Автоматика и телемеханика, № 4, 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)

УСИЛЕННАЯ ОЦЕНКА ФУНКЦИИ БЕЛЛМАНА В ЗАДАЧАХ СТОХАСТИЧЕСКОГО ОПТИМАЛЬНОГО УПРАВЛЕНИЯ С ВЕРОЯТНОСТНЫМ КРИТЕРИЕМ КАЧЕСТВА[1]

Рассматривается задача оптимального управления дискретной стохастической системой с вероятностным терминальным критерием общего вида. С использованием метода динамического программирования и свойств функции Беллмана для этой функции строятся новые двусторонние границы, уточняющие известные ранее. С использованием указанных оценок приводится обоснование применения модифицированной стратегии, оптимальной в двухшаговой задаче управления портфелем ценных бумаг с учетом риска, для решения аналогичной многошаговой задачи. На примере демонстрируются преимущества такой стратегии перед другими известными стратегиями.

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

DOI: 10.1134/S0005231019040032

1. Введение

Задачи оптимального управления стохастическими марковскими системами с дискретным временем и терминальным вероятностным критерием исследуются с середины XX в. Интерес к исследованию данного класса задач мотивирован аэрокосмическими [1–5], экономическими [6–13] и робототехническими [14–16] приложениями. В аэрокосмических приложениях вероятностный критерий исследовался, например, в [1, 5] в рамках задачи оптимальной импульсной коррекции траектории движения искусственного спутника Земли в окрестности геостационарной орбиты. В [1] были сформулированы достаточные условия оптимальности для задач оптимального управления дискретными стохастическими системами с вероятностным критерием в форме рекуррентных соотношений динамического управления. Публикации [6–13] посвящены задачам оптимального управления портфелем ценных бумаг с учетом риска. В модели [6–13] система управления характеризует изменение капитала во времени, за управление принимаются доли капитала, вкладываемые в безрисковый актив, имеющий детерминированную доходность, и

[1] Результаты работы получены в рамках выполнения государственного задания Минобрнауки № 2.2461.2017/ 4.6.

рисковые активы, имеющие случайную доходность с известным распределением. Вероятностный критерий моделирует вероятность достижения капиталом к заданному моменту времени некоторого уровня. В ходе исследования указанных задач был разработан численный алгоритм решения [7]. Интерес к моделям оптимального управления портфелем ценных бумаг в схожих постановках проявляется и в настоящее время, см., например, [17, 18].

Метод динамического программирования (МДП) является основным [1] инструментом решения задач синтеза оптимального управления дискретными стохастическими системами с вероятностным критерием качества. Среди недавних публикаций, посвященных данной проблематике, можно выделить [5, 6, 14–16]. Для ряда случаев АДП позволяет найти оптимальную стратегию в классе измеримых ограниченных функций [7]. Однако в связи с "проклятьем размерности" по сей день нет эффективных численных методов решения уравнения динамического программирования (уравнения Беллмана). Тем не менее в ряде случаев оптимальное управление может быть найдено аналитически с использованием свойств и двусторонних оценок функции Беллмана [5, 6]. Например, в [5] было найдено аналитическое решение задачи оптимального управления билинейной системой со скалярным ограниченным управлением и вероятностным терминальным критерием. Такая модель используется в задачах импульсной коррекции орбиты геостационарного спутника (см., например, [1]). Техника нахождения двусторонних оценок функции Беллмана для синтеза оптимальных и субоптимальных управлений использовалась, например, в [19] для других (не вероятностных) критериев.

Отдельно отметим публикации [14–16], где выделен широкий класс задач оптимального управления с вероятностным критерием, для которого разработаны численные алгоритмы поиска оптимального управления. Указанный класс характеризуется тем, что функция системы (функция перехода), обратная связь по состоянию и точностной функционал являются полиномами, система управления стационарна, а носитель распределения случайных возмущений ограничен. Постулирование полиномиальной обратной связи по состоянию позволило свести бесконечномерную оптимизационную задачу к конечномерной задаче стохастического программирования большой размерности.

Настоящая статья является продолжением [6], где для задач оптимального управления дискретными стохастическими системами с вероятностным терминальным критерием на основе метода динамического программирования и свойств функции Беллмана, исследованных в [5], была найдена двусторонняя оценка для функции Беллмана. В [5] показано, что нижняя и верхняя оценки функции Беллмана как функции состояния определяются в результате решения соответствующих задач стохастического программирования с целевыми функциями в форме функций вероятности. Для задачи оптимального управления портфелем ценных бумаг двусторонняя оценка была найдена в явном виде [6], где, в частности, установлено, что управление, найденное из решений задач стохастического программирования для нижней оценки, является асимптотически оптимальным при стремлении горизонта управления к бесконечности. В настоящей статье находятся усиленные оценки для функции Беллмана, которые уточняют полученные раннее двусторонние оценки. На примере задач оптимального управления портфелем ценных бумаг показано, что полученная оценка является лучше, а стратегия, полученная в результате решения соответствующих задач стохастического программирования, обеспечивает лучшее значение вероятностного критерия по сравнению с известными стратегиями управления в указанной задаче.

2. Постановка задачи

Рассмотрим стохастическую систему управления с дискретным временем

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

$x_k \in \mathbb{R}^n$ – вектор состояния. Начальное состояние X предполагается случайным с известным распределением, $u_k \in \mathbb{R}^m$ – вектор управления, $\xi_k \in \mathbb{R}^s$ – случайный вектор с известным распределением, $f_k : \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^s \to \mathbb{R}^n$ – непрерывная для всех $k = \overline{0, N}$ функция, $N \in \mathbb{N}$ – горизонт управления, причем $N \geqslant 2$.

Пусть $\xi^k = (\xi_0, \ldots, \xi_k)^{\mathtt{T}}$. Предположим, что ξ^N не зависит от X, а ξ_{k+1} не зависит от ξ^k. Обозначим через $\mathcal{U}_k$ множество борелевских функций $u_k = u_k\left(x\right)$, значения которых удовлетворяют ограничениям $u_k\left(x\right) \in U_k$, где $U_k \subset \mathbb{R}^m$ – компактные множества.

На траекториях системы (1) задан функционал вероятности

$$
P_\varphi\left(u\left(\cdot\right)\right) = \mathbf{P}\left(\Phi\left(x_{N+1}\left(u\left(\cdot\right), \zeta\right)\right) \leqslant \varphi\right),
$$

где $\mathbf{P}$ – вероятность, $\Phi : \mathbb{R}^n \to \mathbb{R}$ – ограниченная снизу непрерывная функция, $\varphi \in \mathbb{R}$ – известный скаляр, $u\left(\cdot\right) = \left(u_0^{\mathtt{T}}\left(\cdot\right), \ldots, u_N^{\mathtt{T}}\left(\cdot\right)\right)^{\mathtt{T}}$ – управление, $\zeta = \left(X^{\mathtt{T}}, \xi_0^{\mathtt{T}}, \ldots, \xi_N^{\mathtt{T}}\right)^{\mathtt{T}}$.

Задача стохастического оптимального управления с вероятностным терминальным критерием имеет вид

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

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

Введем в рассмотрение функцию Беллмана $\mathsf{B}_k : \mathbb{R}^n \to [0, 1]$, которая по определению [1] является точной верхней гранью функционала вероятности при фиксированном текущем состоянии $x_k = x$, т.е.

$$
\mathsf{B}_k\left(x\right) = \\
= \sup_{u_k(\cdot) \in \mathcal{U}_k, \ldots, u_N(\cdot) \in \mathcal{U}_N} \mathbf{P}\Big(\Phi\left(x_{N+1}\left(x_k, u_k\left(\cdot\right), \ldots, u_N\left(\cdot\right), \xi_k, \ldots, \xi_N\right)\right) \leqslant \varphi \Big| x_k = x\Big).
$$

В [1] установлено, что если существует стратегия $u^\varphi\left(\cdot\right) \in \mathcal{U}$, удовлетворяющая следующим рекуррентным соотношениям динамического программирования:

$$u_k^{\varphi}(x) = \arg\max_{u_k \in U_k} \mathbf{M}\left[\mathsf{B}_{k+1}\left(f_k\left(x_k, u_k, \xi_k\right)\right)\middle| x_k = x\right],$$
$$\mathsf{B}_k(x) = \max_{u_k \in U_k} \mathbf{M}\left[\mathsf{B}_{k+1}\left(f_k\left(x_k, u_k, \xi_k\right)\right)\middle| x_k = x\right], \quad k = \overline{0, N},$$
$$\mathsf{B}_{N+1}(x) = \mathbf{I}_{(-\infty,\varphi]}\left(\Phi(x)\right),$$

то она оптимальна в задаче (2). Здесь $\mathbf{M}$ – оператор математического ожидания по распределению $\mathbf{P}$, $\mathbf{I}_{\mathcal{A}}(x)$ – индикаторная функция множества $\mathcal{A}$.

В [5] с использованием поверхностей уровней 1 и 0 функции Беллмана, называемых изобеллами, удалось получить модификацию соотношений динамического программирования, на основе которой было получено аналитическое решение некоторых аэрокосмических задач. В разделе 3 приводятся вспомогательные утверждения из [5], составляющие методологическую основу для приводимых далее результатов по уточнению оценок для функции Беллмана.

3. Определение изобелл и модификация уравнения Беллмана

Основной идеей, заложенной в [5], является рассмотрение уравнения Беллмана в разных областях пространства состояний, а именно в таких, в которых функция Беллмана равна единице, нулю и не равна ни единице, ни нулю. Первые две области, а точнее множества, названы в [5] изобеллами уровней 1 и 0. Они определяются следующими соотношениями [5]:

$$\mathcal{I}_k = \left\{x \in \mathbb{R}^n: \quad \mathsf{B}_k(x) = 1\right\}, \quad \mathcal{O}_k = \left\{x \in \mathbb{R}^n: \quad \mathsf{B}_k\Big(x\Big) = 0\right\}, \quad k = \overline{0, N}.$$

Рассмотрим также множество, где функция Беллмана не равна ни 1, ни 0,

$$\mathcal{B}_k = \mathbb{R}^n \setminus \left\{\mathcal{I}_k \cup \mathcal{O}_k\right\}.$$

Отметим, что из определения множеств $\mathcal{I}_k$, $\mathcal{B}_k$, $\mathcal{O}_k$ следует, что

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

В статьях [5, 6] исследованы важные свойства изобелл уровней 1 и 0, с помощью которых установлены новые свойства функции Беллмана. Указанные свойства объединены в одну лемму 1.

Лемма 1. Справедливы следующие утверждения:

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

$$\mathcal{I}_k = \left\{x \in \mathbb{R}^n: \quad \exists u_k \in U_k: \quad \mathbf{P}\left(f_k\left(x, u_k, \xi_k\right) \in \mathcal{I}_{k+1}\right) = 1\right\},$$
$$\mathcal{I}_{N+1} = \left\{x \in \mathbb{R}^n: \quad \Phi(x) \leqslant \varphi\right\};$$

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

$$\mathcal{O}_k = \Big\{x \in \mathbb{R}^n : \quad \forall u_k \in U_k : \quad \mathbf{P}\left(f_k\left(x, u_k, \xi_k\right) \in \mathcal{O}_{k+1}\right) = 1\Big\},$$
$$\mathcal{O}_{N+1} = \Big\{x \in \mathbb{R}^n : \quad \Phi\left(x\right) > \varphi\Big\};$$

3. *Для* $x_k \in \mathcal{I}_k$ *оптимальным управлением на шаге* k *является любой элемент из множества* $U_k^{\mathcal{I}}\left(x_k\right)$:

$$U_k^{\mathcal{I}}\left(x_k\right) = \Big\{u \in U_k : \quad \mathbf{P}\left(f_k\left(x_k, u, \xi_k\right) \in \mathcal{I}_{k+1}\right) = 1\Big\}.$$

4. *Для* $x_k \in \mathcal{O}_k$ *оптимальным управлением на шаге* k *является любой элемент из множества* U_k;

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

$$\mathsf{B}_k\left(x\right) = \max_{u_k \in U_k} \Big\{\mathbf{P}\left(f_k\left(x, u_k, \xi_k\right) \in \mathcal{I}_{k+1}\right) + \mathbf{P}\left(f_k\left(x, u_k, \xi_k\right) \in \mathcal{B}_{k+1}\right) \times \\ \times \mathbf{M}\Big[\mathsf{B}_{k+1}\left(f_k\left(x, u_k, \xi_k\right)\right) \Big| f_k\left(x, u_k, \xi_k\right) \in \mathcal{B}_{k+1}\Big]\Big\}; \tag{3}$$

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

$$\begin{cases} \mathbf{M}\left[\mathsf{B}_{k+1}\left(f_k\left(x, u_k, \xi_k\right)\right)\right] \geqslant \mathbf{P}\left(f_k\left(x, u_k, \xi_k\right) \in \mathcal{I}_{k+1}\right), \\ \mathbf{M}\left[\mathsf{B}_{k+1}\left(f_k\left(x, u_k, \xi_k\right)\right)\right] \leqslant 1 - \mathbf{P}\left(f_k\left(x, u_k, \xi_k\right) \in \mathcal{O}_{k+1}\right); \end{cases} \tag{4}$$

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

$$\underline{\mathsf{B}}_k\left(x\right) \leqslant \mathsf{B}_k\left(x\right) \leqslant \overline{\mathsf{B}}_k\left(x\right), \tag{5}$$

где $\underline{\mathsf{B}}_k\left(x\right)$ *– нижняя, а* $\overline{\mathsf{B}}_k\left(x\right)$ *– верхняя оценки функции Беллмана*

$$\underline{\mathsf{B}}_k\left(x\right) = \sup_{u_k \in U_k} \mathbf{P}\left(f_k\left(x, u_k, \xi_k\right) \in \mathcal{I}_{k+1}\right),$$
$$\overline{\mathsf{B}}_k\left(x\right) = \sup_{u_k \in U_k} \left\{1 - \mathbf{P}\left(f_k\left(x, u_k, \xi_k\right) \in \mathcal{O}_{k+1}\right)\right\},$$

причем $\underline{\mathsf{B}}_N\left(x\right) = \mathsf{B}_N\left(x\right) = \overline{\mathsf{B}}_N\left(x\right)$.

Пункты 1 и 2 леммы 1 устанавливают рекуррентную формулу для вычисления изобелл, достоинством которой является ее “независимость” от функции Беллмана. П. 3 позволяет найти оптимальное управление при $x \in \mathcal{I}_k$, избегая решения сложных оптимизационных задач, путем решения вероятностного уравнения. П. 4 закрывает вопрос об оптимальном управлении при $x \in \mathcal{O}_k$. П. 5 отражает “модифицированное” уравнение Беллмана в области $\mathcal{B}_k$, которое по сути является следствием применения формулы полного математического ожидания в правой части классического уравнения Беллмана. П. 6 является следствием п. 5 и представляет собой двустороннюю оценку для функции правой части уравнения Беллмана, из которой получается двусторонняя оценка для самой функции Беллмана п. 7 (см. [6]).

Отметим, что п. 1–4 леммы 1 позволяют в ряде случаев аналитически найти изобеллы, оптимальное управление при $x \in \mathcal{I}_k$ и $x \in \mathcal{O}_k$ и соответствующие решения задач стохастического программирования для нижней и верхней оценок функции Беллмана [5, 6]. В некоторых частных случаях с использованием оценок п. 5 удается найти также и оптимальное управление при $x \in \mathcal{B}_k$ [5]. Тем не менее даже в задачах с достаточно простой структурой (см., например, задачу оптимального управления портфелем ценных бумаг [20]) с одномерным пространством состояний вопрос об оптимальном управлении при $x \in \mathcal{B}_k$ является нерешенным. Таким образом, к настоящему времени не найдено аналитическое решение указанной задачи для случая заданного горизонта управления (только для случая $N = 1$ [12]). Тем не менее в [6] с использованием результата п. 6 было найдено управление, обеспечивающее асимптотическую оптимальность ($N \to \infty$).

В разделе 2 приводится обобщение результата п. 6, а именно усиленная оценка функции Беллмана. Как будет показано далее, она позволяет найти субоптимальную стратегию в задаче управления портфелем ценных бумаг с заданным горизонтом управления, которая оказывается лучше известных стратегий в смысле функционала вероятности.

4. Основной результат

Поскольку вопрос об оптимальном управлении для областей пространства состояний $\mathcal{I}_k$, $\mathcal{O}_k$ считается закрытым (см. п. 1–4 леммы 1), исследуем новые свойства функции Беллмана для $x \in \mathcal{B}_k$.

Рассмотрим рекуррентные соотношения в дискретном времени $l = \overline{0, k}$ для функции $\underline{\mathsf{B}}_l^k : \mathbb{R}^n \to [0, 1]$

$$(6)\qquad \begin{cases} \underline{\mathsf{B}}_l^k(x) = \sup\limits_{u_k \in U_k} \mathbf{M}\left[\underline{\mathsf{B}}_{l+1}^k\left(f_k\left(x_k, u_k, \xi_k\right)\right) \middle| x_k = x\right], \quad l = \overline{0, k}, \\ \underline{\mathsf{B}}_{k+1}^k(x) = \mathbf{I}_{\mathcal{I}_{k+1}}(x) \end{cases}$$

и функции $\overline{\mathsf{B}}_l^k : \mathbb{R}^n \to [0, 1]$

$$(7)\qquad \begin{cases} \overline{\mathsf{B}}_l^k(x) = \sup\limits_{u_k \in U_k} \mathbf{M}\left[\overline{\mathsf{B}}_{l+1}^k\left(f_k\left(x_k, u_k, \xi_k\right)\right) \middle| x_k = x\right], \quad l = \overline{0, k}, \\ \overline{\mathsf{B}}_{k+1}^k(x) = \mathbf{I}_{\mathcal{I}_{k+1} \cup \mathcal{B}_{k+1}}(x). \end{cases}$$

Граничное условие для указанных последовательностей задано в момент времени k. Нетрудно видеть, что выполнены $\underline{\mathsf{B}}_k^k(x) = \underline{\mathsf{B}}_k(x)$ и $\overline{\mathsf{B}}_k^k(x) = \overline{\mathsf{B}}_k(x)$.

Следующая теорема устанавливает усиленную двустороннюю оценку для функции Беллмана и является основным результатом настоящей статьи.

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

$$\max_{l=\overline{k,N}} \underline{\mathsf{B}}_k^l(x) \leqslant \mathsf{B}_k(x) \leqslant \min_{l=\overline{k,N}} \overline{\mathsf{B}}_k^l(x), \tag{8}$$

где функции $\underline{\mathsf{B}}_k^l(x)$ и $\overline{\mathsf{B}}_k^l(x)$ удовлетворяют рекуррентным соотношениям (6), (7).

Доказательство теоремы приведено в Приложении.

Отметим важные свойства функций в левой и правой частях неравенства (8)

$$\begin{cases} \underline{\mathsf{B}}_l^k(x) = \overline{\mathsf{B}}_l^k(x) = 1, & x \in \mathcal{I}_k, \\ \underline{\mathsf{B}}_l^k(x) \neq 1, \quad \overline{\mathsf{B}}_l^k(x) \neq 0, & x \in \mathcal{B}_k, \\ \underline{\mathsf{B}}_l^k(x) = \overline{\mathsf{B}}_l^k(x) = 0, & x \in \mathcal{O}_k, \end{cases} \qquad \forall l = \overline{0,k}, \quad \forall k = \overline{0,N}.$$

Задачи (6), (7) по сути являются уравнениями метода динамического программирования (МДП) для $2N-2$ задач стохастического оптимального управления с вероятностным терминальным критерием.

Воспользуемся результатами раздела 4 при решении задачи управления портфелем ценных бумаг.

5. Управление портфелем ценных бумаг

Рассмотрим задачу управления портфелем ценных бумаг с вероятностным терминальным критерием [9–12]. Пусть X – размер стартового капитала, x_k – размер капитала на начало k-го года, u_k^1 – доля x_k капитала, вкладываемого в безрисковый инструмент (например, в надежный банк), имеющий доходность b, u_k^2 – доля капитала x_k, вкладываемая в рисковый актив, характеризующийся случайной доходностью $\xi_k \sim \mathcal{R}[-1, a]$, имеющей равномерное распределение на отрезке $[-1, a]$, где $-1 < b < a$. Динамику капитала будем описывать конечно-разностным соотношением [9–12]

$$\begin{cases} x_{k+1} = x_k \left(1 + u_k^1 b + u_k^2 \xi_k\right), \\ x_0 = X, \end{cases} \qquad k = \overline{0,N}. \tag{9}$$

Множество U_k имеет вид

$$U_k = U = \left\{u \in \mathbb{R}^2 : \quad u^1 + u^2 = 1, \quad u^i \geqslant 0, \quad i = \overline{1,2}\right\}, \quad k = \overline{0,N}.$$

Первое ограничение на компоненты вектора u_k моделирует тот факт, что в каждом k-м году осуществляется вложение всего капитала x_k. Второе ограничение на компоненты вектора u_k характеризует запрет операции “short sales” (взятие средств в долг). В отличие от исходных предположений в данном примере начальное состояние $x_0 = X$ является детерминированной величиной.

Рассматривается задача

$$\mathbf{P}\left(-x_{N+1} \leqslant \varphi\right) \to \max_{u(\cdot)\in\mathcal{U}}, \tag{10}$$

где $\varphi < 0$ – фиксированный скаляр. Задача (10) заключается в максимизации вероятности того, что уровень капитала к началу $(N+1)$-го года окажется не меньше фиксированного уровня $(-\varphi)$.

Система (9) рассматривалась в [10–12] для случая $N = 1$ (двухшаговая постановка). При этом в [10] рассмотрен случай усеченного нормального, а в [11, 12] – равномерного распределения случайной величины ξ_k. Настоящая модель соответствует модели из [11, 12], обобщенной на случай заданного горизонта управления N.

Отметим, что в [10] было найдено аналитическое решение задачи (10) для случая $N = 1$, при этом для $k = N$ оптимальное управление и функция Беллмана определяются выражениями:

$$u_N^{\varphi}(x) = v_1(x,\varphi) = \begin{cases} (0,1)^{\mathtt{T}}, & x \in \left(-\infty, -\dfrac{\varphi}{1+b}\right], \\ (1,0)^{\mathtt{T}}, & x \in \left[-\dfrac{\varphi}{1+b}, +\infty\right), \end{cases} \tag{11}$$

$$\mathsf{B}_N(x) = J_1(x,\varphi) = \begin{cases} 0, & x \in \left(-\infty, -\dfrac{\varphi}{1+a}\right], \\ 1 + \dfrac{\varphi}{x(1+a)}, & x \in \left(-\dfrac{\varphi}{1+a}, -\dfrac{\varphi}{1+b}\right), \\ 1, & x \in \left[-\dfrac{\varphi}{1+b}, +\infty\right), \end{cases} \tag{12}$$

а для $k = N-1$ оптимальное управление и функция Беллмана имеют вид:

$$u_{N-1}^{\varphi}(x) = v_2(x,\varphi) = \begin{cases} (0,1)^{\mathtt{T}}, \quad x \in \left(-\infty, -\dfrac{\varphi}{(1+b)^2}\right], \\ \left(1 + \dfrac{\varphi}{(1+a)(1+b)}, -\dfrac{\varphi}{(1+a)(1+b)}\right)^{\mathtt{T}}, \\ \qquad x \in \left(-\dfrac{\varphi}{(1+b)^2}, -\varphi\dfrac{1+\ln\left(\frac{1+a}{1+b}\right)}{(1+b)(1+a)}\right), \\ (1,0)^{\mathtt{T}}, \quad x \in \left[-\varphi\dfrac{1+\ln\left(\frac{1+a}{1+b}\right)}{(1+b)(1+a)}, +\infty\right), \end{cases} \tag{13}$$

(14) $\mathsf{B}_{N-1}(x) = J_2(x,\varphi) =$

$$= \begin{cases} 0, & x \in \left(-\infty, -\dfrac{\varphi}{(1+a)^2}\right], \\ 1 + \dfrac{\varphi\left(1+\ln\left(-\frac{x(1+a)^2}{\varphi}\right)\right)}{x(1+a)^2}, & x \in \left(-\dfrac{\varphi}{(1+a)^2}, -\dfrac{\varphi}{(1+a)(1+b)}\right], \\ 1 + \dfrac{\varphi\left(1+\ln\left(\frac{1+a}{1+b}\right)\right)}{x(1+a)^2}, & x \in \left(-\dfrac{\varphi}{(1+a)(1+b)}, -\dfrac{\varphi\left(1+\ln\left(\frac{1+a}{1+b}\right)\right)}{(1+a)(1+b)}\right], \\ 1 + \dfrac{\varphi(1+b)\left(\ln\left(\frac{1+a}{1+b}\right)\right)}{(\varphi+x(1+a)(1+b))(1+a)}, & x \in \left(-\dfrac{\varphi\left(1+\ln\left(\frac{1+a}{1+b}\right)\right)}{(1+a)(1+b)}, -\dfrac{\varphi}{(1+b)^2}\right), \\ 1, & x \in \left[-\dfrac{\varphi}{(1+b)^2}, +\infty\right). \end{cases}$$

В [6] найдены изобеллы уровней 1 и 0 для функции Беллмана в задаче (10):

$$\mathcal{I}_k = \left[\varphi_k^{\mathcal{I}}, +\infty\right), \quad \varphi_k^{\mathcal{I}} = -\varphi(1+b)^{k-N-1},$$
$$\mathcal{O}_k = \left(-\infty, \varphi_k^{\mathcal{O}}\right], \quad \varphi_k^{\mathcal{O}} = -\varphi(1+a)^{k-N-1}.$$

Из этих соотношений, используя п. 7 леммы 1, получаем верхнюю и нижнюю оценки для функции Беллмана.

Утверждение 1. Справедливы утверждения:

1. *Решение задач стохастического программирования* (6) *для* $k = \overline{0, N-1}$ *и* $l = k$ *имеет вид*

$$\underline{u}_k^k(x) = \arg\max_{u_k \in U_k} \mathbf{M}\left[\underline{\mathsf{B}}_k^k\left(f_k(x_k, u_k, \xi_k)\right) \Big| x_k = x\right] =$$
$$= \arg\max_{u_k \in U} \mathbf{P}\Big(x\left(1 + u_k^1 b + u_k^2 \xi_k\right) \geqslant \varphi_{k+1}^{\mathcal{I}}\Big) = v_1\Big(x, -\varphi_{k+1}^{\mathcal{I}}\Big),$$

а оптимальное значение целевой функции равно

$$\underline{\mathsf{B}}_k^k(x) = \max_{u_k \in U_k} \mathbf{M}\left[\underline{\mathsf{B}}_k^k\left(f_k(x_k, u_k, \xi_k)\right) \Big| x_k = x\right] =$$
$$= \max_{u_k \in U} \mathbf{P}\Big(x\left(1 + u_k^1 b + u_k^2 \xi_k\right) \geqslant \varphi_{k+1}^{\mathcal{I}}\Big) = J_1\Big(x, -\varphi_{k+1}^{\mathcal{I}}\Big);$$

2. *Решение задач стохастического программирования* (7) *для* $k = \overline{0, N-1}$ *и* $l = k$ *имеет вид*

$$\overline{u}_k^k(x) = \arg\max_{u_k \in U_k} \mathbf{M}\left[\overline{\mathsf{B}}_k^k\left(f_k(x_k, u_k, \xi_k)\right) \Big| x_k = x\right] =$$
$$= \arg\max_{u_k \in U} \mathbf{P}\Big(x\left(1 + u_k^1 b + u_k^2 \xi_k\right) \geqslant \varphi_{k+1}^{\mathcal{O}}\Big) = v_1\Big(x, -\varphi_{k+1}^{\mathcal{O}}\Big),$$

а оптимальное значение целевой функции равно

$$\overline{\mathsf{B}}_k^k(x) = \max_{u_k \in U_k} \mathbf{M}\left[\overline{\mathsf{B}}_k^k(f_k(x_k, u_k, \xi_k)) \,\middle|\, x_k = x\right] =$$
$$= \max_{u_k \in U} \mathbf{P}\big(x\left(1 + u_k^1 b + u_k^2 \xi_k\right) \geqslant \varphi_{k+1}^{\mathcal{O}}\big) = J_1\big(x, -\varphi_{k+1}^{\mathcal{O}}\big),$$

где функции $v_1 : \mathbb{R} \times \mathbb{R} \to \mathbb{R}^2$ *и* $J_1 : \mathbb{R} \times \mathbb{R} \to [0, 1]$ *определяются выражениями* (11) *и* (12) *соответственно.*

Доказательство утверждения 1 заключается в подстановке параметров $(-\varphi_{k+1}^{\mathcal{I}})$ и $(-\varphi_{k+1}^{\mathcal{O}})$ вместо параметра φ в выражения для оптимального управления и функции Беллмана при $k = N$ (12).

Как видно из приведенных выражений, верхняя $\overline{\mathsf{B}}_k^k(x)$ и нижняя $\underline{\mathsf{B}}_k^k(x)$ оценки функции Беллмана совпадают с функцией Беллмана при $k = N$ с точностью до параметров φ, $-\varphi_{k+1}^{\mathcal{I}}$ и $-\varphi_{k+1}^{\mathcal{O}}$. Отсюда с учетом рекуррентных соотношений (6) и (7) делаем вывод, что функции $\mathsf{B}_{N-1}(x)$, $\underline{\mathsf{B}}_k^{k+1}(x)$ и $\overline{\mathsf{B}}_k^{k+1}(x)$ тоже совпадают с точностью до тех же параметров. Приведем решение задач стохастического программирования (6) и (7).

Утверждение 2. Справедливы утверждения:

1. *Решение задач стохастического программирования* (6) *для* $k = \overline{0, N-2}$ *и* $l = k + 1$ *имеет вид*

$$\underline{u}_k^{k+1}(x) = \arg\max_{u_k \in U_k} \mathbf{M}\left[\underline{\mathsf{B}}_{k+1}^{k+1}(f_k(x_k, u_k, \xi_k)) \,\middle|\, x_k = x\right] = v_2\Big(x, -\varphi_{k+2}^{\mathcal{I}}\Big),$$

а оптимальное значение целевой функции равно

$$\underline{\mathsf{B}}_k^{k+1}(x) = \max_{u_k \in U_k} \mathbf{M}\left[\underline{\mathsf{B}}_{k+1}^{k+1}(f_k(x_k, u_k, \xi_k)) \,\middle|\, x_k = x\right] = J_2\Big(x, -\varphi_{k+2}^{\mathcal{I}}\Big);$$

2. *Решение задач стохастического программирования* (7) *для* $k = \overline{0, N-2}$ *и* $l = k + 1$ *имеет вид*

$$\overline{u}_k^{k+1}(x) = \arg\max_{u_k \in U_k} \mathbf{M}\left[\overline{\mathsf{B}}_{k+1}^{k+1}(f_k(x_k, u_k, \xi_k)) \,\middle|\, x_k = x\right] = v_2\Big(x, -\varphi_{k+2}^{\mathcal{O}}\Big),$$

а оптимальное значение целевой функции равно

$$\overline{\mathsf{B}}_k^{k+1}(x) = \max_{u_k \in U_k} \mathbf{M}\left[\overline{\mathsf{B}}_{k+1}^{k+1}(f_k(x_k, u_k, \xi_k)) \,\middle|\, x_k = x\right] = J_2\Big(x, -\varphi_{k+2}^{\mathcal{O}}\Big),$$

где функции $v_2 : \mathbb{R} \times \mathbb{R} \to \mathbb{R}^2$ *и* $J_2 : \mathbb{R} \times \mathbb{R} \to [0, 1]$ *определяются выражениями* (13) *и* (14) *соответственно.*

Доказательство утверждения 2 заключается в подстановке параметров $(-\varphi_{k+2}^{\mathcal{I}})$ и $(-\varphi_{k+2}^{\mathcal{O}})$ вместо параметра φ в выражения для оптимального управления и функции Беллмана при $k = N - 1$ (14).

На рис. 1 изображены двусторонние оценки для функции Беллмана для следующих значений параметров: $N = 2$, $k = 0$, $\varphi = -1{,}2$, $a = 1{,}2$, $b = 0{,}05$.

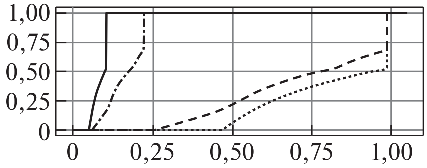

Рис. 1. Двусторонние оценки функции Беллмана. Жирная линия – $\overline{\mathsf{B}}_k^k(x)$, штрихпунктирная – $\overline{\mathsf{B}}_k^{k+1}(x)$, пунктирная – $\underline{\mathsf{B}}_k^{k+1}(x)$, точечная – $\underline{\mathsf{B}}_k^k(x)$.

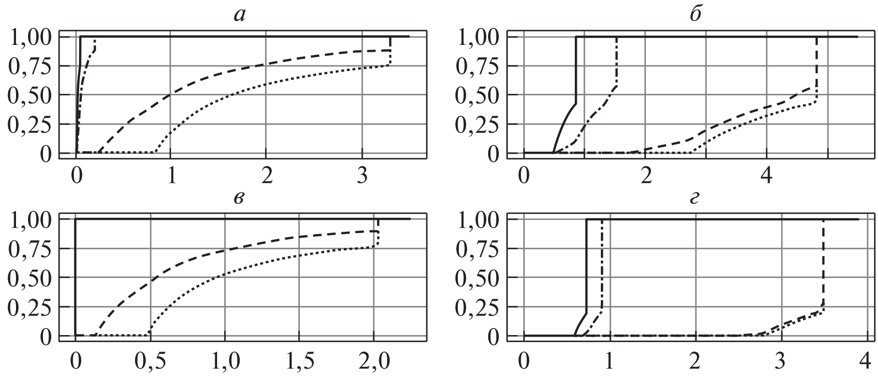

Рис. 2. Двусторонние оценки функции Беллмана.

Из графика видно, что нижняя и верхняя оценки (8) для функции Беллмана при $l = k + 1$ существенно точнее оценок при $k = l$, т.е. найденных в [5]. Другими словами, выполнены $\overline{\mathsf{B}}_k^k(x) > \overline{\mathsf{B}}_k^{k+1}(x)$ и $\underline{\mathsf{B}}_k^k(x) < \underline{\mathsf{B}}_k^{k+1}(x)$. Проведем еще несколько экспериментов при других значениях параметров, занесенных в табл. 1. Графики оценок функции Беллмана для случаев а–г приведены на рис. 2.

Из графиков рис. 2 видно, что в некоторых случаях нижние, а в некоторых – верхние оценки функции Беллмана практически совпадают. Тем не менее во всех случаях двусторонняя оценка (8) при $l = k + 1$ оказывается точнее оценки при $l = k$.

Исследуем теперь усиленную оценку для функции оптимального значения вероятностного критерия. Отметим, что стратегии $\underline{u}_k^k(x)$ и $\underline{u}_k^{k+1}(x)$ являются субоптимальными в исходной задаче, поскольку максимизирует нижние границы (см. (6) при $l = k$ и $l = k + 1$) функции Беллмана. Причем в [6] доказано, что стратегия $\underline{u}_k^k(x)$ при определенных условиях является

Таблица 1. Значение вероятностного критерия

№ эксперимента	N	φ	a	b
а	2	−4	3,2	0,05
б	2	−10	1,12	0,2
в	6	−3	3,5	0,05
г	6	−15	0,5	0,2

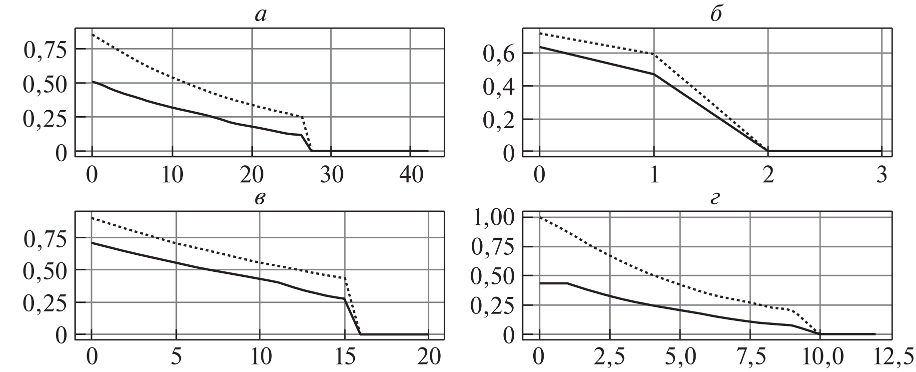

Рис. 3. Функции ΔF^0 (точечная линия) и ΔF^1 (жирная линия) в зависимости от N.

асимптотически оптимальной ($N \to \infty$) и совпадает с так называемой “рисковой стратегией”. Введем в рассмотрение функции $\underline{F}^l : \mathbb{R} \times \mathbb{R} \times \mathbb{N} \to [0,1]$ и $\overline{F}^l : \mathbb{R} \times \mathbb{R} \times \mathbb{N} \to [0,1]$:

$$\underline{F}^l(\varphi, X, N) = \underline{\mathsf{B}}_0^l(X), \quad \overline{F}^l(\varphi, X, N) = \overline{\mathsf{B}}_0^l(X).$$

В [6] было доказано, что

$$\lim_{N\to\infty} \left[\overline{F}^0(\varphi, X, N) - \underline{F}^0(\varphi, X, N)\right] = 0. \tag{15}$$

Используя утверждения 1 и 2, можно получить явный вид функций $\underline{F}^l(\varphi, X, N)$ и $\overline{F}^l(\varphi, X, N)$ при $l = 0$ и $l = 1$.

Проведем сравнение функций

$$\Delta F^0(\varphi, X, N) = \overline{F}^0(\varphi, X, N) - \underline{F}^0(\varphi, X, N) \text{ и}$$
$$\Delta F^1(\varphi, X, N) = \overline{F}^1(\varphi, X, N) - \underline{F}^1(\varphi, X, N),$$

как функций аргумента N, т.е. горизонта управления. Рассмотрим 4 эксперимента, входные данные которых занесены в табл. 2. Отметим, что входные параметры подбирались так, чтобы для всех $N > 2$ верхние границы функции оптимального значения вероятностного критерия не были равны нулю (в противном случае оптимальное значение критерия равно нулю). На рис. 3 изображены их графики функций ΔF^0 и ΔF^1 от аргумента N для случаев а–г.

Таблица 2. Значение вероятностного критерия

№ эксперимента	X	φ	a	b
а	1,01	−4	3,2	0,05
б	1,01	−2,2	1,12	0,2
в	1,01	−2,5	1,5	0,05
г	1,01	−10	5,5	0,2

Таблица 3. Значения вероятностного критерия

Стратегия	$a = 1{,}12$	$a = 1{,}2$	$a = 1{,}3$	$a = 1{,}9$
$\underline{u}_k^{k+1}$	0,611	0,645	0,669	0,783
$\underline{u}_k^{k}$	0,594	0,623	0,664	0,769
u_k^L	0	0,501	0,548	0,592

Из рис. 3 видно, что оценка функции Беллмана при $l = k+1$ "сходится" быстрее оценки функции Беллмана при $l = k$.

Аналитическое исследование свойств найденных стратегий затруднено их нетривиальной структурой, в частности нелинейностью по состоянию. Проведем сравнение известных стратегий управления портфелем ценных бумаг (см. раздел 5) со стратегией $\underline{u}_k^{k+1}(x)$ (п. 1 утверждения 2). Рассматриваются рисковая стратегия [21] (п. 1 утверждения 1)

$$\underline{u}_k^k(x) = u_k^R(x) = \begin{cases} (1,\ 0)^{\mathrm{T}}, & x \geqslant -\varphi(1+b)^{k-N-1}, \\ (0,\ 1)^{\mathrm{T}}, & x < -\varphi(1+b)^{k-N-1}, \end{cases} \qquad k = \overline{0,N},$$

и логарифмическая стратегия Келли [20]

$$(16)\quad u_k^L = \arg\max_{u_k \in U} \mathbf{M}\Big[\ln\left(1 + u_k^1 b + u_k^2 \xi_k\right)\Big] =$$
$$= \arg\max_{u_k \in U} \frac{1}{(1+a)\,u_k^2}\left[\left(1 + b\left(1+u_k^2\right) + a u_k^2\right)\ln\left(\frac{1 + b\left(1+u_k^2\right) + a u_k^2}{e}\right) - \right.$$
$$\left. - \left(u_k^1 + b\left(1+u_k^2\right)\right)\ln\left(\frac{u_k^1 + b\left(1+u_k^2\right)}{e}\right)\right], \quad k = \overline{0,N}.$$

Оптимизационную задачу (16) можно свести [20] к задаче выпуклого программирования со скалярной стратегией

$$(17)\quad \tilde{u}_k^L = \arg\max_{0 \leqslant \tilde{u}_k \leqslant 1} \frac{1}{(1+a)\tilde{u}_k}\left[(1 + b\,(1+\tilde{u}_k) + a\tilde{u}_k)\ln\left(\frac{1 + b(1+\tilde{u}_k) + a\tilde{u}_k}{e}\right) - \right.$$
$$\left. - (1 - \tilde{u}_k + b\,(1+\tilde{u}_k))\ln\left(\frac{1 - \tilde{u}_k + b\,(1+\tilde{u}_k)}{e}\right)\right], \quad k = \overline{0,N},$$

при этом $u_k^L = \left(1 - \tilde{u}_k^L,\ \tilde{u}_k^L\right)^{\mathrm{T}}$.

Зададим следующие числовые значения параметров [12]: $N = 3,\ X = 1$, $b = 0{,}05$, $\varphi = -1{,}2$. Параметр a будем варьировать. Для численного построения оценок вероятностного критерия $P_\varphi(u(\cdot))$ используется метод Монте Карло по выборке из 2000 реализаций, а для численного поиска логарифмической стратегии (задача (17)) используется метод дихотомии. Результаты вычислений сведены в табл. 3.

Из табл. 3 видно, что стратегия $\underline{u}_k^{k+1}$ превосходит известные стратегии по значению вероятностного критерия.

6. Заключение

На основе метода динамического программирования и изобелл для задач оптимального управления дискретными стохастическими системами с вероятностным терминальным критерием получена усиленная двусторонняя оценка для функции Беллмана. Показано, что данная оценка является точнее оценки [6], полученной с использованием изобелл уровней 1 и 0. По аналогии с [6] с использованием усиленной двусторонней оценки выведены соотношения для поиска субоптимальной стратегии, эфективность которой проверена на задаче оптимального управления портфелем ценных бумаг с вероятностным критерием в многошаговой постановке. В данной задаче субоптимальная стратегия найдена в явном виде. Результаты численного эксперимента показали, что полученная стратегия является лучше известных стратегий в смысле значений функционала вероятности.

ПРИЛОЖЕНИЕ

Доказательство теоремы. Рассмотрим шаг $k = N$ метода динамического программирования. С учетом п. 7 леммы 1 получаем, что неравенства для функции Беллмана превращаются в равенства $\underline{\mathsf{B}}_N(x) = \mathsf{B}_N(x) = \overline{\mathsf{B}}_N(x)$.

Перейдем к шагу $k = N-1$ метода динамического программирования. Используя п. 7 леммы 1 получаем оценку

$$\underline{\mathsf{B}}_{N-1}(x) \leqslant \mathsf{B}_{N-1}(x) \leqslant \overline{\mathsf{B}}_{N-1}(x).$$

Легко видеть, что функции левых частей соотношений (6) и (7) в силу граничных условий совпадают с нижей и верхней оценками для функции Беллмана (п. 7 леммы 1).

Пусть теперь $k = N-2$. Выполнив подстановку

$$x = x_{N-1} = f_{N-2}(x_{N-2}, u_{N-2}, \xi_{N-2}),$$

с учетом свойств математического ожидания получим

$$\begin{cases} \mathbf{M}\left[\mathsf{B}_{N-1}\left(f_{N-2}\left(x, u_{N-2}, \xi_{N-2}\right)\right)\right] \geqslant \mathbf{M}\left[\underline{\mathsf{B}}_{N-1}\left(f_{N-2}\left(x, u_{N-2}, \xi_{N-2}\right)\right)\right], \\ \mathbf{M}\left[\mathsf{B}_{N}\left(f_{N-2}\left(x, u_{N-2}, \xi_{N-2}\right)\right)\right] \leqslant \mathbf{M}\left[\overline{\mathsf{B}}_{N-1}\left(f_{N-2}\left(x, u_{N-2}, \xi_{N-2}\right)\right)\right], \end{cases}$$

откуда, взяв супремум по $u_{N-2} \in U_{N-2}$ в обеих частях неравенств, получаем

$$\begin{cases} \mathsf{B}_{N-2}(x) \geqslant \sup\limits_{u_{N-2}\in U_{N-2}} \mathbf{M}\left[\underline{\mathsf{B}}_{N-1}\left(f_{N-2}\left(x, u_{N-2}, \xi_{N-2}\right)\right)\right] = \underline{\mathsf{B}}_{N-2}^{N-1}(x), \\ \mathsf{B}_{N-2}(x) \leqslant \sup\limits_{u_{N-2}\in U_{N-2}} \mathbf{M}\left[\overline{\mathsf{B}}_{N-1}\left(f_{N-2}\left(x, u_{N-2}, \xi_{N-2}\right)\right)\right] = \overline{\mathsf{B}}_{N-2}^{N-1}(x), \end{cases}$$

откуда с учетом неравенства п. 7 леммы 1, а также равенств $\underline{\mathsf{B}}_k^k(x) = \underline{\mathsf{B}}_k(x)$ и $\overline{\mathsf{B}}_k^k(x) = \overline{\mathsf{B}}_k(x)$ для всех $k = \overline{0, N}$ имеем, что

$$\max_{l=\overline{N-2,N}} \underline{\mathsf{B}}_{N-2}^l(x) \leqslant \mathsf{B}_{N-2}(x) \leqslant \max_{l=\overline{N-2,N}} \overline{\mathsf{B}}_{N-2}^l(x). \tag{П.1}$$

Теперь рассмотрим произвольный шаг $k+1$, где $k \leqslant N-2$. Предположим, что с учетом (П.1) выполнено неравенство

$$\max_{l=\overline{k+1,N}} \underline{\mathsf{B}}_{k+1}^{l}(x) \leqslant \mathsf{B}_{k+1}(x) \leqslant \min_{l=\overline{k+1,N}} \overline{\mathsf{B}}_{k+1}^{l}(x), \tag{П.2}$$

тогда получаем эквивалентную систему неравенств

$$\underline{\mathsf{B}}_{k+1}^{l}(x) \leqslant \mathsf{B}_{k+1}(x) \leqslant \overline{\mathsf{B}}_{k+1}^{l}(x), \quad l = \overline{k+1,N}.$$

Выполнив подстановку $x = x_{k+1} = f_k(x_k, u_k, \xi_k)$, с учетом свойств математического ожидания получим систему неравенств для всех $l = \overline{k+1,N}$

$$\begin{cases} \mathbf{M}\left[\mathsf{B}_{k+1}(f_k(x_k,u_k,\xi_k)) \middle| x_k = x\right] \geqslant \mathbf{M}\left[\underline{\mathsf{B}}_{k+1}^{l}(f_k(x_k,u_k,\xi_k)) \middle| x_k = x\right], \\ \mathbf{M}\left[\mathsf{B}_{k+1}(f_k(x_k,u_k,\xi_k)) \middle| x_k = x\right] \leqslant \mathbf{M}\left[\overline{\mathsf{B}}_{k+1}^{l}(f_k(x_k,u_k,\xi_k)) \middle| x_k = x\right], \end{cases}$$

откуда, взяв супремум по $u_k \in U_k$ в обеих частях неравенств, получаем

$$\begin{cases} \mathsf{B}_k(x) \geqslant \sup\limits_{u_k \in U_k} \mathbf{M}\left[\underline{\mathsf{B}}_{k+1}^{l}(f_k(x_k,u_k,\xi_k)) \middle| x_k = x\right] = \underline{\mathsf{B}}_{k}^{l}(x), \\ \mathsf{B}_k(x) \leqslant \sup\limits_{u_k \in U_k} \mathbf{M}\left[\overline{\mathsf{B}}_{k+1}^{l}(f_k(x_k,u_k,\xi_k)) \middle| x_k = x\right] = \overline{\mathsf{B}}_{k}^{l}(x), \end{cases} \qquad l = \overline{k+1,N},$$

откуда имеем

$$\underline{\mathsf{B}}_{k}^{l}(x) \leqslant \mathsf{B}_{k}(x) \leqslant \overline{\mathsf{B}}_{k}^{l}(x), \quad l = \overline{k+1,N}.$$

С учетом равенств $\underline{\mathsf{B}}_{k}^{k}(x) = \underline{\mathsf{B}}_{k}(x)$, $\overline{\mathsf{B}}_{k}^{k}(x) = \overline{\mathsf{B}}_{k}(x)$ и неравенства п. 7 леммы 1 получаем

$$\underline{\mathsf{B}}_{k}^{l}(x) \leqslant \mathsf{B}_{k}(x) \leqslant \overline{\mathsf{B}}_{k}^{l}(x), \quad l = \overline{k,N},$$

откуда окончательно имеем, что

$$\max_{l=\overline{k,N}} \underline{\mathsf{B}}_{k}^{l}(x) \leqslant \mathsf{B}_{k}(x) \leqslant \min_{l=\overline{k,N}} \overline{\mathsf{B}}_{k}^{l}(x). \tag{П.3}$$

В результате доказано, что из (П.2) следует (П.3) для всех k, поэтому с учетом математической индукции из (П.1) следует и (П.3).

Теорема доказана.

СПИСОК ЛИТЕРАТУРЫ

1. *Малышев В.В., Кибзун А.И.* Анализ и синтез высокоточного управления летательными аппаратами. М.: Машиностроение, 1987.
2. *Красильщиков М.Н., Малышев В.В., Федоров А.В.* Автономная реализация динамических операций на геостационарной орбите. I. Формализация задачи управления // Изв. РАН. ТиСУ. 2015. № 6. С. 82–95.

3. *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.

4. *Малышев В.В., Старков А.В., Федоров А.В.* Совмещение задач удержания и уклонения в окрестности опорной геостационарной орбиты // Вестн. Московского городского педагогич. ун-та. Сер. Экономика. 2013. № 1. С 68–74.

5. *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.

6. *Азанов В.М., Кан Ю.С.* Двухсторонняя оценка функции Беллмана в задачах стохастического оптимального управления дискретными системами с вероятностным критерием // 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.

7. *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.

8. *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.

9. *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.

10. *Кибзун А.И., Кузнецов Е.А.* Оптимальное управление портфелем ценных бумаг // АиТ. 2001. № 9. С. 101–113.

    Kibzun A.I., Kuznetsov E.A. Optimal Control of the Portfolio // Autom. Remote Control. 2001. V. 62. No. 9. P. 1489–1501.

11. *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.

12. *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.

13. *Вишняков Б.В., Кибзун А.И.* Оптимизация двухшаговой модели изменения капитала по различным статистическим критериям // АиТ. 2005. № 7. С. 126–143.

    Vishnyakov B.V., Kibzun A.I. A Two-Step Capital Variation Model: Optimization by Different Statistical Criteria // Autom. Remote Control. 2005. V. 66. No. 7. P. 1137–1152.

14. *Jasour A.M., Aybat N.S., Lagoa C.M.* Semidefinite Programming For Chance Constrained Optimization Over Semialgebraic Sets // SIAM J. Optimization. 2015. № 25 (3). P. 1411–1440.

15. *Jasour A.M., Lagoa C.M.* Convex Chance Constrained Model Predictive Control // arXiv preprint arXiv:1603.07413, 2016.

16. *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.

17. *Hsieh Chung-Han, Barmish B. Ross* On Drawdown-Modulated Feedback Control in Stock Trading // IFAC Proc. July 2017. V. 50. No. 1. P. 952–958.

18. *Hsieh Chung-Han, Barmish B. Ross, Gubner J.* At What Frequency Should the Kelly Bettor Bet? // Proc. IEEE Amer. Control Conf. (ACC). Milwaukee. WI. January 2018.

19. *Норкин Б.В.* О численном решении задачи стохастического оптимального управления дивидендной политикой страховой компании // Компьютерная математика. 2014. № 1. С. 131–140.

20. *Кан Ю.С., Кибзун А.И.* Задачи стохастического программирования с вероятностными критериями. М.: Физматлит, 2009.

21. *Kibzun A.I., Kan Yu.S.* Stochastic Programming Problems with Probability and Quantile Functions. Chichester: Wiley, 1996.

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

Поступила в редакцию 22.04.2018

После доработки 29.10.2018

Принята к публикации 08.11.2018