На правах рукописи

Марков Александр Сергеевич

ИДЕНТИФИКАЦИЯ СТОХАСТИЧЕСКИХ СИСТЕМ АВТОРЕГРЕССИОННОГО ТИПА С НЕЛИНЕЙНОСТЯМИ И БЕСКОНЕЧНОЙ ДИСПЕРСИЕЙ ШУМА

05.13.01 – Системный анализ, управление и обработка информации
(в отраслях информатики, вычислительной техники и автоматизации)

Автореферат
диссертации на соискание ученой степени
кандидата физико-математических наук

Томск – 2009

Работа выполнена на кафедре высшей математики и математического моделирования ГОУ ВПО «Томский государственный университет»

Научный руководитель: доктор физико-математических наук,
профессор
Конев Виктор Васильевич

Официальные оппоненты: доктор технических наук,
профессор
Рубан Анатолий Иванович

доктор физико-математических наук,
профессор
Кошкин Геннадий Михайлович

Ведущая организация: ИПУ РАН (г. Москва)

Защита диссертации состоится 17 сентября 2009 г. в 10:30 часов на заседании диссертационного совета Д 212.267.12 при ГОУ ВПО «Томский государственный университет» по адресу: 634050, г. Томск, пр. Ленина, 36.

С диссертацией можно ознакомиться в Научной библиотеке ГОУ ВПО «Томский государственный университет» по адресу: 634050, г. Томск, пр. Ленина, 34а.

Автореферат разослан: 10 июля 2009 г.

Ученый секретарь
диссертационного совета
д.т.н., профессор

В.И. Смагин

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Актуальность проблемы. В прикладных задачах при обработке эмпирических временных рядов центральной проблемой является выбор адекватной математической модели поведения наблюдаемой системы. Как правило, выбор модели поведения системы осуществляется на основе анализа данных по ее функционированию. В этом случае процесс выбора модели называют идентификацией. Выделяют этапы структурной и параметрической идентификации. При структурной идентификации определяется вид функциональных связей между наблюдениями за стохастической динамической системой с точностью до конечного числа неизвестных параметров. Задачей параметрической идентификации является восстановление неизвестных параметров системы по элементам выборки (наблюдениям за системой).

В задачах управления, фильтрации и прогнозирования широко используются линейные модели временных рядов. Применение линейных моделей особенно оправдано в условиях относительно малого объема наблюдений, а также на этапе предварительного исследования структуры изучаемого объекта. При решении различных прикладных задач обширное применение получила линейная модель авторегрессии вида:

$$X_n = \lambda_1 X_{n-1} + \lambda_2 X_{n-2} + \ldots + \lambda_p X_{n-p} + \varepsilon_n,$$

где $\{X_n\}_{n\geq 0}$ – наблюдения за стохастической системой, $\{\varepsilon_i\}_{i\geq 0}$ – ненаблюдаемая последовательность возмущений (шум), $\lambda_1,\ldots,\lambda_p$ – неизвестные параметры модели.

Первоначально исследование характеристик различных процедур оценивания неизвестных параметров $\lambda_1,\ldots,\lambda_p$ проводилось в предположении, что шумы $\{\varepsilon_i\}_{i\geq 0}$ – независимые одинаково распределенные случайные величины с «нормальным» (гауссовым) распределением, а параметры с течением времени не изменяются. Затем был рассмотрен случай, когда распределение шумов неизвестно, а предполагается лишь конечность второго момента. Однако на практике эти ограничения не всегда позволяют отследить динамику реального процесса.

Многочисленные экспериментальные исследования современных систем связи показали, что такие характеристики, как размер файла, время, требуемое процессору для выполнения работ, время соединения, время ожидания между пакетами в сети Ethernet и их размер, данные видео конференций проявляют свойства случайных величин с «тяжелыми» хвостами распределений (бесконечной дисперсией). Это вызвало дополнительный интерес к изучению свойств стохастических динамических систем с шумами, распределение которых имеет бесконечную дисперсию.

Однако наличие шумов с бесконечной дисперсией в динамической системе может привести к тому, что методы идентификации ее параметров, разработанные для случая конечной дисперсии, могут оказаться недостаточно эффективными или качественно поменять свои свойства. В этом случае необходимо дополнительное изучение процедур оценивания ее параметров.

В последние годы проявляется интерес к модели авторегрессии в случае, когда параметры модели не являются постоянными, а изменяются с течением времени. В прикладных задачах нашла применение модель пороговой авторегрессии вида:

$$X_k = \theta_{0j} + \sum_{i=1}^{p_j} \theta_{ij} X_{k-i} + \varepsilon_{kj}, \text{ когда } r_{j-1} \le X_{k-d} < r_j,$$

где $p_j, d \in R^+$, $j = 1, \ldots, m$, $\{\theta_{ij}\}$ – неизвестные параметры модели, $\{r_j\}_{0 \le j \le m+1}$ – пороги, такие, что $-\infty = r_0 < r_1 < \ldots < r_m < r_{m+1} = +\infty$.

Одной из важных характеристик различных процедур идентификации неизвестных параметров является совместное асимптотическое (при неограниченном росте числа наблюдений) распределение ошибок оценивания, поскольку оно позволяет строить доверительные области для неизвестных параметров с заданным уровнем доверия.

Цель диссертационной работы. Построение эффективных процедур оценивания параметров модели авторегрессии с бесконечной дисперсии шумов и нелинейной пороговой авторегрессии. Необходимо рассмотреть как классические процедуры оценивания с фиксированным числом наблюдений, так и последовательные, которые характеризуются тем, что число наблюдений определяется в ходе сбора информации согласно некоторому правилу остановки, и исследовать асимптотическое поведение оценок при неограниченном числе наблюдений.

Методы исследования. Для решения поставленных задач применялись методы теории вероятностей, последовательного анализа, теории мартингалов, теории марковских процессов, методов асимптотического анализа сумм зависимых случайных величин.

Научная новизна. Результаты выносимые на защиту. Научная новизна работы состоит в получении асимптотических свойств оценок параметров модели авторегрессии с бесконечной дисперсией и нелинейностями (пороговой авторегрессии). Результаты выносимые на защиту:

1. Взвешенная оценка по методу наименьших квадратов для модели линейной авторегрессии первого порядка с бесконечной дисперсией шума.
2. Асимптотическое распределение нормированного уклонения взвешенной оценки.
3. Асимптотическое поведение взвешенной оценки в сравнении с обычной оценкой МНК.
4. Предельное распределение МНК оценок взрывной пороговой авторегрессии первого порядка.
5. Последовательная процедура идентификации параметров пороговой авторегрессии.
6. Совместное асимптотическое распределение ошибок последовательных оценок пороговой авторегрессии.

Теоретическая ценность работы состоит в аналитическом решении задачи оценивания параметров модели авторегрессии с бесконечной дисперсией и пороговой авторегрессии.

Практическое значение работы. Построенные процедуры оценивания неизвестных параметров модели авторегрессии можно использовать на практике при построении прогнозов для стохастических динамических систем, описываемых уравнением авторегрессии или пороговой авторегрессии. Найденные предельные распределения оценок могут быть использованы при построении доверительных интервалов для неизвестных параметров.

Реализация и внедрение результатов работы. Результаты диссертации используются в учебном процессе на факультете естественных наук и математики Томского политехнического университета. Предложенные методы оценивания были реализованы посредством ЭВМ и используются при построении прогнозов финансовых индексов в рамках деятельности ООО «Эконофизика – Томск».

Апробация работы. Работа выполнялась в рамках научно-исследовательской работы при финансовой поддержке Российского фонда фундаментальных исследований. Основные результаты диссертации обсуждались на кафедре высшей математики и математического моделирования ТГУ, а также на следующих конференциях:

– на Всероссийской научной конференции молодых ученых «Наука. Технологии. Инновации» в г. Новосибирске, НГТУ, декабрь 2005г.;

– на четырнадцатой Всероссийской школе-коллоквиуме по стохастическим методам, восьмом Всероссийском симпозиуме по прикладной и промышленной математике в г. Сочи-Адлер, сентябрь-октябрь 2007г.;

– на пятнадцатой Международной конференции студентов, аспирантов и молодых ученых «Ломоносов» в г. Москва, МГУ, апрель 2008г.;

– на третьей Международной конференции по передовому управлению и компьютерной информации в г. Далянь, Китай, август 2008г.

Публикации. Основные результаты, полученные в диссертации, опубликованы в 6 работах, в том числе 2 работы в журналах из перечня ВАК.

Структура и объем диссертации. Диссертация состоит из введения, 3-х глав, заключения, списка использованной литературы и 3-х приложений. Работа содержит 117 страниц машинописного текста, 35 рисунков и 12 таблиц.

Краткое содержание работы

Во введении раскрывается актуальность исследуемой проблемы, приводится обзор известных результатов, формулируется цель и содержание работы, обосновывается теоретическая и практическая значимость.

В первой главе решается проблема построения оценки неизвестного параметра модели авторегрессии с бесконечной дисперсией вида:

$$X_k = \lambda X_{k-1} + \varepsilon_k, \; X_0 = 0, \; k \geq 1, \tag{1}$$

где λ – неизвестный параметр, $|\lambda| < 1$, $\{\varepsilon_k\}_{k\geq 1}$ – последовательность независимых одинаково распределенных случайных величин (шум) с симметричной функцией распределения, правильно меняющейся на бесконечности, с параметром $\alpha \in (0,2)$, т.е. для любого $x > 0$

$$P(|\varepsilon_1| > x) = x^{-\alpha} L(x), \tag{2}$$

где $L(x)$ медленно меняется на бесконечности, то есть

$$\lim_{x \to +\infty} L(tx)/L(x) = 1 \text{ для любого } t > 0.$$

К этому типу распределений относятся, например, симметричные устойчивые распределения с характеристической функцией (х.ф.)

$$\psi(t) = e^{-c|t|^{\alpha}}, \ c > 0, \ 0 < \alpha < 2.$$

Оценку неизвестного параметра λ по наблюдениям $X_1, \dots, X_n$ предлагается находить из условия минимума суммы взвешенных квадратов невязок

$$\sum_{k=1}^{n} \omega_k \left(X_k - \lambda X_{k-1} \right)^2,$$

используя весовые коэффициенты $\omega_k = X_{k-1}^{2m}$ при $m \geq 1$. При этом оценка имеет вид:

$$\lambda_{n,m} = \frac{\sum_{k=2}^{n} X_k X_{k-1}^{2m+1}}{\sum_{k=2}^{n} X_{k-1}^{2m+2}}. \tag{3}$$

Решается задача поиска асимптотического распределения оценки (3), а также задача сравнения точностных свойств изучаемой оценки в сравнении с обычной оценкой МНК вида:

$$\lambda_n = \frac{\sum_{k=2}^{n} X_k X_{k-1}}{\sum_{k=2}^{n} X_{k-1}}. \tag{4}$$

Используя уравнение (1), получаем вид ошибки оценивания:

$$\lambda_{n,m} - \lambda = \frac{\sum_{k=2}^{n} \varepsilon_k X_{k-1}^{2m+1}}{\sum_{k=2}^{n} X_{k-1}^{2m+2}}. \tag{5}$$

Показывается справедливость следующей теоремы, которая устанавливает предельное распределение нормированного уклонения взвешенной оценки $\lambda_{n,m}$.

Теорема 1. *Пусть распределение шума* $\{\varepsilon_k\}_{k\ge1}$ *в* (1) *обладает свойством* (2). *Тогда для любого фиксированного* $m \ge 1$ *асимптотическое распределение нормированного уклонения оценки* (5) *задается предельным соотношением:*

$$\lim_{n\to\infty} \frac{K_{2m+1}}{C_{2m+1}(\lambda)} \frac{\sum_{k=2}^{n} X_{k-1}^{2m+2}}{a_n^{2m+1}} \left(\lambda_{n,m} - \lambda\right) \overset{d}{=} Y_{2m+1},$$

где

$$a_n = \inf\left\{x:\ P\left(|\varepsilon_1| > x\right) \le \frac{1}{n}\right\},$$

$$K_p = \left(\frac{2}{\pi}\sin\left(\frac{\alpha\pi}{2p}\right)\Gamma\left(\frac{\alpha}{p}\right)\right)^{\frac{p}{\alpha}}, C_p(\lambda) = \left(E\left|\sum_{k=0}^{\infty}\lambda^k\varepsilon_k\right|^{\frac{\alpha}{p}}\right)^{\frac{p}{\alpha}},$$

а Y_p *– симметричная устойчивая случайная величина с х.ф.*

$$\psi(t) = e^{-|t|^{\frac{\alpha}{p}}},$$

причем $p = 2m+1$. *Здесь и далее символ* $\overset{d}{=}$ *означает, что величины имеют одинаковое распределение.*

Замечание 1. Предельное распределение ошибки оценивания можно использовать для построения доверительных интервалов для неизвестного параметра λ. Хотя теорема 1 верна для любого положительного целого m, целесообразно использовать $m = 1$, так как в этом случае предельное распределение оценки будет иметь наиболее «легкий» хвост.

Следствие 1. В случае, когда шумы имеют симметричное устойчивое распределение, константа $C_p(\lambda)$ принимает вид

$$C_p(\lambda) = \left(1 - |\lambda|^{\alpha}\right)^{-\frac{1}{\alpha}}\left(E|\varepsilon_1|^{\frac{\alpha}{p}}\right)^{\frac{p}{\alpha}}.$$

Представляет интерес сравнить асимптотические свойства взвешенной оценки (3) с обычной оценкой МНК (4). Известно, что предельное распределение нормированного уклонения оценки МНК λ_n от истинного значения λ определяется равенством:

$$\lim_{n\to\infty} \frac{a_n^2}{\tilde{a}_n}(\lambda_n - \lambda) \underset{n\to\infty}{\overset{d}{=}} \frac{1-\lambda^2}{\left(1-\lambda^{\alpha}\right)^{\frac{1}{\alpha}}} \frac{S_1}{S_0}, \tag{6}$$

где

$$a_n = \inf\left\{x:\ P\left(|\varepsilon_1| > x\right) \le 1/n\right\},\ \tilde{a}_n = \inf\left\{x:\ P\left(|\varepsilon_1\varepsilon_2| > x\right) \le 1/n\right\},$$

ε_1 и ε_2 независимы и имеют распределение, удовлетворяющее условию (2), а S_1 и S_0 – некоторые независимые устойчивые случайные величины.

Теорема 2. *В условиях теоремы 1 для любого фиксированного* $m \ge 1$

$$\frac{a_n^2}{\tilde{a}_n}\left(\lambda_{n,m} - \lambda\right) \xrightarrow[n\to\infty]{P} 0. \tag{7}$$

Замечание 2. Из (6) и (7) следует, что

$$\left(\lambda_{n,m} - \lambda\right) = o_p\left(\lambda_n - \lambda\right) \text{ при } n \to \infty,$$

где $o_p(\lambda_n - \lambda)$ обозначает величину более высокого порядка малости по вероятности по сравнению с $(\lambda_n - \lambda)$. Таким образом, взвешенная оценка МНК обладает более высокой асимптотической скоростью сходимости по распределению к истинному значению, чем обычная оценка МНК.

Во второй главе рассматривается модель пороговой авторегрессии первого порядка вида:

$$X_k = \sum_{i=0}^{m} \theta_i X_{k-1} I_{i,k-1} + \varepsilon_k,\ k \ge 1,\ X_0 = 0, \tag{8}$$

где $m \in R^+$, $\{X_k\}_{k\ge0}$ – наблюдаемый процесс, $\{\varepsilon_k\}_{k\ge1}$ – шумовая последовательность, $\{\theta_i\}_{0\le i\le m}$ – неизвестные параметры модели, $I_{i,k} = \chi_{\{r_i < X_k \le r_{i+1}\}}$ ($\chi_{\{\cdot\}}$ – индикаторная функция, $i = \overline{0,m}$), $\{r_i\}_{0\le i\le m+1}$ – известные пороги, такие, что

$$-\infty = r_0 < r_1 < \ldots < r_m < r_{m+1} = +\infty.$$

Рассматривается задача идентификации параметров $\{\theta_i\}_{0\le i\le m}$.

В первом разделе второй главы для оценивания неизвестных параметров модели предлагается использовать последовательную процедуру на основе метода наименьших квадратов. Делается предположение, что последовательность шумов $\{\varepsilon_k\}_{k\ge1}$ – это последовательность независимых одинаково распределенных случайных величин с непрерывной всюду положительной плотностью, причем $E\varepsilon_1 = 0$, $E\varepsilon_1^2 = 1$.

Для построения процедуры оценивания неизвестных параметров θ_1, θ_2 для каждого $H > 0$ вводятся моменты остановки:

$$\tau_j(H) = \inf\left\{n : \sum_{k=2}^{n} y_{j,k-1}^2 \geq H\right\},\ \text{(здесь и далее } j = \overline{0,m}\text{)},$$

при $y_{j,k} = X_k I_{j,k}$. Вводятся $0 < \alpha_j(H) \leq 1$ такие, что

$$\sum_{k=2}^{\tau_j(H)-1} y_{j,k-1}^2 + \alpha_j(H) y_{j,\tau_j(H)-1}^2 = H,$$

и

$$\beta_{j,k}(H) = \chi_{\{k<\tau_j(H)\}} + \alpha_j(H)\chi_{\{k=\tau_j(H)\}},\ k \geq 2.$$

Последовательные МНК – оценки записываются в виде:

$$\hat{\theta}_j(H) = \frac{\sum_{k=2}^{\tau_j(H)} \beta_{j,k}(H) y_{j,k-1} X_k}{\sum_{k=2}^{\tau_j(H)} \beta_{j,k}(H) y_{j,k-1}^2} = \frac{1}{H} \sum_{k=2}^{\tau_j(H)} \beta_{j,k}(H) y_{j,k-1} X_k. \tag{9}$$

При этом ошибка оценивания параметра θ_j записывается в виде:

$$\hat{\theta}_j(H) - \theta_j = \frac{1}{H} \sum_{k=2}^{\tau_j(H)} \beta_{j,k}(H) y_{j,k-1} \varepsilon_k. \tag{10}$$

Здесь коэффициенты $\beta_{j,k}(H)$ вводятся для того, чтобы оценки (9) в отличие от обычных оценок МНК обладали свойством несмещенности, то есть $E\hat{\theta}_j(H) = \theta_j$. Для формулировки основного результата делаются следующие обозначения:

$$\theta = (\theta_0, \theta_1, \ldots, \theta_m)^T,\ \hat{\theta}(H) = \left(\hat{\theta}_0(H), \hat{\theta}_1(H), \ldots, \hat{\theta}_m(H)\right)^T,$$

$$x = (x_0, x_1, \ldots, x_m)^T,\ \Phi_{m+1}(x) = \Phi(x_0)\cdot\Phi(x_1)\cdot\ldots\cdot\Phi(x_m),$$

где $\Phi(\cdot)$ – функция распределения стандартного нормального закона.

Показана справедливость следующей теоремы, которая показывает, что оценки $\hat{\theta}_j(H)$ асимптотически независимы, а их предельное распределение является стандартным гауссовым.

Теорема 3. *Пусть для некоторого компактного множества $K \subset R^{m+1}$ и любого $\delta > 0$ и $j = \overline{0,m}$ выполняются следующие условия:*

$$\lim_{p\to\infty} \sup_{\theta\in K} P_\theta\left(\exists n \geq p :\ y_{j,n}^2 > \delta \sum_{k=2}^{n} y_{j,k-1}^2\right) = 0, \tag{11}$$

$$\lim_{a\to\infty}\sup_{\theta\in K} P\left(y_{j,n}^2 > a\right) = 0 \text{ для любого } n \ge 2. \qquad (12)$$

Тогда ошибки (10) *удовлетворяют предельному соотношению:*

$$\lim_{H\to\infty}\sup_{\theta\in K}\sup_{x\in R^{m+1}}\left|P_\theta\left(\sqrt{H}\left(\hat\theta(H)-\theta\right)\le x\right)-\Phi_{m+1}(x)\right| = 0.$$

Замечание 3. Теорема 3 устанавливает, что оценки (9) обладают свойством **равномерной** по параметрам совместной асимптотической нормальности, когда исходный процесс удовлетворяет условиям (11), (12). Известно, что для обычных оценок МНК данное свойство не выполнено.

Рассматривается частный случай процесса (8) при $m=1$, $r_1=0$, а именно

$$X_k = \begin{cases}\theta_1 X_{k-1}+\varepsilon_k, \text{ при } X_{k-1}\le 0,\\ \theta_2 X_{k-1}+\varepsilon_k, \text{ при } X_{k-1}> 0.\end{cases}$$

Показывается следующая лемма, которая устанавливает, что условия теоремы 3 выполняются для любого компакта из области

$$\Theta = \{(\theta_0,\theta_1):\ \theta_0<1, \theta_1<1, \theta_0\theta_1<1\}.$$

Лемма 1. Для любого компакта $K\subset\Theta$ и любого $\delta>0$

$$\lim_{m\to\infty}\sup_{(\theta_1,\theta_2)\in K} P\left(\exists n\ge m:\ y_{n,j}^2 > \delta\sum_{k=2}^{n} y_{k-1,j}^2\right) = 0,$$

$$\lim_{a\to\infty}\sup_{(\theta_1,\theta_2)\in K} P\left(y_{n,j}^2 > a\right) = 0 \text{ для любого } n\ge 2.$$

Во втором разделе второй главы рассматривается взрывная модель пороговой авторегрессии вида:

$$X_k = \lambda_1 X_{k-1}^+ + \lambda_2 X_{k-1}^- + \varepsilon_k,\ X_0=0,\ k\ge 1,$$

где $\{\varepsilon_k\}_{k\ge1}$ - независимые одинаково распределенные случайные величины, λ_1,λ_2 – неизвестные параметры, удовлетворяющие условию:

$$\lambda_1<0, \lambda_2<0, \lambda_1\lambda_2>1.$$

Рассматривается оценка $\hat\lambda_n$ параметра $\lambda=[\lambda_1,\lambda_2]^T$ по методу наименьших квадратов

$$\hat\lambda_n = M_n^{-1}\sum_{k=1}^{n} Y_{k-1}X_k,$$

где

$$Y_{k-1}=\begin{bmatrix} X_{k-1}^{+} \\ X_{k-1}^{-} \end{bmatrix},\ M_n=\sum_{k=1}^{n} Y_{k-1}Y_{k-1}^{T}=\begin{bmatrix} \sum_{k=1}^{n}\left(X_{k-1}^{+}\right)^2 & 0 \\ 0 & \sum_{k=1}^{n}\left(X_{k-1}^{-}\right)^2 \end{bmatrix}.$$

Или покомпонентно

$$\hat{\lambda}_{1,n}=\frac{1}{S_n^{+}}\sum_{k=1}^{n} X_{k-1}^{+}X_k\ ,\ \hat{\lambda}_{2,n}=\frac{1}{S_n^{-}}\sum_{k=1}^{n} X_{k-1}^{-}X_k\ , \qquad (13)$$

где $S_n^{+}=\sum_{k=1}^{n}\left(X_{k-1}^{+}\right)^2$, $S_n^{-}=\sum_{k=1}^{n}\left(X_{k-1}^{-}\right)^2$.

Ошибки оценивания записываются в виде

$$\hat{\lambda}_{1,n}-\lambda_1=\frac{1}{S_n^{+}}\sum_{k=1}^{n} X_{k-1}^{+}\varepsilon_k\ ,\ \hat{\lambda}_{2,n}-\lambda_2=\frac{1}{S_n^{-}}\sum_{k=1}^{n} X_{k-1}^{-}\varepsilon_k\ .$$

Изучается предельное поведение ошибок оценивания в случае, когда распределение шумов имеет плотность $f(x)$, для которой выполняются условия:

A. Для любого $x\in R \quad f(x)>0$;

B. Для любого $x\in R \quad f(x)=f(-x)$;

C. Для некоторых $b>0,\beta>0,x^{*}>0$

$$F(-x)=1-F(x)\le bx^{\beta} \text{ для } x\ge x^{*}.$$

Теорема 4. *При выполнении условий A-C для оценок* (13) *справедливо предельное соотношение:*

$$\lim_{n\to\infty}\sup_{x,y\in R}\left|P\left(\sqrt{S_n^{+}}\left(\hat{\lambda}_1-\lambda\right)\le x,\sqrt{S_n^{-}}\left(\hat{\lambda}_2-\lambda\right)\le y\right)-F(x)F(y)\right|=0,$$

где $F(\cdot)$ *– функция распределения случайной величины*

$$\xi=\sqrt{1-\theta^2}\sum_{k=0}^{\infty}\varepsilon_k\theta^k\ ,\ \theta=\frac{1}{\lambda_1\lambda_2}.$$

Замечание 4. В случае стандартных гауссовых шумов случайная величина ξ также будет иметь стандартное гауссово распределение.

В третьей главе приводятся результаты имитационного моделирования для:

1. взвешенной оценки МНК параметра авторегрессии первого порядка с бесконечной дисперсией (первый раздел);
2. последовательных оценок МНК параметров пороговой авторегрессии первого порядка (второй раздел).

3. обычной оценки МНК параметров взрывной пороговой авторегрессии первого порядка (третий раздел);

В первом разделе главы 3 представлены результаты численного моделирования взвешенной и обычной оценок МНК параметра модели авторегрессии с бесконечной дисперсией. Имитационное моделирование процесса (1) проводилось с использованием симметричного устойчивого распределения и симметричного распределения типа Парето, для которого $P(|\varepsilon_1| > x) = (1+x)^{-\alpha}$, $x \ge 0$, $\alpha \in (0,2)$. В рамках численного моделирования необходимо определить насколько отличается выборочное распределение взвешенной оценки $\lambda_{n,1}$ от асимптотического. Согласно утверждению теоремы 1 нормированное уклонение оценки $\lambda_{n,1}$ вычислялось по формуле:

$$\frac{K_3}{C_3} \frac{\sum_{k=2}^{n} X_{k-1}^4}{a_n^3} (\lambda_{n,1} - \lambda). \tag{14}$$

На рисунке 1 приводится оценка плотности распределения величины (14), построенная по 10 000 повторений процедуры оценивания (пунктирная линия) и ее теоретическая предельная плотность $p(x)$ с характеристической функцией $\psi(t) = e^{-|t|^\alpha}$ (непрерывная линия) для $n = 50$, $\alpha = 1{,}5$, $\lambda = -0{,}5$.

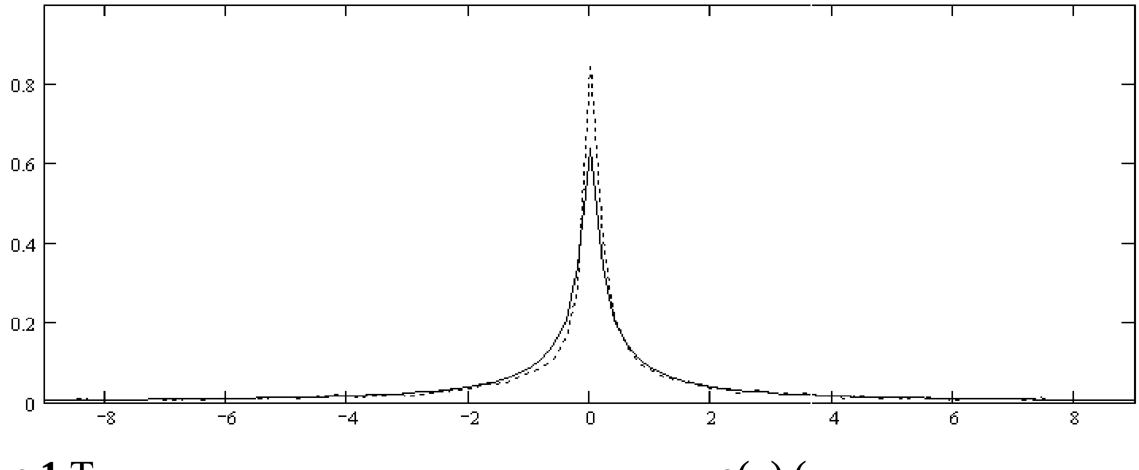

Рис 1 Теоретическая предельная плотность p(x) (непрерывная линия) и оценка плотности нормированного уклонения при $n = 50$, $\alpha = 1{,}5$, $\lambda = -0{,}5$.

Как видно из рисунка, выборочная плотность по сравнению с теоретической имеет более высокий пик в окрестности нуля, а на хвостах плотности практически совпадают. Это говорит о том, что при малом числе наблюдений доверительные границы могут быть несколько завышены.

Для сравнения точности оценивания взвешенной оценки $\lambda_{n,1}$ и обычной оценки МНК λ_n вводится функция относительной эффективности оценок

$$\rho(\lambda,\beta,n)=\frac{E|\lambda_n-\lambda|^\beta}{E|\lambda_{n,1}-\lambda|^\beta},\ 0<\beta<\alpha, \tag{15}$$

которая является отношением абсолютных моментов порядка β для обычной оценки МНК и взвешенной оценки $\lambda_{n,1}$.

Для оценки математических ожиданий в (15) для каждого значения параметра $\lambda=0{,}01i$, $i=-100,\dots,100$, процедуры оценивания запускались 100 000 раз. При этом значение параметра β выбиралось равным $\alpha/2$. На рисунке 2 приводятся оценки функции относительной эффективности для n = 25 и n = 100 при шумах с распределением Коши ($\alpha=1{,}0$).

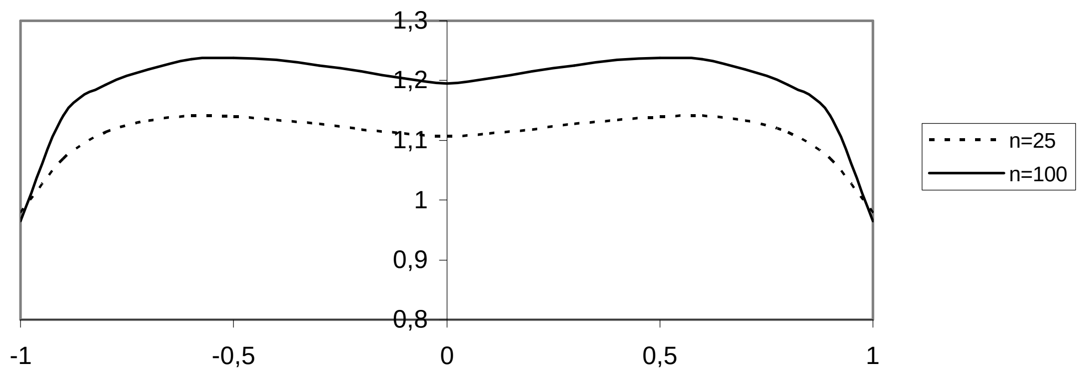

Рис. 2. Функция относительной эффективности оценок при $\alpha=1{,}0$.

Из рисунка 2 видно, что в большинстве точек значение функции относительной эффективности больше единицы для $\alpha=1{,}0$. Это означает, что среднее порядка $\alpha/2$ для абсолютного уклонения оценки МНК больше, чем для взвешенной оценки МНК. Отметим, что значение функции эффективности в целом возрастает при увеличении объема наблюдений с 25 до 100.

Для сравнения точностей взвешенной и обычной оценок МНК в зависимости от «тяжести» хвоста распределения (значения параметра α) наряду с (15) вычислялась функция средней относительной эффективности оценок

$$\rho(\alpha,\beta,n,\delta)=\frac{1}{2-2\delta}\int_{-1+\delta}^{1-\delta}\frac{E|\lambda_n-\lambda|^\beta}{E|\lambda_{n,1}-\lambda|^\beta}d\lambda,\ 0<\beta<\alpha,$$

которая является средним по параметру λ от функции относительной эффективности (15).

На рисунке 3 приводятся выборочные функции средней относительной эффективности при $\beta = \alpha / 2$, $\delta = 0{,}01$ для $n = 25$ и $n = 50$ в зависимости от значений параметра α (построенные по 100 000 повторений процедур оценивания) для устойчивых шумов.

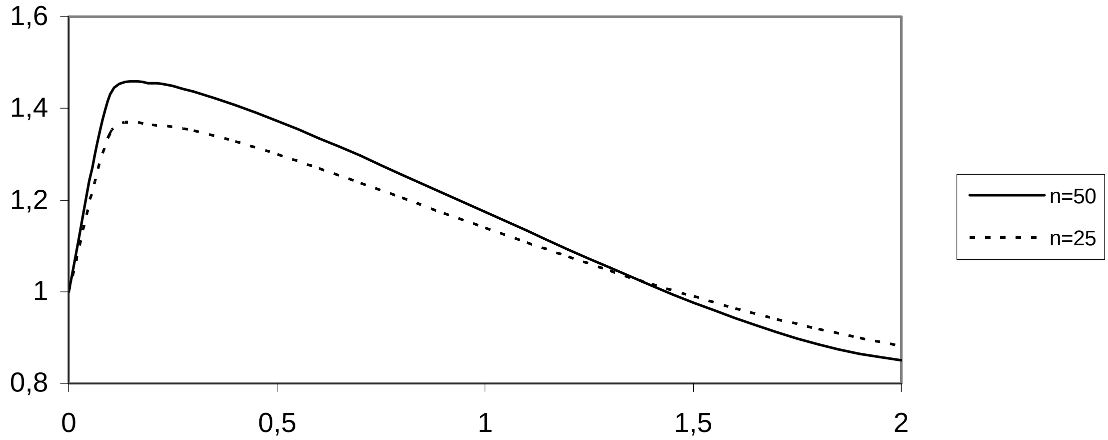

Рис. 3. Функция средней относительной эффективности оценок.

Во втором разделе главы 3 представлены результаты численного моделирования последовательных оценок параметров процесса пороговой авторегрессии первого порядка с одним порогом в нуле. Приводятся оценки плотностей распределений оценок параметров, а также выборочные характеристики:

1. Оценки плотностей распределений нормированных ошибок оценивания $\sqrt{H}\left(\hat{\theta}_j(H) - \theta_j\right)$.
2. $N_j(H)$ – среднее число наблюдений.
3. $D_j(H)$ – выборочные дисперсии ошибок $\left(\hat{\theta}_j(H) - \theta_j\right)$.
4. $\rho(H)$ – выборочная корреляция ошибок $\left(\hat{\theta}_j(H) - \theta_j\right)$.
5. $\Delta(H)$ – максимальное отклонение двумерной эмпирической функции распределения от теоретического предельного распределения.

На рисунке 4 представлены оценки плотностей нормированных уклонений и теоретические предельные плотности (плотность стандартного нормального распределения) для $\theta_1 = 0{,}35$, $\theta_2 = -0{,}45$ при $H = 5$ в предположении, что шумы имеют стандартное нормальное распределение. На рисунке p(x) показывает предельную плотность, а ~p(x) – оценку плотности нормированного уклонения. В таблице 1 приводятся выборочные характеристики 2–5.

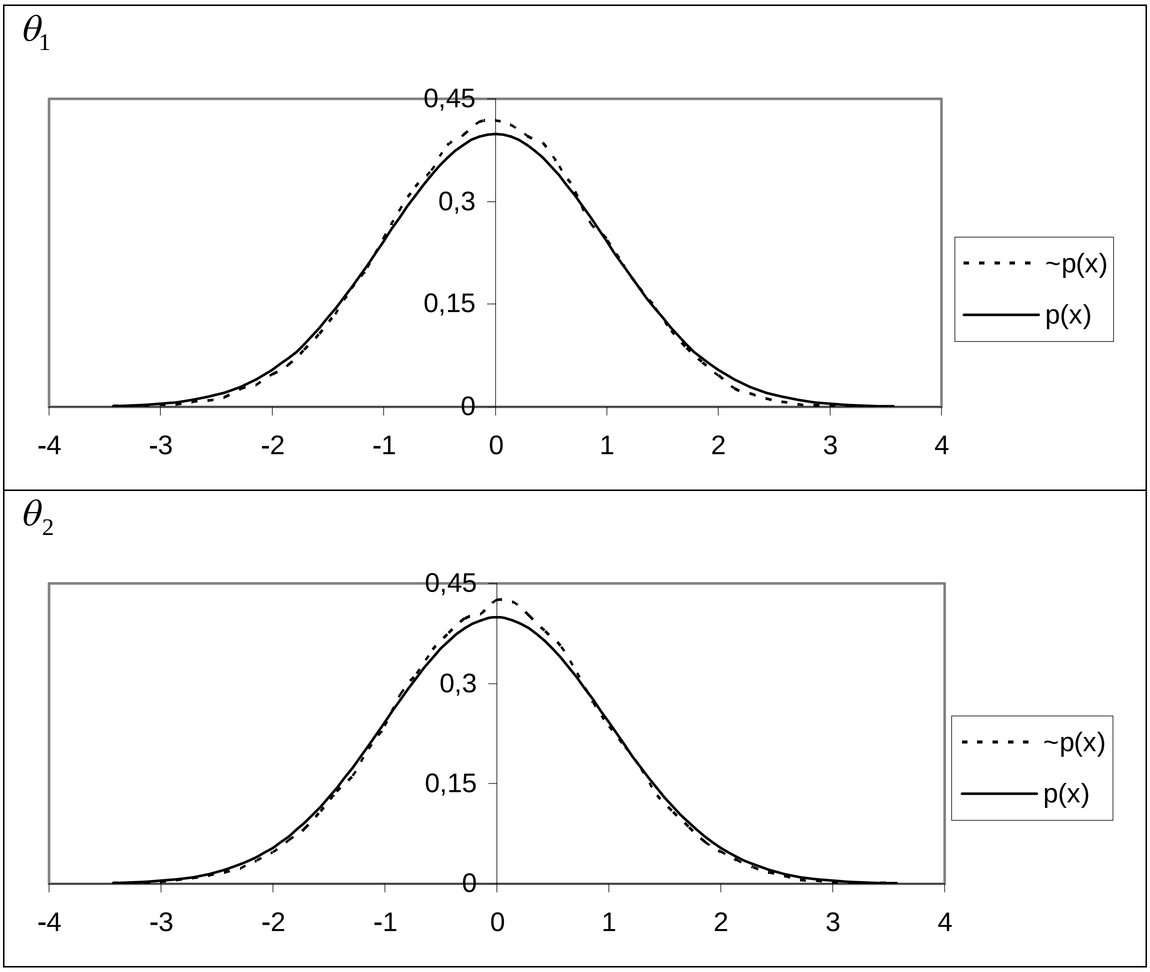

Рис. 4. Теоретическая предельная плотность p(x) и ее оценка ~p(x). Сверху для θ_1. Снизу для θ_2. $\theta_1 = 0{,}35$, $\theta_2 = -0{,}45$, $H = 5$.

Таблица 1. Выборочные характеристики оценок. $\theta_1 = 0{,}35, \theta_2 = -0{,}45, H = 5$.

$N_1(H)$	$N_2(H)$	$D_1(H)$	$D_2(H)$	$\rho(H)$	$\Delta(H)$
5,321710	8,554390	0,395957	0,408710	-0,000996	0,020570

Из рисунка 4 и таблицы 1 видно, что уже при малом пороге *H*, когда среднее число наблюдений по каждому параметру не превышает 9, выборочная плотность практически совпадает с теоретической предельной.

В третьем разделе главы 3 для взрывной модели пороговой авторегрессии приводятся оценки плотностей распределений ошибок оценок параметров по методу наименьших квадратов, построенные по 10 000 повторений процедуры оценивания для различных значений параметров λ_1, λ_2 при нормальном распределении шумов.

На рисунке 5 представлены выборочные плотности распределения ошибок оценивания для $\lambda_1 = -2$, $\lambda_2 = -1$ при $n = 20$.

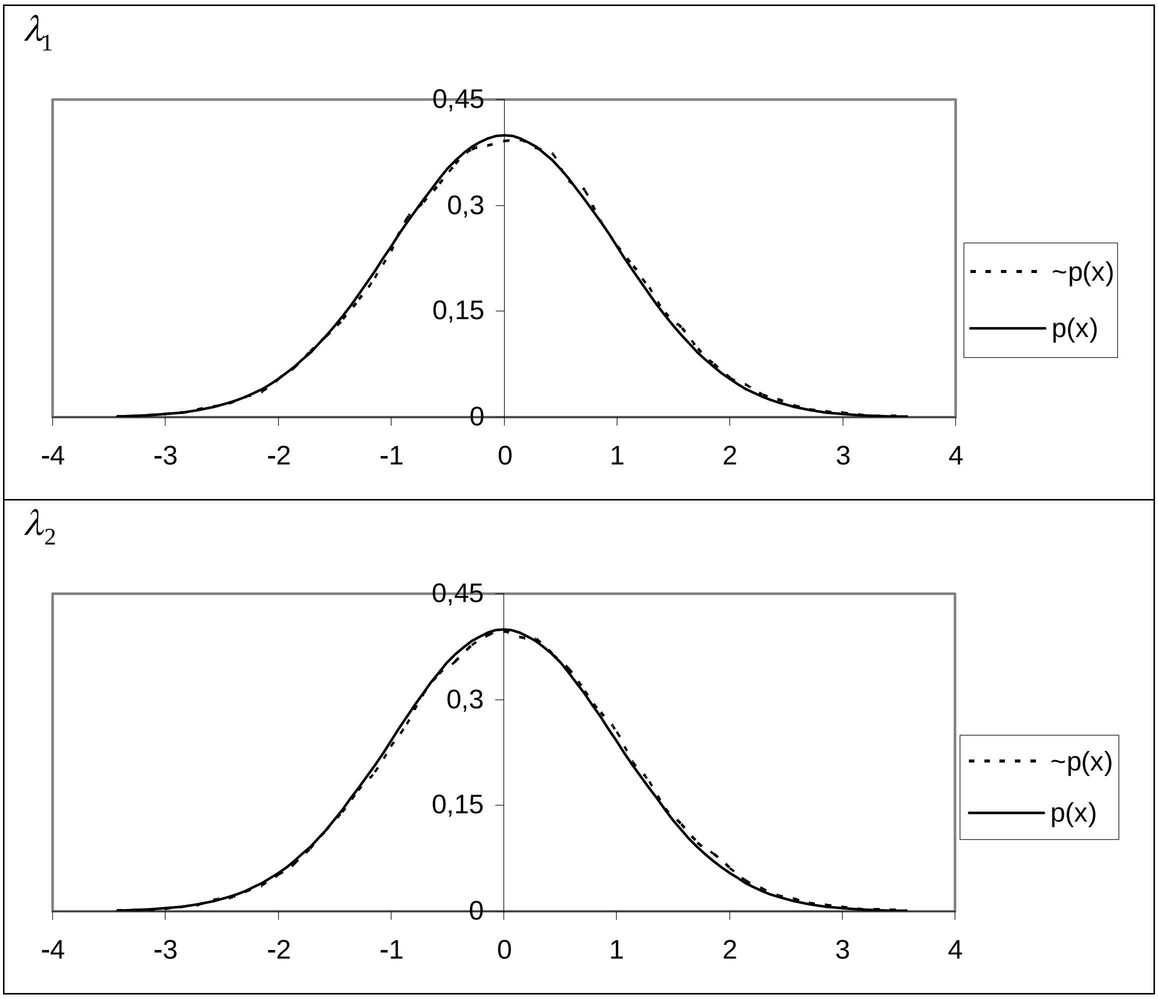

Рис 5. Теоретическая предельная плотность p(x) и ее оценка ~p(x). Сверху для λ_1. Снизу для λ_2. $\lambda_1 = -2$, $\lambda_2 = -1$, $n = 20$.

Как видно из рисунка 5, уже при объеме наблюдений $n = 20$ выборочные плотности нормированных ошибок оценивания практически совпадают с теоретической предельной гауссовой плотностью. Выборочная корреляция при $n = 20$ составила 0,018814.

В заключении формулируются основные теоретические и практические результаты диссертации, которые состоят в следующем:

1) Построена взвешенная оценка по методу наименьших квадратов для оценивания параметра модели авторегрессии с бесконечной дисперсией.
2) Найдено асимптотическое распределение ошибки взвешенной оценки параметра авторегрессии в случае, когда шумы обладают функцией распределения, правильно меняющейся на бесконечности.

3) Установлено, что взвешенная оценка обладает более высокой асимптотической скоростью сходимости к истинному значению параметра в сравнении с обычной оценкой по методу наименьших квадратов.
4) Для оценивания параметров модели пороговой авторегрессии предложены последовательные оценки со специальными моментами остановки.
5) Найдено совместное предельное распределение ошибок последовательных оценок параметров пороговой авторегрессии. Показана асимптотическая независимость оценок.
6) Найдено асимптотическое распределение ошибок оценок параметров взрывной пороговой авторегрессии с одним порогом по методу наименьших квадратов. Во взрывной области исходный процесс не является эргодическим.
7) Проведено имитационное моделирование рассмотренных процедур оценивания неизвестных параметров посредством ЭВМ. Все полученные теоретические результаты были подтверждены с помощью компьютерного моделирования. Проведено численное сравнение рассмотренных процедур оценивания с классическими методами.

ПУБЛИКАЦИИ ПО ТЕМЕ ДИССЕРТАЦИИ

1. **Марков А.С. Оценивание параметра авторегрессии с бесконечной дисперсией шума // Автоматика и Телемеханика. – 2009/ – №1, – С. 104-118.**
2. **Марков А.С. Последовательная идентификация пороговой авторегрессии // Известия ТПУ. – 2009/ – Т. 314, №2/ – С. 21-26.**
3. Марков А.С. Оценивание параметра авторегрессии с бесконечной дисперсией шума // Наука. Технологии. Инновации : материалы Всеросс. науч. конф. молодых ученых. 08-11 декабря 2005. Новосибирск. – Новосибирск : НГТУ. – 2006. – Ч.1. – С. 57-58.
4. Марков А.С. Оценивание параметра авторегрессии при бесконечной дисперсии шумов // Обозрение прикладной и промышленной математики : Четырнадцатая школа-коллоквиум по стохастическим методам. – 2007. – Т.15, вып. 1. – С. 92-93.
5. Марков А.С. Последовательная идентификация пороговой авторегрессии // Материалы докладов XV Международной конференции студентов, аспирантов и молодых ученых «Ломоносов», М. : Изд-во МГУ; СП МЫСЛЬ, 2008. – 1 электрон. опт. диск (CD-ROM). – С. 30-31.
6. Markov A.S. Parameter Estimation in an Autoregression Model with Infinite Variance [Online resource] // 3rd International Conference on Innovative Computing Information and Control. – Dalian, China, 2008. – P. 586-587. – URL: http://doi.ieeecomputersociety.org/10.1109/ICICIC.2008.414.