Прикладная математика и механика. Том 36, 1972

УДК 62—50

ОБ ОДНОЙ ИГРОВОЙ ЗАДАЧЕ КОНФЛИКТНОГО УПРАВЛЕНИЯ

Н. Н. Субботина

(Свердловск)

Рассматривается дифференциальная игра наведения — уклонения, решение которой требуется найти в классе чистых позиционных стратегий. Показано, что в этой задаче введение информационной дискриминации противника существенно искажает содержание исходной игровой проблемы.

Известно [1-3], что дифференциальная игра наведения — уклонения имеет седловую точку в классе чистых позиционных стратегий, если правая часть уравнения, описывающего динамику системы, удовлетворяет условию

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

где максимум и минимум вычисляются по допустимым значениям параметров u и v; s — произвольный n-мерный вектор, символ штрих означает транспонирование. Если же указанное условие нарушается, то, вообще говоря, в классе стратегий не существует ситуаций равновесия. Здесь результат игры существенно зависит от того, имеют ли игроки информацию об управлениях, реализующихся в системе. Типична ситуация, когда игроки не располагают такой информацией; в этом случае представляет интерес задача об отыскании позиционных минимаксных и максиминных чистых стратегий игроков. Ниже с использованием результатов работ [5,6,9] построены такие стратегии в одном примере конфликтного управления.

1. Физическое содержание исследуемой задачи состоит в следующем. Имеется материальная точка, движущаяся в горизонтальной плоскости. Движением этой точки управляют два игрока, которые формируют управляющие воздействия — двумерные векторы $u\,[t]$ и $v\,[t]$. Выбор управления $u\,[t]$ подчинен первому игроку, вектор $v\,[t]$ выбирается вторым игроком, реализации управлений удовлетворяют ограничениям

$$\|\,u\,[t]\,\| \leqslant \mu, \qquad \|\,v\,[t]\,\| \leqslant \nu \tag{1.1}$$

Здесь и в дальнейшем $\|\,x\,\|$ означает евклидову норму вектора x. В системе управления имеется люфт, поэтому вместо управляющей силы $w\,[t] = u\,[t] - v\,[t]$ к точке прикладывается некоторая сила $w_*\,[t] = u_*\,[t] - v_*\,[t]$, где векторы $u_*\,[t]$ и $v_*\,[t]$ отличаются от векторов $u\,[t]$ и $v\,[t]$ поворотом на некоторые углы $\alpha\,[t]$ и $\beta\,[t]$. Причем помехи $\alpha\,[t]$ и $\beta\,[t]$ могут изменяться в пределах

$$|\,\alpha\,[t]\,| \leqslant \alpha_0 < \pi\,/\,2, \quad |\,\beta\,[t]\,| \leqslant \beta_0 < \pi\,/\,2 \tag{1.2}$$

Предполагается, что игрокам известны реализующиеся скорость и геометрические координаты точки. Первая задача, рассматриваемая в данной статье, — задача о минимаксе — состоит в построении такой стратегии первого игрока, которая при любых реализациях помех $\alpha\,[t]$ и $\beta\,[t]$ и управления $v\,[t]$ гарантирует приведение точки в заданное положение за наименьшее возможное время. Вторая задача — задача о максимине, стоящая перед вторым игроком, — заключается в определении стратегии уклонения, формирующей управление $v[t]$ так, чтобы точка не попадала в заданное положение на максимальном промежутке времени; при этом полагаем, что второму игроку неизвестны реализующиеся помехи $\alpha[t]$, $\beta[t]$ и управление $u[t]$.

Уточним постановки этих задач. Пусть $y = \{y_1, y_2\}$ — вектор, составленный из геометрических координат точки, $z = \{z_1, z_2\}$ — скорость точки, $H(\gamma)$ — матрица преобразования поворота на угол γ, т. е.

$$H(\gamma) = \begin{Vmatrix} \cos\gamma & -\sin\gamma \\ \sin\gamma & \cos\gamma \end{Vmatrix}$$

Тогда движение точки можно описать уравнениями

$$\dot{y} = z, \quad \dot{z} = H(\alpha)\, u - H(\beta) v \tag{1.3}$$

Задана начальная позиция игры $\{t_0, x_0\}$; здесь и в дальнейшем $x = \{y, z\}$ — четырехмерный фазовый вектор системы. В качестве платы игры выбрано время до попадания точки $y[t]$ в начало координат $M = \{0, 0\}$. Первый игрок стремится минимизировать плату, второй — максимизирует значение платы игры.

Стратегиями игроков назовем многозначные функции $U = U(t, x)$, $V = V(t, x)$, полунепрерывные сверху относительно включения. Эти функции позиции $\{t, x\}$ ставят в соответствие непустые множества $U(t, x)$, $V(t, x)$, причем элементы u и v этих множеств удовлетворяют условиям (1.1). Движения конфликтно управляемой системы (1.3) определяются так же, как это сделано в работе [6]. Например, движением, порожденным стратегией $U = U(t, x)$, называется всякая абсолютно непрерывная вектор-функция $x[t] = x[t; t_0, x_0, U]$, $(x[t_0] = x_0)$, которая при почти всех $t \geqslant t_0$ удовлетворяет условиям

$$y^{\cdot}[t] = z[t], \quad z^{\cdot}[t] \in \mathrm{co}\,\{H(\alpha)\, u - H(\beta)\, v : \\ : u \in U(t, x[t]),\ |\alpha| \leqslant \alpha_0,\ |\beta| \leqslant \beta_0,\ \|v\| \leqslant \nu\}$$

В рассматриваемых задачах оптимальные стратегии игроков содержатся в классе описанных выше регулярно-разрывных стратегий, поэтому здесь нет необходимости вводить более полный класс позиционных разрывных стратегий [1,2]. Отметим также, что информационная дискриминация, которая иногда вводится для преодоления трудностей решения, существенно исказила бы здесь реальное содержание исходных игровых задач. Действительно, предположение о том, что первому игроку известна реализующаяся помеха $\alpha[t]$, фактически исключало бы воздействие этой помехи на систему управления и привело бы к неверному решению задачи о минимаксе. Аналогичное обстоятельство имело бы место при введении информационной дискриминации в задаче об максимине.

2. Рассмотрим решение задачи о минимаксе. Эта задача состоит в построении стратегии $U_0 = U_0(t, x)$, которая обладает следующим свойством: для любого движения $x[t] = x[t; t_0, x_0, U_0]$ условие $y[t_*] = 0$ осуществится при $t_* \leqslant t_0 + T_0$, где T_0 — некоторое число (минимакс платы), причем не существует способа формирования управления u, использую-

щего лишь информацию о позиции $\{t, x[t]\}$, который гарантирует попадание точки в начало координат $y = 0$ за время меньшее, чем T_0.

В этой задаче первому игроку противодействуют неизвестные ему помехи $\alpha[t]$, $\beta[t]$ и управление второго игрока $v[t]$. При постановке задачи о минимаксе не оговаривается способ формирования этих противодействующих факторов. Тем самым не исключаются такие способы управления, которые используют информацию о реализующихся управлениях первого игрока $u[t]$. Итак, постановка минимаксной задачи, стоящей перед первым игроком, допускает информационную дискриминацию первого игрока. (Лучше переоценить возможности противника, чем свои собственные.)

Перейдем к построению стратегии $U_0 = U_0(t, x)$, воспользовавшись подходом, предложенным в работах [5, 6], введем в рассмотрение минимаксное гипотетическое рассогласование $\varepsilon_0(t, x, \sigma)$, вычисленное для момента σ и исходной позиции $\{t, x\}$. Согласно соотношению (3.9) работы [6], в рассматриваемом случае

$$\varepsilon_0(t, x, \sigma) = \max_l \Big[\int_t^\sigma \min_u \max_{v, \alpha, \beta} l' \{X(\sigma, \tau) f(u, v, \alpha, \beta)\}_m d\tau +$$

$$+ l' \{X(\sigma, t) x\}_m\Big] = \|s_*(t, x, \sigma)\| - {}^1/_2 (\sigma - t)^2 (\mu \cos \alpha_0 - \nu)$$

$$s_* = \{y_1 + (\sigma - t) z_1,\ y_2 + (\sigma - t) z_2\} \tag{2.1}$$

Здесь и дальше l — двумерный единичный вектор, минимум и максимум под интегралом вычисляются соответственно по множеству параметров $\{u: \|u\| \leqslant \mu\}$ и по множеству переменных $\{v, \alpha, \beta: \|v\| \leqslant \nu, |\alpha| \leqslant \alpha_0, |\beta| \leqslant \beta_0\}$; $X(t, \tau)$ — фундаментальная 4×4 матрица системы $y^{\boldsymbol\cdot} = z$, $z^{\boldsymbol\cdot} = 0$; $f(u, v, \alpha, \beta)$ — четырехмерный вектор, первые две компоненты которого — нули, вторые две составляют вектор $H(\alpha)u - H(\beta)v$. Индекс m в формуле (2.1) равен двум и означает, что у вектора, стоящего в фигурных скобках, следует рассматривать первые две компоненты.

Соотношение (2.1) справедливо в области $\{t, x, \sigma\}$, где его правая часть больше нуля, в противном случае полагаем $\varepsilon_0(t, x, \sigma) = 0$.

В области $\varepsilon_0(t, x, \sigma) > 0$ имеет место регулярный случай, т. е. максимум по l в правой части выражения (2.1) достигается на единственном векторе $l_0(t, x, \sigma) = s_*/\|s_*\|$ и $\varepsilon_0(t, x, \sigma)$ — непрерывно дифференцируемая функция переменных t, x при фиксированном значении σ.

Назовем моментом программного поглощения по минимаксу наименьшее значение параметра $\sigma = \vartheta_0(t, x) \geqslant t$, при котором функция $\varepsilon_0(t, x, \sigma)$ обращается в нуль.

Введенные выше величины $\varepsilon_0(t, x, \sigma)$, $l_0(t, x, \sigma)$, $\vartheta_0(t_0, x_0)$ являются основными элементами экстремальной конструкции, определяющей в каждой позиции $\{t, x\}$ значения функции $U_e = U_e(t, x)$ — экстремальной стратегии, которая назначается следующим правилом:

$$U_e(t, x) = \begin{cases} \mu l_0(t, x, \vartheta_0(t_0, x_0)), & \varepsilon_0(t, x, \vartheta_0(t_0, x_0)) > 0 \\ \{u: \|u\| \leqslant \mu\}, & \varepsilon_0(t, x, \vartheta_0(t_0, x_0)) = 0 \end{cases} \tag{2.2}$$

Можно проверить (см., например, аналогичный случай в [5, 6]), что вдоль движений $x[t] = x[t; t_0, U_e]$, порожденных стратегией $U_e(t, x)$, величина $\varepsilon_0[t] = \varepsilon_0(t, x[t], \vartheta_0(t_0, x_0))$ не возрастает; далее, по определению момента $\vartheta_0(t_0, x_0)$, имеем равенство $\varepsilon_0[t_0] = 0$, следовательно $\varepsilon_0[\vartheta_0(t_0, x_0)] = 0$. Из формулы (2.1) теперь получаем равенство $\varepsilon_0[\vartheta_0] = \|y[\vartheta_0]\| = 0$.

Итак, экстремальная стратегия $U_e = U_e(t, x)$ гарантирует первому игроку приведение точки $y[t]$ в положение $M = \{0, 0\}$ к моменту времени ϑ_0. Покажем теперь, что этот результат является наилучшим для первого игрока, т. е. стратегия U_e есть оптимальная минимаксная стратегия.

Для проверки этого положения воспользуемся следующим обстоятельством Пусть T_* — время оптимального быстродействия в задаче о переводе системы

$$y^{\cdot} = z, \quad z^{\cdot} = w \quad (\|w\| \leqslant \mu \cos \alpha_0 - \nu) \tag{2.3}$$

из начального состояния $\{t_0, x_0\}$ в положение $y = 0$. Оказывается, что справедливо равенство

$$T_* = \vartheta_0(t_0, x_0) - t_0 \tag{2.4}$$

Этот факт вытекает из выражения (2.1) и соответствующих условий разрешимости задачи оптимального быстродействия [7,8].

Пусть теперь первый игрок выбирает любой позиционный способ формирования управления $u[t]$; тогда, полагая $v[t] = u[t]_* \nu/\mu$ $\beta[t] = 0$, а помеху $\alpha[t]$ — случайной и равной $+\alpha_0$, $-\alpha_0$ с вероятностью $1/2$, получаем, что среднее значение нормы вектора $H(\alpha[t])\ u[t] - H(\beta[t])\ v[t]$ не будет превосходить величины $\mu \cos \alpha_0 - \nu$, следовательно [7,8], встреча не может произойти раньше, чем за время оптимального быстродействия T_*, т. е. для минимакса платы T_0 справедлива оценка $T_0 \leqslant T_*$. С другой стороны, выше было установлено, что стратегия U_e (2.2) обеспечивает приведение за время $T_0 = \vartheta_0(t_0, x_0) - t_0 = T_*$. Итак, $U_e = U_e(t, x)$ действительно чистая минимаксная стратегия первого игрока.

Класс обобщенных стратегий $\{U\}$, который содержит решение минимаксной задачи наведения $U_0 = U_e$ (2.2) был введен для формального описания разрывных способов управления. Поясним теперь содержательный смысл полученного результата в соответствии с общей содержательной трактовкой, данной в работах [3,5]. Рассмотрим следующую аппроксимационную схему формирования кусочно-постоянных управлений первого игрока $u_\Delta[t]$ $(t \geqslant t_0)$:

$$u_\Delta[t] = u_\Delta[\tau_i] \in U_e(\tau_i, x^{(\Delta)}[\tau_i]), \qquad t \in [\tau_i, \tau_{i+1})$$
$$\tau_{i+1} = \tau_{i+\Delta}, \qquad i = 0, 1, \ldots, \qquad \tau_0 = t_0, \qquad \Delta > 0$$

Здесь $x^{(\Delta)}[t]$ — движение системы (1.3), порожденное управлением $u_\Delta[t]$ и некоторыми реализациями $\alpha[t]$, $\beta[t]$, $v[t]$. Справедливо следующее утверждение: для любой сколь угодно малой окрестности S_ε начала координат $M = \{0, 0\}$ можно указать такое $\Delta_0 > 0$, что при всех $\Delta < \Delta_0$ и для любых реализаций $v[t]$, $\alpha[t]$, $\beta[t]$ аппроксимационное управление $u_\Delta[t]$ обеспечит приведение точки $\{y_1^{(\Delta)}[t], y_2^{(\Delta)}[t]\}$ в окрестность S_ε за время T_0. С другой стороны, при любом $T < T_0$ существует окрестность начала координат $S_\varepsilon(T)$ такая, что с вероятностью, сколь угодно близкой к единице, точка $\{y_1^{(\Delta)}[t], y_2^{(\Delta)}[t]\}$ будет уклоняться на отрезке $[t_0, t_0 + T]$ от попадания в $S_\varepsilon(T)$ при достаточно малых Δ. Движение точки $\{y_1^{(\Delta)}[t], y_2^{(\Delta)}[t]\}$ порождается здесь управлениями $\beta[t] \equiv 0$, $v[t] = \nu/\mu\ u[t]$ (где $u[t]$ — реализация произвольного позиционного способа управления первого игрока), $\alpha_\Delta[t] = \alpha_\Delta[\tau_i]$ $(t \in [\tau_i, \tau_{i+1}))$ — кусочно-постоянная функция, значения которой случайным образом с вероятностью $1/2$ полагаются равными $+\alpha_0$ и $-\alpha_0$. Предполагается, что выбор помехи $\alpha_\Delta[t]$ вероятностно независим с выбором управления $u[t]$.

И наконец, отметим следующий факт. Оптимальный выигрыш в задаче наведения для системы (1.3) при отсутствии помех α и β дает величину, меньшую, чем T_0; эта величина будет в точности равна T_0, если первое ограничение в (1.1) заменить условием $\|u\| \leqslant \mu \cos\alpha_0$. Другими словами, появление помехи α дает тот же эффект, что и уменьшение величины той максимальной силы, с которой первый игрок может действовать на рассматриваемую материальную точку.

3. Рассмотрим теперь задачу о максимине. Оптимальным результатом в задаче о максимине является величина T^0, обладающая следующим свойством. Какую бы стратегию V ни избрал второй игрок, он не может гарантировать выполнение условия $y[t] \neq 0$ при $t_0 \leqslant t \leqslant T^*$ ($T^* > T^0$) для всех движений $x[t] = x[t;\ t_0, x_0, V]$. С другой стороны, существуют стратегии V_δ, для которых все движения $x[t] = x[t;\ t_0, x_0, V_\delta]$ удовлетворяют условию $y[t] \neq 0$ при $t_0 \leqslant t \leqslant T^0 - \delta$, здесь $\delta > 0$ — сколь угодно мало. Таким образом, T^0 — максимин платы игры (точнее sup inf платы) — в классе чистых стратегий. В предлагаемой работе описывается построение стратегии $V_\delta = V_\delta(t, x)$, доставляющей второму игроку результат, сколь угодно близкий к оптимальному.

Подчеркнем еще раз, что стратегии $V_\delta = V_\delta\ (t, x)$ будут формировать управление второго игрока $v[t]$, не используя информацию о реализациях помех $\alpha[t]$, $\beta[t]$. Выше было отмечено, что наличие у второго игрока такой информации существенно искажает смысл исходной игровой задачи. Более того, постановка максиминной задачи уклонения допускает информационную дискриминацию того игрока, в чьих интересах решается эта задача (т. е. второго игрока).

Для построения стратегии $V_\delta = V_\delta(t, x)$ воспользуемся подходом, предложенным в работе [9]. Введем понятие максиминного гипотетического рассогласования $\varepsilon^0(t, x, \sigma)$ [5,6]. Пусть система (1.3) находится в момент t в состоянии $x[t] = x$. Зададимся некоторой измеримой функцией $v[\tau]$ ($t \leqslant \tau \leqslant \sigma$, $\|v[\tau]\| \leqslant \nu$) и рассмотрим область достижимости $G(t, x, \sigma; v[\cdot])$ — множество тех точек y, в которые в момент σ может попасть управляемая точка из начальной позиции $\{t, x\}$ при выборе вторым игроком указанного управления $v[\tau]$, $t \leqslant \tau \leqslant \sigma$, и при всевозможных суммируемых реализациях $u[\tau]$, $\alpha[\tau]$, $\beta[\tau]$ ($t \leqslant \tau \leqslant \sigma$, $\|u[\tau]\| \leqslant \mu$, $|\alpha[t]| \leqslant \alpha_0$, $|\beta[t]| \leqslant \beta_0$). Это множество выпукло, замкнуто, ограничено. Пусть $\varepsilon(t, x, \sigma; v[\cdot])$ — расстояние от точки $M = \{0, 0\}$ до множества $G(t,x, \sigma; v[\cdot])$. Определим теперь максиминное гипотетическое рассогласование $\varepsilon^0(t, x, \sigma)$ как максимальное значение величины $\varepsilon(t, x, \sigma; v[\cdot])$, рассматриваемой при фиксированных значениях t, x, σ на множестве всех программных управлений второго игрока $v[\tau]$ ($t \leqslant \tau \leqslant \sigma$; $\|v[\tau]\| \leqslant \nu$). В данном примере это определение приводит к следующему соотношению для вычисления $\varepsilon^0(t, x, \sigma)$:

$$\varepsilon^0(t, x, \sigma) = \max_{v[\cdot]} \varepsilon(t, x, \sigma; v[\cdot]) = \max_l \Big[\int_t^\sigma \max_v \min_{u,\alpha,\beta} l' \times \tag{3.1}$$

$$\times \{X(\sigma, \tau) f(u, v, \alpha, \beta)\}_m d\tau + l' \{X(\sigma, t) x\}_m\Big] =$$

$$= \|s_*(t, x, \sigma)\| - {}^1/_2 (\sigma - t)^2 (\mu - \nu \cos\beta_0)$$

Все величины, входящие в формулу (3.1), имеют тот же смысл, что и в формуле (2.1), $m = 2$.

Максиминное гипотетическое рассогласование $\varepsilon^0(t, x, \sigma)$ полностью определено соотношением (3.1) в той области пространства $\{t, x, \sigma\}$, где правая часть (3.1) больше нуля; в остальной части пространства полагаем $\varepsilon^0(t, x, \sigma) = 0$. В области $\varepsilon^0(t, x, \sigma) > 0$ функция $\varepsilon^0(t, x, \sigma)$ непрерывно дифференцируема по t, x при фиксированном σ и максимум по $l(\|l\| = 1)$ в правой части (3.1) достигается на единственном векторе $l_0(t, x, \sigma) = s_*/\|s_*\|$, т. е. имеет место так называемый регулярный случай.

Назовем наименьший корень уравнения $\varepsilon^0(t, x, \sigma) = 0$, рассматриваемого при фиксированной исходной позиции $\{t, x\}$, моментом максиминного программного поглощения $\vartheta^0(t, x)$. Оказывается, что максимин платы T^0 в рассматриваемой игре равен $\vartheta^0 - t_0$, где $\vartheta^0 = \vartheta^0(t_0, x_0)$.

Для доказательства этого факта покажем сначала, что ни одна стратегия второго игрока $V = V(t, x)$ не гарантирует уклонения системы (1.3) от попадания в точку $M = \{0, 0\}$ на отрезке $[t_0, \vartheta^0]$, т. е. $T^0 \leqslant \vartheta^0 - t_0$.

Введем в рассмотрение вспомогательную задачу быстродействия для системы (2.3) из начальной позиции $\{t_0, x_0\}$ в состояние $y = 0$, но уже при следующих ограничениях на ресурсы управления $\|w\| \leqslant \mu - \nu \cos \beta_0$. Из равенства (3.1) и условий разрешимости задачи быстродействия вытекает, что время оптимального быстродействия T^* в этом случае равно $\vartheta^0 - t_0$. Пусть теперь на отрезке $[t_0, \vartheta^0]$ рассматривается исходная игра уклонения, и второй игрок выбрал какой-то позиционный способ формирования управления $v[t]$. Рассмотрим следующий способ формирования управления u, помех α и β: $\alpha[t] \equiv 0$, помеха $\beta[t]$ выбирается случайным образом с вероятностью $1/2$ равной $+\beta_0$ или $-\beta_0$, так что в каждый момент времени $t \geqslant t_0$ в системе (1.3) реализуется некоторое усредненное значение управления второго игрока $v_*[t]$, по величине не превосходящее $\nu \cos \beta_0$. Пусть кроме этой силы $v_*[t]$ на рассматриваемую материальную точку действует сила $u[t] = v_*[t] + w^0[t]$, где $w^0[t]$, — решение вспомогательной задачи об оптимальном быстродействии. Отметим, что постановка максиминной задачи не исключает возможности реализации такого управления партнера.

Итак, сила, действующая на точку $H(\alpha)u - H(\beta)v$, оказывается в точности равной вектору $w^0[t]$, следовательно, движение системы (1.3), порожденное этими реализациями, попадает в точку $M = \{0, 0\}$ в момент $t = t_0 + T^* = \vartheta^0$. Поэтому ни одна стратегия V второго игрока не гарантирует уклонения точки $\{y_1[t], y_2[t]\}$ от попадания в состояние $M = \{0,0\}$ до момента $\vartheta^0 = t_0 + T^*$, т. е. действительно, $T^0 \leqslant T^* = \vartheta^0 - t_0$.

С другой стороны, ниже будет построена стратегия V_δ, которая обеспечивает для всех движений $x[t] = x[t; t_0, x_0, V_\delta]$ выполнение условия $y[t] \neq 0$ при $t_0 \leqslant t \leqslant \vartheta^0 - \delta$ ($\delta > 0$ — сколь угодно мало). Следовательно, максимин платы T^0, действительно, совпадает с величиной $\vartheta^0 - t_0$.

Приступим к построению стратегии $V_\delta = V_\delta(t, x)$. Как и в случае линейных систем с разделяющимися управлениями [9], вводится вспомогательная функция

$$L_\delta(t, x) = \int_t^{\vartheta_\delta} [\varepsilon^0(t, x, \sigma)]^{-1} d\sigma, \quad \vartheta_\delta = \vartheta^0 - \delta \tag{3.2}$$

Функция $L_\delta(t, x)$ определена в той части пространства $\{t, x\}$, где $\varepsilon^0(t, x, \sigma) > 0$ при $t \leqslant \sigma \leqslant \vartheta_\delta$. Обратно, из соотношения $L_\delta(t, x) < \infty$ можно вывести, что функция $\varepsilon^0(t, x, \sigma)$ ни при каком $\sigma \in [t_0, \vartheta_\delta]$ не обращается в нуль (в противном случае несобственный интеграл в правой части (3.2) расходится). В частности, из свойства $L_\delta(t, x) < \infty$ следует $\varepsilon^0(t, x, t) =$

$= \|y[t]\| > 0$. Таким образом, искомую стратегию второго игрока V_δ достаточно построить так, чтобы вдоль любого движения $x[t] = x[t; t_0 x_0, V_\delta]$, функция $L_\delta(t, x)$ оставалась ограниченной при всех $t \in [t_0, \vartheta_\delta]$. Заметив, что $L_\delta(t_0, x_0) < \infty$ (т. к. $\varepsilon^0(t, x, \sigma) > 0$ при $t_0 \leqslant \sigma \leqslant \vartheta_\delta = \vartheta^0 - \delta$), будем строить стратегию $V_\delta = V_\delta(t, x)$ так, чтобы обеспечить выполнение соотношения

$$L_\delta(t, x[t]) \leqslant L_\delta(t_0, x_0) \tag{3.3}$$

Вычислим полную производную функции $L_\delta(t, x)$ по переменной t в силу уравнений движения (1.3)

$$\frac{dL}{dt} = \Phi(t, x, u, v, \alpha, \beta) = -\|y\|^{-1} - \int_t^{\vartheta_\delta} [\varepsilon^0(t, x, \sigma)]^{-2}(\sigma - t)(\mu - \nu\cos\beta_0)\,d\sigma -$$

$$-(H(\alpha)u - H(\beta)v)' \int_t^{\vartheta_\delta} [\varepsilon^0(t, x, \sigma)]^{-2}(\sigma - t)\,l_0(t, x, \sigma)\,d\sigma$$

Вычислим теперь величину $\Phi^0(t, x)$

$$\Phi^0(t, x) = \min_v \max_{u,\alpha,\beta} \Phi(t, x, u, v, \alpha, \beta)$$

Получим следующее соотношение:

$$\Phi^0(t, x) = \max_{u,\alpha,\beta} \Phi(t, x, u, v^0(t, x), \alpha, \beta) = \tag{3.4}$$

$$= -\|y\|^{-1} - (\mu - \nu\cos\beta_0)\left(\|p(t, x)\| - \int_t^{\vartheta_\delta} (\sigma - t)[\varepsilon^0(t, x, \sigma)]^{-2}\,d\sigma\right) < 0,$$

$$p(t, x) = \int_t^{\vartheta_\delta} [\varepsilon^0(t, x, \sigma)]^{-2}(\sigma - t)\,l_0(t, x, \sigma)\,d\sigma$$

$$v^0(t, x) = -\nu p(t, x) / \|p(t, x)\|$$

Стратегию V_δ определим теперь следующим образом:

$$V_\delta = v^0(t, x) \tag{3.5}$$

Так же, как и в работе [9], используя соотношение (3.4), можно показать, что стратегия V_δ действительно обеспечивает выполнение неравенства (3.3) и, следовательно, гарантирует выполнение соотношения $y[t] \neq 0$ при $t_0 \leqslant t \leqslant \vartheta_\delta = \vartheta^0 - \delta$. Отметим, что построенная стратегия V_δ (3.5) в области ее определения будет непрерывной вектор-функцией позиции игры.

Содержательный смысл вероятностного механизма, который привлекался при обосновании оптимальности результата T^0, можно пояснить, следуя опять общей трактовке из работы [3]. Пусть помеха β формируется в дискретные моменты времени $t = \tau_i$ ($\tau_i = t_0 + i\Delta$, $\Delta > 0$) и принимает значения $+\beta_0$ или $-\beta_0$ случайным образом с вероятностями, равными $1/2$. Затем полагаем $\beta[t] = \beta[\tau_i]$ при $t \in [\tau_i, \tau_{i+1})$; $\alpha[t] \equiv 0$, $u[t] = w^0[t] + v[t]$, причем выбор значений помехи $\beta[t]$, реализующихся в данный момент в системе (1.3), вероятностно независим с выбором управления $v[t]$. С вероятностью, сколь угодно близкой к единице, движения, порожденные этими реализациями, попадают в сколь угодно малую окрестность точки $M = \{0, 0\}$ не позже момента ϑ^0 при выборе достаточно мелкого разбиения $\Delta > 0$.

В заключение отметим, что минимакс платы T_0 совпадает со значением дифференциальной игры, описываемой уравнением

$$y^{\cdot} = z, \quad z^{\cdot} = u - v \tag{3.6}$$

при следующих ограничениях на управления: $\|u\| \leqslant \mu \cos \alpha_0$, $\|v\| \leqslant \nu$. Максимин платы T^0 есть значение дифференциальной игры, описываемой системой (3.6), при ограничениях на управления $\|u\| \leqslant \mu$, $\|v\| \leqslant \nu \cos \beta_0$. Таким образом, наличие помехи приводит к тому же результату, что и уменьшение ресурсов игрока, с точки зрения которого исследуется игровая задача.

При доказательстве оптимальности построенных выше чистых стратегий U_e, V_δ рассматривались смешанные стратегии противника, а именно, в минимаксной задаче вводилось смешанное управление помехой α, в максиминной задаче — смешанное управление помехой β. Эти понятия были привлечены для вспомогательных рассуждений. Если же поставить задачу об отыскании оптимальных смешанных стратегий в дифференциальной игре, отвечающей уравнениям (1.3), где параметры u, β подчинены первому игроку, а v, α — второму игроку, то можно показать, что в классе смешанных стратегий эта игра преследования — уклонения имеет значение T_c. Величина T_c совпадает со значением дифференциальной игры, описываемой уравнением (3.6) при ограничениях $\|u\| \leqslant \mu \cos \alpha_0$, $\|v\| \leqslant \nu \| \nu \cos \beta_0$.

Приведенные в данной статье постановки задач и построения оптимальных стратегий можно перенести на случай, когда вместо системы (1.3) рассматривается система из контрольного примера [10], в котором, однако, на управляющие воздействия игроков также накладываются помехи типа «люфта».

Автор благодарит Н. Н. Красовского за постановку задачи и ценные советы.

Поступила 31 I 1972

ЛИТЕРАТУРА

1. Красовский Н. Н., Субботин А. И. О структуре дифференциальных игр. Докл. АН СССР, 1970, т. 190, № 3.
2. Красовский Н. Н., Субботин А. И. Экстремальные стратегии в дифференциальных играх. Докл. АН СССР, 1971, т. 196, № 2.
3. Красовский Н. Н., Субботин А. И. О седловой точке позиционной дифференциальной игры. Тр. матем. ин-та, 1972, т. 128.
4. Филиппов А. Ф. Дифференциальные уравнения с разрывной правой частью. Матем. сб., 1970, т. 51(93) вып. 1.
5. Красовский Н. Н. Игровые задачи о встрече движений. М., «Наука», 1970.
6. Красовский Н. Н. Минимаксное поглощение в игре сближения. ПММ, 1971, т. 35, вып. 6.
7. Красовский Н. Н. Теория управления движением. М., «Наука», 1968.
8. Понтрягин Л. С., Болтянский В. Г., Гамкрелидзе Р. В., Мищенко Е. Ф. Математическая теория оптимальных процессов. М., Физматгиз, 1961.
9. Красовский Н. Н., Субботин А. И. Об оптимальных стратегиях в линейной дифференциальной игре. ПММ, 1969, т. 33, вып. 4.
10. Понтрягин Л. С. К теории дифференциальных игр. Усп. матем. н., 1966, т. 21, № 4.