Math-Net.Ru

Общероссийский математический портал

Т. А. Белкина, В. И. Ротарь, Об оптимальности по вероятности и почти наверное для процессов со свойством связности. II. Случай непрерывного времени, *Теория вероятн. и ее примен.*, 2005, том 50, выпуск 2, 209–223

DOI: <https://doi.org/10.4213/tvp104>

Использование Общероссийского математического портала Math-Net.Ru подразумевает, что вы прочитали и согласны с пользовательским соглашением

<http://www.mathnet.ru/rus/agreement>

Параметры загрузки:

IP: 107.22.136.117

26 апреля 2023 г., 16:43:59

ТЕОРИЯ ВЕРОЯТНОСТЕЙ И ЕЕ ПРИМЕНЕНИЯ

Основан в январе 1956 г.
Выходит 4 раза в год
Москва • Наука • ТВП

Том 50
Выпуск 2, 2005 год
апрель, май, июнь

© 2005 г.

БЕЛКИНА Т. А.*, РОТАРЬ В. И.**

ОБ ОПТИМАЛЬНОСТИ ПО ВЕРОЯТНОСТИ И ПОЧТИ НАВЕРНОЕ ДЛЯ ПРОЦЕССОВ СО СВОЙСТВОМ СВЯЗНОСТИ. II. СЛУЧАЙ НЕПРЕРЫВНОГО ВРЕМЕНИ¹⁾

Статья является продолжением работы [1], где рассматриваются условия, при которых стратегия, минимизирующая ожидаемое значение целевого функционала, асимптотически оптимальна почти наверное или по вероятности. Первое из указанных свойств означает, что упомянутая стратегия минимизирует значение самого целевого функционала для всех реализаций управляемого процесса из множества, вероятность которого близка к единице при больших горизонтах планирования. Асимптотическая оптимальность по вероятности определяется естественным аналогичным образом. Главное отличие условий, рассматриваемых в данной работе, от полученных ранее состоит в том, что они связаны не со свойствами функции Беллмана, а с возможностью перевода управляемого процесса из одного состояния в другое за время с конечным математическим ожиданием. В первой части рассматривался случай дискретного времени, данная работа посвящена управляемым диффузионным процессам.

Ключевые слова и фразы: управляемые процессы, управляемые диффузионные процессы, оптимальное управление, асимптотическая оптимальность, оптимальность почти наверное.

* Центральный экономико-математический институт РАН, Нахимовский пр-т, 47, 117418 Москва, Россия; e-mail: tbel@cemi.rssi.ru

** Центральный экономико-математический институт РАН; Department of Mathematics and Statistics of San Diego State University, San Diego, CA 92182-7720, USA; e-mail: rotar@sciences.sdsu.edu

¹⁾ Работа выполнена при частичной поддержке Российского фонда фундаментальных исследований (грант № 03-01-00479) и РАН по итогам 6-го конкурса-экспертизы проектов молодых ученых (грант № 105) и Фонда содействия отечественной науке.

1. Введение и результаты для некоторых простых моделей

Данная статья является продолжением [1]. В [1], так же как и в этой работе, для управляемых процессов достаточно общей природы изучаются стратегии, минимизирующие не только математическое ожидание целевого функционала, но и значение самого функционала с вероятностью, близкой к единице при больших горизонтах планирования. В [1] рассматривался случай дискретного времени²⁾, в данной работе исследуется случай непрерывного времени. Подробное введение и исторические замечания были приведены в [1], поэтому мы не будем повторять их здесь.

Главная цель данной работы — установление для случая непрерывного времени условий, аналогичных полученным в [1] и связанных с возможностями сообщаемости состояний. Простейшее предположение может быть качественно описано следующим образом.

Рассмотрим траекторию, начинающуюся с фиксированного начального состояния и порожденную оптимальной в среднем стратегией (average optimal strategy). Выберем другое начальное состояние и рассмотрим стратегии такие, что траектории, порожденные этими стратегиями, начинающиеся из этого начального состояния, пересекают первую траекторию за некоторое (случайное) время. Тогда условие состоит в том, что среди указанных стратегий существует такая, для которой ожидаемое значение времени слияния траекторий ограничено равномерно по всем начальным состояниям.

Более подробные комментарии можно найти в [1].

Общая ситуация рассматривается в разделе 2. В данном разделе рассмотрим некоторые простые следствия из основной теоремы. Доказательства приводятся в разделе 3.

1.1. Оптимальность почти наверное для однородных процессов. В этом вводном пункте опустим некоторые формальные детали постановки задачи, отнеся их в раздел 2.

Рассмотрим процесс $x(t)$ со значениями в $\mathbf{R}^l$, описываемый соотношением

$$dx(t) = b(x(t), u(t)) dt + \sigma(x(t), u(t)) dw(t). \quad (1.1)$$

Здесь $w(\cdot)$ — стандартный d -мерный винеровский процесс, определенный на полном вероятностном пространстве $(\Omega, \mathcal{F}, \mathbf{P})$; символом $u(t)$ обозначается допустимое управление, точнее, u — случайный процесс со значениями в пространстве $\mathbf{R}^m$, адаптированный относительно семейства $\mathcal{F}_t = \sigma\{w(s), 0 \leq s \leq t\}$ и для которого (1.1) имеет смысл; b и

²⁾ В тексте утверждения 2 вместо слов «условия на стратегию φ^n » следует читать «условия на управление z^n ».

σ — векторная и матричная функции соответственно таких размерностей, при которых (1.1) имеет смысл. Положим $x(0) = x$.

Зафиксируем $\mathcal{U}$ — класс допустимых управлений $u(t)$ на $[0, \infty)$. Подчеркнем, что формально класс $\mathcal{U}$ произволен и, вообще говоря, может содержать не все «мыслимые» управления.

На протяжении всей статьи будем предполагать, что для любого рассматриваемого управления существует сильное решение уравнения (1.1), не рассматривая условия, при которых это имеет место; см. также раздел 2. О сильных решениях и соответствующих условиях см., например, [6].

Для каждого $T > 0$ определим (целевой) функционал

$$J_{xT}(u) = \int_0^T q(x(t), u(t)) dt, \quad (1.2)$$

где $q: \mathbf{R}^l \times \mathbf{R}^m \rightarrow \mathbf{R}$.

Предположим — не описывая конкретные условия, при которых это верно, — что для любого $u \in \mathcal{U}$ отображение $J_{xT}(u)$ как функция $(\omega, T) \in \Omega \times \mathbf{R}$ измеримо относительно $\mathcal{F} \times \mathcal{B}$, где $\mathcal{B}$ — борелевская σ -алгебра. Эти условия являются требованиями на $J(\cdot)$ и $\mathcal{U}$.

Стратегия — это отображение $\psi = \psi(x(\cdot), t)$ такое, что случайная величина $\psi(x(\cdot), t)$ измерима относительно $\sigma\{x(s), 0 \leq s \leq t\}$; см. более подробно об этом определении в разделе 2.

Если стратегия $\psi(x(\cdot), t) = \varphi(x(t), t)$, $t \geq 0$, для измеримой функции $\varphi(\cdot, \cdot)$, то ψ — марковская стратегия; если $\varphi(x, t) = \varphi(x)$ для некоторой $\varphi(x)$, то стратегия стационарная. (Вновь см. подробности в разделе 2.)

Управление $u(t) = \psi(x(\cdot), t)$ и соответствующий процесс называются управлением и процессом, порожденными стратегией ψ . Обозначим $U_{\psi x}(t)$ и $X_{\psi x}(t)$ реализации управления и процесса соответственно в момент t , порожденные стратегией ψ при $x(0) = x$ (см. подробности в разделе 2).

Управление $\hat{u}$ называется оптимальным в среднем на интервале $[0, T]$ при начальном состоянии x , если

$$\mathbf{E}J_{xT}(\hat{u}) = \inf_u \mathbf{E}J_{xT}(u),$$

где инфимум берется по всем сужениям на интервале $[0, T]$ управлений u из $\mathcal{U}$.

Для фиксированного T стратегия $\hat{\psi} = \hat{\psi}^T$ называется оптимальной в среднем, если для каждого начального состояния x управление, порожденное этой стратегией, оптимально в среднем на интервале $[0, T]$.

В дальнейшем мы будем предполагать, что для каждого T стратегия $\hat{\psi}^T$ существует. Реализация управляющего процесса u , порожденного $\hat{\psi}^T$, будет обозначаться ${}_T\hat{U}_x = {}_T\hat{U}_x(t)$.

О п р е д е л е н и е 1. Для данного начального состояния x управление $u^* \in \mathcal{U}$ называется асимптотически оптимальным почти наврное, если для любого $u \in \mathcal{U}$

$$\liminf_{T \rightarrow \infty} T^{-1}(J_{xT}(u) - J_{xT}(u^*)) \geq 0 \quad \text{п.н.} \quad (1.3)$$

(Левая часть (1.3) измерима ввиду полноты $(\Omega, \mathcal{F}, \mathbf{P})$ и сформулированного выше предположения измеримости.)

Пусть $|\cdot|$ обозначает евклидову норму.

Утверждение 1. Пусть функции q и σ ограничены и существует константа C со следующим свойством: для любого T и любой пары состояний z и z' существуют допустимая стратегия φ и случайная величина N (возможно, зависящая от T, z и z') такие, что

$$EN \leq C \quad (1.4)$$

и

$$|z - z'|N \leq t \leq T \implies q(X_{\varphi z}(t), U_{\varphi z}(t)) = q({}_T\hat{X}_{z'}(t), {}_T\hat{U}_{z'}(t)) \quad \text{п.н.} \quad (1.5)$$

Пусть также для данного x существует управление $\tilde{u} \in \mathcal{U}$ (возможно, зависящее от x) такое, что

$$T^{-1}(J_{xT}(\tilde{u}) - J_{xT}({}_T\hat{U}_x)) \rightarrow 0 \quad \text{п.н. при } T \rightarrow \infty.$$

Тогда управление $\tilde{u}$ асимптотически оптимально почти наврное для данного x .

Заметим, что основное условие (1.5), в отличие от соответствующего условия в случае дискретного времени в [1], является инфинитезимальным: время перехода должно уменьшаться пропорционально расстоянию между начальными состояниями. Рассмотрим некоторые примеры.

П р и м е р 1. Пусть в уравнении (1.1) $\sigma(x, u) \equiv \sigma$, где σ — некоторая постоянная матрица, а функция $b(x, u)$ такова, что уравнение $b(x(t), u(t)) = c$ однозначно разрешимо относительно u для любых x и c из $\mathbf{R}^l$. Построим такое управление $u(t)$, что для $t \in [0, |z' - z|]$ выполнено равенство $b(x(t), u(t)) = b(\hat{x}(t), \hat{u}(t)) + \text{sign}(z' - z)$. В этом случае, как нетрудно видеть, $dx(t) = d\hat{x}(t) + \text{sign}(z' - z) dt$, следовательно,

$$x(t) = \hat{x}(t) + (\text{sign}(z' - z))t - (z' - z), \quad (1.6)$$

т.е. при $t = |z' - z|$ получаем слияние траекторий x и $\hat{x}$. Тогда определим $u(t)$ для $t \in [0, |z' - z|]$ из уравнения

$$b(\hat{x}(t) + (\text{sign}(z' - z))t - (z' - z), u(t)) = b(\hat{x}(t), \hat{u}(t)) + \text{sign}(z' - z). \quad (1.7)$$

В силу соотношений (1.6), (1.7) и того факта, что оптимальное управление $\hat{u}$ порождается допустимой марковской стратегией, имеем, что и u порождается допустимой марковской стратегией. Следовательно, условие (1.5) выполнено при $N = 1$.

Пример 2 (аналогичный примеру 3 в [1]). Пусть линейная система описывается соотношением

$$dx(t) = (Ax(t) + Bu(t)) dt + \sigma dw(t), \quad (1.8)$$

где A, B, σ — постоянные матрицы подходящих размерностей. Функция цены q — произвольная ограниченная функция. Положим

$$y(t) = {}_T\hat{X}_{z'}(t) - X_{\varphi z}(t), \quad V(t) = {}_T\hat{U}_{z'}(t) - U_{\varphi z}(t), \quad t \in [0, T], \quad (1.9)$$

где φ — стратегия. Ясно, что

$$dy(t) = (Ay(t) + BV(t)) dt, \quad (1.10)$$

где $y(0) = z - z'$. Известно (см., например, [9]), что если пара (A, B) управляема, то система (1.10) может быть переведена из начального состояния в любое конечное состояние за сколь угодно малое время. Точнее, пусть $\Phi(s, \tau) = e^{A(s-\tau)}$, $W(\theta) = \int_0^\theta \Phi(\theta, \tau) BB' \Phi'(\theta, \tau) d\tau$. Тогда управление

$$V_\theta(t) = \begin{cases} -B' \Phi'(\theta, t) W^{-1}(0, \theta) \Phi(\theta, 0)(z - z') & \text{при } t \leq \theta, \\ 0 & \text{при } \theta < t \leq T \end{cases}$$

переводит систему за время θ из состояния $z - z'$ в нулевое состояние. Если θ фиксировано, то $V_\theta(t)$ неслучайно, и, следовательно, $y_\theta(t)$, являющееся решением (1.10) при $V(\cdot) = V_\theta(\cdot)$, также неслучайно.

Пусть $\hat{\varphi}_T(t, x)$ — марковская стратегия, оптимальная в среднем на $[0, T]$. Тогда, как легко видеть из (1.9), марковская стратегия $f_T(t, x) = \hat{\varphi}_T(t, x + y(t)) - V_{|z-z'|}(t)$ переводит систему (1.8) из состояния z в нулевой момент времени в состояние ${}_T\hat{X}_{z'}(|z - z'|)$ в момент $t = |z - z'|$, если $|z - z'| \leq T$. Следовательно, условие (1.5) вновь выполнено при $N = 1$.

1.2. Оптимальность по вероятности.

Определение 2. Для данного начального состояния x управление $\tilde{u} \in \mathcal{U}$ называется асимптотически оптимальным по вероятности, если для любого $\varepsilon > 0$ и любого управления $u \in \mathcal{U}$

$$\mathbf{P}\{T^{-1}(J_{xT}(\tilde{u}) - J_{xT}(u)) > \varepsilon\} \longrightarrow 0 \quad \text{при } T \rightarrow \infty. \quad (1.11)$$

Утверждение 2. Пусть условия утверждения 1 выполнены с заменой (1.4) на

$$\mathbf{P}\{N \geq t\} = o(t^{-1/2}) \quad \text{при } t \rightarrow \infty \text{ равномерно по } z, z' \text{ и } T. \quad (1.12)$$

Пусть также для данного x существует управление $\tilde{u} \in \mathcal{U}$ (возможно, зависящее от x) такое, что

$$\mathbf{P}\{T^{-1}[J_{xT}(\tilde{u}) - J_{xT}(\hat{U}_x)] > \varepsilon\} \rightarrow 0 \quad \text{при } T \rightarrow \infty. \quad (1.13)$$

Тогда управление $\tilde{u}$ асимптотически оптимально по вероятности для данного x .

Таким образом, как и в случае дискретного времени (см. более подробно [1]), для оптимальности по вероятности не требуется конечности математического ожидания времени перехода: из (1.12) следует, что достаточно предположить, что $\mathbf{E}\sqrt{N} \leq a$, где a — константа.

2. Общая схема

Обобщение рассмотренной выше схемы выглядит следующим образом.

1. Будем рассматривать более общие, в частности неоднородные, процессы.

2. Условия типа (1.5) требуют полного слияния двух определенных траекторий. Ниже, как и в случае дискретного времени в [1], будем предполагать, что эти траектории лишь приближаются друг к другу.

3. Класс допустимых управлений ниже может включать не только процессы на $[0, \infty)$, но и семейства управлений (аналог схемы серий в случае дискретного времени).

Рассмотрим случайный процесс $x(t)$ со значениями в $\mathbf{R}^l$, определенный соотношением

$$dx(t) = b(x(t), u(t), t) dt + \sigma(x(t), u(t), t) dw(t). \quad (2.1)$$

Процессы $w(\cdot)$, $u(\cdot)$ и размерности b и σ определяются так же, как и в п. 1.1. (Мы ограничиваемся здесь евклидовым пространством, чтобы упростить некоторые формулировки; ясно, что приведенные ниже результаты могут быть обобщены на случай метрических пространств более общей природы.)

Положим $x(0) = x$.

Целевой функционал имеет вид

$$J_{xT}(u) = \int_0^T g(x(t), u(t), t) dt, \quad (2.2)$$

где функция $g: \mathbf{R}^l \times \mathbf{R}^m \times \mathbf{R} \rightarrow \mathbf{R}$. Чтобы избежать недоразумения, мы обозначаем функцию цены в частном случае в п. 1.1, см. (1.2), символом q (в этом случае она является функцией двух аргументов), а в общем случае (2.2) — символом g (тогда это функция трех аргументов).

Для каждого T обозначим $\mathcal{U}_T$ фиксированный класс допустимых управлений на $[0, T]$. Не исключается, что процесс $u(\cdot)$ из $\mathcal{U}_T$ зависит от T , т.е., вообще говоря, мы имеем дело с семейством управлений $u = \{u_T(t), 0 \leq t \leq T\}$, $0 \leq T < \infty$. То же касается стратегий, которые также могут зависеть от горизонта планирования.

Положим $\mathcal{U} = \{\mathcal{U}_T, 0 \leq T < \infty\}$ и $u = \{u_T(t), 0 \leq t \leq T\}$, $0 \leq T < \infty$, — семейство управлений. Будем использовать тот же термин «управление» для семейств u .

Обоснование такой постановки состоит в основном не в том, что она позволяет расширить класс допустимых управлений и, в частности, включить в него управления, оптимальные в среднем (хотя это желательно), а в том, что в этом случае мы устанавливаем более строгое свойство оптимальности. А именно, мы показываем, что некоторое управление $\tilde{u}$ является наилучшим в интересующем нас смысле не только среди управлений, порожденных управлениями $u(t)$, определенными на бесконечном интервале $[0, \infty)$, но и среди управлений, порожденных семействами управлений. Ясно, что это более строгое утверждение.

С другой стороны, следует заметить, что схема серий включает традиционную постановку, при которой мы ограничиваемся процессами $u(t)$ из класса $\mathcal{U}_\infty$ управлений на $[0, \infty)$ (чтобы свести первую постановку к последней, достаточно предположить, что класс $\mathcal{U}$ — такой, что $u_T(t) = u_{T'}(t)$ для любого $t \in [0, T]$, $T \leq T'$, и $u = \{u_T(t)\} \in \mathcal{U}$).

Таким образом, если читатель интересуется только традиционной постановкой, он может подразумевать классическую схему, игнорируя второй индекс T там, где он возникает, и это не должно привести к неправильному пониманию.

Ниже мы сохраним предположение измеримости $J_{xT}(u_T)$ из п. 1.1 и иногда будем писать $J_{xT}(u)$ вместо $J_{xT}(u_T)$ при $u_T \in \mathcal{U}_T$.

Определение асимптотической оптимальности почти наверное — то же, что и в (1.3), за исключением того, что теперь, вообще говоря, u и u^* являются семействами управлений из $\mathcal{U}$.

Пусть $C([0, \infty); \mathbf{R}^l)$ — пространство непрерывных $\mathbf{R}^l$ -значных функций на $[0, \infty)$. Как обычно, наделим это пространство топологией равномерной сходимости на компактных множествах. Будем рассматривать только допустимые управления с обратной связью, т.е. случайные процессы $u(\cdot)$, прогрессивно измеримые относительно фильтрации, порожденной решениями $x(\cdot)$ уравнения (2.1). Это предполагает существование соответствующей стратегии, т.е. измеримого отображения $\psi: C([0, \infty); \mathbf{R}^l) \times [0, \infty) \rightarrow \mathbf{R}^m$ такого, что $u(t) = \psi(x(\cdot), t)$ п.н. для любого $t \geq 0$ и случайная величина $\psi(x(\cdot), t)$ измерима относительно $\sigma\{x(s), 0 \leq s \leq t\}$. Управление этого типа и соответствующий процесс называются управлением и процессом, порожденными стратегией ψ .

Если стратегия $\psi(x(\cdot), t) = \varphi(x(t), t)$, $t \geq 0$, для некоторого измеримого отображения $\varphi: \mathbf{R}^l \times [0, \infty) \rightarrow \mathbf{R}^m$, то ψ — марковская стратегия. Если $\varphi(x, t) = \varphi(x)$ для $\varphi: \mathbf{R}^l \rightarrow \mathbf{R}^m$, то стратегия стационарна.

Будем всегда предполагать, что коэффициенты b и σ , а также стратегия ψ таковы, что существует единственное сильное решение уравнения (2.1) с заменой $u(t)$ на $\psi(x(\cdot), t)$. Условия, обеспечивающие существование такого решения хорошо известны; см., например, [6, гл. IV, теорема 3.1], [7, теоремы 5.2.5 и 5.2.9]. Мы не приводим эти условия здесь, но предполагаем существование и единственность сильного решения.

То же касается регулярности условных математических ожиданий: мы предполагаем, что она всегда имеет место; см., например, [3].

Предположим, что для каждого T существует стратегия $\hat{\psi}^T$, оптимальная в среднем на $[0, T]$. Для каждого начального состояния x обозначим ${}_T\hat{U}_x = {}_T\hat{U}_x(t)$ реализацию управления, порожденного $\hat{\psi}^T$.

Для фиксированного T рассмотрим функцию Беллмана для задачи минимизации математического ожидания функционала (2.2), т.е. функцию

$$V(t, x) = \inf_{u(\cdot)} \mathbf{E} \left\{ \int_t^T g(x(s), u(s), s) ds \mid x(t) = x \right\},$$

где инфимум берется по всем управлениям $u(\cdot)$, для которых $x(t) = x$.

При $r \in \mathbf{R}^m$ обозначим $\mathcal{A}^r$ оператор такой, что

$$\mathcal{A}^r V(t, x) = \frac{1}{2} \sum_{i,j=1}^l a_{ij}(x, r, t) \frac{\partial^2 V}{\partial x_i \partial x_j}(t, x) + \sum_{i=1}^l b_i(x, r, t) \frac{\partial V}{\partial x_i}(t, x),$$

где матрица $\{a_{ij}\} = \sigma\sigma'$. Будем предполагать в дальнейшем, что функции $b(\cdot, \cdot, \cdot)$, $g(\cdot, \cdot, \cdot)$ таковы, что:

(а) функция Беллмана существует и удовлетворяет уравнению

$$\inf_{r \in \mathbf{R}^m} \left\{ \frac{\partial V}{\partial t}(t, x) + \mathcal{A}^r V(t, x) + g(x, r, t) \right\} = 0, \quad V(T, x) \equiv 0; \quad (2.3)$$

(б) существует оптимальное в среднем управление на $[0, T]$, и оно порождается допустимой марковской стратегией $\hat{\varphi}_T(x, t)$, которая удовлетворяет уравнению

$$\frac{\partial V}{\partial t}(t, x) + \mathcal{A}^{\hat{\varphi}}(x, t) V(t, x) + g(x, \hat{\varphi}_T(x, t), t) = 0. \quad (2.4)$$

Достаточные условия на функции b и g , при которых последнее требование выполняется, также хорошо известны, см., например, [4].

Для фиксированного t , марковской стратегии φ и состояния z рассмотрим управление

$$u_z(s) = \varphi(x(s), s), \quad s \geq t, \quad (2.5)$$

где $x(t)$ — решение (2.1) при $u(s) = \varphi(x(s), s)$, $s \geq t$, с начальным условием $x(t) = z$. Другими словами, (2.5) — управление, порожденное φ , начиная с момента t , при условии, что процесс в этот момент находился в состоянии z .

Будем обозначать такое управление $U_{\varphi z}[s; t]$, а соответствующий процесс — $X_{\varphi z}[s; t]$, $s \geq t$.

В случае, если при этом мы имеем дело со стратегией $\hat{\varphi}$, оптимальной в среднем на $[0, T]$, будем использовать обозначения ${}_T\hat{U}_z[s; t]$, ${}_T\hat{X}_z[s; t]$, опуская индекс $\hat{\varphi}$. Процессы ${}_T\hat{U}_x[s, 0]$, ${}_T\hat{X}_x[s, 0]$ обозначим ${}_T\hat{U}_x(s)$, ${}_T\hat{X}_x(s)$ соответственно.

Теперь мы можем сформулировать основные условия.

У с л о в и е А. Существует константа A со следующим свойством: для любого T , любой пары (неслучайных) состояний z и z' и всех $t \leq T$ существует марковская стратегия φ такая, что

$$\mathbf{E} \left| \int_t^T \left(g(X_{\varphi z}[\tau; t], U_{\varphi z}[\tau; t], \tau) - g({}_T\hat{X}_{z'}[\tau; t], {}_T\hat{U}_{z'}[\tau; t], \tau) \right) d\tau \right| \leq A|z - z'|. \quad (2.6)$$

Так как, вообще говоря, левая часть (2.6) имеет порядок $T - t$, то неравенство (2.6) может быть выполнено, если значения цены $g(\cdot, \cdot, \tau)$ на двух траекториях, $X_{\varphi z}$ и ${}_T\hat{X}_{z'}$, близки для больших τ . Таким образом, условие А — это условие существования стратегии φ , для которой последнее имеет место. Член $|z - z'|$ в правой части (2.6) означает, что накладываемое условие является версией условия Липшица относительно начальных состояний.

Для оптимальности по вероятности достаточно наложить следующее более слабое условие.

У с л о в и е А'. Существует функция $\delta(T) \rightarrow 0$ при $T \rightarrow \infty$ такая, что для любого T , любой пары состояний z, z' и всех $t \leq T$ найдется допустимая марковская стратегия φ такая, что

$$\mathbf{E} \left| \int_t^T \left(g(X_{\varphi z}[\tau; t], U_{\varphi z}[\tau; t], \tau) - g({}_T\hat{X}_{z'}[\tau; t], {}_T\hat{U}_{z'}[\tau; t], \tau) \right) d\tau \right| \leq \delta(T) T^{1/2} |z - z'|. \quad (2.7)$$

Теорема 1. Пусть матричная функция $\sigma(\cdot)$ ограничена.

(1) Если выполнено условие А', то для любого начального состояния x семейство управлений $\{{}_T\hat{U}_x\}$, $0 \leq T < \infty$, асимптотически оптимально по вероятности.

(2) Пусть выполнено условие А и функция g ограничена снизу. Пусть для заданного x и управления $u^* \in \mathcal{U}$

$$\lim_{T \rightarrow \infty} T^{-1} (J_{xT}(u^*) - J_{x[T]}({}_T\hat{U}_x)) = 0 \quad \text{п.н.}, \quad (2.8)$$

где $[T]$ — целая часть T . Тогда u^* асимптотически оптимально п.н. для данного x .

Обсудим сначала (2.8), которое возникает только потому, что мы рассматриваем схему серий. Это условие означает, в частности, что для проверки асимптотической оптимальности п.н. $T\hat{U}_x$ нужно убедиться, что

$$T^{-1}(J_{xT}(T\hat{U}_x) - J_{x[T]}([T]\hat{U}_x)) \rightarrow 0 \quad \text{п.н. при } T \rightarrow \infty. \quad (2.9)$$

Нетрудно проверить, что ограниченности g недостаточно для того, чтобы (2.9) выполнялось автоматически. Условия (2.8) или (2.9) выглядят достаточно слабыми и носят технический характер, но в общем случае мы не можем от них отказаться.

С другой стороны, если существует управление $\tilde{u}(t)$ на $[0, \infty)$, для которого

$$T^{-1}(J_{xT}(\tilde{u}) - J_{xT}(T\hat{U}_x)) \rightarrow 0 \quad \text{п.н. при } T \rightarrow \infty, \quad (2.10)$$

то (2.9) верно по крайней мере для ограниченных q . В этом случае оба управления, $\tilde{u}$ и $T\hat{U}_x$, асимптотически оптимальны п.н. (Для того, чтобы $\tilde{u}$ было оптимальным, достаточно потребовать выполнения (2.10) только для целых T .)

Следовательно, в традиционной постановке с управлениями $u(t)$ на $[0, \infty)$ из класса $\mathcal{U}_\infty$, если (2.10) выполнено, то условие (2.9) излишне.

3. Доказательства

3.1. Доказательство утверждений 1 и 2. Покажем, что эти утверждения являются следствиями теоремы 1. С учетом замечаний после формулировки теоремы 1 нетрудно видеть, что в данном случае условие (2.8) для $u^* = \tilde{u}$ следует из условий утверждения 1.

Так как процесс однородный, для проверки условий А или А' достаточно положить в (2.6) и (2.7) $t = 0$. Пусть N и φ — случайная величина и стратегия соответственно, определенные в утверждениях 1 и 2. Тогда в силу (1.5) для любых начальных состояний z, z'

$$\left| \int_0^T \left(g(X_{\varphi z}[\tau, 0], U_{\varphi z}[\tau, 0], \tau) - g({}_T\hat{X}_{z'}[\tau, 0], {}_T\hat{U}_{z'}[\tau, 0], \tau) \right) d\tau \right| \leq 2NG|z - z'| \quad \text{п.н.,} \quad (3.1)$$

где $G = \sup_{x,a,t} |g(x, a, t)|$.

Следовательно, если $EN \leq C$, то (2.6) выполнено при $A = 2GC$.

Чтобы получить (2.7) из условий утверждения 2, заметим, что если $N > T$, то левая часть (1.5) — пустое множество и, следовательно, (1.5)

выполнено автоматически. Таким образом, условие (1.5) останется верным, если мы заменим N его усечением: $N^{(T)} = N$, если $N \leq T$, и $N^{(T)} = T$, если $N \geq T$. Это означает, что в правой части (3.1) мы можем заменить N на $N^{(T)}$.

В силу (1.12) существует функция $\delta_0(t) \rightarrow 0$, $t \rightarrow \infty$, не зависящая от z, z' и T и такая, что $P\{N > t\} \leq \delta_0(t)/\sqrt{t}$. Следовательно,

$$EN^{(T)} = E\{N; N \leq T\} + TP\{N > T\} \leq E\{N; N \leq T\} + \delta_0(T) \sqrt{T}.$$

Легко видеть, что (1.12) влечет $E\{N; N \leq T\} \leq \delta_1(T) \sqrt{T}$, где функция $\delta_1(T) \rightarrow 0$ и определяется только функцией $\delta_0(T)$.

Таким образом, (1.5) влечет (2.7) при $\delta(T) = 2G(\delta_1(T) + \delta_0(T))$, и для завершения доказательства осталось использовать утверждение (1) теоремы 1 для $u^* = \tilde{u}$ с учетом соотношения (1.13).

3.2. Доказательство теоремы 1. Приведем сначала многомерный аналог теоремы 2.1 из [11]. Для краткости обозначим здесь оптимальное в среднем управление и порожденный им процесс $\hat{u}_T(t)$ и $\hat{x}_T(t)$ соответственно, не конкретизируя начальное состояние.

Лемма 1. Для произвольного управления u и порожденного им процесса $x(t)$

$$J_{xT}(u) - J_{xT}(\hat{u}_T) \geq \int_0^T \left[\frac{\partial V}{\partial x}(t, x(t)) \sigma(x(t), u(t), t) - \frac{\partial V}{\partial x}(t, \hat{x}_T(t)) \sigma(\hat{x}_T(t), \hat{u}_T(t), t) \right] dw(t) \quad (3.2)$$

(где $\frac{\partial V}{\partial x}$ — вектор $\left(\frac{\partial V}{\partial x_1}, \dots, \frac{\partial V}{\partial x_l}\right)$ и $\frac{\partial V}{\partial x} \sigma dw$ — скалярное произведение векторов $\frac{\partial V}{\partial x} \cdot \sigma dw$).

Доказательство аналогично доказательству в работе [11]. В силу (2.1) и формулы Ито

$$\begin{aligned} & V(T, x(T)) - V(0, x) \\ &= \int_0^T \left[\frac{\partial V}{\partial t}(t, x(t)) + \frac{1}{2} \sum_{i,j=1}^l a_{ij}(x(t), u(t), t) \frac{\partial^2 V}{\partial x_i \partial x_j}(t, x(t)) \right. \\ & \quad \left. + \sum_{i=1}^l b_i(x(t), u(t), t) \frac{\partial V}{\partial x_i}(t, x(t)) \right] dt \\ & \quad + \int_0^T \sum_{i=1}^l \frac{\partial V}{\partial x_i}(t, x(t)) \sum_{j=1}^d \sigma_{ij}(x(t), u(t), t) dw_j(t). \end{aligned} \quad (3.3)$$

В силу (2.3) подынтегральное выражение в первом интеграле в (3.3) не меньше, чем $-g(x(t), u(t), t)$. В силу второго соотношения в (2.3) левая

часть в (3.3) равна $-V(0, x)$. Следовательно,

$$-V(0, x) \geq -\int_0^T g(x(t), u(t), t) dt + \int_0^T \frac{\partial V}{\partial x}(t, x(t)) \sigma(x(t), u(t), t) dw(t).$$

Аналогично из (2.4) и (3.3) следует, что

$$\begin{aligned} -V(0, x) &= -\int_0^T g(\hat{x}_T(t), \hat{u}_T(t), t) dt \\ &\quad + \int_0^T \frac{\partial V}{\partial x}(t, \hat{x}_T(t)) \sigma(\hat{x}_T(t), \hat{u}_T(t), t) dw(t). \end{aligned}$$

Из последних двух соотношений следует (3.2).

Доказательство теоремы 1. Пусть u — управление и $x(t)$ — процесс, порожденный u . Используя ограниченность функции σ , (3.2) и неравенство Дуба для локально квадратично интегрируемых мартингалов (см., например, [10]), получим:

$$\begin{aligned} &\mathbf{P}\{J_{xT}(\hat{u}_T) - J_{xT}(u_T) > \varepsilon T\} \\ &\leq \mathbf{P}\left\{\int_0^T \left(\frac{\partial V}{\partial x}(t, \hat{x}_T(t)) \sigma(\hat{x}_T(t), \hat{u}_T(t), t) - \frac{\partial V}{\partial x}(t, x_T(t)) \sigma(x(t), u(t), t)\right) dw(t) > \varepsilon T\right\} \\ &\leq \frac{2l \sup \|\sigma(\cdot)\|^2}{\varepsilon^2 T^2} \mathbf{E}\left\{\int_0^T \left[\left|\frac{\partial V}{\partial x}(t, \hat{x}_T(t))\right|^2 + \left|\frac{\partial V}{\partial x}(t, x_T(t))\right|^2\right] dt\right\}, \quad (3.4) \end{aligned}$$

где $|\cdot|$ — евклидова норма, $\|\cdot\|$ — норма матрицы.

Ниже будем опускать для краткости индекс T в обозначениях управлений и процессов.

Для $j = 1, \dots, l$ положим $\Delta_j = (0, \dots, 0, \Delta, 0, \dots, 0)$, где число Δ стоит на j -м месте.

Для любого фиксированного x , любого $t \in [0, T]$ и произвольной допустимой стратегии φ

$$\begin{aligned} V(t, x) &\leq \mathbf{E} \int_t^T g(X_{\varphi x}[\tau, t], U_{\varphi x}[\tau, t], \tau) d\tau \\ &= V(t, x + \Delta_j) + \mathbf{E} \left(\int_t^T g(X_{\varphi x}[\tau, t], U_{\varphi x}[\tau, t], \tau) d\tau \right) \\ &\quad - \mathbf{E} \int_t^T g(\hat{X}_{x+\Delta_j}[\tau, t], \hat{U}_{x+\Delta_j}[\tau, t], \tau) d\tau. \end{aligned}$$

Следовательно,

$$\begin{aligned} &V(t, x) - V(t, x + \Delta_j) \\ &\leq \left| \mathbf{E} \int_t^T \left\{ g(X_{\varphi x}[\tau, t], U_{\varphi x}[\tau, t], \tau) - g(\hat{X}_{x+\Delta_j}[\tau, t], \hat{U}_{x+\Delta_j}[\tau, t], \tau) \right\} d\tau \right|. \quad (3.5) \end{aligned}$$

Заменяя в (3.5) φ на другую стратегию $\tilde{\varphi}$ и меняя местами x и $x + \Delta_j$, легко получим, что

$$V(t, x) - V(t, x + \Delta_j) \geq \left| \mathbf{E} \int_t^T \left\{ g(X_{\tilde{\varphi}(x+\Delta_j)}[\tau, t], U_{\tilde{\varphi}(x+\Delta_j)}[\tau, t], \tau) - g(\hat{X}_x[\tau, t], \hat{U}_x[\tau, t], \tau) \right\} d\tau \right|. \quad (3.6)$$

Пусть теперь стратегии φ и $\tilde{\varphi}$ таковы, что порожденные ими управления удовлетворяют свойству A' , где роль пар (z, z') играют $(x, x + \Delta_j)$ и $(x + \Delta_j, x)$ соответственно. Тогда из (3.5) и (3.6) следует, что

$$\begin{aligned} & |V(t, x) - V(t, x + \Delta_j)| \\ & \leq \left| \mathbf{E} \int_t^T \left[g(X_{\varphi x}[\tau, t], U_{\varphi x}[\tau, t], \tau) - g(\hat{X}_{x+\Delta_j}[\tau, t], \hat{U}_{x+\Delta_j}[\tau, t], \tau) \right] d\tau \right| \\ & \quad + \left| \mathbf{E} \int_t^T \left[g(X_{\tilde{\varphi}(x+\Delta_j)}[\tau, t], U_{\tilde{\varphi}(x+\Delta_j)}[\tau, t], \tau) - g(\hat{X}_x[\tau, t], \hat{U}_x[\tau, t], \tau) \right] d\tau \right| \\ & \leq 2\delta(T) T^{1/2} |\Delta_j| \end{aligned}$$

для любых x , Δ и j . Таким образом,

$$\left| \frac{\partial V}{\partial x}(t, x) \right|^2 \leq 4l\delta^2(T) T$$

для любого (вектора) x . Следовательно, то же неравенство верно для любого $x = x(t)$ п.н. (Заметим также, что выше $\partial V / \partial x$ — вектор и $|\cdot|$ — евклидова норма.)

Далее, используя (3.4), для любого семейства управлений $\{u_T(t) \in \mathcal{U}, 0 \leq T < \infty\}$ получим

$$\begin{aligned} \mathbf{P}\{J_{xT}(\hat{u}_T) - J_{xT}(u_T) > \varepsilon T\} & \leq 16 l^2 \sup \|\sigma(\cdot)\|^2 (\varepsilon T)^{-2} \int_0^T \delta^2(T) T dt \\ & = 16 l^2 \sup \|\sigma(\cdot)\|^2 \varepsilon^{-2} \delta^2(T) \longrightarrow 0 \quad \text{при } T \rightarrow \infty. \end{aligned}$$

Первое утверждение теоремы 1 доказано. Обратимся ко второму.

В силу (2.8) достаточно показать, что для любого управления $u \in \mathcal{U}$

$$\limsup_{T \rightarrow \infty} T^{-1} [J_{x[T]}(\hat{u}_{[T]}) - J_{xT}(u)] \leq 0 \quad \text{п.н.}$$

С другой стороны,

$$\begin{aligned} & \mathbf{P}\left\{ \sup_{T > s} [T^{-1} (J_{x[T]}(\hat{u}_{[T]}) - J_{xT}(u))] > \varepsilon \right\} \\ & \leq \sum_{n=[s]}^{\infty} \mathbf{P}\left\{ \sup_{n \leq T < n+1} (J_{xn}(\hat{u}_n) - J_{xT}(u)) > \varepsilon n \right\} \\ & \leq \sum_{n=[s]}^{\infty} \mathbf{P}\{J_{xn}(\hat{u}_n) - J_{xn}(u) + a > \varepsilon n\}, \end{aligned}$$

где $a = |\inf_{x,a,t} g(x, a, t)|$.

Отсюда и из неравенства Маркова легко следует, что достаточно показать сходимость ряда

$$\sum_{n=1}^{\infty} \frac{1}{n^p} \mathbf{E} \left(\int_0^n \left[\frac{\partial V}{\partial x}(t, \hat{x}_n(t)) \sigma(\hat{x}_n(t), \hat{u}_n(t), t) - \frac{\partial V}{\partial x}(t, x(t)) \sigma(x(t), u(t), t) \right] dw(t) \right)^p \quad (3.7)$$

для некоторого $p > 2$.

Пусть $\|\cdot\|_p$ — норма в L_p и $\mathbf{1} = (1, \dots, 1)$. Обозначим $C(\cdot)$ произвольную константу, зависящую только от аргументов в скобках $(\cdot)$ и, возможно, различную в различных формулах.

По неравенству Буркхольдера–Ганди (см., например, [10])

$$\begin{aligned} & \mathbf{E} \left(\int_0^n \left[\frac{\partial V}{\partial x}(t, \hat{x}_n(t)) \sigma(\hat{x}_n(t), \hat{u}_n(t), t) - \frac{\partial V}{\partial x}(t, x(t)) \sigma(x(t), u(t), t) \right] dw(t) \right)^p \\ & \leq C(p) \mathbf{E} \left(\int_0^n \left(\frac{\partial V}{\partial x}(t, \hat{x}_n(t)) \sigma(\hat{x}_n(t), \hat{u}_n(t), t) \mathbf{1} - \frac{\partial V}{\partial x}(t, x(t)) \sigma(x(t), u(t), t) \mathbf{1} \right)^2 dt \right)^{p/2} \\ & \leq 2 C(p) l \sup \|\sigma(\cdot)\|^p \left\| \int_0^n \left(\left| \frac{\partial V}{\partial x}(t, \hat{x}_n(t)) \right|^2 + \left| \frac{\partial V}{\partial x}(t, x(t)) \right|^2 \right) dt \right\|_{p/2}^{p/2}, \quad (3.8) \end{aligned}$$

где $C(p)$ — константа, зависящая только от p , и для случайного вектора X символ $\|X\|_p$ обозначает $(\mathbf{E}|X|^p)^{1/p}$, где $|\cdot|$ — евклидова норма.

Кроме того, проводя рассуждения, аналогичные использованным в первой части доказательства, легко показать, что при условии (2.6)

$$\left| \frac{\partial V}{\partial x}(t, x) \right|^2 \leq 4lA^2,$$

и, следовательно,

$$\begin{aligned} & \mathbf{E} \left(\int_0^n \left[\frac{\partial V}{\partial x}(t, \hat{x}_n(t)) \sigma(\hat{x}_n(t), \hat{u}_n(t), t) - \frac{\partial V}{\partial x}(t, x(t)) \sigma(x(t), u(t), t) \right] dw(t) \right)^p \\ & \leq C(p, l, \sup \|\sigma(\cdot)\|) A^p n^{p/2}. \end{aligned}$$

Таким образом, для $p > 2$ ряд (3.7) сходится. Теорема 1 доказана.

СПИСОК ЛИТЕРАТУРЫ

1. Белкина Т. А., Ротарь В. М. Об оптимальности по вероятности и почти наверное для процессов со свойством связности. I. Случай дискретного времени. — Теория вероятн. и ее примен., 2005, т. 50, в. 1, с. 3–26.
2. Bhattacharia R. N. Asymptotic behavior of several dimensional diffusions. — Stochastic Nonlinear Systems in Physics, Chemistry, and Biology. Ed. by L. Arnold and R. Lefever. Berlin–New York: Springer-Verlag, 1981, p. 86–99.
3. Евстигнеев И. В. Регулярные условные математические ожидания случайных величин, зависящих от параметров. — Теория вероятн. и ее примен., 1986, т. 31, в. 3, с. 586–589.
4. Флеминг У., Ришел Р. Оптимальное управление детерминированными и стохастическими системами. М.: Мир, 1978, 316 с.
5. Гузман И. И., Скороход А. В. Управляемые случайные процессы. Киев: Наукова думка, 1977, 251 с.
6. Ватанабэ С., Икэда Н. Стохастические дифференциальные уравнения и диффузионные процессы. М.: Наука, 1986, 445 с.
7. Karatzas I., Shreve S. I. Brownian Motion and Stochastic Calculus. New York: Springer-Verlag, 1991, 470 p.
8. Kertz R. P. Renewal plans and persistent optimality in countability additive gambling. — Math. Oper. Res., 1982, v. 7, № 3, p. 361–382.
9. Кваскернаак Х., Сиван Р. Линейные оптимальные системы управления. М.: Мир, 1977, 625 с.
10. Липцер Р. Ш., Ширяев А. Н. Теория мартингалов. М.: Наука, 1986, 512 с.
11. Presman E., Rotar' V., Taksar M. Optimality in probability and almost surely. The general scheme and a linear regulator problem. — Stochastics Stochastics Rep., 1993, v. 43, № 3–4, p. 127–137.

Поступила в редакцию
8.IV.2003