

УДК 517.977.58

Р. ТЕННО

## АППРОКСИМАЦИЯ ОПТИМАЛЬНЫХ УПРАВЛЕНИЙ ДИФFUЗИОННЫМ ПРОЦЕССОМ С НЕЛИНЕЙНОЙ ЗАВИСИМОСТЬЮ ОТ НЕИЗВЕСТНЫХ ПОСТОЯННЫХ ПАРАМЕТРОВ

(Представил Н. Алумяэ)

Если управляемый процесс зависит нелинейно от неизвестных параметров, то построение оптимальных управлений наиболее сложно. Тем не менее можно получить сравнительно простые приближения оптимальных управлений при условии, что задача управления в остальном отношении проста: критерий качества

$$v^{\alpha} = M^{\alpha} \left\{ \int_0^T (\xi_t^2 + \alpha_t^T \Upsilon \alpha_t) dt + \xi_T^2 \right\}$$

— квадратичный,  
 управляемый процесс

$$d\xi = [a(\theta)\xi_t + b^T(\theta)\alpha_t] dt + dW,$$

— линейный,  
 неизвестные параметры  $\theta$  — постоянные (не изменяются во времени). Проанализируем этот случай.

### 1. Постановка задачи

Сформулируем аналогичную с [1, 2] задачу управления. Пусть  $a(\theta)$ ,  $b(\theta)$  измеримые функции и пусть на стандартном вероятностном пространстве с фильтрацией  $(\Omega, F, \{F_t, t \in [0, T]\}, P)$  определены винеровский процесс  $(W_t, F_t)$  и независимый от него случайный вектор  $\theta$  с заданной априори плотностью распределения  $\pi$ . Определим случайный процесс

$$\xi_t = \xi_0 + W_t \quad (1)$$

с фиксированным начальным условием  $\xi_0$  и  $\sigma$ -алгебры

$$F_t^{\xi} = \sigma\{\xi_s, 0 \leq s \leq t\}, \quad F_t^{\xi, \theta} = \sigma\{\theta, \xi_s : 0 \leq s \leq t\}.$$

Назовем стратегией  $F_t^{\xi}$  — согласованный случайный процесс  $(\alpha_t)$  со значениями в  $A$  ( $A$  — евклидово пространство или некоторое его подмножество), если он при всех  $\xi \in \mathcal{R}$ ,  $\theta \in \mathcal{R}^k$  удовлетворяет двум условиям

$$P \left\{ \int_0^T [a(\theta)\xi_t + b^T(\theta)\alpha_t]^2 dt < \infty \right\} = 1, \quad (2)$$

$$M \exp \int_0^T [a(\Theta) \xi_t + b^T(\Theta) a_t]^2 dt < \infty. \quad (3)$$

Множество всех стратегий обозначим через  $U^{\xi}$ .  
В принятых выше условиях (2) и (3) функция

$$z^{\alpha}(t, \Theta) = \exp \left\{ \int_0^t [a(\Theta) \xi_s + b^T(\Theta) a_s] dW_s - \frac{1}{2} \int_0^t [a(\Theta) \xi_s + b^T(\Theta) a_s]^2 ds \right\} \quad (4)$$

корректно определена, а случайный процесс  $\{z^{\alpha}(t, \Theta)\}$  является  $F_{t^{\xi, \Theta}}$  — согласованным мартингалом и, в частности (см. теорему 6.1 в [3]),

$$M z^{\alpha}(T, \Theta) = 1. \quad (5)$$

Тем самым  $P^{\alpha}$  является (новой) вероятностной мерой

$$dP^{\alpha} = z^{\alpha}(T, \Theta) dP,$$

а случайный процесс

$$W_t^{\alpha} = W_t - \int_0^t [a(\Theta) \xi_s + b^T(\Theta) a_s] ds \quad (6)$$

винеровским относительно системы  $(F_{t^{\xi, \Theta}})$  и меры  $P^{\alpha}$  (см. теорему 6.2 в [3]). Подставляя (6) в (1), получим

$$\xi_t = \xi_0 + \int_0^t [a(\Theta) \xi_s + b^T(\Theta) a_s] ds + W_t^{\alpha}. \quad (7)$$

Обозначим через  $M^{\alpha}$  математическое ожидание относительно меры  $P^{\alpha}$ . Сформулируем задачу.

Пусть управляемый случайный процесс  $(\xi_t)$  с неизвестными параметрами  $\Theta$  задан уравнением (7). Из множества  $U^{\xi}$  требуется найти такую стратегию, которая минимизирует функционал

$$v^{\alpha} = M^{\alpha} \left\{ \int_0^T (\xi^2 + a^T \Upsilon a) dt + \xi_T^2 \right\}. \quad (8)$$

Из (5) следует: задача (7), (8) эквивалентна задачи минимизации

$$v^{\alpha} = M \left\{ z^{\alpha}(T, \Theta) \left[ \int_0^T (\xi^2 + a^T \Upsilon a) dt + \xi_T^2 \right] \right\} \quad (9)$$

относительно (1) и (4). С помощью элементарных свойств стохастических интегралов из (4), (9) легко вывести

$$v^{\alpha} = M \left\{ \int_0^T z^{\alpha}(t, \Theta) [\xi^2 + a^T \Upsilon a] dt + z^{\alpha}(T, \Theta) \xi_T^2 \right\}, \quad (10)$$

где

$$z^{\alpha}(t, \Theta) = 1 + \int_0^t z^{\alpha}(s, \Theta) [a(\Theta) \xi_s + b^T(\Theta) a_s] dW. \quad (11)$$

Сформулируем задачу (7), (8) в терминах бесконечномерного процесса.

Пусть управляемое случайное поле  $\{z^{\alpha}(t, \Theta)\}$  задано уравнением (11), а неуправляемый случайный процесс — уравнением (1). Из множества  $U^{\xi}$  требуется найти такую стратегию, которая минимизирует функционал (10).

## 2. Задача управления фильтрационной плотностью

Пусть  $a(\Theta)$  и  $b(\Theta)$  — равномерно ограниченные функции, имеющие ограниченные производные до второго порядка. Тогда в слабых условиях на плотность  $\pi$  уравнение

$$\begin{aligned} dq(t, \Theta) &= q(t, \Theta) [a(\Theta) \xi_t + b^T(\Theta) \alpha_t] dW, \\ q(0, \Theta) &= \pi \end{aligned} \quad (12)$$

имеет единственное решение (см. теорему 4.3.2 в [4]), интерпретируемое как ненормированная фильтрационная плотность. При каждом  $t \in [0, T]$  выполняется равенство (см. теорему 6.2.1 в [4])

$$M\{z^\alpha(t, \Theta) | F_t^\xi\} = \int_{R^k} q(t, \Theta) d\Theta.$$

Тем самым стоимость управления (10) равна

$$v^\alpha = M \left\{ \int_0^T (\xi_t^2 + \alpha^T \Gamma \alpha) \int_{R^k} q(t, \Theta) d\Theta dt + \xi_T^2 \int_{R^k} q(T, \Theta) d\Theta \right\} \quad (13)$$

в терминах фильтрационной плотности. Сформулируем задачу «разделенного» управления.

Пусть фильтрационная плотность задана уравнениями (1) и (12). Из множества  $U^{\xi, q}$  требуется найти такую стратегию, которая минимизирует функционал (13).

## 3. Анализ конечномерной задачи управления фильтрационной плотностью

Примем вектор неизвестных параметров  $\Theta$  принадлежащим конечному множеству возможных значений  $(\Theta_i, i=1, 2, \dots, p+1)$  для упрощения задачи и одномерным для удобства обозначений. Введем конечномерную аппроксимацию задачи (1), (12), (13), т. е. задачу минимизации

$$v_p^\alpha = M \left\{ \int_0^T (\xi_t^2 + \alpha^T \Gamma \alpha) h \sum_{i=1}^p q_i(t) dt + \xi_T^2 h \sum_{i=1}^p q_i(T) \right\} \quad (14)$$

относительно (1) и

$$dq_i(t) = q_i(t) [a_i \xi_t + b_i^T \alpha_t] dW, \quad q_i(0) = \pi_i. \quad (15)$$

Здесь

$$h = \Theta_{i+1} - \Theta_i, \quad a_i = a(\Theta_i), \quad b_i = b(\Theta_i), \quad q_i(t) = q(t, \Theta_i).$$

Задача (14), (15) укладывается в схему Н. В. Крылова. Тем самым общие свойства оптимальной стоимости являются такими же, как и в [5], и, в частности, выполняется принцип Беллмана. Введем дополнительные условия.

Пусть  $\pi_i > 0$  при всех  $i=1, \dots, p$ . Тогда, решив уравнение (15), убедимся, что  $q_i(t) > 0$  при всех  $i=1, \dots, p, t \in [0, T]$ . Поэтому условием ниже гарантируется слабая невырожденность процесса  $\{q_i(t)\}$ .

Пусть для каждого  $t \in [0, T]$ ,  $\xi \in R, \lambda \in R^p, |\lambda| \neq 0$  выполняется

$$\sup_{\alpha \in A} \sum_{i,j=1}^p \lambda_i (a_i \xi + b_i^T \alpha) (a_j \xi + b_j^T \alpha) \lambda_j > 0.$$

Тем самым с обобщенными производными уравнение Беллмана имеет единственное решение (см. теоремы 4.7.7 и 5.3.14 в [5]). Более того.

Пусть  $A$  — выпуклое подмножество евклидова пространства. Тогда 1. Существует  $\varepsilon$  — оптимальная допустимая марковская стратегия «разделенного» управления  $\{\alpha(\xi_t, q_i(t))\} \in U_M^{\xi_q}$ . Следовательно, в классе  $U^{\xi_q}$  существует приближенно оптимальная стратегия такая, при котором уравнение (7) имеет сильное решение. 2. Марковская стратегия является достаточной, т. е. множество марковских стратегий  $U_M^{\xi_q}$  плотно в  $U^{\xi_q}$ :

$$\inf_{\alpha \in U_M^{\xi_q}} v^{\alpha} = \inf_{\alpha \in U^{\xi_q}} v^{\alpha}.$$

Утверждения 1 и 2 являются прямым следствием леммы 5.1.1. и теоремы 5.2.5 из [5].

#### 4. Аппроксимация оптимальных управлений

Введем с дискретным параметром времени  $t=0, \Delta, \dots, n\Delta$  задачу минимизации

$$v_{p,n}^{\alpha} = M \left\{ \sum_{t=0}^{(n-1)\Delta} (\xi^2 + \alpha^T \Upsilon \alpha) h \sum_i q_i(t) \Delta + \xi_T^2 h \sum_i q_i(T) \right\}$$

относительно

$$q_i(t+\Delta) = q_i(t) \exp \left\{ (a_i \xi_t + b_i^T \alpha_t) \sqrt{\Delta} \varepsilon_{t+\Delta} - \frac{1}{2} (a_i \xi_t + b_i^T \alpha_t)^2 \Delta \right\}, \quad (16)$$

$$\xi_{t+\Delta} = \xi_t + \sqrt{\Delta} \varepsilon_{t+\Delta},$$

где  $\varepsilon_{t+\Delta} = W_{t+\Delta} - W_t$  — независимая гауссова величина. Аппроксимация (16) строилась с помощью рабастного представления уравнения фильтрации (12).

Решив двухшаговую задачу ( $n=2$ ), получим следующий результат.

Пусть матрица  $\Upsilon$  положительно определена. Тогда оптимальное управление на первом шаге приближенно задается равенством ( $\xi \equiv \xi_0$ ,  $q_i \equiv \pi_i$ )

$$\alpha^+ = - \sum_i q_i (1 + a_i \Delta) c_i \left[ \sum_i q_i (\Upsilon + c_i b_i b_i^T \Delta) \right]^{-1} b_i \xi, \quad (17)$$

где

$$c_i = \Delta + (1 + a_i \Delta)^2 + \Delta (l^T K^{-1} K_i - 2l_i^T) K^{-1} l,$$

$$l = \sum_i q_i l_i, \quad l_i = (1 + a_i \Delta) b_i,$$

$$K = \sum_i q_i K_i, \quad K_i = \Upsilon + b_i b_i^T \Delta,$$

а на втором шаге — равенством ( $\xi \equiv \xi_{\Delta}$ ,  $q_i \equiv q_i(\Delta)$ )

$$\alpha^* = -K^{-1} L \xi. \quad (18)$$

Данный результат (17), (18) получен с помощью несложных, но громоздких вычислений по алгоритму динамического программирования. При этом выяснилось, что с помощью известных функций (в том числе специальных) не удастся исключить оператор математического ожидания из уравнений, которые определяют оптимальное управление на первом шаге. Поэтому задача несколько упростилась — стоимость оптимального управления на втором шаге приближалась линейной функцией по  $q_i(\Delta)$  в окрестности точки  $Mq_i(\Delta) = \pi_i$ .

В общем случае, когда  $n$  — любое натуральное число, можно пользоваться следующим приближением. Пусть  $\{q_i(t), \xi_i\}$  — управляемая последовательность, определяемая из уравнений фильтрации (16). Примем  $q_i = q_i(t)$ ,  $\xi = \xi_t$  в (17) и  $q_i = q_i(T-\Delta)$  в (18). Тогда стратегия управления, задаваемая равенством (17) при каждом  $t=0, \Delta, \dots, (n-2)\Delta$  и равенством (18) при  $t=T-\Delta$ , аппроксимирует стратегию оптимального управления.

#### ЛИТЕРАТУРА

1. Rishel, R. Proc. 19th IEEE Conf. on Decision and Control incl. Symp. on Adaptive Processes, 1, 337—340 (1981).
2. Rishel, R. Abstr. Int. Conf. on Stochastic Optimization. Kiev, 1984.
3. Липцер Р. Ш., Ширяев А. Н. Статистика случайных процессов. Нелинейная фильтрация и смежные вопросы. М., «Наука», 1974.
4. Розовский Б. Л. Эволюционные стохастические системы. М., «Наука», 1983.
5. Крылов Н. В. Управляемые процессы диффузионного типа. М., «Наука», 1977.

Институт кибернетики  
Академии наук Эстонской ССР

Поступила в редакцию  
20/VI 1986

R. TENNO

#### TUNDMATUTE PARAMEETRITEGA MITTELINEAARSE DIFUSIOONI LÄHISOPTIMAALNE JUHTIMINE

On leitud mittelineaarse seose olevate tundmatute parameetritega lineaarse dünaamilise objekti optimaaljuhtimise seaduse aproksimatsioon ruutkriteeriumi korral.

R. TENNO

#### APPROXIMATION OF OPTIMAL CONTROLS FOR DIFFUSION PROCESS WITH UNKNOWN NONLINEAR PARAMETERS

Consider the following control problem of minimizing quadratic cost criteria

$$v\alpha = M\alpha \left\{ \int_0^T (\xi_t^2 + \alpha_t^T \Upsilon \alpha_t) dt + \xi_T^2 \right\} \quad (1)$$

subject to the linear controlled stochastic differential equation

$$d\xi = [a(\theta)\xi + b^T(\theta)\alpha] dt + dW \quad (2)$$

in which the coefficients  $a(\theta)$  and  $b(\theta)$  depend on a vector  $\theta$  of unknown parameters with given prior probability density. In (2)  $(W_t)$  is a standard Wiener process,  $\alpha_t$  is a feedback control based on the observations.

The original control problem with unknown parameters is reformulated as a control problem with complete observations of infinite dimensional stochastic system. This system corresponds to the evolution of the unnormalized conditional probability and is covered by simple Zakai equation.

The approximate solution to the original control problem is obtained by the use of the corresponding finite dimensional discrete time Zakai equation.