Прикладная математика и механика. Том 45, вып. 4, 1981

УДК 62-50

СТОХАСТИЧЕСКИЙ ПРОГРАММНЫЙ СИНТЕЗ ДЛЯ ДЕТЕРМИНИРОВАННОЙ ПОЗИЦИОННОЙ ДИФФЕРЕНЦИАЛЬНОЙ ИГРЫ

Красовский А. Н., Красовский Н. Н., Третьяков В. Е.

Показано, что при определенных, достаточно общих условиях цена позиционной дифференциальной игры может быть найдена на основании вспомогательных программных конструкций, которые включают подходящий случайный процесс. Работа продолжает исследования [1—12].

1. Рассмотрим систему, описываемую дифференциальным уравнением

$$\mathbf{x}^{\boldsymbol{\cdot}} = A(t)\mathbf{x} + B(t)\mathbf{u} + C(t)\mathbf{v}, \quad \mathbf{u} \in P, \quad \mathbf{v} \in Q, \quad t_0 \leqslant t \leqslant \vartheta$$

где $\mathbf{x}$ — n-мерный фазовый вектор объекта, $\mathbf{u}$ и $\mathbf{v}$ — соответственно r-мерный и s-мерный векторы управлений первого и второго игроков, $A(t)$, $B(t)$, $C(t)$ — непрерывные матрицы-функции, P и Q — выпуклые компакты.

Пусть задан функционал

$$\gamma = \gamma(\mathbf{x}(t_*[\cdot]\vartheta), \mathbf{u}(t_*[\cdot]\vartheta), \mathbf{v}(t_*[\cdot]\vartheta)) = \tag{1.1}$$
$$= \int_{t_*}^{\vartheta} [\omega(t, \mathbf{x}[t]) + \omega_1(t, \mathbf{u}[t]) + \omega_2(t, \mathbf{v}[t])]\, dt + \sigma(\mathbf{x}[\vartheta])$$

Здесь и ниже символ $\mathbf{y}(t_*[\cdot]\, t^*)$ обозначает функцию $\{\mathbf{y}[t], t_* \leqslant t \leqslant t^*\}$, $[t_*, t^*] \subset [t_0, \vartheta]$; функции ω, ω_1, ω_2 и σ непрерывны; функции ω и σ удовлетворяют условиям Липшица по $\mathbf{x}$. Содержательно первый игрок должен минимизировать функционал γ, второй — максимизировать.

Игра формализуется следующим образом. В $(r + 1)$-мерном и $(s + 1)$-мерном пространствах соответственно рассмотрим множества

$$P^*(t) = \overline{\mathrm{co}}\, \{\mathbf{u}^* = \{\mathbf{u}, \omega_1(t, \mathbf{u})\}, \ \mathbf{u} \in P\}$$
$$Q^*(t) = \overline{\mathrm{co}}\, \{\mathbf{v}^* = \{\mathbf{v}, \omega_2(t, \mathbf{v})\}, \ \mathbf{v} \in Q\}$$

и введем новые управляющие векторы $\mathbf{u}^* = \{\mathbf{u} = \{u_1^*, \ldots, u_r^*\}, u_{r+1}^*\}$, $\mathbf{v}^* = \{\mathbf{v} = \{v_1^*, \ldots, v_s^*\}, v_{s+1}^*\}$, стесненные условиями

$$\mathbf{u}^* \in P^*(t), \ \mathbf{v}^* \in Q^*(t) \tag{1.2}$$

Назовем стратегией $S(t, \mathbf{x})$ функцию, которая всякой возможной позиции $\{t, \mathbf{x}\}$ ставит в соответствие некоторое множество $S(t, \mathbf{x})$ (может быть, пустое) пар $\mathbf{s} = \{\mathbf{u}^*, \mathbf{v}^*\}$ из векторов $\mathbf{u}^*$ и $\mathbf{v}^*$ (1.2). Движением $\mathbf{x}(t_*[\cdot]t^*)$, порожденным стратегией $S(t, \mathbf{x})$ из позиции $\{t_*, \mathbf{x}_*\}$, будем называть всякую абсолютно непрерывную функцию $\mathbf{x}[t]$, $\mathbf{x}[t_*] = \mathbf{x}_*$, которая при почти всех $t \in [t_*, t^*]$ удовлетворяет условию

(1.3) $$\mathbf{x}^{\boldsymbol\cdot}[t] = A(t)\,\mathbf{x}[t] + B(t)\,\mathbf{u}[t] + C(t)\,\mathbf{v}[t]$$

где

(1.4) $$\{\mathbf{u}^*[t] = \{\mathbf{u}[t], \mathbf{u}^*_{r+1}[t]\}, \mathbf{v}^*[t] = \{\mathbf{v}[t], \mathbf{v}^*_{s+1}[t]\}\} = \mathbf{s}[t] \in S(t, \mathbf{x}[t])$$

При этом принимается, что на данном движении

(1.5) $$\gamma = \gamma(\mathbf{x}(t_*[\cdot]\vartheta), \mathbf{u}^*(t_*[\cdot]\vartheta), \mathbf{v}^*(t_*[\cdot]\vartheta)) = \int\limits_{t_*}^{\vartheta} [\omega(t, \mathbf{x}[t]) + u^*_{r+1}[t] + v^*_{s+1}[t]]\,dt + \sigma(\mathbf{x}[\vartheta])$$

Стратегией первого игрока $S_u(t, \mathbf{x})$ назовем стратегию $S(t, \mathbf{x})$, которая удовлетворяет следующему условию. Каковы бы ни были отрезок $t_* \leqslant t \leqslant t^*$, позиция $\{t_*, \mathbf{x}_*\}$ и измеримая по t допустимая функция $\mathbf{v}^*(t_*[\cdot]t^*)$, найдется измеримая по t допустимая функция $\mathbf{u}^*(t_*[\cdot]t^*)$, такая, что функция $\mathbf{x}(t_*[\cdot]t^*)$, удовлетворяющая уравнению (1.3) и условию $\mathbf{x}[t_*] = \mathbf{x}_*$, будет движением, порожденным стратегией $S(t, \mathbf{x}) = S_u(t, \mathbf{x})$, т. е. для нее будет выполнено условие (1.4) при $S = S_u$ при почти всех $t \in [t_*, t^*]$. Аналогичным образом определяется стратегия $S_v(t, \mathbf{x})$ второго игрока.

Скажем, что стратегии S_u и S_v совместимы, если при всяком выборе $\{t_*, \mathbf{x}_*\}$ и $[t_*, t^*]$ существует функция $\mathbf{x}(t_*[\cdot]t^*)$, которая одновременно является движением, порожденным и стратегией S_u, и стратегией S_v.

Скажем, что совместимые стратегии S_u° и S_v° образуют седловую точку игры на минимакс функционала γ (1.1), (1.5) и дают цену игры $\rho^\circ(t, \mathbf{x})$, если при всякой начальной позиции $\{t_*, x_*\}$ для всякого движения $\mathbf{x}(t_*[\cdot]\vartheta)$, порожденного стратегией S_u°, справедливо неравенство

$$\gamma(\mathbf{x}(t_*[\cdot]\vartheta),\ \mathbf{u}^*(t_*[\cdot]\vartheta),\ \mathbf{v}^*(t_*[\cdot]\vartheta)) \leqslant \rho^\circ(t_*, \mathbf{x}_*)$$

а для всякого движения $\mathbf{x}(t_*[\cdot]\vartheta)$, порожденного стратегией S_v°

$$\gamma(\mathbf{x}(t_*[\cdot]\vartheta),\ \mathbf{u}^*(t_*[\cdot]\vartheta),\ \mathbf{v}^*(t_*[\cdot]\vartheta)) \geqslant \rho^\circ(t_*, \mathbf{x}_*)$$

Для движения, порожденного одновременно стратегиями S_u° и S_v°, выполняется, стало быть, равенство $\gamma = \rho^\circ(t_*, \mathbf{x}_*)$.

Содержательный смысл данной формализации раскрывается через аппроксимационные стратегии. Назовем аппроксимационной стратегией первого (второго) игрока функцию $\mathbf{u}(t, \mathbf{x}, \varepsilon) \in P$ ($\mathbf{v}(t, \mathbf{x}, \varepsilon) \in Q$), где $\varepsilon > 0$ — малый параметр. Пусть выбраны ε, позиция $\{t_*, \mathbf{x}_*\}$, отрезок $[t_*, t^*]$ и разбиение $\Delta = \{\tau_0 = t_*, \tau_{i+1} > \tau_i, \tau_m = t^*\}$. Назовем $\{\varepsilon, \Delta\}$-движением $\mathbf{x}_\Delta^\varepsilon(t_*[\cdot]t^*)$, порожденным стратегией $\mathbf{u}(t, \mathbf{x}, \varepsilon)$, абсолютно непрерывное решение пошагового уравнения

$$\mathbf{x}_\Delta^{\boldsymbol\cdot\varepsilon}[t] = A(t)\,\mathbf{x}_\Delta^\varepsilon[t] + B(t)\,\mathbf{u}(\tau_i, x_\Delta^\varepsilon[\tau_i], \varepsilon) + C(t)\,\mathbf{v}[t]$$
$$\mathbf{x}_\Delta^\varepsilon[t_*] = \mathbf{x}_*, \quad \tau_i \leqslant t \leqslant \tau_{i+1}, \quad i = 0, 1, \ldots, m-1$$

где функция $\mathbf{v}[t] \in Q$ может быть любой измеримой функцией. Аналогичным образом определяются $\{\varepsilon, \Delta\}$-движения, порожденные стратегией $\mathbf{v}(t, \mathbf{x}, \varepsilon)$.

Будем рассматривать только движения $\mathbf{x}(t_*[\cdot]t^*)$ и $\mathbf{x}_\Delta^\varepsilon(t_*[\cdot]t^*)$, начинающиеся в областях

(1.6) $$\mathbf{x}[t_*] = \mathbf{x}_* \in G[t_*] = \{|\mathbf{x}| \leqslant r(t_*)\}$$
$$r(t_*) = [r_0 + (f+g)/L] \exp L[t_* - t_0] - (f+g)/L$$
$$f = \max |B(t)\mathbf{u}|, \quad g = \max |C(t)\mathbf{v}|, \quad L = \max |A(t)|$$

где $|\mathbf{x}|$ — евклидова норма вектора $\mathbf{x}$, $|A(t)|$ — евклидова норма матрицы $A(t)$.

Для таких движений при всех $t \in [t_*, t^*]$ справедливо включение $\mathbf{x}[t] \in G[t]$.

Скажем, что стратегия $\mathbf{u}(t, \mathbf{x}, \varepsilon)$ аппроксимирует стратегию $S_u(t, \mathbf{x})$, если для любого $\zeta > 0$ найдутся $\varepsilon(\zeta) > 0$ и $\delta(\zeta, \varepsilon) > 0$, такие, что для любого $\{\varepsilon, \Delta\}$-движения $\mathbf{x}_\Delta{}^\varepsilon(t_*{}^\varepsilon[\cdot]\vartheta)$, порожденного стратегией $\mathbf{u}(t, x, \varepsilon)$ при $\varepsilon \leqslant \varepsilon(\zeta)$ и $\max_i(\tau_{i+1} - \tau_i) \leqslant \delta(\zeta, \varepsilon)$, найдется движение $\mathbf{x}(t_*[\cdot]\vartheta)$, порожденное стратегией $S_u(t, \mathbf{x})$, удовлетворяющее условиям

$$|\gamma(\mathbf{x}_\Delta{}^\varepsilon(t_*{}^\varepsilon[\cdot]\vartheta), \mathbf{u}(t_*{}^\varepsilon[\cdot]\vartheta), \mathbf{v}(t_*{}^\varepsilon[\cdot]\vartheta)) -$$
$$- \gamma(\mathbf{x}(t_*[\cdot]\vartheta), \mathbf{u}^*(t_*[\cdot]\vartheta), \mathbf{v}^*(t_*[\cdot]\vartheta))| \leqslant \zeta$$
$$|t_* - t_*{}^\varepsilon| \leqslant \zeta, \quad \max_{\tau_* \leqslant t \leqslant \vartheta} |\mathbf{x}_\Delta{}^\varepsilon[t] - \mathbf{x}[t]| \leqslant \zeta, \quad \tau_* = \max(t_*{}^\varepsilon, t_*)$$

Справедливо следующее утверждение

Теорема 1.1. Рассматриваемая игра на минимакс функционала (1.1), (1.5) имеет седловую точку $\{S_u{}^\circ, S_v{}^\circ\}$. Цена игры $\rho^\circ(t, \mathbf{x})$ удовлетворяет условию Липшица по t и $\mathbf{x}$ в области $G = \{\mathbf{x} \in G[t], t_0 \leqslant t \leqslant \vartheta\}$. Оптимальные стратегии $S_u{}^\circ$ и $S_v{}^\circ$ аппроксимируются подходящими оптимальными стратегиями $\mathbf{u}^\circ(t, \mathbf{x}, \varepsilon)$ и $\mathbf{v}^\circ(t, \mathbf{x}, \varepsilon)$.

Аппроксимационные стратегии $\mathbf{u}^\circ(t, \mathbf{x}, \varepsilon)$ и $\mathbf{v}^\circ(t, \mathbf{x}, \varepsilon)$ строятся в соответствии со схемами из работ [7, 8] как стратегии, экстремальные к функции $\rho(t, \mathbf{w}, \mathbf{w}_{n+1}) = \rho^\circ(t, \mathbf{w}) + \mathbf{w}_{n+1}$, где переменные $\mathbf{w}[t]$ и $\mathbf{w}_{n+1}[t]$, описывающие состояние w-модели, изменяются в соответствии с уравнениями

(1.7) $$\dot{\mathbf{w}} = A(t)\mathbf{w} + B(t)\mathbf{u}_* + C(t)\mathbf{v}_*, \quad \mathbf{u}_* \in P, \quad \mathbf{v}_* \in Q$$

(1.8) $$\dot{\mathbf{w}}_{n+1} = \omega(t, \mathbf{w}) + \omega_1(t, \mathbf{u}_*) + \omega_2(t, \mathbf{v}_*)$$

Здесь $\rho(t_*, \mathbf{w}_*, \mathbf{w}_{n+1*})$ — точная верхняя грань значений β, для которых в w-модели (1.7), (1.8) существует ($\beta - Q_{\{t_*, \mathbf{w}_*, \mathbf{w}_{n+1*}\}}$)-процедура [7], обеспечивающая неравенство $\mathbf{w}_{n+1}[\vartheta] + \sigma(\mathbf{w}[\vartheta]) > \beta$ для всякого движения $\{\mathbf{w}[t], \mathbf{w}_{n+1}[t]\}$, $t_* \leqslant t \leqslant \vartheta$, порожденного этой Q-процедурой из начальной позиции $\{t_*, \mathbf{w}_*, \mathbf{w}_{n+1*}\}$.

При этом сопутствующая точка [8] в w-модели $\{\mathbf{w}(t, \mathbf{x}, \varepsilon), c(t, \mathbf{x}, \varepsilon)\}$, отвечающая текущей позиции $\{t, \mathbf{x}\}$, определяется при построении стратегии $\mathbf{u}^\circ(t, \mathbf{x}, \varepsilon)$ из условия $\min_{\{\mathbf{w}, c\}} [\rho^\circ(t, \mathbf{w}) - c] = \rho^\circ[t, \mathbf{w}(t, \mathbf{x}, \varepsilon)] - c(t, \mathbf{x}, \varepsilon)$ при условии

(1.9) $$|\mathbf{x} - \mathbf{w}|^2 + c^2 \leqslant \varepsilon(1 + [t - t_0]) \exp(3L[t - t_0])$$

а при построении стратегии $\mathbf{v}^\circ(t, \mathbf{x}, \varepsilon)$ — из условия
$$\max_{\{\mathbf{w}, c\}} [\rho^\circ(t, \mathbf{w}) - c] = \rho^\circ[t, \mathbf{w}(t, \mathbf{x}, \varepsilon)] - c(t, \mathbf{x}, \varepsilon)$$
при условии (1.9).

В итоге экстремальные стратегии $\mathbf{u}^\circ(t, \mathbf{x}, \varepsilon)$ и $\mathbf{v}^\circ(t, \mathbf{x}, \varepsilon)$ определяются из условий

$$\langle \mathbf{u}^\circ(t,\mathbf{x},\varepsilon)\cdot[\mathbf{x}-\mathbf{w}(t,\mathbf{x},\varepsilon)]\rangle + c(t,\mathbf{x},\varepsilon)\,\omega_1(t,\mathbf{u}^\circ(t,\mathbf{x},\varepsilon)) = \\ = \min\nolimits_{\mathbf{u}\in P} \mathrm{Idem}\,(\mathbf{u}^\circ(t,\mathbf{x},\varepsilon)\to\mathbf{u})$$

$$\langle \mathbf{v}^\circ(t,\mathbf{x},\varepsilon)\cdot[\mathbf{x}-\mathbf{w}(t,\mathbf{x},\varepsilon)]\rangle + c(t,\mathbf{x},\varepsilon)\,\omega_2(t,\mathbf{v}^\circ(t,\mathbf{x},\varepsilon)) = \\ = \min\nolimits_{\mathbf{v}\in Q} \mathrm{Idem}\,(\mathbf{v}^\circ(t,\mathbf{x},\varepsilon)\to\mathbf{v})$$

где $\langle \mathbf{a}\cdot\mathbf{b}\rangle$ — скалярное произведение векторов $\mathbf{a}$ и $\mathbf{b}$. Здесь и далее Idem в правой части равенства означает выражение, совпадающее с левой частью этого равенства при указанной в скобках замене символов.

Стратегии $S_u^\circ(t,\mathbf{x})$ и $S_v^\circ(t,\mathbf{x})$ определяются следующим образом. Стратегия $S_u^\circ(t,\mathbf{x})$ назначает для текущей позиции $\{t,\mathbf{x}\}$ множество всех пар $\mathrm{s}=\{\mathbf{u}^*,\mathbf{v}^*\}$, $\mathbf{u}^*\in P^*(t)$, $\mathbf{v}^*\in Q^*(t)$, удовлетворяющих условию

$$\varlimsup_{\tau\to t-0}\frac{\rho^\circ(\tau,\mathbf{y}[\tau])-\rho^\circ(t,\mathbf{x})}{\tau-t}+\omega(t,\mathbf{x})+u_{r+1}^*+v_{s+1}^*\leqslant 0 \tag{1.10}$$

а стратегия $S_v^\circ(t,\mathbf{x})$ назначает для позиции $\{t,\mathbf{x}\}$ множество всех пар $\mathrm{s}=\{\mathbf{u}^*,\mathbf{v}^*\}$, $\mathbf{u}^*\in P^*(t)$, $\mathbf{v}^*\in Q^*(t)$, удовлетворяющих условию

$$\varliminf_{\tau\to t-0}\frac{\rho^\circ(\tau,\mathbf{y}[\tau])-\rho^\circ(t,\mathbf{x})}{\tau-t}+\omega(t,\mathbf{x})+\mathbf{u}_{r+1}^*+v_{s+1}^*\geqslant 0 \tag{1.11}$$

Здесь $\mathbf{y}[\tau]$ — функция, определенная равенством

$$\mathbf{y}[\tau]=\mathbf{x}+(\tau-t)\,[A(t)\,\mathbf{x}+B(t)\,\mathbf{u}+C(t)\,\mathbf{v}],\quad \tau\leqslant t$$

Условия (1.10), (1.11) заменяют здесь, очевидно, известные соотношения динамического программирования [13, 14], которые имели бы место в случае дифференцируемой цены $\rho^\circ(t,\mathbf{x})$, удовлетворяющей уравнению в частных производных из метода динамического программирования.

2. Построение цены игры $\rho^\circ(t,\mathbf{x})$ на основе Q-процедур, указанное в п. 1, не является, вообще говоря, эффективным. Стало быть, не является эффективным и построение этим путем стратегий $\mathbf{u}^\circ(t,\mathbf{x},\varepsilon)$, $\mathbf{v}^\circ(t,\mathbf{x},\varepsilon)$, $S_u^\circ(t,\mathbf{x})$, $S_v^\circ(t,\mathbf{x})$.

Более эффективен способ построения цены игры $\rho^\circ(t,\mathbf{x})$ и оптимальных стратегий на основе вспомогательных программных конструкций [6, 8]. Однако этот метод дает искомое решение лишь при определенных условиях регулярности [8]. Ниже описывается некоторое развитие метода программных конструкций, позволяющее охватить более широкий круг задач. При этом, однако, в развиваемые вспомогательные программные конструкции вводится некоторый дополнительный элемент в виде подходящего вероятностного процесса. Будем предполагать функции $\omega(t,\mathbf{x})$ и $\sigma(\mathbf{x})$ выпуклыми по $\mathbf{x}$.

Итак, рассмотрим $\mathbf{w}^*$-модель, текущее состояние которой $\mathbf{w}^*=\{\mathbf{w}=\{w_1^*,\ldots,w_n^*\},\ w_{n+1}^*\}$ описывается в соответствии с (1.7), (1.8) уравнениями

$$\dot{\mathbf{w}}=A(t)\,\mathbf{w}+B(t)\,\mathbf{u}+C(t)\,\mathbf{v}$$

$$\dot{w}_{n+1}^*=\omega(t,\mathbf{w})+u_{r+1}^*+v_{s+1}^*,\quad \mathbf{u}^*\in P^*(t),\quad \mathbf{v}^*\in Q^*(t)$$

Пусть выбрана некоторая исходная позиция $\{t_*,\mathbf{w}_*^*\}=\{t_*,\{\mathbf{w}_*,0\}\}$, $t_*\leqslant\vartheta$, $\mathbf{w}_*\in G[t_*]$. Разобьем отрезок $[t_*,\vartheta]$ на части точками $t_i=t_*+$

$+ [\vartheta - t_*] \cdot (i-1)/k$, $i = 1, 2, \ldots, k$, где k — некоторое достаточно большое целое число. Рассмотрим последовательность ξ векторных независимых случайных величин $\{\xi_j^{(i)}, j = 1, 2, \ldots, n\}$, $i = 1, 2, \ldots, k$. Каждая из величин $\xi_j^{(i)}$ может принимать одно из двух значений $\xi_j^{(i)+} = 1$ и $\xi_j^{(i)-} = -1$ с равными вероятностями $p^+ = {}^1/_2$ и $p^- = {}^1/_2$. Назовем стохастической неупреждающей программой $\mathbf{v}^*(t, \xi)$ функцию от t и $\xi = \{\xi_j^{(i)}\}$ со значениями $\mathbf{v}^*(t_i, \xi) \in Q^*(t_i)$, которая обладает следующим свойством: при $t_i \leqslant t < t_{i+1}$, $i = 1, 2, \ldots, k$, $t_{k+1} = \vartheta$ имеем $\mathbf{v}^*(t, \xi) = \mathbf{v}^*(t_i, \xi[t_*, t_i])$, где символ $\xi[t_*, t_i]$ означает реализацию $\{\xi_j^{(s)}, j = 1, 2, \ldots, n, s = 1, 2, \ldots, i\}$. Аналогичным образом определяется стохастическая неупреждающая программа $\mathbf{u}^*(t, \xi) \in P^*(t)$.

Пусть дана исходная позиция $\{t_*, \mathbf{w}_*\}$, выбрано определенное значение k и пара программ $\{\mathbf{u}^*(\cdot, \cdot), \mathbf{v}^*(\cdot, \cdot)\}$. Эти данные определят случайный процесс $\mathbf{w}(t_*[\cdot, \xi]\vartheta)$, который является пошаговым решением дифференциального уравнения

$$\mathbf{w}^{\cdot} = A(t)\,\mathbf{w} + B(t)\,\mathbf{u}(t, \xi) + C(t)\,\mathbf{v}(t, \xi) \tag{2.1}$$

с начальным условием $\mathbf{w}[t_*] = \mathbf{w}_*$. Этот процесс $\mathbf{w}(t_*[\cdot, \xi]\vartheta)$ и управления $\mathbf{u}_{r+1}^*(t, \xi)$, $\mathbf{v}_{s+1}^*(t, \xi)$ определяют случайную величину функционала $\gamma(\xi)$ (1.5):

$$\begin{aligned}\gamma(\xi) &= \gamma(\mathbf{w}(t_*[\cdot, \xi]\vartheta), \mathbf{u}^*(t_*[\cdot, \xi]\vartheta), \mathbf{v}^*(t_*[\cdot, \xi]\vartheta)) = \\ &= \mathbf{w}_{n+1}^*[\vartheta] + \sigma(\mathbf{w}[\vartheta])\end{aligned} \tag{2.2}$$

Рассмотрим функцию

$$\rho_*(t_*, \mathbf{w}_*) = \lim_{k\to\infty} \max_{\mathbf{v}^*(\cdot,\cdot)} \min_{\mathbf{u}^*(\cdot,\cdot)} M\{\gamma(\xi)\}, \tag{2.3}$$

где символ $M\{\gamma\}$ означает математическое ожидание.

Определение функции $\rho_*(t, \mathbf{w})$ (2.3) является корректным. В самом деле, минимум и максимум в правой части (2.3) действительно достигается на некоторых программах $\mathbf{u}^*(\cdot, \cdot)$ и $\mathbf{v}^*(\cdot, \cdot)$, поскольку $M\{\gamma(\xi)\}$ — непрерывная функция конечного числа переменных, заданная на компакте. Существование предела в (2.3) устанавливается в ходе доказательства следующей теоремы 2.1.

Теорема 2.1. Функция $\rho_*(t, \mathbf{w})$ (2.3) есть цена $\rho^\circ(t, \mathbf{w})$ позиционной дифференциальной игры, рассмотренной в п. 1.

Теорема доказывается следующим образом. В области

$$|\mathbf{z}| \leqslant 2r(\vartheta), \quad t_* \leqslant t \leqslant \vartheta \tag{2.4}$$

где $r(\vartheta)$ вычисляется по (1.6), построим функцию

$$\begin{aligned}H_\alpha(\mathbf{p}, \mathbf{z}, t) = \min_{u^*\in P^*(t)} \max_{v^*\in Q^*(t)} [\langle \mathbf{p}\cdot[A(t)\,\mathbf{z} + B(t)\,\mathbf{u} + \\ + C(t)\,\mathbf{v}]\rangle + \omega(t, \mathbf{z}) + u_{r+1}^* + v_{s+1}^* - \alpha|\mathbf{v}^*|^2]\end{aligned}$$

где α — некоторое малое положительное число. Далее построим функцию $F_\alpha(\mathbf{p}, \mathbf{z}, t)$, которая имеет производные всех порядков, удовлетворяет условию Липшица по первому аргументу и вне достаточно большой области G^* в пространстве $\{t, \mathbf{z}\}$, содержащей область (2.4), обращается в нуль.

Кроме того, пусть при этом выполняется условие

$$| H_\alpha(\mathbf{p}, \mathbf{z}, t) - F_\alpha(\mathbf{p}, \mathbf{z}, t) | \leqslant \alpha$$

при всех значениях аргументов $\mathbf{p}$ и $\mathbf{z}$, t из области (2.4).

Рассмотрим дифференциальное уравнение в частных производных для некоторой функции $\rho_\alpha(t, \mathbf{z})$:

$$\frac{\partial\rho_\alpha}{\partial t} + \frac{\alpha^2}{2}\sum_{i=1}^{n}\frac{\partial^2\rho_\alpha}{\partial z_i^2} + F_\alpha(\operatorname{grad}_z \rho_\alpha, \mathbf{z}, t) = 0 \tag{2.5}$$

Пусть $\sigma(\mathbf{z}, \alpha)$ — выпуклая по $\mathbf{z}$ при $| \mathbf{z} | \leqslant 2r(\vartheta)$ функция, имеющая производные всех порядков, удовлетворяющая условию $| \sigma(\mathbf{z}) - \sigma(\mathbf{z}, \alpha) | \leqslant \alpha$ при $| \mathbf{z} | \leqslant 2r(\vartheta)$ и обращающаяся в нуль при всех достаточно больших значениях $| \mathbf{z} |$. Уравнение (2.5) при краевом условии

$$\rho_\alpha(\vartheta, \mathbf{z}) = \sigma(\mathbf{z}, \alpha) \tag{2.6}$$

имеет [15] решение $\rho_\alpha(t, \mathbf{z})$, которое в любой наперед выбранной области $| \mathbf{z} | \leqslant R$, $t_* \leqslant t \leqslant \vartheta$ имеет непрерывные частные производные $\partial\rho_\alpha/\partial t$, $\partial\rho_\alpha/\partial z_i$, $\partial^2\rho_\alpha/\partial z_i\partial z_j$, $i, j = 1, \ldots, n$.

Можно проверить, подобно тому, как это сделано в [16], что для любой позиции $\{t_*, \mathbf{w}_*\}$ из области $| \mathbf{w}_* | \leqslant r(t_*)$, $t_0 \leqslant t_* \leqslant \vartheta$ справедливо предельное соотношение

$$\lim_{\alpha\to 0}\rho_\alpha(t_*, \mathbf{w}_*) = \rho^0(t_*, \mathbf{w}_*) \tag{2.7}$$

Выберем какую-нибудь подпоследовательность чисел $\{k_j, j = 1, 2, \ldots\}$, для которой существует предел

$$\lim_{k_j\to\infty}\max_{\mathbf{v}^*(\cdot,\cdot)}\min_{\mathbf{u}^*(\cdot,\cdot)} M\{\gamma(\xi)\} = \rho^*(t_*, \mathbf{w}_*) \tag{2.8}$$

Зададимся некоторым значением $\varepsilon > 0$. Выберем при каком-либо значении k_j какую-нибудь пару программ $\{\mathbf{v}^*(\cdot, \cdot), \mathbf{u}^*(\cdot, \cdot)\}$, удовлетворяющую условию

$$M\{\gamma(\xi)\} \leqslant \rho^*(t_*, \mathbf{w}_*) + \varepsilon \tag{2.9}$$

где случайная величина $\gamma(\xi)$ (2.2) определена случайным решением $\mathbf{w}(t_*[\cdot, \xi]\vartheta)$ уравнения (2.1) и управлениями $u_{r+1}^*(t, \xi)$, $v_{s+1}^*(t, \xi)$. Для любого $\varepsilon > 0$ найдется $k(\varepsilon)$, такое, что при $k_j \geqslant k(\varepsilon)$ для всякой программы $\mathbf{v}^*(\cdot, \cdot)$ найдется программа $\mathbf{u}^*(\cdot, \cdot)$, такая, что выполнится условие (2.9), как это следует из (2.8).

Выбранной паре программ $\{\mathbf{v}^*(\cdot, \cdot), \mathbf{u}^*(\cdot, \cdot)\}$ поставим в соответствие порожденное ею случайное движение $\mathbf{z}(t_*[\cdot, \xi, \alpha]\vartheta)$, $\mathbf{z}[t_*, \xi, \alpha] = \mathbf{w}_*$, которое является пошаговым решением стохастического дифференциального уравнения ($\delta(t)$ — δ-функция Дирака)

$$\begin{gathered}\mathbf{z}^{\cdot} = A(t)\mathbf{z} + B(t)\mathbf{u}(t, \xi) + C(t)\mathbf{v}(t, \xi) + \\ + \sum_{t_*\leqslant t_i\leqslant t}\alpha\,[(\vartheta - t_*)/k_j]^{1/2}\,\xi^{(i)}\delta(t - t_i) \\ (k_j \geqslant k(\varepsilon),\ \xi^{(i)} = \{\xi_j^{(i)},\ j = 1, 2, \ldots, n\})\end{gathered} \tag{2.10}$$

Это движение $z(t_*[\cdot, \xi, \alpha]\vartheta)$ породит некоторую стохастическую неупреждающую программу $\mathbf{v}^*(\cdot, \cdot, \alpha)^*$, которая определяется из условия

$$\text{(2.11)} \quad \mathbf{v}^*(t, \xi, \alpha)^* = \mathbf{v}^*(t_i, \xi, \alpha)^*, \; t_i \leqslant t < t_{i+1}$$

$$\langle \operatorname{grad}_z \rho_\alpha (t_i, z[t_i, \xi, \alpha]) \cdot C(t_i)\, \mathbf{v}^*(t_i, \xi, \alpha)^* \rangle + v^*_{s+1}(t_i, \xi, \alpha)^* -$$
$$- \alpha \, |\mathbf{v}^*(t_i, \xi, \alpha)^*|^2 = \max_{\mathbf{v}^* \in Q^*(t_i)} \operatorname{Idem} (\mathbf{v}^*(t_i, \xi, \alpha)^* \to \mathbf{v}^*)$$

Такая программа $\mathbf{v}^*(\cdot, \cdot, \alpha)^*$ (2.11) при введенных условиях единственна. Этой программе в свою очередь поставим в соответствие программу $\mathbf{u}^*(\cdot, \cdot, \alpha)^*$, так, чтобы для пары $\{\mathbf{u}^*(\cdot, \cdot, \alpha)^*, \mathbf{v}^*(\cdot, \cdot, \alpha)^*\}$ было выполнено условие (2.9). Таким путем по аналогии с процедурой из работ [6, 17, 18] получаем многозначное отображение всех пар программ $\{\mathbf{u}^*(\cdot, \cdot, \alpha), \mathbf{v}^*(\cdot, \cdot, \alpha)\}$, удовлетворяющих условию (2.9), на подмножество $\{\{\mathbf{u}^*(\cdot, \cdot, \alpha)^*, \mathbf{v}^*(\cdot, \cdot, \alpha)^*\}\}$ таких же пар программ. Как и в работах [17, 18], можно проверить, что это отображение имеет неподвижную точку. Пусть это будет пара программ $\{\mathbf{u}^*(\cdot, \cdot, \alpha)_*, \mathbf{v}^*(\cdot, \cdot, \alpha)_*\}$. Рассмотрим движение $\mathbf{z}(t_*[\cdot, \xi, \alpha]\,\vartheta)_*$, порожденное этой парой программ, как решение стохастического дифференциального уравнения (2.10) при $\mathbf{z}[t_*, \xi, \alpha] = \mathbf{w}_*$. Для этого движения управления $\mathbf{v}^*(t_i, \xi, \alpha)_*$ определяются из условия (2.11). Но тогда, опираясь на тот факт, что функция $\rho_\alpha(t, z)$ есть решение дифференциального уравнения (2.5) при краевом условии (2.6), обычными для метода динамического программирования рассуждениями получим оценку

$$\text{(2.12)} \quad M\{\gamma(\xi, \alpha)_*\} \geqslant \rho_\alpha(t_*, \mathbf{w}_*) - \eta(\alpha, k_j)$$

где $\gamma(\xi, \alpha)_* = \gamma(\mathbf{z}(t_*[\cdot, \xi, \alpha]\,\vartheta)_*, \mathbf{u}^*(\cdot, \cdot, \alpha)_*, \mathbf{v}^*(\cdot, \cdot, \alpha)_*)$ и $\eta(\alpha, k_j) \to 0$ при $k_j \to \infty$, $\alpha \to 0$.

С другой стороны, рассмотрим движение $\mathbf{w}(t_*[\cdot, \xi, \alpha]\,\vartheta)_*$, порожденное той же парой программ $\{\mathbf{u}^*(\cdot, \cdot, \alpha)_*, \mathbf{v}^*(\cdot, \cdot, \alpha)_*\}$, но уже как решение стохастического дифференциального уравнения (2.1). Для этого движения справедливо условие (2.9) при $\gamma(\xi) = \gamma(\xi, \alpha) = \gamma(\mathbf{w}(t_*[\cdot, \xi, \alpha]\,\vartheta)_*, \mathbf{u}^*(\cdot, \cdot, \alpha)_*, \mathbf{v}^*(\cdot, \cdot, \alpha)_*)$. В то же время для полученных так величин $M\{\gamma(\xi, \alpha)\}$ в (2.9) и $M\{\gamma(\xi, \alpha)_*\}$ в (2.12) справедливо соотношение

$$\text{(2.13)} \quad |M\{\gamma(\xi, \alpha)\} - M\{\gamma(\xi, \alpha)_*\}| \leqslant \zeta(\alpha, k_j)$$

где $\zeta(\alpha, k_j) \to 0$ при $k_j \to \infty$, $\alpha \to 0$.

Учитывая теперь (2.7), (2.9), (2.12), (2.13), получаем неравенство

$$\text{(2.14)} \quad \rho^\circ(t_*, \mathbf{w}_*) \leqslant \rho^*(t_*, \mathbf{w}_*)$$

Противоположное неравенство

$$\text{(2.15)} \quad \rho^\circ(t_*, \mathbf{w}_*) \geqslant \rho^*(t_*, \mathbf{w}_*)$$

установим, если будем строить для данной программы $\mathbf{v}^*(t, \xi)$ стохастическую неупреждающую программу $\mathbf{u}^*(t, \xi)$ по шагам $t_i \leqslant t < t_{i+1}$, выбирая управления $\mathbf{u}[t_i, \xi] = \mathbf{u}^\circ[t_i, w[t_i, \xi[t_*, t_{i-1}]], \varepsilon]$ в соответствии с оптимальной аппроксимационной стратегией $\mathbf{u}^\circ(t, w, \varepsilon)$.

Неравенства (2.14) и (2.15) можем получить для любой аналогичной последовательности $\{k_j\}$, для которой существует предел (2.8). Из (2.14) и (2.15) следует, что всякий такой предел $\rho^*(t_*, \mathbf{w}_*)$ должен совпадать с ценой игры ρ° (t_*, w_*). Отсюда вытекает, что предел (2.3) действительно существует и этот предел ρ_* (t_*, w_*) действительно равен цене игры ρ° (t_*, w_*). Этим и завершается доказательство теоремы 2.1.

Авторы благодарят Осипова Ю. С. и Субботина А. И. за обсуждение работы.

ЛИТЕРАТУРА

1. *Понтрягин Л. С., Мищенко Е. Ф.* Задача об убегании одного управляемого объекта от другого.— Докл. АН СССР, 1969, т. 189, № 4, с. 721—723.
2. *Пшеничный Б. Н.* Структура дифференциальных игр.— Докл. АН СССР, 1969, т. 184, № 2, с. 285—291.
3. *Меликян А. А., Черноусько Ф. Л.* Выбор моментов наблюдений в линейной игре сближения.— Изв. АН СССР. Техническая кибернетика, 1979, № 1, с. 11—16.
4. *Никольский М. С.* О некоторых дифференциальных играх с фиксированным временем.— Докл. АН СССР, 1978, т. 240, № 2, с. 272—275.
5. *Флемминг У., Ришел Р.* Оптимальное управление детерминированными и стохастическими системами. М.: Мир, 1978. 316 с.
6. *Красовский Н. Н., Субботин А. И.* Позиционные дифференциальные игры. М.: Наука, 1974. 456 с.
7. *Красовский А. Н.* О позиционном минимаксном управлении.— ПММ, 1980, т. 44, вып. 4, с. 602—610.
8. *Красовский А. Н.* Дифференциальная игра для позиционного функционала.— Докл. АН СССР, 1980, т. 253, № 6, с. 1303—1307.
9. *Красовский Н. Н., Третьяков В. Е.* Седловая точка стохастической дифференциальной игры.— Докл. АН СССР, 1980, т. 254, № 3, с. 534—539.
10. *Осипов Ю. С.* Об условиях стабильности поглощения в дифференциально-разностных играх. I, II.— В кн.: Управляемые системы. Вып. 8. Новосибирск, 1971, I. с. 13—20; II. с. 21—28. (Ин-т матем. СО АН СССР).
11. *Тарлинский С. И.* Об одной линейной дифференциальной игре сближения.— Докл. АН СССР, 1973, т. 209, № 6, с. 1303—1306.
12. *Ченцов А. Г.* Итерационная программная конструкция для дифференциальной игры с фиксированным моментом окончания.— Докл. АН СССР, 1978, т. 240, № 1, с. 36—39.
13. *Айзекс Р.* Дифференциальные игры. М.: Мир, 1967. 479 с.
14. *Беллман Р., Калаба Р.* Динамическое программирование и современная теория управления. М.: Мир, 1969. 118 с.
15. *Ладыженская О. А., Солонников В. А., Уральцева Н. Н.* Линейные и квазилинейные уравнения параболического типа. М.: Наука, 1967. 736 с.
16. *Буслаева Л. Т.* Стохастическое управление в дифференциальной игре.— ПММ, 1978, т. 42, вып. 4, с. 579—592.
17. *Барабанова Н. Н., Субботин А. И.* О классах стратегий в дифференциальных играх уклонения от встречи.— ПММ, 1971, т. 35, вып. 3, с. 385—392.
18. *Субботина Н. Н., Ченцов А. Г.* О существовании функции Беллмана в линейной дифференциальной игре.— Тр. ин-та матем. и механ. УНЦ АН СССР. Свердловск, 1979, вып. 26, с. 80—86.

Свердловск

Поступила в редакцию
18.II.1981