УДК 521.19

В. И. ЛОБАЧ

ОБ ОДНОМ ПОДХОДЕ К РЕШЕНИЮ ЗАДАЧИ УПРАВЛЕНИЯ ЧАСТИЧНО НАБЛЮДАЕМЫМ СЛУЧАЙНЫМ ПРОЦЕССОМ

1. Постановка задачи. Рассмотрим следующую задачу оптимального управления частично наблюдаемым случайным процессом $\{\theta_t, \xi_t, 0 \leqslant t \leqslant T\}$ (θ_t — ненаблюдаемая, ξ_t — наблюдаемая компонента), заданным с помощью системы стохастических дифференциальных уравнений Ито:

$$d\theta_t = (F(t)\theta_t + G(t)u_t)dt + \Phi(t)dw_{1t}, \tag{1}$$

$$d\xi_t = \Theta_t^* \otimes D(t) \otimes \Theta_t dt + \sigma(t)dw_{2t}, \tag{2}$$

θ — гауссовский вектор, $M\{\theta_0\} = \mu_0$, $\operatorname{cov}(\theta_0, \theta_0) = \gamma_0$, $\xi_0 = 0$, где u_t — управление; θ_t — ненаблюдаемый n-мерный случайный процесс; ξ_t — наблюдаемый m-мерный случайный процесс w_{1t}, w_{2t} n, m-мерные независимые винеровские процессы. Через $D(t)$ обозначена

$$D(t) = \begin{bmatrix} D_1(t) & 0 & \dots & 0 \\ 0 & D_2(t) & \dots & 0 \\ \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & & D_m(t) \end{bmatrix}.$$

$*$ — символ транспонирования. Матрицы $F(t)$, $G(t)$, $\Phi(t)$, $D_i(t)$, $i = \overline{1, m}$, $\sigma(t)$ имеют следующие размерности соответственно $(n \times n)$, $(n \times r)$, $(n \times n)$, $(n \times n)$, $(m \times m)$.

$$\Theta_t^* \otimes D(t) \otimes \Theta_t \overset{\Delta}{=} (\Theta_t^* D_1(t)\Theta_t, \ \dots, \ \Theta_t^* D_m(t)\Theta_t)^*.$$

При этих предположениях решение системы уравнений (1), (2) существует и единственно [1].

Требуется выбрать управление u_t как функционал от $\xi_0^t = \{\xi_\tau, 0 \leqslant \tau \leqslant t\}$ так, чтобы было минимально математическое ожидание следующего выражения:

$$\frac{1}{2} \int_0^T (\Theta_t^* A(t)\Theta_t + u_t^* B(t)u_t)dt, \tag{3}$$

где матрицы $A(t)$ и $B(t)$ — симметрические, причем $A(t)$ — неотрицательно определенная, $B(t)$ — положительно определенная.

Сформулированная задача стохастического оптимального управления относится к классу задач управления по неполным данным, так как предполагается, что наблюдается только часть компонент вектора состояния (θ_t, ξ_t). Особенностью задачи (1), (2), (3) является то, что уравнение (2) является нелинейным относительно ненаблюдаемых компонент θ_t, причем нелинейная зависимость квадратичная, т. е. такая, которая не имеет однозначной обратной. Подобная ситуация типична для теории широко распространенных в технике экстремальных систем [2].

Наиболее полно изучены задачи управления в случае, когда уравнения (1), (2) линейны относительно θ_t и u_t. Для таких систем справедливы «теорема разделения» [3], согласно которой задачи управления и фильтрации можно рассматривать в некотором смысле независимо. Оптимальное управление является здесь линейной функцией от наилучшей в среднеквадратическом смысле оценки μ_t процесса θ_t.

При исследовании нелинейных систем методом динамического программирования можно получить уравнение Гамильтона — Якоби — Беллмана для определения оптимального закона управления и минимального значения функционала (3). Основной трудностью на этом пути является то, что функция Беллмана удовлетворяет нелинейному функциональному уравнению и содержит бесконечное число переменных, называемых достаточными координатами. Для того чтобы получить конкретный результат, необходимо делать упрощение в процессе решения задачи.

В статье применяется метод «apparent linearization», рассмотренный в [4] при исследовании детерминированных систем, описываемый обыкновенными дифференциальными уравнениями. В результате получается уравнение Гамильтона — Якоби — Беллмана, которое решается аналитически.

2. Линеаризация задачи (1), (2), (3). Пусть θ_t — какая-либо, не обязательно оптимальная, оценка вектора θ_t по наблюдениям ξ_0^t. Наряду с системой уравнений (1), (2) рассмотрим следующую «линеаризованную»:

$$d\tilde{\Theta}_t = (F(t)\tilde{\Theta}_t + G(t)u_t)\,dt + \Phi(t)dw_{1t} \qquad (1')$$

$$d\tilde{\xi}_t = \hat{\Theta}_t^* \otimes D(t) \otimes \tilde{\Theta}_t dt + \sigma(t)\,dw_{2t}. \qquad (2')$$

Рассмотрим задачу оптимального управления системой (1′), (2′), (3′). Для решения этой задачи применим метод достаточных координат [5]. Пусть $p^u(t, \tilde{\theta}|\tilde{\xi}_0^t)$ — апостериорная плотность вероятностей значений процесса $\tilde{\theta}_t$ в момент времени t при условии, что известна реализация $\tilde{\xi}_0^t$ и фиксировано управление u_t.

Теорема 1. Апостериорная плотность вероятностей $p^u(t, \tilde{\theta}|\tilde{\xi}_0^t)$ является нормальной с математическим ожиданием μ_t^u и матрицей вторых центральных моментов $\Gamma^u(t)$, удовлетворяющими следующим уравнениям:

$$d\mu_t^u = (F(t)\mu_t^u + G(t)u_t)dt + \Gamma^u(t)\hat{\theta}_t^* \otimes D(t) \otimes Q_2^{-1}(d\tilde{\xi}_t - \hat{\theta}_t^* \otimes D(t) \otimes \mu_t dt), \qquad (4)$$

$$\dot{\Gamma}^u = F(t)\Gamma + \Gamma F^*(t) + Q_1(t) - \Gamma^u K(\hat{\theta}_t^*)\Gamma^u, \ \mu_0^u = \mu_0, \ \Gamma^u(0) = \gamma_0, \qquad (5)$$

$$K(\hat{\theta}_t^*) = \begin{bmatrix} \hat{\theta}_t^* D_1(t) Q_2 D_1^*(t)\hat{\theta}_t & & 0 \\ & \ddots & \\ 0 & & \hat{\theta}_t^* D_m Q_2 D_m^* \hat{\theta}_t \end{bmatrix}, \ Q_1 = \Phi\Phi^*, \ Q_2 = \sigma\sigma^*.$$

Доказательство. Очевидно, что $\hat{\theta}_t$ является $F_t^{\tilde{\xi}}$-измеримой функцией, где $F_t^{\tilde{\xi}} = \sigma\{\omega : \tilde{\xi}_s, s \leqslant t\}$ σ — алгебра, порожденная случайными величинами $\tilde{\xi}_s$, $s \leqslant t$. (1′), (2′) линейны относительно $\tilde{\theta}_t$ с коэффициентами, являющимися $F_t^{\tilde{\xi}}$ измеримыми функциями. В [6] показано, что случайный процесс $(\tilde{\theta}_t, \tilde{\xi}_t)$ является условно-гауссовским с параметрами, удовлетворяющими уравнениям (4), (5).

Таким образом, установлено, что достаточными координатами в задаче (1), (2), (3) являются μ_t, $\Gamma(t)$. Обозначим через

$$I(\mu, \Gamma, t) = \min_u M\left\{\frac{1}{2}\int_0^T (\tilde{\theta}_t^* A(t)\tilde{\theta}_t + u_t^* B(t)u_t)\,dt \,|\, \mu_t = \mu, \ \Gamma(t) = \Gamma\right\}.$$

Заметим, что уравнения (4), (5) образуют фильтр Калмана — Бьюси [7].

Теорема 2. Оптимальное управление $u^0(t, \mu(t), \Gamma(t))$ и соответствующее значение функционала $I(\mu, \Gamma, t)$ можно определить из решения следующего уравнения (Гамильтона — Якоби — Беллмана):

$$-\frac{\partial I}{\partial t} = \min_u \left\{\frac{\partial I}{\partial \mu}(F(t)\mu + G(t)u) + \operatorname{tr}\left[\left(\frac{\partial I}{\partial \Gamma} + \frac{1}{2}\frac{\partial^2 I}{\partial \mu^2}\right)\Gamma \cdot K(\hat{\theta}_t^*) \cdot \Gamma\right] + \right.$$

$$\left. + \frac{1}{2}(\mu^* A\mu + u^* Bu = \operatorname{tr}\Gamma A)\right\}, \ I(\mu, \Gamma, T) = 0. \qquad (6)$$

Доказательство. Вывод уравнения (6) проводится стандартным методом динамического программирования (см., например, [3]).

Заметим, что $\Gamma^u(t)$ зависит от u неявно, только через ξ_0^t. Сделаем следующее упрощение: пусть $\Gamma^u(t) \equiv \Gamma(t)$, т. е. имеется только одна достаточная координата μ_t для нахождения оптимального управления. В этом случае уравнение Гамильтона — Якоби — Беллмана примет следующий вид:

$$\frac{\partial I}{\partial t} = \min_u \left\{ \frac{\partial I}{\partial \mu}(F(t)\mu + G(t)u) + \frac{1}{2}\operatorname{tr}\left[\frac{\partial^2 I}{\partial \mu^2}\Gamma \cdot K(\hat{\theta}_t^*)\cdot \Gamma\right] + \right.$$
$$\left. + \frac{1}{2}(\mu^* A\mu + u^* Bu + \operatorname{tr}\Gamma A)\right\}, \quad I(\mu, T) = 0. \tag{7}$$

Приравнивая к нулю производную по u правой части последнего уравнения, получаем выражение для оптимального управления. Исключая в правой части (7) управление u, получаем нелинейное уравнение в частных производных параболического типа для вычисления функции $I(\mu, t)$. Решение этого уравнения можно искать в виде ряда по степеням μ с коэффициентами, зависящими от t: $I(\mu, t) = s_0(t) + s_1^*(t)\mu + \frac{1}{2}\mu^* S(t)\mu + \ldots$. Подставляя такой ряд в уравнение и группируя члены при одинаковых степенях μ, можно записать уравнения для вычисления этих неизвестных коэффициентов. При этом оказывается, что уравнения для коэффициентов при степенях μ, начиная с третьей, однородны и имеют нулевые граничные условия, т. е. эти коэффициенты равны нулю. Коэффициенты при первых трех степенях μ удовлетворяют следующим уравнениям соответственно:

$$-\dot{s}_0 = \frac{1}{2} s_1^* B^{-1}(t)G(t)s_1 + \frac{1}{2}\operatorname{tr}(S\Gamma\cdot K(\theta_t^*)\cdot\Gamma) + \frac{1}{2}\operatorname{tr}\Gamma A, \quad s_0(T) = 0,$$
$$-\dot{s}_1 = F^*(t)s_1, \quad s_1(T) = 0, \tag{8}$$
$$-\dot{S} = F^*(t)S + SF(t) + A(t) - SG(t)B^{-1}(t)G^*(t)S, \quad S(T) = 0.$$

Оптимальное управление имеет вид

$$u^0(t) = -B^{-1}(t)G^*(t)S(t)\mu_t, \tag{9}$$

где μ_t удовлетворяет уравнению (4) при $u_t = u^0(t)$; $S(t)$ — является решением матричного уравнение Рикатти (8).

3. Приближенное решение задачи (1), (2), (3). В соответствии (2′) $\hat{\theta}_t$ — произвольная оценка процесса θ_t по наблюдениям ξ_0^t. В качестве θ_t можно взять оценку μ_t, удовлетворяющую уравнению (4). Таким образом, для вычисления оценки μ_t получается система уравнений:

$$d\mu_t = (F(t)\mu_t + G(t)u_t^0)dt + \Gamma(t)\mu_t^* \otimes D(t) \otimes Q_2^{-1}(t)\,(d\xi_t - \mu_t^* \otimes D(t) \otimes \mu(t)dt),$$
$$\dot{\Gamma} = F^*(t)\Gamma + \Gamma F(t) + Q_1(t) - \Gamma(t)\cdot K(\mu_t^*)\cdot\Gamma(t). \tag{10}$$

В этом случае закон управления (9), где μ_t определяется уравнением (10), будет, вообще говоря, не оптимальным, но близким к оптимальному.

4. Численный пример. Количественный анализ закона управления (9), где μ_t определяется уравнением (10), проводился методом статистического моделирования по 6 реализациям при

$$F(t) = \begin{pmatrix} -0{,}56 & 0 \\ 0 & -0{,}56 \end{pmatrix}, \quad T = 5, \quad D(t) = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix},$$

$$A(t) = \begin{pmatrix} 3 & 0 \\ 0 & 3 \end{pmatrix}, \quad B(t) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad m = 1, \quad \sigma = \sqrt{50}, \quad \mu_0 = \gamma_0 = 1.$$

Номер реализации	Значения функционала (3)		$\frac{\Delta\Phi}{\Phi_0}$
	Φ_0	Φ_1	
1	2,982	3,001	0,006
2	2,997	3,001	0,001
3	1,901	2,052	0,080
4	1,959	2,074	0,058
5	3,026	3,031	0,002
6	3,188	3,211	0,007

Построенный закон управления сравнивался с оптимальным в задаче (1), (3), т. е. в случае, когда вектор θ_t полностью наблюдается. Обозначим через Φ_0 и Φ_1 значения функционала (3) при оптимальном управлении в задаче (1), (3) и при управлении (9) в задаче (1), (2), (3) соответственно. Положим $\Delta\Phi=\Phi_1-\Phi_0$. Результаты моделирования приведены в таблице, из которой видно, что применение управления (9) в задаче (1), (2), (3) является достаточно эффективным.

ЛИТЕРАТУРА

1. Гихман И. И., Скороход А. В. Стохастические дифференциальные уравнения.— Киев, 1968.
2. Медведев Г. А., Тарасенко В. П. Вероятностные методы исследования экстремальных систем.— М., 1967.
3. Вонэм В. М.— Математика, 1973, т. 17, № 4, с. 5.
4. Werni A., Cook C.— Automatica, 1975, v. 11, № 1, p. 75.
5. Параев Ю. И. Введение в статистическую динамику процессов управления и фильтрации.— М., 1976.
6. Липцер В. Ш., Ширяев А. Н. Статистика случайных процессов.— М., 1974.
7. Kalman R. E., Busy R. S.— Trans. ASME J. of Basic Engineering, 1961, v. 83, № 1, p. 123.
8. Брайсон А., Хо Ю-ши. Прикладная теория оптимального управления.— М., 1972.

УДК 517.968.23

А. П. ШИЛИН

ИНТЕГРАЛЬНЫЕ УРАВНЕНИЯ ТИПА СВЕРТКИ, СВОДЯЩИЕСЯ К ЗАДАЧАМ ГИЛЬБЕРТА

Всякая H-непрерывная функция $F(x)$, $-\infty<x<+\infty$, $F(\pm\infty)=0$, может быть единственным образом представлена в виде

$$F(x)=(I^+F)(x)-(I^RF)(x), \tag{1}$$

где $(I^+F)(x)=i\,\mathrm{Im}\,F(x)+\frac{1}{\pi}\int\limits_{-\infty}^{+\infty}\frac{\mathrm{Im}F(\tau)d\tau}{\tau-x}$ есть предельное значение функции, аналитической в верхней полуплоскости, причем $(I^+F)(\pm\infty)=0$, а $(I^RF)(x)=\frac{1}{\pi}\int\limits_{-\infty}^{+\infty}\frac{\mathrm{Im}F(\tau)d\tau}{\tau-x}-\mathrm{Re}F(x)$ — вещественнозначная функция, причем $(I^RF)(\pm\infty)=0$.

Рассмотрим уравнение

$$f(t)+(Lf)(t)=g(t),\quad -\infty<t<+\infty, \tag{2}$$

где $(Lf)(t)=\frac{1}{\sqrt{2\pi}}\int\limits_0^{+\infty}k_1(t-\tau)\,f(\tau)\,d\tau-\frac{1}{\sqrt{2\pi}}\int\limits_{-\infty}^{0}k_1(t+\tau)\overline{f(\tau)}\,d\tau+$ $+\frac{1}{\sqrt{2\pi}}\int\limits_{-\infty}^{0}k_2(t-\tau)f(\tau)d\tau+\frac{1}{\sqrt{2\pi}}\int\limits_{-\infty}^{0}k_2(t+\tau)\overline{f(\tau)}\,d\tau$, а все функции, в том