EESTI NSV TEADUSTE AKADEEMIA TOIMETISED.
FÜÜSIKA * MATEMAATIKA
ИЗВЕСТИЯ АКАДЕМИИ НАУК ЭСТОНСКОЙ ССР. ФИЗИКА * МАТЕМАТИКА
PROCEEDINGS OF THE ACADEMY OF SCIENCES OF THE ESTONIAN SSR.
PHYSICS * MATHEMATICS

1986, 35, 3

https://doi.org/10.3176/phys.math.1986.3.16

УДК 517.977.5

Р. ТЕННО

О РОЛИ СОНДИРОВАНИЯ В ЗАДАЧАХ УПРАВЛЕНИЯ ПО НЕПОЛНЫМ ДАННЫМ. ВЫБОР МОМЕНТА НАБЛЮДЕНИЯ

R. TENNO. VAATLUSMOMENDI VALIKUST VAEGINFORMATSIOONIGA OPTIMAALJUHTIMISE ÜLESANDES

R. TENNO. OPTIMAL SELECTION OF THE OBSERVATION TIME IN THE CONTROL PROBLEM WITH UNKNOWN PARAMETERS

(Представил Н. Алумяэ)

Активное накопление информации характерно для задач оптимального управления по неполным данным особенно в начальный этап, когда управления являются наиболее сондирующими (агрессивными). Ввиду сложности задач в литературе нет оценок длительности интенсивного сондирования, кроме численных примеров (см. напр. [1, 2]). Такие оценки могут быть получены в простых задачах с обозримой структурой. Они позволяют глубже понять ту роль, которую играет (или не играет) сондирование в оптимальном управлении. Рассмотрим одну из таких задач.

Пусть M — математическое ожидание, λ — любая константа, N — натуральное число. Требуется найти пару (τ, α), где τ — момент наблюдения $(1\leqslant\tau<N)$, α — стратегия управления с обратной связью вида $\alpha_t=\alpha(t, \xi_\tau)$, $\tau\leqslant t$, такую, которая минимизирует функционал

$$v^\alpha=M\left\{\sum_{t=1}^{\tau}(\Theta_t-\lambda)^2+\sum_{t=\tau+1}^{N}(\Theta_t-\lambda)^2\right\},\tag{1}$$

если управляемая последовательность задана уравнением

$$\Theta_t=\alpha^\mathrm{T}_{t-1}\beta+\varepsilon_1(t)\tag{2}$$

и наблюдается

$$\xi_\tau=\Theta_\tau+\varepsilon_2(\tau).\tag{3}$$

Здесь $\varepsilon_1(t)$, $t=1, \ldots, N$, $\varepsilon_2(\tau)$ — независимые, центрированные гауссовы величины с дисперсиями σ_1, σ_2 соответственно, $\sigma_1+\sigma_2>0$, β — вектор неизвестных параметров. Заданы гауссова оценка a вектора β и положительно определенная матрица ковариации A оценки a.

Оценка момента наблюдения. В принятых условиях прогноз неизвестных параметров (а также прогноз возмушений $\{\varepsilon(t)\}$) один и тот же для любых упреждений. Поэтому оптимальная стратегия строится следующим образом. С помощью среднего a и ковариации A априорного распределения выбирается управление на первом шаге α_0 и фиксируется на данном уровне $\alpha_t=\alpha_0$ до момента поступления наблюдений τ. С помощью наблюдения ξ_τ вычисляется новое управление $\alpha_\tau=\alpha(\tau, \xi_\tau)$, которое затем фиксируется $\alpha_t=\alpha_\tau$ до конца интервала управления. Такой выбор стратегии позволяет свести задачу (1)—(3) к двухшаговой с критерием оптимальности

$$v^\alpha = \tau S + (N - \tau) R.$$

Здесь

$$S = (\alpha_0^\tau a - \lambda)^2 + \alpha_0^\tau A \alpha_0 + \sigma_1$$

— риск действия,

$$R = M \inf_{\alpha_\tau} M\{(\Theta_{\tau+1} - \lambda)^2 / \xi_\tau\}$$

— риск изучения. Согласно [3] функцию $R(\alpha)$ можно записать в более удобном виде, с помощью решения $w(z)$ дифференциального уравнения

$$\frac{dw}{dz} = \frac{2i}{\sqrt{\pi}} - 2zw(z)$$

от комплексного аргумента $z = x + iy$. Другими словами,

$$R = \sigma_1 + \lambda^2 y \sqrt{\pi} \operatorname{Re} w(z),$$

где (в слу́чае одномерного управления)

$$\sqrt{2}\, x = -\frac{a}{T}, \quad 2y^2 = \frac{A}{T^2} - 1,$$

$$T = A\alpha_\tau / \sqrt{\sigma_1 + \sigma_2 + A\alpha_\tau^2}.$$

Оптимальное управление на втором этапе α_τ^* такое же, как и в случае $N = 2$, а на первом этапе α_0^* и оптимальный момент наблюдения τ^* удовлетворяют трем условиям:

$$S = R, \quad \tau S_\alpha + (N - \tau) R_\alpha = 0, \tag{4}$$

матрица

$$\tau S_{\alpha\alpha} + (N - \tau) R_{\alpha\alpha}$$

неотрицательно определена. Здесь S_α, R_α — первые, а $S_{\alpha\alpha}$, $R_{\alpha\alpha}$ — вторые производные от функции $S(\alpha_0)$, $R(\alpha_0)$. Если управление одномерное и если для простоты отказаться от требования целочисленности τ, то получим простую оценку

$$\tau^* = \frac{R_\alpha}{R_\alpha - S_\alpha} (\alpha_0^*) N.$$

Риск действия и риск изучения изменяют управление в противоположном направлении. Поэтому

$$\frac{S_\alpha}{R_\alpha} (\alpha_0^*) \leqslant 0.$$

В широких пределах изменения параметров $|R_\alpha(\alpha_0^*)| \ll |S_\alpha(\alpha_0^*)|$, тем самым $\tau^* \ll N$.

ЛИТЕРАТУРА

1. *MacRae, E. C.* Ann. of Economic and Social Measurement, **1**, № 4, 437—447 (1972).
2. *Tse, E., Bar-Shalom, Y.* IEEE Trans. Autom. Contr., **18**, № 2, 109—177 (1973).
3. *Тенно Р.* Изв. АН ЭССР. Физ. Матем., **32**, № 1, 11—18 (1983).

Институт кибернетики
Академии наук Эстонской ССР

Поступила в редакцию
26/XII 1985