*УДК 519.2*

*Д.В. ЗЕНЕВИЧ*

# ОБ УСТОЙЧИВОСТИ ПРОГНОЗИРОВАНИЯ АВТОРЕГРЕССИОННЫХ ВРЕМЕННЫХ РЯДОВ ПРИ НАЛИЧИИ ОШИБОК В ЗАДАНИИ ПАРАМЕТРОВ

This paper is devoted to evaluation of robustness of forecasting by autoregressive model in the case of misspecification of model parameters. Dependence of mean square risk on specification error is evaluated.

Прогнозирование временных рядов с использованием модели авторегрессии широко применяется в экономических, технических, медицинских и других приложениях. Классический вариант данной модели исследован достаточно полно [1]. На практике редко выполняются классические модельные предположения – модель искажается из-за наличия выбросов, пропусков, аддитивных искажений, неточного определения параметров модели [2]. В данной работе исследуется влияние последнего фактора на устойчивость традиционных авторегрессионных прогнозов.

## 1. Математическая модель и постановка задачи

Пусть наблюдается временной ряд $\{x_t\}, t=1,\dots,T$, описываемый моделью авторегрессии $AR(p)$:

$$x_t = \mathbf{X}'_{t-1}\theta^0 + \xi_t, \quad t=1,\dots,T, \tag{1}$$

где $\theta^0 \in \mathbb{R}^p$ – неизвестное истинное значение вектора коэффициентов авторегрессии, $p$ – порядок авторегрессии, $\mathbf{X}_{t-1} = (x_{t-1},\dots,x_{t-p})' \in \mathbb{R}^p$, «′» – знак транспонирования, $\{\xi_t\} \sim N(0,\sigma^2)$ – независимые одинаково распределенные случайные величины, начальные значения $\mathbf{X}_0 = (x_0,\dots,x_{1-p})' \in \mathbb{R}^p$ заданы.

На практике гипотетическая модель (1) обычно неизвестна, поэтому для прогнозирования этого ряда на глубину $\tau \geq 1$ используют рекуррентную процедуру [1]:

$$\hat{x}_{T+j} = \hat{\mathbf{X}}'_{T+j-1}\theta, \quad j=\overline{1,\tau}, \tag{2}$$

где $\hat{x}_{T+j}$ – прогноз в момент времени $T+j$, $\theta \in \mathbb{R}^p$ – используемый при прогнозировании вектор коэффициентов, $\hat{\mathbf{X}}_{t-1} = (\hat{x}_{t-1},\dots,\hat{x}_{t-p})' \in \mathbb{R}^p$, $\hat{x}_t = x_t$ для $t < T$. При несовпадении векторов $\theta$ и $\theta^0$ имеем параметрические искажения. В связи с этим важно исследовать устойчивость прогнозов (2) в зависимости от величины параметрического искажения [2, 3]. Для характеристики устойчивости определим матрицу рисков прогнозирования $R(\theta^0,\theta;\tau) = \mathrm{E}\{(\hat{\mathbf{X}}_{T+\tau} - \mathbf{X}_{T+\tau})(\hat{\mathbf{X}}_{T+\tau} - \mathbf{X}_{T+\tau})'\}$ и ее (1,1)-элемент $r(\theta^0,\theta;\tau) = R_{11}(\theta^0,\theta;\tau) = \mathrm{E}\{(\hat{x}_{T+\tau} - x_{T+\tau})^2\}$. Отметим, что величину $r(\cdot)$ в литературе часто называют среднеквадратическим риском прогнозирования.

## 2. Влияние параметрического отклонения

Определим $p$-вектор $\mathbf{U}_t = (\xi_t,0,\dots,0) \in \mathbb{R}^p$ и $(p \times p)$-матрицы:

$$B_0 = \begin{pmatrix} \theta_1^0 & \cdots & \theta_{p-1}^0 & \theta_p^0 \\ 1 & \cdots & 0 & 0 \\ \vdots & & \vdots & \vdots \\ 0 & \cdots & 1 & 0 \end{pmatrix}, \quad B = \begin{pmatrix} \theta_1 & \cdots & \theta_{p-1} & \theta_p \\ 1 & \cdots & 0 & 0 \\ \vdots & & \vdots & \vdots \\ 0 & \cdots & 1 & 0 \end{pmatrix}.$$

**Теорема 1**. Пусть наблюдается временной ряд, удовлетворяющий авторегрессионной модели (1), и для прогнозирования его значений применяется процедура (2). Тогда матрица рисков прогнозирования имеет вид

$$R(\boldsymbol{\theta}^0,\boldsymbol{\theta};\tau)=F_\tau+(B^\tau-B_0^\tau)C_T(B^\tau-B_0^\tau)', \tag{3}$$

где $D=\mathrm{diag}(\sigma^2,0,\ldots,0)$, $F_k=\sigma^2\sum_{l=0}^{k-1}B_0^l D(B_0^l)'$, $C_k=F_k+B_0^k\mathbf{X}_0(B_0^k\mathbf{X}_0)'$.

Доказательство. Утверждение теоремы основано на соотношениях $\mathbf{X}_t=\sum_{i=0}^{t-1}B_0^i\mathbf{U}_{t-i}+B_0^t\mathbf{X}_0$, $\hat{\mathbf{X}}_{T+\tau}=B^\tau\mathbf{X}_T$ и определении риска. Теорема доказана.

*Следствие* 1. Минимальный по $\boldsymbol{\theta}$ риск достигается при $\boldsymbol{\theta}=\boldsymbol{\theta}^0$:

$$r_{\min}(\boldsymbol{\theta}^0;\tau)=r(\boldsymbol{\theta}^0,\boldsymbol{\theta}^0;\tau).$$

Доказательство. Матрицы $D$, $F_k$, $C_k$ неотрицательно определены, значит, $\left((B^\tau-B_0^\tau)C_T(B^\tau-B_0^\tau)'\right)_{11}\geq 0$. Тогда из теоремы 1 следует доказываемое.

Определим $(p\times p)$-матрицу $\beta(\tau)=(\sum_{i=0}^{\tau-1}(B_0^i)_{11}B_0^{\tau-1-i})\times(\sum_{j=0}^{\tau-1}(B_0^j)_{11}B_0^{\tau-1-j})'$, параметрическое искажение $\alpha=\theta-\theta^0\,\mathbb{R}^p$, уровень искажений $\varepsilon=|\boldsymbol{\theta}-\boldsymbol{\theta}^0|>0$, $\mathbf{1}_{p\times p}$ – $(p\times p)$-матрица, все элементы которой равны единице.

**Теорема 2**. В условиях теоремы 1 при $\varepsilon\to 0$ риск прогнозирования удовлетворяет следующему асимптотическому разложению:

$$r(\boldsymbol{\theta}^0,\boldsymbol{\theta};\tau)=r_{\min}(\boldsymbol{\theta}^0;\tau)+\alpha'\beta(\tau)\alpha+\mathrm{O}(\varepsilon^3). \tag{4}$$

Доказательство. Определим $(p\times p)$-матрицу $\Delta=(\alpha,\mathbf{0}_p,\ldots,\mathbf{0}_p)'$, где $\mathbf{0}_p$ – нулевой $p$-вектор. Используя равенство $B=B_0+\Delta$, имеем:

$$R(\boldsymbol{\theta}^0,\boldsymbol{\theta};\tau)=F_\tau+((B_0+\Delta)^\tau-B_0^\tau)C_T((B_0+\Delta)^\tau-B_0^\tau)'=$$

$$=R(\boldsymbol{\theta}^0,\boldsymbol{\theta};\tau)+\sum_{i,j=0}^{\tau-1}B_0^i\Delta B_0^{\tau-1-i}C_T(B_0^{\tau-1-j})'\Delta'(B_0^j)'+\mathrm{O}(\varepsilon^3)\mathbf{1}_{p\times p}.$$

Пусть $Q^{(ij)}=B_0^{\tau-1-i}C_T(B_0^{\tau-1-j})'$ при $i,j=\overline{0,\tau-1}$. Тогда

$$B_0^i\Delta Q^{(ij)}\Delta'(B_0^j)'=B_0^i\begin{pmatrix}\sum_{k=1}^{p}(\sum_{l=1}^{p}q_{lk}^{(ij)}\alpha_l)\alpha_k & 0 & \cdots & 0\\ 0 & 0 & \cdots & 0\\ \vdots & \vdots & & \vdots\\ 0 & 0 & \cdots & 0\end{pmatrix}_{p\times p}(B_0^j)'.$$

Следовательно,

$$r(\boldsymbol{\theta}^0,\boldsymbol{\theta};\tau)=(F_\tau)_{11}+\sum_{i,j=0}^{\tau-1}(\sum_{k,l=1}^{p}q_{lk}^{(ij)}(B_0^i)_{11}(B_0^j)_{11}\alpha_l\alpha_k)+\mathrm{O}(\varepsilon^3). \tag{5}$$

Из (5) и равенства $\sum_{i,j=0}^{\tau-1}q_{lk}^{(ij)}(B_0^i)_{11}(B_0^j)_{11}=\beta_{lk}(\tau)$ получаем (4). Теорема доказана.

Определим гарантированный риск прогнозирования при известном уровне искажений $\varepsilon>0$: $r_+(\boldsymbol{\theta}^0;\tau;\varepsilon)=\sup_{|\alpha|<\varepsilon}r(\boldsymbol{\theta}^0,\boldsymbol{\theta}^0+\alpha;\tau)$. Будем говорить, что для заданного уровня приращения риска $\delta\geq 0$ параметрическое отклонение $\boldsymbol{\theta}-\boldsymbol{\theta}^0$ является $\delta$-допустимым, если $r(\boldsymbol{\theta}^0,\boldsymbol{\theta};\tau)<(1+\delta)r_{\min}(\boldsymbol{\theta}^0;\tau)$.

Вследствие симметричности и неотрицательной определенности $\beta(\tau)$ существует ортогональная матрица $V$ такая, что $\beta(\tau)=V'\Lambda V$, где $\Lambda=\mathrm{diag}(\lambda_1,\ldots,\lambda_p)$, $\lambda_i>0$, $i=\overline{1,p}$ – собственные значения матрицы $\beta(\tau)$, $\lambda_{\max}$ – максимальное из них.

*Следствие* 2. В условиях теоремы 1 для любого $\delta>0$ и уровня искажений $\varepsilon\to 0$ верны утверждения: 1) $r_+(\theta^0;\tau;\varepsilon)=r_{\min}(\theta^0;\tau)+\lambda_{\max}\varepsilon^2+O(\varepsilon^2)$; 2) при использовании главного члена разложения (4) множество $\delta$-допустимых отклонений $\alpha$ образуют эллипсоид в $\mathbb{R}^p$: $\sum_{i=1}^{p}\lambda_i\left((V\alpha)_i\right)^2\leq$ $\leq r_{\min}(\theta^0;\tau)$.

Д о к а з а т е л ь с т в о. Отбросив остаток в формуле (4), имеем:

$$r(\theta^0,\theta;\tau)=r_{\min}(\theta^0;\tau)+(V\alpha)'\Lambda V\alpha=r_{\min}(\theta^0;\tau)+\sum_{i=1}^{p}\lambda_i\left((V\alpha)_i\right)^2. \quad (6)$$

Тогда из неравенства $\sum_{i=1}^{p}\lambda_i\left((V\alpha)_i\right)^2\leq\lambda_{\max}\varepsilon^2$ следует первое утверждение. Второе утверждение следует из (6). Следствие доказано.

**3. Безусловный риск прогнозирования**

Предположим, что параметрическое искажение $\alpha=\theta-\theta^0$ есть случайный вектор. Такая ситуация может возникнуть, например, когда вектор $\theta^0$ оценивается по статистическим данным со случайной погрешностью $\alpha$. Тогда имеет смысл исследовать безусловный риск прогнозирования: $r(\theta^0;\tau)=\mathrm{E}\{r(\theta^0,\theta;\tau)\}$, где усреднение проводится по распределению параметра $\theta$.

**Теорема 3**. Пусть параметрическое искажение $\alpha$ есть случайный вектор с ковариационной матрицей $\Sigma$ и ограниченными моментами 3-го порядка, причем $\varepsilon_+=\left(\max_{i,j,k}\mathbf{E}\{|\alpha_i\alpha_j\alpha_k|\}\right)^{1/3}\to 0$. Тогда безусловный риск прогнозирования удовлетворяет асимптотическому разложению:

$$r(\theta^0;\tau)=r_{\min}(\theta^0;\tau)+tr(\beta(\tau)\Sigma)+O(\varepsilon_+^3).$$

Д о к а з а т е л ь с т в о. Согласно теореме 1 при фиксированных $\theta^0$, $\mathbf{X}_0$, $\sigma^2$, $\tau$ риск есть полином конечной степени от $\theta_i$, $i=\overline{1,p}$. Следовательно, все производные от функции риска существуют и ограничены в конечных областях, т. е. справедливо применение формулы Тейлора:

$$r(\theta,\theta^0;\tau)=r_{\min}(\theta^0;\tau)+\alpha'\beta(\tau)\alpha+\sum_{i,j,l=1}^{p}k_{ijl}r'''(\theta^0,\theta^0+\varsigma_{ijl};\tau)\alpha_i\alpha_j\alpha_l,$$

где $p$-векторы $\{\varsigma_{ijl}\}$ лежат в $p$-мерном параллелепипеде $[0,\alpha_1]\times\ldots\times[0,\alpha_p]$ и справедливо $|\eta(\alpha)|=\left|\sum_{i,j,l=1}^{p}k_{ijl}r'''(\theta^0,\theta^0+\varsigma_{ijl};\tau)\alpha_i\alpha_j\alpha_l\right|\leq\sum_{i,j,l=1}^{p}K_1K_2\left|\alpha_i\alpha_j\alpha_l\right|$, где $K_1$, $K_2$ – константы такие, что $K_1=\max_{i,j,l}|k_{ijl}|$, $|r'''(\cdot)|\leq K_2<\infty$. Осталось заметить, что $|\mathrm{E}\{\eta(\alpha)\}|\leq K_1K_2\sum_{i,j,l=1}^{p}\mathrm{E}\{|\alpha_i\alpha_j\alpha_l|\}\leq K_1K_2p^3\max_{i,j,l}\mathrm{E}\{|\alpha_i\alpha_j\alpha_l|\}=$ $=O(\varepsilon_+^3)$ и $\mathrm{E}\{\alpha'\beta(\tau)\alpha\}=tr(\beta(\tau)\Sigma)$. Теорема доказана.

Пусть по некоторой реализации $\{x_t\}$ временного ряда (1) длины $T_0$ построена ОМП-оценка вектора $\theta^0$, а по независимой от $\{x_t\}$ реализации $\{\tilde{x}_t\}$ ведется прогнозирование. При этом если для любого заданного $\delta>0$ справедливо неравенство $r(\theta^0;\tau)<(1+\delta)\, r_{\min}(\theta^0;\tau)$, то величину $T_0 = T_0(\delta)$ назовем $\delta$-допустимой длительностью наблюдения. Обозначим $F = \sum_{i=0}^{\infty} B_0 D(B_0)'$.

*Следствие* 3. Пусть по некоторой реализации $\{x_t\}$ временного ряда (1) длины $T_0$ построена ОМП-оценка вектора $\boldsymbol{\theta}^0$, а прогнозирование ведется по независимой от $\{x_t\}$ реализации $\{\bar{x}_t\}$. Если собственные значения матрицы $B_0$ находятся в единичном круге, то справедливо асимптотическое разложение безусловного риска прогнозирования:

$$r(\theta^0;\tau) = r_{\min}(\theta^0;\tau) + tr(\beta(\tau)F^{-1})/T_0 + O(T_0^{-3/2}).$$

При этом с точностью до $O(T_0^{-3/2})$ минимальная $\delta$-допустимая длительность наблюдения составляет $T_0(\delta) = tr(\beta(\tau)F^{-1})/(\delta \cdot r_{\min}(\theta^0;\tau))$.

Доказательство. Этот результат есть следствие теоремы 5.5.1 из [1] и теоремы 3. Следствие доказано.

1. Андерсон Т. Статистический анализ временных рядов. М., 1976.
2. Kharin Yu.S., Zenevich D.V. // Computer data analysis and modeling. Minsk, 1998. Vol. 1. P. 120.
3. Dahlhaus R., Wefelmeyer W. // Ann. Stat. 1996. Vol. 24. № 3. P. 952.

Поступила в редакцию 24.03.2000.

***Зеневич Дмитрий Владимирович*** – аспирант кафедры математического моделирования и анализа данных. Научный руководитель доктор физико-математических наук, профессор Ю.С. Харин.

*УДК 519.24*

*Т.В. СОБОЛЕВА*

## ВЫЧИСЛЕНИЕ ПЕРВЫХ ДВУХ МОМЕНТОВ РАСШИРЕННОЙ ПЕРИОДОГРАММЫ ДИСКРЕТНОГО ДЕЙСТВИТЕЛЬНОГО УСТОЙЧИВОГО СТАЦИОНАРНОГО СЛУЧАЙНОГО ПРОЦЕССА

In this article an enlarged periodogram of the discrete real stable stationary random process is being considered, mathematical expectancy and dispersion being calculated for it.

Рассмотрим дискретный действительный устойчивый стационарный случайный процесс $X(t)$, $t\in Z$, с показателем $\alpha$, $0<\alpha<2$, спектральное представление которого имеет следующий вид:

$$X(t) = \int_{-\pi}^{\pi} \cos(\lambda t)\, d\xi(\lambda), \tag{1}$$

где $\xi(\lambda)$ – действительный устойчивый с независимыми приращениями процесс с показателем $\alpha$, $0<\alpha<2$, такой, что

$$\left\{E\left|d\xi(\lambda)\right|^p\right\}^{\alpha/p} = C(p,\alpha)\varphi(\lambda)d\lambda \tag{2}$$

для всех $0<p<\alpha$, где $C(p,\alpha)$ – зависит от $\alpha$, $p$ и не зависит от $\xi(\lambda)$, а $\varphi(\lambda)$, $\lambda\in[-\pi,\pi]$ – неотрицательная функция, которую по аналогии с [1, 2] будем называть спектральной плотностью процесса $X(t)$, $t\in Z$.

По наблюдениям за процессом $X(t)$, $t\in\overline{[-T,T]}$, введем в рассмотрение статистику, являющуюся аналогом расширенного преобразования Фурье для обычных процессов (см. [4]):