XII LIETUVOS MATEMATIKOS RINKINYS 2
ЛИТОВСКИЙ МАТЕМАТИЧЕСКИЙ СБОРНИК

1972

УДК 512.25 519.3:30.115

НЕПРЕРЫВНЫЙ АНАЛОГ ПРОЦЕССА РАСПРЕДЕЛЕНИЯ ДИНАМИЧЕСКОГО ПРОГРАММИРОВАНИЯ

П. Руткаускас, В. Бистрицкас

В настоящей работе исследуется непрерывный аналог функционального уравнения

$$f_{N+1}(x)=\max\left\{\begin{array}{ll} A: & g(x)+f_{N}(ax) \\ B: & h(x)+f_{N}(bx)\end{array}\right\}, \quad N\geqslant 0 \tag{1}$$

при бесконечном N. Предполагается, что параметры a, b, x удовлетворяют неравенствам

$$0\leqslant a,\ b<1, \quad x\geqslant 0$$

функций $g(x)$ и $h(x)$, имеют непрерывные производные и

$$\sum_{N=0}^{\infty} \max_{|x|\leqslant q^N y} [\,|h(x)|,\ \ |g(x)|\,]<\infty,$$

где

$$q=\max(a,\ b).$$

Р. Беллманом (см. [1], стр. 45) для бесконечного N, когда

$$h(x)=cx^d, \quad g(x)=ex^g, \quad c,\ d,\ e,\ g\geqslant 0,$$

и Ю. Багвелом и др. (см. [4]) для конечного N, когда g, $d\geqslant 1$, найдены решения уравнения (1). Решение упомянутого уравнения при первоначальных предположениях о функциях $g(x)$ и $h(x)$ становится довольно сложной задачей. В этом случае структуру решения $f(x)$ исследуем при помощи непрерывного аналога.

§ 1. Дискретный процесс управления

Предположим, что состояние системы S в любой момент времени $t=k\Delta$, $k=0, 1, \ldots$, описывается величиной $x(t)$, $t=0, \Delta, 2\Delta, \ldots$ С течением времени эта система управляется одним из двух допустимых управлений A и B. Используя управление A в интервале $(k\Delta, (k+\delta)\Delta)$, получаем доход $g(x(k\Delta))\,\delta$ и система переходит в состояние $a^\delta x(k\Delta)$. Аналогично приписываются доход $h(x(k\Delta))\,\delta$ и состояние $b^\delta x(k\Delta)$ управлению B в интервале времени $(k\Delta, (k+\delta)\Delta)$.

Предположим, что в каждый момент времени $k\Delta$, $k=0,1, \ldots, N$ мы принимаем решение, какую долю следующего интервала длиной Δ мы отводим соответственно для управлений A и B. Это определяет выбор доли u $(k\Delta)$, которая означает, что в интервале длиной $\big(k\Delta,\ u(k\Delta)(k+1)\Delta\big)$ используется управление A, а в остальной части интервала – управление B.

Задача состоит в определении такой последовательности операций управления $u(k\Delta)$, $k=0, 1, \ldots, N$, которая максимизирует доход.

Обозначим $f_N(z)$ максимальный доход за N шагов при начальном состоянии z, когда во всем интервале длиной Δ используется только одно из допустимых управлений.

Согласно принципу динамического программирования

$$f_N(z)=\max\begin{Bmatrix} g(z)\Delta+f_{N-1}(a^\Delta z)\\ h(z)\Delta+f_{N-1}(b^\Delta z)\end{Bmatrix},\quad N\geqslant 1, \tag{2}$$

где

$$f_0(z)=0,\quad 0\leqslant a,\ b<1.$$

Предельным переходом $N\to\infty$ получаем функциональное уравнение

$$f(z)=\max\begin{Bmatrix} g(z)\Delta+f(a^\Delta z)\\ h(z)\Delta+f(b^\Delta z)\end{Bmatrix},$$

которое эквивалентно (1).

Для полученного уравнения существует единственное решение $f(z)$, ибо оно принадлежит к первому типу (см. [1], стр. 145). Нахождение явного решения данного уравнения является довольно сложной задачей. Поэтому, чтобы получить некоторую информацию о структуре решения упомянутого уравнения, мы рассматриваем его непрерывный аналог.

§ 2. Непрерывный аналог

Обозначим $x(t)$ – капитал в момент времени t, а $f(t)$ – доход, полученный в интервале $[0, t]$ при данном $u(t)$, где $t=0, \Delta, 2\Delta, \ldots$

Таким образом, доход

$$f(t+\Delta)=f(t)+u(t)\cdot\Delta\cdot g\big(x(t)\big)+\big(1-u(t)\big)\,\Delta\cdot h\left(a^{(1-u(t))\,\Delta}\,x(t)\right),$$

$$f(0)=0$$

и капитал

$$x(t+\Delta)=x(t)\,[a^u\,b^{(1-u)}]^\Delta,\quad x(0)=x_0.$$

Разлагая правые стороны последних равенств по степеням Δ, получаем

$$f(t+\Delta)=f(t)+u\Delta g\big(x(t)\big)+(1-u)\,\Delta h\big(x(t)\big)+\sigma(\Delta).$$

$$f(0)=0$$

и

$$x(t+\Delta)=x(t)\,[1+u\Delta\ln a+(1-u)\,\Delta\ln b]+\sigma(\Delta).$$

Предельный переход $\Delta \to 0$ дает уравнения:

$$\frac{df(t)}{dt} = ug(x) + (1-u)\, h(x), \quad f(0) = 0, \tag{3}$$

$$\frac{dx(t)}{dt} = x\,[u \ln a + (1-u) \ln b], \quad x(0) = x_0. \tag{4}$$

Итак, получаем задачу оптимального управления: максимизировать функционал $f(T) = f(T, u)$ при условии (4) и $0 \leqslant u(t) \leqslant 1$.

§ 3. Синтез оптимального управления при $T = \infty$

В этом разделе при помощи принципа максимума (см. [3], стр. 23) решается следующая задача оптимального управления.

Требуется найти такое допустимое управление $u(t)$, чтобы функционал

$$I(u) = -\int_0^\infty [ug(x) + (1-u)\, h(x)]\, dt \tag{5}$$

сходился и принимал наименьшее возможное значение, когда фазовая переменная $x(t)$ удовлетворяет уравнению

$$\frac{dx}{dt} = x(t)\Big[u(t) \ln a + \Big(1 - u(t)\Big) \ln b\Big], \quad x(0) = x_0. \tag{6}$$

Допустимым управлением назовем всякую кусочно-непрерывную функцию $u(t)$, удовлетворяющую неравенству

$$0 \leqslant u(t) \leqslant 1 \quad \text{при} \quad 0 \leqslant t < \infty. \tag{7}$$

Покажем, что при допустимом управлении $u(t)$ интеграл (5) сходится. Из дифференциального уравнения (6) имеем, что

$$x(t) = x_0\, e^{\int_0^t \left[u(t) \ln a + \left(1 - u(t)\right) \ln b\right] dt} \tag{8}$$

Нетрудно видеть, что

$$u(t) \ln a + \Big(1 - u(t)\Big) \ln b < 0.$$

Таким образом, функция $x(t)$, убывая, стремится к нулю, т.е.

$$\lim_{t \to \infty} x(t) = 0. \tag{9}$$

Так как функции $g(x(t))$ и $h(x(t))$ непрерывные и $g(0) = h(0) = 0$, то

$$\lim_{t \to \infty} g\Big(x(t)\Big) = \lim_{t \to \infty} h\Big(x(t)\Big) = 0.$$

Следовательно, существует такое T, при котором функция

$$p\Big(x(t)\Big) = \max\Big[\,\Big|h\Big(x(t)\Big)\Big|, \ \Big|g\Big(x(t)\Big)\Big|\,\Big]$$

не возрастает при $t>T$. Таким образом,

$$\int_T^\infty p\left(x(t)\right)dt=\int_T^\infty \left(p\left(x_0\, e^{\int_0^t [u\ln a+(1-u)\ln b]\,dt}\right)\right)dt\leqslant$$

$$\leqslant \int_T^\infty p(x_0 q^t)\,dt\leqslant \sum_{N=T}^\infty p(x_0 q^N)<\infty.$$

Теперь, ввиду соотношения (9), имеем задачу оптимального управления: среди всех допустимых управлений $0\leqslant u(t)\leqslant 1$ найти такое, которое точку x_0 переводит к точке 0 и функционал (5) принимает наименьшее значение. Далее при решении упомянутой задачи применяем принцип максимума (см. [3], стр. 209).

Для рассматриваемой задачи составляем функцию Гамильтона $H(x, \psi, u)$ (см. [3], стр. 24). Имеем, что

$$H(x,\ \psi,\ u)=\psi_0[uh(x)-ug(x)-h(x)]+\psi_1 x[u\ln a+\ln b-u\ln b]. \tag{10}$$

Отсюда получаем

$$H\left(x(t),\ \psi(t),\ u(t)\right)=A(t)u(t)+B(t), \tag{11}$$

где

$$A(t)=-\psi_0 g\left(x(t)\right)+\psi_0 h\left(x(t)\right)+\psi_1(t)\ x(t)\,(\ln a-\ln b), \tag{12}$$

$$B(t)=-\psi_0 h\left(x(t)\right)+\psi_1(t)\ x(t)\ \ln b. \tag{13}$$

Здесь входящая функция $\psi_1(t)$ удовлетворяет дифференциальному уравнению

$$\frac{d\psi_1}{dt}=-\frac{\partial H}{\partial x}=\psi_0 u(t)\,g'_x\left(x(t)\right)+\left(1-u(t)\right)\psi_0 h'_x\left(x(t)\right)-$$
$$-\psi_1(u\ln a+\ln b-u\ln b), \tag{14}$$

а

$$\psi_0=c_1\leqslant 0. \tag{15}$$

Обозначим через $\bar u(t)$ управление, удовлетворяющее принципу максимума и предположим, что

$$\ln a\, h(x)\ (\neq)\ \ln b\ g(x) \tag{16}$$

(знак „$\neq$“ означает, что неравенство выполняется почти всюду).

При условии (16) докажем несколько лемм.

Лемма 1. *Если удовлетворен принцип максимума, то функция* $A(t)$ $(\neq)0$ *для* $0\leqslant t\leqslant T$, *где*

$$A(t)=c_1\,h(x)-c_1 g(x)+\psi_1 x\,(\ln a-\ln b),\quad T>0. \tag{17}$$

Доказательство. Допустим противное, что

$$A(t)=0\quad \text{для}\quad t_0\leqslant t\leqslant t_0+\varepsilon,\ (\varepsilon>0). \tag{18}$$

Из принципа максимума следует, что

$$B(t)=-c_1\,h(x)+\psi_1 x\ \ln b=0.$$

Отсюда вытекает, что

$$\psi_1(t)=\frac{c_1 h\big(x(t)\big)}{x(t)\ln b} \quad \text{для} \quad t_0 \leqslant t \leqslant t_0+\varepsilon, \quad (\varepsilon>0). \tag{19}$$

Подставляя последнее выражение в соотношение (18), получаем, что

$$c_1\Big[\ln a\, h\big(x(t)\big)-\ln b\, g\big(x(t)\big)\Big]=0 \quad \text{для} \quad t_0 \leqslant t \leqslant t_0+\varepsilon, \ (\varepsilon>0). \tag{20}$$

В таком случае имеет место хотя бы одно из равенств $c_1=0$ и

$$h\big(x(t)\big)\ln a=\ln b\, g\big(x(t)\big), \quad t_0 \leqslant t \leqslant t_0+\varepsilon \ (\varepsilon>0). \tag{21}$$

Если $c_1=0$, то из формулы (19) имеем, что

$$\psi_1(t)=0, \quad t_0 \leqslant t \leqslant t_0+\varepsilon \ (\varepsilon>0).$$

Так как функция $\psi_1(t)$ удовлетворяет линейному дифференциальному уравнению, то из (14) ясно, что $\psi_1(t)=0$ для всех $t \geqslant 0$. Это противоречит предположению, что вектор $\big(\psi_0, \psi_1(t)\big)$ удовлетворяет принципу максимума. Следовательно, имеет место равенство (21), но это противоречит предположению (16). Итак, лемма доказана.

Лемма 2. *Управление $\bar u(t)$, $0 \leqslant t \leqslant T$ равно единице тогда и только тогда, когда $A(t) \geqslant 0$, $0 \leqslant t \leqslant T$ и управление $\bar u(t)$ равно нулю, $0 \leqslant t \leqslant T$, тогда и только тогда, когда $A(t) \leqslant 0$, $0 \leqslant t \leqslant T$, $T>0$.*

Достаточность. Согласно (11)

$$H(x, \psi, u)=A(t)u+B(t).$$

Предположив $A(t) \geqslant 0$, когда $0 \leqslant t \leqslant T$, докажем, что $\bar u(t)=1, 0 \leqslant t \leqslant T$. В силу принципа максимума (см. [3], стр. 209) функция $H(x, \psi, u)$ достигает максимального значения при управлении $\bar u(t)$. Так как $H(x, \psi, u)$ является линейной по u, то управление $\bar u(t)=1$, когда $A(t)>0$. Ввиду леммы 1 функция $A(t)\ (\neq)0$, когда $0 \leqslant t \leqslant T$. Из кусочно-непрерывности функции $u(t)$ следует утверждение.

Аналогично получаем, что

$$\bar u(t)=0 \quad \text{для} \quad A(t) \leqslant 0, \quad 0 \leqslant t \leqslant T.$$

Необходимость. Если $\bar u(t)=1$, $0 \leqslant t \leqslant T$, то из доказанного имеем, что $A(t) \geqslant 0$, ибо $A(t)$ непрерывная функция. Очевидно, что

$$A(t) \leqslant 0, \quad \text{когда} \quad \bar u(t)=0, \quad 0 \leqslant t \leqslant T.$$

Лемма доказана.

Лемма 3. *Если управление $\bar u(t)=1$, $0 \leqslant t \leqslant T$, то функция*

$$\psi_1(t)=\bar\psi_1(t)=\frac{c_1 g\big(x(t)\big)}{x(t)\ln a} \tag{22}$$

удовлетворяет принципу максимума и $x(t)=x_0 a^t$.

Если управление $\bar u(t)=0$, $0 \leqslant t \leqslant T$, то функция

$$\psi_1(t)=\bar\psi_1(t)=\frac{c_1 h\big(x(t)\big)}{x(t)\ln b} \tag{23}$$

удовлетворяет принципу максимума и $x(t)=x_0 b^t$, где $c_1<0$.

Доказательство. В силу принципа максимума

$$H\left(x(t),\ \psi(t),\ u(t)\right)=0,\quad \text{когда}\quad 0\leqslant t\leqslant T.$$

Но, ввиду обозначения (11),

$$H(x,\ \psi,\ \bar{u})=A(t)\,\bar{u}(t)+B(t)=0,\quad \text{когда}\quad 0\leqslant t\leqslant T.$$

Приравнивая управление $\bar{u}$ (t) единице и подставляя выражения A (t) и B (t), имеем:

$$-c_1\,g(x)+c_1\,h(x)+\psi_1\,x\,(\ln a-\ln b)+\psi_1\,x\ln b-c_1\,h(x)=$$
$$=-c_1\,g(x)+\psi_1\,x\ln a=0.$$

Отсюда,

$$\psi_1(t)=\frac{c_1\,g\left(x(t)\right)}{x(t)\ln a},\tag{24}$$

где

$$\mathrm{x}\,(t)=\mathrm{x}_0a^t,\quad 0\leqslant t\leqslant T.$$

Остается показать, что функция $\psi_1(t)$ удовлетворяет дифференциальному уравнению (14). Дифференцируя обе стороны равенства (24) по t, получаем:

$$\frac{d\psi_1}{dt}=\frac{c_1}{\ln a}\ \frac{g'_x(x)\cdot x'_t\cdot x-x'_t\,g(x)}{x^2}=c_1\,g'_x(x)-\psi_1(t).$$

Случай $\bar{u}$ $(t)=0$, $0\leqslant t\leqslant T$, исследуется аналогично.

Теорема. а) *Пусть уравнение*

$$\varphi(x)=h(x)\ln a-g(x)\ln b=0\tag{25}$$

имеет такие решения x_i, $i=1,2,\ \ldots,\ n$ $(n\geqslant 1)$, *что*

1) $0<x_1<X_2<\ \ldots\ <x_n<\bar{X}$;

2) *существуют такие окрестности точек* x_i, *в которых функция*

$$\eta(x)=\frac{d}{dx}\,[h(x)\ln a-g(x)\ln b]\tag{26}$$

имеет постоянные знаки. Тогда точки x_i, $i=1,\ \ldots,\ n$ *являются точками переключения синтеза оптимального управления* u (x), *который принимает значение либо нуля, либо единицы. Если функция* η $(x_1-\varepsilon)<0$ *при достаточно малом* $\varepsilon>0$, *то синтез* $u(x)=1$, *когда* $0\leqslant x\leqslant x_1$ *и, если* η $(x_1-\varepsilon)>0$, *то синтез* u $(x)=0$, *когда* $0\leqslant x\leqslant x_1$.

б) *Пусть уравнение* (25) *в интервале* $[0,\ \bar{x}]$ *не имеет решений, удовлетворяющих условиям* 1) *и* 2). *Тогда, если существует такая точка* $x_0\in$ $\in[0,\ \bar{X}]$, *что* φ $(x_0)<0$, *то синтез*

$$u(x)=0,\quad \textit{когда}\quad 0\leqslant x\leqslant\bar{X},$$

и в противном случае синтез

$$u(x)=1,\quad \textit{когда}\quad 0\leqslant x\leqslant\bar{X}.$$

Доказательство. Докажем утверждение теоремы при условии (16). Сначала покажем, что в случае а) синтез управления u (x) удовлетворяет принципу максимума и условиям регулярности (см. [3], стр. 263). В качестве вектора-функции $\psi(t)$ выбираем вектор $(c_1, \psi_1(t))$. Так как точки переключения управления $u(x(t))$ удовлетворяют уравнению (25), то в силу леммы 3 $\psi(t)$ является непрерывной, ибо

$$\bar{\psi}_1(t_i)=\bar{\bar{\psi}}_1(t_i) \iff h\left(x(t_i)\right)\ln a=\ln b\, g\left(x(t_i)\right).$$

Из предположения $\eta\left(x(t_1)-\varepsilon\right)<0$ следует, что

$$\varphi(x)\geqslant 0, \quad \textit{когда} \quad 0\leqslant x\leqslant x_1.$$

Приравнивая управление $u(t)$ единице, из соотношения (22) получаем, что

$$\psi_1(t)=\frac{c_1 g\left(x(t)\right)}{x(t)\ln a}.$$

Полученное выражение $\psi_1(t)$ подставляем в формулу (17). Имеем, что

$$A(t)=c_1\left[h\left(x(t)\right)\ln a-g\left(x(t)\right)\ln b\right]=c_1\,\varphi\left(x(t)\right)\geqslant 0, \tag{27}$$

когда $t\geqslant t_1$. Теперь лемма 2 утверждает, что управление $u(t)=1$, $t\geqslant t_1$ удовлетворяет принципу максимума. Итак, синтез $u(x)=1$, $0\leqslant x\leqslant x_1$, удовлетворяет принципу максимума. Так как синтез управления $u(x)$ регулярен, то он является и оптимальным (см. [3], стр. 359). Аналогично доказывается оптимальность синтеза управления $u(x)$ в интервалах (x_i, x_{i+1}), $i=1, \ldots, n-1$.

Случай $\eta(x_1-\varepsilon)>0$ исследуется аналогично.

Далее докажем утверждение б) при условии (16). Так как условие 2) имеет место только в конечное число точек интервала $[0, \bar{X}]$, то в этом случае функция $\varphi(x)$ не меняет знака в интервале $[0, \bar{X}]$. Это, как в предыдущем случае, означает, что функция $A(t)$ не меняет знака для всех $t\geqslant 0$. Из леммы 2 получаем утверждение теоремы.

Остается исследовать случай, когда не удовлетворено условие (16), т.е. существуют такие числа

$$a_i,\ b_i\geqslant 0,\ i=1, \ldots, s, \ \text{что}\ a_i<b_i$$

и

$$h(x)\ln a-g(x)\ln b=0 \tag{28}$$

для всех $x\in[a_i, b_i]$, $i=1, \ldots, s$. Покажем, что в этих интервалах любой допустимый синтез управления является оптимальным. Выбираем интервал $[a_i, b_i]$ и исследуем синтез оптимального управления. Получаем задачу оптимального управления:

максимизировать функционал

$$I=\int_0^{\tau}\left[u\,g\left(x(t)\right)+(1-u)\,h\left(x(t)\right)\right]dt,$$

когда фазовая переменная удовлетворяет уравнению

$$\frac{\frac{dx(t)}{dt}}{x(t)}=u(t)\ln a+\left(1-u(t)\right)\ln b, \quad x(0)=a_i, \quad x(\tau)=b_i. \tag{29}$$

Так как, в силу соотношения (28),

$$h(x)\ln a=g(x)\ln b,$$

то функционал I принимает вид

$$I=\ln a \int_0^{\tau} \left[u(t)\ln a+\left(1-u(t)\right)\ln b\right] g\left(x(t)\right) dt.$$

Используя (29), имеем:

$$I=\ln a \int_0^{\tau} \frac{x'(t)}{x(t)}\, g\left(x(t)\right) dt=\ln a \int_{a_i}^{b_i} \frac{g(x)}{x}\, dx.$$

Полученное соотношение означает, что любое допустимое управление $u(t)$, $0\leqslant t\leqslant\tau$, является оптимальным.

Применяя доказанную часть теоремы в остальной части интервала $[0, \tilde{X}]$, не принадлежащей интервалам типа $[a_i, b_i]$, $i=1, 2, \ldots, s$, получаем доказательство теоремы в случае, когда не удовлетворено условие (16).

Итак, теорема доказана.

Авторы признательны И. Ячяускасу за замечание.

Институт физики и математики
Академии наук Литовской ССР

Поступило в редакцию
12. VIII. 1971

Литература

1. Р. Беллман, Динамическое программирование, М., ИИЛ, 1960.
2. Л. С. Понтрягин, Математическая теория оптимальных процессов, М., „Наука", 1969.
3. В. Г. Больтянский, Математические методы оптимального управления, М., „Наука", 1969.
4. J. R. Bagwell, C. S. Beightler, J. P. Stark, On a Class of Conve Allocation Problems, of Math. Anal. and App., 25, 218–231 (1969).
5. Г. М. Фихтенгольц, Курс дифференциального и интегрального исчисления, т. 2, Физматгиз, 1962.

DINAMINIO PROGRAMAVIMO PASISKIRSTYMO PROCESO TOLYDINIS ANALOGAS

P. Rutkauskas, V. Bistrickas

(*Reziumė*)

Darbe nagrinėjamas (1) funkcionalinės lygties tolydinis analogas. Gautas optimalaus valdymo uždavinys: rasti tokią valdymo funkciją $u(t)$, tenkinančią (7) nelygybę, kuri maksimizuotų (5) funkcionalą, kai yra išpildyta (6) sąlyga. Atliekama optimalaus valdymo sintezė.

CONTINUOUS ANALOGUE OF DYNAMIC PROGRAMMING ALLOCATION PROCESS

P. Rutkauskas, V. Bistrickas

(*Summary*)

A continuous version of the equation (1) is investigated. It leads to the optimum control problem: to find the control function u (t) satisfying the inequality (7) which maximizes the integral (5) when the condition (6) is fulfilled.