Прикладная математика и механика. Том 39, 1975

УДК 62—50

МЕТОД МАЛОГО ПАРАМЕТРА ДЛЯ ПОСТРОЕНИЯ ПРИБЛИЖЕННЫХ СТРАТЕГИЙ ОДНОГО КЛАССА ДИФФЕРЕНЦИАЛЬНЫХ ИГР

А. С. Братусь

(Москва)

Рассматривается класс задач, в которых плата есть некоторая функция конечного состояния конфликтно-управляемой системы. В случае, когда возможности одного из игроков малы по отношению к возможностям другого, предложены методы построения приближенных оптимальных стратегий игроков, основанные на решении уравнения Беллмана, содержащего малый параметр. Показано, что приближенные оптимальные стратегии игроков могут быть построены, если известны решения соответствующих задач оптимального управления. Доказаны оценки погрешности методов. Рассмотрены примеры. Приведенные рассуждения опираются на результаты работ [1–6] по теории дифференциальных игр и на работы [7–11], посвященные методам синтеза оптимального управления для систем, подверженных случайным возмущениям малой интенсивности.

1. Постановка задачи. Пусть движение конфликтно-управляемой системы описывается нелинейным уравнением

$$\frac{dx}{dt} = F(x, t, u, v), \quad u \in P, \quad v \in Q_\varepsilon, \quad x[t_0] = x_0, \quad t \in [t_0, T] \tag{1.1}$$

Здесь x — n-мерный вектор, u и v — r-мерные векторы управления первого и второго игроков соответственно, P и Q_ε — ограниченные замкнутые множества, F — непрерывная функция, удовлетворяющая условию Липшица по x и v. Плата представляет собой величину $f[x(T)]$, определяемую в конечный момент времени $t = T$ по реализовавшейся позиции $x(T)$.

Первый игрок старается минимизировать величину $f[x(T)]$ при самом неблагоприятном поведении второго игрока. Задача второго игрока — гарантировать завершение игры с наибольшим возможным значением платы. Предполагается, что возможности одного из игроков малы по сравнению с возможностями другого. А именно: будем считать, что множество Q_ε может быть заключено в r-мерную сферу радиуса ε, малого по отношению к минимальному радиусу сферы, в которую может быть заключено множество P. Далее полагаем, что правая часть системы (1.1) удовлетворяет условию из работы [1]

$$|x'F(x, t, u, v)| \leqslant \lambda(1 + \|x\|^2), \ u \in P, \ v \in Q_\varepsilon, \ \lambda = \text{const} \tag{1.2}$$

Стратегией первого игрока [2,3] назовем функцию $u(x, t)$, которая всякой возможной позиции (x, t) ставит в соответствие замкнутое множество $u(x, t) \subset P$. Предполагается, что множество $u(x, t)$ полунепрерывно сверху относительно включения по x, t. Аналогично определяется класс функций $v(x, t)$, которые задают допустимые стратегии второго игрока.

2. Метод итераций. Рассмотрим класс дифференциальных игр, описанных в п. 1 и имеющих седловую точку. Справедлив следующий результат ([4], теорема 3).

Лемма. Пусть существует непререрывно дифференцируемая функция $S(x, t)$, которая при всех x, t удовлетворяет краевой задаче

$$S_t + \min_u \max_v \{(F(x, t, u, v),\ S_x)\} = S_t + \max_v \min_u \{(F(x, t, u, v),\ S_x)\} = 0 \tag{2.1}$$
$$S(x, T) = f(x), \quad u \in P, \quad v \in Q_\varepsilon .$$

Здесь $S_x = (S_{x_1}, \ldots, S_{x_n})$ — вектор первых частных производных по x, S_t — частная производная по времени t. Тогда множество $u^*(x, t)$ векторов u^* и множество $v^*(x, t)$ векторов v^*, доставляющих в (2.1) минимакс и максимин, таковы, что стратегии $u^*(x, t)$ и $v^*(x, t)$ будут минимаксными и максиминными стратегиями первого и второго игроков соответственно, причем пара стратегий $(u^*(x, t),\ v^*(x, t))$ доставляет седловую точку рассматриваемой игры. Введем обозначение

$$\min_u \max_v \{(F(x, t, u, v), S_x)\} = \max_v \min_u \{(F(x, t, u, v), S_x)\} = H(x, t, u^*, v^*, S_x) \tag{2.2}$$
$$u \in P, \quad v \in Q_\varepsilon$$

Построим следующий итерационный процесс. Выберем стратегию $v^\circ(x, t) \equiv 0$ и рассмотрим краевую задачу

$$S_t^\circ + \min_u \{(F(x, t, u, v^\circ),\ S_x^\circ)\} = 0,\ S^\circ(x, T) = f(x),\ u \in P \tag{2.3}$$

Предполагая, что решение этой задачи существует и непрерывно дифференцируемо, найдем стратегию $u^\circ(x, t)$ из условия

$$\min_u \{(F(x, t, u, v^\circ),\ S_x^\circ)\} = H(x, t, u^\circ, v^\circ, S_x^\circ),\ u \in P \tag{2.4}$$

Новую стратегию $v^1(x, t)$ второго игрока найдем из условия

$$\max_v \{(F(x, t, u^\circ, v),\ S_x^\circ)\} = H(x, t, u^\circ, v^1,\ S_x^\circ), \quad v \in Q_\varepsilon \tag{2.5}$$

Назовем стратегию $u^\circ(x, t)$, определенную в (2.3) и (2.4), приближенной стратегией первого игрока в нулевом приближении. Стратегию $v^1(x, t)$ второго игрока назовем приближенной стратегией второго игрока в первом приближении. Следующий шаг итерационного процесса заключается в решении краевой задачи

$$S_t^1 + \min_u \{(F(x, t, u, v^1),\ S_x^1)\} = 0, S^1(x,\ T) = f(x),\ u \in P \tag{2.6}$$

Как и ранее, предполагая существование и непрерывную дифференцируемость функции S^1, найдем стратегию первого игрока — $u^1(x, t)$ из условия

$$\min_u \{(F(x, t, u, v^1),\ S_x^1)\} = H(x, t, u^1, v^1,\ S_x^1), \quad u \in P \tag{2.7}$$

Стратегию $u^1(x, t)$ назовем приближенной стратегией первого игрока в первом приближении. Новую стратегию $v^2(x, t)$ второго игрока найдем из условия

$$\max_v \{(F(x, t, u^1, v),\ S_x^1)\} = H(x, t, u^1, v^2, S_x^1), \quad v \in Q_\varepsilon \tag{2.8}$$

Используя соотношения (2.5) и (2.7) и рассуждая так же, как в [5] (теорема 1), можно показать, что найденные таким образом стратегии $u^1(x, t)$ и $v^1(x, t)$ являются допустимыми в смысле определения п. 1.

Запишем краевую задачу (2.6), используя (2.7), в виде

$$S_t^1 + H(x, t, u^1, v^1, S_x^1) = 0, \quad S^1(x, T) = f(x) \tag{2.9}$$

Оценим отличие функции Беллмана S от функции S^1. Для этого рассмотрим функцию

$$\alpha_1^\varepsilon(x, t) = H(x, t, u^1, v^1, S_x^1) - H(x, t, u^1, v^2, S_x^1) \leqslant 0$$

Справедливо неравенство

$$|\alpha_1^\varepsilon| = |((F(x, t, u^1, v^1) - F(x, t, u^1, v^2)), S_x^1)| \leqslant \\ \leqslant \|F(x, t, u^1, v^1) - F(x, t, u^1, v^2)\| \|S_x^1\| \leqslant C' \|v^1 - \\ - v^2\| \|S_x^1\| \leqslant C'\varepsilon(1 + \|x\|^2)^m$$

если предположить, что

$$\|S_x^1\| \leqslant C^2(1 + \|x\|^2)^m, \quad c, c^1, c^2 = \mathrm{const} \tag{2.10}$$

Поэтому

$$0 \geqslant \alpha_1^\varepsilon \geqslant -C\varepsilon(1 + \|x\|^2)^m$$

Справедлив следующий результат.

Теорема 2.1. Пусть существуют непрерывно дифференцируемые решения задач (2.1) и (2.9) и выполняется оценка (2.10). Тогда справедливо неравенство

$$0 \leqslant S(x, t) - S^1(x, t) \leqslant (\lambda m)^{-1}\varepsilon C\,[e^{\lambda m(T-t_0)} - 1](1 + \\ + \|x\|^2)^m, \quad \lambda = \mathrm{const} \tag{2.11}$$

Доказательство. Запишем краевую задачу (2.1), учитывая обозначение (2.2), в виде

$$S_t + H(x, t, u^*, v^*, S_x) = 0, \quad S(x, T) = f(x) \tag{2.12}$$

Из результата леммы следует, что пара стратегий $(u^*(x, t), v^*(x, t))$ доставляет седловую точку рассматриваемой игры. Поэтому справедливо неравенство

$$H(x, t, u^*, v^1, S_x) \leqslant H(x, t, u^*, v^*, S_x) \leqslant H(x, t, u^1, v^*, S_x) \tag{2.13}$$

С другой стороны, из определения стратегии $v^2(x, t)$ в (2.8), учитывая введенные обозначения для α_1^ε, получим

$$H(x, t, u^1, v^2, S_x^1) = H(x, t, u^1, v^1, S_x^1) - \alpha_1^\varepsilon(x, t) \geqslant H(x, \\ t, u^1, v^*, S_x^1) \tag{2.14}$$

Далее, из (2.7) следует неравенство

$$H(x, t, u^1, v^1, S_x^1) \leqslant H(x, t, u^*, v^1, S_x^1) \tag{2.15}$$

Используя второе из неравенств (2.13) и уравнение (2.12), получим

$$S_t + H(x, t, u^1, v^*, S_x) \geqslant 0, \quad S(x, T) = f(x) \tag{2.16}$$

Из неравенства (2.14) и уравнения (2.9) следует, что выполняется неравенство

(2.17) $$S_t^1 + H(x, t, u^1, v^*, S_x^1) \leqslant -\alpha_1^\varepsilon(x, t), \quad S^1(x, T) = f(x)$$

Вычитая из (2.16) неравенство (2.17), получим неравенство и краевое условие для функции $Z = S - S^1$

(2.18) $$Z_t + H(x, t, u^1, v^*, Z_x) \geqslant \alpha_1^\varepsilon(x, t), \quad Z(x, T) = 0$$

Это означает (см., например, [4]), что для любого движения $x[t] = x[t, x_0, t_0, u^1, v^*]$ почти при всех $t \in [t_0, T]$ выполняется неравенство

(2.19) $$dZ(x[t], t)/dt \geqslant \alpha_1^\varepsilon(x, t) \geqslant -C\varepsilon(1 + \|x[t]\|^2), \quad Z(x, T) = 0$$

Здесь производная вычисляется вдоль движения $x[t]$. Существование этой производной почти при всех $t \geqslant t_0$ вытекает из непрерывной дифференцируемости функций S и S^1 и абсолютной непрерывности движений [4].

Используя неравенство (1.2) и исходное уравнение (1.1), получим, что справедлива оценка

(2.20) $$(1 + \|x[t]\|^2)^m \leqslant (1 + \|x_0\|^2)^m e^{m\lambda(t-t_0)}, \quad t \in [t_0, T]$$

Здесь λ — постоянная, фигурирующая в оценке (1.2). Интегрируя неравенство (2.19) от t_0 до T и учитывая (2.20) и краевое условие $Z(x, T) = 0$, получим неравенство

(2.21) $$S(x_0, t_0) - S^1(x_0, t_0) \leqslant (\lambda m)^{-1}\varepsilon C[e^{\lambda m(T-t_0)} - 1](1 + \|x_0\|^2)^m$$

Используя первое из неравенств (2.13) и уравнение (2.12), а затем неравенство (2.15) и уравнение (2.9), получим

$$S_t + H(x, t, u^*, v^1, S_x) \leqslant 0, \quad S(x, T) = f(x)$$

$$S_t^1 + H(x, t, u^*, v^1, S_x^1) \geqslant 0, \quad S^1(x, T) = f(x)$$

Вычитая из (2.18) неравенство (2.19), получим неравенство и краевое условие для функции $Z^1 = S^1 - S$

(2.22) $$dZ^1(x[t], t)/dt \geqslant 0, \quad Z^1(x, T) = 0$$

для любых движений $x[t] = x[t, x_0, u^*, v^1]$. Из (2.22) аналогично (2.21) следует неравенство $S^1 - S \leqslant 0$.

Учитывая (2.21), получим неравенство (2.11).

Следствие. Если $\alpha_1^\varepsilon(x, t) \equiv 0$, то пара стратегий $(u^1(x, t), v^1(x, t))$ доставляет седловую точку рассматриваемой игры.

Доказательство. Из неравенства (2.11) в этом случае следует, что $S = S^1$. Поэтому справедливы равенства

$$S_t + H(x, t, u^1, v^1, S_x) = 0, \quad S(x, T) = f(x)$$

Последнее в силу результата леммы означает, что пара стратегий $(u^1(x, t), v^1(x, t))$ доставляет седловую точку рассматриваемой игры.

Замечание 1. Неравенство (2.11) показывает, насколько значение функции Беллмана S в точке (x, t) отличается от минимального значения функционала платы, которое может достичь первый игрок при начальных условиях $x = x_0$, $t = t_0$ в задаче

(1.1) при условии, что второй игрок применяет стратегию $v^1(x, t)$, определенную в (2.5).

Замечание 2. Можно рассмотреть другой итерационный процесс. Для этого выберем некоторую стратегию, $u^\circ(x, t)$. (В качестве $u^\circ(x, t)$ можно, например, выбрать стратегию, полученную на первом шаге ранее рассмотренного итерационного процесса.)

Рассмотрим краевую задачу

$$W_t^\circ + \max_v \{(F(x, t, u^\circ, v), W_x^\circ)\} = 0, \quad W^\circ(x, T) = f(x)$$
$$v \in Q_\varepsilon$$

Стратегию $v^\circ(x, t)$ найдем из условия

$$\max_v \{(F(x, t, u^\circ, v), W_x^\circ)\} = H(x, t, u^\circ, v^\circ, W_x^\circ), \quad v \in Q_\varepsilon$$

Новую стратегию $u^1(x, t)$ первого игрока найдем из условия

$$\min_u \{(F(x, t, u, v^\circ), W_x^\circ)\} = H(x, t, u^1, v^\circ, W_x^\circ), \quad u \in P$$

Аналогично находим функцию W^1 и стратегию v^1 и u^2. Рассмотрим функцию

$$\beta_1^\varepsilon(x, t) = H(x, t, u^1, v^1, W_x^1) - H(x, t, u^2, v^1, W_x^1)$$

Ясно, что $\beta_1^\varepsilon(x, t) \geqslant 0$. Предположим, что справедлива оценка

$$\beta_1^\varepsilon \leqslant K(\varepsilon)(1 + \|x\|^2)^m, \quad K(\varepsilon) \geqslant 0$$

Из построения итерационного процесса ясно, что $K(\varepsilon) = 0$ при $\varepsilon = 0$. Однако в этом случае нельзя гарантировать, что $K(\varepsilon) \to 0$ при $\varepsilon \to 0$, Используя те же соображения, что и в доказательстве теоремы 1.1, можно доказать справедливость неравенства

$$-(\lambda m)^{-1} K(\varepsilon)[e^{\lambda m(T-t_0)} - 1](1 + \|x\|^2)^m \leqslant S - W^1 \leqslant 0$$

Эта оценка показывает, насколько значение функции Беллмана S в точке (x, t) отличается от максимального значения функционала платы, которое может достичь второй игрок при начальных условиях $x = x_0$, $t = t_0$ в задаче (1.1) при условии, что первый применяет стратегию $u^1(x, t)$, полученную на первом шаге описанного итерационного процесса. Как отмечалось ранее, ввиду того, что величина $K(\varepsilon)$ может не стремиться к нулю при $\varepsilon \to 0$, эффективность этой оценки мала.

Замечание 3. Условия непрерывной дифференцируемости функций S и W могут быть ослаблены, если воспользоваться результатом теоремы 2.1 из [6].

3. Метод малого параметра. Предположим, что множество Q_ε, рассмотренное в п. 1, представляет собой сферу радиуса ε в r-мерном пространстве. Здесь ε — достаточно малое число. Задача (1.1) приводится тогда к виду

$$dx/dt = F(x, t, u, \varepsilon v), \quad u \in P, \quad v \in Q_1$$

Здесь Q_1 — единичная сфера в r-мерном пространстве. Для решения этой задачи применим метод, предложенный в работах [6–9].

Рассмотрим минимаксную задачу. Предположим, что существует дважды непрерывно дифференцируемое решение краевой задачи

$$S_t + H(x, t, S_x; \varepsilon) = 0, \quad S(x, T) = f(x) \tag{3.1}$$

$$H(x, t, S_x; \varepsilon) = \min\nolimits_u \max\nolimits_v \{(F(x, t, u, \varepsilon, v), S_x)\}, \quad u \in P \quad v \in Q_1 \tag{3.2}$$

Согласно результату теоремы 1 из [4], множество векторов $u^*(x, t)$, доставляющих минимум в (3.2), определяет минимаксную стратегию первого игрока.

Пусть выполнено условие:

1) функция $H(x, t, S_x; \varepsilon)$ непрерывна вместе со своими производными по S_x и ε до второго порядка включительно. Ищем решение задачи (3.1) в виде разложения по степеням параметра ε

$$S(x, t) = S^\circ(x, t) + \varepsilon S^1(x, t) + \ldots \tag{3.3}$$

Функцию H представим в виде

$$H(x, t, S_x; \varepsilon) = H(x, t, S_x^\circ; 0) + \varepsilon H_\varepsilon(x, t, S_x^\circ; 0) + \varepsilon(\nabla H(x, t, S_x^\circ; 0), S_x^1) + \ldots \tag{3.4}$$

$$H_\varepsilon = \frac{\partial}{\partial \varepsilon} H(x, t, S_x^\circ; \varepsilon)$$

Здесь ∇H — вектор частных производных по компонентам вектора S_x.

Подставляя разложение (3.3) в (3.1) и ограничиваясь лишь членами первого порядка по ε, найдем, что функция S° — решение краевой задачи

$$S_t^\circ + H(x, t, S_x^\circ; 0) = 0, \quad S^\circ(x, T) = f(x) \tag{3.5}$$

При этом предполагается, что

2) решение краевой задачи (3.5) существует и имеет ограниченные по модулю непрерывные производные по x до второго порядка включительно.

Функция S^1 из разложения (3.3) — решение задачи

$$S_t^1 + (\nabla H(x, t, S_x^\circ; 0), S_x^1) + H_\varepsilon(x, t, S_x^\circ; 0) = 0, \quad S^1(x, T) = 0 \tag{3.6}$$

Уравнение первого приближения (3.5) есть уравнение Беллмана исходной задачи, рассматриваемой при $\varepsilon = 0$, как задачи оптимального управления.

Будем предполагать, что эта задача решена, т. е. найдем синтез оптимального управления $u^\circ(x, t)$, функцию $S^\circ(x, t)$ и соответствующее поле оптимальных траекторий

$$x = \psi(t, y) \tag{3.7}$$

Здесь $\psi(t, y)$ — вектор-функция, y — n-мерный вектор произвольных постоянных. В дополнение к уже сделанным предположениям будем полагать, что выполнено следующее условие:

3) равенство (3.7) можно разрешить относительно y, т. е. получить зависимость

$$y = \varphi(x, t) \tag{3.8}$$

Чтобы решить краевую задачу (3.6), определяющую второе приближение, выпишем систему уравнений, определяющую характеристики уравнения (3.6)

$$dx/dt = -\nabla H(x, t, S_x^\circ; 0), \quad dS^1/dt = -H_\varepsilon(x, t, S_x^\circ; 0) \tag{3.9}$$

Учитывая введенное обозначение (3.2) и известное равенство $S_x^\circ = -p$ (p — вектор сопряженных переменных для исходной системы при $\varepsilon = 0$), заметим, что первое уравнение (3.9), согласно принципу максимума, определяет оптимальные траектории исходной системы с $\varepsilon = 0$. Решение первого уравнения из (3.9) задается равенством (3.7), а система первых интегралов — равенством (3.8). Общее решение задачи Коши (3.6), как следует из (3.7)—(3.9), определяется выражением

$$S^1(x, t) = -\int_{t_0}^{t} H_\varepsilon(\xi, \tau, S_\xi^\circ(\xi, \tau); 0)\, d\tau \tag{3.10}$$

$$\xi = \psi(t, \varphi(x, t)) \tag{3.11}$$

Стратегии первого и второго игроков в первом приближении найдем из условия

$$\min_u \max_v \{(F(x, t, u, \varepsilon v), S_x^\circ + \varepsilon S^1)\} = (F(x, t, u^1, \varepsilon v^1), \tag{3.12}$$

$$S_x^\circ + \varepsilon S_x^1) = H(x, t, S_x^\circ + \varepsilon S_x^1; \varepsilon), \quad u \in P, \; v \in Q_1$$

Равенства (3.3), (3.7)—(3.12) определяют в явном виде приближенное решение уравнения Беллмана (3.1) и некоторые стратегии игроков, если известно решение соответствующей задачи оптимального управления.

В следующей теореме будет указана оценка погрешности для функций S° и $S^\circ + \varepsilon S^1$.

Теорема 3.1. Пусть выполняются условия 1) — 3). Тогда для $t \in [t_0, T]$ и $x \in R^n$ справедливы оценки

$$|S - S^\circ| \leqslant C_1\varepsilon, \quad |S - S^\circ - \varepsilon S^1| \leqslant C_2\varepsilon^2 \tag{3.13}$$

с некоторыми постоянными C_k $(k = 1, 2)$.

Доказательство. Из условия 1 и равенства (3.3) вытекает справедливость разложений

$$H(x, t, S_x; \varepsilon) = H(x, t, S_x^\circ; 0) + \varepsilon H_\varepsilon(x, t, S_x^\circ; \varepsilon_1) + (\nabla H(x, t, M_1; \varepsilon), S_x - S_x^\circ) \tag{3.14}$$

$$H(x, t, S_x; \varepsilon) = H(x, t, S_x^\circ; 0) + (\nabla H(x, t, S_x^\circ; 0), S_x - S_x^\circ) + \varepsilon H_\varepsilon(x, t, S_x^\circ; 0) + \varepsilon(\nabla H_\varepsilon(x, t, S_x^\circ; \varepsilon_2), S_x - S_x^\circ) + (N(x, t, M_2; \varepsilon)(S_x - S_x^\circ), S_x - S_x^\circ) + \varepsilon^2 H_{\varepsilon\varepsilon}(x, t, S_x^\circ; \varepsilon_3)$$

Здесь

$$M_i = S_x^\circ + \Theta_i(S_x - S_x^\circ), \quad 0 < \Theta_i < 1, \quad i = 1, 2,$$

$$0 < \varepsilon_k < \varepsilon, \quad k = 1, 2, 3$$

Используя разложение (3.14) из (3.1) и (3.5), найдем, что функция $Z = S - S^{\circ}$ удовлетворяет краевой задаче

$$Z_t + (\nabla H(x, t, M_1; \varepsilon), Z_x) = -\varepsilon H_\varepsilon(x, t, S_x^{\circ}; \varepsilon_1) \tag{3.15}$$
$$Z(x, T) = 0$$

Отсюда, учитывая ограниченность функции H_ε, получим первое из неравенств (3.13).

Покажем теперь, что справедлива оценка

$$|S_x - S_x^{\circ}| \leqslant C'\varepsilon, \quad C' = \mathrm{const} \tag{3.16}$$

Используя условия 1) и 2), продифференцируем по переменной x_i уравнение (3.15) и обозначим $w^i = Z_{x_i}$. Получим краевую задачу для системы уравнений

$$w_t^i + (\nabla H_{x_i}(x, t, M_1; \varepsilon), w) + (\nabla H(x, t, M_1; \varepsilon), w_x^i) = \tag{3.17}$$
$$= \varepsilon \frac{\partial}{\partial x_i} H_\varepsilon(x, t, S_x^{\circ}; \varepsilon_1)$$
$$w^i(x, T) = 0, \quad i = 1, \ldots, n; \quad w = (w^1, \ldots, w^n)$$

Система уравнений (3.17) линейна, поэтому в силу ограниченности функций $\partial H_\varepsilon / \partial x_i$ получим оценки

$$|w^i| \leqslant C^i\varepsilon, \quad i = 1, \ldots, n$$

с некоторыми постоянными C^i. Отсюда и следует неравенство (3.16). Используя разложение (3.14) и уравнения (3.5), (3.6) и (3.1), получим краевую задачу для функции $Z^1 = S - S^{\circ} - \varepsilon S^1$

$$Z_t^1 + (\nabla H(x, t, S_x^{\circ}; 0), Z_x^1) = -\varepsilon^2 H_{\varepsilon\varepsilon}(x, t, S_x^{\circ}; \varepsilon_3) -$$
$$- \varepsilon(\nabla H(x, t, S_x^{\circ}; \varepsilon_2), S_x - S_x^{\circ}) + (N(x, t; M_2; \varepsilon) \times$$
$$\times (S_x - S_x^{\circ}), S_x - S_x^{\circ}), \quad Z^1(x, T) = 0$$

Учитывая доказанное неравенство (3.16), получим, что правая часть последнего уравнения — величина порядка $O(\varepsilon^2)$, следовательно, справедлива вторая оценка (3.13) с некоторой постоянной C_2.

Обозначим через $W(x, t)$ решение краевой задачи

$$W_t + (F(x, t, u^1, \varepsilon v^1), W_x) = 0, \quad W(x, T) = f(x) \tag{3.18}$$

Здесь $u^1(x, t)$ и $v^1(x, t)$ — стратегии первого и второго игроков в первом приближении, найденные из условий (3.12).

Теорема 3.2. Пусть выполняются условия 1)—3). Тогда для функции W справедлива оценка

$$|S - W| \leqslant C\varepsilon^2 \tag{3.19}$$

Доказательство. Обозначим через $E = W - S^{\circ} - \varepsilon S^1$. Из (3.18), (3.5) и (3.6), учитывая обозначения (3.12), получим, что правая часть последнего уравнения — величина порядка ε^2. Поэтому с некоторой постоянной выполняется неравенство

$$|W - S^{\circ} - \varepsilon S^1| \leqslant C_3\varepsilon^2, \quad C_3 = \mathrm{const} \tag{3.20}$$

Используя неравенство (3.20) и доказанное в теореме 3.1 неравенство (3.13), получим справедливость (3.19).

Рассмотрим примеры, иллюстрирующие проведенные построения.

1°. Пусть движение конфликтно управляемого объекта описывается уравнением

$$d^2y/dt^2 = u\,(1 + v), \quad y\,(0) = y_0, \quad y^{\cdot}\,(0) = y_0^{\cdot}, \quad t \in [0,\ T]$$

Здесь y — скаляр, управляющее воздействие принимают значения $|\,u\,| \leqslant 1$, $|\,v\,| \leqslant \varepsilon < 1$. В качестве платы рассматривается величина $[y\,(T)]^2$.

Введем новую переменную $x = y^{\cdot}\,(T - t) + y$. Исходное уравнение запишется в виде

$$dx/dt = (T - t)u\,(1 + v), \quad x\,(0) = x_0$$

Функционал платы при такой замене переменной сохраняет вид — $[x\,(T)]^2$. Построим приближенные стратегии первого и второго игроков, используя результаты п. 2. Положим $v^\circ \equiv 0$. Функция S° и стратегия u° определяются из решения краевой задачи

$$S_\tau{}^\circ = \tau \min_{|u| \leqslant 1} \{uS_x{}^\circ\} = -\tau\,|\,S_x{}^\circ\,|, \quad S^\circ(x,\ 0) = x^2$$

Здесь $T - t = \tau$ — обратное время, $u^\circ = -\operatorname{sign} S_x{}^\circ$. Решение этой задачи имеет вид

$$S^\circ = \begin{cases} (|\,x\,| - \tau^2/\,2)^2, & |\,x\,| \geqslant \tau^2/\,2 \\ 0 & ,\ |\,x\,| < \tau^2/\,2 \end{cases}$$

Стратегия u° однозначно определена в области $|\,x\,| \geqslant \tau^2\,/\,2$, где $u^\circ = -\operatorname{sign} x$, и неоднозначна в остальной части области. Стратегия v^1 находится из условия

$$\max_{|v| \leqslant \varepsilon} \{u^\circ(1 + v)\,S_x{}^\circ\}$$

откуда следует, что $v^1 = -\,\varepsilon$ в области $|\,x\,| \geqslant \tau^2\,/\,2$. В области $|\,x\,| < \tau^2\,/\,2$ стратегия v^1 определена неоднозначно. Согласно п. 2 стратегия v^1 — приближенная оптимальная стратегия второго игрока в первом приближении. Доопределим в области $|\,x\,| < \tau^2\,/\,2$ стратегию v^1, полагая $v^1 = -\,\varepsilon$, и найдем функцию S^1 и стратегию u^1.

Можно убедиться, что функция $\alpha_1{}^\varepsilon\,(x,\ t)$ такова, что

$$\alpha_1{}^\varepsilon\,(x,\ t) = u^1\,(1 + v^1)S_x{}^1 - u^1\,(1 + v^2)S_x{}^1 \equiv 0$$

В силу неравенства (2.11) $S^1 = S$ является функцией Беллмана рассматриваемой задачи. Согласно следствию пара стратегий $(u^1\,(x,\ t),\ v^1\,(x,\ t))$, доставляет седловую точку рассматриваемой задачи.

2°. Рассмотрим плоское управляемое движение с люфтом

$$dx_1/dt = x_3, \quad dx_3/dt = -x_3 + u_1 \cos v - u_2 \sin v$$

$$dx_2/dt = x_4, \quad dx_4/dt = -\,x_4 + u_1 \sin v + u_2 \cos v$$

$$t \in [0,\ T], \quad x_i\,(0) = x_i{}^\circ, \quad i = 1,\ 2,\ 3,\ 4$$

Цель первого игрока — выбрать управление u, подверженное ограничению $u_1^2 + u_2^2 \leqslant 1$, таким образом, чтобы минимизировать значение функционала конечного состояния

$$J = \sqrt{x_1^2(T) + x_2^2(T)}$$

Цель второго игрока, располагая выбором v (люфтом), таким, что $|v| \leqslant \varepsilon < \pi/2$, минимизировать значение функционала J.

Будем решать минимаксную задачу. Применим метод малого параметра, изложенный в п. 3, полагая ε достаточно малым числом. Ищем решение уравнения Беллмана в виде $S = S^\circ + \varepsilon S^1 + \ldots$. Краевая задача для функции S°, являющейся первым приближением, имеет вид

$$S_\tau^\circ = x_3 S_{x_1}^\circ + x_4 S_{x_2}^\circ - x_3 S^\circ_{x_3} - x_4 S_{x_4}^\circ + \min_{u_1^2+u_2^2 \leqslant 1} \{u_1 S_{x_3}^\circ + u_2 S_{x_4}^\circ\}$$

$$S^\circ(x, 0) = \sqrt{x_1^2 + x_2^2}$$

Здесь $T - t = \tau$ — обратное время. Вычисляя минимум, получим, что

$$\min_{u_1^2+u_2^2 \leqslant 1} \{u_1 S_{x_3}^\circ + u_2 S_{x_4}^\circ\} = -\sqrt{[S_{x_3}^\circ]^2 + [S_{x_4}^\circ]^2} = R$$

и достигается на векторе u° с компонентами

$$u_1^\circ = S^\circ_{x_3} / R, \quad u_2^\circ = S^\circ_{x_4} / R$$

Краевая задача для второго приближения имеет вид

$$S_\tau^1 = x_3 S^1_{x_1} + x_4 S^1_{x_2} - x_3 S^1_{x_3} - x_4 S^1_{x_4} + u_1^\circ S^1_{x_3} + u_2^\circ S^1_{x_4}$$

$$S^1(x, 0) = 0$$

Из единственности решения задачи Коши для уравнения в частных производных первого порядка гиперболического типа следует, что $S^1(x, t) \equiv 0$. Непосредственной проверкой убеждаемся, что функция S°, удовлетворяющая краевой задаче для первого приближения, задается выражением

$$S^\circ(x, t) = \{[x_1 + (1 - e^{-\tau}) x_3]^2 + [x_2 + (1 - e^{-\tau}) x_4]^2\}^{1/2} + \\ + (e^{-\tau} + \tau) - 1$$

Отсюда следует, что

$$u_1^\circ = -\frac{x_1 - (1 - e^{-\tau})\, x_3}{W}, \quad u_2^\circ = -\frac{x_2 + (1 - e^{-\tau})\, x_4}{W}$$

$$W = \{[x_1 + (1 - e^{-\tau}) x_3]^2 + [x_2 + (1 - e)^{-\tau} x_4]^2\}^{1/2}$$

Найдем приближенную минимаксную стратегию и стратегию второго игрока из условия (3.11)

$$\min_u \max_v \{\cos v\, (u_1 S_{x_3}^\circ + u_2 S_{x_4}^\circ) + \sin v\, (u_1 S_{x_4}^\circ - u_2 S_{x_3}^\circ)\} = \\ = \min_u \max_v \{\sqrt{[S_{x_3}^\circ]^2 + [S_{x_4}^\circ]^2 (u_1^2 + u_2^2)}\, \sin(v + \alpha)\}$$

Здесь α — угол, такой, что

$$\alpha = \operatorname{arctg}\left[\frac{u_1 S_{x_3}^\circ - u_2 S_{x_4}^\circ}{u_1 S_{x_4}^\circ - u_2 S_{x_3}^\circ}\right]$$

Если $-\pi/2 \leqslant \alpha \leqslant \pi/2 - \varepsilon$, то минимакс написанного выше выражения достигается при $v^1 = \varepsilon$ и $u^1 = u^\circ$. Если же $\pi/2 - \varepsilon < \alpha \leqslant \pi/2$, то $-v^1 = \pi/2 - \alpha$ и $u^1 = 0$. Из физического смысла задачи следует, что второй случай реализуется, когда исходная система приходит в точку ($x_1 = 0$, $x_2 = 0$) при $u = 0$. Множество таких точек задается уравнениями

$$x_1 = (e^{-\tau} - 1)x_3, \quad x_2 = (e^{-\tau} - 1)x_4$$

Учитывая это обстоятельство, получим, что приближенная минимаксная стратегия $u^1 = u^\circ$. Из теоремы 3.2 следует утверждение, что при каждом фиксированном x решение S° отличается от истинного решения уравнения Беллмана на величину порядка $O(\varepsilon^2)$.

Автор благодарит Ф. Л. Черноусько, Г. К. Пожарицкого и Б. Н. Соколова за полезное обсуждение работы и замечания.

Поступила 12 XI 1974

ЛИТЕРАТУРА

1. *Филиппов А. Ф.* Дифференциальные уравнения с разрывной правой частью. Матем. сб., 1960, т. 51 (93), вып. 1.
2. *Красовский Н. Н.* Игровые задачи о встрече движений. М., «Наука», 1970.
3. *Красовский Н. Н.* Игровые задачи динамики. Изв. АН СССР. Техн. кибернетика, 1969, № 5.
4. *Байбазаров М.* Достаточные условия оптимальности в дифференциальных играх. ПММ, 1971, т. 35, вып. 6.
5. *Мак-Кинси Дж.* Введение в теорию игр. М., Физматгиз, 1960.
6. *Байбазиров М.* Минимаксное сближение в регулярном случае. В сб.: Экстремальные стратегии в позиционных дифференциальных играх. Свердловск, 1974.
7. *Черноусько Ф. Л.* Некоторые задачи оптимального управления с малым параметром. ПММ, 1968, т. 32, вып. 1.
8. *Колмановский В. Б., Черноусько Ф. Л.* Задачи оптимального управления при неполной информации. Тр. IV зимней школы по математическому программированию и смежным вопросам, вып. 1, М., 1971.
9. *Соляник А. И., Черноусько Ф. Л.* Приближенный метод синтеза оптимального управления системой, подверженной случайным возмущениям. ПММ, 1972, т. 36, № 5.
10. *Братусь А. С.* Приближенное решение уравнения Беллмана для одного класса задач оптимального управления конечным состоянием. ПММ, 1973, т. 37, № 3.
11. *Братусь А. С.* Метод приближенного решения уравнения Беллмана для задач оптимального управления системой, подверженной случайным возмущениям. ПММ, 1975, т. 39, № 2.