Прикладная математика и механика. Том 44, вып. 4, 1980

УДК 62—50

О ПОЗИЦИОННОМ МИНИМАКСНОМ УПРАВЛЕНИИ

А. Н. Красовский

(Свердловск)

На основе теоретико-игрового подхода, предложенного и развитого в работах [1−6], рассматривается задача об управлении дифференциальной системой при неопределенной помехе. Основной результат состоит в построении седловой точки рассматриваемой дифференциальной игры в форме оптимальных позиционных смешанных стратегий для определенного класса функционалов, названных позиционными. Устанавливается, что оптимальные стратегии могут быть заданы функциями, зависящими только от текущей позиции и от некоторого параметра, введение которого — существенный элемент предлагаемой схемы. Строится устойчивая аппроксимационная схема управления, гарантирующая для игроков результаты, сколь угодно близкие к цене игры с вероятностью, сколь угодно близкой к единице, если только временной шаг будет достаточно мал.

1. Рассматривается объект, описываемый дифференциальным уравнением

$$
\begin{gathered}
y^{\cdot} = f(t, y, u, v), \quad t_0 \leqslant t \leqslant \vartheta, \quad u \in P, \quad v \in Q \\
|f(t, y, u, v)| \leqslant \varkappa \cdot (1 + |y|), \quad \varkappa = \mathrm{const}
\end{gathered} \tag{1.1}
$$

где y — n-мерный вектор, t — время, u и v — векторные управляющие воздействия, P и Q — компакты. Функцию f полагаем непрерывной и в каждой ограниченной области G, удовлетворяющей условию Липшица по y с константой L_G.

Будем рассматривать движения, начинающиеся в заданной ограниченной области G_0. Тогда любые движения на отрезке $[t_0, \vartheta]$, которые встретятся в дальнейшем, не выйдут из некоторой ограниченной области G. При этом все рассматриваемые ниже непрерывные движения будут удовлетворять по t условию Липшица. Рассмотрим задачу об управлениях u и v, которые соответственно минимизируют — максимизируют заданный функционал γ от движения $y_{t_0}[\cdot]_\vartheta = \{y[t], \ t_0 \leqslant t \leqslant \vartheta\}$. Результат состоит в построении оптимальных смешанных стратегий в рамках позиционной дифференциальной игры. В предлагаемой схеме основную роль играют некоторые модели.

2. Состояние x-модели в момент t характеризуется n-мерным вектором $x[t]$. Назовем действием первого игрока $F^{(1)}(t_*, t^*)$ на интервале (t_*, t^*), $t_0 \leqslant t_* < t^* \leqslant \vartheta$, набор измеримых по Борелю функций $\{u_j[t] \in P, \ p_j[t] \geqslant 0, \ j = 1, 2, \ldots, N^{(1)}; \ p_1[t] + \ldots + p_{N^{(1)}}[t] = 1\}$. Действие второго игрока $F^{(2)}(t_*, t^*)$ — набор измеримых функций $\{v_k[t] \in Q, \ q_k[t] \geqslant 0, \ k = 1, 2, \ldots, N^{(2)}; \ q_1[t] + \ldots + q_{N^{(2)}}[t] = 1\}$. Для

данной начальной позиции $\{t_*, x_*\}$ действия $F^{(1)}(t_*, t^*)$ и $F^{(2)}(t_*, t^*)$ порождают движение $x[t]$, $t_* \leqslant t \leqslant t^*$,— абсолютно непрерывное решение уравнения

$$x^{\cdot}[t] = \sum_{j,k=1}^{N^{(1)},\,N^{(2)}} f(t, x[t], u_j[t], v_k[t]) \cdot p_j[t] \cdot q_k[t] \tag{2.1}$$

Действие $F^{(1)}(t_*, t^*)(F^{(2)}(t_*, t^*))$ назовем элементарным, если $u_j[t] = u_j = \text{const}$, $p_j[t] = p_j = \text{const}$ $(v_k[t] = v_k = \text{const}$, $q_k[t] = q_k = \text{const})$.

Стратегией U_x первого игрока (V_x второго игрока) назовем правило, которое любым возможным значениям $\{t, x, \varepsilon > 0\}$ назначает постоянные векторы $\{p_1, \ldots, p_{N(1)}\}$, $\{u_1, \ldots, u_{N(1)}\}$, $u_j \in P$ $(\{q_1, \ldots, q_{N(2)}\}$, $\{v_1, \ldots, v_{N(2)}\}$, $v_k \in Q)$.

Будем строить $\{\varepsilon, \Delta\}$-движение x-модели, порожденное стратегией U_x из позиции $\{t_*, x[t_*]\}$, по шагам $[\tau_i, \tau_{i+1}]$. Пусть выбрано разбиение Δ отрезка $[t_0, \vartheta]$ точками τ_i, так, что $t_0 = \tau_0 < \tau_1 < \ldots < \tau_m = \vartheta$. Зафиксируем $\varepsilon > 0$. Если реализовалась позиция $\{\tau_i, x[\tau_i]\}$, то стратегия U_x по $\{\tau_i, x[\tau_i], \varepsilon\}$ назначает векторы $\{p_j^{[i]}\}$ и $\{u_j^{[i]}\}$ и на промежутке (τ_i, τ_{i+1}) работает соответствующее элементарное действие $F^{(1)}(\tau_i, \tau_{i+1})$. Второй игрок может выбрать любое действие $F^{(2)}(\tau_i, \tau_{i+1})$ (2.2). Эта пара действий реализует движение $x_\Delta^\varepsilon[t]$, $\tau_i \leqslant t \leqslant \tau_{i+1}$,— решение уравнения (2.1) при $t_* = \tau_i$, $t^* = \tau_{i+1}$, $u_j[t] = u_j^{[i]}$, $p_j[t] = p_j^{[i]}$. Аналогично определяется движение x-модели, порожденное стратегией V_x.

3. Будем рассматривать функционалы $\gamma(x_{t_*}[\cdot]_\vartheta)$, $t_0 \leqslant t_* \leqslant \vartheta$, $x_{t_*}[\cdot]_\vartheta = \{x[t], t_* \leqslant t \leqslant \vartheta\}$, определенные на кусочно-непрерывных кривых $x_{t_*}[\cdot]_\vartheta$, имеющих только конечное число точек разрыва первого рода и непрерывных справа. Функционалы непрерывны в следующем смысле: для любого $\varepsilon > 0$ существует $\delta(\varepsilon) > 0$, так, что $|\gamma(x_{t_*}^{(1)}[\cdot]_\vartheta) - \gamma(x_{t_*}^{(2)}[\cdot]_\vartheta)| < \varepsilon$, как только

$$\sup_{t_* \leqslant t \leqslant \vartheta} |x^{(1)}[t] - x^{(2)}[t]| < \delta, \quad x^{(i)}[t] \in G$$

Назовем позиционными функционалы, которые представимы в виде

$$\gamma(x_{t_*}[\cdot]_\vartheta) = \varphi(x_{t_*}[\cdot]_{t^*}, \alpha), \quad \alpha = \gamma(x_{t^*}[\cdot]_\vartheta) \tag{3.1}$$

где функция $\varphi(x_{t_*}[\cdot]_{t^*}, \alpha)$, при фиксированной кривой $x_{t_*}[\cdot]_{t^*}$ непрерывна и не убывает по α. В частности, такими будут функционалы

$$\gamma = \int_{t_0}^{\theta} w(t, x[t])\,dt + \sigma(x[\vartheta])$$

$$\gamma = \inf_{t_0 \leqslant \tau \leqslant \vartheta} \max(\sup_{t_* \leqslant t \leqslant \tau} \omega(t, x, [t]),\ \sigma(\tau, x[\tau]))$$

Если функция $x_{t_*}[\cdot]_\vartheta$ разрывна при $t = t^*$, то в записи $\varphi(x_{t_*}[\cdot]_{t^*}, \alpha)$ символ $x_{t_*}[\cdot]_{t^*}$ обозначает функцию $x[t]$ при $t_* \leqslant t < t^*$.

4. Пусть $X(U, \varepsilon, \delta, t_*, x_*) = \{x_{t_*}[\cdot]_\vartheta\}$ $(X(V, \varepsilon, \delta, t_*, x_*) = \{x_{t_*}[\cdot]_\vartheta\})$ — пучок $\{\varepsilon, \Delta\}$ — движений, порожденных стратегией U_x (V_x) при $(\tau_{i+1} - \tau_i) \leqslant \delta$ из позиции $\{t_*, x_*\}$, $t_* \in [t_0, \vartheta]$. Пусть

$$\inf_U \inf_\varepsilon \inf_\delta \sup [\gamma(x_{t_*}[\cdot]_\vartheta),\ x_{t_*}[\cdot]_\vartheta \in X(U, \varepsilon, \delta, t_*, x_*)] = \gamma^{(1)}(t_*, x[t_*])$$

$$\sup_V \sup_\varepsilon \sup_\delta \inf [\gamma (x_{t_*} [\cdot]_\vartheta), x_{t_*} [\cdot]_\vartheta \in X (V, \varepsilon, \delta, t_*, x_*)] = \\ = \gamma^{(2)} (t_*, x [t_*])$$

Стратегию U_x° назовем оптимальной, если для любого $\zeta > 0$

$$\gamma (x_{t_*} [\cdot]_\vartheta) \leqslant \gamma^{(1)} (t_*, x [t_*]) + \zeta$$

$$x_{t_*} [\cdot]_\vartheta \in X (U_x^0, \varepsilon, \delta, t_*, x_*), \quad \varepsilon \leqslant \varepsilon (\zeta), \quad \delta \leqslant \delta (\varepsilon)$$

Заметим, что для любого выбранного наперед $\zeta > 0$ имеем

$$\gamma (x_{t_0} [\cdot]_\vartheta) \leqslant \varphi (x_{t_0} [\cdot]_{t_*}, \ \gamma^{(1)} (t_*, x [t_*])) + \zeta$$

где $x_{t_0} [\cdot]_{t_*}$ — движение, реализовавшееся к моменту t_*, а $x_{t_0} [\cdot]_\vartheta$ — движение, составленное из $x_{t_0} [\cdot]_{t_*}$ и любого $\{\varepsilon, \Delta\}$ — движения $x_{t_*} [\cdot]_\vartheta \in \in X (U_x^\circ, \varepsilon, \delta, t_*, x [t_*])$ при $\varepsilon \leqslant \varepsilon (\zeta)$, $\delta \leqslant \delta (\varepsilon, \zeta)$.

Стратегию V_x° назовем оптимально-позиционной, если для любого $\zeta > 0$

$$\gamma (x_{t_*} [\cdot]_\vartheta) \geqslant \gamma^{(2)} (t_*, x [t_*]) - \zeta$$

$$x_{t_*} [\cdot]_\vartheta \in X (V_x^\circ, \varepsilon, \delta, t_*, x [t_*]), \quad \varepsilon \leqslant \varepsilon (\zeta), \quad \delta \leqslant \delta (\varepsilon, \zeta).$$

Имеем

$$\gamma (x_{t_0} [\cdot]_\vartheta) \geqslant \varphi (x_{t_0} [\cdot]_{t_*}, \ \gamma^{(2)} (t_*, x [t_*])) - \zeta$$

где $x_{t_0} [\cdot]_{t_*}$ — движение, реализовавшееся к моменту t_*, а $x_{t_0} [\cdot]_\vartheta$ — движение, составленное из $x_{t_0} [\cdot]_{t_*}$ и любого $\{\varepsilon, \Delta\}$ — движения $x_{t_*} [\cdot]_\vartheta \in \in X (V_x^\circ, \varepsilon, \delta, t_*, x [t_*])$, $\varepsilon \leqslant \varepsilon (\zeta)$, $\delta \leqslant \delta (\varepsilon, \zeta)$.

Скажем, что пара $\{U_x^\circ, V_x^\circ\}$ составляет позиционную седловую точку и дает цену игры $\gamma^\circ (t, x)$, если

$$\gamma^{(1)} (t, x) = \gamma^{(2)} (t, x) = \gamma^\circ (t, x)$$

$$\gamma^\circ (t_*, x [t_*]) \geqslant \varphi (x_{t_*} [\cdot]_{t^*}, \ \gamma^\circ (t^*, x [t^*])) - \zeta$$

$$x_{t_*} [\cdot]_\vartheta \in X (U_x^\circ, \ \varepsilon, \delta, t_*, x [t_*])$$

$$\gamma^\circ (t_*, x [t_*]) \leqslant \varphi (x_{t_*} [\cdot]_{t^*}, \ \gamma^\circ (t^*, x [t^*])) + \zeta$$

$$x_{t_*} [\cdot]_\vartheta \in X (V_*^\circ, \ \varepsilon, \delta, t_*, x [t_*])$$

для любого $\zeta > 0$ при $\varepsilon \leqslant \varepsilon (\zeta)$ и $\delta \leqslant \delta (\varepsilon, \zeta)$.

Задача состоит в построении стратегий $\{U_x^\circ, V_x^\circ\}$, составляющих седловую точку.

5. Состояние w-модели характеризуется n-мерным вектором $w [t]$. Движения w-модели порождаются действиями, которые определяются так же, как и для x-модели, и которые для отличия от x-модели, будем отмечать звездочкой $F_*^{(1)} (t_*, t^*)$, $F_*^{(2)} (t_*, t^*)$. Движение w-модели на отрезке $[t_*, t^*]$ — решение уравнения

$$w^\cdot [t] = \sum_{j, k=1}^{N^{(1)}, N^{(2)}} f (t, w [t], u_j [t], v_k [t]) \cdot p_{*j} [t] \cdot q_{*k} [t], \qquad w [t_*] = w_* \tag{5.1}$$

Назовем $Q_{\{t_*, w_*\}}$-процедурой правило, которое в момент $\tau_i \geqslant t_*$ на основании данных $\{\tau_i, F_{*s}^{(1)}, F_{*s}^{(2)}, \ s = 0, 1, \ldots, i - 1\}$, $F_{*s}^{(k)} = F_*^{(k)} (\tau_s, \tau_{s+1})$ назначает $\tau_{i+1} > \tau_i$ и $F_{*i}^{(2)} = F_*^{(2)} (\tau_i, \tau_{i+1})$. Первый игрок назначает $F_{*i}^{(1)} = F_*^{(1)} (\tau_i, \tau_{i+1})$ по известным τ_{i+1} и $F_{*i}^{(2)}$. В обозначении

$Q_{\{t_*, w_*\}}$ символ $\{t_*, w_*\}$ указывает, что движение формируется из позиции $\{t_*, w_*\}$. Допустимы процедуры, у которых число моментов τ_i от $\tau_0 = t_*$ до ϑ конечно для каждой возможной реализации $w_{t_*}[\cdot]_\vartheta$. Для разных реализаций это число может быть различным. Возможны реализации с любым числом моментов τ_i.

Пусть β — некоторое число. Процедуру $Q_{\{t_*, w_*\}}$ назовем $(\beta - Q_{\{t_*, w_*\}})$-процедурой, если для любого движения $w_{t_*}[\cdot]_\vartheta$, которое она породит, будет выполнено условие $\gamma(w_{t_*}[\cdot]_\vartheta) > \beta$. Обозначим $\rho(t_*, w_*)$ — точную верхнюю грань β, для которых существует $(\beta - Q_{\{t_*, w_*\}})$-процедура. Можно показать, что функция $\rho(t, w)$ непрерывна по w.

Справедливы утверждения.

Лемма 5.1. Пусть дана позиция $\{t_*, w_*\}$. Заданы $t^* > t_*$, $\varepsilon_* > 0$ и $F_*^{(2)}(t_*, t^*)$. Тогда найдутся $\varepsilon^* > 0$ и действие $F_*^{(1)}(t_*, t^*)$, которое в паре с $F_*^{(2)}(t_*, t^*)$ породит такое движение $w_{t_*}[\cdot]_{t^*}$, что

$$\varphi(w_{t_*}[\cdot]_{t^*}, \rho(t^*, w^*) + 2\varepsilon^*) \geqslant \rho(t_*, w_*) + 2\varepsilon_* \tag{5.2}$$

Лемма 5.2. Пусть дана позиция $\{t_*, w_*\}$. Заданы $t^* > t_*$, $\varepsilon_* > 0$ и движение $w_{t_*}[\cdot]_{t^*}$, порожденное $(\beta - Q_{\{t_*, w_*\}})$-процедурой при $\beta = \rho(t_*, w_*) - \varepsilon_*$. Тогда найдется такое $\varepsilon^* > 0$, что

$$\varphi(w_{t_*}[\cdot]_{t^*}, \rho(t^*, w^*) - 2\varepsilon^*) \geqslant \rho(t_*, w_*) - 2\varepsilon_*, \quad w^* = w_{t_*}[t^*]_{t^*} \tag{5.3}$$

Рассмотрим функцию

$$\lambda(t, x, w) = |x - w|^2 \exp\{-3L_G \cdot (t - t_0)\} \tag{5.4}$$

Лемма 5.3. Пусть заданы ограниченная область G^* в пространстве $\{x\}$ и $\varepsilon > 0$. Тогда можно указать векторы $\{u_1, \ldots, u_N\}$, $\{v_1, \ldots, v_N\}$, $u_j \in P$, $v_k \in Q$, $N = N(\varepsilon, G^*)$ и число $\delta(\varepsilon, G^*) > 0$, такие, что если для $t_* \in [t_0, \vartheta]$, $x_* \in G^*$ и $w_* \in G^*$ векторы $\{p_1^\circ, \ldots, p_N^\circ\}$ и $\{q_{*1}^\circ, \ldots, q_{*N}^\circ\}$ определены из условий

$$\max_q \sum_{j,k=1}^{N,N} \langle s_* \cdot f(t_*, x_*, u_j, v_k) \cdot p_j^\circ \cdot q_k \rangle = \min_p \operatorname{Idem}(p_j^\circ \to p_j) \tag{5.5}$$

$$\min_{p_*} \sum_{j,k=1}^{N,N} \langle s_* \cdot f(t_*, w_*, u_j, v_k) \cdot p_{*j} \cdot \overset{\circ}{q}_{*k} \rangle = \max_{q_*} \operatorname{Idem}(\overset{\circ}{q}_{*k} \to q_{*k}) \tag{5.6}$$

где $s_* = x_* - w_*$, то при $t_* \leqslant t \leqslant t_* + \delta(\varepsilon, G^*)$ элементарное действие $F^{(1)}(t_*, t)$, соответствующее $\{p_j^\circ\}$, в паре с любым действием $F^{(2)}(t_*, t)$ и элементарное действие $F_*^{(2)}(t_*, t)$, соответствующее $\{q_{*k}^\circ\}$, в паре с любым действием $F_*^{(1)}(t_*, t)$ породят движения $x[t]$ и $w[t]$, для которых

$$\lambda(t, x[t], w[t]) \leqslant \lambda(t_*, x_*, w_*) + \varepsilon \cdot (t - t_*) \tag{5.7}$$

Здесь и далее Idem в правой части равенства означает выражение, совпадающее с левой частью этого равенства при указанной в скобках замене символов.

Лемма 5.4. Формулируется аналогично лемме 5.3 при взаимной замене символов p и q, а также действий $F^{(1)}(t_*, t)$ и $F^{(2)}(t_*, t)$.

6. Построим стратегии U_x^e и V_x^e, которые назовем экстремальными. Выберем $\varepsilon > 0$, $\{t_*, x_*\}$. Обозначим через K_* множество точек w, удовлетворяющих неравенству

(6.1) $$\lambda(t_*, x_*, w) \leqslant \varepsilon \cdot (t_* - t_0)$$

Назовем сопутствующей точку $w_* \in K_*$, удовлетворяющую условию:
а) при построении стратегии U_x^e

(6.2) $$\rho(t_*, w_*) = \min_{w \in K_*} \rho(t_*, w)$$

б) при построении стратегии V_x^e

(6.3) $$\rho(t_*, w_*) = \max_{w \in K_*} \rho(t_*, w)$$

Таких точек w_* может быть несколько. Выберем для каждых данных $\{t_*, x_*, \varepsilon\}$ одну из них. Зафиксируем $\varepsilon^* > 0$ и выберем область G^* так, чтобы при всех $\varepsilon < \varepsilon^*$, $x_* \in G$ было $w_* \in G^*$. Далее выбираем только $\varepsilon < \varepsilon^*$.

Экстремальной стратегией U_x^e (V_x^e) будем называть правило, которое возможным значениям $\{t_*, x_*, \varepsilon\}$ ставит в соответствие векторы $\{u_j\}$ и $\{p_j^\circ\}$ ($\{v_k\}$ и $\{q_k^\circ\}$), связанные условием (5.5) (соответствующим условием из леммы 5.4) и удовлетворяющие оценке (5.7), где w_* — сопутствующая точка. Вследствие теорем об измеримом выборе [7] векторы $\{p_j^\circ\}$ и $\{q_k^\circ\}$ можно выбрать как функции, измеримые по Борелю по x_*, так как условия (5.5), (5.7), соответствующие условия из леммы 5.4 и соотношения (6.1) — (6.3) определяют для выбора $\{p_j^\circ\}$ и $\{q_k^\circ\}$ компактные множества, которые полунепрерывны относительно x_*.

Пусть первый игрок руководствуется стратегией U_x^e, выбраны $\varepsilon > 0$ и разбиение $\Delta = \{\tau_i\}$. Эти данные порождают некоторое движение $x^\varepsilon_{\Delta t_*}[\cdot]_\vartheta$. Пусть $w[\tau_i]$ — сопутствующая точка для $x[\tau_i]$. Движению $x_\Delta^\varepsilon[t]$ поставим в соответствие сопутствующее движение w-модели. На полуинтервале $\tau_i \leqslant t < \tau_{i+1}$ это будет решение уравнения (5.1) с граничным условием $\{\tau_i, w[\tau_i]\}$, порожденное действиями $F_*^{(2)}(\tau_i, \tau_{i+1})$ и $F_*^{(1)}(\tau_i, \tau_{i+1})$ из леммы 5.3 при $x_* = x[\tau_i]$, $w_* = w[\tau_i]$. Движение $w_{t_*}[\cdot]_\vartheta$ в моменты τ_i может претерпевать разрывы. Обозначим $w^{[i]} = \lim w[t]$, $t \to \tau_i - 0$.

Опираясь на лемму 5.3 и свойства экстремальной стратегии, можно показать, что если $\max |\tau_{i+1} - \tau_i| < \delta(\varepsilon, G)$, то будет иметь место оценка $\lambda(t, x_\Delta^\varepsilon[t], w[t]) \leqslant \varepsilon + \varepsilon \cdot (t - t_*)$, $t_* \leqslant t \leqslant \vartheta$, какие бы ни были $F^{[2]}(\tau_i, \tau_{i+1})$ и $F_*^{(1)}(\tau_i, \tau_{i+1})$. Кроме того, $\rho(\tau_{i+1}, w[\tau_{i+1}]) \leqslant \rho(\tau_{i+1}, w^{[i+1]})$. Пусть в сопутствующем движении $w[t]$ действие $F_*^{(1)}(\tau_i, \tau_{i+1})$ на каждом шаге выбирается в соответствии с леммой 5.1. Тогда на каждом шаге имеем

$$\varphi(w_{\tau_i}[\cdot]_{\tau_{i+1}}, \rho(\tau_{i+1}, w^{[i+1]}) + 2\varepsilon_{i+1}) \leqslant \rho(\tau_i, w[\tau_i]) + 2\varepsilon_i$$

откуда получаем $\gamma(w_{\tau_i}[\cdot]_\vartheta) \leqslant \rho(\tau_i, w[\tau_i]) + 2\varepsilon_i$ и, полагая $\varepsilon_0 = \varepsilon$, имеем тогда по индукции

(6.4) $$\gamma(w_{t_*}[\cdot]_\vartheta) \leqslant \rho(t_*, w_*) + 2\varepsilon$$

Пользуясь близостью движения $x^{\varepsilon\Delta^{t_*}}[\cdot, U_x^e]_\vartheta$ к движению $w_{t_*}[\cdot]_\vartheta$, непрерывностью функционала γ и непрерывностью $\rho(t, w)$ по w, получаем следующий результат.

Теорема 6.1. Пусть реализовалась позиция $\{t_*, x_*\}$. Если, начиная с момента t_*, первый игрок использует экстремальную стратегию U_x^e, то для любого сколь угодно малого ε и достаточно малого шага $\delta(\varepsilon)$, для любого $\{\varepsilon, \Delta\}$-движения $x^\varepsilon_{\Delta t_*}[\cdot]_\vartheta$, $x^\varepsilon_{\Delta t_*}[t_*]_\vartheta = x_*$ имеем

$$(6.5)\qquad \gamma(x^\varepsilon_{\Delta t_*}[\cdot]_\vartheta) \leqslant \rho(t_*, x_*) + \xi(\varepsilon), \quad \lim_{\varepsilon\to 0}\xi(\varepsilon) = 0$$

Пусть теперь второй игрок руководствуется экстремальной стратегией V_x^e. Опираясь на свойства этой стратегии и леммы 5.2 и 5.4, можно вывести соотношения $\lambda(t, x^\varepsilon_\Delta[t], w_i^*[t]) \leqslant \varepsilon + \varepsilon\cdot(t - \tau_i)$, $\tau_i \leqslant t \leqslant \tau_{i+1}$, $\rho(\tau_{i+1}, w[\tau_{i+1}]) \geqslant \rho(\tau_{i+1}, w^{[i+1]})$ и затем

$$(6.6)\qquad \gamma(w_{t_*}[\cdot]_\vartheta) \geqslant \rho(t_*, w_*) - 2\varepsilon$$

где $\{t_*, w_* = w[t_*]\}$ — сопутствующая точка для позиции $\{t_*, w_*\}$, $w_{t_*}[\cdot]_\vartheta$ — движение w-модели, которое на каждом шаге порождается действием $F_*^{(1)}(\tau_i, \tau_{i+1})$, выбираемым из условия леммы 5.4, соответствующего условию (5.6), и действием $F_*^{(2)}(\tau_i, \tau_{i+1})$, назначаемым ($\beta$ — $Q_{\{\tau_i w[\tau_i]\}}$)-процедурой при $\beta = \rho(\tau_i, w[\tau_i]) - \varepsilon_i$. Отсюда получаем следующий результат.

Теорема 6.2. Пусть реализовалась позиция $\{t_*, x_*\}$. Если, начиная с момента t_*, второй игрок использует экстремальную стратегию V_x^e, то для любого сколь угодно малого ε и достаточно малого шага $\delta(\varepsilon) > 0$, для любого $\{\varepsilon, \Delta\}$-движения имеем оценку

$$(6.7)\qquad \gamma(x^\varepsilon_{\Delta t_*}[\cdot]_\vartheta) \geqslant \rho(t_*, x_*) - \eta(\varepsilon), \quad \lim_{\varepsilon\to 0}\eta(\varepsilon) = 0$$

Сопоставляя теоремы 6.1 и 6.2, получаем следующее утверждение.

Теорема 6.3. Экстремальная стратегия U_x^e есть оптимальная стратегия первого игрока. Экстремальная стратегия V_x^e есть оптимальная стратегия второго игрока. Пара стратегий $\{U_x^e, V_x^e\}$ образует позиционную седловую точку. Цена игры $\gamma^\circ(t, x) = \rho(t, x)$.

7. Рассмотрим теперь исходную задачу об управлении данным объектом (1.1). Используя метод управления с поводырем [1], где в качестве поводыря выбрана x-модель, определим некоторую объединенную стратегию U первого игрока.

Движение объекта, порожденное из позиции $\{t_0, y_0 \in G_0\}$ стратегией U, строится одновременно с некоторым $\{\varepsilon, \Delta\}$-движением x-модели, порожденным некоторой стратегией U_x, включенной в U. Пусть выбраны U, $\varepsilon > 0$ и $\Delta = \{\tau_i\}$. По данным $\{y[\tau_i], x[\tau_i], \tau_i, \varepsilon\}$ некоторая стратегия U_y, включенная в U, назначает элементарное действие $F_y^{(1)}(\tau_i, \tau_{i+1})$. Производится случайное испытание, соответствующее распределению вероятностей $\{p_1^{[i]}, \ldots, p_{N(1)}^{[i]}\}$ случайной величины $\{u_1, \ldots, u_{N(1)}\}$, определяемому действием $F_y^{(1)}(\tau_i, \tau_{i+1})$. Результат этого испытания $u_{[i]}$ —

управление при $\tau_i < t < \tau_{i+1}$. Второй игрок, пользуясь каким-то своим случайным механизмом, вырабатывает функцию $v_{[i]}[t]$, $\tau_i < t < \tau_{i+1}$. Движение объекта на отрезке $[\tau_i, \tau_{i+1}]$ есть решение уравнения

$$(7.1) \qquad \dot{y}[t] = f(t, y[t], u_{[i]}, v_{[i]}[t])$$

Движение $x_\Delta^\varepsilon[t]$ осуществляется с тем же разбиением $\Delta = \{\tau_i\}$ из некоторой подходящей позиции $\{t_0, x_0\}$. При этом в реальном управлении объединенной системой из y-объекта и x-модели движением x-модели управляет только реальный первый игрок. Однако в связи с материалом из п. 1—6 удобно разделять его действия на действие первого условного игрока, назначающего $F^{(1)}(\tau_i, \tau_{i+1})$ в соответствии с выбранной стратегией U_x, и на действие второго условного игрока, назначающего $F^{(2)}(\tau_i, \tau_{i+1})$ в соответствии с некоторым правилом R_U. Итак, в распоряжении первого реального игрока оказывается выбор $\{U_y, U_x, R_U, x_0\}$, который будем называть его объединенной стратегией U. Движение $\{y[t], x[t]\}$, порожденное U, будем обозначать $\{y_{t_0}[\cdot, U]_\vartheta, x_{t_0}[\cdot, U]_\vartheta\}$. Оно получается случайным, поскольку выбор воздействий определяется случайными испытаниями. Подобным же образом определяются движения объекта и x-модели, порожденные стратегиями V_y и V_x и некоторым правилом R_V, и объединенная стратегия V второго игрока. При этом второй игрок использует свою x-модель.

Опишем построение экстремальной стратегии U^e. Предполагаем, что функции $u[\tau_i]$ и $v[t]$ стохастически независимы при $\tau_i \leqslant t \leqslant \tau_{i+1}$. При этом рассматриваемый процесс $\{y[t], x[t]\}$ формализуется как вероятностный процесс в рамках строгих понятий теории вероятностей. Это не вызывает принципиальных трудностей, так как все проводимые построения дают измеримые по $\{y[\tau_i], x[\tau_i]\}$ функции-управления. Начальное условие x_0 выберем из условия

$$(7.2) \qquad \lambda(t_0, x_0, y_0) \leqslant \varepsilon$$

Движение $x[t]$ x-модели определим экстремальной стратегией U_x^e. Возьмем в качестве правила R_U^e такое, которое назначает действия $F^{(2)}(\tau_i, \tau_{i+1})$, определяемые условием

$$(7.3) \qquad \min_p \sum_{j,k=1}^{N,N} \langle g^{[i]} \cdot f(\tau_i, x[\tau_i], u_j, v_k) \cdot p_j \cdot q_k^\circ \rangle = \max_q \operatorname{Idem}(q_k^\circ \to q_k), \quad g^{[i]} = y[\tau_i] - x[\tau_i]$$

Назовем экстремальной стратегией U_y^e правило, по которому величинам $\{y_*, x_*, t_*, \varepsilon\}$, $y_* \in G^*$, $x_* \in G^*$ назначается элементарное действие $F_y^{(1)}(t_*, t) = \{u_j, p_j^\circ\}$ из условия

$$(7.4) \qquad \max_q \sum_{j,k=1}^{N,N} \langle g_* \cdot f(t_*, y_*, u_j, v_k) \cdot p_j^\circ \cdot q_k \rangle = \min_p \operatorname{Idem}(p_j^\circ \to p_j), \quad g_* = y_* - x_*$$

Аналогично определяется экстремальная стратегия V_y^e.

Назовем стратегию $U^e = \{U_y^e, U_x^e, R_U^e, x_0\}$ экстремальной объединенной стратегией. Аналогично строится объединенная экстремальная стратегия V^e.

Справедливо утверждение.

Теорема 7.1. Для любого числа $\zeta > 0$ и числа $0 \leqslant \chi < 1$ всегда можно подобрать $\varepsilon > 0$ и $\delta(\varepsilon, G^*) > 0$, такие, что если только $\max_i |\tau_{i+1} - \tau_i| < \delta(\varepsilon, G^*)$, то для всех движений, порожденных $U^e = \{U_y^e, U_x^e, R_U^e, x_0\}$, будут иметь место оценки

$$P(\gamma(y_{t_0}[\cdot, U^e]_\theta) \leqslant \rho(t_0, y_0) + \zeta) > \varkappa \tag{7.5}$$

$$P(\gamma(y_{t_*}[\cdot, U^e]_\theta) \leqslant \rho(t_*, y[t_*]) + \zeta) > \varkappa$$

$$P(\varphi(y_{t_0}[\cdot, U^e]_{t_*}, \rho(t, y[t_*])) \leqslant \rho(t_0, y) + \zeta) > \varkappa$$

и для всех движений, порожденных $V^e = \{V_y^e, V_x^e, R_V^e, x_0\}$, будут иметь место оценки, аналогичные (7.5) при замене символов U^e на V^e и ζ на $-\zeta$. Символом P обозначена вероятность соответствующего события. Заметим, что по определению из этих условий вытекает также выполнение условий

$$P(\gamma(y_{t_0}[\cdot, U^e]_\theta) \leqslant \varphi(y_{t_0}[\cdot]_{t_*}, \rho(t_*, y[t_*])) + \zeta) > \varkappa \tag{7.6}$$

$$P(\gamma(y_{t_0}[\cdot, V^e]_\theta) \geqslant \varphi(y_{t_0}[\cdot]_{t_*}, \rho(t_*, y[t_*])) - \zeta) > \varkappa$$

Таким образом, полученные соотношения (7.5), (7.6) позволяют назвать величину $\rho(t, y)$ ценой игры и описанные экстремальные объединенные стратегии U^e и V^e — оптимальными стратегиями, которые дают седловую точку игры.

8. Основная теорема 7.1 доказана в предположении, что функции $u[t]$ и $v[t]$ стохастически независимы при $\tau_i < t < \tau_{i+1}$. Если речь идет об управлении, как об игре с природой, то это условие независимости можно принять как отдельный постулат. Этот постулат без логического противоречия налагает ограничение на неизвестные механизмы, которые формируют помеху $v[t]$. Однако, если рассматривать процесс как игру между двумя реальными игроками, каждый из которых действует на основе своей стратегии U и V со своим разбиением $\Delta_U = \{\tau_i^U\}$ и $\Delta_V = \{\tau_i^V\}$, то такое условие независимости принять как постулат нельзя. Связь между $u[t]$ и $v[t]$ зависит от выбранных стратегий U и V и разбиений Δ_U и Δ_V. Однако эта трудность преодолевается следующим известным образом. Предположим, что воздействия $u[t] = u[\tau_i^U]$ и $\tau_i^U < t < \tau_{i+1}^U$ и $v[t] = v[\tau_i^V]$, $\tau_i^V < t < \tau_{i+1}^V$ на объект (1.1) получаются как результаты случайных испытаний с вероятностными распределениями $\{p_j^\circ\}$ и $\{q_k^\circ\}$, которые отвечают уже не значениям $y[\tau_i^U]$ и $y[\tau_i^V]$, но значениям $y[\tau_i^U - \tau^U]$ и $y[\tau_i^V - \tau^V]$, где $\tau^U > 0$ и $\tau^V > 0$ — постоянные информационные запаздывания. Тогда, опираясь снова на результаты из п. 1—7, получаем следующий результат.

Теорема 8.1. Экстремальные объединенные стратегии U^e и V^e, описанные в п. 7, но работающие на основе запаздывающих значений $y[\tau_i^U - \tau^U]$ и $y[\tau_i^V - \tau^V]$, составляют седловую точку $\{U^\circ, V^\circ\}$ рассматриваемой игры с ценой $\gamma^\circ = \rho\,(t, y)$, т. е. они обеспечивают выполнение условий (7.2) — (7.5), если только будут выполнены условия

$$\tau_{i+1}^U - \tau_i^U < \delta^U < \tau^V, \quad \tau_{i+1}^V - \tau_i^V < \delta^V < \tau^U \tag{8.1}$$

где δ^U, τ^U, δ^V, τ^V — достаточно малые положительные числа.

Важно заметить, что здесь уже оба игрока формируют одновременно одно и то же движение $x\,[t]$, используя каждый свою x-модель и формируя это движение каждый на основе своей объединенной стратегии U° и V° со своим разбиением Δ_U и Δ_V.

9. Основные результаты, данные в теоремах 7.1 и 8.1, сохраняют свою силу при условиях на функцию f несколько более общих, чем условия из п. 1. Именно, предположим, что эта функция ограничена в любой ограниченной области G, при фиксированном t непрерывна по y, u, v и при фиксированных y, u, v измерима по t по Борелю.

Рассмотрим уравнение

$$x\,[\tau] = x\,[t_*] + \int_{t_*}^{\tau} \int_P \int_Q f\,(t, x\,[t], u, v) \cdot \eta\,(dv, du \mid t)\, dt \tag{9.1}$$

где $\eta\,(dv, du \mid t)$ — вероятностная мера на $P \times Q$, слабо измеримая по t по Борелю. Пусть G — ограниченная область. Предположим, что уравнение (9.1) при всяком $x\,[t_*] \in G$ и всякой мере η имеет единственное решение, которое при всех $t_0 \leqslant t \leqslant \vartheta$ содержится в некоторой ограниченной области $G_G{}^*$. Все эти условия, очевидно, выполняются при условиях из п. 1. Однако и при этих более общих условиях можно повторить все леммы и теоремы из п. 1—8, заменяя только в них функцию λ (5.4) на некоторую функцию λ, построенную в соответствии с идеями из работы [6], и заменяя при выборе экстремальных векторов $\{p_j{}^\circ\}$ и $\{q_k{}^\circ\}$ функцию f на подходящую непрерывную функцию f_ε, которая в достаточно большой области G^* аппроксимирует функцию f, так, что

$$\int_{t_0}^{t} \max_{y \in G, u, v} |f_\varepsilon(\tau, y, u, v) - f\,(\tau, y, u, v)|\, d\tau \leqslant \zeta\,[t]$$

$$\xi\,[t] < \varepsilon \cdot (\vartheta - t_0)$$

Автор благодарит Ю. С. Осипова и В. Е. Третьякова за большую помощь в работе.

Поступила 8 X 1979

ЛИТЕРАТУРА

1. *Красовский Н. Н., Субботин А. И.* Позиционные дифференциальные игры. М., «Наука», 1974.
2. *Куржанский А. Б.* Управление и наблюдение в условиях неопределенности. М., «Наука», 1977.
3. *Осипов Ю. С.* Позиционное управление в параболических системах. ПММ, 1977, т. 41, вып. 2.
4. *Batuhtin V. D.* A programmed construction for the positional control. In: Lecture Notes Comput. Sci. Vol. 27. Berlin, Springer-Verlag, 1975, p. 435—439.
5. *Ченцов А. Г.* К игровой задаче наведения. Докл. АН СССР, 1976, т. 226, № 1.
6. *Кряжимский А. В.* К теории позиционных дифференциальных игр сближения — уклонения. Докл. АН СССР, 1978, т. 239, № 4.
7. *Иоффе А. Д., Тихомиров В. М.* Теория экстремальных задач. М., «Наука», 1974.