

Р. ТЕННО, Х. ОЙТ

АДАПТИВНОЕ УПРАВЛЕНИЕ ЛИНЕЙНЫМ ДИНАМИЧЕСКИМ ОБЪЕКТОМ С НЕИЗВЕСТНЫМИ ПАРАМЕТРАМИ

(Представил Н. Алумяэ)

Разработан новый алгоритм адаптивного управления специального вида линейным динамическим объектом, неизвестные параметры которого постоянны или изменяются как авторегрессионный случайный процесс, а выход объекта наблюдается со случайными ошибками.

1. Постановка задачи

Пусть частично наблюдаемый случайный процесс $(x, y) = \{x(t), y(t), t = 1, 2, \dots, N\}$ задан линейными уравнениями

$$x(t) = \sum_{i=1}^q u^T(t-i) \beta_i(t) + \beta_0(t), \quad u(t < 0) = 0, \quad (1)$$

$$y(t) = x(t) + h(t), \quad (2)$$

где $u(t)$, $x(t)$ — вход (k -мерный) и выход управляемого процесса, $y(t)$ — наблюдаемый выход. Допускается, что ошибки наблюдения $\{h(t), t = 1, 2, \dots, N\}$ независимы и распределены по невырожденному нормальному закону с нулевым средним и с дисперсией r . Предполагается, что ненаблюдаемый процесс $\beta = \{\beta(t) : t = 1, 2, \dots\}$ изменения параметров $\beta_i(t)$, $i = 0, \dots, q$ уравнения (1) описывается моделью авторегрессии

$$\beta(t) = \mu + \Phi[\beta(t-1) - \mu] + \alpha(t), \quad \beta(0) = \bar{\beta}, \quad (3)$$

где Φ — матрица параметров авторегрессии, μ — вектор среднего, $\alpha(t)$, $t = 1, 2, \dots$ — случайные независимые векторы (белый шум), распределенные по нормальному закону с нулевым средним и неотрицательно определенной матрицей ковариации D . Векторы $\alpha(t)$, $\bar{\beta}$ и скаляр $h(t)$ являются независимыми для каждого $t = 1, 2, \dots, N$. Также попарно независимыми принимаются векторы $\beta_1(t) \dots \beta_q(t)$ и скаляр $\beta_0(t)$. Вектор начальных условий $\bar{\beta}$ уравнения (3) предполагается гауссовым с математическим ожиданием \bar{b} и с ковариацией Ξ .

Предполагается, что известны параметры Φ , μ модели (3), среднее \bar{b} и ковариация Ξ начальных условий $\bar{\beta}$ и ковариации r и D процессов h и α соответственно. Ставится задача нахождения в виде явной функции оптимальных управлений, минимизирующих функционал,

$$J = M \sum_{t=1}^N W(t), \quad (4)$$

где M — математическое ожидание, N — горизонт управления, $W(t)$ — функция потерь вида

$$W(t) = [x(t) - x^0]^2 + u^T(t-1) \Upsilon u(t-1),$$

x^0 — задающее воздействие, Υ — положительно определенная весовая матрица. Предполагается, что стратегия допустимых управлений $u = \{u(0), \dots, u(t), \dots, u(N-1)\}$ задана непрерывными функциями $u(t) = u(t, y(1), \dots, y(t))$, ограниченными почти наверное.

Если оптимальное управление не выражается в виде явной функции, то ставится задача определения таких управлений, которые при каждом t минимизируют функционал

$$I_t = \sum_{l=0}^{N-t} M\{W(t+l)/F(t-1)\}, \quad (5)$$

где $F(t-1)$ — σ -алгебра, согласованная с процессом $\{y(s), s = 1, \dots, t-1\}$. С помощью (5) построенная стратегия является нейтральной в том смысле, что она направлена в каждый момент времени только на приведение объекта к требуемому состоянию и не имеет специальной функции по изучению объекта [1].

2. Построение стратегии управления

В стохастических системах основным математическим аппаратом для отыскания оптимальных управлений является принцип оптимальности Беллмана, согласно которому оптимальное управление для каждого $t = 1, 2, \dots, N$ определяется из решения уравнения

$$V(t) = \min_{u(t-1)} M\{W(t) + V(t+1)/F(t-1)\}, \quad (6)$$

где $V(t)$ — цена (функция Беллмана). В частично наблюдаемых случайных процессах (β, x, y) строгое применение принципа оптимальности Беллмана наталкивается на значительные трудности, что, в частности, связано с зависимостью условных распределений $\pi(t+1|t), \dots, \pi(N|N-1)$, необходимых для определения цены (7), от управления $u(t-1)$.

$$V(t) = \min_{u(t-1)} \langle W(t) + \min_{u(t)} \langle W(t+1) + \dots + \min_{u(N-1)} \langle W(N), \pi(N|N-1) \rangle, \dots, \pi(t+1|t) \rangle, \pi(t|t-1) \rangle, \quad (7)$$

где

$$\langle W(\tau) + \dots, \pi(\tau|\tau-1) \rangle = \int_{R^{q^{k+1}}} [W(\beta(\tau), u) + \dots] d\pi\{\beta(\tau) | F(\tau-1)\},$$

$\pi\{\beta(\tau) | F(\tau-1)\}, \tau = t, \dots, N$ — условное распределение вероятностей $\beta(\tau)$ относительно $F(\tau-1)$. В общем случае не удастся найти оптимального управления в виде явной функции. Поэтому, согласно (5), будет построена нейтральная стратегия управления с ценой

$$\bar{V}(t) = \min_{u(t-1)} \{\langle W(t), \pi(t|t-1) \rangle + \min_{u(t)} \{\langle W(t+1), \pi(t+1|t-1) \rangle + \dots + \min_{u(N-1)} \langle W(N), \pi(N|t-1) \rangle\}\}, \quad (8)$$

где $\pi(s+l|s) = \pi\{\beta(s+l) | F(s)\}, s = t-1, l = 1, \dots, N-s$ — условное распределение вероятностей $\beta(s+l)$ относительно $F(s)$. Отметим, что в тех случаях, когда применим принцип разделения и оптимальное управление выражается в виде явной функции, задачи (4) и (5) экви-

валентны. Рассмотрим решение задачи (8). В силу принятых условий легко убедиться в том, что случайный процесс (β, y) является условно гауссовым [2]. Поэтому условное распределение $\pi\{\beta(s) | F(s)\}$, $s = t-1$ является гауссовым с вектором среднего $b(s)$ и матрицей ковариации $D(s)$, определяемыми из рекуррентных уравнений

$$b(s) = \mu + \Phi[b(s-1) - \mu] + \Gamma\{y(s) - z^T[\mu + \Phi[b(s-1) - \mu]]\}, \quad b(0) = \bar{b}, \quad (9)$$

$$D(s) = D + \Phi D(s-1) \Phi^T - \Gamma z^T [D + \Phi D(s-1) \Phi^T], \quad D(0) = \Xi, \quad (10)$$

где

$$\Gamma = [D + \Phi D(s-1) \Phi^T] z / \{r + z^T [D + \Phi D(s-1) \Phi^T] z\}, \\ z = [1, u^T(s-1), \dots, u^T(s-q)]^T.$$

Случайный процесс β не зависит от наблюдений y . Следовательно [2], условное распределение $\pi\{\beta(s+l) | F(s)\}$, $s = t-1$ прогнозируемого параметра $\beta(s+l)$ является также гауссовым для каждого $l = 1, \dots, N-s$, причем вектор среднего $b(s+l|s)$ и матрица ковариации $D(s+l|s)$ этого распределения определяются из соотношений

$$b(s+l|s) = \mu + \Phi[b(s+l-1|s) - \mu], \quad b(s|s) = b(s), \quad (11)$$

$$D(s+l|s) = D + \Phi D(s+l-1|s) \Phi^T, \quad D(s|s) = D(s). \quad (12)$$

Примем гипотезу о том, что оценки $b_0(s), b_1(s), \dots, b_q(s)$ попарно независимы, тогда прогнозы параметров также независимы, и из-за гауссовости распределения $\pi(s+l|s)$ величина средних потерь $\langle W(s+l), \pi(s+l|s) \rangle$ для каждого $l = 1, \dots, N-s$ может быть выражена следующим образом

$$\langle W(s+l), \pi(s+l|s) \rangle = [\hat{b}_0(s+l) + \sum_{i=1}^q u^T(s+l-i) \hat{b}_i(s+l) - x^0]^2 + \\ + \hat{D}_0(s+l) + \sum_{i=1}^q u^T(s+l-i) \hat{D}_i(s+l) u(s+l-i) + \\ + u^T(s+l-1) \Upsilon u^T(s+l-1), \quad (13)$$

где $\hat{b}_i(s+l), \hat{D}_i(s+l), i=0, 1, \dots, q$, — упрощенные обозначения соответственно среднего $b_i(s+l|s)$ и ковариации $D_i(s+l|s)$.

С помощью выражений (8) и (13) в Приложении показано, что для каждого $t=1, 2, \dots, N$ нейтральный закон управления может быть найден в явном виде. Для формулировки этого результата введем следующие обозначения и функции:

1. Вектор $\xi^T(t) = [e_2(t), e_3(t+1), \dots, e_{N-t+2}(N)]$,

где

$$e_p(\tau) = \hat{b}_0(\tau) + \sum_{i=p}^{\tau} u^T(\tau-i) \hat{b}_i(\tau) - x^0, \quad \tau=t, \dots, N, \quad p=2, \dots, N-t+2$$

и матрицы

$$B^T(t) = [\hat{b}_1(t), \hat{b}_2(t+1), \dots, \hat{b}_{N-t+1}(N)], \quad \Theta(t) = \begin{pmatrix} 1 & 0^T \\ 0 & R(t+1) \end{pmatrix}$$

где 0 — нулевой вектор, $R(t+1)$ — решение разностного уравнения, определяемого приведенным ниже образом (17).

2. Функции $T(t) = B^T(t) \Theta(t)$, $P(t) = \sum_{i=0}^{N-t} \hat{D}_{t+1}(t+i)$,

$$c(t) = \sum_{i=0}^{N-t} [\hat{D}_0(t+i) + \sum_{j=i+2}^q u^T(t+i-j) \hat{D}_j u(t+i-j)].$$

Примем $\hat{b}_i(\tau) \equiv 0$, $\hat{D}_i(\tau) \equiv 0$, если $i > q$.

Основной результат. Оптимальное в смысле (5) управление для каждого $t = 1, \dots, N$ задается равенством

$$\bar{u}(t-1) = -L^{-1}(t) T(t) \xi(t). \quad (14)$$

Цена управления $\bar{u}(t-1)$ равна

$$\bar{V}(t) = \xi^T(t) R(t) \xi(t) + c(t). \quad (15)$$

Матрицы $L(t)$ и $R(t)$ удовлетворяют рекуррентным уравнениям

$$L(t) = \Upsilon + P(t) + B^T(t) \Theta(t) B(t), \quad (16)$$

$$R(t) = \Theta(t) - T^T(t) L^{-1}(t) T(t), \quad (17)$$

с начальными условиями

$$L(N) = \Upsilon + \hat{D}_1(N) + \hat{b}_1(N) \hat{b}_1^T(N), \quad R(N) = 1 - \hat{b}_1^T(N) L^{-1}(N) \hat{b}_1(N).$$

Если параметры $\beta_i(t)$, $i = 1, \dots, q$ наблюдаемые или случайный процесс β независим во времени, то для каждого $t = 1, \dots, N$ выполняется равенство

$$V_t = \bar{V}_t. \quad (18)$$

Задача, аналогичная рассмотренной выше задаче, была решена в [3], где изучалось управление случайно блуждающими параметрами $\beta_i(t)$, $i = 1, \dots, q$ объектом, причем процесс аддитивных возмущений β_0 являлся независимым. В частности, если управляемый объект имеет чистое запаздывание $q = 1$, то вычисление оптимальных управлений заметно упрощается, а именно, они имеют такой же вид, как в [4]

$$\bar{u}(t-1) = -[\Upsilon + \hat{D}_1(t) + \hat{b}_1(t) \hat{b}_1^T(t)]^{-1} \hat{b}_1(t) [\hat{b}_0(t) - x^0]. \quad (19)$$

3. Обсуждение результатов

Согласно результатам, приведенным в разделе 2, нейтральная стратегия управления вычисляется следующим образом. На основе накопленной к текущему моменту времени t информации $y(1), \dots, y(t-1)$ и (или) начальных условий \bar{b} и Ξ прогнозируются изменения среднего $\hat{b}(t)$ и ковариации $\hat{D}(t)$ параметров $\beta_i(t)$, $i = 1, \dots, q$ и аддитивного возмущения $\beta_0(t)$ на весь оставшийся период управления $t, t+1, \dots, N$. Затем из уравнений (16), (17) определяются матрицы $L(t)$ и $R(t)$. При помощи формулы (14) вычисляется управление $\bar{u}(t-1)$. Реализуется данное управление. С поступлением нового наблюдения $y(t)$ весь цикл вычислений повторяется.

В том случае, если параметры $\beta_i(t)$, $i = 1, \dots, q$ управляемого объекта известны, т. е. мы имеем объект с (не)коррелированными аддитивными возмущениями β_0 , или если последовательность β во времени независима, то управление (14) является оптимальным. В остальных случаях при любом $t = 1, 2, \dots, N$

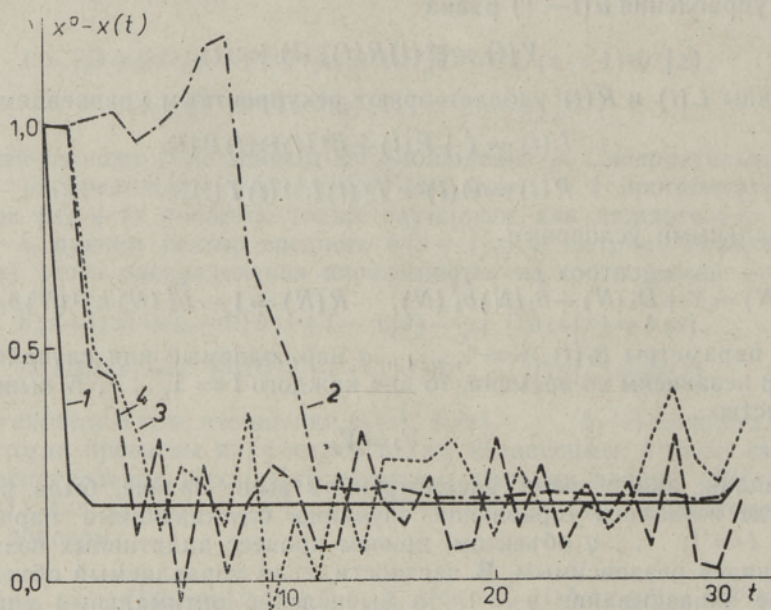
$$M[x(t) - x^0]^2 < \infty.$$

Это следует из условий задачи.

4. Примеры

На рисунке представлена временная зависимость ошибок регулирования $x^0 - x(t)$, $t = 1, \dots, N$ в задачах имитации управления (14) динамическими объектами, параметрами которых являются:

- известные постоянные (см. кривую 1 и задачу 1),
- неизвестные постоянные (см. кривую 2 и задачу 2),
- изменяющиеся во времени величины (см. кривые 3, 4 и задачу 3).



Как видно из рисунка, в задаче 2 ошибка регулирования стремится к оптимальному значению (ср. кривые 1 и 2) по мере уточнения неизвестных параметров, в то время как в задаче 3 алгоритм управления обеспечивает только ограниченность дисперсии ошибок регулирования, поскольку невозможно добиться полного уточнения случайных параметров. Эти результаты получены при решении следующих задач.

Задача 1. Предположим, что процесс задан уравнением

$$x(t) = 0,50u(t-1) + 0,15u(t-2) + 0,05u(t-3) + 0,01u(t-4),$$

и задачу управления решаем при $x^0 = 1$, $\Upsilon = 0$, $N = 31$.

Задача 2. В отличие от задачи 1 допустим, что параметры процесса — неизвестные постоянные, относительно которых имеется следующая априорная информация:

$$\bar{b} = 10^{-2}(1, 1, 1, 1)^T, \quad \Xi = \begin{pmatrix} 0 & 0 \dots 0 \\ 0 & 1 \\ \vdots & \\ 0 \end{pmatrix}.$$

Среднеквадратичное отклонение \sqrt{r} ошибок наблюдения равно 0,05.

Задача 3. Решаем такую же задачу, как и ранее для системы (1)–(3) со следующими данными

$$\mu = (0, 0,50, 0,15, 0,05, 0,01)^T \quad \bar{\beta} = (0, \dots, 0)^T,$$

$$\bar{b} = \mu, \quad D = \Xi = rI, \quad \sqrt{r} = 0,03.$$

Задача 3 решалась для двух случаев:

$$\Phi=0,7 \left(\begin{array}{c|ccc} -1 & 0 & \dots & 0 \\ 0 & & & \\ \vdots & & & \\ 0 & & & \end{array} \middle| \begin{array}{c} I \\ \\ \\ \end{array} \right) \text{ (кривая 3)}, \quad \Phi=0,7 \left(\begin{array}{c|cccc} -1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ \hdashline 0 & 0 & 0 & 0 & 0 \end{array} \right) \text{ (кривая 4)}.$$

В наших экспериментах последовательность управлений $(\bar{u}_0, \dots, \bar{u}_{30})$ вычислялась на мини ЭВМ ЕС-1010 в течение 200—300 с.

Приложение

Доказательство утверждений (14)—(18). Вводя обозначения $\tilde{\xi}^T(t+1) = [e_3(t+1), e_4(t+2), \dots, e_{N-t+2}(N)]$, $\bar{B}^T(t+1) = [\hat{b}_2(t+1), \hat{b}_3(t+2), \dots, \hat{b}_{N-t+1}(N)]$ легко вывести

$$\xi(t+1) = \tilde{\xi}(t+1) + \bar{B}(t+1)u(t-1), \quad (\text{П.1})$$

$$\xi^T(t) = [e_2(t), \tilde{\xi}^T(t+1)], \quad (\text{П.2})$$

$$B^T(t) = [\hat{b}_1(t), \bar{B}^T(t+1)]. \quad (\text{П.3})$$

Очевидно, выражение (8) можно представить следующим образом:

$$\bar{V}(t) = \min_{u(t-1)} \{ \langle W(t), \pi(t|t-1) \rangle + \bar{V}(t+1) \}, \quad (\text{П.4})$$

где, согласно (13),

$$\begin{aligned} \langle W(t), \pi(t|t-1) \rangle &= [\hat{b}_0(t) + \sum_{i=1}^q u^T(t-i) \hat{b}_i(t) - x^0]^2 + \hat{D}_0(t) + \\ &+ \sum_{i=1}^q u^T(t-i) \hat{D}_i(t) u(t-i) + u^T(t-1) \Upsilon u(t-1). \end{aligned}$$

Предположим, что цена управления имеет вид

$$\bar{V}(t) = \xi^T(t) \bar{R}(t) \xi(t) + c(t), \quad (\text{П.5})$$

и выясним, может ли $\bar{V}(t)$ быть решением уравнения (П.4). Если $\bar{V}(t)$ определяется по формуле (П.5), то с помощью соотношений (П1)—(П.3) можно убедиться в том, что

$$\begin{aligned} \langle W(t), \pi(t|t-1) \rangle + \bar{V}(t+1) &= \\ &= u^T(t-1) L(t) u(t-1) + 2u^T(t-1) T(t) \xi(t) + c(t). \end{aligned} \quad (\text{П.6})$$

Вычислив градиент полученного выражения по $u(t-1)$ и приравняв его к нулю, получим, что (П.6) достигает минимума при выполнении (14) и его минимальное значение удовлетворяет (15). Далее, если параметры $\beta_i(t)$ для всех $i=1, \dots, q$, $t=0, 1, \dots, N$ детерминированы или параметры $\beta_j(t)$, $j=0, \dots, q$ чисто случайные так, что $\hat{D}_i(t) = 0$ или $\hat{D}_j(t) = D$ соответственно, то благодаря независимости $\hat{D}_0(t)$ от управлений методом индукции можно показать, что имеет место (18).

ЛИТЕРАТУРА

1. Фельдбум А. А. Основы теории оптимальных автоматических систем. М., «Наука», 1966.
2. Липцер Р. Ш., Ширяев А. Н. Статистика случайных процессов. М., «Наука», 1974.

3. Тенно Р. А., Ойт Х. Э. В кн.: Матер. Всесоюз. конф. «Математическое моделирование сложных химико-технологических систем». Таллин, 1982, 107—109.
4. Wittenmark, B. Intern. J. Control, **21**, 705—730 (1975).

Институт кибернетики
Академии наук Эстонской ССР

Поступила в редакцию
12/IV 1983

R. TENNO, H. OIT

MITTETEADAOLEVATE PARAMEETRITEGA LINEAARSE DÜNAAMLISE OBJEKTI ADAPTIIVNE JUHTIMINE

On leitud mitteteadaolevate parameetritega ja aditiivsete müradega jälgitava olekuga lineaarse dünaamilise objekti juhtimise algoritm.

R. TENNO, H. OIT

ADAPTIVE CONTROL OF LINEAR SYSTEM WITH UNKNOWN PARAMETERS

It is assumed that the output of the linear process

$$x(t) = u^T(t-1)\beta_1(t) + \dots + u^T(t-q)\beta_q(t) + \beta_0(t) \quad (1)$$

is observable by additive «white» noise and the unknown parameters $\beta_i(t)$, $i=0, 1, \dots, q$ are either constants or autoregressive processes. An adaptive control strategy $\{u(t), t=0, 1, \dots, N-1\}$ for the process (1) in case of the quadratic performance index has been found. The proposed control law is based on the technique of separating the functions of parameter estimation and control. The mean and covariance of the unknown parameters are estimated by the usual filtering and prediction methods. A special backward difference equations for calculating the controls has been obtained.

Relationship between optimal and proposed control laws have been discussed. If optimal control law can be obtained in explicit form, then it is shown that proposed control is optimal.

Results on the control of the simulated processes are included.