Прикладная математика и механика. Том 46, вып. 6, 1982

УДК 62—50

О СТОХАСТИЧЕСКОМ ПРОГРАММНОМ СИНТЕЗЕ СТРАТЕГИЙ В ДИФФЕРЕНЦИАЛЬНОЙ ИГРЕ

Красовский Н. Н.

Рассматривается дифференциальная игра [1—17], в которой стратегии формируют управляющие воздействия на основе информации об истории движения. Обсуждается вычисление цены этой игры и построение оптимальных стратегий на основе вспомогательных программных конструкций, которые содержат введенный искусственно случайный элемент. Таким образом, здесь с некоторой общей точки зрения рассматривается метод стохастического программного синтеза, предложенный для дифференциальных игр в работах [18, 19].

1. Рассмотрим систему, описываемую уравнением

$$x^{\cdot} = f(t, x, u, v), \quad u \in R, \quad v \in Q, \quad t_0 \leqslant t \leqslant \vartheta \tag{1.1}$$

где t — время, x — n-мерный фазовый вектор объекта, u — r-мерный вектор управляющего воздействия, v — s-мерный вектор помехи; функция f непрерывна по всем аргументам и удовлетворяет по x условию Липшица

$$| f(t, x^{(1)}, u, v) - f(t, x^{(2)}, u, v) | \leqslant \lambda \, | x^{(1)} - x^{(2)} |$$

R и Q — компакты; символ $|x|$ означает евклидову норму x. Пусть задан функционал

$$\gamma = \gamma(x(t_0[\cdot]\vartheta)) \tag{1.2}$$

определенный на непрерывных функциях $x(t_0[\cdot]\vartheta) = \{x[t], \ t_0 \leqslant t \leqslant \vartheta$ и непрерывный в метрике

$$\| x(t_0[\cdot]\vartheta) \| = \max_t | x[t] |$$

Содержательно задача состоит в построении такого закона управления U, который формирует воздействие u по принципу обратной связи и гарантирует возможно меньшее значение γ. Выберем в качестве информационного элемента для текущего момента времени $\tau \in [t_0, \vartheta]$ историю движения $x(t_0[\cdot]\tau)$, которая реализовалась к этому моменту. Тогда задачу можно формализовать следующим образом. Назовем стратегией всякую функцию $u(x(t_0[\cdot]\tau), \varepsilon)$, определенную для всех возможных историй $x(t_0[\cdot]\tau)$, $\tau \in [t_0, \vartheta]$ и достаточно малых значений параметра точности $\varepsilon > 0$ и удовлетворяющую условию $u(x(t_0[\cdot]\tau), \varepsilon) \in R$. Пусть выбрана некоторая стратегия $u(x(t_0[\cdot]\tau), \varepsilon)$, реализовалась история $x(t_0[\cdot]t_*)$, выбраны значение $\varepsilon > 0$ и разбиение $\Delta\{\tau_i\}$, $i = 0, \ldots, m$ для отрезка $t_* \leqslant t \leqslant \vartheta$ будущего времени, $\tau_0 = t_*, \ldots, \tau_m = \vartheta$. Тогда движение $x(t_0[\cdot]\vartheta)$, продолжающее непрерывно данную историю $x(t_0[\cdot]t_*)$, определяется при $t_* \leqslant t \leqslant \vartheta$ как решение пошагового дифференциального уравнения

$$x^{\cdot}[t] = f(t, x[t], u(x(t_0[\cdot]\tau_i), \varepsilon), v[t]), \quad \tau_i \leqslant t \leqslant \tau_{i+1}, \quad i = 0, \ldots, m-1$$

где реализация помехи $v(t_*[\cdot]\vartheta) = \{v[t] \in Q, \ t_* \leqslant t < \vartheta\}$ может оказаться любой, не зависящей от нашего выбора измеримой по Борелю функцией. Пусть символ Δ_δ, где число $\delta > 0$, означает разбиение $\Delta\{\tau_i\}$, которое удовлетворяет условию $\tau_{i+1} - \tau_i \leqslant \delta$, $i = 0, \ldots, m-1$. Д выбранной стратегии $u(\cdot) = u(x(t_0[\cdot]\tau), \varepsilon)$ и данной исходной истории $x(t_0[\cdot]t_*)$ будем называть гарантированным результатом ρ величину

$$\rho_{u(\cdot)}(x(t_0[\cdot]t_*)) = \overline{\lim_{\varepsilon\to 0}}\lim_{\delta\to 0}\sup_{\Delta_\delta}\sup_{v(t_*[\cdot]\vartheta)}\gamma(x(t_0[\cdot]\vartheta))$$

Будем называть оптимальной стратегию $u^\circ(\cdot)$, которая удовлетворяет условию

$$\rho_{u^\circ(\cdot)}(x(t_0[\cdot]t_*)) = \min_{u(\cdot)}\rho_{u(\cdot)}(x(t_0[\cdot]t_*))$$

для всякой возможной исходной истории $x(t_0[\cdot]t_*)$.

Назовем контрстратегией функцию $v(x(t_0[\cdot]\tau), u, \varepsilon)$, определенную для всех возможных историй $x(t_0[\cdot]\tau)$, достаточно малых $\varepsilon > 0$ и всех $u \in R$, удовлетворяющую условию $v(x(t_0[\cdot]\tau), u, \varepsilon) \in Q$ и измеримую по Борелю по u при каждых фиксированных $x(t_0[\cdot]\tau)$ и ε. Пусть выбрана некоторая контрстратегия $v(x(t_0[\cdot]\tau), u, \varepsilon)$, реализовалась история $x(t_0[\cdot]t_*)$, выбраны $\varepsilon > 0$ и разбиение $\Delta\{\tau_i\}$ для отрезка $[t_*, \vartheta]$. Движение $x(t_0[\cdot]\vartheta)$, продолжающее непрерывно $x(t_0[\cdot]t_*)$, определяется при $t_* \leqslant t \leqslant \vartheta$ как решение пошагового дифференциального уравнения

$$x^\cdot[t] = f(t, x[t], u[t], v(x(t_0[\cdot]\tau_i), u[t], \varepsilon)),$$
$$\tau_i \leqslant t \leqslant \tau_{i+1}, \quad i = 0,\ldots, m-1$$

где реализация $u(t_*[\cdot]\vartheta) = \{u[t] \in R, t_* \leqslant t < \vartheta\}$ может оказаться любой измеримой по Борелю функцией. Для выбранной контрстратегии $v(\cdot) = v(x(t_0[\cdot]\vartheta), u, \varepsilon)$ и данной исходной истории $x(t_0[\cdot]t_*)$ гарантированный результат ρ определяется равенством

$$\rho_{v(\cdot)}(x(t_0[\cdot]t_*)) = \underline{\lim_{\varepsilon\to 0}}\lim_{\delta\to 0}\inf_{\Delta_\delta}\inf_{u(t_*[\cdot]\vartheta)}\gamma(x(t_0[\cdot]\vartheta))$$

Будем называть оптимальной контрстратегию $v^\circ(\cdot)$, которая удовлетворяет условию

$$\rho_{v^\circ(\cdot)}(x(t_0[\cdot]t_*)) = \max_{v(\cdot)}\rho_{v(\cdot)}(x(t_0[\cdot]t_*))$$

для всякой возможной исходной истории $x(t_0[\cdot]t_*)$.

Задачи об оптимальных стратегии $u^\circ(\cdot)$ и контрстратегии $v^\circ(\cdot)$ составляют дифференциальную игру. Будем говорить, что эта игра имеет цену ρ° и седловую точку $\{u^\circ(\cdot), v^\circ(\cdot)\}$, если оптимальные $u^\circ(\cdot)$ и $v^\circ(\cdot)$ существуют и справедливо равенство

$$\rho_{u^\circ(\cdot)}(x(t_0[\cdot]t_*)) = \rho_{v^\circ(\cdot)}(x(t_0[\cdot]t_*)) = \rho^\circ(x(t_0[\cdot]t_*))$$

для всякой возможной исходной истории $x(t_0[\cdot]t_*)$.

Теорема 1.1. Рассматриваемая дифференциальная игра имеет цену и седловую точку.

Теорема доказывается по известной схеме (см., например, [4, 18, 20]). При этом оптимальные $u^\circ(\cdot)$ и $v^\circ(\cdot)$ могут строиться по цене игры ρ° в соответствии с условиями

$$\max_v \langle (x[\tau] - y_*[\tau])\cdot f(\tau, x[\tau], u^\circ, v)\rangle =$$
$$= \min_u\max_v \langle (x[\tau] - y_*[\tau])\cdot f(\tau, x[\tau], u, v)\rangle$$
$$\langle (x[\tau] - y^*[\tau])\cdot f(\tau, x[\tau], u, v^0)\rangle =$$
$$= \min_v \langle (x[\tau] - y^*[\tau])\cdot f(\tau, x[\tau], u, v)\rangle$$

Здесь символ $\langle a\cdot b\rangle$ означает скалярное произведение векторов a и b, $y_*[\tau]$ и $y^*[\tau]$ — значения в момент τ для сопутствующих историй $y_*(t_0[\cdot]\tau)$ и $y^*(t_0[\cdot]\tau)$, которые определяются из условий

(1.3) $\qquad \rho^\circ(y_*(t_0[\cdot]\tau)) = \min \rho^\circ(y(t_0[\cdot]\tau))$

(1.4) $\qquad \rho^\circ(y^*(t_0[\cdot]\tau)) = \max \rho^\circ(y(t_0[\cdot]\tau))$

при ограничении $\| x(t_0[\cdot]\tau) - y(t_0(\cdot)\tau) \| \leqslant \varepsilon \exp 2\lambda(\tau - t_0)$. Обсуждаемое доказательство теоремы 1.1 не является в общем случае конструктивным. Поэтому остается нерешенной задача эффективного вычисления ρ° и построения $u^\circ(\cdot)$ и $v^\circ(\cdot)$.

2. Рассмотрим оценку величины $\rho^\circ(x(t_0[\cdot]t_*))$, опирающуюся на некоторую вспомогательную стохастическую программную конструкцию. Пусть зафиксирована некоторая история $x(t_0[\cdot]t_*)$. Выберем для отрезка $[t_*, \vartheta]$ некоторое разбиение $\Delta\{t_i\}$, $i = 1, \ldots, k+1$, $t_1 = t_*, \ldots, t_{k+1} = \vartheta$. Свяжем с этим разбиением следующее вероятностное пространство $\{\Omega, F, P\}$. В этом пространстве элементарное событие ω есть любой набор $\omega = \{z^{(1)}, \ldots, z^{(k)}, s^{(1)}, \ldots, s^{(k)}\}$, где $z^{(j)}$ и $s^{(j)}$ — n-мерные векторы, $|s^{(j)}| \leqslant 1$, $|z^{(j)}| \leqslant K$, где K — некоторое достаточно большое число, F — борелевская σ-алгебра на Ω, вероятностная мера P порождена независимыми в совокупности равномерными распределениями случайных векторов $z^{(j)}$ и $s^{(j)}$ в соответствующих шарах $|z| \leqslant K$ и $|s| \leqslant 1$. Таким образом, полагаем, что с каждым моментом t_j связана пара случайных величин $\{z^{(j)}, s^{(j)}\}$, распределенных равномерно при $|z| \leqslant K$ и $|s| \leqslant 1$; все величины $z^{(j)}, s^{(j)}$ независимы в совокупности.

Назовем неупреждающими стохастическими программами $v(t, u, \omega)$ и $u(t, \omega)$ неупреждающие (по t) относительно $\xi[t_*, t] = \{z^{(1)}, \ldots, z^{(i)}, s^{(1)}, \ldots, s^{(i)}\}$, $t_i \leqslant t < t_{i+1}$ [21], случайные функции, измеримые по Борелю по всем их аргументам. Стало быть, эти функции удовлетворяют равенствам

$$v(t, u, \omega) = v(t, u, z^{(1)}, \ldots, z^{(i)}, s^{(1)}, \ldots, s^{(i)})$$
$$u(t, \omega) = u(t, z^{(1)}, \ldots, z^{(i)}, s^{(1)}, \ldots, s^{(i)})$$

при $t_i \leqslant t < t_{i+1}$, $i = 1, \ldots, k$, почти наверное по ω. История $x(t_0[\cdot]t_*)$, разбиение $\Delta\{t_i\}$ и какая-либо пара программ $\{u(t, \omega), v(t, u, \omega)\}$ определяют случайное движение $w(t_0[\cdot]\vartheta, \omega)$, которое продолжает непрерывно эту историю при $t_* \leqslant t \leqslant \vartheta$, как решение стохастического дифференциального уравнения

$$w^\cdot[t] = f(t, w[t, \omega], u(t, \omega), v(t, u(t, \omega), \omega)) \tag{2.1}$$

Обозначим через $v_\Delta(\cdot)$ и $u_\Delta(\cdot)$ программы, отвечающие разбиению Δ. Назовем программным максимином ρ^* величину

$$\rho^*(x(t_0[\cdot]t_*)) = \sup_\Delta \sup_{v_\Delta(\cdot)} \inf_{\Delta(\cdot)} \operatorname{vraimax}_\omega \gamma(w(t_0[\cdot]\vartheta, \omega))$$

где величина $\operatorname{vraimax}_\omega \gamma$ вычисляется для случайной величины γ на пространстве $\{\Omega, F, P\}$.

Теорема 2.1. Справедливо равенство

$$\rho^\circ(x(t_0[\cdot]t_*)) = \rho^*(x(t_0[\cdot]t_*)) \tag{2.2}$$

какова бы ни была история $x(t_0[\cdot]t_*)$.

Доказательство теоремы вытекает из следующих лемм.

Лемма 2.1. Каковы бы ни были история $x(t_0[\cdot]t_*)$, разбиение Δ, программа $v_\Delta(\cdot)$ и число $\beta > \rho^\circ(x(t_0[\cdot]t_*))$, существует программа $u_\Delta(\cdot)$, которая для соответствующего движения $w(t_0[\cdot]\vartheta, \omega)$ (2.1) с вероятностью единица обеспечивает неравенство

$$\gamma(w(t_0[\cdot]\vartheta, \omega)) < \beta$$

Справедливость леммы следует из известного свойства u-стабильности функции ρ°: каковы бы ни были история $w(t_0[\cdot]t_i, \omega)$, число $\alpha > 0$ и допустимая функция $v(t_i[\cdot]t_{i+1}, u, \omega)$, найдется допустимая функция $u(t_i[\cdot]t_{i+1}, \omega)$ (все при фиксированном значении так, что для соответствующего движения $w(t_0[\cdot]t_{i+1}, \omega)$ будет выполнено неравенство

(2.3) $$\rho^\circ(w(t_0[\cdot]t_{i+1}, \omega)) \leqslant \rho^\circ(w(t_0[\cdot]t_i, \omega)) + \alpha(t_{i+1} - t_i)$$

Так как функцию $u(t_i[\cdot]t_{i+1}, \omega)$ можно при этом выбрать измеримой по Борелю по t и ω (по t и $\{z^{(1)}, \ldots, z^{(i)}, s^{(1)}, \ldots, s^{(i)}\}$), то лемма доказывается прямо по индукции на основании неравенств (2.3), начиная от $\rho^\circ(x(t_0[\cdot]t_*)) = \rho^\circ(w(t_0[\cdot]t_*))$ и кончая $\rho^\circ(w(t_0[\cdot]\vartheta, \omega)) = \gamma(w(t_0[\cdot]\vartheta, \omega))$.

Лемма 2.2. Каковы бы ни были история $x(t_0[\cdot]t_*)$ и число $\beta < \rho^\circ(x(t_0[\cdot]t_*))$, существует разбиение Δ и программа $v_\Delta(\cdot)$, такие, что при всякой программе $u_\Delta^*(\cdot)$ для соответствующего движения $w(t_0[\cdot]\vartheta, \omega)$ (2.1) будет обеспечено неравенство

(2.4) $$P(\gamma(w(t_0[\cdot]\vartheta, \omega)) > \beta) > \zeta > 0$$

где символ $P(A)$ означает вероятность события A.

В самом деле, будем полагать шаг $\delta = \max_i(t_{i+1} - t_i)$ разбиения $\Delta\{t_i\}$ достаточно малым (оценка подходящей малости δ будет указана ниже). Определим программу $v(t, u, \omega) = v(t_i, u, z^{(i)}, s^{(i)})$, $t_i \leqslant t < t_{i+1}$, $i = 1, \ldots, k$, так, чтобы выполнялось условие

(2.5) $$\langle s^{(i)} \cdot f(t_i, z^{(i)}, u, v(t_i, u, z^{(i)}, s^{(i)}))\rangle = \min_v \langle s^{(i)} \cdot f(t_i, z^{(i)}, u, v)\rangle$$

Из теоремы об измеримом выборе [22] следует возможность построения такой функции $v(t, u, \omega)$, измеримой по аргументам $z^{(i)}, s^{(i)}$ и u при $t_i \leqslant t < t_{i+1}$, $i = 1, \ldots, k$. Рассмотрим случайное движение $w(t_0[\cdot]\vartheta, \omega) = \{w[t, \omega],\ t_0 \leqslant t \leqslant \vartheta, \omega \in \Omega\}$, продолжающее историю $x(t_0[\cdot]t_*)$ и порожденное построенной программой $v_\Delta(\cdot)$ (2.5) и какой-нибудь программой $u_\Delta(\cdot)$. Пусть $y^*[t_i, \omega]$ значение в момент t_i для сопутствующей истории $y^*(t_0[\cdot]t_i, \omega)$, которая определяется из условия (1.4), где $\tau = t_i$, $x(t_0[\cdot]t_i) = w(t_0[\cdot]t_i, \omega)$ и $\varepsilon > 0$ — некоторое достаточно малое число. Выберем некоторое значение $\alpha > 0$. Вследствие независимости случайных величин $z^{(j)}, s^{(j)}$, $j = 1, \ldots, k$ можно утверждать, что справедливо неравенство

(2.6) $$\begin{aligned} P(|z^{(i)}(\omega) - w[t_i, \omega]| \leqslant \alpha, |s^{(i)}(\omega) - (w[t_i, \omega] - \\ - y^*[t_i, \omega])| \leqslant \alpha \mid w(t_0[\cdot]t_i, \omega)) \geqslant \eta(\alpha) > 0 \end{aligned}$$

Здесь символ $P(A \mid \xi)$ означает условную вероятность события A по случайной величине (функции) ξ.

Из условий (2.5), (2.6) вследствие непрерывности функции $f(t, x, u, v)$ и вследствие свойства v-стабильности функции ρ° [4, 20] теперь выводится по известной схеме рассуждений [4, 18, 20], что найдется некоторая (случайная) история $y(t_0[\cdot]t_{i+1}, \omega)$, продолжающая историю $y^*(t_0[\cdot]t_i, \omega)$, и такая, что будет справедливо неравенство

$$\begin{aligned} &P(\|w(t_i[\cdot]t_{i+1}, \omega) - y(t_i[\cdot]t_{i+1}, \omega)\|^2 \leqslant \\ &\leqslant (w[t_i, \omega] - y^*[t_i, \omega])^2(1 + 2\lambda(t_{i+1} - t_i) + \zeta(\delta, \alpha)(t_{i+1} - t_i)), \\ &\rho^\circ(y(t_0[\cdot]t_{i+1}, \omega)) \geqslant \rho^\circ(y^*(t_0[\cdot]t_i, \omega)) - \zeta(\delta, \alpha)(t_{i+1} - t_i) \| \\ &\mid w(t_0[\cdot]t_i, \omega)) \geqslant \eta(\alpha) > 0 \end{aligned}$$

где

$$\lim \zeta(\alpha, \delta) = 0, \alpha \to 0, \delta \to 0$$

Отсюда по индукции доказывается, что, какие бы значения $\zeta^* > 0$ и $\varepsilon > 0$ ни выбрали наперед, можно указать столь малые значения $\delta > 0$ и $\alpha > 0$, что для рассматриваемого движения $w(t_0[\cdot]\vartheta, \omega)$ (2.1) найдется история $y(t_0[\cdot]\vartheta, \omega)$, для которой будет справедливо неравенство

(2.7) $$\begin{aligned} &P(\|w(t_0[\cdot]\vartheta, \omega) - y(t_0[\cdot]\vartheta, \omega)\| \leqslant \varepsilon \exp 2\lambda(\vartheta - t_0), \\ &\rho^\circ(y(t_0[\cdot]\vartheta, \omega)) \geqslant \rho^\circ(y^*(t_0[\cdot]t_*, \omega)) - \zeta^*, \| \\ &\|x(t_0[\cdot]t_*) - y^*(t_0[\cdot]t_*, \omega)\| \leqslant \varepsilon \exp 2\lambda(t_* - t_0)) \geqslant \eta^k(\alpha) \end{aligned}$$

Вследствие непрерывности ρ° и γ из (2.7) вытекает, что, каково бы ни было значение $\beta < \rho^\circ(x(t_0[\cdot]t_*))$ при выборе достаточно малых $\varepsilon > 0$ и $\zeta^* > 0$, при достаточно малых $\delta > 0$ и $\alpha > 0$ построенная программа $v_\Delta(\cdot)$ (2.4) при всякой программе $u_\Delta(\cdot)$ обеспечит неравенство (2.4). Это и доказывает лемму 2.2.

3. В общем случае вычисление цены игры ρ° через программный максимин ρ^* на основе равенства (2.2) вряд ли конструктивно. Однако в не-

которых случаях это равенство полезно для оценки ρ° и построения $u^\circ(\cdot)$ и $v^\circ(\cdot)$. Рассмотрим здесь подробнее случай, когда уравнение движения (1.1) является собственно линейным, т. е.

$$(3.1)\qquad \dot{x} = A(t)\,x + f(t, u, v),\ u \in R,\ v \in Q$$

где $A(t)$ — непрерывная матрица-функция. В этом случае можно строить вероятностное пространство $\{\Omega, F, P\}$ на основе лишь случайных величин $s^{(j)}$, полагая $\omega = \{s^{(1)}, \ldots, s^{(k)}\}$, так как в условии (2.5) в случае (3.1) величины $z^{(j)}$ по сути дела роли не играют. (Вообще, в каждом конкретном случае пространство $\{\Omega, F, P\}$ можно выбирать так или иначе на основе выбора в качестве ω реализаций той или иной случайной функции $\xi(t_0[\cdot]\vartheta, \omega)$, природа которой отвечает данной задаче.)

Выберем теперь на данном вероятностном пространстве $\{\Omega, F, P\}$ некоторое линейное нормированное пространство $L^{(2)}([t_0, \vartheta], \Omega)$ случайных функций $w(t_0[\cdot]\vartheta, \cdot) = \{w(t, \omega),\ t_0 \leqslant t \leqslant \vartheta,\ \omega \in \Omega\}$, содержащее случайные функции $w(t, \omega)$ с непрерывными (почти наверное) реализациями $w(t_0[\cdot]\vartheta, \omega)$. Предположим, что данный функционал γ (1.2) можно так или иначе распространить на реализации $w(t_0[\cdot]\vartheta, \omega)$ элементов из $L^{(2)}$ так, что можно говорить о случайной величине $\gamma[\omega] = \gamma(w(t_0[\cdot]\vartheta, \omega))$. При этом предполагаем выполненным следующее условие. Пусть $W_\beta^{(2)}$ — множество элементов $w(\cdot) = w(t_0[\cdot]\vartheta, \cdot)$ из $L^{(2)}$, которые удовлетворяют условию

$$(3.2)\qquad \operatorname{vrai\,max}_\omega \gamma(w(t_0[\cdot]\vartheta, \omega)) \leqslant \beta$$

и пусть для некоторых значений $\varepsilon > 0$ и $\zeta > 0$ для некоторого элемента $w(\cdot) \in L^{(2)}$ справедливо неравенство

$$(3.3)\qquad P(\gamma(w(t_0[\cdot]\vartheta, \omega)) \geqslant \beta + \varepsilon) \geqslant \zeta$$

Тогда для расстояния φ в $L^{(2)}$ от элемента $w(\cdot)$ до множества $W_\beta^{(2)}$ справедливо неравенство

$$(3.4)\qquad \varphi \geqslant \eta(\beta, \varepsilon, \zeta) > 0$$

В дальнейшем для определенности примем, например, что норма $\| w(\cdot) \|_{(2)}$ в $L^{(2)}$ задается равенством

$$\| w(\cdot) \|_{(2)} = \Big(M \int_{t_0}^{\vartheta} | w(t, \omega) |^2 \mu(dt)\Big)^{1/2} = \Big(\int_\Omega \int_{t_0}^{\vartheta} | w(t, \omega) |^2 \mu(dt)\, P(d\omega)\Big)^{1/2}$$

и норма в сопряженном пространстве $L_*^{(2)}([t_0, \vartheta], \Omega)$ случайных функций $l(\cdot) = l(t_0[\cdot]\vartheta, \omega) = \{l(t, \omega),\ t_0 \leqslant t \leqslant \vartheta,\ \omega \in \Omega\}$ — равенством

$$\| l(\cdot) \|_{(2)}^{*} = \Big(M \int_{t_0}^{\vartheta} | l(t, \omega) |^2 \mu(dt)\Big)^{1/2}$$

где $\mu(dt)$ — некоторая борелевская мера на отрезке $[t_0, \vartheta]$, символ M означает математическое ожидание.

Выберем при данной исходной истории $x(t_0[\cdot]t_*)$ какое-нибудь число $\beta < \rho^0(x(t_0[\cdot]t_*))$. Пусть некоторая функция $w_\beta^{(2)}(t_0[\cdot]\vartheta, \cdot) \in L^{(2)}$ удовлетворяет условию (3.2). Выберем программу $v_\Delta(\cdot)$, построенную в соответствии с условием (2.5), в предположении, что базовое пространство $\{\Omega, F, P\}$ отвечает разбиению Δ отрезка $t_* \leqslant t \leqslant \vartheta$, имеющему достаточно малый шаг δ. Перебирая все возможные стохастические неупреждающие программы $u_\Delta(\cdot)$, получим в пространстве $L^{(2)}$ область достижимости G, составленную из всех возможных случайных движений $w(t_0[\cdot]\vartheta, \cdot)$

(2.1), продолжающих историю $x(t_0[\cdot]t_*)$. Если шаг δ разбиения $\Delta\{t_i\}$ достаточно мал, то согласно лемме 2.2 область G не может содержать функцию $w^{(2)}(t_0[\cdot]\vartheta,\cdot)$, ибо при достаточно малом шаге δ для всех рассматриваемых движений $w(t_0(\cdot)\vartheta,\cdot)$ (2.1) будет выполнено условие (2.4) при некоторых $\zeta>0$ и $\varepsilon>0$. Более того, тогда для расстояния φ в $L^{(2)}$ от любого движения $w(t_0[\cdot]\vartheta,\cdot)$ до элемента $w^{(2)}(t_0[\cdot]\vartheta,\cdot)$ будет выполнено неравенство (3.4).

Известными в теории собственно линейных управляемых систем рассуждениями показывается, что замыкание в $L^{(2)}$ области G совпадает с выпуклой замкнутой оболочкой $W^{(1)}$ в $L^{(2)}$ этой области G. Стало быть, и для расстояния φ^* в $L^{(2)}$ от элемента $w^{(2)}(t_0[\cdot]\vartheta,\cdot)$ до выпуклой замкнутой оболочки области G также будет справедливо неравенство.

$$\varphi^*\geqslant\eta(\beta,\ \varepsilon,\ \zeta)>0 \tag{3.5}$$

Отсюда опять на основе известных рассуждений из теории собственно линейных систем [4], опирающихся на теоремы об отделении выпуклых множеств в $L^{(2)}$, выводится неравенство (штрих означает транспонирование)

$$\begin{aligned}\varphi(x(t_0[\cdot]t_*),\Delta,w^{(2)}(t_0[\cdot]\vartheta,\cdot))&=\\
=\sup_{l(\cdot)}\Big[M\int_{t_0}^{t_*}&<l(t,\omega)\cdot x[t]>\mu(dt)+\\
+M\int_{t_*}^{\vartheta}&<l(t,\omega)\cdot X(\vartheta,t)\,x[t_*]>\mu(dt)+\\
+M\int_{t_*}^{\vartheta}&\min_u<M_\tau\Big[\int^{\vartheta}X'(\vartheta,t)\,l(t,\omega)\,\mu(dt)\Big]\cdot f(\tau,u,v(\tau,u,\omega))>d\tau-\\
-M\int_{t_*}^{\vartheta}&<l(t,\omega)\cdot w^{(2)}(t,\omega)>\mu(dt)\Big]\geqslant\eta(\beta,\varepsilon,\zeta),\ \|l(\cdot)\|^*_{(2)}\leqslant 1\end{aligned} \tag{3.6}$$

ибо величина φ (3.6) и будет расстоянием φ^*. Здесь $X(t,t^*)$ — фундаментальная матрица решений однородного уравнения $\dot{x}=A(t)\,x$ и символ $M_\tau[\xi(\omega)]$ означает условное математическое ожидание

$$M_\tau[\xi(\omega)]=M[\xi(\omega)\,|\,s^{(1)}(\omega),\ldots,s^{(i)}(\omega)],\quad t_i\leqslant\tau<t_{i+1},\quad i=1,\ldots,k$$

Наоборот, если некоторая функция $w^{(2)}(t_0[\cdot]\vartheta,\ \cdot)\in W^{(1)}$, то $\varphi(x(t_0[\cdot]t_*),\ \Delta,\ w^{(2)}(t_0[\cdot]\vartheta,\ \cdot))=0$.

Из этих соотношений с учетом лемм 2.1 и 2.2 заключаем, что программный максимин $\rho^*(x(t_0[\cdot]t_*))$ и равная ему цена игры $\rho^\circ(x(t_0[\cdot]t_*))$ — верхняя грань тех чисел β, для которых справедливо неравенство

$$\sup_\Delta\inf_{w^{(2)}(\cdot)}\varphi(x(t_0[\cdot]t_*),\ \Delta,\ w^{(2)}(t_0[\cdot]\vartheta,\ \cdot))>0 \tag{3.7}$$

при $w^{(2)}(t_0[\cdot]\vartheta,\ \cdot)\in W_\beta^{(2)}$.

Если множества $W_\beta^{(2)}$ выпуклы при всяком β, то величина $\rho^*=\rho^\circ$ — верхняя грань тех β, для которых справедливо неравенство

$$\sup_\Delta\varphi(x,\ (t_0[\cdot]t_*),\ \Delta,\ W_\beta^{(2)})>0 \tag{3.8}$$

где величина φ отличается от величины φ из (3.6) лишь в последнем слагаемом, которое для φ из (3.8) имеет вид

$$\xi=-\sup_{w^{(2)}(\cdot)}M\int_{t_*}^{\vartheta}<l(t,\omega)\cdot w^{(2)}(t,\omega)>\mu(dt)$$

при $w^{(2)}(t_0[\cdot]\vartheta,\ \cdot)\in W_\beta^{(2)}$.

Таким образом, задача сводится к вычислению величины φ для (3.7) или для (3.8). Задача вычисления φ есть задача математического программирования на максимум вогнутого по $l(\cdot)$ функционала при ограничении $\| l(\cdot) \|_{(2)^*} \leqslant 1$. Такая задача не осложнена принципиальными трудностями, однако практические вычисления оказываются часто слишком трудоемкими.

Выше рассмотрен случай, когда исходная дифференциальная игра формализуется в классах {стратегии — контрстратегии}. Совершенно так же можно доказать теорему, подобную теореме 2.1, и в случае, когда игра формализуется в классах смешанных стратегий. При этом лишь в программной стохастической конструкции программы функции $u(t, \omega)$ и $v(t, u, \omega)$ заменяются на программы меры $\mu(t, \omega)$ на R и $\nu(t, \omega)$ на Q.

Сравним в заключение величины φ, которые участвуют в стохастической программной конструкции в случае выпуклых множеств $W_\beta^{(2)}$ с соответствующими величинами в аналогичных задачах, рассмотренных в [18, 19]. Отличие определяется тем, что здесь вычисление базируется всегда на одной и той же универсальной программе $v_\Delta(\cdot)$ (2.5), в то время как в работах [18, 19] программа $v(\cdot)$ выбирается каждый раз как экстремальная программа, отвечающая обстоятельствам задачи на максимум по $l(\cdot)$.

Переход к универсальной программе $v_\Delta(\cdot)$ улучшает некоторые качества задачи о вычислении φ: максимизируемая величина становится функцией, вогнутой по $l(\cdot)$, и т. д. В то же время сама величина φ утрачивает некоторые полезные свойства. Так, вообще говоря, величина φ теряет свойство u-стабильности, оно остается лишь у величины ρ^*. Это получается потому, что, вообще говоря, не предполагается, что природа пространства $L^{(2)}$ отвечает природе функционала γ и природе оценки отклонений случайной величины $\gamma[\omega]$ по величине $\mathrm{vraimax}_\omega \gamma[\omega]$. Так, выше для определенности выбрано гильбертово пространство $L^{(2)}$. Это определяет удобное сопряженное пространство $L_*^{(2)}$. Но во многих случаях величине φ можно вернуть свойство u-стабильности и в случае универсальной программы $v(\cdot)$, если подбирать метрику в $L^{(2)}$ в соответствии с природой функционала γ и с оценкой отклонения $\gamma[\omega]$ по величине $\mathrm{vraimax}_\omega \gamma[\omega]$.

Например, если функционал $\gamma(x(t_0[\cdot]\vartheta)) = \gamma(x[\vartheta])$ и величина $\gamma(x)$ имеет смысл какой-нибудь нормы $\| x \|$ в пространстве $\{x\}$, то в качестве нормы $\| w(\cdot) \|_{(2)} = \| w(\vartheta, \cdot) \|_{(2)}$ можно выбрать величину $\| w(\vartheta, \cdot) \|_{(2)} = \mathrm{vrai\,max}_\omega |w(\vartheta, \omega)\|$. Тогда и величина φ приобретает подходящее свойство u-стабильности. Но вычисление φ осложняется тогда тем, что сопряженным пространством оказывается уже пространство аддитивных функций $\lambda(A)$ от подмножеств $A \in \Omega$. Однако это не обязательно слишком осложняет дело, так как часто потом задача сводится опять к подходящим функциям $l(\omega)$ от точек $\omega \in \Omega$. Если величина γ не имеет смысл нормы, то опять можно стремиться вернуть величине φ свойство u-стабильности, определяя величину φ не через расстояние в $L^{(2)}$, но на основе оценок, которые определяются функциями γ_* (или функционалами γ_*), сопряженными должным образом с величиной γ.

ЛИТЕРАТУРА

1. *Айзекс Р.* Дифференциальные игры. М.: Мир, 1967, 479 с.
2. *Батухтин В. Д.* Экстремальное прицеливание в нелинейной игре сближения.— Докл. АН СССР, 1972, т. 207, № 1, с. 11—14.
3. *Гусятников П. Б., Никольский М. С.* Об оптимальности времени преследования.— Докл. АН СССР, 1969, т. 184, № 3, с. 518—521.

4. *Красовский Н. Н., Субботин А. И.* Позиционные дифференциальные игры. М.: Наука, 1974, 456 с.
5. *Кряжимский А. В.* К теории позиционных дифференциальных игр сближения — уклонения.— Докл. АН СССР, 1978, т. 239, № 4, с. 779—782.
6. *Куржанский А. Б.* Управление и наблюдение в условиях неопределенности. М.: Наука, 1977. 392 с.
7. *Мищенко Е. Ф., Понтрягин Л. С.* Линейные дифференциальные игры.— Докл. АН СССР, 1967, т. 174, № 1, с. 27—29.
8. *Осипов Ю. С.* Позиционное управление в параболических системах.— ПММ, 1977, т. 41, вып. 2, с. 195—201.
9. *Петров Н. Н.* О существовании значения игры преследования.— Докл. АН СССР, 1970, т. 190, № 6, с. 1289—1291.
10. *Петросян Л. А.* Дифференциальные игры преследования. Л.: Изд-во ЛГУ, 1977, 222 с.
11. *Понтрягин Л. С.* К теории дифференциальных игр.— Успехи матем. наук, 1966, т. 21, вып. 4, с. 219—274.
12. *Пшеничный Б. Н.* Структура дифференциальных игр.— Докл. АН СССР, 1969, т. 184, № 2, с. 285—287.
13. *Субботин А. И., Ченцов А. Г.* Оптимизация гарантии в задачах управления. М.: Наука, 1981. 287 с.
14. *Черноусько Ф. Л., Меликян А. А.* Игровые задачи управления и поиска. М.: Наука, 1978. 270 с.
15. *Чикрий А. А.* Квазилинейная задача сближения с участием нескольких лиц.— ПММ, 1979, т. 43, вып. 3, с. 451—455.
16. *Fleming W. H.* A note on differential games of prescribed duration.— Ann. Math. Studies, 1957, v. 39, p. 407—412.
17. *Ho Y. G., Bryson A. E., Baron S.* Differential games and optimal pursuit—evasion strategies.— IEEE Trans. Automat. Control, 1965, v. 10, No. 4, p. 385—389.
18. *Красовский А. Н., Красовский Н. Н., Третьяков В. Е.* Стохастический программный синтез для детерминированной позиционной дифференциальной игры.— ПММ, 1981, т. 45, вып. 4, с. 579—586.
19. *Красовский Н. Н., Третьяков В. Е.* Стохастический программный синтез для позиционной дифференциальной игры.— Докл. АН СССР, 1981, т. 259, № 1, с. 24—27.
20. *Красовский А. Н.* Дифференциальная игра для позиционного функционала.— Докл. АН СССР, 1980, т. 253, № 6, с. 1303—1307.
21. *Липцер Р. Ш., Ширяев А. Н.* Статистика случайных процессов. М.: Наука, 1974. 696 с.
22. *Флиминг У., Ришел Р.* Оптимальное управление детерминированными и стохастическими системами. М.: Мир, 1978. 316 с.

Свердловск

Поступила в редакцию
22.VI.1982