УДК 519.2

Член-корреспондент Ю.С. ХАРИН, В.А. ВОЛОШКО

ПРОГНОЗИРОВАНИЕ СТАЦИОНАРНЫХ ВРЕМЕННЫХ РЯДОВ НА ОСНОВЕ МАЛОПАРАМЕТРИЧЕСКОЙ МОДЕЛИ БЛУМФИЛДА

НИИ прикладных проблем математики и информатики БГУ, Минск *Поступило*

Введение. Точность статистических выводов (оценок, решений, прогнозов) при параметрическом анализе данных, как известно [1,2], зависит от соотношения числа p параметров модели и длительности наблюдения T: $\rho(p,T) = p/T$. Если $p << T$, то есть $\rho(p,T) \to 0$ при $T \to \infty$, то статистические выводы на основе классических методов (МНК, максимального правдоподобия, байесовского метода) оказываются состоятельными и дают приемлемую точность для практики. Если же число параметров модели p сравнимо с длительностью T, то есть $\rho(p,T) \to c,\ c > 0$, то классические методы не применимы, и приемлемую точность можно достичь лишь для весьма специальных случаев [3]. В связи с этим актуальна задача разработки малопараметрических моделей, т.е. моделей с малым коэффициентом $\rho(p,T)$, и статистического анализа на их основе.

В [4] Блумфилдом предложена так называемая экспоненциальная модель *EXP*(*p*) порядка p для стационарных временных рядов и построены статистические оценки параметров этой модели. Данная статья посвящена использованию малопараметрической модели Блумфилда для решения задачи статистического прогнозирования.

Модель Блумфилда и ее связь с авторегрессионной моделью Вольда. Примем обозначения: $\Pi = [-\pi, \pi]$; i – мнимая единица, $e_n(\lambda) = e^{in\lambda} \in \mathrm{C}, \lambda \in \Pi, n \in \mathrm{N}$;

$$E_\Pi\{f\} = (2\pi)^{-1}\int_\Pi f(\lambda)d\lambda\,,\ D_\Pi\{f\} = E_\Pi\{|f|^2\} - |E_\Pi\{f\}|^2,\ f: \Pi \to \mathrm{C}.$$

Пусть наблюдается стационарный в широком смысле временной ряд $x_t, t \in \mathrm{Z}$, с нулевым средним $E\{x_t\} = 0$, имеющий ковариационную функцию σ_τ, корреляционную функцию θ_τ:

$$\sigma_\tau = Cov\{x_t, x_{t+\tau}\} = E\{x_t x_{t+\tau}\},\ \theta_\tau = \sigma_\tau/\sigma_0,\ \tau \in \mathrm{Z}, \tag{1}$$

а также спектральную плотность $S(\lambda)$:

$$S(\lambda) = \sum\nolimits_{\tau\in \mathrm{Z}} \sigma_\tau e_\tau(\lambda),\ \sigma_\tau = E_\Pi\{S(\lambda)e_\tau(\lambda)\},\ \tau \in \mathrm{Z};\ L(\lambda) = \ln S(\lambda),\ \lambda \in \Pi. \tag{2}$$

В дальнейшем для краткости будем называть $S(\lambda)$ и $L(\lambda)$ соответственно *спектром* и *лог-спектром*.

Будем предполагать, что выполняется условие регулярности ***R*1**: $E_\Pi\{S\} < +\infty$, $L(\cdot) \in L_2(\Pi)$. В частности, если $S(\cdot)$, $1/S(\cdot)$ не имеют особенностей на П, то условие ***R*1** выполняется. При выполнении ***R*1** любой временной ряд x_t допускает единственные представления в виде следующих параметрических моделей бесконечного порядка: экспоненциальной модели Блумфилда *EXP*(∞) [4]:

$$S(\lambda) = \exp\Big(\sum\nolimits_{n\in \mathrm{Z}} l_n e_n(\lambda)\Big)\,,\ l(z) = \sum\nolimits_{k\in \mathrm{N}} l_k z^k,\ l_0, l_1, \ldots \in \mathrm{R}, \tag{3}$$

и авторегрессионной модели Вольда *AR*(∞) [1,5]:

$$S(\lambda) = \frac{\sigma^2}{|\beta(e_1(\lambda))|^2},\ \sigma > 0,\ \beta(z) = 1 - \sum\nolimits_{n\in \mathrm{N}} b_n z^n,\ b_1, b_2, \ldots \in \mathrm{R}, \tag{4}$$

где $l(z), \beta(z), z \in \mathrm{C}$, – формальные степенные ряды, а σ – масштабный коэффициент. Из (4) и [1] следует авторегрессионное представление временного ряда $AR(\infty)$: $x_t = \sum_{n \in \mathrm{N}} b_n x_{t-n} +$ $+\xi_t$, $t \in \mathrm{Z}$, где $\{\xi_t\}$ – некоррелированные случайные величины, $E\{\xi_t\} = 0$, $D\{\xi_t\} = \sigma^2$; при этом $\{b_n\}$ называются авторегрессионными коэффициентами. Наряду с условием ***R*1** потребуем выполнения условия регулярности ***R*2** на порождающий процесс $\{\xi_t\}$: $E\{\xi_t \mid \xi_{t-1}, \xi_{t-2}, \ldots\} = 0$, $t \in \mathrm{Z}$. В частности, если $\{\xi_t\}$ – н.о.р.с.в, то ***R*2** выполняется.

Отметим, что в теории цифровой обработки сигналов [6] коэффициент $l_n = l_{-n} = E_{\Pi}\{Le_n\}$ ряда Фурье для лог-спектра принято называть *n*-м кепстральным коэффициентом, а $\beta(\cdot)$ – передаточной функцией спектра $S(\lambda)$. Взаимосвязь параметров моделей $AR(\infty)$ и $EXP(\infty)$ обеспечивают формулы Пурахмади [7]:

$$l_0 = \ln \sigma^2, \ \ \beta(z) = e^{-l(z)} \Leftrightarrow b_n = l_n - n^{-1} \sum_{j=1}^{n-1} j l_j b_{n-j}, \ \ n = 1, 2, \ldots. \tag{5}$$

Установим теперь взаимосвязь между моделями $AR(p)$, $EXP(q)$ в случае конечных p,q.

Лемма 1. *Модель авторегрессии $AR(p)$ порядка $p \in \mathrm{N}$ с некоторыми параметрами σ^2, $\{b_n \in \mathrm{R} : n = 1, 2, \ldots, p\}$ эквивалентна модели Блумфилда $EXP(\infty)$ с параметрами:*

$$l_0 = \ln \sigma^2, \ \ l_n = n^{-1} \sum_{j=1}^{J} r_j z_j^n, \ \ n = 1, 2, \ldots, \tag{6}$$

где $B(z) = z^p - \sum_{n=1}^{p} b_n z^{p-n} = \prod_{j=1}^{J} (z - z_j)^{r_j}, z_j \in \mathrm{C}, |z_j| < 1, r_j \in \mathrm{N}$, – факторизация характеристического полинома $B(z)$ модели $AR(p)$, z_j – j-й корень кратности r_j; J – число различных корней; $\sum_{j=1}^{J} r_j = p$.

Доказательство. Согласно (3)-(5) при $z=e^{i\lambda}$

$$\ln S(\lambda) - \ln \sigma^2 = -2 \ln |B(z)| = -2 \sum_{j=1}^{J} r_j \ln \left| 1 - \frac{z_j}{z} \right| = 2 \operatorname{Re} \left\{ \sum_{n=1}^{\infty} \frac{z^{-n}}{n} \left(\sum_{j=1}^{J} r_j z_j^n \right) \right\} = 2 \operatorname{Re}\{l(z)\},$$

откуда получаем (6). Лемма доказана.

Из Леммы 1 следует, что коэффициенты $n \cdot l_n$ являются аналогами степенных сумм характеристических корней временного ряда $AR(p)$.

Лемма 2. *Модель Блумфилда $EXP(p)$ порядка $p \in \mathrm{N}$ с единственным ненулевым коэффициентом $l_p = -c \in \mathrm{R}$ $(l_0 = l_1 = \ldots = l_{p-1} = 0)$ эквивалентна модели авторегрессии бесконечного порядка $AR(\infty)$ с параметрами:*

$$\sigma^2 = 1, \ \ b_n = \begin{cases} -c^k / k!, & \text{если } n = kp, \ \ k \in \mathrm{N}, \\ 0, & \text{в противном случае}, \end{cases} \quad n = 1, 2, \ldots. \tag{7}$$

Доказательство. Согласно (5) $\sigma^2=\exp(l_0)=1$ и

$$1 - \sum_{n \in \mathrm{N}} b_n z^n = e^{-l(z)} = \exp(c \cdot z^p) = 1 + \sum_{k \in \mathrm{N}} \frac{c^k}{k!} z^{kp}.$$

Приравнивая коэффициенты полиномов в левой и правой частях, получаем (7). Лемма доказана.

Метод прогнозирования на основе модели Блумфилда. Прогноз $\hat{x}_t$ по предыстории $\{x_s : s<t\}$, минимизирующий среднеквадратический риск $r = E\{(\hat{x}_t - x_t)^2\}$, имеет вид [1]: $\hat{x}_t^* = E\{x_t \mid x_{t-1}, \ldots\}$. Из условия регулярности ***R*2** следует, что

$$\widehat{x}_t^* = E\left\{\ \xi_t + \sum_{n\in\mathrm{N}} b_n x_{t-n} \middle| x_{t-1},\ldots\right\} = \sum_{n\in\mathrm{N}} b_n x_{t-n} + E\{\xi_t \mid \xi_{t-1},\ldots\} = \sum_{n\in\mathrm{N}} b_n x_{t-n}\,,$$

т.е. оптимальный прогноз по бесконечной предыстории линеен. Поэтому мы ограничимся рассмотрением линейных прогнозов вида:

$$\widehat{x}_t = \sum_{n\in\mathrm{N}} a_n x_{t-n}, \tag{8}$$

где $\{a_n \in \mathrm{R} : n = 1,2,\ldots\}$ – некоторые коэффициенты прогнозирования.

Для построения алгоритма прогнозирования на основе модели Блумфилда определим следующие операторы приближения спектра $S(\lambda)$ AR- и EXP-моделями порядка $p \in \mathrm{N}$: $\alpha_p(S)(\lambda)$ – спектр процесса $AR(p)$, у которого ковариации до лага p включительно совпадают с $\sigma_0,\ldots,\ \sigma_p$: $E_\Pi\{\alpha_p(S)e_n\}=E_\Pi\{Se_n\}$, $n=0,1,\ldots,p$; $\varepsilon_p(S)(\lambda) = \exp\left(\ \sum_{n=-p}^{p} l_n e_n(\lambda)\right)$, $\lambda \in \Pi$, – спектр процесса $EXP(p)$, который получается из (3) обнулением коэффициентов l_n=0 при $|n|>p$. Отметим, что при $p=+\infty$ в силу (3),(4) $\alpha_\infty(S)=\varepsilon_\infty(S)=S$. Отметим еще, что авторегрессионные коэффициенты $\{\varphi_{p,n}$:$n=1,\ldots,p\}$ спектра $\alpha_p(S)$ зависят только от первых p коэффициентов корреляции $\theta_1,\ldots,\ \theta_p$ спектра S и вычисляются реккурентно по формулам Дурбина-Левинзона [8]:

$$\varphi_{p,p} = \left(\theta_p - \sum_{n=1}^{p-1}\theta_{p-n}\varphi_{p-1,n}\right)\Big/\left(1-\sum_{n=1}^{p-1}\theta_n\varphi_{p-1,n}\right),\ \ \varphi_{p,n} = \varphi_{p-1,n} - \varphi_{p,p}\varphi_{p-1,p-n},\ \ 1 \le n < p. \tag{9}$$

В случае точного знания спектра $S(\lambda)$ оптимальный прогноз (8) для x_{T+1} по $T \in \mathrm{N}$ наблюдениям $x_1,\cdots,x_T$ строится по коэффициентам $a = (\varphi_{T,1},\ldots,\varphi_{T,T},0,0,\ldots)$. Поэтому, если мы вместо истинного спектра $S(\lambda)$ используем приближение Блумфилда $\varepsilon_p(S)$ порядка $p \in \mathrm{N}$, то следуя подстановочному принципу, прогноз строится по коэффициентам $\tilde{a} = (\tilde{\varphi}_{T,1},\ldots,\tilde{\varphi}_{T,T},0,0,\ldots)$, где для вычисления $\{\tilde{\varphi}_{T,n} : n = 1,\ldots,T\}$ в (9) подставляются коэффициенты корреляции спектра $\varepsilon_p(S)$: $\tilde{\theta}_\tau(l_1,\ldots,l_p) = E_\Pi\{\varepsilon_p(S)e_\tau\}\big/E_\Pi\{\varepsilon_p(S)\}$, $\tau = 1,\ldots,T$. Таким образом, прогноз $\widehat{x}_{T+1}$ на основе модели Блумфилда $EXP(p)$ строится по следующей вычислительной схеме:

$$(l_1,\ldots,l_p) \to (\tilde{\theta}_1,\ldots,\tilde{\theta}_T) \xrightarrow{(9)} \{\tilde{\varphi}_{T,1},\ldots,\tilde{\varphi}_{T,T}\} \xrightarrow{(8)} \widehat{x}_{T+1} = \sum_{s=1}^{T} \tilde{\varphi}_{T,s} x_{T+1-s}\,. \tag{10}$$

В дальнейшем прогнозирующую статистику $\widehat{x}_{T+1}$, построенную по схеме (10), будем называть $EXP(p)$-прогнозом.

Риск прогнозирования и его асимптотический анализ.

Теорема 1. *Пусть прогнозируемый временной ряд x_t имеет спектр $S(\lambda)$, а коэффициенты $\{a_n\}$ прогнозирующей статистики* (8) *порождают согласно* (4) *спектр $S_*(\lambda)$ с передаточной функцией $\alpha(z) = 1 - \sum_{n\in\mathrm{N}} a_n z^n$ и некоторым масштабным коэффициентом $\sigma_* > 0$. Тогда для риска $r(S_* \mid S) = E\{(\widehat{x}_t - x_t)^2\}$ прогнозирующей статистики* (8) *справедлива формула:*

$$\ln(r(S_* \mid S)) - l_0(S) = \ln(E_\Pi\{S/S_*\}) - E_\Pi\{\ln(S/S_*)\} \ge 0\,. \tag{11}$$

Доказательство. Используя операторный вид линейной фильтрации [1] и представление (4), имеем представление для погрешности прогноза

$$y_t = x_t - \widehat{x}_t = \alpha(\Delta)x_t = \frac{\alpha(\Delta)}{\beta(\Delta)}\xi_t\,,$$

где ξ_t – порождающий белый шум процесса x_t, $E\{\xi_t\} = 0$, $D\{\xi_t\} = \sigma^2 = \exp(l_0(S))$, Δ – оператор временного сдвига: $\Delta x_t = x_{t-1}$. Тогда спектр для y_t есть

$$S_y(\lambda) = D\{\xi_t\} \cdot \left| \frac{\alpha(e^{i\lambda})}{\beta(e^{i\lambda})} \right|^2 = \sigma_*^2 \cdot \frac{S(\lambda)}{S_*(\lambda)},$$

откуда с учетом (2) и принятых обозначений

$$\ln(r(S_* \mid S)) = \ln D\{y_t\} = \ln(E_\Pi\{S_y\}) = \ln \sigma_*^2 + \ln(E_\Pi\{S/S_*\}).$$

Учитывая, что согласно (5) $\ln \sigma^2 = l_0(S)$, $\ln \sigma_*^2 = l_0(S_*)$, получаем (11). Теорема доказана.

Теорема 1 характеризует увеличение риска прогнозирования за счет ошибки приближения истинного спектра $S(\lambda)$ функцией $S_*(\lambda)$. Отметим, что в силу неравенства Иенсена правая часть (11) всегда неотрицательна, инвариантна к умножению каждой из функций $S(\lambda)$, $S_*(\lambda)$ на положительную константу и обращается в ноль при $S_*(\lambda) = c \cdot S(\lambda)$, $c > 0$. Отметим еще, что минимальный риск прогнозирования $r(S|S)=\exp(l_0(S))=\sigma^2$ и достигается при $S_*(\lambda) = S(\lambda)$, что согласуется с [1].

Следствие 1. *Пусть* $S(\lambda) = S_*(\lambda) + \delta(\lambda)$, *где* $\delta(\lambda)$ – *погрешность приближения спектра* $S(\cdot)$ *функцией* $S_*(\cdot)$. *Если* $m = \sup_{\lambda \in \Pi}\{|\delta(\lambda)/S(\lambda)|\} \to 0$, *то справедливы следующие асимптотические разложения для риска и его абсолютного и относительного приращений:*

$$r(S_* \mid S) = \exp(l_0(S) + 0.5D_\Pi\{S_*/S\}) + O(m^3) \geq \exp(l_0(S)),$$

$$\Delta r = r(S_* \mid S) - r(S \mid S) = \exp(l_0(S)) \cdot 0.5D_\Pi\{S_*/S\} + O(m^3),$$

$$\kappa = \Delta r / r(S \mid S) = 0.5D_\Pi\{S_*/S\} + O(m^3).$$

Доказательство. Обозначим: $m_k=E_\Pi\{(\delta/S)^k\}$. Заметим, что $m_k=O(m^k)$ для $k \geq 0$. Тогда согласно Теореме 1

$$\ln E_\Pi\left\{\frac{S}{S_*}\right\} = \ln E_\Pi\left\{\left(1 - \frac{\delta}{S}\right)^{-1}\right\} = \ln\left(\sum_{k=0}^{\infty} m_k\right) = m_1 + m_2 - \frac{m_1^2}{2} + O(m^3);\ E_\Pi\left\{\ln \frac{S}{S_*}\right\} =$$

$$= E_\Pi\left\{-\ln\left(1 - \frac{\delta}{S}\right)\right\} = m_1 + \frac{m_2}{2} + O(m^3);\ \ln r(S_* \mid S) - l_0(S) = \frac{m_2 - m_1^2}{2} + O(m^3) = \frac{1}{2} D_\Pi\left\{\frac{\delta}{S}\right\} + O(m^3),$$

откуда, учитывая, что $D_\Pi\{\delta/S\} = D_\Pi\{1 - S_*/S\} = D_\Pi\{S_*/S\}$, следуют доказываемые асимптотические разложения. Следствие доказано.

Данный результат характеризует отклонение риска прогнозирования от минимального возможного значения при малых отклонениях используемой модели $S_*(\lambda)$ от истинного спектра $S(\lambda)$.

Для $EXP(p)$-прогноза, определяемого (10), при длительности наблюдений $T \in \mathrm{N}$ имеем $S_* = \alpha_T(\varepsilon_p(S))$. Будем обозначать риск этого прогноза: $r_{T,p} = r(\alpha_T(\varepsilon_p(S)) \mid S)$. Исследуем асимптотику риска $r_{T,p}$ при $T \to +\infty$.

Теорема 2. *Для* $p \in \mathrm{N}$ *риск* $EXP(p)$*-прогноза* $r_{T,p}$ *с увеличением длительности наблюдения* $T \to +\infty$ *сходится к величине*

$$r_{T,p} \to \exp(l_0(S)) \cdot E_\Pi\left\{\exp\left(2\sum_{k=p+1}^{\infty} l_k \cos(k\lambda)\right)\right\} \geq r(S \mid S). \qquad (12)$$

Доказательство. При $T \to +\infty$ имеем $r_{T,p} \to r_{\infty,p} = r(\varepsilon_p(S) \mid S)$, т.к. $\alpha_\infty(\varepsilon_p(S)) = \varepsilon_p(S)$. Поскольку $E_\Pi\{\ln S\} = l_0(S) = l_0(\varepsilon_p(S)) = E_\Pi\{\ln(\varepsilon_p(S))\}$, из (11) получаем

$$r(\varepsilon_p(S)\,|\,S) = \exp(l_0(S)) \cdot E_\Pi \left\{ \frac{S}{\varepsilon_p(S)} \right\} \geq r(S\,|\,S),$$

что совпадает с правой частью (12). Теорема доказана.

Заметим, что скорость убывания правой части (12) при $p \to \infty$ напрямую связана со скоростью убывания коэффициентов l_n. Следующий результат показывает, при каких условиях последовательность $EXP(\infty)$-коэффициентов l_n убывает быстрее, чем последовательность $AR(\infty)$-коэффициентов b_n.

Лемма 3. *Для* $z \in \mathrm{C}, |\operatorname{Im}\{z\}| < \operatorname{Re}\{z\}, y \in \mathrm{R}$, *выполняется равенство:*

$$\int_{\mathrm{R}} \exp\left\{ -\frac{x^2 z^2}{2} + ixy \right\} dx = \frac{\sqrt{2\pi}}{z} \exp\left\{ -\frac{y^2}{2z^2} \right\}. \tag{13}$$

Теорема 3. *Обозначим:* $l(e^{i\lambda}) = l_c(\lambda) = l_r(\lambda) + i l_i(\lambda)$, *где* $l_r(\lambda), l_i(\lambda)$ *– соответственно действительная и мнимая части* $l_c(\lambda)$. *Пусть* $l_c(\lambda)$ *дважды дифференцируема и* $l_r(\lambda)$ *имеет* $M<+\infty$ *минимумов в точках* $\lambda_1, \ldots, \lambda_M \in \Pi$, *причем* $l_r(\lambda_k) = \gamma$, *а вторые производные* $l_c''(\lambda_k) = Q_k e^{i\psi_k}$, $Q_k > 0, |\psi_k| < \pi/2$, $k = 1, \ldots, M$. *Тогда авторегрессионные коэффициенты модели Вольда* $b_n(\omega)$, *отвечающие спектру* $(S(\lambda))^\omega$, *при* $\omega \to +\infty$ *имеют вид* (n=1,2,…)*:*

$$b_n(\omega) = -\frac{e^{-\gamma\omega}}{\sqrt{2\pi\omega}} \left(\sum_{k=1}^{M} F\left(\frac{\omega l_i'(\lambda_k) + n}{\sqrt{\omega}}; Q_k, \psi_k, \omega l_i(\lambda_k) + n\lambda_k \right) + o(1) \right), \tag{14}$$

$$F(x; Q, \psi, a) = \frac{1}{\sqrt{Q}} \exp\left\{ -\frac{x^2}{2Q} \cos\psi \right\} \cos\left(\frac{1}{2}\left(\psi - \frac{x^2}{Q} \sin\psi \right) + a \right).$$

Доказательство. Выделим на отрезке Π попарно непересекающиеся окрестности точек λ_k: $\Pi_k = [\lambda_k - \varepsilon, \lambda_k + \varepsilon] \bmod \Pi$, k=1,2,…,M, где $\forall A \subset \mathrm{R}, A \bmod \Pi = (A + 2\pi \cdot \mathrm{Z}) \bigcap \Pi$, и $\varepsilon \sim \omega^{-0.4}$ при $\omega \to +\infty$. Также обозначим их дополнение: $\Pi_0 = \Pi \setminus \bigcup_{k=1}^{M} \Pi_k$. Тогда согласно (5)

$$-b_n(\omega) = E_\Pi\{\exp(-\omega l_c(\lambda) - in\lambda)\} = \sum\nolimits_{k=0}^{M} I_k, \;\; I_k = E_\Pi\{I_{\Pi_k}(\lambda) \cdot \exp(-\omega l_c(\lambda) - in\lambda)\},$$

где I_A – индикатор множества A. Оценим I_k:

$$|I_0| \leq \max_{\lambda \in \Pi_0}\{\exp(-\omega l_r(\lambda))\} = \exp\left\{ \max_{\lambda \in \Pi_0}\{-\omega l_r(\lambda)\} \right\} =$$

$$= \exp\left\{ -\omega\left(\gamma + \frac{\varepsilon^2}{2} \underline{\min}_{k=1,M}\{l_r''(\lambda_k)\} + o(\varepsilon^2) \right) \right\} = \exp\{-\omega\gamma - \omega^{0.2}(c + o(1))\} = o\left(\frac{e^{-\gamma\omega}}{\sqrt{\omega}} \right),$$

где c>0 по условию Теоремы. Теперь для k=1,2,...,M, $x = y/\sqrt{\omega}$, используя (13), получим:

$$I_k \cdot 2\pi \exp\{\omega l_c(\lambda_k) + in\lambda_k\} = \int_{-\varepsilon}^{+\varepsilon} \exp\left\{ -\omega\left(ixl_i'(\lambda_k) + \frac{x^2}{2} l_c''(\lambda_k) + O(\varepsilon^3) \right) - inx \right\} dx =$$

$$= \frac{1}{\sqrt{\omega}} \left(\int_{-\infty}^{+\infty} \exp\left\{ -iy \frac{\omega l_i'(\lambda_k) + n}{\sqrt{\omega}} - \frac{y^2}{2} l_c''(\lambda_k) + O(\omega^{-0.2}) \right\} dy + o(1) \right) =$$

$$= \sqrt{\frac{2\pi}{\omega}} \left(\frac{1}{\sqrt{Q_k}} \exp\left\{ -\frac{1}{2}\left(\frac{(\omega l_i'(\lambda_k) + n)^2}{\omega Q_k} e^{-i\psi_k} + i\psi_k \right) \right\} + o(1) \right).$$

Объединяя выражения для $I_0, I_1, \ldots, I_M$, и учитывая, что $b_n(\omega) = \operatorname{Re}\{b_n(\omega)\} = -\sum\nolimits_{k=0}^{M} \operatorname{Re}\{I_k\}$, получаем (14). Теорема доказана.

Спектры, сосредоточенные в малых окрестностях конечного числа ярко выраженных максимумов, будем называть “гребенчатыми”. Поскольку такие спектры могут быть получены возведением подходящего многомодального спектра $S(\lambda)$ в достаточно большую степень ω, для них справедливы асимптотические выводы относительно модели S^{ω} при $\omega\to+\infty$.

Согласно Теореме 3, при возведении спектра S в достаточно большую степень ω последовательность коэффициентов авторегрессии спектра S^{ω} принимает форму суммы “осциллирующих гауссовских всплесков”, задаваемых функцией $F(\cdot)$, причем центры этих “всплесков”, как видно из (14), линейно зависят от ω. Поэтому адекватная $AR(p)$-аппроксимация спектра S^{ω} должна иметь порядок $p_{AR}(\omega)=O(\omega)$, чтобы учесть все “всплески”. В то же время в модели Блумфилда $\ln S^{\omega}(\lambda)=\omega L(\lambda)$, и последовательность кепстральных коэффициентов спектра S^{ω} с ростом ω не меняет форму, а только масштабируется коэффициентом ω. По условию Теоремы 3 ряд $l(z)$ не имеет особенностей при $|z|=1$, значит его радиус сходимости $\mu>1$ и $|l_n|^{1/n}\to 1/\mu$. Тогда, если выбор порядка $p(\omega)$ адекватной $EXP(p)$-модели спектра S^{ω} производится из условия: $|l_{p(\omega)}(\omega)|=|\omega\cdot l_{p(\omega)}|=\varepsilon>0$, то $\ln(\omega)-\ln(\varepsilon)=-\ln|l_{p(\omega)}|\sim p(\omega)\ln(\mu)$ и $p_{EXP}(\omega)=O(\ln(\omega))$. Отсюда видно, что $p_{AR}(\omega)/p_{EXP}(\omega)=O(\omega/\ln(\omega))\to\infty$. Следовательно, для “гребенчатых спектров” $EXP(p)$-аппроксимация требует существенно меньшего числа параметров, чем $AR(p)$-аппроксимация.

Обозначим: $f(z)=\sum_{n\in Z}\sigma_n z^n$, $\beta_p(z)=1-\sum_{k=1}^{p}\varphi_{p,k}z^k$, $z\in C$. Следующий результат позволяет аналитически выразить $\beta_p(z)$ через $f(z)$.

Теорема 4. *Для* $z\in C$:

$$\beta_p(z)=\frac{u(z)}{u(0)},\quad u(z)=\int_{\Gamma^p}\Pi(z,z_1,\dots,z_p)\prod_{k=1}^{p}f(z_k)\frac{dz_k}{z_k}, \tag{15}$$

где $\Gamma: t\to e^{it}, t\in\Pi$, – контур единичного круга, и

$$\Pi(z_1,\dots,z_n)=\prod_{1\le j<k\le n}\left(1-\frac{z_j}{z_k}\right),\ n\in N.$$

Доказательство. Согласно [9] для определителя Теплицевой матрицы $\gamma=|\sigma_{i-j}|_{i,j=1}^{p}$ справедливы соотношения:

$$\gamma\beta_p(z_0)=\begin{vmatrix}\sigma_0 & \sigma_1 & \dots & \sigma_{p-1} & z_0^p\\ \sigma_1 & \sigma_0 & \dots & \sigma_{p-2} & z_0^{p-1}\\ \vdots & \vdots & \ddots & \vdots & \vdots\\ \sigma_p & \sigma_{p-1} & \dots & \sigma_1 & 1\end{vmatrix}=\frac{1}{(2\pi i)^p}\int_{\Gamma^p}\begin{vmatrix}1 & z_{p-1} & \dots & z_1^{p-1} & z_0^p\\ z_p^{-1} & 1 & \dots & z_1^{p-2} & z_0^{p-1}\\ \vdots & \vdots & \ddots & \vdots & \vdots\\ z_p^{-p} & z_{p-1}^{1-p} & \dots & z_1^{-1} & 1\end{vmatrix}\times\prod_{k=1}^{p}f(z_k)\frac{dz_k}{z_k}=$$

$$=\frac{1}{(2\pi i)^p}\int_{\Gamma^p}\prod_{0\le j<j'\le p}\left(1-\frac{z_j}{z_{j'}}\right)\prod_{k=1}^{p}f(z_k)\frac{dz_k}{z_k},$$

откуда по условию нормировки ($\beta_p(0)=1$) получаем (15). Теорема доказана.

Формула (11) с учетом (15) дает аналитическое выражение риска $r(S_*\,|\,S)$ прогнозирования через спектр $S(\lambda)$ и его приближение $S_*(\lambda)$. В тех случаях, когда приближение $S_*(\lambda)$ порождается некоторым оператором, действующим на $S(\lambda)$: $S_*=g(S)$, риск прогнозирования представляется в виде функционала от спектра и оператора приближения: $r(S_*\,|\,S)=r(g(S)\,|\,S)=r(g\,|\,S)$. Этот функционал может быть выражен аналитически, если в свою очередь известен аналитический вид оператора g. Тогда задачи исследования и сравнения различных операторов приближения g могут быть представлены в виде, аналитическом относительно спектра $S(\lambda)$, как задачи функционального анализа и решаться соответствующими методами. Примером такой задачи является исследование области выигрыша по риску одного оператора приближения g у альтернативного оператора h: $U(g|h)=\{S(\cdot):r(g|S)<r(h|S)\}$.

Для $EXP(p)$-прогноза (10) при длительности наблюдения T: $S_* = \alpha_T(\varepsilon_p(S))$, поэтому оператор приближения $g = \alpha_T \circ \varepsilon_p$ есть композиция операторов α_T и ε_p. Теорема 4 позволяет аналитически выразить оператор α_n, $n \in \mathrm{N}$, с точностью до умножения на константу: $\alpha_n(S)(\lambda) = c\cdot|\beta(e^{i\lambda})|^{-2}$, $\lambda \in \Pi$, $c>0$, где $\beta_p(e^{i\lambda})$ выражается через $f(e^{ix}) = S(x), x \in \Pi$, согласно (15).

В заключение заметим, что результаты численных экспериментов по сравнению рисков $AR(p)$- и $EXP(p)$-прогнозов показали значительный выигрыш $EXP(p)$-прогноза в величине риска при увеличении ω, что согласуется с Теоремой 3.

Литература

1. Андерсон Т.В. Статистический анализ временных рядов. М., 1976.
2. Харин Ю.С. Оптимальность и робастность в статистическом прогнозировании. Минск, 2008.
3. Serdobolskiy V.I. Multivariate Statistical Analysis. A High-Dimensional Approach. Dordrecht, 2000.
4. Bloomfield P. // Biometrika. 1973. Vol. 60, P. 217-226.
5. Ширяев А.Н. Вероятность. М., 2007.
6. Merhav N. // IEEE Transactions On Signal Processing. 1993. Vol. 41, No. 5, P. 1990-1993.
7. Pourahmadi M. // Communications in Statistics, Part A – Theory and Methods. 1983. Vol. 12, P. 2085-2094.
8. Maronna R.A. Robust statistics: theory and methods. New York, 2006.
9. Сеге Г. Ортогональные многочлены. М., 1962.

УДК 519.2

Харин Ю.С., Волошко В.А. **Прогнозирование стационарных временных рядов на основе малопараметрической модели Блумфилда** // Докл. НАН Беларуси. 2010. Т. 54, № . С.

Построена статистика одношагового прогнозирования стационарного временного ряда на основе малопараметрической модели Блумфилда для случая известной спектральной плотности временного ряда. Получены оценки среднеквадратического риска построенной прогнозирующей статистики в сравнении с риском авторегрессионной прогнозирующей статистики. Установлен класс спектральных плотностей, для которых построенная прогнозирующая статистика имеет наибольший выигрыш по величине риска.

Библиогр. – 9 назв.

KHARIN YU.S., VOLOSHKO V.A.

kharin@bsu.by; valeravoloshko@yandex.ru

STATIONARY TIME SERIES FORECASTING
BASED ON THE SMALL-PARAMETRIC BLOOMFIELD MODEL

Summary

Forecasting of stationary time series based on the Bloomfield model is considered. The mean-square risk of forecasting is analyzed for the situation with known spectral density of the observed time series.

Волошко Валерий Анатольевич – аспирант Факультета прикладной математики и информатики Белорусского государственного университета, раб. тел. 209-55-49, моб. тел. +375-29-144-36-90, email: valeravoloshko@yandex.ru.

Харин Юрий Семенович – директор НИИ прикладных проблем математики и информатики Белорусского государственного университета, раб. тел. 209-51-04, моб. тел. +375-29-621-82-36, email: kharin@bsu.by.