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

1989, 38, 1

https://doi.org/10.3176/phys.math.1989.1.11

УДК 517.977.5

Р. ТЕННО

ЗАКОН ДУАЛЬНОГО УПРАВЛЕНИЯ ЛИНЕЙНЫМ ОБЪЕКТОМ С НЕИЗВЕСТНЫМИ ПАРАМЕТРАМИ

(Представил Н. Алумяэ)

Несмотря на усилия многих специалистов по теории оптимального и адаптивного управления [1], в течение многих лет не удалось получить оптимального закона дуального управления. Сложилось представление, что дуальное управление выражается в очень сложной форме и что оно зависит от всех достаточных статистик неизвестных параметров, благодаря которым обладает свойством пробных сигналов.

Нами в данной статье показано, что существует такая задача управления с непрерывным процессом, в которой оптимальный закон дуального управления выражается в простой форме и что этот закон зависит не от всех достаточных статистик неизвестных параметров, но тем не менее обладает свойством пробных сигналов.

Задача управления решена в предположениях сильной разрешимости уравнений фильтраций и гладкости решения уравнения Беллмана.

1. Постановка задачи

Ставится задача оптимального управления — минимизация квадратичного функционала

$$v^{\alpha}=M\left\{\int_{0}^{T}(\Theta_t-\lambda)^2\,\mathrm{dt}+(\Theta_T-\lambda)^2\right\}\qquad(1)$$

относительно линейного с неизвестными параметрами управляемого процесса

$$\Theta_t=\alpha_t^{\mathrm{T}}\beta,\qquad(2)$$

наблюдаемого с «погрешностью» согласно дифференциальному уравнению

$$d\xi=[A(t,\xi)+\Theta_t]\,\mathrm{dt}+dV,\qquad \xi_0=\xi_0.\qquad(3)$$

Здесь α_t — вектор управлений, λ — задающее воздействие, β — вектор неизвестных параметров, $A(t,\cdot)$ — функция, удовлетворяющая условиям линейного роста и Липшица, (V_t) — винеровский процесс, ξ_0 — фиксированное начальное значение.

На физическом уровне Θ_t обозначает выход управляемого объекта (2), ξ_t — выход измерительного прибора (3).

Допустимые управления заданы всеми $F_t=\sigma(\xi_t, 0\leqslant s\leqslant t)$-измеримыми при каждом $t\in[0,T]$ функциями со значениями в k-мерном евклидовом пространстве такими, которые при всех непрерывных траекториях (ξ_t) удовлетворяют условиям слабой разрешимости уравнения (3):

$$P\left\{\int_0^T |A(t,\xi)+\alpha_t^{\mathrm{T}}\beta|^2\,\mathrm{d}t<\infty\right\}=1, \tag{4}$$

$$M\exp\int_0^T |A(t,\xi)+\alpha_t^{\mathrm{T}}\beta|^2\,\mathrm{d}t<\infty. \tag{5}$$

В этих условиях стоимость управления v^α корректно определена.

Простейшим примером задачи (1)—(3) является оптимальная стабилизация (на уровне λ) выходного напряжения Θ_t делительного устройства с неизвестным коэффициентом передачи β, выход которого наблюдается интегральным (через «шумящий» интегратор) уравнением:

$$\xi_t=\xi_0+\int_0^t\Theta_s\,ds+V_t.$$

2. Оптимальное «разделенное» управление

В том случае, когда заданы m_0 — априорная гауссова оценка вектора параметров β и $\varkappa_0$ — ее ковариация, поставленную выше задачу можно свести [2] к полным данным, т. е. к задаче минимизации функционала

$$v^\alpha=M\left\{\int_0^T[(\alpha^{\mathrm{T}}m-\lambda)^2+\alpha^{\mathrm{T}}\varkappa\alpha]\,\mathrm{d}t+(\alpha_T^{\mathrm{T}}m_T-\lambda)^2+\alpha_T^{\mathrm{T}}\varkappa_T\alpha_T\right\} \tag{6}$$

относительно фильтрационного процесса

$$dm=\varkappa\alpha\,d\overline{V},\qquad \dot{\varkappa}=-\varkappa\alpha\alpha^{\mathrm{T}}\varkappa, \tag{7}$$

где $m=M(\beta/F_t)$ — условное среднее, $\varkappa=\operatorname{cov}(\beta/F_t)$ — условная ковариация, $(\overline{V}_t)$ — винеровский процесс с дифференциалом

$$d\overline{V}=d\xi-[A(t,\xi)+\alpha^{\mathrm{T}}m]\,\mathrm{d}t.$$

Сужением класса допустимых управлений, условием ограниченности

$$|\alpha|<\delta^{-1},\qquad \delta>0, \tag{8}$$

задачу (6), (7) можно уложить в схему Н. В. Крылова [3]. Только в данном случае винеровский процесс $(\overline{V}_t)$ является управляемым и измеримым относительно системы (F_t) более широких σ-алгебр, да и управления разрешается выбирать измеримыми относительно (F_t). Однако оптимальная стоимость

$$v=\inf v^\alpha$$

от такого расширения вероятностного пространства не изменяется. Построим оптимальное «разделенное» управление, используя результаты [3].

Обозначим через $\varkappa_i^{\mathrm{T}}$ $i=1,\ldots,k$ строки, и через $\varkappa_{ij}$, $i,j=1,\ldots,k$ элементы матрицы $\varkappa_t$. Приблизим вырождающийся управляемый процесс $\varkappa_t$ невырождающимся путем добавления в (7) стандартного винеровского процесса

$$d\varkappa_{ij}=-\varkappa_i^{\mathrm{T}}\alpha\alpha^{\mathrm{T}}\varkappa_j\,\mathrm{d}t+\varepsilon dW_{ij}.$$

Если $\varepsilon^2>0$, то почти всюду дифференциальное уравнение Беллмана для задачи (6), (7)

$$\frac{1}{2}\varepsilon^2 \sum_{i,j,l,m=1}^{k} v_{\varkappa_{ij}\varkappa_{lm}}+\inf_{\alpha}\Big\{\frac{1}{2}\sum_{i,j=1}^{k}\alpha^{\tau}\varkappa_i\varkappa_j^{\tau}\alpha v_{m_i m_j}-$$
$$-\sum_{i,j=1}^{k}\alpha^{\tau}\varkappa_i\varkappa_j^{\tau}\alpha v_{\varkappa_{ij}}+(\alpha^{\tau}m-\lambda)^2+\alpha^{\tau}\varkappa\alpha\Big\}=0, \tag{9}$$

с граничным условием

$$v(m_T,\varkappa_T)=\inf_{\alpha}\{(\alpha^{\tau}m-\lambda)^2+\alpha^{\tau}\varkappa\alpha\}, \tag{10}$$

имеет единственное решение [3].

Если матрица

$$L_T=m_T m_T^{\tau}+\varkappa_T,$$

положительно определена, то оптимальное управление в момент $t=T$ выражается из (10) в виде

$$\alpha_T^0=L_T^{-1}m_T\lambda.$$

Аналогично, если матрица

$$L=mm^{\tau}+\varkappa+\sum_{i,j=1}^{k}\varkappa_i\Big(\frac{1}{2}v_{m_i m_j}-v_{\varkappa_{ij}}\Big)\varkappa_j^{\tau}$$

положительно определена, то ε — приближенное оптимальное управление выражается из (9) в виде

$$\alpha_t^{\varepsilon}=L^{-1}m\lambda, \tag{11}$$

где v — решение дифференциального уравнения

$$\lambda^2 m^{\tau}L^{-1}m=\lambda^2+\frac{\varepsilon^2}{2}\sum_{i,j=1}^{k}v_{\varkappa_{ij}\varkappa_{lm}} \tag{12}$$

с граничным условием

$$v(m_T,\varkappa_T)=\frac{\lambda^2}{1+m^{\tau}\varkappa^{-1}m}.$$

Систему уравнений (11), (12) можно переписать в виде

$$\lambda m^{\tau}\alpha_t^{\varepsilon}=\lambda^2+\frac{\varepsilon^2}{2}\sum_{i,j=1}^{k}v_{\varkappa_{ij}\varkappa_{lm}}. \tag{13}$$

Управление

$$\alpha_t^{\varepsilon}=[m^{\tau}m]^{-1}m\Big(\lambda+\frac{\varepsilon^2}{2\lambda}\sum_{i,j=1}^{k}v_{\varkappa_{ij}\varkappa_{lm}}\Big) \tag{14}$$

удовлетворяет (13), оно является единственным решением этого уравнения в классе управлений с минимальной нормой или энергией.

Численное решение уравнения Беллмана методом конечных разностей [4], проведенное нами, показало, что функция $v(m,\varkappa)$ гладкая (имеет все производные, которые входят в уравнение (9)) даже в том случае, если $\varepsilon=0$, а δ — сколь угодно малый параметр. В этом случае управление (14) имеет особенно простой вид

$$\alpha_t^0=[m^{\tau}m]^{-1}m\lambda. \tag{15}$$

Управление (15) от условной ковариации $\varkappa$ не зависит, оно такое, как оптимальное детерминированное управление при заданных параметрах

$$\alpha_t=[\beta^{\tau}\beta]^{-1}\beta\lambda,$$

только в случае (15) неизвестные параметры заменены их оценками.

Несмотря на простоту, управление (15) имеет «богатые» зондирующие свойства.

3. Зондирующие свойства оптимального управления

Построенное «разделенное» управление (15) является решением исходной задачи (1)—(3), если оно выражается через предыдущие наблюдения с помощью измеримого отображения

$$\alpha_t^0 = \varphi(t, \xi_{[0,t]}),$$

другими словами, если управляемое с обратной связью (15) решение системы уравнений (7) сильное (F_t — измеримое при каждом $t \in [0, T]$) [2]. Это выполняется, например, в случае

$$\alpha_t = (m^\tau m + \delta_1)^{-1} m \lambda$$

при любом $\delta_1 > 0$.

Предположим, что система уравнений (7) имеет сильное решение и в случае, когда $\delta_1 = 0$. Обсудим зондирующие свойства замкнутой с обратной связью системы (7) для удобства в одномерном $k=1$ случае

$$dm = \lambda \frac{\varkappa}{m} d\bar{V}, \qquad \dot{\varkappa} = -\lambda^2 \left(\frac{\varkappa}{m}\right)^2. \tag{16}$$

Систему уравнений (16) можно переписать в виде

$$dm = \varkappa_0 \lambda \left\{ \left[1 + \varkappa_0 \int_0^t \left(\frac{\lambda}{m_s}\right)^2 ds \right] m_t \right\}^{-1} d\bar{V}. \tag{17}$$

Из (17) следует, что условная дисперсия

$$\operatorname{var}(m_{t+\Delta t} \mid m_t = m) = \varkappa_0^2 \lambda^2 \left\{ \left[1 + \varkappa_0 \int_0^t \left(\frac{\lambda}{m_s}\right)^2 ds \right] m_t \right\}^{-2} \Delta t + o(\Delta t) \tag{18}$$

случайной величины $m_{t+\Delta t}$ относительно фиксированной $m_t = m$ тем больше, чем меньше фиксированная величина, и разброс тем меньше, чем ближе траектория $(m_s,\ 0 \leqslant s \leqslant t)$ процесса (17) к нулевым значениям. Поэтому управляемый с обратной связью случайный процесс (17) с положительной вероятностью принимает значения, близкие к нулевым. В таком случае произойдет «раскачивание» процесса (2), в следствие которого априорная неопределенность параметра β быстро уменьшается, а сам процесс обучения останавливается $m_t \to \beta$, $\varkappa_t \to 0$. Если же априорная оценка отличается от неизвестных параметров и по знаку, «раскачивание» процесса (2) будет наиболее интенсивным.

На рис. 1 и 2 приведены траектории процессов (m_t), $(\varkappa_t)$, (Θ_t), полученные в двух численных экспериментах.

Относительно дискретного фильтрационного процесса $(m_t, \varkappa_t,\ t = 0, \Delta t, \ldots, N)$ можно доказать следующее утверждение.

Пусть $0 < q \ll 1$. Тогда с положительной вероятностью случайная последовательность (m_t) входит в область Q, где выполняется $m^2 < q\varkappa$, и с вероятностью близкой к единице выходит из нее при первой возможности. Условная дисперсия $\varkappa_t$ при выходе последовательности (m_t) из области Q не менее, чем на q^{-2} раз меньше по сравнению с условной дисперсией последовательности непосредственно до входа в эту область.

В справедливости утверждения легко убедиться вычислив условные дисперсии последовательности (m_t) непосредственно до входа в область Q и после ее выхода из этой области, согласно уравнениям опти-

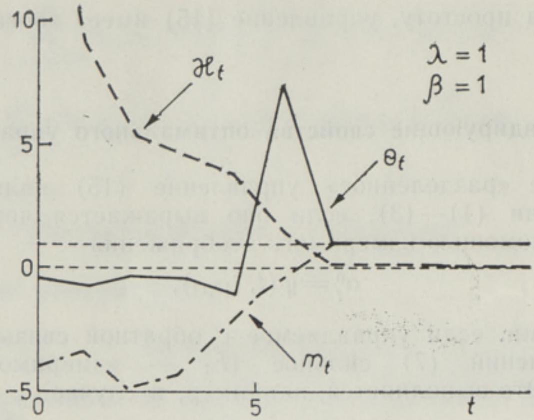

Рис. 1.

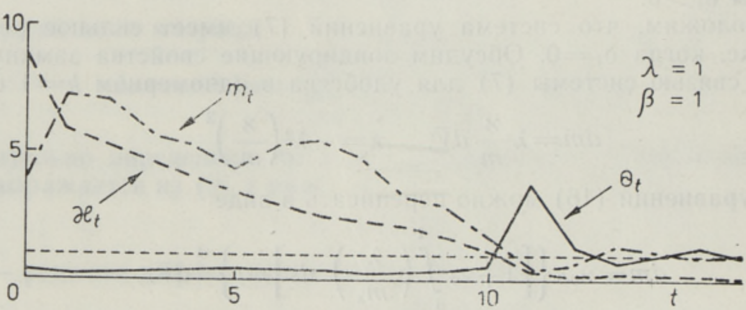

Рис. 2.

мальной нелинейной фильтрации [2].

В частном случае, если $|m_t| > \varepsilon$ для каждого $t \in [0, T]$ ($\varepsilon > 0$ — малый параметр), то стоимость управления (15)

$$v^{\alpha^0} = \lambda^2 M\left(\int_0^T \frac{\varkappa_t}{m_t^2}\,dt + \frac{\varkappa_T}{\varkappa_T + m_T^2}\right) \tag{19}$$

ограничена, а система уравнений (16) имеет единственное сильное решение (условия сильной разрешимости см. в [2]), в силу очевидного неравенства

$$\frac{\varkappa_t}{|m_t|^n} < \frac{\varkappa_0}{\varepsilon^n}$$

для каждого $n = 1, 2, \ldots$. В общем случае легко убедиться в том, что стоимость управления (19) корректно определена ($v^{\alpha^0} < \infty$) при любых траекториях процесса (m_t), начальные значения которых не отличаются от неизвестных параметров по знаку (ограничение на априорную информацию). Действительно, если в марковский момент $\tau \in [0, T]$ случайный процесс принимает значения близкие к нулевым: $|m_\tau| = \varepsilon$, а после этого момента в некоторое время $t \in [\tau, T_1]$ и остается в окрестности нуля: $|m_t| < \varepsilon$, то на этом отрезке времени обновляющий процесс $(\overline{V}_t)$ может быть приближен к детерминированному процессу

$$d\overline{V} = \lambda\left(\frac{\beta}{m_t} - 1\right) dt.$$

Поэтому замкнутая с обратной связью система уравнений (16) аппроксимируется с системой обыкновенных дифференциальных уравнений

$$\dot{m}=\lambda^2 \frac{\varkappa_t}{m_t}\left(\frac{\beta}{m_t}-1\right), \quad m_\tau=\pm\varepsilon, \quad \dot{\varkappa}=-\lambda^2\left(\frac{\varkappa_t}{m_t}\right)^2. \tag{20}$$

Решив систему уравнений (20), получим

$$m_t=\beta-\frac{\varkappa_t}{\varkappa_\tau}(\beta-m_\tau), \quad m_\tau=\pm\varepsilon. \tag{21}$$

Очевидно $\frac{\varkappa_t}{\varkappa_\tau}\leqslant 1$. Если $\operatorname{sgn} m_0=\operatorname{sgn}\beta$, то из (21) следует $|\beta-m_\tau|<|\beta|$. Поэтому $|m_t|>0$ и $\frac{\varkappa_t}{m_t^2}<\infty$ для каждого $t\in[\tau, T_1]$. Следовательно, стоимость управления (19) корректно определена.

ЛИТЕРАТУРА

1. *Asher, R. B., Andrisani, D., Donato, P.* // Proc. IEEE, 1976, **64**, № 8, 1226—1240.
2. *Липцер Р. Ш., Ширяев А. Н.* Статистика случайных процессов. Нелинейная фильтрация и смежные вопросы. М., Наука, 1974.
3. *Крылов Н. В.* Управляемые процессы диффузионного типа. М., Наука, 1977.
4. *Кушнер Г. Дж.* Вероятностные методы аппроксимации в стохастических задачах управления и теории эллиптических уравнений. М., Наука, 1985.

Институт кибернетики
Академии наук Эстонской ССР

Поступила в редакцию
19/V 1987

R. TENNO

TUNDMATUTE PARAMEETRITEGA LINEAARSE OBJEKTI DUAALJUHTIMISE SEADUS

On leitud tundmatute parameetritega, mõõtemüra foonil vaadeldava väljundiga, lineaarse inertsivaba objekti optimaaljuhtimise seadus integraalse ruutkriteeriumi korral.

R. TENNO

AN ACTIVELY ADAPTIVE CONTROL LAW FOR A LINEAR PLANT WITH UNKNOWN PARAMETERS

Consider the following control problem of minimizing quadratic cost criteria

$$v\alpha=M\left\{\int_0^T(\Theta_t-\lambda)^2\,dt+(\Theta_T-\lambda)^2\right\}$$

subject to the linear plant

$$\Theta_t=\alpha_t^{\mathrm{T}}\beta \tag{1}$$

with unknown vector of parameters β. Where λ is a constant reference process, α_t is a feedback control. The observation process is governed by stochastic differential equation

$$d\xi=[A(t,\xi)+\Theta_t]\,dt+dV,$$

where (V_t) is a standard Wiener process and $A(t,\xi)$ is a continuous, bounded, Lipschitz function.

In this paper it is shown that «separated» control law of the form

$$\alpha_t=(m_t^{\mathrm{T}}m_t)^{-1}m_t\lambda \tag{2}$$

is optimal, where m_t is solution of the equations of the nonlinear filter

$$dm=\varkappa_t\alpha_t d\bar{V}, \quad \dot{\varkappa}=-\varkappa_t\alpha_t\alpha_t^{\mathrm{T}}\varkappa_t, \tag{3}$$

$(\bar{V}_t)$ is the innovation process. The control law (2) has special probing and stopping properties of the controlled process (1) and of the estimation process (3).