

© 2020 г. В.М. АЗАНОВ, канд. физ.-мат. наук (azanov59@gmail.com)
(Московский авиационный институт)

ОПТИМАЛЬНОЕ УПРАВЛЕНИЕ ДИСКРЕТНОЙ СТОХАСТИЧЕСКОЙ СИСТЕМОЙ С ВЕРОЯТНОСТНЫМ КРИТЕРИЕМ И НЕФИКСИРОВАННЫМ ВРЕМЕНЕМ ОКОНЧАНИЯ²

Рассматривается задача оптимального управления дискретной стохастической системой с критерием в форме вероятности первого достижения границ заданной области. Формулируются и доказываются достаточные условия оптимальности в форме метода динамического программирования. С помощью поверхностей уровней 1 и 0 функции Беллмана находятся двусторонние оценки функции правой части уравнения метода динамического программирования, функции Беллмана и функции оптимального значения вероятностного критерия, и предлагается способ построения субоптимального управления. Формулируются условия эквивалентности с задачей оптимального управления с вероятностным терминальным критерием. Рассматривается пример.

Ключевые слова: дискретные системы, стохастическое оптимальное управление, вероятностный критерий, метод динамического программирования, функция Беллмана.

DOI: 10.31857/S0005231020120016

1. Введение

Одним из важнейших направлений исследований в области стохастического оптимального управления являются задачи с нефиксированным моментом остановки. Среди них отдельно выделяют задачу стохастического быстрогодействия [1, 2], задачу с бесконечным горизонтом управления [3–6], задачу оптимизации времени пребывания системы в заданной трубке траекторий [1, 7] и задачу оптимизации момента первого достижения границ заданной области [1, 8, 9]. Модели управления с нефиксированным моментом остановки имеют широкое применение в авиационной [10], экономической [11], биологической, робототехнической и энергетической [1] областях. В случае систем с непрерывным временем известность получили методы, основанные на достаточных условиях оптимальности в форме метода динамического программирования, позволяющие искать оптимальные стратегии в классе позиционных. Интересно, что использование именно вероятностного критерия [1, 7]

¹ Статьи данной рубрики являются окончанием тематического выпуска, посвященного В.С. Пугачеву (№ 11, 2020).

² Работа выполнена при поддержке Российского фонда фундаментальных исследований (проект № 18-08-00595).

приводит к конструктивной форме постановки задач управления с нефиксированным временем, для которой удастся записать уравнение Беллмана. Тем не менее до сих пор существует ряд принципиальных проблем численного поиска оптимального управления, а аналитические решения получены лишь для ряда модельных задач [1]. Это обстоятельство связано со следующими трудностями: решение уравнения Беллмана может быть не единственным; даже если решение уравнения Беллмана существует в классе гладких функций, оно может быть недопустимым (например, потому что при нем может не существовать сильного решения уравнения стохастической системы в форме Ито); в классе допустимых управлений не всегда существует такое, при котором достигается точная грань критерия; уравнение Беллмана связано с так называемым “проклятием размерности”.

В случае дискретного времени качественная теория подобных задач изложена в [3]. Известны решения отдельных задач экономики [11] и модельных примеров [9]. В [9] была рассмотрена задача оптимизации вероятности первого достижения окрестности нуля траекториями линейной стохастической системы в канонической форме управляемости Бруновского. С помощью ее сведения к задаче с вероятностным терминальным критерием и дальнейшим использованием метода динамического программирования в форме [12] было найдено ее аналитическое решение.

В настоящей статье исследуется задача оптимального управления дискретной стохастической системой с критерием вероятности первого достижения ее траекториями заданной трубки. Исследуются достаточные условия оптимальности, схожие с [12], и свойства двусторонних границ функции Беллмана [13, 14]. Находятся условия эквивалентности с задачей оптимального управления с вероятностным терминальным критерием [12]. В качестве примера рассматривается задача управления портфелем ценных бумаг.

2. Постановка задачи

Рассмотрим стохастическую систему с дискретным временем

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

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 со значениями на $\mathbb{R}^s$ и известным распределением $\mathbf{P}_{\xi_k}$, $f_k: \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^s \rightarrow \mathbb{R}^n$ — функция перехода (функция системы), $N \in \{0\} \cup \mathbb{N}$ — горизонт управления.

В отношении системы (1) введем *п р е д п о л о ж е н и я*:

1. Известна полная информация о векторе состояния x_k (данный факт позволяет строить управление в классе функций $u_k = \gamma_k(x_k)$, где $\gamma_k(\cdot)$ — некоторая измеримая функция). В данном случае говорят, что “управление ищется в классе полной обратной связи по состоянию”;

2. Начальное состояние $x_0 = X$ является детерминированным вектором из $\mathbb{R}^n$;

3. Функция системы $f_k(x_k, u_k, \xi_k)$ непрерывна для всех k ;

4. Вектор управления u_k формируется следующим образом: $u_k = \gamma_k(x_k)$, где $\gamma_k : \mathbb{R}^n \rightarrow \mathbb{R}^m$ — 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 $u_k = \gamma_k(x_k)$ и в последнюю очередь реализуется случайное возмущение ξ_k ;

6. Управлением называется набор функций $u(\cdot) = (\gamma_0(\cdot), \dots, \gamma_N(\cdot)) \in \mathcal{U}$, классом допустимых управлений называется множество $\mathcal{U} = \mathcal{U}_0 \times \dots \times \mathcal{U}_N$, где $\mathcal{U}_k$ — множество борелевских функций $\gamma_k(\cdot)$ с ограниченными на U_k значениями;

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

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

множества $\mathcal{F}_k$ имеют вид

$$\begin{cases} \mathcal{F}_k = \{x \in \mathbb{R}^n : \Phi_k(x) \leq \varphi\}, & k = \overline{1, N+1}, \\ \mathcal{F}_0 = \mathbb{R}^n, \end{cases}$$

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

Рассматривается задача

$$(2) \quad P(u(\cdot)) \rightarrow \max_{u(\cdot) \in \mathcal{U}},$$

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

Физически задача (2) заключается в поиске управления, максимизирующего вероятность первого достижения трубки траекторий, заданной в виде последовательности множеств $\{\mathcal{F}_{k+1}\}_{k=0}^N$.

Отметим, что метод динамического программирования в форме [12], сформулированный для задач оптимального управления с вероятностным терминальным критерием, неприменим в общем случае к задаче (2). В разделе 3 устанавливаются достаточные условия оптимальности в форме метода динамического программирования, схожие с [12].

3. Метод динамического программирования и двусторонние оценки функции Беллмана

Определим функцию Беллмана $V_k : \mathbb{R}^n \rightarrow [0, 1]$ в задаче (2) как

$$V_k(x) = \sup_{\gamma_k(\cdot) \in \mathcal{U}_k, \dots, \gamma_N(\cdot) \in \mathcal{U}_N} \mathbf{P} \left(\min_{i=k, N} \Phi_{i+1}(x_{i+1}) \leq \varphi \mid x_k = x \right).$$

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

Теорема 1. 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 $X, \xi_0, \dots, \xi_N$ независимы;
- 5) множества $U_0, \dots, U_N$ компактны.

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

$$(3) \quad \gamma_k^*(x) = \arg \max_{u \in U_k} \mathbf{M}_{\xi_k} [\mathbf{I}_{\mathcal{F}_k}(x) + (1 - \mathbf{I}_{\mathcal{F}_k}(x)) \mathbf{B}_{k+1}(f_k(x, u, \xi_k))],$$

$$(4) \quad \mathbf{B}_k(x) = \max_{u \in U_k} \mathbf{M}_{\xi_k} [\mathbf{I}_{\mathcal{F}_k}(x) + (1 - \mathbf{I}_{\mathcal{F}_k}(x)) \mathbf{B}_{k+1}(f_k(x, u, \xi_k))], \quad k = \overline{0, N},$$

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

Доказательства теоремы 1, всех последующих теорем, утверждений и леммы приведены в Приложении.

В теореме 1 $\mathbf{M}_{\xi_k}[\cdot]$ — оператор математического ожидания по распределению $\mathbf{P}_{\xi_k}$ случайного вектора ξ_k , а $\mathbf{I}_{\mathcal{F}_k}(x)$ — индикаторная функция множества $\mathcal{F}_k$. Заметим, что соотношения (3)–(5) отличаются от классического уравнения Беллмана [12] в задаче с вероятностным терминальным критерием наличием в правой части дополнительного слагаемого $\mathbf{I}_{\mathcal{F}_k}(x)$ и множителя $1 - \mathbf{I}_{\mathcal{F}_k}(x)$ под оператором математического ожидания.

Как известно, прямое интегрирование уравнения Беллмана связано с трудностями вычислений кратных интегралов и решения задач стохастического программирования в его правой части. Указанные трудности неоднократно проявлялись даже для систем простейшего вида, например для системы управления портфелем ценных бумаг [15–17] и для системы управления стационарным спутником [12]. Вследствие этого на протяжении почти двадцати лет задачи оптимального управления с вероятностным критерием рассматривались в одношаговой $N = 0$ и двухшаговой $N = 1$ постановках [15–17]. В [13, 14] с использованием поверхностей уровня 1 и 0 были найдены двусторонние границы функции Беллмана, что позволило получить решение отдельных задач оптимального управления с вероятностным критерием для произвольного шага по времени N .

По аналогии с [13, 14] исследуем свойства функции Беллмана для задачи (2) с помощью поверхностей уровня 1 и 0 функции Беллмана.

4. Двусторонние границы функции Беллмана

Введем в рассмотрение поверхности уровней 1 и 0 функции Беллмана

$$\mathcal{I}_k = \{x \in \mathbb{R}^n : \mathbf{B}_k(x) = 1\}, \quad \mathcal{O}_k = \{x \in \mathbb{R}^n : \mathbf{B}_k(x) = 0\}$$

и множество $\mathcal{B}_k = \mathbb{R}^n \setminus \{\mathcal{I}_k \cup \mathcal{O}_k\}$. Для удобства введем обозначение $\overline{\mathcal{F}}_k = \mathbb{R}^n \setminus \mathcal{F}_k$. Нетрудно видеть, что из определения введенных множеств vyplнено, что

$$\mathcal{I}_k \cup \mathcal{B}_k \cup \mathcal{O}_k = \mathbb{R}^n, \quad \begin{cases} \mathbf{B}_k(x) = 1, & x \in \mathcal{I}_k, \\ \mathbf{B}_k(x) \in (0, 1), & x \in \mathcal{B}_k, \\ \mathbf{B}_k(x) = 0, & x \in \mathcal{O}_k. \end{cases}$$

Теорема 2. Справедливы утверждения:

1. Множества $\mathcal{I}_k$, $k = \overline{0, N}$, удовлетворяют рекуррентным соотношениям в обратном времени

$$\mathcal{I}_k = \mathcal{F}_k \cup \{x \in \mathbb{R}^n : \exists u \in U_k : \mathbf{P}_{\xi_k}(f_k(x, u, \xi_k) \in \mathcal{I}_{k+1}) = 1\}, \quad k = \overline{0, N},$$

$$\mathcal{I}_{N+1} = \mathcal{F}_{N+1};$$

2. Множества $\mathcal{O}_k$, $k = \overline{0, N}$, удовлетворяют рекуррентным соотношениям в обратном времени

$$\mathcal{O}_k = \overline{\mathcal{F}}_k \cap \{x \in \mathbb{R}^n : \forall u \in U_k : \mathbf{P}_{\xi_k}(f_k(x, u, \xi_k) \in \mathcal{O}_{k+1}) = 1\}, \quad k = \overline{0, N},$$

$$\mathcal{O}_{N+1} = \overline{\mathcal{F}}_{N+1};$$

3. Для $x \in \mathcal{I}_k$ функция $\gamma_k^*(x)$ принимает любое значение из множества $U_k^{\mathcal{I}}(x)$

$$(6) \quad U_k^{\mathcal{I}}(x) = \{u \in U_k : \mathbf{P}_{\xi_k}(f_k(x, u, \xi_k) \in \mathcal{I}_{k+1}) = 1\};$$

4. Для $x \in \mathcal{O}_k$ функция $\gamma_k^*(x)$ принимает любое значение из множества U_k ;

5. Уравнение Беллмана в области $x \in \mathcal{B}_k$ допускает представление

$$(7) \quad \mathbf{B}_k(x) = \max_{u \in U_k} \left\{ \mathbf{P}_{\xi_k}(f_k(x, u, \xi_k) \in \mathcal{I}_{k+1}) + \right.$$

$$\left. + \mathbf{P}_{\xi_k}(f_k(x, u, \xi_k) \in \mathcal{B}_{k+1}) \mathbf{M}_{\xi_k} \left[\mathbf{B}_{k+1}(f_k(x, u, \xi_k)) \mid f_k(x, u, \xi_k) \in \mathcal{B}_{k+1} \right] \right\};$$

6. Для $x \in \mathcal{B}_k$ и $u \in U_k$ справедлива система неравенств

$$(8) \quad \mathbf{P}_{\xi_k}(f_k(x, u, \xi_k) \in \mathcal{F}_{k+1}) \leq \mathbf{P}_{\xi_k}(f_k(x, u, \xi_k) \in \mathcal{I}_{k+1}) \leq$$

$$\leq \mathbf{M}_{\xi_k}[\mathbf{B}_{k+1}(f_k(x, u, \xi_k))] \leq 1 - \mathbf{P}_{\xi_k}(f_k(x, u, \xi_k) \in \mathcal{O}_{k+1});$$

7. Для $x \in \mathcal{B}_k$ функция Беллмана удовлетворяет двустороннему неравенству

$$(9) \quad \Psi_k(x) \leq \underline{\mathbf{B}}_k(x) \leq \mathbf{B}_k(x) \leq \overline{\mathbf{B}}_k(x),$$

где

$$(10) \quad \Psi_k(x) = \sup_{u \in U_k} \mathbf{P}_{\xi_k}(f_k(x, u, \xi_k) \in \mathcal{F}_{k+1}),$$

$\underline{B}_k(x)$ — нижняя, а $\overline{B}_k(x)$ — верхняя оценки функции Беллмана

$$\underline{B}_k(x) = \sup_{u \in U_k} \mathbf{P}_{\xi_k}(f_k(x, u, \xi_k) \in \mathcal{I}_{k+1}),$$

$$\overline{B}_k(x) = \sup_{u \in U_k} \{1 - \mathbf{P}_{\xi_k}(f_k(x, u, \xi_k) \in \mathcal{O}_{k+1})\},$$

причем $\underline{B}_N(x) = B_N(x) = \overline{B}_N(x)$.

Отличием правой части соотношений п. 1 теоремы 2 от соотношений для поверхности уровня 1 функции Беллмана в задаче с терминальным вероятностным критерием [18] является наличие операции объединения с множеством $\mathcal{F}_k$. Для поверхности уровня 0 функции Беллмана отличие заключается в наличии операции пересечения с множеством $\overline{\mathcal{F}}_k$. Пункты 3 и 4 устанавливают простейшие (относительно (3)) выражения для определения оптимального управления при $x_k \in \mathcal{I}_k \cup \mathcal{O}_k$, которые с точностью до конструкций множеств $\mathcal{I}_k$ совпадают с аналогичными в задаче с терминальным вероятностным критерием. Пункты 6 и 7 теоремы 2 устанавливают двусторонние оценки функции правой части уравнения динамического программирования и функции Беллмана соответственно. При этом выражения для нижних и верхних границ с точностью до конструкций множеств $\mathcal{I}_k$ и $\mathcal{O}_k$ совпадают с аналогичными в задаче с терминальным критерием [13, 18]. Отличием же является наличие дополнительного неравенства в левой части (8) и, как следствие, неравенства $\Psi_k(x) \leq \underline{B}_k(x)$.

Исследуем детально свойства стратегии, максимизирующей на каждом шаге нижнюю границу функции правой части уравнения динамического программирования.

5. Субоптимальная стратегия

Определим стратегию $\underline{u}(\cdot) = (\underline{\gamma}_0(\cdot), \dots, \underline{\gamma}_N(\cdot))$, где

$$(11) \quad \underline{\gamma}_k(x) = \arg \max_{u \in U_k} \mathbf{P}_{\xi_k}(f_k(x, u, \xi_k) \in \mathcal{I}_{k+1}), \quad k = \overline{0, N}.$$

Данная стратегия обладает следующими свойствами:

- При $x \in \mathcal{I}_k \cup \mathcal{O}_k$ для всех $k = \overline{0, N}$ выполнено равенство $\underline{\gamma}_k(x) = \gamma_k^*(x)$;
- Для $k = N$ выполнено равенство $\underline{\gamma}_k(x) = \gamma_k^*(x)$;
- Из $\mathcal{F}_k = \mathcal{I}_k$ для всех $k = \overline{0, N}$ следует $\underline{\gamma}_k(x) = \gamma_k^*(x)$.

Теорема 3. Пусть стратегия $\underline{u}(\cdot)$ существует в классе $\mathcal{U}$. Тогда справедливы утверждения:

1. Значение вероятностного критерия на стратегии $\underline{u}(\cdot)$ имеет вид

$$(12) \quad \begin{aligned} P(\underline{u}(\cdot)) &= \underline{F}(\varphi, N, X) + \\ &+ \sum_{l=1}^{N-1} \mathbf{P} \left(\bigcap_{k=1}^l \{\underline{x}_k \notin \mathcal{I}_k\} \right) \mathbf{P} \left(\bigcup_{k=0}^l \{\underline{x}_{k+1} \in \mathcal{I}_{k+1}\} \middle| \bigcap_{k=1}^l \{\underline{x}_k \notin \mathcal{I}_k\} \right) + \\ &+ \mathbf{P} \left(\bigcap_{k=1}^N \{\underline{x}_k \notin \mathcal{I}_k\} \right) \mathbf{P} \left(\bigcup_{k=0}^N \{\underline{x}_{k+1} \in \mathcal{F}_{k+1}\} \middle| \bigcap_{k=1}^N \{\underline{x}_k \notin \mathcal{I}_k\} \right), \end{aligned}$$

где $\underline{x}_k$ — траектория системы, замкнутой управлением (11)

$$\begin{cases} \underline{x}_{k+1} = f_k(\underline{x}_k, \underline{u}_k, \xi_k), & k = \overline{0, N}, \\ \underline{x}_0 = X, \end{cases}$$

где $\underline{u}_k = \gamma_k(\underline{x}_k)$;

2. Оптимальное значение вероятностного критерия имеет вид

$$(13) \quad \begin{aligned} P(u^*(\cdot)) &= \underline{F}(\varphi, N, X) + \\ &+ \sum_{l=1}^{N-1} \mathbf{P} \left(\bigcap_{k=1}^l \{x_k^* \notin \mathcal{I}_k\} \right) \mathbf{P} \left(\bigcup_{k=0}^l \{x_{k+1}^* \in \mathcal{I}_{k+1}\} \middle| \bigcap_{k=1}^l \{x_k^* \notin \mathcal{I}_k\} \right) + \\ &+ \mathbf{P} \left(\bigcap_{k=1}^N \{x_k^* \notin \mathcal{I}_k\} \right) \mathbf{P} \left(\bigcup_{k=0}^N \{x_{k+1}^* \in \mathcal{F}_{k+1}\} \middle| \bigcap_{k=1}^N \{x_k^* \notin \mathcal{I}_k\} \right), \end{aligned}$$

где x_k^* — траектория системы, замкнутой оптимальным управлением

$$\begin{cases} x_{k+1}^* = f_k(x_k^*, u_k^*, \xi_k), & k = \overline{0, N}, \\ x_0^* = X, \end{cases}$$

где $u_k^* = \gamma_k^*(x_k^*)$;

3. Для любых $\varphi \in \mathbb{R}$, $N \in \mathbb{N}$, $X \in \mathbb{R}^n$ справедлива система неравенств

$$(14) \quad F^{\mathcal{F}}(\varphi, N, X) \leq \underline{F}(\varphi, N, X) \leq P(\underline{u}(\cdot)) \leq P(u^*(\cdot)) \leq \overline{F}(\varphi, N, X),$$

где функции

$$F^{\mathcal{F}}: \mathbb{R} \times \mathbb{N} \times \mathbb{R}^n \rightarrow [0, 1], \quad \underline{F}: \mathbb{R} \times \mathbb{N} \times \mathbb{R}^n \rightarrow [0, 1], \quad \overline{F}: \mathbb{R} \times \mathbb{N} \times \mathbb{R}^n \rightarrow [0, 1]$$

имеют вид

$$F^{\mathcal{F}}(\varphi, N, X) = \Psi_0(X), \quad \underline{F}(\varphi, N, X) = \underline{\mathbb{B}}_0(X), \quad \overline{F}(\varphi, N, X) = \overline{\mathbb{B}}_0(X).$$

Из теоремы 3 можно получить оценку качества субоптимальной стратегии (11) $\underline{u}(\cdot)$

$$P(u^*(\cdot)) - P(\underline{u}(\cdot)) \leq \Delta(\varphi, N, X),$$

справедливаю для всех

$$\varphi \in \mathbb{R}, \quad N \in \{0\} \cup \mathbb{N}, \quad X \in \mathbb{R}^n,$$

где функция $\Delta: \mathbb{R} \times \mathbb{N} \times \mathbb{R}^n \rightarrow [0, 1]$ имеет вид

$$(15) \quad \begin{aligned} \Delta(\varphi, N, X) &= \overline{F}(\varphi, N, X) - \underline{F}(\varphi, N, X) - \\ &- \sum_{l=1}^{N-1} \mathbf{P} \left(\bigcap_{k=1}^l \{\underline{x}_k \notin \mathcal{I}_k\} \right) \mathbf{P} \left(\bigcup_{k=0}^l \{\underline{x}_{k+1} \in \mathcal{I}_{k+1}\} \middle| \bigcap_{k=1}^l \{\underline{x}_k \notin \mathcal{I}_k\} \right) - \\ &- \mathbf{P} \left(\bigcap_{k=1}^N \{\underline{x}_k \notin \mathcal{I}_k\} \right) \mathbf{P} \left(\bigcup_{k=0}^N \{\underline{x}_{k+1} \in \mathcal{F}_{k+1}\} \middle| \bigcap_{k=1}^N \{\underline{x}_k \notin \mathcal{I}_k\} \right). \end{aligned}$$

Исследуем условия эквивалентности задачи (2) и задачи оптимального управления с вероятностным терминальным критерием.

6. Условия эквивалентности с задачей оптимального управления с вероятностным терминальным критерием

Рассмотрим вероятностный терминальный критерий на траекториях системы (1)

$$(16) \quad P_\varphi(u(\cdot)) = \mathbf{P}(x_{N+1} \in \mathcal{F}_{N+1}) = \mathbf{P}(\Phi_{N+1}(x_{N+1}) \leq \varphi)$$

и задачу оптимального управления

$$(17) \quad P_\varphi(u(\cdot)) \rightarrow \max_{u(\cdot) \in \mathcal{U}}.$$

Согласно [12] решение задачи (17) существует в классе $\mathcal{U}$ и определяется в результате решения уравнений динамического программирования

$$(18) \quad \gamma_k^\varphi(x) = \arg \max_{u \in U_k} \mathbf{M}_{\xi_k} [\mathbf{B}_{k+1}^\varphi(f_k(x, u, \xi_k))],$$

$$(19) \quad \mathbf{B}_k^\varphi(x) = \max_{u \in U_k} \mathbf{M}_{\xi_k} [\mathbf{B}_{k+1}^\varphi(f_k(x, u, \xi_k))], \quad k = \overline{0, N},$$

$$(20) \quad \mathbf{B}_{N+1}^\varphi(x) = \mathbf{I}_{\mathcal{F}_{N+1}}(x),$$

где $\mathbf{B}_k^\varphi(x)$ – функция Беллмана в задаче (17).

Сформулируем утверждение об эквивалентности задач (2) и (17).

Лемма. Пусть для всех $k = \overline{0, N}$ выполнено $\mathcal{F}_k \subseteq \Delta \mathcal{I}_k$, где

$$(21) \quad \Delta \mathcal{I}_k = \{x \in \mathbb{R}^n : \exists u \in U_k : \mathbf{P}(f_k(x, u, \xi_k) \in \mathcal{I}_{k+1}) = 1\}.$$

Тогда задача (3) эквивалентна задаче оптимального управления с вероятностным терминальным критерием (17) в смысле равенства оптимальных стратегий $u^*(\cdot) = u^\varphi(\cdot)$, оптимальных значений критериев $P(u^*(\cdot)) = P_\varphi(u^\varphi(\cdot))$ и равенства для всех $k = \overline{0, N+1}$ и $x \in \mathbb{R}^n$ функций Беллмана $\mathbf{B}_k(x) = \mathbf{B}_k^\varphi(x)$.

Отметим, что из леммы для узкого класса систем можно получить более конкретные условия эквивалентности задач (2) и (17), однако это не является целью данной статьи.

Применим полученные результаты для исследования дополнительных свойств задачи оптимального управления портфелем ценных бумаг с вероятностным критерием.

7. Управление портфелем ценных бумаг с нефиксированным временем окончания

Рассмотрим дискретную стохастическую систему управления вида [13, 19]

$$(22) \quad \begin{cases} x_{k+1} = x_k \left(1 + bu_k^1 + \sum_{j=2}^m w_k^j \xi_k^{j-1} \right), & k = \overline{0, N}, \\ x_0 = X, \end{cases}$$

где $n = 1$, m и $s = m - 1$ — размерности векторов состояния, управления и случайных возмущений соответственно, $X > 0$, $b > -1$, $\varphi < 0$ — детерминированные скаляры, $\xi_k = (\xi_k^1, \dots, \xi_k^{m-1})^T$ — случайные векторы с независимыми компонентами, причем ξ_{k+1} и ξ_k являются независимыми для всех $k = \overline{0, N-1}$. Пусть ограничения на управления заданы в виде

$$U_k = U = \left\{ u \in \mathbb{R}^m : \sum_{j=1}^m u^j = 1, u^j \geq 0, \forall j = \overline{1, m} \right\}, \quad k = \overline{0, N}.$$

Предположим, что носитель распределения случайных векторов ξ_k имеет вид $\text{supp} \rho_{\xi}(t) = \bigotimes_{j=1}^{m-1} [\underline{\varepsilon}_j, \overline{\varepsilon}_j]$, причем $\forall j = \overline{1, m-1}$ выполнено $-1 \leq \underline{\varepsilon}_j \leq b \leq \overline{\varepsilon}_j$.

Рассмотрим задачу оптимального управления с нефиксированным временем

$$(23) \quad \mathbf{P} \left(\min_{k=\overline{0, N}} \{-x_{k+1}\} \leq \varphi \right) \rightarrow \sup_{u(\cdot) \in \mathcal{U}}$$

и задачу с вероятностным терминальным критерием

$$(24) \quad \mathbf{P}(-x_{N+1} \leq \varphi) \rightarrow \sup_{u(\cdot) \in \mathcal{U}}.$$

В обозначениях, введенных в статье, имеем:

$$\mathcal{F}_k = \mathcal{F} = [-\varphi, +\infty), \quad \Phi_k(x) = -x, \quad f_k(x, u, \xi) = x \left(1 + bu^1 + \sum_{j=2}^m u^j \xi^{j-1} \right).$$

Если за X принять размер стартового капитала, за x_k — размер капитала на начало k -го года, за u_k^1 — долю x_k капитала, вкладываемого в безрисковый актив (например, в надежный банк), имеющий доходность b , u_k^j — доли капитала, вкладываемые в рискованные активы, характеризующиеся доходностями ξ_k^{j-1} , $j = \overline{2, m}$, то задача (23) заключается в максимизации вероятности достижения размера капитала уровня $(-\varphi)$ за время, ограниченное сверху величиной $N + 1$, а задача (24) — за время $N + 1$ путем инвестиций в некоторые активы.

Задача (24) рассматривалась в двухшаговой постановке $N = 1$ для случая одного рискованного актива $m = 2$ в [15, 16]. В [13] был найден целый класс асимптотически оптимальных (при $N \rightarrow \infty$) стратегий. Задача (23) рассматривается впервые.

Воспользуемся результатами настоящей статьи и проверим условия эквивалентности задач (23) и (24).

Утверждение 1. Для всех $k = \overline{0, N}$ выполнены равенства

$$\mathcal{I}_k = \Delta \mathcal{I}_k = [\varphi_k^I, +\infty), \quad \mathcal{B}_k = (\varphi_k^O, \varphi_k^I), \quad \mathcal{O}_k = (-\infty, \varphi_k^O],$$

где скаляры $\varphi_k^{\mathcal{I}}, \varphi_k^{\mathcal{O}}$ определяются выражениями

$$\begin{aligned}\varphi_k^{\mathcal{I}} &= -\varphi \left(1 + \max \left\{ b, \max_{j=1, m-1} \underline{\varepsilon}_j \right\} \right)^{k-N-1}, \\ \varphi_k^{\mathcal{O}} &= -\varphi \left(1 + \max \left\{ b, \max_{j=1, m-1} \bar{\varepsilon}_j \right\} \right)^{k-N-1}.\end{aligned}$$

Из утверждения 1 вытекает, что условие леммы выполняется $\mathcal{F} \subseteq \Delta \mathcal{I}_k$ и, следовательно, задачи (23) и (24) являются эквивалентными. Из утверждения 1 следует, что управление, оптимальное по вероятностному терминальному критерию, максимизирует вероятность первого достижения капиталом x_k уровня $(-\varphi)$ за время, ограниченное сверху величиной N .

Воспользуемся двусторонними границами функции оптимального значения вероятностного критерия для определения оценки такого момента времени $N^* \in \mathbb{N}$, что

$$\mathbf{P} \left(\max_{k \in \{0, \dots, N^*\}} x_k^* \geq \varphi \right) = 1,$$

где $\{x_k^*\}_{k=0}^N$ – траектории системы (22), замкнутой оптимальным управлением $u^*(\cdot)$. Интересно, что двусторонние границы функции оптимального значения вероятностного критерия (см. теорему 3) позволяют найти такую оценку без нахождения оптимального управления.

Утверждение 2. Пусть $\{\underline{x}_k\}_{k=0}^N$ – траектории системы (22), замкнутой управлением (11), где

$$(25) \quad \underline{\gamma}_k(x) = \arg \max_{u \in U} \mathbf{P} \left(x \left(1 + bu^1 + \sum_{j=2}^m u^j \xi_k^{j-1} \right) \geq \varphi_{k+1}^{\mathcal{I}} \right), \quad k = \overline{0, N}.$$

Тогда существует $\underline{N} \in \mathbb{N}$, такой что

$$\mathbf{P} \left(\max_{k \in \{0, \dots, \underline{N}\}} \underline{x}_{k+1} \geq \varphi \right) = 1,$$

причем для любых $X > 0$, $b > -1$ выполнено $N^ \geq \underline{N}$ и*

$$(26) \quad \underline{N} = \left\lceil \frac{\ln(-\varphi) - \ln(X)}{\ln \left(1 + \max \left\{ b, \max_{j=1, m-1} \underline{\varepsilon}_j \right\} \right)} - 1 \right\rceil.$$

Проведем серию численных экспериментов для проверки адекватности оценки (26). При этом будем рассматривать случай одного рискового актива, для которого в [13] была найдена стратегия (25)

$$\underline{\gamma}_k(x) = \begin{cases} (1, 0)^T, & x \geq -\varphi(1+b)^{k-N-1}, \\ (0, 1)^T, & x < -\varphi(1+b)^{k-N-1}. \end{cases}$$

Значения параметров системы заданы в табл. 1.

Таблица 1. Значения параметров системы

Номер эксперимента	N	φ	X	b	$\underline{\varepsilon}_1$	$\bar{\varepsilon}_1$
а	200	100	15	0,01	-1	0,02
б	200	100	20	0,01	-1	0,02
в	200	100	25	0,01	-1	0,02

Таблица 2. Значения параметров системы

Номер эксперимента	а	б	в
$\underline{N}$	189	160	138
N^*	189	160	139

Для моделирования уровня N^* будем использовать метод Монте-Карло из 50000 наблюдений. Результаты моделирования занесены в табл. 2. Из табл. 2 видно, что $\underline{N}$ является относительно точной оценкой числа N^* .

8. Заключение

В статье рассмотрена задача оптимального управления дискретной стохастической системой с критерием вероятности первого достижения траекториями системы заданной трубки траекторий. Получены достаточные условия оптимальности в форме метода динамического программирования. Найдены двусторонние оценки функции правой части уравнения метода динамического программирования, функции Беллмана и функции оптимального значения вероятностного критерия. Получены аналитические выражения для приближенного определения оптимального управления и найдены оценки точности такого управления. Доказаны условия эквивалентности данной задачи и задачи оптимального управления с вероятностным терминальным критерием. Данные условия были проверены на задаче оптимального управления портфелем ценных бумаг.

ПРИЛОЖЕНИЕ

Доказательство теоремы 1. Введем в рассмотрение функцию $\Phi_0 : \mathbb{R}^n \rightarrow \mathbb{R}$, такую что $\Phi_0(x) = \Phi_1(x)$. Расширим вектор состояния системы путем введения новой переменной $y_k = \min_{i=\overline{0,k}} \Phi_i(x_i)$. Расширенная система управления имеет вид

$$\begin{cases} x_{k+1} = f_k(x_k, \tilde{u}_k, \xi_k), \\ y_{k+1} = \min \{y_k, \Phi_k(x_k)\}, \\ x_0 = X, \\ y_0 = \Phi_0(X), \end{cases} \quad k = \overline{0, N},$$

где $\tilde{u}_k = \tilde{\gamma}_k(x_k, y_k)$. Введем в рассмотрение функцию правой части расширенной системы $\tilde{f} : \mathbb{R}^{n+1} \times \mathbb{R}^m \times \mathbb{R}^s \rightarrow \mathbb{R}^{n+1}$:

$$\tilde{f}_k(x, y, u, \xi) = (f_k(x, u, \xi), \min \{y, \Phi_k(x)\})^T.$$

Тогда эквивалентную задачу оптимального управления можно представить в виде

$$\mathbf{P}(\min \{y_{N+1}, \Phi_{N+1}(x_{N+1})\} \leq \varphi) \rightarrow \sup_{u(\cdot) \in \tilde{\mathcal{U}}}, \quad k = \overline{0, N},$$

где $\tilde{\mathcal{U}} = \tilde{\mathcal{U}}_0 \times \dots \times \tilde{\mathcal{U}}_N$,

$$\tilde{\mathcal{U}}_k = \{\tilde{\gamma} : \mathbb{R}^{n+1} \rightarrow \mathbb{R}^m \mid \tilde{\gamma} \text{ измерима по Борелю, } \forall x \in \mathbb{R}^{n+1} : \tilde{\gamma}(x) \in U_k\}.$$

Эквивалентность выше понимается в смысле равенства критериев

$$P(u(\cdot)) = \mathbf{P}\left(\min_{k=\overline{0, N}} \Phi_{k+1}(x) \leq \varphi\right) = \mathbf{P}(\min \{y_{N+1}, \Phi_{N+1}(x_{N+1})\} \leq \varphi).$$

Уравнение Беллмана для эквивалентной задачи имеет вид [12]:

$$(П.1) \quad \tilde{\gamma}_k^*(x, y) = \arg \max_{u \in U_k} \mathbf{M}_{\xi_k} \left[\tilde{\mathbf{B}}_{k+1} \left(\tilde{f}_k(x, y, u, \xi_k) \right) \right],$$

$$(П.2) \quad \tilde{\mathbf{B}}_k(x, y) = \max_{u \in U_k} \mathbf{M}_{\xi_k} \left[\tilde{\mathbf{B}}_{k+1} \left(\tilde{f}_k(x, y, u_k, \xi_k) \right) \right],$$

$$(П.3) \quad \tilde{\mathbf{B}}_{N+1}(x, y) = \mathbf{I}_{\{\min\{y, \Phi_{N+1}(x)\} \leq \varphi\}}(x, y), \quad k = \overline{0, N}.$$

Согласно [12] если выполнены условия:

- 1) функции $\tilde{f}_k$ непрерывны для всех $k = \overline{0, N}$;
- 2) функция $\min \{y, \Phi_{N+1}(x)\}$ непрерывна и ограничена снизу;
- 3) случайные векторы $X, \xi_0, \dots, \xi_N$ независимы;
- 4) множества $U_0, \dots, U_N$ компактны,

то оптимальное управление существует в классе измеримых функций и определяется в результате решения задач (П.1)–(П.3).

На шаге $k = N$ уравнение Беллмана можно записать в виде

$$\begin{aligned} (П.4) \quad \tilde{\mathbf{B}}_N(x, y) &= \max_{u \in U_N} \mathbf{M}_{\xi_N} \left[\mathbf{I}_{\{\min\{y, \Phi_N(x), \Phi_{N+1}(f(x, u, \xi_k))\} \leq \varphi\}}(x, y) \right] = \\ &= \max_{u \in U_N} \mathbf{M}_{\xi_N} \left[\mathbf{I}_{\{\min\{y, \Phi_N(x)\} \leq \varphi\}}(x, y) + (1 - \mathbf{I}_{\{\min\{y, \Phi_N(x)\} \leq \varphi\}}(x, y)) \times \right. \\ &\quad \left. \times \mathbf{I}_{\{\Phi_{N+1}(f(x, u, \xi_k)) \leq \varphi\}}(x) \right] = \\ &= \max_{u \in U_N} \mathbf{M}_{\xi_N} \left[\mathbf{I}_{\{\min\{y, \Phi_N(x)\} \leq \varphi\}}(x, y) + (1 - \mathbf{I}_{\{\min\{y, \Phi_N(x)\} \leq \varphi\}}(x, y)) \times \right. \\ &\quad \left. \times \mathbf{B}_{N+1}(f_N(x, u, \xi_N)) \right]. \end{aligned}$$

Отсюда легко показать, что для любого $k = \overline{0, N}$ уравнение (П.2) можно представить в виде

$$\begin{aligned} (П.5) \quad \tilde{\mathbf{B}}_k(x, y) &= \max_{u \in U_k} \mathbf{M}_{\xi_k} \left[\mathbf{I}_{\{\min\{y, \Phi_k(x)\} \leq \varphi\}}(x, y) + \right. \\ &\quad \left. + (1 - \mathbf{I}_{\{\min\{y, \Phi_k(x)\} \leq \varphi\}}(x, y)) \mathbf{B}_{k+1}(f_k(x, u, \xi_k)) \right], \end{aligned}$$

поэтому для функции Беллмана в эквивалентной задаче справедливо представление

$$(П.6) \quad \tilde{B}_k(x, y) = \begin{cases} 1, & y \leq \varphi, \\ \max_{u \in U_k} \mathbf{M}_{\xi_k} \left\{ \mathbf{I}_{\mathcal{F}_k}(x) + (1 - \mathbf{I}_{\mathcal{F}_k}(x)) \mathbf{B}_{k+1}(f_k(x, u, \xi_k)) \right\}, & y > \varphi. \end{cases}$$

Отсюда получаем представление и для оптимального управления

$$(П.7) \quad \tilde{\gamma}_k^*(x, y) = \begin{cases} \text{любой элемент из } U_k, & y \leq \varphi, \\ \arg \max_{u \in U_k} \mathbf{M}_{\xi_k} \left\{ \mathbf{I}_{\mathcal{F}_k}(x) + (1 - \mathbf{I}_{\mathcal{F}_k}(x)) \mathbf{B}_{k+1}(f_k(x, u, \xi_k)) \right\}, & y > \varphi. \end{cases}$$

Условия 1–5 теоремы 1 получаются из условий существования оптимального управления для задачи с терминальным критерием. Причем пункты 2 и 3 следуют из условий непрерывности и ограниченности снизу функции $\min \{y, \Phi_{N+1}(x)\}$.

Теорема 1 доказана.

Доказательство теоремы 2. 1. Рассмотрим уравнение Беллмана на некотором шаге k . Для поверхности уровня 1 справедливо соотношение

$$\begin{aligned} \mathcal{I}_k &= \left\{ x \in \mathbb{R}^n \mid \max_{u \in U_k} \mathbf{M}_{\xi_k} [\mathbf{I}_{\mathcal{F}_k}(x) + (1 - \mathbf{I}_{\mathcal{F}_k}(x)) \mathbf{B}_{k+1}(f_k(x, u, \xi_k))] = 1 \right\} = \\ &= \mathcal{F}_k \cup \left\{ x \in \mathbb{R}^n \mid \max_{u \in U_k} \mathbf{M}_{\xi_k} [\mathbf{B}_{k+1}(f_k(x, u, \xi_k))] = 1 \right\}. \end{aligned}$$

Используя формулу полного математического ожидания, получаем

$$\begin{aligned} \mathcal{I}_k &= \mathcal{F}_k \cup \left\{ x \in \mathbb{R}^n \mid \max_{u \in U_k} \left\{ \mathbf{P}_{\xi_k}(f_k(x, u, \xi_k) \in \mathcal{I}_{k+1}) \times \right. \right. \\ &\quad \times \mathbf{M}_{\xi_k} [\mathbf{B}_{k+1}(f_k(x, u, \xi_k)) \mid f_k(x, u, \xi_k) \in \mathcal{I}_{k+1}] + \\ &\quad \left. \left. + (1 - \mathbf{P}_{\xi_k}(f_k(x, u, \xi_k) \in \mathcal{I}_{k+1})) \times \right. \right. \\ &\quad \left. \left. \times \mathbf{M}_{\xi_k} [\mathbf{B}_{k+1}(f_k(x, u, \xi_k)) \mid f_k(x, u, \xi_k) \notin \mathcal{I}_{k+1}] \right\} = 1 \right\}. \end{aligned}$$

Из равенств

$$\begin{aligned} \mathbf{M}_{\xi_k} [\mathbf{B}_{k+1}(f_k(x, u, \xi_k)) \mid f_k(x, u, \xi_k) \in \mathcal{I}_{k+1}] &= 1, \\ \mathbf{M}_{\xi_k} [\mathbf{B}_{k+1}(f_k(x, u, \xi_k)) \mid f_k(x, u, \xi_k) \notin \mathcal{I}_{k+1}] &< 1, \end{aligned}$$

следует, что

$$\begin{aligned} \mathcal{I}_k &= \mathcal{F}_k \cup \left\{ x \in \mathbb{R}^n \mid \max_{u \in U_k} \left\{ \mathbf{P}_{\xi_k}(f_k(x, u, \xi_k) \in \mathcal{I}_{k+1}) + \right. \right. \\ &\quad \left. \left. + (1 - \mathbf{P}_{\xi_k}(f_k(x, u, \xi_k) \in \mathcal{I}_{k+1})) \times \right. \right. \\ &\quad \left. \left. \times \mathbf{M}_{\xi_k} [\mathbf{B}_{k+1}(f_k(x, u, \xi_k)) \mid f_k(x, u, \xi_k) \notin \mathcal{I}_{k+1}] \right\} = 1 \right\} = \\ &= \mathcal{F}_k \cup \left\{ x \in \mathbb{R}^n \mid \max_{u \in U_k} \mathbf{P}_{\xi_k}(f_k(x, u, \xi_k) \in \mathcal{I}_{k+1}) = 1 \right\}. \end{aligned}$$

Утверждение 1 доказано.

2. Аналогично п. 1 получаем соотношение для поверхности уровня 0 функции Беллмана:

$$\begin{aligned}\mathcal{O}_k &= \left\{ x \in \mathbb{R}^n \mid \max_{u \in U_k} \mathbf{M}_{\xi_k} [\mathbf{I}_{\mathcal{F}_k}(x) + (1 - \mathbf{I}_{\mathcal{F}_k}(x)) \mathbf{B}_{k+1}(f_k(x, u, \xi_k))] = 0 \right\} = \\ &= \overline{\mathcal{F}}_k \cap \left\{ x \in \mathbb{R}^n \mid \max_{u \in U_k} \mathbf{M}_{\xi_k} [\mathbf{B}_{k+1}(f_k(x, u, \xi_k))] = 1 \right\}.\end{aligned}$$

Выполним преобразование правой части последнего выражения с использованием формулы полного математического ожидания:

$$\begin{aligned}\mathcal{O}_k &= \overline{\mathcal{F}}_k \cap \left\{ x \in \mathbb{R}^n \mid \max_{u \in U_k} \left\{ \mathbf{P}_{\xi_k}(f_k(x, u, \xi_k) \in \mathcal{I}_{k+1}) \times \right. \right. \\ &\quad \times \mathbf{M}_{\xi_k} [\mathbf{B}_{k+1}(f_k(x, u, \xi_k)) \mid f_k(x, u, \xi_k) \in \mathcal{I}_{k+1}] + \\ &\quad + \mathbf{P}_{\xi_k}(f_k(x, u, \xi_k) \in \mathcal{B}_{k+1}) \times \\ &\quad \times \mathbf{M}_{\xi_k} [\mathbf{B}_{k+1}(f_k(x, u, \xi_k)) \mid f_k(x, u, \xi_k) \in \mathcal{B}_{k+1}] + \\ &\quad + \mathbf{P}_{\xi_k}(f_k(x, u, \xi_k) \in \mathcal{O}_{k+1}) \times \\ &\quad \left. \left. \times \mathbf{M}_{\xi_k} [\mathbf{B}_{k+1}(f_k(x, u, \xi_k)) \mid f_k(x, u, \xi_k) \in \mathcal{O}_{k+1}] \right\} = 0 \right\}.\end{aligned}$$

Из равенств

$$\begin{aligned}\mathbf{M}_{\xi_k} [\mathbf{B}_{k+1}(f_k(x, u, \xi_k)) \mid f_k(x, u, \xi_k) \in \mathcal{I}_{k+1}] &= 1, \\ \mathbf{M}_{\xi_k} [\mathbf{B}_{k+1}(f_k(x, u, \xi_k)) \mid f_k(x, u, \xi_k) \in \mathcal{B}_{k+1}] &\in (0, 1), \\ \mathbf{M}_{\xi_k} [\mathbf{B}_{k+1}(f_k(x, u, \xi_k)) \mid f_k(x, u, \xi_k) \in \mathcal{O}_{k+1}] &= 0\end{aligned}$$

получаем, что

$$\begin{aligned}\mathcal{O}_k &= \overline{\mathcal{F}}_k \cap \left\{ x \in \mathbb{R}^n \mid \max_{u \in U_k} \left\{ \mathbf{P}_{\xi_k}(f_k(x, u, \xi_k) \in \mathcal{I}_{k+1}) + \right. \right. \\ &\quad + \mathbf{P}_{\xi_k}(f_k(x, u, \xi_k) \in \mathcal{B}_{k+1}) \times \\ &\quad \left. \left. \times \mathbf{M}_{\xi_k} [\mathbf{B}_{k+1}(f_k(x, u, \xi_k)) \mid f_k(x, u, \xi_k) \in \mathcal{B}_{k+1}] \right\} = 0 \right\} = \\ &= \overline{\mathcal{F}}_k \cap \left\{ x \in \mathbb{R}^n \mid \forall u \in U_k : \mathbf{P}_{\xi_k}(f_k(x, u, \xi_k) \in \mathcal{I}_{k+1}) = 0, \right. \\ &\quad \left. \mathbf{P}_{\xi_k}(f_k(x, u, \xi_k) \in \mathcal{B}_{k+1}) = 0 \right\} = \\ &= \overline{\mathcal{F}}_k \cap \left\{ x \in \mathbb{R}^n \mid \forall u \in U_k : \mathbf{P}_{\xi_k}(f_k(x, u, \xi_k) \in \mathcal{O}_{k+1}) = 1 \right\}.\end{aligned}$$

Утверждение 2 доказано.

3. Утверждение 3 следует из утверждения 1 теоремы 2.

4. Утверждение 4 выполнено, поскольку при $x \in \mathcal{O}_k$ справедливо равенство $\mathbf{B}_k(x) = 0$.

5. Утверждение 5 следует из утверждения 1 теоремы 2.

6. При $x \in \mathcal{B}_k$ функция правой части уравнения метода динамического программирования представима в виде

$$\begin{aligned} \mathbf{M}_{\xi_k} [\mathbf{B}_{k+1} (f_k(x, u, \xi_k))] &= \mathbf{P}_{\xi_k} (f_k(x, u, \xi_k) \in \mathcal{I}_{k+1}) \times \\ &\times (1 - \mathbf{M}_{\xi_k} [\mathbf{B}_{k+1} (f_k(x, u, \xi_k)) \mid f_k(x, u, \xi_k) \in \mathcal{B}_{k+1}]) + \\ &+ (1 - \mathbf{P}_{\xi_k} (f_k(x, u, \xi_k) \in \mathcal{O}_{k+1})) \times \\ &\times \mathbf{M}_{\xi_k} [\mathbf{B}_{k+1} (f_k(x, u, \xi_k)) \mid f_k(x, u, \xi_k) \in \mathcal{B}_{k+1}], \end{aligned}$$

откуда с использованием двустороннего неравенства для выпуклой комбинации получаем

$$\begin{aligned} \min \left\{ \mathbf{P}_{\xi_k} (f_k(x, u, \xi_k) \in \mathcal{I}_{k+1}), (1 - \mathbf{P}_{\xi_k} (f_k(x, u, \xi_k) \in \mathcal{O}_{k+1})) \right\} &\leq \\ &\leq \mathbf{M}_{\xi_k} [\mathbf{B}_{k+1} (f_k(x, u, \xi_k))] \leq \\ &\leq \max \left\{ \mathbf{P}_{\xi_k} (f_k(x, u, \xi_k) \in \mathcal{I}_{k+1}), (1 - \mathbf{P}_{\xi_k} (f_k(x, u, \xi_k) \in \mathcal{O}_{k+1})) \right\}. \end{aligned}$$

С использованием соотношений $1 - \mathbf{P}_{\xi_k} (f_k(x, u, \xi_k) \in \mathcal{O}_{k+1}) = \mathbf{P}_{\xi_k} (f_k(x, u, \xi_k) \in \mathcal{I}_{k+1}) + \mathbf{P}_{\xi_k} (f_k(x, u, \xi_k) \in \mathcal{B}_{k+1})$ и $\mathcal{F}_k \subseteq \mathcal{I}_k$ завершаем доказательство утверждения 6 теоремы 2.

7. Утверждение 7 следует из предыдущего путем взятия супремума во всех частях неравенства (8).

Доказательство теоремы 3. 1. Введем систему гипотез, образующих полную группу несовместных событий:

$$\left\{ \bigcup_{k=1}^N \{\underline{x}_k \in \mathcal{I}_k\} \right\}, \quad \left\{ \bigcap_{k=1}^N \{\underline{x}_k \notin \mathcal{I}_k\} \right\}.$$

Тогда с учетом формулы полной вероятности справедлива цепочка равенств

$$\begin{aligned} (П.8) \quad P(\underline{u}(\cdot)) &= \mathbf{P} \left(\bigcup_{k=0}^N \{\underline{x}_{k+1} \in \mathcal{F}_{k+1}\} \right) = \\ &= \mathbf{P} \left(\bigcup_{k=1}^N \{\underline{x}_k \in \mathcal{I}_k\} \right) \mathbf{P} \left(\bigcup_{k=0}^N \{\underline{x}_{k+1} \in \mathcal{F}_{k+1}\} \mid \bigcup_{k=1}^N \{\underline{x}_k \in \mathcal{I}_k\} \right) + \\ &+ \mathbf{P} \left(\bigcap_{k=1}^N \{\underline{x}_k \notin \mathcal{I}_k\} \right) \mathbf{P} \left(\bigcup_{k=0}^N \{\underline{x}_{k+1} \in \mathcal{F}_{k+1}\} \mid \bigcap_{k=1}^N \{\underline{x}_k \notin \mathcal{I}_k\} \right). \end{aligned}$$

Исследуем второй множитель первого слагаемого правой части последнего выражения. Из цепочки равенств

$$\mathbf{P}(\underline{x}_{N+1} \in \mathcal{F}_{N+1} \mid \underline{x}_N \in \mathcal{I}_N) = \mathbf{P}(\underline{x}_{N+1} \in \mathcal{I}_{N+1} \mid \underline{x}_N \in \mathcal{I}_N) = 1$$

следует, что

$$\mathbf{P} \left(\bigcup_{k=0}^N \{\underline{x}_{k+1} \in \mathcal{F}_{k+1}\} \mid \bigcup_{k=1}^N \{\underline{x}_k \in \mathcal{I}_k\} \right) = 1.$$

С учетом последнего равенства выражение (П.8) принимает вид

$$(П.9) \quad \begin{aligned} P(\underline{u}(\cdot)) &= \mathbf{P} \left(\bigcup_{k=1}^N \{\underline{x}_k \in \mathcal{I}_k\} \right) + \\ &+ \mathbf{P} \left(\bigcap_{k=1}^N \{\underline{x}_k \notin \mathcal{I}_k\} \right) \mathbf{P} \left(\bigcup_{k=0}^N \{\underline{x}_{k+1} \in \mathcal{F}_{k+1}\} \middle| \bigcap_{k=1}^N \{\underline{x}_k \notin \mathcal{I}_k\} \right). \end{aligned}$$

Введем в рассмотрение систему гипотез, образующих полную группу несовместных событий:

$$\left\{ \bigcup_{k=1}^{N-1} \{\underline{x}_k \in \mathcal{I}_k\} \right\}, \quad \left\{ \bigcap_{k=1}^{N-1} \{\underline{x}_k \notin \mathcal{I}_k\} \right\}.$$

Воспользуемся формулой полной вероятности для первого слагаемого в правой части (П.9):

$$(П.10) \quad \begin{aligned} \mathbf{P} \left(\bigcup_{k=1}^N \{\underline{x}_k \in \mathcal{I}_k\} \right) &= \\ &= \mathbf{P} \left(\bigcup_{k=1}^{N-1} \{\underline{x}_k \in \mathcal{I}_k\} \right) \mathbf{P} \left(\bigcup_{k=1}^N \{\underline{x}_k \in \mathcal{I}_k\} \middle| \bigcup_{k=1}^{N-1} \{\underline{x}_k \in \mathcal{I}_k\} \right) + \\ &+ \mathbf{P} \left(\bigcap_{k=1}^{N-1} \{\underline{x}_k \notin \mathcal{I}_k\} \right) \mathbf{P} \left(\bigcup_{k=1}^N \{\underline{x}_k \in \mathcal{I}_k\} \middle| \bigcap_{k=1}^{N-1} \{\underline{x}_k \notin \mathcal{I}_k\} \right). \end{aligned}$$

Аналогично (П.9) преобразуем правую часть (П.10), откуда получим

$$(П.11) \quad \begin{aligned} \mathbf{P} \left(\bigcup_{k=1}^N \{\underline{x}_k \in \mathcal{I}_k\} \right) &= \mathbf{P} \left(\bigcup_{k=1}^{N-1} \{\underline{x}_k \in \mathcal{I}_k\} \right) + \\ &+ \mathbf{P} \left(\bigcap_{k=1}^{N-1} \{\underline{x}_k \notin \mathcal{I}_k\} \right) \mathbf{P} \left(\bigcup_{k=1}^N \{\underline{x}_k \in \mathcal{I}_k\} \middle| \bigcap_{k=1}^{N-1} \{\underline{x}_k \notin \mathcal{I}_k\} \right). \end{aligned}$$

Выполним подстановку (П.10) в (П.9):

$$(П.12) \quad \begin{aligned} P(\underline{u}(\cdot)) &= \mathbf{P} \left(\bigcup_{k=1}^{N-1} \{\underline{x}_k \in \mathcal{I}_k\} \right) + \\ &+ \mathbf{P} \left(\bigcap_{k=1}^{N-1} \{\underline{x}_k \notin \mathcal{I}_k\} \right) \mathbf{P} \left(\bigcup_{k=1}^N \{\underline{x}_k \in \mathcal{I}_k\} \middle| \bigcap_{k=1}^{N-1} \{\underline{x}_k \notin \mathcal{I}_k\} \right) + \\ &+ \mathbf{P} \left(\bigcap_{k=1}^N \{\underline{x}_k \notin \mathcal{I}_k\} \right) \mathbf{P} \left(\bigcup_{k=0}^N \{\underline{x}_{k+1} \in \mathcal{F}_{k+1}\} \middle| \bigcap_{k=1}^N \{\underline{x}_k \notin \mathcal{I}_k\} \right). \end{aligned}$$

Проводя аналогичные преобразования в отношении первого слагаемого в (П.12) и вводя системы гипотез

$$\left\{ \bigcup_{k=1}^l \{\underline{x}_k \in \mathcal{I}_k\} \right\}, \quad \left\{ \bigcap_{k=1}^l \{\underline{x}_k \notin \mathcal{I}_k\} \right\}, \quad l = \overline{1, \dots, N-2},$$

получаем выражение для значения вероятностного критерия на стратегии $\underline{u}(\cdot)$:

$$(П.13) \quad \begin{aligned} P(\underline{u}(\cdot)) &= \mathbf{P}(\underline{x}_1 \in \mathcal{I}_1) + \\ &+ \sum_{l=1}^{N-1} \mathbf{P}\left(\bigcap_{k=1}^l \{\underline{x}_k \notin \mathcal{I}_k\}\right) \mathbf{P}\left(\bigcup_{k=1}^{l+1} \{\underline{x}_k \in \mathcal{I}_k\} \middle| \bigcap_{k=1}^l \{\underline{x}_k \notin \mathcal{I}_k\}\right) + \\ &+ \mathbf{P}\left(\bigcap_{k=1}^N \{\underline{x}_k \notin \mathcal{I}_k\}\right) \mathbf{P}\left(\bigcup_{k=0}^N \{\underline{x}_{k+1} \in \mathcal{F}_{k+1}\} \middle| \bigcap_{k=1}^N \{\underline{x}_k \notin \mathcal{I}_k\}\right). \end{aligned}$$

Заметим теперь, что для первого слагаемого (П.13) справедлива цепочка равенств

$$\mathbf{P}(\underline{x}_1 \in \mathcal{I}_1) = \mathbf{P}(f_0(X, \underline{u}_0, \xi_0) \in \mathcal{I}_1) = \underline{\mathbf{B}}_0(X) = \underline{F}(\varphi, N, X),$$

откуда следует выражение (14).

Пункт 1 теоремы 3 доказан.

2. Для доказательства п. 2 теоремы 3 достаточно заметить, что при $x_k \in \mathcal{I}_k$ выполнено $u_k^* = \underline{u}_k$ для всех $k = \overline{0, N}$, откуда аналогичным способом можно получить выражение (14) для функции оптимального значения вероятностного критерия на траекториях системы $\{x_k^*\}_{k=1}^{N+1}$, замкнутой оптимальным управлением $u^*(\cdot)$.

Пункт 2 теоремы 3 доказан.

3. Пункт 3 теоремы 3 непосредственно следует из п. 7 теоремы 2 и п. 1 теоремы 3.

Теорема 3 доказана.

Доказательство леммы. Рассмотрим соотношения динамического программирования (3)–(5) и (18)–(20) для задач (2) и (17) соответственно. Поскольку в (19) $\forall x \in \mathbb{R}^n$ выполняется равенство

$$\mathbf{B}_k^\varphi(x) = \max_{u \in U_k} \mathbf{M} \left[\mathbf{I}_{\mathcal{I}_k}^\varphi(x) + \left(1 - \mathbf{I}_{\mathcal{I}_k}^\varphi(x) \right) \mathbf{B}_{k+1}^\varphi(f_k(x, u, \xi_k)) \right], \quad k = \overline{0, N},$$

где $\mathcal{I}_k^\varphi$ – поверхность уровня 1 функции Беллмана $\mathbf{B}_k^\varphi(x)$, то описанные условия эквивалентности верны в том случае, если поверхности уровня 1 функций Беллмана в задачах (2) и (17) равны $\mathcal{I}_k = \mathcal{I}_k^\varphi$. С учетом рекуррентного соотношения для $\mathcal{I}_k$ (см. п. 1 теоремы 2) отмеченное равенство справедливо только в том случае, если

$$(П.14) \quad \mathcal{F}_k \subseteq \{x \in \mathbb{R}^n : \exists u \in U_k : \mathbf{P}(f_k(x, u, \xi_k) \in \mathcal{I}_{k+1}) = 1\}, \quad k = \overline{0, N}.$$

Лемма доказана.

Доказательство утверждения 1. В соответствии с п. 1 теоремы 2 запишем уравнение для поверхностей уровней 1 и 0 функции Беллмана на шаге $k = N$:

$$\begin{aligned}\mathcal{I}_N &= \mathcal{F} \cup \left\{ x \in \mathbb{R} : \exists u \in U : \mathbf{P} \left(x \left(1 + u^1 b + \sum_{i=2}^m u^i \xi_N^i \right) \geq -\varphi \right) = 1 \right\}, \\ \mathcal{O}_N &= \overline{\mathcal{F}} \cap \left\{ x \in \mathbb{R} : \forall u \in U : \mathbf{P} \left(x \left(1 + u^1 b + \sum_{i=2}^m u^i \xi_N^i \right) < -\varphi \right) = 1 \right\}.\end{aligned}$$

Используя результат из [13], где были найдены решения соответствующих уравнений, и введенные в разделе 6 обозначения для границ множеств $\mathcal{I}_k$ и $\mathcal{O}_k$, получаем:

$$\begin{aligned}\mathcal{I}_N &= [\varphi, +\infty) \cup \left[\varphi \left(1 + \max \left\{ b, \max_{j=2, \dots, m} \bar{b}^j \right\} \right)^{-1}, +\infty \right) = [\varphi_N^{\mathcal{I}}, +\infty), \\ \mathcal{O}_N &= (-\infty, \varphi] \cap \left(-\infty, \varphi \left(1 + \max \left\{ b, \max_{j=2, \dots, m} \bar{b}^j \right\} \right)^{-1} \right] = (-\infty, \varphi_N^{\mathcal{O}}].\end{aligned}$$

Отсюда следует, что $\mathcal{I}_N = \Delta \mathcal{I}_N$. Используя п. 1 теоремы 2, получаем, что на шаге $k = N - 1$ уравнение для изобелл примет вид

$$\begin{aligned}\mathcal{I}_{N-1} &= \mathcal{F} \cup \left\{ x \in \mathbb{R} : \exists u \in U : \mathbf{P} \left(x \left(1 + u^1 b + \sum_{i=2}^m u^i \xi_{N-1}^i \right) \geq \varphi_N^{\mathcal{I}} \right) = 1 \right\}, \\ \mathcal{O}_{N-1} &= \overline{\mathcal{F}} \cap \left\{ x \in \mathbb{R} : \forall u \in U : \mathbf{P} \left(x \left(1 + u^1 b + \sum_{i=2}^m u^i \xi_{N-1}^i \right) < \varphi_N^{\mathcal{O}} \right) = 1 \right\}.\end{aligned}$$

По аналогии с шагом $k = N$ получаем

$$\mathcal{I}_{N-1} = \Delta \mathcal{I}_{N-1} = [\varphi_{N-1}^{\mathcal{I}}, +\infty), \quad \mathcal{O}_{N-1} = (-\infty, \varphi_{N-1}^{\mathcal{O}}].$$

Оперируя математической индукцией, заключаем, что для всех $k = \overline{0, N}$ выполнено

$$\mathcal{I}_k = \Delta \mathcal{I}_k = [\varphi_k^{\mathcal{I}}, +\infty), \quad \mathcal{O}_k = (-\infty, \varphi_k^{\mathcal{O}}].$$

Утверждение 1 доказано.

Доказательство утверждения 2. Рассмотрим стратегию управления (11), которая с учетом утверждения 1 принимает вид (25). Из п. 3 теоремы 3 следует, что

$$(П.15) \quad P(\underline{u}(\cdot)) = \mathbf{P} \left(\max_{k \in \{0, \dots, N\}} \underline{x}_{k+1} \geq -\varphi \right) \geq \underline{F}(\varphi, N, X),$$

где функция $\underline{F}$ с учетом

$$\varphi_1^{\mathcal{I}} = -\varphi \left(1 + \max \left\{ b, \max_{j=1, \dots, m-1} \bar{\varepsilon}_j \right\} \right)^{-N}$$

имеет вид

$$(П.16) \quad \underline{F}(\varphi, N, X) = \\ = \max_{u \in U} \mathbf{P} \left(X \left(1 + bu^1 + \sum_{j=2}^m u^j \xi_0^{j-1} \right) \geq - \frac{\varphi}{\left(1 + \max \left\{ b, \max_{j=1, m-1} \varepsilon_j \right\} \right)^N} \right).$$

Нетрудно видеть, что величину $\underline{N} \in \mathbb{N}$ можно определить как корень уравнения $\underline{F}(\varphi, N, X) = 1$, но поскольку таких корней бесконечное множество, то будем искать оценку $\underline{N}$ в виде

$$(П.17) \quad \underline{N} = \min \{ N \in \mathbb{N} : \underline{F}(\varphi, N, X) = 1 \}.$$

Из (П.15) и (П.17) следует, что

$$\mathbf{P} \left(\max_{k \in \{0, \dots, \underline{N}\}} \underline{x}_{k+1} \geq -\varphi \right) = 1.$$

Из определений поверхности уровня 1 функции Беллмана и функции $\underline{F}$ следует, что уравнение $\underline{F}(\varphi, N, X) = 1$ эквивалентно включению $X \in \mathcal{I}_0$, что в свою очередь эквивалентно неравенству

$$X \geq - \frac{\varphi}{\left(1 + \max \left\{ b, \max_{j=1, m-1} \varepsilon_j \right\} \right)^{N+1}}.$$

Путем логарифмирования получаем

$$N \geq \frac{\ln(-\varphi) - \ln(X)}{\ln \left(1 + \max \left\{ b, \max_{j=1, m-1} \varepsilon_j \right\} \right)} - 1.$$

Используя (П.17) окончательно получаем (26).

Утверждение 2 доказано.

СПИСОК ЛИТЕРАТУРЫ

1. Афанасьев В.Н., Колмановский В.Б., Носов В.Р. Математическая теория конструирования систем управления. М.: Высш. шк., 2003.
2. Смирнов И.П. Об одной задаче быстрогодействия для стохастической управляемой системы // Дифференц. уравнения. 1986. Т. 22. № 2. С. 247–254.
3. Бертсекас Д., Шрив С. Стохастическое управление: случай дискретного времени. М.: Наука, 1985.
4. Zhou J. Infinite Horizon Optimal Control Problem for Stochastic Evolution Equations in Hilbert Spaces // J. Dyn. Control Syst. 2016. V. 22. No. 3. P. 531–554.
5. Agram N., Haadem S., Øksendal B., Proske F. A Maximum Principle for Infinite Horizon Delay Equations // SIAM J. Math. Anal. 2013. V. 45. No. 4. P. 2499–2522.

6. *Agram N., Øksendal B.* Infinite Horizon Optimal Control of Forward-Backward Stochastic Differential Equations with Delay // *J. Comput. Appl. Math.* 2014. V. 259. Part B. P. 336–349.
7. *Черноусько Ф.Л., Колмановский В.Б.* Оптимальное управление при случайных возмущениях. М.: Наука, 1978.
8. *Смирнов И.П.* Об управлении вероятностью входа системы в заданную область // *Дифференц. уравнения.* 1990. Т. 26. № 10. С. 1753–1758.
9. *Азанов В.М.* Оптимальное управление линейной дискретной системой по критерию вероятности // *АиТ.* 2014. № 10. С. 39–51.
Azanov V.M. Optimal Control for Linear Discrete Systems with Respect to Probabilistic Criteria // *Autom. Remote Control.* 2014. V. 75. No. 10. P. 1743–1753.
10. *Семаков С.Л.* Выбросы случайных процессов: приложения в авиации. М.: Наука, 2005.
11. *Хаметов В.М., Шелемех Е.А.* Суперхеджирование американских опционов на неполном рынке с дискретным временем и конечным горизонтом // *АиТ.* 2015. № 9. С. 125–149.
Khametov V.M., Shelemekh E.A. Superhedging of American Options on an Incomplete Market with Discrete Time and Finite Horizon // *Autom. Remote Control.* 2015. V. 76. No. 9. P. 1616–1634.
12. *Малышев В.В., Кибзун А.И.* Анализ и синтез высокоточного управления летательными аппаратами. М.: Машиностроение, 1987.
13. *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.
14. *Азанов В.М., Кан Ю.С.* Усиленная оценка функции Беллмана в задачах стохастического оптимального управления с вероятностным критерием качества // *АиТ.* 2019. № 4. С. 53–69.
Azanov V.M., Kan Yu.S. Refined Estimation of the Bellman Function for Stochastic Optimal Control Problems with Probabilistic Performance Criterion // *Autom. Remote Control.* 2019. V. 80. No. 4. P. 634–647.
15. *Григорьев П.В., Кан Ю.С.* Оптимальное управление по квантильному критерию портфелем ценных бумаг // *АиТ.* 2004. № 2. С. 179–197.
Grigor'ev P.V., Kan Yu.S. Optimal Control of the Investment Portfolio with Respect to the Quantile Criterion // *Autom. Remote Control.* 2004. V. 65. No. 2. P. 319–336.
16. *Бунто Т.В., Кан Ю.С.* Оптимальное управление по квантильному критерию портфелем ценных бумаг с ненулевой вероятностью разорения // *АиТ.* 2013. № 5. С. 114–136.
Bunto T.V., Kan Yu.S. Quantile Criterion-based Control of the Securities Portfolio with a Nonzero Ruin Probability // *Autom. Remote Control.* 2013. V. 74. No. 5. P. 811–828.
17. *Кибзун А.И., Игнатов А.Н.* Двухшаговая задача формирования портфеля ценных бумаг из двух рискованных 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.

18. *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.
19. *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.

Статья представлена к публикации членом редколлегии А.И. Кибзуном.

Поступила в редакцию 02.03.2020

После доработки 20.05.2020

Принята к публикации 09.07.2020