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
1984, 33, 4

https://doi.org/10.3176/phys.math.1984.4.07

УДК 517.977.5

Р. ТЕННО

СВОЙСТВА ДВУХШАГОВОГО ОПТИМАЛЬНОГО УПРАВЛЕНИЯ РЕГРЕССИОННОЙ МОДЕЛЬЮ СО СЛУЧАЙНЫМИ КОЭФФИЦИЕНТАМИ

R. TENNO. JUHUSLIKE KOEFITSIENTIDEGA REGRESSIOONVÕRRANDI KAHESAMMULISE OPTIMAALJUHTIMISE OMADUSI

R. TENNO. SOME PROPERTIES OF TWO STEP OPTIMAL CONTROL OF THE RANDOM COEFFICIENT REGRESSION EQUATIONS

(Представил Н. Алумяэ)

Задачи дуального управления [1] приходится решать во многих практических ситуациях, в связи с чем разработано немало приближенных алгоритмов этого управления [2, 3]. Однако анализ свойств предложенных управлений проведен гораздо хуже. По-видимому, это связано с излишней общностью постановки задач. Нами поставлена цель исследовать сравнительно простую задачу дуального управления, имеющую обозримую структуру. При этом рассмотрены асимптотические свойства и достаточные условия единственности оптимального управления, найдена оценка сверху для разности стоимостей одношагового и оптимального управлений.

Задача управления. Предположим, что объект управления с выходом x_t, $t=1, 2, \ldots, N$, и вектором входов (управлением) u_{t-1} описывается регрессионной моделью

$$x_t=u^{\mathrm{T}}_{t-1}\Theta_t+\xi_t$$

со случайными коэффициентами Θ_t. Помеха ξ_t удовлетворяет уравнению

$$\xi_t=\mu_0+\Phi(\xi_{t-1}-\mu_0)+\alpha_t, \qquad \xi_0=\xi,$$

где μ_0, Φ — скалярные параметры, (α_t) — «белый шум», ξ — (случайное) начальное условие. Измеряется

$$y_t=x_t+h_t.$$

Пусть при каждом t величины h_t, α_t, Θ_t, ξ независимы и распределены по нормальному закону со средними

$$Mh_t=0, \qquad M\alpha_t=0, \qquad M\Theta_t=\mu_1, \qquad M\xi=b$$

и (ко-)вариациями

$$\operatorname{var} h_t=r, \qquad \operatorname{var}\alpha_t=\sigma, \qquad \operatorname{cov}\Theta_t=D_1, \qquad \operatorname{var}\xi=p,$$

причем $r>0$, а матрица D_1 положительно определена. Требуется при заданных μ_0, Φ, r, σ, μ_1, D_1, b, p выбрать управление $u \equiv u_0$ на первом шаге такое, которое минимизирует двухшаговый ($N=2$) функционал

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

где x^0 — задающее воздействие, в предположении, что управление $u_1(y_1)$ на втором шаге выбрано оптимальным. Согласно алгоритму динамического программирования оптимальное управление u^* на первом шаге выбирается из условия минимума функции

$$V(u) = S(u) + R(u),$$

где $S(u)$ — риск действия, $R(u)$ — риск изучения. В нашем случае [4]

$$S(u) = (a + u^{\mathrm{T}}\mu_1 - x^0)^2 + G + u^{\mathrm{T}} D_1 u, \tag{1}$$

$$R(u) = A - BT^2, \tag{2}$$

где

$$A = K + (c - x^0)^2/(1 + \mu_1^{\mathrm{T}} D_1^{-1}\mu_1), \quad B = \mu_1^{\mathrm{T}} D_1^{-1}\mu_1/(1 + \mu_1^{\mathrm{T}} D_1^{-1}\mu_1),$$

$$T^2 = (\Phi G)^2/(r + G + u^{\mathrm{T}} D_1 u),$$

a, c — оценки (прогнозы) возмущений ξ_1, ξ_2 (определенные до момента поступления наблюдений); G, K — ковариации оценок a, c.

Одношаговым управлением u^+ назовем такое управление, которое минимизирует функцию $S(u)$.

Асимптотические свойства оптимального управления. Рассмотрим в каком случае оптимальное u^* и одношаговое управление u^+ имеют сходные асимптотические свойства в том смысле, что при стремлении в бесконечность некоторого параметра ϱ можно считать, что процесс изучения объекта происходит одинаково при любом управлении.

Утверждение. *Пусть $\varrho > 0$ и $\varrho \to \infty$. Если $\Phi = \varrho$ или $\sigma = \varrho^2$ и $D_1 = \varrho^2 L$ (где L — положительно определенная матрица), то $u_\varrho^* \to 0$, если $r = \varrho^2$, то $u_\varrho^* \to u^+$.*

Доказательство. Пусть $\Phi = \varrho$. Тогда первые и вторые производные от функции $S(u)$, $R(u)$ можно представить следующим образом:

$$S_u = \varrho \bar{S}_1(u, \varrho), \quad R_u = \varrho^2 \bar{R}_1(u, \varrho), \quad R_{uu} = \varrho^2 \bar{R}_2(u, \varrho),$$

где $\bar{S}_1(u, \cdot)$, $\bar{R}_1(\cdot,\cdot)$ — векторы, $\bar{R}_2(\cdot,\cdot)$ — матрица, координаты (элементы) которых ограниченные функции. Из условия локального минимума функции $V(u)$ получим

$$\varrho^{-1}\bar{S}_1(u, \varrho) + \bar{R}_1(u, \varrho) = 0, \quad \varrho^{-2} S_{uu} + \bar{R}_2(u, \varrho) > 0. \tag{3}$$

Условиям

$$\lim_{\varrho\to\infty} \bar{R}_1(u, \varrho) = 2BD_1 u = 0, \quad \lim_{\varrho\to\infty} \bar{R}_2(u, \varrho) = 2BD_1 > 0$$

удовлетворяет управление $\bar{u} = 0$. Так как функция $\bar{S}_1(u, \cdot)$ ограничена, а матрица S_{uu} положительно определена, то из (3) видно, что $u_\varrho^* \to 0$.

В случае $\sigma = \varrho^2$ и $D_1 = \varrho^2 L$ утверждение доказывается аналогично. В случае $r = \varrho^2$ вместо (3) получим

$$S_u + \varrho^{-4}\bar{R}_1(u, \varrho) = 0, \quad S_{uu} + \varrho^{-4}\bar{R}_2(u, \varrho) > 0. \tag{4}$$

Условиям $S_u = 0$, $S_{uu} > 0$ удовлетворяет одношаговое управление u^+, функции $\bar{R}_1(\cdot,\cdot)$, $\bar{R}_2(\cdot,\cdot)$ ограничены, и поэтому из (4) следует, что $u_\varrho^* \to u^+$.

Замечание. Если $\Phi=\varrho$, то асимптотические свойства оптимального и одношагового управлений отличаются, так как $|u_\rho^+|\to\infty$, при $\varrho\to\infty$. Если $D_1=\varrho^2 L$, то имеем $u_\rho^+\to 0$.

Единственность оптимального управления. В случае одномерного управления можно получить следующее легко проверяемое условие единственности.

Неравенство

$$(\Phi\Xi)^2 B<4, \tag{5}$$

где

$\Xi=G/(r+G)$, гарантирует единственность оптимального управления. Убедимся в этом. Строго выпуклая функция имеет максимум в единственной точке. Если для каждого управления функция $S_{uu}+R_{uu}$ положительна, то $V(u)$ строго выпукла. Поскольку $R_{uu}(u)$ снизу ограничена

$$\min_u R_{uu}(u)=-0{,}5(\Phi\Xi)^2 BD_1,$$

то требуемое выше условие выполняется, если функция

$$S_{uu}+\min_u R_{uu}=2\,\{\mu_1^2+[1-0{,}25(\Phi\Xi)^2 B]D_1\}$$

положительная или если выражение в квадратных скобках положительное, что и является достаточным условием единственности оптимального управления.

Отметим, что если процесс аддитивных возмущений (ξ_t) стационарный, то оптимальное управление единственное. Это следует из условий стационарности $|\Phi|<1$ и из того, что $B<1$, $\Xi<1$.

Оценка разности стоимостей одношагового и оптимального управлений. Ниже показано, что если потери от неоптимального управления большие, то это менее существенно связано с риском изучения, чем с риском действия. Поясним это более подробно.

Пусть $u^+-u=\delta$. Тогда а) из (1) видно, что

$$S(u)-S(u^+)=\delta^{\mathrm{T}}M\delta,$$

где

$$M=\mu_1\mu_1^{\mathrm{T}}+D_1;$$

б) существует постоянная $N\geqslant 0$ такая, что для любого управления

$$|R(u)-R(u^+)|\leqslant N|\delta|. \tag{6}$$

В случае «больших» $|\delta|$ оценку (6) можно заменить более точной

$$|R(u)-R(u^+)|\leqslant K,\quad K>0.$$

Докажем второе утверждение. По теореме о конечных приращениях в интервале (u, u^+) существует вектор v такой, что

$$R(u)-R(u^+)=R_u^{\mathrm{T}}(v)\,\delta.$$

Следовательно, достаточно показать, что

$$|R_u(v)|\leqslant N.$$

После несложных преобразований, получим

$$N=2B[\Phi G/(r+G+\tilde{u}^{\mathrm{T}}D_1\tilde{u})]^2|D_1\tilde{u}|,$$

где $\tilde{u}$ — решение уравнения

$$(r+G+u^{\mathrm{T}}D_1u)D_1=4D_1uu^{\mathrm{T}}D_1. \qquad (7)$$

Так как матрица D_1 положительно определена, то все решения уравнения (7) ограничены. Следовательно, длина вектора $D_1\tilde{u}$ также является ограниченной.

Если $|\delta|$ — «большой» и $R(u^+)\geqslant R(u)$, то из определения (2) получим

$$R(u^+)-R(u)\leqslant B\max_u [T^2(u)-T^2(u+\delta)]\leqslant$$

$$\leqslant B[\max_u T^2(u)-\min_u T^2(u+\delta)]=B\frac{(\Phi G)^2}{r+G}.$$

Если $R(u)\geqslant R(u^+)$, то получим также

$$R(u)-R(u^+)<B\frac{(\Phi G)^2}{r+G}.$$

Определим оценку разности стоимостей одношагового и оптимального управлений. Пусть $\delta=u^+-u^*$. Если $|\delta|$ — «малый», то из утверждений а), б) и определения $V(u)$ следует, что

$$V(u^+)-V(u^*)=R(u^+)-R(u^*)-\delta^{\mathrm{T}}M\delta<N|\delta|.$$

Если $|\delta|$ — «большой», то

$$V(u^+)-V(u^*)\leqslant K.$$

Таким образом, потери оптимальности от одношагового управления ограничены сверху, причем малые потери определяются главным образом риском изучения.

Автор выражает благодарность В. Ольману за полезные замечания.

ЛИТЕРАТУРА

1. *Фельдбаум А. А.* Основы теории оптимальных автоматических систем (2-е изд.). М., «Наука», 1966.
2. *Asher, R. B., Andrisani II D., Donato, P.* Proc IEEE, **64**, 1226—1240 (1976).
3. *Tse, E., Bar-Shalom, Y.* Proc. of the 18th IEEE Conf. on Decision and Control Including the Symposium on Adaptive Processes, Fort Lauderdale Fla., 1979, **1**, New York, 1979, 183—196.
4. *Тенно Р.* Изв. АН ЭССР. Физ. Матем., **32**, № 1, 1—16 (1983).

Институт кибернетики
Академии наук Эстонской ССР

Поступила в редакцию
7/VI 1983

После переработки
23/IV 1984