УДК 519.217.8;330.46

МАКСИМИЗАЦИЯ ВРЕМЕНИ ВЫХОДА УПРАВЛЯЕМОГО СЛУЧАЙНОГО БЛУЖДАНИЯ НА ГРАНИЦУ КВАДРАНТА[1]

С.В. Анулова

В каждый момент времени управление воздействует только на одну из двух координат блуждания, величина его фиксирована. Находится стратегия, оптимальная сразу для ряда критериев (среднее время выхода, вероятность невыхода и др.). Решение уравнения Беллмана для этого не потребовалось благодаря свойствам модели: симметрия, монотонность, возможность декомпозиции. Модель возникла при изучении экономических задач.

Ключевые слова: управляемое случайное блуждание, уравнение Беллмана.

ВВЕДЕНИЕ

В работе рассматривается задача управления марковским процессом. Известен подход к решению таких задач: выписывается уравнение Беллмана, находится его решение — функция выигрыша, отправляясь от функции выигрыша, находят оптимальную стратегию [1]. Главная трудность — найти функцию выигрыша. В некоторых случаях эту трудность можно обойти: напрямую предлагается некоторая стратегия и доказывается ее оптимальность. Это удается сделать, если рассматриваемый объект допускает декомпозицию, обладает свойствами симметрии или монотонности. Именно таким оказался наш объект, и мы реализовали изложенную программу. Общая постановка задачи взята из работы [2].

Рассмотрим две различные государственные политики налогообложения компаний: республиканская политика дает налоговые льготы успешным компаниям, в то время как демократическая политика дает льготы слабым компаниям, чтобы уберечь их от банкротства и не допустить роста безработицы. Какая политика эффективнее? В работе [2] показано, что это зависит от критерия оптимизации. Если требуется минимизировать среднее число обанкротившихся компаний, то численные методы дают оптимальную стратегию в виде комбинации республиканской и демократической политик. Если требуется максимизировать вероятность отсутствия банкротств, то аналитически доказана оптимальность демократической политики. Мы усиливаем последний результат: демократическая политика оптимальна и для целого ряда других критериев. В качестве математической модели мы рассматриваем случайное блуждание в целочисленном неотрицательном квадранте. Первая (вторая) координата выражает размер капитализации первой (второй) фирмы, время дискретно. Управление реализуется как придание блужданию некоторой степени асимметрии (скачок в положительном направлении более вероятен, чем в отрицательном), но только одной из координат в каждый момент времени.

Рассмотрим задачу максимизации среднего дисконтированного времени выхода на границу квадранта (без дисконтирования среднее бесконечно). Цель — найти оптимальную стратегию. В симметричном случае, т. е. когда допустимая степень асимметрии одинакова для обеих координат, мы нашли решение. Оптимальная стратегия — действовать на ту координату, которая ближе к нулю.

При этом мы реализовали идею [2]: вместо того чтобы решать соответствующее уравнение Беллмана, прямо доказать оптимальность определенной стратегии.

В работе [2] задача максимизации вероятности невыхода на границу рассматривалась для винеровского процесса с управляемым сносом, и для уравнения Беллмана было получено аналитическое решение. Для случайного блуждания это вряд ли возможно.

[1] Работа выполнена при финансовой поддержке РФФИ (гранты № 07-08-00739 и 07-01-00541) и Президиума РАН (программа № 29).

1. ОБОЗНАЧЕНИЯ

Обозначим: $\mathbb{Z}^1$ — множество целых чисел, $\mathbb{Z}^1_+$ — множество неотрицательных целых чисел, $\mathbb{Z}^2_+$ — неотрицательный квадрант $\mathbb{Z}^1_+ \times \mathbb{Z}^1_+ = \{(i, j), i, j, \notin \mathbb{Z}^1\}$. Далее, $(\Omega, \mathcal{F}, F, \mathbf{P})$ — вероятностное пространство с потоком σ-алгебр $F = (\mathcal{F}(n), n = 0, 1, 2, ...)$.

Все случайные процессы рассматриваются в дискретном времени $n = 0, 1, ...$ Обозначим через ΔX процесс, составленный из приращений процесса X: $\Delta X(n + 1) = X(n + 1) - X(n)$, $n = 0, 1, ...$

Константа $p \in \left[\frac{1}{2}, 1\right]$ обозначает вероятность скачка вправо асимметричного блуждания.

2. УПРАВЛЯЕМОЕ СЛУЧАЙНОЕ БЛУЖДАНИЕ В КВАДРАНТЕ

Процесс X называется случайным блужданием в неотрицательном квадранте, если $X(n) \in \mathbb{Z}^2_+$, $n = 0, 1, ...$, и $\Delta X(n) \in \{(\pm 1, \pm 1)\}$, $n = 1, ...$ Таким образом, в каждый момент времени каждая из координат переходит в одну из соседних точек.

Перейдем к определению управляемого случайного блуждания. Пусть дан F-согласованный процесс $u(n) \in \{1, 2\}$, $n = 0, 1, ...,$ — стратегия управления. Для $x \in \mathbb{Z}^2_+$ обозначим $X^{x,u}$ процесс, определяемый соотношениями:

$$X^{x,u}(0) = x,$$

$$\mathbf{P}(\Delta X^{x,u}(n + 1) = (1, 1)|\mathcal{F}(n)) = \frac{1}{2}p,$$

$$\mathbf{P}(\Delta X^{x,u}(n + 1) = (-1, -1)|\mathcal{F}(n)) = \frac{1}{2}(1 - p),$$

$$\mathbf{P}(\Delta X^{x,u}(n + 1) = (1, -1)|\mathcal{F}(n)) = \begin{cases} \frac{1}{2}p, \text{ если } u(n) = 1, \\ \frac{1}{2}(1-p), \text{ если } u(n) = 2, \end{cases}$$

$$\mathbf{P}(\Delta X^{x,u}(n + 1) = (-1, 1)|\mathcal{F}(n)) = \begin{cases} \frac{1}{2}(1-p), \text{ если } u(n) = 1, \\ \frac{1}{2}p, \text{ если } u(n) = 2, \end{cases}$$

$n = 0, 1, 2, ...$ Совокупность всех процессов $X^{x,u}$, где u пробегает множество всех стратегий, называется управляемым случайным блужданием в неотрицательном квадранте.

Пусть дана неубывающая функция f: $\mathbb{Z}^1_+ \cup \infty \to \to [0, 1]$. Рассмотрим задачу $\mathbf{E}f(\tau(X^{x,u})) \to \max$, где максимум берется по множеству всех стратегий u, а τ — момент выхода траектории на границу квадранта.

Основная теорема. *Оптимальная стратегия (не единственная на биссектрисе квадранта) имеет вид:*

$$u^*(x) = \begin{cases} i, \text{ если } x_i = \min\{x_1, x_2\}, x_1 \ne x_2, i = 1, 2, \\ 1 \text{ если } x_1 = x_2. \blacklozenge \end{cases}$$

Беря в качестве f функцию, принимающую значение 1 в бесконечности и 0 во всех остальных точках (соответственно, $-e^{-\lambda n}$, где λ — положительная константа), получим задачу максимизации вероятности невыхода на границу (соответственно, среднего дисконтированного времени выхода на границу квадранта). Поясним:

$\mathbf{E}\sum_0^{\tau-1} e^{-\lambda t} = \mathbf{E}\frac{1 - e^{-\lambda\tau}}{1 - e^{-\lambda}}$, максимизация правой части по τ эквивалентна максимизации $-e^{-\lambda\tau}$.

Замечание. Утверждение об оптимальности стратегии u^* нетривиально только в том случае, если выигрыш действительно зависит от стратегии. Например, в задаче максимизации вероятности невыхода на границу гипотетически эта вероятность могла бы равняться нулю для всех стратегий. Покажем, что это не так. Сравним две стратегии: $u \equiv 1$ и $u(2n) = 1$, $u(2n + 1) = 2$, $n = 0, 1, ...$ Для первой стратегии вероятность равна нулю при любом начальном условии, потому что вторая координата — неуправляемое случайное блуждание, которое с вероятностью 1 достигает 0. Для второй стратегии, напротив, при любом начальном условии вероятность положительна: координаты независимы, и каждая подобна винеровскому процессу со сносом [3]. ♦

3. ДЕКОМПОЗИЦИЯ

Введем в квадранте новые координаты $(Y_1, Y_2) = \frac{1}{2}(X_1 - X_2, X_1 + X_2)$. В этих координатах динамика блуждания такова:

если $\Delta X(n) \in \{(-1, 1), (1, -1)\}$,
то $\Delta Y_2(n) = 0$, а $\Delta Y_1(n) = \pm 1$;

если $\Delta X(n) \in \{(1, 1), (-1, -1)\}$,
то $\Delta Y_1(n) = 0$, а $\Delta Y_2(n) = \pm 1$.

Таким образом, координата Y_2 движется с течением времени так: либо стоит на месте, либо совершает скачок ±1. В те (и только те) моменты, когда координата Y_2 стоит, движется координата Y_1, совершая скачок ± 1.

Пусть X — управляемое случайное блуждание. Легко видеть, что в новых координатах (Y_1, Y_2) управление действует только на координату Y_1. Таким образом, задача автоматически свелась к управлению одномерным процессом в множестве Z^1.

4. УПРАВЛЯЕМЫЕ ОДНОМЕРНЫЕ СЛУЧАЙНЫЕ БЛУЖДАНИЯ

Перейдем к определению управляемого одномерного случайного блуждания. Пусть на пространстве $(\Omega, \mathcal{F}, F, \mathbf{P})$ задано простое одномерное F-случайное блуждание $W(n)$, $n = 0, 1, 2, ...$, т. е. F-согласованный процесс со значениями в множестве $\mathbb{Z}^1$ такой, что $W(0) = 0$ и

$$\mathbf{P}(\Delta W(n) = 1|F(n-1)) = $$
$$= \mathbf{P}(\Delta W(n) = -1|F(n-1)) = \frac{1}{2}, \quad n = 1, 2, ...$$

И пусть $\xi(n)$, $n = 1, 2, ...$ — одинаково распределенные F-бернуллиевские случайные величины, т. е. F-согласованный процесс со значениями в множестве $\{0, 1\}$ такой, что для некоторого $p \in [0, 1]$

$$\mathbf{P}(\xi(n) = 1|F(n-1)) = p,$$

$$\mathbf{P}(\xi(n) = 0|F(n-1)) = 1 - p.$$

Наконец, пара $(\Delta W(n), \xi(n))$ независима и не зависит от $F(n-1)$, $n = 1, 2, ...$ Пусть $u(n) \in \{-1, +1\}$, $n = 0, 1, ...$ — F — согласованный процесс (стратегия управления). Для $y \in \mathbb{Z}^1_+$ обозначим $Y^{y,u}$ процесс, определяемый соотношениями: $Y^{y,u}(0) = y$,

$$\Delta Y^{y,u}(n) = \begin{cases} \Delta W(n), \text{если } \xi(n) = 0, \\ u(n-1), \text{если } \xi(n) = 1, \end{cases}$$

$n = 1, 2, ...$ Таким образом, процесс $Y^{y,u}$ совершает скачок вместе с блужданием W, если $\xi = 0$, и совершает скачок в направлении u, если $\xi = 1$ (ξ можно рассматривать как «шанс для управления»). Процесс $Y^{y,u}$ допускает представление

$$Y^{y,u}(n) = y + W(n) + \sum_{k=1}^{n} \xi(n) f(u(n-1), \Delta W(n)),$$
$$n = 0, 1, ...$$

где

$$f(1, 1) = 0, \quad f(1, -1) = 2,$$

$$f(-1, -1) = 0, \quad f(-1, 1) = -2.$$

Таким образом, управляющее воздействие аддитивно по отношению к исходному ведущему блужданию. Ради этой аддитивности мы выбрали такую конструкцию одномерного управляемого блуждания (сравните с конструкцией двумерного), потому что она нужна при доказательстве теоремы 2.

Множество всех стратегий u будет обозначаться U. Для полной строгости следует писать $Y^{y,u}(W, \xi)$ и $Y^{y,u}(W, \xi, n)$, чтобы явно указать ведущие процессы управляемого случайного блуждания.

Для любой функции ψ: $\mathbb{Z}^1_+ \to \{-1, +1\}$ и любого $y \in \mathbb{Z}^1_+$ можно однозначно определить $\psi^1 \in U$ и Y^{y,ψ^1}, удовлетворяющие $\psi^1(n) = y(Y^{y,\psi^1}(n))$, $n = 1, 2, ...$ (рекуррентно). Соответственно, для данного $y \in \mathbb{Z}^1_+$ можно определить стратегию u^*, отвечающую функции $-\mathrm{sgn}$ и зависящую от y. Таким образом, существует единственная стратегия $u^* \in U$, удовлетворяющая соотношению $u^*(n) = -\mathrm{sgn}\, Y^{y,u^*}(n)$, $n = 0, 1, 2, ...$ ($\mathrm{sgn}0 = -1$).

Теорема 1. *Пусть $y_1, y_2 \in \mathbb{Z}^1_+$, $y_1 \leqslant y_2$, $y_1 - y_2$ четно. Для любой стратегии $u \in U$ существуют расширение вероятностного пространства $(\Omega, \mathcal{F}, \mathbf{P})$ с новым потоком F', $F \subseteq F'$, и новое простое одномерное случайное блуждание $W'(n)$, $n = 0, 1, 2, ...$, такие, что*

$$|Y^{y_1, u^*}(W', \xi)| \leqslant |Y^{y_2, u}(W, \xi)|.$$

Это неравенство выполнено на каждом элементарном событии. ♦

5. ОДНОМЕРНОЕ УПРАВЛЯЕМОЕ СЛУЧАЙНОЕ БЛУЖДАНИЕ С ОТРАЖЕНИЕМ

Рассмотрим вспомогательную задачу управления случайным блужданием в полуоси $\mathbb{Z}^1_+$.

Пусть W — траектория блуждания в множестве $\mathbb{Z}^1$. Определим

$$\phi(W)(0) = W(0),$$

$$\phi(W)(n) = W(0) + \sum_{i=1}^{i=n} \Delta\phi(W)(n),$$

где

$$\Delta\phi(W)(n) = \begin{cases} \Delta W(n), \text{если } \phi(W)(n-1) \neq 0, \\ +1, \text{если } \phi(W)(n-1) = 0, \end{cases}$$

$n = 1, 2, ...$ Функционал ϕ называется функционалом отражения (в нуле) Скорохода [5].

Вернемся в парадигму управляемого одномерного случайного блуждания. Определим «оптимальную стратегию» u^*_+ как $u^*_+(n) = -1$, $n = 1, 2, ...$ (она оптимальна для управления блужданием в полуоси $\mathbb{Z}^1_+$, поэтому помечена знаком +).

Теорема 2. *Пусть* $y_1, y_2 \in \mathbb{Z}^1_+$, $y_1 \leqslant y_2$, и разность $y_1 - y_2$ *четна. Тогда для любой стратегии* $u \in U$

$$\phi(Y^{y_1, u^*_+})(n) \leqslant \phi(Y^{y_2, u})(n), \quad n = 0, 1, 2, \ldots \blacklozenge$$

Замечание. Стратегия u^* не обладает соответствующим свойством, поэтому нам и пришлось обратиться к процессам с отражением. Действительно, пусть $y = -1$, $u(0) = -1$, $\xi(1) = 1$, $\xi(2) = \xi(3) = 0$, $\Delta W(2) = \Delta W(3) = 1$. Тогда $Y^{y,u^*}(3) = 1$, а $Y^{y,u}(3) = 0$. ♦

6. ПЕРСПЕКТИВЫ ПРИМЕНЕНИЯ ПОЛУЧЕННЫХ РЕЗУЛЬТАТОВ И ДАЛЬНЕЙШИХ ИССЛЕДОВАНИЙ

Пусть имеются k одинаковых государственных унитарных предприятий. Они ведут хозяйственную деятельность и ежемесячно отчитываются суммой прибыли с начала года $X_i(n)$, $n = 0, 1, 2, \ldots, 12$, $X_i(0) = 0$, $i = 1, \ldots, k$, $X_i(n)$ принимает целочисленные значения в тысячах рублей. Государство резервирует ресурс для поддержки предприятий и направляет его для предотвращения их убыточности. Ежемесячно расходуется фиксированная сумма. Математическая модель: без поддержки государства прибыль с начала года X_i — симметричное случайное блуждание в множестве $\mathbb{Z}^1$, $i = 1, \ldots, k$, причем процессы $\{X_i, i = 1, \ldots, k\}$ независимы в совокупности; применение ресурса смещает распределение скачка в положительном направлении $\{1, 2\}$; критерий оптимизации — минимизировать вероятность случая убыточности среди предприятий в течение года, т. е.

$$\mathbf{P}\Big(\min_{\substack{i = 1, \ldots, k \\ n = 1, \ldots, 12}} X_i(n)\Big) \to \max.$$

В случае $k = 2$ и блуждания со скачками размера ± 1 основная теорема дает оптимальную стратегию при дополнительном предположении: ежемесячно выбирается одно из предприятий и поддержка оказывается только ему. А именно, оптимальная стратегия — поддерживать предприятие с наименьшей суммарной прибылью с начала года.

Обобщение основной теоремы на блуждания в множестве $\mathbb{Z}^k$ (размерности k) со скачками не только единичного размера позволит доказать оптимальность сформулированной стратегии в общем случае. Дальнейшее обобщение — гибкое распределение ресурса между предприятиями.

ЗАКЛЮЧЕНИЕ

Для управляемого случайного блуждания в квадранте аналитически найдена стратегия, максимизирующая время выхода на границу и ряд других критериев. Функция выигрыша при этом не находилась. Рассмотренная модель имеет универсальное применение, она уже использовалась при решении задач математической экономики [2].

Автор весьма признателен профессору R. Pemantle (USA), давшему ему принципиальную идею решения задачи, и намеревается опубликовать окончательные результаты в совместной работе. Автор благодарен рецензенту, посоветовавшему написать § 6.

ПРИЛОЖЕНИЕ

Доказательство теоремы 2. Обозначим для краткости $\phi^*(n) = \phi(Y^{y_1, u^*_+})(n)$ и $\phi(n) = \phi(Y^{y_2, u})(n)$, $n = 0, 1, 2, \ldots$ Заметим, что $(\phi^* - \phi)(n)$ сохраняет четность (как функция n). В самом деле, $(\phi^* - \phi)(n + 1) - (\phi^* - \phi)(n) = \Delta\phi(n + 1) - \Delta\phi(n)$, и поскольку слагаемые в правой части принимают значения ± 1, сама правая часть принимает значения $0, \pm 2$, $n = 0, 1, 2, \ldots$ Поэтому $\min\limits_{n = 0, 1, \ldots} (\phi^* - \phi)(n) < 0$ влечет существование такого n, что $(\phi^* - \phi)(n) = 0$, а $(\phi^* - \phi)(n + 1) < 0$. Но это невозможно:

если $\phi(n) = \phi^*(n) = 0$, то $\phi(n + 1) = \phi^*(n + 1) = 1$;

если $\phi(n) = \phi^*(n) > 0$, то $\phi(n + 1) = \phi^*(n + 1)$ в случае $\xi(n + 1) = 0$ и $\phi(n + 1) \geqslant \phi^*(n + 1)$ в случае $\xi(n + 1) = 1$ — это следует из определения u^*_+.

Для доказательства теоремы 1 нужны следующие две леммы и следствие.

Лемма 1. *Для любых* $y \in \mathbb{Z}^1_+$, $u \in U$ процесс $|Y^{y,u}(W, \xi)|$ *можно представить как* $\phi(Y^{y, \tilde{u}}(\tilde{W}, \xi))$, *где*

$$\tilde{u}(k) = \begin{cases} \operatorname{sgn} Y^{y_2, u}(k) u(k), \text{ если } Y^{y_2, u}(k) \neq 0, \\ u(k), \text{ если } Y^{y_2, u}(k) = 0, \end{cases}$$

$k = 0, 1, \ldots$, *а* $\tilde{W}$ *определяется формулами* $\tilde{W}(0) = 0$ *и*

$$\Delta\tilde{W}(k) = \begin{cases} \operatorname{sgn} Y^{y_2, u}(k - 1)\Delta W(k), \text{ если } Y^{y_2, u}(k - 1) \neq 0; \\ \Delta W(k), \text{ если } Y^{y_2, u}(k - 1) = 0, \end{cases}$$

$k = 1, 2, \ldots$ ♦

Доказательство. Представим $|Y^{y_2, u}|$ как отраженный процесс. Очевидно, что

$$\Delta|Y^{y_2, u}|(k) = \begin{cases} \operatorname{sgn} Y^{y_2, u}(k - 1)\Delta Y^{y_2, u}(k), \text{ если } Y^{y_2, u}(k - 1) \neq 0; \\ 1, \text{ если } Y^{y_2, u}(k - 1) = 0, \end{cases}$$

$k = 1, 2, \ldots$ Далее, если $Y^{y_2, u}(k-1) \neq 0$,

$$\operatorname{sgn} Y^{y_2, u}(k-1)\Delta Y^{y_2, u}(k) =$$

$$= \begin{cases} \operatorname{sgn} Y^{y_2, u}(k-1)\Delta W(k), \text{ если } \xi(k) = 0, \\ \operatorname{sgn} Y^{y_2, u}(k-1)u(k-1), \text{ если } \xi(k) = 1, \end{cases}$$

$k = 1, 2, \ldots$

Теперь

$$\Delta|Y^{y_2, u}|(k) = \begin{cases} \Delta\tilde{W}(k), \text{ если } Y^{y_2, u}(k-1) \neq 0, \xi(k) = 0; \\ \tilde{u}(k), \text{ если } Y^{y_2, u}(k-1) \neq 0, \xi(k) = 1; \\ 1, \text{ если } Y^{y_2, u}(k-1) = 0, \end{cases}$$

$k = 1, 2, \ldots$ Очевидно,

$$|Y^{y_2, u}(W, \xi)| = \phi(Y^{y_2, \tilde{u}}(\tilde{W}, \xi)).$$

Лемма доказана. ♦

Следствие. Для любого $y \in \mathbb{Z}^1_+$ распределения $|Y^{y, u^*}(W, \xi)|$ и $\phi(Y^{y, u^*_+}(W, \xi))$ совпадают.

Лемма 2. *Пусть* $y \in \mathbb{Z}^1_+$ и на $(\Omega, \mathcal{F}, F, \mathbf{P})$ *задан процесс* R, *распределенный как* $|Y^{y, u^*}|$. *Тогда существуют расширение исходного вероятностного пространства и определенные на нем поток* σ*-алгебр* F' *и процесс* W' *такие, что* (W', ξ) *является* F'*-ведущим процессом, и* $R = |Y^{y_1, u^*}(W', \xi)|$. ♦

Доказательство. Превратим пространство последовательностей $(s, \Delta)(n) \in \{-1, 1\} \times \{-1, 1\}$, $n = 1, 2, \ldots$, в вероятностное пространство с мерой так, чтобы $(s, \Delta)(n)$, $n = 1, 2, \ldots$, стали независимыми случайными величинами с независимыми одинаково распределенными координатами, причем S (соответственно, Δ) принимает значение 1 с вероятностью p (соответственно, 1/2). Теперь возьмем прямое произведение построенного вероятностного пространства и исходного и положим

$$\mathcal{F}'(0) = \mathcal{F}(0),$$

$$\mathcal{F}'(n) = \sigma(\mathcal{F}(n), (s, \Delta)(i), i = 1, \ldots, n), \quad n = 1, 2, \ldots$$

Определим последовательность марковских моментов (возможно, конечную) $\tau_0 = 0$, $\tau_{i+1} = \min\{n > \tau_i: R(n) = 0\}$, $i = 0, 1, \ldots$ Заметим, что по определению τ_i, $i = 0, 1, \ldots$, строго возрастает. Определим процессы

$$R'(n) = I_{[\tau_0, \tau_1]}(n)R(n) +$$

$$+ \sum_{i=1}^{\infty} I_{(\tau_i, \tau_{i+1}]}(n)R(n)(I_{\{0\}}(\xi_{\tau_i+1})\operatorname{sgn} s(i) + I_{\{1\}}(\xi_{\tau_i+1})),$$

$$W'(n) = \sum_{i=1}^{n} (I_{\{0\}}(\xi(i))\Delta R'(i) + I_{\{1\}}(\xi(i))\Delta(i)),$$

$n = 0, 1, \ldots$ Построенные объекты удовлетворяют всем требованиям леммы. Действительно, после каждого попадания процесса R в 0 мы разыгрываем его экскурсию (отрезок траектории до следующего попадания в 0) для построения экскурсии процесса R' в соответствии со стратегией u^*: если $\xi_{\tau_i+1} = 0$, то экскурсии распределены симметрично (относительно нуля). А если $\xi_{\tau_i+1} = 1$, то процесс R' детерминированно переходит в 1, и его экскурсия совпадает с экскурсией R. Далее, процесс Δ позволяет заменить скачки ведущего случайного блуждания, искаженные управлением (когда имелся «шанс для управления», $\xi(i) = 1$). ♦

Доказательство теоремы 1. Идея состоит в том, чтобы представить $|Y^{y,u}|$ как отраженный процесс и воспользоваться теоремой 2. Согласно лемме 1, $|Y^{y,u}(W, \xi)| = \phi(Y^{y, \tilde{u}}(\tilde{W}, \xi))$. Введем процесс $R = \phi(Y^{y, u^*_+}(\tilde{W}, \xi))$. В силу теоремы 1 $|R| \leqslant |Y^{y,u}|$. Но в силу следствия леммы 1 процесс R распределен как процесс $|Y^{y,u^*}|$. Применим к процессу R лемму 2 и получим утверждение теоремы. ♦

Доказательство основной теоремы. Перейдем к координатам (Y_1, Y_2) (см. § 3). Любой стратегии u соответствует стратегия u_1 управления процессом Y_1 (в те моменты, когда он совершает скачки):

$$u_1(n) = \begin{cases} +1, \text{ если } u(n) = 1, \\ -1, \text{ если } u(n) = 2. \end{cases}$$

Таким образом, заявленной в теореме оптимальной стратегии соответствует стратегия u^* теоремы 1. Применим теорему 1: для любой альтернативной стратегии u можно расширить исходный стохастический базис так, что для процессов $X^{x,u}$ и X^{x,u^*} выполнено: $\tau(X^{x,u}) \leqslant \tau(X^{x,u^*})$, что доказывает теорему. ♦

ЛИТЕРАТУРА

1. *Беллман Р.* Динамическое программирование. — М.: Изд-во иностр. лит. — 1960. — 400 с.
2. *McKean H.P., Shepp L.A.* The Advantages of Capitalism vs. Socialism Depends on the Criterion. Presented April 25—29, 2005 at the Linnik meeting, St. Petersberg, Russia // Записки научных семинаров ПОМИ (СПб.). — 2005. — Т. 328. — С. 160—168.
3. *Бородин А.Н., Салминен П.* Справочник по броуновскому движению. Факты и формулы. — СПб.: Лань. — 2000. — 639 с.
4. *Ватанабэ С., Икэда Н.* Стохастические дифференциальные уравнения и диффузионные процессы. — М.: Наука, 1986. — Гл. 4, § 7.

Статья представлена к публикации членом редколлегии Е.Я. Рубиновичем.

Анулова Светлана Владимировна — канд. физ.-мат. наук, ст. науч. сотрудник, Институт проблем управления им. В.А. Трапезникова РАН, г. Москва,
☎ (495) 334-90-49, ✉ anulovas@ipu.rssi.ru.