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

Алексеева Александра Александровна

Выпускная квалификационная работа бакалавра

О методах решения линейно-квадратичных дифференциальных игр

Направление 01.03.02

Прикладная математика и информатика

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

Научный руководитель,
доктор физ.-мат.наук,
доцент
Громова Е. В.

Рецензент,
кандидат физ.-мат. наук,
доцент
Егоров А. В.

Санкт-Петербург
2020

Содержание

Введение

Дифференциальные игры являются удобным математическим аппаратом для изучения различного рода конфликтно-управляемых задач. В современной экономике огромное число разнообразных процессов может быть описано именно с помощью дифференциальных игр, очень часто имеющих линейно-квадратичную структуру. В частности, такие разноплановые, но актуальные в настоящий момент экономические задачи, как управление объемом вредных выбросов или инвестирование компанией средств в рекламу, описываются именно дифференциальными линейно-квадратичными играми.

В связи с широким использованием упомянутых моделей, неизбежно, что решение игр данного типа носит приоритетный характер. В линейно-квадратичных моделях можно выделить отдельные классы игр, которые могут быть решены за счет некоторых своих структурных особенностей. Именно таким классам игр, а также методам, позволяющим упростить поиск их аналитического решения, и посвящена данная работа.

В главе I осуществляется формальная постановка задачи оптимального управления эксплуатацией ресурсов, объединяющей в себе в том числе и задачу управления объемом вредных выбросов, и задачу управления инвестициями в рекламу. Рассматривается теоретическая база, используемая для решения задач данного типа: принцип максимума Понтрягина и уравнение Гамильтона—Якоби—Беллмана. Дополнительно рассмотрены метод упрощения интегрального функционала линейно-квадратичной дифференциальной игры и метод потенциала. Данные методы позволяют облегчить поиск аналитического решения линейно-квадратичных дифференциальных игр в тех случаях, когда применение принципа максимума Понтрягина и уравнения Гамильтона—Якоби—Беллмана затруднено.

В главе II более подробно рассматривается метод, позволяющий упростить интегральный функционал линейно-квадратичной дифференциальной игры путем замены фазовой переменной. Детально изучается применение данного метода на примере задачи оптимального управления вредными выбросами в кооперативной постановке для числа игроков $n = 3$. Аналитическое решение данной задачи будет найдено двумя способами: с помощью принципа максимума Понтрягина и с помощью уравнения Гамильтона—Якоби—Беллмана. Также рассматривается численный пример задачи, в котором использованы значения параметров, полученные экспериментальным путем.

В главе III рассмотрено применение метода потенциала при решении линейно-квадратичных дифференциальных игр на примере задачи оптимального управления инвестициями в рекламу. Игра изучается в некооперативной постановке с дисконтированием $e^{-\rho t}$ на бесконечном промежутке времени для числа игроков $n = 2$. Рассматривается также численный пример задачи.

В главе IV рассмотрен дополнительный метод решения линейно-квадратичной дифференциальной игры, применимый в том случае, когда известен вид искомого оптимального управления. Метод будет применен для решения задачи оптимального управления вредными выбросами в некооперативной постановке для числа игроков $n = 3$.

Обзор литературы

Линейно-квадратичные дифференциальные игры являются одним из наиболее изученных классов дифференциальных игр в целом. В первую очередь они нашли свое применение в ряде экономических задач [1, 2].

Решение задачи оптимального управления инвестициями в рекламу с помощью теоретико-игрового подхода было рассмотрено в работах [1, 3]. Данная задача была также решена методом динамического программирования [12, 13] в классе позиционных стратегий [5] уже для случая трех симметричных игроков в статье [4], где игра рассматривалась как в кооперативной постановке [3], так и в некооперативной [1]. В работах [8, 9, 19] было предложено решение кооперативного варианта задачи оптимального управления ресурсами в рекламную кампанию с использованием преобразования фазовой переменной, позволяющем упростить подынтегральную функцию путем сокращения ее линейной составляющей.

Широко изучена также теоретико-игровая модель управления вредными выбросами, представленная в работе [2]. В статье [6] данная дифференциальная игра была решена для случая двух игроков в некооперативной постановке. Кооперативный случай игры управления вредными выбросами был изучен в работе [7], где решение задачи было найдено с использованием метода максимума Понтрягина [10].

Особый интерес из себя представляют потенциальные игры [14]. Статические потенциальные игры впервые появились в работе [15]. Позднее потенциальные игры в динамическом варианте были исследованы в работах [16, 17]. В играх такого типа поиск равновесия по Нэшу [3] существенно упрощается с помощью использования потенциала динамической игры. В работе [18] была изучена потенциальная линейно-квадратичная теоретико-игровая модель экологического менеджмента в дискретном времени с бесконечным горизонтом планирования, где было показано, что потенциал для линейно-квадратичной игры представляет собой квадратичную форму. В исследовании [14] были обоснованы достаточные условия существования потенциала дифференциальной динамической игры, а также были приведены примеры потенциальных дифференциальных игр.

Глава 1. Линейно-квадратичные дифференциальные игры эксплуатации ресурсов

Общее описание игры

Введем в рассмотрение линейно-квадратичную дифференциальную игру эксплуатации некоторого ресурса для n лиц. Пусть момент времени t_0, где $t_0 \geq 0$, служит моментом начала игры, и пусть игра завершается в момент времени T. В начальный момент времени состояния игроков определяются как x_0.

В общем виде динамика процесса накопления (убывания) ресурса описывается системой дифференциальных уравнений с начальным условием:

$$\dot{x} = f(t, x, u), \quad x(t_0) = x^0, \tag{1}$$

где $x \in X \subseteq \mathbb{R}^m$, $u \in U = U_1 \times \ldots \times U_n \subseteq \mathbb{R}^n$, $t \in [t_0, T]$. Предполагается, что функция $f : \mathbb{R}^1 \times \mathbb{R}^m \times \mathbb{R}^n \longrightarrow \mathbb{R}^m$ является липшицевой по x, непрерывной на множестве $[t_0, T] \times X \times U_1 \times \ldots \times U_n$ и дифференцируемой.

Интегральный выигрыш i-го игрока в игре определяется как:

$$K_i(t_0, x_0, u) = \int_{t_0}^{T} L_i(t, x(t), u(t))dt + h_i(x(T)), \quad i = \overline{1, n}, \tag{2}$$

где $h_i(x(T))$ - терминальный выигрыш i-го игрока, а функция $L_i(t, x(t), u(t))$ имеет линейно-квадратичную структуру и в общем виде определяется как:

$$L_i(\cdot) = r_i^T u + u^T R_i u + q_i^T x + x^T Q_i x, \quad i = \overline{1, n}.$$

Здесь R_i, Q_i - матрицы соответствующих размерностей; r_i, q_i - вектор-столбцы.

В случае, если игра рассматривается на бесконечном промежутке времени, интегральный выигрыш $K_i(t_0, x_0, u)$ определяется как:

$$K_i(t_0, x_0, u) = \int_{t_0}^{\infty} L_i(\cdot)\Theta(t)dt, \quad i = \overline{1, n}, \tag{3}$$

где $\Theta(t)$: $\Theta(0) = 1$ — невозрастающая дисконтирующая функция. Чаще всего в задачах встречается дисконтирующий множитель вида:

$$\Theta(t) = e^{-\rho t}, \quad \text{где } \rho > 0.$$

В некооперативной постановке задачи каждый игрок стремится максимизировать свою функцию выигрыша, действуя независимо от остальных игроков, т.е. выбранная стратегия игрока i в игре должна обеспечивать ему максимально возможный выигрыш при любых неизвестных ему стратегиях, выбранных остальными игроками. Таким образом, решение поставленной задачи (равновесие по Нэшу) находится в результате одновременного решения системы оптимизационных задач:

$$\begin{aligned}
&K_1(t_0, x_0, u) \to \max_{u_1}, \\
&\vdots \\
&K_n(t_0, x_0, u) \to \max_{u_n},
\end{aligned}$$

Для кооперативного случая игры, когда игроки соглашаются действовать сообща для максимизации общего выигрыша, соответствующую оптимизационную задачу можно записать в виде

$$\sum_{i=1}^{n} K_i(t_0, x_0, u) \to \max_{u_1 \ldots u_n}$$

Модель оптимального управления объемом вредных выбросов

Экологическая задача оптимального управления объемом вредных выбросов представляет один из типов линейно-квадратичных задач оптимальной эксплуатации ресурса. Здесь в качестве игроков могут выступать страны, компании, предприятия и т.п. Ограничимся случаем, когда эксплуатируется только один ресурс, т.е. $x \in X \subseteq \mathbb{R}^1$. Под $x(t)$ будем понимать объем вредных выбросов, выделенных в атмосферу к моменту времени t. Интенсивность выброса вредных веществ каждым из игроков в момент времени t обозначим как $u_i(t)$.

Динамика накопления вредных выбросов в атмосфере (1) в этом случае задается следующим образом:

$$\dot{x} = \sum_{i=1}^{n} u_i - \alpha x, \quad x(t_0) = x^0, \quad x^0 \geqslant 0,$$

где α служит коэффициентом естественного рассеивания вредных веществ.

Моментальный выигрыш i-го игрока в данной задаче определяется как разница между мгновенным доходом от производства игрока i, вычисляемым как $h_i(u_i) = (A - \frac{1}{2}u_i)u_i$, и затратами на сокращение объема вредных выхлопов $p_i(x) = k_i x^l$, где показатель степени $l \in \{1; 2\}$ (его значение определяется для каждой конкретной задачи):

$$L_i(x, u) = (A - \frac{1}{2}u_i)u_i - p_i(x), \quad i = \overline{1, n}.$$

Следовательно, интегральный выигрыш (2) i-го игрока в игре определяется как:

$$K_i(t_0, x_0, u_i) = \int_{t_0}^{T} ((A - \frac{1}{2}u_i)u_i - p_i(x))dt, \quad i = \overline{1, n}.$$

Модель оптимального управления инвестициями в рекламу

Задача оптимального управления инвестициями в рекламную кампанию является еще одним представителем класса линейно-квадратичных задач эксплуатации ресурса. В качестве игроков в данном случае рассматриваются участники рынка, конкурирующие за его доли. Под $x_i(t)$ будем понимать деловую репутацию игрока i (гудвилл) в момент времени t. Объем средств, вложенных в рекламу i-м игроком за единицу времени, обозначим как $u_i(t)$.

Динамика процесса накопления гудвилла (1) игроком i задается следующим дифференциальным уравнением с начальными условиями:

$$\dot{x}_i = k_i u_i - \delta x_i, \quad x_i(0) = x_0^i, \quad x_0^i > 0, \quad i = \overline{1, n}.$$

В данной модели $k_i \geqslant 0$ - коэффициент, показывающий степень восприимчивости рынка к воздействию на него рекламы; $\delta \geqslant 0$ - коэффициент амортизации.

В данной задаче моментальный выигрыш i-го игрока определяется как разница между мгновенным доходом игрока i, вычисляемым как $s_i(x) = \pi(\beta - \sum_{k=1}^{n} x_k)x_i$, и затратами на рекламную кампанию $\frac{c}{2}u_i^2$:

$$L_i(x, u) = \pi(\beta - \sum_{k=1}^{n} x_k)x_i - \frac{c}{2}u_i^2, \quad i = \overline{1, n}.$$

Здесь π - прибыль игрока за единицу проданной продукции в единицу времени; β - требуемый (максимальный) объем продаж (мы предполагаем, что объем продаж пропорционален деловой репутации фирмы).

Игра рассматривается на бесконечном промежутке времени, т.е. $t \in [0, \infty)$. В качестве дисконтирующего множителя рассматривается функция $\Theta(t) = e^{-\rho t}, \quad$ где $\rho > 0$.

Таким образом, интегральный выигрыш (3) i-го игрока в игре примет вид:

$$K_i = \int_0^{\infty} e^{-\rho t}(\pi(\beta - \sum_{k=1}^{n} x_k)x_i - \frac{c}{2}u_i^2)dt \to \max_{u_i}, \quad i = \overline{1, n}.$$

Задача оптимального управления

Задача максимизации функционала вида (2), где подынтегральная функция является дифференцируемой, при динамике изменения фазовой переменной (1) является классической задачей оптимального управления с одним критерием.

Управление $u^*(t) = (u_1^*(t), \ldots, u_n^*(t)) \in U$, $t \in [t_0, T]$, являющееся решением данной задачи, т.е. доставляющее максимум функционалу (1), называют оптимальным управлением.

Принцип максимума Понтрягина

Для того чтобы управление $u^*(t)$ и соответствующее ему решение $x^*(t)$ были оптимальны в задаче (1), (2), для $i = \overline{1,n}$, необходимо, чтобы существовало n сопряженных вектор-функций $\lambda^i(t) = (\lambda_1^i(t), \ldots, \lambda_m^i(t)) : [t_0, T] \to \mathbb{R}^m$, таких, что выполнялись бы условия:

$$\begin{cases} H_i(x, u, \lambda^i) = L_i(t, x^*, u) + \langle \lambda^i(t), f(t, x^*, u) \rangle, \\ u_i^*(t) = \arg\max_{u_i} H_i(x^*, u, \lambda^i), \\ \dot{x}^*(t) = f(t, x^*, u^*), \quad x^*(t_0) = x_0 \\ \dot{\lambda}_j^i = -\frac{\partial H_i}{\partial x_j}, \qquad j = \overline{1,m} \\ \lambda^i(T) = \frac{\partial}{\partial x^*} h_i(x^*(T)). \end{cases} \tag{4}$$

Заметим, что система (4) соответствует решению оптимизационной задачи для i-го игрока, $i = \overline{1,n}$. В случае некооперативной постановки игры необходимо одновременно решить n задач вида (4).

В случае, если рассматривается игра с бесконечной продолжительностью и с функцией дисконтирования $\Theta(t) = e^{-\rho t}$, $\rho > 0$, получаем задачу максимизации функционала (3) при ограничениях (1). Тогда формулировка принципа максимума Понтрягина примет следующий вид:

Для того чтобы управление $u^*(t)$ и соответствующее ему решение $x^*(t)$ были оптимальны в задаче (1), (3), для $i = \overline{1,n}$, необходимо, чтобы существовало n сопряженных вектор-функций $\lambda^i(t) = (\lambda_1^i(t), \ldots, \lambda_m^i(t)) : [t_0, T] \to \mathbb{R}^m$, $i = \overline{1,n}$, таких, что выполнялись бы условия:

$$\begin{cases} H_i(x, u, \lambda^i) = L_i(t, x^*, u) + \langle \lambda^i(t), f(t, x^*, u) \rangle, \\ u_i^*(t) = \arg\max_{u_i} H_i(x^*, u, \lambda^i) \\ x^*(t) = f(t, x^*, u^*), \qquad x^*(t_0) = x_0 \\ \dot{\lambda}_j^i = -\frac{\partial H_i}{\partial x_j} + \rho \lambda_j^i(t), \qquad j = \overline{1,m} \\ \lim\limits_{t \to +\infty} e^{-\rho t} H_i(x^*(t), u^*(t), \lambda^i(t)) = 0. \end{cases}$$

Уравнение Гамильтона—Якоби—Беллмана

Управление $u^*(x,t)$ и соответствующее ему решение $x^*(t)$ оптимальны в задаче (1), (2), для $i = \overline{1,n}$, если существует функция $V_i(t,x)$: $[t_0, T] \times \mathbb{R}^m \to \mathbb{R}$ - непрерывно-дифференцируемая по своим аргументам и удовлетворяющая уравнению Гамильтона – Якоби – Беллмана с краевым условием:

$$-\frac{\partial V_i(t,x)}{\partial t} = \max_u \left[\frac{\partial V_i(t,x)}{\partial x} f(t,x,u) + L_i(t,x,u) \right]$$

$$V_i(T, x(T)) = 0$$

Упрощение интегрального функционала кооперативной линейно-квадратичной дифференциальной игры

Вновь обратимся к линейно-квадратичной дифференциальной игре в кооперативной постановке в общем виде:

$$\dot{x} = f(t,x,u), \quad x(t_0) = x_0, \tag{5}$$

$$\int_{t_0}^{T} \Big(r^T u + \frac{1}{2} u^T R u + q^T x + \frac{1}{2} x^T Q x \Big) dt \to \max_u \tag{6}$$

Здесь R, Q - симметричные матрицы соответствующих размерностей; r,q - вектор-столбцы.

Функция моментального выигрыша, имеющая линейно-квадратичную структуру, может быть упрощена с помощью аффинного преобразования следующего вида: $u = Tv - \tau$, где T - невырожденная симметричная матрица, τ-вектор-столбец.

Тогда получим:

$$r^T u + \frac{1}{2} u^T R u = \frac{1}{2} v^T T^T R T v + (r - \frac{1}{2} \tau^T R) T v - \frac{1}{2} (\tau^T R)^T (TV)^T + \frac{1}{2} \tau^T R \tau - r^T \tau =$$

$$= \frac{1}{2} v^T T^T R T v + (r^T - \tau^T R) T v + (\frac{1}{2} \tau^T R - r^T) \tau$$

Если $r^T - \tau^T R = 0$, то линейное слагаемое сократится.

Аналогичным образом с помощью замены переменных $x = My - \mu$ может быть проведено преобразование выражения $q^T x + \frac{1}{2} x^T Q x$.

В результате дифференциальная игра (5),(6) оказывается сведена к следующему виду:

$$\dot{y}(t) = \tilde{f}(t, y, v), \quad y(t_0) = M^{-1}(x^0 + \mu),$$

$$\int_{t_0}^{T} \Big(\frac{1}{2}v^T T^T R T v + \frac{1}{2}y^T M^T Q M y - \frac{1}{2}r^T \tau - \frac{1}{2}q^T \mu\Big) dt \to \max_{u_i}$$

Потенциальные игры

Рассматривается некооперативная дифференциальная игра в общем виде на бесконечном промежутке времени:

$$\dot{x} = f(t, x, u), \quad x(t_0) = x_0, \tag{7}$$

$$\int_{t_0}^{\infty} e^{-\rho t} L_i(t, x(t), u(t)) dt \to \max_{u_i}, \quad i = \overline{1, n} \tag{8}$$

где $x \in X$, $u \in U = U_1 \times \ldots \times U_n \subseteq \mathbb{R}^n$, $t \in T = [t_0, \infty)$

Будем использовать стандартные обозначения:

$$u_{-i} = (u_1, \ldots, u_{i-1}, u_{i+1}, \ldots, u_n)$$
$$x_{-i} = (x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)$$

Дифференциальная игра (7),(8) называется потенциальной динамической дифференциальной игрой, если существует функция $P(t, x, u)$: $T \times X \times U \to \mathbb{R}$ такая, что для всех $u_i, v_i \in U_i$ и для любого $i = \overline{1, n}$ выполняется следующее условие:

$$\int_{t_0}^{\infty} e^{-\rho t}\Big(L_i(t, x, u_i, u_{-i}) - L_i(t, x, v_i, u_{-i})\Big) dt =$$
$$= \int_{t_0}^{\infty} e^{-\rho t}\Big(P(t, x, u_i, u_{-i}) - P(t, x, v_i, u_{-i})\Big) dt.$$

Заметим, что в данном случае для каждого $i = \overline{1, n}$ мы варьируем только i-ю компоненту вектора $u(t)$, для чего и используется введенное выше обозначение $u_{-i}(t)$.

Функция $P(t, x, u)$ называется потенциалом игры (7),(8).

Достаточные условия существования потенциала в дифференциальной игре

Рассмотрим достаточные условия существования потенциала дифференциальной игры, которые были приведены в работе [14].

Пусть рассматривается дифференциальная игра (7),(8), существует функция $p(t,x,u) : T \times X \times U \to \mathbb{R}$ и пусть для каждого $i = \overline{1,n}$ выполняется одно из следующих условий:

1. Существует функция $c^i : T \times U_{-i} \to \mathbb{R}$ такая, что выполняется:

$$L_i(t,x(t),u(t)) = p(t,x,u) + c^i(t,u_{-i})$$

2. Существуют функции $c^i : T \times X \times U_{-i} \to \mathbb{R}$ и $g^i : T \times X \to X_i$ такие, что выполняется:

$$L_i(t,x(t),u(t)) = p(t,x,u) + c^i(t,x,u_{-i})$$

$$f(t,x(t),u(t)) = g^i(t,x)$$

3. Существуют функции $c^i : T \times X_{-i} \times U_{-i} \to \mathbb{R}$ и $g^i : T \times X_i \times U_i \to X_i$ такие, что выполняется:

$$L_i(t,x(t),u(t)) = p(t,x,u) + c^i(t,x_{-i},u_{-i})$$

$$f(t,x(t),u(t)) = g^i(t,x_i,u_i)$$

Тогда дифференциальная игра (7),(8) является потенциальной дифференциальной игрой с потенциалом $P(\cdot) = p(\cdot)$.

Альтернативная формулировка достаточных условий существования потенциала в дифференциальной игре

В работе [14] также была приведена альтернативная формулировка достаточных условий существования потенциала дифференциальной игры. Рассмотрим ее.

Пусть рассматривается дифференциальная игра (7),(8), существует функция $p(t,x,u) : T \times X \times U \to \mathbb{R}$ и пусть для каждого $i = \overline{1,n}$ выполняется одно из следующих условий:

1. Существует функция $c^i : T \times U_i \to \mathbb{R}$ такая, что выполняется:

$$L_i(t,x(t),u(t)) = p(t,x,u) + c^i(t,u_i)$$

2. Существуют функции $c^i : T \times X \times U_i \to \mathbb{R}$ и $g^i : T \times X \to X_i$ такие, что выполняется:

$$L_i(t,x(t),u(t)) = p(t,x,u) + c^i(t,x,u_i)$$

$$f(t,x(t),u(t)) = g^i(t,x)$$

3. Существуют функции $c^i : T \times X_i \times U_i \to \mathbb{R}$ и $g^i : T \times X_i \times U_i \to X_i$

такие, что выполняется:

$$L_i(t, x(t), u(t)) = p(t, x, u) + c^i(t, x_i, u_i)$$

$$f(t, x(t), u(t)) = g^i(t, x_i, u_i)$$

Тогда дифференциальная игра (7),(8) является потенциальной дифференциальной игрой с потенциалом $P(\cdot) = p(\cdot) + \sum_{j=1}^{n} c^j(\cdot)$.

Глава 2. Об упрощении интегрального функционала кооперативной линейно-квадратичной игры на примере задачи оптимального управления объемом вредных выбросов

Постановка задачи

Рассмотрим дифференциальную игру [6] оптимального управления объемом вредных выбросов с линейно-квадратичной функцией выигрыша и с числом игроков $n = 3$. В данном случае в качестве игроков могут выступать страны, компании, предприятия и т.п. Динамика игры имеет линейный вид:

$$\dot{x} = u_1 + u_2 + u_3 - \alpha x \tag{9}$$

В начальный момент времени $x(t_0) = x_0, \quad x_0 > 0$. Пусть $t \in [t_0, T]$. Мгновенный выигрыш i-го игрока в момент времени t имеет вид:

$$L_i(u_i(t)) = u_i(A - \frac{u_i}{2}) - \frac{kx^2}{2}, \quad i = \overline{1,3}$$

Терминальный выигрыш i-го игрока определяется как:

$$H_i(x(T)) = d_i x(T), \quad i = \overline{1,3}$$

Тогда для ожидаемого выигрыша игроков в кооперативной игре можно записать:

$$K = \sum_{i=1}^{3} K_i(x_0, T - t_0, u_i) = \sum_{i=1}^{3} \int_{t_0}^{T} (u_i(A - \frac{u_i}{2}) - \frac{kx^2}{2})dt + \sum_{i=1}^{3} d_i x(T) \to \max \tag{10}$$

Упрощение функции выигрыша в линейно-квадратичной дифференциальной игре

Рассматриваем случай кооперативной игры, в которой игроки действуют в соответствии с некоторым заранее определенным принципом, позволяющим максимизировать собственную выгоду. Функцию выигрыша игроков (10) можно представить в матричной форме следующим образом:

$$K = \int_{t_0}^{T} (r^T u + \frac{1}{2} u^T R u - \frac{1}{2} x^T Q x)dt + d^T x(T) \to \max,$$

где $R = diag(-1, ..., -1)_{[n\times n]}$, $Q = diag(k, ..., k)_{[n\times n]}$, $r = A \cdot 1_{[n\times 1]}$, $d = \begin{pmatrix} d_1 & d_2 & d_3 \end{pmatrix}^T$. Здесь $1_{[n\times 1]}$ - вектор-столбец размерности $[n \times 1]$, составленный из единиц.

Согласно [8], функция выигрыша линейно-квадратичной дифференциальной игры может быть упрощена с помощью аффинного преобразования следующего вида: $u = Tv - \tau$, где T - невырожденная симметричная матрица, τ-вектор-столбец.
Тогда аналогично [9] получим:

$$r^T u + \frac{1}{2}u^T R u = \frac{1}{2}v^T T^T R T v + (r - \frac{1}{2}\tau^T R)Tv - \frac{1}{2}(\tau^T R)^T \cdot (TV)^T + \frac{1}{2}\tau^T R \tau - r^T \tau =$$

$$= \frac{1}{2}v^T T^T R T v + (r^T - \tau^T R)Tv + (\frac{1}{2}\tau^T R - r^T)\tau$$

Если $r^T - \tau^T R = 0$, то линейное слагаемое сократится. Для выполнения этого условия необходимо, чтобы $\tau_i = -A$. Получаем окончательный вид подынтегральной функции после преобразования $u = Tv + A \cdot 1_{[n\times 1]}$:

$$r^T u + \frac{1}{2}u^T R u - \frac{1}{2}x^T Q x = \frac{1}{2}v^T T^T R T v + \frac{1}{2}A r^T \cdot 1_{[n\times 1]} - \frac{1}{2}x^T Q x$$

Матрица R симметрическая, следовательно, существует ортогональное преобразование T такое, что матрица $\tilde{R} = T^T R T$ - диагональная.

Если взять, например, матрицу $T = \begin{pmatrix} \sqrt{2} & 2 & 0 \\ \sqrt{2} & -1 & \sqrt{3} \\ \sqrt{2} & -1 & -\sqrt{3} \end{pmatrix}$, тогда
$\tilde{R} = \begin{pmatrix} -6 & 0 & 0 \\ 0 & -6 & 0 \\ 0 & 0 & -6 \end{pmatrix}$.

Таким образом, свели исходную задачу к задаче, записанной в скалярной форме, вида:

$$\sum_{i=1}^{3} K_i = \sum_{i=1}^{3} \int_{t_0}^{T} (-\frac{kx^2}{2} - 3v_i^2 + \frac{3}{2}A^2)dt + \sum_{i=1}^{3} d_i x(T) \to \max. \tag{11}$$

$$\dot{x} = 3\sqrt{2}v_1 - \alpha x + 3A. \tag{12}$$

Решение кооперативной игры методом максимума Понтрягина

Найдем решение кооперативной игры (9),(10) методом максимума Понтрягина [10], используя результат (11),(12) упрощения максимизируемого функционала линейно-квадратичной дифференциальной игры.

Для использования метода максимума Понтрягина потребуем, чтобы множество допустимых управлений $u_i, i = 1, 2, 3$ было компактом. Иными словами, необходимо, чтобы найденные далее оптимальные управления $u_i^*(t) \in [0, A]$. Забегая вперед, стоит отметить, что при определенных значениях параметров модели данное требование будет выполняться при $\forall t \geq 0$. Например, этому удовлетворяют следующие значения параметров: $A = 34; k = 28; \alpha = 15; x_0 = 10; t_0 = 0; T = 10; d_i = -1, i = 1, 2, 3$.

Запишем гамильтониан системы:

$$H(x, v, \psi) = \sum_{i=1}^{3}(-\frac{kx^2}{2} - 3v_i^2 + \frac{3}{2}A^2) + \psi(3\sqrt{2}v_1 - \alpha x + 3A), \quad i = \overline{1,3}$$

Найдём максимум функции H по v:

$$\frac{\partial H}{\partial v_1} = -6v_1 + 3\sqrt{2}\psi = 0$$

$$\frac{\partial H}{\partial v_i} = -6v_i = 0, i = 2, 3$$

$$\frac{\partial^2 H}{\partial v^2} = -6E_3 < 0,$$

где E_n – единичная матрица размерности $[n \times n]$.

Получаем, что функция $H(x, v, \psi)$ достигает максимума при $v_1 = \frac{\sqrt{2}}{2}\psi$, $v_2 = v_3 = 0$.

Обозначим $D = d_1 + d_2 + d_3$. Учитывая начальные условия задачи и условие трансверсальности $\psi(T) = -D$, получаем систему уравнений:

$$\begin{cases} \frac{dx}{dt} = 3\psi - \alpha x + 3A \\ \frac{d\psi}{dt} = 3kx + \alpha\psi \\ x(t_0) = x_0 \\ \psi(T) = -D \end{cases}$$

Получаем решение:

$$\begin{cases} \psi(t) = -G_1e^{-tF} - G_2e^{tF} - \frac{9Ak}{F^2} \\ x(t) = \frac{F+\alpha}{3k}G_1e^{-tF} - \frac{F-\alpha}{3k}G_2e^{tF} + \frac{9Ak\alpha}{3F^2k} \end{cases},$$

где

$$F = \sqrt{\alpha^2 + 9k}$$

$$G_1 = \frac{27k^2x_0e^{TF} - \alpha^3De^{t_0F} - 9\alpha Dke^{t_0F} + \alpha^2DFe^{t_0F} + 3\alpha^2kx_0e^{TF}}{F^2(Fe^{TF-t_0F} + Fe^{-TF+t_0F} + \alpha e^{TF-t_0F} - \alpha e^{-TF+t_0F})} + \\ + \frac{-9A\alpha ke^{TF} - 9AkFe^{t_0F} + 9A\alpha ke^{t_0F} + 9DkFe^{t_0F}}{F^2(Fe^{TF-t_0F} + Fe^{-TF+t_0F} + \alpha e^{TF-t_0F} - \alpha e^{-TF+t_0F})}$$

$$G_2 = \frac{27k^2x_0e^{-TF} + \alpha^3De^{-t_0F} + 9\alpha Dke^{-t_0F} + \alpha^2DFe^{-t_0F} - 3\alpha^2kx_0e^{-TF}}{F^2(Fe^{TF-t_0F} + Fe^{-TF+t_0F} + \alpha e^{TF-t_0F} - \alpha e^{-TF+t_0F})} + \\ + \frac{9A\alpha ke^{-TF} - 9AkFe^{-t_0F} - 9A\alpha ke^{-t_0F} + 9DkFe^{-t_0F}}{F^2(Fe^{TF-t_0F} + Fe^{-TF+t_0F} + \alpha e^{TF-t_0F} - \alpha e^{-TF+t_0F})}$$

Имеем решение задачи (11),(12):

$$v^*(t) = \begin{pmatrix} -\frac{\sqrt{2}}{2}(G_1e^{-tF} + G_2e^{tF} + \frac{9Ak}{F^2}) \\ 0 \\ 0 \end{pmatrix}$$

$$x^*(t) = \frac{F+\alpha}{3k}G_1e^{-tF} - \frac{F-\alpha}{3k}G_2e^{tF} + \frac{9Ak\alpha}{3kF^2}$$

Тогда в исходной игре (9),(10) получаем, что оптимальные управления и оптимальная траектория имеют вид:

$$u^*(t) = \begin{pmatrix} -G_1e^{-tF} - G_2e^{tF} - \frac{9Ak}{F^2} + A \\ -G_1e^{-tF} - G_2e^{tF} - \frac{9Ak}{F^2} + A \\ -G_1e^{-tF} - G_2e^{tF} - \frac{9Ak}{F^2} + A \end{pmatrix}$$

$$x^*(t) = \frac{F+\alpha}{3k}G_1e^{-tF} - \frac{F-\alpha}{3k}G_2e^{tF} + \frac{9Ak\alpha}{3kF^2}$$

Численный пример

Пусть коэффициенты в модели (9),(10) принимают следующие значения, полученные экспериментальным путем по данным некоторых промышленных предприятий Восточной Сибири [11]:

$$A = 59035.12$$

$$d_1 = 525.06; d_2 = 351.64; d_3 = 112.71$$

$$k = 1$$

Пусть $t_0 = 0$, $T = 0.4$ и пусть $\alpha = 0.2$ (данный коэффициент естественного рассеивания вредных веществ в атмосфере соответствует нормальным погодным условиям).

Тогда для кооперативного случая получим следующие оптимальные управления при указанных значениях параметров:

$$u^*(t) = \begin{pmatrix} 15781.29 \cdot e^{3.006 \cdot t} + 17473.79 \cdot e^{-3.006 \cdot t} + 261.217 \\ 15781.29 \cdot e^{3.006 \cdot t} + 17473.79 \cdot e^{-3.006 \cdot t} + 261.217 \\ 15781.29 \cdot e^{3.006 \cdot t} + 17473.79 \cdot e^{-3.006 \cdot t} + 261.217 \end{pmatrix}$$

Соответствующая оптимальная траектория примет вид:

$$x^*(t) = 14764.23 \cdot e^{3.006 \cdot t} - 18677.49 \cdot e^{-3.066 \cdot t} + 3918.26$$

График оптимального управления игрока i при указанных значениях параметров модели (Рис.1)

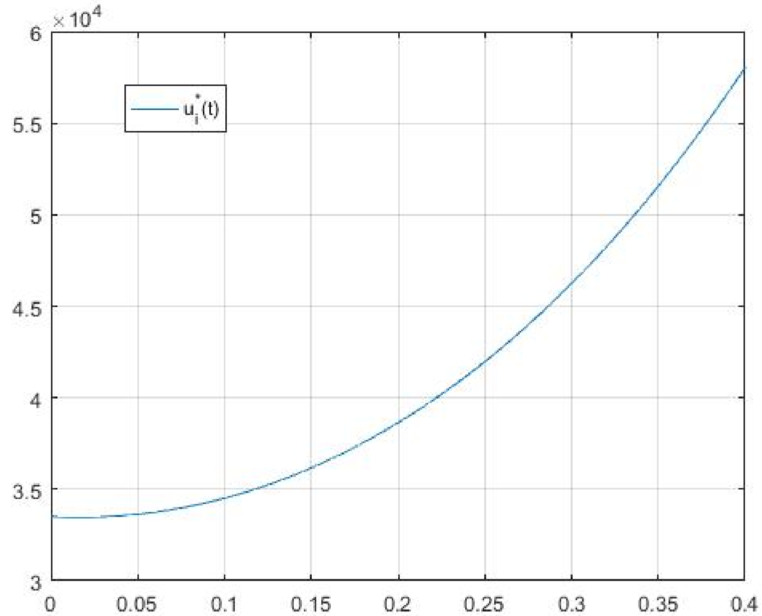

Рис. 1: Оптимальное управление $u_i^*(t)$

График соответствующей оптимальной траектории игрока i при указанных значениях параметров модели (Рис.2)

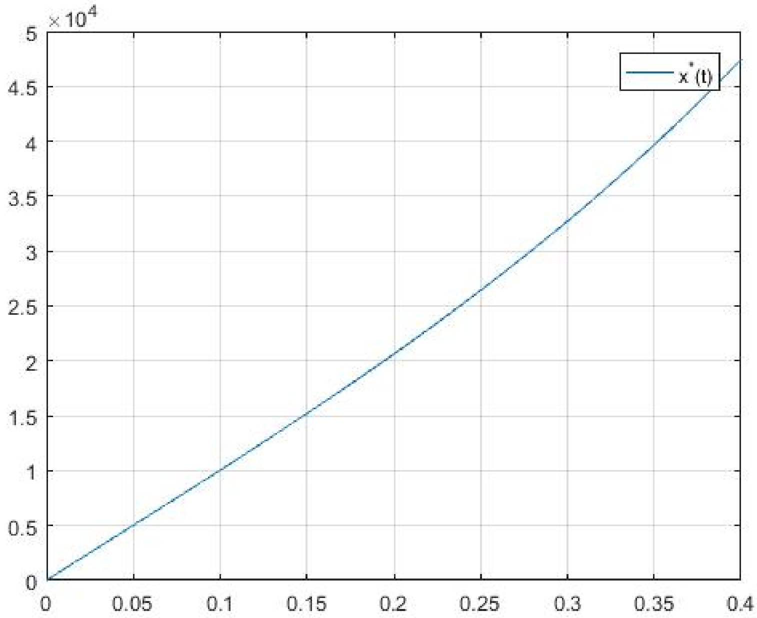

Рис. 2: Оптимальная траектория $x_i^*(t)$

Решение кооперативной игры с помощью уравнения Гамильтона—Якоби—Беллмана

Найдем теперь решение той же кооперативной игры (9),(10) с помощью уравнения Гамильтона—Якоби—Беллмана [12, 13], используя результат (11),(12) упрощения выигрыша линейно-квадратичной дифференциальной игры.

Данная задача эквивалентна задаче:

$$\sum_{i=1}^{3} K_i = \sum_{i=1}^{3} \int_{t_0}^{T} (-\frac{kx^2}{2} - 3v_i^2)dt + \sum_{i=1}^{3} d_i x(T) \to \max .$$

$$\dot{x} = 3\sqrt{2}v_1 - \alpha x + 3A$$

Запишем уравнение Гамильтона—Якоби—Беллмана для этой задачи:

$$-\frac{\partial W}{\partial t} = \max_{v_i} \Big(\sum_{i=1}^{3} (-\frac{kx^2}{2} - 3v_i^2) + \frac{\partial W}{\partial x}(3\sqrt{2}v_1 + 3A - \alpha x) \Big) \tag{13}$$

$$W(x, T) = \sum_{i=1}^{3} d_i x(T)$$

Максимизирующее управление имеет вид: $v_1 = \frac{\sqrt{2}}{2} W_x$, $v_2 = 0$, $v_3 = 0$
При подстановке v_i в уравнение (13) получаем:

$$-W_t = \sum_{i=1}^{3} (-\frac{kx^2}{2}) - \frac{3\sqrt{2}}{2} W_x + W_x(3W_x + 3A - \alpha x) \tag{14}$$

Вид искомого решения определим как:

$$W(t,x) = L(t)x^2 + M(t)x + N(t)$$

Следовательно, имеем $W_x = 2L(t)x + M(t), W_t = \dot{L}(t)x^2 + \dot{M}(t)x + \dot{N}(t)$. Подставляя W_t, W_x в уравнение (14), получим систему уравнений для определения функций $L(t), M(t), N(t)$:

$$\begin{cases} \dot{L}(t) + 12L^2(t) - 2\alpha L(t) - \frac{3k}{2} = 0 \\ \dot{M}(t) + 12L(t)M(t) + 6AL(t) - 3\sqrt{2}L(t) - \alpha M(t) = 0 \\ \dot{N}(t) + 3M^2(t) - 3AM(t) - \frac{3\sqrt{2}}{2} M(t) = 0 \\ L(T) = 0 \\ M(T) = d_1 + d_2 + d_3 \\ N(T) = 0 \end{cases}$$

Получаем:

$$L(t) = \frac{\alpha}{12} - \frac{S}{12} \tan\left(2S(\frac{t}{2} + \frac{1}{2S} \arctan\frac{\alpha}{S} - \frac{1}{2})\right)$$

$$M(t) = \frac{2S^2D - 9k(2A - \sqrt{2})(1 - \cos(S(T-t)))}{2S(\alpha \sin(S(T-t)) + S\cos(S(T-t))},$$

где $S = \sqrt{-\alpha^2 - 18k}$, $D = d_1 + d_2 + d_3$.
$N(t)$ определяется из соответствующего дифференциального уравнения $\dot{N}(t) + 3M^2(t) - 3AM(t) - \frac{3\sqrt{2}}{2} M(t) = 0$ и начального условия $N(T) = 0$.

Таким образом, получаем максимизирующее управление

$$v(t) = \begin{pmatrix} \frac{\alpha}{12} - \frac{S}{12} \tan\left(2S(\frac{t}{2} + \frac{1}{2S} \arctan\frac{\alpha}{S} - \frac{1}{2})\right)x + \frac{2S^2D - 9k(2A-\sqrt{2})(1-\cos(S(T-t)))}{2S(\alpha\sin(S(T-t)) + S\cos(S(T-t))}) \\ 0 \\ 0 \end{pmatrix}$$

С помощью полученного управления $v(t)$ и дифференциального уравнения динамики игры (12) можно определить оптимальную траекторию $x^*(t)$ и оптимальное управление $v^*(t)$ игры (11),(12). Затем, учитывая преобразование $u = Tv + A \cdot 1_{[n\times 1]}$ и вид выбранной матрицы T, можно получить оптимальное управление исходной игры (9),(10):

$$u(t) = \begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix},$$

где

$$u_i = \frac{\sqrt{2}\alpha}{4} - \frac{\sqrt{2}S}{4} \tan\left(2S(\frac{t}{2} + \frac{1}{2S}\arctan\frac{\alpha}{S} - \frac{1}{2})\right)x + \\ + \frac{6S^2D - 27k(2A - \sqrt{2})(1 - \cos(S(T-t)))}{\sqrt{2}S(\alpha\sin(S(T-t)) + S\cos(S(T-t))} + A, \quad i = \overline{1,3}.$$

Глава 3. О решении линейно-квадратичной дифференциальной игры методом потенциала на примере задачи оптимального управления инвестициями в рекламу

Постановка задачи

Рассмотрим некооперативную дифференциальную игру $\Gamma_{G_0}^{\infty}$ из [4] с линейно-квадратичной функцией выигрыша и с числом игроков n. Динамика игры имеет линейный вид:

$$\dot{G}_i = a_i - \delta G_i, \quad i = \overline{1, n} \tag{15}$$

Начальные условия задачи определяются как $G_i(0) = G_0^i$, $G_0^i > 0$. Пусть $t \in [0, \infty)$.
В данной модели $a_i \geq 0$ - управление игрока i;
G_i - значение деловой репутации (Гудвилл) игрока i;
$\delta \geq 0$ - коэффициент амортизации.

Выигрыш игрока i в игре задается следующим образом:

$$J_i = \int\limits_0^{\infty} e^{-\rho t}(\pi(\beta - \sum_{h=1}^{n} G_h)G_i - \frac{c}{2}a_i^2)dt \to \max_{u_i}, \quad i = \overline{1, n}. \tag{16}$$

Потенциал игры

Согласно п.3) альтернативной формулировки достаточных условий [14] существования потенциала в дифференциальной игре, данная дифференциальная игра $\Gamma_{x_0}^{\infty}$ является потенциальной со следующим потенциалом:

$$P(t, G, a) = -\pi G_i \sum_{h=1}^{i-1} G_h - \pi G_i \sum_{h=i+1}^{n} G_h + \sum_{j=1}^{n}(\pi\beta G_j - \pi G_j^2 - \frac{c}{2}a_j^2). \tag{17}$$

Потенциал $P(t, G, a)$ для линейно-квадратичной игры представляет собой квадратичную форму [18].

Действительно, в принятых условных обозначениях можем записать:

$$L_i(t, G, a) = \pi(\beta - \sum_{h=1}^{n} G_h)G_i - \frac{c}{2}a_i^2 =$$

$$= \pi\beta G_i - \pi G_i^2 - \pi G_i \sum_{h=1}^{i-1} G_h - \pi G_i \sum_{h=i+1}^{n} G_h - \frac{c}{2}a_i^2.$$

Соответственно, имеем

$$f(t,G,a) = a_i - \delta G_i = g^i(t,G_i,a_i),$$

$$c^i(t,G_i,a_i) = \pi\beta G_i - \pi G_i^2 - \frac{c}{2}a_i^2$$

и

$$p(t,G,a) = -\pi G_i \sum_{h=1}^{i-1} G_h - \pi G_i \sum_{h=i+1}^{n} G_h.$$

Тогда получаем: $L^i(t,G,a) = p(t,G,a) + c^i(t,G_i,a_i)$, т.е. действительно существуют такие функции $c^i(t,G_i,a_i), g^i(t,G_i,a_i), p(t,G,a)$, что достаточные условия существования потенциала выполнены, и, значит, потенциал рассматриваемой игры имеет вид (17).

Применение метода потенциала в решении дифференциальной игры

Таким образом, получили, что рассматриваемая некооперативная игра является потенциальной, поэтому для нахождения равновесия по Нэшу мы можем рассмотреть соответствующую задачу оптимизации:

$$\dot{G}_i = g^i(t,G_i,a_i)$$

$$J = \int_0^\infty e^{-\rho t} P(t,G,a)dt \to \max_a$$

Итак, исходная система оптимизационных задач 16 оказывается сведена к одной задаче оптимизации, что облегчает дальнейшие вычисления.

Подставим соответствующие функции, получаем следующую задачу:

$$\dot{G}_i = a_i - \delta G_i, \quad i = \overline{1,n}, \tag{18}$$

$$\int_0^\infty e^{-\rho t}(-\pi G_i \sum_{h=1}^{i-1} G_h - \pi G_i \sum_{h=i+1}^{n} G_h + \sum_{j=1}^{n}(\pi\beta G_j - \pi G_j^2 - \frac{c}{2}a_j^2))dt \to \max_a. \tag{19}$$

Для решения задачи (18),(19) воспользуемся принципом максимума Понтрягина.

Запишем гамильтониан системы:

$$H_i(G,a,\psi) = -\pi G_i \sum_{h=1}^{i-1} G_h - \pi G_i \sum_{h=i+1}^{n} G_h + \sum_{j=1}^{n}(\pi\beta G_j - \pi G_j^2 - \frac{c}{2}a_j^2) + \psi_i(t)(a_i - \delta G_i)$$

Найдём точку максимума функции H_i, используя условия Куна—Таккера [20]. Введем в рассмотрение функционал:

$$L_i(a_i) = -H_i + \lambda(-a_i)$$

.

Для того чтобы искомое управление a_i^N было оптимально, необходимо выполнение условий:

1. Стационарность: $\min_{a_i} L_i(a_i) = L_i(a_i^N)$
2. Дополняющая нежесткость: $\lambda(-a_i^N) = 0$
3. Неотрицательность: $\lambda \geq 0$

Согласно п.1, получаем:

$$\frac{\partial L_i}{\partial a_i} = -\frac{\partial H_i}{\partial a_i} - \lambda = -ca_i + \psi_i(t) - \lambda = 0 \tag{20}$$

Рассмотрим возможные случаи:

1. $\lambda \neq 0$. Из условия неотрицательности в этом случае следует, что $\lambda > 0$. Тогда для выполнения условия дополняющей нежесткости должно выполняться: $a_i^N = 0$. Следовательно, уравнение (20) запишется как $\lambda(t) = -\psi_i(t)$. Далее в работе будет показано, что $\psi_i(t) \geq 0$ при условии, что параметры модели удовлетворяют выражению: $g_0 \leq E$, где коэффициент E также определен далее. Итак, получаем, что если $g_0 \leq E$, то $\lambda < 0$, что противоречит условию неотрицательности. Если же $g_0 > E$, то условие неотрицательности оказывается выполненным и в этом случае имеем $a_i^N = 0$.

2. $\lambda = 0$, $g_0 \leq E$. Выполнение условия дополняющей нежесткости обеспечено. Уравнение (20) примет вид $ca_i + \psi_i(t) = 0$. Получаем, что функция $H_i(G, a, \psi)$ достигает максимума при $a_i^N = \frac{1}{c}\psi_i(t)$.

Условие трансверсальности для задач, рассматриваемых на бесконечном промежутке времени, согласно [19], имеет вид:

$$\lim_{t \to +\infty} e^{-\rho t} H_i(G^N(t), a^N(t), \psi(t)) = 0, \quad i = \overline{1, n}$$

Учитывая начальные условия задачи и указанные условия трансерсальности, а также тот факт, что игра рассматривается на бесконечном промежутке времени, получаем систему уравнений:

$$\begin{cases} \frac{dG_i}{dt} = \frac{1}{c}\psi_i(t) - \delta G_i(t), \quad i = \overline{1,n} \\ \frac{d\psi_i}{dt} = -\frac{\partial H_i}{\partial G_i} + \rho\psi_i(t), \quad i = \overline{1,n} \\ G_i(0) = G_0^i, \quad i = \overline{1,n} \\ \lim\limits_{t\to+\infty} e^{-\rho t} H_i(G^N(t), a^N(t), \psi^N(t)) = 0, \quad i = \overline{1,n} \end{cases}$$

Рассмотрим случай для количества игроков $n = 2$. Тогда систему уравнений относительно переменных состояния и сопряженных переменных можно переписать следующим образом:

$$\begin{cases} \frac{dG_1}{dt} = \frac{1}{c}\psi_1(t) - \delta G_1(t) \\ \frac{dG_2}{dt} = \frac{1}{c}\psi_2(t) - \delta G_2(t) \\ \frac{d\psi_1}{dt} = \pi G_2(t) - \pi\beta + 2\pi G_1(t) + \delta\psi_1(t) + \rho\psi_1(t) \\ \frac{d\psi_2}{dt} = \pi G_1(t) - \pi\beta + 2\pi G_2(t) + \delta\psi_2(t) + \rho\psi_2(t) \\ G_1(0) = G_0^1 \\ G_2(0) = G_0^2 \\ \lim\limits_{t\to+\infty} e^{-\rho t} H_1(G^N(t), a^N(t), \psi^N(t)) = 0 \\ \lim\limits_{t\to+\infty} e^{-\rho t} H_2(G^N(t), a^N(t), \psi^N(t)) = 0 \end{cases}$$

Далее будем использовать следующие обозначения:

$$D = \sqrt{c(4c\delta^2 + 4c\delta\rho + c\rho^2 + 12\pi)}$$
$$S = \sqrt{c(4c\delta^2 + 4c\delta\rho + c\rho^2 + 4\pi)}$$
$$Q_1 = \frac{2c\delta - S + c\rho}{2c\pi}$$
$$Q_2 = \frac{2c\delta + S + c\rho}{2c\pi}$$
$$F_1 = \frac{2c\delta - D + c\rho}{6c\pi}$$
$$F_2 = \frac{2c\delta + D + c\rho}{6c\pi}$$
$$E = \frac{2\beta c\pi - \frac{4\beta c^2\delta\pi}{D - c\rho}}{2D} \cdot \frac{2c\delta + D + c\rho}{6\pi c} - \frac{2\beta c\pi + \frac{4\beta c^2\delta\pi}{D + c\rho}}{2D} \cdot \frac{2c\delta + S + c\rho}{6\pi c}$$
$$Z = \frac{2\beta c^2\delta\pi}{(D + c\rho)D} + \frac{2\beta c^2\delta\pi}{(D - c\rho)D}$$

Тогда общее решение рассматриваемой системы имеет вид:

$$\begin{cases} \psi_1^N(t) = -C_1 e^{\frac{S+c\rho}{2c}t} - C_2 e^{-\frac{S-c\rho}{2c}t} + C_3 e^{\frac{D+c\rho}{2c}t} + C_4 e^{-\frac{D-c\rho}{2c}t} + Z \\ \psi_2^N(t) = C_1 e^{\frac{S+c\rho}{2c}t} + C_2 e^{-\frac{S-c\rho}{2c}t} + C_3 e^{\frac{D+c\rho}{2c}t} + C_4 e^{-\frac{D-c\rho}{2c}t} + Z \\ G_1^N(t) = Q_1 C_1 e^{\frac{S+c\rho}{2c}t} + Q_2 C_2 e^{-\frac{S-c\rho}{2c}t} - F_1 C_3 e^{\frac{D+c\rho}{2c}t} - F_2 C_4 e^{-\frac{D-c\rho}{2c}t} + E \\ G_2^N(t) = -Q_1 C_1 e^{\frac{S+c\rho}{2c}t} - Q_2 C_2 e^{-\frac{S-c\rho}{2c}t} - F_1 C_3 e^{\frac{D+c\rho}{2c}t} - F_2 C_4 e^{-\frac{D-c\rho}{2c}t} + E \end{cases}$$

Частное решение может быть найдено с помощью системы уравнений:

$$\begin{cases} G_1(0) = Q_1C_1 + Q_2C_2 - F_1C_3 - F_2C_4 + E = G_0^1 \\ G_2(0) = -Q_1C_1 - Q_2C_2 - F_1C_3 - F_2C_4 + E = G_0^2 \\ \lim\limits_{t\to+\infty} e^{-\rho t}(-\pi G_1^N G_2^N + \pi\beta G_1^N - \pi(G_1^N)^2 - \frac{1}{2c}(\psi_1^N)^2 + \pi\beta G_2^N - \pi(G_2^N)^2 - \\ -\frac{1}{2c}(\psi_2^N)^2 + \psi_1^N(t)(\frac{1}{c}\psi_1^N - \delta G_1^N(t))) = 0 \\ \lim\limits_{t\to+\infty} e^{-\rho t}(-\pi G_1^N G_2^N + \pi\beta G_1^N - \pi(G_1^N)^2 - \frac{1}{2c}(\psi_1^N)^2 + \pi\beta G_2^N - \pi(G_2^N)^2 - \\ -\frac{1}{2c}(\psi_2^N)^2 + \psi_2^N(t)(\frac{1}{c}\psi_2^N - \delta G_2^N(t))) = 0 \end{cases} \tag{21}$$

Далее для простоты вычислений рассмотрим случай, когда игроки в рассматриваемой дифференциальной игре $\Gamma_{G_0}^{\infty}$ полностью симметричны, т.е. $G_1(0) = G_2(0) = g_0$, $a_1^N(t) = a_2^N(t) = a^N(t)$, $G_1^N(t) = G_2^N(t) = G^N(t)$.

В этом случае система уравнений (21) значительно упроститься, а именно получим:

$$\begin{cases} C_1 = C_2 = 0 \\ F_1C_3 + F_2C_4 = E - g_0 \\ -3\pi F_1^2 C_3^2 - \delta F_1 C_3^2 = 0 \\ -6\pi F_1F_2C_3C_4 - (F_1+F_2)C_3C_4 = 0 \\ 6\pi EF_1C_3 - 2\pi\beta F_1C_3 + 2\pi\beta EC_3(1-F_1) = 0 \end{cases}$$

Решив полученную систему, найдем следующие значения коэффициентов: $C_3 = 0, C_4 = \frac{E-g_0}{F_2}$.

Таким образом, для игры (15),(16) с симметричными игроками окончательно получим решение:

$$\begin{pmatrix} G_1^N \\ G_2^N \end{pmatrix} = \begin{pmatrix} (g_0 - E)e^{\frac{c\rho - D}{2c}t} + E \\ (g_0 - E)e^{\frac{c\rho - D}{2c}t} + E \end{pmatrix}$$

$$\begin{pmatrix} a_1^N \\ a_2^N \end{pmatrix} = \begin{pmatrix} \frac{E-g_0}{cF_2}e^{\frac{c\rho - D}{2c}t} + \frac{1}{c}Z \\ \frac{E-g_0}{cF_2}e^{\frac{c\rho - D}{2c}t} + \frac{1}{c}Z \end{pmatrix}$$

Численный пример

Пусть коэффициенты в модели (15),(16) принимают следующие значения: $\delta = \frac{1}{2}, c = 1, \rho = 2, g_0 = 5, \beta = 17, \pi = 20$.

Тогда для случая двух симметричных игроков при условии отсутствия их кооперации получим следующие оптимальные управления при указанных значениях параметров:

$$\begin{pmatrix} a_1^N \\ a_2^N \end{pmatrix} = \begin{pmatrix} 4.67 \cdot e^{-7.76 \cdot t} + 5.64 \\ 4.67 \cdot e^{-7.76 \cdot t} + 5.64 \end{pmatrix}$$

Оптимальные траектории игроков примут вид:

$$\begin{pmatrix} G_1^N \\ G_2^N \end{pmatrix} = \begin{pmatrix} -0.64 \cdot e^{-7.76 \cdot t} + 2.82 \\ -0.64 \cdot e^{-7.76 \cdot t} + 2.82 \end{pmatrix}$$

График оптимального управления игрока i при указанных значениях параметров модели (Рис.3)

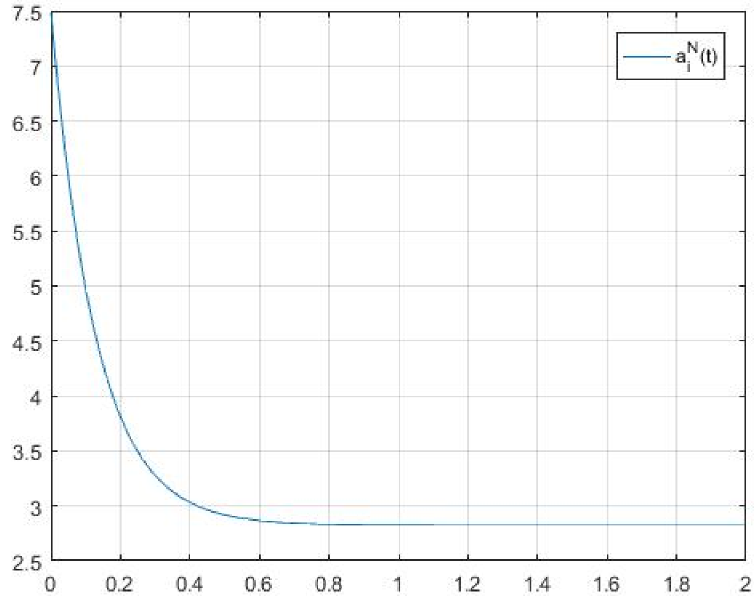

Рис. 3: Оптимальное управление $a_i^N(t)$

График соответствующей оптимальной траектории игрока i при указанных значениях параметров модели (Рис.4)

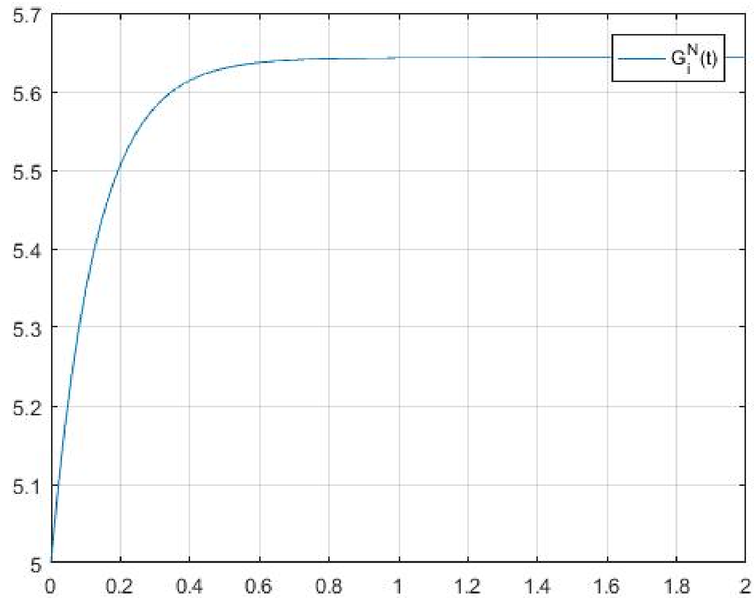

Рис. 4: Оптимальная траектория $G_i^N(t)$

Глава 4. О решении линейно-квадратичной дифференциальной игры при известном виде оптимального управления

Вновь обратимся к некооперативной дифференциальной игре управления выбросами в атмосферу. Рассмотрим в данной главе модель [7], несущественно отличающуюся от изученной в главе II:

$$\dot{x} = u_1 + u_2 + u_3, \quad x(0) = x_0 \tag{22}$$

$$K_i(x_0, T, u) = \int_0^T (u_i(b_i - \frac{u_i}{2}) - k_i x)dt \to \max_{u_i}, \quad i = 1, 2, 3 \tag{23}$$

Здесь $k_i > 0$. Решение будет найдено в программных стратегиях. Для задач оптимального управления такого типа искомое оптимальное управление будет иметь линейную структуру [21]: $u_i^N = A_i t + C_i$.

Тогда для (22),(23) можем записать:

$$\dot{x} = \sum_{i=1}^{3} A_i t + \sum_{i=1}^{3} C_i, \quad x(0) = x_0$$

$$K_i(x_0, T, u) = \int_0^T \Big(b_i(A_i t + C_i) - \frac{1}{2}(A_i t + C_i)) - k_i x\Big)dt \to \max_{A_i, C_i}, \quad i = 1, 2, 3$$

Из уравнения динамики игры получим:

$$x = \frac{1}{2}\sum_{i=1}^{3} A_i t^2 + \sum_{i=1}^{3} C_i t + x_0$$

Подставим полученное выражение в максимизируемый функционал:

$$K_i(x_0, T, u) = \int_0^T \Big(b_i(A_i t + C_i) - \frac{1}{2}(A_i t + C_i)) - k_i(\frac{1}{2}\sum_{i=1}^{3} A_i t^2 + \sum_{i=1}^{3} C_i t + x_0)\Big)dt =$$

$$= \frac{1}{2}b_i A_i T^2 + b_i C_i T - \frac{1}{6}A_i^2 T^3 - \frac{1}{2}C_i^2 T - \frac{1}{2}A_i C_i T^2 - \frac{1}{6}k_i \sum_{i=1}^{3} A_i T^3 -$$

$$-\frac{1}{2}k_i \sum_{i=1}^{3} C_i T^2 - k_i x_0 T, \quad i = 1, 2, 3$$

Найдем значения коэффициентов A_i, C_i, $i = \overline{1, 3}$, доставляющих максимум функционалу (23):

$$\begin{cases} \frac{dK_i}{dA_i} = \frac{1}{2}b_iT^2 - \frac{1}{3}A_iT^3 - \frac{1}{2}C_iT^2 - \frac{1}{6}T^3 = 0 \\ \frac{dK_i}{dC_i} = b_iT - C_iT - \frac{1}{2}A_iT^2 - \frac{1}{2}k_iT^2 = 0 \end{cases}$$

Получаем:

$$\begin{cases} A_i = k_i \\ C_i = b_i - k_iT, \quad i = 1, 2, 3 \end{cases}$$

Таким образом, окончательно получим, что для задачи (22),(23) оптимальное управление и траектория имеют вид:

$$u^N(t) = \begin{pmatrix} b_1 - k_1(T - t) \\ b_2 - k_2(T - t) \\ b_3 - k_3(T - t) \end{pmatrix}$$

$$x^*(t) = x_0 + \sum_{i=1}^{3} b_i - \sum_{i=1}^{3} d_i(t^2 - t)$$

Найденное оптимальное управление должно удовлетворять условию $u_i^N \in [0, b_i]$, $i = 1, 2, 3$ в силу того, что множество допустимых управлений должно являться компактом. Данное требование выполнено для $\forall t \in [0, T]$ в случае, когда параметры модели удовлетворяют следующему условию: $b_i \geq k_iT$.

Полученный в данной главе результат совпадает с решением, полученным ранее с помощью принципа максимума Понтрягина в работе [7].

Выводы

В ходе решения задачи оптимального управления (10) с динамикой изменения фазовой переменной, задаваемой уравнением (9), было проделано преобразование фазовой переменной, позволившее упростить дальнейший поиск решения задачи. Решение игры в кооперативной постановке было реализовано двумя методами. Поиск оптимальной траектории и оптимального управления с помощью уравнения Гамильтона—Якоби—Беллмана оказался более затратным с вычислительной точки зрения по сравнению с методом максимума Понтрягина. Тем не менее в обоих случаях решение игры было найдено в аналитическом виде.

В работе также была рассмотрена потенциальная динамическая дифференциальная игра (15),(16) для числа игроков $n = 2$. Поиск оптимальных траекторий и оптимальных управлений игроков был выполнен с использованием принципа максимума Понтрягина. В ходе решения задачи возникли сложности с решением системы уравнений, поэтому было использовано предположение о полной симметричности игроков.

Дополнительно был рассмотрен метод решения линейно-квадратичной дифференциальной игры в случае, когда вид искомого оптимального управления заведомо известен. Данный метод требует последующей проверки принадлежности найденного оптимального управления компакту, что может осложнить решение задачи.

Заключение

В ходе проделанной работы были рассмотрены различные методы решения дифференциальных игр с линейно-квадратичной функцией выигрыша на примере игр эксплуатации ресурсов.

Было проделано преобразование фазовой переменной для кооперативной игры управления выбросами в атмосферу, позволившее структурно упростить интегральный функционал игры. Показано, что применение данного метода значительно упрощает поиск решения рассмотренной задачи.

Также в работе был предложен новый путь решения дифференциальной игры управления ресурсами в рекламную кампанию. Было показано, что игра данного типа принадлежит классу потенциальных игр, решение которых может быть упрощено с помощью нахождения потенциала игры и сведения задачи к соответствующей задаче оптимального управления.

Список литературы

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

[2] Haurie A., Krawczyk J. B., Zaccour G. Games and Dynamic Games. Singapore: World Scientific Publishing, 2012. 465 p.

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

[4] Jørgensen S., Gromova E. *Sustaining cooperation in a differential game of advertising goodwill accumulation.* European Journal of Operational Research. 2016. P. 294–303.

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

[6] de Frutos J., Martín-Herrán G. *Selection of a Markov Perfect Nash Equilibrium in a Class of Differential Games.* Dyn Games Appl (2018) 8: 620

[7] Gromova E. (2016) *The Shapley Value as a Sustainable Cooperative Solution in Differential Games of Three Players.* In: Petrosyan L., Mazalov V. (eds) Recent Advances in Game Theory and Applications. Static & Dynamic Game Theory: Foundations & Applications. Birkhäuser, Cham

[8] Gromov D., Gromova E. *On a class of hybrid differential games.* Dynamic games and applications. 2016. Vol.7, № 2. P. 266-288.

[9] Лахина Ю.Э. *Оптимальное управление в задаче эксплуатации нескольких ресурсов.* Выпускная квалификационная работа бакалавра.

[10] Pontryagin L.S., Boltyanskii V.G., Gamkrelidze R.V., Mishenko E. *Matematicheskaya teoriya optimal'nykh protsessov.* Nauka, 1983.

[11] Gromova E.V., Tur A.V., Barsuk P.I. *A pollution control problem for the aluminum production in Eastern Siberia: Differential game approach.* submitted to SCP 2020

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

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

[14] Gonzalez-Sanchez D., Hernandez-Lerma O. *Potential Differential Games.* Dyn Games Appl (2018) 8: 254-279

[15] MondererD.,ShapleyL.S.*Potential games*. Games and Economic Behavior. 1996. V. 14. P. 124–143.

[16] Dechert W.D. *Optimal control problems from second order difference equations*. J. Econ. Theory. 1978. V. 19. P. 50–63.

[17] Dechert W.D. *Noncooperative dynamic games: a control theoretic approach.* University of Houston, 1997.

[18] Владимир В. Мазалов, Анна Н. Реттиева, Константин Е. Авраченков, *Линейно-квадратичные динамические потенциальные игры в дискретном времени*, МТИП, 9:1 (2017), 95–107; Autom. Remote Control, 78:8 (2017), 1537–1544

[19] Gromova E. *Теоретико-игровые задачи со случайной продолжительностью.*

[20] Алексеев В.М., Тихомиров В.М., Фомин С.В. *Оптимальное управление*, М.: Наука, 1979, 432 с.

[21] Basar T. and Olsder G.J. *Dynamic Noncooperative Game Theory*, 2nd edition, Classics in Applied Mathematics, SIAM, Philadelphia, 1999. 536 p.