© 2020 г. А.В. БОСОВ, д-р техн. наук (abosov@frccsc.ru)
(Институт проблем информатики Федерального исследовательского центра “Информатика и управление” РАН, Москва;
Московский авиационный институт)

ПРИМЕНЕНИЕ УСЛОВНО-ОПТИМАЛЬНОГО ФИЛЬТРА ДЛЯ СИНТЕЗА СУБОПТИМАЛЬНОГО УПРАВЛЕНИЯ В ЗАДАЧЕ ОПТИМИЗАЦИИ ВЫХОДА НЕЛИНЕЙНОЙ ДИФФЕРЕНЦИАЛЬНОЙ СТОХАСТИЧЕСКОЙ СИСТЕМЫ[1]

Предложено субоптимальное решение задачи управления для диффузионного процесса Ито и линейного управляемого выхода с квадратичным критерием качества для случая косвенных наблюдений за состоянием. Используется полученное ранее решение задачи с полной информацией, концепция разделения задач управления и фильтрации и метод условно-оптимальной фильтрации В.С. Пугачева. Предлагается альтернатива традиционному практическому подходу к синтезу субоптимального управления в задаче с неполной информацией, состоящему в формальной замене в решении состояния на его оценку. Вместо задачи оптимизации выхода, порождаемого исходной моделью дифференциального уравнения, в качестве состояния используется оценка условно-оптимального фильтра. Предложен вариант численной реализации предлагаемого алгоритма на основе метода Монте-Карло и компьютерного моделирования.

Ключевые слова: стохастическое дифференциальное уравнение; стохастическая дифференциальная система; оптимальное управление; стохастическая фильтрация; условно-оптимальная фильтрация; метод Монте-Карло.

DOI: 10.31857/S0005231020110033

1. Введение

Рассматриваемая в статье задача может быть отнесена к традиционной задаче оптимального управления дифференциальными стохастическими системами. Не претендуя на полный обзор, упомянем некоторые результаты и методы, важные в контексте данной статьи. Заметим, что рассматриваемая задача в постановке с полной информацией о состоянии решается классическими методами, так же как и основанными на них приближенными методами поиска оптимальных управлений (см., например, [1–3]). Наибольшей полнотой и результативностью обладает задача управления линейно-гауссовскими стохастическими системами по квадратичному критерию [4]. Классические модели и методы сохраняют актуальность и составляют источники вполне

[1] Работа выполнена при частичной поддержке Российского фонда фундаментальных исследований (проект 19-07-00187-А).

современных исследований. Например, метод динамического программирования применяется в [5] для сравнительно нового класса логико-динамических систем, квадратичный критерий для квазилинейных систем используется, например, в [6], традиционные описания нелинейных систем управления в пространстве состояний дополняются альтернативными методами, например спектральным [7].

Существенно меньшее внимание привлекают постановки, в которых доступность состояния ограничивается. Здесь имеются результаты для моделей, в которых известными предполагается лишь часть координат вектора состояния (пионерская публикация [8], также см., например, [9, 10]). И, конечно, основополагающее значение имеет теорема разделения задач управления и фильтрации состояния в классической линейно-гауссовской задаче с минимальными обобщениями на функционал [11] (см. также обзор [12]). Следует упомянуть также о результатах применения принципа разделения в задачах управления пучками траекторий детерминированных систем [13, 14]. Общая теория стохастического управления для случая косвенных наблюдений основана на уравнении Дункана–Мортенсена–Закаи, описывающего эволюцию апостериорной плотности вероятности и уравнение динамического программирования в вариационных производных [15], развивался этот подход в [16–18]. Исследования в этой области остаются актуальными – один из примеров современных работ в данной области это решение, полученное в [19] для модели управляемой марковской цепи.

Практически востребованными являются именно задачи с неполной информацией о состоянии, с косвенными наблюдениями. В этих задачах ключевую роль играют уже методы стохастической фильтрации. Задача фильтрации, конечно, имеет собственное значение как инструмент обработки результатов экспериментов, идентификации параметров математических моделей. Но важна и вспомогательная функция фильтров, оценками которых заменяют значения фазовых координат в синтезированных по полной информации управлениях. Используемые при этом оценки фильтрации, как правило, не являются оптимальными. Их характеризуют термином “субоптимальность”, который означает, что эти оценки, хотя и не являются оптимальными, но обладают точностью, достаточной для решения некоторого класса практических задач оценивания. Такой подход к управлению уместно охарактеризовать как применение принципа разделения, т.е. обособленного решения задач управления и оценивания состояний с последующим объединением решений, подсказываемого теоремой разделения [11]. При этом понятие субоптимальности оценок фильтрации уместно распространить и на управление. Разделение, однако, далеко не всегда имеет место, и вопрос о потерях качества при формальном разделении останется, по-видимому, открытым в обозримом будущем. И это обстоятельство делает актуальным любые альтернативные подходы к синтезу субоптимальных управлений в задачах оптимизации нелинейных стохастических систем.

Одна из таких задач и альтернативный формальному разделению подход рассматриваются в данной статье. Это задача оптимизации линейного выхода, порождаемого состоянием, моделью которого является нелинейное стохастическое дифференциальное уравнение Ито, по квадратичному критерию качества. Используемое оптимальное решение этой задачи [20] в данной статье анализируется в предположении отсутствия информации о состоянии и интерпретации имеющегося выхода как косвенных наблюдений. Свойства квадратичного критерия выглядят хорошим основанием для того, чтобы попытаться разделить задачи управления и фильтрации в стиле классического подхода [11]. Однако структура оценки оптимальной фильтрации [21] этого сделать не позволяет, что делает логичным отказ от использования уравнений оптимальной фильтрации. Применение же вместо оптимального фильтра оценки по методу условно-оптимальной фильтрации (УОФ) В.С. Пугачева [22] позволяет не только получить практически реализуемый подход к синтезу субоптимального управления в рассматриваемой постановке на основе разделения, но и предложить альтернативный подход, заменив исходную задачу задачей оптимизации выхода, порождаемого оценкой УОФ. Описание соответствующего алгоритма и варианта его численной реализации и составляют основную цель статьи.

Изложение организовано следующим образом. Используемый результат решения рассматриваемой задачи управления для случая полной информации кратко приведен в разделе 3, вопросам разделения и оптимальной фильтрации в задаче с косвенными наблюдениями посвящен раздел 4, в разделе 5 описан предлагаемый подход к синтезу субоптимального управления на основе УОФ. В разделе 2 кратко обсуждается предполагаемая область практического применения обсуждаемой постановки, в заключении сформулированы задачи ближайшей перспективы.

2. Перспективная область практического применения задачи управления выходом

Задачи стохастического оптимального управления изучаются давно и имеют множество прикладных интерпретаций. Сравнительно новую область применения обеспечило развитие информационных технологий (ИТ) и исследования математических моделей для информационно-телекоммуникационных систем. Прообразом современных исследований можно, видимо, считать публикации [23, 24], предложившие модели для описания функционирования телефонных сетей. Эти простые модели можно рассматривать как прототип для моделей на основе цепей Маркова [25, 19], используемых для описания процессов пересылки пакетов в компьютерных сетях, управляемых протоколами стека TCP/IP, в том числе и моделей протоколов на основе скачкообразных марковских процессов [26, 27]. Использование таких моделей существенно усложняется с ростом числа возможных состояний, поэтому применение в этой области нашли диффузионные аппроксимации [28, 29].

Но одними телекоммуникациями область ИТ не исчерпывается. К другому кругу относятся, например, задачи распределения ресурсов программных систем. Надо отметить, что понятие ИТ-ресурса имеет довольно специфическое содержание. Так говорят про сайты, базы данных и банки знаний, наконец, про вычислительные ресурсы, узлы распределенных систем, компоненты центров обработки данных и т.д. Не все такие ресурсы материальны, часть – виртуальны по сути (программа, сайт, виртуальная машина, банк данных, поисковый запрос и т.п.). Все такие виртуальные ресурсы, как нетрудно заметить, объединяет то, что реализуются они некоторыми программными средствами, системами: или сами являются элементами программ или ими обслуживаются. В рамках таких программных систем процедуры управления ресурсами реализуются повсеместно. Традиционные примеры – это менеджеры задач, оптимизаторы запросов в системах управления базами данных, системы балансировки нагрузки [30–32]. В качестве моделей для постановки задач оптимизации функционирования программных систем, конечно, наиболее естественными представляются системы массового обслуживания, т.е. уже упомянутые модели на основе цепей Маркова. Аргумент против этих моделей тот же – большие значения для числа состояний, принимающие в реальных задачах значения порядка сотен тысяч и более, например, если речь идет об обработке пользовательских запросов в часы пиковых нагрузок. К применению альтернативного описания процессов в таких задачах в форме диффузионной аппроксимации подталкивают хорошо известные классические результаты о сходимости марковских процессов к диффузионным [34]. А практическую применимость такого подхода к моделированию процессов в ИТ показывают, например, уже цитированные публикации [28, 29], инициированные тем же обстоятельством – значительным ростом числа состояний цепи из-за роста числа пакетов, обслуживаемых сетевым протоколом.

Таким образом, обозначая распределение ресурсов программной системы в качестве области для практического применения рассматриваемой в данной статье задачи оптимизации, можно обоснованно использовать в качестве модели состояния такой системы (переменой, характеризующей число пользователей, запросов, буферов памяти, процессорного времени и т.п.) диффузионный процесс, описываемый стохастическим дифференциальным уравнением Ито. При этом под объемом ресурсов можно понимать линейное выражение, связывающее состояние и переменную, определяющую активное действие программной системы – объем выделенных или освобожденных ресурсов. Наконец, надо заметить, что рассматриваемая задача порождена именно такой практической интерпретацией, представленной в [35] для случая дискретного времени.

3. Постановка и решение задачи управления выходом

Решение рассматриваемой в данной статье задачи управления линейным выходом стохастической дифференциальной системы по квадратичному критерию качества при наличии полной информации получено в [20] для скалярного случая. Используя здесь этот результат, отметим, что метод УОФ ограничений на размерности системных переменных не предъявляет, поэтому рассмотрение скалярной постановки не ограничит его применимость в дальнейшем, а для целей данной статьи скалярного случая достаточно. В обсуждаемой постановке задачи управления выходом используются описываемые уравнениями Ито переменная состояния y_t и связанный с ней линейно выход z_t:

$$dy_t = A_t\left(y_t\right)dt + \Sigma_t\left(y_t\right)dv_t, \quad y_0 = Y, \tag{1}$$

$$dz_t = a_t y_t dt + b_t z_t dt + c_t u_t dt + \sigma_t dw_t, \quad z_0 = Z, \tag{2}$$

где v_t, w_t — стандартные независимые винеровские процессы, начальные условия Y, Z — независимые друг от друга и от v_t, w_t случайные величины с конечным вторым моментом. Предполагается, что функции A_t, Σ_t удовлетворяют условиям Ито [3]

$$\begin{aligned}
&|A_t(y)| + |\Sigma_t(y)| \le C(1+|y|) \quad \text{для всех} \quad 0 \le t \le T, \quad y \in \mathbb{R}^1,\\
&|A_t(y_1) - A_t(y_2)| + |\Sigma_t(y_1) - \Sigma_t(y_2)| \le C|y_1 - y_2|\\
&\qquad\qquad\qquad\qquad\qquad \text{для всех} \quad 0 \le t \le T \quad \text{и} \quad y_1, y_2 \in \mathbb{R}^1,
\end{aligned}$$

обеспечивающим существование единственного решения уравнения (1), функции a_t, b_t, c_t, σ_t являются ограниченными, процесс управления u_t – допустимым неупреждающим [3]. Предположение наличия полной информации о состоянии y_t и выходе z_t означает, что допустимое управление ищется в классе $\mathcal{F}_t^{y,z}$-измеримых неупреждающих функций (далее через $\mathcal{F}_t^x$ обозначается σ-алгебра, порожденная компонентами x_s, $0 \le s \le t$), обеспечивающих существование решения (2).

Оптимизируется квадратичный целевой функционал вида:

$$(3) \quad J\left(U_0^T\right) = \mathrm{E}\left\{\int_0^T \Big(S_t(s_t y_t - g_t z_t - h_t u_t)^2 + G_t z_t^2 + H_t u_t^2\Big)dt + \right.$$
$$\left. + S_T(s_T y_T - g_T z_T)^2 + G_T z_T^2\right\}, \quad U_0^T = \{u_t,\, 0 \le t \le T\},$$

где S_t, G_t, H_t – неотрицательные ограниченные функции.

Такой функционал отражает практическое содержание обсуждаемой выше задачи распределения ресурсов. Так, (3) позволяет ставить задачи отслеживания выходом состояния $(y_t - z_t)^2$ или управлением выхода $(z_t - u_t)^2$, учитывая при этом расходы на управляющее воздействие u_t^2 и/или значение выходной переменной z_t^2.

Решение задачи дает метод динамического программирования. Удается показать, что функция Беллмана имеет вид $V_t(y,z) = \alpha_t z^2 + \beta_t(y)z + \gamma_t(y)$. Задача, таким образом, сводится к выводу уравнений для коэффициентов α_t, $\beta_t(y)$, $\gamma_t(y)$, который выполнен в [20]. Именно: для коэффициента α_t получено уравнение Риккати, для коэффициентов $\beta_t(y)$, $\gamma_t(y)$ – линейные дифференциальные уравнения в частных производных второго порядка параболического типа, при этом оптимальным оказалось управлением с обратной связью

$$(4) \quad u_t^* = u_t^*(y_t, z_t) = -\frac{1}{2}\big(S_t h_t^2 + H_t\big)^{-1}(c_t(2\alpha_t z_t + \beta_t(y_t)) + 2S_t(s_t y_t - g_t z_t)h_t).$$

Постановка рассматриваемой далее задачи включает те же элементы (1)–(3) с тем отличием, что допустимыми считаются управления из класса $\mathcal{F}_t^z$-измеримых неупреждающих функций, т.е. в отношении переменной состояния y_t отсутствует точная информация, а выходная переменная z_t интерпретируется как косвенное наблюдение.

4. Управление на основе принципа разделения

В данном разделе дополнительно используется обозначение $\mathrm{E}\{\bullet\,|\mathcal{F}_t^x\}$ – условное математическое ожидание относительно $\mathcal{F}_t^x$.

Согласно классической теореме разделения в линейно-гауссовской стохастической системе [11] задачу управления можно решать в два этапа: сначала решить задачу с полной информацией, определившись с законом управления в форме обратной связи от текущего состояния, затем решить задачу фильтрации и заменить полученной оценкой состояние в решении задачи управления по полной информации. В рассматриваемой здесь задаче имеется результат для случая полной информации $u_t^* = u_t^*(y_t, z_t)$, оценка фильтрации пусть будет обозначена через $\widetilde{y}_t = \mathrm{E}\{y_t\,|\mathcal{F}_t^z\}$, а соответствующее управление $u_t^s = u_t^*(\widetilde{y}_t, z_t)$ следует интерпретировать как субоптимальное решение рассматриваемой задачи с неполной информацией, полученное согласно принципу разделения (индекс s принят от английского separation). Обозначение $\widetilde{y}_t$ для оптимальной оценки вместо традиционного $\widehat{y}_t$ использовано здесь с целью сохранить последнее для обозначения оценки УОФ. При этом задача фильтрации, т.е. вычисления $\widetilde{y}_t$, не совсем "отделена" от задачи управления, так как имеется зависимость оценки $\widetilde{y}_t$ от реализуемого закона управления, поскольку u_t явно входит в уравнение (2) для z_t. Более того, это делает возможным наличие в задаче дуального эффекта [36], т.е. влияния закона управления на качество оценивания состояния. Этот аспект в задаче оценивания состояния (1) по наблюдениям (2) требует отдельного обсуждения. В рамках обсуждения, во-первых, предлагается выполнить замену переменных в (2), избавляясь от слагаемых $b_t z_t dt$ и $c_t u_t dt$. Для выполнения замены обозначим

$$B_t = \exp\left\{-\int_0^t b_s ds\right\} \quad \text{и} \quad \widetilde{z}_t = B_t z_t - \int_0^t B_s c_s u_s ds.$$

Учитывая далее, что $dB_t = -b_t B_t dt$, получаем $d\widetilde{z}_t = a_t B_t y_t dt + \sigma_t B_t dw_t$ или

$$d\widetilde{z}_t = \widetilde{a}_t y_t dt + \widetilde{\sigma}_t dw_t \tag{5}$$

с дополнительными обозначениями $\widetilde{a}_t = a_t B_t$, $\widetilde{\sigma}_t = \sigma_t B_t$. Выполненные преобразования при этом не влияют на решение задачи фильтрации в том смысле, что $\widetilde{y}_t = \mathrm{E}\{y_t\,|\mathcal{F}_t^z\} = \mathrm{E}\left\{y_t\,\middle|\mathcal{F}_t^{\widetilde{z}}\right\}$, поскольку использованная для получения $\widetilde{z}_t$ замена является линейным невырожденным преобразованием z_t.

Таким образом, $\widetilde{y}_t$ не зависит от u_t, и в качестве наблюдений можно использовать $\widetilde{z}_t$, получая одну и ту же оценку состояния для любого допустимого управления. Соответственно вместо задачи оценивания состояния по наблюдениям z_t, зависящим от реализуемого закона управления, можно рассматривать эквивалентную задачу оценивания y_t по наблюдениям $\widetilde{z}_t$, описываемым уравнением (5) и не зависящим от u_t.

Второй обсуждаемый аспект разделения – это вычисление оптимальной оценки состояния системы (1) по наблюдениям (5). Наложенные выше ограничения обеспечивают существование этой оценки и принципиальную возможность использования для нее общих уравнений нелинейной фильтрации на основе обновляющих процессов [21], имеющих в рассматриваемом случае вид

$$d\widetilde{y}_t = \widetilde{A}_t dt + \frac{\widetilde{\Sigma}_t}{\widetilde{\sigma}_t} d\widetilde{v}_t, \quad \widetilde{y}_0 = \mathrm{E}\left\{Y\right\}, \tag{6}$$

где

$$\widetilde{A}_t = \mathrm{E}\left\{A_t\left(y_t\right) \,\Big|\, \mathcal{F}_t^{\widetilde{z}}\right\}, \quad \widetilde{\Sigma}_t = \widetilde{a}_t \mathrm{E}\left\{y_t\left(y_t - \widetilde{y}_t\right) \,\Big|\, \mathcal{F}_t^{\widetilde{z}}\right\},$$
$$d\widetilde{v}_t = \frac{1}{\widetilde{\sigma}_t}\Big(d\widetilde{z}_t - \widetilde{a}_t\widetilde{y}_t dt\Big).$$

Обновляющий процесс $\widetilde{v}_t$ является стандартным винеровским относительно $\mathcal{F}_t^{\widetilde{z}}$.

Разделяя задачи, следует рассматривать (6) как уравнение состояния, учитывая, что наблюдения (5) переписывается в виде

$$d\widetilde{z}_t = \widetilde{a}_t\widetilde{y}_t dt + \widetilde{\sigma}_t d\widetilde{v}_t, \quad \widetilde{z}_0 = Z. \tag{7}$$

Таким образом, имеются уравнение состояния (6) для переменной $\widetilde{y}_t$ и линейное уравнение наблюдения (7) для переменной $\widetilde{z}_t$, которые можно использовать для постановки эквивалентной задачи оптимизации линейного выхода по полной информации. Поддерживает это и возможность представления целевого функционала (3) в виде

$$J\left(U_0^T\right) = \mathrm{E}\left\{\int_0^T \Big(S_t\left(s_t\widetilde{y}_t - g_t z_t - h_t u_t\right)^2 + G_t z_t^2 + H_t u_t^2\Big)\, dt + \right.$$
$$\left. + S_T\left(s_T\widetilde{y}_T - g_T z_T\right)^2 + G_T z_T^2 \right\} +$$
$$+ \mathrm{E}\left\{\int_0^T S_t s_t^2\left(y_t - \widetilde{y}_t\right)^2 dt + S_T s_T^2\left(y_t - \widetilde{y}_T\right)^2\right\}.$$

Здесь оставлена без замены переменная $z_t = B_t^{-1}\left(\widetilde{z}_t + \int_0^t B_s c_s u_s ds\right)$, которую не трудно выполнить, дополнив выход вспомогательной переменной, но главное – это выделено второе слагаемое, определяющее вклад в целевой функционал ошибки оценивания, который можно исключить из эквивалентной оптимизационной постановки.

Перечисленное исчерпывает обсуждение принципа разделения в рассматриваемой задаче, поскольку, даже сохранив структуру наблюдений и целевого функционала, получить уравнение оценки в дифференциальной форме (1) не удастся, а значит, отсутствует главное условие разделения – готовое решение задачи управления по полной информации. Единственным результатом обсуждения можно считать предположение о целесообразности использования в рассматриваемой задаче управления оценки фильтрации и если эта оценка будет неоптимальной, то результат управления будет тем лучше, чем ближе оценка к оптимальной. В отношении же управления u_t^s остается заметить, что оно было бы оптимальным для задачи с полной информацией с состоянием y_t вида $dy_t = A_t(y_t)\,dt + \frac{\Sigma_t(y_t)}{\sigma_t} dv_t$, выходом (2) и целевым функционалом (3). Такое уравнение состояния можно интерпретировать как нулевое приближение решения задачи оптимальной фильтрации (6).

5. Управление на основе условно-оптимального фильтра

Основной проблемой реализации принципа разделения в рассматриваемой задаче все-таки представляются сложности реализации оптимальной оценки $\widetilde{y}_t$. Отказаться от нее предлагается в пользу решения задачи оценивания методом УОФ В.С. Пугачева [22]. Условно-оптимальная оценка $\widehat{y}_t$ состояния системы (1) по наблюдениям (5), не зависящим от выбора допустимого управления, ищется в виде

$$(8) \qquad d\widehat{y}_t = \widehat{\alpha}_t \xi_t(\widehat{y}_t, \widetilde{z}_t)\,dt + \widehat{\beta}_t \zeta_t(\widehat{y}_t, \widetilde{z}_t)\,dz_t + \widehat{\gamma}_t dt, \quad \widehat{y}_0 = \mathrm{E}\{Y\},$$

где $\widehat{\alpha}_t,\ \widehat{\beta}_t,\ \widehat{\gamma}_t$ – ограниченные функции времени, выполняющие роль параметров, функции $\xi_t(y,z)$ и $\zeta_t(y,z)$ – заданные структурные функции, выбираемые из эмпирических соображений. Будем предполагать, что для обеспечения условий существования решения (8) на структурные функции ξ_t, ζ_t наложены те же ограничения, что и в (1), т.е.

$$|\xi_t(y,z)| + |\zeta_t(y,z)| \le C(1 + |y| + |z|) \quad \text{для всех} \quad 0 \le t \le T, \quad y, z \in \mathbb{R}^1,$$
$$|\xi_t(y_1,z_1) - \xi_t(y_1,z_1)| + |\zeta_t(y_1,z_1) - \zeta_t(y_2,z_2)| \le C(|y_1 - y_2| + |z_1 - z_2|)$$
$$\text{для всех} \quad 0 \le t \le T \quad \text{и} \quad y_1, y_2, z_1, z_2 \in \mathbb{R}^1.$$

Рекомендации в части выбора структурных функций для рассматриваемой модели наблюдения можно, например, сформулировать на основании структуры оптимальной оценки (6), а именно положить

$$\xi_t = A_t(\widehat{y}_t) - \frac{\Sigma_t(\widehat{y}_t)}{\widetilde{\sigma}_t^2} \widetilde{a}_t \widehat{y}_t$$

и

$$\zeta_t = \frac{\Sigma_t(\widehat{y}_t)}{\widetilde{\sigma}_t^2},$$

т.е. искать оценку УОФ в виде

$$(9) \quad d\widehat{y}_t = \widehat{\alpha}_t \left(A_t(\widehat{y}_t) - \frac{\Sigma_t(\widehat{y}_t)}{\widetilde{\sigma}_t^2} \widetilde{a}_t \widehat{y}_t \right) dt + \widehat{\beta}_t \frac{\Sigma_t(\widehat{y}_t)}{\widetilde{\sigma}_t^2} d\widetilde{z}_t + \widehat{\gamma}_t dt, \quad \widehat{y}_0 = \mathrm{E}\{Y\}.$$

При этом коэффициенты $\widehat{\alpha}_t,\ \widehat{\beta}_t,\ \widehat{\gamma}_t$ УОФ выбираются так, что обеспечивают оценке фильтра (9) несмещенность и гарантированное качество, лучшее на некотором классе допустимых фильтров [22]. Такой выбор структурных функций обосновывается представлением оценки УОФ (9) в виде

$$d\widehat{y}_t = \left(\widehat{\alpha}_t A_t\left(\widehat{y}_t\right) + \widehat{\gamma}_t\right) dt + \frac{\Sigma_t\left(\widehat{y}_t\right)}{\widetilde{\sigma}_t}\left(\frac{1}{\widetilde{\sigma}_t}\left(\widehat{\beta}_t d\widetilde{z}_t - \widehat{\alpha}_t \widetilde{a}_t \widehat{y}_t dt\right)\right),$$

воспроизводящем структуру оптимального фильтра (6). Другое представление (9)

$$\begin{aligned} d\widehat{y}_t = &\left(\widehat{\alpha}_t A_t\left(\widehat{y}_t\right) + \left(\widehat{\beta}_t - \widehat{\alpha}_t\right)\frac{\Sigma_t\left(\widehat{y}_t\right)}{\widetilde{\sigma}_t^2}\widetilde{a}_t\widehat{y}_t + \widehat{\gamma}_t\right) dt + \\ &+ \widehat{\beta}_t\frac{\Sigma_t\left(\widehat{y}_t\right)}{\widetilde{\sigma}_t}\left(\frac{1}{\widetilde{\sigma}_t}\left(d\widetilde{z}_t - \widetilde{a}_t\widehat{y}_t dt\right)\right) \end{aligned} \tag{10}$$

позволяет предложить замену в рассматриваемой задаче управления уравнению состояния, положив

$$dy_t = \left(\widehat{\alpha}_t A_t\left(y_t\right) + \left(\widehat{\beta}_t - \widehat{\alpha}_t\right)\frac{\Sigma_t\left(y_t\right)}{\widetilde{\sigma}_t^2}\widetilde{a}_t y_t + \widehat{\gamma}_t\right) dt + \widehat{\beta}_t\frac{\Sigma_t\left(y_t\right)}{\widetilde{\sigma}_t} dv_t \tag{11}$$

при сохранении выхода (2) и целевого функционала (3) и предполагая наличие полной информации. Такое представление уравнения состояния основано на аппроксимации процесса $d\widehat{v}_t = \frac{1}{\widetilde{\sigma}_t}\left(d\widetilde{z}_t - \widetilde{a}_t\widehat{y}_t dt\right)$ стандартным винеровским процессом v_t. Последнее точно имеет место в случае оптимального фильтра $\widetilde{y}_t$, а значит, можно рассчитывать на качественную аппроксимацию в случае, когда оценка УОФ $\widehat{y}_t$ близка к оптимальной $\widetilde{y}_t$.

Таким образом, по аналогии с управлением u_t^s получается еще один вариант неформального применения принципа разделения – субоптимальное управление u_t^c вида $u_t^c = u_t^*(\widehat{y}_t, z_t)$, где $\widehat{y}_t$ – оценка УОФ (10), u_t^* определено соотношением (4) – решением исходной задачи оптимизации с полной информацией для состояния y_t, заданного уравнением (11), т.е. в уравнениях для коэффициентов α_t, $\beta_t\left(y\right)$ вместо $A_t\left(y_t\right)$ используется $\widehat{\alpha}_t A_t\left(y_t\right) + \left(\widehat{\beta}_t - \widehat{\alpha}_t\right) \times$ $\times \frac{\Sigma_t(y_t)}{\widetilde{\sigma}_t^2}\widetilde{a}_t y_t + \widehat{\gamma}_t$, а вместо $\Sigma_t\left(y_t\right)$ используется $\widehat{\beta}_t\frac{\Sigma_t(y_t)}{\widetilde{\sigma}_t}$.

Нетрудно видеть, что для использования имеющегося решения (4) требуется, чтобы коэффициенты $\widehat{\alpha}_t$, $\widehat{\beta}_t$, $\widehat{\gamma}_t$ в уравнении (11) были заданы. Соответствующие соотношения для УОФ приведены в [22]. Соотношения представляют собой комбинации моментных характеристик состояния y_t, наблюдений $\widetilde{z}_t$, структурных функций ξ_t, ζ_t и не зависят, как показано выше, от реализуемого закона управления u_t, т.е. могут быть вычислены заранее, отдельно от расчетов в целях управления.

6. Численная реализация управления u_t^c

Подведем итог рассуждениям раздела 5, перечислив формальные шаги, выполняемые для реализации варианта управления u_t^c для субоптимального решения исходной задачи оптимизации, включающей уравнение (1) ненаблюдаемого состояния, уравнение (2) выхода, псевдонаблюдения (5) и целевой функционал (3), минимизируемый на классе $\mathcal{F}_t^z$-измеримых неупреждающих управлений.

Шаг 1. Рассматривается вспомогательная задача УОФ для системы наблюдения и фильтра вида:

$$\begin{gathered}
dy_t = A_t\left(y_t\right)dt + \Sigma_t\left(y_t\right)dv_t, \quad y_0 = Y,\\
d\widetilde{z}_t = \widetilde{a}_t y_t dt + \widetilde{\sigma}_t dw_t, \quad z_0 = Z,\\
d\widehat{y}_t = \widehat{\alpha}_t \xi_t\left(\widehat{y}_t\right)dt + \widehat{\beta}_t \zeta_t\left(\widehat{y}_t\right)d\widetilde{z}_t + \widehat{\gamma}_t dt, \quad \widehat{y}_0 = \mathrm{E}\left\{Y\right\},\\
\xi_t\left(\widehat{y}_t\right) = A_t\left(\widehat{y}_t\right) - \frac{\Sigma_t\left(\widehat{y}_t\right)}{\widetilde{\sigma}_t^2}\widetilde{a}_t\widehat{y}_t, \quad \zeta_t\left(\widehat{y}_t\right) = \frac{\Sigma_t\left(\widehat{y}_t\right)}{\widetilde{\sigma}_t^2},
\end{gathered} \tag{12}$$

где обозначено $\widetilde{a}_t = a_t B_t$, $\widetilde{\sigma}_t = \sigma_t B_t$, $B_t = \exp\left\{-\int_0^t b_s ds\right\}$.

Решение – коэффициенты $\widehat{\alpha}_t$, $\widehat{\beta}_t$, $\widehat{\gamma}_t$ УОФ определяются системой уравнений (это частный случай решения задачи УОФ для системы наблюдения общего вида [22])

$$\left\{\begin{array}{l}
\widehat{\gamma}_t = \mathrm{E}\left\{A_t\right\} - \widehat{\alpha}_t\mathrm{E}\left\{\xi_t\right\} - \widehat{\beta}_t\widetilde{a}_t\mathrm{E}\left\{y_t\right\},\\
\widehat{\beta}_t = \widetilde{a}_t\mathrm{E}\left\{\left(y_t - \widehat{y}_t\right)y_t\zeta_t\right\}/\widetilde{\sigma}_t^2\mathrm{E}\left\{\zeta_t^2\right\},\\
\widehat{\alpha}_t = \dfrac{\mathrm{E}\left\{\begin{array}{c}
\left(y_t - \widehat{y}_t\right)\left(\dfrac{\partial\xi_t}{\partial t} + \dfrac{1}{2}\dfrac{\partial^2\xi_t}{\partial\left(\widehat{y}_t\right)^2}(\widehat{\beta}_t\zeta_t\widetilde{\sigma}_t)^2\right) +\\
+\left(y_t - \widehat{y}_t\right)\dfrac{\partial\xi_t}{\partial\widehat{y}_t}\left(\mathrm{E}\left\{A_t\right\} + \widehat{\beta}_t\widetilde{a}_t\overline{\zeta_t y_t}\right) -\\
-\dfrac{\partial\xi_t}{\partial\widehat{y}_t}(\widehat{\beta}_t\zeta_t\widetilde{\sigma}_t)^2 + \xi_t\left(\overline{A}_t - \widehat{\beta}_t\widetilde{a}_t\overline{\zeta_t y_t}\right)
\end{array}\right\}}{\mathrm{E}\left\{\xi_t\overline{\xi}_t - \left(y_t - \widehat{y}_t\right)\frac{\partial\xi_t}{\partial\widehat{y}_t}\overline{\xi}_t\right\}},
\end{array}\right. \tag{13}$$

где $A_t = A_t\left(y_t\right)$, $\overline{A}_t = A_t - \mathrm{E}\left\{A_t\right\}$, $\xi_t = \xi_t\left(\widehat{y}_t\right)$, $\overline{\xi}_t = \xi_t - \mathrm{E}\left\{\xi_t\right\}$, $\zeta_t = \zeta_t\left(\widehat{y}_t\right)$, $\overline{\zeta_t y_t} = \zeta_t y_t - \mathrm{E}\left\{\zeta_t y_t\right\}$.

В качестве численной реализации вычислений (13) предлагается использовать метод Монте-Карло и компьютерное моделирование. Для этого интервал управления $[0, T]$ следует разбить на отрезки равной (для простоты) малой длины Δt и смоделировать пучок траекторий для y_t и $\widetilde{z}_t$, заменив уравнения (12) их разностным аналогом, используя любую (предпочтительно явную) схему численного интегрирования [37]. Для вычисления коэффициентов $\widehat{\alpha}_t$, $\widehat{\beta}_t$, $\widehat{\gamma}_t$ в (13) в правой части использовать уже вычисленные $\widehat{\alpha}_{t-\Delta t}$, $\widehat{\beta}_{t-\Delta t}$, $\widehat{\gamma}_{t-\Delta t}$, операции $\mathrm{E}\left\{\bullet\right\}$ заменять их статистическим аналогом, для чего использовать смоделированный пучок траекторий переменных y_t, $\widetilde{z}_t$ и вычисленные с помощью $\widehat{\alpha}_{t-\Delta t}$, $\widehat{\beta}_{t-\Delta t}$, $\widehat{\gamma}_{t-\Delta t}$ приближенные оценки $\widehat{y}_t$.

Шаг 2. Рассматривается вспомогательная задача управления с целевым функционалом (3), наблюдаемым состоянием и линейным выходом вида

$$dy_t = \left(\widehat{\alpha}_t A_t(y_t) + \left(\widehat{\beta}_t - \widehat{\alpha}_t\right)\frac{\Sigma_t(y_t)}{\widetilde{\sigma}_t^2}\widetilde{a}_t y_t + \widehat{\gamma}_t\right)dt + \widehat{\beta}_t \frac{\Sigma_t(y_t)}{\widetilde{\sigma}_t} dv_t, \tag{14}$$
$$dz_t = a_t y_t dt + b_t z_t dt + c_t u_t dt + \sigma_t dw_t,$$

в которой используются вычисленные на шаге 1 коэффициенты $\widehat{\alpha}_t$, $\widehat{\beta}_t$, $\widehat{\gamma}_t$. Решение – коэффициенты α_t, $\beta_t(y)$, формирующие закон управления $u_t^* = u_t^*(y_t, z_t)$ согласно (4) вычисляются любым численным методом [38].

Шаг 3. Формируется расширенная модель для исходной задачи, включающая уравнение состояния (1), линейный выход (2) и псевдонаблюдения (6). Предложенный вариант управления u_t^c вычисляется по формуле (4) $u_t^c = u_t^*(\widehat{y}_t, z_t)$, где используются параметры управления α_t, $\beta_t(y)$, вычисленные на шаге 2, и оценка УОФ $\widehat{y}_t$ по наблюдениям $\widetilde{z}_t$ с коэффициентами $\widehat{\alpha}_t$, $\widehat{\beta}_t$, $\widehat{\gamma}_t$, вычисленными на шаге 1.

Заметим, что в [38] имеется значительный иллюстративный материал и обсуждаются различные практические детали реализации данного алгоритма в постановке с полной информацией.

7. Заключение

В статье на примере решенной задачи оптимизации линейного выхода нелинейной дифференциальной системы по квадратичному критерию обсуждается возможность приближенного решения аналогичной задачи для случая неполной информации о состоянии. На основе концепции разделения задач управления и фильтрации предложено два варианта субоптимального управления: путем формального разделения задач и на основании альтернативного представления переменной состояния, использующего метод условно-оптимальной фильтрации состояний стохастических дифференциальных систем наблюдения В.С. Пугачева. Любой из предложенных вариантов потребует существенных усилий для численной реализации, детализации и адаптации приведенного в статье принципиального алгоритма расчета управления u_t^c и значительных вычислительных ресурсов. Практическая реализация описанных алгоритмов и апробация их для оптимизации функционирования программных систем, хотя бы в рамках модельных экспериментов, – ближайшая перспектива.

СПИСОК ЛИТЕРАТУРЫ

1. *Kushner H.J., Dupuis P.G.* Numerical methods for stochastic control problems in continuous time. N.Y.: Springer-Verlag, 2001.
2. *Bertsekas D.P.* Dynamic programming and optimal control. Cambridge: Athena Scientific, 2013.
3. *Флеминг У., Ришел Р.* Оптимальное управление детерминированными и стохастическими системами / Пер. с англ. М.: Мир, 1978.

   Fleming W.H., Rishel R.W. Deterministic and stochastic optimal control. N.Y.: Springer-Verlag, 1975.

4. *Athans M.* Editorial on the LQG Problem // IEEE Trans. Automat. Control. 1971. V. 16. No. 6. P. 528–552.

5. *Бортаковский А.С.* Достаточные условия оптимальности управления переключаемыми системами // Изв. РАН. Теория и системы управления. 2017. № 4. С. 86–103.

   Bortakovskii A.S. Sufficient Optimality Conditions for Controlled Switched Systems // J. Comput. Syst. Sci. 2017. V. 56. No. 4. P. 636–651.

6. *Хрусталёв М.М., Онегин Е.Е.* Необходимые и достаточные условия в задаче оптимальной стабилизации квазилинейных стохастических систем // АиТ. 2019. No. 7. С. 89–104.

   Khrustalev M.M., Onegin E.E. Necessary and Sufficient Conditions for Optimal Stabilization of Quasi-linear stochastic Systems // Autom. Remote Control. 2019. V. 80. No. 7. P. 1252–1264.

7. *Пантелеев А.В., Рыбаков К.А.* Синтез оптимальных нелинейных стохастических систем управления спектральным методом // Информатика и ее применения. 2011. Т. 5. Вып. 2. С. 69–81.

8. *Fleming W.H.* Stochastic Control of Partially Observable Diffusions // SIAM J. Control. 1968. V. 6. No. 2. P. 194–214.

9. *Хрусталев М.М.* Синтез оптимальных и устойчивых управляемых стохастических систем при неполной информации о состоянии на неограниченном интервале времени // АиТ. 2011. No. 11. С. 174–190.

   Khrustalev M.M. Optimal and Stable Controllable Stochastic Systems Synthesis with Incomplete State Information on an Unbounded Time Interval // Autom. Remote Control. 2011. V. 72. No. 11. P. 2379–2394.

10. *Пантелеев А.В., Рыбаков К.А.* Приближенный синтез оптимальных непрерывных стохастических систем управления с неполной обратной связью // АиТ. 2018. No. 1. С. 130–146.

    Panteleev A.V., Rybakov K.A. Optimal Continuous Stochastic Control Systems with Incomplete Feedback: Approximate Synthesis // Autom. Remote Control. 2018. V. 79. No. 1. P. 103–116.

11. *Wonham W.M.* On the Separation Theorem of Stochastic Control // SIAM J. Control. 1968. V. 6. No. 2. P. 312–326.

12. *Georgiou T.T., Lindquist A.* The Separation Principle in Stochastic Control, redux // IEEE Trans. Automat. Control. 2013. V. 58. No. 10. P. 2481–2494.

13. *Бортаковский А.С., Немыченков Г.И.* Оптимальное в среднем управление детерминированными переключаемыми системами при наличии дискретных измерений // Изв. РАН. Теория и системы управления. 2019. No. 1. С. 52–77.

    Bortakovskii A.S., Nemychenkov G.I. Optimal in the Mean Control of Deterministic Switchable Systems Given Discrete Inexact Measurements // J. Comput. Syst. Sci. 2019. V. 58. No. 1. P. 50–74.

14. *Давтян Л.Г., Пантелеев А.В.* Метод параметрической оптимизации нелинейных непрерывных систем совместного оценивания и управления //Изв. РАН. Теория и системы управления. 2019. No. 3. С. 34–47.

    Davtyan L.G., Panteleev A.V. Method of Parametric Optimization of Nonlinear Continuous Systems of Joint Estimation and Control // J. Comput. Syst. Sci. 2019. V. 58. No. 3. C. 360–373.

15. *Mortensen R.E.* Stochastic Optimal Control with Noisy Observations // Int. J. Control. 1966. V. 4. No. 5. P. 455–464.

16. *Bensoussan A.* Stochastic control of partially observable systems. Cambridge: Cambridge University Press, 1992.

17. *Davis M.H.A., Varaiya P.P.* Dynamic Programming Conditions for Partially Observable Stochastic Systems // SIAM J. Control. 1973. V. 11. No. 2. P. 226–262.

18. *Benes V.E., Karatzas I.* On the Relation of Zakai's and Mortensen's Equations // SIAM J. Control Optim. 1983. V. 21. No. 3. P. 472–489.

19. *Миллер Б.М., Авраченков К.Е., Степанян К.В., Миллер Г.Б.* Задача оптимального стохастического управления потоком данных по неполной информации // Пробл. передачи информ. 2005. Т. 41. No. 2. С. 89–110.

   Miller B.M., Avrachenkov K.E., Stepanyan K.V., Miller G.B. The Problem of Optimal Stochastic Data Flow Control Based Upon Incomplete Information // Problems Inform. Transmission. 2005. V. 41. No. 2. P. 150–170.

20. *Босов А.В., Стефанович А.И.* Управление выходом стохастической дифференциальной системы по квадратичному критерию. I. Оптимальное решение методом динамического программирования // Информатика и ее применения. 2018. Т. 12. Вып. 3. С. 99–106.

21. *Липцер 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.

22. *Пугачев В.С., Синицын И.Н.* Стохастические дифференциальные системы. Анализ и фильтрация. М.: Наука, 1990.

   Pugachev V.S., Sinitsyn I.N. Stochastic differential systems. Analysis and filtering. Chichester–N.Y.: Wiley & Sons, 1987.

23. *Gilbert E.N.* Capacity of a Burst-noise Channel // Bell Syst. Tech. J. 1960. V. 39. P. 1253–1265.

24. *Elliott E.O.* Estimates of Error Rates for Codes on Burst-noise Channels // Bell Syst. Tech. J. 1963. V. 42. P. 1977–1997.

25. *Altman E., Avrachenkov K., Barakat C.* TCP in Presence of Bursty Losses // Perform. Evaluation. 2000. V. 42. P. 129–147.

26. *Borisov A., Bosov A., Miller G.* Modeling and Monitoring of RTP Link on the Receiver Side // Lect. Notes Comput. Sci. 2015. V. 9247. P. 229–241.

27. *Борисов А.В.* Применение методов оптимальной фильтрации для оперативного оценивания состояний сетей массового обслуживания // АиТ. 2016. No. 2. С. 115–141.

   Borisov A.V. Application of Optimal Filtering Methods for On-line of Queueing Network States // Autom. Remote Control. 2016. V. 77. No. 2. P. 277–296.

28. *Whitt W.* Stochastic-process limits. An introduction to stochastic-process limits and their application to queues. N.Y.: Springer, 2002.

29. *Bohacek S.* A Stochastic Model of TCP and Fair Video Transmission // 22nd Annual Joint Conf. of the IEEE Computer and Communications (INFOCOM) Proc. IEEE, 2003. V. 2. P. 1134–1144.

30. *Таненбаум Э.С., Вудхалл А.С.* Операционные системы. Разработка и реализация / Пер. с англ. 3-е изд. СПб.: Питер, 2007.

   Tanenbaum A.S., Woodhull A.S. Operating systems: design and implementation. 3rd ed. Upper Saddle River. NJ: Prentice Hall, 2006.

31. *Дейт К.Дж.* Введение в системы баз данных / Пер. с англ. 8-е изд. М.: Вильямс, 2005.

Date C.J. An introduction to database systems. 8th ed. Reading–MA: Addison-Wesley, 2004.

32. *Elsässer R., Monien B., Preis R.* Diffusion Schemes for Load Balancing on Heterogeneous Networks // Theor. Comput. Syst. 2002. V. 35. No. 3. P. 305–320.

33. *Welzl M.* Network congestion control. N.Y.: Wiley, 2005.

34. *Боровков А.А.* Асимптотические методы в теории массового обслуживания. М.: Наука, 1980.

Borovkov A.A. Asymptotic methods in queuing theory. N.Y.: Wiley, 1984.

35. *Босов А.В.* Управление линейным выходом дискретной стохастической системы по квадратичному критерию // Изв. РАН. Теория и системы управления. 2016. No. 3. С. 19–35.

Bosov A.V. Discrete Stochastic System Linear Output Control with Respect to a Quadratic Criterion // J. Comput. Syst. Sci. 2016. V. 55. No. 3. P. 349–364.

36. *Фельдбаум А.А.* Основы теории оптимальных автоматических систем. Изд. 2, испр. и доп. М.: Наука, 1966.

37. *Kloden P.E., Platen E.* Numerical solution of stochastic differential equations. Berlin–Heidelberg: Springer-Verlag, 1992.

38. *Босов А.В., Стефанович А.И.* Управление выходом стохастической дифференциальной системы по квадратичному критерию. IV. Альтернативное численное решение // Информатика и ее применения. 2020. Т. 14. Вып. 14. С. 24–30.

Статья представлена к публикации членом редколлегии А.И. Кибзуном.

Поступила в редакцию 02.03.2020

После доработки 20.05.2020

Принята к публикации 09.07.2020