ОБ ОДНОМ МЕТОДЕ РОБАСТНОГО ОЦЕНИВАНИЯ ПАРАМЕТРОВ AR(p) ПРИ НАЛИЧИИ “ВЫБРОСОВ” И “ПРОПУСКОВ”

Ю.С. Харин, В.А. Волошко

1. Введение. Математическая модель

В приложениях часто в качестве вероятностной модели наблюдаемого временного ряда используется модель *AR(p)* гауссовского авторегрессионного временного ряда y_t, задаваемая на вероятностном пространстве (Ω,F,P) стохастическим разностным уравнением порядка $p \in N$ [1]:

$$y_t + b_1 y_{t-1} + \ldots + b_p y_{t-p} = \xi_t, t \in Z, \qquad (1)$$

где $b = (b_1,\ldots,b_p)' \in R$ - вектор-столбец неизвестных коэффициентов авторегрессии, удовлетворяющий условию стационарности [1], N,Z – множества натуральных и целых чисел соответственно, $\{\xi_t : t \in Z\}$ - последовательность независимых в совокупности одинаково распределенных гауссовских случайных величин с законом распределения $L\{\xi_t\} = N_1(0,\sigma^2), 0 < \sigma^2 < +\infty$ - неизвестная дисперсия инновационного процесса.

На практике наблюдаемый временной ряд $x_t, t \in Z,$ не соответствует гипотетической модели (1) из-за часто встречающихся искажений [2,3,4,5]. В данной статье рассматривается случай двух одновременно влияющих типов искажений: “выбросы” и “пропуски”. Здесь рассматривается случай аддитивных “выбросов”:

$$z_t = (1-\eta_t) y_t + \eta_t \upsilon_t, t \in Z, \qquad (2)$$

где $\{\eta_t \in \{0,1\}\}$ - независимые в совокупности случайные величины Бернулли,

$$P\{\eta_t = 1\} = 1 - P\{\eta_t = 0\} = \varepsilon, \qquad (3)$$

$0 \le \varepsilon << 1$ - достаточно малая вероятность появления выброса, $\{\upsilon_t \in R\}$ - независимые в совокупности одинаково распределенные случайные величины с некоторой неизвестной симметричной функцией распределения $H(\upsilon)$, нулевым средним $E\{\upsilon_t\} = 0$ и дисперсией $D\{\upsilon_t\} >> D\{y_t\}$.

Искажения, порождаемые “пропусками”, условимся характеризовать при помощи неслучайной двоичной индексной последовательности

$$o_t = \{1, \text{если } z_t \text{ наблюдается}; 0, \text{если } z_t \text{ пропущено}\}, t \in Z. \qquad (4)$$

Обозначим $\{t_k\}$ - упорядоченную в порядке возрастания последовательность моментов времени, для которых наблюдения доступны; $t_- = \min t_k$, $t_+ = \max t_k$ (без потери общности $t_- = 1$, $t_+ = T$). Тогда наблюдаемый временной ряд представим в виде

$$x_k = z_{t_k}, k = 1,\ldots,K,$$

где K – число наблюдений. Задача заключается в статистическом оценивании коэффициентов b по наблюдениям $\{x_k\}$.

Дадим краткий обзор результатов по робастному оцениванию параметров *AR(p)*. Основной подход заключается в робастном оценивании ковариационной функции $\sigma_\tau = Cov\{y_t, y_{t+\tau}\} = E\{y_t y_{t+\tau}\}, \tau = 0,1,\ldots,p,$ и нахождении оценок для b при решении уравнений Юла-Уокера. Известны следующие робастные оценки ковариаций: оценка Хьюбера [2], медианный коэффициент корреляции и его обобщения [6], оценки, основанные на непараметрических мерах (Blomqvist, 1950), на отсеивании выбросов, на робастных ковариациях [2] и др. Отличительной особенностью данной статьи является использование новых (параметрических) робастных оценок корреляционной функции с использованием дополнительной априорной информации о гауссовском распределении ξ_t в (1) и специальных свойств распределения вероятностей Коши.

2. Робастное оценивание корреляционной функции

Обозначим:

$$\theta_\tau = Corr\{y_t, y_{t+\tau}\} = \sigma_\tau / \sigma_0, \tau \in Z, \text{ -} \quad (5)$$

подлежащая оцениванию корреляционная функция; $med\{z_1,...,z_T\}$ - выборочная медиана. Согласно [2], робастная оценка коэффициента корреляции θ_τ может быть вычислена при наличии пропусков с помощью следующего алгоритма.

1. Вычисляем вспомогательные коэффициенты:
   $a = 1/med\{|z_t| : o_t = 1, t = 1,...,T-\tau\}$, $b = 1/med\{|z_t| : o_t = 1, t = 1+\tau,...,T\}$.
2. Находим
   $S_+ = med\{|az_t + bz_{t+\tau}| : o_t = o_{t+\tau} = 1, t = 1,...,T-\tau\}$, $S_- = med\{|az_t - bz_{t+\tau}| : o_t = o_{t+\tau} = 1, t = 1,...,T-\tau\}$.
3. Вычисляем оценку Хьюбера:

$$\hat{\theta}_\tau = S_+^2 - S_-^2 / S_+^2 + S_-^2 . \quad (6)$$

Для построения (параметрической) робастной оценки θ_τ воспользуемся специальным свойством распределения Коши.

Лемма 1. Если $y_t, t \in Z$, - стационарный гауссовский временной ряд с нулевым средним и некоторой корреляционной функцией (5), то для любого $\tau \neq 0$ отношение случайных величин $\zeta = y_t / y_{t+\tau}$ имеет распределение Коши $C(\theta_\tau, \sqrt{1-\theta_\tau^2})$.

Доказательство. По условию имеем:

$$L\left\{\begin{pmatrix} y_t \\ y_{t+\tau} \end{pmatrix}\right\} = N_2\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma^2 & \theta_\tau \sigma^2 \\ \theta_\tau \sigma^2 & \sigma^2 \end{pmatrix}\right).$$

В силу известного регрессионного свойства многомерного нормального распределения [7] справедливо представление:

$$y_t = \theta_\tau y_{t+\tau} + \sqrt{1-\theta_\tau^2} \cdot \sigma \cdot \eta_t ,$$

где η_t не зависит от $y_{t+\tau}$ и имеет стандартное распределение $N_1(0,1)$. Следовательно, $\zeta = \theta_\tau + \sqrt{1-\theta_\tau^2} \cdot \eta_t / (y_{t+\tau}/\sigma)$, где случайная величина $\zeta_0 = \eta_t / (y_{t+\tau}/\sigma)$ имеет стандартное распределение Коши $C(0,1)$, поэтому согласно свойствам этого распределения [8] линейное преобразование $\zeta = \theta_\tau + \sqrt{1-\theta_\tau^2}\zeta_0$ приводит к распределению $C(\theta_\tau, \sqrt{1-\theta_\tau^2})$. ■

Лемма 2. Если $\varphi : R^2 \to R$ - ограниченная, нечетная по обеим переменным функция, а z_t - временной ряд, определяемый моделью (1) - (3), то

$$E\{\varphi(z_t, z_{t+\tau})\} = (1-\varepsilon)^2 E\{\varphi(y_t, y_{t+\tau})\}. \quad (7)$$

Доказательство. В силу (2) по формуле полного математического ожидания

$$E\{\varphi(z_t, z_{t+\tau})\} = (1-\varepsilon)^2 E\{\varphi(y_t, y_{t+\tau})\} + \varepsilon(1-\varepsilon)\big(E\{\varphi(y_t, \upsilon_{t+\tau})\} + E\{\varphi(\upsilon_t, y_{t+\tau})\}\big) + \varepsilon^2 E\{\varphi(\upsilon_t, \upsilon_{t+\tau})\}.$$

Так как y_t - стационарный гауссовский процесс с симметричным относительно $y=0$ распределением $L\{y_t\} = N_1(0, D\{y_t\})$, а υ_t также имеет симметричное относительно $\upsilon=0$ распределение и $\{y_t\},\{\upsilon_t\}$ - независимы, то в силу нечетности $\varphi(\cdot)$ все математические ожидания, кроме первого, в этом выражении равны нулю. ■

Выберем функцию $\varphi(\cdot)$ в (7) следующим специальным образом:

$$\varphi(z_t, z_{t+\tau}) = \psi(z_t / z_{t+\tau}), \quad (8)$$

где $\psi(\cdot)$ - некоторая пока произвольная ограниченная, нечетная функция: $|\psi(u)| \leq c_0 < +\infty$, $\psi(-u) = -\psi(u)$, $u \in R$. С помощью этой функции и Леммы 1 построим ограниченную, нечетную функцию [-1,+1] →R:

$$f_\psi(\theta_\tau) ::= E\{\psi(\zeta)\} = \frac{\sqrt{1-\theta_\tau^2}}{\pi} \int_{-\infty}^{+\infty} \frac{\psi(z)}{1-\theta_\tau^2 + (z-\theta_\tau)^2} dz = \frac{4\theta_\tau\sqrt{1-\theta_\tau^2}}{\pi} \int_0^{+\infty} \frac{z\psi(z)}{z^4 - 2z^2(2\theta_\tau^2 - 1) + 1} dz. \quad (9)$$

Прежде чем привести основной результат, докажем две вспомогательные леммы.

Лемма 3. Если $\tau \in N, f(\cdot): R^{\tau+1} \to R$ - произвольная борелевская функция, а $\{z_t : t \in N\}$ - произвольная стационарная случайная последовательность, обладающая сильным перемешиванием с коэффициентом $\alpha_z(n)$, то случайная последовательность $s_t = f(z_t,...,z_{t+\tau})$ обладает сильным перемешиванием с коэффициентом

$$\alpha_s(n) \le \begin{cases} 1/4,\ 0 \le n < \tau, \\ \alpha_z(n-\tau),\ n \ge \tau. \end{cases}$$

Доказательство. По определению сильного перемешивания [9,10]

$$\alpha_s(n) = \sup_{t\ge1} \sup_{A\in G_1^t, B\in G_{t+n}^{+\infty}} |P(AB) - P(A)P(B)|,$$

где $G_a^b = \sigma\{s_t : t \in [a,b]\}$ - σ – алгебра, порожденная траекторией s_t на отрезке *[a,b]*. Пусть $F_a^b = \sigma\{z_t : t \in [a,b]\}$. Тогда по построению $G_1^t \subset F_1^{t+\tau}, G_{t+n}^{+\infty} \subset F_{t+n}^{+\infty}$, так что при $n \ge \tau$ справедливо неравенство: $\alpha_s(n) \le \alpha_z(n-\tau)$. Так как всегда $|P(AB) - P(A)P(B)| \le 1/4$, то $\alpha_s(n) \le 1/4$ при $n > \tau$. ■

Лемма 4. Если случайная последовательность y_t обладает сильным перемешиванием с коэффициентом $\alpha_y(n)$, $\{\mu_t\}$ – независимые одинаково распределенные случайные величины, не зависящие от $\{y_t\}$, а $g(\cdot)$ - произвольная борелевская функция, то $z_t = g(y_t, \mu_t)$ также обладает сильным перемешиванием с коэффициентом $\alpha_z(n) \le \alpha_y(n), n \in N$.

Доказательство. Указанный результат следует из Теоремы 5.2 [9]. ■

Теорема 1. Если имеет место модель (1)-(4) и функция $\psi(\cdot)$ такова, что функция $f_\psi(\cdot)$, определяемая (9), имеет непрерывную обратную функцию $f_\psi^{-1}(\cdot)$, и $N_{T\tau} = \sum_{t=1}^{T-\tau} o_t o_{t+\tau} \to \infty$ при $T \to \infty$, то

$$\widehat{\theta}_\tau = f_\psi^{-1}\left((1-\varepsilon)^{-2} \sum_{t=1}^{T-\tau} \psi(z_t / z_{t+\tau}) o_t o_{t+\tau} \Big/ \sum_{t=1}^{T-\tau} o_t o_{t+\tau} \right),\ 0 < \tau < T,$$

является состоятельной оценкой коэффициента корреляции θ_τ:

$$\widehat{\theta}_\tau \xrightarrow{P} \theta_\tau. \tag{10}$$

Доказательство. В силу непрерывности $f_\psi^{-1}(\cdot)$ и свойства сходимости по вероятности функционально преобразованной сходящейся последовательности достаточно доказать справедливость следующего закона больших чисел:

$$S_T = \sum_{t=1}^{T-\tau} \psi(z_t / z_{t+\tau}) o_t o_{t+\tau} \Big/ N_{T\tau} \xrightarrow{P} E\{\psi(z_t / z_{t+\tau})\}. \tag{11}$$

Действительно, в силу Леммы 2, (8), (9), $E\{\psi(z_t / z_{t+\tau})\} = (1-\varepsilon)^2 f_\psi(\theta_\tau)$, и для получения (10) достаточно обратить (11).

Для доказательства (11) проверим условие Маркова [11]. Исследуем дисперсию

$$D\{S_T\} = \frac{1}{N_{T\tau}^2} \left(\sum_{t=1}^{T-\tau} o_t o_{t+\tau} \sigma_\psi(0) + 2 \sum_{t=1}^{T-\tau-1} \sum_{t'=t+1}^{T-\tau} o_t o_{t+\tau} o_{t'} o_{t'+\tau} \sigma_\psi(t'-t) \right),$$

где $\sigma_\psi(t'-t) = Cov\{\psi(z_t / z_{t+\tau}), \psi(z_{t'} / z_{t'+\tau})\}$. Так как $|\psi(u)| \le c_0 < +\infty$, то справедлива оценка сверху:

$$D\{S_T\} \le \frac{1}{N_{T\tau}^2} \left(N_{T\tau} \cdot c_0^2 + 2 \sum_{t=1}^{T-\tau-1} o_t o_{t+\tau} \sum_{t'=t+1}^{T-\tau} |\sigma_\psi(t'-t)| \right) =$$

$$= \frac{1}{N_{T\tau}^2} \left(N_{T\tau} c_0^2 + 2 \sum_{t=1}^{T-\tau-1} o_t o_{t+\tau} \sum_{i=1}^{T-\tau-t} |\sigma_\psi(i)| \right). \tag{12}$$

В силу Леммы 1.2.1 [12] $|\sigma_\psi(t'-t)|\le 4c_0^2\alpha_s(|t'-t|)$, где $\alpha_s(n), n\in N$, - коэффициент перемешивания [12] для стационарной случайной последовательности $s_t=\psi(z_t/z_{t+\tau})$. Согласно Лемме 3, если $\alpha_z(n)$ - коэффициент перемешивания для стационарной случайной последовательности z_t, то

$$\alpha_s(n)\le\begin{cases}1/4,\ 0\le n<\tau,\\ \alpha_z(n-\tau),\ n\ge\tau.\end{cases}$$

Наконец, в силу Леммы 4 $\alpha_z(n)\le\alpha_y(n)$, где $\alpha_y(n)$ - коэффициент перемешивания для *AR(p)* – временного ряда (1).

Согласно [10], для *AR(p)* - временного ряда (1), удовлетворяющего условию стационарности, справедлива экспоненциальная асимптотическая оценка коэффициента перемешивания (при $n\to+\infty$): $\alpha_y(n)=O(e^{-bn})$, где $b>0$ - некоторая константа.

Объединяя приведенные выше оценки и подставляя результат в (12), получаем искомую оценку дисперсии: $D\{S_T\}\le\left(c_0^2+2\cdot S\right)/N_{T\tau}$, где ряд $S=\sum_{i=1}^{+\infty}|\sigma_\psi(i)|\le c_*<+\infty$ сходится в силу установленной асимптотики. В результате находим $D\{S_T\}\le\left(c_0^2+2\cdot c_*\right)/N_{T\tau}$, что и означает справедливость (11), влекущую (10). ■

Выбирая различные функции $\psi(\cdot)$ в (9), с помощью Теоремы 1 можно получить семейство состоятельных оценок для θ_τ:

<u>S – оценка</u>: $\psi(x)=sign(x)$, $f_\psi(y)::=\mathrm{E}\{sign(\zeta)\}=\int_R\frac{\sqrt{1-y^2}\,sign(x)}{\pi(x^2-2xy+1)}dx=\frac{2}{\pi}\arcsin y$,

$$\hat{\theta}_\tau=\sin\left(\frac{\pi}{2(1-\varepsilon)^2}\sum_{t=1}^{T-\tau}sign(z_t/z_{t+\tau})o_to_{t+\tau}\Big/\sum_{t=1}^{T-\tau}o_to_{t+\tau}\right). \qquad (13)$$

<u>A – оценка</u>: $\psi(x)=\arctan(x)$, $f_\psi(y)::=\mathrm{E}\{\arctan\zeta\}=\int_R\frac{\sqrt{1-y^2}\,\arctan x}{\pi(x^2-2xy+1)}dx=\frac{1}{2}\arcsin y$,

$$\hat{\theta}_\tau=\sin\left(\frac{2}{(1-\varepsilon)^2}\sum_{t=1}^{T-\tau}\arctan(z_t/z_{t+\tau})o_to_{t+\tau}\Big/\sum_{t=1}^{T-\tau}o_to_{t+\tau}\right).$$

<u>T – оценка</u>: $\psi(x)=2x/(1+x^2)$, $f_\psi(y)::=\int_R\frac{\sqrt{1-y^2}\cdot 2x}{\pi(x^2-2xy+1)\cdot(1+x^2)}dx=\frac{1-\sqrt{1-y^2}}{y}$,

$$\hat{\theta}_\tau=\psi\left(\frac{1}{(1-\varepsilon)^2}\sum_{t=1}^{T-\tau}\arctan(z_t/z_{t+\tau})o_to_{t+\tau}\Big/\sum_{t=1}^{T-\tau}o_to_{t+\tau}\right).$$

Исследуем свойства S – оценки (13). Для произвольной последовательности $f:Z\to R$ обозначим $\Sigma^f=\sum_{s\in Z}f(s)$.

<u>Лемма 5.</u> Пусть $a_t=sign(z_tz_{t+\tau})$, $u_t=sign(y_ty_{t+\tau})$, $\sigma(s)=cov\{a_t,a_{t+s}\}$, $\rho(s)=cov\{u_t,u_{t+s}\}$. Тогда справедливо неравенство:

$$\Sigma^\sigma\le\varepsilon(2-\varepsilon)(1+3(1-\varepsilon)^2)+(1-\varepsilon)^4\Sigma^\rho.$$

<u>Доказательство.</u> Пусть $s>0, s\neq\tau$. Согласно (2) $\eta_t=1$, если z_t - выброс. Очевидно $E\{a_t|\eta_t+\eta_{t+\tau}>0\}=0$, т.е. если произошел выброс в момент t или $t+\tau$. Легко видеть также, что при таком условии имеет место условная независимость a_t от a_{t+s}. Обозначим событие $\Omega_t=\{\eta_t+\eta_{t+\tau}>0\}$. Тогда по формуле полного матожидания $E\{*\}=P(\Omega_t)E\{*|\Omega_t\}+$ $+P(\Omega_{t+s})E\{*|\Omega_{t+s}\}+P(\overline{\Omega}_t\overline{\Omega}_{t+s})E\{*|\overline{\Omega}_t\overline{\Omega}_{t+s}\}-P(\Omega_t\,\Omega_{t+s})E\{*|\Omega_t\Omega_{t+s}\}$, где $\overline{A}$ значит дополнение A. Вместо $*$ подставим a_ta_{t+s}: $E\{a_ta_{t+s}|\Omega_t\}=E\{a_t|\Omega_t\}E\{a_{t+s}|\Omega_t\}=0=E\{a_ta_{t+s}|\Omega_{t+s}\}=$ $=E\{a_ta_{t+s}|\Omega_t\Omega_{t+s}\}$. Тогда $E\{a_ta_{t+s}\}=(1-\varepsilon)^4E\{a_ta_{t+s}|\overline{\Omega}_t\overline{\Omega}_{t+s}\}=\sigma(s)+E^2\{a_t\}$. Поскольку a_t и Ω_{t+s}

– независимы и $E\{a_t|\Omega_t\}=0$, имеем $E\{a_t\}=P(\overline{\Omega}_t)E\{a_t|\overline{\Omega}_t\}=P(\overline{\Omega}_t)E\{a_t|\overline{\Omega}_t\overline{\Omega}_{t+s}\}=P(\overline{\Omega}_{t+s})$ $E\{a_{t+s}|\overline{\Omega}_t\overline{\Omega}_{t+s}\}$. Получаем $\sigma(s)=(1-\varepsilon)^4 cov\{a_t,a_{t+s}|\overline{\Omega}_t\overline{\Omega}_{t+s})=(1-\varepsilon)^4\rho(s)$.

При $s=0$ имеем $\sigma(0)=D\{a_t\}=1-E^2\{a_t\}=1-(1-\varepsilon)^4E^2\{u_t\}=1+(1-\varepsilon)^4(\rho(0)-1)$. Наконец $\sigma(\tau)=E\{a_ta_{t+\tau}\}-E^2\{a_t\}=E\{sign(z_tz_{t+2\tau})\}-(1-\varepsilon)^4E^2\{u_t\}=(1-\varepsilon)^2E\{u_tu_{t+\tau}\}-(1-\varepsilon)^4E^2\{u_t\}=((1-\varepsilon)^2-(1-\varepsilon)^4)E\{u_tu_{t+\tau}\}+(1-\varepsilon)^4\rho(\tau)$. Тогда $\sigma(0)+2\sigma(\tau)-(\rho(0)+2\rho(\tau))(1-\varepsilon)^4<1+2(1-\varepsilon)^2-3(1-\varepsilon)^4=\varepsilon(2-\varepsilon)(1++3(1-\varepsilon)^2)$. Учитывая, что $\sigma(s)=(1-\varepsilon)^4\rho(s)$ для $s\neq 0,\pm\tau$, получаем требуемое неравенство для сумм рядов. ■

Теорема 2. Если имеет место модель (1)-(4) и $N_{T\tau}\underset{T\to\infty}{\to}+\infty$, то S – оценка (13) асимптотически нормально распределена:

$$\sqrt{N_{T\tau}}\cdot\left(\widehat{\theta}_\tau-\theta_\tau\right)\xrightarrow[T\to\infty]{D}N_1\left(0,\Sigma^\sigma\cdot\pi^2(1-\theta_\tau^2)\big/4(1-\varepsilon)^4\right).$$

Доказательство. Для простоты будем считать без потери общности, что выбросы отсутствуют и $N_{T\tau}=T-\tau$. Пусть снова $a_t=sign(z_tz_{t+\tau})$. Обозначим $S_\tau=(a_1+\ldots+a_{T-\tau})/(T-\tau)=E\{a_1\}+\eta$. Согласно Теореме 5.19 [13], если существуют такие $C>0$, $C_0>0$, $K>0$, $0<\delta<1$, $\beta>1$, что $\max\limits_{k\le n}E\,|\,a_k-Ea_k\,|^{2+\delta}\le C, T\cdot D\{\eta\}\ge C_0$ и коэффициент перемешивания $\alpha_a(n)\le Kn^{-\beta(2+\delta)/\delta}$, то η имеет асимптотическое нормальное распределение при $T\to\infty$. Первое условие, очевидно, выполняется. Второе следует из того, что $D\{\eta\}\underset{T\to\infty}{\sim}T^{-1}\cdot\Sigma^\sigma$. Теперь из доказательства Теоремы 1 $\alpha_y(n)=O(e^{-bn})$, откуда по Лемме 4 $\alpha_z(n)=O(e^{-bn})$ и по Лемме 3 $\alpha_a(n)=O(e^{-bn})\le Kn^{-\beta(2+\delta)/\delta}$ для некоторого K. Значит $\sqrt{T}\cdot\eta\xrightarrow{D}N_1\left(0,\Sigma^\sigma\right)$. Поскольку S – оценка θ_τ вычисляется как $\widehat{\theta}_\tau=Sin\left(\pi\cdot S_\tau\big/2(1-\varepsilon)^2\right)$, то из теоремы о функциональном преобразовании асимптотически нормальной случайной величины получаем

$$\sqrt{T}\cdot\left(\widehat{\theta}_\tau-\theta_\tau\right)\xrightarrow{D}N_1\left(0,Cos^2\left(\arcsin\left(\theta_\tau\right)\right)\cdot\Sigma^\sigma\cdot\pi^2\big/4(1-\varepsilon)^4\right),$$

что эквивалентно утверждению теоремы. ■

3. Робастные оценки коэффициентов авторегрессии и их свойства

Уравнения Юла-Уокера для модели (1) имеют вид:

$$\begin{cases}\sigma_0+b_1\sigma_1+\ldots+b_p\sigma_p=\sigma^2,\\ \sigma_{-1}+b_1\sigma_0+\ldots+b_p\sigma_{p-1}=0,\\ \ldots\\ \sigma_{-p}+b_1\sigma_{1-p}+\ldots+b_p\sigma_0=0.\end{cases}$$

где $\sigma_\tau=Cov\{y_t,y_{t+\tau}\}$. Заметим, что решение $\{b_i\}$ этой системы уравнений не зависит от σ, поэтому вместо ковариации σ_τ можно подставить коэффициент корреляции $Corr\{y_0,y_\tau\}=\sigma_\tau/\sigma_0$ и без потери общности полагать $\sigma=1$. Если вместо коэффициентов корреляции $\{\theta_\tau\}$ подставить их оценки $\{\widehat{\theta}_\tau\}$, то решением будет оценка $\widehat{b}=(\widehat{b}_i)$ вектора коэффициентов. Сформулируем алгоритм оценивания вектора $b=(b_i)$.

Алгоритм

Шаг 1) Если неизвестна доля выбросов, строим её оценку $\widehat{\varepsilon}$ (см. раздел 4).

Шаг 2) Согласно (13) строим S – оценки коэффициентов корреляции $\widehat{\theta}=(\widehat{\theta}_\tau)$, подставляя в них либо ε (если известно), либо $\widehat{\varepsilon}$.

Шаг 3) Строим оценку матрицы системы Юла-Уокера: $\widehat{\Xi}_{ij}=\widehat{\theta}_{i-j},\widehat{\theta}_0=1,i,j=\overline{0,p}$.

Шаг 4) Вычисляем обратную матрицу и оценки параметров авторегрессии:

$$\widehat{b}_i=\left(\widehat{\Xi}^{-1}\right)_{i0}\Big/\left(\widehat{\Xi}^{-1}\right)_{00},i=\overline{1,p}.$$

Шаг 5) Находим численным методом наибольший по модулю корень $\lambda_{\max}$ характеристического многочлена $\widehat{B}(\lambda)=\lambda^p+\widehat{b}_1\lambda^{p-1}+...+\widehat{b}_p$ и, если нарушено условие стационарности и $|\lambda_{\max}|\ge 1$, нормируем оценки следующим образом: $\widehat{b}_i \to \widehat{b}_i/(|\lambda_{\max}|+\Delta)^i$, где $\Delta=0.1/|\lambda_{\max}|$. Этим мы масштабируем корни $\widehat{B}(\lambda)$, поместив их в пределы единичного круга $|\lambda|<1$.

<u>Теорема 3.</u> Обозначим $\Xi(n)\in R^{n+1\times n+1}, \Xi_{ij}(n)=\theta_{i-j}$. Тогда, если $\min\limits_{0\le\tau\le p} N_{T\tau}\to+\infty$, имеет место следующая сходимость

$$P\left(\|\widehat{b}-b\|\le K\cdot\|\widehat{\theta}-\theta\|\right)\xrightarrow[T\to\infty]{}1,\ K=\left(1+\sqrt{\frac{2p}{\pi}\cdot\int\limits_{-\pi}^{\pi}|B(e^{ix})|^2\,dx}\right)\Big/\lambda_{\min}\left(\Xi(p-1)\right),\quad(14)$$

где $B(\lambda)=\lambda^p+b_1\lambda^{p-1}+...+b_p$ - характеристический многочлен для (1).

<u>Доказательство.</u> Согласно приведенному выше алгоритму, $\widehat{b}=\Psi\left(\widehat{\theta}\right)$, $\Psi: R^p\to R^p$ Ψ ограничена. Поскольку $\widehat{\theta}\xrightarrow[T\to\infty]{P}\theta$ и Ψ дифференцируема в окрестности θ, получаем, что $\Delta b=\widehat{b}-b=\dfrac{\partial\Psi}{\partial\theta}\cdot\Delta\theta+o(\|\Delta\theta\|)$ и $\|\Delta b\|\le\|\dfrac{\partial\Psi}{\partial\theta}\|\cdot\|\Delta\theta\|+o(\|\Delta\theta\|)$. Отсюда следует, что $\forall\delta>0,\ \mathrm{P}\left(\|\Delta b\|\le\left(\|\dfrac{\partial\Psi}{\partial\theta}\|+\delta\right)\cdot\|\Delta\theta\|\right)\ge\mathrm{P}\left(o(\|\Delta\theta\|)\le\delta\|\Delta\theta\|\right)\underset{T\to\infty}{\to}1$.

Докажем теперь, что $\|\dfrac{\partial\Psi}{\partial\theta}\|\le K$, где константа *K* определяется (14). Обозначим $\widetilde{b}=(1,b_1,...,b_p)$ и $e_0\in R^{p+1}$ - первый столбец единичной матрицы. Тогда из системы Юла-Уокера имеем: $\Xi(p)\cdot\widetilde{b}=f(\theta)\cdot e_0$. Дифференцируя по θ_i, получаем: $\partial\Xi(p)/\partial\theta_i\cdot\widetilde{b}+\Xi(p)\cdot$ $\cdot\partial\widetilde{b}/\partial\theta_i=\partial f(\theta)/\partial\theta_i\cdot e_0$. Это матричное равенство является системой из *p+1* линейных уравнений. Отбросив первое из них, имеем: $\partial\theta/\partial\theta_i+\partial\Xi(p-1)/\partial\theta_i\cdot b+\Xi(p-1)\cdot\partial b/\partial\theta_i=$ $=0_p$. Собрав эти уравнения для $i=1,...,p$, получим: $I_p+\left(B^{(1)}+B^{(2)}\right)+\Xi(p-1)\cdot\partial\Psi(\theta)/\partial\theta$, где матрицы $B^{(j)}$ состоят из нулей и коэффициентов $\{b_i\}$. Значит $\|\partial\Psi(\theta)/\partial\theta\|\le\|I_p+$ $+\left(B^{(1)}+B^{(2)}\right)\|\cdot\|\Xi^{-1}(p-1)\|$. Далее, как известно [14], $\|\Xi^{-1}(p-1)\|=\lambda_{\min}^{-1}(\Xi(p-1))$, $\|B^{(j)}\|\le$ $\le\sqrt{p}\cdot\max\limits_{1\le i\le p}\|B_i^{(j)}\|\le\sqrt{p}\cdot\sqrt{1+b_1^2+...+b_p^2}$. Выражение под корнем равно коэффициенту при нулевой степени в разложении $B(\lambda)B(\lambda^{-1})$ в ряд по λ. Отсюда $\|I_p+B^{(1)}+B^{(2)}\|\le 1+2\sqrt{p}\cdot$ $\cdot(2\pi)^{-0.5}\left(\int\limits_{-\pi}^{\pi}|B(e^{ix})|^2\,dx\right)^{0.5}$. Умножаем правую часть на $\|\Xi^{-1}(p-1)\|$ и получаем выражение для *K*, входящее в (14). ■

4. Оценки доли выбросов ε и их свойства

S – оценка (13) использует долю выбросов ε как априорную информацию о модели. Это ограничивает ее применение на практике, т.к. такая информация редко бывает заранее известной.

Укажем способ статистического оценивания ε. Будем дополнительно предполагать, что выбросы $\{\upsilon_t\}$ имеют гауссовское распределение: $L\{\upsilon_t\}=N_1(0,\beta_1\sigma^2)$, где $\beta_1>0$ - коэффициент, определяющий уровень выбросов. Тогда распределение наблюдений $\{z_t\}$ является смесью нормальных распределений с нулевыми средними и дисперсиями $\sigma_1^2=D\{\upsilon_t\}=\sigma^2\beta_1$ и $\sigma_2^2=D\{y_t\}=\sigma^2\beta_2$, где $\beta_2=\dfrac{1}{2\pi}\int\limits_{-\pi}^{\pi}|B(e^{ix})|^{-2}\,dx$. Тогда соответствующая характеристическая функция имеет вид: $f_z(\lambda)=\varepsilon\exp\{-0.5\sigma_1^2\lambda^2\}+(1-\varepsilon)\exp\{-0.5\cdot$

$\cdot\sigma_2^2\lambda^2\}$. Если $\beta_1 >> \beta_2$, то при достаточно больших значениях λ можно использовать приближение

$$f_z(\lambda) \approx (1-\varepsilon)\exp\{-0.5\sigma_2^2\lambda^2\}. \qquad (15)$$

Это приближение тем точнее, чем больше уровень выбросов β_1. Далее, используя (15), имеем приближение: $f_z(\lambda_2)/f_z(\lambda_1) = \exp\{-0.5\sigma_2^2(\lambda_2^2-\lambda_1^2)\}$. Возьмем некоторые конкретные значения λ_1 и λ_2. Тогда в левой части соотношения можно статистически оценить отдельно числитель и знаменатель, и затем из правой части получить оценку $\hat{\sigma}_2$. Подставив ее вместо σ_2 в формулу (15), получим равенство, из которого и найдем оценку $\hat{\varepsilon}$. Таким образом, алгоритм оценивания ε имеет вид:

1) выбираем λ_1 и λ_2;
2) строим выборочную характеристическую функцию и вычисляем
$$\hat{f}_z(\lambda_i) = \sum_{t=1}^{T} o_t \cos\{\lambda_i z_t\} \Big/ \sum_{t=1}^{T} o_t \ ;$$
3) строим оценку σ_2^2: $\hat{\sigma}_2^2 = 2(\lambda_1^2-\lambda_2^2)^{-1}\ln \hat{f}_z(\lambda_2)/\hat{f}_z(\lambda_1)$;
4) в формуле (15) делаем подстановки: $\sigma_2 \to \hat{\sigma}_2, \lambda \to \hat{\sigma}_2^{-1}$ и вычисляем
$$\hat{\varepsilon} = 1-\sqrt{e}\cdot \hat{f}_z(\hat{\sigma}_2^{-1}),\ \text{где}\ \hat{f}_z(\hat{\sigma}_2^{-1}) = \sum_{t=1}^{T} o_t \cos(z_t/\hat{\sigma}_2) \Big/ \sum_{t=1}^{T} o_t .$$

В этом алгоритме остается одна неопределенность – выбор λ_1 и λ_2. В численных экспериментах хорошие результаты показал алгоритм со следующим подбором этих параметров: $\lambda_1^2 = \sum_{t=1}^{T} o_t \Big/ \sum_{t=1}^{T} o_t z_t^2 \approx (\varepsilon\sigma_1^2 + (1-\varepsilon)\sigma_2^2)^{-1}$, $\lambda_2 = 2\lambda_1$.

5. Численные результаты

Для двух моделей авторегрессии M_1, M_2 с выбросами с параметрами $T=500$, $\varepsilon=0.1$, $\beta=10$, $\sigma=1$ и порождающими уравнениями:

M_1: $y_t + 0.3y_{t-1} + 0.2y_{t-2} = u_t, p = 2$,

M_2: $y_t - 0.2y_{t-1} + 0.1y_{t-2} - 0.4y_{t-4} = u_t, p = 4$,

были смоделированы серии по ***10000*** реализаций. Для каждой реализации коэффициенты авторегрессии b оценивались шестью разными процедурами: робастных на основе t-оценки и s-оценки (отдельно с известным и не известным ε), процедурой, использующей оценки Хьюбера (6) и классической МНК-оценкой [1]. В таблицах 1 - 4 приведены численные результаты для оценок $\{\hat{\theta}_\tau\},\{\hat{b}_i\}$ (*Sm* – смещение, *Var* – вариация). Гистограммы для абсолютных уклонений оценок $\delta = \max_i |b_i - \hat{b}_i|$ представлены на рис.1 и расположены сверху вниз в том же порядке, что и перечисленные выше процедуры оценивания.

Таблица 1, Модель M_1

Процедура	*Var(θ1)*	*Sm(θ1)*	*Var(θ2)*	*Sm(θ2)*
роб. t-оц. ε - изв.	0.013	0.005	0.015	0.007
роб. t-оц. ε - неизв.	0.014	0.001	0.017	-0.002
роб. s-оц. ε - изв.	0.017	0.007	0.023	0.009
роб. s-оц. ε - неизв.	0.023	0.004	0.019	0.004
Хьюбера	0.029	0.012	0.031	0.017
МНК	0.197	-0.059	0.092	0.132

Таблица 2, Модель M_1

Процедура	*Var(b1)*	*E(b1)*	*Var(b2)*	*E(b2)*
роб. t-оц. ε - изв.	0.007	0.298	0.009	0.203
роб. t-оц. ε - неизв.	0.009	0.305	0.008	0.201
роб. s-оц. ε - изв.	0.011	0.301	0.013	0.207

роб. s-оц. ε - неизв.	0.015	0.312	0.012	0.217
Хьюбера	0.014	0.284	0.015	0.188
МНК	0.097	-0.011	0.058	0.007

Таблица 3, Модель M_2

Процедура	*Var(θ1)*	*Sm(θ1)*	*Var(θ2)*	*Sm(θ2)*	*Var(θ3)*	*Sm(θ3)*	*Var(θ4)*	*Sm(θ4)*
роб. t-оц. ε - изв.	0.011	0.004	0.014	0.004	0.009	-0.003	0.014	0.007
роб. t-оц. ε - неизв.	0.010	0.006	0.016	-0.002	0.010	0.004	0.015	0.003
роб. s-оц. ε - изв.	0.019	-0.009	0.026	0.004	0.017	0.007	0.029	-0.009
роб. s-оц. ε - неизв.	0.017	-0.007	0.021	0.003	0.014	0.003	0.022	-0.008
Хьюбера	0.027	-0.014	0.031	0.008	0.022	0.012	0.034	-0.013
МНК	0.073	0.071	0.085	0.143	0.068	-0.052	0.273	0.039

Таблица 4, Модель M_2

Процедура	*Var(b1)*	*E(b1)*	*Var(b2)*	*E(b2)*	*Var(b3)*	*E(b3)*	*Var(b4)*	*E(b4)*
роб. t-оц. ε - изв.	0.006	-0.202	0.004	0.109	0.006	0.003	0.009	-0.412
роб. t-оц. ε - неизв.	0.005	-0.206	0.003	0.105	0.004	0.007	0.008	-0.403
роб. s-оц. ε - изв.	0.013	-0.209	0.011	0.115	0.010	0.008	0.017	-0.384
роб. s-оц. ε - неизв.	0.012	-0.210	0.013	0.114	0.009	0.004	0.015	-0.390
Хьюбера	0.016	-0.189	0.017	0.121	0.013	0.003	0.021	-0.426
МНК	0.042	0.053	0.023	0.014	0.053	-0.02	0.191	0.008

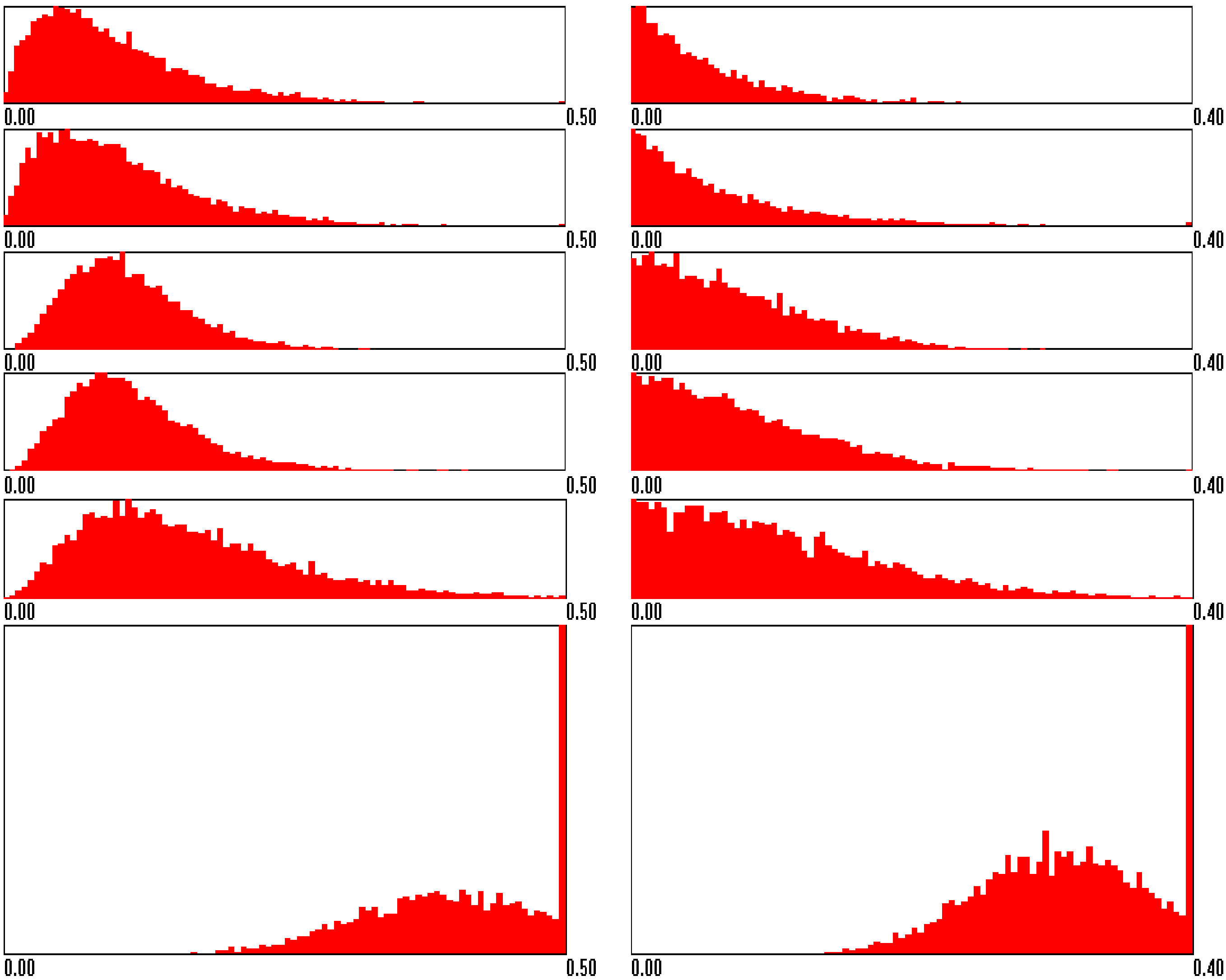

Рис.1. Гистограммы величин δ для моделей M_1(справа), M_2(слева)

Литература

1. Андерсон Т.В. Статистический анализ временных рядов. М.: Мир, 1976.
2. П. Хьюбер. Робастность в статистике. М.: Мир, 1984.
3. Хампель Ф. Робастность в статистике: подход на основе функций влияния. М.: Мир, 1989.
4. Мостеллер Ф., Тьюки Дж. Анализ данных и регрессия. Москва: Статистика, 1982.
5. Kharin Yu. Robustness in statistical pattern recognition. Dordrecht-Boston-London: Kluwer Academic Publishers, 1996.
6. Shevlyakov G.L., Vilchevski N.O. Robustness Data Analysis: criteria and methods. N.Y.: VSP, 2002.
7. Андерсон Т.В. Введение в многомерный статистический анализ данных. М.: ФМЛ, 1963.
8. Вадзинский Р.Н. Справочник по вероятностным распределениям. М.: Наука, 2001.
9. Bradley R.C. Introducion to strong mixing conditions. N.Y.: Kendrick Press, 2007.
10. Doukhan P. Mixing: properties and examples. N.Y.: Springer-Verlag, 1994.
11. Ширяев А.Н. Вероятность. М.: Наука, 1980.
12. Lin Zh., Lu Ch. Limit theory for mixing dependent random variables. Dordrecht-Boston-London: Kluwer Academic Publishers, 1996.
13. Харин Ю.С., Зуев Н.М. Теория вероятностей. Минск: БГУ, 2004.
14. Гантмахер Ф.Р. Теория матриц. М.: Наука, 1966.
15. Королюк В.С. Справочник по теории вероятностей и математической статистике. М.: Наука, 1985.