Автоматика и телемеханика, № 1, 2019

Стохастические системы

© 2019 г. **В.М. АЗАНОВ (azanov59@gmail.com),**
Ю.С. КАН, д-р физ.-мат. наук (yu_kan@mail.ru)
(Московский авиационный институт)

ОБ ОПТИМАЛЬНОМ УДЕРЖАНИИ ТРАЕКТОРИИ ДИСКРЕТНОЙ СТОХАСТИЧЕСКОЙ СИСТЕМЫ В ТРУБКЕ[1]

Рассматривается задача синтеза оптимального управления дискретной стохастической системой общего вида с критерием в форме вероятности пребывания вектора состояния в каждый момент времени на заданных множествах. Выводятся соотношения метода динамического программирования, позволяющие найти оптимальное решение в классе марковских стратегий без расширения вектора состояния с последующим сведением к эквивалентной задаче с вероятностным терминальным критерием. Рассмотрена задача однопараметрической коррекции траектории движения летательного аппарата. Получено аналитическое решение.

Ключевые слова: дискретные системы, стохастическое оптимальное управление, вероятностный критерий, метод динамического программирования, однопараметрическая импульсная коррекция, управление движением летательного аппарата.

DOI: 10.1134/S0005231019010033

1. Введение

Дискретные стохастические системы обсуждаются во многих прикладных задачах синтеза оптимального управления, например, в аэрокосмической области при импульсной коррекции положения летательного аппарата [1–7], в экономической области в проблеме оптимального управления портфелем ценных бумаг [8–10], в области робототехники в задачах автоматизированного управления [11–13], в информатике в задачах оптимизации распределения внешних и внутренних вычислительных ресурсов программных систем [14–16].

Критерий в форме функционала вероятности характерен больше для аэрокосмических [1–4], экономических [8–10] и робототехнических [11–13] задач, в которых требуется оптимизировать вероятность обеспечения желаемой точности системы при ограничении на ресурс управления, при этом точностной функционал задан на траекториях системы в терминальный момент времени. Альтернативой вероятностному критерию служит среднеквадратический критерий, нацеленный на выполнение точностного ограничения с учетом неравенства Чебышева. Недостатки такого подхода, связанные с грубостью неравенства Чебышева, подробно освещены в [1].

[1] Работа, за исключением раздела 4, выполнена при поддержке Российского научного фонда (проект № 16-11-00062). Результаты раздела 4 получены при поддержке Российского фонда фундаментальных исследований (проект № 18-08-00595).

Во многих задачах в качестве критерия рассматривается точностной функционал, зависящий не только от терминального состояния, но и от состояний во все моменты дискретного времени. Критерий такого вида играет важную роль в приложениях как с точки зрения задач оптимизации точности (см., например, [17, 18]), так и оптимизации ресурса управления (см., например [14–16]).

В настоящей статье исследуется задача оптимального управления дискретной стохастической системой по критерию максимума вероятности попадания траектории системы в заданную трубку. Доказывается теорема о существовании оптимальной стратегии в классе марковских и выводится уравнение Беллмана. На примере задачи однопараметрической коррекции сравниваются два подхода: первый — с использованием уравнения Беллмана и второй — с использованием приема расширения пространства состояний и сведения исходной задачи к эквивалентной в определенном смысле задаче с вероятностным терминальным критерием.

2. Постановка задачи

Пусть динамика объекта управления описывается разностным уравнением

$$\begin{cases} x_{k+1} = f_k\left(x_k, u_k, \xi_k\right), \\ x_0 = X, \end{cases} \qquad k = \overline{0, N}, \tag{1}$$

где $x_k \in \mathbb{R}^n$ — вектор состояния, $u_k \in U_k \subset \mathbb{R}^m$ — вектор управления, U_k — множество ограничений на управление, ξ_k — вектор случайных возмущений с значениями на $\mathbb{R}^s$, $f_k : \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^s \to \mathbb{R}^n$ — функция перехода (функция системы), $N \in \mathbb{N}$ — горизонт управления.

В отношении системы (1) введем ряд предположений:

1) известна полная информация о векторе состояния x_k (данный факт позволяет строить управление в классе функций $u_k = \gamma_k\left(x_k\right)$, где $\gamma_k\left(\cdot\right)$ — некоторая измеримая функция. В данном случае говорят, что "управление ищется в классе полной обратной связи по состоянию";
2) начальное состояние $x_0 = X$ является в общем случае случайным вектором с значениями в $\mathbb{R}^n$ и с известным распределением $\mathbf{P}_X$;
3) функция системы $f_k\left(x_k, u_k, \xi_k\right)$ непрерывна для всех k;
4) вектор управления u_k формируется следующим образом: $u_k = \gamma_k\left(x_k\right)$, где $\gamma_k : \mathbb{R}^n \to \mathbb{R}^m$ — измеримая функция с ограниченными значениями $u_k \in U_k$, причем U_k — компактное множество;
5) вектор состояния x_{k+1} формируется следующим образом: на шаге k реализуется вектор x_k, далее формируется вектор управления $u_k = \gamma_k\left(x_k\right)$ и в последнюю очередь реализуется случайное возмущение ξ_k;
6) управлением называется набор функций $u(\cdot) = (\gamma_0(\cdot), \ldots, \gamma_N(\cdot)) \in \mathcal{U}$, классом допустимых управлений называется множество $\mathcal{U} = \mathcal{U}_0 \times \ldots \times \mathcal{U}_N$, где $\mathcal{U}_k$ — множество борелевских функций $\gamma_k\left(\cdot\right)$ с ограниченными на U_k значениями;

7) случайный вектор ξ_k является непрерывным с значениями в $\mathbb{R}^s$ и известным распределением $\mathbf{P}_k$, причем компоненты вектора $\zeta = (X, \xi_0, \ldots, \xi_N)$ независимы.

Заметим, что система (1) является марковской, т.е. ее поведение в будущем не зависит от прошлого и полностью определяется текущим состоянием.

На траекториях системы (1) определим функционал вероятности

$$P_\varphi(u(\cdot)) = \mathbf{P}\left(\bigcap_{k=0}^{N}\{x_{k+1} \in \mathcal{F}_{k+1}\}\right),$$

множества $\mathcal{F}_k$ имеют вид

$$\mathcal{F}_k = \{x \in \mathbb{R}^n : \quad \Phi_k(x) \leqslant \varphi\},$$

где $\varphi \in \mathbb{R}$ — известный скаляр, $\Phi_k : \mathbb{R}^n \to \mathbb{R}$ — непрерывные функции, $k = 1, \ldots, N+1$, причем $\Phi_{N+1}(x)$ ограничена снизу.

Рассматривается задача

$$P_\varphi(u(\cdot)) \to \max_{u(\cdot)\in\mathcal{U}}, \tag{2}$$

где $\mathcal{U} = \mathcal{U}_0 \times \ldots \times \mathcal{U}_N$.

Задача (2) является задачей синтеза оптимального управления в классе позиционных стратегий. Отметим, что если принять

$$\mathcal{F}_k = \mathbb{R}^n, \quad k = \overline{1, N},$$

то получаем задачу синтеза оптимального управления с вероятностным терминальным критерием в классической постановке [1, 3].

Для синтеза оптимальной стратегии может быть использован прием, примененный, например, в [19], сведения задачи (2) к задаче оптимального управления с вероятностным терминальным критерием с расширенным пространством состояний размерности $n+1$ (сказанное поясняется в разделе 4 на примере простейшей задачи импульсной коррекции траектории движения летательного аппарата). Такой подход влечет за собой усложнение как численного, так и аналитического синтеза оптимальной стратегии. Кроме того, в некоторых случаях аналитическое решение "приведенной задачи" с вероятностным терминальным критерием и вовсе затруднено, в то время как решение исходной задачи возможно.

В разделе 3 выводится уравнение Беллмана для задачи (2).

3. Условия оптимальности в форме уравнения Беллмана

Для удобства примем

$$\mathcal{F}_0 = \mathbb{R}^n.$$

Определим функцию Беллмана $\mathsf{B}_k : \mathbb{R}^n \to [0,1]$ в задаче (2) как

$$\mathsf{B}_k(x) =$$
$$= \sup_{\gamma_k(\cdot)\in\mathcal{U}_k,\ldots,\gamma_N(\cdot)\in\mathcal{U}_N} \mathbf{P}\left(\max_{i=\overline{k,N}} \Phi_{i+1}(x_{i+1}(x_k,\gamma_k(\cdot),\ldots,\gamma_i(\cdot),\xi_k,\ldots,\xi_i)) \leqslant \varphi \Big| x_k = x\right).$$

Принимая во внимание сделанные в разделе 2 предположения, сформулируем теорему об уравнении Беллмана для задачи (2) в пространстве состояний размерности n.

Теорема. Пусть выполнены условия:

1) *функции* $f_k(x_k,u_k,\xi_k)$*непрерывны для всех* $k=\overline{0,N}$*;*
2) *функции* $\Phi_k(x_k)$ *непрерывны для всех* $k=\overline{1,N+1}$*;*
3) *функция* $\Phi_{N+1}(x_{N+1})$ *ограничена снизу;*
4) *случайные векторы* $X,\xi_0,\ldots,\xi_N$ *независимы;*
5) *множества* $U_0,\ldots,U_N$ *компактны.*

Тогда оптимальная стратегия в задаче (2) *существует в классе измеримых функций* $u^*(\cdot)\in\mathcal{U}$ *и определяется в результате решения следующих задач*:

(3) $$u_k^* = \arg\max_{u_k\in U_k} \mathbf{M}\left[\mathbf{I}_{\mathcal{F}_k}(x_k)\,\mathsf{B}_{k+1}(f_k(x_k,u_k,\xi_k))\,\Big|x_k\right],$$

(4) $$\mathsf{B}_k(x) = \max_{u_k\in U_k} \mathbf{M}\left[\mathbf{I}_{\mathcal{F}_k}(x_k)\,\mathsf{B}_{k+1}(f_k(x_k,u_k,\xi_k))\,\Big|x_k = x\right], \quad k=\overline{0,N},$$

(5) $$\mathsf{B}_{N+1}(x) = \mathbf{I}_{\mathcal{F}_{N+1}}(x).$$

Доказательство теоремы вынесено в Приложение.

Таким образом, получено уравнение Беллмана в пространстве состояний размерности n. Оно отличается от уравнения Беллмана для задачи с вероятностным терминальным критерием наличия в правой части сомножителя в виде индикаторной функции $\mathbf{I}_{\mathcal{F}_k}(x_k)$ множества $\mathcal{F}_k$.

4. Задача однопараметрической коррекции траектории летательного аппарата

4.1. Постановка задачи с использованием разных подходов

На примере простой одномерной задачи сравнивается подход к решению задачи (2), описанный в разделе 3 с другим подходом, основанным на сведении задачи (2) к эквивалентной (в определенном смысле) задаче оптимального управления с вероятностным терминальным критерием и, как следствие, на использовании метода динамического программирования в классической форме [1].

Пусть $n=m=s=1$ — размерности вектора состояния, вектора управления и вектора случайного возмущения $\xi_k \sim \mathcal{R}[-\varepsilon,\varepsilon]$, $\varepsilon\in(0,1)$, $\{r_k\}_{k=1}^{N+1}$ — убывающая последовательность вещественных чисел, $r_k \geqslant r_{k+1}$, $r_{N+1}>0$.

Рассмотрим задачу

$$\mathbf{P}\left(\bigcap_{k=0}^{N}\left\{|x_{k+1}|\leqslant r_{k+1}\right\}\right)\to\max_{u(\cdot)}, \tag{6}$$

$$\begin{cases} x_{k+1}=x_k+u_k\left(1+\xi_k\right), \\ x_0=X, \end{cases} \qquad k=\overline{0,N}, \tag{7}$$

и задачу

$$\mathbf{P}\left(y_{N+1}\leqslant 1\right)\to\max_{u(\cdot)}, \tag{8}$$

$$\begin{cases} x_{k+1}=x_k+u_k\left(1+\xi_k\right), \\ y_{k+1}=\max\left\{y_k, \quad \left|x_k+u_k\left(1+\xi_k\right)\right|/r_{k+1}\right\}, \\ x_0=X, \\ y_0=\left|X+u_0\left(1+\xi_0\right)\right|/r_1, \end{cases} \qquad k=\overline{0,N}, \tag{9}$$

которые, очевидно, являются эквивалентными в смысле критериев

$$\mathbf{P}\left(y_{N+1}\leqslant 1\right)=\mathbf{P}\left(\max_{k=\overline{0,N}}\left\{\frac{|x_{k+1}|}{r_{k+1}}\leqslant 1\right\}\right)=\mathbf{P}\left(\bigcap_{k=0}^{N}\left\{|x_{k+1}|\leqslant r_{k+1}\right\}\right).$$

Задача (6) относится к классу задач, описанных в разделах 2 и 3 настоящей статьи, а задача (8) — к классу задач оптимального управления с вероятностным терминальным критерием [1]. Отметим, что в задаче (8) управление ищется в более широком классе функций $u_k=\gamma_k\left(x_k,y_k\right)$, при этом для поиска оптимальной стратегии примени́м метод динамического программирования в классической форме [1], а также его модификации (см. [3, 20]).

Важно отметить, что задача (6) возникает естественным образом в аэрокосмических приложениях [1, 3] в рамках задач однопараметрической импульсной коррекции траектории движения космических аппаратов. Принципиальным отличием от постановок с терминальным критерием качества $\mathbf{P}\left(|x_{N+1}|\leqslant\varphi\right)\to\max_{u(\cdot)}$ является то, что корректирующие импульсы должны обеспечивать “равномерную” во времени коррекцию параметра движения, что в постановке (6) формализуется в виде ограничения $x_{k+1}\leqslant r_{k+1}$ в каждый момент времени, где убывающая положительная последовательность чисел $\{r_{k+1}\}_{k=0}^{N}$ имеет смысл максимально допустимого промаха состояния x_{k+1}. Подобный смысл в среднеквадратической задаче $\mathbf{M}\left[\left(x_{N+1}\right)^2\right]\to\min_{u(\cdot)}$ несут в себе изопериметрические ограничения на вектор состояния в каждый момент времени $\mathbf{M}\left[\left(x_{k+1}\right)^2\right]\leqslant\tilde{r}_{k+1}$. Отметим, что задача с терминальным критерием в более общей постановке с векторной системой и ограниченным скалярным управлением аналитически решена в [2], а одномерный случай без ограничений рассматривался в [1, 3].

4.2. Решение задачи (6)

Перейдем к решению задачи (6). В исходных обозначениях имеем

$$f_k(x_k, u_k, \xi_k) = x_k + u_k(1 + \xi_k), \quad \Phi_k(x) = |x|, \quad \mathcal{F}_k = [-r_k, r_k].$$

В утверждении 1 с использованием теоремы найдено аналитическое решение задачи (6) для случая, когда граница носителя распределения случайной величины ξ_k связана естественными ограничениями с границей множества $\mathcal{F}_k$.

Утверждение 1. *Пусть выполнено*

$$\varepsilon \leqslant \min_{k=\overline{1,N}} \frac{r_{k+1}}{r_k}, \tag{10}$$

тогда функция Беллмана в задаче (6) *имеет вид*

$$\mathsf{B}_k(x) = \begin{cases} \mathbf{I}_{[-r_k, r_k]}(x), & k = \overline{1,N}, \\ \min\left\{1, \quad \frac{r_1}{r_1 + |x|}\frac{1+\varepsilon}{\varepsilon}\right\}, & k = 0, \end{cases}$$

и оптимальное управление определяется выражением

$$u_k^* = \begin{cases} \text{любое из } U_k^*(x_k), & |x_k| \in [0, r_k], \\ \text{любое из } \mathbb{R}, & |x_k| \in (r_k, +\infty), \end{cases} \qquad k = \overline{1,N},$$

$$u_0^* = \begin{cases} \text{любое из } U_0^*(X), & |X| \in \left[0, r_1\varepsilon^{-1}\right], \\ \dfrac{-\mathsf{sign}(X)\, r_1 - X}{1+\varepsilon}, & |X| \in \left(r_1\varepsilon^{-1}, +\infty\right), \end{cases} \qquad k = 0,$$

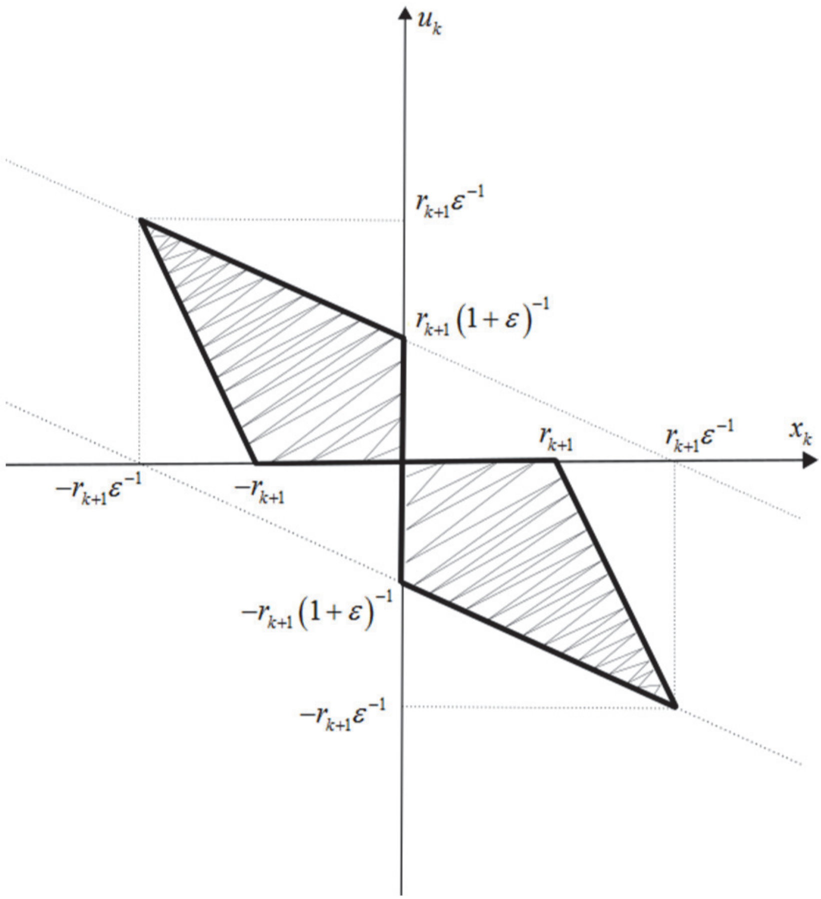
где

$$U_k^*(x_k) = \Big\{u_k \in \mathbb{R}: \quad \mathsf{sign}(u_k) = -\mathsf{sign}(x_k),$$
$$\max\left\{0, \ \frac{-r_{k+1} + |x_k|}{1-\varepsilon}\right\} \leqslant |u_k| \leqslant \frac{r_{k+1} + |x_k|}{1+\varepsilon}\Big\}.$$

Доказательство утверждения 1 вынесено в Приложение.

Основным условием утверждения 1 является неравенство (10). Оно очевидно выполнено в случае, когда $r_k = \mathrm{const}$ для всех k, так как из постановки задачи следует, что $\varepsilon < 1$. Как видно из утверждения 1, функция Беллмана является индикаторной (на шагах $k = \overline{1,N}$) и выполнено $\mathcal{F}_k = \{x \in \mathbb{R}: \quad \mathsf{B}_k(x) = 1\}$, что означает, что цели управления выполняются с вероятностью единица на шагах $k = \overline{1,N}$, при этом оптимальным является целое множество управлений на указанных шагах. Важно отметить, что оптимальным при $x_k \in [-r_k, r_k]$ является любой элемент из множества $U_k^*(x_k)$. Указанное множество изображено на рисунке.

Интересным является тот факт, что при $k = 0$ функция Беллмана представляет собой максимальную вероятность попадания траекторией системы



Множество $U_k^*(x_k)$.

на множество $\mathcal{F}_1$, т.е. $\mathsf{B}_0(x) = \max_{u_0} \mathbf{P}(f_0(x, u_0, \xi_0) \in \mathcal{F}_1)$, что означает, что при фиксированном начальном состоянии максимальная вероятность пребывания системы на множествах $\mathcal{F}_{k+1}$, $k = \overline{0, N}$, равна максимальной вероятности попадания системы на множество $\mathcal{F}_1$. Упомянутый факт обусловлен тем, что при выполнении неравенства (10) любым управлением из $U_k^*(x_k)$ удается с вероятностью единица перевести систему из $\mathcal{F}_k$ на множество $\mathcal{F}_{k+1}$.

4.3. Решение задачи (8)

Перейдем теперь к решению задачи (8). В соответствии с [1] функция Беллмана $\tilde{\mathsf{B}}_k : \mathbb{R}^{n+1} \to [0,1]$ для (8) равна:

$$\tilde{\mathsf{B}}_k(x,y) =$$
$$= \sup_{\gamma_k(\cdot),\ldots,\gamma_N(\cdot)} \mathbf{P}\left(y_{N+1}(x_k, y_k, \gamma_k(\cdot), \ldots, \gamma_N(\cdot), \xi_k, \ldots, \xi_N) \leqslant 1 \middle| x_k = x,\ y_k = y\right),$$

а уравнения метода динамического программирования имеют вид:

$$\tilde{u}_k^* = \arg\max_{u_k} \mathbf{M}\left[\tilde{\mathsf{B}}_{k+1}\left(\tilde{f}_k\left(x_k, y_k, u_k, \xi_k\right)\right)\Big| x_k, y_k\right], \quad k = \overline{0, N},$$

$$\tilde{\mathsf{B}}_k\left(x, y\right) = \max_{u_k} \mathbf{M}\left[\tilde{\mathsf{B}}_{k+1}\left(\tilde{f}_k\left(x_k, y_k, u_k, \xi_k\right)\right)\Big| x_k = x,\ y_k = y\right],$$

$$\tilde{\mathsf{B}}_{N+1}\left(x, y\right) = \mathbf{I}_{(-\infty,1]}\left(y\right),$$

где $\tilde{f}_k : \mathbb{R}^{n+1} \to \mathbb{R}^{n+1}$ — функция правой части “расширенной” системы управления

$$\tilde{f}_k\left(x_k, y_k, u_k, \xi_k\right) = \begin{pmatrix} x_k + u_k\left(1 + \xi_k\right) \\ \max\left\{y_k, \quad \dfrac{\left|x_k + u_k\left(1 + \xi_k\right)\right|}{r_{k+1}}\right\} \end{pmatrix}.$$

В следующем утверждении 2 найдено аналитическое решение задачи (8) при условии (10): оптимальное управление и функция Беллмана.

Утверждение 2. Пусть для задачи (8) *выполнено условие* (10)*. Тогда фукнция Беллмана имеет вид*

$$\tilde{\mathsf{B}}_k\left(x, y\right) = \begin{cases} \mathbf{I}_{(-\infty,1]}\left(y\right)\min\left\{1, \quad \dfrac{r_{k+1}}{r_{k+1} + |x|}\dfrac{1+\varepsilon}{\varepsilon}\right\}, & k = \overline{1, N}, \\ \min\left\{1, \quad \dfrac{r_1}{r_1 + |x|}\dfrac{1+\varepsilon}{\varepsilon}\right\}, & k = 0, \end{cases}$$

и оптимальное управление определяется выражением

$$\tilde{u}_k^* = \begin{cases} \text{любое из } U_k^*\left(x_k\right), & |x_k| \leqslant r_{k+1}\varepsilon^{-1},\ y_k \leqslant 1, \\ \dfrac{-\mathsf{sign}\left(x_k\right) r_{k+1} - x_k}{1+\varepsilon}, & |x_k| > r_{k+1}\varepsilon^{-1},\ y_k \leqslant 1, \\ \text{любое из } \mathbb{R}, & y_k > 1, \end{cases} \quad k = \overline{1, N},$$

$$u_0^*\left(X\right) = \begin{cases} \text{любое из } U_0^*\left(X\right), & |X| \in \left[0, r_1\varepsilon^{-1}\right], \\ \dfrac{-\mathsf{sign}\left(X\right) r_1 - X}{1+\varepsilon}, & |X| \in \left(r_1\varepsilon^{-1}, +\infty\right), \end{cases} \quad k = 0,$$

где множество $U_k^*\left(x_k\right)$ определяется в соответствии с утверждением 1.

Доказательство утверждения 2 вынесено в Приложение.

Нетрудно видеть, что функция Беллмана в задаче (8) при $k = 0$ не зависит от y_k и совпадает с функцией Беллмана в задаче (6), что подтверждает факт эквивалентности задач (6) и (8) в смысле оптимальных значений критериев. Покажем теперь, что в силу специфики оптимальных управлений в обеих задачах (а именно: для некоторых областей пространства состояний оптимальным является целое множество стратегий) эквивалентность задач (6) и (8) может быть усилена.

4.4. Сравнение подходов

Из утверждений 1 и 2 видно, что функция Беллмана в задаче (8) при $k = 0$ не зависит от y_k и совпадает с функцией Беллмана в задаче (6). Таким образом, функции оптимальных значений вероятностных критериев в этих задачах совпадают. Более того, эти функции зависят лишь от распределения начального состояния X, параметра ε и r_1. Покажем теперь, что оптимальное управление в задаче (8) можно выбрать таким образом (в силу того, что оптимальным является целое множество управлений при $y_k \leqslant 1$), что оно совпадает с оптимальным управлением в задаче (6). Для этого выберем управление (с учетом специфики множества $U_k^*(x_k)$) в задаче (6) в следующем виде:

$$(11) \qquad u_k^* = \begin{cases} 0, & |x_k| \in [0, r_k], \\ \dfrac{\mathsf{sign}(x_k)\, r_{k+1} - x_k}{1-\varepsilon}, & |x_k| \in \left(r_k, r_k\varepsilon^{-1}\right], \\ \dfrac{-\mathsf{sign}(x_k)\, r_{k+1} - x_k}{1+\varepsilon}, & |x_k| \in \left(r_k\varepsilon^{-1}, +\infty\right), \end{cases} \qquad k = \overline{0, N}.$$

Нетрудно видеть, что в задаче (8) в области $y_k > 1$, $k = \overline{1, N}$, оптимальным является любое управление из $\mathbb{R}$, и, следовательно, оно может быть выбрано как

$$(12) \quad \tilde{u}_k^* = u_k^* = \begin{cases} 0, & |x_k| \in [0, r_{k+1}], \\ \dfrac{\mathsf{sign}(x_k)\, r_{k+1} - x_k}{1-\varepsilon}, & |x_k| \in \left(r_{k+1}, r_{k+1}\varepsilon^{-1}\right], \\ \dfrac{-\mathsf{sign}(x_k)\, r_{k+1} - x_k}{1+\varepsilon}, & |x_k| \in \left(r_{k+1}\varepsilon^{-1}, +\infty\right), \end{cases} \quad k = \overline{0, N}.$$

Таким образом, оптимальным в задаче (8) может быть выбрано управление, оптимальное в задаче (6) и не зависящее явно от y_k.

5. Заключение

В статье рассмотрена задача оптимального управления дискретной стохастической системой с критерием в форме вероятности пребывания траектории системы на заданных множествах. Найдены достаточные условия существования оптимального управления в классе измеримых функций, а также выведены соотношения метода динамического программирования, отличающиеся от “классических”, справедливых для задач с вероятностным терминальным критерием. С помощью этих соотношений аналитически решена задача однопараметрической коррекции траектории движения летательного аппарата в постановке с ограничениями на траекторию системы. Показана эквивалентность в смысле оптимальных стратегий и оптимальных значений критериев указанной задачи и аналогичной задачи с вероятностным терминальным критерием и расширенным пространством состояний.

Доказательство теоремы. Покажем, что уравнения МДП естественным образом получаются из эквивалентной задачи с вероятностным терминальным критерием и расширенным пространством состояний. Для этого введем в рассмотрение переменную $y_k \in \mathbb{R}$:

$$y_k = \max_{i=\overline{0,k}} \Phi_i\left(x_i\right).$$

Рассмотрим расширенную систему управления

$$\begin{cases} x_{k+1} = f_k\left(x_k, u_k, \xi_k\right), \\ y_{k+1} = \max\left\{y_k,\ \Phi_k\left(x_k\right)\right\}, \\ x_0 = X, \\ y_0 = \Phi_0\left(X\right), \end{cases} \qquad k = \overline{0, N},$$

где $\Phi_0 : \mathbb{R}^n \to (-\infty, \varphi]$ — некоторая ограниченная сверху функция (отсюда следует, что $\mathcal{F}_0 = \mathbb{R}^n$). На траекториях расширенной системы рассмотрим задачу

$$\text{(П.1)} \qquad \mathbf{P}\left(\max\left\{y_{N+1},\ \Phi_{N+1}\left(x_{N+1}\right)\right\} \leqslant \varphi\right) \to \max_{\tilde{u}(\cdot)\in\tilde{\mathcal{U}}},$$

где управление ищется в более широком классе функций $\tilde{u}_k = \tilde{u}_k\left(x_k, y_k\right) \in \tilde{\mathcal{U}}_k$, $\tilde{\mathcal{U}} = \tilde{\mathcal{U}}_0 \times \ldots \times \tilde{\mathcal{U}}_N$, $\tilde{\mathcal{U}}_k$ — множество борелевских функций с ограниченными на U_k значениями. Запишем уравнения МДП для задачи (П.1):

$$\text{(П.2)} \qquad \tilde{u}_k^* = \arg\max_{u_k\in U_k} \mathbf{M}\left[\tilde{\mathsf{B}}_{k+1}\left(\tilde{f}_k\left(x_k, y_k, u_k, \xi_k\right)\right) \Big| x_k,\ y_k\right],$$

$$\text{(П.3)} \qquad \tilde{\mathsf{B}}_k\left(x, y\right) = \max_{u_k\in U_k} \mathbf{M}\left[\tilde{\mathsf{B}}_{k+1}\left(\tilde{f}_k\left(x_k, y_k, u_k, \xi_k\right)\right) \Big| x_k = x,\ y_k = y\right],$$

$$\text{(П.4)} \qquad \tilde{\mathsf{B}}_{N+1}\left(x, y\right) = \mathbf{I}_{(-\infty,\varphi]}\left(\max\left\{y,\ \Phi_{N+1}\left(x\right)\right\}\right).$$

В [8, теорема 3, с. 143] приведены достаточные условия, при которых оптимальная стратегия существует в классе измеримых функций для задачи оптимального управления с вероятностным терминальным критерием. В соответствии с этими достаточными условиями функция правой части системы должна быть непрерывной для всех $k = \overline{0, N}$, случайные векторы $X, \xi_0, \ldots, \xi_N$ независимы, функция терминального состояния (так называемый точностной функционал) полунепрерывен снизу, а множества $U_0, \ldots, U_N$ компактны. Применим эти условия к задаче (П.1). Получаем, что если выполнены условия:

1) функции $f_k\left(x_k, u_k, \xi_k\right)$ непрерывны для всех $k = \overline{0, N}$;
2) функции $\Phi_k\left(x_k\right)$ непрерывны для всех $k = \overline{1, N+1}$;
3) функция $\Phi_{N+1}\left(x_{N+1}\right)$ ограничена снизу;
4) случайные векторы $X, \xi_0, \ldots, \xi_N$ независимы;
5) множества $U_0, \ldots, U_N$ компактны,

то оптимальная стратегия существует в классе измеримых функций $\tilde{u}^*(\cdot) \in \tilde{\mathcal{U}}$ и определяется в результате решения уравнения Беллмана в форме (П.2). Покажем теперь, что само уравнение Беллмана допускает упрощение. Пусть $k = N$, тогда

$$\begin{aligned}\tilde{\mathsf{B}}_N(x,y) &= \max_{u_N \in U_N} \mathbf{P}\left(\max\left\{y,\ \Phi_N(x),\ \Phi_{N+1}\left(f_N(x,u_N,\xi_N)\right)\right\} \leqslant \varphi\right) = \\ &= \max_{u_N \in U_N} \mathbf{I}_{(-\infty,\varphi]}(y)\, \mathbf{I}_{(-\infty,\varphi]}\left(\Phi_N(x)\right) \mathbf{P}\left(\Phi_{N+1}\left(f_N(x,u_N,\xi_N)\right) \leqslant \varphi\right) = \\ &= \mathbf{I}_{(-\infty,\varphi]}(y)\, \mathsf{B}_N(x),\end{aligned}$$

при этом

$$\tilde{u}_N^* = \begin{cases} u_N^*, & y_N \leqslant \varphi, \\ \text{любое из } U_N, & y_N > \varphi, \end{cases}$$

где u_N^* определяется соотношением (3).

Пусть теперь $k = N - 1$, тогда имеем

$$\begin{aligned}\tilde{\mathsf{B}}_{N-1}(x,y) &= \max_{u_{N-1} \in U_{N-1}} \mathbf{M}\Big[\mathbf{I}_{(-\infty,\varphi]}\left(\max\left\{y_{N-1},\ \Phi_{N-1}(x_{N-1})\right\}\right) \times \\ &\qquad \times \mathsf{B}_N\left(f_{N-1}(x_{N-1},u_{N-1},\xi_{N-1})\right) \Big| x_{N-1} = x,\ y_{N-1} = y\Big] = \\ &= \max_{u_{N-1} \in U_{N-1}} \mathbf{I}_{(-\infty,\varphi]}(y)\, \mathbf{M}\Big[\mathbf{I}_{(-\infty,\varphi]}\left(\Phi_{N-1}(x_{N-1})\right) \times \\ &\qquad \times \mathsf{B}_N\left(f_{N-1}(x_{N-1},u_{N-1},\xi_{N-1})\right) \Big| x_{N-1} = x,\ y_{N-1} = y\Big] = \\ &= \mathbf{I}_{(-\infty,\varphi]}(y)\, \mathsf{B}_{N-1}(x).\end{aligned}$$

Продолжая аналогичные размышления, получаем, что для всех $k = \overline{0, N}$ выполнено равенство

$$\tilde{\mathsf{B}}_k(x,y) = \mathbf{I}_{(-\infty,\varphi]}(y)\, \mathsf{B}_k(x),$$

причем

$$\mathsf{B}_k(x) = \max_{u_k \in U_k} \mathbf{M}\left[\mathbf{I}_{(-\infty,\varphi]}\left(\Phi_k(x_k)\right) \mathsf{B}_{k+1}\left(f_k(x_k,u_k,\xi_k)\right) \Big| x_k = x\right]$$

и

$$\tilde{u}_k^* = \begin{cases} u_k^*, & y_k \leqslant \varphi, \\ \text{любое из } U_k, & y_k > \varphi. \end{cases}$$

Таким образом, доказана справедливость соотношений (3). Заметим теперь, что компонента y_k не влияет на выбор оптимального управления и, более того, в силу того, что при $y_k > \varphi$ оптимальным является любое управление из U_k и выполнено $\tilde{\mathsf{B}}_k(x, y) = 0$, получаем, что управление $u_k^*(x_k)$, $k = \overline{0, N}$, является оптимальным в задаче (П.1).

Теорема доказана.

Доказательство утверждения 1. Примéним метод динамического программирования из теоремы к решению задачи (6). На шаге $k = N$ получаем задачу стохастического программирования

$$\mathsf{B}_N(x) = \max_{u_N} \mathbf{I}_{[-r_N, r_N]}(x)\,\mathbf{P}\left(\frac{|x + u_N(1+\xi_N)|}{r_{N+1}} \leqslant 1\right),$$

решение которой (при $|x| \leqslant r_N$) известно (см. [3, утверждение 1, с. 117]) с точностью до параметров:

$$u_N^* = \begin{cases} \text{любое из } U_N^*(x_N), & |x_N| \leqslant \min\left\{r_N,\ r_{N+1}\varepsilon^{-1}\right\}, \\ \dfrac{-\mathsf{sign}(x_N)\,r_{N+1} - x_N}{1+\varepsilon}, & |x_N| > r_{N+1}\varepsilon^{-1},\ |x_N| \leqslant r_N, \\ \text{любое из } \mathbb{R}, & |x_N| > r_N, \end{cases}$$

$$\mathsf{B}_N(x) = \mathbf{I}_{[-r_N, r_N]}(x)\min\left\{1,\quad \frac{r_{N+1}}{r_{N+1} + |x|}\frac{1+\varepsilon}{\varepsilon}\right\}.$$

Последнее выражение допускает следующее представление

$$\mathsf{B}_N(x) = \mathbf{I}_{[-r_N, r_N]}(x)\Bigg(\mathbf{I}_{[-r_{N+1}\varepsilon^{-1}, r_{N+1}\varepsilon^{-1}]}(x) + {} \\ + \mathbf{I}_{(-\infty, -r_{N+1}\varepsilon^{-1}) \cup (r_{N+1}\varepsilon^{-1}, +\infty)}(x)\,\frac{r_{N+1}}{r_{N+1} + |x|}\frac{1+\varepsilon}{\varepsilon}\Bigg),$$

откуда с учетом неравенства (10) (из которого следует $r_N \leqslant r_{N+1}\varepsilon^{-1}$) получаем:

$$u_N^* = \begin{cases} \text{любое из } U_N^*(x_N), & |x_N| \leqslant r_N, \\ \text{любое из } \mathbb{R}, & |x_N| > r_N, \end{cases}$$

$$\mathsf{B}_N(x) = \mathbf{I}_{[-r_N, r_N]}(x).$$

Перейдем теперь к шагу $k = N - 1$, на котором в соответствии с МДП получаем

$$\mathsf{B}_{N-1}(x) = \max_{u_{N-1}} \mathbf{I}_{[-r_{N-1}, r_{N-1}]}(x)\,\mathbf{P}\left(\frac{|x + u_{N-1}(1+\xi_{N-1})|}{r_N} \leqslant 1\right).$$

Нетрудно видеть, что правая часть уравнения МДП при $k = N - 1$ совпадает с правой частью уравнения МДП при $k = N$ с точностью до параметров r_k, $k = \overline{N-1, N+1}$, откуда с учетом (10) получаем аналогичное решение и функцию Беллмана. Продолжая рассуждения по аналогии, заключаем, что оптимальное управление и функция Беллмана на шагах $k = \overline{1, N}$ имеют вид:

$$u_k^* = \begin{cases} \text{любое из } U_k^*(x_k), & |x_k| \in [0, r_k], \\ \text{любое из } \mathbb{R}, & |x_k| \in (r_k, +\infty), \end{cases}$$

$$\mathsf{B}_k(x) = \mathbf{I}_{[-r_k, r_k]}(x).$$

Перейдем к шагу $k = 0$. Учитывая тот факт, что $\mathcal{F}_0 = \mathbb{R}^n = \mathbb{R}$ (см. начало раздела 3) получаем задачу стохастического программирования

$$\mathsf{B}_0(x) = \max_{u_0} \mathbf{P}\left(\frac{|x + u_0(1 + \xi_0)|}{r_1} \leqslant 1\right),$$

откуда с учетом [3] получаем:

$$u_0^* = \begin{cases} \text{любое из } U_0^*(X), & |X| \in \left[0, r_1\varepsilon^{-1}\right], \\ \dfrac{-\mathsf{sign}(X)\, r_1 - X}{1 + \varepsilon}, & |X| \in \left(r_1\varepsilon^{-1}, +\infty\right), \end{cases}$$

$$\mathsf{B}_0(x) = \min\left\{1, \quad \frac{r_1}{r_1 + |x|}\frac{1 + \varepsilon}{\varepsilon}\right\}.$$

Утверждение 1 доказано.

Доказательство утверждения 2. Примéним метод динамического программирования к решению задачи (8). На шаге $k = N$ получаем задачу стохастического программирования

$$\tilde{\mathsf{B}}_N(x, y) = \max_{u_N} \mathbf{I}_{(-\infty, 1]}(y)\, \mathbf{P}\left(\frac{|x + u_N(1 + \xi_N)|}{r_{N+1}} \leqslant 1\right),$$

решение которой (в области $y \leqslant 1$) известно [3] с точностью до параметров:

$$\tilde{u}_N^\varphi = \begin{cases} \text{любое из } U_N^*(x_N), & |x_N| \leqslant r_{N+1}\varepsilon^{-1},\ y_N \leqslant 1, \\ \dfrac{-\mathsf{sign}(x_N)\, r_{N+1} - x_N}{1 + \varepsilon}, & |x_N| > r_{N+1}\varepsilon^{-1},\ y_N \leqslant 1, \\ \text{любое из } \mathbb{R}, & y_N > 1, \end{cases}$$

$$\tilde{\mathsf{B}}_N(x, y) = \mathbf{I}_{(-\infty, 1]}(y) \min\left\{1, \quad \frac{r_{N+1}}{r_{N+1} + |x|}\frac{1 + \varepsilon}{\varepsilon}\right\}.$$

Перейдем теперь к шагу $k = N - 1$. Получаем, что

$$\tilde{\mathsf{B}}_{N-1}(x, y) = \max_{u_{N-1}} \mathbf{I}_{(-\infty, 1]}(y)\, \mathbf{M}\Bigg[\mathbf{I}_{(-\infty, 1]}\left(\frac{|x + u_{N-1}(1 + \xi_{N-1})|}{r_N}\right) \times$$

$$\times \min\left\{1, \quad \frac{r_{N+1}}{r_{N+1} + |x + u_{N-1}(1 + \xi_{N-1})|}\frac{1 + \varepsilon}{\varepsilon}\right\}\Bigg].$$

Последнее выражение может быть представлено в виде

$$\tilde{\mathsf{B}}_{N-1}(x,y)=\max_{u_{N-1}}\mathbf{I}_{(-\infty,1]}(y)\,\mathbf{M}\Bigg[\mathbf{I}_{(-\infty,\,r_N]}\left(|x+u_{N-1}(1+\xi_{N-1})|\right)\times$$

$$\times\Bigg(\mathbf{I}_{(-\infty,\,r_{N+1}\varepsilon^{-1}]}\left(|x+u_{N-1}(1+\xi_{N-1})|\right)+$$

$$+\mathbf{I}_{(r_{N+1}\varepsilon^{-1},\,+\infty)}\left(|x+u_{N-1}(1+\xi_{N-1})|\right)\frac{r_{N+1}}{r_{N+1}+|x+u_{N-1}(1+\xi_{N-1})|}\frac{1+\varepsilon}{\varepsilon}\Bigg)\Bigg],$$

откуда с учетом неравенства (10), из которого получаем, что $r_N\leqslant r_{N+1}\varepsilon^{-1}$,

$$\tilde{\mathsf{B}}_{N-1}(x,y)=\max_{u_{N-1}}\mathbf{I}_{(-\infty,1]}(y)\,\mathbf{M}\Big[\mathbf{I}_{(-\infty,\,r_N]}\left(|x+u_{N-1}(1+\xi_{N-1})|\right)\Big]=$$

$$=\max_{u_{N-1}}\mathbf{I}_{(-\infty,1]}(y)\,\mathbf{P}\left(\frac{|x+u_{N-1}(1+\xi_{N-1})|}{r_N}\leqslant 1\right).$$

Видно, что полученное выражение совпадает с выражением для функции Беллмана для $k=N$ с точностью до параметров r_N и r_{N+1}. Продолжая аналогичные рассуждения для $k=\overline{1,N}$, получаем выражения для оптимального управления и функции Беллмана на указанных шагах:

$$\tilde{u}_k^{\varphi}=\begin{cases}\text{любое из } U_k^*(x_k), & |x_k|\leqslant r_{k+1}\varepsilon^{-1},\ y_k\leqslant 1,\\ \dfrac{-\mathsf{sign}\,(x_k)\,r_{k+1}-x_k}{1+\varepsilon}, & |x_k|>r_{k+1}\varepsilon^{-1},\ y_k\leqslant 1,\\ \text{любое из } \mathbb{R}, & y_k>1,\end{cases}$$

$$\tilde{\mathsf{B}}_k(x,y)=\mathbf{I}_{(-\infty,1]}(y)\min\left\{1,\quad\frac{r_{k+1}}{r_{k+1}+|x|}\frac{1+\varepsilon}{\varepsilon}\right\}\quad k=\overline{1,N}.$$

Рассмотрим теперь шаг $k=0$. С учетом $y_0=x_1=|X+u_0(1+\xi_0)|$ справедливо

$$\tilde{\mathsf{B}}_0(x,y)=\max_{u_0}\mathbf{M}\Bigg[\mathbf{I}_{(-\infty,\,r_1]}\left(|x+u_0(1+\xi_0)|\right)\mathbf{I}_{(-\infty,\,r_1]}\left(|x+u_0(1+\xi_0)|\right)\times$$

$$\times\Bigg(\mathbf{I}_{(-\infty,\,r_2\varepsilon^{-1}]}\left(|x+u_0(1+\xi_0)|\right)+$$

$$+\mathbf{I}_{(r_2\varepsilon^{-1},\,+\infty)}\left(|x+u_0(1+\xi_0)|\right)\frac{r_2}{r_2+|x+u_0(1+\xi_0)|}\frac{1+\varepsilon}{\varepsilon}\Bigg)\Bigg]=$$

$$=\max_{u_0}\mathbf{M}\left[\mathbf{I}_{(-\infty,\,r_1]}\left(|x+u_0(1+\xi_0)|\right)\right]=\max_{u_0}\mathbf{P}\left(|x+u_0(1+\xi_0)\leqslant r_1|\right)=$$

$$=\min\left\{1,\quad\frac{r_1}{r_1+|x|}\frac{1+\varepsilon}{\varepsilon}\right\}.$$

Утверждение 2 доказано.

СПИСОК ЛИТЕРАТУРЫ

1. *Малышев В.В., Кибзун А.И.* Анализ и синтез высокоточного управления летательными аппаратами. М.: Машиностроение, 1987.
2. *Азанов В.М., Кан Ю.С.* Синтез оптимальных стратегий в задачах управления стохастическими дискретными системами по критерию вероятности // АиТ. 2017. № 6. С. 57–83.

   Azanov V.M., Kan Yu.S. Design of Optimal Strategies in the Problems of Discrete System Control by the Probabilistic Criterion // Autom. Remote Control. 2017. V. 78. No. 6. P. 1006–1027.
3. *Азанов В.М., Кан Ю.С.* Однопараметрчиеская задача оптимальной коррекции траектории летательного аппарата по критерию вероятности // Изв. РАН. ТиСУ. 2016. № 2. С. 1–13.
4. *Азанов В.М., Кан Ю.С.* Оптимизация коррекции околокруговой орбиты искусственного спутника Земли по вероятностному критерию // Тр. ИСА РАН. 2015. № 2. С. 18–26.
5. *Красильщиков М.Н., Малышев В.В., Федоров А.В.* Автономная реализация динамических операций на геостационарной орбите. I // Изв. РАН. ТиСУ. 2015. № 6. С. 82–95.
6. *Малышев В.В., Старков А.В., Федоров А.В.* Синтез оптимального управления при решении задачи удержания космического аппарата в орбитальной группировке // Космонавтика и ракетостроение. 2012. № 4. С. 150–158.
7. *Малышев В.В., Старков А.В., Федоров А.В.* Совмещение задач удержания и уклонения в окрестности опорной геостационарной орбиты // Вестн. Москов. город. пед. ун-та. Сер. Экономика. 2013. № 1. С. 68–74.
8. *Кибзун А.И., Игнатов А.Н.* О существовании оптимальных стратегий в задаче управления стохастической системой с дискретным временем по вероятностному критерию // АиТ. 2017. № 10. С. 139–154.

   Kibzun A.I., Ignatov A.N. On the Existence of Optimal Strategies in the Control Problem for a Stochastic Discrete Time System with Respect to the Probability Criterion // Autom. Remote Control. 2017. V. 78. No. 10. P. 1845–1856.
9. *Кибзун А.И., Игнатов А.Н.* Сведение двухшаговой задачи стохастического оптимального управления с билинейной моделью к задаче смешанного целочисленного линейного программирования. // АиТ. 2016. № 12. С. 89–111.

   Kibzun A.I., Ignatov A.N. Reduction of the Two-Step Problem of Stochastic Optimal Control with Bilinear Model to the Problem of Mixed Integer Linear Programming // Autom. Remote Control. 2016. V. 77. No. 12. P. 2175–2192.
10. *Кибзун А.И., Игнатов А.Н.* Двухшаговая задача формирования портфеля ценных бумаг из двух рисковых активов по вероятностному критерию. // АиТ. 2015. № 7. С. 78–100.

    Kibzun A.I., Ignatov A.N. The Two-Step Problem of Investment Portfolio Selection from two Risk Assets via the Probability Criterion // Autom. Remote Control. 2015. V. 76. No. 7. P. 1201–1220.
11. *Jasour A.M., Aybat N.S., Lagoa C.M.* Semidefinite Programming For Chance Constrained Optimization Over Semialgebraic Sets // SIAM J. Optim. 2015. № 25 (3). P. 1411–1440.
12. *Jasour A.M., Lagoa C.M.* Convex Chance Constrained Model Predictive Control // 2016. arXiv preprint arXiv:1603.07413.
13. *Jasour A.M., Lagoa C.M.* Convex Relaxations of a Probabilistically Robust Control Design Problem // 52nd IEEE Conf. on Decision and Control. 2013. P. 1892–1897.

14. *Босов А.В.* Обобщенная задача распределения ресурсов программной системы // Информ. и ее применение. 2014. Т. 8. Вып. 2. С. 39–47.

15. *Босов А.В.* Задачи анализа и оптимизации для модели пользовательской активности. Ч. 3. Оптимизация внешних ресурсов // Информ. и ее применение. 2012. Т. 6. Вып. 2. С. 14–21.

16. *Босов А.В.* Задачи анализа и оптимизации для модели пользовательской активности. Ч. 2. Оптимизация внутренних ресурсов // Информ. и ее применение. 2012. Т. 6. Вып. 1. С. 19–26.

17. *Ярошевский В.А., Петухов С.В.* Оптимальная однопараметрическая коррекция траекторий космических аппаратов // Космич. исследования. 1970. Вып. 8. № 4. С. 515–525.

18. *Ярошевский В.А., Парышева Г.В.* Оптимальное распределение корректирующих импульсов при однопараметрической коррекции // Космич. исследования. 1965. Т. III. Вып. 6; 1966. Т. IV. Вып. 1.

19. *Азанов В.М.* Оптимальное управление линейной дискретной системой по критерию вероятности // АиТ. 2014. № 10. С. 39–51.

    Azanov V.M., Kan Yu.S. Optimal Control for Linear Discrete Systems with Respect to Probabilistic Criteria // Autom. Remote Control. 2014. V. 75. No. 10. P. 1743–1753.

20. *Азанов В.М., Кан Ю.С.* Двухсторонняя оценка функции Беллмана в задачах стохастического оптимального управления дискретными системами по вероятностному критерию качества // АиТ. 2018. № 2. С. 3–18.

    Azanov V.M., Kan Yu.S. Bilateral Estimation of the Bellman Function in the Problems of Optimal Stochastic Control of Discrete Systems by the Probabilistic Performance Criterion // Autom. Remote Control. 2018. V. 79. No. 2. P. 203–215.

Статья представлена к публикации членом редколлегии Б.М. Миллером.

Поступила в редакцию 29.01.2018

После доработки 26.06.2018

Принята к публикации 08.11.2018