Прикладная математика и механика. Том 50, вып. 1, 1986

УДК 62-50

ОПТИМАЛЬНОЕ УПРАВЛЕНИЕ СКАЧКООБРАЗНЫМИ ПРОЦЕССАМИ С ПЕРИОДИЧЕСКИМИ ХАРАКТЕРИСТИКАМИ

Пиуновский А. Б.

Решается задача оптимального управления скачкообразным марковским процессом с периодическими характеристиками, который не является непрерывным по вероятности. Доказана достаточность периодических марковских стратегий управления, получено уравнение оптимальности, приведены примеры решения практических задач.

Построение оптимальных стратегий управления случайными процессами является актуальной прикладной задачей [1—10]. Наряду со стохастически непрерывными [1, 2, 5, 7—9] и чисто разрывными [3, 4, 6] моделями управляемых процессов представляют интерес задачи, в которых управляемый случайный процесс носит смешанный характер. Так, в работах [1, 10] изучались модели, включающие диффузионную и скачкообразную компоненты, а также некоторые другие взаимодействующие марковские системы. Одна из разновидностей таких комбинированных моделей, включающая цепь с дискретным временем и стохастически непрерывный скачкообразный процесс, рассматривается в настоящей работе. Проблемы оптимального управления такими системами на конечном отрезке времени исследовались в [10]. В данной статье рассматриваются задачи синтеза на бесконечном временном интервале при условии, что все характеристики управляемой модели — периодические функции времени.

1. Обозначения и определения. Рассматривается двухкомпонентный скачкообразный марковский случайный процесс (ξ_t, ψ_t) на бесконечном временном интервале $I = [0, \infty)$. Компонента ξ_t образует стохастически непрерывный процесс, скачки компоненты ψ_t происходят в известные моменты $\tau, 2\tau, \ldots$. Обозначим X — пространство состояний компоненты ξ_t, Y — пространство состояний компоненты ψ_t — конечные или счетные множества. Под состоянием процесса (ξ_t, ψ_t) в момент $t \in I$ понимается упорядоченная пара $(x, y) \in X \times Y$. Траектории полагаются непрерывными справа и имеющими предел слева.

Предположим, что для каждой пары $(x, y) \in X \times Y$ задано борелевское множество допустимых управлений $A\ (x, y)$. Если в момент t процесс (ξ_t, ψ_t) принял значение (x, y) и выбрано управление $a \in A\ (x, y)$, то $\lambda_{x,z}\ (a, y, t)$ — интенсивность перехода компоненты ξ_t из состояния x в состояние $z \in X$ в момент t. В этом случае скорость дохода в момент t составляет $e^{-\alpha t} r_t\ (x, y, a)$. Здесь и ниже $\alpha > 0$ — коэффициент дисконтирования. Далее рассматриваются только периодические модели, для которых $\lambda_{x,z}\ (a, y, t + \tau) = \lambda_{x,z}\ (a, y, t)$; $r_{t+\tau}\ (x, y, a) = r_t\ (x, y, a)$.

Если $(\xi_{n\tau-0}, \psi_{n\tau-0}) = (x, y)$ и выбрано управление $a \in A\ (x, y)$, то $P_{y,z}\ (a, x)$ — вероятность перехода компоненты ψ_t из состояния y в состояние $z \in Y$. В этом случае доход в момент $n\tau$ составляет $e^{-\alpha n\tau} R(x, y, a)$.

Определение. Моделью называется множество $Z = \{I, X, Y, A, \lambda, r, P, R\}$, где $A = \bigcup\limits_{(x, y) \in X \times Y} A\ (x, y)$.

Обозначим символом $(x, y)_0{}^t$ траекторию процесса (ξ_t, ψ_t) на интервале $[0, t] \subset I$.

Определение. Стратегией управления процессом (ξ_t, ψ_t) называется измеримое отображение π, ставящее в соответствие каждой траектории

$(x, y)_0^t$ точку $a = \pi[(x, y)_0^t] \in A(x_t, y_t)$, где (x_t, y_t) — состояние процесса (ξ_t, ψ_t) в момент t. Стратегия называется марковской, если управление зависит лишь от времени и конечного состояния: $\pi[(x, y)_0^t] = \varphi(t, x_t, y_t)$. Марковская стратегия (МС) называется периодической, если $\varphi(t + \tau, x, y) = \varphi(t, x, y)$. При фиксированной стратегии π используется обозначение $a_t = \pi[(x, y)_0^t]$.

Определение. Оценкой стратегии π называется средний доход на интервале I:

$$\omega_0(x, y, \pi) = M_{(x, y)}^{\pi} \left\langle \int_0^{\infty} e^{-\alpha t} r_t(\xi_t, \psi_t, a_t)\, dt + \sum_{n=1}^{\infty} e^{-\alpha n\tau} R(\xi_{n\tau-0}, \psi_{n\tau-0}, a_{n\tau-0}) \right\rangle \tag{1.1}$$

где (x, y) — начальное состояние процесса в момент $t = 0$, $M_{(x, y)}^{\pi}$ — символ математического ожидания по мере $P_{(x, y)}^{\pi}$ на пространстве траекторий процесса (ξ_t, ψ_t), начинающегося из состояния (x, y), при фиксированной стратегии управления π. Оценкой модели Z называется величина

$$v_0(x, y) = \sup_{\pi} \omega_0(x, y, \pi) \tag{1.2}$$

Наряду с моделью Z будет рассматриваться «производная модель» $Z_t^T(S)$, заданная на интервале $I' = [t, T]$, с финальным доходом $S : X \times Y \to \mathbf{R}^1$. Оценки стратегии и модели для $Z_t^T(S)$ в дальнейшем обозначаются символами $\omega_t^T(x, y, \pi, S)$ и $v_t^T(x, y, S)$. Если $T = \infty$ и $S = 0$, то эти аргументы опускаются.

Определение. Стратегия π называется ε-оптимальной, если $\omega_0(x, y, \pi) \geqslant v_0(x, y) - \varepsilon$ для всех $(x, y) \in X \times Y$; 0-оптимальная стратегия называется оптимальной.

Будем считать, что выполнены следующие условия:

1°. Функция $\lambda_{x, z}(a, y, t)$ измерима и ограничена равномерно по всем аргументам.

2°. $\sum_{z \in X} \lambda_{x, z}(a, y, t) = 0$

3°. $\| \sup_{a \in A(x, y)} |r_t(x, y, a)| \| + \| \sup_{a \in A(x, y)} |R(x, y, a)| \| < \infty$

Здесь и ниже

$$\| f_t(x, y) \| = \sup_{t \in [0, \tau),\, (x, y) \in X \times Y} |f_t(x, y)|; \quad \| f(x, y) \| = \sup_{(x, y) \in X \times Y} |f(x, y)|$$

2. Основные теоремы. Введем следующие обозначения:

$$D^a g_t(x, y) = e^{-\alpha t} r_t(x, y, a) + \sum_{z \in X} \lambda_{x, z}(a, y, t)\, g_t(z, y)$$

$$E^a g(x, y) = e^{-\alpha\tau} \left[R(x, y, a) + \sum_{z \in Y} P_{y, z}(a, x)\, g(x, z) \right]$$

где g — действительная функция на $[0, \tau) \times X \times Y$, или на $X \times Y$ соответственно.

Основные математические результаты данной работы содержатся в следующих двух теоремах.

Теорема 1. Для всякого $\varepsilon > 0$ существует периодическая ε-оптимальная стратегия. Если $A(x, y)$ — компакты, функции r и R полунепрерывны сверху по a, функции λ и P непрерывны по a, то существует оптимальная периодическая стратегия.

Теорема 2. а) Оценка модели $v_t(x, y)$ при $t < \tau$ является единственным абсолютно непрерывным решением уравнения оптимальности

$$(2.1)\qquad v_t(x,y) = \sup_{a \in A(x,\,y)} E^a v_0(x,y) + \int_t^\tau \sup_{a \in A(x,\,y)} D^a v_\theta(x,y)\,d\theta$$

б) $v_{t+\tau}(x,y) = e^{-\alpha\tau} v_t(x,y)$

в) периодическая стратегия φ^* оптимальна тогда и только тогда, когда при всех $(x, y) \in X \times Y$

$$(2.2)\qquad (d/dt + D^{a_t^*})\, v_t(\xi_t, \psi_t) = 0, \quad dt \times P^{\varphi^*}_{(x,\,y)} \text{ — п. н.}$$

$$(2.3)\qquad E^{a^*_{\tau-0}} v_0(\xi_{\tau-0}, \psi_{\tau-0}) = v_{\tau-0}(\xi_{\tau-0}, \psi_{\tau-0}), \quad P^{\varphi^*}_{(x,\,y)} \text{ — п. н.}$$

Теорема 1 позволяет при решении задач оптимального управления ограничиться классом периодических стратегий. Уравнение (2.1) дает возможность численно строить оптимальные периодические стратегии.

Для доказательства теорем потребуются следующие вспомогательные построения.

Пусть $\Phi(y)$ — множество всех измеримых функций из $[0, \tau) \times X$ в $A(x, y)$. Очевидно, каждая МС задается последовательностью $[f_0, f_1, \ldots]$, где f_n — отображение, ставящее в соответствие каждому $y \in Y$ некоторый элемент $f_n(y) \in \Phi(y)$. Если $\varphi = [f_0, f_1, \ldots]$, то символом φ^n будет обозначаться последовательность $[f_n, f_{n+1}, \ldots]$.

Пусть f — отображение множества Y в $\Phi = \bigcup_{y \in Y} \Phi(y)$, такое, что $f(y) \in \Phi(y)$; u: $X \times Y \to \mathbf{R}^1$ — некоторая равномерно ограниченная функция. Символом $L(f)\,u$ будет обозначаться решение (при $t = 0$) следующей задачи Коши:

$$(2.4)\qquad (d/dt + D^{f(y)(t,\,x)})\, g_t(x, y) = 0$$

$$(2.5)\qquad g_\tau(x, y) = E^{f(y)(\tau-0,\,x)}\, u(x, y)$$

Наряду с операторами $L(f)$ будет рассматриваться оператор U, задаваемый формулой

$$(2.6)\qquad Uu = \sup_{f:\, Y \to \Phi} L(f)\, u$$

Лемма 1. Для всякого $f: Y \to \Phi$:

а) если $u_1 \geqslant u_2$, то $L(f)\,u_1 \geqslant L(f)\,u_2$ и $Uu_1 \geqslant Uu_2$;

б) $L(f)(u + c) = L(f)\,u + e^{-\alpha\tau}c$; $U(u + c) = Uu + e^{-\alpha\tau}c$, где c — произвольная функция-константа, заданная на $X \times Y$;

в) операторы $L(f)$ и U — сжимающие, причем

$$\| L(f)\,u_1 - L(f)\,u_2 \| \leqslant e^{-\alpha\tau} \| u_1 - u_2 \|; \quad \| Uu_1 - Uu_2 \| \leqslant e^{-\alpha\tau} \| u_1 - u_2 \|;$$

г) для всякого $\varepsilon > 0$ и для любой функции $u(x, y)$ существует отображение $f: Y \to \Phi$, такое, что $L(f)\,u \geqslant Uu - \varepsilon$.

Доказательство. Заметим, что матрица $\exp\left(\int_t^\tau \Lambda(f, y, \theta)\,d\theta\right)$ — стохастическая [11]. Тогда утверждения пп. а), б) следуют непосредственно из (2.4)—(2.6). Доказательство п. в) совпадает с приведенным в [3].

Положим

$$(2.7)\qquad h(x, y) = \sup_{a \in A(x,\,y)} E^a u(x, y)$$

Из результатов, полученных в [5], следует, что $Uu(x, y) = v_0^\tau(x, y, h)$. Там же доказано, что при любом фиксированном y в модели $Z_0^\tau(h)$ существует $\varepsilon/2$-оптимальная МС $f(y)(t, x)$. Кроме того, согласно (2.7), $\exists a^*(x) = f(y)(\tau - 0; x)$: $E^{a^*}u > h - \frac{\varepsilon}{2}$. Повторяя эти рассуждения для всех $y \in Y$, получаем отображение f: $Y \to \Phi$. Для доказательства неравенства $L(f)u \geqslant Uu - \varepsilon$ достаточно заметить, что $\omega_0^\tau(x, y, f(y), h)$ — единственное решение (2.4) с начальным условием $g_\tau = h$ (см. [5]). Лемма доказана.

Лемма 2. Пусть $\varphi = [f_0, f_1, \ldots]$ — произвольная МС. Тогда

а) $\omega_{n\tau}(x, y, \varphi) = e^{-\alpha n\tau}\,\omega_0(x, y, \varphi^n)$

б) $\omega_0(\varphi) = L(f_0)\ldots L(f_n)\,\omega_0(\varphi^{n+1})$

Доказательство. Утверждение п. а) следует из (1.1) и периодичности функций λ, r. Согласно [5], $\omega_t^\tau(x, y, \varphi, \omega_{\tau-0})$ — единственное решение (2.4) при $f = f_0$ с начальным условием $g_\tau(x, y) = \omega_{\tau-0}(x, y, \varphi)$. Из п. а) следует, что $\omega_0(\varphi) = L(f_0)\,\omega_0(\varphi^1)$. Продолжая эти рассуждения, получаем требуемое: $\omega_0(\varphi) = L(f_0)\ldots L(f_n)\,\omega_0(\varphi^{n+1})$.

Лемма 3. Пусть u^* — неподвижная точка оператора U. Тогда для всякого f: $Y \to \Phi$ $L^n(f)\,u^* \xrightarrow[n\to\infty]{} \omega_0(f^\infty)$. Здесь и ниже f^∞ — периодическая стратегия $\varphi = [f, f, \ldots]$.

Доказательство. Согласно лемме 2, $\omega_0(f^\infty) = L(f)\,\omega_0(f^\infty)$. Воспользовавшись п. в) леммы 1, получаем $\| L(f)\ldots L(f)\,u^* - \omega_0(f^\infty)\| \leqslant \leqslant e^{-\alpha(n+1)\tau}\,\|\omega_0(f^\infty) - u^*\| \xrightarrow[n\to\infty]{} 0$, что и требовалось.

Доказательство теоремы 1. Докажем, что для всякого $\varepsilon > 0$ в модели существует ε-оптимальная МС. Из условия 3° следует, что для всякого $\delta > 0$ существует N, такое, что при $n \geqslant N$ справедливо неравенство

$$e^{-\alpha n\tau}\Big\{\frac{1}{\alpha}(\delta + \|\sup_{a\in A(x,\,y)}|r_t(x, y, a)|\|) + \frac{1}{1 - e^{-\alpha\tau}}(\delta + \|\sup_{a\in A(x,\,y)}|R(x, y, a)|\|)\Big\} \leqslant \frac{\varepsilon}{3} \tag{2.8}$$

Известно [1], что в модели $Z_0^{N\tau}(0)$ существует $\varepsilon/3$-оптимальная МС $\varphi_1(t, x, y)$. Наконец, при $t \geqslant N\tau$ выберем стратегию $\varphi_2(t, x, y)$ так, чтобы выполнялись неравенства

$$\begin{aligned} &r_t(x, y, \varphi_2(t, x, y)) \geqslant -\sup_{a\in A(x,y)}|r_t(x, y, a)| - \delta \\ &R(x, y, \varphi_2(n\tau - 0, x, y)) \geqslant -\sup_{a\in A(x,y)}|R(x, y, a)| - \delta \end{aligned} \tag{2.9}$$

Воспользовавшись определениями (1.1), (1.2) и неравенствами (2.8), (2.9), нетрудно показать, что МС $\varphi(t, x, y)$, образованная функциями φ_1 и φ_2, ε-оптимальна.

Для доказательства достаточности периодических стратегий заметим, что $v_0(x, y) \leqslant u^*(x, y)$, где u^* — неподвижная точка оператора U. Действительно, в силу п. в) леммы 1 и п. б) леммы 2 имеем $\omega_0(\varphi) \leqslant L(f_0)\ldots \ldots L(f_n)\,u^* + e^{-\alpha(n+1)\tau}(\|\omega_0(\varphi)\| + \|u^*\|) \leqslant u^* + e^{-\alpha(n+1)\tau}(\|u^*\| + + \|\omega_0(\varphi)\|) \xrightarrow[n\to\infty]{} u^*$. Пусть $L(f)\,u^* \geqslant Uu^* - \varepsilon'$, где $\varepsilon' = \varepsilon(1 - e^{-\alpha\tau})$. Существование отображения f следует из п. г) леммы 1. Пользуясь утверждениями а), б) леммы 1, нетрудно показать методом полной математической индукции справедливость неравенства: $L^n(f)\,u^* \geqslant u^* - \varepsilon$. С помощью

[1] *Пиуновский А. Б.* Оптимальное управление непрерывно-дискретным скачкообразным марковским процессом с полной информацией. М., 1980—17 с. Деп. в ВИНИТИ 27.10.80; № 4512-80.

предельного перехода получаем $\omega_0(f^\infty) \geqslant u^* - \varepsilon \geqslant v_0 - \varepsilon$, что и требовалось. Здесь использована лемма 3.

Вторая часть теоремы доказывается аналогично.

Доказательство теоремы 2. Положим $h_{n\tau}(x, y) = e^{-\alpha(n-1)\tau} \sup\limits_{a \in A(x, y)} E^a \times \times v_0(x, y)$. Согласно лемме 2, при $t < \tau$ $v_t(x, y) = v_t^\tau(x, y, h_\tau)$. Интегральное уравнение (2.1) следует непосредственно из результатов работы [5] для модели $Z_t^\tau(h_\tau)$. Аналогичным образом получаются интегральные представления для $v_{\tau+t} = v_{\tau+t}^{(n+1)\tau}(x, y, h_{(n+1)\tau})$ и $v_t(x, y) = v_t^{n\tau}(x, y, h_{n\tau})$ при $(n - 1)\tau \leqslant t < n\tau$. Утверждение п. б) проверяется подстановкой; существование и единственность решения интегрального уравнения (2.1) и ему подобных доказывается стандартным методом сжимающих отображений. Доказательство п. в) проводится с использованием п. а) леммы 2 и следующего тождества:

$$\omega_t(x, y, \pi) = F_t(x, y) + \mathrm{M}^\pi_{(x, y)}\Big\langle \omega_{\tau-0}(\xi_{\tau-0}, y, \pi) - $$
$$- F_{\tau-0}(\xi_{\tau-0}, y) + \int_t^\tau \{d/d\theta + D^{a_\theta}\} F_\theta(\xi_\theta, y)\, d\theta \Big\rangle \quad (t < \tau)$$

справедливого в силу условий 1, 2 для любой стратегии π и произвольной абсолютно непрерывной функции $F_t(x, y)$ на $[0, \tau) \times X \times Y$. В качестве функции $F_t(x, y)$ следует брать решение уравнения (2.1).

3. Примеры. *Задача о регулируемой подсистеме управления.* Рассмотрим прибор, способный обслуживать требования двух типов, предварительно запасенные в бункерах 1, 2. Прерывать начатое обслуживание запрещается, а выбор очередного требования производится по усмотрению обслуживающего прибора. Время обслуживания требования любого типа полагается экспоненциально распределенным с параметром λ. Руководство работой обслуживающего прибора со стороны подсистемы управления более высокого уровня заключается в следующем. В начале каждого интервала $[0, \tau]$, $[\tau, 2\tau]$; . . . поступает указание, какие требования следует обслуживать в первую очередь (т. е. назначается приоритет $y = 1$ или $y = 2$); в конце соответствующего интервала обслуживающий прибор получает доход R или наказывается штрафом r, в зависимости от того, занят он обслуживанием более или менее приоритетного требования. Пусть $P_1 = 0{,}5 + q$ и $P_2 = 0{,}5 - q$ — вероятности назначения приоритетов 1 и 2 соответственно $(-0{,}5 \leqslant q \leqslant 0{,}5)$. Для заданных λ, τ, R, r, q, α (коэффициент дисконтирования) требуется найти оптимальное поведение обслуживающего прибора в смысле суммарного дисконтированного дохода на бесконечном интервале.

Согласно принятым обозначениям, $X = \{1, 2\}$; $Y = \{1, 2\}$, причем $\xi_t = x \in X$ означает, что прибор занят обслуживанием требования типа x; $\psi_t = y \in Y$ означает, что требованиям типа y приписан более высокий приоритет. Пусть $A(x, y) = \{0, 1\}$, где $a = 1$ $(a = 0)$ означает, что принято решение переключиться (не переключаться) на противоположный бункер. Пусть $\lambda_{x,z}(a, y, t) = (-1)^{x+z+1} a\lambda$;

$$r_t(x, y, a) = 0; \quad R(x, y, a) = \begin{cases} R, & x = y, \\ -r, & x \neq y; \end{cases} \quad P_{y,z}(a, x) = P_z$$

Уравнение оптимальности (2.1) можно записать в следующем виде:

$$d/dt v_t(x, y) = -\sup_{a \in \{0, 1\}} \{a\lambda\,[v_t(x', y) - v_t(x, y)]\} \tag{3.1}$$

$$(3.2)\qquad v_{\tau-0}(x,y)=e^{-\alpha\tau}[R_{x,y}+P_1v_0(x,1)+P_2v_0(x,2)]$$

$$x'=\begin{cases}1, & x=2\\ 2, & x=1\end{cases};\quad R_{x,y}=\begin{cases}R, & x=y\\ -r, & x\neq y\end{cases}$$

Предположим, что

$$(3.3)\qquad v_{\tau-0}(1,1)\geqslant v_{\tau-0}(2,1),\ v_{\tau-0}(1,2)\leqslant v_{r-0}(2,2)$$

Можно убедиться, что в этом случае решение системы (3.1) имеет вид $v_t(x,y)=v_{\tau-0}(x,y)$ $(x=y)$; $v_t(x,y)=e^{-\lambda(\tau-t)}[v_{\tau-0}(x,y)-v_{\tau-0}(x',y)]+v_{\tau-0}(x',y)$ $(x\neq y)$. Подставив найденные выражения при $t=0$ в (3.2) и заметив, что при $x\neq y$ $v_{\tau-0}(x,y)=v_{\tau-0}(x,y')-e^{-\alpha\tau}(R+r)$, получаем систему линейных уравнений для $v_{\tau-0}(1,1)$ и $v_{\tau-0}(2,2)$, откуда $v_{\tau-0}(1,1)=\Delta_1/\Delta$; $v_{\tau-0}(2,2)=\Delta_2/\Delta$. Эти рассуждения справедливы лишь в случае, когда выполнены неравенства (3.3), эквивалентные неравенству $|\Delta_1-\Delta_2|\leqslant e^{-\alpha\tau}(R+r)\Delta$, решение которого имеет вид

$$(3.4)\qquad (\alpha+\lambda)\tau\geqslant\ln(1+2|q|)$$

Оптимальная стратегия в этом случае будет следующей:

$$\varphi(t,x,y)=\begin{cases}0, & x=y\\ 1, & x\neq y\end{cases}$$

Таким образом, если выполнено (3.4), то оптимальная стратегия предписывает выполнять поступающие распоряжения.

Если вместо (3.3) рассмотреть гипотезы

$$(3.5)\qquad v_{\tau-0}(1,1)\geqslant v_{\tau-0}(2,1);\quad v_{\tau-0}(1,2)\geqslant v_{\tau-0}(2,2)$$

$$(3.6)\qquad v_{\tau-0}(1,1)\leqslant v_{\tau-0}(2,1);\quad v_{\tau-0}(1,2)\leqslant v_{\tau-0}(2,2)$$

то с помощью аналогичных рассуждений получаем оптимальные стратегии

$$\varphi(t,x,y)=\begin{cases}0, & x=1\\ 1, & x=2\end{cases};\quad \varphi(t,x,y)=\begin{cases}0, & x=2\\ 1, & x=1\end{cases}$$

соответственно, причем (3.5) эквивалентно неравенствам $q\geqslant 0$; $(\alpha+\lambda)\tau\leqslant\ln(1+2q)$, а (3.6) — неравенствам $q\leqslant 0$; $(\alpha+\lambda)\tau\leqslant\ln(1-2q)$.

Следовательно, оптимальное поведение обслуживающего прибора имеет вид, представленный на фиг. 1, где область *а* отвечает решению «принимать к обслуживанию только требования первого типа», область *б* — «выполнять поступающие распоряжения», область *в* — «принимать к обслуживанию только требования второго типа». Линия переключения ACB задается уравнением $(\alpha+\lambda)\tau=\ln(1+2|q|)$.

Одноканальная СМО с отказами. Рассмотрим одноканальную систему массового обслуживания (СМО), на которую поступает пуассоновский входной поток с интенсивностью $\lambda(t)=b+d\sin^2(\pi t/\tau)$. Предположим, что имеется два режима работы СМО, характеризующихся интенсивностями обслуживания μ_1 и μ_2, причем $\mu_1<\mu_2$. Скорость потерь, связанных с обслуживанием, равна r_1 в первом режиме и r_2 — во втором. Потеря требования приводит к штрафу R. Требуется построить оптимальную стратегию управления системой, т. е. для каждого момента времени указать наилучший режим работы СМО.

Математической моделью описанной системы является стохастически непрерывный управляемый марковский процесс с периодическими характеристиками, являющийся частным случаем исследованной непрерывно-дискретной модели.

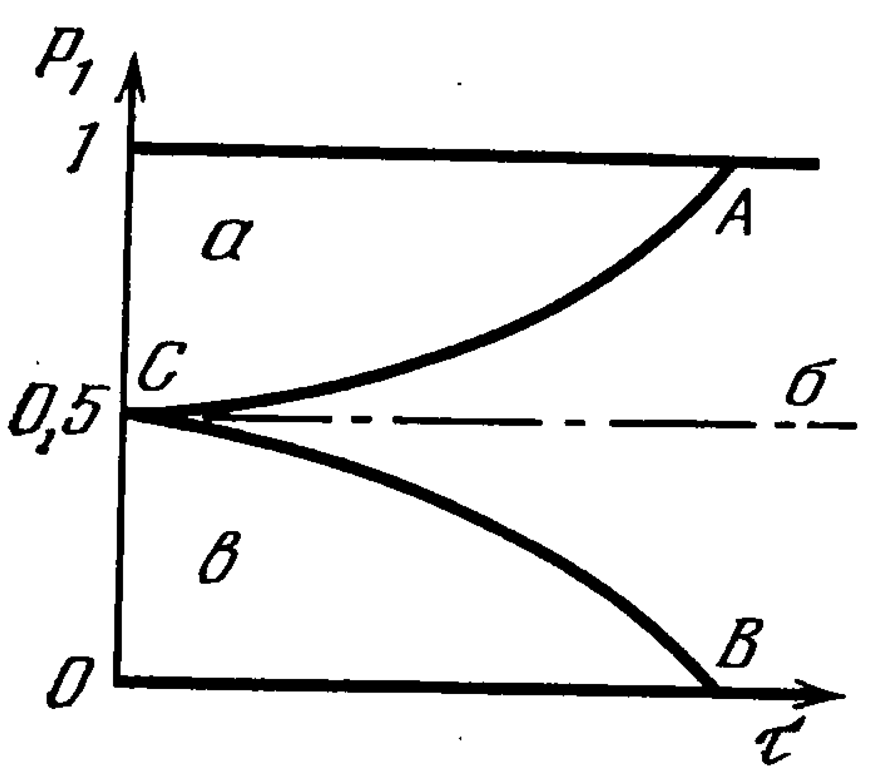
Согласно принятым обозначениям, $X = \{1, 2\}$, причем $\xi_t = 1$ означает, что канал свободен; $\xi_t = 2$ — канал занят. Пространство управлений $A = \{1, 2\}$ состоит из двух элементов: 1 — первый режим работы; 2 — второй режим. Компонента ψ_t отсутствует (т. е. $Y = \{1\}$; $R\,(x, y, a) = 0$), и в дальнейшем аргумент y опускается. Инфинитезимальная матрица:

$$\Lambda(a, t) = \begin{Vmatrix} -b - d\sin^2\left(\frac{\pi t}{\tau}\right) & b + d\sin^2\left(\frac{\pi t}{\tau}\right) \\ \mu_a & -\mu_a \end{Vmatrix}$$

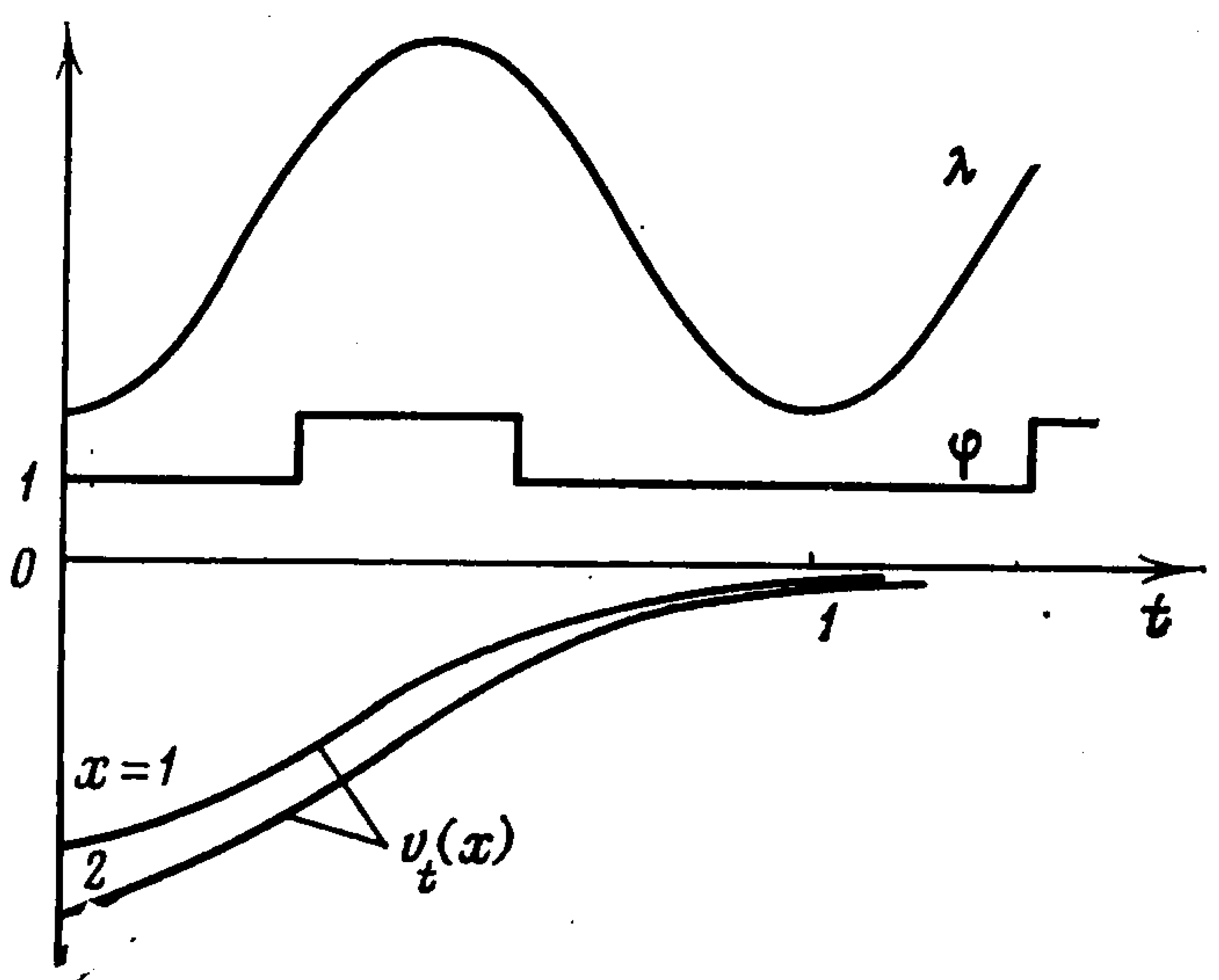
Интенсивность дохода равна

$$r_t(x, a) = \begin{cases} 0, & x = 1 \\ -r_a - R\,(b + d\sin^2(\pi t/\tau)), & x = 2 \end{cases}$$

Задача решалась при следующих значениях параметров: $\tau = 1$, $\alpha = 2$, $r_1 = 20$, $r_2 = 120$, $R = 2$, $\mu_1 = 20$, $\mu_2 = 100$, $b = 2$, $d = 10$. Начальное



Фиг. 1



Фиг. 2

приближение для $v_0\,(x)$ было выбрано нулевым. Уравнение (2.1) решалось итерационным методом, и на пятом шаге оценка модели $v_0\,(x)$ была вычислена с точностью 0,01. Графики оценки модели и оптимальной стратегии управления приведены на фиг. 2. Как видно, второй режим работы СМО следует выбирать на интервале (0, 3; 0,6]. Это объясняется возрастанием интенсивности входного потока. В дальнейшем, на интервалах [1, 2]; [2, 3]; . . . стратегия периодически повторяется.

4. Частные случаи. Рассмотрим несколько частных случаев общей модели управляемой системы, описанной в пп. 1, 2. Пусть $m\,(X) = 1$; $r_t\,(x, y, a) = 0$. Здесь и ниже $m\,(D)$ — мощность конечного множества D. Тогда исследованная модель превращается в хорошо изученный дисконтированный марковский процесс принятия решений [3]. В этом случае приведенные в п. 2 утверждения совпадают с известными результатами.

С другой стороны, если положить $m\,(Y) = 1$; $R\,(x, y, a) = 0$, то свойство стохастической непрерывности восстанавливается. Такие модели изучались в предположении, что $\lambda_{x,z}\,(a, t) \equiv \lambda_{x,z}\,(a)$; $r_t\,(x, a) \equiv r\,(x, a)$. Доказано [7], что в подобных моделях достаточно ограничиться стационарными МС: $\varphi\,(t, x) \equiv \varphi\,(t)$. Этот результат хорошо согласуется с теоремой 1, если заметить, что в качестве τ можно взять произвольное действительное число.

Предположим, что $m\,(Y) = 1$; $R\,(x, y, a) = 0$, но процесс не стационарен. Тогда изученная модель дает важный частный случай управляемых стохастически непрерывных скачкообразных марковских процессов, а именно: модель с периодическими функциями $\lambda_{x,z}$ (a, t) и r_t (x, a). Полученные результаты позволяют утверждать, что в таких моделях достаточно ограничиться периодическими стратегиями с тем же периодом τ. Согласно п. б) теоремы 2, оценка модели — периодическая экспоненциально затухающая функция.

При $\tau \to \infty$ исследованная модель превращается в стохастически непрерывный дисконтированный управляемый марковский процесс. При этом понятия марковской и периодической стратегий управления сливаются и сформулированные в п. 2 утверждения совпадают с результатами работы [5].

Автор благодарит Г. Е. Колосова за интерес к работе и советы.

ЛИТЕРАТУРА

1. *Черноусько Ф. Л., Колмановский В. Б.* Оптимальное управление при случайных возмущениях. М.: Наука, 1978. 351 с.
2. *Крылов Н. В.* Управляемые процессы диффузионного типа. М.: Наука, 1977. 399 с.
3. *Майн Х., Осаки С.* Марковские процессы принятия решений. М.: Наука, 1977. 175 с.
4. *Дашевский М. Л., Липцер Р. Ш.* Моделирование стохастических дифференциальных уравнений, связанных с задачей о «разладке» на аналоговой машине.— Автоматика и телемеханика, 1966, № 4, с. 142—150.
5. *Юшкевич А. А.* Управляемые марковские модели со счетным множеством состояний и непрерывным временем.— Теория вероятностей и ее применения, 1977, т. 22, вып. 2, с. 222—241.
6. *De Leve G., Federgruen A., Tijms H. C.* A general Markov decision method I: Model and techniques.— Adv. Appl. Probability, 1977, v. 9, No. 2, p. 296—315.
7. *Юшкевич А. А., Файнберг Е. А.* Об однородных управляемых марковских моделях с непрерывным временем и конечным или счетным множеством состояний.— Теория вероятностей и ее применения, 1979, т. 24, вып. 1, с. 155—160.
8. *Колосов Г. Е.* Синтез оптимальных стохастических систем управления методом последовательных приближений.— ПММ, 1979, т. 43, вып. 1, с. 7—16.
9. *Колмановский В. Б., Майзенберг Т. Л.* Об управлении временем достижения области случайным движением.— ПММ, 1980, т. 44, вып. 1, с. 13—21.
10. *Пиуновский А. Б., Хаметов В. М.* Об оптимальном управлении непрерывно-дискретными скачкообразными процессами.— Изв. АН СССР. Техн. кибернетика, 1983, № 3, с. 56—61.
11. *Гихман И. И., Скороход А. В.* Введение в теорию случайных процессов. М.: Наука, 1977. 567 с.

Москва

Поступила в редакцию
26.VII.1983