Синтез оптимального стохастического управления системой массового обслуживания с учетом ограничений*

Б. М. Миллер
Monash University, Victoria, Australia and the Institute for Information Transmission Problems
bmiller@iitp.ru

Д. В. Мясников
Московский физико-технический институт
dmitry.myasnikov@phystech.edu

К. В. Семенихин
Московский авиационный институт
siemenkv@gmail.com

Аннотация

Рассматривается задача оптимального стохастического управления одноканальной системой массового обслуживания по критерию минимума среднего времени полного обслуживания с учетом ограничений на среднее число отраженных заявок и объем использованных ресурсов. Роль контролируемых параметров выполняют вероятность отражения заявки и интенсивность обслуживания. Математическая постановка проблемы основана на понятии управляемого марковского процесса и его мартингальном разложении. Для решения задачи с ограничениями используется принцип множителей Лагранжа. Законность его применения обоснована выпуклостью множества достижимости, т.е. множества векторов, составленных из значений оптимизируемого функционала и функционалов, задающих ограничения. В работе установлено, что при определенных условиях оптимальное управление по расширенному функционалу в задаче без ограничений дает оптимум в задаче с ограничениями, если множители Лагранжа образуют решение двойственной задачи. Приведены результаты численного моделирования динамики системы при использовании оптимального стохастического управления и детерминированного управления, рассчитанного на оптимизацию системы в стационарном режиме.

1. Введение

Эффективность управления потоками данных остается актуальной проблемой теории и практики инфокоммуникационных систем. Несмотря на большое количество разработанных протоколов передачи данных и инженерных методик их оптимизации, математическую основу управления потоками данных составляет теория массового обслуживания, методы которой обычно ограничены анализом стационарного случая [1–3]. Однако применение методов, рассчитанных на асимптотическое поведение, в реальных системах массового обслуживания (СМО) сопряжено с множеством ограничением. Например, передача данных по каналам мобильной или спутниковой связи, управление водными ресурсами осложнены нестационарностью или периодичностью процессов, описывающих эволюцию этих систем. Поэтому в последнее время для разработки эффективных стратегий управления инфокоммуникационными системами используют методы оптимального стохастического управления скачкообразными марковскими процессами.

Для процесса с конечным числом состояний задача оптимального управления без ограничений была решена достаточно давно [4–6]. Общий подход к решению подобных задач с использованием мартингального представления был позднее развит для более широкого класса управляемых процессов [7–9]. Оптимизация управляемых марковских процессов с учетом ограничений в рамках стационарных постановок посвящены работы [2, 10–12]. В [13, 14] исследована задача управления марковским процессом на фиксированном промежутке времени по интегральному критерию качеству при наличии конечного числа ограничений типа неравенств, задаваемых посредством функционалов той же структуры, что и целевой функционал. Основной прием, позволяющий синтезировать оптимальное стохастическое управление в задаче с ограничениями, состоит в пе-

*Работа поддержана грантом РФФИ № 13-01-00406 и частично грантом РФФИ № 13-07-00408.

реходе к минимаксной формулировке, для решения которой требуется определить оптимальный набор множителей Лагранжа, а затем построить оптимальное управление в задаче с соответствующим расширенным функционалом.

Настоящая работа посвящена развитию этого подхода применительно к задаче управления одноканальной СМО по критерию минимума среднего времени полного обслуживания с учетом ограничений на среднее число отраженных заявок и объем использованных ресурсов. Для управления СМО по принципу обратной связи вероятность отклонения заявки и интенсивность обслуживания предполагаются функциями от предыстории системы. В работе приведены результаты численного моделирования динамики системы при использовании оптимального стохастического управления и детерминированного управления, рассчитанного на оптимизацию системы в стационарном режиме.

2. Постановка задачи

Рассмотрим систему, состоящую из одного прибора обслуживания и очереди на $N-1$ место. Состояние СМО будем описывать процессом $\nu(t) \in \mathsf{N} = \{0,1,\dots,N\}$, равным числу заявок в системе, где $t \in [0,T]$. Поступающие заявки образуют нестационарный пуассоновский поток, интенсивность которого $\alpha(t)$ известна и является непрерывной функцией. Заявка, прибывающая в момент t, такой что $\nu(t) < N$, отклоняется с вероятностью $R(t) \in [0,1]$. На обслуживание заявки, находящейся на приборе в момент t, тратится случайное время, которое будет больше h с вероятностью

$$\exp\Big\{-\int\limits_t^{t+h} M(s)\,ds\Big\}, \quad \text{где} \quad M(t) \in [\underline{\mu},\overline{\mu}],$$

причем $0 < \underline{\mu} < \overline{\mu} < \infty$ — известные границы.

Данная СМО рассматривается, как система управления с обратной связью, в которой вероятность отражения заявки $R(t)$ и интенсивность обслуживания $M(t)$ играют роль контролируемых параметров. Поэтому управление $U(t) = (R(t), M(t))$ — это случайный процесс, предсказуемый относительно потока сигма-алгебр $\{\mathscr{F}_t\}$, порожденных процессом $\nu(t)$. Иначе говоря, если τ_k и ν_k обозначают соответственно момент k-го скачка и значение процесса $\nu(t)$ в этот момент (где $k \geq 0$, $\tau_0 = 0$), то определены борелевские функции

$$\mathbf{u}_k\colon \underbrace{[0,T]\times\mathsf{N}\times\ldots\times[0,T]\times\mathsf{N}}_{k+1} \to \mathsf{U} = [0,1]\times[\underline{\mu},\overline{\mu}],$$

такие что $U(t) = \mathbf{u}_k(t,\nu_0,\tau_1,\nu_1,\dots,\tau_k,\nu_k)$ $\forall t \in (\tau_k,\tau_{k+1}]$. Тем самым между скачками $U(t)$ — программное управление. В случае

$$R(t) = r_{\nu(t-)}(t), \qquad M(t) = \mu_{\nu(t-)}(t), \tag{1}$$

где $r_n(t) \in [0,1]$, $\mu_n(t) \in [\underline{\mu},\overline{\mu}]$, $n \in \mathsf{N}$, — некоторые детерминированные борелевские функции, $U(t)$ представляет собой марковское управление. Если же α, $R = r$, $M = \mu$ — константы, то получаем «классический» случай: $M_\lambda\,|\,M_\mu\,|\,1\,|\,N-1$, где $\lambda = (1-r)\alpha$.

Для описания качества работы СМО введем три функционала. Основной оптимизируемой характеристикой системы будет среднее время полного обслуживания:

$$J_0(U) = \frac{1}{T}\int\limits_0^T \mathbf{E}\left\{\frac{\nu(t)}{M(t)}\right\} dt.$$

Рассмотрим также среднее количество отраженных заявок

$$J_1(U) = \int\limits_0^T \mathbf{E}\{\mathrm{I}\{\nu(t)=N\} + R(t)\,\mathrm{I}\{\nu(t)<N\}\}\,\alpha(t)\,dt,$$

где $\mathrm{I}\{\dots\}$ — индикатор случайного события. Третий функционал, равный среднему объему использованных ресурсов, будем считать пропорциональным интенсивности обслуживания:

$$J_2(U) = \int\limits_0^T \mathbf{E}\{M(t)\,\mathrm{I}\{\nu(t)>0\}\}\,dt.$$

Итак, задача синтеза оптимального стохастического управления СМО состоит в построении такого предсказуемого управления $\hat{U}(t)$, которое доставляло бы минимум

$$\begin{cases} J_0(U) \to \min\limits_{U\in\mathscr{U}} \quad \text{при ограничениях:} \\ J_1(U) \leq \overline{J}_1, \quad J_2(U) \leq \overline{J}_2, \end{cases} \tag{2}$$

по множеству $\mathscr{U}$ всех предсказуемых управлений $U(t)$, где $\overline{J}_1, \overline{J}_2$ — заданные верхние границы.

3. Оптимальное управление СМО без учета ограничений

Для незамкнутой системы, т.е. при постоянном детерминированном управлении $R(t) = r$, $M(t) = \mu$, $\nu(t)$ будет марковским процессом с множеством состояний N и трехдиагональной матрицей интенсивностей $A(t,r,\mu) = \{a_{nm}(t,r,\mu)\}_{n,m\in\mathsf{N}}$. На главной диагонали стоят со знаком минус интенсивности выхода

$$a_n(t,r,\mu) = \begin{cases} \alpha(t)(1-r), & n = 0, \\ \alpha(t)(1-r)+\mu, & n = 1,\dots,N-1, \\ \mu, & n = N, \end{cases}$$

а на побочных диагоналях стоят интенсивности переходов $n \to n \pm 1$:

$$a_{n,n+1}(t,r,\mu) = \alpha(t)(1-r), \qquad n = 0,1,\ldots,N-1,$$
$$a_{n,n-1}(t,r,\mu) = \mu, \qquad n = 1,2,\ldots,N.$$

Определение замкнутой системы основано на понятии управляемого марковского процесса $\nu(t)$. Традиционный подход к его описанию состоит в задании условных распределений моментов переключений τ_k и состояний $\nu_k = \nu(\tau_k)$, $k \ge 0$. Рассмотрим для простоты марковское управление (1). Тогда

$$\mathbf{P}\big\{\tau_{k+1} > t \mid \mathscr{F}_{\tau_k}\big\} = \exp\Big\{-\int_{\tau_k}^{t} a_n(s, r_n(s), \mu_n(s))\,ds\Big\}\Big|_{n=\nu_k}$$
$$\mathbf{P}\big\{\nu_{k+1} = \nu_k \pm 1 \mid \mathscr{F}_{\tau_k}, \tau_{k+1}\big\} = \\ = \frac{a_{n,n\pm1}(s, r_n(s), \mu_n(s))}{a_n(s, r_n(s), \mu_n(s))}\Big|_{n=\nu_k,\, s=\tau_{k+1}}$$

где $\mathscr{F}_{\tau_k}$ — сигма-алгебра, порожденная величинами ν_0, τ_1, ν_1, …, τ_k, ν_k.

Описанный способ удобен для численного моделирования замкнутой системы. Однако для исследования задачи оптимального стохастического управления более плодотворен подход, основанный на мартингальном разложении векторного процесса $X(t) = e^{(\nu(t))} \in \mathbb{R}^{N+1}$, где $e^{(n)}$ — единичный вектор-столбец:

$$e^{(n)} = [\underbrace{0,\ldots,0}_{n}, 1, 0, \ldots, 0]^*, \quad n = 0,1,\ldots,N. \tag{3}$$

Пользуясь представлением

$$X(t) = X(0) + \int_0^t A^*(s, U(s))X(s)\,ds + \mathscr{M}(t),$$

где $\mathscr{M}(t)$ — мартингал, а $U(t) = (R(t), M(t))$ — марковское управление (1), легко вывести систему уравнений

$$\dot{\pi}_m(t) = \sum_{n=0}^{N} a_{nm}(t, r_n(t), \mu_n(t))\pi_n(t). \tag{4}$$

для одномерного распределения

$$\pi(t) = \{\pi_n(t)\}_{n\in\mathsf{N}} = \mathbf{E}X(t), \quad \pi_n(t) = \mathbf{P}\{\nu(t) = n\}.$$

Кроме того, если ввести обозначение

$$f_{ln}(t,r,\mu) = \begin{cases} n/(T\mu), & l = 0,\\ (\mathrm{I}\{n=N\} + r\,\mathrm{I}\{n<N\})\alpha(t), & l = 1,\\ \mu\,\mathrm{I}\{n>0\}, & l = 2,\end{cases}$$

то определенные выше функционалы могут быть записаны в виде

$$J_l(U) = \int_0^T \sum_{n=0}^{N} f_{ln}(t, r_n(t), \mu_n(t))\pi_n(t)\,dt, \quad l = 0,1,2.$$

Рассмотрим следующую постановку задачи оптимального управления: найти предсказуемое управление $\hat{U}(t,\lambda)$, которое доставляет минимум линейной комбинации функционалов

$$\langle \lambda, J(U)\rangle = \lambda_0 J_0(U) + \lambda_1 J_1(U) + \lambda_2 J_2(U) \to \min_{U\in\mathscr{U}}, \tag{5}$$

где $\lambda = [\lambda_0, \lambda_1, \lambda_2]^*$ — заданный набор множителей, $J(\cdot) = [J_0(\cdot), J_1(\cdot), J_2(\cdot)]^*$ — векторный функционал.

Определим вектор-функцию $H(\cdot) = \{H_n(\cdot)\}_{n\in\mathsf{N}}$:

$$H_n(t,x,r,\mu,\lambda) = (A(t,r,\mu)x)_n + \\ + \sum_{l=0}^{2} \lambda_l f_{ln}(t,r,\mu), \quad x \in \mathbb{R}^{N+1}.$$

Тогда система дифференциальных уравнений метода динамического программирования

$$\dot{\varphi}_n(t,\lambda) = -\min_{(r,\mu)\in\mathsf{U}} H_n(t, \varphi(t,\lambda), r, \mu, \lambda), \quad n \in \mathsf{N}, \tag{6}$$

как система с липшицевой правой частью, на отрезке $[0,T]$ будет иметь (и притом единственное) решение $\varphi(t,\lambda) = \{\varphi_n(t,\lambda)\}_{n\in\mathsf{N}}$, удовлетворяющее терминальному условию

$$\varphi(T,\lambda) = 0. \tag{7}$$

В силу непрерывности $H_n(\cdot)$ и компактности U определены борелевские функции $\hat{r}_n(t,\lambda)$, $\hat{\mu}_n(t,\lambda)$, на которых реализуется минимум в (6).

Т е о р е м а 1. *Начальное значение функции $\varphi(t,\lambda)$, усредненное в соответствии с начальным распределением $\pi(0)$, совпадает с минимумом в (5), т.е. $\langle \varphi(0,\lambda), \pi(0)\rangle = \min_{U\in\mathscr{U}} \langle \lambda, J(U)\rangle$, а марковское управление*

$$\hat{U}(t,\lambda) = (\hat{r}_{\nu(t-)}(t,\lambda), \hat{\mu}_{\nu(t-)}(t,\lambda)) \tag{8}$$

образует решение задачи оптимального управления с расширенным функционалом (5).

Нетрудно найти аналитические выражения для функций (8):

$$\hat{r}_n(t,\lambda) = \mathrm{I}\{-\varphi_n(t,\lambda) + \varphi_{n+1}(t,\lambda) - \lambda_1 > 0\}$$

при $n < N$ и $\hat{r}_N(t,\lambda) = 0$, а также

$$\hat{\mu}_n(t,\lambda) = S(\varphi_{n-1}(t,\lambda) - \varphi_n(t,\lambda) + \lambda_2, \lambda_0 n/T)$$

при $n > 0$ и $\hat{\mu}_0(t,\lambda) = 0$, где

$$S(a,b) = \begin{cases} \underline{\mu}, & a > 0,\ \sqrt{b/a} < \underline{\mu},\\ \sqrt{b/a}, & a > 0,\ \sqrt{b/a} \in [\underline{\mu}, \overline{\mu}],\\ \overline{\mu}, & a > 0,\ \sqrt{b/a} > \overline{\mu} \text{ или } a \le 0.\end{cases} \tag{9}$$

4. Множество достижимости

Ключевым результатом, позволяющим решать задачу оптимального управления с ограничениями, является теорема о выпуклости множества достижимости, т.е. множества значений векторного функционала:

$$\mathsf{J} = \{J(U) \colon U \in \mathscr{U}\}.$$

В [13, 14] эта теорема была доказана в предположении выпуклости множеств

$$\{B(t,u)e^{(n)} \colon u \in \mathsf{U}\}, \quad n \in \mathsf{N}, \quad t \in [0,T], \tag{10}$$

где $B(t,u) = \begin{pmatrix} A^*(t,u) \\ F(t,u) \end{pmatrix}$, а $F(t,u)$ — матричная функция в представлении векторного функционала

$$J(U) = \int\limits_0^T \mathbf{E}\{F(t,U(t))X(t)\}\,dt. \tag{11}$$

Предположение такого вида известно в теории оптимального управления детерминированными системами, как достаточное условие существования решения [15]. Отметим, что в теории управляемых марковских процессов условие выпуклости используется в терминах переходных вероятностей [10].

Для рассматриваемой СМО матрица $F(t,u)$ при $u = (r,\mu)$ состоит из элементов $f_{ln}(t,r,\mu)$, $l = 0,1,2$, $n \in \mathsf{N}$. Тогда если из вектора $B(t,u)e^{(n)}$ взять всего две компоненты, например, $y_l = f_{ln}(t,r,\mu)$ при $l = 0,2$ и $n = 1$, то при произвольном выборе $(r,\mu) \in \mathsf{U}$ пары (y_0, y_2) будут составлять точки гиперболы, что противоречит выпуклости множества (10). Тем не менее, эта кривая ограничивает снизу выпуклую область, поэтому вместо (10) выпуклыми будут множества

$$\left\{ \begin{pmatrix} x \\ y \end{pmatrix} \colon \exists u \in \mathsf{U} \colon \left\{ \begin{array}{l} x = A^*(t,u)e^{(n)} \\ y \ge F(t,u)e^{(n)} \end{array} \right\}, \quad n \in \mathsf{N}. \tag{12}$$

Этот факт немедленно следует из того, что элементы матриц $A^*(t,r,\mu)$ и $F(t,r,\mu)$ представляют собой соответственно линейные и выпуклые функции переменных $r \in [0,1]$, $\mu > 0$.

Теорема 2. *Пусть для управляемого марковского процесса $\{X(t),\, t \in [0,T]\}$ со значениями в множестве S единичных вектор-столбцов (3), матрица интенсивностей $A(t,u)$ непрерывна по $(t,u) \in [0,T] \times \mathsf{U}$, где U — компакт.*

Если векторный функционал имеет вид (11), в котором $F(t,u)$ — непрерывная матричная функция, и для каждого $t \in [0,T]$ множества (12) выпуклы, то выпуклым будет множество

$$\widetilde{\mathsf{J}} = \{y \colon \exists U \in \mathscr{U} \colon y \ge J(U)\}.$$

Отметим, что у множества достижимости J и его расширения $\widetilde{\mathsf{J}}$ одна и та же «нижняя граница»: точнее, для каждого вектора λ, составленного из неотрицательных коэффициентов, справедливо

$$\operatorname*{arg\,min}_{y \in \mathsf{J}} \langle \lambda, y \rangle = \mathsf{J} \cap \operatorname*{arg\,min}_{y \in \widetilde{\mathsf{J}}} \langle \lambda, y \rangle. \tag{13}$$

На рис. 1–2 приведена графическая иллюстрация факта, установленного в теореме 2. Для сравнения нижней границы множества достижимости $\{J(\hat{U}(t,\lambda)) \colon \lambda \ge 0\}$ была построена поверхность $\{J(r,\mu) \colon (r,\mu) \in \mathsf{U}\}$, образованная значениями векторного функционала при постоянных управлениях r, μ.

5. Оптимальное управление в задаче с ограничениями

Согласно теореме 2 задача оптимального управления с ограничениями (2) может быть сформулирована в виде задачи выпуклого программирования

$$\begin{cases} y_0 \to \min\limits_{(y_0,y_1,y_2) \in \widetilde{\mathsf{J}}} \\ y_1 \le \bar{J}_1, \ y_2 \le \bar{J}_2, \end{cases}$$

к которой применима теорема Куна—Таккера [16]. Если предположить, что выполнено условие Слейтера

$$\exists \overline{U} \in \mathscr{U} \colon \ J_l(\overline{U}) < \bar{J}_l, \quad l = 1,2, \tag{14}$$

то управление $\hat{U} \in \mathscr{U}$ будет оптимальным в задаче с ограничениями (2) тогда и только тогда, когда найдутся числа $\hat{\lambda}_1 \ge 0$, $\hat{\lambda}_2 \ge 0$, такие что

$$\hat{U} \in \operatorname*{arg\,min}_{U \in \mathscr{U}} \langle \hat{\lambda}, J(U) \rangle, \quad \hat{\lambda} = [1, \hat{\lambda}_1, \hat{\lambda}_2]^*,$$
$$J_l(\hat{U}) \le \bar{J}_l, \quad \hat{\lambda}_l(J_l(\hat{U}) - \bar{J}_l) = 0, \quad l = 1,2.$$

Последнее условие можно эквивалентным образом записать в виде того, что пара $(\hat{U}, \hat{\lambda})$ образует седловую точку функционала $\langle \lambda, J(U) \rangle$ на произведении конуса множителей Лагранжа

$$\Lambda = \{[1, \lambda_1, \lambda_2]^* \colon \lambda_1 \ge 0, \lambda_2 \ge 0\}$$

и множества предсказуемых управлений $\mathscr{U}$, т.е.

$$\langle \lambda, J(\hat{U}) \rangle \le \langle \hat{\lambda}, J(\hat{U}) \rangle \le \langle \hat{\lambda}, J(U) \rangle \quad \forall (\lambda, U) \in \Lambda \times \mathscr{U}.$$

Таким образом, нахождение оптимального управления в задаче с ограничениями (2) равносильно решению минимаксной задачи

$$\hat{U} \in \operatorname*{arg\,min}_{U \in \mathscr{U}} \sup_{\lambda \in \Lambda} \langle \lambda, J(U) \rangle.$$

Поэтому для определения $\hat{U}$ можно предложить следующую схему [17, 18]:

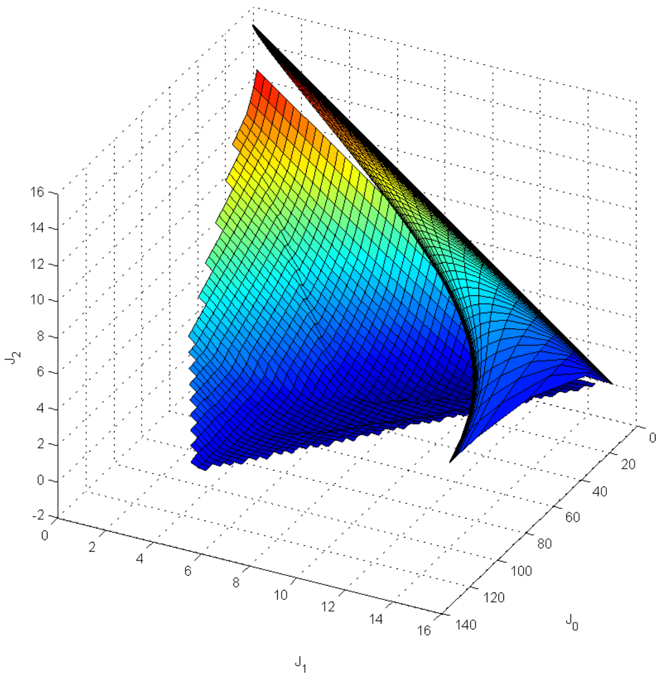

Рис. 1. Нижняя граница множества достижимости J и поверхность, образованная векторами $J(r,\mu)$ при выборе постоянных управлений r,μ.

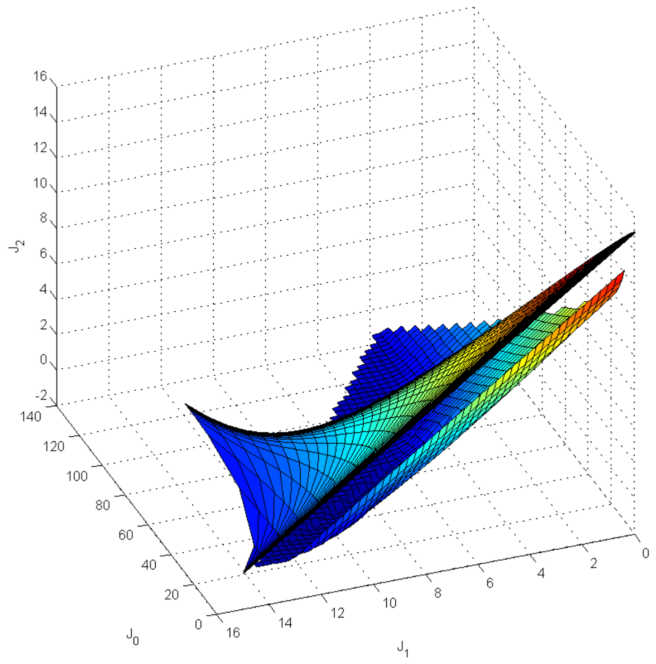

Рис. 2. Нижняя граница множества достижимости J и поверхность, образованная векторами $J(r,\mu)$ при выборе постоянных управлений r,μ.

1) найти решение двойственной задачи

$$\hat{\lambda} \in \arg\max_{\lambda\in\Lambda} \min_{U\in\mathscr{U}} \langle \lambda, J(U)\rangle; \tag{15}$$

2) синтезировать оптимальное управление, соответствующее оптимальному вектору множителей $\hat{\lambda}$:

$$\hat{U} \in \arg\min_{U\in\mathscr{U}} \langle \hat{\lambda}, J(U)\rangle.$$

Перечислим условия, при которых указанный метод действительно приводит к решению минимаксной задачи [19]:

а) Λ — выпуклое множество;

б) $\langle \lambda, J(U)\rangle$ — выпуклый функция по λ;

в) для каждого $\lambda \in \Lambda$ определено

$$\hat{U}(\cdot,\lambda) \in \arg\min_{U\in\mathscr{U}} \langle \lambda, J(U)\rangle;$$

г) существует решение двойственной задачи (15);

д) управление $\hat{U}(\cdot,\lambda)$ непрерывно зависит от λ в том смысле, что

$$\lim_{s\to\infty} J(\hat{U}(\cdot,\lambda^s)) = J(\hat{U}(\cdot,\hat{\lambda})) \quad \text{при} \quad \lambda^s \to \hat{\lambda}.$$

Условия а), б) очевидно выполнены, в) следует из теоремы 1, а условие г) в предположении (14) доказано в [14]. Остановимся более подробно на доказательстве условия д).

Правая часть дифференциального уравнения (6), как максимум непрерывной функции по компакту, тоже будет непрерывной функцией. Следовательно, $\varphi(t,\lambda)$, как решение системы (6)–(7), непрерывно зависит от λ.

Запишем векторный функционал в виде

$$J(\hat{U}(\cdot,\lambda)) = \int_0^T g(t,\hat{u}(t,\lambda),\pi(t,\lambda))\,dt,$$

где $\hat{u}(t,\lambda) = \{\hat{r}_n(t,\lambda),\hat{\mu}_n(t,\lambda)\}_{n\in\mathsf{N}}$ обозначает оптимальное управление, соответствующее вектору λ, а $\pi(t,\lambda)$ удовлетворяет системе (4), в которую подставлено указанное управление, причем $\pi(0,\lambda)=\pi(0)$. Тогда функцию $g(\cdot)$ можно считать липшицевой

$$|g(t,u,p)-g(t,v,q)| \le C(|u-v|+|p-q|).$$

Поэтому для доказательства непрерывной зависимости $J(\hat{U}(\cdot,\lambda))$ от λ достаточно проверить, что имеет место сходимость в $\mathsf{L}_1[0,T]$:

$$\int_0^T |\hat{u}(t,\lambda^s)-\hat{u}(t,\hat{\lambda})|\,dt \to 0, \tag{16}$$

$$\int_0^T |\pi(t,\lambda^s)-\pi(t,\hat{\lambda})|\,dt \to 0. \tag{17}$$

Отметим, что (17) вытекает из (16) (см., например, [20]). Непрерывная зависимость $\hat{\mu}_n(t,\lambda)$ от λ (при каждом фиксированном t) следует из непрерывности функции (9) и непрерывности $\varphi(t,\lambda)$. Поэтому сходимость $\hat{\mu}_n(t,\lambda^s) \to \hat{\mu}_n(t,\hat{\lambda})$ в $\mathsf{L}_1[0,T]$ вытекает из теоремы Лебега.

Теперь рассмотрим

$$\int_0^T |\hat{r}_n(t,\lambda^s) - \hat{r}_n(t,\hat{\lambda})|\,dt = \\ = \mathrm{mes}\{t\colon (\psi(t,\lambda^s), \psi(t,\hat{\lambda})) \in B\}, \quad (18)$$

где $\psi(t,\lambda) = -\varphi_n(t,\lambda) + \varphi_{n+1}(t,\lambda) - \lambda_1$,

$$B = \{(x_1,x_2) \in \mathbb{R}^2 \colon x_1 > 0 \geq x_2 \text{ или } x_1 \leq 0 < x_2\}.$$

Так как $\psi(t,\lambda^s) \to \psi(t,\hat{\lambda})$ при каждом t, для перехода к пределу под знаком меры (18) достаточно:

$$\mathrm{mes}\{t\colon (\psi(t,\hat{\lambda}), \psi(t,\hat{\lambda})) \in \partial B\} = 0, \qquad (19)$$

где ∂B — граница множества B (см., например, [21]). Заметим, что условие (19) означает

$$-\varphi_n(t,\hat{\lambda}) + \varphi_{n+1}(t,\hat{\lambda}) - \hat{\lambda}_1 \neq 0 \quad \text{п.в.}, \quad n < N, \qquad (20)$$

а после предельного перехода в (18) получаем

$$\mathrm{mes}\{t\colon (\psi(t,\hat{\lambda}), \psi(t,\hat{\lambda})) \in B\} = 0,$$

что и требовалось.

Итак, доказана

Т е о р е м а 3. *Если для управляемой СМО выполнено условие Слейтера* (14), *то существует решение* $\hat{\lambda}$ *двойственной задачи* (15). *Если к тому же верно* (20), *то марковское управление, описываемое соотношениями* (8)–(9) *при* $\lambda = \hat{\lambda}$, *является оптимальным в задаче с ограничениями* (2).

Отметим, что двойственная задача представляет собой задачу выпуклого программирования и может быть записана следующим образом:

$$\hat{\lambda} \in \operatorname*{arg\,max}_{\lambda \in \Lambda} \langle \varphi(0,\lambda), \pi \rangle,$$

а множество Λ может быть выбрано компактным

$$\Lambda = \{[1,\lambda_1,\lambda_2]^* \colon \lambda_1 \geq 0, \lambda_2 \geq 0, \\ \lambda_1(\bar{J}_1 - J_1(\overline{U})) + \lambda_2(\bar{J}_2 - J_2(\overline{U})) \leq J_0(\overline{U})\}.$$

Список литературы

[1] *Serfozo R.* Optimal control of randomwalks, birth and death processes, and queues // *Advances in Applied Probability.* 1981. Vol. 13, no. 1. Pp. 61–83.

[2] *Hordijk A., Spieksma F.* Constrained admission control to a queueing system // *Advances in Applied Probability.* 1989. Vol. 21, no. 2. Pp. 409–431.

[3] *Low S. H., Paganini F., Doyle J. C.* Internet congestion control // *IEEE Control Systems Magazine.* 2002. Vol. 22, no. 1. Pp. 28–43.

[4] *Howard R.* Dynamic programming and Markov processes. New York: J. Wiley & Sons, 1960.

[5] *Дынкин Е. Б., Юшкевич А. А.* Управляемые марковские процессы и их приложения. М.: Наука, 1975.

[6] *Pliska S.* Controlled jump processes // *Stochastic processes and their applications.* 1975. Vol. 3. Pp. 259–282.

[7] *Elliott R. J., Aggoun L., Moore J. B.* Hidden Markov models. Estimation and control. New York: Springer-Verlag, 1995.

[8] *Liptser R. S., Shiryayev A. N.* Statistics of random processes. 3rd ed. edition. New York: Springer-Verlag, 2005.

[9] *Miller B. M.* Optimization of queuing system via stochastic control // *Automatica.* 2009. Vol. 45. Pp. 1423–1430.

[10] *Altman E.* Constrained Markov decision processes. Boca Raton, FL: Chapman & Hall/CRC, 1999.

[11] *Пиуновский А. Б.* Управляемые случайные последовательности: методы выпуклого анализа и задачи с функциональными ограничениями // *Успехи матем. наук.* 1998. Т. 53, № 6. С. 129–192.

[12] *Piunovskiy A. B.* Bicriteria optimization of a queue with a controlled input stream // *Queueing Systems.* 2004. Vol. 48. Pp. 159–184.

[13] *Miller B., Miller G., Siemenikhin K.* Towards the optimal control of Markov chains with constraints // *Automatica.* 2010. Vol. 46, no. 9. Pp. 1495–1502.

[14] *Миллер Б. М., Миллер Г. Б., Семенихин К. В.* Методы синтеза оптимального управления марковским процессом с конечным множеством состояний при наличии ограничений // *Автоматика и телемеханика.* 2011. № 2. С. 111–130.

[15] *Флеминг У., Ришел Р.* Оптимальное управление детерминированными и стохастическими системами. М.: Мир, 1978.

[16] *Иоффе А. Д., Тихомиров В. М.* Теория экстремальных задач. М.: Наука, 1974.

[17] *Verdú S., Poor H. V.* On minimax robustness: A general approach and applications // *IEEE Trans. Inform. Theory.* 1984. Vol. IT-30, no. 2. Pp. 328–340.

[18] *Соловьёв В. Н.* Двойственные экстремальные задачи и их применение к задачам минимаксного оценивания // *Успехи матем. наук.* 1997. Т. 52, № 4. С. 49–86.

[19] *Семенихин К. В.* Минимаксное оценивание в неопределенно-стохастических моделях линейной регрессии. М.: Изд-во МАИ, 2011.

[20] *Лебедев М. В., Семенихин К. В.* Минимаксная фильтрация в стохастической дифференциальной системе с нестационарными возмущениями неизвестной интенсивности // *Известия РАН. Теория и системы управления.* 2007. № 2. С. 45–56.

[21] *Ширяев А. Н.* Вероятность. М.: Наука, 1980.