Математическая Теория Игр и её Приложения, т.14, в.2, с. 31–50

УДК 517.977.8, 519.83
ББК 22.18

КООПЕРАТИВНЫЕ ДИФФЕРЕНЦИАЛЬНЫЕ ИГРЫ СО СЛУЧАЙНЫМ МОМЕНТОМ ПЕРЕКЛЮЧЕНИЯ ФУНКЦИИ ПОЛЕЗНОСТИ

Анастасия П. Заремба
Факультет прикладной математики –
процессов управления
Санкт-Петербургский государственный университет
198504, Санкт-Петербург, Университетский пр., 35
e-mail: st011335@student.spbu.ru

В статье формулируется дифференциальная игра n лиц, в которой функции полезности игроков имеют гибридный вид, а именно изменяются в случайный момент времени. При помощи интегрирования по частям упрощен вид функционала выигрыша. Для кооперативной постановки задачи изучена проблема динамической устойчивости выбранного игроками принципа оптимальности и предложено решение в виде адаптированной процедуры распределения дележа. В качестве примера рассмотрена дифференциальная игра инвестирования.

Ключевые слова: кооперативные игры, дифференциальные игры, устойчивая кооперация, гибридные дифференциальные игры, случайная продолжительность, процедура распределения дележа.

Поступила в редакцию: 19.01.22 *После доработки*: 07.04.22 *Принята к публикации*: 16.05.22

1. Введение

Теория кооперативных дифференциальных игр занимается изучением конфликтно-управляемых процессов со многими участниками [1]. Во главу угла ставится вопрос построения кооперативного решения для рассматриваемой игры. Часто перед игроками ставится задача максимизации суммарного выигрыша, определения кооперативной траектории, вдоль которой достигается выигрыш, и выбора принципа оптимальности, позволяющего распределить выигрыш между игроками. Впервые подробно описать дифференциальные игры в кооперативной постановке удалось в работе Л.А. Петросяна и Н.Н. Данилова [6]. Немаловажной является проблема динамической устойчивости выбранного решения: возможность реализуемости данного принципа оптимальности во времени.

В классической постановке процесс рассматривается на конечном фиксированном промежутке времени. В работах [4], [3], [7] Петросяном Л.А. был изучен вопрос динамической устойчивости принципа оптимальности для игр с предписанной продолжительностью. Им был предложен аппарат для решения проблемы нереализуемости дележа во времени путем введения специальной системы выплат, называемой процедурой распределения дележа [5],[2].

Естественным расширением класса дифференциальных игр с предписанной продолжительностью является предположение о случайности момента окончания игры. В работе Петросяна Л.А., Шевкопляс Е.В. [8] в общем виде была описана кооперативная постановка дифференциальной игры со случайной продолжительностью, в том числе был изучен вопрос динамической устойчивости принципа оптимальности для данного класса игр. Дифференциальные игры со случайным моментом окончания с гибридной функцией распределения были изучены в работах [12],[13], в которых решения модельных задач были получены при помощи гибридного принципа максимума [17], [18], [9]. Проблема нахождения кооперативного решения в случае разрывов первого рода для функции распределения момента окончания игры была изучена и решена в работах [14], [16], [19]. Кроме того, отметим, что интересной также представляется игра, в которой момент начала является случайной величиной. Формализация такой игры описана в работе [15].

В данной работе рассматривается класс дифференциальных игр с гибридными и случайными компонентами, в некотором смысле объединяющий в себе как игры со случайным началом, так и случайным окончанием. А именно, предполагается, что в некоторый случайный момент времени подынтегральные функции выигрышей (функции полезности) игроков меняют свой вид.

Похожая постановка игры была также описана в работе [10]. Авторами была рассмотрена модель разработки возобновляемых ресурсов в условиях изменения режима работы экосистемы. Функционирование рассматриваемой экосистемы было описано в виде дифференциальной игры с гибридной структурой и случайным моментом изменения функции полезности и уравнения динамики. В данной работе приводится общая формализация класса игр со случайным моментом переключения функции полезности, а также изучается проблема устойчивости во времени выбранного принципа оптимальности.

В разделе 2 приводится формальное описание игры, делаются основные предположения, допускающие наличие последующих выводов, упрощается функция выигрыша до вида, позволяющего использовать принцип максимума при дальнейшей оптимизации. Раздел 3 посвящен численному примеру, иллюстрирующему возможность поиска оптимального решения. Также в данном разделе решение рассматриваемой модели сравнивается с решением в случае случайного момента окончания игры. Раздел 4 посвящен форме выигрыша в подыгре, начинающейся после момента времени t_0. Раздел 5 описывает проблему динамической устойчивости выбранного принципа оптимальности и вариант ее решения с помощью принципа распределения дележа. Результат формулируется в виде теоремы.

2. Описание игры

Рассмотрим дифференциальную игру n лиц $\Gamma^T(x_0, t_0, T_f)$. Множество игроков обозначим как $N = \{1, 2, \ldots, n\}$, $|N| = n$.

Пусть игра $\Gamma^T(x_0, t_0, T_f)$ начинается в момент t_0 из состояния x_0 и развивается на конечном фиксированном промежутке времени $[t_0, T_f]$.

Предположим, что в случайный момент времени $T \in [t_0, T_f]$ происходит изменение функции мгновенного выигрыша для каждого игрока. Назовем этот момент моментом переключения. Для рассматриваемой случайной величины задана функция распределения $F(t)$ на отрезке $[t_0, T_f]$, т.е. случайная величина нормирована на рассматриваемом отрезке:

$$\int_{t_0}^{T_f} dF(t) = 1.$$

Положим, что динамика процесса описывается системой обыкновенных дифференциальных уравнений:

$$\dot{x} = g(x, u_1, \ldots, u_n), \quad x(t_0) = x_0.$$

Также положим, что переменная состояния системы принадлежит некоторому открытому подмножеству конечномерного вещественного евклидова пространства: $x(t) \in X \subset \mathbb{R}^r$ для всех $t \in [t_0, T_f]$.

Пусть управления u_i выбираются из множеств допустимых управлений $\mathcal{U}_i$, которые состоят из всех измеримых функций на отрезке $[t_0, T_f]$, принимающих значения в множестве допустимых значений управления $\mathbb{U}_i$, которое в свою очередь представляет собой выпуклое компактное подмножество конечномерного пространства $\mathbb{R}^k$.

Для удобства записи, в дальнейшем мы будем использовать сокращенные обозначения: $u = (u_1, \ldots, u_n)$, $\mathbb{U} = \prod_{i=1}^n \mathbb{U}_i$, $\mathcal{U} = \prod_{i=1}^n \mathcal{U}_i$. Таким образом, $u \in \mathcal{U}$, $u(t) \in \mathbb{U}$, $t \in [t_0; T_f]$.

Обозначим за $h_{i1}(x(\tau), u(\tau))$ – функцию мгновенного выигрыша в игре $\Gamma^T(x_0, t_0, T_f)$ игрока i в момент времени τ до или во время момента переключения T, то есть для $\tau \in [t_0, T]$. При этом, положим, что функция $h_{i1}(x(\tau), u(\tau))$ определена и интегрируема на всем рассматриваемом промежутке времени $[t_0, T_f]$.

Аналогично, обозначим функцию мгновенного выигрыша игрока после момента переключения. Пусть $h_{i2}(x(\tau), u(\tau))$ – функцию мгновенного выигрыша в игре $\Gamma^T(x_0, t_0, T_f)$ игрока i в момент времени τ после момента переключения T, то есть для $\tau \in (T, T_f]$. При этом, положим, что функция $h_{i2}(x(\tau), u(\tau))$ определена интегрируема на всем рассматриваемом промежутке времени $[t_0, T_f]$.

Далее, для простоты записи, можем использовать сокращения $h_{ik}(x(\tau), u(\tau)) = h_{ik}(\tau), k = \{1, 2\}$.

Выигрыш игрока i с учетом случайной компоненты модели можем записать как математическое ожидание выигрыша:

$$K_i(x_0,t_0,T_f,u) = \mathbb{E}\left(\int\limits_{t_0}^{t} h_{i1}(x(\tau),u(\tau))d\tau + \int\limits_{t}^{T_f} h_{i2}(x(\tau),u(\tau))d\tau\right) =$$

$$= \int\limits_{t_0}^{T_f}\left[\int\limits_{t_0}^{t} h_{i1}(x(\tau),u(\tau))d\tau + \int\limits_{t}^{T_f} h_{i2}(x(\tau),u(\tau))d\tau\right] dF(t) =$$

$$= \int\limits_{t_0}^{T_f}\int\limits_{t_0}^{t} h_{i1}(x(\tau),u(\tau))d\tau dF(t) + \int\limits_{t_0}^{T_f}\int\limits_{t}^{T_f} h_{i2}(x(\tau),u(\tau))d\tau dF(t).$$

Далее проинтегрируем по частям, для этого обозначим

$$H_{i1}(t) = \int_{t_0}^{t} h_{i1}(\tau)d\tau;$$

$$H_{i2}(t) = \int_{t}^{T_f} h_{i2}(\tau)d\tau;$$

Так как $H_{i1}(T_f)$ и $H_{i2}(t_0)$ - конечные значения, $F(t_0) = 0$, $F(T_f) = 1$, $H_{i1}(t_0) = 0$, $H_{i2}(T_f) = 0$ функционал $K_i(\cdot)$ может быть упрощен путем интегрирования по частям следующим образом:

$$K_i(x_0,t_0,T_f,u) = [H_{i1}(t)F(t)]_{t_0}^{T_f} - \int_{t_0}^{T_f} F(t)dH_{i1}(t) +$$

$$+ [H_{i2}(t)F(t)]_{t_0}^{T_f} - \int_{t_0}^{T_f} F(t)dH_{i2}(t) =$$

$$= H_{i1}(T_f)F(T_f) - H_{i1}(t_0)F(t_0) - \int_{t_0}^{T_f} F(t)dH_{i1}(t) +$$

$$+ H_{i2}(T_f)F(T_f) - H_{i1}(t_0)F(t_0) - \int_{t_0}^{T_f} F(t)dH_{i2}(t) =$$

$$= H_{i1}(T_f)\cdot 1 - 0\cdot 0 - \int_{t_0}^{T_f} F(t)dH_{i1}(t) + 0\cdot 1 - H_{i2}(t_0)\cdot 0 - \int_{t_0}^{T_f} F(t)dH_{i2}(t).$$

Согласно формуле Лейбница можем вычислить производную от $H_{i1}(t)$ и $H_{i2}(t)$.

Таким образом,

$$K_i(x_0,t_0,T_f,u)=\int_{t_0}^{T_f} h_{i1}(t)dt-\int_{t_0}^{T_f} F(t)h_{i1}(t)dt-\int_{t_0}^{T_f} F(t)(-h_{i2}(t))dt= \\ =\int_{t_0}^{T_f}(1-F(t))h_{i1}(t)dt+\int_{t_0}^{T_f} F(t)h_{i2}(t)dt. \quad (2.1)$$

Можем также записать функционал (2.1) следующим образом:

$$K_i(x_0,t_0,T_f,u)=\int_{t_0}^{T_f}\left[(1-F(t))h_{i1}(t)+F(t)h_{i2}(t)\right]dt.$$

$$K_i(x_0,t_0,T_f,u)=\int_{t_0}^{T_f}\left[h_{i1}(t)+F(t)(h_{i2}(t)-h_{i1}(t))\right]dt.$$

Таким образом, доказана следующая теорема:

Теорема 2.1. *Пусть в игре $\Gamma^T(x_0,t_0,T_f)$ функции мгновенной полезности $h_{ik}(x(\tau),u(\tau))=h_{ik}(\tau), k=\{1,2\}$ непрерывны по всем своим аргументам в области определения, тогда выигрыш игрока i может быть представлен в двух эквивалентных видах*

$$K_i(x_0,t_0,T_f,u)=\int_{t_0}^{T_f}\left[(1-F(t))h_{i1}(t)+F(t)h_{i2}(t)\right]dt. \quad (2.2)$$

$$K_i(x_0,t_0,T_f,u)=\int_{t_0}^{T_f}\left[h_{i1}(t)+F(t)(h_{i2}(t)-h_{i1}(t))\right]dt. \quad (2.3)$$

Заметим, что при условии неизменности функции полезности, то есть, когда $h_{i1}(t)$ и $h_{i2}(t)$ совпадают, функционалы (2.2) и (2.3) совпадают с видом функционала выигрыша для класса дифференциальных игр с предписанной продолжительностью [6].

Выигрыш, записанный в виде (2.2), может быть привычен для работы в связи с тем, что два интеграла по слагаемым имеют схожий вид с выигрышем в задаче со случайным моментом окончания и случайным моментом начала. Вид выигрыша (2.3) позволяет упростить вычисления в случае несущественных изменений функции полезности.

3. Пример игры инвестирования

Рассмотрим модификацию задачи инвестирования, описанной в [11], в рамках постановки Раздела 2.

Пусть n экономических агентов инвестирует в акции некоторой индустрии. Переменная состояния $x(t)$ отвечает за объем инвестирования в момент времени t, управления $u_i(t)$ есть инвестиционная стратегия игрока i на момент времени t, т.е. $u_i(t)$ – количество приобретаемых (продаваемых) акций в портфеле на момент времени t.

Динамика имеет следующий вид:

$$\dot{x}(t) = \sum_{i=1}^{n} u_i(t), \quad x \in R, \quad u_i \in U \subseteq R, \quad x(t_0) = x_0.$$

Для простоты положим $t_0 = 0$.

Положим, что функции мгновенных выигрышей являются линейными по переменной состояния и квадратичными по управлению и отличаются между собой лишь значением параметров. Такой вид функции выигрыша продиктован линейной зависимостью дохода от объема инвестирования в текущий момент времени $x(t)$ с некоторым коэффициентом q_{i1}, q_{i2} и расходами на приобретение активов, растущими квадратично с коэффициентом r_{i1}, r_{i2}. Подробнее с экономическим обоснованием модели читатель может ознакомиться в [11].

$$h_{i1}(t) = q_{i1}x(t) - r_{i1}u_i^2(t).$$

$$h_{i2}(t) = q_{i2}x(t) - r_{i2}u_i^2(t).$$

Пусть момент переключения распределен равномерно:

$$F(t) = \frac{t}{T_f}, t \in [0, T_f].$$

Выбор данного вида распределения обусловлен предположением, что момент переключения может быть осуществлен с равной вероятностью в любой момент времени промежутка развития процесса $[0, T_f]$.

Рассмотрим кооперативную форму игры, и как следствие следующую задачу максимизации суммарного выигрыша (математического ожидания выигрыша):

$$\int_0^{T_f} \sum_{i=1}^{n} [(1-F(t))h_{i1}(t) + F(t)h_{i2}(t)]dt \to \max_{u_i}.$$

В рамках рассматриваемой постановки имеем:

$$\int_0^{T_f} \sum_{i=1}^{n} [(1-F(t))(q_{i1}x(t) - r_{i1}u_i^2(t)) + F(t)(q_{i2}x(t) - r_{i2}u_i^2(t))]dt \to \max_{u_i}.$$

С детальными вычислениями можно ознакомиться в Приложении. Оптимальная траектория и оптимальное управление i-го игрока примут вид (3.1) и (3.2) соответственно

$$x^*(t) = \sum_{i=1}^{n} \frac{1}{8(r_{i1}-r_{i2})^3} \left[(\hat{q}_2 - \hat{q}_1)(r_{i1}-r_{i2})^2 t^2 - \right.$$
$$-2T_f(\hat{q}_1(r_{i1}^2 - r_{i1}r_{i2} + r_{i2}^2) + \hat{q}_2 r_{i2}(r_{i2} - 2r_{i1})) \ln \frac{r_{i1}(T_f - t) + r_{i2}t}{r_{i1}T_f} +$$
$$\left. +2(r_{i1}-r_{i2})T_f t(\hat{q}_2 r_{i1} - \hat{q}_1 r_{i2})\right] + x_0. \quad (3.1)$$

$$u_i^*(t) = \frac{-2\hat{q}_1 t T_f + (\hat{q}_1 - \hat{q}_2)t^2 + (\hat{q}_1 + \hat{q}_2)T_f^2}{4r_{i1}(T_f - t) + 4r_{i2}t}, \quad (3.2)$$

где $\hat{q}_1 = \sum_{i=1}^{n} q_{i1}, \hat{q}_2 = \sum_{i=1}^{n} q_{i2}$.

Поведение оптимальной траектории и оптимального управления на рассматриваемом промежутке времени зависит от значения параметров $q_{ik}, r_{ik}, k = \overline{1,2}$. В общем виде нельзя утверждать о монотонности функций. Однако, сравнительный анализ в отношении игры со случайным моментом окончания можно провести зафиксировав значения параметров одинаковыми для обеих постановок.

Рисунок 1 позволяет сравнить вид оптимальной траектории в четырех разных постановках описанной модели при одинаковых параметрах q_{i1}, r_{i1}, но разной продолжительности и изменении параметров r_{i2} в гибридной постановке при переключении.

Рисунок 2 аналогичным образом демонстрирует различие в форме оптимального управления одного из игроков для четырех модификаций модели.

На Рисунке 3 изображены графики подынтегральной функции выигрыша гранд коалиции для каждой из модификаций модели.

Красная траектория на Рисунке 1 и управление на Рисунке 2 отвечает процессу с предписанной продолжительностью без переключения. Зеленая траектория и управление отвечает игре со случайной продолжительностью, где момент окончания игры распределен равномерно на отрезке $[0, T_f]$. Синяя траектория и управление соответствует гибридной поставке, описанной в данной работе, в предположении, что после переключения расходы на приобретение активов (комиссия брокера) выросли в 2 раза: $r_{i2} = 2r_{i1}, i = \overline{1,n}$, при неизменных прочих параметрах. Оранжевая траектория и управление также соответствует гибридной постановке, но в предположении уменьшения расходов: $r_{i2} = 0.5r_{i1}, i = \overline{1,n}$, при неизменных прочих параметрах.

Объем позиции портфеля на момент времени (количество активов, которыми владеет экономический агент в момент времени), которому отвечает переменная состояния возрастает во всех случаях, кроме гибридной модели с увеличением издержек в случайный момент (Рисунок 1). Управление, в свою очередь, есть инвестиционная стратегия игрока: объем купленных акций в момент времени, влияющих на скорость изменения общей позиции. По Рисунку 2 следует заключить об уменьшении темпа инвестиций во времени. Наибольшему объему портфеля соответствует модель с предписанной продолжительностью. Можно сделать вывод, что дополнительный риск в виде случайности в отношении окончания процесса или увеличения расходов мотивируют игроков к более осторожному поведению, которое проявляется в меньшей величине позиции и, как следствие, меньшему мгновенному выигрышу. Наиболее ярко это видно при сравнении гибридной модели в случае увеличения издержек после переключения в случайный момент времени и уменьшения издержек после переключения в случайный момент времени. Риск увеличения издержек заставляет приобретать меньше, и, по итогу, иметь меньший выигрыш (площадь под графиком).

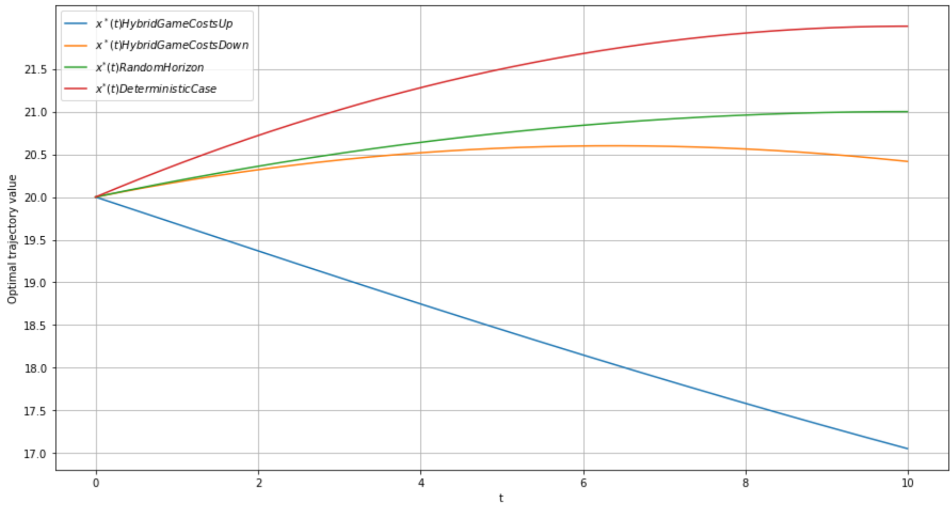

Рисунок 1. Оптимальная траектория в случае гибридной игры в сравнении с другими моделями

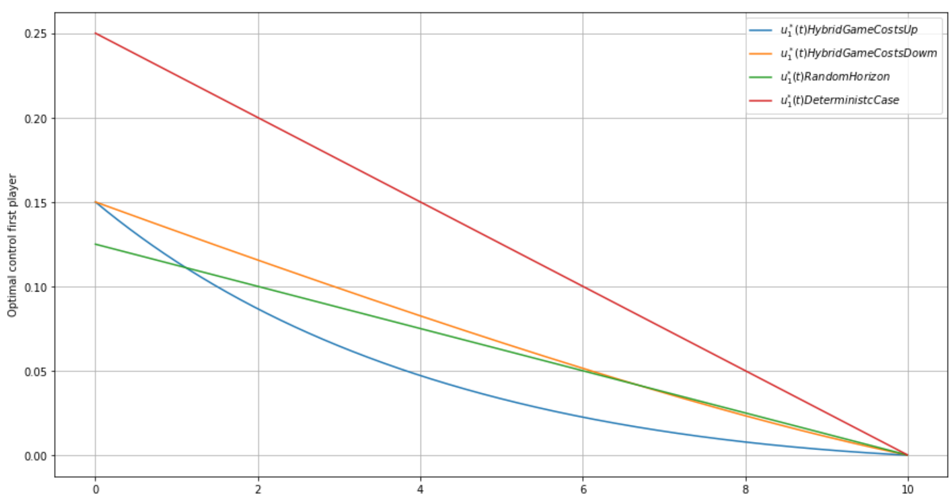

Рисунок 2. Оптимальное управление в случае гибридной игры в сравнении с другими моделями

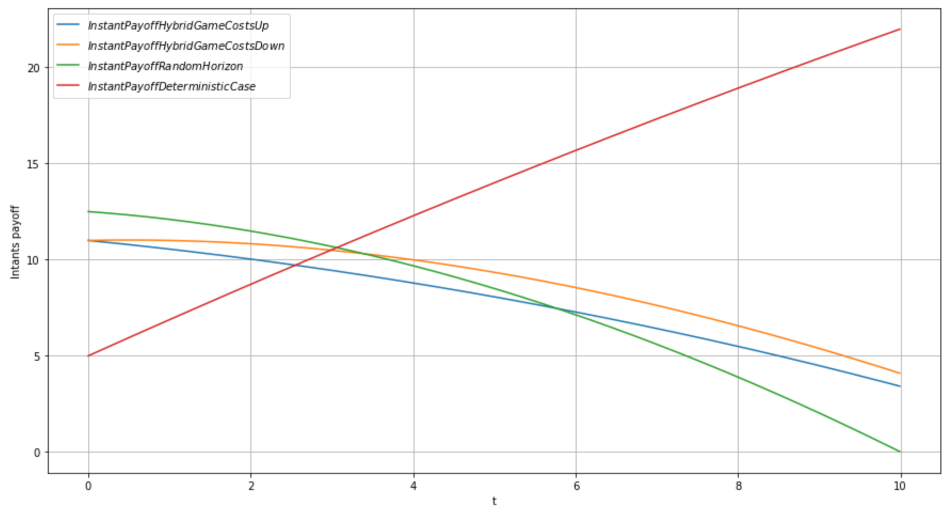

Рисунок 3. Подынтегральная функция выигрыша игрока в случае гибридной игры в сравнении с другими моделями

4. Выигрыш в подыгре

Рассмотрим некоторый момент времени $\theta \in [t_0, T_f]$. Начиная с момента времени θ, игроки попадают в подыгру $\Gamma^T(x^*(\theta), \theta, T_f)$.

Рассмотрим полную группу событий: первая игра закончилась до θ (вероятность события $F(\theta)$), первая игра не закончилась до θ (вероятность события $1-F(\theta)$). Выпишем вероятность события «в подыгре первая игра закончилась до t» согласно формуле полной вероятности

$$F_\theta(t) = \frac{F(t) - F(\theta)}{1 - F(\theta)}.$$

Тогда математическое ожидание выигрыша в подыгре может быть выписано как

$$K_i(x^*(\theta), \theta, T_f, u) = \int_\theta^{T_f} \left[\int_\theta^t h_{i1}(x(\tau), u(\tau)) d\tau + \right.$$
$$\left. + \int_t^{T_f} h_{i2}(x(\tau), u(\tau)) d\tau \right] dF_\theta(t).$$

Аналогично, интегрирую по частям, получим

$$H_{i1}^{\theta}(t) = \int_{\theta}^{t} h_{i1}(\tau)d\tau;$$

$$H_{i2}^{\theta}(t) = \int_{t}^{T_f} h_{i2}(\tau)d\tau;$$

Так как $H_{i1}^{\theta}(T_f)$ и $H_{i2}^{\theta}(\theta)$ - конечные значения, $F_\theta(T_f) = 1$, $F_\theta(\theta) = 0$, $H_{i2}(T_f) = 0$ функционал $K_i(\cdot)$ может быть упрощен путем интегрирования по частям следующим образом:

$$K_i(x^*(\theta), \theta, T_f, u) = \left[H_{i1}^{\theta}(t)F_\theta(t)\right]_{\theta}^{T_f} - \int_{\theta}^{T_f} F_\theta(t)dH_{i1}^{\theta}(t) +$$

$$+ \left[H_{i2}^{\theta}(t)F_\theta(t)\right]_{\theta}^{T_f} - \int_{\theta}^{T_f} F_\theta(t)dH_{i2}^{\theta}(t) =$$

$$= H_{i1}^{\theta}(T_f)F_\theta(T_f) - H_{i1}^{\theta}(\theta)F_\theta(\theta) - \int_{\theta}^{T_f} F_\theta(t)dH_{i1}^{\theta}(t) +$$

$$+ H_{i2}^{\theta}(T_f)F_\theta(T_f) - H_{i1}^{\theta}(\theta)F_\theta(\theta) - \int_{\theta}^{T_f} F_\theta(t)dH_{i2}^{\theta}(t) =$$

$$= H_{i1}^{\theta}(T_f)\cdot 1 - 0\cdot 0 - \int_{\theta}^{T_f} F_\theta(t)dH_{i1}^{\theta}(t) + 0\cdot 1 - H_{i2}^{\theta}(\theta)\cdot 0 - \int_{\theta}^{T_f} F_\theta(t)dH_{i2}^{\theta}(t) =$$

$$= \int_{\theta}^{T_f} h_{i1}(t)dt - \int_{\theta}^{T_f} F_\theta(t)h_{i1}(t)dt - \int_{\theta}^{T_f} F_\theta(t)(-h_{i2}(t))dt =$$

$$= \int_{\theta}^{T_f} (1 - F_\theta(t))h_{i1}(t)dt + \int_{\theta}^{T_f} F_\theta(t)h_{i2}(t)dt. \qquad (4.1)$$

Можем также записать функционал (4.1) следующим образом:

$$K_i(x^*(\theta), \theta, T_f, u) = \frac{1}{1 - F(\theta)} \int_{\theta}^{T_f} \left[(1 - F(t))h_{i1}(t) + F(t)h_{i2}(t)\right] dt.$$

$$K_i(x^*(\theta), \theta, T_f, u) = \frac{1}{1 - F(\theta)} \int_{\theta}^{T_f} \left[h_{i1}(t) + F(t)(h_{i2}(t) - h_{i1}(t))\right] dt.$$

Таким образом, доказана следующая теорема:

Теорема 4.1. *В подыгре $\Gamma^T(x^*(\theta), \theta, T_f)$ выигрыш игрока i может быть представлен в двух эквивалентных видах*

$$K_i(x^*(\theta), \theta, T_f, u) = \frac{1}{1 - F(\theta)} \int_\theta^{T_f} \left[(1 - F(t))h_{i1}(t) + F(t)h_{i2}(t)\right] dt.$$

$$K_i(x^*(\theta), \theta, T_f, u) = \frac{1}{1 - F(\theta)} \int_\theta^{T_f} \left[h_{i1}(t) + F(t)(h_{i2}(t) - h_{i1}(t))\right] dt.$$

Нетрудно заметить, что при условии $\theta = t_0$ результат Теоремы 4.1 совпадает с результатом Теоремы 2.1.

5. Проблема динамической устойчивости

Пусть в игре $\Gamma^T(x_0, t_0, T_f)$ задана некоторая характеристическая функция $V^T(x_0, t_0, T_f)$. Таким образом, можем сказать, что задана игра в форме характеристической функции

$$\Gamma_V^T(x_0, t_0, T_f) = < N, V^T(x_0, t_0, T_f, \cdot) > .$$

При движении вдоль выбранной игроками кооперативной траектории $x^*(t)$ может возникнуть ситуация, когда одному или нескольким игрокам (некоторой коалиции) будет выгодно отклониться от изначально выбранного профиля стратегий и принципа оптимальности, т.е. возникает проблема динамической устойчивости выбранного решения.

Положим, что до начала игры игроки договорились о распределении суммарного выигрыша согласно некоторому принципу оптимальности. Обозначим соответствующий выбранному принципу оптимальности дележ как $\xi(x_0, t_0, T_f) = (\xi_1(x_0, t_0, T_f), \ldots, \xi_n(x_0, t_0, T_f))$.

Введем определение процедуры распределения дележа для игры $\Gamma^T(x_0, t_0, T_f)$.

Определение 5.1. *Рассмотрим набор вектор-функций* $\{\beta_1(t), \beta_2(t)\} = \{(\beta_{11}(t), \ldots, \beta_{n1}(t)), (\beta_{12}(t), \ldots, \beta_{n2}(t))\}$, $\beta_{ij}(t) \geq 0 \quad \forall i = \overline{1, n}, \quad j = \overline{1, 2}$, *такой, что компоненты вектора дележа в игре $\Gamma^T(x_0, t_0, T_f)$ удовлетворяют условию*

$$\xi_i(x_0, t_0, T_f) = \int_{t_0}^{T_f} \left[(1 - F(\tau))\beta_{i1}(\tau) + F(\tau)\beta_{i2}(\tau)\right] d\tau, \ i = \overline{1, n}.$$

Набор вектор- функций $\{\beta_1(t), \beta_2(t)\}$ будем называть процедурой распределения дележа (ПРД).

Определение 5.2. *Будем называть принцип оптимальности и соответствующий ему дележ $\xi(x_0, t_0, T_f) = \{\xi_i(x_0, t_0, T_f)\}$ в дифференциальной игре $\Gamma_V^T(x_0, t_0, T_f)$ динамически устойчивым, если существует такая ПРД $\{\beta_1(t), \beta_2(t)\}$, что для любого $\theta \in [t_0, T_f)$ вектор $\xi^\theta = (\xi_1^\theta, \ldots, \xi_N^\theta)$, где*

$$\xi_i^\theta = \frac{1}{1 - F(\theta)} \left[\int_{t_0}^{T_f} [(1 - F(t))\beta_{i1}(t) + F(t)\beta_{i2}(t)]\, dt \right].$$

для всех $i = 1, \ldots, N$, принадлежит тому же изначально выбранному принципу оптимальности, в подыгре $\Gamma^T(x^*(\theta), \theta, T_f)$, т.е. ξ^θ также является дележом в $\Gamma^T(x^*(\theta), \theta, T_f)$;

Определим соотношение между дележом ξ и процедурой распределения этого дележа $\{\beta_1(t), \beta_2(t)\}$.

Лемма 5.1. *Если $t_0 \le \theta \le T_f$, для всех $i = 1, \ldots, n$, компоненты вектора дележа ξ могут быть записаны следующим образом:*

$$\xi_i = \int_{t_0}^{\theta} [(1 - F(t))\beta_{i1}(t) + F(t)\beta_{i2}(t)]\, dt + (1 - F(\theta))\xi_i^\theta. \tag{5.1}$$

Доказательство. Можем записать:

$$\xi_i = \int_{t_0}^{T_f} [(1 - F(t))\beta_{i1}(t) + F(t)\beta_{i2}(t)]\, dt =$$

$$= \int_{t_0}^{\theta} [(1 - F(t))\beta_{i1}(t) + F(t)\beta_{i2}(t)]\, dt +$$

$$+ \int_{\theta}^{T_f} [(1 - F(t))\beta_{i1}(t) + F(t)\beta_{i2}(t)]\, dt =$$

$$= \int_{t_0}^{\theta} [(1 - F(t))\beta_{i1}(t) + F(t)\beta_{i2}(t)]\, dt + (1 - F(\theta))\xi_i^\theta,$$

что приводит к выражению (5.1). □

Лемма 5.2. *Если $\theta \in [t_0, T_f)$, тогда для всех $i = 1, \ldots, n$, справедливо следующее соотношение:*

$$\beta_{i1}(\theta) + \frac{F(\theta)}{1-F(\theta)}\beta_{i2}(\theta) = \frac{f(\theta)}{1-F(\theta)}\xi_i^\theta - (\xi_i^\theta)',$$

где $f(\theta) = F'(\theta)$ – функция плотности распределения.

Доказательство. Пусть $\theta \in [t_0,\ T_f)$, продифференцируем выражение (5.1) по θ, получим:

$$0 = \beta_{i1}(\theta)(1-F(\theta)) + F(\theta)\beta_{i2}(\theta) - f(\theta)\xi_i^\theta + (1-F(\theta))(\xi_i^\theta)'.$$

Откуда,

$$\beta_{i1}(\theta) + \frac{F(\theta)}{1-F(\theta)}\beta_{i2}(\theta) = \frac{f(\theta)}{1-F(\theta)}\xi_i^\theta - (\xi_i^\theta)'.$$

□

Выше полученные результаты могут быть объединены в качестве теоремы:

Теорема 5.1. *Пусть дележ $\xi(x^*(t), t, T_f)$ в игре $\Gamma_V^T(x*(t), t, T_f)$ есть абсолютно непрерывная функция t, $t \in [t_0,\ T_f]$. Если существует соответствующая ПРД, которая для $\tau \in [t_0,\ T_f]$ удовлетворяет следующему соотношению:*

$$\beta_{i1}(\tau) + \frac{F(\tau)}{1-F(\tau)}\beta_{i2}(\tau) = \frac{f(\tau)}{1-F(\tau)}\xi_i(x^*(\tau), \tau, T_f) - \xi_i'(x^*(\tau), \tau, T_f), \tag{5.2}$$

тогда $\xi(x_0, t_0, T_f)$ является динамически устойчивым дележом.

Данная теорема позволяет убедиться в состоятельности во времени выбранного принципа оптимальности, построив специальную систему выплат во времени (ПРД), используя условие (5.2) наряду с Определением 5.1. Однако, Теорема 5.1 не гарантирует существование такой ПРД.

Обратим внимание, что при условии отсутствия случайности, когда функция распределения тождественна равна 1 на всем промежутке определения, результат Теоремы 5.1 совпадает с результатом, полученным для класса дифференциальных игр с предписанной продолжительностью [3].

6. Заключение

В работе был описан класс дифференциальных игр со случайным моментом переключения функции полезности. Для данной постановки был упрощен функционал выигрыша. Нахождение оптимальных управлений продемонстрировано на примере дифференциальной игры инвестирования. Результаты были проиллюстрированы на графиках, позволяющих сравнить вид оптимальной траектории и оптимального управления при наличии и отсутствии переключения функции полезности. Для описанной игры также была изучена проблема динамической устойчивости выбранного решения в кооперативной постановке и предложен вариант ее разрешения путем построения системы платежей во времени (процедуры распределения дележа).

7. Приложение

Воспользуемся принципом максимума Понтрягина, при условии закрепленного левого конца: $x(t) = x_0$, гамильтониан для задачи (3)

$$H(x(t), u(t), \psi(t)) = \psi(t) \sum_{i=1}^{n} u_i(t) +$$
$$+ \sum_{i=1}^{n} [\frac{T_f - t}{T_f}(q_{i1}x(t) - r_{i1}u_i^2(t)) + \frac{t}{T_f}(q_{i2}x(t) - r_{i2}u_i^2(t))]. \quad (7.1)$$

Обозначим $\hat{q}_1 = \sum_{i=1}^{n} q_{i1}, \hat{q}_2 = \sum_{i=1}^{n} q_{i2}$

Таким образом, из (7.1), можем получить выражение для оптимального управления.

$$u_i^*(t) = \frac{\psi(t)T_f}{2r_{i1}(T_f - t) + 2r_{i2}t}. \quad (7.2)$$

Вторая производная даст подтверждение тому, что это максимум

$$\frac{\partial^2 H}{\partial u_i^2(t)} = -2r_{i1}\frac{T_f - t}{T_f} - 2r_{i2}\frac{t}{T_f} < 0.$$

Уравнение для сопряженной переменной $\psi(t)$ примет вид

$$\frac{\partial\psi}{\partial t} = -\frac{\partial H}{\partial x} = -\hat{q}_1\frac{T_f - t}{T_f} - \hat{q}_2\frac{t}{T_f}, \quad \psi(0) = \psi_0, \psi(T_f) = 0.$$

Откуда в конечном счете путем интегрирования получаем

$$\psi(t) = -\hat{q}_1 t + \frac{(\hat{q}_1 - \hat{q}_2)t^2}{2T_f} + \frac{\hat{q}_1 + \hat{q}_2}{2}T_f. \tag{7.3}$$

Таким образом, можем подставить полученный вид сопряженной переменной (7.3) в (7.2) оптимальное управление окончательно принимает вид

$$u_i^*(t) = \frac{-2\hat{q}_1 t T_f + (\hat{q}_1 - \hat{q}_2)t^2 + (\hat{q}_1 + \hat{q}_2)T_f^2}{4r_{i1}(T_f - t) + 4r_{i2}t}. \tag{7.4}$$

Используя выражение (7.4), запишем уравнение для переменной состояния

$$\dot{x}(t) = \sum_{i=1}^{n} u_i(t) = \sum_{i=1}^{n} \frac{-2\hat{q}_1 t T_f + (\hat{q}_1 - \hat{q}_2)t^2 + (\hat{q}_1 + \hat{q}_2)T_f^2}{4r_{i1}(T_f - t) + 4r_{i2}t}, \quad x(0) = x_0.$$

Путем интегрирования получим оптимальную траекторию системы

$$\begin{aligned} x(t) = \sum_{i=1}^{n} \frac{1}{8(r_{i1} - r_{i2})^3} \Big[& (\hat{q}_2 - \hat{q}_1)(r_{i1} - r_{i2})^2 t^2 - \\ & -2T_f(\hat{q}_1(r_{i1}^2 - r_{i1}r_{i2} + r_{i2}^2) + \hat{q}_2 r_{i2}(r_{i2} - 2r_{i1})) \ln\frac{r_{i1}(T_f - t) + r_{i2}t}{r_{i1}T_f} + \\ & +2(r_{i1} - r_{i2})T_f t(\hat{q}_2 r_{i1} - \hat{q}_1 r_{i2}) \Big] + x_0. \end{aligned}$$

СПИСОК ЛИТЕРАТУРЫ

1. Айзекс Р. *Дифференциальные игры* // М.: Мир. 1967.

2. Петросян Л.А. *О новых сильно динамически устойчивых принципах оптимальности в кооперативных дифференциальных играх* // М:Труды математического института им. Стеклова «Оптимальное управление и дифференциальные уравнения». 1995. Т. 211. С. 370–376.

3. Петросян Л.А. *Сильно динамически устойчивые дифференциальные принципы оптимальности* // Вестник ЛГУ, Серия 1: математика, механика, астрономия. 1993. № 4. С. 35–40.

4. Петросян Л.А. *Устойчивость решений в дифференциальных играх со многими участниками* // Вестник ЛГУ. 1977. № 4. С. 46–52.

5. Петросян Л.А. *Характеристические функции кооперативных дифференциальных игр* // Вестник СПбГУ, сер. 1: Математика, механика, астрономия. 1995. № 1. С. 48–52.

6. Петросян Л.А., Данилов Н.Н. *Устойчивые решения неантагонистических дифференциальных игр с трансферабельными выигрышами* // Вестник ЛГУ. 1979. № 1. С. 46–54.

7. Петросян Л.А., Зенкевич Н.А. *Принципы устойчивой кооперации* // Математическая теория игр и ее приложения. 2009. С. 106–123.

8. Петросян Л.А., Шевкопляс Е.В. *Кооперативные дифференциальные игры со случайной продолжительностью* // Вестник СПбГУ. 2000. Сер. 1. Вып. 4. С. 18–23.

9. Azhmyakov V., Attia S.A., Gromov D., Raisch J. *Necessary optimality conditions for a class of hybrid optimal control problems* // HSCC 2007. LNCS. 2007. Vol. 4416. P. 637–640.

10. De Zeeuw A., He X. *Managing a renewable resource facing the risk of a regime shift in the ecological system* // Resour. EnergyEcon. 2017. P. 42–54.

11. Dockner E.J., Jorgensen S., Long N.V., Sorger G. *Differential Games in Economics and Management Science*. Cambridge, UK: Cambridge University Press, 2000.

12. Gromov D.V., Gromova E.V. *Differential games with random duration: a hybrid systems formulation* // Contributions to game theory and management. 2014. P. 104–119.

13. Gromov D.V., Gromova E.V. *On a class of hybrid differential games* // Dynamic Games and Applications. 2017. P. 266–288.

14. Gromova E., Malakhova A., Palestini A. *Payoff Distribution in a Multi-Company Extraction Game with Uncertain Duration* // Mathematics. 2018. Vol. 6. P. 165.

15. Malakhova A.P., Gromova E.V. *Dynamic programming equations for the game-theoretical problem with random initial time* // Lecture Notes in Control and Information Sciences – Proceedings. Springer. 2020.

16. Malakhova A.P., Gromova E.V. *Strongly Time-Consistent Core in Differential Games with Discrete Distribution of Random Time Horizon* // Math. Appl. 2018. Vol. 46. P. 197–209.

17. Riedinger P., Iung C., Kratz F. *An optimal control approach for hybrid systems* // European Journal of Control. 2003. Vol. 9(5). P. 449–458.

18. Shaikh M.S., Caines P.E. *On the hybrid optimal control problem: Theory and algorithms* // IEEE Transactions on Automatic Control. 2007. Vol. 52(9). P. 1587–1603.

19. Zaremba A., Gromova E., Tur A. *A Differential Game with Random Time Horizon and Discontinuous Distribution*// Mathematics. 2020. Vol. 8. P. 2185.

COOPERATIVE DIFFERENTIAL GAMES WITH THE UTILITY FUNCTION SWITCHED AT A RANDOM TIME MOMENT

Anastasiia P. Zaremba, Saint Petersburg State University, PhD student (malakhova.a.p@gmail.com).

Abstract: This paper describes a differential game of n persons in which the utility functions of the players have a hybrid form, namely, they are changed at a random moment in time. With the help of integration in parts, the form of the payoff functional is simplified. For the cooperative scenario the problem of time-consistency of the optimality principle chosen by the players is studied and a solution is proposed in the form of an adapted imputation distribution procedure. The differential investment game is considered as an example.

Keywords: cooperative games, differential games, time-consistency problem, random duration, hybris differential games, imputation distribution procedure.