Прикладная математика и механика. Том 33, 1969

ОБ ОПТИМАЛЬНЫХ СТРАТЕГИЯХ В ЛИНЕЙНОЙ ДИФФЕРЕНЦИАЛЬНОЙ ИГРЕ

Н. Н. Красовский, А. И. Субботин

(Свердловск)

Для игровой задачи о приведении на заданное множество управляемого объекта, движение которого описывается линейными дифференциальными уравнениями, выводятся условия, при которых в рассматриваемой дифференциальной игре существует седловая точка в классе обобщенных стратегий. Предлагается конструкция обобщенных оптимальных стратегий игроков.

§ 1. Рассматривается игровая задача о приведении на заданное множество M управляемого объекта, движение которого описывается системой дифференциальных уравнений

$$\frac{dx}{dt} = A(t)x + B(t)u - C(t)v \tag{1.1}$$

Здесь $x = \{x_1, \ldots, x_n\}$ — n-мерный фазовый вектор, u — r-мерное управляющее воздействие, выбором которого распоряжается первый игрок; v — s-мерное управление второго игрока; $A(t)$, $B(t)$, $C(t)$ — непрерывные матрицы соответствующих размерностей. Предполагается, что на мгновенные значения управляющих воздействий наложены ограничения вида

$$u \in U^*, \qquad v \in V^* \tag{1.2}$$

где U^*, V^* — выпуклые, ограниченные, замкнутые множества в пространствах E_r и E_s соответственно. Относительно множества M ниже предполагается, что выполнено одно из следующих условий:

1) множестве M — выпуклое, ограниченное, множество в E_n;

2) множество M — линейное подпространство пространства E_n.

Платой в рассматриваемой игре будет величина ϑ — момент попадания фазового вектора $x[t]$ на заданное множество M. При этом первый игрок стремится минимизировать величину ϑ, целью второго игрока является максимизация момента ϑ.

Подобные игровые задачи рассматривались в работах [1-11], элементы конструкций, предложенных в этих статьях, используются в данной работе.

Будем придерживаться следующих определений стратегий игроков и решения системы (1.1).

Определение 1.1. Стратегией или управлением $U = U(x, t)$ ($V = V(x,t)$) называется закон, по которому $n+1$-мерному вектору $\{x, t\}$ ставится в соответствие выпуклое замкнутое множество $U(x, t) \subset U^*$

($V(x, t) \subset V^*$). Ниже предполагается, что стратегии полунепрерывны сверху по включению [1].

Определение 1.2. Решением системы (1.1), порожденным некоторой парой стратегий $U = U(x, t)$, $V = V(x, t)$, будем называть всякую абсолютно непрерывную вектор-функцию $x = x[t]$, удовлетворяющую условиям

$$x[t_0] = x^\circ, \quad \frac{dx[t]}{dt} \in A(t)x[t] + B(t)U(x[t], t) - C(t)V(x[t], t) \tag{1.3}$$

почти для всех $t \geqslant t_0$.

Здесь t_0 — начальный момент времени, x^0 — исходное значение фазового вектора, $Ax + BU - CV$ — алгебраическая сумма вектора Ax и множеств BU и $-CV$, т. е. совокупность всех векторов вида $Ax + Bu - Cv$, где $u \in U$, $v \in V$.

Заметим, что в определении 1.1 отражен характер информации, которой располагают игроки в процессе игры, а именно, в каждый момент времени $t = \tau \geqslant t_0$ каждый из игроков знает реализовавшееся значение фазового вектора $x = x[\tau]$, при этом игрокам неизвестно будущее поведение противника при $t > \tau$. Существование решения системы (1.1) в смысле определения 1.2 можно доказать переходом к пределу от соответствующих ломаных Эйлера (см., например, [10–12]).

В предлагаемой работе описывается построение стратегии $V = V_\alpha(x,t)$, гарантирующей второму игроку результат, сколь угодно близкий к наилучшему. Формулировка условий, используемых при построении стратегии $V_\alpha(x, t)$, приводится ниже в § 3.

§ 2. В этом параграфе приводится вспомогательная теорема, используемая в § 3 при построении стратегии $V = V_\alpha(x, t)$.

Пусть задан некоторый промежуток времени $[t_0, \varkappa]$, в течение которого второй игрок рассчитывает избежать попадания точки $x[t]$ на множества M. Обозначим через G ограниченное множество в E_n такое, что для решения $x = x[t]$ системы (1.1), отвечающего любым суммируемым управлениям $u = u(t)$, $v = v(t)$, удовлетворяющим ограничениям (1.2), выполняется условие $x[t] \in G$ при $t_0 \leqslant t \leqslant \varkappa$. Такое множество G всегда можно указать.

Предположим, что существует функция $L = L(x, t)$, удовлетворяющая следующим требованиям.

(1). Функция $L = L(x, t)$ определена и непрерывно дифференцируема по x и t в некоторой открытой области D, причем точка $\{x^0, t_0\} \in D$.

(2). Если $\{x, t\} \in D$, то $x \overline{\in} M$.

(3). Пусть $\Gamma_{a,b}$ — множество точек $\{x, t\}$, для которых выполняются следующие условия: $x \in G$, $t_0 \leqslant t \leqslant \varkappa - a$, $L(x, t) \leqslant b$. Третье требование, накладываемое на функцию $L = L(x, t)$, состоит в следующем: для любых значений параметров $a > 0$, $b \geqslant L(x^0, t_0)$ можно указать

[1] Множество $U(x, t)$ полунепрерывно сверху по включению в точке $\{x^*, t^*\}$, если для любого $\varepsilon > 0$ найдется такое $\delta > 0$, что для всякого вектора $u \in U(x, t)$, где $\|\{x,t\} - \{x^*, t^*\}\| \leqslant \delta$ найдется вектор $u^* \in U(x^*, t^*)$ такой, что $\|u - u^*\| \leqslant \varepsilon$.

такое число $\varepsilon(a, b) > 0$, что $\Gamma_{a,b}^{\varepsilon(a,b)} \subset D$, где Γ^ε — ε-окрестность множества Γ (т. е. множество точек вида $q_1 + q_2$, где $q_1 \in \Gamma$, $\|q_2\| \leqslant \varepsilon$, символ $\|q\|$ означает евклидову норму вектора q).

(4). Рассмотрим величину

$$\Phi(x, t, u, v) = (\operatorname{grad}_x L(x, t))' \ A(t)x + B(t)u - C(t)v) + \frac{\partial L(x, t)}{\partial t} \tag{2.1}$$

(символ штрих означает транспонирование), определенную для точек $\{x, t\} \in D$. Поскольку $\Phi(x, t, u, v)$ можно представить в виде суммы двух слагаемых, первое из которых зависит лишь от u, а второе — от v, то справедливо равенство

$$\min_v = \max_u \Phi(x, t, u, v) = \max_u \min_v \Phi(x, t, u, v) = \Phi^0(x, t)$$
$$(u \in U^*, \ v \in V^*) \tag{2.2}$$

Последнее требование, накладываемое на $L(x, t)$, состоит в том что должно выполняться неравенство

$$\Phi^0(x, t) \leqslant c, \text{ при } \{x, t\} \in D \cap (G \times [t_0, \varkappa]) \quad (c = \text{const}) \tag{2.3}$$

Обозначим через $V^0(x, t)$ множество векторов, доставляющих

$$\max_v (\operatorname{grad}_x L(x, t))' C(t) v \ \text{ при } \ v \in V^* \tag{2.4}$$

(в случае $(\operatorname{grad}_x L(x, t))'C(t) = 0$ полагаем $V^0(x, t) = V^*$). Можно проверить, что при $\{x, t\} \in D$ стратегия $V^0(x, t)$ удовлетворяет определению 1.1.

Справедливо следующее утверждение.

Теорема 2.1. Если существует функция $L = L(x, t)$, удовлетворяющая требованиям (1) — (4), то любое решение системы

$$dx/dt \in A(t)x + B(t)U(x, t) - C(t)V^0(x, t), \ x[t_0] = x^0 \tag{2.5}$$

где $U(x, t)$ — произвольная допустимая стратегия, определенная при $\{x, t\} \in G \times [t_0, \varkappa]$, продолжимо до момента времени $t = \varkappa$, причем для любого решения системы (2.5) имеет место условие $x[t] \overline{\in} M$ при всех $t_0 \leqslant t < \varkappa$.

Доказательство. Известно, что решение $x = x[t]$ системы

$$dx / dt \in A(t)x + W(x, t), \qquad x[t_0] = x^0 \tag{2.6}$$

продолжимо до момента времени $t = t^*$, пока для этого решения выполняется условие $\{x[t], t\} \in N$ при всех $t_0 \leqslant t < t^*$, где N — открытое множество точек $\{x, t\}$, в котором $W(x, t)$ полунепрерывно сверху по включению, выпукло, ограничено и замкнуто. Поэтому для того, чтобы доказать, что любое решение системы (2.5) продолжимо до момента $t = \varkappa$, достаточно показать, что для всякого решения $x = x[t]$ системы (2.5) выполняется условие

$$\{x[t], t\} \in D \text{ при } t_0 \leqslant t < \varkappa \tag{2.7}$$

Предположим противное. Пусть $t = t_* < \varkappa$ — момент времени, когда для некоторого решения $x = x[t]$ системы (2.5) имеет место соотношение

$$\lim_{t \to t_* - 0} \{x[t], t\} = \{x^*, t_*\} \overline{\in} D \tag{2.8}$$

Заметим, что для рассматриваемого решения выполняется включение

$$\{x[t], t\} \in D \ \text{ при } t_0 \leqslant t < t_* \tag{2.9}$$

кроме того, вдоль рассматриваемого решения для абсолютно непрерывной функции $L[t] = L(x[t], t)$ в силу (2.1) — (2.4) при $t < t_*$ имеет место неравенство

$$dL[t] / dt \leqslant c \tag{2.10}$$

Отсюда вытекает

$$L(x[t], t) \leqslant L(x^0, t_0) + c(t - t_0) \text{ при } t < t_* \tag{2.11}$$

Полагая

$$a = \varkappa - t_*, \quad b = L(x^0, t_0) + c(\varkappa - t_0) \tag{2.12}$$

определим в силу условия (3) соответствующее число $\varepsilon = \varepsilon(a, b) > 0$. Пусть $t' < t_*$ — момент времени такой, что

$$\| \xi[t'] - \xi^* \| \leqslant {}^1/_2\, \varepsilon \tag{2.13}$$

где $\xi[t'] = \{x[t'], t'\}$, $\xi^* = \{x^*, t_*\}$. Поскольку $t' < t_*$, то в силу (2.11), (2.12) из условия (3) имеем

$$S_\varepsilon(\xi[t']) \subset D \tag{2.14}$$

где $S_\varepsilon(\xi)$ — ε-окрестность точки ξ. Из (2.14), (2.13) вытекает, что $S_{\varepsilon/2}(\xi^*) \subset D$, что противоречит предположению (2.8). Таким образом, справедливость условия (2.7) установлена. Из (2.7) в силу (2) вытекает, что для любого решения системы (2.5) имеет место условие $x[t] \in M$ при $t_0 \leqslant t < \varkappa$. Итак, утверждения теоремы 2.1 доказаны.

§ 3. Сформулируем условия, при выполнении которых функцию $L = L(x, t)$ можно построить так, чтобы управление $V = V^0(x, t)$, конструкция которого описана в § 2, обеспечивало второму игроку результат, сколь угодно близкий к наилучшему. Некоторые замечания относительно класса задач (1.1), (1.2), удовлетворяющих сформулированным ниже условиям, будут сделаны в конце этого параграфа.

При построении функции $L = L(x, t)$ здесь используется величина $\varepsilon^0(x, t, \sigma)$, которая определяется следующим образом.

Пусть $G_1(\tau, \sigma)$ — множество точек $g \in E_n$, для каждой из которых существует допустимое программное управление $u = u(t)$, $\tau \leqslant t \leqslant \sigma$, переводящее систему

$$dx/dt = A(t)x + B(t)u \tag{3.1}$$

Из состояния $x(\tau) = 0$ в положение $x(\sigma) = g$. Здесь допустимым программным управлением называется суммируемая вектор-функция $u = u(t)$, $\tau \leqslant t \leqslant \sigma$, почти всюду удовлетворяющая условию (1.2). Множество $G_1(\tau, \sigma)$ называется областью достижимости объекта (3.1), отвечающей начальному и конечному моментам τ и σ соответственно.

Аналогичным образом вводится понятие области достижимости $G_2(\tau, \sigma)$ управляемого объекта, движение которого описывается системой

$$dx/dt = A(t)x + C(t)v \tag{3.2}$$

Пусть M^ε — замкнутая ε-окрестность множества M, т. е. множество точек вида $g_1 + g_2$, где $g_1 \in M$, $\| q_2 \| \leqslant \varepsilon$; $X(\sigma, t)$ — $n \times n$ — матрица, удовлетворяющая уравнению $\partial X(\sigma, t) / \partial\sigma = A(\sigma) X(\sigma, t)$;

$X(t, t) = E$ — единичная матрица. Определим $\varepsilon^0(x, t, \sigma)$ как наименьшее значение параметра ε, при котором имеет место включение

$$G_2(\tau, \sigma) \subset G_1(\tau, \sigma) + X(\sigma, \tau)x - M^\varepsilon \tag{3.3}$$

Здесь, как и выше, выражение, стоящее справа, будет алгебраической суммой множеств $G_1, -M^\varepsilon$ и вектора Xx. Множество $G(x, \tau, \sigma) = G_1(\tau, \sigma) + X(\sigma, \tau)x - M^\varepsilon$ можно интерпретировать как область достижимости, построенную для конечного момента $\sigma \geqslant \tau$, управляемого объекта, движение которого описывается системой

$$dx^{(1)}/dt = A(t)\, x^{(1)} + B(t)\, u(t) - p\delta(\sigma - t), \quad x(\tau) = x \tag{3.4}$$

Здесь, наряду с допустимым программным управлением $u = u(t)$, вводится фиктивное импульсное управление вида $p\delta(\sigma - t)$, где $p \in M^{\varepsilon}$. Тогда величина $\varepsilon^0(x, t, \sigma)$ может быть определена здесь так же, как и в [10], как наименьшее значение параметра ε, при котором имеет место ε-поглощение одного управляемого процесса другим (в данном случае процесс (3.4) поглощает процесс (3.2)). При этом исходную задачу можно рассматривать как задачу об игровой встрече управляемых объектов (3.4), (3.2), где уравнение (3.4) описывает движение преследователя, а уравнение (3.2) — преследуемого игрока, начальное положение преследователя $x^{(1)}[t_0] = x^0$, преследуемого — $x[t_0] = 0$.

Можно показать, что в случае, когда M — выпуклое, ограниченное, замкнутое множество, условие (3.3) эквивалентно неравенству

$$\min_l \{\rho_1(l, \tau, \sigma) + \rho(l) + (X(\sigma, \tau)x)'\, l + \varepsilon \| l \| - \rho_2(l, \tau, \sigma)\} \geqslant 0$$

$$\text{при } \|l\| \leqslant 1 \tag{3.5}$$

где ρ_1, ρ_2, ρ — опорные функции множеств G_1, G_2,— M соответственно [1].

В случае, когда M — линейное подпространство пространства E_n, условие (3.3) равносильно неравенству

$$\min_l \{\rho_1^*(l, \tau, \sigma) + \varepsilon \| l \| + l'(X(\sigma, \tau)\, x)^* - \rho^*_2(l, \tau, \sigma)\} \geqslant 0 \tag{3.6}$$

$$\text{при } \| l \| \leqslant 1$$

Здесь ρ_1^*, ρ_2^* — опорные функции множеств G_1^*, G_2^*-проекций множеств G_1, G_2 на ортогональное дополнение подпространства M, обозначаемое в дальнейшем через Q; g^* — проекция n-мерного вектора g на подпространство Q; размерность вектора l совпадает с размерностью подпространства Q.

Требования, накладываемые на рассматриваемую систему (1.1), (1.2), формулируются следующим образом.

Условие 3.1. При всех допустимых значениях x, τ, σ, удовлетворяющих неравенству $\varepsilon^0(x, \tau, \sigma) > 0$, минимум в (3.5) (или (3.6)) достигается на единственном векторе $l = l^0(x, \tau, \sigma)$.

Определим функцию $L = L(x, t)$ по формуле

$$L(x, t) = \int_t^{\varkappa} \frac{1}{\varepsilon^0(x, t, \sigma)}\, d\sigma \qquad (\varkappa = \vartheta^\circ - \alpha) \tag{3.7}$$

Здесь $\alpha > 0$ сколь угодно мало, ϑ^0 — наименьший корень уравнения

$$\varepsilon^0(x^0, t_0, \sigma) = 0 \quad (\sigma > t_0) \tag{3.8}$$

При выполнении условия 3.1 функция (3.7) удовлетворяет требованиям (1) — (3), сформулированным в § 2.

---

[1] Для произвольного выпуклого замкнутого множества $N \subset E_n$ опорная функция $\rho_N(l)$ определяется по формуле $\rho_N(l) = \max l'x$ при $x \in N$, где $l = \{l_1, \ldots, l_n\}$ — n-мерный вектор.

Опишем кратко основные положения, используемые при проверке этого утверждения. Определим область D как совокупность точек $\{x, t\}$, для которых имеет место неравенство

$$\varepsilon^0 (x, t, \sigma) > 0 \text{ при } t \leqslant \sigma \leqslant \varkappa \tag{3.9}$$

Тогда дифференцируемость функции (3.7) в области D будет следствием дифференцируемости по x и t функции $\varepsilon^0 = \varepsilon^0 (x, t, \sigma)$ в окрестности тех точек $\{x, t, \sigma\}$, для которых выполняется неравенство $\varepsilon^0 (x, t, \sigma) > 0$, при этом существование производных функции $\varepsilon^0 (x, t, \sigma)$ можно вывести из формулы (3.5) (или (3.6)), используя условие 3.1 (см., например, [6,13]).

Включение $\{x^0, t_0\} \in D$ выполняется в силу определений величины $\varkappa$ и области D. Второе требование, сформулированное в § 2, в данном случае вытекает из определений величины $\varepsilon^0 (x, t, \sigma)$ и области D.

Третье требование также выполняется для функции $L = L (x, t)$, определяемой (3.7). Это можно проверить, используя неравенство

$$|\varepsilon^0 (x, t, \sigma') - \varepsilon^0 (x, t, \sigma'')| \leqslant k|\sigma' - \sigma''| \text{ при } \{x, t\} \in G \times [t_0, \varkappa] \tag{3.10}$$

$$\delta', \delta'' \leqslant \varkappa$$

которое выводится непосредственно из определения величины $\varepsilon^0 (x, t, \sigma)$.

Таким образом, при условии 3.1 функция (3.7) удовлетворяет требованиям (1) — (3) § 2. Будем предполагать также, что система (1.1) и ограничения (1.2) таковы, что выполняется дополнительное условие.

Условие 3.2. При $\varkappa < \vartheta^0$ функция (3.7) удовлетворяет неравенству (2.3).

Если для рассматриваемой игровой задачи выполняются условия 3.1 и 3.2, то функция (3.7) удовлетворяет всем требованиям, сформулированным в § 2, следовательно, в силу теоремы 2.1 можно построить стратегию $V = V_\alpha (t, x)$, определяемую соотношениями (2.4), (3.7), которая обеспечивает второму игроку приведение точки $x = x [t]$ на M не раньше, чем к моменту $t = \varkappa = \vartheta^0 - \alpha$, где $\alpha > 0$ сколь угодно мало. Заметим, что при выполнении условия 3.1 можно построить управление $U^0(x, t)$, гарантирующее первому игроку приведение объекта (1.1) на множество M к моменту $t = \vartheta^0$. Конструкция этого управления $U^0(x, t)$, предположенная в работах [9,10] для задач преследования, целиком переносится и на рассматриваемую игровую задачу. Для проверки этого факта нет необходимости повторять выкладки, приведенные в работах [9,10], поскольку, как отмечалось выше, рассматриваемая здесь задача может быть интерпретирована как задача об игровой встрече управляемых объектов (3.2) (3.4).

Таким образом, при выполнении условий 3.1 и 3.2 существуют стратегии $U^0 (x, t)$, $V_\alpha (x, t)$, первая из которых гарантирует приведение фазового вектора $x = x [t]$ на M при любом выборе стратегии $V (x, t)$ не позже, чем к моменту $t = \vartheta^0$, а вторая стратегия $V_\alpha (x, t)$ обеспечивает уклонение траектории $x = x [t]$ от попадания на множество M до момента $\vartheta^0 - \alpha$ (где α —сколь угодно мало) при любом выборе стратегии $U (x,t)$. Поэтому стратегия $U^0 (x, t)$ будет оптимальной стратегией первого игрока, а стратегии $V_\alpha (x, t)$ гарантируют второму игроку результат $\vartheta^0 - \alpha$, сколь угодно близкий (при малых α) к наилучшему, равному моменту ϑ^0.

Эффективная проверка условий 3.1 и 3.2 в общем случае затруднительна. Ниже приводится формулировка ограничений, которые достаточно наложить на коэффициенты системы (1.1) и на множества U^*, V^*, M для того, чтобы выполнялись условия 3.1 и 3.2.

1) Множества

$$U_* (\sigma, \tau) = X (\sigma, \tau) B (\tau) U^*, \quad V_* (\sigma, \tau) = X (\sigma, \tau) C (\tau) V^*$$

удовлетворяют условию

$$U_* (\sigma, \tau) = V_* (\sigma, \tau) + W (\sigma, \tau) \tag{3.11}$$

где $W (\sigma, \tau)$ — некоторое выпуклое множество.

2) Для любого вектора $u \in U^*$ найдется вектор $v \in V^*$, такой, что

$$X (\sigma, \tau) (B (\tau) u - C (\tau) v) \in W (\sigma, \tau) \quad \text{при} \quad \tau \leqslant \sigma \leqslant \varkappa \tag{3.12}$$

Проверка приведенных выше соотношений проще проверки условий 3.1 и 3.2. В частности, указанным условиям (3.11) и (3.12) удовлетворяет контрольный пример, приведенный в работе [3], а также те игровые задачи (1.1), (1.2), которые могут быть интерпретированы, как задачи об игровой встрече однотипных объектов.

Заметим, что соотношения (3.11), (3.12) совпадают с приведенными в докладе [7] условиями, при которых управление, построенное в работе [4], будет оптимальным. В случае, когда уравнение $\varepsilon^0 (x^0, t_0, \sigma) = 0$ не имеет решения, описанная выше конструкция обеспечивает второму игроку уклонение фазовой точки от попадания на многообразия M до любого сколь угодно большого момента времени $\varkappa$.

Поступила 27 XII 1968

ЛИТЕРАТУРА

1. Айзекс Р. Дифференциальные игры. М., «Мир», 1967.
2. Nardzewski C. R. A theory of pursuit and evasion. Adv. in game theory, Ann. of Math. Studies, 1964, vol. 35, No. 52.
3. Понтрягин Л. С. К теории дифференциальных игр. Усп. матем. наук, 1966, т. 21, вып. 4.
4. Понтрягин Л. С. О линейных дифференциальных играх. 2. Докл. АН СССР. 1967, т. 175, № 4.
5. Мищенко Е. Ф., Понтрягин Л. С. Линейные дифференциальные игры. Докл. АН СССР, 1967, т. 174, № 1.
6. Пшеничный Б. Н. Линейные дифференциальные игры. Автоматика и телемеханика, 1968, № 1.
7. Гусятников П. Б., Никольский М. С. Об оптимальности времени преследования. I Всес. конференц. по теории игр. Тезисы докл. и науч. сообщ., ноябрь, 1968.
8. Красовский Н. Н., Репин Ю. М., Третьяков В. Е. О некоторых игровых ситуациях в теории управляемых систем. Изв. АН СССР, Техн. кибернетика, 1965, № 4.
9. Красовский Н. Н. К задаче об игровой встрече движений. ПММ, 1968, т. 32, вып. 5.
10. Красовский Н. Н. О дифференциальной игре на сближение. Докл. АН СССР, 1968, т. 182, № 6.
11. Красовский Н. Н., Субботин А. И. Задача о сближении управляемых объектов. ПММ, 1968, т. 32, вып. 4.
12. Филиппов А. Ф. Дифференциальные уравнения с разрывной правой частью. Матем. сб., 1960, т. 51 (93), вып. 1.
13. Субботин А. И. Регуляризация одной задачи о встрече движений. Дифференциальные уравнения, 1968, т. 6, № 5.