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

1988, 37, 1

УДК 517.977.58

Р. ТЕННО

АППРОКСИМАЦИЯ ОПТИМАЛЬНЫХ УПРАВЛЕНИЙ БИЛИНЕЙНЫМ ДИФФУЗИОННЫМ ПРОЦЕССОМ

R. TENNO. BILINEAARSE DIFUSIOONIPROTSESSI OPTIMAALJUHTIMISE APROKSIMATSIOON

R. TENNO. APPROXIMATION FOR BILINEAR QUADRATIC STOCHASTIC CONTROL PROBLEM

(Представил Н. Алумяэ)

Предложен вероятностный метод приближенного решения уравнения Беллмана для задачи оптимального управления билинейным диффузионным процессом, наблюдаемым вместе с погрешностями измерений.

1. Постановка задачи

Ставится задача оптимального управления — минимизации квадратичного функционала

$$v^{\alpha}=M\left\{\int_0^T (\Theta_t^2+\alpha_t^{\tau}\Upsilon\alpha_t)\,\mathrm{dt}+\Theta_T^2\right\}, \tag{1}$$

относительно частично наблюдаемого случайного процесса с дифференциалом

$$\begin{aligned} d\Theta&=[g(\alpha_t)\Theta_t+\alpha_t^{\tau}a_0]\,\mathrm{dt}+dW, \quad \Theta_0=\tilde{\Theta},\\ d\xi&=[A(t,\xi)+C\Theta_t]\,\mathrm{dt}+dV, \quad \xi_0=0. \end{aligned} \tag{2}$$

Здесь α — вектор управлений, Υ — положительно определенная матрица, (Θ_t) — ненаблюдаемый одномерный процесс, $g(\alpha)$ — линейная функция

$$g(\alpha)=a_1+\alpha^{\tau}a_2,$$

$\tilde{\Theta}$ — гауссовое начальное значение; (ξ_t) — наблюдаемый векторный процесс, $A(t,\cdot)$ — функция, удовлетворяющая условиям линейного роста и Липшица, (W_t) и (V_t) — независимые винеровские процессы.

Допустимые управления заданы всеми неупреждающими функциями $\alpha(t,\xi)$, такими, при которых система уравнений (2) имеет слабое решение, например, всеми неупреждающими ограниченными функциями. В этих случаях стоимость управления v^{α} корректно определена [1].

2. Оптимальное управление

С помощью уравнений оптимальной нелинейной фильтрации [2] задачу управления по неполным данным (1), (2) можно свести к полным, т. е. к задаче минимизации функционала

$$v^{\alpha}=M\left\{\int_0^T (m_t^2+\gamma_t+\alpha_t^{\tau}\Upsilon\alpha_t)\,\mathrm{dt}+m_T^2+\gamma_T\right\}\tag{3}$$

относительно фильтрационного процесса (m_t, γ_t) с уравнениями

$$\begin{aligned} dm&=[g(\alpha)m_t+\alpha_t^{\tau}a_0]\,\mathrm{dt}+\gamma_t C^{\tau}\,d\overline{W}, \qquad m_0=m_0,\\ \dot{\gamma}&=2g(\alpha)\gamma_t+1-\gamma_t^2 C^{\tau}C, \qquad \gamma_0=\gamma_0,\end{aligned}\tag{4}$$

где $(\overline{W}_t)$ — винеровский процесс с дифференциалом

$$d\overline{W}=d\xi-[A(t,\xi)+Cm_t]\,\mathrm{dt}.$$

Предположим, что существуют все производные, которые входят в уравнения Беллмана для задачи (3), (4), и что это уравнение имеет единственное решение *. Если оптимальное управление существует **, то из уравнения Беллмана нетрудно вывести, что оно задается равенством

$$\alpha_t^*=-\Upsilon^{-1}L(v),\tag{5}$$

где

$$L(v)=\frac{1}{2}(a_0+a_2m)\frac{\partial v}{\partial m}+a_2\gamma\frac{\partial v}{\partial\gamma},$$

$v=v(m_t,\gamma_t)$ — решение дифференциального уравнения

$$\frac{\gamma^2}{2}C^{\tau}C\frac{\partial^2 v}{\partial m^2}+a_1m\frac{\partial v}{\partial m}+[2a_1\gamma+1-\gamma^2C^{\tau}C]\frac{\partial v}{\partial\gamma}+m^2+\gamma-L^{\tau}\Upsilon^{-1}L=0\tag{6}$$

с граничным условием

$$v(m_T,\gamma_T)=m_T^2+\gamma_T.$$

Ввиду сложности уравнение (6) не может быть решено аналитическим методом. Численным методом конечных разностей [3] можно найти его приближенное решение в случае одномерного управления. Однако без сглаживания это численное решение не может быть применено для вычисления производных функций $v(m_t,\gamma_t)$, а тем самым и для вычисления управлений (5). Сглаживание требует подробного анализа результатов вычислений в каждом случае. Оценивание вида и параметров функции $v(m_t,\gamma_t)$ по дискретным расчетным (неточным) значениям дело не простое.

Приведем другой, более удобный метод приближенного решения уравнения Беллмана (6) и аппроксимации оптимальных управлений (5).

3. Метод аппроксимации

Пусть $\Delta t>0$ и $T/\Delta t=N$ — целое число. Определим последовательность $(m_t,\ \gamma_t,\ t=0,\ \Delta t,\ \ldots,\ N\Delta t)$ с уравнениями

$$\begin{aligned} m_{t+\Delta t}&=m_t+[g(\alpha)m_t+\alpha_t^{\tau}a_0]\Delta t+\sigma\sqrt{\Delta t}\,\varepsilon_{t+\Delta t},\\ \gamma_{t+\Delta t}&=(1+g(\alpha)\Delta t)^2\gamma_t+\Delta t-\sigma^2\Delta t,\end{aligned}\tag{7}$$

---

* Уравнение Беллмана имеет единственное решение (см. теоремы 7.7.4 и 14.3.5 в [1]) в случае приближения управляемого процесса (4) к невырождающимся.

** В условии выпуклости множества допустимых управлений существует ε-оптимальное управление (см. лемму 1.1.5 в [1]).

где (ε_t) — обновляющая последовательность из независимых нормированных гауссовых величин, σ — усиление фильтра

$$\sigma = \frac{1+g(\alpha)\Delta t}{\sqrt{1+C^{\tau}\gamma C\Delta t}}\gamma C^{\tau},$$

и функционал

$$v_N^{\alpha} = M\left\{\sum_{t=0}^{(N-1)\Delta t}(m_t^2+\gamma_t+\alpha_t^{\tau}\Upsilon\alpha_t)\Delta t + m_{N\Delta t}^2 + \gamma_{N\Delta t}\right\}. \quad (8)$$

Сформулируем задачу управления последовательностью, как задачу определения таких управлений (α_t), которые минимизируют функционал (8) относительно системы (7).

Данная задача является аппроксимационной ввиду следующего. Первый

$$M(m_{t+\Delta t}-m_t/m_t=m) = [g(\alpha)m+\alpha^{\tau}a_0]\Delta t$$

и второй момент

$$\mathrm{var}\,(m_{t+\Delta t}-m_t/m_t=m) = \sigma^2\Delta t$$

условного (гауссового) распределения $P(m_{t+\Delta t}-m_t/m_t=m)$ состоятельны (при $\Delta t\to 0$) относительно соответствующих моментов диффузионного процесса (m_t) с дифференциалом (4). Тем самым ступенчатый процесс, построенный путем интерполяции последовательности (7), сходится к случайному процессу с дифференциалом (4) в слабом смысле, т. е. в смысле мер, а оптимальная стоимость управления последовательностью

$$v_N = \inf v_N^{\alpha}$$

к оптимальной стоимости управления случайным процессом (см. теорему 3.8 в [4])

$$v = \inf v^{\alpha}.$$

Если $N=3$, то задача (7), (8) поддается точному аналитическому решению [5]. Оптимальная стоимость v_3 выражается через начальные условия γ_0, m_0 уравнений фильтраций (7). Если в качестве начальных условий принимать текущие (в зависимости $v_3(m_0,\gamma_0)$), то стоимость $v_3(m_t,\gamma_t)$ аппроксимирует стоимость, оптимальную управлению последовательностью в текущий момент времени

$$v_N(m_t,\gamma_t) = v_3(m_t,\gamma_t)+\mu+\varepsilon(m_t,\gamma_t) \quad (9)$$

с точностью до неизвестного постоянного μ с малой погрешностью ε [6]. А что особенно важно, $v_3(m_t,\gamma_t)$ аппроксимирует стоимость оптимального управления $v(m_t,\gamma_t)$ непрерывным процессом (4) в текущий момент времени или, что то же самое, аппроксимирует решение уравнения Беллмана (6) с некоторыми другими постоянным μ и погрешностью ε.

Оптимальное управление (5) от константы μ не зависит — оно зависит только от производных функции $v(m_t,\gamma_t)$. Для аппроксимации оптимального управления зависимостью типа (9) достаточно гладкости функции $\varepsilon(m_t,\gamma_t)$ по первым производным. Гладкость функции $\varepsilon(m_t,\gamma_t)$ подтверждается нашими результатами численного решения уравнения (6) и гладкостью функции $v_2(m_t,\gamma_t)$ [5].

ЛИТЕРАТУРА

1. *Крылов Н. В.* Управляемые процессы диффузионного типа. М., «Наука», 1977.
2. *Липцер Р. Ш., Ширяев А. Н.* Статистика случайных процессов. Нелинейная фильтрация и смежные вопросы. М., «Наука», 1974.
3. *Кушнер Г. Дж.* Вероятностные методы аппроксимации в стохастических задачах управления и теории эллиптических уравнений. М., «Наука», 1985.
4. *Гихман И. И., Скороход А. В.* Управляемые случайные процессы. Киев, «Наукова думка», 1977.
5. *Тенно Р.* Изв. АН ЭССР. Физ. Матем., 35, № 4, 435—436 (1986).
6. *Тенно Р.* Изв. АН ЭССР. Физ. Матем., 35, № 1, 62—70 (1986).

Институт кибернетики
Академии наук Эстонской ССР

Поступила в редакцию
10/VI 1987