EESTI NSV TEADUSTE AKADEEMIA TOIMETISED. 19. KÖIDE
FÜÜSIKA * MATEMAATIKA. 1970, NR. 4

ИЗВЕСТИЯ АКАДЕМИИ НАУК ЭСТОНСКОЙ ССР. ТОМ 19
ФИЗИКА * МАТЕМАТИКА. 1970, № 4

https://doi.org/10.3176/phys.math.1970.4.01

Л. МЫТУС

УПРАВЛЕНИЕ ОБЪЕКТОМ СО СЛУЧАЙНО ИЗМЕНЯЮЩИМСЯ ПАРАМЕТРОМ

Рассматривается одномерный объект управления, параметр которого изменяется случайно в случайные моменты времени. Приведены рекурсивные формулы для вычисления приближенных апостериорных функций плотности вероятности (ф. п. в.) неизвестного параметра, сохраняющие свою функциональную форму при увеличении числа наблюдений. Найдена оптимальная стратегия, минимизирующая суммарный риск (3). Стратегия определяет значения управляющего воздействия на этапах и условия для окончания этапов. Приведены также результаты моделирования системы управления на ЦВМ.

Отметим, что нахождение оценок для случайно изменяющегося параметра по фиксированному числу наблюдений рассмотрено в [1].

Отметим еще некоторые работы, связанные с рассматриваемыми в данной статье вопросами. Управление случайным процессом с учетом стоимости отклонения процесса от заданного режима и затрат, связанных с управлением и наблюдениями, рассмотрены в [2]. Рекурсивное оценивание медленно изменяющегося марковского параметра рассмотрено в [3, 4].

1. Постановка задачи

Пусть дискретно-непрерывный объект управления описывается уравнениями:

$$\begin{cases} x_i = \theta_i + u_i, \\ \theta_i = \theta_{i-1} + h_i z_i, \\ y_i = x_i + n_i, \\ i = 1, 2, \ldots, N, \end{cases} \tag{1}$$

где y_i — значение выхода, x_i — состояние объекта, θ_i — неизвестный параметр, u_i — управляющее воздействие, h_i, z_i и n_i — независимые случайные величины на i-м такте.

Обозначим через $L(x)$ ф. п. в. случайной величины x, а через $\mathrm{Nor}\,(a, b)$ — ф. п. в. нормального распределения со средним a и дисперсией b.

Пусть ф. п. в. для z_i и n_i известны и пусть

$$L(z_i) = \mathrm{Nor}\,(\overline{z}, w^2) \quad \text{и} \quad L(n_i) = \mathrm{Nor}\,(0, 1).$$

Пусть для h и θ_0 заданы априорные ф. п. в.:

$$L(\theta_0) = \mathrm{Nor}\,(\overline{\theta}_0, v_0^2) \quad \text{и} \quad L(h) = b\,(1, 1, p_0),$$

где $b\;(1, 1, p_0)$ — биномиальное распределение. Другими словами, в каждый момент времени с вероятностью p происходит изменение значения параметра θ.

Функция потерь имеет вид

$$F = \sum_{j=1}^{l} \Bigl[\sum_{i=k_{j-1}}^{k_j} (x_i - \tilde{x}_i)^2 + c \Bigr], \tag{2}$$

где $k_0 = 1$, $k_l = N$, N — число тактов, l — число этапов, c — цена изменения управляющего воздействия, $\tilde{x}$ — заданное состояние объекта. Так как x — случайная величина, то мы минимизируем математическое ожидание функции потерь (суммарный риск). Обозначим

$$R_j = E_x \Bigl\{ \sum_{i=k_{j-1}}^{k_j} (x_i - \tilde{x}_i)^2 \Bigr\} + c = \sum_{i=k_{j-1}}^{k_j} r_i + c,$$

где

$$r_i = E_x (x_i - \tilde{x}_i)^2 = \int (\theta_i + u_i - \tilde{x}_i)^2 L(\theta_i / y^{i-1})\, d\theta_i,$$

а E_x — символ условного математического ожидания по x при фиксированном $y^{i-1} = (y_1, y_2, \ldots, y_{i-1})$.

Суммарный риск

$$R = E \sum_{j=1}^{l} R_j = R(u_1, u_2, \ldots, u_l, l), \tag{3}$$

где математическое ожидание взято по наблюдениям.

Поскольку одинаковой продолжительности этапов не требуется, число этапов не фиксировано и условия для окончания этапов определяются из условия минимума суммарного риска.

2. Нахождение апостериорных ф. п. в.

Для нахождения ф. п. в. $L(\theta_i / y^{i-1})$, необходимой для вычисления математического ожидания от (2), найдем сперва $L(\theta_i / y^i)$ и $L(h / y^i)$.

Начинаем с ф. п. в. $L(\theta_i / y^i)$. По формуле Байеса имеем

$$L(\theta_i / y^i, h_i) = \frac{L(\theta_i, y_i / y^{i-1}, h_i)}{L(y_i / y^{i-1}, h_i)},$$

где $L(\theta_i, y_i / y^{i-1}, h_i) = \int L(y_i / \theta_i) L(\theta_i / \theta_{i-1}, h_i) L(\theta_{i-1} / y^{i-1})\, d\theta_{i-1}$

и $$L(y_i / y^{i-1}, h_i) = \int L(\theta_i, y_i / y^{i-1}, h_i)\, d\theta_i. \tag{4}$$

Из уравнений (1) имеем

$L(y_i / \theta_i) = \mathrm{Nor}(\theta_i + u_i, 1)$ и $L(\theta_i / \theta_{i-1}, h_i) = \mathrm{Nor}(\theta_{i-1} + h_i \bar{z}, h_i^2 w^2)$.

Теперь легко найти

$$\begin{aligned} L(\theta_i / y^i) &= E_h L(\theta_i / y^i, h_i) = \\ &= p_i L(\theta_i / y^i, h_i = 1) + q_i L(\theta_i / y^i, h_i = 0), \end{aligned} \tag{*}$$

где $p_i = L(h = 1 / y^i)$ и $q_i = 1 - p_i$.

Так как найденная ф. п. в. изменяет свою функциональную форму при увеличении числа наблюдений, в дальнейшем воспользуемся линейной частью разложения выражения (*) в ряд Тейлора по $\bar{z}$ и w^2, т. е. пренебрежем членом $o(\bar{z}) + o(w^2)$. Полученная таким образом приближенная ф. п. в. сохраняет при увеличении числа наблюдений свою функциональную форму

$$L(\theta_i/y^i) = (2\pi v_i^2)^{-0,5} \exp\left[-\frac{(\theta_i-\overline{\theta}_i)^2}{2v_i^2}\right]\left\{1 + a_i\overline{z}(\theta_i-\overline{\theta}_i) + \right. \tag{5}$$

$$\left. + e_i w^2(\theta_i-\overline{\theta}_i) + b_i w^2(\theta_i-\overline{\theta}_i)^2 - b_i v_i^2 w^2\right\},$$

где $a_i = a_{i-1} + p_i/v_{i-1}^2$, $\quad b_i = b_{i-1} + p_i/2v_{i-1}^4$,

$$e_i = e_{i-1} + 2b_i v_i^2(y_i - u_i - \overline{\theta}_{i-1}), \quad \overline{\theta}_i = (\overline{\theta}_{i-1} + (y_i - u_i)v_{i-1}^2)(1 + v_{i-1}^2)^{-1},$$

$$v_i^2 = v_{i-1}^2(1 + v_{i-1}^2)^{-1} \quad \text{и} \quad a_0 = b_0 = e_0 = 0.$$

Формулу (5) можно доказать при помощи индукции. Отметим, что если $p_0 = 0$, то (5) дает точную апостериорную ф. п. в.

Для определения $L(h/y^i)$ имеем

$$p_i = p_{i-1}\frac{L(y_i/y^{i-1}, h_i = 1)}{\sum\limits_h L(h/y^{i-1})L(y_i/y^{i-1}, h)}, \tag{6}$$

где $p_i = L(h = 1/y^i)$.

Вывод более удобной формы для (6) дан в приложении.

И, наконец, для определения ф. п. в. $L(\theta_i/y^{i-1})$ имеем

$$L(\theta_i/y^{i-1}, h_i) = \int L(\theta_i/\theta_{i-1}, h_i)L(\theta_{i-1}/y^{i-1})d\theta_{i-1}.$$

Разлагая $E_h L(\theta_i/y^{i-1}, h_i)$ в ряд Тейлора и пренебрегая членом $o(\overline{z}) + o(w^2)$, получим приближенную ф. п. в.

$$L(\theta_i/y^{i-1}) = (2\pi v_{i-1}^2)^{-0,5}\exp\left[-\frac{(\theta_i - \overline{\theta}_{i-1})^2}{2v_{i-1}^2}\right]\{1 + \hat{a}_i\overline{z}(\theta_i - \overline{\theta}_{i-1}) +$$

$$+ e_{i-1}w^2(\theta_i - \overline{\theta}_{i-1}) + \hat{b}_i w^2(\theta_i - \overline{\theta}_{i-1})^2 - \hat{b}_i w^2 v_{i-1}^2\},$$

где $\hat{a}_i = a_{i-1} + p_{i-1}/v_{i-1}^2$ и $\hat{b}_i = b_{i-1} + p_{i-1}/2v_{i-1}^4$.

Математическое ожидание $\tilde{\theta}_i$ и дисперсия $\tilde{v}_i^2$ случайной величины θ_i с ф. п. в. $L(\theta_i/y^{i-1})$ соответственно равны

$$\tilde{\theta}_i = \overline{\theta}_{i-1} + v_{i-1}^2(\hat{a}_i\overline{z} + e_{i-1}w^2),$$

$$\tilde{v}_i^2 = v_{i-1}^2[1 - v_{i-1}^2(\hat{a}_i\overline{z} + e_{i-1}w^2)^2 + 2\hat{b}_i v_{i-1}^2 w^2].$$

3. Нахождение оптимальной стратегии

Выражение (3) нужно минимизировать по $u_1, u_2, \ldots, u_l, l$. Стратегия определяет значения управляющего воздействия на этапах и условия для окончания этапов.

Решаем задачу при помощи динамического программирования. Будущий риск, когда до конца остается один такт, определяется согласно принципу оптимальности выражением

$$V_1 = \min\{r_N(u_{N-1}), r_N^* + c\},$$

где $r_N^* = \min\limits_{u_N} r_N = \min\limits_{u_N} E_\theta(\theta_N + u_N - \tilde{x}_N)^2$, а $\quad r_N(u_{N-1}) = E_\theta(\theta_N + u_{N-1} - \tilde{x}_N)^2$.

Для предыдущего такта имеем

$$V_2 = \min \{r_{N-1}(u_{N-2}) + E_y V_1, \min_{u_{N-1}} [r_{N-1} + E_y V_1] + c\}.$$

Покажем, что система является нейтральной в смысле Фельдбаума [5], т. е. что $E_y V_1$ не зависит от u_{N-1}. Рассмотрим выражение $r_N^* + c$.

Перепишем r_N^*, чтобы выделить члены, содержащие y_{N-1} и u_{N-1}:

$$r_N^*(y_{N-1} - u_{N-1}) = v_{N-1}^{2}[1 - v_{N-1}^2\{[a_{N-2} + p_{N-1}(y_{N-1} - u_{N-1}) \times$$
$$\times [v_{N-2}^{-2} + v_{N-1}^{-2}]]\overline{z} + [e_{N-2} + [2b_{N-2} + p_{N-1}(y_{N-1} - u_{N-1}) v_{N-2}^{-4}] \times$$
$$\times v_{N-1}^2[y_{N-1} - u_{N-1} - \overline{\theta}_{N-2}]]w^2\}^2 + 2[b_{N-2} + 0{,}5 p_{N-1}(y_{N-1} - u_{N-1}) \times$$
$$\times [v_{N-2}^{-4} + v_{N-1}^{-4}]] v_{N-1}^2 w^2].$$

Сделав в интеграле $\int r_N^*(y_{N-1} - u_{N-1}) L(y_{N-1}/y^{N-2})\, dy_{N-1}$ замену переменных $y_{N-1} - u_{N-1} = g_{N-1}$, видим, что $E_y(r_N^* + c)$ не зависит от u_{N-1}. Аналогично и $E_y r_N(u_{N-1})$ не зависит от u_{N-1}. Можно показать, что система является нейтральной и при точных ф. п. в.

Воспользуясь нейтральностью системы, получим

$$V_2 = \min \{r_{N-1}(u_{N-2}) + E_y V_1,\ r_{N-1}^* + E_y V_1 + c\}$$

и, если $r_{N-1}(u_{N-2}) < r_{N-1}^* + c$, то в $(N-1)$-м такте управляющее воздействие не меняется.

Введя множество A_{ij} значений θ, при которых на i-м такте можно продолжать применение управляющего воздействия j-го такта ($i > j$, $i = 2, 3, \ldots, N$), получим

$$A_{ij} = \{\theta : r_i(u_j^*) < r_i^* + c\},$$

где $u_j^* = \tilde{x}_j - \tilde{\theta}_j$ найдено из уравнения $\dfrac{\partial r_j}{\partial u_j} = 0$.

Учитывая, что

$$r_i(u_j^*) = \tilde{v}_i^2 + (\tilde{x}_j - \tilde{x}_i - \tilde{\theta}_j + \tilde{\theta}_i)^2$$

и что $\quad r_i^* = \tilde{v}_i^2$, получим

$$A_{ij} = \{\theta : \tilde{\theta}_j + \tilde{x}_i - \tilde{x}_j - \sqrt{c} < \tilde{\theta}_i < \tilde{\theta}_j + \tilde{x}_i - \tilde{x}_j + \sqrt{c}\}. \tag{7}$$

Теперь можно сформулировать оптимальную стратегию:

$$u_i = u_j^* \text{ при } \tilde{\theta}_i \in A_{ij} \text{ и } u_i = u_i^* \text{ при } \tilde{\theta}_i \notin A_{ij}, \tag{8}$$

где $i > j$, $i = 2, 3, \ldots, N$ и $u_1 = u_1^*$.

4. Результаты моделирования

Целью моделирования являлось исследование влияния упрощений, сделанных при нахождении ф. п. в., на оценку параметра и сравнение стратегии (7)—(8), где продолжительность этапов была случайной, со стратегией, этапы которой имели фиксированные одинаковые длины.

Для усреднения результатов было сделано 500 экспериментов с каждой из трех ф. п. в. z: 1) $\bar{z} = 0{,}02\,\bar{\theta}_0$, $w^2 = 0{,}02\,\bar{\theta}_0$; 2) $\bar{z} = 0{,}02\,\bar{\theta}_0$, $w^2 = 0{,}1\,\bar{\theta}_0$ и 3) $\bar{z} = 0{,}1\,\bar{\theta}_0$, $w^2 = 0{,}02\,\bar{\theta}_0$.

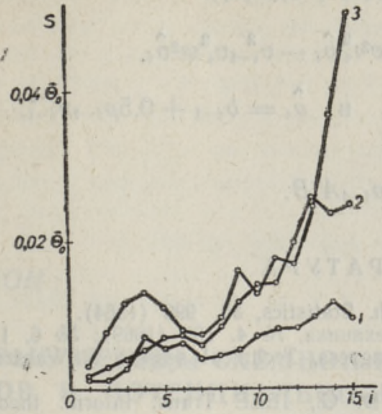

Рис. 1.

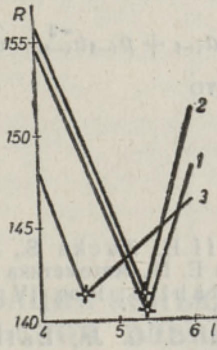

Рис. 2.

На рисунках приведены два графика: 1) среднеквадратичное отклонение s оценки параметра от действительного значения параметра; 2) зависимость суммарного риска от числа этапов. Крестиком обозначен суммарный риск, соответствующий оптимальному числу этапов. Для большей наглядности отдельные точки на графиках соединены линиями.

Выводы: 1) Оценивание при помощи приближенных ф. п. в. дает удовлетворительные результаты. Если $\bar{z} \leqslant 0{,}1\,\bar{\theta}_0$ и число тактов небольшое ($N = 15$), то $s \leqslant 0{,}05\,\bar{\theta}_0$.

2) Оценка параметра (математическое ожидание случайной величины θ_i с ф. п. в. $L(\theta_i/y^i)$) более чувствительна к изменениям $\bar{z}$, чем к изменениям w^2.

3) Стратегия (7)—(8) дает почти всегда меньший суммарный риск, чем стратегия с фиксированными продолжительностями этапов. Только в двух случаях из восемнадцати стратегия (7)—(8) оказалась хуже. (Оба случая с третьей ф. п. в. z).

Приложение

Для определения p_i имеем формулу (6). Из (4) следует, что

$$L(y_i/y^{i-1}, h_i) = \iint L(y_i/\theta_i)\, L(\theta_i/\theta_{i-1}, h_i)\, L(\theta_{i-1}/y^{i-1})\, d\theta_i\, d\theta_{i-1},$$

где $L(y_i/\theta_i)$ и $L(\theta_i/\theta_{i-1}, h_i)$ получаем из уравнений (1), а $L(\theta_{i-1}/y^{i-1})$ определено выражением (5).

Получим, что

$$L(y_i/y^{i-1}, h_i) = (2\pi(1 + v_{i-1}^2 + h_i^2 w^2))^{-0,5} \exp\left[-\frac{(t_i - h_i\bar{z})^2}{2(1 + v_{i-1}^2 + h_i^2 w^2)}\right]\Big\{1 + a_{i-1}\bar{z}v_{i-1}^2 m_i + e_{i-1}w^2 v_{i-1}^2 m_i + b_{i-1}w^2 v_{i-1}^4 m_i - b_{i-1}w^2 v_{i-1}^4 (1 + v_{i-1}^2 + h_i^2 w^2)^{-1}\Big\},$$

где $m_i = (t_i - h_i\bar{z})(1 + v_{i-1}^2 + h_i^2 w^2)^{-1}$, $\quad t_i = y_i - u_i - \bar{\theta}_{i-1}$.

Подставляя приведенные выражения в (6) и обозначая

$$A = 1 + v_i^2 \bar{z} t_i (a_{i-1} + v_{i-1}^{-2}) + w^2 e_{i-1} v_i^{2'} t_i + v_i^{4'} w^2 t_i^2 (b_{i-1} + 0{,}5 v_{i-1}^{-4}) -$$
$$- v_{i-1}^{2'} v_i^2 w^2 (b_{i-1} + 0{,}5 v_{i-1}^{-4})$$

и $B = 1 + v_i^2 \bar{z} t_i \hat{a}_i + w^2 \hat{e}_i v_i^2 t_i + v_i^{4'} w^2 t_i^2 \hat{b}_i - v_{i-1}^{2'} v_i^2 w^2 \hat{b}_i,$

где $\hat{a}_i = a_{i-1} + p_{i-1} v_{i-1}^{-2}$, $\hat{e}_i = e_{i-1}$ и $\hat{b}_i = b_{i-1} + 0{,}5 p_{i-1} v_{i-1}^{-4}$,

получим, что

$$p_i = p_{i-1} A/B.$$

ЛИТЕРАТУРА

1. Chernoff H., Zacks S., Ann. Math. Statistics, 35, 999 (1964).
2. Маслов Е. П., Автоматика и телемеханика, № 4, 183 (1969); № 6, 191 (1969).
3. Segerståhl B., Proc. IV IFAC Congress, Technical session 5, Warszawa, 1969. p. 62.
4. Hilborn C. G. Jr., Lainiotis D. G., IEEE Trans. inform. theory, IT-14, 514 (1968).
5. Фельдбаум А. А., Основы теории оптимальных автоматических систем, М., 1966.

Институт кибернетики
Академии наук Эстонской ССР

Поступила в редакцию
3/XI 1969

L. MÕTUS

JUHUSLIKULT MUUTUVA PARAMEETRIGA OBJEKTI JUHTIMISEST

Esitatakse juhuslikul hetkel juhuslikult muutuva parameetriga objekti ligikaudsed taastuvad aposterioorsed jaotustihedused. On välja töötatud summaarset riski minimiseeriv strateegia, mis määrab juhttoime väärtuse etappidel ja etappide lõputingimused. Imiteerimistulemused kinnitavad, et teatud tingimustel on ligikaudsete jaotustiheduste kasutamine õigustatud.

L. MÕTUS

ON THE CONTROL OF A RANDOM PARAMETER PLANT

The parameter of a plant is subjected to change randomly at random time instants. Recursive formulae are presented for computing self-reproducing approximate a posteriori probability densities for solving control problem. Optimal strategy is developed to determine values of the control variable at stages and moments for changing these values. Simulation shows that the method gives satisfactory results under certain conditions.