Санкт-Петербургский Государственный Университет
Факультет прикладной математики — процессов управления
Кафедра математической теории игр и статистических решений

Лахина Юлия Эдуардовна

Выпускная квалификационная работа

Решение задачи длительной эксплуатации нескольких ресурсов методами дифференциальных игр

Уровень образования: магистратура
Направление: 01.04.02 «Прикладная математика и информатика»
Основная образовательная программа: ВМ.5504 «Исследование операций и системный анализ»

Заведующий кафедрой,
доктор физ.-мат.н.,
профессор
Петросян Л. А.

Научный руководитель,
доктор физ.-мат.н.,
профессор
Громова Е. В.

Рецензент,
кандидат физ.-мат.н.,
доцент
Соловьева Н. А.

Санкт-Петербург
2020

Оглавление

Введение

В настоящее время модели управления ресурсами имеют широкое применение в таких важных областях как экономика и экология, поскольку позволяют решать актуальные проблемы.

Грамотная рекламная политика является основным методом привлечения целевой аудитории, и, как следствие, неотъемлемой частью успешного функционирования компании, поэтому вопрос об управлении инвестициями в рекламную кампанию становится актуальным. Также как и вопрос об управлении объемами вредных выбросов в окружающую среду в связи с современной неблагоприятной экологической ситуацией.

В данной работе исследуются теоретико-игровые модели управления ресурсами с многомерной фазовой переменной на бесконечном временном промежутке с постоянной ставкой дисконтирования. В качестве ресурсов рассматриваются инвестиции в рекламу и количество выбросов в окружающую среду.

В первой главе приводятся основные модели и методы, используемые в ВКР. Формулируется постановка задачи оптимального управления на бесконечном промежутке времени с интегральной формой выигрыша и дисконтированием подынтегральной функци, рассматривается метод ее решения, а также критерии для отбора допустимых решений. Поиск оптимальных управлений в кооперативной дифференциальной игре также может быть описан в форме задачи оптимального управления, поскольку игроки объединяются с целью максимизации общего суммарного выигрыша.

Проблема распределения суммарного выигрыша между игроками, в том числе, с учетом фактора времени, описана с точки зрения проблемы динамической устойчивости. Распределение компонент дележа игроков предлагается осуществить согласно процедуре распределения дележа (ПРД).

Кроме того, в главе I формализуется постановка сетевой дифференциальной игры.

Во второй главе изучена кооперативная дифференциальная игра управления инвестициями в рекламную кампанию для случая n симметричных игроков, которые конкурируют за объем собственных продаж некоторого однородного продукта с учетом амортизации, свойственной рынку. В третьей главе предложена модель управления объемами вредных выбросов при производстве взаимозаменяемых товаров для двух симметричных игроков при отсутствии абсорбции. Дифференциальная игра изучается как в кооперативной постановке, для которой также находится распределение общего выигрыша игроков, так и в некооперативной, в которой рассматривается вопрос существования квадратичного решения задачи. Показывается, что решение в обеих моделях является неединственным, поэтому осуществляется отбор допустимых решений с помощью критериев, предложенных в первой главе. В главах II и III также изучается проблема динамической устойчивости, т.е. реализации при долгосрочном процессе. Для решения данной проблемы используются схемы ПРД из главы I.

В последней главе математическая модель управления вредными выбросами формулируется как дифференциальная игра на сети, в которой находится равновесие по Нэшу.

Основные цели и задачи

Основной целью данной работы является исследование линейно - квадратичной дифференциальной игры с интегральным функционалом при условии бесконечного временного горизонта, дисконтированием функции полезности и, что существенно отличает данную работу от большинства известных широко изученных приложений линейно-квадратичных игр, с **многомерной** фазовой переменной. Поскольку в задачах такого типа возникает неединственность решений, достаточно широко встречающаяся в экономических работах (см., например, [27]), необходимо рассмотреть вопрос отбора допустимых решений.

Кроме того, актуальным вопросом является распределения полученного суммарного выигрыша в случае кооперации игроков между игроками, а также реализация выбранного игроками способа распределения выигрыша (кооперативного решения) на всем временном промежутке. Таким образом, при длительной эксплуатации ресурсов проблема динамической устойчивости выбранного кооперативного решения становится очень важной и назревает необходимость исследования и решения данной проблемы.

Еще одной целью работы являлась формализация сетевой постановки для дифференциальной игры, сформулированной Л.А. Петросяном в [29] в общем виде, для предложенной модели управления объемами вредных выбросов. Предполагается, что некоторые предприятия объединены “связями”, т.е. общими интересами, которые могут быть основанием для рассмотрения сетевой структуры.

В связи с поставленной целью формулируются следующие задачи:

1) Изучить основные методы, используемые для решения рассматриваемых моделей управления различными типами ресурсов.

2) Формализовать задачу оптимального управления инвестициями в рекламу в кооперативной постановке для n игроков, изучить различные способы отбора допустимых решений из множества полученных решений, а также рассмотреть вопрос распределения между игроками выигрыша, полученного в результате кооперации, в соответствии с принципом динамической устойчивости. Получить аналитические выражения для оптимальных управлений, траектории, значения функционала суммарного выигрыша, а также формулы для процедуры распределения дележа для случая полностью симметричных игроков.

3) Сформулировать модель управления объемами вредных выбросов как дифференциальную игру с многомерной фазовой переменной для двух игроков, рассмотреть игру в некооперативной и кооперативной постановке, найти решение задачи в аналитическом виде. Интерпретировать полученные результаты графически. Получить аналитические выражения для процедуры распределения дележа как средства решения проблемы динамической устойчивости кооперативного решения для случая симметричных игроков.

4) Построить сетевую модель управления объемами вредных выбросов трех игроков в случае образования коалиции двоих из них. Найти управления, максимизирующие выигрыш коалиции.

Обзор литературы

Изучение линейно-квадратичных дифференциальных игр [1,2] с многомерной фазовой переменной [3–7] представляет особый интерес, поскольку как в кооперативной, так и в некооперативной постановке задач возникает неединственность решения. Данный вопрос часто находит отражение в экономических приложениях [8–10].

В статье [11], отражающей результаты автора настоящей магистерской диссертации, рассматривается модель управления инвестициями в рекламу с n симметричными игроками в двух постановках: кооперативной, в которой были найдены оптимальные управления и некооперативной, в которой находилось равновесие по Нэшу [12]. Решение линейно-квадратичной задачи [13] разыскивается в классе позиционных стратегий [2,14]. Для определения оптимального поведения игроков использован теоретико-игровой подход [12, 15, 16].

В классе линейно-квадратичных дифференциальных игр возникают вопросы существования равновесия по Нэшу [3, 11, 14, 17, 18, 20–22]. Кроме того, проблема существования и единственности оптимального управления была рассмотрена в [23–25]. Модель управления выбросами была предложена в работе [26], где в некооперативной постановке было показано, что не существует квадратичных решений задачи нахождения равновесия по Нэшу.

В работе [27] предложен подход отбраковки несостоятельных решений в случае их неединственности, используемый в экономических приложениях. Также в книге [28] описан еще один метод отбора допустимого решения — классический метод, используемый для линейно-квадратичных задач оптимизации (LQR).

В статье Л.А. Петросяна [29] формулируется сетевая дифференциальная игра n игроков, где на выигрыш игрока влияют лишь те игроки, которые связаны с ним дугой. Кроме того, в [29] были введены классы стратегий, включающих в себя возможность прерывания связи игроком с другими игроками в процессе игры.

В данной работе рассматриваются математические модели управления ресурсами с многомерной фазовой переменной, где в качестве ресурсов рассматриваются инвестиции в рекламу [11] и объемы производства продукции, пропорциональные выбросам в окружающую среду [26]. Задача решается методом динамического программирования [14, 30]. Кооперативный выигрыш распределяется в соответствии с концепцией динамической устойчивости согласно процедуре распределения дележа (ПРД), впервые введенной Л.А. Петросяном в 1977 г. и описанным для задач с дисконтированием и бесконечной продолжительностью в [12].

Результаты, полученные в ходе подготовки магистерской диссертации, частично были опубликованы в работах:

1. Громова Е. В., Громов Д. В., Лахина Ю. Э. О дифференциальной игре управления инвестициями в рекламную кампанию. Труды института математики и механики УРО РАН. Т. 24, № 2. 2018. С. 64-75.

2. Gromova E., Lakhina Yu. On the selection of the Nash equilibria in a linear-quadratic differential game of pollution control. Frontiers of Dynamic Games. Ed. by L. Petrosyan, V. Mazalov, N. Zenkevich. St. Petersburg. 2019. P. 37–48.

3. Лахина Ю. Э. Оптимальное управление инвестициями в рекламу на рынке однородной продукции. Процессы управления и устойчивость. 2018. Т. 5. № 1. С. 470–474.

4. Лахина Ю. Э. О существовании равновесия по Нэшу в некооперативной дифференциальной игре управления объемами вредных выбросов. Процессы управления и устойчивость. 2020.

Глава 1.

Основные методы

Будем рассматривать дифференциальную игру [12] $\Gamma(t_0, x_0, \infty)$ n лиц на бесконечном промежутке времени, $N = \{1, \ldots, n\}$, $|N| = n$, как конфликтно-управляемый процесс, где n—количество игроков, t_0 — начальный момент времени, $t_0 \geqslant 0$ и x^0 — начальное состояние игры.

Динамика игры задается системой обыкновенных дифференциальных уравнений с начальным условием:

$$\dot{x}(t) = f(x(t), a_1(t), \ldots, a_n(t)), \quad x(t_0) = x^0, \tag{1.1}$$

где $x(t) \in X \subseteq \mathbb{R}^m$, $t \in [t_0, \infty)$, а a_i — управление i-го игрока из его множества допустимых управлений $\mathcal{U}_i$. $\mathcal{U}_i$, $i = \overline{1, n}$ состоят из множества всех измеримых функций на $[t_0, \infty)$ в U_i. Множество допустимых значений управлений U_i i-го игрока представляет собой выпуклое компактное подмножество $\mathbb{R}^k$, такое что $\{0\} \in U_i$.

Введем обозначения $a(t) = (a_1(t), \ldots, a_n(t))$, $\mathcal{U} = \mathcal{U}_1 \times \ldots \times \mathcal{U}_n$, $U = U_1 \times \ldots \times U_n$, то есть $a \in \mathcal{U}$, $a(t) \in U$, $t \in [t_0, \infty)$.

Функция $f(x(t), a(t))$ непрерывна на множестве $X \times U_1 \times \ldots \times U_n$, липшицева по x, дифференцируема. Тогда будем говорить, что выполнены условия существования и единственности решений [31] системы дифференциальных уравнений (1.1) для любого набора допустимых управлений $a_1, \ldots, a_n$ [2].

В различных постановках задачи игроки могут конкурировать между собой или объединяться в коалиции для того, чтобы получить некоторый выигрыш, то есть увеличить свою прибыль. С этой целью они эксплуатируют некие ресурсы. В последующих главах будем рассматривать задачу нескольких ресурсов $x_1, x_2, ..., x_n$. В таком случае динамика игры будет описываться системой дифференциальных уравнений —

1) в случае возобновляемых ресурсов:

$$\begin{aligned} \dot{x}_1(t) &= a_1 - \delta_1 x_1, \qquad & x_1(t_0) = x_1^0, \\ \dot{x}_2(t) &= a_2 - \delta_2 x_2, \qquad & x_2(t_0) = x_2^0, \\ &\ldots & \end{aligned}$$

$$\dot{x}_n(t) = a_n - \delta_n x_n, \qquad x_n(t_0) = x_n^0.$$

2) так называемых «независимых движений»:

$$\begin{aligned} \dot{x}_1(t) &= a_1, \qquad & x_1(t_0) &= x_1^0, \\ \dot{x}_2(t) &= a_2, \qquad & x_2(t_0) &= x_2^0, \\ &\dots \\ \dot{x}_n(t) &= a_n, \qquad & x_n(t_0) &= x_n^0. \end{aligned}$$

Мгновенный выигрыш i-го игрока в момент времени $t \in [t_0, , \infty)$ определяется следующим образом: $h_i(x(t), a_1(t), \dots, a_n(t))$, где $h_i(\cdot)$ — непрерывная функция, а $x(t)$ — решение задачи Коши для (1.1). Функцию $h_i(\cdot)$ также называют функцией полезности.

§ 1.1. Постановка задачи оптимального управления

Рассмотрим задачу оптимального управления [12] с одним игроком на бесконечном промежутке времени с постоянным дисконтированием $\rho > 0$:

$$J(a, x) = \max_a \int_{t_0}^{\infty} e^{-\rho t} h(x(t), a(t)) dt, \tag{1.2}$$

когда переменная состояния изменяется в соответствии с дифференциальным уравнением:

$$\dot{x}(t) = f(x(t), a(t)), \quad x(t_0) = x^0. \tag{1.3}$$

Функции $h(x(t), a(t))$ и $f(x(t), a(t))$ будем считать дифференцируемыми.

Управление $a^*(t)$, доставляющее максимум функционалу (1.2), будем называть оптимальным управлением.

Рассмотрим случай, когда оптимальное управление разыскивается в классе позиционных стратегий $a(x, t)$, то есть стратегия игрока зависит от текущего состояния игры x и текущего момента времени t. В таком случае оптимальное управление игрока находится с использованием уравнения Гамильтона–Якоби–Беллмана [2].

Определим функцию Беллмана для задачи (1.2)-(1.3)

$$V(t, x) = \max_a \left\{ \int_{t_0}^{\infty} e^{-\rho t} h(x(t), a(t)) dt \,\middle|\, x(t) = x = x_t^* \right\},$$

где x_t^* — состояние системы в момент t на оптимальной траектории.

Теорема 1. *Управление* $a(t) = \varphi^*(t, x)$ *является оптимальным решением задачи управления на бесконечном промежутке времени, если существует непрерывно-дифференцируемая функция* $V(t, x)$, *определенная на* $[0, \infty) \times \mathbb{R}^m \to \mathbb{R}$ *и удовлетворяющая следующему уравнению*

$$\rho V(t,x) = \max_a \left\{ h(x,a) + \frac{\partial V(t,x)}{\partial x} f(x,a) \right\} =$$

$$= h(x, \varphi^*(t,x)) + \frac{\partial V(t,x)}{\partial x} f(x, \varphi^*(t,x)). \tag{1.4}$$

При этом значение функции Беллмана, вычисленной в начальный момент времени $V(t_0, x(t_0)) = J(a^*, x^0)$ равно выигрышу игрока в игре.

В случае кооперации игроков в дифференциальной игре для нахождения максимума суммарного выигрыша можно использовать аппарат теории управления (т.к. игроки действуют как один игрок).

Уравнение (1.4) называется уравнением Гамильтона–Якоби–Беллмана.

§ 1.2. Критерии отбора допустимых решений в случае их неединственности

Поскольку для линейно-квадратичных задач с многомерной фазовой переменной неединственность решений возникает как в кооперативной, так и в некооперативной постановке, то встает вопрос отбора из множества полученных решений. Для этого рассмотрим следующие критерии.

1.2.1. Экономический критерий Басса

В экономических приложениях существует достаточно общий подход [27] для отбора решений в том случае, если равновесие по Нэшу не является единственным. Он заключается в том, что полученные решения проверяются на выполнение следующего условия с ясной экономической интерпретацией: если предельная (маргинальная) прибыль, т.е. прибыль за продажу единицы товара в единицу времени, равна нулю, то и общая прибыль фирмы на всем промежутке игры должна равняться нулю. Это означает, что заменяя в полученных решениях определенный параметр модели нулем, отбраковываются те решения, для которых полученные значения функции Беллмана не равны нулю.

1.2.2. Математический метод для линейно-квадратичных задач оптимизации

Рассмотрим классический случай линейно-квадратичной задачи оптимизации (LQR) [28].

Будем рассматривать управляемую систему

$$\dot{x} = Ax + Ba, \quad x(t_0) = x^0, \tag{1.5}$$

где $x \in \mathbb{R}^n$ — вектор состояний, $a \in \mathbb{R}^m$ — вектор управлений, $A \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{n\times m}$, пара (A, B) управляема, $x^0 \in \mathbb{R}^n$ — начальное условие.

Задача оптимального управления заключается в построении управления вида линейной обратной связи по состоянию

$$\bar{a} = Kx, \quad K \in \mathbb{R}^{m\times n}, \tag{1.6}$$

которое доставляет наименьшее значение квадратичному функционалу качества $\forall x^0$:

$$J(a, x^0) = \int\limits_{t_0}^{\infty} (x^T Px + a^T Ra)dt, \tag{1.7}$$

где $P \in \mathbb{R}^{n\times n}$ — симметрическая матрица, $R \in \mathbb{R}^{m\times m}$ — положительно определенная матрица.

Классический метод решения основан на рассмотрении алгебраического уравнения Риккати.

Теорема 2. *Задача оптимального управления (1.6)-(1.7) разрешима тогда и только тогда, когда существует симметричная матрица Q, удовлетворяющая уравнению Риккати:*

$$A^T Q + QA - QBS^{-1}B^T Q + R = 0,$$

а управление $a = -R^{-1}B^T Qx$ является оптимальным.

Алгебраическое уравнение Риккати имеет два решения: положительно и отрицательно определенное. Такие решения существуют в предположении об управляемости системы и невырожденности матриц P, R. В качестве функции Беллмана выбирается положительно определенное решение Q^*, которое также является функцией Ляпунова V^* для замкнутой системы (1.5).

V^* — решение матричного уравнения Ляпунова при $K^* = -R^{-1}B^T V^*$:

$$(A + BK)^T V + V(A + BK) = -P - K^T RK.$$

Подставляя оптимальное управление $\bar{a}$ и функцию Ляпунова V^* в (1.7), получим минимальное значение функционала качества

$$J^*(\bar{a}, x^0) = {x^0}^T V^* x^0.$$

§ 1.3. Принцип оптимальности в кооперативной игре

Рассмотрим кооперативную дифференциальную игру [12] $\Gamma(t_0, x_0, \infty)$ n игроков с уравнением динамики:

$$\dot{x}(t) = f(t, x(t), a_1(t), \ldots, a_n(t)), \quad x(t_0) = x^0,$$

где $x(t) = [x_1(t), \ldots, x_n(t)]^T$. Функция выигрыша игрока i:

$$J(x_i, a_i) = \int\limits_{t_0}^{\infty} e^{-\rho t} h_i(t, x(t), a_1(t), \ldots, a_n(t))dt, \quad i = \overline{1, n}.$$

Будем считать, что игроки до начала игры договорились действовать в соответствии с принципом оптимальности, который состоит из:
1) выбора оптимальных управлений $a^* = (a_1^*, \ldots, a_n^*)$, которые максимизируют суммарный выигрыш;
2) выбора справедливого распределения суммарного выигрыша между игроками.

Пусть игроки получили общий выигрыш $V(N, x_0, t_0) = \sum_{i=1}^{n} J(x_i^*, a_i^*)$, где $x^*(t) = [x_1^*(t), \ldots, x_n^*(t)]^T$ — оптимальная кооперативная траектория. Рассмотрим вопрос раздела между игроками этой величины.

1.3.1. Дележ. Вектор Шепли

Рассмотрим кооперативную игру $\Gamma_V(t_0, x_0)$ в форме характеристической функции [29], где $V(S, x_0, t_0)$ — характеристическая функция, отражающая максимальный гарантированный выигрыш коалиции S, $s = |S|$.

Определение 1. *Вектор $\xi(x_0, t_0) = [\xi_1(x_0, t_0), \ldots, \xi_n(x_0, t_0)]$, удовлетворяющий условиям:*
1) индивидуальной рациональности

$$\xi_i(x_0, t_0) \geq V(\{i\}, x_0, t_0), \ i \in N,$$

2) групповой рациональности

$$\sum_{j \in N} \xi_j(x_0, t_0) = V(N, x_0, t_0),$$

называется дележом в игре $\Gamma_V(t_0, x_0)$.

Обозначим за $\mathrm{E}_V(t_0, x_0)$ множество дележей в игре $\Gamma_V(t_0, x_0)$.

Пусть в качестве механизма распределения суммарного выигрыша игроки договорились использовать вектор Шепли — дележ, определенный в начальный момент игры. Тогда компонента дележа игрока i имеет вид:

$$Sh_i(x_0, t_0) = \sum_{\substack{S\subset N\\ i\in S}} \frac{(n-s)!(s-1)!}{n!}[V(S, x_0, t_0) - V(S\backslash\{i\}, x_0, t_0)], \quad i = \overline{1, n}. \tag{1.8}$$

1.3.2. Принцип динамической устойчивости. Процедура распределения дележа

Рассмотрим изменение дележа при движении вдоль кооперативной траектории [12]. Пусть в кооперативной игре $\Gamma_V(t_0, x_0)$ в форме характеристической функции в качестве принципа оптимальности выбран вектор Шепли, который, при применении к игре из начальных состояний $x(t_0) = x_0$, $t = t_0$ определяет некоторое подмножество множества дележей
$\mathrm{W}_V(t_0, x_0) \subseteq \mathrm{E}_V(t_0, x_0)$, $\mathrm{W}_V(t_0, x_0) \neq \emptyset$, и оптимальную траекторию $\{x^*(t)\}$, $t \in [t_0, \infty)$, которая максимизирует

$$\sum_{i=1}^{n} \int_{t_0}^{\infty} e^{-\rho t} h_i(t, x^*(t), a^*(t)) dt. \tag{1.9}$$

Определение 2. *Пусть существует вектор-функция*

$$\beta(t) = \{\beta_i(t) \geq 0\}_{i=\overline{1,n}},$$

такая что компоненты вектора Шепли $Sh(t_0, x_0) = \{Sh_i(t_0, x_0)\}_{i=\overline{1,n}}$ в игре $\Gamma_V(t_0, x_0)$ представимы в виде

$$Sh_i(t_0, x_0) = \int_{t_0}^{\infty} e^{-\rho t} \beta_i(t) dt, \ i = \overline{1, n},$$

Вектор-функцию $\beta(t) = \{\beta_i(t)\}$ будем называть процедурой распределения дележа (ПРД) в игре $\Gamma_V(t_0, x_0)$.

Определение 3. *Будем называть вектор Шепли $\{\bar{S}h(t_0, x_0)\}$ динамически устойчивым вектором Шепли, если существует такая ПРД*

$$\{\beta_i(t) \geq 0\}, \ t \in [t_0, \infty),$$

что вектор $\bar{S}h(x^*(\tau),\tau) = \{\bar{S}h_i^\tau\}$, $\forall \tau \in [t_0, \infty)$, *вычисленный по формуле*

$$\bar{S}h(x^*(\tau),\tau) = e^{\rho\tau} \int\limits_{\tau}^{\infty} e^{-\rho t}\beta_i(t)dt, \ i = \overline{1,n},$$

также является вектором Шепли в соответствующей подыгре $\Gamma_V(x^*(\tau),\tau)$, $\tau \in [t_0, \infty)$.

Согласно определению (3), при распределении дележа $\{\bar{S}h(t_0, x_0)\}$ во времени при помощи выплат согласно ПРД $\{\beta_i(\tau)\}$ в каждый текущий момент времени τ, $\tau \in [t_0, \infty)$ ожидаемый дележ $\bar{S}h(x^*(\tau),\tau)$ в оставшейся подыгре $\Gamma_V(x^*(\tau),\tau)$ также является вектором Шепли.
Таким образом, имеет место **динамическая устойчивость**, то есть игрокам будет невыгодно нарушать соглашение о кооперации, которое они заключили перед началом игры.

Теорема 3. *[19] Пусть для каждой подыгры* $\Gamma_V(x^*(\tau),\tau)$, $\tau \in [t_0, \infty)$ *вектор-функция* $\bar{S}h(x^*(\tau),\tau)$ *является абсолютно непрерывной функцией времени* τ, $\tau \in [t_0, \infty)$. *Пусть*

$$\beta_i(\tau) = \rho\bar{S}h_i^\tau - (\bar{S}h_i^\tau)', \ i = \overline{1,n}, \tag{1.10}$$

где ρ *— дисконтирующий множитель в (1.9). Тогда в игре* $\Gamma_V(t_0, x_0)$ *вектор Шепли* $Sh(t_0, x_0)$ *является динамически устойчивым дележом с процедурой распределения дележа (ПРД) (1.10).*

Таким образом, в игре $\Gamma_V(t_0, x_0)$ ПРД определяет правило, по которому компоненты ожидаемого дележа распределяются во времени $[t_0, \infty)$. Игроки, использующие ПРД для распределения вектора Шепли во времени, не имеют оснований для отклонения от кооперативного соглашения.

§ 1.4. Дифференциальная игра на сети

Будем рассматривать дифференциальную игру на бесконечном временном промежутке n игроков (узлов), $n = |N|$, $N = \{1, \ldots, n\}$, объединенных в сеть [29]. Пусть $L = \{(i,j),\ i \in N,\ j \in N\}$ — множество всех ребер сети N. Считаем, что сеть не ориентирована, то есть если $(i,j) \in L$, то и $(j,i) \in L$.

Определение 4. *Подмножество* $S \subset N$ *называется коалицией, если для* $i \in S$, $j \in S$ *существует путь из* i *в* j *в сети* N.

Пусть $K(i,t) = \{(i,j),\ j \in N,\ (i,j) \in L\}$ —- множество всех игроков, с которыми у игрока i в данной сети имеется связь в момент времени $t \in$

$[t_0, \infty)$. За a_i обозначим управление игрока i, тогда $U^S = \{a_i, i \in S\}$ — набор управляющих переменных игроков из коалиции S.

Уравнение движения для игрока i, не входящего ни в одну в коалицию, определяется следующим образом:

$$\dot{x}_i = f_i(x_i, a_i), \quad x_i(t_0) = x_i^0 \geqslant 0, \ i \in N. \tag{1.11}$$

Уравнение движения для игрока i, входящего в коалицию S:

$$\dot{x}_i^S = f_i^S(x_i, a_i), \quad x_i(t_0) = x_i^0 \geqslant 0, \ i \in S. \tag{1.12}$$

Будем полагать, что в процессе игры игроки в одностороннем порядке могут ликвидировать ребра, соединяющие их с другими игроками, но восстанавливать их или создавать новые ребра не могут.

Пусть $x = (x_1, \ldots, x_n)$ — вектор состояния игроков. На множестве допустимых траекторий системы (1.11) для каждого игрока определим вектор-функцию

$$\varphi_i(x(t), t) = \{\varphi_i^1(x(t), t), \ \ldots, \ \varphi_i^n(x(t), t)\}, \ i = \overline{1, n},$$

компоненты которой могут принимать только два значения: 0 или 1. Также полагаем, что $\varphi_i^j(x^0, t_0) = 1$, если дуга $(i, j) \in L$.

Вектор-функция $\varphi_i(x(t), t)$ указывает на наличие дуг, связывающих игрока i с другими игроками. Для удобства положим $\varphi_i^i(x(t), t) \equiv 1$, то есть игрок i всегда связан с собой дугой. Поскольку игроки могут удалять ребра, то $\varphi_i(x(t), t)$ будет меняться во времени, $t \in [t_0, \infty)$.

Определение 5. *Стратегия игрока $i \in N$ есть пара $a_i(\cdot) = (a_i(t), \varphi_i(x, t))$.*

Обозначим за $\bar{U}_i$ множество всех стратегий игрока i.

Наличие зависимости от позиции x второй компоненты $\varphi_i(x, t))$ стратегии $a_i(\cdot)$ игрока i дает ему возможность прекратить связь с теми игроками, с которыми связь у него была, поэтому в процессе игры может меняться множество $K(i, t)$ игрока i. $L(a(\cdot))$ представляет собой динамику изменения множества дуг в зависимости от ситуации в игре, реализованной в момент времени $t \in [t_0, \infty)$:

$$a(\cdot) = (a_1(\cdot), \ldots, a_n(\cdot)) = (a_1(t), \varphi_1(x, t); \ldots; a_n(t), \varphi_n(x, t)).$$

Определим **функции выигрыша** игроков $i = \overline{1, n}$. Положим, что функция полезности игрока i: $h_i(x_i, x_i) = 0, \ h_i > 0$.

Функция выигрыша игрока i, который не входит ни в какую коалицию, зависит от переменной его состояния и переменных состояния игроков из множества $K(i,t)$:

$$H_i(x_i^0, x^0_{K(i,t_0)}; a_i, a_{K(i,t_0)}) = \int\limits_0^{\infty} e^{-\rho t} \sum_{j \in K(i,t)} h_i(x_i, x_j) d\tau,$$

где $x^0_{K(i,t_0)} = \{x_j^0;\ j \in K(i,t_0)\}$, $x_i(\tau),\ x_j(\tau)$ — решения (1.11), $i \neq j$.
Выигрыш коалиции $S \subset N$:

$$H(S) = \int\limits_0^{\infty} e^{-\rho t} \sum_{i \in S} \sum_{j \in K(i,t) \cap S} h_i(x_i, x_j) d\tau, \tag{1.13}$$

где $x_i(\tau),\ x_j(\tau)$ — решения (1.12), $i, j \in S$.

В данной задаче будем искать **равновесие по Нэшу**.

Теорема 4. *[29] В рассматриваемой игре найдется такая ситуация равновесия по Нэшу, что набор выигрышей игроков вида*

$$H_i(x_i^0, x^0_{K(i)}; a_i(\cdot), a_{K(i,t_0)}(\cdot)) = H_i(x_i^0, x^0_{K(i,t_0)}; (a_i(t), 1), (a_{K(i)}(t), 1)),\ i = \overline{1, n},$$

реализуем в этой ситуации. Здесь $(a_{K(i)}(t), 1) = \{(a_j(t), 1), j \in K(i,t)\}$.

Теорема означает, что любой набор выигрышей в построенной игре реализуем в некоторой ситуации равновесия по Нэшу.

Глава 2.

Теоретико-игровая модель управления инвестициями в рекламу

§ 2.1. Постановка задачи

Рассмотрим задачу нахождения оптимального управления инвестициями в рекламу. Отметим, что в ходе подготовки диссертации основные результаты данной главы были опубликованы в работе [11]. Кроме того, постановка задачи управления капиталовложениями в рекламную кампанию была изучена в бакалаврской работе автора [34] для двух игроков [32], а при отборе допустимого решения ранее рассматривался только численный пример для конкретного набора коэффициентов. В данной работе проблема решается для n игроков в общем виде.

Предположим, что на рынке находится n фирм, которые конкурируют за доли рынка, пропорциональные их гудвиллу.

Будем считать, что динамика накопления гудвилла G_i фирмы i описывается ОДУ

$$\dot{G}_i = ka_i - \delta G_i, \quad G_i(0) = G_0^i > 0, \qquad i = 1, \ldots, n, \tag{2.1}$$

где $k \geq 0$ — характеристика рынка, отвечающая за влияние рекламы на накопление гудвилла (чем больше k, тем более восприимчив рынок к воздействию рекламы); a_i — управление игрока i (объем инвестиций в рекламу в единицу времени), $a_i \geq 0$; δ — коэффициент амортизации, $\delta \geq 0$; G_0^i — значение гудвилла игрока i в начальный момент времени.

Положим $k = 1$.

Пусть $G(t) = \begin{bmatrix} G_1(t) & \ldots & G_n(t) \end{bmatrix}^T$. Будем рассматривать случай идентичных фирм.

Функция выигрыша игрока i имеет следующий вид:

$$J_i(a_i) = \int\limits_0^{\infty} e^{-\rho t} \left(\pi \left[\beta - \sum_{h=1}^{n} G_h \right] G_i - \frac{c}{2} a_i^2 \right) dt, \; i = 1, \ldots n, \tag{2.2}$$

где π — предельная прибыль фирмы (прибыль за продажу единицы товара в единицу времени), $\pi > 0$; $\left[\beta - \sum_{h=1}^{n} G_h(t)\right] G_i(t)$ — объем продаж фирмы i в момент времени t, $\beta > 0$, $\frac{c}{2}a_i^2$ — стоимость рекламных усилий (затраты), $c > 0$. Выигрыш игрока дисконтируется при помощи функции $e^{-\rho t}$, а игра рассматривается на бесконечном временном промежутке.

§ 2.2. Кооперативный случай

Рассмотрим кооперативный вариант игры [12]. Пусть до начала игры фирмы договорились об использовании оптимальных управлений $a^* = \{a_i^*\}$, максимизирующих их суммарный выигрыш. Тогда общая сумма выигрышей запишется в следующем виде:

$$J(G, a) = \sum_{i=1}^{n} J_i(G, a_i) = \sum_{i=1}^{n} \int_0^{\infty} e^{-\rho t} \left(\pi \left[\beta - \sum_{h=1}^{n} G_h \right] G_i - \frac{c}{2} a_i^2 \right) dt. \tag{2.3}$$

Будем искать решение данной линейно-квадратичной задачи в классе позиционных стратегий. Воспользуемся уравнением типа Гамильтона–Якоби–Беллмана:

$$\rho V(G) = \max_{\{a_1, \ldots, a_n\}} \left\{ \pi \left[\beta - \sum_{h=1}^{n} G_h \right] \sum_{i=1}^{n} G_i - \frac{c}{2} \sum_{i=1}^{n} a_i^2 + \sum_{i=1}^{n} \frac{\partial V}{\partial G_i} \left(a_i - \delta G_i \right) \right\}, \tag{2.4}$$

где $V(G)$ — непрерывно-дифференцируемая функция Беллмана, соответствующая искомому максимуму в (2.3).

Из условия максимизации правой части (2.4) имеем $a_i^* = \frac{1}{c}\frac{\partial V(G)}{\partial G_i}$, $a_i^* > 0$, $\forall i \in N$. Выберем функцию Беллмана в виде линейно-квадратичной функции:

$$V(G) = \alpha + \frac{\epsilon}{2} \sum_{i} G_i^2 + \frac{\eta}{2} \sum_{i \neq j} G_i G_j + \gamma \sum_{i} G_i. \tag{2.5}$$

Подставляя $V(G)$ и оптимальные управления $a_i^* = \frac{1}{c}\frac{\partial V(G)}{\partial G_i}$, $i = 1, \ldots, n$ в уравнение (2.4), получаем алгебраическое выражение, которое должно быть равно нулю при всех значениях G_i. Собирая члены при подобных слагаемых, приходим к следующей системе алгебраических уравнений, которые решаются относительно коэффициентов α, ϵ, η и γ:

$$
\begin{aligned}
0 &= 2\pi c + 2c\delta\epsilon + \epsilon\rho - \epsilon^2 - 2\eta^2, \\
0 &= \epsilon\gamma + 2\eta\gamma + \pi\beta c - c\delta\gamma - c\gamma\rho, \\
0 &= 2\pi c - 2\epsilon\eta - \eta^2 + 2c\delta\eta + c\eta\rho, \\
0 &= 2\alpha\rho c - 3\gamma^2.
\end{aligned} \tag{2.6}
$$

Вычитая третье уравнение из первого в (2.6), получаем

$$
(\epsilon - \eta)\left(\eta - \epsilon + 2c\delta + c\rho\right) = 0,
$$

откуда следует, что коэффициенты ϵ и η удовлетворяют одному из двух условий:

$$
\epsilon = \eta, \tag{2.7a}
$$

$$
\epsilon = \eta + 2c\delta + c\rho. \tag{2.7b}
$$

Коэффициент γ выражается из второго уравнения в (2.6) и имеет вид $\gamma = -\frac{\pi\beta c}{\epsilon+2\eta-c\delta-c\rho}$. С учетом указанных условий (2.7a), (2.7b) получаем следующие решения системы уравнений (2.6):

$$
\eta = \frac{1}{6}\begin{bmatrix} 2c\delta + c\rho - D \\ 2c\delta + c\rho + D \\ -2c\delta - c\rho - D \\ -c\rho - 2c\delta + D \end{bmatrix}, \quad \epsilon = \frac{1}{6}\begin{bmatrix} 2c\delta + c\rho - D \\ 2c\delta + c\rho + D \\ 10c\delta + 5c\rho - D \\ 10c\delta + 5c\rho + D \end{bmatrix}, \quad \gamma = \begin{bmatrix} \frac{2\pi\beta c}{c\rho + D} \\ \frac{2\pi\beta c}{c\rho - D} \\ \frac{2\pi\beta c}{c\rho + D} \\ \frac{2\pi\beta c}{c\rho - D} \end{bmatrix},
$$

$$
\alpha = \begin{bmatrix} \frac{6\pi^2\beta^2 c}{\rho\left(c\rho + D\right)^2} \\ \frac{6\pi^2\beta^2 c}{\rho\left(c\rho - D\right)^2} \\ \frac{6\pi^2\beta^2 c}{\rho\left(c\rho + D\right)^2} \\ \frac{6\pi^2\beta^2 c}{\rho\left(c\rho - D\right)^2} \end{bmatrix} \tag{2.8}
$$

где $D = \sqrt{(2c\delta + c\rho)^2 + 24c\pi}$. Легко заметить, что первые две компоненты решения удовлетворяют условию (2.7a), а последние две – условию (2.7b).

§ 2.3. Применение критериев для отбора допустимых решений

Было получено четыре различных решения данной задачи. Необходимо отобрать из них допустимые, для этого воспользуемся критериями (1.2.1), (1.2.2).

2.3.1. Математический метод для линейно-квадратичных задач оптимизации

Заметим, что функция – кандидат $V(G)$ (2.5) представляет собой сумму квадратичной формы, линейного слагаемого и постоянной величины. Исследуем свойства матрицы M квадратичной формы:

$$M = \frac{1}{2}\begin{bmatrix} \epsilon & \eta & \dots & \eta \\ \eta & \epsilon & \dots & \eta \\ \vdots & \vdots & \ddots & \eta \\ \eta & \dots & \eta & \epsilon \end{bmatrix}.$$

Матрица M имеет $n-1$ одинаковых собственных значений вида $\lambda_i = \epsilon - \eta$, $i = 1, \dots n-1$, с соответствующими собственными векторами вида $v_i = \begin{bmatrix} -1 & 0 & \dots & 1_{i+1} & \dots & 0 \end{bmatrix}^T$, где 1_k соответствует единице в k-й позиции. Последнее собственное значение соответствует собственному вектору $v_n = \begin{bmatrix} 1 & \dots & 1 \end{bmatrix}^T$ и равно $\lambda_n = \epsilon + (n-1)\eta$.

Рассмотрим структуру матрицы M для различных компонент решения (2.8). Заметим, что для первых двух компонент (2.8) матрица M вырождена с $n-1$ нулевыми собственными значениями и одним ненулевым, которое равно

$$\lambda_n = n\epsilon = n\eta = \frac{n}{6}\left(2c\delta + c\rho \mp \sqrt{(2c\delta + c\rho)^2 + 24c\pi}\right),$$

где знак минус соответствует первой компоненте в (2.8), а плюс – второй. Отметим, что выполняется неравенство $\sqrt{(2c\delta + c\rho)^2 + 24c\pi} > 2c\delta + c\rho$, поэтому в первом случае мы имеем $\lambda_n < 0$. Для третьей и четвертой компонент собственные значения имеют следующий вид (мы используем обозначение μ_i для собственных чисел, соответствующих последним двум компонентам решения):

$$\mu_i = 2c\delta + c\rho > 0, \qquad i = 1, \dots n-1$$
$$\mu_n = n\eta + 2c\delta + c\rho = \mp\frac{n}{6}\sqrt{(2c\delta + c\rho)^2 + 24c\pi} + \frac{6-n}{6}(2c\delta + c\rho),$$

где минус соответствует третьей компоненте, а плюс – четвертой. Легко проверить, что в первом случае (т.е. для третьей компоненты решения) $\mu_n < 0$ для любого числа игроков n.

Таким образом, мы имеем четыре функции–кандидата, соответствующие четырем компонентам решения (2.8). Эта ситуация схожа с классическим случаем линейно-квадратичной задачи оптимизации (LQR) [28].

Поскольку, в отличие от задачи LQR, в рассматриваемой постановке решается задача максимизации, в качестве функции Беллмана необходимо выбрать отрицательно полу-определенную функцию, соответствующую первой компоненте решения (2.8). Обозначим ее $\bar{V}(G)$.

2.3.2. Экономический критерий Басса

В экономических приложениях для отбора решения из полученного (конечного) множества решений часто используется подход, предложенный в [27].

Идея данного подхода заключается в том, что при нулевой маргинальной прибыли (т.е. прибыли за продажу единицы товара) общий доход должен быть равен нулю.

Используя обозначения нашей модели, сформулируем данный критерий.

Критерий (Bass et.al., 2005) 1 ([27]). *Пусть в задаче (2.1, 2.2) маргинальная полезность $\pi = 0$. Тогда с необходимостью имеем общий доход $V(G) = 0$, где $V(G)$ — функция Беллмана, соответствующая искомому максимуму в (2.3).*

Будем использовать для полученных решений простой "тест" из Критерия 1.

Зафиксируем $\pi = 0$. Получаем

$$\eta = \frac{1}{6}\begin{bmatrix} 0 \\ 4c\delta + 2c\rho \\ -4c\delta - 2c\rho \\ 0 \end{bmatrix} \quad \epsilon = \frac{1}{6}\begin{bmatrix} 0 \\ 4c\delta + 2c\rho + D \\ 8c\delta + 4c\rho \\ 12c\delta + 6c\rho \end{bmatrix} \quad \gamma = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}. \tag{2.9}$$

Таким образом, получаем, что только функция–кандидат $V(G)$ (2.5), коэффициенты которой соответствуют первой компоненте решения (2.8), удовлетворяет критерию, предложенному в [27].

Отметим, что полученный с экономической точки зрения результат соответствует результатам проведенного выше математического анализа. Результатом обоих подходов является выбор функции Беллмана $\bar{V}(G)$ (2.5), где коэффициенты имеют следующий вид:

$$
\begin{aligned}
\eta &= \frac{1}{6}(2c\delta + c\rho - \sqrt{(2c\delta + c\rho)^2 + 24c\pi}),\\
\epsilon &= \frac{1}{6}(2c\delta + c\rho - \sqrt{(2c\delta + c\rho)^2 + 24c\pi}),\\
\gamma &= \frac{2\pi\beta c}{c\rho + \sqrt{(2c\delta + c\rho)^2 + 24c\pi}},\\
\alpha &= \frac{6\pi^2\beta^2 c}{\rho\left(c\rho + \sqrt{(2c\delta + c\rho)^2 + 24c\pi}\right)^2}.
\end{aligned} \tag{2.10}
$$

Соответствующие оптимальные управления имеют вид $a_i^* = \frac{1}{c}(\eta \sum\limits_{i=1}^{n} G_i + \gamma)$, $i = 1, \ldots, n$. Также можно вычислить стационарные значения G_i^*, $i = 1, \ldots, n$, переменных состояния, соответствующих оптимальному решению:

$$
G_i^* = \frac{6\pi\beta c}{12n\pi c + 2nc^2\delta^2 - 3\delta\sqrt{(2c\delta + c\rho)^2 + 24c\pi} + nc\delta\sqrt{(2c\delta + c\rho)^2 + 24c\pi} + 3c\delta\rho + nc^2\delta\rho}.
$$

§ 2.4. Распределение кооперативного выигрыша

В результате кооперации все игроки заработали максимальную сумму $V(N, x_0)$ — значение функции Беллмана в точке x_0.

Пусть для раздела полученной суммы игроки договорились использовать вектор Шепли (1.8). Поскольку игроки симметричные, игрок i получит выигрыш:

$$
Sh_i = \frac{V(N, G_0)}{n} = \frac{6\pi^2\beta^2 c}{n\rho\left(c\rho + \sqrt{(2c\delta + c\rho)^2 + 24c\pi}\right)^2} +
$$

$$
+ \frac{(\sum_i \sum_j G_i^0 G_j^0)(2c\delta + c\rho - \sqrt{(2c\delta + c\rho)^2 + 24c\pi})}{12n} + \frac{2\pi\beta c \sum_i G_i^0}{n(c\rho + \sqrt{(2c\delta + c\rho)^2 + 24c\pi})}.
$$

Процедура распределения дележа для задачи с бесконечным горизонтом и дисконтированием с коэффициентом ρ (1.10)

$$
\beta_i(t)n = \rho V(N, G(t)) - (V(N, G(t))' = \rho \int_t^\infty \sum_{i=1}^{n} h_i(G^*, a^*) e^{-\rho\tau} d\tau -
$$

$$-\left(\int_t^\infty \sum_{i=1}^n h_i(G^*,a^*)e^{-\rho\tau}d\tau\right)' = 2\sum_{i=1}^n h_i(G^*,a^*)e^{-\rho t}. \qquad (2.11)$$

Учитывая, что $h_i(G^*,a^*) = \pi\left[\beta - \sum\limits_{h=1}^n G_h^*\right]G_i^* - \frac{c}{2}a_i^{*2}$

$$\beta_i(t)n = e^{-\rho t}\sum_{i=1}^n\left\{2\pi\left[\beta - \sum_{h=1}^n G_h^*\right]G_i^* - \frac{1}{c}\left(\eta\sum_{i=1}^n G_i + \gamma\right)^2\right\} =$$

$$= e^{-\rho t}\left(2\pi n G_i^*(\beta - n) - \frac{n^2}{c}(\eta n G_i^* + \gamma)^2\right).$$

Тогда каждый игрок в каждый момент времени t получит:

$$\beta_i(t) = e^{-\rho t}\left(2\pi G_i^*(\beta - n) - \frac{n}{c}(\eta n G_i^* + \gamma)^2\right) = e^{-\rho t}\Bigg(2\pi G_i^*(\beta - n) -$$

$$-n\left\{\frac{nG_i^*}{6}\left[\sqrt{c}(2\delta+\rho) - \sqrt{c(2\delta+\rho)^2 + 24\pi}\right] + \frac{2\pi\beta\sqrt{c}}{c\rho + \sqrt{(2c\delta + c\rho)^2 + 24c\pi}}\right\}^2\Bigg).$$

При распределении компонент дележа согласно полученной вектор-функции (ПРД) $\{\beta_i(t)\}$ у игроков не возникнет оснований для отклонения от кооперативного решения в течение реализации всего долгосрочного проекта по вложению средств в рекламную кампанию. Следовательно, и сама кооперативная траектория $G^*(t)$ будет устойчивой.

Графическая интерпретация полученных результатов приводится на рис. (2.1), (2.2).

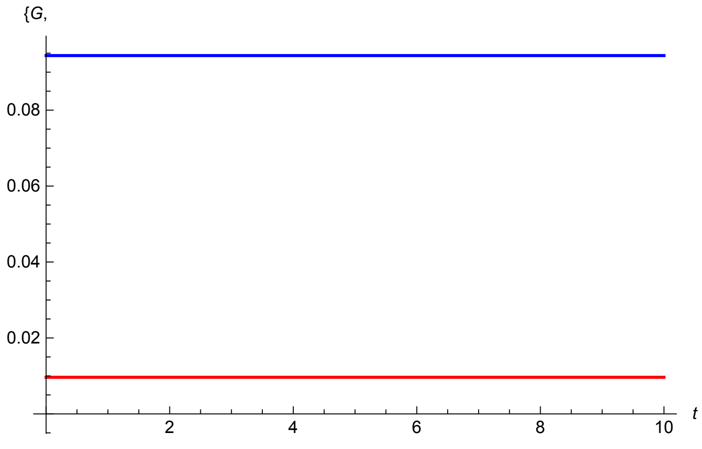


Рис. 2.1: Стационарные оптимальные траектория $G_i^*(t)$ (выше) и управление $a_i^*(t)$ при $n = 3, c = 14, \rho = 10, \beta = 1, \delta = 0,1$

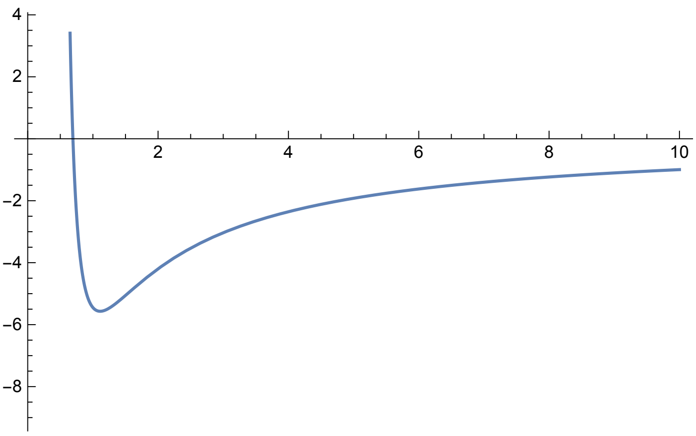


Рис. 2.2: $\beta_i(t)$ при $n = 3, c = 14, \rho = 10, \beta = 1, \delta = 0,1$

Глава 3.

Теоретико-игровая модель управления объемами вредных выбросов

§ 3.1. Постановка задачи

Сформулируем модель управления объемами вредных выбросов для случая, когда фазовая переменная, описывающая накопленный в окружающей среде объем загрязнений, не является одномерной.

Отметим, что данная модель в некооперативной постановке была описана автором в работах [34] и [26]. В данной работе уточняется равновесие по Нэшу в общем виде, а также решается задача в кооперативной форме.

Пусть динамика объема выбросов x_i в окружающую среду фирмы i описывается ОДУ

$$\dot{x}_i = a_i, \quad i = \overline{1,n}, \quad x_i(0) = x_0^i \geqslant 0, \tag{3.1}$$

где a_i — поток выбросов игрока i (количество выбросов за единицу времени), $a_i \geqslant 0$, x_i — общее количество выбросов, x_0^i — начальное количество выбросов игрока i. Пусть $y_i = m_i a_i$ — поток продукта, произведенного игроком i. Функция выигрыша игрока i имеет следующий вид:

$$J_i(a,x) = \int\limits_0^\infty e^{-\rho t}\left(\pi\left[\left(k - \sum_{h=1}^n y_h\right) + \beta\right] y_i - d_i x_i\right) dt, \tag{3.2}$$

где π — коэффициент пропорциональности, k — требуемый (максимальный) поток продукта, $\pi\left[\left(k - \sum\limits_{h=1}^n y_h\right) + \beta\right] y_i(t)$ — цена продукта, β — базовая цена продукта в условиях полного насыщения рынка, а $d_i x_i$ — затраты на борьбу с загрязнением.

Для простоты положим $m_i = 1$, $d_i = 1$, $\beta = 0$. Тогда интегральный функционал (3.2) примет вид:

$$J_i(a,x) = \int\limits_0^\infty e^{-\rho t}\left(\pi\left(k-\sum_{h=1}^n a_h\right)a_i - x_i\right)dt. \tag{3.3}$$

§ 3.2. Равновесие по Нэшу

Рассмотрим некооперативную игру. Пусть игроки не смогли договориться до начала игры. В варианте конкуренции будем искать равновесие по Нэшу.

Определение 6. *Набор управлений $a^{NE} = \{a_1^{NE},\ldots,a_n^{NE}\}$ называется равновесием по Нэшу если*

$$J_i(a^{NE}) \geqslant J_i(a^{NE}||a_i),$$

где $a^{NE}||a_i = \{a_1^{NE},\ldots,a_{i-1}^{NE},a_i,a_{i+1}^{NE},\ldots,a_n^{NE}\}$, $a_i\in U$, $i\in N$.

Рассматриваем случай двух игроков, то есть считаем, что $n=2$.

Пусть $V_i(x)$ — непрерывно-дифференцируемая функция Беллмана, которая является решением уравнения Гамильтона – Якоби – Беллмана:

$$\rho V_i(x) = \max_{a_i\geqslant 0, a_j^{NE}}\left\{\pi\left(k-\sum_{h=1}^2 a_h\right)a_i - x_i + \sum_{i=1}^2\frac{\partial V_i}{\partial x_i}a_i\right\},\ i,j=\overline{1,2}, j\neq i \tag{3.4}$$

Предполагается, что второй игрок использует управление из равновесия по Нэшу a^{NE}.

Выберем следующий вид функции Беллмана

$$V_i(x) = x'\begin{bmatrix} q_{i1} & \frac{q_{i2}}{2} \\ \frac{q_{i2}}{2} & q_{i3}\end{bmatrix}x + x'[w_{i1},\, w_{i2}]' + z_i, \quad i=\overline{1,2}, \tag{3.5}$$

или

$$V_1(x) = q_{11}x_1^2 + q_{12}x_1x_2 + q_{13}x_2^2 + x_1w_{11} + x_2w_{12} + z_1,$$

$$V_2(x) = q_{21}x_1^2 + q_{22}x_1x_2 + q_{23}x_2^2 + x_1w_{21} + x_2w_{22} + z_2,$$

и эти функции имеют частные производные

$$\frac{\partial V_1(x)}{\partial x_1} = 2q_{11}x_1 + q_{12}x_2 + w_{11}; \quad \frac{\partial V_2(x)}{\partial x_2} = q_{22}x_1 + 2q_{23}x_2 + w_{22}.$$

Подставляя функцию $V_i(x)$,$i=\overline{1,2}$ и ее частные производные в (3.4), получаем

$$\rho\Bigg(q_{11}x_1^2+q_{12}x_1x_2+q_{13}x_2^2+x_1w_{11}+x_2w_{12}+z_1\Bigg)=\max_{a_1\geq 0,a_2^{NE}}\Bigg\{\pi(k-a_1-a_2)a_1-$$

$$-x_1+a_1(2q_{11}x_1+q_{12}x_2+w_{11})+a_2(q_{22}x_1+2q_{23}x_2+w_{22})\Bigg\};$$

$$\rho\Bigg(q_{21}x_1^2+q_{22}x_1x_2+q_{23}x_2^2+x_1w_{21}+x_2w_{22}+z_2\Bigg)=\max_{a_2\geq 0,a_1^{NE}}\Bigg\{\pi(k-a_1-a_2)a_2-$$

$$-x_2+a_1(2q_{11}x_1+q_{12}x_2+w_{11})+a_2(q_{22}x_1+2q_{23}x_2+w_{22})\Bigg\}. \tag{3.6}$$

Рассмотрим функции

$$F(a_1,a_2^{NE})=\pi[k-a_1-a_2]a_1-x_1+a_1(2q_{11}x_1+q_{12}x_2+w_{11})+a_2(q_{22}x_1+2q_{23}x_2+w_{22});$$

$$F(a_2,a_1^{NE})=\pi[k-a_1-a_2]a_2-x_2+a_1(2q_{11}x_1+q_{12}x_2+w_{11})+a_2(q_{22}x_1+2q_{23}x_2+w_{22}),$$

и найдем ее частные производные

$$\frac{\partial F(a_1,a_2^{NE})}{\partial a_1}=\pi k-2\pi a_1-\pi a_2+2q_{11}x_1+q_{12}x_2+w_{11},$$

$$\frac{\partial F(a_2,a_1^{NE})}{\partial a_2}=\pi k-\pi a_1-2\pi a_2+q_{22}x_1+2q_{23}x_2+w_{22}.$$

С необходимостью приравняем их к нулю и получим оптимальные управления следующего вида:

$$a_1^*=\frac{k}{3}+\frac{(4q_{11}-q_{22})x_1+2(q_{12}-q_{23})x_2+2w_{11}-w_{22}}{3\pi}; \tag{3.7}$$

$$a_2^*=\frac{k}{3}+\frac{2(q_{22}-q_{11})x_1+(4q_{23}-q_{12})x_2+2w_{22}-w_{11}}{3\pi}. \tag{3.8}$$

Подставим (3.7) и (3.8) в исходные уравнения (3.6), раскроем скобки, приведем подобные и методом неопределенных коэффициентов придем к следующим двум системам уравнений. Из первого уравнения (3.6):

$$x_1^2: 16{q_{11}}^2-q_{11}(14q_{22}+9\pi\rho)+7{q_{22}}^2=0;$$

$$x_1x_2: 4q_{11}(4q_{12}-7q_{23})+7q_{22}(4q_{23}-q_{12})=9\pi q_{12}\rho;$$

$$x_1 : \pi k(8q_{11} + q_{22}) - 9\pi = w_{11}(9\pi\rho - 16q_{11} + 7q_{22}) + 14w_{22}(q_{11} - q_{22}); \quad (3.9)$$

$$x_2^2 : 4{q_{12}}^2 - 14q_{12}q_{23} + 28{q_{23}}^2 = 9\pi q_{13}\rho;$$

$$x_2 : 9\pi\rho w_{12} - 2\pi k(2q_{12} + q_{23}) = 2w_{11}(4q_{12} - 7q_{23}) + 7w_{22}(4q_{23} - q_{12}); \quad (3.10)$$

$$x_1^0x_2^0 : \pi k(4w_{11} + w_{22}) - 7w_{11}w_{22} + 4{w_{11}}^2 + 7{w_{22}}^2 = 9\pi\rho z_1 - \pi^2k^2. \quad (3.11)$$

Из второго уравнения (3.6) получим

$$x_1^2 : 28{q_{11}}^2 - 14q_{11}q_{22} + 4{q_{22}}^2 = 9\pi q_{21}\rho;$$

$$x_1x_2 : 28q_{11}(q_{23} - q_{12}) = q_{22}(16q_{23} - 7q_{12} - 9\pi\rho);$$

$$x_1 : 9\pi\rho w_{21} - 2\pi k(q_{11} + 2q_{22}) = 7w_{11}(4q_{11} - q_{22}) + 2w_{22}(4q_{22} - 7q_{11}); \quad (3.12)$$

$$x_2^2 : 16{q_{23}}^2 - q_{23}(14q_{12} + 9\pi\rho) + 7{q_{12}}^2 = 0;$$

$$x_2 : \pi k(q_{12} + 8q_{23}) - 9\pi = 14w_{11}(q_{23} - q_{12}) + w_{22}(7q_{12} - 16q_{23} + 9\pi\rho); \quad (3.13)$$

$$x_1^0x_2^0 : \pi k(w_{11} + 4w_{22}) - 7w_{11}w_{22} + 7{w_{11}}^2 + 4{w_{22}}^2 = 9\pi\rho z_2 - \pi^2k^2. \quad (3.14)$$

Рассмотрим систему из шести уравнений, которые зависят только от $q_{11}, q_{12}, q_{13}, q_{21}, q_{22}, q_{23}$:

$$16{q_{11}}^2 - q_{11}(14q_{22} + 9\pi\rho) + 7{q_{22}}^2 = 0; \quad (3.15)$$

$$4q_{11}(4q_{12} - 7q_{23}) + 7q_{22}(4q_{23} - q_{12}) = 9\pi q_{12}\rho; \quad (3.16)$$

$$4{q_{12}}^2 - 14q_{12}q_{23} + 28{q_{23}}^2 = 9\pi q_{13}\rho; \quad (3.17)$$

$$28{q_{11}}^2 - 14q_{11}q_{22} + 4{q_{22}}^2 = 9\pi q_{21}\rho; \quad (3.18)$$

$$28q_{11}(q_{23} - q_{12}) = q_{22}(16q_{23} - 7q_{12} - 9\pi\rho); \quad (3.19)$$

$$16{q_{23}}^2 - q_{23}(14q_{12} + 9\pi\rho) + 7{q_{12}}^2 = 0. \quad (3.20)$$

Из (3.15) и (3.20) выразим

$$q_{11} = \frac{1}{32}(14q_{22} + 9\pi\rho \pm \sqrt{-252{q_{22}}^2 + 252q_{22}\pi\rho + 81\pi^2\rho^2}),$$

$$q_{23} = \frac{1}{32}(14q_{12} + 9\pi\rho \pm \sqrt{-252{q_{12}}^2 + 252q_{12}\pi\rho + 81\pi^2\rho^2}).$$

Введем следующие обозначения:

$$D_{22} = -252{q_{22}}^2 + 252q_{22}\pi\rho + 81\pi^2\rho^2, \quad D_{12} = -252{q_{12}}^2 + 252q_{12}\pi\rho + 81\pi^2\rho^2.$$

Сложим (3.16) и (3.19)

$$4q_{11}q_{12} - 4q_{22}q_{23} + 3\pi\rho(q_{12} - q_{22}) = 0,$$

и подставим в это равенство q_{11} и q_{23}. Тогда получим:

$$33\pi\rho(q_{12} - q_{22}) \pm 8q_{12}\sqrt{D_{22}} = \pm 8q_{22}\sqrt{D_{12}},$$

и возведем обе части этого равенства в квадрат:

$$(q_{12}-q_{22})(1089\pi^2\rho^2(q_{12}-q_{22})\pm 528\pi\rho q_{12}\sqrt{D_{22}}+16128\pi\rho q_{12}q_{22}+5184\pi^2\rho^2(q_{12}+q_{22})) =$$

Получаем, что либо $q_{12} = q_{22}$, либо

$$363\pi\rho(q_{12} - q_{22}) + 84q_{12}q_{22} + 27\pi\rho(q_{12} + q_{22}) = \mp 22q_{12}\sqrt{D_{22}}. \qquad (3.21)$$

Для простоты вычислений рассмотрим случай, когда $q_{12} = q_{22}$. Тогда из этого следует, что $q_{11} = q_{23}$, но тогда и $q_{13} = q_{21}$. Учитывая это, получаем, что равенства (3.15) и (3.20), (3.16) и (3.19), (3.17) и (3.18) идентичные, тогда система (3.15)-(3.20) из шести уравнений сведется к системе из трех уравнений:

$$\begin{aligned} 16{q_{11}}^2 - q_{11}(14q_{22} + 9\pi\rho) + 7{q_{22}}^2 = 0; \\ 44q_{11}q_{22} - 28q_{11}^2 - 7q_{22}^2 = 9\pi q_{22}\rho; \\ 4{q_{22}}^2 - 14q_{22}q_{11} + 28{q_{11}}^2 = 9\pi q_{13}\rho. \end{aligned} \qquad (3.22)$$

Из второго уравнения этой системы выразим

$$q_{11} = \frac{1}{14}(11q_{22} \pm 3\sqrt{8q_{22}^2 - 7q_{22}\pi\rho}), \qquad (3.23)$$

и, подставив (3.23) в первое уравнение системы (3.22), найдем q_{22}.

Тогда при подстановке q_{22} в (3.23) получим, что каждому значению q_{22} соответствуют два значения q_{11}.

Учитывая q_{22} и q_{11}, из последнего уравнения системы (3.22) находим q_{13}.

Учитывая равенства $q_{12} = q_{22}$, $q_{11} = q_{23}$ и уравнения (3.9), (3.13), получаем, что $w_{11} = w_{22} = \dfrac{k\pi(q_{22} + 8q_{11}) - 9\pi}{9\pi\rho - 7q_{22} - 2q_{11}}$, и, подставляя q_{22} и q_{11}, находим w_{11}.

Учитывая равенства $q_{12} = q_{22}$, $q_{11} = q_{23}$, $w_{11} = w_{22}$ и уравнения (3.10), (3.12), получаем, что $w_{12} = w_{21} = \dfrac{2k\pi(2q_{22} + q_{11}) + w_{11}(14q_{11} + q_{22})}{9\pi\rho}$, и, подставляя q_{22}, q_{11} и w_{11}, находим w_{12}.

Учитывая равенство $w_{11} = w_{22}$ и уравнения (3.11), (3.14), получаем, что $z_1 = z_2 = \dfrac{k^2\pi^2 + 5k\pi w_{11} + 4w_{11}^2}{9\pi\rho}$, и, подставляя w_{11}, находим z_1.

$$q_{11} = q_{23} = \begin{bmatrix} 0 \\ \frac{9\pi\rho}{16} \\ \frac{7\pi\rho}{16} \\ \frac{\pi\rho}{8} \\ \frac{7\pi\rho}{32} \\ \frac{63\pi\rho}{64} \end{bmatrix}, \quad q_{12} = q_{22} = \begin{bmatrix} 0 \\ 0 \\ \pi\rho \\ -\frac{\pi\rho}{4} \\ -\frac{\pi\rho}{4} \\ \frac{9\pi\rho}{8} \end{bmatrix}, \quad q_{13} = q_{21} = \begin{bmatrix} 0 \\ \frac{63\pi\rho}{64} \\ \frac{23\pi\rho}{64} \\ \frac{\pi\rho}{8} \\ \frac{67\pi\rho}{256} \\ \frac{1899\pi\rho}{1024} \end{bmatrix},$$

$$w_{11} = w_{22} = \begin{bmatrix} -\frac{1}{\rho} \\ \frac{4\pi k\rho - 8}{7\rho} \\ \frac{4\pi k\rho - 8}{\rho} \\ \frac{\pi k\rho - 12}{14\rho} \\ \frac{8\pi k\rho - 48}{55\rho} \\ \frac{32 - 32\pi k\rho}{3\rho} \end{bmatrix}, \quad w_{12} = w_{21} = \begin{bmatrix} 0 \\ \frac{5\pi k\rho - 8}{8\rho} \\ \frac{89\pi k\rho - 152}{24\rho} \\ -\frac{\pi k\rho + 2}{14\rho} \\ -\frac{3\pi k\rho + 48}{176\rho} \\ \frac{1696 - 1627\pi k\rho}{96\rho} \end{bmatrix}, \quad (3.24)$$

$$z_1 = z_2 = \begin{bmatrix} \frac{\pi^2k^2\rho^2 - 5\pi k\rho + 4}{9\rho^3\pi} \\ \frac{253\pi^2k^2\rho^2 - 536\pi k\rho + 256}{441\rho^3\pi} \\ \frac{85\pi^2k^2\rho^2 - 296\pi k\rho + 256}{9\rho^3\pi} \\ \frac{15\pi^2k^2\rho^2 - 52\pi k\rho + 32}{98\rho^3\pi} \\ \frac{609\pi^2k^2\rho^2 - 1808\pi k\rho + 1024}{3025\rho^3\pi} \\ \frac{2454193\pi^2k^2\rho^2 - 5315264\pi k\rho + 2876416}{20736\rho^3\pi} \end{bmatrix}.$$

Получили решения (3.24) систем уравнений (3.9)-(3.20).

3.2.1. Применение критериев для отбора допустимых решений

Рассмотрим критерии отбоа (1.2.1) и (1.2.2). Покажем, что полученные решения не удовлетворяют ни одному из критериев.

Экономический критерий Басса

Сформулируем данный критерий в рамках нашей модели.

Критерий (Bass et.al., 2005) 2 ([27]). *Пусть в задаче маргинальная полезность* $\pi = 0$. *Тогда с необходимостью имеем доход игрока* i $V_i(x) = 0$, $i = \overline{1, n}$, *где* $V_i(x)$ — *функция Беллмана, соответствующая искомому максимуму в (3.3).*

Пусть $\pi = 0$ в (3.24). Получим

$$q_{11} = q_{23} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \quad q_{12} = q_{22} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \quad q_{13} = q_{21} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \quad w_{11} = w_{22} = \begin{bmatrix} -\frac{1}{\rho} \\ -\frac{8}{7\rho} \\ -\frac{8}{\rho} \\ -\frac{6}{7\rho} \\ -\frac{48}{55\rho} \\ \frac{32}{3\rho} \end{bmatrix},$$

$$w_{12} = w_{21} = \begin{bmatrix} 0 \\ -\frac{1}{\rho} \\ -\frac{19}{3\rho} \\ -\frac{1}{7\rho} \\ -\frac{3}{11\rho} \\ \frac{53}{3\rho} \end{bmatrix}, \quad z_1 = z_2 = \begin{bmatrix} -\frac{5k}{9\rho^2} \\ -\frac{536k}{441\rho^2} \\ -\frac{296k}{9\rho^2} \\ -\frac{26k}{49\rho^2} \\ -\frac{1808k}{3025\rho^2} \\ -\frac{83051k}{324\rho^2} \end{bmatrix}. \qquad (3.25)$$

Таким образом, ни одно из полученных решений задачи не удовлетворяет данному критерию.

Классический метод для линейно-квадратичных задач оптимизации

Рассмотрим матрицу квадратичной формы функции Беллмана $V_i(x)$:

$$M = \begin{bmatrix} q_{i1} & \frac{q_{i2}}{2} \\ \frac{q_{i2}}{2} & q_{i3} \end{bmatrix},$$

которая имеет два собственных числа

$$\lambda_1 = \frac{1}{2}\left(q_{i1} + q_{i3} + \sqrt{q_{i1}^2 - 4q_{i2}^2 + q_{i3}^2 + 6q_{i1}q_{i3}}\right),$$

$$\lambda_2 = \frac{1}{2}\left(q_{i1} + q_{i3} - \sqrt{q_{i1}^2 - 4q_{i2}^2 + q_{i3}^2 + 6q_{i1}q_{i3}}\right).$$

В зависимости от знаков λ_1, λ_2 квадратичная форма в функции Беллмана $V_i(x)$ будет либо положительно определенной, либо отрицательно определенной. В нашем случае функция Беллмана должна быть отрицательно полу-определенной (2.3.1).

Поскольку λ_1 строго положительна, за исключением первой компоненты решения, где λ_1=λ_2=0, ни одна из компонент решения не удовлетворяет данному критерию.

Оба подхода дали одинаковый результат.

§ 3.3. Модификация модели

Предположим, что в модели (3.1, 3.3) игроки не тратятся на борьбу с загрязнением [33], то есть полагаем $d_i = 0$, $i = \overline{1, n}$. тогда функция выигрыша (3.3) игрока i перепишется в виде:

$$J_i(a, x) = \int\limits_0^\infty e^{-\rho t}\left(\pi\left(k - \sum_{h=1}^{n} a_h\right)a_i\right)dt. \tag{3.26}$$

Уравнение Гамильтона–Якоби–Беллмана примет вид:

$$\rho V_i(x) = \max_{a_i \geqslant 0, a_j^{NE}}\left\{\pi\left(k - \sum_{h=1}^{2} a_h\right)a_i + \sum_{i=1}^{2}\frac{\partial V_i}{\partial x_i}a_i\right\}, \quad i = \overline{1,2}, j \neq i. \tag{3.27}$$

Управления, доставляющие максимум правой части (3.27), также будут иметь вид (3.7, 3.8). Функцию Беллмана также выбираем в виде (3.5). Решая задачу (3.1, 3.26) аналогичным образом, придем к тому, что в системах (3.9)-(3.20) перепишутся только уравнения, получаемые относительно первой степени x:

$$x_1 : \pi k(8q_{11} + q_{22}) = w_{11}(9\pi\rho - 16q_{11} + 7q_{22}) + 14w_{22}(q_{11} - q_{22}); \quad (3.28)$$

$$x_2 : \pi k(q_{12} + 8q_{23}) = 14w_{11}(q_{23} - q_{12}) + w_{22}(7q_{12} - 16q_{23} + 9\pi\rho); \quad (3.29)$$

Решая обе системы, получим следующие значения коэффициентов функции Беллмана:

$$q_{11} = \begin{bmatrix} 0 \\ \frac{9\pi\rho}{16} \\ \frac{7\pi\rho}{16} \\ \frac{\pi\rho}{8} \\ \frac{7\pi\rho}{32} \\ \frac{63\pi\rho}{64} \end{bmatrix}, \quad q_{12} = \begin{bmatrix} 0 \\ 0 \\ \pi\rho \\ -\frac{\pi\rho}{4} \\ -\frac{\pi\rho}{4} \\ \frac{9\pi\rho}{8} \end{bmatrix}, \quad q_{13} = \begin{bmatrix} 0 \\ \frac{63\pi\rho}{64} \\ \frac{23\pi\rho}{64} \\ \frac{\pi\rho}{8} \\ \frac{67\pi\rho}{256} \\ \frac{1899\pi\rho}{1024} \end{bmatrix},$$

$$w_{11} = \begin{bmatrix} 0 \\ \frac{4\pi k}{7} \\ 4\pi k \\ \frac{\pi k}{14} \\ \frac{8\pi k}{55} \\ \frac{-32\pi k}{3} \end{bmatrix}, \quad w_{12} = \begin{bmatrix} 0 \\ \frac{5\pi k}{8} \\ \frac{89\pi k}{24} \\ -\frac{\pi k}{14} \\ -\frac{3\pi k}{176} \\ \frac{-1627\pi k}{96} \end{bmatrix}, \quad z_1 = \begin{bmatrix} \frac{\pi k^2}{9\rho} \\ \frac{253\pi k^2}{441\rho} \\ \frac{85\pi k^2}{9\rho} \\ \frac{15\pi k^2}{98\rho} \\ \frac{609\pi k^2}{3025\rho} \\ \frac{3625\pi k^2}{81\rho} \end{bmatrix}. \quad (3.30)$$

Видим, что $q_{i,j}$, $i = 1, 2$, $j = 1, 2, 3$ не изменятся.

Применим к полученным решениям критерии (1.2.1), (1.2.2).

Пусть в задаче коэффициент пропорциональности $\pi = 0$. Поскольку $k \geqslant 0$, а $\rho > 0$, то функции Беллмана с положительными коэффициентами соответствует только третий компонент решения:

$$q_{11} = \frac{7\pi\rho}{16};\ q_{12} = \pi\rho;\ q_{13} = \frac{\pi\rho}{8};\ w_{11} = 4\pi k;\ w_{12} = \frac{89\pi k}{24};\ z_1 = \frac{85\pi k^2}{9\rho},$$

который удовлетворяет экономическому критерию.

Рассмотрим матрицу квадратичной формы функции Беллмана $V_i(x)$ (3.5):

$$M = \begin{bmatrix} q_{i1} & \frac{q_{i2}}{2} \\ \frac{q_{i2}}{2} & q_{i3} \end{bmatrix},$$

которая имеет два собственных числа

$$\lambda_{1,2} = \frac{1}{2}\left(q_{i1} + q_{i3} \pm \sqrt{q_{i1}^2 - 4q_{i2}^2 + q_{i3}^2 + 6q_{i1}q_{i3}}\right).$$

Поскольку решается задача максимизации, функция Беллмана должна быть отрицательно полу-определенной, что невозможно, учитывая, что коэффициенты q_{11} и q_{13} положительны. Таким образом, ни одна из компонент решения не удовлетворяет данному критерию.

3.3.1. Графическая интерпретация

Рассмотрим допустимое решение, полученное с помощью экономического критерия. Равновесные по Нэшу управления

$$a_1^{NE} = \frac{5k}{3} + \frac{\rho}{8}(2x_1 + 3x_2),$$
$$a_2^{NE} = \frac{5k}{3} + \frac{\rho}{8}(3x_1 + 2x_2).$$

Подставим их в систему (3.1) и решим ее с начальными условиями $x_1(0) = 0,\ x_2(0) = 2$. Построим графики переменных в зависимости от времени $t,\ t \in [0, \infty)$, при фиксированных коэффициентах $k = 0,46$, $\rho = 0,27$. На рис. (3.1) изображены оптимальные траектории $x_1^{NE}(t),\ x_2^{NE}(t)$. На рис. (3.2) изображены равновесные по Нэшу управления $a_1^{NE}(t),\ a_2^{NE}(t)$. Графики зависимости выигрышей $J_1(x^{NE}, a^{NE}),\ J_2(x^{NE}, a^{NE})$ игроков от $\rho = \overline{0,001; 0,27}$ при $k = 0,46$ изображены на рисунке (3.3).

Графики иллюстрируют выводы, полученные в ходе применения к отбору допустимых решений LQR метода, которые говорят о том, что уравнения Гамильтона–Якоби–Беллмана не имеют квадратичных решений. Таким образом, в данной задаче не удалось отобрать допустимые равновесия по Нэшу в рассматриваемом классе функций.

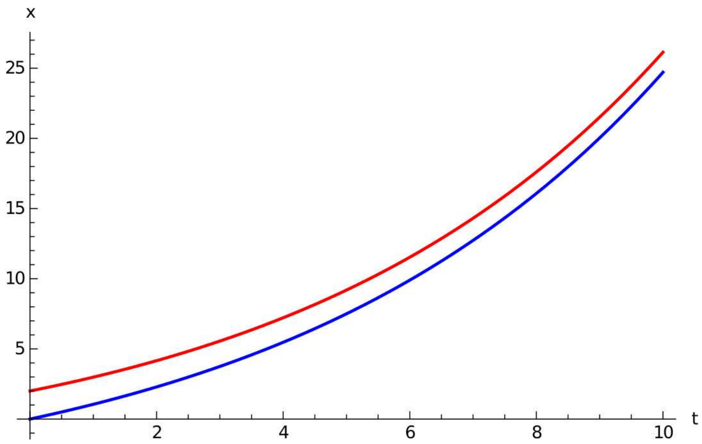


Рис. 3.1: Оптимальные траектории $x_1^{NE}(t),\ x_2^{NE}(t)$

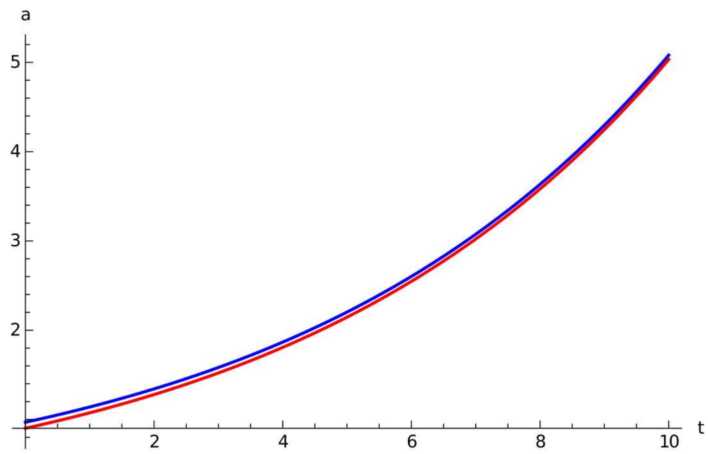


Рис. 3.2: Равновесие по Нэшу $a_1^{NE}(t),\ a_2^{NE}(t)$

§ 3.4. Кооперативный случай

Рассмотрим случай кооперации игроков, тогда их суммарный выигрыш представим в следующем виде:

$$J(a,x) = \int_0^{\infty} e^{-\rho t}\left(\pi\left(k - \sum_{h=1}^{n} a_h\right)\sum_{i=1}^{n} a_i - \sum_{i=1}^{n} x_i\right)dt. \tag{3.31}$$

Решаем задачу с помощью уравнения Гамильтона–Якоби–Беллмана, которое в кооперативной постановке будет иметь вид:

$$\rho V(x) = \max_{a_1,\ldots,a_n}\left\{\pi\left(k - \sum_{h=1}^{n} a_h\right)\sum_{i=1}^{n} a_i - \sum_{i=1}^{n} x_i + \sum_{i=1}^{n}\frac{\partial V}{\partial x_i}a_i\right\}. \tag{3.32}$$

Будем рассматривать игру для двух симметричных игроков, предположим, что $a_1 = a_2 = a$, а функция Беллмана, которая является решением урав-

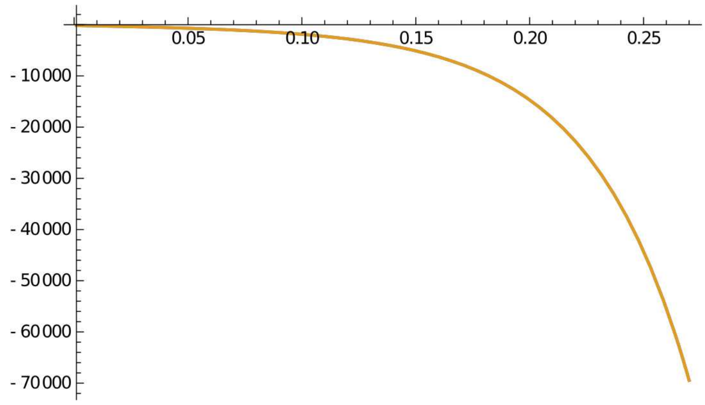


Рис. 3.3: Выигрыши игроков $J_1(x^{NE}, a^{NE})$, $J_2(x^{NE}, a^{NE})$

нения (3.32), выбирается в виде

$$V(x) = \alpha + \frac{\varepsilon}{2}\left(x_1^2 + x_2^2\right) + \varepsilon x_1 x_2 + \gamma(x_1 + x_2). \tag{3.33}$$

Поскольку функция $V(x)$ непрерывно-дифференцируемая, то ее производная по переменным состояния в случае симметричных игроков:

$$\frac{\partial V(x)}{\partial x_1} = \frac{\partial V(x)}{\partial x_2} = \varepsilon(x_1 + x_2) + \gamma. \tag{3.34}$$

Подставляя эту производную в (3.32), получим:

$$\rho V(x) = \max_a \left\{ -4\pi a^2 + 2\pi k a - (x_1 + x_2) + 2\frac{\partial V}{\partial x_1} a \right\}. \tag{3.35}$$

Рассмотрим функцию под максимумом правой части предыдущего уравнения:

$$F(a) = -4\pi a^2 + 2\pi k a - (x_1 + x_2) + 2(\varepsilon(x_1 + x_2) + \gamma)a. \tag{3.36}$$

Найдем a^*, которое доставляет максимум функции $F(a)$:

$$\frac{\partial F(a)}{\partial a} = 2\pi k - 8\pi a + 2(\varepsilon(x_1 + x_2) + \gamma) = 0 \ \Rightarrow$$

$$a^* = \frac{k}{4} + \frac{\varepsilon(x_1 + x_2) + \gamma}{4\pi}.$$

подставляем его в уравнение Гамильтона–Якоби–Беллмана.

Распишем полученное уравнение Гамильтона–Якоби–Беллмана:

$$\rho\left(\alpha + \frac{\varepsilon}{2}\left(x_1^2 + x_2^2\right) + \varepsilon x_1 x_2 + \gamma(x_1 + x_2)\right) = -4\pi\left(\frac{k}{4} + \frac{\varepsilon(x_1 + x_2) + \gamma}{4\pi}\right)^2$$

$$+2\pi k(\frac{k}{4}+\frac{\varepsilon(x_1+x_2)+\gamma}{4\pi})-(x_1+x_2)+2\left(\varepsilon(x_1+x_2)+\gamma\right)\left(\frac{k}{4}+\frac{\varepsilon(x_1+x_2)+\gamma}{4\pi}\right).$$

Методом неопределенных коэффициентов придем к системе уравнений:

$$x_1^2, x_2^2 : \rho\frac{\varepsilon}{2} = \frac{\varepsilon^2}{4\pi}; \tag{3.37}$$

$$x_1 x_2 : \rho\varepsilon = \frac{\varepsilon^2}{2\pi}; \tag{3.38}$$

$$x_1, x_2 : \gamma\rho = \frac{8\varepsilon\gamma - 16\pi + 8\varepsilon k\pi}{16\pi}; \tag{3.39}$$

$$x_1^0 x_2^0 : \alpha\rho = \frac{4\gamma^2 + 8\gamma k\pi + 4k^2\pi^2}{16\pi}. \tag{3.40}$$

Рассмотрим уравнение (3.37). Решая его относительно ε, получим, что либо $\varepsilon = 2\pi\rho$, либо $\varepsilon = 0$. Аналогичные решения получаем и из (3.38). Рассмотрим первое решение. Пусть $\varepsilon = 2\pi\rho$, подставим его в (3.39). Видим, при таком ε не можем найти выражение для γ, из (3.39) находим только условие, накладываемое на коэффициенты модели: $k\pi\rho = 1$.

Тогда рассмотрим второе решение (3.37). Пусть $\varepsilon = 0$, тогда из (4.7) получаем $\gamma = -\frac{1}{\rho}$.

Подставим $\gamma = -\frac{1}{\rho}$ в последнее уравнение системы (3.40) и найдем, что $\alpha = \frac{(1+k\pi\rho)^2}{4\rho^3\pi}$.

Таким образом, заработанный двумя игроками суммарный выигрыш:

$$V(2, x_0) = \frac{(1+k\pi\rho)^2}{4\rho^3\pi} - \frac{(x_1+x_2)}{\rho}, \tag{3.41}$$

при выборе стационарных оптимальных управлений:

$$a_1^* = a_2^* = \frac{k}{4} - \frac{1}{4\pi\rho}, \tag{3.42}$$

которые не зависят от переменной состояния. Учитывая, что управления игроков мы предполагаем неотрицательными, накладывается условие на параметры модели:

$$k\pi\rho \geqslant 1. \tag{3.43}$$

3.4.1. Графическая интерпретация

Возьмем в качестве начальных значений объема выбросов игроков $x_0^1 = 2,\ \ x_0^2 = 0$. Подставим данные начальные условия и оптимальные управления $a_1^*,\ a_2^*$ в виде (3.42) в систему дифференциальных уравнений

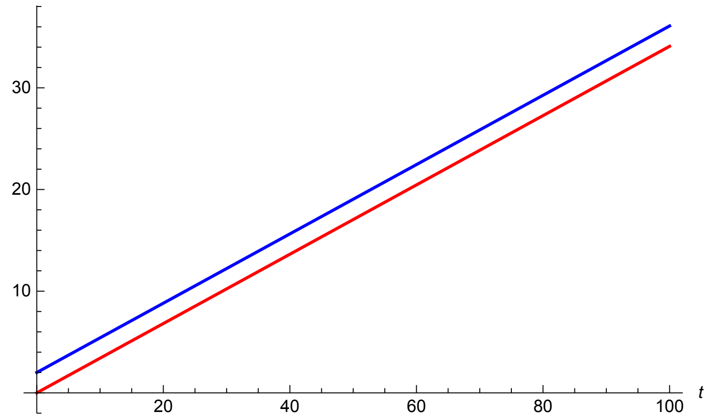


Рис. 3.4: Оптимальные траектории игроков $x_1^*(t),\ x_2^*(t)$

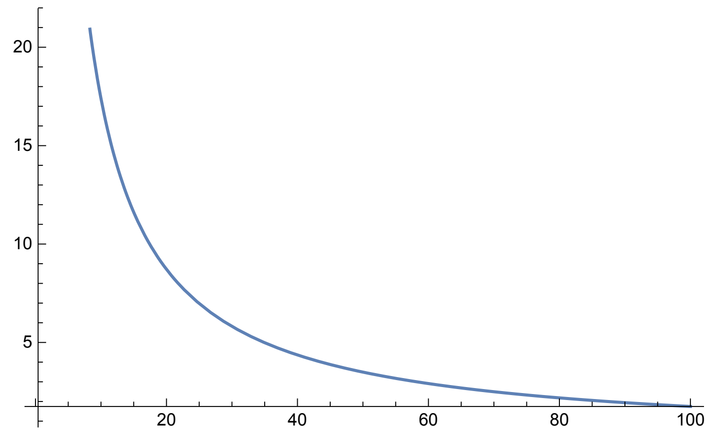


Рис. 3.5: Кооперативный выигрыш $J(a^*, x^*)$

(3.1) при $i = 1, 2$. Решая ее, найдем оптимальные траектории:

$$x_1^* = \frac{t(k\pi\rho - 1)}{4\pi\rho} + 2, \quad x_2^* = \frac{t(k\pi\rho - 1)}{4\pi\rho},$$

которые изображены в зависимости от времени $t,\ t \in [0, \infty)$ на рис. (3.4) при фиксированных коэффициентах $k = 15,\ \rho = 0,4$, удовлетворяющих условию (3.43). Подставим оптимальные управления $a_1^*,\ a_2^*$ и оптимальные траектории $x_1^*(t),\ x_2^*(t)$ в (3.31) и найдем общую сумму выигрыша игроков:

$$J(a^*, x^*) = \frac{1 + \pi\rho(k(k\pi\rho - 2) - 8\rho}{4\pi\rho^3}.$$

На рис. (3.5) изображен график зависимости кооперативного выигрыша игроков $J(a^*)$ в зависимости от $\rho \geqslant 0,5$ при $k = 15$.

Полученные графики объективно интерпретируют процесс с экономической точки зрения. В процессе игры при производстве объем вредных выбросов возрастает, но, поскольку управления игроков постоянны $\forall t,\ t \in [0, \infty)$, то общая сумма выигрыша уменьшается при постоянной ставке дисконтирования.

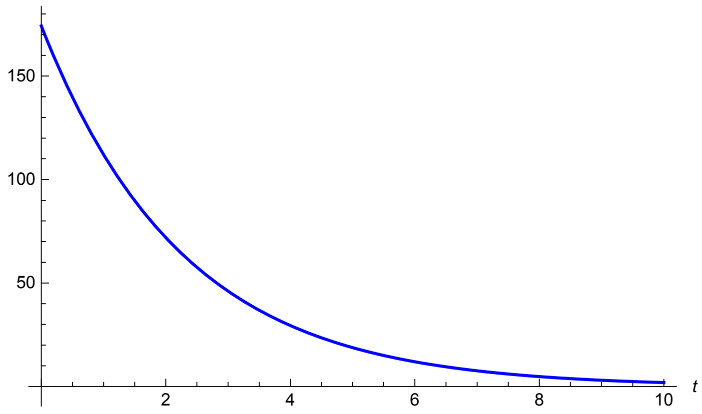


Рис. 3.6: $\beta_i(t),\ i=\{1,2\}$ при $k=15,\ \rho=0,4$

3.4.2. Распределение кооперативного выигрыша

В результате кооперации оба игрока заработали максимальную сумму (3.41). В качестве принципа оптимальности рассмотрим вектор Шепли (1.8). В силу симметричности игроков, каждый игрок получит выигрыш:

$$Sh_i(x_0,0)=\frac{V(2,x_0)}{2}=\frac{(1+k\pi\rho)^2}{8\rho^3\pi}-\frac{1}{\rho}.$$

Согласно процедуре распределения дележа (2.11), учитывая, что

$$h_i(x^*,a^*)=\pi(k-(a_1+a_2))a_i-x_i,\ i=\{1,2\},$$

каждый игрок в каждый момент времени t получит:

$$\beta_i(t)=e^{-\rho t}\sum_{i=1}^{2}h_i(x^*,a^*)=e^{-\rho t}\left(\frac{\pi}{4}\left(k^2-\frac{1}{\pi^2\rho^2}\right)-\frac{t(k\pi\rho-1)}{2\pi\rho}-2\right).$$

При распределении компонент дележа согласно полученной вектор-функции (ПРД) $\{\beta_i(t)\}$ у игроков не возникнет оснований для отклонения от кооперативного решения в течение реализации всего долгосрочного проекта по контролю объемов вредных загрязнений в окружающую среду. Следовательно, и сама кооперативная траектория $x^*(t)$, соответствующая накоплению загрязнений, будет устойчивой и может быть спрогнозирована.

Графическая интерпретация полученных результатов приводится на рис. (3.6).

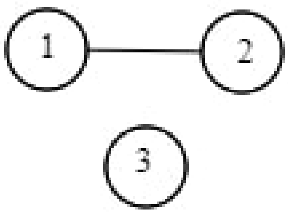
Глава 4.

Сетевая дифференциальная игра управления объемами вредных выбросов

Поскольку в уравнениях динамики модели управления вредными выбросами, рассмотренной в главе 3, имеют место «независимые движения», то для данной задачи можно построить сетевую модель [29].
Будем рассматривать дифференциальную игру трех игроков в предположении, что на выигрыш игрока влияют лишь те игроки, которые связаны с ним дугой.

§ 4.1. Постановка задачи

Пусть $N = \{1, 2, 3\}$ — множество игроков (узлов) объединенных в сеть 1.4. Предположим, что в сети есть только одно ребро, то есть $L = \{(1, 2)\}$. Рассмотрим следующую неориентированную сеть.



Будем считать, что в определенной выше игре на сети существует коалиция $S = \{1, 2\}$. Для каждого игрока i, $i = \overline{1, 3}$ определим множества $K(i, 0)$ всех игроков, с которыми у игрока i в данной сети имеется связь в начальный момент времени:

$$K(1, 0) = \{1, 2\}, \quad K(2, 0) = \{1, 2\}, \quad K(3, 0) = \{3\}.$$

Как было определено в предыдущей главе 3, a_i — управление игрока i. Тогда $U^S = \{a_1, a_2\}$ — набор управляющих переменных игроков из коалиции

S, а $U^3 = \{a_3\}$ состоит только из управления игрока 3, поскольку он не входит ни в какую коалицию (то есть входит в коалицию, состоящую из него самого). Пусть $x = (x_1, x_2, x_3)$ вектор состояния игроков.

Уравнение движения для игрока i, $i = \overline{1,3}$ определяется согласно предыдущей главе 3:

$$\dot{x}_i = a_i, \quad x_i(0) = x_i^0 \geqslant 0, \tag{4.1}$$

Вспомним из 1.4, что $\varphi_i^i(x(t), t) \equiv 1$, тогда на множестве допустимых траекторий системы (4.1) для каждого игрока i определим начальные значения вектор-функции $\varphi_i(x^0, 0)$, которая указывает на наличие дуг, связывающих игрока i с другими игроками в момент времени $t = 0$:

$$\varphi_1(x^0, 0) = \{1, 1, 0\}, \quad \varphi_2(x^0, 0) = \{1, 1, 0\}, \quad \varphi_3(x^0, 0) = \{0, 0, 1\}.$$

В процессе игры могут изменяться множество дуг L и множество $K(i, t)$.

§ 4.2. Функции выигрыша

Определим функции выигрыша игроков $i = \overline{1,3}$. Положим, что функция полезности игрока $i = \overline{1,3}:\ h_i(x_i, x_i) = 0,\ h_i > 0$.
Функция выигрыша игрока 3, который не входит ни в какую коалицию, зависит от переменной его состояния и переменных состояния игроков из множества $K(i, t)$:

$$H_3(x_3^0, x_{K(3,0)}^0; a_3, a_{K(3,0)}) = \int\limits_0^\infty e^{-\rho t} \sum_{j \in K(3,t)} h_3(x_3, x_j) d\tau = \int\limits_0^\infty e^{-\rho t} h_3(x_3, x_3) d\tau = 0,$$

так как $K(3, t) = \{3\},\ \forall t,\ t \in [0, \infty)$. Здесь $x_{K(i,0)}^0 = \{x_j^0;\ j \in K(i, 0)\}$.
Выигрыш коалиции $S \subset N$, согласно (1.13):

$$H(S) = \int\limits_0^\infty e^{-\rho t} \sum_{i \in S} \sum_{j \in K(i,t) \cap S} h_i(x_i, x_j) d\tau, \tag{4.2}$$

где функция полезности определяется согласно предыдущей главе 3:

$$h_i(x_i, x_j) = \pi(k - a_j) a_i - x_i.$$

Множества $K(i, t)$ для игроков i коалиции S меняются в процессе игры. Перед началом игры множества $K(i, t)$ содержат по два элемента. Поскольку множество дуг $L = \{(1, 2)\}$ сети состоит только из одного элемента, игроки в процессе игры могут удалить только дугу $(1, 2)$. В этом случае их множества $K(i, t)$ будут состоять только из одного элемента — самого игрока, для которого определено данное множество. Тогда функция полезности будет иметь вид $h_i(x_i, x_i)$. А выше было определено, что $h_i(x_i, x_i) = 0$. С учетом вышесказанного выигрыш коалиции (4.2) запишется в следующем виде:

$$H(S) = -\pi(a_1^2 + a_2^2) + \pi k(a_1 + a_2) - 2\pi a_1 a_2 - (x_1 + x_2). \tag{4.3}$$

§ 4.3. Решение коалиционной игры

Заметим, что при определении выигрыша коалиции S мы используем тот факт, что игроки, входящие в коалицию S могут разрывать связи с игроками из $N \setminus S$, поскольку те могут действовать против игроков из S.

Из (4.3) заметим, что задачу можно свести к задаче максимизации выигрыша для одного игрока, если положим $x_1 + x_2 = x$ — переменная состояния игрока, $a_1 + a_2 = a$ — его управление. В этом случае функция выигрыша (4.3) перепишется следующим образом:

$$H(S) = -\pi a^2 + \pi k a - x.$$

Для решения коалиционной игры воспользуемся уравнением Гамильтона–Якоби–Беллмана, которое в рамках данной задачи запишется в виде

$$\rho V(x) = \max_{a \geqslant 0} \left\{ -\pi a^2 + \pi k a - x + \frac{\partial V}{\partial x} a \right\}. \tag{4.4}$$

Решение этого уравнения $V(x)$ — функция Беллмана, выберем ее в виде:

$$V(x) = \varepsilon x^2 + \gamma x + \alpha.$$

Поскольку функция $V(x)$ непрерывно-дифференцируемая, то ее производная по переменной состояния:

$$\frac{\partial V(x)}{\partial x} = 2\varepsilon x + \gamma,$$

Подставляя эту производную в (4.4), получим:

$$\rho(\varepsilon x^2 + \gamma x + \alpha) = \max_{a \geqslant 0} \left\{ -\pi a^2 + \pi k a - x + (2\varepsilon x + \gamma) a \right\}. \tag{4.5}$$

Рассмотрим функцию под максимумом правой части предыдущего уравнения:

$$F(a) = -\pi a^2 + \pi k a - x + (2\varepsilon x + \gamma) a.$$

Найдем управление a^*, которое доставляет максимум функции $F(a)$:

$$\frac{\partial F(a)}{\partial a} = -2\pi a + \pi k + 2\varepsilon x + \gamma = 0.$$

$$a^* = \frac{k}{2} + \frac{2\varepsilon x + \gamma}{2\pi}.$$

Подствим его в уравнение Гамильтона–Якоби–Беллмана (4.5):

$$\rho(\varepsilon x^2 + \gamma x + \alpha) = -\pi\left(\frac{k}{2} + \frac{2\varepsilon x + \gamma}{2\pi}\right)^2 + \pi k\left(\frac{k}{2} + \frac{2\varepsilon x + \gamma}{2\pi}\right) -$$

$$-x + (2\varepsilon x + \gamma)\left(\frac{k}{2} + \frac{2\varepsilon x + \gamma}{2\pi}\right).$$

Применяя метод неопределенных коэффициентов к полученному уравнению, мы придем к системе уравнений относительно коэффициентов ε, γ, α:

$$x^2 : \rho\varepsilon = \frac{\varepsilon^2}{\pi}; \tag{4.6}$$

$$x : \gamma\rho = \frac{\varepsilon\gamma - \pi + \varepsilon k\pi}{\pi}; \tag{4.7}$$

$$x^0 : \alpha\rho = \frac{\gamma^2 + 2\gamma k\pi + k^2\pi^2}{4\pi}. \tag{4.8}$$

Рассмотрим уравнение (4.6). Решая его относительно ε, получим, что либо $\varepsilon = \pi\rho$, либо $\varepsilon = 0$. Пусть $\varepsilon = \pi\rho$, тогда из (4.7) находим только, что $k\pi\rho = 1$.

Тогда рассмотрим второе решение (4.6). Пусть $\varepsilon = 0$, тогда из (4.7) получаем $\gamma = -\frac{1}{\rho}$.

Подставим $\gamma = -\frac{1}{\rho}$ в (4.6) и найдем $\alpha = \frac{(1+k\pi\rho)^2}{4\rho^2\pi}$.

Таким образом, заработанный двумя игроками из коалиции S выигрыш:

$$V(S, x_0) = \frac{(1 + k\pi\rho)^2}{4\rho^2\pi} - \frac{x}{\rho}.$$

при выборе оптимального управления:

$$a^* = \frac{k}{2} - \frac{1}{2\pi\rho},$$

которое не зависит от переменной состояния. Учитывая, что управление мы предполагаем неотрицательным, накладывается условие на параметры модели: $k\pi\rho \geqslant 1$.

Третий игрок действует независимо от коалиции S и использует максимизирующее управление a_3^*.

Решение кооперативной игры находится аналогично приведенному выше методу решения.

Вывод

В данной работе были рассмотрены математические модели управления ресурсами с многомерной фазовой переменной (2.1)-(2.2) и (3.1)-(3.2). Под ресурсами понимались как объемы выпуска продукции, пропорциональные объемам вредных выбросов в окружающую среду, так и вложения в рекламную кампанию различными фирмами. Методика решения и отбора решений в случае их неединственности использовалась аналогичная.

Дифференциальные игры управления ресурсами рассматривались на бесконечном временном промежутке $t \in [t_0, \infty)$ с функцией дисконтирования $e^{-\rho t}$ с постоянной ставкой дисконтирования ρ, кроме того, спецификой задачи являлся линейно-квадратичный вид задачи для случая нескольких фазовых переменных. Для нахождения решений данных задач в классе позиционных стратегий было использовано уравнение Гамильтона–Якоби–Беллмана в кооперативном случае (2.4) и система уравнений в некооперативном (3.4). Решением этого уравнения является функция Беллмана, рассмотренная в данной работе в линейно-квадратичном виде для кооперативного (2.5) случая. Для некооперативного случая равновесие по Нэшу является результатом решения системы уравнений, в которых соответствующие функции Беллмана также выбираются в линейно-квадратичной форме (3.5). В результате решения задачи в данном классе стратегий решение определяется неединственным образом. В связи с чем была предпринята попытка использовать два метода для отбора допустимых решений: математический подход (1.2.2) и экономический критерий (1.2.1). Графическая интерпретация отражает полученные результаты.

Важным свойством, необходимым для реализации кооперативных решений во времени, является свойство динамической устойчивости кооперативных решений. Для случая симметричных игроков данная проблема была решена для моделей из Главы II, Главы III при помощи введения так называемой процедуры распределения дележа (ПРД), согласно которой компоненты выигрыша распределены во времени. Все выражения для ПРД были получены в аналитическом виде в общем случае.

Кроме того, была изучена модель дифференциальной игры с сетевой структурой на основе модели, сформулированной в Главе III. Выражения для управления игроков из коалиции S, связанных в сети, были получены в аналитическом виде.

Дальнейшие перспективы исследования заключаются в поиске кооперативного решения на сети и распределения компонент дележа согласно ПРД.

Заключение

В ходе проделанной работы были рассмотрены модели управления в дифференциальных играх двумя типами ресурсов: инвестирование в рекламу (в кооперативной постановке для n игроков) и выбросы в окружающую среду (в кооперативной и некооперативной постановке для двух игроков).

Показано, что уравнение Гамильтона–Якоби–Беллмана имеет неединственное решение, которое требует изучения для отбраковки несостоятельных решений. В ходе применения экономического критерия и классического метода, используемого для линейно-квадратичных задач оптимизации (LQR), было установлено:

дифференциальная игра управления рекламой имеет допустимое решение, которое удовлетворяет обоим критериям;

игра управления вредными выбросами в некооперативной постановке не имеет решений в указанных классах функций, полученные результаты также были проинтерпретированы графически.

Кроме того, в случае кооперативной постановки задачи в рассматриваемых моделях было найдено распределение общего выигрыша между игроками в соответствии с выбранными принципами оптимальности, а именно, вектором Шепли, а затем распределение его компонент во времени согласно процедуре распределения дележа (ПРД).

Также для модели управления выбросами была построена дифференциальная игра на сети для трех игроков. Выражения для управлений игроков в коалиционной игре были получены в аналитическом виде.

Таким образом, поставленные цели и задачи были достигнуты.

Список литературы

[1] Беллман Р. *Динамическое прграммирование* М.:И.Л., 1960.

[2] Красовский Н. Н., Субботин А. И. Позиционные дифференциальные игры. М.: Наука, 1974. 455 с.

[3] Frutos J., Martin-Herran G. Selection of a Markov perfect Nash equilibrium in a class of differential games. Dynamic Games and Applications. Vol. 8, № 3. 2018. P. 620–636.

[4] Fershtman C., Nitzan S. Dynamic voluntary provision of public goods. European Economic Review. Vol. 35, № 5. 1991. P. 1057–1067.

[5] Wirl F. Dynamic voluntary provision of public goods: extension for nonlinear strategies. European Journal of Political Economy. Vol. 12, № 3. 1996. P. 555–560.

[6] Dockner E. J. Optimal pricing in a dynamic duopoly game model. Zeitschrift fur Operations Research. Basar T. Informationally nonunique equilibrium solutions in differential games. 1985. P. 1–16.

[7] Leitmann G., and Schmitendorf W. E. Profit maximization through advertising: nonzero-sum differential game approach. IEEE Transactions on Automatic Control. Vol. 23, № 4. 1978. P. 645–650.

[8] Basar T. Informationally nonunique equilibrium solutions in differential games. SIAM Journal on Control and Optimization. Vol. 15, № 4. 1977. P. 636—660.

[9] Driskill R. Durable goods monopoly, increasing marginal cost and depreciation. Economica. Vol. 64, № 253. 1997. P. 137—154.

[10] Cartigny P., Michel P. On the Selection of One Feedback Nash Equilibrium in Discounted Linear-Quadratic Games. Journal of Optimization Theory and Applications. Vol. 117, № 2. 2003. P. 231—243.

[11] Громова Е. В., Громов Д. В., Лахина Ю. Э. О дифференциальной игре управления инвестициями в рекламную кампанию. Труды института математики и механики УРО РАН. Т. 24, № 2. 2018. С. 64–75.

[12] Петросян Л. А., Зенкевич Н. А., Шевкопляс Е. В. Теория игр. СПБ.: Изд-во БХВ-Петербург, 2012. 432 с.

[13] Жуковский В. И., Чикрий А. А. Линейно-квадратичные дифференциальные игры. Киев: Науковая думка, 1994. 320 с.

[14] Olsder G. G., Basar T. Dynamic Noncooperative Game Theory, 2nd Edition. New York: Academic Press, 1999. 519 p.

[15] Dockner E. J., Jorgensen S., Long N. V., Sorger G. Differential Games in Economics and Management Science. Cambridge: Cambridge University Press, 2000. 396 p.

[16] Jorgensen S., Zaccour G. Differential Games in Marketing. Boston: Kluwer Academic Publisher. 2004. 159 p.

[17] Papavassilopoulos G., Olsder G. J. On the linear-quadratic, closedloop, nomemory Nash game. Journal of Optimization Theory and Applications. Vol. 42, № 4. 1984. P. 551—560.

[18] Reddy P. V., Zaccour G.: Feedback Nash equilibria in linear-quadratic difference games with constraints. IEEE Transactions on Automatic Control. Vol. 62, № 2. 2017. P. 590—604.

[19] Шевкопляс Е. В. Устойчивая кооперация в дифференциальных играх со случайной продолжительностью. УБС, 31.1. 2010. С. 162—190.

[20] Singh R., Wisznieszka-Matyszkiel A. Linear-quadratic game of exploitation of common renewable resources with inherent constraints. Topological Methods in Nonlinear Analysis. Vol. 51, № 1. 2018. P. 23—54.

[21] Dockner E.J., Sorger G.: Existence and properties of equilibria for a dynamic game on productive assets. J. Economic Theory. Vol. 71, № 1. 1996. P. 209—227.

[22] Singh R., Wisznieszka-Matyszkiel A.: Discontinuous Nash Equilibria in a Two Stage Linear-Quadratic Dynamic Game with Linear Constraints. IEEE Transactions on Automatic Control. Vol. 64, № 7. 2018. P. 3074–3079.

[23] Kononenko A.: The structure of the optimal strategy in controlled dynamic systems. USSR Computational Mathematics and Mathematical Physics. Vol. 20, № 5. 1980. P. 13–24.

[24] Malafeev O.: Stationary strategies in differential games. USSR Computational Mathematics and Mathematical Physics. Vol. 17, № 1. 1977. P. 37–46.

[25] Basar T.: On the uniqueness of the Nash solution in linear-quadratic differential games. International Journal of Game Theory. Vol. 5, № 2-3. 1976. P. 65—90.

[26] Gromova E., Lakhina Yu. On the selection of the Nash equilibria in a linear-quadratic differential game of pollution control. Frontiers of Dynamic Games. Ed. by L. Petrosyan, V. Mazalov, N. Zenkevich. St. Petersburg. 2018. P. 37–48.

[27] Bass F. M., Krishnamoorthy A., Prasad A., Sethi S. P. Generic and brand advertising strategies in a dynamic duopoly. Marketing Science. Vol. 24, № 4. 2005. P. 556–568.

[28] Александров А. Г. Оптимальные и адаптивные системы: Учеб. пособие для вузов по спец. «Автоматика и упр. в техн. системах». М.: Высшая школа, 1989. 263 с.

[29] Петросян Л. А. Кооперативные дифференциальные игры на сетях. Тр. Ин-та математики и механики УрО РАН. Т. 16, № 5, 2010. С. 143–150.

[30] Мазалов В. В. Математическая теория игр и приложения. Спб.: Изд-во Лань, 2010. 448 с.

[31] Жабко А. П., Котина Е. Д., Чижова О. Н. Дифференциальные уравнения и устойчивость. Спб.: Изд-во Лань, 2015. 310 с.

[32] Лахина Ю. Э. Оптимальное управление инвестициями в рекламу на рынке однородной продукции. Процессы управления и устойчивость. 2018. Т. 5. № 1. С. 470–474.

[33] Лахина Ю. Э. О существовании равновесия по Нэшу в некооперативной дифференциальной игре управления объемами вредных выбросов. Процессы управления и устойчивость. 2020.

[34] Лахина Ю. Э. Оптимальное управление в задаче эксплуатации нескольких ресурсов. Архив открытого доступа Санкт-Петербургского государственного университета. 2018. 45 с.