STATISTICAL FORECASTING BASED ON BLOOMFIELD EXPONENTIAL MODEL

V.A. Voloshko*, Yu.S. Kharin**
Belarusian State University
Minsk, BELARUS
e-mail: *valeravoloshko@yandex.ru, **kharin@bsu.by

Abstract

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 parameters of the model.

1 Introduction

The accuracy of statistical inferences (estimates, decisions, forecasts) in parametric data analysis, as it is known [1, 2], depends on the ratio of the number p of model parameters and the observation length T: $\rho(p,T) = p/T$. If $p << T$, i.e. $\rho(p,T) \to 0$ at $T \to \infty$, then statistical inferences based on classical methods (least squares, maximum likelihood, Bayesian method) appear consistent and give an acceptable accuracy for practice. On the other hand, if the number of model parameters p is comparable with T, i.e. $\rho(p,T) \to c$, $0 < c < 1$, then classical methods appear inapplicable, and acceptable accuracy is reached only for some special cases. As a result, the problem of development and statistical analysis of small-parametric models, i.e. models with small $\rho(p,T)$, is very topical. In [3] Bloomfield proposed the so-called exponential model EXP(p) of order p for stationary time series and built some estimators of its parameters. This paper is devoted to using this model for statistical forecasting.

2 Bloomfield model and its properties

Introduce the notation: $\Pi = [-\pi, \pi]$; $e_n(\lambda) = e^{in\lambda}, \lambda \in \Pi, n \in \mathbb{Z}$; $E_\Pi\{f\} = (2\pi)^{-1} \times \times \int_\Pi f(\lambda)d\lambda$, $D_\Pi\{f\} = E_\Pi\{|f|^2\} - |E_\Pi\{f\}|^2$, $f : \Pi \to \mathbb{C}$.

Let $\{x_t\}, t \in \mathbb{Z}$, be a real valued stationary time series with zero mean $E\{x_t\} = 0$, a covariance function $\sigma_\tau = E\{x_t x_{t+\tau}\}$, the correspondent correlation function $\theta_\tau = \sigma_\tau/\sigma_0$, $\tau \in \mathbb{Z}$, and the spectral density function $S(\lambda) = \sum_{\tau\in\mathbb{Z}} \sigma_\tau e_\tau(\lambda)$, $L(\lambda) = \ln S(\lambda)$, $\lambda \in \Pi$. For brevity we call $S(\lambda)$ and $L(\lambda)$ the *spectrum* and the *log-spectrum* respectively.

Under the assumption $L(\cdot) \in L_2(\Pi)$, any time series x_t is uniquely representable by the following models of infinite order: the Bloomfield exponential model EXP(∞) [3]:

$$S(\lambda) = \exp\left\{l_0 + 2\mathrm{Re}\{l(e^{i\lambda})\}\right\},\ l(z) = \sum_{n\in\mathbb{N}} l_n z^n,\ l_0, l_1, \dots \in \mathbb{R}, \quad (1)$$

and the Wold autoregressive model AR(∞) [1]:

$$S(\lambda) = \frac{\sigma^2}{|\beta(e^{i\lambda})|^2}, \ \sigma > 0, \ \beta(z) = 1 - \sum_{n \in \mathbb{N}} b_n z^n, \ b_1, b_2, \dots \in \mathbb{R}, \tag{2}$$

where $l(z), \beta(z), z \in \mathbb{C}$, are formal power series. Note, that the coefficient $l_n = l_{-n} = E_\Pi\{Le_n\}$ is called the n-th cepstral coefficient in the theory of digital signal processing, and $\beta(\cdot)$ is called the transfer function of the spectrum $S(\lambda)$. Pourahmadi formulas [4] provide relation between parameters of the model AR(∞) and the model EXP(∞): $l_0 = \ln \sigma^2$, $\beta(z) = e^{-l(z)} \Leftrightarrow b_n = l_n - n^{-1} \sum_{j=1}^{n-1} j l_j b_{n-j}$, $n = 1, 2, \dots$.

3 Forecasting based on the Bloomfield model

As we use for forecasting only information on the second order moments $\sigma_{t-s} = E\{x_t x_s\}$ of the process x_t, let us consider linear (by the prehistory $\{x_s : s < t\}$) forecasts [1]:

$$\hat{x}_t = \sum_{n=1}^{\infty} a_n x_{t-n}, \tag{3}$$

where $\{a_n \in \mathbb{R} : n = 1, 2, \dots\}$ are some forecasting coefficients. In this case the risk $r = E\{(\hat{x}_t - x_t)^2\}$ depends on $\{\sigma_\tau : \tau \in \mathbb{Z}\}$ only.

To construct a forecasting algorithm based on the Bloomfield model, we define the following operators of approximation for the spectrum $S(\lambda)$ by AR- and EXP-models of order $p \in \mathbb{N}$: $\alpha_p(S)(\lambda)$ is a spectrum of the AR(p)-process with first covariances till lag p coinciding with $\sigma_0, \dots, \sigma_p$: $E_\Pi\{\alpha_p(S)e_n\} = E_\Pi\{Se_n\} = \sigma_n, n = 0, 1, \dots, p$; $\varepsilon_p(S)(\lambda) = \exp\left\{\sum_{|n| \le p} l_n e_n(\lambda)\right\}$, $\lambda \in \Pi$, is a spectrum of the EXP(p)-process with parameters $l_n = 0$ at $|n| > p$. Note, that due to (1) and (2) $\alpha_\infty(S) = \varepsilon_\infty(S) = S$. Autoregression coefficients $\{\varphi_{p,n} : n = 1, \dots, p\}$ of the spectrum $\alpha_p(S)$ depend only on the first p correlation coefficients $\theta_1, \dots, \theta_p$ of the spectrum S. They are calculated recurrently by the Durbin-Levinson formulas [5]:

$$\phi_{p,p} = \frac{\theta_p - \sum_{n=1}^{p-1} \theta_{p-n} \phi_{p-1,n}}{1 - \sum_{n=1}^{p-1} \theta_n \phi_{p-1,n}}, \ \phi_{p,n} = \phi_{p-1,n} - \phi_{p,p} \phi_{p-1,p-n}, \ 1 \le n < p. \tag{4}$$

If the spectrum $S(\lambda)$ is known exactly, then the optimal forecast (3) of the depth $T \in \mathbb{N}$ is built by coefficients $a = (\varphi_{T,1}, \dots, \varphi_{T,T}, 0, 0, \dots)$. Therefore, if we consider the Bloomfield approximation $\varepsilon_p(S)$ of the order $p \in \mathbb{N}$ to be an exact model of the time series x_t, then the optimal forecast (3) of the depth T is built by coefficients $\widetilde{a} = (\widetilde{\varphi}_{T,1}, \dots, \widetilde{\varphi}_{T,T}, 0, 0, \dots)$, where $\{\widetilde{\varphi}_{T,n} : n = 1, \dots, T\}$ are calculated by putting into (4) the numerically found EXP(p)-approximations of the correlation coefficients: $\widetilde{\theta}_\tau = E_\Pi\{\varepsilon_p(S)e_\tau\}/E_\Pi\{\varepsilon_p(S)\}, \tau = 1, \dots, T$. In other words, the EXP(p)-forecast $\hat{x}_t$ is built by the following computing scheme:

$$(l_1, \dots, l_p) \to (\widetilde{\theta}_1, \dots, \widetilde{\theta}_T) \stackrel{(4)}{\to} (\widetilde{\varphi}_{T,1}, \dots, \widetilde{\varphi}_{T,T}) \stackrel{(3)}{\to} \hat{x}_t = \sum_{s=1}^{T} \widetilde{\varphi}_{T,s} x_{t-s}. \tag{5}$$

4 Risk of forecasting and its asymptotic analysis

Theorem 1. *Let the forecasted time series x_t have the spectrum $S(\lambda)$, the coefficients $\{a_n\}$ of the forecasting statistic (3) generate the spectrum $S_*(\lambda) = \sigma_*^2/|\alpha(e^{i\lambda})|^2$ with the transfer function $\alpha(\cdot)$ specified by the series: $\alpha(z) = 1 - \sum_{n=1}^{\infty} a_n z^n$, and $\sigma = \sigma_*$. Then the risk $r(S_*|S) = E\{(\hat{x}_t - x_t)^2\}$ of the forecasting statistic (3) satisfies the relation:*

$$\ln\{r(S_*|S)\} - l_0(S) = \ln\{E_\Pi\{S/S_*\}\} - E_\Pi\{\ln\{S/S_*\}\} \geq 0. \tag{6}$$

Theorem 1 characterizes an increment of the risk of forecasting caused by an error of the approximation of the spectrum $S(\lambda)$ by the function $S_*(\lambda)$. Note, that by the Jensen inequality the right side of (6) is always nonnegative, invariant to the scaling of each of the functions $S(\lambda), S_*(\lambda)$ by some positive factor and vanishes at $S_*(\lambda) = c \cdot S(\lambda),\ c > 0$. Also note, that the minimal risk of forecasting $r(S|S) = \exp(l_0(S)) = \sigma^2$ is reached at $S_*(\cdot) = S(\cdot)$, that fits [1].

Corollary 1. *Let $S(\lambda) = S_*(\lambda) + \delta(\lambda)$, where $\delta(\lambda)$ is an approximation error. If $m = \sup_{\lambda \in \Pi}\{|\delta(\lambda)/S(\lambda)|\} \to 0$, then the following asymptotic expressions hold:*

$$\begin{gathered} r(S_*|S) = e^{l_0(S)} \cdot e^{0.5 D_\Pi\{\delta/S\}} + O(m^3) \geq e^{l_0(S)}, \\ \Delta r = r(S_*|S) - r(S|S) = e^{l_0(S)} \cdot 0.5 D_\Pi\{\delta/S\} + O(m^3), \\ \kappa = \Delta r / r(S|S) = 0.5 D_\Pi\{S_*/S\} + O(m^3). \end{gathered}$$

This result characterizes the risk deviation from the minimal one, when the deviation of the model $S_*(\lambda)$ from the true model $S(\lambda)$ is small.

Due to the scheme (5), for the EXP(p)-forecast of the depth $T \in \mathbb{N}$ we have $S_* = \alpha_T(\varepsilon_p(S))$. Denote the risk of this forecast: $r_{T,p} = r(\alpha_T(\varepsilon_p(S))|S)$, and investigate its asymptotics.

Theorem 2. *For $p \in \mathbb{N}$ and increasing depth $T \to +\infty$ the risk of the EXP(p)-forecast $r_{T,p}$ satisfies the asymptotics:*

$$r_{T,p} \to \exp\{l_0(S)\} \cdot E_\Pi \left\{ \exp\left\{ 2 \sum_{k=p+1}^{\infty} l_k \cos(k\lambda) \right\} \right\}. \tag{7}$$

Note, that the rate of decrease of the right side of (7) at $p \to \infty$ is directly related to the rate of decrease of coefficients l_n at $n \to \infty$. The following result formulate the conditions, when the sequence of the EXP(∞)-coefficients l_n decreases faster, than the sequence of the AR(∞)-coefficients b_n.

Theorem 3. *Let $l_c(\lambda) = l(e^{i\lambda}), \lambda \in \Pi$, be twice differentiable and $Re\{l_c(\lambda)\}$ have $M < +\infty$ minimums l_r^- at points $\lambda_1, \ldots, \lambda_M$, while the second derivatives $l_c''(\lambda_k) = R_k e^{i\varphi_k}, R_k > 0, |\varphi_k| < \pi/2, k = 1, \ldots, M$. Then the coefficients of the Wold autoregression $b_n(\gamma)$ for the spectrum $(S(\lambda))^\gamma$ have the following form at $\gamma \to +\infty$*

$(l_i(\lambda) = Im\{l_c(\lambda)\})$:

$$b_n(\gamma) = -\frac{e^{-\gamma l_r^-}}{\sqrt{2\pi\gamma}} \left(\sum_{k=1}^{M} F\left(\frac{\gamma l_i'(\lambda_k) + n}{\sqrt{\gamma}}; R_k, \varphi_k, \gamma l_i(\lambda_k) + n\lambda_k \right) + o(1) \right), \; n \in \mathbb{N},$$

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

Let us consider *polymodal* spectrums that are concentrated in tiny neighborhoods of finite high-contrast maximums. The asymptotic conclusions about models S^γ at $\gamma \to +\infty$ lead to the conclusions for polymodal spectrums, since they can be produced by exponentiating a suitable nonpolymodal spectrum $S(\lambda)$ with a large degree γ.

According to Theorem 3, the autoregression coefficients $b_n(\gamma)$ of the spectrum S^γ for large γ are significant at positions $n \sim c \cdot \gamma, c > 0$, whence the adequate AR(p)-approximation of the spectrum S^γ should have an order $p(\gamma) \sim c \cdot \gamma$. At the same time the sequence of the cepstral coefficients $l_n(\gamma)$ of the spectrum S^γ is only getting scaled when increasing γ. In the worst case of slow decrease of cepstral coefficients of the spectrum S through the power low: $l_n \sim c \cdot n^{-q}, c > 0, q \geq 1$. If the choice of the order $p(\gamma)$ for the adequate EXP(p)-model of the spectrum S^γ is made from the condition: $|l_{p(\gamma)}(\gamma)| = |\gamma \cdot l_{p(\gamma)}| = \varepsilon > 0$, we get: $\gamma \cdot cp(\gamma)^{-q} \sim \varepsilon \Rightarrow p(\gamma) \sim (\gamma c/\varepsilon)^{1/q}$. Comparing the rates of increasing for $p(\gamma)$ at $\gamma \to +\infty$ for AR(p)- and EXP(p)-approximations, we get, taking into account $1/q \leq 1$, that an adequate EXP(p)-approximation of the polymodal spectrum requires significantly less of parameters, then the AR(p)'s one.

In the numerical experiments the risks of the AR(p)- and the EXP(p)-forecasts of the depth $T = 80$ were compared. Some spectrum $S(\lambda)$ was chosen and the risk was computed for $\gamma = 1, 2, \ldots, 6$ and $p = 1, 2, \ldots, 80$. The results of experiments show an increasing gain of the EXP(p)-forecast w.r.t. the AR(p)-forecast for increasing γ, that fits the results of Theorem 3.

References

[1] Anderson T.W. (1994). *The Statistical Analysis of Time Series*. Wiley-Interscience, New York.

[2] Kharin Yu.S. (2008). *Optimality and robustness in statistical forecasting (in Russian)*. BSU, Minsk.

[3] Bloomfield P. (1973). An exponential model for the spectrum of a scalar time series. *Biometrika*. Vol. **60**, pp. 217-226.

[4] Pourahmadi M. (1983). Exact Factorization of the Spectral Density and Its Application to Forecasting and Time Series Analysis. *Communications in Statistics, Part A – Theory and Methods*. Vol. **12**, pp. 2085-2094.

[5] Maronna R.A., Martin R.D., Yohai V.J. (2006). *Robust statistics: theory and methods*. Wiley, New York.