $y = 1$ своего минимума. Это означает, что

$$H(x', y_0) \leq H(x', y) \quad (2.5.21)$$

при всех $y \in [0, 1]$. С другой стороны, из (2.5.17) следует, что

$$H(x, y_0) \leq H(x', y_0) \quad (2.5.22)$$

при всех $x \in [0, 1]$. Неравенства (2.5.21), (2.5.22) показывают, что (x', y_0) — ситуация равновесия.

Случай $y_0 = 0$ исследуется аналогично. Перейдем к случаю 3. Если $0 < y_0 < 1$, то имеются две уравнивающие стратегии x' и x'' , удовлетворяющие (2.5.18), (2.5.19) соответственно.

Рассмотрим функцию

$$\varphi(\beta) = \beta H'_y(x', y_0) + (1 - \beta) H'_y(x'', y_0).$$

Из (2.5.18), (2.5.19) следует, что $\varphi(0) \geq 0$, $\varphi(1) \leq 0$. Функция $\varphi(\beta)$ непрерывна, поэтому найдется $\alpha \in [0, 1]$, для которого $\varphi(\alpha) = 0$.

Рассмотрим смешанную стратегию μ_0 игрока 1, заключающуюся в выборе стратегии x' с вероятностью α и стратегии x'' с вероятностью $(1 - \alpha)$. Функция

$$K(\mu_0, y) = \alpha H(x', y) + (1 - \alpha) H(x'', y)$$

выпукла по y . Ее производная по y в точке $y = y_0$ равна

$$K'_y(\mu_0, y_0) = \alpha H'_y(x', y_0) + (1 - \alpha) H'_y(x'', y_0) = 0.$$

Следовательно, в точке y_0 функция $K(\mu_0, y)$ достигает минимума. Отсюда, учитывая (2.5.17), имеем

$$K(\mu_0, y_0) \leq K(\mu_0, y), \quad K(\mu_0, y) = H(x, y_0) = v = \max_x H(x, y_0) \geq H(x, y_0)$$

при всех $x \in [0, 1]$ и $y \in [0, 1]$, что и доказывает оптимальность стратегий μ_0 и y_0 .

2.5.7. Теорема п. 2.5.6 дает способ отыскания оптимальных стратегий, который мы проиллюстрируем на примере.

Пример 12. Рассмотрим игру на единичном квадрате с функцией выигрыша $H(x, y) = (x - y)^2$. Это есть одномерный аналог примера 11, только в качестве функции выигрыша здесь взят квадрат расстояния. Поэтому естественно ожидать, что значение v игры будет равно $v = 1/4$, оптимальной стратегией игрока 2 является середина отрезка $1/2$, а оптимальной стратегией игрока 1 — выбор с вероятностью $1/2$ крайних точек 0 и 1 интервала $[0, 1]$. Покажем это, используя теорему п. 2.5.6.

Заметим, что $\partial^2 H(x, y)/\partial y^2 = 2 > 0$, так что игра Γ — строго выпуклая, поэтому игрок 2 имеет единственную оптимальную стратегию, которая является чистой (теорема п. 2.5.5). Пусть y — фиксированная стратегия игрока 2. Тогда

$$\max_x (x - y)^2 = \begin{cases} (1 - y)^2, & y \leq 1/2, \\ y^2, & y \geq 1/2. \end{cases}$$

Таким образом, из (2.5.16) получаем

$$v = \min \left\{ \min_{0 \leq y \leq \frac{1}{2}} (1 - y)^2, \min_{\frac{1}{2} \leq y \leq 1} y^2 \right\}.$$

Оба внутренних минимума достигаются на $y_0 = 1/2$ и принимают значение $1/4$. Поэтому $v = 1/4$, а $y_0 = 1/2$ — единственная оптимальная стратегия игрока 2.

Найдем оптимальную стратегию игрока 1. Для этого заметим, что $0 < y_0 < 1$ ($y_0 = 1/2$). Найдем существенные стратегии игрока 1. Уравнение (2.5.17) в данном случае принимает вид $(x - 1/2)^2 = 1/4$. Откуда $x_1 = 0$ и $x_2 = 1$, т. е. существенными для игрока 1 являются крайние точки отрезка $[0, 1]$.

Вычислим производные

$$H'_y(x_1, y_0) = 1 > 0, \quad H'_y(x_2, y_0) = -1 < 0.$$

Составим уравнение (2.5.20) относительно α . Имеем $2\alpha - 1 = 0$, откуда $\alpha = 1/2$. Таким образом, оптимальная стратегия игрока 1 состоит в выборе им чистых стратегий 0 и 1 с вероятностью $1/2$.

2.5.8. В заключение параграфа приведем результат, аналогичный п. 2.5.6 для вогнутой игры.

Теорема. Пусть Γ — вогнутая игра на единичном квадрате с функцией выигрыша H , дифференцируемой по x при любом фиксированном y , x_0 — чистая оптимальная стратегия игрока 1, а v — значение игры. Тогда:

- 1) если $x_0 = 1$, то среди оптимальных стратегий игрока 2 имеется чистая стратегия y' , для которой выполняется неравенство:

$$H'_x(x_0, y') \geq 0; \tag{2.5.23}$$

- 2) если $x_0 = 0$, то среди оптимальных стратегий игрока 2 имеется чистая стратегия y'' , для которой

$$H'_x(x_0, y'') \leq 0; \tag{2.5.24}$$

- 3) если $0 < x_0 < 1$, то среди оптимальных стратегий игрока 2 найдется такая, которая является смесью двух существенных стратегий y' и y'' , удовлетворяющих (2.5.23), (2.5.24), с вероятностями β и $1 - \beta$. При этом число $\beta \in [0, 1]$ является решением уравнения

$$\beta H'_x(x_0, y') + (1 - \beta) H'_x(x_0, y'') = 0.$$

§ 2.6. Одновременные игры преследования

В этом параграфе приведено решение некоторых одновременных игр преследования, у которых функция выигрыша или множества стратегий игроков невыпуклые. К таким играм не применимы результаты §5, поэтому решение для обоих игроков находится в классе смешанных стратегий. Существование решения в этом классе гарантируется теоремой п.2.4.4.

2.6.1. Пример 13 (Одновременная игра преследования в кольце). Эта игра является частным случаем примера 1 п. 2.1.2, когда множества $S_1 = S_2 = S$; S представляет собой кольцо. Радиусы внешней и внутренней окружностей кольца S обозначим соответственно R и r , $R > r$.

Покажем, что оптимальными стратегиями игроков 1 и 2 являются выборы точек с равномерным распределением на внутренней (для игрока 2) и внешней (для игрока 1) окружностях кольца S . Обозначим эти стратегии μ^* (для игрока 1) и ν^* (для игрока 2). При указанных стратегиях среднее значение выигрыша (расстояния) равно

$$\begin{aligned} K(\mu^*, \nu^*) &= \frac{1}{4\pi^2} \int_0^{2\pi} \int_0^{2\pi} \sqrt{R^2 + r^2 - 2Rr \cos(\varphi - \psi)} d\varphi d\psi = \\ &= \frac{1}{2\pi} \int_0^{2\pi} \sqrt{R^2 + r^2 - 2Rr \cos \xi} d\xi = \Phi(r, R) \end{aligned} \quad (2.6.1)$$

где ψ и φ — полярные углы чистых стратегий игроков 1 и 2 соответственно. Если игрок 1 выбирает точку x с полярными координатами ρ, ψ , то ожидаемое расстояние (игрок 2 придерживается стратегии ν^*) равно

$$K(x, \nu^*) = \Phi(r, \rho) = \frac{1}{2\pi} \int_0^{2\pi} \sqrt{r^2 + \rho^2 - 2r\rho \cos \xi} d\xi.$$

При $r \leq \rho \leq R$ функция $\varphi(\rho) = \rho^2 + r^2 - 2r\rho \cos \xi$ монотонно возрастает. В частности, $\varphi(\rho) \leq \varphi(R)$ при $r \leq \rho \leq R$. Отсюда имеем $\Phi(r, \rho) \leq \Phi(r, R)$. Поэтому для любой стратегии игрока 1 ожидаемое расстояние не больше $\Phi(r, R)$.

Рассмотрим теперь ситуацию (μ^*, y) , в которой $y \in S$, а ρ и φ — полярные координаты точки y . Имеем

$$K(\mu^*, y) = \Phi(\rho, R) = \frac{1}{2\pi} \int_0^{2\pi} \sqrt{R^2 + \rho^2 - 2R\rho \cos \xi} d\xi(\rho), \quad r \leq \rho \leq R.$$

Зафиксируем R и рассмотрим функцию $\Phi(\rho, R)$ на отрезке $0 \leq \rho \leq R$. Дифференцируя по ρ , можно убедиться, что

$$\frac{\partial \Phi(0, R)}{\partial \rho} = 0, \quad \frac{\partial^2 \Phi(\rho, R)}{\partial \rho^2} > 0, \quad 0 \leq \rho \leq R.$$

Поэтому функция $\Phi(\rho, R)$ монотонно возрастает по ρ , следовательно, $\Phi(r, R) \leq \Phi(\rho, R)$ и

$$K(x, \nu^*) \leq K(\mu^*, \nu^*) \leq K(\mu^*, y)$$

для всех $x, y \in S$. Таким образом, оптимальность стратегий μ^* и ν^* доказана, а значение игры v равно $K(\mu^*, \nu^*)$, где $K(\mu^*, \nu^*)$ определяется (2.6.1). В частности, если S — окружность радиуса R (случай $r = R$), то значение игры равно $4R/\pi$.

2.6.2. Пример 14. Рассмотрим одновременную игру, когда игрок 2 выбирает пару точек $y = \{y_1, y_2\}$, где $y_1 \in S$, $y_2 \in S$, а игрок 1, не зная выбора игрока 2, — точку $x \in S$. Выигрыш игрока 1 полагаем равным $\min_{i=1,2} \rho^2(x, y_i)$. Приведем решение для случая, когда множество R представляет собой круг радиуса R с центром в начале координат (точке O): $S = S(O, R)$.

Рассмотрим функцию $\Phi(r, \rho) = r^2 + \rho^2 - 4r\rho/\pi$, где r и ρ принимают значения из промежутка $r, \rho \in [0, R]$. Установим свойства функции $\Phi(r, \rho)$.

Лемма 1. *Функция $\Phi(r, R)$ (как функция переменной r) является строго выпуклой и достигает абсолютного минимума в единственной точке $r_0 = 2R/\pi$.*

Доказательство. Имеем $\partial^2 \Phi / \partial r^2 = 2 > 0$. Следовательно, функция $\Phi(r, \rho)$, $r \in [0, R]$ строго выпуклая, а производная

$$\frac{\partial \Phi(r, R)}{\partial r} = 2r - \frac{4R}{\pi} \quad (2.6.2)$$

строго монотонна. Очевидно, что функция (2.6.2) в единственной точке $r_0 = 2R/\pi$ обращается в нуль. В силу строгой выпуклости, $\Phi(r, R)$ точка r_0 является единственной точкой абсолютного минимума. Лемма доказана.

Лемма 2. *Функция $\Phi(r_0, \rho)$ строго выпукла по ρ и достигает абсолютного максимума в точке $\rho_0 = R$.*

Доказательство. В силу симметрии функция $\Phi(r, \rho)$ является строго выпуклой по ρ . Поэтому максимум этой функции достигается в одной из точек 0 или R . Тогда имеем

$$\Phi(r_0, R) - \Phi(r_0, 0) = r_0^2 + R^2 - 4r_0R/\pi - r_0^2 = R^2 - \frac{4}{\pi} \left(\frac{2R}{\pi} \right) R = \frac{R^2(\pi^2 - 8)}{\pi^2} > 0.$$

Лемма доказана.

Из лемм 1, 2 вытекает, что пара (r_0, R) является седловой точкой функции Φ :

$$\Phi(r_0, \rho) \leq \Phi(r_0, R) \leq \Phi(r, R).$$

Теорема. *Оптимальными смешанными стратегиями являются: для игрока 2 — выбор точки y_1 с равномерным распределением на окружности $S(O, r_0)$ с центром в точке O и радиусом r_0 ($y_1 = -y_2$), для игрока 1 — выбор точки x с равномерным распределением на окружности $S(O, R)$. Значение игры равно величине $\Phi(r_0, R)$.*

Доказательство. Указанные в теореме стратегии обозначим через μ^* и ν^* для игроков 1 и 2 соответственно. Пусть игрок 1 придерживается стратегии μ^* , а игрок 2 — произвольной чистой стратегии $y = \{y_1, y_2\}$, $y_i = (r_i \cos \varphi_i, r_i \sin \varphi_i)$, $i = 1, 2$. Рассмотрим сначала случай, когда $y_1 = y_2$.

Обозначим через r число $r_1 + r_2$, а через φ — угол $\varphi_1 = \varphi_2$. Выигрыш игрока 1 равен

$$K(\mu^*, y) = \frac{1}{2\pi} \int_0^{2\pi} [R^2 + r^2 - 2Rr \cos(\psi - \varphi)] d\psi = R^2 + r^2 \geq R^2 + r^2 - \frac{4}{\pi} (Rr) = \Phi(r, R). \quad (2.6.3)$$

Тогда, по лемме 1, имеем $K(\mu^*, y) \geq \Phi(r_0, R)$.

В дальнейшем будем предполагать, что $y_1 \neq y_2$. Введем на плоскости полярную систему координат следующим образом. За начало координат возьмем точку O , за полярную ось — луч, выходящий из точки O перпендикулярно хорде AB (множеству равноудаленных от y_1 и y_2 точек круга $S(O, R)$). Для простоты записи предположим,

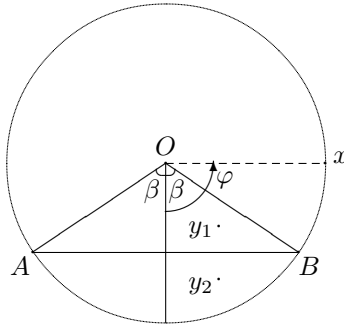


Рис. 2.5. Одновременная игра преследования

что и относительно новой системы координат точка y_i имеет те же координаты $(r_i \cos \varphi_i, r_i \sin \varphi_i)$. Тогда (рис. 2.5) выигрыш первого игрока равен

$$\begin{aligned} K(\mu^*, y) &= \frac{1}{2\pi} \int_0^{2\pi} \min_{i=1,2} [R^2 + r_i^2 - 2Rr_i \cos(\psi - \varphi_i)] d\psi = \\ &= \frac{1}{2\pi} \int_{-\beta}^{\beta} [R^2 + r_2^2 - 2Rr_2 \cos(\psi - \varphi_2)] d\psi + \frac{1}{2\pi} \int_{\beta}^{2\pi-\beta} [R^2 + r_1^2 - 2Rr_1 \cos(\psi - \varphi_1)] d\psi. \end{aligned}$$

Пусть

$$F_1(\varphi) = [(R^2 + r_2^2)\beta - 2Rr_2 \sin \beta \cos \varphi]/\pi, \quad -\beta \leq \varphi \leq \beta,$$

$$F_2(\varphi) = [(R^2 + r_1^2)(\pi - \beta) + 2Rr_1 \sin \beta \cos \varphi]/\pi, \quad \beta \leq \varphi \leq 2\pi - \beta.$$

Стационарными точками функций F_1 и F_2 являются 0 и π соответственно, так как имеем $0 < \beta < \pi/2$ и $F_1'(\varphi) = \frac{2}{\pi} Rr_2 \sin \beta \sin \varphi$, $F_2'(\varphi) = -\frac{2}{\pi} Rr_1 \sin \beta \sin \varphi$, причем 0 и π — точки абсолютного минимума функций F_1 и F_2 , ($F_1'(\varphi) < 0$ при $\varphi \in (-\beta, 0)$, $F_1'(\varphi) > 0$ при $\varphi \in (0, \beta)$; $F_2'(\varphi) < 0$ при $\varphi \in (\beta, \pi)$, $F_2'(\varphi) > 0$ при $\varphi \in (\pi, 2\pi - \beta)$). Следовательно,

$$\begin{aligned} K(\mu^*, y) &= F_1(\varphi_2) + F_2(\varphi_1) \geq F_1(0) + F_2(\pi) = \\ &= \frac{1}{2\pi} \int_{-\beta}^{\beta} (R^2 + r_2^2 - 2Rr_2 \cos \psi) d\psi + \frac{1}{2\pi} \int_{\beta}^{2\pi-\beta} (R^2 + r_1^2 - 2Rr_1 \cos(\psi - \pi)) d\psi, \quad (2.6.4) \end{aligned}$$

т. е. игрок 1 при использовании игроком 2 стратегии $y_1 = \{-r_1, 0\}$, $y_2 = \{r_2, 0\}$ получит меньший выигрыш, чем при использовании стратегии

$$\bar{y}_i = |r_i \cos \varphi_i, r_i \sin \varphi_i|, \quad i = 1, 2.$$

Пусть теперь точки y_1 и y_2 лежат на диаметре круга $S(O, R)$ и расстояние между ними $2\bar{r}$. Обозначим через 2α центральный угол, опирающийся на дугу, стягиваемую хордой AB (рис. 2.6). Предположим, что $y_1 = \{R \cos \alpha - \bar{r}, 0\}$, $y_2 = \{R \cos \alpha + \bar{r}, 0\}$.

Тогда выигрыш первого игрока равен

$$\begin{aligned}
 \psi(\alpha, \bar{r}) &= \frac{1}{2\pi} \int_{-\alpha}^{\alpha} [(R \cos \psi - R \cos \alpha - \bar{r})^2 + R^2 \sin^2 \psi] d\psi + \\
 &+ \frac{1}{2\pi} \int_{\alpha}^{2\pi-\alpha} [(R \cos \psi - R \cos \alpha + \bar{r})^2 + R^2 \sin^2 \psi] d\psi = \\
 &= \frac{1}{2\pi} \int_{-\alpha}^{\alpha} [R^2 - 2R \cos \psi (R \cos \alpha + \bar{r}) + (R \cos \alpha + \bar{r})^2] d\psi + \\
 &+ \frac{1}{2\pi} \int_{\alpha}^{2\pi-\alpha} [R^2 - 2R \cos \psi (R \cos \alpha - \bar{r}) + (R \cos \alpha - \bar{r})^2] d\psi = \\
 &= \frac{1}{\pi} \{ [R^2 + (R \cos \alpha + \bar{r})^2] \alpha - 2R \sin \alpha (R \cos \alpha + \bar{r}) + \\
 &+ [R^2 + (R \cos \alpha - \bar{r})^2] (\pi - \alpha) + 2R \sin \alpha (R \cos \alpha - \bar{r}) \}.
 \end{aligned}$$

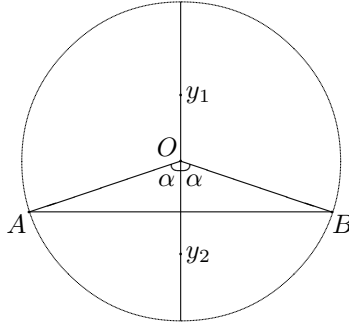


Рис. 2.6. Точки y_1 и y_2 лежат на диаметре круга $S(O, R)$

Покажем, что функция \bar{r} при фиксированном $\psi(\alpha, \bar{r})$ достигает минимума по α при $\alpha = \pi/2$. В результате элементарных вычислений получим $\partial\psi/\partial\alpha = \{2R \sin \alpha[(\pi - 2\alpha)\bar{r} - \pi R \cos \alpha]\}/\pi$, поэтому для достаточно малых значений α имеем $\partial\psi(\alpha, \bar{r})/\partial\alpha < 0$, т. к. $\sin \alpha > 0$, $\bar{r}(\pi - 2\alpha) - \pi R \cos \alpha < 0$ (в предельном случае $\bar{r}\pi - \pi R < 0$). Вместе с тем $\partial\psi(\pi/2, r)/\partial\alpha = 0$.

При каждом фиксированном r функция $\partial\psi(\alpha, r)/\partial\alpha$ не имеет нулей по α , кроме $\alpha = \pi/2$. Предположим противное. Пусть α_1 — нуль этой функции в интервале $(0, \pi/2)$. Тогда функция $G(\alpha) = (\pi - 2\alpha)r - \pi R \cos \alpha$ обратится в нуль при $\alpha = \alpha_1$. Таким образом, $G(\alpha_1) = G(\pi/2) = 0$.

Очевидно, что $G(\alpha) > 0$ при всех $\alpha \in (\alpha_1, \pi/2)$. Это противоречит выпуклости функции $G(\alpha)$ ($G''(\alpha) = \pi R \cos \alpha > 0$). Следовательно, $\partial\psi(\alpha, r)/\partial\alpha < 0$ при $\alpha \in (0, \pi/2)$ и $\partial\psi(\pi/2, r)/\partial\alpha = 0$. Тогда функция $\psi(\alpha, r)$ достигает абсолютного минимума по α при $\alpha = \pi/2$: $\psi(\alpha, r) \geq \psi(\pi/2, r)$. Значит, и в этом случае имеем

$$K(\mu^*, y) = \psi(\alpha, r) \geq \psi(\pi/2, r) = \Phi(r, R) \geq \Phi(r_0, R). \quad (2.6.5)$$

Из (2.6.3)–(2.6.5) вытекает, что для любой чистой стратегии $y = \{y_1, y_2\}$ справедливо неравенство

$$K(\mu^*, y) \geq \Phi(r_0, R). \quad (2.6.6)$$

Пусть игрок 2 применяет стратегию ν^* , а игрок 1 — произвольную чистую стратегию $x = \{\rho \cos \psi, \rho \sin \psi\}$. Тогда игрок 1 получает выигрыш

$$\begin{aligned} K(x, \nu^*) &= \frac{1}{2\pi} \int_0^{2\pi} \min[\rho^2 + r_0^2 - 2\rho r_0 \cos(\psi - \varphi), \rho^2 + r_0^2 + 2\rho r_0 \cos(\psi - \varphi)] d\varphi = \\ &= \frac{1}{2\pi} \int_0^{2\pi} \min(\rho^2 + r_0^2 - 2\rho r_0 \cos \xi, \rho^2 + r_0^2 + 2\rho r_0 \cos \xi) d\xi = \Phi(r_0, \rho) \end{aligned}$$

и по лемме 2 имеем

$$K(x, \nu^*) = \Phi(r_0, \rho) \leq \Phi(r_0, R). \quad (2.6.7)$$

Из неравенств (2.6.6) и (2.6.7) получаем, что μ^* и ν^* — являются оптимальными стратегиями игроков, а $\Phi(r_0, R)$ — значение игры. Теорема доказана.

2.6.3. Пример 15. Пусть игрок 2 выбирает набор из m точек $y = \{y_1, \dots, y_m\}$, где $y_i \in S, i = 1, \dots, m$, а игрок 1 одновременно с ним — точку $x \in S$. Выигрыш игрока 1 полагаем равным $\min_{i=1, \dots, m} \rho(x, y_i)$. Решим игру в случае, когда множество S совпадает с отрезком $[-1, 1]$.

Теорема. *Оптимальная смешанная стратегия ν^* игрока 2 заключается в равновероятном выборе двух наборов из m точек*

$$\left\{-1 + \frac{4i}{2m-1}, i = 0, 1, \dots, m-1\right\},$$

$$\left\{1 - \frac{4i}{2m-1}, i = 0, 1, \dots, m-1\right\}.$$

Оптимальная стратегия μ^ игрока 1 состоит в выборе точек*

$$\left\{\frac{2m-2i-1}{2m-1}, i = 0, 1, \dots, 2m-1\right\}$$

с вероятностями $1/(2m)$. Значение игры равно $1/(2m-1)$.

Доказательство. Пусть μ^* и ν^* — смешанные стратегии игроков 1 и 2 соответственно, оптимальность которых нужно доказать. Введем следующие обозначения:

$$l_i = \left[\frac{2m-2i-1}{2m-1}, \frac{2m-2i+1}{2m-1} \right], \quad i = 1, 2, \dots, 2m-1.$$

Покажем сначала, что $K(x, \nu^*) \leq \frac{1}{(2m-1)}$ при всех $x \in [-1, 1]$. Действительно, при $x \in l_i$ имеем

$$\begin{aligned} K(x, \nu^*) &= \frac{1}{2} \min_i \left| \frac{2m-4i-1}{2m-1} - x \right| + \\ &\quad + \frac{1}{2} \min_i \left| \frac{-2m+4i+1}{2m-1} - x \right| = \\ &= \frac{1}{2} \left(x - \frac{2m-2j-1}{2m-1} \right) + \frac{1}{2} \left(\frac{2m-2j+1}{2m-1} - x \right) = \frac{1}{2m-1}. \end{aligned} \quad (2.6.8)$$

Пусть теперь игрок 1 выбирает смешанную стратегию μ^* , а игрок 2 — произвольную чистую стратегию $y = \{y_1, \dots, y_m\}$.

Обозначим

$$x_j = \frac{2m - 2j - 1}{2m - 1}, \quad j = 0, 1, \dots, 2m - 1.$$

Тогда

$$\begin{aligned} K(\mu^*, y) &= \sum_{j=0}^{2m-1} \min_{1 \leq i \leq m} \rho(x_j, y_i) \frac{1}{2m} = \\ &= \frac{1}{2m} \sum_{j=1}^m \left[\min_{1 \leq i \leq m} \rho(x_{2j-1}, y_i) + \min_{1 \leq i \leq m} \rho(x_{2j-2}, y_i) \right] \geq \frac{1}{2m} m \frac{2}{2m-1} = \frac{1}{2m-1}. \end{aligned}$$

Из неравенств (2.6.8), (2.6.9) вытекает утверждение теоремы.

§ 2.7. Один класс игр с разрывной функцией выигрыша

Для игр, у которых функции выигрыша разрывны, нельзя гарантировать существование значения игры в смешанных стратегиях (см. пример п. 2.4.12). Однако часто именно разрывность функции выигрыша позволяет найти оптимальные стратегии и значение игры. Нахождению решения помогают также эмпирические предположения о виде оптимальных стратегий игроков.

2.7.1. В данном параграфе будут исследованы игры с выбором момента времени или игры типа дуэли (см. примеры 4,5 п. 2.1.2). Основной особенностью этого класса игр на квадрате является разрывность функции выигрыша $H(x, y)$ вдоль диагонали $x = y$.

Рассмотрим игру на единичном квадрате с функцией выигрыша [Карлин, 1964]

$$H(x, y) = \begin{cases} \psi(x, y), & \text{если } x < y, \\ \varphi(x), & \text{если } x = y, \\ \theta(x, y), & \text{если } x > y, \end{cases} \quad (2.7.1)$$

где $\psi(x, y)$ определена и непрерывна на множестве $0 \leq x \leq y \leq 1$, функция φ непрерывна на $[0, 1]$, а $\theta(x, y)$ определена и непрерывна на множестве $0 \leq y \leq x \leq 1$. Предположим, что игра $\Gamma = (X, Y, H)$, где $X = Y = [0, 1]$, H задана (2.7.1), имеет оптимальные смешанные стратегии μ^* и ν^* игроков 1 и 2 соответственно. Более того, предположим, что оптимальные смешанные стратегии μ^*, ν^* являются распределениями вероятностей, которые имеют непрерывные плотности $f^*(x)$ и $g^*(x)$ соответственно.

Далее в этом параграфе будем обозначать искомую стратегию f (соответственно g), понимая под этим плотность распределения. Выясним свойства оптимальных стратегий.

Пусть f — стратегия игрока 1. Для $y \in [0, 1]$ имеем

$$K(f, y) = \int_0^y \psi(x, y) f(x) dx + \int_y^1 \theta(x, y) f(x) dx. \quad (2.7.2)$$

Предположим, что f и g — оптимальные стратегии игроков 1 и 2 соответственно. Тогда для любой точки y_0 , в которой

$$g(y_0) > 0 \quad (2.7.3)$$

(точки спектра стратегии g), выполняется

$$K(f, y_0) = v, \quad (2.7.4)$$

где v — значение игры. Но неравенство (2.7.3) строгое, поэтому существует $\delta > 0$ такое, что для всех $y : |y - y_0| < \delta$, неравенство (2.7.3) сохраняется. Таким образом, для этих y сохраняется и равенство (2.7.4), т. е. выполняется равенство $K(f, y) = v$. Это означает, что

$$\frac{\partial K(f, y)}{\partial y} = 0. \quad (2.7.5)$$

Уравнение (2.7.5) перепишем в виде

$$[\theta(y, y) - \psi(y, y)]f(y) = \int_0^y \psi_y(x, y)f(x)dx + \int_y^1 \theta_y(x, y)f(x)dx, \quad y \in S(y_0, \delta). \quad (2.7.6)$$

Таким образом, мы получили интегральное уравнение (2.7.6) относительно искомой стратегии f .

2.7.2. Пример 16. Рассмотрим беспумную дуэль, сформулированную в примере 5 п. 2.1.2. Функция выигрыша в игре имеет вид (2.7.1), где

$$\psi(x, y) = x - y + xy, \quad (2.7.7)$$

$$\theta(x, y) = x - y - xy, \quad (2.7.8)$$

$$\varphi(x) = 0. \quad (2.7.9)$$

Заметим, что данная игра является симметричной, поскольку $H(x, y) = -H(y, x)$ (кососимметричная функция выигрыша). Поэтому анализ, аналогичный проведенному в п. 1.9.2 гл. 1, показывает, что значение v игры, если оно существует, равно нулю, а оптимальные стратегии игроков (если они также существуют) должны быть одинаковыми.

Имеем: $\psi_y(x, y) = -1 + x$, $\theta_y(x, y) = -1 - x$, $\theta(y, y) - \psi(y, y) = -2y^2$ и интегральное уравнение (2.7.6) принимает вид

$$-2y^2 f(y) = \int_0^y (x - 1)f(x)dx - \int_y^1 (x + 1)f(x)dx. \quad (2.7.10)$$

Будем искать стратегию f в классе дифференцируемых плотностей распределения, принимающих положительные значения в интервале $(\alpha, \beta) \subset [0, 1]$ (интервал (α, β) — спектр стратегии f). Тогда (2.7.10) можно записать следующим образом:

$$-2y^2 f(y) = \int_\alpha^y (x - 1)f(x)dx - \int_y^\beta (x + 1)f(x)dx. \quad (2.7.11)$$

Дифференцируя обе части (2.7.11) по y , получим дифференциальное уравнение вида

$$-4yf - 2y^2 f' = (y - 1)f + (y + 1)f$$

или

$$yf' = -3f \quad (y \neq 0). \quad (2.7.12)$$

Интегрируя уравнение (2.7.12), имеем

$$f(y) = \gamma y^{-3}, \quad (2.7.13)$$

где γ — некоторая константа.

Теперь осталось найти α, β и γ . Напомним, что оптимальные стратегии игроков в рассматриваемой игре одинаковы. Из нашего предположения о спектре стратегии f следует, что

$$K(f, y) = 0 \quad (2.7.14)$$

для всех $y \in (\alpha, \beta)$.

Пусть $\beta < 1$. Поскольку функция $K(f, y)$ непрерывна по y , из (2.7.14) имеем $K(f, \beta) = 0$. Следовательно,

$$\int_{\alpha}^{\beta} (x - \beta + \beta x) f(x) dx = 0. \quad (2.7.15)$$

Однако в случае $\beta < 1$ из (2.7.15) следует

$$K(f, 1) = \int_{\alpha}^{\beta} (x - 1 + x) f(x) dx < 0,$$

что противоречит оптимальности стратегии f . Таким образом, $\beta = 1$ и $K(f, 1) = 0$. Тогда, подставляя (2.7.13) в (2.7.15) при $\beta = 1$, получаем

$$\gamma \int_{\alpha}^1 \frac{2x - 1}{x^3} dx = 0, \quad \gamma \neq 0.$$

Откуда вытекает

$$3\alpha^2 - 4\alpha + 1 = 0. \quad (2.7.16)$$

Решая уравнение (2.7.16), найдем два корня $\alpha = 1$ и $\alpha = 1/3$, первый из которых посторонний. Следовательно, $\alpha = 1/3$. Коэффициент γ находится из условия нормировки $f(y)$

$$\int_{1/3}^1 f(y) dy = \gamma \int_{1/3}^1 y^{-3} dy = 1,$$

откуда $\gamma = 1/4$.

Таким образом, получено решение игры примера 5 п. 2.1.2: значение игры равно $v = 0$, оптимальные стратегии f и g обоих игроков (как плотности распределения) равны между собой и имеют вид

$$f(x) = \begin{cases} 0, & x < 1/3, \\ 1/(4x^3), & x > 1/3. \end{cases}$$

2.7.3. Пример 17. Найдем решение игры «шумная дуэль» (см. пример 4 п. 2.1.2) для функций меткости $p_1(x) = x$ и $p_2(y) = y$. Функция выигрыша $H(x, y)$ в игре имеет вид (2.7.1), где

$$\psi(x, y) = 2x - 1, \quad (2.7.17)$$

$$\theta(x, y) = 1 - 2y, \quad (2.7.18)$$

$$\varphi(x) = 0. \quad (2.7.19)$$

Игра является симметричной, поэтому $v = 0$, а оптимальные стратегии игроков совпадают. Здесь оба игрока имеют чистую оптимальную стратегию $x^* = y^* = 1/2$. Действительно, $H(1/2, y) = \theta(1/2, y) = 1 - 2y > 0$ при $y < 1/2$; $H(1/2, y) = \varphi(1/2) = 0$ при $y = 1/2$; $H(1/2, y) = \psi(1/2, y) = 0$ при $y > 1/2$.

С точки зрения интерпретации игры решение предписывает дуэлянтам стрелять одновременно, когда каждый пройдет половину дистанции до барьера.

В заключение следует отметить, что класс игр с выбором момента времени хорошо изучен (см. [Давыдов, 1978, Karlin, 1959, Воробьев, 1984]).

§ 2.8. Решение бесконечных одновременных игр поиска

В этом параграфе будет приведено решение игр поиска с бесконечным числом стратегий, сформулированных в п. 2.1.2. Первая из рассматриваемых игр интересна тем, что в ней оба игрока имеют оптимальные смешанные стратегии с конечным спектром.

2.8.1. Пример 18 (Поиск на отрезке). [Дюбин, Суздаль, 1981]. Рассмотрим задачу поиска на отрезке (см. пример 2 п. 2.1.1), которая моделируется игрой на единичном квадрате с функцией выигрыша $H(x, y)$ вида

$$H(x, y) = \begin{cases} 1, & \text{если } |x - y| \leq l, \quad l \in (0, 1), \\ 0, & \text{в противном случае.} \end{cases} \quad (2.8.1)$$

Заметим, что при $l \geq 1/2$ у игрока 1 имеется чистая оптимальная стратегия $x^* = 1/2$ и значение игры равно единице, поскольку в этом случае $H(x^*, y) = H(1/2, y) = 1$, так как $|1/2 - y| \leq 1/2 \leq l$ для всех $y \in [0, 1]$. Предположим, что $l < 1/2$. Заметим, что стратегия $\bar{x} = l$ доминирует все чистые стратегии $x < l$, а стратегия $\bar{x} = 1 - l$ — все стратегии $x > 1 - l$. Действительно,

$$H(\bar{x}, y) = H(l, y) = \begin{cases} 1, & y \in [0, 2l], \\ 0, & \text{в противном случае,} \end{cases}$$

и если $x < l$, то

$$H(x, y) = \begin{cases} 1, & y \in [0, l + x], \\ 0, & \text{в противном случае.} \end{cases}$$

Таким образом, при $x < l$: $H(x, y) \leq H(l, y)$ для всех $y \in [0, 1]$.

Аналогично имеем

$$H(x, y) = H(1 - l, y) = \begin{cases} 1, & y \in [1 - 2l, 1], \\ 0, & \text{в противном случае,} \end{cases}$$

а если $x \in [1 - l, 1]$, то

$$H(x, y) = \begin{cases} 1, & y \in [x - l, 1], \\ 0, & \text{в противном случае.} \end{cases}$$

Таким образом, при $x \in [1 - l, 1]$, $H(x, y) \leq H(1 - l, y)$ для всех $y \in [0, 1]$.

Рассмотрим следующую смешанную стратегию μ^* игрока 1. Пусть $l = x_1 < x_2 < \dots < x_m = 1 - l$ — точки, для которых расстояние между любой парой соседних точек не превосходит $2l$. Стратегия μ^* выбирает каждую из этих точек с равными вероятностями $1/m$. Очевидно, что при этом любая точка $y \in [0, 1]$ попадает в l -окрестность хотя бы одной точки x_k . Следовательно,

$$K(\mu^*, y) \geq 1/m. \quad (2.8.2)$$

Пусть теперь ν^* — стратегия игрока 2, которая состоит в равновероятном выборе точек $0 = y_1 < y_2 < \dots < y_n = 1$, причем расстояние между парой соседних точек больше $2l$. Тогда, очевидно, существует не более одной точки y_k , в l -окрестности которой содержится точка x . Следовательно,

$$K(x, \nu^*) \leq 1/n. \quad (2.8.3)$$

Если бы удалось построить стратегии μ^*, ν^* так, чтобы $m = n$, то величина $1/n$ была бы значением игры, а стратегии μ^*, ν^* — оптимальными стратегиями игроков. Оказывается, такие стратегии действительно можно построить.

Для этого достаточно взять

$$m = n = \begin{cases} 1/(2l), & \text{если } 1/(2l) \text{ — целое,} \\ [1/(2l)] + 1, & \text{в противном случае.} \end{cases} \quad (2.8.4)$$

Здесь $[a]$ — целая часть числа a . Точки

$$x_i = l + \frac{1 - 2l}{n - 1}(i - 1), \quad i = 1, 2, \dots, n, \quad (2.8.5)$$

отстоят друг от друга не более чем на $2l$, а расстояние между соседними точками

$$y_j = \frac{j - 1}{n - 1}, \quad j = 1, 2, \dots, n, \quad (2.8.6)$$

строго больше $2l$. Таким образом, $1/n$ — значение игры, а оптимальные стратегии μ^*, ν^* являются равновероятными смесями чистых стратегий, определяемых формулами (2.8.5), (2.8.6).

2.8.2. Пример 19. Рассмотрим обобщение предыдущей задачи в том случае, когда игрок 1 (ищущий) выбирает систему из s точек x_1, \dots, x_s , $x_i \in [0, 1]$, $i = 1, \dots, s$, а игрок 2 (прячущийся) выбирает независимо и одновременно с игроком 1 точку $y \in [0, 1]$. Игрок 2 считается обнаруженным, если находится такое $j \in \{1, \dots, s\}$, что $|y - x_j| \leq l$, $l > 0$.

В соответствии с этим функция выигрыша (выигрыш игрока 1) определяется следующим образом:

$$H(x_1, \dots, x_s, y) = \begin{cases} 1, & \min_j |y - x_j| \leq l, \\ 0, & \text{в противном случае.} \end{cases} \quad (2.8.7)$$

Предположим, что игрок 1 располагает точки x_1, \dots, x_s в точках $\bar{x}_i = l + (1 - 2l)(i - 1)/(n - 1)$, $1 \leq i \leq n$, являющихся точками спектра стратегии μ^* из предыдущего примера. Очевидно, что располагать две точки x_{j_1}, x_{j_2} в одной точке отрезка $[0, 1]$ (т. е. выбирать совпадающие точки) невыгодно. Пусть μ_s^* — стратегия игрока 1, выбирающая равновероятно любые s -наборы не равных друг другу точек $\{\bar{x}_i\}$. Если $s \geq n$, то, расположив в каждой из точек x_j по точке \bar{x}_i , игрок 1 полностью покрывает отрезок $[0, 1]$ интервалами длины $2l$ с центрами в точках \bar{x}_i , и тем самым обеспечит, что для любой точки $y \in [0, 1]$ будет иметь место $\min_j |x_j - y| \leq l$, т. е. в этом случае значение игры равно единице. Поэтому будем считать, что $s < n$. Число всевозможных различных выборов s -наборов точек из множества $\{\bar{x}_i\}$ равно C_n^s . Имеем

$$K(\mu_s^*, y) = \sum H(\bar{x}_{i_1}, \dots, \bar{x}_{i_s}, y) \left(\frac{1}{C_n^s} \right) \geq \frac{C_{n-1}^{s-1}}{C_n^s} = \frac{s}{n}.$$

Действительно, точка y обнаруживается, если она попадает в l -окрестность хотя бы одной из выбранных стратегией μ_s^* точек $\{\bar{x}_i\}$. Для того, чтобы это произошло, игроку 1 необходимо выбрать точку \bar{x}_i из l -окрестности точки y . Число наборов, удовлетворяющих этому требованию, не менее C_{n-1}^{s-1} .

Предположим теперь, что игрок 2 использует стратегию ν^* из предыдущего примера, а игрок 1 — произвольную чистую стратегию $x = (x_1, \dots, x_s)$. Тогда

$$K(x_1, \dots, x_s, \nu^*) = \sum_{j=1}^n H(x_1, \dots, x_s, y_j) \frac{1}{n} \leq \frac{s}{n}.$$

Таким образом, значение игры равно s/n , а μ_s^*, ν^* — оптимальные стратегии игроков. Значение игры линейно зависит от количества выбираемых ищущим игроком точек.

2.8.3. Пример 20 (Поиск на сфере). Рассмотрим игру поиска на сфере (см. пример 3 п. 2.1.2). Функция выигрыша $H(x, y)$ имеет вид

$$H(x, y) = \begin{cases} 1, & y \in M_x, \\ 0, & \text{в противном случае,} \end{cases} \quad (2.8.8)$$

где $x = (x_1, \dots, x_s)$ — набор s точек на сфере C и $M_x = \cup_{j=1}^s S(x_j, r)$; $S(x_j, r)$ — r -сферическая окрестность точки x_j . Множество смешанных стратегий игрока 1 представляет собой семейство вероятностных мер $\{M\}$, определенных на декартовом произведении s сфер $C \times C \times \dots \times C = \Omega$, т. е. на $\Omega = C^s$.

Множество смешанных стратегий игрока 2 определим как семейство вероятностных мер $\{\nu\}$, определенных на сфере C .

Рассмотрим конкретную пару стратегий (μ^*, ν^*) . В качестве стратегии ν^* выберем равномерную меру на сфере C , т. е. потребуем, чтобы

$$\int_A d\nu^* = \frac{L(A)}{4\pi R^2}, \quad (2.8.9)$$

где $L(A)$ — лебегова мера (площадь) множества A .

Будем предполагать, что параметры игры s , r и R таковы, что можно выбрать систему точек $x = (x_1, x_2, \dots, x_s)$, удовлетворяющих условию

$$L(M_x) = \sum_{j=1}^s L(S(x_j, r)) \quad (2.8.10)$$

(сферические сегменты $S(x_j, r)$ не пересекаются).

Зафиксируем фигуру M_x на некоторой сфере C' . Тогда смешанная стратегия μ^* порождается случайным бросанием этой фигуры M_x на сферу C . Для этого в фигуре M_x фиксируется некоторая внутренняя точка z , с которой жестко связываются два неколлинеарных вектора a, b (с углом $\varphi > 0$ между ними), расположенных в касательной плоскости к M_x в точке z .

Точка z «бросается» на сферу C в соответствии с равномерным распределением, т. е. плотностью $1/(4\pi R^2)$. Пусть в результате реализуется точка $z' \in C$. Фигура M_x с фиксированными на ней векторами параллельно переносится на сферу C так, чтобы точки z и z' совпали. Таким образом, векторы a, b будут лежать в касательной плоскости к сфере C в точке z' .

Затем на промежутке $[0, 2\pi]$ выбирают в соответствии с равномерным распределением угол φ' , и вектор b в касательной плоскости поворачивают вместе со связанной с ним фигурой M_x на угол φ' по часовой стрелке. В результате фигура M_x и вектор b переходят в новое положение на сфере C . Случайное размещение множества M_x на сфере в соответствии с описанной двухэтапной процедурой и порождает случайный выбор точек $x'_1, x'_2, \dots, x'_s \in C$, соответствующих смешанной стратегии μ^* , а именно: игрок 1 выбирает точки $x'_1, x'_2, \dots, x'_s \in C$, в которых оказались центры x_1, \dots, x_s сферических окрестностей $S(x_j, r)$, составляющих множество M_s .

Мера μ^* , построенная таким образом, оказывается инвариантной, т. е. вероятность покрытия множеством M_x любой точки $y \in C$ не зависит от y . Действительно, найдем вероятность этого события. Пусть $\bar{\Omega} = \{\omega\}$ — пространство всевозможных размещений M_x на сфере C . Тогда средняя площадь, покрываемая на сфере C при бросании на нее множества M_x (математическое ожидание площади), равна $L(M_x)$. В то же время

$$L(M_x) = \int_{\bar{\Omega}} \int_C J(y, \omega) dy d\mu^*, \quad (2.8.11)$$

где $J(y, \omega)$ — характеристическая функция множества на сфере C , покрываемого областью M_x . По теореме Фубини имеем

$$\int_{\bar{\Omega}} \int_C J(y, \omega) dy d\mu^* = \int_C \int_{\bar{\Omega}} J(y, \omega) d\mu^* dy. \quad (2.8.12)$$

Однако в силу инвариантности меры μ^* интеграл $\int_{\bar{\Omega}} J(y, \omega) d\mu^*$, совпадающий с вероятностью покрытия точки y множеством M_x , не зависит от y и равен p . Тогда из (2.8.11), (2.8.12) получаем

$$p = \frac{L(M_x)}{4\pi R^2} = \frac{\sum_{j=1}^s L(S(x_j, r))}{4\pi R^2}. \quad (2.8.13)$$

Обозначим через $K(\mu, \nu)$ математическое ожидание выигрыша при использовании игроками смешанных стратегий $\mu \in \{\mu\}$ и $\nu \in \{\nu\}$. Если один из игроков использует чистую стратегию, то

$$K(x, \nu) = \int_C H(x, y) d\nu = \int_{M_x} d\nu = \Pr(y \in M_x),$$

$$K(\mu, y) = \int_{\Omega} H(x, y) d\mu = \int_{\Omega} J(x, y) d\mu = \Pr(y \in M_x),$$

и в этом случае математические ожидания соответственно имеют смысл вероятностей попадания случайной точки в фиксированную область и накрытия случайной областью фиксированной точки. Для всех y и $x = (x_1, \dots, x_s)$ в силу условий (2.8.9) и (2.8.13) имеем

$$K(x, \nu^*) = \frac{L(M_x)}{4\pi R^2} \leq \frac{\sum_{j=1}^s L(S(x_j, r))}{4\pi R^2} = \frac{s}{2} \left(1 - \sqrt{1 - \left(\frac{r}{R}\right)^2} \right),$$

$$K(\mu^*, y) = \frac{\sum_{j=1}^s L(S(x_j, r))}{4\pi R^2} = \frac{s}{2} \left(1 - \sqrt{1 - \left(\frac{r}{R}\right)^2} \right),$$

так как $L(S(x_j, r)) = 2\pi R(R - \sqrt{R^2 - r^2})$.

Из определения ситуации равновесия и полученного неравенства $K(\mu^*, y) \geq K(x, \nu^*)$ следует, что смешанные стратегии μ^* и ν^* являются оптимальными и

$$K(\mu^*, \nu^*) = \frac{s}{2} \left(1 - \sqrt{1 - \left(\frac{r}{R}\right)^2} \right)$$

— значение рассмотренной игры поиска.

2.8.4. Рассмотрим вариант предыдущей игры, полагая, что игрок 2 выбирает некоторое односвязное множество $Y \subset C$ и целью игрока 1 является максимизация площади пересечения

$$L(Y \cap M_x) = L(Y \cap \cup_{j=1}^s S(x_j, r)).$$

Цель игрока 2 противоположна. В остальном игра совпадает с игрой, рассмотренной в начале параграфа. Стратегия μ^* игрока 1 совпадает с таковой в предыдущей игре. Смешанная стратегия ν^* игрока 2 строится аналогично стратегии μ^* и заключается в случайном бросании множества Y на сферу (в предыдущем случае игрок 2 случайно выбирал точки $y \in C$). Таким образом, ν^* строится как инвариантная мера, которая состоит из случайного (в соответствии с равномерным распределением на C) выбора одной из фиксированных точек множества Y на C и далее поворота Y вокруг этой точки на случайный угол (в соответствии с равномерным распределением на $[0, 2\pi]$). Пусть $K(x, \nu)$, $K(\mu, y)$ соответствуют математическим ожиданиям площади пересечения $L(Y \cap M_x)$. Тогда

$$K(\mu^*, y) = K(x, \nu^*) = K(\mu^*, \nu^*) = \frac{L(Y)L(M_x)}{4\pi R^2}.$$

Если Y — r -окрестность точки y , то значение игры равно

$$K(\mu^*, \nu^*) = \pi s (R - \sqrt{R^2 - r^2})^2.$$

§ 2.9. Покер

2.9.1. *Модель покера с одним кругом ставок и одним размером ставки* (см. [Беллман, 1960; Karlin, 1959]).

Модель [Karlin, 1959], рассмотренная в данном разделе, является частным случаем модели п. 2.9.2, допускающей n возможных размеров ставки.

Модель. В начале партии каждый из двух игроков A и B ставит по единице. После того, как каждый из игроков получит карту, ходит игрок A : он может или поставить еще a единиц или спасовать и потерять свою начальную ставку. Если A ставит, то у B две альтернативы: он может или спасовать (теряя при этом свою начальную ставку), или уравнивать, поставив a единиц. Если B уравнивает, то игроки открывают свои карты и игрок с лучшей картой выигрывает единицу (банк).

Будем обозначать карту игрока A через ξ , а карту игрока B через η , при этом предполагаем, что случайные величины ξ и η имеют равномерное распределение на единичном интервале. Положим $L(\xi, \eta) = \text{sign}(\xi - \eta)$.

Стратегии и выигрыши. Стратегии строятся так. Пусть

$\phi(\xi)$ — вероятность того, что если A получит ξ , то он поставит a ,

$1 - \phi(\xi)$ — вероятность того, что если A получит ξ , то он спасует,

$\psi(\eta)$ — вероятность того, что если B получит η , то он уравнивает ставку a ,

$1 - \psi(\eta)$ — вероятность того, что если B получит η , то он спасует.

Если игроки применяют эти стратегии, то ожидаемый чистый выигрыш $K(\phi, \psi)$ представляет собой сумму выигрышей, соответствующих трем взаимно исключающим возможностям: A пасует, A ставит a единиц и B уравнивает; A ставит и B пасует. Таким образом,

$$\begin{aligned} K(\phi, \psi) = & (-1) \int_0^1 [1 - \phi(\xi)] d\xi + (a + 1) \int_0^1 \int_0^1 \phi(\xi) \psi(\eta) L(\xi, \eta) d\xi d\eta + \\ & + \int_0^1 \int_0^1 \phi(\xi) [1 - \psi(\eta)] d\xi d\eta. \end{aligned}$$

Выигрыш игрока A можно переписать следующим образом:

$$K(\phi, \psi) = -1 + \int_0^1 \phi(\xi) \left(2 + a \int_0^\xi \psi(\eta) d\eta - (a + 2) \int_\xi^1 \psi(\eta) d\eta \right) d\xi \quad (2.9.1)$$

или

$$K(\phi, \psi) = -1 + 2 \int_0^1 \phi(\xi) d\xi + \int_0^1 \psi(\eta) \left(-(a + 2) \int_0^\eta \phi(\xi) d\xi + a \int_\eta^1 \phi(\xi) d\xi \right) d\eta. \quad (2.9.2)$$

Для решения игры нам надо найти такую пару стратегий (ϕ^*, ψ^*) , что

$$K(\phi, \psi^*) \leq K(\phi^*, \psi^*) \leq K(\phi^*, \psi) \quad (2.9.3)$$

для всех стратегий ϕ и ψ соответственно. Таким образом, ϕ^* максимизирует $K(\phi, \psi^*)$, в то время как ψ^* минимизирует $K(\phi^*, \psi)$. Поэтому мы будем искать стратегию ϕ^* , которая максимизирует выигрыш (2.9.1), где ψ заменено на ψ^* ; точно так же мы будем

искать стратегию ψ^* , которая минимизирует (2.9.2), где ϕ заменено на ϕ^* . Поскольку постоянные слагаемые не играют роли, задача состоит здесь в том, чтобы найти экстремумы:

$$\max_{\phi} \int_0^1 \phi(\xi) \left(2 + a \int_0^{\xi} \psi^*(\eta) d\eta - (a+2) \int_{\xi}^1 \psi^*(\eta) d\eta \right) d\xi \quad (2.9.4)$$

и

$$\min_{\psi} \int_0^1 \psi(\eta) \left(-(a+2) \int_0^{\eta} \phi^*(\xi) d\xi + a \int_{\eta}^1 \phi^*(\xi) d\xi \right) d\eta. \quad (2.9.5)$$

Нетривиальность задачи состоит в том, чтобы проверить совместность наших результатов, т. е. что функция ϕ^* , которая максимизирует величину (2.9.4), есть та же самая функция ϕ^* , которая входит в выражение (2.9.5), и аналогично для ψ^* ; если эти утверждения верны, то выполнено соотношение (2.9.3), и мы нашли решение игры.

Интуитивные соображения подсказывают вид решения, которое мы ищем. Так как игрок B не имеет возможности блефовать, то $\psi^*(\eta) = 1$ для значений η больших, чем некоторое критическое число c , и $\psi^*(\eta) = 0$ в противном случае; кроме того, поскольку B — минимизирующий игрок, функция $\psi^*(\eta)$ должна быть равна 1, если коэффициент при $\psi(\eta)$ в выражении (2.9.5) отрицателен. Но этот коэффициент представляет собой убывающую функцию η , и пусть c есть то значение, при котором она впервые обращается в нуль. Следовательно,

$$-(a+2) \int_0^c \phi^*(\xi) d\xi + a \int_c^1 \phi^*(\xi) d\xi = 0. \quad (2.9.6)$$

При таком выборе $\psi^*(\eta)$ мы видим, что коэффициент при $\phi(\xi)$ в выражении (2.9.4) является константой при $\xi \leq c$. Если мы предположим, что эта константа равна нулю, то получим при $\xi = c$

$$2 + 0 - (a+2)(1-c) = 0,$$

или

$$c = \frac{a}{a+2}. \quad (2.9.7)$$

Причина, по которой мы определили постоянную c так, чтобы обратить в нуль коэффициент на интервале $[0, c]$, состоит в следующем. При нахождении максимума в (2.9.4) мы, очевидно, вынуждены положить $\phi^*(\xi) = 1$, если коэффициент при этой функции положителен, и $\phi^*(\xi) = 0$, если этот коэффициент отрицателен. Значение $\phi^*(\xi)$ может быть произвольным только в тех случаях, когда соответствующий коэффициент равен нулю. Но можно ожидать, что A будет пытаться частично блефовать, имея на руках плохую карту, откуда следует, что для этих раскладов $\phi(\xi)$, вероятно, удовлетворяет условию $0 < \phi^*(\xi) < 1$. Как указывалось, это может быть, если коэффициент при $\phi(\xi)$ равен нулю. При определении c , согласно соотношению (2.9.7), коэффициент при $\phi(\xi)$ в выражении (2.9.4) равен нулю при $\xi \leq c$ и положителен при $\xi > c$. В этих условиях из (2.9.4) следует, что максимизирующий игрок вынужден положить $\phi^*(\xi) = 1$ при $\xi > c$, в то время как значения $\phi^*(\xi)$ при $\xi \leq c$ могут быть произвольны, ибо они не влияют на выигрыш.

Однако для того, чтобы равенство (2.9.6) при таком выборе ϕ^* удовлетворялось, мы должны иметь

$$-(a+2) \int_0^c \phi^*(\xi) d\xi + a(1-c) = 0,$$

или

$$\int_0^c \phi^*(\xi) d\xi = \frac{a(1-c)}{a+2} = \frac{2a}{(a+2)^2},$$

и это условие может быть выполнено при $\phi^*(\xi) < 1$. Теперь легко проверить, что

$$\psi^*(\eta) = \begin{cases} 0, & 0 \leq \eta \leq \frac{a}{a+2}, \\ 1, & \frac{a}{a+2} < \eta \leq 1 \end{cases}$$

и

$$\phi^*(\xi) = \begin{cases} \text{произвольна между 0 и 1,} & \text{но удовлетворяет условию} \\ \int_0^c \phi^*(\xi) d\xi = \frac{2a}{(a+2)^2}, & \text{при } 0 \leq \xi \leq \frac{a}{a+2}, \\ 1, & \text{при } \frac{a}{a+2} < \xi < 1, \end{cases}$$

то ϕ^* максимизирует (2.9.4) для ψ^* , а ψ^* минимизирует (2.9.5) для ϕ^* .

Интерпретация решения представляет некоторый интерес.

(1) Оба игрока при высоких картах делают ставку или уравнивают. Особое значение имеет тот факт, что оба игрока используют одну и ту же критическую точку $a/(a+2)$ для того, чтобы различать хорошие и плохие карты.

(2) Элемент блефа для игрока 1 проявляется только в том, что определена доля карт, на которой он мог бы блефовать; он может поставить a при получении карты из промежутка $[0, a/(a+2)]$ при единственном ограничении на вероятность ставки

$$\int_0^{a/(a+2)} \phi^*(\xi) d\xi = \frac{2a}{(a+2)^2}.$$

2.9.2. Модель покера с несколькими размерами ставки [Karlin, Restrepo, 1957].

Излагаемая в этом пункте модель является обобщением рассмотренной в п.2.9.1.

Как и раньше, в качестве представления всех возможных раскладов, которые могут быть сданы игроку, берется единичный интервал. Все расклады считаются равновероятными, так что операция сдачи расклада игроку может считаться эквивалентной выбору случайного числа из единичного интервала согласно равномерному распределению. Естественно, расклад ξ_1 считается ниже расклада ξ_2 тогда и только тогда, когда $\xi_1 < \xi_2$. Игра происходит следующим образом. Два игрока A и B выбирают соответственно точки ξ и η из единичного интервала согласно равномерному распределению. Оба игрока ставят по одной единице. Игрок A , зная свое значение ξ , ходит первым и имеет возможность либо сразу спасовать, теряя свою ставку в пользу игрока B , либо поставить любую из величин a_1, a_2, \dots, a_n , где $1 < a_1 < a_2 < \dots < a_n$. Игрок B должен в ответ или сразу спасовать, или уравнять ставку. В первом случае игрок A выигрывает ставку игрока B . Если же B уравнивает, то расклады ξ и η сравниваются и игрок с лучшим раскладом выигрывает банк. Если $\xi = \eta$, то никаких платежей не производится.

Стратегия игрока A может быть описана как n -мерная вектор-функция

$$\phi(\xi) = \langle \phi_1(\xi), \phi_2(\xi), \dots, \phi_n(\xi) \rangle,$$

где $\phi_i(\xi)$ есть вероятность того, что A поставит величину a_i , если его расклад равен ξ . Функции $\phi_i(\xi)$ должны удовлетворять условиям $\phi_i(\xi) \geq 0$ и

$$\sum_{i=1}^n \phi_i(\xi) \leq 1.$$

Вероятность того, что игрок A сразу спасует, равна

$$1 - \sum_{i=1}^n \phi_i(\xi).$$

Стратегия игрока B может быть представлена вектором

$$\psi(\eta) = \langle \psi_1(\eta), \psi_2(\eta), \dots, \psi_n(\eta) \rangle,$$

где $\psi_i(\eta)$ означает вероятность того, что B уравнивает ставку в a_i единиц, если он имеет расклад η . Вероятность того, что B спасует, если A поставил a_i , равна $1 - \psi_i(\eta)$. Каждая функция $\psi_i(\eta)$ удовлетворяет условию

$$0 \leq \psi_i(\eta) \leq 1.$$

Если A применяет стратегию ϕ , а B — стратегию ψ , то ожидаемый выигрыш игрока A обозначается через $K(\phi, \psi)$. Перечисляя все возможности, мы можем написать

$$\begin{aligned} K(\phi, \psi) = & (-1) \int_0^1 \left[1 - \sum_{i=1}^n \phi_i(\xi) \right] d\xi + \int_0^1 \int_0^1 \sum_{i=1}^n \phi_i(\xi) [1 - \psi_i(\eta)] d\xi d\eta + \\ & + \sum_{i=1}^n (a_i + 1) \int_0^1 \int_0^1 \phi_i(\xi) \psi_i(\eta) L(\xi, \eta) d\xi d\eta, \end{aligned}$$

где $L(\xi, \eta) = \text{sign}(\xi - \eta)$.

По определению ситуации равновесия, пара оптимальных стратегий ϕ^* и ψ^* удовлетворяет неравенствам

$$K(\phi^*, \psi) \geq K(\phi^*, \psi^*) \text{ для всех } \psi \quad (2.9.8)$$

и

$$K(\phi, \psi^*) \leq K(\phi^*, \psi^*) \text{ для всех } \phi. \quad (2.9.9)$$

Верно и обратное: если эти неравенства выполнены, то стратегии ϕ^*, ψ^* оптимальны. Таким образом, ϕ^* максимизирует $K(\phi, \psi^*)$, а ψ^* минимизирует $K(\phi^*, \psi)$.

Преобразуя выражение для $K(\phi, \psi)$, мы можем написать

$$K(\phi, \psi^*) = -1 + \sum_{i=1}^n \int_0^1 \phi_i(\xi) \left[2 + a_i \int_0^\xi \psi_i^*(\eta) d\eta - (a_i + 2) \int_\xi^1 \psi_i^*(\eta) d\eta \right] d\xi \quad (2.9.10)$$

и

$$K(\phi^*, \psi) = -1 + 2 \sum_{i=1}^n \int_0^1 \phi_i^*(\xi) d\xi + \sum_{i=1}^n \int_0^1 \psi_i(\eta) \left[-(a_i + 2) \int_0^\eta \phi_i^*(\xi) d\xi + a_i \int_\eta^1 \phi_i^*(\xi) d\xi \right] d\eta. \quad (2.9.11)$$

Соотношения (2.9.10) и (2.9.11) имеют вид

$$K(\phi, \psi^*) = C_1 + \sum_{i=1}^n \int_0^1 \phi_i(\xi) L_i(\xi) d\xi \quad (2.9.12)$$

и

$$K(\phi^*, \psi) = C_2 + \sum_{i=1}^n \int_0^1 \psi_i(\eta) K_i(\eta) d\eta, \quad (2.9.13)$$

где C_1 и C_2 не зависят соответственно от ϕ и ψ , а L_i и K_i обозначают соответственно выражения, стоящие в равенствах (2.9.10) и (2.9.11) во внешних скобках. Принимая во внимание ограничения, налагаемые на функции ϕ_1, \dots, ϕ_n , ясно, что для максимизации (2.9.10) или (2.9.12), игрок A должен выбирать $\phi_i(\xi) = 1$, если $L_i(\xi)$ положительно и превосходит все $L_j(\xi)$ для $j \neq i$; далее, он должен выбирать $\phi_i(\xi) = 0$, если $L_i(\xi) < 0$; наконец, если $L_i(\xi) = 0$, а остальные коэффициенты $L_j(\xi)$ ($j \neq i$) неположительны, то он может максимизировать $K(\phi, \psi^*)$, выбирая функцию $\phi_i(\xi)$ произвольно, лишь бы она удовлетворяла условию $0 \leq \phi_i(\xi) \leq 1$ (или, если более чем одно из чисел $L_i(\xi)$ равно нулю, условию $\sum \phi_i(\xi) \leq 1$, где суммирование распространяется по индексам, соответствующим $L_i(\xi) = 0$). Аналогично, для того, чтобы минимизировать величину (2.9.11) или (2.9.13), B должен выбирать $\psi_i(\eta) = 1$, если $K_i(\eta) < 0$, и $\psi_i(\eta) = 0$, если $K_i(\eta) > 0$. Если $K_i(\eta) = 0$, то значения $\psi_i(\eta)$ на выигрыш не влияют.

Руководствуясь интуитивными соображениями, построим две стратегии ϕ^* и ψ^* . Далее проверим, удовлетворяют ли эти стратегии условиям (2.9.8) и (2.9.9), и тем самым — оптимальны ли они. Главная задача построения состоит в том, чтобы гарантировать совместность этих стратегий, т. е. чтобы функция ϕ^* , которая максимизирует правую часть равенства (2.9.10), была той же функцией, которая содержится в выражении (2.9.11), и аналогичный факт должен иметь место для функции ψ^* .

Так как игрок B не имеет возможности блефовать, мы можем ожидать, что для некоторых b_i

$$\psi_i^*(\eta) = \begin{cases} 0, & \eta < b_i, \\ 1, & \eta > b_i. \end{cases} \quad (2.9.14)$$

Это утверждение справедливо, поскольку каждая из функций $K_j(\eta)$ является невозрастающей.

С другой стороны, мы можем ожидать, что игрок A , когда его расклад плох, иногда будет блефовать. Для того, чтобы учесть эту возможность, найдем критические числа b_i , определяющие функцию $\psi_i^*(\eta)$ так, чтобы коэффициент $L_i(\xi)$ при ϕ_i обращался для

$\xi < b_i$ в нуль. Это возможно, так как функция $L_i(\xi)$ на этом интервале постоянна. Тогда получаем

$$b_i = \frac{a_i}{2 + a_i}, \quad (2.9.15)$$

и $b_1 < b_2 \dots < b_n < 1$. Коэффициент $L_i(\xi)$ при ϕ_i на интервале $(0, b_i)$ равен нулю, а вне его представляет собой такую линейную функцию ξ , что

$$L_i(1) = 4\left(1 - \frac{1}{a_i + 2}\right).$$

Отсюда мы заключаем, что функции $L_i(\xi)$ и $L_j(\xi)$ совпадают в точке

$$c_{ij} = 1 - \frac{2}{(2 + a_i)(2 + a_j)}. \quad (2.9.16)$$

Очевидно, что c_{ij} является строго возрастающей функцией i и j . Положим $c_1 = b_1$ и $c_i = c_{i-1, i}$ для $i = 2, \dots, n$ и $c_{n+1} = 1$. Ясно, что при $j \neq i$ ξ $L_i(\xi) > L_j(\xi) \geq 0$ для ξ из интервала (c_i, c_{i+1}) . Согласно предыдущим рассуждениям, если функция ϕ^* максимизирует $K(\phi, \psi^*)$, то $\phi_i^*(\xi) = 1$ при $c_i < \xi < c_{i+1}$. Для определенности мы также положим $\phi_i^*(c_i) = 1$; это не существенно, так как если стратегия (функция) изменяется только в конечном числе точек (или на множестве лебеговой меры нуль), то выигрыш $K(\phi, \psi)$ остается неизменным.

Таким образом, мы показали, что если функция ψ^* определена согласно соотношениям (2.9.14), где

$$b_i = \frac{a_i}{2 + a_i},$$

то $K(\phi, \psi^*)$ максимизируется любой стратегией ϕ^* вида

$$\phi_i^*(\xi) = \begin{cases} \text{произвольна, } \xi < c_1 = b_1, \\ 0, & c_1 \leq \xi < c_i, \\ 1, & c_i \leq \xi < c_{i+1}, \\ 0, & c_{i+1} \leq \xi \leq 1, \end{cases} \quad (2.9.17)$$

где

$$\sum_{i=1}^n \phi_i^*(\xi) \leq 1, \quad \phi_i^*(\xi) \geq 0.$$

Значения $\phi_i^*(\xi)$ на интервале $0 \leq \xi < c_1$ все еще не определены в виду соотношения $L_i(\xi) \equiv 0$, которое имеет место на этом интервале.

Остается показать, что так построенная функция ψ^* действительно минимизирует $K(\phi^*, \psi)$. Для того, чтобы ψ^* заведомо обладало этим свойством, необходимо наложить на ϕ_i^* еще некоторые условия; для этого мы воспользуемся произволом, допущенным в определении ϕ^* , когда ξ меняется в интервале $[0, c_1)$. Для того, чтобы показать, что ψ^* минимизирует $K(\phi^*, \psi)$, достаточно установить, что коэффициент $K_i(\eta)$ при $\psi_i(\eta)$ неотрицателен для $\eta < b_i$ и неположителен для $\eta > b_i$. Так как $K_i(\eta)$ — непрерывная монотонно убывающая функция, последнее условие эквивалентно соотношению

$$-(a_i + 2) \int_0^{b_i} \phi_i^*(\xi) d\xi + a_i \int_{b_i}^1 \phi_i^*(\xi) d\xi = 0. \quad (2.9.18)$$

Подставляя функцию ϕ^* , определенную в (2.9.17) в соотношение (2.9.18) получаем равенства

$$2 \int_0^{b_1} \phi_1^*(\xi) d\xi = b_1(1 - b_1)(b_2 + 1), \quad 2 \int_0^{b_1} \phi_n^*(\xi) d\xi = b_n(1 - b_n)(1 - b_{n-1}), \quad (2.9.19)$$

и

$$2 \int_0^{b_1} \phi_i^*(\xi) d\xi = b_i(1 - b_i)(b_{i+1} - b_{i-1}), \quad i = 2, \dots, n-1. \quad (2.9.20)$$

Поскольку

$$\sum_{i=1}^n \phi_i^*(\xi) \leq 1,$$

эти равенства выполняются тогда и только тогда, когда

$$2 \int_0^{b_1} \sum_{i=1}^n \phi_i^*(\xi) d\xi \leq 2b_1. \quad (2.9.21)$$

Но сумма правых частей равенств (2.9.19) и (2.9.21) не превосходит $(2 + b_n - b_1)/4$, так как всегда $b_i(1 - b_i) \leq 1/4$. Из неравенства $b_1 \geq 1/3$, мы получаем

$$\frac{1}{4}(2 + b_n - b_1) \leq \frac{1}{4}(3 - b_1) \leq \frac{2}{3} \leq 2b_1.$$

Таким образом, условия (2.9.19)–(2.9.21) могут быть выполнены. Итак, неравенства (2.10.8) и (2.9.9) для стратегий ϕ^* и ψ^* доказаны, т. е. оптимальными стратегиями являются следующие функции

$$\psi_i^*(\eta) = \begin{cases} 0, & \eta < b_i, \\ 1, & \eta \geq b_i, \end{cases} \quad (2.9.22)$$

$$\phi_i^*(\xi) = \begin{cases} \text{произвольна, но удовлетворяет условию} \\ \int_0^{b_1} \phi_i^*(\xi) d\xi = b_i(c_{i+1} - c_i), & 0 \leq \xi < b_1 \\ 0, & b_1 \leq \xi < c_i \text{ или } c_{i+1} \leq \xi \leq 1, \\ 1, & c_i \leq \xi < c_{i+1}, \end{cases} \quad (2.9.23)$$

где

$$b_i = \frac{a_i}{2 + a_i}, \quad c_1 = b_1,$$

$$c_i = 1 - \frac{2}{(2 + a_i)(2 + a_{i-1})}, \quad i = 2, \dots, n, \quad c_{n+1} = 1$$

и

$$\sum_{i=1}^n \phi_i^*(\xi) \leq 1.$$

2.9.3. Модель покера с двумя кругами ставок (см. [Беллман, 1960, Karlin, Restrepo, 1957]). В этом пункте мы распространим модель покера, описанную в предыдущем пункте, на случай двух кругов ставок. В то же время, мы ограничимся случаем, когда допускается только один размер ставки. Мы снова предполагаем для удобства, что каждому игроку сдаются случайные расклады, равномерно распределенные на единичном интервале.

Выигрыш и стратегии. После того, как сделана начальная единичная ставка, игрок A ходит первым и имеет две альтернативы: он может пасовать или поставить a единиц. Затем ходит игрок B , который располагает тремя возможностями: он может пасовать, уравнивать ставку игрока A или, наконец, повысить, поставив $a + b$ единиц. Если B повысил, то на долю A остается либо пасовать, либо уравнивать ставку B .

Если A и B получили соответственно карты ξ и η , то их стратегии могут быть описаны следующим образом:

$\phi_1(\xi)$ — вероятность того, что игрок A ставит a и, если игрок B повышает, пасует.

$\phi_2(\xi)$ — вероятность того, что игрок A ставит a и, если игрок B повышает, то A уравнивает ставку B .

$1 - \phi_1(\xi) - \phi_2(\xi)$ — вероятность того, что A сразу пасует.

$\psi_1(\eta)$ — вероятность того, что игрок B уравнивает начальную ставку.

$\psi_2(\eta)$ — вероятность того, что игрок B повышает.

$1 - \psi_1(\eta) - \psi_2(\eta)$ — вероятность того, что игрок B пасует.

Ожидаемый выигрыш при этом равен

$$\begin{aligned} K(\phi, \psi) = & - \int [1 - \phi_1(\xi) - \phi_2(\xi)] d\xi + \int \int [\phi_1(\xi) + \phi_2(\xi)] [1 - \psi_1(\eta) - \psi_2(\eta)] d\xi d\eta + \\ & + (a + 1) \int \int \phi_1(\xi) \psi_1(\eta) L(\xi, \eta) d\xi d\eta - (a + 1) \int \int \phi_1(\xi) \psi_2(\eta) d\xi d\eta \\ & + (a + 1) \int \int \phi_2(\xi) \psi_1(\eta) L(\xi, \eta) d\xi d\eta + (1 + a + b) \int \int \phi_2(\xi) \psi_2(\eta) L(\xi, \eta) d\xi d\eta, \end{aligned}$$

где $L(\xi, \eta) = \text{sign}(\xi - \eta)$. (Этот ожидаемый выигрыш получается путем рассмотрения взаимно исключающих возможностей: A пасует; A ставит и B пасует; A применяет стратегию ϕ_1 и B уравнивает или повышает; A применяет стратегию ϕ_2 и B уравнивает или повышает.)

Если мы обозначим оптимальные стратегии через

$$\phi^*(\xi) = \langle \phi_1^*(\xi), \phi_2^*(\xi) \rangle, \psi^*(\eta) = \langle \psi_1^*(\eta), \psi_2^*(\eta) \rangle$$

и перегруппируем члены в выражении $K(\phi, \psi)$, как это делалось в предыдущих примерах, то получим

$$\begin{aligned} K(\phi, \psi^*) = & -1 + \int_0^1 \phi_1(\xi) \left[2 + a \int_0^\xi \psi_1^*(\eta) d\eta - (a + 2) \int_\xi^1 \psi_1^*(\eta) d\eta - \right. \\ & \left. - (a + 2) \int_0^1 \psi_2^*(\eta) d\eta \right] d\xi + \int_0^1 \phi_2(\xi) \left[2 + a \int_0^\xi \psi_1^*(\eta) d\eta - (a + 2) \int_\xi^1 \psi_1^*(\eta) d\eta + \right. \\ & \left. + (a + b) \int_0^\xi \psi_2^*(\eta) d\eta - (a + b + 2) \int_\xi^1 \psi_2^*(\eta) d\eta \right] d\xi, \end{aligned} \quad (2.9.24)$$

$$\begin{aligned}
K(\phi^*, \psi) = & \int_0^1 [-1 + 2\phi_1^*(\xi) + 2\phi_2^*(\xi)]d\xi + \int_0^1 \psi_1(\eta) \left[-(a+2) \int_0^\eta [\phi_1^*(\xi) + \phi_2^*(\xi)]d\xi + \right. \\
& + a \int_\eta^1 [\phi_1^*(\xi) + \phi_2^*(\xi)]d\xi \Big] d\eta + \int_0^1 \psi_2(\eta) \left[-(a+2) \int_0^1 \phi_1^*(\xi)d\xi - \right. \\
& \left. -(a+b+2) \int_0^\eta \phi_2^*(\xi)d\xi + (a+b) \int_\eta^1 \phi_2^*(\xi)d\xi \right] d\eta.
\end{aligned} \tag{2.9.25}$$

Поиск оптимальных стратегий. Будем, как и раньше, искать функции $\phi^*(\xi)$, которые максимизируют (2.9.24) и функции $\psi^*(\eta)$, которые минимизируют выражение (2.9.25). Интуитивные соображения подсказывают, что игрока A может в некоторых случаях блефовать с плохими картами, намереваясь спасовать, если B повышает; он будет также в некотором промежуточном интервале, скажем, при $c < \xi < e$, выбирать $\phi_1^*(\xi) = 1$; он будет выбирать стратегию $\phi_2^*(\xi) = 1$ для $e \leq \xi \leq 1$. Игрок B мог бы иногда блефовать, повышая ставки в интервале $0 \leq \eta < c$; он мог бы выбирать $\psi_1^* = 1$ в интервале $c \leq \eta < d$, и $\psi_2^* = 1$ в интервале $d \leq \eta \leq 1$.

Отсюда вовсе не следует, что это единственный возможный вид оптимальных стратегий. Действительно, далее мы увидим, что существуют оптимальные стратегии и иного вида. Однако, после того, как определена одна пара оптимальных стратегий, сравнительно легко найти и все остальные решения. Поэтому, прежде всего мы попытаемся найти оптимальные стратегии указанного вида. Для этого нам нужны такие значения c , d и e , которые порождают решения требуемого типа. Из построения ψ_1^* мы сразу видим, что для того случая, когда $\xi \leq c$ коэффициент при $\phi_1(\xi)$ в выражении (2.9.24) есть константа; составляя его выражение и приравнявая его к нулю, мы получаем

$$2 - (a+2) \int_0^1 \psi_2^*(\eta)d\eta - (a+2)(d-c) = 0. \tag{2.9.26}$$

Коэффициенты при ψ_1^* и ψ_2^* должны быть равны в точке d , в которой игрок B изменяет характер своего поведения

$$(2a+2) \int_d^1 \phi_1^*(\xi)d\xi = -b \int_0^d \phi_2^*(\xi)d\xi + b \int_d^1 \phi_2^*(\xi)d\xi. \tag{2.9.27}$$

Аналогичное условие при $\xi = e$ требует, чтобы выполнялось равенство

$$(2a+b+2) \int_0^e \psi_2^*(\eta)d\eta = b \int_e^1 \psi_2^*(\eta)d\eta. \tag{2.9.28}$$

В точке $\eta = c$, в которой игрок B начинает применять стратегии ψ_1^* и ψ_2^* без блефа, соответствующие коэффициенты (которые являются убывающими функциями) изменяют знак с плюса на минус, т. е. они должны при $\eta = c$ обращаться в нуль. Следовательно,

$$\begin{aligned}
-(a+2) \int_0^c [\phi_1^*(\xi) + \phi_2^*(\xi)]d\xi + a \int_c^1 [\phi_1^*(\xi) + \phi_2^*(\xi)]d\xi &= 0, \\
-(a+2) \int_0^1 \phi_1^*(\xi)d\xi + (a+b) \int_c^1 \phi_2^*(\xi)d\xi &= 0.
\end{aligned} \tag{2.9.29}$$

(При выводе (2.9.29) мы положили $\phi_2^*(\xi) = 0$ для $0 \leq \xi \leq c$; это интуитивно ясно.) Введем теперь обозначения:

$$m_1 = \int_0^c \phi_1^*(\xi)d\xi, \quad m_2 = \int_0^c \psi_2^*(\eta)d\eta.$$

Вспоминая предположение о виде решения и предполагая, что $c < e < d$, мы можем уравнения (2.9.26)–(2.9.29) записать следующим образом:

$$2 = (a + 2)(m_2 + 1 - c), \quad (2.9.30)$$

$$1 - d = d - e \text{ или } 2(1 - d) = (1 - e), \quad (2.9.31)$$

$$(2a + b + 2)m_2 = b(1 - d), \quad (2.9.32)$$

$$(a + 2)m_1 = a(1 - c), \quad (2.9.33)$$

$$(a + 2)(m_1 + e - c) = (a + b)(1 - c). \quad (2.9.34)$$

Мы получили систему из пяти уравнений с пятью неизвестными: m_1, m_2, c, d, e ; докажем теперь, что эта система уравнений имеет решение, согласующееся с предположениями, сделанными ранее, а именно, что $0 < c < e < d < 1$, $0 < m_1 < 0$, $0 < m_2 < c$.

Решение уравнений (2.9.30)–(2.9.34). Эта система может быть решена в явном виде следующим образом. Перепишем последнее уравнение в виде

$$(a + 2)(m_1 + 1 - c) = (2a + b + 2)(1 - e).$$

Исключая с помощью уравнений (2.9.33) и (2.9.31) m_1 и $(1 - e)$, мы получим

$$(a + 1)(1 - c) = (2a + b + 2)(1 - d). \quad (2.9.35)$$

Из оставшихся уравнений исключаем m_2 ; тогда

$$\frac{2}{a + 2} - (1 - c) = \frac{b}{2a + b + 2}(1 - d). \quad (2.9.36)$$

Следовательно,

$$(1 - d) \left(\frac{b}{2a + b + 2} + \frac{2a + b + 2}{a + 1} \right) = \frac{2}{a + 2}. \quad (2.9.37)$$

Найдя $(1 - d)$, мы можем разрешить (2.9.36) относительно $(1 - c)$, а тогда значения остальных неизвестных находятся из исходных уравнений.

Для того, чтобы показать, что это решение согласуется с поставленными условиями, заметим прежде всего, что из соотношения (2.9.37) следует $(1 - d) > 0$. Уравнение (2.9.35) показывает, что $(1 - c) > (1 - d)$, и, следовательно, $c < d$. Кроме того, из соотношения (2.9.36) следует, что

$$c = \left(\frac{b}{2a + b + 2} \right)(1 - d) + \left(1 - \frac{2}{a + 2} \right) > 0. \quad (2.9.38)$$

Так как $2(a + 1)(1 - c) = (2a + b + 2)(1 - e)$, мы заключаем, что $(1 - e) < (1 - c)$, или $c < e$; а так как $2d = 1 + e$, мы должны иметь $e < d$. Итак, мы показали, что $0 < c < e < d < 1$. Для проверки двух оставшихся условий заметим, что из уравнения (2.9.30) следует, что

$$m_2 = c - \left(1 - \frac{2}{a + 2} \right),$$

так как $m_2 < c$, и согласно (2.9.38), $m_2 > 0$.

Наконец, используя уравнение (2.9.33) и (2.9.30), мы заключаем, что

$$\begin{aligned} m_1 &= \frac{a}{a+2}(1-c) = (1-c) - \frac{2}{a+2}(1-c) = \\ &= (1-c)[1 - (m_2 + 1 - c)] = (1-c)(c - m_2), \end{aligned}$$

и поэтому $0 < m_1 < c$.

Оптимальность стратегий ϕ^* и ψ^* . Итак, мы можем выразить ϕ^* и ψ^* через найденные выше значения величин c, e, d, m_1, m_2 :

$$\phi_1^*(\xi) = \begin{cases} 1, & c \leq \xi < e, \\ 0, & e \leq \xi \leq 1, \end{cases} \quad \phi_2^*(\xi) = \begin{cases} 0, & 0 \leq \xi < e, \\ 1, & e \leq \xi \leq 1, \end{cases} \quad (2.9.39)$$

$$\psi_1^*(\eta) = \begin{cases} 0, & 0 \leq \eta < c, \\ 1, & c \leq \eta < d, \\ 0, & d \leq \eta \leq 1, \end{cases} \quad \psi_2^*(\eta) = \begin{cases} 0, & c \leq \eta < d, \\ 1, & d \leq \eta \leq 1. \end{cases}$$

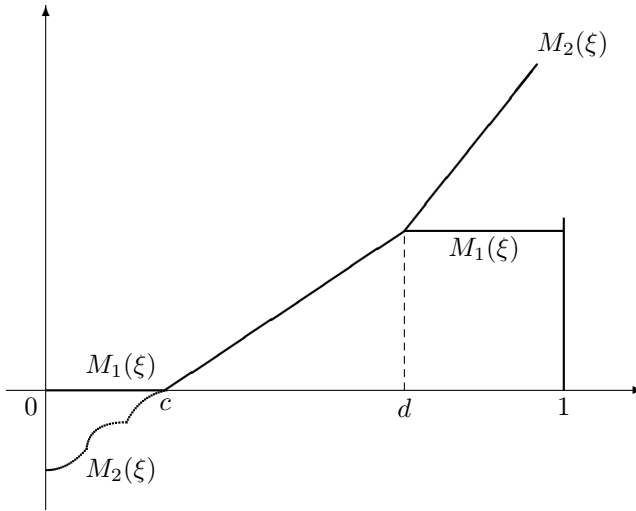
В оставшемся интервале $0 \leq \eta < c$ функции $\phi_1^*(\xi)$ и $\psi_2^*(\eta)$ выбираются произвольно, но с учетом того, что они заключены между нулем и единицей и удовлетворяют условиям

$$\int_0^c \phi_1^*(\xi) d\xi = m_1 \text{ и } \int_0^c \psi_2^*(\eta) d\eta = m_2.$$

Остается проверить, что стратегии ϕ^* и ψ^* , описанные выше, соответственно максимизируют $K(\phi, \psi^*)$ (2.9.24) и минимизируют $K(\phi^*, \psi)$ (2.9.25). Для этого исследуем сначала коэффициенты $M_1(\xi)$ и $M_2(\xi)$ при ϕ_1 и ϕ_2 в выражении для $K(\phi, \psi^*)$. По построению, коэффициент $M_1(\xi)$ при ϕ_1 на интервале $[0, c]$ равен тождественно нулю, линейно возрастает на интервале $[c, d]$, и в остальной части остается постоянным. Кроме того, функция $M_1(\xi)$ непрерывна во всем промежутке $[0, 1]$. Заметим теперь, что функция $M_2(\xi)$ линейна на $[c, d]$ и имеет тот же наклон, что и $M_1(\xi)$. Кроме того, эти функции совпадают при $\xi = e$ в промежутке $[c, d]$ согласно (2.9.28), поэтому $M_1 = M_2$ для $\xi \in [c, d]$.

Из определения ψ^* мы можем сразу заключить, что функция M_2 строго возрастает во всем промежутке $[0, 1]$ (см. рис. 2.7). Используя эти факты, легко осуществить максимизацию $K(\phi, \psi^*)$. Ясно, что максимум достигается на любой функции ϕ , обладающей следующими свойствами: (а) $\phi_2 = 0$ и ϕ_1 произвольна ($0 \leq \phi_1 \leq 1$) для $\xi \in [0, c]$; (б) функции удовлетворяют условию $\phi_1 + \phi_2 = 1$, а в остальном произвольны для $\xi \in [c, d]$; (в) $\phi_2 = 1$ для $\xi \in [d, 1]$. Очевидно, что определенная выше функция ϕ^* этим условиям удовлетворяет.

Исследование коэффициентов $N_1(\eta)$ и $N_2(\eta)$ при $\psi_1(\eta)$ и $\psi_2(\eta)$ в выражении для $K(\phi^*, \psi)$ показывает, что они имеют вид, указанный на рис. 2.8. Заметим, что функция $K(\phi^*, \psi)$ минимизируется любой функцией ψ , обладающей свойствами: (а') $\psi_1 = 0$ и ψ_2 произвольна ($0 \leq \psi_2 \leq 1$) для $\eta \in [0, c]$; (б') $\psi_1 = 1$ для $\eta \in [c, d]$; (в') $\psi_2 = 1$ для $\eta \in [d, 1]$. Ясно, что описанная выше функция ψ^* этим требованиям удовлетворяет. Доказательство оптимальности стратегий ϕ^* и ψ^* завершено.

Рис. 2.7. Коэффициенты $M_1(\xi)$ и $M_2(\xi)$

В качестве иллюстрации приведем пример. Пусть $a = b = 2$; тогда

$$\phi_1^*(\xi) = \begin{cases} \int_0^{19/35} \phi_1^* d\xi = 8/35, & 0 \leq \xi < 19/35, \\ 1, & 19/35 \leq \xi < 23/35, \\ 0, & 23/35 \leq \xi \leq 1, \end{cases}$$

$$\phi_2^*(\xi) = \begin{cases} 0, & 0 \leq \xi < 23/35, \\ 1, & 23/35 \leq \xi \leq 1, \end{cases}$$

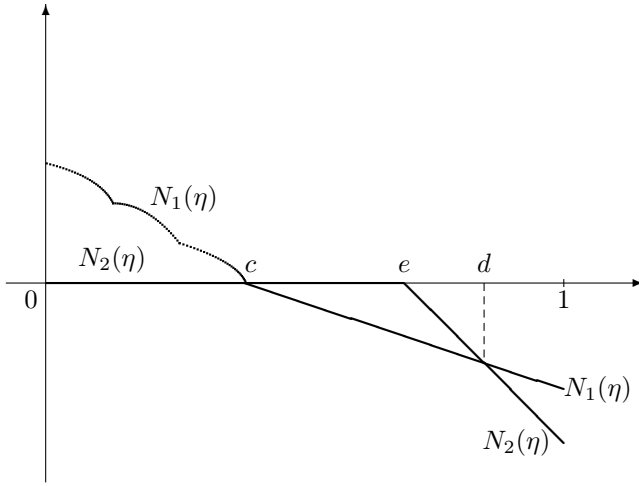
$$\psi_1^*(\eta) = \begin{cases} 0, & 0 \leq \eta < 19/35, \\ 1, & 19/35 \leq \eta < 29/35, \\ 0, & 29/35 < \eta \leq 1, \end{cases}$$

$$\psi_2^*(\eta) = \begin{cases} \int_0^{19/35} \psi_2^* d\xi = 3/70, & 0 \leq \eta < 19/35, \\ 0, & 19/35 \leq \eta < 29/35, \\ 1, & 29/35 \leq \eta \leq 1. \end{cases}$$

Значение этой игры равно $-11/35$.

Общее решение. Из рассмотрения рис. 2.8 непосредственно следует, что минимизирующее решение не может иметь никакого иного вида, кроме указанного в соотношениях (2.9.39). Однако при максимизации функции $K(\phi, \psi^*)$ мы нашли, что на интервале $[c, d]$ единственное необходимое условие, которому должно удовлетворять максимизирующее ϕ , состоит в том, что $\phi_1 + \phi_2 = 1$. Изменяя ϕ_1 и ϕ_2 на этом интервале так, чтобы выполнялось это условие, мы можем найти все возможные оптимальные стратегии игрока А. С этой целью мы определим $\tilde{\phi}_1(\xi)$ на $[0, c]$ так, чтобы выполнялось соотношение

$$\int_0^c \tilde{\phi}_1 d\xi = m_1,$$

Рис. 2.8. Коэффициенты $N_1(\eta)$, $N_2(\eta)$

(значение m_1 вычислено выше) и $\tilde{\phi}_2(\xi) = 1$ на $[d, 1]$. Для ξ из $[c, d]$ мы потребуем только, чтобы $\tilde{\phi}_1 + \tilde{\phi}_2 = 1$. Выписывая условия, при которых $K(\tilde{\phi}, \psi)$ минимизируется функцией ψ^* , мы получаем

$$1 - d = \int_c^d \tilde{\phi}_2(\xi) d\xi \text{ и } \int_\eta^d \tilde{\phi}_1(\xi) d\xi \leq \frac{b(d - \eta)}{a + b + 1} \text{ для } \eta \in [c, d], \quad (2.9.40)$$

где c и d те же, что и выше. Мы получаем эти условия, приравнявая коэффициенты при $\psi_1(\eta)$ и $\psi_2(\eta)$ в выражении (2.9.25) в точке $\eta = d$ и требуя, чтобы для η из $[c, d]$ имело место неравенство $N_1(\eta) \leq N_2(\eta)$.

Легко видеть, что это условие необходимо и достаточно для того, чтобы стратегия $\tilde{\phi}$ была оптимальной.

2.9.4. Модель покера с k повышениями [Karlin, Restrepo, 1957].

В этом разделе мы найдем оптимальные стратегии в модели покера с несколькими кругами ставок. Метод исследования представляет собой, в сущности, обобщение метода, использованного в предыдущем пункте. Значительно более сложные детали доказательства мы приводить не будем.

Правила, стратегии и выигрыш. Два игрока ставят по единице каждый и получают независимо случайные расклады ξ и η (которые отождествляются с точками единичного интервала), распределенные равномерно. Имеется $k + 1$ кругов ставок («круг» в этом пункте означает одно действие того и иного из игроков). В первом круге игрок A может или спасовать (и потерять единицу), или поставить a единиц. A и B ходят поочередно. В каждом последующем круге игрок может либо спасовать, либо уравнивать ставку противника (в этих случаях игра заканчивается), либо повысить ставку на a единиц. В последнем круге каждый игрок может только пасовать или уравнивать. Если k четно, то последний возможный круг заканчивает игрок A ; если k нечетно, то последний возможный круг заканчивает игрок B .

Стратегию игрока A можно описать набором k функций $\phi = \langle \phi_1(\xi), \phi_2(\xi), \dots, \phi_k(\xi) \rangle$. Эти функции указывают способ действия игрока A , если он получает расклад ξ . Точнее,

$$1 - \sum_{i=1}^k \phi_i(\xi)$$

есть вероятность того, что игрок A сразу спасует, а

$$\sum_{i=1}^k \phi_i(\xi)$$

вероятность того, что игрок A в первом круге сделает ставку. Пусть далее,

$\phi_1(\xi)$ = вероятность того, что игрок A спасует на своем втором круге;

$\phi_2(\xi)$ = вероятность того, что игрок A уравнивает на своем втором круге;

$\sum_{i=3}^k \phi_i(\xi)$ = вероятность того, что игрок A повысит на своем втором круге, если есть возможность повышать, т. е. если игрок B повысил на своем первом круге и продолжает игру. Аналогично, если игра продолжается до r -го круга игрока A , то пусть

$\phi_{2r-3}(\xi)$ = вероятность того, что игрок A будет пасовать на своем r -ом круге;

$\phi_{2r-2}(\xi)$ = вероятность того, что игрок A будет уравнивать на своем r -ом круге;

$\sum_{i=2r-1}^k \phi_i(\xi)$ = вероятность того, что игрок A повысит на своем r -ом круге.

Аналогично стратегия игрока B может быть описана набором k функций:

$$\psi = \langle \psi_1(\eta), \dots, \psi_k(\eta) \rangle$$

который указывает образ действий игрока B , если он получает расклад η . Вероятность того, что игрок B при первой возможности будет пасовать, равна

$$\psi_0(\eta) = 1 - \sum_{j=1}^k \psi_j(\eta).$$

Если игра продолжается до r -го круга игрока B , то

$\psi_{2r-2}(\eta)$ — вероятность того, что игрок B спасует на своем r -ом круге;

$\psi_{2r-1}(\eta)$ — вероятность того, что игрок B будет уравнивать на своем r -ом круге;

$\sum_{j=2r}^k \psi_j(\eta)$ — вероятность того, что игрок B повысит на своем r -ом круге.

Если игроки получают расклады ξ и η и выбирают соответственно стратегии ϕ и ψ , то выигрыш игрока A может быть вычислен, как и в предыдущих примерах, путем рассмотрения взаимно исключающих случаев, которыми могут закончиться ставки.

Выигрыш игрока A имеет следующий вид:

$$\begin{aligned}
 P[\phi(\xi), \psi(\eta)] = & (-1) \left(1 - \sum_{i=1}^k \phi_i(\xi) \right) + \sum_{i=1}^k \phi_i(\xi) \left[1 - \sum_{j=1}^k \psi_j(\eta) + (a+1)\psi_1(\eta)L(\xi, \eta) \right] + \\
 & + \sum_{j=2}^k \psi_j(\eta) \{ -(a+1)\phi_1(\xi) + (2a+1)\phi_2(\xi)L(\xi, \eta) \} + \\
 & + \sum_{j=2r-2}^k \psi_j(\eta) \{ -[(2r-3)a+1]\phi_{2r-3}(\xi) + [(2r-2)a+1]\phi_{2r-2}(\xi)L(\xi, \eta) \} + \\
 & + \sum_{i=2r-1}^k \phi_i(\xi) \{ [(2r-2)a+1]\psi_{2r-2}(\eta) + [(2r-1)a+1]\psi_{2r-1}(\eta)L(\xi, \eta) \} + \\
 & + \sum_{j=2r}^k \psi_j(\eta) \{ -[(2r-1)a+1]\phi_{2r-1}(\xi) + (2ra+1)\phi_{2r}(\xi)L(\xi, \eta) \} + \\
 & + \sum_{i=2r+1}^k \phi_i(\xi) \{ (2ra+1)\psi_{2r}(\eta) + [(2r+1)a+1]\psi_{2r+1}(\eta)L(\xi, \eta) \},
 \end{aligned}$$

где $L(\xi, \eta) = \text{sign}(\xi - \eta)$. Математическое ожидание выигрыша равно

$$K(\phi, \psi) = \int_0^1 \int_0^1 P[\phi(\xi), \psi(\eta)] d\xi d\eta. \quad (2.9.41)$$

Описание оптимальных стратегий. Можно показать, что существуют оптимальные стратегии ϕ^* и ψ^* , характеризуемые $2k+1$ числами $b, c_1, \dots, c_k, d_1, \dots, d_k$. Когда игрок получает расклад ξ из промежутка $(0, b)$, он будет часть времени блефовать, а также некоторое время пасовать. Мы будем обозначать через

$$m_i = \int_0^b \phi_i^*(\xi) d\xi, \quad i = 1, 3, 5, \dots \quad (2.9.42)$$

и

$$n_j = \int_0^b \psi_j^*(\eta) d\eta, \quad j = 2, 4, 6, \dots \quad (2.9.43)$$

вероятности блефа на различных кругах ставок. Если игрок A получает расклад ξ в интервале (c_{i-1}, c_i) , где $c_0 = b$, то он будет выбирать $\phi_i^*(\xi) = 1$ и $\phi_l^*(\xi) = 0$ для $l \neq i$. Аналогично, если игрок B получает расклад η в интервале (d_{j-1}, d_j) , где $d_0 = b$, то он будет выбирать $\psi_j^*(\eta) = 1$ и $\psi_l^*(\eta) = 0$ для $l \neq j$. Решение изображено на рис. 2.9. Тот факт, что

$$c_{2r-1} < d_{2r-1} < d_{2r} < c_{2r}, \quad r = 1, 2, \dots$$

имеет большое значение.

Постоянные c_i, d_j, m_i , и n_j определяются путем решения громоздкой системы уравнений, аналогичной системе (2.9.30)-(2.9.34). Точнее, если k четно, то числа b, c_i, d_j, m_i , и n_j находятся как решения следующих уравнений:

$$\begin{aligned}
 [(4r-1)a+2] \sum_{j=2r}^k n_j &= a(1-d_{2r-1}), \quad 2r=2,4,\dots,k, \\
 a(c_{2r-2}-d_{2r-2}) &= a(1-c_{2r-2}) + [(4r-3)a+2] \sum_{j=2r}^k n_j, \quad 2r=4,6,\dots,k, \\
 [(4r-3)a+2] \sum_{i=2r-1}^k m_i &= a(1-c_{2r-2}), \quad 2r=2,4,\dots,k, \\
 a(d_{2r-1}-c_{2r-1}) &= a(1-d_{2r-1}) + [(4r-1)a+2] \sum_{i=2r+1}^k m_i, \quad 2r=2,4,\dots,k, \\
 (4ra+2)(c_{2r}-d_{2r-1}) &= [(4r+2)a+2](c_{2r}-d_{2r}), \quad 2r=2,4,\dots,k, \\
 [(4r-2)a+2](d_{2r-1}-c_{2r-2}) &= (4ra+2)(d_{2r-1}-c_{2r-1}), \quad 2r=2,4,\dots,k, \\
 2 &= (a+2) \int_0^1 \sum_{j=1}^k \psi_j^*(\eta) d\eta.
 \end{aligned}$$

Аналогичная система может быть составлена и для нечетного k . Получаемое при этом решение согласуется со схемой рис. 2.9.

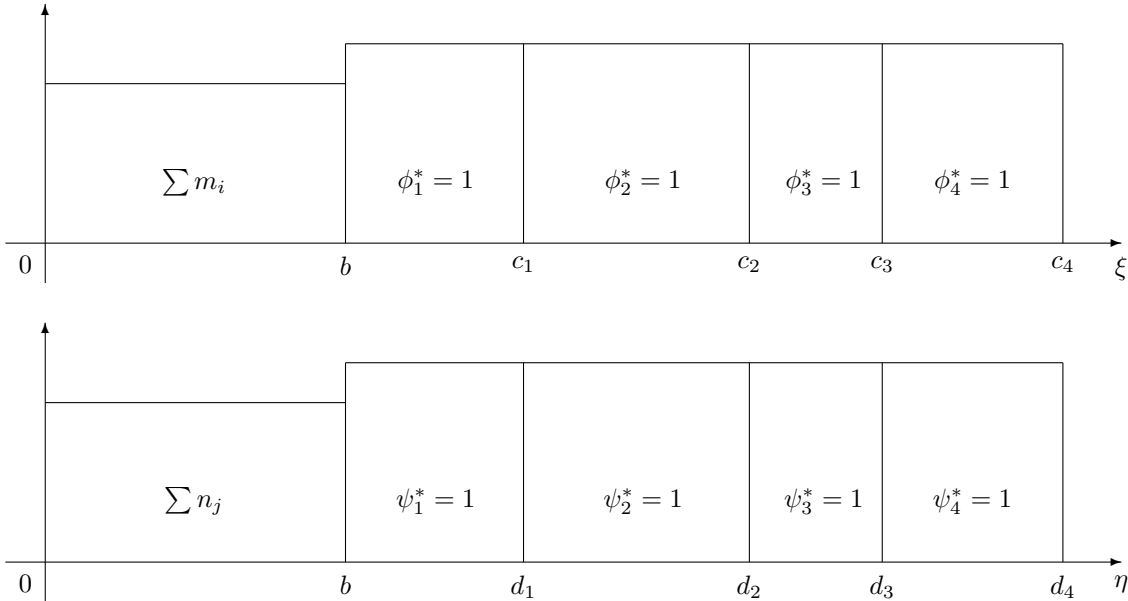


Рис. 2.9. Оптимальные стратегии ϕ^*, ψ^*

2.9.5. Пожер с одновременными ходами. [von Neumann, Morgenstern (1944), Karlin (1959)]. Два игрока A и B после получения случайных равномерно распределенных раскладов одновременно делают ставки. Начальная ставка может быть или b (низкая ставка) или a (высокая ставка). Если обе ставки равны, то игрок с более высоким раскладом выигрывает. Если один игрок поставил высокую ставку, а другой - низкую, то поставивший низкую ставку имеет выбор: либо спастись (теряя свою ставку), либо уравнивать высокую ставку, поставив дополнительно $-b$.

Если игрок, поставивший ставку, уравнивает, то игрок с более высоким раскладом выигрывает банк.

Так как игра симметрична, достаточно описать стратегии одного из игроков. Если игрок A получает расклад ξ , то мы будем обозначать через

$\phi_1(\xi)$ = вероятность того, что игрок A поставит низкую ставку и спасует, если игрок B поставит высокую ставку;

$\phi_2(\xi)$ = вероятность того, что игрок A поставит низкую ставку и затем уравнивает ставку B ;

$\phi_3(\xi)$ = вероятность того, что игрок A поставит высокую ставку.

Разумеется, эти функции удовлетворяют условиям

$$\phi_i(\xi) \geq 0, \quad \sum_{i=1}^3 \phi_i(\xi) = 1,$$

Ожидаемый выигрыш игрока A , если он использует стратегию ϕ , а игрок B использует стратегию ψ , выражается в виде

$$\begin{aligned} K(\phi, \psi) &= b \int_0^1 \int_0^1 [\phi_1(\xi) + \phi_2(\xi)][\psi_1(\eta) + \psi_2(\eta)]L(\xi, \eta)d\xi d\eta - \\ &- b \int_0^1 \int_0^1 [\phi_1(\xi)\psi_3(\eta)d\xi d\eta + b \int_0^1 \int_0^1 \phi_3(\xi)\psi_1(\eta)d\xi d\eta + \\ &+ a \int_0^1 \int_0^1 \phi_2(\xi)\psi_3(\eta)L(\xi, \eta)d\xi d\eta + \\ &+ a \int_0^1 \int_0^1 \phi_3(\xi)\psi_2(\eta)L(\xi, \eta)d\xi d\eta + \\ &+ a \int_0^1 \int_0^1 \phi_3(\xi)\psi_3(\eta)L(\xi, \eta)d\xi d\eta. \end{aligned}$$

Ввиду симметрии игры мы можем заменить в этом выражении стратегию $\psi(\eta)$ игрока B оптимальной стратегией $\phi^*(\eta)$. Сделаем довольно правдоподобное предположение о том, что в условиях этой стратегии $\phi_2^*(\eta) = 0$, ибо, по-видимому, нет веских оснований сначала делать низкую ставку, а затем уравнивать. Это предположение далее будет строго обосновано.

Так как $\phi_2^*(\eta) = 0$, мы можем записать функцию $K(\phi, \phi^*)$ следующим образом:

$$\begin{aligned}
 K(\phi, \phi^*) &= b \int_0^1 \int_0^1 [\phi_1(\xi) + \phi_2(\xi)] \phi_1^*(\eta) L(\xi, \eta) d\xi d\eta - b \int_0^1 \int_0^1 \phi_1(\xi) \phi_3^*(\eta) d\xi d\eta + \\
 &\quad + b \int_0^1 \int_0^1 \phi_3(\xi) \phi_1^*(\eta) d\xi d\eta + \\
 &\quad + a \int_0^1 \int_0^1 [1 - \phi_1(\xi) - \phi_3(\xi)] \phi_3^*(\eta) L(\xi, \eta) d\xi d\eta + a \int_0^1 \int_0^1 \phi_3(\xi) \phi_3^*(\eta) L(\xi, \eta) d\xi d\eta = \\
 &= \int_0^1 \phi_1(\xi) \left[b \int_0^\xi \phi_1^*(\eta) d\eta - b \int_\xi^1 \phi_1^*(\eta) d\eta - b \int_0^1 \phi_3^*(\eta) d\eta - \right. \\
 &\quad \left. - a \int_0^\xi \phi_3^*(\eta) d\eta + a \int_\xi^1 \phi_3^*(\eta) d\eta \right] d\xi + \int_0^1 \phi_2(\xi) \left[b \int_0^\xi \phi_1^*(\eta) d\eta - b \int_\xi^1 \phi_1^*(\eta) d\eta \right] d\xi + \\
 &\quad + \int_0^1 \phi_3(\xi) \left[b \int_0^1 \phi_1^*(\eta) d\eta \right] d\xi + \int_0^1 \int_0^1 \phi_3^*(\eta) L(\xi, \eta) d\xi d\eta, \tag{2.9.44}
 \end{aligned}$$

или

$$K(\phi, \phi^*) = \int_0^1 \sum \phi_i(\xi) T_i(\xi) d\xi + Z,$$

где Z не зависит от ϕ_i . Стратегия ϕ , максимизирующая $K(\phi, \phi^*)$, определяется выбором компоненты ϕ_i так, чтобы она была как можно больше, если $T_i(\xi) = \max_j T_j(\xi)$. Если максимум $T_j(\xi)$ достигается одновременно на двух из T_i , то соответствующие ϕ_i могут принимать любые положительные значения при условии, что их сумма равна единице.

При низких раскладах иногда рекомендуется блеф. Это означает, что при $\xi < \xi_0$ мы должны иметь $T_1(\xi) = T_3(\xi) > T_2(\xi)$. Дифференцируя тождество $T_1(\xi) = T_3(\xi)$, и помня, что на интервале $[0, \xi_0]$

$\phi_1^*(\xi) + \phi_3^*(\xi) = 1$, мы заключаем, что

$$\phi_1^*(\xi) = \frac{a}{a+b} \text{ почти всюду для } \xi < \xi_0.$$

Выбирая ϕ_1^* указанным образом и полагая $\phi_3^*(\xi) = 1$ для $\xi > \xi_0$, мы получаем, что равенство $T_1(\xi) = T_3(\xi)$ возможно, если

$$\xi_0 = \frac{a-b}{a}.$$

Таким образом, мы приходим к следующему решению:

$$\phi^* = \begin{cases} \phi_1^*(\xi) = \frac{a}{a+b}, & \phi_3^*(\xi) = \frac{b}{a+b}, & \text{при } \xi < \xi_0, \\ \phi_3^*(\xi) = 1, & & \text{при } \xi > \xi_0. \end{cases} \tag{2.9.45}$$

Проверка оптимальности так построенной стратегии ϕ^* осуществляется обычными приемами. Ясно, что $T_1(\xi) = T_3(\xi) > T_2(\xi)$ для $\xi < \xi_0$, и, следовательно, максимум достигается лишь при условии $\phi_1 + \phi_3 = 1$, которое, несомненно, выполнено для ϕ^* вида (2.9.45). Далее, мы видели, что равенство $T_1(\xi) = T_3(\xi)$ единственным образом определяет ϕ^* в соответствии с формулой (2.9.45) для $\xi < \xi_0$.

Для $\xi > \xi_0$, исследуя (2.9.44), мы находим, что $T_2(\xi) = T_3(\xi) > T_1(\xi)$. Следовательно, максимизация функции $K(\phi, \phi^*)$ требует, чтобы ϕ удовлетворяла условию $\phi_2 + \phi_3 = 1$. Но, если на этом интервале $\phi_2^* > 0$, то несложное вычисление показывает, что $T_1(\xi) < T_2(\xi)$ для $\xi \geq \xi_1$, где $\xi_1 < \xi_0$. В совокупности все эти доводы доказывают, что стратегия ϕ^* вида (2.9.45) есть единственная оптимальная стратегия игры.

§ 2.10. Упражнения и задачи

1. Игра нападения — защиты. Игрок 1 силами A единиц намерен атаковать один из объектов C_1, \dots, C_n , ценность которых определяется числами $\tau_1 > 0, \tau_2 > 0, \dots, \tau_n > 0$, причем $\tau_1 \geq \tau_2 \geq \dots \geq \tau_n$. Чистой стратегией x игрока 1 является вектор $x = (\xi_1, \dots, \xi_n)$, $\sum_{i=1}^n \xi_i = A$, где ξ_i — часть сил, выделенных для атаки объекта C_i . Суммарные силы обороняющейся стороны (игрок 2) равны B . Чистой стратегией y игрока 2 является выбор набора неотрицательных чисел $y = (\eta_1, \dots, \eta_n)$, удовлетворяющих условию $\sum_{i=1}^n \eta_i = B$, где η_i — часть сил, предназначенных для защиты объекта C_i . Результат атаки на объект C_i пропорционален разности $\xi_i - \eta_i$, если силы атакующих превосходят силы защищающихся, а в остальных случаях он равен нулю. Построить функцию выигрыша.

2. Игра на единичном квадрате имеет функцию выигрыша

$$H(x, y) = xy - \frac{1}{3}x - \frac{1}{2}y.$$

Показать, что $(1/2, 1/3)$ — ситуация равновесия в этой игре.

3. Показать, что игра на единичном квадрате с функцией выигрыша

$$H(x, y) = \text{sign}(x - y)$$

имеет седловую точку.

4. Показать, что игра на единичном квадрате типа дуэли с функцией выигрыша

$$H(x, y) = \begin{cases} -1/x^2, & x > y, \\ 0, & x = y, \\ 1/y^2, & x < y \end{cases}$$

имеет ситуацию равновесия $(0, 0)$.

5. Показать, что игра на единичном квадрате с функцией выигрыша $H(x, y) = (x - y)^2$ не имеет ситуаций равновесия в чистых стратегиях.

6. Показать, что в игре на единичном квадрате с функцией выигрыша

$$H(x, y) = \begin{cases} x + y, & x \neq 1, y \neq 0, \\ 1/2 + y, & x = 1, y \neq 0, \\ 1/2 + x, & x \neq 1, y = 0, \\ 2, & x = 1, y = 0 \end{cases}$$

пара $(x_\varepsilon, y_\varepsilon)$, где $x_\varepsilon = 1 - \varepsilon$, $y_\varepsilon = \varepsilon$, является ситуацией ε -равновесия. Имеет ли эта игра значение?

7. Решить игру «поиска шумного объекта», сформулированную в примере 6 п. 2.1.2.

8. Вычислить выигрыш игрока 1 в игре на единичном квадрате с функцией выигрыша $H(x, y)$ в ситуации $(F(x), G(y))$ (F и G — функции распределения), если:

(а) $H(x, y) = (x + y)/(4xy)$, $F(x) = x^2$, $G(y) = y^2$;

(б) $H(x, y) = |x - y|(1 - |x - y|)$, $F(x) = x$, $G(y) = y$;

(с) $H(x, y) = (x - y)^2$, $F(x) = 1/2I_0(x) + 1/2I_1(x)$, $G(y) = I_{1/2}(x)$, где $I_k(x)$ — ступенчатая функция.

9. Игра дискретного поиска. Рассматривается следующая бесконечная игра. Стратегия игрока 2 заключается в выборе точки, равномерно распределенной на окружности радиуса y , где y может принимать значения из интервала $[0, 1]$. Игрок 1 может просмотреть в единичном круге односвязную область $a(Q) = a = \text{const}$, где $a < A$, $A = \pi$ — площадь единичного круга. Его стратегия x заключается в выборе формы области Q , имеющей площадь a , которая целиком лежит в единичном круге. Выигрыш $H(x, y)$ игрока 1 равен вероятности обнаружения, т. е. $H(x, y) = \text{Pr}(y \in Q)$. Под смешанной стратегией $g(y)$ игрока 2 будем понимать функцию плотности распределения случайной величины $y \in [0, 1]$. Найти решение игры.

10. Доказать теорему Хелли п. 2.5.4.

11. Рассмотрим непрерывный аналог игры «обороны города» (п. 1.1.3 гл. 1). Игрок 1 должен направить силы x , $x \in [0, 1]$ в наступление на первую позицию и силы $1 - x$ — в наступление на вторую позицию. Игрок 2 должен направить силы y , $y \in [0, 1]$ для обороны первой позиции и силы $1 - y$ — для обороны второй, на которой уже расположены постоянные оборонительные силы размером $1/2$. Один игрок платит другому единицу на каждой позиции, если его силы на этой позиции меньше сил противника, и ничего не платит, если их силы равны.

Построить функцию выигрыша $H(x, y)$ для игры на единичном квадрате. Показать, что данная игра не имеет решения в смешанных стратегиях. Указание. Воспользоваться результатом примера 10 п. 2.4.12.

12. Показать, что в непрерывной игре с функцией выигрыша

$$H(x, y) = [1 + (x + y)^2]^{-1}$$

стратегии $F^*(x) = I_{1/2}(x)$, $G^*(y) = 1/2I_0(y) + 1/2I_1(y)$ оптимальны для игроков 1 и 2 соответственно.

13. Доказать, что значение симметричной непрерывной игры на единичном квадрате равно нулю, а оптимальные смешанные стратегии совпадают (игра симметричная), если функция выигрыша кососимметрична, т. е. $H(x, y) = -H(y, x)$.

14. Определить оптимальные стратегии и значение игры на единичном квадрате с функцией выигрыша $H(x, y) = y^3 - 3xy + x^3$.

15. Показать, что в игре с функцией выигрыша

$$H(x, y) = e^{\gamma y - x} \sqrt{1 - x^2/y^2}, \quad x \in [x_0, x_1], \quad y \in [y_0, y_1], \quad \gamma > 0$$

игрок 2 имеет оптимальную чистую стратегию. Выяснить вид этой стратегии в зависимости от параметра. Что можно сказать об оптимальной стратегии игрока 1?

16. Проверить, что функция выигрыша из примера 11 п. 2.5.5, $H(x, y) = \rho(x, y)$, $x \in S(0, l)$, $y \in S(0, l)$, где $S(0, l)$ — круг с центром в 0 и радиусом l , $\rho(\cdot)$ — расстояние в R^2 , строго выпукла по y при любом фиксированном x .

17. Показать, что сумма двух выпуклых функций выпукла.

18. Доказать, что если выпуклая функция $\varphi : [\alpha, \beta] \rightarrow R^1$ ограничена, то она непрерывна в любой точке $x \in (\alpha, \beta)$. Вместе с тем на концах α и β промежутка (α, β) , выпуклая функция φ полунепрерывна сверху, т. е.

$$\overline{\lim}_{x \rightarrow \alpha} \varphi(x) \leq \varphi(\alpha)$$

(аналогично при $x \rightarrow \beta$).

19. Пусть дана игра $\Gamma = (X, Y, H)$, $X = Y = [0, 1]$ с выпуклой ограниченной функцией выигрыша $H(x, \cdot) : [0, 1] \rightarrow R^1$. Показать, что игрок 2 в этой игре имеет либо оптимальную чистую стратегию, либо для каждого $\varepsilon > 0$ чистую ε -оптимальную стратегию. Относительно игрока 1 справедлив результат теоремы п. 2.5.6. Указание. Использовать результат упр. 18 и рассмотреть вспомогательную игру $\Gamma_0 = (X, Y, H_0)$, где

$$H_0(x, y) = \begin{cases} H(x, y), & \text{если } y \in (0, 1), \\ \lim_{y_n \rightarrow y} H(x, y_n), & \text{если } y = 0 \text{ или } y = 1. \end{cases}$$

20. Решить игру «нападение — защита», сформулированную в упр. 1.

21. Рассматривается одновременная игра преследования на плоскости (см. пример 1 п. 2.1.2), когда множество стратегий $S_1 = S_2 = S$, где S — некоторое замкнутое выпуклое ограниченное множество.

(а) Показать, что значение рассматриваемой игры равно R , где R — радиус минимального круга $S(O, R)$, содержащего S , а оптимальная стратегия игрока 2 является чистой и заключается в выборе центра O круга $S(O, R)$.

(б) Показать, что оптимальная стратегия игрока 1 является смешанной и является смесью либо двух диаметрально противоположных точек касания множества S с кругом $S(O, R)$ (если такие точки x_1 и x_2 существуют), либо таких трех точек касания x'_1, x'_2, x'_3 , что точка O лежит внутри треугольника, вершинами которого являются данные точки.

22. Решить одновременную игру преследования на плоскости, рассмотренную в упр. 21, в предположении, что игрок 2 выбирает не одну точку $y \in S$, а m точек $y_1, \dots, y_m \in S$. Функция выигрыша игры имеет вид

$$H(x, y) = \frac{1}{m} \sum_{i=1}^m \rho^2(x, y_i),$$

где $\rho(\cdot)$ — расстояние в R^2 .

23. Игрок 1 выбирает системы x из m точек промежутка $[-1, 1]$, т. е. $x = (\xi_1, \dots, \xi_m)$, $\xi_i \in [-1, 1]$, $i = 1, \dots, m$. Одновременно и независимо от него игрок 2 выбирает систему y из n точек того же промежутка $[-1, 1]$, т. е. $y = (\eta_1, \dots, \eta_n)$, $\eta_j \in [-1, 1]$, $j = 1, 2, \dots, n$. Функция выигрыша $H(x, y)$ имеет вид

$$H(x, y) = \frac{1}{2} (\max_i \min_j |\xi_i - \eta_j| + \max_j \min_i |\xi_i - \eta_j|).$$

Найти решение игры.

24. Рассмотреть обобщение задачи п. 2.8.3, а именно игру поиска, в которой игрок 2 выбирает систему k точек $y = (y_1, \dots, y_k)$ на сфере C , а игрок 1, как и прежде, — систему x из s точек $x = (x_1, \dots, x_s)$ на сфере C . Функция выигрыша имеет вид

$$H(x, y) = \{M \mid M = |\{y_i\}| : y_i \in S(x_j, r); j = 1, \dots, s\},$$

где $S(x_j, r)$ — сферический сегмент с вершиной в точке x_j и радиусом основания r ; запись $|\{y_i\}|$ означает количество точек множества $\{y_i\}$. Точка y_i считается обнаруженной, если $y_i \in S(x_j, r)$ хотя бы для одного x_j . Таким образом, значение функции выигрыша имеет смысл числа обнаруженных точек в ситуации (x, y) .

Найти решение игры.

Глава 3

Неантагонистические игры

§ 3.1. Определение бескоалиционной игры в нормальной форме

3.1.1. В предыдущих главах были рассмотрены антагонистические игры двух лиц, т. е. игры, в которых интересы сторон прямо противоположны. Однако реальные задачи принятия решения в условиях конфликта характеризуются большим числом участников и, как следствие этого, неантагонистичностью конфликтной ситуации. Если говорить о конфликте двух лиц и его моделях, то можно заметить, что он также не исчерпывается только антагонистическим случаем. Дело в том, что интересы игроков могут пересекаться, но не быть обязательно противоположными. Это, в частности, может приводить к ситуациям, взаимовыгодным обоим игрокам (в антагонистическом конфликте это невозможно), что делает осмысленным кооперирование (выбор согласованного решения), приводящее к увеличению выигрыша обоих игроков. Однако возможны такие конфликты, когда кооперация или соглашение невозможны по правилам игры. Поэтому в неантагонистических играх различают бескоалиционное поведение, когда соглашения между игроками запрещены правилами (см. § 3.1 – 3.8), и кооперативное поведение игроков, когда разрешается кооперация типа выбора совместных стратегий (см. § 3.9 – 3.10) и совершения побочных платежей (см. § 3.11 – 3.14). Рассмотрим сначала бескоалиционное поведение.

3.1.2. Определение. Система

$$\Gamma = (N, \{X_i\}_{i \in N}, \{H_i\}_{i \in N}),$$

в которой $N = \{1, 2, \dots, n\}$ — множество игроков, X_i — множество стратегий игрока i , H_i — функция выигрыша игрока i , определенная на декартовом произведении множеств стратегий игроков $X = \prod_{i=1}^n X_i$ (множество ситуаций игры), называется *бескоалиционной игрой*.

Бескоалиционная игра n лиц происходит следующим образом. Игроки одновременно и независимо друг от друга выбирают свои стратегии x_i из множеств стратегий X_i , $i = 1, 2, \dots, n$, в результате чего формируется ситуация $x = (x_1, \dots, x_n)$, $x_i \in X_i$. После этого каждый игрок i получает выигрыш $H_i(x) = H_i(x_1, \dots, x_n)$ и игра заканчивается.

Если множества чистых стратегий игроков X_i конечны, то игра называется *конечной бескоалиционной игрой* n лиц.

3.1.3. Бескоалиционная игра Γ , в которой принимают участие два игрока, называется *игрой двух лиц*. Таким образом, бескоалиционная игра двух лиц Γ в нормальной

форме определяется системой $\Gamma = (X_1, X_2, H_1, H_2)$, где X_1 — множество стратегий первого игрока, X_2 — множество стратегий второго игрока, $X_1 \times X_2$ — множество ситуаций игры, а $H_1 : X_1 \times X_2 \rightarrow R^1$, $H_2 : X_1 \times X_2 \rightarrow R^1$ — функции выигрыша соответственно 1 и 2 игроков. Конечная бескоалиционная игра двух лиц называется *биматричной*. Это объясняется тем, что, перенумеровав множества чистых стратегий игроков числами $1, 2, \dots, \bar{m}$ и $1, 2, \dots, \bar{n}$ соответственно, функции выигрыша можно записать в виде двух матриц

$$H_1 = A = \begin{bmatrix} \alpha_{11} & \dots & \alpha_{1n} \\ \dots & \dots & \dots \\ \alpha_{m1} & \dots & \alpha_{mn} \end{bmatrix} \quad \text{и} \quad H_2 = B = \begin{bmatrix} \beta_{11} & \dots & \beta_{1n} \\ \dots & \dots & \dots \\ \beta_{m1} & \dots & \beta_{mn} \end{bmatrix}.$$

При этом элементы α_{ij} и β_{ij} матриц A, B являются соответственно выигрышами игроков 1 и 2 в ситуации (i, j) , $i \in \bar{M}$, $j \in \bar{N}$, $\bar{M} = \{1, \dots, \bar{m}\}$, $\bar{N} = \{1, \dots, \bar{n}\}$.

В соответствии с изложенным выше биматричная игра происходит следующим образом. Первый игрок выбирает номер i строки, а второй (одновременно и независимо) номер j столбца матрицы. Тогда игрок 1 получает выигрыш $\alpha_{ij} = H_1(x_i, y_j)$, а игрок 2 — выигрыш $\beta_{ij} = H_2(x_i, y_j)$.

Заметим, что биматричную игру с матрицами A и B можно также задать $(\bar{m} \times \bar{n})$ матрицей (A, B) , каждый элемент которой есть пара $(\alpha_{ij}, \beta_{ij})$, $i = 1, 2, \dots, \bar{m}$, $j = 1, 2, \dots, \bar{n}$. Игру, определяемую матрицами A и B , будем обозначать $\Gamma(A, B)$.

Если бескоалиционная игра Γ двух лиц такова, что $H_1(x, y) = -H_2(x, y)$ для всех $x \in X_1$, $y \in X_2$, то Γ оказывается антагонистической игрой, рассмотренной в предыдущих главах. В частном случае, когда в биматричной игре $\alpha_{ij} = -\beta_{ij}$, мы получаем матричную игру, рассмотренную в гл. 1.

3.1.4. Пример 1 («Семейный спор»). Рассматривается биматричная игра с матрицей

$$(A, B) = \begin{matrix} & \begin{matrix} \beta_1 & \beta_2 \end{matrix} \\ \begin{matrix} \alpha_1 \\ \alpha_2 \end{matrix} & \begin{bmatrix} (4, 1) & (0, 0) \\ (0, 0) & (1, 4) \end{bmatrix} \end{matrix}.$$

Имеются различные интерпретации этой игры, но наиболее известная [Льюис и Райфа, 1961] следующая. Муж (игрок 1) и жена (игрок 2) могут выбрать одно из двух вечерних развлечений: футбольный матч (α_1, β_1) или театр (α_2, β_2) . Если они имеют разные желания (α_1, β_2) или (α_2, β_1) , то остаются дома. Муж предпочитает футбольный матч, а жена — театр. Однако обоим гораздо важнее провести вечер вместе, чем участвовать в развлечении (хотя и предпочтительном) одному.

Пример 2 (Игра «перекресток») [Мулен, 1985]. Два автомобилиста двигаются по двум взаимно перпендикулярным дорогам и одновременно встречаются на перекрестке. Каждый из них может остановиться (1-я стратегия α_1 или β_1) и ехать (2-я стратегия α_2 или β_2).

Предполагается, что каждый из игроков предпочитает остановиться, а не пострадать в аварии и проехать, если другой сделал остановку. Этот конфликт может быть формализован биматричной игрой с матрицей

$$(A, B) = \begin{matrix} & \begin{matrix} \beta_1 & \beta_2 \end{matrix} \\ \begin{matrix} \alpha_1 \\ \alpha_2 \end{matrix} & \begin{bmatrix} (1, 1) & (1 - \varepsilon, 2) \\ (2, 1 - \varepsilon) & (0, 0) \end{bmatrix} \end{matrix}.$$

Здесь неотрицательное число ε соответствует неудовольствию от того, что игрок остановился и пропустил партнера.

Пример 3 (Выбор способа передвижения по городу) [Мулен, 1985]. Пусть число игроков n велико и каждое из множеств X_i состоит из двух элементов: $X_i = \{0, 1\}$ (для определенности: 0 — воспользоваться автомобилем, 1 — использовать общественный транспорт). Функция выигрыша определяется следующим образом:

$$H_i(x_1, \dots, x_n) = \begin{cases} a(t), & \text{при } x_i = 1, \\ b(t), & \text{при } x_i = 0, \end{cases}$$

где $t = \frac{1}{n} \sum_{j=1}^n x_j$.

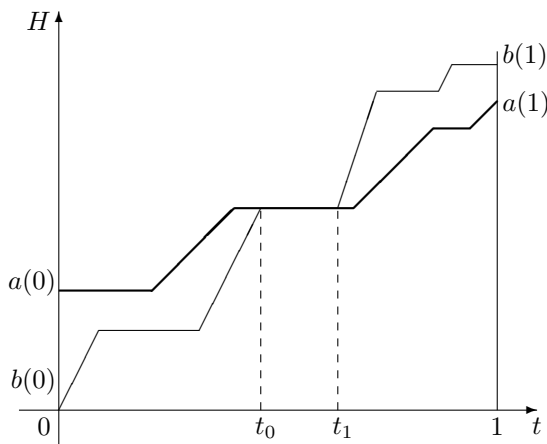


Рис. 3.1. Функции $a(t)$, $b(t)$ (выбор способа передвижения по городу)

Пусть a и b имеют вид, изображенный на рис. 3.1. Из вида функций $a(t)$ и $b(t)$ следует, что если доля игроков, выбирающих 1, больше t_1 , то уличное движение настолько свободно, что водитель чувствует себя лучше, чем пассажир в общественном транспорте. Если же доля автомобилистов больше $1 - t_0$, то движение настолько интенсивное (при естественном приоритете общественного транспорта), что сравнение теперь в пользу пассажиров общественного транспорта.

Пример 4 (Распределение ограниченного ресурса с учетом интересов потребителей). Предположим, что n потребителей имеют возможность расходовать (накапливать) некоторый ресурс, объем которого ограничен величиной $A > 0$. Обозначим объем ресурса, который расходует (накапливает) i -й потребитель, через x_i . В зависимости от значений вектора $x = (x_1, x_2, \dots, x_n)$ потребители получают выигрыш, который оценивается для i -го потребителя функцией $h_i(x_1, x_2, \dots, x_n)$, если общий объем израсходованного (накопленного) ресурса не превосходит заданной положительной величины $\Theta < A$, т. е.

$$\sum_{i=1}^n x_i \leq \Theta, \quad x_i \geq 0.$$

Если выполняется противоположное неравенство, то выигрыш i -го потребителя вычисляется с помощью функции $g_i(x_1, x_2, \dots, x_n)$. При этом предполагается, что полезность ресурса резко снижается, если $\sum_{i=1}^n x_i > \Theta$, т. е.

$$g_i(x_1, x_2, \dots, x_n) < h_i(x_1, x_2, \dots, x_n).$$

Рассмотрим неантагонистическую игру в нормальной форме

$$\Gamma = (N, \{X_i\}_{i \in N}, \{H_i\}_{i \in N}),$$

в которой функции выигрыша игроков имеют вид

$$H_i(x_1, x_2, \dots, x_n) = \begin{cases} h_i(x_1, \dots, x_n), & \sum_{i=1}^n x_i \leq \Theta \\ g_i(x_1, \dots, x_n), & \sum_{i=1}^n x_i > \Theta, \end{cases}$$

$$X_i = [0, a_i], \quad 0 \leq a_i \leq A, \quad \sum_{i=1}^n a_i = A, \quad N = \{1, 2, \dots, n\}.$$

Игроками в этой игре являются потребители ресурса.

Пример 5 (Теоретико-игровая модель охраны воздушного бассейна от загрязнений) [Петросян, Захаров, 1986]. В промышленном районе расположено n предприятий, каждое из которых имеет один источник, выбрасывающий в атмосферу вредную примесь. В районе имеется экологически значимая зона Ω , уровень загрязнения в которой не должен превышать предельно допустимого значения. Усредненное по времени и области значение концентрации вредной примеси в атмосфере при наличии n источников можно приближенно рассчитать по формуле

$$q = \sum_{i=1}^n c_i x_i, \quad 0 \leq x_i \leq a_i, \quad i = 1, 2, \dots, n.$$

Пусть $\Theta < \sum_{i=1}^n c_i a_i$ — значение предельно допустимой концентрации (ПДК) вредной примеси.

Считая предприятия игроками, построим игру, моделирующую конфликтную ситуацию загрязнения атмосферы. Предположим, что каждое предприятие i может снижать свои эксплуатационные расходы, увеличивая выброс x_i , однако если в зоне Ω уровень загрязнения превышает ПДК, на предприятие накладывается штраф $s_i > 0$.

Пусть игрок i (предприятие) имеет возможность выбирать значения x_i из множества $X_i = [0, a_i]$. Функции выигрыша игроков имеют вид

$$H_i(x_1, \dots, x_n) = \begin{cases} h_i(x_1, x_2, \dots, x_n), & q \leq \Theta, \\ h_i(x_1, x_2, \dots, x_n) - s_i, & q > \Theta, \end{cases}$$

где $h_i(x_1, x_2, \dots, x_n)$ — непрерывные и возрастающие по аргументу x_i функции.

Пример 6 (Аукцион делимого товара) [Зенкевич, 1994]. Два игрока участвуют в аукционе, на котором предлагаются q единиц товара с минимальной ценой p_0 . Предполагается, что игроки 1, 2 имеют бюджет M_1, M_2 соответственно. Они запрашивают количество товара q_1, q_2 соответственно (q_1, q_2 — целые числа) и предлагают цены p_1, p_2 за единицу товара, причем делают это независимо и одновременно.

При этом

$$q_1 + q_2 > q, \quad 0 < q_1 < q, \quad 0 < q_2 < q, \quad p_1 \in [p_0, \overline{p_1}], \quad p_2 \in [p_0, \overline{p_2}],$$

где $\overline{p_1} = M_1/(q-1)$, $\overline{p_2} = M_2/(q-1)$.

Согласно правилам аукциона, игрок, предложивший наибольшую цену, покупает запрашиваемое количество единиц товара по предложенной им цене. Другой игрок покупает остатки товара по предложенной им самим цене. Если предложенные игроками цены за единицу товара совпадают, то преимущество имеет игрок 1. Каждый игрок стремится максимизировать свой выигрыш.

Данный аукцион может быть описан как неантагонистическая игра двух лиц в нормальной форме $\Gamma = (X, Y, H_1, H_2)$, где множества стратегий определяются следующим образом:

$$X = \{p_1 | p_1 \in [p_0, \overline{p_1}]\}, \quad Y = \{p_2 | p_2 \in [p_0, \overline{p_2}]\},$$

а функциями выигрыша являются функции

$$H_1(p_1, p_2) = \begin{cases} (\overline{p_1} - p_1)q_1, & p_1 \geq p_2, \\ (\overline{p_1} - p_1)(q - q_2), & p_1 < p_2, \end{cases}$$

$$H_2(p_1, p_2) = \begin{cases} (\overline{p_2} - p_2)q_2, & p_1 < p_2, \\ (\overline{p_2} - p_2)(q - q_1), & p_1 \geq p_2. \end{cases}$$

§ 3.2. Принципы оптимальности в бескоалиционных играх

3.2.1. Известно, что для антагонистических игр принципы минимакса, максимина и равновесия совпадают (если они реализуемы, т. е. существует равновесие, а максимин и минимакс достигаются). В таком случае они определяют единое понятие оптимальности и решения игры. В теории неантагонистических игр нет единого подхода к выработке принципов оптимальности. По существу имеется целое множество таких принципов, каждый из которых основывается на некоторых дополнительных предположениях о поведении игроков и структуре игр.

Естественно предположить, что в игре Γ каждый из игроков стремится к достижению ситуации x , в которой значение его функции выигрыша было бы наибольшим. Однако функция выигрыша H_i зависит не только от стратегии i -го игрока, но и от стратегий, выбираемых другими игроками, поэтому ситуации $\{x^i\}$, дающие большее значение выигрыша для i -го игрока, могут не быть таковыми для других игроков. Таким образом, так же, как и в случае антагонистической игры, стремление игроков получить наибольший выигрыш носит конфликтный характер и сама формулировка того, какое поведение является «хорошим» или оптимальным в игре, является проблематичной. Здесь имеется несколько подходов. Одним из них является равновесие по Нэшу и его различные обобщения. В случае, когда игра Γ является антагонистической, равновесие по Нэшу совпадает с понятием равновесия, которое представляет собой основной принцип оптимальности в антагонистической игре.

Пусть $x = (x_1, \dots, x_{i-1}, x_i, x_{i+1}, \dots, x_n)$ — произвольная ситуация в игре Γ , а x_i — некоторая стратегия игрока i . Построим ситуацию, которая отлична от x только тем, что стратегия x_i игрока i заменена на стратегию x'_i . В результате мы получаем ситуацию $(x_1, \dots, x_{i-1}, x'_i, x_{i+1}, \dots, x_n)$, которую будем обозначать через $(x \| x'_i)$. Очевидно, что если x_i и x'_i совпадают, то $(x \| x'_i) = x$.

Определение. Ситуация $x^* = (x_1^*, \dots, x_i^*, \dots, x_n^*)$ называется ситуацией равновесия по Нэшу, если для всех $x_i \in X_i$ и $i = 1, \dots, n$ имеет место неравенство

$$H_i(x^*) \geq H_i(x^* \| x_i). \quad (3.2.1)$$

Пример 7. Рассмотрим игру примера 3 п. 3.1.4. Равновесными по Нэшу здесь являются ситуации, для которых выполняется условие

$$t_0 \leq t^* - 1/n, \quad t^* + 1/n \leq t_1, \quad (3.2.2)$$

где $t^* = \frac{1}{n} \sum_{j=1}^n x_j^*$. Из условия (3.2.2) следует, что переключение каждого отдельного игрока с одной чистой стратегии на другую при условии, что другие игроки своих стратегий не изменяют, не влияет на его выигрыш.

Пусть в игре реализовалась ситуация x , которой соответствует $t = \frac{1}{n} \sum_{j=1}^n x_j$, $t \in [t_0, t_1]$, и пусть величина δ — доля игроков, решивших переключиться со стратегии 0 на стратегию 1. Заметим, что если δ таково, что $b(t) = a(t) < a(t + \delta)$, то выигрыши этих игроков увеличиваются при таком переключении, если стратегии остальных игроков останутся прежними. Однако если это переключение действительно произойдет, то у тех же игроков возникает желание переключиться со стратегии 1 на стратегию 0, поскольку выполнено условие $a(t + \delta) < b(t + \delta)$. Если же это желание осуществится, то доля $\frac{1}{n} \sum_{j=1}^n x_j$ игроков уменьшится и вновь попадет на отрезок $[t_0, t_1]$.

Аналогично, пусть δ — доля игроков, переключившихся по каким-либо причинам (например, из-за случайных ошибок) со стратегии 1 на стратегию 0, причем $t - \delta < t_0$. Тогда в силу условия $b(t - \delta) < a(t - \delta)$, у игроков появится желание переключиться обратно на стратегию 1. При осуществлении этого желания доля $\frac{1}{n} \sum_{j=1}^n x_j$ увеличится и вновь вернется на отрезок $[t_0, t_1]$.

3.2.2. Из определения ситуации равновесия по Нэшу следует, что ни один из игроков i не заинтересован в отклонении от стратегии x_i^* , входящей в эту ситуацию (согласно (3.2.1) его выигрыш при использовании стратегии x_i вместо x_i^* разве лишь уменьшится при условии, что остальные игроки придерживаются стратегий, образующих ситуацию равновесия x^*). Таким образом, если игроки договорились предварительно об использовании стратегий, входящих в ситуацию равновесия x^* , то индивидуальное отклонение от договора невыгодно отклонившемуся игроку.

Определение. Стратегия $x_i^* \in X_i$ называется равновесной, если она входит хотя бы в одну ситуацию равновесия по Нэшу.

Для бескоалиционной игры двух лиц $\Gamma = (X_1, X_2, H_1, H_2)$ ситуация (x^*, y^*) является ситуацией равновесия, если неравенства

$$H_1(x, y^*) \leq H_1(x^*, y^*), \quad H_2(x^*, y) \leq H_2(x^*, y^*) \quad (3.2.3)$$

выполняются для всех $x \in X_1$ и $y \in X_2$

В частности, для биматричной $(m \times n)$ -игры $\Gamma(A, B)$ пара (i^*, j^*) будет ситуацией равновесия по Нэшу, если неравенства

$$\alpha_{ij^*} \leq \alpha_{i^*j^*}, \quad \beta_{i^*j} \leq \beta_{i^*j^*} \quad (3.2.4)$$

выполняются для всех номеров строк $i \in M$ и столбцов $j \in N$. Так, в примере 1 равновесными являются ситуации (α_1, β_1) и (α_2, β_2) , а в примере 2 — (α_1, β_2) и (α_2, β_1) .

Напомним, что для антагонистической игры $\Gamma = (X_1, X_2, H)$ пара $(x^*, y^*) \in X_1 \times X_2$ является ситуацией равновесия, если

$$H(x, y^*) \leq H(x^*, y^*) \leq H(x^*, y), \quad x \in X_1, y \in X_2.$$

При этом имеют место следующие основные свойства антагонистических игр.

1⁰. Игроку невыгодно информировать своего противника о стратегии (чистой или смешанной), которую он собирается применить. (Конечно, если игрок собирается использовать оптимальную стратегию, то его выигрыш не уменьшится от того, что он объявит об этом, но он ничего и не выигрывает.)

2⁰. Если $(x, y) \in Z(\Gamma)$, $(x', y') \in Z(\Gamma)$ — ситуации равновесия в игре Γ , а v — значение игры, то

$$(x', y) \in Z(\Gamma), \quad (x, y') \in Z(\Gamma), \quad (3.2.5)$$

$$v = H(x, y) = H(x', y') = H(x, y') = H(x', y). \quad (3.2.6)$$

3⁰. Игроки не заинтересованы в общении перед началом игры для выработки совместных действий.

4⁰. Если в игре Γ существует ситуация равновесия, а x и y — максиминная и минимаксная стратегии соответственно, то $(x, y) \in Z(\Gamma)$ — ситуация равновесия, и наоборот.

Выясним, выполняются ли эти свойства для биматричных игр.

Пример 8. Рассмотрим игру «семейный спор» (см. пример 1 и п. 3.1.4). Как уже отмечалось, в ней есть две равновесные ситуации (α_1, β_1) и (α_2, β_2) . Однако 1-я ситуация выгодна игроку 1, а 2-я — игроку 2. Это противоречит (3.2.6), поскольку выигрыши игроков в этих ситуациях различны. Далее заметим, что, несмотря на равновесность ситуаций (α_1, β_1) , (α_2, β_2) , пары (α_1, β_2) и (α_2, β_1) — не являются ситуациями равновесия по Нэшу, т. е. не выполнено свойство 2⁰ (см. (3.2.5)).

Если игрок 1 информирует партнера о намерении выбрать стратегию α_1 и если игрок 2 убежден, что тот будет упрямствовать, то ему ничего не остается, как объявить первую стратегию β_1 . Аналогичные рассуждения можно провести и за игрока 2. Таким образом, каждому из игроков выгодно первому объявить свою стратегию, что противоречит свойству 1⁰ для антагонистических игр.

Предположим, что игроки не общаются до начала игры, а делают выбор одновременно и независимо друг от друга (как и предусмотрено правилами бескоалиционной игры). Проведем рассуждения за игрока 1. Ему выгодно, чтобы реализовалась ситуация (α_1, β_1) , в то время как игроку 2 выгодна ситуация (α_2, β_2) . Поэтому, если игрок 1 выберет стратегию α_1 , то игрок 2 может выбрать стратегию β_2 и они оба проигрывают (вектор выигрышей $(0, 0)$). Тогда игроку 1 имеет смысл выбрать стратегию α_2 , поскольку в ситуации (α_2, β_2) он получает выигрыш 1. Но игрок 2 может рассуждать аналогично и выбрать β_1 , тогда в ситуации (α_2, β_1) они оба опять проигрывают.

Таким образом, имеет место случай, когда ситуация выгодна (и поэтому неустойчива) для игрока 1. Аналогично (с точки зрения игрока 2) можно исследовать ситуацию (α_2, β_2) . Следовательно, игрокам выгодно общаться перед началом игры и договариваться о совместном плане действий, что противоречит свойству 3⁰. Затруднения возникают также из-за того, что пара максиминных стратегий не является равновесной.

Таким образом, мы имеем пример игры, когда не выполнено ни одно из свойств 1⁰ — 4⁰ антагонистической игры.

Итак, в различных ситуациях равновесия по Нэшу векторы выигрышей игроков могут быть различны. Кроме того, множество ситуаций равновесия по Нэшу в отличие

от множества ситуаций равновесия в антагонистической игре не является прямоугольным. Если $x = (x_1, \dots, x_i, \dots, x_n)$ и $x' = (x'_1, \dots, x'_i, \dots, x'_n)$ — две различные ситуации равновесия, то ситуация x'' , состоящая из стратегий, которые образуют ситуации x и x' и не совпадающая ни с одной из этих ситуаций, равновесной может не являться. Ситуация равновесия по Нэшу является множественным принципом оптимальности в том смысле, что различные ситуации равновесия могут быть в разной степени предпочтительными для различных игроков. Таким образом, остается не решенным вопрос: какую из ситуаций равновесия можно принять как устраивающий всех игроков принцип оптимальности? В дальнейшем будет показано, что множественность принципа оптимальности является существенной характерной чертой оптимального поведения в конфликтных управляемых процессах со многими участниками.

Заметим также, что в отличие от антагонистического случая равновесная стратегия i -го игрока x_i^* далеко не всегда обеспечивает получение, по крайней мере, выигрыша $H_i(x^*)$ в ситуации равновесия по Нэшу, поскольку это существенно зависит от того, выберут ли остальные игроки стратегии, входящие в данную ситуацию равновесия по Нэшу. Поэтому равновесную стратегию не следует трактовать как оптимальную стратегию i -го игрока. Такая трактовка осмыслена только для набора стратегий игроков, т. е. для ситуаций.

3.2.3. Важная особенность ситуации равновесия по Нэшу заключается в том, что отклонение от нее двух игроков и более может привести к увеличению выигрыша одного из отклонившихся игроков. Пусть $S \subset N$ — некоторое подмножество множества игроков (коалиция) и пусть $x = (x_1, \dots, x_n)$ — ситуация в игре Γ . Обозначим через $(x \| x'_S)$ ситуацию, которая получается из ситуации x при замене в ней стратегий x_i , $i \in S$ на стратегии $x'_i \in X_i$, $i \in S$. Иными словами, в ситуации $(x \| x'_S)$ игроки, входящие в коалицию S , заменяют свои стратегии x_i на x'_i . Если x^* — ситуация равновесия по Нэшу, то из (3.2.1) вовсе не следует, что

$$H_i(x^*) \geq H_i(x^* \| x_S) \quad \text{для всех } i \in S. \quad (3.2.7)$$

Это будет показано далее на простых примерах.

Можно усилить понятие равновесия по Нэшу, потребовав выполнения условия (3.2.7) или ослабленного условия (3.2.7) хотя бы для одного из игроков $i \in S$. Тогда мы приходим к следующему определению.

Определение. Ситуация x^* называется *сильно равновесной*, если для любой коалиции $S \subset N$ и $x_S \in \prod_{i \in S} X_i$, существует игрок $i_0 \in S$, для которого выполняется строгое неравенство:

$$H_{i_0}(x^*) > H_{i_0}(x^* \| x_S). \quad (3.2.8)$$

Условие (3.2.8) гарантирует нецелесообразность соглашения между игроками с целью вступления в некоторую коалицию S , так как в любой коалиции находится игрок i_0 , которого это соглашение не устраивает. Любая сильно равновесная ситуация является равновесной.

Если бы сильное равновесие существовало в достаточно широком классе игр, то оно могло бы явиться приемлемым принципом оптимальности в бескоалиционной игре. Однако оно существует крайне редко.

Пример 9 («Дилемма заключенного»). Рассмотрим биматричную игру с матрицей

$$(A, B) = \begin{matrix} & \begin{matrix} \beta_1 & \beta_2 \end{matrix} \\ \begin{matrix} \alpha_1 \\ \alpha_2 \end{matrix} & \begin{bmatrix} (5, 5) & (0, 10) \\ (10, 0) & (1, 1) \end{bmatrix} \end{matrix}.$$

Здесь одна ситуация равновесия (α_2, β_2) (не сильно равновесная), которая дает игрокам вектор выигрышей $(1, 1)$. Однако если оба игрока сыграют (α_1, β_1) , то они получат вектор выигрышей $(5, 5)$, что выгодно обоим. Эта ситуация не является равновесной, но она лучшая для обоих игроков. Таких парадоксов в антагонистических играх не бывает. Если говорить об этом конкретном случае, то данный результат является следствием того, что при одновременном отклонении от равновесной стратегии каждый из игроков может выиграть еще больше.

3.2.4. Пример 8 приводит к мысли о возможности других принципов оптимальности в бескоалиционной игре, приводящих к ситуациям, более выгодным обоим участникам, чем в случае равновесных ситуаций. Таким принципом оптимальности является *оптимальность по Парето*.

Рассмотрим множество векторов $\{H(x)\} = \{H_1(x), \dots, H_n(x)\}$, $x \in X$, $X = \prod_{i=1}^n X_i$, т. е. множество значений вектор-выигрышей игроков во всех возможных ситуациях $x \in X$.

Определение. Ситуация \bar{x} в бескоалиционной игре Γ называется *оптимальной по Парето*, если не существует ситуации $x \in X$, для которой имеют место неравенства:

$$H_i(x) \geq H_i(\bar{x}) \text{ для всех } i \in N \text{ и}$$

$$H_{i_0}(x) > H_{i_0}(\bar{x}) \text{ хотя бы для одного } i_0 \in N.$$

Множество всех ситуаций, оптимальных по Парето, будем обозначать через X^P .

Содержательно принадлежность ситуации \bar{x} множеству X^P означает, что не существует другой ситуации x , которая была бы предпочтительнее ситуации \bar{x} для всех игроков.

Отметим содержательное различие понятий ситуации равновесия и ситуации, оптимальной по Парето. В первой ситуации ни один игрок, действуя в одиночку, не может увеличить своего выигрыша, во второй — все игроки, действуя совместно, не могут (даже не строго) увеличить выигрыш каждого.

Заметим также, что соглашение о выборе фиксированной ситуации равновесия удерживает каждого индивидуального игрока от отклонения от нее. В оптимальной по Парето ситуации отклонившийся игрок может в некоторых случаях получить существенно больший выигрыш. В то же время сильно равновесная ситуация безусловно является и оптимальной по Парето. Так, в примере 9 («Дилемма заключенного») ситуация (α_2, β_2) равновесна, но не оптимальна по Парето. Вместе с тем ситуация (α_1, β_1) , наоборот, оптимальна по Парето, но не является равновесной. В игре «семейный спор» обе равновесные ситуации (α_1, β_1) , (α_2, β_2) сильно равновесны и оптимальны по Парето, но, как уже отмечено в примере 8, не являются взаимозаменяемыми. Такая же картина имеет место и в следующем примере.

Пример 10. Рассмотрим игру «перекресток» (см. пример 2 п. 3.1.4). Ситуации (α_2, β_1) , (α_1, β_2) равновесны и оптимальны по Парето (ситуация (α_1, β_1) оптимальна по Парето, но не равновесна). Для каждого игрока равновесной является стратегия α_1, β_1 «остановиться», если другой игрок решил проехать перекресток, и, наоборот,

выгодно выбрать стратегию α_2, β_2 «ехать», если другой игрок остановился. Однако выигрыш в две единицы каждый из игроков получает только при выборе стратегии α_2, β_2 — «ехать», поэтому здесь неизбежна борьба за лидерство, т. е. каждый из игроков заинтересован первым заявить, что он выбрал стратегию «ехать».

Заметим, что точно к такому же выводу мы пришли при анализе игры «семейный спор» (см. пример 8).

3.2.5. Проанализируем поведение типа лидер — ведомый в игре двух лиц $\Gamma = (X_1, X_2, H_1, H_2)$. Обозначим как Z^1, Z^2 множества наилучших ответов игроков 1 и 2 соответственно, где

$$Z^1 = \{(x_1, x_2) | H_1(x_1, x_2) = \sup_{y_1} H_1(y_1, x_2)\}, \quad (3.2.9)$$

$$Z^2 = \{(x_1, x_2) | H_2(x_1, x_2) = \sup_{y_2} H_2(x_1, y_2)\} \quad (3.2.10)$$

(предполагается, что супремумы в (3.2.9) и (3.2.10) достигаются).

Определение. Назовём ситуацию $(x_1, x_2) \in X_1 \times X_2$ *i-равновесием по Штакельбергу* в игре двух лиц Γ , а \bar{H}_i — *i-выигрышем*, если $(x_1, x_2) \in Z^j$ и выполняется равенство

$$\bar{H}_i = H_i(x_1, x_2) = \sup_{(y_1, y_2) \in Z^j} H_i(y_1, y_2), \quad (3.2.11)$$

где $i = 1, 2, i \neq j$.

Понятие *i-равновесия* можно интерпретировать следующим образом. Игрок 1 (лидер) знает функции выигрыша обоих игроков H_1, H_2 , а тем самым и множество наилучших ответов Z^2 игрока 2 (ведомого) на любую стратегию x_1 игрока 1. Тогда он, обладая этой информацией, максимизирует свой выигрыш, выбирая стратегию x_1 из условия (3.2.11). Таким образом, \bar{H}_i — это выигрыш *i-го* игрока, действующего оптимально в качестве «лидера» в игре Γ .

Лемма. Пусть $Z(\Gamma)$ — множество ситуаций равновесия по Нэшу в игре двух лиц Γ . Тогда

$$Z(\Gamma) = Z^1 \cap Z^2, \quad (3.2.12)$$

где Z^1, Z^2 — множества наилучших ответов (3.2.9), (3.2.10) игроков 1, 2 в игре Γ .

Доказательство. Пусть $(x_1, x_2) \in Z(\Gamma)$ — ситуация равновесия по Нэшу. Тогда неравенства

$$H_1(x'_1, x_2) \leq H_1(x_1, x_2), \quad H_2(x_1, x'_2) \leq H_2(x_1, x_2)$$

выполняются для всех $x'_1 \in X_1$ и $x'_2 \in X_2$. Отсюда следует:

$$H_1(x_1, x_2) = \sup_{x'_1} H_1(x'_1, x_2), \quad (3.2.13)$$

$$H_2(x_1, x_2) = \sup_{x'_2} H_2(x_1, x'_2). \quad (3.2.14)$$

Таким образом, $(x_1, x_2) \in Z^1$ и $(x_1, x_2) \in Z^2$, т. е. $(x_1, x_2) \in Z^1 \cap Z^2$.

Обратное включение непосредственно следует из (3.2.13), (3.2.14). Лемма доказана.

Определение [Мулен, 1985]. Будем говорить, что в игре двух лиц $\Gamma = (X_1, X_2, H_1, H_2)$ имеет место борьба за лидерство, если не существует такой ситуации $(x_1, x_2) \in X_1 \times X_2$, что

$$\bar{H}_i \leq H_i(x_1, x_2), \quad i = 1, 2. \quad (3.2.15)$$

Теорема [Мулен, 1985]. Если игра двух лиц $\Gamma = (X_1, X_2, H_1, H_2)$ имеет по крайней мере две оптимальных по Парето и равновесных по Нэшу ситуации (x_1, x_2) , (y_1, y_2) с различными векторами выигрышей

$$(H_1(x_1, x_2), H_2(x_1, x_2)) \neq (H_1(y_1, y_2), H_2(y_1, y_2)), \quad (3.2.16)$$

то в игре Γ имеет место борьба за лидерство.

Доказательство. В силу (3.2.12) для всякой ситуации равновесия по Нэшу $(z_1, z_2) \in Z(\Gamma)$ справедливы неравенства

$$H_i(z_1, z_2) \leq \bar{H}_i, \quad i = 1, 2.$$

Предположим противное, т. е. что в игре Γ нет борьбы за лидерство. Тогда существует ситуация $(z_1, z_2) \in X_1 \times X_2$, для которой

$$H_i(x_1, x_2) \leq \bar{H}_i \leq H_i(z_1, z_2), \quad (3.2.17)$$

$$H_i(y_1, y_2) \leq \bar{H}_i \leq H_i(z_1, z_2), \quad (3.2.18)$$

$i = 1, 2$. Однако (x_1, x_2) , (y_1, y_2) — ситуации, оптимальные по Парето. Поэтому неравенства (3.2.17), (3.2.18) выполняются как равенства, что противоречит (3.2.16). Теорема доказана.

В заключение заметим, что игры «семейный спор» и «перекресток» (п. 3.1.4) удовлетворяют условиям теоремы п. 3.2.5, поэтому в них имеет место борьба за лидерство.

§ 3.3. Смешанное расширение бескоалиционной игры

3.3.1. Рассмотрим бескоалиционную игру двух лиц $\Gamma = (X_1, X_2, H_1, H_2)$. В антагонистическом случае мы уже убедились, что ситуация равновесия в обычных чистых стратегиях, вообще говоря, не существует. Даже матричные игры в общем случае имеют ситуацию равновесия лишь в смешанных стратегиях. Поэтому естественно искать равновесие по Нэшу в бескоалиционной игре в классе смешанных стратегий. Как и в случае антагонистических игр, смешанную стратегию игрока мы отождествляем с вероятностным распределением на множестве чистых стратегий. Предположим для простоты, что множества стратегий X_i конечны, и введем понятие смешанного расширения игры. Пусть

$$\Gamma = (N, \{X_i\}_{i \in N}, \{H_i\}_{i \in N}) \quad (3.3.1)$$

— произвольная конечная бескоалиционная игра. Для определенности предположим, что игрок i в игре Γ имеет m_i стратегий.

Обозначим через μ_i произвольную смешанную стратегию игрока i , т. е. некоторое вероятностное распределение на множестве стратегий X_i , которые назовем чистыми стратегиями. Через $\mu_i(x_i)$ будем обозначать вероятность, которую стратегия μ_i приписывает конкретной чистой стратегии $x_i \in X_i$. Множество всех смешанных стратегий игрока i будем обозначать через \bar{X}_i .

Пусть каждый из игроков $i \in N$ применяет свою смешанную стратегию μ_i , т. е. выбирает чистые стратегии с вероятностями $\mu_i(x_i)$. Будем предполагать, что вероятность появления ситуации $x = (x_1, \dots, x_n)$ равна произведению вероятностей выборов составляющих ее стратегий, т. е.

$$\mu(x) = \mu_1(x_1) \times \mu_2(x_2) \times \dots \times \mu_n(x_n). \quad (3.3.2)$$

Формула (3.3.2) определяет вероятностное распределение на множестве всех ситуаций $X = \prod_{i=1}^n X_i$, определяемое смешанными стратегиями $\mu_1, \mu_2, \dots, \mu_n$. Набор $\mu = (\mu_1, \dots, \mu_n)$ называется ситуацией в смешанных стратегиях. Ситуация в смешанных стратегиях μ реализует различные ситуации в чистых стратегиях с некоторыми вероятностями, поэтому значение функции выигрыша каждого из игроков оказывается случайной величиной. В качестве значения функции выигрыша 2-го игрока в ситуации μ принимается математическое ожидание этой случайной величины:

$$\begin{aligned} K_i(\mu) &= \sum_{x \in X} H_i(x) \mu(x) = \\ &= \sum_{x_1 \in X_1} \dots \sum_{x_n \in X_n} H_i(x_1, \dots, x_n) \times \mu_1(x_1) \times \dots \times \mu_n(x_n), \\ &\quad i \in N, \quad x = (x_1, \dots, x_n) \in X. \end{aligned} \quad (3.3.3)$$

Введем обозначение

$$K_i(\mu \| x'_j) = \sum_{x_1 \in X_1} \dots \sum_{x_{j-1} \in X_{j-1}} \sum_{x_{j+1} \in X_{j+1}} \dots \sum_{x_n \in X_n} H_i(x \| x'_j) \prod_{k \neq j} \mu_k(x_k). \quad (3.3.4)$$

Пусть μ'_j — произвольная смешанная стратегия игрока j в игре Γ . Умножая (3.3.4) на $\mu'_j(x'_j)$ и суммируя по всем $x'_j \in X_j$, получаем

$$\sum_{x'_j \in X_j} K_i(\mu \| x'_j) \mu'_j(x'_j) = K_i(\mu \| \mu'_j).$$

Определение. Игра $\bar{\Gamma} = (N, \{\bar{X}_i\}_{i \in N}, \{K_i\}_{i \in N})$, в которой N — множество игроков, \bar{X}_i — множество смешанных стратегий каждого игрока i , а функция выигрыша определяется равенством (3.3.3), называется смешанным расширением игры Γ .

Если неравенство $K_j(\mu \| x_i) \leq a$ выполняется для любой чистой стратегии x_i игрока i , то для любой смешанной стратегии μ_i^* справедливо неравенство $K_j(\mu \| \mu_i^*) \leq a$. Доказательство этого факта вытекает из (3.3.3) и (3.3.4) стандартным переходом к смешанным стратегиям.

3.3.2. Для биматричной $(m \times n)$ -игры $\Gamma(A, B)$ можно определить множества смешанных стратегий X_1, X_2 1 и 2 игрока, соответственно, в виде

$$X_1 = \{x \mid xu = 1, x \geq 0, x \in R^m\},$$

$$X_2 = \{y \mid yw = 1, y \geq 0, y \in R^n\},$$

где $u = (1, \dots, 1) \in R^m$, $w = (1, \dots, 1) \in R^n$. Также определим выигрыши игроков K_1 и K_2 в смешанных стратегиях в ситуации (x, y) как математическое ожидание выигрыша

$$K_1(x, y) = xAy, \quad K_2(x, y) = xBy, \quad x \in X_1, \quad y \in X_2.$$

Следовательно, формально построено смешанное расширение $\bar{\Gamma}(A, B)$ игры $\Gamma(A, B)$, т. е. бескоалиционная игра двух лиц $\bar{\Gamma}(A, B) = (X_1, X_2, K_1, K_2)$.

Для биматричной игры (как и для матричной) множество $M_x = \{i \mid \xi_i > 0\}$ будем называть спектром смешанной стратегии $x = (\xi_1, \dots, \xi_m)$ игрока 1, а стратегию x , для которой $M_x = M$, $M = \{1, 2, \dots, m\}$, вполне смешанной. Аналогично, $N_y = \{j \mid \eta_j > 0\}$

будем называть спектром смешанной стратегии $y = \{\eta_1, \dots, \eta_n\}$ в биматричной $(m \times n)$ -игре $\Gamma(A, B)$. Ситуацию (x, y) , в которой обе стратегии x и y являются вполне смешанными, будем называть *вполне смешанной*.

Покажем на примере игры «семейный спор», что введение смешанных стратегий не снимает те трудности, которые возникают при анализе бескоалиционной игры (см. пример 8 п. 3.2.2).

Пример 11. Пусть в игре «семейный спор» игрок 1 хочет максимально увеличить свой гарантированный выигрыш. Это означает, что он намерен выбрать смешанную стратегию $x^0 = (\xi^0, 1 - \xi^0)$, $0 \leq \xi^0 \leq 1$ так, чтобы максимально увеличить наименьшую из двух величин $K_1(x, \beta_1)$ и $K_1(x, \beta_2)$, т. е.

$$\max_x \min\{K_1(x, \beta_1), K_1(x, \beta_2)\} = \min\{K_1(x^0, \beta_1), K_1(x^0, \beta_2)\}.$$

Максиминная стратегия x^0 игрока 1 имеет вид $x^0 = (1/5, 4/5)$ и дает ему средний гарантированный выигрыш $4/5$. Если игрок 2 выберет стратегию β_1 , то выигрыши игроков будут равны $(4/5, 1/5)$, если же он воспользуется стратегией β_2 , то $(4/5, 16/5)$.

Таким образом, если игрок 2 догадается, что его партнер придерживается стратегии x^0 , то он выберет β_2 и получит выигрыш $16/5$. (Если игрок 1 может обосновать выбор β_2 за игрока 2, то он может улучшить и свой выбор.) Аналогично, пусть игрок 2 придерживается максиминной стратегии (она имеет вид $y^0 = (4/5, 1/5)$), и если игрок 1 выбирает стратегию α_1 , то выигрыши игроков равны $(16/5, 4/5)$. Если же игрок 1 выберет α_2 , то выигрыши игроков равны $(1/5, 4/5)$. Следовательно, против максиминной стратегии y^0 ему выгодно применять стратегию α_1 .

Если оба игрока будут рассуждать таким образом, то они приходят к ситуации (α_1, β_2) , в которой вектор выигрышей $(0, 0)$. Здесь ситуация (x^0, y^0) в максиминных смешанных стратегиях не является ситуацией равновесия по Нэшу.

3.3.3. Определение. Ситуация μ^* называется ситуацией равновесия по Нэшу в смешанных стратегиях в игре Γ , если для любого игрока i и для любой его смешанной стратегии μ_i выполняется следующее неравенство:

$$K_i(\mu^* \parallel \mu_i) \leq K_i(\mu^*), \quad i = 1, \dots, n.$$

Как показывает пример 11, ситуация в максиминных смешанных стратегиях не обязательно является ситуацией равновесия по Нэшу в смешанных стратегиях.

Пример 12. В игре «перекресток» (см. пример 10 п.3.2.4) имеются две ситуации равновесия по Нэшу в чистых стратегиях: (α_1, β_2) и (α_2, β_1) . Эти же ситуации оптимальны по Парето. В смешанном расширении игры возникает еще одна ситуация равновесия, а именно пара (x^*, y^*) :

$$x^* = y^* = \frac{1-\varepsilon}{2-\varepsilon}u_1 + \frac{1}{2-\varepsilon}u_2,$$

где $u_1 = (1, 0)$, $u_2 = (0, 1)$ или $x^* = y^* = ((1-\varepsilon)/(2-\varepsilon), 1/(2-\varepsilon))$.

Действительно, имеем

$$K_1(\alpha_1, y^*) = \frac{1-\varepsilon}{2-\varepsilon} + \frac{1-\varepsilon}{2-\varepsilon} = 1 - \frac{\varepsilon}{2-\varepsilon},$$

$$K_1(\alpha_2, y^*) = 2\frac{1-\varepsilon}{2-\varepsilon} = 1 - \frac{\varepsilon}{2-\varepsilon}.$$

Более того, так как для любых смешанных стратегий $x = (\xi, 1 - \xi)$ и $y = (\eta, 1 - \eta)$ выполняются равенства

$$K_1(x, y^*) = \xi K_1(\alpha_1, y^*) + (1 - \xi) K_1(\alpha_2, y^*) = 1 - \frac{\varepsilon}{2 - \varepsilon},$$

$$K_2(x^*, y) = \eta K_2(x^*, \beta_1) + (1 - \eta) K_2(x^*, \beta_2) = 1 - \frac{\varepsilon}{2 - \varepsilon},$$

то получаем

$$K_1(x, y^*) = K_1(x^*, y^*), \quad K_2(x^*, y) = K_2(x^*, y^*)$$

для всех смешанных стратегий $x \in X_1$ и $y \in X_2$. Следовательно, (x^*, y^*) — ситуация равновесия по Нэшу. Более того, это вполне смешанная ситуация равновесия. Однако ситуация (x^*, y^*) не является оптимальной по Парето, так как вектор $K(x^*, y^*) = (1 - \varepsilon/(2 - \varepsilon), 1 - \varepsilon/(2 - \varepsilon))$ строго меньше (покомпонентно) вектора выигрышей $(1, 1)$ в ситуации (α_1, β_1) .

Пусть $K(\mu^*) = \{K_i(\mu^*)\}$ — вектор выигрышей в некоторой ситуации равновесия по Нэшу. Обозначим $v_i = K_i(\mu^*)$ и $v = \{v_i\}$. Заметим, что если в антагонистических играх значение v функции выигрыша в ситуации равновесия было одним и тем же для всех ситуаций равновесия, а следовательно, осуществлялось единственным образом для каждой антагонистической игры, в которой существовала ситуация равновесия, то в неантагонистических играх вектор v определяется неоднозначно. Таким образом, здесь можно говорить лишь о равновесном выигрыше $v_i = K_i(\mu^*)$ игрока i в ситуации равновесия μ^* , $\mu^* \in \bar{X}$, $\bar{X} = \prod_{i=1}^n \bar{X}_i$.

Так, в игре «перекресток» в ситуации равновесия (α_1, β_2) вектор равновесных выигрышей (v_1, v_2) имеет вид $(1 - \varepsilon, 2)$, а в ситуации (x^*, y^*) он равен $(1 - \varepsilon/(2 - \varepsilon), 1 - \varepsilon/(2 - \varepsilon))$ (см. пример 12).

3.3.4. Если в бескоалиционной игре $\Gamma = (X_1, X_2, H_1, H_2)$ пространства стратегий бесконечны, например, $X_1 \subset R^m$, $X_2 \subset R^n$, то, как и в случае бесконечных антагонистических игр, смешанные стратегии игроков отождествляются с вероятностными мерами, заданными на борелевских σ -алгебрах множеств X_1 и X_2 . Если μ и ν — смешанные стратегии игроков 1 и 2 соответственно, то выигрышем игрока i в этой ситуации является $K_i(\mu, \nu)$ — математическое ожидание выигрыша, т. е.

$$K_i(\mu, \nu) = \int_{X_1} \int_{X_2} H_i(x, y) d\mu(x) d\nu(y), \quad (3.3.5)$$

где интегралы понимаются в смысле Лебега–Стилтьеса. Заметим, что в ситуациях (x, y) и (μ, y) выигрыши игроков имеют вид

$$K_i(x, \nu) = \int_{X_2} H_i(x, y) d\nu(y),$$

$$K_i(\mu, y) = \int_{X_1} H_i(x, y) d\mu(x), \quad i = 1, 2.$$

(Предполагается, что интегралы существуют.)

Таким образом, формально смешанное расширение бескоалиционной игры Γ двух лиц может быть задано системой $\bar{\Gamma} = (\bar{X}_1, \bar{X}_2, K_1, K_2)$, где $\bar{X}_1 = \{\mu\}$, $\bar{X}_2 = \{\nu\}$, а K_1 и K_2 определяются (3.3.5). Игра $\bar{\Gamma}$ является бескоалиционной игрой двух лиц, поэтому ситуация (μ^*, ν^*) равновесна тогда и только тогда, когда выполнены неравенства, аналогичные (3.2.3).

§ 3.4. Существование ситуации равновесия по Нэшу

3.4.1. В теории антагонистических игр для существования ситуации равновесия в смешанных стратегиях было достаточно непрерывности функции выигрыша и компактности множеств стратегий (см. п. 2.4.4). Оказывается, что этих условий достаточно и для существования ситуации равновесия по Нэшу в смешанных стратегиях для бескоалиционной игры двух лиц. Вместе с тем вопрос о существовании ситуации равновесия в бескоалиционной игре двух лиц является правомерным. Уже приводился пример антагонистической игры, которая не имеет ситуации равновесия в смешанных стратегиях (см. п. 2.4.12).

Сначала докажем существование ситуации равновесия в смешанных стратегиях для биматричной игры. Это доказательство опирается на известную теорему Какутани о неподвижной точке, которую приведем без доказательства в следующей формулировке (см. также п. 3.5.5).

Теорема. Пусть S — компактное выпуклое множество в R^n и ψ — многозначное отображение, переводящее точки S в компактные выпуклые подмножества S и удовлетворяющее условию: если $x_n \in S$, $x_n \rightarrow x$, $y_n = \psi(x_n)$, $y_n \rightarrow y$, то $y \in \psi(x)$. Тогда существует такое $x^* \in S$, что $x^* \in \psi(x^*)$.

Теорема. Пусть $\Gamma(A, B)$ — биматричная $(m \times n)$ -игра. Тогда существуют смешанные стратегии $x^* \in X_1$ и $y^* \in X_2$ игроков 1 и 2 соответственно, такие, что пара (x^*, y^*) является ситуацией равновесия по Нэшу.

Доказательство. Множества смешанных стратегий X_1 и X_2 игроков 1 и 2 — выпуклые многогранники, поэтому множество ситуаций $X_1 \times X_2$ — компактное выпуклое множество.

Пусть ψ — многозначное отображение,

$$\psi : X_1 \times X_2 \rightarrow X_1 \times X_2,$$

определяемое соотношением

$$\psi : (x_0, y_0) \rightarrow \left\{ (x', y') \left| \begin{array}{l} K_1(x', y_0) = \max_{x \in X_1} K_1(x, y_0), \\ K_2(x_0, y') = \max_{y \in X_2} K_2(x_0, y), \end{array} \right. \right.$$

т. е. образ отображения ψ состоит из пар наилучших ответов игроков на стратегии y_0 и x_0 соответственно.

Функции K_1 и K_2 как математические ожидания выигрышей в ситуации (x, y) билинейны по x и y , а следовательно, образ $\psi(x_0, y_0)$ ситуации (x_0, y_0) при отображении ψ представляет собой выпуклое компактное подмножество в $X_1 \times X_2$. Более того, если последовательности пар $\{(x_0^n, y_0^n)\}$, $(x_0^n, y_0^n) \in X_1 \times X_2$ и $\{(x'_n, y'_n)\}$, $(x'_n, y'_n) \in \psi(x_0^n, y_0^n)$ имеют предельные точки, т. е.

$$\lim_{n \rightarrow \infty} (x_0^n, y_0^n) = (x_0, y_0), \quad \lim_{n \rightarrow \infty} (x'_n, y'_n) = (x', y'),$$

то в силу билинейности функций K_1 и K_2 и компактности множеств X_1 и X_2 , имеем, что $(x', y') \in \psi(x_0, y_0)$.

Тогда по теореме Какутани существует ситуация $(x^*, y^*) \in X_1 \times X_2$, для которой $(x^*, y^*) \in \psi(x^*, y^*)$, т. е.

$$K_1(x^*, y^*) \geq K_1(x, y^*), \quad K_2(x^*, y^*) \geq K_2(x^*, y)$$

для всех $x \in X_1$ и $y \in X_2$. Теорема доказана.

3.4.2. Предыдущая теорема может быть обобщена на случай непрерывных функций выигрыша H_1 и H_2 . При доказательстве этого результата требуется хорошо известная теорема о неподвижной точке, принадлежащая Брауэру [Партхасаратхи, Рагхаван, 1974]. Приведем без доказательства соответствующие теоремы.

Теорема 1. Пусть $\Gamma = (X_1, X_2, H_1, H_2)$ — бескоалиционная игра двух лиц, пространства стратегий $X_1 \subset R^m$, $X_2 \subset R^n$ — компактные выпуклые подмножества, а множество $X_1 \times X_2$ имеет внутренность. Пусть также функции выигрыша $H_1(x, y)$ и $H_2(x, y)$ непрерывны на $X_1 \times X_2$, причем $H_1(x, y)$ вогнута по x при каждом фиксированном y , а функция $H_2(x, y)$ вогнута по y при каждом фиксированном x . Тогда в игре Γ существует ситуация равновесия по Нэшу (x^*, y^*) .

Теорема 2. Пусть $\Gamma = (X_1, X_2, H_1, H_2)$ — бескоалиционная игра двух лиц, где H_1 и H_2 — непрерывные функции на $X_1 \times X_2$; X_1, X_2 — компактные подмножества конечномерных евклидовых пространств. Тогда игра Γ имеет ситуацию равновесия (μ, ν) в смешанных стратегиях.

Не будем подробно останавливаться на построении смешанных стратегий в бескоалиционных играх n лиц с бесконечным числом стратегий и доказательстве существования ситуации равновесия по Нэшу. Отметим только, что если функции выигрыша $H_i(x)$ непрерывны на декартовом произведении $X = \prod_{i=1}^n X_i$ компактных множеств чистых стратегий, то в такой бескоалиционной игре всегда существует ситуация равновесия по Нэшу в смешанных стратегиях. Для существования ситуаций, оптимальных по Парето, достаточно компактности множества $\{H(x)\}$, $x \in X$, что, в свою очередь, может быть обеспечено компактностью в некоторой топологии множества всех ситуаций X и непрерывностью в этой же топологии всех функций выигрыша K_i , $i = 1, 2, \dots, n$. Очевидно, что для конечных бескоалиционных игр это всегда выполнено.

§ 3.5. Доказательство существования ситуации равновесия в конечной игре n лиц

3.5.1. Читатель может изучать параграф 3.5 независимо от предыдущего параграфа 3.4. Рассмотрим игру в нормальной форме $\Gamma = \langle N, \{X_i\}_{i \in N}, \{H_i\}_{i \in N} \rangle$, где N ($|N| = n$) — множество игроков, X_i — конечное множество стратегий игрока i , H_i — функция выигрыша игрока i . Под смешанной стратегией игрока i , как и ранее, будем понимать вероятностное распределение на множестве чистых стратегий X_i . Обозначим как \bar{X}_i множество всех возможных смешанных стратегий игрока i . Для того, чтобы подчеркнуть отличие от смешанных стратегий, стратегии из множества X_i далее будем называть чистыми.

Ситуацией в смешанных стратегиях будем называть любой вектор, компонентами которого являются смешанные стратегии всех игроков. Таким образом, множество всех ситуаций в смешанных стратегиях представляет собой декартово произведение $\bar{X} = \prod_{i=1}^n \bar{X}_i$. Поэтому $\mu = (\mu_1, \dots, \mu_n)$ является ситуацией в смешанных стратегиях тогда и только тогда, когда каждому игроку $i \in N$ и каждой чистой стратегии $x_i \in X_i$ ситуация

μ ставит в соответствие неотрицательное вещественное число $\mu_i(x_i)$, представляющее собой вероятность выбора чистой стратегии x_i игроком i таким образом, что

$$\sum_{x_i \in X_i} \mu_i(x_i) = 1.$$

Если согласно ситуации μ игроки выбирают свои чистые стратегии независимо, то вероятность того, что они выберут ситуацию в чистых стратегиях $x = (x_1, \dots, x_i, \dots, x_n)$ равна произведению вероятностей $\prod_{i=1}^n \mu_i(x_i)$.

Для любой ситуации в смешанных стратегиях μ обозначим через $K_i(\mu)$ математическое ожидание выигрыша игрока i , который будет получен при независимом выборе игроками чистых стратегий согласно μ . Пусть $X = \prod_{i=1}^n X_i$ (X_i — множество всех возможных ситуаций в чистых стратегиях). Тогда

$$K_i(\mu) = \sum_{x \in X} \left(\prod_{j \in N} \mu_j(x_j) \right) H_i(x), \text{ для всех } i \in N.$$

Обозначим как $(\mu \| \tau_i)$ для всех $\tau_i \in \bar{X}_i$ ситуацию в смешанных стратегиях, где i -й компонентой является τ_i , а остальные компоненты такие же, как в μ . Таким образом,

$$K_i(\mu \| \tau_i) = \sum_{x \in X} \left(\prod_{j \in N \setminus \{i\}} \mu_j(x_j) \right) \tau_i(x_i) H_i(x).$$

Далее мы не будем использовать никакого специального обозначения для смешанной стратегии $\mu_i \in \bar{X}_i$, которая означает выбор чистой стратегии x_i с вероятностью 1, а просто будем обозначать ее через x_i , используя то же обозначение, что и для соответствующей чистой стратегии.

Если игрок i использует чистую стратегию x_i , в то время как все остальные игроки действуют независимо и выбирают стратегии согласно ситуации в смешанных стратегиях μ , то математическое ожидание выигрыша игрока i вычисляется следующим образом:

$$K_i(\mu \| x_i) = \sum_{x_j \in X_j, j \neq i} \left(\prod_{j \in N \setminus \{i\}} \mu_j(x_j) \right) H_i(\mu \| x_i).$$

3.5.2. Определение. Ситуация в смешанных стратегиях μ является ситуацией равновесия по Нэшу в смешанных стратегиях, если

$$K_i(\mu \| \tau_i) \leq K_i(\mu), \text{ для всех } \tau_i \in \bar{X}_i, i \in N.$$

3.5.3. Лемма. Для любого $\mu \in \prod_{i=1}^n \bar{X}_i$ и любого игрока $i \in N$ выполнено равенство

$$\max_{x_i \in X_i} K_i(\mu \| x_i) = \max_{\tau_i \in \bar{X}_i} K_i(\mu_i \| \tau_i).$$

Кроме того, $p_i \in \arg \max_{\tau_i \in \bar{X}_i} K_i(\mu \| \tau_i)$ тогда и только тогда, когда $p_i(x_i) = 0$ при всех x_i , таких что $x_i \notin \arg \max_{x_i \in X_i} K_i(\mu \| x_i)$.

Доказательство. Заметим, что для любого $\tau_i \in \bar{X}_i$ справедливо следующее равенство:

$$K_i(\mu \| \tau_i) = \sum_{x_i \in X_i} \tau_i(x_i) K_i(\mu \| x_i).$$

Здесь $K_i(\mu \parallel \tau_i)$ является математическим ожиданием для $K_i(\mu \parallel x_i)$. Следовательно, $K_i(\mu \parallel \tau_i)$ не может быть больше, чем максимальное значение случайной величины $K_i(\mu \parallel x_i)$, и оно строго меньше, чем это максимальное значение, всегда когда значение $K_i(\mu \parallel x_i)$ вычисляется для положительной вероятности ($\tau_i(x_i) \geq 0$, $\sum_{x_i \in X_i} \tau_i(x_i) = 1$).

Таким образом, наибольший ожидаемый выигрыш, который игрок i может получить при любых комбинациях смешанных стратегий других игроков, один и тот же, в независимости от того, использует игрок i смешанную стратегию или нет.

3.5.4. Как было показано выше, для случая игры двух лиц теорема Какутани о неподвижной точке является удобным математическим аппаратом для доказательства существования различных решений игры (принципов оптимальности), в том числе — равновесия по Нэшу. Перед тем как сформулировать теорему Какутани о неподвижной точке, напомним некоторые определения.

Множество S из конечномерного векторного пространства R^m называется *замкнутым* тогда и только тогда, когда для любой сходящейся последовательности векторов $\{x^j\}$, $j = 1, \dots, \infty$ выполнено свойство: если $x^j \in S$ при всех j , то и $\lim_{j \rightarrow \infty} x^j \in S$.

Множество S называется *ограниченным*, если существует такое положительное число K , что для любого $x \in S$, выполнено $\sum_{i=1}^m \xi_i^2 \leq K$ (здесь $x = \{\xi_i\}$, ξ_i являются компонентами x).

Заметим, что ограниченное, замкнутое множество в конечномерном евклидовом пространстве R^m является *компактом*.

Точечно-множественное отображение $F : X \rightarrow Y$ — это такое отображение, которое каждой точке x из X ставит в соответствие множество $F(x)$, являющееся подпространством Y . Предположим, что X и Y — метрические пространства, тогда понятия сходимости и предела должны быть определены для последовательностей точек из X и Y . Точечно-множественное отображение $F : X \rightarrow Y$ полунепрерывно сверху при выполнении следующего свойства для любой последовательности $x^j, y^j, j = 1, \dots, \infty$: если $x^j \in S$ и $y^j \in F(x^j)$ при всех j , причем последовательность $\{x^j\}$ сходится к некоторой точке \bar{x} и последовательность $\{y^j\}$ сходится к некоторой точке \bar{y} , то $\bar{y} \in F(\bar{x})$. Таким образом, $F : X \rightarrow Y$ является полунепрерывным сверху отображением, если множество $\{(x, y) : x \in X, y \in F(x)\}$ — замкнутое подмножество множества $X \times Y$.

Неподвижной точкой отображения $F : S \rightarrow S$ называется любая такая точка x из S , что $x \in F(x)$.

3.5.5. Теорема Какутани. Пусть S — непустое выпуклое компактное подмножество конечномерного векторного пространства R^m . Пусть $F : S \rightarrow S$ — любое полунепрерывное сверху точечно-множественное отображение, такое что для всех x из S $F(x)$ является непустым выпуклым подмножеством множества S . Тогда существует точка \bar{x} из S , такая что $\bar{x} \in F(\bar{x})$.

С доказательством теоремы Какутани о неподвижной точке читатель может ознакомиться в книге [Партхасаратхи, Рагхаван, 1974].

3.5.6. Теперь при помощи теоремы Какутани мы можем доказать следующий фундаментальный результат.

Теорема. В любой конечной игре n лиц в нормальной форме Γ существует по крайней мере одно равновесие по Нэшу в смешанных стратегиях.

Доказательство. Пусть Γ — конечная игра в нормальной форме

$$\Gamma = (N, \{X_i\}_{i \in N}, \{H_i\}_{i \in N}).$$

Множество ситуаций в смешанных стратегиях $\prod_{i=1}^n \bar{X}_i$ является непустым, выпуклым, замкнутым и ограниченным подмножеством конечномерного векторного пространства.

Это множество удовлетворяет приведенному выше определению ограниченного множества ($K = |N|$) и является подмножеством R^m , где $m = \sum_{i=1}^n |X_i|$ (здесь $|A|$ означает число элементов в конечном множестве A).

Введем следующее обозначение для всех $\mu \in \prod_{i=1}^n \bar{X}_i$ и всех игроков $j \in N$:

$$R_j(\mu) = \arg \max_{\tau_j \in \bar{X}_j} K_j(\mu \| \tau_j).$$

Множество $R_j(\mu)$ представляет собой множество наилучших ответов из \bar{X}_j на набор независимых смешанных стратегий $(\mu_1, \dots, \mu_{j-1}, \mu_{j+1}, \dots, \mu_N)$ других игроков. По предыдущей лемме $R_j(\mu)$ — множество всех распределений вероятностей ρ_j на X_j , таких что

$$\rho_j(x_j) = 0 \text{ при всех } x_j \notin \arg \max_{y_j \in \bar{X}_j} K_j(\mu \| y_j).$$

Таким образом, $R_j(\mu)$ выпукло, поскольку оно является подмножеством из \bar{X}_j , которое определено системой линейных уравнений. Кроме того, $R_j(\mu)$ непусто, т. к. оно содержит x_j из множества $\arg \max_{y_j \in \bar{X}_j} K_j(\mu \| y_j)$, являющегося непустым.

Пусть $R : \prod_{i=1}^n \bar{X}_i \rightarrow \prod_{i=1}^n \bar{X}_i$ — точно-множественное отображение, такое что

$$R(\mu) = \prod_{i=1}^n R_i(\mu), \text{ для всех } \mu \in \prod_{i=1}^n \bar{X}_i.$$

Значит, $\tau \in R(\mu)$ тогда и только тогда, когда $\tau_j \in R_j(\mu)$ для всех $j \in N$. Для любого μ , $R(\mu)$ непусто и выпукло, поскольку $R(\mu)$ является декартовым произведением непустых выпуклых множеств.

Для того, чтобы показать, что R является полунепрерывным сверху положим, что $\{\mu^k\}$ и $\{\tau^k\}$, $k = 1, \dots, \infty$ — сходящиеся последовательности, $\mu^k \in \prod_{i \in N} \bar{X}_i$, $k = 1, 2, \dots$; $\tau_k \in R(\mu_k)$, $k = 1, 2, \dots$; $\bar{\mu} = \lim_{k \rightarrow \infty} \mu^k$, $\bar{\tau} = \lim_{k \rightarrow \infty} \tau^k$.

Докажем, что $\bar{\tau} \in R(\bar{\mu})$. Для каждого игрока $j \in N$ и каждого $\rho_j \in \bar{X}_j$ справедливо неравенство

$$K_j(\mu^k \| \tau_j^k) \geq K_j(\mu^k \| \rho_j), \quad k = 1, 2, \dots$$

В силу непрерывности математического ожидания $K_j(\mu)$ по $\prod_{i=1}^n \bar{X}_i$, можно последовательно показать, что для всех $j \in N$ и $\rho_j \in \bar{X}_j$,

$$K_j(\bar{\mu} \| \bar{\tau}_j) \geq K_j(\bar{\mu} \| \rho_j).$$

Таким образом, $\bar{\tau}_j \in R_j(\bar{\mu})$ при всех $j \in N$, а по определению $R(\bar{\mu})$, выполнено и $\bar{\tau} \in R(\bar{\mu})$. Мы уже доказали, что $R : \prod_{i \in N} \bar{X}_i \rightarrow \prod_{i \in N} \bar{X}_i$ — полунепрерывное сверху отображение.

По теореме Какутани о неподвижной точке существует ситуация в смешанных стратегиях μ из $\prod_{i \in N} \bar{X}_i$, такая что $\mu \in R(\mu)$. Следовательно, $\mu_j \in R_j(\mu)$ при всех $j \in N$. Поэтому $K_j(\mu) \geq K_j(\mu \| \tau_j)$ при всех $j \in N$, $\tau_j \in \bar{X}_j$, т. е. μ — равновесие по Нэшу в смешанных стратегиях в игре Γ .

§ 3.6. Модификации концепции равновесия по Нэшу

3.6.1. Как было показано выше, равновесие по Нэшу обладает некоторыми недостатками, поэтому было предпринято много попыток усовершенствовать это понятие.

В настоящее время сложно выделить наиболее перспективные из новых определений. В этом параграфе мы приведем только некоторые из них. Для более глубокого изучения данного вопроса советуем обратиться к книге [van Damme, 1991].

3.6.2. В основе одного из подходов для модификации понятия равновесия по Нэшу лежит идея о том, что каждый игрок совершает ошибки с некоторой маленькой вероятностью и, как следствие, каждая чистая стратегия может быть выбрана с некоторой положительной (возможно, маленькой) вероятностью. Эта идея позаимствована из игр с «дрожанием руки», в которых игроки должны использовать только смешанные стратегии.

Пусть $\Gamma = \langle N, X_1, \dots, X_n, H_1, \dots, H_n \rangle$ — игра n лиц в нормальной форме. Обозначим, как и ранее, через \bar{X}_i множество смешанных стратегий игрока i , а через K_i — математическое ожидание выигрыша игрока i в смешанных стратегиях. Определим следующим образом $\eta_i(x_i)$, $x_i \in X_i$ и $\bar{X}_i(\eta_i)$ для $i \in N$:

$$\bar{X}_i(\eta_i) = \{\mu_i \in \bar{X}_i : \mu_i(x_i) \geq \eta_i(x_i), \text{ для всех } x_i \in X_i, \text{ где } \eta_i(x_i) > 0, \sum_{x_i \in X_i} \eta_i(x_i) < 1\}.$$

Пусть $\eta(x) = (\eta_1(x_1), \dots, \eta_n(x_n))$, $x_i \in X_i$, $i = 1, \dots, n$ и $\bar{X}[\eta(x)] = \prod_{i=1}^n \bar{X}_i(\eta_i(x_i))$. Игра с «дрожанием руки» (Γ, η) — это бесконечная игра в нормальной форме вида

$$\Gamma = \langle N, \bar{X}_1(\eta_1(x_1)), \dots, \bar{X}_n(\eta_n(x_n)), K_1(\mu_1, \dots, \mu_n), \dots, K_n(\mu_1, \dots, \mu_n) \rangle$$

определенная на множествах стратегий $\bar{X}_i(\eta_i(x_i))$ с функциями выигрышей $K_i(\mu_1, \dots, \mu_n)$, $\mu_i \in \bar{X}_i(\eta_i(x_i))$, $i = 1, \dots, n$.

3.6.3. Нетрудно заметить, что игра «дрожащей руки» (Γ, η) удовлетворяет условиям, при которых можно использовать теорему Какутани п. 3.5.5 о неподвижной точке. Следовательно, данная игра имеет по крайней мере одно равновесие. Очевидно, что для такого равновесия чистая стратегия, которая не является наилучшим ответом на действия других игроков, должна быть выбрана с минимальной вероятностью. Поэтому справедлив следующий результат.

Лемма. Ситуация в смешанных стратегиях $\mu \in \bar{X}(\eta)$ является равновесием по Нэшу в игре (Γ, η) тогда и только тогда, когда выполнено следующее условие:

$$\text{если } K_i(\mu \| x_k) < K_i(\mu \| x_l), \text{ то } \mu_i(x_k) = \eta_i(x_k), \text{ для всех } i, x_k, x_l.$$

3.6.4. Определение. Пусть Γ — игра в нормальной форме. Равновесие μ в игре Γ называется совершенным равновесием в игре Γ , если μ — предельная точка последовательности $\{\mu(\eta)\}_{\eta \rightarrow +0}$, где $\mu(\eta)$ — равновесие по Нэшу в игре «дрожащей руки» (Γ, η) для всех η .

Для того, чтобы равновесие μ в игре Γ было совершенным равновесием, достаточно, чтобы для некоторой ассоциированной игры «дрожащей руки» (Γ, η) при η близких к нулю существовало равновесие, близкое к μ , что не обязательно означает существование такого равновесия для всех игр «дрожащей руки» (Γ, η) при η близких к нулю. Пусть $\{(\Gamma, \eta^k)\}$, $k = 1, \dots, \infty$ — последовательность игр «дрожащей руки», для которых $\eta^k \rightarrow 0$ при $k \rightarrow \infty$. Поскольку в каждой игре (Γ, η^k) имеется по крайней мере одно равновесие μ^k , и поскольку μ является элементом компактного множества $\bar{X} = \prod_{j=1}^n \bar{X}_j$, существует предельная точка $\{\mu^k\}$. Нетрудно заметить, что эта точка является равновесием в игре Γ , которое удовлетворяет определению совершенного равновесия. Поэтому доказана теорема.

Теорема. Для любой игры в нормальной форме существует по крайней мере одно совершенное равновесие [Selten, 1975].

3.6.5. Пример 13. Рассмотрим биматричную игру Γ

$$\begin{array}{cc} & \begin{array}{cc} L_2 & R_2 \end{array} \\ \begin{array}{c} L_1 \\ R_1 \end{array} & \left[\begin{array}{cc} (1, 1) & (0, 0) \\ (0, 0) & (0, 0) \end{array} \right]. \end{array}$$

В этой игре есть два равновесия — (L_1, L_2) и (R_1, R_2) . Рассмотрим ассоциированную игру дрожащей руки (Γ, η) . В ситуации (R_1, R_2) в игре с дрожанием руки стратегии R_1 и R_2 выбираются с вероятностями $1 - \eta_1(L_1)$ и $1 - \eta_2(L_2)$ соответственно, а стратегии L_1 и L_2 выбираются с вероятностями $\eta_1(L_1)$ и $\eta_2(L_2)$. Таким образом, выигрыш $K_1^\eta(R_1, R_2)$ в игре (Γ, η) будет равен

$$K_1^\eta(R_1, R_2) = \eta_1(L_1) \cdot \eta_2(L_2).$$

В ситуации (L_1, R_2) стратегии L_1 и R_2 выбираются с вероятностями $(1 - \eta_1(R_1))$ и $(1 - \eta_2(L_2))$. Поэтому

$$K_1^\eta(L_1, R_2) = (1 - \eta_1(R_1))\eta_2(L_2).$$

Поскольку η мало, получаем

$$K_1^\eta(L_1, R_2) > K_1^\eta(R_1, R_2).$$

Поэтому в игре дрожащей руки ситуация (R_1, R_2) не является равновесной и (R_1, R_2) не является совершенным равновесием в изначальной игре. Непосредственной проверкой можно показать, что ситуация (L_1, L_2) является совершенным равновесием.

Пример 14. Теперь рассмотрим другую игру с матрицей

$$\begin{array}{cc} & \begin{array}{cc} L_2 & R_2 \end{array} \\ \begin{array}{c} L_1 \\ R_1 \end{array} & \left[\begin{array}{cc} (1, 1) & (10, 0) \\ (0, 10) & (10, 10) \end{array} \right]. \end{array}$$

Очевидно, что в этой игре есть два различных равновесия (L_1, L_2) и (R_1, R_2) . Равновесие (L_1, L_2) строго доминируется равновесием (R_1, R_2) . Рассмотрим ассоциированную игру дрожащей руки (Γ, η) . Покажем, что (L_1, L_2) — совершенное равновесие в игре (Γ, η) . Заметим, что

$$K_1(L_1, L_2) = (1 - \eta_1(R_1))(1 - \eta_2(R_2)) + 10(1 - \eta_1(R_1))\eta_2(R_2) + 10\eta_1(R_1)\eta_2(R_2),$$

$$K_1(R_1, L_2) = 10(1 - \eta_1(L_1))\eta_2(R_2) + 1 \cdot \eta_1(L_1)(1 - \eta_2(R_2)) + 10\eta_1(L_1)\eta_2(R_2).$$

При малых η_1, η_2 получаем неравенство:

$$K_1(L_1, L_2) > K_1(R_1, L_2).$$

Аналогичным образом можно показать, что справедливо следующее неравенство:

$$K_2(L_1, L_2) > K_2(L_1, R_2),$$

т. е. (L_1, L_2) является совершенным равновесием в игре Γ .

Для ситуации (R_1, R_2) в игре (Γ, η) выполнено

$$\begin{aligned} K_1(R_1, R_2) &= 10(1 - \eta_1(L_1))(1 - \eta_2(L_2)) + 10(1 - \eta_2(L_2))\eta_1(L_1) + \eta_1(L_1)\eta_2(L_2) = \\ &= 10(1 - \eta_2(L_2)) + \eta_1(L_1)\eta_2(L_2), \\ K_1(L_1, R_2) &= 10(1 - \eta_1(R_1))(1 - \eta_2(L_2)) + 10\eta_1(R_1)(1 - \eta_2(L_2)) + (1 - \eta_1(R_1))\eta_2(L_2) = \\ &= 10(1 - \eta_2(L_2)) + (1 - \eta_1(R_1))\eta_2(L_2). \end{aligned}$$

При малых η выполняется неравенство $K_1^\eta(L_1, R_2) > K_1^\eta(R_1, R_2)$. Таким образом, (R_1, R_2) не является равновесием в игре (Γ, η) и, следовательно, не является совершенным равновесием в игре Γ .

Поэтому (L_1, L_2) является равновесием в игре (Γ, η) и, следовательно, единственным совершенным равновесием в игре Γ , но выигрыши в этом равновесии доминируются ситуацией (R_1, R_2) . Таким образом, мы видим, что попытка улучшения равновесия по Нэшу приводит к потере равновесия с более привлекательными выигрышами. В то же самое время, усовершенствованная концепция не приводит к исключению всех интуитивно неоправданных равновесий.

В работе [Myerson, 1978] изучалась игра с матрицей

$$\begin{array}{c|cc} & L_2 & R_2 & A_2 \\ \begin{array}{c} L_1 \\ R_1 \\ A_1 \end{array} & \begin{bmatrix} (1, 1) \\ (0, 0) \\ (-2, -1) \end{bmatrix} & \begin{bmatrix} (0, 0) \\ (0, 0) \\ (-2, 0) \end{bmatrix} & \begin{bmatrix} (-1, -2) \\ (0, -2) \\ (-2, -2) \end{bmatrix} \end{array},$$

которая является развитием примера 13 путем добавления строго доминируемых строк A_1 и столбца A_2 .

Доказано, что в этой игре (R_1, R_2) также является совершенным равновесием в дополнение к (L_1, L_2) . Это следует из доказательства того, что если игроки договорились использовать равновесие (R_1, R_2) и каждый игрок ожидает, что ошибка A будет появляться с большей вероятностью, чем ошибка L , то оптимальным поведением для каждого игрока будет выбор стратегии R . Поэтому добавление строго доминирующих стратегий может изменять множество совершенных равновесий.

3.6.6. Существует другое понятие равновесия, введенное Р. Майерсоном в [Myerson, 1978], которое исключает некоторые «неразумные» совершенные равновесия, такие как равновесие (R_1, R_2) в последнем примере.

Это равновесие называется *правильным равновесием*. Основная идея, лежащая в определении этого равновесия, состоит в том, что игрок, совершающий ошибку, будет более упорно пытаться предотвратить более «дорогие» ошибки, чем «менее» дорогие, т. е. предполагается наличие некоторой рациональности в механизме совершения ошибок. В результате наиболее «дорогие» ошибки будут случаться с меньшей вероятностью, чем более «дешевые» ошибки.

3.6.7. Определение 1. Пусть $\langle N, \bar{X}_1, \dots, \bar{X}_n, K_1, \dots, K_n \rangle$ — игра n лиц в нормальной форме в смешанных стратегиях. Положим $\varepsilon > 0$ и пусть $\mu^\varepsilon \in \prod_{i=1}^n \bar{X}_i$. Будем говорить, что ситуация в смешанных стратегиях μ^ε является ε -правильным равновесием в игре Γ , если μ^ε — вполне смешанная и удовлетворяет следующему условию

$$\text{если } K_i(\mu^\varepsilon \| x_k) < K_i(\mu^\varepsilon \| x_l), \text{ то } \mu_i^\varepsilon(x_k) < \varepsilon \mu_i^\varepsilon(x_l) \text{ для всех } i, k, l.$$

Определение 2. Будем говорить, что $\mu \in \prod_{i=1}^n \bar{X}_i$ — правильное равновесие в игре Γ , если μ — предельная точка для последовательности μ^ε при $\varepsilon \rightarrow +0$, где μ^ε является ε -правильным равновесием в игре Γ .

Справедлива следующая теорема [Myerson, 1978].

Теорема. *В каждой игре в нормальной форме существует по крайней мере одно правильное равновесие.*

3.6.8. В представленных выше понятиях совершенного и правильного равновесий основная идея усовершенствования концепции Нэша основывалась на том, что «разумные» равновесия должны быть устойчивыми против небольших отклонений в равновесных стратегиях. Существуют также модификации равновесия по Нэшу, основанные на идее устойчивости равновесий при возмущениях в выигрышах. Однако здесь мы не можем привести все эти концепции и рекомендуем читателям самостоятельно ознакомиться с книгой Эрика ван Дамма [van Damme, 1991] для более глубокого изучения материала.

§ 3.7. Свойства оптимальных решений

3.7.1. Приведем свойства ситуации равновесия, которые помогают находить решение бескоалиционной игры двух лиц.

Теорема. *Для того, чтобы ситуация (μ^*, ν^*) в смешанных стратегиях в игре $\Gamma = (X_1, X_2, H_1, H_2)$ была ситуацией равновесия, необходимо и достаточно, чтобы для всех чистых стратегий $x \in X_1$ и $y \in X_2$ игроков 1 и 2, соответственно, выполнялись следующие неравенства:*

$$K_1(x, \nu^*) \leq K_1(\mu^*, \nu^*), \quad (3.7.1)$$

$$K_2(\mu^*, y) \leq K_2(\mu^*, \nu^*). \quad (3.7.2)$$

Доказательство. Необходимость очевидна, поскольку каждая чистая стратегия является частным случаем смешанной и, следовательно, должны быть выполнены неравенства (3.7.1), (3.7.2). Для доказательства достаточности необходимо перейти к смешанным стратегиям игроков 1 и 2, соответственно, в неравенствах (3.7.1), (3.7.2).

Эта теорема (как и в случае антагонистических игр) показывает, что для доказательства равновесности ситуации в смешанных стратегиях достаточно проверить неравенства (3.7.1), (3.7.2) только для чистых стратегий партнера. Для биматричной $(m \times n)$ -игры $\Gamma(A, B)$ эти неравенства принимают вид:

$$K_1(i, y^*) = a_i y^* \leq x^* A y^* = K_1(x^*, y^*), \quad (3.7.3)$$

$$K_2(x^*, j) = x^* b^j \leq x^* B y^* = K_2(x^*, y^*), \quad (3.7.4)$$

где $a_i(b^j)$ — строки (столбцы) матрицы $A(B)$, $i = 1, \dots, m$, $j = 1, \dots, n$.

3.7.2. Напомним, что для матричных игр каждая существенная чистая стратегия уравнивает любую оптимальную стратегию противника (см. п. 1.7.6). Аналогичный результат справедлив и для биматричных игр.

Теорема. *Пусть $\Gamma(A, B)$ — биматричная $(m \times n)$ -игра, а $(x, y) \in Z(\Gamma)$ — ситуация равновесия по Нэшу в смешанных стратегиях. Тогда выполняются равенства*

$$K_1(i, y) = K_1(x, y), \quad (3.7.5)$$

$$K_2(x, j) = K_2(x, y) \quad (3.7.6)$$

для всех $i \in M_x$ и $j \in N_y$, где $M_x(N_y)$ — спектр смешанной стратегии $x(y)$.

Доказательство. По теореме п. 3.7.1 имеем

$$K_1(i, y) \leq K_1(x, y) \quad (3.7.7)$$

для всех $i \in M_x$. Пусть выполняется хотя бы одно строгое неравенство в (3.7.7), т. е.

$$K_1(i_0, y) < K_1(x, y), \quad (3.7.8)$$

где $i_0 \in M_x$. Обозначим как ξ_i компоненты вектора $x = (\xi_1, \dots, \xi_m)$. Тогда $\xi_{i_0} > 0$ и

$$K_1(x, y) = \sum_{i=1}^m \xi_i K_1(i, y) = \sum_{i \in M_x} \xi_i K_1(i, y) < K_1(x, y) \sum_{i \in M_x} \xi_i = K_1(x, y).$$

Противоречие доказывает справедливость (3.7.5). Равенства (3.7.6) доказываются аналогично.

Данная теорема дает способ нахождения оптимальных смешанных стратегий игроков в игре $\Gamma(A, B)$. Действительно, предположим, что мы ищем ситуацию равновесия (x, y) , считая спектры стратегий M_x, N_y заданными. Тогда оптимальные стратегии должны удовлетворять системе линейных уравнений

$$y a_i = v_1, \quad x b^j = v_2, \quad (3.7.9)$$

где $i \in M_x, j \in N_y, v_1, v_2$ — некоторые числа. Если же ситуация равновесия (x, y) вполне смешанная, то система уравнений (3.7.9) принимает вид

$$A y = v_1 u, \quad x B = v_2 w, \quad (3.7.10)$$

где $u = (1, \dots, 1), w = (1, \dots, 1)$ — векторы соответствующей размерности, составленные из единиц; числа $v_1 = x A y, v_2 = x B y$ — выигрыши игроков в ситуации равновесия (x, y) .

3.7.3. Теорема. Пусть $\Gamma(A, B)$ — биматричная $(m \times n)$ -игра, где A, B — невырожденные. Если игра Γ имеет вполне смешанную ситуацию равновесия, то она единственная и вычисляется по формулам

$$x = v_2 u B^{-1}, \quad (3.7.11)$$

$$y = v_1 A^{-1} u, \quad (3.7.12)$$

где

$$v_1 = 1/(u A^{-1} u), \quad v_2 = 1/(u B^{-1} u). \quad (3.7.13)$$

Обратно, если для векторов $x, y \in R^m$, определяемых равенствами (3.7.11)–(3.7.13) справедливо $x \geq 0, y \geq 0$, то пара (x, y) образует ситуацию равновесия в смешанных стратегиях в игре $\Gamma(A, B)$ с вектором равновесных выигрышей (v_1, v_2) .

Доказательство. Если (x, y) — вполне смешанная ситуация равновесия, то x и y с необходимостью удовлетворяют системе (3.7.10). Умножая первое из равенств (3.7.10) на A^{-1} , а второе на B^{-1} , получаем (3.7.11), (3.7.12). С другой стороны, поскольку $x u = 1$ и $y u = 1$, находим значения для v_1 и v_2 . Единственность вполне смешанной ситуации (x, y) следует из единственности решения системы (3.7.10) в условиях теоремы.

Докажем обратное утверждение теоремы. По построению векторов x, y согласно (3.7.11)–(3.7.13), имеем $x u = y u = 1$. Отсюда и из условия $x \geq 0, y \geq 0$ следует, что (x, y) ситуация в смешанных стратегиях в игре Γ .

Согласно теореме п. 3.7.1 для того, чтобы ситуация (x, y) являлась ситуацией равновесия в смешанных стратегиях в игре $\Gamma(A, B)$, достаточно выполнения условий

$$a_i y = K_1(i, y) \leq x A y, \quad i = \overline{1, m},$$

$$x b^j = K_2(x, j) \leq x B y, \quad j = \overline{1, m},$$

или

$$A y \leq (x A y) u, \quad x B \leq (x B y) u.$$

Проверим справедливость этих соотношений для $x = \frac{u B^{-1}}{u B^{-1} u}$ и $y = \frac{A^{-1} u}{u A^{-1} u}$. Имеем

$$A y = \frac{u}{u A^{-1} u} = \frac{(u B^{-1} A A^{-1} u) u}{(u B^{-1} u)(u A^{-1} u)} = (x A y) u,$$

$$x B = \frac{u}{u B^{-1} u} = \frac{(u B^{-1} B A^{-1} u) u}{(u B^{-1} u)(u A^{-1} u)} = (x B y) u,$$

что и требовалось доказать.

Проиллюстрируем применение теоремы на примере игры «семейный спор» п. 3.1.4. Рассмотрим смешанное расширение игры. Множество точек, соответствующих векторам выигрышей в смешанных стратегиях, можно изобразить графически (рис. 3.2). Нетрудно заметить, что игра удовлетворяет условиям теоремы, поэтому здесь имеется единственная вполне смешанная ситуация равновесия (x, y) вычисляемая по формулам (3.7.11)–(3.7.13): $x = (4/5, 1/5)$, $y = (1/5, 4/5)$, $(v_1, v_2) = (4/5, 4/5)$.

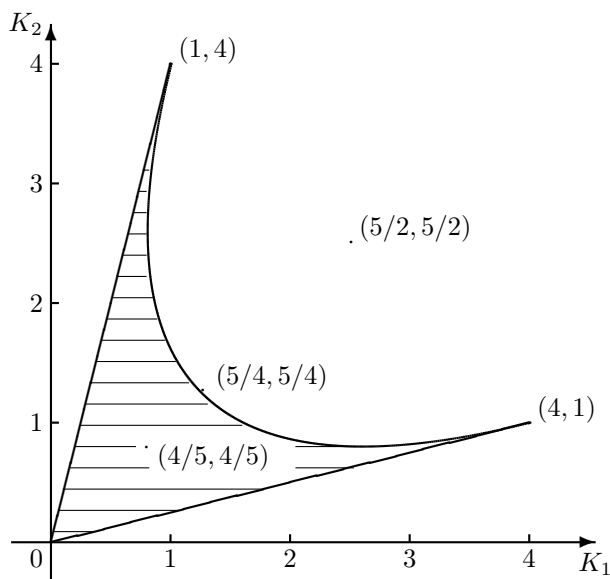


Рис. 3.2. Выигрыши в смешанных стратегиях

3.7.4. Рассмотрим свойства различных принципов оптимальности. Заметим, что определения оптимальности ситуации по Парето и Нэшу, приведенные в п.3.3.2, касаются произвольной бескоалиционной игры (в частности, двух лиц), поэтому они справедливы и для смешанного расширения $\bar{\Gamma}$. Следовательно, для игры двух лиц

$$Z(\bar{\Gamma}) = \bar{Z}^1 \cap \bar{Z}^2,$$

где $Z(\bar{\Gamma})$ — множество ситуаций равновесия по Нэшу, \bar{Z}^1 и \bar{Z}^2 — множества наилучших ответов игроков 1 и 2 в игре $\bar{\Gamma}$ и справедлива теорема о борьбе за лидерство (см. п. 3.2.5).

В более сложном отношении находятся ситуации, равновесные по Нэшу и оптимальные по Парето. Из примеров п. 3.3.2 следует, что возможны случаи, когда ситуация равновесна по Нэшу, но не оптимальна по Парето, и наоборот. Вместе с тем возможно, что одна и та же ситуация оптимальна и в том и в другом смысле (п. 3.2.4).

В примере 12 п. 3.3.3 было показано, что дополнительная ситуация равновесия, возникающая в смешанном расширении игры Γ , не является оптимальной по Парето в смешанном расширении Γ . Оказывается, что это довольно распространенное свойство биматричных игр.

Теорема. Пусть $\Gamma(A, B)$ — биматричная $(m \times n)$ -игра. Тогда почти для всех $(m \times n)$ -игр (за исключением не более чем счетного множества игр) справедливо следующее утверждение.

Ситуации равновесия по Нэшу в смешанных стратегиях, которые не являются равновесными в исходной игре, не являются оптимальными по Парето в смешанном расширении.

Доказательство теоремы основано на том, что ее результат справедлив для множества Ω так называемых регулярных игр, которое открыто и всюду плотно в множестве биматричных $(m \times n)$ -игр. Полное доказательство этой теоремы можно найти в [Мулен, 1985].

3.7.5. В заключение параграфа рассмотрим пример решения биматричных игр с малым числом стратегий, который во многом поучителен.

Пример 16 (Биматричные (2×2) -игры) [Мулен, 1985]. Рассмотрим игру $\Gamma(A, B)$, в которой у каждого из игроков по две чистые стратегии. Пусть

$$(A, B) = \begin{matrix} & \begin{matrix} \tau_1 & \tau_2 \end{matrix} \\ \begin{matrix} \delta_1 \\ \delta_2 \end{matrix} & \begin{bmatrix} (\alpha_{11}, \beta_{11}) & (\alpha_{12}, \beta_{12}) \\ (\alpha_{21}, \beta_{21}) & (\alpha_{22}, \beta_{22}) \end{bmatrix} \end{matrix}.$$

Здесь индексами $\delta_1, \delta_2, \tau_1, \tau_2$ обозначены чистые стратегии игроков 1 и 2 соответственно.

Предположим для простоты, что числа $\alpha_{11}, \alpha_{12}, \alpha_{21}, \alpha_{22}, (\beta_{11}, \beta_{12}, \beta_{21}, \beta_{22})$ различны.

Случай 1. В исходной игре Γ , по крайней мере, один игрок, (пусть игрок 1) имеет строго доминирующую стратегию, скажем δ_1 (см. п. 3.1.8). Тогда игра Γ и ее смешанное расширение $\bar{\Gamma}$ имеют единственную ситуацию равновесия по Нэшу. Действительно, неравенства $\alpha_{11} > \alpha_{21}, \alpha_{12} > \alpha_{22}$ приводят к тому, что в игре $\bar{\Gamma}$ чистая стратегия δ_1 строго доминирует все остальные смешанные стратегии первого игрока. Поэтому ситуацией равновесия является пара (δ_1, τ_1) , если $\beta_{11} > \beta_{12}$ или пара (δ_1, τ_2) , если $\beta_{11} < \beta_{12}$.

Случай 2. Игра Γ не имеет ситуации равновесия по Нэшу в чистых стратегиях. Здесь возможны два взаимоисключающих случая а) или б):

$$a) \alpha_{21} < \alpha_{11}, \alpha_{12} < \alpha_{22}, \beta_{11} < \beta_{12}, \beta_{22} < \beta_{21},$$

$$b) \alpha_{11} < \alpha_{21}, \alpha_{22} < \alpha_{12}, \beta_{12} < \beta_{11}, \beta_{21} < \beta_{22},$$

причем $\det A \neq 0, \det B \neq 0$ поэтому выполняются условия теоремы п. 3.7.3. Поэтому в игре существует ситуация равновесия (x^*, y^*) , где

$$x^* = \left(\frac{\beta_{22} - \beta_{21}}{\beta_{11} + \beta_{22} - \beta_{21} - \beta_{12}}, \frac{\beta_{11} - \beta_{12}}{\beta_{11} + \beta_{22} - \beta_{21} - \beta_{12}} \right), \quad (3.7.14)$$

$$y^* = \left(\frac{\alpha_{22} - \alpha_{12}}{\alpha_{11} + \alpha_{22} - \alpha_{21} - \alpha_{12}}, \frac{\alpha_{11} - \alpha_{21}}{\alpha_{11} + \alpha_{22} - \alpha_{21} - \alpha_{12}} \right) \quad (3.7.15)$$

а соответствующие равновесные выигрыши v_1 и v_2 определяются по формулам

$$v_1 = \frac{\alpha_{11}\alpha_{22} - \alpha_{12}\alpha_{21}}{\alpha_{11} + \alpha_{22} - \alpha_{21} - \alpha_{12}}, \quad v_2 = \frac{\beta_{11}\beta_{22} - \beta_{12}\beta_{21}}{\beta_{11} + \beta_{22} - \beta_{12} - \beta_{21}}.$$

Случай 3. Игра Γ имеет две ситуации равновесия по Нэшу. Этому случаю соответствует выполнение одно из условий:

$$a) \alpha_{21} < \alpha_{11}, \alpha_{12} < \alpha_{22}, \beta_{12} < \beta_{11}, \beta_{21} < \beta_{22},$$

$$b) \alpha_{11} < \alpha_{21}, \alpha_{22} < \alpha_{12}, \beta_{11} < \beta_{22}, \beta_{12} < \beta_{21}.$$

В случае а) равновесными будут ситуации (δ_1, τ_1) , (δ_2, τ_2) , а в случае б) — ситуации (δ_1, τ_2) , (δ_2, τ_1) . Однако в смешанном расширении есть еще одна вполне смешанная ситуация равновесия (x^*, y^*) , определенная формулами (3.7.14), (3.7.15).

Рассмотренные случаи исчерпывают изучение (2×2) -игры при условии, что элементы в матрицах различны.

§ 3.8. Эволюционно устойчивые стратегии

3.8.1. Пусть $\Gamma = (X, Y, A, B)$ — биматричная игра. Игра Γ называется симметричной, если множества X и Y совпадают ($X = Y$) и $\alpha_{ij} = \beta_{ji}$ для всех i, j .

Данное определение симметрии не является инвариантным по отношению к перестановкам множеств стратегий. Положим $|X| = |Y| = m$, а чистые стратегии будем обозначать как i или j . В теории эволюционных игр матрица A игрока 1 называется «фитнес» матрицей игры. Поскольку в случае симметричных игр матрица A полностью определяет игру, то будем далее идентифицировать ее с игрой Γ и называть игрой A . Смешанные стратегии x, y определим стандартным способом. Математическое ожидание выигрыша игрока 1 в ситуации (x, y) равно

$$E(x, y) = xAy = \sum_{i=1}^m \sum_{j=1}^m \alpha_{ij} \xi_i \eta_j.$$

Для любой смешанной стратегии $p = \{\xi\}$ обозначим через $C(p)$ носитель стратегии p и через $B(p)$ — множество наилучших ответов в чистых стратегиях против стратегии p в игре A :

$$C(p) = \{i : \xi_i > 0\}, \quad B(p) = \{i : E(i, p) = \max_j E(j, p)\}.$$

3.8.2. В работе [Maynard and Price (1973)] был рассмотрен пример игры, получившей название «Ястреб–голубь», который приводит к определению *эволюционно устойчивой стратегии* с общепринятым обозначением ESS.

Пример 17. Игра «Ястреб–голубь» является 2×2 симметричной биматричной игрой со следующими матрицами:

$$A = \begin{matrix} & \begin{matrix} H & D \end{matrix} \\ \begin{matrix} H \\ D \end{matrix} & \begin{bmatrix} 1/2(V-C) & V \\ 0 & 1/2V \end{bmatrix} \end{matrix}, \quad B = \begin{matrix} & \begin{matrix} H & D \end{matrix} \\ \begin{matrix} H \\ D \end{matrix} & \begin{bmatrix} 1/2(V-C) & 0 \\ V & 1/2V \end{bmatrix} \end{matrix}. \quad (3.8.1)$$

Предположим, что двое животных соперничают за некоторый ресурс (например, территорию в благоприятном для проживания месте) со значением V , т. е. получившее ресурс животное увеличивает ожидаемое число потомков на V . Для простоты мы предполагаем, что существует всего две чистых стратегии — «ястреб» (H) и «голубь» (D). Животное, применяющее стратегию «ястреб», всегда сражается за территорию в полную силу и уходит, только получив серьезные травмы. «Голубь» же только угрожает противнику некоторым удобным для него способом и отступает, когда получает серьезный отпор, не успев получить ранения. Два голубя могут разделить ресурс миролюбиво, но два ястреба обязательно будут сражаться, пока один из них не будет ранен и будет вынужден отступить. Предполагается, что ранение уменьшает ожидаемое число потомков на величину C . Далее будем предполагать, что у особей нет различий в размерах и возрасте, которые влияют на вероятность ранения. Тогда данный конфликт может быть описан при помощи биматричной игры (3.8.1).

Если $V > C$, то игра Ястреб–Голубь имеет единственное равновесие по Нэшу (H, H), т. е. для игрока всегда разумно сражаться. В популяции голубей и ястребов ястребы имеют больший репродуктивный успех, чем голуби, которые будут постепенно вымирать и в долгосрочной перспективе полностью вымрут.

Если $V < C$, то (H, H) не является равновесием, т. е. популяция, состоящая только из ястребов, не является устойчивой. В такой популяции мутанты, т. е. особи, использующие стратегию «голубь», имеют больший репродуктивный успех, что приводит к тому, что голуби будут распространяться в популяции, изначально состоящей из одних ястребов. Аналогично, популяция голубей также может быть завоевана ястребами, т. к. ситуация (D, D) не является равновесием по Нэшу.

Если $V < C$, то игра имеет единственное симметричное равновесие по Нэшу в смешанных стратегиях

$$\bar{x} = (\xi, 1 - \xi), \quad \bar{y} = (\xi, 1 - \xi),$$

где $\xi = V/C$.

Кроме того, в игре есть два асимметричных равновесия (H, D) и (D, H).

3.8.3. Предположим, что в мономорфной популяции, т. е. популяции, состоящей из особей (игроков) одного типа, игроки используют смешанные стратегии p в игре с репродуктивной матрицей A , а, кроме того, появляются мутанты, использующие стратегию q . Также будем считать, что популяция находится в неустойчивом состоянии, означающем, что малая часть ε игроков использует стратегию q .

Популяция вернется к своему изначальному состоянию, если количество потомков особей, использующих стратегию q будет меньше, чем количество потомков особей, использующих стратегию p .

Предположим, что (p, p) — симметричное равновесие по Нэшу в симметричной биматричной игре. Пусть второй игрок вместо стратегии p решает использовать смесь

двух смешанных стратегий p и q с вероятностями $1 - \varepsilon$, ε соответственно, где ε — малая величина. Тогда в общем случае для новой смешанной стратегии $y = (1 - \varepsilon)p + \varepsilon q$ множество наилучших ответов игрока 1 на стратегию y не обязательно будет содержать стратегию p .

Кроме того, может так случиться, что q будет лучшим ответом против y , чем p . Однако, если для любого q существует такое $\varepsilon > 0$, что p — лучший ответ на стратегию $y = (1 - \varepsilon)p + \varepsilon q$, чем q , то стратегия p игрока 1 будет в некотором роде устойчивой против малых возмущений, в том смысле, что для любого q найдется $\varepsilon > 0$, такое что

$$qA((1 - \varepsilon)p + \varepsilon q) < pA((1 - \varepsilon)p + \varepsilon q). \quad (3.8.1)$$

Если ситуация (p, p) является строгим равновесием ($pAp > qAp$ при всех q), то всегда имеет место неравенство (3.8.1). Существует также эволюционная интерпретация неравенства (3.8.1), основанная на примере игры «Ястреб–голубь».

Если выполняется неравенство (3.8.1), то

$$(1 - \varepsilon)qAp + \varepsilon qAq < (1 - \varepsilon)pAp + \varepsilon pAq. \quad (3.8.2)$$

Из (3.8.2) получаем, что неравенство $qAp > pAp$ невозможно, поскольку в этом случае неравенство (3.8.1) не будет выполняться для малых $\varepsilon > 0$. Тогда из (3.8.2) следует, что

$$qAp < pAp, \quad (3.8.3)$$

или

$$\text{если } qAp = pAp, \text{ то } qAq < pAq. \quad (3.8.4)$$

Из неравенств (3.8.3), (3.8.4) тривиальным образом следует выполнение (3.8.2) при достаточно малых $\varepsilon > 0$ (здесь ε зависит от q).

3.8.4. Определение. Будем называть смешанную стратегию p эволюционно устойчивой (ESS), если ситуация (p, p) является равновесием по Нэшу и выполнены следующие условия устойчивости:

$$\text{если } q \neq p \text{ и } qAp = pAp, \text{ то } qAq < pAq. \quad (3.8.5)$$

3.8.5. Рассмотрим отображение

$$p \rightarrow \{q \in Y, C(q) \subset B(p)\}.$$

Данное отображение удовлетворяет условиям теоремы Какутани о неподвижной точке. Поэтому существует такая точка p^* , что

$$p^* \in \{q \in Y, C(q) \subset B(p)\},$$

откуда

$$C(p^*) \subset B(p^*). \quad (3.8.6)$$

Из включения (3.8.6) следует, что

$$p^*Ap^* \geq qAp^*$$

при всех $q \in Y$ и поэтому (p^*, p^*) является симметричным равновесием по Нэшу. Тем самым мы доказали следующую теорему.

Теорема [Nash, 1951]. В каждой симметричной биматричной игре существует симметричное равновесие по Нэшу

3.8.6. Как уже было показано, если (p, p) является строгим равновесием по Нэшу, то p — эволюционно устойчивая стратегия (ESS), что также следует непосредственно из определения эволюционно устойчивой стратегии, поскольку в этом случае не существует $q \in Y$, таких что $qAp = pAp$.

Не все биматричные игры имеют эволюционно устойчивые стратегии. Например, если в матрице A элементы равны $\alpha_{ij} = \alpha$ для всех i, j , то невозможно выполнение неравенства (3.8.5).

3.8.7. Теорема. Пусть A — 2×2 -матрица с элементами $\alpha_{11} \neq \alpha_{21}$ и $\alpha_{12} \neq \alpha_{22}$. Тогда игра с матрицей A имеет эволюционно устойчивую стратегию.

Если $\alpha_{11} > \alpha_{21}$ и $\alpha_{22} > \alpha_{12}$, то A имеет два сильных равновесия $(1, 1)$, $(2, 2)$ и они являются эволюционно устойчивыми. Если же $\alpha_{11} < \alpha_{21}$, $\alpha_{22} < \alpha_{12}$, то игра с матрицей A имеет единственное симметричное равновесие (p, p) , которое является вполне смешанным ($C(p) = B(p) = \{1, 2\}$). При $q \neq p$ имеем $qAq - pAp = (q - p)A(q - p)$. Если $q = (\eta_1, \eta_2)$, $p = (\xi_1, \xi_2)$, то

$$(q - p)A(q - p) = (\eta_1 - \xi_1)^2(\alpha_{11} - \alpha_{21} + \alpha_{22} - \alpha_{12}) < 0.$$

Поскольку выполнено (3.8.5), следовательно, p является эволюционно устойчивой стратегией.

3.8.8. Рассмотрим игру с матрицей A

$$A = \begin{bmatrix} b & a & -a & a & -a \\ -a & b & a & -a & a \\ a & -a & b & a & -a \\ -a & a & -a & b & a \\ a & -a & a & -a & b \end{bmatrix}. \quad (3.8.8)$$

Если $0 < b < a$, то эта игра не имеет эволюционно устойчивой стратегии. Действительно, эта игра имеет единственное симметричное равновесие $p = (1/5, 1/5, 1/5, 1/5, 1/5)$, $pAp = b/5$, и каждая стратегия является наилучшим ответом на стратегию p , для всех i , $e_i Ae_i = \alpha_{ii} = b > b/5 = pAp = pAe_i$ (где $e_i = (0, \dots, 0, 1_i, 0, \dots, 0)$), поэтому условие (3.8.5) эволюционной устойчивости не выполняется.

Таким образом, для игр с более, чем двумя чистыми стратегиями, теорема не выполняется.

3.8.9. Интересен тот факт, что число эволюционно устойчивых стратегий в игре всегда конечно (хотя и может быть равно нулю).

Если (p, p) и (q, q) — равновесия по Нэшу в игре с матрицей A , причем $q \neq p$ и $C(q) \subset B(p)$, то p не может быть эволюционно устойчивой стратегией, поскольку q — наилучший ответ на p и q .

Теорема. Если p является эволюционно устойчивой стратегией в игре A и (q, q) — симметричное равновесие по Нэшу в игре A при $C(q) \subset B(p)$, то $p = q$.

3.8.10. Пусть (p^n, p^n) — последовательность симметричных равновесий по Нэшу в игре A , такая что $\lim_{n \rightarrow \infty} p^n = p$. Тогда из определения предела получаем, что существует такое N , что для всех $n \geq N$ справедливо

$$C(p) \subset C(p^n) \subset B(p^n) \subset B(p).$$

Из предыдущей теоремы получаем, что $p^n = p$ при $n \geq N$. Следовательно, любая эволюционно устойчивая стратегия является изолированной во множестве симметричных равновесных стратегий.

Из компактности множества ситуаций в смешанных стратегиях получаем, что если бы было бесконечное множество эволюционно устойчивых стратегий, то существовала бы и некоторая точка сгущения, но предыдущие рассуждения показывают, что это невозможно. Поэтому справедлива следующая теорема.

Теорема. Число эволюционно устойчивых стратегий конечно [Haigh, 1975].

§ 3.9. Равновесие в совместных смешанных стратегиях

3.9.1. Продолжим изучение игр двух лиц. Как уже отмечалось в §3.2, даже если ситуация равновесия является недоминируемой (оптимальной по Парето), возможны случаи, когда одна ситуация равновесия выгодна игроку 1, а другая — игроку 2. Это затрудняет нахождение взаимоприемлемого решения возникающего неантагонистического конфликта на уровне формализации бескоалиционной игры. Поэтому исследуем неантагонистический конфликт в формализации, разрешающей игрокам принимать совместные решения. Проиллюстрируем этот подход на примере игры «семейный спор» (см. пример 12 п. 3.1.4).

Пример 18. Рассмотрим смешанное расширение игры «семейный спор». Множество точек, соответствующих векторам выигрышей в смешанных стратегиях в игре, можно изобразить графически (см. рис. 3.2 п.3.5.3). На рисунке изображены две ситуации равновесия по Нэшу с векторами выигрышей $(1, 4)$, $(4, 1)$ в чистых стратегиях и одна вполне смешанная равновесная ситуация с вектором выигрышей $(4/5, 4/5)$ (ищется с использованием теоремы п. 3.5.3), которая менее предпочтительна для игроков, чем каждая из ситуаций равновесия в чистых стратегиях. Напомним, что равновесными здесь являются ситуации: (α_1, β_1) , (α_2, β_2) , (x^*, y^*) , где $x^* = (4/5, 1/5)$, $y^* = (1/5, 4/5)$, а ситуации (α_1, β_1) , (α_2, β_2) также оптимальны по Парето.

Если игра повторяется многократно, то игрокам имеет смысл сделать совместный выбор: с вероятностью $1/2$ выбирать ситуацию (α_1, β_1) или (α_2, β_2) . Тогда средний ожидаемый выигрыш игроков будет $(5/2, 5/2)$. Однако эта точка не лежит в множестве точек, соответствующих возможным ситуациям бескоалиционной игры (рис. 3.2), т. е. не может быть реализована, если игроки выбирают смешанные стратегии независимо.

Под *совместной смешанной стратегией* игроков будем понимать вероятностное распределение на множестве всевозможных пар (i, j) (ситуаций в чистых стратегиях), не обязательно порожденное независимыми случайными выборами чистых стратегий игроками 1 и 2. Такие стратегии могут быть реализованы посредником до начала игры.

Обозначим через M совместную смешанную стратегию в игре $\Gamma(A, B)$. Тогда ожидаемые выигрыши $K_1(M)$, $K_2(M)$ игроков 1 и 2 при использовании совместной смешанной стратегии соответственно равны

$$K_1(M) = \sum_{i,j} \alpha_{ij} \mu_{ij}, \quad K_2(M) = \sum_{i,j} \beta_{ij} \mu_{ij},$$

где $A = \{\alpha_{ij}\}$, $B = \{\beta_{ij}\}$ — матрицы выигрышей игроков, $M = \{\mu_{ij}\}$, при этом $uMw = 1$, $M \geq 0$, $u = (1, \dots, 1) \in R^m$, $w = (1, \dots, 1) \in R^n$. Геометрически множество точек, соответствующее множеству векторов выигрышей в совместных смешанных стратегиях, — это выпуклая оболочка множества точек возможных выигрышей в чистых стратегиях. Для игры примера 14 оно примет вид, как на рис. 3.3.

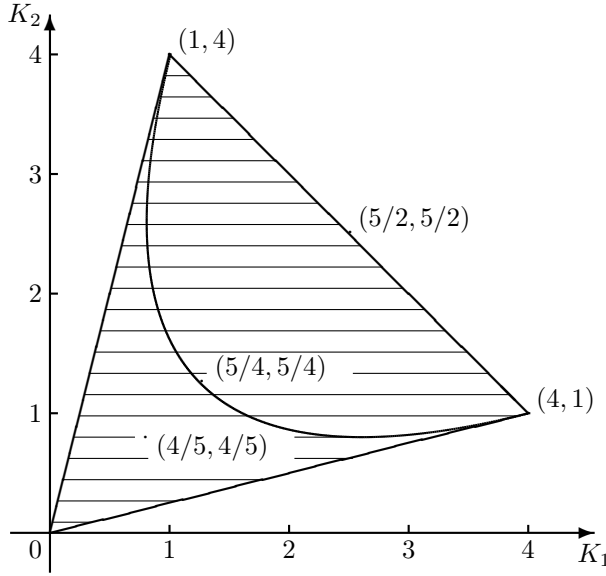


Рис. 3.3. Выигрыши в совместных смешанных стратегиях

Заметим, что совместная смешанная стратегия $M^* = \begin{bmatrix} 1/2 & 0 \\ 0 & 1/2 \end{bmatrix}$ является оптимальной по Парето и ей соответствует вектор выигрышей $(5/2, 5/2)$. Таким образом, M^* может быть рекомендована в качестве решения игры «семейный спор».

Определение. Для биматричной $(m \times n)$ -игры $\Gamma(A, B)$, обозначим через $M = \{\mu_{ij}\}$ совместное вероятностное распределение на парах (i, j) , $i = 1, \dots, m$, $j = 1, \dots, n$. Через $\mu_i(j)$ обозначим условную вероятность реализации стратегии j при условии, что реализовалась стратегия i . Аналогично, через $\nu_j(i)$ обозначим условную вероятность реализации стратегии i при условии, что реализовалась стратегия j .

$$\mu_i(j) = \begin{cases} \mu_{ij} / \sum_{j=1}^n \mu_{ij}, & \text{если } \sum_{j=1}^n \mu_{ij} \neq 0, \\ 0, & \text{если } \sum_{j=1}^n \mu_{ij} = 0, \quad j = 1, \dots, n, \end{cases}$$

$$\nu_j(i) = \begin{cases} \mu_{ij} / \sum_{i=1}^m \mu_{ij}, & \text{если } \sum_{i=1}^m \mu_{ij} \neq 0, \\ 0, & \text{если } \sum_{i=1}^m \mu_{ij} = 0, \quad i = 1, \dots, m. \end{cases}$$

Будем говорить, что $M^* = \{\mu_{ij}^*\}$ — ситуация равновесия в совместных смешанных стратегиях в игре $\Gamma(A, B)$, если выполнены следующие неравенства:

$$\sum_{j=1}^n \alpha_{ij} \mu_i^*(j) \geq \sum_{j=1}^n \alpha_{i'j} \mu_i^*(j), \quad \sum_{i=1}^m \beta_{ij} \nu_j^*(i) \geq \sum_{i=1}^m \beta_{ij'} \nu_j^*(i) \quad (3.9.1)$$

для всех $i, i' \in \{1, 2, \dots, m\}$ и $j, j' \in \{1, 2, \dots, n\}$.

3.9.2. Игру $\Gamma(A, B)$ в совместных смешанных стратегиях можно интерпретировать следующим образом. Пусть игроки договорились об использовании стратегии

$M^* = \{\mu_{ij}^*\}$ и пусть также в результате реализации случайного механизма выпала пара (i, j) , т. е. первый (второй) игрок получил номер $i(j)$ стратегии. Заметим, что каждый из игроков знает только свою реализацию. Этот игрок, вообще говоря, может не согласиться с реализацией $i(j)$, соответственно совместной стратегии и выбрать стратегию $i'(j')$. Тогда, если M^* — равновесная ситуация, то каждому из игроков невыгодно отклоняться от предложенной реализации $i(j)$, соответственно, что следует из (3.9.1), где в левой части неравенства стоит ожидаемый выигрыш игрока 1 (игрока 2) в случае согласия с реализацией $i(j)$.

Теперь предположим, что стратегия i игрока 1 такова, что $\mu_{ij} = 0$ для всех $j = 1, 2, \dots, n$. Тогда первое из неравенств (3.9.1), очевидно, выполняется. Аналогично, если $\mu_{ij} = 0$ для всех $i = 1, \dots, m$, то второе из неравенств (3.9.1) выполняется. Подставим выражения для $\mu_i(j)$ и $\nu_j(i)$ через μ_{ij} в формулы (3.9.1). Тогда получаем, что необходимым и достаточным условием равновесности ситуации $M^* = \{\mu_{ij}^*\}$ является выполнение неравенств

$$\sum_{j=1}^n \alpha_{ij} \mu_{ij}^* \geq \sum_{j=1}^n \alpha_{i'j} \mu_{ij}^*, \quad \sum_{i=1}^m \sum_{j=1}^n \mu_{ij}^* = 1, \quad \sum_{i=1}^m \beta_{ij} \mu_{ij}^* \geq \sum_{i=1}^m \beta_{ij'} \mu_{ij}^*, \quad \mu_{ij}^* \geq 0 \quad (3.9.2)$$

для всех $i, i' \in \{1, 2, \dots, m\}$ и $j, j' \in \{1, 2, \dots, n\}$.

Обозначим через $Z_c(\Gamma)$ множество равновесных ситуаций в совместных смешанных стратегиях.

Теорема. *Справедливы следующие утверждения.*

- 1) Множество $Z_c(\Gamma)$ равновесных ситуаций в совместных смешанных стратегиях в биматричной $(m \times n)$ -игре $\Gamma(A, B)$ является непустым выпуклым компактом пространства $R^{m \times n}$.
- 2) Если (x, y) — ситуация в смешанных стратегиях игры $\Gamma(A, B)$, то определяемая по ней ситуация $M = \{\mu_{ij}\}$ в совместных смешанных стратегиях будет равновесной тогда и только тогда, когда (x, y) — ситуация равновесия по Нэшу в смешанных стратегиях в игре $\Gamma(A, B)$.

Доказательство. Пусть (x, y) , $x = (\xi_1, \dots, \xi_m)$, $y = (\eta_1, \dots, \eta_n)$ — ситуация в смешанных стратегиях игры $\Gamma(A, B)$, а $M = \{\mu_{ij}\}$ — ситуация в смешанных стратегиях игры $\mu_{ij} = \xi_i \cdot \eta_j$, $i = 1, \dots, m$, $j = 1, \dots, n$. Необходимым и достаточным условием равновесности M является система неравенств (3.9.2), т. е.

$$\xi_i K_1(i, y) \geq \xi_i K_1(i', y), \quad \eta_j K_2(x, j) \geq \eta_j K_2(x, j'), \quad (3.9.3)$$

где $i, i' \in \{1, 2, \dots, m\}$, $j, j' \in \{1, \dots, n\}$. Если $\xi_i = 0$ ($\eta_j = 0$), то неравенства очевидны. Поэтому система неравенств (3.9.3) эквивалентна следующей:

$$K_1(i, y) \geq K_1(i', y), \quad K_2(x, j) \geq K_2(x, j'), \quad (3.9.4)$$

$i, i' \in \{1, \dots, m\}$, $j, j' \in \{1, \dots, n\}$, где i и j принадлежит спектрам стратегий x и y . Предположим, что (x, y) — ситуация равновесия по Нэшу в смешанных стратегиях в игре $\Gamma(A, B)$. Тогда согласно теореме п. 3.5.2

$$K_1(i, y) = K_1(x, y), \quad K_2(x, j) = K_2(x, y)$$

для всех i и j из спектров оптимальных стратегий. Поэтому неравенства (3.9.4) выполнены и $M \in Z_c(\Gamma)$.

Обратно, если (3.9.3) выполнено, то, суммируя неравенства (3.9.3) по i и по j , соответственно, и применяя теорему п. 3.5.1, получаем, что ситуация (x, y) равновесна по Нэшу.

Выпуклость и компактность множества $Z_c(\Gamma)$ следует из того, что $Z_c(\Gamma)$ — множество решений системы линейных неравенств (3.9.2), которое ограничено, а непустота — из существования ситуации равновесия по Нэшу в смешанных стратегиях (см. п. 3.4.1).

Теорема доказана.

Отметим, что совместная смешанная стратегия $M^* = \begin{bmatrix} 1/2 & 0 \\ 0 & 1/2 \end{bmatrix}$ равновесна в игре «семейный спор» (см. пример 1 п. 3.1.4), что просто установить проверкой неравенств (3.9.2).

§ 3.10. Задача о переговорах

3.10.1. Основным вопросом, который мы рассмотрим в данном параграфе, заключается в том, как разумным игрокам прийти к соглашению при совместном выборе решения в ходе переговоров. Перед тем как сформулировать задачу, еще раз вернемся к игре «семейный спор».

Пример 19. Рассмотрим множество R , соответствующее возможным векторам выигрышей в совместных смешанных стратегиях для игры «семейный спор» (область, заштрихованная на рис. 3.4). Действуя совместно, игроки могут реализовать любой выигрыш в смешанных стратегиях в области R . Однако это не означает, что они мо-

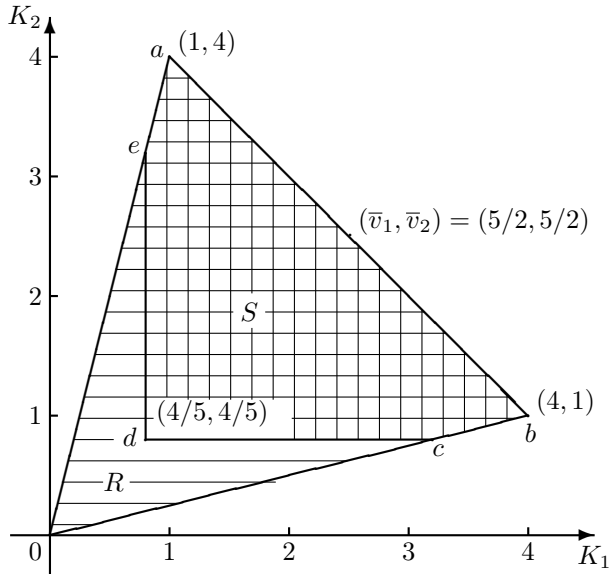


Рис. 3.4. Выигрыши в игре «семейный спор»

гут договориться о любом исходе игры. Так, игроку 1 наиболее предпочтительна точка $(4, 1)$, а игроку 2 — точка $(1, 4)$. Ни один из игроков не согласится с результатами переговоров, если его выигрыш будет меньше максиминного значения, поскольку этот выигрыш он может получить самостоятельно (независимо от партнера). Максиминные смешанные стратегии игроков в этой игре $x^0 = (1/5, 4/5)$ и $y^0 = (4/5, 1/5)$, соответственно, а вектор выигрышей в максиминных стратегиях (v_1^0, v_2^0) равен $(4/5, 4/5)$. Поэтому множество S , возможное для переговоров, ограничено точками a, b, c, d, e (см. рис. 3.4). Назовем его *переговорным множеством игры*. Далее, действуя совместно, игроки всегда могут договориться выбирать точки на отрезке \overline{ab} , поскольку это выгодно обоим (отрезок \overline{ab} соответствует ситуациям, оптимальным по Парето).

3.10.2. Назовем задачу выбора точки (\bar{v}_1, \bar{v}_2) из S в результате переговоров *задачей о переговорах*. Таким образом, мы пришли к следующей проблеме. Пусть для биматричной игры $\Gamma(A, B)$ задано переговорное множество S и вектор максиминных выигрышей (v_1^0, v_2^0) . Требуется найти правило, решающее задачу о переговорах, т. е. необходимо найти функцию φ , такую, что

$$\varphi(S, v_1^0, v_2^0) = (\bar{v}_1, \bar{v}_2). \quad (3.10.1)$$

Точку (v_1^0, v_2^0) будем называть точкой «статус кво».

Оказывается, что при некоторых разумных предположениях задача (3.10.1) разрешима в силу справедливости следующей теоремы.

Теорема. Пусть S — выпуклый компакт в R^2 , (v_1^0, v_2^0) — вектор максиминных выигрышей в игре $\Gamma(A, B)$. Множество S , пара (\bar{v}_1, \bar{v}_2) и функция φ удовлетворяют следующим условиям:

- 1) $(\bar{v}_1, \bar{v}_2) \geq (v_1^0, v_2^0)$.
- 2) $(\bar{v}_1, \bar{v}_2) \in S$.
- 3) Если $(v_1, v_2) \in S$ и $(v_1, v_2) \geq (\bar{v}_1, \bar{v}_2)$, то $(v_1, v_2) = (\bar{v}_1, \bar{v}_2)$.
- 4) Если $(\bar{v}_1, \bar{v}_2) \in \bar{S} \subset S$ и $(\bar{v}_1, \bar{v}_2) = \varphi(S, v_1^0, v_2^0)$, то $(\bar{v}_1, \bar{v}_2) = \varphi(\bar{S}, v_1^0, v_2^0)$.
- 5) Пусть T получается из S с помощью линейного преобразования $v'_1 = \alpha_1 v_1 + \beta_1$, $v'_2 = \alpha_2 v_2 + \beta_2$; $\alpha_1 > 0$, $\alpha_2 > 0$. Если $\varphi(S, v_1^0, v_2^0) = (\bar{v}_1, \bar{v}_2)$, то $\varphi(T, \alpha_1 v_1^0 + \beta_1, \alpha_2 v_2^0 + \beta_2) = (\alpha_1 \bar{v}_1 + \beta_1, \alpha_2 \bar{v}_2 + \beta_2)$.
- 6) Если из $(v_1, v_2) \in S$ следует $(v_2, v_1) \in S$; $v_1^0 = v_2^0$ и $\varphi(S, v_1^0, v_2^0) = (\bar{v}_1, \bar{v}_2)$, то $\bar{v}_1 = \bar{v}_2$.

Тогда существует единственная функция φ , удовлетворяющая 1) – 6), такая, что

$$\varphi(S, v_1^0, v_2^0) = (\bar{v}_1, \bar{v}_2).$$

Функция φ , которая отображает игру с переговорами (S, v_1^0, v_2^0) в множество векторов выигрышей (\bar{v}_1, \bar{v}_2) и удовлетворяет условиям 1) – 6), называется *арбитражной схемой Нэша* [Оуэн, 1971], условия 1) – 6) — *аксиомами Нэша*, а вектор (\bar{v}_1, \bar{v}_2) — *арбитражным вектором выигрышей*. Таким образом, арбитражная схема — это реализуемый принцип оптимальности в игре с переговорами.

Прежде чем перейти к доказательству теоремы, обсудим ее условия на примере игры «семейный спор» (см. рис. 3.4). Условия 1 и 2 означают, что вектор выигрышей (\bar{v}_1, \bar{v}_2)

находится в множестве, ограниченном точками a, b, c, d, e . Ограничение 3 показывает, что (\bar{v}_1, \bar{v}_2) лежит в множестве точек, оптимальных по Парето. Условие 4 говорит о независимости функции φ от посторонних стратегий, т. е. если (\bar{v}_1, \bar{v}_2) — арбитражный вектор выигрышей для множества \bar{S} , то при расширении множества переговоров до S решением будет либо (\bar{v}_1, \bar{v}_2) , либо другая точка, но не принадлежащая \bar{S} . Ограничение 5 говорит о том, что если функции выигрыша отличаются лишь масштабом измерения и началом отсчета, то также отличаются и результаты переговоров. Свойство 6 указывает на равноправность обоих игроков.

Доказательство теоремы п. 3.10.2 основано на следующих вспомогательных результатах.

3.10.3. Лемма. *Если существуют точки $(v_1, v_2) \in S$, такие что $v_1 > v_1^0$ и $v_2 > v_2^0$, то существует единственная точка (\bar{v}_1, \bar{v}_2) , максимизирующая функцию*

$$\theta(v_1, v_2) = (v_1 - v_1^0)(v_2 - v_2^0)$$

на подмножестве $S_1 \subset S$, $S_1 = \{ (v_1, v_2) \mid (v_1, v_2) \in S, v_1 \geq v_1^0 \}$.

Доказательство. По условию S_1 — непустой компакт, а θ — непрерывная функция, поэтому она достигает на нем своего максимума $\bar{\theta}$. По предположению, $\bar{\theta}$ положительно.

Пусть существуют две точки максимума (v'_1, v'_2) и (v''_1, v''_2) функции θ на S_1 . Заметим, что $v'_1 \neq v''_1$, поскольку в противном случае из вида функции θ следует $v'_2 = v''_2$.

Если $v'_1 < v''_1$, то $v'_2 > v''_2$. Так как множество S_1 выпукло, то $(\bar{v}_1, \bar{v}_2) \in S_1$, где $\bar{v}_1 = (v'_1 + v''_1)/2$, $\bar{v}_2 = (v'_2 + v''_2)/2$. Имеем

$$\begin{aligned} \theta(\bar{v}_1, \bar{v}_2) &= \frac{(v'_1 - v_1^0) + (v''_1 - v_1^0)}{2} \frac{(v'_2 - v_2^0) + (v''_2 - v_2^0)}{2} = \\ &= \frac{(v'_1 - v_1^0)(v'_2 - v_2^0)}{2} + \frac{(v''_1 - v_1^0)(v''_2 - v_2^0)}{2} + \frac{(v'_1 - v''_1)(v''_2 - v'_2)}{4}. \end{aligned}$$

Каждое из первых двух слагаемых последней суммы равно $\bar{\theta}/2$, а третье слагаемое положительно, что невозможно, поскольку $\bar{\theta}$ — максимум функции θ . Таким образом, точка (\bar{v}_1, \bar{v}_2) , максимизирующая функцию θ на множестве S_1 , единственна.

3.10.4. Лемма. *Пусть S удовлетворяет условиям леммы п.3.10.3, а (\bar{v}_1, \bar{v}_2) — точка максимума функции $\theta(v_1, v_2)$ и пусть*

$$\delta(v_1, v_2) = (\bar{v}_2 - v_1^0)v_1 + (\bar{v}_1 - v_1^0)v_2.$$

Если $(v_1, v_2) \in S$, то выполняется неравенство:

$$\delta(v_1, v_2) \leq \delta(\bar{v}_1, \bar{v}_2).$$

Доказательство. Предположим, что существует такая точка $(v_1, v_2) \in S$, что $\delta(v_1, v_2) > \delta(\bar{v}_1, \bar{v}_2)$. Из выпуклости S имеем: $(v'_1, v'_2) \in S$, где $v'_1 = \bar{v}_1 + \varepsilon(v_1 - \bar{v}_1)$ и $v'_2 = \bar{v}_2 + \varepsilon(v_2 - \bar{v}_2)$, $0 < \varepsilon < 1$. В силу линейности $\delta(v_1 - \bar{v}_1, v_2 - \bar{v}_2) > 0$. Имеем

$$\theta(v'_1, v'_2) = \theta(\bar{v}_1, \bar{v}_2) + \varepsilon\delta(v_1 - \bar{v}_1, v_2 - \bar{v}_2) + \varepsilon^2(v_1 - \bar{v}_1)(v_2 - \bar{v}_2).$$

Последнее слагаемое — бесконечно малая величина порядка $o(\varepsilon)$. Поэтому при достаточно малом $\varepsilon > 0$ получаем неравенство $\theta(v'_1, v'_2) > \theta(\bar{v}_1, \bar{v}_2)$, но это противоречит максимальности $\theta(\bar{v}_1, \bar{v}_2)$.

3.10.5. Перейдем к доказательству теоремы п. 3.10.2. Для этого покажем, что точка (\bar{v}_1, \bar{v}_2) , которая максимизирует $\theta(v_1, v_2)$, является решением задачи о переговорах.

Доказательство. Предположим, что выполнены условия леммы п. 3.10.3. Тогда определена точка (\bar{v}_1, \bar{v}_2) , которая максимизирует $\theta(v_1, v_2)$. Можно проверить, что (\bar{v}_1, \bar{v}_2) удовлетворяет условиям 1) – 4) теоремы п. 3.10.2. Она также удовлетворяет условию 5) этой теоремы, так как если $v'_1 = \alpha_1 v_1 + \beta_1$ и $v'_2 = \alpha_2 v_2 + \beta_2$, то

$$\theta'(v'_1, v'_2) = [v'_1 - (\alpha_1 v_1^0 + \beta_1)][v'_2 - (\alpha_2 v_2^0 + \beta_2)] = \alpha_1 \alpha_2 \theta(v_1, v_2),$$

и если (\bar{v}_1, \bar{v}_2) максимизирует $\theta(v_1, v_2)$, то (\bar{v}'_1, \bar{v}'_2) максимизирует $\theta'(v'_1, v'_2)$. Покажем, что (\bar{v}_2, \bar{v}_1) удовлетворяет условию 6). Пусть множество S симметрично в смысле условия 6). Тогда $(\bar{v}_2, \bar{v}_1) \in S$ и $\theta(\bar{v}_1, \bar{v}_2) = \theta(\bar{v}_2, \bar{v}_1)$. Так как (\bar{v}_1, \bar{v}_2) — единственная точка, которая максимизирует $\theta(v_1, v_2)$ на S_1 , то $(\bar{v}_1, \bar{v}_2) = (\bar{v}_2, \bar{v}_1)$, т. е. $\bar{v}_1 = \bar{v}_2$.

Таким образом, точка (\bar{v}_1, \bar{v}_2) удовлетворяет условиям 1)– 6). Покажем, что это единственное решение задачи о переговорах. Рассмотрим множество

$$R = \{ (v_1, v_2) \mid \delta(v_1, v_2) \leq \delta(\bar{v}_1, \bar{v}_2) \}. \quad (3.10.2)$$

По лемме п. 3.10.4 имеет место включение $S \subset R$. Пусть T получается из R с помощью преобразования

$$v'_1 = \frac{v_1 - v_1^0}{\bar{v}_1 - v_1^0}, \quad v'_2 = \frac{v_2 - v_2^0}{\bar{v}_2 - v_2^0}. \quad (3.10.3)$$

Выражая v_1 и v_2 из (3.10.3) и подставляя в (3.10.2), получаем, что

$$T = \{ (v'_1, v'_2) \mid v'_1 + v'_2 \leq 2 \}$$

и $v_1^0 = v_2^0 = 0$. Поскольку T симметрично, то из свойства 6) имеем, что решение (если оно существует) должно лежать на прямой $v'_1 = v'_2$, а согласно свойству 3) оно должно быть точкой $(1, 1) = \varphi(T, 0, 0)$. Обращая преобразование (3.10.3) и применяя свойство 5), получаем, что $(\bar{v}_1, \bar{v}_2) = \varphi(R, v_1^0, v_2^0)$. Так как $(\bar{v}_1, \bar{v}_2) \in S$, а $S \subset R$, на основании свойства 4 пара (\bar{v}_1, \bar{v}_2) является решением для (S, v_1^0, v_2^0) .

Предположим теперь, что условия леммы п. 3.10.3 не выполнены, т. е. не существует точек $(v_1, v_2) \in S$, для которых $v_1 > v_1^0$ и $v_2 > v_2^0$. Тогда возможны следующие случаи.

а) Существуют точки, у которых $v_1 > v_1^0$ и $v_2 = v_2^0$. Тогда в качестве (\bar{v}_1, \bar{v}_2) возьмем точку в S , которая максимизирует v_1 при ограничении $v_2 = v_2^0$.

б) Существуют точки, у которых $v_1 = v_1^0$ и $v_2 > v_2^0$. В этом случае в качестве (\bar{v}_1, \bar{v}_2) возьмем точку в S , которая максимизирует v_2 при ограничении $v_1 = v_1^0$.

с) Переговорное множество S вырождается в точку (v_1^0, v_2^0) максимальных выигрышей (например, случай матричных игр). Полагаем $\bar{v}_1 = v_1^0$, $\bar{v}_2 = v_2^0$.

Непосредственно можно проверить, что эти решения удовлетворяют свойствам 1) – 6), при этом из свойств 1) – 3) следует единственность. Теорема доказана.

В игре «семейный спор» (см. пример 14) схема Нэша дает арбитражный выигрыш $(\bar{v}_1, \bar{v}_2) = (5/2, 5/2)$ (см. рис. 3.4).

3.10.6. В этом разделе дан краткий обзор аксиоматической теории переговоров n игроков. Хотя модификации концепции решения Нэша появились вскоре после публикации работы Нэша [Nash, 1951], заметим, что до середины 1970-х годов решения Нэша часто использовались экономистами и специалистами в области теории игр как

главное, если не единственное, решение задачи о переговорах. Поскольку все существующие решения в действительности являются инвариантными относительно параллельного переноса, будет удобно рассматривать как допустимые только те задачи, в которых точка «статус кво» может быть без ограничения общности перенесена в начало координат. Следовательно, всегда можно считать выполненным условие $v^0 = (v_1^0, \dots, v_n^0) = (0, \dots, 0) \in R^n$ и стандартная задача просто обозначается как S вместо $(S, 0)$. Кроме того, далее все задачи рассматриваются на подмножестве R_+^n (вместо R^n). Это означает, что все альтернативы, которые будут приносить игроку меньше, чем он получает в точке «статус кво» $v^0 = 0$, не принимаются во внимание.

Определение. Будем говорить, что определено арбитражное решение Нэша ϕ , если для всех выпуклых, компактных подмножеств $S \subset R_+^n$, содержащих по крайней мере один вектор с положительными координатами (обозначим $S \in \Sigma^n$), $\phi(S)$ совпадает с максимизирующим вектором $v \in S$ в «произведении Нэша» $\prod_{i=1}^n v_i$.

Арбитражное решение Нэша базируется на выполнении следующих аксиом:

1⁰. *Парето оптимальность:* Если $v \geq \varphi(S)$ и $v \neq \varphi(S)$, то $v \notin S$ [$\varphi(S) \in PO(S)$] для всех $S \in \Sigma^n$, $v \in R^n$.

Некоторым ослаблением этого условия является следующая аксиома:

2⁰. *Слабая Парето оптимальность:* Если $v > \varphi(S)$, то $v \notin S$ для всех $S \in \Sigma^n$, $v \in R^n$.

Пусть $\Pi^n : \{1, \dots, n\} \rightarrow \{1, \dots, n\}$ — класс перестановок порядка n . Для $\Pi \in \Pi^n$ и $v \in R^n$, положим $\pi(v) = (v_{\pi(1)}, \dots, v_{\pi(n)})$. Кроме того, пусть для $S \subset R^n$, выполнено $\pi(S) = \{v' \in R^n \mid \exists v \in S \text{ при } v' = \pi(v)\}$.

3⁰. *Симметричность:* Пусть $S \in \Sigma^n$. Если для всех $\pi \in \Pi^n$, $\pi(S) = S$, то $\varphi_i(S) = \varphi_j(S)$ для всех i, j (отметим, что $\pi(S) \in \Sigma^n$).

Пусть $L^n : R^n \rightarrow R^n$ — класс положительных независимых взаимнооднозначных линейных преобразований порядка n . Каждое $l \in L^n$ характеризуется n положительными числами α_i , такими что для $v \in R^n$, $l(v) = (\alpha_1 v_1, \dots, \alpha_n v_n)$. Далее, пусть для $S \subset R^n$ выполнено $l(S) = \{v' \in R^n \mid \exists v \in S \text{ при } v' = l(v)\}$.

4⁰. *Масштабная инвариантность:* Для всех $S \in \Sigma^n$ и $l \in L^n$, выполняется $\varphi(l(S)) = l(\varphi(S))$ [отметим, что $l(S) \in \Sigma^n$].

5⁰. *Независимость от посторонних альтернатив:* Если $S' \subset S$ и $\varphi(S) \in S'$, то $\varphi(S') = \varphi(S)$ для всех $S, S' \subset \Sigma^n$.

В предыдущем разделе мы доказали теорему Нэша для $n = 2$ и получили, что только одно решение удовлетворяет сформулированным аксиомам. Этот результат непосредственно обобщается на произвольное n .

Теорема. Решение $\varphi(S)$, $S \in \Sigma^n$ удовлетворяет аксиомам 1⁰, 3⁰, 4⁰, 5⁰ тогда и только тогда, когда оно является арбитражным решением Нэша.

Эта теорема является основой аксиоматической теории переговоров. Она доказывает, что в каждой задаче существует единственная точка, соответствующая справедливому компромиссу. В середине 1970 годов результаты Нэша получили большую известность, и аксиомы были тщательно исследованы несколькими авторами.

6⁰. *Сильная индивидуальная рациональность:* Для всех $S \in \Sigma^n$, выполнено $\varphi(S) > 0$.

Теорема. [Roth, 1977]. Решение $\varphi(S)$, $S \in \Sigma^n$ удовлетворяет аксиомам 3⁰, 4⁰, 5⁰, 6⁰ тогда и только тогда, когда оно является арбитражным решением Нэша.

Если отбросить аксиому 3⁰ из набора аксиом в теореме 3.10.6, то появляется небольшой дополнительный набор решений.

3.10.7. Определение. Для $a = (\alpha_1, \dots, \alpha_n)$, $\alpha_i > 0$, $i = 1, \dots, n$, $\sum_{i=1}^n \alpha_i = 1$, определим асимметричное решение Нэша N^a с весами a если для всех $S \in \sum^n$ выполняется $N^a(S) = \arg \max \prod_{i=1}^n v_i^{\alpha_i}$, $v \in S$.

Данное решение было предложено в работе [Harsanyi, Selten, 1972].

Теорема. Решение $\varphi(S)$, $S \in \sum^n$ удовлетворяет аксиомам $4^0, 5^0, 6^0$ тогда и только тогда, когда оно является асимметричным решением Нэша.

Если отказаться от аксиомы 6^0 , то появляется еще несколько решений.

3.10.8. Определение. Для $i \in \{1, \dots, n\}$ определим диктаторское решение D^i , если для всех $S \in \sum^n$, $D^i(S)$ совпадает с максимизирующей точкой множества S в направлении i -го единичного вектора.

3.10.9. Решение Калаи–Смородинского. Новым толчком для развития аксиоматической теории переговоров послужила работа [Kalai, Smorodinsky, 1975], в которой было предложено следующее решение (см. Рис. 3.5).

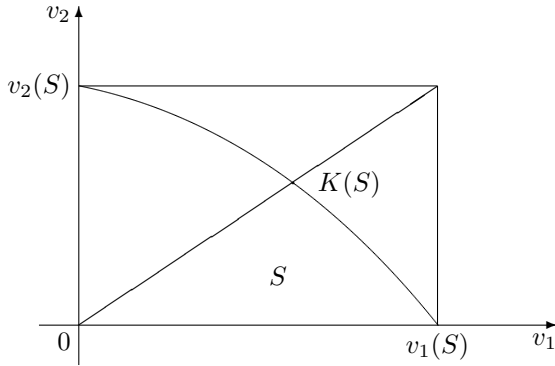


Рис. 3.5. Решение Калаи–Смородинского

Определение. Решение называется решением Калаи–Смородинского K , если для всех $S \in \sum^n$, $K(S)$ является максимизирующей точкой множества S на отрезке, соединяющем изначальную точку «статус кво» (в примере $(0; 0)$) с идеальной точкой S , которая определяется как $v_i(S) = \max\{v_i \mid v \in S\}$ для всех i .

3.10.10. Эгалитарное решение. Теперь вернемся к третьему решению, которое наиболее существенно отличается от первых двух.

Определение. Эгалитарным решением E будем называть решение, которое для всех $S \in \sum^n$, $E(S)$ является максимизирующей точкой множества S среди точек с одинаковыми координатами (см. Рис. 3.6).

Отличительной чертой этого решения является то, что оно удовлетворяет условию монотонности, которое является очень сильным условием. В самом деле, эта аксиома может применяться для простого определения решения.

8^0 . Строгая монотонность: Если $S \subset S'$, то $\varphi(S) \leq \varphi(S')$ для всех $S, S' \in \sum^n$.

Теорема. Решение $\varphi(S)$, $S \in \sum^n$ удовлетворяет аксиомам $2^0, 3^0, 8^0$ тогда и только тогда, когда оно является эгалитарным решением.

3.10.11. Утилитарное решение. Закончим наш обзор определением утилитарного решения.

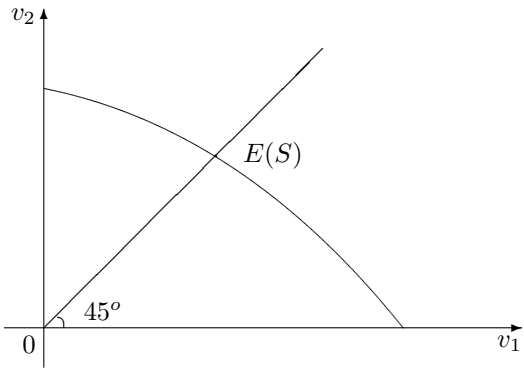


Рис. 3.6. Эгалитарное решение

Определение. Утилитарное решение U определяется выбором для каждого $S \in \sum^n$ точек, максимизирующих сумму $\sum_{i=1}^n v_i$ для $v \in S$ (см. Рис. 3.7).

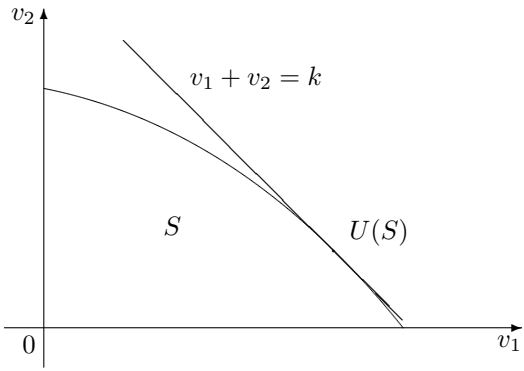


Рис. 3.7. Утилитарное решение

Очевидно, все утилитарные решения удовлетворяют аксиоме 1⁰. Однако, никакие утилитарные решения не удовлетворяют 4⁰. Также никакие утилитарные решения не удовлетворяют аксиоме 5⁰. Утилитарное решение было введено Майерсоном (1981).

Остальные решения обсуждаются в книгах [Льюис и Райфа, 1961; Perles, Mashler, 1981]. В этом разделе мы использовали работу [Thomson and Lensberg, 1990], в которой читатель может ознакомиться с доказательством теорем.

§ 3.11. Игры в форме характеристической функции

В § 3.9–3.10 на примере игр двух лиц было показано, как, используя возможность согласованного выбора стратегий, игроки могут прийти к взаимоприемлемому решению возникающего неантагонистического конфликта (стратегический подход). Теперь будем считать, что условия игры допускают совместные действия игроков и перераспределение выигрыша. Это предполагает, что полезности различных игроков могут быть оценены единой шкалой (трансферабельные выигрыши), и поэтому взаимное перераспределение выигрышей не искажает содержательной постановки первоначальной задачи. Представляется естественным, что объединение игроков в максимальную коалицию (в коалицию, состоящую из всех игроков) с целью получения максимального суммарного выигрыша приведет к наилучшим результатам также и с точки зрения каждого игрока, при этом нас будет интересовать не столько как коалиция игроков добывается своего суммарного выигрыша, сколько как он будет распределен между членами коалиции (кооперативный подход).

В § 3.11–3.14 рассмотрена кооперативная теория игр n лиц. В ней исследуются условия, при которых объединение игроков в максимальную коалицию является целесообразным, а отдельные игроки не будут иметь желания создавать меньшие группировки или действовать индивидуально.

3.11.1. Пусть $N = \{1, \dots, n\}$ — множество всех игроков. Любое непустое подмножество $S \subset N$ называется *коалицией*.

Определение. *Характеристической функцией игры n лиц будем называть вещественную функцию v , определенную на коалициях $S \subset N$, при этом для любых непересекающихся коалиций T, S ($T \subset N, S \subset N$) выполняется неравенство*

$$v(T) + v(S) \leq v(T \cup S), \quad v(\emptyset) = 0. \quad (3.11.1)$$

Свойство (3.11.1) называется *свойством супераддитивности*. Оно необходимо для содержательной интерпретации числа $v(T)$ как гарантированного выигрыша коалиции T в случае, когда она действует независимо от остальных игроков. При такой интерпретации неравенство (3.11.1) означает, что коалиция $S \cup T$ имеет не меньше возможностей, чем две непересекающиеся коалиции S и T , действующие независимо.

Из супераддитивности v получаем, что для любых непересекающихся коалиций S_1, \dots, S_k справедливо

$$\sum_{i=1}^n v(S_i) \leq v(N).$$

Отсюда, в частности, следует, что не существует такого разбиения множества N на коалиции, чтобы суммарный гарантированный выигрыш этих коалиций превышал максимальный выигрыш всех игроков $v(N)$.

3.11.2. Рассмотрим бескоалиционную игру $\Gamma = (N, \{X_i\}_{i \in N}, \{H_i\}_{i \in N})$.

Пусть игроки, составляющие некоторую коалицию $S \subset N$, объединяют свои усилия с целью увеличения своего суммарного выигрыша. Установим, какой наибольший выигрыш они могут себе гарантировать. Совместные действия игроков из коалиции S означают, что коалиция S , действуя от имени своих членов как один игрок (обозначим его 1), имеет в качестве множества чистых стратегий всевозможные комбинации стратегий, составляющих ее игроков из S , т. е. элементы декартового произведения

$$X_S = \prod_{i \in S} X_i.$$

Общность интересов игроков из S означает, что выигрыш коалиции S (игрока 1) есть сумма выигрышей игроков из S , т. е.

$$H_S(x) = \sum_{i \in S} H_i(x),$$

где $x \in X_N$, $x = (x_1, \dots, x_n)$ — ситуация в чистых стратегиях.

Нас интересует тот наибольший выигрыш, который игроки из S могут себе гарантировать. В худшем случае для S (игрока 1) оставшиеся игроки из $N \setminus S$ могут также объединиться в коллективного игрока 2 с множеством стратегий $X_{N \setminus S} = \prod_{i \in N \setminus S} X_i$ и интересом, диаметрально противоположным игроку 1 (т. е. выигрыш игрока 2 в ситуации x равен $-H_S(x)$). В результате таких рассуждений вопрос о наибольшем гарантированном выигрыше коалиции S превратился в вопрос о наибольшем гарантированном выигрыше игрока 1 в антагонистической игре $\Gamma_S = (X_S, X_{N \setminus S}, H_S)$. В смешанном расширении $\bar{\Gamma}_S = (\bar{X}_S, \bar{X}_{N \setminus S}, K_S)$ игры Γ_S гарантированный выигрыш $v(S)$ игрока 1 может разве лишь увеличиться по сравнению с игрой Γ_S , поэтому в дальнейшем будем рассматривать смешанное расширение игры Γ_S . Заметим, в частности, что при такой интерпретации $v(S)$ совпадает со значением игры $\bar{\Gamma}_S$ (если оно существует), а $v(N)$ — максимальный суммарный выигрыш игроков. Очевидно, что $v(S)$ зависит в результате только от коалиции S (и еще от самой исходной бескоалиционной игры, которая в наших рассуждениях остается одной и той же), являясь ее функцией. Убедимся, что эта функция является характеристической функцией бескоалиционной игры. Для этого достаточно показать выполнение условия (3.11.1).

Заметим, что для каждой бескоалиционной игры, построенной выше, $v(\emptyset) = 0$.

Лемма о супераддитивности. Для бескоалиционной игры $\Gamma = (N, \{X_i\}_{i \in N}, \{H_i\}_{i \in N})$, построим функцию $v(S)$ следующим образом:

$$v(S) = \sup_{\mu_S} \inf_{\nu_{N \setminus S}} K_S(\mu_S, \nu_{N \setminus S}), \quad S \subset N, \quad (3.11.2)$$

где $\mu_S \in \bar{X}_S$, $\nu_{N \setminus S} \in \bar{X}_{N \setminus S}$ и $\bar{\Gamma}_S = (\bar{X}_S, \bar{X}_{N \setminus S}, K_S)$ — смешанное расширение антагонистической игры Γ_S . Тогда для всех $S, T \subset N$, для которых $S \cap T = \emptyset$, имеет место неравенство:

$$v(S \cup T) \geq v(S) + v(T). \quad (3.11.3)$$

Доказательство. Заметим, что

$$v(S \cup T) = \sup_{\mu_{S \cup T}} \inf_{\nu_{N \setminus (S \cup T)}} \sum_{i \in S \cup T} K_i(\mu_{S \cup T}, \nu_{N \setminus (S \cup T)}),$$

где $\mu_{S \cup T}$ — смешанные стратегии коалиции $S \cup T$, т. е. произвольные вероятностные меры на $X_{S \cup T}$, $\nu_{N \setminus (S \cup T)}$ — вероятностные меры на $X_{N \setminus (S \cup T)}$, K_i — выигрыш игрока i в смешанных стратегиях. Если ограничиться только такими вероятностными мерами на $X_{S \cup T}$, которые являются произведениями независимых распределений μ_S и μ_T на декартовом произведении $X_S \times X_T$, то область изменения переменной, по которой производится максимизация, сузится и супремум разве лишь уменьшится. Таким образом, имеем

$$v(S \cup T) \geq \sup_{\mu_S} \sup_{\mu_T} \inf_{\nu_{N \setminus (S \cup T)}} \sum_{i \in S \cup T} K_i(\mu_S \times \mu_T, \nu_{N \setminus (S \cup T)}).$$

Следовательно,

$$\begin{aligned} v(S \cup T) &\geq \inf_{\nu_{N \setminus (S \cup T)}} \sum_{i \in S \cup T} K_i(\mu_S \times \mu_T, \nu_{N \setminus (S \cup T)}) = \\ &= \inf_{\nu_{N \setminus (S \cup T)}} \left(\sum_{i \in S} K_i(\mu_S \times \mu_T, \nu_{N \setminus (S \cup T)}) + \sum_{i \in T} K_i(\mu_S \times \mu_T, \nu_{N \setminus (S \cup T)}) \right). \end{aligned}$$

Так как сумма инфимумов не превосходит инфимум суммы, имеем

$$v(S \cup T) \geq \inf_{\nu_{N \setminus (S \cup T)}} \sum_{i \in S} K_i(\mu_S \times \mu_T, \nu_{N \setminus (S \cup T)}) + \inf_{\nu_{N \setminus (S \cup T)}} \sum_{i \in T} K_i(\mu_S \times \mu_T, \nu_{N \setminus (S \cup T)}).$$

Минимизация первого слагаемого в правой части неравенства по μ_T , а второго — по μ_S (для единообразия переименуем их соответственно ν_T и ν_S), приводит к соотношениям

$$\begin{aligned} v(S \cup T) &\geq \inf_{\nu_T} \inf_{\nu_{N \setminus (S \cup T)}} \sum_{i \in S} K_i(\mu_S \times \nu_T, \nu_{N \setminus (S \cup T)}) + \inf_{\nu_S} \inf_{\nu_{N \setminus (S \cup T)}} \sum_{i \in T} K_i(\nu_S \times \mu_T, \nu_{N \setminus (S \cup T)}) \\ &\geq \inf_{\nu_{N \setminus S}} \sum_{i \in S} K_i(\mu_S, \nu_{N \setminus S}) + \inf_{\nu_{N \setminus T}} \sum_{i \in T} K_i(\mu_T, \nu_{N \setminus T}). \end{aligned}$$

Последнее неравенство справедливо при любых значениях мер μ_S в первом слагаемом и μ_T — во втором. Следовательно, по этим мерам можно перейти к супремумам

$$v(S \cup T) \geq \sup_{\mu_S} \inf_{\nu_{N \setminus S}} \sum_{i \in S} K_i(\mu_S, \nu_{N \setminus S}) + \sup_{\mu_T} \inf_{\nu_{N \setminus T}} \sum_{i \in T} K_i(\mu_T, \nu_{N \setminus T}),$$

откуда, используя (3.11.2), получаем

$$v(S \cup T) \geq v(S) + v(T).$$

Таким образом, супераддитивность доказана.

Заметим, что неравенство (3.11.3) также справедливо, если функция $v(S)$ строится по правилу

$$v(S) = \sup_{x_S} \inf_{x_{N \setminus S}} H_S(x_S, x_{N \setminus S}), \quad S \subset N,$$

где $x_S \in X_S$, $x_{N \setminus S} \in X_{N \setminus S}$, $\Gamma_S = (X_S, X_{N \setminus S}, H_S)$, при этом доказательство дословно повторяет приведенное выше, т. е. если $X_S, X_{N \setminus S}$ — множества чистых стратегий коалиций S и $N \setminus S$ соответственно.

3.11.3. Определение. Бескоалиционная игра $\Gamma = (N, \{X_i\}_{i \in N}, \{H_i\}_{i \in N})$ называется игрой с постоянной суммой, если

$$\sum_{i \in N} H_i(x) = c = \text{const}$$

для всех $x \in X_N$, $X_N = \prod_{i \in N} X_i$.

Лемма. Пусть $\Gamma = (N, \{X_i\}_{i \in N}, \{H_i\}_{i \in N})$ — бескоалиционная игра с постоянной суммой, функция $v(S)$, $S \subset N$, определена, как в лемме п. 3.11.2, а игры Γ_S , $S \subset N$, имеют значения в смешанных стратегиях. Тогда

$$v(N) = v(S) + v(N \setminus S), \quad S \subset N.$$

Доказательство. Из определения игры с постоянной суммой получаем, что

$$v(N) = \sum_{i \in N} H_i(x) = \sum_{i \in N} K_i(\mu) = c$$

для всех ситуаций x в чистых и μ — в смешанных стратегиях. С другой стороны,

$$\begin{aligned} v(S) &= \sup_{\mu_S} \inf_{\nu_{N \setminus S}} \sum_{i \in S} K_i(\mu_S, \nu_{N \setminus S}) = \sup_{\mu_S} \inf_{\nu_{N \setminus S}} \left(c - \sum_{i \in N \setminus S} K_i(\mu_S, \nu_{N \setminus S}) \right) = \\ &= c - \inf_{\nu_{N \setminus S}} \sup_{\mu_S} \sum_{i \in N \setminus S} K_i(\mu_S, \nu_{N \setminus S}) = c - v(N \setminus S), \end{aligned}$$

что и требовалось доказать.

3.11.4. В дальнейшем под кооперативной игрой будем понимать просто пару (N, v) , где v — характеристическая функция, удовлетворяющая неравенству (3.11.1), поскольку содержательная интерпретация характеристической функции, обосновывающая свойство (3.11.1), не имеет принципиального значения.

Пример 20 (Игра «джаз-оркестр») [Мулен, 1985]. Директор клуба обещает 100 усл.ед. певцу S , пианисту P и ударнику D за совместное выступление. Дуэт певца и пианиста он оценивает в 80 усл. ед., ударника и пианиста в 65 усл. ед. и одного пианиста — в 30 усл. ед. Другие дуэты и солисты не рассматриваются, поскольку присутствие фортепиано директор клуба считает обязательным. Дуэт певец — ударник зарабатывает 50 усл. ед., а певец — в среднем 20 усл. ед. за вечер. Ударник один ничего не может заработать.

Обозначая цифрами 1, 2, 3 игроков S , P и D соответственно, мы имеем дело с кооперативной игрой (N, v) , где $N = \{1, 2, 3\}$, $v(1, 2, 3) = 100$, $v(1, 3) = 50$, $v(1) = 20$, $v(1, 2) = 80$, $v(2, 3) = 65$, $v(2) = 30$, $v(3) = 0$.

Основная задача кооперативной теории игр n лиц заключается в построении реализуемых принципов оптимального распределения максимального суммарного выигрыша $v(N)$ между игроками.

Пусть α_i — сумма, которую получает игрок i при распределении максимального суммарного выигрыша $v(N)$, $N = \{1, 2, \dots, n\}$.

Определение 1. Вектор $\alpha = (\alpha_1, \dots, \alpha_n)$, удовлетворяющий условиям

$$\alpha_i \geq v(\{i\}), \quad i \in N, \quad (3.11.4)$$

$$\sum_{i=1}^n \alpha_i = v(N), \quad (3.11.5)$$

где $v(\{i\})$ — значение характеристической функции для одноэлементной коалиции $S = \{i\}$, называется *дележом*.

Условие (3.11.4) называется условием индивидуальной рациональности и означает, что, участвуя в коалиции, каждый игрок получает по меньшей мере столько, сколько он мог бы получить, действуя самостоятельно и не заботясь о поддержке каких-либо других игроков. Должно также выполняться условие (3.11.5), так как в случае $\sum_{i \in N} \alpha_i < v(N)$ существует распределение α' , при котором каждый игрок $i \in N$ получает больше, чем его доля α_i . Если же $\sum_{i \in N} \alpha_i > v(N)$, то игроки из N делят между

собой нереализуемый выигрыш, и поэтому вектор α неосуществим. Следовательно, вектор α может считаться допустимым только при выполнении условия (3.11.5), которое называется условием коллективной (или групповой) рациональности.

На основании условий (3.11.4), (3.11.5), для того, чтобы вектор $\alpha = (\alpha_1, \dots, \alpha_n)$ был дележом в кооперативной игре (N, v) , необходимо и достаточно выполнение равенства

$$\alpha_i = v(\{i\}) + \gamma_i, \quad i \in N,$$

причем

$$\gamma_i \geq 0, \quad i \in N, \quad \sum_{i \in N} \gamma_i = v(N) - \sum_{i \in N} v(\{i\}).$$

Определение 2. Игра (N, v) называется *существенной*, если

$$\sum_{i \in N} v(\{i\}) < v(N), \quad (3.11.6)$$

в противном случае игра называется *несущественной*.

Для любого дележа α будем обозначать величину $\sum_{i \in S} \alpha_i$ через $\alpha(S)$, а множество всех дележей — через D . Несущественная игра имеет единственный дележ $\alpha = (v(\{1\}), v(\{2\}), \dots, v(\{n\}))$.

Во всякой существенной игре с более чем одним игроком множество дележей бесконечно. Поэтому будем анализировать такие игры с помощью отношения доминирования.

Определение 3. Дележ α доминирует дележ β по коалиции S (обозначение $\alpha \overset{S}{\succ} \beta$), если

$$\alpha_i > \beta_i, \quad i \in S, \quad \alpha(S) = \sum_{i \in S} \alpha_i \leq v(S). \quad (3.11.7)$$

Первое из условий в определении (3.11.7) означает, что дележ α лучше дележа β для всех членов коалиции S , а второе отражает реализуемость дележа α коалицией S (т. е. коалиция S на самом деле может предложить каждому из игроков $i \in S$ величину α_i).

Определение 4. Говорят, что дележ α доминирует дележ β , если существует коалиция S , для которой $\alpha \overset{S}{\succ} \beta$. Доминирование дележа β дележом α обозначается как $\alpha \succ \beta$.

Доминирование невозможно по одноэлементной коалиции и множеству всех игроков N . Действительно, из $\alpha \overset{i}{\succ} \beta$ следовало бы $\beta_i < \alpha_i \leq v(\{i\})$, что противоречит условию (3.11.5).

3.11.5. Объединение кооперативных игр в те или иные классы существенно упрощает их последующее рассмотрение. В качестве таких классов можно рассмотреть классы эквивалентных игр.

Определение. Кооперативная игра (N, v) называется *эквивалентной* игре (N, v') , если существует положительное число k , а также n таких произвольных вещественных чисел c_i , $i \in N$, что для любой коалиции $S \subset N$ выполняется равенство

$$v'(S) = kv(S) + \sum_{i \in S} c_i. \quad (3.11.8)$$

Эквивалентность игр (N, v) и (N, v') будем обозначать как $(N, v) \sim (N, v')$ или $v \sim v'$.

Очевидно, что $v \sim v$. Чтобы убедиться в этом, достаточно положить в формуле (3.11.8) $c_i = 0$, $k = 1$, $v' = v$. Такое свойство называется рефлексивностью.

Докажем симметрию отношения, т. е. что из условия $v \sim v'$ следует $v' \sim v$. Действительно, полагая $k' = 1/k$, $c'_i = -c_i/k$, получим

$$v(S) = k'v'(S) + \sum_{i \in S} c'_i,$$

т. е. $v' \sim v$.

Наконец, если $v \sim v'$ и $v' \sim v''$, то $v \sim v''$. Это свойство называется транзитивностью. Оно проверяется последовательным применением формулы (3.11.8).

Так как отношение эквивалентности рефлексивно, симметрично и транзитивно, оно разбивает множество всех игр n лиц на взаимонепересекающиеся классы эквивалентных игр.

Теорема. Если две игры v и v' эквивалентны, то отображение $\alpha \rightarrow \alpha'$, где

$$\alpha'_i = k\alpha_i + c_i, \quad i \in N,$$

устанавливает также взаимно однозначное отображение множества всех дележей игры v на множество дележей игры v' , так что из $\alpha \succ^S \beta$ следует $\alpha' \succ^S \beta'$.

Доказательство. Проверим, что α' является дележом в игре (N, v') . Действительно,

$$\alpha'_i = k\alpha_i + c_i \geq kv(\{i\}) + c_i = v(\{i\}),$$

$$\sum_{i \in N} \alpha'_i = \sum_{i \in N} (k\alpha_i + c_i) = kv(N) + \sum_{i \in N} c_i = v'(N).$$

Следовательно, для α' условия (3.11.4), (3.11.5) выполнены. Далее, если $\alpha \succ^S \beta$, то

$$\alpha_i > \beta_i, \quad i \in S, \quad \sum_{i \in S} \alpha_i \leq v(S),$$

и поэтому

$$\alpha'_i = k\alpha_i + c_i > k\beta_i + c_i = \beta'_i \quad (k > 0),$$

$$\sum_{i \in S} \alpha'_i = k \sum_{i \in S} \alpha_i + \sum_{i \in S} c_i \leq kv(S) + \sum_{i \in S} c_i = v'(S),$$

т. е. $\alpha' \succ^S \beta'$. Взаимная однозначность соответствия следует из существования обратного отображения (оно было использовано при доказательстве симметрии отношения эквивалентности). Теорема доказана.

3.11.6. При разбиении множества кооперативных игр на попарно непересекающиеся классы эквивалентности возникает задача выбора наиболее простых представителей из каждого класса.

Определение. Игра (N, v) называется игрой в $(0-1)$ – редуцированной форме, если для всех $i \in N$

$$v(\{i\}) = 0, \quad v(N) = 1.$$

Теорема. Каждая существенная кооперативная игра эквивалентна некоторой игре в $(0-1)$ -редуцированной форме.

Доказательство. Пусть

$$k = \frac{1}{v(N) - \sum_{i \in N} v(\{i\})} > 0,$$

$$c_i = -\frac{v(\{i\})}{v(N) - \sum_{i \in N} v(\{i\})}, \quad v'(S) = kv(S) + \sum_{i \in S} c_i.$$

Тогда $v'(\{i\}) = 0$, $v'(N) = 1$. Теорема доказана.

Из теоремы следует, что свойства игр, включающие понятие доминирования, можно изучать на играх в $(0-1)$ -редуцированной форме. Если v — характеристическая функция произвольной существенной игры (N, v) , то

$$v'(S) = \frac{v(S) - \sum_{i \in S} v(\{i\})}{v(N) - \sum_{i \in N} v(\{i\})}, \quad (3.11.9)$$

есть $(0-1)$ -нормализация, соответствующая функции v . При этом дележом оказывается любой вектор $\alpha = (\alpha_1, \dots, \alpha_n)$, компоненты которого удовлетворяют условиям

$$\alpha_i \geq 0, \quad i \in N, \quad \sum_{i \in N} \alpha_i = 1, \quad (3.11.10)$$

т. е. дележи можно рассматривать как точки $(n-1)$ -мерного симплекса, порожденного ортами $w_j = (0, \dots, 0, 1, 0, \dots, 0)$, $j = \overline{1, n}$ пространства R^n .

§ 3.12. C -ядро и NM -решение

Перейдем к рассмотрению принципов оптимального поведения в кооперативных играх. Как уже отмечалось в п. 3.11.4, речь будет идти о принципах оптимального распределения максимального суммарного выигрыша между игроками.

3.12.1. Возможен следующий подход. Пусть игроки в кооперативной игре (N, v) пришли к такому соглашению о распределении выигрыша всей коалиции N (дележу α^*), при котором ни один из дележей не доминирует α^* . Тогда такое распределение устойчиво в том смысле, что ни одной из коалиций S невыгодно отделиться от других игроков и распределить между членами коалиции выигрыш $v(S)$. Это рассуждение наводит на мысль о целесообразности рассмотрения множества недоминируемых дележей.

Определение. Множество недоминируемых дележей кооперативной игры (N, v) называется ее C -ядром.

Имеет место следующая теорема, которая характеризует C -ядро.

Теорема. Для того чтобы дележ α принадлежал C -ядру, необходимо и достаточно выполнение для всех $S \subset N$ неравенств

$$v(S) \leq \alpha(S) = \sum_{i \in S} \alpha_i \quad (3.12.1)$$

Доказательство. Для несущественных игр теорема очевидна, и в силу теоремы п. 3.11.6 достаточно провести ее доказательство для игр в $(0-1)$ -редуцированной форме.

Докажем достаточность утверждения теоремы. Пусть для дележа α выполнено условие (3.12.1). Покажем, что дележ α принадлежит C -ядру. Пусть это не так. Тогда найдется такой дележ β , что $\beta \succ^S \alpha$, т. е. $\beta(S) > \alpha(S)$ и $\beta(S) \leq v(S)$. Однако это противоречит (3.12.1).

Покажем необходимость условия (3.12.1). Для любого дележа α , не удовлетворяющего (3.12.1), существует такая коалиция S , что $\alpha(S) < v(S)$. Пусть S — такая коалиция, тогда построим вектор β по следующему правилу:

$$\beta_i = \alpha_i + \frac{v(S) - \alpha(S)}{|S|}, \quad i \in S,$$

$$\beta_i = \frac{1 - v(S)}{|N| - |S|}, \quad i \notin S,$$

где $|S|$ — число элементов множества S . Легко видеть, что $\beta(N) = 1$, $\beta_i \geq 0$ и $\beta \succ^S \alpha$. Отсюда следует, что α не принадлежит C -ядру.

Из теоремы п. 3.12.1 следует, что C -ядро является замкнутым, выпуклым подмножеством множества всех дележей (C -ядро может быть пустым множеством).

3.12.2. Пусть игроки договариваются о выборе кооперативного соглашения. Из супераддитивности v следует, что такое соглашение приводит к образованию коалиции N всех игроков. Решается вопрос о способе дележа суммарного дохода $v(N)$, т. е. о выборе вектора $\alpha \in R^n$, для которого $\sum_{i \in N} \alpha_i = v(N)$.

Минимальным требованием для получения согласия игроков выбрать вектор α является индивидуальная рациональность этого вектора, т. е. условие $\alpha_i \geq v(\{i\})$, $i \in N$. Пусть игроки договариваются о выборе конкретного дележа α . Против выбора дележа может возражать некоторая коалиция S , требующая для себя более выгодного распределения. Коалиция S выдвигает это требование, угрожая в противном случае нарушить общую кооперацию (это вполне реальная угроза, так как для достижения дохода $v(N)$ требуется единодушное согласие всех игроков). Предположим, что остальные игроки $N \setminus S$ реагируют на эту угрозу объединенными действиями против коалиции S . Тогда максимальный гарантированный доход коалиции S оценивается числом $v(S)$. Условие (3.12.1) означает существование стабилизирующей угрозы коалиции S со стороны коалиции $N \setminus S$. Таким образом, C -ядром игры (N, v) является множество устойчивых в смысле коалиционных угроз распределений максимального суммарного дохода $v(N)$.

Приведем еще один критерий принадлежности дележа C -ядру.

Лемма. Пусть α — дележ игры (N, v) . Тогда α принадлежит C -ядру в том и только в том случае, когда для всех коалиций $S \subset N$ выполняется неравенство

$$\sum_{i \in S} \alpha_i \leq v(N) - v(N \setminus S). \quad (3.12.2)$$

Доказательство. Поскольку $\sum_{i \in N} \alpha_i = v(N)$, то приведенное выше неравенство можно записать в виде

$$v(N \setminus S) \leq \sum_{i \in N \setminus S} \alpha_i.$$

Теперь утверждение леммы следует из (3.12.1).

Из условия (3.12.1) видно, что если дележ α принадлежит C -ядру, то ни одна коалиция S не может гарантировать себе выигрыш, превосходящий $\sum_{i \in S} \alpha_i = \alpha(S)$, т. е.

суммарный выигрыш, который обеспечивается членам коалиции дележом α . Это делает нецелесообразным существование коалиций S , отличных от максимальной коалиции N .

Интересное необходимое и достаточное условие непустоты C -ядра было получено О. Н. Бондаревой (см. [Бондарева, 1963]), которое теперь носит название теоремы Бондаревой–Шепли.

Во многих случаях C -ядро может оказаться пустым, а в других случаях оно представляет собой множественный принцип оптимальности, и остается вопрос, какой дележ из C -ядра необходимо выбрать в конкретном случае.

Пример 21. Рассмотрим игру «джаз-оркестр» (см. пример 20 п. 3.11.4). Суммарный доход трех музыкантов максимален (и равен 100 руб.) в случае их совместного выступления. Если певец выступает отдельно от пианиста с ударником, то все втроем они получают 65 + 20 руб., если пианист выступает один, то 30 + 50 руб. Наконец, суммарный доход равен 80 руб., если пианист и певец отказываются от участия ударника. Какое распределение максимального общего дохода следует признать разумным, учитывая описанные возможности игроков в смысле частичной кооперации и индивидуального поведения?

Вектор $\alpha = (\alpha_1, \alpha_2, \alpha_3)$ в игре «джаз-оркестр» принадлежит C -ядру тогда и только тогда, когда

$$\begin{cases} \alpha_1 \geq 20, \alpha_2 \geq 30, \alpha_3 \geq 0, \\ \alpha_1 + \alpha_2 + \alpha_3 = 100, \\ \alpha_1 + \alpha_2 \geq 80, \alpha_2 + \alpha_3 \geq 65, \alpha_1 + \alpha_3 \geq 50. \end{cases}$$

Это множество является выпуклой оболочкой следующих трех дележей: (35, 45, 20), (35, 50, 15), (30, 50, 20). Таким образом, выигрыши всех игроков определяются с точностью до 5 руб. Типичным представителем ядра является центр (среднеарифметическое крайних точек) C -ядра, а именно: $\alpha^* = (33.3, 48.3, 18.3)$. Для дележа α^* характерно, что все двуэлементные коалиции имеют одинаковый дополнительный доход: $\alpha_i + \alpha_j - v(\{i, j\}) = 1.6$. Дележ α^* является справедливым компромиссом внутри C -ядра.

3.12.3. Из того, что C -ядро пусто, не следует невозможность кооперации всех игроков N . Это просто означает, что никакой дележ не может быть стабилизирован с помощью простых угроз, описанных выше. Пустота ядра имеет место тогда, когда промежуточные коалиции слишком сильны. Это утверждение поясняется следующим образом.

Пример 22 (Симметричные игры) [Мулен, 1985]. В симметричной игре коалиции с одинаковым числом игроков имеют одинаковый выигрыш. Характеристическая функция v имеет следующий вид:

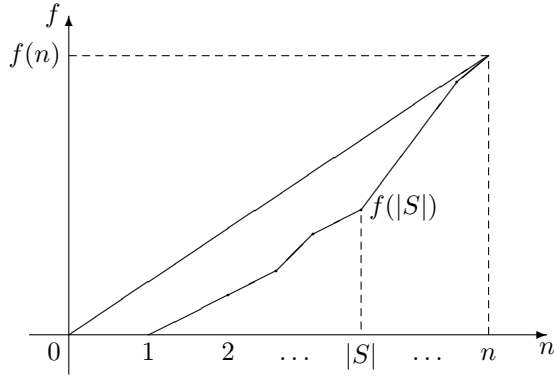
$$v(S) = f(|S|)$$

для всех $S \subset N$, где $|S|$ — число элементов множества S .

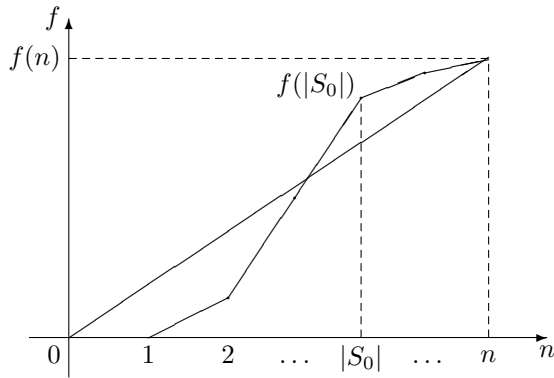
Предположим без потери общности, что $f(1) = 0$ и $N = \{1, \dots, n\}$. Тогда множеством дележей игры (N, v) является следующий симплекс в R^n

$$\sum_{i=1}^n \alpha_i = f(n) = v(N), \alpha_i \geq 0, i = 1, \dots, n.$$

C -ядром является подмножество множества дележей, определенное линейными неравенствами (3.12.1), т. е. это выпуклый многогранник. В силу симметричности $v(S)$ C -ядро также симметрично, т. е. инвариантно относительно любой перестановки компонент $\alpha_1, \dots, \alpha_n$. Учитывая, кроме того, выпуклость C -ядра, можно показать, что оно

Рис. 3.8. Непустое C -ядро

не пусто в том и только в том случае, когда содержит центр α^* множества всех дележей ($\alpha_i^* = f(n)/n$, $i = 1, \dots, n$). Возвращаясь к системе (3.12.1), получаем, что C -ядро не пусто тогда и только тогда, когда для всех $|S| = 1, \dots, n$ имеет место неравенство $(1/|S|)f(|S|) \leq (1/n)f(n)$. Таким образом, C -ядро не пусто тогда и только тогда, когда

Рис. 3.9. Пустое C -ядро

не существует промежуточной коалиции S , в которой средняя доля каждого игрока больше соответствующей величины в коалиции N . Рис. 3.8(3.9) соответствует случаю, когда C -ядро не пусто (пусто).

3.12.4. Пример 23 [Воробьев, 1985]. Рассмотрим общую игру трех лиц в $(0 - 1)$ -редуцированной форме. Для ее характеристической функции имеем $v(\emptyset) = v(1) = v(2) = v(3) = 0$, $v(1, 2, 3) = 1$, $v(1, 2) = c_3$, $v(1, 3) = c_2$, $v(2, 3) = c_1$, где $0 \leq c_i \leq 1$, $i = 1, 2, 3$.

На основании теоремы п. 3.9.1, чтобы дележ α принадлежал C -ядру, необходимо и достаточно выполнение следующих неравенств:

$$\alpha_1 + \alpha_2 \geq c_3, \quad \alpha_1 + \alpha_3 \geq c_2, \quad \alpha_2 + \alpha_3 \geq c_1$$

или

$$\alpha_3 \leq 1 - c_3, \quad \alpha_2 \leq 1 - c_2, \quad \alpha_1 \leq 1 - c_1. \quad (3.12.3)$$

Складывая неравенства (3.12.3), получаем

$$\alpha_1 + \alpha_2 + \alpha_3 \leq 3 - (c_1 + c_2 + c_3),$$

или, поскольку сумма всех α_i , $i = 1, 2, 3$, тождественно равна единице,

$$c_1 + c_2 + c_3 \leq 2. \quad (3.12.4)$$

Последнее неравенство является необходимым условием существования в рассматриваемой игре непустого C -ядра. С другой стороны, если (3.12.4) выполняется, то существуют такие неотрицательные ξ_1, ξ_2, ξ_3 , что

$$\sum_{i=1}^3 (c_i + \xi_i) = 2, \quad c_i + \xi_i \leq 1, \quad i = 1, 2, 3.$$

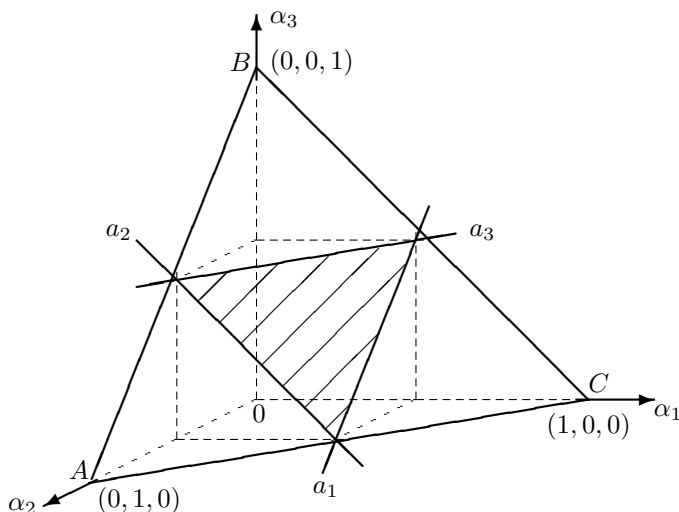


Рис. 3.10. C -ядро игры

Пусть $\beta_i = 1 - c_i - \xi_i$, $i = 1, 2, 3$. Числа β_i удовлетворяют неравенствам (3.12.3), так что дележ $\beta = (\beta_1, \beta_2, \beta_3)$ принадлежит C -ядру игры. Таким образом, соотношение (3.12.4) является также достаточным для существования непустого C -ядра. Геометрически множество дележей в рассматриваемой игре есть симплекс: $\alpha_1 + \alpha_2 + \alpha_3 = 1$, $\alpha_i \geq 0$, $i = 1, 2, 3$ (треугольник ABC , рис. 3.10). Непустое C -ядро представляет собой

пересечение множества дележей ($\triangle ABC$) и выпуклого многогранника (параллелепипеда) $0 \leq \alpha_i \leq 1 - c_i$, $i = 1, 2, 3$. Это часть треугольника ABC , вырезаемая линиями пересечения плоскостей

$$\alpha_i = 1 - c_i, \quad i = 1, 2, 3 \quad (3.12.5)$$

с плоскостью $\triangle ABC$. На рис. 3.10 через α_i , $i = 1, 2, 3$ обозначена прямая, образованная пересечением плоскостей $\alpha_i = 1 - c_i$ и $\alpha_1 + \alpha_2 + \alpha_3 = 1$. Точка пересечения двух прямых α_i и α_j принадлежит треугольнику ABC , если неотрицательна k -я координата этой точки ($k \neq i$, $k \neq j$), в противном случае она находится за пределами $\triangle ABC$ (рис. 3.11а, 3.11б). Таким образом, C -ядро имеет вид треугольника, если совместное решение любой пары уравнений (3.12.5) и уравнения $\alpha_1 + \alpha_2 + \alpha_3 = 1$ состоит из неотрицательных чисел. Это требование выполняется при

$$c_1 + c_2 \geq 1, \quad c_1 + c_3 \geq 1, \quad c_2 + c_3 \geq 1. \quad (3.12.6)$$

В зависимости от различных случаев (а всего их может быть восемь) C -ядро будет приобретать тот или иной вид. Например, если не выполняется ни одно из трех неравенств (3.12.6), то C -ядро оказывается шестиугольником (рис. 3.11б).

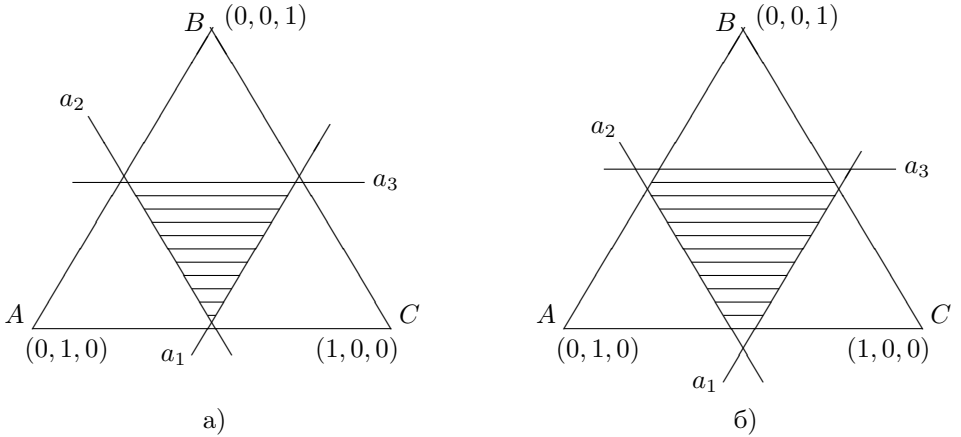


Рис. 3.11. Виды C -ядра

3.12.5. Другим принципом оптимальности в кооперативных играх является NM -решение. NM -решение, так же как и C -ядро, является множественным принципом оптимальности в множестве всех дележей. Хотя элементы C -ядра и не доминируются никакими другими дележами, однако нельзя утверждать, что в C -ядре для любого наперед заданного дележа α найдется доминирующий его дележ. Поэтому оказывается целесообразной формулировка принципа оптимальности, который бы учитывал и это последнее обстоятельство.

Определение. Подмножество дележей L кооперативной игры (N, v) называется NM -решением, если:

- 1) из $\alpha \succ \beta$ следует, что либо $\alpha \notin L$ либо $\beta \notin L$ (внутренняя устойчивость);
- 2) для любого $\alpha \notin L$ существует такой дележ $\beta \in L$, что $\beta \succ \alpha$ (внешняя устойчивость).

К сожалению, применение понятия NM -решения на практике не легко. Оно несет скорее философский, нежели практический смысл.

Между C -ядром кооперативной игры и ее NM -решением имеется известная связь. Например, если C -ядро не пусто и NM -решение существует, то оно содержит C -ядро. Действительно, пусть дележ α принадлежит C -ядру; тогда, если бы он не принадлежал NM -решению L , то согласно свойству 2) нашелся бы такой дележ α' , что $\alpha' \succ \alpha$. Однако это противоречит принадлежности α C -ядру как множеству недоминируемых дележей.

Теорема. Если для характеристической функции игры (N, v) в $(0-1)$ -редуцированной форме $(|N| = n)$ выполняются неравенства

$$v(S) \leq \frac{1}{n - |S| + 1},$$

где $|S|$ — число игроков в коалиции S , то C -ядро этой игры не пусто и является ее NM -решением.

Доказательство. Возьмем произвольный дележ α , лежащий вне C -ядра. Тогда существует непустое множество коалиций S , по которым можно доминировать α , т. е. это те и только те коалиции, для которых $\alpha(S) < v(S)$. Множество $\{S\}$ частично упорядочено по включению, т. е. $S_1 \supset S_2$, если $S_2 \subset S_1$. Возьмем в нем какой-нибудь минимальный элемент S_0 , который, очевидно, существует. Пусть k — число игроков в коалиции S_0 . Очевидно, что $2 \leq k \leq n - 1$. Построим дележ β следующим образом:

$$\beta_i = \begin{cases} \alpha_i + \frac{v(S_0) - \alpha(S_0)}{k}, & i \in S_0, \\ \frac{1 - v(S_0)}{n - k}, & i \notin S_0. \end{cases}$$

Так как $\beta(S_0) = v(S_0)$, $\beta_i > \alpha_i$, $i \in S_0$, то β доминирует α по коалиции S_0 . Докажем, что β содержится в C -ядре. Для этого достаточно показать, что $\beta(S) \geq v(S)$ при произвольном S . Пусть сначала $|S| \leq k$. Заметим, что β не доминируется по $S \subset S_0$, поскольку $\beta_i > \alpha_i$ ($i \in S_0$) и не может доминироваться ни по какой коалиции, так как S_0 — минимальная коалиция, по которой можно доминировать α . Если же хоть один игрок из S не содержится в S_0 , то

$$\beta(S) \geq \frac{1 - v(S_0)}{n - k} \geq \frac{1 - \frac{1}{n - k + 1}}{n - k} = \frac{1}{n - k + 1} \geq \frac{1}{n - |S| + 1} \geq v(S).$$

Таким образом, β не доминируется ни по какой коалиции, содержащей не более k игроков.

Пусть теперь $|S| > k$. Если $S_0 \subset S$, то

$$\begin{aligned} \beta(S) &= \frac{(|S| - k)(1 - v(S_0))}{n - k} + v(S_0) \geq \frac{|S| - k}{n - k} \\ &\geq \frac{|S| - k + k - |S| + 1}{n - k + k - |S| + 1} = \frac{1}{n - |S| + 1} \geq v(S). \end{aligned}$$

Если же S не содержит S_0 , то число игроков множества S , не содержащихся в S_0 , не меньше $|S| - k + 1$; следовательно,

$$\beta(S) \geq \frac{(|S| - k + 1)(1 - v(S_0))}{n - k} \geq \frac{|S| - k + 1}{n - k + 1} \geq \frac{1}{n - |S| + 1} \geq v(S).$$

Таким образом, β не доминируется ни по какой коалиции S . Следовательно, β содержится в C -ядре. Кроме того, β доминирует α . Итак, доказано, что C -ядро непусто и удовлетворяет свойству 2, характеризующему множество NM -решений. Свойству 1 C -ядро удовлетворяет автоматически в силу определения. Теорема доказана.

3.12.6. Определение. Игра (N, v) в $(0-1)$ -редуцированной форме называется простой, если для любых $S \subset N$ $v(S)$ принимает лишь одно из двух значений 0 или 1. Кооперативная игра называется простой, если проста ее $(0-1)$ -редуцированная форма.

Пример 24 [Воробьев, 1985]. Рассмотрим простую игру трех лиц в $(0-1)$ -редуцированной форме, в которой коалиция, состоящая из двух и трех игроков, выигрывает ($v(S) = 1$), а коалиция, включающая только одного игрока, проигрывает ($v(\{i\}) = 0$). Для этой игры рассмотрим три дележа:

$$\alpha_{12} = (1/2, 1/2, 0), \alpha_{13} = (1/2, 0, 1/2), \alpha_{23} = (0, 1/2, 1/2). \quad (3.12.7)$$

Ни один из этих трех дележей не доминирует никакого другого. Множество дележей (3.12.7) имеет и следующее свойство, любой дележ (кроме трех дележей α_{ij}) доминируется одним из дележей α_{ij} . Чтобы это проверить, рассмотрим какой-нибудь дележ $\alpha = (\alpha_1, \alpha_2, \alpha_3)$. Так как мы рассматриваем игру в $(0-1)$ -редуцированной форме, то $\alpha_i \geq 0$ и $\alpha_1 + \alpha_2 + \alpha_3 = 1$. Следовательно, не более двух компонент вектора α могут быть не меньше $1/2$. Если их действительно две, то каждая из них равна $1/2$, в то время как третья равна 0. Но это означает, что α совпадает с одним из α_{ij} . Если же α — какой-нибудь иной дележ, то он имеет не более одной компоненты, не меньшей чем $1/2$. Значит, по крайней мере две компоненты, например, α_i и α_j ($i < j$), меньше $1/2$.

Но в этом случае $\alpha_{ij} \succ^{ij} \alpha$. Таким образом, три дележа (3.12.7) образуют NM -решение. Но это не единственное NM -решение.

Пусть c — любое число из отрезка $[0, 1/2]$; легко проверить, что множество

$$L_{3,c} = \{(a, 1-c-a, c) \mid 0 \leq a \leq 1-c\}$$

также является NM -решением. Действительно, в это множество входят дележи, при которых игрок 3 получит постоянную c , а игроки 1 и 2 делят остаток во всевозможных пропорциях. Внутренняя устойчивость следует из того, что для любых двух дележей α и β из этого множества имеем: если $\alpha_1 > \beta_1$, то $\alpha_2 < \beta_2$. Однако доминирование по коалиции, состоящей из единственного участника, невозможно. Чтобы доказать внешнюю устойчивость $L_{3,c}$, возьмем какой-либо дележ $\beta \notin L_{3,c}$. Это означает, что либо $\beta_3 > c$ либо $\beta_3 < c$. Пусть $\beta_3 > c$, например, $\beta_3 = c + \varepsilon$. Определим дележ α следующим образом:

$$\alpha_1 = \beta_1 + \varepsilon/2, \alpha_2 = \beta_2 + \varepsilon/2, \alpha_3 = c.$$

Тогда $\alpha \in L_{3,c}$ и $\alpha \succ \beta$ по коалиции $\{1, 2\}$. Пусть теперь $\beta_3 < c$. Ясно, что либо $\beta_1 \leq 1/2$ либо $\beta_2 \leq 1/2$ (ибо в противном случае их сумма была бы больше 1). Пусть $\beta_1 \leq 1/2$. Положим $\alpha = (1-c, 0, c)$. Так как $1-c > 1/2 \geq \beta_1$, то $\alpha \succ \beta$ по коалиции $\{1, 3\}$. Очевидно, что $\alpha \in L_{3,c}$. Однако, если $\beta_2 \leq 1/2$, то можно аналогично показать, что $\gamma \succ \beta$, где $\gamma = (0, 1-c, c)$. Итак, кроме симметричного NM -решения, рассматриваемая игра имеет еще целое семейство решений, при которых игрок 3 получает фиксированное количество c из отрезка $0 \leq c \leq 1/2$. Эти NM -решения называются дискриминирующими; говорят, что игрок 3 при этом дискриминирован. В случае множества $L_{3,0}$ говорят, что игрок 3 полностью дискриминирован или исключен.

Из соображений симметрии очевидно, что существуют также два семейства NM -решений, $L_{1,c}$ и $L_{2,c}$, в которых дискриминируются игроки 1 и 2 соответственно.

Предшествующий пример показывает, что у игры может быть чрезвычайно много NM -решений. Совершенно неясно, какое из них следует выбрать. Когда же NM -решение выбрано, остается непонятным, какой из него выбрать дележ.

Более подробно с NM -решением и его свойствами можно ознакомиться в [Дюбин, Суздаль, 1981]. В теории кооперативных игр важную роль играют одноточечные решения. Это прежде всего вектор Шепли и его различные обобщения, N -ядро, различные виды пропорциональных решений и др. (см. [Печерский, Яновская, 2004]). Далее мы подробно остановимся на векторе Шепли.

§ 3.13. Вектор Шепли

3.13.1. Множественность рассмотренных ранее принципов оптимальности C -ядра и NM -решения в кооперативных играх, а также жесткие условия существования этих принципов стимулируют попытки поиска принципов оптимальности, существование и единственность которых были бы обеспечены в каждой кооперативной игре. К таким принципам оптимальности относится вектор Шепли. Вектор Шепли определяется аксиоматически.

Определение 1. *Носителем игры (N, v) называется такая коалиция T , что $v(S) = v(S \cap T)$ для любой коалиции $S \subset N$.*

Содержательно определение утверждает, что любой игрок, не принадлежащий носителю, является «болваном», т. е. не может ничего внести ни в какую коалицию.

Рассмотрим произвольную перестановку P упорядоченного множества игроков $N = \{1, 2, \dots, n\}$. С этой перестановкой связана подстановка π , т. е. такая взаимно однозначная функция $\pi : N \rightarrow N$, что для $i \in N$ значение $\pi(i) \in N$ представляет собой элемент из N , в который переходит $i \in N$ в перестановке P .

Определение 2. *Пусть (N, v) — игра n лиц, P — перестановка множества N , а π — соответствующая ей подстановка. Тогда через $(N, \pi v)$ обозначим такую игру (N, u) , что для любой коалиции $S \subset N$, $S = \{i_1, i_2, \dots, i_s\}$*

$$u(\{\pi(i_1), \pi(i_2), \dots, \pi(i_s)\}) = v(S).$$

По существу игра $(N, \pi v)$ отличается от игры (N, v) лишь тем, что в последней игроки поменялись ролями в соответствии с перестановкой P .

С помощью этих определений можно изложить аксиоматику Шепли. Сначала заметим, что так как кооперативные игры n лиц, в сущности, отождествляются с вещественными (характеристическими) функциями, то можно говорить о сумме двух или большего числа игр, а также о произведении игры на число.

3.13.2. Поставим в соответствие каждой кооперативной игре (N, v) вектор $\varphi(v) = (\varphi_1[v], \dots, \varphi_n[v])$, компоненты которого будем интерпретировать как выигрыши, полученные игроками в результате соглашения или решения арбитра. При этом будем считать, что указанное соответствие удовлетворяет следующим аксиомам.

Определение 3. *Аксиомы Шепли.*

1. Если S — любой носитель игры (N, v) , то

$$\sum_{i \in S} \varphi_i[v] = v(S).$$

2. Для любой подстановки π и $i \in N$

$$\varphi_{\pi(i)}[\pi v] = \varphi_i[v].$$

3. Если (N, u) и (N, v) — две любые кооперативные игры, то

$$\varphi_i[u + v] = \varphi_i[u] + \varphi_i[v].$$

Определение 4. Пусть φ — функция, ставящая в соответствие согласно аксиомам 1 — 3 каждой игре (N, v) вектор $\varphi[v]$. Тогда $\varphi[v]$ называется вектором значений или вектором Шепли игры (N, v) .

Оказывается, что этих аксиом достаточно для определения единственным образом значения для всех игр n лиц.

Теорема. Существует единственная функция φ , определенная для всех игр (N, v) и удовлетворяющая аксиомам 1—3.

3.13.3. Доказательство теоремы опирается на следующие результаты.

Лемма. Пусть для любой коалиции $S \subset N$ игра (N, w_S) определяется следующим образом:

$$w_S(T) = \begin{cases} 0, & S \not\subset T, \\ 1, & S \subset T. \end{cases} \quad (3.13.1)$$

Тогда для игры (N, w_S) аксиомы 1, 2 однозначно определяют вектор $\varphi[w_S]$:

$$\varphi_i[w_S] = \begin{cases} 1/s, & i \in S, \\ 0, & i \notin S, \end{cases} \quad (3.13.2)$$

где $s = |S|$ — число игроков в S .

Доказательство. Ясно, что S — носитель w_S , как и любое множество T , содержащее множество S . Тогда по аксиоме 1, если $S \subset T$, то

$$\sum_{i \in T} \varphi_i[w_S] = 1.$$

Но это означает, что $\varphi_i[w_S] = 0$ при $i \notin S$. Далее, если π — любая подстановка, которая переводит S в себя, то $\pi w_S = w_S$. Следовательно, в силу аксиомы 2 для любых $i, j \in S$ имеет место равенство $\varphi_i[w_S] = \varphi_j[w_S]$. Так как этих величин всего $s = |S|$, а сумма их равна 1, то $\varphi_i[w_S] = 1/s$, если $i \in S$.

Игра с характеристической функцией w_S , определяемой (3.13.1), называется *простой игрой n лиц*. Таким образом, лемма утверждает, что для простой игры (N, w_S) вектор Шепли определяется формулой (3.13.2). Вектор Шепли для игры (N, w_S) определяется единственным образом.

Следствие. Если $c \geq 0$, то

$$\varphi_i[cw_S] = \begin{cases} c/s, & i \in S, \\ 0, & i \notin S. \end{cases}$$

Доказательство очевидно. Таким образом, $\varphi[cw_S] = c\varphi[w_S]$ при $c \geq 0$.

Теперь покажем, что если $\sum_S c_S w_S$ является характеристической функцией, то

$$\varphi_i\left(\sum_S c_S w_S\right) = \sum_S \varphi_i(c_S w_S) = \sum_S c_S \varphi_i(w_S). \quad (3.13.3)$$

В случае $c_S \geq 0$ первое равенство в (3.13.3) постулируется аксиомой 3, второе следует из следствия. Далее, если u, v и $u - v$ — характеристические функции, то согласно аксиоме 3 имеем $\varphi[u - v] = \varphi[u] - \varphi[v]$. Отсюда следует справедливость (3.13.3) для любых c_S . Действительно, если $\sum_S c_S w_S$ — характеристическая функция, то

$$v = \sum_S c_S w_S = \sum_{S|c_S \geq 0} c_S w_S - \sum_{S|c_S < 0} (-c_S) w_S,$$

поэтому

$$\begin{aligned} \varphi[v] &= \varphi\left[\sum_{S|c_S \geq 0} c_S w_S\right] - \varphi\left[\sum_{S|c_S < 0} (-c_S) w_S\right] = \\ &= \sum_{S|c_S \geq 0} c_S \varphi[w_S] - \sum_{S|c_S < 0} (-c_S) \varphi[w_S] = \sum_S c_S \varphi[w_S]. \end{aligned}$$

3.13.4. Лемма. Пусть (N, v) — любая игра, тогда найдутся $2^n - 1$ вещественных чисел c_S , таких что

$$v = \sum_{S \subset N} c_S w_S, \quad (3.13.4)$$

где w_S определены (3.13.1), а суммирование ведется по всем подмножествам S множества N , исключая пустое множество. При этом представление (3.13.4) единственно.

Доказательство. Положим

$$c_S = \sum_{T|T \subset S} (-1)^{s-t} v(T) \quad (3.13.5)$$

(здесь t — число элементов в T). Покажем, что эти числа c_S удовлетворяют условиям леммы. Действительно, если U — произвольная коалиция, то

$$\begin{aligned} \sum_{S|S \subset U} c_S w_S(U) &= \sum_{S|S \subset U} c_S \\ &= \sum_{S|S \subset U} \left(\sum_{T|T \subset S} (-1)^{s-t} v(T) \right) = \sum_{T|T \subset U} \left[\sum_{S|T \subset S \subset U} (-1)^{s-t} \right] v(T). \end{aligned}$$

Рассмотрим теперь величину в квадратных скобках в последнем выражении. Для каждого значения s между t и u имеется C_{u-t}^{u-s} таких множеств S с s элементами, что $T \subset S \subset U$. Следовательно, выражение в скобках можно заменить следующим выражением:

$$\sum_{s=t}^u C_{u-t}^{u-s} (-1)^{s-t} = \sum_{s=t}^u C_{u-t}^{s-t} (-1)^{s-t},$$

но это биномиальное разложение $(1 - 1)^{u-t}$. Следовательно, для всех $t < u$ оно равно 0, а для $t = u$ равно 1. Поэтому для всех $U \subset N$

$$\sum_{S|S \subset N} c_S w_S(U) = v(U).$$

Докажем единственность представления (3.13.4). Любой характеристической функции v соответствует элемент пространства R^{2^n-1} . Действительно, упорядочим коалиции $T \subset U$. Тогда каждой непустой коалиции $T \subset U$ соответствует компонента вектора, равная $v(T)$. Эти векторы будем обозначать, как и функции, через v . Очевидно, что простейшим характеристическим функциям w_S соответствуют векторы, у которых компоненты равны либо нулю, либо единице. Докажем, что простейшие характеристические функции (точнее, соответствующие им векторы) линейно независимы. Действительно, пусть

$$\sum_{S \subset N} \lambda_S w_S(T) = 0 \text{ для всех } T \subset N.$$

Тогда для $T = \{i\}$ имеем $w_S(\{i\}) = 0$, если $S \neq \{i\}$, и $w_S(\{i\}) = 1$, если $S = \{i\}$. Следовательно, $\lambda_{\{i\}} = 0$ для всех $i \in N$. Продолжим доказательство методом индукции. Пусть $\lambda_S = 0$ для всех $S \subset T$, $S \neq T$. Покажем, что $\lambda_T = 0$. Действительно,

$$\sum_{S \subset N} \lambda_S w_S(T) = \sum_{S \subset T} \lambda_S w_S(T) = \lambda_T = 0.$$

Таким образом, мы имеем $2^n - 1$ линейно независимых вектора в R^{2^n-1} , поэтому любой вектор, а значит и любая характеристическая функция v единственным образом выражается в виде линейной комбинации (3.13.4) простейших характеристических функций w_S . Лемма доказана.

3.13.5. Перейдем к доказательству теоремы п. 3.13.2. Лемма п. 3.13.4 показывает, что любая игра может быть представлена в виде линейной комбинации игр w_S , причем представление (3.13.4) единственно. Согласно п. 3.13.3, функция $\varphi[v]$ единственным образом определяется соотношениями (3.13.2), (3.13.3).

Пусть (N, v) — произвольная игра. Получим теперь выражение для вектора $\varphi[v]$. Согласно п.3.13.3, 3.13.4,

$$\varphi_i[v] = \sum_{S|S \subset N} c_S \varphi_i[w_S] = \sum_{S|i \in N} c_S (1/s),$$

но c_S определены формулой (3.13.5). Подставляя (3.13.5) в это выражение, получаем

$$\varphi_i[v] = \sum_{S|i \in N} (1/s) \left[\sum_{T|T \subset S} (-1)^{s-t} v(T) \right] = \sum_{T|T \subset N} \left[\sum_{S|T \cup i \subset S \subset N} (-1)^{s-t} (1/s) v(T) \right].$$

Положим

$$\gamma_i(T) = \sum_{S|T \cup i \subset S \subset N} (-1)^{s-t} (1/s). \quad (3.13.6)$$

Если $i \notin T'$ и $T = T' \cup \{i\}$, то $\gamma_i(T') = -\gamma_i(T)$. Действительно, все члены в правой части (3.13.6) в обоих случаях одни и те же, и только $t = t' + 1$; следовательно, они отличаются лишь знаком. Таким образом, имеем

$$\varphi_i[v] = \sum_{T|i \in T \subset N} \gamma_i(T) [v(T) - v(T \setminus \{i\})].$$

Далее, если $i \in T$, то существует ровно C_{n-t}^{s-t} таких коалиций S с s элементами, что $T \subset S$. В результате получаем хорошо известный определенный интеграл:

$$\begin{aligned}\gamma_i(T) &= \sum_{s=t}^n (-1)^{s-t} C_{n-t}^{s-t} (1/s) = \sum_{s=t}^n (-1)^{s-t} C_{n-t}^{s-t} \int_0^1 x^{s-1} dx = \\ &= \int_0^1 \sum_{s=t}^n (-1)^{s-t} C_{n-t}^{s-t} x^{s-1} dx = \int_0^1 x^{t-1} \sum_{s=t}^n (-1)^{s-t} C_{n-t}^{s-t} x^{s-t} dx = \\ &= \int_0^1 x^{t-1} (1-x)^{n-t} dx.\end{aligned}$$

Таким образом, имеем (бета-функция)

$$\gamma_i(T) = (t-1)!(n-t)!/(n!)$$

и, следовательно,

$$\varphi_i[v] = \sum_{T|i \in T \subset N} \frac{(t-1)!(n-t)!}{n!} [v(T) - v(T \setminus \{i\})]. \quad (3.13.7)$$

Формула (3.13.7) определяет компоненты вектора Шепли в явном виде. Это выражение удовлетворяет аксиомам 1–3 п. 3.13.2.

Заметим, кроме того, что вектор $\varphi[v]$ всегда является дележом. Действительно, в силу супераддитивности функции v ,

$$\begin{aligned}\varphi_i[v] &\geq v(\{i\}) \sum_{T|i \in T \subset N} \frac{(t-1)!(n-t)!}{n!} = \\ &= v(\{i\}) \sum_{t=1}^n C_{n-1}^{t-1} \frac{(t-1)!(n-t)!}{n!} = v(\{i\}).\end{aligned}$$

3.13.6. Если отвлечься от аксиоматического определения, то вектору Шепли, выраженный формулой (3.13.7), можно дать следующее содержательное истолкование. Предположим, что игроки (элементы множества N) решили встретиться в определенном месте в определенное время. Естественно, что из-за случайных отклонений все они будут прибывать в различные моменты времени; однако предполагается, что все порядки прибытия игроков (т. е. их перестановки) имеют одну и ту же вероятность, а именно $1/(n!)$. Предположим, что если игрок i , прибывая, застаёт на месте членов коалиции $T \setminus \{i\}$ (и только их), то он получает выигрыш $v(T) - v(T \setminus \{i\})$; иначе говоря, его выигрышем является предельная величина, которую он вносит в коалицию. Тогда компонента вектора Шепли $\varphi_i[v]$ представляет собой математическое ожидание выигрыша игрока i в условиях этой рандомизационной схемы.

3.13.7. Для простой игры (п. 3.11.6), формула для вектора Шепли особенно наглядна. Действительно, $v(T) - v(T \setminus \{i\})$ всегда равно либо 0, либо 1, причем это выражение равно 1, если T — выигрывающая коалиция, а коалиция $T \setminus \{i\}$ не является выигрывающей. Следовательно, имеем

$$\varphi_i[v] = \sum_T (t-1)!(n-t)!/n!,$$

где суммирование распространяется на все такие выигрывающие коалиции $T \supset i$, для которых коалиция $T \setminus \{i\}$ не является выигрывающей.

Пример 25 (Игра с главным игроком) [Воробьев, 1985]. В игре участвуют n игроков, один из которых называется «главным». Коалиция S выигрывает 1, если она либо содержит главного игрока и хотя бы одного кроме него, либо всех $n-1$ «неглавных». Если главный игрок имеет номер n , то характеристическая функция этой игры записывается в следующем виде:

$$v(S) = \begin{cases} 1, S \supset \{i, n\}, i \neq n, \\ 1, S \supset \{1, \dots, n-1\}, \\ 0, \text{ в остальных случаях.} \end{cases}$$

Ясно, что для всякой коалиции $T \supset \{n\}$ условия $v(T) = 1$ и $v(T \setminus \{n\}) = 0$ выполняются тогда и только тогда, когда $2 \leq |T| \leq n-1$. Следовательно,

$$\varphi_n[v] = \sum_{t=2}^{n-1} C_{n-1}^{t-1} \frac{(t-1)!(n-t)!}{n!} = \frac{n-2}{n}.$$

Поскольку игра имеет $(0-1)$ -редуцированную форму,

$$\sum_{i=1}^{n-1} \varphi_i[v] = 1 - \varphi_n[v] = 2/n.$$

Все неглавные игроки равноправны, поэтому в силу симметрии

$$\varphi_i[v] = \frac{2}{n(n-1)}, i = 1, \dots, n-1.$$

Таким образом «монопольное» положение главного игрока обеспечивает ему в $(n-1)(n-2)/2$ раз больший выигрыш, чем «рядовым» участникам игры.

3.13.8. Пример 26 («Помещик и батраки») [Воробьев, 1985]. Предположим, что имеются $n-1$ батраков (игроки $i = 1, \dots, n-1$) и помещик (игрок n). Помещик, наняв k батраков, получит от урожая доход $f(k)$ ($f(k)$ монотонно возрастает), а батраки сами дохода получить не могут. Это описывается следующей характеристической функцией:

$$v(S) = \begin{cases} f(|S| - 1), \{n\} \in S, \\ 0, \text{ в противном случае.} \end{cases}$$

Здесь для всех $T \supset \{n\}$, $|T| > 1$, $v(T) - v(T \setminus \{n\}) = f(t-1)$, где $t = |T|$, и из (3.13.7) следует

$$\varphi_n[v] = \sum_{t=2}^n C_{n-1}^{t-1} \frac{(t-1)!(n-t)!}{n!} f(t-1) = \frac{1}{n} \sum_{t=1}^{n-1} f(t).$$

На основании условия эффективности и симметрии всех батраков

$$\varphi_i[v] = \frac{1}{n-1} (f(n-1) - \frac{1}{n} \sum_{t=1}^{n-1} f(t)), i = 1, \dots, n-1.$$

§ 3.14. Вектор Шепли и потенциал

3.14.1. Как и ранее, рассмотрим кооперативную игру n лиц с трансферабельными выигрышами в форме характеристической функции. Мы рассматривали различные решения (или различные принципы оптимальности) игры. Как было показано, некоторые из них являются подмножествами множества дележей (как, например, C -ядро и NM -решение), а вектор Шепли представляет собой принцип оптимальности, состоящий из единственного дележа. В данном параграфе, следуя Харту и Мас-Колеллу [Hart, Mas-Colell, 1988], мы введем число, описывающее кооперативную игру. Используем принцип «маргинальных вкладов» и введем вектор выигрышей, состоящий из маргинальных вкладов игроков.

3.14.2. Кооперативной игрой с трансферабельными выигрышами называется пара (N, v) , где N — конечное множество игроков, а $v : 2^N \rightarrow R$ — характеристическая функция, удовлетворяющая условию $v(\emptyset) = 0$. Подмножество $S \subset N$ называется коалицией, а $v(S)$ — доходом коалиции S . Рассмотрим игру (N, v) и коалицию $S \subset N$. Через (S, v) обозначим подыгру, полученную сужением v на множество S (и его подмножества), то есть, область определения функции v сужается на 2^S .

3.14.3. Обозначим через $\hat{\Gamma}$ множество всех игр. Рассмотрим функцию $P : \hat{\Gamma} \rightarrow R$, ставящую в соответствие каждой игре (N, v) вещественное число $P(N, v)$. Маргинальным вкладом игрока i в игру (N, v) будем называть

$$D^i P(N, v) = P(N, v) - P(N \setminus \{i\}, v),$$

где $i \in N$. Здесь игра $(N \setminus \{i\}, v)$ представляет собой сужение (N, v) на $(N \setminus \{i\}, v)$.

Функция $P : \hat{\Gamma} \rightarrow R$ такая, что $P(\emptyset, v) = 0$, называется *функцией потенциала*, если она удовлетворяет следующему условию:

$$\sum_{i \in N} D^i P(N, v) = v(N) \quad (3.14.1)$$

для всех игр (N, v) . Таким образом, для функции потенциала вектор маргинальных вкладов игроков всегда эффективен, т. е. сумма его компонент равна доходу максимальной коалиции, состоящей из всех игроков.

3.14.4. Теорема. *Существует единственная функция потенциала P . Для любой игры (N, v) вектор маргинальных вкладов $(D^i P(N, v))_{i \in N}$ совпадает с вектором Шепли. Кроме того, потенциал игры (N, v) единственным образом определяется соотношением (3.14.1), примененным к самой игре (N, v) и ее подыграм (S, v) , $S \subset N$.*

Доказательство. Равенство (3.14.1) можно переписать в виде

$$P(N, v) = \frac{1}{|N|} \left(v(N) + \sum_{i \in N} P(N \setminus \{i\}, v) \right). \quad (3.14.2)$$

Начиная с равенства $P(\emptyset, v) = 0$, соотношение (3.14.2) рекурсивно определяет значение $P(N, v)$. Это доказывает существование и единственность функции потенциала P , а также то, что $P(N, v)$ единственным образом определяется соотношением (3.14.1) (или (3.14.2)), примененным к (S, v) для всех $S \subset N$.

Остается показать, что $D^i P(N, v) = Sh^i(N, v)$ для всех игр (N, v) и всех игроков $i \in N$, где P — (единственная) функция потенциала, а $Sh^i(N, v)$ — компонента вектора Шепли для игрока i в игре (N, v) . Для этого покажем, что $D^i P$ удовлетворяет

всем аксиомам, однозначно характеризующим вектор Шепли. Эффективность следует из (3.14.1); выполнение остальных трех аксиом — фиктивного (нулевого) игрока, симметричности и аддитивности, — докажем при помощи индукции по числу игроков, используя (3.14.2). В самом деле, пусть игрок i является нулевым игроком в игре (N, v) (т. е., $v(S) = v(S \setminus \{i\})$ для всех S). Покажем, что $P(N, v) = P(N \setminus \{i\}, v)$, а, следовательно, и $D^i P(N, v) = 0$. Предположим, что утверждение верно для всех игр с числом игроков, меньшим $|N|$; в частности, $P(N \setminus \{j\}, v) = P(N \setminus \{j, i\}, v)$ для всех $j \neq i$. Теперь вычтем равенство (3.14.2) для $N \setminus \{i\}$ из того же равенства (3.14.2) для N . Мы получим

$$|N|[P(N, v) - P(N \setminus \{i\}, v)] = [v(N) - v(N \setminus \{i\})] + \\ + \sum_{j \neq i} [P(N \setminus \{j\}, v) - P(N \setminus \{j, i\}, v)] = 0.$$

Теперь предположим, что игроки i и j взаимозаменяемы в игре (N, v) . Из этого следует, что $P(N \setminus \{i\}, v) = P(N \setminus \{j\}, v)$ (с помощью (3.14.2) и того факта, что i и j взаимозаменяемы в подыграх $(N \setminus \{k\}, v)$ для всех $k \neq i, j$). Таким образом, $D^i P(N, v) = D^j P(N, v)$. Кроме того, при помощи (3.14.2) по индукции можно доказать, что $P(N, v + w) = P(N, v) + P(N, w)$, откуда следует аддитивность.

Рассмотрим другой подход к интерпретации понятия «потенциал». В общем случае в игре (N, v) распределение маргинальных вкладов (т. е., $v(N) - v(N \setminus \{i\})$ для игрока i), не является эффективным. Одним из путей решения возникшей проблемы является добавление игрока, например, игрока 0, и расширение игры $N_0 = N \cup \{0\}$ таким образом, чтобы распределение маргинальных вкладов в расширенной игре являлось эффективным. Формально, пусть (N_0, v_0) — расширение игры (N, v) (т. е., $v_0(S) = v(S)$ для всех $S \subset N$). Тогда условие эффективности можно записать следующим образом:

$$v_0(N_0) = \sum_{i \in N_0} [v_0(N_0) - v_0(N_0 \setminus \{i\})] = \\ = [v_0(N_0) - v(N)] + \sum_{i \in N} [v_0(N_0) - v_0(N_0 \setminus \{i\})]. \quad (3.14.3)$$

Упрощая, получаем

$$v(N) = \sum_{i \in N} [v_0(N_0) - v_0(N_0 \setminus \{i\})]. \quad (3.14.4)$$

С учетом этого можно переформулировать результат теоремы п. 3.14.4.

3.14.5. Следствие. *Существует единственное расширение v_0 характеристической функции v , при котором распределение маргинальных вкладов всегда эффективно (точнее, (3.14.3) выполняется для игры и всех подыгр) и оно задается следующим образом: $v_0(S \cup \{0\}) = P(S, v)$ для всех $S \subset N$, где P — функция потенциала. Здесь*

$$P(S, v) = \frac{1}{|S|} \left(v(S) + \sum_{i \in S} P(S \setminus \{i\}, v) \right).$$

Отметим, что выигрыши первоначальных игроков (из коалиции N) в сумме дают $v(N)$ (3.14.4); они равны компонентам вектора Шепли. Игрока 0, чей выигрыш равен остатку $P(N, v) - v(N)$, можно рассматривать как «скрытого игрока».

Выражения (3.14.1) и (3.14.2) задают потенциал только неявным образом. Получим формулы в явном виде. Рассмотрим *простейшую* игру u_T (где T — непустое конечное

множество), определяемую как $u_T(S) = 1$ если $S \supset T$, и $u_T(S) = 0$ в противном случае. Известно, что такие игры составляют линейный базис для Γ : каждая игра (N, v) имеет единственное представление (см. [Shapley, 1953] и п. 3.13.4.) и § 3.8)

$$v = \sum_{T \subset N} \alpha_T u_T,$$

где для всех $T \subset N$ имеет место

$$\alpha_T \equiv \alpha_T(N, v) = \sum_{S \subset T} (-1)^{|T|-|S|} v(S). \quad (3.14.5)$$

3.14.6. Теорема. *Функция потенциала P удовлетворяет равенству*

$$P(N, v) = \sum_{T \subset N} \frac{1}{|T|} \alpha_T$$

для всех игр (N, v) , где α_T задается (3.14.5).

Доказательство. Обозначим правую часть в предыдущей формуле через $Q(N, v)$. Тогда $Q(\emptyset, v) = 0$, и $Q(N, v) - Q(N \setminus \{i\}, v) = \sum_{i \in T} \alpha_T / |T|$. Суммируя по i , получаем, что Q удовлетворяет (3.14.1). Поэтому, по теореме п. 3.11.4, Q совпадает с единственной функцией потенциала P .

Число $\delta_T = \alpha_T / |T|$ называется *дивидендом* каждого члена коалиции T и $Sh^i(N, v) = \sum_{i \in T} \delta_T$ [Harsanyi, 1963].

3.14.7. Теорема. *Функция потенциала P может быть вычислена по формуле*

$$P(N, v) = \sum_{S \subset N} \frac{(s-1)!(n-s)!}{n!} v(S),$$

где $n = |N|$, а $s = |S|$.

Доказательство. Нетрудно видеть, что вектор маргинальных вкладов для функции в правой части совпадает с вектором Шепли.

Для интерпретации последней формулы рассмотрим следующую вероятностную модель выбора случайной непустой коалиции $S \subset N$. Сначала выберем размер коалиции $s = 1, 2, \dots, n = |N|$ равновероятно (т. е. с вероятностью $1/n$ каждый). Затем выберем подмножество S размерности s , снова равновероятно (т. е. каждое из C_n^s подмножеств имеет одинаковую вероятность). Или, что эквивалентно, выберем случайный порядок следования n элементов из N (с вероятностью $1/n!$ каждый), а затем выберем граничную точку s ($1 \leq s \leq n$), а коалицией S будут первые s элементов. Вероятность выбора множества S при $|S| = s$ равна

$$\pi_S = \frac{s!(n-s)!}{n \cdot n!} = \frac{s}{n} \frac{(s-1)!(n-s)!}{n!}.$$

Следовательно, формула из п. 3.14.7 может быть переписана в виде

$$P(N, v) = \sum_{S \subset N} \pi_S \frac{n}{s} v(S) = E \left[\frac{|N|}{|S|} v(S) \right], \quad (3.14.6)$$

где E обозначает математическое ожидание S в вышеупомянутой вероятностной модели.

§ 3.15. Упражнения и задачи

1. Два объединения производят разведку полезных ископаемых на n месторождениях. Фонды средств на разведку у 1-го и 2-го объединения составляют α и β соответственно. Прибыль от добычи полезных ископаемых на i -м месторождении равна $\gamma_i > 0$, она распределяется между объединениями пропорционально доле средств, которые они вложили в i месторождение. При этом если в i месторождение обоими не вложено никаких средств, то и прибыли, полученные обоими объединениями на i -м месторождении, также равны нулю.

(а) Описать указанный конфликт в виде игры двух лиц, считая выигрышем каждого объединения суммарную прибыль, полученную от добычи полезных ископаемых на всех месторождениях.

(б) Найти ситуацию равновесия по Нэшу.

Указание. Воспользоваться вогнутостью функций H_1 по x и H_2 по y .

2. В экологически значимом районе имеется n промышленных предприятий, на каждом из которых один источник загрязнения. Значение концентрации q_i вредной примеси, выбрасываемой i -м предприятием, пропорционально величине выброса $0 \leq x_i \leq a_i$, $i = 1, \dots, n$, этого предприятия. Потери i -го предприятия складываются из расходов на переработку отходов производства ($f_i(x_i)$) и налога за загрязнение, который пропорционален суммарной концентрации q вредной примеси от выброса всех предприятий. Величина Θ не должна превышать q — значения предельно допустимой концентрации вредной примеси. В противном случае i -е предприятие дополнительно платит штраф s_i .

Описать указанный конфликт в виде бескоалиционной игры n лиц, считая проигрышем каждого предприятия суммарные затраты на природоохранные мероприятия.

Указание. Воспользоваться результатом примера 5, п.3.1.4.

3. Найти множества всех ситуаций равновесия по Нэшу (в чистых стратегиях) в следующих $(m \times n)$ -биматричных играх с матрицами $A = \{\alpha_{ij}\}$ и $B = \{\beta_{ij}\}$.

(а) Матрицы A и B — диагональные и положительные, т. е. $m = n$, $\alpha_{ij} = \beta_{ij} = 0$, $i \neq j$ и $\alpha_{ii} > 0$, $\beta_{ii} > 0$, $i = 1, \dots, m$, $j = 1, \dots, n$.

(б)

$$A = \begin{bmatrix} 2 & 0 & 5 \\ 2 & 2 & 3 \end{bmatrix}, \quad B = \begin{bmatrix} 2 & 2 & 1 \\ 0 & 7 & 8 \end{bmatrix}.$$

(с)

$$A = \begin{bmatrix} 3 & 8 & -1 \\ 4 & 0 & 2 \\ 1 & 2 & 3 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 3 & 4 \\ 2 & 1 & 8 \\ 2 & 3 & 0 \end{bmatrix}.$$

4. Показать, что в биматричной игре с матрицами

$$A = \begin{bmatrix} 1 & 2 & 0 \\ 1 & 3 & 1 \\ 2 & 2 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} 3 & 4 & 0 \\ 1 & 3 & 2 \\ 1 & 3 & 0 \end{bmatrix}$$

ситуация (2, 2) является равновесной. Является ли она сильно равновесной?

5. В биматричной игре с матрицами

$$A = \begin{bmatrix} 4 & 1 & 0 \\ 2 & 7 & 5 \\ 6 & 0 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} 0 & 5 & 6 \\ 7 & 0 & 2 \\ 2 & 6 & 1 \end{bmatrix}.$$

найти все ситуации, оптимальные по Парето в чистых стратегиях. Есть ли в этой игре равновесные ситуации в чистых стратегиях?

6. Изобразить графически в координатах (K_1, K_2) множество всевозможных векторов выигрышей в смешанных стратегиях в игре «семейный спор» (см. п. 3.1.4).

Указание. Произвольные смешанные стратегии x и y игроков 1 и 2 соответственно могут быть записаны в виде $x = (\xi, 1 - \xi)$, $y = (\eta, 1 - \eta)$, $\xi, \eta \in [0, 1]$. Записывая функции выигрыша K_1 и K_2 в смешанных стратегиях и исключая один из параметров, получаем однопараметрическое семейство отрезков, объединение которых и есть искомое множество (см. рис. 3.2). Криволинейная часть границы представляет собой огибающую этого семейства отрезков и является частью параболы: $5K_1^2 + 5K_2^2 - 10K_1K_2 - 18(K_1 + K_2) + 45 = 0$.

7. В биматричной игре с матрицами

$$A = \begin{bmatrix} 6 & 0 & 2 \\ 0 & 4 & 3 \\ 7 & 0 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 6 & 0 & 7 \\ 0 & 4 & 0 \\ 2 & 3 & 0 \end{bmatrix}$$

найти вполне смешанную ситуацию равновесия по Нэшу. Имеет ли эта игра еще ситуации равновесия в смешанных стратегиях?

Указание. Найти сначала вполне смешанную ситуацию равновесия (x, y) , $x = (\xi_1, \xi_2, \xi_3)$, $y = (\eta_1, \eta_2, \eta_3)$, затем такую равновесную ситуацию, для которой $\xi_1 = 0$ и т. д.

8. «Игра на оригинальность» [Воробьев, 1984]. Рассматривается бескоалиционная игра n лиц $\Gamma = (N, \{X_i\}_{i \in N}, \{H_i\}_{i \in N})$, где $X_i = \{0, 1\}$, $H_i(0, \dots, 0 \| 1) = g_i > 0$, $H_i(1, \dots, 1 \| i0) = h_i > 0$, $H_i(x) = 0$ в остальных случаях, где $\|_i$ означает, что замена производится на i -м месте.

(а) Интерпретировать игру в терминах рекламного дела.

(б) Найти вполне смешанную ситуацию равновесия.

9. В п. 1.10.1 было показано, что игры двух лиц с нулевой суммой можно решать методом «фиктивного разыгрывания». Рассматривая биматричную игру показать, что этот метод не может быть использован для нахождения ситуаций с матрицами

$$A = \begin{bmatrix} 2 & 0 & 1 \\ 1 & 2 & 0 \\ 0 & 1 & 2 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 0 & 2 \\ 2 & 1 & 0 \\ 0 & 2 & 1 \end{bmatrix},$$

равновесия в биматричных играх.

10. Игра «музыкальные стулья» [Мулен, 1985]. Имеются два игрока и три стула, помеченные цифрами 1, 2, 3. Стратегия игрока состоит в выборе номера стула. Оба игрока несут потери при выборе одного и того же стула. Если же их выборы различны, то тот игрок i , чей стул следует сразу за стулом игрока j , выигрывает вдвое больше, чем игрок j (предполагается, что стул 1 следует за стулом 3). Получаем биматричную игру $\Gamma(A, B)$,

$$(A, B) = \begin{bmatrix} (0, 0) & (1, 2) & (2, 1) \\ (2, 1) & (0, 0) & (1, 2) \\ (1, 2) & (2, 1) & (0, 0) \end{bmatrix}.$$

(а) Показать, что единственное вполне смешанное равновесие по Нэшу состоит в равновесном выборе стульев каждым игроком.

(б) Показать, что равновесие в совместных смешанных стратегиях имеет вид

$$L(i, j) = \begin{cases} 1/6, & \text{если } i \neq j, \\ 0, & \text{если } i = j. \end{cases}$$

(с) Показать, что выигрыши в ситуации равновесия по Нэшу не являются оптимальными по Парето, а равновесие в совместных смешанных стратегиях приводит к выигрышам $(3/2, 3/2)$, оптимальным по Парето.

11. Равновесие в совместных смешанных стратегиях не обязывает игроков придерживаться чистых стратегий, реализовавшихся в результате принятой совместной смешанной стратегии (см. определение п. 3.9.1). Если же мы обязаны придерживаться результатов конкретной реализации совместной смешанной стратегии, то можно обобщить «равновесие в совместных

смешанных стратегиях». Для всех $i \in N$ обозначим через $\mu(N \setminus \{i\})$ сужение распределения μ на множество $X_{N \setminus \{i\}} = \prod_{i \in N \setminus \{i\}} X_i$, а именно

$$\mu(N \setminus \{i\}) = \sum_{x_i \in X_i} \mu(x \| x_i)$$

для всех $x \in \prod_{i \in N} X_i$. Будем говорить, что μ есть слабое равновесие в совместных смешанных стратегиях, если выполнены следующие неравенства для всех $i \in N$ и $y_i \in X_i$:

$$\sum_{x \in \prod_{i \in N} X_i} H_i(x) \mu(x) \geq \sum_{x \in \prod_{i \in N} X_i} H(x \| y_i) \mu(N \setminus \{i\}).$$

(а) Доказать, что всякое равновесие в совместных смешанных стратегиях является слабым равновесием в совместных смешанных стратегиях.

(б) Пусть $\mu = (\mu_1, \dots, \mu_n)$ — векторная ситуация в смешанных стратегиях в игре Г. Показать, что вероятностная мера $\bar{\mu} = \prod_{i \in N} \mu_i$ на множестве $X = \prod_{i \in N} X_i$ является слабым равновесием в совместных смешанных стратегиях и равновесием в совместных стратегиях тогда и только тогда, когда ситуация $\mu = (\mu_1, \dots, \mu_n)$ равновесна по Нэшу.

12. (а) Доказать, что в игре, сформулированной в упр. 10, множество ситуаций равновесия по Нэшу, множество ситуаций равновесия в совместных стратегиях и множество ситуаций слабо равновесных в совместных смешанных стратегиях различны и не совпадают между собой.

(б) Показать, что множество вектор-выигрышей, оптимальных по Парето среди выигрышей в ситуации равновесия в совместных смешанных стратегиях, покрывает отрезок $[(5/3, 4/3), (4/3, 5/3)]$, а выигрыши, оптимальные по Парето среди слабо равновесных в совместных смешанных стратегиях, покрывают отрезок $[(2, 1), (1, 2)]$.

13. Найти арбитражное решение биматричной игры с матрицами $A = \begin{bmatrix} 2 & -1 \\ -1 & 1 \end{bmatrix}$, $B = \begin{bmatrix} 1 & -1 \\ -1 & 2 \end{bmatrix}$, используя схему Нэша.

14. Рассмотрим биматричную (2×2) -игру с матрицей

$$(A, B) = \begin{matrix} & \begin{matrix} \beta_1 & \beta_2 \end{matrix} \\ \begin{matrix} \alpha_1 \\ \alpha_2 \end{matrix} & \begin{bmatrix} (1, 1) & (1, 2) \\ (2, 1) & (-5, 0) \end{bmatrix} \end{matrix}.$$

Это модификация игры «перекресток» (см. пример 2 п. 3.1.4), отличие которой заключается лишь в следующем. Водитель легкового автомобиля (игрок 1) и грузового (игрок 2) по-разному оценивают результаты аварии (ситуация (α_2, β_2)). Показать, что анализ игры в стратегиях угроз предписывает ситуацию (α_1, β_2) , т. е. грузовому автомобилю «ехать», а легковому — «остановиться».

15. Пусть ядро имеет непустое пересечение со всеми гранями $x_i = v(\{i\})$ множества дележей. Показать, что в этом случае оно является единственным NM -решением.

16. Для кооперативной игры (N, v) определим полудележ как вектор $\alpha = (\alpha_1, \dots, \alpha_n)$, для которого $\alpha_i \geq v(\{i\})$ и $\sum_{i=1}^n \alpha_i \leq v(N)$. Показать, что если L является NM -решением игры (N, v) , а α — полудележ, не принадлежащий L , то существует такой дележ $\beta \in L$, что $\beta > \alpha$.

17. Для игры (N, v) определим β_i равенством

$$\beta_i = \max_{S \subset N \setminus \{i\}} [v(S \cup \{i\}) - v(S)].$$

Показать, что если найдется i , для которого $\alpha_i > \beta_i$, то дележ α не может принадлежать ни ядру, ни одному из NM -решений.

18. Пусть (N, v) — простая игра в $(0, -1)$ -редуцированной форме (см. п. 3.10.6) Игрок i называется «вето»-игроком, если $v(N \setminus \{i\}) = 0$.

(а) Доказать, что для того, чтобы C -ядро в простой игре было непустым, необходимо и достаточно, чтобы в игре существовал хотя бы один «вето»-игрок.

(б) Пусть S — множество всех «вето»-игроков. Показать, что дележ $\alpha = (\alpha_1, \dots, \alpha_n)$ принадлежит C -ядру, если $\sum_{i \in S} \alpha_i = 1$, $\alpha_i \geq 0$ для $i \in S$ и $\alpha_i = 0$ для $i \notin S$.

19. В игре (N, v) под квазидележом будем понимать вектор $\alpha = (\alpha_1, \dots, \alpha_n)$, такой, что $\sum_{i \in N} \alpha_i = v(N)$. Для каждого $\varepsilon > 0$ определим строгое ε -ядро $C_\varepsilon(v)$, как множество квазидележей, таких что для каждой коалиции

$$\sum_{i \in S} \alpha_i \geq v(S) - \varepsilon.$$

(а) Показать, что если $\varepsilon < \varepsilon'$, то $C_\varepsilon(v) \subset C_{\varepsilon'}(v)$.

(б) Показать, что существует наименьшее число, для которого $C_\varepsilon(v) \neq \emptyset$. При таком ε множество $C_\varepsilon(v)$ называется минимальным ε -ядром и обозначается через $MC(v)$.

(с) Найти минимальное ε -ядро в игре (N, v) , где $N = \{1, 2, 3\}$, $v(\{i\}) = 0$, $v(\{1, 2\}) = 50$, $v(\{1, 3\}) = 80$, $v(\{2, 3\}) = 90$, $v(\{N\}) = 100$.

(д) Пусть (N, v) , (N, v') — две кооперативные игры и для некоторых ε и ε' выполняется равенство $C_\varepsilon(v) = C_{\varepsilon'}(v') \neq \emptyset$. Показать, что в этом случае справедливо $C_{\varepsilon-\delta}(v) = C_{\varepsilon'-\delta}(v')$ для всех $\delta > 0$, $\delta < \min[\varepsilon, \varepsilon']$.

20. Показать, что если (N, v) — игра с постоянной суммой (см. п. 3. 9.3), то вектор Шепли Sh определяется по формуле

$$Sh_i(v) = 2 \sum_{S: S \subset N, i \in S} \left[\frac{(n-s)!(s-1)!}{n!} v(S) \right] - v(N).$$

21. Игра (N, v) называется выпуклой, если для всех $S, T \subset N$

$$v(S \cup T) + v(S \cap T) \geq v(S) + v(T).$$

(а) Доказать, что выпуклая игра имеет непустое C -ядро и вектор Шепли принадлежит C -ядру.

(б) Показать, что (N, v) — выпуклая игра, если

$$v(S) = \left(\sum_{i \in S} m_i \right)^2, \quad S \subset N,$$

а $m = (m_1, \dots, m_n)$ — неотрицательный вектор.

22. Рассмотрим простую игру (N, v) в $(0-1)$ -редуцированной форме. Под «скачком» игрока i будем понимать такое множество $S \subset N$, для которого $v(S) = 1$ и $v(S \setminus \{i\}) = 0$. Обозначим через Θ_i число скачков игрока i в игре. Тогда вектор $\beta(v) = (\beta_1(v), \dots, \beta_n(v))$, где $\beta_i(v) = \Theta_i / \sum_{j=1}^n \Theta_j$ называется *вектором Банзафа* для простой игры.

(а) Показать, что $\Theta_1 = 6$, $\Theta_2 = \Theta_3 = \Theta_4 = 2$, и, следовательно, $\beta(v) = (1/2, 1/6, 1/6, 1/6)$ для простой игры четырех лиц (N, v) , в которой коалиция S выигрывает, если она состоит либо из двух игроков и игрока $\{1\} \in S$, либо из трех игроков или четырех игроков.

(б) Показать, что $\beta(v)$ совпадает с вектором Шепли.

23. Пусть (N, v) — простая игра трех лиц, в которой коалиции $(1, 2)$, $(1, 3)$, $(1, 2, 3)$ являются единственными выигрывающими коалициями. Показать, что в этой игре $\Theta_1 = 3$, $\Theta_2 = \Theta_3 = 1$, и, следовательно, вектор Банзафа имеет вид $\beta(v) = (3/5, 1/5, 1/5)$, а вектор Шепли равен $Sh[v] = (2/3, 1/6, 1/6)$.

24. Рассмотрим неотрицательный вектор $p = (\pi_1, \dots, \pi_n)$ и число $\Theta > 0$. Пусть $0 < \Theta \leq \sum_{i=1}^n \pi_i$. Взвешенной игрой большинства будем называть простую игру (N, v) , в которой характеристическая функция v определяется по правилу

$$v(S) = \begin{cases} 0, & \sum_{i \in S} \pi_i < \theta, \\ 1, & \sum_{i \in S} \pi_i \geq \theta. \end{cases}$$

Пусть $\Theta = 8$ и $p = (4, 3, 3, 2, 2, 1)$, $n = 6$. Вычислить вектор Шепли и вектор Банзафа для простой взвешенной игры большинства.

Глава 4

Многошаговые игры

§ 4.1. Определение динамической игры с полной информацией

4.1.1. Ранее нами рассматривалась игра в нормальной форме. К такой форме в принципе может быть сведен динамический (т. е. происходящий в течение некоторого промежутка времени) конфликтно-управляемый процесс формальным введением понятия чистой стратегии. В тех немногочисленных случаях, когда мощность пространства стратегий невелика и имеется возможность численного нахождения решений, такой подход является вполне допустимым. Однако в большинстве задач поиска оптимального поведения участников конфликтно-управляемого процесса переход к нормальной форме, т. е. сведение задачи к однократному выбору чистых стратегий как элементов пространств больших размерностей или функциональных пространств, не приводит к эффективным способам нахождения решений, хотя и позволяет наглядно иллюстрировать те или иные принципы оптимальности. В ряде случаев общие теоремы существования решения для игр в нормальной форме не позволяют находить или даже конкретизировать оптимальное поведение в играх, нормализацией которых они являются. Как будет показано ниже, в «шаматах» существует решение в классе чистых стратегий. Однако этот результат невозможно получить непосредственным исследованием матричной игры. Еще более наглядно это обстоятельство проявляется при исследовании дифференциальных игр, для которых в ряде случаев удается находить решения в явной форме. Однако нормальная форма дифференциальной игры является настолько общей, что получение конкретных результатов оказывается практически невозможным.

4.1.2. Математические модели конфликтов, учитывающие динамику, исследуются в теории *позиционных* игр. Наиболее простым классом позиционных игр является класс *конечношаговых игр с полной информацией*. Для определения конечношаговой игры n лиц с полной информацией потребуются элементарные сведения из теории графов.

Пусть X — некоторое конечное множество. Правило f , ставящее в соответствие каждому элементу $x \in X$ элемент $f(x) \in X$, называется однозначным отображением X в X_n или функцией, определенной на X и принимающей значения в X . Многозначное отображение F множества X в X — это правило, которое каждому элементу $x \in X$ ставит в соответствие некоторое подмножество $F_x \subset X$ (при этом не исключается возможность $F_x = \emptyset$). В дальнейшем для простоты будем употреблять термин «отображение», понимая под ним «многозначное отображение».

Пусть F — отображение X в X , и $A \subset X$. Под образом множества A будем понимать множество $F(A) \triangleq \bigcup_{x \in A} F_x$.

Полагаем $F(\emptyset) = \emptyset$. Можно убедиться в том, что если $A_i \subset X$, $i = 1, \dots, n$, то

$$F\left(\bigcup_{i=1}^n A_i\right) = \bigcup_{i=1}^n F(A_i), \quad F\left(\bigcap_{i=1}^n A_i\right) \subset \bigcap_{i=1}^n F(A_i).$$

Определим отображения $F^2, F^3, \dots, F^k, \dots$, следующим образом:

$$F_x^2 \triangleq F(F_x), \quad F_x^3 \triangleq F(F_x^2), \dots, \quad F_x^k \triangleq F(F_x^{k-1}), \dots$$

Отображение \hat{F} множества X в X называется *транзитивным замыканием* отображения F , если $\hat{F}_x \triangleq \{x\} \cup F_x \cup F_x^2 \cup \dots \cup F_x^k \dots$.

Отображение F^{-1} , обратное отображению F , определяется как

$$F_y^{-1} \triangleq \{x | y \in F_x\}.$$

Другими словами, F_y^{-1} — это множество тех точек x , образ которых содержит точку y . Аналогично отображению F_x^k определяется отображение $(F^{-1})_y^k$:

$$(F^{-1})_y^2 = F^{-1}((F^{-1})_y),$$

$$(F^{-1})_y^3 = F^{-1}((F^{-1})_y^2), \dots, (F^{-1})_y^k = F^{-1}((F^{-1})_y^{k-1}).$$

Если $B \subset X$, то полагаем $F^{-1}(B) \triangleq \{x | F_x \cap B \neq \emptyset\}$.

Определение. Пара (X, F) называется *графом*, если X — некоторое конечное множество, а F — отображение X в X .

Граф (X, F) будем обозначать символом G . В дальнейшем элементы множества X будем изображать точками на плоскости, а пары точек x и y , для которых $y \in F_x$, соединять непрерывной линией со стрелкой, направленной от x к y . Тогда каждый элемент множества X называется *вершиной* или *узлом* графа, а пара элементов (x, y) , в которой $y \in F_x$ — *дугой* графа. Для дуги $p = (x, y)$ вершины x и y называются *граничными вершинами* дуги, причем x — начало, а y — конец дуги. Две дуги p и q называются *смежными*, если они различны и имеют общую граничную точку.

Множество дуг в графе будем обозначать P . Задание множества дуг в графе $G = (X, F)$ определяет отображение F и, наоборот, отображение F определяет множество P . Поэтому граф G можно записывать как в виде $G = (X, F)$, так и в виде $G = (X, P)$.

Путь в графе $G = (X, F)$ называется такая последовательность $p = (p_1, p_2, \dots, p_k, \dots)$ дуг, что конец каждой предыдущей дуги совпадает с началом следующей. Длина пути $p = (p_1, \dots, p_k)$ есть число $l(p) = k$ дуг последовательности. В случае бесконечного пути p полагаем $l(p) \triangleq \infty$.

Ребром графа $G = (X, F)$ называется множество из двух элементов $x, y \in X$, для которых или $(x, y) \in P$, или $(y, x) \in P$. В отличие от дуги в ребре ориентация роли не играет. Ребра будем обозначать буквами p, q , а множество ребер — P . Под *цепью* будем понимать последовательность ребер (p_1, p_2, \dots) , в которой у каждого ребра p_k

одна из граничных вершин является также граничной для p_{k-1} , а другая — граничной для p_{k+1} .

Цикл — это конечная цепь, начинающаяся в некоторой вершине и оканчивающаяся в той же вершине. Граф называется связным, если любые две его вершины можно соединить цепью.

Дерево или *древовидный граф*, по определению, есть конечный связный граф без циклов, имеющий не менее двух вершин, в котором существует единственная вершина x_0 , такая, что $\hat{F}_{x_0} = X$. Вершина x_0 называется *начальной вершиной* графа G .

4.1.3. Пусть $z \in X$. *Подграфом* G_z древовидного графа $G = (X, F)$ называется граф вида (X_z, F_z) , где $X_z = \hat{F}_z$, а $F_z x = F_x \cap X_z$.

В древовидном графе для всех $x \in X_z$ множество F_x и множество $F_z x$ совпадают, т. е. отображение F_z является сужением отображения F на множество X_z . Поэтому для подграфов древовидного графа будем использовать обозначение $G_z = (X_z, F)$.

4.1.4. Перейдем к определению *многошаговой игры с полной информацией на древовидном конечном графе*.

Определение. Пусть $G = (X, F)$ — древовидный граф. Рассмотрим разбиение множества вершин X на $n+1$ множество X_1, \dots, X_n, X_{n+1} , $\bigcup_{i=1}^{n+1} X_i = X$, $X_k \cap X_e = \emptyset$, $k \neq l$, где $F_x = \emptyset$ для $x \in X_{n+1}$. Множество X_i , $i = 1, \dots, n$ называется множеством очередности игрока i , а множество X_{n+1} — множеством окончательных позиций. На множестве окончательных позиций X_{n+1} определены n вещественных функций $H_1(x), \dots, H_n(x)$, $x \in X_{n+1}$. Функция $H_i(x)$, $i = 1, \dots, n$, называется *выигрышем* игрока i .

Игра происходит следующим образом. Задано множество N игроков, перенумерованных натуральными числами $1, 2, \dots, i, \dots, n$ (в дальнейшем $N = \{1, 2, \dots, n\}$). Пусть $x_0 \in X_{i_1}$, тогда в вершине (позиции) x_0 «ходит» игрок i_1 и выбирает вершину $x_1 \in F_{x_0}$. Если $x_1 \in X_{i_2}$, то в вершине x_1 «ходит» игрок i_2 и выбирает следующую вершину (позицию) $x_2 \in F_{x_1}$, и т. д. Таким образом, если на k -м шаге вершина (позиция) $x_{k-1} \in X_{i_k}$, то в ней «ходит» игрок i_k и выбирает следующую вершину (позицию) из множества $F_{x_{k-1}}$. Игра прекращается как только достигается окончательная вершина (позиция) $x_l \in X_{n+1}$, т. е. такая, для которой $F_{x_l} = \emptyset$.

4.1.5. В результате последовательного выбора позиций однозначно реализуется некоторая последовательность $x_0, \dots, x_k, \dots, x_l$, определяющая путь в древовидном графе G , исходящий из начальной позиции x_0 и достигающий одной из окончательных позиций игры. Такой путь в дальнейшем будем называть *партией*. Из-за древовидности графа G каждая партия заканчивается в окончательной позиции x_l и, наоборот, окончательная позиция x_l однозначно определяет партию. В позиции x_l каждый из игроков $i = 1, 2, \dots, n$ получает выигрыш $H_i(x_l)$. Будем предполагать, что игрок i при совершении выбора в позиции $x \in X_i$ знает эту позицию x , а следовательно, из-за древовидности графа G может восстановить и все предыдущие позиции. В таком случае говорят, что игроки имеют полную информацию. Примером игр с полной информацией служат шахматы и шашки, поскольку в них игроки могут записывать ходы и поэтому можно считать, что они знают предысторию игры при совершении каждого очередного хода.

Определение. Однозначное отображение u_i , которое каждой вершине (позиции) $x \in X_i$ ставит в соответствие некоторую вершину (позицию) $y \in F_x$, называется *стратегией* игрока i .

Множество всевозможных стратегий игрока i будем обозначать через U_i . Таким образом, стратегия игрока i предписывает ему в любой позиции x из множества его очередности X_i однозначный выбор следующей позиции.

Упорядоченный набор $u = (u_1, \dots, u_i, \dots, u_n)$, где $u_i \in U_i$, называется *ситуацией в игре*, а декартово произведение $U = \prod_{i=1}^n U_i$ — *множеством ситуаций*. Каждая ситуация $u = (u_1, \dots, u_i, \dots, u_n)$ однозначно определяет партию игры, а следовательно, и выигрыши игроков. Действительно, пусть $x_0 \in X_{i_1}$. Тогда в ситуации $u = (u_1, \dots, u_i, \dots, u_n)$ следующая позиция x_1 определяется однозначно по правилу $u_{i_1}(x_0) = x_1$. Пусть теперь $x_1 \in X_{i_2}$. Тогда x_2 определяется однозначно по правилу $u_{i_2}(x_1) = x_2$. Если теперь на k -м шаге реализовалась позиция $x_{k-1} \in X_{i_k}$, то x_k определяется однозначно по правилу $x_k = u_{i_k}(x_{k-1})$ и т. д.

Пусть ситуации $u = (u_1, \dots, u_i, \dots, u_n)$ в указанном смысле соответствует партия x_0, x_1, \dots, x_l . Тогда можно ввести понятие *функции выигрыша* K_i игрока i , положив ее значение в каждой ситуации u равным значению выигрыша H_i в окончательной позиции партии x_0, x_1, \dots, x_l , соответствующей ситуации $u = (u_1, \dots, u_n)$, т. е. $K_i(u_1, \dots, u_i, \dots, u_n) = H_i(x_l)$, $i = 1, \dots, n$. Функции K_i , $i = 1, \dots, n$, определены на множестве ситуаций $U = \prod_{i=1}^n U_i$. Таким образом, мы получаем некоторую игру в нормальной форме $\Gamma = (N, \{U_i\}_{i \in N}, \{K_i\}_{i \in N})$, где $N = \{1, \dots, i, \dots, n\}$ — множество игроков, U_i — множество стратегий игрока i , K_i — функция выигрыша игрока i .

§ 4.2. Равновесие по Нэшу

4.2.1. Повторим определение п. 1.5.2, сформулировав его в обозначениях, принятых в теории динамических игр с полной информацией. Ситуация $u^* = (u_1^*, \dots, u_i^*, \dots, u_n^*)$ называется равновесием по Нэшу, если имеет место неравенство

$$K_i(u_1^*, \dots, u_{i-1}^*, u_i^*, u_{i+1}^*, \dots, u_n^*) \geq K_i(u_1^*, \dots, u_{i-1}^*, u_i, u_{i+1}^*, \dots, u_n^*)$$

для всех $u_i \in U_i$, $i \in N$.

Для дальнейшего исследования игры Γ необходимо ввести понятие подыгры, т. е. игры на подграфе графа G основной игры.

Пусть $z \in X$. Рассмотрим подграф $G_z = (X_z, F)$, с которым свяжем подыгру Γ_z следующим образом. Множество очередности игроков в подыгре Γ_z определяется по правилу $Y_i^z = X_i \cap X_z$, $i = 1, 2, \dots, n$, множество окончательных позиций $Y_{n+1}^z = X_{n+1} \cap X_z$, выигрыш $H_i^z(x)$ игрока i в подыгре полагается равным

$$H_i^z(x) = H_i(x), \quad x \in Y_{n+1}^z, \quad i = 1, \dots, n.$$

В соответствии с этим стратегия u_i^z i -го игрока в подыгре Γ_z определена как сужение стратегии u_i игрока i в игре Γ на множество Y_i^z , т. е. $u_i^z = u_i(x)$, $x \in Y_i^z = X_i \cap X_z$, $i = 1, \dots, n$. Множество всех стратегий i -го игрока в подыгре обозначается через U_i^z . В результате с каждым подграфом G_z мы связываем подыгру в нормальной форме $\Gamma_z = (N, \{U_i^z\}, \{K_i^z\})$, где функции выигрыша K_i^z , $i = 1, \dots, n$, определены на декартовом произведении $U^z = \prod_{i=1}^n U_i^z$.

Определение. Ситуация равновесия по Нэшу основной игры $u^* = (u_1^*, \dots, u_n^*)$ называется ситуацией абсолютного равновесия по Нэшу в игре Γ , если для любого $z \in X$ ситуация $(u^*)^z = ((u_1^*)^z, \dots, (u_n^*)^z)$, где $(u_i^*)^z$ — сужение стратегии u_i^* на подыгру Γ_z , является ситуацией равновесия по Нэшу в подыгре Γ_z .

Имеет место следующая основная теорема.

Теорема. В любой многошаговой игре с полной информацией на конечном древовидном графе существует ситуация абсолютного равновесия по Нэшу.

Прежде чем перейти к ее доказательству, введем понятие длины игры. Под длиной игры Γ будем понимать длину наибольшего пути в графе $G = (X, F)$.

Доказательство проведем индукцией по длине игры. Если длина игры Γ равна единице, то может ходить лишь один из игроков, который, выбирая следующую вершину из условия максимизации своего выигрыша, будет действовать согласно стратегии, образующей абсолютное равновесие по Нэшу.

Пусть теперь игра Γ имеет длину k и $x_0 \in X_{i_1}$ (т. е. в начальной позиции x_0 ходит игрок i_1). Рассмотрим семейство подыгр Γ_z , $z \in F_{x_0}$, длина каждой из которых не превосходит $k-1$. Предположим, что теорема справедлива для всех игр, длина которых не превосходит $k-1$, и докажем ее для игры k . Поскольку подыгры Γ_z , $z \in F_{x_0}$, имеют длину не более $k-1$, по предположению индукции для них теорема справедлива и тем самым существует ситуация абсолютного равновесия по Нэшу. Обозначим для каждой подыгры Γ_z , $z \in F_{x_0}$ эту ситуацию через

$$(\bar{u})^z = [(\bar{u}_1)^z, \dots, (\bar{u}_n)^z]. \quad (4.2.1)$$

Используя ситуации абсолютного равновесия в подыграх Γ_z , построим ситуацию абсолютного равновесия в игре Γ . Пусть $\bar{u}_i(x) = (\bar{u}_i(x))^z$ для $x \in X_i \cap X_z$, $z \in F_{x_0}$, $i = 1, \dots, n$, $\bar{u}_{i_1}(x_0) = \bar{z}$, где \bar{z} находится из условия

$$K_{i_1}^{\bar{z}}[(\bar{u})^{\bar{z}}] = \max_{z \in F_{x_0}} K_{i_1}^z[(\bar{u})^z]. \quad (4.2.2)$$

Функция \bar{u}_i определена на множестве X_i , $i = 1, \dots, n$ очередности i -го игрока, а при каждом фиксированном $x \in X_i$ значение $\bar{u}_i(x) \in F_x$. Таким образом, \bar{u}_i , $i = 1, \dots, n$, является стратегией i -го игрока в игре Γ , т. е. $\bar{u}_i \in U_i$. По построению, сужение $(\bar{u}_i)^z$ стратегии \bar{u}_i на множество $X_i \cap X_z$ является стратегией, входящей в абсолютное равновесие по Нэшу игры Γ_z , $z \in F_{x_0}$. Следовательно, для завершения доказательства теоремы достаточно показать, что стратегии \bar{u}_i , $i = 1, \dots, n$, образуют ситуацию равновесия по Нэшу в игре Γ . Пусть $i \neq i_1$. По построению стратегии \bar{u}_{i_1} , после выбора игроком i_1 позиции \bar{z} на первом шаге игра Γ переходит в подыгру $\Gamma_{\bar{z}}$, поэтому

$$K_i(\bar{u}) = K_i^{\bar{z}}\{(\bar{u})^{\bar{z}}\} \geq K_i^{\bar{z}}\{(\bar{u}||u_i)^{\bar{z}}\} = K_i(\bar{u}||u_i), \quad (4.2.3)$$

$$u_i \in U_i, \quad i = 1, \dots, n, \quad i \neq i_1,$$

так как $(\bar{u})^z$ — ситуация абсолютного равновесия в подыгре Γ_z . Пусть $u_{i_1} \in U_{i_1}$ — произвольная стратегия игрока i_1 в игре Γ . Обозначим $z_0 = u_{i_1}(x_0)$. Тогда

$$\begin{aligned} K_{i_1}(\bar{u}) &= K_{i_1}^{\bar{z}}\{(\bar{u})^{\bar{z}}\} = \max_{z \in F_{x_0}} K_{i_1}^z\{(\bar{u})^z\} \geq K_{i_1}^{z_0}\{(\bar{u})^{z_0}\} \geq \\ &\geq K_{i_1}^{z_0}\{(\bar{u}||u_{i_1})^{z_0}\} = K_{i_1}(\bar{u}||u_{i_1}). \end{aligned} \quad (4.2.4)$$

Утверждение теоремы следует теперь из (4.2.3), (4.2.4).

Действительно, обозначим через $v_1(x)$, $v_2(x)$ выигрыши в подыгре Γ_x в некоторой фиксированной ситуации абсолютного равновесия. Сначала решаем подыгры $\Gamma_{1,6}$, $\Gamma_{1,7}$, $\Gamma_{2,7}$. Как легко убедиться, $v_1(1.6) = 6$, $v_2(1.6) = 2$, $v_1(1.7) = 2$, $v_2(1.7) = 4$, $v_1(2.7) = 1$, $v_2(2.7) = 8$. Далее решаем подыгры $\Gamma_{2,5}$, $\Gamma_{2,6}$, $\Gamma_{1,8}$. В подыгре $\Gamma_{2,5}$ два равновесия по Нэшу, поскольку игроку 2 безразлично, какую альтернативу выбрать. Вместе с тем его выбор оказывается существенным для игрока 1, поскольку при выборе игроком 2 левой дуги первый игрок выигрывает +1, а при выборе игроком 2 второй дуги +6. Отметим это обстоятельство и предположим, что игрок 2 «благожелателен» и выбирает в позиции (2.5) правую дугу. Тогда $v_1(2.5) = v_1(1.6) = 6$, $v_2(2.5) = v_2(1.6) = 2$, $v_1(2.6) = v_1(1.7) = 2$, $v_2(2.6) = v_2(1.7) = 4$, $v_1(1.8) = 2$, $v_2(1.8) = 3$. Далее решаем игры $\Gamma_{1,3}$, $\Gamma_{1,4}$, $\Gamma_{2,3}$, $\Gamma_{1,5}$, $\Gamma_{2,4}$. В подыгре $\Gamma_{1,3}$ два равновесия по Нэшу, поскольку игроку 1 безразлично, какую альтернативу выбрать. Вместе с тем его выбор оказывается существенным для игрока 2, так как при выборе игроком 1 левой альтернативы он выигрывает 1, а при выборе правой — 10. Предположим, что игрок 1 «благожелателен» и выбирает в позиции (1.3) правую альтернативу. Тогда $v_1(1.3) = 5$, $v_2(1.3) = 10$, $v_1(1.4) = v_1(2.5) = 6$, $v_2(1.4) = v_2(2.5) = 2$, $v_1(1.5) = v_1(2.6) = 2$, $v_2(1.5) = v_2(2.6) = 4$, $v_1(2.3) = 0$, $v_2(2.3) = 6$, $v_1(2.4) = 3$, $v_2(2.4) = 5$. Далее решаем игры $\Gamma_{2,1}$, $\Gamma_{1,2}$, $\Gamma_{2,2}$: $v_1(2.1) = v_1(1.3) = 5$, $v_2(2.1) = v_2(1.3) = 10$, $v_1(1.2) = v_1(2.4) = 3$, $v_2(1.2) = v_2(2.4) = 5$, $v_1(2.2) = -5$, $v_2(2.2) = 6$. Теперь решаем игру $\Gamma_{1,1}$. Здесь $v_1(1.1) = v_1(2.1) = 5$, $v_2(1.1) = v_2(2.1) = 10$.

В результате получаем ситуацию абсолютного равновесия по Нэшу (u_1^*, u_2^*) , где

$$u_1^* = (1, 2, 2, 2, 2, 3, 2, 1), \quad u_2^* = (1, 3, 2, 2, 2, 1, 2). \quad (4.2.5)$$

В ситуации (u_1^*, u_2^*) игра развивается по пути (1.1), (2.1), (1.3). В процессе построения было замечено, что стратегии u_i^* , $i = 1, 2$ «доброжелательны» в том смысле, что игрок i при совершении своего хода, будучи в равной степени заинтересован в выборе последующих альтернатив, выбирает ту из них, которая более благоприятна для другого игрока.

В игре Γ существуют ситуации абсолютного равновесия, в которых выигрыши игроков будут другими. Для построения таких равновесий достаточно снять условие «доброжелательности» игроков и заменить его обратным условием «недоброжелательности». Обозначим через $\bar{v}_1(x)$, $\bar{v}_2(x)$ выигрыши игроков в подыгре Γ_x при использовании игроками «недоброжелательного» равновесия. Тогда имеем $v_1(1.6) = \bar{v}_1(1.6) = 6$, $v_2(1.6) = \bar{v}_2(1.6) = 2$, $v_1(1.7) = \bar{v}_1(1.7) = 2$, $v_2(1.7) = \bar{v}_2(1.7) = 4$, $\bar{v}_1(2.7) = -2$, $v_2(2.7) = \bar{v}_2(2.7) = 8$. Как уже отмечалось, в подыгре $\Gamma_{2,5}$ два равновесия по Нэшу. В отличие от предыдущего случая предположим, что игрок 2 «недоброжелателен» и выбирает ту из вершин, в которой при его максимальном выигрыше выигрыш игрока 1 минимален. Тогда $\bar{v}_1(2.5) = 1$, $\bar{v}_2(2.5) = 2$, $\bar{v}_1(2.6) = v_1(1.7) = 2$, $\bar{v}_2(2.6) = v_2(1.7) = 4$, $\bar{v}_1(1.8) = v_1(1.8) = 2$, $\bar{v}_2(1.8) = v_2(1.8) = 3$. Далее ищем решение игр $\Gamma_{1,3}$, $\Gamma_{1,4}$, $\Gamma_{1,5}$, $\Gamma_{2,3}$, $\Gamma_{2,4}$. В подыгре $\Gamma_{1,3}$ два равновесия по Нэшу. Как и в предыдущем случае, выберем «недоброжелательные» действия игрока 1. Тогда имеем: $\bar{v}_1(1.3) = v_1(1.3) = 5$, $\bar{v}_2(1.3) = 1$, $\bar{v}_1(1.4) = 2$, $\bar{v}_2(1.4) = 3$, $\bar{v}_1(1.5) = v_1(2.6) = v_1(1.5) = 2$, $\bar{v}_2(1.5) = v_2(2.6) = v_2(2.6) = 4$, $\bar{v}_1(2.3) = v_1(2.3) = 0$, $\bar{v}_2(2.3) = v_2(2.3) = 6$, $\bar{v}_1(2.4) = v_1(2.4) = 3$, $\bar{v}_2(2.4) = v_2(2.4) = 5$. Далее решаем игры $\Gamma_{2,1}$, $\Gamma_{1,2}$, $\Gamma_{2,2}$. Имеем $\bar{v}_1(2.1) = \bar{v}_1(1.5) = 2$, $\bar{v}_2(2.1) = \bar{v}_2(1.5) = 4$, $\bar{v}_1(1.2) = \bar{v}_1(2.4) = 3$, $\bar{v}_2(1.2) = \bar{v}_2(2.4) = 5$, $\bar{v}_2(2.2) = v_2(2.2) = 6$, $v_1(2.2) = \bar{v}_1(2.2) = -5$. Теперь решаем игру $\Gamma = \Gamma_{1,1}$. Здесь $\bar{v}_1(1.1) = \bar{v}_1(1.2) = 3$, $\bar{v}_2(1.1) = \bar{v}_2(1.2) = 5$.

Таким образом, получена новая ситуация равновесия по Нэшу

$$\bar{u}_1^* = (2, 2, 1, 1, 2, 3, 2, 1), \quad \bar{u}_2^* = (3, 3, 2, 2, 1, 1, 3). \quad (4.2.6)$$

Выигрыши обоих игроков в ситуации (4.2.6) меньше таковых в ситуации (4.2.5). Ситуация (4.2.6), так же как и ситуация (4.2.5), является ситуацией абсолютного равновесия.

§ 4.3. Основные функциональные уравнения

4.3.1. Рассмотрим *многошаговые антагонистические игры с полной информацией*. Если в условиях определения § 4.1. множество игроков состоит из двух элементов $N = \{1, 2\}$ и $H_2(x) = -H_1(x)$ для всех $x \in X_3$ (X_3 — множество окончательных позиций в игре Γ), то $\Gamma = \langle N, U_i, K_i \rangle$ оказывается антагонистической многошаговой игрой с полной информацией. Очевидно, что этим же свойством обладают и все подыгры Γ_z игры Γ . Так как из условия $H_2(x) = -H_1(x)$ немедленно следует, что $K_2(u_1, u_2) = -K_1(u_1, u_2)$ для всех $u_1 \in U_1, u_2 \in U_2$, то в ситуации равновесия по Нэшу (u_1^*, u_2^*) выполняются неравенства $K_1(u_1, u_2^*) \leq K_1(u_1^*, u_2^*) \leq K_1(u_1^*, u_2)$ для всех $u_1 \in U_1, u_2 \in U_2$. Пару (u_1^*, u_2^*) в этом случае будем называть *ситуацией равновесия* или *седловой точкой*, а стратегии, образующие ситуацию равновесия, *оптимальными*. Значения функции выигрыша в ситуации равновесия обозначим v и назовем значением игры Γ .

Из п. 4.2.1 следует, что в антагонистической многошаговой игре с полной информацией на конечном древовидном графе существует ситуация абсолютного равновесия, т. е. такая ситуация (u_1^*, u_2^*) , сужение которой на любую подыгру Γ_z игры Γ образует в Γ_z ситуацию равновесия. Для любой подыгры Γ_y можно также определить число $v(y)$, представляющее значение функции выигрыша в ситуации равновесия этой подыгры и называемое *значением подыгры* Γ_y . Можно легко показать, что значение антагонистической игры (т. е. значение функции выигрыша игрока 1 в ситуации равновесия) определяется единственным образом, поэтому функция $v(y)$ определена для всех $y \in X_1, y \in X_2$ и является однозначной функцией.

4.3.2. Выведем функциональные уравнения для вычисления функции $v(y)$. Из определения $v(y)$ следует, что

$$v(y) = K_1^y((u_1^*)^y, (u_2^*)^y) = -K_2^y((u_1^*)^y, (u_2^*)^y),$$

где $((u_1^*)^y, (u_2^*)^y)$ — ситуация равновесия в подыгре Γ_y , являющаяся сужением ситуации абсолютного равновесия (u_1^*, u_2^*) .

Пусть $y \in X_1$ и $z \in F_y$. Тогда имеем

$$v(y) = \max_{z \in F_y} K_1^z((u_1^*)^z, (u_2^*)^z) = \max_{z \in F_y} v(z). \quad (4.3.1)$$

Для $y \in X_2$ аналогично получаем

$$\begin{aligned} v(y) &= -K_2^y((u_1^*)^y, (u_2^*)^y) = -\max_{z \in F_y} K_2^z((u_1^*)^z, (u_2^*)^z) = \\ &= -\max_{z \in F_y} (-v(z)) = \min_{z \in F_y} v(z). \end{aligned} \quad (4.3.2)$$

Из (4.3.1) и (4.3.2) окончательно имеем

$$v(y) = \max_{z \in F_y} v(z), \quad y \in X_1; \quad (4.3.3)$$

$$v(y) = \min_{z \in F_y} v(z), \quad y \in X_2. \quad (4.3.4)$$

Уравнения (4.3.3), (4.3.4) решаются при граничном условии

$$v(y)|_{y \in X_3} = H_1(y). \quad (4.3.5)$$

Система уравнений (4.3.3), (4.3.4) с граничным условием (4.3.5) позволяет осуществить попятную рекуррентную процедуру нахождения значения игры и оптимальных стратегий игроков. Действительно, пусть значения всех подыгр Γ_z длиной $l(z) \leq k - 1$ известны и равны $v(z)$, пусть Γ_y — некоторая подыгра длиной $l(y) = k$. Тогда, если $y \in X_1$, то $v(y)$ определяется по формуле (4.3.3), если же $y \in X_2$, то $v(y)$ находится по формуле (4.3.4). При этом значения функций $v(z)$ в формулах (4.3.3), (4.3.4) известны, поскольку соответствующие подыгры имеют длину не более чем $k - 1$. Эти формулы указывают способ построения стратегий игроков. Действительно, если $y \in X_1$, то игрок 1 (максимизирующий) должен выбрать в точке y вершину $z \in F_y$, для которой значение следующей подыгры максимально. Если же $y \in X_2$, то игрок 2 (минимизирующий) должен выбрать позицию $z \in F_y$, для которой значение следующей подыгры минимально.

В случае, когда выборы игроков в антагонистической многошаговой игре чередуются (поочередная игра), уравнения (4.3.3), (4.3.4) могут быть записаны в виде одного уравнения. Действительно, рассмотрим подыгру Γ_x и пусть, для определенности, $x \in X_1$. Тогда в следующей позиции ходит игрок 2 или эта позиция является (игра поочередная) окончательной, т. е. $F_x \subset X_2 \cup X_3$. Поэтому можно записать

$$v(x) = \max_{y \in F_x} v(y), \quad x \in X_1; \quad (4.3.6)$$

$$v(y) = \min_{z \in F_y} v(z), \quad y \in F_x \subset X_2 \cup X_3. \quad (4.3.7)$$

Подставляя (4.3.7) в (4.3.6), получаем

$$v(x) = \max_{y \in F_x} [\min_{z \in F_y} v(z)], \quad x \in X_1. \quad (4.3.8)$$

Если $x \in X_2$, то аналогично имеем

$$v(x) = \min_{y \in F_x} [\max_{z \in F_y} v(z)]. \quad (4.3.9)$$

Уравнения (4.3.8), (4.3.9) эквивалентны и должны рассматриваться с начальным условием $v(x)|_{x \in X_3} = H_1(x)$.

4.3.3. Теорема п. 4.2.1 о существовании абсолютного равновесия по Нэшу, рассматриваемая применительно к антагонистическим поочередным многошаговым играм, позволяет утверждать существование ситуации равновесия в «шаматах», «шашках», в классе чистых стратегий, а уравнения (4.3.8), (4.3.9) показывают путь для нахождения значения игры. Вместе с тем очевидно, что никогда в обозримом будущем решение

указанных функциональных уравнений для нахождения значения игры и оптимальных стратегий не будет реализовано на ЭВМ и мы так и не узнаем, может ли какой-либо игрок, «белый» или «черный», гарантировать победу в любой партии или всегда возможна «ничья»? Однако в шахматах и шашках делаются небезуспешные попытки построения приближенно оптимальных решений путем создания программ, думающих на несколько шагов вперед, и использования всевозможных (полученных, как правило, эмпирическим путем) функций оценки текущих позиций. Такой подход возможен и при исследовании общих антагонистических многошаговых игр с полной информацией. Последовательное итерирование оценочных функций на несколько шагов вперед может привести к желаемым результатам.

§ 4.4. Иерархические игры

Важнейшим подклассом неантагонистических многошаговых игр являются иерархические игры. Иерархические игры моделируют конфликтно-управляемые системы с иерархической структурой. Такая структура определяется последовательностью уровней управления, следующих друг за другом в порядке определенного приоритета. В математической постановке иерархические игры классифицируются по числу уровней и характеру вертикальных связей. Простейшей из них является двухуровневая система, схема которой изображена на рис. 4.2.

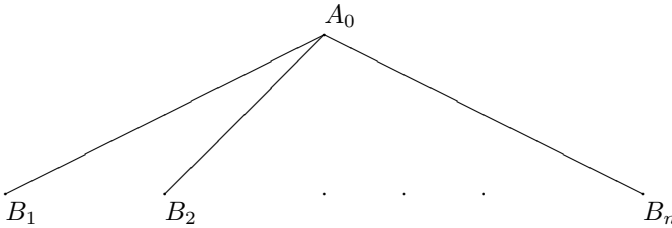


Рис. 4.2. Древоподобная структура управления

4.4.1. *Двухуровневая конфликтно управляемая система* функционирует следующим образом. Управляющий (координирующий) центр A_0 , находящийся в первом уровне иерархии, выбирает вектор $u = (u_1, \dots, u_n)$ из заданного множества управлений U , где u_i — управляющее воздействие центра на подчиненные ему подразделения B_i , $i = 1, \dots, n$, находящиеся на втором уровне иерархии. В свою очередь, B_i , $i = 1, \dots, n$ выбирают управления $v_i \in V_i(u_i)$, где $V_i(u_i)$ — множество управлений подразделения B_i , предопределенное управлением u центра A_0 . Таким образом, управляющий центр имеет право первого хода и может ограничивать возможности подчиненных ему подразделений, направляя их действия в нужное русло. Цель центра A_0 заключается в максимизации по u функционала $K_0(u, v_1, \dots, v_n)$, а подразделения B_i , $i = 1, \dots, n$, обладая собственными целями, стремятся максимизировать по v_i функционалы $K_i(u_i, v_i)$.

4.4.2. Формализуем эту задачу как бескоалиционную игру $(n+1)$ -го лица Γ (административного центра A_0 и производственных подразделений B_1, \dots, B_n) в нормальной форме.

Пусть игрок A_0 выбирает вектор $u \in U$, где

$$U = \{u = (u_1, \dots, u_n) : u_i \geq 0, u_i \in R^l, i = 1, \dots, n, \sum_{i=1}^n u_i \leq b\}, b \geq 0$$

— множество стратегий игрока A_0 в игре Γ . Вектор u_i будем интерпретировать как набор ресурсов l наименований, выделяемых центром A_0 для i -го производственного подразделения.

Пусть в исходной задаче п. 4.4.4 каждый из игроков B_i , зная выбор A_0 , выбирает вектор $v_i \in V_i(u_i)$, где

$$V_i(u_i) = \{v_i \in R^m : v_i A_i \leq u_i + \alpha_i, v_i \geq 0\}. \quad (4.4.1)$$

Вектор v_i , интерпретируется как производственная программа i -го производственного подразделения по различным видам продукции; A_i — производственная или технологическая матрица i -го производственного подразделения ($A_i \geq 0$); α_i — вектор наличных ресурсов i -го производственного подразделения ($\alpha_i \geq 0$).

Под стратегиями игрока B_i в игре Γ будем понимать множество функций $v_i(\cdot)$, ставящих в соответствие каждому элементу $u_i : (u_1, \dots, u_i, \dots, u_n) \in U$ вектор $v_i(u_i) \in V_i(u_i)$. Множество таких функций будем обозначать через V_i , $i = 1, \dots, n$.

Определим функции выигрышей игроков в игре Γ . Для игрока A_0 функция выигрыша имеет вид

$$K_0(u, v_1(\cdot), \dots, v_n(\cdot)) = \sum_{i=1}^n a_i v_i(u_i),$$

где $a_i \geq 0$, $a_i \in R^m$ — фиксированный вектор, $i = 1, \dots, n$; а $a_i v_i(u_i)$ — скалярное произведение векторов a_i и $v_i(u_i)$. Функцию выигрыша игрока B_i полагаем равной

$$K_i(u, v_1(\cdot), \dots, v_n(\cdot)) = c_i v_i(u_i),$$

где $c_i \geq 0$, $c_i \in R^m$ — фиксированный вектор, $i = 1, \dots, n$.

Таким образом, игра Γ имеет вид $\Gamma = (U, V_1, \dots, V_n, K_0, K_1, \dots, K_n)$.

4.4.3. Построим ситуацию равновесия по Нэшу в игре Γ . Пусть $v_i^*(u_i) \in V_i(u_i)$ — решение задачи параметрического линейного программирования (параметром является вектор u_i)

$$\max_{v_i \in V_i(u_i)} c_i v_i = c_i v_i^*(u_i), \quad i = 1, \dots, n, \quad (4.4.2)$$

и пусть $u^* \in U$ — решение задачи

$$\max_{u \in U} K_0(u, v_1^*(\cdot), \dots, v_n^*(\cdot)). \quad (4.4.3)$$

Для простоты предполагаем, что максимумы в (4.4.2) и (4.4.3) достигаются. Заметим, что (4.4.3) — задача нелинейного программирования с существенно разрывной целевой функцией (максимизация ведется по u , и $v_i^*(u_i)$, вообще говоря, — разрывные функции параметра u_i). Покажем, что точка $(u^*, v_1^*(\cdot), \dots, v_n^*(\cdot))$ является ситуацией равновесия в игре Γ . Действительно,

$$K_0(u^*, v_1^*(\cdot), \dots, v_n^*(\cdot)) \geq K_0(u, v_1^*(\cdot), \dots, v_n^*(\cdot)), \quad u \in U.$$

Далее, при всех $i = 1, \dots, n$ справедливо неравенство

$$\begin{aligned} K_i(u^*, v_1^*(\cdot), \dots, v_n^*(\cdot)) &= c_i v_i^*(u_i^*) \geq c_i v_i(u_i^*) = \\ &= K_i(u^*, v_1^*(\cdot), \dots, v_{i-1}^*(\cdot), v_i(\cdot), v_{i+1}^*(\cdot), \dots, v_n^*(\cdot)) \end{aligned}$$

для любой $v_i(\cdot) \in V_i$. Таким образом, никому из игроков A_0, B_1, \dots, B_n невыгодно в одностороннем порядке отклоняться от ситуации $(u^*, v_1^*(\cdot), \dots, v_n^*(\cdot))$, т. е. она является равновесной. Заметим, что эта ситуация также устойчива против отклонения от нее любой коалиции $S \subset \{B_1, \dots, B_n\}$, поскольку выигрыш K_i i -го игрока не зависит от стратегий $v_j(\cdot)$, $j \in \{1, \dots, n\}$, $j \neq i$.

§ 4.5. Иерархические игры (кооперативный вариант)

В этом параграфе рассматривается кооперативный вариант ряда простейших иерархических игр (в том числе игры, определенной в п. 4.5.1, 4.5.2). Строятся характеристические функции и исследуются условия существования непустого C -ядра.

4.5.1. Исходя из содержательного смысла задачи п. 4.4.1, 4.4.2 и с использованием стратегий, образующих равновесие по Нэшу, для каждой коалиции $S \subset N = \{A_0, B_1, \dots, B_n\}$ определим ее гарантированный доход $v(S)$ следующим образом:

$$v(S) = \begin{cases} 0, & \text{при } S = \{A_0\}, \\ \sum_{i: B_i \in S} c_i v_i^*(0), & \text{при } A_0 \notin S, \\ \max_{\{u \in U: \sum_{i: B_i \in S} u_i = b\}} \sum_{i: B_i \in S} (a_i + c_i) v_i^*(u_i), & \text{при } A_0 \in S, \end{cases} \quad (4.5.1)$$

где $v_i^*(u_i)$, $i = 1, \dots, n$ — решение задачи параметрического линейного программирования (4.4.2). Первое из (4.5.1) имеет место, поскольку коалиция $\{B_1, \dots, B_n\}$ может добиться получения нулевого выигрыша игроком A_0 , выбирая все $v_i = 0$, $i = 1, \dots, n$; второе справедливо, так как игрок A_0 всегда может гарантировать для S выигрыш не более чем $\sum_{i: B_i \in S} c_i v_i^*(0)$, направляя каждому $B_i \in S$ нулевой ресурс; третье равенство (4.5.1) имеет место, поскольку коалиция S , содержащая в своем составе A_0 , всегда может обеспечить распределение всего ресурса только между своими членами.

Пусть S — произвольная коалиция, содержащая A_0 . Обозначим через $u^S = (u_1^S, \dots, u_n^S)$ вектор, доставляющий максимум в задаче нелинейного программирования

$$v(S) = \max_{\{u \in U: \sum_{i: B_i \in S} u_i = b\}} \sum_{i: B_i \in S} (a_i + c_i) v_i^*(u_i)$$

(для $i: B_i \notin S$ выполнено условие $u_i^S = 0$). Тогда для любой коалиции $\bar{S} \subset S$, $\bar{S} \neq A_0$, $A_0 \in S$ справедливо следующее неравенство:

$$\begin{aligned} \sum_{i: B_i \in S} (a_i + c_i) v_i^*(u_i^S) &\geq \sum_{i: B_i \in S} (a_i + c_i) v_i^*(u_i^{\bar{S}}) = \\ &= \sum_{i: B_i \in \bar{S}} (a_i + c_i) v_i^*(u_i^{\bar{S}}) + \sum_{i: B_i \in S \setminus \bar{S}} (a_i + c_i) v_i^*(0). \end{aligned}$$

Пусть $S, R \subset N$, $S \cap R = \emptyset$ и $A_0 \in S \neq A_0$. Тогда $A_0 \notin R$. Принимая во внимание условия $a_i \geq 0$, $c_i \geq 0$, $v_i \geq 0$, $i = 1, \dots, n$, имеем

$$\begin{aligned} v(S \cup R) &= \sum_{i: B_i \in S \cup R} (a_i + c_i) v_i^*(u_i^{S \cup R}) \geq \sum_{i: B_i \in S \cup R} (a_i + c_i) v_i^*(u_i^S) = \sum_{i: B_i \in S} (a_i + c_i) v_i^*(u_i^S) + \\ &+ \sum_{i: B_i \in R} (a_i + c_i) v_i^*(0) = v(S) + v(R) + \sum_{i: B_i \in R} a_i v_i^*(0) \geq v(S) + v(R), \end{aligned}$$

где $\sum_{i: B_i \in R} a_i v_i^*(0) \geq 0$ — выигрыш центра A_0 от «нефинансируемых» предприятий. В случаях $A_0 \notin S \cup R$ или $S = A_0 \notin R$ неравенство $v(S \cup R) \geq v(S) + v(R)$ очевидно.

Таким образом, функция $v(S)$, определяемая (4.5.1), супераддитивна и можно рассмотреть кооперативную игру $(\{A_0, B_1, \dots, B_n\}, v)$ в форме характеристической функции.

4.5.2. Рассмотрим $(n+1)$ -мерный вектор

$$\xi = (\sum_{i=1}^n a_i v_i^*(\bar{u}_i), c_1 v_1^*(\bar{u}_1), \dots, c_n v_n^*(\bar{u}_n)), \quad (4.5.2)$$

где $\bar{u} = u^N$. Вектор ξ является дележом, поскольку выполнены следующие соотношения:

$$1) \sum_{k=0}^n \xi_k = \sum_{i=1}^n (a_i + c_i) v_i^*(\bar{u}_i) = v(N);$$

$$2) \xi_0 = \sum_{i=1}^n a_i v_i^*(\bar{u}_i) \geq 0 = v(A_0),$$

$$\xi_i = c_i v_i^*(\bar{u}_i) \geq c_i v_i^*(0) = v(B_i), \quad i = 1, \dots, n.$$

Напомним условие принадлежности дележа С-ядру. Согласно теореме п. 3.12.1 необходимым и достаточным условием принадлежности дележа $(\xi_0, \xi_1, \dots, \xi_n)$ С-ядру является выполнение неравенства

$$\sum_{i \in S} \xi_i \geq v(S) \quad (4.5.3)$$

для всех коалиций $S \subset \{A_0, B_1, \dots, B_n\}$.

Выведем условие, при котором дележ ξ принадлежит С-ядру. Если $S = \{A_0\}$ или $S \subset \{B_1, \dots, B_n\}$, то условие (4.5.21) выполнено, поскольку

$$\xi_0 = \sum_{i=1}^n a_i v_i^*(\bar{u}_i) \geq 0 = v(\{A_0\}),$$

$$\sum_{i \in S} \xi_i = \sum_{i: B_i \in S} c_i v_i^*(\bar{u}_i) \geq \sum_{i: B_i \in S} c_i v_i^*(0) = v(S).$$

Если $A_0 \in S \neq A_0$, то условие (4.5.21) можно записать в виде

$$\begin{aligned} \sum_{i=1}^n a_i v_i^*(\bar{u}_i) + \sum_{i: B_i \in S} c_i v_i^*(\bar{u}_i) &= \sum_{i: B_i \in S} a_i v_i^*(\bar{u}_i) + \sum_{i: B_i \in S} c_i v_i^*(\bar{u}_i) + \\ &+ \sum_{i: B_i \notin S} a_i v_i^*(\bar{u}_i) \geq \sum_{i: B_i \in S} (a_i + c_i) v_i^*(u_i^S). \end{aligned}$$

Следовательно, дележ (4.5.2) принадлежит C -ядру, если для всех $S : A_0 \in S$, выполнено неравенство

$$\sum_{i: B_i \notin S} a_i v_i^*(\bar{u}_i) \geq \sum_{i: B_i \in S} (a_i + c_i) [v_i^*(u_i^S) - v_i^*(\bar{u}_i)].$$

Заметим, что в данном случае мы определили характеристическую функцию игры, используя выигрыш в ситуации равновесия по Нэшу, и величина $v(N) = \max_u \sum_{i=1}^n (a_i + c_i) v_i^*(u_i)$, вообще говоря, меньше максимального суммарного выигрыша всех игроков, равного

$$\max_{u \in U} \max_{v_k \in V_k(u_k)} \left[\sum_{k=1}^n (a_k + c_k) v_k(u_k) \right],$$

(в этом отличие от принятого в гл. 3 определения характеристической функции).

4.5.3. Характеристическую функцию игры можно построить и обычным способом, а именно: для каждой коалиции S определить ее как значение антагонистической игры между этой коалицией и коалицией остальных игроков $N \setminus S$. Построим теперь характеристическую функцию именно таким образом. При этом несколько обобщим предыдущую задачу, введя в рассмотрение произвольные функции выигрышей участников игры. Как и ранее, будем предполагать, что центр A_0 распределяет ресурсы между подразделениями B_1, \dots, B_n , которые используют эти ресурсы для производства продукции. Выигрыши управляющего центра A_0 и «производственных» подразделений B_1, \dots, B_n зависят от продукции, производимой B_1, \dots, B_n . Вектор ресурсов, имеющийся в распоряжении центра A_0 , обозначим через b . Центр (игрок) A_0 выбирает систему n векторов $u = (u_1, \dots, u_n)$ из множества

$$U = \{u = (u_1, \dots, u_n) : u_k \geq 0, u_k \in R^l, \sum_{k=1}^n u_k \leq b, k = 1, \dots, n\}.$$

Здесь u_k интерпретируется как вектор ресурса, выделяемый центром A_0 производственному подразделению B_k . Возможности предприятия (игрока) B_k определяются ресурсом u_k , получаемым от A_0 , т. е. предприятие B_k выбирает свою производственную программу x_k из множества $B_k(u_k) \subset R^m$ неотрицательных векторов. Будем предполагать, что множества $B_k(u_k)$ при всех u_k содержат нулевой вектор и монотонно возрастают по включению, т. е. из $u'_k > u_k$ следует $B_k(u'_k) \supset B_k(u_k)$. Кроме того, выполнено условие $B_k(0) = \emptyset$ (невозможность производства при отсутствии ресурсов).

Пусть $x = (x_1, \dots, x_n)$. Выигрыш игрока A_0 определяется с помощью неотрицательной функции $l_0(x) \geq 0$, а выигрыши игроков B_k полагаем равными $l_k(x_k) \geq 0$, $k = 1, \dots, n$ (выигрыш игрока B_k зависит лишь от производственной программы). Для простоты будем считать, что выигрыш центра A_0 удовлетворяет условию

$$l_0(x) = \sum_{k=1}^n l_k(x_k),$$

где слагаемое $l(x_k)$ интерпретируется, как выигрыш игрока A_0 , получаемый от игрока B_k . Предположим также, что $l(x_k) \geq 0$ для всех $x_k \in B_k(u_k)$ и $l_k(0) = 0$, $l(0) = 0$, $k = 1, \dots, n$.

Подобно тому, как это сделано в § 4.5, представим иерархическую игру п. 4.5.3 в виде бескоалиционной игры $(n+1)$ лица в нормальной форме, где стратегиями игрока

A_0 будут векторы $u \in U$, а стратегиями игроков B_k — функции из соответствующих множеств. Построим характеристическую функцию $\bar{v}(\cdot)$ этой игры, следуя п. 3.11.2. Для каждого подмножества S игроков $\bar{v}(S)$ будет равно значению (оно существует в условиях п. 4.5.3) антагонистической игры между коалициями S и $N \setminus S$, в которой выигрыш коалиции S определяется как сумма выигрышей, принадлежащих множеству S игроков. Пусть $N = \{A_0, B_1, \dots, B_n\}$. Тогда

$$\bar{v}(N) = \sup_{\{u \in U : \sum_{k=1}^n u_k = b\}} \sup_{x_k \in B_k(u_k), k=1, \dots, n} \left\{ \sum_{k=1}^n [l(x_k) + l_k(x_k)] \right\}.$$

Заметим, что для всех $S \subset \{B_1, \dots, B_n\}$, $\bar{v}(S) = 0$, поскольку игрок A_0 всегда может распределить весь ресурс b среди членов коалиции $N \setminus S$, в которую он входит, лишив, таким образом, коалицию S ресурсов (т. е. A_0 всегда может положить $u_k \equiv 0$ для $k : B_k \in S$, что приводит к $B_k(0) = \emptyset$ для всех $B_k \in S$). Рассуждая аналогично, имеем $\bar{v}(A_0) = 0$, поскольку игроки B_1, \dots, B_n всегда могут сделать выигрыш центра A_0 равным нулю, полагая $x_k = 0$ для $k = 1, \dots, n$ (не производя продукции). В том случае, когда коалиция S содержит центр A_0 , очевидно, что A_0 будет распределять весь ресурс среди членов коалиции. Это соображение приводит к следующей формуле:

$$\bar{v}(S) = \sup_{\{u \in U : \sum_{k: B_k \in S} u_k = b\}} \sup_{x_k \in B_k(u_k), k: B_k \in S} \left\{ \sum_{k: B_k \in S} [l(x_k) + l_k(x_k)] \right\}$$

для $S : A_0 \in S$.

Можно показать, что при таком определении характеристической функции, C -ядро множества дележей

$$\{\alpha = (\alpha_0, \alpha_1, \dots, \alpha_n) : \alpha_i \geq 0, i = 0, 1, \dots, n, \sum_{i=0}^n \alpha_i = \bar{v}(N)\}$$

всегда непусто.

4.5.4. Иерархические системы с подразделениями двойного подчинения называются ромбовидными (рис. 4.3). Управление подразделения двойного подчинения C зависит от управления B_1 и от управления B_2 .

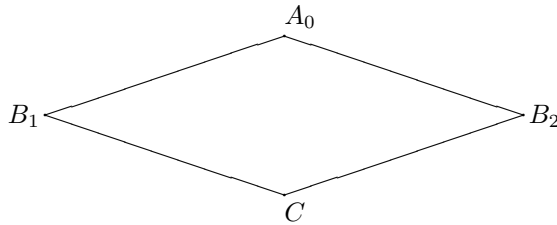


Рис. 4.3. Ромбовидная структура управления

Можно представить ситуацию, в которой центр B_1 представляет интересы отрасли, а B_2 — региональные интересы, включающие вопросы охраны окружающей среды. Простая ромбовидная система управления является примером иерархической системы

с тремя уровнями принятия решений. На высшем уровне находится административный центр, располагающий материальными и трудовыми ресурсами. Он воздействует на деятельность двух подчиненных ему центров, принадлежащих следующему уровню. От решений, принимаемых этими центрами, зависит объем производства предприятия, находящегося на последнем уровне иерархической системы. Будем рассматривать этот процесс принятия решений, как некоторую игру четырех лиц. Обозначим ее через Γ . Переходя к игровой постановке, условимся считать, что на 1-м шаге ходит игрок A_0 и выбирает элемент (стратегию) $u = (u_1, u_2)$ из некоторого множества U , где U — множество стратегий игрока A_0 . Элемент $u \in U$ ограничивает возможности выборов игроков B_1 и B_2 на следующем шаге. Другими словами, множество выборов игрока B_1 оказывается функцией параметра u_1 (обозначим его через $B_1(u_1)$), и, аналогично, множество выборов игрока B_2 оказывается функцией параметра u_2 (обозначим его через $B_2(u_2)$). Через $\omega_1 \in B_1(u_1)$ и $\omega_2 \in B_2(u_2)$ обозначим элементы множества выборов игроков B_1 и B_2 соответственно. Параметры ω_1 и ω_2 , выбираемые игроками B_1 и B_2 , задают ограничения на множество выборов игрока C на 3-м шаге игры, т. е. это множество оказывается функцией параметров ω_1 и ω_2 . Обозначим его через $C(\omega_1, \omega_2)$, а элементы этого множества (производственные программы) — через v .

Пусть выигрыши всех игроков A_0, B_1, B_2, C зависят только от производственной программы v , выбираемой игроком C , и равны соответственно $l_1(v), l_2(v), l_3(v), l_4(v)$, где $l_i(v) \geq 0$.

Такую иерархическую игру можно представить как бескоалиционную игру четырех лиц в нормальной форме, если считать стратегиями игрока A_0 элементы $u = (u_1, u_2) \in U$, а стратегиями игроков B_1, B_2 и C — функции $\omega_1(u_1), \omega_2(u_2)$ и $v(\omega_1, \omega_2)$ со значениями в множествах $B_1(u_1), B_2(u_2), C(\omega_1, \omega_2)$, соответственно (обозначим множества таких функций через B_1, B_2, C) которые каждому возможному выбору игрока (или игроков), находящегося на более высоком уровне, ставят в соответствие выбор данного игрока. Полагая

$$K_i(u, \omega_1(\cdot), \omega_2(\cdot), v(\cdot)) = l_i(v(\omega_1(u_1), \omega_2(u_2))), \quad i = \overline{1, 4},$$

получим нормальную форму игры Γ

$$\Gamma = (U, B_1, B_2, C, K_1, K_2, K_3, K_4).$$

4.5.5. Будем искать ситуацию равновесия по Нэшу в игре Γ . Для этого выполним вспомогательные построения. Для каждой фиксированной пары (ω_1, ω_2) , $(\omega_1, \omega_2) \in \cup_{u \in U} B_1(u_1) \times B_2(u_2)$ обозначим через $v^*(\omega_1, \omega_2)$ решение параметрической экстремальной задачи

$$\max_{v \in C(\omega_1, \omega_2)} l_4(v) = l_4(v^*(\omega_1, \omega_2)). \quad (4.5.4)$$

(Считаем, что максимум в (4.5.4) достигается.) Решение $v^*(\cdot) = v^*(\omega_1, \omega_2)$ задачи (4.5.4) оказывается функцией параметров ω_1, ω_2 и $v^*(\cdot) \in C$.

Рассмотрим вспомогательную параметрическую (с параметрами u_1, u_2) неантагонистическую игру $\Gamma'(u_1, u_2) = \{B_1(u_1), B_2(u_2), l_2, l_3\}$, где $l_2 = l_2(v^*(\omega_1, \omega_2))$, $l_3 = l_3(v^*(\omega_1, \omega_2))$. Стратегиями игрока B_1 в $\Gamma'(u_1, u_2)$ являются элементы $\omega_1 \in B_1(u_1)$, стратегиями игрока B_2 являются элементы $\omega_2 \in B_2(u_2)$. Предположим, что в игре $\Gamma'(u_1, u_2)$ существует ситуация равновесия по Нэшу, которую обозначим $(\omega_1^*(u_1), \omega_2^*(u_2))$. Отметим, что $\omega_i^*(\cdot)$ является функцией параметра u_i и $\omega_i^*(\cdot) \in B_i$, $i = 1, 2$.

Пусть, далее, $u^* = (u_1, u_2)$ — решение следующей экстремальной задачи:

$$\max_{u \in U} l_1(v^*(\omega_1^*(u_1), \omega_2^*(u_2))). \quad (4.5.5)$$

Лемма. Совокупность $(u^*, \omega_1^*(\cdot), \omega_2^*(\cdot), v^*(\cdot))$ является ситуацией равновесия по Нэшу в игре Γ .

Доказательство. Согласно определению u^* из (4.5.5) следует соотношение

$$\begin{aligned} K_1(u^*, \omega_1^*(\cdot), \omega_2^*(\cdot), v^*(\cdot)) &= \max_{u \in U} l_1(v^*(\omega_1^*(u_1), \omega_2^*(u_2))) \geq \\ &\geq l_1(v^*(\omega_1^*(u_1), \omega_2^*(u_2))) = K_1(u, \omega_1^*(\cdot), \omega_2^*(\cdot), v^*(\cdot)) \end{aligned}$$

для всех $u \in U$. Поскольку $\omega_1^*(u_1^*), \omega_2^*(u_2^*)$ образуют ситуацию равновесия по Нэшу во вспомогательной игре $\Gamma(u_1^*, u_2^*)$, для любой функции $\omega_1(\cdot) \in B_1$, $\omega_1(u^*) = \tilde{\omega}_1 \in B_1(u_1^*)$ выполняются соотношения

$$\begin{aligned} K_2(u^*, \omega_1^*(\cdot), \omega_2^*(\cdot), v^*(\cdot)) &= l_2(v^*(\omega_1^*(u_1), \omega_2^*(u_2))) \geq \\ &\geq l_2(v^*(\tilde{\omega}_1(u_1^*), \omega_2^*(u_2^*))) = K_2(u^*, \omega_1(\cdot), \omega_2^*(\cdot), v^*(\cdot)) \end{aligned}$$

Аналогичное неравенство справедливо и для игрока B_2 .

По определению функции v^* из (4.5.4) имеем:

$$\begin{aligned} K_4(u^*, \omega_1^*(\cdot), \omega_2^*(\cdot), v^*(\cdot)) &= l_4(v^*(\omega_1^*(u_1^*), \omega_2^*(u_2^*))) = \\ &= \max_{v \in C(\omega_1^*(u_1^*), \omega_2^*(u_2^*))} l_4(v) \geq l_4(\tilde{v}) = K_4(u^*, \omega_1^*(\cdot), \omega_2^*(\cdot), v(\cdot)) \end{aligned}$$

для любой функции $v(\cdot) \in C$, $v(\omega_1^*(u_1^*), \omega_2^*(u_2^*)) = \tilde{v} \in C(\omega_1^*(u_1^*), \omega_2^*(u_2^*))$.

Лемма доказана.

4.5.6. Применяя максиминный подход, для каждой коалиции $S \subset \{A_0, B_1, B_2, C\}$ определим $v'(S)$ как наибольший гарантированный выигрыш S в антагонистической игре между коалицией S , выступающей в качестве максимизирующего игрока, и коалицией $S' = \{A_0, B_1, B_2, C\} \setminus S$. Предположим, что существует такое $v_0 \in C(\omega_1, \omega_2)$ для всех ω_1, ω_2 , что $l_i(v_0) = 0$, $i = 1, 2, 3, 4$.

Будем различать два вида коалиций: 1) $S : C \notin S$; 2) $S : C \in S$.

В первом случае $S \subset \{A_0, B_1, B_2\}$ и игрок C , являющийся членом коалиции $N \setminus S$, может выбрать стратегию $v_0 : l_i(v_0) = 0$, $i = 1, 2, 3, 4$. Поэтому $v'(S) = 0$.

Во втором случае определим характеристическую функцию $v'(S)$ следующими равенствами:

а) $S = \{C\}$

$$v'(S) = \min_{u \in U} \min_{\omega_1 \in B_1(u_1)} \min_{\omega_2 \in B_2(u_2)} \max_{v \in C(\omega_1, \omega_2)} l_4(v),$$

(здесь и далее предполагаем, что все \max и \min достигаются);

б) $S = \{A_0, C\}$

$$v'(S) = \max_{u \in U} \min_{\omega_1 \in B_1(u_1)} \min_{\omega_2 \in B_2(u_2)} \max_{v \in C(\omega_1, \omega_2)} (l_1(v) + l_4(v));$$

в) $S = \{B_1, C\}$

$$v'(S) = \min_{u \in U} \max_{\omega_1 \in B_1(u_1)} \min_{\omega_2 \in B_2(u_2)} \max_{v \in C(\omega_1, \omega_2)} (l_2(v) + l_4(v));$$

г) $S = \{B_2, C\}$

$$v'(S) = \min_{u \in U} \max_{\omega_2 \in B_2(u_2)} \min_{\omega_1 \in B_1(u_1)} \max_{v \in C(\omega_1, \omega_2)} (l_3(v) + l_4(v));$$

$$д) S = \{B_1, B_2, C\}$$

$$v'(S) = \min_{u \in U} \max_{\omega_1 \in B_1(u_1)} \max_{\omega_2 \in B_2(u_2)} \max_{v \in C(\omega_1, \omega_2)} \sum_{i=2,3,4} l_i(v);$$

$$е) S = \{A_0, B_1, C\}$$

$$v'(S) = \max_{u \in U} \max_{\omega_1 \in B_1(u_1)} \min_{\omega_2 \in B_2(u_2)} \max_{v \in C(\omega_1, \omega_2)} \sum_{i=1,2,4} l_i(v);$$

$$ж) S = \{A_0, B_2, C\}$$

$$v'(S) = \max_{u \in U} \max_{\omega_2 \in B_2(u_2)} \min_{\omega_1 \in B_1(u_1)} \max_{v \in C(\omega_1, \omega_2)} \sum_{i=1,3,4} l_i(v);$$

$$з) S = \{A_0, B_1, B_2, C\}$$

$$v'(S) = \max_{u \in U} \max_{\omega_1 \in B_1(u_1)} \max_{\omega_2 \in B_2(u_2)} \max_{v \in C(\omega_1, \omega_2)} \sum_{i=1}^4 l_i(v).$$

При таком определении характеристическая функция обладает свойством супераддитивности, т. е. для любых $S, R \subset \{A_0, B_1, B_2, C\}$, для которых $S \cap R = \emptyset$, имеет место неравенство $v(S \cup R) \geq v(S) + v(R)$.

§ 4.6. Многошаговые игры с неполной информацией

4.6.1. В предыдущих параграфах рассматривались многошаговые игры с полной информацией, определенные на конечном древовидном графе $G = (X, F)$, в которых каждый из игроков в момент совершения своего хода точно знал, в какой позиции или в какой вершине дерева он находится. Именно поэтому удалось ввести понятие стратегии игрока i , как однозначной функции $u_i(x)$, определенной на множестве очередности X_i со значениями в множестве F_x . Однако если попытаться исследовать многошаговую игру, в которой игроки при совершении своих выборов не знают точно позиции, в которой они совершают ход, или могут лишь предполагать, что эта позиция принадлежит некоторому подмножеству A множества очередности X_i , то реализация стратегии игрока как функции от позиции $x \in X_i$ окажется невозможной. Таким образом, желание усложнить информационную структуру игры неизбежно приводит к изменению понятия стратегии. Для точных формулировок необходимо в первую очередь формализовать понятие информации в игре. Важную роль здесь играет понятие информационного множества. Проиллюстрируем это на нескольких простейших, ставших классическими в учебной литературе по теории игр, примерах [Мак—Кинси, 1960].

Пример 2 (Игра антагонистическая). Делая 1-й ход, игрок 1 выбирает число из множества $\{1, 2\}$. Второй ход делает игрок 2. Зная выбор игрока 1, он выбирает число из множества $\{1, 2\}$. Третий ход опять делает игрок 1. Зная выбор игрока 2 и помня свой выбор, он выбирает число из множества $\{1, 2\}$. На этом игра прекращается, и игрок 1 получает выигрыш H (игрок 2 получает выигрыш $(-H)$, т. е. игра антагонистическая).

Функция H определяется следующим образом:

$$H(1, 1, 1) = -3, \quad H(2, 1, 1) = 4, \quad (4.6.1)$$

$$H(1, 1, 2) = -2, \quad H(2, 1, 2) = 1, \quad (4.6.2)$$

$$H(1, 2, 1) = 2, \quad H(2, 2, 1) = 1, \quad (4.6.3)$$

$$H(1, 2, 2) = -5, \quad H(2, 2, 2) = 5 \quad (4.6.4)$$

Граф $G = (X, F)$ игры изображен на рис. 4.4. Кругами на графе изображены позиции, в которых ходит игрок 1, а квадратами — позиции, в которых ходит игрок 2.

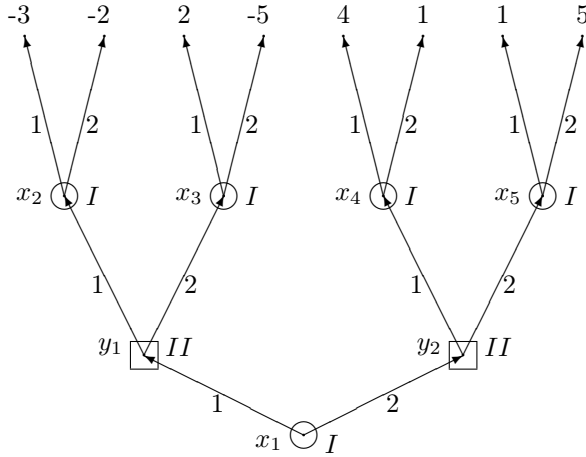


Рис. 4.4. Дерево антагонистической игры

Если множество X_1 обозначить через X , множество X_2 — через Y и элементы этих множеств соответственно — через $x \in X$, $y \in Y$, то стратегия игрока 1, $u_1(\cdot)$, задается пятимерным вектором $u_1(\cdot) = \{u_1(x_1), u_1(x_2), u_1(x_3), u_1(x_4), u_1(x_5)\}$, предписывающим выбор одного из двух чисел $\{1, 2\}$ в каждой позиции множества X . Аналогично стратегия $u_2(\cdot)$ игрока 2 представляет собой двумерный вектор $u_2(\cdot) = \{u_2(y_1), u_2(y_2)\}$, предписывающий выбор одного из двух чисел $\{1, 2\}$ в каждой из позиций множества Y . Таким образом, у игрока 1 в этой игре 32 стратегии, а у игрока 2 — 4 стратегии. Соответствующая нормальная форма игры имеет матрицу размера 32×4 , которая однако (это следует из теоремы п. 2.1), имеет ситуацию равновесия в чистых стратегиях. Можно убедиться, что значение рассматриваемой игры равно 4. Игрок 1 имеет четыре оптимальные чистые стратегии: $(2, 1, 1, 1, 2)$, $(2, 1, 2, 1, 2)$, $(2, 2, 1, 1, 2)$, $(2, 2, 2, 1, 2)$, у игрока 2 — две оптимальные стратегии: $(1, 1)$, $(2, 1)$.

4.6.2. Пример 3. Несколько изменим информационные условия примера 2. Игра антагонистическая. Делая первый ход, игрок 1 выбирает число из множества $\{1, 2\}$. Второй ход делает игрок 2. Зная выбор игрока 1, он выбирает число из множества $\{1, 2\}$. Третий ход делает игрок 1. Не зная выбора игрока 2 и забыв свой выбор, он выбирает число из множества $\{1, 2\}$. На этом игра прекращается и выигрыш определяется по формуле (6.1), так же как и в игре примера 2.

Граф $G = (X, F)$ игры не изменяется, однако, находясь в узлах x_2, x_3, x_4, x_5 (на 3-м ходе игры), игрок 1 не может определить, в каком из этих узлов он на самом деле

находится, но, зная очередность хода (3-й ход), он может быть уверен, что не находится в узле x_1 . На графе G мы обведем узлы x_2, x_3, x_4, x_5 пунктирной линией (рис. 4.5).

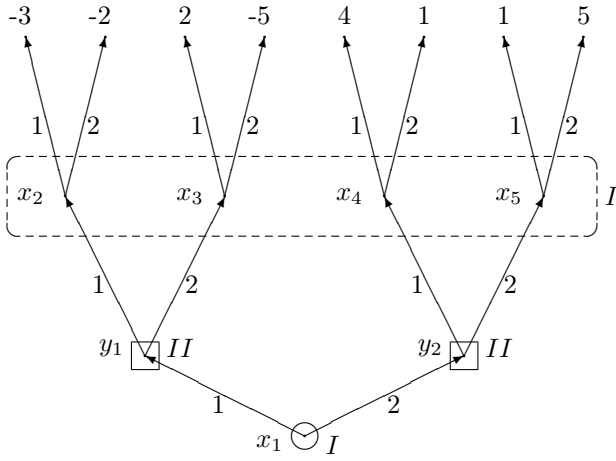


Рис. 4.5. Дерево игры примера 3

В результате узел x_1 оказался обведенным кружком, что можно интерпретировать как точное знание игроком 1 этого узла, когда он в нем находился. Узлы y_1, y_2 обведены квадратами, что также означает, что игрок 2, находясь в одном из них, при совершении своего хода может отличить его от другого. Объединяя узлы x_2, x_3, x_4, x_5 в одно множество, мы иллюстрируем факт их неразличимости для игрока 1.

Множества, на которые разбиты узлы, будем называть *информационными множествами*.

Перейдем теперь к описанию стратегий. Состояние информации игрока 2 не изменилось, поэтому множество его стратегий то же, что и в примере 2, т. е. оно состоит из четырех векторов $(1, 1), (1, 2), (2, 1), (2, 2)$. Информационное состояние игрока 1 изменилось. На 3-м шаге игры он знает лишь номер этого шага, но не знает позиции, в которой находится. Следовательно, он не может реализовать выбор следующей вершины (или выбор числа из множества $\{1, 2\}$) в зависимости от позиции, в которой находится на третьем шаге. Поэтому на 3-м шаге ему остается независимо от в действительности реализовавшейся позиции выбирать одно из двух чисел $\{1, 2\}$. Поэтому его стратегия представляет собой пару чисел (i, j) , $i, j \in \{1, 2\}$, где число i выбирается в позиции x_1 , а число j на 3-м шаге одинаково во всех позициях x_2, x_3, x_4, x_5 . Таким образом, выбор числа j оказывается функцией множества и может быть записан как $u\{x_2, x_3, x_4, x_5\} = j$. В данной игре у обоих игроков по четыре стратегии и матрица игры имеет вид

$$\begin{array}{cc}
 & \begin{matrix} (1.1) & (1.2) & (2.1) & (2.2) \end{matrix} \\
 \begin{matrix} (1.1) \\ (1.2) \\ (2.1) \\ (2.2) \end{matrix} & \begin{bmatrix} -3 & -2 & 2 & -5 \\ 4 & 1 & 1 & 5 \\ 1 & 5 & 1 & 5 \end{bmatrix}
 \end{array}$$

В этой игре нет ситуации равновесия в чистых стратегиях. Значение игры равно $19/7$, оптимальная смешанная стратегия игрока 1 есть вектор $(0, 0, 4/7, 3/7)$, а опти-

мальная смешанная стратегия игрока 2 равна $(4/7, 3/7, 0, 0)$. По сравнению с примером 2 гарантированный выигрыш игрока 1 уменьшается. Это вызвано ухудшением его информационного состояния.

Интересно заметить, что матрица игры примера 3 имеет размер 4×4 , в то время как матрица игры примера 2 имеет размер 32×4 . Таким образом, уменьшение доступной информации уменьшает размер матрицы выигрышей, следовательно, и облегчает решение самой игры, что противоречит распространенному мнению о том, что уменьшение информации приводит к усложнению принятия решений.

Изменяя информационные условия, можно получить другие варианты игры, описанной в примере 2.

4.6.3. Пример 4. Делая первый ход, игрок 1 выбирает число из множества $\{1, 2\}$. Второй ход делает игрок 2, который, не зная выбора игрока 1, выбирает число из множества $\{1, 2\}$. Далее, совершая 3-й ход, игрок 1 выбирает число из множества $\{1, 2\}$, зная выбор игрока 2 и помня свой выбор на первом шаге. Выигрыш определяется так же, как и в примере 2 (рис. 4.6). Поскольку при совершении третьего хода игрок знает позицию, в которой он находится, позиции третьего уровня обведены кружками, два узла, в которых ходит игрок 2, мы обвели штриховой линией, включив их в одно информационное множество.

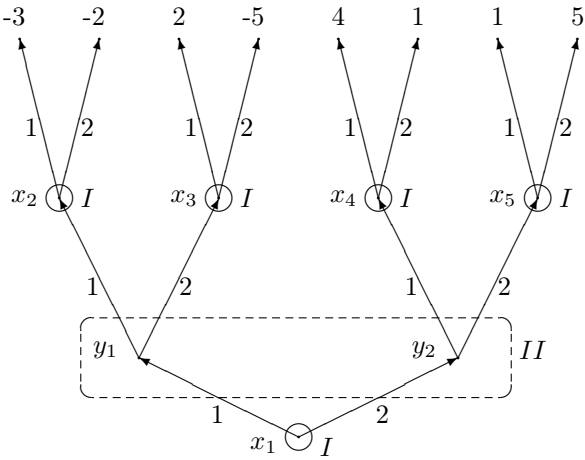


Рис. 4.6. Дерево игры примера 4

Пример 5. Делая первый ход, игрок 1 выбирает число из множества $\{1, 2\}$ Второй ход делает игрок 2, не зная выбора игрока 1. Далее, совершая третий ход, игрок 1 выбирает число из множества $\{1, 2\}$, не зная выбора игрока 2 и не помня свой выбор на 1-м шаге. Выигрыш определяется так же, как в игре из примера 2 (рис. 4.7).

Здесь стратегия игрока 1 состоит из пары чисел (i, j) , где i — выбор на 1-м шаге, а j — на 3-м шаге игры. Стратегия игрока 2 есть выбор числа j на 2-м шаге игры. Таким образом, у игрока 1 — четыре стратегии, а у игрока 2 — две стратегии. Игра в

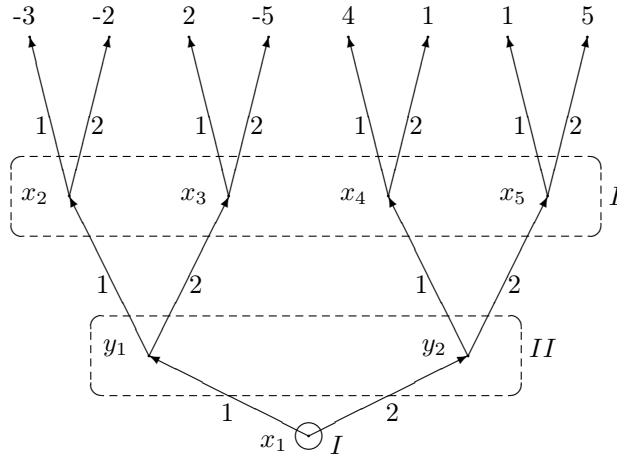


Рис. 4.7. Дерево игры примера 5

нормальной форме имеет матрицу размера 4×2 :

$$\begin{array}{cc}
 & \begin{array}{cc} 1 & 2 \end{array} \\
 \begin{array}{c} (1.1) \\ (1.2) \\ (2.1) \\ (2.2) \end{array} & \begin{bmatrix} -3 & 2 \\ -2 & -5 \\ 4 & 1 \\ 1 & 5 \end{bmatrix}
 \end{array}$$

Значение игры равно $19/7$, оптимальная смешанная стратегия игрока 1 $(0, 0, 4/7, 3/7)$, оптимальная стратегия игрока 2 $(4/7, 3/7)$.

В этой игре значение оказалось таким же, как и в игре из примера 3, т. е. оказалось, что ухудшение информационных условий игрока 2 не улучшило состояние игрока 1. Это обстоятельство в данном случае носит случайный характер и вызвано спецификой функции выигрыша.

4.6.4. Пример 6. В предыдущем примере игроки не различают позиции, находящиеся на одном уровне дерева игры, однако они все таки знают, какой ход совершают. Можно построить игру, в которой игроки проявляют большее незнание.

Рассмотрим антагонистическую игру двух лиц, в которой игрок 1 — один человек, а игрок 2 — команда из двух человек A и B . Все трое изолированы друг от друга (находятся в изолированных помещениях) и не могут общаться между собой. В начале игры посредник входит в помещение, где находится игрок 1, и предлагает ему выбрать число из множества $\{1, 2\}$. Если игрок 1 выбирает 1, то посредник заходит сначала в помещение, где находится A , и предлагает ему выбрать число из множества $\{1, 2\}$, затем заходит к B и предлагает ему сделать выбор из множества $\{1, 2\}$. Если же игрок 1 выбирает 2, то посредник предлагает игроку B сделать выбор первому. После того как три числа выбраны, игрок 1 выигрывает величину $K(x, y, z)$, где x, y, z — выборы игрока 1 и членов команды 2 A и B , соответственно. Функция $K(x, y, z)$ определяется следующим образом:

$$K(1, 1, 1) = 1, \quad K(1, 1, 2) = 3,$$

$$\begin{aligned} K(1, 2, 1) &= 7, \quad K(1, 2, 2) = 9, \\ K(2, 1, 1) &= 5, \quad K(2, 1, 2) = 1, \\ K(2, 2, 1) &= 6, \quad K(2, 2, 2) = 7. \end{aligned}$$

Из правил игры следует, что, когда одному из членов команды *A* или *B* предлагается сделать выбор, он не знает, совершает ли он выбор на 2-м или 3-м шаге игры. Структура игры изображена на рис. 4.8.

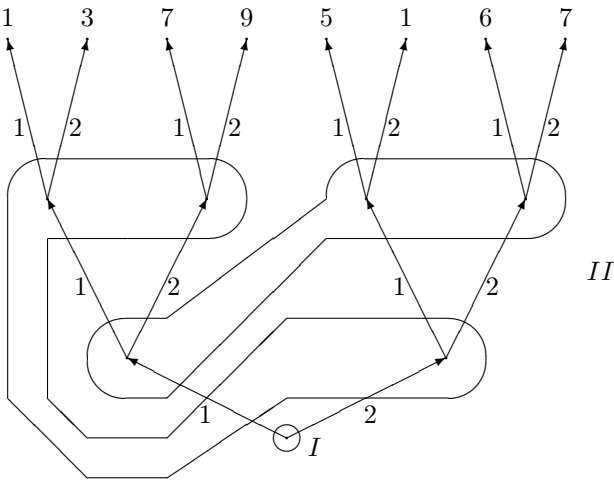


Рис. 4.8. Дерево игры примера 6

Таким образом информационные множества игрока 2 содержат вершины разного уровня, что соответствует незнанию номера хода в игре. Здесь игрок 1 имеет две стратегии. Игрок 2 имеет четыре стратегии, они состоят из всевозможных комбинаций выборов членов команды *A* и *B*, его стратегии суть пары (1, 1), (1, 2), (2, 1), (2, 2).

Для того чтобы понять, как определяются элементы матрицы выигрышей, рассмотрим ситуацию (2, (2, 1)). Так как игрок 1 выбрал 2, то посредник идет в комнату к *B*, который согласно стратегии (2, 1) выбирает 1. Далее он идет к *A*, который выбирает 2. Таким образом, выигрыш в ситуации (2, (2, 1)) равен $K(2, 1, 2) = 1$. Матрица выигрышей для игры в нормальной форме имеет вид

$$\begin{matrix} & \begin{matrix} (1.1) & (1.2) & (2.1) & (2.2) \end{matrix} \\ \begin{matrix} 1 \\ 2 \end{matrix} & \begin{bmatrix} 1 & 3 & 7 & 9 \\ 5 & 6 & 1 & 7 \end{bmatrix} \end{matrix}.$$

Значение игры равно $17/5$, и оптимальные смешанные стратегии игроков 1 и 2 соответственно равны $(2/5, 3/5)$, $(3/5, 0, 2/5, 0)$.

Заметим, что в многошаговых играх с полной информацией (см. теорему §4.2) существует ситуация равновесия по Нэшу в классе чистых стратегий, а в случае антагонистических многошаговых игр — просто ситуация равновесия в чистых стратегиях. Вместе с тем во всех играх с неполной информацией, рассмотренных в примерах 2 – 6, ситуации равновесия в чистых стратегиях не существует.

4.6.5. Дадим теперь формальное определение *многошаговой позиционной игры*.

Определение [Kuhn, 1953]. Многошаговая позиционная игра n лиц определяется:

- 1) Заданием древовидного графа $G = (X, F)$ с начальной вершиной x_0 , называемой начальной позицией игры.
- 2) Разбиением множества всех вершин X на $n + 1$ множество $X_1, X_2, \dots, X_n, X_{n+1}$, где множество X_i называется множеством очередности i -го игрока, $i = 1, \dots, n$, а множество $X_{n+1} = \{x : F_x = \emptyset\}$ — множеством окончательных позиций.
- 3) Заданием вектор-функции $K(x) = (K_1(x), \dots, K_n(x))$ на множестве окончательных позиций $x \in X_{n+1}$; функция $K_i(x)$ называется выигрышем i -го игрока.
- 4) Подразбиением каждого множества X_i , $i = 1, \dots, n$ на непересекающиеся подмножества X_i^j , называемые информационными множествами i -го игрока. При этом для любых позиций одного и того же информационного множества множество следующих за ними вершин должно содержать одно и то же число вершин, т. е. для любых $x, y \in X_i^j$ $|F_x| = |F_y|$ ($|F_x|$ — число элементов множества F_x), и никакая вершина информационного множества не должна следовать за некоторой другой вершиной этого же множества, т. е. если $x \in X_i^j$, то не существует другой вершины $y \in X_i^j$ такой, что $y \in \hat{F}_x$.

Определение многошаговой игры с полной информацией отличается от приведенного здесь лишь условием 4, где вводятся дополнительные разбиения множеств очередности игроков X_i на информационные множества. Как видно из примеров, содержательный смысл такого разбиения заключается в том, что при совершении своего хода в позиции $x \in X_i$, игрок в условиях неполной информации не знает самой позиции x , а знает лишь принадлежность некоторому множеству $X_i^j \subset X_i$ ($x \in X_i^j$). На информационные множества игрока условие 4 накладывает определенные ограничения. Требование $|F_x| = |F_y|$ для любых двух вершин одного информационного множества вводится для того, чтобы вершины $x, y \in X_i^j$ были неразличимы. Действительно, при $|F_x| \neq |F_y|$ игрок i мог бы различить между собой вершины $x, y \in X_i^j$ по числу выходящих из них дуг. Если бы в одном информационном множестве существовали две такие вершины x, y , что $y \in \hat{F}_x$, то это означало бы, что партия игры может пересекать дважды одно информационное множество, а это, в свою очередь, равносильно тому, что игрок j не помнит номера своего хода в данной партии, что трудно представимо в реальной игре.

§ 4.7. Стратегия поведения

Продолжим исследование многошаговой игры с неполной информацией и покажем, что в случае полной памяти у всех игроков она имеет ситуацию равновесия в стратегиях поведения.

4.7.1. Для дальнейшего исследования необходимо ввести ряд дополнительных понятий.

Определение. Альтернативами в вершине x , т. е. $\{(x, y) : y \in F_x\}$, называются дуги, инцидентные с вершиной $x \in X$.

Если $|F_x| = k$, то в вершине x имеется k альтернатив. Будем считать, что если в x имеется k альтернатив, то они нумеруются целыми числами $1, \dots, k$, причем вершина x обходится по часовой стрелке (рис. 4.9). В вершине x_0 первая альтернатива может быть указана произвольно. Если некоторая вершина $x \neq x_0$ обходится по часовой стрелке, то

первой альтернативой в x считается та, которая следует за единственной дугой (F_x^{-1}, x) , входящей в x .

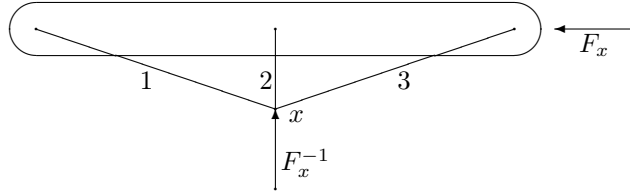


Рис. 4.9. Альтернативы в вершине x

Будем считать, что в игре Γ все альтернативы перенумерованы указанным способом. Пусть A_k — множество всех вершин $x \in X$, имеющих ровно k альтернатив, т. е. $A_k = \{x : |F_x| = k\}$. Пусть $I_i = \{X_i^j : X_i^j \subset X_i\}$ — множество всех информационных множеств игрока i . Под *чистой стратегией* игрока i будем понимать функцию u_i , отображающую I_i в множество положительных чисел так, что $u_i(X_i^j) \leq k$, если $X_i^j \subset A_k$. Будем говорить, что стратегия u_i выбирает альтернативу l в позиции $x \in X_i^j$, если $u_i(X_i^j) = l$, где l — номер альтернативы.

Так же как это было сделано в п. 4.1.4–4.1.5, можно показать, что каждой ситуации $u(\cdot) = (u_1(\cdot), \dots, u_n(\cdot))$ единственным образом соответствует партия ω , следовательно, и выигрыш в окончательной позиции этой партии.

Пусть $x \in X_{n+1}$ — некоторая окончательная позиция и ω — единственный путь $(F$ — дерево), ведущий из x_0 в x . Условие принадлежности позиции y пути ω будем записывать в виде $y \in \omega$ или $y < x$.

Определение. Позиция $x \in X$ называется *возможной для $u_i(\cdot)$* , если существует ситуация $u(\cdot)$, содержащая $u_i(\cdot)$, такая, что в ситуации x реализуется путь ω , который содержит позицию x , т. е. $x \in \omega$. Информационное множество X_i^j называется *существенным для $u_i(\cdot)$* , если некоторая позиция $x \in X_i^j$ возможна для $u_i(\cdot)$.

Множество позиций, возможных для $u_i(\cdot)$, обозначим через $Possu_i(\cdot)$, а семейство информационных множеств, существенных для $u_i(\cdot)$ — через $Relu_i(\cdot)$.

Лемма. Позиция $x \in X$ возможна для $u_i(\cdot)$ тогда и только тогда, когда $u_i(\cdot)$ выбирает альтернативы, лежащие на отрезке партии ω_x от x_0 до x во всех своих информационных множествах, пересекающих ω_x .

Доказательство. Пусть $x \in Possu_i(\cdot)$. Тогда существует ситуация $u(\cdot)$, содержащая $u_i(\cdot)$, такая, что партия ω , реализовавшаяся в этой ситуации, проходит через x , а это и означает, что на своих информационных множествах, пересекающих отрезок партии ω_x , стратегия $u_i(\cdot)$ выбирает альтернативы (дуги), принадлежащие ω_x .

Пусть теперь $u_i(\cdot)$ выбирает все альтернативы игрока i в ω_x . Для того чтобы доказать возможность x для $u_i(\cdot)$, необходимо построить ситуацию $u(\cdot)$, содержащую $u_i(\cdot)$, в которой партия проходила бы через x . Для игрока $k \neq i$ построим стратегию $u_k(\cdot)$, которая на информационных множествах X_k^j , пересекающих отрезок пути ω_x , выбирает альтернативы (дуги), лежащие на этом пути, а в остальном произвольна. Поскольку каждое информационное множество пересекает путь ω лишь однажды, это всегда можно сделать. В полученной ситуации $u(\cdot)$ партия ω обязательно пройдет через x . Следовательно, мы показали, что $x \in Possu_i(\cdot)$.

4.7.2. Смешанные стратегии в многошаговой игре с неполной информацией Γ определяются так же, как и в п. 1.4.2 для конечных игр.

Определение. Смешанной стратегией μ_i игрока i называется вероятностное распределение на множестве чистых стратегий игрока i , которое каждой его чистой стратегии $u_i(\cdot)$ ставит в соответствие вероятность $q_{u_i}(\cdot)$ (в дальнейшем для простоты будем писать просто q_{u_i}).

Ситуация $\mu = (\mu_1, \dots, \mu_n)$ в смешанных стратегиях определяет распределение вероятностей на всех партиях ω (следовательно, и на окончательных позициях X_{n+1}) по формуле

$$P_\mu(\omega) = \sum_u q_{u_1} \dots q_{u_n} P_u(\omega),$$

где $P_u(\omega) = 1$, если партия ω реализуется в ситуации $u(\cdot)$ и $P_u(\omega) = 0$ в противном случае.

Лемма. Обозначим через $P_\mu(x)$ вероятность реализации позиции x в ситуации μ . Тогда имеет место формула

$$P_\mu(x) = \sum_{\{u(\cdot): x \in \text{Poss}_{u_i}(\cdot), i=1, \dots, n\}} q_{u_1} \dots q_{u_n} = \prod_{i=1}^n \sum_{\{u_i: x \in \text{Poss}_{u_i}\}} q_{u_i}. \quad (4.7.1)$$

Доказательство этого утверждения непосредственно следует из леммы п. 4.7.1. Математическое ожидание выигрыша $E_i(\mu)$ игрока i в ситуации μ равно

$$E_i(\mu) = \sum_{x \in X_{n+1}} K_i(x) P_\mu(x), \quad (4.7.2)$$

где $P_\mu(x)$ вычисляется по формуле (4.7.1).

Определение. Позиция $x \in X$ называется возможной для μ_i , если существует ситуация μ в смешанных стратегиях, содержащая μ_i такая, что $P_\mu(x) > 0$. Информационное множество X_i^j игрока i называется существенным для μ_i если некоторое $x \in X_i^j$ является возможным для μ_i .

Множество возможных для μ_i позиций обозначим через $\text{Poss} \mu_i$, а множество существенных для μ_i информационных множеств — через $\text{Rel} \mu_i$.

4.7.3. Исследуя многошаговые игры с полной информацией, мы показали, что выбор стратегии может осуществляться на каждом шаге в соответствующей позиции игры, а при решении конкретных задач необязательно (да и практически невозможно) определять заранее стратегию, т. е. полный набор рекомендуемого поведения во всех позициях (информационных множествах), поскольку такое правило (см. пример п. 4.2.2) «страдает сильной избыточностью». Можно ли сделать аналогичное упрощение в играх с неполной информацией, т. е. строить стратегию не как заранее фиксированное правило выбора во всех информационных множествах, а формировать ее по мере попадания в соответствующее информационное множество? Оказывается, что в общем случае этого сделать нельзя. Однако существует класс игр с неполной информацией, где такое упрощение возможно. Для этого введем понятие стратегии поведения.

Определение. Под стратегией поведения β_i игрока i будем понимать правило, которое каждому информационному множеству $X_i^j \subset A_k$ игрока i ставит в соответствие систему из k чисел $b(X_i^j, \nu) \geq 0$, $\nu = 1, \dots, k$, таких что

$$\sum_{\nu=1}^k b(X_i^j, \nu) = 1, \quad \text{где } A_k = \{x : |F_x| = k\}.$$

Числа $b(X_i^j, \nu)$ могут интерпретироваться как вероятности выбора альтернативы ν в информационном множестве $X_i^j \subset A_k$, каждая позиция которого содержит ровно k альтернатив.

Любой набор $\beta = (\beta_1, \dots, \beta_n)$ стратегий поведения для n игроков определяет вероятностное распределение на партиях игры и окончательных позициях следующим образом:

$$P_\beta(\omega) = \prod_{X_i^j \cap \omega \neq \emptyset, \nu \in \omega} b(X_i^j, \nu). \quad (4.7.3)$$

Здесь произведение берется по всем X_i^j, ν таким, что $X_i^j \cup \omega \neq \emptyset$ и выбор в точке $X_i^j \cap \omega$ альтернативы с номером ν приводит в позицию, принадлежащую пути ω .

В дальнейшем под понятием «путь» удобно подразумевать не только набор составляющих его позиций, но и набор соответствующих альтернатив (дуг).

Ожидаемый выигрыш $E_i(\beta)$ в ситуации $\beta = (\beta_1, \dots, \beta_n)$ в стратегиях поведения определяется как математическое ожидание

$$E_i(\beta) = \sum_{x \in X_{n+1}} K_i(x) P_\beta(\omega_x), \quad i = 1, \dots, n,$$

где ω_x — партия, завершающаяся позицией $x \in X_{n+1}$.

4.7.4. Каждой смешанной стратегии μ_i можно сопоставить некоторую стратегию поведения β_i .

Определение. Стратегией поведения β_i , соответствующей смешанной стратегии $\mu_i = \{q_{u_i}\}$ игрока i , называется стратегия поведения, определенная следующим образом:

Если $X_i^j \in \text{Rel}\mu_i$, то

$$b(X_i^j, \nu) = \frac{\sum_{\{u_i: X_i^j \in \text{Rel}u_i, u_i(X_i^j) = \nu\}} q_{u_i}}{\sum_{\{u_i: X_i^j \in \text{Rel}u_i\}} q_{u_i}}. \quad (4.7.4)$$

Если $X_i^j \notin \text{Rel}\mu_i$, то на множестве X_i^j стратегию β_i можно определить произвольным, отличным от (4.7.4) образом. В случае $X_i^j \notin \text{Rel}\mu_i$ знаменатель в выражении (4.7.4) обращается в нуль. Для определенности в этом случае будем полагать

$$b(X_i^j, \nu) = \sum_{\{u_i: u_i(X_i^j) = \nu\}} q_{u_i}. \quad (4.7.5)$$

Приведем без доказательства следующий результат.

Лемма. Пусть β_i — стратегия поведения игрока i , а $\mu_i = \{q_{u_i}\}$ смешанная стратегия, определяемая формулой

$$q_{u_i} = \prod_{X_i^j} b(X_i^j, u_i(X_i^j)).$$

Тогда β_i — стратегия поведения, соответствующая μ_i .

4.7.5. Определение [Kuhn, 1953]. Игра Γ называется игрой с полной памятью для i -го игрока, если для любых $u_i(\cdot), X_i^j, x$ из условий $X_i^j \in \text{Rel}u_i$ и $x \in X_i^j$ следует, что $x \in \text{Poss}u_i$.

Из определения следует, что в игре с полной памятью для i -го игрока любая позиция из существенного для $u_i(\cdot)$ информационного множества является возможной для $u_i(\cdot)$. Термин «полная память» подчеркивает то обстоятельство, что очутившись в любом своем информационном множестве, i -й игрок может точно восстановить, какие альтернативы (т. е. номера) он выбирал во всех своих предыдущих ходах (в силу однозначного соответствия). Игра с полной памятью для всех игроков превращается в игру с полной информацией, если все ее информационные множества содержат по одной вершине.

4.7.6. Лемма. Пусть Γ — игра с полной памятью для всех игроков; ω — некоторая партия в Γ . Пусть $x \in X_i^j$ — последняя позиция в пути ω , в которой ходит игрок i , и пусть он выбирает в x дугу ν . Положим

$$T_i(\omega) = \{u_i : X_i^j \in Relu_i, u_i(X_i^j) = \nu\}.$$

Если в ω нет позиций из X_i , то через $T_i(\omega)$ обозначим множество всех чистых стратегий игрока i . Тогда партия ω реализуется в тех и только тех ситуациях $u(\cdot) = (u_1(\cdot), \dots, u_n(\cdot))$, для которых $u_i \in T_i(\omega)$.

Доказательство. Достаточность. Достаточно доказать, что если $u_i \in T_i(\omega)$, то стратегия u_i выбирает все дуги (альтернативы) игрока i , входящие в партию ω (если, конечно, игрок i вообще имеет ход в ω). Однако, если $u_i \in T_i(\omega)$, то $X_i^j \in Relu_i$, и поскольку игра Γ имеет полную память, то $x \in Possu_i$ ($x \in \omega$). Значит, согласно лемме п. 4.7.1, стратегия u_i выбирает все альтернативы игрока i , входящие в партию ω .

Необходимость. Предположим, что партия ω реализуется в ситуации $u(\cdot)$, у которой $u_i \notin T_i(\omega)$ для некоторого i . Поскольку $X_i^j \in Relu_i$, то это означает, что $u_i(X_i^j) \neq \nu$. Но тогда путь ω не реализуется. Полученное противоречие завершает доказательство леммы.

4.7.7. Лемма. Пусть Γ — игра с полной памятью для всех игроков. Пусть ν — альтернатива (дуга) в партии ω , инцидентная $x \in X_i^j$, где $x \in \omega$, и следующая позиция игрока i (если она существует) в пути ω есть $y \in X_i^k$. Рассмотрим множества S и T , где

$$S = \{u_i : X_i^j \in Relu_i, u_i(X_i^j) = \nu\},$$

$$T = \{u_i : X_i^k \in Relu_i\}.$$

Тогда $S = T$.

Доказательство. Пусть $u_i \in S$. Тогда $X_i^j \in Relu_i$, и поскольку Γ имеет полную память, то $x \in Possu_i$. Следовательно, по лемме п. 4.7.1 стратегия u_i выбирает все дуги, инцидентные к позициям игрока i на пути от x_0 до x , и $u_i(X_i^j) = \nu$. Таким образом u_i выбирает все дуги, инцидентные к позициям игрока i на пути от x_0 до y , т. е. $y \in Possu_i$, $X_i^k \in Relu_i$ и $u_i \in T$.

Пусть $u_i \in T$. Тогда $X_i^k \in Relu_i$, и поскольку Γ имеет полную память, то $y \in Possu_i$. Однако это означает, что $x \in Possu_i$ и $u_i(X_i^j) = \nu$, т. е. $u_i \in S$. Лемма доказана.

4.7.8. Теорема. Пусть β — ситуация в стратегиях поведения, соответствующая ситуации в смешанных стратегиях μ в игре Γ (в которой все позиции имеют по крайней мере две альтернативы). Тогда для того чтобы

$$E_i(\beta) = E_i(\mu), \quad i = 1, \dots, n,$$

необходимо и достаточно, чтобы Γ была игрой с полной памятью для всех игроков.

Доказательство. Достаточность. Пусть Γ — игра с полной памятью для всех игроков. Фиксируем произвольное μ . Достаточно показать, что $P_\beta(\omega) = P_\mu(\omega)$ для всех партий ω . Если в ω существует позиция игрока i , принадлежащая несущественному для μ_i информационному множеству, то найдется $X_i^j \in Rel\mu_i$, $X_i^j \cap \omega \neq \emptyset$ такое, что для стратегии поведения β_i , соответствующей μ_i выполняется равенство $b(X_i^j, \nu) = 0$, где $\nu \in \omega$. Отсюда имеем $P_\beta(\omega) = 0$. Справедливость соотношения $P_\mu(\omega) = 0$ в этом случае очевидна.

Будем теперь считать, что все информационные множества i -го игрока, через которые проходит партия ω , существенны для μ_i , $i = 1, 2, \dots, n$. Пусть игрок i в партии ω ходит по порядку в позициях, принадлежащих множествам X_i^1, \dots, X_i^s и выбирает в множестве X_i^j альтернативу ν_j , $j = 1, \dots, s$. Тогда согласно формуле (4.7.3) и лемме п. 4.7.7 имеем

$$\prod_{j=1}^s b(X_i^j, \nu_j) = \sum_{u_i \in T_i(\omega)} q_{u_i}.$$

Действительно, поскольку в партии ω игрок i свой 1-й ход делает из множества X_i^1 , оно является существенным для всех $u_i(\cdot)$, поэтому знаменатель в формуле (4.7.4) для $b(X_i^1, \nu_1)$ равен единице. Далее в силу леммы п. 4.7.7 в формулах (4.7.4) числитель $b(X_i^j, \nu_j)$ равен знаменателю $b(X_i^{j+1}, \nu_{j+1})$, $i = 1, \dots, s$. Согласно формуле (4.7.3) окончательно получим

$$P_\beta(\omega) = \prod_{i=1}^n \sum_{u_i \in T_i(\omega)} q_{u_i},$$

где $T_i(\omega)$ определено в лемме п. 4.7.6.

В то же время на основании леммы п. 4.7.6

$$P_\mu(\omega) = \sum_{u(\cdot)} q_{u_1} \dots q_{u_n} P_u(\omega) = \sum_{u: u_i \in T_i(\omega), i=1, \dots, n} q_{u_1} \dots q_{u_n},$$

т. е. $P_\mu(\omega) = P_\beta(\omega)$ и достаточность доказана.

Необходимость. Пусть Γ не является игрой с полной памятью для всех игроков. Тогда существуют игрок i , стратегия u_i , информационное множество $X_i^j \in Relu_i$ и две позиции $x, y \in X_i^j$ такие, что $x \in Possu_i$, $y \notin Possu_i$. Пусть u'_i — стратегия игрока i , для которой $y \in Possu'_i$ и ω — соответствующая партия, проходящая через y в ситуации u' . Обозначим через μ_i смешанную стратегию игрока i , которая предписывает с вероятностью $1/2$ выбирать стратегию u_i или u'_i . Тогда $P_{u' \parallel \mu_i}(y) = P_{u' \parallel \mu_i}(\omega) = 1/2$ (здесь $u' \parallel \mu_i$ — ситуация, в которой чистая стратегия u'_i заменена на смешанную μ_i). Из условия $y \notin Possu_i$ следует, что путь $\bar{\omega}$, реализующийся в ситуации $u' \parallel \mu_i$ не проходит через y . Это означает, что существует X_i^k такое, что $X_i^k \cap \omega = X_i^k \cap \bar{\omega} \neq \emptyset$ и $u_i(X_i^k) \neq u'_i(X_i^k)$. Отсюда, в частности, следует $X_i^k \in Relu_i$, $X_i^k \in Relu'_i$. Пусть β_i — стратегия поведения, соответствующая μ_i . Тогда $b(X_i^k, u'_i(X_i^k)) = 1/2$. Не ограничивая общности, можно считать, что $u_i(X_i^j) \neq u'_i(X_i^j)$. Тогда $b(X_i^j, u'_i(X_i^j)) = 1/2$. Обозначим через β ситуацию в стратегиях поведения, соответствующую ситуации в смешанных стратегиях $u' \parallel \mu_i$. Тогда $P_\beta(\omega) \leq 1/4$, в то время как $P_{u' \parallel \mu_i}(\omega) = 1/2$. Теорема доказана.

Из теоремы п. 4.7.8 следует, что для нахождения ситуации равновесия в играх с полной памятью достаточно ограничиться классом стратегий поведения.

§ 4.8. Функциональные уравнения для одновременных многошаговых игр

Теорема о стратегиях поведения, доказанная в предыдущем параграфе, в общем случае не дает возможности непосредственно решать многошаговые игры с полной памятью. Однако при простой структуре информационных множеств она обосновывает вывод функциональных уравнений для значения игры и основанные на этих уравнениях методы нахождения оптимальных стратегий. Наиболее простыми играми с полной памятью, не считая игр с полной информацией, являются так называемые *одновременные многошаговые игры*. Выведем функциональное уравнение для значения таких игр и рассмотрим несколько широко известных [Дюбин, Суздаль, 1981; Оуэн, 1971] примеров, где эти уравнения поддаются решению.

4.8.1. Содержательно одновременная многошаговая игра представляет собой антагонистическую многошаговую игру, в которой на каждом шаге игры игроки 1 и 2 выбирают свои действия одновременно, т. е. не имея информации о выборе противником позиции в этот момент. После того как выборы сделаны, они становятся известными обоим игрокам, и игроки вновь совершают одновременный выбор и т. д.

Условно такую игру будем изображать с помощью графа, имеющего одно из двух представлений а) или б) рис. 4.10.

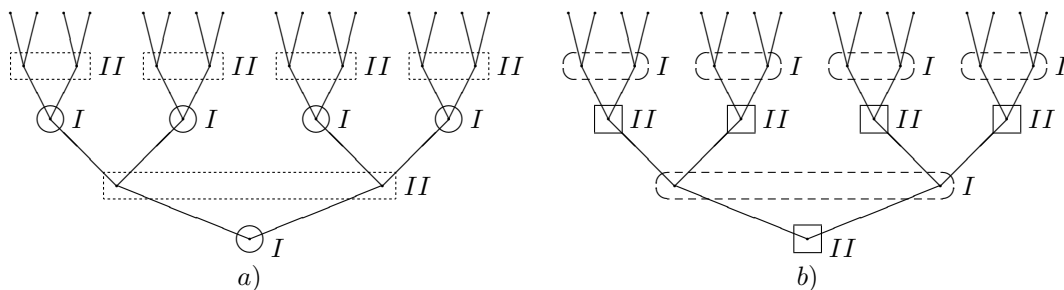


Рис. 4.10. Дерево одновременной многошаговой игры

Граф изображает поочередную игру с четным числом ходов, в которой информационные множества игрока, совершающего первый ход, являются одноэлементными, а информационные множества другого игрока двухэлементными. В такой игре Γ оба игрока обладают полной памятью, поэтому в ней согласно теореме п. 4.7.8 при отыскании ситуации равновесия можно ограничиться классом стратегий поведения.

Пусть, для определенности, в игре Γ первым ходит игрок 1. С каждой вершиной $x \in X_1$ связывается подыгра Γ_x с той же информационной структурой, что и игра Γ . Нормальная форма любой антагонистической конечно-шаговой игры с неполной информацией представляет собой матричную игру, т. е. антагонистическую игру с конечным числом стратегий. Поэтому во всех подыграх Γ_x , $x \in X_1$ (включая игру $\Gamma = \Gamma_{x_0}$) существует ситуация равновесия в классе смешанных стратегий. Согласно теореме п. 4.7.8 такая ситуация равновесия существует и в классе стратегий поведения и значения игр (т. е. значения функции выигрыша в ситуации равновесия в классе смешанных стратегий и в классе стратегий поведения) равны между собой.

Обозначим значение игры Γ_x через $v(x)$, $x \in X_1$ и составим функциональные уравнения для $v(x)$.

Для каждого $x \in X_1$ следующая позиция x' , в которой ходит игрок 1 (если таковая вообще существует), принадлежит множеству F_x^2 . Позиция x' реализуется в результате двух последовательных выборов: игроком 1 — дуги, инцидентной к вершине x , и игроком 2 — дуги в позициях $y \in F_x$, образующих информационные множества игрока 2. Поэтому можно считать, что позиция x' получается в результате отображения T_x , зависящего от выборов α, β игроков 1 и 2, т. е.

$$x' = T_x(\alpha, \beta).$$

Поскольку число различных альтернатив α и β конечно, то можно рассмотреть для каждого $x \in X_1$ матричную игру с матрицей выигрышей $A_x = \{v[T_x(\alpha, \beta)]\}$. Пусть $\beta_I^*(x) = \{b_I^*(x, \alpha)\}$, $\beta_{II}^*(x) = \{b_{II}^*(x, \beta)\}$ — оптимальные смешанные стратегии в игре с матрицей A_x . Тогда имеет место следующая теорема о структуре оптимальных стратегий в игре Γ_x .

Теорема. В игре Γ оптимальная стратегия поведения игрока 1 в точке x (каждое информационное множество игрока 1 в игре Γ состоит из одной позиции $x \in X_1$) предписывает каждой альтернативе α вероятность в соответствии со смешанной оптимальной стратегией игрока 1 в матричной игре $A_x = \{v(T_x(\alpha, \beta))\}$, т. е.

$$b_1(x, \alpha) = b_I^*(x, \alpha).$$

Оптимальная стратегия поведения $\{b_2(X_2^j, \beta)\}$ игрока 2 в игре Γ предписывает каждой альтернативе β вероятность в соответствии с оптимальной смешанной стратегией игрока 2 в игре с матрицей A_x , т. е.

$$b_2(X_2^j, \beta) = b_{II}^*(x, \beta),$$

где $x = F_y^{-1}$, если $y \in X_2^j$.

Значение игры удовлетворяет следующему функциональному уравнению:

$$v(x) = \text{Val}\{v[T_x(\alpha, \beta)]\}, \quad x \in X_1, \quad (4.8.1)$$

с граничным условием

$$v(x)|_{x \in X_3} = H(x). \quad (4.8.2)$$

Здесь запись $\text{Val}A$ обозначает значение игры с матрицей A .

Доказательство проводится по индукции и вполне аналогично доказательству теоремы п. 4. 2.1.

4.8.2. Пример 7 (Игра инспектирования) [Дюбин, Суздаль, 1981]. Игрок E (Нарушитель) хочет совершить некоторое запрещенное действие. Имеется N периодов времени, в которые это действие может быть осуществлено. Игрок P (инспектор), желающий предотвратить это действие, может провести только одну инспекцию в любой из этих периодов времени. Выигрыш игрока E равен 1, если запрещенное действие произошло и осталось необнаруженным, и равен (-1) , если нарушитель пойман (это будет в том случае, когда для совершения действия он выбирает тот же самый период времени, что и инспектор для проверки); выигрыш равен нулю, если нарушитель не действует вовсе. Обозначим такую N -шаговую игру через Γ_N .

В первом периоде (на 1-м шаге) каждый игрок имеет две альтернативы. Игрок E может предпринимать действие или не предпринимать его; игрок P может инспектировать или не инспектировать. Если игрок E действует и игрок P инспектирует, то игра заканчивается и выигрыш равен -1 . Если игрок E действует, а игрок P не инспектирует, то игра заканчивается и выигрыш равен 1 . Если игрок E не действует, а игрок P инспектирует, то игрок E может предпринять действие в следующий период времени (в предположении, что $N > 1$) и выигрыш также равен 1 . Если игрок E не действует и игрок P не инспектирует, то переходят к следующему шагу игры, который отличается от предыдущего только тем, что до конца игры остается меньшее число периодов времени, т. е. попадают в подыгру Γ_{N-1} . Следовательно, матрица для 1-го шага игры выглядит следующим образом:

$$\begin{bmatrix} -1 & 1 \\ 1 & v_{N-1} \end{bmatrix}. \quad (4.8.3)$$

Уравнение (4.8.1) в этом случае принимает вид

$$v_N = Val \begin{bmatrix} -1 & 1 \\ 1 & v_{N-1} \end{bmatrix}. \quad (4.8.4)$$

Здесь $v(x)$ одинаково для всех позиций игры одного уровня и поэтому зависит только от числа периодов до конца игры. Поэтому вместо $v(x)$ записано v_N . Далее будет показано, что $v_{N-1} < 1$, следовательно, матрица в (4.8.4) не имеет седловой точки, т. е. игра с матрицей (4.8.4) является вполне смешанной. Отсюда получаем (см. п. 1.9.1) рекуррентное уравнение

$$v_N = \frac{v_{N-1} + 1}{-v_{N-1} + 3}, \quad (4.8.5)$$

которое вместе с начальным условием

$$v_1 = Val \begin{pmatrix} -1 & 1 \\ 0 & 0 \end{pmatrix} = 0 \quad (4.8.6)$$

определяет v_N . Преобразуем уравнение (4.8.5) с помощью подстановки $t_N = \frac{1}{v_{N-1}}$. Получим новое рекуррентное уравнение $t_N = t_{N-1} - 1/2$, $t_1 = -1$. Это уравнение имеет очевидное решение $t_N = -(N+1)/2$, откуда имеем

$$v_N = \frac{N-1}{N+1}. \quad (4.8.7)$$

Теперь можно вычислить оптимальные стратегии поведения на каждом шаге игры. Действительно, матрица игры (4.8.4) принимает вид

$$\begin{bmatrix} -1 & 1 \\ 1 & [N-2]/N \end{bmatrix}$$

и оптимальные стратегии поведения таковы:

$$b_1^N = \left(\frac{1}{N+1}, \frac{N}{N+1} \right), \quad b_0^N = \left(\frac{1}{N+1}, \frac{N}{N+1} \right).$$

Пример 8 (Теоретико-игровые особенности оптимального расхода ресурса). Пусть первоначально игроки 1 и 2 имеют соответственно r и $R-r$ единиц некоторого ресурса,

а также по две чистые стратегии. Допустим, что если игроки выберут одинаковые по номеру чистые стратегии, то ресурс игрока 2 уменьшится на единицу. Если же игроки выберут разные по номеру чистые стратегии, то на единицу уменьшится ресурс игрока 1. Игра заканчивается после того, как ресурс одного из игроков станет равным нулю. При этом игрок 1 получает выигрыш, равный 1, если ресурс игрока 2 станет равным нулю, и выигрыш -1 , если станет равным нулю его собственные ресурсы.

Обозначим через $\Gamma_{k,l}$ многошаговую игру, в которой игрок 1 имеет k ($k = 1, 2, \dots, r$) единиц, а игрок 2 — l ($l = 1, \dots, R - r$) единиц ресурса. Тогда

$$Val\Gamma_{k,l} = Val \begin{bmatrix} Val\Gamma_{k,l-1} & Val\Gamma_{k-1,l} \\ Val\Gamma_{k-1,l} & Val\Gamma_{k,l-1} \end{bmatrix},$$

где $Val\Gamma_{k,0} = 1$, $Val\Gamma_{0,l} = -1$.

Рассмотрим 1-й от конца шаг, т. е. когда у обоих игроков осталось по одной единице ресурсов. Очевидно, что на этом шаге разыгрывается следующая матричная игра:

$$\Gamma_{1,1} = \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}.$$

Игра $\Gamma_{1,1}$ и является симметричной, ее значение, которое мы обозначим через $v_{1,1}$ равно нулю, а оптимальные стратегии игроков совпадают и равны $(1/2, 1/2)$.

На 2-м от конца шаге, т. е. когда у игроков осталось три единицы ресурсов, разыгрывается одна из двух матричных игр: $\Gamma_{1,2}$ или $\Gamma_{2,1}$. При этом

$$v_{1,2} = Val\Gamma_{1,2} = Val \begin{bmatrix} v_{1,1} & -1 \\ -1 & v_{1,1} \end{bmatrix} = \frac{v_{1,1} - 1}{2} = -\frac{1}{2},$$

$$v_{2,1} = Val\Gamma_{2,1} = Val \begin{bmatrix} 1 & v_{1,1} \\ v_{1,1} & 1 \end{bmatrix} = \frac{v_{1,1} + 1}{2} = \frac{1}{2}.$$

На 3-м от конца шаге (т. е. когда у игроков имеется в общей сложности четыре единицы ресурса) разыгрывается одна из следующих трех игр: $\Gamma_{1,3}$, $\Gamma_{2,2}$, $\Gamma_{3,1}$. При этом

$$v_{1,3} = Val\Gamma_{1,3} = Val \begin{bmatrix} v_{1,2} & -1 \\ -1 & v_{1,2} \end{bmatrix} = \frac{v_{1,2} - 1}{2} = -\frac{3}{4},$$

$$v_{2,2} = Val\Gamma_{2,2} = Val \begin{bmatrix} v_{2,1} & v_{1,2} \\ v_{1,2} & v_{2,1} \end{bmatrix} = \frac{v_{2,1} + v_{1,2}}{2} = 0,$$

$$v_{3,1} = Val\Gamma_{3,1} = Val \begin{bmatrix} 1 & v_{2,1} \\ v_{2,1} & 1 \end{bmatrix} = \frac{v_{2,1} + 1}{2} = \frac{3}{4}.$$

Продолжая аналогичные вычисления далее до N -го шага от конца, получим следующее выражение для значения исходной игры:

$$v_{r,R-r} = Val\Gamma_{r,R-r} = Val \begin{bmatrix} v_{r,R-r-1} & v_{r-1,R-r} \\ v_{r-1,R-r} & v_{r,R-r-1} \end{bmatrix}.$$

В силу симметричности матрицы выигрышей игры $\Gamma_{r,R-r}$ имеем

$$v_{r,R-r} = \frac{1}{2}(v_{r,R-r-1} + v_{r-1,R-r}),$$

оптимальные стратегии поведения игроков на каждом шаге совпадают и равны $(1/2, 1/2)$.

Пример 9. В шуточной игре играют две команды: игрок 1 (m_1 женщин и m_2 кошек) и игрок 2 (n_1 мышей и n_2 мужчин). На каждом шаге каждый из игроков выбирает своего представителя. Один из двух выбранных представителей «устраняется» согласно следующим правилам: женщина «устраняет» мужчину; мужчина «устраняет» кошку; кошка «устраняет» мышью; мышья «устраняет» женщину. Игра продолжается до тех пор, пока в одной из групп не останутся игроки только одного типа. Когда группа не имеет больше выбора, другая группа, очевидно, выигрывает.

Обозначим значение исходной игры $v(m_1, m_2, n_1, n_2)$. Будем полагать

$$v(m_1, m_2, n_1, 0) = v(m_1, m_2, 0, n_2) = 1, \text{ при } m_1, m_2 > 0. \quad (4.8.8)$$

$$v(m_1, 0, n_1, n_2) = v(0, m_2, n_1, n_2) = -1, \text{ при } n_1, n_2 > 0. \quad (4.8.9)$$

Введем следующие обозначения: $v(m_1 - 1) = v(m_1 - 1, m_2, n_1, n_2)$, $v(m_2 - 1) = v(m_1, m_2 - 1, n_1, n_2)$, $v(n_1 - 1) = v(m_1, m_2, n_1 - 1, n_2)$, $v(n_2 - 1) = v(m_1, m_2, n_1, n_2 - 1)$. Согласно теореме п. 1.9.1 справедливо соотношение

$$v(m_1, m_2, n_1, n_2) = Val \begin{bmatrix} v(m_1 - 1) & v(n_2 - 1) \\ v(n_1 - 1) & v(m_2 - 1) \end{bmatrix}.$$

Можно показать, что рассматриваемая игра является вполне смешанной. Согласно теореме п. 1.9.1 имеем

$$v(m_1, m_2, n_1, n_2) = \frac{v(m_1 - 1)v(m_2 - 1) - v(n_1 - 1)v(n_2 - 1)}{v(m_1 - 1) + v(m_2 - 1) - v(n_1 - 1) - v(n_2 - 1)}.$$

Учитывая граничные условия (4.8.8), отсюда получаем

$$v(m_1, 1, 1, 1) = \frac{v(m_1 - 1) + 1}{v(m_1 - 1) + 3}$$

и $v(1, 1, 1, 1) = 0$. Но эти уравнения совпадают с уравнениями (4.8.5), (4.8.6), следовательно, $v(m, 1, 1, 1) = (m - 1)/(m + 1)$ и оптимальные стратегии в этом случае также совпадают с приведенными в примере 8.

4.8.3. Повторяющиеся эволюционные игры. Для определения повторяющихся эволюционных игр необходимо определить понятие симметрии для позиционных игр. Мы ограничимся изучением игр двух лиц с полной памятью. Следуя [Selten (1983)], симметрия игры Γ определяется как отображение $(\cdot)^T$ из множества альтернатив на множество альтернатив со следующими свойствами. Пусть M_i обозначает множество альтернатив (выборов) игрока i в игре Γ .

Если $m \in M_i$, то $m^T \in M_j$ ($i \neq j \in \{1, 2\}$),

$$(m^T)^T = m \text{ для всех } m.$$

Для каждого информационного множества u существует такое информационное множество u^T , что любая альтернатива из u отображается на выбор в u^T , а для каждой конечной точки $x \in X_{n+1}$ существует такая конечная точка $x^T \in X_{n+1}$, что если x достигается последовательностью m_1, m_2, \dots, m_n , то x^T достигается (перестановкой) $m_1^T, m_2^T, \dots, m_n^T$, и выигрыши равны $H_1(x) = H_2(x^T)$ для каждой конечной точки $x \in X_{n+1}$, $x^T \in X_{n+1}$.

Симметричная позиционная игра задается парой (Γ, T) где Γ — позиционная игра, а T — симметрия в игре Γ . Если b — стратегия поведения игрока 1 в (Γ, T) , то симметричный образ b является стратегией поведения b^T игрока 2, которая определяется как

$$b_u^T(m) = b_{u_T}(m^T) \quad (u \in U_1, m \in M_u).$$

Если b_1, b_2 — стратегии поведения игрока 1, то вероятность того, что конечная точка x достигается при использовании стратегий (b_1, b_2^T) равна вероятности того, что x^T достигается при использовании (b_2, b_1^T) . Следовательно, ожидаемый выигрыш игрока 1 при использовании стратегий (b_1, b_2^T) равен ожидаемому выигрышу игрока 2 при использовании стратегий (b_2, b_1^T) :

$$E_1(b_1, b_2^T) = E_2(b_2, b_1^T).$$

Это уравнение определяет симметричную игру в нормальной форме при ограничении, что в игре (Γ, T) используются только чистые стратегии. Следуя [van Damme, 1991], определим эволюционно устойчивую стратегию (ESS) в игре (Γ, T) как стратегию поведения \bar{b} игрока 1, которая удовлетворяет условию:

$$E_1(\bar{b}, \bar{b}^T) = \max_{b \in B_1} E_1(b, \bar{b}^T),$$

и если $b \in B_2, b \neq \bar{b}$ и $E_1(b, \bar{b}^T) = E_1(\bar{b}, \bar{b}^T)$, то

$$E_1(b, b^T) < E_1(\bar{b}, b^T).$$

Здесь B_i — множество стратегий поведения игрока i). В работе [van Damme, 1991] отмечалось, что во многих играх интуитивно приемлемые решения не будут удовлетворять указанному выше условию ESS.

Далее мы приведем некоторые модификации этого определения, которые не будут исключать интуитивно приемлемые решения. Сначала рассмотрим следующий пример.

4.8.4. Пример 10 (Повторяющаяся игра Ястреб–Голубь).

Данная биматричная игра задается следующими матрицами:

$$A = \begin{matrix} & \begin{matrix} H & D \end{matrix} \\ \begin{matrix} H \\ D \end{matrix} & \begin{bmatrix} 1/2(V-C) & V \\ 0 & 1/2V \end{bmatrix} \end{matrix}, \quad A^T = \begin{matrix} & \begin{matrix} H & D \end{matrix} \\ \begin{matrix} H \\ D \end{matrix} & \begin{bmatrix} 1/2(V-C) & 0 \\ V & 1/2V \end{bmatrix} \end{matrix}.$$

Если $v > c$, то (H, H) является эволюционно устойчивой стратегией ESS в заданной биматричной игре Γ . Дерево игры представлено на рис. 4.11.

Для игры с двумя шагами стратегией игрока 1(2) является правило, которое определяет выбор H или D в каждом из информационных множеств игрока. Игрок 1(2) имеет пять информационных множеств, таким образом он имеет 32 стратегии, которые представляют собой последовательности вида

$$(H, H, D, H, D).$$

Обозначим стратегию игрока через $u(\cdot)$.

Рассмотрим стратегию $u(\cdot) = (H, H, H, H, H)$, которая сформирована из ESS стратегий ($v > c$) в каждой подыгре (одношаговой игре). Было бы хорошо, если бы эта

§ 4.9. Построение единственного равновесия по Нэшу

Как и прежде, пусть N — множество игроков, X_1, \dots, X_n — множества очередности и X_{n+1} — множество окончательных позиций. Для каждого $i \in N$ определим вектор предпочтений i -го игрока $F_i = \{f_i(j)\}$, $j \in N \setminus \{i\}$ таким образом, что n -мерный вектор $\{f_i(j)\}$, $j \in N$ представляет собой перестановку чисел $1, \dots, n$.

Можно представить себе следующую интерпретацию вектора предпочтений $f_i(j)$, $j \in N \setminus \{i\}$, $i \in N$: если $f_i(j) = k > 0$, тогда игрок j является «другом» игрока i уровня k , если $f_i(j) = k < 0$, тогда игрок j является «врагом» игрока i уровня k и $f_i(i) = 1$.

Каждый вектор F_i определяет отношение игрока i к другим игрокам или «тип» игрока (здесь мы используем терминалогию, впервые введенную в [Dresher, 1961], из-за некоторой отдаленной аналогии).

Предположим, что абсолютное равновесие в игре Γ найдено методом математической индукции. Сначала рассмотрим подыгры наименьшей длины Γ (одношаговые подыгры). Далее разбиваем Γ на подыгры меньшей длины, предполагая, что в подыграх используются равновесные по Нэшу стратегии. При построении равновесия по Нэшу может случиться, что в некоторой подыгре Γ_x , $x \in X_i$ игры Γ игрок i , совершающий выбор в x , обнаруживает, что его выигрыш (при условии продолжения игры в соответствии с данным равновесием по Нэшу в подыграх) не зависит от того, какую альтернативу в x он выберет. В этом случае, используя вектор предпочтений, он обратит внимание на игрока j , для которого $|f_i(j)| = 2$. Если $f_i(j) > 0$, он выберет альтернативу в x так, чтобы максимизировать выигрыш игрока j , а если $f_i(j) < 0$, — так, чтобы его минимизировать.

Если выигрыш игрока j во всех подыграх, непосредственно следующих из Γ_x , один и тот же, то игрок i смотрит на игрока j , для которого $|f_i(j)| = 3$, и ведет себя аналогичным образом, и т. д. Стратегию, включающую такой способ поведения, будем называть *тип-стратегией*. Абсолютное равновесие, получаемое методом динамического программирования с использованием *тип-стратегий*, назовем *равновесием по Нэшу в тип-стратегиях*. Обозначим его через $T = (T_1, \dots, T_n)$ и через $K_i(T)$, $i \in N$, соответствующие выигрыши.

Дадим определение равновесия по Нэшу в тип-стратегиях индукцией по длине l дерева игры. Обозначим через $x(k; y)$ позицию, непосредственно следующую за y , если в y выбрана альтернатива (дуга) k (рис. 4.12).

Если $l = 1$, то в игре Γ имеется только один ход в x_0 , который заканчивается окончательной позицией. Пусть $x_0 \in X_{i_1}$. Выигрыши игроков $h_i\{x(k; x_0)\}$, $i = 1, \dots, n$, определены в окончательных позициях $x(k; x_0) \in X_{n+1}$.

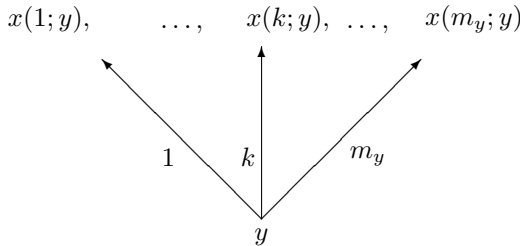


Рис. 4.12. Альтернативы в позиции y

Пусть $F(i_1)$ — вектор предпочтений игрока i_1 . Введем последовательность i_1, i_2, \dots, i_n таким образом, что $|f_{i_1}(i_k)| = 1 + |f_{i_1}(i_{k-1})|$, $k = 2, \dots, n$ ($f_{i_1}(i_1) = 1$).

Обозначим через $\varphi(j) = \text{sign} f_{i_1}(i_j)$, $j = 1, \dots, n$. Тогда $f_{i_1}(i_k) = k[\text{sign} f_{i_1}(i_k)]$.

Определим семейство множеств $A_{i_1}[\varphi(j)]$, $j = 2, \dots, n$ по правилу:

$$A_{i_1}[\varphi(1)] = \arg \left[\max_k h_{i_1} \{x(k; x_0)\} \right],$$

$$A_{i_1}[\varphi(j)] = \begin{cases} \arg \left[\max_{k \in A_{i_1}[\varphi(j-1)]} h_{i_j} \{x(k; x_0)\} \right], & \text{если } \varphi(j) > 0, \\ \arg \left[\min_{k \in A_{i_1}[\varphi(j-1)]} h_{i_j} \{x(k; x_0)\} \right], & \text{если } \varphi(j) < 0. \end{cases}$$

Отсюда $A_{i_1}[\varphi(j)] \subset A_{i_1}[\varphi(j-1)]$, $i_1 \in N$.

Обозначим $A_{i_1}^* = \bigcap_{j=1}^n A_{i_1}[\varphi(j)]$. В одношаговой игре Γ в $x_0 \in X_{i_1}$ тип-стратегия T_{i_1} выбирает любую альтернативу $k \in A_{i_1}^*$, т. е.

$$T_{i_1}(x_0) = k, \quad k \in A_{i_1}^* \quad (4.9.1)$$

Лемма. Для всех $k_1, k_2 \in A^*$ имеем

$$h_j \{x(k_1; x_0)\} = h_j \{x(k_2; x_0)\}, \quad j = 1, \dots, n. \quad (4.9.2)$$

Доказательство. Предположим, что (4.9.2) не выполняется для некоторого $j \in N$. Тогда существует \bar{j} , такое, что

$$h_{i_j} \{x(k_1; x_0)\} = h_{i_j} \{x(k_2; x_0)\}, \quad i_j < i_{\bar{j}}, \quad (4.9.3)$$

$$h_{i_{\bar{j}}} \{x(k_1; x_0)\} \neq h_{i_{\bar{j}}} \{x(k_2; x_0)\}. \quad (4.9.4)$$

Поскольку $k_1, k_2 \in A^*$,

$$k_1, k_2 \in A[\varphi(\bar{j})] = \begin{cases} \arg \left[\max_{k \in A_{i_1}[\varphi(\bar{j}-1)]} h_{i_{\bar{j}}} \{x(k; x_0)\} \right], & \text{если } \varphi(\bar{j}) > 0, \\ \arg \left[\min_{k \in A_{i_1}[\varphi(\bar{j}-1)]} h_{i_{\bar{j}}} \{x(k; x_0)\} \right], & \text{если } \varphi(\bar{j}) < 0. \end{cases} \quad (4.9.5)$$

В (4.9.5) \max (\min) функции $h_{i_{\bar{j}}}$ определен на одном множестве $A_{i_1}[\varphi(\bar{j}-1)]$. Пусть $\varphi(\bar{j}) > 0$, тогда по определению

$$\max_{k \in A_{i_1}[\varphi(\bar{j}-1)]} h_{i_{\bar{j}}} \{x(k; x_0)\} = h_{i_{\bar{j}}} \{x(k_1; x_0)\} = h_{i_{\bar{j}}} \{x(k_2; x_0)\},$$

что противоречит (4.9.3)—(4.9.4). Так же можно рассмотреть и случай $\varphi(\bar{j}) < 0$. Лемма доказана.

Следствие. Из леммы § 4.9. следует, что в случае $l = 1$ в любых двух ситуациях в тип-стратегиях T'_{i_1} , T''_{i_2} соответствующие вектор-выигрыши совпадают.

Пусть в Γ $x_0 \in X_{i_1}$. Все подыгры Γ_x , начинающиеся из вершин (позиций), непосредственно следующих за x_0 , имеют длину $< l$. Пусть тип-стратегии уже определены по индукции в подыграх. Для каждого x зафиксируем ситуацию $T^x = (T_1^x, \dots, T_n^x)$ в тип-стратегиях в Γ_x . Пусть $v(x) = (v_1(x), \dots, v_n(x))$ — соответствующий вектор выигрышей в Γ_x в ситуации T^x . Если вершина x соответствует альтернативе k в x_0 , то будем писать

$$v(x) = v\{x(k; x_0)\}, \quad v_i(x) = v_i\{x(k; x_0)\}, \quad i = 1, \dots, n.$$

Определим семейство множеств $A_{i_1}[\varphi(j)]$, $j = 2, \dots, n$,

$$A_{i_1}[\varphi(1)] = \arg \left[\max_k v_{i_1}\{x(k; x_0)\} \right],$$

$$A_{i_1}[\varphi(j)] = \begin{cases} \arg \left[\max_{k \in A_{i_1}[\varphi(j-1)]} v_{i_j}\{x(k; x_0)\} \right], & \text{если } \varphi(j) > 0, \\ \arg \left[\min_{k \in A_{i_1}[\varphi(j-1)]} v_{i_j}\{x(k; x_0)\} \right], & \text{если } \varphi(j) < 0. \end{cases}$$

Обозначим $A_{i_1}^* = \bigcap_{j=1}^n A_{i_1}[\varphi(j)]$. Определим тип-стратегию в игре Γ :

$$T_{i_1}(x_0) = k, \quad k \in A_{i_1}^*, \quad T_{i_1}(y) = T_{i_1}^x(y) \quad \text{при } y \in X_{i_1}^x,$$

где $X^x = \{X_1^x, \dots, X_{i_1}^x, \dots, X_n^x\}$ — множество очередности в подыгре Γ_x , начинающейся с позиции x , непосредственно следующей за x_0 :

$$T_i(y) = T_i^x(y) \quad \text{для } y \in X_i^x, \quad i \neq i_1.$$

Доказательство следующей леммы полностью повторяет предыдущее, если мы заменим $h_i\{x(k; x_0)\}$ на $v_i\{x(k; x_0)\}$.

Лемма. Для всех $k_1, k_2 \in A^*$ имеет место:

$$v_j\{x(k_1; x_0)\} = v_j\{x(k_2; x_0)\}, \quad j = 1, \dots, n. \quad (4.9.6)$$

Теорема. Если игрокам известны свой тип и типы остальных игроков, и все об этом знают, то в каждой игре Γ существует абсолютное равновесие по Нэшу в тип-стратегиях, и в любых различных ситуациях $T = (T_1, \dots, T_n)$, $T' = (T'_1, \dots, T'_n)$ в тип-стратегиях выигрыши игроков совпадают, т. е.

$$K_i(T_1, \dots, T_n) = K_i(T'_1, \dots, T'_n), \quad i \in N. \quad (4.9.7)$$

Доказательство. Рассмотрим две различные ситуации в тип-стратегиях $T = (T_1, \dots, T_n)$, $T' = (T'_1, \dots, T'_n)$. Из построения T (T') следует, что она образует абсолютное равновесие в Γ .

Предположим, что (4.9.7) не выполняется при некотором $i_1 \in N$:

$$K_{i_1}(T_1, \dots, T_n) \neq K_{i_1}(T'_1, \dots, T'_n). \quad (4.9.8)$$

Рассмотрим два различных случая:

1. $T_{i_1}(x_0) = T'_{i_1}(x_0) = k$;
2. $T_{i_1}(x_0) \neq T'_{i_1}(x_0)$.

В случае 1 в ситуациях T и T' после выбора в позиции x_0 игроки попадают в одну и ту же подыгру Γ_x , где $x = x(k; x_0)$.

По рекуррентному построению ситуаций T , T' выигрыши в подыгре Γ_x в этих ситуациях соответственно равны:

$$K_i^x(T_1^x, \dots, T_n^x) = K_i^x(T_1'^x, \dots, T_n'^x) = v_i(k; x_0), \quad (4.9.9)$$

$i = 1, \dots, n$, где T_i^x ($T_i'^x$) — след стратегии T_i (T_i') в подыгре Γ_x .

Однако в случае 1

$$K_i(T_1, \dots, T_n) = K_i^x(T_1^x, \dots, T_n^x), K_i(T_1', \dots, T_n') = K_i^x(T_1'^x, \dots, T_n'^x), \quad (4.9.10)$$

$i = 1, \dots, n$, и теорема следует из (4.9.9), (4.9.10).

Пусть в случае 2 $T_{i_1}(x_0) = k_1$, $T'_{i_1}(x_0) = k_2$ ($k_1 \neq k_2$) и $x_1 = x(k_1; x_0)$, $x_2 = x(k_2; x_0)$. Тогда по лемме 2

$$\begin{aligned} K_i^{x_1}(T_1^{x_1}, \dots, T_n^{x_1}) &= v_i(k_1; x_0) = \\ &= v_i(k_2; x_0) = K_i^{x_2}(T_1'^{x_2}, \dots, T_n'^{x_2}), \quad i = 1, \dots, n. \end{aligned} \quad (4.9.11)$$

Однако

$$\begin{aligned} K_i(T_1, \dots, T_n) &= K_i^{x_1}(T_1^{x_1}, \dots, T_n^{x_1}), \\ K_i(T_1', \dots, T_n') &= K_i^{x_2}(T_1'^{x_2}, \dots, T_n'^{x_2}), \end{aligned} \quad (4.9.12)$$

$i = 1, \dots, n$, и в случае 2 теорема следует из (4.9.11), (4.9.12).

Положение намного сложнее, если игроки не знают типов противника или знают лишь некоторые вероятностные распределения на типах.

Пример 11. Рассмотрим игру двух лиц (рис. 4.13). Элементы множества X_1 обозначим кружками, а множества X_2 — квадратиками. Вершины дерева игры перенумерованы двойными индексами.

В позициях (1.2) и (1.3) оба выбора игрока 1 приводят к одинаковому для него выигрышу. В то же время, если $f_1(2) = 2$ (игрок 2 «друг» игрока 1), то он выберет альтернативу 2 в позиции (1.2) и альтернативу 1 в позиции (1.3). Если $f_1(2) = -2$ (игрок 2 «враг» игрока 1), он выберет 1 в позиции (1.2) и 2 в позиции (1.3).

В случае $F_1 = (1, 2)$ ($f_1(2) = 2$) единственное абсолютное равновесие по Нэшу будет $\{(2, 2, 1, 1, 1), (2, 1)\}$ с выигрышами $(2, 1)$, если оба игрока знают вектор предпочтений F_1 . Если это не имеет места, то положение усложняется. Пусть игрок 2 думает, что вектор предпочтений 1 есть $F_1 = (1, -2)$ ($f_1(2) = -2$) (хотя в действительности $f_1(2) = 2$), и игрок 1 знает о мнении игрока 2 о себе. Тогда, используя метод динамического программирования, мы получим пару стратегий $\{(1, 2, 1, 1, 1), (1, 1)\}$ с выигрышами $(5, 4)$, что не есть равновесие по Нэшу.

Если игрок 1 не знает мнение игрока 2 о себе, решение будет другим.

Также интересно отметить, что в случае, когда вектор F_1 известен обоим игрокам и $f_1(2) = -2$, единственное равновесие по Нэшу в тип-стратегиях будет $\{(1, 1, 2, 1, 1), (1, 1)\}$ с выигрышами $(5, 3)$. Это довольно необычно, так как оказывается, что дружеские отношения между игроками могут иногда приводить к меньшим выигрышам для обоих игроков, чем враждебные.

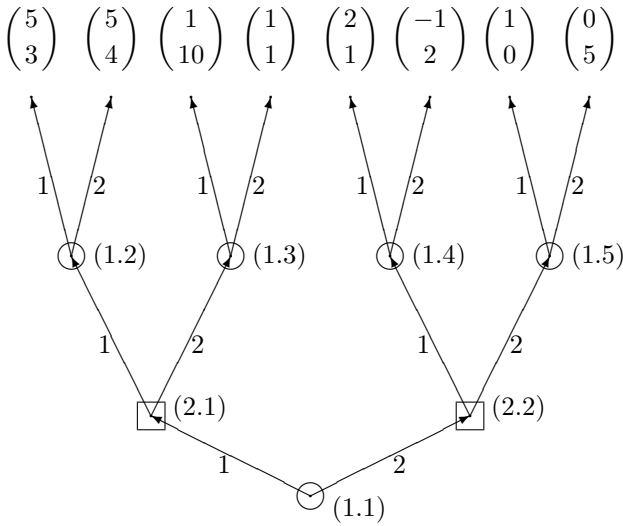


Рис. 4.13. Пример, когда выигрыши в «благожелательном» равновесии меньше, чем в «неблагожелательном»

§ 4.10. Структура множества абсолютных равновесий по Нэшу

4.10.1. В п. 4.2.1 была доказана теорема о существовании ситуации абсолютно-го равновесия по Нэшу в конечной игре с полной информацией в чистых стратегиях. В данном параграфе впервые дается полное конструктивное описание всех ситуаций абсолютного равновесия по Нэшу в играх с полной информацией на древовидном графе [Мамкина, Петросян, 2004]. Для этого построен специальный класс абсолютных равновесий по Нэшу в конечной игре с полной информацией в стратегиях поведения, предусматривающих возможность случайного выбора альтернатив, и показано, что любая ситуация абсолютного равновесия принадлежит данному классу.

Пусть задана многошаговая позиционная игра с полной информацией на конечном древовидном графе (см. определение п. 4.1.4).

Для упрощения дальнейшего изложения введем некоторые дополнительные обозначения. Пусть x — некоторая вершина (позиция), $G(x)$ — дерево подыгры Γ_x , т. е. поддереву дерева игры G с началом в вершине x . Обозначим через $Z(x)$ множество вершин, непосредственно следующих за x , $Z(x) = F_x$. Альтернативами в вершине x называются вершины y , следующие за x ($y \in Z(x)$). Игрока i , принимающего решение в позиции x (выбирающего следующую альтернативу в вершине x), будем обозначать через $i(x)$. Обозначим выбор игрока $i(x)$ в позиции x через $\bar{x} \in Z(x)$, а подыгру Γ_x для удобства через $\Gamma(x)$.

Нам также удобнее будет в этом параграфе множества X_i обозначить через P_i ($X_i \equiv P_i$), $i = 1, \dots, n + 1$ (см. определение п. 4.1.4). Заметим, что для простоты изложения в этой главе мы рассматриваем лишь игры с терминальным выигрышем. Однако все

результаты могут быть перенесены и на игры, выигрыши в которых определены во всех позициях дерева G .

Обычные стратегии, определенные в п. 4.1.5, будем называть чистыми стратегиями, т. е. чистой стратегией игрока i как и ранее называется однозначное отображение $u_i(\cdot)$, которое каждой позиции $x \in P_i$ ставит в соответствие некоторую альтернативу $y \in Z(x)$.

Множество всех чистых стратегий игрока i будем обозначать через U_i .

4.10.2. Определение. Стратегией поведения игрока i называется однозначное отображение $b_i(\cdot)$, которое каждой позиции $x \in P_i$ ставит в соответствие некоторое вероятностное распределение

$$p_x = p_{x_1}(y), \dots, p_{x_{|Z(x)|}}(y),$$

$$p_{x_k}(y) \geq 0, 1 \leq k \leq |Z(x)|$$

и

$$\sum_{y \in Z(x)} p_{x_k}(y) = 1$$

на множестве альтернатив $Z(x)$ (здесь $|Z(x)|$ — число элементов множества $Z(x)$).

Множество всевозможных стратегий поведения игрока i будем обозначать через B_i .

Введение класса стратегий поведения позволяет удобным образом описать множество всех ситуаций абсолютного равновесия в конечной игре с полной информацией. В данном параграфе мы построим специальный класс абсолютных равновесий по Нэшу в конечной игре с полной информацией в стратегиях поведения и покажем, что любая ситуация абсолютного равновесия принадлежит данному классу.

4.10.3. Под длиной игры $\Gamma(x_0)$ будем понимать длину наибольшего пути (число вершин содержащихся в пути) на дереве $G(x_0)$.

Пусть длина игры $\Gamma(x_0)$ равна $T+1$. Рассмотрим разбиение множества всех позиций дерева игры $G(x_0)$ на $T+1$ множество $X_0, X_1, \dots, X_T = \{x_0\}$, где множество X_t состоит из позиций, достигаемых из начальной позиции x_0 в точности за $T-t$ ходов. Обозначим множество позиций, принадлежащих X_t , через x_t , $t = 0, \dots, T$.

При построении абсолютного равновесия по Нэшу методом обратной индукции может случиться, что в некоторой подыгре $\Gamma(x)$ на поддереве $G(x)$ игрок $i(x)$, совершающий выбор в позиции x , обнаруживает, что при продолжении игры выбором некоторых альтернатив $y \in Z(x)$ его выигрыш, при условии что игроки и далее будут придерживаться данного фиксированного абсолютного равновесия по Нэшу, принимает равные значения (т. е. для игрока $i(x)$ выбор любой из таких альтернатив представляет одинаковый интерес). Именно это обстоятельство и приводит к множественности ситуаций абсолютного равновесия по Нэшу в конечных играх с полной информацией. Следующее определение вполне аналогично определению п. 4.2.1 и приводится для стратегий поведения.

Определение. Ситуация равновесия по Нэшу в стратегиях поведения $b(\cdot) = (b_1(\cdot), \dots, b_n(\cdot))$ называется абсолютным равновесием в игре $\Gamma(x_0)$, если ее сужение в любой подыгре $\Gamma(x)$ является ситуацией равновесия в этой подыгре.

4.10.4. Приведем алгоритм построения ситуации абсолютного равновесия по Нэшу в стратегиях поведения $((\bar{b}_1(\cdot), \dots, \bar{b}_n(\cdot))) = \bar{b}(\cdot)$ для конечной игры с полной информацией и покажем, что любая ситуация абсолютного равновесия в $\Gamma(x_0)$ может быть получена с помощью этого алгоритма.

Рассмотрим множество позиций X_0 . Так как длина игры равна $T + 1$, то $X_0 \subset P_{n+1}$ и выигрыши игроков уже определены и равны соответственно $h_i(x_0)$, $x_0 \in X_0$, $i = 1, \dots, n$.

Шаг 1. Перейдем от позиций X_0 к позициям $x_1 \in X_1$. Если $x_1 \notin P_{n+1}$, то в позиции x_1 ходит игрок $i(x_1)$. Алгоритм предписывает игроку $i(x_1)$ выбрать любую из позиций (альтернатив) $x \in Z(x_1)$ из условия:

$$\max_{x \in Z(x_1)} h_{i(x_1)}(x) = h_{i(x_1)}(\bar{x}_0). \quad (4.10.1)$$

Однако, максимум в (4.10.1) может достигаться не в одной точке. Обозначим через

$$\tilde{Z}_{i(x_1)}(x_1) = \arg \max_{x \in Z(x_1)} h_i(x),$$

т. е.

$$\tilde{Z}_{i(x_1)}(x_1) = \left\{ y : h_{i(x_1)}(y) = \max_{x_0 \in Z(x_1)} h_{i(x_1)}(x_0) \right\}. \quad (4.10.2)$$

В позиции $x_1 \in X_1$ стратегия поведения $\bar{b}_{i(x_1)}$ игрока $i(x_1)$ предписывает выбор любой из альтернатив множества $y \in \tilde{Z}_{i(x_1)}(x_1)$ с некоторыми вероятностями $p_{x_1}(y) \geq 0$,

$\sum_{y \in \tilde{Z}_{i(x_1)}(x_1)} p_{x_1}(y) = 1$, которые определяются стратегией $\bar{b}_{i(x_1)}$. Если максимум (4.10.1)

достигается в единственной точке \bar{x}_0 , т. е. $|\tilde{Z}_{i(x_1)}(x_1)| = 1$, то стратегия поведения $\bar{b}_{i(x_1)}$ игрока $i(x_1)$ в x_1 предписывает ему выбор точки $x_0 \in Z(x_1)$ с вероятностью 1. При этом выигрыши игроков $i = 1, \dots, n$ равны

$$\sum_{x_0 \in \tilde{Z}_{i(x_1)}(x_1)} p_{x_1}(x_0) h_i(x_0).$$

Если $x_1 \in P_{n+1}$, то выигрыши игроков уже определены и равны соответственно $h_i(x_1)$, $i = 1, \dots, n$.

Заметим, что из-за возможного попадания игрока $i(x_1)$ в вершину $x_1 \in \tilde{Z}_{i(x_1)}(x_1)$, $|\tilde{Z}_{i(x_1)}(x_1)| > 1$, предполагаемая схема выбора не определяет путь (дугу) однозначно, и мы получим некоторое «поддерево» дерева $G(x_1)$. В случае если $|\tilde{Z}_{i(x_1)}(x_1)| = 1$, то «поддерево» будет состоять из единственной дуги (или одной вершины, если $x_1 \in P_{n+1}$).

Применяя аналогичные рассуждения, можно построить «поддерево» с началом в $x_1 \in X_1$ для каждой позиции $x_1 \in X_1$. Таким образом, на каждом поддереве $G(x_1)$, $x_1 \in X_1$ фиксируется позиция \bar{x}_0 , являющаяся предполагаемой окончательной позицией строящегося «поддерева» игры $\Gamma(x_0)$, или вероятностное распределение на $\tilde{Z}_{i(x_1)}(x_1)$, если $|\tilde{Z}_{i(x_1)}(x_1)| > 1$. Поэтому, зная поведение игроков на поддеревьях $G(x_1)$, $x_1 \in X_1$, мы можем ввести аналог функций Беллмана

$$H_i^1 : X_1 \rightarrow R^1, i = 1, \dots, n.$$

Здесь H_i^1 — ожидаемый выигрыш игрока i в позиции $x_1 \in X_1$, в предположении, что на поддереве $G(x_1)$ игроки действуют согласно предложенному алгоритму:

$$H_i^1(x_1) = \begin{cases} h_i(x_1), & \text{если } x_1 \in P_{n+1}, \\ \sum_{y \in \tilde{Z}_{i(x_1)}(x_1)} p_{x_1}(y) h_i(y), & \text{если } x_1 \notin P_{n+1}, \end{cases} \quad (4.10.3)$$

где

$$\sum_{y \in |\tilde{Z}_{i(x_1)}(x_1)|} p_{x_i}(y) = 1. \quad (4.10.4)$$

Стратегии $\bar{b}_i(\cdot)$ строим в позиции $x_1 \in X_1 \cap P_i$ по правилу:

$$\bar{b}_i(\cdot) = \begin{cases} \bar{x}_0, \bar{x}_0 \in \arg \max_{y \in Z(x_1)} h_{i(x_1)}(y), & |\tilde{Z}_{i(x_1)}(x_1)| = 1, \\ p_{x_1} = \{p_{x_1}(y), y \in \tilde{Z}_{i(x_1)}(x_1)\}, & |\tilde{Z}_{i(x_1)}(x_1)| > 1. \end{cases} \quad (4.10.5)$$

Здесь $\sum_{y \in \tilde{Z}_{i(x_1)}(x_1)} p_{x_1}(y) = 1$, $p_{x_1}(y) \geq 0$ — некоторое фиксированное вероятностное распределение на множестве альтернатив $\tilde{Z}_{i(x_1)}(x_1)$.

Предположим, что функции $H_i^l(x_l)$ и стратегии $\bar{b}_i(\cdot)$, $i \in N$ построены для всех $l < t$.

Рассмотрим шаг t . Предположим, что при движении к корню игры мы достигли позиции $x_t \in X_t$. Пусть функции

$$H_i^{t-1}(x_{t-1}) : X_{t-1} \rightarrow R^1, i \in N$$

определяют, какие выигрыши получают игроки $i \in N$ в подыгре $\Gamma(x_{t-1})$ после выбора ими в позиции $x_t \in X_t$, $l < t$ предложенных нами решений.

Шаг t . Пусть далее

$$\tilde{Z}_{i(x_t)}(x_t) = \arg \max_{y \in Z(x_1)} H_i^{t-1}(y). \quad (4.10.6)$$

В позиции $x_t \in X_t$ стратегия поведения $\bar{b}_i(x_t)$ игрока $i(x_t)$ выбирает любую из альтернатив множества $y \in \tilde{Z}_{i(x_t)}(x_t)$ с некоторыми вероятностями

$$p_{x_t}(y) \geq 0, \quad \sum_{y \in |\tilde{Z}_{i(x_1)}(x_1)|} p_{x_t}(y) = 1.$$

При этом выигрыши игроков $i \in N$ в подыграх $\Gamma(x_t)$, $x_t \in X_t$ равны

$$\sum_{x_{t-1} \in \tilde{Z}_{i(x_t)}(x_1)} p_{x_t}(x_{t-1}) H_i^{t-1}(x_{t-1}), i \in N.$$

Понятно, что если максимум (4.10.6) достигается в единственной точке \bar{x}_t , то $\bar{b}_i(x_t)$ выбирает эту точку с вероятностью 1.

Если $x_t \in P_{n+1}$, то выигрыши игроков уже определены и равны соответственно $h_i(x_t)$, $i = 1, \dots, n$.

Функции $H_i^t(x_t) : X_t \rightarrow R^1$ $i \in N$ зададим следующим образом:

$$H_i^1(x_t) = \begin{cases} h_i(x_t), & \text{если } x_t \in P_{n+1}, \\ \sum_{y \in |\tilde{Z}_{i(x_t)}(x_t)|} p_{x_t}(y) h_i(y), & \text{если } x_t \notin P_{n+1}. \end{cases} \quad (4.10.7)$$

Стратегии $\bar{b}_i(\cdot)$ строим в позиции $x_t \in X_t \cap P_i$, $i \in N$ по правилу:

$$\bar{b}_i(\cdot) = \begin{cases} \bar{x}_{t-1}, \bar{x}_{t-1} \in \arg \max_{y \in Z(x_t)} h_{i(x_t)}(y), & |\tilde{Z}_{i(x_t)}(x_t)| = 1, \\ p_{x_t} = \{p_{x_t}(y), y \in \tilde{Z}_{i(x_t)}(x_t)\}, & |\tilde{Z}_{i(x_t)}(x_t)| > 1. \end{cases} \quad (4.10.8)$$

Здесь $\sum_{y \in \tilde{Z}_{i(x_t)}(x_t)} p_{x_t}(y) = 1$, $p_{x_t}(y) \geq 0$.

Обозначим через $\bar{b}^{x_t}(\cdot) = (\bar{b}_1^{x_t}(\cdot), \dots, \bar{b}_n^{x_t}(\cdot))$ ситуацию в подыгре $\Gamma(x_t)$, построенную на первых t шагах алгоритма.

Продолжая спускаться по дереву игры $\Gamma(x_0)$ к начальной позиции O и последовательно определяя выборы игроков в оставшихся множествах X_τ , $\tau = t+1, \dots, T$, мы построим поддерево, соответствующее ситуации $\bar{b}^{x_0}(\cdot) = (\bar{b}_1^{x_0}(\cdot), \dots, \bar{b}_n^{x_0}(\cdot)) = \bar{b}(\cdot)$, и соответствующие выигрыши, которые реализуются в игре $\Gamma(x_0)$.

4.10.5. Теорема. Построенная ситуация $\bar{b}(\cdot) = (\bar{b}_1(\cdot), \dots, \bar{b}_n(\cdot))$ образует ситуацию абсолютного равновесия по Нэшу в $\Gamma(x_0)$.

Доказательство теоремы вполне аналогично доказательству теоремы п. 4.2.1 и при первом чтении может быть опущено.

Доказательство. Обозначим через

$$K_i(x, \bar{b}_1^x(\cdot), \dots, \bar{b}_n^x(\cdot)) = K_i(x; \bar{b}^x(\cdot))$$

математическое ожидание выигрышей игроков $i \in N$ в подыгре $\Gamma(x)$ в ситуации $(\bar{b}_1^x(\cdot), \dots, \bar{b}_n^x(\cdot)) = b^x(\cdot)$ в стратегиях поведения. Тогда, по определению, в ситуации $(\bar{b}_1^{x_t}(\cdot), \dots, \bar{b}_n^{x_t}(\cdot)) = b^{x_t}(\cdot)$ имеет место равенство

$$K_i(x; \bar{b}^{x_t}(\cdot)) = H_i^t(x_t), i \in N, x_t \in X_t.$$

Для доказательства теоремы достаточно показать, что

$$K_i(x; \bar{b}^{x_t}(\cdot)) \geq K_i(x_t; \bar{b}^{x_t}(\cdot) \| b_i^{x_t}(\cdot)) \quad (4.10.9)$$

для всех $x_t \in X_t$, $i \in N$, $b_i^{x_t}(\cdot) \in B_i^{x_t}$. Здесь $B_i^{x_t}$ — множество стратегий поведения игрока i в подыгре $\Gamma(x_t)$, представляющее сужение множества B_i на позиции подыгры $\Gamma(x_t)$.

Проведем доказательство индукцией по длине игры. Если длина игры равна 1, то теорема справедлива, поскольку все позиции в игре окончательные и игроки в них ходов не делают. Пусть теперь игра Γ имеет длину T . Рассмотрим семейство подыгр $\Gamma(x_{t-1})$ на поддереве $G(x_{t-1})$, $x_{t-1} \in Z(x_t)$. Длина каждой из этих подыгр не превосходит $T-1$. Предположим, что теорема справедлива для всех игр, длина которых не превосходит $T-1$, и докажем ее для игры длины T .

Пусть $i = i(x_T)$. Поскольку $\bar{b}_i^{x_{T-1}}(\cdot)$ по индукционному предположению есть ситуация абсолютного равновесия в $\Gamma(x_{T-1})$, то имеем:

$$\begin{aligned} K_{i(x_T)}(x_T; \bar{b}^{x_T}(\cdot)) &= H_{i(x_T)}^T(x_T) = \max_{x_{T-1} \in Z(x_T)} H_{i(x_T)}^{T-1}(x_{T-1}) = \\ &= H_{i(x_T)}^{T-1}(\bar{x}_{T-1}) \geq H_{i(x_T)}^{T-1}(x_{T-1}) = K_{i(x_T)}(x_{T-1}; \bar{b}_i^{x_{T-1}}(\cdot)) \geq \\ &\geq K_{i(x_T)}(x_{T-1}; \bar{b}^{x_{T-1}}(\cdot) \| b_{i(x_T)}(\cdot)) = \\ &= K_{i(x_T)}(x_T; \bar{b}^{x_T}(\cdot) \| b_{i(x_T)}(\cdot)). \end{aligned}$$

Если $i \neq i(x_T)$, то

$$\begin{aligned} K_i(x_{T-1}; \bar{b}^{x_T}(\cdot)) &= \sum_{y \in \tilde{Z}_{i(x_T)}(x_T)} p_{x_T}(y) H_i^{T-1}(y) = \\ &= \sum_{y \in \tilde{Z}_{i(x_T)}(x_T)} p_{x_T}(y) K_i(x_{T-1}; \bar{b}^y(\cdot)) \geq \\ &\geq \sum_{y \in \tilde{Z}_{i(x_T)}(x_T)} p_{x_T}(y) K_i(x_{T-1}; \bar{b}^y(\cdot) \| b_i(y)) = \\ &= K_i(x_T; \bar{b}^{x_T}(\cdot) \| b_i(\cdot)) \end{aligned}$$

и теорема доказана.

4.10.6. Докажем теперь в некотором смысле обратное утверждение, а именно, докажем следующую теорему.

Теорема. Любая ситуация абсолютного равновесия по Нэшу в $\Gamma(x_0)$ может быть получена в результате реализации построенного алгоритма при соответствующим образом выбранных вероятностных распределениях

$$p_x(y) \geq 0, \quad x \in P_i, \quad \sum_{y \in \tilde{Z}_{i(x)}(x)} p_x(y) = 1. \quad (4.10.10)$$

Доказательство. Обозначим через B^* множество всевозможных ситуаций абсолютного равновесия по Нэшу в $\Gamma(x_0)$. И покажем, что $B^* \subset \bar{B}$. Пусть $b^* \in B^*$. Покажем, что существует такая ситуация в стратегиях поведения $\bar{b} \in \bar{B}$, предписывающая поведения

$$p_x(y) \geq 0, x \in P_i, \quad \sum_{y \in \tilde{Z}_{i(x)}(x)} p_x(y) = 1, i \in N, \text{ что } b^* = \bar{b};$$

здесь \bar{B} — класс абсолютных равновесий по Нэшу в $\Gamma(x_0)$, полученный в результате реализации предложенного алгоритма. Обозначим через $\bar{B}_i^{x_t}$ и B_i^* сужение множества \bar{B}_i и B_i^* соответственно на позиции подыгры $\Gamma(x_t)$. Рассмотрим семейство подыгр $\Gamma(x_1)$. Пусть $b^* \in B^*$ и пусть не существует $\bar{b} \in \bar{B}$, порождающей поведения (вероятностные распределения), определенные формулой (4.10.10) для позиции x_1 : $b^* = \bar{b}$, $\bar{b} \in \bar{B}$. Так как длина игры $\Gamma(x_1)$ не превышает 2, то в игре $\Gamma(x_1)$ ходит только один игрок $i(x_1)$. Следовательно, ситуации b^{*x_1} и \bar{b}^{x_1} отличаются разве лишь стратегиями игрока $i(x_1)$ $b_{i(x_1)}^{*x_1}(x_1)$ и $\bar{b}_{i(x_1)}^{x_1}(x_1)$ соответственно. т. е. \bar{b}^{x_1} может быть представлено как $b^{*x_1} \| b_{i(x_1)}^{x_1}$ для некоторого $b_{i(x_1)}(x_1) = \bar{b}_{i(x_1)}(x_1)$.

Так как b^* — ситуация абсолютного равновесия по Нэшу в $\Gamma(x_0)$, то ее сужение на любую из подыгр игры $\Gamma(x_0)$ также является ситуацией абсолютного равновесия в соответствующей подыгре. Поэтому

$$K_{i(x_1)}(x_1; b^{*x_1}(\cdot)) \geq K_{i(x_1)}(x_1; b^{*x_1}(\cdot) \| b_{i(x_1)}^{x_1}(\cdot)) = K_{i(x_1)}(x_1; \bar{b}^{x_1}(\cdot)). \quad (4.10.11)$$

Если $|\tilde{Z}_{i(x_1)}(x_1)| = |Z(x_1)|$, то, очевидно, что в этих позициях b^* имеет структуру \bar{b} , т. е. $b^* \in \bar{B}$. Пусть существует $x \in Z(x_1)$: $x \notin \tilde{Z}_{i(x_1)}(x_1)$. Если b^{*x_1} не имеет структуры

\bar{b} , то это означает, что вероятности $p_x(y)$, порожденные b^{*x_1} , должны предписывать положительную меру вершинам $x \in Z(x_1)$, $x \notin \tilde{Z}_{i(x_1)}(x_1)$. Тогда

$$\begin{aligned} K_{i(x_1)}(x_1; \bar{b}^{x_1}(\cdot)) &= H_{i(x_1)}^1(x_1) = \max_{x \in Z(x_1)} h_{i(x_1)}(x) > \\ &> \sum_{y \in Z(x_1)} p_{x_1}(y) h_{i(x_1)}(y) = K_{i(x_1)}(x_1; b^{*x_1}(\cdot)). \end{aligned}$$

Из (4.10.11) и последней цепочки равенств и неравенств следует, что наше утверждение неверно и $b^{*x_1} = \bar{b}^{x_1}$ для некоторого $\bar{b} \in \bar{B}$.

Рассмотрим семейство подыгр $\Gamma(x_2)$. Мы показали, что $b^{*x_1} = \bar{b}^{x_1}$ для некоторого $\bar{b} \in \bar{B}$. Следовательно, ситуации b^{*x_2} и \bar{b}^{x_2} отличаются разве лишь стратегиями игрока $i(x_2)$ в позиции x_2 .

$$K_{i(x_2)}(x_2; b^{*x_2}(\cdot)) \geq K_{i(x_2)}(x_2; b^{*x_2}(\cdot) \| b_{i(x_2)}^{x_2}(\cdot)) = K_{i(x_2)}(x_2; \bar{b}^{x_2}(\cdot)), \quad (4.10.12)$$

так как b^{*x_2} является ситуацией абсолютного равновесия в $\Gamma(x_2)$.

Если $|\tilde{Z}_{i(x_1)}(x_1)| = |Z(x_1)|$, то утверждение теоремы тривиально.

Пусть существует $x \in Z(x_2) : x \notin \tilde{Z}_{i(x_2)}(x_2)$. Если b^{*x_2} не имеет структуры \bar{b} , то это означает, что вероятности $p_x(y)$, порожденные b^{*x_2} , должны предписывать положительную меру вершинам $x \in Z(x_2)$, $x \notin \tilde{Z}_{i(x_2)}(x_2)$. Тогда

$$\begin{aligned} K_{i(x_2)}(x_2; \bar{b}^{x_2}(\cdot)) &= H_{i(x_2)}^1(x_2) = \max_{x \in Z(x_2)} H_{i(x_2)}^1(x) > \\ &> \sum_{y \in Z(x_2)} p_{x_2}(y) H_{i(x_2)}^1(y) = K_{i(x_2)}(x_2; b^{*x_2}(\cdot)). \end{aligned} \quad (4.10.13)$$

Это противоречит тому, что $b^* \in B^*$. Из (4.10.12) и (4.10.13) следует, что $b^{*x_2} = \bar{b}^{x_2}$ для некоторого вероятностного распределения, определенного условием (4.10.10) для позиции x_2 (для некоторого $\bar{b} \in \bar{B}$).

Предположим, что утверждение теоремы справедливо для всех подыгр длины k , т. е. существует вероятностное распределение, определенное условием (4.10.10) для всех позиций x_k , такое что $b^{*x_k} = \bar{b}^{x_k}$. Рассмотрим семейство подыгр $\Gamma(x_{k+1})$ и покажем, что $b^{*x_{k+1}} = \bar{b}^{x_{k+1}}$ для некоторого $\bar{b} \in \bar{B}$. Пусть это не так. Тогда имеем:

$$\begin{aligned} K_{i(x_{k+1})}(x_{k+1}; b^{*x_{k+1}}(\cdot)) &\geq K_{i(x_{k+1})}(x_{k+1}; b^{*x_{k+1}}(\cdot) \| b_{i(x_{k+1})}^{x_{k+1}}(\cdot)) = \\ &= K_{i(x_{k+1})}(x_{k+1}; \bar{b}^{x_{k+1}}(\cdot)). \end{aligned}$$

Пусть существует $x \in Z(x_{k+1})$, $x \notin \tilde{Z}_{i(x_{k+1})}(x_{k+1})$, которое выбирается с положительной вероятностью стратегией $b^{*x_{k+1}}$:

$$\begin{aligned} K_{i(x_{k+1})}(x_{k+1}; \bar{b}^{x_{k+1}}(\cdot)) &= H_{i(x_{k+1})}^{k+1}(x_{k+1}) = \max_{x \in Z(x_{k+1})} H_{i(x_{k+1})}^k(x) > \\ &> \sum_{y \in Z(x_{k+1})} p_{x_{k+1}}(y) H_{i(x_{k+1})}^k(y) = K_{i(x_{k+1})}(x_{k+1}; b^{*x_{k+1}}(\cdot)). \end{aligned}$$

Но это противоречит тому, что $b^* \in B^*$. Поэтому $b^{*x_{k+1}} = \bar{b}^{x_{k+1}}$ для некоторого $\bar{b} \in \bar{B}$. Для случая, когда $|\tilde{Z}_{i(x_{k+1})}(x_{k+1})| = |Z(x_{k+1})|$ утверждение теоремы тривиально. Теорема доказана.

§ 4.11. Индифферентное равновесие в позиционных играх

4.11.1. Как мы видели ранее в § 4.9—4.10, в позиционных играх ситуация абсолютного равновесия по Нэшу может не являться единственной и зависит от «доброжелательности» игроков в том смысле, что один из игроков, будучи в равной степени заинтересован в выборе последующих альтернатив, может выбрать любую из таких вершин, руководствуясь своими личными соображениями. Например, в случае «доброжелательности», игрок во множестве своих личных позиций, в которых выбор последующих альтернатив принесет ему одинаковый максимальный выигрыш, выбирает из них ту, которая более благоприятна для другого игрока, или же ту, которая является неблагоприятной для какого-либо из игроков (см. § 4.9). В настоящем параграфе мы предложим другой подход выделения абсолютного равновесия по Нэшу, которое может быть получено только при использовании смешанных стратегий.

Пример 12. На рис. 4.14 представлена позиционная игра двух лиц с полной информацией на древовидном графе. Выигрыши игроков записаны в окончательных позициях. Причем, выигрыш первого игрока соответствует верхнему числу, второго — нижнему.

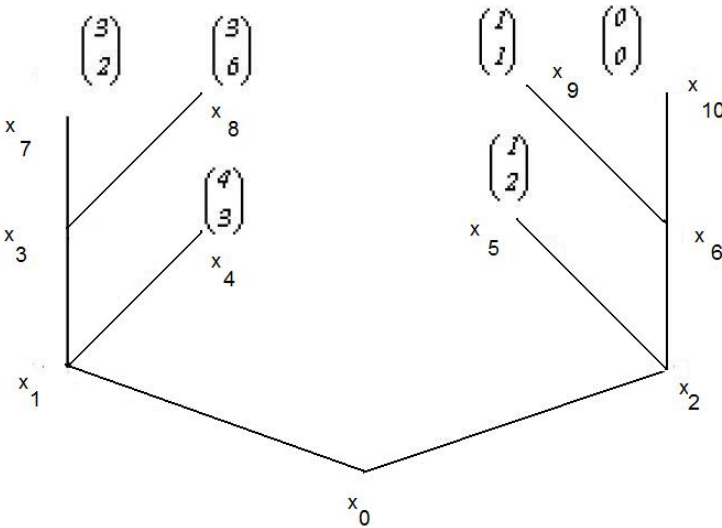


Рис. 4.14. Дерево игры

Множество очередности игрока 1 представляет собой множество $P_1 = \{x_0, x_3, x_6\}$, второго игрока — $P_2 = \{x_1, x_2\}$, множество окончательных позиций равно $P_3 = \{x_4, x_5, x_7, x_8, x_9, x_{10}\}$.

В позиции x_3 ходит игрок 1. Поскольку максимум

$$\max_{x \in Z(x_3)} h_1(x)$$

достигается в 2-х точках, построим множество

$$\tilde{Z}_1(x_3) = \arg \max_{x \in Z(x_3)} h_1(x) = \{x_7, x_8\}.$$

Предположим, что игрок 1 настроен доброжелательно по отношению к игроку 2. Это означает, что в позиции x_3 он выбирает позицию $x_8 \in \tilde{Z}_1(x_3)$, которая является более благоприятной для игрока 2, чем позиция $x_7 \in \tilde{Z}_1(x_3)$. Тогда выигрыши игроков в позиции x_3 в подыгре $\Gamma(x_3)$ составят $H(x_3) = (3, 6)^*$.

В позиции x_6 ходит игрок 1. Поскольку максимум

$$\max_{x \in Z(x_6)} h_1(x) = 1$$

достигается в x_9 , следовательно, в позиции x_6 игрок 1 выберет альтернативу x_9 . Тогда в подыгре $\Gamma(x_6)$ выигрыши игроков составят $H(x_6) = (1, 1)^*$.

В позиции x_1 ходит игрок 2. Так как максимум

$$\max_{x \in Z(x_1)} h_2(x) = 6$$

достигается в x_3 , следовательно, в позиции x_1 игрок 2 выберет альтернативу x_3 . Тогда в подыгре $\Gamma(x_1)$ выигрыши игроков составят

$$H(x_1) = (3, 6)^*.$$

В позиции x_2 ходит игрок 2. Максимум выигрыша второго игрока

$$\max_{x \in Z(x_2)} h_2(x) = 2$$

достигается в x_5 . Поэтому в позиции x_2 игрок 2 выберет альтернативу x_5 . Тогда в подыгре $\Gamma(x_2)$ выигрыши игроков составят $H(x_2) = (1, 2)^*$.

Рассмотрим позицию x_0 . В позиции x_0 ходит игрок 1 и выбирает альтернативу $x_1 \in Z(x_0)$ из условия

$$\max_{x \in Z(x_0)} h_1(x) = 3.$$

Тогда выигрыш игроков во всей игре $\Gamma(x_0)$ составит $H(x_0) = (3, 6)^*$.

Заметим, что в позиции x_3 множество $\tilde{Z}_1(x_3)$ состоит более чем из одного элемента, что порождает неоднозначность выбора альтернативы в позиции x_3 . Выше мы строили абсолютное равновесие по Нэшу в предположении «доброжелательности» первого игрока. Рассмотрим случай, когда игрок 1 настроен «недоброжелательно» по отношению к игроку 2. Это означает, что в позиции x_3 игрок 1 выбирает позицию $x_7 \in \tilde{Z}_1(x_3)$,

$$\tilde{Z}_1(x_3) = \arg \max_{x \in Z(x_3)} h_1(x) = \{x_7, x_8\},$$

которая является менее благоприятной для игрока 2, чем позиция $x_8 \in \tilde{Z}_1(x_3)$. Тогда выигрыши игроков в позиции x_3 в подыгре $\Gamma(x_3)$ составят $H(x_3) = (3, 2)^*$.

В позиции x_6 ходит игрок 1. Так как максимум

$$\max_{x \in Z(x_6)} h_1(x) = 1$$

достигается в x_9 , то в позиции x_6 игрок 1 выберет альтернативу x_9 . Тогда в подыгре $\Gamma(x_6)$ выигрыши игроков составят $H(x_6) = (1, 1)^*$.

В позиции x_1 ходит игрок 2. Так как максимум

$$\max_{x \in Z(x_1)} h_2(x) = 3$$

и достигается в x_4 , следовательно, в позиции x_1 игрок 2 выберет альтернативу x_4 . Тогда в подыгре $\Gamma(x_1)$ выигрыши игроков составят

$$H(x_1) = (4, 3)^*.$$

В позиции x_2 ходит игрок 2. Максимум выигрыша второго игрока

$$\max_{x \in Z(x_2)} h_2(x) = 2$$

достигается в x_5 . А значит, в позиции x_2 игрок 2 выберет альтернативу x_5 . Тогда в подыгре $\Gamma(x_2)$ выигрыши игроков составят $H(x_1) = (1, 2)^*$.

Рассмотрим позицию x_0 . В позиции x_0 ходит игрок 1 и выбирает альтернативу $x_1 \in Z(x_0)$ из условия

$$\max_{x \in Z(x_0)} h_1(x) = 4.$$

Тогда выигрыш игроков во всей игре $\Gamma(x_0)$ составит $H(x_0) = (4, 3)^*$.

4.11.2. Как видно из рассмотренного примера, при различном настрое одного из игроков (доброжелательном и недоброжелательном) абсолютные равновесия в игре $\Gamma(x_0)$ и соответствующие им выигрыши различны.

Как было упомянуто ранее, при построении абсолютного равновесия по Нэшу методом обратной индукции, в подыгре может случиться, что один из игроков обнаруживает, что его выигрыш (при условии продолжения игры в соответствии с данным равновесием по Нэшу в подыграх) не зависит от того, какую альтернативу он выберет. Для устранения проблемы многозначности абсолютных равновесий по Нэшу введем понятие *индифферентного равновесия*, отражающее «безразличие» при выборе альтернативы в указанных позициях. В позициях $x \in G(x_0)$, в которых принимающему решению игроку $i(x)$ безразлично, какую из альтернатив $y \in \tilde{Z}_{i(x)}(x)$ выбрать, предпишем игроку $i(x)$ выбрать вершины $y \in \tilde{Z}_{i(x)}(x)$ с равными вероятностями, т. е. вероятность выбора каждой из альтернатив $y_k \in \tilde{Z}_{i(x)}(x)$, $k = 1, \dots, |\tilde{Z}_{i(x)}(x)|$ в позиции $x \in G(x_0)$ равна $\frac{1}{|\tilde{Z}_{i(x)}(x)|}$.

Процедура построения индифферентного равновесия по Нэшу отличается от классического построения абсолютного равновесия методом обратной индукции только выбором стратегий и определением выигрыша игрока в позициях $y \in \tilde{Z}_{i(x)}(x)$. А именно,

$$H_i(x) = \sum_{y \in \tilde{Z}_{i(x)}(x)} p_x(y) H_i(y), i \in N,$$

где $H_i(y)$, $y \in \tilde{Z}_{i(x)}(x)$ — выигрыш игрока $i \in N$ в подыгре $\Gamma(y)$. Очевидно, что если $\tilde{Z}_{i(x)}(x)$ состоит из одного элемента, то выбор альтернативы $y \in \tilde{Z}_{i(x)}(x)$ осуществляется с вероятностью 1 и процедура построения индифферентного равновесия в позициях, в которых $|\tilde{Z}_{i(x)}(x)| = 1$, совпадает с классической процедурой построения абсолютного равновесия по Нэшу методом обратной индукции.

4.11.3. Вернемся к примеру 12 (рис. 4.14) и найдем для него индифферентное равновесие.

В позиции x_3 $\tilde{Z}_1(x_3) = \{x_7, x_8\}$. Согласно предложенному выше алгоритму, альтернатива из $\tilde{Z}_1(x_3)$ в x_3 выбирается с равными вероятностями. И выигрыши игроков в x_3 составят

$$H(x) = \left(\frac{1}{2}(3+3), \frac{1}{2}(2+6)\right)^* = (3, 4)^*.$$

В позиции x_6 ходит игрок 1. Поскольку максимум

$$\max_{x \in Z(x_6)} h_1(x) = 1$$

достигается в x_9 , то $|\tilde{Z}_1(x_6)| = 1$, и в позиции x_6 игрок 1 выберет альтернативу x_9 с вероятностью 1. Тогда в подыгре $\Gamma(x_6)$ выигрыши игроков составят $H(x_6) = (1, 1)^*$.

В позиции x_1 ходит игрок 2. Так как максимум

$$\max_{x \in Z(x_1)} H_2(x) = 4$$

достигается в x_3 , то $|\tilde{Z}_1(x_1)| = 1$, и в позиции x_1 игрок 2 выберет альтернативу x_3 с вероятностью 1. Тогда в подыгре $\Gamma(x_1)$ выигрыши игроков составят $H(x_1) = (3, 4)^*$.

Аналогично в позиции x_2 выигрыши игроков составят $H(x_1) = (1, 2)^*$. И выигрыш игроков во всей игре $\Gamma(x_0)$ составит $H(x_0) = (3, 4)^*$.

Как видно из примера 2, выигрыши игроков при индифферентном равновесии отличны от выигрышей в рассмотренных нами выше других равновесиях.

Вернемся к описанному нами в параграфе § 4.10 классу всех абсолютных равновесий по Нэшу. Полагая в нашем примере вероятности на $\tilde{Z}_1(x_3)$ равными $p_{x_3}(x_8) = 1$, $p_{x_3}(x_7) = 0$, получим абсолютное равновесие по Нэшу при «благоприятном» настрое игрока 1. Полагая $p_{x_3}(x_8) = 0$, $p_{x_3}(x_7) = 1$ получим абсолютное равновесие по Нэшу при «неблагоприятном» настрое игрока 1. В случае, когда $p_{x_3}(x_8) = p_{x_3}(x_7) = \frac{1}{2}$ мы получим индифферентное равновесие в стратегиях поведения.

Понятно, что результаты § 4.10—4.11 сохраняют силу и для случая, когда выигрыши игроков заданы не только в окончательных позициях, но и во всех вершинах дерева игры и суммируются вдоль партии игры.

§ 4.12. Стратегии наказания и «народные теоремы»

4.12.1. В п. 4.2.1 доказана теорема о существовании ситуации абсолютного равновесия (по Нэшу) в многошаговых играх с полной информацией на конечном древовидном графе. В то же время при исследовании конкретных игр этого класса можно обнаружить целое семейство ситуаций равновесия, сужения которых необязательно являются ситуациями равновесия во всех подыграх исходной игры. К числу таких ситуаций равновесия относятся равновесия в стратегиях наказания. Проиллюстрируем это понятие на примере.

Пример 13. Пусть игра Γ происходит на графе, изображенном на рис. 4.15. Множество $N = \{1, 2\}$ состоит из двух игроков. Кружками изображены вершины, составляющие множество X_1 , квадратами — множество X_2 . Вершины графа перенумерованы двойными индексами, дуги — одинарными.

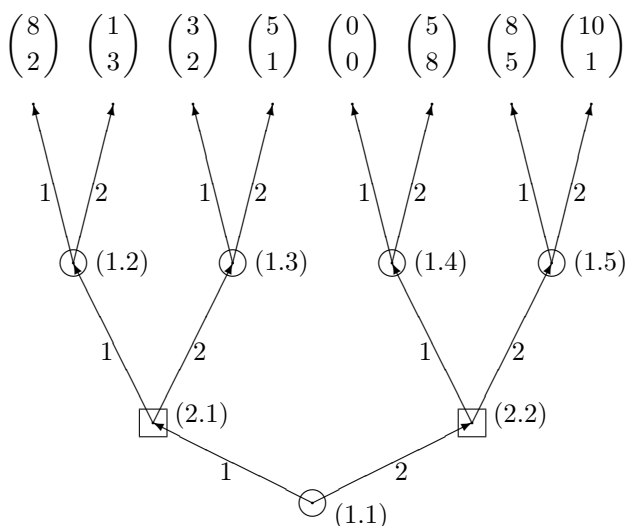


Рис. 4.15. Абсолютное равновесие и стратегии наказания

Нетрудно убедиться в том, что ситуация $u_1^* = (1, 1, 2, 2, 2)$, $u_2^* = (1, 1)$ является абсолютно равновесной в игре Γ . При этом выигрыши игроков равны 8 и 2 соответственно. Рассмотрим теперь ситуацию $\bar{u}_1 = (2, 1, 2, 1, 2)$, $\bar{u}_2 = (2, 2)$. В этой ситуации выигрыши игроков равны соответственно 10 и 1, тем самым игрок 1 получает больше, чем в ситуации (u_1^*, u_2^*) . Ситуация (\bar{u}_1, \bar{u}_2) является равновесной в игре Γ , но не является абсолютно равновесной. Действительно, в подыгре $\Gamma_{1,4}$ сужение стратегии \bar{u}_1 диктует игроку 1 выбор левой дуги, что не является для него оптимальным в позиции (1.4). Такое действие игрока 1 в позиции (1.4) можно интерпретировать как угрозу «наказания» игрока 2, если он отклонится от желательного для игрока 1 выбора дуги 2 в позиции (2.2), лишив тем самым игрока 1 максимального выигрыша, равного 10. Однако по существу такую угрозу «наказания» едва ли следует считать действенной, поскольку наказывающий игрок при этом сам может потерять в выигрыше 5 единиц (действуя неоптимально в игре $\Gamma_{1,4}$).

4.12.2. Дадим строгое определение стратегий наказания. Для простоты ограничимся случаем неантагонистической игры двух лиц.

Пусть задана неантагонистическая игра двух лиц

$$\Gamma = \langle U_1, U_2, K_1, K_2 \rangle.$$

С игрой Γ свяжем две антагонистические игры Γ_1 и Γ_2 следующим образом. Игра Γ_1 — это антагонистическая игра, построенная на основе игры Γ , в которой игрок 2 играет против игрока 1, т. е. $K_2 = -K_1$. Игра Γ_2 — это антагонистическая игра, построенная на основе игры Γ , в которой игрок 1 играет против игрока 2, т. е. $K_1 = -K_2$. Графы игр Γ_1 , Γ_2 , Γ и множества стратегий в них совпадают. Обозначим через (u_{11}^*, u_{21}^*) и (u_{12}^*, u_{22}^*) ситуации абсолютного равновесия в играх Γ_1 и Γ_2 соответственно. Пусть Γ_{1x} , Γ_{2x} — подыгры игр Γ_1 , Γ_2 ; $v_1(x)$, $v_2(x)$ — значения этих подыгр. Тогда ситуации $\{u_{11}^{*x}, u_{21}^{*x}\}$ и $\{u_{12}^{*x}, u_{22}^{*x}\}$ являются равновесными в играх Γ_{1x} , Γ_{2x} соответственно и $v_1(x) = K_1^x(u_{11}^{*x}, u_{21}^{*x})$, $v_2(x) = K_2^x(u_{12}^{*x}, u_{22}^{*x})$.

Рассмотрим произвольную пару (u_1, u_2) стратегий в игре Γ . Разумеется, эта пара стратегий является таковой и в играх Γ_1, Γ_2 . Пусть $Z = (x_0 = z_0, z_1, \dots, z_l)$ — путь, реализуемый в ситуации (u_1, u_2) .

Определение. Стратегия $\tilde{u}_1(\cdot)$ называется стратегией наказания игрока 1, если:

$$\begin{aligned} \tilde{u}_1(z_k) &= z_{k+1} \quad \text{для} \quad z_k \in Z \cap X_1, \\ \tilde{u}_1(y) &= u_{12}^*(y) \quad \text{для} \quad y \in X_1, \quad y \notin Z. \end{aligned} \quad (4.12.1)$$

Стратегия $\tilde{u}_2(\cdot)$ называется стратегией наказания игрока 2, если:

$$\begin{aligned} \tilde{u}_2(z_k) &= z_{k+1} \quad \text{для} \quad z_k \in Z \cap X_2, \\ \tilde{u}_2(y) &= u_{21}^*(y) \quad \text{для} \quad y \in X_2, \quad y \notin Z. \end{aligned} \quad (4.12.2)$$

Из определения стратегий наказания сразу получаем следующие свойства:

1. $K_1(\tilde{u}_1(\cdot), \tilde{u}_2(\cdot)) = H_1(z_e), K_2(\tilde{u}_1(\cdot), \tilde{u}_2(\cdot)) = H_2(z_e)$.

2. Пусть один из игроков, например игрок 1, используя стратегию $u_1(\cdot)$, для которой позиция $z_k \in Z \cap X_1$ является первой в пути Z , где $u_1(\cdot)$ диктует выбор следующей позиции z'_{k+1} , отличной от выбора, диктуемого стратегией $\tilde{u}_1(\cdot)$, т. е. $z'_{k+1} \neq z_{k+1}$. Тогда из определения наказывающей стратегии $\tilde{u}_2(\cdot)$ следует, что

$$K_1(u_1(\cdot), \tilde{u}_2(\cdot)) \leq v_1(z_k). \quad (4.12.3)$$

Аналогично, если игрок 2 использует стратегию $u_2(\cdot)$, для которой позиция $z_k \in Z \cap X_2$ является первой в пути Z , где $u_2(\cdot)$ диктует выбор следующей позиции z'_{k+1} , отличной от диктуемой стратегией $\tilde{u}_2(\cdot)$, т. е. $z'_{k+1} \neq z_{k+1}$, то из определения наказывающей стратегии $\tilde{u}_1(\cdot)$ следует, что

$$K_2(\tilde{u}_1(\cdot), u_2(\cdot)) \leq v_2(z_k). \quad (4.12.4)$$

Отсюда, в частности, получаем следующую теорему.

Теорема. Пусть $(\tilde{u}_1(\cdot), \tilde{u}_2(\cdot))$ — ситуация в стратегиях наказания. Для равновесности ситуации $(\tilde{u}_1(\cdot), \tilde{u}_2(\cdot))$ достаточно, чтобы для всех $k = 0, 1, \dots, l-1$ выполнялись неравенства

$$K_1(\tilde{u}_1(\cdot), \tilde{u}_2(\cdot)) \geq v_1(z_k), \quad K_2(\tilde{u}_1(\cdot), \tilde{u}_2(\cdot)) \geq v_2(z_k), \quad (4.12.5)$$

где z_0, z_1, \dots, z_l — путь, реализовавшийся в ситуации $(\tilde{u}_1(\cdot), \tilde{u}_2(\cdot))$.

4.12.3. Пусть $u_{11}^*(\cdot)$ и $u_{22}^*(\cdot)$ — оптимальные стратегии игроков 1 и 2 во вспомогательных антагонистических играх Γ_1 и Γ_2 соответственно и $\bar{Z} = \{\bar{z}_0, \bar{z}_1, \dots, \bar{z}_l\}$ — путь, соответствующий ситуации $(u_{11}^*(\cdot), u_{22}^*(\cdot))$. Предположим, что стратегии наказания $\tilde{u}_1(\cdot)$ и $\tilde{u}_2(\cdot)$ таковы, что $\tilde{u}_1(\bar{z}_k) = u_{11}^*(\bar{z}_k)$ для $\bar{z}_k \in \bar{Z} \cap X_1$ и $\tilde{u}_2(\bar{z}_k) = u_{22}^*(\bar{z}_k)$ для $\bar{z}_k \in \bar{Z} \cap X_2$. Тогда ситуация $(\tilde{u}_1(\cdot), \tilde{u}_2(\cdot))$ образует ситуацию равновесия по Нэшу в стратегиях наказания. Для доказательства этого утверждения достаточно показать, что

$$\begin{aligned} K_1(u_{11}^*(\cdot), u_{22}^*(\cdot)) &= K_1(\tilde{u}_1(\cdot), \tilde{u}_2(\cdot)) \geq v_1(\bar{z}_k), \\ K_2(u_{11}^*(\cdot), u_{22}^*(\cdot)) &= K_2(\tilde{u}_1(\cdot), \tilde{u}_2(\cdot)) \geq v_2(\bar{z}_k), \quad k = \overline{0, l-1}, \end{aligned} \quad (4.12.6)$$

и воспользоваться теоремой из предыдущего пункта. Неравенства (4.12.6) следуют из оптимальности стратегий $u_{11}^*(\cdot)$ и $u_{22}^*(\cdot)$ в играх Γ_1 и Γ_2 соответственно, их обоснование предлагаем в качестве упражнения. Таким образом, получена следующая теорема.

Теорема. В игре Γ всегда существует ситуация равновесия в стратегиях наказания, в которой выигрыши равны $K_i(u_{11}^*(\cdot), u_{22}^*(\cdot))$, где $u_{11}^*(\cdot)$ и $u_{22}^*(\cdot)$ — оптимальные стратегии игроков 1 и 2 во вспомогательных антагонистических играх Γ_1 и Γ_2 соответственно.

4.12.4. Смысл стратегий наказания заключается в том, что игрок заставляет партнера придерживаться определенного пути в игре (определенных выборов), используя постоянную угрозу переключения на стратегию, оптимальную в антагонистической игре против партнера. Множество ситуаций равновесия в классе стратегий наказания достаточно представительно, однако эти стратегии не следует считать очень «хорошими», поскольку, наказывая партнера, игрок может еще сильнее наказать самого себя.

Пример 14. Рассмотрим игру N лиц, изображенную на рис. 4.16. В этой игре игроки ходят один за другим по одному разу, имея возможность в каждой позиции выбрать одну из двух альтернатив: A или D . Выигрыши игроков записаны в окончательных позициях. Легко убедиться, действуя по индукции с конца игры, что ситуация $u_i^* = A, i = 1, \dots, N, u^* = (A, A, \dots, A)$ является ситуацией равновесия по Нэшу и абсолютно равновесной ситуацией с выигрышами $(2, 2, \dots, 2)$.

Действительно, пусть игрок i выбирает $u_i = D$. Тогда в ситуации $(u^*|u_i) = (u^*|D) = (A, A, \dots, A, D, A, \dots, A)$ выигрыши всех игроков равны соответственно $(1/i, 1/i, \dots, 1/i)$, т. е.

$$2 = K_i(u^*) > K_i(u^*|u_i = D) = \frac{1}{i},$$

и u^* есть равновесие по Нэшу. Очевидно, что это же рассуждение можно провести для любой подыгры, начиная с шага k .

В то же время эта ситуация не является устойчивой в том смысле, что при большом числе игроков нельзя быть уверенным (первым игрокам), что какой-то из игроков не «ошибется» и вместо A выберет D . Тогда все игроки (не только тот, который «ошибся») потеряют в выигрыше.

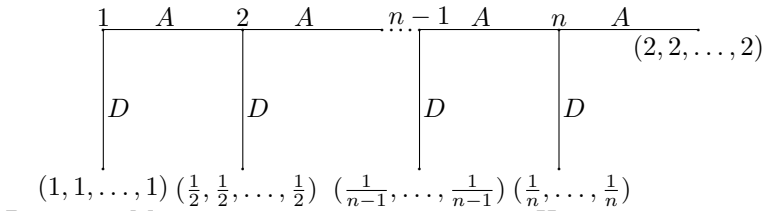


Рис. 4.16. Множественность равновесий по Нэшу в простейшей игре с двумя альтернативами у каждого игрока

В игре имеется богатое множество ситуаций равновесия в стратегиях наказания. Ситуация, в которой первый и любой другой игрок выбирает D , является ситуацией равновесия по Нэшу, т. е. равновесной оказываются ситуации вида

$$\bar{u} = (D, A, \dots, D, \dots, A).$$

(с D на первом месте и еще одним D на любом другом). Выигрыши во всех таких ситуациях одни и те же и равны $(1, 1, \dots, 1)$.

Действительно, пусть в ситуации \bar{u} второй игрок, выбирающий D , имеет номер $k > 1$, если игрок $i \geq 2$ выбирает u_i , отличную от той, которая входит в ситуацию \bar{u} , то

выигрыши игроков не меняются, так как выбор игроком 1 на первом шаге альтернативы D гарантирует завершение игры на этом шаге, при котором все игроки получают выигрыш 1. Если игрок $i = 1$ выбирает вместо D альтернативу A , то благодаря наличию в ситуации \bar{u} еще одного игрока, выбирающего D (игрока с номером k), выигрыш игрока 1 уменьшится и станет равным $1/k$.

Выигрыши в ситуации \bar{u} , конечно, меньше выигрышей в ситуации $u^* = (A, A, \dots, A)$, но не зависят от ошибок большого числа игроков.

§ 4.13. Кооперация в многошаговых играх

4.13.1. Рассмотрим игру Γ с полной информацией на конечном графе, как определено в § 4.1 и в обозначениях этого параграфа. Здесь будет незначительное различие в определении выигрышей игроков. В § 4.1 предполагалось, что выигрыши определены только на множестве X_{n+1} окончательных позиций игры. Здесь будем предполагать, что для каждого $x \in X$ определены n действительных чисел $h_i(x)$, $i = 1, \dots, n$, и для каждого пути игры $z = (z_0, z_1, \dots, z_l)$, $z_l \in X_{n+1}$ выигрыш i -го игрока определяется как

$$H_i(z_0) = \sum_{k=0}^l h_i(z_k), \quad h_i \geq 0.$$

Если $h_i(x) = 0$, $x \in X_i$, $i = 1, \dots, n$, то мы имеем в точности игру, определенную в § 4.1. Как принято в теории классических кооперативных игр, предполагаем, что перед началом игры игроки соглашаются выбрать такой n -мерный вектор стратегий

$$\bar{u}(\cdot) = (\bar{u}_1(\cdot), \dots, \bar{u}_i(\cdot), \dots, \bar{u}_n(\cdot)),$$

который максимизирует сумму выигрышей игроков (см. § 3.11). Если $\bar{z} = (\bar{z}_0, \dots, \bar{z}_k, \dots, \bar{z}_l)$, $z_l \in X_{n+1}$ — путь (траектория), реализованный ситуацией $\bar{u}(\cdot) = (\bar{u}_1, \dots, \bar{u}_i, \dots, \bar{u}_n)$, то по определению ситуации $\bar{u}(\cdot)$ имеем

$$\max_{z_0, \dots, z_i, \dots, z_l} \sum_{i=1}^n \sum_{k=0}^l h_i(z_k) = \sum_{i=1}^n \sum_{k=0}^l h_i(\bar{z}_k). \quad (4.13.1)$$

Кооперативная игра Γ развивается вдоль траектории $\bar{z} = (\bar{z}_0, \dots, \bar{z}_k, \dots, \bar{z}_l)$, которую мы будем называть *оптимальной траекторией*.

Ясно, что в игре Γ мы можем иметь целое семейство «оптимальных траекторий», каждая из которых дает одинаковый максимальный общий выигрыш. В этом параграфе для простоты предполагаем также, что в игре Γ оптимальная траектория единственна. Определим в игре Γ характеристическую функцию (см. п. 3.11.1).

4.13.2. Характеристическая функция может быть введена аксиоматически или как значение игры с нулевой суммой, происходящей между коалициями $S \subset N$ и $N \setminus S$.

Как известно, важным является выполнение условия

$$V(N) = \sum_{i=1}^n \sum_{k=0}^l h_i(\bar{z}_k), \quad N = \{1, \dots, n\}, \quad (4.13.2)$$

и для $S_1 \subset N$, $S_2 \subset N$, $S_1 \cap S_2 = \emptyset$

$$V(S_1 \cup S_2) \geq V(S_1) + V(S_2), \quad V(\emptyset) = 0.$$

Если характеристическая функция определена, то мы можем определить множество дележей

$$C = \{\xi = (\xi_i) : \sum_{i=1}^n \xi_i = V(N), \quad \xi_i \geq V(\{i\}), \quad i = 1, \dots, n\},$$

C -ядро

$$M = \{\xi = (\xi_i) : \sum_{i \in S} \xi_i \geq V(S), \quad S \subset N\} \subset C,$$

вектор Шепли и другие принципы оптимальности классической теории кооперативных игр. В дальнейшем будем обозначать через $M \subset C$ любой из этих *принципов оптимальности*.

4.13.3. Предположим, что в начале игры игроки соглашаются использовать принцип оптимальности $M \subset C$ как основу для выбора «оптимального» дележа $\bar{\xi} \in M$. Это означает, что, играя кооперативно, выбирая стратегии, максимизирующие общий выигрыш, игроки ожидают, что каждый из них получит доход $\bar{\xi}_i$ из оптимального дележа $\bar{\xi} \in M$ после окончания игры (после того как максимальный общий выигрыш $V(N)$ действительно будет получен всеми игроками).

Но когда игра Γ действительно развивается вдоль оптимальной траектории $\bar{z} = (\bar{z}_0, \bar{z}_1, \dots, \bar{z}_k, \dots, \bar{z}_l)$, в каждой вершине \bar{z}_k игроки находятся в новой многошаговой игре с полной информацией $\Gamma_{\bar{z}_k}$, $k = 0, \dots, l$, которая является подыгрой исходной игры Γ , начинающейся из вершины \bar{z}_k с выигрышами

$$H_i(\bar{z}_k) = \sum_{j=k}^l h_i(\bar{z}_j), \quad h_i \geq 0, \quad i = 1, \dots, n.$$

Важно заметить, что для выражения (4.13.1) принцип оптимальности Беллмана выполняется и часть $\bar{z}^k = (\bar{z}_k, \dots, \bar{z}_j, \dots, \bar{z}_l)$ траектории \bar{z} , начинающейся с \bar{z}_k , максимизирует сумму выигрышей в подыгре $\Gamma_{\bar{z}_k}$, т. е.

$$\max_{z_k, \dots, z_j, \dots, z_l} \sum_{i=1}^n \sum_{j=k}^l h_i(z_j) = \sum_{i=1}^n \sum_{j=k}^l h_i(\bar{z}_j). \quad (4.13.3)$$

Это означает, что траектория $\bar{z}^k = (\bar{z}_k, \dots, \bar{z}_j, \dots, \bar{z}_l)$ также «оптимальная» в подыгре $\Gamma_{\bar{z}_k}$.

Перед попаданием в подыгру $\Gamma_{\bar{z}_k}$ каждый из игроков i уже заработал выигрыш

$$H_i^{\bar{z}^k} = \sum_{j=0}^{k-1} h_i(\bar{z}_j). \quad (4.13.4)$$

В то же время в начале игры $\Gamma = \Gamma(x_0) = \Gamma(\bar{z}_0)$ игрок i стремился получить доход $\bar{\xi}_i$ — i -ю компоненту «оптимального» дележа $\bar{\xi} \in M \subset C$. Отсюда следует, что в подыгре $\Gamma_{\bar{z}_k}$ он ожидает получить доход, равный

$$\bar{\xi}_i - H_i^{\bar{z}_k} = \bar{\xi}_i^{\bar{z}_k}, \quad i = 1, \dots, n, \quad (4.13.5)$$

и тогда возникает вопрос, останется ли новый вектор $\bar{\xi}^{\bar{z}_k} = (\bar{\xi}_1^{\bar{z}_k}, \dots, \bar{\xi}_i^{\bar{z}_k}, \dots, \bar{\xi}_n^{\bar{z}_k})$ оптимальным в том же смысле в подыгре $\Gamma_{\bar{z}_k}$, как вектор $\bar{\xi}$ был в игре $\Gamma(\bar{z}_0)$. Если нет, это будет означать, что игроки в подыгре $\Gamma_{\bar{z}_k}$ не будут ориентироваться на тот же принцип оптимальности, что и в игре $\Gamma(\bar{z}_0)$, что может побудить их выйти из кооперации изменением выбора кооперативных стратегий $\bar{u}_i(\cdot)$, $i = 1, \dots, n$, и, следовательно, изменением оптимальной траектории \bar{z}_k в подыгре $\Gamma_{\bar{z}_k}$. Постараемся теперь обосновать справедливость этих доводов.

4.13.4. Введем в подыгре $\Gamma_{\bar{z}_k}$, $k = 1, \dots, l$, характеристическую функцию $V(S; \bar{z}_k)$, $S \subset N$, таким же образом, каким она была введена в игре $\Gamma = \Gamma(z_0)$. Основываясь на характеристической функции $V(S; \bar{z}_k)$, можно ввести множество дележей

$$C(\bar{z}_k) = \{\xi = (\xi_i) : \sum_{i=1}^n \xi_i = V(N; \bar{z}_k), \xi_i \geq V(\{i\}; \bar{z}_k), i = 1, \dots, n\},$$

ядро

$$M = \{\xi = (\xi_i) : \sum_{i \in S} \xi_i \geq V(S; \bar{z}_k), S \subset N\} \subset C(\bar{z}_k),$$

вектор Шепли и другие принципы оптимальности классической теории игр. Обозначим через $M(\bar{z}_k) \subset C(\bar{z}_k)$ принцип оптимальности $M \subset C$ (который был выбран игроками в игре $\Gamma(z_0)$), рассматриваемый в подыгре $\Gamma_{\bar{z}_k}$.

Если мы предположим, что игроки в игре $\Gamma(z_0)$, двигаясь вдоль оптимальной траектории $(\bar{z}_0, \dots, \bar{z}_k, \dots, \bar{z}_l)$, следуют одинаковой идеологии оптимального поведения, то вектор $\bar{\xi}^{\bar{z}_k} = \bar{\xi} - H^{\bar{z}_k}$ должен принадлежать множеству $M(\bar{z}_k)$ — соответствующему принципу оптимальности в кооперативной игре $\Gamma_{\bar{z}_k}$, $k = 0, \dots, l$.

Ясно видно, что очень трудно найти игры и соответствующие принципы оптимальности, для которых это условие выполняется. Постараемся проиллюстрировать это на следующем примере.

Предположим, что в игре Γ $h_i(z) \neq 0$ только для $z \in X_{n+1}$ (игра Γ — игра с терминальным выигрышем из §4.1). Тогда последнее условие дает $\bar{\xi} = \bar{\xi}^{\bar{z}_k} \in M(\bar{z}_k)$, $k = 0, \dots, l$, откуда получается

$$\bar{\xi} \in \bigcap_{k=0}^l M(\bar{z}_k). \quad (4.13.6)$$

Для $k = l$ имеем $\bar{\xi} \in M(\bar{z}_l)$. Но $M(\bar{z}_l) = C(\bar{z}_l) = \{h_i(\bar{z}_l)\}$, и это условие должно иметь место для всех дележей множества $M(\bar{z}_0)$ и для всех принципов оптимальности $M(z_0) \subset C(z_0)$, это означает, что в кооперативной игре с терминальным выигрышем единственно разумным принципом оптимальности будет $\bar{\xi} = \{h_i(\bar{z}_l)\}$ — вектор выигрыша, получаемый в конечной точке кооперативной траектории в игре $\Gamma(z_0)$. В то же время простейшие примеры показывают, что включение (4.13.78), кроме случая «блванов», не имеет место для игры с терминальными выигрышами.

4.13.5. Как же преодолеть эту трудность? Правдоподобным способом является введение специального правила выплат (пошаговая плата) на каждом шаге игры таким образом, чтобы выплаты на каждом шаге не превышали общей суммы, заработанной

игроками на этом шаге, и выплаты, получаемые игроками, начиная с шага k (в подыгре $\Gamma_{\bar{z}_k}$), принадлежали бы тому же принципу оптимальности, что и дележ $\bar{\xi}$, на который игроки согласились в игре Γ_{z_0} в начале игры. Является это возможным или нет, об этом и пойдет речь ниже.

Введем понятие процедуры распределения дележа.

Определение. Предположим, что $\xi = \{\xi_1, \dots, \xi_i, \dots, \xi_n\} \in M(z_0)$. Всякая матрица $\beta = \{\beta_{ik}\}$, $i = 1, \dots, n$, $k = 0, \dots, l$, такая, что

$$\xi_i = \sum_{k=0}^l \beta_{ik}, \quad \beta_{ik} \geq 0, \quad (4.13.7)$$

называется процедурой распределения дележа (ПРД).

Обозначим $\beta_k = (\beta_{1k}, \dots, \beta_{nk})$, $\beta(k) = \sum_{m=0}^{k-1} \beta_m$. Интерпретация ПРД β следующая: β_{ik} — выплата игроку i на шаге k игры Γ_{z_0} , т. е. на первом шаге подыгры $\Gamma_{\bar{z}_k}$. Из определения (4.13.7) следует, что в игре Γ_{z_0} каждый игрок i получает доход $\bar{\xi}_i$, $i = 1, \dots, n$, который он ожидает получить как i -ю компоненту оптимального дележа $\xi_i \in M(z_0)$ в игре Γ_{z_0} .

Интерпретация $\beta_i(k)$ следующая: $\beta_i(k)$ — сумма, получаемая игроком i на первых k шагах игры Γ_{z_0} .

4.13.6. Определение. Принцип оптимальности $M(z_0)$ называется динамически устойчивым, если для каждого $\xi \in M(z_0)$ существует ПРД β , такая, что

$$\xi^k = \xi - \beta(k) \in M(z_k), \quad k = 1, \dots, l. \quad (4.13.8)$$

Определение. Принцип оптимальности $M(z_0)$ называется сильно динамически устойчивым, если для каждого $\xi \in M(z_0)$ существует ПРД β , такая, что $\beta(k) \oplus M(\bar{z}_k) \subset M(z_0)$, $k = 1, \dots, l$, где $a \oplus A = \{a + a' : a' \in A, a \in R^n, A \subset R^n\}$.

Динамическая устойчивость принципа оптимальности $M(z_0)$ подразумевает, что для каждого дележа $\xi \in M$ существует такая ПРД β , что если выплаты в каждой позиции \bar{z}_k на оптимальной траектории \bar{z} будут сделаны игрокам в соответствии с ПРД β , то в каждой подыгре $\Gamma_{\bar{z}_k}$ игроки могут ожидать выплат $\bar{\xi}^k$, которые являются оптимальными в подыгре $\Gamma_{\bar{z}_k}$ в том же смысле, в каком они были оптимальными в исходной игре Γ_{z_0} .

Сильная динамическая устойчивость означает, что если выплаты сделаны в соответствии с ПРД β , то, заработав на первых k шагах сумму $\beta(k)$, игроки (если они ориентировались в подыгре $\Gamma_{\bar{z}_k}$ на тот же принцип оптимальности, что и в Γ_{z_0}), пересматривая оптимальный дележ в этой подыгре (заменяя один оптимальный дележ другим), все равно получают в результате в игре Γ_{z_0} выплаты в соответствии с некоторым дележом, оптимальным в предыдущем смысле, т. е. дележом, принадлежащим множеству $M(z_0)$.

4.13.7. Если мы опустим условие неотрицательности $\beta_{ik} \geq 0$, налагаемое на компоненты ПРД β , то для любого принципа оптимальности $M(z_0) \subset C(z_0)$ и для каждого $\bar{\xi} \in M(z_0)$ можно определить β_{ik} по следующим формулам:

$$\bar{\xi}_i^{\bar{z}_k} - \bar{\xi}_i^{\bar{z}_{k+1}} = \beta_{ik}, \quad \bar{\xi}_i^{\bar{z}_l} = \beta_{il}, \quad i = 1, \dots, n, \quad k = 0, \dots, l-1. \quad (4.13.9)$$

Из определения следует, что

$$\sum_{k=0}^l \beta_{ik} = \sum_{k=0}^{l-1} (\bar{\xi}_i^{\bar{z}_k} - \bar{\xi}_i^{\bar{z}_{k+1}}) + \bar{\xi}_i^{\bar{z}_l} = \bar{\xi}_i^{\bar{z}_0} = \bar{\xi}_i.$$

В то же время $\bar{\xi} - \beta(k) = \xi^{\bar{z}_k} \in M(\bar{z}_k)$, $k = 0, \dots, l$. Последнее включение может означать динамическую устойчивость $M(z_0)$, если мы будем уверены, что

$$\beta_{ik} = \bar{\xi}_i^{\bar{z}_k} - \bar{\xi}_i^{\bar{z}_{k+1}} \geq 0, \quad i = 1, \dots, n, \quad k = 0, \dots, l. \quad (4.13.10)$$

К сожалению, последнее неравенство не может быть гарантировано даже в простейших случаях. Можно убедиться в этом, рассматривая игры с терминальными выигрышами. Для таких игр условие (4.13.10) практически никогда не выполняется. Если же рассмотреть свойство сильной динамической устойчивости, то оно более сильное. Для него мы не можем даже вывести формулу типа (4.13.10).

4.13.8. Проведем регуляризацию (улучшение) классических принципов оптимальности, которая приведет нас к сильной динамической устойчивости. Более того, регуляризация порождает динамическую устойчивость новых принципов оптимальности.

Рассмотрим следующие функции:

$$\begin{aligned} \beta_i^0 &= \frac{\xi_i^0 \sum_{i=1}^n h_i(\bar{z}_0)}{V(N; z_0)}, \quad \xi^0 \in C(z_0); \\ \beta_i^1 &= \frac{\xi_i^1 \sum_{i=1}^n h_i(\bar{z}_1)}{V(N, \bar{z}_1)}, \quad \xi^1 \in C(\bar{z}_1); \\ &\dots \\ \beta_i^k &= \frac{\xi_i^k \sum_{i=1}^n h_i(\bar{z}_k)}{V(N, \bar{z}_k)}, \quad \xi^k \in C(\bar{z}_k); \\ &\dots \\ \beta_i^l &= \frac{\xi_i^l \sum_{i=1}^n h_i(\bar{z}_l)}{V(N, \bar{z}_l)}, \quad \xi^l \in C(\bar{z}_l). \end{aligned} \quad (4.13.11)$$

Определим ПРД $\bar{\beta}^k = \{\beta_i^k, i = 1, \dots, n\}$, $k = 0, \dots, l$. Легко видеть, что $\bar{\beta}^k \geq 0$. Рассмотрим формулу (4.13.11). Для различных дележей $\xi^k \in C(\bar{z}_k)$ получаем разные значения β_i^k и, следовательно, разные $\bar{\beta}$. Пусть B^k — множество всевозможных $\bar{\beta}^k$ для всех $\xi^k \in C(\bar{z}_k)$, $k = 1, \dots, l$.

Рассмотрим множества $\tilde{C}(z_0) = \{\bar{\xi} : \bar{\xi} = \sum_{k=0}^l \bar{\beta}^k, \bar{\beta}^k \in B^k\}$ и $\tilde{C}(\bar{z}_k) = \{\bar{\xi}^k : \bar{\xi}^k = \sum_{m=k}^l \bar{\beta}^m, \bar{\beta}^m \in B^m\}$.

Множество $\tilde{C}(z_0)$ называется *регуляризованным принципом оптимальности* (ПО) $\bar{C}(z_0)$ и, соответственно, $\tilde{C}(\bar{z}_k)$ — регуляризованный ПО $\bar{C}(\bar{z}_k)$.

Будем считать $\tilde{C}(z_0)$ новым принципом оптимальности в игре $\Gamma(z_0)$.

Теорема. Если ПРД β определена как $\bar{\beta}$, $k = 1, \dots, l$, то всегда выполняется $\beta(k) \oplus \tilde{C}(\bar{z}_k) \subset \tilde{C}(z_0)$, т. е. ПО $\tilde{C}(z_0)$ — сильно динамически устойчивый $\left(\beta(k) = \sum_{m=0}^{k-1} \bar{\beta}^m \right)$. Здесь множество $\beta(k) \oplus \tilde{C}(\bar{z}_k)$ есть множество всех векторов $\beta(k) + \bar{\xi}^k$, где $\bar{\xi}^k \in \tilde{C}(\bar{z}_k)$.

Доказательство. Пусть $\bar{\xi} \in \beta(k) \oplus \tilde{C}(\bar{z}_k)$, тогда $\bar{\xi} = \bar{\xi}^{\bar{z}_k} + \sum_{m=0}^{k-1} \bar{\beta}^m$ для некоторых $\beta^m \in B^m$, $m = k, \dots, l$.

Но $\bar{\xi}^{\bar{z}_k} = \sum_{m=0}^{k-1} \beta'^m$ для некоторых $\beta'^m \in B^m$, $m = 0, \dots, k-1$.

Рассмотрим

$$(\beta'')^m = \begin{cases} \beta'^m, & m = 0, \dots, k-1, \\ \beta^m, & m = k, \dots, l, \end{cases}$$

тогда $(\beta'')^m \in B^m$, $\bar{\xi} = \sum_{m=0}^l (\beta'')^m$ и, следовательно, $\bar{\xi} \in \tilde{C}(z_0)$. Теорема доказана.

Мы имеем также:

$$\sum_{i=1}^n \beta_i^k = \sum_{i=1}^n h_i(\bar{z}_k), \quad k = 0, \dots, l,$$

и, следовательно, при $0 \leq \Theta \leq n$

$$\sum_{k=0}^{\Theta} \beta_i^k = \sum_{i=1}^n \sum_{k=0}^{\Theta} h_i(\bar{z}_k), \quad (4.13.12)$$

что является действительным выигрышем, который можно распределить между игроками на первых $\Theta + 1$ шагах и, как это видно из (4.13.12), в точности равен сумме, заработанной ими на этих шагах.

4.13.9. Пример 15. Рассмотрим игру трех лиц Γ , изображенную на рис. 4.17.

В данном пункте исследуем динамическую устойчивость одного классического принципа оптимальности вектора Шепли. Не вдаваясь в подробную интерпретацию этого принципа оптимальности, отметим только, что это дележ, определяемый по формуле

$$\text{Sh}_i = \xi_i = \sum_{S: S \subset N, \{i\} \in S} \frac{(s-1)!(n-s)!}{n!} [V(S) - V(S \setminus \{i\})],$$

где s — число элементов множества S .

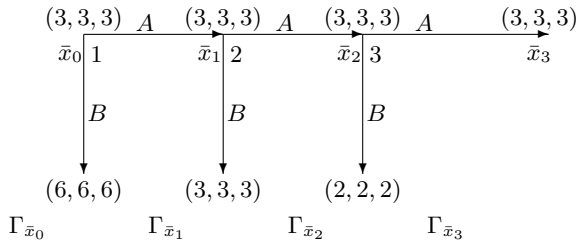


Рис. 4.17. Простейшая игра, в которой выигрыши накапливаются в процессе игры

Здесь в скобках записаны выигрыши игроков 1, 2, 3 в каждом из узлов графа, $N = \{1, 2, 3\}$. Множества стратегий игроков $X_1 = X_2 = X_3 = \{A, B\}$ состоят из двух элементов A и B .

Выигрыши игроков равны $H_1(A, A, A) = H_2(A, A, A) = H_3(A, A, A) = (12, 12, 12)$; $H_1(B, x_2, x_3) = H_2(B, x_2, x_3) = H_3(B, x_2, x_3) = (9, 9, 9)$ при любых $x_2 \in X_2, x_3 \in X_3$; $H_1(A, B, x_3) = H_2(A, B, x_3) = H_3(A, B, x_3) = (9, 9, 9)$ при любых $x_3 \in X_3$; $H_1(A, A, B) = H_2(A, A, B) = H_3(A, A, B) = (11, 11, 11)$.

Правила вычисления выигрыша можно проследить, используя рис. 4.18. Если реализована последовательность A, A, A , то выигрыш игрока равен сумме его выигрышей в каждой вершине графа; если реализованный путь включает выбор B одним из игроков, то выигрыш каждого игрока равен сумме выигрышей вдоль пути, заканчивающегося первым выбором стратегии B .

Вычислим значения характеристической функции V . Здесь возможны следующие коалиции: $\{1, 2, 3\}$, $\{1, 2\}$, $\{1, 3\}$, $\{2, 3\}$, $\{1\}$, $\{2\}$, $\{3\}$;

$V(\{1, 2, 3\}) = 36$, так как, выбирая A , каждый игрок получает 12 (а вся коалиция $12 \times 3 = 36$).

$V(\{1, 2\}) = 22$, так как игроки 1, 2, выбирая A , могут обеспечить себе в худшем случае 22. Худший случай имеет место, если игрок 3 выберет B .

$V(\{1, 3\}) = 18$, так как игрок 1, выбирая A , может обеспечить себе и игроку 3 в худшем случае выигрыш 18. Худший случай имеет место, если игрок 2 выберет B .

$V(\{2, 3\}) = 18$, так как игрок 1, выбирая B , всегда может сделать так, что выигрыш игроков $\{2, 3\}$ будет не превосходить 18. При другом выборе игрока 1 игроки $\{2, 3\}$, очевидно, получают больший выигрыш.

$V(\{1\}) = 9$, этот выигрыш достигается выбором B игроком 1 и выбором B игроком 2. Очевидно, что выбором B игрок 2 ограничивает выигрыш игрока 1 числом 3.

$V(\{2\}) = 9$, здесь игрок 1, выбирая B , ограничивает выигрыш игрока 2 числом 9. При другом выборе игрока 1 выигрыш игрока 2 может быть больше.

$V(\{3\}) = 9$, этот случай подобен предыдущему.

Вычислим вектор Шепли: $Sh_1 = 76/6$, $Sh_2 = 76/6$, $Sh_3 = 64/6$, т. е. вектор Шепли предписывает одинаковые выигрыши игрокам 1 и 2.

4.13.10. Пример 16. Найдем характеристическую функцию для примера, изображенного на рис. 4.18 (здесь выигрыши игроков заданы в окончательных позициях).

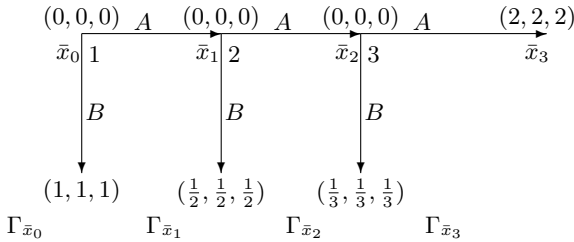


Рис. 4.18. Кооперативный вариант простейшей игры с двумя альтернативами у каждого игрока

Характеристическая функция имеет вид

$$V(\{1, 2, 3\}) = 6, \quad V(\{1, 2\}) = 2, \quad V(\{1, 3\}) = 2, \quad V(\{2, 3\}) = 2,$$

$$V(\{1\}) = 1, \quad V(\{2\}) = \frac{1}{2}, \quad V(\{3\}) = \frac{1}{2}.$$

Для вектора Шепли получаем:

$$\text{Sh}_1 = 26/12, \quad \text{Sh}_2 = 23/12, \quad \text{Sh}_3 = 23/12.$$

Данный пример используем для иллюстрации последующих результатов.

Рассмотрим позиции игры Γ вдоль оптимальной траектории \bar{x} (траектории, максимизирующей суммарный выигрыш игроков), т. е. подыгры $\Gamma = \Gamma_{\bar{x}_1}, \Gamma_{\bar{x}_2}, \dots, \Gamma_{\bar{x}_l}$. Очевидно, что отрезок траектории \bar{x} вида $\bar{x}^k = (\bar{x}_k, \dots, \bar{x}_i, \dots, \bar{x}_l)$, рассмотренный в подыгре $\Gamma_{\bar{x}_k}$, является оптимальной траекторией в этой подыгре (принцип оптимальности Беллмана). Обозначим через $V(S, k)$, $S \subset N$, характеристическую функцию в подыгре $\Gamma_{\bar{x}_k}$. В частности, $V(S, 0) = V(S)$, $S \subset N$, есть характеристическая функция игры Γ . Зная характеристические функции подыгры, можно построить вектор Шепли для подыгры $\Gamma_{\bar{x}_k}$. Обозначим его через $\text{Sh}(k) = \{\text{Sh}_i(k), i = 1, \dots, n\}$. Предположим, что в кооперативной игре $\Gamma_{x_0} = \Gamma$ в качестве принципа оптимальности выбран вектор Шепли. Это означает, что игроки, договорившись о выборе набора стратегий $\bar{u} = (\bar{u}_1, \dots, \bar{u}_i, \dots, \bar{u}_n)$, гарантирующего максимальный суммарный выигрыш игроков, рассчитывают получить выигрыши, определяемые вектором Шепли для игры Γ_{x_0} . Именно это и служит основой для их кооперации.

В ситуации \bar{u} игра развивается вдоль оптимальной траектории $\bar{x} = (x_0, \bar{x}_1, \dots, \bar{x}_l)$. После первого шага игра переходит в вершину \bar{x}_1 , и фактически игроки играют в новую игру $\Gamma_{\bar{x}_1}$, которая является подыгрой игры Γ_{x_0} . Вектор Шепли в этой подыгре $\text{Sh}(1)$ отличается, вообще говоря, от вектора Шепли $\text{Sh}(0) = \text{Sh}$ в игре Γ_{x_0} . Поэтому, если мы желаем произвести выплаты в Γ_{x_0} согласно $\text{Sh}(0)$, что являлось основой для кооперации в Γ_{x_0} , то в $\Gamma_{\bar{x}_1}$ для сохранения кооперации игроки должны ожидать получение выигрышей в соответствии с $\text{Sh}(1)$ — вектором Шепли, рассчитанным для $\Gamma_{\bar{x}_1}$. Вопрос заключается в том, можно ли, осуществляя выплаты на каждом шаге игры, добиться того, чтобы оставшиеся выплаты представляли собой компоненты вектора Шепли для подыгры, начинающейся с данного шага.

Предыдущие определения ПРД п. 4.13.5 и динамической устойчивости п. 4.13.6 для вектора Шепли примут следующий вид (заметим, что из единственности вектора Шепли следует, что понятие динамической устойчивости и сильной динамической устойчивости совпадают).

Определение. Вектор $\beta^i = (\beta_0^i, \beta_1^i, \dots, \beta_{l-1}^i, \beta_l^i)$, $i = 1, \dots, n$, называется *процедурой распределения вектора Шепли (ПДРШ)*, если

$$1. \quad \sum_{k=0}^n \beta_k^i = \text{Sh}_i(0) = \text{Sh}_i, \quad (4.13.13)$$

$$2. \quad \beta^i(k) + \text{Sh}_i(k) = \text{Sh}_i(0), \quad (4.13.14)$$

$$\text{где } \beta^i(k) = \sum_{m=0}^{k-1} \beta_m^i.$$

Определение. Вектор Шепли в игре Γ называется *динамически устойчивым (состоятельным во времени)*, если существует неотрицательная процедура распределения вектора Шепли.

Таким образом, если вектор Шепли динамически устойчив, то осуществляя на каждом шаге траектории выплаты игрокам в соответствии с ПДРШ (т. е. выплачивая на шаге k игроку i сумму β_k^i), можно добиться того, чтобы вектор Шепли для подыгры $\Gamma_{\bar{x}_k}$ как раз соответствовал выигрышам, которые игрокам осталось получить в подыгре

$\Gamma_{\bar{x}_k}$. Если не требовать неотрицательности β_k^i , $i = 1, \dots, n$, $k = 1, \dots, l$, то соотношение (4.13.8) всегда может быть выполнено, однако отрицательные β_k^i не имеют экономического смысла, так как игроки едва ли согласятся отдавать средства во имя кооперации.

4.13.11. Пример 17. Рассмотрим игру Γ из примера 16 как многошаговую игру с полной информацией. Здесь $V(N) = 36$, $\bar{u} = (A, A, A)$ и оптимальная траектория имеет вид $(x_0, \bar{x}_1, \bar{x}_2, \bar{x}_3) = \bar{x}$. Характеристическая функция для игры $\Gamma = \Gamma_{x_0}$ была найдена в примере 16:

$$V(\{1, 2, 3\}, 0) = 36, \quad V(\{1, 2\}, 0) = 22, \quad V(\{1, 3\}, 0) = 18,$$

$$V(\{2, 3\}, 0) = 18, \quad V(\{1\}, 0) = 9, \quad V(\{2\}, 0) = 9, \quad V(\{3\}, 0) = 9.$$

Вектор Шепли имеет вид $\text{Sh}(0) = \{76/6, 76/6, 64/6\}$. Характеристические функции для игр $\Gamma_{\bar{x}_1}$, $\Gamma_{\bar{x}_2}$, $\Gamma_{\bar{x}_3}$ вычисляются аналогично и равны соответственно

$$V(\{1, 2, 3\}, 1) = 27, \quad V(\{1, 2\}, 1) = 16, \quad V(\{1, 3\}, 1) = 12,$$

$$V(\{2, 3\}, 1) = 18, \quad V(\{1\}, 1) = 6, \quad V(\{2\}, 1) = 8, \quad V(\{3\}, 1) = 6.$$

$$V(\{1, 2, 3\}, 2) = 18, \quad V(\{1, 2\}, 2) = 10, \quad V(\{1, 3\}, 2) = 12,$$

$$V(\{2, 3\}, 2) = 12, \quad V(\{1\}, 2) = 5, \quad V(\{2\}, 2) = 5, \quad V(\{3\}, 2) = 6;$$

$$\text{Sh}(1) = \left\{ \frac{22}{3}, \frac{34}{3}, \frac{25}{3} \right\}, \quad \text{Sh}(2) = \left\{ 5\frac{1}{2}, 5\frac{1}{2}, 7 \right\}, \quad \text{Sh}(3) = \{3, 3, 3\}$$

$$\text{Sh}(0) = \left(\frac{16}{3}, \frac{4}{3}, \frac{7}{3} \right) + \text{Sh}(1), \quad \text{Sh}(1) = \left(\frac{11}{6}, \frac{32}{6}, \frac{8}{6} \right) + \text{Sh}(2),$$

$$\text{Sh}(2) = \left(2\frac{1}{2}, 2\frac{1}{2}, 4 \right) + \text{Sh}(3), \quad \text{Sh}(3) = (3, 3, 3),$$

т. е. $\beta_k^i > 0$, $i = 1, 2, 3$, $k = 1, 2, 3$, и вектор Шепли в Γ является динамически устойчивым. Следующий пример показывает, что это далеко не всегда имеет место.

Пример 18. Рассмотрим игру из примера 16 как многошаговую игру с полной информацией. Здесь $V(N) = 6$, $\bar{u} = (A, A, A)$ и оптимальная траектория, как и в предыдущем случае, имеет вид $(x_0, \bar{x}_1, \bar{x}_2, \bar{x}_3) = \bar{x}$. Характеристическая функция для игры $\Gamma = \Gamma_{\bar{x}_0}$ была найдена в примере 8:

$$V(\{1, 2, 3\}, 0) = 6, \quad V(\{1, 2\}, 0) = 2, \quad V(\{1, 3\}, 0) = 2, \quad V(\{2, 3\}, 0) = 2,$$

$$V(\{1\}, 0) = 1, \quad V(\{2\}, 0) = 1/2, \quad V(\{3\}, 0) = 1/2.$$

Вектор Шепли имеет вид $\text{Sh}(0) = (26/12, 23/12, 23/12)$. Характеристические функции для игр $\Gamma_{\bar{x}_1}$, $\Gamma_{\bar{x}_2}$, $\Gamma_{\bar{x}_3}$ вычисляются аналогично и равны соответственно

$$V(\{1, 2, 3\}, 1) = 6, \quad V(\{1, 2\}, 1) = 1, \quad V(\{1, 3\}, 1) = 1,$$

$$V(\{2, 3\}, 1) = 4, \quad V(\{1\}, 1) = 1/3, \quad V(\{2\}, 1) = 1/2, \quad V(\{3\}, 1) = 1/2.$$

$$\text{Sh}(1) = \left\{ 34/36, 91/36, 91/36 \right\},$$

$$V(\{1, 2, 3\}, 2) = 6, \quad V(\{1, 2\}, 2) = \frac{2}{3}, \quad V(\{1, 3\}, 2) = 4,$$

$$V(\{2, 3\}, 2) = 4, \quad V(\{1\}, 2) = \frac{1}{3}, \quad V(\{2\}, 2) = \frac{1}{3}, \quad V(\{3\}, 2) = 2,$$

$$\text{Sh}(2) = \left\{ \frac{21}{18}, \frac{21}{18}, \frac{66}{18} \right\}, \quad \text{Sh}(3) = \{2, 2, 2\},$$

$$\text{Sh}(0) = \left(\frac{44}{36}, -\frac{22}{36}, -\frac{22}{36} \right) + \text{Sh}(1), \quad \text{Sh}(1) = \left(-\frac{8}{36}, \frac{49}{36}, -\frac{41}{36} \right) + \text{Sh}(2),$$

$$\text{Sh}(2) = \left(-\frac{15}{18}, -\frac{15}{18}, \frac{30}{18} \right) + \text{Sh}(3), \quad \text{Sh}(3) = (2, 2, 2).$$

В данном примере вектор Шепли динамически неустойчив, так как среди величин β_k^i имеются отрицательные. Как мы замечали ранее, это обстоятельство является типичным для игр с терминальными выигрышами, т. е. когда игроки получают выигрыши лишь в окончательных позициях игры.

В данном параграфе невозможно охватить все вопросы, связанные с динамической устойчивостью принципов оптимальности. Заметим только, что понятие динамической устойчивости решений дифференциальных игр было впервые введено и исследовано Л.А. Петросяном в работах [Петросян, 1977, 1979, 1992, 1993, 1997]. Ф. Кидланд и Е. Прескотт [Kidland, Prescott, 1977] обратили внимание на данное свойство в экономическом контексте и предложили назвать его *time-consistency* (состоятельность во времени). Невыполнение динамической устойчивости вектора Шепли делает невозможным реальное применение этого принципа дележа в динамической кооперативной игре, оказывается невозможным организовать поэтапные выплаты игрокам таким образом, чтобы они могли рассчитывать на справедливое распределение выигрышей (распределение в соответствии с вектором Шепли, который они выбрали в качестве принципа оптимальности) в каждой текущей подыгре. К сожалению, многие из принципов оптимальности классической кооперативной теории оказываются динамически неустойчивыми.

Важнейшей проблемой современной теории динамических игр является построение и исследование новых динамически устойчивых принципов оптимальности.

§ 4.14. Кооперативные стохастические игры

4.14.1. Пусть задан древовидный бесконечный граф $G = (X, F)$, где X — множество вершин графа, а F — точечно-множественное отображение, заданное на множестве X , со значениями в множестве подмножеств множества X , то есть $F(x) \subset X, (x \in X)$. Предполагаем, что число элементов множества $F(x)$ равномерно ограничено некоторым числом $M > 0$. Вершина x_0 — начальная вершина древовидного графа G . Древовидный граф с начальной вершиной x_0 будем обозначать через $G(x_0)$.

Пусть в каждой вершине $x \in X$ графа $G(x_0)$ задан игровой элемент — одновременная игра n лиц в нормальной форме

$$\Gamma(x) = \langle N, U_1^x, \dots, U_n^x, K_1^x, \dots, K_n^x \rangle,$$

где $N = \{1, 2, \dots, n\}$ — множество игроков, одинаковое для всех вершин $x \in X$; U_i^x — множество стратегий i -го игрока в вершине x , предполагаем, что оно конечно для любых $x \in X$ и $i \in N$, $K_i^x(u_1^x, \dots, u_n^x)$ — функция выигрыша игрока i ($i \in N, u_i^x \in U_i^x$). Набор стратегий $u^x = (u_1^x, \dots, u_n^x)$, $u_i^x \in U_i^x, i \in N$ называется ситуацией в игровом

элементе $\Gamma(x)$, $u^x \in U^x = \prod_{i \in N} U_i^x$ — множество ситуаций в одновременной игре $\Gamma(x)$.

Предполагается, что $K_i^x(u^x) \geq 0$ для всех $x \in X$, $u^x \in U^x$ и любого игрока $i \in N$.

Для каждой вершины $x \in X$ в зависимости от ситуации u^x , реализовавшейся в игровом элементе $\Gamma(x)$, определены вероятности перехода в следующие вершины $y \in F(x)$ графа $G(x_0)$

$$p(x, y; u_1^x, \dots, u_n^x) = p(x, y; u^x) \geq 0,$$

$$\sum_{y \in F(x)} p(x, y; u^x) = 1,$$

где $p(x, y; u^x)$ — вероятность того, что реализуется игровой элемент $\Gamma(y)$ ($y \in F(x)$), если на предыдущем шаге (в одновременной игре $\Gamma(x)$) реализовалась ситуация $u^x = (u_1^x, \dots, u_n^x)$.

Кроме того, в каждой вершине $x \in X$ задана вероятность q_k того, что игра закончится на шаге k , $0 < q_k \leq 1$, шаг k в вершине $x \in X$ определяется из условия $x \in F^k(x_0)$.

Определение. Стохастической игрой $\bar{G}(x_0)$ со случайной продолжительностью будем называть систему

$$\langle N, G(x_0), \{\Gamma(x)\}_{x \in X}, \{q_k\}_{k=0}^\infty, \{p(x, y; u^x)\}_{x \in X, y \in F(x); u^x \in U^x} \rangle,$$

где x_0 — начальная вершина древовидного графа $G(x_0)$.

Из определения стохастической игры со случайной продолжительностью понятно, что считается заданным изначально древовидный граф $G(x_0)$ и множество игровых элементов $\{\Gamma(x)\}_{x \in X}$, а случайными в описанной структуре стохастической игры $\bar{G}(x_0)$ являются переходы из одних вершин графа $G(x_0)$ в другие, а также момент окончания игры.

Стохастическая игра со случайной продолжительностью $\bar{G}(x_0)$ происходит следующим образом.

- 1) В вершине x_0 графа $G(x_0)$ осуществляется игровой элемент $\Gamma(x_0)$. Пусть в нем реализуется некоторая ситуация $u^{x_0} \in U^{x_0}$. Далее игра $\bar{G}(x_0)$ либо прекращается с вероятностью q_0 , $0 < q_0 \leq 1$, либо с вероятностью $(1 - q_0)$ игра $\bar{G}(x_0)$ продолжается и переходит в вершину $y \in F(x_0)$ графа $G(x_0)$ с вероятностью $p(x_0, y; u^{x_0})$, зависящей от ситуации u^{x_0} , реализовавшейся в игровом элементе $\Gamma(x_0)$.
- 2) Предположим, что на k -ом шаге игровой процесс находится в вершине $x_k \in X$, где задан игровой элемент $\Gamma(x_k)$, и в этом игровом элементе реализуется ситуация $u^{x_k} \in U^{x_k}$. Далее игра либо заканчивается с вероятностью q_k , $0 < q_k \leq 1$, либо с вероятностью $1 - q_k$ продолжается и переходит в вершину графа $x_{k+1} \in F(x_k)$ с вероятностью $p(x_k, x_{k+1}; u^{x_k})$, зависящей от ситуации u^{x_k} , реализовавшейся в игровом элементе $\Gamma(x_k)$.

4.14.2. Обозначим через $\bar{G}(x)$ подыгру игры $\bar{G}(x_0)$, берущую начало в вершине $x \in X$ графа $G(x_0)$ (игрового элемента $\Gamma(x)$), которая естественно также является стохастической игрой со случайной продолжительностью.

Получим основные функциональные уравнения для стохастической игры со случайной продолжительностью. Предположим, что в стохастической игре со случайной продолжительностью $\bar{G}(x_0)$ реализовалась последовательность ситуаций $u^{x_0}, u^{x_1}, \dots, u^{x_l}$,

где $u^0 \in U^{x_0}$, $u^1 \in U^{x_1}$, ..., $u^l \in U^{x_l}$, ..., $x_1 \in F(x_0)$, $x_2 \in F(x_1)$, ..., $x_l \in F(x_{l-1})$, а $F(x_l) = \emptyset$. Тогда выигрыш i -го игрока определяется следующим образом:

$$K_i(x_0) = \sum_{j=0}^{\infty} q_j \left(\prod_{\substack{k < j \\ k \geq 0}} (1 - q_k) \right) \left(\sum_{m=0}^j K_i^{x_m}(u^{x_m}) \right).$$

Введем определение стратегии i -го игрока в стохастической игре со случайной продолжительностью $\bar{G}(x_0)$, которую обозначим через $u_i(\cdot)$, т. е. $u_i(\cdot)$ — это стратегия i -го игрока в игре $\bar{G}(x_0)$ или правило, по которому для каждого игрового элемента $\Gamma(x)$ ($x \in X$) определяется, какую стратегию в игровом элементе $\Gamma(x)$ выбрать,

$$u_i(x) = u_i^x$$

для всех $x \in X$, а $u_i^x \in U_i^x$.

Если $u_i(\cdot)$ — стратегия i -го игрока в игре $\bar{G}(x_0)$, то усечение этой стратегии, рассмотренное на подграфе $G(x)$ графа $G(x_0)$, которое обозначим через $u_i^x(\cdot)$, будет стратегией i -го игрока в подыгре $\bar{G}(x)$ игры $\bar{G}(x_0)$.

Пусть $u(\cdot) = (u_1(\cdot), \dots, u_n(\cdot))$ — ситуация в игре $\bar{G}(x_0)$, а $u^x(\cdot) = (u_1^x(\cdot), \dots, u_n^x(\cdot))$ — ситуация в подыгре $\bar{G}(x)$, являющаяся сужением ситуации $u(\cdot)$ на подыгру $\bar{G}(x)$.

Так как игра имеет стохастическую структуру, то в качестве выигрыша необходимо рассматривать математическое ожидание выигрыша:

$$E_i(x_0, u(\cdot)) = E(K_i(x_0), u(\cdot)).$$

Здесь $K_i(x_0, u(\cdot)) = K_i(x_0)$ есть функция выигрыша игрока i в ситуации $u(\cdot)$.

Математическое ожидание выигрыша i -го игрока $E_i(x_0)$ удовлетворяет функциональному уравнению

$$\begin{aligned} E_i(x_0, u(\cdot)) &= q_0 K_i^{x_0}(u^{x_0}) + (1 - q_0) \left[K_i^{x_0}(u^{x_0}) + \sum_{y \in F(x_0)} p(x_0, y; u^{x_0}) E_i(y, u^y(\cdot)) \right] = \\ &= K_i^{x_0}(u^{x_0}) + (1 - q_0) \sum_{y \in F(x_0)} p(x_0, y; u^{x_0}) E_i(y, u^y(\cdot)), \end{aligned}$$

где $E_i(y, u^y(\cdot))$ — математическое ожидание выигрыша i -го игрока в подыгре $\bar{G}(y)$, начинающейся в вершине $y \in X$, $y \in F(x_0)$, графа $G(x_0)$ при условии реализации ситуации $u^y(\cdot)$. Предположим, что $E_i(x, u(\cdot))$ конечно для любой вершины $x \in X$ и $u(\cdot) \in U(\cdot)$, и равномерно ограничено.

Пусть $x \in F^k(x_0)$, то есть игровой процесс на k -м шаге попадает в вершину $x \in X$. Тогда для математического ожидания выигрыша i -го игрока в подыгре $\bar{G}(x)$ справедлива формула

$$\begin{aligned} E_i(x, u^x(\cdot)) &= q_k K_i^x(u^x(\cdot)) + (1 - q_k) \left[K_i^x(u^x(\cdot)) + \sum_{y \in F(x)} p(x, y; u^x(\cdot)) E_i(y, u^y(\cdot)) \right] = \\ &= K_i^x(u^x(\cdot)) + (1 - q_k) \sum_{y \in F(x)} p(x, y; u^x(\cdot)) E_i(y, u^y(\cdot)). \end{aligned}$$

В стохастической игре со случайной продолжительностью $\bar{G}(x_0)$ в качестве смешанных стратегий игроков рассмотрим стратегии поведения. В кооперативной теории стратегии игроков используют лишь для нахождения кооперативного пути, то есть пути, который максимизирует суммарный выигрыш игроков. В случае стохастических игр это поддерево с заданными вероятностями перехода, на которых достигается максимум математического ожидания суммарного выигрыша игроков. Однако, максимум математического ожидания суммарного выигрыша игроков в классе смешанных стратегий поведения равен максимуму математического ожидания суммарного выигрыша игроков в классе чистых стратегий поведения, поэтому для нахождения кооперативного поведения в стохастической игре можно ограничиться классом чистых стратегий.

4.14.3. Построим кооперативный вариант стохастической игры со случайной продолжительностью. Обозначим через

$$\bar{u}(\cdot) = (\bar{u}_1(\cdot), \dots, \bar{u}_n(\cdot))$$

ситуацию в чистых стратегиях поведения в стохастической игре $\bar{G}(x_0)$, которая максимизирует сумму математических ожиданий выигрышей игроков

$$V(N, x_0) = \max_{u(\cdot)} \left[\sum_{i \in N} E_i(x_0, u(\cdot)) \right].$$

Назовем такую ситуацию *кооперативным решением*. Можем определить кооперативное решение для любой подыгры $\bar{G}(x)$, $x \in X$, начинающейся с игрового элемента $\Gamma(x)$ (см. [Петросян, Баранова, Шевкопляс, 2004; Petrosyan, 2006]).

Для определения кооперативного варианта стохастической игры необходимо определить характеристическую функцию для каждого подмножества S (коалиции) множества игроков N . Характеристическую функцию, вычисленную для подыгры $\bar{G}(x)$ ($x \in X$), обозначим через $V(S, x)$, где $S \subset N$.

Сначала найдем максимум суммарного выигрыша коалиции N в стохастической игре $\bar{G}(x_0)$. С этой целью выпишем уравнение Беллмана для максимума суммы математических ожиданий выигрышей игроков:

$$\begin{aligned} V(N, x_0) &= \max_{\substack{u_i^{x_0} \in U_i^{x_0} \\ i \in N}} \left[\sum_{i \in N} K_i^{x_0}(u^{x_0}) + (1 - q_0) \sum_{y \in F(x_0)} p(x_0, y; u^{x_0}) V(N, y) \right] = \\ &= \sum_{i \in N} K_i^{x_0}(\bar{u}^{x_0}) + (1 - q_0) \sum_{y \in F(x_0)} p(x_0, y; \bar{u}^{x_0}) V(N, y) \end{aligned} \quad (4.14.1)$$

с граничным условием

$$V(N, x) = \max_{\substack{u_i^x \in U_i^x \\ i \in N}} \sum_{i \in N} K_i^x(u^x), \quad x \in \{x : F(x) = \emptyset\}. \quad (4.14.2)$$

В дальнейшем, в этой главе, для определенности будем предполагать, что $x \in F^k(x_0)$.

Для подыгры $\bar{G}(x)$ ($x \in X$) уравнение (4.14.1) с начальным условием (4.14.2) принимает вид:

$$\begin{aligned} V(N, x) &= \max_{\substack{u_i^x \in U_i^x \\ i \in N}} \left[\sum_{i \in N} K_i^x(u^x) + (1 - q_k) \sum_{y \in F(x)} p(x, y; u^x) V(N, y) \right] = \\ &= \sum_{i \in N} K_i^x(\bar{u}^x) + (1 - q_k) \sum_{y \in F(x)} p(x, y; \bar{u}^x) V(N, y) \end{aligned} \quad (4.14.3)$$

с граничным условием

$$V(N, x) = \max_{\substack{u_i^x \in U_i^x \\ i \in N}} \sum_{i \in N} K_i^x(u^x), \quad x \in \{x : F(x) = \emptyset\}. \quad (4.14.4)$$

Ситуация в чистых стратегиях $\bar{u}(\cdot) = (\bar{u}_1(\cdot), \dots, \bar{u}_n(\cdot))$ в стохастической игре $\bar{G}(x_0)$ порождает вероятностные распределения на множестве X вершин графа $G(x_0)$.

Определение. Подграф графа $G(x_0)$, который состоит из вершин $x \in X$ графа $G(x_0)$, имеющих положительную вероятность реализации, порожденную ситуацией $\bar{u}(\cdot)$ (кооперативным решением), назовем кооперативным поддеревом и обозначим через $\hat{G}(x_0)$.

Очевидно, что подграф $\hat{G}(x_0)$ является древовидным графом. Множество вершин в графе $\hat{G}(x_0)$ обозначим через $CX \subset X$.

4.14.4. Определим кооперативную стохастическую игру со случайной продолжительностью, построенную на основе стохастической игры со случайной продолжительностью $\bar{G}(x_0)$, описанной выше. Для этого, для каждой вершины $x \in CX$ определим вспомогательную игру с нулевой суммой, которую обозначим через $G_S(x)$. Это антагонистическая игра между коалицией $S \subset N$, выступающей в качестве максимизирующего игрока, и коалицией $N \setminus S$, выступающей в качестве минимизирующего игрока, где выигрыш коалиции S определяется как сумма выигрышей игроков, входящих в коалицию S . Тогда значение характеристической функции $V(S, x)$ зададим как нижнее значение антагонистической игры $G_S(x)$ в чистых стратегиях (аналогично нижнему значению матричной игры).

Функция $V(S, x)$, $x \in CX$, удовлетворяет функциональному уравнению

$$\begin{aligned} V(S, x) &= \max_{u_S^x \in U_S^x} \min_{u_{N \setminus S}^x \in U_{N \setminus S}^x} \left[\sum_{i \in S} K_i^x(u_S^x, u_{N \setminus S}^x) + \right. \\ &\quad \left. + (1 - q_k) \sum_{y \in F(x)} p(x, y; u_S^x, u_{N \setminus S}^x) V(S, y) \right] \end{aligned} \quad (4.14.5)$$

с граничным условием

$$V(S, x) = \max_{u_S^x \in U_S^x} \min_{u_{N \setminus S}^x \in U_{N \setminus S}^x} \sum_{i \in S} K_i^x(u_S^x, u_{N \setminus S}^x), \quad x \in \{x : F(x) = \emptyset\}, \quad (4.14.6)$$

где игроки $i_1, i_2, \dots, i_k \in S$, $u_S^x = (u_{i_1}^x, \dots, u_{i_r}^x)$ — стратегия коалиции S , а $U_S^x = \prod_{j=1, r} U_{i_j}^x$ — множество стратегий коалиции S , игроки i_{r+1}, \dots, i_n образуют коалицию $N \setminus S$

($\{i_1, i_2, \dots, i_r\} \cup \{i_{r+1}, \dots, i_n\} = N$), и $u_{N \setminus S}^x = (u_{i_{r+1}}^x, \dots, u_{i_n}^x)$ — стратегия коалиции $N \setminus S$, а $U_{N \setminus S}^x = \prod_{j=r+1, n} U_{i_j}^x$ — множество стратегий коалиции $N \setminus S$.

Для всех $x \in CX$ естественно предположить, что

$$V(\emptyset, x) = 0. \quad (4.14.7)$$

Таким образом, для каждой подыгры $\bar{G}(x)$, $x \in CX$, мы определили характеристическую функцию $V(S, x)$, $S \subset N$, $V(S, x)$ определяется уравнением Беллмана (4.14.3) с граничным условием (4.14.4), а также уравнением (4.14.5) с граничным условием (4.14.6) и уравнением (4.14.7).

Характеристическая функция $V(S, x)$, определенная формулами (4.14.3), (4.14.4), (4.14.5), (4.14.6) и (4.14.7), супераддитивна по S .

Определение. Кооперативной стохастической игрой со случайной продолжительностью $\bar{\bar{G}}(x_0)$, основанной на стохастической игре $\bar{G}(x_0)$, назовем пару $(N, V(S, x_0))$, где $V(S, x_0)$ — характеристическая функция, определенная по формуле (4.14.1) с граничным условием (4.14.2) для коалиции N , по формуле (4.14.5) с граничным условием (4.14.6) для коалиции $S \neq \emptyset$ и для коалиции $S = \emptyset$ по формуле (4.14.7).

Определение. Дележом в кооперативной стохастической игре $\bar{\bar{G}}(x_0)$ будем называть вектор $\xi(x_0) = (\xi_1(x_0), \dots, \xi_n(x_0))$, удовлетворяющий свойствам:

- 1) $\sum_{i \in N} \xi_i(x_0) = V(N, x_0)$,
- 2) $\xi_i(x_0) \geq V(\{i\}, x_0)$, для всех $i \in N$.

Множество дележей в кооперативной стохастической игре $\bar{\bar{G}}(x_0)$ обозначим через $I(x_0)$.

Определение. Решением кооперативной стохастической игры $\bar{\bar{G}}(x_0)$ будем называть любое фиксированное подмножество $C(x_0)$ множества дележей $I(x_0)$.

4.14.5. Предположим, что решение $C(x_0)$ кооперативной стохастической игры $\bar{\bar{G}}(x_0)$ является непустым подмножеством множества дележей $I(x_0)$.

Кооперативной подыгрой $\bar{\bar{G}}(x)$, $x \in X$, кооперативной стохастической игры $\bar{\bar{G}}(x_0)$, основанной на подыгре $\bar{G}(x)$ стохастической игры $\bar{G}(x_0)$, будем называть пару $(N, V(S, x))$, где $V(S, x)$ — характеристическая функция, определенная по формуле (4.14.3) с граничным условием (4.14.4) для коалиции N , по формуле (4.14.5) с граничным условием (4.14.6) для коалиции $S \neq \emptyset$ и для коалиции $S = \emptyset$ по формуле (4.14.7).

Определение. Дележом в кооперативной подыгре $\bar{\bar{G}}(x)$ будем называть вектор $\xi(x) = (\xi_1(x), \dots, \xi_n(x))$, удовлетворяющий свойствам:

- 1) $\sum_{i \in N} \xi_i(x) = V(N, x)$,
- 2) $\xi_i(x) \geq V(\{i\}, x)$, для всех $i \in N$.

Множество дележей в кооперативной подыгре $\bar{\bar{G}}(x)$ обозначим через $I(x)$.

Определение. Решением кооперативной подыгры $\bar{\bar{G}}(x)$ будем называть любое фиксированное подмножество $C(x)$ множества дележей $I(x)$.

Предположим, что решение $C(x)$ кооперативной подыгры $\bar{\bar{G}}(x)$ является непустым подмножеством множества дележей $I(x)$ для всех $x \in CX$.

Если $C(x_0)$ — решение кооперативной стохастической игры $\bar{\bar{G}}(x_0)$, то далее под решением $C(x)$ кооперативной подыгры $\bar{\bar{G}}(x)$ будем понимать решение, построенное

по тем же правилам, что и $C(x_0)$. Например, если $C(x_0)$ — вектор Шепли для стохастической игры $\bar{G}(x_0)$, то $C(x)$ — вектор Шепли, вычисленный для кооперативной подыгры $\bar{G}(x)$. Предполагается, что игроки выбирают для себя какое-то фиксированное подмножество множества дележей, обладающее «оптимальными» для них свойствами, т. е. игроки, объединившись в коалицию N , собираются следовать некоторому правилу распределения выигрышей в течение всего игрового процесса.

4.14.6. Далее будем предполагать, что $C(x)$ является непустым подмножеством множества $I(x)$ для любого $x \in CX$, то есть для каждой вершины $x \in CX$ существует по крайней мере один дележ

$$\xi(x) = (\xi_1(x), \dots, \xi_n(x)) \in C(x) \subset I(x).$$

Определение. Вектор-функцию $\beta(x) = (\beta_1(x), \dots, \beta_n(x))$, где $x \in CX$, назовем кооперативной процедурой распределения дележа (ПРД) в вершине x [Петросян, 2004], если

$$\sum_{i \in N} \beta_i(x) = \sum_{i \in N} K_i^x(\bar{u}_1^x, \dots, \bar{u}_n^x) = \sum_{i \in N} K_i^x(\bar{u}^x), \quad (4.14.8)$$

где $\bar{u}^x = (\bar{u}_1^x, \dots, \bar{u}_n^x)$ — ситуация в игровом элементе $\Gamma(x)$, реализовавшаяся при кооперативном решении $\bar{u} = (\bar{u}_1(\cdot), \dots, \bar{u}_n(\cdot))$.

Определение. Путем в стохастической игре будем называть последовательность ситуаций $u^{x_0}, u^{x_1}, \dots, u^{x_l}, \dots$, где u^{x_i} — это ситуация, реализовавшаяся в игровом элементе $\Gamma(x_i)$, $x_i \in F(x_{i-1})$, $i = 1, \dots, l, \dots$

Пусть $x \in CX$ и $x \in F^k(x_0)$. В любой кооперативной подыгре $\bar{G}(x)$ игрок может связать с отрезком пути $\bar{u}^x, \bar{u}^y, \dots = \bar{u}^{x,y,\dots}$, реализовавшимся при кооперативном решении $\bar{u} = (\bar{u}_1(\cdot), \dots, \bar{u}_n(\cdot))$ (очевидно, что все $x, y, \dots \in CX$, так как CX — множество вершин кооперативного поддерева, и ситуация \bar{u} фиксирована), случайную величину — сумму величин $\beta_i(x)$, вычисленных вдоль этого отрезка пути $\bar{u}^{x,y,\dots}$. Будем предполагать, что игрок i на каждом шаге игры, то есть в каждой вершине пути, получает выплаты $\beta_i(x), \beta_i(y), \dots$. Математическое ожидание сумм таких выплат, посчитанных вдоль такого отрезка пути $\bar{u}^{x,y,\dots}$ в кооперативной подыгре $\bar{G}(x)$, обозначим через $B_i(x)$. Величины $B_i(x)$ удовлетворяют следующему функциональному уравнению:

$$B_i(x) = \beta_i(x) + (1 - q_k) \sum_{y \in F(x)} p(x, y; u^x) B_i(y) \quad (4.14.9)$$

с граничным условием

$$B_i(x) = \beta_i(x) \text{ для } x \in \{x : F(x) = \emptyset\}. \quad (4.14.10)$$

4.14.7. Для каждой кооперативной подыгры $\bar{G}(x)$ ($x \in CX$) составим функциональное уравнение для компонент $\xi_i(x)$ дележа $\xi(x) \in C(x) \subset I(x)$.

Определим величины $\gamma_i(x)$ из уравнения

$$\xi_i(x) = \gamma_i(x) + (1 - q_k) \sum_{y \in F(x)} p(x, y; u^x) \xi_i(y), \quad (4.14.11)$$

где $\xi(y) = (\xi_1(y), \dots, \xi_n(y))$ — некоторый дележ, принадлежащий решению $C(y)$ кооперативной подыгры $\bar{G}(y)$.

Лемма. Вектор $\gamma(x) = (\gamma_1(x), \dots, \gamma_n(x))$, определяемый формулой (4.14.11), является кооперативной процедурой распределения дележа.

Доказательство. Из (4.14.11) выразим величины $\gamma_i(x)$ и суммируем их по $i \in N$, получаем

$$\sum_{i \in N} \gamma_i(x) = \sum_{i \in N} \xi_i(x) - (1 - q_k) \sum_{i \in N} \left(\sum_{y \in F(x)} p(x, y; u^x) \xi_i(y) \right). \quad (4.14.12)$$

Так как

$$\xi(x) = (\xi_1(x), \dots, \xi_n(x)) \in C(x) \subset I(x),$$

$$\xi(y) = (\xi_1(y), \dots, \xi_n(y)) \in C(y) \subset I(y),$$

то из (4.14.12) получаем

$$\sum_{i \in N} \gamma_i(x) = V(N, x) - (1 - q_k) \sum_{y \in F(x)} p(x, y; u^x) V(N, y). \quad (4.14.13)$$

Из (4.14.13) и (4.14.3) следует, что $\sum_{i \in N} \gamma_i(x) = \sum_{i \in N} K_i^x(\bar{u}^x)$ для ситуации $\bar{u}^x = (\bar{u}_1^x, \dots, \bar{u}_n^x)$, которая реализовалась в игровом элементе $\Gamma(x)$ при использовании игроками кооперативного решения $\bar{u} = (\bar{u}_1(\cdot), \dots, \bar{u}_n(\cdot))$. Получается, что $\gamma_i(x)$ удовлетворяет условию (4.14.8), то есть лемма доказана.

Игроки перед началом игры приходят к соглашению о кооперации, то есть договариваются максимизировать математическое ожидание суммарного выигрыша и рассчитывают получить дележ $\xi(x_0) \in C(x_0)$. Развитию игры во времени соответствует движение вдоль вершин кооперативного поддерева $\hat{G}(x_0)$. Однако, поскольку стохастическая структура игры подразумевает неоднозначность в реализации вершин кооперативного поддерева, то движение вдоль некоторого пути, то есть вдоль вершин кооперативного поддерева, еще не обеспечивает сохранение кооперации. Действительно, при движении вдоль пути игроки попадают в кооперативные подыгры с текущими начальными состояниями, в которых один и тот же игрок имеет различные возможности. Это является вполне естественным обстоятельством, поскольку со временем изменяются условия конфликта и возможности участвующих в нем сторон. То есть, в некоторый момент времени, в вершине $x \in CX$, сумма оставшихся выплат для игрока i может не равняться i -ой компоненте дележа из решения $C(x)$ кооперативной подыгры $\bar{G}(x)$, являющегося решением подыгры $\bar{G}(x)$. Следовательно, в вершине $x \in CX$ перед игроком i может встать вопрос о целесообразности придерживаться далее намеченного перед началом игры соглашения действовать «совместно оптимально», то есть игрок i может пожелать отклониться от договоренного кооперативного решения и получить больший выигрыш. Если такое отклонение будет выгодно хотя бы для одного из игроков, то это и будет означать позиционную несостоятельность дележа $\xi(x_0) \in C(x_0)$ и, соответственно, самого движения вдоль вершин кооперативного поддерева.

4.14.8. Определение. Дележ $\xi(x_0) \in C(x_0)$ называется позиционно состоятельным в кооперативной стохастической игре $\bar{G}(x_0)$, если для каждой вершины $x \in CX \cap F^k(x_0)$ существует неотрицательная ПРД $\beta(x) = (\beta_1(x), \dots, \beta_n(x))$ такая, что

$$\xi_i(x) = \beta_i(x) + (1 - q_k) \sum_{y \in F(x)} p(x, y; u^x) \xi_i(y), \quad (4.14.14)$$

и

$$\xi_i(x) = \beta_i(x), \quad x \in \{x : F(x) = \emptyset\}, \quad (4.14.15)$$

где $x \in F^k(x_0)$, $\xi(y) = (\xi_1(y), \dots, \xi_n(y))$ — некоторый дележ, принадлежащий решению $C(y)$ кооперативной подыгры $\bar{G}(y)$.

Определение. Будем говорить, что кооперативная стохастическая игра со случайной продолжительностью $\bar{G}(x_0)$ имеет позиционно состоятельное решение $C(x_0)$, если являются позиционно состоятельными все дележи $\xi(x_0) \in C(x_0)$.

Отсюда получаем, что если дележ $\xi(x_0)$ позиционно состоятелен в игре $\bar{G}(x_0)$, то ПРД может быть определена для всех $x \in CX$ и таких, что $x \notin \{x : F(x) = \emptyset\}$ по формуле

$$\beta_i(x) = \xi_i(x) - (1 - q_k) \sum_{y \in F(x)} p(x, y; u^x) \xi_i(y), \quad (4.14.16)$$

а для $x \in \{x : F(x) = \emptyset\}$ по формуле (4.14.15). Однако из (4.14.16) следует, что в общем случае невозможно гарантировать неотрицательность $\beta_i(x)$ для всех вершин $x \in CX$.

Найдем значения величин $B_i(x)$ для всех $x \in CX$, то есть математические ожидания сумм $\beta_i(x)$ из (4.14.9) и (4.14.10), $x \in CX$, вдоль путей, реализовавшихся в кооперативной подыгре $\bar{G}(x)$ игры $\bar{G}(x_0)$ (вдоль путей подграфа $\hat{G}(x_0)$) при использовании игроками кооперативного решения $\bar{u}(\cdot)$.

Лемма. Имеет место равенство $B_i(x) = \xi_i(x)$ для всех $x \in CX$ и всех $i \in N$.

Доказательство. Доказательство является очевидным и следует из того, что $B_i(x)$ и $\xi_i(x)$ удовлетворяют одним и тем же функциональным уравнениям (4.14.9) и (4.14.14) с одними и теми же граничными условиями (4.14.10) и (4.14.15).

Если выплаты игрокам производить не в соответствии с их выигрышами в игровых элементах, по которым проходит кооперативный путь, а в соответствии с кооперативной процедурой распределения дележа $\beta(x) = (\beta_1(x), \dots, \beta_n(x))$, определенной формулами (4.14.14), (4.14.15) для всех $x \in CX$, где $\beta_i(x)$ — это выплата i -му игроку в вершине $x \in CX$, то математическое ожидание всех выплат i -му игроку будет совпадать с математическим ожиданием i -ой компоненты выбранного игроками дележа из решения, что следует из леммы § 4.14. Таким образом, игроки могут пойти и на получение в каких-то вершинах отрицательных выплат, чтобы гарантировать сохранение коалиции на протяжении всей игры и получение компонент заранее выбранного дележа $\xi(x_0)$, принадлежащего решению $C(x_0)$ кооперативной стохастической игры $\bar{G}(x_0)$.

Предлагаемый способ реализации дележа обладает важным свойством: в каждой вершине пути игроки ориентируются на один и тот же «принцип оптимальности» и, в этом смысле, не имеют оснований для нарушения ранее принятого кооперативного поведения, то есть реализации кооперативного решения.

В случае, когда нельзя гарантировать неотрицательность $\beta_i(x)$ для всех вершин $x \in CX$, можно пойти по пути построения нового позиционно состоятельного решения на основе решения из классической теории кооперативных игр. Покажем, как это делается, когда в качестве решения рассматривается множество $C(x_0) \subset I(x_0)$. Заметим, что данная процедура может быть применена для дележей, известных в классической статической кооперативной теории (C -ядро, N -ядро, вектор Шепли и т. д.).

Для всех вершин $x \in CX$ определим новую ПРД по формуле

$$\beta_i(x) = \frac{\sum_{i \in N} K_i(\bar{u}_1^x, \dots, \bar{u}_n^x)}{V(N, x)} \xi_i(x), \quad (4.14.17)$$

где $\xi(x) = (\xi_1(x), \dots, \xi_n(x)) \in C(x)$, а $\bar{u}^x = (\bar{u}_1^x, \dots, \bar{u}_n^x)$ — реализация кооперативного решения $\bar{u} = (\bar{u}_1(\cdot), \dots, \bar{u}_n(\cdot))$ в вершине $x \in CX$, максимизирующего сумму математических ожиданий выигрышей игроков в стохастической игре $\bar{G}(x_0)$, $V(N, x)$ — значение характеристической функции для коалиции N , вычисленное для кооперативной подыгры $\bar{G}(x)$.

Поскольку $K_i(u^x) \geq 0$ для всех вершин $x \in CX$ и всех $i \in N$, то $\beta_i(x) \geq 0$ для любой вершины $x \in CX$. Из того, что $\sum_{i \in N} \xi_i(x) = V(N, x)$, и из (4.14.17) следует также, что текущая выплата $\beta_i(x)$ игроку i в игровом элементе $\Gamma(x)$ должна быть пропорциональна i -ой компоненте дележа $\xi(x) \in C(x)$ в кооперативной подыгре $\bar{G}(x)$ стохастической игры $\bar{G}(x_0)$.

Определим компоненты нового дележа в кооперативной подыгре $\bar{G}(x)$, где $x \in CX$, и $x \in F^k(x_0)$ на основе «старого» дележа $\xi(x)$ как решение функционального уравнения

$$\hat{\xi}_i(x) = \frac{\sum_{i \in N} K_i(\bar{u}^x)}{V(N, x)} \xi_i(x) + (1 - q_k) \sum_{y \in F(x)} p(x, y; \bar{u}^x) \hat{\xi}_i(y) \quad (4.14.18)$$

с граничным условием

$$\hat{\xi}_i(x) = \frac{\sum_{i \in N} K_i(\bar{u}^x)}{V(N, x)} \xi_i(x) = \xi_i(x) \quad (4.14.19)$$

для $x \in \{x : F(x) = \emptyset\}$.

Построим новую характеристическую функцию $\hat{V}(S, x)$ для каждой кооперативной подыгры $\bar{G}(x)$ для всех $x \in CX$, используя функциональное уравнение

$$\hat{V}(S, x) = \frac{\sum_{i \in N} K_i(\bar{u}^x)}{V(N, x)} V(S, x) + (1 - q_k) \sum_{y \in F(x)} p(x, y; \bar{u}^x) \hat{V}(S, y) \quad (4.14.20)$$

с граничным условием

$$\hat{V}(S, x) = V(S, x) \text{ для } x \in \{x : F(x) = \emptyset\}. \quad (4.14.21)$$

Функции $\hat{V}(S, x)$ и $V(S, x)$ супераддитивны, и $\hat{V}(N, x) = V(N, x)$.

Для всех позиционно несостоятельных дележей $\xi(x) \in C(x)$ для всех $x \in CX$ вычислим регуляризированные дележи $\hat{\xi}(x)$ и найдем множество $\hat{C}(x)$ следующим образом:

$$\hat{C}(x) = \{\hat{\xi}(x) : \hat{\xi}_i(x) = \frac{\sum_{i \in N} K_i(\bar{u}^x)}{V(N, x)} \xi_i(x) + (1 - q_k) \sum_{y \in F(x)} p(x, y; \bar{u}^x) \hat{\xi}_i(y) \quad (4.14.22)$$

с граничным условием $\hat{\xi}_i(x) = \xi_i(x)$ для $x \in \{x : F(x) = \emptyset\}$,
где $\xi(x) \in C(x)$.

Множество $\hat{C}(x_0)$, определенное формулой (4.14.22), назовем регуляризованным решением кооперативной стохастической игры со случайной продолжительностью $\bar{G}(x_0)$.

Таким образом, у игроков есть возможность регуляризовать выбранное ими в начале игры решение, чтобы в каждой вершине стохастической игры $\bar{G}(x_0)$ «новое» решение

$\widehat{C}(x_0)$ было позиционно состоятельным. Но дележ, принадлежащий новому регуляризованному решению $\widehat{C}(x_0)$, вообще говоря, не будет являться дележом для кооперативной игры с характеристической функцией $V(S, x_0)$, определенной по формулам (4.14.5), (4.14.6), а будет являться дележом для кооперативной стохастической игры с новой характеристической функцией $\widehat{V}(S, x_0)$, построенной по формулам (4.14.20), (4.14.21).

4.14.9. Теорема. Дележ $\widehat{\xi}(x) = (\widehat{\xi}_1(x), \dots, \widehat{\xi}_n(x))$, определенный формулой (4.14.18) с граничным условием (4.14.19), является позиционно состоятельным дележом в кооперативной игре (N, \widehat{V}) , где характеристическая функция $\widehat{V}(S, x)$ задана функциональным уравнением (4.14.20) с граничным условием (4.14.21).

Доказательство. Позиционная состоятельность следует из построения «нового дележа» $\widehat{\xi}(x)$. Для доказательства необходимо показать неотрицательность $\frac{\sum_{i \in N} K_i(\bar{u}^x)}{V(N, x)} \xi_i(x)$, а это очевидно, поскольку в определении стохастической игры было задано условие:

$$K_i^x(u_1^x, \dots, u_n^x) \geq 0$$

для всех $x \in X$ и любого игрока $i \in N$.

Теперь покажем, что $\widehat{\xi}(x) = (\widehat{\xi}_1(x), \dots, \widehat{\xi}_n(x))$ обладает свойствами дележа для кооперативной игры с характеристической функцией $\widehat{V}(S, x)$, которая задана функциональным уравнением (4.14.20) с граничным условием (4.14.21). Для этого необходимо показать свойства:

$$1) \sum_{i \in N} \widehat{\xi}_i(x) = \widehat{V}(N, x),$$

$$2) \widehat{\xi}_i(x) \geq \widehat{V}(\{i\}, x)$$

для любого игрока $i \in N$ и любой вершины $x \in CX$.

Первое свойство очевидно для вершин $x \in \{x : F(x) = \emptyset\}$ и таких, что $x \in CX$. Теперь докажем его для вершин $x \in \{x : F(x) \ni y \text{ и } F(y) = \emptyset\}$ и таких, что $x \in CZ$.

$$\begin{aligned} \sum_{i \in N} \widehat{\xi}_i(x) &= \frac{\sum_{i \in N} K_i(\bar{u}^x)}{V(N, x)} \sum_{i \in N} \xi_i(x) + (1 - q_k) \sum_{y \in F(x)} \left(p(x, y; \bar{u}^x) \sum_{i \in N} \widehat{\xi}_i(y) \right) = \\ &= \frac{\sum_{i \in N} K_i(\bar{u}^x)}{V(N, x)} V(N, x) + (1 - q_k) \sum_{y \in F(x)} \left(p(x, y; \bar{u}^x) \widehat{V}(N, y) \right) = \\ &= V(N, x) = \widehat{V}(N, x), \end{aligned}$$

т. к. вершина $y \in \{y : F(y) = \emptyset\}$.

Второе свойство также очевидно для вершин $x \in \{x : F(x) = \emptyset\}$. Для вершин $x \in \{x : F(x) \ni y \text{ и } F(y) = \emptyset\}$ покажем, что $\widehat{\xi}_i(x) - \widehat{V}(\{i\}, x) \geq 0$, используя формулы (4.14.20) и (4.14.18):

$$\begin{aligned} \widehat{\xi}_i(x) - \widehat{V}(\{i\}, x) &= \frac{\sum_{i \in N} K_i(\bar{u}^x)}{V(N, x)} \xi_i(x) + (1 - q_k) \sum_{y \in F(x)} p(x, y; \bar{u}^x) \widehat{\xi}_i(y) - \\ &\quad - \left\{ \frac{\sum_{i \in N} K_i(\bar{u}^x)}{V(N, x)} V(\{i\}, x) + (1 - q_k) \sum_{y \in F(x)} p(x, y; \bar{u}^x) \widehat{V}(\{i\}, y) \right\} = \\ &= \frac{\sum_{i \in N} K_i(\bar{u}^x)}{V(N, x)} (\xi_i(x) - V(\{i\}, x)) + (1 - q_k) \sum_{y \in F(x)} p(x, y; \bar{u}^x) (\widehat{\xi}_i(y) - \widehat{V}(\{i\}, y)) \geq 0. \end{aligned}$$

Первое слагаемое неотрицательно, поскольку $\xi(x)$ является дележом кооперативной подыгры $\overline{G}(x)$, а второе слагаемое неотрицательно, потому что вершина $y \in \{y : F(y) = \emptyset\}$. Используя метод математической индукции, свойства 1 и 2 могут быть доказаны для всех вершин.

§ 4.15. Марковские игры

4.15.1. Пусть задано конечное множество одновременных игр n лиц в нормальной форме (игровых элементов) $\{\Gamma^1, \dots, \Gamma^t\}$:

$$\Gamma^j = \langle N, U_1^j, \dots, U_n^j, K_1^j, \dots, K_n^j \rangle,$$

где N — множество игроков, одинаковое для всех игр Γ^j , $j = 1, \dots, t$, U_i^j ($i = 1, \dots, n$) — множество стратегий i -го игрока в игре Γ^j , множество U_i^j конечно для любого игрока $i \in N$ и любого игрового элемента Γ^j , $j = 1, \dots, t$, $K_i^j(u_1^j, \dots, u_n^j)$ — функция выигрыша i -го игрока в игре Γ^j . Причем, $K_i^j(u_1^j, \dots, u_n^j) \geq 0$ для всех $i \in N$, для любого $j = 1, \dots, t$ и любого $u_i^j \in U_i^j$. Набор стратегий $u^j = (u_1^j, \dots, u_n^j)$ называется ситуацией в игре Γ^j . Предполагаем также, что множество стратегий каждого из игроков конечно для всех игровых элементов из множества $\{\Gamma^1, \dots, \Gamma^t\}$.

Для каждого игрового элемента Γ^j , $j = 1, \dots, t$, в зависимости от ситуации u^j , реализовавшейся в этом игровом элементе, определены вероятности перехода в следующие игровые элементы $\Gamma^1, \dots, \Gamma^t$:

$$p(j, k; u^j) \geq 0,$$

$$\sum_{k=1}^t p(j, k; u^j) = 1,$$

где $p(j, k; u^j)$ — вероятность того, что состоится одновременная игра Γ^k , если на предыдущем шаге (в игровом элементе Γ^j) реализовалась ситуация $u^j = (u_1^j, \dots, u_n^j)$.

Задана вероятность окончания игры на каждом шаге — q ($0 < q \leq 1$). Также будем считать, что задан вектор $\pi = (\pi_1, \dots, \pi_n)$ начального распределения вероятностей на

множестве игровых элементов $\{\Gamma^1, \dots, \Gamma^t\}$, где π_j ($j = 1, \dots, t$) — вероятность того, что на «нулевом» шаге (то есть перед началом игрового процесса) «случай» выбирает игровой элемент Γ^j , $\sum_{j=1}^t \pi_j = 1$.

Определение. Марковской игрой \bar{G} со случайной продолжительностью называется следующий набор

$$G = \left\langle N, \{\Gamma^j\}_{j=1}^t, q, \pi, \{p(j, k; u^j)\}_{j=\overline{1, t}, k=\overline{1, t}, u^j \in \prod_{l=1}^n U_l^j} \right\rangle. \quad (4.15.1)$$

Марковская игра \bar{G} происходит следующим образом.

- 1) Перед началом игры «случай» выбирает игровой элемент в соответствии с вектором начального распределения вероятностей π . Другими словами, с вероятностью π_1 марковская игра \bar{G} начнется с игрового элемента Γ^1 , с вероятностью π_2 — с игрового элемента Γ^2 и так далее. Пусть реализуется вероятность π_j , то есть «случай» выбрал игровой элемент $\Gamma^j \in \{\Gamma^1, \dots, \Gamma^t\}$,
- 2) Пусть на первом шаге марковской игры \bar{G} в игровом элементе Γ^j реализуется ситуация $u^j \in U^j = \prod_{k=1}^n U_k^j$. Далее стохастическая игра с конечным числом игровых элементов \bar{G} либо прекращается с вероятностью q , $0 < q \leq 1$, либо с вероятностью $(1 - q)$ переходит на следующий шаг.
- 3) На втором шаге игры \bar{G} происходит одна из одновременных игр $\Gamma^1, \dots, \Gamma^t$ с вероятностями $p(j, 1; u^j), \dots, p(j, t; u^j)$ соответственно. Далее игра \bar{G} может прекратиться с вероятностью q , $0 < q \leq 1$, либо с вероятностью $(1 - q)$ переходит на третий шаг.
- 4) Аналогично предыдущему реализуется третий шаг и т. д.

Подыгру стохастической игры с конечным числом игровых элементов, начинающуюся с k -го шага, обозначим через $\bar{G}(k)$.

4.15.2. Замечание. Марковская игра, рассматриваемая в данном параграфе, является частным случаем игры, рассматриваемой в параграфе §4.14, поскольку в данной постановке задачи предполагается, что множество игровых элементов конечно. Марковская игра \bar{G} представляет собой марковскую цепь с конечным числом состояний и вектором начального распределения π , конечным множеством состояний $\{\Gamma^1, \dots, \Gamma^t\}$ и матрицей вероятностей перехода, состоящей из элементов

$$\Pi(u) = \{p(j, k; u^j)\}_{j=\overline{1, t}, k=\overline{1, t}, u^j \in U^j}.$$

4.15.3. Выпишем основные функциональные уравнения для марковской игры. Так как игра \bar{G} имеет стохастическую структуру, поэтому в качестве выигрышей игроков будем рассматривать математическое ожидание выигрышей. Для математического ожидания выигрыша i -го игрока в игре \bar{G} имеет место формула

$$\bar{E}_i = \pi_1 E_i^1 + \dots + \pi_t E_i^t = \pi E_i,$$

где $E_i = (E_i^1, \dots, E_i^t)$, а E_i^j ($j = 1, \dots, t$) — математическое ожидание выигрыша i -го игрока в подыгре $\bar{G}(1)$ при условии, что игра начинается с игрового элемента Γ^j .

Для условного математического ожидания выигрыша игрока i будет верна формула

$$E_i^j(u^j) = K_i^j(u^j) + (1 - q) \sum_{k=1}^t p(j, k; u^j) E_i^k(u^k), \quad (4.15.2)$$

где $u^j \in U^j = \prod_{k=1}^n U_k^j$ — ситуация, которая реализовалась в игровом элементе Γ^j , $u^k \in U^k$ — ситуация, которая реализовалась в игровом элементе Γ^k .

Стратегия $\eta_i(\cdot)$ i -го игрока в игре \bar{G} — это правило, по которому для каждого игрового элемента $\Gamma^j \in \{\Gamma^1, \dots, \Gamma^t\}$ ($j = \overline{1, t}$) определяется, какую стратегию в одновременной игре Γ^j выбрать.

Определение марковской игры предполагает, что игровой процесс на каждом шаге может попадать в одни и те же игровые элементы, поэтому в данной постановке стохастической игры целесообразно использовать класс стационарных стратегий.

Определение. Стратегия $\eta_i(\cdot)$ игрока i в игре \bar{G} называется стационарной, если выбор стратегии в каждом игровом элементе из множества $\{\Gamma^1, \dots, \Gamma^t\}$ на каждом шаге зависит только от того, какой игровой элемент реализуется на этом шаге, то есть

$$\eta_i : \Gamma^j \mapsto u_i^j \in U_i^j, j = \overline{1, t}.$$

В марковской игре, рассматриваемой в классе стационарных стратегий, выбор игроками стратегии в игровых элементах не зависит от ситуаций, реализовавшихся на предыдущих шагах, а зависит только от самого игрового элемента, в котором на данном шаге игра \bar{G} находится.

Марковская игра \bar{G} рассматривается в классе стационарных стратегий, и множество одновременных игр $\{\Gamma^1, \dots, \Gamma^t\}$ конечно, поэтому достаточно рассмотреть t подыгр игры \bar{G} , обозначенных через $\bar{G}^1, \dots, \bar{G}^t$, начинающихся с игровых элементов $\Gamma^1, \dots, \Gamma^t$ соответственно.

Пусть $\eta(\cdot) = (\eta_1(\cdot), \dots, \eta_n(\cdot))$ — ситуация в стационарных стратегиях, такая что $\eta_i(\Gamma^j) = u_i^j \in U_i^j$ где $j = \overline{1, t}$, $i \in N$. Так же, как и в предыдущем параграфе, будем рассматривать класс чистых стационарных стратегий.

4.15.4. Множество чистых стационарных стратегий i -го игрока в марковской игре \bar{G} обозначим через Ξ_i . Очевидно, что стационарная стратегия i -го игрока в игре \bar{G} будет являться стационарной стратегией в любой подыгре $\bar{G}^1, \dots, \bar{G}^t$.

Рассмотрим матрицу вероятностей перехода из одних игровых элементов в другие, которая зависит от ситуации в стационарных стратегиях $\eta(\cdot)$, реализовавшейся в марковской игре \bar{G} :

$$\Pi(\eta(\cdot)) = \begin{pmatrix} p(1, 1; u^1) & \dots & p(1, t; u^1) \\ p(2, 1; u^2) & \dots & p(2, t; u^2) \\ \dots & \dots & \dots \\ p(t, 1; u^t) & \dots & p(t, t; u^t) \end{pmatrix}, \quad (4.15.3)$$

где $\eta(\cdot) = (\eta_1(\cdot), \dots, \eta_n(\cdot))$, а $\eta_i(\cdot) : \eta_i(\Gamma^1) = u_i^1 \in U_i^1, \dots, \eta_i(\Gamma^t) = u_i^t \in U_i^t$, а ситуация, реализовавшаяся в игровом элементе Γ^j , ($j = \overline{1, t}$), — это ситуация $u^j = (u_1^j, \dots, u_n^j)$.

Тогда можно записать формулу для вычисления математического ожидания выигрыша i -го игрока в любой подыгре марковской игры \bar{G} при реализации ситуации в чистых стационарных стратегиях $\eta(\cdot) \in \prod_{i \in N} \Xi_i$, описанной выше. Имеет место формула

$$E_i(\eta(\cdot)) = K_i(\eta(\cdot)) + (1 - q)\Pi(\eta(\cdot))E_i(\eta(\cdot)), \quad (4.15.4)$$

где $E_i(\eta(\cdot)) = (E_i^1(\eta(\cdot)), \dots, E_i^t(\eta(\cdot)))$ а $E_i^j(\cdot)$ — математическое ожидание выигрыша i -го игрока в подыгре \bar{G}^j , то есть подыгре, начинающейся с игрового элемента Γ^j , $K_i(\eta(\cdot)) = (K_i^1(u^1), \dots, K_i^t(u^t))$, где $K_i^j(u^j)$ — это выигрыш i -го игрока в игровом элементе Γ^j при условии, что в этом игровом элементе реализовалась ситуация $u^j \in U^j$. Если $\det(E - (1 - q)\Pi(\eta(\cdot))) \neq 0$, то из (4.15.4) получаем

$$E_i(\eta(\cdot)) = (E - (1 - q)\Pi(\eta(\cdot)))^{-1} K_i(\eta(\cdot)). \quad (4.15.5)$$

Формула (4.15.5) позволяет в явном виде вычислять математическое ожидание выигрыша для любой подыгры марковской игры \bar{G} для каждой ситуации $\eta(\cdot) \in \prod_{i \in N} \Xi_i$.

Для марковской игры \bar{G} математическое ожидание выигрыша i -го игрока обозначим через $\bar{E}_i(\eta(\cdot))$, оно зависит от вектора начального распределения вероятностей π и математического ожидания выигрышей i -го игрока $E_i^1(\cdot), \dots, E_i^t(\cdot)$ в подыграх $\bar{G}^1, \dots, \bar{G}^t$ соответственно, и может быть найдено по формуле

$$\bar{E}_i(\eta(\cdot)) = \pi E_i(\eta(\cdot)) = \pi (E - (1 - q)\Pi(\eta(\cdot)))^{-1} K_i(\eta(\cdot)). \quad (4.15.6)$$

4.15.5. Построим кооперативный вариант марковской игры. Обозначим через $\bar{\eta}(\cdot) = (\bar{\eta}_1(\cdot), \dots, \bar{\eta}_n(\cdot))$ ситуацию в чистых стационарных стратегиях, максимизирующую сумму математических ожиданий выигрышей игроков в марковской игре \bar{G} , то есть

$$\max_{\eta(\cdot) \in \prod_{i \in N} \Xi_i} \sum_{i \in N} \bar{E}_i(\eta(\cdot)) = \sum_{i \in N} \bar{E}_i(\bar{\eta}(\cdot)). \quad (4.15.7)$$

Будем называть ситуацию $\bar{\eta}(\cdot)$ *кооперативным решением*.

Определим характеристическую функцию $\bar{V}(N)$ в марковской игре \bar{G} следующим образом. Сначала определим значения характеристической функции для каждой подыгры $\bar{G}^1, \dots, \bar{G}^t$ марковской игры \bar{G} . Сформируем вектор, состоящий из значений характеристических функций для коалиции N для подыгр игры \bar{G} : $V(N) = (V^1(N), \dots, V^t(N))$, где $V^j(N)$ — максимальный суммарный выигрыш коалиции N в подыгре \bar{G}^j , $j = 1, \dots, t$.

Для вектора $V(N)$ уравнение Беллмана принимает вид:

$$\begin{aligned} V(N) &= \max_{\eta(\cdot) \in \prod_{i \in N} \Xi_i} \left[\sum_{i \in N} K_i(\eta(\cdot)) + (1 - q)\Pi(\eta(\cdot))V(N) \right] = \\ &= \sum_{i \in N} K_i(\bar{\eta}(\cdot)) + (1 - q)\Pi(\bar{\eta}(\cdot))V(N), \end{aligned}$$

где $\bar{\eta}(\cdot)$ — ситуация в чистых стационарных стратегиях, которая удовлетворяет условию (4.15.7). Если $\det(E - (1 - q)\Pi(\bar{\eta}(\cdot))) \neq 0$, то

$$V(N) = (E - (1 - q)\Pi(\bar{\eta}(\cdot)))^{-1} \sum_{i \in N} K_i(\bar{\eta}(\cdot)). \quad (4.15.8)$$

Для стохастической игры с конечным числом игровых элементов \bar{G} максимальный суммарный выигрыш коалиции N обозначим через $\bar{V}(N)$ и вычислим по формуле:

$$\bar{V}(N) = \pi V(N) = \pi (E - (1 - q)\Pi(\bar{\eta}(\cdot)))^{-1} \sum_{i \in N} K_i(\bar{\eta}(\cdot)). \quad (4.15.9)$$

Для каждой подыгры \bar{G}^j ($j = 1, \dots, t$) марковской игры \bar{G} определим вспомогательную марковскую игру с нулевой суммой \bar{G}_S^j между коалицией $S \subset N$, выступающей в качестве максимизирующего игрока, и коалицией $N \setminus S$, выступающей в качестве минимизирующего игрока. Значение функции $V^j(S)$ для подыгры \bar{G}^j зададим как нижнее значение антагонистической марковской игры \bar{G}_S^j , найденное в чистых стратегиях (фактически как нижнее значение матричной игры):

$$V^j(S) = \max_{\eta_S(\cdot)} \min_{\eta_{N \setminus S}(\cdot)} \sum_{i \in S} E_i^j(\eta_S(\cdot), \eta_{N \setminus S}(\cdot)), j = \overline{1, t}, \quad (4.15.10)$$

$$V(\emptyset) = 0,$$

где пара $(\eta_S(\cdot), \eta_{N \setminus S}(\cdot))$ образует некоторую ситуацию в чистых стационарных стратегиях, а $\eta_S(\cdot) = (\eta_{i_1}(\cdot), \dots, \eta_{i_r}(\cdot))$ — вектор стационарных стратегий игроков $i_1, \dots, i_r \in S$, $i_1 \cup \dots \cup i_r = S$, $\eta_S(\cdot) \in \prod_{j=1}^r \Xi_{i_j}$ — множество чистых стратегий коалиции $S \subset N$, а $\eta_{N \setminus S}(\cdot)$ — вектор стационарных стратегий игроков $i_{r+1}, \dots, i_n \in N \setminus S$, $i_{r+1} \cup \dots \cup i_n = N \setminus S$, $\prod_{j=r+1}^n \Xi_{i_j}$ — множество чистых стратегий коалиции $N \setminus S$. Обозначим через $V(S)$ вектор $(V^1(S), \dots, V^t(S))$, где все $V^j(S)$, $j = 1, \dots, t$, удовлетворяют условию (4.15.10).

Тогда значение характеристической функции для коалиции S для марковской игры \bar{G} обозначим через $\bar{V}(S)$ и вычислим по формуле:

$$\bar{V}(S) = \pi V(S), \quad (4.15.11)$$

$$\bar{V}(\emptyset) = 0.$$

Определение. Кооперативной марковской игрой, основанной на игре \bar{G} , называется кооперативная игра $\bar{\bar{G}} = \langle N, \bar{V} \rangle$, где N — это множество игроков, а $\bar{V} : S \rightarrow R$ — характеристическая функция, определенная по формулам (4.15.9) и (4.15.11).

Определение. Кооперативной марковской подыгрой, основанной на подыгре \bar{G}^j , назовем кооперативную игру $\bar{\bar{G}}^j = \langle N, V \rangle$, где N — это множество игроков, а $V : S \rightarrow R$ — характеристическая функция, определенная по формулам (4.15.8) и (4.15.10).

Характеристическая функция $\bar{V}(S)$, определяемая формулами (4.15.9) и (4.15.11), и функция $V(S)$, определенная формулами (4.15.8) и (4.15.10), являются супераддитивными характеристическими функциями.

4.15.6. Определение. Дележом в кооперативной марковской игре $\bar{\bar{G}}$ будем называть вектор $\bar{\xi} = (\bar{\xi}_1, \dots, \bar{\xi}_n)$, удовлетворяющий свойствам:

$$1) \sum_{i \in N} \bar{\xi}_i = \bar{V}(N),$$

$$2) \bar{\xi}_i \geq \bar{V}(\{i\}), \quad i \in N.$$

Множество дележей в кооперативной марковской игре $\bar{\bar{G}}$ обозначим через \bar{I} .

Определение. Решением кооперативной марковской игры $\bar{\bar{G}}$ будем называть любое фиксированное подмножество \bar{I} множества дележей \bar{I} .

Решением может быть одно из классических решений статической теории кооперативных игр (C -ядро, N -ядро, вектор Шепли). Предположим, что решение \bar{C} кооперативной марковской игры \bar{G} является непустым подмножеством множества дележей \bar{I} .

Вектором Шепли кооперативной марковской игры \bar{G} будем называть вектор $\bar{Sh} = (\bar{Sh}_1, \dots, \bar{Sh}_n)$, где элемент \bar{Sh}_i вычисляется по формуле

$$\bar{Sh}_i = \sum_{\substack{S \subset N \\ S \ni i}} \frac{(|S| - 1)!(n - |S|)!}{n!} (\bar{V}(S) - \bar{V}(S \setminus \{i\})),$$

$i \in N$ и $|S|$ — мощность множества S .

Определения дележа, множества дележей, решения, вектора Шепли для любой кооперативной подыгры \bar{G}^j ($j = 1, \dots, t$) кооперативной марковской игры \bar{G} вводятся аналогичным образом, только в качестве характеристической функции выступает соответствующая функция, определенная формулами (4.15.8) и (4.15.10).

Обозначим через Sh_i вектор (Sh_i^1, \dots, Sh_i^t) , $i \in N$. Для игрока i вектор Sh_i , состоящий из i -ых компонент векторов Шепли, рассчитанных для кооперативных подыгр $\bar{G}^1, \dots, \bar{G}^t$ соответственно, и i -ая компонента вектора Шепли \bar{Sh} , рассчитанного для марковской игры \bar{G} связаны следующим соотношением

$$\bar{Sh}_i = \pi Sh_i.$$

4.15.7. Определение. Вектор-функцию $\beta^j = (\beta_1^j, \dots, \beta_n^j)$, $j = 1, \dots, t$, назовем кооперативной процедурой распределения дележа (ПРД) в игровом элементе Γ^j , если

$$\sum_{i \in N} \beta_i^j = \sum_{i \in N} K_i^j(\bar{u}_1^j, \dots, \bar{u}_n^j), \quad (4.15.12)$$

где выражение в правой части — суммарный выигрыш игроков в игровом элементе Γ^j , вычисленный при реализации кооперативного решения $\bar{\eta}(\cdot)$ в чистых стационарных стратегиях, удовлетворяющего условию (4.15.7) и такой, что $\bar{\eta}(\Gamma^j) = \bar{u}^j$, $j = 1, \dots, t$.

Определение. Путем в кооперативной марковской игре \bar{G} будем называть последовательность реализовавшихся ситуаций $u(1), \dots, u(j), \dots$, где $u(j)$ — это ситуация, реализовавшаяся на j -ом шаге игры.

Пусть \bar{u} — это путь в игре \bar{G} , полученный при реализации кооперативного решения $\bar{\eta}(\cdot)$, определенного формулой (4.15.7). В любой марковской подыгре \bar{G}^j ($j = \bar{1}, \bar{t}$) путем будет являться усечение последовательности \bar{u} . Так как множество игровых элементов конечно, и марковская игра \bar{G} рассматривается в классе стационарных стратегий, то для подыгры \bar{G}^j (начинающейся с одновременной игры Γ^j) путем при условии, что игроки придерживаются кооперативного решения $\bar{\eta}(\cdot)$, будет последовательность ситуаций, которую обозначим через $\bar{u}^{\{j\}}$.

Предположим, что игрок i в подыгре \bar{G}^j вдоль пути \bar{u}^j получает некоторые выплаты — β_i^j, \dots . Тогда в кооперативной марковской подыгре \bar{G}^j , $j = 1, \dots, t$, игрок i может связать с реализовавшимся путем $\bar{u}^{\{j\}}$ в ситуации $\bar{\eta}(\cdot) = (\bar{\eta}_1(\cdot), \dots, \bar{\eta}_n(\cdot))$ случайную величину — сумму величин β_i , вычисленных вдоль этого пути. Математическое ожидание таких β_i , посчитанных вдоль пути $\bar{u}^{\{j\}}$ в кооперативной подыгре \bar{G}^j , обозначим

через B_i^j . Величины B_i^1, \dots, B_i^t , вычисленные для подыгр $\bar{G}^1, \dots, \bar{G}^t$ соответственно, образуют вектор $B_i = (B_i^1, \dots, B_i^t)$, и B_i удовлетворяет следующему уравнению

$$B_i = \beta_i + (1 - q)\Pi(\bar{\eta}(\cdot))B_i,$$

где $\beta_i = (\beta_i^1, \dots, \beta_i^t)$, а β_i^j — это выплата i -му игроку на том шаге игры, на котором реализуется игровой элемент Γ^j . Если $\det(E - (1 - q)\Pi(\bar{\eta}(\cdot))) \neq 0$, то можно написать:

$$B_i = (E - (1 - q)\Pi(\bar{\eta}(\cdot)))^{-1}\beta_i. \quad (4.15.13)$$

Математическое ожидание суммы выплат i -му игроку, посчитанной вдоль пути \bar{u} в кооперативной марковской игре \bar{G} обозначим через \bar{B}_i , для этой величины верна формула:

$$\bar{B}_i = \pi B_i = \pi(E - (1 - q)\Pi(\bar{\eta}(\cdot)))^{-1}\beta_i. \quad (4.15.14)$$

Определим вектор $\gamma_i = (\gamma_i^1, \dots, \gamma_i^t)$ из уравнения

$$Sh_i = \gamma_i + (1 - q)\Pi(\bar{\eta}(\cdot))Sh_i. \quad (4.15.15)$$

Лемма. Вектор $\gamma_i = (\gamma_i^1, \dots, \gamma_i^t)$, определяемый формулой (4.15.15), состоит из i -ых компонент ПРД игровых элементов $\Gamma^1, \dots, \Gamma^t$ соответственно.

Доказательство. Для доказательства достаточно проверить выполнение неравенства (4.15.12). Из уравнения (4.15.15) следует:

$$\gamma_i = (E - (1 - q)\Pi(\bar{\eta}(\cdot)))Sh_i.$$

Суммируя эти равенства по всем игрокам $i \in N$, получаем

$$\sum_{i \in N} \gamma_i = (E - (1 - q)\Pi(\bar{\eta}(\cdot)))V(N).$$

Учитывая равенство (4.15.8), получаем

$$\sum_{i \in N} \gamma_i = (E - (1 - q)\Pi(\bar{\eta}(\cdot))) (E - (1 - q)\Pi(\bar{\eta}(\cdot)))^{-1} \sum_{i \in N} K_i(\bar{\eta}(\cdot)),$$

следовательно, верно равенство:

$$\sum_{i \in N} \gamma_i = \sum_{i \in N} K_i(\bar{\eta}(\cdot)),$$

что и доказывает лемму.

4.15.8. Игроки перед началом игры договариваются о выборе набора стратегий, гарантирующего максимальный суммарный выигрыш и рассчитывают получить компоненты вектора $\bar{Sh} = (\bar{Sh}_1, \dots, \bar{Sh}_n)$, $\bar{Sh}_i = \pi Sh_i$. Развитию игры во времени соответствует движение вдоль некоторого случайного пути \bar{u} , который получается при реализации кооперативного решения $\bar{\eta}(\cdot)$. После первого шага игра переходит в новое состояние, являющееся начальным для подыгры, начинающейся со второго шага, то есть, фактически, игроки попадают в новую стохастическую игру, которая является подыгрой игры \bar{G} . Для сохранения кооперации на этом шаге игроки должны ожидать получение выигрышей в соответствии с вектором Шепли, рассчитанным для этой

подыгры (фактически, рассчитанным для одной из подыгр $\{\bar{G}^1, \dots, \bar{G}^t\}$). К сожалению, осуществляя выплаты на каждом шаге игры в соответствии с выигрышами в игровых элементах, реализовавшихся на этих шагах, невозможно добиться того, чтобы оставшиеся выплаты представляли собой компоненты вектора Шепли для подыгры, начинающейся с данного шага. Это и есть проявление позиционной несостоятельности вектора Шепли. Требуется перераспределять выигрыши игроков в каждом игровом элементе, чтобы позиционная несостоятельность вектора Шепли была преодолена.

Определение. [Baranova, Petrosjan, 2006]. Вектор Шепли $\bar{Sh} = (\bar{Sh}_1, \dots, \bar{Sh}_n)$, $\bar{Sh}_i = \pi Sh_i$ ($i \in N$) назовем *позиционно состоятельным* в марковской игре \bar{G} , если для каждого игрового элемента Γ^j ($j = 1, \dots, t$) существует неотрицательная ПРД $\beta^j = (\beta_1^j, \dots, \beta_n^j)$ такая, что

$$Sh_i = \beta_i + (1 - q)\Pi(\bar{\eta}(\cdot))Sh_i \quad (4.15.16)$$

для любого $i \in N$. Здесь $Sh_i = (Sh_i^1, \dots, Sh_i^t)$, $\beta_i = (\beta_i^1, \dots, \beta_i^t)$ и β_i^j — это i -й элемент кооперативной процедуры распределения дележа для игрового элемента Γ^j , $\Pi(\bar{\eta}(\cdot))$ — матрица вероятностей перехода, построенная согласно (4.15.3), $\bar{\eta}(\cdot)$ — кооперативное решение, удовлетворяющее условию (4.15.7).

Если вектор Шепли позиционно состоятелен, то осуществляя на каждом шаге пути выплаты игрокам в соответствии с их ПРД, можно добиться того, чтобы эти выплаты были неотрицательными и чтобы математическое ожидание вектора Шепли, вычисленного для кооперативной марковской подыгры \bar{G}^j , ($j = 1, \dots, t$), совпадало бы с математическим ожиданием выигрышей, которые игрокам осталось получить в подыгре \bar{G}^j марковской игры \bar{G} .

Если $\det(E - (1 - q)\Pi(\bar{\eta}(\cdot))) \neq 0$, то формула (4.15.16) примет вид

$$Sh_i = (E - (1 - q)\Pi(\bar{\eta}(\cdot)))^{-1}\beta_i. \quad (4.15.17)$$

Это уравнение имеет единственное решение относительно β_i , если $\det(E - (1 - q)\Pi(\bar{\eta}(\cdot))) \neq 0$. Получаем формулу для вычисления величин выплат β_i i -му игроку, $i \in N$:

$$\beta_i = (E - (1 - q)\Pi(\bar{\eta}(\cdot)))Sh_i. \quad (4.15.18)$$

В общем случае невозможно гарантировать неотрицательность элементов вектора $\beta_i = (\beta_i^1, \dots, \beta_i^t)$, таким образом невозможно гарантировать позиционную состоятельность вектора Шепли \bar{Sh} в кооперативной стохастической игре \bar{G} .

Лемма. Имеют место равенства $B_i = Sh_i$, $\bar{B}_i = \bar{Sh}_i$ для всех $i \in N$.

Доказательство аналогично доказательству леммы п. 4.14.7.

Лемма п. 4.15.9 говорит о том, что математическое ожидание сумм β_i , рассчитанных вдоль пути $u^{\{j\}}$ для подыгры \bar{G}^j , $j = 1, \dots, t$, равно B_i^j , где β_i представляет собой вектор выплат игроку i , производимых вдоль реализовавшегося пути $u^{\{j\}}$ подыгры \bar{G}^j , когда игроки придерживаются кооперативного решения $\bar{\eta}(\cdot)$, равно математическому ожиданию выигрыша i -го игрока в этой подыгре (то есть i -ой компоненте вектора Шепли Sh^j). Таким образом, представлен конструктивный способ построения реальных выплат игрокам на каждом шаге игры, причем, исходя из леммы § 4.15., можно утверждать, что игроки заинтересованы в перераспределении своих выигрышей, так как, получая $\beta_i^1, \dots, \beta_i^t$ в игровых элементах $\Gamma^1, \dots, \Gamma^t$ соответственно, игрок i в игре \bar{G} получит столько же (с точки зрения математического ожидания), сколько и планировал

получить в начале игры (то есть \overline{Sh}_i), и оставшиеся выплаты будут соответствовать тому же «принципу оптимальности» (в нашем случае, вектору Шепли). Это означает, что на каждом шаге игры \overline{G} оставшиеся выплаты будут рассчитаны по тем же «правилам», что и в начале игры (в нашем случае, по аксиомам Шепли).

4.15.10. Пример 19. Рассмотрим кооперативную игру передачи данных в беспроводных сетях [Sagduyu, Ephremides, 2006; Michiardi et al., 2003; Srinivasan et al., 2003]. Беспроводная сеть представлена на рис. 4.19.

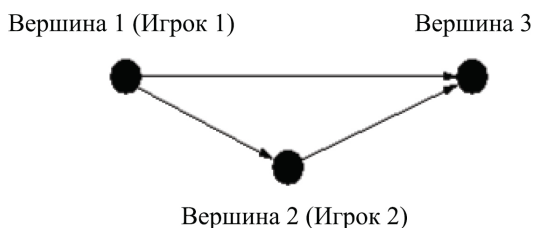


Рис. 4.19. Простая схема беспроводной сети

Рассмотрим систему, в которой приемники (вершины 1 и 2) независимо генерируют пакеты данных на каждом промежутке времени с вероятностями a_1 и a_2 соответственно. Пакет данных может появиться в вершине 1 (2) с вероятностью a_1 (a_2) только, если в конце предыдущего временного промежутка очередь в вершине 1 (2) пуста. Сделаем некоторые предположения:

- 1) вершины 1 и 2 (игроки 1 и 2 соответственно) стремятся послать пакеты данных, скопившихся у них, в конечный пункт назначения - вершину 3;
- 2) максимальная емкость буфера каждой вершины равна единице. Вершина 3 не может принять одновременно два пакета данных в один промежуток времени. В данной постановке исключается многопакетная передача данных, а также исключается одновременное отправленное и принятие пакетов никакой из вершин в любой промежуток времени;
- 3) если игроки одновременно пересылают пакеты в вершину 3, то эти пакеты отклоняются и возвращаются в начальные вершины, таким образом, в следующий промежуток времени ни один новый пакет не может появиться в вершинах 1 и 2;
- 4) все пересылаемые пакеты данных имеют одинаковый размер, и доставка одного пакета данных из одной вершины в другую, которые имеют прямое соединение друг с другом, занимает один промежуток времени;
- 5) игрок 1 (вершина 1) выбирает одну из двух стратегий: переслать пакет данных напрямую в вершину 3 или переслать этот пакет в вершину 2, чтобы тот послал этот пакет в вершину 3 в следующий промежуток времени;
- 6) если игрок 1 (вершина 1) пересылает пакет данных игроку 2, который уже имеет в данный промежуток времени пакет в своей очереди, игрок 2 отклоняет переданный ему пакет. В противном случае игрок 2 решает принять или отклонить пакет, переданный игроком 1.

Предположим, что в вышеописанной системе передачи данных введена следующая схема поощрений и наказаний:

1. величина $f \geq 0$ — это премия, которую получает игрок 1 или 2 за каждую успешную передачу одного пакета данных в вершину 3;

2. игрок 1 получает премию в размере $c \geq 0$ от игрока 2 за передачу одного пакета данных игроку 2, который, в свою очередь, может рассчитывать на премию размером f только после успешной передачи этого пакета в конечный пункт (вершину 3) в следующий промежуток времени;

3. задержка пакета данных в вершине 1 или 2 на один промежуток времени приносит игроку, находящемуся в этой вершине, издержки в размере $d \geq 0$, независимо от того, по какой причине произошла задержка;

4. величина D_{ij} — это издержки по пересылке одного пакета данных из вершины i в вершину j , которые несет игрок i .

Процесс передачи данных может остановиться в любой промежуток времени с вероятностью $0 < q < 1$. Вероятность q , по сути, является дисконт-фактором. Модель передачи данных в беспроводных сетях может быть представлена марковской игрой. Игроки, находящиеся в вершинах 1 и 2, стремятся максимизировать ожидаемый суммарный выигрыш с последующим разделом этого выигрыша с помощью вектора Шепли.

Обозначим через (Q_1, Q_2) состояние в беспроводной сети, где Q_i — это число пакетов данных, находящееся в очереди игрока $i = 1, 2$. Число Q_i может принимать значения 0 или 1, если ни одного или один пакет данных находится в данный промежуток времени в очереди игрока i , соответственно.

В марковской игре передачи данных в беспроводных сетях возможно 4 состояния:

$$T = \{(0, 0); (0, 1); (1, 0); (1, 1)\}.$$

Предположим, что игроки имеют информацию о состоянии не только своей очереди, но и очереди другого игрока. Это предположение разумно, поскольку мы пытаемся найти кооперативное решение, которое подразумевает совместные действия, включая обмен информацией о состоянии очередей обоих игроков.

Определим, исходя из условия задачи игровые элементы во всех состояниях системы.

1. Игровой элемент $\Gamma(0, 0)$.

У игрока 1 имеется одна стратегия W (ожидать), у игрока 2 — одна стратегия W (ожидать). Выигрыши игроков будут $(0, 0)$.

2. Игровой элемент $\Gamma(0, 1)$.

У игрока 1 имеется одна стратегия W (ожидать), у игрока 2 — одна стратегия $\xrightarrow{3}$ (послать пакет в вершину 3). Выигрыши игроков будут следующими: $(0, f - D_{23})$.

3. Игровой элемент $\Gamma(1, 0)$.

Игрок 1 имеет две стратегии: 1) $\xrightarrow{3}$ (послать пакет в вершину 3), 2) $\xrightarrow{2}$ (послать пакет в вершину 2); игрок 2 имеет две стратегии: 1) Ac (принять пакет от игрока 1), 2) Rej (не принять пакет от игрока 1). Выигрыши игроков будут следующими:

$$\begin{pmatrix} (f - D_{13}, 0) & (f - D_{13}, 0) \\ (c - D_{12}, -c) & (-d - D_{12}, 0) \end{pmatrix}.$$

4. Игровой элемент $\Gamma(1, 1)$.

Игрок 1 имеет две стратегии: 1) $\xrightarrow{3}$ (послать пакет в вершину 3), 2) W (ожидать); игрок 2 имеет две стратегии: 1) $\xrightarrow{3}$ (послать пакет в вершину 3), 2) W (ожидать). Выигрыши игроков будут следующими:

$$\begin{pmatrix} (-d - D_{13}, -d - D_{23}) & (f - D_{13}, -d) \\ (-d, f - D_{23}) & (-d, -d) \end{pmatrix}.$$

Без потери общности прибавим число

$$z = \max\{0, f - D_{13}, f - D_{23}, c - D_{12}, -c\}$$

ко всем выигрышам игроков во всех игровых элементах, чтобы сделать все выигрыши неотрицательными.

Будем решать описанную выше марковскую игру в классе стационарных стратегий. Обозначим через X_i множество смешанных стационарных стратегий игрока i , $i = 1, 2$.

В соответствии со структурой марковской игры передачи данных в беспроводной сети смешанная стационарная стратегия игрока 1 диктует ему выбрать стратегию W с вероятностью 1 в состояниях $(0, 0)$, $(0, 1)$, стратегию $\xrightarrow{3}$ с вероятностью p_{11} в состоянии $(1, 0)$, и стратегию $\xrightarrow{3}$ с вероятностью p_{12} в состоянии $(1, 1)$. Смешанная стационарная стратегия игрока 2 диктует ему выбрать стратегию W с вероятностью 1 в состояниях $(0, 0)$, стратегию $\xrightarrow{3}$ в состоянии $(0, 1)$, стратегию Ac с вероятностью p_{21} в состоянии $(1, 0)$, и стратегию $\xrightarrow{3}$ с вероятностью p_{22} в состоянии $(1, 1)$.

Обозначим через $u_i = (p_{i1}, p_{i2})$ смешанную стационарную стратегию игрока i и множество смешанных стационарных стратегий игрока i через U_i , $i = 1, 2$. Получаем ситуацию в стационарных стратегиях $u = (u_1, u_2) = (p_{11}, p_{12}, p_{21}, p_{22})$. Матрица переходных вероятностей в ситуации в стационарных стратегиях u будет следующей:

$$\Pi(u) = \begin{pmatrix} \alpha_{11} & \alpha_{12} & \alpha_{13} & \alpha_{14} \\ \alpha_{21} & \alpha_{22} & \alpha_{23} & \alpha_{24} \\ \alpha_{31} & \alpha_{32} & \alpha_{33} & \alpha_{34} \\ \alpha_{41} & \alpha_{42} & \alpha_{43} & \alpha_{44} \end{pmatrix},$$

где

$$\begin{aligned} \alpha_{11} &= (1 - a_1)(1 - a_2), \\ \alpha_{12} &= (1 - a_1)a_2, \\ \alpha_{13} &= a_1(1 - a_2), \\ \alpha_{14} &= a_1a_2, \\ \alpha_{21} &= (1 - a_1)(1 - a_2), \\ \alpha_{22} &= (1 - a_1)a_2, \\ \alpha_{23} &= a_1(1 - a_2), \\ \alpha_{24} &= a_1a_2, \\ \alpha_{31} &= p_{11}(1 - a_1)(1 - a_2), \\ \alpha_{32} &= p_{11}(1 - a_1)a_2 + (1 - p_{11})p_{21}(1 - a_1), \\ \alpha_{33} &= p_{11}a_1(1 - a_2) + (1 - p_{11})(1 - p_{21})(1 - a_2), \\ \alpha_{34} &= p_{11}a_1a_2 + (1 - p_{11})p_{21}a_1 + (1 - p_{11})(1 - p_{21})a_2, \\ \alpha_{41} &= 0, \\ \alpha_{42} &= p_{12}(1 - p_{22})(1 - a_1), \\ \alpha_{43} &= (1 - p_{12})p_{22}(1 - a_2), \\ \alpha_{44} &= p_{12}p_{22} + (1 - p_{12})(1 - p_{22}) + p_{12}(1 - p_{22})a_1 + (1 - p_{12})p_{22}a_2. \end{aligned}$$

Если в описанной выше марковской игре реализуется ситуация в стационарных стратегиях u , выигрыши игрока 1 в соответствующих состояниях будут следующими:

$$K_1(u) = \begin{pmatrix} z \\ z \\ p_{11}(z + f - D_{13}) + (1 - p_{11})p_{21}(z + c - D_{12}) + \\ + (1 - p_{11})(1 - p_{21})(z - d - D_{12}) \\ p_{12}p_{22}(z - d - D_{13}) + \\ + p_{12}(1 - p_{22})(z + f - D_{13}) + (1 - p_{12})(z - d) \end{pmatrix},$$

у игрока 2 будут следующими:

$$K_2(u) = \begin{pmatrix} z \\ z + f - D_{23} \\ (1 - p_{11})p_{21}(z - c) \\ p_{12}p_{22}(z - d - D_{23}) + \\ + (1 - p_{12})p_{22}(z + f - D_{23}) + (1 - p_{22})(z - d) \end{pmatrix}.$$

Кооперативную игру передачи данных в беспроводной сети будем рассматривать в классе чистых стационарных стратегий. Обозначим через Ξ_i множество чистых стационарных стратегий игрока i , $i = 1, 2$. Например, чистая стационарная стратегия $\eta_1 = (1, 0)$ игрока 1 диктует ему выбирать стратегию $\xrightarrow{3}$ в состоянии $(1, 0)$ и стратегию W в состоянии $(1, 1)$. Каждый игрок имеет 4 чистых стационарных стратегии, т. е. мы получаем 16 ситуаций в чистых стационарных стратегиях. Для ситуации в чистых стационарных стратегиях $\eta = (\eta_1, \eta_2)$ можно записать в упрощенном виде матрицу переходных вероятностей $\Pi(\eta)$.

Например, для ситуации $\eta^1 = (1, 1, 1, 1)$ матрица переходных вероятностей будет иметь вид:

$$\Pi(\eta^1) = \begin{pmatrix} (1 - a_1)(1 - a_2) & (1 - a_1)a_2 & a_1(1 - a_2) & a_1a_2 \\ (1 - a_1)(1 - a_2) & (1 - a_1)a_2 & a_1(1 - a_2) & a_1a_2 \\ (1 - a_1)(1 - a_2) & (1 - a_1)a_2 & a_1(1 - a_2) & a_1a_2 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Для каждой ситуации в чистых стационарных стратегиях $\eta \in \Xi = \prod_{i=1}^2 \Xi_i$ мы можем посчитать математическое ожидание выигрышей игроков для каждой подыгры, начинающейся с определенного состояния:

$$E_i = (E_i^{(0,0)}, E_i^{(0,1)}, E_i^{(1,0)}, E_i^{(1,1)}).$$

Математическое ожидание выигрышей игрока i для подыгр можно рассчитать по следующей формуле:

$$E_i(\eta) = (E - (1 - q)\Pi(\eta))^{-1}K_i(\eta),$$

где $K_i(\eta)$, $\Pi(\eta)$ определены выше.

Математическое ожидание выигрыша игрока i во всей марковской игре, включая ход «случая» (т. е. выбор начального состояния) может быть рассчитано по формуле:

$$\bar{E}_i(\eta) = \pi E_i(\eta),$$

где $\pi = (\pi_{(0,0)}, \pi_{(0,1)}, \pi_{(1,0)}, \pi_{(1,1)})$ — вектор начальных вероятностей, и π_k — вероятность того, что первое состояние в марковской игре будет $k \in T$.

Для кооперативного решения этой марковской игры необходимо найти кооперативной решение $\bar{\eta} = (\bar{\eta}_1, \bar{\eta}_2)$, т. е. ситуацию в чистых стационарных стратегиях такую, что

$$\sum_{i \in \{1,2\}} \bar{E}_i(\bar{\eta}) = \max_{\eta \in \Xi} \sum_{i \in \{1,2\}} \bar{E}_i(\eta).$$

Значение характеристических функций для подыгр

$$V(S) = (V^{(0,0)}(S), V^{(0,1)}(S), V^{(1,0)}(S), V^{(1,1)}(S))$$

можно рассчитать по формуле (4.15.10), а для всей марковской игры — по формуле (4.15.11).

Кооперативной марковской игрой передачи данных в беспроводной сети будет пара $\langle \{1, 2\}, \bar{V}(S) \rangle$, $\bar{V}(\emptyset) = 0$ и $\bar{V}(\{1, 2\}) = \sum_{i \in \{1,2\}} \bar{E}_i(\bar{\eta})$.

В качестве дележа максимального суммарного математического ожидания выигрыша игроков рассмотрим вектор Шепли. Обозначим через $Sh = (Sh_1, Sh_2)$, где

$$Sh_i = (Sh_i^{(0,0)}, Sh_i^{(0,1)}, Sh_i^{(1,0)}, Sh_i^{(1,1)})$$

вектор Шепли, рассчитанный для подыгр и через $\bar{Sh} = (\bar{Sh}_1, \bar{Sh}_2)$ вектор Шепли, рассчитанный для всей марковской игры.

Игроки перед началом игры договариваются о кооперации и ожидают получить совместный выигрыш $\bar{V}(\{1, 2\})$ и соответствующие компоненты \bar{Sh}_1 и \bar{Sh}_2 вектора Шепли.

Было бы естественно, если выплаты игрокам в игровых элементах, соответствующих состояниям марковской игры, были бы равны выигрышам игроков в одновременных играх, что эквивалентно условию:

$$Sh_i = K_i(\bar{\eta}) + (1 - q)\Pi(\bar{\eta})Sh_i.$$

Но это условие не всегда выполняется, поэтому используем кооперативную процедуру распределения дележа

$$\beta_i = (\beta_i^{(0,0)}, \beta_i^{(0,1)}, \beta_i^{(1,0)}, \beta_i^{(1,1)}).$$

Это вектор реальных выплат игроку i в состояниях $(0, 0)$, $(0, 1)$, $(1, 0)$, $(1, 1)$, соответственно, такой, что

$$\sum_{i=1}^2 \beta_i^k = \sum_{i \in N} K_i^k(\bar{\eta})$$

для любого $k \in \{(0, 0), (0, 1), (1, 0), (1, 1)\}$. Посчитаем вектор β_i по формуле (4.15.18).

Можно потребовать, чтобы выплаты i -му игроку β_i^k были неотрицательными для любого состояния $k \in K$ и любого игрока $i \in N$, что эквивалентно тому, чтобы система уравнений относительно $\beta_i = (\beta_i^{(0,0)}, \beta_i^{(0,1)}, \beta_i^{(1,0)}, \beta_i^{(1,1)})$

$$Sh_i = (E - (1 - q)\Pi(\bar{\eta}(\cdot)))^{-1}\beta_i$$

имела бы неотрицательное решение. Тем самым мы проверяем условие позиционной состоятельности вектора Шепли.

Далее рассмотрим численный пример кооперативной игры передачи данных в беспроводной сети, поскольку в общем виде представить расчеты матрицы $(E - (1 - q)\Pi(\bar{\eta}))^{-1}$ в книге не имеется возможности. Пусть параметры игры принимают следующие значения:

$$\begin{aligned} a_1 &= 0.5, & q &= 0.01, \\ a_2 &= 0.1, & f &= 1, \\ D_{12} &= 0.1, & d &= 0.1, \\ D_{13} &= 0.6, & c &= 0.3, \\ D_{23} &= 0.2, & \pi &= \left(\frac{1}{4}, \frac{1}{4}, \frac{1}{4}, \frac{1}{4}\right) \end{aligned}$$

Таблица 4.1 показывает для каждой ситуации в чистых стационарных стратегиях η рассчитанные значения математических ожиданий. Здесь

$E_1(\eta) = (E_1^{(0,0)}(\eta), E_1^{(0,1)}(\eta), E_1^{(1,0)}(\eta), E_1^{(1,1)}(\eta))$ — вектор математических ожиданий выигрышей игрока 1 в подыграх;

$E_2(\eta) = (E_2^{(0,0)}(\eta), E_2^{(0,1)}(\eta), E_2^{(1,0)}(\eta), E_2^{(1,1)}(\eta))$ — вектор математических ожиданий выигрышей игрока 2 в подыграх;

$\pi \sum_{i \in \{1,2\}} \bar{E}_i(\eta)$ — математическое ожидание суммарного выигрыша игроков.

Таблица 4.1. Ожидаемые значения

η	$E_1(\eta)$	$E_2(\eta)$	$\pi(E_1(\eta) + E_2(\eta))$
$\eta^1 = (1, 1, 1, 1)$	14.75966387	45.70756302	55.65042017
	14.75966387	46.50756302	
	15.15966387	45.70756302	
	0	40.	
$\eta^2 = (1, 1, 1, 0)$	89.80000000	76.24887286	166.56447240
	89.80000000	77.04887286	
	90.20000000	76.24887286	
	90.20000000	76.71127141	
$\eta^3 = (1, 0, 1, 1)$	88.22724883	77.92000000	166.66781694
	88.22724883	78.72000000	
	88.62724883	77.92000000	
	88.30952131	78.72000000	
$\eta^4 = (1, 0, 1, 0)$	64.67563026	62.34621849	125.56638655
	64.67563026	63.14621849	
	65.07563026	62.34621849	
	60.	60.	
$\eta^5 = (1, 1, 0, 1)$	14.75966387	45.70756302	55.65042017
	14.75966387	46.50756302	
	15.15966387	45.70756302	
	0.	40.	
$\eta^6 = (1, 1, 0, 0)$	89.80000000	76.24887286	166.56447250
	89.80000000	77.04887286	
	90.20000000	76.24887286	
	90.20000000	76.71127141	

Окончание таблицы 4.1

η	$E_1(\eta)$	$E_2(\eta)$	$\pi(E_1(\eta) + E_2(\eta))$
$\eta^7 = (1, 0, 0, 1)$	88.22724883 88.22724883 88.62724883 88.30952131	77.92000000 78.72000000 77.92000000 78.72000000	166.66781694
$\eta^8 = (1, 0, 0, 0)$	64.67563026 64.67563026 65.07563026 60.	62.34621849 63.14621849 62.34621849 60.	125.56638655
$\eta^9 = (0, 1, 1, 1)$	3.870077599 3.870077599 2.815688411 0.	41.81391498 42.61391498 41.29388792 40.	44.06939037
$\eta^{10} = (0, 1, 1, 0)$	85.30045000 85.30045000 85.58955000 85.78955000	82.15225000 82.95225000 82.29775000 82.49775000	167.97000000
$\eta^{11} = (0, 0, 1, 1)$	75.28276807 75.28276807 75.40659007 75.23559576	93.56491025 94.36491025 93.89870345 94.52135936	169.38940133
$\eta^{12} = (0, 0, 1, 0)$	60.82133034 60.82133034 60.70655852 60.	60.79766590 61.59766590 60.59084462 60.	121.33384890
$\eta^{13} = (0, 1, 0, 1)$	7.05150332 7.05150332 6.42201835 0	43.10048870 43.90048870 42.75229358 40.	47.56957398
$\eta^{14} = (0, 1, 0, 0)$	76.00869440 76.00869440 76.06878134 76.68178956	75.10793102 75.90793102 75.07981035 75.59292249	151.61413864
$\eta^{15} = (0, 0, 0, 1)$	69.01980198 69.01980198 69.01000000 68.91000000	77.92000001 78.72000001 77.92000000 78.72000000	147.30990100
$\eta^{16} = (0, 0, 0, 0)$	61.00735762 61.00735762 60.91743119 60.	61.08577346 61.88577346 60.91743119 60.	121.70528113

В примере 19 п. 4.15.10 кооперативным решением будет ситуация

$$\eta_{11} = ((W, W, \xrightarrow{2}, W), (W, \xrightarrow{3}, Ac, \xrightarrow{3})).$$

Максимум математического ожидания суммарного выигрыша игроков в этой марковской игре будет следующим:

$$\max_{\eta \in \Xi} \sum_{i \in N} \bar{E}_i(\eta) = 169.39.$$

Значения характеристических функций для подыгр будут следующими:

$$V(\{1\}) = (64.68, 64.68, 65.08, 60),$$

$$V(\{2\}) = (61.09, 61.89, 60.92, 60),$$

$$V(\{1, 2\}) = (168.85, 169.65, 169.31, 169.76),$$

и для всей марковской игры:

$$\bar{V}(\{1\}) = 63.61,$$

$$\bar{V}(\{2\}) = 60.97,$$

$$\bar{V}(\{1, 2\}) = 169.39.$$

Компоненты вектора Шепли, рассчитанного для подыгр марковской игры передачи данных, будут следующими:

$$Sh_1 = (86.22, 86.22, 86.73, 84.88),$$

$$Sh_2 = (82.63, 83.43, 82.57, 84.88).$$

Компоненты вектора Шепли, рассчитанного для всей марковской игры передачи данных, будут иметь значения:

$$\bar{Sh}_1 = 86.01,$$

$$\bar{Sh}_2 = 83.38.$$

Процедуры распределения дележа для игроков будут иметь следующие значения:

$$\beta_1 = (0.7, 0.7, 2.04, -0.8),$$

$$\beta_2 = (0.7, 1.5, -0.74, 2.9).$$

Если игроки в каждый промежуток времени хотят в оставшейся части игры передачи данных в беспроводной сети получать компоненты вектора Шепли, то выплаты игрокам должны быть

- в состоянии $(1, 0)$:
2.04 игроку 1 вместо 0.9,
-0.74 игроку 2 вместо 0.4,
- в состоянии $(1, 1)$:
-0.8 игроку 1 вместо 0.6,
2.9 игроку 2 вместо 1.5.

В нашем численном примере марковской игры передачи данных вектор Шепли $\bar{Sh} = (86.01, 83.38)$ не является позиционно состоятельным.

§ 4.16. Упражнения и задачи

1. Найти все ситуации абсолютного равновесия по Нэшу в примере п.4.2.1.
2. Доказать, что в неантагонистической конечношаговой игре двух лиц с полной информацией выигрыши во всех «благожелательных» («неблагожелательных») ситуациях равновесия по Нэшу равны между собой.
3. Пусть $v_1(x), v_2(x), \dots, v_n(x)$ — значения функций выигрыша игроков $1, 2, \dots, n$ в подыгре Γ_x в ситуации абсолютного равновесия в игре Γ .
 - а) Показать, что функции $v_i(x)$, $i = 1, 2, \dots, n$, удовлетворяют следующей системе функциональных уравнений:

$$v_i(x) = \max_{x' \in \Gamma_x} v_i(x'), \quad x \in X_i, \quad i = 1, 2, \dots, n, \quad (4.16.1)$$

при граничном условии

$$v_i(x)|_{x \in X_{n+1}} = H_i(x). \quad (4.16.2)$$

- б) Привести пример игры, в которой выигрыши игроков в ситуации равновесия в стратегиях наказания не удовлетворяют системе функциональных уравнений (4.16.1) при граничном условии (4.16.2).

4. Построить пример неантагонистической многошаговой игры двух лиц, в которой в ситуации равновесия в «стратегиях наказания» наказывающий игрок при наказании противника за отклонение от выбранного пути еще сильнее наказывает самого себя.

5. Построить Парето-оптимальные множества в игре из примера п.4.2.1.

6. Построить пример многошаговой неантагонистической игры, в которой ни одна из ситуаций равновесия по Нэшу не приводит к Парето-оптимальному решению.

7. Построить отображение T , которое каждой подыгре Γ_z игры Γ ставит в соответствие некоторое подмножество ситуаций U_z в этой подыгре. Пусть $T(\Gamma) = U_{x_0}$.

Будем говорить, что отображение T динамически устойчиво, если из $u(\cdot) \in U_{x_0}$ следует, что $u^{z_k}(\cdot) \in U_{z_k}$, где $u^{z_k}(\cdot) = (u_1^{z_k}(\cdot), \dots, u_n^{z_k}(\cdot))$ — сужение ситуации $u(\cdot)$ на подыгру Γ_{z_k} , $\omega_0 = \{x_0, z_1, \dots, z_k\}$ — партия, реализовавшаяся в ситуации $u(\cdot) \in U_{x_0}$.

Показать, что если отображения T каждой подыгре Γ_{z_k} ставит в соответствие множество Парето-оптимальных ситуаций U_x^P , то оно динамически устойчиво.

8. Отображение T , определенное в упр. 7, называется сильно динамически устойчивым, если для любой ситуации $u(\cdot) \in U_{x_0}$, любого $z_k \in \{z_i\} = \omega$, где $\{z_i\} = \omega$ — партия в ситуации $u(\cdot)$, ситуации $\hat{u}^{z_k}(\cdot) \in U_{z_k}$ существует ситуация $\hat{u}(\cdot) \in U_{x_0}$, для которой ситуация $\hat{u}^{z_k}(\cdot)$ является ее сужением на позициях игры Γ_{z_k} и позиция z_k возможна в ситуации $\hat{u}(\cdot)$. Показать, что если отображение T каждой подыгре Γ_{z_k} ставит в соответствие множество ситуаций равновесия по Нэшу, то оно сильно динамически устойчиво.

9. Построить пример, когда отображение T , ставящее в соответствие каждой подыгре Γ_z множество Парето-оптимальных ситуаций равновесия, сильно динамическим не является.

10. Для каждой подыгры Γ_z введем в рассмотрение величины $v(\{i\}, z)$, $i = 1, \dots, n$, представляющие собой гарантированный выигрыш i -го игрока в подыгре Γ_z , т. е. $v(\{i\}, z)$ — значение антагонистической игры, построенной на графе подыгры Γ_z между игроком i и игроками $N \setminus i$, действующими как один игрок. При этом множество стратегий коалиции игроков $N \setminus i$ есть декартово произведение множества стратегий каждого из игроков $k \in \{N, i\}$, $u_{\{N, i\}} \in \prod_{k \in \{N, i\}} u_i$, функция выигрыша игрока i в ситуации $(u_i, u_{N \setminus i})$ определяется как $H_i^z(u_i, u_{N \setminus i})$, а функция выигрыша коалиции $N \setminus i$ полагается равной $-H_i^z(u_i, u_{N \setminus i})$.

11. Показать, что если в некоторой многошаговой неантагонистической игре Γ с неотрицательными выигрышами ($H_i \geq 0, i = 1, \dots, n$), $v(\{i\}, z) = 0$ для всех $i = 1, \dots, n$, и $z \in \bigcup_{i=1}^n X_i$, то любая партия может быть реализована в некоторой ситуации равновесия в стратегиях наказания.

12. Формализовать k -уровневую древовидную систему управления в виде иерархической игры, в которой управляющий центр, находящийся на i -м уровне ($i = 1, \dots, k-1$), распределяет ресурсы между подчиненными ему управляющими центрами следующего уровня при $i < k-1$ и между подчиненными ему производственными подразделениями при $i = k-1$. Выигрыш каждого производственного подразделения зависит только от своего производства, а выигрыш управляющих центров — от подчиненных им производственных подразделений.

13. Найти ситуацию равновесия по Нэшу в построенной в упр. 12 k -уровневой иерархической древовидной игре.

14. Показать, что вектор выигрышей $\alpha = \{v(\bar{N}), 0, \dots, 0\}$ принадлежит C -ядру двухуровневой иерархической древовидной игры с характеристической функцией $\bar{v}(S)$. Показать, что ситуация равновесия, построенная в двухуровневой древовидной иерархической игре, является также ситуацией сильного равновесия.

15. В ромбовидной иерархической игре построить характеристическую функцию, используя ситуацию равновесия по Нэшу.

16. Описать множество всех ситуаций равновесия по Нэшу в двухуровневой древовидной иерархической игре. Учесть возможность «наказания» центра A_0 игроками B_1, \dots, B_n (например, прекращение выпуска продукции при распределении ресурсов, не отвечающих интересам игрока i).

17. Построить матрицу выигрышей игроков в игре примера 6 п. 4.7.1. Найти оптимальные чистые стратегии и значение получившейся матричной игры.

18. Привести к матричной форме и решить игру из примера 8 п. 4.7.1.

19. Рассмотрим следующую антагонистическую многошаговую игру с задержкой информации о местоположении одного из игроков. Имеются два игрока: мишень E и стрелок P . Мишень может двигаться только по точкам оси Ox с координатами $0, 1, 2, \dots$, причем если игрок E находится в точке i , то в следующий момент времени он может переместиться только в точки $i+1$, $i-1$ или остаться на месте. Стрелок P имеет j патронов, $j = 0, 1, \dots$, и может производить не более одного выстрела в каждый момент времени. Считается, что стрелок попадает в ту точку, в которую целится.

В каждый момент времени игрок P знает только точное местоположение игрока E на предыдущем шаге, т. е. если E находился на предыдущем шаге в точке i , то игроку P необходимо целиться в точки $i+1$, i и $i-1$. Игрок E знает количество патронов, которые имеет игрок P в каждый момент времени, но не знает, куда целится игрок P . Выигрыш стрелка P равен числу попаданий в мишень. Таким образом, цель стрелка P — максимизировать количество попаданий в мишень E до того, как она достигнет «бункера». Цель мишени противоположна. Здесь под «бункером» понимается точка O , в которой мишень недостижима для стрелка P .

Обозначим символом $\Gamma_{(i,j)}$ описанную выше игру при условии, что мишень E в начальный момент времени находилась в точке с координатой i , а стрелок P имел j патронов. Символом $v(i, j)$ обозначим значение игры $\Gamma_{(i,j)}$ (если оно существует). Нетрудно заметить, что $v(i, 0) = 0$, $i = 1, 2, \dots$, $v(1, j) = 0$, $j = 1, 2, \dots$. На каждом шаге игры $\Gamma_{i,j}$, $i = 2, 3, \dots$, $j = 1, 2, \dots$, стрелок имеет четыре стратегии (на самом деле больше, но они неразумны), а игрок E — три стратегии. Стратегии стрелка P таковы: выстрелить в точку $i-1$, выстрелить в точку i , выстрелить в точку $i+1$, не стрелять на данном шаге. Стратегии мишени: передвинуться в точку $i-1$, оставаться в точке i , передвинуться в точку $i+1$. Следовательно, на каждом шаге игры разыгрывается матричная игра с матрицей выигрышей

$$A = \begin{bmatrix} 1 + v(i-1, j-1), & v(i, j-1), & v(i+1, j-1), \\ v(i-1, j-1), & 1 + v(i, j-1), & v(i+1, j-1), \\ v(i-1, j), & v(i, j), & 1 + v(i+1, j-1), \\ v(i-1, j), & v(i, j), & v(i+1, j) \end{bmatrix}.$$

Символами $x_1(i, j)$, $x_2(i, j)$, $x_3(i, j)$, $x_4(i, j)$ обозначим вероятности, с которыми стрелок P использует свои 1, 2, 3 и 4-ю стратегии, а символами $y_1(i, j)$, $y_2(i, j)$, $y_3(i, j)$ — вероятности,

с которыми мишень E использует свою 1, 2, и 3-ю стратегии (стратегии поведения игроков P и E , соответственно, есть функции информационных множеств $\{i, j\}$).

а) Показать, что значение игры $v(i, j)$ и оптимальные стратегии поведения стрелка $P(x_1(i, j), x_2(i, j), x_3(i, j), x_4(i, j))$ и мишени $E(y_1(i, j), y_2(i, j), y_3(i, j))$ связаны между собой следующими соотношениями:

$$\begin{aligned} (1 + v(i-1, j-1))x_1 + v(i-1, j-1)x_2 + v(i-1, j-1)x_3 + v(i-1, j)x_4 &\geq v(i, j), \\ v(i, j-1)x_1 + (1 + v(i, j-1))x_2 + v(i, j-1)x_3 + v(i, j)x_4 &\geq v(i, j), \\ v(i+1, j-1)x_1 + v(i+1, j-1)x_2 + (1 + v(i+1, j-1))x_3 + v(i+1, j)x_4 &\geq v(i, j), \\ x_1 + x_2 + x_3 + x_4 = 1, x_1 \geq 0, x_2 \geq 0, x_3 \geq 0, x_4 \geq 0; \\ (1 + v(i-1, j-1))y_1 + v(i, j-1)y_2 + v(i+1, j-1)y_3 &\leq v(i, j), \\ v(i-1, j-1)y_1 + (1 + v(i, j-1))y_2 + v(i+1, j-1)y_3 &\leq v(i, j), \\ v(i-1, j-1)y_1 + v(i, j-1)y_2 + (1 + v(i+1, j-1))y_3 &\leq v(i, j), \\ v(i-1, j)y_1 + v(i, j)y_2 + v(i+1, j)y_3 &\leq v(i, j), \\ y_1 + y_2 + y_3 = 1, y_1 \leq 0, y_2 \leq 0, y_3 \leq 0. \end{aligned}$$

Указание. Трудность решения этой игры состоит в том, что для определения $v(i, j)$ необходимо знать $v(i+1, j)$, для определения $v(i+1, j)$ необходимо знать $v(i+2, j)$ и т. д. В приводимых ниже упражнениях дано решение игры $\Gamma_{(i, j)}$ и приводятся некоторые его свойства.

б) Пусть $\phi(i, j), i = 1, 2, \dots, j = 0, 1, \dots$ — двойная последовательность, определяемая соотношениями

$$\begin{aligned} \phi(i, 0) = 0, i = 1, 2, \dots; \phi(1, j) = 0, j = 1, 2, \dots, \\ \phi(i, j) = \min\{(1 + \phi(i-1, j-1) + \phi(i, j-1) + \phi(i+1, j-1))/3, \\ (1 + \phi(i-1, j-1) + \phi(i, j-1))/2\}. \end{aligned}$$

1) Доказать, что $v(i, j) = \phi(i, j)$, и если $v(i, j) = (1 + v(i-1, j-1) + v(i, j-1) + v(i+1, j-1))/3$,

то

$$\begin{aligned} x_1(i, j) = v(i, j) - v(i-1, j-1), \quad x_2(i, j) = v(i, j) - v(i, j-1), \\ x_3(i, j) = v(i, j) - v(i+1, j-1), \quad x_4(i, j) = 0, \\ y_1(i, j) = y_2(i, j) = y_3(i, j) = 1/3; \end{aligned}$$

2) Доказать, что $v(i, j) = \phi(i, j)$, и если $v(i, j) = (1 + v(i-1, j-1) + v(i, j-1))/2$, то $x_i(i, j) = v(i, j) - v(i-1, j-1)$, $x_2(i, j) = v(i, j) - v(i, j-1)$, $x_3(i, j) = x_4(i, j) = 0$, $y_1(i, j) = y_2(i, j) = 1/2$; $y_3(i, j) = 0$.

в) Доказать, что при любом $j = 0, 1, 2, \dots$ справедливы следующие соотношения:

- 1) $v(i, j) = j/3$, $i = j+1, j+2, \dots$;
- 2) $v(i, j) \leq v(i+1, j)$, $i = 1, 2, \dots$;
- 3) $v(i, j) \leq v(i, j+1)$, $i = 2, 3, \dots$;
- 4) $v(i, j) + v(i+2, j) \leq 2v(i+1, j)$, $i = 1, 2, \dots$;

г) Доказать, что:

- 1) $\lim_{j \rightarrow +\infty} v(i, j) = j/3$ при любом фиксированном $j = 0, 1, 2, \dots$;
- 2) $\lim_{j \rightarrow -\infty} v(i, j) = i-1$ при любом фиксированном $j = 1, 2, \dots$;

20. Рассмотрим обобщение игры о стрелке и мишени, когда мишень E , занимая положение i , может из него передвинуться максимум на k единиц вправо или влево, т. е. перейти в каждую из следующих точек $i-k, i-k+1, \dots, i, i+1, \dots, i+k$. Остальные цели и возможности стрелка P и мишени E остаются прежними с учетом нового определения стратегии игрока E . Символом $G(i, j)$ обозначим игру при условии, что мишень в начальный момент времени занимает i -ю точку, а стрелок имеет j патронов. Символом $v(i, j)$ обозначим значение игры $G(i, j)$. Из определения $G(i, j)$ имеем

$$v(i, 0) = 0, \quad i = 1, 2, \dots, \quad v(i, j) = 0, \quad i = 1, 2, \dots, \quad k; \quad j = 1, 2, \dots$$

На каждом шаге игры $G(i, j)$, $i = k+1, \dots, j = 1, \dots$ стрелок P имеет $2k+2$ чистые стратегии, а мишень $E(2k+1)$ — чистую стратегию. Чистыми стратегиями игрока P являются: стрельба в точку $i-k$, стрельба в точку $i-k+1, \dots$, стрельба в точку $i+k$, отказ от выстрелов на данном шаге. Стратегиями E являются: перемещение в точку $i-k$, перемещение в точку $i-k+1, \dots$, перемещение в точку $i+k$.

Таким образом, на каждом шаге игры разыгрывается игра с матрицей $\{\alpha_{mn}(i, j)\}$ размера $(2k+2) \times (2k+1)$, где

$$\{\alpha_{mn}(i, j)\} = \begin{bmatrix} 1 + v(i + n - k - 1, j - 1), & \text{если} & m = n = 1, \dots, 2k + 1, \\ v(i + n - k - 1, j - 1), & \text{если} & m \neq n; m, n = 1, \dots, 2k + 1, \\ v(i + n - k - 1, j), & \text{если} & m = 2k + 2, n = 1, \dots, 2k + 1 \end{bmatrix}.$$

а) Показать, что игра $G(i, j)$ имеет значение, равное $v(i, j)$, в том и только в том случае, если существуют $(x_1, x_2, \dots, x_{2k+2}), (y_1, y_2, \dots, y_{2k+1})$ такие, что:

$$\begin{aligned} \sum_{m=1}^{2k+2} a_{mn}(i, j) x_m &\geq v(i, j), \quad n = 1, \dots, 2k + 1, \\ \sum_{m=1}^{2k+2} x_m &= 1, \quad x_m \geq 0, \quad m = 1, \dots, 2k + 2, \\ \sum_{m=1}^{2k+1} a_{mn}(i, j) y_n &\leq v(i, j), \quad m = 1, \dots, 2k + 1, \\ \sum_{n=1}^{2k+1} y_n &= 1, \quad y_n \geq 0, \quad n = 1, \dots, 2k + 1. \end{aligned}$$

Указание. Обозначим символами $x_1(i, j), x_2(i, j), \dots, x_{2k+2}(i, j)$ оптимальные стратегии поведения (если они существуют), с которыми стрелок P использует свои 1-ю, 2-ю, ..., $(2k+2)$ -ю стратегии в информационном состоянии (i, j) , а символами $y_1(i, j), y_2(i, j), \dots, y_{2k+1}(i, j)$ — оптимальные стратегии поведения, с которыми мишень E использует свою 1-ю, 2-ю, ..., $(2k+1)$ -ю стратегии в информационном состоянии (i, j) . В упражнениях ниже приведено решение игры $G(i, j)$ и его свойства.

б) Символом $\phi(i, j), j = 0, 1, \dots; i = 1, 2, \dots$, обозначим следующую двойную последовательность:

$$\begin{aligned} \phi(i, 0) &= 0, \quad i = 1, 2, \dots; \\ \phi(i, j) &= 0, \quad i = 1, 2, \dots, k; \quad j = 1, 2, \dots; \\ \phi(i, j) &= \min_{r=1, \dots, k+1} ((1 + \sum_{t=1}^r k + rv(i + t - k - 1, j - 1)) / (k + 2)), \\ &\quad i = k + 1, k + 2, \dots, \quad j = 1, 2, \dots, \end{aligned} \tag{4.16.3}$$

Доказать, что

1) $v(i, j) = \phi(i, j)$;
 2) при $i = k + 1, \dots; j = 1, 2, \dots$, имеем $x_m(i, j) = v(i, j) - v(i + m - k - 1, j - 1)$ при $m = 1, \dots, k + r^*$, иначе $x_m(i, j) = 0, y_n(i, j) = 1/(k + r^*)$ при $n = 1, \dots, k + r^*$, иначе $y_n = 0$.
 Здесь $r = r^*$ — точка, в которой достигается минимум в (4.16.3).

в) Доказать, что при $j = 0, 1, \dots$:

- 1) $v(i, j) \geq 0, \quad j = 1, 2, \dots;$
- 2) $v(i, j) = j/(2k + 1), \quad i = kj + 1, kj + 2, \dots;$
- 3) $v(i, j) \leq v(i + 1, j), \quad i = 1, 2, \dots;$
- 4) $v(i, j) \leq v(i, j + 1), \quad i = k + 1, k + 2, \dots;$
- 5) $v(i, j + 1) \leq v(i, j) + 1/(2k + 1), \quad i = 1, 2, \dots;$

г) Игра $G(i, \infty)$. Доказать, что $\lim_{j \rightarrow \infty} v(i, j) = w(i)$ при каждом $i = 1, 2, \dots$, где $w(i)$ —

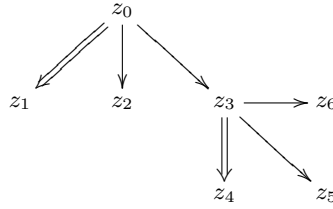
решение линейного разностного уравнения

$$kw(i) - \sum_{p=1}^k (p-1)w(i-p) = 1, \quad i = k + 1, k + 2, \dots$$

с начальными условиями:

$$w(1) = w(2) = \dots = w(k) = 0.$$

21. Задана стохастическая игра $\overline{\overline{G}}(z_0)$ на графе $G(z_0)$, который имеет вид:



Множество вершин графа $G(z_0)$: $Z = \{z_0, \dots, z_6\}$. Множество игроков $N = \{1, 2\}$. В каждой вершине графа $G(z_0)$ задан игровой элемент – одновременная игра двух лиц $\Gamma(z)$, $z \in Z$.

Игровые элементы:

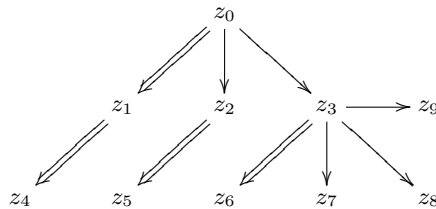
$$\begin{aligned} \Gamma(z_0) &: \begin{pmatrix} (5, 5) & (0, 8) \\ (8, 0) & (1, 1) \end{pmatrix}, & \Gamma(z_2) &: \begin{pmatrix} (3, 0) & (6, 4) \\ (5, 6) & (2, 2) \end{pmatrix}, \\ \Gamma(z_3) &: \begin{pmatrix} (1, 11) & (4, 2) \\ (1, 3) & (1, 1) \end{pmatrix}, & \Gamma(z_4) &: \begin{pmatrix} (1, 1) & (0, 2) \\ (2, 0) & (1, 2) \end{pmatrix}, \\ \Gamma(z_5) &: \begin{pmatrix} (5, 5) & (6, 1) \\ (1, 6) & (6, 6) \end{pmatrix}, & \Gamma(z_6) &: \begin{pmatrix} (4, 2) & (3, 4) \\ (5, 6) & (1, 5) \end{pmatrix}, \\ \Gamma(z_1) &: \begin{pmatrix} (0, 0) & (1, 0) \\ (1, 0) & (0, 1) \end{pmatrix}. \end{aligned}$$

Вероятности перехода определены следующим образом. Если в игровом элементе $\Gamma(z_0)$ реализуется ситуация (2,2), игра $\overline{\overline{G}}(z_0)$ переходит в вершину z_2 с вероятностью $\frac{1}{3}$ и в вершину z_3 с вероятностью $\frac{2}{3}$, если реализуется ситуация, отличная от (2,2) (стрелка \implies означает детерминированный переход), то игра $\overline{\overline{G}}(z_0)$ переходит в вершину z_1 . Если в игре $\Gamma(z_3)$ реализуется ситуация (2,2), то стохастическая игра $\overline{\overline{G}}(z_0)$ переходит в вершины z_5 и z_6 с равными вероятностями $\frac{1}{2}$, из остальных ситуаций совершается переход с вероятностью 1 в вершину z_4 . Вероятности q_k того, что игра закончится на k -м шаге:

$$q_1 = \frac{1}{8}, q_2 = 0, q_3 = 1.$$

Найти вектор Шепли в стохастической игре $\overline{\overline{G}}(z_0)$ и проверить, является ли он позиционно состоятельным. В случае отсутствия позиционной состоятельности провести его регуляризацию.

22. Стохастическая игра $\overline{\overline{G}}(z_0)$ задана на графе $G(z_0)$, который имеет вид:



Множество вершин графа $G(z_0)$: $Z = \{z_0, \dots, z_9\}$. Множество игроков $N = \{1, 2\}$. В каждой вершине графа $G(z_0)$ задан игровой элемент – одновременная игра двух лиц $\Gamma(z)$, $z \in Z$.

Игровые элементы:

$$\begin{aligned} \Gamma(z_0) &: \begin{pmatrix} (6, 6) & (0, 8) \\ (8, 0) & (2, 2) \end{pmatrix}, \quad \Gamma(z_2) : \begin{pmatrix} (3, 0) & (7, 3) \\ (4, 7) & (2, 2) \end{pmatrix}, \\ \Gamma(z_3) &: \begin{pmatrix} (1, 12) & (4, 2) \\ (2, 3) & (1, 1) \end{pmatrix}, \quad \Gamma(z_7) : \begin{pmatrix} (1, 1) & (1, 2) \\ (3, 0) & (1, 2) \end{pmatrix}, \\ \Gamma(z_8) &: \begin{pmatrix} (4, 6) & (6, 1) \\ (1, 6) & (7, 6) \end{pmatrix}, \quad \Gamma(z_9) : \begin{pmatrix} (4, 2) & (3, 4) \\ (6, 6) & (1, 7) \end{pmatrix}, \end{aligned}$$

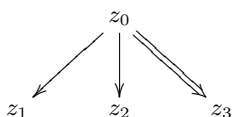
$$\Gamma(z_1), \Gamma(z_4), \Gamma(z_5), \Gamma(z_6) : \begin{pmatrix} (0, 0) & (1, 0) \\ (1, 0) & (0, 1) \end{pmatrix}.$$

Вероятности перехода определены следующим образом. Если в игровом элементе $\Gamma(z_0)$ реализуется ситуация (2,2), то игра $\bar{G}(z_0)$ переходит в вершину z_2 с вероятностью $\frac{1}{4}$ и в вершину z_3 с вероятностью $\frac{3}{4}$, если же реализуется ситуация, отличная от (2,2) (стрелка \Rightarrow означает детерминированный переход), то игра $\bar{G}(z_0)$ переходит в вершину z_1 . В вершинах z_1, z_2 при реализации любой ситуации стохастическая игра $\bar{G}(z_0)$ переходит в вершины z_4 и z_5 соответственно. Если в одновременной игре $\Gamma(z_3)$ реализуется ситуация (2,2), то стохастическая игра $\bar{G}(z_0)$ переходит в вершины z_8 и z_9 с равными вероятностями $\frac{1}{2}$, а если в игре $\Gamma(z_3)$ реализуется ситуация (2,1), то игра с вероятностью 1 перейдет в вершину z_7 , из остальных ситуаций совершается переход в вершину z_6 с вероятностью 1 (стрелка \Rightarrow означает детерминированный переход). Вероятности q_k того, что игра закончится на k -м шаге

$$q_1 = \frac{1}{10}, \quad q_2 = \frac{1}{100}, \quad q_3 = 1.$$

Найти N -ядро в стохастической игре $\bar{G}(z_0)$ и проверить, является ли оно позиционно состоятельным. В случае отсутствия позиционной состоятельности провести его регуляризацию.

23. Стохастическая игра $\bar{G}(z_0)$ задана на графе $G(z_0)$, который имеет вид:



Множество вершин графа $G(z_0)$: $Z = \{z_0, \dots, z_3\}$. Множество игроков $N = \{1, 2\}$. В каждой вершине графа $G(z_0)$ задан игровой элемент – одновременная игра двух лиц $\Gamma(z)$, $z \in Z$.

Игровые элементы:

$$\begin{aligned} \Gamma(z_0), \Gamma(z_1) &: \begin{pmatrix} (10, 3) & (0, 8) \\ (8, 0) & (1, 1) \end{pmatrix}, \quad \Gamma(z_2) : \begin{pmatrix} (2, 0) & (7, 3) \\ (5, 8) & (3, 1) \end{pmatrix}, \\ \Gamma(z_3) &: \begin{pmatrix} (1, 12) & (3, 3) \\ (2, 2) & (1, 0) \end{pmatrix}. \end{aligned}$$

Вероятности перехода определены следующим образом. Если в игровом элементе $\Gamma(z_0)$ реализуется ситуация (2,2), то игра $\bar{G}(z_0)$ переходит в вершину z_1 с вероятностью $\frac{1}{5}$ и в вершину z_2 с вероятностью $\frac{4}{5}$, если же реализуется ситуация, отличная от (2,2) (стрелка \Rightarrow означает детерминированный переход), то игра $\bar{G}(z_0)$ переходит в вершину z_3 . Вероятности q_k того, что игра закончится на k -м шаге

$$q_1 = \frac{1}{10}, \quad q_2 = 1.$$

Найти C -ядро в стохастической игре $\overline{\overline{G}}(z_0)$ и проверить, является ли оно позиционно состоятельным. В случае отсутствия позиционной состоятельности провести его регуляризацию.

24. Марковская игра $\overline{\overline{G}}$ задана следующим образом. Множество игроков $N = \{1, 2\}$. Множество игровых элементов $\{\Gamma^1\}$, где

$$\Gamma^1 : \begin{pmatrix} (5, 5) & (0, 10) \\ (10, 0) & (2, 2) \end{pmatrix}.$$

Первый игрок в одновременной игре Γ^1 имеет две стратегии, т. е. $X_1^1 = \{x_1, x_2\}$, второй игрок — также две стратегии, т. е. $X_2^1 = \{y_1, y_2\}$.

Вероятность окончания игры $\overline{\overline{G}}$ на каждом шаге равна $q = \frac{1}{2}$.

В этом случае марковская игра представляет собой повторяющуюся игру с постоянным дисконтированием. Найти C -ядро и проверить его на позиционную состоятельность. В случае отсутствия позиционной состоятельности провести его регуляризацию.

25. Марковская игра $\overline{\overline{G}}$ задана следующим образом. Множество игроков $N = \{1, 2\}$. Множество игровых элементов $\{\Gamma^1, \Gamma^2\}$, где

$$\Gamma^1 : \begin{pmatrix} (6, 6) & (0, 9) \\ (9, 0) & (2, 2) \end{pmatrix}, \quad \Gamma^2 : \begin{pmatrix} (2, 0) & (7, 4) \\ (6, 5) & (3, 2) \end{pmatrix}.$$

Первый игрок в одновременной игре Γ^1 имеет две стратегии, т. е. $X_1^1 = \{x_1, x_2\}$, а в игре Γ^2 — $X_1^2 = \{a_1, a_2\}$. Второй игрок в одновременной игре Γ^1 имеет две стратегии, т. е. $X_2^1 = \{y_1, y_2\}$, а в игре Γ^2 — $X_2^2 = \{b_1, b_2\}$.

Вектор начального распределения вероятностей $(1; 0)$. Вероятность окончания игры $\overline{\overline{G}}$ на каждом шаге равна $q = \frac{5}{6}$.

Вероятности перехода:

$$\begin{array}{llll} p_1^1(x_1, y_1) = \frac{1}{2}, & p_1^1(x_1, y_2) = \frac{1}{6}, & p_2^1(a_1, b_1) = \frac{1}{3}, & p_2^1(a_1, b_2) = \frac{1}{2}, \\ p_1^2(x_1, y_1) = \frac{1}{2}, & p_1^2(x_1, y_2) = \frac{5}{6}, & p_2^2(a_1, b_1) = \frac{2}{3}, & p_2^2(a_1, b_2) = \frac{1}{2}, \\ p_1^1(x_2, y_1) = \frac{5}{6}, & p_1^1(x_2, y_2) = \frac{1}{3}, & p_2^1(a_2, b_1) = \frac{1}{2}, & p_2^1(a_2, b_2) = \frac{2}{3}, \\ p_1^2(x_2, y_1) = \frac{1}{6}, & p_1^2(x_2, y_2) = \frac{2}{3}, & p_2^2(a_2, b_1) = \frac{1}{2}, & p_2^2(a_2, b_2) = \frac{1}{3}. \end{array}$$

Найти вектор Шепли и проверить его на позиционную состоятельность. В случае отсутствия позиционной состоятельности провести его регуляризацию.

26. Найти N -ядро в кооперативной игре из примера 3.10.1 в случае, когда параметры игры следующие:

$$\begin{array}{ll} a_1 = 0.7, & q = 0.001, \\ a_2 = 0.2, & f = 1, \\ D_{12} = 0.1, & d = 0.1, \\ D_{13} = 0.5, & c = 0.3, \\ D_{23} = 0.1, & \pi = (1, 0, 0, 0). \end{array}$$

Проверить N -ядро на позиционную состоятельность. В случае отсутствия позиционной состоятельности провести его регуляризацию.

Глава 5

Антагонистические дифференциальные игры

§ 5.1. Антагонистические дифференциальные игры с предписанной продолжительностью

Дифференциальные игры являются обобщением многошаговых игр на случай, когда число шагов в игре становится бесконечным (континуум), и игроки 1 и 2 (будем обозначать их буквами E и P), соответственно, имеют возможность принимать решения непрерывно. В такой постановке траектории движения игроков представляют собой решения систем дифференциальных уравнений, правые части которых зависят от параметров, которые находятся под контролем игроков.

5.1.1. Пусть $x \in R^n$, $y \in R^n$, $u \in U \subset R^k$, $v \in V \subset R^l$, $f(x, u), g(y, v)$ — вектор-функции размерности n , заданные на $R^n \times U$ и $R^n \times V$ соответственно. Рассмотрим две системы обыкновенных дифференциальных уравнений

$$\dot{x} = f(x, u), \quad (5.1.1)$$

$$\dot{y} = g(y, v) \quad (5.1.2)$$

с начальными условиями x_0, y_0 . Игрок P (E) начинает движение из фазового состояния x_0 (y_0) и перемещается в фазовом пространстве R^n согласно (5.1.1) или (5.1.2), выбирая в каждый момент времени значение параметра $u \in U$ ($v \in V$) в соответствии со своими целями и информацией, доступной в каждом текущем состоянии.

Наиболее просто поддается описанию случай полной информации. В дифференциальной игре это означает, что игрокам в каждый момент времени t при выборе параметров $u \in U$, $v \in V$ известно время t и фазовые состояния — свое и противника. Иногда требуют знание одним из игроков, например, игроком P , в каждый текущий момент t значения параметра $v \in V$, выбранного игроком E в этот же момент. В таком случае говорят, что игрок E дискриминирован, а сама игра называется *игрой с дискриминацией* игрока E .

Параметры $u \in U$, $v \in V$ называются управлениями игроков P и E соответственно. Функции $x(t), y(t)$, удовлетворяющие уравнениям (5.1.1), (5.1.2) и начальным условиям, называются *траекториями* движения игроков P, E .

5.1.2. Цели в дифференциальной игре определяются с помощью выигрыша, который может различным образом зависеть от реализовавшихся траекторий $x(t), y(t)$. Например, предполагается, что процесс игры продолжается некоторое заранее предписанное время T . Пусть $x(T), y(T)$ — фазовые состояния игроков P и E в момент окончания игры T . Тогда выигрыш игрока E полагается равным $H(x(T), y(T))$, где $H(x, y)$ — некоторая функция, заданная на $R^n \times R^n$. В частном случае, когда

$$H(x(T), y(T)) = \rho(x(T), y(T)), \quad (5.1.3)$$

где $\rho(x(T), y(T)) = \sqrt{\sum_{i=1}^n (x_i(T) - y_i(T))^2}$ — евклидово расстояние между точками $x(T), y(T)$, игра описывает процесс преследования, в котором целью игрока E является уклонение от игрока P к моменту окончания игры на максимальное расстояние. В этой главе будем предполагать дифференциальную игру антагонистической. В случае выполнения условия (5.1.3) это означает, что цель игрока P — максимальное сближение с игроком E к моменту окончания игры T .

При таком определении выигрыш зависит лишь от конечных состояний процесса и каждому игроку не засчитываются результаты, достигнутые им в процессе игры до момента T . Поэтому логичной является и такая постановка задачи, в которой выигрыш игрока E определяется как минимальное расстояние между игроками в процессе игры:

$$\min_{0 \leq t \leq T} \rho(x(t), y(t)).$$

Существуют игры, в которых ограничение на продолжительность игры не является существенным и игра продолжается до достижения игроками определенного результата. Пусть в R^{2n} задана m -мерная поверхность F , которую будем называть терминальной.

Положим

$$t_n = \{\min t : (x(t), y(t)) \in F\}, \quad (5.1.4)$$

т. е. t_n — первый момент попадания точки $(x(t), y(t))$ на поверхность F . Если при всех $t \geq 0$ точка $(x(t), y(t)) \notin F$, то t_n полагаем равным $+\infty$. Для реализовавшихся траекторий $x(t), y(t)$ выигрыш игрока E полагаем равным t_n (выигрыш игрока P равен $-t_n$). В частности, если F представляет собой сферу радиуса $l \geq 0$, заданную уравнением

$$\sqrt{\sum_{i=1}^n (x_i - y_i)^2} = l,$$

то имеет место задача преследования, в которой целью игрока P является скорейшее сближение с игроком E на расстояние $l \geq 0$. Если $l = 0$, то под встречей понимается совпадение фазовых координат игроков P и E , при этом игрок E стремится оттянуть момент встречи. Игры преследования этого типа будем называть *играми преследования на быстроедействие*.

В теории дифференциальных игр рассматриваются также задачи определения множества начальных состояний игроков, из которых игрок P может обеспечить встречу с игроком E на расстоянии l , и определения множества начальных состояний игроков, из которых игрок E может гарантировать, что встреча с игроком P на расстоянии l за конечное время не произойдет. Первое множество называется *областью встречи*

или захвата и обозначается (C, Z) , второе — областью убегания и обозначается (E, Z) . Очевидно, что эти области не пересекаются, однако важным является вопрос, покрывает ли объединение замыканий областей встречи и убегания все фазовое пространство? Ответ на этот вопрос будет дан ниже, а пока заметим, что для адекватного описания такого процесса достаточно определить выигрыш следующим образом. Если существует $t_n < \infty$ (см. (5.1.4)), то выигрыш игрока E полагаем равным -1 . Если же $t_n = \infty$, то выигрыш равен $+1$ (выигрыш игрока P равен выигрышу игрока E с противоположным знаком, так как игра антагонистическая).

Игры преследования с таким выигрышем называются *играми преследования качества*.

5.1.3. Фазовые ограничения. Если дополнительно потребовать, чтобы в процессе игры фазовая точка (x, y) не покидала некоторого множества $F \subset R^{2n}$, то получим дифференциальную игру с *фазовыми ограничениями*. Частным случаем такой игры является игра с «линией жизни». Она является антагонистической игрой качества, в которой выигрыш игрока E полагается равным $+1$, если ему удастся достичь границы множества F («линии жизни») до встречи с игроком P . Таким образом, целью игрока E является достижение границы множества F до встречи с игроком P (сближение с игроком P на расстояние $l \geq 0$), цель же игрока P — сближение с игроком E на расстояние $l \geq 0$, пока последний еще находится внутри множества F . Предполагается, что в процессе игры игрок P не может покинуть множества F .

5.1.4. Пример 1 (Простое движение.) Игра происходит на плоскости. Движение игроков P и E описывается системой дифференциальных уравнений

$$\begin{aligned} \dot{x}_1 &= u_1, \quad \dot{x}_2 = u_2, \quad u_1^2 + u_2^2 \leq \alpha^2, \\ \dot{y}_1 &= v_1, \quad \dot{y}_2 = v_2, \quad v_1^2 + v_2^2 \leq \beta^2, \\ x_1(0) &= x_1^0, \quad x_2(0) = x_2^0, \quad y_1(0) = y_1^0, \quad y_2(0) = y_2^0, \quad \alpha \geq \beta. \end{aligned} \quad (5.1.5)$$

С физической точки зрения уравнения (5.1.5) означают, что игроки P и E перемещаются в плоскости с ограниченными скоростями, при этом максимальные скорости α и β постоянны по величине и максимальная скорость игрока E не превосходит скорость игрока P .

Выбирая в каждый момент времени управление $u = (u_1, u_2)$, стесненное ограничением $u_1^2 + u_2^2 \leq \alpha^2$ (множество U), игрок P может изменять направление движения (направление вектора скорости). Аналогично, игрок E , выбирая в каждый момент времени управление $v = (v_1, v_2)$, стесненное ограничением $v_1^2 + v_2^2 \leq \beta^2$ (множество V), может также в каждый момент времени изменить направление движения. Очевидно, что если $\alpha > \beta$, то множество захвата (C, Z) совпадает со всем пространством, т. е. игрок P всегда может гарантировать l -встречу с игроком E за конечное время для любого l . Для этого достаточно выбрать движение с максимальной скоростью α и в каждый момент времени t направлять вектор скорости на преследуемую точку $y(t)$, т. е. осуществлять преследование по погонной линии. Если $\alpha \leq \beta$, то множество убегания (E, Z) совпадает со всем пространством игры за вычетом точек (x, y) , для которых $\rho(x, y) \leq l$. Действительно, если в начальный момент $\rho(x_0, y_0) > l$, то игрок E всегда может гарантировать избежание захвата, удаляясь от игрока P вдоль прямой, соединяющей начальные точки x_0, y_0 с максимальной скоростью β .

Здесь проявляется характерное свойство, которое будет встречаться и в дальнейшем. Для формирования управления, гарантирующего игроку E избежание захвата,

достаточно знать лишь начальные состояния x_0, y_0 , в то время как игроку P в случае $\alpha > \beta$ для формирования управления, гарантирующего встречу с игроком E , необходимо иметь информацию о своем состоянии и состоянии противника в каждый текущий момент времени.

Пример 2. Игроки P и E представляют собой материальные точки с единичными массами, которые перемещаются на плоскости под действием ограниченных по модулю сил и силы трения. Уравнения движения игроков имеют вид

$$\begin{aligned}\dot{x}_1 &= x_3, \quad \dot{x}_2 = x_4, \quad \dot{x}_3 = \alpha u_1 - k_P x_3, \\ \dot{x}_4 &= \alpha u_2 - k_P x_4, \quad u_1^2 + u_2^2 \leq \alpha^2, \\ \dot{y}_1 &= y_3, \quad \dot{y}_2 = y_4, \quad \dot{y}_3 = \beta v_1 - k_E y_3, \\ \dot{y}_4 &= \beta v_2 - k_E y_4, \quad v_1^2 + v_2^2 \leq \beta^2,\end{aligned}\tag{5.1.6}$$

где $(x_1, x_2), (y_1, y_2)$ — геометрические координаты, $(x_3, x_4), (y_3, y_4)$ — импульсы точек P и E соответственно, k_P и k_E — коэффициенты трения, α и β — максимальные силы, которые могут быть приложены к материальным точкам P и E . Движение начинается из состояний $x_i(0) = x_i^0, y_i(0) = y_i^0, i = 1, 2, 3, 4$. Здесь под состоянием понимается не геометрическое местоположение игроков P и E , а их фазовое состояние в пространстве координат и импульсов. Множества U, V — представляют собой круги $U = \{u = (u_1, u_2) : u_1^2 + u_2^2 \leq \alpha^2\}, V = \{v = (v_1, v_2) : v_1^2 + v_2^2 \leq \beta^2\}$. Это означает, что игроки P и E в каждый момент времени могут выбирать направления прилагаемых сил, однако максимальные значения этих сил ограничены константами α и β . В такой постановке, как это будет показано в дальнейшем, условия $\alpha > \beta$ (превосходство в силе) недостаточно для завершения преследования игроком P из любого начального состояния.

5.1.5. Пока не указан способ выбора управлений $u \in U, v \in V$ игроками P и E в процессе игры в зависимости от поступающей информации. Иначе говоря, не дано определение понятия стратегии в дифференциальной игре.

Существует несколько разных подходов к определению этого понятия. Остановимся на тех интуитивно очевидных теоретико-игровых качествах, которыми оно должно обладать. Как уже отмечалось в гл. 4, стратегия должна характеризовать поведение игрока во всех информационных состояниях, в которых он может оказаться в процессе игры. В дальнейшем будем определять информационное состояние каждого игрока фазовыми векторами $x(t), y(t)$ в текущий момент t и временем $t - t_0$, прошедшим с момента начала игры. Тогда естественно было бы рассматривать стратегию игрока $P(E)$ как функцию $u(x, y, t) (v(x, y, t))$ со значениями в множестве управлений $U(V)$. Именно таким образом определяется стратегия в [Айзекс, 1967]. Стратегии этого типа будем называть *синтезирующими*. Однако этот способ определения стратегии обладает рядом существенных недостатков. Действительно, пусть игроки P и E выбрали стратегии $u(x, y, t)$ и $v(x, y, t)$ соответственно. Тогда для определения траектории движения игроков, следовательно, и выигрыша (который зависит от траекторий) подставим функции $u(x, y, t), v(x, y, t)$ в уравнения (5.1.1), (5.1.2) вместо управляющих параметров u, v и попытаемся их проинтегрировать при начальных условиях x_0, y_0 на отрезке времени $[0, T]$. Получим следующую систему обыкновенных дифференциальных уравнений:

$$\dot{x} = f(x, u(x, y, t)), \quad \dot{y} = g(y, v(x, y, t)).\tag{5.1.7}$$

Для существования и единственности решения системы (5.1.7) необходимо наложить определенные условия на функции $f(x, u)$, $g(y, v)$ и стратегии $u(x, y, t)$, $v(x, y, t)$. Первая группа условий не ограничивает стратегических возможностей игроков, относится к постановочной части задачи и оправдывается физической природой рассматриваемого процесса. По-иному обстоит дело с ограничениями на класс функций (стратегий) $u(x, y, t)$, $v(x, y, t)$. Ограничения возможностей игроков не согласуются с принятым в теории игр представлением о свободе выбора поведения и приводят в ряде случаев к существенному «оскудению» множеств стратегий. Например, если ограничиться лишь непрерывными функциями $u(x, y, t)$, $v(x, y, t)$, то встречаются задачи, в которых не существует решения в классе непрерывных функций. Допущение же более широкого класса стратегий приводит к невозможности обеспечить существование единственного решения системы (5.1.7) на отрезке $[0, T]$. Иногда для преодоления этой трудности рассматривают множества таких стратегий $u(x, y, t)$, $v(x, y, t)$, при которых система (5.1.7) имеет единственное решение, продолжимое на отрезок $[0, T]$. Однако такой подход (помимо неконструктивности определения множества стратегий) не является достаточно обоснованным, поскольку множество всех пар стратегий $u(x, y, t)$, $v(x, y, t)$, при которых система (5.1.7) имеет единственное решение, оказывается прямоугольным.

5.1.6. В качестве стратегий в дифференциальной игре будем рассматривать *кусочно-программные стратегии*.

Кусочно-программная стратегия $u(\cdot)$ игрока P состоит из пары $\{\sigma, a\}$, где σ — некоторое разбиение $0 = t'_0 < t'_1 < \dots < t'_n < \dots$ полуоси времени $[0, \infty)$ точками t'_k , не имеющими конечных точек сгущения; a — отображение, ставящее в соответствие каждой точке t'_k и фазовым состояниям $x(t'_k)$, $y(t'_k)$ некоторое измеримое программное управление $u(t) \in U$ при $t \in [t'_k, t'_{k+1})$ (измеримую функцию $u(t)$, принимающую значения из множества U). Аналогично, кусочно-программная стратегия $v(\cdot)$ игрока E состоит из пары $\{\tau, b\}$, где τ — некоторое разбиение $0 = t''_0 < t''_1 < \dots < t''_n < \dots$ полуоси времени $[0, \infty)$ точками t''_k , не имеющими конечных точек сгущения; b — отображение, ставящее в соответствие каждой точке t''_k и позициям $x(t''_k)$, $y(t''_k)$ некоторое измеримое программное управление $v(t) \in V$ на отрезке $[t''_k, t''_{k+1})$ (измеримую функцию $v(t)$, принимающую значения из множества V). Используя кусочно-программную стратегию, игрок реагирует на изменение информации не непрерывно во времени, а через интервал $[t_k, t_{k+1})$, длину которого он определяет сам.

Обозначим множество всех кусочно-программных стратегий игрока P через P , а множество всех возможных кусочно-программных стратегий игрока E — через E .

Пусть $u(t)$, $v(t)$ — пара измеримых программных управлений игроков P и E (измеримых функций со значениями в множествах управлений U, V). Рассмотрим систему обыкновенных дифференциальных уравнений

$$\dot{x} = f(x, u(t)), \quad \dot{y} = g(y, v(t)), \quad t \geq 0. \quad (5.1.8)$$

На правые части систем (5.1.8) наложим следующие ограничения. Вектор-функции $f(x, u)$, $g(y, v)$ — непрерывны по всем аргументам, т. е. $f(x, u)$ непрерывна на множестве $R^n \times U$, а $g(y, v)$ непрерывна на множестве $R^n \times V$. Кроме того, вектор-функции $f(x, u)$ и $g(y, v)$ удовлетворяют условию Липшица по x и y , соответственно, независимо от u и v , т. е.

$$\begin{aligned} \|f(x_1, u) - f(x_2, u)\| &\leq \alpha_1 \|x_1 - x_2\|, \quad u \in U, \\ \|g(y_1, v) - g(y_2, v)\| &\leq \beta_1 \|y_1 - y_2\|, \quad v \in V, \end{aligned}$$

где $U \subset R^k, V \subset R^l$ — компактные множества в соответствующих евклидовых пространствах.

Из теорем существования и единственности Каратеодори следует, что при выполнении указанных условий для любых начальных состояний x_0, y_0 , любых измеримых программных управлений $u(t), v(t)$, заданных на отрезке $[t_1, t_2]$, $0 \leq t_1 < t_2$, существуют единственные абсолютно непрерывные вектор-функции $x(t), y(t)$, которые удовлетворяют почти всюду (т. е. всюду, за исключением множества меры нуль) в промежутке $[t_1, t_2]$ системе дифференциальных уравнений

$$\dot{x}(t) = f(x(t), u(t)), \quad \dot{y}(t) = g(y(t), v(t)) \quad (5.1.9)$$

и начальному условию $x(t_1) = x_0, y(t_1) = y_0$ (см. [Колмогоров, Фомин, 1981; Сансоне, 1954]).

5.1.7. Пусть (x_0, y_0) — пара начальных условий для уравнений (5.1.8). Система $S = \{x_0, y_0; u(\cdot), v(\cdot)\}$, где $u(\cdot) \in P, v(\cdot) \in E$, называется *ситуацией* в дифференциальной игре. Каждой ситуации S единственным образом соответствует пара траекторий $x(t), y(t)$ таких, что $x(0) = x_0, y(0) = y_0$ и при почти всех $t \in [0, T]$, $T > 0$ выполнены соотношения (5.1.9).

Действительно, пусть $u(\cdot) = \{\delta, a\}, v(\cdot) = \{\tau, b\}$. Далее, пусть $0 = t_0 < t_1 < \dots < t_k < \dots$ — разбиение полуоси $[0, \infty)$, являющееся объединением разбиений δ, τ . Решение системы (5.1.9) строится следующим образом. На каждом отрезке $[t_k, t_{k+1})$, $k = 0, 1, \dots$, образы отображений a, b представляют собой измеримые программные управления $u(t), v(t)$. Поэтому на отрезке $[t_0, t_1)$ система уравнений (5.1.9) при $x(0) = x_0, y(0) = y_0$ имеет единственное решение. На отрезке $[t_1, t_2)$, взяв в качестве начальных условий $x(t_1) = \lim_{t \rightarrow t_1-0} x(t), y(t_1) = \lim_{t \rightarrow t_1-0} y(t)$ строим решение (5.1.9), вторично используя измеримость управлений $u(t), v(t)$ как образов отображений a и b на отрезках $[t_k, t_{k+1})$, $k = 1, 2, \dots$. Полагая $x(t_2) = \lim_{t \rightarrow t_2-0} x(t), y(t_2) = \lim_{t \rightarrow t_2-0} y(t)$ продолжаем этот процесс, в результате чего находим единственное решение $x(t), y(t)$, такое что $x(0) = x_0, y(0) = y_0$. Любую траекторию $x(t)$ ($y(t)$), соответствующую некоторой ситуации $\{x_0, y_0, u(\cdot), v(\cdot)\}$, будем называть траекторией игрока P (игрока E).

5.1.8. Функция выигрыша. Как уже было показано, каждая ситуация $S = \{x_0, y_0; u(\cdot), v(\cdot)\}$ в кусочно-программных стратегиях однозначно определяет траектории $x(t), y(t)$ игроков P и E . Степень предпочтительности этих траекторий будем оценивать функцией выигрыша K , которая каждой ситуации ставит в соответствие некоторое вещественное число — выигрыш игрока E . Выигрыш игрока P равен $(-K)$ (это означает, что игра антагонистическая, поскольку сумма выигрышей игроков P и E в каждой ситуации равна нулю). Будем рассматривать игры с функцией выигрыша четырех видов.

Терминальный выигрыш. Заданы некоторое число $T > 0$ и непрерывная по (x, y) функция $H(x, y)$. Выигрыш в каждой ситуации $S = \{x_0, y_0; u(\cdot), v(\cdot)\}$ определяется следующим образом:

$$K(x_0, y_0; u(\cdot), v(\cdot)) = H(x(T), y(T)),$$

где $x(T) = x(t)|_{t=T}, y(T) = y(t)|_{t=T}$ (здесь $x(t), y(t)$ — траектории игроков P и E , соответствующие ситуации S). В случае, когда функция $H(x, y)$ представляет собой евклидово расстояние между точками x и y , имеет место задача преследования.

Минимальный результат. Пусть $H(x, y)$ — вещественная непрерывная функция. В ситуации $S = \{x_0, y_0; u(\cdot), v(\cdot)\}$ выигрыш игрока E полагается равным $\min_{0 \leq t \leq T} H(x(t), y(t))$, где $T > 0$ — заданное число.

Если $H(x, y) = \rho(x, y)$, то игра описывает процесс преследования.

Интегральный выигрыш. В $R^n \times R^n$ заданы некоторое многообразие F размерности m и непрерывная функция $H(x, y)$. Пусть в ситуации $S = \{x_0, y_0; u(\cdot), v(\cdot)\}$, t_n — первый момент попадания траектории $(x(t), y(t))$ на F . Тогда

$$K(x_0, y_0; u(\cdot), v(\cdot)) = \int_0^{t_n} H(x(t), y(t)) dt,$$

(если $t_n = \infty$, то $K = \infty$), где $x(t), y(t)$ — траектории игроков P и E , соответствующие ситуации S . В случае $H \equiv 1, K = t_n$, имеет место задача преследования на быстродействие.

Качественный выигрыш. Функция выигрыша K может принимать только одно из следующих трех значений: $+1, 0, -1$ в зависимости от расположения $(x(t), y(t))$ в $R^n \times R^n$. В $R^n \times R^n$ заданы два многообразия F и L размерности m_1 и m_2 соответственно. Пусть в ситуации $S = \{x_0, y_0; u(\cdot), v(\cdot)\}$, t_n — первый момент попадания траектории $(x(t), y(t))$ на F . Тогда

$$K(x_0, y_0; u(\cdot), v(\cdot)) = \begin{cases} -1, & \text{если } (x(t_n), y(t_n)) \in L, \\ 0, & \text{если } t_n = \infty, \\ +1, & \text{если } (x(t_n), y(t_n)) \notin L. \end{cases}$$

5.1.9. Определив множества стратегий игроков P и E и функцию выигрыша, можно определить дифференциальную игру как игру в нормальной форме. В п. 1.1.1 под нормальной формой Γ мы понимали тройку $\Gamma = \langle X, Y, K \rangle$, где $X \times Y$ — пространство пар всевозможных стратегий в игре Γ , а K — функция выигрыша, определенная на $X \times Y$. В рассматриваемом случае функция выигрыша определена не только на множестве пар всевозможных стратегий в игре, но и на множестве всех пар начальных позиций x_0, y_0 . Поэтому каждой паре $(x_0, y_0) \in R^n \times R^n$ соответствует своя игра в нормальной форме, т. е. фактически определяется некоторое семейство игр в нормальной форме, зависящее от параметров $(x_0, y_0) \in R^n \times R^n$.

Определение. Под нормальной формой дифференциальной игры $\Gamma(x_0, y_0)$, заданной на пространстве пар стратегий $P \times E$, будем понимать систему

$$\Gamma(x_0, y_0) = \langle x_0, y_0; P, E, K(x_0, y_0; u(\cdot), v(\cdot)) \rangle,$$

где $K(x_0, y_0; u(\cdot), v(\cdot))$ — функция выигрыша, определенная любым из четырех описанных выше способов.

Если функция выигрыша K в игре Γ терминальная, то соответствующая игра Γ называется игрой с терминальным выигрышем. Если функция K определяется вторым способом, то имеем игру на достижение минимального результата. Если функция K в игре Γ является интегральной, то соответствующая игра Γ называется игрой с интегральным выигрышем. Когда функция выигрыша в игре Γ качественная, соответствующая игра Γ называется игрой качества.

5.1.10. Естественно, что в классе кусочно-программных стратегий (в виду некомпактности множества) оптимальных стратегий может не существовать. Однако удастся показать, что в достаточно большом числе случаев для любого $\varepsilon > 0$ существуют ситуации ε -равновесия. Напомним определение ситуации ε -равновесия (см. п. 2.2.3).

Определение. Пусть задано некоторое $\varepsilon > 0$. Ситуация $s_\varepsilon = \{x_0, y_0; u_\varepsilon(\cdot), v_\varepsilon(\cdot)\}$ называется ситуацией ε -равновесия в игре $\Gamma(x_0, y_0)$, если для всех $u(\cdot) \in P$, $v(\cdot) \in E$ выполнено неравенство

$$K(x_0, y_0; u(\cdot), v_\varepsilon(\cdot)) + \varepsilon \geq K(x_0, y_0; u_\varepsilon(\cdot), v_\varepsilon(\cdot)) \geq K(x_0, y_0; u_\varepsilon(\cdot), v(\cdot)) - \varepsilon. \quad (5.1.10)$$

Стратегии $u_\varepsilon(\cdot), v_\varepsilon(\cdot)$, определенные в (5.1.10), называются ε -оптимальными стратегиями игроков P и E соответственно.

Следующая лемма является перефразировкой теоремы п. 2.2.5 для дифференциальных игр.

Лемма. Пусть в игре $\Gamma(x_0, y_0)$ для каждого $\varepsilon > 0$ существует ситуация ε -равновесия $s_\varepsilon = \{x_0, y_0, u_\varepsilon(\cdot), v_\varepsilon(\cdot)\}$. Тогда существует предел

$$\lim_{\varepsilon \rightarrow 0} K(x_0, y_0; u_\varepsilon(\cdot), v_\varepsilon(\cdot)).$$

Определение. Функция $V(x, y)$, определенная в каждой точке (x, y) некоторого множества $D \subset R^n \times R^n$ по правилу

$$\lim_{\varepsilon \rightarrow 0} K(x, y; u_\varepsilon(\cdot), v_\varepsilon(\cdot)) = V(x, y), \quad (5.1.11)$$

называется функцией значения игры $\Gamma(x, y)$ на множестве начальных условий $(x, y) \in D$.

Существование при любом $\varepsilon > 0$ ситуации ε -равновесия в игре $\Gamma(x_0, y_0)$ эквивалентно выполнению равенства

$$\sup_{v(\cdot) \in E} \inf_{u(\cdot) \in P} K(x_0, y_0; u(\cdot), v(\cdot)) = \inf_{u(\cdot) \in P} \sup_{v(\cdot) \in E} K(x_0, y_0; u(\cdot), v(\cdot)).$$

Если в игре $\Gamma(x_0, y_0)$ для любого $\varepsilon > 0$ существуют ε -оптимальные стратегии игроков P и E , то будем говорить, что игра $\Gamma(x_0, y_0)$ имеет решение.

Определение. Пусть $u^*(\cdot), v^*(\cdot)$ — пара таких стратегий, что

$$K(x_0, y_0; u(\cdot), v^*(\cdot)) \geq K(x_0, y_0; u^*(\cdot), v^*(\cdot)) \geq K(x_0, y_0; u^*(\cdot), v(\cdot)) \quad (5.1.12)$$

для всех $u(\cdot) \in P$ и $v(\cdot) \in E$. Тогда ситуация $s^* = (x_0, y_0; u^*(\cdot), v^*(\cdot))$ называется ситуацией равновесия в игре $\Gamma(x_0, y_0)$. Стратегии $u^*(\cdot) \in P$ и $v^*(\cdot) \in E$ из (5.1.12) называются оптимальными стратегиями игроков P и E соответственно.

Существование ситуации равновесия в игре $\Gamma(x_0, y_0)$ эквивалентно (см. п. 1.3.4) выполнению равенства

$$\max_{v(\cdot) \in E} \inf_{u(\cdot) \in P} K(x_0, y_0; u(\cdot), v(\cdot)) = \min_{u(\cdot) \in P} \sup_{v(\cdot) \in E} K(x_0, y_0; u(\cdot), v(\cdot)).$$

Очевидно, что если существует ситуация равновесия, то для любого $\varepsilon > 0$ она является и ситуацией ε -равновесия, т. е. функция $V(x, y)$ в данном случае просто совпадает с $K(x, y; u^*(\cdot), v^*(\cdot))$ (см. п. 2.2.3).

5.1.11. Рассмотрим синтезирующие стратегии.

Определение. Пара $(u^*(x, y, t), v^*(x, y, t))$ называется ситуацией равновесия в дифференциальной игре в синтезирующих стратегиях, если имеет место неравенство

$$\begin{aligned} K(x_0, y_0; u(x, y, t), v^*(x, y, t)) &\geq K(x_0, y_0; u^*(x, y, t), v^*(x, y, t)) \geq \\ &\geq K(x_0, y_0; u^*(x, y, t), v(x, y, t)) \end{aligned} \quad (5.1.13)$$

для всех ситуаций $(u(x, y, t), v^*(x, y, t))$ и $(u^*(x, y, t), v(x, y, t))$, для которых существует единственное, продолжимое на $[0, \infty)$ решение системы (5.1.7) из начальных состояний x_0, y_0 . Стратегии $u^*(x, y, t), v^*(x, y, t)$ называются оптимальными стратегиями игроков P и E .

Установим различие понятий ситуации равновесия в кусочно-программных и синтезирующих стратегиях. Заметим, что определить ситуацию равновесия в обычном смысле в классе функций $u(x, y, t), v(x, y, t)$ невозможно из-за непрямоугольности пространства ситуаций, т. е. в синтезирующих стратегиях невозможно потребовать выполнения неравенства (5.1.13) для всех стратегий $u(x, y, t), v(x, y, t)$, поскольку некоторые пары $(u^*(x, y, t), v(x, y, t)), (u(x, y, t), v^*(x, y, t))$ могут не быть допустимыми (система уравнений (5.1.7) в соответствующей ситуации может не иметь решения вообще или не иметь единственного решения).

В дальнейшем, если специально не будет оговорено, во всех случаях будем рассматривать классы кусочно-программных стратегий. Прежде чем перейти к доказательству существования ситуации ε -равновесия в дифференциальной игре, рассмотрим один вспомогательный класс многошаговых игр с полной информацией.

§ 5.2. Многошаговые игры с полной информацией и бесконечным числом альтернатив

5.2.1. Рассмотрим класс многошаговых игр с полной информацией, представляющих собой обобщение игр с полной информацией из п. 4. 3. 1. Игра происходит в n -мерном евклидовом пространстве R^n . Будем обозначать через $x \in R^n$ местоположение (позицию) игрока 1, а через $y \in R^n$ — местоположение игрока 2. Пусть для каждого $x \in R^n, y \in R^n$ определены множества U_x, V_y соответственно, которые будем предполагать компактными множествами евклидова пространства R^n . Игра начинается из позиции x_0, y_0 . На 1-м шаге игроки 1 и 2 выбирают точки $x_1 \in U_{x_0}$ и $y_1 \in V_{y_0}$. При этом выбор игрока 2 сообщается игроку 1 до выбора им точки $x_1 \in U_{x_0}$. В точках x_1, y_1 игроки 1 и 2 выбирают точки $x_2 \in U_{x_1}$ и $y_2 \in V_{y_1}$, и выбор игрока 2 сообщается игроку 1 перед выбором им точки $x_2 \in U_{x_1}$ и т. д. На k -м шаге в позициях x_{k-1}, y_{k-1} игроки выбирают $x_k \in U_{x_{k-1}}$ и выбор игрока 2 сообщается игроку 1 перед выбором им точки $x_k \in U_{x_{k-1}}$. Процесс заканчивается на N -м шаге выбором $x_N \in U_{x_{N-1}}, y_N \in V_{y_{N-1}}$ и переходом в состояние x_N, y_N .

Семейства множеств $U_x, V_y, x \in R^n, y \in R^n$ предполагаются непрерывными в метрике Хаусдорфа по x, y . Это означает, что для любого $\varepsilon > 0$ найдется такое $\delta > 0$ что при $|x - x_0| < \delta$ ($|y - y_0| < \delta$)

$$\begin{aligned}(U_{x_0})_\varepsilon &\supset U_x, & (U_x)_\varepsilon &\supset U_{x_0}; \\ (V_{y_0})_\varepsilon &\supset V_y, & (V_y)_\varepsilon &\supset V_{y_0}.\end{aligned}$$

Здесь U_ε (V_ε) — ε -окрестность множества U (V).

Следующий результат хорошо известен в анализе (см. [Petrosyan, 1993]).

Лемма. Пусть $f(x', y')$ — непрерывная функция на декартовом произведении $U_x \times V_y$. Тогда если семейства $\{U_x\}, \{V_y\}$ непрерывны по Хаусдорфу по x, y , то функционалы

$$F_1(x, y) = \max_{y' \in V_y} \min_{x' \in U_x} f(x', y'),$$

$$F_2(x, y) = \min_{x' \in U_x} \max_{y' \in V_y} f(x', y')$$

непрерывны по x, y .

Пусть $\bar{x} = (x_0, \dots, x_N)$ и $\bar{y} = (y_0, \dots, y_N)$ — траектории игроков 1 и 2 соответственно, реализовавшиеся в процессе игры. Выигрышем игрока 2 является величина

$$\max_{0 \leq k \leq N} f(x_k, y_k) = F(\bar{x}, \bar{y}), \quad (5.2.1)$$

где $f(x, y)$ — непрерывная функция от x, y . Выигрыш игрока 1 равен $-F$ (игра антагонистическая).

Будем предполагать, что данная игра является игрой с полной информацией, т. е. в каждый момент времени (на каждом шаге) игрокам известны позиции x_k, y_k и момент времени k , а игроку 1, кроме того, известен выбор y_{k+1} игрока 2 в этот момент.

Стратегиями игрока 1 являются всевозможные функции $u(x, y, t)$ такие, что $u(x_{k-1}, y_k, k) \in U_{x_{k-1}}$. Стратегиями игрока 2 — всевозможные функции $v(x, y, t)$ такие, что $v(x_{k-1}, y_{k-1}, k) \in V_{y_{k-1}}$. Эти стратегии будем называть чистыми стратегиями (в отличие от смешанных).

Пусть игроки 1 и 2 применяют чистые стратегии $u(x, y, t), v(x, y, t)$. В ситуации $(u(\cdot), v(\cdot))$ игра происходит следующим образом. На 1-м шаге игрок 2 из состояния y_0 переходит в состояние $y_1 = v(x_0, y_0, 1)$, а игрок 1 — из состояния x_0 в состояние $x_1 = u(x_0, y_1, 1) = u(x_0, v(x_0, y_0, 1), 1)$ (поскольку игрок 1 знает выбор игрока 2). На 2-м шаге игроки переходят в состояния $y_2 = v(x_1, y_1, 2)$, $x_2 = u(x_1, y_2, 2) = u(x_1, v(x_1, y_1, 2), 2)$ и т. д. На k -м шаге игроки 1 и 2 переходят из состояний x_{k-1}, y_{k-1} в состояния $y_k = v(x_{k-1}, y_{k-1}, k)$, $x_k = u(x_{k-1}, y_k, k) = u(x_{k-1}, v(x_{k-1}, y_{k-1}, k), k)$. Таким образом, каждой ситуации $(u(\cdot), v(\cdot))$ однозначно соответствуют траектории игроков 1 и 2: $\bar{x} = (x_0, x_1, \dots, x_k, \dots, x_N)$ и $\bar{y} = (y_0, y_1, \dots, y_k, \dots, y_N)$; следовательно, и выигрыш $K(u(\cdot), v(\cdot)) = F(x, y)$, определяемый по формуле (5.2.1).

Рассматриваемая игра зависит от двух параметров: начальных позиций x_0, y_0 и продолжительности N , поэтому будем обозначать ее через $\Gamma(x_0, y_0, N)$. Для дальнейшего исследования каждую игру $\Gamma(x_0, y_0, N)$ удобно отнести к семейству игр $\Gamma(x, y, T)$, зависящих от параметров x, y, T .

5.2.2. Справедлив следующий результат, являющийся обобщением теоремы п. 4.2.1 для конечных игр с полной информацией.

Теорема. В игре $\Gamma(x_0, y_0, N)$ существует ситуация равновесия в чистых стратегиях и значение игры $V(x_0, y_0, N)$ удовлетворяет рекуррентному соотношению

$$V(x_0, y_0, k) = \max\{f(x_0, y_0), \max_{y \in V_{y_0}} \min_{x \in U_{x_0}} V(x, y, k-1)\}, \quad k = 1, \dots, N;$$

$$V(x, y, 0) = f(x, y). \quad (5.2.2)$$

Доказательство проведем методом индукции по числу шагов игры. Пусть $N = 1$. Определим стратегии $u^*(\cdot), v^*(\cdot)$ игроков в игре $\Gamma(x_0, y_0, 1)$ следующим образом:

$$\min_{x \in U_{x_0}} f(x, y) = f(u^*(x_0, y, 1), y), \quad y \in V_{y_0}.$$

Если $\max_{y \in V_{y_0}} \min_{x \in U_{x_0}} f(x, y) = f(u^*(x_0, y^*, 1), y^*)$, то $v^*(x_0, y_0, 1) = y^*$. Тогда

$$K(u^*(\cdot), v^*(\cdot)) = \max\{f(x_0, y_0), \max_{y \in V_{y_0}} \min_{x \in U_{x_0}} f(x, y)\}$$

и для любых стратегий $u(\cdot), v(\cdot)$ игроков в игре $\Gamma(x_0, y_0, 1)$ справедливы соотношения:

$$K(u^*(\cdot), v(\cdot)) \leq K(u^*(\cdot), v^*(\cdot)) \leq K(u(\cdot), v^*(\cdot)).$$

Тем самым утверждение теоремы справедливо при $N \leq 1$.

Предположим теперь, что утверждение теоремы справедливо при $N \leq n$ и докажем ее для $N = n + 1$, т. е. для игры $\Gamma(x_0, y_0, n + 1)$. Рассмотрим семейство игр $\Gamma(x, y, n)$, $x \in U_{x_0}, y \in V_{y_0}$. Обозначим через $\bar{u}_{xy}^n(\cdot), \bar{v}_{xy}^n(\cdot)$ ситуацию равновесия в игре $\Gamma(x, y, n)$. Тогда $K(\bar{u}_{xy}^n(\cdot), \bar{v}_{xy}^n(\cdot)) = V(x, y, n)$, где $V(x, y, n)$ определено соотношениями (5.2.2). Используя непрерывность функции $f(x, y)$ и лемму п. 5.2.1, нетрудно доказать непрерывность функции $V(x, y, n)$ по x, y . Определим стратегии $\bar{u}^{n+1}(\cdot), \bar{v}^{n+1}(\cdot)$ игроков в игре $\Gamma(x_0, y_0, n + 1)$ следующим образом:

$$\min_{x \in U_{x_0}} V(x, y, n) = V(\bar{u}^{n+1}(x_0, y, 1), y, n), \quad y \in V_{y_0}.$$

Если $\max_{y \in V_{y_0}} \min_{x \in U_{x_0}} V(x, y, n) = V(\bar{u}^{n+1}(x_0, \tilde{y}, 1), \tilde{y}, n)$, то $\bar{v}^{n+1}(x_0, y_0, 1) = \tilde{y}$ (для $x \neq x_0, y \neq y_0$ функции $\bar{v}^{n+1}(x, y, 1)$ и $\bar{u}^{n+1}(x, y, 1)$ определим произвольно):

$$\bar{u}^{n+1}(\cdot, k) = \bar{u}_{x_1 y_1}^n(\cdot, k-1), \quad k = 2, \dots, n+1,$$

$$\bar{v}^{n+1}(\cdot, k) = \bar{v}_{x_1 y_1}^n(\cdot, k-1), \quad k = 2, \dots, n+1.$$

Здесь $x_1 \in U_{x_0}, y_1 \in V_{y_0}$ — позиции, которые реализовались после 1-го шага в игре $\Gamma(x_0, y_0, n + 1)$. По построению,

$$K(\bar{u}^{n+1}(\cdot), \bar{v}^{n+1}(\cdot)) = \max\{f(x_0, y_0), \max_{y \in V_{y_0}} \min_{x \in U_{x_0}} V(x, y, n)\}. \quad (5.2.3)$$

Фиксируем произвольную стратегию $u(\cdot)$ игрока 1 в игре $\Gamma(x_0, y_0, n + 1)$. Пусть $u(x_0, \tilde{y}, 1) = x_1$, где $\tilde{y} = \bar{v}^{n+1}(x_0, y_0, 1)$ и $u_{xy}^n(\cdot)$ — сужение стратегии $u(\cdot)$ на игру $\Gamma(x, y, n)$, $x \in U_{x_0}, y \in V_{y_0}$.

Справедливы следующие соотношения:

$$\begin{aligned} K(\bar{u}^{n+1}(\cdot), \bar{v}^{n+1}(\cdot)) &\leq \max\{f(x_0, y_0), V(x_1, \tilde{y}, n)\} = \\ &= \max\{f(x_0, y_0), K(\bar{u}_{x_1 \tilde{y}}^n(\cdot), \bar{v}_{x_1 \tilde{y}}^n(\cdot))\} \leq \\ &\leq \max\{f(x_0, y_0), K(u_{x_1 \tilde{y}}^n(\cdot), \bar{v}_{x_1 \tilde{y}}^n(\cdot))\} = K(u(\cdot), \bar{v}^{n+1}(\cdot)). \end{aligned} \quad (5.2.4)$$

Аналогично доказывается неравенство

$$K(\bar{u}^{n+1}(\cdot), \bar{v}^{n+1}(\cdot)) \geq K(\bar{u}^{n+1}(\cdot), v(\cdot)) \quad (5.2.5)$$

для любой стратегии $v(\cdot)$ игрока 2 в игре: $\Gamma(x_0, y_0, n+1)$. Из соотношений (5.2.3)–(5.2.5) следует справедливость утверждения теоремы для $N = n+1$. Тем самым доказательство теоремы по индукции закончено.

Рассмотрим теперь игру $\bar{\Gamma}(x_0, y_0, N)$, которая отличается от игры $\Gamma(x_0, y_0, N)$ тем, что в ней сообщает свой выбор игрок 1. Таким образом, в игре $\bar{\Gamma}(x_0, y_0, N)$ на каждом шаге k игрок 2 кроме состояний x_{k-1}, y_{k-1} и шага k знает состояние $x_k \in U_{x_{k-1}}$, выбранное игроком 1. Игрок 1 на каждом шаге k знает лишь x_{k-1}, y_{k-1} . Аналогично теореме п. 5.2.5 можно показать, что в игре $\bar{\Gamma}(x_0, y_0, N)$ существует ситуация равновесия в чистых стратегиях и значение игры $\bar{V}(x_0, y_0, N)$ удовлетворяет рекуррентному уравнению

$$\begin{aligned} \bar{V}(x_0, y_0, k) &= \max\{f(x_0, y_0), \min_{x \in U_{x_0}} \max_{y \in V_{y_0}} \bar{V}(x, y, k-1)\}, \\ k &= 1, \dots, N, \quad \bar{V}(x, y, 0) = f(x, y). \end{aligned} \quad (5.2.6)$$

5.2.3. Рассмотрим игры $\Gamma'(x_0, y_0, N)$ и $\bar{\Gamma}'(x_0, y_0, N)$, которые отличаются от игр $\Gamma(x_0, y_0, N)$ и $\bar{\Gamma}(x_0, y_0, N)$ соответственно лишь видом функции выигрыша. Предположим, что в этих играх выигрыш игрока 2 равен расстоянию между ним и игроком 1 на последнем шаге игры, т. е. $\rho(x_N, y_N)$. Тогда утверждение теоремы п. 5.2.2 и ее следствие сохраняют силу и вместо рекуррентных уравнений (5.2.2), (5.2.6), справедливы уравнения

$$\begin{aligned} V'(x, y, k) &= \max_{y' \in V_y} \min_{x' \in U_x} V'(x', y', k-1), \quad k = 1, \dots, N, \\ V'(x, y, 0) &= \rho(x, y); \end{aligned} \quad (5.2.7)$$

$$\begin{aligned} \bar{V}'(x, y, k) &= \min_{x' \in U_x} \max_{y' \in V_y} \bar{V}'(x', y', k-1), \quad k = 1, \dots, N, \\ \bar{V}'(x, y, 0) &= \rho(x, y). \end{aligned} \quad (5.2.8)$$

Пример 3. Рассмотрим дискретную игру преследования, в которой множества U_x представляют собой круги радиуса α с центром в точке x , а множества V_y — круги радиуса β с центром в точке y ($\alpha > \beta$). Это соответствует игре, в которой игрок 2 (убегающий) перемещается на плоскости со скоростью, не превосходящей β , а игрок 1 (преследователь) — со скоростью, не превосходящей α . Скорость преследователя превосходит скорость убегающего, и игрок 1 ходит вторым. Игра такого типа называется дискретной игрой «простое преследование» с дискриминацией убегающего игрока. Игра продолжается N шагов, и выигрыш игрока 2 равен расстоянию между игроками на последнем шаге.

Найдем значение игры и оптимальные стратегии игроков, используя функциональное уравнение (5.2.7). Имеем

$$V(x, y, 1) = \max_{y' \in V_y} \min_{x' \in U_x} \rho(x', y'). \quad (5.2.9)$$

Поскольку U_x и V_y — круги с центрами в x и y с радиусами α и β , то $V(x, y, l) = 0$, если $U_x \supset V_y$.

Если же $U_x \not\supset V_y$, то $V(x, y, 1) = \rho(x, y) + \beta - \alpha = \rho(x, y) - (\alpha - \beta)$ (см. пример 8 в п. 2.2.6). Таким образом,

$$V(x, y, 1) = \begin{cases} 0, & \text{если } U_x \supset V_y, \text{ т. е. } \rho(x, y) - (\alpha - \beta) \leq 0, \\ \rho(x, y) - (\alpha - \beta), & \text{если } U_x \not\supset V_y, \end{cases}$$

или, что то же самое,

$$V(x, y, 1) = \max[0, \rho(x, y) - (\alpha - \beta)]. \quad (5.2.10)$$

Докажем, применив индукцию по числу шагов k , что имеет место следующая формула:

$$V(x, y, k) = \max[0, \rho(x, y) - k(\alpha - \beta)], \quad k \geq 2. \quad (5.2.11)$$

Пусть (5.2.11) выполнено при $k = m - 1$. Покажем, что формула справедлива для $k = m$. Воспользовавшись уравнением (5.2.7) и соотношениями (5.2.9), (5.2.10), получим

$$\begin{aligned} V(x, y, m) &= \max_{y' \in V_y} \min_{x' \in U_x} V(x', y', m - 1) = \\ &= \max_{y' \in V_y} \min_{x' \in U_x} \{\max[0, \rho(x', y') - (m - 1)(\alpha - \beta)]\} = \\ &= \max[0, \max_{y' \in V_y} \min_{x' \in U_x} \{\rho(x', y')\} - (m - 1)(\alpha - \beta)] = \\ &= \max[0, \max\{0, \rho(x, y) - (\alpha - \beta)\} - (m - 1)(\alpha - \beta)] = \max[0, \rho(x, y) - m(\alpha - \beta)], \end{aligned}$$

что и требовалось доказать.

Если $V(x_0, y_0, m) = \rho(x_0, y_0) - m(\alpha - \beta)$, т. е. $\rho(x_0, y_0) - m(\alpha - \beta) > 0$, то оптимальная стратегия игрока 2 диктует ему выбирать на k -м шаге игры точку y_k пересечения линии центров x_{k-1}, y_{k-1} с границей $V_{y_{k-1}}$, наиболее удаленную от x_{k-1} . Здесь x_{k-1}, y_{k-1} — позиции игроков после $(k - 1)$ -го шага, $k = 1, \dots, N$. Оптимальная стратегия игрока 1 диктует ему на k -м шаге игры выбирать точку из множества $U_{x_{k-1}}$, наиболее близкую к точке y_k . Если оба игрока действуют оптимально, то последовательность выбранных точек $x_0, x_1, \dots, x_N, y_0, y_1, \dots, y_N$ лежит на прямой, проходящей через x_0, y_0 . Если $V(x_0, y_0, m) = 0$, то оптимальная стратегия игрока 2 произвольна, а игрока 1 — та же. При этом после некоторого шага k выполняется равенство $\max_{y \in V_{y_k}} \min_{x \in U_{x_k}} \rho(x, y) = 0$. Поэтому, начиная с шага $(k + 1)$, выбор игрока 1 повторяет выбор игрока 2.

Многошаговые игры с неполной информацией детально изучены в работе [Слобожанин, 2002].

§ 5.3. Существование ситуаций ε -равновесия в дифференциальных играх с предписанной продолжительностью

5.3.1. В данном параграфе будет доказано существование ситуаций ε -равновесия в дифференциальных играх преследования с предписанной продолжительностью в классе кусочно-программных стратегий, определенных в п. 5.1.6. Рассмотрим подробно случай, когда выигрыш игрока E представляет собой расстояние $\rho(x(T), y(T))$ в последний момент игры T .

Пусть динамика игры задается следующими дифференциальными уравнениями:

$$\text{для } P: \dot{x} = f(x, u); \quad (5.3.1)$$

$$\text{для } E: \dot{y} = g(y, v). \quad (5.3.2)$$

Здесь $x, y \in R^n$, $u \in U$, $v \in V$, где U, V — компактные множества евклидовых пространств R^k и R^l соответственно, $t \in [0, \infty)$. Пусть выполнены все требования п. 5.1.6.

Определение. Обозначим через $C_P^t(x_0)$ множество точек $x \in R^n$ для которых существует измеримое программное управление $u(t) \in U$, переводящее точку x_0 в x за время t , т. е. $x(t_0) = x_0$, $x(t_0 + t) = x$. Множество $C_P^t(x_0)$ называется множеством достижимости игрока P из начального состояния x_0 за время t .

Аналогично определяется множество достижимости $C_E^t(y_0)$ игрока E из начального состояния y_0 за время t .

Предположим, что функции f, g таковы, что множества достижимости $C_P^t(x_0)$, $C_E^t(y_0)$ игроков P и E соответственно удовлетворяют следующим условиям:

- 1) $C_P^t(x_0)$, $C_E^t(y_0)$ определены при всяких $x_0, y_0 \in R^n$, $t_0, t \in [0, \infty)$ ($t_0 \leq t$) и являются компактными множествами пространства R^n ;
- 2) отображение $C_P^t(x_0)$ непрерывно по совокупности аргументов в метрике Хаусдорфа, т. е. для любых $\varepsilon > 0$, $x'_0 \in R^n$, $t \in [0, \infty)$ существует такое $\delta > 0$, что если $|t - t'| < \delta$, $\rho(x_0, x'_0) < \delta$, то $\rho^*(C_P^t(x_0), C_P^{t'}(x'_0)) < \varepsilon$. То же выполняется для $C_E^t(y_0)$.

Напомним, что метрика Хаусдорфа ρ^* в пространстве компактных подмножеств R^n задается так:

$$\rho^*(A, B) = \max(\rho'(A, B), \rho'(B, A)), \quad \rho'(A, B) = \max_{a \in A} \rho(a, B)$$

и $\rho(a, B) = \min_{b \in B} \rho(a, b)$, где ρ — стандартная метрика в R^n .

Теорему существования будем доказывать для игры преследования $\Gamma(x_0, y_0, T)$ с предписанной продолжительностью, где $x_0, y_0 \in R^n$ — начальные позиции игроков P и E соответственно, а T — продолжительность игры. Игра $\Gamma(x_0, y_0, T)$ протекает следующим образом. Игроки P и E в момент времени $t_0 = 0$ начинают перемещаться из позиций x_0, y_0 в соответствии с выбранными кусочно-программными стратегиями. В момент времени $t = T$ игра заканчивается, при этом игрок E получает от игрока P выигрыш, равный $\rho(x(T), y(T))$ (см. п. 5.1.8). В каждый момент времени $t \in [0, T]$ игры $\Gamma(x_0, y_0, T)$ обоим игрокам известны момент времени t , своя позиция и позиция противника. Обозначим через $P(x_0, t_0, t)$ ($E(y_0, t_0, t)$) множество траекторий системы (5.3.1) ((5.3.2)), исходящих из точки $x_0(y_0)$ и определенных на промежутке $[t_0, t]$.

5.3.2. Фиксируем некоторое натуральное $n \geq 1$. Положим $\delta = T/2^n$ и введем в рассмотрение вспомогательные по отношению к игре $\Gamma(x_0, y_0, T)$ игры $\Gamma_i^\delta(x_0, y_0, T)$, $i = 1, 2, 3$.

Игра $\Gamma_1^\delta(x_0, y_0, T)$ протекает следующим образом. На 1-м шаге игрок E , находясь в позиции y_0 , выбирает y_1 из множества $C_E^\delta(y_0)$, а игрок P , находясь в позиции y_1 и зная выбор y_1 игрока E на этом шаге, выбирает точку $x_1 \in C_P^\delta(x_0)$. На k -м шаге, $k = 2, 3, \dots, 2^n$, игрок E , зная позицию $x_{k-1} \in C_P^\delta(x_{k-2})$ игрока P и свою позицию $y_{k-1} \in C_E^\delta(y_{k-2})$, выбирает точку $y_k \in C_E^\delta(y_{k-1})$. Игрок P , зная x_{k-1}, y_{k-1}, y_k , выбирает $x_k \in C_P^\delta(x_{k-1})$. На 2^n -м шаге игра заканчивается, и игрок E получает выигрыш, равный $\rho(x(T), y(T))$, где $x(T) = x^{2^n}$, $y(T) = y^{2^n}$.

Отметим, что выбор игроками на k -м шаге точек x_k, y_k из множеств достижимости $C_P^\delta(x_{k-1})$, $C_E^\delta(y_{k-1})$ можно трактовать как выбор ими соответствующих траекторий из множеств $P((x_{k-1}, k-1)\delta, k\delta)$, $E((y_{k-1}, k-1)\delta, k\delta)$, оканчивающихся в точках x_k, y_k в

момент $t = k\delta$ (или выбор управлений $u(\cdot), v(\cdot)$ на $[(k-1)\delta, k\delta]$ которым эти траектории соответствуют согласно (5.3.1), (5.3.2)).

Игра $\Gamma_2^\delta(x_0, y_0, T)$ отличается от игры $\Gamma_1^\delta(x_0, y_0, T)$ тем, что на k -м шаге игрок P выбирает $x_k \in C_P^\delta(x_{k-1})$, зная x_{k-1}, y_{k-1} , а игрок E , зная, кроме того, x_k , выбирает $y_k \in C_E^\delta(y_{k-1})$.

Игра $\Gamma_3^\delta(x_0, y_0, T)$ отличается от игры $\Gamma_2^\delta(x_0, y_0, T)$ тем, что на 2^n -м шаге игрок P выбирает $x_{2^n} \in C_P^\delta(x_{2^n-1})$, после чего игра заканчивается и игрок E получает выигрыш $\rho(x(T), y(T-\delta))$, где $x(T) = x_{2^n}$, $y(T-\delta) = y_{2^n-1}$.

5.3.3. Лемма. В играх $\Gamma_i^\delta(x_0, y_0, T)$, $i = 1, 2, 3$, существуют ситуации равновесия при всех $x_0, y_0, T < \infty$ и значение игры $Val\Gamma_i^\delta(x_0, y_0, T)$ есть непрерывная функция по $x_0, y_0 \in R^n$. При всяком $n \geq 0$ выполняется неравенство

$$Val\Gamma_1^\delta(x_0, y_0, T) \leq Val\Gamma_2^\delta(x_0, y_0, T), \quad T = 2^n\delta. \quad (5.3.3)$$

Доказательство. Игры $\Gamma_i^\delta(x_0, y_0, T)$, $i = 1, 2, 3$ принадлежат классу многошаговых игр, определенных в п. 5.2.1. Существование ситуации равновесия в играх $\Gamma_i^\delta(x_0, y_0, T)$ и непрерывность функций $Val\Gamma_i^\delta(x_0, y_0, T)$ по x_0, y_0 непосредственно следует из теоремы п. 5.2.2 и ее следствия. Для значений игр $\Gamma_i^\delta(x_0, y_0, T)$, $i = 1, 2$,

$$Val\Gamma_1^\delta(x_0, y_0, T) = \max_{y \in C_E^\delta(y_0)} \min_{x \in C_P^\delta(x_0)} Val\Gamma_1^\delta(x, y, T-\delta),$$

$$Val\Gamma_2^\delta(x_0, y_0, T) = \min_{x \in C_P^\delta(x_0)} \max_{y \in C_E^\delta(y_0)} Val\Gamma_2^\delta(x, y, T-\delta),$$

при начальном условии $Val\Gamma_1^\delta(x, y, 0) = Val\Gamma_2^\delta(x, y, 0) = \rho(x, y)$. Применяя последовательно лемму п. 1.2.2, убеждаемся в справедливости неравенства (5.3.3).

5.3.4. Лемма. При любом целом $n \geq 0$ справедливы неравенства:

$$Val\Gamma_1^{\delta n}(x_0, y_0, T) \leq Val\Gamma_1^{\delta n+1}(x_0, y_0, T),$$

$$Val\Gamma_2^{\delta n}(x_0, y_0, T) \geq Val\Gamma_2^{\delta n+1}(x_0, y_0, T),$$

где $\delta_k = T/2^k$

Доказательство. Покажем справедливость первого из неравенств. Второе неравенство доказывается аналогично. Во избежание громоздкости обозначений будем далее полагать $C^k(y_i) = C_E^{\delta_k}(y_i)$, $C^k(x_i) = C_P^{\delta_k}(x_i)$, $i = 0, 1, \dots, 2^n - 1$. Имеем

$$\begin{aligned} Val\Gamma_1^{\delta n+1}(x_0, y_0, T) &= \\ &= \max_{y_1 \in C^{n+1}(y_0)} \min_{x_1 \in C^{n+1}(x_0)} \max_{y_2 \in C^{n+1}(y_1)} \min_{x_2 \in C^{n+1}(x_1)} Val\Gamma_1^{\delta n+1}(x_2, y_2, T - 2\delta_{n+1}) \geq \\ &\geq \max_{y_1 \in C^{n+1}(y_0)} \max_{y_2 \in C^{n+1}(y_1)} \min_{x_1 \in C^{n+1}(x_0)} \min_{x_2 \in C^{n+1}(x_1)} Val\Gamma_1^{\delta n+1}(x_2, y_2, T - 2\delta_{n+1}) = \\ &= \max_{y_1 \in C^n(y_0)} \min_{x_1 \in C^n(x_0)} Val\Gamma_1^{\delta n+1}(x_1, y_1, T - \delta_n). \end{aligned}$$

Продолжая этот процесс, получим

$$\begin{aligned} Val\Gamma_1^{\delta n+1}(x_0, y_0, T) &\geq \max_{y_1 \in C^n(y_0)} \min_{x_1 \in C^n(x_0)} \dots \max_{y_{2^n} \in C^n(y_{2^n-1})} \min_{x_{2^n} \in C^n(x_{2^n-1})} \rho(x_{2^n}, y_{2^n}) = \\ &= Val\Gamma_1^{\delta n}(x_0, y_0, T). \end{aligned}$$

5.3.5. Теорема. При всех $x_0, y_0 \in R^n$, $T < \infty$ справедливо равенство пределов:

$$\lim_{n \rightarrow \infty} Val\Gamma_1^{\delta_n}(x_0, y_0, T) = \lim_{n \rightarrow \infty} Val\Gamma_2^{\delta_n}(x_0, y_0, T),$$

где $\delta_n = T/2^n$.

Доказательство. Фиксируем некоторое $n \geq 0$. Пусть $u(\cdot), v(\cdot)$ — пара стратегий в игре $\Gamma_2^{\delta_n}(x_0, y_0, T)$. Эта пара является таковой и в игре $\Gamma_3^{\delta_n}(x_0, y_0, T)$. Пусть в ситуации $u(\cdot), v(\cdot)$ реализуется последовательность $x_0, x_1, \dots, x_{2^n}, y_0, y_1, \dots, y_{2^n}$. Обозначим функции выигрышей в играх $\Gamma_2^{\delta_n}(x_0, y_0, T)$, $\Gamma_3^{\delta_n}(x_0, y_0, T)$ как $K_2(u(\cdot), v(\cdot)) = \rho(x_{2^n}, y_{2^n})$, $K_3(u(\cdot), v(\cdot)) = \rho(x_{2^n}, y_{2^n-1})$ соответственно. Тогда

$$K_2(u(\cdot), v(\cdot)) \leq K_3(u(\cdot), v(\cdot)) + \rho(y_{2^n-1}, y_{2^n}).$$

Тогда в силу произвольности $u(\cdot), v(\cdot)$ имеем:

$$Val\Gamma_2^{\delta_n}(x_0, y_0, T) \leq Val\Gamma_3^{\delta_n}(x_0, y_0, T) + \max_{y \in C_E^{T-\delta_n}(y_0)} \max_{y' \in C_E^{\delta_n}(y)} \rho(y, y'). \quad (5.3.4)$$

Пусть $y_1^{\delta_n} \in C_E^{\delta_n}(y_0)$. Тогда $C_E^{T-\delta_n}(y_1^{\delta_n}) \subset C_E^T(y_0)$. Запишем неравенство (5.3.4) для игр с начальным состоянием $x_0, y_1^{\delta_n}$. Учитывая предыдущее включение, получим

$$Val\Gamma_2^{\delta_n}(x_0, y_1^{\delta_n}, T) \leq Val\Gamma_3^{\delta_n}(x_0, y_1^{\delta_n}, T) + \max_{y \in C_E^T(y_0)} \max_{y' \in C_E^{\delta_n}(y)} \rho(y, y'). \quad (5.3.5)$$

Из определения игр $\Gamma_1^{\delta_n}(x_0, y_0, T)$ и $\Gamma_3^{\delta_n}(x_0, y_0, T)$ вытекает равенство

$$Val\Gamma_1^{\delta_n}(x_0, y_0, T) = \max_{y_1^{\delta_n} \in C_E^{\delta_n}(y_0)} Val\Gamma_3^{\delta_n}(x_0, y_1^{\delta_n}, T). \quad (5.3.6)$$

В силу непрерывности по t функции $C_E^t(y)$ и выполнения условия $C_E^0(y) = y$ второе слагаемое в (5.3.5) стремится к нулю при $n \rightarrow \infty$. Обозначим его через $\varepsilon_1(n)$. Из (5.3.5), (5.3.6) получаем

$$Val\Gamma_1^{\delta_n}(x_0, y_0, T) \geq Val\Gamma_2^{\delta_n}(x_0, y_1^{\delta_n}, T) - \varepsilon_1(n). \quad (5.3.7)$$

В силу непрерывности функции $Val\Gamma_2^{\delta_n}(x_0, y_0, T)$ из (5.3.7) имеем неравенство

$$Val\Gamma_1^{\delta_n}(x_0, y_0, T) \geq Val\Gamma_2^{\delta_n}(x_0, y_0, T) - \varepsilon_1(n) - \varepsilon_2(n), \quad (5.3.8)$$

где $\varepsilon_2(n) \rightarrow 0$ при $n \rightarrow \infty$. Переходя в (5.3.8) к пределу при $n \rightarrow \infty$ (что возможно на основании лемм п. 5.3.3, 5.3.4 и теоремы о существовании предела у монотонной ограниченной последовательности), получаем

$$\lim_{n \rightarrow \infty} Val\Gamma_1^{\delta_n}(x_0, y_0, T) \geq \lim_{n \rightarrow \infty} Val\Gamma_2^{\delta_n}(x_0, y_0, T). \quad (5.3.9)$$

Из леммы п. 5.3.3 вытекает противоположное неравенство. Следовательно, оба предела в (5.3.9) совпадают.

5.3.6. Утверждение теоремы п. 5.3.5 доказано в предположении, что последовательность разбиений интервала $[0, T]$

$$\sigma_n = \{t_0 = 0 < t_1 < \dots < t_N = T\}, \quad n = 1, \dots,$$

удовлетворяет условию $t_{j+1} - t_j = T/2^n$, $j = 0, 1, \dots, 2^n - 1$. Утверждения теоремы п. 5.3.5 и лемм п. 5.3.3, 5.3.4 справедливы для всякой последовательности σ_n измельчающихся разбиений интервала $[0, T]$, т. е. такой, что $\sigma_{n+1} \supset \sigma_n$ (это означает, что разбиение σ_{n+1} получается из σ_n добавлением новых точек)

$$\gamma(\sigma_n) = \max_i (t_{i+1} - t_i) \rightarrow_{n \rightarrow +\infty} 0.$$

Рассмотрим теперь такие любые последовательности разбиений $\{\sigma_n\}$ и $\{\sigma'_n\}$ промежутка $[0, T]$.

Лемма. *Имеет место равенство*

$$\lim_{n \rightarrow \infty} \text{Val} \Gamma_1^{\sigma_n}(x_0, y_0, T) = \lim_{n \rightarrow \infty} \text{Val} \Gamma_1^{\sigma'_n}(x_0, y_0, T),$$

где $x_0, y_0 \in R^n$. $T < \infty$.

Доказательство проведем от противного. Допустим, что утверждение леммы неверно, и предположим для определенности, что выполняется неравенство

$$\lim_{n \rightarrow \infty} \text{Val} \Gamma_1^{\sigma_n}(x_0, y_0, T) > \lim_{n \rightarrow \infty} \text{Val} \Gamma_1^{\sigma'_n}(x_0, y_0, T).$$

Тогда, по теореме 5.3.5, имеем

$$\lim_{n \rightarrow \infty} \text{Val} \Gamma_1^{\sigma_n}(x_0, y_0, T) > \lim_{n \rightarrow \infty} \text{Val} \Gamma_2^{\sigma'_n}(x_0, y_0, T).$$

Отсюда найдутся натуральные числа m_1, n_1 такие, что выполнено неравенство:

$$\text{Val} \Gamma_1^{\sigma_{m_1}}(x_0, y_0, T) > \text{Val} \Gamma_2^{\sigma'_{n_1}}(x_0, y_0, T).$$

Обозначим через $\bar{\sigma}$ разбиение интервала $[0, T]$ точками, принадлежащими как разбиению σ_{m_1} , так и разбиению σ'_{n_1} . Для него выполняется неравенство

$$\text{Val} \Gamma_2^{\bar{\sigma}}(x_0, y_0, T) \leq \text{Val} \Gamma_2^{\sigma'_{n_1}}(x_0, y_0, T) < \text{Val} \Gamma_1^{\sigma_{m_1}}(x_0, y_0, T) \leq \text{Val} \Gamma_1^{\bar{\sigma}}(x_0, y_0, T).$$

Откуда

$$\text{Val} \Gamma_2^{\bar{\sigma}}(x_0, y_0, T) < \text{Val} \Gamma_1^{\bar{\sigma}}(x_0, y_0, T).$$

Это противоречит (5.3.3), следовательно, сделанное предположение неверно и утверждение леммы справедливо.

5.3.7. Теорема. *При всех $x_0, y_0, T < \infty$ в игре $\Gamma(x_0, y_0, T)$ существует ситуация ε -равновесия для любого $\varepsilon > 0$. При этом*

$$\text{Val} \Gamma(x_0, y_0, T) = \lim_{n \rightarrow \infty} \text{Val} \Gamma_1^{\sigma_n}(x_0, y_0, T), \quad (5.3.10)$$

где $\{\sigma_n\}$ — любая последовательность измельчающихся разбиений интервала $[0, T]$.

Доказательство. Зададим произвольно выбранное число $\varepsilon > 0$ и покажем, что найдутся такие стратегии $u_\varepsilon(\cdot)$ и $v_\varepsilon(\cdot)$ игроков P и E соответственно, что для всех стратегий $u(\cdot) \in P$ и $v(\cdot) \in E$ выполняются неравенства:

$$K(x_0, y_0, u_\varepsilon(\cdot), v(\cdot)) - \varepsilon \leq K(x_0, y_0, u_\varepsilon(\cdot), v_\varepsilon(\cdot)) \leq K(x_0, y_0, u(\cdot), v_\varepsilon(\cdot)) + \varepsilon. \quad (5.3.11)$$

В силу теоремы п. 5.3.5 найдется такое разбиение σ интервала $[0, T]$, что

$$Val\Gamma_2^\sigma(x_0, y_0, T) - \lim_{n \rightarrow \infty} Val\Gamma_2^{\sigma_n}(x_0, y_0, T) < \varepsilon/2,$$

$$\lim_{n \rightarrow +\infty} Val\Gamma_1^{\sigma_n}(x_0, y_0, T) - Val\Gamma_1^\sigma(x_0, y_0, T) < \varepsilon/2.$$

Пусть $u^\varepsilon(\cdot) = (\sigma, a_{u^\varepsilon})$, $v^\varepsilon(\cdot) = (\sigma, b_{v^\varepsilon})$, где a_{u^ε} , b_{v^ε} — оптимальные стратегии игроков P и E , соответственно, в играх $\Gamma_2^\sigma(x_0, y_0, T)$ и $\Gamma_1^\sigma(x_0, y_0, T)$.

Тогда справедливы соотношения:

$$\begin{aligned} K(x_0, y_0, u^\varepsilon(\cdot), v(\cdot)) &\leq Val\Gamma_2^\sigma(x_0, y_0, T) \\ &< \lim_{n \rightarrow \infty} Val\Gamma_2^{\sigma_n}(x_0, y_0, T) + \frac{\varepsilon}{2}, \quad v(\cdot) \in E, \end{aligned} \quad (5.3.12)$$

$$\begin{aligned} K(x_0, y_0, u(\cdot), v^\varepsilon(\cdot)) &\geq Val\Gamma_1^\sigma(x_0, y_0, T) \\ &> \lim_{n \rightarrow \infty} Val\Gamma_1^{\sigma_n}(x_0, y_0, T) - \frac{\varepsilon}{2}, \quad u(\cdot) \in P. \end{aligned} \quad (5.3.13)$$

Из соотношений (5.3.12), (5.3.13) и теоремы 5.3.5, получаем

$$-\frac{\varepsilon}{2} < K(x_0, y_0, u^\varepsilon(\cdot), v^\varepsilon(\cdot)) - \lim_{n \rightarrow \infty} Val\Gamma_1^{\sigma_n}(x_0, y_0, T) < \frac{\varepsilon}{2}. \quad (5.3.14)$$

Из (5.3.12)–(5.3.14) следует (5.3.11). В силу произвольности ε из (5.3.14) следует (5.3.10). Теорема доказана.

5.3.8. Замечание. При доказательстве теоремы существования нигде не был использован специфический вид выигрыша $\rho(x(T), y(T))$. Существенной является лишь непрерывная зависимость выигрыша от реализованных траекторий. Поэтому теорема п. 5.3.7 остается справедливой, если вместо $\rho(x(T), y(T))$ рассмотреть любой непрерывный функционал траекторий $(x(t), y(t))$. В частности, таким функционалом может быть $\min_{0 \leq t < T} \rho(x(t), y(t))$ т. е. минимальное расстояние между игроками в процессе игры. Поэтому результат данного параграфа остается в силе и для дифференциальной игры преследования на достижение минимального результата с предписанной продолжительностью.

§ 5.4. Дифференциальные игры преследования на быстроедействие

5.4.1. Дифференциальные игры преследования на быстроедействие представляют собой частный случай дифференциальных игр с интегральным выигрышем, определенных в п. 5.1.8. Классы стратегий игроков P и E те же, что и в игре с предписанной продолжительностью. Предположим, что в $R^n \times R^n$ задано множество $F = \{(x, y) : \rho(x, y) \leq l, l > 0\}$, и пусть $x(t), y(t)$ — траектории игроков P и E в ситуации $(u(\cdot), v(\cdot))$ из начальных состояний x_0, y_0 . Обозначим

$$t_n(x_0, y_0; u(\cdot), v(\cdot)) = \min\{t : (x(t), y(t)) \in F\}. \quad (5.4.1)$$

Если не существует такого t , что $(x(t), y(t)) \in F$, то $t_n(x_0, y_0; u(\cdot), v(\cdot))$ полагается равным $+\infty$. В дифференциальной игре преследования на быстроедействие выигрыш игрока E полагают равным

$$K(x_0, y_0; u(\cdot), v(\cdot)) = t_n(x_0, y_0; u(\cdot), v(\cdot)). \quad (5.4.2)$$

Игра зависит от начальных состояний x_0, y_0 , поэтому будем обозначать ее через $\Gamma(x_0, y_0)$.

Из определения функции выигрыша (5.4.2) следует, что в игре $\Gamma(x_0, y_0)$ целью игрока E является максимизация времени сближения с игроком P на заданное расстояние $l \geq 0$. Игрок P , наоборот, стремится минимизировать это время.

5.4.2. Между игрой преследования на быстродействие $\Gamma(x_0, y_0, T)$ и игрой преследования с предписанной продолжительностью на достижение минимального результата существует прямая связь. Пусть $\Gamma(x_0, y_0, T)$ — игра преследования с предписанной продолжительностью T на достижение минимального результата (выигрыш игрока E равен $\min_{0 \leq t < T} \rho(x(t), y(t))$). Было показано, что для игр этого типа при любом $\varepsilon > 0$ в классе кусочно-программных стратегий существует ситуация ε -равновесия (см. п. 5.3.8). Пусть $V(x_0, y_0, T)$ — значение такой игры $\Gamma(x_0, y_0, T)$, а $V(x_0, y_0)$ — значение игры $\Gamma(x_0, y_0)$, если оно существует.

Лемма. При фиксированных x_0, y_0 функция $V(x_0, y_0, T)$ непрерывна и не возрастает по T на $[0, \infty]$.

Доказательство. Пусть $T_1 > T_2 > 0$. Обозначим через $v_\varepsilon^{T_1}$ стратегию игрока E в игре $\Gamma(x_0, y_0, T)$, которая гарантирует игроку E , что расстояние между ним и игроком P на отрезке $[0, T_1]$ будет не меньше $\max[0, V(x_0, y_0, T_1) - \varepsilon]$. Следовательно, она тем более гарантирует расстояние $\max[0, V(x_0, y_0, T_1) - \varepsilon]$ между ними на отрезке $[0, T_2]$, где $T_2 < T_1$. Следовательно,

$$V(x_0, y_0, T_2) \geq \max[0, V(x_0, y_0, T_1) - \varepsilon] \quad (5.4.3)$$

(ε -оптимальная в игре $\Gamma(x_0, y_0, T_1)$ стратегия не обязательно ε -оптимальна в игре $\Gamma(x_0, y_0, T_2)$). Поскольку ε может быть выбрано произвольным, из (5.4.3) следует второе утверждение леммы. Непрерывность $V(x_0, y_0, T)$ по T доказывать не будем. Отметим лишь, что это свойство можно получить, используя непрерывность $V(x_0, y_0, T)$ по x_0, y_0 .

5.4.3. Рассмотрим уравнение

$$V(x_0, y_0, T) = l \quad (5.4.4)$$

относительно T . Возможны следующие три случая:

- 1) уравнение (5.4.4) не имеет корней;
- 2) имеет единственный корень;
- 3) имеет более одного корня.

В случае 3) из невозрастания и непрерывности функции $V(x_0, y_0, T)$ по T следует, что уравнение (5.4.4) имеет целый сегмент корней, т. е. функция $V(x_0, y_0, T)$, как функция от T имеет промежуток постоянства.

Рассмотрим каждый случай отдельно.

Случай 1. В этом случае возможно:

- а) $V(x_0, y_0, T) < l$ для всех $T \geq 0$;
- б) $\inf_{T \geq 0} V(x_0, y_0, T) > l$;
- в) $\inf_{T \geq 0} V(x_0, y, T) = l$.

В случае а) имеем:

$$V(x_0, y_0, 0) = \rho(x_0, y_0) < l,$$

т. е. $t_n(x_0, y_0; u(\cdot), v(\cdot)) = 0$ для всех $u(\cdot), v(\cdot)$. Тогда значение игры $\Gamma(x_0, y_0)$ равно $V(x_0, y_0) = 0$.

В случае б) выполняется равенство:

$$\inf_{T \geq 0} V(x_0, y_0, T) = \lim_{T \rightarrow \infty} V(x_0, y_0, T) > l.$$

Отсюда для любого $T > 0$ (сколь угодно большого) у игрока E найдется соответствующая стратегия $v^T(\cdot) \in E$, которая гарантирует ему избежание l -встречи на отрезке $[0, T]$. Но тогда игрок P не имеет стратегии, которая бы гарантировала ему l -встречу с игроком E за конечное время. В то же время нельзя утверждать, что игрок E обладает стратегией, гарантирующей избежание l -встречи за любое время. Вопрос о нахождении начальных состояний, в которых такая стратегия существует, сводится к решению игры качества для игрока E . Таким образом, при $l < \lim_{T \rightarrow \infty} V(x_0, y_0, T)$ можно лишь утверждать, что значение игры $\Gamma(x_0, y_0)$, если оно существует, больше любого наперед заданного T , т. е. равно $+\infty$.

в) рассмотрим совместно со случаем 3).

Случай 2. Пусть T_0 — единственный корень уравнения (5.4.4). Тогда из невозрастания и непрерывности по T функции $V(x_0, y_0, T)$ следует, что

$$V(x_0, y_0, T) > V(x_0, y_0, T_0) \text{ при всех } T < T_0,$$

$$V(x_0, y_0, T) < V(x_0, y_0, T_0) \text{ при всех } T > T_0, \quad (5.4.5)$$

$$\lim_{T \rightarrow T_0} V(x_0, y_0, T) = V(x_0, y_0, T_0). \quad (5.4.6)$$

Фиксируем произвольное $T > T_0$. Рассмотрим игру преследования $\Gamma(x_0, y_0, T)$. Она обладает ситуацией ε -равновесия в классе кусочно-программных стратегий для любого $\varepsilon > 0$. Это означает, в частности, что для любого $\varepsilon > 0$ существует стратегия $u_\varepsilon(\cdot) \in P$ игрока P , которая гарантирует ему сближение с игроком E на расстояние $V(x_0, y_0, T) + \varepsilon$, т. е.

$$K(u_\varepsilon(\cdot), v(\cdot)) \leq V(x_0, y_0, T) + \varepsilon, \quad v(\cdot) \in E, \quad (5.4.7)$$

где $K(u(\cdot), v(\cdot))$ — функция выигрыша в игре $\Gamma(x_0, y_0, T)$. Тогда из (5.4.5), (5.4.6) следует существование $\bar{\varepsilon} > 0$ такого, что для любого $\varepsilon < \bar{\varepsilon}$ найдется число $\tilde{T}(\varepsilon)$, $T_0 < \tilde{T}(\varepsilon) \leq T$, для которого

$$\varepsilon = V(x_0, y_0, T_0) - V(x_0, y_0, \tilde{T}(\varepsilon)). \quad (5.4.8)$$

Из (5.4.7), (5.4.8) следует, что для любого $\varepsilon < \bar{\varepsilon}$

$$K(u_\varepsilon(\cdot), v(\cdot)) \leq V(x_0, y_0, T) + \varepsilon \leq V(x_0, y_0, \tilde{T}(\varepsilon)) + \varepsilon = V(x_0, y_0, T_0) = l, \quad v(\cdot) \in E,$$

т. е. стратегия $u_\varepsilon(\cdot)$ обеспечивает l -встречу за время T . Отсюда, в силу произвольности $T > T_0$, следует, что для любого $T > T_0$ найдется отвечающая ему стратегия $u^T(\cdot) \in P$, которая гарантирует l -встречу за время T . Иными словами, для любого $\delta > 0$ существует $u_\delta(\cdot) \in P$ такая, что

$$t_n(x_0, y_0; u_\delta(\cdot), v(\cdot)) \leq T_0 + \delta \text{ для всех } v(\cdot) \in E. \quad (5.4.9)$$

Аналогично доказывается существование $v_\delta(\cdot) \in E$ такого, что

$$t_n(x_0, y_0; u(\cdot), v_\delta(\cdot)) \geq T_0 - \delta \text{ для всех } u(\cdot) \in P. \quad (5.4.10)$$

Из (5.4.9), (5.4.10) следует, что в игре преследования на быстродействие $\Gamma(x_0, y_0)$ для любого $\varepsilon > 0$ существует ситуация ε -равновесия в кусочно-программных стратегиях и значение игры равно T_0 , где T_0 — единственный корень уравнения (5.4.4).

Случай 3. Обозначим через T_0 минимальный корень уравнения (5.4.4). Теперь, вообще говоря, мы не можем утверждать, что значение игры $Val\Gamma(x_0, y_0) = T_0$. Действительно, из $V(x_0, y_0, T_0) = l$ следует лишь, что в игре $\Gamma(x_0, y_0, T_0)$ для любого $\varepsilon > 0$ у игрока P существует стратегия $u_\varepsilon(\cdot)$, гарантирующая ему за время T_0 встречу с игроком E на расстоянии не более чем $l + \varepsilon$, а из существования более одного корня уравнения (5.4.4) и монотонности $V(x_0, y_0, T)$ по T получаем существование промежутка постоянства функции $V(x_0, y_0, T)$ по $T \in [T_0, T_1]$. Поэтому увеличение продолжительности игры $\Gamma(x_0, y_0, T_0)$ на δ , где $\delta < T_1 - T_0$, не приводит к уменьшению гарантированного сближения с игроком E , т. е. для всех $T \in [T_0, T_1]$ игрок P может лишь обеспечить сближение с игроком E на расстояние $l + \varepsilon$ (для любого $\varepsilon > 0$), и нет основания считать, что при каком-то $T \in [T_0, T_1]$ величина ε окажется равной нулю. Если бы в игре $\Gamma(x_0, y_0, T_0)$ существовала ситуация равновесия (а не ситуация ε -равновесия), то значение игры $\Gamma(x_0, y_0)$ было бы равно T_0 и в случае 3.

5.4.4. Модифицируем понятие ситуации равновесия в игре $\Gamma(x_0, y_0)$. Далее в этом параграфе удобнее использовать запись $\Gamma(x_0, y_0, l)$ вместо $\Gamma(x_0, y_0)$, подчеркивая, что игра $\Gamma(x_0, y_0, l)$ заканчивается при сближении игроков на расстояние l .

Пусть $t_n^l(x_0, y_0; u(\cdot), v(\cdot))$ — время до момента сближения на расстояние l в ситуации $(u(\cdot), v(\cdot))$ и заданы $\varepsilon \geq 0$, $\delta \geq 0$.

Определение. Будем говорить, что пара стратегий $\bar{u}_\varepsilon^\delta(\cdot), \bar{v}_\varepsilon^\delta(\cdot)$ образует ситуацию ε, δ -равновесия в игре $\Gamma(x_0, y_0, l)$, если

$$t_n^{l+\delta}(x_0, y_0; u(\cdot), \bar{v}_\varepsilon^\delta(\cdot)) + \varepsilon \geq t_n^{l+\delta}(x_0, y_0; \bar{u}_\varepsilon^\delta(\cdot), \bar{v}_\varepsilon^\delta(\cdot)) \geq t_n^{l+\delta}(x_0, y_0; \bar{u}_\varepsilon^\delta(\cdot), v(\cdot)) - \varepsilon$$

для всех стратегий $u(\cdot) \in P$, $v(\cdot) \in E$.

Определение. Пусть существует такая последовательность $\{\delta_k\}$, $\delta_k \geq 0$, $\delta_k \rightarrow 0$, что во всех играх $\Gamma(x_0, y_0, l + \delta_k)$ для любого $\varepsilon > 0$ существуют ситуации ε -равновесия. Тогда предел

$$\lim_{k \rightarrow \infty} V(x_0, y_0, l + \delta_k) = V'(x_0, y_0, l)$$

называется значением игры $\Gamma(x_0, y_0, l)$ в обобщенном смысле.

Заметим, что величина $V'(x_0, y_0, l)$ не зависит от выбора последовательности $\{\delta_k\}$ вследствие монотонного убывания функции $V(x_0, y_0, l)$ по l .

Определение. Будем говорить, что игра $\Gamma(x_0, y_0, l)$ имеет значение в обобщенном смысле, если существует такая последовательность $\{\delta_k\}$, $\delta_k \rightarrow 0$, что для любого $\varepsilon > 0$ и $\delta_k \in \{\delta_k\}$ в игре $\Gamma(x_0, y_0, l)$ существует ε, δ_k -равновесие.

Можно показать, что если игра $\Gamma(x_0, y_0, l)$ имеет значение в обычном смысле, то ее значение $V'(x_0, y_0, l)$ (в обобщенном смысле) существует и равно

$$\lim_{\varepsilon \rightarrow 0, \delta_k \rightarrow 0} t_n^{l+\delta_k}(x_0, y_0; \bar{u}_\varepsilon^\delta(\cdot), \bar{v}_\varepsilon^\delta(\cdot)) = V'(x_0, y_0, l).$$

Из определения значения и решения игры $\Gamma(x_0, y_0, l)$ (в обобщенном смысле) вытекает, что если в игре $\Gamma(x_0, y_0, l)$ для любого $\varepsilon > 0$ существует ε -ситуация равновесия в обычном смысле (т. е. решение в обычном смысле), то $V(x_0, y_0, l) = V'(x_0, y_0, l)$ (достаточно взять последовательность $\delta_k \equiv 0$ для всех k).

Теорема. Пусть уравнение (5.4.4) имеет более одного корня и T_0 — наименьший корень, $T_0 < \infty$. Тогда существует значение $V'(x_0, y_0, l)$ (в обобщенном смысле) игры преследования на быстродействие $\Gamma(x_0, y_0, l)$ и $V'(x_0, y_0, l) = T_0$.

Доказательство. Из монотонности и непрерывности функции $V(x_0, y_0, T)$ по T следует существование такой последовательности $T_k \rightarrow T_0$ слева, что $V(x_0, y_0, T_k) \rightarrow V(x_0, y_0, T_0) = l$ и функция $V(x_0, y_0, T_k)$ строго монотонна в точках T_k . Пусть

$$\delta_k = V(x_0, y_0, T_k) - l \geq 0.$$

Из строгой монотонности функции $V(x_0, y_0, T)$ в точках T_k вытекает, что уравнение $V(x_0, y_0, T) = l + \delta_k$ имеет единственный корень T_k . Это означает, что для любого $\delta_k \in \{\delta_k\}$ в играх $\Gamma(x_0, y_0, l + \delta_k)$ существует ситуация ε -равновесия для любого $\varepsilon > 0$ (см. случай 2 в п. 5.4.3). Значит, в игре $\Gamma(x_0, y_0, l)$ существует решение в обобщенном смысле:

$$\lim_{k \rightarrow \infty} V(x_0, y_0, l + \delta_k) = \lim_{k \rightarrow \infty} T_k = T_0 = V'(x_0, y_0, l).$$

Теорема доказана.

Рассмотрим теперь случай с) п. 5.4.3. Имеем: $\inf_T V(x_0, y_0, T) = l$. Пусть $T_k \rightarrow \infty$. Тогда $\lim_{k \rightarrow \infty} V(x_0, y_0, T_k) = l$. Из монотонности и непрерывности $V(x_0, y_0, T)$ по T следует, что последовательность $\{T_k\}$ можно выбрать так, что в точках T_k функция $V(x_0, y_0, T)$ строго монотонна. Тогда как и при доказательстве теоремы п. 5.4.4 можно показать, что существует такая последовательность $\{\delta_k\}$, что

$$\lim_{k \rightarrow \infty} V(x_0, y_0, l + \delta_k) = \lim_{k \rightarrow \infty} T_k = T_0 = \infty.$$

Таким образом, и в данном случае обобщенное решение существует, а обобщенное значение игры $\Gamma(x_0, y_0, l)$ равно бесконечности.

5.4.5. Часто оказывается важным определить, может ли игрок P гарантировать l -встречу из данных начальных позиций x, y за фиксированное время T , а если это невозможно, то может ли игрок E гарантировать избежание l -встречи в течение заданного времени.

Пусть $V(x, y, T)$ — значение игры с предписанной продолжительностью T из начальных состояний $x, y \in R^n$ с выигрышем $\min_{0 \leq t \leq T} \rho(x(t), y(t))$. Тогда возможны следующие альтернативы: 1) $V(x, y, T) > l$; 2) $V(x, y, T) \leq l$.

Случай 1. Из определения функции $V(x, y, T)$ следует, что для любого $\varepsilon > 0$ найдется такая стратегия игрока E , что для всех стратегий $u(\cdot)$ справедливо неравенство

$$K(x, y; u(\cdot), v_\varepsilon^*(\cdot)) \geq V(x, y, T) - \varepsilon.$$

Выбрав ε достаточно малым, можно добиться выполнения неравенства

$$K(x, y; u(\cdot), v_\varepsilon^*(\cdot)) \geq V(x, y, T) - \varepsilon > l$$

для всех стратегий $u(\cdot) \in P$ игрока P . Из вида функции выигрыша K следует, что, используя стратегию $v_\varepsilon^*(\cdot)$, игрок E может гарантировать выполнение неравенства $\min_{0 \leq t \leq T_0} \rho(x(t), y(t)) > l$ независимо от действий игрока P , т. е. в рассматриваемом случае игрок E гарантирует избежание l -встречи на отрезке времени $[0, T]$ независимо от действий игрока P .

Случай 2. Пусть T_0 — минимальный корень уравнения $V(x, y, T) = l$ при фиксированных x, y (если $\rho(x, y) < l$, то T_0 полагаем равным 0). Тогда из определения $V(x, y, T_0)$

следует, что в игре $\Gamma(x, y, T_0)$ для любого $\varepsilon > 0$ у игрока P существует стратегия u_ε^* , гарантирующая выполнение неравенства

$$K(x, y; u_\varepsilon^*(\cdot), v(\cdot)) \leq V(x, y; T_0) + \varepsilon = l + \varepsilon$$

для всех стратегий $v(\cdot) \in E$ игрока E . Из вида функции выигрыша K следует, что, используя стратегию $u_\varepsilon^*(\cdot)$, игрок P может гарантировать выполнение неравенства $\min_{0 \leq t \leq T} \rho(x(t), y(t)) \leq l + \varepsilon$ независимо от действий игрока E . Продолжая произвольным образом стратегию $u_\varepsilon^*(\cdot)$ на отрезок $[T_0, T]$, получаем, что в случае 2 игрок P при любом $\varepsilon > 0$ может гарантировать $(l + \varepsilon)$ -встречу с игроком E за время T независимо от действий последнего.

Фактически доказана следующая теорема (об альтернативе).

Теорема. Для любых $x, y \in R^n$, $T > 0$ справедливо, по крайней мере, одно из следующих утверждений:

- 1) из начальных состояний x, y игрок E может в течение времени T гарантировать избежание l -встречи независимо от действий игрока P ;
- 2) для любого $\varepsilon > 0$ игрок P может гарантировать $(l + \varepsilon)$ -встречу с игроком E из начальных состояний x, y за время T независимо от действий игрока E .

5.4.6. Для каждого фиксированного $T > 0$ все пространство $R^n \times R^n$ делится на три непересекающиеся области: область $A = \{x, y : V(x, y, T) < l\}$, которую будем называть зоной захвата; область $B = \{x, y : V(x, y, T) > l\}$, которую естественно назвать зоной избежания захвата, и область $C = \{x, y : V(x, y, T) = l\}$ — зона нейтрального исхода.

Пусть $x, y \in A$. По определению A , при любом $\varepsilon > 0$ игрок P обладает такой стратегией $u_\varepsilon^*(\cdot)$, что

$$K(x, y; u_\varepsilon^*(\cdot), v(\cdot)) \leq V(x, y, T) + \varepsilon$$

при всех стратегиях $v(\cdot)$ игрока E . Выбрав подходящим образом $\varepsilon > 0$, можно обеспечить выполнение неравенства:

$$K(x, y; u_\varepsilon^*(\cdot), v(\cdot)) \leq V(x, y, T) + \varepsilon < l.$$

Последнее означает, что стратегия u_ε^* игрока P гарантирует ему l -встречу с игроком E из начальных состояний x, y за время T .

В результате получаем следующее уточнение теоремы п. 5.4.5.

Теорема. Для любого фиксированного $T > 0$ все пространство делится на три непересекающиеся области A, B, C , обладающие следующими свойствами:

- 1) при любых $x, y \in A$ игрок P обладает стратегией $u_\varepsilon^*(\cdot)$, которая гарантирует l -встречу с игроком E на отрезке $[0, T]$ независимо от действий последнего;
- 2) для $x, y \in B$ игрок E обладает стратегией $v_\varepsilon^*(\cdot)$, которая гарантирует избежание l -встречи с игроком P на отрезке $[0, T]$ независимо от действий последнего;
- 3) если $x, y \in C$ и $\varepsilon > 0$, то игрок P обладает стратегией $u_\varepsilon^*(\cdot)$, гарантирующей $(l + \varepsilon)$ -встречу с игроком E за время T независимо от действий последнего.

§ 5.5. Необходимые и достаточные условия существования оптимальной программной стратегии убегающего

5.5.1. Важным подклассом игр преследования являются игры, в которых оптимальная стратегия убегающего игрока является только функцией времени (так называемый регулярный случай).

Ограничимся рассмотрением игры преследования с предписанной продолжительностью, хотя все результаты могут быть перенесены и на игры преследования по быстродействию. Пусть $C_P^T(x)(C_E^T(y))$ — множество достижимости игрока $P(E)$ из начального состояния $x(y)$ к моменту времени T , т. е. множество тех позиций, в которые может попасть игрок $P(E)$ из начального состояния $x(y)$ в момент T , используя всевозможные измеримые программные управления $u(t), (v(t))$, $t \in [0, T]$ при условии, что движение происходит в соответствии с системой $\dot{x} = f(x, u)$ ($\dot{y} = g(y, v)$). Введем в рассмотрение величину

$$\hat{\rho}_T(x_0, y_0) = \max_{y \in C_E^T(y_0)} \min_{x \in C_P^T(x_0)} \rho(x, y), \quad (5.5.1)$$

называемую иногда (см. [Красовский, 1970; Красовский, Субботин, 1974; Красовский, 1985]) *гипотетическим рассогласованием множеств* $C_E^T(y_0)$ и $C_P^T(x_0)$ (см. пример 6 в п. 2.2.6).

Функция $\hat{\rho}_T(x_0, y_0)$ обладает следующими свойствами:

1⁰. $\hat{\rho}_T(x_0, y_0) \geq 0$, $\hat{\rho}_T(x_0, y_0)|_{T=0} = \rho(x_0, y_0)$;

2⁰. $\hat{\rho}_T(x_0, y_0) = 0$, если $C_P^T(x_0) \supset C_E^T(y_0)$;

3⁰. Если $V(x_0, y_0, T)$ — значение игры $\Gamma(x_0, y_0, T)$ с предписанной продолжительностью и терминальным выигрышем $\rho(x(T), y(T))$, то

$$V(x_0, y_0, T) \geq \hat{\rho}_T(x_0, y_0).$$

Действительно, свойство 1⁰ следует из неотрицательности функции $\rho(x, y)$. Пусть $C_P^T(x_0) \supset C_E^T(y_0)$. Тогда для любого $y' \in C_E^T(y_0)$ существует такое $x' \in C_P^T(x_0)$, что $\rho(x', y') = 0$, ($x' = y'$), откуда следует 2⁰. Свойство 3⁰ следует из того, что игрок E , выбирая направление движения на точку $M \in C_E^T(y_0)$, для которой

$$\hat{\rho}_T(x_0, y_0) = \min_{x \in C_P^T(x_0)} \rho(x, M),$$

всегда гарантирует получение выигрыша $\hat{\rho}_T(x_0, y_0)$. Точка M называется *центром преследования*.

5.5.2. Пусть $\Gamma_\delta(x_0, y_0, T)$ — дискретная игра преследования с шагом δ ($\delta = t_{k+1} - t_k$), предписанной продолжительностью T , дискриминацией игрока E и начальными состояниями x_0, y_0 . Тогда справедлива следующая теорема.

Теорема. Для того, чтобы для любых $x_0, y_0 \in R^n$ и $T = \delta \cdot k$, $k = 1, 2, \dots$, выполнялось равенство:

$$\hat{\rho}_T(x_0, y_0) = \text{Val} \Gamma_\delta(x_0, y_0, T), \quad (5.5.2)$$

необходимо и достаточно, чтобы для всех $x_0, y_0 \in R^n$, $\delta > 0$ и $T = \delta \cdot k$, $k = 1, 2, \dots$, выполнялось соотношение

$$\hat{\rho}_T(x_0, y_0) = \max_{y \in C_E^\delta(y_0)} \min_{x \in C_P^\delta(x_0)} \hat{\rho}_{T-\delta}(x, y) \quad (5.5.3)$$

($\text{Val} \Gamma_\delta(x_0, y_0, T)$ — значение игры $\Gamma_\delta(x_0, y_0, T)$).

Доказательство теоремы опирается на следующий результат.

Лемма. Следующее неравенство справедливо для любых $x_0, y_0 \in R^n$, $T \geq \delta$:

$$\hat{\rho}_T(x_0, y_0) \leq \max_{y \in C_E^\delta(y_0)} \min_{x \in C_P^\delta(x_0)} \hat{\rho}_{T-\delta}(x, y).$$

Доказательство. По определению функции $\hat{\rho}_T$ имеем

$$\max_{y \in C_E^\delta(y_0)} \min_{x \in C_P^\delta(x_0)} \hat{\rho}_{T-\delta}(x, y) = \max_{y \in C_E^\delta(y_0)} \min_{x \in C_P^\delta(x_0)} \max_{\bar{y} \in C_E^{T-\delta}(y)} \min_{\bar{x} \in C_P^{T-\delta}(x)} \rho(\bar{x}, \bar{y}).$$

Для всех $x \in C_P^\delta(x_0)$ справедливо включение $C_P^{T-\delta}(x) \subset C_P^T(x_0)$. Следовательно, для любых $x \in C_P^\delta(x_0)$, $\bar{y} \in C_E^{T-\delta}(y)$.

$$\min_{\bar{x} \in C_P^{T-\delta}(x)} \rho(\bar{x}, \bar{y}) \geq \min_{\bar{x} \in C_P^T(x_0)} \rho(\bar{x}, \bar{y}).$$

Тогда для всех $x \in C_P^\delta(x_0)$, $y \in C_E^\delta(y_0)$

$$\max_{\bar{y} \in C_E^{T-\delta}(y)} \min_{\bar{x} \in C_P^{T-\delta}(x)} \rho(\bar{x}, \bar{y}) \geq \max_{\bar{y} \in C_E^{T-\delta}(y)} \min_{\bar{x} \in C_P^T(x_0)} \rho(\bar{x}, \bar{y})$$

и

$$\min_{x \in C_P^\delta(x_0)} \max_{\bar{y} \in C_E^{T-\delta}(y)} \min_{\bar{x} \in C_P^{T-\delta}(x)} \rho(\bar{x}, \bar{y}) \geq \max_{\bar{y} \in C_E^{T-\delta}(y)} \min_{\bar{x} \in C_P^T(x_0)} \rho(\bar{x}, \bar{y}).$$

Беря от обоих частей этого неравенства $\max_{y \in C_E^\delta(y_0)}$, получаем

$$\begin{aligned} \max_{y \in C_E^\delta(y_0)} \min_{x \in C_P^\delta(x_0)} \hat{\rho}_{T-\delta}(x, y) &\geq \max_{y \in C_E^\delta(y_0)} \max_{\bar{y} \in C_E^{T-\delta}(y)} \min_{\bar{x} \in C_P^T(x_0)} \rho(\bar{x}, \bar{y}) \\ &= \max_{y \in C_E^\delta(y_0)} \min_{x \in C_P^T(x_0)} \rho(x, y) = \hat{\rho}_T(x_0, y_0). \end{aligned}$$

Лемма доказана.

Перейдем к доказательству теоремы.

Необходимость. Пусть выполняется условие (5.5.2) и не выполняется условие (5.5.3). Тогда согласно лемме существуют такие $\delta > 0$, $x_0, y_0 \in R^n$, $T_0 = \delta k_0$, $k_0 \geq 1$, что

$$\hat{\rho}_{T_0}(x_0, y_0) < \max_{y \in C_E^\delta(y_0)} \min_{x \in C_P^\delta(x_0)} \hat{\rho}_{T_0-\delta}(x, y). \quad (5.5.4)$$

Пусть $u^0(\cdot)$ — оптимальная стратегия игрока P в игре $\Gamma_\delta(x_0, y_0, T_0)$ и на 1-м шаге игры игрок E выбирает точку $y^* \in C_E^\delta(y_0)$, для которой

$$\min_{x \in C_P^\delta(x_0)} \hat{\rho}_{T_0-\delta}(x, y^*) = \max_{y \in C_E^\delta(y_0)} \min_{x \in C_P^\delta(x_0)} \hat{\rho}_{T_0-\delta}(x, y). \quad (5.5.5)$$

Пусть $x^0(\delta)$ — состояние, в которое переходит P на 1-м шаге при использовании стратегии $u^0(\cdot)$, а $\bar{v}^0(\cdot)$ — оптимальная стратегия игрока E в игре $\Gamma_\delta(x_0(\delta), y^*, T_0 - \delta)$. Рассмотрим следующую стратегию $\tilde{v}(\cdot)$ игрока E в игре $\Gamma_\delta(x_0, y_0, T_0)$: в момент $t = 0$ он выбирает точку y^* , а начиная с момента $t = \delta$ использует стратегию $\bar{v}^0(\cdot)$.

Обозначим через $\hat{u}^0(\cdot)$ сужение стратегии $u^0(\cdot)$ на отрезке $[\delta, T_0]$. Из (5.5.2), (5.5.4), (5.5.5) (согласно формуле (5.5.2) $\hat{\rho}_T(x_0, y_0)$ — значение игры $\Gamma_\delta(x_0, y_0, T)$) находим

$$\begin{aligned} \hat{\rho}_{T_0}(x_0, y_0) &\geq K(u^0(\cdot), \tilde{v}(\cdot); x_0, y_0, T_0) = K(\hat{u}^0(\cdot), \bar{v}^0(\cdot); x^0(\delta), y^*, T_0 - \delta) = \\ &= \hat{\rho}_{T_0-\delta}(x^0(\delta), y^*) \geq \min_{x \in C_P^\delta(x_0)} \hat{\rho}_{T_0-\delta}(x, y^*) = \max_{y \in C_E^\delta(y_0)} \min_{x \in C_P^\delta(x_0)} \hat{\rho}_{T_0-\delta}(x, y) > \hat{\rho}_{T_0}(x_0, y_0). \end{aligned}$$

Полученное противоречие доказывает необходимость условия (5.5.3).

Достаточность. Заметим, что условие (5.5.3) совместно с условием $\hat{\rho}_T(x_0, y_0)|_{T=0} = \rho(x_0, y_0)$, показывает, что функция $\hat{\rho}_T(x_0, y_0)$ удовлетворяет функциональному уравнению для функции значения игры $\Gamma_\delta(x_0, y_0, T)$. Как следует из доказательства теоремы п. 5.2.2, это условие является достаточным для того, чтобы $\hat{\rho}_T(x_0, y_0)$ было значением игры $\Gamma_\delta(x_0, y_0, T)$.

5.5.3. Справедлива следующая лемма.

Лемма. Для того чтобы в игре $\Gamma(x_0, y_0, T)$ существовала оптимальная программная стратегия игрока E (т. е. стратегия, являющаяся функцией только времени), необходимо и достаточно, чтобы

$$Val\Gamma(x_0, y_0, T) = \hat{\rho}_T(x_0, y_0). \quad (5.5.6)$$

Доказательство. Достаточность. Пусть $v^*(t)$, $t \in [0, T]$ — допустимое управление игрока E , переводящее точку y_0 в точку M такую, что

$$\hat{\rho}_T(x_0, y_0) = \min_{x \in C_P^T(x_0)} \rho(x, M).$$

Обозначим $v^*(\cdot) = \{\sigma, v^*(t)\}$, где разбиение σ отрезка $[0, T]$ состоит из двух точек $t_0 = 0, t_1 = T$. Очевидно, что $v^*(\cdot) \in E$. Согласно теореме п.1.3.4, $v^*(\cdot) \in E$ — оптимальная стратегия игрока E в игре $\Gamma(x_0, y_0, T)$, если

$$Val\Gamma(x_0, y_0, T) = \inf_{u(\cdot) \in P} K(u(\cdot), v^*(\cdot); x_0, y_0, T).$$

Но это равенство следует из (5.5.6), поскольку

$$\inf_{u(\cdot) \in P} K(u(\cdot), v^*(\cdot); x_0, y_0, T) = \hat{\rho}_T(x_0, y_0).$$

Необходимость. Пусть в игре $\Gamma(x_0, y_0, T)$ существует оптимальная программная стратегия игрока E , тогда

$$\begin{aligned} Val\Gamma(x_0, y_0, T) &= \sup_{v(\cdot) \in E} \inf_{u(\cdot) \in P} K(u(\cdot), v(\cdot); x_0, y_0, T) \\ &= \max_{y \in C_E^\delta(y_0)} \inf_{u(\cdot) \in P} \rho(x(T), y) = \hat{\rho}_T(x_0, y_0). \end{aligned}$$

Лемма доказана.

Теорема. Для того чтобы при любых $x_0, y_0 \in R^n$, $T > 0$ в игре $\Gamma(x_0, y_0, T)$ игрок E имел оптимальную программную стратегию, необходимо и достаточно, чтобы для любых $\delta > 0$, $x_0, y_0 \in R^n$, $T \geq \delta$, выполнялось равенство

$$\hat{\rho}_T(x_0, y_0) = \max_{y \in C_E^\delta(y_0)} \min_{x \in C_P^\delta(x_0)} \hat{\rho}_{T-\delta}(x, y). \quad (5.5.7)$$

Доказательство. Достаточность. Согласно теореме п. 5.5.2, из условия (5.5.7) следует соотношение (5.5.2), из которого предельным переходом (см. теорему п. 5.3.7) получим

$$\hat{\rho}_T(x_0, y_0) = Val\Gamma(x_0, y_0, T).$$

Отсюда согласно лемме п. 5.5.3 следует существование оптимальной программной стратегии игрока E .

Необходимость условия (5.5.7) следует из теоремы п. 5.5.1, поскольку существование оптимальной программной стратегии игрока E в игре $\Gamma(x_0, y_0, T)$ влечет существование таковой во всех играх $\Gamma_\delta(x_0, y_0, T)$, $T = \delta k$, $k \geq 1$, и справедливость соотношения (5.5.3).

§ 5.6. Основное уравнение

В данном параграфе будет показано, что при определенных условиях функция значения дифференциальной игры удовлетворяет уравнению в частных производных, которое называется основным. Впервые в монографической литературе оно было рассмотрено Р. Айзексом [Айзекс, 1967] и его часто называют уравнением Айзекса—Беллмана.

5.6.1. Используя теорему п. 5.5.3, выведем уравнение в частных производных для функции значения дифференциальной игры. Будем предполагать, что для игры $\Gamma(x, y, T)$ выполнены условия теоремы п. 5.5.3. Тогда функция $\hat{\rho}_T(x, y)$ представляет собой значение игры $\Gamma(x, y, T)$ продолжительностью T из начальных состояний x, y .

Пусть в некоторой области Ω пространства $R^n \times R^n \times [0, \infty)$ функция $\hat{\rho}_T(x, y)$ имеет непрерывные частные производные по всем переменным. Покажем, что в этом случае функция $\hat{\rho}_T(x, y)$ в области Ω удовлетворяет экстремальному дифференциальному уравнению

$$\frac{\partial \hat{\rho}}{\partial T} - \max_{v \in V} \sum_{i=1}^n \frac{\partial \hat{\rho}}{\partial y_i} g_i(y, v) - \min_{u \in U} \sum_{i=1}^n \frac{\partial \hat{\rho}}{\partial x_i} f_i(x, u) = 0, \quad (5.6.1)$$

где функции $f_i(x, u), g_i(y, v), i = 1, \dots, n$ определяют закон движения игроков в игре Γ (см. (5.3.1), (5.3.2)).

Предположим, что (5.6.1) не выполняется в некоторой точке $(x, y, T) \in \Omega$. Пусть, для определенности,

$$\frac{\partial \hat{\rho}}{\partial T} - \max_{v \in V} \sum_{i=1}^n \frac{\partial \hat{\rho}}{\partial y_i} g_i(y, v) - \min_{u \in U} \sum_{i=1}^n \frac{\partial \hat{\rho}}{\partial x_i} f_i(x, u) < 0$$

и пусть $\bar{v} \in V$ таково, что в рассматриваемой точке $(x, y, T) \in \Omega$ выполнено соотношение

$$\sum_{i=1}^n \frac{\partial \hat{\rho}}{\partial y_i} g_i(y, \bar{v}) = \max_{v \in V} \sum_{i=1}^n \frac{\partial \hat{\rho}}{\partial y_i} g_i(y, v).$$

Тогда при любом $u \in U$ в точке $(x, y, T) \in \Omega$ выполнено неравенство:

$$\frac{\partial \hat{\rho}}{\partial T} - \sum_{i=1}^n \frac{\partial \hat{\rho}}{\partial y_i} g_i(y, \bar{v}) - \sum_{i=1}^n \frac{\partial \hat{\rho}}{\partial x_i} f_i(x, u) < 0. \quad (5.6.2)$$

Из непрерывной дифференцируемости функции $\hat{\rho}$ по всем переменным следует, что неравенство (5.6.2) выполняется и в некоторой окрестности S точки (x, y, T) . Выберем число $\delta > 0$ настолько малым, чтобы точка $(x(\tau), y(\tau), T - \tau) \in S$ при всех $\tau \in [0, \delta]$. Здесь

$$x(\tau) = x + \int_0^\tau f(x(t), u(t)) dt,$$

$$y(\tau) = y + \int_0^\tau g(y(t), \bar{v}(t)) dt$$

— траектории систем (5.3.1), (5.3.2), отвечающие некоторому допустимому управлению $u(t)$ и $\bar{v}(t) \equiv \bar{v}$ соответственно и начальным условиям $x(0) = x, y(0) = y$.

Определим функцию

$$G(\tau) = \frac{\partial \hat{\rho}}{\partial T} \Big|_{(x(\tau), y(\tau), T-\tau)} - \sum_{i=1}^n \frac{\partial \hat{\rho}}{\partial y_i} \Big|_{(x(\tau), y(\tau), T-\tau)} g_i(y(\tau), \bar{v}) - \\ - \sum_{i=1}^n \frac{\partial \hat{\rho}}{\partial x_i} \Big|_{(x(\tau), y(\tau), T-\tau)} f_i(x(\tau), u(\tau)), \quad \tau \in [0, \delta].$$

Функция $G(\tau)$ непрерывна по τ , поэтому найдется число $c < 0$, такое, что $G(\tau) \leq c$ при $\tau \in [0, \delta]$. Отсюда имеем

$$\int_0^\delta G(\tau) d\tau \leq c\delta. \quad (5.6.3)$$

Нетрудно убедиться в том, что

$$G(\tau) = -\frac{d\hat{\rho}}{d\tau} \Big|_{(x(\tau), y(\tau), T-\tau)}.$$

Из (5.6.3) получаем

$$\hat{\rho}_T(x, y) - \hat{\rho}_{T-\delta}(x(\delta), y(\delta)) \leq c\delta.$$

Отсюда, учитывая $c < 0$, в силу произвольности $u(t)$ следует

$$\hat{\rho}_T(x, y) < \max_{y' \in C_E^\delta(y)} \min_{x' \in C_P^\delta(x)} \hat{\rho}_{T-\delta}(x', y'),$$

что противоречит (5.5.7).

Таким образом, мы показали, что в том случае, когда у игрока E в игре $\Gamma(x, y, T)$ при любых $x, y \in R^n$, $T > 0$, существует оптимальная программная стратегия, значение игры $V(x, y, T)$ (оно совпадает с $\hat{\rho}_T(x, y)$ по лемме п. 5.5.3) в области пространства $R^n \times R^n \times [0, \infty)$, где существуют непрерывные частные производные у этой функции, удовлетворяет уравнению

$$\frac{\partial V}{\partial T} = \max_{v \in V} \sum_{i=1}^n \frac{\partial V}{\partial y_i} g_i(y, v) + \min_{u \in U} \sum_{i=1}^n \frac{\partial V}{\partial x_i} f_i(x, u) \quad (5.6.4)$$

при начальном условии $V(x, y, T)|_{T=0} = \rho(x, y)$. Предположим, что каким-то образом удастся определить \bar{u}, \bar{v} , доставляющие \max и \min в (5.6.4) как функции от x, y и $\frac{\partial V}{\partial x}$, т. е.

$$\bar{u} = \bar{u}(x, \frac{\partial V}{\partial x}), \quad \bar{v} = \bar{v}(y, \frac{\partial V}{\partial y}). \quad (5.6.5)$$

Подставляя выражения (5.6.5) в (5.6.4), получаем

$$\sum_{i=1}^n \frac{\partial V}{\partial y_i} g_i\left(y, \bar{v}(y, \frac{\partial V}{\partial y})\right) + \sum_{i=1}^n \frac{\partial V}{\partial x_i} f_i\left(x, \bar{u}(x, \frac{\partial V}{\partial x})\right) = \frac{\partial V}{\partial T} \quad (5.6.6)$$

при условии

$$V(x, y, T)|_{T=0} = \rho(x, y). \quad (5.6.7)$$

Таким образом, для определения $V(x, y, T)$ имеем задачу Коши для уравнения в частных производных первого порядка (5.6.6) при начальном условии (5.6.7).

Замечание. При выводе функциональных уравнений (5.6.4), (5.6.6) и доказательстве теоремы п. 5.5.3 мы не использовали конкретный вид функции выигрыша, поэтому теорема остается справедливой для любого непрерывного терминального выигрыша $H(x(T), y(T))$. Однако в этом случае вместо величины $\hat{\rho}_T(x, y)$ необходимо рассмотреть величину

$$\hat{H}_T(x, y) = \max_{y' \in C_E^T(y)} \min_{x' \in C_P^T(x)} H(x', y').$$

Уравнение (5.6.4) также справедливо для значения дифференциальной игры с предписанной продолжительностью и любым терминальным выигрышем, т. е. если в дифференциальной игре с предписанной продолжительностью $\Gamma(x, y, T)$ и терминальным выигрышем $H(x(T), y(T))$ у игрока E существует оптимальная программная стратегия, то значение игры $V(x, y, T)$ в области пространства $R^n \times R^n \times [0, \infty)$, где существуют непрерывные частные производные, удовлетворяет уравнению (5.6.4) при начальном условии $V(x, y, T)|_{T=0} = H(x, y)$ или уравнению (5.6.6) с тем же начальным условием.

5.6.2. Рассмотрим теперь игры преследования, в которых функция выигрыша равна времени до момента встречи. Предположим, для определенности, что терминальное многообразие F является сферой $\rho(x, y) = l$, $l > 0$. Будем предполагать, что множества $C_P^t(x)$ и $C_E^t(y)$ — непрерывны по t в нуле равномерно относительно x и y .

Пусть имеет смысл величина

$$\theta(x, y, l) = \max_{v(t)} \min_{u(t)} t_n^l(x, y; u(t), v(t)),$$

где $t_n^l(x, y; u(t), v(t))$ — время сближения на l -расстояние игроков P и E , движущихся из начальных точек x, y при использовании измеримых программных управлений $u(t)$ и $v(t)$ соответственно. Предположим также, что функция $\theta(x, y, l)$ непрерывна по совокупности аргументов.

Игру на быстроедействие будем обозначать как $\Gamma(x_0, y_0)$. Так же, как это было сделано в §§ 5.4, 5.5, можно вывести необходимые и достаточные условия существования оптимальной программной стратегии игрока E в игре преследования на быстроедействие. Справедлива следующая теорема.

Теорема. Для того чтобы игрок E при любых $x_0, y_0 \in R^n$ в игре $\Gamma(x_0, y_0)$ имел оптимальную программную стратегию, необходимо и достаточно, чтобы при любом $\delta > 0$ и любых $x_0, y_0 \in R^n$

$$\theta(x_0, y_0, l) = \delta + \max_{y' \in C_E^\delta(y_0)} \min_{x' \in C_P^\delta(x_0)} \theta(x', y', l).$$

Для игры преследования по быстроедействию уравнение (5.6.4) принимает вид

$$\max_{v \in V} \sum_{i=1}^n \frac{\partial \theta}{\partial y_i} g_i(y, v) + \min_{u \in U} \sum_{i=1}^n \frac{\partial \theta}{\partial x_i} f_i(x, u) = -1 \quad (5.6.8)$$

при начальном условии

$$\theta(x, y, l)|_{\rho(x, y)=l} = 0. \quad (5.6.9)$$

Здесь предполагается существование непрерывных частных производных первого порядка функции $\theta(x, y, l)$ по x, y .

Полагая, что каким-то образом можно определить \bar{u}, \bar{v} , доставляющие \max и \min (5.6.8) как функции от $x, y, \partial\theta/\partial x, \partial\theta/\partial y$, т. е. $\bar{u} = \bar{u}(x, \frac{\partial\theta}{\partial x})$, $\bar{v} = \bar{v}(y, \frac{\partial\theta}{\partial y})$, перепишем уравнение (5.6.8) в виде

$$\sum_{i=1}^n \frac{\partial\theta}{\partial y_i} g_i \left(y, \bar{v}(y, \frac{\partial\theta}{\partial y}) \right) + \sum_{i=1}^n \frac{\partial\theta}{\partial x_i} f_i \left(x, \bar{u}(x, \frac{\partial\theta}{\partial x}) \right) = -1 \quad (5.6.10)$$

при условии

$$\theta(x, y, l)|_{\rho(x, y)=l} = 0. \quad (5.6.11)$$

Вывод уравнения (5.6.8) аналогичен выводу уравнения (5.6.4) для игры преследования с предписанной продолжительностью.

Обе задачи Коши (5.6.4), (5.6.7) и (5.6.8), (5.6.9) являются нелинейными относительно частных производных, поэтому при их решении возникают значительные трудности.

5.6.3. Перейдем теперь к выводу уравнений характеристик для уравнения (5.6.4). Предположим, что функция $V(x, y; T)$ имеет непрерывные вторые смешанные производные на всем пространстве, функции $g_i(y, v)$, $f_i(x, u)$ и функции $\bar{u} = \bar{u}(x, \frac{\partial V}{\partial x})$, $\bar{v} = \bar{v}(y, \frac{\partial V}{\partial y})$ имеют непрерывные первые производные по всем переменным, а множества U, V представляют собой параллелепипеды $a_m \leq u_m \leq b_m$, $m = 1, \dots, k$ и $c_q \leq v_q \leq d_q$, $q = 1, \dots, l$, где $u = (u_1, \dots, u_k) \in U$, $v = (v_1, \dots, v_l) \in V$. Введем обозначение:

$$B(x, y, T) = \frac{\partial V}{\partial T} - \sum_{i=1}^n \frac{\partial V}{\partial x_i} f_i(x, \bar{u}) - \sum_{i=1}^n \frac{\partial V}{\partial y_i} g_i(y, \bar{v}).$$

Функция $B(x, y, T) \equiv 0$. Поэтому беря частные производные по x_1, \dots, x_n , получим

$$\begin{aligned} \frac{\partial B}{\partial x_k} &= \frac{\partial^2 V}{\partial T \partial x_k} - \sum_{i=1}^n \frac{\partial^2 V}{\partial x_i \partial x_k} f_i - \sum_{i=1}^n \frac{\partial V}{\partial x_i} \frac{\partial f_i}{\partial x_k} - \sum_{i=1}^n \frac{\partial^2 V}{\partial y_i \partial x_k} g_i - \sum_{m=1}^k \frac{\partial}{\partial u_m} \left(\sum_{i=1}^n \frac{\partial V}{\partial x_i} f_i \right) \frac{\partial \bar{u}_m}{\partial x_k} \\ &\quad - \sum_{i=1}^l \frac{\partial}{\partial v_q} \left(\sum_{i=1}^n \frac{\partial V}{\partial y_i} g_i \right) \frac{\partial \bar{v}_q}{\partial x_k} = 0, \quad k = 1, \dots, n. \end{aligned} \quad (5.6.12)$$

Для каждой фиксированной точки $(x, y, T) \in R^n \times R^n \times [0, \infty)$ максимизирующее значение \bar{v} и минимизирующее значение \bar{u} в (5.6.4) лежат либо внутри, либо на границе области ограничений. Если это внутренняя точка, то

$$\frac{\partial}{\partial u_m} \left(\sum_{i=1}^n \frac{\partial V}{\partial x_i} f_i \right) \Big|_{u=\bar{u}} = 0, \quad \frac{\partial}{\partial v_q} \left(\sum_{i=1}^n \frac{\partial V}{\partial y_i} g_i \right) \Big|_{v=\bar{v}} = 0.$$

Если же $\bar{u}(\bar{v})$ лежит на границе, то здесь могут представиться два случая. Исследуем их подробно для одной из компонент $\bar{u}_m(x, \frac{\partial V}{\partial x})$ вектора \bar{u} . Исследование остальных компонент вектора \bar{u} и компонент вектора \bar{v} проводится аналогично. Для простоты предположим, что в некоторой точке (x', y', T')

$$\bar{u}_m = \bar{u}_m \left(x', \frac{\partial V(x', y', T')}{\partial x} \right) = a_m.$$

Случай 1. Существует шар в пространстве R^n с центром в точке x' , для всех точек x которого выполняется равенство:

$$\bar{u}_m = \bar{u}_m \left(x, \frac{\partial V(x, y', T')}{\partial x} \right) = a_m.$$

Функция \bar{u}_m на этом шаре принимает постоянное значение, поэтому в точке x' имеем

$$\frac{\partial \bar{u}_m}{\partial x_i} = 0, \quad i = 1, \dots, n.$$

Случай 2. Такого шара не существует. Тогда найдется последовательность x_r , $\lim_{r \rightarrow \infty} x_r = x'$ такая, что

$$\bar{u}_m \left(x_r, \frac{\partial V(x_r, y', T')}{\partial x} \right) \neq a_m.$$

Следовательно,

$$\frac{\partial}{\partial u_m} \left(\sum_{i=1}^n \frac{\partial V}{\partial x_i} \Big|_{(x_r, y', T')} f_i(x_r, \bar{u}) \right) = 0.$$

Из непрерывности производных $\partial V / \partial x_i$, $\partial f_i / \partial u_m$ и функции $\bar{u} = \bar{u}(x, \frac{\partial V(x, y, T)}{\partial x})$ следует, что предыдущее равенство выполняется и в точке (x', y', T') .

Таким образом, два последних слагаемых в (5.6.12) равны нулю, и при всех $(x, y, T) \in R^n \times [0, \infty)$ выполняется равенство:

$$\begin{aligned} \frac{\partial B}{\partial x_k} &= \frac{\partial^2 V}{\partial T \partial x_k} - \sum_{i=1}^n \frac{\partial^2 V}{\partial x_i \partial x_k} f_i(x, \bar{u}) - \\ &- \sum_{i=1}^n \frac{\partial V}{\partial x_i} \frac{\partial f_i}{\partial x_k} - \sum_{i=1}^n \frac{\partial^2 V}{\partial y_i \partial x_k} g_i(y, \bar{v}) = 0, \quad k = 1, 2, \dots, n. \end{aligned}$$

Пусть $\bar{x}(t), \bar{y}(t)$, $t \in [0, T]$ — решение системы

$$\begin{aligned} \dot{x} &= f \left(x, \bar{u} \left(x, \frac{\partial V(x, y, T-t)}{\partial x} \right) \right), \\ \dot{y} &= g \left(y, \bar{v} \left(y, \frac{\partial V(x, y, T-t)}{\partial y} \right) \right) \end{aligned}$$

с начальными условиями $x(0) = x_0$, $y(0) = y_0$. Вдоль решения $\bar{x}(t), \bar{y}(t)$ имеем

$$\begin{aligned} &\frac{\partial^2 V(\bar{x}(t), \bar{y}(t), T-t)}{\partial T \partial x_k} - \sum_{i=1}^n \frac{\partial^2 V(\bar{x}(t), \bar{y}(t), T-t)}{\partial x_i \partial x_k} f_i(\bar{x}(t), \bar{u}(t)) - \\ &- \sum_{i=1}^n \frac{\partial V(\bar{x}(t), \bar{y}(t), T-t)}{\partial x_i} \frac{\partial f_i(\bar{x}(t), \bar{u}(t))}{\partial x_k} - \\ &- \sum_{i=1}^n \frac{\partial^2 V(\bar{x}(t), \bar{y}(t), T-t)}{\partial y_i \partial x_k} g_i(\bar{y}(t), \bar{v}(t)) = 0, \quad k = 1, \dots, n, \end{aligned} \quad (5.6.13)$$

где

$$\begin{aligned} \bar{u}(t) &= \bar{u} \left(\bar{x}(t), \frac{\partial V(\bar{x}(t), \bar{y}(t), T-t)}{\partial x} \right), \\ \bar{v}(t) &= \bar{v} \left(\bar{y}(t), \frac{\partial V(\bar{x}(t), \bar{y}(t), T-t)}{\partial y} \right). \end{aligned}$$

Однако,

$$\begin{aligned} \frac{d}{dt} \left(\frac{\partial V(\bar{x}(t), \bar{y}(t), T-t)}{\partial x_k} \right) &= \sum_{i=1}^n \frac{\partial^2 V(\bar{x}(t), \bar{y}(t), T-t)}{\partial x_k \partial x_i} f_i(\bar{x}(t), \tilde{u}(t)) + \\ &+ \sum_{i=1}^n \frac{\partial^2 V(\bar{x}(t), \bar{y}(t), T-t)}{\partial x_k \partial y_i} g_i(\bar{y}(t), \tilde{v}(t)) - \\ &- \frac{\partial^2 V(\bar{x}(t), \bar{y}(t), T-t)}{\partial x_k \partial T}, \quad k = 1, \dots, n. \end{aligned} \quad (5.6.14)$$

Заметим, что у дважды непрерывно дифференцируемой функции можно менять порядок дифференцирования. Перепишем (5.6.13) с учетом (5.6.14) в виде

$$\frac{d}{dt} \left(\frac{\partial V(\bar{x}(t), \bar{y}(t), T-t)}{\partial x_k} \right) = - \sum_{i=1}^n \frac{\partial V(\bar{x}(t), \bar{y}(t), T-t)}{\partial x_i} \frac{\partial f_i(\bar{x}(t), \tilde{u}(t))}{\partial x_k}, \quad k = 1, \dots, n.$$

Аналогичным образом получим уравнения

$$\frac{d}{dt} \left(\frac{\partial V(\bar{x}(t), \bar{y}(t), T-t)}{\partial y_i} \right) = - \sum_{j=1}^n \frac{\partial V(\bar{x}(t), \bar{y}(t), T-t)}{\partial y_j} \frac{\partial g_j(\bar{y}(t), \tilde{v}(t))}{\partial y_i}, \quad i = 1, \dots, n.$$

Так как при $t \in [0, T]$

$$V(\bar{x}(t), \bar{y}(t), T-t) = H(\bar{x}(T), \bar{y}(T)),$$

то

$$\frac{d}{dt} \left(\frac{\partial V(\bar{x}(t), \bar{y}(t), T-t)}{\partial T} \right) = 0.$$

Введем следующие обозначения:

$$\begin{aligned} V_{x_i}(t) &= \frac{\partial V(\bar{x}(t), \bar{y}(t), T-t)}{\partial x_i}, \\ V_{y_i}(t) &= \frac{\partial V(\bar{x}(t), \bar{y}(t), T-t)}{\partial y_i}, \quad i = 1, \dots, n; \\ V_x(t) &= \{V_{x_i}(t)\}, \quad V_y(t) = \{V_{y_i}(t)\}, \\ V_T(t) &= \frac{\partial V(\bar{x}(t), \bar{y}(t), T-t)}{\partial T}. \end{aligned}$$

В результате получим следующую систему обыкновенных дифференциальных уравнений для функций $x(t), y(t), V_x(t), V_y(t)$:

$$\begin{aligned} \dot{x}_i &= f_i(x, \bar{u}(x, V_x)), \\ \dot{y}_i &= g_i(y, \bar{v}(y, V_y)), \\ \dot{V}_{x_k} &= - \sum_{i=1}^n V_{x_i} \frac{\partial f_i(x, \bar{u}(x, V_x))}{\partial x_k}, \\ \dot{V}_{y_k} &= - \sum_{i=1}^n V_{y_i} \frac{\partial g_i(y, \bar{v}(y, V_y))}{\partial y_k}, \\ V_T &= 0, \quad i, k = 1, \dots, n, \end{aligned} \quad (5.6.15)$$

и, кроме того, согласно (5.6.6) имеем

$$V_T = \sum_{i=1}^n V_{y_i} g_i(y, \bar{v}(y, V_y)) + \sum_{i=1}^n V_{x_i} f_i(x, \bar{u}(x, V_x)).$$

Для решения системы нелинейных уравнений (5.6.15) относительно функций $x(t), y(t), V_{x_k}(t), V_{y_k}(t), V_T(t)$ необходимо определить начальные условия. Для функции $V(\bar{x}(t), \bar{y}(t), T - t)$ они заданы в момент времени $t = T$. Поэтому введем переменную $\tau = T - t$ и запишем уравнение характеристик в регрессивной форме. Введем обозначения $\overset{\circ}{x} = -\dot{x}, \overset{\circ}{y} = -\dot{y}$. Уравнения характеристик принимают следующий вид:

$$\begin{aligned} \overset{\circ}{x}_i &= -f_i(x, \bar{u}), \\ \overset{\circ}{y}_i &= -g_i(y, \bar{v}), \\ \overset{\circ}{V}_{x_k} &= \sum_{i=1}^n V_{x_i} \frac{\partial f_i(x, \bar{u})}{\partial x_k}, \\ \overset{\circ}{V}_{y_k} &= \sum_{i=1}^n V_{y_i} \frac{\partial g_i(y, \bar{v})}{\partial y_k}, \\ \overset{\circ}{V}_T &= 0. \end{aligned} \tag{5.6.16}$$

При задании начальных условий для системы (5.6.16) используется соотношение $V(x, y, T)|_{T=0} = H(x, y)$. Пусть $x|_{\tau=0} = s, y|_{\tau=0} = s'$. Тогда

$$\begin{aligned} V_{x_i}|_{\tau=0} &= \frac{\partial H}{\partial x_i} \Big|_{x=s, y=s'}, \\ V_{y_i}|_{\tau=0} &= \frac{\partial H}{\partial y_i} \Big|_{x=s, y=s'}, \end{aligned} \tag{5.6.17}$$

$$V_T|_{\tau=0} = \sum_{i=1}^n V_{y_i}|_{\tau=0} g_i(s', \bar{v}(s', V_y|_{\tau=0})) + \sum_{i=1}^n V_{x_i}|_{\tau=0} f_i(s, \bar{u}(s, V_x|_{\tau=0})).$$

Подробные исследования возможных путей решения системы (5.6.16)–(5.6.17) см. в [Айзекс, 1967].

Аналогичным образом, используя уравнение (5.6.8), можно записать уравнение характеристик для задачи преследования на быстродействие.

§ 5.7. Методы последовательных приближений для решения дифференциальных игр преследования

5.7.1. Пусть $\Gamma_\delta(x, y, T)$ — дискретная форма дифференциальной игры $\Gamma(x, y, T)$ продолжительностью $T > 0$ с фиксированным шагом разбиения δ и дискриминацией игрока E на время $\delta > 0$ вперед. Обозначим через $V_\delta(x, y, T)$ значение игры $\Gamma_\delta(x, y, T)$.¹ Тогда

$$\lim_{\delta \rightarrow 0} V_\delta(x, y, T) = V(x, y, T)$$

¹ Терминальный выигрыш равен $\rho(x(T), y(T))$, где $\rho(x, y)$ — расстояние в R^n .

и оптимальные стратегии в игре $\Gamma_\delta(x, y, T)$ при достаточно малых δ могут быть эффективно использованы для построения ситуаций ε -равновесия в игре $\Gamma(x, y, T)$.

5.7.2. Идея численного метода состоит в построении алгоритма нахождения решения игры $\Gamma_\delta(x, y, T)$. Перейдем непосредственно к изложению метода.

Нулевое приближение. За нулевое приближение функции значения игры $V_\delta(x, y, T)$ принимаем функцию

$$V_\delta^0(x, y, T) = \max_{\eta \in C_E^T(y)} \min_{\xi \in C_P^T(x)} \rho(\xi, \eta), \quad (5.7.1)$$

где $C_P^T(x), C_E^T(y)$ — множества достижимости игроков P и E из начальных состояний $x, y \in R^n$ к моменту времени T .

Выбор функции $V_\delta^0(x, y, T)$ в качестве начального приближения оправдан тем, что в достаточно широком классе игр (так называемый регулярный случай) она оказывается значением игры $\Gamma(x, y, T)$. Следующие приближения строятся по правилу:

$$\begin{aligned} V_\delta^1(x, y, T) &= \max_{\eta \in C_E^\delta(y)} \min_{\xi \in C_P^\delta(x)} V_\delta^0(\xi, \eta, T - \delta), \\ V_\delta^2(x, y, T) &= \max_{\eta \in C_E^\delta(y)} \min_{\xi \in C_P^\delta(x)} V_\delta^1(\xi, \eta, T - \delta), \\ &\dots \\ V_\delta^k(x, y, T) &= \max_{\eta \in C_E^\delta(y)} \min_{\xi \in C_P^\delta(x)} V_\delta^{k-1}(\xi, \eta, T - \delta) \end{aligned} \quad (5.7.2)$$

при $T > \delta$ и $V_\delta^k(x, y, T) = V_\delta^0(x, y, T)$ при $T \leq \delta, k \geq 1$.

Как видно из формул (5.7.2), операция $\max \min$ берется по множествам достижимости $C_E^\delta(y), C_P^\delta(x)$ за время δ , т. е. за один шаг дискретной игры $\Gamma_\delta(x, y, T)$.

5.7.3. Теорема. При фиксированных x, y, T, δ числовая последовательность $\{V_\delta^k(x, y, T)\}$ не убывает с ростом k .

Доказательство. Докажем сначала неравенство

$$V_\delta^1(x, y, T) \geq V_\delta^0(x, y, T).$$

Для всех $\xi \in C_P^\delta(x)$ справедливо: $C_P^{T-\delta}(\xi) \subset C_P^T(x)$. Для любых $\bar{\eta} \in C_E^{T-\delta}(\eta), \xi \in C_P^\delta(x)$ имеем

$$\min_{\bar{\xi} \in C_P^{T-\delta}(\xi)} \rho(\bar{\xi}, \bar{\eta}) \geq \min_{\bar{\xi} \in C_P^T(x)} \rho(\bar{\xi}, \bar{\eta}).$$

Следовательно,

$$\begin{aligned} V_\delta^1(x, y, T) &= \max_{\eta \in C_E^\delta(y)} \min_{\xi \in C_P^\delta(x)} \max_{\bar{\eta} \in C_E^{T-\delta}(\eta)} \min_{\bar{\xi} \in C_P^{T-\delta}(\xi)} \rho(\bar{\xi}, \bar{\eta}) \\ &\geq \max_{\eta \in C_E^\delta(y)} \max_{\bar{\eta} \in C_E^{T-\delta}(\eta)} \min_{\bar{\xi} \in C_P^T(x)} \rho(\bar{\xi}, \bar{\eta}) = \max_{\eta \in C_E^T(y)} \min_{\xi \in C_P^T(x)} \rho(\xi, \eta) = V_\delta^0(x, y, T). \end{aligned}$$

Предположим теперь, что для $l \leq k$ справедливо неравенство

$$V_\delta^l(x, y, T) \geq V_\delta^{l-1}(x, y, T), \quad (5.7.3)$$

и докажем его для $l = k + 1$.

Из соотношений (5.7.2) и (5.7.3) следует, что

$$\begin{aligned} V_{\delta}^{k+1}(x, y, T) &= \max_{\eta \in C_E^{\delta}(y)} \min_{\xi \in C_P^{\delta}(x)} V_{\delta}^k(\xi, \eta, T - \delta) \geq \\ &\geq \max_{\eta \in C_E^{\delta}(y)} \min_{\xi \in C_P^{\delta}(x)} V_{\delta}^{k-1}(\xi, \eta, T - \delta) = V_{\delta}^k(x, y, T). \end{aligned}$$

Таким образом, в случае $T > \delta$, по индукции утверждение теоремы доказано. В случае $T \leq \delta$ утверждение теоремы очевидно.

5.7.4. Теорема. *Последовательность $\{V_{\delta}^k(x, y, T)\}$ сходится за конечное число шагов N , при этом справедлива оценка $N \leq \lceil \frac{T}{\delta} \rceil + 1$, где квадратные скобки означают целую часть.*

Доказательство. Пусть $N = \lceil T/\delta \rceil + 1$. Покажем, что

$$V_{\delta}^N(x, y, T) = V_{\delta}^{N+1}(x, y, T). \quad (5.7.4)$$

Равенство (5.7.4) легко получить из построения последовательности $\{V_{\delta}^k(x, y, T)\}$. Действительно,

$$\begin{aligned} V_{\delta}^N(x, y, T) &= \max_{\eta^1 \in C_E^{\delta}(y)} \min_{\xi^1 \in C_P^{\delta}(x)} V_{\delta}^{N+1}(\xi^1, \eta^1, T - \delta) = \\ &= \max_{\eta^1 \in C_E^{\delta}(y)} \min_{\xi^1 \in C_P^{\delta}(x)} \max_{\eta^2 \in C_E^{\delta}(\eta^1)} \dots \max_{\eta^{N-1} \in C_E^{\delta}(\eta^{N-2})} \min_{\xi^{N-1} \in C_P^{\delta}(\xi^{N-2})} V_{\delta}^1(\xi^{N-1}, \eta^{N-1}, T - (N-1)\delta). \end{aligned}$$

Аналогично имеем

$$\begin{aligned} V_{\delta}^{N+1}(x, y, T) &= \\ &= \max_{\eta^1 \in C_E^{\delta}(y)} \min_{\xi^1 \in C_P^{\delta}(x)} \max_{\eta^2 \in C_E^{\delta}(\eta^1)} \dots \max_{\eta^{N-1} \in C_E^{\delta}(\eta^{N-2})} \min_{\xi^{N-1} \in C_P^{\delta}(\xi^{N-2})} V_{\delta}^2(\xi^{N-1}, \eta^{N-1}, T - (N-1)\delta). \end{aligned}$$

Однако $T - (N-1)\delta = \alpha < \delta$, поэтому

$$V_{\delta}^1(\xi^{N-1}, \eta^{N-1}, \alpha) = V_{\delta}^2(\xi^{N-1}, \eta^{N-1}, \alpha) = V_{\delta}^0(\xi^{N-1}, \eta^{N-1}, \alpha),$$

откуда и следует равенство (5.7.4).

Совпадение членов последовательности V_{δ}^k при $k \geq N$ выводится из (5.7.4) индукцией. Теорема доказана.

5.7.5. Теорема. *Предел последовательности $\{V_{\delta}^k(x, y, T)\}$ совпадает со значением игры $\Gamma_{\delta}(x, y, T)$.*

Доказательство. Данная теорема является, по существу, следствием теоремы п. 5.7.4. Действительно, обозначим

$$V_{\delta}(x, y, T) = \lim_{k \rightarrow \infty} V_{\delta}^k(x, y, T).$$

Сходимость происходит за конечное число шагов, не превосходящее $N = \lceil T/\delta \rceil + 1$, поэтому в рекуррентном уравнении (5.7.2) можно перейти к пределу при $k \rightarrow \infty$. Предельная функция $V_{\delta}(x, y, T)$ удовлетворяет уравнению

$$V_{\delta}(x, y, T) = \max_{\eta \in C_E^T(y)} \min_{\xi \in C_P^T(x)} V_{\delta}(\xi, \eta, T - \delta) \quad (5.7.5)$$

при начальном условии

$$V_{\delta}(x, y, T)|_{0 \leq T \leq \delta} = \max_{\eta \in C_E^T(y)} \min_{\xi \in C_P^T(x)} \rho(\xi, \eta), \quad (5.7.6)$$

что и является достаточным условием для того, чтобы функция $V_\delta(x, y, T)$ была значением игры $\Gamma_\delta(x, y, T)$.

5.7.6. Приведем модификацию метода последовательных приближений, изложенного выше.

В качестве начального приближения возьмем функцию $\tilde{V}_\delta^0(x, y, T) = V_\delta^0(x, y, T)$, где $V_\delta^0(x, y, T)$ определена равенством (5.7.1). Следующие приближения строим по правилу:

$$\tilde{V}_\delta^{k+1}(x, y, T) = \max_{i \in [1:N]} \max_{\eta \in C_E^{i\delta}(y)} \min_{\xi \in C_P^{i\delta}(x)} \tilde{V}_\delta^k(\xi, \eta, T - i\delta)$$

при $T > \delta$, где $N = [T/\delta]$, и $\tilde{V}_\delta^{k+1}(x, y, T) = \tilde{V}_\delta^0(x, y, T)$ при $T \leq \delta$.

Для последовательности функций $\{\tilde{V}_\delta^k(x, y, T)\}$ так же, как и для последовательности функций $\{V_\delta^k(x, y, T)\}$, справедливы утверждения теорем п. 5.7.3–5.7.5.

Доказательство этих утверждений для последовательности функций $\{\tilde{V}_\delta^k(x, y, T)\}$ почти дословно повторяет аналогичные рассуждения для последовательности функций $\{V_\delta^k(x, y, T)\}$. В области $\{(x, y, T) | T > \delta\}$ функциональное уравнение для функции значения игры $\Gamma_\delta(x, y, T)$ принимает вид

$$V_\delta(x, y, T) = \max_{i \in [1:N]} \max_{\eta \in C_E^{i\delta}(y)} \min_{\xi \in C_P^{i\delta}(x)} V_\delta(\xi, \eta, T - i\delta), \quad (5.7.7)$$

где $N = [T/\delta]$, а начальное условие остается прежним, т. е. имеет вид (5.7.6).

5.7.7. Докажем эквивалентность уравнений (5.7.5) и (5.7.7).

Теорема. Уравнения (5.7.5) и (5.7.7) с начальным условием (5.7.6) являются эквивалентными.

Доказательство. Пусть функция $V_\delta(x, y, T)$ удовлетворяет уравнению (5.7.5) и начальному условию (5.7.6). Покажем, что она удовлетворяет уравнению (5.7.7) в области $\{(x, y, T) | T > \delta\}$.

Действительно, справедливы следующие соотношения:

$$\begin{aligned} V_\delta(x, y, T) &= \max_{\eta \in C_E^\delta(y)} \min_{\xi \in C_P^\delta(x)} V_\delta(\xi, \eta, T - \delta) = \max_{\eta \in C_E^\delta(y)} \min_{\xi \in C_P^\delta(x)} \max_{\bar{\eta} \in C_E^\delta(\eta)} \min_{\bar{\xi} \in C_P^\delta(\xi)} V_\delta(\bar{\xi}, \bar{\eta}, T - 2\delta) \geq \\ &\geq \max_{\eta \in C_E^\delta(y)} \max_{\bar{\eta} \in C_E^\delta(\eta)} \min_{\xi \in C_P^\delta(x)} \min_{\bar{\xi} \in C_P^\delta(\xi)} V_\delta(\bar{\xi}, \bar{\eta}, T - 2\delta) = \\ &= \max_{\eta \in C_E^{2\delta}(y)} \min_{\xi \in C_P^{2\delta}(x)} V_\delta(\xi, \eta, T - 2\delta) \geq \dots \geq \max_{\eta \in C_E^{i\delta}(y)} \min_{\xi \in C_P^{i\delta}(x)} V_\delta(\xi, \eta, T - i\delta) \geq \dots \end{aligned}$$

При $i = 1$, имеем

$$V_\delta(x, y, T) = \max_{\eta \in C_E^\delta(y)} \min_{\xi \in C_P^\delta(x)} V_\delta(\xi, \eta, T - \delta),$$

поэтому справедливо равенство

$$V_\delta(x, y, T) = \max_{i \in [1:N]} \max_{\eta \in C_E^{i\delta}(y)} \min_{\xi \in C_P^{i\delta}(x)} V_\delta(\xi, \eta, T - i\delta),$$

где $N = [T/\delta]$, что и доказывает требуемое утверждение.

Пусть теперь функция $V_\delta(x, y, T)$ в области $\{(x, y, T) | T > \delta\}$ удовлетворяет уравнению (5.7.7) и начальному условию (5.7.6). Покажем, что она удовлетворяет также уравнению (5.7.5). Предположим противное. Тогда в области $\{(x, y, T) | T > \delta\}$ должно выполняться неравенство:

$$V_\delta(x, y, T) > \max_{\eta \in C_E^\delta(y)} \min_{\xi \in C_P^\delta(x)} V_\delta(\xi, \eta, T - \delta).$$

Однако,

$$\begin{aligned}
 \max_{\eta \in C_E^\delta(y)} \min_{\xi \in C_P^\delta(x)} V_\delta(\xi, \eta, T - \delta) &= \max_{\eta \in C_E^\delta(y)} \min_{\xi \in C_P^\delta(x)} \max_{i \in [1:N-1]} \max_{\bar{\eta} \in C_E^{i\delta}(\eta)} \min_{\bar{\xi} \in C_P^{i\delta}(\xi)} V_\delta(\bar{\xi}, \bar{\eta}, T - (i+1)\delta) \geq \\
 &\geq \max_{\eta \in C_E^\delta(y)} \max_{i \in [1:N-1]} \max_{\bar{\eta} \in C_E^{i\delta}(\eta)} \min_{\xi \in C_P^\delta(x)} \min_{\bar{\xi} \in C_P^{i\delta}(\xi)} V_\delta(\bar{\xi}, \bar{\eta}, T - (i+1)\delta) = \\
 &= \max_{i \in [1:N-1]} \max_{\eta \in C_E^\delta(y)} \max_{\bar{\eta} \in C_E^{i\delta}(\eta)} \min_{\xi \in C_P^\delta(x)} \min_{\bar{\xi} \in C_P^{i\delta}(\xi)} V_\delta(\bar{\xi}, \bar{\eta}, T - (i+1)\delta) = \\
 &= \max_{i \in [2:N]} \max_{\eta \in C_E^\delta(y)} \min_{\xi \in C_P^\delta(x)} V_\delta(\xi, \eta, T - i\delta) = V_\delta(x, y, T).
 \end{aligned}$$

Поскольку при $i = 1$ выполнено строгое неравенство, полученное противоречие доказывает теорему.

§ 5.8. Примеры решения дифференциальных игр преследования

5.8.1. Пример 4 (Простое движение.) Рассмотрим дифференциальную игру $\Gamma(x_0, y_0, T)$, в которой движение игроков P и E происходит в Евклидовом пространстве R^n согласно следующим уравнениям:

$$\begin{aligned}
 \text{для } P: \dot{x} &= \alpha u, \quad \|u\| < 1, \quad x(0) = x_0, \\
 \text{для } E: \dot{y} &= \beta v, \quad \|v\| < 1, \quad y(0) = y_0,
 \end{aligned} \tag{5.8.1}$$

где α, β — константы, $\alpha > \beta > 0$, $x, y, u, v \in R^n$. Для Выигрыш игрока E равен

$$H(x(T), y(T)) = \|x(T) - y(T)\|.$$

Пусть $\Gamma_\delta(x, y, T)$ — дискретная форма дифференциальной игры $\Gamma(x, y, T)$ с шагом разбиения $\delta > 0$ и дискриминацией игрока E . Игра $\Gamma_\delta(x, y, T)$ имеет N шагов, где $N = T/\delta$. Согласно результатам §5.2 (см. пример п. 5.2.3) игра $\Gamma_\delta(x, y, T)$ имеет значение

$$V_\delta(x, y, T) = \max\{0, \|x - y\| - N \cdot \delta \cdot (\alpha - \beta)\} = \max\{0, \|x - y\| - T(\alpha - \beta)\},$$

а оптимальное движение игроков происходит по прямой, соединяющей начальные состояния x, y .

Согласно результатам §5.3 значение исходной дифференциальной игры

$$V(x, y, T) = \lim_{\delta \rightarrow 0} V_\delta(x, y, T) = \max\{0, \|x - y\| - T(\alpha - \beta)\}. \tag{5.8.2}$$

Можно убедиться, что

$$V(x, y, T) = \max_{y' \in C_E^T(y)} \min_{x' \in C_P^T(x)} \|x' - y'\| = \hat{\rho}_T(x, y),$$

где $C_E^T(y) = S(y, \beta T)$ — шар в R^n радиуса βT с центром в точке y , аналогично $C_P^T(x) = S(x, \alpha T)$. Тем самым согласно лемме п. 5.5.3 у игрока E в игре $\Gamma(x_0, y_0, T)$ существует оптимальная программная стратегия $v^*(t)$, $t \in [0, T]$, которая приводит траекторию игрока в точку $y^* \in C_E^T(y_0)$, для которой

$$\hat{\rho}_T(x_0, y_0) = \min_{x' \in C_P^T(x_0)} \|x' - y^*\|.$$

Очевидно,

$$v^*(t) \equiv v^* = \begin{cases} \frac{y_0 - x_0}{\|y_0 - x_0\|}, & \text{при } y_0 \neq x_0, \\ v, & \text{при } y_0 = x_0, \end{cases}$$

где $v \in R^n$ — произвольный вектор такой, что $\|v\| = 1$. Из результатов 5.6 следует, что в области Δ

$$\Delta = \{(x, y, T) : \|x - y\| - T(\alpha - \beta) > 0\},$$

где существуют непрерывные частные производные

$$\frac{\partial V}{\partial T} = -(\alpha - \beta), \quad \frac{\partial V}{\partial x} = -\frac{\partial V}{\partial y} = \frac{x - y}{\|x - y\|},$$

функция $V(x, y, T)$ удовлетворяет уравнению (5.6.4):

$$\frac{\partial V}{\partial T} - \alpha \min_{\|u\| \leq 1} \left(\frac{\partial V}{\partial x}, u \right) - \beta \max_{\|v\| \leq 1} \left(\frac{\partial V}{\partial y}, v \right) = 0. \quad (5.8.3)$$

В уравнении (5.8.3) минимум и максимум достигаются при управлениях

$$\bar{u}(x, \frac{\partial V}{\partial x}) = -\frac{\frac{\partial V}{\partial x}}{\|\frac{\partial V}{\partial x}\|} = \frac{y - x}{\|y - x\|}; \quad (5.8.4)$$

$$\bar{v}(y, \frac{\partial V}{\partial y}) = \frac{\frac{\partial V}{\partial y}}{\|\frac{\partial V}{\partial y}\|} = \frac{y - x}{\|y - x\|}. \quad (5.8.5)$$

Стратегии (5.8.4), (5.8.5) являются оптимальными в дифференциальной игре (5.8.1). Стратегию $\bar{u}(x, y)$, определяемую соотношением (5.8.4), называют «погонной стратегией», так как в каждый момент времени вектор скорости игрока P при использовании этой стратегии нацелен на преследуемого игрока E .

5.8.2. Пример 5 (Игра преследования при наличии сил трения). Преследование происходит на плоскости. Уравнения движения имеют следующий вид:

для P :

$$\begin{aligned} \dot{q}_i &= p_i, \\ \dot{p}_i &= \alpha u_i - k_P p_i, \quad i = 1, 2, \quad \|u\| \leq 1; \end{aligned} \quad (5.8.6)$$

для E :

$$\begin{aligned} \dot{r}_i &= s_i, \\ \dot{s}_i &= \beta v_i - k_E s_i, \quad i = 1, 2, \quad \|v\| \leq 1; \\ q_i(0) &= q_i^0, \quad p_i(0) = p_i^0, \quad r_i(0) = r_i^0, \\ s_i(0) &= s_i^0, \quad i = 1, 2, \quad \alpha, \beta, k_E, k_P > 0. \end{aligned} \quad (5.8.7)$$

Здесь $q = (q_1, q_2)$ и $r = (r_1, r_2)$ — местоположение на плоскости игроков P и E соответственно; $p = (p_1, p_2)$ и $s = (s_1, s_2)$ — их импульсы; k_P, k_E — некоторые константы, интерпретируемые как коэффициенты трения.

Выигрыш игрока E полагается равным

$$H(q(T), r(T)) = \|q(T) - r(T)\| = \sqrt{[q_1(T) - r_1(T)]^2 + [q_2(T) - r_2(T)]^2}.$$

В плоскости $q = (q_1, q_2)$ множество достижимости $C_P^T(q^0, p^0)$ игрока P из начальных состояний $p(0) = p^0, q(0) = q^0$ за время T представляет собой круг (см. упр. 18) радиуса

$$R_P(T) = \frac{\alpha}{k_P^2} (e^{-k_P T} + k_P T - 1)$$

с центром в точке

$$a(q^0, p^0, T) = q^0 + p^0 \frac{1 - e^{-k_P T}}{k_P}.$$

Аналогично, множество $C_E^T(r^0, s^0)$ — круг радиуса

$$R_E(T) = \frac{\beta}{k_E^2} (e^{-k_E T} + k_E T - 1)$$

с центром в точке

$$b(r^0, s^0, T) = r^0 + s^0 \frac{1 - e^{-k_E T}}{k_E}.$$

Для величины $\hat{\rho}_T(q^0, p^0, r^0, s^0)$, определяемой соотношением (5.5.1), в данной дифференциальной игре выполняется равенство

$$\hat{\rho}_T(q^0, p^0, r^0, s^0) = \max_{r \in C_E^T(r^0, s^0)} \min_{q \in C_P^T(q^0, p^0)} \|q - r\|.$$

Отсюда (см. формулу (5.2.10)) имеем

$$\begin{aligned} \hat{\rho}_T(q, p, r, s) &= \max\{0, \|a(q, p, T) - b(r, s, T)\| - (R_P(T) - R_E(T))\} = \\ &= \max\left\{0, \sqrt{\sum_{i=1}^2 \left(q_i - r_i + p_i \frac{1 - e^{-k_P T}}{k_P} - s_i \frac{1 - e^{-k_E T}}{k_E}\right)^2} - \right. \\ &\quad \left. - \left(\alpha \frac{e^{-k_P T} + k_P T - 1}{k_P^2} - \beta \frac{e^{-k_E T} + k_E T - 1}{k_E^2}\right)\right\}. \end{aligned} \quad (5.8.9)$$

В частности, условий $\alpha > \beta, \frac{\alpha}{k_P} > \frac{\beta}{k_E}$ достаточно, чтобы для любых начальных состояний q, p, r, s нашлось отвечающее им T , при котором $\hat{\rho}_T(q, p, r, s) = 0$.

Функция $\hat{\rho}_T(q, p, r, s)$ удовлетворяет дифференциально-экстремальному уравнению (5.6.1) в области $\Omega = \{(q, p, r, s, T) : \hat{\rho}_T(q, p, r, s) > 0\}$. Действительно, в области Ω существуют непрерывные частные производные

$$\frac{\partial \hat{\rho}}{\partial T}, \frac{\partial \hat{\rho}}{\partial q_i}, \frac{\partial \hat{\rho}}{\partial p_i}, \frac{\partial \hat{\rho}}{\partial r_i}, \frac{\partial \hat{\rho}}{\partial s_i}, \quad i = 1, 2. \quad (5.8.10)$$

Уравнение (5.6.1) принимает вид

$$\begin{aligned} \frac{\partial \hat{\rho}}{\partial T} - \sum_{i=1}^2 \left(\frac{\partial \hat{\rho}}{\partial q_i} p_i + \frac{\partial \hat{\rho}}{\partial r_i} s_i - \frac{\partial \hat{\rho}}{\partial p_i} k_P p_i - \frac{\partial \hat{\rho}}{\partial s_i} k_E s_i \right) - \\ - \beta \max_{\|v\| \leq 1} \sum_{i=1}^2 \frac{\partial \hat{\rho}}{\partial s_i} v_i - \alpha \min_{\|u\| \leq 1} \sum_{i=1}^2 \frac{\partial \hat{\rho}}{\partial p_i} u_i = 0. \end{aligned} \quad (5.8.11)$$

Здесь экстремумы достигаются на управлениях \bar{u}, \bar{v} , определяемых следующими формулами:

$$\bar{u}_i = - \frac{\frac{\partial \hat{\rho}}{\partial p_i}}{\sqrt{(\frac{\partial \hat{\rho}}{\partial p_1})^2 + (\frac{\partial \hat{\rho}}{\partial p_2})^2}}, \quad (5.8.12)$$

$$\bar{v}_i = \frac{\frac{\partial \hat{\rho}}{\partial s_i}}{\sqrt{(\frac{\partial \hat{\rho}}{\partial s_1})^2 + (\frac{\partial \hat{\rho}}{\partial s_2})^2}}, \quad i = 1, 2. \quad (5.8.13)$$

Подставляя эти управления в (5.8.11), получим нелинейное уравнение в частных производных первого порядка

$$\begin{aligned} \frac{\partial \hat{\rho}}{\partial T} - \sum_{i=1}^2 \left(\frac{\partial \hat{\rho}}{\partial q_i} p_i + \frac{\partial \hat{\rho}}{\partial r_i} s_i - \frac{\partial \hat{\rho}}{\partial p_i} k_P p_i - \frac{\partial \hat{\rho}}{\partial s_i} k_E s_i \right) - \beta \sqrt{(\frac{\partial \hat{\rho}}{\partial s_1})^2 + (\frac{\partial \hat{\rho}}{\partial s_2})^2} + \\ + \alpha \sqrt{(\frac{\partial \hat{\rho}}{\partial p_1})^2 + (\frac{\partial \hat{\rho}}{\partial p_2})^2} = 0. \end{aligned} \quad (5.8.14)$$

Вычисляя частные производные (5.8.10), убеждаемся, что функция $\hat{\rho}_T(q, p, r, s)$ в области Ω удовлетворяет уравнению (5.8.14). Отметим, что величина $\hat{\rho}_T(q^0, p^0, r^0, s^0)$ является значением дифференциальной игры (5.8.6)–(5.8.8), а управления, определяемые соотношениями (5.8.12), (5.8.13), оптимальные в области Ω .

Из формул (5.8.12), (5.8.13), (5.8.9) находим

$$\bar{u}_i = \frac{r_i - q_i + s_i \frac{1-e^{-k_E T}}{k_E} - p_i \frac{1-e^{-k_P T}}{k_P}}{\sqrt{\sum_{i=1}^2 (r_i - q_i + s_i \frac{1-e^{-k_E T}}{k_E} - p_i \frac{1-e^{-k_P T}}{k_P})^2}}, \quad \bar{v}_i = \bar{u}_i, \quad i = 1, 2. \quad (5.8.15)$$

В ситуации \bar{u}, \bar{v} направление действия силы каждого из игроков параллельно линии, соединяющей центры кругов достижимости (как это следует из формулы (5.8.15)), и остается постоянным, поскольку в этой ситуации центры кругов достижимости перемещаются вдоль прямой линии.

§ 5.9. Игры преследования с задержкой информации у преследователя

5.9.1. Ранее в этой главе рассматривались конфликтно-управляемые процессы, в которых каждый из участников (игроков) имел полную информацию, т. е. в каждый текущий момент игры $P(E)$ знал свое состояние $x(t)[y(t)]$ и состояние противника $y(t)[x(t)]$. Были получены теоремы о существовании ситуаций ε -равновесия в чистых стратегиях в таких играх и проиллюстрированы различные методы построения движения. Это оказалось возможным, поскольку дифференциальные игры с полной информацией представляют собой предельный случай многошаговых игр с полной информацией, когда промежуток времени между двумя последовательными ходами стремится к нулю. Иначе обстоит дело с дифференциальными играми с неполной информацией, где применение смешанных стратегий играет существенную роль. Не останавливаясь на анализе всей проблемы, рассмотрим только случай игры преследования с предписанной

продолжительностью, терминальным выигрышем и задержкой поступления информации игроку P о фазовом состоянии игрока E на время $l > 0$.

5.9.2. Пусть задано некоторое число $l > 0$, называемое временем задержки информации. При $0 \leq t \leq l$, преследователь P в каждый момент времени t знает свое состояние $x(t)$, время t и начальное местоположение y_0 убегающего E . При $l \leq t \leq T$, игрок P в каждый момент t знает свое состояние $x(t)$, время t и состояние $y(t-l)$ игрока E в момент $t-l$. Игрок E в каждый момент времени t знает свое состояние $y(t)$, состояние противника $x(t)$ и время t . Его выигрыш равен расстоянию между игроками в момент времени T , выигрыш игрока P равен выигрышу E с обратным знаком (игра антагонистическая). Обозначим эту игру $\Gamma(x_0, y_0, T)$.

Определение. Под кусочно-программной чистой стратегией $v(\cdot)$ игрока E будем понимать пару $\{\tau, b\}$, где τ — разбиение отрезка времени $[0, T]$ конечным числом точек $0 = t_1 < \dots < t_k = T$, и b — отображение, которое каждому состоянию $x(t_i), y(t_i), t_i$ ставит в соответствие отрезок измеримого программного управления $v(t)$ игрока E при $t \in [t_i, t_{i+1})$.

Определение. Под кусочно-программной чистой стратегией $u(\cdot)$ игрока P будем понимать пару $\{\sigma, a\}$, где σ — произвольное разбиение отрезка времени $[0, T]$ конечным числом точек $0 = t'_1 < t'_2 < \dots < t'_s = T$, и a — отображение, которое каждому состоянию $x(t'_i), y(t'_i - l), t'_i$ при $l \leq t'_i$ ставит в соответствие отрезок измеримого программного управления $u(t)$ игрока P при $t \in [t'_i, t'_{i+1})$. При $t'_i \leq l$, отображение a каждому состоянию $x(t'_i), y_0, t'_i$ ставит в соответствие отрезок измеримого управления $u(t)$ при $t \in [t'_i, t'_{i+1})$.

Множества всех кусочно-программных чистых стратегий игроков P и E будем обозначать соответственно через P и E . Уравнения движения имеют вид

$$\begin{aligned}\dot{x} &= f(x, u), \quad u \in U \subset R^p, \quad x \in R^n, \\ \dot{y} &= g(y, v), \quad v \in V \subset R^q, \quad y \in R^n.\end{aligned}\tag{5.9.1}$$

Полагаем выполненными все условия, обеспечивающие существование и единственность решения системы (5.9.1) для любой пары измеримых программных управлений $u(t), v(t)$ при заданных начальных условиях x_0, y_0 . Это гарантирует существование единственного решения системы (5.9.1) в случае использования игроками P и E кусочно-программных стратегий $u(\cdot) \in P, v(\cdot) \in E$ при заданных начальных условиях x_0, y_0 . Таким образом, в любой ситуации $(u(\cdot), v(\cdot))$ при заданных начальных условиях x_0, y_0 функция выигрыша игрока E определяется однозначно

$$K(x_0, y_0; u(\cdot), v(\cdot)) = \rho(x(T), y(T)),\tag{5.9.2}$$

где $x(t), y(t)$ — решение системы (5.9.1) при начальных условиях x_0, y_0 в ситуации $(u(\cdot), v(\cdot))$, и ρ — евклидово расстояние.

5.9.3. Можно на простейших примерах показать, что в рассматриваемой игре $\Gamma(x_0, y_0, T)$ ситуации ε -равновесия существуют не для всех чисел $\varepsilon > 0$. Поэтому для построения ситуаций равновесия воспользуемся подходом, предложенным Ф. Нейманом и О. Моргенштерном для конечных позиционных игр с неполной информацией [Фон Нейман и Моргенштейн, 1970]. Расширим пространства стратегий игроков P и E до так называемых смешанных кусочно-программных стратегий поведения (СКПСР), которые предполагают возможность случайного выбора управления на каждом шаге.

Пример 6. Уравнения движения имеют вид

$$\begin{aligned} \text{для } P : \dot{x} &= u, \quad \|u\| \leq \alpha, \\ \text{для } E : \dot{y} &= v, \quad \|v\| \leq \beta, \\ \alpha > \beta > 0, \quad x, y, &\in R^2, \quad u, v \in R^2. \end{aligned} \quad (5.9.3)$$

Выигрыш игрока E равен $\rho(x(T), y(T))$, где $x(t), y(t)$ — решение системы (5.9.3) при начальных условиях $x(t_0) = x_0, y(t_0) = y_0$. Игрок P в течение игры знает лишь начальное состояние y_0 противника, а игрок E имеет полную информацию о состоянии игрока P ($l = T$).

Пусть $\bar{v}(x, y, t)$ — некоторая кусочно-программная стратегия игрока E . Для каждой стратегии \bar{v} существует стратегия $\bar{u}(x, t)$ игрока P , использующая только информацию о начальном положении игрока E , своем текущем положении и времени, прошедшем с момента начала игры, гарантирующая выигрыш $\rho(x(T), y(T)) \leq \varepsilon$ при $T \geq \rho(x_0, y_0)/(\alpha - \beta)$. Действительно, пусть $u^*(x, y, t)$ — стратегия игрока P в игре с полной информацией, имеющая следующую структуру: до момента встречи t_n осуществляется погонное преследование игрока E , а при $t_n \leq t \leq T$ точка $x(t)$ сохраняется в некоторой ε -окрестности убегающей точки. Такая стратегия в игре с полной информацией может быть легко описана аналитически (см. пример 4 п. 5.8.1). Построим траектории $\bar{x}(t), \bar{y}(t)$ движения игроков в ситуации $(u^*(x, y, t), \bar{v}(x, y, t))$ из начальных состояний x_0, y_0 . Для этого достаточно проинтегрировать систему

$$\begin{aligned} \dot{x} &= u^*(x, y, t), \quad \bar{x}(t_0) = x_0, \\ \dot{y} &= \bar{v}(x, y, t), \quad \bar{y}(t_0) = y_0. \end{aligned} \quad (5.9.4)$$

По построению $\rho(\bar{x}(T), \bar{y}(T)) \leq \varepsilon$. Пусть теперь $\tilde{u}(t) = u^*(\bar{x}(t), \bar{y}(t), t)$ и хотя стратегия $u^*(x, y, t)$, использующая для выработки управления информацию о положении E , недопустима, стратегия $\tilde{u}(t)$ является допустимой, поскольку использует лишь информацию о времени, прошедшем с момента начала игры и о начальном состоянии игрока E . Очевидно, что в ситуациях $(\tilde{u}(t), \bar{v}(x, y, t))$ и $(u^*(x, y, t), \bar{v}(x, y, t))$ траектории игроков совпадают, поскольку стратегия $\bar{v}(x, y, t)$ одинаково реагирует как на стратегию $u^*(x, y, t)$, так и на стратегию $\tilde{u}(t)$ выбором управления $\bar{v}(\bar{x}(t), \bar{y}(t), t)$.

Таким образом, мы показали, что для каждой стратегии $\bar{v}(x, y, t)$ существует программное управление $\tilde{u}(t)$, являющееся допустимой стратегией в игре с неполной информацией, и такое, что $\rho(\bar{x}(T), \bar{y}(T)) \leq \varepsilon$, где $\bar{x}(t), \bar{y}(t)$ — соответствующие траектории. Выбор $\bar{v}(x, y, t)$ произволен, поэтому отсюда следует, что

$$\sup \inf \rho(x(T), y(T)) = 0, \quad (5.9.5)$$

где $\sup \inf$ берется по множествам стратегии игроков в игре с неполной информацией.

Вместе с тем для любой стратегии $u(x, t)$ игрока P можно построить такую стратегию $v(x, y, t)$ для игрока E , что в ситуации $(u(x, t), v(x, y, t))$ выигрыш ρ игрока E превзойдет βT . Действительно, пусть $\bar{u}(x, t)$ — некоторая стратегия игрока P . Так как его движение не зависит от $y(t)$, то траектория движения игрока P может быть получена интегрированием системы

$$\dot{x} = \bar{u}(x, t), \quad x(t_0) = x_0, \quad (5.9.6)$$

независимо от движения игрока E . Пусть $\bar{x}(t)$ — траектория, получившаяся в результате интегрирования системы (5.9.6). Соединим точки $\bar{x}(T)$ и y_0 и направим движение игрока E по прямой $[\bar{x}(T), y_0]$ в направлении от точки $\bar{x}(T)$ с максимальной скоростью. Очевидно, что такое движение игрока E обеспечивает расстояние между ним и точкой $\bar{x}(T)$ большее или равное βT . Обозначим построенную таким образом стратегию игрока E через $\bar{v}(t)$. Тогда получим, что в ситуации $(\bar{u}(x, t), \bar{v}(t))$, выигрыш игрока E больше или равен величине βT . Отсюда следует, что

$$\inf \sup \rho(x(T), y(T)) \geq \beta T, \quad (5.9.7)$$

где $\inf \sup$ берется по множествам стратегий игроков в игре с неполной информацией.

Из (5.9.5) и (5.9.7) следует, что значение игры в классе чистых стратегий в рассматриваемой игре не существует.

5.9.4. Определение. Под смешанной кусочно-программной стратегией поведения (СКПСП) игрока P будем понимать пару $\mu(\cdot) = \{\tau, d\}$, где τ — произвольное разбиение отрезка времени $[0, T]$ конечным числом точек $0 = t_1 < t_2 < \dots < t_k = T$, а d — отображение, ставящее в соответствие состоянию $x(t_i), y(t_i - l), t_i$ при $t_i > l$ и состоянию $x(t_i), y_0, t_i$ при $t_i \leq l$ вероятностное распределение $\mu'_i(\cdot)$, сосредоточенное на конечном числе измеримых программных управлений $u(t)$ при $t \in [t_i, t_{i+1})$.

Аналогично под СКПСП игрока E будем понимать пару $\nu(\cdot) = \{\sigma, c\}$, где σ — произвольное разбиение отрезка времени $[0, T]$ конечным числом точек $0 = t'_1 < t'_2 < \dots < t'_s = T$, а c — отображение, ставящее в соответствие состоянию $x(t'_i), y(t'_i), t'_i$ вероятностное распределение $\nu'_i(\cdot)$, сосредоточенное на конечном числе измеримых программных управлений $v(t)$ при $t \in [t_i, t_{i+1})$.

СКПСП игроков P и E будем обозначать соответственно через \bar{P} и \bar{E} (ср. со «стратегиями поведения» п. 4.8.3).

Каждая пара СКПСП $\mu(\cdot), \nu(\cdot)$ индуцирует распределение вероятностей на пространстве траекторий $x(t), x(0) = x_0; y(t), y(0) = y_0$. Поэтому под выигрышем $\bar{K}(x_0, y_0; \mu(\cdot), \nu(\cdot))$ в СКПСП будем понимать математическое ожидание выигрыша $K(x_0, y_0; \mu(\cdot), \nu(\cdot))$, усредненное по распределениям на пространствах траекторий, которые индуцируются СКПСП $\mu(\cdot), \nu(\cdot)$. Определив пространства стратегий \bar{P}, \bar{E} и выигрыш \bar{K} , мы определили смешанное расширение $\bar{\Gamma}(x_0, y_0, T)$ игры $\Gamma(x_0, y_0, T)$.

5.9.5. Обозначим через $C_P^T(x)$ и $C_E^T(y)$ соответственно множества достижимости игроков P и E из начальных состояний x и y в момент времени T , и через $\bar{C}_E^T(y)$ выпуклую оболочку множества $C_E^T(y)$. Предположим, что множества достижимости компактны, и введем в рассмотрение величину

$$\gamma(y, T) = \min_{\xi \in \bar{C}_E^T(y)} \max_{\eta \in C_E^T(y)} \rho(\xi, \eta).$$

Пусть $\gamma(y, T) = \rho(\tilde{y}, \bar{y})$, где $\tilde{y} \in \bar{C}_E^T(y)$, $\bar{y} \in C_E^T(y)$. Из определения точки \tilde{y} следует, что она является центром минимальной сферы, содержащей множество $C_E^T(y)$. Отсюда получаем, что эта точка единственна. В то же время существуют по крайней мере две точки касания множества $C_E^T(y)$ с минимальной содержащей его сферой, которые совпадают с точками \bar{y} .

Пусть $y(t)$ — некоторая траектория ($y(0) = y_0$) игрока E при $0 \leq t \leq T$. При перемещении игрока E вдоль этой траектории величина $\gamma(y(t), T - t)$ изменяется, меняется также и точка \tilde{y} . Пусть $\tilde{y}(t)$ — траектория точки \tilde{y} , соответствующая траектории $y(t)$.

Назовем точку $M \in C_E^{T-l}(y_0)$ центром преследования, если

$$\gamma(M, l) = \max_{y' \in C_E^{T-l}(y_0)} \gamma(y', l).$$

5.9.6. Рассмотрим вспомогательную одновременную антагонистическую игру преследования на выпуклой оболочке множества $C_E^T(y)$. Преследователь выбирает некоторую точку $\xi \in \bar{C}_E^T(y)$, а убегающий — точку $\eta \in C_E^T(y)$. Выбор совершается одновременно, и игрок P при выборе точки ξ не знает выбора η игрока E , и наоборот. Игрок E получает выигрыш $\rho(\xi, \eta)$. Обозначим значение этой игры через $V(y, T)$, чтобы подчеркнуть зависимость значения игры от параметров y и T , которые определяют множества стратегий $\bar{C}_E^T(y)$ и $C_E^T(y)$ игроков P и E соответственно. Игру в нормальной форме запишем следующим образом:

$$\Gamma(y, T) = \langle \bar{C}_E^T(y), C_E^T(y), \rho(y', y'') \rangle.$$

Множество стратегий минимизирующего игрока P выпукло, функция $\rho(y', y'')$ также выпукла по своим аргументам и непрерывна. Для таких игр мы можем применить теорему п. 2.5.5, поэтому в игре $\Gamma(y, T)$ существует ситуация равновесия в смешанных стратегиях. Оптимальная стратегия игрока P чистая, а оптимальная стратегия игрока E предписывает положительную вероятность не более чем $(n+1)$ точке из множества $C_E^T(y)$, причем $V(y, T) = \gamma(y, T)$. Оптимальная стратегия игрока P в игре $\Gamma(y, T)$ заключается в выборе центра минимальной сферы \tilde{y} , содержащей множество $C_E^T(y)$. Оптимальная стратегия игрока E предписывает положительные вероятности не более чем $(n+1)$ точке из точек касания указанной сферы с множеством $C_E^T(y)$ (здесь n — размерность пространства y). Значение игры равно радиусу этой сферы (см. пример 11 п. 2.5.5).

5.9.7. Рассмотрим одновременную игру $\Gamma(M, l)$, где M — центр преследования. Обозначим через $\bar{y}_1(M), \dots, \bar{y}_{n+1}(M)$ точки из множества $C_E^l(M)$, которые входят в спектр оптимальной смешанной стратегии игрока E в игре $\Gamma(M, l)$ и через $y(M)$ — оптимальную стратегию игрока P в этой игре.

Определение. Траектория $y^*(t)$ называется условно-оптимальной, если $y^*(0) = y_0$, $y^*(T-l) = M$, $y^*(T) = \bar{y}_i(M)$ для некоторого i из чисел $1, \dots, n+1$.

Для каждого i может существовать несколько условно-оптимальных траекторий игрока E .

Теорема. Пусть $T \geq l$ и для любого числа $\varepsilon > 0$ игрок P к моменту времени T может гарантировать ε -встречу с центром $y(T)$ минимальной сферы, содержащей множество $C_E^l(y(T-l))$. Тогда игра $\bar{\Gamma}(x_0, y_0, T)$ имеет значение $\gamma(M, l)$, и ε -оптимальная стратегия игрока P чистая и совпадает с любой его стратегией, гарантирующей $\varepsilon/2$ -встречу с точкой $y(T)$. Оптимальная стратегия игрока E смешанная: в течение времени $0 \leq t \leq T-l$ он должен перемещаться в точку M по любой условно-оптимальной траектории $y^*(t)$ и далее с вероятностями p_1, \dots, p_{n+1} (оптимальная стратегия игрока E в игре $\Gamma(M, l)$), выбрать одну из условно-оптимальных траекторий, переводящих точку $y^*(T-l) = M$ в точки $\bar{y}_i(M)$, $i = 1, \dots, n+1$, входящие в спектр оптимальной смешанной стратегии игрока E в игре $\Gamma(M, l)$.

Доказательство. Обозначим через $u_\varepsilon(\cdot), \nu_*(\cdot)$ указанные в теореме стратегии, оптимальность которых требуется доказать. Для доказательства теоремы достаточно убедиться в справедливости следующих соотношений:

$$\begin{aligned} \bar{K}(x_0, y_0; \mu(\cdot), \nu_*(\cdot)) + \varepsilon &\geq \bar{K}(x_0, y_0; u_\varepsilon(\cdot), \nu_*(\cdot)) \geq \\ &\geq \bar{K}(x_0, y_0; u_\varepsilon(\cdot), \nu(\cdot)) - \varepsilon, \quad \mu(\cdot) \in \bar{P}, \quad \nu(\cdot) \in \bar{E}, \end{aligned} \quad (5.9.8)$$

$$\lim_{\varepsilon \rightarrow 0} \bar{K}(x_0, y_0; u_\varepsilon(\cdot), \nu_*(\cdot)) = \gamma(M, l). \quad (5.9.9)$$

Левая часть неравенства (5.9.8) следует из определения стратегии $u_\varepsilon(\cdot)$, в силу которого для любой кусочно-программной стратегии $u(\cdot) \in P$

$$\bar{K}(x_0, y_0; u(\cdot), \nu_*(\cdot)) + \varepsilon \geq \bar{K}(x_0, y_0; u_\varepsilon(\cdot), \nu_*(\cdot)).$$

Обозначим через $x^*(t)$ траекторию преследователя в ситуации $(u_\varepsilon(\cdot), \nu_*(\cdot))$. Тогда

$$\bar{K}(x_0, y_0; u_\varepsilon(\cdot), \nu_*(\cdot)) = \sum_{i=1}^{n+1} p_i \rho(x^*(T), \bar{y}_i(M)). \quad (5.9.10)$$

Пусть R — радиус минимальной сферы, содержащей множество $C_E^l(M)$, т. е. $R = \gamma(M, l)$. Тогда $R - \varepsilon/2 \leq \rho(x^*(T), \bar{y}_i(M)) \leq R + \varepsilon/2$ для всех $i = 1, \dots, n+1$, поскольку точка $x^*(T)$ принадлежит $\varepsilon/2$ -окрестности точки $y(M)$. Так как $\sum_{i=1}^{n+1} p_i = 1$, $p_i \geq 0$, из (5.9.10) получаем

$$R - \varepsilon/2 \leq \bar{K}(x_0, y_0; u_\varepsilon(\cdot), \nu_*(\cdot)) \leq R + \varepsilon/2, \quad (5.9.11)$$

что доказывает (5.9.9).

Пусть состояния $x(T), y(T-l)$ реализовались в ситуации $(u_\varepsilon(\cdot), \nu(\cdot))$ и $Q(\cdot)$ — вероятностная мера, индуцированная на множестве $C_E^l(y(T-l))$. Из оптимальности смешанной стратегии $p = (p_1, \dots, p_{n+1})$ в игре $\Gamma(M, l)$ имеем

$$\begin{aligned} \sum_{i=1}^{n+1} p_i \rho(x(T), \bar{y}_i(M)) &\geq \gamma(y(T-l), l) = Val \Gamma(y(T-l), l) \geq \\ &\geq \int_{C_E^l(y(T-l))} \rho(y[y(T-l)], y) dQ, \end{aligned} \quad (5.9.12)$$

где $y[y(T-l)]$ — центр минимальной сферы, содержащей множество $C_E^l(y(T-l))$.

Однако, $\rho(x(T), y[y(T-l)]) \leq \varepsilon/2$, следовательно, при $y \in C_E^l(y(T-l))$ имеем

$$\rho(x(T), y) \leq \frac{\varepsilon}{2} + \rho(y[y(T-l)], y) \leq R + \varepsilon/2. \quad (5.9.13)$$

Из неравенства (5.9.11)–(5.9.13) вытекает, что

$$\bar{K}(x_0, y_0; u_\varepsilon(\cdot), \nu_*(\cdot)) \geq \int_{C_E^l(y(T-l))} \rho(x(T), y) dQ - \varepsilon. \quad (5.9.14)$$

Однако

$$\int_{C_E^l(y(T-l))} \rho(x(T), y) dQ = \bar{K}(x_0, y_0; u_\varepsilon(\cdot), \nu(\cdot)). \quad (5.9.15)$$

Из формул (5.9.14) и (5.9.15) получаем правую часть неравенства (5.9.8). Теорема доказана.

При $T < l$ решение игры существенно не отличается от случая $T \geq l$ и теорема сохраняет силу, если вместо $C_E^l(y_0)$, $\bar{C}_E^l(y_0)$, $\gamma(M, l)$, $y(T - l)$, y_0 рассматривать $C_E^T(y_0)$, $\bar{C}_E^T(y_0)$, $\gamma(M, T)$.

Диаметр множества $C_E^l(M)$ стремится к нулю при $l \rightarrow 0$, что, в свою очередь, вызывает стремление к нулю значения вспомогательной игры $\Gamma(M, l)$. Однако значение этой вспомогательной игры равно значению $V_l(x_0, y_0, T)$ игры преследования с задержкой информации $\bar{\Gamma}(x_0, y_0, T)$ (здесь индекс l означает время задержки информации). Смешанная оптимальная стратегия игрока E в $\Gamma(M, l)$, сосредоточивающая свою массу на не более чем в $n + 1$ точке из $C_E^l(M)$, в пределе сосредоточивает всю массу в одной точке M , т. е. превращается в чистую стратегию. Это вполне согласуется с тем, что при $l \rightarrow 0$ игра $\Gamma(x_0, y_0, T)$ превращается в игру с полной информацией.

Пример 7. Уравнения движения имеют вид

$$\dot{x} = u, \quad \|u\| \leq \alpha; \quad \dot{y} = v, \quad \|v\| \leq \beta, \quad \alpha > \beta, \quad x, y \in R^2.$$

Пусть время T удовлетворяет условию $T > \rho(x_0, y_0)/(\alpha - \beta) + l$. Множество достижимости $C_E^l(y_0) = \bar{C}_E^l(y_0)$ и совпадает с кругом радиуса βl с центром y_0 . Значение игры $\Gamma(y, l)$ равно радиусу круга $C_E^l(y)$, т. е. $V(y, l) = \beta l$.

Так как величина $V(y, l)$ в данном случае не зависит от y , то любая точка множества $C_E^{T-l}(y_0)$ может быть центром преследования M . Оптимальная стратегия игрока P в игре $\Gamma(y, l)$ заключается в выборе точки y , а оптимальная стратегия игрока E — смешанная и заключается в выборе двух любых диаметрально противоположных точек круга $C_E^l(y)$ с вероятностями $(1/2, 1/2)$. В соответствии с этим оптимальная стратегия преследователя в игре $\bar{\Gamma}(x_0, y_0, T)$ заключается в погонном преследовании точки $y(t - l)$ при $l \leq t \leq T$ (точки y_0 при $0 \leq t \leq l$) до встречи с этой точкой, кроме того, до момента T следует оставаться в $\varepsilon/2$ -окрестности этой точки. Оптимальная стратегия игрока E — смешанная кусочно-программная стратегия поведения — и заключается в переходе из точки y_0 в произвольную точку $M \in C_E^{T-l}(y_0)$ в течение времени $T - l$, далее в равновероятном выборе направления на одну из двух диаметрально противоположных точек круга $C_E^l(M)$. В этом случае $Val \bar{\Gamma}(x_0, y_0, T) = \beta l$.

§ 5.10. Упражнения и задачи

1. Построить множество достижимости в игре «простое движение» для игрока P и игрока E .

2. Пусть игрок E перемещается из точки $y^0 = (y_1^0, y_2^0)$ с постоянной по величине и направлению скоростью β . Показать, что для каждого такого движения существует единственное движение игрока P из точки $x^0 = (x_1^0, x_2^0)$ с постоянной скоростью α ($\alpha > \beta$), которое осуществляет встречу (l -встречу) с игроком E за минимальное время. Такое движение игрока P будем называть быстроедействием в точку встречи.

3. Пусть игрок E перемещается из точки $y^0 = (y_1^0, y_2^0)$ с постоянной по величине и направлению скоростью β , а игрок P осуществляет быстроедействие в точку встречи из точки $x^0 = (x_1^0, x_2^0)$. Для каждой такой пары движений игроков E и P построить точку встречи. Показать, что полученное геометрическое место точек встречи игроков E и P представляет собой окружность Апполония, и написать ее уравнение.

4. В условиях предыдущего упражнения построить множество точек l -встречи игроков E и P .

5. Обозначим через $A(x_0, y_0)$ множество точек встречи относительно начальных состояний x_0, y_0 игроков P и E (окружность Апполония). Пусть до некоторого момента τ (τ меньше времени до момента встречи) игроки E и P перемещаются прямолинейно с максимальными скоростями в точку встречи M . Построим новое множество точек встречи $A(x(\tau), y(\tau))$ относительно состояний $x(\tau), y(\tau)$ как начальных в момент времени τ . Это некоторая новая окружность Апполония. Показать, что окружности $A(x_0, y_0)$ и $A(x(\tau), y(\tau))$ касаются в точке M , следовательно, $A(x(\tau), y(\tau))$ содержится в круге $A(x_0, y_0)$, ограниченном окружностью $A(x_0, y_0)$.

6. Пусть игрок E перемещается из точки y_0 вдоль некоторой гладкой кривой $y(t)$ с максимальной скоростью β . Игрок P движется с максимальной скоростью α , в каждый момент времени τ зная местоположение $y(\tau)$ игрока E и направление вектора скорости $v(\tau) = \{v_1(\tau), v_2(\tau)\}$, $v_1^2(\tau) + v_2^2(\tau) = \beta^2$. Построим П-стратегию игрока P . Согласно этой стратегии он выбирает направление вектора скорости на точку встречи M в предположении, что игрок E будет на отрезке времени $[\tau, \infty)$ придерживаться постоянного направления движения $\{v_1(\tau), v_2(\tau)\}$ (перемещаться вдоль луча с постоянной скоростью β).

Показать, что если игрок P использует П-стратегию, то отрезок $[x(\tau), y(\tau)]$, соединяющий текущие местоположения игроков, останется до момента встречи параллельным отрезку $[x_0, y_0]$.

7. Пусть игрок E перемещается из y_0 вдоль некоторой гладкой кривой $y(\tau)$ с максимальной скоростью β . Написать аналитическое выражение для П-стратегии игрока P .

8. Показать, что при использовании П-стратегии игроком P точка встречи всегда содержится во множестве $A(x_0, y_0)$, ограниченном окружностью Апполония $A(x_0, y_0)$.

Указание. Доказательство провести сначала для движений игрока E вдоль k -вершинных ломаных, используя утверждение упр. 5, а затем совершить предельный переход.

9. (Игра «шофер—убийца»). Чтобы записать уравнения движения игроков в этой игре, достаточно задать пять фазовых координат: по две координаты для обозначения местоположения игроков P (автомобиль) и E (пешеход) и еще одну для обозначения направления движения преследователя. Обозначим их через $x_1, x_2, y_1, y_2, \vartheta$ (рис. 5.1). Задание этих фазовых координат полностью и однозначно определяет состояние игры в каждый момент.

Управление для игрока E выглядит просто. Для описания направления его движения достаточно задать угол ψ (см. рис. 5.1).

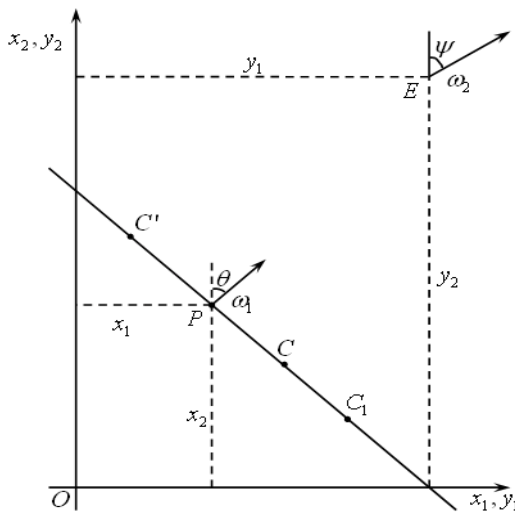


Рис. 5.1. Игра «шофер—убийца»

Теперь выберем управление для игрока P . Проведем через точку P прямую CC' ($|C'P| = |PC| = R$), перпендикулярную вектору скорости преследования. По своему желанию игрок P выбирает мгновенный центр кривизны своей траектории в любой точке, например, в точке C_1 , лежащей на этой прямой вне интервала $C'C$.

Управление u будем считать равным по абсолютной величине $R/|PC_1|$, положительным для точек C_1 , лежащих слева от P , и отрицательным — справа от P ; таким образом, $-1 \leq u \leq 1$. Доказать, что уравнения движения имеют следующий вид:

$$\begin{aligned} \dot{x}_1 &= \omega_1 \sin \vartheta, & \dot{x}_2 &= \omega_1 \cos \vartheta, \\ \dot{y}_1 &= \omega_2 \sin \psi, & \dot{y}_2 &= \omega_2 \cos \psi, \\ \dot{\vartheta} &= \omega_1 / Ru. \end{aligned}$$

10. (Игра «шофер-убийца». Понижение размерности.) Предположим, что на плоскости выбрана подвижная система координат, связанная с автомобилем P . Координаты пешехода y_1, y_2 можно рассматривать в этой системе как составляющие единственного переменного вектора x ; ось x_2 будем считать всегда направленной вдоль вектора скорости автомобиля.

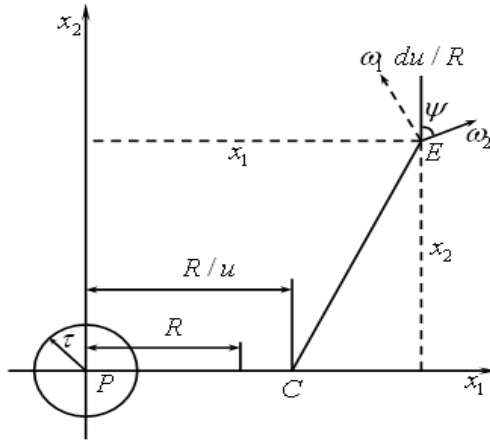


Рис. 5.2. Понижение размерности

Пусть игрок P в момент t выбирает центр кривизны своей траектории в точке $C = (R/u, 0)$ и пусть расстояние CE равно d (рис. 5.2). Тогда вращение игрока P вокруг точки C эквивалентно вращению x вокруг C в противоположном направлении, но с той же угловой скоростью. Таким образом, вектор x движется со скоростью, равной по модулю $\omega_1 (du/R)$, в направлении, перпендикулярном CE . Составляющие его скорости получаются умножением модуля соответственно на $-x_2/d$ и $(x_1 - R/\psi)/d$.

Показать, что уравнения движения имеют вид:

$$\begin{aligned} \dot{x}_1 &= -\frac{\omega_1}{R} x_2 u + \omega_2 \sin \psi, & \dot{x}_2 &= \frac{\omega_1}{R} x_1 u - \omega_1 + \omega_2 \cos \psi, \\ -1 &\leq u \leq +1, & 0 &\leq \psi \leq 2\pi. \end{aligned}$$

11. Пусть a и b — такие числа, что $\rho = \sqrt{a^2 + b^2} > 0$. Показать, что $\max_{\psi} (a \cos \psi + b \sin \psi)$ достигается на таком $\bar{\psi}$, что $\cos \bar{\psi} = a/\rho$, $\sin \bar{\psi} = b/\rho$ и этот максимум равен ρ .

12. Записать и проиллюстрировать уравнение (5.6.6) для игры на «перетягивание»

$$\dot{x}_1 = u + v, \quad |u| \leq \alpha, \quad \dot{x}_2 = u + v, \quad |v| \leq \beta, \quad x(0) = x_0,$$

с терминальным выигрышем $\rho(x(T), A)$, где A — некоторая точка, $A \in R^2$, лежащая вне множества достижимости системы к моменту времени T из начального состояния x_0 .

13. Записать явные выражения для оптимальных стратегий в игре упр. 12 и для ее модификации, когда продолжительность игры не фиксируется заранее, а выигрыш игрока E полагается равным времени попадания в начало координат.

14. Доказать, что множество достижимости управляемой системы

$$q_i(0) = q_i^0, \quad \dot{q}_i = p_i, \quad \dot{p}_i = \alpha u_i - k p_i, \quad u_1^2 + u_2^2 \leq 1, \quad i = 1, 2$$

в пространстве геометрических координат (q_1, q_2) — круг с центром в точке $\bar{q} = q^0 + p^0(1 - e^{-kT})/k$ и радиусом $R = \alpha(e^{-kT} + kT - 1)/k^2$.

15. Преследование происходит в плоскости, уравнения движения имеют следующий вид: для P

$$\dot{q}_i = p_i, \quad \dot{p}_i = \alpha u_i - k p_i, \quad u_1^2 + u_2^2 \leq 1, \quad i = 1, 2,$$

для E

$$\dot{y}_i = \beta v_i, \quad v_1^2 + v_2^2 \leq 1, \quad i = 1, 2.$$

Здесь q и y — местоположения игроков P и E соответственно, p — импульс игрока P . Таким образом, в рассмотренном случае игрок E движется согласно «простому движению», а игрок P , представляющий собой материальную точку единичной массы, перемещается под действием силы трения α .

Выигрыш игрока определим как расстояние между геометрическими местоположениями игроков в момент окончания игры T .

$$H(q(T), y(T)) = p(q(T), y(T)) = \sqrt{\sum_{i=1}^2 (q_i(T) - y_i(T))^2}.$$

Вычислить величину $\hat{\rho}_T(q, y)$.

16. Вывести уравнение (5.6.6) для задачи из упр. 15.

17. Рассмотреть игру «простое преследование» с предписанной продолжительностью T в полуплоскости F , т. е. при дополнительном предположении, что игроки в процессе преследования не могут покинуть множество F . Построить области достижимости игроков.

18. Вычислить величину $\hat{\rho}_T(x, y)$ для игры «простое преследование» на полуплоскости с предписанной продолжительностью.

19. Рассмотреть антагонистическую игру «простое преследование» с предписанной продолжительностью между двумя преследователями $P = \{P_1, P_2\}$ действующими как один игрок, и преследуемым игроком E . Уравнения движения имеют следующий вид:

$$\begin{aligned} \dot{x}^1 &= u^1, & |u^1| &\leq \alpha_1, \\ \dot{x}^2 &= u^2, & |u^2| &\leq \alpha_2, \\ x^1, x^2, y &\in R^2, & \beta &< \min\{\alpha_1, \alpha_2\}, \\ \dot{y} &= v, & |v| &\leq \beta, \quad u^1, u^2, v \in R^2, \\ x^1(0) &= x_0^1, \quad x^2(0) = x_0^2, \quad y(0) = y_0. \end{aligned}$$

Выигрыш игрока E равен $\min_{i=1,2} \rho(x^i(T), y(T))$, т. е. игрок E заинтересован в максимизации расстояния до ближайшего из преследователей к моменту окончания игры.

Построить множества достижимости игроков и геометрически определить максимальное расстояние $\hat{\rho}_T(x_0^1, x_0^2, y)$ между этими множествами.

20. Обобщить теорему п. 5.5.9 на случай, когда в преследовании участвуют несколько преследователей P_1, \dots, P_m , действующих как один игрок, и один убегающий игрок E .

Глава 6

Неантагонистические дифференциальные игры

§ 6.1. Принцип динамического программирования

6.1.1. Рассмотрим классическую задачу оптимального управления с одним критерием или с одним лицом, принимающим решение:

$$\max_u \left\{ \int_{t_0}^T g[s, x(s), u(s)] ds + q(x(T)) \right\}, \quad (6.1.1)$$

при условии, когда переменная состояния x изменяется в соответствии с дифференциальным уравнением:

$$\dot{x}(s) = f[s, x(s), u(s)], \quad x(t_0) = x_0, \quad (6.1.2)$$

где $x(s) \in X \subset R^m$ и управление $u \in U$. Функции $f[s, x, u]$, $g[s, x, u]$ и $q(x)$ предполагаются дифференцируемыми.

Программное управление $u^*(t)$, доставляющее максимум функционала (6.1.1), будем называть *оптимальным управлением*.

Для определения оптимального управления задачи (6.1.1)–(6.1.2) используется техника динамического программирования и принцип максимума. Техника динамического программирования впервые была введена Р. Беллманом [Bellman, 1957]. Его суть содержится в следующей теореме.

Теорема. Управление $u^*(t) = \phi^*(t, x)$ образует оптимальное решение задачи управления (6.1.1)–(6.1.2), если существует непрерывно-дифференцируемая функция $V(t, x)$, определенная на $[t_0, T] \times R^m \rightarrow R$ и удовлетворяющая следующему уравнению (уравнению Белмана):

$$\begin{aligned} -V_t(t, x) &= \max_u \{g[t, x, u] + V_x(t, x) f[t, x, u]\} = \\ &= \{g[t, x, \phi^*(t, x)] + V_x(t, x) f[t, x, \phi^*(t, x)]\}, \\ V(T, x) &= q(x). \end{aligned}$$

Здесь $V_t(t, x)$ и $V_x(t, x)$ — частные производные функции $V(t, x)$ по t и x соответственно.

Доказательство. Введем функцию Беллмана как значение максимизируемого функционала (6.1.1), (6.1.2) при оптимальном управлении, если задача решается из начального состояния x и начального момента времени t . А именно,

$$\begin{aligned} V(t, x) &= \max_u \left[\int_t^T g(s, x(s), u(s)) ds + q(x(T)) \right] = \\ &= \int_t^T g[s, x^*(s), \phi^*(s, x^*(s))] ds + q(x^*(T)) \end{aligned}$$

при граничном условии

$$V(T, x^*(T)) = q(x^*(T)),$$

и управляемой динамике

$$\dot{x}^*(s) = f[s, x^*(s), \phi^*(s, x^*(s))], \quad x^*(t_0) = x_0.$$

Если $u(s) \in U$ — некоторое другое управление, реализующее траекторию $x(s)$, то из условий теоремы 1 следует, что

$$\begin{aligned} g(t, x, u) + V_x(t, x) f(t, x, u) + V_t(t, x) &\leq 0, \\ g(t, x^*, u^*) + V_{x^*}(t, x^*) f(t, x^*, u^*) + V_t(t, x^*) &= 0. \end{aligned}$$

Интегрируя указанное выражение на промежутке $[t_0, T]$, получаем

$$\begin{aligned} \int_{t_0}^T g(s, x(s), u(s)) ds + V(T, x(T)) - V(t_0, x_0) &\leq 0, \\ \int_{t_0}^T g(s, x^*(s), u^*(s)) ds + V(T, x^*(T)) - V(t_0, x_0) &= 0. \end{aligned}$$

Исключая $V(t_0, x_0)$, имеем

$$\int_{t_0}^T g(s, x(s), u(s)) ds + q(x(T)) \leq \int_{t_0}^T g(s, x^*(s), u^*(s)) ds + q(x^*(T)),$$

Откуда следует, что $u^* = \phi^*(t, x)$ есть оптимальное управление.

Подставляя оптимальное управление (стратегию) $\phi^*(t, x)$ в (6.1.2), получаем уравнение для оптимальной траектории в виде

$$\dot{x}(s) = f[s, x(s), \phi^*(s, x(s))], \quad x(t_0) = x_0. \quad (6.1.3)$$

Обозначим через $x^*(t)$, $t \in [t_0, T]$ решение уравнения (6.1.3). Оптимальная траектория $\{x^*(t)\}_{t=t_0}^T$ может быть записана в виде

$$x^*(t) = x_0 + \int_{t_0}^t f[s, x^*(s), \phi^*(s, x^*(s))] ds, \quad t \in [t_0, T]. \quad (6.1.4)$$

Для упрощения обозначений в дальнейшем помимо обозначения $x^*(t)$ для оптимальной траектории мы будем использовать обозначение x_t^* .

Функция Беллмана или функция значения $V(t, x)$ из начального состояния на оптимальной траектории, т. е. при $x = x_t^*$, может быть записана в виде

$$V(t, x_t^*) = \int_t^T g[s, x^*(s), \phi^*(s, x^*(s))] ds + q(x^*(T)).$$

6.1.2. Пример 1. Рассмотрим следующую задачу оптимального управления

$$\max_u \left\{ \int_0^T \exp[-rs] [-x(s) - cu(s)^2] ds + \exp[-rT] qx(T) \right\} \quad (6.1.5)$$

при условии

$$\dot{x}(s) = a - u(s)(x(s))^{1/2}, \quad x(0) = x_0, \quad u(s) \geq 0, \quad (6.1.6)$$

где a, c, x_0, r, q — положительные вещественные числа.

Используя теорему п. 6.1.1, получаем

$$\begin{aligned} -V_t(t, x) &= \max_u \left\{ [-x - cu^2] \exp[-rt] + V_x(t, x) [a - ux^{1/2}] \right\}, \\ V(T, x) &= \exp[-rT] qx(T). \end{aligned} \quad (6.1.7)$$

Легко видеть, что максимизирующее управление в (6.1.7) имеет вид

$$\phi(t, x) = \frac{-V_x(t, x) x^{1/2}}{2c} \exp[rt].$$

Подставляя $\phi(t, x)$ в (6.1.7) и решая, получаем

$$V(t, x) = \exp[-rt] [A(t)x + B(t)],$$

где $A(t)$ и $B(t)$ удовлетворяют системе уравнений

$$\begin{aligned} \dot{A}(t) &= rA(t) - \frac{A(t)^2 e^{rt}}{4c} + 1, \\ \dot{B}(t) &= rB(t) - aA(t) e^{rt}, \end{aligned}$$

при граничном условии

$$A(T) = q, \quad B(T) = 0.$$

Таким образом, оптимальное управление может быть найдено в явном виде и определяется по формуле

$$\phi(t, x) = \frac{-A(t) x^{1/2}}{2c}.$$

6.1.3. Рассмотрим теперь задачу оптимизации на бесконечном промежутке времени с постоянным дисконтированием

$$\max_u \left\{ \int_{t_0}^{\infty} g[x(s), u(s)] \exp[-r(s - t_0)] ds \right\}, \quad (6.1.8)$$

при условии

$$\dot{x}(s) = f[x(s), u(s)], \quad x(t_0) = x_0. \quad (6.1.9)$$

Параметр s не входит явным образом в $g[x(s), u(s)]$, $f[x(s), u(s)]$. Такие задачи называются *автономными задачами управления*. Рассмотрим теперь семейство задач подобного рода из любых начальных состояний, т. е. задачу

$$\max_u \int_t^\infty g[x(s), u(s)] \exp[-r(s-t)] ds, \quad (6.1.10)$$

при условии

$$\dot{x}(s) = f[x(s), u(s)], \quad x(t) = x. \quad (6.1.11)$$

Задача (6.1.10)–(6.1.11) не зависит от выбора начального момента времени t , а зависит лишь от состояния, в котором система находится в данный момент, а именно от x .

Определим теперь функцию значений или функцию Беллмана для задачи (6.1.8)–(6.1.9) по аналогии с предыдущим случаем по формуле

$$V(t, x) = \max_u \left\{ \int_t^\infty g[x(s), u(s)] \exp[-r(s-t_0)] ds \mid x(t) = x = x_t^* \right\},$$

где x_t^* — состояние системы в момент t на оптимальной траектории. Более того, мы можем записать

$$\begin{aligned} V(t, x) &= \\ &= \exp[-r(t-t_0)] \max_u \left\{ \int_t^\infty g[x(s), u(s)] \exp[-r(s-t)] ds, \right. \\ &\quad \left. x(t) = x = x_t^* \right\}. \end{aligned}$$

Поскольку задача

$$\max_u \left\{ \int_t^\infty g[x(s), u(s)] \exp[-r(s-t)] ds \mid x(t) = x = x_t^* \right\}$$

зависит лишь от текущего состояния x , мы можем написать

$$W(x) = \max_u \left\{ \int_t^\infty g[x(s), u(s)] \exp[-r(s-t)] ds \mid x(t) = x = x_t^* \right\}.$$

Таким образом, получаем

$$\begin{aligned} V(t, x) &= \exp[-r(t-t_0)] W(x), \\ V_t(t, x) &= -r \exp[-r(t-t_0)] W(x), \\ V_x(t, x) &= \exp[-r(t-t_0)] W_x(x). \end{aligned} \quad (6.1.12)$$

Подставляя выражение (6.1.12) в основную формулу теоремы п. 6.1.1, получаем

$$rW(x) = \max_u \{g[x, u] + W_x(x) f[x, u]\}. \quad (6.1.13)$$

Поскольку время в (6.1.13) явно не присутствует, управление u будет функцией лишь x . Таким образом, мы получаем следующую теорему.

Теорема. Управление $u = \phi^*(x)$ представляет собой оптимальное решение задачи на бесконечном промежутке времени (6.1.10)–(6.1.11), если существует непрерывно-дифференцируемая функция $W(x)$, определенная на $R^m \rightarrow R$ и удовлетворяющая следующему уравнению

$$\begin{aligned} rW(x) &= \max_u \{g[x, u] + W_x(x) f[x, u]\} \\ &= \{g[x, \phi^*(x)] + W_x(x) f[x, \phi^*(x)]\}. \end{aligned}$$

Подставляя оптимальное управление, полученное с использованием теоремы § 6.1 в (6.1.9), имеем уравнение для оптимальной траектории

$$\dot{x}(s) = f[x(s), \phi^*(x(s))], \quad x(t_0) = x_0.$$

Отсюда можно получить выражение для оптимальной траектории $\{x^*(t)\}_{t \geq t_0}$ в виде

$$x^*(t) = x_0 + \int_{t_0}^t f[x^*(s), \phi^*(x^*(s))] ds, \quad \text{при } t \geq t_0.$$

Как и раньше будем обозначать $x^*(t)$ через x_t^* . Оптимальное программное управление в задаче на бесконечном промежутке времени (6.1.8) представимо в виде функции $\phi^*(x_t^*)$, $t \in [t_0, \infty)$.

6.1.4. Пример 2. Рассмотрим задачу оптимального управления на бесконечном промежутке времени

$$\max_u \int_0^\infty \exp[-rs] [-x(s) - cu(s)^2] ds \quad (6.1.14)$$

при условии (6.1.6).

Используя теорему п. 6.1.3, получаем

$$rW(x) = \max_u \left\{ [-x - cu^2] + W_x(x) [a - ux^{1/2}] \right\}. \quad (6.1.15)$$

Найдем максимизирующую функцию в (6.1.15). Она будет иметь вид

$$\phi^*(x) = \frac{-W_x(x)x^{1/2}}{2c}.$$

Подставляя $\phi^*(x)$ в (6.1.15) и решая, получаем

$$W(x) = Ax + B,$$

где A и B удовлетворяют уравнениям

$$0 = rA - \frac{A^2}{4c} + 1, \quad B = \frac{a}{r}A.$$

Решая квадратное уравнение, найдем A .

$$A = 2c \left[r \pm (r^2 + c^{-1})^{1/2} \right].$$

Максимум достигается при отрицательном корне в выражении для A . Оптимальное управление получается в следующем виде

$$\phi^*(x) = \frac{-Ax^{1/2}}{2c}.$$

Подставляя $\phi^*(x) = -Ax^{1/2}/(2c)$ в (6.1.6), получаем формулу для оптимальной траектории

$$\dot{x}(s) = a + \frac{A}{2c}(x(s)), \quad x(0) = x_0.$$

Решая это уравнение, получаем оптимальную траекторию $\{x^*(t)\}_{t \geq t_0}$ в явном виде:

$$x^*(t) = \left[x_0 + \frac{2ac}{A} \right] \exp\left(\frac{A}{2c}t\right) - \frac{2ac}{A} = x_t^*, \quad t \geq t_0.$$

Тогда оптимальное управление для задачи (6.1.14)–(6.1.15) будет иметь вид

$$\phi^*(x_t^*) = \frac{-A(x_t^*)^{1/2}}{2c}.$$

Также представляется интересным рассмотрение задачи оптимального управления (6.1.1), (6.1.2) на промежутке $[t_0, T]$, где T — случайная величина с известной функцией распределения $F(t)$, $t \in [t_0; \infty)$ [Шевкопляс, 2009].

§ 6.2. Принцип максимума Понтрягина

6.2.1. Принцип максимума был предложен Л. С. Понтрягиным и его учениками [Понтрягин, Болтянский, Гамкрелидзе, Мищенко, 1976]. Рассмотрим вновь задачу оптимального управления (6.1.1)–(6.1.2)

Теорема (принцип максимума Понтрягина). *Для оптимальности управления $u^*(s) = \zeta^*(s, x_0)$ и соответствующей траектории $\{x^*(s), t_0 \leq s \leq T\}$ необходимо существование сопряженных функций $\Lambda(s) : [t_0, T] \rightarrow R^m$, удовлетворяющих условиям:*

$$\begin{aligned} \zeta^*(s, x_0) &\equiv u^*(s) = \arg \max_u \{g[s, x^*(s), u(s)] + \\ &\quad + \Lambda(s) f[s, x^*(s), u(s)]\}, \\ \dot{x}^*(s) &= f[s, x^*(s), u^*(s)], \quad x^*(t_0) = x_0, \\ \dot{\Lambda}(s) &= -\frac{\partial}{\partial x} \{g[s, x^*(s), u^*(s)] + \Lambda(s) f[s, x^*(s), u^*(s)]\}, \\ \Lambda(T) &= \frac{\partial}{\partial x^*} q(x^*(T)). \end{aligned}$$

Заметим, что принцип максимума Понтрягина дает необходимые условия оптимальности, однако если предположить существование дважды непрерывно дифференцируемой функции Беллмана, то эти условия оказываются и достаточными.

Доказательство. Доказательство проведем в предположении дважды непрерывной дифференцируемости функции Беллмана. Определим функцию (гамильтониан)

$$H(t, x, u) = g(t, x, u) + V_x(t, x) f(t, x, u).$$

Из теоремы п. 6.1.1 имеем

$$-V_t(t, x) = \max_u H(t, x, u).$$

Используя выражение для оптимального управления u^* , получаем

$$H(t, x, u^*) + V_t(t, x) \equiv 0,$$

что является тождеством по x . Возьмем частные производные от этого выражения по x . Это дает нам

$$V_{tx}(t, x) + g_x(t, x, u^*) + V_x(t, x) f_x(t, x, u^*) + V_{xx}(t, x) f(t, x, u^*) + [g_u(t, x, u^*) + V_x(t, x) f_u(t, x, u^*)] \frac{\partial u^*}{\partial x} = 0.$$

Если u^* является внутренней точкой ограничений на управления, тогда $[g_u(t, x, u^*) + V_x(t, x) f_u(t, x, u^*)] = 0$, что следует из условия $-V_t(t, x) = \max_u H(t, x, u)$. Если u^* не является внутренней точкой, то можно показать что

$$[g_u(t, x, u^*) + V_x(t, x) f_u(t, x, u^*)] \frac{\partial u^*}{\partial x} = 0,$$

поскольку из оптимальности следует ортогональность векторов $[g_u(t, x, u^*) + V_x(t, x) f_u(t, x, u^*)]$ и $\partial u^* / \partial x$ (в отдельных случаях может быть даже $\partial u^* / \partial x = 0$). Выражение

$$V_{tx}(t, x) + V_{xx}(t, x) f(t, x, u^*) \equiv V_{tx}(t, x) + V_{xx}(t, x) \dot{x}$$

может быть записано в виде $[dV_x(t, x)](dt)^{-1}$. Таким образом, мы получаем

$$\frac{dV_x(t, x)}{dt} + g_x(t, x, u^*) + V_x(t, x) f_x(t, x, u^*) = 0.$$

Вводя сопряженный вектор $\Lambda(t) = V_{x^*}(t, x^*)$, где x^* — траектория, соответствующая управлению u^* , получаем

$$\frac{dV_x(t, x^*)}{dt} = \dot{\Lambda}(s) = -\frac{\partial}{\partial x} \{g[s, x^*(s), u^*(s)] + \Lambda(s) f[s, x^*(s), u^*(s)]\}.$$

И, наконец, граничные условия для $\Lambda(t)$ определяются из граничного условия для задачи оптимального управления и теоремы п. 6.1.1

$$\Lambda(T) = \frac{\partial V(T, x^*)}{\partial x} = \frac{\partial q(x^*)}{\partial x}.$$

Таким образом, теорема доказана.

6.2.2. Пример 3. Рассмотрим задачу, содержащуюся в примере 1 п. 6.1.2. Используя теорему п. 6.2.1 найдем управление $u(s)$ из условия

$$\arg \max_u \left\{ \left[-x^*(s) - cu(s)^2 \right] \exp[-rs] + \Lambda(s) \left[a - u(s) x^*(s)^{1/2} \right] \right\}.$$

Проводя оптимизацию, получаем

$$u^*(s) = \frac{-\Lambda(s) x^*(s)^{1/2}}{2c} \exp[rs]. \quad (6.2.1)$$

Мы также получаем уравнение для сопряженного вектора (в данном случае одномерного)

$$\dot{\Lambda}(s) = \exp[-rs] + \frac{1}{2}\Lambda(s)u^*(s)x^*(s)^{-1/2}. \quad (6.2.2)$$

Подставляя выражение $u^*(s)$, полученное в (6.2.1), в (6.1.6) и (6.2.2), получаем пару дифференциальных уравнений

$$\begin{aligned} \dot{x}^*(s) &= a + \frac{1}{2c}\Lambda(s)(x^*(s))\exp[rs], \\ \dot{\Lambda}(s) &= \exp[-rs] - \frac{1}{4c}\Lambda(s)^2\exp[rs], \end{aligned} \quad (6.2.3)$$

с начальным и граничным условиями

$$x^*(0) = x_0, \quad \Lambda(T) = \exp[-rT]q.$$

Решая (6.2.3), получаем

$$\begin{aligned} \Lambda(s) &= 2c \left(\theta_1 - \theta_2 \frac{q - 2c\theta_1}{q - 2c\theta_2} \exp\left[\frac{\theta_1 - \theta_2}{2}(T - s)\right] \right) \exp(-rs) : \\ &: \left(1 - \frac{q - 2c\theta_1}{q - 2c\theta_2} \exp\left[\frac{\theta_1 - \theta_2}{2}(T - s)\right] \right), \\ x^*(s) &= \varpi(0, s) \left[x_0 + \int_0^s \varpi^{-1}(0, t) a dt \right], \quad s \in [0, T], \end{aligned}$$

где

$$\begin{aligned} \theta_1 &= r - \sqrt{r^2 + \frac{1}{c}}, \quad \theta_2 = r + \sqrt{r^2 + \frac{1}{c}}; \\ \varpi(0, s) &= \exp\left[\int_0^s H(\tau) d\tau\right], \\ H(\tau) &= \left(\theta_1 - \theta_2 \frac{q - 2c\theta_1}{q - 2c\theta_2} \exp\left[\frac{\theta_1 - \theta_2}{2}(T - \tau)\right] \right) : \\ &: \left(1 - \frac{q - 2c\theta_1}{q - 2c\theta_2} \exp\left[\frac{\theta_1 - \theta_2}{2}(T - \tau)\right] \right). \end{aligned}$$

Подставляя $\Lambda(s)$ и $x^*(s)$ в (6.2.1), получаем выражение для $u^*(s) = \zeta^*(s, x_0)$, которая оказывается функцией от s и x_0 .

Рассмотрим теперь задачу оптимального управления на бесконечном промежутке времени (6.1.8)–(6.1.9). Гамильтониан может быть записан в виде

$$H(t, x, u) = g(x, u) \exp[-r(t - t_0)] + \Lambda(t) f(x, u).$$

Введем функцию $\lambda(t) = \Lambda(t) \exp[r(t - t_0)]$. Тогда формула для текущего значения гамильтониана принимает вид

$$\begin{aligned} \hat{H}(t, x, u) &= H(t, x, u) \exp[r(t - t_0)] = \\ &= g(x, u) + \lambda(t) f(x, u). \end{aligned} \quad (6.2.4)$$

6.2.3. Используя (6.2.4), можно переформулировать теорему п. 6.2.1 для задачи (6.1.10)–(6.1.11).

Теорема. Управление $u^*(s) = \zeta^*(s, x_t)$ является оптимальным решением задачи управления на бесконечном промежутке времени (6.1.10)–(6.1.11) и $\{x^*(s), s \geq t\}$ — соответствующей оптимальной траекторией, если существуют такие сопряженные функции $\lambda(s) : [t, \infty) \rightarrow R^m$, что выполнены следующие условия:

$$\begin{aligned}\zeta^*(s, x_t) &\equiv u^*(s) = \arg \max_u \{g[x^*(s), u(s)] + \lambda(s) f[x^*(s), u(s)]\}, \\ \dot{x}^*(s) &= f[x^*(s), u^*(s)], \quad x^*(t) = x_t, \\ \dot{\lambda}(s) &= r\lambda(s) - \frac{\partial}{\partial x} \{g[x^*(s), u^*(s)] + \lambda(s) f[x^*(s), u^*(s)]\}.\end{aligned}$$

Пример 4. Рассмотрим задачу, предложенную в примере 2, на бесконечном промежутке времени. Используя теорему п. 6.1.3, получаем

$$\begin{aligned}\zeta^*(s, x_t) &\equiv u^*(s) = \\ &= \arg \max_u \left\{ [-x^*(s) - cu(s)]^2 + \lambda(s) [a - u(s)x^*(s)]^{1/2} \right\}, \\ \dot{x}^*(s) &= a - u^*(s)(x^*(s))^{1/2}, \quad x^*(t) = x_t, \\ \dot{\lambda}(s) &= r\lambda(s) + \left[1 + \frac{1}{2}\lambda(s)u^*(s)x^*(s)^{-1/2} \right].\end{aligned}\tag{6.2.5}$$

Проведя максимизацию, находим оптимальное управление

$$u^*(s) = \frac{-\lambda(s)x^*(s)^{1/2}}{2c}.$$

Подставляя $u^*(s)$ в (6.2.5), получаем

$$\begin{aligned}\dot{x}^*(s) &= a + \frac{\lambda(s)}{2c}u^*(s)x^*(s), \quad x^*(t) = x_t, \\ \dot{\lambda}(s) &= r\lambda(s) + \left[1 - \frac{1}{4c}\lambda(s)^2 \right].\end{aligned}\tag{6.2.6}$$

Решая (6.2.6) аналогично тому как мы делали в примере 3 п. 6.2.2, получаем явные выражения для $x^*(s)$ и $\lambda(s)$.

Подставляя эти выражения в $u^*(s)$, получаем оптимальное решение нашей задачи.

§ 6.3. Равновесие по Нэшу в программных стратегиях

6.3.1. Наиболее сложной и в тоже время плодотворной областью теории игр являются динамические или дифференциальные игры, в которых принятие решения конкурирующими сторонами происходит непрерывно во времени, и при этом делаются дополнительные предположения, касающиеся информационной структуры, неопределенности и возможности предварительных соглашений. Дифференциальные игры возникли в 50-х годах, когда Р. Айзекс сформулировал задачу перехвата самолета управляемой ракетой в терминах навигационных переменных (состояния и управления) и вывел основополагающие уравнения для решения задачи ([Isaacs, 1965]). Вклад Айзекса вместе с

классическим исследованием Беллмана в области динамического программирования и Понтрягина в области теории управления создали основы детерминистической теории дифференциальных игр. Первые работы в области дифференциальных игр были направлены на использование результатов теории оптимального управления для решения игровых задач. Л. Берковиц [Berkovitz, 1964] развил вариационный подход к дифференциальным играм. Г. Лейтман и Г. Мон [Leitmann, Mon, 1967] изучали геометрические аспекты дифференциальных игр. Л. Понтрягин [Pontryagin, 1966] предложил решать дифференциальные игры с использованием принципа максимума. Важнейшие результаты в области обоснования и методов нахождения решения антагонистических дифференциальных игр были предложены в [Красовский, 1970; Красовский, Субботин, 1974]. Особо следует отметить работы [Клейменов, 1990; Клейменов, 1993; Петров, 1970а; Петров, 1970b; Жуковский, Чикрий, 1994; Вайсборд, Жуковский, 1980; Тынянский, Жуковский, 1979].

Работы в области дифференциальных игр продолжают появляться в различных областях и разделах математики, в частности, в приложениях к экономике и менеджменту развиваются достаточно быстро, и подробный анализ указанных работ можно найти в работе [Dockner, Jorgensen, 2000].

6.3.2. В дифференциальных играх или играх с непрерывным временем изучается класс проблем принятия решений, в которых эволюция состояния описывается дифференциальным уравнением, и игроки действуют на заданном временном промежутке.

В частности, в общей дифференциальной игре n лиц игрок i стремится

$$\max_{u_i} \int_{t_0}^T g^i[s, x(s), u_1(s), u_2(s), \dots, u_n(s)] ds + q^i(x(T)) \quad (6.3.1)$$

для $i \in N = \{1, 2, \dots, n\}$,

при условии, что состояние системы развивается в соответствии с дифференциальным уравнением

$$\dot{x}(s) = f[s, x(s), u_1(s), u_2(s), \dots, u_n(s)], \quad x(t_0) = x_0, \quad (6.3.2)$$

где $x(s) \in X \subset R^m$ означает состояние системы в игре, и $u_i \in U^i$ — управление игрока i при $i \in N$. Функции $f[s, x, u_1, u_2, \dots, u_n]$, $g^i[s, \cdot, u_1, u_2, \dots, u_n]$ и $q^i(\cdot)$ при $i \in N$ и $s \in [t_0, T]$ предполагаются дифференцируемыми.

Определим понятие стратегии в дифференциальной игре. Вспомним как определялась стратегия u_i игрока i в динамической игре с полной информацией (см. определение п. 4.1.1). Там под стратегией понималось некоторое правило, которое каждой позиции игрока $x \in X_i$ из его множества очередности ставило в соответствие следующую позицию $y \in F_x$. Заметим, что реализация такой стратегии предполагает знание игроком позиции, в которой он находится при совершении очередного хода. Однако, поскольку граф игры является деревом, то знание позиции эквивалентно знанию всей предыстории игры до данной позиции. Стратегии в дифференциальной игре определяются аналогичным образом, при этом различаются два вида информационных состояний в игре: знание в каждый момент времени $s \in [t_0, T]$ начального состояния x_0 , момента s или знание в каждый момент времени $s \in [t_0, T]$ начального состояния игры x_0 , момента s и состояния игры в момент s , т. е. знание $x(s)$.

Под стратегией игрока $i \in N$ мы будем понимать функцию $u_i(s)$, ставящую в соответствие информации, имеющейся у игрока i в момент s , управление u_i из множе-

ства возможных управлений, т. е. в зависимости от состояния информации стратегии могут быть функциями времени $v_i(s) = u_i(x_0, s)$ и функциями состояния $v_i(s) = u_i(x_0, x(s), s)$. В первом случае мы будем говорить о программных стратегиях, во втором — о позиционных. К сожалению, такое определение стратегии математически не достаточно корректно (хотя им пользуются большинство авторов [Isaacs, 1965; Jorgensen, 1985; Haurie, 1976; Leitmann, 1974] и др.), поскольку не всегда при подстановке стратегий в правую часть уравнения движения (6.3.2) можно гарантировать существование, единственность и продолжимость решения на весь отрезок $[t_0, T]$. Именно поэтому имеются различные уточнения понятия стратегий, свободные от указанного недостатка. Наиболее удачными являются стратегии, введенные в работах [Красовский, Субботин, 1974; Красовский, 1985; Клейменов, 1993; Чистяков, 1992; Малафеев, 1982]. Однако их определение и использование требует глубоких знаний соответствующего математического аппарата, изложение которого в данной книге не представляется возможным. Для наших целей введенное нами определение стратегий вполне достаточно.

6.3.3. Определение. *Набор стратегий (ситуация) $\{v_1^*(s), v_2^*(s), \dots, v_n^*(s)\}$ образует равновесие по Нэшу в дифференциальной игре (6.3.1)–(6.3.2) n лиц, если следующие неравенства выполнены для всех $v_i(s) \in U^i$, $i \in N$*

$$\begin{aligned}
 & \int_{t_0}^T g^1[s, x^*(s), v_1^*(s), v_2^*(s), \dots, v_n^*(s)] ds + q^1(x^*(T)) \geq \\
 & \geq \int_{t_0}^T g^1[s, x^{[1]}(s), v_1(s), v_2^*(s), \dots, v_n^*(s)] ds + q^1(x^{[1]}(T)), \\
 & \int_{t_0}^T g^2[s, x^*(s), v_1^*(s), v_2^*(s), \dots, v_n^*(s)] ds + q^2(x^*(T)) \geq \\
 & \geq \int_{t_0}^T g^2[s, x^{[2]}(s), v_1^*(s), v_2(s), v_3^*(s), \dots, v_n^*(s)] ds + q^2(x^{[2]}(T)), \\
 & \qquad \qquad \qquad \vdots \qquad \qquad \qquad \vdots \\
 & \int_{t_0}^T g^n[s, x^*(s), v_1^*(s), v_2^*(s), \dots, v_n^*(s)] ds + q^n(x^*(T)) \geq \\
 & \geq \int_{t_0}^T g^n[s, x^{[n]}(s), v_1^*(s), \dots, v_{n-1}^*(s), v_n(s)] ds + q^n(x^{[n]}(T));
 \end{aligned}$$

и на отрезке времени $s \in [t_0, T]$:

$$\begin{aligned} \dot{x}^*(s) &= f[s, x^*(s), v_1^*(s), v_2^*(s), \dots, v_n^*(s)], \quad x^*(t_0) = x_0, \\ \dot{x}^{[1]}(s) &= f[s, x^{[1]}(s), v_1(s), v_2^*(s), \dots, v_n^*(s)], \quad x^{[1]}(t_0) = x_0, \\ \dot{x}^{[2]}(s) &= f[s, x^{[2]}(s), v_1^*(s), v_2(s), v_3^*(s), \dots, v_n^*(s)], \quad x^{[2]}(t_0) = x_0, \\ &\vdots \\ \dot{x}^{[n]}(s) &= f[s, x^{[n]}(s), v_1^*(s), \dots, v_{n-1}^*(s), v_n(s)], \quad x^{[n]}(t_0) = x_0. \end{aligned}$$

Если игроки строят свои стратегии, изначально не предполагая изменять управлений в будущем в зависимости от состояния, информационная структура может быть рассмотрена как программная. Эти стратегии являются функциями начального состояния x_0 и времени s и могут быть записаны в виде $\{u_i(s) = \vartheta_i(s, x_0), i \in N\}$. В частности, равновесие по Нэшу в программных стратегиях в игре (6.3.1) и (6.3.2) описывается следующей теоремой.

Теорема. Для того, чтобы набор стратегий $\{u_i^*(s) = \zeta_i^*(s, x_0), i \in N\}$ являлся равновесием по Нэшу в программных стратегиях в игре (6.3.1)–(6.3.2), и $\{x^*(s), t_0 \leq s \leq T\}$ была бы соответствующей оптимальной траекторией необходимо существование сопряженных функций $\Lambda^i(s) : [t_0, T] \rightarrow R^m$ при $i \in N$, удовлетворяющих условиям:

$$\begin{aligned} \zeta_i^*(s, x_0) &\equiv u_i^*(s) = \\ &= \arg \max_{u_i \in U^i} \{g^i[s, x^*(s), u_1^*(s), \dots, u_{i-1}^*(s), u_i(s), u_{i+1}^*(s), \dots, u_n^*(s)] + \\ &+ \Lambda^i(s) f[s, x^*(s), u_1^*(s), \dots, u_{i-1}^*(s), u_i(s), u_{i+1}^*(s), \dots, u_n^*(s)]\}, \\ \dot{x}^*(s) &= f[s, x^*(s), u_1^*(s), u_2^*(s), \dots, u_n^*(s)], \quad x^*(t_0) = x_0, \\ \dot{\Lambda}^i(s) &= -\frac{\partial}{\partial x^*} \{g^i[s, x^*(s), u_1^*(s), u_2^*(s), \dots, u_n^*(s)] + \\ &+ \Lambda^i(s) f[s, x^*(s), u_1^*(s), u_2^*(s), \dots, u_n^*(s)]\}, \\ \Lambda^i(T) &= \frac{\partial}{\partial x^*} q^i(x^*(T)). \end{aligned}$$

Доказательство. Из условий теоремы фактически следует, что программное управление $u_i^*(s) = \zeta_i^*(s, x_0)$ максимизирует выражение

$$\int_{t_0}^T g^i[s, x(s), u_1^*(s), \dots, u_{i-1}^*(s), u_i(s), u_{i+1}^*(s), \dots, u_n^*(s)] ds + q^i(x(T)),$$

на множестве всех программных управлений $u_i(s) \in U^i$ при условии

$$\begin{aligned} \dot{x}(s) &= f[s, x(s), u_1^*(s), \dots, u_{i-1}^*(s), u_i(s), u_{i+1}^*(s), \dots, u_n^*(s)], \\ x(t_0) &= x_0, \quad i \in N. \end{aligned}$$

Это стандартная задача оптимального управления для игрока i , поскольку $u_j^*(s)$ при $j \in N$ и $j \neq i$ — фиксированные программные управления, не зависящие от $u_i^*(s)$. Поэтому теорема прямо следует из принципа максимума Понтрягина, сформулированного в теореме п. 6.2.3.

§ 6.4. Равновесие по Нэшу в позиционных стратегиях

6.4.1. Предположим что информационная структура такова, что в каждый момент времени игрокам известны переменная состояния в этот момент времени, время и начальные условия. Стратегии игрока (позиционные стратегии) оказываются функциями начальных условий x_0 , текущего состояния $x(s)$ и текущего времени s и могут быть представлены в виде $\{u_i(s) = \vartheta_i(s, x, x_0) \text{ для } i \in N\}$. Следующая теорема содержит необходимые условия, которым должно удовлетворять равновесие по Нэшу в таком классе стратегий.

Теорема. Для того, чтобы набор стратегий $\{u_i(s) = \vartheta_i(s, x, x_0), \text{ для } i \in N\}$ образовывал равновесие по Нэшу в позиционных стратегиях в игре (6.3.1)–(6.3.2), где $\{x^*(s), t_0 \leq s \leq T\}$ — соответствующая оптимальная траектория, необходимо существование n сопряженных функций $\Lambda^i(s) : [t_0, T] \rightarrow R^m$ при $i \in N$, таких что выполнены следующие условия

$$\begin{aligned} \vartheta_i^*(s, x^*, x_0) &\equiv u_i^*(s) = \\ &= \arg \max_{u_i \in U^i} \{g^i[s, x^*(s), u_1^*(s), \dots, u_{i-1}^*(s), u_i(s), u_{i+1}^*(s), \dots, u_n^*(s)] + \\ &\quad + \Lambda^i(s) f[s, x^*(s), u_1^*(s), \dots, u_{i-1}^*(s), u_i(s), u_{i+1}^*(s), \dots, u_n^*(s)]\}, \\ \dot{x}^*(s) &= f[s, x^*(s), u_1^*(s), u_2^*(s), \dots, u_n^*(s)], \quad x^*(t_0) = x_0, \\ \dot{\Lambda}^i(s) &= -\frac{\partial}{\partial x^*} \{g^i[s, x^*(s), \vartheta_1^*(s, x^*, x_0), \vartheta_2^*(s, x^*, x_0), \dots \\ &\quad \dots, \vartheta_{i-1}^*(s, x^*, x_0), u_i^*(s), \vartheta_{i+1}^*(s, x^*, x_0), \dots, \vartheta_n^*(s, x^*, x_0)] + \\ &\quad + \Lambda^i(s) f[s, x^*(s), \vartheta_1^*(s, x^*, x_0), \vartheta_2^*(s, x^*, x_0), \dots \\ &\quad \dots, \vartheta_{i-1}^*(s, x^*, x_0), u_i^*(s), \vartheta_{i+1}^*(s, x^*, x_0), \dots, \vartheta_n^*(s, x^*, x_0)]\}, \\ \Lambda^i(T) &= \frac{\partial}{\partial x^*} q^i(x^*(T)). \end{aligned}$$

Доказательство. Рассмотрим i -ое равенство в условиях теоремы, в котором зафиксированы все стратегии игроков (кроме стратегии i -го игрока) и при $j \neq i$ и $j \in N$ $u_j^*(s) = \vartheta_j^*(s, x^*, x_0)$ — есть оптимальное управление игрока i . Тогда вышеприведенные условия следуют из принципа максимума Понтрягина, и игрок i максимизирует функционал

$$\int_{t_0}^T g^i[s, x(s), u_1^*(s), \dots, u_{i-1}^*(s), u_i(s), u_{i+1}^*(s), \dots, u_n^*(s)] ds + q^i(x(T))$$

на множестве выборов $u_i(s) \in U^i$ при условии

$$\begin{aligned} \dot{x}(s) &= f[s, x(s), u_1^*(s), \dots, u_{i-1}^*(s), u_i(s), u_{i+1}^*(s), \dots, u_n^*(s)], \\ x(t_0) &= x_0, \quad i \in N. \end{aligned}$$

Заметим, что частные производные по переменной x в сопряженных уравнениях теоремы п. 6.4.1 вычисляются при условии, что оставшиеся $n - 1$ стратегий игроков зависят от значений текущего состояния x . Это обстоятельство отсутствует в сопряженных уравнениях теоремы п. 6.3.3.

6.4.2. Рассмотрим теперь *состоятельное равновесие по Нэшу*. Заметим, что определенное нами позиционное равновесие и равновесие в программных стратегиях зависят от начальных условий задачи. Однако во многих случаях удастся построить равновесные решения, имеющие абсолютный характер, независящий от начального состояния процесса. Назовем такие равновесия *состоятельными позиционными равновесиями*.

В дифференциальной игре n лиц (6.1.1)–(6.1.2) набор стратегий $\{u_i^*(s) = \phi_i^*(s, x) \in U^i, i \in N\}$ образует состоятельное позиционное равновесие, если существуют функционалы (аналоги функции Беллмана) $V^i(t, x)$, определенные на $[t_0, T] \times R^m$ и удовлетворяющие следующим условиям

$$\begin{aligned} V^i(T, x) &= q^i(x), \\ V^i(t, x) &= \int_t^T g^i[s, x^*(s), \phi_1^*(s, x), \dots, \phi_n^*(s, x)] ds + q^i(x^*(T)) \geq \\ &\geq \int_t^T g^i[s, x^{[i]}(s), \phi_1^*(s, x), \dots, \phi_{i-1}^*(s, x), \phi_i(s, x), \phi_{i+1}^*(s, x), \dots, \phi_n^*(s, x)] ds + \\ &\quad + q^i(x^{[i]}(T)), \quad x \in R^m, \end{aligned}$$

при всех $\phi_i(s, x)$, $i \in N$. При этом на промежутке $[t_0, T]$ имеет место

$$\begin{aligned} \dot{x}^{[i]}(s) &= f[s, x^{[i]}(s), \phi_1^*(s, x), \dots, \phi_{i-1}^*(s, x), \\ &\quad \phi_i(s, x), \phi_{i+1}^*(s, x), \dots, \phi_n^*(s, x)], \quad x^{[1]}(t) = x; \\ \dot{x}^*(s) &= f[s, x^*(s), \phi_1^*(s, x), \dots, \phi_n^*(s, x)], \quad x(s) = x. \end{aligned}$$

Одно важное свойство введенного нами определения состоятельного равновесия по Нэшу заключается в том, что если набор $\{\phi_i^*; i \in N\}$ образует состоятельное равновесие по Нэшу в дифференциальной игре на отрезке $[t_0, T]$, то его сужение на промежутке $[t, T]$ образует состоятельное позиционное равновесие по Нэшу в той же самой дифференциальной игре, но определенной на более коротком промежутке $[t, T]$ из начального состояния $x(t)$. Это имеет место для всех $t_0 \leq t \leq T$. Отсюда сразу же следует, что стратегия, образующая состоятельное позиционное равновесие, зависит лишь от времени и текущего состояния, но не зависит от предыстории (включая начальное состояние x_0). Поэтому стратегии игроков могут быть записаны в виде $\{u_i(s) = \phi_i(s, x)$, для $i \in N\}$. Следующая теорема определяет необходимое условие существования состоятельных позиционных равновесий для игры (6.3.1)–(6.3.2).

Теорема. Набор стратегий $\{u_i^*(s) = \phi_i^*(t, x) \in U^i, i \in N\}$ является состоятельным позиционным равновесием по Нэшу в игре (6.3.1)–(6.3.2), если существуют непрерывно-дифференцируемые функции $V^i(t, x) : [t_0, T] \times R^m \rightarrow R$, $i \in N$, удовлетворяющие следующей системе дифференциальных уравнений в частных производных:

$$\begin{aligned} &-V_t^i(t, x) = \\ &= \max_{u_i} \{g^i[t, x, \phi_1^*(t, x), \phi_2^*(t, x), \dots, \phi_{i-1}^*(t, x), u_i(t, x), \phi_{i+1}^*(t, x), \dots, \phi_n^*(t, x)] + \end{aligned}$$

$$\begin{aligned}
& + V_x^i(t, x) f[t, x, \phi_1^*(t, x), \phi_2^*(t, x), \dots, \phi_{i-1}^*(t, x), u_i(t, x), \phi_{i+1}^*(t, x), \dots, \phi_n^*(t, x)] \} = \\
& = \{g^i[t, x, \phi_1^*(t, x), \phi_2^*(t, x), \dots, \phi_n^*(t, x)] + V_x^i(t, x) f[t, x, \phi_1^*(t, x), \phi_2^*(t, x), \dots, \phi_n^*(t, x)]\}, \\
& V^i(T, x) = q^i(x), \quad i \in N.
\end{aligned}$$

Доказательство. По теореме п. 6.3.3 $V^i(t, x)$ являются функциями значения (функциями Беллмана) в задаче оптимального управления для игрока $i \in N$. Из определения п. 6.3.3 следует, что условие теоремы п. 6.4.2 являются условиями равновесия по Нэшу.

6.4.3. Рассмотрим теперь антагонистический вариант игры (6.3.1)–(6.3.2), в которой имеется всего два игрока и выигрыш игрока 1 равен выигрышу игрока 2 с противоположным знаком. В этом случае состоятельное позиционное равновесие характеризуется следующей теоремой.

Теорема. Пара стратегий $\{\phi_i^*(t, x); i = 1, 2\}$ образует ситуацию равновесия (седловую точку) в антагонистическом варианте игры (6.3.1)–(6.3.2), если существует непрерывно-дифференцируемая функция $V : [t_0, T] \times R^m \rightarrow R$, удовлетворяющая следующему уравнению в частных производных

$$\begin{aligned}
-V_t(t, x) &= \min_{u_1} \max_{u_2} \{g[t, x, u_1(t), u_2(t)] + V_x f[t, x, u_1(t), u_2(t)]\} = \\
&= \max_{u_2} \min_{u_1} \{g[t, x, u_1(t), u_2(t)] + V_x f[t, x, u_1(t), u_2(t)]\} = \\
&= \{g[t, x, \phi_1^*(t, x), \phi_2^*(t, x)] + V_x f[t, x, \phi_1^*(t, x), \phi_2^*(t, x)]\}, \\
V(T, x) &= q(x).
\end{aligned}$$

Доказательство. Этот результат является частным случаем теоремы п. 6.4.2, если взять $n = 2$, $g^1(\cdot) = -g^2(\cdot) \equiv g(\cdot)$ и $q^1(\cdot) = -q^2(\cdot) \equiv q(\cdot)$. В этом случае $V^1 = -V^2 \equiv V$, и существование седловой точки (ситуации равновесия) эквивалентно взаимозаменяемости операций минимума и максимума.

§ 6.5. Конкурентная реклама с двумя участниками

В этом параграфе мы рассмотрим приложение теории дифференциальных игр к конкуренции в области рекламы.

Рассмотрим динамическую игру конкурентной рекламы, впервые предложенную Г. Зоргером [Sorger, 1989]. На рынок выходят две фирмы, при этом функции дохода фирмы 1 и фирмы 2 имеют соответственно вид

$$\begin{aligned}
& \int_0^T \left[q_1 x(s) - \frac{c_1}{2} u_1(s)^2 \right] \exp(-rs) ds + \exp(-rT) S_1 x(T), \\
& \int_0^T \left[q_2 (1 - x(s)) - \frac{c_2}{2} u_2(s)^2 \right] \exp(-rs) ds + \exp(-rT) S_2 [1 - x(T)],
\end{aligned} \tag{6.5.1}$$

где r, q_i, c_i, S_i при $i \in \{1, 2\}$ — положительные постоянные, $x(s)$ — доля рынка фирмы 1 в момент s , $[1 - x(s)]$ — доля рынка фирмы 2, $u_i(s)$ — вложения в рекламу фирмы $i \in \{1, 2\}$.

Предполагается, что емкость рынка не изменяется со временем. Единственным рыночным инструментом, которым пользуются фирмы, является реклама. Реклама влияет на динамику изменения доли рынка каждой из фирм. Динамика изменения доли

рынка первой фирмы определяется следующим дифференциальным уравнением

$$\dot{x}(s) = u_1(s)[1 - x(s)]^{1/2} - u_2(s)x(s)^{1/2}, \quad x(0) = x_0. \quad (6.5.2)$$

Предположим, что в качестве принципа оптимальности фирмы выбрали равновесие по Нэшу в программных стратегиях. Использование программных стратегий требует, чтобы фирмы определили свои управления в начальный момент времени. Такое предположение реалистично лишь в том случае, если есть четкие соглашения, касающиеся рекламы. Согласно теореме п. 6.3.3 решение в программных стратегиях игры (6.5.1)–(6.5.2) должно удовлетворять следующим условиям:

$$\begin{aligned} u_1^*(s) &= \arg \max_{u_1} \left\{ \left[q_1 x^*(s) - \frac{c_1}{2} u_1(s)^2 \right] \exp(-rs) + \right. \\ &\quad \left. + \Lambda^1(s) \left(u_1(s)[1 - x^*(s)]^{1/2} - u_2(s)x^*(s)^{1/2} \right) \right\}, \\ u_2^*(s) &= \arg \max_{u_2} \left\{ \left[q_2(1 - x^*(s)) - \frac{c_2}{2} u_2(s)^2 \right] \exp(-rs) + \right. \\ &\quad \left. + \Lambda^2(s) \left(u_1(s)[1 - x^*(s)]^{1/2} - u_2(s)x^*(s)^{1/2} \right) \right\}, \\ \dot{x}^*(s) &= u_1^*(s)[1 - x^*(s)]^{1/2} - u_2^*(s)x^*(s)^{1/2}, \quad x^*(0) = x_0, \\ \dot{\Lambda}^1(s) &= -q_1 \exp(-rs) + \Lambda^1(s) \left(\frac{1}{2} u_1^*(s)[1 - x^*(s)]^{-1/2} + \frac{1}{2} u_2^*(s)x^*(s)^{-1/2} \right), \\ \dot{\Lambda}^2(s) &= q_2 \exp(-rs) + \Lambda^2(s) \left(\frac{1}{2} u_1^*(s)[1 - x^*(s)]^{-1/2} + \frac{1}{2} u_2^*(s)x^*(s)^{-1/2} \right), \\ \Lambda^1(T) &= \exp(-rT) S_1, \\ \Lambda^2(T) &= -\exp(-rT) S_2. \end{aligned} \quad (6.5.3)$$

Из (6.5.3) мы получаем

$$\begin{aligned} u_1^*(s) &= \frac{\Lambda^1(s)}{c_1} [1 - x^*(s)]^{1/2} \exp(rs), \\ u_2^*(s) &= \frac{\Lambda^2(s)}{c_2} [x^*(s)]^{1/2} \exp(rs). \end{aligned}$$

Подставляя $u_1^*(s)$ и $u_2^*(s)$ в (6.5.3), получаем

$$\begin{aligned} \dot{\Lambda}^1(s) &= \left\{ -q_1 \exp(-rs) + \left(\frac{[\Lambda^1(s)]^2}{2c_1} + \frac{\Lambda^1(s)\Lambda^2(s)}{2c_2} \right) \right\}, \\ \dot{\Lambda}^2(s) &= \left\{ q_2 \exp(-rs) + \left(\frac{[\Lambda^2(s)]^2}{2c_2} + \frac{\Lambda^1(s)\Lambda^2(s)}{2c_1} \right) \right\}, \end{aligned} \quad (6.5.4)$$

с граничными условиями

$$\Lambda^1(T) = \exp(-rT) S_1, \quad \Lambda^2(T) = -\exp(-rT) S_2.$$

Тогда уравнение движения в ситуации равновесия будет иметь вид

$$\begin{aligned}\dot{x}^*(s) &= \frac{\Lambda^1(s) \exp(rs)}{c_1} [1 - x^*(s)] - \frac{\Lambda^2(s) \exp(rs)}{c_2} x^*(s), \\ x^*(0) &= x_0.\end{aligned}\quad (6.5.5)$$

Решая систему дифференциальных уравнений (6.5.4)–(6.5.5), мы получаем выражения для $x^*(s)$, $\Lambda^1(s)$ и $\Lambda^2(s)$.

Подставляя эти выражения в $u_1^*(s)$ и $u_2^*(s)$, получаем программные равновесные стратегии.

Попробуем теперь найти позиционно-состоятельное решение в задаче конкурентной рекламы. Позиционное решение позволяет игрокам строить свои уровни рекламы в зависимости от состояния игры (состояния рынка), что является более реалистичным подходом к решению задачи (6.5.1)–(6.5.2). Используя теорему § 6.4 получаем, что равновесное решение игры (6.5.1)–(6.5.2) должно удовлетворять следующим условиям

$$\begin{aligned}-V_t^1(t, x) &= \max_{u_1} \left\{ \left[q_1 x - \frac{c_1}{2} u_1^2 \right] \exp(-rt) + \right. \\ &\quad \left. + V_x^1(t, x) \left(u_1 [1 - x]^{1/2} - \phi_2^*(t, x) x^{1/2} \right) \right\}, \\ -V_t^2(t, x) &= \max_{u_2} \left\{ \left[q_2 (1 - x) - \frac{c_2}{2} u_2^2 \right] \exp(-rt) + \right. \\ &\quad \left. + V_x^2(t, x) \left(\phi_1^*(t, x) [1 - x]^{1/2} - u_2 x^{1/2} \right) \right\}, \\ V^1(T, x) &= \exp(-rT) S_1 x, \\ V^2(T, x) &= \exp(-rT) S_2 (1 - x).\end{aligned}\quad (6.5.6)$$

Нахождение максимизирующих стратегий в (6.5.6) дает нам

$$\begin{aligned}\phi_1^*(t, x) &= \frac{V_x^1(t, x)}{c_1} [1 - x]^{1/2} \exp(rt), \\ \phi_2^*(t, x) &= \frac{V_x^2(t, x)}{c_2} [x]^{1/2} \exp(rt).\end{aligned}$$

Подставляя $\phi_1^*(t, x)$ и $\phi_2^*(t, x)$ в (6.5.6) и решая уравнение, получаем следующие функции значения (функции Беллмана)

$$\begin{aligned}V^1(t, x) &= \exp[-r(t)] [A_1(t) x + B_1(t)], \\ V^2(t, x) &= \exp[-r(t)] [A_2(t) (1 - x) + B_2(t)],\end{aligned}\quad (6.5.7)$$

где $A_1(t)$, $B_1(t)$, $A_2(t)$ и $B_2(t)$ удовлетворяют уравнениям

$$\begin{aligned}\dot{A}_1(t) &= r A_1(t) - q_1 + \frac{A_1(t)^2}{2c_1} + \frac{A_1(t) A_2(t)}{2c_2}, \\ \dot{A}_2(t) &= r A_2(t) - q_2 + \frac{A_2(t)^2}{2c_2} + \frac{A_1(t) A_2(t)}{2c_1}, \\ A_1(T) &= S_1, \quad B_1(T) = 0, \quad A_2(T) = S_2; \quad B_2(T) = 0.\end{aligned}$$

Подставляя соответствующие частные производные функций $V^1(t, x)$ и $V^2(t, x)$ из (6.5.7) в (6.5.6), получаем состоятельное позиционное равновесие по Нэшу

$$\phi_1^*(t, x) = \frac{A_1(t)}{c_1} [1 - x]^{1/2}, \quad \phi_2^*(t, x) = \frac{A_2(t)}{c_2} [x]^{1/2}. \quad (6.5.8)$$

§ 6.6. Игры с бесконечной продолжительностью

6.6.1. Рассмотрим дифференциальную игру с бесконечной продолжительностью и с постоянным параметром дисконтирования. Будем считать, что игрок $i \in N$ стремится

$$\max_{u_i} \int_{t_0}^{\infty} g^i[x(s), u_1(s), u_2(s), \dots, u_n(s)] \exp[-r(s - t_0)] ds, \quad (6.6.1)$$

при ограничении

$$\dot{x}(s) = f[x(s), u_1(s), u_2(s), \dots, u_n(s)], \quad x(t_0) = x_0, \quad (6.6.2)$$

где r — постоянный параметр дисконтирования.

Рассмотрим теперь подыгры игры (6.6.1)–(6.6.2), начинающиеся в момент времени t

$$\max_{u_i} \int_t^{\infty} g^i[x(s), u_1(s), u_2(s), \dots, u_n(s)] \exp[-r(s - t)] ds, \quad i \in N, \quad (6.6.3)$$

при условии

$$\dot{x}(s) = f[x(s), u_1(s), u_2(s), \dots, u_n(s)], \quad s \geq t, \quad x(t) = x. \quad (6.6.4)$$

Игра (6.6.3)–(6.6.4), заданная на бесконечном промежутке времени, не зависит от выбора начального момента t , а зависит лишь от состояния x в момент начала игры.

Как мы ранее показали, управление в задаче на бесконечном промежутке времени в случае автономной системы является функцией только состояния.

Определение. В дифференциальной игре n лиц (6.6.1)–(6.6.2) набор стратегий

$$\{u_i^*(s) = \phi_i^*(\eta_s) \in U^i, \text{ для } i \in N\}$$

образует состоятельное позиционное равновесие по Нэшу, если существуют функции $V^i(t, x)$, определенные на $[t_0, \infty) \times R^m$ и удовлетворяющие следующим соотношениям

$$\begin{aligned} V^i(t, x) &= \int_t^{\infty} g^i[x^*(s), \phi_1^*(\eta_s), \dots, \phi_n^*(\eta_s)] \exp[-r(s - t_0)] ds \geq \\ &\geq \int_t^{\infty} g^i[x(s), \phi_1^*(\eta_s), \dots, \phi_{i-1}^*(\eta_s) \phi_i(\eta_s) \phi_{i+1}^*(\eta_s), \dots, \phi_n^*(\eta_s)] \exp[-r(s - t_0)] ds, \end{aligned}$$

$$\forall \phi_i(\cdot, \cdot) \in \Gamma^i, \quad x \in R^n,$$

для всех $\varphi_i(\eta_s)$, $i \in N$, при этом на интервале $[t_0, \infty)$ имеет место

$$\begin{aligned} \dot{x}(s) &= f[s(s), \phi_1^*(\eta_s), \dots, \phi_{i-1}^*(\eta_s) \phi_i(\eta_s) \phi_{i+1}^*(\eta_s), \dots, \phi_n^*(\eta_s)], \\ x^{[1]}(t) &= x; \end{aligned}$$

$$\dot{x}^*(s) = f[x^*(s), \phi_1^*(\eta_s), \phi_2^*(\eta_s), \dots, \phi_n^*(\eta_s)], \quad x^*(t_0) = x_0.$$

Здесь символ η_s означает либо информационное состояние $\{s, x_0\}$, либо информационное состояние $\{x(s), x_0\}$ в зависимости от того, какая информационная структура рассматривается.

Мы можем записать

$$V^i(t, x) = \exp[-r(t - t_0)] \int_t^\infty g^i[x^*(s), \phi_1^*(\eta_s), \dots, \phi_n^*(\eta_s)] \exp[-r(s - t)] ds,$$

где

$$x(t) = x = x_t^* = x^*(t).$$

Поскольку

$$\int_t^\infty g^i[x^*(s), \phi_1^*(\eta_s), \phi_2^*(\eta_s), \dots, \phi_n^*(\eta_s)] \exp[-r(s - t)] ds$$

зависит лишь от текущего состояния x , можно написать

$$W^i(x) = \int_t^\infty g^i[x^*(s), \phi_1^*(\eta_s), \dots, \phi_n^*(\eta_s)] \exp[-r(s - t)] ds.$$

Отсюда получаем для $i \in N$

$$\begin{aligned} V^i(t, x) &= \exp[-r(t - t_0)] W^i(x), \\ V_t^i(t, x) &= -r \exp[-r(t - t_0)] W^i(x), \\ V_x^i(t, x) &= \exp[-r(t - t_0)] W_x^i(x). \end{aligned} \quad (6.6.5)$$

Состоятельное позиционное равновесие по Нэшу для автономной игры на бесконечном промежутке времени (6.6.3)–(6.6.4) может быть охарактеризовано следующим образом.

6.6.2. Теорема. *Набор стратегий $\{u_i^*(s) = \phi_i^*(\cdot) \in U^i; \text{ для } i \in N\}$ является состоятельным позиционным равновесием по Нэшу в игре на бесконечном интервале времени (6.6.3)–(6.6.4), если существуют непрерывно-дифференцируемые функции $W^i(x) : R^m \rightarrow R, i \in N$, удовлетворяющие следующей системе дифференциальных уравнений в частных производных*

$$\begin{aligned} rW^i(x) &= \\ &= \max_{u_i} \{g^i[x, \phi_1^*(x), \phi_2^*(x), \dots, \phi_{i-1}^*(x), u_i, \phi_{i+1}^*(x), \dots, \phi_n^*(x)] + \\ &\quad + W_x^i(x) f[x, \phi_1^*(x), \dots, \phi_{i-1}^*(x), u_i, \phi_{i+1}^*(x), \dots, \phi_n^*(x)]\} = \\ &= \{g^i[x, \phi_1^*(x), \phi_2^*(x), \dots, \phi_n^*(x)] + \\ &\quad + W_x^i(x) f[x, \phi_1^*(x), \phi_2^*(x), \dots, \phi_n^*(x)]\}, \quad \text{для } i \in N. \end{aligned}$$

Доказательство. По теореме п. 6.1.3 функция $W^i(x)$ является функцией значения соответствующей задачи оптимизации игрока $i \in N$. Вместе с неравенством в определении п. 6.3.3 условия теоремы п. 6.6.2 дают равновесие по Нэшу.

Поскольку s явно не входит в уравнение в частных производных, из теоремы п. 6.3.3 мы получаем независимость состоятельного позиционного равновесия по Нэшу от времени $\{u_i^* = \phi_i^*(x), i \in N\}$. Подставляя равновесные стратегии из теоремы п. 6.3.3 в (6.6.2), получаем равновесную траекторию

$$\dot{x}(s) = f[x(s), \phi_1^*(x(s)), \phi_2^*(x(s)), \dots, \phi_n^*(x(s))], \quad x(t_0) = x_0.$$

Решая дифференциальные уравнения, получаем оптимальную траекторию $\{x^*(t)\}_{t \geq t_0}$ в виде

$$x^*(t) = x_0 + \int_{t_0}^t f[x^*(s), \phi_1^*(x^*(s)), \dots, \phi_n^*(x^*(s))] ds, \quad t \geq t_0.$$

Как и ранее обозначаем выражение $x^*(t)$ через x_t^* . Тогда состоятельное позиционное равновесие для игры на бесконечном отрезке времени (6.6.1)–(6.6.2) может быть получено в виде

$$[\phi_1^*(x_t^*), \phi_2^*(x_t^*), \dots, \phi_n^*(x_t^*)], \quad t \geq t_0.$$

6.6.3. Следуя теореме п. 6.2.1 и п. 6.3.3, охарактеризуем равновесие в программных стратегиях для игры с бесконечной временно продолжительностью (6.6.3) и (6.6.4).

Теорема. Для того, чтобы набор стратегий $\{u_i^*(s) = \zeta_i^*(s, x_t), i \in N\}$ образовывал равновесие по Нэшу в программных стратегиях для игры на бесконечном промежутке времени (6.6.3)–(6.6.4), где $\{x^*(s), t \leq s \leq T\}$ — соответствующая оптимальная траектория, необходимо существование n таких сопряженных функций $\lambda^i(s) : [t, T] \rightarrow R^m$ при $i \in N$, что выполнены следующие условия

$$\begin{aligned} \zeta_i^*(s, x) &\equiv u_i^*(s) = \\ &= \arg \max_{u_i \in U^i} \{g^i[x^*(s), u_1^*(s), \dots, u_{i-1}^*(s), u_i(s), u_{i+1}^*(s), \dots, u_n^*(s)] + \\ &+ \lambda^i(s) f[x^*(s), u_1^*(s), \dots, u_{i-1}^*(s), u_i(s), u_{i+1}^*(s), \dots, u_n^*(s)]\}, \\ x^*(s) &= f[x^*(s), u_1^*(s), u_2^*(s), \dots, u_n^*(s)], \quad x^*(t) = x_t, \\ \dot{\lambda}^i(s) &= r\lambda(s) - \frac{\partial}{\partial x^*} \{g^i[x^*(s), u_1^*(s), u_2^*(s), \dots, \\ &\dots, u_n^*(s)] + \lambda^i(s) f[x^*(s), u_1^*(s), u_2^*(s), \dots, u_n^*(s)]\}. \end{aligned}$$

Доказательство. Рассмотрим i -ое равенство в условиях теоремы, из которого следует, что $u_i^*(s) = \zeta_i^*(s, x_t)$ максимизирует

$$\int_{t_0}^{\infty} g^i[x(s), u_1^*(s), \dots, u_{i-1}^*(s), u_i(s), u_{i+1}^*(s), \dots, u_n^*(s)] ds,$$

на множестве $u_i(s) \in U^i$ при ограничении

$$\begin{aligned} \dot{x}(s) &= f[x(s), u_1^*(s), \dots, u_{i-1}^*(s), u_i(s), u_{i+1}^*(s), \dots, u_n^*(s)], \\ x(t) &= x_t, \quad i \in N. \end{aligned}$$

Это есть задача оптимального управления для игрока i на бесконечном промежутке времени. Поскольку $u_j^*(s)$, при $j \in N$ и $j \neq i$, — программные управления и, следовательно, не зависят от $u_i^*(s)$, эти результаты следуют из теоремы п. 6.2.1.

§ 6.7. Модель конкуренции с бесконечной продолжительностью

Рассмотрим динамическую дуополию, в которой две фирмы продают однородный продукт. Будем следовать [Tsutsui, Mino, 1990] и предположим что

$$\dot{P}(s) = k[a - u_1(s) - u_2(s) - P(s)], \quad P(t_0) = P_0, \quad (6.7.1)$$

где $P(s)$ — рыночная цена в момент s , $u_i(s)$ — объем выпуска продукции фирмой $i \in \{1, 2\}$, и текущий спрос определяется по формуле $P(s) = [a - u_1(s) - u_2(s)]$. Коэффициент $k > 0$ характеризует скорость изменения цены.

Выигрыш фирмы i определяется как суммарный доход, дисконтированный на момент начала процесса

$$\int_{t_0}^{\infty} \left\{ P(s) u_i(s) - c u_i(s) - (1/2) [u_i(s)]^2 \right\} \exp[-r(s - t_0)] ds, \\ i \in \{1, 2\}, \quad (6.7.2)$$

где $c u_i(s) + (1/2) [u_i(s)]^2$ есть затраты на производство продукции в объеме $u_i(s)$, а r — процентная ставка.

Рассмотрим подыгры

$$\max_{u_i} \int_t^{\infty} \left\{ P(s) u_i(s) - c u_i(s) - (1/2) [u_i(s)]^2 \right\} \exp[-r(s - t)] ds, \\ i \in \{1, 2\}, \quad (6.7.3)$$

при условии

$$\dot{P}(s) = k[a - u_1(s) - u_2(s) - P(s)], \quad P(t) = P. \quad (6.7.4)$$

Игра на бесконечном промежутке времени (6.7.3)–(6.7.4) является автономной с постоянной процентной ставкой, поэтому можно применить теорему п. 6.6.3 для характеристики состоятельного позиционного равновесия по Нэшу. Поэтому можем написать следующие уравнения

$$r W^i(P) = \max_{u_i} \left\{ \left[P u_i - c u_i - (1/2) (u_i)^2 \right] + \right. \\ \left. + W_P^i [k(a - u_i - \phi_j^*(P) - P)] \right\}, \quad i \in \{1, 2\}. \quad (6.7.5)$$

Проводя максимизацию в (6.7.5), получаем

$$\phi_i^*(P) = P - c - k W_P^i(P), \quad i \in \{1, 2\}. \quad (6.7.6)$$

Поставляя (6.7.6) в (6.7.5) и решая уравнение, получаем

$$W^i(P) = \frac{1}{2} A P^2 - B P + C, \quad (6.7.7)$$

где

$$A = \frac{r + 6k - \sqrt{(r + 6k)^2 - 12k^2}}{6k^2}, \\ B = \frac{-akA + c - 2kcA}{r - 3k^2A + 3k}, \\ C = \frac{c^2 + 3k^2B^2 - 2kB(2c + a)}{2r}.$$

Далее мы можем еще раз убедиться в том, что $W^i(P)$, определяемая (6.7.7), действительно решает (6.7.5), подставляя $W^i(P)$ и ее производные в (6.7.5) и (6.7.6).

Равновесные стратегии могут быть записаны в виде:

$$\phi_i^*(P) = P - c - k(AP - B), \quad i \in \{1, 2\}.$$

Подстановка равновесных стратегий в (6.7.1) дает возможность записать равновесную траекторию игры (6.7.1)–(6.7.2) в виде

$$\dot{P}(s) = k[a - 2(c + kB) - (3 - kA)P(s)], \quad P(t_0) = P_0.$$

Решение этого дифференциального уравнения дает нам выражение для оптимальной траектории

$$P^*(t) = \left[P_0 - \frac{k[a - 2(c + kB)]}{k(3 - kA)} \right] \exp[-k(3 - kA)t] + \frac{k[a - 2(c + kB)]}{k(3 - kA)}.$$

Обозначим для простоты $P^*(t)$ через P_t^* . Состоятельное позиционное равновесие для игры (6.7.1)–(6.7.2) может быть записано в виде

$$\phi_i^*(P_t^*) = P_t^* - c - k(AP_t^* - B), \quad i \in \{1, 2\}.$$

Нахождение программного равновесия оставим читателю.

§ 6.8. Упражнения и задачи

1. Рассмотрим следующую задачу оптимального управления:

$$\begin{aligned} \max_u \int_{t_0}^T \left[u(s)^{1/2} - \frac{c}{x(s)^{1/2}} u(s) \right] \exp[-r(s - t_0)] ds \\ + \exp[-r(T - t_0)] qx(T)^{1/2}, \end{aligned}$$

при условии

$$\dot{x}(s) = [ax(s)^{1/2} - bx(s) - u(s)], \quad x(t_0) = x_0 \in X.$$

Найти оптимальное управление, используя принцип динамического программирования Беллмана.

2. Используя постановку задачи из упр.1, выполнить следующие задания.

- а) Пусть $c = 1$, $q = 2$, $r = 0.01$, $t_0 = 0$, $T = 5$, $a = 0.5$, $b = 1$ и $x_0 = 20$. Найти оптимальное управление, оптимальную траекторию и выписать выражение для сопряженной функции $\Lambda(s)$.
- б) Пусть $c = 1$, $q = 2$, $r = 0.01$, $t_0 = 0$, $T = 5$, $a = 0.5$, $b = 1$ и $x_0 = 30$. Найти оптимальное управление, оптимальную траекторию и выписать выражение для сопряженной функции $\Lambda(s)$.

3. Рассмотрим игру, в которой игрок i максимизирует выигрыш

$$\int_{t_0}^{\infty} \{P(s)u_i(s) - cu_i(s) - (1/2)[u_i(s)]^2\} \exp[-r(s - t_0)] ds, \quad \text{где } i \in \{1, 2\},$$

при условии

$$\dot{P}(s) = k[a - u_1(s) - u_2(s) - P(s)], \quad P(t_0) = P_0.$$

Найти равновесие по Нэшу в программных стратегиях.

4. Рассмотрим игру

$$\max_{u_i} \left\{ \int_0^{10} \left[10u_i(s) - \frac{u_i(s)^2}{x(s)} \right] \exp[-0.05s] ds + \exp(-0.5) 2x(T) \right\},$$

где $i \in \{1, 2, \dots, 6\}$

при условии

$$\dot{x}(s) = 15 - \frac{1}{2}x(s) - \sum_{j=1}^6 u_j(s), \quad x(0) = 25.$$

а) Найти равновесие по Нэшу в программных стратегиях.

б) Найти равновесие по Нэшу в позиционных стратегиях.

5. Найти равновесие по Нэшу в программных стратегиях для модели конкуренции с бесконечной продолжительностью (см. § 6.5).

6. Рассмотрим модель конкурентной рекламы с двумя участниками (см. § 6.7). Будем предполагать, что задача рассматривается на бесконечном временном промежутке.

Найти равновесие по Нэшу в позиционных стратегиях.

7. Рассматривается теоретико-игровая модель сокращения вредных выбросов в атмосферу [Jorgensen, Martin-Herran, Zaccour, 2003]. Пусть $s(t)$ — поток загрязнений, $e_i(t)$ — выбросы i -й страны, $i = 1, \dots, n$. Динамика игры задана уравнением:

$$\dot{s}(t) = \sum_{i=1}^n \beta_i e_i(t) - \delta s, \quad s(0) = s_0 \geq 0,$$

где β_i — неотрицательный параметр.

Доход $R_i(e_i)$ определяется по формуле:

$$R_i(e_i) = \gamma_i \log(\alpha_i e_i), \quad \gamma_i > 0, \alpha_i > 0.$$

Функция затрат $D_i(s)$ линейна и определяется по формуле:

$$D_i(s) = \phi_i s, \quad \phi_i > 0.$$

Выигрыш игрока i имеет вид

$$K_i(s_0, e_1, \dots, e_n) = \int_{t_0}^{\infty} (R_i(e_i) - D_i(s)) e^{-\rho t} dt.$$

Найти равновесие по Нэшу в позиционных стратегиях.

Указание. Функцию Беллмана искать в виде $V_i = A_i s + B_i$. Обратить внимание на симметричность игроков.

8. Вывести уравнение типа Беллмана для задачи оптимального управления (6.1.1), (6.1.2) при $[t_0, T]$, где T — случайная величина с известной функцией распределения $F(t), t \in [t_0; \infty)$ [Шевкопляс, 2009]. В качестве интегрального функционала взять математическое ожидание функционала (6.1.1).

9. Используя решение упражнения 8, найти оптимальные управления в примере 1, предполагая, что задача рассматривается при $[t_0, T]$, где T — случайная величина с известной функцией распределения $F(t), t \in [t_0; \infty)$.

Найти решение при предположении, что случайная величина T распределена по закону Вейбулла ($F(t) = 1 - e^{-\lambda t^\gamma}, t \geq 0$).

10. Рассмотрим модель конкурентной рекламы с двумя участниками (см. § 6.5). Будем предполагать, что рассматривается игра со случайной продолжительностью (см. упражнение 8).

Найти равновесие по Нэшу в позиционных стратегиях.

Глава 7

Кооперативные дифференциальные игры в форме характеристической функции

§ 7.1. Определение кооперативной игры

7.1.1. Рассмотрим общую дифференциальную игру $\Gamma(x_0, T - t_0)$ для n лиц с уравнением движением вида

$$\dot{x}(s) = f[s, x(s), u_1(s), u_2(s), \dots, u_n(s)], \quad x(t_0) = x_0, \quad (7.1.1)$$

Выигрыш игрока i определяется так:

$$\int_{t_0}^T g^i[s, x(s), u_1(s), u_2(s), \dots, u_n(s)] ds + q^i(x(T)), \quad (7.1.2)$$
$$i \in N = \{1, 2, \dots, n\},$$

где $x(s) \in X \subset R^m$ — позиционная переменная игры и $u_i \in U^i$ — управление игрока $i \in N$. Выигрыши игроков предполагаются трансферабельными. Используя теорему п. 6.4.1, состоятельное позиционное равновесие может быть построено в предположении, что игроки в игре не кооперируются.

А теперь рассмотрим случай, когда игроки согласились на кооперацию. Обозначим через $\Gamma_c(x_0, T - t_0)$ кооперативную игру с игровой структурой игры $\Gamma(x_0, T - t_0)$, в которой игроки согласились действовать в соответствии с некоторым заранее определенным принципом оптимальности. Соглашение о том, как кооперироваться и как разделить выигрыш, получившийся в результате кооперации, составляет содержание принципа оптимальности в кооперативной игре.

Таким образом, принцип оптимальности в кооперативной игре $\Gamma_c(x_0, T - t_0)$ состоит из:

- 1) соглашения о множестве кооперативных стратегий (управлений),
- 2) механизма распределения общего выигрыша между игроками.

7.1.2. Принцип оптимальности должен сохранять свою эффективность вдоль кооперативной траектории $\{x_s^*\}_{s=t_0}^T$. Более того, групповая рациональность требует, чтобы игроки ориентировались на множество стратегий (управлений), дающих оптимальное по Парето решение. В дополнении к этому, принцип распределения выигрыша должен удовлетворять свойству индивидуальной рациональности в том смысле, что ни один из игроков не должен ухудшить свое состояние в результате кооперации.

Для выполнения свойства групповой рациональности в случае трансферабельных выигрышей игроки должны стремиться к максимизации суммарного выигрыша

$$\sum_{j=1}^N \left\{ \int_{t_0}^T g^j[s, x(s), u_1(s), u_2(s), \dots, u_n(s)] ds + q^j(x(T)) \right\}, \quad (7.1.3)$$

при условии (7.1.1).

7.1.3. Используя принцип максимума, множество оптимальных управлений $u^*(s) = [u_1^*(s), u_2^*(s), \dots, u_n^*(s)]$ может быть описано теоремой п. 6.2.1. Подставляя этот набор оптимальных управлений в (7.1.1), получаем оптимальную траекторию $\{x^*(t)\}_{t=t_0}^T$, где

$$x^*(t) = x_0 + \int_{t_0}^t f[s, x^*(s), u^*(s)] ds, \quad t \in [t_0, T]. \quad (7.1.4)$$

Как и ранее мы будем использовать как обозначение $x^*(t)$, так и обозначение x_t^* . Обозначим величину

$$\sum_{j=1}^n \left\{ \int_{t_0}^T g^j[s, x^*(s), u^*(s)] ds + q^j(x^*(T)) \right\}$$

через $v(N; x_0, T - t_0)$. Пусть $S \subseteq N$, и $v(S; x_0, T - t_0)$ — характеристическая функция, отражающая максимальный гарантированный выигрыш коалиции S . Величина $v(S; x_0, T - t_0)$ означает максимальный выигрыш коалиции S в случае, когда оставшиеся игроки из коалиции $N \setminus S$ играют против S . Используя супераддитивность характеристической функции, получаем $v(S; x_0, T - t_0) \geq v(S'; x_0, T - t_0)$, если $S' \subset S \subseteq N$. Поэтому игрокам выгодно создавать максимальную коалицию N для получения максимально возможного суммарного выигрыша $v(N; x_0, T - t_0)$ в данной игре.

§ 7.2. Дележи

7.2.1. Важнейшей частью исследования кооперативных игр является исследование возможности создания коалиций и предложение «приемлемого» распределения суммарного кооперативного выигрыша между игроками. В действительности, аппарат характеристических функций показывает возможности коалиций и является основой для построения схем распределения суммарного выигрыша, которые приемлемы для игроков-участников.

Мы будем использовать символ $\Gamma_v(x_0, T - t_0)$ для обозначения кооперативной дифференциальной игры в форме характеристической функции.

Исследуем сейчас различные решения игры $\Gamma_v(x_0, T - t_0)$.

Определение. Вектор $\xi(x_0, T - t_0) = [\xi_1(x_0, T - t_0), \xi_2(x_0, T - t_0), \dots, \xi_n(x_0, T - t_0)]$ удовлетворяющий условиям:

$$1) \quad \xi_i(x_0, T - t_0) \geq v(\{i\}; x_0, T - t_0), \quad i \in N,$$

$$2) \quad \sum_{j \in N} \xi_j(x_0, T - t_0) = v(N; x_0, T - t_0),$$

называется дележом в игре $\Gamma_v(x_0, T - t_0)$.

Условие 1 определения § 7.2. гарантирует индивидуальную рациональность в том смысле, что игрок в условиях данного дележа получает по крайней мере столько же, сколько бы он мог получить, если бы играл против всех игроков. Условие 2 гарантирует оптимальность по Парето, а, следовательно, и групповую рациональность.

Теорема. Предположим, что функция $w : 2^n \times R^m \times R^1 \rightarrow R^1$ аддитивна на $S \in 2^n$, т. е. для любых $S, A \in 2^n$, $S \cap A = \emptyset$ мы имеем $w(S \cup A; x_0, T - t_0) = w(S; x_0, T - t_0) + w(A; x_0, T - t_0)$. Тогда в игре $\Gamma_w(x_0, T - t_0)$ существует единственный дележ $\xi_i(x_0, T - t_0) = w(\{i\}; x_0, T - t_0)$ для всех $i \in N$.

Доказательство. Из аддитивности функции w мы немедленно получаем

$$w(N; x_0, T - t_0) = w(\{1\}; x_0, T - t_0) + \dots + w(\{n\}; x_0, T - t_0),$$

Отсюда и в силу индивидуальной рациональности дележа следует теорема § 7.2.

7.2.2. Игры с аддитивной характеристической функцией называются несущественными, в отличие от игр с супераддитивной функцией, которые называются существенными. В существенных играх $\Gamma_v(x_0, T - t_0)$ множество дележей бесконечно. Действительно, любой вектор вида

$$[v(\{1\}; x_0, T - t_0) + \alpha_1, v(\{2\}; x_0, T - t_0) + \alpha_2, \dots, v(\{n\}; x_0, T - t_0) + \alpha_n], \quad (7.2.1)$$

где $\alpha_i \geq 0$, $i \in N$ и

$$\sum_{i \in N} \alpha_i = v(N; x_0, T - t_0) - \sum_{i \in N} v(\{i\}; x_0, T - t_0),$$

является дележом в игре $\Gamma_v(x_0, T - t_0)$.

Обозначим множество дележей в игре $\Gamma_v(x_0, T - t_0)$ через $E_v(x_0, T - t_0)$.

Определение. Будем говорить, что дележ $\xi(x_0, T - t_0)$ доминирует дележ $\eta(x_0, T - t_0)$ по коалиции S или писать $\xi(x_0, T - t_0) \overset{S}{\succ} \eta(x_0, T - t_0)$, если

$$1) \quad \xi_i(x_0, T - t_0) > \eta_i(x_0, T - t_0), \quad i \in S,$$

$$2) \quad \sum_{i \in S} \xi_i(x_0, T - t_0) \leq v(S; x_0, T - t_0).$$

Дележ $\xi(x_0, T - t_0)$ доминирует дележ $\eta(x_0, T - t_0)$, или $\xi(x_0, T - t_0) \succ \eta(x_0, T - t_0)$, если существует коалиция $S \subset N$, такая что $\xi(x_0, T - t_0) \stackrel{S}{\succ} \eta(x_0, T - t_0)$. Из определения следует, что доминирование по одноэлементной и максимальной коалиции невозможно.

7.2.3. Определение. Множество недоминируемых дележей игры $\Gamma_v(x_0, T - t_0)$ называется C -ядром игры и обозначается через $C_v(x_0, T - t_0)$.

Определение. Множество $L_v(x_0, T - t_0) \subset E_v(x_0, T - t_0)$ называется решением по Нейману и Моргенштерну (NM -решением) игры $\Gamma_v(x_0, T - t_0)$, если

$$1) \quad \begin{aligned} &\xi(x_0, T - t_0), \eta(x_0, T - t_0) \in L_v(x_0, T - t_0), \text{ влечет} \\ &\xi(x_0, T - t_0) \not\succ \eta(x_0, T - t_0), \end{aligned}$$

$$2) \quad \begin{aligned} &\text{если } \eta(x_0, T - t_0) \notin L_v(x_0, T - t_0), \text{ то существует} \\ &\text{такой дележ } \xi(x_0, T - t_0) \in L_v(x_0, T - t_0), \text{ что} \\ &\xi(x_0, T - t_0) \succ \eta(x_0, T - t_0). \end{aligned}$$

Заметим, что NM -решение всегда содержит C -ядро, если последнее не пусто.

7.2.4. Определение. Вектор $\Phi^v(x_0, T - t_0) = \{\Phi_i^v(x_0, T - t_0), i = 1, \dots, n\}$ называется вектором Шепли, если он представим в виде

$$\begin{aligned} \Phi_i^v(x_0, T - t_0) &= \\ &= \sum_{S \subset N(S \ni i)} \frac{(n-s)!(s-1)!}{n!} [v(S; x_0, T - t_0) - v(S \setminus i; x_0, T - t_0)], \\ &\quad i \in N. \end{aligned}$$

Компоненты вектора Шепли являются выигрышами игроков от кооперации. Вектор Шепли всегда единственный и является дележом. Интересно отметить, что вектор Шепли, в отличие от ядра и NM -решения, определяет способ распределения суммарного выигрыша $v(N; x_0, T - t_0)$ без использования понятия доминирования.

§ 7.3. Дележи в динамике

7.3.1. В п. 7.2.1 мы построили дележи, определенные в начальный момент игры. В динамике игроков естественно интересует как происходит изменение дележей при развитии игры вдоль кооперативной траектории. В этом параграфе мы сосредоточим свое внимание на динамике дележей, подчиненных определенным принципам оптимальности.

Пусть в игре $\Gamma_v(x_0, T - t_0)$ выбран некоторый принцип оптимальности. Этот принцип оптимальности, примененный к игре из начальных состояний $x(t_0) = x_0$, $t = t_0$, определяет некоторое подмножество множества дележей $W_v(x_0, T - t_0) \subseteq E_v(x_0, T - t_0)$ и оптимальную траекторию $\{x^*(t)\}_{t=t_0}^T$, которая максимизирует

$$\sum_{j=1}^n \left\{ \int_{t_0}^T g^j[s, x^*(s), u^*(s)] ds + q^j(x^*(T)) \right\}.$$

Мы предполагаем также, что $W_v(x_0, T - t_0) \neq \emptyset$.

Определение. Любая траектория $\{x^*(t)\}_{t=t_0}^T$ системы (7.1.1), для которой имеет место

$$\sum_{j=1}^n \left\{ \int_{t_0}^T g^j[s, x^*(s), u^*(s)] ds + q^j(x^*(T)) \right\} = v(N; x_0, T - t_0),$$

называется оптимальной траекторией игры $\Gamma_v(x_0, T - t_0)$ или оптимальной кооперативной траекторией.

Из определения п. 7.3.1 следует, что вдоль оптимальной траектории игроки получают максимальный суммарный выигрыш. Предположим, что такая траектория существует. Рассмотрим поведение множества $W_v(x_0, T - t_0)$ вдоль оптимальной траектории $\{x^*(t)\}_{t=t_0}^T$. Для каждого текущего состояния $x^*(t) \equiv x_t^*$ на оптимальной траектории текущая подыгра $\Gamma_v(x_t^*, T - t)$ определяется следующим образом. В момент времени t и состоянии $x^*(t)$ определим характеристическую функцию

$$v(S; x_t^*, T - t) = \begin{cases} 0, & S = \emptyset, \\ \text{Val } \Gamma_S(x_t^*, T - t), & S \subset N, \\ K_N(x_t^*(\cdot), u^*(\cdot), T - t), & S = N, \end{cases} \quad (7.3.1)$$

где

$$K_N(x_t^*, u^*(\cdot), T - t) = \sum_{j=1}^n \left\{ \int_t^T g^j[s, x^*(s), u^*(s)] ds + q^j(x^*(T)) \right\}$$

представляет собой суммарный выигрыш игроков на промежутке времени $[t, T]$ вдоль оптимальной траектории $\{x^*(s)\}_{s=t}^T$, и $\text{Val } \Gamma_S(x_t^*, T - t)$ — значение антагонистической игры $\Gamma_S(x_t^*, T - t)$ между коалициями S и $N \setminus S$ из начального состояния $x^*(t) \equiv x_t^*$ продолжительностью $T - t$, в которой коалиция S является максимизирующей.

Множество дележей в игре $\Gamma_v(x_t^*, T - t)$ имеет вид

$$E_v(x_t^*, T - t) = \left\{ \xi \in R^n \left| \begin{aligned} &\xi_i \geq v(\{i\}; x_t^*, T - t), \quad i = 1, 2, \dots, n; \\ &\sum_{i \in N} \xi_i = v(N; x_t^*, T - t) \end{aligned} \right. \right\}, \quad (7.3.2)$$

где

$$\begin{aligned} v(N; x_t^*, T - t) &= \\ &= v(N; x_0, T - t_0) - \sum_{j=1}^n \left\{ \int_{t_0}^t g^j[s, x^*(s), u^*(s)] ds + q^j(x^*(T)) \right\}. \end{aligned}$$

Величина $\sum_{j=1}^n \left\{ \int_{t_0}^t g^j[s, x^*(s), u^*(s)] ds + q^j(x^*(T)) \right\}$ представляет собой кооперативный выигрыш игроков на промежутке времени $[t_0, t]$ вдоль траектории $\{x^*(s)\}_{s=t_0}^T$.

7.3.2. Рассмотрим семейство текущих игр $\{\Gamma_v(x_t^*, T - t), t_0 \leq t \leq T\}$ и их решений $W_v(x_t^*, T - t) \subset E_v(x_t^*, T - t)$, порожденных тем же принципом оптимальности, который определял решение $W_v(x_0, T - t_0)$ в начальный момент.

Лемма. Множество $W_v(x_T^*, 0)$ является решением текущей игры $\Gamma_v(x_T^*, 0)$ в момент T и состоит из единственного дележа

$$\begin{aligned} q(x^*(T)) &= \{q^1(x^*(T)), q^2(x^*(T)), \dots, q^n(x^*(T))\} = \\ &= \{q^1(x_T^*), q^2(x_T^*), \dots, q^n(x_T^*)\}. \end{aligned}$$

Доказательство. Поскольку игра $\Gamma_v(x_T^*, 0)$ имеет нулевую продолжительность, для всех $i \in N$ имеет место $v(\{i\}; x_T^*, 0) = q^i(x_T^*)$. Поэтому $\sum_{i \in N} v(\{i\}; x_T^*, 0) = \sum_{i \in N} q^i(x_T^*) = v(N; x_T^*, 0)$, и характеристическая функция игры $\Gamma_v(x_T^*, 0)$ аддитивна по S . Тогда согласно теореме п. 7.2.1 имеем

$$E_v(x_T^*, 0) = q(x_T^*) = W_v(x_T^*, 0),$$

что завершает доказательство леммы.

§ 7.4. Принцип динамической устойчивости

7.4.1. Построение оптимального поведения игроков является основным элементом теории кооперативных игр. Поведение игроков, удовлетворяющее определенному принципу оптимальности, образует решение игры. Иными словами, решения кооперативных игр порождаются целой совокупностью принципов оптимальности (например, вектор Шепли [Shapley, 1953], решение Неймана-Моргенштерна [Neumann, Morgenstern, 1944], арбитражное решение Нэша [Nash, 1953]). В динамических играх должно выполняться еще одно важное требование: выбранный принцип оптимальности должен генерировать то же решение в любой подыгре, возникающей вдоль оптимальной траектории, выбранной игроками в начальный момент игры. Это условие носит название динамической устойчивости или временной состоятельности. Предположим, что в начале игры игроки выбрали некоторый принцип оптимальности (который включает в себя соглашение о выборе траектории, максимизирующей суммарный выигрыш игроков). Когда игра развивается вдоль оптимальной траектории может оказаться, что выбранный принцип оптимальности или будет порождать пустое множество решений, или решение отличное от того, которое соответствовало ему в начальный момент игры. В этом случае некоторые игроки могут посчитать для себя более выгодным отклониться от кооперативной траектории (траектории, максимизирующей суммарный выигрыш игроков), выбранной ими в начале игры. Если такое происходит, то это ведет к неустойчивости процесса и, как следствие, к нереализуемости первоначально выбранного решения о кооперации. В частности, динамическая устойчивость (временная состоятельность) принципа оптимальности означает, что когда игра развивается вдоль кооперативной траектории, в каждый момент времени игроки, ориентируясь на один и тот же принцип оптимальности, следуют одним и тем же решениям и поэтому не имеют оснований отказаться от принципа оптимальности, принятого ими в начале игры и тем самым не имеют оснований отказаться от кооперации.

7.4.2. Вопрос динамической устойчивости в кооперативных дифференциальных играх подробно обсуждался последние десятилетия в научной литературе. А. Ори [Naurie, 1976] заметил возможную динамическую неустойчивость арбитражной схемы Нэша при попытке ее переноса на дифференциальные игры. Л. Петросян [Petrosjan, 1977] формализовал понятие динамической устойчивости для кооперативных дифференциальных

игр. Петросян и Данилов [Petrosjan, Danilov, 1979; 1982] ввели понятие процедуры распределения дележей во времени для кооперативных решений, которая позволяет в определенном смысле преодолеть проблему динамической устойчивости. В. Толвинский и др. [Tolwinski, 1986] рассмотрели кооперативное равновесие в дифференциальной игре с использованием стратегий угроз, которое позволяло игрокам оставаться на кооперативной траектории. В книге Л. Петросяна и Н. Зенкевича [Petrosjan, Zenkevich, 1996] приведен детальный анализ динамической устойчивости в дифференциальных играх. В частности, предложен метод регуляризации для построения динамически устойчивых (состоятельных во времени) принципов оптимальности. Д. Янг и Л. Петросян [Yeung, Petrosyan, 2001] построили состоятельные во времени (динамически устойчивые) решения дифференциальной игры и вывели условия, которым должен удовлетворять состоятельный во времени (динамически устойчивый) оптимальный дележ. Л. Петросян [Petrosyan, 2003] использовал метод регуляризации для построения состоятельных во времени (динамически устойчивых) арбитражных решений.

§ 7.5. Динамически устойчивые решения

7.5.1. Пусть заданы решения подыгр $W_v(x_t^*, T - t) \neq \emptyset$, $t_0 \leq t \leq T$ вдоль оптимальной кооперативной траектории (максимизирующей суммарный выигрыш игроков) $\{x^*(t)\}_{t=t_0}^T$. Если условие непустоты решений не выполнено, то игроки не могут следовать выбранному принципу оптимальности, поскольку в первый же момент времени t , когда $W_v(x_t^*, T - t) = \emptyset$, игроки не будут иметь возможности выбора решения, соответствующего первоначальному принципу оптимальности. Предположим, что в начальном состоянии x_0 игры игроки согласились на выбор дележа

$$\xi(x_0, T - t_0) = [\xi_1(x_0, T - t_0), \dots, \xi_n(x_0, T - t_0)] \in W_v(x_0, T - t_0).$$

Это означает, что игроки договорились на такой дележ суммарного выигрыша, при котором выигрыш i -го игрока на промежутке времени $[t_0, T]$ составляет $\xi_i(x_0, T - t_0)$. Если в соответствии с дележом игрок i должен получить выигрыш $\varpi_i[\xi(x_0, T - t_0); x^*(t), t - t_0]$ на отрезке времени $[t_0, t]$, то на оставшемся промежутке $[t, T]$ он должен получить в соответствии с дележом $\xi(x_0, T - t_0)$ выигрыш, равный

$$\eta_i[\xi(x_0, T - t_0); x^*(t), T - t] = \xi_i(x_0, T - t_0) - \varpi_i[\xi(x_0, T - t_0); x^*(t), t - t_0]. \quad (7.5.1)$$

7.5.2. Определение. Пусть $\eta[\xi(x_0, T - t_0); x^*(t), T - t]$ — вектор с компонентами

$$\eta_i[\xi(x_0, T - t_0); x^*(t), T - t], \quad \text{для } i \in \{1, 2, \dots, n\}.$$

Для того, чтобы первоначально выбранный принцип оптимальности $\xi(x_0, T - t_0)$ сохранял свою значимость в текущий момент t на оптимальной траектории, необходимо, чтобы вектор

$$\eta[\xi(x_0, T - t_0); x^*(t), T - t] \in W_v(x_t^*, T - t), \quad (7.5.2)$$

и тогда $\eta[\xi(x_0, T - t_0); x^*(t), T - t]$ будет действительно решением текущей подыгры $\Gamma_v(x_t^*, T - t)$. Если это условие выполняется в каждый момент времени $t \in [t_0, T]$ вдоль траектории $\{x^*(t)\}_{t=t_0}^T$, тогда дележ $\xi(x_0, T - t_0)$ динамически устойчив.

Вдоль траектории $x^*(t)$ на отрезке времени $[t, T]$, $t_0 \leq t \leq T$ коалиция состоящая из всех игроков N получает выигрыш

$$v(N; x^*(t), T - t) = \sum_{j=1}^n \left\{ \int_t^T g^j[s, x^*(s), u^*(s)] ds + q^j(x^*(T)) \right\}. \quad (7.5.3)$$

Тогда разность

$$v(N; x_0, T - t_0) - v(N; x^*(t), T - t) = \sum_{j=1}^n \left\{ \int_{t_0}^t g^j[s, x^*(s), u^*(s)] ds \right\}$$

есть выигрыш коалиции N на промежутке $[t_0, t]$.

Динамическая устойчивость (состоятельность во времени) дележа $\xi(x_0, T - t_0)$ гарантирует, что принцип оптимальности, породивший этот дележ, примененный к начальным условиям на оптимальной траектории в более поздние моменты времени приводит к дележу аналогичной структуры. Более того, сохраняется групповая и индивидуальная рациональность решения. Для выполнения этого условия необходимо ввести некоторый механизм реализации дележа во времени (механизм временных выплат).

§ 7.6. Процедура распределения дележа

7.6.1. Мы определим процедуру распределения дележа (ПРД), впервые введенную Л. Петросяном в 1978 г. таким образом, чтобы первоначально выбранный принцип оптимальности мог бы быть в действительности реализован в игре. Предположим, что выигрыш, получаемый игроком i на интервале времени $[t_0, t]$, может быть представлен в виде

$$\varpi_i[\xi(x_0(\cdot), T - t_0); x^*(\cdot), t - t_0] = \int_{t_0}^t B_i(s) ds, \quad (7.6.1)$$

где

$$\sum_{j \in N} B_j(s) = \sum_{j \in N} g^j[s, x^*(s), u^*(s)], \quad t_0 \leq s \leq t \leq T.$$

Из (7.6.1) получаем

$$\frac{d\varpi_i}{dt} = B_i(t). \quad (7.6.2)$$

Эта величина может интерпретироваться как мгновенный выигрыш игрока i в момент времени t . Очевидно, что вектор $B(t) = [B_1(t), B_2(t), \dots, B_n(t)]$ определяет распределение суммарного мгновенного выигрыша между игроками коалиции N . Правильным выбором функций $B(t)$ можно достичь того, чтобы игроки не были заинтересованы в моменты времени $t \in [t_0, T]$ отклониться от первоначального соглашения о дележе $\xi(x_0, T - t_0)$.

Определение. Дележ $\xi(x_0, T - t_0) \in W_v(x_0, T - t_0)$ динамически устойчив (состоятелен во времени) в игре $\Gamma_v(x_0, T - t_0)$, если выполнены следующие условия:

- 1) существует оптимальная траектория $\{x^*(t)\}_{t=t_0}^T$, вдоль которой $W_v(x^*(t), T - t) \neq \emptyset$, $t_0 \leq t \leq T$,

2) существуют функции $B(t) = [B_1(t), B_2(t), \dots, B_n(t)]$, интегрируемые на отрезке $[t_0, T]$ и такие, что

$$\sum_{j \in N} B_j(t) = \sum_{j \in N} g^j[t, x^*(t), u^*(t)], \quad t_0 \leq s \leq t \leq T,$$

$$\xi(x_0, T - t_0) \in \bigcap_{t_0 \leq t \leq T} (\varpi[\xi(x_0(\cdot), T - t_0); x^*(t), t - t_0] \oplus W_v(x^*(t), T - t)),$$

где вектор $\varpi[\xi(x_0(\cdot), T - t_0); x^*(t), t - t_0]$ с компонентами $\varpi_i[\xi(x_0(\cdot), T - t_0); x^*(t), t - t_0]$, $i \in N$; $W_v(x^*(t), T - t)$ — решение текущей подыгры $\Gamma_v(x^*(t), T - t)$ вдоль оптимальной траектории, и оператор \oplus означает следующее: для $\eta \in R^n$ и $A \subset R^n$, $\eta \oplus A = \{\eta + a | a \in A\}$.

7.6.2. Мы будем говорить, что кооперативная дифференциальная игра $\Gamma_v(x_0, T - t_0)$ имеет динамически устойчивое (состоятельное во времени) решение $W_v(x_0, T - t_0)$, если все дележи $\xi(x_0, T - t_0) \in W_v(x_0, T - t_0)$ динамически устойчивы (состоятельны во времени).

Из определения п. 7.6.1 получаем

$$\xi(x_0, T - t_0) \in (\varpi[\xi(x_0(\cdot), T - t_0); x^*(T), T - t_0] \oplus W_v(x^*(T), 0)),$$

где $W_v(x^*(T), 0) = q(x^*(T))$ — решение игры $\Gamma_v(x^*(T), 0)$. Поэтому можно написать

$$\xi(x_0, T - t_0) = \int_{t_0}^T B(s) ds + q(x^*(T)).$$

Динамически устойчивый дележ $\xi(x_0, T - t_0) \in W_v(x_0, T - t_0)$ может быть реализован следующим образом. Из определения п. 7.6.1 следует, что в каждый момент времени $t_0 \leq t \leq T$ мы имеем включение:

$$\xi(x_0, T - t_0) \in (\varpi[\xi(x_0, T - t_0); x^*(t), t - t_0] \oplus W_v(x^*(t), T - t)), \quad (7.6.3)$$

где $\varpi[\xi(x_0, T - t_0); x^*(t), t - t_0] = \int_{t_0}^t B(s) ds$ — вектор выигрышей игроков на отрезке времени $[t_0, t]$.

Выигрыш игрока i на этом же отрезке времени может быть выражен в виде:

$$\varpi_i[\xi(x_0, T - t_0); x^*(t), t - t_0] = \int_{t_0}^t B_i(s) ds.$$

Когда игра развивается на отрезке времени $[t_0, t]$, игроки делят суммарный заработанный на этом отрезке выигрыш

$$\int_{t_0}^t \sum_{j \in N} g^j[s, x^*(s), u^*(s)] ds$$

таким образом, что выполняется включение:

$$\xi(x_0, T - t_0) - \varpi[\xi(x_0, T - t_0); x^*(t), t - t_0] \in W_v(x^*(t), T - t). \quad (7.6.4)$$

Условие (7.6.4) обеспечивает существование вектора $\xi(x_t^*, T-t) \in W_v(x^*(t), T-t)$, удовлетворяющего соотношению

$$\xi(x_0, T-t_0) = \varpi[\xi(x_0, T-t_0); x^*(t), t-t_0] + \xi(x_t^*, T-t).$$

Таким образом, после выбора вектора $B(s)$ вектор выигрышей, получаемый игроками на оставшемся отрезке времени $[t, T]$, удовлетворяет условию

$$\begin{aligned} \xi(x_t^*, T-t) &= \xi(x_0, T-t_0) - \varpi[\xi(x_0, T-t_0); x^*(t), t-t_0] = \\ &= \int_t^T B(s) ds + q(x^*(T)), \end{aligned}$$

где

$$\begin{aligned} \sum_{j \in N} B_j(s) &= \sum_{j \in N} g^j[s, x^*(s), u^*(s)], \quad t \leq s \leq T, \\ \xi(x_t^*, T-t) &\in W_v(x^*(t), T-t). \end{aligned}$$

Изменяя вектор $\varpi[\xi(x_0, T-t_0); x^*(t), t-t_0]$ при условии

$$\sum_{j \in N} \varpi_j[\xi(x_0, T-t_0); x^*(t), t-t_0] = \int_{t_0}^t \sum_{j \in N} g^j[s, x^*(s), u^*(s)] ds,$$

игроки гарантируют расположение множества

$$(\varpi[\xi(x_0, T-t_0); x^*(t), t-t_0] \oplus W_v(x^*(t), T-t))$$

таким образом, то условие (7.6.3) выполняется.

7.6.3. Реализуя свои выигрыши с использованием ПРД $B(\tau)$, удовлетворяющей условиям (7.6.3)–(7.6.4) в каждый момент времени $t_0 \leq t \leq T$, игроки ориентированы на один и тот же принцип оптимальности, приводящий к одному и тому же дележу $\xi(x_t^*, T-t) \in W_v(x^*(t), T-t)$ в течение всей игры и поэтому не имеют оснований для пересмотра первоначального решения.

Динамическая неустойчивость решений кооперативной дифференциальной игры приводит к обесцениванию принципа оптимальности, породившего данное решение, поскольку выбранный в начале игры дележ из решения $W_v(x_0, T-t_0)$ не остается в этом решении, когда игра заканчивается. Именно поэтому мы считаем, что множество $W_v(x_0, T-t_0)$ может называться решением игры $\Gamma_v(x_0, T-t_0)$, если оно динамически устойчиво. В противном случае мы вынуждены констатировать, что игра $\Gamma_v(x_0, T-t_0)$ не имеет решения в смысле данного принципа оптимальности.

§ 7.7. Управление загрязнением окружающей среды

7.7.1. Рассмотрим модель загрязнения окружающей среды, предложенную Л. Петросяном и Д. Закуром в [Petrosjan, Zaccour, 2003]. Обозначим через N множество стран-участниц соглашения о сокращении вредных выбросов в атмосферу (игроков). Загрязнение страны (игрока) $i \in \{1, 2, \dots, n\} = N$ в момент времени $t \in [0, \infty)$ обозначим через $m_i(t)$. Пусть $x(t)$ — загрязнение, накопленное к моменту t с момента начала

игры. Изменение совокупной величины загрязнения происходит в соответствии с дифференциальным уравнением

$$\frac{dx(t)}{dt} = \dot{x}(t) = \sum_{i \in I} m_i(t) - \delta x(t), \quad x(0) = x_0, \quad (7.7.1)$$

где δ означает долю естественной очистки загрязнения.

Каждый из игроков стремится минимизировать общую дисконтированную сумму затрат на сокращение вредных выбросов и затрат, возникающих из-за потерь, вызванных загрязнением атмосферы. Последнее зависит от накопленного загрязнения. В дальнейшем, для упрощения обозначений, мы опускаем аргумент, означающий время в тех случаях, когда это не приводит к возможному непониманию текста. Обозначим через $C_i(m_i)$ затраты на сокращение выбросов игрока i при условии, что он ограничил свои выбросы величиной m_i , а через $D_i(x)$ обозначим затраты, возникающие из-за потерь, вызванных загрязнением атмосферы. Предположим, что обе функции непрерывно дифференцируемы и выпуклы, причем $C'(m_i) < 0$ и $D'(x) > 0$. Таким образом, задача каждой страны (игрока) i состоит в минимизации функционала

$$\min_{m_i} J^i(m, x) = \int_0^\infty \exp(-rs) \{C_i(m_i(s)) + D_i(x(s))\} ds \quad (7.7.2)$$

при условии (7.7.1), где $m = (m_1, m_2, \dots, m_n)$, и r — единый для всех участников коэффициент дисконтирования.

7.7.2. Наша модель выбрана исходя из следующих предпосылок. Во-первых, упрощенная динамика в рассмотренной эколого-экономической задаче позволяет выделить проблему распределения затрат между участниками соглашения и позволяет построить механизм распределения этих затрат во времени. Далее, в этой постановке присутствует основная особенность рассматриваемой проблемы, а именно то, что затраты каждого игрока зависят от общего уровня выбросов и накопленного к данному моменту загрязнения. Условие выпуклости рассматриваемых функций и условия на знак производных кажутся нам также вполне естественными. Например, условие выпуклости функции $C_i(e_i)$ означает, что прирост затрат на сокращение выбросов выше при невысоких уровнях выброса (см. [Germain, 1998]). Для простоты математических выкладок предполагается, что страны дисконтируют свои затраты одинаковым образом. В данной постановке для упрощения задачи мы считаем, что снижение уровня загрязнения происходит только за счет снижения уровня вредных выбросов в атмосферу и никак не учитываем возможность использования очистных мероприятий. Последний случай исследовался в работе [Krawczyk, Zassour, 1999].

7.7.3. Рассмотрим вопрос о распределении вектора Шепли на временном интервале. Для решения задачи распределения затрат на сокращение вредных выбросов использовалась методология теории кооперативных дифференциальных игр. Выделим основные шаги использования методологии:

- 1) Вычисление значений характеристической функции кооперативной игры.
- 2) Распределение между игроками суммарных кооперативных затрат в соответствии с вектором Шепли.
- 3) Распределение затрат, определяемых компонентами вектора Шепли для каждого игрока на отрезке времени с целью обеспечения динамической устойчивости (временной состоятельности) вектора Шепли.

Мы используем вектор Шепли в качестве принципа оптимальности по двум причинам: единственностью и возможностью построения для любых игровых задач. Первые два шага нашего подхода являются классическими (может быть кроме, частично, шага 1, поскольку способ вычисления значений характеристической функции в нашем случае не будет традиционным). Третий шаг предполагает распределение затрат на временном интервале состоятельным во времени способом. Ниже мы покажем, как это предполагается сделать.

7.7.4. Состояние игры определяется парой (t, x) . Тогда кооперативная подыгра, берущая начало из этого состояния будет обозначаться через $\Gamma_v(x, t)$. Обозначим через $x^N(t)$ траекторию (траекторию развития уровня загрязнения) при полной кооперации (большая коалиция N). В дальнейшем мы будем использовать два обозначения для кооперативной траектории $x^N(t)$ и x_t^N . Пусть $\Gamma_v(x_t^N, t)$ — подыгра с началом на кооперативной траектории. Значение характеристической функции, вычисленной для подкоалиции $K \subseteq N$ в подыгре $\Gamma_v(x, t)$, определяется как минимальные затраты этой подкоалиции и обозначается как $v(K; x, t)$ (далее мы подробно объясним, как вычисляются эти минимальные затраты). Следуя этому определению, общие кооперативные затраты, которые должны быть распределены между игроками, равны $v(N; x, 0)$, что и составляет минимальные затраты коалиции N и совпадает со значением характеристической функции для большой коалиции в игре $\Gamma_v(x, 0)$. Обозначим через $\Phi^v(x, t) = [\Phi_1^v(x, t), \Phi_2^v(x, t), \dots, \Phi_n^v(x, t)]$ вектор Шепли в подыгре $\Gamma_v(x, t)$. Обозначим далее через $B_i(t)$ затраты игрока i в момент времени t , и пусть $B(t) = (B_1(t), \dots, B_n(t))$.

Вектор $B(t) = [B_1(t), B_2(t), \dots, B_n(t)]$ представляет собой процедуру распределения дележа (ПРД), так что имеет место

$$\Phi_1^v(x, 0) = \int_0^\infty \exp(-rt) B_i(t) dt, \quad i = 1, \dots, n. \quad (7.7.3)$$

Интерпретация этого определения очевидна. Функция времени $B_i(t)$, являясь ПРД, распределяет затраты игрока i , определяемые компонентой вектора Шепли, вычисленной для всей игры $\Gamma_v(x, 0)$. Вектор $B(t)$ является динамически устойчивой (состоятельной во времени) ПРД, если для (x_t^N, t) и $t \in [0, \infty)$ выполнено следующее условие:

$$\Phi_i^v(x_0, 0) = \int_0^t \exp(-r\tau) B_i(\tau) d\tau + \exp(-rt) \Phi_i^v(x_t^N, t). \quad (7.7.4)$$

Для того, чтобы интерпретировать условие (7.7.4) предположим, что игроки пожела-ли пересмотреть кооперативное соглашение о затратах в игре $\Gamma_v(x, 0)$ в какой-то произвольный промежуточный момент времени t . В этот момент состояние системы будет $x^N(t)$, что означает, что кооперация игроков происходила до момента t и что каждый из игроков понес затраты, определяемые первым слагаемым в формуле (7.7.4). Если то, что он уже потратил до момента t плюс то, что он должен еще потратить, начиная с этого момента, ориентируясь на тот же принцип оптимальности (компоненту вектора Шепли в подыгре, начинающейся с этого момента) совпадает с компонентой вектора Шепли, тогда изменение первоначального соглашения не имеет смысла. Если можно найти такую ПРД $B(t) = [B_1(t), B_2(t), \dots, B_n(t)]$, что (7.7.4) выполнено, тогда эта ПРД динамически устойчива (состоятельна во времени). Мы предложим алгоритм построения такой ПРД в предположении дифференцируемости вектора Шепли. Можно убедиться в глубоком содержательном смысле формул для ПРД $B(t) = [B_1(t), B_2(t), \dots, B_n(t)]$.

7.7.5. Перейдем к формулировке алгоритма решения задачи. На первых трех шагах алгоритма вычисляются элементы, необходимые для вычисления характеристической функции, которое происходит на четвертом шаге. На последних двух шагах вычисляются вектор Шепли и функции ПРД $B_i(t)$, $i = 1, 2, \dots, n$.

Шаг 1. Минимизация суммарных затрат большой коалиции.

Большая коалиция решает задачу минимизации суммарных затрат методом динамического программирования при ограничениях, задаваемых динамикой развития накопленного загрязнения, т. е.

$$\min_{m_1, m_2, \dots, m_n} \sum_{i \in N} \int_t^{\infty} \exp[-r(\tau - t)] \{C_i(m_i(\tau)) + D_i(x(\tau))\} d\tau$$

при условии $\dot{x}(s) = \sum_{i \in N} m_i(s) - \delta x(s)$, $x(t) = x^N(t)$.

Обозначим через $W(N, x, t)$ функцию Беллмана для этой задачи оптимизации. В результате решения получаем вектор выбросов $m^N(x^N(\tau)) = [m_1^N(x^N(\tau)), \dots, m_n^N(x^N(\tau))]$ и соответствующее накопленное загрязнение при полной кооперации $x^N(\tau)$.

Шаг 2. Вычисляем состоятельное позиционное равновесие.

Поскольку игра происходит на бесконечном отрезке времени рассматриваются лишь стационарные стратегии. Для получения равновесия по Нэшу, при условии непрерывной дифференцируемости функций значения (выигрышей в равновесии по Нэшу) необходимо решить следующую систему нелинейных уравнений Айзекса–Беллмана (см. теорему § 6.6.)

$$r\bar{V}^i(x) = \min_{m_i} \left\{ C_i(m_i) + D_i(x) + \bar{V}_x^i(x) \left[\sum_{i \in I} m_i - \delta x \right] \right\}, \quad i \in N.$$

Обозначим через $m^*(x) = [m_1^*(x), m_2^*(x), \dots, m_n^*(x)]$ любое состоятельное позиционное равновесие кооперативной игры. Этот набор можно рассматривать как обычную стратегию выбросов при отсутствии кооперации. В дальнейшем для конкретного случая мы получим выражения этих стратегий в явном виде. Теперь же мы только заметим, что используя эти стратегии, мы можем получить равновесные затраты в игре $\Gamma_v(x_0, 0)$, которые мы обозначим через $V^i(0, x_0) = \bar{V}^i(x_0)$ и равновесные затраты в подыгре $\Gamma_v(x_t^N, t)$ которые мы обозначим через $V^i(t, x_t^N) = \bar{V}^i(x_t^N)$.

Шаг 3. Вычисление затрат для подкоалиций.

Для получения затрат для любого подмножества игроков (коалиций) мы поступим следующим образом. Затратами каждой коалиции будет сумма затрат игроков, входящих в коалицию. В функции затрат игроков и в правую часть дифференциального уравнения мы в качестве управлений игроков, не входящих в коалицию, подставим стратегии, полученные на шаге 2 (т. е. стратегии, входящие в равновесие по Нэшу). Обозначим через $W(K, x, t)$ значение этих затрат, вычисленные для коалиции K . Это

значение определяется из решения следующей задачи:

$$\begin{aligned} W(K, x, t) &= \\ &= \min_{m_i, i \in K} \sum_{i \in K} \left\{ \int_t^\infty \exp[-r(\tau - t)] \{C_i(m_i(\tau)) + D_i(x(\tau))\} d\tau \right\}, \\ \dot{x}(s) &= \sum_{i \in N} m_i(s) - \delta x(s), \quad x(t) = x^N(t). \\ m_j &= m_j^N, \quad \text{для } j \in I \setminus K. \end{aligned}$$

Шаг 4. Определим теперь характеристическую функцию.

Характеристическая функция $v(K; x, t)$ определяется следующим образом:

$$\begin{aligned} v(\{i\}; x, t) &= V^i(x, t) = \bar{V}^i(x), \quad i = 1, \dots, n; \\ v(K; x, t) &= W(K; x, t), \quad K \subseteq I. \end{aligned}$$

Шаг 5. Вычисление вектора Шепли.

Обозначим через $\Phi^v(x, t) = [\Phi_1^v(x, t), \Phi_2^v(x, t), \dots, \Phi_n^v(x, t)]$ вектор Шепли в игре $\Gamma_v(x, t)$; i -ая компонента вектора Шепли определяется по формуле

$$\Phi_i^v(x, t) = \sum_{K \ni i} \frac{(n-k)!(k-1)!}{n!} [W(K; x, t) - W(K \setminus \{i\}; x, t)],$$

где k — число игроков в коалиции K . Если кооперация продолжается в течение всей игры, то затраты игрока i определяются компонентой вектора Шепли в игре $\Gamma_v(x_0, 0)$ и равны

$$\Phi_i^v(x_0, 0) = \sum_{K \ni i} \frac{(n-k)!(k-1)!}{n!} [W(K; x_0, 0) - W(K \setminus \{i\}; x_0, 0)].$$

Обоснование нестандартного подхода к определению характеристической функции будет дано позже.

Шаг 6. Построение состоятельной во времени (динамически устойчивой) ПРД.

Распределяем затраты игрока i , $i = 1, \dots, n$ на отрезке времени $t \in [0, \infty)$ в соответствии с формулой

$$B_i(t) = r\Phi_i^v(x_t^N, t) - \frac{d}{dt}\Phi_i^v(x_t^N, t). \quad (7.7.5)$$

Формула (7.7.5) предписывает в каждый момент времени t затраты игрока i в соответствии с его будущими затратами минус производную будущих затрат.

7.7.6. Утверждение. Вектор $B(t) = (B_1(t), \dots, B_n(t))$, где компоненты $B_i(t)$ определяются формулой (7.7.5), является динамически устойчивой ПРД.

Доказательство. Покажем что вектор (7.7.5) является ПРД, т. е. $\int_0^\infty \exp(-rt) B_i(t) dt = \Phi_i^v(x_0, 0)$.

Умножая (7.7.5) на дисконт $\exp(-rt)$ и интегрируя, получаем

$$\begin{aligned} \int_0^\infty \exp(-rt) B_i(t) dt &= \int_0^\infty \exp(-rt) \left[r\Phi_i^v(x_t^N, t) - \frac{d}{dt}\Phi_i^v(x_t^N, t) \right] dt = \\ &= -\exp(-rt) \Phi_i^v(x_t^N, t) \Big|_0^\infty = \Phi_i^v(x_0, 0). \end{aligned}$$

Применяя аналогичным образом интегрирование для $\Phi_i^v(x_t^N, t)$, можно показать, что $\Phi_i^v(x_0, 0) = \int_0^t \exp(-r\tau) B_i(\tau) d\tau + \exp(-rt) \Phi_i^v(x_t^N, t)$, что и есть условие динамической устойчивости.

7.7.7. Перейдем теперь к обоснованию алгоритма и особого вида характеристической функции. Как было ранее отмечено в работе [Petrosjan, Zaccour, 2003], при формулировке алгоритма решения задачи важнейшим элементом теории переговоров является определение точки статус-кво, которая показывает, что могут получить игроки, если переговоры проваливаются. Эта точка определяет индивидуальную силу игрока, когда он действует в одиночку. Та же идея применима и к подмножеству игроков. Для измерения силы некоторого подмножества игроков (коалиции) необходимо обратиться к понятию характеристической функции — математическому аппарату, в точности созданному для измерения подобной силы. Все известные решения кооперативной теории (C -ядро, вектор Шепли, NM -решение и др.) используют аппарат характеристических функций для определения множеств дележей, определяющих то или иное решение. В частности, C -ядро состоит из дележей, выбор которых не может быть оспорен ни одной из коалиций, а вектор Шепли представляет собой дележ, удовлетворяющий определенной системе аксиом. Если полученное таким образом множество дележей не состоит из единственного дележа, игроки могут вести переговоры о выборе единственного дележа из множества дележей, составляющих решение. В динамической (дифференциальной) игре дележи соответствуют выигрышам (в нашем случае — дисконтированным суммам), получаемым игроками в течение всей игры. Важнейшим условием является то, чтобы распределение выигрыша во времени было бы допустимым, т. е. чтобы интегрированный выигрыш, получаемый им в течение всей игры, совпадал бы с его компонентой дележа-решения (см. определение процедуры распределения дележа (ПРД) § 7.6.) Очевидно, что можно построить бесконечное число подобных распределений дележа на временном интервале игры, но далеко не все распределения могут быть концептуально и интуитивно обоснованы. Основной принцип, заложенный нами в распределение индивидуальных затрат на временном интервале, заключается в том, что если игроки захотят пересмотреть первоначальное соглашение о выбранном дележе в любой промежуточный момент времени на оптимальной кооперативной траектории, то они получат тот же результат (тот же дележ).

Заметим, что в данном случае вычисление характеристической функции производится нестандартным путем. Предполагается, что игроки, не входящие в коалицию, а именно, игроки из коалиции $N \setminus K$, используют свои равновесные по Нэшу стратегии, при вычислении значения характеристической функции для коалиции K . Заметим, что в теории игр имеется не так много подходов к определению характеристической функции.

Классический подход был предложен Д. фон Нейманом и О. Моргенштерном [Neumann, Morgenstern, 1944], в котором они предполагают, что игроки, не входящие в коалицию, стремятся, объединившись, максимизировать затраты игроков, входящих в коалицию. Этот подход, который приводит к минимальным гарантированным затратам, не совсем подходит при рассмотрении нашей задачи. Действительно, маловероятно, что если страны объединились в коалицию для решения своих проблем, связанных с охраной окружающей среды, другие страны объединятся в антикоалицию с тем, чтобы препятствовать их усилиям. Мы, конечно, признаем возможность использования подхода фон Неймана и Моргенштерна в кооперативной теории, однако в нашем случае он не применим.

Другим подходом является аксиоматический подход, при котором значения характеристической функции приписываются коалициям из некоторых внешних соображений, показывающих силу коалиции, как это было предложено Д. Филаром и П. Гертнером [Filar, Gaertner, 1997]. Этот подход применим в тех случаях, когда сила коалиции может быть оценена экспертным путем без анализа индивидуальных возможностей игроков. Кроме того, такой подход сильно затруднен при использовании в дифференциальных играх, поскольку не позволяет найти зависимость характеристической функции от начальных состояний игры. Здесь мы используем более традиционный путь для определения значений характеристической функции, который основывается непосредственно на исследовании игрового процесса как такового.

Третий подход заключается в предположении, что значения характеристической функции вычисляются как выигрыши в равновесии по Нэшу в игре между этой коалицией и антикоалицией. Здесь мы сталкиваемся с вычислительными трудностями и трудностями, связанными с неединственностью равновесия. Действительно, для определения значений характеристической функции таким образом нам пришлось бы решить $2^n - 2$ дифференциальных неантагонистических игр (число, равное числу непустых коалиций в игре). В нашем же случае мы решаем лишь одну дифференциальную игру (находим равновесия по Нэшу). Все остальные задачи сводятся к стандартным задачам оптимального управления. Такой подход значительно проще, поскольку решение дифференциальной игры существенно сложнее задачи оптимального управления. Вспомним теперь, что в качестве оптимального дележа нами выбран вектор Шепли. Поэтому нашей целью является вычисление i -ой компоненты этого вектора. Для ее вычисления нам нужно вычислить маргинальные вклады игрока во все коалиции, т. е. величины $v(K, S, t) - v(K \setminus \{i\}, S, t)$. Если бы нам пришлось пользоваться третьим подходом к определению характеристической функции, то нам пришлось бы находить равновесия по Нэшу в игре между коалицией K и контрcoalцией $N \setminus K$. При этом если в какой-то из этих $2^n - 2$ игр равновесие по Нэшу было бы неединственным, то мы столкнулись бы с исключительно сложной проблемой выбора.

При нашем подходе игроки не входящие в коалицию придерживаются равновесных стратегий (выбросов) из одного и того же равновесия по Нэшу, которое было первоначально найдено в игре. Если окажется так, что на шаге 2 алгоритма мы получим несколько равновесий по Нэшу, то мы просто можем подсчитать вектор Шепли для каждого из них, минуя при этом проблему выбора.

7.7.8. Пример 1. Рассмотрим здесь некоторые частные виды функций, фигурирующих в нашей задаче. Пусть в (7.7.2) имеем

$$C_i(m_i) = \frac{\gamma}{2} [m_i - \bar{m}_i]^2, \quad 0 \leq m_i \leq \bar{m}_i, \quad \gamma > 0 \text{ и } i \in \{1, 2, 3\};$$

$$D_i(x) = \pi x, \quad \pi > 0.$$

Вычисление оптимальных затрат большой коалиции (Шаг 1).

Функция значения $W(N, x, t)$ удовлетворяет следующему уравнению Беллмана:

$$rW(N, x, t) =$$

$$= \min_{m_1, m_2, m_3} \left\{ \sum_{i=1}^3 \left(\frac{\gamma}{2} [m_i - \bar{m}_i]^2 + \pi x \right) + W_x(N, x, t) \left[\sum_{i=1}^3 m_i - \delta x \right] \right\}.$$

(7.7.6)

Проводя операцию минимизации в (7.7.6), получаем

$$m_i^N = \bar{m}_i - \frac{1}{\gamma} W_x(N, x, t), \quad \text{при } i \in \{1, 2, 3\}.$$

Подставляя m_i^N в (7.7.6) и решая, получаем

$$W(N, x, t) = \bar{W}(N, x) = \frac{3\pi}{r(r+\delta)} \left\{ \left[\sum_{i=1}^3 \bar{m}_i - \frac{3^2\pi}{2\gamma(r+\delta)} \right] + rx \right\}, \text{ и} \quad (7.7.7)$$

$$m_i^N = \bar{m}_i - \frac{3\pi}{\gamma(r+\delta)}, \text{ при } i \in \{1, 2, 3\}. \quad (7.7.8)$$

Оптимальная траектория выбросов имеет вид

$$x^N(t) = \exp(-\delta t) x(0) + \frac{1}{\delta} \left\{ \left[\sum_{i=1}^3 m_i^N \right] [1 - \exp(-\delta t)] \right\}. \quad (7.7.9)$$

Вычисление равновесия по Нэшу (Шаг 2).

Для нахождения состоятельного позиционного равновесия по Нэшу используем теорему § 6.6. и получаем следующее уравнение Беллмана

$$r\bar{V}^i(x) = \min_{m_i} \left\{ \frac{\gamma}{2} [m_i - \bar{m}_i]^2 + \pi x + \bar{V}_x^i(x) \left[\sum_{\substack{j \in \{1, 2, 3\} \\ i \neq j}} m_j^* + m_i - \delta x \right] \right\},$$

при $i \in \{1, 2, 3\}.$ (7.7.10)

Находя минимум в правой части (7.7.10), получаем

$$m_i^* = \bar{m}_i - \frac{1}{\gamma} \bar{V}_x^i(x), \quad i \in \{1, 2, 3\}. \quad (7.7.11)$$

Подставляя (7.7.11) в (7.7.10) и решая уравнение, находим

$$\bar{V}^i(x) = \frac{\pi}{r(r+\delta)} \left\{ \frac{\pi}{2\gamma(r+\delta)} + \sum_{i=1}^3 \bar{m}_i - \frac{3\pi}{\gamma(r+\delta)} + rx \right\},$$

при $i \in \{1, 2, 3\}.$ (7.7.12)

Тогда равновесные по Нэшу выбросы имеют вид

$$m_i^* = \bar{m}_i^* - \frac{\pi}{\gamma(r+\delta)}, \quad i \in \{1, 2, 3\}. \quad (7.7.13)$$

Разница между выбросами по Нэшу и выбросами при кооперации состоит в том, что в кооперативном случае игрок учитывает маргинальные затраты всех членов большой коалиции, а не только свои.

Вычисление оптимальных затрат для промежуточных коалиций (Шаг 3).

Функция значения $W(K, x, t)$ для каждой коалиции K , состоящей из двух игроков, должна удовлетворять следующему уравнению Беллмана:

$$\begin{aligned} rW(K, x, t) = \\ = \min_{m_1, i \in K} \left\{ \sum_{i \in K} \left(\frac{\gamma}{2} [m_i - \bar{m}_i]^2 + \pi x \right) + W_x(K, x, t) \left[\sum_{i \in K} m_i + m_j^* - \delta x \right] \right\}, \end{aligned} \quad (7.7.14)$$

где $j \notin K$.

Пользуясь тем же способом решения, который мы использовали для нахождения значения для большой коалиции, получаем

$$\begin{aligned} W(K, x, t) = \bar{W}(K, x) = \\ = \frac{2\pi}{r(r+\delta)} \left\{ \sum_{i \in K} \bar{m}_i - \frac{4\pi}{2\gamma(r+\delta)} - \frac{\pi}{\gamma(r+\delta)} + rx \right\}. \end{aligned} \quad (7.7.15)$$

Соответствующие выбросы для коалиции K будут равны

$$m_i^K = \bar{m}_i - \frac{2\pi}{\gamma(r+\delta)}, \quad i \in K. \quad (7.7.16)$$

Определение характеристической функции (Шаг 4).

$$\begin{aligned} v(\{i\}; x, t) = V^i(x, t) = \bar{V}^i(x) = \\ = \frac{\pi}{r(r+\delta)} \left\{ \frac{\pi}{2\gamma(r+\delta)} + \sum_{i=1}^3 \bar{m}_i - \frac{3\pi}{\gamma(r+\delta)} + rx \right\}, \quad i = 1, 2, 3; \\ v(K; x, t) = W(K, x, t) = \bar{W}(K, x) = \\ = \frac{2\pi}{r(r+\delta)} \left\{ \sum_{i \in K} \bar{m}_i - \frac{4\pi}{2\gamma(r+\delta)} - \frac{\pi}{\gamma(r+\delta)} + rx \right\}, \\ K \subseteq \{1, 2, 3\}. \end{aligned}$$

Вычисление вектора Шепли (Шаг 5).

Если значения \bar{m}_i симметричны, то вектор Шепли вычисляется в явном виде

$$\begin{aligned} \Phi_i^v(x, t) &= \sum_{K \ni i} \frac{(n-k)!(k-1)!}{n!} [v(K; x, t) - v(K \setminus \{i\}; x, t)] = \\ &= \frac{1}{2r(r+\delta)} \left\{ 2\pi \left(\sum_{i=1}^3 \bar{m}_i + \rho S \right) - \frac{9\pi^2}{\gamma(r+\delta)} \right\}, \\ &i = 1, 2, 3. \end{aligned} \quad (7.7.17)$$

Вычисление функций ПРД (Шаг 6).

Заметим, что согласно (7.7.9) функции ПРД имеют вид

$$B_i(t) = r\Phi_i^v(x_t^N, t) - \frac{d}{dt}\Phi_i^v(x_t^N, t).$$

Прямые вычисления дают нам

$$B_i(t) = \pi x^N(t) + \frac{9\pi^2}{2\gamma(r+\delta)^2}, \quad i = 1, 2, 3. \quad (7.7.18)$$

Умножая обе части (7.7.18) на дисконт-фактор и интегрируя, получаем

$$\int_0^\infty \exp(-rt) B_i(t) dt = \int_0^\infty \exp(-rt) \left[\pi x^N(t) + \frac{9\pi^2}{2\gamma(r+\delta)^2} \right] dt, \\ i = 1, 2, 3, \quad (7.7.19)$$

и из (7.7.8)–(7.7.9) получаем

$$x^N(t) = \exp(-\delta t) x(0) + \frac{1}{\delta} \left\{ \left[\sum_{j=1}^3 \left(\bar{m}_j - \frac{3\pi}{2\gamma(r+\delta)} \right) \right] [1 - \exp(-\delta t)] \right\}.$$

Подстановка $x^N(t)$ в (7.7.19) дает нам

$$\int_0^\infty \exp(-rt) \beta_i(t) dt = \\ = \int_0^\infty \exp[-(r+\delta)t] \pi x_0 dt + \int_0^\infty \exp(-rt) \frac{\pi}{\delta} \left(\sum_{i=1}^3 \bar{m}_i - \frac{9\pi}{\gamma(r+\delta)} \right) dt + \\ + \int_0^\infty \exp[-(r+\delta)t] \frac{\pi}{\delta} \left(\sum_{i=1}^3 \bar{m}_i - \frac{9\pi}{\gamma(r+\delta)} \right) dt + \int_0^\infty \exp(-rt) \frac{9\pi^2}{2\gamma(r+\delta)^2} dt.$$

И, интегрируя, получаем

$$\int_0^\infty \exp(-rt) \beta_i(t) dt = \\ = \frac{1}{2r(r+\delta)} \left\{ 2\pi \left(\sum_{i=1}^3 \bar{m}_i + rx(0) \right) - \frac{9\pi^2}{\gamma(r+\delta)} \right\} = \Phi_i^v(x_0, 0), \\ i = 1, 2, 3.$$

§ 7.8. Упражнения и задачи

1. Рассмотрим игру со следующей характеристической функцией:

$$\begin{aligned} v(\{i\}) &= 0, \text{ при } i \in \{1, 2\}; \\ v(\{3\}) &= 50; \\ v(\{1, 2\}) &= 100; \quad v(\{1, 3\}) = 200; \quad v(\{2, 3\}) = 300; \\ v(\{1, 2, 3\}) &= 500. \end{aligned}$$

а) Вычислить вектор Шепли.

б) Вычислить вектор Шепли, предположив, что $v(\{2, 3\}) = 200$ вместо 300.

2. Рассмотрим игру со следующей характеристической функцией:

$$\begin{aligned} v(\{i\}) &= 0, \text{ при } i \in \{1, 2, 3\}; \quad v(\{4\}) = 0.15; \\ v(\{i, j\}) &= 0.1, \text{ при } i, j \in \{1, 2, 3\}, \quad i \neq j; \\ v(\{i, 4\}) &= 0.2, \text{ при } i \in \{1, 2, 3\}; \\ v(\{i, j, k\}) &= 0.5, \text{ при } i, j, k \in \{1, 2, 3\}, \quad i \neq j \neq k; \\ v(\{1, 2, 4\}) &= 0.55; \quad v(\{1, 3, 4\}) = 0.6; \quad v(\{2, 3, 4\}) = 0.65; \\ v(\{1, 2, 3, 4\}) &= 1. \end{aligned}$$

а) Вычислить вектор Шепли.

б) Вычислить вектор Шепли, предположив, что $v(\{4\}) = 0.2$ вместо 0.15.

3. Рассмотрим игру со следующей характеристической функцией:

$$\begin{aligned} v(\{i\}) &= w_i, \quad i = 1, 2, 3, 4; \\ v(\{i, j\}) &= w_{ij}, \quad i, j \in \{1, 2, 3, 4\}, \quad i < j; \\ v(\{i, j, k\}) &= w_{ijk}, \quad i, j, k \in \{1, 2, 3, 4\}, \quad i < j < k; \\ v(\{1, 2, 3, 4\}) &= W; \end{aligned}$$

причем

$$\begin{aligned} w_i + w_j &\leq w_{ij}, \quad \text{при } i, j \in \{1, 2, 3, 4\}, \quad i \neq j; \\ w_{ij} + w_k &\leq w_{ijk}, \quad \text{при } i, j, k \in \{1, 2, 3, 4\}, \quad i \neq j \neq k; \\ w_{ij} + w_{kl} &\leq W, \quad \text{при } i, j, k, l \in \{1, 2, 3, 4\}, \quad i \neq j \neq k \neq l; \\ w_{ijk} + w_l &\leq W, \quad \text{при } i, j, k, l \in \{1, 2, 3, 4\}, \quad i \neq j \neq k \neq l; \\ w_i &\geq 0, \quad \text{при } i \in \{1, 2, 3, 4\}. \end{aligned}$$

Вычислить вектор Шепли.

4. Рассмотрим следующую динамическую эколого-экономическую модель игры 3 лиц. Выбросы страны $i \in \{1, 2, 3\}$ в момент времени t ($t \in [0, \infty)$) обозначим как $m_i(t)$. Пусть $x(t)$ — уровень накопленных загрязнений к моменту t . Уравнение динамики имеет следующий вид:

$$\frac{dx(t)}{dt} = \dot{x}(t) = \sum_{i=1}^3 m_i(t) - \delta x(t), \quad x(0) = 100,$$

где $\delta = 0.05$ — коэффициент абсорбции.

Каждая страна стремится к минимизации расходов на сокращение выбросов и устранение последствий загрязнения:

$$\min_{m_i} J^i(m, x) = \int_0^\infty \exp(-rs) \{C_i(m_i(s)) + D_i(x(s))\} ds,$$

где

$$\begin{aligned} C_i(m_i) &= \frac{1}{2} [m_i - \bar{m}_i]^2, \quad 0 \leq m_i \leq \bar{m}_i, \quad \gamma > 0, \\ D_i(x) &= x. \end{aligned}$$

Предположим, что страны договорились о сотрудничестве и решили разделить общие расходы согласно вектору Шепли. Следуя подходу Петросяна и Заккура

[Petrosyan, Zaccour, 2003], построить характеристическую функцию. Вычислить вектор Шепли и процедуру распределения дележа (ПРД).

5. Рассмотреть кооперативный вариант игры управления вредными выбросами в модели [Jorgensen, Martin-Herran, Zaccour, 2003] (см. упр. 7 главы 6). Предположим, что страны договорились о сотрудничестве и решили разделить общие расходы согласно вектору Шепли. Следуя подходу Петросяна и Заккура [Petrosyan, Zaccour, 2003], построить характеристическую функцию. Вычислить вектор Шепли и процедуру распределения дележа (ПРД).

6. Изучить модель управления загрязнениями (см. § 7.7) при предположении, что игра рассматривается со случайной продолжительностью (см. упр. 8–10 главы 6). Построить характеристическую функцию и вычислить вектор Шепли.

Глава 8

Кооперативные дифференциальные игры двух лиц с дисконтированием

§ 8.1. Постановка задачи

8.1.1. Рассмотрим неантагонистическую дифференциальную игру двух лиц из начального состояния x_0 продолжительности $T - t_0$ в пространстве состояний $X \subset R^m$ с множеством допустимых траекторий $\{x(s), t_0 \leq s \leq T\}$. Состояние игры изменяется в соответствии с системой дифференциальных уравнений вида

$$\dot{x}(s) = f[s, x(s), u_1(s), u_2(s)], \quad x(t_0) = x_0. \quad (8.1.1)$$

В момент времени $s \in [t_0, T]$ мгновенный выигрыш (плотность выигрыша) игрока $i \in \{1, 2\}$ обозначим через $g^i[s, x(s), u_1(s), u_2(s)]$. Кроме того, в момент T завершения игры игрок i получает терминальный выигрыш $q^i(x(T))$. Выигрыши предполагаются трансферабельными как между игроками, так и во времени. Предполагается заданным параметр дисконтирования $r(s)$ в каждый момент времени $s \in [t_0, T]$ и поэтому выигрыши, получаемые в момент t после момента начала игры t_0 , умножаются на величину $\exp\left[-\int_{t_0}^t r(y) dy\right]$. Таким образом, в момент t_0 выигрыш игрока $i \in \{1, 2\}$ определяется по формуле

$$\begin{aligned} & \int_{t_0}^T g^i[s, x(s), u_1(s), u_2(s)] \exp\left[-\int_{t_0}^s r(y) dy\right] ds + \\ & + \exp\left[-\int_{t_0}^T r(y) dy\right] q^i(x(T)). \end{aligned} \quad (8.1.2)$$

Рассмотрим случай некооперативного поведения игроков. Обозначим через $\Gamma(x_0, T - t_0)$ игру (8.1.1)–(8.1.2). Используя теорему п. 6.4.2, мы можем охарактеризовать решение игры $\Gamma(x_0, T - t_0)$ следующим образом.

Теорема. *Множество стратегий $\left\{u_1^{(t_0)*}(t) = \phi_1^{(t_0)*}(t, x), u_2^{(t_0)*}(t) = \phi_2^{(t_0)*}(t, x)\right\}$ образует состоятельное позиционное равновесие по Нэ*

шу в игре $\Gamma(x_0, T - t_0)$, если существуют непрерывно дифференцируемые функции $V^{(t_0)1}(t, x) : [t_0, T] \times R^m \rightarrow R$ и $V^{(t_0)2}(t, x) : [t_0, T] \times R^m \rightarrow R$, удовлетворяющие следующей системе уравнений Айзекса–Беллмана:

$$-V_t^{(t_0)i}(t, x) = \max_{u_i} \left\{ g^i \left[t, x, u_i, \phi_j^{(t_0)*}(t, x) \right] \exp \left[- \int_{t_0}^t r(y) dy \right] + \right. \\ \left. + V_x^{(t_0)i}(t, x) f \left[t, x, u_i, \phi_j^{(t_0)*}(t, x) \right] \right\},$$

и

$$V^{(t_0)i}(T, x) = \exp \left[- \int_{t_0}^T r(y) dy \right] q^i(x), \\ i \in \{1, 2\}, \quad j \in \{1, 2\}, \quad j \neq i.$$

Состоятельные позиционные стратегии являются марковскими и зависят лишь от текущего момента времени t и текущего состояния x и поэтому не зависят от предыстории игры.

Рассмотрим подыгру $\Gamma(x_\tau, T - \tau)$ со структурой выигрышей вида (8.1.1) и с динамикой (8.1.2), начинающуюся в момент времени $\tau \in [t_0, T]$ из состояния $x_\tau \in X$. Следуя теореме п. 8.1.1, обозначим равновесие по Нэшу в состоятельных позиционных стратегиях в игре $\Gamma(x_\tau, T - \tau)$ через $\{\phi_1^{(\tau)*}(t, x), \phi_2^{(\tau)*}(t, x)\}$ и соответствующую функцию значения для игрока $i \in \{1, 2\}$ (выигрыш в ситуации равновесия как функция начального состояния) через $V^{(\tau)i}(t, x_t) : [\tau, T] \times R^n \rightarrow R$. Тогда функции $V^{(\tau)1}(t, x)$ и $V^{(\tau)2}(t, x)$ удовлетворяют следующей системе уравнений:

$$-V_t^{(\tau)i}(t, x) = \max_{u_i} \left\{ g^i \left[t, x, u_i(t, x), \phi_j^{(\tau)*}(t, x) \right] \exp \left[- \int_{\tau}^t r(y) dy \right] + \right. \\ \left. + V_x^{(\tau)i}(t, x) f \left[t, x, u_i(t, x), \phi_j^{(\tau)*}(t, x) \right] \right\}, \\ V^{(\tau)i}(T, x) = \exp \left[- \int_{\tau}^T r(y) dy \right] q^i(x), \\ i \in \{1, 2\}, \quad j \in \{1, 2\}, \quad j \neq i. \quad (8.1.3)$$

8.1.2. Замечание 1. Заметим, что равновесные стратегии являются марковскими в том смысле, что они зависят от текущего состояния и времени. Сравнивая уравнения Айзекса–Беллмана (8.1.3) для различных значений $\tau \in [t_0, T]$, можно заметить, что

$$\phi_i^{(\tau)*}(s, x(s)) = \phi_i^{(t_0)*}(s, x(s)), \quad s \in [\tau, T],$$

$$V^{(\tau)i}(\tau, x_\tau) = \exp \left[\int_{t_0}^{\tau} r(y) dy \right] V^{(t_0)i}(\tau, x_\tau),$$

$$V^{(t)i}(t, x_t) = \exp \left[\int_{\tau}^t r(y) dy \right] V^{(\tau)i}(t, x_t),$$

$$t_0 \leq \tau \leq t \leq T, \quad i \in \{1, 2\}.$$

В равновесии по Нэшу игры $\Gamma(x_\tau, T - \tau)$ настоящее значение выигрыша i -го игрока на промежутке $[t, T]$ при $x(t) = x_t$ и $t \in [\tau, T]$ определяется по формуле

$$\begin{aligned} V^{(\tau)i}(t, x_t) = & \\ = \int_t^T g^i[s, x(s), \phi_1^{(\tau)*}(s, x(s)), \phi_2^{(\tau)*}(s, x(s))] \exp\left[-\int_\tau^s r(y) dy\right] ds + & \\ + \exp\left[-\int_\tau^T r(y) dy\right] q^i(x(T)) \Big|_{x(t) = x_t}, \quad i \in \{1, 2\}. & \end{aligned}$$

Траектория игры в ситуации равновесия получается из решения системы дифференциальных уравнений

$$\dot{x}(s) = f\left[s, x(s), \phi_1^{(\tau)*}(s, x(s)), \phi_2^{(\tau)*}(s, x(s))\right], \quad x(t) = x_\tau. \quad (8.1.4)$$

8.1.3. Пример 1. Рассмотрим игру добычи ограниченного ресурса, в которой две фирмы имеют лицензию на проведение работ на отрезке времени $[t_0, T]$. Запасы ресурса $x(s) \in X \subset R$ изменяются в соответствии с дифференциальным уравнением

$$\dot{x}(s) = ax(s)^{1/2} - bx(s) - u_1(s) - u_2(s), \quad x(t_0) = x_0 \in X, \quad (8.1.5)$$

где $u_1(s)$ — уровень добычи игрока 1 и $u_2(s)$ — уровень добычи игрока 2. Мгновенные выигрыши в момент времени $s \in [t_0, T]$ игроков 1 и 2 соответственно равны

$$\left[u_1(s)^{1/2} - \frac{c_1}{x(s)^{1/2}} u_1(s) \right], \quad \left[u_2(s)^{1/2} - \frac{c_2}{x(s)^{1/2}} u_2(s) \right],$$

где c_1 и c_2 константы и $c_1 \neq c_2$.

В момент времени T каждый добывающий получает дополнительный выигрыш в размере

$$q^i(x(T)) = qx(T)^{1/2},$$

который зависит от величины запаса ресурса в момент окончания игры.

Выигрыши трансферабельны во времени и между игроками. Задан параметр дисконтирования r и поэтому выигрыши, получаемые в момент t после момента начала игры t_0 , умножаются на множитель $\exp[-r(t - t_0)]$.

В момент времени t_0 выигрыши игроков 1 и 2 соответственно равны

$$\begin{aligned} \int_{t_0}^T \left[u_1(s)^{1/2} - \frac{c_1}{x(s)^{1/2}} u_1(s) \right] \exp[-r(t - t_0)] ds + \\ + \exp[-r(T - t_0)] qx(T)^{1/2}, \end{aligned}$$

и

$$\begin{aligned} \int_{t_0}^T \left[u_2(s)^{1/2} - \frac{c_2}{x(s)^{1/2}} u_2(s) \right] \exp[-r(t - t_0)] ds + \\ + \exp[-r(T - t_0)] qx(T)^{1/2}. \end{aligned} \quad (8.1.6)$$

Пусть $[\phi_1^{(t_0)*}(t, x), \phi_2^{(t_0)*}(t, x)]$, $t \in [t_0, T]$ — равновесие по Нэшу в игре $\Gamma(x_0, T - t_0)$ и $V^{(t_0)i}(t, x) : [t_0, T] \times R^n \rightarrow R$ — функции значения игроков $i \in \{1, 2\}$. Тогда они должны удовлетворять уравнениям Айзекса–Беллмана (теорема п. 6.4.2) вида

$$\begin{aligned} -V_t^{(t_0)i}(t, x) &= \max_{u_i} \left\{ \left[u_i(t)^{1/2} - \frac{c_i}{x^{1/2}} u_i(t) \right] \exp[-r(t - t_0)] + \right. \\ &\quad \left. + V_x^{(t_0)i}(t, x) \left[ax^{1/2} - bx - u_i(t) - \phi_j^{(t_0)*}(t, x) \right] \right\}, \\ V^{(t_0)i}(T, x) &= \exp[-r(T - t_0)] qx(T)^{1/2}, \\ i &\in \{1, 2\}, \quad j \in \{1, 2\}, \quad j \neq i. \end{aligned} \quad (8.1.7)$$

Максимизируя правую часть равенства (8.1.7), получаем

$$\phi_i^{(t_0)*}(t, x) = \frac{x}{4 \left[c_i + V_x^{(t_0)i} \exp[r(t - t_0)] x^{1/2} \right]^2}. \quad (8.1.8)$$

Утверждение. *Функция значения игрока $i \in \{1, 2\}$ (выигрыш в ситуации равновесия по Нэшу) в игре $\Gamma(x_0, T - t_0)$ равна*

$$V^{(t_0)i}(t, x) = \exp[-r(t - t_0)] \left[A_i(t) x^{1/2} + B_i(t) \right].$$

Здесь $i, j \in \{1, 2\}$ и $i \neq j$, $A_i(t)$, $B_i(t)$, $A_j(t)$ и $B_j(t)$ удовлетворяют уравнениям

$$\begin{aligned} \dot{A}_i(t) &= \left[r + \frac{b}{2} \right] A_i(t) - \frac{1}{2[c_i + A_i(t)/2]} + \frac{c_i}{4[c_i + A_i(t)/2]^2} \\ &\quad + \frac{A_i(t)}{8[c_i + A_i(t)/2]^2} + \frac{A_i(t)}{8[c_j + A_j(t)/2]^2}, \\ \dot{B}_i(t) &= rB_i(t) - \frac{a}{2} A_i(t) \quad \text{и} \quad A_i(T) = q, \quad \text{и} \quad B_i(T) = 0. \end{aligned}$$

Доказательство получается подстановкой $\phi_1^{(t_0)*}(t, x)$ и $\phi_2^{(t_0)*}(t, x)$ в (8.1.7) и решением уравнения (8.1.7).

Используя теорему п. 8.1.1 можно получить выражения для равновесных по Нэшу стратегий в виде

$$\phi_1^{(t_0)*}(t, x) = \frac{x}{4[c_1 + A_1(t)/2]^2}, \quad \phi_2^{(t_0)*}(t, x) = \frac{x}{4[c_2 + A_2(t)/2]^2}. \quad (8.1.9)$$

8.1.4. Рассмотрим теперь подыгру $\Gamma(x_\tau, T - \tau)$ со структурой выигрыша (8.1.6) и динамикой (8.1.5), которая начинается в момент времени $\tau \in [t_0, T]$ из начального состояния $x_\tau \in X$. Из предыдущих рассмотрений следует, что функция значения $V^{(\tau)i}(t, x) : [\tau, T] \times R \rightarrow R$ при $i \in \{1, 2\}$ и $\tau \in [t_0, T]$ подыгры $\Gamma(x_\tau, T - \tau)$ может быть определена следующим образом.

Утверждение. *Функция значения игрока $i \in \{1, 2\}$ в подыгре $\Gamma(x_\tau, T - \tau)$ имеет вид*

$$V^{(\tau)i}(t, x) = \exp[-r(t - \tau)] \left[A_i(t) x^{1/2} + B_i(t) \right],$$

где $i, j \in \{1, 2\}$ и $i \neq j$, $A_i(t)$, $B_i(t)$, $A_j(t)$ и $B_j(t)$ определяются таким же образом, как и в утверждении п. 8.1.3.

Доказательство аналогично теореме п. 8.1.1.

Равновесные по Нэшу стратегии в подыгре $\Gamma(x_\tau, T - \tau)$ соответственно равны

$$\phi_1^{(\tau)*}(t, x) = \frac{x}{4[c_1 + A_1(t)/2]^2}, \quad \phi_2^{(\tau)*}(t, x) = \frac{x}{4[c_2 + A_2(t)/2]^2}. \quad (8.1.10)$$

Заметим, что условия замечания п. 8.1.2 выполняются.

8.1.5. Предположим теперь, что игроки решили кооперироваться. Обозначим через $\Gamma_c(x_0, T - t_0)$ кооперативный вариант игры $\Gamma(x_0, T - t_0)$, в которой игроки выбрали совместно некоторый принцип оптимальности. Соглашение о кооперации и о том, как распределить между собой общий суммарный выигрыш (выигрыши предполагаются трансферабельными), и составляет принцип оптимальности при кооперативном поведении. В частности, принцип оптимальности в кооперативной игре $\Gamma_c(x_0, T - t_0)$ включает в себя:

- 1) соглашение о кооперативных стратегиях (управлениях);
- 2) механизм распределения суммарного выигрыша между игроками.

Принцип оптимальности должен сохранять свою оптимальность вдоль кооперативной траектории $\{x_s^*\}_{s=t_0}^T$. Более того, групповая рациональность требует, чтобы игроки использовали оптимальную по Парето траекторию. В дополнение к этому принцип распределения выигрыша должен удовлетворять условию индивидуальной рациональности в том смысле, что ни один из игроков не мог бы получить больший выигрыш, действуя индивидуально.

Исследуем условие групповой рациональности на оптимальной траектории. Поскольку выигрыши трансферабельны, групповая рациональность сводится к максимизации игроками суммарного выигрыша. Рассмотрим кооперативную игру $\Gamma_c(x_0, T - t_0)$. Как мы уже говорили, игроки должны при кооперации действовать таким образом, чтобы максимизировать суммарный выигрыш

$$\begin{aligned} \max_{u_1, u_2} & \left\{ \int_{t_0}^T \sum_{j=1}^2 g^j[s, x(s), u_1(s), u_2(s)] \exp \left[- \int_{t_0}^s r(y) dy \right] ds + \right. \\ & \left. + \exp \left[- \int_{t_0}^T r(y) dy \right] \sum_{j=1}^2 q^j(x(T)) \right\} \end{aligned} \quad (8.1.11)$$

при условии (8.1.1).

8.1.6. Обозначим задачу оптимального управления (8.1.11) и (8.1.1) через $\Psi(x_0, T - t_0)$. Для решения этой задачи можно воспользоваться как теорией динамического программирования, так и принципом максимума Л.С. Понтрягина. Для простоты в данном случае применим технику динамического программирования. Используя теорему п. 6.1.1, получаем:

Теорема. Набор управлений $\left\{ \left[\psi_1^{(t_0)*}(t, x), \psi_2^{(t_0)*}(t, x) \right] \text{ при } t \in [t_0, T] \right\}$ образует оптимальное управление в задаче $\Psi(x_0, T - t_0)$, если существует непрерывно-дифференцируемая функция $W^{(t_0)}(t, x) : [t_0, T] \times R^m \rightarrow R$, удовлетворяющая уравнению Беллмана

$$-W_t^{(t_0)}(t, x) = \max_{u_1, u_2} \left\{ \sum_{j=1}^2 g^j[t, x, u_1, u_2] \exp \left[- \int_{t_0}^t r(y) dy \right] + W_x^{(t_0)} f[t, x, u_1, u_2] \right\},$$

при граничном условии

$$W^{(t_0)}(T, x) = \exp \left[- \int_{t_0}^T r(y) dy \right] \sum_{j=1}^2 q^j(x).$$

Пусть игроки используют кооперативные управления $[\psi_1^{(t_0)*}(t, x), \psi_2^{(t_0)*}(t, x)]$, при $t \in [t_0, T]$. При кооперации вопрос неединственности кооперативных управлений легко решается, поскольку все кооперативные управления обеспечивают один и тот же максимальный суммарный выигрыш, и с этой точки зрения безразлично, какое именно кооперативное управление будет на самом деле использовано. Подставляя выбранные кооперативные управления в уравнения (8.1.1), получаем уравнение для кооперативной траектории

$$\dot{x}(s) = f[s, x(s), \psi_1^{(t_0)*}(s, x(s)), \psi_2^{(t_0)*}(s, x(s))], \quad x(t_0) = x_0. \quad (8.1.12)$$

Пусть $x^*(t)$, $t \in [t_0, T]$ — решение уравнения (8.1.12). Оптимальная траектория $\{x^*(t)\}_{t=t_0}^T$ может быть записана в виде

$$x^*(t) = x_0 + \int_{t_0}^t f[s, x^*(s), \psi_1^{(t_0)*}(s, x^*(s)), \psi_2^{(t_0)*}(s, x^*(s))] ds. \quad (8.1.13)$$

Для удобства мы будем одновременно использовать обозначения $x^*(t)$ и x_t^* . Кооперативные программные управления имеют вид:

$$\left\{ \left[\psi_1^{(t_0)*}(t, x^*(t)), \psi_2^{(t_0)*}(t, x^*(t)) \right], \quad t \in [t_0, T] \right\}. \quad (8.1.14)$$

Рассмотрим кооперативную подыгру $\Gamma_c(x_\tau^*, T - \tau)$ с динамикой (8.1.1) и некооперативным выигрышем (8.1.2) с началом в момент времени $\tau \in [t_0, T]$ из состояния x_τ^* на кооперативной траектории. В момент τ групповая рациональность требует, чтобы игроки решали задачу

$$\begin{aligned} \max_{u_1, u_2} \left\{ \int_{\tau}^T \sum_{j=1}^2 g^j[s, x(s), u_1(s), u_2(s)] \exp \left[- \int_{\tau}^s r(y) dy \right] ds + \right. \\ \left. + \exp \left[- \int_{\tau}^T r(y) dy \right] \sum_{j=1}^2 q^j(x(T)) \right\} \end{aligned} \quad (8.1.15)$$

при условии

$$\dot{x}(s) = f[s, x(s), u_1(s), u_2(s)], \quad x(\tau) = x_\tau^*.$$

8.1.7. Замечание 2. Из теории динамического программирования следует, что кооперативное управление в игре $\Gamma_c(x_\tau^*, T - \tau)$ на отрезке времени $[\tau, T]$ совпадает с кооперативным управлением в игре $\Gamma_c(x_0, T - t_0)$ на том же отрезке времени. Поэтому оптимальная траектория в игре $\Gamma_c(x_\tau^*, T - \tau)$ является отрезком оптимальной траектории игры $\Gamma_c(x_0, T - t_0)$ на соответствующем промежутке времени.

Кроме того, вдоль оптимальной траектории $\{x^*(s)\}_{s=t_0}^T$ имеет место следующее тождество:

$$\begin{aligned}
 W^{(t_0)}(t, x_t^*) &= \\
 &= \left\{ \int_t^T \sum_{j=1}^2 g^j \left[s, x(s), \psi_1^{(t_0)*}(s, x^*(s)), \psi_2^{(t_0)*}(s, x^*(s)) \right] \times \exp \left[- \int_{t_0}^s r(y) dy \right] ds + \right. \\
 &\quad \left. + \exp \left[- \int_{t_0}^T r(y) dy \right] \sum_{j=1}^2 q^j(x^*(T)) \right\} = \\
 &= \exp \left[- \int_{t_0}^t r(y) dy \right] \times \\
 &\times \left\{ \int_t^T \sum_{j=1}^2 g^j \left[s, x(s), \psi_1^{(\tau)*}(s, x^*(s)), \psi_2^{(\tau)*}(s, x^*(s)) \right] \times \exp \left[- \int_t^s r(y) dy \right] ds + \right. \\
 &\quad \left. + \exp \left[- \int_t^T r(y) dy \right] \sum_{j=1}^2 q^j(x^*(T)) \right\} = \\
 &= \exp \left[- \int_{t_0}^\tau r(y) dy \right] W^{(\tau)}(t, x_t^*), \quad \tau \in [t_0, T], \quad t \geq \tau.
 \end{aligned}$$

8.1.8. Пример 2. Рассмотрим задачу оптимального управления $\Psi(x_0, T - t_0)$ максимизации суммы выигрышей игроков 1 и 2 из примера 1 п. 8.1.3.

$$\begin{aligned}
 &\int_{t_0}^T \left(\left[u_1(s)^{1/2} - \frac{c_1}{x(s)^{1/2}} u_1(s) \right] + \left[u_2(s)^{1/2} - \frac{c_2}{x(s)^{1/2}} u_2(s) \right] \right) \exp[-r(t - t_0)] ds + \\
 &+ 2 \exp[-r(T - t_0)] q x(T)^{1/2},
 \end{aligned} \tag{8.1.16}$$

при условии (8.1.5).

Пусть $[\psi_1^{(t_0)*}(t, x), \psi_2^{(t_0)*}(t, x)]$ — управления, решающие задачу максимизации $\Psi(x_0, T - t_0)$, и $W^{(t_0)}(t, x) : [t_0, T] \times R^n \rightarrow R$ — функция Беллмана, удовлетворяющая уравнению (см. теорему п. 8.1.6)

$$\begin{aligned}
 &-W_t^{(t_0)}(t, x) = \\
 &= \max_{u_1, u_2} \left\{ \left(\left[u_1^{1/2} - \frac{c_1}{x^{1/2}} u_1 \right] + \left[u_2^{1/2} - \frac{c_2}{x^{1/2}} u_2 \right] \right) \exp[-r(t - t_0)] + \right. \\
 &\quad \left. + W_x^{(t_0)}(t, x) \left[a x^{1/2} - b x - u_1 - u_2 \right] \right\}, \\
 &W^{(t_0)}(T, x) = 2 \exp[-r(T - t_0)] q x^{1/2}.
 \end{aligned} \tag{8.1.17}$$

Проведя операцию максимизации в правой части (8.1.17), получаем выражения для оптимальных управлений:

$$\begin{aligned}\psi_1^{(t_0)*}(t, x) &= \frac{x}{4 \left[c_1 + W_x^{(t_0)} \exp[r(t - t_0)] x^{1/2} \right]^2}, \\ \psi_2^{(t_0)*}(t, x) &= \frac{x}{4 \left[c_2 + W_x^{(t_0)} \exp[r(t - t_0)] x^{1/2} \right]^2}.\end{aligned}$$

Подставляя $\psi_1^{(t_0)*}(t, x)$ и $\psi_2^{(t_0)*}(t, x)$ в (8.1.17), получаем явное выражение для функции Беллмана:

$$\begin{aligned}W^{(t_0)}(t, x) &= \exp[-r(t - t_0)] \left[\hat{A}(t) x^{1/2} + \hat{B}(t) \right], \\ \dot{\hat{A}}(t) &= \left[r + \frac{b}{2} \right] \hat{A}(t) - \frac{1}{2 \left[c_1 + \hat{A}(t)/2 \right]} - \frac{1}{2 \left[c_2 + \hat{A}(t)/2 \right]} + \\ &+ \frac{c_1}{4 \left[c_1 + \hat{A}(t)/2 \right]^2} + \frac{c_2}{4 \left[c_2 + \hat{A}(t)/2 \right]^2} + \\ &+ \frac{\hat{A}(t)}{8 \left[c_1 + \hat{A}(t)/2 \right]^2} + \frac{\hat{A}(t)}{8 \left[c_2 + \hat{A}(t)/2 \right]^2}, \\ \dot{\hat{B}}(t) &= r\hat{B}(t) - \frac{a}{2}\hat{A}(t), \quad \hat{A}(T) = 2q, \quad \hat{B}(T) = 0.\end{aligned}$$

Оптимальные кооперативные управления могут быть представлены в виде:

$$\psi_1^{(t_0)*}(t, x) = \frac{x}{4 \left[c_1 + \hat{A}(t)/2 \right]^2}, \quad \psi_2^{(t_0)*}(t, x) = \frac{x}{4 \left[c_2 + \hat{A}(t)/2 \right]^2}.$$

Подставляя эти управления в (8.1.5), получаем уравнение для кооперативной траектории

$$\begin{aligned}\dot{x}(s) &= ax(s)^{1/2} - bx(s) - \frac{x(s)}{4 \left[c_1 + \hat{A}(s)/2 \right]^2} - \frac{x(s)}{4 \left[c_2 + \hat{A}(s)/2 \right]^2}, \\ x(t_0) &= x_0.\end{aligned}\tag{8.1.18}$$

Решая (8.1.18), получаем кооперативную траекторию $\Gamma_c(x_0, T - t_0)$ в виде

$$x^*(s) = \varpi(t_0, s)^2 \left[x_0^{1/2} + \int_{t_0}^s \varpi^{-1}(t_0, t) H_1 dt \right]^2, \quad s \in [t_0, T],\tag{8.1.19}$$

где

$$\begin{aligned}\varpi(t_0, s) &= \exp \left[\int_{t_0}^s H_2(\tau) d\tau \right], \quad H_1 = \frac{1}{2}a, \\ H_2(s) &= - \left[\frac{1}{2}b + \frac{1}{8 \left[c_1 + \hat{A}(s)/2 \right]^2} + \frac{1}{8 \left[c_2 + \hat{A}(s)/2 \right]^2} \right].\end{aligned}$$

Кооперативные программные управления в игре $\Gamma_c(x_0, T - t_0)$ на отрезке времени $[t_0, T]$ имеют вид:

$$\psi_1^{(t_0)*}(t, x_t^*) = \frac{x_t^*}{4 [c_1 + \hat{A}(t)/2]^2}, \quad \psi_2^{(t_0)*}(t, x_t^*) = \frac{x_t^*}{4 [c_2 + \hat{A}(t)/2]^2}. \quad (8.1.20)$$

Заметим, что при этом выполнены условия замечания 1 п. 8.1.3.

В подыгре $\Gamma(x_\tau^*, T - \tau)$ программные управления вдоль оптимальной траектории имеют вид

$$\psi_1^{(\tau)*}(t, x_t^*) = \frac{x_t^*}{4 [c_1 + \hat{A}(t)/2]^2}, \quad \text{и} \quad \psi_2^{(\tau)*}(t, x_t^*) = \frac{x_t^*}{4 [c_2 + \hat{A}(t)/2]^2}.$$

Оптимальная кооперативная траектория определяется по формуле:

$$x^*(s) = \varpi(\tau, s)^2 \left[(x_\tau^*)^2 + \int_\tau^s \varpi^{-1}(\tau, t) H_1 dt \right]^2, \quad s \in [\tau, T], \quad (8.1.21)$$

где

$$\varpi(\tau, s) = \exp \left[\int_\tau^s H_2(\zeta) d\zeta \right], \quad H_1 = \frac{1}{2}a,$$

$$H_2(s) = - \left[\frac{1}{2}b + \frac{1}{8 [c_1 + \hat{A}(s)/2]^2} + \frac{1}{8 [c_2 + \hat{A}(s)/2]^2} \right].$$

Выражения (8.1.21) и (8.1.19) совпадают на промежутке $[\tau, T]$, т. е. $x^*(s)$ в (8.1.21) является подмножеством множества из (8.1.19). Таким образом, вдоль кооперативной траектории $\{x^*(s)\}_{s=t_0}^T$ групповая рациональность сохраняется в каждый момент времени $t \in [t_0, T]$.

8.1.9. Индивидуальная рациональность. Предположим, что в момент t_0 в состоянии x_0 выбранный принцип оптимальности предписывает выбор дележа $\xi(x_0, T - t_0) = [\xi^1(x_0, T - t_0), \xi^2(x_0, T - t_0)]$. Это означает, что игроки согласились на такой дележ суммарного выигрыша, при котором выигрыш игрока i на отрезке времени $[t_0, T]$ равен $\xi^i(x_0, T - t_0)$.

Индивидуальная рациональность дележа означает, что должно иметь место

$$\xi^i(x_0, T - t_0) \geq V^{(t_0)i}(t_0, x_0), \quad i \in \{1, 2\}.$$

Используя тот же принцип оптимальности в момент τ в состоянии x_τ^* на кооперативной траектории, принцип оптимальности предписывает выбор дележа

$$\xi(x_\tau^*, T - \tau) = [\xi^1(x_\tau^*, T - \tau), \xi^2(x_\tau^*, T - \tau)].$$

Это означает, что игроки согласились на такой дележ суммарного выигрыша (в подыгре с началом в момент τ), при котором выигрыш игрока i на отрезке времени $[\tau, T]$ равен $\xi^i(x_\tau^*, T - \tau)$. Индивидуальная рациональность будет иметь место, если

$$\xi^i(x_\tau^*, T - \tau) \geq V^{(\tau)i}(\tau, x_\tau^*), \quad i \in \{1, 2\}.$$

В динамической постановке индивидуальная рациональность должна выполняться в каждый момент времени $\tau \in [t_0, T]$ на оптимальной траектории $\{x^*(t)\}_{t=t_0}^T$.

8.1.10. Как было нами ранее указано в §§ 7.4–7.5, важнейшим дополнительным требованием к решению динамических кооперативных игр является динамическая устойчивость или состоятельность во времени. Свойство динамической устойчивости решения заключается в том, что когда игра развивается вдоль кооперативной траектории, в каждый момент времени игроки руководствуются одним и тем же принципом оптимальности и, следовательно, не имеют оснований для отклонения от принципа оптимальности, выбранного в начале игры. В этом разделе мы исследуем свойство динамической устойчивости или временной состоятельности в неантагонистических кооперативных играх двух лиц с дисконтированным выигрышем.

Рассмотрим кооперативную игру $\Gamma_c(x_0, T - t_0)$, в которой игроки решили максимизировать совместными действиями суммарный выигрыш и применить некоторый механизм дележа этого выигрыша между собой. Для достижения максимального суммарного выигрыша игроки применяют кооперативные управления $[\psi_1^{(t_0)*}(t, x), \psi_2^{(t_0)*}(t, x)]$, полученные в теореме п. 8.1.6. При этом кооперативная траектория $\{x^*(s)\}_{s=t_0}^T$ определяется уравнением (8.1.13).

В момент t_0 в начальном состоянии x_0 компонента дележа $\xi^{(t_0)i}(t_0, x_0)$ представляет собой часть максимального суммарного выигрыша (получаемого на отрезке времени $[t_0, T]$), на которую игрок i ориентирован согласно выбранному и согласованному принципу оптимальности.

Рассмотрим теперь кооперативную подыгру $\Gamma_c(x_\tau^*, T - \tau)$, начинающуюся в промежуточный момент времени $\tau \in [t_0, T]$ из начального состояния x_τ^* , в которой используется тот же принцип оптимальности что и в основной игре. Пусть $\xi^{(\tau)i}(\tau, x_\tau^*)$ — компонента дележа (часть кооперативного выигрыша) игрока i на промежутке времени $[\tau, T]$ в этой подыгре.

Векторы $\xi^{(\tau)}(\tau, x_\tau^*) = [\xi^{(\tau)1}(\tau, x_\tau^*), \xi^{(\tau)2}(\tau, x_\tau^*)]$, $\tau \in [t_0, T]$ будут допустимыми дележами, если выполнены следующие условия.

Определение. Вектор $\xi^{(\tau)}(\tau, x_\tau^*)$ является дележом кооперативной игры $\Gamma_c(x_\tau^*, T - \tau)$ при $\tau \in [t_0, T]$, если выполнены следующие условия:

$$1) \xi^{(\tau)}(\tau, x_\tau^*) = [\xi^{(\tau)1}(\tau, x_\tau^*), \xi^{(\tau)2}(\tau, x_\tau^*)] - \text{парето-оптимальный дележ};$$

$$2) \xi^{(\tau)i}(\tau, x_\tau^*) \geq V^{(\tau)i}(\tau, x_\tau^*), i \in \{1, 2\}.$$

8.1.11. Следуя [Petrosjan, 1997] и [Yeung, Petrosjan, 2004], построим процедуру распределения дележа (ПРД), чтобы согласованный в начале принцип оптимальности мог бы реально реализоваться в игре. Пусть $B^\tau(s) = [B_1^\tau(s), B_2^\tau(s)]$ — мгновенные выигрыши (плотности выигрышей) при кооперации в момент времени $s \in [\tau, T]$ в кооперативной игре $\Gamma_c(x_\tau^*, T - \tau)$. Иными словами, игрок $i \in \{1, 2\}$ получает выигрыш $B_i^\tau(s)$ в момент времени s . Терминальный выигрыш $q^i(x_T^*)$ получается игроком i в момент T .

В частности, $B_i^\tau(s)$ и $q^i(x_T^*)$ образуют ПРД в игре $\Gamma_c(x_\tau, T - \tau)$, если величина $\xi^{(\tau)i}(\tau, x_\tau^*)$ будет равна

$$\left\{ \left(\int_\tau^T B_i^\tau(s) \exp \left[- \int_\tau^s r(y) dy \right] ds + \right. \right. \\ \left. \left. + q^i(x_T^*) \exp \left[- \int_\tau^T r(y) dy \right] \right) \middle| x(\tau) = x_\tau^* \right\}, \quad (8.1.22) \\ i \in \{1, 2\}, \quad \tau \in [t_0, T].$$

Определим для $i \in \{1, 2\}$ и $t \in [\tau, T]$ вспомогательную величину $\xi^{(\tau)i}(t, x_t^*)$, равную

$$\left\{ \left(\int_t^T B_i^\tau(s) \exp \left[- \int_\tau^s r(y) dy \right] ds + \right. \right. \\ \left. \left. + q^i(x_T^*) \exp \left[- \int_\tau^T r(y) dy \right] \right) \middle| x(t) = x_t^* \right\}, \quad (8.1.23)$$

для обозначения оценки (с учетом дисконтирования) кооперативного выигрыша игрока i на отрезке времени $[t, T]$ из состояния x_t^* в момент времени $t \in [\tau, T]$, когда игрок находится в начале подыгры, стартующей в момент τ из состояния x_τ^* .

Определение. Вектор $\xi^{(\tau)}(\tau, x_\tau^*) = [\xi^{(\tau)1}(\tau, x_\tau^*), \xi^{(\tau)2}(\tau, x_\tau^*)]$, определенный в (8.1.22) и (8.1.23), есть состоятельный во времени дележ в подыгре $\Gamma_c(x_\tau^*, T - \tau)$ при $\tau \in [t_0, T]$, если $\xi^{(\tau)}(\tau, x_\tau^*)$ — оптимальный по Парето дележ, при $t \in [\tau, T]$ имеет место неравенство $\xi^{(\tau)i}(t, x_t^*) \geq V^{(t)i}(t, x_t^*)$ для $i \in \{1, 2\}$, $t \in [\tau, T]$, и выполнено условие

$$\xi^{(\tau)i}(t, x_t^*) = \exp \left[- \int_\tau^t r(y) dy \right] \xi^{(t)i}(t, x_t^*), \\ \text{для } \tau \leq t \leq T, \quad i \in \{1, 2\}.$$

Определение п. 8.1.11 означает, что часть выигрыша, который игрок в начальный момент рассчитывает получить в подыгре при использовании некоторого выбранного заранее принципа оптимальности (с точностью до процедуры дисконтирования), совпадает с выигрышем, который он реально может получить, попав в эту подыгру при использовании того же принципа оптимальности, но уже примененного к этой подыгре. Кроме того, групповая и индивидуальная рациональность выполняются на всем отрезке времени на котором происходит игра.

Как отмечалось в [Jorgensen, Zassour, 2002], условия, гарантирующие временную состоятельность кооперативных решений, могут быть достаточно сложными и аналитически трудно представимыми. Для того чтобы достичь временной состоятельности или динамической устойчивости решения $\xi^{(\tau)}(\tau, x_\tau^*)$, удовлетворяющего определению п. 8.1.11, необходимо определить дополнительный механизм выплат на игровом отрезке времени, то есть механизм реализации дележей. В работе [Yeung, Petrosjan, 2004] получена общая теорема, с помощью которой такой механизм выплат (ПРД) распределения дележа во времени может быть получен. Ниже мы используем этот подход для дифференциальной игры.

Процедура распределения дележа (ПРД) в кооперативной игре, определенная в (8.1.22) и (8.1.23), должна быть сформулирована таким образом, чтобы первоначально согласованный дележ мог бы быть реализован.

Для того чтобы условия, сформулированные в определении п. 8.1.11 выполнялись, необходимо потребовать, чтобы $B_i^\tau(s) = B_i^t(s)$ при $i \in \{1, 2\}$ и $\tau \in [t_0, T]$ и $t \in [t_0, T]$ и $\tau \neq t$. Используя обозначение $B_i^\tau(s) = B_i^t(s) = B_i(s)$ и определение п. 8.1.11, ПРД для состоятельного во времени (динамически устойчивого) дележа должна удовлетворять условиям:

1.
$$\sum_{i=1}^2 B_i(s) = \sum_{j=1}^2 g^j \left[s, x_s^*, \psi_1^{(\tau)*}(s, x_s^*), \psi_2^{(\tau)*}(s, x_s^*) \right], \quad s \in [t_0, T];$$
2.
$$\int_\tau^T B_i(s) \exp \left[- \int_\tau^s r(y) dy \right] ds + q^i(x^*(T)) \exp \left[- \int_\tau^T r(y) dy \right] \geq$$

$$\geq V^{(\tau)i}(\tau, x_\tau^*), \quad i \in \{1, 2\}, \quad \tau \in [t_0, T];$$
3.
$$\xi^{(\tau)i}(\tau, x_\tau^*) = \int_\tau^{\tau+\Delta t} B_i(s) \exp \left[- \int_\tau^s r(y) dy \right] ds +$$

$$+ \exp \left[- \int_\tau^{\tau+\Delta t} r(y) dy \right] \xi^{(\tau+\Delta t)i}(\tau + \Delta t, x_\tau^* + \Delta x_\tau^*),$$

$$\tau \in [t_0, T], \quad i \in \{1, 2\};$$

где

$$\Delta x_\tau^* = f \left[\tau, x_\tau^*, \psi_1^{(\tau)*}(\tau, x_\tau^*), \psi_2^{(\tau)*}(\tau, x_\tau^*) \right] \Delta t + o(\Delta t).$$

8.1.12 Рассмотрим следующее условие, касающееся $\xi^{(\tau)}(t, x_t^*)$ для $\tau \in [t_0, T]$ и $t \in [\tau, T]$.

Условие. При $i \in \{1, 2\}$, $t \geq \tau$ и $\tau \in [t_0, T]$ величины $\xi^{(\tau)i}(t, x_t^*)$ непрерывно дифференцируемы по t и x_t^* .

Если величины $\xi^{(\tau)}(t, x_t^*)$ при $\tau \in [t_0, T]$ удовлетворяют условию п. 8.1.12, то можно получить следующее соотношение

$$\begin{aligned} & \int_\tau^{\tau+\Delta t} B_i(s) \exp \left[- \int_\tau^s r(y) dy \right] ds = \\ & = \xi^{(\tau)i}(\tau, x_\tau^*) - \exp \left[- \int_\tau^{\tau+\Delta t} r(y) dy \right] \xi^{(\tau+\Delta t)i}(\tau + \Delta t, x_\tau^* + \Delta x_\tau^*) = \\ & = \xi^{(\tau)i}(\tau, x_\tau^*) - \xi^{(\tau)i}(\tau + \Delta t, x_\tau^* + \Delta x_\tau^*), \end{aligned}$$

$$\text{для всех } \tau \in [t_0, T] \text{ и } i \in \{1, 2\}. \quad (8.1.24)$$

При $\Delta t \rightarrow 0$ из (8.1.24) получаем

$$\begin{aligned} B_i(\tau) \Delta t = & - \left[\xi_t^{(\tau)i}(t, x_t^*) \Big|_{t=\tau} \right] \Delta t - \\ & - \left[\xi_{x_t^*}^{(\tau)i}(t, x_t^*) \Big|_{t=\tau} \right] f \left[\tau, x_\tau^*, \psi_1^{(\tau)*}(\tau, x_\tau^*), \psi_2^{(\tau)*}(\tau, x_\tau^*) \right] \Delta t - \\ & - o(\Delta t). \end{aligned} \quad (8.1.25)$$

Таким образом, получаем следующую теорему.

Теорема. Если оптимальный дележ $\xi^{(\tau)i}(\tau, x_\tau^*)$, при $i \in \{1, 2\}$ и $\tau \in [t_0, T]$ удовлетворяет условиям определения п. 8.1.11 и условию п. 8.1.12 с терминальной выплатой $q^i(x_T^*)$ в момент T и мгновенными выплатами в моменты времени $\tau \in [t_0, T]$:

$$B_i(\tau) = - \left[\xi_t^{(\tau)i}(t, x_t^*) \Big|_{t=\tau} \right] - \\ - \left[\xi_{x_t^*}^{(\tau)i}(t, x_t^*) \Big|_{t=\tau} \right] f \left[\tau, x_\tau^*, \psi_1^{(\tau)*}(\tau, x_\tau^*), \psi_2^{(\tau)*}(\tau, x_\tau^*) \right], \\ \text{при } i \in \{1, 2\},$$

то он представляет состоятельное во времени решение игры $\Gamma_c(x_0, T - t_0)$.

8.1.13. Исследуем теперь динамически устойчивые (состоятельные во времени) решения для конкретных принципов оптимальности.

Рассмотрим кооперативную игру $\Gamma_c(x_0, T - t_0)$, в которой игроки максимизируют суммарный выигрыш и делят его в соответствии с арбитражной схемой Нэша. Эта схема также совпадает с вектором Шепли для кооперативной игры двух лиц. Известно, что при таком подходе дележ имеет вид, представленный в следующем утверждении.

Утверждение. В игре $\Gamma_c(x_0, T - t_0)$ компонента дележа

$$\xi^{(t_0)i}(t_0, x_0) = V^{(t_0)i}(t_0, x_0) + \frac{1}{2} \left[W^{(t_0)}(t_0, x_0) - \sum_{j=1}^2 V^{(t_0)j}(t_0, x_0) \right],$$

предписывается игроку $i \in \{1, 2\}$ выплаты согласно арбитражной схеме Нэша, и в подыгре $\Gamma_c(x_\tau^*, T - \tau)$, при $\tau \in (t_0, T]$ компонента дележа

$$\xi^{(\tau)i}(\tau, x_\tau^*) = V^{(\tau)i}(\tau, x_\tau^*) + \frac{1}{2} \left[W^{(\tau)}(\tau, x_\tau^*) - \sum_{j=1}^2 V^{(\tau)j}(\tau, x_\tau^*) \right],$$

предписывается игроку $i \in \{1, 2\}$ выплаты согласно той же схеме.

Заметим, что при таком дележе каждый из игроков получает свой некооперативный выигрыш (выигрыш в равновесии по Нэшу) плюс половину дополнительного выигрыша, получаемого в результате кооперации на отрезке времени $[\tau, T]$, $\tau \in [t_0, T]$.

Можно убедиться, что $\xi^{(\tau)i}(\tau, x_\tau^*)$ удовлетворяет определению п. 8.1.11. Более того, используя замечание 1 п. 8.1.2 и замечание 2 п. 8.1.7, получаем

$$\xi^{(t)i}(t, x_t^*) = \exp \left[\int_\tau^t r(y) dy \right] \times \\ \times \left\{ V^{(\tau)i}(t, x_t^*) + \frac{1}{2} \left[W^{(\tau)}(t, x_t^*) - \sum_{j=1}^2 V^{(\tau)j}(t, x_t^*) \right] \right\} = \\ = \exp \left[\int_\tau^t r(y) dy \right] \xi^{(\tau)i}(\tau, x_\tau^*), \\ \text{при } t_0 \leq \tau \leq t. \tag{8.1.26}$$

Таким образом $\xi^{(\tau)i}(\tau, x_\tau^*)$, как это следует из утверждения п. 8.1.13, является состоятельным во времени дележом в кооперативной игре $\Gamma_c(x_0, T - t_0)$. Используя теорему п. 8.1.6, получаем результат.

Следствие. ПРД с терминальной выплатой $q^i(x(T))$ в момент времени T и мгновенными выплатами (плотностью выплат) в моменты времени $\tau \in [t_0, T]$:

$$\begin{aligned}
 B_i(\tau) = & \frac{-1}{2} \left[\left[V_t^{(\tau)i}(t, x_t) \right]_{t=\tau} \right] + \\
 & + \left[V_{x_t}^{(\tau)i}(t, x_t) \right]_{t=\tau} f \left[\tau, x_\tau, \psi_1^{(\tau)*}(\tau, x_\tau), \psi_2^{(\tau)*}(\tau, x_\tau) \right] - \\
 & - \frac{1}{2} \left[\left[W_t^{(\tau)}(t, x_t) \right]_{t=\tau} \right] + \\
 & + \left[W_{x_t}^{(\tau)}(t, x_t) \right]_{t=\tau} f \left[\tau, x_\tau, \psi_1^{(\tau)*}(\tau, x_\tau), \psi_2^{(\tau)*}(\tau, x_\tau) \right] + \\
 & + \frac{1}{2} \left[\left[V_t^{(\tau)j}(t, x_t) \right]_{t=\tau} \right] + \\
 & + \left[V_{x_t}^{(\tau)j}(t, x_t) \right]_{t=\tau} f \left[\tau, x_\tau, \psi_1^{(\tau)*}(\tau, x_\tau), \psi_2^{(\tau)*}(\tau, x_\tau) \right], \\
 & \text{при } i, j \in \{1, 2\} \text{ и } i \neq j,
 \end{aligned} \tag{8.1.27}$$

представляет собой состоятельное во времени (динамически устойчивое) решение кооперативной игры $\Gamma_c(x_0, T - t_0)$, в которой игроки согласились делить свой кооперативный выигрыш в соответствии с утверждением п. 8.1.13.

8.1.14. Пример 3. Построим состоятельную во времени ПРД в игре добычи полезного ресурса, когда игроки договорились делить кооперативный выигрыш в соответствии с утверждением п. 8.1.13. Поэтому рассмотрим пример 1 п. 8.1.13, в котором две добывающие фирмы согласились максимизировать суммарный выигрыш и разделить его между собой согласно утверждению п. 8.1.13.

Используя результаты, полученные в примерах 1 п. 8.1.3, 2 п. 8.1.8 и теореме п. 8.1.12, мы получаем следующий результат.

Следствие. ПРД с терминальным выигрышем $q^i(x(T))$ в момент T и мгновенными выплатами (плотностью выплат) в моменты $\tau \in [t_0, T]$ имеет вид:

$$\begin{aligned}
 B_i(\tau) = & \frac{-1}{2} \left\{ \left(\left[\dot{A}_i(\tau) (x_\tau^*)^{1/2} + \dot{B}_i(\tau) \right] + r \left[A_i(\tau) (x_\tau^*)^{1/2} + B_i(\tau) \right] \right) + \right. \\
 & + \left[\frac{1}{2} A_i(\tau) (x_\tau^*)^{-1/2} \right] \left[a (x_\tau^*)^{1/2} - b x_\tau^* - \frac{x_\tau^*}{4 [c_i + \hat{A}(\tau)/2]^2} - \frac{x_\tau^*}{4 [c_j + \hat{A}(\tau)/2]^2} \right] \left. \right\} - \\
 & - \frac{1}{2} \left\{ \left(\left[\dot{\hat{A}}(\tau) (x_\tau^*)^{1/2} + \dot{\hat{B}}(\tau) \right] + r \left[\hat{A}(\tau) (x_\tau^*)^{1/2} + \hat{B}(\tau) \right] \right) + \right. \\
 & + \left[\frac{1}{2} \hat{A}(\tau) (x_\tau^*)^{-1/2} \right] \left[a (x_\tau^*)^{1/2} - b x_\tau^* - \frac{x_\tau^*}{4 [c_i + \hat{A}(\tau)/2]^2} - \frac{x_\tau^*}{4 [c_j + \hat{A}(\tau)/2]^2} \right] \left. \right\} + \\
 & + \frac{1}{2} \left\{ \left(\left[\dot{A}_j(\tau) (x_\tau^*)^{1/2} + \dot{B}_j(\tau) \right] + r \left[A_j(\tau) (x_\tau^*)^{1/2} + B_j(\tau) \right] \right) + \right. \\
 & + \left[\frac{1}{2} A_j(\tau) (x_\tau^*)^{-1/2} \right] \left[a (x_\tau^*)^{1/2} - b x_\tau^* - \frac{x_\tau^*}{4 [c_i + \hat{A}(\tau)/2]^2} - \frac{x_\tau^*}{4 [c_j + \hat{A}(\tau)/2]^2} \right] \left. \right\}, \\
 & \text{при } i, j \in \{1, 2\} \text{ и } i \neq j,
 \end{aligned}$$

и является состоятельным во времени решением кооперативной игры $\bar{\Gamma}_c(x_0, T - t_0)$, в которой игроки решили делить кооперативный выигрыш в соответствии с утверждением § 8.1.

§ 8.2. Кооперативные игры с бесконечной продолжительностью

8.2.1. Во многих конкретных задачах продолжительность игры T либо слишком велика, либо неизвестна участникам. Например, сегодняшняя стоимость фирмы определяется дисконтированной суммой ее будущих доходов. Никто не знает когда фирма прекратит существовать. Как справедливо замечено в работе [Dockner, Jorgensen, 2000], в этом случае предположение $T = \infty$ может быть хорошим приближением. Важнейшими примерами задач такого рода являются добыча возобновляемых ресурсов, природоохранный менеджмент и оценка корпоративной собственности.

Рассмотрим неантагонистическую дифференциальную игру с выигрышами

$$\int_{t_0}^{\infty} g^i[x(s), u_1(s), u_2(s)] \exp[-r(s - t_0)] ds, \quad i \in \{1, 2\} \quad (8.2.1)$$

и динамикой

$$\dot{x}(s) = f[x(s), u_1(s), u_2(s)], \quad x(t_0) = x_0. \quad (8.2.2)$$

Поскольку время s не присутствует явно в $g^i[x(s), u_1(s), u_2(s)]$, выигрышах и уравнениях динамики, игра (8.2.1)–(8.2.2) является автономной. Рассмотрим подыгру $\Gamma(x)$:

$$\max_{u_i} \int_t^{\infty} g^i[x(s), u_1(s), u_2(s)] \exp[-r(s - t)] ds, \quad i \in \{1, 2\}$$

при условии

$$\dot{x}(s) = f[x(s), u_1(s), u_2(s)], \quad x(t) = x.$$

Автономная задача $\Gamma(x)$ с бесконечной продолжительностью не зависит от t , а зависит лишь от x .

Используя теорему п. 6.6.2 можно охарактеризовать некооперативное позиционное равновесие по Нэшу.

Теорема. Пара стратегий $\{\phi_1^*(x), \phi_2^*(x)\}$ образует позиционное равновесие по Нэшу в игре $\Gamma(x)$, если существуют непрерывно-дифференцируемые функции $V^1(x) : R^m \rightarrow R$ и $V^2(x) : R^m \rightarrow R$, удовлетворяющие следующей системе уравнений в частных производных:

$$rV^i(x) = \max_{u_i} \{g^i[x, u_i, \phi_j^*(x)] + V_x^i(x) f[x, u_i, \phi_j^*(x)]\}, \\ i \in \{1, 2\}, \quad j \in \{1, 2\}, \quad j \neq i.$$

Здесь

$$V^i(x) = \int_t^\infty g^i[x(s), \phi_1^*(s), \phi_2^*(s)] \exp[-r(s-t)] ds$$

есть выигрыш игрока i в ситуации позиционного равновесия в подыгре, начинающейся в момент $t \in [t_0, \infty]$ из состояния x .

8.2.2. Рассмотрим теперь случай, когда игроки решили кооперироваться. Пусть $\Gamma_c(x)$ — кооперативная игра, определенная на структуре игры $\Gamma(x)$ из начального состояния x . Предположим, что игроки предварительно договорились о максимизации суммарного выигрыша, то есть

$$\max_{u_1, u_2} \left\{ \int_t^\infty \sum_{j=1}^2 g^j[x(s), u_1(s), u_2(s)] \exp[-r(s-t)] ds \right\}, \quad (8.2.3)$$

$$\dot{x}(s) = f[x(s), u_1(s), u_2(s)], \quad x(t) = x.$$

Следуя теореме п. 6.1.1, получаем результат.

Теорема. Набор управлений $\{[\psi_1^*(x), \psi_2^*(x)]\}$ решает задачу (8.2.3), связанную с игрой $\Gamma_c(x)$, если существует непрерывно-дифференцируемая функция $W(x) : R^m \rightarrow R$, удовлетворяющая уравнению Беллмана для задачи с бесконечной продолжительностью

$$rW(x) = \max_{u_1, u_2} \left\{ \sum_{j=1}^2 g^j[x, u_1, u_2] + W_x f[x, u_1, u_2] \right\} = \\ = \sum_{j=1}^2 g^j[x, \psi_1^*, \psi_2^*] + W_x f[x, \psi_1^*, \psi_2^*].$$

Обозначим соответствующие кооперативные управления через $[\psi_1^*(x), \psi_2^*(x)]$. Подставляя эти управления в (8.2.2), получаем уравнение для оптимальной кооперативной траектории

$$\dot{x}(s) = f[x(s), \psi_1^*(x(s)), \psi_2^*(x(s))], \quad x(t) = x.$$

Если игра начинается в момент t_0 , $x(t_0) = x_0$, то имеем

$$\dot{x}(s) = f[x(s), \psi_1^*(x(s)), \psi_2^*(x(s))], \quad x(t_0) = x_0. \quad (8.2.4)$$

Пусть $x^*(t)$ — решение (8.2.4). Как и раньше будем использовать обозначения $x^*(t)$ и x_t^* в зависимости от обстоятельств.

8.2.3. Предположим, что в момент времени $t \geq t_0$ в начальном состоянии x_t^* первоначально согласованный принцип оптимальности предписывает дележ $\xi(x_t^*) = [\xi^1(x_t^*), \xi^2(x_t^*)]$. Это означает, что игроки договариваются о таком дележе, когда доля игрока i в игре равна $\xi^i(x_t^*)$.

Условие индивидуальной рациональности требует выполнения неравенства

$$\xi^i(x_t^*) \geq V^i(x_t^*), \quad \text{при } i \in \{1, 2\}.$$

Следуя [Petrosjan, 1997] и [Yeung, Petrosjan, 2004], будем обозначать через $B(s) = [B_1(s), B_2(s)]$ мгновенные выигрыши игроков (плотности выигрышей) при кооперации в момент времени $s \in [t_0, \infty)$ в кооперативной игре $\Gamma_c(x_{t_0}^*)$.

Тогда вдоль кооперативной траектории $\{x^*(t)\}_{t \geq t_0}$ будем иметь

$$\begin{aligned} \xi^i(x_\tau^*) &= \int_\tau^\infty B_i(s) \exp[-r(s - \tau)] ds, \quad i \in \{1, 2\}, \\ \xi^i(x_t^*) &= \int_t^\infty B_i(s) \exp[-r(s - t)] ds, \quad i \in \{1, 2\} \quad t \geq \tau. \end{aligned}$$

Далее определим величины

$$\begin{aligned} \gamma^i(\tau; \tau, x_\tau^*) &= \int_\tau^\infty B_i(s) \exp[-r(s - \tau)] ds = \xi^i(x_\tau^*), \\ \gamma^i(\tau; t, x_t^*) &= \int_t^\infty B_i(s) \exp[-r(s - \tau)] ds, \\ i &\in \{1, 2\}, \quad \tau \in [t_0, \infty). \end{aligned}$$

Заметим, что

$$\begin{aligned} \gamma^i(\tau; t, x_t^*) &= \exp[-r(t - \tau)] \int_t^\infty B_i^*(s) \exp[-r(s - t)] ds = \\ &= \exp[-r(t - \tau)] \xi^i(x_t^*) = \\ &= \exp[-r(t - \tau)] \gamma^i(t; t, x_t^*), \quad i \in \{1, 2\}. \end{aligned} \quad (8.2.5)$$

Условие (8.2.5) гарантирует динамическую устойчивость дележа в том смысле, что применение согласованного в начале игры принципа оптимальности к подыгре, стартовой позже из начальных условий на оптимальной траектории сохраняет оптимальность выбранного дележа.

Следуя схеме доказательства теоремы п. 8.1.12, получаем

$$\begin{aligned} \gamma^i(\tau; \tau, x_\tau^*) &= \int_\tau^{\tau+\Delta t} B_i(s) \exp[-r(s - \tau)] ds + \\ &\quad + \exp[-r(\Delta t)] \gamma^i(\tau + \Delta t; \tau + \Delta t, x_\tau^* + \Delta x_\tau^*), \\ &\text{для } \tau \in [t_0, T], \quad i \in \{1, 2\}; \end{aligned} \quad (8.2.6)$$

где

$$\Delta x_\tau^* = f[x_\tau^*, \psi_1^*(x_\tau^*), \psi_2^*(x_\tau^*)] \Delta t + o(\Delta t).$$

Поскольку $x_\tau^* + \Delta x_\tau^* = x_{\tau+\Delta t}^*$, то из (8.2.6) следует

$$\begin{aligned} \gamma^i(\tau; \tau + \Delta t, x_{\tau+\Delta t}^*) &= \exp[-r(\Delta t)] \xi^i(x_{\tau+\Delta t}^*) = \\ &= \exp[-r(\Delta t)] \gamma^i(\tau + \Delta t; \tau + \Delta t, x_{\tau+\Delta t}^*) = \\ &= \exp[-r(\Delta t)] \gamma^i(\tau + \Delta t; \tau + \Delta t, x_\tau^* + \Delta x_\tau^*). \end{aligned} \quad (8.2.7)$$

Поэтому (8.2.6) переписывается в виде

$$\begin{aligned} \gamma^i(\tau; \tau, x_\tau^*) &= \int_\tau^{\tau+\Delta t} B_i(s) \exp[-r(s-\tau)] ds + \gamma^i(\tau; \tau + \Delta t, x_{\tau+\Delta t}^*), \\ \tau &\in [t_0, T], \quad i \in \{1, 2\}. \end{aligned} \quad (8.2.8)$$

Можно получить следующее соотношение:

$$\begin{aligned} \int_\tau^{\tau+\Delta t} B_i(s) \exp[-r(s-\tau)] ds &= \gamma^i(\tau; \tau, x_\tau^*) - \gamma^i(\tau; \tau + \Delta t, x_{\tau+\Delta t}^*), \\ \text{для всех } \tau &\in [t_0, T] \text{ и } i \in \{1, 2\}. \end{aligned} \quad (8.2.9)$$

При $\Delta t \rightarrow 0$ условие (8.2.9) переходит в

$$\begin{aligned} B_i(\tau) \Delta t &= -[\gamma_t^i(\tau; t, x_t^*)|_{t=\tau}] \Delta t - \\ &- [\gamma_{x_t^*}^i(\tau; t, x_t^*)|_{t=\tau}] f[x_\tau^*, \psi_1^*(x_\tau^*), \psi_2^*(x_\tau^*)] \Delta t - o(\Delta t). \end{aligned} \quad (8.2.10)$$

Деля обе части (8.2.10) на Δt и устремляя $\Delta t \rightarrow 0$, получаем

$$B_i(\tau) = -[\gamma_t^i(\tau; t, x_t^*)|_{t=\tau}] - [\gamma_{x_t^*}^i(\tau; t, x_t^*)|_{t=\tau}] f[x_\tau^*, \psi_1^*(x_\tau^*), \psi_2^*(x_\tau^*)]. \quad (8.2.11)$$

Используя (8.2.5), получаем $\gamma^i(\tau; t, x_t^*) = \exp[-r(t-\tau)] \xi^i(x_t^*)$ и $\gamma^i(\tau; \tau, x_\tau^*) = \xi^i(x_\tau^*)$. Тогда из (8.2.11) получаем следующую теорему.

Теорема. Мгновенные выплаты игрокам в моменты времени $\tau \in [t_0, T]$ в соответствии с ПРД

$$\begin{aligned} B_i(\tau) &= r\xi^i(x_\tau^*) - \xi_{x_\tau^*}^i(x_\tau^*) f[x_\tau^*, \psi_1^*(x_\tau^*), \psi_2^*(x_\tau^*)], \\ \text{для всех } \tau &\in [t_0, T] \text{ и } i \in \{1, 2\}, \end{aligned} \quad (8.2.12)$$

обеспечивают динамическую устойчивость решения в кооперативной игре $\Gamma_c(x_0)$.

8.2.4. Рассмотрим теперь динамическую устойчивость (временную состоятельность) решений для конкретных принципов оптимальности.

Рассмотрим кооперативную игру $\Gamma_c(x_0)$, в которой игроки договорились о максимизации суммы выигрышей и последующего ее раздела в соответствии с арбитражной схемой Нэша. В этом случае дележ должен иметь вид, описываемый в следующем утверждении.

Утверждение. В игре $\Gamma_c(x_0)$ в момент времени t_0 компонента дележа

$$\xi^i(x_0) = V^i(x_0) + \frac{1}{2} \left[W(x_0) - \sum_{j=1}^2 V^j(x_0) \right]$$

приписывается игроку i , $i \in \{1, 2\}$. И в момент времени $\tau \in (t_0, \infty)$ компонента дележа

$$\xi^i(x_\tau^*) = V^i(x_\tau^*) + \frac{1}{2} \left[W(x_\tau^*) - \sum_{j=1}^2 V^j(x_\tau^*) \right],$$

приписывается игроку i при $i \in \{1, 2\}$.

Используя теорему п. 8.2.3, можно получить формулу для ПРД в виде

$$\begin{aligned} B_i(\tau) = B_i(x_\tau^*) &= \frac{1}{2} \left\{ rV^i(x_\tau^*) - V_{x_\tau^*}^i(x_\tau^*) f[x_\tau^*, \psi_1^*(x_\tau^*), \psi_2^*(x_\tau^*)] \right\} + \\ &+ \frac{1}{2} \left\{ rW(x_\tau^*) - W_{x_\tau^*}(x_\tau^*) f[x_\tau^*, \psi_1^*(x_\tau^*), \psi_2^*(x_\tau^*)] \right\} - \\ &- \frac{1}{2} \left\{ rV^j(x_\tau^*) - \xi_{x_\tau^*}^j(x_\tau^*) f[x_\tau^*, \psi_1^*(x_\tau^*), \psi_2^*(x_\tau^*)] \right\}, \\ &\text{при } i, j \in \{1, 2\} \text{ и } i \neq j, \end{aligned} \quad (8.2.13)$$

ПРД, определенная в (8.2.13), обеспечивает динамическую устойчивость (временную состоятельность) решения кооперативной игры $\Gamma_c(x_0)$, в которой игроки условились о разделе максимального суммарного выигрыша в соответствии с утверждением п. 8.2.4.

8.2.5. Пример 4. Рассмотрим игру добычи ограниченного ресурса из примера 1, в которой продолжительность игры бесконечна. В момент t_0 выигрыши игроков 1 и 2 соответственно равны

$$\begin{aligned} &\int_{t_0}^{\infty} \left[u_1(s)^{1/2} - \frac{c_1}{x(s)^{1/2}} u_1(s) \right] \exp[-r(t - t_0)] ds, \\ &\int_{t_0}^{\infty} \left[u_2(s)^{1/2} - \frac{c_2}{x(s)^{1/2}} u_2(s) \right] \exp[-r(t - t_0)] ds. \end{aligned} \quad (8.2.14)$$

Запас ресурса $x(s) \in X \subset R$ изменяется в соответствии с уравнением (8.1.5).

Равновесие по Нэшу в игре (8.1.5)–(8.2.14) характеризуется следующим образом

$$\begin{aligned} rV^i(x) &= \max_{u_i} \left\{ u_i^{1/2} - \frac{c_i}{x^{1/2}} u_i + V_x^i(x) [ax^{1/2} - bx - u_i - \phi_j^*(x)] \right\}, \\ &i, j \in \{1, 2\}, \quad i \neq j. \end{aligned} \quad (8.2.15)$$

Проводя максимизацию правой части в (8.2.15), получаем

$$\phi_i^*(x) = \frac{x}{4[c_i + V_x^i(x)x^{1/2}]^2}, \quad i \in \{1, 2\}.$$

Подставляя $\phi_1^*(x)$ и $\phi_2^*(x)$ в (8.2.15) и решая (8.2.15), получаем функции значения игроков $i \in \{1, 2\}$ в виде

$$V^i(t, x) = [A_i x^{1/2} + B_i],$$

где при $i, j \in \{1, 2\}$ и $i \neq j$, A_i , B_i , A_j и B_j удовлетворяют уравнениям

$$\left[r + \frac{b}{2}\right] A_i - \frac{1}{2[c_i + A_i/2]} + \frac{c_i}{4[c_i + A_i/2]^2} + \frac{A_i}{8[c_i + A_i/2]^2} + \frac{A_i}{8[c_j + A_j/2]^2} = 0,$$

$$B_i = \frac{a}{2} A_i.$$

Равновесные стратегии имеют вид

$$\phi_1^*(x) = \frac{x}{4[c_1 + A_1/2]^2}, \quad \phi_2^*(x) = \frac{x}{4[c_2 + A_2/2]^2}.$$

Рассмотрим теперь случай, когда игроки приходят к соглашению максимизировать суммарный выигрыш и разделить его потом в соответствии с утверждением п. 8.2.4. В этом случае они должны решить следующую задачу максимизации

$$\int_{t_0}^{\infty} \left(\left[u_1(s)^{1/2} - \frac{c_1}{x(s)^{1/2}} u_1(s) \right] + \left[u_2(s)^{1/2} - \frac{c_2}{x(s)^{1/2}} u_2(s) \right] \right) \exp[-r(t - t_0)] ds \quad (8.2.16)$$

при условии (8.1.5).

Используя теорему п. 8.2.3, получаем

$$rW(x) = \max_{u_1, u_2} \left\{ \left(\left[u_1^{1/2} - \frac{c_1}{x^{1/2}} u_1 \right] + \left[u_2^{1/2} - \frac{c_2}{x^{1/2}} u_2 \right] \right) + W_x(x) [ax^{1/2} - bx - u_1 - u_2] \right\}.$$

Используя стандартные выкладки, получаем

$$W(x) = [\hat{A}x^{1/2} + \hat{B}],$$

где

$$\left[r + \frac{b}{2}\right] \hat{A} - \frac{1}{2[c_1 + \hat{A}/2]} - \frac{1}{2[c_2 + \hat{A}/2]} + \frac{c_1}{4[c_1 + \hat{A}/2]^2} +$$

$$+ \frac{c_2}{4[c_2 + \hat{A}/2]^2} + \frac{\hat{A}}{8[c_1 + \hat{A}/2]^2} + \frac{\hat{A}}{8[c_2 + \hat{A}/2]^2} = 0,$$

$$\hat{B} = \frac{a}{2r} \hat{A}.$$

Оптимальные кооперативные решения тогда могут быть получены в виде

$$\psi_1^*(x) = \frac{x}{4[c_1 + \hat{A}/2]^2}, \quad \psi_2^*(x) = \frac{x}{4[c_2 + \hat{A}/2]^2}. \quad (8.2.17)$$

Подстановка этих управлений в (8.1.5) дает уравнения движения в случае кооперации

$$\dot{x}(s) = ax(s)^{1/2} - bx(s) - \frac{x(s)}{4[c_1 + \hat{A}/2]^2} - \frac{x(s)}{4[c_2 + \hat{A}/2]^2}, \quad x(t_0) = x_0. \quad (8.2.18)$$

Решая (8.2.18), получаем кооперативную траекторию игры $\Gamma_c(x_0)$ в виде

$$x^*(s) = \left[\frac{a}{2H} + \left(x_0^{1/2} - \frac{a}{2H} \right) \exp[-H(s - t_0)] \right]^2,$$

где

$$H = - \left[\frac{b}{2} + \frac{1}{8 [c_1 + \hat{A}/2]^2} + \frac{1}{8 [c_2 + \hat{A}/2]^2} \right].$$

Используя (8.2.13), получаем

$$\begin{aligned} B_i(\tau) = B_i(x_\tau^*) = & \frac{1}{2} \left\{ r [A_i(x_\tau^*)^{1/2} + B_i] + r [\hat{A}(x_\tau^*)^{1/2} + \hat{B}] - r [A_j(x_\tau^*)^{1/2} + B_j] \right\} - \\ & - \frac{1}{4} \left\{ A_i(x_\tau^*)^{-1/2} + \hat{A}(x_\tau^*)^{-1/2} - A_j(x_\tau^*)^{-1/2} \right\} \times \\ & \times \left[a(x_\tau^*)^{1/2} - bx_\tau^* - \frac{x_\tau^*}{4 [c_1 + \hat{A}/2]^2} - \frac{x_\tau^*}{4 [c_2 + \hat{A}/2]^2} \right], \\ & \text{при } i, j \in \{1, 2\} \text{ и } i \neq j. \end{aligned} \quad (8.2.19)$$

§ 8.3. Игры с нетрансферабельными выигрышами

8.3.1. Выигрыши игроков в игре можно считать трансферабельными, если существует некоторая приемлемая разменная единица. В то же время существует множество случаев, когда выигрыши нельзя считать трансферабельными — рассмотрим хотя бы случаи, касающиеся политической стабильности, религиозной терпимости и суверенитета. Решения соответствующих кооперативных дифференциальных игр в этом случае существенно усложняются. В этом разделе мы будем рассматривать кооперативные дифференциальные игры с нетрансферабельными выигрышами.

Рассмотрим неантагонистическую дифференциальную игру двух лиц с динамикой (8.1.1) и выигрышами (8.1.2). Очевидно, что некооперативное решение этой игры то же, что и в игре с трансферабельными выигрышами.

Рассмотрим парето-оптимальные траектории при кооперации. В случае нетрансферабельных выигрышей необходимыми условиями оптимальности, также как и в случае трансферабельных выигрышей, является групповая и индивидуальная рациональность. Для достижения групповой рациональности необходимо ограничиться управлениями, которые приводят к парето-оптимальным исходам.

8.3.2. Рассмотрим кооперативную игру $\Gamma_c(x_0, T - t_0)$, в которой выигрыши нетрансферабельны. Парето-оптимальные исходы в $\Gamma_c(x_0, T - t_0)$ эквивалентны выбору весового коэффициента $\alpha_1 \in (0, \infty)$ и решению следующей задачи максимизации

(см. [Yeung, Petrosyan, 2005]):

$$\begin{aligned}
 & \max_{u_1, u_2} \{J^1(t_0, x_0) + \alpha_1 J^2(t_0, x_0)\} \equiv \\
 & \equiv \max_{u_1, u_2} \left\{ \int_{t_0}^T (g^1[s, x(s), u_1(s), u_2(s)] + \right. \\
 & \quad \left. + \alpha_1 g^2[s, x(s), u_1(s), u_2(s)]) \exp \left[- \int_{t_0}^s r(y) dy \right] ds + \right. \\
 & \quad \left. + [q^1(x(T)) + \alpha_1 q^2(x(T))] \exp \left[- \int_{t_0}^T r(y) dy \right] \middle| x(t_0) = x_0 \right\}, \quad (8.3.1)
 \end{aligned}$$

при условии (8.1.1). Заметим, что оптимальные управления для задачи $\max_{u_1, u_2} \{J^1(t_0, x_0) + \alpha_1 J^2(t_0, x_0)\}$ совпадают с оптимальными управлениями для задачи $\max_{u_1, u_2} \{J^2(t_0, x_0) + \alpha_2 J^1(t_0, x_0)\}$, если $\alpha_1 = 1/\alpha_2$.

Пусть в $\Gamma_c(x_0, T - t_0)$ весовой коэффициент α_1^0 выбран игроками в соответствии с некоторым согласованным принципом оптимальности. Используя теорему п. 6.1.1, мы получаем:

Теорема. *Пара управлений*

$$\left\{ [\psi_1^{\alpha_1^0(t_0)}(t, x), \psi_2^{\alpha_1^0(t_0)}(t, x)], \text{ при } t \in [t_0, T] \right\}$$

представляет собой оптимальное решение задачи

$$\max_{u_1, u_2} \{J^1(t_0, x_0) + \alpha_1^0 J^2(t_0, x_0)\},$$

если существует такая непрерывно-дифференцируемая функция $W^{\alpha_1^0(t_0)}(t, x) : [t_0, T] \times R^m \rightarrow R$, которая удовлетворяет следующему уравнению Беллмана:

$$\begin{aligned}
 & -W_t^{\alpha_1^0(t_0)}(t, x) = \\
 & = \max_{u_1, u_2} \left\{ (g^1[t, x, u_1, u_2] + \alpha_1^0 g^2[t, x, u_1, u_2]) \exp \left[- \int_{t_0}^t r(y) dy \right] + \right. \\
 & \quad \left. + W_x^{\alpha_1^0(t_0)} f[t, x, u_1, u_2] \right\} = \\
 & = \left(g^1[t, x, \psi_1^{\alpha_1^0}(t, x), u_2] + \alpha_1^0 g^2[t, x, \psi_1^{\alpha_1^0}(t, x), u_2] \right) \exp \left[- \int_{t_0}^t r(y) dy \right], \\
 & \text{где} \\
 & W^{\alpha_1^0(t_0)}(T, x) = \\
 & = \exp \left[- \int_{t_0}^T r(y) dy \right] \sum_{j=1}^2 q^j(x) [q^1(x) + \alpha_1 q^2(x)] \exp \left[- \int_{t_0}^T r(y) dy \right].
 \end{aligned}$$

Подставляя $\psi_1^{\alpha_1^0(t_0)}(t, x)$ и $\psi_2^{\alpha_1^0(t_0)}(t, x)$ в (8.1.1), получаем уравнение для парето-оптимальной траектории, соответствующей весовому коэффициенту α_1^0 :

$$\dot{x}(s) = f \left[s, x(s), \psi_1^{\alpha_1^0(t_0)}(s, x(s)), \psi_2^{\alpha_1^0(t_0)}(s, x(s)) \right], \quad x(t_0) = x_0. \quad (8.3.2)$$

Решение (8.3.2) $x^{\alpha_1^0}(t)$ может быть представлено в виде

$$x^{\alpha_1^0}(t) = x_0 + \int_{t_0}^t f \left[s, x^{\alpha_1^0}(s), \psi_1^{\alpha_1^0(t_0)}(s, x^{\alpha_1^0}(s)), \psi_2^{\alpha_1^0(t_0)}(s, x^{\alpha_1^0}(s)) \right] ds, \\ \text{при } t \in [t_0, T]. \quad (8.3.3)$$

Путь $\left\{ x^{\alpha_1^0}(t) \right\}_{t=t_0}^T$ представляет собой оптимальную траекторию в задаче

$$\max_{u_1, u_2} \left\{ J^1(t_0, x_0) + \alpha_1^0 J^2(t_0, x_0) \right\}.$$

Обозначения $x^{\alpha_1^0}(t)$ и $x_t^{\alpha_1^0}$ мы будем использовать в дальнейшем в зависимости от степени удобства.

Для обеспечения групповой рациональности кооперативные управления

$$\left[\psi_1^{\alpha_1^0(t_0)}(t, x), \psi_2^{\alpha_1^0(t_0)}(t, x) \right]$$

должны быть использованы на отрезке времени $[t_0, T]$.

8.3.3. Рассмотрим теперь кооперативную игру $\Gamma_c(x_{\tau}^{\alpha_1^0}, T - \tau)$ при $\tau \in [t_0, T]$. Пусть α_1^{τ} — весовой коэффициент, выбранный в соответствии с согласованным в начале игры принципом оптимальности. Мы будем использовать обозначения $\left\{ \left[\psi_1^{\alpha_1^{\tau}(\tau)}(t, x), \psi_2^{\alpha_1^{\tau}(\tau)}(t, x) \right], t \in [\tau, T] \right\}$ для определения множества оптимальных управлений и $W^{\alpha_1^{\tau}(\tau)}(t, x) : [\tau, T] \times R^n \rightarrow R$ для функции значения из теоремы п. 8.3.2.

Можно показать, что

$$\left[\psi_1^{\alpha_1^{\tau}(\tau)}(t, x), \psi_2^{\alpha_1^{\tau}(\tau)}(t, x) \right] = \left[\psi_1^{\alpha_1^0(t_0)}(t, x), \psi_2^{\alpha_1^0(t_0)}(t, x) \right] \text{ при } \alpha_1^{\tau} = \alpha_1^0.$$

Замечание 3. Групповая рациональность будет иметь место только в случае, если $\alpha_1^{\tau} = \alpha_1^0$ будет весовым коэффициентом во всех играх $\Gamma_c(x_{\tau}, T - \tau)$ при $\tau \in [t_0, T]$.

8.3.4. Рассмотрим теперь индивидуальные выигрыши игроков при кооперации. Для проверки индивидуальной рациональности выигрыша игрока выпишем выигрыши игроков при кооперации вдоль кооперативной траектории. Для того чтобы это сделать, подставим кооперативные управления $\psi_1^{\alpha_1^0(t_0)}(t, x)$ и $\psi_2^{\alpha_1^0(t_0)}(t, x)$ в выражения для выигрышей (8.1.2) и получим ожидаемые выигрыши игроков в игре $\Gamma_c(x_0, T - t_0)$ с весовым коэффициентом α_1^0 . Следуя [Yeung, 2004], приведем определение.

Определение. Определим кооперативный выигрыш игрока i на отрезке времени $[t, T]$ как

$$\hat{W}^{\alpha_1^0(t_0)i} \left(t, x_t^{\alpha_1^0} \right) = \\ = \int_t^T g^i \left[s, x^{\alpha_1^0}(s), \psi_1^{\alpha_1^0(t_0)}(s, x^{\alpha_1^0}(s)), \psi_2^{\alpha_1^0(t_0)}(s, x^{\alpha_1^0}(s)) \right] \exp \left[- \int_{t_0}^s r(y) dy \right] ds + \\ + \exp \left[- \int_{t_0}^T r(y) dy \right] q^i \left(x^{\alpha_1^0}(T) \right), \quad i \in \{1, 2\},$$

где

$$\begin{aligned}\dot{x}^{\alpha_1^0}(s) &= f\left[s, x^{\alpha_1^0}(s), \psi_1^{\alpha_1^0(t_0)}\left(s, x^{\alpha_1^0}(s)\right), \psi_2^{\alpha_1^0(t_0)}\left(s, x^{\alpha_1^0}(s)\right)\right], \\ x^{\alpha_1^0}(t) &= x_t^{\alpha_1^0}.\end{aligned}$$

Заметим, что при $\Delta t \rightarrow 0$ выражение для $\hat{W}^{\alpha_1^0(t_0)i}\left(t, x_t^{\alpha_1^0}\right)$ принимает вид

$$\begin{aligned}\hat{W}^{\alpha_1^0(t_0)i}\left(t, x_t^{\alpha_1^0}\right) &= \\ &= \int_t^{t+\Delta t} g^i\left[s, x^{\alpha_1^0}(s), \psi_1^{\alpha_1^0(t_0)}\left(s, x^{\alpha_1^0}(s)\right), \psi_2^{\alpha_1^0(t_0)}\left(s, x^{\alpha_1^0}(s)\right)\right] \times \\ &\quad \times \exp\left[-\int_{t_0}^s r(y) dy\right] ds + \hat{W}^{\alpha_1^0(t_0)i}\left(t + \Delta t, x^{\alpha_1^0} + \Delta x^{\alpha_1^0}\right),\end{aligned}\quad (8.3.4)$$

где $x^{\alpha_1^0} = f\left[t, x_t^{\alpha_1^0}, \psi_1^{\alpha_1^0(t_0)}\left(t, x_t^{\alpha_1^0}\right), \psi_2^{\alpha_1^0(t_0)}\left(t, x_t^{\alpha_1^0}\right)\right] \Delta t$.

Используя теорему Тейлора, получаем

$$\begin{aligned}\hat{W}^{\alpha_1^0(t_0)i}\left(t, x_t^{\alpha_1^0}\right) &= \\ &= g^i\left[t, x_t^{\alpha_1^0}, \psi_1^{\alpha_1^0(t_0)}\left(t, x_t^{\alpha_1^0}\right), \psi_2^{\alpha_1^0(t_0)}\left(t, x_t^{\alpha_1^0}\right)\right] \exp\left[-\int_{t_0}^t r(y) dy\right] \Delta t + \\ &\quad + \hat{W}^{\alpha_1^0(t_0)i}\left(t, x_t^{\alpha_1^0}\right) + \hat{W}_t^{\alpha_1^0(t_0)i}\left(t, x_t^{\alpha_1^0}\right) \Delta t + \\ &\quad + \hat{W}_x^{\alpha_1^0(t_0)i}\left(t, x_t^{\alpha_1^0}\right) f\left[t, x_t^{\alpha_1^0}, \psi_1^{\alpha_1^0(t_0)}\left(t, x_t^{\alpha_1^0}\right), \psi_2^{\alpha_1^0(t_0)}\left(t, x_t^{\alpha_1^0}\right)\right] \Delta t + \\ &\quad + o(\Delta t), \quad i \in \{1, 2\}.\end{aligned}\quad (8.3.5)$$

Деля на Δt и полагая $\Delta t \rightarrow 0$, получаем

$$\begin{aligned}-\hat{W}_t^{\alpha_1^0(t_0)i}\left(t, x_t^{\alpha_1^0}\right) &= \\ &= g^i\left[t, x_t^{\alpha_1^0}, \psi_1^{\alpha_1^0(t_0)}\left(t, x_t^{\alpha_1^0}\right), \psi_2^{\alpha_1^0(t_0)}\left(t, x_t^{\alpha_1^0}\right)\right] \exp\left[-\int_{t_0}^t r(y) dy\right] + \\ &\quad + \hat{W}_x^{\alpha_1^0(t_0)i}\left(t, x_t^{\alpha_1^0}\right) f\left[t, x_t^{\alpha_1^0}, \psi_1^{\alpha_1^0(t_0)}\left(t, x_t^{\alpha_1^0}\right), \psi_2^{\alpha_1^0(t_0)}\left(t, x_t^{\alpha_1^0}\right)\right], \\ &\quad i \in \{1, 2\}.\end{aligned}\quad (8.3.6)$$

Граничные условия дают нам

$$\hat{W}^{\alpha_1^0(t_0)i}\left(T, x_T^{\alpha_1^0}\right) = \exp\left[-\int_{t_0}^T r(y) dy\right] q^i\left(x_T^{\alpha_1^0}\right), \quad i \in \{1, 2\}.\quad (8.3.7)$$

Теорема [Yeung, 2004]. Если существуют непрерывно-дифференцируемые функции $\hat{W}^{\alpha_1^0(t_0)i} \left(t, x_t^{\alpha_1^0} \right) : [t_0, T] \times R^m \rightarrow R$, $i \in \{1, 2\}$, удовлетворяющие

$$\begin{aligned} -\hat{W}_t^{\alpha_1^0(t_0)i} \left(t, x_t^{\alpha_1^0} \right) = \\ = g^i \left[t, x_t^{\alpha_1^0}, \psi_1^{\alpha_1^0(t_0)} \left(t, x_t^{\alpha_1^0} \right), \psi_2^{\alpha_1^0(t_0)} \left(t, x_t^{\alpha_1^0} \right) \right] \exp \left[- \int_{t_0}^t r(y) dy \right] + \\ + \hat{W}_x^{\alpha_1^0(t_0)i} \left(t, x_t^{\alpha_1^0} \right) f \left[t, x_t^{\alpha_1^0}, \psi_1^{\alpha_1^0(t_0)} \left(t, x_t^{\alpha_1^0} \right), \psi_2^{\alpha_1^0(t_0)} \left(t, x_t^{\alpha_1^0} \right) \right], \end{aligned}$$

и

$$\hat{W}^{\alpha_1^0(t_0)i} \left(T, x_T^{\alpha_1^0} \right) = \exp \left[- \int_{t_0}^T r(y) dy \right] q^i \left(x_T^{\alpha_1^0} \right), \quad i \in \{1, 2\},$$

то $\hat{W}^{\alpha_1^0(t_0)i} \left(t, x_t^{\alpha_1^0} \right)$ есть кооперативный выигрыш игрока i на отрезке времени $[t, T]$ при весовом коэффициенте α_1^0 .

8.3.5. Повторим предыдущие рассуждения для подыгры $\Gamma_c \left(x_\tau^{\alpha_1^0}, T - \tau \right)$ с началом в момент времени τ из состояния $x_\tau^{\alpha_1^0}$ с выбранным весовым коэффициентом $\alpha_1^\tau = \alpha_1^0$. Легко можно видеть, что

$$\hat{W}^{\alpha_1^0(t_0)i} \left(\tau, x_\tau^{\alpha_1^0} \right) \exp \left[\int_{t_0}^\tau r(y) dy \right] = \hat{W}^{\alpha_1^0(\tau)i} \left(\tau, x_\tau^{\alpha_1^0} \right), \quad i \in \{1, 2\}. \quad (8.3.8)$$

Замечание 4. Для выполнения условия индивидуальной рациональности в течение всей игры выбранный коэффициент α_1^0 должен удовлетворять условию

$$\begin{aligned} \hat{W}^{\alpha_1^0(t_0)i} \left(\tau, x_\tau^{\alpha_1^0} \right) \exp \left[\int_{t_0}^\tau r(y) dy \right] = \hat{W}^{\alpha_1^0(\tau)i} \left(\tau, x_\tau^{\alpha_1^0} \right) \geq \\ \geq V^{(\tau)i} \left(\tau, x_\tau^{\alpha_1^0} \right), \quad i \in \{1, 2\}. \end{aligned}$$

Исследуем теперь состоятельность во времени (динамическую устойчивость) решения. Временная состоятельность принципа оптимальности гарантирует сохранение оптимальности первоначально выработанного решения в каждый момент времени, когда игра развивается вдоль выбранной в начале игры оптимальной траектории. В дополнение к этому должны сохраняться парето-оптимальность и индивидуальная рациональность решения вдоль оптимальной траектории. В случае игры $\Gamma_c(x_0, T - t_0)$ с нетрансферабельными выигрышами временная состоятельность сводится к выполнению следующих условий.

8.3.6. Условие.

1. Дележ $\left[\hat{W}^{\tau(\alpha_1^\tau)1} \left(\tau, x_\tau^{\alpha_1^0} \right), \hat{W}^{\tau(\alpha_1^\tau)2} \left(\tau, x_\tau^{\alpha_1^0} \right) \right]$, при $\tau \in [t_0, T]$ должен оставаться парето-оптимальным в подыграх вдоль оптимальной траектории.

2.

$$\hat{W}^{\alpha_1^\tau(\tau)i} \left(\tau, x_\tau^{\alpha_1^0} \right) \geq V^{(\tau)i} \left(\tau, x_\tau^{\alpha_1^0} \right),$$

для $i \in \{1, 2\}$ и $\tau \in [t_0, T]$; и

3.

$$\hat{W}^{\alpha_1^0(t_0)i}(\tau, x_\tau^{\alpha_1^0}) \exp \left[\int_{t_0}^{\tau} r(y) dy \right] = \hat{W}^{\alpha_1^\tau(\tau)i}(\tau, x_\tau^{\alpha_1^0}),$$

для $i \in \{1, 2\}$ и $\tau \in [t_0, T]$.

Условие 1) обеспечивает парето-оптимальность на всем отрезке игры, условие 2) обеспечивает индивидуальную рациональность на всем отрезке игры, и условие 3) гарантирует временную состоятельность.

8.3.7. Предположим, что в начальный момент t_0 игроки согласились на некоторый принцип оптимальности в игре $\Gamma_c(x_0, T - t_0)$, который заключается:

- 1) в выборе весового коэффициента α_1^0 , и соответствующей пары кооперативных управлений $[\psi_1^{\alpha_1^0(t_0)}(t, x), \psi_2^{\alpha_1^0(t_0)}(t, x)]$
- 2) и выборе на основе этого дележа

$$[\hat{W}^{\alpha_1^0(t_0)1}(t_0, x_0), \hat{W}^{\alpha_1^0(t_0)2}(t_0, x_0)].$$

В последующие моменты времени $\tau \in (t_0, T]$, используя тот же принцип оптимальности, игроки соглашаются

- 1) выбрать весовой коэффициент α_1^τ и пару кооперативных управлений $[\psi_1^{\alpha_1^\tau(\tau)}(t, x), \psi_2^{\alpha_1^\tau(\tau)}(t, x)]$, и
- 2) на основе этого дележ $[\hat{W}^{\alpha_1^\tau(\tau)1}(\tau, x_\tau^{\alpha_1^0}), \hat{W}^{\alpha_1^\tau(\tau)2}(\tau, x_\tau^{\alpha_1^0})]$.

Теорема. Принцип оптимальности, на основе которого игроки выбирают один и тот же весовой коэффициент α_1^0 во всех подыграх $\Gamma_c(x_\tau^{\alpha_1^0}, T - \tau)$ и такой, что

$$\hat{W}^{\alpha_1^0(\tau)1}(\tau, x_\tau^{\alpha_1^0}) \geq V^{(\tau)1}(\tau, x_\tau^{\alpha_1^0}), \quad \hat{W}^{\alpha_1^0(\tau)2}(\tau, x_\tau^{\alpha_1^0}) \geq V^{(\tau)2}(\tau, x_\tau^{\alpha_1^0}),$$

представляет собой состоятельное во времени (динамически устойчивое) решение кооперативной игры $\Gamma_c(x_0, T - t_0)$.

Доказательство. Если игроки выбирают один и тот же весовой коэффициент α_1^0 во всей игре, то это означает, что в течение игры $\Gamma_c(x_\tau^{\alpha_1^0}, T - \tau)$ будет использована одна и та же пара кооперативных управлений $[\psi_1^{\alpha_1^0(t_0)}(t, x), \psi_2^{\alpha_1^0(t_0)}(t, x)]$ при $[t_0, T]$. Групповая оптимальность при этом гарантирована, и дележ

$$\xi^{(\tau)}(x_\tau, T - \tau) = [\hat{W}^{\alpha_1^0(\tau)1}(\tau, x_\tau^{\alpha_1^0}), \hat{W}^{\alpha_1^0(\tau)2}(\tau, x_\tau^{\alpha_1^0})], \text{ при } \tau \in [t_0, T],$$

действительно оптимален по Парето во всех подыграх и в основной игре. Поэтому часть 1) условия п. 8.3.6 выполнена. Индивидуальная рациональность следует из условия теоремы

$$\hat{W}^{\alpha_1^0(\tau)1}(\tau, x_\tau^{\alpha_1^0}) \geq V^{(\tau)1}(\tau, x_\tau^{\alpha_1^0}) \text{ и } \hat{W}^{\alpha_1^0(\tau)2}(\tau, x_\tau^{\alpha_1^0}) \geq V^{(\tau)2}(\tau, x_\tau^{\alpha_1^0}),$$

при $\tau \in [t_0, T]$,

и часть 2) условия п. 8.3.6 также выполнена.

Более того, из условия (8.3.8) имеем:

$$\hat{W}^{\alpha_1^0(t_0)i}(\tau, x_\tau^{\alpha_1^0}) \exp \left[\int_{t_0}^{\tau} r(y) dy \right] = \hat{W}^{\alpha_1^0(\tau)i}(\tau, x_\tau^{\alpha_1^0}), \quad i \in \{1, 2\}.$$

Часть 3) условия п. 8.3.6 также выполнена.

8.3.8. Пример 5. Рассмотрим неантагонистическую дифференциальную игру двух лиц из примера 1 п. 8.1.3 в предположении, что выигрыши нетрансферабельны. Необходимыми условиями, которым должен отвечать выбранный принцип оптимальности, являются групповая рациональность и индивидуальная рациональность.

Парето-оптимальные исходы в игре $\Gamma_c(x_0, T - t_0)$ определяются решением следующей задачи оптимального управления, включая выбор весового множителя $\alpha_1^0 \in (0, \infty)$:

$$\begin{aligned} \max_{u_1, u_2} \{ & J^1(t_0, x_0) + \alpha_1^0 J^2(t_0, x_0) \} \equiv \\ \equiv \max_{u_1, u_2} \{ & \int_{t_0}^T \left(\left[u_1(s)^{1/2} - \frac{c_1}{x(s)^{1/2}} u_1(s) \right] + \right. \\ & + \alpha_1 \left[u_2(s)^{1/2} - \frac{c_2}{x(s)^{1/2}} u_2(s) \right] \Big) \exp[-r(s - t_0)] ds + \\ & \left. + \exp[-r(T - t_0)] \left(q_1 x(T)^{1/2} + \alpha_1^0 q_2 x(T)^{1/2} \right) \Big| x(t_0) = x_0 \right\}, \end{aligned} \quad (8.3.9)$$

при ограничениях (8.1.5).

Пусть $[\psi_1^{\alpha_1^0(t_0)}(t, x), \psi_2^{\alpha_1^0(t_0)}(t, x)]$, $t \in [t_0, T]$ — пара управлений, которая доставляет максимум в задаче управления $\max_{u_1, u_2} \{ J^1(t_0, x_0) + \alpha_1^0 J^2(t_0, x_0) \}$, и $W^{(t_0)}(t, x) : [t_0, T] \times R^n \rightarrow R$ — функция значения, удовлетворяющая следующему уравнению (теорема п. 8.3.2):

$$\begin{aligned} -W_t^{\alpha_1^0(t_0)}(t, x) = \\ = \max_{u_1, u_2} \left\{ \left(\left[u_1^{1/2} - \frac{c_1}{x^{1/2}} u_1 \right] + \alpha_1^0 \left[u_2^{1/2} - \frac{c_2}{x^{1/2}} u_2 \right] \right) \exp[-r(t - t_0)] + \right. \\ \left. + W_x^{\alpha_1^0(t_0)}(t, x) \left[ax^{1/2} - bx - u_1 - u_2 \right] \right\}, \\ W^{\alpha_1^0(t_0)}(T, x) = \exp[-r(T - t_0)] \left[q_1 x^{1/2} + \alpha_1^0 q_2 x^{1/2} \right]. \end{aligned} \quad (8.3.10)$$

Максимизация правой части в (8.3.10) дает

$$\begin{aligned} \psi_1^{\alpha_1^0(t_0)}(t, x) &= \frac{x}{4 \left[c_1 + x^{1/2} W_x^{\alpha_1^0(t_0)}(t, x) \exp[r(t - t_0)] \right]^2}, \\ \psi_2^{\alpha_1^0(t_0)}(t, x) &= \frac{x}{4 \left[c_2 + x^{1/2} W_x^{\alpha_1^0(t_0)}(t, x) \exp[r(t - t_0)] / \alpha_1 \right]^2}, \\ t &\in [t_0, T]. \end{aligned} \quad (8.3.11)$$

Подставляя $\psi_1^{\alpha_1^0(t_0)}(t, x)$ и $\psi_2^{\alpha_1^0(t_0)}(t, x)$ из (8.3.11) в (8.3.10), получаем выражение для функции значения

$$W^{\alpha_1^0(t_0)}(t, x) = \exp[-r(t - t_0)] \left[A^{\alpha_1^0}(t) x^{1/2} + B^{\alpha_1^0}(t) \right], \quad t \in [t_0, T], \quad (8.3.12)$$

где $A^{\alpha_1^0}(t)$ и $B^{\alpha_1^0}(t)$ удовлетворяют уравнениям:

$$\begin{aligned} \dot{A}^{\alpha_1^0}(t) &= \left[r + \frac{b}{2} \right] A^{\alpha_1^0}(t) - \frac{1}{4[c_1 + A^{\alpha_1^0}(t)/2]} - \frac{\alpha_1^0}{4[c_2 + A^{\alpha_1^0}(t)/2\alpha_1^0]}, \\ \dot{B}^{\alpha_1^0}(t) &= rB^{\alpha_1^0}(t) - \frac{a}{2}A^{\alpha_1^0}(t), \\ A^{\alpha_1^0}(T) &= q_1 + \alpha_1^0 q_2, \quad B^{\alpha_1^0}(T) = 0. \end{aligned}$$

Подставляя частные производные $W_x^{\alpha_1^0(t_0)}(t, x)$ из (8.3.12) в $\psi_1^{\alpha_1^0(t_0)}(t, x)$ и $\psi_2^{\alpha_1^0(t_0)}(t, x)$, получаем оптимальное управление в задаче $\max_{u_1, u_2} \{J^1(t_0, x_0) + \alpha_1^0 J^2(t_0, x_0)\}$. Подставляя эти управления в (8.1.5), получаем парето-оптимальную траекторию, соответствующую весовому коэффициенту α_1^0 в виде

$$\begin{aligned} \dot{x}(s) &= ax(s)^{1/2} - bx(s) - \frac{x(s)}{4[c_1 + A^{\alpha_1^0}(s)/2]^2} - \frac{x(s)}{4[c_2 + A^{\alpha_1^0}(s)/2\alpha_1^0]^2}, \\ x(t_0) &= x_0. \end{aligned} \quad (8.3.13)$$

Решая (8.3.13), получаем явное выражение для парето-оптимальной траектории, соответствующее весовому коэффициенту α_1^0 в виде

$$x^{\alpha_1^0}(s) = \left\{ \Phi(t_0, s) \left[x_0^{1/2} + \int_{t_0}^s \Phi^{-1}(t_0, t) \frac{a}{2} dt \right] \right\}^2, \quad s \in [t_0, T], \quad (8.3.14)$$

где $\Phi(t_0, s) = \exp \left[\int_{t_0}^s H_2(\tau) d\tau \right]$ и

$$H_2(s) = - \left[\frac{b}{2} + \frac{1}{8[c_1 + A^{\alpha_1^0}(s)/2]^2} + \frac{1}{8[c_2 + A^{\alpha_1^0}(s)/2\alpha_1^0]^2} \right].$$

Кооперативное управление, соответствующее весовому коэффициенту α_1^0 на временном промежутке $[t_0, T]$, определяется по формуле

$$\begin{aligned} \psi_1^{\alpha_1^0(t_0)}(t, x_t^{\alpha_1^0}) &= \frac{x_t^{\alpha_1^0}}{4[c_1 + A^{\alpha_1^0}(t)/2]^2} \\ \psi_2^{\alpha_1^0(t_0)}(t, x_t^{\alpha_1^0}) &= \frac{x_t^{\alpha_1^0}}{4[c_2 + A^{\alpha_1^0}(t)/2\alpha_1^0]^2}, \end{aligned} \quad \text{при } t \in [t_0, T]. \quad (8.3.15)$$

Найдем индивидуальные выигрыши игроков. Подставим $\psi_1^{\alpha_1^0(t_0)}(t, x_t^{\alpha_1^0})$ и $\psi_2^{\alpha_1^0(t_0)}(t, x_t^{\alpha_1^0})$, определенные в (8.3.15), в выражение для функции выигрыша на отрезке времени $[t, T]$. Тогда получаем

$$\begin{aligned}
& \hat{W}^{\alpha_1^0(t_0)1} \left(t, x_t^{\alpha_1^0} \right) = \\
& = \int_t^T \left[\frac{x^{\alpha_1^0}(s)^{1/2}}{2[c_1 + A^{\alpha_1^0}(s)/2]} - \frac{c_1 x^{\alpha_1^0}(s)^{1/2}}{4[c_1 + A^{\alpha_1^0}(s)/2]^2} \right] \exp[-r(s - t_0)] ds + \\
& \quad + \exp[-r(T - \tau)] q_1 x^{\alpha_1^0}(T)^{1/2}; \\
& \hat{W}^{\alpha_1^0(t_0)2} \left(t, x_t^{\alpha_1^0} \right) = \\
& = \int_t^T \left[\frac{x^{\alpha_1^0}(s)^{1/2}}{2[c_2 + A^{\alpha_1^0}(s)/2\alpha_1^0]} - \frac{c_2 x^{\alpha_1^0}(s)^{1/2}}{4[c_2 + A^{\alpha_1^0}(s)/2\alpha_1^0]^2} \right] \exp[-r(s - t_0)] ds + \\
& \quad + \exp[-r(T - \tau)] q_2 \left(x_T^{\alpha_1^0} \right)^{1/2}.
\end{aligned}$$

Из теоремы п. 8.3.4 имеем, что если существуют непрерывно-дифференцируемые функции

$$\hat{W}^{\alpha_1^0(t_0)i} \left(t, x_t^{\alpha_1^0} \right) : [t_0, T] \times R^m \rightarrow R, i \in \{1, 2\},$$

удовлетворяющие уравнениям

$$\begin{aligned}
& \hat{W}_t^{\alpha_1^0(t_0)1} \left(t, x_t^{\alpha_1^0} \right) = \left[\frac{\left(x_t^{\alpha_1^0} \right)^{1/2}}{2[c_1 + A^{\alpha_1^0}(t)/2]} - \frac{c_1 \left(x_t^{\alpha_1^0} \right)^{1/2}}{4[c_1 + A^{\alpha_1^0}(t)/2]^2} \right] \exp[-r(t - t_0)] + \\
& + \hat{W}_x^{\alpha_1^0(t_0)1} \left(t, x_t^{\alpha_1^0} \right) \left[a \left(x_t^{\alpha_1^0} \right)^{1/2} - b x_t^{\alpha_1^0} - \frac{x_t^{\alpha_1^0}}{4[c_1 + A^{\alpha_1^0}(t)/2]^2} - \frac{x_t^{\alpha_1^0}}{4[c_2 + A^{\alpha_1^0}(t)/2\alpha_1^0]^2} \right], \\
& \hat{W}^{\alpha_1^0(t_0)1} \left(T, x_T^{\alpha_1^0} \right) \exp[-r(T - \tau)] q_1 x^{\alpha_1^0}(T)^{1/2} - \hat{W}_t^{\alpha_1^0(t_0)2} \left(t, x_t^{\alpha_1^0} \right) = \\
& = \left[\frac{\left(x_t^{\alpha_1^0} \right)^{1/2}}{2[c_2 + A^{\alpha_1^0}(t)/2\alpha_1^0]} - \frac{c_2 \left(x_t^{\alpha_1^0} \right)^{1/2}}{4[c_2 + A^{\alpha_1^0}(t)/2\alpha_1^0]^2} \right] \exp[-r(t - t_0)] + \\
& + \hat{W}_x^{\alpha_1^0(t_0)2} \left(t, x_t^{\alpha_1^0} \right) \left[a \left(x_t^{\alpha_1^0} \right)^{1/2} - b x_t^{\alpha_1^0} - \frac{x_t^{\alpha_1^0}}{4[c_1 + A^{\alpha_1^0}(t)/2]^2} - \frac{x_t^{\alpha_1^0}}{4[c_2 + A^{\alpha_1^0}(t)/2\alpha_1^0]^2} \right], \\
& \hat{W}^{\alpha_1^0(t_0)2} \left(T, x_T^{\alpha_1^0} \right) \exp[-r(T - \tau)] q_2 \left(x_T^{\alpha_1^0} \right)^{1/2}. \tag{8.3.16}
\end{aligned}$$

Тогда $\hat{W}^{\alpha_1^0(t_0)i} \left(t, x_t^{\alpha_1^0} \right)$ есть кооперативный выигрыш игрока i на отрезке $[t, T]$ с весовым коэффициентом α_1^0 .

Утверждение. Функции $\hat{W}^{\alpha_1^0(t_0)1}(t, x) : [\tau, T] \times R \rightarrow R$, удовлетворяющие (8.3.16), имеют вид:

$$\hat{W}^{\alpha_1^0(t_0)1} \left(t, x_t^{\alpha_1^0} \right) = \exp[-r(t - t_0)] \left[\hat{A}_1^{\alpha_1^0}(t) x^{1/2} + \hat{B}_1^{\alpha_1^0}(t) \right], \tag{8.3.17}$$

где

$$\begin{aligned}\dot{\hat{A}}_1^{\alpha_0}(t) &= \left[r + \frac{b}{2}\right] \hat{A}_1^{\alpha_0}(t) - \frac{1}{2[c_1 + A^{\alpha_0}(t)/2]} + \frac{c_1}{4[c_1 + A^{\alpha_0}(t)/2]^2} + \\ &+ \frac{\hat{A}_1^{\alpha_0}(t)}{8[c_1 + A^{\alpha_0}(t)/2]^2} + \frac{\hat{A}_1^{\alpha_0}(t)}{8[c_2 + A^{\alpha_0}(t)/2\alpha_1^0]^2}, \\ \dot{\hat{B}}_1^{\alpha_0}(t) &= r\hat{B}_1^{\alpha_0}(t) - \frac{a}{2}\hat{A}_1^{\alpha_0}(t), \quad \hat{A}_1^{\alpha_0}(T) = q_1, \quad \hat{B}_1^{\alpha_0}(T) = 0.\end{aligned}$$

Доказательство. Вычисляя производные $\hat{W}_t^{\alpha_1^0(t_0)1}(t, x_t^{\alpha_1^0})$ и $\hat{W}_{x_t}^{\alpha_1^0(t_0)1}(t, x_t^{\alpha_1^0})$ и подставляя их в (8.3.16), получаем утверждение п. 8.3.8.

8.3.9. Утверждение. Функции $\hat{W}^{\alpha_1^0(t_0)2}(t, x) : [\tau, T] \times R \rightarrow R$, удовлетворяющие (8.3.16), имеют вид:

$$\hat{W}^{\alpha_1^0(t_0)2}(t, x_t^{\alpha_1^0}) = \exp[-r(t - t_0)] \left[\hat{A}_2^{\alpha_1^0}(t) x^{1/2} + \hat{B}_2^{\alpha_1^0}(t) \right], \quad (8.3.18)$$

где

$$\begin{aligned}\dot{\hat{A}}_2^{\alpha_1^0}(t) &= \left[r + \frac{b}{2}\right] \hat{A}_2^{\alpha_1^0}(t) - \frac{1}{2[c_2 + A^{\alpha_1^0}(t)/2\alpha_1^0]} + \frac{c_2}{4[c_2 + A^{\alpha_1^0}(t)/2\alpha_1^0]^2} + \\ &+ \frac{\hat{A}_2^{\alpha_1^0}(t)}{8[c_1 + A^{\alpha_1^0}(t)/2]^2} + \frac{\hat{A}_2^{\alpha_1^0}(t)}{8[c_2 + A^{\alpha_1^0}(t)/2\alpha_1^0]^2}, \\ \dot{\hat{B}}_2^{\alpha_1^0}(t) &= r\hat{B}_2^{\alpha_1^0}(t) - \frac{a}{2}\hat{A}_2^{\alpha_1^0}(t), \quad \hat{A}_2^{\alpha_1^0}(T) = q_2, \quad \hat{B}_2^{\alpha_1^0}(T) = 0.\end{aligned}$$

Доказательство. Вычисляя производные $\hat{W}_t^{\alpha_1^0(t_0)2}(t, x_t^{\alpha_1^0})$ и $\hat{W}_{x_t}^{\alpha_1^0(t_0)2}(t, x_t^{\alpha_1^0})$ и подставляя их в (8.3.16), получаем утверждение п. 8.3.9.

Используя (8.3.8) и повторяя предыдущее утверждение для подыгр, получаем

$$\begin{aligned}\hat{W}^{\alpha_1^0(t_0)2}(t, x_t^{\alpha_1^0}) \exp[-r(\tau - t_0)] &= \hat{W}^{\alpha_1^0(\tau)1}(t, x_t^{\alpha_1^0}) = \\ &= \exp[-r(t - \tau)] \left[\hat{A}_1^{\alpha_1^0}(t) (x_t^{\alpha_1^0})^{1/2} + \hat{B}_1^{\alpha_1^0}(t) \right], \\ \hat{W}^{\alpha_1^0(t_0)2}(t, x_t^{\alpha_1^0}) \exp[-r(\tau - t_0)] &= \hat{W}^{\alpha_1^0(\tau)2}(t, x_t^{\alpha_1^0}) = \\ &= \exp[-r(t - \tau)] \left[\hat{A}_2^{\alpha_1^0}(t) (x_t^{\alpha_1^0})^{1/2} + \hat{B}_2^{\alpha_1^0}(t) \right].\end{aligned} \quad (8.3.19)$$

Для выполнения свойства индивидуальной рациональности выбранный весовой коэффициент α_1^0 должен удовлетворять условиям

$$\hat{W}^{\alpha_1^0(\tau)1}(\tau, x_\tau^{\alpha_1^0}) \geq V^{(\tau)1}(\tau, x_\tau^{\alpha_1^0}) \quad \text{и} \quad \hat{W}^{\alpha_1^0(\tau)2}(\tau, x_\tau^{\alpha_1^0}) \geq V^{(\tau)2}(\tau, x_\tau^{\alpha_1^0}).$$

Предложим состоятельное во времени (динамически устойчивое) решение кооперативной игры $\Gamma_c(x_0, T - t_0)$. Как это следует из теоремы § 8.3., принцип оптимальности,

согласно которому игроки договариваются выбрать один и тот же весовой коэффициент α_1^0 во всех играх $\Gamma_c \left(x_\tau^{\alpha_1^0}, T - \tau \right)$ и при этом имеют место неравенства

$$\hat{W}^{\alpha_1^0(\tau)1} \left(\tau, x_\tau^{\alpha_1^0} \right) \geq V^{(\tau)1} \left(\tau, x_\tau^{\alpha_1^0} \right), \quad \hat{W}^{\alpha_1^0(\tau)2} \left(\tau, x_\tau^{\alpha_1^0} \right) \geq V^{(\tau)2} \left(\tau, x_\tau^{\alpha_1^0} \right),$$

определяет состоятельное во времени решение кооперативной игры $\Gamma_c(x_0, T - t_0)$.

Условие $\hat{W}^{\alpha_1^0(\tau)i} \left(\tau, x_\tau^{\alpha_1^0} \right) \geq V^{(\tau)i} \left(\tau, x_\tau^{\alpha_1^0} \right)$, $i \in \{1, 2\}$ порождает следующие ограничения на выбор весового коэффициента α_1^0 :

$$\left[\hat{A}_i^{\alpha_1^0}(\tau) \left(x_\tau^{\alpha_1^0} \right)^{1/2} + \hat{B}_i^{\alpha_1^0}(\tau) \right] \geq \left[A_i(\tau) \left(x_\tau^{\alpha_1^0} \right)^{1/2} + B_i(\tau) \right], \quad (8.3.20)$$

при $i \in \{1, 2\}$ и $\tau \in [t_0, T]$.

Заметим, что из утверждений п. 8.3.8 и п. 8.3.9 можно получить $\hat{B}_i^{\alpha_1^0}(\tau) = (a/2r) \hat{A}_i^{\alpha_1^0}(\tau)$, и из утверждения п. 8.3.8 можно получить $B_i(\tau) = (a/2r) A_i(\tau)$. Поэтому, если $\hat{A}_i^{\alpha_1^0}(\tau) \geq A_i(\tau)$, то $\hat{B}_i^{\alpha_1^0}(\tau) \geq B_i(\tau)$, при $i \in \{1, 2\}$ и $\tau \in [t_0, T]$.

Для выполнения (8.3.20) необходимо, чтобы

$$\hat{A}_i^{\alpha_1^0}(\tau) \geq A_i(\tau), \quad i \in \{1, 2\} \text{ и } \tau \in [t_0, T]. \quad (8.3.21)$$

8.3.10. Определение. Обозначим множество всех весовых коэффициентов α_1^t , удовлетворяющих условию $\hat{A}_i^{\alpha_1^t}(t) \geq A_i(t)$, при $i \in \{1, 2\}$ в момент $t \in [t_0, T]$ через S_t . Обозначим через $\underline{\alpha}_1^t$ точную нижнюю границу значений α_1 в S_t , и через $\bar{\alpha}_1^t$ — точную верхнюю границу. Когда t стремится к T , используем обозначения $\bar{\alpha}_1^{T-}$ для $\lim_{t \rightarrow T-} \bar{\alpha}_1^t$, и $\underline{\alpha}_1^{T-}$ для $\lim_{t \rightarrow T-} \underline{\alpha}_1^t$.

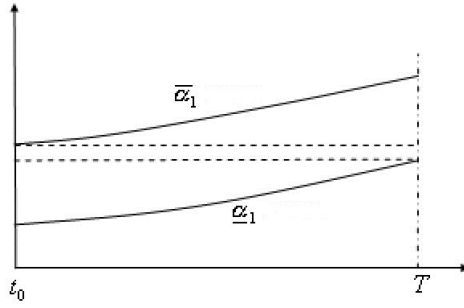
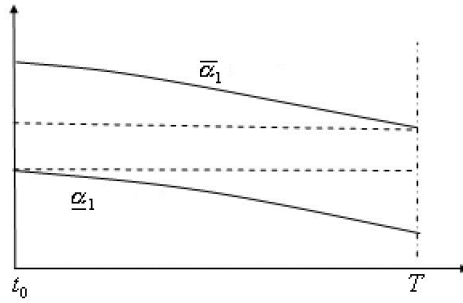
Определение. Определим множество $S^T = \bigcap_{\tau \leq t < T} S_t$ при $\tau \in [t_0, T]$.

S_t — множество весовых коэффициентов α_1 , для которых выполнено условие индивидуальной рациональности в момент $t \in [t_0, T]$ и S^T — множество весовых коэффициентов α_1 , для которых выполнено условие индивидуальной рациональности на всем промежутке $[\tau, T]$. В общем случае $S_\tau^T \neq S_t^T$ при $\tau, t \in [t_0, T]$, где $\tau \neq t$.

Рассмотрим типичные конфигурации S_t . Для нахождения типичных конфигураций S_t при $t \in [t_0, T]$ в игре $\Gamma_c(x_0, T - t_0)$ были проведены численные эксперименты, изменяя параметры $a, b, \sigma, c_1, c_2, q_1, q_2, T, r, x_0$ в широких пределах. Мы вычислили функции $A_1(t), B_1(t), A_2(t)$ и $B_2(t)$ при $t \in [t_0, T]$ из примера 1. Далее были выбраны весовые коэффициенты α_1^t и были вычислены функции $\hat{A}_1^{\alpha_1^t}(t), \hat{A}_2^{\alpha_1^t}(t), \hat{B}_1^{\alpha_1^t}(t)$ и $\hat{B}_2^{\alpha_1^t}(t)$ из утверждений п. 8.3.8 и п. 8.3.9 при $t \in [t_0, T]$. В каждый момент времени $t \in [t_0, T]$ выделяется множество тех весовых коэффициентов α_1^t , при которых имеет место $\hat{A}_i^{\alpha_1^t}(t) \geq A_i(t)$ при $i \in \{1, 2\}$ с целью построения множеств S_t при $t \in [t_0, T]$.

Расположение значений $\underline{\alpha}_1^t$ при $t \in [t_0, T]$ представлено в виде кривой $\underline{\alpha}_1$, и расположение значений $\bar{\alpha}_1^t$ — в виде кривой $\bar{\alpha}_1$. Отметим две основные особенности:

- 1) функции $\underline{\alpha}_1$ и $\bar{\alpha}_1$ непрерывны и, при этом, или обе монотонно возрастают, или обе монотонно убывают;
- 2) множество $S_{t_0}^T$ может оказаться как пустым, так и непустым.

Рис. 8.1. Множество $S_t = [\underline{\alpha}_1^t, \bar{\alpha}_1^t]$ при $S_{t_0}^T \neq \emptyset$ Рис. 8.2. Множество $S_t = [\underline{\alpha}_1^t, \bar{\alpha}_1^t]$ при $S_{t_0}^T \neq \emptyset$

Типичные случаи изображены на рис. 8.1–8.3.

Замечание 5. Рассмотрим случай $S_{t_0}^T \neq \emptyset$. Если $\underline{\alpha}_1$ и $\bar{\alpha}_1$ монотонно убывают, условия $\underline{\alpha}_1^{T-} \notin S_\tau^T$ и $\bar{\alpha}_1^{T-} \in S_\tau^T$ при $\tau \in [t_0, T)$ встречаются чаще. Если $\underline{\alpha}_1$ и $\bar{\alpha}_1$ монотонно возрастают, условия $\underline{\alpha}_1^{T-} \in S_\tau^T$ и $\bar{\alpha}_1^{T-} \notin S_\tau^T$ при $\tau \in [t_0, T)$ встречаются чаще.

Представим ниже одно состоятельное во времени решение игры $\Gamma_c(x_0, T - t_0)$.

Теорема. Если $S_{t_0}^T \neq \emptyset$, принцип оптимальности, выбранный игроками и предпринимающий им выбор весового коэффициента

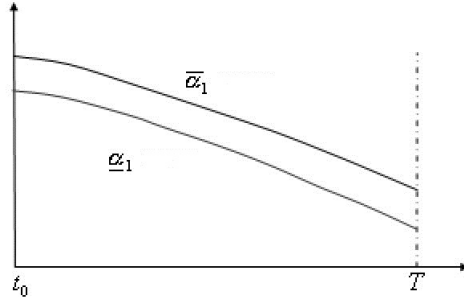
$$\alpha_1^* = \begin{cases} \underline{\alpha}_1^{T-}, & \text{если } \underline{\alpha}_1^{T-} \in S_\tau^T \text{ и } \bar{\alpha}_1^{T-} \notin S_\tau^T, \text{ при } \tau \in [t_0, T], \\ \bar{\alpha}_1^{T-}, & \text{если } \bar{\alpha}_1^{T-} \in S_\tau^T \text{ и } \underline{\alpha}_1^{T-} \notin S_\tau^T, \text{ при } \tau \in [t_0, T], \end{cases} \quad (8.3.22)$$

представляет собой состоятельное во времени решение кооперативной игры $\Gamma_c(x_0, T - t_0)$.

Доказательство. Если хотя бы одно из двух взаимно-исключающих условий, определяющих выбор α_1^* согласно (8.3.22), выполнено, то согласно теореме п. 8.3.7 единственный весовой коэффициент α_1^* может быть выбран для всех подыгр $\Gamma_c(x_\tau^{\alpha_1^*}, T - \tau)$ при $t_0 \leq \tau \leq T$. Вектор

$$\xi^{(\tau)}(x_\tau^{\alpha_1^*}, T - \tau) = \left[\hat{W}^{\tau(\alpha_1^*)1}(\tau, x_\tau), \hat{W}^{\tau(\alpha_1^*)2}(\tau, x_\tau) \right],$$

при $\tau \in [t_0, T]$,

Рис. 8.3. Множество $S_t = [\underline{\alpha}_1^t, \bar{\alpha}_1^t]$ при $S_{t_0}^T \neq \emptyset$

представляет собой парето-оптимальную пару дележей. Поэтому часть 1 условия п. 8.3.6 выполнена.

Поскольку

$$\alpha_1^* \in S_{t_0}^T, \left[\hat{A}_i^{\alpha_1^*}(t) \left(x_t^{\alpha_1^*} \right)^{1/2} + \hat{B}_i^{\alpha_1^*}(t) \right] \geq \left[A_i(t) \left(x_t^{\alpha_1^*} \right)^{1/2} + B_i(t) \right],$$

при $i \in \{1, 2\}$.

Часть 2 условия п. 8.3.6 выполнена.

И наконец из условия (8.3.19) имеем

$$\hat{W}^{\alpha_1^0(t_0)2} \left(t, x_t^{\alpha_1^0} \right) \exp[-r(\tau - t_0)] = \hat{W}^{\alpha_1^0(\tau)1} \left(t, x_t^{\alpha_1^0} \right),$$

$i \in \{1, 2\}, \tau \in [t_0, T],$

т. е. часть 3 условия п. 8.3.6 выполнена. Это и доказывает теорему п. 8.3.10.

§ 8.4. Упражнения и задачи

1. Рассмотрим игру разработки ресурсов двумя игроками. Предполагаем, что игра имеет предписанную продолжительность $[t_0, T]$. Поток ресурсов $x(s) \in X \subset R$ описывается следующим уравнением динамики:

$$\dot{x}(s) = a - bx(s) - u_1(s) - u_2(s), \quad x(t_0) = x_0 \in X,$$

где a и b — константы, $u_1(s), u_2(s)$ — скорости разработки ресурса для игроков 1 и 2 соответственно.

В момент t_0 функции выигрышей игроков определяются следующим образом:

$$\int_{t_0}^T \left[u_1(s) - \frac{c_1}{x(s)} u_1^2(s) \right] \exp[-r(t - t_0)] ds + \exp[-r(T - t_0)] qx(T),$$

$$\int_{t_0}^T \left[u_2(s) - \frac{c_2}{x(s)} u_2^2(s) \right] \exp[-r(t - t_0)] ds + \exp[-r(T - t_0)] qx(T).$$

где q, c_1 и c_2 — константы, причем $c_1 \neq c_2$. Выигрыши трансферабельные, r — ставка дисконтирования.

- а) Найти равновесие по Нэшу в позиционных стратегиях.
- б) Рассмотреть кооперативную форму игры. Найти оптимальные управления.
- в) Рассмотреть кооперативную форму игры. Найти оптимальную траекторию.

2. Рассмотрим кооперативную форму игры из упр.1. Предположим, что игроки договорились о максимизации суммы выигрышей и последующего ее раздела в соответствии с арбитражной схемой Нэша. ПРД должна удовлетворять условию, сформулированному в утверждении п. 8.1.1.

$$\xi^{(\tau)i}(\tau, x_\tau^*) = V^{(\tau)i}(\tau, x_\tau^*) + \frac{1}{2} \left[W^{(\tau)}(\tau, x_\tau^*) - \sum_{j=1}^2 V^{(\tau)j}(\tau, x_\tau^*) \right],$$

- а) Вычислить ПРД.
- б) Показать, что выполнены Парето-оптимальность и индивидуальная рациональность.

3. Рассмотрим игру на бесконечном промежутке времени, в которой игроки 1 и 2 максимизируют свои выигрыши:

$$\int_{t_0}^{\infty} \left[u_1(s) - \frac{c_1}{x(s)} u_1^2(s) \right] \exp[-r(t - t_0)] ds,$$

$$\int_{t_0}^{\infty} \left[u_2(s) - \frac{c_2}{x(s)} u_2^2(s) \right] \exp[-r(t - t_0)] ds.$$

Выигрыши игроков трансферабельные и поток реуров $x(s) \in X \subset R$ определяется динамикой:

$$\dot{x}(s) = a - bx(s) - u_1(s) - u_2(s), \quad x(t_0) = x_0 \in X.$$

- а) Найти равновесие по Нэшу в позиционных стратегиях.
 - б) Найти оптимальные стратегии для кооперативной формы игры.
 - в) Для кооперативной формы игры найти оптимальную траекторию.
 - г) Рассмотрим кооперативную игру, в которой игроки максимизируют суммарный выигрыш и делят его в соответствии с арбитражной схемой Нэша. Эта схема также совпадает с вектором Шепли для кооперативной игры двух лиц. Найти ПРД.
- 4.** Рассмотреть игру из упр.3, но предположить, что выигрыши нетрансферабельные.
- а) Найти парето-оптимальное решение.
 - б) Предложить динамически устойчивое решение, удовлетворяющее аксиоме симметрии.

Литература

Основная

а) учебники

Айзекс Р. Дифференциальные игры. — М.: Мир, 1967.

Васин А. А., Краснощеков П. С., Морозов В. В. Исследование операций. — М.: Изд. центр «Академия», 2008, 464 с.

Васин А. А., Морозов В. В. Теория игр и модели математической экономики. — М.: Изд-во Моск. ун-та, 2005, 272 с.

Воробьев Н. Н. Теория игр для экономистов-кибернетиков. — М.: Наука, 1985.

Дюбин Г. Н., Суздаль В. Г. Введение в прикладную теорию игр. — М.: Наука, 1981, с. 311.

Зенкевич Н. А., Петросян Л. А., Янг Д. В. К. Динамические игры и их приложения в менеджменте. — СПб: Изд-во ВШМ, 2009, 416 с.

Карлин С. Математические методы в теории игр, программировании и экономике. — М.: Мир, 1964, с. 840.

Красовский Н. Н. Управление динамической системой. Задача о минимуме гарантированного результата. — М.: Наука, 1985, с. 469.

Красовский Н. Н., Субботин А. И. Позиционные дифференциальные игры. — М.: Наука, 1974, с. 456.

Мазалов В. В. Математическая теория игр и приложения. — СПб.; Москва; Краснодар: Лань, 2010, 446 с.

Мак-Кинси Дж. К. Введение в теорию игр. — Москва, 1960, с. 371.

Мулен Э. Теория игр с примерами из математической экономики. — М.: Мир, 1985.

Оуэн Г. Теория игр. — Москва, 1971, с. 230.

Петросян Л. А., Данилов Н. Н. Кооперативные дифференциальные игры и их приложения. — Томск, 1985.

Петросян Л. А., Кузютин Д. В. Игры в развернутой форме: оптимальность и устойчивость. — СПб.: Изд-во СПбГУ, 2000.

Петросян Л. А. Дифференциальные игры преследования. — Л., 1977.

Петросян Л. А., Зенкевич Н. А., Семина Е. А. Теория игр. — М.: Высшая школа, 1998.

Печерский С. Л., Яновская Е. Б. Кооперативные игры: решения и аксиомы. — СПб.: Изд-во Европ. ун-та в С.-Петербурге, 2004, с. 459.

Petrosjan L. A., Zenkevich N. A. Game theory. — Singapore, London: World Scientific Publisher, 1996.

Yeung D. W. K., Petrosjan L. A. Cooperative Stochastic Differential Games. — Springer Science+Business Media, Inc., 2006.

б) сборники задач

Коваленко А. А. Сборник задач по теории игр. — Львов: Высшая школа, 1974.

Морозов В. В., Сухарев А. Г. Исследование операций в задачах и упражнениях. — М.: Высшая школа, 1986.

Дополнительная

в) монографии и учебные пособия

Ауман Р., Шепли Л. Значение для неатомических игр. — Принстон: Изд-во Принстонского ун-та, 1974, с. 283.

Ашманов С. А. Линейное программирование. — М.: Наука, 1981, с. 198.

Беллман Р. Динамическое программирование. — М.: ИЛ., 1960.

Берж К. Общая теория игр нескольких лиц. — М.: Физматгиз, 1961, с. 114.

Блекуэлл Д., Гиршик М. Теория игр и статистических решений. — М.: ИЛ., 1958, с. 330.

Бондарева О. Н. О теоретико-игровых моделях в экономике. — Л.: Изд-во ЛГУ, 1974, с. 115.

Воробьев Н. Н. Основы теории игр. Бескоалиционные игры. — М.: Наука, 1984.

Григоренко Н. Л. Дифференциальные игры преследования несколькими объектами. — М.: Изд-во МГУ, 1983, с. 217.

Данилов Н. Н. Игровые модели принятия решений. — Кемерово: КГТУ, 1981.

Данскин Дж. М. Теория максимина. — М.: Сов. радио, 1970, с. 126.

Демьянов В. Ф., Малоземов В. Н. Введение в минимакс. — М.: Наука, 1972.

Гейл Д. Теория линейных экономических моделей. — М.: Мир, 1963, с. 330.

Горелик В. А., Кононенко А. Ф. Теоретико-игровые модели принятия решений в эколого-экономических системах. — М., 1982.

Данчилов Н.Н., Зенкевич Н.А. Неантагонистические игры двух лиц. — Кемерово: Изд-во КемГУ, 1990.

Давыдов Э. Г. Методы и модели теории антагонистических игр. — М.: Изд-во МГУ, 1978, с. 135.

Дрешер М. Стратегические игры. Теория и приложения. — М.: Сов. радио, 1964, с. 186.

Дюбин Г. Н., Суздаль В. Г. Введение в прикладную теорию игр. — М.: Наука, 1981, с. 311.

Жуковский В.И., Салуквадзе М.Е. Некоторые игровые задачи управления и их приложения. — Тбилиси, 1998, с. 462.

Жуковский В.И., Чикрий А.А. Линейно-квадратичные дифференциальные игры. — Киев: Наукова думка, 1994.

Жуковский В.И., Тынянский Н.Т. Равновесные управления многокритериальных динамических систем. — М.: Изд-во МГУ, 1984.

Зенкевич Н.А., Еськова В.А. Конечные антагонистические игры. — Кемерово: Изд-во КемГУ, 1989.

Зубов В.И. Динамика управляемых систем. — М., 1982.

Зубов В.И., Петросян Л.А. Математические методы в планировании. — Л.: Изд-во ЛГУ, 1982.

Клейменов А.Ф. Неантагонистические позиционные дифференциальные игры. — Екатеринбург: Наука, 1993.

Колмогоров А.Н., Фомин С.В. Элементы теории функций и функционального анализа. — М.: Наука, 1981, с. 389.

Кононенко А.Ф., Хафизов А.Д., Чумаков В.В. Принятие решений в условиях неопределенности. — М., 1991.

Красовский Н.Н. Игровые задачи о встрече движений. — М.: Наука, 1970.

Кузютин В.Ф., Зенкевич Н.А., Еремеев В.В. Геометрия: Учебник для вузов. — СПб.: Лань, 2003.

Кукушкин Н.Н., Морозов В.В. Теория неантагонистических игр. — М.: Изд-во МГУ, 1977.

Куржанский А.Б. Управление и наблюдение в условиях неопределенности. — М.: Наука, 1977, с. 325.

Льюис Р.Д., Райфа Х. Игры и решения. Введение и критический обзор. — М.: И.Л., 1961, с. 509.

Малафеев О.А. Управляемые конфликтные системы. — СПб.: Изд-во СПбГУ, 2000, 280 с.

- Мулен Э. Кооперативное принятие решений: аксиомы и модели. — М.: Мир, 1993.
- Партхасаратхи Т., Рагхаван Т. Некоторые вопросы теории игр двух лиц. — Москва, 1974, с. 259.
- Петросян Л. А., Гарнаев А. Ю. Игры поиска. — Л.: Изд-во ЛГУ, 1992, с. 217.
- Петросян Л. А., Захаров В. В. Введение в математическую экологию. — Л.: Изд-во ЛГУ, 1986, с. 295.
- Петросян Л. А., Зенкевич Н. А. Оптимальный поиск в условиях конфликта. — Л.: Изд-во ЛГУ, 1986, с. 96.
- Петросян Л. А., Кузютин Д. В. Устойчивые решения позиционных игр (монография). — СПб.: Изд-во СПбГУ, 2008.
- Петросян Л. А., Томский Г. В. Динамические игры и их приложения. — Л., 1982.
- Петросян Л. А., Томский Г. В. Геометрия простого преследования. — Новосибирск, 1983.
- Печерский С. Л., Беляева А. А. Теория игр для экономистов. Вводный курс. — СПб.: Изд-во Европейского университета, 2001.
- Печерский С. Л., Соболев А. И. Проблема оптимального распределения в социально-экономических задачах и кооперативные игры. — Л.: Наука, 1983, с. 176.
- Подиновский В. В., Ногин В. Д. Парето-оптимальные решения многокритериальных задач. — М., 1982.
- Понтрягин Л. С. К теории дифференциальных игр // Успехи математических наук, 21(4), 1966.
- Понтрягин Л. С., Болтянский В. Г., Гамкрелидзе Р. В., Мищенко Е. Ф. Математическая теория оптимальных процессов. — М.: Наука, 1976.
- Прохоров Ю. В., Рязанов Ю. А. Теория вероятностей. Основные понятия. Предельные теоремы. Случайные процессы. — М.: Наука, 1967, с. 358.
- Рокафеллар Р. Т. Выпуклый анализ. — М.: Мир, 1973, с. 470.
- Розенмюллер И. Кооперативные игры и рынки. — М.: Мир, 1974, с. 115.
- Сансоне, Дж. Обыкновенные дифференциальные уравнения. Т. 2. — М.: ИЛ., 1954, с. 269.
- Слобожанин Н. М. Информация и управление в динамических играх. — СПб.: Изд-во СПбГУ, 2002, с. 308.
- Субботин А. И., Ченцов А. Г. Оптимизация гарантии в задачах управления. — М.: Наука, 1981, 288 с.
- Субботин А. И. Обобщенные решения дифференциальных уравнений 1-го порядка. Перспективы динамической оптимизации. — М.: Ижевск, 2003, с. 336.

Тироль Ж. Рынки и рыночная власть: теория организации промышленности. В 2-х т.; пер. с англ.; под ред. Гальперина В. М. и Зенкевича Н. А. — СПб.: Экономическая школа, 2000.

Тынянский Н. Т., Жуковский В. И. Итоги науки и техники. Математический анализ. Т. 10. — М.: ВИНТИ, 1979.

Феллер В. Введение в теорию вероятностей и ее приложения. Т. 1, 2. — М.: Мир, 1984, с. 1230.

Фон Нейман, Дж. и О. Моргенштейн. Теория игр и экономическое поведение. — М.: Наука, 1970, с. 625.

Харшаньи Дж., Зельтен Р. Общая теория выбора равновесия в играх / пер. с англ.; под ред. Зенкевича Н. А. — СПб.: Экономическая школа, 2001.

Ху Т. Целочисленное программирование и потоки в сетях. — М.: Мир, 1974, с. 411.

Черноусько Ф. Л., Меликян А. А. Игровые задачи управления и поиска. — М.: Наука, 1978.

Aumann R. J. Acceptable points in general cooperative n -person games // Contributions to the Theory of Games IV ed. by Luce R.D., Tucker A.W. — Princeton: Princeton University Press, 1959, P. 287-324.

Aumann R. J. Game theory // The New Palgrave, A Dictionary of Economics. — London Macmillan, 1987, Vol. 2, P. 460-482.

Basar T., Olsder G. J. Dynamic Noncooperative Game Theory, 2nd Edn. — London: Academic Press, 1995.

Bellman R. Dynamic Programming. — Princeton, Princeton University Press, NJ, 1957.

Bierman H. S., Fernandez L. Game Theory with Economic Applications. — Addison, Wesley Publishing Company, Inc., Massachusetts, USA, 1993.

Binmore K. Fun and Games. A Text on Game Theory. D. C. Heath and Company. — Lexington, Massachusetts, USA, 1992.

Brams, S. J. Theory of Moves. — Cambridge Univ. Press, 1994.

Burger, E. Introduction to the Theory of Games. — Prentice-Hall, Englewood Cliffs, N.Y., 1963, p. 211.

Dockner E., Jorgensen S., Long N.V., Sorger, G. Differential Games in Economics and Management Science. — Cambridge University Press, Cambridge, 2000.

Dresher M. Games of strategy, theory and applications. — N.Y.: Prentice-Hall, Englewood Cliffs, 1961.

Friedman, A. Differential Games. — Wiley, N.Y., 1971, p. 350.

Fishburn P. C. Utility theory for decision making. — New York; London; Sydney; Toronto: Wiley, 1970.

- Fleming W. H., Rishel R. W.* Deterministic and Stochastic Optimal Control. — Springer-Verlag, New York, 1975.
- Fudenberg D., Tirole J.* Game theory. — Mass: MIT Press, 1991.
- Gibbons R.* Game Theory for Applied Economists. — Princ. Univ. Press, Princeton, New Jersey, 1992.
- Isaacs R.* Differential Games. — Wiley, New York, 1965.
- Karlin, S.* Reduction of certain classes of games to integral equations. — Princeton (N.Y.), 1953.
- Karlin, S.* Mathematical Methods and Theory in Games, Programming and Economics. — Pergamon Press, London, 1959, p. 840.
- Kohlberg E.* Refinement of Nash equilibrium: The main ideas. Mimeo. — Harvard University, 1989.
- Krasovskii A. N., Krasovskii N. N.* Control under lack of information. — Birkhauser, Boston, 1995, p. 320.
- Kreps D. M.* Game theory and economic modeling. — Oxford: Oxford Univ. Press, 1990.
- Leitmann G.* Cooperative and Non-Cooperative Many Players Differential Games. — Springer-Verlag, New York, 1974.
- Maynard, S. J. and G. R. Price.* The logic of animal conflict. — Nature, London, 1973.
- McMillan J.* Games, Strategies and Managers. — Oxford University Press, 1992.
- Myerson R. B.* Game theory. — Cambridge: Massachusetts, Harvard University Press, 1991.
- Neumann J. von, Morgenstern O.* Theory of games and economic behavior. — Princeton: Princeton University Press, 1944.
- Owen G.* Game theory. — Philadelphia: W.B. Saunders Company, 1968.
- Peleg, B. and P. Sudholter.* Introduction to the theory of cooperative games. — Springer, Second ed., 2007, 328 pp.
- Petrosyan L. A.* Differential Games of Pursuit. — World Scientific, Singapore, 1993.
- Rosenmüller J.* The theory of games and markets. — Amsterdam, 1981.
- Roth A. E.* Axiomatic models of bargaining. — Berlin: Springer-Verlag, 1979.
- Schelling T. C.* The strategy of conflict. — Mass.: Harvard University Press, 1960.
- Shubik M.* Strategy and Market Structure. — New York: Wiley, 1959.
- Thomson W., Lensberg T.* Bargaining theory: Axiomatic approach. — San Diego: Academic Press, 1990.

van Damme, E. Stability and Perfection of Nash equilibria. — Springer-Verlag, Berlin, 1991, p. 215.

Weibull J.W. Evolutionary Game Theory. — MIT Press, Cambridge, 1995.

г) статьи

Бондарева О. Н. Некоторые применения методов линейного программирования к теории кооперативных игр // Проблемы кибернетики. Вып. 10, 1963, С. 119–139.

Э. М. Вайсборд, В. И. Жуковский. Введение в дифференциальные игры нескольких лиц и их приложения. — М.: Сов. радио, 1980, с. 303.

Воробьев Н. Н. Современное состояние теории игр // Успехи мат. наук. 1970. — Т. 25. — № 2. — С. 69–90.

Зенкевич Н. А., Козловская Н. В. Устойчивый вектор Шепли в задаче экологического производства // Математическая теория игр и ее приложения. Вып. 1, 2010.

Клейменов А. Ф. К кооперативной теории бескоалиционных позиционных дифференциальных игр // Докл. АН СССР, 1990. Т. 32.— № 1.— С. 32–35.

Клейменов А. Ф. О решениях в неантагонистической позиционной дифференциальной игре // Прикладная математика и механика, 61:5, 1997, С. 739–746.

А. Ф. Кононенко. О равновесных позиционных стратегиях в неантагонистических дифференциальных играх // ДАН СССР, 231(2), 1976.

Красовский Н. Н., Котельникова А. Н. О дифференциальной игре на перехват // Труды математического института им. В. А. Стеклова, 2010. — Т. 268. — С. 168–214.

Мазалов В. В., Сакагучи М. Равновесие в бескоалиционной игре n лиц с выбором момента времени // Математическая теория игр и ее приложения. — Т. 1. Вып. 1. — 2009, С. 65–85.

О. А. Малафеев. О существовании ситуации равновесия в дифференциальных бескоалиционных играх двух лиц с независимыми движениями // Вестник ЛГУ, № 7, 1980.

Меликян А. А. Сингулярные характеристики в граничных условиях уравнения Гамильтона – Якоби // Доклады Академии Наук, 2005. — Т. 404. — Вып. 4.

Местников С. В. Аппроксимация области неопределенности в дифференциальных играх поиска // Дифференциальные уравнения, 1992. — Т. 28.— № 6.— С. 967–972.

Никитин Ф. Ф., Чистяков С. В. О дифференциальных антагонистических играх с неограниченной продолжительностью // Вестник СПбГУ. Сер. 1. — Вып. 3. — 2004. — С. 37–43.

Париллина Е. М. Кооперативная игра передачи данных в беспроводной сети // Математическая теория игр и ее приложения. — Т. 1. — Вып. 4. — 2009. — С. 93–110.

Пек Дж., Далмидж А.Л. Игры на компактном множестве // Сборник статей. — С.3.

Петров Н.Н. О существовании значения игры преследования // ДАН СССР. — Т. 190. — № 6. — 1970. — С. 1289–1291.

Петров Н.Н. Доказательство существования значения игры преследования с ограниченным временем // Дифференциальные уравнения. Т. 6. — № 5. — 1970. — С. 784–797.

Петросян Л.А. Устойчивость решений в дифференциальных играх со многими участниками // Вестн. ЛГУ. Сер. 1. — Вып. 4. — № 19. — 1977. — С. 46–52.

Петросян Л.А. Характеристические функции кооперативных дифференциальных игр // Вестн. СПбГУ. Сер. 1. — Вып. 1. — 1995. — С. 48–52.

Петросян Л.А., Баранова Е.М., Шевкопляс Е.В. Многошаговые кооперативные игры со случайной продолжительностью // Сборник науч. трудов «Оптимальное управление и дифференциальные игры» в Тр. Инст. мат-ки и мех-ки. — Екатеринбург: Изд-во УрО РАН. — Т. 10. — № 2. — 2004. — С. 116–130.

Петросян Л.А., Данилов Н.Н. Устойчивость решений в неантагонистических дифференциальных играх с трансферабельными выигрышами // Вестн. ЛГУ. — № 1. — 1979. — С. 46–54.

Л. А. Петросян, Н. А. Зенкевич. Принципы устойчивой кооперации // Математическая теория игр и ее приложения. Т. 1. — Вып. 1. — 2009. — С. 102–117.

Петросян Л.А., Мамкина С.И. Игры с переменным коалиционным разбиением // Вестник СПбГУ. Сер. 1. — Вып. 3. — 2004. — С. 60–69.

Петросян Л.А., Шевкопляс Е.В. Кооперативные дифференциальные игры со случайной продолжительностью // Вестн. СПбГУ. Сер. 1. — Вып. 4. — 2000. — С. 23–28.

Садовский, А. Л. Монотонный итеративный алгоритм решения матричных игр // ДАН СССР. — 238(3). — 1978.

Сайон М. Некоторые общие теоремы о минимаксах // Сборник статей. — Сер. 3, С. 40 — 46.

Сайон М., Вульф Ф. Об игре, не обладающей значением // Сборник статей. — Сер. 4, С. 290 — 300.

Чистяков С.В. О построении сильно динамически устойчивых решений кооперативных дифференциальных игр // Вестн. СПбГУ. Сер. 1. — № 1. — 1992. — С. 50–54.

Чистяков С.В. Динамический аспект решения классических кооперативных игр // Докл. РАН. Т. 330. — № 6. — 1993. — С. 707–709.

Шевкопляс Е.В. Уравнение Гамильтона-Якоби-Беллмана в дифференциальных играх со случайной продолжительностью // Управление большими системами. 26.1. — М.: ИПУ РАН. — 2009. — С. 385–408.

Яновская Е. Б. О существовании значения антагонистических игр с полунепрерывными функциями выигрыша // Изв. АН СССР. — Техн. кибернетика. — № 6. — 1973. — С. 56–60.

Яновская Е. Б. Антагонистические игры // Проблемы кибернетики. Вып. 34. — М.: Наука. — 1978. — С. 221–246.

Baranova E. M., Petrosjan L. A. Cooperative Stochastic Games in Stationary Strategies // Game theory and Applications. — Nova Science Publishers. Vol. 11. — 2006. — P. 7–17.

Breton M., Zaccour G., Zahaf M. A differential game of joint implementation of environmental projects // Automatica. Vol. 41. — № 10. — 2005, P. 1737–1749.

Domansky V. Dynkin games with randomized optimal stopping rules // Annals of the International Society of Dynamic Games. Vol. 7. — Birkhauser, Boston, MA, USA, 2004, P. 247–262.

Harsanyi, J. C. International Economic Review, 4, 1963.

Harsanyi, J. C. and R. Selten. Management Science, 18, 1972.

Hart, S. and A. Mas-Colell. In A. E. Roth, editor, The Shapley Value. — Cambridge Univ. Press, Cambridge, 1988.

Haurie A. A note on nonzero-sum differential games with bargaining solutions // Journal of Optimization Theory and Application, 1976, 18, P. 31–39.

Haurie A., Zaccour G. Differential game models of global environmental management // Annals of Dynamic Games. — Boston, 1994, P. 124–132.

Haigh, J. Adv. Applied Prob., 7, 1975.

Jorgensen S., Yeung D. W. K. Stochastic differential game model of a common property fishery // Journal of Optimization Theory and Applications, 1996, 90, P. 391–403.

Jorgensen S., Yeung D. W. K. Inter and intragenerational renewable resource extraction. — Annals of Operations Research, 1999, 88, P. 275–289.

S. Jorgensen, G. Martin-Herran, G. Zaccour. Agreeability and Time Consistency in Linear-State Differential Games // Journal of Optimization Theory and Applications, Vol. 119, №1, 2003, P. 49–63.

Kaitala V. Equilibria in a stochastic resource management game under imperfect information // European Journal of Operational Research, 71, 1993, P. 439–453.

Kalai, E. and M. Smorodinsky. Econometrica, 43, 1975.

Karlin, S. and R. Restrepo. In H. Kuhn and A. Tucker, editors, Contributions to the Theory of Games. — N.Y.: Princeton Univ. Press, Princeton, 1957.

Kidland F. E., Prescott E. C. Rules rather than decisions: the inconsistency of optimal plans // J. of Political Economy, 1977, Vol. 85, P. 473–490.

Kohlberg E., Mertens J.-F. On the strategic stability of equilibria // *Econometrica*. Vol. 54. — 1986. — P. 1003–1039.

Krawczyk J., Zaccour G. Management of pollution from decentralized agents by the local government // *International Journal of Environment and Pollution*. V. 12. — №2/3. — 1999. — P. 343–357.

Kreps D. M., Ramey G. Structural consistency, consistency and sequential rationality // *Econometrica*. Vol. 55. — 1987. — P. 1331–1348.

Kuhn H. W. Extensive games and the problem of information // *Annals of Mathematics Studies*. Vol. 28. — P. 193–216.

Leitmann G., Schmitendorf W. E. Profit maximization through advertising: nonzero-sum differential game approach // *IEEE Transactions on Automatic Control*, Vol. 23. — 1978. — P. 645–650.

Marín-Solano J. and Navas J. Non-constant discounting in finite horizon: the free terminal time case // *Journal of Economic Dynamics and Control*, 33, 2009, 666–675.

Marin-Solano J. and Shevkoplyas E. V. Non-constant discounting in differential games with random time horizon // *Automatica*, 2011.

Melikyan A., Olsder G., Akhmetzhanov A. Solution of a Nonzero-Sum Game via Reduction to a Zero-Sum Game // *International Game Theory Review*. Vol. 10(4). — 2008. — P. 437–459.

Myerson R. B. Refinements of the Nash equilibrium concept // *International Journal of Game Theory*. Vol. 7. — 1978. — P. 73–80.

Nash J. F. Equilibrium points in n -person games // *Proc. Nat. Acad. Sci., USA*. Vol. 36. — 1950. — P. 48–49.

Nash J. F. The bargaining problem // *Econometrica*. Vol. 18. — 1950. — P. 155–162.

Nash, J. Non-cooperative games // *Annals of Mathematics*. Vol. 54. — 1951. — P. 286–295.

Nash J. F., Jr. Two-person cooperative games // *Econometrica*. V. 21. — 1953. — P. 128–140.

Neumann J. von. Zur Theorie der Gesellschaftsspiele // *Math. Ann.* Vol. 100. — 1928. — P. 295–320.

Perles, M. A. and M. Mashler. *International Journal of Game Theory*, 10, 1981.

Petrosjan, L. A., A. Azamov and H. Satimov. *Controlled Systems*, 13, 1974.

Petrosjan L. A. The Regularization of NB-scheme in differential games // *Dynamics and Control*. Vol. 5. — 1995. — P. 31–35.

Petrosjan L. A. The Shapley value for differential games // *New Trends in Dynamic Games and Applications* / eds G.Y. Olsder. — Boston: Birkhauser, 1996.

Petrosyan L. A. Agreeable solutions in differential games // International Journal of Mathematics, Game Theory and Algebra. Vol. 7. — 1997. — P. 65–177.

Petrosjan L. A. The time-consistency problem in nonlinear dynamics // RBCM — J. of the Braz. Soc. Mechanical Sciences, 1997. — Vol. 19. — № 2. — P. 281–303.

Petrosyan L. A. Bargaining in dynamic games. In: Petrosyan, L., Yeung, D. (ed) ICM Millennium Lectures on Games, 2003. — Springer-Verlag, Berlin. — P. 139–143.

Petrosjan L. A. Cooperative Stochastic Games // Advances in Dynamic Games, Annals of the International Society of Dynamic Games, Application to Economics, Engineering and Environmental Management, ed. by A. Haurie, S. Muto, L. A. Petrosjan, T.E.S. Raghavan, 2006. — P. 139–146.

Petrosjan L., Mamkina S. Dynamic games with coalitional structures // International Game Theory Review. Vol. 8. — 2006. — №2. — P. 295–307.

Petrosjan L.A., Shevkoplyas E. V. Cooperative Solutions for Games with Random Duration // Game Theory and Applications. Vol.IX, — 2003. — Nova Science Publishers. — P. 125–139.

Petrosjan L. A., Zaccour G. Time-consistent Shapley value allocation of pollution lost reduction // Journal of Economic Dynamic & Control. — Vol. 27. — 2003. — P. 381–398.

Roth, A. E. Mathematics of operations research, 2, 1977.

Robinson, G. B. An iteration method of solving a game. — RAND Corp. Vol. P-154. — 1950. — P. 9.

Rosen J. B. Existence and uniqueness of equilibrium points for concave n -person games // Econometrica. Vol. 33. — 1965. — P. 520–534.

Sakaguchi, M. Oper. Res. Soc. Jap., 16, 1973.

Selten, R. Reexamination of the perfectness concept for equilibrium points in extensive games // International Journal of game theory, 4, 1975.

Shapley L. S. A value for n -person games // Contributions to the Theory of Games II / eds Luce R.D. and Tucker A.W. — Princeton: N.J. Princeton University Press, 1953, P. 307–317.

Shevkoplyas E. V. The Shapley Value in cooperative differential games with random duration // Advances in Dynamic Games. Vol. 11. — Edt. by M.Breton and K. Szajowski. — Springer's imprint Birkhauser, Boston, 2011.

Subbotina N.N. The method of characteristics for Hamilton-Jacobi equations and applications to dynamical optimization // Journal of math. sciences. Vol. 135. — №3, 2955–3091, DOI: 10.1007/s10958-006-0146-2.

Van Damme E. E. C. A relation between perfect equilibria in extensive form games and proper equilibria in normal form games // Intern. J. Game Theory. Vol. 13. — 1984. — P. 1–13.

Yanovskaya E. Lexicographical Maximin Core Solutions. In: Constructing Scalar-Valued Objective Functions, Lecture Notes in Economics and Math. Systems, 450, Springer-Verlag, 1996, P. 250–261.

Yanovskaya E. Nonsymmetric consistent surplus sharing methods // International Journal of Mathematics, Game Theory and Algebra. Vol. 14. — 2004. — № 3. — P. 189–203.

Yeung D. W. K. On differential games with a feedback Nash equilibrium // Journal of Optimization Theory and Applications, 1994. 82. — № 1. — P. 181–188.

Yeung D. W. K. Infinite horizon stochastic differential games with branching payoffs // Journal of Optimization Theory and Applications. Vol. 111. — № 2. — 2001. — P. 445–460.

Yeung D. W. K. Nontransferable individual payoff functions under stochastic dynamic cooperation // International Game Theory Review, Vol. 6. — 2004. — P. 281–289.

Yeung D. W. K., Petrosyan L. Proportional time-consistent solution in differential games. In: Yanovskaya E.B. (ed) International Conference on Logic, Game Theory and Social Choice. — St Petersburg State University, 2001. — P. 254–256.

Yeung D. W. K. An irrational-behavior-proofness condition in cooperative differential games // Int. J. of Game Theory Rev. Vol. 9. — № 1. — 2007.— P. 256–273.

Zenkevich, N. A. and S. N. Voznyuk. In M. Breton and G. Zaccour, editors, 6th International Symposium on Dynamic Games and Applications, Preprint Vol. — Montréal, Canada: École des Hautes Études Commerciales, 1994.

Предметный указатель

C -ядро, 165

NM -решение, 170

ε , δ -равновесие, 304

i -равновесие по Штакельбергу, 128

ESS, 146, 222

А

Айзекса–Беллмана уравнение, 310

альтернативы, 210

арбитражная схема Нэша, 153

асимметричное равновесие Нэша, 157

Б

бескоалиционная игра, 119

бесконечные игры, 55

бесшумная дуэль, 57

биматричная игра, 120

В

вектор Банзафа, 185

вектор Шепли, 174

верхнее значение, 17

верхнее значение игры, 58

вогнутая игра, 76

вполне смешанная стратегия, 43, 130

выпуклая игра, 76

выпуклая оболочка, 26

выпуклая функция выигрыша, 76

выпуклое многогранное множество, 26

выпуклое множество, 26

выпуклый конус, 27

Г

гипотетическое рассогласование, 307

Д

диктаторское решение, 157

дифференциальная игра с предписанной продолжительностью, 296

дуэль, 14, 91, 93

З

задача о переговорах, 153

значение игры, 19, 291

И

игра двух лиц, 120

игра качества, 290

игра на достижение минимального результата, 290

игра на уклонение, 14

игра поиска, 15, 55, 93

игра преследования, 285

игра преследования на быстроедействие, 301

игра преследования при наличии сил трения, 321

игра с интегральным выигрышем, 290

игра с нулевой суммой, 12

игра с полной памятью, 213

игра с постоянной суммой, 161

игра с терминальным выигрышем, 290

игры на единичном квадрате, 58

интегральный выигрыш, 290

информационное множество, 206

К

коалиция, 159

конус, 27

крайняя точка, 26

кусочно-программные стратегии, 288

М

максиминная стратегия, 16
 матричные игры, 12
 минимальный результат, 289
 многошаговая позиционная игра, 210
 множество достижимости, 297

Н

непрерывная игра, 70
 нижнее значение игры, 16, 58
 носитель игры, 173

О

одновременные многошаговые игры, 216
 оптимальная программная стратегия, 309

П

Парето оптимальность, 127
 принцип максимума Понтрягина, 338
 переговорное множество, 153
 повторяющиеся эволюционные игры, 220
 подыгра, 12
 поиск, 93
 покер, 98
 правильное равновесие, 140
 принцип максимина, 16
 простое движение, 286, 320

Р

равновесие в совместных смешанных стратегиях, 150
 равновесие по Нэшу, 124, 131
 равновесная ситуация, 18
 решение Калаи–Смородинского, 157

С

СКПСР, 324
 седловая точка, 18, 28
 сильное равновесие, 126
 симметричная игра, 46, 221
 симметрия в игре, 221
 синтезирующие стратегии, 287, 292
 ситуация в игре, 12
 смешанная стратегия, 23, 63, 129

смешанные кусочно-программные стратегии поведения (СКПСР), 324
 совершенное равновесие, 138
 совместная смешанная стратегия, 149
 спектр смешанной стратегии, 130
 стратегия поведения, 212
 существенная стратегия, 36
 существование ситуации равновесия, 20

У

условно-оптимальная траектория, 327
 утилитарное решение, 157

Ф

функция выигрыша, 12, 55, 289
 функция потенциала, 179

Х

характеристическая функция, 159

Ч

чистая стратегия, 211

Э

эволюционно устойчивая стратегия (ESS), 146, 222
 эволюционные игры, 220
 эгалитарное решение, 157