Владикавказский математический журнал
январь–март, 2007, Том 9, Выпуск 1

UDC 517.984

SOME ASYMPTOTIC PROPERTIES OF A KERNEL SPECTRUM ESTIMATE WITH DIFFERENT MULTITAPERS

Teamah A. A. M., Bakouch H. S.

Let $X(t)$, $t = 0, \pm 1, \ldots$, be a zero mean real-valued stationary time series with spectrum $f_{XX}(\lambda)$, $-\pi \leqslant \lambda \leqslant \pi$. Given the realization $X(1)$, $X(2), \ldots, X(N)$, we construct L different multitapered periodograms $I_{XX}^{(mt)_j}(\lambda)$, $j = 1, 2, \ldots, L$, on non-overlapped and overlapped segments $X^{(j)}(t)$, $1 \leqslant t < N$. Also, we give asymptotic expressions of the mean and variance of the average of these different multitapered periodograms. We obtain an estimate of $f_{XX}(\lambda)$ via $I_{XX}^{(mt)_j}(\lambda)$ and different kernels $W_{\beta}^{(j)}(\alpha)$, $j = 1, 2, \ldots, L$; $-\pi < \alpha \leqslant \pi$; β is a bandwidth. We find asymptotic expressions of the first and second-order moments of this estimate. Moreover, we propose a choice of the considered bandwidth. An asymptotic expression of the integrated relative mean squared error (IMSE) of the estimate is formulated.

Mathematics Subject Classification (2000): 62M10, 62M15.

Key words: Stationary time series; Non-overlapped and overlapped segments; Multitapering; Kernels; Bandwidth; Spectrum estimate.

1. Introduction

Multitapering method maintains the good bias properties that tapering provides and at the same time produces an estimate with less variability (see [7, 11, 12, 14]). Some asymptotic statistical properties of spectral estimates were studied by several authors (see [1, 3, 4]) using a tapered data. The authors of this paper argued in [9, 10] the asymptotic expressions of the first and second-order moments of some spectral estimates, on non-overlapped and overlapped segments via different tapers and different weight functions (kernels) for both continuous time and discrete time stationary processes.

In this paper we study the problem of estimating a spectral density function (spectrum) on non-overlapped and overlapped segments using different multitapers and different kernels with a bandwidth parameter, for a discrete parameter stationary time series. In section 2 we introduce an estimate of the spectral density function using different multitapers and different kernels. Moreover, we give asymptotic expressions of the mean and variance of the average of the constructed different multitapered periodograms. In section 3 we obtain the asymptotic expressions of the mean and variance for the suggested estimate, assuming that direct spectral estimates are uncorrelated. Also, we obtain an optimal choice of the bandwidth. Furthermore, we formulate an asymptotic expression of the integrated relative mean squared error of the estimate.

2. The model

Suppose that $X(1), X(2), \dots, X(N)$ is a realization of N observations from a real-valued stationary and discrete parameter process $X(t)$, $t = 0, \pm 1, \dots$, with a zero mean. The spectral density function of $X(t)$ is

$$f_{XX}(\lambda) = \frac{1}{2\pi} \sum_{\tau=-\infty}^{\infty} C_{XX}(\tau) e^{-i\lambda\tau}, \quad -\pi \leqslant \lambda \leqslant \pi;\ i = \sqrt{-1}, \tag{1}$$

where $C_{XX}(\tau)$ is the autocovariance function of $X(t)$ and given by

$$C_{XX}(\tau) = \int_{-\pi}^{\pi} f_{XX}(\lambda) e^{i\lambda\tau} d\lambda, \tag{2}$$

provided that $\sum_{\tau=-\infty}^{\infty} |C_{XX}(\tau)| < \infty$.

If the process $X(t)$ is invertible, then the inverse spectral density function is defined by

$$f_{XX}^{-1}(\lambda) = \frac{1}{2\pi} \sum_{\tau=-\infty}^{\infty} d_{XX}(\tau) e^{-i\lambda\tau}, \tag{3}$$

where $d_{XX}(\tau)$ is the inverse autocovariance of $X(t)$ (see [2]) and given by

$$d_{XX}(\tau) = \int_{-\pi}^{\pi} f_{XX}^{-1}(\lambda) e^{i\lambda\tau}\, d\lambda,$$

such that $\sum_{\tau=-\infty}^{\infty} |d_{XX}(\tau)| < \infty$.

We construct L segments by dividing the given observations:

$$X^{(j)}(t) = X[(j-1)M + t], \quad j = 1, 2, \dots, L;\ t = 1, 2, \dots, M + q;\ 0 \leqslant q < M, \tag{4}$$

where $X^{(j)}(t)$ is the set of observations in the $j^{\underline{th}}$ segment. If $N = LM + q$, $0 < q < M$, then the number of overlapped segments $L = (N - q)/M$ and each segment contains $M + q$ observations. Also, if $q = 0$, then the number of non-overlapped segments $L = N/M$.

Now, we define the average of different multitapered periodograms as an estimate of $f_{XX}(\lambda)$:

$$\hat{f}_{XX}^{(mt)}(\lambda) = \frac{1}{L} \sum_{j=1}^{L} I_{XX}^{(mt)_j}(\lambda), \tag{5}$$

where $I_{XX}^{(mt)_j}(\lambda)$ is the multitapered periodogram of $X^{(j)}(t)$ and given by

$$I_{XX}^{(mt)_j}(\lambda) = \frac{1}{K} \sum_{k=1}^{K} I_{XX}^{(d)_{j,k}}(\lambda)$$

with

$$I_{XX}^{(d)_{j,k}}(\lambda) = \left[2\pi \sum_{t=1}^{M+q} h_{j,k}^2(t) \right]^{-1} \left| \sum_{t=1}^{M+q} h_{j,k}(t) X^{(j)}(t) e^{-i\lambda t} \right|^2.$$

$h_{j,k}(t)$ is called the data taper for the $(j,k)^{\underline{th}}$ direct spectral estimate $I_{XX}^{(d)_{j,k}}(\lambda)$ and equals zero outside the interval $[1, M+q]$ and K is the number of components of multitaper in each segment.

Let $h_{j,k}(t)$ be a set of orthonormal tapers, that is

$$\sum_{t=1}^{M+q} h_{j,r}(t)h_{j,k}(t) = \begin{cases} 1, & \text{if } r = k; \\ 0, & \text{otherwise,} \end{cases}$$

(see [12]). From the orthonormality we deduce that direct spectral estimates $I_{XX}^{(d)_{j,k}}(\lambda)$, $j = 1, 2, \ldots, L$; $k = 1, 2, \ldots, K$, are uncorrelated (see [13]), and then

$$I_{XX}^{(d)_{j,k}}(\lambda) = \frac{1}{2\pi} \left| \sum_{t=1}^{M+q} h_{j,k}(t) X^{(j)}(t) e^{-i\lambda t} \right|^2 .$$

In fact, the direct spectral estimate $I_{XX}^{(d)_{j,k}}(\lambda)$ has the asymptotic properties (see [8, 13])

$$E\left[I_{XX}^{(d)_{j,k}}(\lambda)\right] \approx f_{XX}(\lambda), \quad \mathrm{Var}\left[I_{XX}^{(d)_{j,k}}(\lambda)\right] \approx f_{XX}^2(\lambda). \tag{6}$$

Equations (5), (6) and the uncorrelation of direct spectral estimate $I_{XX}^{(d)_{j,k}}(\lambda)$ imply $E\left[\hat{f}_{XX}^{(mt)}(\lambda)\right] \approx f_{XX}(\lambda)$ and $\mathrm{Var}\left[\hat{f}_{XX}^{(mt)}(\lambda)\right] \approx f_{XX}^2(\lambda)/KL$. Obviously, $\mathrm{Var}\left[\hat{f}_{XX}^{(mt)}(\lambda)\right]$ becomes less variability as at least K or L increases. The case when $L = 1$ was investigated in [13].

Smoothing the multitapered periodograms $I_{XX}^{(mt)_j}(\lambda)$, $j = 1, 2, \ldots, L$, in equation (5) by the different kernels (weight functions) $W_\beta^{(j)}(\alpha)$, $-\pi < \alpha \leqslant \pi$, and taking their average, we get

$$\hat{f}_{XX}^{(mt)_{sp}}(\lambda) = \frac{1}{\beta L} \sum_{j=1}^{L} \int_{-\pi}^{\pi} W^{(j)}\left(\frac{\lambda - \mu}{\beta}\right) I_{XX}^{(mt)_j}(\mu)\, d\mu, \tag{7}$$

which is a smoothed estimate (kernel estimate) of $f_{XX}(\lambda)$ with $W_\beta^{(j)}(\alpha) = \frac{1}{\beta} W^{(j)}(\frac{\alpha}{\beta})$, such that $\int_{-\pi}^{\pi} W^{(j)}(\alpha) d\alpha = 1$; $W^{(j)}(-\alpha) = W^{(j)}(\alpha)$. β is called the bandwidth. Also, we can deduce that

$$\hat{f}_{XX}^{(mt)_{sp}}(\lambda) = \frac{1}{L} \sum_{j=1}^{L} \hat{f}_{XX}^{(mt)_j}(\lambda), \tag{8}$$

where

$$\hat{f}_{XX}^{(mt)_j}(\lambda) = \frac{1}{\beta} \int_{-\pi}^{\pi} W^{(j)}\left(\frac{\lambda - \mu}{\beta}\right) I_{XX}^{(mt)_j}(\mu)\, d\mu$$

is the $j^{\underline{th}}$ smoothed multitapered periodogram of $X^{(j)}(t)$.

3. Statistical properties of $\hat{f}_{XX}^{(mt)_{sp}}(\lambda)$

In this section we obtain the asymptotic expressions of expectation, variance and integrated relative mean squared error of the smoothed (kernel) spectrum estimate $\hat{f}_{XX}^{(mt)_{sp}}(\lambda)$:

3.1. Expected value. Taking expectation of equation (7), we get

$$E\left[\hat{f}_{XX}^{(mt)_{sp}}(\lambda)\right] \approx \frac{1}{\beta L} \sum_{j=1}^{L} \int_{-\pi}^{\pi} W^{(j)}\left(\frac{\lambda - \mu}{\beta}\right) f_{XX}(\mu)\, d\mu. \tag{9}$$

Making use of the transformation $\mu = \lambda + \beta\alpha$, with small β; $\lambda \in (-\pi, \pi]$, then equation (9) becomes

$$E\left[\hat{f}_{XX}^{(mt)_{sp}}(\lambda)\right] \approx \frac{1}{L}\sum_{j=1}^{L}\int_{-\infty}^{\infty} W^{(j)}(\alpha) f_{XX}(\lambda+\beta\alpha)\,d\alpha, \tag{10}$$

from a Taylor expansion for $f_{XX}(\lambda+\beta\alpha)$ about λ, equation (10) has the form:

$$E\left[\hat{f}_{XX}^{(mt)_{sp}}(\lambda)\right] = \frac{1}{L}\sum_{j=1}^{L}\int_{-\infty}^{\infty} W^{(j)}(\alpha)\left[f_{XX}(\lambda) + \beta\alpha f'_{XX}(\lambda) + \frac{\beta^2\alpha^2}{2} f''_{XX}(\lambda) + O(\beta^2)\right] d\alpha =$$

$$= f_{XX}(\lambda) + \frac{\beta^2}{2L} f''_{XX}(\lambda)\sum_{j=1}^{L}\int_{-\infty}^{\infty} \alpha^2 W^{(j)}(\alpha)\,d\alpha + O(\beta^2), \tag{11}$$

where $f''_{XX}(\lambda)$ is the second derivative of the spectrum $f_{XX}(\lambda)$. Therefore

$$\operatorname{Bias}\left[\hat{f}_{XX}^{(mt)_{sp}}(\lambda)\right] \approx \frac{\beta^2}{2L} f''_{XX}(\lambda)\sum_{j=1}^{L}\int_{-\infty}^{\infty} \alpha^2 W^{(j)}(\alpha)\,d\alpha. \tag{12}$$

It is clear that the bias of $\hat{f}_{XX}^{(mt)_{sp}}(\lambda)$ is of the order $\beta^2 L^{-1}$.

3.2. Variance. Since the tapered periodogram ordinates $I_{XX}^{(d)_{j,k}}(\mu_\ell)$, $1 \leqslant \ell < (M+q)/2$, are asymptotically independent (see [6, 8]), then equation (8) can be put in the form:

$$\hat{f}_{XX}^{(mt)_j}(\lambda) = \frac{1}{\beta K}\sum_{\ell}\sum_{k=1}^{K}\int_{\mu_{\ell-1}}^{\mu_\ell} W^{(j)}\left(\frac{\lambda-\mu}{\beta}\right) I_{XX}^{(d)_{j,k}}(\mu)\,d\mu \approx$$

$$\approx \frac{2\pi}{\beta K(M+q)}\sum_{\ell}\sum_{k=1}^{K} W^{(j)}\left(\frac{\lambda-\mu_\ell}{\beta}\right) I_{XX}^{(d)_{j,k}}(\mu_\ell).$$

Using equations (6) and (8), we get

$$\operatorname{Var}\left[\hat{f}_{XX}^{(mt)_{sp}}(\lambda)\right] \approx \frac{4\pi^2}{\beta^2 L^2 K(M+q)^2}\sum_{j=1}^{L}\sum_{\ell}\left[W^{(j)}\left(\frac{\lambda-\mu_\ell}{\beta}\right)\right]^2 f_{XX}^2(\mu_\ell) \approx$$

$$\approx \frac{2\pi}{\beta^2 L^2 K(M+q)}\sum_{j=1}^{L}\int_{-\pi}^{\pi}\left[W^{(j)}\left(\frac{\lambda-\mu}{\beta}\right)\right]^2 f_{XX}^2(\mu)\,d\mu.$$

Putting $\mu = \lambda + \beta\alpha$,

$$\begin{aligned}\operatorname{Var}\left[\hat{f}_{XX}^{(mt)_{sp}}(\lambda)\right] &= \frac{2\pi}{\beta K(M+q)L^2}\sum_{j=1}^{L}\int_{-\infty}^{\infty}\left[W^{(j)}(\alpha)\right]^2 f_{XX}^2(\lambda+\beta\alpha)\,d\alpha \approx \\ &\approx \frac{2\pi}{\beta K(M+q)L^2} f_{XX}^2(\lambda)\sum_{j=1}^{L}\int_{-\infty}^{\infty}\left[W^{(j)}(\alpha)\right]^2 d\alpha,\end{aligned} \tag{13}$$

which is of the order $\left[\beta K L^2(M+q)\right]^{-1}$.

From equations (12) and (13) the mean squared error (MSE)of $\hat{f}_{XX}^{(mt)_{sp}}(\lambda)$ is

$$\mathrm{MSE}\left[\hat{f}_{XX}^{(mt)_{sp}}(\lambda)\right] = \left(\mathrm{Bias}\left[\hat{f}_{XX}^{(mt)_{sp}}(\lambda)\right]\right)^2 + \mathrm{Var}\left[\hat{f}_{XX}^{(mt)_{sp}}(\lambda)\right].$$

The MSE of an estimate can be small only if both bias term and variance term are small. We show that the two terms are of the orders $\beta^2 L^{-1}$ and $[\beta K L^2(M+q)]^{-1}$. Then it follows that the variance and the squared bias terms of $\hat{f}_{XX}^{(mt)_{sp}}(\lambda)$ are balanced for $[\beta K L^2(M+q)]^{-1} \approx \beta^4 L^{-2}$. This implies an optimal choice of bandwidth equals to $\beta \approx [K(M+q)]^{-1/5}$. Hence, $\beta \to 0$ as $M \to \infty$. Using equations (12), (13) and the optimal choice of β, we get

$$\mathrm{Bias}\left[\hat{f}_{XX}^{(mt)_{sp}}(\lambda)\right], \quad \mathrm{Var}\left[\hat{f}_{XX}^{(mt)_{sp}}(\lambda)\right] \to 0, \quad M \to \infty, \tag{14}$$

that is $\hat{f}_{XX}^{(mt)_{sp}}(\lambda)$ is a consistent estimate of $f_{XX}(\lambda)$ as $M \to \infty$. Also, $\mathrm{Var}\left[\hat{f}_{XX}^{(mt)_{sp}}(\lambda)\right]$ becomes less variability as at least K or L increases.

3.3. Integrated relative mean squared error. We take IMSE as a measure for the goodness of fit of a spectral estimate. IMSE of $\hat{f}_{XX}^{(mt)_{sp}}(\lambda)$ is defined by (see [5]):

$$\begin{aligned}\mathrm{IMSE}\left[\hat{f}_{XX}^{(mt)_{sp}}(\lambda)\right] &= \int_{-\pi}^{\pi} E\left[\frac{\hat{f}_{XX}^{(mt)_{sp}}(\lambda) - f_{XX}(\lambda)}{f_{XX}(\lambda)}\right]^2 d\lambda = \\ &= \int_{-\pi}^{\pi}\left[\frac{E\left[\hat{f}_{XX}^{(mt)_{sp}}(\lambda)\right]}{f_{XX}(\lambda)} - 1\right]^2 d\lambda + \int_{-\pi}^{\pi} \mathrm{Var}\left[\frac{\hat{f}_{XX}^{(mt)_{sp}}(\lambda)}{f_{XX}(\lambda)}\right] d\lambda.\end{aligned} \tag{15}$$

Hence,

$$\begin{aligned}\mathrm{IMSE}\left[\hat{f}_{XX}^{(mt)_{sp}}(\lambda)\right] &= 2\pi + \int_{-\pi}^{\pi}\left[f_{XX}^{-1}(\lambda)\right]^2\left(E\left[\hat{f}_{XX}^{(mt)_{sp}}(\lambda)\right]\right)^2 d\lambda - \\ &- 2\int_{-\pi}^{\pi} f_{XX}^{-1}(\lambda) E\left[\hat{f}_{XX}^{(mt)_{sp}}(\lambda)\right] d\lambda + \int_{-\pi}^{\pi}\left[f_{XX}^{-1}(\lambda)\right]^2 \mathrm{Var}\left[\hat{f}_{XX}^{(mt)_{sp}}(\lambda)\right] d\lambda.\end{aligned}$$

From equations (3), (11) and (13), then formula (15) has the form:

$$\begin{aligned}\mathrm{IMSE}\left[\hat{f}_{XX}^{(mt)_{sp}}(\lambda)\right] &\approx 2\pi + (2\pi)^{-2}\sum_{\tau=-\infty}^{\infty}\sum_{r=-\infty}^{\infty} d_{XX}(\tau)\, d_{XX}(r) \times \\ &\times \int_{-\pi}^{\pi}\left[f_{XX}(\lambda) + \frac{\beta^2}{2L} f''_{XX}(\lambda)\sum_{j=1}^{L}\int_{-\infty}^{\infty}\alpha^2 W^{(j)}(\alpha)d\alpha\right]^2 e^{-i\lambda(\tau+r)} d\lambda - \\ &- \pi^{-1}\sum_{\tau=-\infty}^{\infty} d_{XX}(\tau)\int_{-\pi}^{\pi}\left[f_{XX}(\lambda) + \frac{\beta^2}{2L} f''_{XX}(\lambda)\sum_{j=1}^{L}\int_{-\infty}^{\infty}\alpha^2 W^{(j)}(\alpha)d\alpha\right] e^{-i\lambda\tau} d\lambda + \\ &+ \frac{(2\pi)^{-1}}{\beta K L^2(M+q)}\sum_{\tau=-\infty}^{\infty}\sum_{r=-\infty}^{\infty} d_{XX}(\tau) d_{XX}(r)\int_{-\infty}^{\infty}\int_{-\pi}^{\pi} f_{XX}^2(\lambda)\sum_{j=1}^{L}\left[W^{(j)}(\alpha)\right]^2 e^{-i\lambda(\tau+r)}\, d\lambda\, d\alpha.\end{aligned} \tag{16}$$

If we consider the optimal choice of β, $\beta \approx [K(M+q)]^{-1/5}$, then equation (16) can be put in the form:

$$\mathrm{IMSE}\left[\hat{f}_{XX}^{(mt)_{sp}}(\lambda)\right] \approx 2\pi + (2\pi)^{-2} \sum_{\tau=-\infty}^{\infty} \sum_{r=-\infty}^{\infty} d_{XX}(\tau) d_{XX}(r) \left[\int_{-\pi}^{\pi} f_{XX}^2(\lambda) e^{-i\lambda(\tau+r)} d\lambda\right] -$$

$$- \pi^{-1} \sum_{\tau=-\infty}^{\infty} d_{XX}(\tau) \left[\int_{-\pi}^{\pi} f_{XX}(\lambda) e^{-i\lambda\tau} d\lambda\right] \approx 2\pi - \pi^{-1} \sum_{\tau=-\infty}^{\infty} d_{XX}(\tau) C_{XX}(-\tau) +$$

$$+ (2\pi)^{-2} \sum_{\tau=-\infty}^{\infty} \sum_{r=-\infty}^{\infty} d_{XX}(\tau) d_{XX}(r) \left[\int_{-\pi}^{\pi} f_{XX}^2(\lambda) e^{-i\lambda(\tau+r)} d\lambda\right], \; M \to \infty. \quad (17)$$

References

1. *Brillinger D. R.* Time Series: Data Analysis and Theory.—New York: Holt, Rinehart and Winston, 1975.
2. *Cleveland W. S.* The inverse autocorrelations of a time // Series and their applications. Technometrics.—1972.—V. 14.—P. 277–293.
3. *Ghazal M. A., Farag E. A.* Asymptotic distribution of spectral density estimate of cotinuous time series on crossed intervals // The Egyptian Statistical J.—Cairo, 1998.—V. 42, № 2.—P. 197–214.
4. *Ghazal M. A.* On a spectral density estimate on non-crossed intervals observation // International J. of Applied Math.—1999.—V. 1, № 8.—P. 875–882.
5. *Hurvich C. M.* A mean squared error criterion for time series data windows // Biometrika.—1988.—V. 75, № 3.—P. 485–490.
6. *Koopmans L. H.* The Spectral Analysis of Time Series.—New York: Acad. press, 1974.
7. *McCoy E. J.* Some New Statistical Approaches to the Analysis of Long Memory Processes: Ph. D. Thesis.—Imperial College, Univ. of London, 1994.
8. *Smith R. L.* Time Series, a course of Statistics.—Chapel Hill: North Carolina Univ., 1999.
9. *Teamah A. A. M., Bakouch H. S.* Statistical analysis on the average of periodograms with different tapers and spectral estimates of continuous time processes // AMSE Periodicals, Advances in Modeling and Analysis, Ser. D—2003.—V. 8, № 2.—P. 1–12.
10. *Teamah A. A. M., Bakouch H. S.* Asymptotic statistical properties of spectral estimates with different tapers for discrete time processes // Applied Mathematics and computation.—2004.—V. 150.—P. 681–695.
11. *Thomson D. J.* Spectrum estimation and harmonic analysis // Proc. IEEE.—1982.—V. 70.—P. 1055–1096.
12. *Walden A. T.* Some advances in non-parametric multiple time series and spectral analysis // Environmetrics.—1994.—V. 5.—P. 281–295.
13. *Walden A. T., McCoy E. J., Percival D. B.* The variance of multitaper spectrum estimates for real Gaussian processes // IEEE Trans. Sig. Proces.—1994.—V. 42, № 2.—P. 479–482.
14. *Walden A. T., McCoy E. J., Percival D. B.* The effective bandwidth of a multitaper spectral estimator // Biometrika.—1995.—V. 82, № 1.—P. 201–214.

Received 12 March, 2006

Teamah A. A. M.
Department of Mathematics, Faculty of Science,
Tanta University, Tanta, EGYPT
E-mail: abdelmoneim2002uk@yahoo.co.uk

Hassan Salah Bakouch
Department of Mathematics, Faculty of Science,
Tanta University, Tanta, EGYPT
E-mail: hbakouch@yahoo.com