Прикладная математика и механика. Том 42, 1978

УДК 62—50

О ПРИБЛИЖЕННОМ СИНТЕЗЕ ОПТИМАЛЬНОГО УПРАВЛЕНИЯ СТОХАСТИЧЕСКИМИ КВАЗИЛИНЕЙНЫМИ СИСТЕМАМИ С ПОСЛЕДЕЙСТВИЕМ

В. Б. Колмановский, Л. Е. Шайхет

(Москва, Донецк)

Задачи управления квазилинейными детерминированными системами без запаздывания рассматривались в [1, 2]. В данной работе изучается управление квазилинейными стохастическими уравнениями, теория которых изложена в [3-6]. Вопрос о приближенном синтезе управления стохастическими системами с последействием имеет важное значение, поскольку построение точного оптимального управления ими удается получить лишь в исключительных случаях [7, 8]. В работе предложен алгоритм приближенного синтеза оптимального управления и развит метод получения оценок погрешности, отличающийся от полученных ранее [9, 10].

1. Пусть $\{Q, \sigma, P\}$ — фиксированное вероятностное пространство; $\{Q_t, t \geqslant 0\}$ — монотонно неубывающее семейство σ-алгебр, $Q_t \subset \sigma$, $W(t) = (W_1(t), \ldots, W_N(t))$ — N-мерный стандартный винеровский процесс; $\nu^\circ(t, A)$ — центрированная пуассоновская мера с параметром $t\Pi(A)$; процесс $W(t)$ и мера $\nu^\circ(t, A)$ независимы между собой и Q_t-измеримы при $t \geqslant 0$. Мера $\Pi(A)$ определена на борелевских множествах евклидова пространства R^n; H_0 — множество детерминированных функций $\varphi(s)$, $(-h \leqslant s \leqslant 0)$ со значениями в R^n, имеющих пределы слева, а также непрерывных справа при $s < 0$.

Норма в H_0 определена равенством

$$\|\varphi\| = \sup_{-h\leqslant s\leqslant 0} |\varphi(s)|$$

Встречающиеся далее в работе функционалы, заданные на $[0, T] \times H_0$, считаются измеримыми относительно σ-алгебры борелевских множеств пространства $[0, T] \times H_0$.

Обозначим через θ_t, $0 \leqslant t \leqslant T$ семейство операторов, ставящих в соответствие произвольной функции $\xi(s_1)$, $-h \leqslant s_1 \leqslant T$ функцию $\theta_t\xi = \xi(t + s)$. Здесь s при каждом фиксированном t пробегает значение $-h \leqslant s \leqslant 0$.

Цель работы — построение приближенного оптимального управления и получение оценок погрешности для стохастической системы вида

($\xi(t) \in R^n$ — фазовый вектор, $u \in R^l$ — управление)

$$d\xi(t) = (\varepsilon f(t, \theta_t\xi) + B(t) u)\, dt + d\eta(t) \tag{1.1}$$

$$d\eta(t) = \sum_{r=1}^{N} b_r(t)\, dW_r(t) + \int_{R^n} c(z,t)\, \nu^\circ(dt, dz), \quad \theta_0\xi = \varphi_0 \in H_0$$

Начальное условие φ_0 и число T заданы, $\varepsilon \geqslant 0$ — малый параметр. Функционал $f(t, \varphi)$ измерим и определен на $[0, T] \times H_0$. Предполагается, что существует неотрицательная и неубывающая по τ функция $r(t, \tau)$. при которой справедливы неравенства

$$|f(t,\varphi)|^2 \leqslant a_0 + \int_0^h |\varphi(-\tau)|^2\, d_\tau r(t,\tau) \tag{1.2}$$

$$|f(t,\varphi) - f(t,\psi)|^2 \leqslant \int_0^h |\varphi(-\tau) - \psi(-\tau)|^2\, d_\tau r(t,\tau)$$

$$\sup_{0\leqslant t\leqslant T} \int_0^h d_\tau r(t,\tau) < \infty$$

Матрица $B(t)$ размерности $n \times l$ и n-мерные векторы $b_r(t)$ $(r = 1, \ldots, N)$, $C(z, t)$ $(z \in R^n)$ измеримы и ограничены на $[0, T]$.

Отметим, что система

$$d\xi_1(t) = (A(t)\,\xi_1(t) + \varepsilon f(t, \theta_t\xi_1) + B(t)\,u)\, dt + d\eta(t)$$

легко сводится к системе вида (1.1) с помощью замены переменных $\xi_1(t) = Z(t)\,\xi(t)$. Здесь $Z(t)$ — решение матричного дифференциального уравнения $Z^\cdot = AZ$ с начальным условием $Z(0) = I$, где I — единичная матрица.

Пусть D — класс функционалов $V(t, \varphi)$ на $[0, T] \times H_0$, таких, что для любой фиксированной при $-h \leqslant \tau < 0$ функции $\varphi(\tau)$ и произвольного вектора $x = \varphi(0)$ функция $V_\varphi(t, x) = V(t, \varphi)$ дважды непрерывно дифференцируема по x и имеет непрерывную производную по t при почти всех t из $[0, T]$.

Свяжем с системой (1.1) интегро-дифференциальный оператор L_u, определенный на D и имеющий вид

$$L_u V(t,\varphi) = L_0 V_\varphi(t,x) + (\varepsilon f(t,\varphi) + B(t)\,u)' \nabla V_\varphi(t,x)$$

$$L_0 V_\varphi(t,x) = \frac{\partial V_\varphi(t,x)}{\partial t} + \frac{1}{2}\sum_{r=1}^{N} b_r'(t)\, \nabla^2 V_\varphi(t,x)\, b_r(t) +$$

$$+ \int_{R^n} [V_\varphi(t, x + C(z,t)) - V_\varphi(t,x) - C'(z,t)\, \nabla V_\varphi(t,x)]\, \Pi(dz)$$

Здесь штрих — знак транспонирования и $\partial V_\varphi/\partial t$ — частная производная по t, а ∇V_φ и $\nabla^2 V_\varphi$ — соответственно вектор первых и матрица вторых частных производных по $x = \varphi(0)$ функции $V_\varphi(t, x) = V(t, \varphi)$ при фиксированной на $-h \leqslant \tau < 0$ функции $\varphi(\tau)$.

Произвольное управление u назовем допустимым, если при этом управлении система (1.1) имеет решение (не обязательно единственное) и конечен функционал $G(0, u)$, где

$$G(t, u) = M_\varphi \left\{ H(\xi^u(T)) + \int_t^T F(s, \theta_s\xi, u(s))\, ds \right\}$$

Здесь M_φ — математическое ожидание, вычисляемое при условии, что траектория процесса $\xi^u(s)$ на $[t-h, t]$ фиксирована и совпадает с заданной функцией $\varphi \in H_0$. Функции $H(x)$ и $F(t, \varphi, u)$ заданы и неотрицательны.

Пусть U — класс допустимых управлений. Задача оптимального управления состоит в выборе из U такого управления u, при котором функционал $G(0, u)$ минимален. Оптимальное управление зависит, вообще говоря, от времени t и от траектории $\theta_t\xi^u$ управляемого процесса до момента времени t, т. е. имеет вид измеримого на $[0, T] \times H_0$ функционала $u(t, \varphi)$, такого, что $u(t) = u(t, \theta_t\xi) \in U$.

Справедлива следующая теорема [7].

Теорема. Пусть существуют функционал $V(t, \varphi) \in D$ и управление $v = v(t, \varphi) \in U$, удовлетворяющие при почти всех $t \in [0, T]$, всех $\varphi \in \in H_0$ и всех $u \in U$ условиям

$$L_uV(t, \varphi) + F(t, \varphi, u) \geqslant 0, \quad L_vV(t, \varphi) + F(t, \varphi, u(t, \varphi)) = = 0, \quad V(T, \varphi) = H(\varphi(0)) \tag{1.3}$$

Тогда управление v является оптимальным в смысле критерия качества $G(t, u)$, причем для любых $t \geqslant 0$ и $\varphi \in H_0$ справедливо соотношение

$$G(t, v) = \inf_{u \in V} G(t, u) = V(t, \varphi)$$

В дальнейшем предполагается, что

$$F(t, \varphi, u) = \varphi'(0) F(t) \varphi(0) + u'N(t)u, \quad H(x) = x'Hx$$

где $N(t)$ и $F(t)$ измеримы и ограничены, $N(t)$ — равномерно положительно-определенная, а $F(t)$ и H — неотрицательно-определенные матрицы.

Пусть $V(t, \varphi) = V_\varphi(t, x)$ — минимальное значение функционала качества при начальном условии $\theta_t\xi = \varphi$. При этом условия (1.3) объединяются в одно соотношение, являющееся аналогом уравнения Беллмана рассматриваемой задачи

$$\inf_{u \in V} [L_0V_\varphi(t, x) + (\varepsilon f(t, \varphi) + B(t)u)'\nabla V_\varphi(t, x) + x'F(t)x + + u'N(t)u] = 0, \quad x = \varphi(0)$$

Отсюда следует, что $V_\varphi(t, x)$ определяется уравнением

$$\begin{gathered} L_0V_\varphi(t, x) + \varepsilon f'(t, \varphi)\nabla V_\varphi(t, x) + x'F(t)x = \\ = {}^1/_4 \nabla V_\varphi'(t, x) B_1(t) \nabla V_\varphi(t, x) \\ V_\varphi(T, x) = x'Hx, \quad x = \varphi(0), \quad B_1 = BN^{-1}B' \end{gathered} \tag{1.4}$$

Оптимальное управление $v(t, \varphi)$ равняется

$$v(t, \varphi) = -{}^1/_2 N^{-1}(t) B'(t) \nabla V(t, \varphi)$$

2. Известно [11], что при $\varepsilon = 0$ уравнение (1.4) имеет точное решение вида $V_0(t, \varphi) = \varphi'(0) P(t) \varphi(0) + P_1(t)$. Здесь матрицы P, P_1 ограничены, неотрицательно определены и зависят лишь от параметров системы (1.1) и функционала качества. При $\varepsilon = 0$ оптимальное управление имеет вид

$$u_0(t, \varphi) = u_0(t, \varphi(0)) = -N^{-1}(t) B'(t) P(t) \varphi(0)$$

Покажем, что это управление является нулевым приближением к оптимальному, т. е. дает погрешность по функционалу качества порядка ε.

Введем следующие обозначения: ξ_ε^u — решение системы (1.1) при $\varepsilon > 0$ и управлении u; ξ_0^u — решение системы (1.1) при $\varepsilon = 0$ и управлении u; u и v — соответственно управление и оптимальное управление системы (1.1) при $\varepsilon > 0$. Положим

$$I_\varepsilon(u) = M_{\varphi_0}\left[\xi'(T) H \xi(T) + \int_0^T (\xi'(s) F(s) \xi(s) + u'(s) N(s) u(s))\, ds\right]$$

$$V_\varepsilon(\varphi_0) = I_\varepsilon(v), \quad V_0(\varphi_0) = I_0(u_0)$$

Выбирая в качестве управления Q_t измеримые случайные процессы $a(t) = u_0(t, \xi_0^{u_0}(t))$ и $b(t) = v(t, \theta_t \xi_\varepsilon^v)$, заключаем, что

$$V_\varepsilon(\varphi_0) = \inf_{u \in U} I_\varepsilon(u) \leqslant I_\varepsilon(a) = V_0(\varphi_0) + [I_\varepsilon(a) - I_0(a)]$$
$$V_0(\varphi_0) = \inf_{u \in U} I_0(u) \leqslant I_0(b) = V_\varepsilon(\varphi_0) + [I_0(b) - I_\varepsilon(b)]$$

Следовательно

$$|V_0(\varphi_0) - V_\varepsilon(\varphi_0)| \leqslant \max\,[|I_0(a) - I_\varepsilon(a)|, |I_0(b) - I_\varepsilon(b)|] \tag{2.1}$$

Далее, используя рассуждения [6], стр. 141—145, можно показать, что для любого Q_t-измеримого случайного процесса $\gamma(t)$, при котором

$$\int_0^T M_{\varphi_0} |\gamma(t)|^2 dt \leqslant C(1 + \|\varphi_0\|^2) \equiv C^\circ \tag{2.2}$$

имеет место неравенство

$$M_{\varphi_0}\{\sup_{0 \leqslant t \leqslant T} |\xi_\varepsilon^\gamma(t)|^2\} \leqslant C^\circ \tag{2.3}$$

Здесь и всюду в дальнейшем C — некоторые различные положительные постоянные, зависящие от параметров задачи управления и не зависящие от начального условия системы (1.1).

Аналогично можно доказать и неравенство

$$M_{\varphi_0}\{\sup_{0 \leqslant t \leqslant T} |\xi_\varepsilon^\gamma(t) - \xi_0^\gamma(t)|^2\} \leqslant \varepsilon^2 C^\circ \tag{2.4}$$

Из (2.3) и (2.4) следует, что

$$(2.5)\qquad |I_\varepsilon(\gamma)-I_0(\gamma)| = \Big|M_{\varphi_0}\Big[(\xi_\varepsilon^\gamma(T)-\xi_0^\gamma(T))'H(\xi_\varepsilon^\gamma(T)+\xi_0^\gamma(T))+$$
$$+\int_0^T(\xi_\varepsilon^\gamma(s)-\xi_0^\gamma(s))'F(s)(\xi_\varepsilon^\gamma(s)+\xi_0^\gamma(s))\,ds\Big|\leqslant$$
$$\leqslant C\,[M_{\varphi_0}\{\sup\nolimits_{0\leqslant t\leqslant T}|\xi_\varepsilon^\gamma(t)+\xi_0^\gamma(t)|\}\times$$
$$\times M_{\varphi_0}\{\sup\nolimits_{0\leqslant t\leqslant T}|\xi_\varepsilon^\gamma(t)-\xi_0^\gamma(t)|\}]^{1/2}\leqslant \varepsilon C^\circ$$

Покажем, что управления $a(t)$ и $b(t)$ удовлетворяют неравенству вида (2.2). Пусть

$$q=\inf\nolimits_{t\in[0,T],\,|u|=1}u'N(t)u$$

Так как $N(t)$ равномерно по t положительно определена, то $q>0$. Значит

$$\int_0^T M_{\varphi_0}|a(t)|^2dt=\int_0^T M_{\varphi_0}|u_0(t,\xi_0^{u_0}(t))|^2dt\leqslant\frac{1}{q}\int_0^T M_{\varphi^\circ}\times$$
$$\times u_0'(t,\xi_0^{u_0}(t))N(t)u_0(t,\xi_0^{u_0}(t))\,dt\leqslant\frac{1}{q}I_0(u_0)\leqslant$$
$$\leqslant\frac{1}{q}I_0(0)\leqslant CM_{\varphi_0}\{\sup\nolimits_{0\leqslant t\leqslant T}|\xi_0^\circ(t)|^2\}\leqslant C^\circ$$

Аналогично для $b(t)$. Следовательно, для управлений $a(t)$ и $b(t)$ справедливы оценки вида (2.3) — (2.5).

Отсюда и из (2.1) следует, что

$$|V_0(\varphi_0)-V_\varepsilon(\varphi_0)|\leqslant\varepsilon C^\circ$$

Оценки вида (2.3) — (2.5) можно получить аналогично и для управления u_0, используя его линейность по $\xi(t)$. Следовательно

$$|J_\varepsilon(u_0)-J_0(u_0)|\leqslant\varepsilon C^\circ$$

Таким образом

$$(2.6)\qquad 0\leqslant J_\varepsilon(u_0)-J_\varepsilon(v)\leqslant|J_\varepsilon(u_0)-J_0(u_0)|+|V_0(\varphi_0)-$$
$$-V_\varepsilon(\varphi_0)|\leqslant\varepsilon C^\circ$$

что и требовалось доказать.

Отметим, что в приведенном доказательстве существенно использовалась некоторая вспомогательная управляемая система, для которой управление u_0 является оптимальным, а функционал $V_0(t,\varphi)=V_\varphi^\circ(t,\varphi(0))$ — функцией Беллмана.

Этот способ доказательства используется и в дальнейшем. Именно, на каждом шаге строится вспомогательная управляемая система, для которой очередное приближение u_k к оптимальному управлению само будет оптимальным, а некоторый функционал $Q_k(t,\varphi)=Q_\varphi^k(t,\varphi(0))$ — функцией Беллмана. В качестве вспомогательной управляемой системы при $k\geqslant1$ выбирается уравнение (1.1), а функционал Q_k отличается от функции Беллмана исходной задачи на величину порядка ε^{k+1}. При этом отпадает необходимость в оценках вида (2.4). Кроме того, при обосновании оценок погрешности не делается никаких предположений об уравнении Беллмана исходной задачи.

Для иллюстрации сказанного приведем еще одно доказательство оценки (2.6), отличное от предыдущего. Именно это доказательство и будет впоследствии обобщено на последовательные приближения более высокого порядка.

Прежде всего заметим, что уравнение

$$(2.7)\qquad \begin{gathered} L_0 W_\varphi (t, x) + \varepsilon f' (t, \varphi) (\nabla W_\varphi (t, x) - 2P (t) x) + \\ + x'F (t) x = {}^1/_4 \nabla W_\varphi{}' (t, x) B_1 (t) \nabla W_\varphi (t, x) \\ W_\varphi (T, x) = x'Hx, \quad x = \varphi (0) \end{gathered}$$

определяет функцию Беллмана задачи оптимального управления с уравнением движения (1.1) и функционалом качества I_ε, равным

$$(2.8)\qquad I_\varepsilon (u) = J_\varepsilon (u) - 2\varepsilon M_{\varphi_0} \int_0^T f' (s, \theta_s, \xi_\varepsilon^u) P (s) \xi_\varepsilon^u (s) ds$$

Из соотношений (1.4) при $\varepsilon = 0$ следует, что функционал $V_0(t, \varphi) = \varphi' (0)P(t)\varphi(0) + P_1 (t)$ будет решением уравнения] (2.7) при любом $\varepsilon > 0$. Аналогично [9] можно показать, что при достаточно малых ε решение уравнения (2.7) единственно. Таким образом, V_0 есть функция] Беллмана, а u_0 — оптимальное управление и для задачи (1.1), (2.8).

Пусть теперь $W_\varepsilon (\varphi_0) = I_\varepsilon (u_0)$, $C (t) = u_0 (t, \xi_\varepsilon^{u_0} (t))$. Как и прежде, имеем

$$\begin{gathered} 0 \leqslant J_\varepsilon (u_0) - J_\varepsilon (v) \leqslant | J_\varepsilon (u_0) - I_\varepsilon (u_0) | + \\ + | W_\varepsilon (\varphi_0) - V_\varepsilon (\varphi_0) | \leqslant | J_\varepsilon (u_0) - I_\varepsilon (u_0) | + \\ + \max [| J_\varepsilon (C) - I_\varepsilon (C) |, | J_\varepsilon (b) - I_\varepsilon (b) |] \end{gathered}$$

Положим

$$\alpha (u) = M_{\varphi_0} \{\sup_{0 \leqslant t \leqslant T} | \xi_\varepsilon^u (t) |^2\}$$

Из (2.8) и (1.2) для любого $u \in U$ следует, что

$$\begin{gathered} | I_\varepsilon (u) - J_\varepsilon (u) | \leqslant \varepsilon C [\alpha (u) (1 + \| \varphi_0 \|^2 + \\ + \alpha (u))]^{1/2} \leqslant \varepsilon C (1 + \| \varphi_0 \|^2 + \alpha (u)) \end{gathered}$$

Отсюда и из того, что для управлений u_0, C и b имеют место оценки (2.3), вытекает (2.6).

3. Перейдем теперь к вопросу о высших ($k \geqslant 1$) приближениях к оптимальному управлению. Алгоритм построения этих приближений состоит в следующем. Представим функционал V в виде ряда

$$V (t, \varphi) = V_0 (t, \varphi) + \varepsilon V_1 (t, \varphi) + \varepsilon^2 V_2 (t, \varphi) + \ldots$$

где $V_i \in D$.

Подставим это разложение в уравнение (1.4) и приравняем нулю коэффициенты при одинаковых степенях ε. Учитывая уравнение (1.4) для $V_0 (t, \varphi)$ при $\varepsilon = 0$, получим, что функционалы $V_i (t, \varphi) = V_\varphi^i (t, x)$

($i = 1, 2, \ldots$) определяются рекуррентными уравнениями

$$L_0 V_\varphi^i(t, x) + f'(t, \varphi) \nabla V_\varphi^{i-1}(t, x) = \frac{1}{4} \sum_{j=0}^{i} \nabla V_\varphi^{j\prime}(t, x) B_1(t) \nabla V_\varphi^{i-j}(t, x) \tag{3.1}$$

$$V_\varphi^i(T, x) = 0, \quad x = \varphi(0)$$

Определив отсюда $V_i(t, \varphi)$ ($i = 1, \ldots, k$), зададим k-е приближение к оптимальному управлению в виде

$$u_k(t, \varphi) = -{}^1/_2 N^{-1}(t) B'(t) [\nabla V_0(t, \varphi) + \ldots + \varepsilon^k \nabla V_k(t, \varphi)]$$

Эффективность приведенного алгоритма зависит от умения вычислять функционалы $V_i(t, \varphi)$. Для $V_0(t, \varphi)$ есть явная формула. При $i \geqslant 1$ в силу (3.1) имеем

$$LV_i(t, \varphi) + S_i(t, \varphi) = 0, \quad V_i(T, \varphi) = 0 \tag{3.2}$$

$$S_i(t, \varphi) = f'(t, \varphi) \nabla V_{i-1}(t, \varphi) - \frac{1}{4} \sum_{j=1}^{i-1} \nabla V_j'(t, \varphi) B_1(t) \nabla V_{i-j}(t, \varphi)$$

В формуле (3.2) через L обозначен производящий оператор стохастического дифференциального уравнения без запаздывания

$$d\xi(s) = -B_1(s) P(s) \xi(s) ds + d\eta(s), \quad s \in [t, T], \quad \theta_t \xi = \varphi \tag{3.3}$$

При выводе формулы (3.2) было использовано следующее тождество:

$$\sum_{j=0}^{i} \nabla V_\varphi^{j\prime}(t, x) B_1(t) \nabla V_\varphi^{i-j}(t, x) =$$

$$= 4B_1(t) P(t) x + \sum_{j=1}^{i-1} \nabla V_\varphi^{j\prime}(t, x) B_1(t) \nabla V_\varphi^{i-j}(t, x)$$

Лемма 1. [7]. Пусть $V(t, \varphi) \in D$ и L_1 — производящий оператор системы

$$d\xi(s) = a(s, \theta_s \xi) ds + d\eta(s), \quad \theta_t \xi = \varphi \quad (t \leqslant s \leqslant T) \tag{3.4}$$

Тогда для любых t_1, t_2

$$M_\varphi V(t_2, \theta_{t_2}\xi) - M_\varphi V(t_1, \theta_{t_1}\xi) = \int_{t_1}^{t_2} M_\varphi L_1 V(s, \theta_s \xi) ds \quad (t \leqslant t_1 \leqslant t_2 \leqslant T)$$

Из леммы 1 вытекает

Лемма 2. Пусть $V(t, \varphi) \in D$ и для любого $t \in [0, T]$

$$L_1 V(t, \varphi) + r(t, \varphi) = 0, \quad V(T, \varphi) = 0$$

где L_1 — производящий оператор системы (3.4). Тогда функционал $V(t, \varphi)$ представим в виде

$$V(t, \varphi) = M_\varphi \int_t^T r(s, \theta_s \xi) ds$$

где $\xi(s)$ — решение системы (3.4).

Запишем уравнение (3.2) при $i = 1$ в виде

$$LV_\varphi^1(t, x) + 2f'(t, \varphi) P(t) x = 0 \tag{3.5}$$

$$V_\varphi^1(T, x) = 0, \quad x = \varphi(0)$$

Если V_φ^1 и W_φ^1 — два решения уравнения (3.5), то для $R_\varphi^1 = V_\varphi^1 - W_\varphi^1$ из $LR_\varphi^1(t, x) = 0$ и $R_\varphi^1(T, x) = 0$ следует [4], что $R_\varphi^1(t, x) \equiv 0$, т. е. решение уравнения (3.5) единственно. Аналогично, методом математической индукции можно доказать единственность решения уравнения (3.2) для всех $i \geqslant 0$.

На основании леммы 2 отсюда и из соотношений (3.2), (3.3) вытекает представление

$$V_i(t, \varphi) = M_\varphi \int_t^T S_i(\tau, \theta_\tau \xi)\, d\tau \tag{3.6}$$

Здесь $\xi(\tau)$ — решение уравнения (3.3), причем при $\tau \leqslant t$ процесс $\xi(\tau)$ определен равенством $\xi(\tau) = \varphi(\tau)$, где $\varphi(\tau)$ — заданная детерминированная функция.

В некоторых случаях вычисление правой части (3.6) сводится к квадратуре. Пусть, например, $f(t, \theta_t\xi) = f(t, \xi(t-h))$, где $h \geqslant 0$ — заданная постоянная и $p(t, x, s, y)$ — плотность вероятности перехода процесса, заданного уравнением (3.3). Тогда при $i = 1$ представление (3.6) можно записать в виде $(0 \leqslant t + h \leqslant T)$

$$V_1(t, \varphi) = 2 \int_t^{t+h} \int_{R^n} f'(s, \varphi(s-h)) P(s) y p(t, \varphi(t), s, y)\, dy\, ds +$$

$$+ 2 \int_{t+h}^T \int_{R^n} \int_{R^n} f'(s, z) P(s) y p(t, \varphi(t), s-h, z) p(s-h, z, s, y)\, dz\, dy\, ds$$

Отметим, что фигурирующая в последней формуле плотность $p(t, x, s, y)$ для некоторых систем вида (3.3) также вычисляется в явном аналитическом виде.

Оценки погрешности по функционалу, допускаемые при управлении u_k для $k \geqslant 1$, устанавливаются аналогично. Поэтому подробное доказательство оценки погрешности приведем лишь для важного с практической точки зрения управления первого приближения u_1.

Основная идея доказательства, как уже отмечалось, состоит в построении вспомогательной задачи управления, для которой управление u_1 будет оптимальным, а функционал Q_1, равный $Q_1 = V_0 + \varepsilon V_1$, — функцией Беллмана. Построим вспомогательную задачу управления. Сложим уравнение (1.4) при $\varepsilon = 0$ и умноженное на ε уравнение (3.1) при $i = 1$. Тогда для функционала $Q_1(t, \varphi) = Q_\varphi^1(t, x)$ получим соотношения

$$L_0 Q_\varphi^1(t, x) + x'F(t)x + \varepsilon f'(t, \varphi) Q_\varphi^1(t, x) + \frac{\varepsilon^2}{4} \nabla V_\varphi^{1\prime}(t, x) B_1(t) \nabla V_\varphi^1(t, x) -$$

$$- \varepsilon^2 f'(t, \varphi) \nabla V_\varphi^1(t, x) =$$

$$= \frac{1}{4} \nabla Q_\varphi^{1\prime}(t, x) B_1(t) \nabla Q_\varphi^1(t, x), \qquad Q_\varphi^1(T, x) = x'Hx, \qquad x = \varphi(0)$$

Следовательно, управление u_1 является оптимальным, а $Q_1(t, \varphi)$ является функцией Беллмана задачи управления с уравнением движения (1.1) и функционалом качества

$$I_\varepsilon^1(u) = J_\varepsilon(u) + \varepsilon^2 M_{\varphi_0} \int_0^T \delta_1(s, \theta_s \xi_\varepsilon^u)\, ds$$

$$\delta_1 = {}^1/_4 \nabla V_1' B_1 \nabla V_1 - f' \nabla V_1$$

Предположим, что функционал $V_1(t, \varphi)$ удовлетворяет неравенству

$$|\nabla V_1(t, \varphi)|^2 \leqslant C(1 + \|\varphi\|^2) \tag{3.7}$$

Положим

$$\alpha(u) = M_{\varphi_0}\{\sup_{0 \leqslant t \leqslant T} |\xi_\varepsilon^u(t)|^2\}, \qquad u \in U$$

Тогда, как и выше, нетрудно установить неравенство

$$|I^1_\varepsilon(u) - J_\varepsilon(u)| \leqslant \varepsilon^2 C(1 + \|\varphi_0\|^2 + \alpha(u))$$

Кроме того, при $V_\varepsilon^1(\varphi_0) = I_\varepsilon^1(u_1)$ и $c_1(t) = u_1(t, \theta_t \xi_\varepsilon^{u_1})$ будет

$$|V_\varepsilon(\varphi_0) - V_\varepsilon^1(\varphi_0)| \leqslant \max[|I_\varepsilon^1(c_1) - J_\varepsilon(c_1)|, |J_\varepsilon^1(b) - J_\varepsilon(b)|]$$

Для управлений $b(t)$ и $c_1(t)$ выполняется оценка вида (2.3). Следовательно

$$|V_\varepsilon(\varphi_0) - V_\varepsilon^1(\varphi_0)| \leqslant \varepsilon^2 C^\circ$$

Используя (3.7), можно показать, что оценка вида (2.3) справедлива и для управления u_1. Таким образом

$$0 \leqslant J_\varepsilon(u_1) - J(v) \leqslant |J_\varepsilon(u_1) - I_\varepsilon^1(u_1)| + |V_\varepsilon^1(\varphi_0) - V_\varepsilon(\varphi_0)| \leqslant \varepsilon^2 C^\circ$$

где постоянная C может быть оценена через параметры исходной задачи.

Итак, показано, что управление u_1 для исходной задачи управления дает погрешность по функционалу порядка ε^2.

Для завершения доказательства остается показать, что функционал V_1, действительно, удовлетворяет условию (3.7). Из (3.6) при $i = 1$ имеем

$$V_1(t, \varphi) = 2M_\varphi \int_t^T f'(s, \theta_s \xi) P(s) \xi(s)\, ds$$

где $\xi(s)$ — решение уравнения (3.3).

Пусть теперь $\xi(s)$ — решение уравнения (3.3) при $\theta_t \xi = \varphi$, а $\xi_1(s)$ — при $\theta_t \xi_1 = \varphi_1$. Тогда

$$|V_1(t, \varphi) - V_1(t, \varphi_1)|^2 = |M(V_1(t, \varphi) - V_1(t, \varphi_1))|^2 =$$

$$= 4\left|M\left(M_\varphi \int_t^T f'(s, \theta_s \xi) P(s) \xi(s)\, ds - M_{\varphi_1} \int_t^T f'(s, \theta_s \xi_1) P(s) \xi_1(s)\, ds\right)\right|^2 =$$

$$= 4\left|M \int_t^T (f'(s, \theta_s \xi) P(s) \xi(s) - f'(s, \theta_s \xi_1) P(s) \xi_1(s))\, ds\right|^2 \leqslant$$

$$\leqslant 8T\left[\int_t^T M((f'(s, \theta_s \xi) - f'(s, \theta_s \xi_1)) P(s) \xi(s))^2 ds + \right.$$

$$\left. + \int_t^T M(f'(s, \theta_s \xi_1) P(s)(\xi(s) - \xi_1(s)))^2 ds\right] \leqslant$$

$$\leqslant C\left[\int_t^T M|\xi(s)|^2 M|f(s,\theta_1\xi)-f(s,\theta_s\xi_1)|^2 ds+\right.$$

$$\left.+\int_t^T M|f(s,\theta_s\xi_1)|^2 M|\xi(s)-\xi_1(s)|^2 ds\right]\leqslant$$

$$\leqslant C[(1+\|\varphi\|^2)\|\varphi-\varphi_1\|^2+(1+\|\varphi_1\|^2)\|\varphi-\varphi_1\|^2]\leqslant$$

$$\leqslant C(1+\|\varphi\|^2+\|\varphi_1\|^2)\|\varphi-\varphi_1\|^2$$

Пусть $\varphi_1(s)=\varphi(s)$ при $-h\leqslant s<0$, $\varphi(0)=x$, $\varphi_1(0)=x+\Delta x$. Тогда $\|\varphi-\varphi_1\|^2=|\Delta x|^2$ и

$$|\nabla V_1(t,\varphi)|^2=|\nabla V_\varphi^1(t,x)|^2=\lim_{\Delta x\to 0}\frac{|V_\varphi^1(t,x)-V_\varphi^1(t,x+\Delta x)|^2}{|\Delta x|^2}=$$

$$=\lim_{\Delta x\to 0}\frac{|V_1(t,\varphi)-V_1(t,\varphi_1)|^2}{|\Delta x|^2}\leqslant\lim_{\Delta x\to 0}C(1+\|\varphi\|^2+\|\varphi_1\|^2)\leqslant$$

$$\leqslant C(1+\|\varphi\|^2)$$

откуда и следует (3.7).

4. В случае нулевого и первого приближений были построены вспомогательные задачи управления с уравнением движения (1.1) и функционалами качества, отличающимися от функционала качества исходной задачи на величины порядка ε и ε^2 соответственно.

Построим вспомогательную задачу, для которой управление u_k будет оптимальным, а функционал

$$Q_k=V_0+\varepsilon V_1+\ldots+\varepsilon^k V_k \tag{4.1}$$

при всех $k\geqslant 1$ является функцией Беллмана.

Просуммируем уравнение (1.4) при $\varepsilon=0$ и уравнения (3.1), умноженные на ε^i ($i=1,\ldots,k$). В полученном равенстве добавим и вычтем выражение $1/4\ \nabla Q_k' B_1\nabla Q_k$. В результате этого при $x=\varphi(0)$ получим

$$L_0 Q_\varphi^k(t,x)+x'F(t)x+\varepsilon f'(t,\varphi)\nabla Q_\varphi^k(t,x)+$$

$$+\frac{1}{4}\left[\nabla Q_\varphi^{k\prime}(t,x)B_1(t)\nabla Q_\varphi^k(t,x)-\right.$$

$$\left.-\sum_{i=1}^k\varepsilon^i\sum_{j=0}^i\nabla V_\varphi^{j\prime}(t,x)B_1(t)\nabla V_\varphi^{i-j}(t,x)\right]-$$

$$-\varepsilon^{k+1}f'(t,\varphi)\nabla V_\varphi^k(t,x)=\frac{1}{4}\nabla Q_\varphi^{k\prime}(t,x)B_1(t)\nabla Q_\varphi^k(t,x)$$

Используя (4.1), преобразуем выражение в квадратных скобках следующим образом:

$$\sum_{m=0}^k\sum_{n=0}^k\varepsilon^{n+m}\nabla V_m' B_1\nabla V_n-\sum_{i=0}^k\varepsilon^i\sum_{j=0}^i\nabla V_j' B_1\nabla V_{i-j}=$$

$$=\sum_{j=0}^k\sum_{i=j}^{k+j}\varepsilon^i\nabla V_j' B_1\nabla V_{i-j}-\sum_{j=0}^k\sum_{i=j}^k\varepsilon^i\nabla V_j' B_1\nabla V_{i-j}=$$

$$=\sum_{j=1}^k\sum_{i=k+1}^{k+j}\varepsilon^i\nabla V_j' B_1\nabla V_{i-j}=\varepsilon^{k+1}\sum_{j=1}^k\sum_{i=0}^{j-1}\varepsilon^i\nabla V_j' B_1\nabla V_{i-j}$$

Таким образом, для функционала $Q_k(t, \varphi)$ получено уравнение

$$L_0 Q_\varphi^k(t, x) + \varepsilon f'(t, \varphi) \nabla Q_\varphi^k(t, x) + x' F(t) x +$$
$$+ \varepsilon^{k+1} \delta_k(t, \varphi) = {}^1/_4 \nabla Q_\varphi^{k\prime}(t, x) B_1 \nabla Q_\varphi^k(t, x)$$
$$\delta_k = \frac{1}{4} \sum_{j=1}^{k} \sum_{i=0}^{j-1} \varepsilon^i \nabla V_j' B_1 \nabla V_{k+1+i-j} - f' \nabla V_k$$

Следовательно, $Q_k(t, \varphi)$ является функцией Беллмана задачи управления с уравнением движения (1.1) и минимизируемым функционалом

$$I_\varepsilon^k(u) = J_\varepsilon(u) + \varepsilon^{k+1} M_{\varphi_0} \int_0^T \delta_k(s, \theta_s \xi_\varepsilon^u) ds$$

Представление (3.6) позволяет установить некоторые достаточные условия и ограничения на $f(t, \varphi)$, при выполнении которых функционалы $V_i(t, \varphi)$ удовлетворяют оценке вида (3.7). После этого аналогично предыдущему можно доказать, что

$$0 \leqslant J_\varepsilon(u_k) - J_\varepsilon(v) \leqslant \varepsilon^{k+1} C^\circ$$

В заключение отметим, что полученные результаты нетрудно обобщить и на системы вида

$$d\xi(t) = (\varepsilon f(t, \theta_t \xi) + B(t) u) dt + d\eta(t) + d\eta_1(t) \xi(t)$$

где $\eta_1(t)$ — матричный процесс с независимыми приращениями, а остальные параметры имеют тот же смысл, что и в уравнении (1.1), а также на системы с шумами при управлении.

Поступила 27 IX 1977

ЛИТЕРАТУРА

1. *Красовский Н. Н.* Теория управления движением. М., «Наука», 1967.
2. *Альбрехт Э. Г.* Об управлении движением нелинейных систем. Дифференциальные уравнения, 1966, т. 2, № 3.
3. *Гихман И. И., Скороход А. В.* Введение в теорию случайных процессов. М., «Наука», 1965.
4. *Гихман И. И., Скороход А. В.* Стохастические дифференциальные уравнения. Киев, «Наукова думка», 1968.
5. *Гихман И. И., Скороход А. В.* Теория случайных процессов, т. 3, М., «Наука», 1975.
6. *Гихман И. И., Скороход А. В.* Управляемые случайные процессы. Киев, «Наукова думка», 1977.
7. *Колмановский В. Б., Майзенберг Т. Л.* Оптимальное управление стохастическими системами с последействием. Автоматика и телемеханика, 1973, № 1.
8. *Шайхет Л. Е.* Решение задачи стабилизации управляемых стохастических систем с запаздыванием методом функционалов Ляпунова. В сб.: Теория случайных процессов, вып. 4. Киев, «Наукова думка», 1976.
9. *Колмановский В. Б.* О приближенном синтезе некоторых стохастических квазилинейных систем. Автоматика и телемеханика, 1975, № 1.
10. *Колмановский В. Б.* Об оптимальном управлении некоторыми квазилинейными стохастическими системами. ПММ, 1975, т. 39, вып. 4.
11. *Ройтенберг Я. Н.* Некоторые задачи управления движением. М., Физматгиз, 1963