

Стохастические системы

© 2022 г. А.В. БОСОВ, д-р техн. наук (AVosov@frccsc.ru)
(Федеральный исследовательский центр
“Информатика и управление” РАН, Москва)

СТАБИЛИЗАЦИЯ И СЛЕЖЕНИЕ ЗА ТРАЕКТОРИЕЙ ЛИНЕЙНОЙ СИСТЕМЫ СО СКАЧКООБРАЗНО ИЗМЕНЯЮЩИМСЯ ДРЕЙФОМ¹

Для управляемой линейной стохастической дифференциальной системы рассматривается задача слежения за скачкообразно изменяющимся состоянием аддитивного входного воздействия, определяющего текущее направление стабилизации (дрейф). Цель слежения — стабилизация системы возле изменяющегося дрейфа — формализуется квадратичным функционалом качества. Входное воздействие задает цепь Маркова с непрерывным временем. Задача рассмотрена в вариантах с полной и неполной информацией. В обоих случаях для решения используется динамическое программирование. Решение уравнения Беллмана в первом случае получается благодаря свойствам конечномерной цепи, во втором — принципу разделения задач управления и оценивания состояния, обеспечиваемому оценкой фильтра Вонэма и свойствам квадратичного критерия качества. Численный эксперимент использует прикладную модель, описывающую положение простого механического привода. Приводятся и подробно обсуждаются результаты расчетов, подтверждающие применимость полученных решений, а также способы преодоления трудностей их численной реализации.

Ключевые слова: слежение за целью, управление линейной дифференциальной системой, квадратичный функционал качества, динамическое программирование, фильтр Вонэма, принцип разделения.

DOI: 10.31857/S0005231022040031, EDN: AAGWVY

1. Введение

Важным стимулом для представленного в статье исследования стал один из примеров задачи управления, рассмотренный в [1, раздел 11.6. Risk-Sensitive LQG Control]. В этом примере целью управления линейной системой, часть из координат которой составляет наблюдаемый выход, объявляется следование этим выходом за некоторой заданной желаемой траекторией (desired trajectory). Формализуется цель управления с помощью квадратичной функции, зависящей от разности наблюдаемого выхода и целевой траектории, и в

¹ Работа выполнена при поддержке проекта № 075-15-2020-799 Министерства науки и высшего образования Российской Федерации. Работа выполнялась с использованием инфраструктуры Центра коллективного пользования “Высокопроизводительные вычисления и большие данные” (ЦКП “Информатика” ФИЦ ИУ РАН, Москва).

форме чувствительного к риску критерия, когда квадратичная форма помещается в показатель экспоненты. Рисксовая составляющая для дальнейшего не важна, а вот способ формализации задачи слежения оказался продуктивным. Варианты задач слежения за целью (target tracking) весьма многообразны [2]. Возможная постановка с хорошими прикладными перспективами состоит в том, чтобы вместо заданной априорно траектории следовать за некоторым скачкообразно изменяющимся входом — марковской цепью. Разные состояния такой цепи могут задавать разные направления дрейфа, которому должна следовать управляемая система, причем эти направления заранее неизвестны и постоянно меняются.

Поиск публикаций по такого рода тематике показал, что близкие постановки ранее уже привлекали внимание. Наиболее общий результат получен в [3], где задача поставлена с помощью классического квадратичного критерия, вход описывается произвольным марковским процессом, цепь возможна как частный случай. Более ранние результаты [4, 5] использовали квадратичный критерий частного вида (см. также [6]), включающий только слагаемые с управлением. В фокусе внимания всех этих публикаций не столько само управление, оно получается в форме решения классической линейно-квадратичной задачи, сколько борьба со входом, который интерпретируется как сложное возмущение и соответственно не наблюдается, т.е. фактически в цитированных публикациях решается задача стабилизации наблюдаемой выходной переменной в условиях неполной информации о входном воздействии. Подход к решению обеспечивается общими уравнениями нелинейной фильтрации [7], а теоретическую проблему составляет выяснение действенности принципа разделения. Любопытно, что задача с полной информацией не рассматривалась, равно как не обсуждались и возможные приложения результатов.

Данная статья комбинирует формализацию квадратичной функцией задачи слежения за целью [1] и результат для классического квадратичного критерия [3]. Но в отличие от [1] желаемая траектория не предполагается известной, а описывается марковской цепью с непрерывным временем. При этом в отличие от [3] в критерий включено слагаемое с разностью входа и выхода, отвечающее за слежение, что дает формально другие уравнения для оптимального управления. Кроме того, рассмотрена и задача с полной информацией, т.е. в предположении, что текущее состояние входного процесса известно.

Большое внимание уделено прикладному примеру и вычислительному эксперименту. Предложенная в статье прикладная модель описывает поведение простого механического привода, целью функционирования которого является не простая стабилизация возле нулевого положения, а стабилизация возле постоянно дрейфующего положения, модель которого задается марковской цепью. Вопросы численной реализации представлены весьма подробно и иллюстрируют не только применимость полученных теоретических результатов, но и сопутствующие реализационные проблемы.

2. Модель и постановка задачи

На каноническом вероятностном пространстве $(\Omega, \mathcal{F}, \mathcal{P}, \mathcal{F}_t)$, $t \in [0, T]$, рассмотрим линейную дифференциальную стохастическую систему с управляемым вектором выхода $z_t \in \mathbb{R}^{n_z}$:

$$(1) \quad dz_t = a_t y_t dt + b_t z_t dt + c_t u_t dt + \sigma_t dw_t, \quad z_0 = Z.$$

Систематическую составляющую траекториям z_t (направление дрейфа) обеспечивает марковский скачкообразный процесс y_t — цепь 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} . Вероятностные характеристики y_t предполагаются известными:

- распределение начального состояния $y_0 = Y$ обозначается π_0 ;
- матрица интенсивностей переходов Λ_t .

Остальные величины в уравнении (1):

- $w_t \in \mathbb{R}^{n_w}$ — стандартный векторный винеровский процесс;
- $Z \in \mathbb{R}^{n_z}$ — гауссовская случайная величина с известными математическим ожиданием и ковариацией, w_t, y_t, Y, Z независимы в совокупности;
- $u_t \in \mathbb{R}^{n_u}$ — управление — случайный процесс с конечным вторым моментом;
- $a_t \in \mathbb{R}^{n_z \times n_y}$, $b_t \in \mathbb{R}^{n_z \times n_z}$, $c_t \in \mathbb{R}^{n_z \times n_u}$, $\sigma_t \in \mathbb{R}^{n_z \times n_w}$ — заданные матричные функции.

Система (1) будет рассмотрена для двух случаев. Первый — случай полной информации, когда наблюдаются и выход z_t , и входное воздействие y_t . Второй — случай неполной информации, когда скачкообразный процесс y_t доступен только по косвенным наблюдениям, в роли которых выступает выход z_t .

Управление u_t выбирается из класса допустимых управлений, который для простоты определяется как класс управлений с полной обратной связью по выходу, что с учетом марковского характера задачи [8] не ограничивает общности рассуждений. Таким образом, $u_t = U_t(y_t, z_t)$ в постановке с полной информацией и $u_t = U_t(z_t)$ в случае косвенных наблюдений за y_t . Соответственно определение закона управления $U_t = U_t(y, z)$ или $U_t = U_t(z)$, $y \in \mathbb{R}^{n_y}$, $z \in \mathbb{R}^{n_z}$, является целью оптимизации функционала качества вида

$$(2) \quad J(U_0^T) = \mathbb{E} \left\{ \int_0^T \|P_t y_t + Q_t z_t + R_t u_t\|_{S_t}^2 dt + \|P_T y_T + Q_T z_T\|_{S_T}^2 \right\},$$

где $U_0^T = \{U_t(y, z) \text{ или } U_t(z), 0 \leq t \leq T\}$, $P_t \in \mathbb{R}^{n_J \times n_y}$, $Q_t \in \mathbb{R}^{n_J \times n_z}$, $R_t \in \mathbb{R}^{n_J \times n_u}$, $S_t \in \mathbb{R}^{n_J \times n_J}$, $S_t \geq 0$, $S_t = S'_t$, $0 \leq t \leq T$, — заданные ограниченные матричные функции, весовая функция $\|x\|_S^2 = x' S x$ для симметричной неотрицательно определенной матрицы S , единичной матрице $S = \mathbf{1}$ соответствует евклидова норма $\|x\|_1^2 = |x|^2$, x' — транспонированная матрица x .

Относительно параметров задачи предполагаются выполненными следующие условия:

- ограниченность параметров модели выхода: $|a_t| + |b_t| + |c_t| + |\sigma_t| \leq C$ для всех $0 \leq t \leq T$ обеспечивает существование решения уравнения (1) для любого допустимого управления u_t ;
- кусочная непрерывность матричных функций $\Lambda_t, a_t, b_t, c_t, \sigma_t, P_t, Q_t, R_t, S_t$ обеспечит выполнение типовых условий существования решений обыкновенных дифференциальных уравнений, получаемых далее;
- штраф за неограниченные управления обеспечивает обычное условие невырожденности $R'_t S_t R_t > 0$;
- возможность использования выхода z_t в качестве наблюдений в постановке задачи с неполной информацией обеспечивает невырожденность ошибок наблюдений, т.е. условие $\sigma_t \sigma'_t > 0$.

Отметим, что принципиальные результаты по оптимизации системы (1) по классическому квадратичному критерию качества, процитированные во введении, получены именно для второго случая, задача с полной информацией отдельно не рассматривалась. Возможное объяснение этому видится в том обстоятельстве, что процесс y_t интерпретировался исключительно как возмущение, с которым надо бороться и исключить его влияние. Для этого использовался классический функционал качества

$$(3) \quad J(U_0^T) = \mathbb{E} \left\{ \int_0^T |Q_t z_t + R_t u_t|^2 dt + |Q_T z_T|^2 \right\}.$$

Здесь можно напомнить, что представление управляемой системы в форме вход-выход, предполагающее, что входная переменная является неуправляемой, а задачей является регулирование выхода, используется давно. Одним из первых было применение специальным образом записанной классической постановки линейно-квадратичного управления для решения задачи “доведения выхода до нуля” или, другими словами, для стабилизации выхода около нуля, описанное в [9]. Рассматриваемая модель (1) исходит из интерпретации переменной z_t в качестве управляемого выхода, а y_t — из неуправляемого входа, который может быть как наблюдаемым, так и ненаблюдаемым. Функционал качества (2) исходит из конструктивного понимания y_t как воздействия, требующего ответной реакции управляемой системы, изменения направления движения, выведения на другой координатный уровень, компенсации изменившегося силового воздействия и т.п. Для этого функционалом (2) обеспечивается возможность включать в критерий слагаемые вида $|z_t - P_t y_t|^2$, т.е. ставить задачу слежения за траекторией входа (состояния стохастической системы наблюдения), слежения за дрейфом выхода, формируемым скачкообразным процессом y_t . Эта возможность позволяет в дополнение к традиционной стабилизации решать другие практически значимые задачи маневрирования, в том числе представленную далее задачу управления механическим приводом.

3. Основные утверждения

Приведем сначала решение задачи с полной информацией, т.е. когда в системе (1), оптимизируемой по критерию (2), допустимые управления относятся к классу $\mathcal{F}_t^{y,z}$ -измеримых процессов, где σ -алгебра $\mathcal{F}_t^{y,z}$ порождена наблюдаемыми величинами $\{y_\tau, z_\tau, 0 \leq \tau \leq t\}$, так что $\mathcal{F}_t^{y,z} \subseteq \mathcal{F}_t \subseteq \mathcal{F}$. Решение задачи представлено в следующем утверждении.

Теорема 1. Решение задачи оптимизации

$$(U^*)^T = \{U_t^*(y, z), 0 \leq t \leq T\} \in \operatorname{argmin} J(U_0^T)$$

для целевого функционала $J(U_0^T)$, заданного в (2), определяется соотношениями:

$$(4) \quad U_t^* = U_t^*(y, z) = -\frac{1}{2}(R_t' S_t R_t)^{-1} (c_t' (2\alpha_t z + \beta_t y) + 2R_t' S_t (P_t y + Q_t z)),$$

$$(5) \quad \frac{d\alpha_t}{dt} - (M_t^\alpha \alpha_t + \alpha_t' (M_t^\alpha)') + N_t^\alpha - \alpha_t' c_t (R_t' S_t R_t)^{-1} c_t' \alpha_t = 0, \\ \alpha_T = Q_T' S_T Q_T,$$

$$(6) \quad \frac{d\beta_t}{dt} + \beta_t \Lambda_t' + M_t^\beta - N_t^\beta \beta_t = 0, \quad \beta_T = 2Q_T' S_T P_T,$$

где

$$M_t^\alpha = Q_t' S_t R_t (R_t' S_t R_t)^{-1} c_t', \quad N_t^\alpha = Q_t' (S_t - S_t R_t (R_t' S_t R_t)^{-1} R_t' S_t) Q_t, \\ M_t^\beta = 2 \left((a_t' - P_t' S_t R_t (R_t' S_t R_t)^{-1} c_t') \alpha_t + P_t' (S_t - S_t R_t (R_t' S_t R_t)^{-1} R_t' S_t) Q_t \right), \\ N_t^\beta = Q_t' S_t R_t (R_t' S_t R_t)^{-1} c_t' + \alpha_t c_t (R_t' S_t R_t)^{-1} c_t'.$$

Траекторию оптимального управления далее будем обозначать $u_t^* = U_t^*(y_t, z_t^*)$, где через z_t^* обозначена оптимальная траектория выхода — решение (1), отвечающее $u_t = u_t^*$.

Формальный вывод соотношений (4)–(6) представлен в [10], а именно: получено решение задачи минимизации целевого функционала $J(U_0^T)$ для выхода z_t , заданного уравнением (1), и входного воздействия y_t , описываемого диффузионным уравнением

$$(7) \quad dy_t = \Phi_t(y_t) dt + \Sigma_t(y_t) dV_t, \quad y_0 = Y,$$

где $V_t \in \mathbb{R}^{nv}$ — стандартный векторный винеровский процесс, не зависящий от w_t, Y, Z .

Для решения использовалось уравнение Беллмана [3], получить соотношения для оптимального управления позволяет предположение о представлении функции Беллмана

$$V_t = V_t(y, z) = \inf_{U_t^T} \mathbb{E} \left\{ \int_t^T \|P_s y_s + Q_s z_s + R_s u_s\|_{S_s}^2 ds + \|P_T y_T + Q_T z_T\|_{S_T}^2 \right\}$$

в виде

$$(8) \quad V_t = z' \alpha_t z + z' B_t(y) + \Gamma_t(y).$$

Уравнения для $V_t(y)$ уточняются для частного случая линейного сноса в (7), т.е. линейной функции $\Phi_t(y) = \varphi_t y$. Функция $V_t(y)$ также оказывается линейной $V_t(y) = \beta_t y$.

Соотношения (4)–(6) теоремы 1 — это уравнения для α_t и β_t из [10] с учетом того, что вместо диффузионного процесса (7) используется мартингалное представление скачкообразного y_t [1]

$$(9) \quad dy_t = \Lambda'_t y_t dt + d\Lambda_t^y, \quad y_0 = Y,$$

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

$$\langle \Lambda^y, \Lambda^y \rangle_t = \int_0^T (\text{diag}(\Lambda'_s y_s) - \Lambda'_s \text{diag}(y_s) - \text{diag}(y_s) \Lambda_s) ds.$$

Соответственно вместо функции сноса $\Phi_t(y) = \varphi_t y$ из (7) используется $\Lambda'_t y$, вместо ковариации винеровского процесса — квадратичная характеристика $\langle \Lambda^y, \Lambda^y \rangle_t$, так что с учетом формулы Ито для семимартингалов [11] и независимости y_t и w_t , точнее Λ_t^y и w_t , уравнение Беллмана, решаемое в [10], могло бы быть записано в виде

$$(10) \quad \frac{\partial V_t}{\partial t} + \frac{1}{2} \left(\text{tr} \left\{ \frac{\partial^2 V_t}{\partial y^2} \langle \Lambda^y, \Lambda^y \rangle_t \right\} + \text{tr} \left\{ \sigma'_t \frac{\partial^2 V_t}{\partial z^2} \sigma_t \right\} \right) + \\ + \min_u \left\{ y' \Lambda_t \frac{\partial V_t}{\partial y} + (a_t y + b_t z + c_t u)' \frac{\partial V_t}{\partial z} + \|P_t y + Q_t z + R_t u\|_{S_t}^2 \right\} = 0.$$

Однако в рассматриваемой задаче модель входа даже проще, чем диффузия $\Sigma_t(y_t) dV_t$ в (7), поскольку цепь y_t принимает значения единичных векторов $\{e_1, \dots, e_{n_y}\}$, что позволяет уточнить функцию Беллмана (8), задав линейным слагаемое $\Gamma_t(y) = \gamma_t y + \delta_t$:

$$(11) \quad V_t = z' \alpha_t z + z' \beta_t y + \gamma_t y + \delta_t.$$

Таким образом, уравнение Беллмана для рассматриваемой задачи принимает вид

$$(12) \quad \frac{\partial V_t}{\partial t} + \frac{1}{2} \text{tr} \left\{ \sigma'_t \frac{\partial^2 V_t}{\partial z^2} \sigma_t \right\} + \\ + \min_u \left\{ y' \Lambda_t \frac{\partial V_t}{\partial y} + (a_t y + b_t z + c_t u)' \frac{\partial V_t}{\partial z} + \|P_t y + Q_t z + R_t u\|_{S_t}^2 \right\} = 0$$

и решается с начальным условием

$$(13) \quad V_T = \|P_T y + Q_T z\|_{S_T}^2.$$

Решить (12) можно непосредственно, подставив в него V_t из (11). После этой подстановки получается квадратичная форма, минимум которой достигается в U_t^* из (4). В итоге (12) примет вид квадратичной формы относительно z , а соотношения для коэффициентов $\alpha_t, \beta_t, \gamma_t, \delta_t$ из (11) получаются приравниванием к нулю коэффициентов при $z'z, z'$ и z^0 . Удобнее уравнения для $\alpha_t, \beta_t, \gamma_t, \delta_t$ записать, воспользовавшись общим решением из [10]. Для α_t, β_t так получаются (5), (6), а уравнения для γ_t и δ_t имеют вид

$$(14) \quad \begin{aligned} \frac{\partial \gamma_t}{\partial t} + y' \Lambda_t \Lambda_t' + M_t^\gamma y &= 0, & \gamma_T y &= y' P_T' S_T P_T y, \\ \frac{\partial \delta_t}{\partial t} + \text{tr} \{ \sigma_t' \alpha_t \sigma_t \} &= 0, & \delta_T &= 0, \\ M_t^\gamma y = N_t^\gamma (y) &= y' \beta_t' \left(a_t - c_t (R_t' S_t R_t)^{-1} R_t' S_t P_t \right) y + \\ &+ y' P_t' \left(S_t - S_t R_t (R_t' S_t R_t)^{-1} R_t' S_t \right) P_t y - \frac{1}{4} y' \beta_t' c_t (R_t' S_t R_t)^{-1} c_t' \beta_t y. \end{aligned}$$

Для подтверждения сделанного предположения о линейности $\Gamma_t(y) = \gamma_t y + \delta_t$ в (14) должна иметь место линейность $N_t^\gamma(y)$. Несмотря на наличие множителей $y'y$, это условие выполнено благодаря ограниченной единичными векторами области значений процесса y_t , поскольку имеет место представление $M_t^\gamma y = (N_t^\gamma(e_1), \dots, N_t^\gamma(e_{n_y})) y$. Аналогично равенство $\gamma_T y = y' P_T' S_T P_T y$, записанное в виде $\gamma_T y = (e_1' P_T' S_T P_T e_1, \dots, e_{n_y}' P_T' S_T P_T e_{n_y}) y$, дает начальное условие для вычисления γ_t . Отметим, что такое свойство типично для задач, использующих модель марковской цепи с областью значений $\{e_1, \dots, e_{n_y}\}$. Так, в [12] это свойство обеспечило линейность и целевого функционала, и функции Беллмана в задаче управления марковским процессом с конечным числом состояний и оптимизируемой матрицей интенсивностей.

Отметим одно интересное свойство рассмотренной задачи (1), (2), (9). Полученное решение аналогично решению классической задачи линейно-квадратичного (LQ) управления. Действительно, управляемая система (1), (9) описывается линейными уравнениями, целевой функционал (2) является квадратичным по переменным состояния (входа и выхода) и управления, оптимальное управление получается линейным, а функция цены — квадратичной функцией состояния. Нетрудно проверить, что если вместо (9) использовать для y_t модель с винеровским процессом V_t вместо мартингала Λ_t^y , т.е. линейный вариант (7) вида $dy_t = \Lambda_t^y y_t dt + dV_t$, то получится частный случай классической LQ задачи. При этом решение этой задачи дает те же соотношения (4)–(6). Но есть и отличие. Функция Беллмана в классическом случае отличается от полученного представления (11), в ее составе будет еще одно слагаемое с множителем $y'y$, поскольку цепи нет, а значит, нет и линейности по y . Соответственно вместо уравнений (14) будут другие уравнения для трех, а не двух, коэффициентов квадратичной формы по y .

Перейдем к решению задачи с неполной информацией, т.е. будем оптимизировать систему (1) по критерию (2), выбирая допустимые управления из класса \mathcal{F}_t^z -измеримых процессов, где σ -алгебра \mathcal{F}_t^z порождена наблюдаемыми

величинами $\{z_\tau, 0 \leq \tau \leq t\}$, так что $\mathcal{F}_t^z \subseteq \mathcal{F}_t^{y,z} \subseteq \mathcal{F}_t \subseteq \mathcal{F}$. Отметим, что в этом случае выход z_t интерпретируется также в качестве косвенных наблюдений за y_t .

Решить задачу для этого случая позволяют два классических результата. Во-первых, это фильтр Вонэма [1], который позволяет заменить процесс y_t , описываемый уравнением состояния (9), оценкой $\hat{y}_t = \mathbb{E}\{y_t | \mathcal{F}_t^z\}$, описываемой стохастическим дифференциальным уравнением с винеровским процессом. Во-вторых, это теорема разделения [13], условиям которой удовлетворяет квадратичный функционал качества.

Фильтр Вонэма для цепи y_t , заданной уравнением (9), и процесса наблюдений z_t , заданного уравнением (1), имеет вид

$$(15) \quad d\hat{y}_t = \Lambda'_t \hat{y}_t dt + (\text{diag}(\hat{y}_t) - \hat{y}_t \hat{y}'_t) a'_t (\sigma_t \sigma'_t)^{-1/2} \times \\ \times (\sigma_t \sigma'_t)^{-1/2} (dz_t - a_t \hat{y}_t dt - b_t z_t dt - c_t u_t dt), \quad \hat{y}_0 = \mathbb{E}\{Y\},$$

где

$$(16) \quad dW_t = (\sigma_t \sigma'_t)^{-1/2} (dz_t - a_t \hat{y}_t dt - b_t z_t dt - c_t u_t dt)$$

— стохастический дифференциал \mathcal{F}_t^z -измеримого стандартного векторного винеровского процесса $W_t \in \mathbb{R}^{n_z}$. Соответственно (15) и (16) можно записать в виде

$$(17) \quad d\hat{y}_t = \Lambda'_t \hat{y}_t dt + \Sigma_t(\hat{y}_t) dW_t, \quad \hat{y}_0 = \mathbb{E}\{Y\}, \\ dz_t = a_t \hat{y}_t dt + b_t z_t dt + c_t u_t dt + \hat{\sigma}_t dW_t, \quad z_0 = Z,$$

где $\Sigma_t(y) = (\text{diag}(y) - yy') a'_t (\sigma_t \sigma'_t)^{-1/2}$, $\hat{\sigma}_t = (\sigma_t \sigma'_t)^{1/2}$.

Важно отметить здесь, что оценка \hat{y}_t согласно (15) не зависит от реализуемого управления u_t , что обеспечивает отсутствие в задаче дуального эффекта, т.е. влияния закона управления на точность оценивания будущих состояний [14]. Благодаря этому можно преобразовать целевой функционал $J(U_0^T)$ к виду

$$J(U_0^T) = \mathbb{E} \left\{ \int_0^T \|P_t \hat{y}_t + Q_t z_t + R_t u_t\|_{S_t}^2 dt + \|P_T \hat{y}_T + Q_T z_T\|_{S_T}^2 + \right. \\ \left. + \int_0^T \|P_t (y_t - \hat{y}_t)\|_{S_t}^2 dt + \|P_T (y_T - \hat{y}_T)\|_{S_T}^2 \right\}$$

и утверждать, что второй интеграл, содержащий слагаемые, определяемые точностью фильтра (15), не зависит от реализуемого управления, а значит, может быть исключен. Таким образом, для системы с полной информацией (17), эквивалентной системе наблюдения (9), (1) в смысле разделения задач управления и фильтрации, можно применять функционал

$$(18) \quad J(U_0^T) = \mathbb{E} \left\{ \int_0^T \|P_t \hat{y}_t + Q_t z_t + R_t u_t\|_{S_t}^2 dt + \|P_T \hat{y}_T + Q_T z_T\|_{S_T}^2 \right\}.$$

Заметим, что здесь имеет место полное (или сильное, как названо в [3]) разделение в том смысле, что в результате разделения остались такими же и уравнение наблюдений, и выражение целевого функционала.

Выполненные преобразования дают возможность сформулировать следующее утверждение.

Теорема 2. Решением задачи оптимизации системы (9), (1)

$$(U^{**})_0^T = \{U_t^{**}(z), 0 \leq t \leq T\} \in \operatorname{argmin} J(U_0^T)$$

для целевого функционала $J(U_0^T)$, заданного в (2), является решение задачи оптимизации системы (17) для целевого функционала $J(U_0^T)$, заданного в (18).

Данный результат означает, что траектории оптимального управления в задаче с неполной информацией $u_t^{**} = U_t^{**}(z_t^{**})$, где через z_t^{**} обозначена оптимальная траектория выхода — решение (1), отвечающее $u_t = u_t^{**}$, может быть вычислено как $u_t^{**} = U_t^{**}(\hat{y}_t, z_t^{**})$ и закон управления $U_t^{**}(y, z)$ определяется теми же соотношениями (4)–(6), т.е., как и предполагает принцип сильного разделения, управление U_t^{**} для случая неполной информации получается из управления U_t^* для случая полной информации заменой переменной, соответствующей состоянию y_t , на переменную, соответствующую оптимальной оценке фильтрации состояния \hat{y}_t .

Поскольку состояние \hat{y}_t описывается уравнением Ито с винеровским процессом, то задачу управления для (17), (18) можно решать как частный случай задачи, решенной в [10], причем для варианта с линейным сносом $\Phi_t(y) = \varphi_t y$. Единственной особенностью при этом является зависимость возмущений в уравнениях состояния \hat{y}_t и наблюдения z_t в (17). Нетрудно видеть, что эта зависимость, точнее равенство, поскольку возмущения в обоих уравнениях описываются одним дифференциалом dW_t , отражается на уравнении Беллмана, которое, в отличие от (10), примет вид

$$(19) \quad \frac{\partial V_t}{\partial t} + \frac{1}{2} \operatorname{tr} \left\{ \Sigma_t' \frac{\partial^2 V_t}{\partial z^2} \Sigma_t + \hat{\sigma}_t' \frac{\partial^2 V_t}{\partial z^2} \hat{\sigma}_t + 2 \Sigma_t' \frac{\partial^2 V_t}{\partial y \partial z} \hat{\sigma}_t \right\} + \\ + \min_u \left\{ y' \Lambda_t \frac{\partial V_t}{\partial y} + (a_t y + b_t z + c_t u)' \frac{\partial V_t}{\partial z} + \|P_t y + Q_t z + R_t u\|_{S_t}^2 \right\} = 0.$$

Здесь обозначено $\Sigma_t = \Sigma_t(y)$ и учтено, что $\Sigma_t \in \mathbb{R}^{n_y \times n_z}$, $\hat{\sigma}_t \in \mathbb{R}^{n_z \times n_z}$.

Решение (19) аналогично случаю с независимыми возмущениями находится в виде

$$(20) \quad V_t = z' \alpha_t z + z' \beta_t y + \Gamma_t(y)$$

и отличается тем, что последнее слагаемое $\Gamma_t(y)$ упростить не удастся. Для $\Gamma_t(y)$ получается уравнение

$$(21) \quad \frac{\partial \Gamma_t}{\partial t} + \frac{1}{2} \operatorname{tr} \left\{ \Sigma_t' \frac{\partial^2 \Gamma_t}{\partial y^2} \Sigma_t \right\} + \operatorname{tr} \left\{ \hat{\sigma}_t' \alpha_t \hat{\sigma}_t + \Sigma_t' \beta_t' \hat{\sigma}_t \right\} + y' \Lambda_t \frac{\partial \Gamma_t}{\partial y} + N_t^\gamma = 0.$$

В этом уравнении фигурирует тот же коэффициент $N_t^\gamma = N_t^\gamma(y)$ из (14), а упростить его нельзя, во-первых, из-за того, что y здесь принимает произвольные значения, а во-вторых, из-за существенно нелинейного характера $\Sigma_t(y)$. Большого значения отличия в выражениях между $\Gamma_t = \Gamma_t(y)$ из (21) и $\Gamma_t(y) = \gamma_t y$ из (14) не имеют, так как соотношения и для γ_t , и для Γ_t имеют исключительно академический интерес и не влияют на расчеты при практической реализации соответствующих управлений.

Отметим, что полученное решение можно рассматривать как обобщение результата публикации [3] для критерия (2), т.е. для приложения к задаче слежения, но терминологические отличия несколько значительнее. В [3] и более ранних [4, 5] отправной точкой следует считать общие уравнения оптимальной фильтрации на основе обновляющих процессов [7], обеспечивающие возможность разделения задач управления и фильтрации. Рассматриваемая постановка в большей степени порождена задачей с полной информацией [10] и конструктивной интерпретацией ненаблюдаемого процесса в качестве состояния системы управления, определяющего цель управления, а не возмущения, препятствующего достижению цели. Возможно, по этой причине в упомянутых публикациях отсутствуют прикладные примеры, а оставшийся материал статьи посвящен детальному рассмотрению именно такого примера.

4. Модель механического привода и демонстрация возможностей управления

В качестве технической системы, ставшей источником для используемой модели механического привода, выступил мостовой кран или его более простой вариант — кран-балка. Моделируется управление перемещением по рельсу крана тележки с закрепленным талем для подъема груза. Перемещаемая тележка обладает значительной инерцией, а целевых положений для ее размещения конечное число, например линий погрузки-разгрузки или железнодорожных путей. В модели состояние тележки-привода описывается скалярной переменной x_t , задающей положение на балке, и регулируемой скоростью v_t , определяемой силой, которая линейно зависит от текущего положения, текущей скорости, неконтролируемого входного воздействия y_t (номера линии погрузки) и управления u_t :

$$(22) \quad \begin{aligned} dx_t &= v_t dt, \quad t \in (0, T], \\ dv_t &= ax_t dt + bv_t dt + cy_t dt + hu_t dt + \sqrt{g} dw_t. \end{aligned}$$

Марковская цепь y_t имеет три состояния, т.е. $n_y = 3$, $y_t \in \{e_1, e_2, e_3\}$, и постоянную матрицу интенсивностей $\Lambda_t = \Lambda$. Во всех расчетах использовалась модель простого процесса рождения-гибели с матрицей интенсивностей

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

и начальным распределением $\pi = (1, 0, 0)'$, т.е. $y_0 = Y = e_1$.

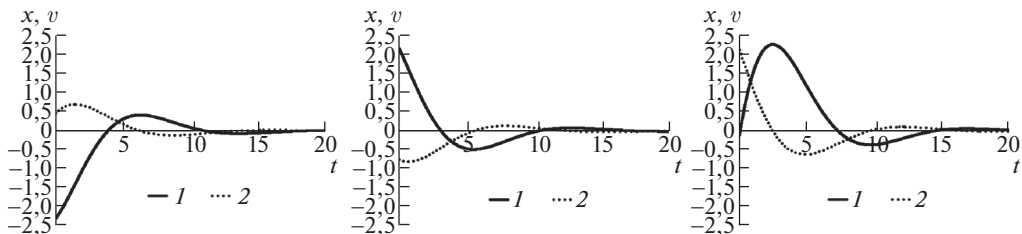


Рис. 1. Примеры траекторий для модели привода без внешних воздействий: 1 — положение x_t , 2 — скорость v_t .

Скаляры a, b, h, g и строка (c_1, c_2, c_3) — известные постоянные; w_t — стандартный винеровский процесс. Начальные условия x_0 и v_0 предполагаются независимыми гауссовскими случайными величинами с нулевым средним и дисперсиями σ_x^2 и σ_v^2 соответственно.

Нетрудно видеть, что система (22) является устойчивой, если $b < 0$ и $b^2 + 4a < 0$, поскольку b и $b^2 + 4a$ — собственные числа матрицы $\begin{pmatrix} 0 & 1 \\ a & b \end{pmatrix}$. Сначала рассмотрим поведение именно устойчивого привода, а в разделе 6 оценим возможность управления неустойчивой системой.

Для первого из рассмотренных примеров выбирались следующие параметры: $a = -0,25$, $b = -0,5$, $\sigma_x^2 = 1$, $\sigma_v^2 = 1$, $c_1 = c_2 = c_3 = 0$, $h = 0$.

Этот набор параметров дает иллюстрацию поведения привода, на который не действуют внешние воздействия, помимо начального положения/скорости и возмущения w_t , влияние которого сделаем минимальным, положив $g = 0,001$ (это позволит дальше лучше иллюстрировать влияние неконтролируемого входа y_t).

Интегрирование в этом примере и во всех следующих системы (22) выполнено методом Эйлера с шагом 0,001, неявный метод Эйлера с таким же шагом использовался для приближенного интегрирования (5) и (6) при вычислении коэффициентов оптимальных управлений.

На рис. 1 показаны примеры траекторий x_t и v_t , рассчитанных для $T = 20$. Такое значение горизонта для модельного примера несколько избыточно, но в этом примере лучше иллюстрирует завершение переходного процесса.

Теперь изменим параметры, положив $c_1 = -1,5$, $c_2 = -0,5$, $c_3 = 0,5$.

Чтобы показать влияние на траектории x_t и v_t значений, принимаемых цепью y_t , на графиках с иллюстрациями расчетов следующего примера показана траектория процесса Cy_t для $C = \left(-\frac{c_1}{a}, -\frac{c_2}{a}, -\frac{c_3}{a}\right) = (-6, -2, 2)$.

Именно этот процесс представляет ненулевой дрейф x_t , который определяет входное воздействие y_t . Примеры траекторий показаны на рис. 2.

Показанные на рис. 2 траектории v_t , реагируя на смену состояния y_t , действительно направляют x_t в сторону дрейфа Cy_t . Если бы цепь y_t пребывала в одном состоянии длительное время, то каждый промежуточный переходной процесс, связанный со сменой состояния y_t , приводил бы к стабилизации x_t около Cy_t , а v_t — около 0. Можно предполагать, ссылаясь на рис. 1, что такая

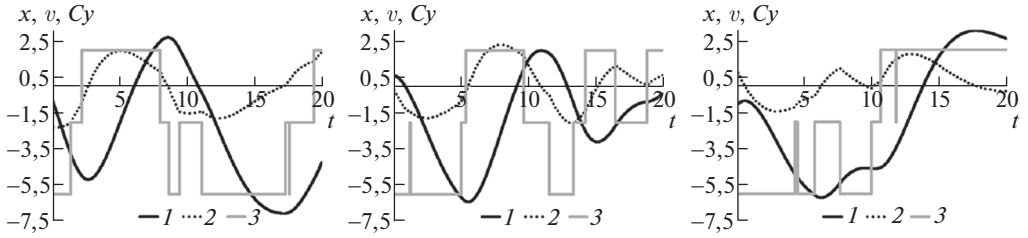


Рис. 2. Примеры траекторий для модели привода с внешним скачкообразным воздействием: 1 — положение x_t , 2 — скорость v_t , 3 — дрейф Cy_t .

промежуточная стабилизация достигала бы результата в течение 10–15 с, но характеристики цепи таковы, что ее динамика дождаться этого результата не дает, на рис. 2 видна скорее тенденция к слежению за дрейфом Cy_t , чем результативность реагирования на изменяющийся вход.

Теперь сформулированную цель слежения в работе механического привода можно формализовать в целевой функции управления вида

$$(23) \quad J(U_0^T) = \mathbb{E} \left\{ \int_0^T (|Cy_t - x_t|^2 + R|u_t|^2) dt \right\}.$$

Использованные в приведенных примерах параметры уточняются значениями $h = 1$, $R = 0,00001$, $T = 10$.

Модель опять исследуется в некоторых “идеальных” условиях малого возмущения $\sqrt{g}dw_t$ и фактически отсутствующего штрафа за энергию $R|u_t|^2$, затрачиваемую на управление. Вычисление в этих условиях траекторий x_t^* , v_t^* , u_t^* в соответствии с теоремой 1, т.е. в условиях полной информации о состоянии цепи y_t , и траекторий x_t^{**} , v_t^{**} , u_t^{**} в соответствии с теоремой 2, т.е. использование вместо y_t оценки фильтра Вонэма \hat{y}_t , покажут потенциал достижения цели слежения за положением x_t , выраженной функционалом (23). Заметим, что формальное нарушение системой (22) условия невырожденности ошибок наблюдений $\sigma_t \sigma_t' > 0$, требуемое для вычисления оценки фильтрации по формуле (15), не препятствует расчетам, поскольку в рассматриваемой модели $\hat{y}_t = \mathbb{E}\{y_t | \mathcal{F}_t^z\} = \mathbb{E}\{y_t | \mathcal{F}_t^v\}$, т.е. фильтр Вонэма строится по одному наблюдению v_t .

Иллюстрации траекторий x_t^* , x_t^{**} и v_t^* , v_t^{**} , u_t^* , u_t^{**} приведены на рис. 3 и рис. 4 соответственно. Представленные реализации целевой траектории Cy_t одинаковы в примерах на обоих рисунках.

Отметим, что если на рис. 3 визуальнo отличить положения x_t^* и x_t^{**} , отвечающие случаям полной и неполной информации, возможно, то скорости v_t^* , v_t^{**} и управления u_t^* , u_t^{**} визуальнo неразличимы, поэтому на рис. 4 они показаны одними “общими” линиями. Также на рис. 4 ограничены величины, представленные по вертикальной оси, потому что u_t^* и u_t^{**} в моменты изменения состояний цепи y_t достигают значений порядка ± 2000 , что, конечно, не имеет практического смысла и объясняется малым штрафом R в (23).

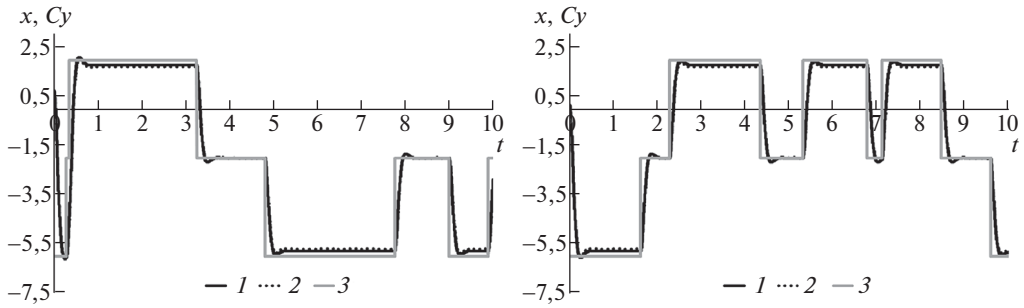


Рис. 3. Примеры траекторий положения для модели привода в идеальных условиях: 1 — положение x_t^* , 2 — положение x_t^{**} , 3 — дрейф Cy_t .

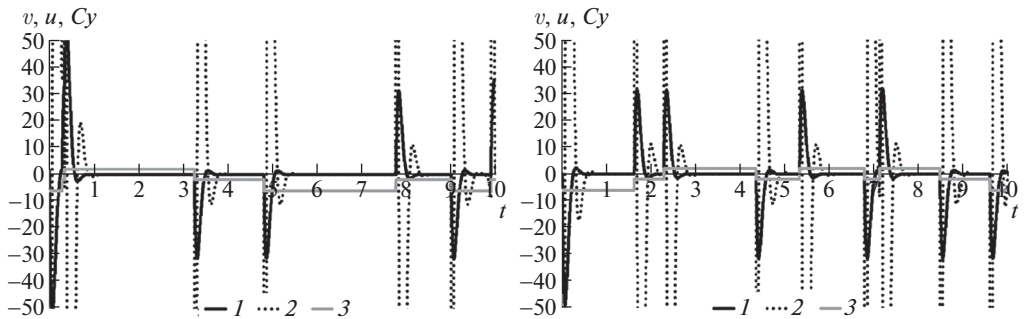


Рис. 4. Примеры траекторий скорости и управления для модели привода в идеальных условиях: 1 — скорости v_t^* , v_t^{**} , 2 — управления u_t^* , u_t^{**} , 3 — дрейф Cy_t .

При этом рисунки ожидаемо показывают, что если оптимальному управлению “не мешать” возмущениями и ограничениями на затраты энергии, то дрейф, определяемый скачками цепи y_t , отслеживается очень эффективно, в том числе в случае неполной информации при малых ошибках наблюдений оценка фильтрации очень незначительно ухудшает качество оптимального управления.

Для качественной оценки результатов расчета в “обычных” условиях, представленного в разделе 5, помогут следующие значения, вычисленные путем моделирования и осреднения по пучку из 1000 траекторий, примеры которых показаны на рис. 2–4: $J(U_0^T | u_t = u_t^*) = 10,9$, $J(U_0^T | u_t = u_t^{**}) = 11,3$, $J(U_0^T | u_t = 0) = 129,2$, дисперсия ошибки фильтрации $\mathbb{E} \left\{ |cy_t - c\hat{y}_t|^2 \right\}$, усредненная по $t \leq T$, равна 0,0025.

5. Численный пример. Управление траекторией устойчивой системы

В примере 4 вычисления выполнялись в “идеальных” условиях, поэтому получился результат, иллюстрирующий потенциал управления для модели

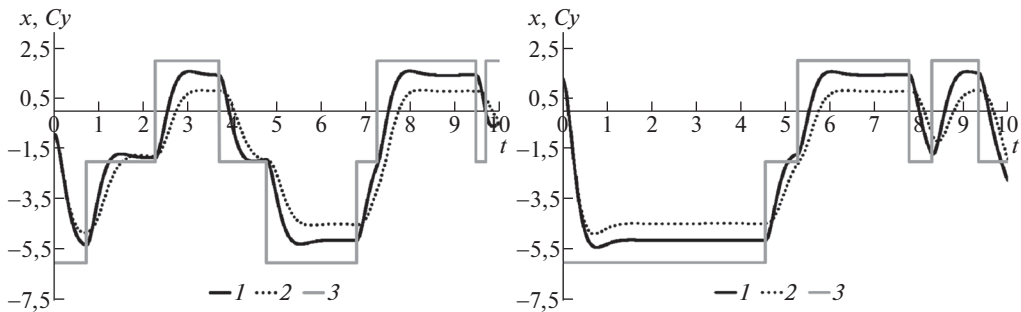


Рис. 5. Примеры траекторий положения для модели привода в обычных условиях: 1 — положение x_t^* , 2 — положение x_t^{**} , 3 — дрейф Cy_t .

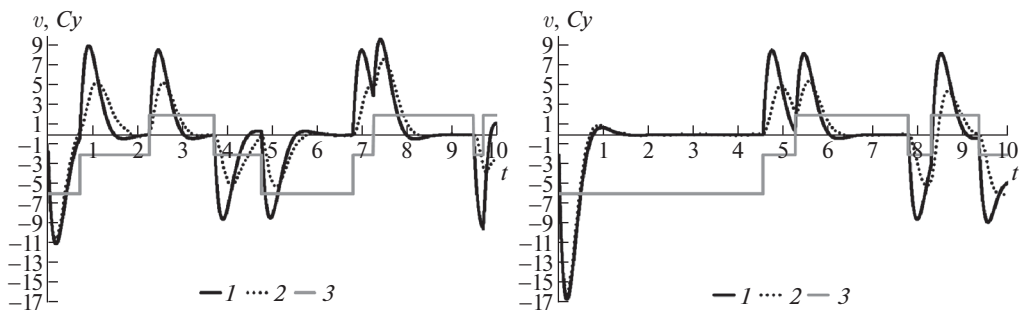


Рис. 6. Примеры траекторий скорости для модели привода в обычных условиях: 1 — скорость v_t^* , 2 — скорость v_t^{**} , 3 — дрейф Cy_t .

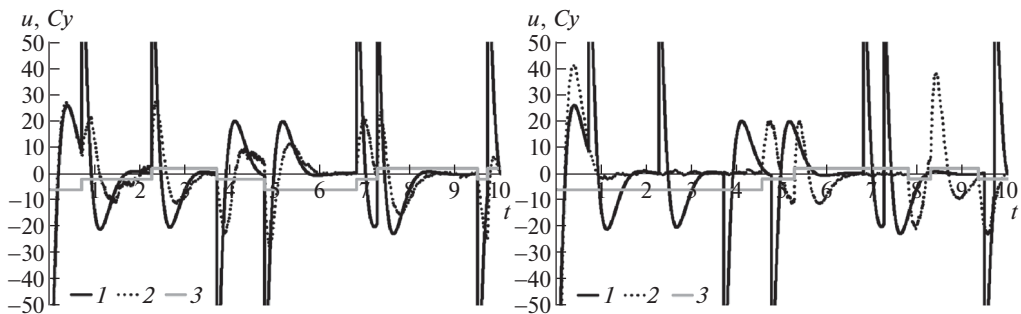


Рис. 7. Примеры траекторий управления для модели привода в обычных условиях: 1 — управление u_t^* , 2 — управление u_t^{**} , 3 — дрейф Cy_t .

механического привода (22). В данном разделе модельный расчет выполнен в более реалистичных предположениях и сравнение формальных характеристик качества управляемой и неуправляемой систем уже более содержательно.

Параметры предыдущего расчета уточнены величинами $g = 0,1$, $R = 0,001$, т.е. погрешность наблюдений и штраф за затраты энергии увели-

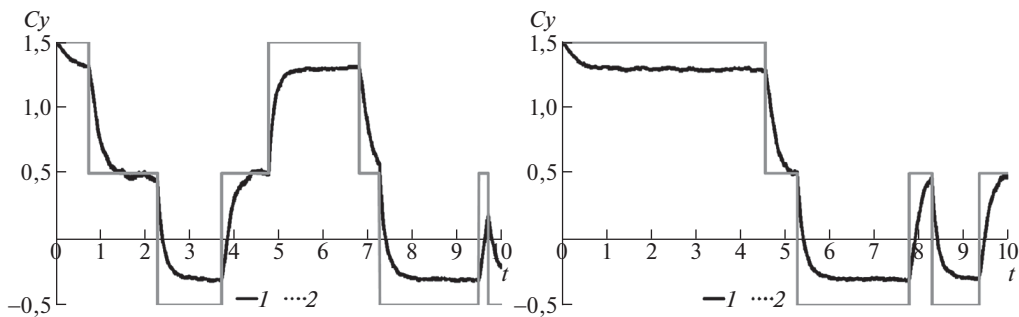


Рис. 8. Примеры траекторий цепи и оценки фильтрации: 1 — процесс cy_t , 2 — оценка $\hat{c}y_t$.

чены в 100 раз в сравнении с идеальными условиями раздела 4. Остальные параметры модели и расчета алгоритмов оставлены неизменными. Моделирование выполнено 1000 раз, т.е. приведенные характеристики качества управления и фильтрации получены путем осреднения соответствующих величин по пучку из 1000 траекторий.

Результаты расчета иллюстрируют рис. 5–7, представляющие по два примера типичных траекторий x_t^* и x_t^{**} , v_t^* и v_t^{**} , u_t^* и u_t^{**} соответственно. Оба набора примеров отвечают одним и тем же траекториям цепи y_t , целевой дрейф $C_t y_t$ повторен на каждом из рисунков.

Анализируя приведенные примеры, можно отметить, что:

- реализации траекторий положений привода на рис. 5 подтверждают возможность достаточно эффективно отслеживать дрейф Cy_t в условиях и полной, и неполной информации, но во втором случае разница в результатах уже более чувствительна, чем в предыдущем примере в “идеальных” условиях;
- примеры траекторий скоростей привода на рис. 6 объясняют лучшую результативность управления по полной информации тем, что оно более динамично отслеживает изменения состояния цепи y_t ;
- примеры траекторий управлений на рис. 7 дополняют предыдущий тезис тем, что демонстрируют тенденцию компенсации запаздывающей реакции u_t^{**} на смену состояния y_t более длительным интервалом активного воздействия; при этом если цепь прибывает в одном состоянии достаточно долго, то оба управления успевают стабилизировать привод около нужной величины дрейфа, а управляющие воздействия привести к нулю (это особенно хорошо видно в начальной фазе второго примера).

Отметим также, что максимальные абсолютные значения принимаются u_t^* в моменты изменения состояний цепи y_t и достигают значений порядка ± 100 , максимальные значения u_t^{**} меньше, порядка ± 30 , но их длительность, т.е. период активного управления, существенно больше.

Завершит анализ этого примера рис. 8, иллюстрирующий работу фильтра Вонэма для этих же реализаций y_t . Заметим, что цепь на рис. 8 представлена

сверткой cy_t , задающей силовое воздействие согласно (22), а не формируемым в результате этого воздействия дрейфом Cy_t .

Наконец, интегральные характеристики качества управлений в рассмотренном примере:

$$J(U_0^T | u_t = u_t^*) = 32,5, \quad J(U_0^T | u_t = u_t^{**}) = 43,0, \quad J(U_0^T | u_t = 0) = 129,6,$$

дисперсия ошибки фильтрации $\mathbb{E} \left\{ |cy_t - c\hat{y}_t|^2 \right\}$, усредненная по $t \leq T$, равна 0,077.

6. Пример управления траекторией неустойчивой системы

Вопрос этого раздела статьи — можно ли использовать полученное управление в системе (22), если она неустойчива, т.е. отказавшись от условий $b < 0$ и $b^2 + 4a < 0$. Не обсуждая физический смысл отказа от этих условий, отметим, что формальных трудностей для этого случая нет из-за линейного представления (22). Например, если в рассмотренном примере вместо $a = -0,25$, $b = -0,5$ положить $a = 0,25$, $b = 0,5$, то “вернуться” от неустойчивой системы с такими a, b к рассмотренному устойчивому варианту можно простой заменой v_t на $(-v_t)$, остальные переменные изменятся также зеркально, а значение целевого функционала останется прежним. Это подтверждает и практическая проверка — моделирование в условиях раздела 5, но с параметрами $a = 0,25$, $b = 0,5$, приводит с учетом зеркалирования к тем же результатам, что проиллюстрированы на рис. 5–7, оценки критериев $J(U_0^T | u_t = u_t^*)$ и $J(U_0^T | u_t = u_t^{**})$ совпадают. Из-за неустойчивости системы, как и следует из формальных соображений, существенно увеличивается только $J(U_0^T | u_t = 0)$, траектории x_t и v_t нестабильны, абсолютные значения растут. При этом скорость, с которой “разваливается” система не слишком велика, при достижении горизонта $T = 10$ значения величин положения и скорости оказываются порядка нескольких сотен, часть траекторий за время расчета вообще не успевают сформировать тенденцию к расхождению.

Однако, несмотря на успешное моделирование примера с $a = 0,25$, $b = 0,5$, потенциальные трудности при практической реализации управления неустойчивой системой (22) не представляются исчерпанными. Подтверждает это следующий пример, в котором применены параметры $a = 0,25$, $b = 5$, а все остальные оставлены неизменными. Сохранение величины a обеспечит этому примеру с учетом знака те же величины c и C , т.е. тот же порядок значений для целевого функционала, что даст возможность качественного сравнения этого примера с предыдущим. Выбор значения $b = 5$ приводит к тому, что разойтись за время $T = 10$ успеют все моделируемые траектории, именно: расчет показал, что по достижении горизонта все смоделированные траектории положения достигают величин порядка 10^{20} , 10^{21} , все траектории скорости — величин порядка 10^{21} , 10^{22} . При этом проблем с реализацией оптимального управления u_t^* нет, но возникли проблемы с реализацией управления u_t^{**} ,

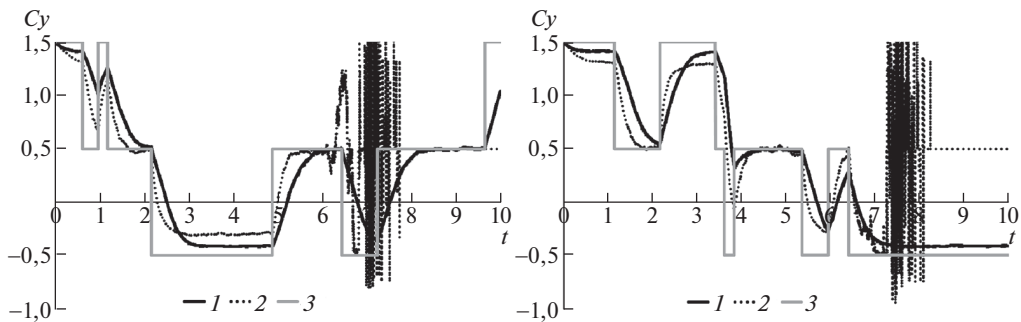


Рис. 9. Примеры траекторий цепи и оценки фильтрации: 1 — оценка $c\hat{y}_t$, 2 — оценка $c\check{y}_t$, 3 — процесс cy_t .

численная схема для которого также оказалась неустойчивой. Причиной этого оказалась неустойчивость использованной процедуры приближенного расчета оценки фильтра Вонэма \hat{y}_t . Вообще при большом объеме расчетов расходимость нескольких траекторий аппроксимаций \hat{y}_t методом Эйлера наблюдалась и в предыдущих расчетах, но такие траектории носили исключительный характер и в окончательных расчетах отсутствовали. Действительно, простая аппроксимация фильтра Вонэма (15) не гарантирует ни неотрицательности компонентам вектора \hat{y}_t , ни выполнения условия нормировки. О необходимости использования устойчивых схем для аппроксимации \hat{y}_t хорошо известно [15]. В данной статье использовалось решение, представленное в публикациях [16–18], а именно: аналитическая аппроксимация непрерывной системы наблюдения дискретной с ограничением учитываемого числа скачков и применением для приближения интегралов квадратур Гаусса. Соответствующая оценка обозначена \check{y}_t .

Сначала на рис. 9 проиллюстрированы результаты фильтрации.

Поведение траекторий \check{y}_t на рис. 9 подтверждает применимость выбранного метода в рассматриваемой задаче, а кроме того, иллюстрирует качественное отличие \check{y}_t от простой аппроксимации \hat{y}_t , состоящее в том, что траектории \check{y}_t выглядят гладкими в сравнении с траекториями \hat{y}_t , обладающими характерным “дрожанием”. вполне удовлетворительна и объективная характеристика качества \check{y}_t — дисперсия ошибки фильтрации $\mathbb{E} \left\{ |cy_t - c\check{y}_t|^2 \right\}$, усредненная по $t \leq T$, равна 0,11, что сопоставимо с погрешностью фильтра Вонэма в устойчивом случае раздела 5.

Главное, что видно из рис. 9 — это расхождение траекторий фильтра Вонэма. Примеры на рис. 9 характеризуют поведение всех без исключения смоделированных траекторий \hat{y}_t . Раньше или позже, но каждая из траекторий теряет устойчивость и в конечном итоге нарушает условие неотрицательности и нормировки. В представленном расчете, чтобы не допустить полное расхождение фильтра, выполнение этих условий проверялось дополнительно и при их нарушении оценке \hat{y}_t искусственно присваивалось значение предельного распределения $\pi_\infty = \left(\frac{1}{3}, \frac{1}{3}, \frac{1}{3} \right)'$, т.е. $c\hat{y}_t = 0,5$. Оказалось, что дальнейшие на-

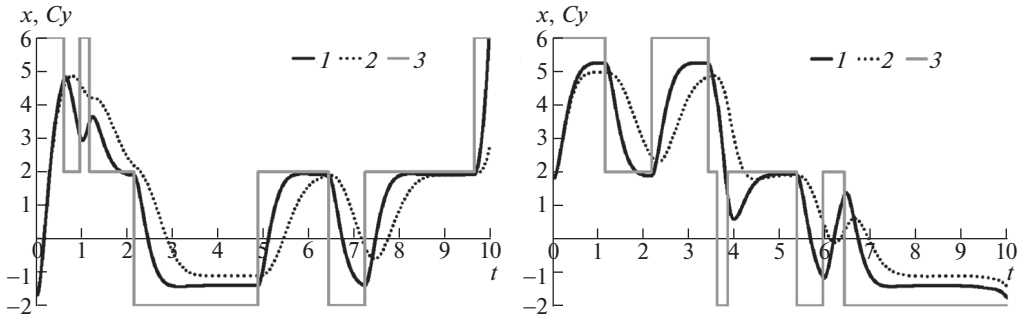


Рис. 10. Примеры траекторий положения для неустойчивой модели привода:
 1 — положение x_t^* , 2 — положение x_t^{**} , 3 — дрейф Cy_t .

блюдения на такую оценку больше не влияют, что на рис. 9 отвечает прямой линии на заключительной части траектории \hat{y}_t — превращение ее в тривиальную априорную оценку.

Результаты использования управлений u_t^* и u_t^{**} с заменой в u_t^{**} оценки \hat{y}_t на \check{y}_t иллюстрируют рис. 10.

Визуально эффективность u_t^* и u_t^{**} с точки зрения влияния на положение привода представляется аналогичной рассмотренному устойчивому примеру, т.е. рис. 5. Объективные характеристики качества управлений в рассмотренном примере принимают вполне ожидаемые значения: $J(U_0^T | u_t = u_t^*) = 35,1$, $J(U_0^T | u_t = u_t^{**}) = 48,4$. Величина $J(U_0^T | u_t = 0)$ имеет порядок 10^{41} , 10^{42} , т.е. практически бессмысленна.

7. Заключение

В постановке задачи, рассмотренной в данной статье, используются типовые понятия теории оптимального управления — линейная система управления, скачкообразный марковский процесс, квадратичный критерий, а при решении — традиционные методы — динамическое программирование, принцип разделения. Особенность обеспечивается конструктивной интерпретацией марковской цепи, которая фактически рассматривается как состояние стохастической системы, наблюдаемое прямо или косвенно, а не как сложное возмущение. Это позволяет придать нетипичное содержание квадратичному критерию и использовать его для решения задачи слежения. Действенность и хорошие прикладные перспективы полученных результатов подтверждаются численным экспериментом, использующим практически содержательную модель функционирования механического привода.

СПИСОК ЛИТЕРАТУРЫ

1. *Elliott R.J., Aggoun L., Moore J.B.* Hidden Markov Models: Estimation and Control. N.Y.: Springer-Verlag, 1995.
2. *Bar-Shalom Y., Willett P.K., Tian X.* Tracking and Data Fusion: a Handbook of Algorithms. Storrs, Conn.: YBS Publishing, 2011.

3. *Rishel R.* A Strong Separation Principle for Stochastic Control Systems Driven by a Hidden Markov Model // *SIAM J. Control and Optimization*. 1994. V. 32. No. 4. P. 1008–1020.
4. *Beneš V.* Quadratic Approximation by Linear Systems Controlled from Partial Observations / *Stochastic Analysis*. Mayer-Wolf E.; Merzbach E.; Shwartz A., Eds.; Academic Press, 1991. P. 39–50.
5. *Helmes K., Rishel R.* The Solution of a Partially Observed Stochastic Optimal Control Problem in Terms of Predicted Miss // *IEEE Trans. Autom. Control*. 1992. V. 37. No. 9. P. 1462–1464.
6. *Benes V., Karatzas I., Ocone D., Wang H.* Control with Partial Observations and an Explicit Solution of Mortensen's Equation // *Appl Math Optim*. 2004. No. 49. P. 217–239. <https://doi.org/10.1007/s00245-003-0788-0>
7. *Лунцер Р.Ш., Ширяев А.Н.* Статистика случайных процессов (нелинейная фильтрация и смежные вопросы). М.: Наука, 1974.
Liptser R.S., Shiryaev A.N. Statistics of random processes. II. Applications. Berlin: Springer-Verlag, 2001.
8. *Fleming W.H., Rishel R.W.* Deterministic and Stochastic Optimal Control. N.Y.: Springer-Verlag, 1975.
9. *Athans M., Falb P.L.* Optimal Control: An Introduction to the Theory and Its Applications. N.Y.–Sydney: McGraw-Hill, 1966.
10. *Босов А.В.* Задача управления линейным выходом нелинейной неуправляемой стохастической дифференциальной системы по квадратичному критерию // *Изв. РАН. Теория и системы управления*. 2021. № 5. С. 52–73.
Bosov A.V. The Problem of Controlling the Linear Output of a Nonlinear Uncontrollable Stochastic Differential System by the Square Criterion // *J. Computer and Systems Sciences International*. 2021. V. 60. No. 5. P. 719–739.
11. *Лунцер Р.Ш., Ширяев А.Н.* Теория мартингалов. М.: Наука, 1986.
12. *Миллер Б.М., Миллер Г.Б., Семенихин К.В.* Методы синтеза оптимального управления марковским процессом с конечным множеством состояний при наличии ограничений // *АиТ*. 2011. № 2. С. 111–130.
Miller B.M., Miller G.B., Semnikhin K.V. Methods to Design Optimal Control of Markov Process with Finite State Set in the Presence of Constraints // *Autom. Remote Control*. 2011. V. 72. No. 2. P. 323–341.
13. *Wonham W.M.* On the Separation Theorem of Stochastic Control // *SIAM J. Control*. 1968. V. 6. No. 2. P. 312–326.
14. *Фельдбаум А.А.* Основы теории оптимальных автоматических систем. 2-е изд., испр. и доп. М.: Наука, 1966.
15. *Yin G., Zhang Q., Liu Y.* Discrete-time Approximation of Wonham Filters // *J. Control Theory Applications*. 2004. No. 2. P. 1–10.
16. *Борисов А.В.* Численные схемы фильтрации марковских скачкообразных процессов по дискретизованным наблюдениям II: случай аддитивных шумов // *Информ. и ее примен.* 2020. Т. 14. № 1. С. 17–23.
17. *Борисов А.В.* L1-оптимальная фильтрация марковских скачкообразных процессов II: численный анализ конкретных схем // *АиТ*. 2020. № 12. С. 24–49.
Borisov A.V. L1-Optimal Filtering of Markov Jump Processes. II. Numerical Analysis of Particular Realizations Schemes // *Autom. Remote Control*. 2020. V. 81. No. 12. P. 2160–2180.

18. *Borisov A., Sokolov I.* Optimal Filtering of Markov Jump Processes Given Observations with State-Dependent Noises: Exact Solution and Stable Numerical Schemes // *Mathematics*. 2020. V. 8. No. 4 (506).

Статъа представена к публикации членом редколлегии Б.М. Миллером.

Поступила в редакцию 25.11.2021

После доработки 22.12.2021

Принята к публикации 30.12.2021