



Общероссийский математический портал

А. В. Босов, Управление линейным выходом марковской цепи по квадратичному критерию, *Информ. и её примен.*, 2021, том 15, выпуск 2, 3–11

DOI: 10.14357/19922264210201

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

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

Параметры загрузки:

IP: 18.97.14.89

19 февраля 2025 г., 13:49:18



УПРАВЛЕНИЕ ЛИНЕЙНЫМ ВЫХОДОМ МАРКОВСКОЙ ЦЕПИ ПО КВАДРАТИЧНОМУ КРИТЕРИЮ*

А. В. Босов¹

Аннотация: Решена задача оптимального управления выходом стохастической системы наблюдения, в которой состояние определяет ненаблюдаемый марковский скачкообразный процесс, а линейные наблюдения задаются системой дифференциальных уравнений Ито с винеровским процессом. В наблюдении аддитивно входит управление, так что формируется управляемый выход системы. Цель оптимизации задается квадратичным критерием общего вида. Для решения задачи управления сформулирована теорема разделения, использующая решение задачи оптимальной фильтрации, обеспечиваемое фильтром Вонэма. В результате разделения формируется эквивалентная задача управления выходом диффузионного процесса частного вида, а именно: с линейным сносом и нелинейной диффузией. Решение этой задачи обеспечивается непосредственным применением метода динамического программирования.

Ключевые слова: марковский скачкообразный процесс; стохастическая дифференциальная система Ито; оптимальное управление; квадратичный критерий; стохастическая фильтрация; фильтр Вонэма

DOI: 10.14357/19922264210201

1 Введение

В теории управления есть результаты, обладающие особой выразительностью, определяющие не частные решения, а принципы и концепции. К ним, безусловно, относится LQG-задача (LQG — linear-quadratic-Gaussian) — управление линейно-гауссовской стохастической системой по квадратичному критерию [1]. Для целей данной статьи особо важен результат LQG-управления в постановке с неполной информацией, известный как теорема разделения задач управления и фильтрации состояния [2]. Этот результат породил, по-видимому, наиболее действенный подход к синтезу управлений в нелинейных системах наблюдения, который называют принципом разделения.

Считается, что классическими методами, равно как и основанными на них приближенными методами поиска оптимальных управлений [3–5], хорошо решаются задачи с полной информацией о состоянии. Но более востребованными для практики представляются именно задачи с неполной информацией, постановки для систем наблюдения, прежде всего стохастических. Общая теория для таких систем и соответствующих задач основана на уравнении Дункана–Мортенсена–Закаи, описывающего эволюцию апостериорной плотности вероятности и уравнение динамического программирования в вариационных производных [6], развивался этот подход в [7–9]. Но получить опти-

мальные решения удается крайне редко. Один из примеров дает результат, полученный в [10] для модели управляемой марковской цепи, там же можно оценить сопровождающие такие решения технические трудности.

Ясно, что в задачах с неполной информацией ключевую роль играют методы стохастической фильтрации, поскольку результатами фильтрации заменяют значения фазовых координат в синтезированных по полной информации управлениях. Такая замена и есть принцип разделения, но если в LQG-задаче его реализация обеспечивается формальной теоремой разделения и приводит к оптимальному решению, то в общем случае объединение решений задач управления и фильтрации носит интуитивный характер, а вопрос о потерях качества при постулировании разделения остается открытым. Такие задачи, как в [10], единичны, и это обстоятельство делает ценными любые результаты в задачах оптимизации нелинейных стохастических систем по неполной информации о состоянии. Примеру такой задачи посвящена настоящая работа.

2 Постановка задачи управления выходом цепи

На каноническом вероятностном пространстве $(\Omega, \mathcal{F}, \mathcal{P}, \mathcal{F}_t)$, $t \in [0, T]$, рассмотрим стохастическую

* Работа выполнена при частичной поддержке РФФИ (проект 19-07-00187-А).

¹ Институт проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук, ABosov@frcsc.ru

систему наблюдения с вектором состояния y_t и связанным с ним линейно управляемым выходом $z_t^{(0)}$ (верхний индекс здесь и в других обозначениях введен для удобства дальнейших обозначений):

$$dy_t = \Lambda_t^T y_t dt + d\Lambda_t^y, \quad y_0 = Y; \quad (1)$$

$$dz_t^{(0)} = a_t^{(0)} y_t dt + b_t^{(0)} z_t^{(0)} dt + c_t^{(0)} u_t dt + \sigma_t^{(0)} dw_t, \quad z_0^{(0)} = Z^{(0)}. \quad (2)$$

Уравнение (1) определяет марковский скачкообразный процесс — цепь с конечным числом состояний и значениями в множестве $\{e_1, \dots, e_{n_y}\}$, состоящем из единичных координатных векторов в евклидовом пространстве \mathbb{R}^{n_y} . Предполагается, что начальное состояние Y имеет известное распределение π , Λ_t — матрица интенсивностей переходов и Λ_t^y — \mathcal{F}_t -согласованный мартингал с квадратичной характеристикой [11]

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

Уравнение (2) представляют косвенные наблюдения $z_t^{(0)} \in \mathbb{R}^{n_z}$ за состоянием цепи y_t и одновременно линейный управляемый выход системы. Порождаемую наблюдениями σ -алгебру будем обозначать $\mathcal{F}_t^{z^{(0)}}$ и предполагать, что

$$\mathcal{F}_t^{z^{(0)}} \subseteq \mathcal{F}_t \subseteq \mathcal{F}.$$

Далее, здесь $w_t \in \mathbb{R}^{n_w}$ — не зависящий от Λ_t^y , Y и $Z^{(0)}$ стандартный векторный винеровский процесс; $Z^{(0)} \in \mathbb{R}^{n_z}$ — гауссовская случайная величина с известными моментами, не зависящая от Λ_t^y и Y ; $u_t \in \mathbb{R}^{n_u}$, $u_t = U_t(z_t^{(0)})$, $U_t = U_t(z)$, $z \in \mathbb{R}^{n_z}$, — допустимое управление с обратной связью [3]. Для использования выхода $z_t^{(0)}$ в качестве наблюдений будем предполагать невырожденность ошибок наблюдений, т. е. $\sigma_t^{(0)}(\sigma_t^{(0)})^T > 0$.

Качество управления

$$U_0^T = \{U_t(z), 0 \leq t \leq T\}$$

определяется целевым функционалом следующего вида:

$$J(U_0^T) = \mathbb{E} \left\{ \int_0^T \left\| P_t y_t + Q_t^{(0)} z_t^{(0)} + R_t u_t \right\|_{S_t}^2 dt + \left\| P_T y_T + Q_T^{(0)} z_T^{(0)} \right\|_{S_T}^2 \right\}, \quad (3)$$

где $u_t = U_t(z_t^{(0)})$; $P_t \in \mathbb{R}^{n_y \times n_y}$; $Q_t^{(0)} \in \mathbb{R}^{n_y \times n_z}$; $R_t \in \mathbb{R}^{n_y \times n_u}$; $S_t \in \mathbb{R}^{n_y \times n_y}$, $S_t \geq 0$, $S_t = S_t^T$,

$0 \leq t \leq T$, — заданные ограниченные матричные функции. Весовая функция $\|x\|_S^2 = x^T S x$ для симметричной неотрицательно определенной матрицы S , единичной матрице $S = \mathbf{1}$ соответствует евклидова норма $\|x\|_1^2 = |x|^2$. Для исключения возможности отсутствия штрафа для отдельных компонентов вектора управления u_t , которая делает целевой функционал (3) физически некорректным, предполагается выполненным обычное условие невырожденности, в данных обозначениях принимающее вид:

$$R_t^T S_t R_t > 0.$$

Функционал (3) отражает одну из традиционных задач автоматического управления — регулирование выхода или управление по выходу. В [12] в качестве практического приложения такой постановки рассматривается задача «доведения выхода до нуля». В других практических целях функционал (3) можно использовать, например, для формализации задачи отслеживания выходом состояния $|y_t - z_t^{(0)}|^2$ или управлением — выхода $|z_t^{(0)} - u_t|^2$, учитывая при этом расходы на управляющее воздействие $|u_t|^2$ и/или значение выходной переменной $|z_t^{(0)}|^2$.

Матричные функции $a_t^{(0)} \in \mathbb{R}^{n_z \times n_y}$, $b_t^{(0)} \in \mathbb{R}^{n_z \times n_z}$, $c_t^{(0)} \in \mathbb{R}^{n_z \times n_u}$ и $\sigma_t^{(0)} \in \mathbb{R}^{n_z \times n_w}$ предполагаются ограниченными:

$$\left| a_t^{(0)} \right| + \left| b_t^{(0)} \right| + \left| c_t^{(0)} \right| + \left| \sigma_t^{(0)} \right| \leq C$$

для всех $0 \leq t \leq T$. Таким образом обеспечивается существование решения уравнения (2) для любого допустимого управления u_t . Более того, будем предполагать, что все используемые функции времени Λ_t , $a_t^{(0)}$, $b_t^{(0)}$, $c_t^{(0)}$, $\sigma_t^{(0)}$, P_t , $Q_t^{(0)}$, R_t и S_t кусочно-непрерывны, чтобы обеспечить выполнение типовых условий существования решений обыкновенных дифференциальных уравнений, получаемых далее.

Задачу составляет поиск u_t^* — допустимого управления, доставляющего минимум квадратичному функционалу $J(U_0^T)$:

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

в предположении существования этого минимума.

3 Разделение задач управления и фильтрации

Для решения поставленной задачи потребуется выполнить некоторые преобразования, которые

позволят показать наличие в рассматриваемой задаче двух важных свойств, необходимых для успешного решения. Заметим, что для системы наблюдения (1), (2) известно решение задачи фильтрации, а именно: уравнение для условного математического ожидания $\mathbb{E}\{y_t | \mathcal{F}_t^{z^{(0)}}\}$, описываемое фильтром Вонэма [11]. Однако записать этот фильтр здесь требуется так, чтобы показать отсутствие в задаче управления дуального эффекта, т. е. влияния выбора оценки состояния на качество управления [13]. Возможно, точнее эту первую промежуточную цель выполняемых преобразований выразит предложение показать независимость качества оптимальной фильтрации от реализуемого закона управления. Вторая цель — преобразовать целевой функционал (3) так, чтобы были выделены независимые слагаемые, определяющие по отдельности качество управления и качество оценивания.

Вначале предлагается выполнить замену переменных в (2), избавляясь от слагаемых $b_t^{(0)} z_t^{(0)} dt$ и $c_t^{(0)} u_t dt$. Для такой замены обозначим через $B_t \in \mathbb{R}^{n_z \times n_z}$ решение уравнения

$$dB_t = -B_t b_t^{(0)} dt,$$

т. е. матричную экспоненту

$$B_t = \exp \left\{ - \int_0^t b_s^{(0)} ds \right\},$$

и через $z_t^{(1)} \in \mathbb{R}^{n_z}$ линейное преобразование выхода

$$z_t^{(1)} = B_t z_t^{(0)} - \int_0^t B_s c_s^{(0)} u_s ds.$$

Далее, дифференцируя, получаем:

$$\begin{aligned} dz_t^{(1)} &= -B_t b_t^{(0)} z_t^{(0)} dt + \\ &+ B_t \left(a_t^{(0)} y_t dt + b_t^{(0)} z_t^{(0)} dt + c_t^{(0)} u_t dt + \sigma_t^{(0)} dw_t \right) - \\ &- B_t c_t^{(0)} u_t dt = B_t a_t^{(0)} y_t dt + B_t \sigma_t^{(0)} dw_t, \end{aligned}$$

или

$$dz_t^{(1)} = a_t^{(1)} y_t dt + \sigma_t^{(1)} dw_t \quad (4)$$

с дополнительными обозначениями $a_t^{(1)} = B_t a_t^{(0)}$, $\sigma_t^{(1)} = B_t \sigma_t^{(0)}$. Выполненные преобразования при этом не влияют на решение задачи фильтрации в том смысле, что

$$\mathbb{E} \left\{ y_t | \mathcal{F}_t^{z^{(0)}} \right\} = \mathbb{E} \left\{ y_t | \mathcal{F}_t^{z^{(1)}} \right\},$$

поскольку использованная для получения $z_t^{(1)}$ замена является линейным невырожденным преобразованием $z_t^{(0)}$. Таким образом, оценка оптимального

фильтра не зависит от u_t , и в качестве наблюдений можно использовать $z_t^{(1)}$, получая одну и ту же оценку состояния для любого допустимого управления. Соответственно, вместо задачи оценивания состояния по наблюдениям $z_t^{(0)}$, зависящим от реализуемого закона управления, можно рассматривать эквивалентную задачу оценивания y_t по наблюдениям $z_t^{(1)}$, описываемым уравнением (4) и не зависящим от u_t .

Обозначив оптимальную оценку

$$Y_t = \mathbb{E} \left\{ y_t | \mathcal{F}_t^{z^{(1)}} \right\},$$

запишем уравнение для нее, т. е. фильтр Вонэма [11]:

$$\begin{aligned} dY_t &= \Lambda_t^T Y_t dt + (\text{diag}(Y_t) - Y_t Y_t^T) \left(a_t^{(1)} \right)^T \times \\ &\times \left(\sigma_t^{(1)} \left(\sigma_t^{(1)} \right)^T \right)^{-1/2} \times \\ &\times \left(\sigma_t^{(1)} \left(\sigma_t^{(1)} \right)^T \right)^{-1/2} \left(dz_t^{(1)} - a_t^{(1)} Y_t dt \right), \\ Y_0 &= \mathbb{E}\{Y\}. \quad (5) \end{aligned}$$

Сделать это позволяет предположение о невырожденности $\sigma_t^{(1)} \left(\sigma_t^{(1)} \right)^T > 0$.

Однако использовать уравнение (5) в качестве уравнения состояния пока нельзя, так как требуется вторая замена, нужная для преобразования целевого функционала (3). Эта замена состоит во введении дополнительной переменной в вектор наблюдений, а именно: через $z_t^{(2)} \in \mathbb{R}^{n_z}$ обозначим $\int_0^t B_s c_s u_s ds$, т. е. дополним модель наблюдений уравнением

$$dz_t^{(2)} = B_t c_t u_t dt, \quad z_0^{(2)} = 0.$$

Новый вектор наблюдений обозначим

$$Z_t = \begin{pmatrix} z_t^{(1)} \\ z_t^{(2)} \end{pmatrix} \in \mathbb{R}^{n_z}, \quad n_Z = 2n_z.$$

Ясно, что такая замена ничего не изменит в отношении оценки оптимального фильтра, т. е.

$$\mathbb{E} \left\{ y_t | \mathcal{F}_t^{z^{(0)}} \right\} = \mathbb{E} \left\{ y_t | \mathcal{F}_t^{z^{(1)}} \right\} = \mathbb{E} \left\{ y_t | \mathcal{F}_t^{z^{(1)}, z^{(2)}} \right\}.$$

Уравнение для Z_t имеет вид:

$$dZ_t = \begin{pmatrix} a_t^{(1)} \\ \mathbf{0} \end{pmatrix} y_t dt + \begin{pmatrix} \mathbf{0} \\ B_t c_t \end{pmatrix} u_t dt + \begin{pmatrix} \sigma_t^{(1)} \\ \mathbf{0} \end{pmatrix} dw_t. \quad (6)$$

Обозначая блочные матрицы

$$a_t = \begin{pmatrix} a_t^{(1)} \\ \mathbf{0} \end{pmatrix}; \quad c_t = \begin{pmatrix} \mathbf{0} \\ B_t c_t \end{pmatrix}; \quad \sigma_t^{(2)} = \begin{pmatrix} \sigma_t^{(1)} \\ \mathbf{0} \end{pmatrix},$$

где $\mathbf{0}$ — нулевая матрица подходящей размерности, перепишем уравнение (6):

$$dZ_t = a_t y_t dt + c_t u_t dt + \sigma_t^{(2)} dw_t, \quad (7)$$

причем заметим, что

$$\begin{aligned} z_t^{(0)} &= B_t^{-1} \left(z_t^{(1)} + \int_0^t B_s c_s^{(0)} u_s ds \right) = \\ &= B_t^{-1} (z_t^{(1)} + z_t^{(2)}) = B_t^{-1} (\mathbf{1} \mathbf{1}) Z_t, \end{aligned} \quad (8)$$

где использовано обозначение $(\mathbf{1} \mathbf{1})$ для блочной матрицы, составленной из двух единичных $\mathbf{1} \in \mathbb{R}^{n_z \times n_z}$. Здесь использован тот факт, что матричная экспонента B_t невырожденная [14], что позволяет применять обратную матрицу B_t^{-1} . Полученное равенство (8) далее используется для замены переменных в целевом функционале (3).

Теперь можно получить уравнение для оптимальной оценки

$$Y_t = \mathbb{E} \{ y_t | \mathcal{F}_t^Z \},$$

используя уравнения (5) и (7). Записать фильтр Вонэма непосредственно для наблюдений Z_t нельзя, поскольку $\sigma_t^{(2)} (\sigma_t^{(2)})^T$ вырождена, но вопрос легко решается, поскольку запись (7) носит технический характер, а фактически нужно использовать наблюдения (4), которые легко получить из (7) с помощью блочной матрицы $I = (\mathbf{1} \mathbf{0})$, составленной из единичной и нулевой матриц $\mathbf{1}, \mathbf{0} \in \mathbb{R}^{n_z \times n_z}$. С учетом $\sigma_t^{(0)} (\sigma_t^{(0)})^T > 0$ и $\sigma_t^{(1)} = I \sigma_t^{(2)} = B_t \sigma_t^{(0)}$ для Y_t имеем:

$$\begin{aligned} dY_t &= \Lambda_t^T Y_t dt + (\text{diag}(Y_t) - Y_t Y_t^T) (I a_t)^T \times \\ &\quad \times \left(I \sigma_t^{(2)} (I \sigma_t^{(2)})^T \right)^{-1/2} \times \\ &\quad \times \left(I \sigma_t^{(2)} (I \sigma_t^{(2)})^T \right)^{-1/2} I (dZ_t - a_t Y_t dt), \\ Y_0 &= \mathbb{E} \{ Y \}, \end{aligned}$$

где $dW_t = (I \sigma_t^{(2)} (I \sigma_t^{(2)})^T)^{-1/2} I (dZ_t - a_t Y_t dt)$ определяет \mathcal{F}_t^Z -согласованный стандартный векторный винеровский процесс W_t [11]. Это обстоятельство позволяет переписать уравнение наблюдений (7) в виде:

$$I dZ_t = I a_t Y_t dt + \left(I \sigma_t^{(2)} (I \sigma_t^{(2)})^T \right)^{1/2} dW_t.$$

Здесь было учтено, что

$$I (dZ_t - a_t Y_t dt - c_t u_t dt) = I (dZ_t - a_t Y_t dt).$$

Окончательно, обозначив

$$\begin{aligned} \sigma_t &= I^T \left(I \sigma_t^{(2)} (I \sigma_t^{(2)})^T \right)^{1/2}; \\ \Sigma_t &= \Sigma_t(Y_t) = \\ &= (\text{diag}(Y_t) - Y_t Y_t^T) (I a_t)^T \left(I \sigma_t^{(2)} (I \sigma_t^{(2)})^T \right)^{-1/2}, \end{aligned}$$

получаем систему управления с полной информацией следующего вида:

$$\left. \begin{aligned} dY_t &= \Lambda_t^T Y_t dt + \Sigma_t(Y_t) dW_t, \quad Y_0 = \mathbb{E} \{ Y \}; \\ dZ_t &= a_t Y_t dt + c_t u_t dt + \sigma_t dW_t, \\ Z_0 &= Z = \begin{pmatrix} Z^{(0)} \\ \mathbf{0} \end{pmatrix}. \end{aligned} \right\} \quad (9)$$

Для окончательного формирования эквивалентной исходной задачи управления остается записать в выбранных переменных имеющийся целевой функционал (3). Обеспечивает это формула полного математического ожидания [15] и замена переменной $z_t^{(0)}$, заданная в (8):

$$\begin{aligned} J(U_0^T) &= \mathbb{E} \left\{ \int_0^T \left\| P_t (y_t - Y_t + Y_t) + \right. \right. \\ &\quad \left. \left. + Q_t^{(0)} B_t^{-1} (\mathbf{1} \mathbf{1}) Z_t + R_t u_t \right\|_{S_t}^2 dt + \right. \\ &\quad \left. + \left\| P_T (y_T - Y_T + Y_T) + Q_T^{(0)} B_T^{-1} (\mathbf{1} \mathbf{1}) Z_T \right\|_{S_T}^2 \right\} = \\ &= \mathbb{E} \left\{ \int_0^T \left\| P_t Y_t + Q_t Z_t + R_t u_t \right\|_{S_t}^2 + \right. \\ &\quad \left. + \left\| P_T Y_T + Q_T Z_T \right\|_{S_T}^2 + \int_0^T \left\| P_t (y_t - Y_t) \right\|_{S_t}^2 dt + \right. \\ &\quad \left. + \left\| P_T (y_T - Y_T) \right\|_{S_T}^2 \right\}, \quad (10) \end{aligned}$$

где учтено, что $Y_t = \mathbb{E} \{ y_t | \mathcal{F}_t^Z \}$ и дополнительно обозначено

$$Q_t = Q_t^{(0)} B_t^{-1} (\mathbf{1} \mathbf{1}).$$

Поскольку последнее слагаемое в выражении (10) не зависит от U_0^T , а характеризует только точность оптимальной оценки фильтрации Y_t , то его можно из целевого функционала исключить. Таким образом, новый целевой функционал имеет вид:

$$J(U_0^T) = \mathbb{E} \left\{ \int_0^T \|P_t Y_t + Q_t Z_t + R_t u_t\|_{S_t}^2 dt + \|P_T Y_T + Q_T Z_T\|_{S_T}^2 \right\}. \quad (11)$$

Формулировка окончательного утверждения такова.

Утверждение 1 (теорема разделения). Решение задачи оптимального управления с полной информацией ($\mathcal{F}_t^{Y,Z}$ -измеримого) системой (9) с целевым функционалом (11) является оптимальным решением задачи управления системой (1), (2) с косвенными наблюдениями ($\mathcal{F}_t^{z^{(0)}}$ -измеримого) с целевым функционалом (3).

Заметим, что управление $(U^*)_0^T$, минимизирующее функционал (11), минимизирует и (3), поэтому для целевой функции не имеет смысла вводить новое обозначение.

4 Решение задачи управления ВЫХОДОМ

Полученное далее решение основано на подходе, представленном в [16] в задаче управления выходом системой с состоянием, описываемым уравнением Ито, и работе [17], где проанализирован частный случай этой же задачи в случае коррелированных возмущений. Найти решение, как и в [16, 17], сформулированной выше задачи управления выходом Z_t формируемым состоянием Y_t , с квадратичной функцией цены $J(U_0^T)$ удастся классическим методом динамического программирования [3, 5]. Обозначив функцию Беллмана через

$$V_t = V_t(y, z), \quad y \in \mathbb{R}^{n_y}, \quad z \in \mathbb{R}^{n_z}, \quad \Sigma_t = \Sigma_t(y),$$

получим уравнение динамического программирования

$$\frac{\partial V_t}{\partial t} + \frac{1}{2} \text{tr} \left\{ \Sigma_t^T \frac{\partial^2 V_t}{\partial y^2} \Sigma_t + \sigma_t^T \frac{\partial^2 V_t}{\partial z^2} \sigma_t + 2 \Sigma_t^T \frac{\partial^2 V_t}{\partial y \partial x} \sigma_t \right\} +$$

$$+ \min_u \left\{ y^T \Lambda_t \frac{\partial V_t}{\partial y} + (a_t y + c_t u)^T \frac{\partial V_t}{\partial z} + \|P_t y + Q_t z + R_t u\|_{S_t}^2 \right\} = 0,$$

$$V_T = \|P_T y + Q_T z\|_{S_T}^2. \quad (12)$$

Существование решения уравнения (12) является достаточным условием оптимальности, оптимальное управление при этом — точка минимума соответствующего слагаемого. Нетрудно видеть, что при сделанном выше предположении $R_t^T S_t R_t > 0$ этот минимум доставляет оптимальное управление (в предположении существования решения (12)):

$$u_t^* = u_t^*(y, z) = -\frac{1}{2} (R_t^T S_t R_t)^{-1} \left(c_t^T \frac{\partial V_t}{\partial z} + 2 R_t^T S_t (P_t y + Q_t z) \right). \quad (13)$$

Подставляя u_t^* в (12) и перегруппировывая слагаемые, получаем:

$$\begin{aligned} & \frac{\partial V_t}{\partial t} + \frac{1}{2} \text{tr} \left\{ \Sigma_t^T \frac{\partial^2 V_t}{\partial y^2} \Sigma_t + \sigma_t^T \frac{\partial^2 V_t}{\partial z^2} \sigma_t + \right. \\ & \left. + 2 \Sigma_t^T \frac{\partial^2 V_t}{\partial y \partial z} \sigma_t \right\} + y^T \Lambda_t \frac{\partial V_t}{\partial y} + (y^T a_t^T - \\ & - (R_t^T S_t (P_t y + Q_t z))^T (R_t^T S_t R_t)^{-1} c_t^T) \frac{\partial V_t}{\partial z} + (P_t y + \\ & + Q_t z)^T (S_t - S_t R_t (R_t^T S_t R_t)^{-1} R_t^T S_t) (P_t y + \\ & + Q_t z) - \frac{1}{4} \left(\frac{\partial V_t}{\partial z} \right)^T c_t (R_t^T S_t R_t)^{-1} c_t^T \frac{\partial V_t}{\partial z} = 0. \quad (14) \end{aligned}$$

Рассматривая полученное уравнение, заметим, что линейное вхождение Z_t в уравнение наблюдений (9) и квадратичное в целевой функционал (11) позволяет предположить, что решение уравнения (14) может быть представлено в виде (векторная форма представления, предложенного в [16]):

$$V_t = V_t(y, z) = z^T \alpha_t z + z^T \beta_t(y) + \gamma_t(y), \quad (15)$$

что сводит поиск оптимального решения к поиску уравнений относительно симметричной матричной функции α_t , векторной функции $\beta_t(y)$ и скалярной функции $\gamma_t(y)$, причем явный вид функции $\gamma_t(y)$ для реализации оптимального управления не требуется, нужна только производная

$$\frac{\partial V_t}{\partial z} = 2\alpha_t z + \beta_t(y).$$

Представление функции Беллмана (15) можно упростить дальше, используя то, что слагаемое с производной $\partial V_t / \partial y$ в (14) содержит только множитель $y^T \Lambda_t$. Это позволяет предположить, что $\beta_t(y)$ является аффинным преобразованием y (адаптированная к рассматриваемой задаче форма, предложенная в [17]):

$$\beta_t(y) = \beta_t^{MN} y, \quad (16)$$

где матрица $\beta_t^{MN} \in \mathbb{R}^{n_z \times n_y}$. Выбранное здесь обозначение β_t^{MN} будет ясно из дальнейшего (см. далее (21), где уравнение для β_t^{MN} определяется матрицами M_t^β и N_t^β). Граничное условие при этом принимает вид:

$$\begin{aligned} V_T &= \|P_T y + Q_T z\|_{S_T}^2 = \\ &= z^T Q_T^T S_T Q_T z + 2z^T Q_T^T S_T P_T y + y^T P_T^T S_T P_T y, \end{aligned}$$

т. е.

$$\left. \begin{aligned} \alpha_T &= Q_T^T S_T Q_T; \\ \beta_T^{MN} &= 2Q_T^T S_T P_T; \\ \gamma_T(y) &= y^T P_T^T S_T P_T y, \end{aligned} \right\} \quad (17)$$

а оптимальное управление (13) окончательно:

$$\begin{aligned} u_t^* &= u_t^*(y, z) = \\ &= -\frac{1}{2} (R_t^T S_t R_t)^{-1} (c_t^T (2\alpha_t z + \beta_t^M y) + \\ &\quad + 2R_t^T S_t (P_t y + Q_t z)). \end{aligned}$$

Таким образом, u_t^* содержит два слагаемых: первый терм, линейный по z , т. е. по наблюдениям, точнее по переменной выхода; второй терм, линейный по y , т. е. по состоянию, точнее по оценке состояния, заданной фильтром Вонэма. Чтобы сказанное было верно, нужно получить уравнения для α_t, β_t^{MN} и $\gamma_t(y)$, показав, что сделанное предположение относительно функции V_t с учетом аффинности $\beta_t(y)$ позволяет решить уравнение (14). Для этого подставляем (15) в уравнение (14) и учитываем, что производные $\beta_t(y)$ согласно (16) равны

$$\frac{\partial^2 \beta_t(y)}{\partial y^2} = 0; \quad \frac{\partial \beta_t(y)}{\partial y} = (\beta_t^{MN})^T,$$

а также что

$$\begin{aligned} y^T \Lambda_t \frac{\partial V_t}{\partial y} &= \left(\frac{\partial V_t}{\partial y} \right)^T \Lambda_t^T y = z^T \beta_t^{MN} \Lambda_t^T y + y^T \Lambda_t \frac{\partial \gamma_t}{\partial y}; \\ \frac{\partial^2 V_t}{\partial y \partial z} &= (\beta_t^{MN})^T. \end{aligned}$$

После этой подстановки и незначительных преобразований получаются уравнения для α_t — коэф-

фициента при z^T и z , $\beta_t(y) = \beta_t^{MN} y$ — коэффициента при z^T и y , а также для $\gamma_t = \gamma_t(y)$ — оставшихся слагаемых — функций y :

$$\begin{aligned} \frac{d\alpha_t}{dt} - (M_t^\alpha \alpha_t + \alpha_t^T (M_t^\alpha)^T) + N_t^\alpha - \\ - \alpha_t^T c_t (R_t^T S_t R_t)^{-1} c_t^T \alpha_t = 0; \quad (18) \end{aligned}$$

$$\frac{d\beta_t^{MN}}{dt} y + \beta_t^{MN} \Lambda_t^T y + M_t^\beta y - N_t^\beta \beta_t^{MN} y = 0; \quad (19)$$

$$\begin{aligned} \frac{\partial \gamma_t}{\partial t} + \frac{1}{2} \text{tr} \left\{ \Sigma_t^T \frac{\partial^2 \gamma_t}{\partial y^2} \Sigma_t \right\} + \\ + \text{tr} \left\{ \sigma_t^T \alpha_t \sigma_t + \Sigma_t^T (\beta_t^{MN})^T \sigma_t \right\} + \\ + y^T \Lambda_t \frac{\partial \gamma_t}{\partial y} + M_t^\gamma = 0, \quad (20) \end{aligned}$$

где обозначено

$$\begin{aligned} M_t^\alpha &= Q_t^T S_t R_t (R_t^T S_t R_t)^{-1} c_t^T; \\ N_t^\alpha &= Q_t^T (S_t - S_t R_t (R_t^T S_t R_t)^{-1} R_t^T S_t) Q_t; \\ M_t^\beta &= 2 \left((a_t^T - P_t^T S_t R_t (R_t^T S_t R_t)^{-1} c_t^T) \alpha_t + \right. \\ &\quad \left. + P_t^T (S_t - S_t R_t (R_t^T S_t R_t)^{-1} R_t^T S_t) Q_t \right); \\ N_t^\beta &= Q_t^T S_t R_t (R_t^T S_t R_t)^{-1} c_t^T + \\ &\quad + \alpha_t c_t (R_t^T S_t R_t)^{-1} c_t^T; \\ M_t^\gamma &= M_t^\gamma(y) = \beta_t^T (a_t - c_t (R_t^T S_t R_t)^{-1} R_t^T S_t P_t) y + \\ &\quad + y^T P_t^T (S_t - S_t R_t (R_t^T S_t R_t)^{-1} R_t^T S_t) P_t y - \\ &\quad - \frac{1}{4} \beta_t^T c_t (R_t^T S_t R_t)^{-1} c_t^T \beta_t. \end{aligned}$$

Наконец, из (19) получаются уравнения для β_t^{MN} :

$$\frac{d\beta_t^{MN}}{dt} + \beta_t^{MN} \Lambda_t^T + M_t^\beta - N_t^\beta \beta_t^{MN} = 0. \quad (21)$$

Решаются полученные уравнения с граничными условиями (17). При этом уравнение (18) — это матричное уравнение Риккати для квадратной симметричной матрицы α_t . Сделанных выше предположений в отношении кусочной непрерывности коэффициентов этого уравнения и условия $R_t^T S_t R_t > 0$ достаточно для существования единственного неотрицательного решения для всех $0 \leq t \leq T$. Действительно, такое уравнение имеется в классической линейно-квадратичной задаче и, как известно [18], существует единственное оптимальное управление — линейное с обратной связью

по выходу Z_t , с коэффициентом усиления, описываемым этим уравнением Риккати. Соотношения (21) представляют собой систему обыкновенных линейных дифференциальных уравнений относительно элементов матрицы β_t^{MN} с кусочно-непрерывными коэффициентами, так что существование и единственность решения обеспечиваются обычной для линейных дифференциальных уравнений теоремой. Уравнение (20) для $\gamma_t(y)$ является линейным дифференциальным уравнением в частных производных параболического типа. Предполагается существование решения у этого уравнения, и это предположение интерпретируется как достаточное условие, такое же как использованное уравнение Беллмана (12). Подводит итог рассуждениям раздела следующее утверждение.

Утверждение 2 (управление выходом для случая полной информации). Если существует решение уравнения динамического программирования (12), то это решение может быть представлено в виде (15), (16) так, что коэффициенты α_t , $\beta_t(y) = \beta_t^M y$ и $\gamma_t(y)$ определяются уравнениями (18), (21) и (20) соответственно, а оптимальное управление

$$\begin{aligned} u_t^* &= u_t^*(Y_t, Z_t) = - \\ &= \frac{1}{2} (R_t^T S_t R_t)^{-1} (c_t^T (2\alpha_t Z_t + \beta_t^{MN} Y_t) + \\ &\quad + 2R_t^T S_t (P_t Y_t + Q_t Z_t)). \end{aligned} \quad (22)$$

5 Заключение

Принципиальный результат статьи представляет, во-первых, найденная явная форма (22) оптимального управления в рассматриваемой задаче, во-вторых, тот факт, что это оптимальное решение оформлено в виде управления с обратной связью, в котором ключевую роль играет решение вспомогательной задачи оптимальной фильтрации — условное математическое ожидание $Y_t = \mathbb{E}\{y_t | \mathcal{F}_t^Z\}$. Это и дало основание сформулировать основной результат, назвав его теоремой разделения. Принципиальное отличие полученного решения от классической LQG-задачи состоит в том, что измененная для разделения задача не похожа на исходную, а именно: мартингалное представление марковской цепи (1) заменяется совсем иным объектом — стохастическим уравнением Ито с винеровским процессом (9).

Таким образом, ключом к решению оказалось уникальное свойство фильтра Вонэма, представившего оценку марковской цепи по косвенным, зашумленным гауссовским шумом наблюдениям в виде классического уравнения Ито с винеровским процессом.

Литература

1. Athans M. The role and use of the stochastic linear-quadratic-Gaussian problem in control system design // IEEE T. Automat. Contr., 1971. Vol. 16. No. 6. P. 529–552.
2. Wonham W. M. On the separation theorem of stochastic control // SIAM J. Control, 1968. Vol. 6. No. 2. P. 312–326.
3. Флеминг У., Рушел Р. Оптимальное управление детерминированными и стохастическими системами / Пер. с англ. — М.: Мир, 1978. 316 с. (Fleming W. H., Rishel R. W. Deterministic and stochastic optimal control. — New York, NY, USA: Springer-Verlag, 1975. 222 p.).
4. Kushner H. J., Dupuis P. G. Numerical methods for stochastic control problems in continuous time. — New York, NY, USA: Springer-Verlag, 2001. 476 p.
5. Bertsekas D. P. Dynamic programming and optimal control. — Cambridge: Athena Scientific, 2017. 576 p.
6. Mortensen R. E. Stochastic optimal control with noisy observations // Int. J. Control, 1966. Vol. 4. No. 5. P. 455–464.
7. Davis M. H. A., Varaiya P. P. Dynamic programming conditions for partially observable stochastic systems // SIAM J. Control, 1973. Vol. 11. No. 2. P. 226–262.
8. Benes V. E., Karatzas I. On the relation of Zakai's and Mortensen's equations // SIAM J. Control Optim., 1983. Vol. 21. No. 3. P. 472–489.
9. Bensoussan A. Stochastic control of partially observable systems. — Cambridge: Cambridge University Press, 1992. 364 p.
10. Miller B. M., K. E. Avrachenkov, K. V. Stepanyan, G. B. Miller. The problem of optimal stochastic data flow control based upon incomplete information // Probl. Inf. Transm., 2005. Vol. 41. No. 2. P. 150–170.
11. Elliott G. J., Aggoun L., Moore J. B. Hidden Markov models: Estimation and control. — New York, NY, USA: Springer-Verlag, 1995. 382 p.
12. Athans M., Falb P. L. Optimal control: An introduction to the theory and its applications. — New York, NY, USA: Dover Publications, 2007. 879 p.
13. Фельдбаум А. А. Основы теории оптимальных автоматических систем. — 2-е изд. — М.: Наука, 1966. 624 с.
14. Bhatia R. Matrix analysis. — Graduate texts in mathematics ser. — New York, NY, USA: Springer-Verlag, 1997. Vol. 169. 349 p.
15. Ширяев А. Н. Вероятность. — 2-е изд. — М.: Наука, 1989. 640 с.
16. Босов А. В., Стефанович А. И. Управление выходом стохастической дифференциальной системы по квадратичному критерию. I. Оптимальное решение

методом динамического программирования // Информатика и её применения, 2018. Т. 12. Вып. 3. С. 99–106.

17. Босов А. В. О некоторых частных случаях в задаче управления выходом стохастической дифференциальной системы по квадратичному критерию // Ин-

форматика и её применения, 2021. Т. 15. Вып. 1. С. 11–17.

18. Девис М. Х. А. Линейное оценивание и стохастическое управление / Пер. с англ. — М.: Наука, 1984. 206 с. (Davis M. H. A. Linear estimation and stochastic control. — London: Chapman and Hall, 1977. 224 p.)

Поступила в редакцию 24.12.2020

LINEAR OUTPUT CONTROL OF MARKOV CHAINS BY THE QUADRATIC CRITERION

A. V. Bosov

Institute of Informatics Problems, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation

Abstract: The problem of optimal output control of a stochastic observation system, in which the state determines an unobservable Markov jump process and linear observations are given by a system of Ito differential equations with a Wiener process, is solved. Observations additively include control vector, so that a controlled output of the system is formed. The optimization goal is set by a general quadratic criterion. To solve the control problem, a separation theorem is formulated that uses the solution to the optimal filtering problem provided by the Wonham filter. As a result of the separation, an equivalent problem of output control of a diffusion process of a particular type, namely, with linear drift and nonlinear diffusion, is formed. The solution of this problem is provided by direct application of the dynamic programming method.

Keywords: Markov jump process; Ito stochastic differential system; optimal control; quadratic criterion; stochastic filtering; Wonham filter

DOI: 10.14357/19922264210201

Acknowledgments

This work was partially supported by the Russian Foundation for Basic Research (grant 19-07-00187-A).

References

- Athans, M. 1971. The role and use of the stochastic linear-quadratic-gaussian problem in control system design. *IEEE T. Automat. Contr.* 16(6):529–552.
- Wonham, W. M. 1968. On the separation theorem of stochastic control. *SIAM J. Control* 6(2):312–326.
- Fleming, W. H., and R. W. Rishel. 1975. *Deterministic and stochastic optimal control*. New York, NY: Springer-Verlag. 222 p.
- Kushner, H. J., and P. G. Dupuis. 2001. *Numerical methods for stochastic control problems in continuous time*. New York, NY: Springer-Verlag. 476 p.
- Bertsekas, D. P. 2017. *Dynamic programming and optimal control*. Cambridge: Athena Scientific. 576 p.
- Mortensen, R. E. 1966. Stochastic optimal control with noisy observations. *Int. J. Control* 4(5):455–464.
- Davis, M. H. A., and P. P. Varaiya. 1973. Dynamic programming conditions for partially observable stochastic systems. *SIAM J. Control* 11(2):226–262.
- Benes, V. E., and I. Karatzas. 1983. On the relation of Zakai's and Mortensen's equations. *SIAM J. Control Optim.* 21(3):472–489.
- Bensoussan, A. 1992. *Stochastic control of partially observable systems*. Cambridge: Cambridge University Press. 364 p.
- Miller, B. M., K. E. Avrachenkov, K. V. Stepanyan, and G. B. Miller. 2005. The problem of optimal stochastic data flow control based upon incomplete information. *Probl. Inf. Transm.* 41(2):150–170.
- Elliott, R. J., L. Aggoun, and J. B. Moore. 1995. *Hidden Markov models: Estimation and control*. New York, NY: Springer-Verlag. 382 p.
- Athans, M., and P. L. Falb. 2007. *Optimal control: An introduction to the theory and its applications*. New York, NY: Dover Publications. 879 p.
- Feldbaum, A. A. 1966. *Osnovy teorii optimal'nykh avtomaticheskikh sistem* [Foundations of theory of optimal automatic systems]. Moscow: Nauka. 624 p.

14. Bhatia, R. 1997. *Matrix analysis*. Graduate texts in mathematics ser. New York, NY: Springer-Verlag. Vol. 169. 349 p.
15. Shiryaev, A. N. 1996. *Probability*. New York, NY: Springer Verlag. 624 p.
16. Bosov, A. V., and A. I. Stefanovich. 2018. Upravlenie vykhodom stokhasticheskoy differentsial'noy sistemy po kvadratichnomu kriteriyu. I. Optimal'noe reshenie metodom dinamicheskogo programmirovaniya [Stochastic differential system output control by the quadratic criterion. I. Dynamic programming optimal solution]. *Informatika i ee Primeneniya — Inform. Appl.* 12(3):99–106.
17. Bosov, A. V. 2021. O nekotorykh chastnykh sluchayakh v zadache upravleniya vykhodom stokhasticheskoy differentsial'noy sistemy po kvadratichnomu kriteriyu [On some special cases in the problem of stochastic differential system output control by the quadratic criterion]. *Informatika i ee Primeneniya — Inform. Appl.* 15(1):11–17.
18. Davis, M. H. A. 1977. *Linear estimation and stochastic control*. London: Chapman and Hall. 224 p.

Received December 24, 2020

Contributor

Bosov Alexey V. (b. 1969) — Doctor of Science in technology, principal scientist, Institute of Informatics Problems, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation; AVBosov@ipiran.ru