

УДК 517.91

МАТЕМАТИКА

А. И. СУББОТИН

ЭКСТРЕМАЛЬНЫЕ СТРАТЕГИИ В ДИФФЕРЕНЦИАЛЬНЫХ ИГРАХ С ПОЛНОЙ ПАМЯТЬЮ

(Представлено академиком Н. Н. Красовским 22 II 1972)

Рассматриваются позиционные дифференциальные игры, в качестве платы выбран некоторый полунепрерывный снизу (либо сверху) функционал, вычисляемый на движениях конфликтно управляемой системы. Описаны конструкции, доказывающие существование оптимальных чистых и смешанных стратегий игроков. Предполагается, что игроки располагают информацией о реализующемся фазовом положении и помнят предыдущие состояния системы. Основу подхода, который используется здесь при исследовании дифференциальных игр с полной памятью, составляет экстремальная конструкция, описанная в работах ⁽¹⁻³⁾.

Пусть n -мерный фазовый вектор x конфликтно управляемой системы Σ изменяется в соответствии с уравнением

$$dx/dt = f(t, x, u, v), \quad x(t_0) = x_0; \quad (1)$$

здесь $f(t, x, u, v)$ — непрерывная вектор-функция, удовлетворяющая условию Липшица по x ; u, v — управляющие воздействия первого и второго игроков, реализации которых стеснены условиями

$$u[t] \in P, \quad v[t] \in Q, \quad (2)$$

где P и Q — компакты в соответствующих векторных пространствах.

Исход игры оценивается функционалом $\gamma(x[\cdot])$, который определен на движениях $x[t]$, $t_0 \leq t \leq T$. Будем полагать, что плата игры — функционал $\gamma(x[\cdot])$ зависит от вектор-функций $x[t]$, $t_0 \leq t \leq T$, полунепрерывно снизу, т. е. из условий $x_k[\cdot] \rightrightarrows x_*[\cdot]$, $\gamma(x_k[\cdot]) \rightarrow \gamma$, при $k \rightarrow \infty$ вытекает неравенство $\gamma(x_*[\cdot]) \leq \gamma$. Здесь символ $x_k[\cdot] \rightrightarrows x_*[\cdot]$ обозначает равномерную сходимость $\max\{\|x_k[t] - x_*[t]\| : t \in [t_0, T]\} \rightarrow 0$ при $k \rightarrow \infty$, символом $\|x\|$ обозначается евклидова норма вектора x . (Условие полунепрерывности снизу можно заменить условием полунепрерывности сверху, последний случай легко сводится к первому.) Отметим, что условию полунепрерывности удовлетворяет плата игры в типичных игровых задачах динамики, в частности, в игре преследования время до встречи зависит от движений системы полунепрерывно снизу.

Как обычно, будем считать, что первый игрок заинтересован в исходе игры, при котором функционал $\gamma(x[\cdot])$ принимает наименьшее возможное значение. Второй игрок, напротив, стремится максимизировать значение платы при самом неблагоприятном поведении партнера. Причем каждому из игроков наперед неизвестны управляющие воздействия, которые будут реализованы партнером в процессе игры. При постановке многих игровых задач динамики предполагается, что игрокам в каждый текущий момент времени $t \geq t_0$ становится известным реализованное значение вектора $x[t]$. Такое предположение об информированности игроков позволяет в этих задачах определить полный класс позиционных стратегий игроков, который содержит оптимальные стратегии, доставляющие ситуации типа седловой точки ⁽⁴⁾. В данном случае, когда в качестве платы

выбран произвольный полунепрерывный функционал $\gamma(x[\cdot])$, этой информации о состоянии системы, вообще говоря, недостаточно для построения полного класса стратегий. Поэтому в рассматриваемых здесь игровых задачах вводится следующее предположение об информированности игроков. В каждый текущий момент времени $t \geq t_0$ игрокам известно не только реализованное состояние системы Σ — вектор $x[t]$, но игроки помнят все предыдущие состояния $x[\tau]$, $t_0 \leq \tau < t$.

Итак, позицией игры в момент времени $t \geq t_0$ будем называть совокупность, состоящую из скаляра t и вектор-функции $x[\tau]$, $t_0 \leq \tau \leq t$, — движения системы Σ , реализовавшего на промежутке времени $[t_0, t]$; эти движения обозначаются в дальнейшем символом $x[\cdot; t_0, t]$.

Введем теперь некоторые обозначения и определения. Обозначим через $R(t_0, T)$ компактное подмножество пространства непрерывных вектор-функций $x[\tau]$, $t_0 \leq \tau \leq T$, которое содержит всякое движение системы Σ , выходящее из точки $\{x_0, t_0\}$. Такое множество всегда можно указать. Символом $R(t_0, t)$ обозначим множество, которое составляют вектор-функции $x[\tau]$ из $R(t_0, T)$, рассматриваемые на отрезке $[t_0, t]$, $t \leq T$.

Определение 1. Стратегией первого (второго) игрока будем называть функцию $U = U(t, x[\cdot; t_0, t])$ ($V = V(t, x[\cdot; t_0, t])$), которая точкам $\{t, x[\cdot; t_0, t]\}$, $t \in [t_0, T]$, $x[\cdot; t_0, t] \in R(t_0, t)$, ставит в соответствие непустые замкнутые множества $U(t, x[\cdot; t_0, t]) \subset P(V(t, x[\cdot; t_0, t]) \subset Q)$.

Определение 2. Пусть U — некоторая стратегия первого игрока, вектор-функция $x_*[\cdot; t_0, t_*]$ содержится во множестве $R(t_0, t_*)$, совокупность Γ полуинтервалов (τ_i, τ_{i+1}) , $\tau_0 = t_*$, $i = 0, 1, 2, \dots$, покрывает отрезок $[t_*, T]$. Аппроксимационным движением $x_a[t] = x_a[t; t_*, x_*[\cdot; t_0, t_*], U, \Gamma]$ системы Σ , порожденным на отрезке $[t_*, T]$ стратегией U , будем называть всякую абсолютно непрерывную вектор-функцию $x_a[t]$, которая удовлетворяет условиям

$$\begin{aligned} dx_a[t]/dt &\in F(t, x_a[t], u[\tau_i]), \quad \tau_i \leq t < \tau_{i+1}, \quad i = 0, 1, 2, \dots, \\ u[\tau_i] &\in U(\tau_i, x_a[\cdot; t_0, \tau_i]), \quad F(t, x, u) = \text{co}\{f(t, x, u, v) : v \in Q\}, \\ x_a[\cdot; t_0, \tau_i] &= \{x_*[\tau], \quad t_0 \leq \tau \leq t_*; \quad x_a[\tau], \quad t_* \leq \tau \leq \tau_i\}. \end{aligned}$$

Отметим, что вектор-функция $x_*[\cdot; t_0, t_*]$ играет роль начального условия для движения $x_a[t]$, $t \geq t_*$.

Определение 3. Движением системы Σ $x[t] = x[t; t_*, x_*[\cdot; t_0, t_*], U]$, порожденным на отрезке $[t_*, T]$ стратегией U и удовлетворяющим начальному условию $x[\cdot; t_0, t_*] = x_*[\cdot; t_0, t_*] \in R(t_0, t_*)$, будем называть всякую абсолютно непрерывную вектор-функцию $x[t]$, $t_* \leq t \leq T$, для которой можно указать последовательность аппроксимационных движений $x_a^{(k)}[t] = x_a[t; t_*, x_*^{(k)}[\cdot; t_0, t_*], U, \Gamma_k]$, $t_* \leq t \leq T$, такую, что

$$\begin{aligned} x_a^{(k)}[t] &\rightrightarrows x[t], \quad t_* \leq t \leq T, \\ x_*^{(k)}[\cdot; t_0, t_*] &\rightrightarrows x_*[\cdot; t_0, t_*], \quad \Delta(\Gamma_k) \rightarrow 0 \quad \text{при } k \rightarrow \infty, \end{aligned}$$

где $\Delta(\Gamma) = \sup(\tau_{i+1} - \tau_i)$, $i = 0, 1, 2, \dots$.

Аналогичным образом определяются движения системы Σ , порожденные стратегиями V второго игрока.

Отметим, что совокупность движений $x[t; t_*, x_*[\cdot; t_0, t_*], U]$ ($x[t; t_*, x_*[\cdot; t_0, t_*], V]$) содержит всякое движение системы Σ , которое порождается стратегией U (V) в паре с любой суммируемой реализацией $v[t] \in Q$ ($u[t] \in P$) управления второго (первого) игрока. Поэтому выполнение некоторого свойства для всех движений $x[t; t_*, x_*[\cdot; t_0, t_*], U]$ ($x[t; t_*, x_*[\cdot; t_0, t_*], V]$) будет означать, что стратегия U (V) гарантирует осуществление этого свойства при любом допустимом поведении противника.

Перейдем теперь к описанию экстремальной конструкции для рассматриваемого здесь случая дифференциальных игр с полной памятью. Будем сначала предполагать, что для точек $\{t, x\} \in [t_0, T] \times E_n$ и для любых векторов $s \in E_n$ выполняется равенство

$$\max_u \min_v s'f(t, x, u, v) = \min_v \max_u s'f(t, x, u, v), \quad u \in P, \quad v \in Q, \quad (3)$$

где символ штрих означает транспонирование.

Определение 4. Пусть задана некоторая система непустых замкнутых множеств $W(t) \subset R(t_0, t)$, $t_0 \leq t \leq T$. Эту систему будем называть *и-стабильной*, если, каковы бы ни были числа $t_0 \leq t_* \leq T$, вектор-функция $x_*[\cdot; t_0, t_*] \in W(t_*)$ и стратегия V , существует движение $x_*[t] = x_*[t, t_*, x_*[\cdot; t_0, t_*], V]$ такое, что вектор-функция $\tilde{x}[\cdot; t_0, t^*) = \{x_*[\tau], t_0 \leq \tau \leq t_*; x^*[\tau], t_* \leq \tau \leq t^*\}$ содержится во множестве $W(t^*)$.

Рассмотрим неоднозначную функцию $U_e = U_e(t, x[\cdot; t_0, t])$ следующего вида. Если $x[\cdot; t_0, t] \in W(t)$, то $U_e(t, x[\cdot; t_0, t]) = P$. Если $x[\cdot; t_0, t] \notin W(t)$, то множество $U_e(t, x[\cdot; t_0, t])$ составляют точки u_e , которые удовлетворяют условию максимума

$$\max_u \min_v s'f(t, x, u, v) = \min_v s'f(t, x, u_e, v), \quad (4)$$

$$u \in P, \quad v \in Q.$$

хотя бы при одном из векторов s вида $s = w_0[t] - x[t]$, где $w_0[\cdot; t_0, t]$ — точки множества $W(t)$, ближайšie к точке $x[\cdot; t_0, t]$.

Лемма 1. Если $x_*[\cdot; t_0, t_*] \in W(t_*)$, система $W(t)$, $t_0 \leq t \leq T$, и-стабильна и выполняется условие (3), то для любого движения $x[t] = x[t; t_*, x_*[\cdot; t_0, t_*], U_e]$ выполняется включение $\tilde{x}[\cdot; t_0, T] \in W(T)$, где $\tilde{x}[\tau] = \{x_*[\tau], t_0 \leq \tau \leq t_*; x[\tau], t_* \leq \tau \leq T\}$.

Основу доказательства этого утверждения составляют оценки, приведенные в работе ⁽¹⁾, стр 1013.

Определение 5. Множеством поглощения $W(t_*, c)$, $t_0 \leq t_* \leq T$, назовем совокупность элементов $x_*[\cdot; t_0, t_*] \in R(t_0, t_*)$, удовлетворяющих условию: для любой стратегии V найдется движение $x[t] = x[t; t_*, x_*[\cdot; t_0, t_*], V]$ такое, что для вектор-функции $\tilde{x}[\tau] = \{x_*[\tau], t_0 \leq \tau \leq t_*; x[\tau], t_* \leq \tau \leq T\}$ будет справедливо неравенство $\gamma(\tilde{x}[\cdot]) \leq c$.

Используя подход, изложенный в работах ^(1, 4), можно показать, что для числа c_0 , определенного соотношением

$$c_0 = \inf \{c: x_0 \in W(t_0, c)\}, \quad (5)$$

справедливо включение $x_0 \in W(t_0, c_0)$, причем множества $W(t, c_0)$, $t_0 \leq t \leq T$, замкнуты и непусты, система этих множеств и-стабильна. Поэтому из определения 5 в силу леммы 1 вытекает следующее утверждение.

Теорема 1. Если выполняется условие (3), то рассматриваемая дифференциальная игра с полной памятью в классе стратегий $\{U\}$ и $\{V\}$ имеет значение, равное числу c_0 (5). Причем стратегия U_e , экстремальная к системе множеств $W(t, c_0)$, $t_0 \leq t \leq T$, гарантирует завершение игры с платой $\gamma(x[\cdot]) \leq c_0$. Для любого числа $c < c_0$ существует стратегия V_c , которая обеспечивает выполнение неравенства $\gamma(x[\cdot]) > c$.

Пусть теперь условие (3) не выполняется. В этом случае значение игры будет существовать в классе смешанных стратегий $\{\tilde{U}\}$ и $\{\tilde{V}\}$. Эти стратегии ставят в соответствие точкам $\{t, x[\cdot; t_0, t]\} \in [t_0, T] \times R(t_0, t)$ непустые слабо замкнутые множества вероятностных мер $\{\mu(du)\}$ и $\{\nu(dv)\}$, заданных на компактах P и Q . При определении движений множества $F(t, x_a[t], u[\tau], \nu[\tau])$ следует заменить $\tilde{F}(t; x_a[t], \mu(du; \tau))$, где

$$\mu(du; \tau_i) \in \mathcal{U}(\tau_i, x_i[\cdot; t_0, \tau_i]),$$

$$F(t, x, \mu) = \text{co} \left\{ \int f(t, x, u, v) \mu(du): v \in Q \right\}.$$

При определении экстремальной смешанной стратегии условие максимума (4) заменяется соотношением

$$\max_{\mu} \min_v \iint s'f(t, x, u, v) \mu(du) v(dv) = \min_v \iint s'f(t, x, u, v) \mu_0(du) v(dv).$$

В определениях 4 и 5 и в формулировке леммы 1 чистые стратегии следует заменить смешанными.

Теорема 2. В дифференциальной игре с полной памятью в классе смешанных стратегий $\{U\}$, $\{V\}$ существует значение, равное числу $\tilde{c}_0 = \inf \{c: x_0 \in W(t_0, c)\}$, где $W(t, c)$ — множество поглощения, определенное для смешанных стратегий. Причем смешанная стратегия $\tilde{U}_c$, экстремальная к системе множеств $\tilde{W}(t, \tilde{c}_0)$, $t_0 \leq t \leq T$, гарантирует завершение игры с платой $\gamma(x[\cdot]) \leq \tilde{c}_0$. Для любого числа $c < \tilde{c}_0$ существует смешанная стратегия $\tilde{V}_c$, которая обеспечивает выполнение неравенства $\gamma(x[\cdot]) > c$.

Наконец, в соответствии с подходом, развитым в статье ⁽²⁾, можно решить задачу о нахождении минимаксных и максиминных чистых стратегий в случае, когда условие (3) не выполняется. Устанавливается, что класс чистых стратегий $\{U\}$, $\{V\}$ является полным, т. е. стратегии из этих классов доставляют игрокам наилучшие результаты, которых можно достичь при выборе чистых позиционных стратегий.

Результаты данной работы можно перенести на случай, когда движение системы Σ описывается уравнением

$$dx[t]/dt = f(t, x[\cdot; \eta, t], u, v), \quad t \geq t_0,$$

и удовлетворяет начальному условию $x[\tau] = \varphi(\tau)$ при $\eta \leq \tau \leq t_0$. Здесь оператор f отображает множество $D = \{t \in [t_0, T]\} \times R(\eta, t) \times P \times Q$ в пространство E_n и удовлетворяет некоторым условиям непрерывности. Основные элементы экстремальной конструкции и приведенные выше результаты легко переформулируются для этого случая.

Примечание. Множества $W(t, c)$ здесь было удобно определить как множества поглощения (см. определение 5). Оказывается, что эти множества можно также определять с помощью понятных процедур, которые подобны конструкциям, описанным в работах ⁽⁵⁻⁷⁾ при исследовании дифференциальных игр преследования.

Институт математики и механики
Уральского научного центра Академии наук СССР
Свердловск

Поступило
25 XII 1971

ЦИТИРОВАННАЯ ЛИТЕРАТУРА

- ¹ Н. Н. Красовский, А. И. Субботин, ПММ, 34, № 6, 1005 (1970).
² Н. Н. Красовский, ДАН, 201, № 2, 270 (1971). ³ Н. Н. Красовский, Ю. С. Осипов, ДАН, 197, № 4, 777 (1971). ⁴ Н. Н. Красовский, А. И. Субботин, ПММ, 35, № 1, 110 (1971). ⁵ Л. С. Понтрягин, ДАН, 175, № 4, 764 (1967).
⁶ Б. Н. Пшеничный, ДАН, 184, № 2, 285 (1969). ⁷ Р. П. Федоренко, Журн. вычислит. матем. и матем. физ., 9, № 5, 1036 (1969).