STATISTICAL ANALYSIS OF THE SPECTRAL DENSITY ESTIMATE OBTAINED VIA COIFMAN SCALING FUNCTION

N.V. Semenchuk
Grodno State University
Grodno, BELARUS
e-mail: `n.semenchuk@grsu.by`

Abstract

Spectral density built as Fourier transform of covariance sequence of stationary random process is determining the process characteristics and makes for analysis of it's structure. Thus, one of the main problems in time series analysis is constructing consistent estimates of spectral density via successive, taken after equal periods of time observations of stationary random process. This article is devoted to investigation of problems dealing with application of wavelet analysis methods for solving task of spectral density nonparametric estimating of stationary random process with discrete time.

1 Introduction

Wavelet analysis is a new trend in data processing, which was developed independently in the fields of mathematics, physics and engineering. Interchanges between these fields during the last decades have led to many new wavelet applications, in particular, applications in mathematical statistics. There are few investigations done in the branch of spectral density estimating of stationary random process via wavelets. The idea of using periodogram expansion in series via periodic scaling functions and wavelets for spectral density estimate of one-dimensional stationary random process with discrete time was first suggested in article by M.H. Newmann [3]. In researches of M.H. Newmann, D.Donoho, I. Jonstone [2] behavior of nonlinear spectral density estimates constructed via wavelet coefficients thresholding methods is investigated. At the same time, analysis of statistical properties of linear wavelet estimates of spectral density and opportunity of their practical application with various wavelets are not adequately explored.

2 Wavelet spectral density estimates of stationary random processes with discrete time

Let's consider the problem of estimating unknown spectral density $f(\lambda),\ \lambda \in \Pi = [-\pi;\pi]$ via T successive observations $X(0), X(1), \ldots, X(T-1)$ of stationary random process $X(t),\ MX(t) = 0,\ t \in \mathbb{Z}$, obtained after equal time intervals. For the spectral density estimate we'll consider following statistic:

$$\hat{f}(\lambda) = \sum_{k=1}^{2^J} \hat{\alpha}_{J,k} \tilde{\varphi}_{J,k}(\lambda), \tag{1}$$

$$\hat{\alpha}_{J,k} = \int_{\Pi} I_T^{(h)}(\alpha) \tilde{\varphi}_{J,k}(\alpha) d\alpha, \tag{2}$$

$$\tilde{\varphi}_{J,k}(\lambda) = \sum_{n \in \mathbb{Z}} (2\pi)^{-1/2} \varphi_{J,k} \left((2\pi)^{-1} \lambda + n \right), \tag{3}$$

$\varphi_{J,k}(x) = 2^{J/2} \varphi \left(2^J x - k \right)$, $\varphi(x)$ – scaling function [1,3], $J \in \mathbb{N}$, $k = \overline{1, 2^J}$; $I_T^{(h)}(\lambda)$ – modified periodogram

$$I_T^{(h)}(\lambda) = \frac{1}{2\pi H_2^{(T)}(0)} \left| \sum_{t=0}^{T-1} h_T(t) X(t) e^{-i\lambda t} \right|^2, \tag{4}$$

$$H_2^{(T)}(\lambda) = \sum_{t=0}^{T-1} (h_T(t))^2 e^{-i\lambda t}, \tag{5}$$

function $h_T(t) = h\left(\frac{t}{T}\right)$, $h : [0,1] \to \mathbb{R}$ – data taper, $T \in \mathbb{N}$.

Statistic (1) is called linear wavelet estimate of spectral density. It is proved that estimate (1) is spectral density estimate of stationary random process consistent in mean-square sense.

In this article we'll used Coifman scaling function $\varphi(x) \in L_2(\mathbb{R})$ of the order M, $M \in \mathbb{N}$ which is of real value and has the following properties [1]:

$$\text{supp}\varphi(x) \subset [-M; 2M-1],$$

$$\int_{\mathbb{R}} \varphi(x) dx = 1, \qquad \int_{\mathbb{R}} x^m \varphi(x) dx = 0,$$

$m = \overline{1, M-1}$.

As Coifman scaling function has vanishing moments, coefficients $\hat{\alpha}_{J,k}$ may be approximated by evaluations of the function $I_T^{(h)}$ at dyadic points:

$$\hat{\alpha}_{J,k} \approx \frac{2\pi}{2^{J/2}} I_T^{(h)} \left(\frac{2\pi k}{2^J} \right), \tag{6}$$

$k = \overline{1, 2^J}$, $J \in \mathbb{N}$. Taking this fact into consideration, we'll substitute the coefficients $\hat{\alpha}_{J,k}$ with the approximations (6) in defining statistic $\hat{f}(\lambda)$ and shall define statistic

$$\hat{\hat{f}}(\lambda) = \frac{\sqrt{2\pi}}{2^{\frac{J}{2}}} \sum_{k=1}^{2^J} I_T^{(h)} \left(\frac{2\pi k}{2^J} \right) \tilde{\varphi}_{J,k}(\lambda) \tag{7}$$

$I_T^{(h)}(\lambda)$, $\tilde{\varphi}_{J,k}(\lambda)$ are defined by (3) and (4) correspondingly $J \in \mathbb{N}$, $k = \overline{1, 2^J}$.

3 Statistic behavior of the wavelet estimate (7)

Let us consider the behavior of the moments of suggested estimate (7). For the further investigation the following definition will be used:

Definition 1. *Given M, K and L, let $C_M(K,L)$ denote the set of functions $f(\lambda)$, $\lambda \in \Pi$, having $M+1$ bounded derivatives:*

$$\left|f^{(M+1)}(\lambda)\right| \leq K < \infty$$

$M \in \mathbb{N}_0$ and satisfying Lipschitz condition with the constant L (independent of λ).

Theorem 1. *Let $f(\lambda) \in C_M(L,K)$, $\lambda \in \Pi$, data tapers have total variation bounded by $V > 0$, Coifman scaling function $\varphi(x)$ of the order $M+1$ is absolutely bounded by the constant A, then the following inequality is valid:*

$$\left|E\hat{f}(\lambda) - E\hat{\hat{f}}(\lambda)\right| \leq R_T(M),$$

$$R_T(M) = \left(\frac{2C_1^2V^2L}{2\pi H_2^{(T)}(0)}\left[2\ln(\pi T)+1\right] + \frac{(2\pi)^{M+1}K(3M+3)^{M+2}A}{(M+2)!2^{J(M+1)}}\right)\int\limits_{\mathbb{R}}|\varphi(y)|dy, \quad (8)$$

$H_2^{(T)}(0)$ is defined by (5), $0 < C_1 \leq \pi$.

Theorem 2. *If conditions of Theorem 1 are satisfied, then for the bias of estimator (7) the following inequality takes place:*

$$\left|E\hat{\hat{f}}(\lambda) - f(\lambda)\right| \leq \frac{C_1^2V^2LD}{2\pi H_2^{(T)}(0)}\left[2\ln(\pi T)+1\right] +$$

$$+\frac{6\pi AL(M+1)}{2^J}\left[\int\limits_{\mathbb{R}}|z|\,|\varphi(z)|\,dz + 3(M+1)\int\limits_{\mathbb{R}}|\varphi(z)|\,dz\right] + R_T(M),$$

$$D = \sum_{m\in\mathbb{Z}}\left|\varphi\left(\frac{2^J\lambda}{2\pi} - m\right)\right|\int\limits_{\mathbb{R}}|\varphi(y)|\,dy < +\infty,$$

$H_2^{(T)}(0)$ and $R_T(M)$ are given by (5) and (8) correspondingly, $0 < C_1 \leq \pi$.

As an rate of convergence for dispersion of estimate (7) one can use relations obtained for dispersion of estimate (1) [3].

It is necessary to emphasize advantages of use of estimations (7). First of all, at calculation of estimates (7) it is possible to take advantage of algorithm fast Fourier transform. It will allow to improve speed of calculations, is especial at great values T. Second, the statistic (7) allows to receive a consistent estimate of spectral density in any point $\lambda \in \Pi$ as against similar classical estimates where we have a consistent estimate of spectral density only in points $\lambda_k = \frac{2\pi k}{T}$, for the integer k, $-\left[\frac{T}{2}\right] \leq k \leq \left[\frac{T}{2}\right]$, $\left[\frac{T}{2}\right]$ – the integer part of $\frac{T}{2}$.

4 Example

Let's consider an autoregressive process of p-th order $p \in \mathbb{N} = \{1, 2, ...\}$:

$$\sum_{j=0}^{p} a_j X_{t-j} = \varepsilon_t, \ (a_0 = 1), \tag{9}$$

where ε_t is a sequence of independent, identically distributed random variates, $t \in \mathbb{Z} = \{0, \pm1, \pm2, ...\}$. We'll assume that $\varepsilon_t \sim N(0, \sigma^2)$. For model (9) we can rewrite:

$$\sum_{j=0}^{p} a_j z^j = \prod_{j=1}^{p} (1 - u_j z), \ u_j \in \mathbb{C}.$$

If $u_j = \theta_j e^{-i\mu_j}$, then at $\theta_j \to 1$ spectral density $f(\lambda), \ \lambda \in \Pi$ of the random process (9) has peaks on a frequency $\mu_j \in \Pi, \ j = \overline{1, p}, \ p \in \mathbb{N}$. Let's consider the autoregressive processes of the 4^{th} order AR(4), where $\varepsilon_t \sim N(0, 1)$ and

$$u_1 = 0,9e^{-i0,8\pi}, \ u_2 = \overline{u_1}, \ u_3 = 0,9e^{-i0,12\pi}, \ u_4 = \overline{u_3}.$$

On figure 1 theoretical spectral density of the considering process and wavelet estimate (6) obtained via Coifman scaling function of 8^{th} order with $T = 256$ are shown. Data taper Risse, Bokhner and Parsen's, $h(x) = 1 - x^2, \ x \in [-1, 1]$, is used.

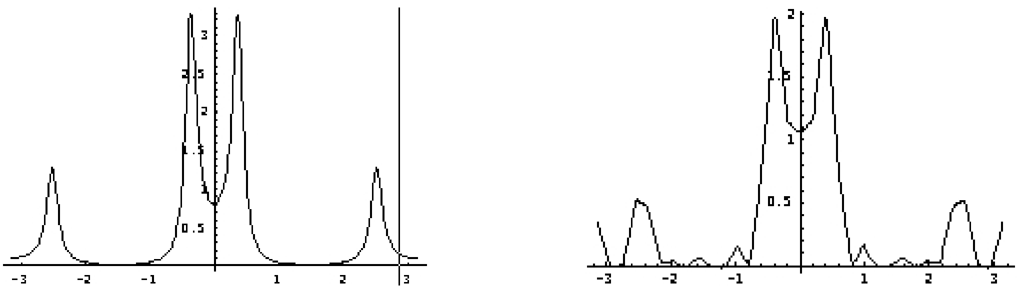

Figure 1: Theoretical spectral density and it's wavelet estimate for the process AR(4).

Thus, by numerical experiments we provide results which corroborate the theory.

References

[1] Daubechies I. (2001) *Ten lectures on wavelets*. Scientific-Research Centre "Regular and chaotic dynamic". Izhevsk.

[2] Donoho D., Johnstone I. (1994) Ideal spatial adaptation by wavelet shrinkage. *Biometrika* Vol. **81**, pp. 425-455.

[3] Neumann M.H. (1996) Spectral density estimation via nonlinear wavelet methods for stationary non-Gaussian time series. *J. Time Ser. Anal.* Vol. **17**, pp. 601-633.