

(ОЗНАКОМИТЕЛЬНЫЙ ФРАГМЕНТ)

БЕЛОРУССКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ

УДК 519.2

ВОЛОШКО

Валерий Анатольевич

РОБАСТНОЕ СТАТИСТИЧЕСКОЕ ОЦЕНИВАНИЕ ПАРАМЕТРОВ И ПРОГНОЗИРОВАНИЕ ВРЕМЕННЫХ РЯДОВ НА ОСНОВЕ МАЛОПАРАМЕТРИЧЕСКИХ МОДЕЛЕЙ

АВТОРЕФЕРАТ

диссертации на соискание ученой степени
кандидата физико-математических наук

по специальности 01.01.05 — теория вероятностей
и математическая статистика

(ОЗНАКОМИТЕЛЬНЫЙ ФРАГМЕНТ)

Работа выполнена в Белорусском государственном университете

- Научный руководитель -- Харин Юрий Семенович,
доктор физико-математических наук,
профессор, член-корреспондент Национальной академии наук Беларуси, заведующий кафедрой математического моделирования и анализа данных Белорусского государственного университета.
- Официальные оппоненты: Леваков Анатолий Афанасьевич,
доктор физико-математических наук,
профессор, профессор кафедры высшей математики Белорусского государственного университета;
- Марковская Наталья Вацлавовна,
кандидат физико-математических наук,
доцент, доцент кафедры математического и информационного обеспечения экономических систем УО "Гродненский государственный университет имени Янки Купалы".
- Опонирующая организация — ГНУ "Институт математики Национальной академии наук Беларуси".

Защита состоится 31 января 2014 года в 10.00 на заседании совета по защите диссертаций Д 02.01.07 при Белорусском государственном университете по адресу: 220030, Республика Беларусь, г. Минск, ул. Ленинградская, 8 (юридический факультет), ауд. 407, тел. ученого секретаря: (017) 209-57-09.

С диссертацией можно ознакомиться в Фундаментальной библиотеке Белорусского государственного университета.

Автореферат разослан "12" декабря 2013 года.

Ученый секретарь совета
по защите диссертаций Д 02.01.07,
доктор физ.-мат. наук, профессор



Н. В. Лазакович

(ОЗНАКОМИТЕЛЬНЫЙ ФРАГМЕНТ) ФРАГМЕНТ В ВЕДЕНИИ

В статистическом анализе временных рядов (ВР) можно выделить две основные задачи: прогнозирование¹ и восстановление информации о вероятностном распределении, (вероятностной модели)²; вторая задача включает в себя подзадачи: оценивание параметров модели и проверка гипотез. При решении этих задач в каждом конкретном случае, как правило, выбираются два определяющих математических объекта: целевой функционал (например, среднеквадратическая ошибка оценивания, вероятность ошибочного решения, риск прогнозирования) и априорные модельные предположения, после чего целевой функционал минимизируется (или максимизируется) на пространстве выбранных гипотетических модельных предположений. Однако зачастую построенный таким образом оптимальный статистический вывод, примененный на реальных данных, дает неадекватный результат. Это связано с тем, что для постулируемых гипотетических модельных предположений, при незначительном их нарушении значение целевого функционала может "резко ухудшиться"³. Поэтому, кроме оптимальности в смысле достижения минимума целевого функционала, появляется такой критерий качества статистического вывода, как робастность (robustness), введенный в математическую статистику швейцарским математиком П. Хьюбером и означающий устойчивость значения целевого функционала к малым отклонениям от модельных предположений⁴.

При решении математической задачи прогнозирования, впервые рассмотренной А. Н. Колмогоровым, в качестве целевого функционала обычно используют среднеквадратический риск: в этом случае оптимальной прогнозирующей статистикой является условное математическое ожидание прогнозируемой величины в будущий момент времени при фиксированных наблюдениях в прошлые и настоящий моменты времени. Кроме того, для гауссовских стационарных ВР оптимальная в среднеквадратичном прогнозирующая статистика линейна по статистическим данным и для ее построения достаточно знать функцию спектральной плотности (спектр) ВР. Поэтому в таких условиях задача робастного прогнозирования сопряжена с задачей робастной идентификации, состоящей, согласно Боксу и Дженкинсу⁵, из двух этапов: 1) выбор модели M (системы координат $\varphi : \text{Spec}_M \rightarrow \mathbb{R}^\infty$

¹ Колмогоров, А. Н. Интерполяция и экстраполяция стационарных случайных последовательностей / А. Н. Колмогоров // Известия АН СССР. Сер. матем. — 1941. — № 5. — С. 3–14.

² Dzhaparidze, K. Parameter estimation and hypothesis testing in spectral analysis of stationary time series / K. Dzhaparidze. — Springer-Verlag, 1985. — 330 p.

³ Робастность в статистике. Подход на основе функций влияния / Ф. Хампель [и др.]. — М.: Мир, 1989. — 512 с.

⁴ Хьюбер, П. Робастность в статистике / П. Хьюбер. — М.: Мир, 1984. — 304 с.

⁵ Box, G. Time series analysis: Forecasting and control / G. Box, G. Jenkins. — San Francisco: Holden-Day, 1976. — 575 p.

(ОЗНАКОМИТЕЛЬНЫЙ ФРАГМЕНТ)

на подмножестве $\text{Spec}_M \subset \text{Spec}$ множества всех допустимых спектров Spec);
2) робастное статистическое оценивание модели M по данным (fitting to data⁶).
На втором этапе выбирается *параметр сглаживания* $\eta \in \mathbb{N}$ и последовательность $\varphi = (\varphi_j)_{j \in \mathbb{N}}$ параметров модели M разбивается на две части: $\varphi = (\varphi_0, \varphi_+)$, $\varphi_0 = (\varphi_1, \dots, \varphi_\eta) \in \mathbb{R}^\eta$, $\varphi_+ = (\varphi_j)_{j > \eta} \in \mathbb{R}^\infty$. Затем по наблюдениям строится робастная оценка $\hat{\varphi}_0$ вектора φ_0 первых η параметров модели M , по которой вычисляется оценка $\hat{\varphi}_+$ последовательности оставшихся параметров φ_+ : $\hat{\varphi}_+ = \Phi_+(\hat{\varphi}_0)$. Бокс и Дженкинс разработали детальную стратегию построения моделей, в которой подчеркивается важность так называемой *малопараметричности* последних⁷. Важность этого свойства для задач статистического прогнозирования может быть обоснована следующим рассуждением: допустим вместо истинного спектра $S \in \text{Spec}$ используется его оценка $\hat{S} = \varphi^{-1}(\hat{\varphi}) \in \text{Spec}_M$, и подставляется в оптимальную прогнозирующую статистику (эту статистику принято называть *подстановочной*). Риск такой прогнозирующей статистики зачастую превышает минимальный возможный риск на величину порядка $\mathcal{O}(\eta/T)$ при $\eta > \eta_-(T)$, где η — описанный выше параметр сглаживания, $T \in \mathbb{N}$ — *число наблюдений*, использованных для оценивания параметров φ . Функция $\eta_-(\cdot)$ зависит от вероятностного распределения прогнозируемого ВР и выбранной процедуры прогнозирования: она определяет число параметров, при котором *ошибка аппроксимации* спектра данной моделью становится сравнима со *статистической погрешностью оценивания* модели. Отсюда появляется необходимость в малопараметрических моделях M . Мы определяем малопараметрические модели как модели с “медленно растущей” границей η_- на некотором *классе* спектров $\text{Spec}_M^* \subset \text{Spec}$, к которому, по априорному предположению, принадлежит спектр прогнозируемого ВР. При этом модели M и классы спектров Spec_M^* , на которых эти модели малопараметричны, взаимно определяют друг друга.

Указанные актуальные проблемы — построения малопараметрических моделей стационарных временных рядов и робастных оценивающих и прогнозирующих статистик на их основе — определяют направление исследований настоящей диссертационной работы.

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Связь работы с крупными научными программами (проектами) и темами

Тема диссертации соответствует направлениям научных исследований “Методы математического и компьютерного моделирования, компьютерные тех-

⁶Amari, S. Methods of Information Geometry / S. Amari, H. Nagaoka. — Oxford University Press, 2000. — 206 p.

⁷Ledolter, J. Parsimony and its importance in time series forecasting / J. Ledolter, B. Abraham // Technometrics. — 1981. — Vol. 23, № 4. — P. 411–414.

(ОЗНАКОМИТЕЛЬНЫЙ ФРАГМЕНТ)

нологии и интеллектуальные системы поддержки принятия решений” и “Математические и интеллектуальные методы, информационные технологии и системы распознавания и обработки образов, сигналов, речи и мультимедийной информации”, определенным Перечнем приоритетных направлений научных исследований Республики Беларусь на 2011–2015 годы. Результаты диссертационного исследования использованы при выполнении в Белорусском государственном университете следующих научно-исследовательских работ:

1) НИР “Статистический анализ данных сложной структуры и его применения в компьютерных системах моделирования и обработки информации” (2006–2010 гг.), номер госрегистрации 20062641;

2) Государственная программа фундаментальных исследований “Математические модели”, НИР “Разработка вероятностных моделей, методов и алгоритмов статистического анализа временных рядов с конечным пространством состояний и “длинной памятью” (2006–2010 гг.), номер госрегистрации 20062202;

3) Грант Министерства Образования Республики Беларусь, НИР “Робастный статистический анализ временных рядов на основе малопараметрических моделей” (2011 г.), номер госрегистрации 20112193;

4) Грант БГУ, НИР “Разработка методов робастного статистического анализа случайных последовательностей и полей при наличии неоднородностей и цензурирования” (2011 г.), номер госрегистрации 20111047;

5) НИР “Робастные статистические выводы и их применение в компьютерных системах моделирования и анализа данных” (2011–2015 гг.), номер госрегистрации 20120355;

6) Государственная комплексная программа научных исследований “Инфотех”, НИР “Разработка методов, алгоритмов и программных средств статистического прогнозирования процессов в сложных стохастических системах” (2006–2010 гг.), номер госрегистрации 20062205.

Цель и задачи исследования

Целью диссертационной работы является построение и асимптотический анализ робастных оценок параметров и прогнозирующих статистик для стационарных временных рядов на основе малопараметрических моделей спектра (спектральной плотности). Для достижения поставленной цели требуется решить следующие основные задачи:

1) установление свойств малопараметрических моделей спектра и соответствующих им классов спектров;

2) построение и анализ робастных оценивающих и прогнозирующих статистик на основе малопараметрической модели Блумфилда спектральной плотности;

(ОЗНАКОМИТЕЛЬНЫЙ ФРАГМЕНТ)

3) Построены и статистически оценены мощностные и прогнозирующие статистики на основе малопараметрической ковариационной модели спектральной плотности.

Объектом исследования являются стационарные временные ряды с искажениями, порожденными выбросами, неоднородностями, пропусками. Предмет исследования — малопараметрические модели спектра, робастные оценивающие и прогнозирующие статистики на их основе, вероятностные характеристики прогнозирующих статистик.

Положения, выносимые на защиту

1. Построенный класс спектров на основе малопараметрической модели Блумфилда и интегральная формула для функции Сеге.

2. Построенные асимптотические разложения смешанных моментов логарифма периодограммы и подстановочных оценок спектральных коэффициентов в асимптотике растущей длительности наблюдения временного ряда.

3. Доказанные асимптотические свойства и оценка робастности риска подстановочной прогнозирующей статистики при аппроксимации спектра малопараметрической моделью Блумфилда.

4. Построенное семейство робастных пси-оценок коэффициентов корреляции гауссовского стационарного временного ряда при наличии “выбросов” и “пропусков”, доказанные свойства состоятельности, асимптотической нормальности и робастности найденной оптимальной пси-оценки.

Личный вклад соискателя

Основные результаты, приведенные в диссертации, получены автором самостоятельно. Научному руководителю в совместных работах принадлежат выбор направлений исследования, предметные постановки задач, обсуждение результатов.

Апробация результатов диссертации

Основные результаты диссертации были представлены и обсуждались на заседаниях Республиканского научного семинара кафедры математического моделирования и анализа данных БГУ и Научно-исследовательского института прикладных проблем математики и информатики “Математическое моделирование сложных систем, анализ данных и защита информации”, научном семинаре Института математики и информатики Вильнюсского университета (29 апреля 2013 года), научных конференциях студентов и аспирантов Белгосуниверситета, 11-й Республиканской научной конференции студентов и аспирантов “Новые математические методы и компьютерные технологии в проектировании, производстве и научных исследованиях” (17–19 марта 2008 года, Гомель), Международной научной конференции “Теория вероятностей,

(ОЗНАКОМИТЕЛЬНЫЙ ФРАГМЕНТ)

случайные процессы, математическая статистика и приложения” (15–19 сентября 2008 года, Минск), 2-й Международной научно-практической конференции “Современные информационные компьютерные технологии” (26–28 апреля 2010 года, Гродно), 9-й Международной конференции “Computer Data Analysis and Modeling” (7–11 сентября 2010 года, Минск), 6-й Международной научной конференции “Информационные системы и технологии” (24–25 ноября 2010 года, Минск), Республиканской научной конференции студентов и аспирантов высших учебных заведений Республики Беларусь (18 октября 2011 года, Минск), Международном конгрессе по информатике “Информационные системы и технологии” (31 октября–3 ноября 2011 года, Минск), Международной конференции “Modeling and Simulation” (2–4 мая 2012 года, Минск), 11-й Белорусской математической конференции (5–9 ноября 2012 года, Минск), 10-й Международной конференции “Computer Data Analysis and Modeling” (10–14 сентября 2013 года, Минск).

Опубликованность результатов диссертации

Результаты диссертационной работы опубликованы в 17 научных работах. Из них 6 статей в научных журналах в соответствии с пунктом 18 Положения о присуждении ученых степеней и присвоении ученых званий в Республике Беларусь (общим объемом 7 авторских листов), 1 статья в сборнике научных трудов, 5 статей в сборниках материалов научных конференций, 5 тезисов докладов. Результаты диссертационной работы также вошли в состав четырех заключительных отчетов о НИР.

Структура и объем диссертации

Диссертационная работа состоит из перечня условных обозначений, введения, общей характеристики работы, трех глав, заключения, библиографического списка. Полный объем диссертации составляет 99 страниц, в том числе 11 рисунков на 9 страницах, библиографический список из 99 наименований, включая 17 публикаций соискателя, на 8 страницах.

ОСНОВНОЕ СОДЕРЖАНИЕ

В **главе 1** представлен обзор известных параметрических моделей спектров стационарных в широком смысле ВР, введено понятие робастности и найден класс спектров, для которых модель Блумфилда (Блумфилд, 1973) малопараметрична.

В **разделе 1.1** проведен аналитический обзор основных результатов по робастному оцениванию параметров и прогнозированию стационарных ВР. **Раздел 1.2** содержит описание основных параметризаций спектра и их свойств. *Временной ряд* — последовательность $\{x_t = x_t(\omega) : \omega \in \Omega, t \in \mathbb{Z}\}$

позволившие построить робастную прогнозирующую статистику, использованную в ППП "РОСТАПРО" [1, 3, 7, 13].

Рекомендации по практическому использованию результатов

Работа носит теоретический характер. Полученные результаты могут быть применены для развития теории робастного статистического оценивания параметров и прогнозирования, а также для разработки программных средств компьютерного анализа данных и прогнозирования в экономике, финансах, обработке сигналов. Результаты диссертации могут быть использованы в учебном процессе на спецкурсах по статистическому анализу временных рядов.

(ОЗНАКОМИТЕЛЬНЫЙ ФРАГМЕНТ)

ПО ТЕМЕ ДИССЕРТАЦИИ

Статьи в научных журналах

1. Харин, Ю. С. Об одном методе робастного оценивания параметров авторегрессии при наличии “выбросов” / Ю. С. Харин, В. А. Волошко // Весці НАН Беларусі. Серыя фізіка-матэматычных навук. — 2009. — № 1. — С. 4–12.
2. Харин, Ю. С. Прогнозирование стационарных временных рядов на основе малопараметрической модели Блумфилда / Ю. С. Харин, В. А. Волошко // Доклады НАН Беларуси. — 2010. — Т. 54, № 6. — С. 27–32.
3. Kharin, Yu. S. Robust estimation of AR coefficients under simultaneously influencing outliers and missing values / Yu. S. Kharin, V. A. Voloshko // Journal of Statistical Planning and Inference. — 2011. — Vol. 141, № 9. — P. 3276–3288.
4. Волошко, В. А. Об асимптотических свойствах оценок кепстральных коэффициентов для гауссовского временного ряда / В. А. Волошко, Ю. С. Харин // Доклады НАН Беларуси. — 2011. — Т. 55, № 6. — С. 23–28.
5. Kharin, Yu. S. On asymptotic properties of the plug-in cepstrum estimator for Gaussian time series / Yu. S. Kharin, V. A. Voloshko // Mathematical Methods of Statistics. — 2012. — Vol. 21, № 1. — P. 43–60.
6. Волошко, В. А. Риск прогнозирования на основе модели Блумфилда при наличии обучающей последовательности / В. А. Волошко // Вестник БГУ. Серия 1. — 2012. — № 3. — С. 72–78.

Статьи в сборниках научных трудов и материалов научных конференций

7. Харин, Ю. С. Робастное оценивание коэффициентов авторегрессии в условиях “выбросов” и “пропусков” / Ю. С. Харин, В. А. Волошко // Теория вероятностей, случайные процессы, математическая статистика и приложения: сборник научных статей Международной научной конференции, Минск, 15–19 сентября 2008 г. / БГУ; редкол.: Н. Н. Труш (отв. ред.) [и др.]. — Минск, 2008. — С. 379–387.
8. Voloshko, V. A. On inequalities of information geometry / V. A. Voloshko // Computer Data Analysis and Modeling: proceedings of 9th International conference, Minsk, 7–11 september 2010: in 2 vol. / Ministry of Education of Republic of Belarus; ed. board: Prof. Dr. S. Aivazian, Prof. Dr. P. Filzmoser, Prof. Dr. Yu. Kharin. — Minsk, 2010. — Vol. 2. — P. 79–82.
9. Voloshko, V. A. Statistical forecasting based on Bloomfield exponential model / V. A. Voloshko, Yu. S. Kharin // Computer Data Analysis and Modeling:

proceedings of 9th International conference, Minsk, 11-12 september 2010; in 2 vol. / Ministry of Education of Republic of Belarus; ed. board: Prof. Dr. S. Aivazian, Prof. Dr. P. Filzmoser, Prof. Dr. Yu. Kharin. — Minsk, 2010. — Vol. 1. — P. 268–271.

10. *Волошко, В. А.* О прогнозировании стационарных временных рядов на основе экспоненциальной модели Блумфилда с малым числом параметров / В. А. Волошко, Ю. С. Харин // Стохастическое и компьютерное моделирование систем и процессов: сборник научных статей / ГрГУ им. Я. Купалы; редкол.: Л. В. Рудикова (гл. ред.) [и др.]. — Гродно : ГрГУ, 2011. — С. 213–217.

11. *Волошко, В. А.* Асимптотические свойства кепстра гауссовской последовательности / В. А. Волошко // Информационные системы и технологии: материалы Международного научного конгресса, Минск, 31 октября–3 ноября 2011 г.: в 2 ч. / БГУ, ОИПИ НАН Беларуси, НТА “Инфопарк”; редкол.: С. В. Абламейко (отв. ред.) [и др.]. — Минск, 2011. — Ч. 1. — С. 49–54.

12. *Valoshka, V. A.* On comparison of classical and cepstrum-based forecasts / V. A. Valoshka, Yu. S. Kharin // Computer Data Analysis and Modeling: proceedings of 10th International conference, Minsk, 10–14 september 2013; in 2 vol. / Ministry of Education of Republic of Belarus; ed. board: Prof. Dr. S. Aivazian, Prof. Dr. P. Filzmoser, Prof. Dr. Yu. Kharin. — Minsk, 2013. — Vol. 1. — P. 221–224.

Тезисы докладов

13. *Волошко, В. А.* Робастное оценивание параметров модели авторегрессии $AR(p)$ при наличии пропусков и выбросов / В. А. Волошко // Новые математические методы и компьютерные технологии в проектировании, производстве и научных исследованиях: материалы XI Республиканской научной конференции студентов и аспирантов, Гомель, 17–19 марта 2008 г.: в 2 ч. / Министерство образования Республики Беларусь, ГГУ им. Ф. Скорины; редкол.: О. М. Демиденко (гл. ред.) [и др.]. — Гомель, 2008. — Ч. 1. — С. 154–155.

14. *Волошко, В. А.* Риск кепстр-прогноза стационарной последовательности / В. А. Волошко // Информационные системы и технологии: материалы VI Международной конференции, Минск, 24–25 ноября 2010 г. / БГУ, БГУИР, НТА “Инфопарк”; редкол.: А. Н. Курбацкий (отв. ред.) [и др.]. — Минск, 2010. — С. 144–145.

15. *Волошко, В. А.* Асимптотические свойства кепстральных коэффициентов для гауссовского случайного процесса / В. А. Волошко, Ю. С. Харин // Сборник тезисов докладов Республиканской научной конференции студентов и аспирантов Республики Беларусь, Минск, 18 октября 2011 г. / БГУ и др.; редкол.: С. В. Абламейко [и др.]. — Минск, 2011. — С. 79.

(ОЗНАКОМИТЕЛЬНЫЙ ФРАГМЕНТ)

Voloshko, V. A. Risk of forecasting with fitting Bloomfield model to independent training sample / V. Voloshko // Modeling and Simulation: proceedings of International Conference, Minsk, 2-4 may 2012. / AMSE, BSU, BSIAPR, BSUIR; ed. board: Prof. Dr. V. Krasnoproshin [and oth.]. — Minsk, 2012. — P. 173–174.

17. Волошко, В. А. Риск прогнозирования на основе идентификации модели Блумфилда по обучающим данным / В. А. Волошко // Материалы XI Белорусской Математической Конференции, Минск, 5–9 ноября 2012 г. / БГУ, ИМ НАН Беларуси; редкол.: С. Г. Красовский [и др.]. — Минск, 2012. — С. 52–53.

(ОЗНАКОМИТЕЛЬНЫЙ ФРАГМЕНТ)

РЭЗІУМЭ

Валашко Валерый Анатолевіч

Рабаснае статыстычнае ацэньванне параметраў і прагназаванне часовых шэрагаў на аснове малапараметрычных мадэляў

Ключавыя словы: гаусаўскі стацыянарны часовы шэраг, рабасная прагназуючая статыстыка, малапараметрычная мадэль, асімптатычны расклад, кспстр.

Мэтай дысертацыйнай працы з'яўляецца пабудаванне і даследаванне імавернасных уласцівасцей рабасных статыстычных ацэнак параметраў і прагназуючых статыстык на аснове малапараметрычных мадэляў спектральнай шчыльнасці (спектра) гаусаўскага стацыянарнага часовага шэрагу. Пры даследаванні выкарыстоўваліся метады матэматычнага аналізу, тэорыі імавернасцей, матэматычнай статыстыкі, кампутарнага мадэлявання.

У працы атрыманы наступныя новыя навуковыя рэзультаты. Дзеля малапараметрычнай мадэлі Блумфілда: пабудаваны асацыяваны з ёй клас спектраў і асімптатычныя расклады змешаных момантаў падстаноўных ацэнак кепстра, якія дазволілі абгрунтаваць рабаснасць гэтых ацэнак; даказаны асімптатычныя ўласцівасці рызыкі падстаноўнай прагназуючай статыстыкі і знойдзены клас спектраў, дзеля якіх значэнне рызыкі гэтай прагназуючай статыстыкі не перавышае значэнне рызыкі аўтарэгрэсійнай прагназуючай статыстыкі; пабудаваны асімптатычны расклад рызыкі пры статыстычным ацэньванні спектра па “навучальнаму” часоваму шэрагу. Дзеля малапараметрычнай каварыяцыйнай мадэлі пабудавана новае сямейства рабасных псі-ацэнак каэфіцыентаў карэляцыі пры сумесным уплыве скажэнняў двух тыпаў (“выкіды” і “пропускі”) і даказаны уласцівасці слушнасці, асімптатычнай нармальнасці і рабаснасці гэтых ацэнак, дазволіўшыя пабудоваць рабасную прагназуючую статыстыку.

Атрыманыя тэарэтычныя рэзультаты могуць быць ужыты пры распрацоўцы праграмных сродкаў кампутарнага аналізу дадзеных і прагназавання у эканоміцы, фінансах, апрацоўцы сігналаў, а таксама у вучэбным працэсе на спецкурсах па статыстычным аналізе часовых шэрагаў.

(ОЗНАКОМИТЕЛЬНЫЙ ФРАГМЕНТ)

РЕЗЮМЕ

Волошко Валерий Анатольевич

Робастное статистическое оценивание параметров и прогнозирование временных рядов на основе малопараметрических моделей

Ключевые слова: гауссовский стационарный временной ряд, робастная прогнозирующая статистика, малопараметрическая модель, асимптотическое разложение, кепстр.

Целью диссертационной работы является построение и исследование вероятностных свойств робастных статистических оценок параметров и прогнозирующих статистик на основе малопараметрических моделей спектральной плотности (спектра) гауссовского стационарного временного ряда. При исследовании использовались методы математического анализа, теории вероятностей, математической статистики, компьютерного моделирования.

В работе получены следующие новые научные результаты. Для малопараметрической модели Блумфилда: построен ассоциированный с ней класс спектров и асимптотические разложения смешанных моментов подстановочных оценок кепстра, позволившие обосновать робастность этих оценок; доказаны асимптотические свойства риска подстановочной прогнозирующей статистики и найден класс спектров, для которых значение риска этой прогнозирующей статистики не превышает значение риска авторегрессионной прогнозирующей статистики; построено асимптотическое разложение риска при статистическом оценивании спектра по “обучающему” временному ряду. Для малопараметрической ковариационной модели построено новое семейство робастных пси-оценок коэффициентов корреляции при совместном влиянии искажений двух типов (“выбросы” и “пропуски”) и доказаны свойства состоятельности, асимптотической нормальности и робастности этих оценок, позволившие построить робастную прогнозирующую статистику.

Полученные теоретические результаты могут быть применены при разработке программных средств компьютерного анализа данных и прогнозирования в экономике, финансах, обработке сигналов, а также в учебном процессе на спецкурсах по статистическому анализу временных рядов.

(ОЗНАКОМИТЕЛЬНЫЙ ФРАГМЕНТ)

SUMMARY

Valashko Valery Anatolyevich

Robust statistical parameter estimating and forecasting of time series
on the base of parsimonious models

Key words: Gaussian stationary time series, robust forecasting statistics, parsimonious model, asymptotic expansion, cepstrum.

The objective of dissertation is building and researching of probabilistic properties of robust statistical parameter estimators and forecasting statistics on the base of parsimonious models of spectral density (spectrum) of Gaussian stationary time series. *The techniques used include* mathematical analysis, probability theory, methods of mathematical statistics and computer simulation.

In this dissertation the following new scientific results are obtained. For the parsimonious Bloomfield model: the associated class of spectrums is built; the asymptotic expansions of mixed moments of plug-in cepstrum estimators, allowed to prove the robustness of these estimators, are built; the asymptotic properties of risk of plug-in forecasting statistics are proved and the class of spectrums providing the gain by risk of this forecasting statistics with respect to autoregressive forecasting statistics is found; the asymptotic expansion of the risk under statistical spectrum fitting to "training" time series is built. For the parsimonious covariance model the new family of robust psi-estimators of correlation coefficients is build under simultaneously influencing outliers and missing values. The properties of consistency, asymptotical normality and robustness of these estimators, allowed to build the robust forecasting statistics, are proved.

The obtained theoretical results can be applied in the development of computer software for data analysis in economics, finance, signal processing, as well as in teaching process on the special courses of time series statistical analysis.

