ON COMPARISON OF CLASSICAL AND CEPSTRUM-BASED FORECASTS

V.A. Valoshka*, Yu.S. Kharin**
Research Institute for Applied Problems of Mathematics and Informatics
Minsk, BELARUS
e-mail: *valoshka@bsu.by, **kharin@bsu.by

Abstract

The alternative cepstrum-based approach to forecasting is experimentally compared with the classical correlation-based one. For the class of time series with the polymodal spectral densities the alternative forecast is shown to be sufficiently more accurate then the classical forecast, when both forecasts use the same small number of estimated parameters.

1 Introduction

The usual approach in linear forecasting of a wide-sense stationary time series with unknown spectral density is to estimate some correlation coefficients and to compute estimates of the optimal forecasting coefficients. When the correlation function vanishes fast, its enough to estimate a small number of correlation coefficients, and the traditional methods bring a good result. But for the long-memory processes, when the large number of correlations should be estimated, it is useful to have some suitable parsimonious model of the spectral density. Then we can accurately estimate a few parameters of this model and transform it to the estimates of a necessary number of correlations. This paper presents some experimental results, demonstrating the benefit of using the parsimonious Bloomfield model [2] for forecasting time series with polymodal spectral density.

2 Mathematical model and algorithms

Let $\{x_t\}_{t\in\mathbb{Z}}$ be the Gaussian stationary time series with zero mean $E\{x_t\} = 0$, covariance function $\sigma_\tau = E\{x_t x_{t+\tau}\}$, $\tau \in \mathbb{Z}$, and the spectral density $S(\lambda) = \sum_{\tau\in\mathbb{Z}} \sigma_\tau \cos(\tau\lambda)$, $\lambda \in \Pi = [-\pi, \pi]$, which is known to be non negative. The correlation coefficients (or correlations) are $\theta_\tau = \sigma_\tau/\sigma_0$. The theoretical cepstral coefficient [2] is the fourier coefficient of the logarithm of spectral density:

$$l_\tau = \frac{1}{2\pi}\int_\Pi \ln S(\lambda)\cos(\tau\lambda)d\lambda,\ \tau \in \mathbb{Z}.$$

For the time series $x_1, \ldots, x_T$ of the length $T \in \mathbb{N}$ we should fix some smooth parameter $p \in \mathbb{N}$ and compare two approaches of one step forward forecasting for x_{T+1} on the base of the mean squared error, or risk:

$$\mathcal{R} = E\left\{(x_{T+1} - \hat{x}_{T+1})^2\right\}.$$

In the traditional approach the first p correlations $\theta_1, \ldots, \theta_p$ are straightly estimated by the sample covariances, and then the linear forecast of depth p is computed. In the alternative approach the first p cepstrum coefficients $l_1, \ldots, l_p$ are estimated, and then the estimates of the first T correlations $\theta_1, \ldots, \theta_T$ and finally the linear forecast of depth T are computed. In [2] it was suggested that the alternative approach reaches a sufficient gain w.r.t. the traditional one for the time series with polymodal spectral densities, which are concentrated in the thin neighborhoods of a few values of $\lambda \in \Pi$. Here some experimental results supporting this hypothesis are presented.

Introduce the algorithms. The biased sample covariances and classical estimates of correlations [1]:

$$\hat{\sigma}_\tau = \hat{\sigma}_{-\tau} = \frac{1}{T} \sum_{t=1}^{T-\tau} x_t x_{t+\tau}, \ \hat{\theta}_\tau^C = \hat{\sigma}_\tau / \hat{\sigma}_0, \ \tau = 0, 1, \ldots, T-1.$$

The smoothed periodogram and the separated from zero one:

$$\hat{S}(\lambda) = \sum_{\tau=1-K}^{K-1} \hat{\sigma}_\tau f(\tau/K) \cos(\tau\lambda), \ \hat{S}^+(\lambda) = \max\left\{\hat{S}(\lambda), \varepsilon\hat{\sigma}_0\right\}, \ \lambda \in \mathbb{R}. \tag{1}$$

Here $f : \mathbb{R} \to \mathbb{R}$, $K \in \mathbb{N}$ and $\varepsilon > 0$ are respectively the kernel of smoothing, the width of the smoothing window and zero separator. It will be clear from the results of experiments that varying these three parameters can sufficiently increase the accuracy of the alternative (cepstrum-based) forecast. Introduce the plug-in cepstrum estimates [1] and the alternative estimates of the correlations:

$$\hat{l}_\tau = \frac{1}{T} \sum_{t=1}^{T} \hat{S}^+\left(\frac{2\pi t}{T}\right) \cos\left(\frac{2\pi\tau t}{T}\right),$$

$$\hat{\theta}_\tau^A = \frac{\Theta_\tau}{\Theta_0}, \ \Theta_\tau = \frac{1}{2\pi} \int_\Pi \exp\left(2\sum_{j=1}^{p} \hat{l}_j \cos(j\lambda)\right) \cos(\tau\lambda) d\lambda,$$

where the integrals for Θ_τ in the algorithm are computed approximately by numerical methods. Finally using Durbin-Levinson formulas [2] the classical estimates of the first p correlations $\hat{\theta}_1^C, \ldots, \hat{\theta}_p^C$ are transformed into the estimates of the optimal forecasting coefficients $a_1^C, \ldots, a_p^C$ at the depth of p. Similarly, the alternative estimates of the first T correlations $\hat{\theta}_1^A, \ldots, \hat{\theta}_T^A$ are transformed into the estimates of the optimal forecasting coefficients $a_1^A, \ldots, a_T^A$ at the depth of T. The classical and the alternative forecasting statistics for x_{T+1} are:

$$\hat{x}_{T+1}^C = \sum_{t=T-p+1}^{T} x_t a_{T+1-t}^C, \ \hat{x}_{T+1}^A = \sum_{t=1}^{T} x_t a_{T+1-t}^A,$$

and the corresponding risks to be examined:

$$\mathcal{R}^C(S, T, p) = E\left\{\left(\hat{x}_{T+1}^C - x_{T+1}\right)^2\right\}, \tag{2}$$

$$\mathcal{R}^A(S, T, p; f, K, \varepsilon) = E\left\{\left(\hat{x}_{T+1}^A - x_{T+1}\right)^2\right\}. \tag{3}$$

3 Computer experiments

For illustration of the gain of alternative forecast w.r.t. the classical one we used the Gaussian time series with the polymodal spectrum specified by the Bloomfield model of fifth order:

$$S(\lambda) = \exp\left(2\sum_{j=1}^{5} l_j \cos(j\lambda)\right), \tag{4}$$

$$(l_1, \dots, l_5) = (-0.00864, -1.15869, -1.97555, 0.06574, -3.36081).$$

The time series x_t was generated as $x_t = \sum_{j=0}^{123} c_j \xi_{t+j}$, using the standard discrete Gaussian white noise ξ_t and 124 pre-calculated moving average coefficients c_j, associated with the spectral density (4): $S(\lambda) = \left|\sum_{j\geq 0} c_j e^{ij\lambda}\right|^2$, $\lambda \in \mathbb{R}$, $i = \sqrt{-1}$. The plots of moving average coefficients c_j, spectral density (4) and time series realization x_t are presented in Figure 1.

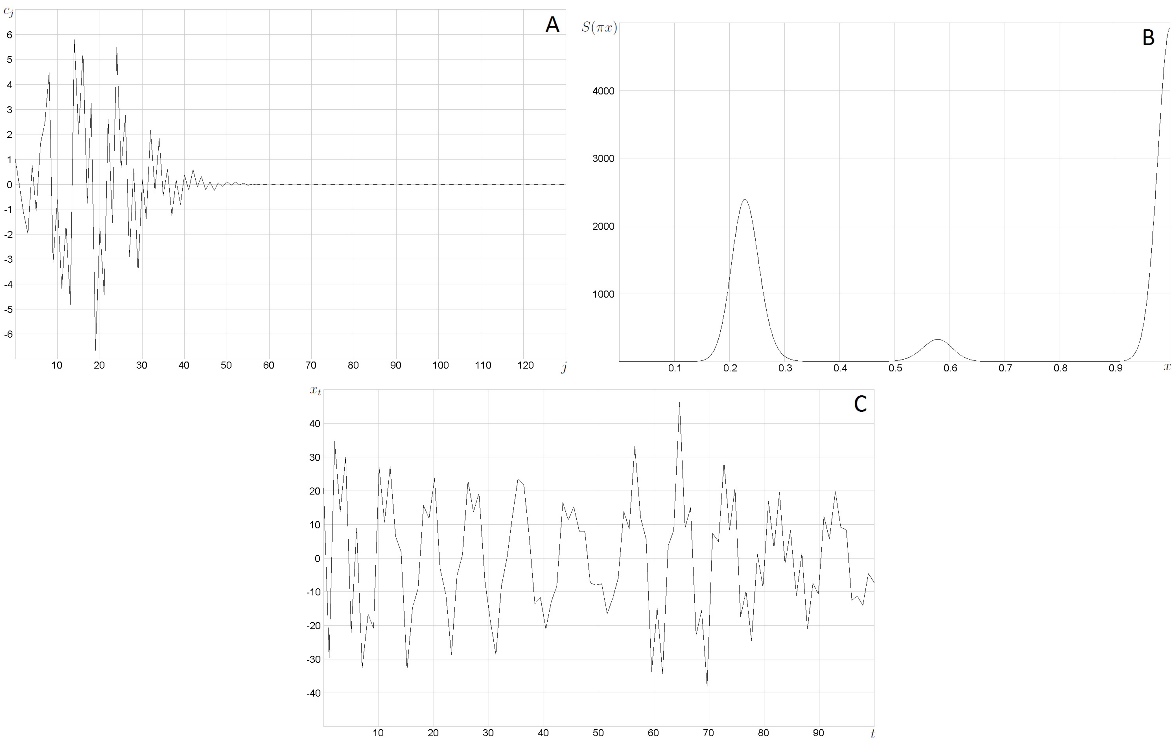

Figure 1: Moving average coefficients (A), spectral density (B) and time series realization (C)

The risks (2) and (3) was computed by Monte-Carlo method with 3000 replications. The number p of estimated parameters was set to five for both classical and alternative forecasts. In (1) the following kernels of smoothing were used [1]:

$$f_{\cos}(x) = \frac{1+\cos(\pi x)}{2}, \;\; f_{\text{const}}(x) \equiv 1, \;\; f_{\text{Parzen(q)}}(x) = 1 - |x|^q.$$

The optimal widths K of smoothing windows was manually selected for every kernel. The value $\varepsilon = 10^{-11}$ was chosen in a similar way. The plots of the Monte-Carlo estimates of the risks (2) and (3) on a logarithmic scale are presented in Figure 2; they illustrate significant gain of the risk (3) for the Bloomfield forecasting statistic w.r.t. the risk (2) for the classical one. Moreover, the best of alternative forecasts has the smallest optimal width of the smoothing window ($K = 12$) and the kernel $f(x)$, smooth both at $x = 0$ and at $x = 1$. The average two of alternative forecasts have close optimal widths of the smoothing windows ($K = 17$ and $K = 20$) and the kernels $f(x)$, smooth at $x = 0$ and nonsmooth at $x = 1$. Finally, the worst one of alternative forecasts have the largest optimal width of the smoothing window ($K = 77$) and the kernel $f(x)$, nonsmooth both at $x = 0$ and at $x = 1$. This trend suggests an obvious assumption on preferability of using kernels $f(x)$, totally smooth on $\mathbb{R}$.

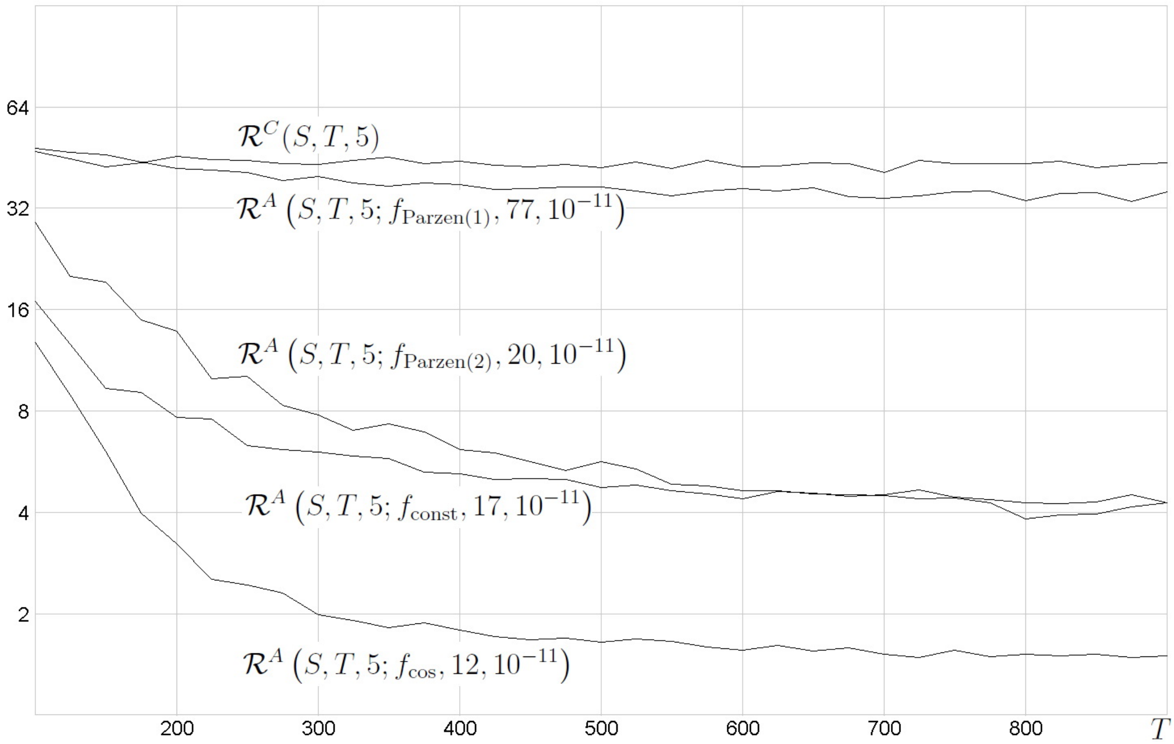

Figure 2: Risks of classical and alternative forecasts (logarithmic scale)

References

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

[2] Kharin Yu.S., Voloshko V.A. (2010). Stationary time series forecasting based on the small-parameter Bloomfield model. *Proceedings of the BAS*. Vol. **54**, No. **6**, pp. 27-32 (in russian).