Прикладная математика и механика. Том 45, вып. 2, 1981

УДК 62—50

АСИМТОТИЧЕСКОЕ ПОВЕДЕНИЕ ФУНКЦИИ БЕЛЛМАНА В ОДНОЙ СТОХАСТИЧЕСКОЙ ЗАДАЧЕ ОПТИМАЛЬНОГО УПРАВЛЕНИЯ

Овсеевич А. И.

Путем изучения асимптотики функции Беллмана задачи оптимального управления движением системы, возмущаемой белым шумом при ограниченных энергетических ресурсах управления, находится квазиоптимальное управление, дающее близкий к оптимальному результат при малых энергетических ресурсах. Определение асимптотики функции Беллмана и квазиоптимального управления требует решения введенных в [1] уравнений в частных производных. Найден сходящийся итерационный процесс их решения.

1. Постановка задачи. Пусть движение точки в n-мерном пространстве R^n описывается стохастическим дифференциальным уравнением

$$dx_t = u_t dt + dw_t, \quad x_s = x \tag{1.1}$$

где w_t — n-мерное броуновское движение. Предполагается, что управление u_t зависит от траектории x_τ до момента t и таково, что ($| u_t |$ — евклидова норма вектора u_t)

$$\int_s^T | u_t |^2 dt \leqslant Q^2 \tag{1.2}$$

Пусть, кроме того, $F(x)$ — некоторая фиксированная функция. Требуется минимизировать, математическое ожидание $E^u_{x,s}(F(x_T))$ случайной величины $F(x_T)$ при условии $x_s = x$. Такая постановка может рассматриваться как модельная задача оптимального управления движением при ограниченной энергетике управления, а величина Q интерпретируется как начальный запас энергетических ресурсов.

Эта задача и близкие к ней рассматривались в ряде работ [1, 2]. Ниже исследуется асимптотика функции Беллмана $S(x, t, Q) = \inf E^u_{x,t}(F(x_T))$ при $Q \to 0$. Именно выводится асимптотическое представление

$$S(x, t, Q) = S_0(x, t) + QS_1(x, t) + O(Q^2)$$

и уравнения в частных производных, которым удовлетворяют функции S_0 и S_1, а также сходящийся итерационный процесс для решения этих уравнений. Уравнения имеют следующий вид:

$$\frac{\partial S_0}{\partial t} + \frac{1}{2}\Delta S_0 = 0, \quad S_0(x, T) = F(x) \tag{1.3}$$

$$\frac{\partial S_1}{\partial t} + \frac{1}{2}\Delta S_1 + \min\left((u, \operatorname{grad} S_0) - \frac{|u|^2}{2} S_1\right) = 0, \quad S_1(x, T) = 0 \tag{1.4}$$

где Δ и grad — соответственно лапласиан и градиент в R^n, и называются в дальнейшем уравнением нулевого и первого приближения соответственно.

Далее, по S_0, S_1 строится такое квазиоптимальное управление u, что

$$S = E^u (F(x_T)) + O(Q^2)$$

Основная идея работы состоит в том, что функция $S_0 + QS_1$, так же как и S, является функцией Беллмана некоторой стохастической экстремальной задачи, причем множества допустимых управлений в обеих задачах совпадают, а функционалы различаются на величину порядка Q^2.

В дальнейшем, для простоты обозначений, всюду считается, что $n = 1$, т. е. x_t, u_t, w_t — скаляры.

2. Уточнение постановки задачи. Множество $U(Q)$ допустимых управлений с ресурсом Q состоит из таких функционалов $u_t = u(t, x_{\tau \leqslant t})$, которые измеримо зависят от траектории x_τ, $\tau \leqslant t$, случайного процесса x_t (таким образом, u_t — неантисипативный функционал в терминологии работы [3]) и для которых с вероятностью единица

$$\int_1^T u_t^2\, dt \leqslant Q^2$$

Относительно функции риска $F(x)$ предполагаем, что она бесконечно гладкая, с экспоненциально ограниченными производными, т. е. любая производная оценивается сверху через (const) exp $(\lambda\,|x|)$, где $\lambda = \mathrm{const}$. Это предположение запишем в виде $F \in C^\infty$ (exp). Определим функционалы

$$I(u) = E^u_{x,s}(F(x_T))$$

$$K(u) = E_{x,s}\Big(F(w_T)\Big(1 + \int_s^T u_t\, dw_t\Big)\Big)$$

Броуновское движение w_t, фигурирующее в записи $K(u)$, предполагается выходящим из x в момент s, а управление $u_t = u(t, w_{\tau \leqslant t})$ вычислено по траектории w_τ.

Теорема 1. $|I(u) - K(u)| \leqslant CQ^2$ при $u \in U(Q)$ и достаточно малом Q, причем $C = C(x, s)$ — функция экспоненциального роста (т. е. $C(x, s) \leqslant$ (const) exp $(\lambda |x|)$ при s, лежащем в ограниченном интервале).

Доказательство. Идея состоит в том, чтобы применить формулу Камерона — Мартина (см. [3—5])

$$I(u) = E(F(w_T) \exp X), \quad X = \int_s^T u_t\, dw_t - \frac{1}{2}\int_s^T u_t^2\, dt \tag{2.1}$$

разложить экспоненту в ряд и оставить в нем только члены первого порядка по Q, которые составляют функционал $K(u)$. Формула (2.1) верна для $u \in U(Q)$, поскольку в силу теорем 6.1 и 4.13 из [4] достаточно, чтобы

$$E\Big(\exp \frac{1}{2}\int_s^T u_t^2\, dt\Big) < \infty$$

но это следует из (1.2).

Таким образом, реализована часть программы, намеченной в конце п. 1. Остается решить задачу на минимум для упрощенного функционала $K(u)$. Начнем с приведения $K(u)$ к более удобному виду.

3. Преобразование функционала $K(u)$. Определим функцию S_0 формулой $S_0(x, s) = E_{x,s}(F(w_T))$, где броуновское движение w_t начинается в момент s из x.

Тогда, как известно (см., например, [5]), S_0 — гладкое решение уравнения нулевого приближения (1.3). Положим $\varphi = \partial S_0 / \partial x$.

При помощи формулы Ито (см. [3]), примененной к $S_0(w_t, t)$, с учетом (1.3) можно показать, что

$$E\left(F(w_T)\int_s^T u_t dw_t\right) = E\int_s^T \varphi(w_t, t)\, u_t\, dt \tag{3.1}$$

Обозначим правую часть (3.1) через $J(u)$. Тогда

$$K(u) = S_0(x, s) + J(u)$$

Для полной реализации программы, намеченной в конце п. 1, остается показать, что существует минимум $J(u)$ по $u \in U(Q)$ и $\min J(u) = QS_1(x, s)$, где S_1 удовлетворяет уравнению первого приближения (1.4).

4. Регуляризация экстремальной задачи. Дополним случайный процесс $x_t = w_t$ компонентой

$$Q_t = \left(Q^2 - \int_s^t u_\tau^2\, d\tau\right)^{1/2}$$

(см. [1, 6, 7]) и получим систему

$$dx_t = dw_t,\quad x_s = x$$
$$dQ_t = -(u_t^2/(2Q_t))\,dt,\quad Q_s = Q$$

Обозначим $\inf J(u)$ по $u \in U(Q)$ через $S_+(x, s, Q)$. Тогда естественно ожидать, что S_+ удовлетворяет уравнению Беллмана

$$\frac{\partial S_+}{\partial t} + \frac{1}{2}\frac{\partial^2 S_+}{\partial x^2} + \min\left(u\varphi - \frac{u^2}{2Q}\frac{\partial S_+}{\partial Q}\right) = 0 \tag{4.1}$$

$S_+ = 0$ при $Q = 0$ или при $t = T$.

Полагая $S_+(x, s, Q) = QS_1(x, S)$, получим для S_1 уравнение первого приближения

$$\frac{\partial S_1}{\partial t} + \frac{1}{2}\frac{\partial^2 S_1}{\partial x^2} + \min\left(u\varphi - \frac{u^2}{2}S_1\right) = 0 \tag{4.2}$$
$$S_1(x, T) = 0.$$

Заметим, что $\min(u\varphi - (u^2/2)s_1) = \varphi^2/(2S_1)$ при $S_1 < 0$. Вместе с граничным условием $S_1(x, T) = 0$ это показывает, что уравнение (4.2) сингулярно при $t = T$. Поэтому начнем с решения регуляризованной задачи

$$\frac{\partial \Sigma_\varepsilon}{\partial t} + \frac{1}{2}\frac{\partial^2 \Sigma_\varepsilon}{\partial x^2} + \min\left(u\varphi - \frac{u^2}{2}\Sigma_\varepsilon\right) = 0 \tag{4.3}$$
$$\Sigma_\varepsilon(x, T) = -\varepsilon,\quad \varepsilon > 0$$

которая оказывается связанной с минимизацией функционала $J_\varepsilon(u) = J(u) - \varepsilon E(Q_T)$. Решение регуляризованной задачи (4.3) получится с помощью стандартного беллмановского итерационного процесса.

5. Метод Беллмана. Положим $\Phi_0 = -\varepsilon$, $u_0 = 0$. Если Φ_n определено, то полагаем $u_{n+1} = \varphi / \Phi_n$ (поскольку $\varphi u_{n+1} - (u_{n+1}^2/2)\Phi_n = \min(\varphi u - (u^2/2)\Phi_n)$) и определяем Φ_{n+1} как решение задачи Коши

$$\frac{\partial \Phi_{n+1}}{\partial t} + \frac{1}{2}\frac{\partial^2 \Phi_{n+1}}{\partial x^2} + u_{n+1}\varphi - \frac{u_{u+1}^2}{2}\Phi_{n+1} = 0 \tag{5.1}$$

$$\Phi_{n+1}(x, T) = -\varepsilon$$

Заметим, что функции $\varphi = \partial S_0/\partial x$ и $S_0 = E(F(W_T))$ — бесконечно гладкие с экспоненциально ограниченными производными, поскольку F обладает этими свойствами (в обозначениях п. 2 $\varphi \in C^\infty(\exp)$). Следующая лемма обеспечивает возможность неограниченного продолжения итераций.

Лемма 1. Пусть u, $\psi \in C^\infty(\exp)$, $\psi \leqslant 0$. Тогда задача Коши

$$\frac{\partial \Phi}{\partial t} + \frac{1}{2}\frac{\partial^2 \Phi}{\partial x^2} + \psi - \frac{u^2}{2}\Phi = 0, \quad \Phi(x, T) = -\varepsilon$$

имеет единственное решение $\Phi \in C^\infty(\exp)$, причем $\Phi \leqslant -\varepsilon$.

Лемма 2. Последовательность Φ_n монотонно убывает и сходится к функции $\Sigma_\varepsilon \in C^\infty(\exp)$ — единственному решению задачи:

$$\frac{\partial \Sigma_\varepsilon}{\partial t} + \frac{1}{2}\frac{\partial^2 \Sigma_\varepsilon}{\partial x^2} + \frac{1}{2}\frac{\varphi^2}{\Sigma_\varepsilon} = 0, \quad \Sigma_\varepsilon(x, T) = -\varepsilon \tag{5.2}$$

Доказательство. Положим $W = \Phi_{n+1} - \Phi_n$. Тогда

$$W(x, T) = 0, \quad \frac{\partial W}{\partial t} + \frac{1}{2}\frac{\partial^2 W}{\partial x^2} - \frac{u_{n+1}^2}{2} W \geqslant 0.$$

Неравенство следует из того, что

$$\frac{\partial \Phi_n}{\partial t} + \frac{1}{2}\frac{\partial^2 \Phi_n}{\partial x^2} + u_{n+1}\varphi - \frac{u_{n+1}^2}{2}\Phi_n \leqslant \frac{\partial \Phi_n}{\partial t} + \frac{1}{2}\frac{\partial^2 \Phi_n}{\partial x^2} + u_n\varphi - \frac{u_n^2}{2}\Phi_n$$

из-за специального выбора u_{n+1}. Поэтому принцип максимума дает $W \leqslant 0$, что и требовалось. Для доказательства сходимости Φ_n достаточно установить нижнюю границу для Φ_n.

Положим теперь $\Sigma_\varepsilon^+(x, t, Q) = Q\Sigma_\varepsilon(x, t)$ и покажем, что Σ_ε^+ обладает некоторым «квазиоптимальным свойством».

Лемма 3. 1) $\Sigma_\varepsilon^+(x, s, Q) = \min J_\varepsilon(u)$, где функционал $J_\varepsilon(u) = J(u) - \varepsilon E(Q_T)$ определен в п. 4, а минимум берется по $u \in U(Q)$.

2) Минимум в 1) достигается с помощью управления

$$u_t^\varepsilon = (\varphi/\Sigma_\varepsilon)(w_t, t)\, Q_t^\varepsilon$$

$$Q_t^\varepsilon = Q \exp\left(-\frac{1}{2}\int_s^t \left(\frac{\varphi}{\Sigma_\varepsilon}\right)^2 (w_\tau, \tau)\, d\tau\right)$$

6. Решение уравнения первого приближения. Следующая теорема завершает выполнение программы, намеченной в конце п. 1.

Теорема 2. 1) Функции Σ_ε монотонно возрастают при $\varepsilon \downarrow 0$ и предельная функция $S_1 = \lim \Sigma_\varepsilon$ является обобщенным решением задачи Коши:

$$\frac{\partial S_1}{\partial t} + \frac{1}{2}\frac{\partial^2 S_1}{\partial x^2} + \min\left(u\varphi - \frac{u^2}{2} S_1\right) = 0, \quad S_1(x, T) = 0 \tag{6.1}$$

2) Пусть $S_+(x, t, Q) = QS_1(x, t)$. Тогда $S_+(x, t, Q) = \min J(u)$, где функционал $J(u)$ определен в п. 3, а минимум берется по $u \in U(Q)$.

3) Минимум в 2) достигается с помощью управления

$$u_t = Q_t(\varphi/S_1)(w_t, t)$$

Доказательство. 1) Пусть $\varepsilon_1 < \varepsilon_2$, $V_1 = \Sigma_{\varepsilon_1}$, $V_2 = \Sigma_{\varepsilon_2}$, $W = V_1 - V_2$. Тогда

$$\frac{\partial W}{\partial t} + \frac{1}{2}\frac{\partial^2 W}{\partial x^2} - \frac{1}{2}\frac{\varphi^2}{V_1 V_2} W = 0, \quad W(x, T) \geqslant 0$$

Отсюда по принципу максимума получим $W \geqslant 0$, что и требовалось.

Для доказательства 2) нужно установить некоторые априорные оценки для Σ_ε, равномерные по ε.

Пусть $\rho \in C^\infty$, $\rho(x) = (\mathrm{const})\,|x|$ при $|x| \geqslant 1$. Тогда, умножая (5.2) на $\Sigma_\varepsilon e^{-\rho}$, интегрируя по частям и применяя лемму Гронуолла, получим

$$\int_s^T \int \left|\frac{\partial \Sigma_\varepsilon}{\partial x}\right|^2 e^{-\rho}\, dx\, dt + \int |\Sigma_\varepsilon(x, s)|^2 e^{-\rho}\, dx \leqslant C\left(\int_s^T \int |\varphi|^2 e^{-\rho}\, dx\, dt + 1\right) \tag{6.2}$$

где C не зависит от ε (слагаемое 1 в правой части (6.2) происходит от того, что $\Sigma_\varepsilon(x, T) \neq 0$). Умножая (5.2) на $e^{-\rho}$, интегрируя по частям и пользуясь (6.2), получим, что

$$\int_s^T \int \frac{\varphi^2}{|\Sigma_\varepsilon|} e^{-\rho}\, dx\, dt \leqslant C\left(\int_s^T \int \varphi^2 e^{-\rho}\, dx\, dt + 1\right)^{1/2} \tag{6.3}$$

где C снова не зависит от ε.

Оценки (6.2), (6.3) позволяют перейти к пределу при $\varepsilon \to 0$ в равенстве

$$\int_s^T \int \Sigma_\varepsilon \left(-\frac{\partial f}{\partial t}\right) dx\, dt + \frac{1}{2}\int_s^T \int \Sigma_\varepsilon \frac{\partial^2 f}{\partial x^2}\, dx\, dt +$$

$$+ \frac{1}{2}\int_s^T \frac{\varphi^2}{\Sigma_\varepsilon} f\, dx\, dt = \int \Sigma_\varepsilon(x, s) f(x, s)\, dx - \varepsilon \int f(x, T)\, dx$$

где f — гладкая финитная функция, и получить, что $S_1 = \lim \Sigma_\varepsilon$ — обобщенное решение уравнения (6.1). Утверждение 1) доказано. Утверждение 2) теоремы сразу следует из леммы 3.

Доказательство утверждения 3) сводится к обоснованию возможности перехода к пределу при $\varepsilon \to 0$ в равенстве $\Sigma_\varepsilon^+ = J_\varepsilon(u^\varepsilon)$ из леммы 3. Напомним, что

$$u_t^\varepsilon = (\varphi/\Sigma_\varepsilon)(w_t, t)\, Q_t^\varepsilon$$

где

$$Q_t^\varepsilon = Q \exp\left(-\frac{1}{2}\int_s^t \left(\frac{\varphi}{\Sigma_\varepsilon}\right)^2 (w_\tau, \tau)\, d\tau\right)$$

Нам нужно показать, что u_t^ε сходится к $u\,(w_t,\, t,\, Q_t)$ при $\varepsilon \to 0$ и функции $\varphi\,(w_t,\, t)\, u_t^\varepsilon$ равномерно интегрируемы по произведению меры Лебега отрезка $[s,\, T]$ на винеровскую меру.

Для доказательства равномерной интегрируемости достаточно оценить

$$E \int_s^T |\varphi\,(w_t,\, t)\, u_t^\varepsilon|^2\, dt$$

сверху равномерно по ε. Это выражение оценивается через

$$E\left((\sup|\varphi\,(w_t, t)|^2) \int_s^T (u_t^\varepsilon)^2\, dt\right) \leqslant Q^2 E\,(\sup|\varphi\,(w_t,\, t)|^2), \quad s \leqslant t \leqslant T$$

поскольку $u^\varepsilon \in U\,(Q)$. Далее φ имеет экспоненциальный рост и поэтому остается заметить, что для любого λ

$$E\,(\sup \exp\,(\lambda\, |\, w_t\,|)) < \infty, \; s \leqslant t \leqslant T$$

7. *Пример*. Рассмотрим управляемое движение твердого тела, испытывающего случайные возмущения, вокруг неподвижной точки O. Уравнение движения имеет вид

$$dM_t = ([M_t,\, \omega_t] + u_t)dt + dw_t,\, M_s = x$$

где $M = J\omega$ — вектор кинетического момента относительно O в теле, ω — вектор угловой скорости в теле, J — тензор инерции, u — управление, w_t — трехмерное броуновское движение.

Требуется минимизировать математическое ожидание $E_{x,s}^u\,(|\,M_T\,|^2)$ квадрата модуля момента в фиксированный момент T при ограничении на управление вида (1.2).

Из леммы Ито следует, что скаляр $r_t = |\,M_t\,|$ удовлетворяет уравнению вида

$$dr_t = \left(\left(u_t,\, \frac{M_t}{|M_t|}\right) + \frac{2}{r_t}\right)dt + d\xi_t, \quad r_s = |\,x\,| \tag{7.1}$$

где ξ_t — скалярный винеровский процесс. Из той же леммы Ито получаем, что если векторный процесс x_t удовлетворяет уравнению

$$dx_t = u_t\, dt + dw_t,\, x_s = x$$

то процесс $|\,x_t\,|$ удовлетворяет уравнению, аналогичному (7.1), и задачи оптимизации $E_{x,s}^u\,(|\,M_T\,|^2)$ и $E_{x,s}^u\,(|\,x_T\,|^2)$ при ограничении (1.2) эквивалентны. Поэтому из теоремы 2 получаем квазиоптимальное, при малых значениях ресурса Q, управление, заданное в форме синтеза

$$u_t = u\,(t,\, M_t,\, Q_t) = Q_t \frac{2M_t}{S_1\,(t,\, |\,M_t\,|)}$$

где $S_1\,(t,\, r)$ — неположительное решение уравнения

$$\frac{\partial S_1}{\partial t} + \frac{1}{2}\left(\frac{\partial^2 S_1}{\partial r^2} + \frac{2}{r}\,\frac{\partial S_1}{\partial r}\right) + \frac{2r^2}{S_1} = 0, \quad S_1\,(T,\, r) = 0$$

Функция S_1 записывается в автомодельных переменных в виде $S_1\,(t,\, r) = r^2 f\,(\xi)$, где $\xi = r\,(T - t)^{-1/2}$, а $f\,(\xi)$ удовлетворяет уравнению

$$\tfrac{1}{2} f''\,(\xi)\,\xi^2 + f'\,(\xi)\,(\tfrac{1}{2}\xi^3 + 3\xi) + 3f\,(\xi) + 2/f\,(\xi) = 0$$

Методы численного решения уравнений такого типа изложены в [6].

Автор благодарит Черноусько Ф. Л. за постановку задачи.

ЛИТЕРАТУРА

1. *Черноусько Ф. Л.* Автомодельные решения уравнения Беллмана для задач оптимальной коррекции случайных возмущений.— ПММ, 1971, т. 35, вып. 2, с. 333.
2. *Черноусько Ф. Л., Колмановский В. Б.* Оптимальное управление при случайных возмущениях. М.: Наука, 1978. 351 с.
3. *Струк Д. В., Варадан С. Р.* Диффузионные процессы с непрерывными коэффициентами. Математика. Период. сб. переводов иностр. статей. М.: Мир, (1971), т. 15, вып. 6, с. 66; (1972), т. 16, вып. 1, с. 100.
4. *Липцер Р. Ш., Ширяев А. Н.* Статистика случайных процессов. М.: Наука, 1973. 696 с.
5. *Маккин Г. П.* Стохастические интервалы. М.: Мир, 1973. 182 с.
6. *Красовский Н. Н., Третьяков В. Е.* К задаче о встрече движений. Докл. АН СССР, 1967, т. 173 вып. 2, с. 285.
7. *Ушаков В. Н.* Экстремальные стратегии в дифференциальных играх с интегральными ограничениями.— ПММ, 1972, т. 36, вып. 1, с. 15.

Москва

Поступила в редакцию
2.II.1979