

ОПТИМАЛЬНОЕ  
УПРАВЛЕНИЕ

УДК 519.857, 519.21

ЗАДАЧА УПРАВЛЕНИЯ ЛИНЕЙНЫМ ВЫХОДОМ НЕЛИНЕЙНОЙ  
НЕУПРАВЛЯЕМОЙ СТОХАСТИЧЕСКОЙ ДИФФЕРЕНЦИАЛЬНОЙ  
СИСТЕМЫ ПО КВАДРАТИЧНОМУ КРИТЕРИЮ<sup>1</sup>

© 2021 г. А. В. Босов

МАИ (национальный исследовательский университет), Москва, Россия

e-mail: AVBosov@mail.ru

Поступила в редакцию 17.11.2020 г.

После доработки 14.02.2021 г.

Принята к публикации 29.03.2021 г.

Решена задача оптимального управления по квадратичному критерию качества общего вида для неуправляемого диффузионного процесса Ито и линейного управляемого выхода. Управление с обратной связью в предположении наличия полной информации ищется методом динамического программирования. Функция Беллмана представлена в виде квадратичной формы относительно переменной выхода с коэффициентами, нелинейными относительно переменной состояния. Оптимальное управление формируется двумя слагаемыми: линейной функцией выхода и нелинейной функцией состояния. Коэффициенты, не зависящие от состояния, определены матричным уравнением Риккати, коэффициенты-функции состояния — системой уравнений параболического типа. Изучение свойств параболических уравнений приводит к стохастическому представлению описываемых ими коэффициентов в форме условного математического ожидания терминального состояния некоторой вспомогательной стохастической системы. Это представление обеспечивается формулой Фейнмана–Каца. Показано, что в частном случае линейного сноса оптимальное управление является линейным и по состоянию, а коэффициенты определяются системой обыкновенных дифференциальных уравнений. Этот случай иллюстрируется моделью Кокса–Ингерсона–Росса и численным примером для показателя времени отклика сетевого протокола TCP. В рамках численного эксперимента выполнен сравнительный анализ приближенных решений, полученных традиционным методом конечных разностей и альтернативным методом на основе стохастического представления Фейнмана–Каца.

DOI: 10.31857/S000233882104003X

**Введение.** Рассматриваемая в работе задача может быть отнесена к традиционной задаче оптимального управления дифференциальными стохастическими системами. Задача решается классическими методами, равно как и основанными на них приближенными методами поиска оптимальных управлений [1–3]. Ключевые результаты, составляющие классическую теорию управления, получены более 40 лет назад. Наибольшую полноту и результативность имеет задача управления линейно-гауссовскими стохастическими системами по квадратичному критерию (LQG) [4]. Эта задача дает исходный базис, на котором основано множество успешно исследованных и решенных задач стохастического управления и оптимизации. Множество результатов в этой области как в части моделей, так и в части математического аппарата, имевших место в последующие годы, существенно обогатили теорию управления. Но и до настоящего времени линейные модели и квадратичный критерий, несмотря на всю справедливую критику в отношении их адекватности и гибкости, сохраняют исследовательский интерес и находят современные области приложения.

Не претендуя на сколько-нибудь полное обоснование последнего тезиса, приведем несколько примеров, показавшихся наиболее интересными. Так, в [5] решается линейно-квадратичная задача в игровой постановке с запаздыванием. В близкой по модели работе [6] задача управления ставится в терминах  $H_\infty$ -робастности. Точнее называть эту тематику  $H_2/H_\infty$ -управлением, и работ по этой теме очень много. Аккуратности ради следует уточнить, что под линейными в этой

<sup>1</sup> Работа выполнена при частичной финансовой поддержке РФФИ (проект № 19-07-00187-А).

теории понимаются модели с мультипликативными по состоянию возмущениями. Совсем другой класс моделей, особо популярных в последние годы, составляют скачкообразные процессы. Например, линейные уравнения в сочетании с пуассоновскими скачками в [7] используются в моделях, описывающих различные показатели функционирования сетевых протоколов передачи данных транспортного уровня. Надо заметить, что телекоммуникации представляют в последние годы популярный прикладной материал для исследований, работ по этой проблематике множество, математические техники привлекаются самые разные и самые современные, но и линейным моделям место находится. В этой сравнительно новой прикладной области имеются приложения и для методов стохастического оптимального управления [8].

Близкой к области телекоммуникаций являются задачи управления ресурсами программных систем, в качестве которых могут выступать сайты, базы данных и банки знаний, а также традиционные вычислительные ресурсы, такие, как узлы распределенных систем, компоненты центров обработки данных и т.д. Такие ресурсы управляются программами – менеджерами задач, оптимизаторами запросов в системах управления базами данных, системами балансировки нагрузки. Моделями для постановки задач оптимизации функционирования программных систем являются системы массового обслуживания или, при значительном росте числа состояний, их диффузионные аппроксимации [9]. Соответствующая практическая интерпретация в качестве состояния оптимизируемой системы может использовать число пользователей, запросов, буферов памяти, процессорного времени и т.п., описывая его диффузионным процессом. Объем ресурсов при этом – это линейный выход, учитывающий и состояние, и активное действие программной системы, такое, как объем выделенных или освобожденных ресурсов. Отметим, что такая прикладная интерпретация изучаемой задачи ранее была представлена для случая дискретного времени [10, 11].

Исследуемая в статье модель следует традиционному представлению управляемой системы в форме вход-выход. Входная переменная предполагается неуправляемой, оптимизируется выход. Такого рода задачи известны в числе многих исследуемых постановок задач оптимального управления под терминами управление (регулирование) выходом или управление по выходу. Одно из первых применений этой терминологии обнаруживается в [12], где специальным образом записанная классическая LQG-задача используется для математической формализации инженерной задачи “доведения выхода до нуля”.

В качестве примера более современных постановок в терминологии управления выходом можно привести работу [13] (также указан ряд релевантных ссылок аналогичной тематики). Неопределенность там не имеет стохастической природы, а описывается простыми геометрическими ограничениями, что указывает на робастный характер применяемого метода инвариантных эллипсоидов. Изучаемая система при этом является линейной и автономной. Аналогичную модель исследуют авторы [14] с помощью аппарата линейных матричных неравенств. В указанной и других работах оптимизируется квадратичный критерий и используются свойства линейных регуляторов. Исследований, посвященных нелинейным моделям систем с управляемым выходом, обнаруживается немного. Так, довольно популярны постановки в терминах нелинейного адаптивного управления, предлагающие различные варианты задачи стабилизации [15].

В данной работе, как и во многих других, эксплуатируются привлекательные свойства линейных моделей и квадратичного критерия. Обобщение затронуло модель динамики системы. Оптимизируемая динамическая система описывается двумя системами уравнений. Состояние задается системой нелинейных стохастических дифференциальных уравнений Ито, не содержащих управляемой переменной. Возмущения здесь описываются стандартным векторным винеровским процессом, накладываются простые условия существования и единственности решения. Поскольку состояние не управляется, то возможно его интерпретировать еще и как сложное внешнее возмущение. Вторая переменная – управляемый выход – задается системой линейных стохастических дифференциальных уравнений. Цель оптимизации выхода формируется квадратичным критерием общего вида. Формальная постановка задачи дана в разд. 2. Для решения задачи применяется метод динамического программирования, решается уравнение Беллмана [3], используемые общие положения приведены в разд. 1. Соответственно в результате получаются аналитические выражения и для оптимального управления, и для значения функционала качества. Технически традиционный, стандартный подход к задаче обременен, пожалуй, единственной проблемой – поиском верного представления структуры функции Беллмана. Справиться с этой проблемой в большей степени удастся за счет результата, полученного при решении дискретного по времени аналога рассматриваемой постановки [11]. Конечные соотношения для оптимального решения, как и во всех подобных задачах, включая классическую LQG,

содержат решения определенных дифференциальных уравнений (обыкновенных и в частных производных). Вывод этих уравнений представлен в разд. 3. Решение дает функция Беллмана вида  $V(t, y, z) = z^T \alpha(t)z + z^T \beta(t, y) + \gamma(t, y)$ , в котором переменная  $y$  соответствует вектору состояния, переменная  $z$  – выходу, а коэффициенты  $\alpha(t)$ ,  $\beta(t, y)$ ,  $\gamma(t, y)$  представлены решениями полученных дифференциальных уравнений (обыкновенного Риккати и системы линейных уравнений в частных производных параболического типа). При определении функции Беллмана получено и выражение для оптимального управления в виде линейной функции  $z$  и коэффициента  $\beta(t, y)$ . Соответственно естественная численная реализация оптимального решения – это приближенное решение методом конечных разностей соответствующих дифференциальных уравнений. При этом применение сеточных методов для параболических уравнений осложняется тем обстоятельством, что имеются только начальные условия для уравнений, определяющих  $\beta(t, y)$ ,  $\gamma(t, y)$ , оснований для указания граничных условий, нужных для метода конечных разностей, нет. Из-за этого граничные условия в модельном эксперименте, которому посвящен раздел 6, выбираются довольно волюнтаристски, обсуждается экспериментальный анализ их влияния на решение. И хотя удастся получить удовлетворительный результат, поиск альтернативного метода для приближенного вычисления коэффициентов функции Беллмана сохраняет актуальность.

Поскольку обсуждаемые коэффициенты описываются линейными уравнениями в частных производных второго порядка параболического типа, то перспективным представлялось опираться в исследовании этой задачи на известные факты, ассоциирующие такие уравнения с уравнениями стохастическими. Видимо самая известная ассоциация такого рода – это уравнение А.Н. Колмогорова [16], которое и было использовано в качестве инструмента для реализации альтернативного метода приближенного вычисления коэффициентов  $\beta(t, y)$ ,  $\gamma(t, y)$ . Точнее применялось утверждение, которое определяет связь между решением задачи Коши для уравнения А.Н. Колмогорова, с одной стороны, и теоретико-вероятностным представлением ее терминального условия. Использованное в работе интегральное представление этого результата известно как формула Фейнмана–Каца [17]. Такой метод для численного решения параболических уравнений и уравнений даже более общей структуры успешно реализован и исследован [18], так что примененный в статье подход можно считать апробированным. Обсуждению этого результата посвящен раздел 4. Соответственно стохастическая интерпретация  $\beta(t, y)$  дает возможность определения этого коэффициента как статистической оценки математического ожидания терминальной переменной вспомогательной динамической системы, формируемой из интегрального представления Фейнмана–Каца. Таким образом, альтернативный метод аппроксимации оптимального решения рассматриваемой задачи сводится к использованию компьютерного моделирования и метода Монте-Карло. Результаты применения этого метода и сравнение с методом конечных разностей приведены в разд. 6, представляющем численный эксперимент.

Практически важный частный случай рассматриваемой задачи, случай линейного сноса, но нелинейной диффузии, обычно характеризуемый как модель с мультипликативными возмущениями, рассмотрен в разд. 5. Оптимальное решение в этом случае не требует интегрирования параболического уравнения, управление линейно по состоянию и выходу, а численная реализация сводится к решению только обыкновенных дифференциальных уравнений. Это дает возможность, во-первых, получить подходящий эталонный пример для сравнительного анализа двух других методов аппроксимации оптимального решения – сеточного и имитационного. Во-вторых, позволяет в полном объеме рассмотреть модельный пример, связанный с выбранной целевой прикладной областью телекоммуникаций. Именно в разд. 6 обсуждаются результаты численного эксперимента, выполненного предложенными методами для модели Кокса–Ингерсона–Росса, описывающей показатель времени сетевого отклика RTT (round-trip time) основного сетевого протокола Интернета TCP (transmission control protocol).

**1. Общий подход к решению.** Предваряя основной результат, кратко опишем положения, привлекаемые далее к решению задачи, следуя обозначениям и терминологии [3]. Будем рассматривать задачу оптимального управления состоянием  $x(t) \in \mathbb{R}^n$  стохастической динамической системы по полной информации, применяя метод динамического программирования. Допустимыми будем считать управления с обратной связью, т.е. функции вида  $u = u(t, x) \in \mathbb{R}^k$ ,  $x \in \mathbb{R}^n$ , множество допустимых управлений обозначается  $U_0^T = \{u(t, x), 0 \leq t \leq T, x \in \mathbb{R}^n\}$ , такие, что реализациями управления  $u(t) = u(t, x(t))$  обеспечиваются выполненными условия существования  $x(t)$  для  $u \in U_0^T$ . Заметим, что использование обозначений  $u = u(t, x)$  для управления с обратной

связью и  $u(t) = u(t, x(t))$  для его реализации, соответствующей траектории  $x(t)$ , далее будет везде понятно из контекста и не приведет к неоднозначности.

Минимизируемый целевой функционал имеет вид

$$J(U_0^T) = \mathbb{E} \left\{ \int_0^T L(t, x(t), u(t)) dt + l(x(T)) \right\}, \quad (1.1)$$

где  $\mathbb{E}\{\cdot\}$  – оператор математического ожидания,  $x(t)$  – случайный процесс, описываемый системой стохастических дифференциальных уравнений Ито:

$$dx(t) = m(t, x(t), u(t)) dt + \lambda(t, x(t)) dW(t), \quad 0 \leq t \leq T, \quad x(0) = X, \quad (1.2)$$

где  $W(t) \in \mathbb{R}^m$  – стандартный векторный винеровский процесс,  $X \in \mathbb{R}^n$  – случайный вектор с конечным вторым моментом,  $m(t, x, u), \lambda(t, x)$  – векторная и матричная функции соответствующих размерностей.

Относительно функций сноса  $m(t, x, u)$  и диффузии  $\lambda(t, x)$  в (1.2) будем предполагать выполненными какие-либо условия существования и единственности решения для заданного допустимого управления. Так, учитывая, что допустимыми являются управления с обратной связью, будем считать, что  $m(t, x, u(t, x))$  и  $\lambda(t, x)$  удовлетворяют условию линейного роста по  $x$  и локальному условию Липшица по  $x$  равномерно по  $t$  (т.е. условиям Ито). Также относительно функций из целевого функционала  $L(t, x, u)$  и  $l(x)$  предполагаются выполненными типовые условия существования, например ограничения полиномиального роста по  $x$  и  $u$ .

Для поиска оптимального управления, минимизирующего  $J(U_0^T)$ , рассматривается функция Беллмана

$$V(t, x) = \inf_{U_t^T} \mathbb{E} \left\{ \int_t^T L(\tau, x(\tau), u(\tau)) d\tau + l(x(T)) \middle| \mathcal{F}_t^x \right\}, \quad (1.3)$$

где  $\mathcal{F}_t^x$  –  $\sigma$ -алгебра, порожденная  $x(\tau), 0 \leq \tau \leq t$ ;  $\mathbb{E}\{\cdot | \mathcal{F}_t^x\}$  – оператор условного математического ожидания относительно  $\mathcal{F}_t^x$ ;  $V(t, x)$  предполагается непрерывно дифференцируемой по  $t$  и дважды непрерывно дифференцируемой по  $x$ . В качестве достаточного условия оптимальности рассматривается уравнение динамического программирования (используется представление из [3], записанное в векторной форме):

$$\frac{\partial V}{\partial t} + 1/2 \text{tr} \left\{ \lambda^\top \frac{\partial^2 V}{\partial x^2} \lambda \right\} + \min_u \left\{ m^\top \frac{\partial V}{\partial x} + L \right\} = 0, \quad V(T, x) = l(x), \quad (1.4)$$

где  $\top$  – операция транспонирования,  $\text{tr}$  – след матрицы,  $\partial^2 V / \partial x^2$  – матрица Гессе  $V$ ,  $\partial V / \partial x$  – градиент  $V$ ,  $m = m(t, x, u), \lambda = \lambda(t, x), V = V(t, x)$ . Далее также будут опускаться аргументы функций, если это возможно сделать однозначно.

Традиционно в рамках применения метода динамического программирования будем предполагать, что функции  $L, l, m, \lambda$  обеспечивают существование хотя бы одного решения уравнения (1.4), а следовательно, и оптимального управления

$$(U^*)_0^T = \{u^*(t, x), 0 \leq t \leq T, x \in \mathbb{R}^n\}, \quad u^* = \underset{u}{\operatorname{argmin}} \left\{ m^\top \frac{\partial V}{\partial x} + L \right\},$$

и доставляющего минимум целевому функционалу (1.1). Задача оптимизации далее решается путем указания конкретных выражений для  $L, l, m, \lambda$ .

**2. Постановка задачи управления выходом.** Рассмотрим стохастическую дифференциальную систему, состояние которой представляет диффузионный процесс  $y(t) \in \mathbb{R}^{n_y}$ , описываемый системой нелинейных стохастических дифференциальных уравнений Ито

$$dy(t) = A(t, y(t)) dt + \Sigma(t, y(t)) dv(t), \quad y(0) = Y, \quad (2.1)$$

где  $v(t) \in \mathbb{R}^{n_v}$  – стандартный векторный винеровский процесс,  $Y \in \mathbb{R}^{n_y}$  – случайная величина с конечным вторым моментом, векторная функция  $A(t, y) : \mathbb{R}^{n_y} \rightarrow \mathbb{R}^{n_y}$  и матричная функция  $\Sigma(t, y) : \mathbb{R}^{n_y} \rightarrow \mathbb{R}^{n_y \times n_y}$  удовлетворяют условиям Ито:

$$|A(t, y)| + |\Sigma(t, y)| \leq C(1 + |y|), \quad 0 \leq t \leq T,$$

$$|A(t, y_1) - A(t, y_2)| + |\Sigma(t, y_1) - \Sigma(t, y_2)| \leq C|y_1 - y_2|, \quad 0 \leq t \leq T, \quad y_1, y_2 \in \mathbb{R}^{n_y},$$

обеспечивающим существование единственного потраекторного решения уравнения [3] (здесь и далее  $|\cdot|$  обозначает евклидову норму вектора или матрицы).

Будем считать, что  $y(t)$  описывает состояние некоторой динамической системы. Под выходом этой системы понимается процесс  $z(t) \in \mathbb{R}^{n_z}$ , линейно связанный с состоянием:

$$dz(t) = a(t)y(t)dt + b(t)z(t)dt + c(t)u(t)dt + \sigma(t)dw(t), \quad z(0) = Z, \quad (2.2)$$

где  $w(t) \in \mathbb{R}^{n_w}$  – не зависящий от  $v(t), Y, Z$  стандартный векторный винеровский процесс;  $Z \in \mathbb{R}^{n_z}$  – случайная величина с конечным вторым моментом, не зависящая от  $v(t), Y$ ;  $u(t) \in \mathbb{R}^{n_u}$  – управление. Задача формулируется в предположении наличия полной информации о состоянии  $y(t)$  и выходе  $z(t)$  (соответствующую  $\sigma$ -алгебру обозначим  $\mathcal{F}_t^{y,z}$ ) и управление  $u(t) \mathcal{F}_t^{y,z}$  – измеримо. Класс  $U_0^T$  допустимых управлений составляют управления с полной обратной связью, т.е. функции вида  $u = u(t, y, z) \in \mathbb{R}^{n_u}$ ,  $y \in \mathbb{R}^{n_y}$ ,  $z \in \mathbb{R}^{n_z}$ , в предположении, что соответствующая реализация  $u(t) = u(t, y(t), z(t))$  обеспечивает выполнение условий существования  $y(t), z(t)$  для  $u \in U_0^T$ .

Качество управления  $U_0^T$  определяется целевым функционалом следующего вида:

$$J(U_0^T) = \mathbb{E} \left\{ \int_0^T \|P(t)y(t) + Q(t)z(t) + R(t)u(t)\|_{S(t)}^2 dt + \|P(T)y(T) + Q(T)z(T)\|_{S(T)}^2 \right\}, \quad (2.3)$$

где  $P(t) \in \mathbb{R}^{n_y \times n_y}$ ,  $Q(t) \in \mathbb{R}^{n_z \times n_z}$ ,  $R(t) \in \mathbb{R}^{n_y \times n_u}$ ,  $S(t) \in \mathbb{R}^{n_y \times n_y}$ ,  $S(t) \geq 0$ ,  $S(t) = S^T(t)$ ,  $0 \leq t \leq T$ , – заданные ограниченные матричные функции, весовая функция  $\|x\|_S^2 = x^T S x$  для симметричной неотрицательно определенной матрицы  $S$ , единичной матрице  $S = E$  соответствует евклидова норма  $\|x\|_E^2 = |x|^2$ . Для исключения возможности отсутствия штрафа для отдельных компонентов вектора управления  $u(t)$ , которая делает целевой функционал (2.3) физически некорректным, предполагается выполненным обычное условие невырожденности, в данных обозначениях принимающее вид  $R(t)^T S(t) R(t) > 0$ .

Матричные функции  $a(t) \in \mathbb{R}^{n_z \times n_y}$ ,  $b(t) \in \mathbb{R}^{n_z \times n_z}$ ,  $c(t) \in \mathbb{R}^{n_z \times n_u}$ ,  $\sigma(t) \in \mathbb{R}^{n_z \times n_w}$  ограничены:  $|a(t)| + |b(t)| + |c(t)| + |\sigma(t)| \leq C$  для всех  $0 \leq t \leq T$ . Таким образом, обеспечивается существование решения уравнения (2.2) для любого допустимого управления  $u$ . Более того, будем предполагать, что все используемые функции времени  $a(t), b(t), c(t), \sigma(t), P(t), Q(t), R(t), S(t)$  кусочно-непрерывны, чтобы обеспечить выполнение типовых условий существования решений дифференциальных уравнений, получаемых далее.

Задачу составляет поиск  $(U^*)_0^T = \{u^*(t, y, z), 0 \leq t \leq T, y \in \mathbb{R}^{n_y}, z \in \mathbb{R}^{n_z}\}$  – допустимого управления с обратной связью, доставляющего минимум квадратичному функционалу  $J(U_0^T)$ .

Функционал (2.3) формализует типичную задачу управления выходом. В [12], как уже упоминалось, в качестве примера практического приложения такого типа целевой функции рассматривается задача “доведения выхода до нуля”. Предлагаемый более общий вид целевого функционала отражает практическое содержание упомянутой выше задачи распределения ресурсов. Так, частный вид функционала (2.3)

$$J(U_0^T) = \mathbb{E} \left\{ \int_0^T (|y(t) - z(t)|^2 + Q(t)|z(t)|^2 + R(t)|u(t)|^2) dt + Q(T)|z(T)|^2 \right\}$$

позволяет ставить задачи слежения выходной переменной за состоянием, учитывая расходы на управляющее воздействие и/или значение выходной переменной, функционал

$$J(U_0^T) = \mathbb{E} \left\{ \int_0^T (|z(t) - u(t)|^2 + R(t)|u(t)|^2) dt \right\}$$

позволяет ставить задачи слежения управлением за выходной переменной, учитывая расходы на управляющее воздействие.

Поставленная задача очевидным образом формулируется в терминах выражений (1.1)–(1.3), для этого нужно обозначить

$$x(t) = (y^T(t), z^T(t))^T, \quad m(t, x(t), u(t)) = (A^T(t, y(t)), (a(t)y(t) + b(t)z(t) + c(t)u(t))^T)^T,$$

$$\lambda(t, x(t)) = \begin{pmatrix} \Sigma(t, y(t)) & 0_{n_y \times n_w} \\ 0_{n_z \times n_v} & \sigma(t) \end{pmatrix}, \quad W(t) = (v^T(t), w^T(t))^T$$

для записи уравнения состояния в виде (1.2) и

$$L(t, x(t), u(t)) = L(t, y(t), z(t), u(t)) = \|P(t)y(t) + Q(t)z(t) + R(t)u(t)\|_{S(t)}^2,$$

$$l(x(T)) = l(y(T), z(T)) = \|P(T)y(T) + Q(T)z(T)\|_{S(T)}^2$$

для записи целевого функционала в виде (1.1). В записи блочной матрицы выше через  $0_{n \times m}$  обозначена нулевая матрица размерности  $n \times m$ .

Функция Беллмана (1.3) принимает вид  $V = V(t, y, z)$ , а уравнение (1.4) с обозначениями  $\Sigma = \Sigma(t, y)$ ,  $A = A(t, y)$ ,  $a = a(t)$ ,  $b = b(t)$ ,  $c = c(t)$ ,  $u = u(t)$ ,  $\sigma = \sigma(t)$ ,  $P = P(t)$ ,  $Q = Q(t)$ ,  $R = R(t)$ ,  $S = S(t)$ , записывается как

$$\frac{\partial V}{\partial t} + 1/2 \text{tr} \left\{ \Sigma^T \frac{\partial^2 V}{\partial y^2} \Sigma + \sigma^T \frac{\partial^2 V}{\partial z^2} \sigma \right\} +$$

$$+ \min_u \left\{ A^T \frac{\partial V}{\partial y} + (ay + bz + cu)^T \frac{\partial V}{\partial z} + \|Py + Qz + Ru\|_S^2 \right\} = 0, \quad (2.4)$$

$$V(T, y, z) = \|P(T)y + Q(T)z\|_{S(T)}^2.$$

Существование решения уравнения (2.4) является достаточным условием оптимальности, оптимальное управление при этом получается при минимизации соответствующего слагаемого в фигурных скобках.

**3. Динамическое программирование и оптимальное управление.** В рассматриваемой постановке линейность выхода и квадратичность критерия дают те же преимущества, что и в классической LQG-задаче [4], а именно позволяют сразу определить вид оптимального управления и фактические условия его существования. Действительно, сохраняя в (2.4) под знаком  $\min_u$  только члены, зависящие от  $u$ , получаем

$$\frac{\partial V}{\partial t} + 1/2 \text{tr} \left\{ \Sigma^T \frac{\partial^2 V}{\partial y^2} \Sigma + \sigma^T \frac{\partial^2 V}{\partial z^2} \sigma \right\} +$$

$$+ A^T \frac{\partial V}{\partial y} + (ay + bz)^T \frac{\partial V}{\partial z} + \|Py + Qz\|_S^2 +$$

$$+ \min_u \left\{ u^T \left( c^T \frac{\partial V}{\partial z} + 2R^T S (Py + Qz) \right) + \|Ru\|_S^2 \right\} = 0. \quad (3.1)$$

Далее из (3.1) с учетом сделанного предположения  $R^T S R > 0$ , обозначения “ $A^{-1}$ ” для обратной матрицы  $A$  и выражения для минимума квадратичной формы получаем равенство для оптимального управления в предположении существования решения (3.1):

$$u^* = u^*(t, y, z) = -1/2(R^T S R)^{-1} \left( c^T \frac{\partial V}{\partial z} + 2R^T S (Py + Qz) \right). \quad (3.2)$$

Это управление доставляет минимум последнему слагаемому в (3.1):

$$\begin{aligned} & u^T \left( c^T \frac{\partial V}{\partial z} + 2R^T S (Py + Qz) \right) + \|Ru\|_S^2 \Big|_{u=u^*} = \\ & = -1/4 \left( c^T \frac{\partial V}{\partial z} + 2R^T S (Py + Qz) \right)^T (R^T S R)^{-1} \left( c^T \frac{\partial V}{\partial z} + 2R^T S (Py + Qz) \right). \end{aligned}$$

Отметим, что оптимальное управление получается из двух слагаемых: одно слагаемое линейное по состоянию и выходу (как в классической линейно-квадратичной задаче), второе слагаемое нелинейное, определяется производной  $\partial V / \partial z$ .

Таким образом, функция Беллмана описывается следующим дифференциальным уравнением:

$$\begin{aligned} & \frac{\partial V}{\partial t} + 1/2 \text{tr} \left\{ \Sigma^T \frac{\partial^2 V}{\partial y^2} \Sigma + \sigma^T \frac{\partial^2 V}{\partial z^2} \sigma \right\} + \\ & + A^T \frac{\partial V}{\partial y} + (ay + bz)^T \frac{\partial V}{\partial z} + \|Py + Qz\|_S^2 - \\ & - 1/4 \left( c^T \frac{\partial V}{\partial z} + 2R^T S (Py + Qz) \right)^T (R^T S R)^{-1} \left( c^T \frac{\partial V}{\partial z} + 2R^T S (Py + Qz) \right) = 0. \end{aligned} \quad (3.3)$$

Преобразуя последнее слагаемое в (3.3), перепишем его в виде

$$\begin{aligned} & \frac{\partial V}{\partial t} + 1/2 \text{tr} \left\{ \Sigma^T \frac{\partial^2 V}{\partial y^2} \Sigma + \sigma^T \frac{\partial^2 V}{\partial z^2} \sigma \right\} + A^T \frac{\partial V}{\partial y} + \\ & + ((ay + bz)^T - (R^T S (Py + Qz))^T (R^T S R)^{-1} c^T) \frac{\partial V}{\partial z} + \\ & + (Py + Qz)^T (S - S R (R^T S R)^{-1} R^T S) (Py + Qz) - \\ & - 1/4 \left( \frac{\partial V}{\partial z} \right)^T c (R^T S R)^{-1} c^T \frac{\partial V}{\partial z} = 0. \end{aligned} \quad (3.4)$$

Рассматривая полученное уравнение, предположим, что его решение может быть представлено в как

$$V = V(t, y, z) = z^T \alpha(t) z + z^T \beta(t, y) + \gamma(t, y), \quad (3.5)$$

т.е. решением является функция, квадратичная по переменной  $z$ , и, таким образом, поиск оптимального решения сводится к поиску уравнений относительно симметричной матричной функции  $\alpha = \alpha(t) \in \mathbb{R}^{n_z \times n_z}$ ,  $\alpha = \alpha^T$ , векторной функции  $\beta = \beta(t, y) \in \mathbb{R}^{n_z}$  и скалярной функции  $\gamma = \gamma(t, y) \in \mathbb{R}^1$ . Отметим сразу, что явный вид функции  $\gamma$  для реализации оптимального управления не требуется, однако далее будет предложен вариант вычисления и этой функции. Источником для предложения (3.5) является уже упоминавшаяся аналогичная задача для случая дискретного времени [10, 11]. В той задаче выражение для функции Беллмана получается формально без дополнительных усилий. При этом форма (3.5) обнаруживается как свойство оптимального решения. В рассматриваемом случае непрерывного времени (3.5) является предположением, правильность которого подтверждается далее результирующими уравнениями для  $\alpha, \beta, \gamma$ . Кроме того, данное предположение представляется вытекающим из линейной структуры задачи в отношении переменной  $z$ , в частности тем фактом, что такой вид функции Беллмана обеспечивает линейность оптимального управления (3.2) по  $z$ .

Граничное условие при сделанном предположении (3.5) принимает вид

$$\begin{aligned} & V(T, y, z) = \|P(T)y + Q(T)z\|_{S(T)}^2 = \\ & = z^T Q^T(T) S(T) Q(T) z + 2z^T Q^T(T) S(T) P(T) y + y^T P^T(T) S(T) P(T) y, \end{aligned}$$

т.е.

$$\begin{aligned} & \alpha(T) = Q^T(T) S(T) Q(T), \quad \beta(T, y) = 2Q^T(T) S(T) P(T) y, \\ & \gamma(T, y) = y^T P^T(T) S(T) P(T) y. \end{aligned} \quad (3.6)$$

При этом поскольку (3.5) дает  $\partial V/\partial z = 2\alpha z + \beta$ , то само оптимальное управление, определенное выражением (3.2),

$$u^* = u^*(t, y, z) = -1/2 (R^T S R)^{-1} (c^T (2\alpha z + \beta) + 2R^T S (P y + Q z)). \quad (3.7)$$

Дальнейшие преобразования удобнее выполнить, представив  $V$  в виде

$$V = z^T \alpha z + \sum_{i=1}^{n_z} z^{(i)} \beta^{(i)} + \gamma,$$

обозначая  $z = (z^{(1)}, \dots, z^{(n_z)})^T$ ,  $\beta = (\beta^{(1)}, \dots, \beta^{(n_z)})^T$ ,  $\beta^{(i)} = \beta^{(i)}(t, y)$ , чтобы не сталкиваться при дифференцировании с необходимостью использования тензорных обозначений для производных векторной функции. Теперь следует подставить  $V$ , представленную в виде (3.5), в уравнение (3.4):

$$\begin{aligned} & z^T \frac{d\alpha}{dt} z + \sum_{i=1}^{n_z} z^{(i)} \frac{\partial \beta^{(i)}}{\partial t} + \frac{\partial \gamma}{\partial t} + 1/2 \text{tr} \left\{ \Sigma^T \left( \sum_{i=1}^{n_z} z^{(i)} \frac{\partial^2 \beta^{(i)}}{\partial y^2} + \frac{\partial^2 \gamma}{\partial y^2} \right) \Sigma + 2\sigma^T \alpha \sigma \right\} + \\ & + A^T \left( \sum_{i=1}^{n_z} z^{(i)} \frac{\partial \beta^{(i)}}{\partial y} + \frac{\partial \gamma}{\partial y} \right) + ((ay + bz)^T - (R^T S (P y + Q z))^T (R^T S R)^{-1} c^T) (2\alpha z + \beta) + \\ & + (P y + Q z)^T (S - S R (R^T S R)^{-1} R^T S) (P y + Q z) - 1/4 (2\alpha z + \beta)^T c (R^T S R)^{-1} c^T (2\alpha z + \beta) = 0. \end{aligned}$$

Раскрывая скобки и формируя слагаемые при  $z^T z, z$  (каждой компоненте  $z^{(i)}$ ) и  $z^0 = 1$ , имеем

$$\begin{aligned} & z^T \frac{d\alpha}{dt} z + \sum_{i=1}^{n_z} z^{(i)} \frac{\partial \beta^{(i)}}{\partial t} + \frac{\partial \gamma}{\partial t} + 1/2 \left( \sum_{i=1}^{n_z} z^{(i)} \text{tr} \left\{ \Sigma^T \frac{\partial^2 \beta^{(i)}}{\partial y^2} \Sigma \right\} \right) + \\ & + 1/2 \text{tr} \left\{ \Sigma^T \frac{\partial^2 \gamma}{\partial y^2} \Sigma \right\} + \text{tr} \{ \sigma^T \alpha \sigma \} + \sum_{i=1}^{n_z} z^{(i)} A^T \frac{\partial \beta^{(i)}}{\partial y} + A^T \frac{\partial \gamma}{\partial y} + \\ & + 2z^T (b^T - Q^T S R (R^T S R)^{-1} c^T) \alpha z + \\ & + (\beta^T (b - c (R^T S R)^{-1} R^T S Q) + 2y^T (a^T - P^T S R (R^T S R)^{-1} c^T) \alpha) z + \\ & + \beta^T (a - c (R^T S R)^{-1} R^T S P) y + z^T Q^T (S - S R (R^T S R)^{-1} R^T S) Q z + \\ & + 2y^T P^T (S - S R (R^T S R)^{-1} R^T S) Q z + y^T P^T (S - S R (R^T S R)^{-1} R^T S) P y - \\ & - z^T \alpha^T c (R^T S R)^{-1} c^T \alpha z - \beta^T c (R^T S R)^{-1} c^T \alpha z - 1/4 \beta^T c (R^T S R)^{-1} c^T \beta = 0. \end{aligned}$$

Окончательно требуемая группировка слагаемых записывается с использованием обозначения  $[A]^{(ji)}$  для элемента  $j$ -й строки  $i$ -го столбца матрицы  $A$ :

$$\begin{aligned} & z^T \left( \begin{aligned} & \frac{d\alpha}{dt} + 2(b^T - Q^T S R (R^T S R)^{-1} c^T) \alpha + \\ & + Q^T (S - S R (R^T S R)^{-1} R^T S) Q - \alpha^T c (R^T S R)^{-1} c^T \alpha \end{aligned} \right) z + \\ & + \sum_{i=1}^{n_z} z^{(i)} \left( \begin{aligned} & \frac{\partial \beta^{(i)}}{\partial t} + 1/2 \text{tr} \left\{ \Sigma^T \frac{\partial^2 \beta^{(i)}}{\partial y^2} \Sigma \right\} + A^T \frac{\partial \beta^{(i)}}{\partial y} + \\ & + \sum_{j=1}^{n_z} \beta^{(j)} [b - c (R^T S R)^{-1} R^T S Q - c (R^T S R)^{-1} c^T \alpha]^{(ji)} + \\ & + 2 \sum_{j=1}^{n_y} y^{(j)} [(a^T - P^T S R (R^T S R)^{-1} c^T) \alpha + P^T (S - S R (R^T S R)^{-1} R^T S) Q]^{(ji)} \end{aligned} \right) + \\ & + \left( \begin{aligned} & \frac{\partial \gamma}{\partial t} + 1/2 \text{tr} \left\{ \Sigma^T \frac{\partial^2 \gamma}{\partial y^2} \Sigma \right\} + \text{tr} \{ \sigma^T \alpha \sigma \} + A^T \frac{\partial \gamma}{\partial y} + \beta^T (a - c (R^T S R)^{-1} R^T S P) y + \\ & + y^T P^T (S - S R (R^T S R)^{-1} R^T S) P y - 1/4 \beta^T c (R^T S R)^{-1} c^T \beta \end{aligned} \right) = 0. \end{aligned}$$



Поскольку полученное равенство выполняется для всех  $z \in \mathbb{R}^{n_z}$ ,  $y \in \mathbb{R}^{n_y}$ , то каждое из трех слагаемых должно быть равно 0, что дает следующие уравнения:

– матричное дифференциальное уравнение для функции  $\alpha = \alpha(t)$  (с учетом предположения  $\alpha = \alpha^T$ ):

$$\begin{aligned} \frac{d\alpha}{dt} + (b^T - Q^T SR(R^T SR)^{-1} c^T)\alpha + \alpha(b - c(R^T SR)^{-1} R^T SQ) + \\ + Q^T(S - SR(R^T SR)^{-1} R^T S)Q - \alpha c(R^T SR)^{-1} c^T \alpha = 0, \quad \alpha(T) = Q^T(T)S(T)Q(T), \end{aligned} \quad (3.8)$$

– систему дифференциальных уравнений в частных производных для функций  $\beta^{(i)} = \beta^{(i)}(t, y), i = \overline{1, n_z}$ :

$$\begin{aligned} \frac{\partial \beta^{(i)}}{\partial t} + 1/2 \text{tr} \left\{ \Sigma^T \frac{\partial^2 \beta^{(i)}}{\partial y^2} \Sigma \right\} + A^T \frac{\partial \beta^{(i)}}{\partial y} + \\ + \sum_{j=1}^{n_z} \beta^{(j)} [b - c(R^T SR)^{-1} R^T SQ - c(R^T SR)^{-1} c^T \alpha]^{(ji)} + \\ + 2 \sum_{j=1}^{n_y} y^{(j)} [(a^T - P^T SR(R^T SR)^{-1} c^T)\alpha + P^T(S - SR(R^T SR)^{-1} R^T S)Q]^{(ji)} = 0, \end{aligned} \quad (3.9)$$

$$\beta^{(i)}(T, y) = 2 \sum_{j=1}^{n_y} y^{(j)} [Q^T(T)S(T)P(T)]^{(ji)},$$

– дифференциальное уравнение в частных производных для функции  $\gamma = \gamma(t, y)$ :

$$\begin{aligned} \frac{\partial \gamma}{\partial t} + 1/2 \text{tr} \left\{ \Sigma^T \frac{\partial^2 \gamma}{\partial y^2} \Sigma \right\} + \text{tr} \{ \sigma^T \alpha \sigma \} + A^T \frac{\partial \gamma}{\partial y} + \beta^T (a - c(R^T SR)^{-1} R^T SP)y + \\ + y^T P^T (S - SR(R^T SR)^{-1} R^T S)Py - 1/4 \beta^T c(R^T SR)^{-1} c^T \beta = 0, \\ \gamma(T, y) = y^T P^T(T)S(T)P(T)y. \end{aligned} \quad (3.10)$$

Вводя дополнительно обозначения

$$M = M(t) = 2((a^T - P^T SR(R^T SR)^{-1} c^T)\alpha + P^T(S - SR(R^T SR)^{-1} R^T S)Q),$$

$$N = N(t) = b - c(R^T SR)^{-1} R^T SQ - c(R^T SR)^{-1} c^T \alpha,$$

$$K = K(t, y) = \beta^T (a - c(R^T SR)^{-1} R^T SP)y + y^T P^T (S - SR(R^T SR)^{-1} R^T S)Py - 1/4 \beta^T c(R^T SR)^{-1} c^T \beta,$$

для  $\beta(t, y)$  и  $\gamma(t, y)$  вместо (3.9) и (3.10) окончательно получаем

$$\frac{\partial \beta^{(i)}}{\partial t} + 1/2 \text{tr} \left\{ \Sigma^T \frac{\partial^2 \beta^{(i)}}{\partial y^2} \Sigma \right\} + A^T \frac{\partial \beta^{(i)}}{\partial y} + \sum_{j=1}^{n_y} y^{(j)} [M]^{(ji)} + \sum_{j=1}^{n_z} \beta^{(j)} [N]^{(ji)} = 0, \quad (3.11)$$

$$\frac{\partial \gamma}{\partial t} + 1/2 \text{tr} \left\{ \Sigma^T \frac{\partial^2 \gamma}{\partial y^2} \Sigma \right\} + \text{tr} \{ \sigma^T \alpha \sigma \} + A^T \frac{\partial \gamma}{\partial y} + K = 0. \quad (3.12)$$

Уравнение (3.8) – это матричное уравнение Риккати, точнее его частный случай для квадратной симметричной матрицы  $\alpha$ . Сделанных выше предположений достаточно для существования единственного неотрицательного решения для всех  $0 \leq t \leq T$ . Этот факт требует дополнительно комментария. Рассмотрим исходную задачу при дополнительных условиях:  $a(t) = 0_{n_z \times n_y}$ ,  $P(t) = 0_{n_y \times n_y}$ ,  $S(t) = E$  для всех  $0 \leq t \leq T$ . Это дает

$$J(U_0^T) = \mathbb{E} \left\{ \int_0^T (\|z\|_Q^2 + \|u\|_R^2) dt + \|z(T)\|_{Q(T)}^2 \right\}$$

и сводит рассматриваемую постановку к классической LQG-задаче. Для такого упрощения целевого функционала требуется наличие блочной структуры матриц  $Q, R$  из (2.3), чтобы исклю-

чить в упрощенном  $J(U_0^T)$  слагаемое с множителями  $z$  и  $u$ . В этой упрощенной форме матрицы  $Q(t) \in \mathbb{R}^{n_z \times n_z}$  и  $R(t) \in \mathbb{R}^{n_u \times n_u}$  представляют соответствующие блоки одноименных матриц из (2.3).

Теперь уравнение (3.8) для  $\alpha(t)$  принимает хорошо известный вид

$$\frac{d\alpha}{dt} + b^T \alpha + \alpha b + Q^T Q - \alpha c (R^T R)^{-1} c^T \alpha = 0, \quad \alpha(T) = Q^T(T) Q(T). \quad (3.13)$$

В таком случае, как известно [19], существует единственное оптимальное управление – линейное с обратной связью по выходу  $z(t)$ , с коэффициентом усиления, описываемым уравнением Риккати (3.13). Именно этот результат дают уравнения (3.8)–(3.10) и описываемая ими функция Беллмана (3.5), так как из  $a(t) = 0, P(t) = 0$  немедленно следует, что  $\beta^{(i)}(t, y) = 0, i = \overline{1, n_z}$ , откуда, в свою очередь, с учетом независимости решения от  $y$  следует, что  $\gamma$  не зависит от  $y$  и задается уравнением  $d\gamma/dt + \text{tr}\{\sigma^T \alpha \sigma\} = 0, \gamma(T) = 0$ . Оптимальное управление при этом  $u^* = -(R^T R)^{-1} c^T \alpha z$ , т.е. все полностью совпадает с известным классическим решением.

С уравнениями (3.11) и (3.12) ситуация, естественно, обстоит сложнее. Это линейные уравнения в частных производных второго порядка параболического типа, поскольку  $\Sigma^T \Sigma \geq 0$ . Уравнения неоднородные, причем в уравнении (3.12) для  $\gamma(t, y)$  участвуют все  $\beta^{(i)}(t, y)$ , т.е. сначала должны быть решены уравнения (3.11) и результат подставлен в (3.12). Также в уравнения (3.11) нужно подставить  $\alpha(t)$  – решение уравнения (3.8). Далее, в каждом из уравнений (3.11) неоднородность для  $\beta^{(i)}(t, y)$  формируется неизвестными функциями  $\beta^{(j)}(t, y), j \neq i$ , т.е. параболические уравнения объединены в систему и по отдельности решаться не могут, так что актуален еще и вопрос их приближенного решения. Кроме того, отсутствуют конструктивные условия, гарантирующие существование решений (требовать, чтобы все фигурирующие в уравнениях коэффициенты были представлены аналитическими функциями на всем пространстве значений, вряд ли целесообразно, а в имеющейся системе уравнений, может быть, и не достаточно). Поэтому далее будем предполагать, что уравнения (3.11) и (3.12) имеют на интервале  $0 \leq t \leq T$  хотя бы одно гладкое (непрерывно дифференцируемое по  $t$  и дважды непрерывно дифференцируемое по  $y$ ) решение и именно эти условия можно считать достаточными условиями существования оптимального решения рассматриваемой задачи.

Таким образом, доказана следующая теорема.

**Теорема.** Пусть для диффузионного процесса (2.1) выполнены условия Ито, для выхода (2.2) ограничены коэффициенты, уравнения (3.8)–(3.10) имеют гладкие решения для  $0 \leq t \leq T$ . Тогда минимум функционалу (2.3) доставляет  $(U^*)_0^T$  – оптимальное управление с полной обратной связью, определенное (3.7).

**4. Стохастическое представление коэффициентов функции Беллмана.** Для реализации оптимального управления и оценки его качества потребуются решать полученные уравнения (3.8), (3.11), (3.12) приближенно. Решение уравнения (3.8) не представляет трудностей и может быть выполнено любым из множества эффективных численных методов решения обыкновенных дифференциальных уравнений. Несколько сложнее обстоит ситуация с уравнениями (3.11), (3.12). Для уравнений параболического типа также имеются эффективные численные схемы, основанные на методе конечных разностей. Но в рассматриваемом случае трудности в реализации возникают при определении граничных условий, необходимых для этого метода. В рассматриваемой задаче физические основания для определения границы и, тем более граничных условий как для  $\beta(t, y)$ , так и для  $\gamma(t, y)$  отсутствуют. Вариант возможного преодоления этих сложностей обсуждается далее в рамках представления численного эксперимента. Более подходящим и лишены таких трудностей представляется решение параболического уравнения на основе его ассоциации с уравнением А.Н. Колмогорова, точнее связи между решением задачи Коши с терминальным условием для уравнения А.Н. Колмогорова и теоретико-вероятностным представлением этого терминального условия [16]. Выражением этого результата является формула Фейнмана–Каца [17], непосредственно применимая к уравнению (3.12). С учетом введенных выше

обозначений интегральное представление Фейнмана–Каца решения дифференциального уравнения (3.12) имеет вид

$$\gamma(t, y) = \mathbb{E} \left\{ y^T(T) P^T(T) S(T) P(T) y(T) + \int_t^T K(\tau, y(\tau)) d\tau \middle| \mathcal{F}_t^y \right\}. \quad (4.1)$$

В предположении, что  $\alpha(t)$  и  $\beta(t, y)$  известны (вычислены), (4.1) дает основание для применения компьютерного моделирования для аппроксимации  $\gamma(t, y)$ . Многократно моделируя решение уравнения (2.1), проходящее через точку  $(t, y)$  (именно это обозначает  $\sigma$ -алгебра  $\mathcal{F}_t^y$  в условном математическом ожидании), значение  $\gamma(t, y)$  можно получить как статистическую оценку функции терминального условия, согласно (4.1). Подробнее этот метод описан далее. Следует отметить, что как в формулировке для уравнения А.Н. Колмогорова, так и в формуле Фейнмана–Каца утверждается существование процесса  $y$ , обеспечивающего равенство (4.1), т.е. формально это другой процесс и задан он на другом пространстве, чем процесс, участвующий в поставленной выше задаче управления. При этом описывается этот процесс таким же уравнением (2.1) и имеет те же вероятностные характеристики, что дает основание использовать уже имеющуюся дифференциальную систему, не вводя новые обозначения.

Такой же подход хотелось бы применить для приближенного вычисления  $\beta(t, y)$ . Более того, в скалярном случае  $n_z = 1$  интегральное представление Фейнмана–Каца решения дифференциального уравнения (3.11) записывается без дополнительных усилий и имеет вид

$$\beta(t, y) = \mathbb{E} \left\{ \begin{array}{l} 2 \exp \left\{ \int_t^T N(\tau) d\tau \right\} y^T(T) P^T(T) S(T) Q(T) + \\ + \int_t^T \exp \left\{ \int_t^\tau N(s) ds \right\} y^T(\tau) M(\tau) d\tau \end{array} \middle| \mathcal{F}_t^y \right\}. \quad (4.2)$$

Далее выводится аналог (4.2) для  $n_z > 1$ . Для этого введем новую переменную  $\bar{\beta} = \bar{\beta}(t, y) \in \mathbb{R}^{n_z}$  и выполним замену  $\beta(t, y) = I(t) \bar{\beta}(t, y)$ , где матрица  $I = I(t) \in \mathbb{R}^{n_z \times n_z}$  является решением уравнения  $dI(t)/dt + N^T(t)I(t) = 0$ . Такая замена позволяет исключить из уравнений (3.11) слагаемые с  $\beta^{(j)}$ . Действительно, для слагаемых

$$\frac{\partial \beta^{(i)}}{\partial t} + \sum_{j=1}^{n_z} \beta^{(j)} [N]^{(ji)}, \quad i = \overline{1, n_z},$$

записанных в векторной форме  $\frac{\partial \bar{\beta}}{\partial t} + N^T \bar{\beta}$ , замена переменной на  $\bar{\beta}$  дает  $I \frac{\partial \bar{\beta}}{\partial t} + dI/dt \bar{\beta} + N^T I \bar{\beta} = I \frac{\partial \bar{\beta}}{\partial t}$ . Необходимая для такой замены матрица  $I(t)$  существует и равна матричной экспоненте

$$I(t) = \exp \left\{ - \int_0^t N^T(\tau) d\tau \right\},$$

причем  $I(t)$  невырожденная [20], что позволяет далее использовать обратную матрицу  $I^{-1}(t)$ . Остальные слагаемые в (3.11) с дифференцированием по  $t$  не связаны и линейны по  $\bar{\beta}$ , что дает

$$\sum_{k=1}^{n_z} \frac{\partial \bar{\beta}^{(k)}}{\partial t} [I]^{(ik)} + \sum_{k=1}^{n_z} 1/2 \text{tr} \left\{ \Sigma^T \frac{\partial^2 \bar{\beta}^{(k)}}{\partial y^2} \Sigma \right\} [I_t]^{(ik)} + \sum_{k=1}^{n_z} A^T \frac{\partial \bar{\beta}^{(k)}}{\partial y} [I]^{(ik)} + \sum_{j=1}^{n_y} y^{(j)} [M]^{(ji)} = 0.$$

Следующим действием левая часть полученного равенства разрешается относительно эллиптической части и  $\frac{\partial \bar{\beta}}{\partial t}$  умножением на  $I^{-1}(t)$

$$\frac{\partial \bar{\beta}^{(i)}}{\partial t} + 1/2 \text{tr} \left\{ \Sigma^T \frac{\partial^2 \bar{\beta}^{(i)}}{\partial y^2} \Sigma \right\} + A^T \frac{\partial \bar{\beta}^{(i)}}{\partial y} + \sum_{j=1}^{n_y} y^{(j)} [I^{-1} M^T]^{(ij)} = 0, \quad i = \overline{1, n_z}. \quad (4.3)$$

Теперь для каждого из уравнений (4.3) может быть записана формула Фейнмана–Каца:

$$\bar{\beta}^{(i)}(t, y) = \mathbb{E} \left\{ \begin{array}{l} 2 \sum_{j=1}^{n_y} y^{(j)}(T) [I^{-1}(T) Q^T(T) S(T) P(T)]^{(ij)} + \\ + \int_t^T \sum_{j=1}^{n_y} y^{(j)}(\tau) [I^{-1}(\tau) M^T(\tau)]^{(ij)} d\tau \end{array} \middle| \mathcal{F}_t^y \right\}$$

или в векторном виде

$$\bar{\beta}(t, y) = \mathbb{E} \left\{ 2I^{-1}(T) Q^T(T) S(T) P(T) y(T) + \int_t^T I^{-1}(\tau) M^T(\tau) y(\tau) d\tau \middle| \mathcal{F}_t^y \right\}.$$

Наконец, обратная замена  $\bar{\beta}(t, y) = I^{-1}(t) \beta(t, y)$  дает

$$\beta(t, y) = \mathbb{E} \left\{ \begin{array}{l} 2I^{-1}(t) I^{-1}(T) Q^T(T) S(T) P(T) y(T) + \\ + I^{-1}(t) \int_t^T I^{-1}(\tau) M^T(\tau) y(\tau) d\tau \end{array} \middle| \mathcal{F}_t^y \right\}. \quad (4.4)$$

Приводя выражения, нужные для расчета по формулам (4.1) и (4.4), в дифференциальный вид, окончательно получаем следующую систему:

$$\begin{aligned} dy(\tau) &= A(\tau, y(\tau)) d\tau + \Sigma(\tau, y(\tau)) dv(\tau), \quad y_t = y, \quad t \leq \tau \leq T, \\ \frac{dI^{-1}(\tau)}{d\tau} - N^T(\tau) I^{-1}(\tau) &= 0, \quad I^{-1}(t) = \exp \left\{ \int_0^t N^T(s) ds \right\}, \\ \frac{dy^\beta(\tau)}{d\tau} &= I^{-1}(\tau) M^T(\tau) y(\tau), \quad y^\beta(\tau) = 0, \\ \beta(t, y) &= \mathbb{E} \{ I^{-1}(t) I^{-1}(T) Q^T(T) S(T) P(T) y(T) + I^{-1}(t) y^\beta(T) \}, \\ \frac{dy^\gamma(\tau)}{d\tau} &= K(\tau, y(\tau)), y^\gamma(t) = 0, \\ \gamma(t, y) &= \mathbb{E} \{ y^T(T) P^T(T) S(T) P(T) y(T) + y^\gamma(T) \}. \end{aligned} \quad (4.5)$$

Здесь для удобства интегрирования введены дополнительные переменные  $y^\beta(\tau) \in \mathbb{R}^{n_z}$ ,  $y^\gamma(\tau) \in \mathbb{R}^1$  и учтено, что

$$I^{-1}(\tau) = \exp \left\{ \int_0^\tau N^T(s) ds \right\}.$$

Полученная дифференциальная система (4.5) дает альтернативный вариант приближенного расчета оптимального управления (3.7) в отношении расчета коэффициентов  $\beta(t, y)$  и определения его качества  $J((U^*)_0^T) = V(0, Y, Z) = \mathbb{E} \{ Z^T \alpha(0) Z + Z^T \beta(0, Y) + \gamma(0, Y) \}$  в отношении расчета коэффициента  $\gamma(0, y)$ . Именно многократно моделируя решение системы (4.5) для различных начальных условий  $(t, y)$ , указанные коэффициенты можно оценивать по методу Монте-Карло, заменяя операторы  $\mathbb{E}\{\cdot\}$  в (4.5) статистическим средним. Детально соответствующий алгоритм приведен в [21]. Заметим, что альтернативным предлагаемый метод квалифицируется по отношению к традиционному способу решения параболических уравнений (3.11), (3.12) методом конечных разностей.

**5. Частный случай линейного сноса.** Прежде чем перейти к вопросам практической реализации и обсуждению модельного примера, рассмотрим один частный случай, когда оптимальное управление (3.7) может быть найдено аналитически, т.е. без численного решения параболического уравнения (3.11). Очевидно, что частным случаем рассматриваемой задачи является классическая линейно-гауссовская постановка, о чем уже упоминалось выше. В таком случае функция Беллмана (3.5) является квадратичной, что означает линейность коэффициентов  $\beta(t, y)$  и

квадратичность коэффициента  $\gamma(t, y)$  по  $y$ , т.е. решения (3.11), (3.12) могут быть найдены в виде  $\beta(t, y) = \beta^a(t)y + \beta^b(t)$  и  $\gamma(t, y) = y^T \gamma^a(t)y + y^T \gamma^b(t) + \gamma^c(t)$  соответственно, где матрица  $\beta^a(t) \in \mathbb{R}^{n_z \times n_y}$ , вектор  $\beta^b(t) \in \mathbb{R}^{n_z}$ , матрица  $\gamma^a(t) \in \mathbb{R}^{n_y \times n_y}$ , вектор  $\gamma^b(t) \in \mathbb{R}^{n_y}$  и функция  $\gamma^c(t) \in \mathbb{R}^1$  не зависят от  $y$ . Такой случай с точки зрения новизны интереса не представляет.

Однако рассматриваемая постановка включает другой более интересный частный вариант. Из-за независимости состояния  $y(t)$ , заданного уравнением (2.1), от управления  $u(t)$ , которое входит линейно только в уравнение (2.2) выхода  $z(t)$ , для линейности коэффициентов  $\beta(t, y)$  достаточно линейного сноса  $A(t, y)$  в (2.1), а линейность диффузии  $\Sigma(t, y)$  не требуется. В отличие от случая LQG функция Беллмана при этом не будет квадратичной, решение уравнения (3.12) в виде  $\gamma(t, y) = y^T \gamma^a(t)y + y^T \gamma^b(t) + \gamma^c(t)$  не представляется именно из-за нелинейности  $\Sigma(t, y)$ , но простое линейное представление оптимального управления (3.7) в этом практически важном частном случае получается.

Итак, в предположении, что  $A(t, y) = A^a(t)y + A^b(t)$ , где матрица  $A^a(t) \in \mathbb{R}^{n_y \times n_y}$  и вектор  $A^b(t) \in \mathbb{R}^{n_y}$  не зависят от  $y$ , найдем решение (3.11) в виде  $\beta(t, y) = \beta^a(t)y + \beta^b(t)$ . При этом предположении

$$\frac{\partial^2 \beta}{\partial y^2} = 0, \quad \frac{\partial \beta}{\partial y} = (\beta^a)^T.$$

Кроме того, учтем, что для слагаемых с  $\partial \beta^{(i)} / \partial y$  в (3.11) имеет место

$$A^T \frac{\partial \beta^{(i)}}{\partial y} = \left( \frac{\partial \beta^{(i)}}{\partial y} \right)^T A = \left( \frac{\partial \beta^{(i)}}{\partial y} \right)^T (A^a(t)y + A^b(t)),$$

так что

$$A^T \frac{\partial \beta}{\partial y} = \beta^a(t)(A^a(t)y + A^b(t)),$$

а уравнения (3.11), записанные в векторном виде, дают

$$\frac{d\beta^a}{dt}y + \frac{d\beta^b}{dt} + \beta^a A^a y + \beta^a A^b + M^T y + N^T \beta^a y + N^T \beta^b = 0,$$

откуда получаются уравнения для  $\beta^a(t)$  и  $\beta^b(t)$ :

$$\frac{d\beta^a}{dt} + \beta^a A^a + M^T + N^T \beta^a = 0, \quad \frac{d\beta^b}{dt} + \beta^a A^b + N^T \beta^b = 0, \quad (5.1)$$

т.е. система обыкновенных линейных дифференциальных уравнений. Оптимальное управление (3.7) принимает вид

$$u^* = u^*(t, y, z) = -1/2 (R^T S R)^{-1} (2(c^T \alpha + R^T S Q)z + (c^T \beta^a + 2R^T S P)y + c^T \beta^b). \quad (5.2)$$

**6. Вопросы численной реализации и модельный пример.** В этом разделе обсуждаются некоторые аспекты практической реализации полученного оптимального управления. С этой целью рассматривается прикладной пример из области телекоммуникаций – популярного современного источника приложений для различных теоретических задач. Далее использована простая модель для показателя RTT сетевого протокола TCP [22]. Рассматривается модельный пример, поэтому параметры подобраны для лучшей выразительности иллюстративного материала. Переменная состояния RTT описывается следующей диффузией:

$$dy(t) = (1 - y(t))dt + 2.5\sqrt{y(t)}dv(t), \quad y(0) = Y \sim \mathbb{N}(15, 9). \quad (6.1)$$

Здесь  $\mathbb{N}(M, D)$  – нормальное распределение со средним  $M$  и дисперсией  $D$ .

Надо отметить, что такое уравнение (с точностью до варьируемых параметров) изначально известно как модель Кокса–Ингерсона–Росса (Cox–Ingersoll–Ross model) [23] и описывает эволюцию процентных ставок. Конечно, для описания реальных данных RTT модель типа (6.1)

представляет не более чем начальное приближение, но оно полезно уже потому, что дает основу для дальнейших обобщений, уточнений модели (см., например, [7]). Свойства процесса  $y(t)$  из (6.1) хорошо изучены, в частности имеется свойство неотрицательности выборочных значений  $y(t) > 0$ , наличие эргодичности, известно предельное распределение и переходная вероятность. Отметим, что начальные условия в (6.1) выбраны так, что  $M, D$  значительно отличаются от моментных характеристик предельного распределения, т.е. управление ведется в рамках переходного процесса, а не в стационарном режиме.

Выход для состояния (6.1) задается уравнением

$$dz(t) = y(t)dt - z(t)dt + u(t)dt + 2.5dw(t), \quad z_0 = Z \sim \mathbb{N}(9,9), \quad (6.2)$$

целевой функционал

$$J(U_0^T) = \mathbb{E} \left\{ \int_0^T ((y(t) - z(t))^2 + z^2(t) + u^2(t))dt + (y(T) - z(T))^2 + z^2(T) \right\}. \quad (6.3)$$

Такой выход имитирует выделение некоторого временного ресурса, складывающегося из текущего показателя РГТ, управляющего воздействия и возмущения. Соответственно функционал (6.3) ставит целью отслеживание выходной переменной состояния с минимизацией затрат на управляющие воздействия и абсолютное значение выхода. Таким образом, понятная для практики задача отслеживания состояния дополняется учетом значения выхода, которое не позволит ему неограниченно расти в случае роста РГТ, и платой за применение управляющих воздействий, которые в реальности связаны с некоторыми изменениями конфигурации оборудования или программного обеспечения.

Моделировалось  $N = 1000$  траекторий  $y(t), z^*(t), u^*(t)$  для  $T = 5$ . Здесь через  $z^*(t)$  обозначен выход (6.2), рассчитанный для оптимального управления с обратной связью  $u^* = u^*(t, y, z)$ . Кроме того, моделировались траектории  $z^{(prog)}(t), u^{(prog)}(t)$  для программного управления  $u^{(prog)}(t) = \mathbb{E}\{u^*(t)\}$  и  $z^{(0)}(t), u^{(0)}(t)$  для неуправляемого выхода (6.2), т.е. для  $u(t) = 0$ . В каждом случае путем осреднения по пучку траекторий оценивались величины  $J(U_0^T)$ . Соответственно в результате можно увидеть не только конечные значения целевой функции для разных вариантов управления  $J((U^*)_0^T), J((U^{(prog)})_0^T), J((U^{(0)})_0^T)$ , но и их формирование в динамике.

Описание численных процедур начнем с моделирования. Траектории  $y(t), z(t)$  моделировались приближенно, для чего уравнения (6.1), (6.2) дискретизировались по методу Эйлера с шагом  $\delta_t = 0.001$  (тем же, что далее использовался при численной аппроксимации  $\beta(t, y)$ ). Соответствующие смоделированные значения  $y(t - \delta_t), z^*(t - \delta_t)$  применялись в (3.7) для вычисления реализаций управления  $u^*(t) = u^*(t, y(t), z^*(t))$ .

Далее заметим, что уравнение Риккати (3.8) для  $\alpha(t)$  с учетом постоянных коэффициентов в выбранной модели выхода (6.2) и целевом функционале (6.3) решается точно. Именно

$$\alpha(t) = \frac{C_\alpha e^{2\sqrt{3}t}(1 + \sqrt{3}) - 1 + \sqrt{3}}{1 - C_\alpha e^{2\sqrt{3}t}}, \quad C_\alpha = \frac{3 - \sqrt{3}}{3 + \sqrt{3}} e^{-10\sqrt{3}}.$$

Наконец, оптимальное управление  $u^*$  рассчитывалось тремя способами. Во-первых, поскольку снос в модели (6.1) линейный, то в рассматриваемом примере можно вычислить решение, не интегрируя общее уравнение (3.11) для  $\beta(t, y)$ , а решая уравнения (5.1) и вычисляя далее управление, согласно (5.2). Система (5.1) линейных уравнений для  $\beta^a(t)$  и  $\beta^b(t)$  решалась неявным методом Эйлера с тем же шагом  $\delta_t = 0.001$ . Соответствующее приближенное оптимальное управление обозначается далее  $u^{(ref)}$ , решение уравнение (3.11) —  $\beta^{(ref)}(t, y)$ . Эти аппроксимации в данном примере носят референсный (эталонный) характер.

Второй способ аппроксимации  $\beta(t, y)$  — использование традиционного метода конечных разностей для решения параболического уравнения (3.11). Реализованные варианты и особенности применения этого метода ниже обсуждаются подробнее, а соответствующее приближенное решение обозначается  $u^{(net)}$ , решение уравнения (3.11) —  $\beta^{(net)}(t, y)$ .

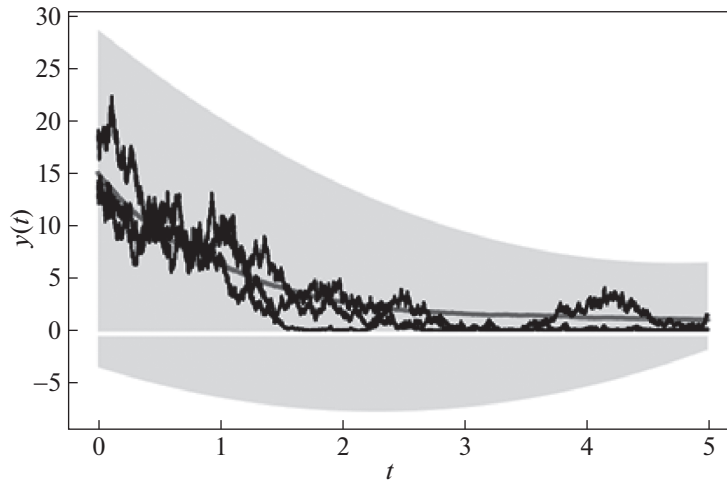


Рис. 1. Выборочные траектории и характеристика границы

Третий способ – моделирование решений системы (4.5) и замена математического ожидания статистической оценкой (метод Монте-Карло). Соответствующие приближенные решения –  $u^{(МК)}$  и  $\beta^{(МК)}(t)$ .

Оказалось, что на графических иллюстрациях далее визуализировать разницу между реализациями управлений  $u^{(ref)}(t)$ ,  $u^{(net)}(t)$ ,  $u^{(МК)}(t)$  не удастся, поэтому в этих иллюстрациях использовано общее обозначение  $u^*(t)$ .

Вопросы применения метода конечных разностей для уравнения параболического типа давно и хорошо изучены [24]. Первое, что надо сделать для расчета  $\beta^{(net)}(t, y)$ , – это ограничить область значения аргумента  $y$  для формирования численной схемы расчета. С этой целью использовались  $N$  смоделированных траекторий (6.1): выборочные значения использовались для оценивания функций математического ожидания и дисперсионной, с помощью которой определялась  $3\sigma$ -трубка. Рисунок 1 иллюстрирует характерные траектории (6.1), среднее значение процесса и вычисленные границы (закрашенная серым область). В расчетах граница значений  $y$  задавалась добавлением еще 10% к границе  $3\sigma$ -трубки, чтобы уменьшить влияние выбора граничных условий, а также учитывалась положительность  $y(t)$ , поэтому далее область интегрирования  $y \in [\underline{y}, \bar{y}] = [0, 40]$ .

Определение границы, нужное для инициирования любой схемы метода конечных разностей, выявляет проблему отсутствия граничных условий для решаемого уравнения (3.11). Этих условий нет, и нет физических оснований для их обоснованного выбора. Из-за этого граница  $\bar{y}$  была сдвинута дополнительно, а сами граничные условия приходилось выбирать волюнтаристски в надежде на нечувствительность решения в отношении этих граничных условий.

Выбор граничного условия и анализ его влияния на решение выполнялся следующим образом. Были взяты два традиционных варианта граничных условий – в задаче Дирихле:  $\beta(t, y)|_{y=\bar{y}} = 0$ , т.е. условие поглощения, и в задаче Неймана:  $\partial\beta(t, y)/\partial y|_{y=\bar{y}} = 0$ , т.е. условие отражения. Причем наличие в данном примере эталонного решения  $\beta^{(ref)}(t, y)$  позволяет видеть, что оба варианта далеки от точного. Далее уравнение (3.11) решалось методом конечных разностей с использованием явной и неявной численных схем сначала для одного граничного условия, затем для второго и полученные решения сравнивались. Устойчивость в таком случае означает совпадение (близость) решений в области, выделенной по результатам моделирования, т.е. для  $y \in [0, 30]$ .

Для численного интегрирования в выбранной области  $[0, T] \cup [\underline{y}, \bar{y}]$  формировалась сетка  $\{t_k, y_i\}$  с шагами  $\delta_t$  (для интегрирования по переменной  $t$ ) и  $\delta_y$  (для интегрирования по

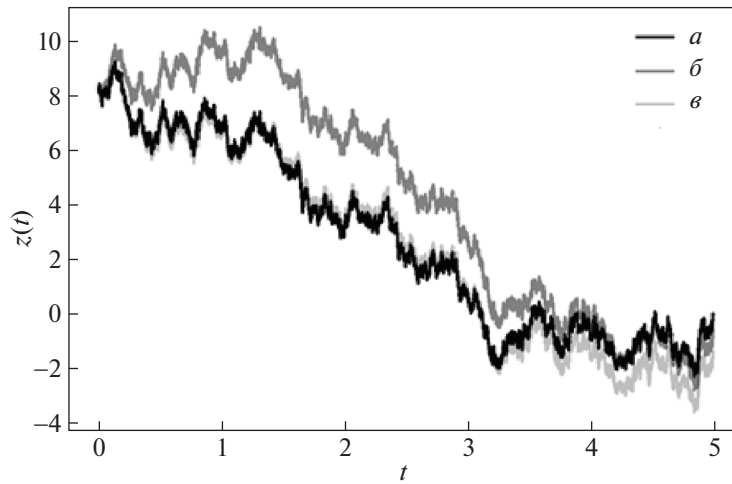


Рис. 2. Выборочные траектории:  $a - z^*(t)$ ,  $b - z^{(prog)}(t)$ ,  $e - z^{(0)}(t)$

переменной  $y$ ). С использованием обозначения  $\beta_i^k$  для значения  $\beta(t, y)$  в узле  $(t_k, y_i)$  производные аппроксимировались следующим образом:

$$\frac{\partial \beta(t, y)}{\partial t} \approx \frac{\beta_i^{k+1} - \beta_i^k}{\delta_t}, \quad \frac{\partial \beta(t, y)}{\partial y} \approx \frac{\beta_{i+1}^k - \beta_{i-1}^k}{2\delta_y}, \quad \frac{\partial^2 \beta(t, y)}{\partial y^2} \approx \frac{\beta_{i+1}^k - 2\beta_i^k + \beta_{i-1}^k}{\delta_y^2},$$

где  $K = k$  для явной схемы и  $K = k + 1$  — для неявной.

Шаги интегрирования  $\delta_t, \delta_y$  при анализе влияния граничных условий и устойчивости схем аппроксимации выбирались разными, например явная схема обшчитывалась для  $\delta_t = 0.0002$  и  $\delta_y = 0.1$ , в том числе с учетом того, что для нее самой вопрос устойчивости ограничивает вариации шагов. Главное, что в итоге подтвердилось отсутствие влияния выбора граничных условий на результаты численного интегрирования уравнения (3.11) для  $\beta(t, y)$ , по крайней мере, для выбранной модели, а также для комбинации шагов  $\delta_t, \delta_y$ , обеспечивающей устойчивость схем аппроксимации. Итоговый расчет выполнен для неявной схемы с шагами  $\delta_t = 0.001$  и  $\delta_y = 0.01$ , и именно это решение обозначено  $\beta^{(net)}(t, y)$ , т.е. сетка содержала  $N_T = 5000$  точек разбиения отрезка  $[0, T]$  и  $N_y = 4000$  точек разбиения отрезка  $[y, \bar{y}]$ .

Наконец, последний третий метод расчета на основании стохастического представления коэффициентов функции Беллмана дал приближенное решение  $\beta^{(MK)}(t, y)$ . Использовалась та же сетка  $\{t_k, y_i\}$  с  $\delta_t = 0.001$  и  $\delta_y = 0.01$ , что и в итоговом расчете методом конечных разностей, и выборка из  $N \cdot N_T \cdot N_y = 2 \times 10^{10}$  фрагментов траекторий  $y(t)$ , аппроксимирующих (4.5) согласно явному методу Эйлера.

Результаты расчетов иллюстрируют рис. 2–4, на которых приведены примеры траекторий для выхода и управлений  $z^*(t), u^*(t), z^{(prog)}(t), u^{(prog)}(t)$  и  $z^{(0)}(t), u^{(0)}(t)$ .

Рисунок 4 демонстрирует ожидаемый проигрыш неуправляемой системы  $z^{(0)}(t), u^{(0)}(t)$  и хорошее качество программного управления  $u^{(prog)}(t)$ . Последнее обстоятельство, конечно, объясняется исключительно выбранными параметрами рассматриваемого модельного примера.

Остается сравнить результаты трех методов аппроксимации оптимального управления  $u^*$ . Лучшей иллюстрацией является сравнение аппроксимаций  $\beta(t, y)$ , т.е. поверхностей  $\beta^{(ref)}(t, y), \beta^{(net)}(t, y), \beta^{(MK)}(t, y)$ . Выполненные расчеты иллюстрируются на рис. 5–7, представляющих эти поверхности в точках  $t = 0, 1, 2, 3, 4, 4.1, 4.2, 4.3, 4.4, 4.5, 4.6, 4.7, 4.8, 4.9, 4.95, 5.0$  (точки выбраны



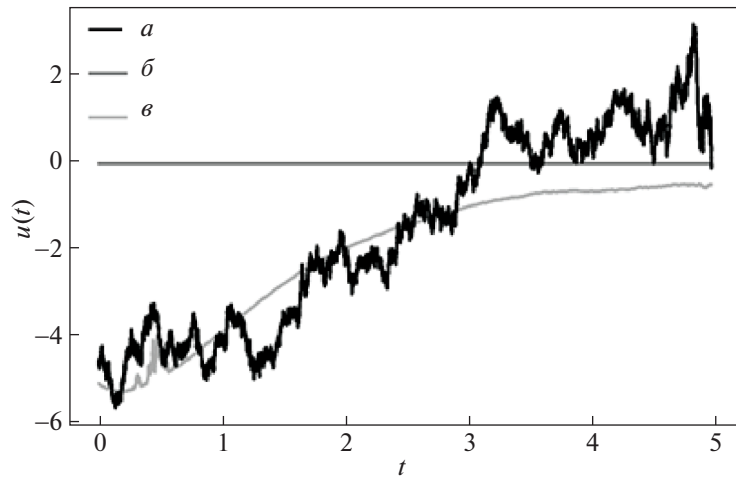


Рис. 3. Выборочные траектории:  $a - u^*(t)$ ,  $b - u^{(prog)}(t)$ ,  $v - u^{(0)}(t)$

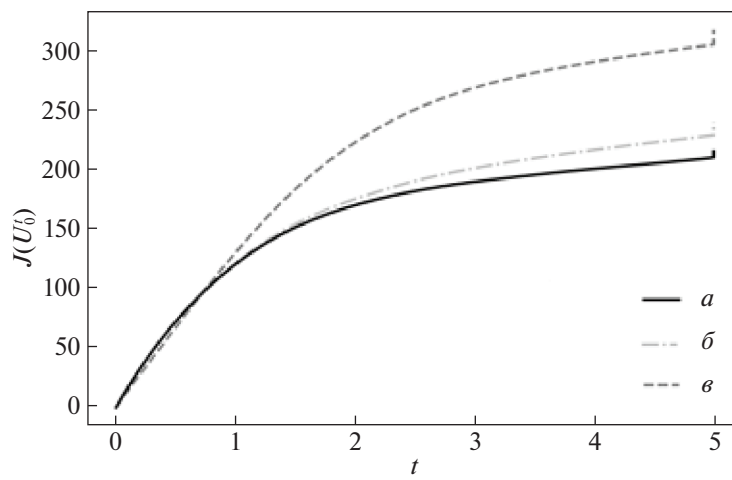


Рис. 4. Динамика целевого функционала  $a - J((U^*)_0^t)$ ,  $b - J((U^{(prog)})_0^t)$ ,  $v - J((U^{(0)})_0^t)$

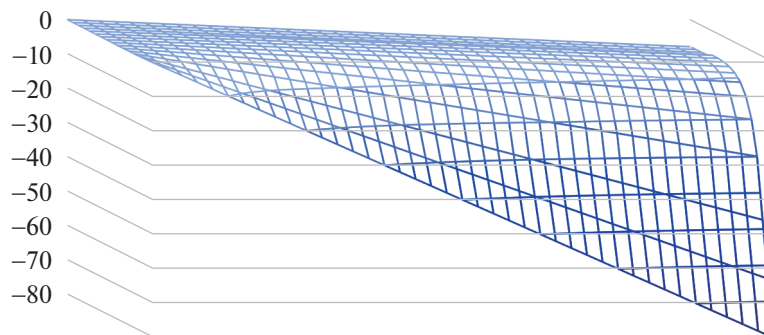


Рис. 5. Поверхность  $\beta^{(ref)}(t, y)$

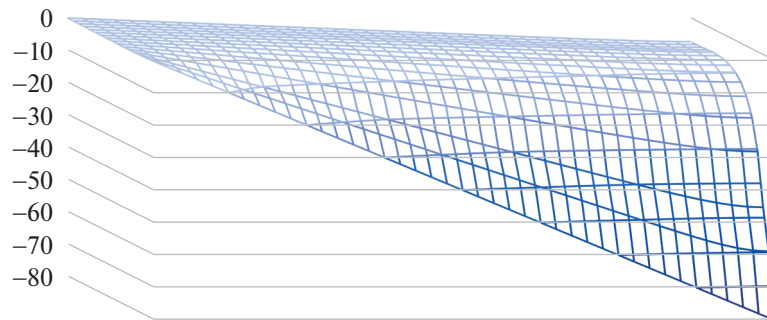


Рис. 6. Поверхность  $\beta^{(net)}(t, y)$

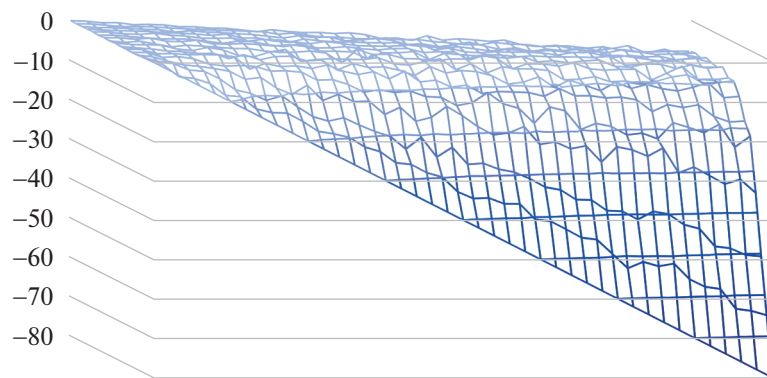


Рис. 7. Поверхность  $\beta^{(MK)}(t, y)$

там, где изменения поверхности более выражены и иллюстрации более выразительны) и целочисленных точках отрезка  $y \in [0, 30]$ .

Как уже отмечалось, визуально оценить данные на этих рисунках затруднительно, но качественно можно отметить, что, с одной стороны, поверхности похожи, с другой – поверхность  $\beta^{(MK)}(t, y)$  отличается некоторым “дрожанием” в сравнении с гладкими  $\beta^{(ref)}(t, y)$  и  $\beta^{(net)}(t, y)$ , что является очевидным следствием аппроксимации значений  $\beta(t, y)$  статистическими оценками по моделируемым данным.

Несколько более информативны графические материалы, приведенные на рис. 8–10. Это сечения поверхностей  $\beta^{(ref)}(t, y)$ ,  $\beta^{(net)}(t, y)$  и  $\beta^{(MK)}(t, y)$  в тех же точках  $t$ . Причем в точном решении, представленном  $\beta^{(ref)}(t, y)$ , это прямые линии. Также прямой  $\beta(T, y) = -2y$  отвечает терминальное условие.

Числовые значения, иллюстрирующие отклонения  $|\beta^{(net)}(t, y) - \beta^{(ref)}(t, y)|$  и  $|\beta^{(MK)}(t, y) - \beta^{(ref)}(t, y)|$  в некоторых точках  $t, y$ , представлены в табл. 1 и 2. В табл. 3 приведены значения  $\beta^{(ref)}(t, y)$ , чтобы можно было оценить относительные отклонения. По всей совокупности полученных данных можно сказать, что отклонения  $\beta^{(MK)}(t, y)$  от эталонной поверхности  $\beta^{(ref)}(t, y)$  несколько (2–3%) меньше в большинстве точек, чем отклонения  $\beta^{(net)}(t, y)$ . Но поверхность, рассчитанная методом конечных разностей, и ее отклонения от эталона носят более регулярный характер. Точнее, в некоторых узлах относительные отклонения  $|\beta^{(MK)}(t, y) - \beta^{(ref)}(t, y)|$  составляют до 10% и в целом, как и самой поверхности  $\beta^{(MK)}(t, y)$ , им свойственно некоторое “дрожание”, характерное для оценок Монте-Карло. Объективную оценку проводимого сравнения, конечно, дает значение

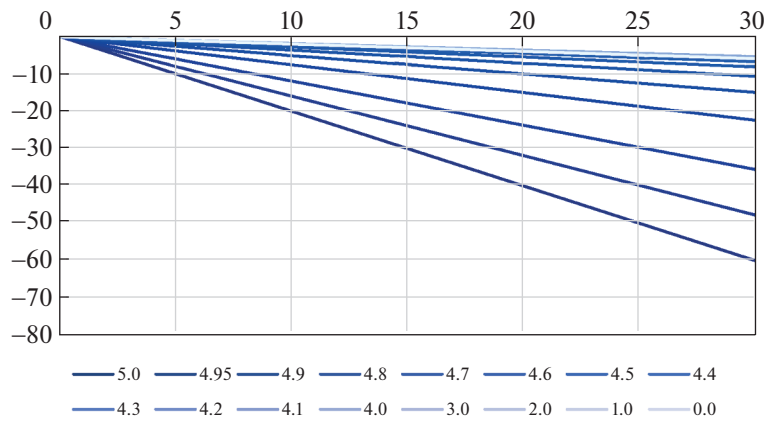


Рис. 8. Сечения поверхности  $\beta^{(ref)}(t, y)$

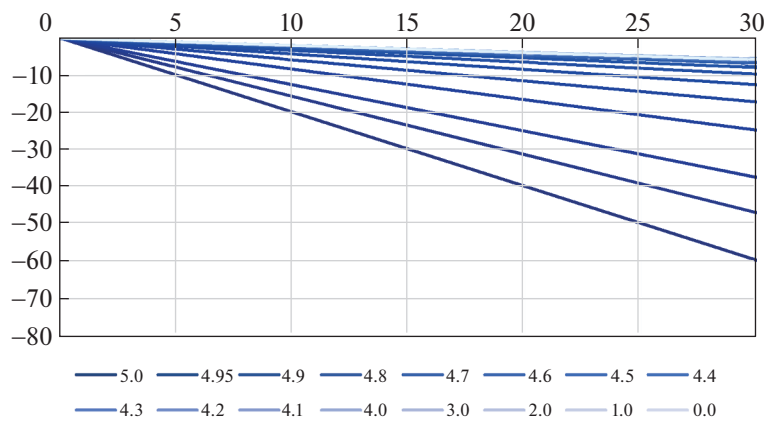


Рис. 9. Сечения поверхности  $\beta^{(net)}(t, y)$

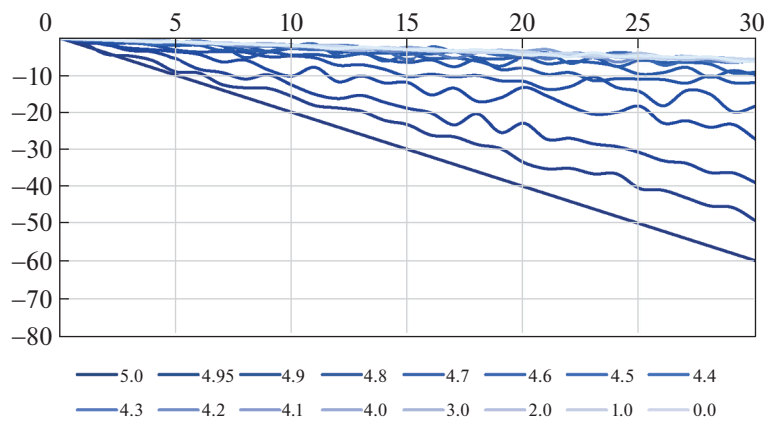


Рис. 10. Сечения поверхности  $\beta^{(MK)}(t, y)$

целевого функционала (6.3), доставляемое управлениями  $u^{(ref)}$ ,  $u^{(net)}$  и  $u^{(MK)}$ . Эти данные приведены в табл. 4 и подтвердили не только высокую точность расчетов, но и превосходство метода аппроксимации, основанного на стохастическом представлении коэффициентов функции Беллмана. И хотя последний вывод сделан в отношении цели решаемой задачи управления, т.е.

Таблица 1. Отклонения  $|\beta^{(net)}(t, y) - \beta^{(ref)}(t, y)|$ 

$t, y$	0.0	3.0	5.0	10.0	15.0	25.0
4.95	0.0233	0.0404	0.0829	0.1892	0.2955	0.5080
4.9	0.0196	0.2237	0.3599	0.7001	1.0404	1.7208
4.8	0.0329	0.2804	0.4454	0.8580	1.2706	2.0936
4.7	0.0409	0.2735	0.4285	0.8162	1.2038	1.9742
4.6	0.0448	0.2429	0.3749	0.7051	1.0353	1.6890
4.5	0.0457	0.2056	0.3122	0.5788	0.8453	1.3713
4.4	0.0445	0.1691	0.2521	0.4598	0.6673	1.0759
4.3	0.0420	0.1364	0.1993	0.3566	0.5138	0.8225
4.2	0.0386	0.1084	0.1550	0.2713	0.3875	0.6151
4.1	0.0349	0.0853	0.1190	0.2030	0.2870	0.4508
4.0	0.0310	0.0666	0.0903	0.1496	0.2089	0.3237
3.0	0.0061	0.0049	0.0042	0.0022	0.0002	0.0060
2.0	0.0009	0.0005	0.0003	0.0003	0.0010	0.0046
1.0	0.0001	0.0001	0.0001	0.0000	0.0001	0.0027
0.0	0.0000	0.0000	0.0000	0.0000	0.0001	0.0025

Таблица 2. Отклонения  $|\beta^{(MK)}(t, y) - \beta^{(ref)}(t, y)|$ 

$t, y$	0.0	3.0	5.0	10.0	15.0	25.0
4.95	0.0285	0.0105	1.0764	0.2761	0.5671	0.6918
4.9	0.0089	0.2988	0.4090	0.7194	1.0914	1.1994
4.8	0.1142	0.3755	0.1529	2.6940	0.7033	0.3248
4.7	0.1229	0.2741	0.6190	0.1041	2.8318	1.9781
4.6	0.0350	0.1291	0.3174	0.9699	1.2111	2.1372
4.5	0.0879	0.4242	0.5385	0.0298	0.4110	2.8989
4.4	0.0574	0.1915	0.7325	0.0092	0.7438	0.4403
4.3	0.1234	0.3271	0.2571	0.7245	2.7593	0.3966
4.2	0.0195	0.4084	0.1320	0.0601	0.7345	0.2707
4.1	0.1129	0.2158	0.1640	0.0178	0.6478	0.6370
4.0	0.0300	0.0933	0.0546	0.5945	0.2144	0.9019
3.0	0.0243	0.0657	0.0551	0.3454	0.0940	0.1782
2.0	0.0468	0.0200	0.1190	0.0636	0.1032	0.8721
1.0	0.0626	0.1975	0.1085	0.1386	0.0987	0.4386
0.0	0.0036	0.0828	0.0301	0.0288	0.0860	0.3083

Таблица 3. Поверхность  $\beta^{(ref)}(t, y)$ 

$t, y$	0.0	3.0	5.0	10.0	15.0	25.0
4.95	-0.0592	-4.8340	-8.0172	-15.9752	-23.9332	-39.8493
4.9	-0.1183	-3.6679	-6.0344	-11.9504	-17.8665	-29.6985
4.8	-0.1658	-2.3944	-3.8802	-7.5947	-11.3091	-18.7380
4.7	-0.1805	-1.6625	-2.6506	-5.1208	-7.5909	-12.5312
4.6	-0.1798	-1.2315	-1.9326	-3.6855	-5.4384	-8.9442
4.5	-0.1723	-0.9759	-1.5116	-2.8510	-4.1903	-6.8689
4.4	-0.1625	-0.8256	-1.2676	-2.3727	-3.4778	-5.6880
4.3	-0.1525	-0.7392	-1.1302	-2.1079	-3.0856	-5.0411
4.2	-0.1435	-0.6917	-1.0572	-1.9709	-2.8845	-4.7119
4.1	-0.1358	-0.6678	-1.0224	-1.9090	-2.7957	-4.5690
4.0	-0.1295	-0.6577	-1.0099	-1.8904	-2.7708	-4.5317
3.0	-0.1118	-0.6910	-1.0772	-2.0425	-3.0079	-4.9386
2.0	-0.1127	-0.7006	-1.0925	-2.0723	-3.0520	-5.0116
1.0	-0.1132	-0.7016	-1.0939	-2.0746	-3.0553	-5.0167
0.0	-0.1132	-0.7017	-1.0940	-2.0748	-3.0555	-5.0170

Таблица 4. Динамика функционалов  $J((U^{(ref)})_0^T)$ ,  $J((U^{(net)})_0^T)$ ,  $J((U^{(MK)})_0^T)$ 

$t$	$J((U^{(ref)})_0^t)$	$J((U^{(net)})_0^t)$	$J((U^{(MK)})_0^t)$	$\frac{ J((U^{(ref)})_0^t) - J((U^{(net)})_0^t) }{J((U^{(ref)})_0^t)} 100\%$	$\frac{ J((U^{(ref)})_0^t) - J((U^{(MK)})_0^t) }{J((U^{(ref)})_0^t)} 100\%$
1	112.7	120.6	116.0	7.0	2.9
2	157.4	170.3	161.0	8.2	2.3
3	174.6	190.0	179.9	8.8	3.1
4	186.4	200.9	192.1	7.8	3.1
5	202.1	215.7	208.6	6.7	3.2

качества аппроксимации управления  $u^*$ , данные в табл. 4 можно интерпретировать и как косвенное свидетельство о более высокой точности аппроксимации  $\beta(t, y)$ .

**Заключение.** Представленные в работе результаты в полном объеме описывают решение задачи управления линейным выходом векторного диффузионного процесса Ито по квадратичному критерию качества. Оптимальное решение – управление с полной обратной связью – задано замкнутой системой уравнений, приближенное решение которых может быть получено известными численными методами решения обыкновенных дифференциальных уравнений и уравнений в частных производных параболического типа. Свойства оптимального управления описаны стохастическим представлением коэффициентов найденной функции Беллмана на основе формулы Фейнмана–Каца. Это стохастическое представление позволило реализовать альтернативный метод приближенного решения задачи, базирующийся на имитационном компьютерном моделировании оптимизируемой системы. Численный пример показал работоспособность обоих методов аппроксимации и небольшое преимущество метода, основанного на стохастическом представлении коэффициентов функции Беллмана.

Для дальнейшего изучения важны следующие два направления. Во-первых, интересна задача с неполной информацией, т.е. постановка, в которой диффузионный процесс (2.1) предполагается ненаблюдаемым непосредственно, а доступным только косвенным наблюдениям (2.2). В этом направлении решение в общем виде получить не удастся. Причиной являются трудности в решении возникающей в такой постановке вспомогательной задачи оценивания – оптимальной фильтрации состояния  $y(t)$  по наблюдениям  $z(t)$ . Субоптимальная фильтрация позволяет

предлагать приближенные решения, практически пригодные, но не очень содержательные для теоретического изучения [25].

Второе перспективное направление развития исследованной задачи – это обобщение модели диффузии на случай смешанных непрерывно-дискретных возмущений. Это важное обобщение сделает постановку более практически содержательной. Так в использованном здесь модельном примере обобщается и становится более соответствующей потребностям практики модель РТТ.

## СПИСОК ЛИТЕРАТУРЫ

1. *Kushner H.J., Dupuis P.G.* Numerical Methods for Stochastic Control Problems in Continuous Time. N.Y.: Springer-Verlag, 2001.
2. *Bertsekas D.P.* Dynamic Programming and Optimal Control. Cambridge: Athena Scientific, 2013.
3. *Флеминг У., Ришел Р.* Оптимальное управление детерминированными и стохастическими системами / Пер. с англ. М.: Мир, 1978.
4. *Athans M.* The Role and Use of the Stochastic Linear-Quadratic-Gaussian Problem in Control System Design // IEEE Trans. Automat. Control. 1971. V. 16. № 6. P. 529–552.
5. *Wu Z.* Forward-backward Stochastic Differential Equations, Linear Quadratic Stochastic Optimal Control and Nonzero Sum Differential Games // J. Syst. Sci. Complexity. 2005. V. 18. № 2. P. 179–192.
6. *Chen B.S., Zhang W.* Stochastic H2/H1 Control with State-dependent Noise // IEEE Trans. Automat. Control. 2004. V. 49. № 1. P. 45–56.
7. *Bohacek S.* A Stochastic Model of TCP and Fair Video Transmission // Proc. IEEE INFOCOM. 2003. P. 1134–1144.
8. *Миллер Б.М., Авраченко К.Е., Степанян К.В., Миллер Г.Б.* Задача оптимального стохастического управления потоком данных по неполной информации // Пробл. передачи информ. 2005. Т. 41. № 2. С. 89–110.
9. *Боровков А.А.* Асимптотические методы в теории массового обслуживания. М.: Наука, 1980.
10. *Босов А.В.* Обобщенная задача распределения ресурсов программной системы // Информатика и ее применения. 2014. Т. 8. Вып. 2. С. 39–47.
11. *Босов А.В.* Управление линейным выходом дискретной стохастической системы по квадратичному критерию // Изв. РАН. ТиСУ. 2016. № 3. С. 19–35.
12. *Athans M., Falb P.L.* Optimal Control: An Introduction to the Theory and Its Applications. N.Y., Sydney: McGraw-Hill, 1966.
13. *Поляк Б.Т., Топунов М.В.* Подавление ограниченных внешних возмущений: управление по выходу // АиТ. 2008. № 5. С. 72–90.
14. *Баландин Д.В., Коган М.М.* Минимаксный подход к синтезу оптимального управления при неопределенных начальных условиях // АиТ. 2009. № 11. С. 3–12.
15. *Мирошник И.В., Никифоров В.О., Фрадков А.Л.* Нелинейное и адаптивное управление сложными динамическими системами. СПб.: Наука, 2000.
16. *Gihman I.I., Skorohod A.V.* The Theory of Stochastic Processes III. N.Y.: Springer-Verlag, 2012.
17. *Øksendal B.* Stochastic Differential Equations. An Introduction with Applications. N.Y.: Springer-Verlag, 2003.
18. *Fahim A., Touzi N., Warin, X.* A Probabilistic Numerical Method for Fully Nonlinear Parabolic Pdes // Annals of Applied Probability. 2011. V. 21. № 4. P. 1322–1364.
19. *Девис М.Х.А.* Линейное оценивание и стохастическое управление / Пер. с англ. М.: Наука, 1984.
20. *Bhatia R.* Matrix Analysis. Т. 169. N.Y.: Springer, 1997.
21. *Босов А.В., Стефанович А.И.* Управление выходом стохастической дифференциальной системы по квадратичному критерию. IV. Альтернативное численное решение // Информатика и ее применения. 2020. Т. 14. Вып. 1. С. 24–30.
22. *Bohacek S., Rozovskii B.* A Diffusion Model of Roundtrip Time // Computational Statistics & Data Analysis. 2004. V. 45. Iss. 1. P. 25–50.
23. *Cox J.C., Ingersoll J.E., Ross S.A.* A Theory of the Term Structure of Interest Rates // Econometrica. 1985. V. 53. Iss. 2. P. 385–407.
24. *Саульев В.К.* Интегрирование уравнений параболического типа методом сеток. М.: Физматлит, 1960.
25. *Босов А.В.* Применение условно-оптимального фильтра для синтеза субоптимального управления в задаче оптимизации выхода нелинейной дифференциальной стохастической системы // АиТ. 2020. № 11. С. 34–47.