Прикладная математика и механика. Том 51, вып. 4, 1987

ДИФФЕРЕНЦИАЛЬНАЯ ИГРА НЕОГРАНИЧЕННОЙ ПРОДОЛЖИТЕЛЬНОСТИ

Адиатулина Р. А., Тарасьев А. М.

Исследуются свойства функции цены (ФЦ) в дифференциальной игре неограниченной продолжительности с обесценивающимся функционалом качества и проводится сравнение двух методов аппроксимации ФЦ. Изучаемая ФЦ не удовлетворяет условию Липшица, что связано с видом функционала. Поэтому не удается показать в общем случае справедливость дифференциальных неравенств для производных по направлению традиционного вида. Для преодоления этой трудности предлагается обобщение понятия производной по направлению непрерывной функции. Оно заключается в «размывании» направления приращениями более высокого порядка малости, чем приращение аргумента. В терминах производной по направлению и сопряженных производных получены необходимые и достаточные условия, которым должна удовлетворять ФЦ игры. Указано, что дифференциальные неравенства, используемые для определения вязких решений, и неравенства данной работы эквивалентны в каждой позиции. Показано также, что метод дискретной аппроксимации стационарного уравнения Гамильтона—Якоби для задач управления применим и для задач теории дифференциальных игр. Устанавливается эквивалентность этого метода с ранее известной попятной процедурой.

Задачи исследуемого типа возникают, например, при моделировании процессов с обесценивающимся показателем качества. Ранее [1—3] подобные задачи оптимального программного управления изучались в отсутствие неконтролируемой помехи, которую необходимо рассматривать как игрока-противника.

1. Постановка задачи и предварительные результаты. Рассматривается управляемая система

$$x^{\cdot} = f(x, u, v),\quad x \in R^n,\quad u \in P \subset R^p,\quad v \in Q \subset R^q \tag{1.1}$$

с функционалом качества

$$J(x(\cdot), u(\cdot), v(\cdot)) = \int_{t_0}^{+\infty} e^{-\lambda\tau} g(x(\tau), u(\tau), v(\tau))\, d\tau,\qquad \lambda > 0 \tag{1.2}$$

Здесь P и Q — компакты; функции $f(\cdot)$ и $g(\cdot)$ непрерывны по совокупности переменных, удовлетворяют условию Липшица по x с постоянной L и ограничены постоянной K; u — управление, v — помеха, $t_0 \in [0, +\infty)$ — начальный момент времени.

Будем исследовать игру (1.1), (1.2) в рамках формализации [4, 5]. Добавим к системе (1.1) $(n+1)$-е уравнение

$$y^{\cdot} = \left\| \begin{matrix} x^{\cdot} \\ z^{\cdot} \end{matrix} \right\| = \left\| \begin{matrix} f(x, u, v) \\ e^{-\lambda t} g(x, u, v) \end{matrix} \right\| \tag{1.3}$$

$$y(t_0) = \left\| \begin{matrix} x(t_0) \\ z(t_0) \end{matrix} \right\| = y_0 = \left\| \begin{matrix} x_0 \\ z_0 \end{matrix} \right\|$$

$$x \in R^n,\ z \in R^1,\ t \in [t_0, +\infty),\ u \in P,\ v \in Q,\ \lambda > 0$$

которое в дифференциальной форме задает функционал (1.2). Функционал платы определим соотношением

$$J^*(y(\cdot)) = \lim_{T\to\infty} z(T) \tag{1.4}$$

Здесь $z(T)$ — значение $(n+1)$-й координаты движения $y(\cdot)$ системы (1.3) в момент времени T. Отметим, что при $z_0 = 0$ значения функционалов J (1.2) и J^* (1.4) совпадают.

Рассмотрим игру (1.3), (1.4) в классах позиционных стратегий первого игрока $(t, y) \mapsto U(t, y) : [0, +\infty) \times R^{n+1} \mapsto P$ и контрстратегий $(t, y, u) \mapsto V(t, y, u) : [0, +\infty) \times R^{n+1} \times P \mapsto Q$ второго игрока. Множества стратегий U и контрстратегий V обозначаются $\mathbf{U}$ и $\mathbf{V}$ соответственно. Основываясь на известных результатах [4, 5], можно доказать, что игра (1.3), (1.4) имеет цену

$$\omega^\circ(t_0, y_0) = \inf_U \sup_{y(\cdot)} \lim_{T\to+\infty} z(T) = \sup_V \inf_{y(\cdot)} \lim_{T\to+\infty} z(T) \tag{1.5}$$

$$U \in \mathbf{U},\ V \in \mathbf{V}; \qquad y(\cdot) = \begin{Vmatrix} x(\cdot) \\ z(\cdot) \end{Vmatrix}$$

Супремум (инфимум) по $y(\cdot)$ вычисляется на множестве $Y(t_0, y_0, U)$ (множестве $Y(t_0, y_0, V)$), элементами которого являются движения $y(\cdot)$ системы (1.3), порожденные стратегией U (контрстратегией V).

Укажем некоторые свойства ФЦ $(t, y) \mapsto \omega^\circ(t, y)$.

Свойство 1. Пусть $\omega_T^\circ : [0, T] \times R^{n+1} \mapsto R^1$ — ФЦ в игре конечной продолжительности T $(T \in [0, +\infty))$ с динамикой (1.3) и функционалом платы

$$J_T^*(y(\cdot)) = z(T) \tag{1.6}$$

Тогда

$$\sup_{(t, y) \in [0, T] \times R^{n+1}} |\omega^\circ(t, y) - \omega_T^\circ(t, y)| \leqslant K\lambda^{-1}e^{-\lambda T} \tag{1.7}$$

Свойство 2. ФЦ ω° представима в виде

$$\begin{gathered} \omega^\circ(t, y) = \omega^\circ(t, x, z) = z + e^{-\lambda t} w^\circ(x) \\ w^\circ(x) = \omega^\circ(0, x, 0) \\ x \in R^n,\ z \in R^1,\ t \in [0, +\infty) \end{gathered} \tag{1.8}$$

В силу (1.8) функция $w^\circ : R^n \mapsto R^1$ — ФЦ игры

$$\begin{gathered} \dot{x} = f(x, u, v),\ x(0) = x_0,\ u \in P,\ v \in Q \\ J_0(x(\cdot), u(\cdot), v(\cdot)) = \int_0^{+\infty} e^{-\lambda\tau} g(x(\tau), u(\tau), v(\tau))\, d\tau, \qquad \lambda > 0 \end{gathered} \tag{1.9}$$

В дальнейшем будем рассматривать функцию w°.

Свойство 3. ФЦ w° ограничена

$$\sup_{x \in R^n} |w^\circ(x)| \leqslant K\lambda^{-1} \tag{1.10}$$

Из оценки (1.7), справедливости условия Липшица (см. [4, 5]) для функции ω_T° и соотношения (1.8) следует неравенство

$$\begin{gathered} |w^\circ(x_1) - w^\circ(x_2)| \leqslant L\, \|x_1 - x_2\| \int_0^T e^{(L-\lambda)\tau} d\tau + \\ + K\lambda^{-1}e^{-\lambda T}; \qquad x_1, x_2 \in R^n, \qquad T \in [0, +\infty) \end{gathered} \tag{1.11}$$

Свойство 4. ФЦ w° непрерывна по Гельдеру

$$\sup_{x_1 \neq x_2} \frac{|w^\circ(x_1) - w^\circ(x_2)|}{\|x_1 - x_2\|^\gamma} \leqslant C \tag{1.12}$$

$$\gamma = \begin{cases} 1, & \mu > 1 \\ (0, 1), & \mu = 1, \\ \mu, & \mu < 1 \end{cases} \qquad C = \begin{cases} (\mu - 1)^{-1}, & \mu > 1 \\ (\mu(1-\gamma))^{-1}(K/L)^{1-\gamma}, & \mu = 1 \\ (\mu(1-\mu))^{-1}(K/L)^{1-\mu}, & \mu < 1, \end{cases} \quad \mu = \lambda/L$$

Замечание 1. Из оценок (1.12) следует, что в случае $\mu > 1$ ФЦ w° удовлетворяет условию Липшица. В случае же $\mu \leqslant 1$ можно привести примеры игр вида (1.3), (1.4), т. е. так подобрать функции $f(\cdot)$ и $g(\cdot)$, что ФЦ w° не будет удовлетворять условию Липшица.

2. Дифференциальные неравенства. Наиболее важными свойствами ФЦ являются так называемые свойства стабильности (u- и v-стабильности), которые можно трактовать как принцип оптимальности (суб-, супероптимальности) динамического программирования. Были предложены [6, 7] дифференциальные неравенства, которые выражают свойства стабильности в инфинитезимальной форме для ФЦ, удовлетворяющих условию Липшица, в игре ограниченной продолжительности. Аналогичные неравенства можно получить для непрерывной по Гельдеру ФЦ в игре (1.9) неограниченной продолжительности.

Введем следующие обозначения и определения.

Символом IS обозначим класс функций $r(\cdot): [0, +\infty) \mapsto R^n$, таких, что $\lim \| r(\delta) \| \delta^{-1} = 0$ при $\delta \downarrow 0$.

Определение. Нижнюю и верхнюю производные непрерывной функции w: $R^n \mapsto R^1$ в точке x по направлению h определим соответственно равенствами

$$(2.1)\qquad \begin{aligned} &\partial_- w(x) \,|\, (h) = \inf_{r(\cdot)\in \mathrm{IS}} \liminf_{\delta\downarrow 0} \Delta w \delta^{-1} \\ &\partial_+ w(x) \,|\, (h) = \sup_{r(\cdot)\in \mathrm{IS}} \limsup_{\delta\downarrow 0} \Delta w \delta^{-1} \\ &\Delta w = w(x + \delta h + r(\delta)) - w(x) \end{aligned}$$

(если функция w удовлетворяет условию Липшица, то в (2.1) можно полагать $r(\cdot) \equiv 0$ и опустить операции inf, sup по $r(\cdot) \in \mathrm{IS}$; такое определение производных по направлению использовалось в [6—8]).

Пусть (co A — выпуклая оболочка множества A)

$$\begin{aligned} &H_1(x, u) = \mathrm{co}\left\{\begin{Vmatrix} f(x, u, v) \\ g(x, u, v) \end{Vmatrix} : v \in Q\right\}, \quad u \in P \\ &H_2(x, v) = \mathrm{co}\left\{\begin{Vmatrix} f(x, u, v) \\ g(x, u, v) \end{Vmatrix} : u \in P\right\}, \quad v \in Q \\ &H(x, l) = \max_{v\in Q} \min_{u\in P} (\langle l, f(x, u, v)\rangle + g(x, u, v)) \\ &x \in R^n, \qquad l \in R^n \end{aligned}$$

Функция $H(\cdot)$ называется гамильтонианом игры (1.9).

Теорема 1. Для того чтобы функция w: $R^n \mapsto R^1$ была ФЦ игры (1.9), необходимо и достаточно выполнение следующих условий:

1) функция w ограничена и непрерывна по Гельдеру с оценками (1.10), (1.12);

2) при всех $x \in R^n$ справедливы неравенства

$$(2.2)\qquad \begin{aligned} &\lambda w(x) \geqslant \sup_v \inf_h (h_2 + \partial_- w(x) \,|\, (h_1)) \\ &v \in Q, \qquad h = \begin{Vmatrix} h_1 \\ h_2 \end{Vmatrix} \in H_2(x, v) \\ &\lambda w(x) \leqslant \inf_u \sup_h (h_2 + \partial_+ w(x) \,|\, (h_1)) \\ &u \in P, \qquad h = \begin{Vmatrix} h_1 \\ h_2 \end{Vmatrix} \in H_1(x, u) \end{aligned}$$

или при всех $(x, l) \in R^n \times R^n$ справедливы эквивалентные неравенства

$$(2.3)\qquad \sup_{h \in R^n} (\langle l, h\rangle - \partial_- w(x) \mid (h)) \geqslant -\lambda w(x) + H(x, l)$$

$$\inf_{h \in R^n} (\langle l, h\rangle - \partial_+ w(x) \mid (h)) \leqslant -\lambda w(x) + H(x, l)$$

Величины, стоящие в левой части неравенств (2.3), называются соответственно верхней и нижней сопряженными производными функции w, вычисленными в точке x [7].

Замечание 2. В точках дифференцируемости функции w условия (2.2) и (2.3) обращаются в стационарное уравнение Гамильтона—Якоби

$$(2.4)\qquad -\lambda w(x) + H(x, \operatorname{grad} w(x)) = 0$$

Если функция w удовлетворяет условию Липшица, то она почти всюду дифференцируема и, следовательно, почти всюду удовлетворяет уравнению (2.4). Ранее [9, 10] было введено понятие вязкого решения уравнения (2.4) и доказаны существование и единственность такого решения. Определение этого понятия содержит отличные по форме от (2.2) и (2.3) дифференциальные неравенства в терминах суб- и супердифференциалов [10]

$$(2.5)\qquad -\lambda w(x) + H(x, d) \leqslant 0, \quad x \in R^n, \quad d \in D_* w(x)$$

$$-\lambda w(x) + H(x, d) \geqslant 0, \quad x \in R^n, \quad d \in D^* w(x)$$

(множества $D_* w(x)$ и $D^* w(x)$ — соответственно суб- и супердифференциалы функции w в точке x). Можно показать, что соотношения (2.5) эквивалентны неравенствам (2.3), т. е. ФЦ w° игры (1.9) — вязкое решение уравнения (2.4).

3. Аппроксимация функции цены. Рассмотрим два аппроксимационных метода. Первый метод — известная [5, 11] попятная процедура, на каждом шаге которой работает оператор программного максимина в той или иной форме. Второй метод — метод дискретной аппроксимации стационарного уравнения Гамильтона — Якоби (2.4) был предложен [1, 2] для построения функции оптимального результата в задаче управления вида ($u(\cdot)$ — измеримое управление)

$$(3.1)\qquad \inf\left\{ \int_0^{+\infty} e^{-\lambda\tau} g^*(x(\tau), u(\tau))\, d\tau : u(\cdot) : [0, +\infty) \mapsto P \right\}$$

где $x(\cdot)$ определяется дифференциальным уравнением

$$(3.2)\qquad \dot{x}(t) = f^*(x(t), u(t)), \quad t \in [0, +\infty), \quad x(0) = x_0$$

Техника доказательства сходимости дискретной аппроксимации уравнения

$$(3.3)\qquad -\lambda w(x) + \min_{u \in P} (\langle \operatorname{grad} w(x), f^*(x, u)\rangle + g^*(x, u)) = 0$$

отвечающего задаче (3.1), (3.2), и оценка этой сходимости основаны [1, 2] на понятии вязкого решения. Так как вязкое решение уравнения (3.3) и ФЦ w° — понятия, эквивалентные в силу замечания 2, то эта техника с соответствующими изменениями может быть применена в игре (1.9) для получения аналогичных результатов.

Перейдем к более подробному рассмотрению упомянутых конструкций.

Начнем с попятной процедуры. Для ее применения воспользуемся свойствами 1, 2 ФЦ игры (1.3), (1.4).

Из (1.7), (1.8) имеем

$$(3.4)\qquad \sup | w^\circ(x) - \omega_T^\circ(0, x, 0)| \leqslant K\lambda^{-1} e^{-\lambda T}, \quad T \in [0, +\infty)$$

Здесь и всюду далее операция sup осуществляется по $x \in R^n$

Для аппроксимации ФЦ $(t, x, z) \mapsto \omega_T^\circ(t, x, z)$: $[0, T] \times R^n \times R^1 \mapsto R^1$, $\omega^\circ(T, x, z) = z$ в игре (1.3), (1.6) с фиксированным моментом окончания будем использовать попятную процедуру. Она заключается в следующем. Отрезок времени $[0, T]$ разбивается на m равных частей с шагом $h > 0$, так что $T = m \cdot h$. В момент времени T аппроксимация $\omega_T^h(T, \cdot, \cdot)$ ФЦ $\omega_T^\circ(T, \cdot, \cdot)$ определяется формулой

$$\omega_T^h(T, x, z) = z, \ x \in R^n, \ z \in R^1 \tag{3.5}$$

Предполагается далее, что в момент $(i+1)h$ аппроксимация $\omega_T^h((i+1)h, \cdot, \cdot)$ $(i = 0, \ldots, m-1)$ построена. Для нахождения аппроксимации $\omega_T^h(ih, \cdot, \cdot)$ применяется формула

$$\begin{aligned}\omega_T^h(ih, x, z) = \max_{v \in Q} \min_{u \in P} \{&\omega_T^h((i+1)h, x + hf(x, u, v), \\ &z + he^{-\lambda(ih)}g(x, u, v))\} \\ &x \in R^n, \ z \in R^1, \ i = 0, \ldots, m-1\end{aligned} \tag{3.6}$$

Оператор, стоящий в правой части соотношения (3.6), является оператором программного максимина, который обычно используется в попятных конструкциях [5, 11].

Из (3.5), (3.6) следует, что

$$\omega_T^h(ih, x, z) = z + e^{-\lambda(ih)}w_T^h(ih, x), \qquad i = 0, \ldots, m \tag{3.7}$$

Здесь

$$\begin{aligned}\omega_T^h(ih, x) = \max_{v \in Q} \min_{u \in P} \{&e^{-\lambda h}w_T^h((i+1)h \\ &x + hf(x, u, v)) + hg(x, u, v)\}, \ i = 0, \ldots, m-1\end{aligned} \tag{3.8}$$

$$w_T^h(mh, x) = 0, \ x \in R^n \tag{3.9}$$

Пользуясь техникой получения оценки (15.1) из [5] для попятной процедуры (3.5) — (3.9), можно доказать следующее утверждение.

Теорема 2. Аппроксимация $w_T^h(0, \cdot)$ равномерно сходится к ФЦ $\omega_T^\circ(0, \cdot, 0)$, когда $h \downarrow 0$. При этом справедлива оценка

$$\sup |\omega_T^\circ(0, x, 0) - w_T^h(0, x)| \leqslant h^{1/2}L \int_0^T e^{(L-\lambda)\tau}\, d\tau \tag{3.10}$$

Из (3.4), (3.10) следует соотношение

$$\begin{aligned}&\sup |w^\circ(x) - w_T^h(0, x)| \leqslant h^{1/2}L \int_0^T e^{(L-\lambda)\tau}\, d\tau + K\lambda^{-1}e^{-\lambda T}, \\ &T \in [0, +\infty)\end{aligned} \tag{3.11}$$

Неравенство (3.11) влечет оценку

$$\sup |w^\circ(x) - w_{T_*}^h(0, x)| \leqslant Ch^{\gamma/2} \tag{3.12}$$

где числа γ и C определяются соотношением (1.12), $h \in (0, 1)$, момент времени T_* зависит от h и определяется из условия минимума правой части в (3.11).

Рассмотрим теперь метод дискретной аппроксимации [1, 2] в применении к стационарному уравнению Гамильтона — Якоби (2.4), соответствующему игре (1.9).

Дискретной аппроксимацией уравнения (2.4) называется уравнение

$$\begin{aligned}-w^h(x) + \max_{v \in Q} \min_{u \in P} (&(1 - \lambda h)w^h(x + hf(x, u, v)) + \\ &+ hg(x, u, v)) = 0, \quad x \in R^n, \quad h \in (0, 1/\lambda)\end{aligned} \tag{3.13}$$

Справедливы следующие утверждения, которые можно доказать аналогично теоремам 2.1, 2.2 из [1].

Теорема 3. Пусть $h \in (0, 1/\lambda)$. Тогда (3.13) имеет единственное решение w^h в классе ограниченных, непрерывных по Гельдеру функций. Причем для функции w^h выполняются оценки (1.10), (1.12).

Доказательство теоремы 3 основано на том, что оператор П, определяемый соотношением

$$\Pi w(x) = \max_{v \in Q} \min_{u \in P} ((1 - \lambda h) w(x + hf(x, u, v)) + hg(x, u, v)) \tag{3.14}$$

есть оператор сжатия с коэффициентом сжатия $(1 - \lambda h)$. По принципу сжимающих отображений существует единственное ограниченное и непрерывное по Гельдеру решение w^h уравнения (3.13) и итерационная процедура $\overline{w}_p^h = \Pi \overline{w}_{p-1}^h$ с начальным приближением $\overline{w}_0^h$, удовлетворяющим условиям (1.10), (1.12), сходится к нему равномерно.

В частности, при $\overline{w}_0^h \equiv 0$ справедлива оценка

$$\sup | w^h(x) - \overline{w}_p^h(x) | \leqslant K\lambda^{-1}(1 - \lambda h)^p, \quad p = 1, 2, \ldots \tag{3.15}$$

Теорема 4. При $h \downarrow 0$ решение w^h уравнения (3.13) сходится локально равномерно к вязкому решению w° уравнения (2.4), которое в силу замечания 2 является ФЦ игры (1.9).

Теорема 5. Справедлива оценка

$$\sup | w^\circ(x) - w^h(x) | \leqslant Bh^{\gamma/2} \tag{3.16}$$

где $h \in (0, \min\{1/\lambda, 1\})$, гельдерова экспонента γ определяется соотношением (1.12), B — некоторая постоянная.

Из (3.15), (3.16) при $h \in (0, \min\{1/\lambda, 1\})$ получаем

$$\sup | w^\circ(x) - \overline{w}_p^h(x) | \leqslant Bh^{\gamma/2} + K\lambda^{-1}(1 - \lambda h)^p, \quad p = 1, 2, \ldots \tag{3.17}$$

Неравенство (3.17) влечет оценку

$$\sup | w^\circ(x) - (w^h{}_{p_*}(x) | \leqslant Gh^{\gamma/2} \tag{3.18}$$

где число γ определяется соотношением (1.12), G — некоторая постоянная $h \in (0, \min\{1/\lambda, 1\})$, число p_* зависит от h и определяется из условия $(1 - \lambda h)^{p_*} \leqslant h^{\gamma/2}$.

Из сравнения соотношений (3.8) и (3.14), (3.12) и (3.18) можно сделать вывод, что попятная процедура и метод дискретной аппроксимации, отличаясь по форме и способам доказательства сходимости и оценок, по существу дают одну и ту же процедуру аппроксимации ФЦ w° игры (1.9) с оценкой сходимости порядка $h^{\gamma/2}$.

ЛИТЕРАТУРА

1. *Dolcetta I. C.* On a discrete approximation of the Hamilton — Jacobi equation of dynamic programming // Appl. Math. Optimiz. 1983. V. 10. No. 4. P. 367—377.
2. *Dolcetta I. C., Ishii H.* Approximate solution of the Bellman equation of deterministic control theory // Appl. Math. Optimiz. 1984. V. 11. No. 2. P. 161—181.
3. *Lions P.-L., Souganidis P, E.* Differential games, optimal and directional derivatives of viscosity solutions of Bellman's and Isaac's Equations // SIAM J. Control & Optimiz. 1985. V. 23. No. 4. P. 566—583.
4. *Красовский Н. Н.* Управление динамической системой. М.: Наука. 1985. 518 с.
5. *Красовский Н. Н., Субботин А. И.* Позиционные дифференциальные игры. М.: Наука. 1974. 456 с.

6. *Субботин А. И., Субботина Н. Н.* Свойства потенциала дифференциальной игры // ПММ. 1982. Т. 46. Вып. 2. С. 204—211.
7. *Субботин А. И., Тарасьев А. М.* Сопряженные производные функции цены дифференциальной игры // Докл. АН СССР (ДАН СССР). 1985. Т. 283. № 3. С. 559—564.
8. *Субботин А. И., Тарасьев А. М.* Свойства стабильности функции цены дифференциальной игры и вязкие решения уравнений Гамильтона—Якоби // Пробл. управления и теории информ. 1986. Т. 15. № 6. С. 451—463.
9. *Crandall M. G., Lions P.-L.* Viscosity solutions of Hamilton — Jacobi equations // Trans. Amer. Math. Soc. 1983. V. 277. No. 4. P. 1—42.
10. *Crandall M. G., Evans L. C., Lions P.-L.* Some properties of viscosity solutions of Hamilton — Jacobi equations // Trans. Amer. Math. Soc. 1984. V. 282. No. 2. P. 487—502.
11. *Friedman A.* Differential games. N. Y.: Wiley. 1971. 350 p.

Свердловск

Поступила в редакцию
14.VIII.1986