ТРУДЫ ИНСТИТУТА МАТЕМАТИКИ И МЕХАНИКИ УрО РАН

Том 26 № 1 2020

УДК 517.977

О МАКСИМАЛЬНОМ ГАРАНТИРОВАННОМ ВЫИГРЫШЕ В НЕКОТОРЫХ ЗАДАЧАХ КОНФЛИКТНОГО УПРАВЛЕНИЯ МНОГОШАГОВЫМИ ПРОЦЕССАМИ

М. С. Никольский

В статье рассматриваются многошаговые конфликтно управляемые процессы с двумя управляющими субъектами. Продолжительность процесса фиксирована и нет ограничения на правый конец дискретной траектории. Первый игрок стремится к максимизации терминального функционала, причем информация о будущем поведении второго игрока отсутствует. В статье изучается важное понятие максимального гарантированного выигрыша первого игрока с помощью идей беллмановского метода динамического программирования. В теореме 1 при широких предположениях относительно изучаемого конфликтно управляемого процесса с помощью метода динамического программирования получена формула для искомого максимального гарантированного выигрыша. В теореме 2 получены достаточные условия, обеспечивающие липшицевость соответствующих функций беллмановского типа. Для иллюстрации рассмотрено два примера.

Ключевые слова: многошаговые управляемые процессы, конфликт, динамическое программирование.

M. S. Nikol'skii. On the maximal guaranteed payoff in some problems of conflict control of multistep processes.

We consider multistep conflict-controlled processes with two controlling partners. The duration of the process is fixed, and there are no constraints on the right end of the discrete trajectory. The first player aims to maximize the terminal functional without information about the future behavior of the second player. We study the important notion of maximal guaranteed payoff of the first player using the ideas of Bellman's dynamic programming method. Based on this method, a formula for the maximal guaranteed payoff is derived in Theorem 1 under broad assumptions on the conflict-controlled process. In Theorem 2, we obtain sufficient conditions under which the corresponding functions of Bellman type are Lipschitz. Two examples are considered.

Keywords: discrete controlled processes, conflict, dynamical programming.

MSC: 42C10, 47A58

DOI: 10.21538/0134-4889-2020-26-1-167-172

Введение

Теория многошаговых управляемых процессов является важным разделом современной математической теории управления. Имеются определенные связи и аналоги этой теории с теорией оптимального управления непрерывными динамическими системами. Многошаговые управляемые процессы с успехом используются, например, при моделировании экономических процессов, развивающихся во времени, так как в экономике текущее время часто квантуется по периодам день, неделя и так далее. Помимо управляемых многошаговых процессов в теории управления интересным и актуальным объектом изучения являются конфликтно управляемые многошаговые процессы. Они оказываются полезными, в частности, при изучении управляемых многошаговых процессов при наличии разного рода возмущений, поведение которых не предсказуемо.

Отметим, что современная теория дифференциальных игр тесно связана с теорией конфликтно управляемых многошаговых процессов.

Многошаговые конфликтно управляемые процессы (см., например, [1–3]) представляют значительный интерес при моделировании конфликтных процессов, развивающихся во времени. При их изучении оказываются полезными общие концепции теории игр и теории многошаговых управляемых процессов. Следует также отметить глубокие связи между теорией многошаговых конфликтно управляемых процессов и теорией дифференциальных игр (см., например, [1; 2; 4]).

В статье изучается важное понятие максимального гарантированного выигрыша с точки зрения первого игрока в играх двух лиц, описываемых конфликтно управляемым многошаговым процессом. В теореме 1 при широких предположениях с помощью метода динамического программирования, восходящего к работам Р. Беллмана, получена формула для искомого максимального гарантированного результата. В теореме 2 найдены конструктивные достаточные условия, при которых соответствующие функции беллмановского типа, участвующие в построении максимального гарантированного результата, удовлетворяют условию Липшица. Этот результат полезен при приближенных вычислениях искомого максимального гарантированного результата.

1. Условимся символом $\mathbb{R}^k$ $(k \geqslant 1)$ обозначать арифметическое евклидово пространство, элементами которого являются упорядоченные столбцы из k чисел, со стандартным скалярным произведением $\langle\cdot,\cdot\rangle$. Для $y \in \mathbb{R}^k$ под символом $|y|$ будем понимать стандартную длину вектора y.

Далее рассматривается многошаговый конфликтно управляемый процесс (см., например, [1–3]) вида

$$x_{t+1} = f(x_t, u_t, v_t), \tag{1}$$

где $t = 0, \ldots, N-1$ $(N \geqslant 1)$, $x_t \in \mathbb{R}^n$, $u_t \in P \subset \mathbb{R}^p$, P — компакт в $\mathbb{R}^p$; $v_t \in Q \subset \mathbb{R}^q$, Q — компакт в $\mathbb{R}^q$; $f(x,u,v)$ — n-мерная векторная функция, определенная и непрерывная по совокупности переменных на $\mathbb{R}^n \times P \times Q$.

Для конфликтно управляемого процесса (1) фиксировано начальное условие

$$x(0) = x_0. \tag{2}$$

Процесс движения вектора x_t (см. (1), (2)) рассматривается с точки зрения первого игрока, распоряжающегося выбором вектора $u \in P$, причем он стремится к максимизации терминального функционала $\varphi(x_N)$, где функция $\varphi(x)$ определена и непрерывна на $\mathbb{R}^n$. Предполагается, что первый игрок знает функции $f(x,u,v)$, $\varphi(x)$ и в каждый момент $t = 0, \ldots, N-1$ знает текущий фазовый вектор x_t. На основании доступной ему информации он старается выбрать функции $u_t(x) \in P$ $(t = 0, \ldots, N-1,\ x \in \mathbb{R}^n)$ так, чтобы обеспечить как можно большее значение критерия $\varphi(x_N)$. Предполагается далее, что второй игрок знает функцию $f(x,u,v)$ и в каждый момент $t = 0, \ldots, N-1$ знает текущий фазовый вектор x_t. В качестве стратегий второго игрока выступают всевозможные функции $v_t(x) \in Q$ $(t = 0, \ldots, N-1,\ x \in \mathbb{R}^n)$. О целях второго игрока никаких специальных предположений не делается.

Отметим, что об аналитических свойствах функций $u_t(x)$, $v_t(x)$ по переменным (t,x) также никаких специальных предположений не делается.

Заметим, что для конфликтно управляемого процесса (1), (2) при произвольных допустимых стратегиях

$$U = \{u_0(x), \ldots, u_{N-1}(x)\},\ \ V = \{v_0(x), \ldots, v_{N-1}(x)\}$$

на выходе мы получаем однозначно определенный вектор $x_N(U,V)$ и, следовательно, значение функции выигрыша первого игрока $\varphi(x_N(U,V))$.

Качество (гарантированный выигрыш) произвольной допустимой стратегии U естественно оценить величиной

$$\alpha(U) = \inf_V \varphi(x_N(U,V)). \tag{3}$$

Здесь V — произвольная допустимая стратегия второго игрока. Предельные возможности первого игрока можно оценить величиной (максимальный гарантированный выигрыш)

$$\beta = \sup_U \alpha(U), \tag{4}$$

где U — произвольная допустимая стратегия первого игрока. Подчеркнем, что из непрерывности функции $f(x,u,v)$ на $\mathbb{R}^n \times P \times Q$ при любых парах стратегий (U,V) игроков получаем равномерную оценку

$$|x_N(U,V)| \leqslant c_1,$$

где c_1 — некоторая неотрицательная константа. Следовательно, в силу непрерывности $\varphi(x)$ на $\mathbb{R}^n$ имеем равномерную оценку при любых допустимых стратегиях U, V вида

$$|\varphi(x_N(U,V))| \leqslant c_2,$$

где c_2 — некоторая неотрицательная константа. Из непрерывности $\varphi(x)$ и сказанного вытекает, что величины $\alpha(U)$, β (см. (3), (4)) являются конечными.

Проблема вычисления величины β представляет большой интерес для теории многошаговых конфликтно управляемых процессов. Этой проблеме посвящен п. 2.

2. Нам понадобятся следующие леммы.

Лемма 1. *Пусть скалярная функция $g(x,u,v)$ определена и непрерывна на $\mathbb{R}^n \times P \times Q$, где P — компакт из $\mathbb{R}^p$, Q — компакт из $\mathbb{R}^q$. Тогда функция*

$$h(x,u) = \min_{v\in Q} g(x,u,v) \tag{5}$$

непрерывна на $\mathbb{R}^n \times P$.

Д о к а з а т е л ь с т в о леммы 1 при наложенных условиях следует из равномерной непрерывности функции, непрерывной на компакте. □

Лемма 2. *В условиях леммы* 1 *функция*

$$k(x) = \max_{u\in P} h(x,u) \tag{6}$$

определена и непрерывна на $\mathbb{R}^n$.

Д о к а з а т е л ь с т в о. Лемма 2 вытекает из леммы 1 и свойств рассматриваемых функций. □

Для нахождения величины β, определенной формулой (4), рассмотрим последовательность скалярных функций беллмановского типа на $\mathbb{R}^n$:

$$\begin{gathered} L_N(x) = \varphi(x), \\ L_t(x) = \max_{u\in P} \min_{v\in Q} (L_{t+1}(f(x,u,v))), \end{gathered} \tag{7}$$

где $t = N-1,\dots,0$. Отметим, что согласно леммам 1, 2 функции $L_{N-1}(x),\dots,L_0(x)$ при сделанных предположениях о $f(x,u,v)$, $\varphi(x)$ определены и непрерывны на всем $\mathbb{R}^n$.

Обозначим при $x \in \mathbb{R}^n$ через $\tilde u_t(x)$ $(t = 0,\dots,N-1)$ один из максимизаторов по $u \in P$ функции $\min\limits_{v\in Q}(L_{t+1}(f(x,u,v)))$, а через $\tilde v_t(x)$ $(t = 0,\dots,N-1)$ — один из минимизаторов по $v \in Q$ функции $L_{t+1}(f(x,\tilde u_t(x),v))$. Наборы функций $\tilde u_t(x)$, $\tilde v_t(x)$, $t = 0,\dots,N-1$, образуют стратегии $\tilde U$, $\tilde V$.

Теорема 1. *Для конфликтно управляемого процесса* (1), (2) *имеют место следующие равенства*

$$\beta = L_0(x_0), \quad \beta = \alpha(\tilde U),$$

где β, $L_0(x_0)$ и $\alpha(U)$ определены в (3), (4) *и* (7).

Д о к а з а т е л ь с т в о. Для доказательства равенства $\beta = L_0(x_0)$ можно использовать схему доказательства соответствующего утверждения в [3, с. 140–141]. При этом оказывается, что

$$\varphi(x_N(\tilde U,V)) \geqslant L_0(x_0) \tag{8}$$

для любой допустимой стратегии V, причем

$$\varphi(x_N(\tilde{U},\tilde{V})) = L_0(x_0) \tag{9}$$

и, далее,

$$\varphi(x_N(U,\hat{V}(U))) \leqslant L_0(x_0) \tag{10}$$

для любой допустимой стратегии U. Здесь стратегия

$$\hat{V}(U) = (\hat{v}_0(x),\dots,\hat{v}_{N-1}(x))$$

при $t = 0,\dots,N-1$ определяется как один из минимизаторов функции $L_{t+1}(f(x,u_t(x),v)$ по $v \in Q$, где $x \in \mathbb{R}^n$. Из равенства $\beta = L_0(x_0)$ и соотношений (8)–(10) вытекает равенство $\beta = \alpha(\tilde{U})$. $\square$

3. Приведем два примера вычисления величины β.

П р и м е р 1. Пусть многошаговый конфликтно управляемый процесс (1) имеет вид

$$x_{t+1} = Ax_t + Bu_t + Cv_t,$$

где $x_t \in \mathbb{R}^n$, $u_t \in P \subset \mathbb{R}^p$ $(p \geqslant 1)$, $v \in Q \subset \mathbb{R}^q$ $(q \geqslant 1)$, $t = 0,\dots,N-1$ $(N \geqslant 1)$, причем P и Q — компакты, A, B, C — матрицы размерности $n \times n$, $n \times p$, $n \times q$ соответственно. Пусть функция $\varphi(x)$ имеет вид $\varphi(x) = \langle a, x\rangle$, где $a \in \mathbb{R}^n$ — фиксированный вектор. Отметим, что для произвольных допустимых стратегий U, V справедлива формула

$$\begin{aligned} x_N(U,V) = A^N x_0 + A^{N-1}(Bu_0(x_0) + Cv_0(x_0)) + \dots \\ \dots + (Bu_{N-1}(x_{N-1}) + Cv_{N-1}(x_{N-1})), \end{aligned} \tag{11}$$

которая доказывается по индукции. Несложные рассуждения позволяют последовательно вычислить функции $L_t(x)$, $t = N-1,\dots,0$ вида (7), при $\varphi(x) = \langle a, x\rangle$ и в результате (с учетом (11)) получаем

$$\begin{aligned} L_0(x) = \langle a, A^N x\rangle + (\max_{u\in P}\langle a, A^{N-1}Bu\rangle + \min_{v\in Q}\langle a, A^{N-1}Cv\rangle) + \dots \\ \dots + (\max_{u\in P}\langle a, Bu\rangle + \min_{v\in Q}\langle a, Cv\rangle), \end{aligned}$$

т. е. $L_0(x) = \langle a, A^N x\rangle + \gamma$, где γ — некоторая эффективно вычислимая константа.

П р и м е р 2. Пусть конфликтно управляемый процесс (1) имеет вид

$$x_{t+1} = u_t x_t + v_t,$$

где $x_t \in \mathbb{R}^1$, $u_t \in P = [-1,1]$, $v_t \in Q = [-\mu,\mu]$ $(\mu > 0)$, $N = 2$. Пусть функция $\varphi(x) = x$. В этом примере

$$\begin{gathered} L_2(x) = x, \quad L_1(x) = |x| - \mu, \\ L_0(x) = \max_{u\in P}\min_{v\in Q}(|ux+v| - \mu). \end{gathered} \tag{12}$$

Нетрудно видеть, что при $|x| \geqslant \mu$ $L_0(x) = |x| - 2\mu$, а при $|x| \leqslant \mu$ $L_0(x) = -\mu$.

4. В этом пункте мы кратко остановимся на проблеме вычисления функций $L_t(x)$, $t = N-1,\dots,0$, при $N \geqslant 1$ (7). В п. 2 мы уже отмечали, что при сделанных ранее предположениях функции $L_N(x), L_{N-1}(x),\dots,L_0(x)$ определены и непрерывны на $\mathbb{R}^n$. Это обстоятельство открывает определенные возможности при приближенных вычислениях этих функций сеточным методом. Оказывается, что при некотором усилении требований на функции $\varphi(x)$, $f(x,u,v)$ можно гарантировать липшицевость функций $L_N(x), L_{N-1}(x),\dots,L_0(x)$ на $\mathbb{R}^n$ (см. далее теорему 2). Для доказательства теоремы 2 нам понадобятся дополнительные вспомогательные утверждения.

Лемма 3. *Пусть скалярная функция $g(x,u,v)$ определена и непрерывна на $\mathbb{R}^n \times P \times Q$, где P, Q — компакты из $\mathbb{R}^p$, $\mathbb{R}^q$ соответственно. Пусть функция $g(x,u,v)$ липшицева по $x \in \mathbb{R}^n$ равномерно относительно $u \in P$, $v \in Q$, т. е. существует такая константа $l \geqslant 0$, что при произвольных $x', x'' \in \mathbb{R}^n$ и произвольных $u \in P$, $v \in Q$ имеет место неравенство*

$$|g(x',u,v) - g(x'',u,v)| \leqslant l|x' - x''|. \tag{13}$$

Тогда функция $h(x,u)$, определенная в (5), является непрерывной на $\mathbb{R}^n \times P$ и липшицевой по $x \in \mathbb{R}^n$ равномерно относительно $u \in P$ с константой l.

Доказательство. Непрерывность функции $h(x,u)$ на $\mathbb{R}^n \times P$ была установлена в лемме 1. Пусть x', x'' — произвольные векторы из $\mathbb{R}^n$ и u — произвольный вектор из P. Имеем

$$h(x',u) - h(x'',u) = g(x',u,v(x',u)) - g(x'',u,v(x'',u)), \tag{14}$$

где $v(x',u)$ — один из минимизаторов функции $g(x',u,v)$ по $v \in Q$ и $v(x'',u)$ — один из минимизаторов функции $g(x'',u,v)$ по $v \in Q$. Из (13), (14) вытекают неравенства

$$h(x',u) - h(x'',u) \leqslant g(x',u,v(x'',u)) - g(x'',u,v(x'',u)) \leqslant l|x' - x''|. \tag{15}$$

Аналогичным образом получаем оценку

$$h(x'',u) - h(x',u) \leqslant l|x' - x''|. \tag{16}$$

Из (15), (16) следует утверждение леммы 3 с липшицевой константой l. □

Лемма 4. *Пусть выполнены условия леммы 3. Тогда функция $k(x)$, определенная в (6), удовлетворяет на $\mathbb{R}^n$ условию Липшица с константой l.*

Доказательство. Непрерывность функции $k(x)$ на $\mathbb{R}^n$ вытекает из леммы 1. Пусть x', x'' — произвольные векторы из $\mathbb{R}^n$. Имеем

$$k(x') - k(x'') = h(x',u(x')) - h(x'',u(x'')), \tag{17}$$

где $u(x')$ — один из максимизаторов функции $h(x',u)$ по $u \in P$, $u(x'')$ — один из максимизаторов функции $h(x'',u)$ по $u \in P$. Из (6), (17) следуют соотношения

$$k(x') - k(x'') \leqslant h(x',u(x')) - h(x'',u(x')) \leqslant l|x' - x''|. \tag{18}$$

Аналогично получаем соотношение

$$k(x'') - k(x') \leqslant l|x' - x''|. \tag{19}$$

Из (18), (19) вытекает искомая липшицевость $k(x)$ на $\mathbb{R}^1$ с константой l. □

Нетрудно доказать следующее утверждение.

Лемма 5. *Пусть скалярная функция $m(x)$ удовлетворяет условию Липшица на $\mathbb{R}^n$ с константой $l_1 \geqslant 0$. Пусть векторная функция $f(x,u,v)$ из (1) непрерывна на $\mathbb{R}^n \times P \times Q$ и удовлетворяет условию Липшица по $x \in \mathbb{R}^n$ равномерно относительно пары (u,v), где $u \in P$, $v \in Q$, с константой $l_2 \geqslant 0$. Тогда суперпозиция функций $m(f(x,u,v))$ непрерывна на $\mathbb{R}^n \times P \times Q$ и удовлетворяет условию Липшица по $x \in \mathbb{R}^n$ равномерно по (u,v), где $u \in P$, $v \in Q$, с константой $l_1 l_2$.*

Лемма 6. *Пусть скалярная функция $m(x)$ и векторная функция $f(x,u,v)$ удовлетворяют условиям леммы 5. Тогда функция*

$$p(x) = \max_{u \in P} \min_{v \in Q} m(f(x,u,v))$$

удовлетворяет на $\mathbb{R}^n$ условию Липшица с константой $l_1 l_2$.

Д о к а з а т е л ь с т в о леммы 6 при сделанных предположениях производится очевидным образом с помощью лемм 3–5. □

Справедлива следующая теорема.

Теорема 2. *Пусть скалярная функция $\varphi(x)$ удовлетворяет на $\mathbb{R}^n$ условию Липшица, а векторная функция $f(x,u,v)$ непрерывна на $\mathbb{R}^n \times P \times Q$ и удовлетворяет там условию Липшица по $x \in \mathbb{R}^n$ равномерно относительно пары (u,v), где $u \in P$, $v \in Q$. Тогда функции $L_0(x)$, $L_1(x)$, …, $L_N(x)$ (см. (7)) удовлетворяют условию Липшица на $\mathbb{R}^n$.*

Д о к а з а т е л ь с т в о Используя формулы (7) и лемму 6, нетрудно обосновать липшицевость функции $L_{N-1}(x)$ на $\mathbb{R}^n$. При $N > 1$ липшицевость функций $L_{N-2}(x)$,…, $L_0(x)$ на $\mathbb{R}^n$ доказывается последовательно с помощью формул (7) и математической индукции. □

СПИСОК ЛИТЕРАТУРЫ

1. **Петросян Л.А., Зенкевич Н.А., Семина Е.А.** Теория игр. М.: Изд-во Высшая школа, 1998. 302 с.
2. **Жуковский В.И., Салуквадзе М.Е.** Оценка рисков и многошаговые позиционные конфликты. М.: Юрайт, 2018. 304 с.
3. **Горелов М.А., Кононенко А.Ф.** Динамические модели конфликтов. I. Язык моделирования // Автоматика и телемеханика. 2014. Вып. 11. С. 127–149.
4. **Красовский Н.Н., Субботин А.И.** Позиционные дифференциальные игры. М.: Наука, 1974. 456 с.

Поступила 4.11.2019
После доработки 5.02.2020
Принята к публикации 10.02.2020

Никольский Михаил Сергеевич
д-р физ.-мат. наук, ведущий науч. сотрудник
Математический институт им. В. А. Стеклова РАН
г. Москва
e-mail: mni@mi-ras.ru

REFERENCES

1. Petrosyan L.A., Zenkevich N.A., Semina E.A. *Teoriya igr* [Game theory]. Moscow: Vysshaya Shkola Publ., 1998, 302 p. ISBN: 5-06-001005-8 .
2. Zhukovskii V.I., Salukvadze M.E. *Otsenka riskov i mnogoshagovye pozitsionnye konflikty* [Estimations of risks and multistep position conflicts]. Moscow: Urait Publ., 2018, 304 p. ISBN: 978-5-534-08782-6 .
3. Gorelov M.A., Kononenko A.F. Dynamical conflict models. I. Language of modeling. *Autom. Remote Control*, 2014, vol. 75, no. 11, pp. 1996–2013. doi: 10.1134/S0005117914110083 .
4. Krasovskii N.N., Subbotin A.I. *Game-theoretical control problems.* N Y: Springer, 1988, 517 p. ISBN: 978-1-4612-8318-8 . Original Russian text published in Krasovskii N.N., Subbotin A.I. *Pozitsionnye differentsial'nye igry.* Moscow: Nauka Publ., 1974, 456 p.

Received November 4, 2019
Revised February 5, 2020
Accepted February 10, 2020

Mikhail Sergeevich Nikolskii, Dr. Phys.-Math. Sci, Steklov Mathematical Institute of the Russian Academy of Science, Moscow, 119991 Russia, e-mail: mni@mi-ras.ru .

Cite this article as: M. S. Nikolskii. About maximal guaranteed payoff in some problems of conflict control by discrete processes, *Trudy Instituta Matematiki i Mekhaniki URO RAN*, 2020, vol. 26, no. 1, pp. 167–172 .