EESTI NSV TEADUSTE AKADEEMIA TOIMETISED. 32. KÖIDE
FÜÜSIKA * MATEMAATIKA. 1983, NR. 1

ИЗВЕСТИЯ АКАДЕМИИ НАУК ЭСТОНСКОЙ ССР. ТОМ 32
ФИЗИКА * МАТЕМАТИКА. 1983, № 1

УДК 681.513.5

Р. ТЕННО

ПРИМЕР РЕШЕНИЯ ДВУХШАГОВОЙ ЗАДАЧИ УПРАВЛЕНИЯ ЧАСТИЧНО НАБЛЮДАЕМЫМ СЛУЧАЙНЫМ ПРОЦЕССОМ

(Представил Н. Алумяэ)

Задачи управления частично наблюдаемым случайным процессом, которые не могут быть сведены к задаче с полными наблюдениями, мало изучены. Ниже рассмотрена двухшаговая задача оптимального (дуального) управления линейным безынерционным объектом, неизвестные параметры которого постоянны или изменяются случайным образом. Показана разрешимость этой задачи. Найден явный вид управления на втором шаге и выведены уравнения оптимального управления на первом шаге. Полученные результаты предлагается использовать в качестве приближенного решения многошаговой задачи управления.

1. Постановка задачи и основной результат

1. Пусть на вероятностном пространстве $(\Omega, \mathcal{F}, P)$ задан частично наблюдаемый случайный процесс $\{\beta_t, x_t, y_t\colon t=1, 2, \ldots, N\}$, определяемый уравнениями

$$\varphi(\beta_t-\mu)=\Theta\alpha_t, \quad \beta_s=\tilde{\beta}_s, \quad \alpha_s=0, \quad s\leqslant 0, \tag{1}$$

$$x_t=u_t^{\mathrm{T}}\beta_{1,t}+\beta_{0,t}, \tag{2}$$

$$y_t=x_t+h_t, \tag{3}$$

где $\beta_t=(\beta_{0,t}, \beta_{1,t}^{\mathrm{T}})^{\mathrm{T}}$; $\alpha_t, \beta_t \in E^{k+1}$, $u_t \in E^k$, $y_t, x_t \in E$, x_t — состояние, u_t — управление, y_t — наблюдаемое состояние, h_t — ошибка наблюдения, β_t — неизвестный параметр управляемого объекта; α_t, $t=1, 2, \ldots, N$, — импульсы обновления случайного процесса $\{\beta_t\}$. φ, Θ — линейные операторы авторегрессии и скользящего среднего соответственно, равные

$$\varphi=\mathrm{I}-\varphi_1 B-\varphi_2 B^2-\ldots-\varphi_m B^m, \quad \Theta=\mathrm{I}-\Theta_1 B-\Theta_2 B^2-\ldots-\Theta_m B^m.$$

Здесь I — единичная матрица, B^k — оператор сдвига назад: $B^k\xi_t=\xi_{t-k}$. φ_i, Θ_i — симметричные матрицы параметров, μ — вектор среднего процесса $\{\beta_t\}$ (существует в стационарном случае, в нестационарном $\mu\equiv 0$).

Пусть ошибки наблюдения h_t, $t=1, 2, \ldots, N$, импульсы обновления α_t, $t=1, 2, \ldots, N$, и начальные условия $\tilde{\beta}=(\tilde{\beta}_0^{\mathrm{T}}, \tilde{\beta}_{-1}^{\mathrm{T}}, \ldots, \tilde{\beta}_{1-m}^{\mathrm{T}})^{\mathrm{T}}$ рекуррентного уравнения (1) попарно и во времени независимы и распределены по нормальному закону, т. е.

$$h_t \sim IN(0, r),\ r>0; \quad \alpha_t \sim IN(0, D), \quad \tilde{\beta} \sim N(\tilde{b}, \tilde{P}).$$

Попарно независимыми принимаются также аддитивное возмущение $\beta_{0,t}$ и параметр наклона $\beta_{1,t}$ уравнения (2), $t = 1, 2, \dots, N$.
2. Ставится задача оптимального управления

$$J = M\{\sum_{t=1}^{N} W_t\} \to \inf_u, \tag{4}$$

где M — оператор усреднения по неизвестным (в момент управления) величинам, N — горизонт управления (при формулировке основного результата принимается $N = 2$), W_t — квадратичная функция потерь, равная

$$W_t = (x_t - x_t^0)^2,$$

x_t^0 — известная функция времени такая, что $(x_t^0)^2 < \infty$, $t = 1, 2, \dots, N$.
3. Допустимая стратегия управления $u = (u_1, \dots, u_t, \dots, u_N)$ задана непрерывными (по наблюдениям $y_1, \dots, y_{t-1}$) функциями

$$u_t = u_t(y_1, \dots, y_{t-1}) \in E^k.$$

4. При формулировке основного результата будем пользоваться следующими упрощенными обозначениями векторов оценок неизвестных параметров

$$a = M\{\beta_{N-1} | \mathcal{F}^y_{N-2}\}, \quad b = M\{\beta_N | \mathcal{F}^y_{N-1}\}, \quad c = M\{\beta_N | \mathcal{F}^y_{N-2}\}$$

и матриц ковариации

$$G = \operatorname{cov}\{\beta_{N-1} | \mathcal{F}^y_{N-2}\}, \quad P = \operatorname{cov}\{\beta_N | \mathcal{F}^y_{N-1}\}, \quad K = \operatorname{cov}\{\beta_N | \mathcal{F}^y_{N-2}\},$$

где $\mathcal{F}^y_t = \sigma\{y_1, \dots, y_t\}$.

Определим функции

$$\Pi = \mathrm{I} - \pi_1 B - \pi_2 B^2 - \dots = \Theta^{-1}\varphi, \tag{5}$$

$$G = D + \pi_1 \bar{P} \pi_1, \quad K = D + \pi_1 G \pi_1,$$

$$T = \pi_1 G v / \sqrt{r + v^{\mathrm{T}} G v}, \quad v = (1, u^{\mathrm{T}}_{N-1})^{\mathrm{T}},$$

$$P = K - TT^{\mathrm{T}}, \tag{6}$$

$$\sqrt{2}\, X = -\tau^{-1} T_1^{\mathrm{T}} P_1^{-1} c_1, \quad \tau = T_1^{\mathrm{T}} P_1^{-1} T_1, \quad \sqrt{2}\, L = (c_0 - x_N^0)/T_0 + \sqrt{2}\, X,$$

$$2Y^2 = \tau^{-1}(c_1^{\mathrm{T}} P_1^{-1} c_1 + 1) - 2X^2, \tag{7}$$

$$w(z) = \mathrm{e}^{-z^2}(1 + \pi^{-1/2} 2i \int_0^z \mathrm{e}^{t^2} dt), \quad z = X + iY.$$

Здесь и далее индексами 0 и 1 обозначены соответственно первая и все остальные координаты рассматриваемых векторов a, b, c, T. Элементы матриц G, P, K обозначаются аналогично.

5. Примем следующие условия:

А. Векторы $P_{1,0}$, $P^{\mathrm{T}}_{0,1}$ — нулевые, т. е. оценки b_0, b_1 некоррелированы.

Б. Матрицы G_1, P_1, K_1 положительно определены.

Замечаем, что нам не требуется положительной определенности ни матрицы D_1, ни тем более матрицы D. При требовании положительной определенности матрицы $\bar{P}_1$ условие Б выполняется автоматически. Следовательно, принятые условия не исключают из рассмотрения случая, когда неизвестные параметры управляемого объекта постоянны. Положительная определенность матриц G_1, P_1, K_1 в последнем случае обусловлена начальной неопределенностью.

Введем обозначения: R_u, R_{uu} — первая и вторая частные производные по u_{N-1} от функции

$$R=\begin{cases} P_0+[1-c_1^{\mathrm{T}}(K_1+c_1c_1^{\mathrm{T}})^{-1}c_1][c_0-x_N^0)^2+T_0^2], & \text{если } |T_1|=0, \\ P_0+T_0^2\{1-2\sqrt{\pi}[L\,\mathrm{Im}\,w(z)+(Y-L^2/Y)\,\mathrm{Re}\,w(z)/2]\}/\tau, & \text{если } |T_1|\neq 0. \end{cases}$$

Теорема. *Пусть $N=2$ и пусть выполняются условия А и Б. Тогда задача оптимального управления* (4) *имеет решение.*

Оптимальное управление u^*_{N-1} на предпоследнем шаге удовлетворяет при $u_{N-1}=u^*_{N-1}$ двум условиям:

а) $(a_0+u^{\mathrm{T}}_{N-1}a_1-x^0_{N-1})a_1^{\mathrm{T}}+u^{\mathrm{T}}_{N-1}G_1+R_u/2=0$,

б) $a_1a_1^{\mathrm{T}}+G_1+R_{uu}/2$ — *неотрицательно определенная матрица.*

Оптимальное управление u^*_N на последнем шаге выражается в виде

$$u^*_N=-(P_1+b_1b_1^{\mathrm{T}})^{-1}b_1(b_0-x_N^0).$$

Цена управления u^*_N равна

$$V^*_N=P_0+[1-b_1^{\mathrm{T}}(P_1+b_1b_1^{\mathrm{T}})^{-1}b_1](b_0-x_N^0)^2.$$

Основные моменты доказательства теоремы даны во 2-м разделе.

Замечание 1. На предпоследнем шаге оптимальное управление u^*_{N-1} отличается от сепарированного $\bar{u}_{N-1}$, определяемого по формуле

$$\bar{u}_{N-1}=-(G_1+a_1a_1^{\mathrm{T}})^{-1}a_1(a_0-x^0_{N-1}).$$

Результаты моделирования задачи управления на ЭВМ [1] показали, что во многих случаях различие между управлениями u^*_{N-1} и $\bar{u}_{N-1}$ мало. Поэтому предлагаемое нами управление на предпоследнем шаге можно привести с помощью формулы Тейлора к виду

$$\tilde{u}_{N-1}=\bar{u}_{N-1}-(G_1+a_1a_1^{\mathrm{T}}+R_{uu}(\bar{u}_{N-1})/2)^{-1}R_u(\bar{u}_{N-1})/2$$

при условии, что матрица $G_1+a_1a_1^{\mathrm{T}}+R_{uu}(\bar{u}_{N-1})/2$ положительно определена. Причем, если $|u^*_{N-1}-\bar{u}_{N-1}|<\delta$, то $\tilde{V}_{N-1}=V^*_{N-1}+C\delta^2$. Здесь $\tilde{V}_{N-1}$ и V^*_{N-1} — цены управления $\tilde{u}_{N-1}$ и u^*_{N-1} соответственно.

Замечание 2. Пусть выполняется одно из следующих условий:
а) $\{\beta_t : t=1, 2, \ldots, N\}$ — независимый случайный процесс;
б) параметры $\beta_{1,t}$ уравнения (2) — известные константы.
Тогда риск R не зависит от управления u_{N-1}

а) $R=D_0+[1-c_1^{\mathrm{T}}(D_1+c_1c_1^{\mathrm{T}})^{-1}c_1](c_0-x_N^0)^2$,
б) $R=K_0-T_0^2+[1-\beta_1^{\mathrm{T}}(\beta_1\beta_1^{\mathrm{T}})^{-1}\beta_1][(c_0-x_N^0)^2+T_0^2]$,
где $T_0^2=\pi^2_{0,1}G_0^2/(r+G_0)$.

Следовательно, задачи управления и оценивания коэффициентов $\{\beta_t\}$ разделяются. Оптимальное управление от времени t, $t=1, 2, \ldots, N$, явно не зависит.

2. Двухшаговое оптимальное управление

1. В [2] показано, что путем расширения вектора состояния система (1)—(3) может быть приведена к виду

$$\bar{\beta}_{t+1}=A+F\bar{\beta}_t+\Sigma\dot{\alpha}_{t+1}, \tag{8}$$

$$y_{t+1}=H_{t+1}(A+F\bar{\beta}_t+\Sigma\dot{\alpha}_{t+1})+\sigma\dot{h}_{t+1}, \tag{9}$$

где $\dot{\alpha}_t\sim IN(0, \mathrm{I})$, $\dot{h}_t\sim IN(0, 1)$, $A=(\mathrm{I}-F)\Lambda$, $\sigma=\sqrt{r}$,

$$\bar{\beta}_t=\begin{bmatrix}\beta_t^1\\ \beta_t^2\\ \vdots\\ \beta_t^{m+1}\end{bmatrix},\quad \Lambda=\begin{bmatrix}\mu\\ \mu\\ \vdots\\ \mu\end{bmatrix},\quad F=\left[\begin{array}{c:c}\varphi_1 & \\ \varphi_2 & \\ \vdots & \mathrm{I}\\ \varphi_m & \\ \hdashline 0 & 0\end{array}\right],\quad \Sigma=\begin{bmatrix}\mathrm{I}\\ -\Theta_1\\ \vdots\\ -\Theta_m\end{bmatrix}D^{1/2},$$

$$H_{t+1}=[1,\ u^{\mathrm{T}}_{t+1},\ 0,\ \ldots,\ 0],$$

0 — нулевая матрица, I — единичная матрица.

С помощью условий, сформулированных в [3], убеждаемся в том, что процесс $\{\bar{\beta}_t,\ y_t : t=1, 2, \ldots, N\}$ условно гауссов, т. е. условное распределение

$$P(\bar{\beta}_1\leqslant q_1,\ \ldots,\ \bar{\beta}_t\leqslant q_t\,|\,\mathcal{F}^y_t)$$

гауссово для каждого $t=1, 2, \ldots, N$. Следовательно, вектор среднего $\bar{b}_t=M\{\bar{\beta}_t\,|\,\mathcal{F}^y_t\}$ и матрица ковариации $\bar{P}_t=\mathrm{cov}\{\bar{\beta}_t\,|\,\mathcal{F}^y_t\}$ определяются рекуррентными уравнениями

$$\bar{b}_{t+1}=A+F\bar{b}_t+\Gamma_{t+1}[y_{t+1}+H_{t+1}(A+F\bar{b}_t)] \tag{10}$$

и

$$\bar{P}_{t+1}=\bar{P}_{t+1/t}-\Gamma_{t+1}H_{t+1}\bar{P}_{t+1/t}, \tag{11}$$

где $\Gamma_{t+1}=\bar{P}_{t+1/t}H^{\mathrm{T}}_{t+1}/(r+H_{t+1}\bar{P}_{t+1/t}H^{\mathrm{T}}_{t+1})$.

Согласно условиям задачи, накопление информации запаздывает на один такт. Поэтому прогнозируемое распределение вероятностей $P(\bar{\beta}_{t+1}\leqslant q_{t+1}/\mathcal{F}^y_t\}$ тоже оказывается гауссовым с вектором средних $\bar{b}_{t+1/t}=M\{\bar{\beta}_{t+1}\,|\,\mathcal{F}^y_t\}$ и матрицей ковариации $\bar{P}_{t+1/t}=\mathrm{cov}\{\bar{\beta}_{t+1}\,|\,\mathcal{F}^y_t\}$:

$$\bar{b}_{t+1/t}=A+F\bar{b}_t, \tag{12}$$

$$\bar{P}_{t+1/t}=\Sigma\Sigma^{\mathrm{T}}+F\bar{P}_tF^{\mathrm{T}}. \tag{13}$$

Легко видеть, что из-за обратной связи $u_{t+1}(y_1, \ldots, y_t)$ случайными оказываются не только оценки средних $\bar{b}_{t+1/t}$, $\bar{b}_t$, но и матрицы ковариации $\bar{P}_{t+1/t}$, $\bar{P}_t$.

Определяем

$$\bar{T}_t=F\bar{P}_{t/t-1}H^{\mathrm{T}}_t/\sqrt{r+H_tP_{t/t-1}H^{\mathrm{T}}_t},$$

$$\omega_t=[y_t-H_t(A+F\bar{b}_{t-1})]/\sqrt{r+H_t\bar{P}_{t/t-1}H^{\mathrm{T}}_t}.$$

Обозначая $\bar{b}_{t+1/t-1}=M\{\bar{\beta}_{t+1}/\mathcal{F}^y_{t-1}\}$ $\bar{P}_{t+1/t-1}=\mathrm{cov}\{\bar{\beta}_{t+1}/\mathcal{F}^y_{t-1}\}$, убеждаемся в том, что

$$\bar{b}_{t+1/t}=\bar{b}_{t+1/t-1}+\bar{T}_t\omega_t, \quad \bar{P}_{t+1/t}=\bar{P}_{t+1/t-1}-\bar{T}_t\bar{T}^{\mathrm{T}}_t,$$

где ω_t — независимая нормированная случайная гауссова величина, т. е.

$$\omega_t \sim IN(0,1).$$

2. Используя результаты п. 1, можем выразить условный риск управления $r_t=M\{W_t|\mathcal{F}^y_{t-1}\}$ в виде

$$r_t=(H_t\bar{b}_{t/t-1}-x^0_t)^2+H_t\bar{P}_{t/t-1}H^{\mathrm{T}}_t. \tag{14}$$

В п. 4 раздела 1 через b и P мы обозначали условное среднее и ковариацию первой координаты расширенного вектора $\bar{\beta}_N=((\beta^1_N)^{\mathrm{T}}, \ldots, (\beta^m_N)^{\mathrm{T}})^{\mathrm{T}}$. Поэтому вектор b и матрица P могут быть определены путем выделения соответствующих элементов из вектора $\bar{b}_{N/N-1}$ и матрицы $\bar{P}_{N/N-1}$. Их можно определить и с помощью оператора (5). В таком случае P, b вычисляются по уравнениям (6) и

$$b=c+T\omega. \tag{15}$$

Лемма 1. *Пусть матрица $S_1=P_1+b_1b_1^{\mathrm{T}}$ положительно определена. Тогда*

$$u^*_N=-S_1^{-1}(b\xi+P_{1,0}), \tag{16}$$

где $\xi=b_0-x_N{}^0$, есть оптимальное управление (на последнем шаге) с ценой

$$V^*_N=P_0+\xi^2-[b_1\xi+P_{1,0}]^{\mathrm{T}}S_1^{-1}[b_1\xi+P_{1,0}]. \tag{17}$$

Доказательство. Из (14) следует, что

$$\begin{aligned} r_N &= (b_0+u^{\mathrm{T}}_N b_1-x^0_N)^2+P_0+2u^{\mathrm{T}}_N P_{1,0}+u^{\mathrm{T}}_N P_1 u_N= \\ &= u^{\mathrm{T}}_N S_1 u_N+2u^{\mathrm{T}}_N(b_1\xi+P_{1,0})+\xi^2+P_0. \end{aligned} \tag{18}$$

Так как матрица $S_1{}^{-1}$ существует, то из условия минимальности квадратной формы (18) получаем (16) и (17).

Рассмотрим далее более узкий случай, когда P_1 положительно определена и $P_{1,0}$, $P^{\mathrm{T}}_{0,1}=0$. Тогда выражение (17) упрощается:

$$V^*_N=P_0+(b_0-x^0_N)^2/(1+b^{\mathrm{T}}_1P_1^{-1}b_1). \tag{19}$$

3. Придерживаясь обозначений и определений п. 4 раздела 1, выразим удельный риск управления $R=M\{V^*_N|\mathcal{F}^y_{N-2}\}$ на предпоследнем шаге. Пусть матрица K_1 положительно определена. Если $|T_1|=0$, то риск R выражается просто:

$$R=P_0+[1-c^{\mathrm{T}}_1(K_1+c_1c^{\mathrm{T}}_1)^{-1}c_1][(c_0-x^0_N)^2+T^2_0].$$

Лемма 2. *Пусть $|T_1|\neq 0$ и пусть матрица P_1 положительно определена. Тогда*

$$R=P_0+T^2_0\{1-2\sqrt{\pi}\,[L\,\mathrm{Im}\,w(z)+(Y-L^2/Y)\,\mathrm{Re}\,w(z)/2]\}/\tau. \tag{20}$$

Доказательство. Согласно (15) и (19),

$$R=P_0+(2\pi)^{-1/2}\int_{-\infty}^{\infty}[(b_0-x^0_N)^2/(1+b^{\mathrm{T}}_1P_1^{-1}b_1)]e^{-\omega^2/2}\,d\omega, \tag{21}$$

где $b_i=c_i+T_i\omega$, $i=0, 1$.

Если $|\bar{T}_1| \neq 0$, то интеграл (21) преобразуется в сумму из двух интегралов: 7.4.13, 7.4.14 [4]. Если $Y > 0$, то интеграл (21) выражается аналитически. После несложных преобразований получаем (20). Доказываем, что в наших условиях неравенство $Y > 0$ всегда выполняется. Из определения (7) получаем

$$2Y^2 = \{1 + [T_1^{\tau} P_1^{-1} T_1 c_1^{\tau} P_1^{-1} c_1 - (T_1^{\tau} P_1^{-1} c_1)^2]/\tau\}/\tau.$$

По неравенству Коши—Шварца, выражение в квадратных скобках неотрицательно. В силу положительной определенности матрицы P_1^{-1} константа $\tau > 0$, $\tau = T_1^{\tau} P_1^{-1} T_1$. Отсюда $Y > 0$.

4. Согласно принципу оптимальности Беллмана,

$$V_t^* = \inf_u M\{W_t + V_{t+1}^* \,|\, \mathcal{F}_{t-1}^y\}, \tag{22}$$

где V_s — цена управления $u = (u_s, \dots, u_N)$. Поэтому оптимальное управление u^*_{N-1} определяется из условия минимума функции

$$V_{N-1}(u_{N-1}) = (a_0 + u_{N-1}^{\mathrm{T}} a_1 - x_{N-1}^0)^2 + G_0 + u_{N-1}^{\mathrm{T}} G_1 u_{N-1} + R(u_{N-1}). \tag{23}$$

Так как матрица G_1 положительно определена, то первые три слагаемых формулы (23) выражены в положительной квадратной форме. Функция $R(u_{N-1})$ непрерывна и $R(u_{N-1}) \geqslant 0$. Следовательно, $V_{N-1}(u_{N-1})$ имеет точку минимума. В этой точке

а) $\dfrac{\partial V_{N-1}}{\partial u_{N-1}}(u_{N-1}^*) = 0,$

б) $\dfrac{\partial}{\partial u_{N-1}} \dfrac{\partial V_{N-1}}{\partial u_{N-1}^{\mathrm{T}}}(u_{N-1}^*)$ — неотрицательно определенная матрица.

3. Примеры

Если случайный процесс $\{\beta_{0,t} : t = 1, 2, \dots, N\}$ аддитивных возмущений независим, то оптимальная стратегия управления заметно упрощается, так как $T_0 = 0$: отпадает необходимость «делить» информацию между двумя оцениваемыми коэффициентами $\beta_{0,t}$, $\beta_{1,t}$. Поэтому

$$R = P_0 + \sqrt{\pi}\, Y (c_0 - x_N^0)^2 \operatorname{Re} w(z).$$

Принимая, что $u_t \in E$, поясним решение задачи управления двумя примерами.

Пример 1. Пусть частично наблюдаемый управляемый случайный процесс $\{x_t, y_t : t = 1, 2\}$ описывается уравнениями

$$\begin{aligned} x_t &= u_t \beta + \alpha_{0,t}, \\ y_t &= x_t + h_t, \end{aligned} \tag{24}$$

где $\alpha_{0,t} \sim IN(0, D_0)$, $h_t \sim IN(0, r)$; β — неизвестный постоянный параметр управляемого объекта. Предполагается, что нам известны несмещенная гауссова начальная оценка $\bar{b}_1$ параметра β и ее дисперсия $\bar{P}_1 > 0$.

Пример 2. Пусть управляемый процесс $\{\beta_t, x_t, y_t : t = 1, 2\}$ описывается уравнениями

$$\begin{aligned} \beta_t &= \Phi \beta_{t-1} + \alpha_{1,t}, \quad \beta_0 = \tilde{\beta}, \\ x_t &= u_t \beta_t + \alpha_{0,t}, \\ y_t &= x_t + h_t, \end{aligned} \tag{25}$$

где $\alpha_{it} \sim IN(0, D_i)$, $i = 0, 1$, $D_1 > 0$; $h_t \sim IN(0, r)$; β_t — ненаблюдаемый, изменяющийся параметр управляемого объекта. Предполагается, что нам известны несмещенная гауссова оценка $\bar{b}_1$ начального условия $\tilde{\beta}$ и ее дисперсия $\bar{P}_1$.

Решение задачи управления

1. При помощи уравнений оптимальной фильтрации (10), (11) и прогнозирования (12), (13) определяем достаточные статистики неизвестных параметров. Убеждаемся в том, что $a_0=b_0=c_0=0$, $P_0=D_0$, $b_1=c_1+\Gamma(y_1-u_1a_1)$, $P_1=G_1-\Gamma u_1G_1$, $\Gamma=G_1u_1/(r+G_0+u_1^2G_1)$, и в том, что для первого примера

$$a_1=c_1=\bar{b}_1, \quad G_1=\bar{P}_1$$

и для второго —

$$a_1=\Phi\bar{b}_1, \quad c_1=\Phi^2\bar{b}_1, \quad G_1=D_1+\Phi^2\bar{P}_1.$$

2. Теорема утверждает, что оптимальное управление $u_1=u_1^*$ удовлетворяет условию

$$(u_1a_1-x_1^0)a_1+u_1G_1+(x_2^0)^2\Xi\{2Y^2+1-\sqrt{\pi}[X(4Y^2+1)\operatorname{Im}w(z)-$$
$$-[2Y(X^2-Y^2)-1/2Y]\operatorname{Re}w(z)]\}/2u_1=0, \tag{26}$$

где $\Xi=(r+G_0)/(r+G_0+u_1^2G_1)$, $\sqrt{2}X=-c_1/T_1$, $2Y^2=P_1/T_1^2$, $T_1=G_1u_1/\sqrt{r+G_0+u_1^2G_1}$ (для первого примера), $T_1=\Phi G_1u_1/\sqrt{r+G_0+u_1^2G_1}$ (для второго примера).

Решая уравнение (26), определяем оптимальное управление u_1^* на первом шаге.

3. Оптимальное управление u_2^* на втором шаге и цена управления V_2^* определяются следующим образом:

$$u_2^*=b_1x_2^0/(P_1+b_1^2), \quad V_2^*=D_0+P_1(x_2^0)^2/(P_1+b_1^2).$$

Автор благодарен Т. Тобиасу за обсуждение работы.

ЛИТЕРАТУРА

1. *Тенно Р., Ойт Х.* Изв. АН ЭССР, Физ. Матем., (в печати).
2. *Indjehagopian, J.-P.* Rev. franç. inform. et rech. opér., **15**, № 1, 39—49 (1981).
3. *Липцер Р. Ш., Ширяев А. Н.* Статистика случайных процессов. М., «Наука», 1974.
4. Справочник по специальным функциям (под ред. М. Абрамовича, И. Стичана). М., «Наука», 1979.

Институт кибернетики
Академии наук Эстонской ССР

Поступила в редакцию
16 апреля 1982

R. TENNO

OSALISELT VAADELDAVA JUHUSLIKU PROTSESSI KAHESAMMULISE OPTIMAALJUHTIMISÜLESANDE LAHENDAMISE NÄIDE

On vaadeldud kahesammulist mitteseparееruvat optimaaljuhtimisülesannet juhul, kui juhitava protsessi parameetrid pole eelnevalt teada, on konstantsed või ajas muutuvad juhuslikud suurused ning protsessi olek pole vahetult jälgitav. On tõestatud lahendi olemasolu, leitud optimaalsustingimused esimesel etapil ja optimaalne juhtimisseadus teisel etapil.

R. TENNO

AN EXAMPLE OF TWO-STEP OPTIMAL CONTROL OF PARTIALLY OBSERVED STOCHASTIC PROCESS

It is assumed that a partially observed stochastic process $(x, y)=\{x_t,\ y_t : t=1,\ 2\}$ is described by the equations

$$x_t=u^{\mathrm{T}}_t\beta^1_t+\beta^0_t, \qquad y_t=x_t+h_t,$$

where y is an observable process, $\beta=(\beta^0,\ \beta^{1\mathrm{T}})^{\mathrm{T}}$ is a multivariate autoregressive moving average random process, h is an idependent nonsingular Gaussian process. The problem is to choose the controls, u_1, $u_2(y_1)$, which minimize the functional

$$J(u)=M\{(x_1-x^0)^2+(x_2-x^0)^2\}.$$

The existence of the solution for the two-step control problem is proved. Sufficient conditions for optimality for the first step, and the optimal control law for the second step are obtained. Examples are given to illustrate the expressions obtained. The argumenting is the following: at first it is shown that the process $(\beta,\ x,\ y)$ is conditional Gaussian. Then, using the exact solution to the conditional Gaussian filtration problem, the Bellman optimum principle is applied.

The computational complexity of the two-step optimal control is due to the numerical integration in calculating the probability integral.