

Стохастические системы

© 2019 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),
Г.Б. МИЛЛЕР, канд. физ.-мат. наук (gmiller@frccsc.ru)
(Институт проблем информатики Федерального исследовательского
центра “Информатика и управление” РАН, Москва)

УСЛОВНО-МИНИМАКСНЫЙ НЕЛИНЕЙНЫЙ ФИЛЬТР И СИГМА-ТОЧЕЧНЫЕ ФИЛЬТРЫ: ПРАКТИЧЕСКИЙ АНАЛИЗ И СРАВНЕНИЕ¹

Представлены результаты анализа и сравнения свойств двух концепций в задачах фильтрации состояний нелинейных стохастических динамических систем наблюдения с дискретным временем — сигма-точечного фильтра Калмана, основанного на дискретной аппроксимации непрерывных распределений, и условно-минимаксного нелинейного фильтра, реализующего метод условно-оптимальной фильтрации на основе имитационного моделирования. Краткое обсуждение структуры и свойств оценок и обоснований соответствующих алгоритмов сопровождается значительным объемом модельных примеров, иллюстрирующих как положительные варианты применения, так и ограничения работоспособности процедур оценивания. Простота и наглядность рассмотренных примеров (скалярные автономные регрессии в уравнении состояния и линейные наблюдения) позволяют объективно охарактеризовать рассматриваемые методики оценивания. Предлагается новая модификация нелинейного фильтра, сочетающая идеи обоих рассмотренных направлений.

Ключевые слова: нелинейная стохастическая система наблюдения, сигма-точечное преобразование, сигма-точечный фильтр, условно-оптимальная фильтрация, условно-минимаксный нелинейный фильтр, имитационное моделирование.

DOI: 10.1134/S000523101907002X

1. Введение

Стохастическая фильтрация, оценивание состояний и идентификация параметров динамических систем наблюдения — важные и для множества приложений, и для фундаментальной теории задачи, привлекающие значительный исследовательский интерес на протяжении многих лет. Самостоятельное место в этих исследованиях занимают нелинейные модели и алгоритмы фильтрации. В силу того что нелинейные системы крайне трудны для изучения

¹ Работа выполнена при финансовой поддержке Российского фонда фундаментальных исследований (проект № 19-07-00187-а).

общими строгими методами, в этой области допускается определенная математическая свобода, заключающаяся в возможности неформального обоснования предлагаемых решений соображениями разумности, инженерными пояснениями и приближенными соотношениями. Полученные таким образом результаты в целом определяются понятием “субоптимальная фильтрация”. В отличие от фундаментальных результатов в области оптимальной стохастической фильтрации субоптимальные фильтры имеют реальные перспективы практического применения, например обеспечивают возможность проведения расчетов для сложных многомерных моделей. И хотя полученные с их помощью оценки не обладают гарантированными свойствами, характеризующими их качество, имеются примеры, подтверждающие практическую применимость и полезность. Простой классический пример субоптимального алгоритма — расширенный фильтр Калмана — успешно применяется в ряде нелинейных задач, равно как и демонстрирует расходимость в других менее “удачных” примерах. Детальный обзор субоптимальных алгоритмов фильтрации не является задачей данной статьи, основное внимание будет уделено лишь одному успешному и развитому направлению, получившему значительную популярность — сигма-точечной фильтрации.

Концепция данной модификации фильтра Калмана (unscented Kalman Filter, сигма-точечный, или СТ-фильтр), впервые изложенная в публикации [1], существенных изменений в дальнейшем не претерпела и успешно эксплуатируется по настоящее время, превратившись в целое направление, находящее применение в самых разных областях. Основная идея СТ-фильтра состоит в аппроксимации непрерывных распределений параметров нелинейной системы наблюдения дискретными с использованием специально подбираемого “базиса” — набора сигма-точек. Данная аппроксимация позволяет приближенно рассчитывать моментные характеристики параметров, которые используются в рекуррентных уравнениях фильтра калмановской структуры. Хороший обзор приемов, методов и модификаций, касающихся сигма-точечных алгоритмов фильтрации, можно найти в [2], где сделаны значительные усилия по систематизации результатов в данной области. Отдельного упоминания заслуживает публикация [3], дополняющая исходную методику вариантом решения проблемы большой размерности (scaled unscented transformation), а также [4, 5], исследующие свойства сигма-точечного фильтра Калмана для систем специального вида, подходы к робастификации фильтра [6, 7], оптимизации параметров [8–10], повышению вычислительной эффективности [11] и обобщению на непрерывное время [12]. Популярность направления подтверждается большим числом публикаций практической направленности, посвященных применению СТ-фильтра и его адаптации к новым моделям [13–18].

Эвристический подход, характерный для направления субоптимальной фильтрации, имеет альтернативы, с одной стороны, обеспечивающие строгое математическое обоснование, а с другой — допускающие широкое применение на практике из-за простоты реализации. Такой альтернативой является теория условно-оптимальной фильтрации Пугачева и ее развитие в отношении динамических систем с дискретным временем — теория условно-минимаксной нелинейной фильтрации (УМНФ) [19–22]. В настоящее время

эти результаты имеют менее широкую область применения по сравнению с направлением СТ-фильтрации. Цель данной статьи — качественное сравнение концепций, лежащих в основе двух практически значимых подходов, и иллюстрация результатов этого анализа серий модельных экспериментов. Отдельной задачей является развитие условно-минимаксного подхода на основе идей, заложенных в методе СТ-фильтрации.

2. Основная модель

Рассмотрим нелинейную систему наблюдения, описываемую следующими разностными уравнениями:

$$(1) \quad \begin{aligned} x_t &= \varphi_t^{(1)}(x_{t-1}) + \varphi_t^{(2)}(x_{t-1})w_t, \quad t = 1, 2, \dots, \quad x_0 = \eta, \\ y_t &= \psi_t^{(1)}(x_t) + \psi_t^{(2)}(x_t)v_t. \end{aligned}$$

Здесь $x_t \in \mathbb{R}^p$ — случайный процесс, определяющий состояние системы; $y_t \in \mathbb{R}^q$ — косвенные наблюдения; w_t и v_t — дискретные белые шумы второго порядка; случайный вектор начальных условий η также имеет конечный второй момент; процессы w_t и v_t и вектор η независимы в совокупности. Их математические ожидания и ковариации далее будем обозначать $m_w(t)$, $D_w(t)$, $m_v(t)$, $D_v(t)$, m_η , D_η соответственно.

Рассматривается задача оценивания состояния x_t по наблюдениям y_τ , $\tau = 1, \dots, t$, критерий точности оценки \hat{x}_t — среднеквадратический:

$$E \left\{ \|x_t - \hat{x}_t\|^2 \right\}.$$

Модель (1) — компромиссное описание системы наблюдения, подходящее и для сигма-точечных, и для условно-минимаксных фильтров. Более того, практически важным часто оказывается дополнительное предположение $\psi_t^{(2)}(x) = 1$ и/или требование существования всех моментов у процессов w_t и v_t . Для целей данной статьи достаточно отметить, что при линейных ограничениях на скорость роста функций на бесконечности и наличии вторых моментов у всех возмущений процессы, описываемые (1), являются гильбертовыми. Соответственно, в таких условиях рассматриваемая задача фильтрации является корректной и можно, таким образом, говорить о ее решении как об оптимальном — условном математическом ожидании x_t относительно всех наблюдений, так и о каких-то вариантах аппроксимации решения.

3. Исследуемые концепции и методы

В основе обеих выбранных для исследования концепций синтеза алгоритмов нелинейной фильтрации лежат две простые статические задачи.

Сначала рассмотрим задачу, решение которой иллюстрирует суть метода УМНФ. Пусть есть случайный вектор $z = \text{col}(x, y)$, $x \in \mathbb{R}^p$, $y \in \mathbb{R}^q$ с известным математическим ожиданием $m_z = E\{z\} = \text{col}(m_x, m_y)$ и ковариации

ей $D_z = \text{cov}(z, z) = \begin{pmatrix} D_x & D_{xy} \\ D_{yx} & D_y \end{pmatrix}$. Закон распределения \mathcal{F}_z вектора z неизвестен, предполагается, что $\mathcal{F}_z \in \Phi(m_z, D_z)$ — классу всех вероятностных распределений со средним m_z и ковариацией D_z . Требуется найти оценку $\hat{x} = \theta(y)$ ненаблюдаемого вектора x по наблюдениям y , исходя из среднеквадратического критерия и заданного множества неопределенности Φ . Допустимыми считаются любые оценщики — измеримые по Борелю функции $\theta(y)$. Точность оценивания, обеспечиваемую оценкой $\theta(y)$ при условии, что вектор z имеет распределение \mathcal{F}_z , обозначим через $J(\theta, \mathcal{F}_z) = E\{\|x - \hat{x}\|^2\}$.

Решение поставленной задачи дает седловая точка, определяемая неравенствами $J(\theta^*, \mathcal{F}_z) \leq J(\theta^*, \mathcal{F}_z^*) \leq J(\theta, \mathcal{F}_z^*)$, где \mathcal{F}_z^* — гауссовское распределение с параметрами $m_z, D_z, x^* = \theta^*(y) = D_{xy}D_y^+y + (m_x - D_{xy}D_y^+m_y)$ — наилучшая в среднем квадратическом линейная оценка x по наблюдению y [21, 22]. При этом $J(\theta^*, \mathcal{F}_z^*) = D_x - D_{xy}D_y^+D_{yx}$. Выше и далее обозначены: $\text{cov}(x, y)$ — ковариационная матрица случайных векторов x и y , $E\{x\}$ — математическое ожидание случайного вектора x , $+$ — операция псевдообращения матрицы.

Из данного результата следует, что при формировании оценки в указанной задаче оценивания x по наблюдению y минимаксно обоснованным является линейный оценщик, параметры которого определяются моментными характеристиками совместного распределения. Этот результат лежит в основе метода УМНФ, а именно: оценка фильтрации \hat{x}_t состояния x_t получается в результате решения следующих минимаксных задач для определения прогноза \tilde{x}_t и его коррекции:

$$(2) \quad \tilde{x}_t = \tilde{\theta}_t(\xi_t), \quad \tilde{\theta}_t = \underset{\theta_t}{\text{argmin}} \max_{\mathcal{F}_z} E\{\|x_t - \theta_t(\xi_t)\|^2\}, \quad z = \text{col}(x_t, \xi_t),$$

$$\hat{x}_t = \tilde{x}_t + \hat{\theta}_t(\zeta_t), \quad \hat{\theta}_t = \underset{\theta_t}{\text{argmin}} \max_{\mathcal{F}_z} E\{\|x_t - \tilde{x}_t - \theta_t(\zeta_t)\|^2\}, \quad z = \text{col}(x_t - \tilde{x}_t, \zeta_t).$$

Здесь и далее $\xi_t = \xi_t(x)$ и $\zeta_t = \zeta_t(x, y)$ — некоторые структурные функции фильтра, варианты выбора которых составляют практический багаж метода.

Далее собственно описание алгоритма УМНФ таково. Пусть имеется \hat{x}_{t-1} — оценка УМНФ состояния x_{t-1} по наблюдениям $y_\tau, \tau = 1, \dots, t-1$. Прогноз \tilde{x}_t ищется в виде

$$(3) \quad \tilde{x}_t = F_t \xi_t + f_t, \quad \xi_t = \xi_t(\hat{x}_{t-1}), \quad \xi_t(x) = \varphi_t^{(1)}(x) + \varphi_t^{(2)}(x) m_w(t).$$

Структура прогноза “в силу системы” определяется структурной функцией $\xi_t(x)$ и коэффициентами F_t и f_t (матрица и вектор соответствующей размерности), определяемыми в результате решения задачи (2), а именно $\tilde{\theta}_t(\xi_t) = F_t \xi_t + f_t$.

Оценка \hat{x}_t состояния x_t ищется в виде

$$(4) \quad \begin{aligned} \hat{x}_t &= \tilde{x}_t + H_t \zeta_t + h_t, \quad \zeta_t = \zeta_t(\tilde{x}_t, y_t), \\ \zeta_t(x, y) &= y - \psi_t^{(1)}(x) - \psi_t^{(2)}(x) m_v(t). \end{aligned}$$

Структура коррекции в форме невязки определяется структурной функцией $\zeta_t(x, y)$ (невязкой) и коэффициентами H_t и h_t (матрица и вектор соответствующей размерности), определяемыми в результате решения задачи (2), а именно $\hat{\theta}_t(\zeta_t) = H_t\zeta_t + h_t$.

Решения обеих задач в (2) при условии конечности необходимых вторых моментов, очевидно, существуют, но могут быть не единственными. Единственность обеспечивается дополнительным требованием о минимуме евклидовой нормы решения и, следовательно, использованием операции псевдообращения по Муру–Пенроузу. Таким образом, искомые коэффициенты задаются в виде:

$$(5) \quad \begin{aligned} F_t &= \text{cov}(x_t, \xi_t) \text{cov}^+(\xi_t, \xi_t), \quad f_t = E\{x_t\} - F_t E\{\xi_t\}, \\ H_t &= \text{cov}(x_t - \tilde{x}_t, \zeta_t) \text{cov}^+(\zeta_t, \zeta_t), \quad h_t = -H_t E\{\zeta_t\}. \end{aligned}$$

Прогноз \tilde{x}_t и оценка фильтрации \hat{x}_t при этом являются несмещенными и обеспечивают следующее качество оценивания:

$$(6) \quad \begin{aligned} \tilde{K}_t &= \text{cov}(x_t - \tilde{x}_t, x_t - \tilde{x}_t) = \text{cov}(x_t, x_t) - F_t \text{cov}(\xi_t, x_t), \\ \hat{K}_t &= \text{cov}(x_t - \hat{x}_t, x_t - \hat{x}_t) = \tilde{K}_t - H_t \text{cov}(\zeta_t, x_t - \tilde{x}_t), \end{aligned}$$

т.е. обеспечивается нерасходимость и “содержательность” оценок, подчеркивая словом “содержательность” гарантированное преимущество и прогноза, и оценки по сравнению с тривиальной оценкой, поскольку $F_t \text{cov}(\xi_t, x_t) \geq 0$, $H_t \text{cov}(\zeta_t, x_t - \tilde{x}_t) \geq 0$.

Соотношения (3)–(5) определяют условно-оптимальный фильтр Пугачева, а концепция УМНФ дополняет их минимаксным обоснованием структуры фильтра. Кроме того, неизменным атрибутом УМНФ является способ практического определения коэффициентов F_t , f_t , H_t и h_t методом Монте-Карло, т.е. компьютерное имитационное моделирование. Сам фильтр получается заменой в (5) математических ожиданий и ковариаций их статистическими оценками, полученными в результате компьютерного моделирования. По остальным вопросам, включая условия существования фильтров, можно обратиться к публикации [23], в которой, кроме того, дан более подробный обзор метода и направлений его развития. Отметим возможность гибкого определения структурных функций фильтра — прогнозирующей $\xi_t(x)$ и корректирующей $\zeta_t(x, y)$: представленный вариант прогноза “в силу системы” и коррекции в форме невязки, конечно, является простейшим, и в зависимости от свойств конкретной системы наблюдения $\xi_t(x)$ и $\zeta_t(x, y)$ могут существенно изменяться.

Отметим еще, что УМНФ, в принципе, можно отнести к так называемым *нелинейным фильтрам Калмана* или, как упоминалось в разделе 1, фильтрам калмановской структуры. Поясним это, записав самый простой из таких фильтров — расширенный фильтр Калмана (РФК). Здесь и далее не вводятся дополнительные обозначения, а используются уже имеющиеся, например \tilde{x}_t и \hat{x}_t для обозначения прогноза и оценки фильтрации, ξ_t для прогноза “в силу системы” и ζ_t для невязки. При этом из контекста понятно, какие именно

алгоритмы имеются в виду. Для РФК соотношения таковы:

$$\begin{aligned}
 \tilde{x}_t &= \xi_t, \quad \xi_t = \xi_t(\hat{x}_{t-1}), \quad \xi_t(x) = \varphi_t^{(1)}(x) + \varphi_t^{(2)}(x)m_w(t), \\
 \tilde{K}_t &= f_t \hat{K}_{t-1} f_t^T + \varphi_t^{(2)} D_w(t) \varphi_t^{(2)T}, \quad f_t = \left. \frac{\partial \varphi_t^{(1)}(x)}{\partial x} \right|_{x = \tilde{x}_t}, \\
 \varphi_t^{(2)} &= \varphi_t^{(2)}(\tilde{x}_t), \quad \hat{x}_t = \tilde{x}_t + H_t \zeta_t, \quad \zeta_t = y - \psi_t^{(1)}(x) - \psi_t^{(2)}(x)m_v(t), \\
 (7) \quad H_t &= \tilde{K}_t h_t^T \left(h_t \tilde{K}_t h_t^T + \psi_t^{(2)} D_v(t) \psi_t^{(2)T} \right)^+, \\
 h_t &= \left. \frac{\partial \psi_t^{(1)}(x)}{\partial x} \right|_{x = \tilde{x}_t}, \quad \psi_t^{(2)} = \psi_t^{(2)}(\tilde{x}_t), \\
 \hat{K}_t &= \tilde{K}_t - H_t h_t \tilde{K}_t.
 \end{aligned}$$

Другие повторно использованные обозначения также имеют смысл. Так, f_t, h_t — некоторые оценки моментов, точнее, их аппроксимации, полученные путем линеаризации, H_t имеет тот же смысл, что и в (5), \tilde{K}_t и \hat{K}_t — ковариации ошибки прогноза и ошибки оценки, только здесь в отличие от (6) это аппроксимации в силу элементарной линеаризации нелинейностей $\varphi_t^{(1)}(x)$ и $\psi_t^{(1)}(x)$, а в УМНФ — точные значения или их статистические оценки.

Теперь обратимся к СТ-фильтрации. Калмановская структура фильтра в этом методе сохраняется, а основная идея принципиально другая. Именно: концептуальную основу СТ-фильтров составляет сигма-точечное преобразование (unscented transformation), состоящее в следующем. Пусть, как и выше, есть вектор $z = \text{col}(x, y) \in \mathbb{R}^{p+q}$, такой что $y = \varphi(x)$. Относительно распределения \mathcal{F}_x вектора x предполагается, что $\mathcal{F}_x \in \Phi(m_x, D_x)$, т.е. известны два первых момента. Неформально относительно \mathcal{F}_x можно сказать, что это распределение предполагается непрерывным с формой, “близкой” к гауссовской. Более того, во многих публикациях по данной тематике, следуя исходным предложениям [1], обсуждение СТ-фильтрации иллюстрируется примером расчета именно с гауссовским вектором x . Отметим, что рассуждения здесь, как и в публикациях, непосредственно посвященных данной тематике, не претендуют на формальность и не предполагают математической строгости условий и результатов. Как и в любых исследованиях по субоптимальному оцениванию, важна именно концептуальная составляющая — идея, обоснованная эмпирическими соображениями и подкрепленная модельными примерами. Искомое преобразование состоит в представлении \mathcal{F}_x его дискретным аналогом. Именно: вместо предположения о том, что x может принимать любые значения, будем считать, что область значений x исчерпывается конечным числом сигма-точек x^0, x^1, \dots, x^{2p} , с каждой из которых связана соответствующая вероятность W^0, W^1, \dots, W^{2p} . Обратим внимание на связь размерности p вектора x и числа сигма-точек $2p + 1$. Это принципиальный момент, подчеркивающий существенную ограниченность числа сигма-точек, которых должно быть немного. В противном случае, выбирая неограниченно

много x^i и подбирая значения W^i , можно, очевидно, добиться любой точности аппроксимации. Введенные обозначения дают простые приближенные соотношения для моментов вектора y :

$$(8) \quad m_y \approx \hat{y} = \sum_{i=0}^{2p} W^i y^i, \quad D_y \approx \hat{D}_y = \sum_{i=0}^{2p} W^i (y^i - \hat{y})(y^i - \hat{y})^T,$$

$$D_{xy} \approx \hat{D}_{xy} = \sum_{i=0}^{2p} W^i (x^i - m_x)(y^i - \hat{y})^T, \quad y^i = \varphi(x^i).$$

Близость моментов m_y и \hat{y} , D_y и \hat{D}_y , D_{xy} и \hat{D}_{xy} в (8) и есть цель СТ-преобразования. И эта цель становится яснее с позиции следующего шага — оценивания вектора x по наблюдениям y . Действительно, наилучшая линейная оценка x по y имеет вид $\theta^*(y) = D_{xy}D_y^+y + (m_x - D_{xy}D_y^+m_y)$, СТ-преобразование даст оценку $\hat{x} = \hat{D}_{xy}\hat{D}_y^+y + (m_x - \hat{D}_{xy}\hat{D}_y^+\hat{y})$. При этом интуитивно понятно, что чем точнее величинами \hat{D}_{xy} , \hat{D}_y , \hat{y} аппроксимируются моменты D_{xy} , D_y , m_y , тем точнее будет и оценка \hat{x} . Другое дело, что метод сигма-точек не обременен какими-либо оптимизационными постановками в отношении наилучшего выбора самих сигма-точек и весов, хотя и содержит изрядное число разумных, физически обоснованных и практически проиллюстрированных вариантов их выбора (весьма полезен в этой части обзор [2]).

Заметим, что предложенная интерпретация СТ-преобразования нужна только для иллюстрации. Обычно в публикациях по СТ-фильтрации не сообщается о дискретном распределении, а речь идет только о сигма-точках и “весах”, которые могут быть даже отрицательными. Выбор этих весов может быть весьма разнообразным и неуниверсальным даже в рамках одной модели и различается в зависимости от цели преобразования. Один из вариантов, который использовался в представленных далее расчетах, таков. Зададим параметры преобразования α, β, γ (рекомендации по их выбору и физическая интерпретация даны, например, в [24]). Определим “параметр масштаба” $\lambda = \alpha^2(p + \gamma) - p$. Сигма-точки сформируются следующим образом:

$$(9) \quad x^0 = m_x, \quad x^i = m_x + \left(\sqrt{(p + \lambda) D_x}\right)_i, \quad i = 1, \dots, p,$$

$$x^i = m_x - \left(\sqrt{(p + \lambda) D_x}\right)_i, \quad i = p + 1, \dots, 2p,$$

где $\left(\sqrt{(p + \lambda) D_x}\right)_i$ — i -й столбец матрицы $\sqrt{(p + \lambda) D_x}$, под корнем от матричнозначного аргумента понимается результат разложения Холецкого. Веса определяются так:

$$(10) \quad W^0 = \frac{\lambda}{p + \lambda} \quad \text{для вычисления } \hat{y},$$

$$W^0 = \frac{\lambda}{p + \lambda} + 1 - \alpha^2 + \beta \quad \text{для вычисления } \hat{D}_y \text{ и } \hat{D}_{xy},$$

$$W^i = \frac{1}{2(p + \lambda)}, \quad i = 1, \dots, 2p.$$

В примерах далее упомянутые рекомендации из [24] были использованы для выбора базовых значений параметров α^* , β^* , γ^* . Кроме того, в каждой задаче дополнительно выполнялась оптимизация, в которой базовые значения выступали в роли начальных условий, расчеты многократно повторялись для разных значений параметров и выбирались те, что обеспечивали СТ-фильтру наибольшую точность. Реализовать такую операцию в рамках данной статьи удалось благодаря тому, что рассматривались только простые примеры модельного характера, скалярные состояния и наблюдения, и не ограничивалось время, затрачиваемое на эксперименты. На практике, конечно, такая оптимизация параметров не всегда окажется возможной.

Сделаем два важных замечания. Во-первых, представив оценку СТ-преобразования в виде $\hat{x} = m_x + \hat{D}_{xy}\hat{D}_y^+(y - \hat{y})$, обратим внимание на сходство, функциональную идентичность разности $y - \hat{y}$ и невязки ζ_t , а также коэффициентов $\hat{D}_{xy}\hat{D}_y^+$ и H_t в соотношениях (7) РФК. Это сходство и есть эмпирическое обоснование использования СТ-преобразования в соотношениях нелинейного фильтра Калмана.

Второе замечание касается интерпретации моментов D_{xy} , D_y , m_y и соответствующих аппроксимаций \hat{D}_{xy} , \hat{D}_y , \hat{y} . Данное описание СТ-преобразования основано на статической модели оценивания, в которой эти параметры — просто моменты первого и второго порядков. Если же иметь в виду, что роль вектора x будет играть состояние x_t динамической системы (1), а предположение о его распределении будет изменяться в зависимости от наблюдений y_t , выполняющих роль y , то тогда \mathcal{F}_x уместно характеризовать как условное распределение x , форма которого неизвестна, но среднее значение и ковариация — результаты оценивания, выполненного по ранее поступившим наблюдениям. В этом смысле D_{xy} , D_y , m_y следует интерпретировать как условные моменты.

Теперь изложение можно подытожить соотношениями СТ-фильтра для системы (1). Эти соотношения приведены в соответствии с [24].

Пусть имеется \hat{x}_{t-1} — оценка СТ-фильтра состояния x_{t-1} по наблюдениям y_τ , $\tau = 1, \dots, t-1$, и \hat{K}_{t-1} — оценка условной ковариационной матрицы ошибки $x_{t-1} - \hat{x}_{t-1}$. Выберем сигма-точки \hat{x}_{t-1}^i в соответствии с (9) и веса W^i согласно (10), используя $D_x = \hat{K}_{t-1}$, и пересчитаем сигма-точки $\tilde{x}_t^i = \varphi_t^1(\hat{x}_{t-1}^i) + \varphi_t^2(\hat{x}_{t-1}^i)m_w(t)$, $i = 0, \dots, 2p$. Прогноз \tilde{x}_t ищется в виде

$$(11) \quad \tilde{x}_t = \sum_{i=0}^{2p} W^i \tilde{x}_t^i,$$

$$\tilde{K}_t = \sum_{i=0}^{2p} W^i (\tilde{x}_t^i - \tilde{x}_t) (\tilde{x}_t^i - \tilde{x}_t)^T + \varphi_t^2(\hat{x}_{t-1}^i) D_w(t) (\varphi_t^2(\hat{x}_{t-1}^i))^T,$$

где \tilde{K}_t — СТ-оценка условной ковариационной матрицы ошибки прогноза $x_t - \tilde{x}_t$. Далее пересчитываются следующие сигма-точки

$$y_t^i = \psi_t^1(\tilde{x}_t^i) + \psi_t^2(\tilde{x}_t^i)m_v(t).$$

Оценка \hat{x}_t состояния x_t ищется в виде

$$\begin{aligned} \hat{y}_t &= \sum_{i=0}^{2p} W^i y_t^i, & \hat{K}_t^y &= \sum_{i=0}^{2p} W^i (y_t^i - \hat{y}_t) (y_t^i - \hat{y}_t)^T + \psi_t^2(\tilde{x}_t^i) D_v(t) (\psi_t^2(\tilde{x}_t^i))^T, \\ (12) \quad \hat{K}_t^{xy} &= \sum_{i=0}^{2p} W^i (\tilde{x}_t^i - \tilde{x}_t) (y_t^i - \hat{y}_t)^T, \\ H_t &= \hat{K}_t^{xy} (\hat{K}_t^y)^+, & \hat{x}_t &= \tilde{x}_t + H_t \zeta_t, & \zeta_t &= y_t - \hat{y}_t, & \hat{K}_t &= \tilde{K}_t - H_t \hat{K}_t^y H_t^T. \end{aligned}$$

Здесь отметим величины \hat{K}_t^y и \hat{K}_t^{xy} , которые являются СТ-оценками условных ковариации невязки ζ_t и смешанной ковариации прогноза \tilde{x}_t и невязки ζ_t соответственно. Следует уточнить, что положения об условных моментах выше опираются на специфическое условие — предположение о том, что состояние x_{t-1} имеет распределение $\mathcal{F}_{x_{t-1}} \in \Phi(\hat{x}_{t-1}, \hat{K}_{t-1})$, т.е. вычисленные на предыдущем шаге фильтрации величины \hat{x}_{t-1} и \hat{K}_{t-1} интерпретируются как условное математическое ожидание и условная ковариация, по крайней мере являются их аппроксимациями. Именно из-за этого соображения ранее было сделано замечание в отношении интерпретации аппроксимаций \hat{D}_{xy} , \hat{D}_y , \hat{y} , получаемых в результате СТ-преобразования. На взгляд авторов, это крайне важное замечание в отношении методологии СТ-фильтрации, объясняющее практическую результативность алгоритмов фильтрации и высокую точность оценок в конкретных экспериментах. При этом термин “условный” здесь никак не связан с “действительно” условным распределением x_t относительно наблюдений y_τ , $\tau = 1, \dots, t$, или какого-то иного набора случайных величин. Но интуитивно содержание этого предложения понятно, а по форме оно иллюстрируется строгими соотношениями линейного фильтра Калмана и теоремой о нормальной корреляции [25].

Наконец, отметим, что использованные в (11) и (12) обозначения специально выбирались аналогичными обозначениям, использованным в соотношениях (3)–(6) УМНФ и (7) РФК, для того чтобы подчеркнуть сходство и даже идентичность в их интерпретации.

4. Модельные расчеты

4.1. Статическая регрессия и полярные координаты

Рассмотрим серию модельных примеров, позволяющих проиллюстрировать наличие как “хороших”, так и “неудачных” условий применения СТ-фильтров и проанализировать в таких же условиях поведение УМНФ. Отметим, что, в целом, концепция СТ-оценивания представляется вполне жизнеспособной и полезной и в ней имеются видимые достоинства и даже перспективы, которые пока не реализованы. Но вместе с этим есть и ограничения, систематическому изучению которых должно уделяться со стороны специалистов в этой области не уделялось. Одна из целей данной статьи — выявить ограничения СТ-фильтров хотя бы на примерах академического ха-

рактера и убедиться, что в таких моделях возможность практической фильтрации имеется (для этого привлекается УМНФ).

Рассмотрим простой статической пример, модель которого во многих публикациях используется для иллюстрации результатов СТ-преобразования (полностью такие рассуждения хорошо проиллюстрированы в [24]). Пусть имеется гауссовский вектор $\text{col}(x, y)$, величины x, y интерпретируются как декартовы координаты точки на плоскости. Будем считать эти координаты независимыми, неизвестными и наблюдаемыми косвенно. Через ρ и ϕ обозначим наблюдения соответствующих полярных координат x и y , зашумленных гауссовским аддитивным шумом:

$$(13) \quad \rho = \sqrt{x^2 + y^2} + v_\rho, \quad \phi = \arctg(x/y) + v_\phi,$$

где v_ρ, v_ϕ — независимые центрированные величины со среднеквадратическими отклонениями $\sigma_{v_\rho} = 30$ м, $\sigma_{v_\phi} = 5$ град. Требуется по ρ и ϕ оценить x и y . К оцениванию привлекаются следующие алгоритмы.

1. *Наилучшая линейная оценка*

$$\theta^*(\rho, \phi) = D_{x,y,\rho,\phi} D_{\rho,\phi}^+ \text{col}(\rho, \phi) + (\text{col}(m_x, m_y) - D_{x,y,\rho,\phi} D_{\rho,\phi}^+ \text{col}(m_\rho, m_\phi)),$$

$$D_{x,y,\rho,\phi} = \begin{pmatrix} \text{cov}(x, \rho) & \text{cov}(x, \phi) \\ \text{cov}(y, \rho) & \text{cov}(y, \phi) \end{pmatrix}, \quad D_{\rho,\phi} = \begin{pmatrix} \text{cov}(\rho, \rho) & \text{cov}(\rho, \phi) \\ \text{cov}(\phi, \rho) & \text{cov}(\phi, \phi) \end{pmatrix}.$$

Все фигурирующие в этих соотношениях моментные характеристики вычисляются методом Монте-Карло по выборке объема 10000.

2. *Оценка СТ-фильтра.* Вычисляется в точности по тем же формулам с той лишь разницей, что моменты $m_\rho, m_\phi, D_{x,y,\rho,\phi}, D_{\rho,\phi}$ вычисляются по формулам СТ-фильтра (12) (прогноз (11) здесь не требуется в силу статичности задачи), нужные для этого сигма-точки — по формулам (9), веса — по (10), при этом параметры α, β, γ подвергаются дополнительной оптимизации, а именно: выбираются наилучшие значения этих параметров с точки зрения качества оценивания в рассматриваемой задаче.

Замечание. Здесь и далее, говоря об оптимизации параметров α, β и γ СТ-фильтра, имеется в виду следующая вычислительная процедура. В соответствии с рекомендациями [24] в качестве базовых выбираются следующие значения параметров: $\alpha^* = 0,5, \beta^* = 2,0, \gamma^* = 3 - p$. Вокруг базовых значений формируется равномерная сетка возможных значений α, β, γ , из которых выбирается лучшая в смысле минимума второго момента ошибки оценки СТ-фильтра на последнем шаге фильтрации (единственном в статической задаче пересчета полярных координат). В окрестности этой комбинации параметров затем производится оптимизация методом Нелдера–Мида. В итоге рассматривается порядка 10000–20000 комбинаций возможных значений α, β, γ . Следует отметить, что в некоторых случаях вариация параметров СТ-преобразования относительно базовых $\alpha^*, \beta^*, \gamma^*$ не приводит к видимым изменениям качества оценивания. В таких случаях, а именно: когда разница между минимальным и максимальным значениями второго момента ошибки оценки на последнем шаге фильтрации не превышает 0,1%, для оценивания с помощью СТ-фильтра используются базовые значения.

Таблица

Сравнение для модели (13)		“Близко”	“Далеко”
1. Линейная оценка	$D[x - \hat{x}]$	585	562
	$D[y - \hat{y}]$	456	514
2. Сигма-точечная оценка	$D[x - \hat{x}]$	614 (444)	562 (553)
	$D[y - \hat{y}]$	467 (395)	514 (508)
3. Оценка УМНФ	$D[x - \hat{x}]$	276	564
	$D[y - \hat{y}]$	323	516

3. *Оценка УМНФ.* Заметим, во-первых, что наилучшая линейная оценка в данной задаче также является и оценкой УМНФ, если считать, что структурная функция (в данном статическом примере структурная функция только одна) $\zeta = \text{col}(\rho, \phi)$. Но в качестве “действительно” условно-минимаксной оценки здесь будем использовать $\zeta = \text{col}(\zeta_x, \zeta_y)$, $\zeta_x = \rho \cos \phi$, $\zeta_y = \rho \sin \phi$, т.е. соотношения перехода от полярных координат к декартовым, “обратные” в смысле модели наблюдения (13). Это характерный пример подбора структуры оценки УМНФ с использованием физического смысла модели.

Неопределенными в модели (13) остаются только условия на расположение исходной точки, т.е. параметры $\text{col}(m_x, m_y)$ и $\begin{pmatrix} D_x & 0 \\ 0 & D_y \end{pmatrix}$. Вторые моменты принципиального разнообразия в эксперимент не внесут, поэтому всюду полагаем $D_x = D_y = 30^2$. А вот значения m_x и m_y будут отражать понимание физического свойства “близко–далеко”. В качестве “близко” полагалось $m_x = 30$ и $m_y = 40$; “далеко” — $m_x = 300$, $m_y = 400$. Полученные результаты представлены в таблице, где через $D[x - \hat{x}]$ и $D[y - \hat{y}]$ обозначены дисперсии ошибок оценок координат x и y соответствующими алгоритмами. Для СТ-фильтра дополнительно в скобках указывается оценка соответствующей дисперсии, полученная усреднением величин \hat{K}_t из (12). Аналогичные “теоретические” значения для линейной оценки и УМНФ не приводятся, так как они близки к приведенным в таблице “статистическим” и становятся тем ближе, чем больше экспериментов проводится в рамках метода Монте-Карло.

Для расчетов в модели “близко” были определены следующие параметры СТ-преобразования: $\alpha = 1,06$, $\beta = 1,35$, $\gamma = 0,91$. Эти параметры существенно улучшают точность оценивания по сравнению с базовыми, для которых $D[x - \hat{x}] = 711$, $D[y - \hat{y}] = 477$. Оптимизация в модели “далеко” не оказала существенного влияния на качество оценивания, поэтому для СТ-преобразования были использованы базовые значения α^* , β^* и γ^* .

Первое, чем обращают на себя внимание приведенные результаты — это идентичность свойств оценок в модели “далеко”, небольшой проигрыш УМНФ в рамках статистической погрешности. Достаточно близки при этом и оценки точности СТ-фильтра, что позволяет считать, что используемые в (12) оценки условных моментов также хорошо аппроксимируются с помощью сигма-точек. Второе — это более чем двукратный проигрыш СТ-фильтра в модели “близко”, сопровождаемый существенной погрешностью аппроксимации дисперсии ошибки оценки. При этом чуть меньше, но также слишком велика ошибка линейной оценки. УМНФ справляется с задачей существенно луч-

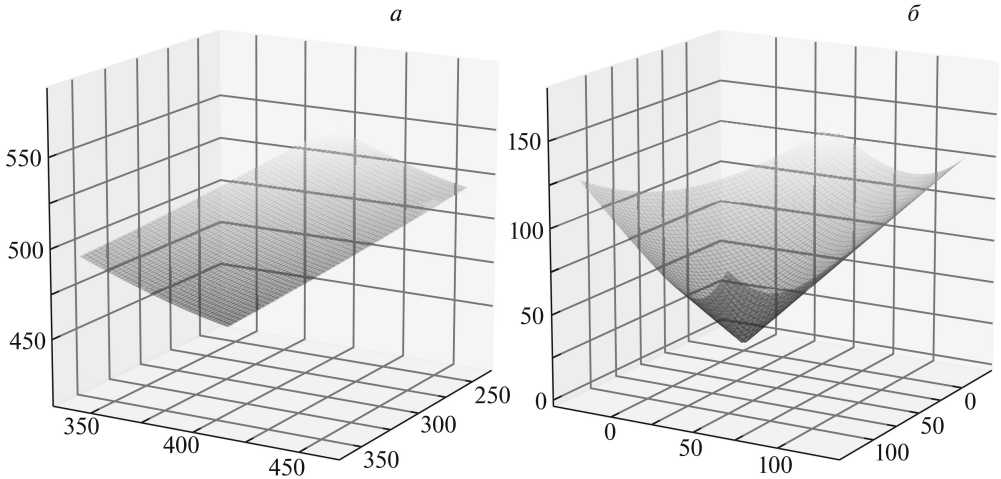


Рис. 1. Фрагменты поверхности $\rho = \sqrt{x^2 + y^2}$: *a* – модель “далеко”; *б* – модель “близко”.

ше и, что особо любопытно, значительно успешнее, чем в модели “далеко”. Причина таких результатов, на взгляд авторов, довольно проста. Модель “далеко” показывает поведение, “близкое к линейному”, модель “близко” — существенно нелинейна. Подтверждает это рис. 1, на котором изображены фрагменты поверхности $\rho = \sqrt{x^2 + y^2}$ для $x \in [m_x - 2\sqrt{D_x}, m_x + 2\sqrt{D_x}]$ и $y \in [m_y - 2\sqrt{D_y}, m_y + 2\sqrt{D_y}]$ и видно, что в модели “далеко” (рис. 1, *a*) этот фрагмент визуально неотличим от плоскости, а в модели “близко” (рис. 1, *б*) — это совсем не плоскость, и эта нелинейность, как демонстрирует УМНФ, весьма информативна.

Заметим, что такое поведение СТ-оценки не является неожиданным, на него обращают внимание многие исследователи, обозначая условие эффективности СТ-преобразования “консервативностью” ковариации ошибки оценки, которая в данном примере обеспечивается линейным поведением модели “далеко” и нарушается — нелинейным поведением модели “близко”.

4.2. Дробно-рациональная регрессия и кубический сенсор

Далее будем рассматривать примеры автономных скалярных динамических систем. В таких примерах легко анализировать четко выраженные свойства модели и их влияние, в конечном итоге, на точность оценок. Первая модель в этом ряду такова:

$$(14) \quad \begin{aligned} x_t &= \frac{x_{t-1}}{1 + x_{t-1}^2} + w_t, \quad t = 1, \dots, T, \quad T = 50, \quad x_0 = \eta, \\ y_t &= x_t + x_t^3 + v_t. \end{aligned}$$

Шумы w_t и v_t будем предполагать стандартными гауссовскими, начальное условие η тоже гауссовским с $m_\eta = 0,1$ и $D_\eta = 1$. Наблюдения в (14) хорошо известны под названием “кубический сенсор”, они очень удобны (информативны) для оценивания, так как при больших значениях x_t существенно

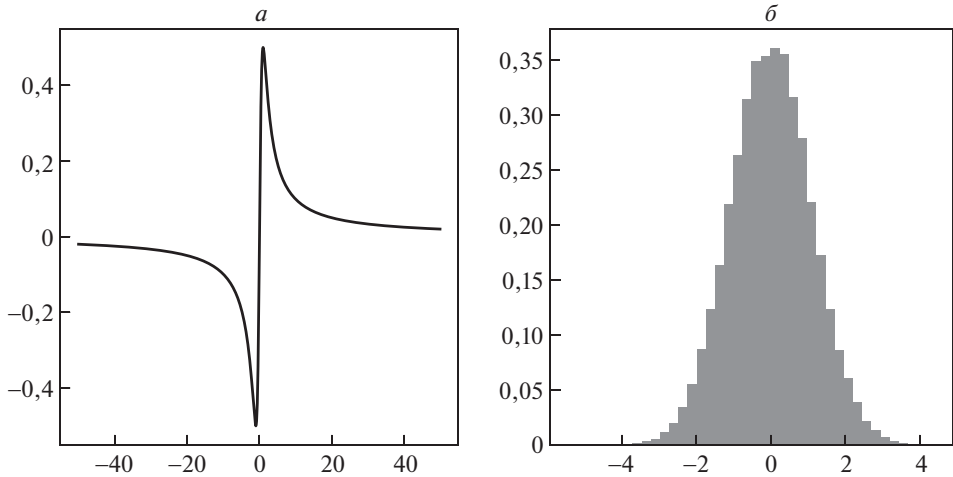


Рис. 2. Характеризация модели (14): a – функция регрессии; b – гистограмма для x_T , $E[x_T] = 0$, $D[x_T] = 1,16$.

увеличивают отношение сигнал/шум, а при малых — близки к линейным. Что касается динамики, то функция регрессии $x/(1+x^2)$ хотя и нелинейная (рис. 2, a), но демонстрирует довольно инертное поведение — обладает всего двумя экстремумами, гладкими производными и асимптотически приближается к нулю на бесконечности.

Отметим также, что фазовый процесс x_t обладает одним полезным свойством. Он, очевидно, является эргодическим (нетрудно показать это, следуя базовым результатам по нелинейной регрессии, изложенным, например, в [26]), и при этом сходимость к предельному распределению практически мгновенная: уже на втором шаге при начальных условиях, сильно отличающихся от предельных, получаются моменты, характерные для предельного распределения $E[x_T] = 0$, $D[x_T] = 1,16$. (на рис. 2, b предельное распределение иллюстрируется гистограммой x_T). Таким образом, у системы (14) фактически отсутствует переходной процесс, поэтому она представляется идеальной как для СТ-фильтрации, так и для УМНФ.

Здесь и всюду далее СТ-фильтр реализован в соответствии с соотношениями (11), (12), дополненными упомянутой в первом примере оптимизацией параметров α , β , γ . Структура УМНФ базовая: $\xi_t(x) = x/(1+x^2)$, т.е. прогноз строится “в силу системы”, $\zeta_t(x, y) = y - x - x^3$ — коррекция в форме невязки. Параметры УМНФ вычислялись методом Монте-Карло по пучку 10^5 траекторий (его уместно называть обучающим), этот же пучок использовался для оптимизации параметров СТ-фильтра. Качество оценок фильтрации и моментные характеристики x_t оценивались по второму пучку из 10^6 траекторий. Приведенные условия расчетов применялись и в следующих примерах.

На рис. 3 показаны: примеры характерных траекторий (состояния, наблюдения — рис. 3, a), ошибок оценок СТ-фильтра и УМНФ — рис. 3, b , показатели качества оценок — средние и дисперсии ошибок оценок и процесса пока-

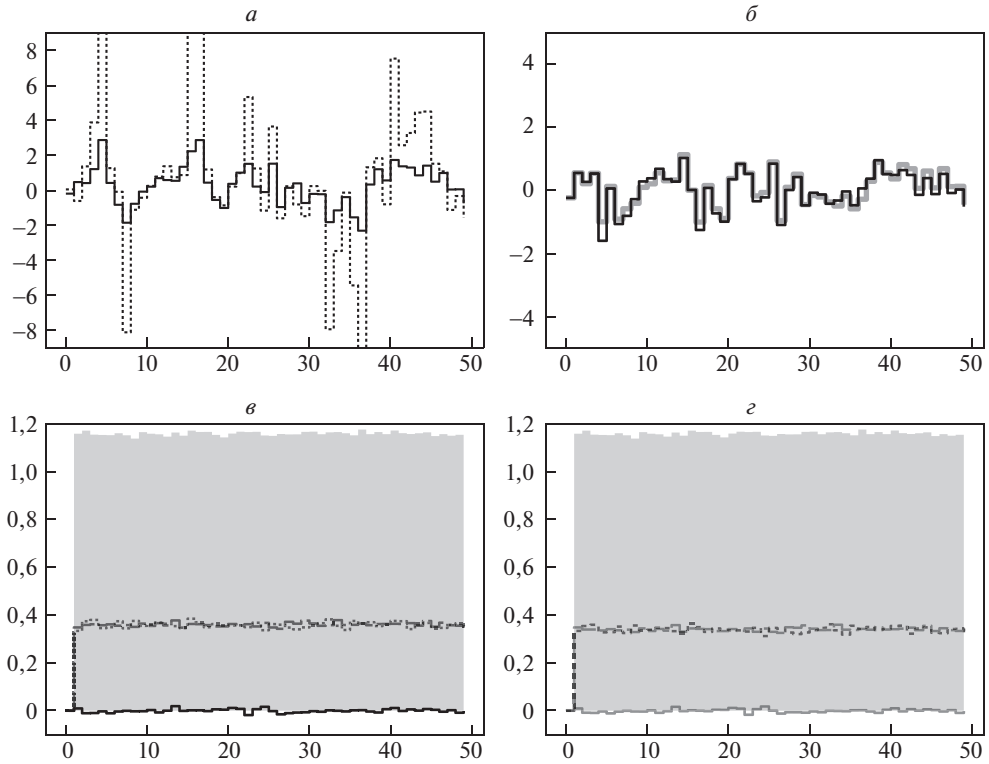


Рис. 3. Примеры траекторий и показатели качества оценок в модели (14): *а* – характерные траектории x_t (сплошная линия), y_t (точечная линия); *б* – траектории ошибок фильтрации $\hat{x}_t - x_t$ УМНФ (черная линия), СТ-фильтра (серая линия); *в, г* – показатели точности УМНФ и СТ-фильтра на фоне дисперсии процесса $D[x_t]$ (верхняя граница серой области): $E[x_t - \hat{x}_t]$ (сплошные линии), $D[x_t - \hat{x}_t]$ (пунктирные линии), \hat{K}_t (точечные линии).

заны для УМНФ на рис. 3, *в*, для СТ-фильтра – на рис. 3, *г*. Выборочные моменты предельных распределений ошибок оценивания равны $E[x_T - \hat{x}_T] = 0$, $D[x_T - \hat{x}_T] = 0,35$ для УМНФ и $E[x_T - \hat{x}_T] = 0$, $D[x_T - \hat{x}_T] = 0,34$ для СТ-фильтра. Дополнительно на рис. 3, *в* показана априорная оценка точности УМНФ, т.е. величина \hat{K}_t из (6), вычисленная методом Монте-Карло, $\hat{K}_T = 0,35$; на рис. 3, *г* показана “теоретическая” точность – усредненная величина \hat{K}_t из (12), $\hat{K}_T = 0,11$. Наилучшие найденные значения параметров СТ-преобразования в данной модели $\alpha = 0,20$, $\beta = 1,99$, $\gamma = 1,23$. Для базовых значений параметров $\alpha^* = 0,5$, $\beta^* = 2,0$, $\gamma^* = 2,0$ СТ-оценка оказалась хуже тривиальной, обеспечивая точность $D[x_T - \hat{x}_T] = 1,57$ при $D[x_T] = 1,16$.

Как видно, обе оценки демонстрируют близкие качества, обе несмещенные и содержательные в том смысле, что дисперсия ошибки оценки меньше, причем существенно меньше, дисперсии самого процесса и дисперсии шума в наблюдениях. Отметим также, что СТ-фильтр немного превосходит УМНФ, но последний имеет самую простую структуру (прогноз “в силу системы”, коррекция в форме невязки) и никаких вариантов ее улучшения не предпри-

малось, хотя они и легко просматриваются. Например, к улучшению качества УМНФ приведет увеличение размерности базовой коррекции $\zeta_t(x, y)$ за счет “дополнительных” наблюдений y_t^2, y_t^3 . Также отметим, что “теоретическая” точность СТ-фильтра далека от реальности, что не мешает ему справляться с задачей.

4.3. Обратная регрессия

Перейдем теперь к моделям, представляющим не столь комфортные условия для нелинейного оценивания. При этом состояния, как и ранее, описываются скалярными автономными регрессиями с гауссовскими возмущениями, обладающими эргодическим свойством. В данном подразделе рассмотрим сразу две модели, использующих функцию обратной пропорциональности и наблюдения — прямые линейные:

$$(15) \quad x_t = \frac{1}{\sqrt[3]{x_{t-1}}} + 100w_t, \quad t = 1, \dots, T, \quad T = 50, \quad x_0 = \eta,$$

$$y_t = x_t + 100v_t,$$

$$(16) \quad x_t = \min\left(10^5, \frac{1}{x_{t-1}^2}\right) + 100w_t, \quad t = 1, \dots, T, \quad T = 50, \quad x_0 = \eta,$$

$$y_t = x_t + 100v_t.$$

Шумы w_t и v_t будем предполагать стандартными гауссовскими, начальное условие η — тоже гауссовским с $m_\eta = 3$ и $D_\eta = 1$. Соотношения (15), (16) необходимо для корректности дополнить равенством $x_t = 10^5$, если $x_{t-1} = 0$. В обеих моделях регрессионная функция одного класса $1/x^\delta$, $\delta = \frac{1}{3}$ в (15), $\delta = 2$ в (16). Разница, легко видеть, в том, что степенью δ в (15) обеспечивается существование вторых моментов у x_t , а в (16) — напротив, нет даже первого момента. Именно поэтому, чтобы сделать задачу фильтрации корректной, в (16) добавлена верхняя граница 10^5 , которая обеспечивает существование всех моментов. Предельные распределения процессов x_t в (15) и (16) характеризуются типовыми унимодальными плотностями вероятности, предельные выборочные средние значения $E[x_T]$ равны 0,18 и 6,06, дисперсии $D[x_T]$ равны 10314,41 и 361201,04 соответственно.

Оба процесса x_t имеют аналогичное поведение: близкое к нулевому математическое ожидание и “много” реализаций около нуля, большая дисперсия, обеспечиваемая периодически появляющимися “большими” значениями, когда реализации $x_{t-1} \approx 0$. Заметим, что обе модели не попадают в сферу робастного оценивания, здесь нет “исчезнувших” или “ложных” наблюдений, выбросов и т.п. Робастификация в традиционном смысле не даст в этой модели результата, так как редкие всплески требуется оценивать, поскольку они представляют фактическое состояние оцениваемого процесса и не могут быть проигнорированы. Результаты проведенных расчетов иллюстрируются следующими рисунками. Для (15) приведены примеры характерных траекторий (состояния, наблюдения — на рис. 4,а), ошибок оценок СТ-фильтра и УМНФ — на рис. 4,б), показатели качества оценок — средние и дисперсии ошибок оценок и процесса показаны для УМНФ на рис. 4,в,

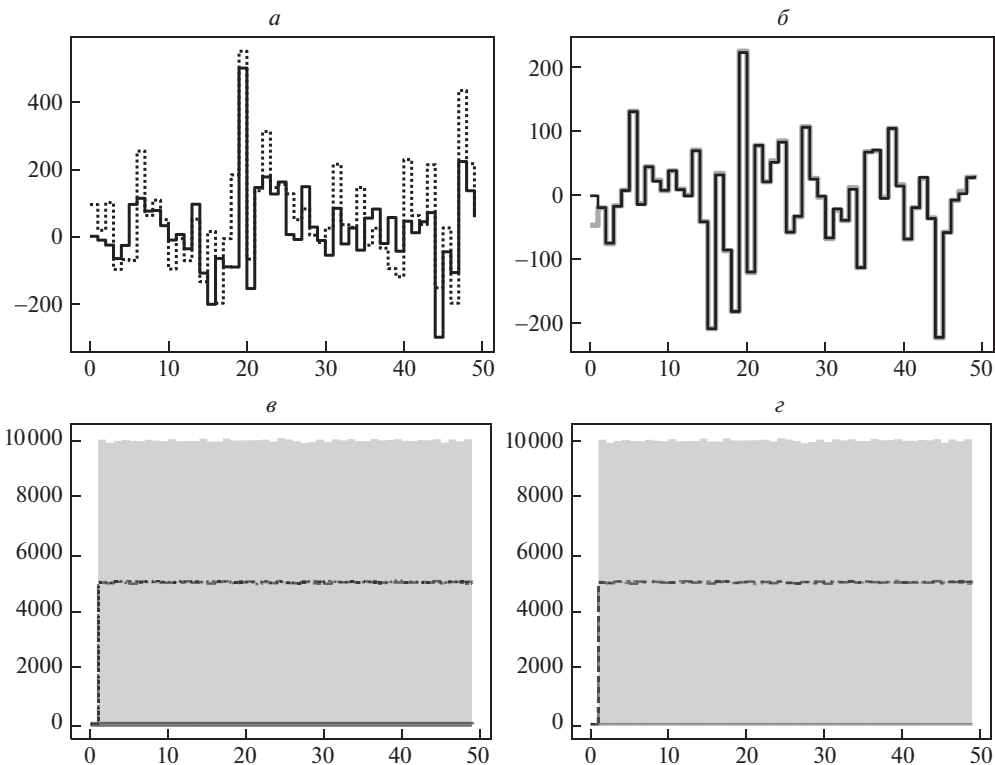


Рис. 4. Примеры траекторий и показатели качества оценок в модели (15): *а* – характерные траектории x_t (сплошная линия), y_t (точечная линия); *б* – траектории ошибок фильтрации $\hat{x}_t - x_t$ УМНФ (черная линия), СТ-фильтра (серая линия); *в, г* – показатели точности УМНФ и СТ-фильтра на фоне дисперсии процесса $D[x_t]$ (верхняя граница серой области): $E[x_t - \hat{x}_t]$ (сплошные линии), $D[x_t - \hat{x}_t]$ (пунктирные линии), \hat{K}_t (точечные линии).

для СТ-фильтра – на рис. 4, *г*. Выборочные моменты предельных распределений ошибок оценивания равны $E[x_T - \hat{x}_T] = -1,87$, $D[x_T - \hat{x}_T] = 5107,98$ для УМНФ и $E[x_T - \hat{x}_T] = -1,62$, $D[x_T - \hat{x}_T] = 5109,28$ для СТ-фильтра. Дополнительно на рис. 4, *в* показана величина \hat{K}_t из (6), $\hat{K}_T = 5000,11$; на рис. 4, *г* – усредненная величина \hat{K}_t из (12), $\hat{K}_T = 5000,04$. Для модели (16) аналогичные результаты приведены на рис. 5. Выборочные моменты предельных распределений ошибок оценивания здесь равны $E[x_T - \hat{x}_T] = -0,78$, $D[x_T - \hat{x}_T] = 9667,38$ для УМНФ и $E[x_T - \hat{x}_T] = 2,51$, $D[x_T - \hat{x}_T] = 92521,33$ для СТ-фильтра. Дополнительно на рис. 5, *в* показана величина \hat{K}_t из (6), $\hat{K}_T = 9804,96$; на рис. 5, *г* – усредненная величина \hat{K}_t из (12), $\hat{K}_T = 5007,19$. В обеих моделях (15), (16) вариация базовых значений параметров СТ-преобразования α^* , β^* , γ^* не приводит к улучшению качества оценки.

Полученные результаты соответствуют ожиданиям: в комфортных условиях модели (15) обе оценки ведут себя одинаково хорошо, в “неудобной” модели (16) работает только УМНФ, его точность хоть и ненамного, но выше точности прямых наблюдений. Причина этого не в плохой реализации СТ-филь-

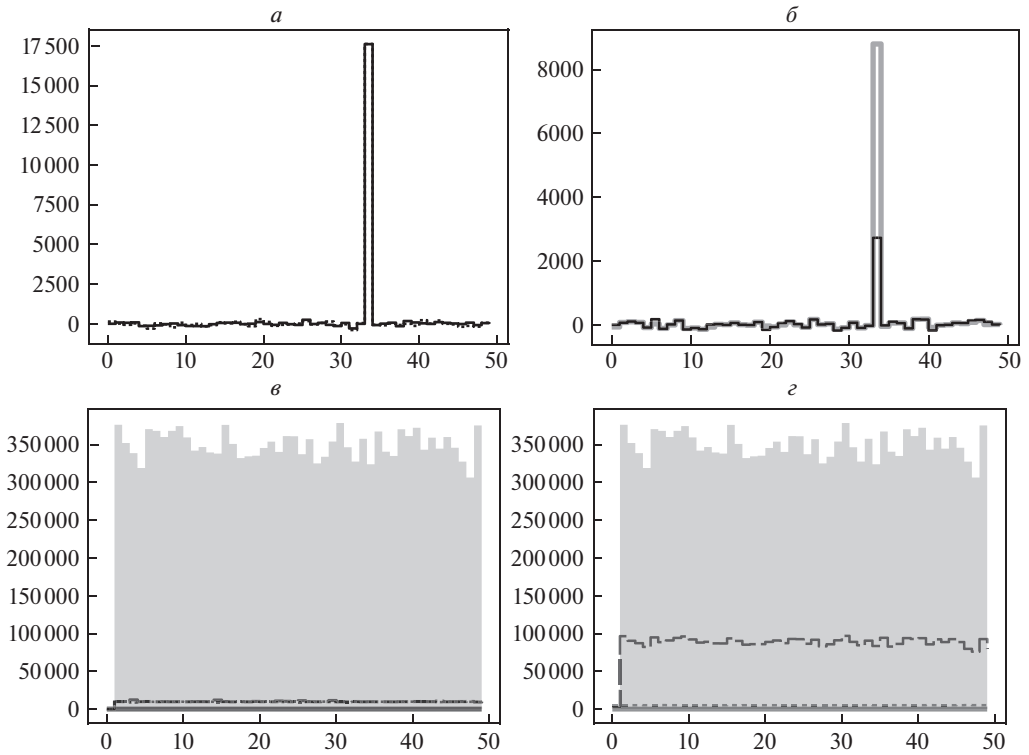


Рис. 5. Примеры траекторий и показатели качества оценок в модели (16): *a* – характерные траектории x_t (сплошная линия), y_t (точечная линия); *б* – траектории ошибок фильтрации $\hat{x}_t - x_t$ УМНФ (черная линия), СТ-фильтра (серая линия); *в, г* – показатели точности УМНФ и СТ-фильтра на фоне дисперсии процесса $D[x_t]$ (верхняя граница серой области): $E[x_t - \hat{x}_t]$ (сплошные линии), $D[x_t - \hat{x}_t]$ (пунктирные линии), \hat{K}_t (точечные линии).

тра, а в неприменимости в этой модели самой концепции сигма-точек, поскольку такую нелинейность, как в (16), охарактеризовать небольшим конечным набором образцов преобразований (сигма-точек), по-видимому, нельзя. Но самое неприятное, на взгляд авторов, для СТ-фильтра в этой модели следующее обстоятельство. В данном простом модельном примере заранее предсказать его плохое поведение можно, просто визуализируя характер нелинейности в уравнении состояния. На практике же такую особенность вряд ли удастся предсказать априорно, поскольку рассчитывать на успешный визуальный анализ уравнений, чуть более сложных, чем представленная скалярная автономная регрессия, вряд ли возможно.

4.4. Логистическая модель

Следующей рассмотрим модель, основанную на дискретном аналоге хорошо известной логистической модели (модели для численности популяции) — логистическом отображении или отображении Фейгенбаума $x_t = \delta x_{t-1}(1 - x_{t-1})$ [27]. Добавив в данную модель аддитивное гауссовское воз-

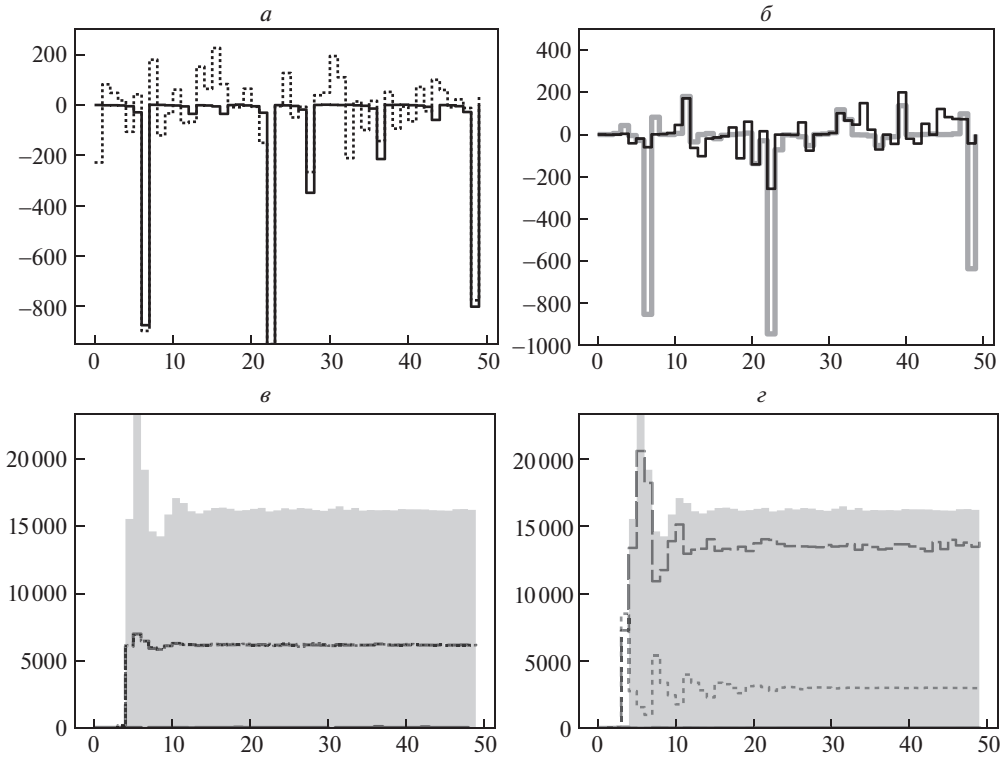


Рис. 6. Примеры траекторий и показатели качества оценок в модели (17): *а* – характерные траектории x_t (сплошная линия), y_t (точечная линия); *б* – траектории ошибок фильтрации $\hat{x}_t - x_t$ УМНФ (черная линия), СТ-фильтра (серая линия); *в, г* – показатели точности УМНФ и СТ-фильтра на фоне дисперсии процесса $D[x_t]$ (верхняя граница серой области): $E[x_t - \hat{x}_t]$ (сплошные линии), $D[x_t - \hat{x}_t]$ (пунктирные линии), \hat{K}_t (точечные линии).

мушение, получим динамику системы наблюдения. Модель при этом получается неустойчивой и, как и ранее, ее приходится искусственно ограничить. Причем при достижении границы процесс предлагается возвращать в исходное состояние, чтобы взрывные явления повторялись раз за разом. Таким образом, в этом разделе будем рассматривать следующую систему наблюдения:

$$\begin{aligned}
 x_t &= \begin{cases} x_{t-1}(1 - x_{t-1}), & \text{если } x_{t-1}(1 - x_{t-1}) > -10^3, \\ 0, & \text{иначе} \end{cases} + w_t, \quad t = 1, \dots, 50, \quad x_0 = \eta, \\
 (17) \quad y_t &= x_t + 100v_t.
 \end{aligned}$$

Здесь в основе динамики логистическое отображение с $\delta = 1$, поэтому если бы в (17) не фигурировала нижняя граница (-10^3), то имел бы место “взрыв” в отрицательную область: $x_t \rightarrow -\infty$. Но при этом, даже если убрать из (17) границу (-10^3), формально у состояния x_t есть все конечные моменты в любой момент времени, так что решение задачи фильтрации есть для любого t . Другое дело, что рост моментов и абсолютных значений x_t характеризуется

степенной скоростью, поэтому при компьютерных расчетах за короткое время любая траектория в модели без границы превысит компьютерный максимум.

Полученные результаты иллюстрирует рис. 6. Выборочные моменты предельных распределений ошибок оценивания равны

$$E[x_T - \hat{x}_T] = -0,83, \quad D[x_T - \hat{x}_T] = 6208,12 \quad \text{для УМНФ}$$

и

$$E[x_T - \hat{x}_T] = -17,07, \quad D[x_T - \hat{x}_T] = 13527,82 \quad \text{для СТ-фильтра,}$$

тогда как выборочная дисперсия процесса $D[x_T] = 16341,25$. Дополнительно на рис. 6,в показана величина \hat{K}_t из (6), $\hat{K}_T = 6300,53$; на рис. 6,г — усредненная величина \hat{K}_t из (12), $\hat{K}_T = 3068,76$. В данной модели оптимизация параметров СТ-преобразования также не дает результата, поэтому используются базовые α^* , β^* , γ^* .

Данный пример представляется яркой иллюстрацией потенциальных сложностей практического применения алгоритмов СТ-фильтрации. Он представляет собой простую полиномиальную регрессию, обеспечивающую формальное существование всех моментов, и информативные линейные наблюдения, но фазовый процесс при этом имеет “взрывной” характер, уследить за которым сигма-точечным методом не удастся. СТ-фильтр оказывается существенно смещенным, точность оценки близка к тривиальной и не выдерживает сравнения с точностью УМНФ. Ошибка оценки УМНФ также велика, но оценивание условно-минимаксным методом имеет смысл, поскольку дисперсия ошибки определено меньше дисперсии шума в наблюдениях.

4.5. Модель переключающихся каналов наблюдения

Эта модель основана на идее дополнительного внешнего влияния на характеристики наблюдателя, впервые привлечшей внимание около 50 лет назад (см., например, [28]) и в дальнейшем обобщенной в моделях систем случайной структуры и скрытых марковских моделях. Здесь рассматривается один из самых простых вариантов.

Пусть имеется не зависящий от состояния x_t белый шум δ_t , сечениями которого являются дискретные случайные величины, принимающие значения δ^i , $i = 1, \dots, n$, с известным распределением $P\{\delta_t = \delta^i\} = p^i$. Для удобства будем считать, что δ^i — n -мерные единичные векторы, т.е. $\delta^i = (0, \dots, 1, \dots, 0)^T$, где 1 стоит на i -м месте. Пусть d , $\sigma \in \mathbb{R}^n$. Выражение $d^T \delta_t$ дает i -ю координату вектора $d = (d^1, \dots, d^n)^T$. Будем считать, что состояние x_t в момент t может наблюдаться одним из n каналов линейных наблюдений $y_t = d^i x_t + \sigma^i v_t$, реализация которых определяется величиной δ_t . С помощью введенных обозначений такие наблюдения записываются в виде $y_t = d^T \delta_t x_t + \sigma^T \delta_t v_t$. Состояние x_t опишем простой линейной регрессией, получив, таким образом, следующую систему наблюдения:

$$(18) \quad \begin{aligned} x_t &= ax_{t-1} + b + cw_t, \quad t = 1, \dots, T, \quad T = 50, \quad x_0 = \eta, \\ y_t &= d^T \delta_t x_t + \sigma^T \delta_t v_t, \end{aligned}$$

где η , w_t , v_t — стандартные гауссовские.

Для проведения расчетов выберем $n = 3$, $p^1 = p^3 = 0,25$, $p^2 = 0,5$, $d^T = (4; 1; 0,5)$, $\sigma^T = (1; 1; 3)$, $a = 0,8$, $b = 0,2$, $c = 6,0$. Предельное распределение x_t , легко видеть, является гауссовским со средним единица и дисперсией 100.

В данном примере нельзя использовать ни соотношения (11), (12) СТ-фильтра, ни соотношения (1)–(4) УМНФ из-за формального несоответствия моделей (18) и (1). Устраняется это несоответствие следующим образом. В дополнение к шуму w_t введем еще один независимый гауссовский белый шум $\{w_t^\delta\}$ и непересекающиеся интервалы Δ^i : $P\{w_t^\delta \in \Delta^i\} = p^i$, $i = 1, 2, 3$. Далее дополним состояние x_t второй координатой

$$x_t^\delta = \left(I(w_t^\delta \in \Delta^1), I(w_t^\delta \in \Delta^2), I(w_t^\delta \in \Delta^3) \right) (1, 2, 3)^T,$$

т.е. номером реализации δ_t . В уравнении наблюдений вместо $d^T \delta_t$ будем использовать функцию $\psi_t^{(1)}(x_t, x_t^\delta) = d^i x_t$, если $x_t^\delta = i$, $i = 1, 2, 3$, и аналогично вместо $\sigma^T \delta_t$ — функцию $\psi_t^{(2)}(x_t, x_t^\delta) = \sigma^i x_t$, если $x_t^\delta = i$. Эти обозначения формально сводят модель (18) к виду (1), состояние при этом становится двумерным, наблюдения остаются скалярными. Таким образом, в реализации СТ-фильтра будет использоваться пять сигма-точек в соответствии с (9). Заметим, что дополнительных соображений в формировании структуры СТ-фильтра рассматриваемая модель не дает. Иная ситуация с УМНФ. Специфическая структура системы наблюдения требует учета в структуре фильтра, в его корректирующем члене $\zeta_t(x, y)$. Действительно, сформировав стандартный прогноз \tilde{x}_t “в силу системы” и определив его точность \tilde{K}_t согласно (3) и (6), шаг коррекции можно рассматривать как вспомогательную задачу оценивания x_t по одному наблюдению $y_t = d^T \delta_t x_t + \sigma^T \delta_t v_t$, решение которой при дополнительном предположении, что распределение x_t — гауссовское с параметрами \tilde{x}_t и \tilde{K}_t , легко записать в виде взвешенной суммы линейных оценок x_t по каждому возможному наблюдению $y_t = d^i x_t + \sigma^i v_t$, веса в которой образуют отношения правдоподобия. Этот подход дает следующую базовую коррекцию:

$$\zeta_t(x, y) = \frac{\sum_{i=1}^3 p^i N\left(y; d^i \tilde{x}_t, d^{i2} \tilde{K}_t + \sigma^{i2}\right) \left(d^i \tilde{K}_t \left(d^{i2} \tilde{K}_t + \sigma^{i2}\right)^{-1} (y - d^i x)\right)}{\sum_{i=1}^3 p^i N\left(y; d^i \tilde{x}_t, d^{i2} \tilde{K}_t + \sigma^{i2}\right)},$$

где через $N(y; m, \sigma^2)$ обозначена гауссовская плотность с параметрами m, σ^2 , вычисленная в точке y .

Комментируя это предложение, можно сказать, что выражения

$$\frac{p^i N\left(y_t; d^i \tilde{x}_t, d^{i2} \tilde{K}_t + \sigma^{i2}\right)}{\sum_{i=1}^3 p^i N\left(y_t; d^i \tilde{x}_t, d^{i2} \tilde{K}_t + \sigma^{i2}\right)}$$

представляют оценку апостериорной вероятности события $\{\delta_t = \delta^i\}$, величины $d^i \tilde{K}_t \left(d^{i2} \tilde{K}_t + \sigma^{i2}\right)^{-1} (y - d^i \tilde{x}_t)$, по сути, являются корректирующими сла-

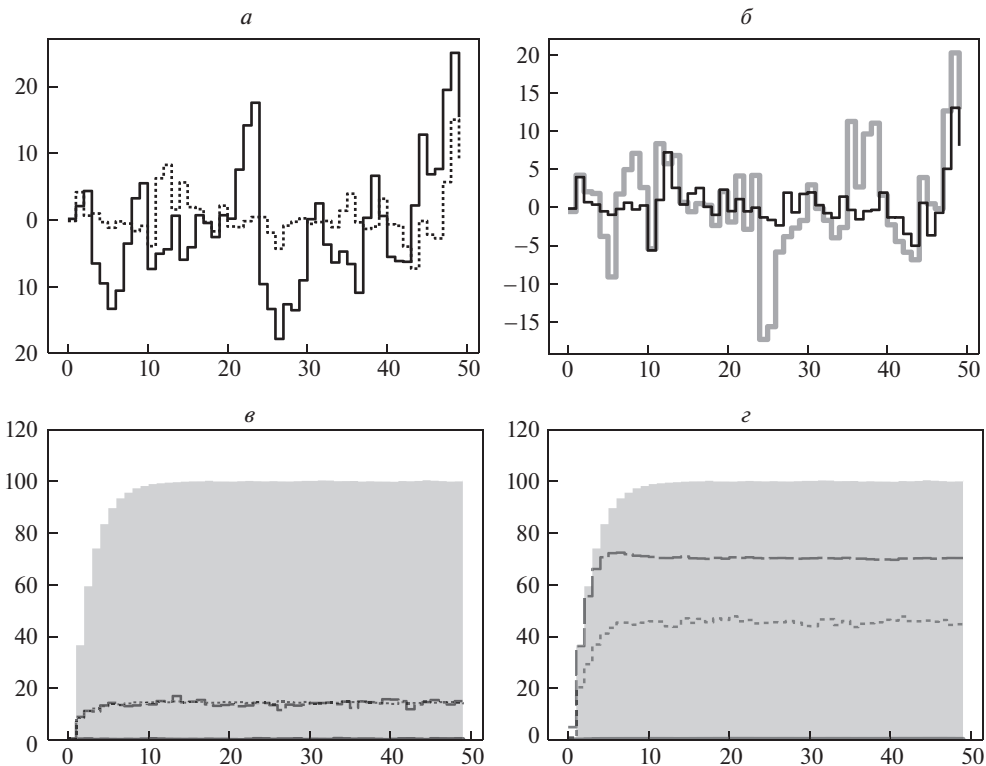


Рис. 7. Примеры траекторий и показатели качества оценок в модели (18): *а* – характерные траектории x_t (сплошная линия), y_t (точечная линия); *б* – траектории ошибок фильтрации $\hat{x}_t - x_t$ УМНФ (черная линия), СТ-фильтра (серая линия); *в, г* – показатели точности УМНФ и СТ-фильтра на фоне дисперсии процесса $D[x_t]$ (верхняя граница серой области): $E[x_t - \hat{x}_t]$ (сплошные линии), $D[x_t - \hat{x}_t]$ (пунктирные линии), \hat{K}_t (точечные линии).

гаемыми линейного фильтра Калмана при условии, что в момент t реализовался i -й канал наблюдения.

В данной модели была дополнительно предпринята попытка оптимизации значений параметров СТ-преобразования отдельно на шагах прогнозирования и коррекции, т.е. дополнительного пересчета точек y_t^i в (12) в соответствии с (9). Наилучшие найденные значения таковы: $\alpha = 2,46$, $\beta = 3,82$, $\gamma = 2,59$ для шага прогнозирования и $\alpha = 0,52$, $\beta = 4,15$, $\gamma = 1,24$ для шага коррекции. При этом добиться серьезного улучшения точности СТ-фильтра не удастся: для базовых значений параметров α^* , β^* , γ^* точность оценивания составляет $D[x_T - \hat{x}_T] = 71,81$.

Результаты расчетов иллюстрирует рис. 7. Выборочные моменты предельных распределений ошибок оценивания равны

$$E[x_T - \hat{x}_T] = 0,18, \quad D[x_T - \hat{x}_T] = 14,30 \quad \text{для УМНФ}$$

и

$$E[x_T - \hat{x}_T] = 0,29, \quad D[x_T - \hat{x}_T] = 70,18 \quad \text{для СТ-фильтра.}$$

Дополнительно на рис. 7,6 показана величина \hat{K}_t из (6), $\hat{K}_T = 14,41$; на рис. 7,2 — усредненная величина \hat{K}_t из (12), $\hat{K}_T = 45,27$.

Результаты демонстрируют ожидаемый значительный проигрыш СТ-фильтра по отношению к УМНФ. Ясно, что высокое качество оценки УМНФ обеспечивается выбором коррекции $\zeta_t(x, y)$, отражающим специфику модели. Иллюстрация гибкости метода УМНФ в отношении структуры фильтра — одна из целей данного примера и его отличие от всех предыдущих. СТ-фильтр, наоборот, поставлен в заведомо проигрышные условия формальным сведением модели (18) к виду (1). В данном примере нетрудно увидеть и реализовать гораздо более качественный СТ-фильтр, выбрав сигма-точки с учетом дискретного характера распределения δ_t , а именно для каждой из трех сигма-точек, выбранных отдельно для первой координаты x_t , определить три двумерных сигма-вектора (x_t^j, δ^j) , $j = 1, 2, 3$, сформировав таким образом окончательный набор из девяти сигма-точек для состояния (x_t, δ_t) . Такой вариант не был реализован по следующим причинам. Во-первых, он нарушает существенную часть концепции СТ-фильтрации — предположение о небольшом по отношению к размерности задачи числе сигма-точек (в двумерном случае стандартно рекомендуется пять сигма-точек). Во-вторых, такой подход для СТ-фильтрации не будет универсальным. Он применим в данной частной ситуации с тремя переключающимися каналами, так как 9 сигма-точек — не слишком большое число, но с ростом числа каналов станет неприемлем. Наконец, слишком больших вычислительных ресурсов требует используемая во всех примерах процедура оптимизации параметров СТ-фильтрации α, β, γ из соотношений (9), (10).

5. Модифицированный УМНФ

Результаты выполненных экспериментов с СТ-фильтрами, их сравнение с УМНФ в целом подтверждают эффективность обеих концепций. Тем не менее, СТ-фильтр в любой практической задаче — это всегда вызов с непредсказуемым результатом, и получение хороших результатов — это всегда оригинальная манипуляция параметрами (сигма-точками). Условно-минимаксный фильтр работает всегда, структура подбирается естественно и несложно, но при этом оценка может оказаться слишком консервативной, что характерно для минимаксных задач. Перспективным представляется комбинирование двух концепций. Возможный вариант в этом направлении — модификация УМНФ за счет использования в структуре фильтра оценок условных моментных характеристик, ковариаций, вычисляемых вместе с оценками состояния на каждом шаге фильтрации по аналогии с тем, как это делается в СТ-фильтре. Но в отличие от СТ-фильтров для этого предлагается использовать не сигма-точечное преобразование, а метод Монте-Карло, т.е. имитационное моделирование.

Для описания алгоритма модифицированного УМНФ (МУМНФ) вернемся к соотношениям (5), (6), в которых коэффициенты УМНФ выражаются через математические ожидания и ковариации состояния x_t , базового прогноза ξ_t , ошибки прогноза $x_t - \tilde{x}_t$ и базовой коррекции ζ_t . В реализациях УМНФ операции $E\{\cdot\}$ и $\text{cov}(\cdot, \cdot)$ заменялись соответствующими статистиче-

скими оценками, рассчитанными априорно по смоделированному пучку траекторий. Обозначим соответствующие операторы усреднения $\bar{E}\{\cdot\}$ и $\overline{\text{cov}}(\cdot, \cdot)$. Например, в (5) вместо $E\{x_t\}$ используется $\bar{E}\{x_t\} = \frac{1}{N} \sum_{i=1}^N x_t^i$, в (6) вместо $\text{cov}(\zeta_t, x_t - \tilde{x}_t)$ используется $\overline{\text{cov}}(\zeta_t, x_t - \tilde{x}_t) = \frac{1}{N} \sum_{i=1}^N \zeta_t^i (x_t^i - \tilde{x}_t^i)^T$, где N — число смоделированных траекторий, а $x_t^i, \tilde{x}_t^i, \zeta_t^i, i = 1, \dots, N$, — реализации. В модифицированном фильтре предлагается исключить априорное моделирование и выполнять оценивание параметров фильтра в процессе расчетов для каждой траектории. Соответственно алгоритм МУМНФ таков.

Пусть в момент t имеется \hat{x}_{t-1} — оценка МУМНФ состояния x_{t-1} по наблюдениям $y_\tau, \tau = 1, \dots, t-1, \hat{K}_{t-1}$ — оценка условной ковариации ошибки. На шаге t выполняется моделирование выборок: $\{x_{t-1}^i\}_{i=1}^N$ — из гауссовского распределения со средним \hat{x}_{t-1} и ковариацией \hat{K}_{t-1} , $\{w_t^i, v_t^i\}_{i=1}^N$ — дискретных белых шумов в соответствии с моделью (1), и рассчитываются наборы $\{x_t^i, y_t^i\}_{i=1}^N$ по формулам (1). Прогноз МУМНФ \tilde{x}_t ищется в виде $\tilde{x}_t = \bar{E}\{x_t\}$.

Заметим, что здесь не нужна функция базового прогноза ξ_t , коэффициент $F_t = 0$, поскольку не моделируется выборка оценок \hat{x}_{t-1} , и расчет выполняется также, как и на первом шаге УМНФ. Точность прогноза определяется как $\tilde{K}_t = \overline{\text{cov}}(x_t - \tilde{x}_t, x_t - \tilde{x}_t)$. Далее формируется набор $\{\zeta_t^i\}_{i=1}^N$ и вычисляются коэффициенты $H_t = \overline{\text{cov}}(x_t - \tilde{x}_t, \zeta_t) \overline{\text{cov}}^+(\zeta_t, \zeta_t)$, $h_t = -H_t \bar{E}\{\zeta_t\}$, $\zeta_t^i = \zeta_t(x_t^i, y_t^i)$ и оценка точности $\hat{K}_t = \tilde{K}_t - H_t \overline{\text{cov}}(\zeta_t, x_t - \tilde{x}_t)$. Оценка \hat{x}_t состояния x_t , как и в УМНФ, ищется в виде $\hat{x}_t = \tilde{x}_t + H_t \zeta_t + h_t, \zeta_t = \zeta_t(\tilde{x}_t, y_t)$.

Таким образом, идея МУМНФ состоит в аппроксимации условного распределения x_{t-1} относительно $y_\tau, \tau = 1, \dots, t-1$, гауссовским, а соответствующих моментов — результатами фильтрации, и в вычислении параметров фильтра через оценки математических ожиданий и ковариаций, полученных имитационным моделированием. Заметим, что предположение о гауссовской аппроксимации условного распределения можно считать таким же минимаксным контекстом, что и в УМНФ. Также важно отметить, что размеры моделируемых выборок в УМНФ и МУМНФ должны принципиально отличаться, поскольку расчет коэффициентов УМНФ выполняется априорно и, таким образом, ничем не ограничен, расчет коэффициентов МУМНФ выполняется в процессе фильтрации, т.е. предполагает реализуемость в режиме реального времени. По этой причине в экспериментах для УМНФ моделировались пучки объемов 10^5 и в отдельных примерах 10^6 траекторий, далее в расчете МУМНФ N полагается равным 10^3 .

Для проверки работоспособности МУМНФ и демонстрации продуктивности комбинирования концепций УМНФ и СТ-фильтрации рассмотрим модель регрессии с порогами. Из множества проведенных экспериментов это самая простая модель, которая демонстрирует идентичное поведение УМНФ и СТ-фильтра, с одной стороны, и преимущество МУМНФ — с другой. Основана эта модель на следующей идее. Пусть для скалярного x_t определены n интервалов $(x^{i-1}, x^i], i = 1, \dots, n: -\infty = x^0 < x^1 < \dots < x^{n-1} < x^n = +\infty$. Через $e(x_t)$ обозначим n -мерный единичный вектор: $e(x_t) = (0, \dots, 1, \dots, 0)^T$, где единица стоит на i -м месте, если $x_t \in (x^{i-1}, x^i]$. Пусть $a, b, c \in \mathbb{R}^n$. Вы-

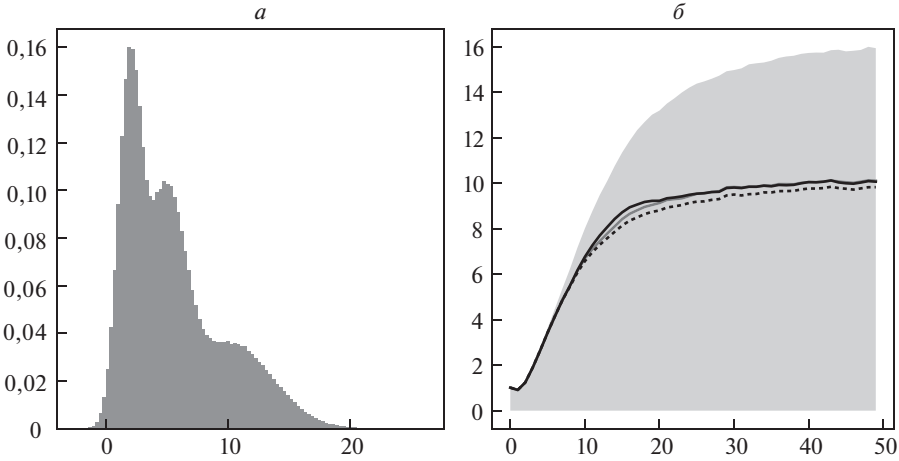


Рис. 8. Модель (19): *a* – гистограмма для x_T , $E[x_t] = 5,56$, $D[x_T] = 15,93$; *б* – дисперсия ошибки оценки СТ-фильтра (серая линия), УМНФ (черная сплошная линия) и МУМНФ (черная точечная линия) на фоне дисперсии процесса $D[x_t]$ (верхняя граница серой области).

ражение $a^T e(x_t)$ дает i -ю координату вектора a , номер которой определяется принадлежностью x_t i -му отрезку. С помощью этого простого приема легко записать, например, n линейных регрессий: i -я регрессия будет иметь вид $a_i x + b_i + c_i w_t$. Окончательная модель имеет вид $x_t = a^T e(x_{t-1}) x_{t-1} + b^T e(x_{t-1}) + c^T e(x_{t-1}) w_t$, т.е. представляет собой комбинацию n авторегрессий первого порядка. Дополнив приведенные соотношения простыми линейными наблюдениями, получаем следующую систему наблюдения:

$$(19) \quad \begin{aligned} x_t &= a^T e(x_{t-1}) x_{t-1} + b^T e(x_{t-1}) + c^T e(x_{t-1}) w_t, & t = 1, 2, \dots, & \quad x_0 = \eta, \\ y_t &= x_t + 10v_t, \end{aligned}$$

где η , w_t , v_t – стандартные гауссовские.

Для проведения расчетов выберем $n = 3$, интервалы $-\infty < 3 < 7 < +\infty$, $a^T = (0,3; 0,4; 0,7)$, $b^T = (1,4; 3,0; 3,0)$, $c^T = (0,9; 1,5; 2,5)$. Как и во всех моделях ранее, и здесь процесс x_t – эргодический, его предельное распределение проиллюстрировано рис. 8,*a*, а выборочные моменты равны $E[x_T] = 5,56$, $D[x_T] = 15,93$.

У этого распределения легко видеть три “пика”, соответствующие трем устойчивым авторегрессиям $x_t = a_i x_{t-1} + b_i + c_i w_t$. Предельное распределение каждой их регрессий, рассмотренной в отдельности от других, будет гауссовским со средним $m_i = b_i / (1 - a_i)$ и дисперсией $\sigma_i^2 = c_i^2 / (1 - a_i^2)$. Образуя вместе с заданными порогами процесс x_t в соответствии с (19), выбранные параметры обеспечивают предельную плотность, показанную на рис. 8,*a*, которой отвечают следующие вероятности $p^1 = P(x_t \leq 3,0) \approx 0,33$, $p^2 = P(3,0 < x_t \leq 7,0) \approx 0,37$, $p^3 = P(x_t > 7,0) \approx 0,3$.

Кроме мультимодального предельного распределения модель (19) еще позволяет продемонстрировать возможность гибкого формирования структуры оценки УМНФ. Причем здесь, в отличие от модели с переключающимися ка-

налами наблюдения (18), структуру фильтра определяет прогноз. Используемая, как и в большинстве примеров приведенных ранее, коррекция в форме невязки естественна для линейных наблюдений. Прогнозирующую структурную функцию $\xi_t(x)$ определим по аналогии с коррекцией для модели (18) на основе отношений правдоподобия, используя обозначение $N(x; m, \sigma^2)$ для гауссовской плотности:

$$\xi_t(x) = \frac{\sum_{i=1}^3 p^i N(x; m_i, \sigma_i^2)(a_i x + b_i)}{\sum_{i=1}^3 p^i N(x; m_i, \sigma_i^2)}.$$

Комментируя такой базовый прогноз, можно сказать, что выражение

$$\sum_{i=1}^3 p^i N(x; m_i, \sigma_i^2)$$

следует рассматривать как элементарную оценку предельной плотности, показанной на рис. 8,а, а отношение

$$p^i N(x; m_i, \sigma_i^2) / \sum_{i=1}^3 p^i N(x; m_i, \sigma_i^2)$$

– как оценку условной вероятности попадания x_{t-1} в i -й интервал.

Результаты расчетов иллюстрирует рис. 8,б. Все три исследуемых фильтра демонстрируют хорошее качество оценивания, МУМНФ с предельной выборочной дисперсией ошибки оценки $D[x_T - \hat{x}_T] = 9,82$ имеет небольшое преимущество порядка 2–3%, СТ-фильтр с $D[x_T - \hat{x}_T] = 10,11$ совсем незначительно проигрывает УМНФ с $D[x_T - \hat{x}_T] = 10,06$. При этом напомним, что для СТ-фильтра используется предварительная оптимизация параметров, УМНФ также требует выполнения значительного объема априорных расчетов, тогда как для МУМНФ нет необходимости в априорных вычислениях, а значит, отсутствуют ограничения в отношении горизонта оценивания. Наилучшие найденные значения параметров СТ-преобразования в данной модели: $\alpha = 0,71$, $\beta = 1,45$, $\gamma = 2,56$ для шага прогнозирования и $\alpha = 0,41$, $\beta = 1,80$, $\gamma = 1,84$ для шага коррекции, они позволяют улучшить точность оценивания СТ-фильтра $D[x_T - \hat{x}_T] = 11,77$ для базовых значений α^* , β^* , γ^* .

В заключение отметим, что подбор простой модели для иллюстрации результатов МУМНФ оказался довольно затруднительным. Продемонстрированный результат обеспечен, по-видимому, характерным предельным распределением (рис. 8,а). Большого эффекта удастся добиться, если отказаться от скалярных уравнений.

6. Заключение

В целом, совокупность экспериментов, проведенных со всеми рассмотренными фильтрами, позволяет на качественном уровне дать характеристику

вычислительной трудоемкости и ресурсоемкости алгоритмов. Необходимость априорного расчета и хранения коэффициентов УМНФ делает этот фильтр самым ресурсоемким. При этом объем хранимых параметров определяется задаваемым горизонтом фильтрации, а значительные затраты на априорное моделирование компенсируются крайней простотой вычисления апостериорных оценок и характеризуются той же вычислительной сложностью, что имеет классический фильтр Калмана в линейной задаче. “Оптимизированный” СТ-фильтр в скалярном случае предполагает проведение такого же небольшого объема вычислений, но этот объем растет с ростом размерности задачи, т.е. с ростом числа $2p + 1$ сигма-точек $x_t^i \in \mathbb{R}^p$, усложняется в связи с необходимостью выполнения вычисления матричнозначного корня на каждом шаге оценивания. Кроме того, надо учитывать, что отсутствие априорных расчетов для СТ-фильтра не гарантирует даже “удачного” выбора его параметров, а оптимизация, выполненная в рассмотренных экспериментах, требует тех же ресурсов, что и моделирование параметров УМНФ в скалярных примерах. Вычислительная сложность такой оптимизации экспоненциально растет с ростом размерности, так что возможность оптимального подбора параметров α , β , γ СТ-фильтра в практически значимых задачах представляется сомнительной. Отсутствие априорных расчетов и затрат на хранение параметров — свойство и преимущество МУМНФ. Достигается оно за счет переноса шагов моделирования в апостериорный расчет. Объем моделирования при этом меньше, чем для априорного расчета коэффициентов УМНФ (в примерах 10^3 против 10^5 и 10^6), но все равно значительный, и представляется целесообразным только при наличии соответствующих вычислительных возможностей, которыми обладают современные компьютерные системы.

СПИСОК ЛИТЕРАТУРЫ

1. Julier S.J., Uhlmann J.K., Durrant-Whyte H.F. A New Approach for Filtering Nonlinear Systems // Proc. IEEE Amer. Control Conf. (ACC'95). 1995. P. 1628–1632.
2. Menegaz H.M.T., Ishihara J.Y., Borges G.A., Vargas A.N. A Systematization of the Unscented Kalman Filter Theory // IEEE Trans. Autom. Control. 2015. V. 60. No. 10. P. 2583–2598.
3. Julier S.J. The Scaled Unscented Transformation // Proc. IEEE Amer. Control Conf. (ACC'02). 2002. P. 4555–4559.
4. Xu L., Ma K., Fan H. Unscented Kalman Filtering for Nonlinear State Estimation with Correlated Noises and Missing Measurements // Int. J. Control Autom. Syst. 2018. V. 16. No. 3. P. 1011–1020.
5. Li L., Xia Y. Stochastic Stability of the Unscented Kalman Filter with Intermittent Observations // Automatica. 2012. V. 48. No. 5. P. 978–981.
6. Lee D., Vukovich G., Lee R. Robust Unscented Kalman Filter for Nanosat Attitude Estimation // Int. J. Control Autom. Syst. 2017. V. 15. No. 53. P. 2161–2173.
7. Zhao Y., Gao S.-S., Zhang J., Sun Q.-N. Robust Predictive Augmented Unscented Kalman Filter // Int. J. Control Autom. Syst. 2014. V. 12. No. 5. P. 996–1004.
8. Scardua L.A., da Cruz J.J. Complete Offline Tuning of the Unscented Kalman Filter // Automatica. 2017. V. 80. P. 54–61.

9. *Straka O., Dunik J., Simandl M.* Unscented Kalman Filter with Advanced Adaptation of Scaling Parameter // *Automatica*. 2014. V. 50. No. 10. P. 2657–2664.
10. *Dunik J., Simandl M., Straka O.* Unscented Kalman Filter: Aspects and Adaptive Setting of Scaling Parameter // *IEEE Trans. Autom. Control*. 2012. V. 57. No. 9. P. 2411–2416.
11. *Biswas S.K., Qiao L., Dempster A.G.* A Novel a Priori State Computation Strategy for the Unscented Kalman Filter to Improve Computational Efficiency // *IEEE Trans. Autom. Control*. 2017. V. 62. No. 4. P. 1852–1864.
12. *Sarkka S.* On Unscented Kalman Filtering for State Estimation of Continuous-Time Nonlinear Systems // *Trans. Autom. Control*. 2007. V. 52. No. 9. P. 1631–1641.
13. *Li X., Liu A., Yu C., Su F.* Widely Linear Quaternion Unscented Kalman Filter for Quaternion-Valued Feedforward Neural Network // *IEEE Signal Process. Lett.* 2017. V. 24. No. 9. P. 1418–1422.
14. *Bhotto M.Z.A., Bajic I.V.* Constant Modulus Blind Adaptive Beamforming Based on Unscented Kalman Filtering // *IEEE Signal Process. Lett.* 2015. V. 22. No. 4. P. 474–478.
15. *Li L., Xia Y.* Unscented Kalman Filter over Unreliable Communication Networks with Markovian Packet Dropouts // *Trans. Autom. Control*. 2013. V. 58. No. 12. P. 3224–3230.
16. *Wu P., Li X., Bo Y.* Iterated Square Root Unscented Kalman Filter for Maneuvering Target Tracking Using TDOA Measurements // *Int. J. Control Autom. Syst.* 2013. V. 11. No. 4. P. 761–767.
17. *Jochmann G., Kerner S., Tasse S., Urbann O.* Efficient Multi-Hypotheses Unscented Kalman Filtering for Robust Localization // *Lect. Notes Comput. Sci.* (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics) 7416 LNCS. 2012. P. 222–233.
18. *Leven W.F., Lanterman A.D.* Unscented Kalman Filters for Multiple Target Tracking with Symmetric Measurement Equations // *Trans. Autom. Control*. 2009. V. 54. No. 2. P. 370–375.
19. *Пугачев В.С.* Рекуррентное оценивание переменных и параметров в стохастических системах, описываемых разностными уравнениями // *ДАН СССР*. 1978. Т. 243. № 5. С. 1131–1133.
20. *Пугачев В.С.* Оценивание переменных и параметров в дискретных нелинейных системах // *АиТ*. 1979. № 6. С. 63–79.
Pugachev V.S. Estimation of Variables and Parameters in Discrete-Time Nonlinear Systems // *Autom. Remote Control*. 1979. V. 40. No. 4. P. 39–50.
21. *Панков А.Р.* Рекуррентная условно-минимаксная фильтрация процессов в разностных нелинейных стохастических системах // *Изв. РАН. Теория и системы управления*. 1992. № 3. С. 63–70.
Pankov A.R. Recurrent Conditionally Minimax Filtering of Processes In Nonlinear Difference Stochastic Systems // *J. Comput. Syst. Sci. Int.* 1993. V. 31. No. 4. P. 54–60.
22. *Pankov A.R., Bosov A.V.* Conditionally Minimax Algorithm for Nonlinear System State Estimation // *IEEE Trans. Autom. Control*. 1994. V. 39. No. 8. P. 1617–1620.
23. *Борисов А.В., Босов А.В., Кибзун А.И., Миллер Г.Б., Семенихин К.В.* Метод условно-оптимальной нелинейной фильтрации и современные подходы к оцениванию состояний нелинейных стохастических систем // *АиТ*. 2018. № 1. С. 3–17.

- Borisov A.V., Bosov A.V., Kibzun A.I., Miller G.B., Semenikhin K.V.* The Conditionally Minimax Nonlinear Filtering Method and Modern Approaches to State Estimation in Nonlinear Stochastic Systems // Autom. Remote Control. 2018. V. 79. No. 1. P. 1–11.
24. *Wan E.A., Van der Merwe, R.* The Unscented Kalman Filter/ Kalman Filtering And Neural Networks / Haykin S. (ed). N.Y.: John Wiley & Sons. 2001. P. 221–280.
25. *Шуряев А.Н.* Вероятность. М.: Наука, 1989.
26. *Bhattacharya R.N., Lee C.* Ergodicity of Nonlinear First Order Autoregressive Models // J. Theor. Probab. 1995. V. 8. No. 1. P. 207–219.
27. *May R.M.* Simple Mathematical Models with Very Complicated Dynamics // Nature. 1976. V. 261. P. 459–467.
28. *Nahi N.* Optimal Recursive Estimation with Uncertain Observation // IEEE Trans. Inform. Theory. 1969. V. 15. No. 4. P. 457–462.

Статья представлена к публикации членом редколлегии М.М. Хрустальевым.

Поступила в редакцию 15.11.2018

После доработки 05.02.2019

Принята к публикации 07.02.2019