

УПРАВЛЕНИЕ В СТОХАСТИЧЕСКИХ СИСТЕМАХ И В УСЛОВИЯХ НЕОПРЕДЕЛЕННОСТИ

УДК 519.857, 519.21

СТАБИЛИЗАЦИЯ ЛИНЕЙНОГО УПРАВЛЯЕМОГО ВЫХОДА АВТОНОМНОЙ СТОХАСТИЧЕСКОЙ ДИФФЕРЕНЦИАЛЬНОЙ СИСТЕМЫ НА БЕСКОНЕЧНОМ ГОРИЗОНТЕ¹

© 2024 г. А. В. Босов^a, *

^aФИЦ ИУ РАН, Москва, Россия

*e-mail: ABosov@frcsc.ru

Поступила в редакцию 10.04.2004 г.

После доработки 18.05.2024 г.

Принята к публикации 15.07.2024 г.

Рассматривается задача управления линейным выходом автономной нелинейной стохастической дифференциальной системы. Бесконечный горизонт и квадратичный функционал позволяют интерпретировать цель управления как стабилизацию выхода около положения, определяемого состоянием, которое описывается нелинейным стохастическим дифференциальным уравнением. Решение получено для двух вариантов модели — с точными измерениями и в предположении, что линейный выход представляет собой косвенные наблюдения за состоянием. В случае косвенных наблюдений в качестве модели состояния используется непрерывная цепь Маркова, что позволяет разделить задачи управления и фильтрации и применить фильтр Вонэма. В обоих вариантах достаточные условия существования оптимального решения состоят из типовых для линейных систем требований, обеспечивающих существование предельного решения уравнения Риккати. Дополнительные требования из-за нелинейных элементов — эргодичность нелинейной динамики и существование предела в формуле Фейнмана-Каца для коэффициентов нелинейной части управления. Приведены и проанализированы результаты численного эксперимента.

Ключевые слова: стохастическая дифференциальная система; оптимальное управление; стохастическая фильтрация; фильтр Вонэма; принцип разделения; уравнение Риккати

DOI: 10.31857/S0002338824050042, EDN: TTKBCP

STABILIZATION OF THE LINEAR CONTROLLED OUTPUT OF AN AUTONOMOUS STOCHASTIC DIFFERENTIAL SYSTEM ON AN INFINITE HORIZON

A. V. Bosov^a, *

^aFederal Research Center “Computer Science and Control” of the Russian Academy of Sciences,

44/2 Vavilova Str., 119333 Moscow, Russia

*e-mail: ABosov@frcsc.ru

The control problem of the linear output of an autonomous nonlinear stochastic differential system is considered. The infinite horizon and the quadratic objective make it possible to interpret the control goal as stabilization of the output near the position determined by the state, which is described by a nonlinear stochastic differential equation. The solution is obtained for two variants of the model: with accurate measurements and under the assumption that the linear output represents indirect observations of the state. In the case of indirect observations, a continuous Markov chain is used as a state model, which makes it possible to separate the control and filtering tasks and apply the Wonham filter. In both variants, sufficient conditions for the existence of an optimal solution consist of typical requirements for linear systems that ensure the existence of a limiting solution to the Riccati equation. Additional requirements due to nonlinear elements are the ergodicity of nonlinear dynamics and the existence of a limit in the Feynman-Katz formula for the coefficients of the nonlinear part of the control. The results of the numerical experiment are presented and analyzed.

¹ Работа выполнялась с помощью инфраструктуры Центра коллективного пользования “Высокопроизводительные вычисления и большие данные” (ЦКП “Информатика” ФИЦ ИУ РАН, Москва).

Keywords: stochastic differential system; optimal control; stochastic filtering; Wonham filter; separation principle; Riccati equation

Введение. Классический квадратичный критерий управления динамической системой допускает разные прикладные интерпретации, среди которых важное место занимает стабилизация состояния системы в некотором положении [1–3]. В свою очередь, задача стабилизации может иметь разный смысл: приведение системы в устойчивое «нулевое» состояние [1], движение выходной переменной по заданной траектории [4] или отслеживание меняющегося состояния цели по косвенным наблюдениям за ней [5]. В каждой такой задаче самостоятельного рассмотрения заслуживает постановка с моделью автономной динамической системы и бесконечным горизонтом управления [6, 7]. Действительно, когда нет оснований для определения времени завершения управления, естественным будет считать это время бесконечным. На систему управления в этом случае приходится накладывать довольно жесткие ограничения. Помимо формальной автономности, т.е. независимости от переменной времени всех функций как в модели состояния, так и в модели выхода и/или наблюдений и в целевом функционале, требуется устойчивость процессов, описывающих состояние, выход и косвенные наблюдения, и стабилизируемость управляемых переменных. Эти свойства хорошо иллюстрирует классическая линейноквадратичная гауссовская (LQG) задача управления с бесконечным горизонтом, в которой устойчивость и стабилизируемость означают конечность целевого функционала и принципиальную возможность управления выходной переменной [7]. Еще одно важное положение LQG-задачи с бесконечным временем — ограничение класса допустимых управлений линейными функциями выхода. Это вполне естественное предположение является следствием решения соответствующей задачи с конечным горизонтом, включая линейность оптимального управления, разделение и двойственность задач управления и фильтрации в постановке с косвенными наблюдениями, свойства фильтра Калмана-Бьюси [8] и, наконец, принципиальную роль уравнения Риккати, к исследованию свойств которого в конечном итоге сводится задача управления автономной линейной системой [6]. Оптимальное автономное управление получается как предельный случай оптимального управления в соответствующей задаче с конечным горизонтом.

Рассматриваемая в статье задача управления линейным выходом неуправляемой стохастической дифференциальной системы на бесконечном горизонте имеет и сходства, и отличия от задачи LQG. Так, используется квадратичный целевой функционал и линейный выход, дающий в оптимальном управлении линейное слагаемое. При этом есть нелинейная неуправляемая динамика, из-за которой в оптимальном управлении появляется нелинейная часть. Наконец, детально исследовано решение задачи с конечным горизонтом [9], содержащее в том числе описание нелинейной части управления, которое представляет собой параболическое дифференциальное уравнение в частных производных. Следовательно, имеются все исходные данные для того, чтобы получить решение соответствующей задачи с бесконечным горизонтом в предельной форме оптимального решения с конечным временем.

Первой в статье рассматривается постановка с полной информацией. Модель состояния в форме нелинейного уравнения Ито дает, как уже упоминалось, нелинейное слагаемое в оптимальном управлении, поэтому оптимальное автономное управление также остается нелинейным. При этом предельный переход от конечного к бесконечному горизонту управления вместо системы параболических уравнений приведет к обыкновенным дифференциальным уравнениям, хотя и нелинейным.

Вариант с косвенными наблюдениями, рассмотренный вторым, носит более частный вид. На первый план в этом варианте выходит задача нелинейной фильтрации, что исключает использование фильтра Калмана-Бьюси. Применение вместо линейных общих уравнений фильтрации на основе обновляющих процессов [10] в аналогичной задаче прекрасно продемонстрировано в [11]. Но этот очень важный для данной области результат не оставляет шансов для практической реализации. При этом автономный характер задачи управления практически никак не облегчает трудности решения задачи фильтрации, так как реализовать оптимальный фильтр продолжает оставаться гораздо сложнее, чем оптимальное управление. В качестве приемлемого промежуточного варианта можно ограничиться моделью состояния в форме цепи Маркова и, следовательно, оптимальной оценкой, задаваемой фильтром Вонэма [12]. Поведение цепи Маркова на бесконечности прекрасно изучено и изложено, например, в [13], требование эргодичности хорошо согласуется с бесконечным горизонтом управления. Таким образом, имеющееся решение задачи в постановке с неполной информацией [14] будет дополнено вариантом с бесконечным горизонтом управления.

1. Постановка задачи. На каноническом вероятностном пространстве $(\Omega, \mathcal{F}, \mathcal{P}, \mathcal{F}_t)$ с бесконечным временем $t \in [0, \infty)$ рассмотрим автономную стохастическую динамическую систему с состоянием (входом) $y_t \in \mathbb{R}^{n_y}$, которое формирует управляемый линейный выход $z_t \in \mathbb{R}^{n_z}$:

$$dz_t = ay_t dt + bz_t dt + cu_t dt + \sigma dw_t, z_0 = Z, t \in [0, \infty), \quad (1.1)$$

где $w_t \in \mathbb{R}^{n_w}$ — стандартный векторный винеровский процесс; $Z \in \mathbb{R}^{n_z}$ — гауссовский случайный вектор с известными математическим ожиданием и ковариацией; управление $u_t \in \mathbb{R}^{n_u}$ — случайный процесс из класса допустимых управлений $\mathcal{U}_0^\infty$, определенного далее; $a \in \mathbb{R}^{n_z \times n_y}$, $b \in \mathbb{R}^{n_z \times n_z}$, $c \in \mathbb{R}^{n_z \times n_u}$, $\sigma \in \mathbb{R}^{n_z \times n_w}$ — заданные матрицы.

Для описания состояния y_t возможны две модели. Первая — модель системы с полной информацией, согласно которой без ошибок наблюдаются как значения выхода z_t , так и состояния y_t . В этой модели y_t определяется системой автономных нелинейных стохастических дифференциальных уравнений Ито:

$$dy_t = A(y_t)dt + \Sigma(y_t)dv_t, y_0 = Y, t \in [0, \infty), \quad (1.2)$$

где $v_t \in \mathbb{R}^{n_v}$ — стандартный векторный винеровский процесс; $Y \in \mathbb{R}^{n_y}$ — случайный вектор с конечным вторым моментом; v_t, w_t, Y, Z независимы в совокупности; векторная функция $A = A(y): \mathbb{R}^{n_y} \rightarrow \mathbb{R}^{n_y}$ и матричная функция $\Sigma = \Sigma(y): \mathbb{R}^{n_y} \rightarrow \mathbb{R}^{n_y \times n_v}$ удовлетворяют стандартным условиям Ито:

$$|A(y)| + |\Sigma(y)| \leq C_1(1 + |y|),$$

$$|A(y_1) - A(y_2)| + |\Sigma(y_1) - \Sigma(y_2)| \leq C_2|y_1 - y_2|,$$

$$C_1 = \text{const}, C_2 = \text{const}, y, y_1, y_2 \in \mathbb{R}^{n_y},$$

обеспечивающим существование единственного поттраекторного решения (1.2) на любом конечном интервале $t \in [0, T]$ [15] (здесь и далее $|\cdot|$ обозначает евклидову норму вектора или матрицы). Чтобы рассматривать решения (1.2) на интервале $[0, \infty)$, дополнительно предполагается, что процесс y_t является эргодическим [13].

Во второй модели состояние y_t не наблюдается напрямую, доступна только косвенная информация. Это значит, что выход z_t выполняет функцию наблюдений. В этой модели y_t определяется марковским скачкообразным процессом — цепью с конечным числом состояний и значениями во множестве $\{e_1, \dots, e_{n_y}\}$, состоящем из единичных координатных векторов в евклидовом пространстве $\mathbb{R}^{n_y}$. Поскольку цепь y_t эргодическая, матрица интенсивностей переходов Λ не зависит от времени, предельное распределение обозначается π_* , начальное распределение $y_0 = Y$, как и для диффузии (1.2). y_t, w_t, Y, Z являются независимыми в совокупности.

Если представить процесс y_t его мартингалным разложением [4]

$$dy_t = \Lambda^T y_t dt + dv_t^y, y_0 = Y, t \in [0, \infty), \quad (1.3)$$

где v_t^y — $\mathcal{F}_t$ -согласованный мартингал с квадратичной характеристикой

$$\langle v^y, v^y \rangle_t = \int_0^t (\text{diag}(\Lambda^T y_s) - \Lambda^T \text{diag}(y_s) - \text{diag}(y_s) \Lambda) ds,$$

то вторую модель состояния можно считать частным случаем первой модели (1.2) с той разницей, что вместо винеровского процесса v_t фигурирует мартингал v_t^y , снос $A(y_t)$ является линейной функцией $\Lambda^T y_t$, а диффузия $\Sigma(y_t)$ определяется квадратичной характеристикой $\langle v^y, v^y \rangle_t$. Здесь и далее T — операция транспонирования.

Принципиальным отличием моделей с уравнениями состояния (1.2) или (1.3) является предположение о доступности y_t прямым или косвенным наблюдениям. В первой модели значение y_t известно и может участвовать в формировании управления u_t . Во второй модели значение y_t можно только оценить, используя вектор выхода z_t в качестве косвенных наблюдений. Возникающая таким образом задача фильтрации для (1.3) имеет известное решение —

фильтр Вонэма [12], для практической (численной) реализации которого в рассматриваемой задаче есть эффективные инструменты [16].

С другой стороны, если не ограничиваться частным случаем модели состояния (1.3), а в постановке с косвенными наблюдениями оставить модель (1.2), то для решения возникающей вспомогательной задачи фильтрации можно воспользоваться общим уравнением оптимальной фильтрации в терминах обновляющих процессов [10], в рассматриваемом случае имеющем следующий вид:

$$d\hat{y}_t = \hat{A}_t dt + \hat{\Sigma}_t a^T (\sigma \sigma^T)^{-1} (dz_t - a\hat{y}_t dt - bz_t dt - cu_t dt), \hat{y}_0 = \mathcal{E}\{Y\}, \quad (1.4)$$

где $\hat{A}_t = \mathcal{E}\{A(y_t) | \mathcal{F}_t^z\}$, $\hat{\Sigma}_t = \mathcal{E}\{(y_t - \hat{y}_t)y_t^T | \mathcal{F}_t^z\}$. Здесь и далее обозначены $\mathcal{E}\{\cdot\}$ и $\mathcal{E}\{\cdot | \mathcal{F}\}$ – операторы математического и условного относительно σ -алгебры $\mathcal{F}$ математического ожидания, $(\mathcal{F}_t^z - \sigma$ -алгебра, порожденная процессом z_t , так что $\mathcal{F}_t^z \subseteq \mathcal{F}_t \subseteq \mathcal{F}$). Таким образом, даже линейность наблюдений (1.1) не делает общее уравнение оптимальной фильтрации существенно проще, так как уравнение (1.4) остается слишком сложными для практической реализации, поскольку для вычисления условных математических ожиданий $\hat{A}_t$ и $\hat{\Sigma}_t$ все равно требуются апостериорные распределения y_t относительно $\mathcal{F}_t^z$. Устраняется эта трудность при дальнейшем уточнении (1.4) для марковской цепи y_t , состояния которой образуют множество $\{e_1, \dots, e_{n_y}\}$. В этом случае удастся явно записать $\hat{A}_t$ и $\hat{\Sigma}_t$, что, как известно, приводит к фильтру Вонэма [4,12]:

$$d\hat{y}_t = \Lambda^T \hat{y}_t dt + (\text{diag}(\hat{y}_t) - \hat{y}_t \hat{y}_t^T) a^T (\sigma \sigma^T)^{-1} \times \\ \times (dz_t - a\hat{y}_t dt - bz_t dt - cu_t dt), \hat{y}_0 = \mathcal{E}\{Y\}. \quad (1.5)$$

Именно, структура (1.5) оценки $\hat{y}_t = \mathcal{E}\{y_t | \mathcal{F}_t^z\}$ и дополнительное предположение об эргодичности y_t обеспечивают далее решение задачи управления для второй рассматриваемой модели. Для удобства записи функциональную зависимость оценки $\hat{y}_t$ от наблюдений, определяемую уравнением (1.5), обозначим $\hat{y} = \hat{y}(z)$, так что можно записать $\hat{y}_t = \hat{y}(z_t)$.

Для постановки задачи управления определим уже упомянутый класс $\mathcal{U}_0^\infty$ допустимых управлений. В варианте задачи с полной информацией, т.е. с состоянием (1.2) и выходом (1.1), допустимое управление u_t предполагается $\mathcal{F}_t^{y,z}$ -измеримым $\mathcal{F}_t^{y,z} - \sigma$ -алгебра, порожденная y_t и z_t , $\mathcal{F}_t^{y,z} \subseteq \mathcal{F}_t \subseteq \mathcal{F}$ и $\mathcal{U}_0^\infty$ составляют автономные (не зависящие прямо от времени) управления с полной обратной связью. Другими словами, это функции вида $u_t = u(y, z) \in \mathbb{R}^{n_u}$, $y \in \mathbb{R}^{n_y}$, $z \in \mathbb{R}^{n_z}$, при условии, что соответствующая реализация $u_t = u(y_t, z_t)$ обеспечивает существование y_t, z_t . Поскольку состояние y_t от управления u_t не зависит, а выход z_t описывается линейным автономным уравнением с винеровским процессом, то данное формальное требование означает наличие у допустимых управлений конечного второго момента $\mathcal{E}\{|u_t|^2\} < \infty$, что обеспечивает существование решения (1.1) на любом конечном интервале $t \in [0, T]$. Условия для существования управления на интервале $[0, \infty)$ обсуждаются далее.

Для варианта задачи с косвенными наблюдениями, т.е. с состоянием (1.3) и наблюдениями (1.1), допустимое управление u_t предполагается $\mathcal{F}_t^z$ -измеримым и $\mathcal{U}_0^\infty$ составляют автономные управления с полной обратной связью по выходу, т.е. функции вида $u_t = u(z) \in \mathbb{R}^{n_u}$, $z \in \mathbb{R}^{n_z}$, в предположении, что соответствующая реализация $u_t = u(z_t)$ обеспечивает выполнение условий существования z_t . На самом деле, обеспечиваемое приведенным далее квадратичным функционалом качества разделение задач управления и фильтрации на конечном горизонте управления $t \in [0, T]$ [14], этот вариант задачи сводит к первому варианту, в котором роль уравнения состояния выполняет фильтр Вонэма (1.5). При этом обсуждаемые далее условия для существования управления на интервале $[0, \infty)$ оказываются даже проще, чем в задаче с первым вариантом модели состояния.

Качество управления $U_0^\infty = \{u_t, t \in [0, \infty)\} \in \mathcal{U}_0^\infty$ определяется функционалом:

$$J(U_0^\infty) = \lim_{T \rightarrow \infty} J(U_0^T), J(U_0^T) = \mathcal{E}\left\{\frac{1}{T} \int_0^T P y_t + Q z_t + R u_t^2 dt\right\}, \quad (1.6)$$

где $U_0^T = \{u_t, t \in [0, T]\}$; $P \in \mathbb{R}^{n_y \times n_y}$, $Q \in \mathbb{R}^{n_z \times n_z}$, $R \in \mathbb{R}^{n_u \times n_u}$, $S \in \mathbb{R}^{n_y \times n_z}$ ($S \geq 0$, $S = S'$) – заданные матрицы; весовая функция $\|x\|_S^2 = x^T S x$, единичной матрице $S = E$ соответствует ев-

клидова норма $\|x\|_E^2 = |x|^2$. Кроме того, предполагается выполненным обычное условие невырожденности $R^T S R > 0$.

Задача состоит, таким образом, в поиске $(U^*)_0^\infty = \{u_t^*, t \in [0, \infty)\} \in \mathcal{U}_0^\infty$ — допустимого управления, реализации которого $u_t^* = u^*(y_t, z_t^*)$ или $u_t^* = u^*(z_t^*), t \in [0, \infty)$ соответственно доставляют минимум квадратичному функционалу $J(U_0^\infty)$:

$$(U^*)_0^\infty = \operatorname{argmin}_{U_0^\infty \in \mathcal{U}_0^\infty} J(U_0^\infty). \quad (1.7)$$

Далее в обоих вариантах задачи через z_t^* обозначается решение (1.1), отвечающее оптимальному управлению u_t^* , и учитывается, что y_t от u_t не зависит.

Надо отметить, что физический смысл поставленной задачи, определенный в заголовке термином стабилизация, обеспечивается именно целевым функционалом (1.6). Такая форма записи квадратичного функционала нужна, прежде всего, чтобы формализовать цель управления, состоящую в формировании такого значения выходной переменной z_t , которое задается состоянием y_t . Типичная интерпретация, которую обеспечивает слагаемое $P y_t + Q z_t$, — это задача слежения за целью. Поскольку горизонт слежения бесконечный и характеристики положения цели не изменяются, отсюда получается стабилизация выхода. В варианте с цепью физическая интерпретация даже содержательнее. Так как цепь y_t после очередной смены состояния e_i на состояние e_j некоторое время пребывает в последнем, не изменяясь, то в рамках этого периода постоянства состояние e_j вносит в уравнение (1.1) систематическую (кусочно-постоянную) составляющую, которая отличается от величины e_i систематической составляющей, вносимой на предыдущем интервале постоянства. Аналогично в целевом функционале формируются ступенчатые слагаемые на интервалах постоянства y_t . Таким образом, (1.6) формализует цель управления, состоящую в стабилизации выхода z_t около скачкообразно изменяющихся кусочно-постоянных положений, задаваемых состоянием цепи y_t . По-другому эти положения можно назвать направлениями дрейфа, а цель управления — обеспечение дрейфа выхода системы z_t в заданном и периодически изменяющемся направлении.

Кроме того, следует обратить внимание на важное отличие предложенной постановки задачи управления от классической задачи LQG с бесконечным горизонтом [7], в которой класс допустимых управлений включает только линейные функции состояния и выхода или только наблюдений. В LQG такое предположение не является ограничением, так как в соответствующей задаче с конечным горизонтом оптимальное решение получилось линейным на классе всех управлений. В рассматриваемой задаче получается иначе, поскольку и состояние (1.2) или (1.3), и, главное, решение соответствующей задачи с конечным горизонтом [9] нелинейные. Далее формулируются условия, при которых искомое решение представляется в виде предельной формы оптимального управления с конечным горизонтом и остается в классе нелинейных, сохраняя допредельную структуру.

2. Оптимальное автономное управление (полная информация). Решение задачи (1.7) в постановке с полной информацией получается предельным переходом в соответствующей задаче с конечным горизонтом [9]. Для этого требуется выполнение ряда условий, объединенных в следующем утверждении.

Т е о р е м а 1. Решение u^* задачи (1.7) для системы (1.1), (1.2) на классе управлений $\mathcal{U}_0^\infty = \{u_t = u(y, z), y \in \mathbb{R}^{n_y}, z \in \mathbb{R}^{n_z}, t \in [0, \infty)\}$ имеет вид

$$u^*(y, z) = -\frac{1}{2} (R^T S R)^{-1} (c^T (2\alpha_* z + \beta_*(y)) + 2R^T S (P y + Q z)), \quad (2.1)$$

где $\alpha_* \in \mathbb{R}^{n_z \times n_z}$ — симметричная неотрицательно определенная матрица, удовлетворяющая матричному уравнению:

$$\begin{aligned} (M^\alpha)^T \alpha_* + \alpha_* M^\alpha + N^\alpha - \alpha_* c (R^T S R)^{-1} c^T \alpha_* &= 0, \\ M^\alpha &= b - c (R^T S R)^{-1} R^T S Q, \\ N^\alpha &= Q^T \left(S - S R (R^T S R)^{-1} R^T S \right) Q, \end{aligned} \quad (2.2)$$

и $\beta_* = \beta_*(y) = \left(\beta_*^{(1)}(y), \dots, \beta_*^{(n_z)}(y) \right)^T \in \mathbb{R}^{n_z}$ — вектор-функция, удовлетворяющая системе обыкновенных дифференциальных уравнений:

$$\begin{aligned} \frac{1}{2} \operatorname{tr} \left\{ \Sigma^T \frac{\partial^2 \beta_*^{(i)}}{\partial y^2} \Sigma \right\} + A^T \frac{\partial \beta_*^{(i)}}{\partial y} + \sum_{j=1}^{n_y} y^{(j)} [M_*^\beta]^{(ji)} + \sum_{j=1}^{n_z} \beta_*^{(j)} [N_*^\beta]^{(ji)} &= 0, \\ M_*^\beta &= 2 \left(\left(a^T - P^T S R (R^T S R)^{-1} c^T \right) \alpha_* + P^T \left(S - S R (R^T S R)^{-1} R^T S \right) Q \right), \\ N_*^\beta &= b - c \left(R^T S R \right)^{-1} R^T S Q - c \left(R^T S R \right)^{-1} c^T \alpha_*, \quad i = 1, \dots, n_z, \end{aligned} \quad (2.3)$$

если для параметров системы (1.1), (1.2) и целевого функционала (1.6) выполнены следующие условия:

- 1) матрица b устойчива (все собственные значения квадратной матрицы $b \in \mathbb{R}^{n_z \times n_z}$ имеют отрицательные действительные части);
- 2) матрицы (M^α, c) стабилизируемы (в терминах [7, разд. 5.4], пара (M^α, c) является стабилизирующей);
- 3) матрицы $\left((M^\alpha)^T, K \right)$, $K = Q^T S^{\frac{1}{2}} \left(E - S^{\frac{1}{2}} R (R^T S R)^{-1} R^T S^{\frac{1}{2}} \right)$, стабилизируемы (в терминах [7, разд. 5.4], пара (K^T, M^α) является выявляющей);
- 4) для любого $y \in \mathbb{R}^{n_y}$ существует и не зависит от t конечный предел:

$$\lim_{T \rightarrow \infty} \mathcal{E} \left\{ I_t^{-1} \int_t^T I_\tau^{-1} (M_\tau^\beta)^T y_\tau d\tau \middle| \mathcal{F}_t^y \right\}, \quad (2.4)$$

где $I_\tau^{-1} = \exp \left\{ \left(N_\tau^\beta \right)^T \tau \right\}$, y_τ — решение уравнения (1.2) с переменной времени $\tau \in [t, \infty)$ и начальным условием $y_t = y$.

Уравнения (2.3) записаны с использованием обозначений $y = \left(y^{(1)}, \dots, y^{(n_y)} \right)^T$ для элементов вектора y , $[A]^{(ji)}$ — для элемента j -й строки i -го столбца матрицы A , $\operatorname{tr} \{A\}$ — для следа матрицы A . Доказательство теоремы 1 приведено в Приложении.

Как и в случае LQG-управления, решение задачи с бесконечным временем базируется на решении соответствующей задачи с конечным горизонтом. Это управление $(U^\#)_0^T = \arg \min_{U_0^T \in \mathcal{U}_0^T} J(U_0^T)$ из класса допустимых управлений $\mathcal{U}_0^T = \left\{ u_t = u_t(y, z, T), y \in \mathbb{R}^{n_y}, z \in \mathbb{R}^{n_z}, t \in [0, T] \right\}$ получено в [9] и с учетом использованных обозначений может быть записано в виде

$$u_t^\#(y, z, T) = -\frac{1}{2} \left(R^T S R \right)^{-1} \left(c^T (2\alpha_t z + \beta_t) + 2R^T S (P y + Q z) \right), \quad (2.5)$$

где матричный коэффициент $\alpha_t = \alpha_t(T) \in \mathbb{R}^{n_z \times n_z}$ является решением задачи Коши для уравнения Риккати (см. Приложение), а векторный коэффициент $\beta_t = \beta_t(y, T) = \left(\beta_t^{(1)}(y, T), \dots, \beta_t^{(n_z)}(y, T) \right)^T \in \mathbb{R}^{n_z}$ — задачи Коши для системы параболических дифференциальных уравнений в частных производных (см. Приложение). Соответственно уравнения (2.2), (2.3) получают предельным переходом $\alpha_* = \lim_{T \rightarrow \infty} \alpha_t(T)$ и $\beta_*(y) = \lim_{T \rightarrow \infty} \beta_t(y, T)$, обеспечиваемым условиями теоремы.

Отметим, что как уже упоминалось выше, оптимальное автономное управление (2.1) остается нелинейным по переменной y как в допредельной постановке. Это главное отличие рассматриваемой задачи от классической задачи LQG. Здесь же надо указать и на сходство задач: слагаемые, формируемые переменной выхода z , входят в управление (2.1) линейно и определяются теми же соотношениями, что и в задаче LQG. Кроме того, конструктивными условиями существования (2.1) можно считать только традиционные требования к коэффициентам уравнения Риккати. Проверять последнее условие (2.4) без компьютерного моделирования

затруднительно. Зато если ограничиться только численными экспериментами, то можно применять как минимум два разных эффективных метода численного решения параболических уравнений, определяющих $\beta_t(y, T)$ [17, 18], а при наличии готовых расчетов анализ сходимости $\lim_{T \rightarrow \infty} \beta_t(y, T)$ труда не составляет. Подтверждающий это пример подробно обсуждался в [9], где роль состояния (1.2) исполняет модель эволюции процентных ставок Кокса–Ингерсона–Росса (cox–ingersoll–ross model) [19], использованная для описания показателя RTT (round-trip time) сетевого протокола TCP (transmission control protocol) [20], и поверхности $\beta_t(y, T)$ вычислены двумя имеющимися численными методами и аналитически.

3. Оптимальная стабилизация по косвенным наблюдениям. Решение задачи (1.7) для модели с косвенно наблюдаемой марковской цепью так же, как и в случае с полной информацией, получается предельным переходом в соответствующей задаче с конечным горизонтом [14], условия которого сформулированы в следующем утверждении.

Т е о р е м а 2. Решение u^* задачи (1.7) для системы (1.2), (1.3) на классе управлений $\mathcal{U}_0^\infty = \{u_t = u(z), z \in \mathbb{R}^{n_z}, t \in [0, \infty)\}$ имеет вид

$$u^*(z) = -\frac{1}{2}(R^T S R)^{-1} \left(c^T (2\alpha_* z + \beta_* \hat{y}(z)) + 2R^T S (P \hat{y}(z) + Qz) \right), \quad (3.1)$$

где $\hat{y}(z)$ задано (1.5), $\alpha_* \in \mathbb{R}^{n_z \times n_z}$ – симметричная неотрицательно определенная матрица, удовлетворяющая матричному уравнению (2.2), и $\beta_* \in \mathbb{R}^{n_z \times n_y}$ – прямоугольная матрица, являющаяся решением матричного уравнения:

$$\beta_* \Lambda^T + M_*^\beta - N_*^\beta \beta_* = 0, \quad (3.2)$$

где M_*^β и N_*^β определены в (2.3), если для параметров системы (1.1), (1.3) и целевого функционала (1.6) выполнены условия 1)–3) теоремы 1.

Доказательство теоремы 2 приведено в Приложении.

Как и в постановке с полной информацией решение задачи с бесконечным временем базируется на решении соответствующей задачи с конечным горизонтом. Это решение можно обозначить так же, как и выше, $(U^\#)^T_0 = \arg \min_{U_0^T \in \mathcal{U}_0^T} J(U_0^T)$, имея в виду класс допустимых управлений $\mathcal{U}_0^T = \{u_t = u_t(z, T), z \in \mathbb{R}^{n_z}, t \in [0, T]\}$. Оно получено в [14] и с учетом использованных обозначений может быть записано в виде:

$$u_t^\#(z, T) = -\frac{1}{2}(R^T S R)^{-1} \left(c^T (2\alpha_t z + \beta_t \hat{y}) + 2R^T S (P \hat{y} + Qz) \right). \quad (3.3)$$

Оценки $\hat{y}(z_t^\#)$ на оптимальных траекториях $z_t^\#$ задает фильтр Вонэма (1.5), матричный коэффициент $\alpha_t = \alpha_t(T)$ является решением той же задачи Коши для уравнения Риккати, что и в теореме 1 (см. Приложение), а матричный коэффициент $\beta_t = \beta_t(T) \in \mathbb{R}^{n_z \times n_y}$ – решением задачи Коши для системы дифференциальных уравнений в частных производных параболического типа (см. Приложение).

Отметим, что предельная функция $\beta_*(y)$ в первой модели и предельная матрица β_* во второй связаны равенством $\beta_*(y) = \beta_* y$. С учетом этого еще обратим внимание, что от условия 4) теоремы 1, обеспечивавшего существование предельной функции $\beta_*(y)$ в первой модели, здесь осталась только предельная матрица M_*^β , существование которой обеспечивает предел α_* . Так получается благодаря структуре фильтра Вонэма. Уравнение (1.5) имеет линейный снос и нелинейную диффузию, благодаря чему в оптимальном управлении второе слагаемое, формируемое состоянием y_t (а точнее, его оценкой $\hat{y}_t$), как и первое, формируемое выходом z_t , оказывается линейным. Такой частный случай в задаче управления линейным дифференциальным выходом общего вида рассмотрен в [9], но здесь снос $\Lambda^T \hat{y}_t$ не только линейный, но и аффинный, без свободного члена, что еще больше упрощает вид управления. Таким образом, нелинейность управлению u дает только оценка фильтрации $\hat{y}_t = \hat{y}(z_t)$. При этом если предположить, что цепь y_t доступна прямым наблюдениям, то оптимальное управление будет иметь тот же вид (3.3), где вместо $\hat{y}_t$ будет использоваться точное значение y_t . Такое свойство управления в [11] названо сильным принципом разделения.

4. Численный анализ предельного режима. Упомянутые выше расчеты с моделью Кокса—Ингерсона—Росса, выполненные в [9], хорошо иллюстрируют первую модель и результативность как оптимального автономного управления, так и приближенных методов, которые можно применять для его реализации. По этой причине численный эксперимент здесь выполнен для второй модели. Именно для иллюстрации того, как работает оптимальное автономное управление $u^*(z)$ (3.1) в сравнении с оптимальным неавтономным управлением $u_t^\#(z, T)$ (3.3) использована предложенная в [21] модель простого механического привода

$$\begin{aligned} dx_t &= v_t dt, \quad t \in (0, T], \\ dv_t &= ax_t dt + bv_t dt + cy_t dt + hu_t dt + \sqrt{g} dw_t. \end{aligned} \quad (4.1)$$

Буквенные обозначения выполнены полужирным шрифтом, чтобы исключить совпадения с использованными выше обозначениями и подчеркнуть физический смысл величин: v_t — скорость перемещения привода, x_t — его координатное положение, w_t — погрешность в задании ускорения и т.д. Не касается это только цепи y_t . Предполагается, что она имеет размерность $n_y = 3$, начальное распределение $Y = (1, 0, 0)^T$ и

$$\Lambda = \begin{pmatrix} -1/2 & 1/2 & 0 \\ 1/2 & -1 & 1/2 \\ 0 & 1/2 & -1/2 \end{pmatrix},$$

три состояния цепи — это три целевых положения привода. Начальные условия x_0 и v_0 — независимые гауссовские случайные величины с нулевым средним и дисперсиями $\sigma_x^2 = \sigma_v^2 = 1$. Остальные параметры: $a = -1$, $b = -1/2$, $h = 10$, $g = 0.01$ и $c = (c_1, c_2, c_3) = (-1, 0, 1)$. Система (4.1) устойчива при $b < 0$ и $b^2 + 4a = -3.75 < 0$, поскольку b и $b^2 + 4a$ — собственные числа матрицы системы $\begin{pmatrix} 0 & 1 \\ a & b \end{pmatrix}$. Обсуждаемые в [21] расчеты выполнены для $T = 10$ и целевого функционала:

$$J(U_0^T) = \mathcal{E} \left\{ \int_0^T (|Cy_t - x_t|^2 + R|u_t|^2) dt \right\}, \quad (4.2)$$

где

$$C = \left(-\frac{c_1}{a}, -\frac{c_2}{a}, -\frac{c_3}{a} \right) = (-1, 0, 1) = c, \quad R = 0.001.$$

Цель управления, минимизирующего (4.2), состоит в поддержании привода в положении x_t , задаваемом величиной Cy_t , т.е. в обеспечении дрейфа в направлении состояния цепи. Причем неуправляемый вариант системы (4.1) при $u_t = 0$ дрейфует в направлении текущего состояния цепи в силу своей устойчивости, но эффективность этого дрейфа в сравнении с управляемой системой невелика. Проиллюстрирован этот вариант системы рис. 1, где приведена одна реализация из смоделированного пучка 1000 траекторий положения привода x_t , сформированная тремя управлениями: 1) «нулевым» $u_t = u_t^{(0)} = 0$; 2) оптимальным $u_t = u_t^\# = u_t^\#(z_t^\#, T)$, определенным соотношением (3.3); 3) оптимальным автономным $u_t = u_t^* = u^*(z_t^*)$, определенным соотношением (3.1).

Рисунок иллюстрирует высокое качество управления, минимизирующего (4.2), и «обозначение» неуправляемой системой (4.1) передвижения привода в направлении дрейфа Cy_t . Главный вывод — это практически полное совпадение траекторий $x_t^\#$ и x_t^* . Небольшая разница видна только на самом конце траектории (показана на выноске), где собственно и происходит формирование коэффициентов α_* и β_* оптимального автономного управления из коэффициентов α_t и β_t . Переходной процесс, который хотелось проиллюстрировать, протекает очень быстро, что подтвердил расчет с начальным значением $x_0 = 100$. Фактические значения целевого функционала (4.2) для этого примера приведены ниже в таблице вместе с остальными примерами.

Добиться видимой разницы в результативности автономного u_t и оптимального $u_t^\#$ управлений оказалось довольно сложно. Дело в том, что даже сильное удаление системы (4.1) от предельного состояния, растягивающее переходной процесс при отсутствии управления, что и могло бы стать предметом для анализа, при наличии любого из двух управлений u_t или $u_t^\#$ не сильно сказывается на общем результате. Любое из управлений, минимизирующее (4.2), очень быстро стабилизирует положение x_t , уменьшая время перехода системы в предельный режим.

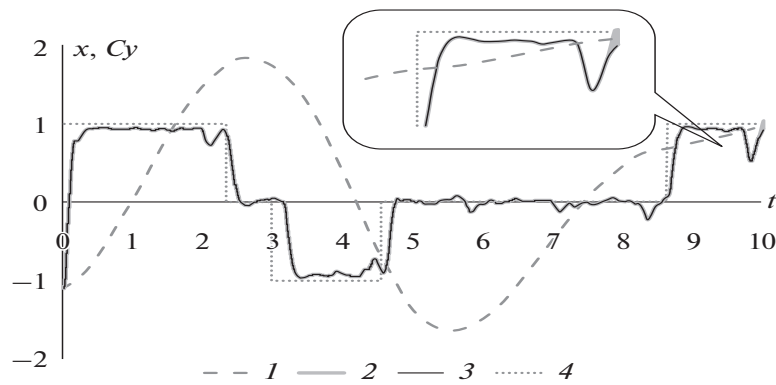


Рис. 1. Типовые траектории положения привода x_t , модель $T = 10$ и $R = 0.001$; управления: 1 – $u_t^{(0)}$, 2 – $u_t^\#$, 3 – u_t^* , 4 – направление дрейфа Cy_t .

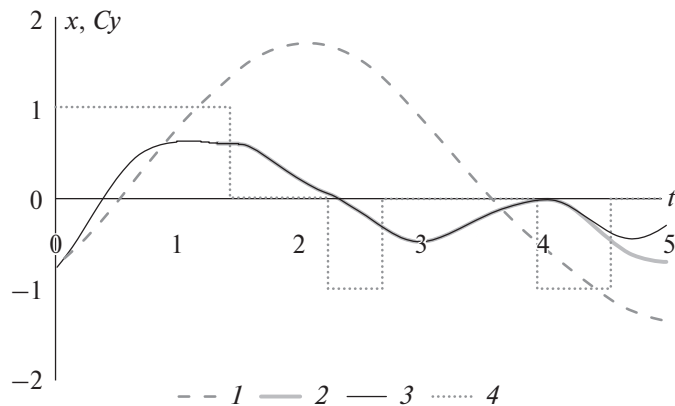


Рис. 2. Типовые траектории положения привода x_t , модель $T = 5$ и $R = 1$; управления: 1 – $u_t^{(0)}$, 2 – $u_t^\#$, 3 – u_t^* , 4 – направление дрейфа Cy_t , 5 – направление дрейфа Cy_t .

Немного уменьшить этот эффект удастся только за счет увеличения цены величины управляющего воздействия R и уменьшения времени стабилизации. Два следующих примера рассчитаны для $R = 1$ и двух значений $T = 5$ и $T = 1$. Иллюстрируют эти расчеты рис. 2 и 3 соответственно, на которых приведены по одной реализации из множества смоделированных траекторий.

Таблица 1. Оценки качества управлений $J(\mathcal{U}_0^T)$ по результатам моделирования

Модель	$u_t^{(0)}$	$u_t^\# = u_t^\#(z_t^\#, T)$	$u_t^* = u^*(z_t^*)$
$T = 10, R = 0.001$	8.418	1.093	1.094
$T = 5, R = 1$	5.148	2.128	2.160
$T = 1, R = 1$	1.600	0.926	0.987
$T = 5, R = 0.001, x_0 = 100$	11551.04	0.630	0.631

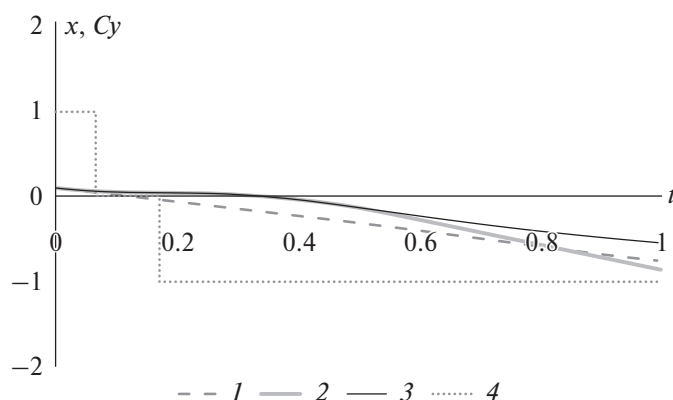


Рис. 3. Типовые траектории положения привода x_r , модель $T = 1$ и $R = 1$; управления: 1 – $u_t^{(0)}$, 2 – $u_t^\#$, 3 – u_t^* , 4 – направление дрейфа Cy_r

Формальные выводы к приведенным иллюстрациям дает следующая таблица, в которой указаны значения (4.2), вычисленные приближенно путем компьютерного моделирования 1000 траекторий системы (4.1) с шагом $\Delta t = 0.001$ и всех трех управлений $u_t^{(0)}$, $u_t^\#$ и u_t^* .

Таким образом, даже в тех примерах, где разница по критерию между автономным и оптимальным управлением есть, величина ее невелика (вторая-третья значащая цифра). С учетом сравнительной простоты расчета u_t по отношению к $u_t^\#$ и его эффективности как средства ускорения переходного процесса и достижения предельного режима, преимущества автономного управления представляются очевидными.

Заключение. Решенная в статье задача — это логичное развитие задачи оптимального управления линейным выходом стохастической дифференциальной системы по квадратичному критерию на случай бесконечного времени. Основные представленные результаты в полной мере эксплуатируют классические методы исследования уравнения Риккати, давно разработанные для классической LQG-задачи управления и автономной линейной динамической системы управления/наблюдения. Однако в рассмотренных постановках эта теория обеспечивает существование только «части» искомого оптимального автономного управления, которая отвечает за линейные элементы модели. При этом остаются, как и в допредельном варианте с конечным горизонтом, нелинейные элементы. Полезным с практической точки зрения оказалось, что параметры этих нелинейных слагаемых в оптимальном автономном управлении вычислить существенно проще, чем в случае конечного горизонта, поскольку автономная система дала возможность исключить из решения сложные в реализации схемы решения параболических уравнений, заменив их на обыкновенные дифференциальные. Кроме того, бесконечное время позволило сформулировать прикладную цель полученного управления, состоящую в стабилизации выхода системы в положениях, задаваемых ее состоянием. На конечном горизонте такая интерпретация задачи выглядела несколько хуже.

Наконец, можно попробовать определить варианты дальнейшего развития. Условия, в которых решена задача стабилизации, предполагают, что полностью известны все вероятностные характеристики модели, чего на практике не бывает. Таким образом, перспективным развитием могут стать любые варианты адаптации представленных алгоритмов управления, обеспечивающие их нечувствительность к априорной неопределенности модели, не зависимо от того, будет ли это использование робастных методов оценивания или нечувствительных к риску управлений.

СПИСОК ЛИТЕРАТУРЫ

1. Athans M. The Role and Use of the Stochastic Linear-Quadratic-Gaussian Problem in Control System Design // IEEE T. Automat. Contr. 1971. V. 16. № 6. P. 529–552.
2. Astrom K.J. Introduction to Stochastic Control Theory. N.Y.: Acad. Press, 1970.
3. Lindquist A. On Feedback Control of Linear Stochastic Systems // SIAM J. Control. 1973. V. 11. № 2. P. 323–343.
4. Elliott R.J., Aggoun L., Moore J.B. Hidden Markov Models: Estimation and Control. N.Y.: Springer-Verlag, 1995.
5. Bar-Shalom Y., Willett P.K., Tian X. Tracking and Data Fusion: a Handbook of Algorithms. Storrs, Conn.: YBS Publishing, 2011.

6. *Wonham W.M.* Linear Multivariable Control. A Geometric Approach. Lecture Notes in Economics and Mathematical Systems, V. 101. Berlin: Springer-Verlag, 1974.
7. *Девис М.Х.А.* Линейное оценивание и стохастическое управление / Пер. с англ. М.: Наука, 1984.
8. *Kalman R.E., Bucy R.S.* New Results in Linear Filtering and Prediction Theory // Trans. ASME J. Basic Eng. 1965. № 83. P. 95–107.
9. *Босов А.В.* Задача управления линейным выходом нелинейной неуправляемой стохастической дифференциальной системы по квадратичному критерию // Изв. РАН. ТиСУ. 2021. № 5. С. 52–73.
10. *Липцер Р.Ш., Ширяев А.Н.* Статистика случайных процессов (нелинейная фильтрация и смежные вопросы). М.: Наука, 1974.
11. *Rishel R.* A Strong Separation Principle for Stochastic Control Systems Driven by a Hidden Markov Model // SIAM J. Control and Optimization. 1994. V. 32. P. 1008–1020.
12. *Wonham W.M.* Some Applications of Stochastic Differential Equations to Optimal Non-linear Filtering // SIAM J. Control. 1965. V. 2. P. 347–369.
13. *Ширяев А.Н.* Вероятность. 2-е изд. М.: Наука, 1989.
14. *Borisov A., Bosov A., Miller G.* Optimal Stabilization of Linear Stochastic System with Statistically Uncertain Piecewise Constant Drift // Mathematics. 2022. V. 10. № 2 (184).
15. *Флеминг У., Ришел Р.* Оптимальное управление детерминированными и стохастическими системами / Пер. с англ. М.: Мир, 1978.
16. *Bosov A., Borisov A.* Comparative Study of Markov Chain Filtering Schemas for Stabilization of Stochastic Systems under Incomplete Information // Mathematics. 2022. V. 10. № 18 (338).
17. *Босов А.В., Стефанович А.И.* Управление выходом стохастической дифференциальной системы по квадратичному критерию. II. Численное решение уравнений динамического программирования // Информатика и ее применения. 2019. Вып. 1. Т. 13. С. 9–15.
18. *Босов А.В., Стефанович А.И.* Управление выходом стохастической дифференциальной системы по квадратичному критерию. IV. Альтернативное численное решение // Информатика и ее применения. 2020. Вып. 1. Т. 14. С. 24–30.
19. *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.
20. *Bohacek S., Rozovskii B.* A Diffusion Model of Roundtrip Time // Computational Statistics & Data Analysis. 2004. V. 45. Iss. 1. P. 25–50.
21. *Босов А.В.* Стабилизация и слежение за траекторией линейной системы со скачкообразно изменяющимся дрейфом // АиТ. 2022. № 4. С. 27–46.
22. *Øksendal B.* Stochastic Differential Equations. An Introduction with Applications. N.Y.: Springer-Verlag, 2003.

ПРИЛОЖЕНИЕ

Д о к а з а т е л ь с т в о т е о р е м ы 1. Условия теоремы применяются к решению соответствующей задачи с конечным горизонтом. Это управление, обозначенное как $(U^\#)_0^T = \arg \min_{U_0^T \in \mathcal{U}_0^T} J(U_0^T)$, $u_t^\# = u_t^\#(y_t, z_t^\#, T)$, оптимальное на классе допустимых управлений

$$\mathcal{U}_0^T = \left\{ u_t = u_t(y, z, T), y \in \mathbb{R}^{n_y}, z \in \mathbb{R}^{n_z}, t \in [0, T] \right\}, \text{ может быть записано в виде [9]}$$

$$u_t^\#(y, z, T) = -\frac{1}{2} \left(R^T S R \right)^{-1} \left(c^T (2\alpha_t z + \beta_t) + 2R^T S (P y + Q z) \right), \quad (\text{П.1})$$

где матричный коэффициент $\alpha_t = \alpha_t(T) \in \mathbb{R}^{n_z \times n_z}$ является решением задачи Коши для уравнения Риккати:

$$\frac{d\alpha_t}{dt} + (M^\alpha)^T \alpha_t + \alpha_t M^\alpha + N^\alpha - \alpha_t c \left(R^T S R \right)^{-1} c^T \alpha_t = 0, \alpha_T = Q^T S Q, \quad (\text{П.2})$$

а векторный коэффициент $\beta_t = \beta_t(y, T) = \left(\beta_t^{(1)}(y, T), \dots, \beta_t^{(n_z)}(y, T) \right)^T \in \mathbb{R}^{n_z}$ — задачей Коши для системы дифференциальных уравнений в частных производных:

$$\begin{aligned} \frac{\partial \beta_t^{(i)}}{\partial t} + \frac{1}{2} \text{tr} \left\{ \Sigma^T \frac{\partial^2 \beta_t^{(i)}}{\partial y^2} \Sigma \right\} + A^T \frac{\partial \beta_t^{(i)}}{\partial y} + \sum_{j=1}^{n_y} y^{(j)} \left[M_t^\beta \right]^{(ji)} + \sum_{j=1}^{n_z} \beta_t^{(j)} \left[N_t^\beta \right]^{(ji)} &= 0 \\ M_t^\beta = M_t^\beta(T) &= 2 \left(\left(a^T - P^T S R \left(R^T S R \right)^{-1} c^T \right) \alpha_t + P^T \left(S - S R \left(R^T S R \right)^{-1} R^T S \right) Q \right) \\ N_t^\beta = N_t^\beta(T) &= b - c \left(R^T S R \right)^{-1} R^T S Q - c \left(R^T S R \right)^{-1} c^T \alpha_t, \\ \beta_T^{(i)} &= 2 \sum_{j=1}^{n_y} y^{(j)} \left[Q^T S P \right]^{(ji)}, \quad i = 1, \dots, n_z. \end{aligned} \quad (\text{П.3})$$

Искомое оптимальное управление $(U^*)_0^\infty = \arg \min_{U_0^\infty \in \mathcal{U}_0^\infty} J(U_0^\infty)$, $u_t^* = u_t^*(y_t, z_t^*, T)$, является пределом (П.1):

$$u^*(y, z) = \lim_{T \rightarrow \infty} u_t^\#(y, z, T) = -\frac{1}{2} \left(R^T S R \right)^{-1} \left(c^T (2\alpha_* z + \beta_*) + 2R^T S (P y + Q z) \right)$$

если существуют α_* и β_* .

Требования 1)–3) теоремы — это условия существования предельного решения $\alpha_* = \lim_{T \rightarrow \infty} \alpha_t$ матричного уравнения Риккати (П.2), неотрицательно определенного и не зависящего от граничного условия α_T , согласно формулировке теоремы 12.2 из [6]. Уточнить требуется только условие 3), которое сформулировано с учетом представления матрицы $N^\alpha = K K^T$, так как

$$\begin{aligned} &\left(E - S^{\frac{1}{2}} R \left(R^T S R \right)^{-1} R^T S^{\frac{1}{2}} \right) \left(E - S^{\frac{1}{2}} R \left(R^T S R \right)^{-1} R^T S^{\frac{1}{2}} \right)^T = \\ &= \left(E - S^{\frac{1}{2}} R \left(R^T S R \right)^{-1} R^T S^{\frac{1}{2}} \right)^2 = E - S^{\frac{1}{2}} R \left(R^T S R \right)^{-1} R^T S^{\frac{1}{2}}. \end{aligned}$$

Умножение справа и слева этого равенства на $Q^T S^{\frac{1}{2}}$ и $S^{\frac{1}{2}} Q$ дает разложение N^α , используемое в формулировке условия 3).

Требование 1) вместе с исходным предположением об эргодичности Y_t обеспечивает, кроме того, конечность целевого функционала (1.6). Кроме предела α_* упомянутая теорема 12.2 из [6] гарантирует устойчивость матрицы N_*^β — условие, без которого не может выполняться (2.4).

Наконец, предельный переход $\beta_*(y) = \lim_{T \rightarrow \infty} \beta_t(y, T)$ обеспечивается условием 4) теоремы 1, поскольку решение может быть представлено с помощью формулы Фейнмана-Каца [22]:

$$\beta_t(y, T) = \mathcal{E} \left\{ \begin{aligned} & 2I_t^{-1} I_T^{-1} Q^T S P y_T + \\ & + I_t^{-1} \int_t^T I_\tau^{-1} (M_\tau^\beta)^T y_\tau d\tau \end{aligned} \middle| \mathcal{F}_t^y \right\}, I_t^{-1} = \exp \left\{ (N_t^\beta)^T t \right\}. \quad (\text{П.4})$$

Учитывая, что матрица $N_*^\beta = \lim_{T \rightarrow \infty} N_t^\beta(T)$ (предел существует и не зависит от t , т.к. $N_t^\beta(T)$ выражается линейно через $\alpha_t(T)$) является устойчивой, первое слагаемое в (П.4) обращается в ноль, что вместе с условием 4) теоремы 1 обеспечивает существование предела при $T \rightarrow \infty$ в (П.3), не зависящего от t и граничного условия $\beta_T^{(i)}(y, T)$, что и дает уравнения (2.3), завершая доказательство.

Доказательство теоремы 2. Следует повторить рассуждения теоремы 1, используя вместо оптимального управления (2.5) решение задачи $(U^\#)_0^T = \arg \min_{U_0^T \in \mathcal{U}_0^T} J(U_0^T)$, $u_t^\# = u_t^\#(z_t^\#, T)$, оптимальное на классе $\mathcal{U}_0^T = \{u_t = u_t(z, T), z \in \mathbb{R}^{n_z}, t \in [0, T]\}$, т.е. для постановки с косвенными наблюдениями. Это управление может быть записано в виде [14]

$$u_t^\#(z, T) = -\frac{1}{2} (R^T S R)^{-1} (c^T (2\alpha_t z + \beta_t \hat{y}) + 2R^T S (P\hat{y} + Qz)). \quad (\text{П.5})$$

Оценки $\hat{y}(z_t^\#)$ на оптимальных траекториях $z_t^\#$ задает фильтр Вонэма (1.5), матричный коэффициент $\alpha_t = \alpha_t(T)$ является решением задачи Коши для уравнения Риккати (П.2), а матричный коэффициент $\beta_t = \beta_t(T) \in \mathbb{R}^{n_z \times n_y}$ — решением задачи Коши для линейного уравнения:

$$\frac{d\beta_t}{dt} = \beta_t \Lambda^T + M_t^\beta - N_t^\beta \beta_t = 0, \beta_T = 2Q_T^T S_T P_T, \quad (\text{П.6})$$

где M_t^β и N_t^β определены в (П.3).

Искомое оптимальное управление $(U^*)_0^\infty = \arg \min_{U_0^\infty \in \mathcal{U}_0^\infty} J(U_0^\infty)$, $u_t^* = u_t^*(z_t^*)$, является пределом (П.5):

$$u^*(z) = \lim_{T \rightarrow \infty} u_t^\#(z, T) = -\frac{1}{2} (R^T S R)^{-1} (c^T (2\alpha_* z + \beta_* \hat{y}) + 2R^T S (P\hat{y} + Qz))$$

если существуют α_* и β_* .

Отличие в рассуждениях, таким образом, касается только уравнения (П.6), т.е. существование предельной матрицы β_* — решения уравнения (3.2). Устойчивость (П.6) обеспечивает устойчивость матрицы N_*^β и изначально предполагаемая моделью (1.3) эргодичность цепи y_t .

Это замечание завершает доказательство.