На правах рукописи

Лебедев Максим Витальевич

АЛГОРИТМЫ ОПТИМАЛЬНОГО ОЦЕНИВАНИЯ В СТОХАСТИЧЕСКИХ СИСТЕМАХ В УСЛОВИЯХ АПРИОРНОЙ НЕОПРЕДЕЛЕННОСТИ

05.13.01 — Системный анализ, управление и обработка информации
(авиационная и ракетно-космическая техника)

АВТОРЕФЕРАТ
диссертации на соискание ученой степени
кандидата физико-математических наук

Москва, 2008

Работа выполнена на кафедре теории вероятностей Московского авиационного института (государственного технического университета).

Научный руководитель: кандидат физико-математических наук Семенихин Константин Владимирович

Официальные оппоненты: доктор физико-математических наук, профессор Кузнецов Евгений Борисович

кандидат физико-математических наук, н.с. Миллер Григорий Борисович

Ведущая организация: Институт проблем управления РАН

Защита состоится “19” декабря 2008 г. в 10 ч. 00 мин. на заседании Диссертационного совета Д 212.125.04 при Московском авиационном институте (государственном техническом университете) по адресу: 125993, Москва, А-80, ГСП-3, Волоколамское ш., 4, Ученый совет МАИ.

С диссертацией можно ознакомиться в библиотеке Московского авиационного института (государственного технического университета).

Отзыв на автореферат, заверенный гербовой печатью организации, просьба направлять по указанному адресу в двух экземплярах.

Автореферат разослан “___” __________ _____ г.

Ученый секретарь Диссертационного совета Д 212.125.04,
кандидат физико-математических наук М.В. Ротанина

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Объект исследования. В диссертационной работе исследованы задачи оптимального оценивания в стохастических системах при наличии существенной априорной неопределенности.

Актуальность темы. На сегодняшний день с развитием информационных технологий все большее значение приобретают математические методы, связанные с обработкой и анализом эмпирической информации. Несомненно, среди этих методов — алгоритмы оценивания и фильтрации являются ключевыми.

Зачастую в реальных практических задачах не удается построить полностью определенную адекватную математическую модель, в которой можно было бы воспользоваться известными методами для оценивания параметров и сигналов. Иногда вообще отсутствует информация о природе тех или иных процессов, а известны лишь некоторые достаточно широкие ограничения на их поведение. Таким образом, возникает задача оптимального оценивания и фильтрации в стохастических системах при наличии априорной неопределенности.

К настоящему моменту в этой области обработки информации сформировались два основных подхода: робастный и адаптивный. В соответствии с адаптивным подходом недостающая априорная информация извлекается из нарастающего массива эмпирических данных. Исследованию адаптивного подхода в задачах оценивания и фильтрации посвящены работы Я.З. Цыпкина, Б.Т. Поляка, В.Н. Фомина, С.П. Урясьева, А.В. Назина, L. Ljung'а и др.

В представляемой диссертации исследуются постановки, имеющие дело с фиксированным ограниченным набором наблюдений. В такой ситуации альтернативой адаптивным методам являются методы, основанные на робастном или минимаксном подходе. В нашей стране к основоположникам этого подхода относятся Н.Н. Красовский, А.Б. Куржанский, М.Л. Лидов, Б.Т. Поляк, П.Е. Эльясберг. Среди зарубежных специалистов можно отметить C.J. Martin, M. Mintz, H.V. Poor, P. Huber, S. Verdu, A. Wald. Дальнейшее развитие данной тематики связано с работами Б.И, Ананьева, Б.Ц. Бахшияна, Г.А. Голубева, М.И. Гусева, И.Я. Каца, А.И. Матасова, А.Г. Наконечного, А.Р. Панкова, Ю.П. Пытьева, Е.Я. Рубиновича, В.Н. Соловьева, Г.А. Тимофеевой, В.И. Ширяева, S. Boyd'а, L.El Ghaoui, G. Calafiore.

Итак, при робастном подходе требуется (по ограниченному набору наблюдений) указать фиксированную оценку, чье наихудшее качество на заданном классе неопределенности будет наилучшим по сравнению с другими допустимыми оценками. Тем самым, задачу робастного оценивания можно сформулировать в виде игровой постановки, в которой критерий (погрешность оценивания) зависит от пары элементов, выбираемых из пары заданных множеств, содержащих соответственно допустимые операторы оценивания и возможные характеристики модели наблюдения.

Основная задача диссертационной работы — синтез алгоритмов минимаксного оценивания для нескольких типов моделей наблюдения. В первой главе минимаксная оценка неопределенно-стохастического вектора построена аналитически при наличии ошибок наблюдений неизвестной ковариационной структуры. Во второй главе диссертации искомая оценка находится итерационно при поэлементных ограничениях на ковариации ошибок наблюдения. И наконец, третья глава посвящена процедуре минимаксной фильтрации в стохастической дифференциальной системе с неопределенными интенсивностями нестационарных возмущений. Таким образом, в первой главе неопределенность описывается обширным множеством ковариационных матриц, во второй главе неопределенность носит алгебраический характер (заранее неизвестно регулярна ли модель наблюдения или нет), в третьей главе неопределенность непараметрическая, т.е. неизвестными являются функции. Общим для указанных моделей наблюдения является то, что непосредственное применение к ним существующих алгоритмов минимаксного оценивания оказывается чрезвычайно трудоемким. С целью подчеркнуть эту особенность рассматриваемых моделей в работе использован термин «существенная априорная неопределенность».

Во всех неопределенно-стохастических системах, рассматриваемых в диссертационной работе, прямой синтез минимаксного алгоритма оценивания является трудноразрешимой задачей. Поэтому используется подход, основанный на переходе к двойственной (максиминной) задаче. Применение такого подхода к задачам минимаксного оценивания было отмечено ранее в работах В.Б. Меласа, S. Verdu, H.V. Poor'а, И.Ф. Пинелиса, В.Н. Соловьева, А.Р. Панкова, К.В. Семенихина.

Цель работы — исследование и решение задач оптимального оценивания по среднеквадратическому критерию в стохастических системах в условиях существенной априорной неопределенности различного типа.

Достижение указанной цели подразумевает выполнение следующих основных этапов данной работы:

1) аналитический синтез минимаксной оценки случайного вектора в присутствии случайных ошибок двух видов: белошумной и произвольно

коррелированной;

2) сравнение качества минимаксного оценивания при различных ограничениях на произвольно коррелированные ошибки;

3) формулировка и обоснование алгоритма оптимального оценивания в стохастической линейной многомерной модели наблюдения при наличии алгебраической неопределенности относительно ковариационной структуры случайных ошибок;

4) статистическое моделирование и сравнительный анализ разработанного алгоритма минимаксного оценивания на примере модели с поэлементными ограничениями на ковариацию погрешности наблюдения;

5) обоснование алгоритма минимаксной фильтрации в стохастической дифференциальной системе с нестационарными возмущениями неизвестной интенсивности;

6) аналитический синтез минимаксного фильтра скалярного состояния стохастической дифференциальной системы при произвольных нестационарных белошумных возмущениях с неизвестной взаимной корреляцией.

Методы исследования. В диссертационной работе используются методы теории оптимизации, выпуклого анализа, теории двойственности, линейной алгебры, теории вероятностей, математической статистики, функционального анализа и теории управления, численные методы выпуклого программирования, а также современные средства компьютерного моделирования.

Научная новизна.

1. В работе сформулирована задача минимаксного оценивания вектора состояния конечномерной стохастической системы по среднеквадратичному критерию при наличии интегрального ограничения на дисперсию произвольно коррелированных ошибок. Для указанных ошибок найдена наименее благоприятная ковариация, на основе которой построена искомая минимаксная оценка вектора состояния. Проведено сравнение качества минимаксной оценки при различных видах ограничений на ковариацию случайных ошибок.

2. Решена задача минимаксного оценивания вектора состояния конечномерной стохастической системы при наличии поэлементных ограничений на ковариационную матрицу вектора ошибок. Для данной модели наблюдения сформулирован алгоритм совместного решения минимаксной и двойственной задачи.

3. Описано решение задачи минимаксной фильтрации в стохастической дифференциальной системе с нестационарными возмущениями

неизвестной интенсивности. Получен явный вид минимаксного фильтра скалярного состояния в модели наблюдения с белошумными возмущениями произвольной интенсивности.

Практическая ценность и теоретическая значимость. Полученные результаты составляют теоретическую базу для решения многих практических задач обработки информации в отсутствии точных математических моделей, описывающих структуру случайных возмущений и помех наблюдений. Среди этих задач можно отметить следующие: оценивание параметров движения летательных аппаратов, статистическая обработка внешнетраекторных наблюдений и результатов летных испытаний авиационной и ракетно-космической техники.

Результаты диссертации позволяют провести сравнение используемых на практике алгоритмов оценивания и фильтрации с оптимальными методами, обеспечивающими гарантированное качество оценок.

Апробация работы. Результаты диссертационной работы обсуждались на научных конференциях и симпозиумах: 44th IEEE Conference on Decision and Control and European Control Conference (CDC-ECC) (2005, Spain, Seville); «Научная сессия МИФИ 2008» (2008, Россия, Москва); «Информационные технологии в авиационной и космической технике-2008» (2008, Россия, Москва); 13-ая Международная конференция «Системный анализ и управление космическими комплексами» (2008, Украина, Евпатория); 7-ая Международная конференция «Авиация и космонавтика - 2008» (2008, Россия, Москва), а также на научных семинарах под руководством проф. А.И. Кибзуна (МАИ) и проф. Б.Т. Поляка (ИПУ РАН).

Публикации. Основные результаты диссертации опубликованы в пяти статьях [1–5] ([1, 2] — в журналах из перечня ВАК, [3–5] — в сборниках научных трудов), а также в тезисах научных конференций [6–9].

Структура и объем диссертации. Диссертация состоит из введения, 3 глав, заключения и списка литературы (95). Объем диссертации — 80 м.п.с.

СОДЕРЖАНИЕ РАБОТЫ

Во введении представлен краткий обзор современного состояния исследуемой проблемы, сформулированы цель и задачи диссертационной работы, перечислены полученные в диссертации новые результаты, обоснованы их актуальность и практическая ценность, описана структура диссертации.

В первой главе представлено аналитическое решение задачи минимаксного оценивания в многомерной линейной стохастической системе при наличии произвольно коррелированных ошибок наблюдений:

$$X = A\theta, \qquad Y = B\theta + \eta + \alpha\,\upsilon, \tag{1.1}$$

где $X \in \mathbb{R}^m$ — вектор состояния, который требуется восстановить по наблюдениям $\{Y_1, \ldots, Y_n\}$, образующим вектор $Y \in \mathbb{R}^n$; $\theta \in \mathbb{R}^p$ — вектор случайных параметров системы; $\upsilon \in \mathbb{R}^n$ и $\eta \in \mathbb{R}^n$ — векторы случайных помех наблюдений; A и B — известные матрицы соответствующего размера; α — заданное положительное число. Известно, что θ, η, υ центрированы и некоррелированы между собой, а также имеют ковариационные матрицы, удовлетворяющие условиям

$$\mathsf{cov}\{\theta,\theta\} \leqslant T, \qquad \mathrm{tr}\big[\mathsf{cov}\{\eta,\eta\}\big] \leqslant \sigma^2, \qquad \mathsf{cov}\{\upsilon,\upsilon\} = I, \tag{1.2}$$

где положительно определенная матрица $T \in \mathbb{R}_+^{p\times p}$ и число $\sigma > 0$ заданы. Здесь и далее $\mathbb{R}_+^{n\times n}$ — семейство симметричных неотрицательно определенных матриц размера $n \times n$, а неравенство между матрицами $S \leqslant T$ означает $T - S \in \mathbb{R}_+^{n\times n}$.

Цель данной главы — построить наилучшую оценку случайного вектора X в модели наблюдения (1.1) при наличии описанной выше априорной неопределенности.

Рассмотрим произвольную линейную оценку $\widetilde{X} = FY$, $F \in \mathbb{R}^{m\times n}$ и предположим, что ее точность определяется величиной с.к. погрешности $\mathsf{E}_\Pi\|X - \widetilde{X}\|^2$, где Π обозначает совместное распределение случайных векторов θ, η и υ. Для наилучшего оценивания достаточно найти оператор оценивания $\hat{F} \in \mathbb{R}^{m\times n}$, на котором указанный с. к. критерий достигает наименьшего значения. Однако искомый оператор оценивания $\hat{F}$ будет зависеть от распределения Π. Поскольку оно известно не точно, а задано лишь посредством условий (1.2), найти наилучший оператор оценивания без привлечения дополнительной априорной информации о модели наблюдения невозможно. Поэтому для корректной формулировки задачи оптимального оценивания используется минимаксный подход.

Определение 1.1. Оценка $\hat{X} = \hat{F}Y$ называется оптимальной в минимаксном смысле (*или просто минимаксной*), *если*

$$\hat{F} \in \operatorname*{arg\,min}_{F\in\mathbb{R}^{m\times n}} \max_{\Pi\in\mathcal{P}} \mathsf{E}_\Pi\|X - FY\|^2,$$

где $\mathcal{P}$ — класс всех совместных распределений векторов θ, η, υ, удовлетворяющих условиям (1.2).

Итак, искомая оценка доставляет минимум гарантированному (т. е. наихудшему) значению среднеквадратического критерия.

В работе показано, что

$$\max_{\Pi\in\mathcal{P}} \mathsf{E}_\Pi \|X - FY\|^2 = \max_{R\in\mathcal{R}} J(F,R),$$

где обозначено

$$J(F,R) = \mathrm{tr}\big[(FB-A)T(FB-A)^* + F(R+\alpha^2 I)F^*\big]\,,$$
$$\mathcal{R} = \{R\in\mathbb{R}_+^{n\times n}:\ \mathrm{tr}R \leqslant \sigma^2\}. \tag{1.3}$$

Таким образом, задача минимаксного оценивания принимает вид

$$\hat F \in \arg\min_{F\in\mathbb{R}^{m\times n}} \max_{R\in\mathcal{R}} J(F,R). \tag{1.4}$$

Наряду с (1.4) в данной главе рассматривается также задача, двойственная по отношению к указанной, а именно, проблема нахождения наименее благоприятной ковариации $\hat R$:

$$\hat R \in \arg\max_{R\in\mathcal{R}} \underline{J}(R), \quad \underline{J}(R) = \min_{F\in\mathbb{R}^{m\times n}} J(F,R). \tag{1.5}$$

Для построения минимаксной оценки $\hat X = \hat F Y$ используется метод, основанный на решении двойственной задачи. Суть его заключается в том, что минимаксный оператор оценивания $\hat F$ ищется в виде решения задачи оптимального оценивания

$$F_R \in \arg\min_{F\in\mathbb{R}^{m\times n}} J(F,R) \quad \text{при} \quad R = \hat R,$$

где $\hat R$ — наименее благоприятная ковариационная матрица (1.5). Тем самым оценка $F_{\hat R}Y$ построена по наилучшему алгоритму оценивания в той ситуации, когда для модели наблюдения реализовался наихудший случай, определяемый решением $\hat R$ двойственной задачи (1.5).

Отметим, что метод, использующий непосредственную минимизацию максимума с. к. погрешности, к задаче (1.4) оказывается неприменим.

В общем случае структура минимаксной оценки и наименее благоприятной ковариации описывается в следующей теореме.

Теорема 1.1. *Для сформулированной задачи минимаксного оценивания* (1.4) *справедливы следующие утверждения*:

1) *решения минимаксной и двойственной задач* (*см.* (1.4) *и* (1.5) *соответственно*) *существуют и образуют седловую точку*

$$J(\hat F, R) \leqslant J(\hat F, \hat R) \leqslant J(F, \hat R) \quad \forall\, F\in\mathbb{R}^{m\times n},\ R\in\mathcal{R};$$

2) $F_{\hat{R}} = ATB^*(BTB^* + \hat{R} + \alpha^2 I)^{-1}$ — *минимаксный оператор оценивания, если* $\hat{R}$ — *решение двойственной задачи* (1.5);

3) *если симметричная матрица* $\hat{R}$ *представляет собой решение задачи* (1.5), *то существует число* $\lambda > 0$ *такое, что*

$$\mathrm{tr}\hat{R} = \sigma^2, \quad \hat{R} \geqslant O, \quad \|QF^*_{\hat{R}}F_{\hat{R}}Q\| \leqslant \lambda, \quad P\big(\lambda I - F^*_{\hat{R}}F_{\hat{R}}\big)P = O, \tag{1.6}$$

где $P = \hat{R}\hat{R}^+$ *и* $Q = I - P$;

4) *если же матрицы* Q *и* $F^*_{\hat{R}}F_{\hat{R}}$ *перестановочны, то приведенные условия* (1.6) *являются достаточными для того, чтобы* $\hat{R}$ *была решением двойственной задачи* (1.5).

В работе для искомой минимаксной оценки удалось найти аналитическое выражение при следующих предположениях:

а) вектор состояния системы при отсутствии возмущений восстанавливается безошибочно, иначе говоря,

$$A = F_0 B, \quad \text{где} \quad F_0 = AB^+; \tag{1.7}$$

б) матрицы $V = BTB^*$ и $U = F_0^* F_0$ перестановочны т. е. $UV = VU$ или, что тоже самое, они допускают одинаковое спектральное разложение

$$V = \sum_{i=1}^{l} v_i P_i, \qquad U = \sum_{i=1}^{l} u_i P_i, \qquad I = \sum_{i=1}^{l} P_i, \tag{1.8}$$

где v_i и u_i — собственные значения ($v_i \geqslant 0$ и $u_i \geqslant 0$); P_i — ортопроекторы на взаимно ортогональные подпространства; $1 \leqslant l \leqslant p$.

Первое предположение известно в регрессионном анализе как условие идентифицируемости и может быть записано в виде $\ker B \subseteq \ker A$. Оно заведомо выполнено, если B — матрица полного ранга.

Второе условие означает, что при отсутствии помех вектор состояния — есть результат преобразования вектора наблюдений в спектральной области. В частности, данное предположение будет выполнено, если $X = \theta$ или $X = B\theta$.

Без ограничения общности можно предположить, что всем парам собственных значений v_i и u_i приписаны индексы $i = 1, \ldots, l$ так, что

$$\frac{v_1^2 u_1}{(v_1 + \alpha^2)^2} \geqslant \ldots \geqslant \frac{v_l^2 u_l}{(v_l + \alpha^2)^2} \geqslant 0.$$

Решения минимаксной и двойственной задач представлены в приведенной ниже теореме. Для ее формулировки введем некоторые вспомогательные обозначения: пусть даны индекс $k \in \{1, \ldots, l\}$ и положительные числа λ, $r_1, \ldots, r_k$, удовлетворяющие соотношениям

$$\frac{v_k^2 u_k}{(v_k+\alpha^2)^2} > \lambda \geqslant \frac{v_{k+1}^2 u_{k+1}}{(v_{k+1}+\alpha^2)^2}, \tag{1.9}$$

$$\frac{v_i^2 u_i}{(v_i+r_i+\alpha^2)^2} = \lambda, \qquad i=1,\ldots,k, \quad \sum_{i=1}^{k} d_i r_i = \sigma^2, \tag{1.10}$$

где d_i — ранг ортопроектора P_i (т. е. $d_i = \operatorname{tr} P_i$) и $v_{l+1} = u_{l+1} = 0$.

Теорема 1.2. *Пусть в модели наблюдения* (1.1)–(1.2) *выполнены условия* (1.7) *и* (1.8). *Тогда минимаксный оцениватель, наименее благоприятная ковариация и оптимальное гарантированное значение с. к. погрешности имеют следующий вид*:

$$\hat{F} = ATB^*(V+\hat{R}+\alpha^2 I)^{-1} = \sum_{i=1}^{l} \frac{1}{v_i+r_i+\alpha^2} ATB^* P_i, \tag{1.11}$$

$$\hat{R} = \sum_{i=1}^{k} r_i P_i, \qquad r_{k+1} = \ldots = r_l = 0, \tag{1.12}$$

$$\hat{J} = J(\hat{F},\hat{R}) = \sum_{i=1}^{l} \frac{v_i u_i d_i (r_i+\alpha^2)}{v_i+r_i+\alpha^2}, \tag{1.13}$$

причем $k, \lambda, r_1, \ldots, r_k$, *удовлетворяющие* (1.9)–(1.10), *действительно существуют и определены единственным образом.*

Также в данной главе представлено несколько примеров для иллюстрации полученной минимаксной оценки.

В примерах 1.1, 1.2 проведено сравнение качества оценивания в частной модели наблюдения (1.1) при различных ограничениях на ковариационную матрицу вектора ошибок наблюдения.

В примерах 1.3, 1.4 для двух многомерных моделей наблюдения при помощи теоремы 1.2 найдены явные представления для минимаксных оценок.

Во второй главе получен итерационный алгоритм минимаксного оценивания вектора состояния в конечномерной линейной неопределенно-стохастической системе

$$X = A\theta, \qquad Y = B\theta + \eta,$$

при наличии поэлементных ограничений на ковариационную матрицу вектора ошибок η

$$\mathsf{cov}\{\eta,\eta\} \in \mathcal{R}_e = \{R \in \mathbb{R}_s^{n\times n} : |R_{i,j} - R^0_{i,j}| \leqslant \Delta_{i,j},\ i,j = 1,\ldots,n\}, \tag{2.14}$$

где $\Delta_{i,j} \geqslant 0$ и $\Delta \in \mathbb{R}_s^{n\times n}$, а $R^0 \in \mathbb{R}_+^{n\times n}$ — некоторая опорная матрица. $\mathbb{R}_s^{n\times n}$ — семейство симметричных матриц размера $n \times n$.

Исследуемая задача минимаксного оценивания имеет вид

$$\hat{F} \in \arg\min_{F\in\mathbb{R}^{m\times n}} \max_{R\in\mathcal{R}} J(F,R), \tag{2.15}$$
$$J(F,R) = \operatorname{tr}[(FB-A)T(FB-A)^* + FRF^*]$$

где множество неопределенности имеет следующую структуру: $\mathcal{R} = \mathbb{R}_+^{n\times n} \cap \mathcal{R}_e$.

Рассматриваемая в данной главе модель наблюдения может оказаться сингулярной, поэтому для решения задачи минимаксного оценивания (2.15) применяется метод регуляризации Тихонова. Для этого вместо функционала $J(F,R)$ рассматривается его регуляризованный вариант $J^\varepsilon(F,R) = J(F,R) + \varepsilon \operatorname{tr}[FF^*]$, где $\varepsilon > 0$ — некоторая константа. В этом случае задача (2.15) минимаксного оценивания принимает вид

$$\hat{F}^\varepsilon \in \arg\min_{F\in\mathbb{R}^{m\times n}} \max_{R\in\mathcal{R}} J^\varepsilon(F,R). \tag{2.16}$$

Тогда в соответствии с методом двойственной оптимизации

$$\hat{F}^\varepsilon \in \arg\min_{F\in\mathbb{R}^{m\times n}} J^\varepsilon(F,\hat{R}),$$

где $\hat{R}$ — решение регуляризованной двойственной задачи

$$\hat{R} \in \arg\max_{R\in\mathcal{R}} \underline{J}^\varepsilon(R), \quad \underline{J}^\varepsilon(R) = \inf_{F\in\mathbb{R}^{m\times n}} J^\varepsilon(F,R). \tag{2.17}$$

Искомый алгоритм оценивания существенно упрощается, если расширить множество неопределенности и заменить исходный критерий $J^\varepsilon(F,R)$ функцией Лагранжа $L^\varepsilon(F,R,\Lambda) = J^\varepsilon(F,R) + \operatorname{tr}[\Lambda R]$. Тогда новая двойственная задача примет следующий вид:

$$(\hat{R},\hat{\Lambda}) \in \arg\max_{R\in\mathcal{R}_e,\, O\leqslant\Lambda\leqslant\overline{\lambda}I} \underline{L}^\varepsilon(R,\Lambda), \tag{2.18}$$

где $\underline{L}^\varepsilon(R,\Lambda) = \underline{J}^\varepsilon(R) + \operatorname{tr}[\Lambda R]$, $\overline{\lambda} > 0$.

В следующей теореме показано, что задача (2.18) эквивалентна исходной двойственной задаче (2.17).

Т е о р е м а 2.3. *Пусть множество $\mathcal{R}$ удовлетворяет условию Слейтера $\exists \overline{R} \in \mathbb{R}_s^{n\times n}: \quad \overline{R} > 0,\ |\overline{R}_{i,j} - R^0_{i,j}| < \Delta_{i,j}$ и $\overline{\lambda} = \dfrac{\operatorname{tr}[ATA^*] - \underline{J}^\varepsilon(\overline{R})}{n\min\sigma(\overline{R})}$. Тогда $\hat{R}$ является решением задачи* (2.17) *тогда и только тогда, когда существует матрица $\hat{\Lambda}$: $O \leqslant \hat{\Lambda} \leqslant \overline{\lambda}I$ такая, что $\hat{\Lambda}$ и $\hat{R}$ образуют решение задачи* (2.18).

В сформулированной теореме $\min \sigma(R)$ — минимальное собственное значение матрицы R.

В итоге для поиска минимаксной оценки в работе сформулирован итеративный алгоритм, основанный на методе условного градиента.

Алгоритм 2.1. 1) выбрать матрицу $\overline{R}$ в качестве начальной $R^{(0)} = \overline{R}$ и положить $s = 0$;

2) задать число

$$\overline{\lambda} = \frac{\mathrm{tr}[ATA^*] - \underline{J}^{\varepsilon}(\overline{R})}{n \min \sigma(R^0)};$$

3) найти оператор оценивания

$$F^{(s)} = ATB^*(BTB^* + R^{(s)} + \varepsilon^2 I)^{-1};$$

4) взять матрицу $\Lambda^{(s)} = \overline{\lambda} P^{(-)}_{R^{(s)}}$, где $P^{(-)}_{R^{(s)}}$ — ортопроектор на подпространство, соответствующее отрицательным собственным числами $R^{(s)}$;

5) составить матрицу $\tilde{R}^{(s)}$ такую, что

$$\tilde{R}^{(s)}_{i,j} = R^0_{i,j} + \Delta_{i,j} \,\mathrm{sign}\left(F^{(s)^*} F^{(s)} + \Lambda^{(s)}\right)_{i,j}, \quad i, j = 1, \ldots, n;$$

6) если $\delta^{(s)} = 0$, где $\delta^{(s)} = L^{\varepsilon}(F^{(s)}, \tilde{R}^{(s)}, \Lambda^{(s)}) - L^{\varepsilon}(F^{(s)}, R^{(s)}, \Lambda^{(s)})$, то $\hat{R} = R^{(s)}$ и $\hat{F} = F^{(s)}$, иначе перейти к шагу 7);

7) определить число $\gamma^{(s)}$ из условия

$$\gamma^{(s)} \in \arg \max_{\gamma \in [0,1]} \underline{L}^{\varepsilon}((1 - \gamma) R^{(s)} + \gamma \tilde{R}^{(s)});$$

8) положить $R^{(s+1)} = (1 - \gamma^{(s)}) R^{(s)} + \gamma^{(s)} \tilde{R}^{(s)}$, увеличить s на 1 и перейти к шагу 3).

В конце данной главы представлены примеры, иллюстрирующие применение алгоритма 2.1 к нескольким частным моделям наблюдения.

Пример 2.1 демонстрирует применение разработанного алгоритма к задаче определения параметров движения летательного аппарата по результатам траекторных наблюдений при наличии немоделируемых возмущений в характеристиках ошибок наблюдений.

В примере 2.2. рассматривается модель простой линейной регрессии

$$Y_j = \theta_1 + x_j \theta_2 + \eta_j + \varepsilon \upsilon_j, \quad j = 1, \ldots, n,$$

в которой требуется восстановить вектор случайных параметров θ_1, θ_2 по скалярным измерениям Y_j в присутствии случайных помех η_j и υ_j.

Данная модель измерений рассматривается в следующих предположениях: θ случайный вектор, у которого компоненты θ_1 и θ_2 являются некоррелированными и имеют нулевые средние и одинаковые дисперсии $D > 0$; $\{\eta_1, \dots, \eta_n\}$ — центрированные случайные величины с неопределенной ковариационной матрицей R; $\{\upsilon_1, \dots, \upsilon_n\}$ — некоррелированные нормированные случайные величины; векторы θ, η, υ — попарно не коррелированы.

Кроме того, для случайного вектора η априорная неопределенность описывается поэлементными ограничениями на его ковариационную матрицу: $|R_{i,j}| \leqslant \sigma_\eta^2, \quad i, j = 1, \dots, n$. Этим ограничениям соответствует множество вида (2.14):

$$\mathcal{R} = \{R \in \mathbb{R}_+^{n\times n} : |R_{i,j} - R_{i,j}^0| \leqslant \Delta_{i,j}, \quad i, j = 1, \dots, n\},$$

где $R^0 = \dfrac{\sigma_\eta^2}{2} I, \quad \Delta_{i,j} = \sigma_\eta^2/2 \quad$ при $i = j$ и $\Delta_{i,j} = \sigma_\eta^2 \quad$ при $i \neq j$.

Для представленной модели было проведено численное моделирование в следующих условиях: векторы θ, η и υ являлись независимыми в совокупности и распределенными по нормальному закону, а именно,

$$\theta \sim \mathcal{N}(0, DI), \qquad \eta \sim \mathcal{N}(0, \widetilde{R}), \qquad \upsilon \sim \mathcal{N}(0, I),$$

где

$$\widetilde{R} = \sigma_\eta^2 \begin{bmatrix} 1 & 0 & 1 & \dots & 0 \\ 0 & 1 & 0 & \dots & 1 \\ 1 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 1 & 0 & \dots & 1 \end{bmatrix}.$$

Наряду с истинной ковариационной матрицей $\widetilde{R}$ и наименее благоприятной $\hat{R}$ был взят еще один пример допустимой ковариационной матрицы $R_0 = \sigma_\eta^2 I$. Случай $R = R_0$ можно рассматривать как гипотезу о структуре случайных ошибок наблюдения. В работе показано, что расчет на эту неверную гипотезу может привести к недопустимой переоценке качества используемых решений.

В данном примере рассматриваются следующие оценки вектора θ: $\hat{\theta}$ — минимаксная оценка для множества $\mathcal{R}$; $\widetilde{\theta} = F_{\widetilde{R}} Y$ — оптимальная оценка (она построена в условиях полной априорной информации $R = \widetilde{R}$); $\widetilde{\theta}_0 = F_{R_0} Y$ — оценка, называемая эвристической, поскольку она найдена при гипотетической ковариации R_0.

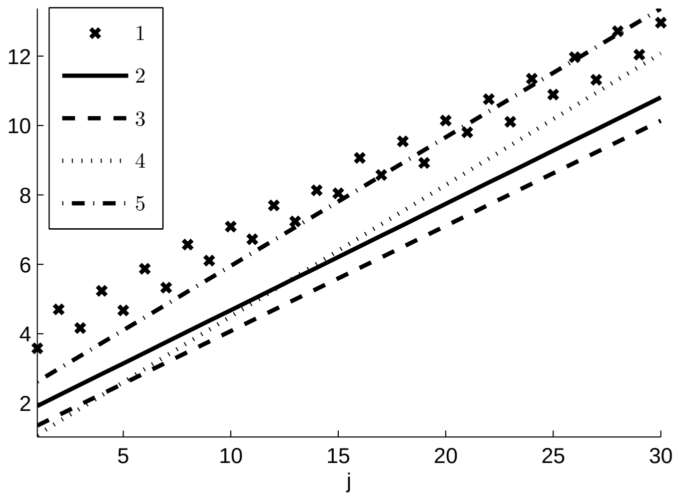

Рис. 2.1. Линии регрессии и наблюдения

На рис. 2.1 изображены результаты измерений Y_j и построены линии регрессии для параметров θ_i и их оценок (1 — наблюдения; 2 — полезный сигнал; 3 — оптимальная оценка; 4 — минимаксная оценка; 5 — эвристическая оценка).

Как видно, минимаксная оценка уступает оптимальной, но превосходит в точности эвристическую. В этом состоит одновременное проявление оптимальных и гарантирующих свойств минимаксных алгоритмов оценивания.

В третьей главе рассмотрена задача минимаксной фильтрации в стохастической дифференциальной системе с нестационарными возмущениями неизвестной интенсивности.

Рассматривается следующая стохастическая дифференциальная система:

$$\begin{cases} \xi(t) = \xi(0) + \int\limits_0^t a(s)\xi(s)\,ds + \int\limits_0^t b(s)\,d\mu(s), \\ \eta(t) = \int\limits_0^t A(s)\xi(s)\,ds + \int\limits_0^t B(s)\,d\nu(s), \end{cases} \tag{3.19}$$

в которой $\{\xi(t),\ t \in [0,T]\}$ — m-мерный случайный процесс, описывающий состояние системы, а $\{\eta(t),\ t \in [0,T]\}$ — n-мерный случайный процесс, доступный непосредственному наблюдению. Начальное состояние $\xi(0)$ предполагается центрированным случайным вектором с матрицей ковариации $D = \mathsf{cov}\{\xi(0), \xi(0)\}$, принадлежащей заданному множеству неотрицательно определенных симметричных матриц $\mathcal{D}$. Возмущения $\mu(t) \in \mathbb{R}^p$ и $\nu(t) \in \mathbb{R}^q$ образуют процесс $w(t) = \mathrm{col}[\mu(t), \nu(t)]$, который является центрированным, некоррелированным с вектором $\xi(0)$ и имеющим ортогональные приращения. Последнее предположение означает, что ковариационная функция процесса $w(t)$ имеет вид:

$$\mathsf{cov}\{w(t), w(s)\} = \int\limits_0^{\min(t,s)} U(\tau)\, d\tau, \quad t, s \in [0, T].$$

Функция интенсивности $U(\tau)$ принадлежит следующему классу:

$$\mathcal{U} = \{U\colon [0,T] \to \mathbb{R}^{u\times u}\colon U(\tau) \text{ измерима и } U(\tau) \in \mathcal{V}, \tau \in [0,T]\}, \quad (3.20)$$

где $\mathcal{V}$ — некоторое фиксированное множество неотрицательно определенных симметричных матриц размера $u \times u$, $u = p + q$. Для элементов множества $\mathcal{V}$ используется запись в виде блочной матрицы: $V = \begin{pmatrix} V_\mu & V_{\mu\nu} \\ V_{\nu\mu} & V_\nu \end{pmatrix}$.

Коэффициенты системы (3.19) $a(s) \in \mathbb{R}^{m\times m}$, $A(s) \in \mathbb{R}^{n\times m}$, $b(s) \in \mathbb{R}^{m\times p}$, $B(s) \in \mathbb{R}^{n\times q}$ предполагаются известными измеримыми ограниченными функциями.

Следующие допущения играют ключевую роль при построении минимаксного фильтра:

А) множества $\mathcal{D}$ и $\mathcal{V}$ выпуклы и компактны;

Б) выполнено условие регулярности:

$$\exists\, \varepsilon > 0\colon \quad B(s)V_\nu B^*(s) \geqslant \varepsilon I, \quad s \in [0,T] \quad \forall\, V_\nu \in \mathcal{V} \quad (3.21)$$

Отметим, что (3.21) означает равномерную невырожденность матрицы интенсивности шумов в уравнениях процесса наблюдения.

Случайный вектор $\widetilde{\xi}(t)$ называется допустимой оценкой вектора состояния $\xi(t)$ по наблюдениям процесса $\{\eta(t),\ t \in [0,T]\}$, если существует измеримая функция $g\colon [0,T] \times [0,T] \to \mathbb{R}^{m\times n}$, такая что $\int\limits_0^T \int\limits_0^T \|g(t,s)\|^2\, dtds < \infty$ и

$$\widetilde{\xi}(t) = \int\limits_0^t g(t,s)\, d\eta(s) \quad \forall\, t \in [0,T]. \quad (3.22)$$

Таким образом, рассматриваются только линейные неупреждающие преобразования наблюдаемого процесса.

Пусть далее Ξ — класс всех допустимых оценок (фильтров), т.е. случайных процессов $\{\widetilde{\xi}(t),\, t \in [0,T]\}$ вида (3.22).

Точность фильтра $\widetilde{\xi} \in \Xi$ будем измерять с помощью среднеквадратического критерия

$$L_t(\widetilde{\xi} \mid D, U) = \mathsf{E}\Big\{ \|\xi(t) - \widetilde{\xi}(t)\|^2 \mid D, U \Big\}, \ t \in [0,T], \quad (3.23)$$

где $\mathsf{E}\{\cdot \mid D, U\}$ обозначает математическое ожидание, вычисленное в предположении, что процессы $\xi(t)$ и $\eta(t)$ удовлетворяют уравнениям (3.19) с ковариационной матрицей D начального состояния $\xi(0)$ и функцией интенсивности возмущений U.

Наряду с локальным критерием (3.23) в работе рассматривается также интегральный критерий

$$I_T(\widetilde{\xi} \mid D, U) = \int\limits_0^T L_t(\widetilde{\xi} \mid D, U)\, dt.$$

Определение 3.2. *Пусть $J_T(\cdot)$ обозначает один из критериев $L_T(\cdot)$ или $I_T(\cdot)$. Допустимый фильтр $\hat{\xi}$ называется минимаксным относительно $J_T(\cdot)$, если*

$$\hat{\xi} \in \arg\min_{\widetilde{\xi} \in \Xi} \sup_{D \in \mathcal{D}, U \in \mathcal{U}} J_T(\widetilde{\xi} \mid D, U).$$

При этом

$$\hat{J}_T = \inf_{\widetilde{\xi} \in \Xi} \sup_{D \in \mathcal{D}, U \in \mathcal{U}} J_T(\widetilde{\xi} \mid D, U)$$

называется оптимальным гарантированным значением критерия $J_T(\cdot)$.

В данной главе для построения минимаксного фильтра используется метод двойственной оптимизации. Суть его заключается в том, что минимаксный фильтр ищется в виде решения задачи оптимальной фильтрации:

$$\hat{\xi} \in \arg\min_{\widetilde{\xi} \in \Xi} J_T(\widetilde{\xi} \mid \hat{D}, \hat{U}),$$

где $\hat{D}$ и $\hat{U}$ являются наименее благоприятными характеристиками модели наблюдения. Последнее означает, что $\hat{D}$, $\hat{U}$ образуют решение следующей задачи максимизации:

$$(\hat{D}, \hat{U}) \in \arg\max_{D \in \mathcal{D}, U \in \mathcal{U}} \underline{J}_T(D, U), \tag{3.24}$$

где

$$\underline{J}_T(D, U) = \inf_{\widetilde{\xi} \in \Xi} J_T(\widetilde{\xi} \mid D, U). \tag{3.25}$$

Задача (3.24) и функционал (3.25) называются двойственными, если выполнено следующее равенство:

$$\inf_{\widetilde{\xi} \in \Xi} \sup_{D \in \mathcal{D}, U \in \mathcal{U}} J_T(\widetilde{\xi} \mid D, U) = \sup_{D \in \mathcal{D}, U \in \mathcal{U}} \underline{J}_T(D, U). \tag{3.26}$$

Точная формулировка описанного метода приведена в следующей теореме.

Теорема 3.4. Пусть $J_t(\cdot)$ обозначает один из критериев $L_t(\cdot)$ или $I_t(\cdot)$. Если условия А) *и* Б) *выполнены, то справедливы следующие утверждения.*

1) *Выполнено соотношение двойственности* (3.26).

2) *Существует решение двойственной задачи* (3.24), *где*

$$\underline{J}_t(D,U) = \begin{cases} \mathrm{tr}[\gamma(t\,|\,D,U)] & \text{при} \quad J_t(\cdot) = L_t(\cdot), \\ \int\limits_0^t \mathrm{tr}[\gamma(s\,|\,D,U)]\,ds & \text{при} \quad J_t(\cdot) = I_t(\cdot), \end{cases} \tag{3.27}$$

а $\gamma(s\,|\,D,U)$ удовлетворяет уравнению Риккати:

$$\frac{d\gamma(s\,|\,D,U)}{ds} = \Gamma(s,\gamma(s\,|\,D,U),U(s)), \qquad \gamma(0\,|\,D,U) = D, \tag{3.28}$$

где $\Gamma(s,\gamma,V) = a(s)\gamma + \gamma a^*(s) + b(s)V_\mu b^*(s) - (b(s)V_{\mu\nu}B^*(s) + \gamma A^*(s))\times$ $\times(B(s)V_\nu B^*(s))^{-1}(B(s)V_{\nu\mu}b^*(s) + A(s)\gamma)$.

При этом $\underline{J}_T(\hat{D},\hat{U})$ — оптимальное гарантированное значение критерия.

3) *Минимаксный фильтр совпадает с фильтром Калмана:*

$$\hat{\xi}(t) = \int\limits_0^t \bigl\{a(s)\hat{\xi}(s)\,ds + [b(s)\hat{U}_{\mu\nu}(s)B^*(s) + \hat{\gamma}(s)A^*(s)]\times$$
$$\times\,(B(s)\hat{U}_\nu(s)B^*(s))^{-1}[d\eta(s) - A(s)\hat{\xi}(s)\,ds]\bigr\}, \quad \text{где } \hat{\gamma}(s) = \gamma(s\,|\,\hat{D},\hat{U}).$$

4) *Минимаксный фильтр $\hat{\xi}$ и наименее благоприятные характеристики $(\hat{D},\hat{U})$ образуют седловую точку функционала $J_T(\cdot)$ на произведении Ξ и $\mathcal{D}\times\mathcal{U}$:*

$$J_T(\hat{\xi}\,|\,D,U) \leqslant J_T(\hat{\xi}\,|\,\hat{D},\hat{U}) \leqslant J_T(\widetilde{\xi}\,|\,\hat{D},\hat{U}) \quad \forall\,\widetilde{\xi}\in\Xi \quad \text{и} \quad \forall\,(D,U)\in\mathcal{D}\times\mathcal{U}.$$

В силу теоремы 3.4 двойственную задачу (3.24) можно рассматривать как задачу оптимального управления системой (3.28) по терминальному или интегральному критерию (3.27) с состоянием γ и управлением U, подчиняющимся ограничениям (3.20).

Приведенное ниже утверждение описывает результат применения принципа максимума к вариационной задаче (3.24).

Для краткости изложения нам потребуются следующие обозначение:

$$\Psi(s,\gamma,V,\psi)=\psi\,\{a(s)-K(s,\gamma,V)A(s)\}\{a(s)-K(s,\gamma,V)A(s)\}^*\psi,$$

где $K(s,\gamma,V)=(b(s)V_{\mu\nu}B^*(s)+\gamma A^*(s))(B(s)V_\nu B^*(s))^{-1}$.

Т е о р е м а 3.5. *Пусть выполнены условия* А) *и* Б), *а множество* $\mathcal{D}$ *содержит максимальный элемент* $\bar{D}$, *т.е.* $D\leqslant\bar{D}\ \forall D\in\mathcal{D}$.

1) *При* $J_t(\cdot)=L_t(\cdot)$ *пара* $(\bar{D},\hat{U})$ *является решением двойственной задачи* (3.24) *тогда и только тогда, когда*

$$\hat{U}(s)\in\arg\max_{V\in\mathcal{V}}\operatorname{tr}[\psi(s)\Gamma(s,\hat{\gamma}(s),V)]\,,$$

$$\dot{\hat{\gamma}}(s)=\Gamma(s,\hat{\gamma}(s),\hat{U}(s)),\quad \hat{\gamma}(0)=\bar{D},$$

$$\dot{\psi}(s)=-\Psi(s,\hat{\gamma}(s),\hat{U}(s),\psi(s)),\quad s\in[0,t],\quad \psi(t)=I.$$

2) *В случае* $J_T(\cdot)=I_T(\cdot)$ *пара* $(\bar{D},\hat{U})$ — *решение двойственной задачи* (3.24) *тогда и только тогда, когда соотношения* (1), (1) *и*

$$\dot{\psi}(s)=-\Psi(s,\hat{\gamma}(s),\hat{U}(s),\psi(s))-I,\quad s\in[0,T],\quad \psi(T)=O.$$

Решение двойственной задачи может быть существенно упрощено, если в системе (3.19) процесс состояния принимает скалярные значения.

С л е д с т в и е 3.1. *Если* $m=1$, *т.е.* $\xi(t)\in\mathbb{R}$, *то в условиях теоремы 3.5 решение двойственной задачи* (3.24) *определяется следующим образом*:

$$\hat{U}(s)\in\arg\max_{V\in\mathcal{V}}\Gamma(s,\hat{\gamma}(s),V),\quad \dot{\hat{\gamma}}(s)=\Gamma(s,\hat{\gamma}(s),\hat{U}(s)),\quad \hat{\gamma}(0)=\bar{D}.$$

В конце данной главы представлен пример 3.1, в котором для частного случая модели наблюдения (3.19) удалось найти аналитическое представление для минимаксного фильтра. В этом примере

$$\begin{gathered} m=n=p=q=1,\qquad \mathcal{D}=[0,\bar{D}],\\ \mathcal{V}=\{V\geqslant O\colon\ V_\mu\leqslant 1,\,V_\nu\leqslant 1,\,|V_{\mu\nu}|\leqslant\bar{r}\}, \end{gathered}\tag{3.29}$$

где $\bar{D}\geqslant 0$, $0\leqslant\bar{r}\leqslant 1$ — заданные параметры.

Таким образом, в рассматриваемой системе оцениваемый, наблюдаемый и возмущающие процессы являются скалярными, а дисперсия начального состояния и функция взаимной интенсивности шумов неизвестны, но ограничены.

Относительно коэффициента диффузии в уравнении процесса наблюдения предпологается $\exists\,\varepsilon > 0\colon \quad B^2(s) \geqslant \varepsilon \quad \forall\, s \in [0, T]$.

В работе для модели (3.29) получены следующие уравнения минимаксного фильтра:

$$d\hat{\xi}(t) = a(t)\hat{\xi}(t)\,dt + \frac{\operatorname{sign}(A(t))(\hat{\gamma}(t)|A(t)| - \bar{r}|b(t)B(t)|)_+}{B^2(t)} \times$$
$$\times\, [d\eta(t) - A(t)\hat{\xi}(t)\,dt], \quad \hat{\xi}(0) = 0,$$

где наихудшая дисперсия ошибки оценивания

$$\dot{\hat{\gamma}}(t) = 2a(t)\hat{\gamma}(t) + b^2(t) - \Big(\frac{(\hat{\gamma}(t)|A(t)| - \bar{r}|b(t)B(t)|)_+}{B(t)}\Big)^2, \quad \hat{\gamma}(0) = \overline{D},$$

где $r_+ = \max(r, 0)$.

Кроме того, для данного примера представлены результаты численного моделирования и сравнительный анализ различных способов фильтрации.

ПОЛОЖЕНИЯ ДИССЕРТАЦИИ, ВЫНОСИМЫЕ НА ЗАЩИТУ

1. Аналитический синтез минимаксной оценки случайного вектора при наличии произвольно коррелированных ошибок наблюдения с интегрально ограниченной дисперсией (теоремы 1.1, 1.2).

2. Разработка и реализация алгоритма оптимального оценивания в стохастической многомерной модели регрессии при поэлементных ограничениях на ковариационную матрицу вектора ошибок (алгоритм 2.1, теорема 2.3).

3. Обоснование алгоритма минимаксной фильтрации в стохастической дифференциальной системе с нестационарными возмущениями неизвестной интенсивности (теоремы 3.4, 3.5).

4. Решение задачи минимаксной фильтрации скалярного состояния в линейной стохастической дифференциальной системе (следствие 3.1, пример 3.1).

СПИСОК РАБОТ, ОПУБЛИКОВАННЫХ ПО ТЕМЕ ДИССЕРТАЦИИ

1. Лебедев М.В., Семенихин К.В. Минимаксная оценка случайного вектора при наличии произвольно коррелированных помех // Вестник МАИ, 2008, Т.15, №2, С.90–104.

2. Лебедев М.В., Семенихин К.В. Минимаксная фильтрация в стохастической дифференциальной системе с нестационарными возмущениями неизвестной интенсивности // Известия РАН. Теория и системы управления, 2007, №2, С.45–56.

3. Лебедев М.В., Семенихин К.В. Минимаксное оценивание в линейных неопределенно-стохастических динамических системах с непрерывным временем // В кн. Проектирование, конструирование и производство авиационной техники. М: МАИ, 2005, С.103–108.

4. Siemenikhin K.V., Lebedev M.V., Platonov E.N. Kalman filtering by minimax criterion with uncertain noise intensity functions // Proceedings of the 44th IEEE Conference on Decision and Control (CDC-ECC'2005). Seville (Spain): 2005, Pp.1929–1934.

5. Siemenikhin K.V., Lebedev M.V. Minimax estimation of random elements: theory and applications // Proceedings of the 43th IEEE Conference on Decision and Control (CDC'2004). Nassau (Bahamas): 2004, Pp.3581–3586.

6. Лебедев М.В. Минимаксное оценивание в сингулярных регрессионных моделях наблюдения // Тезисы докладов конференции «Авиация и космонавтика - 2008», МАИ, 20–23 октября 2008, С.92–93.

7. Лебедев М.В., Семенихин К.В. Минимаксная фильтрация в стохастической модели с неизвестными функциями интенсивности возмущений // Тезисы докладов конференции «Системный анализ, управление и навигация», Крым, Евпатория, 29 июня–6 июля 2008, С.255.

8. Лебедев М.В. Оптимальное оценивание параметров стохастических систем в присутствии произвольно коррелированных возмущений // Тезисы докладов конференции «Информационные технологии в авиационной и космической технике - 2008», МАИ, 21–24 апреля 2008, С.115.

9. Лебедев М.В., Семенихин К.В. Восстановление параметров и состояний стохастических систем в присутствии произвольно коррелированных возмущений // Тезисы докладов конференции «Научная сессия МИФИ - 2008», МИФИ, 21–27 января 2008, С.88–89.

Лебедев Максим Витальевич

АЛГОРИТМЫ ОПТИМАЛЬНОГО ОЦЕНИВАНИЯ
В СТОХАСТИЧЕСКИХ СИСТЕМАХ В УСЛОВИЯХ
АПРИОРНОЙ НЕОПРЕДЕЛЕННОСТИ

05.13.01 — Системный анализ, управление и обработка информации
(авиационная и ракетно-космическая техника)

Тираж — 100 экз.