مجلهٔ پژوهش‌های آماری ایران
پژوهشکدهٔ آمار دورهٔ ۲، شمارهٔ ۲، پاییز و زمستان ۱۳۸۴، صص ۱۹۸-۲۱۸

J. Statist. Res. Iran **2** (2005): 107-127

Spectral Estimation of Stationary Time Series: Recent Developments

A. R. Nematollahi

Shiraz University

Abstract. Spectral analysis considers the problem of determining (the art of recovering) the spectral content (i.e., the distribution of power over frequency) of a stationary time series from a finite set of measurements, by means of either nonparametric or parametric techniques. This paper introduces the spectral analysis problem, motivates the definition of power spectral density functions, and reviews some important and new techniques in nonparametric and parametric spectral estimation. We also consider the problem in the context of multivariate time series.

Keywords. spectral density; Capon's estimate; high resolution estimate; block-Toeplitz matrix; windows.

1 Introduction

Time series analysis is permeated with both engineering and statistical concepts and terminology, the former being associated with the "spectral" or "frequency domain" approach, and the latter with the "correlation" or "time domain" approach, to the analysis of time series. Many statisticians find it difficult to work with the ideas of energy, power and frequency, while engineers may find it equally different to challenge with the statistical inference. In this paper we have chosen to emphasize the spectral approach to time series. Spectral analysis considers the problem of determining the distribution of total power over

frequency, and the spectral estimation problem tries to estimate this distribution from a finite record of a data sequence, by means of either nonparametric or parametric techniques.

Spectral analysis have applications in many diverse fields. In economics, meteorology, astronomy and geology the spectral analysis may reveal hidden periodicities, which are to be associated with cyclic behavior or recurring processes. In radar and sonar systems, the spectral contents of the received signals provide information on the location of the sources (or targets) situated in the field of view. In medicine, spectral analysis of various signals measured from a patient, such as electrocardiogram (ECG) or electroencephalogram (EEG) signals, can provide useful material for diagnosis. In seismology, the spectral analysis of the signals recorded prior to and during a seismic event (such as a volcano eruption or an earthquake) gives useful information on the ground movement associated with such events and may help in predicting them. Seismic spectral estimation is also used to predict subsurface geologic structure in gas and oil exploration. In control systems, there is a resurging interest in spectral analysis methods as a means of characterizing the dynamical behavior of a given system. In hydrology, the effect of monitoring time intervals on computed hydrologic delay times of the karstic system is important for analysis and study of behaviors of different karstic hydrological systems. The previous and other applications of spectral analysis are reviewed in Kay (1988), Marple (1987), Bloomfield (1976), Bracewell (1986), Haykin (1991), Haykin (1995), Koopmans (1974), Priestley (1989), Percival and Walden (1993), Porat (1994), Scharf (1991), Therrien (1992), Proakis et al. (1992), Larson et al. (2003), Scargle (1997), Jakowatz, et al. (1996), DeGraaf (1998), Gini and Lombardini (2002), Rahnemaee et al. (2005), and Stoica and Moses (1997, 2005).

The history of spectral analysis as an established discipline started more than one century ago with the work by Schuster (1898) on detecting cyclic behavior and hidden periodicities in time series, and so he called his statistics the periodogram. Marple (1987) notes that the word "spectrum" was apparently introduced by Newton in relation to his studies of the decomposition of white light into a band of light colors, when passed through a glass prism. This word appears to be a variant of the Latin word "specter" which means "ghostly apparition". The contemporary English word that has the same meaning as the original Latin word is "spectre" (Stoica and Moses, 1997). For more information about spectral analysis one can refer to Marple (1987), Kay (1988), Stoica and Moses (1997), Priestly (1989), Brockwell and Davis (1991), Brillinger (1981), Chatfield (1975), and Stoica and Moses (2005).

In general, the methods of estimation of the spectrum can be grouped into two categories: non-parametric methods and parametric methods (see Stoica and Moses, 2005; Priestley, 1989; and Brockwell and Davis, 1991). The AR and

ARMA spectral estimates and the maximum entropy methods of estimation come into the second category, i.e. parametric methods. Here the spectrum is estimated via a model and the main problem is the identification of the model given the data. Windows (kernel) estimates are non-parametric. Here the main problems involved are the choice of bandwidth and the choice of a suitable windows.

This paper introduces the spectral analysis problem, motivates the definition of power spectral density functions, and reviews some important and new techniques in nonparametric and parametric spectral estimation. We also consider the problem in the context of multivariate time series.

2 Preliminaries

Let $\{X_t\}$ be a discrete parameter, zero mean and real stationary time series and let $E|X_t|^2 < \infty$, $t \in \mathbb{Z}$, where $\mathbb{Z}$ stands for all integers. Suppose $R(\tau) = E(X_\tau X_{t+\tau})$, $\tau \in \mathbb{Z}$, is the autocovariance function of X_t satisfying $\sum_{\tau=-\infty}^{\infty} |R(\tau)| < \infty$. The power spectral density function of $\{X_t\}$ is defined by

$$h(\omega) = \frac{1}{2\pi} \sum_{\tau=-\infty}^{\infty} R(\tau) \exp\{-i\omega\tau\}, \quad -\infty < \omega < \infty \tag{1}$$

where ω, number of radians per unit times, is sometimes called the angular frequency, but in keeping with most authors we will simply call ω the frequency. Some authors refer to frequency as $f = \omega/2\pi$, number of cycles per unit time, since it is much easier to interpret from a physical point of view. Here, we will usually use ω, angular frequency, in mathematical formulas for conciseness, and use f, frequency, for the interpretation of data. The period is clearly $T = 2\pi/\omega = 1/f$. The absolute summability of $R(\tau)$, $(\sum_{\tau=-\infty}^{\infty} |R(\tau)| < \infty)$, implies that the above series converges absolutely and $h(\omega)$ is a bounded uniformly continuous function. It is easy to show that $h(\omega)$ is non-negative, even, and also a periodic function with period 2π. Hence $h(\omega)$ is completely described by its variation in the interval $\omega \in [0, \pi]$. Also the relation (1) may be inverted and the autocovariance function $R(\tau)$ expressed as

$$R(\tau) = \int_{-\pi}^{\pi} h(\omega) \exp\{i\omega\tau\} d\omega, \qquad \tau \in \mathbb{Z} \tag{2}$$

In particular, setting $\tau = 0$ give

$$R(0) = \mathrm{var}(X_t) = \int_{-\pi}^{\pi} h(\omega)\ d\omega. \tag{3}$$

The adjective power, which is sometimes prefixed to spectral density function derives from the engineer's use of the word in connection with the passage of an electric current through a resistance. For a sinusoidal input, the power is directly proportional to the squared amplitude of the oscillation. For a more general input, the power spectral density describes how the power is distributed over frequency. Mathematically, the area under $h(\omega)$ represent the average power, as indicated by (3), (Chatfield, 1975). Therefore $h(\omega)$ is a density function (power per unit of frequency) that represents the distribution of power with frequency (Brillinger, 1981). The term power spectral density function is often shortened to spectrum. Note that, the physical meaning of the spectrum is that $h(\omega)d\omega$ represents the contribution to variance of components with frequencies in the range $(\omega, \omega + d\omega)$. According to equation (3), the total area underneath the curve is equal to the variance of the process. A peak in the spectrum indicates an important contribution to variance at frequencies in the appropriate region. Note that the autocovariance function and the spectrum are equivalent ways of describing a stationary stochastic process. From a practical point of view, they are complementary to each other. Both functions contain the same information, but express it in different ways. In some situations a time-domain approach based on the autocovariance function is more useful while in other situations a frequency-domain approach based on spectrum is preferable.

3 Nonparametric Methods

A common spectral estimator is based on a function called periodogram. One of the first uses of periodogram has been in determining possible hidden periodicities in time series, which may be seen as a motivation for the name of this method (Schuster, 1898). The basic ideas underlying periodogram analysis may be explained heuristically as follows. For each positive integer n define

$$h_n(\omega) = \frac{1}{2\pi n} E\left(\left| \sum_{t=1}^{n} X_t \exp\{-it\omega\} \right|^2 \right). \tag{4}$$

It can be shown that $h_n(\omega) \to \frac{1}{2\pi} \sum_{\tau=-\infty}^{\infty} R(\tau) \exp\{-i\omega\tau\} = h(\omega)$ as $n \to \infty$ (see Brockwell and Davis, 1991, p. 343). So $I_n(\omega) = \frac{1}{n} \left| \sum_{t=1}^{n} X_t e^{-it\omega} \right|^2$, which is called periodogram, is asymptotically unbiased estimator of $2\pi h(\omega)$, a natural estimate.

Let $I_n(\omega) = \frac{1}{n} \left| \sum_{t=1}^{n} X_t \exp\{-it\omega\} \right|^2$ be the periodogram of data set

$\{X_1, X_2, \ldots, X_n\}$. Then for each ω, we may write $I_n(\omega)$ in the alternative form

$$\begin{aligned} I_n(\omega) &= \frac{1}{n}\left|\sum_{t=1}^{n} X_t \exp\{-it\omega\}\right|^2 \\ &= \frac{1}{n}\left(\sum_{t=1}^{n} X_t \exp\{-it\omega\}\right)\left(\sum_{r=1}^{n} X_r \exp\{ir\omega\}\right) \\ &= \sum_{\tau=-n-1}^{n-1} \widehat{R}(\tau)\exp\{-i\tau\omega\}, \qquad \tau = t-r \\ &= \widehat{R}(0) + 2\sum_{\tau=1}^{n-1} \widehat{R}(\tau)\exp\{-i\tau\omega\}, \qquad \tau = t-r \end{aligned} \tag{5}$$

where $\widehat{R}(\tau)$ is the (biased) sample autocovariance function of $R(\tau)$ at lag τ, i.e.

$$\widehat{R}(\tau) = \frac{1}{n}\sum_{t=1}^{n-|\tau|} X_t X_{t+|\tau|}. \tag{6}$$

From (5) we have that the periodogram is the discrete Fourier transform of the complete sample autocovariance function. The striking resemblance between (5) and the expression $h(\omega) = \frac{1}{2\pi}\sum_{\tau=-\infty}^{\infty} R(\tau)\exp\{-i\omega\tau\}$ for the spectral density of a stationary time series with $\sum_{\tau=-\infty}^{\infty}|R(\tau)| < \infty$ suggests the potential value of the periodogram for spectral density estimation (indeed, if the sample autocovariance function, $\widehat{R}(\tau)$, of the observations $\{X_1, X_2, \ldots, X_n\}$ can be regarded as a sample analogue of $R(\tau)$, so can the periodogram, $I_n(\cdot)$, of the observations be regarded as a sample analogue of $2\pi h(\omega)$.

A critical disadvantage of the periodogram as an estimate of the power spectrum $h(\omega)$ is that its variance is approximately $h^2(\omega)$, under reasonable regularity conditions, even when based on a lengthy stretch of data (Brillinger, 1981). In particular, for each $\omega \in [0, \pi]$, and $\varepsilon > 0$,

$$P\left(|I_n(\omega) - 2\pi h(\omega)| > \varepsilon\right) \to p > 0,$$

as $n \to \infty$. This means $I_n(\omega)$ is not a consistent estimator of $2\pi h(\omega)$. Thus $I_n(\omega)$ is an asymptotically unbiased estimator of $2\pi h(\omega)$, but of course, it is not consistent. So no matter how large n is taken, the variance will tend to remain at the level $h^2(\omega)$, (the variance does not decrease as n increases), and if an estimate with a variance smaller than this is desired, it is not to be obtained by simply increasing the sample length and continuing to use the periodogram. The lack of consistency is perhaps not too surprising, because the graph of

the periodogram usually fluctuates wildly. So for a process with a continuous spectrum, the periodogram provides a poor estimate and needs to be modified.

One way of obtaining an expression with a reduced variance is simply to omit some of the terms in (5). If we do this, we will certainly reduce the variance, but on the other hand the terms omitted will affect the expected value of the new expression, and the general effect will be to increase the bias. However, we know that if the process has a purely continuous spectrum, then $R(\tau) \to 0$, as $|\tau| \to \infty$, and hence if we omit only those terms which correspond to the tail of the sample autocovariance function, then hopefully the bias will not be affected too seriously. These ideas suggest that we might consider as an estimate of $h(\omega)$ an expression of the form

$$\widehat{h}_0(\omega) = \frac{1}{2\pi} \sum_{\tau=-M}^{M} \widehat{R}(\tau) \exp\{-i\tau\omega\}, \tag{7}$$

where $M(< n)$ is some integer whose precise value is as yet unspecified, and is called the truncated point or bandwidth (Priestley, 1989). The precision of the $\widehat{R}(\tau)$ decreases as τ increases, so that it would seem intuitively reasonable to give less weight to the values of $\widehat{R}(\tau)$ as τ increases. An estimator with this property is

$$\widehat{h}(\omega) = \frac{1}{2\pi} \sum_{\tau=-M}^{M} \lambda(\tau)\widehat{R}(\tau) \exp\{-i\tau\omega\}, \tag{8}$$

where $\{\lambda(\tau)\}$ is a set of weights called the lag windows. The reason for this rather unusual terminology stems from the fact that we often think $\lambda(\tau)$ as being effectively zero outside a small interval, say $(-\varepsilon, \varepsilon)$, and hence the summation (8) may be regarded as giving a view of $\widehat{R}(\tau)\exp\{-i\tau\omega\}$ through a narrow window. The term window was first introduced by Blackman and Tukey (1959). The lag window is an even, piecewise continuous function of λ satisfying the conditions $\lambda(0) = 1$, $|\lambda(\tau)| \leqslant 1$ for all τ, and $\lambda(\tau) = 0$ for $|\tau| > M$. Also the bandwidth M satisfy $M \to \infty$ and $M/n \to 0$ as $n \to \infty$, which simply mean that the number of terms in the weighted series (8) goes to ∞ as $n \to \infty$, while at the same time the width of the frequency interval over which the average is taken goes to zero. The estimator $\widehat{h}(\omega)$ given by (8) is called the lag window spectral density estimator.

In order to use the above estimator, we must choose a suitable lag window and a suitable truncation point. Some lag windows which are in common use today are:

i) The truncated or rectangular window:

$$\lambda(\tau) = \begin{cases} 1, & |\tau| \leqslant M \\ 0, & |\tau| > M \end{cases}$$

Here, all the autocovariances up to lag M are given equal weight. Using this window, (8) reduces to (7).

ii) The Bartlett or triangular window:

$$\lambda(\tau) = \begin{cases} 1 - \frac{|\tau|}{M}, & |\tau| \leqslant M \\ 0, & |\tau| > M \end{cases}$$

Here, we apply linearly decreasing weights to the autocovariances up to lag M and zero weights thereafter.

iii) The Tukey-Hamming or Blackman-Tukey window:

$$\lambda(\tau) = \begin{cases} 0.54 + 0.46\cos\left(\frac{\pi\tau}{M}\right), & |\tau| \leqslant M \\ 0, & |\tau| > M \end{cases}$$

iv) The Parzen window:

$$\lambda(\tau) = \begin{cases} 1 - 6\left(\frac{\tau}{M}\right)^2 + 6\left(\frac{|\tau|}{M}\right)^3, & |\tau| \leqslant \frac{M}{2} \\ 2\left(\frac{1-|\tau|}{M}\right)^3, & \frac{M}{2} \leqslant |\tau| \leqslant M \\ 0, & |\tau| > M \end{cases}$$

v) The Bartlett-Priestly window:

$$\lambda(\tau) = \frac{3M^2}{(\pi\tau)^2}\left(\frac{\sin\left(\frac{\pi\tau}{M}\right)}{\frac{\pi\tau}{M}} - \cos\left(\frac{\pi\tau}{M}\right)\right).$$

The Bartlett-Priestly window was derived by minimizing an approximate expression for the relative mean square error of the estimated spectral density function with respect to the functional form of the window (Priestley, 1962; Bartlett, 1963). Priestly showed that this window is optimal within a somewhat restricted class of windows, but its optimality with respect to the relative mean square error was later established more generally by Epanechnikov (1969).

The choice of the truncation point (bandwidth), M, is rather difficult and little clear-cut advice is available in the literature. The smaller the value of M, the smaller will be the variance of $\widehat{h}(\omega)$, but the larger will be the bias. If M is too small, important features may be smoothed out, while if M is too large, the behaviour of $\widehat{h}(\omega)$ becomes more like that of periodogram with erratic variation, i.e.

$$\text{as } M\uparrow,\ \text{variance}\uparrow,\ \text{bias}\downarrow; \qquad \text{as } M\downarrow,\ \text{variance}\downarrow,\ \text{bias}\uparrow$$

(see Priestley, 1989, p. 517). Thus a compromise value must be chosen. A useful rough guide is to choose M to be about $2\sqrt{n}$. This choice of M ensures the asymptotic situation that as $n \to \infty$, so also does $M \to \infty$ but in such a way that $M/n \to 0$. But, as Hannan (1970, p. 311) says, experience is the real teacher and that cannot be got from a book.

The above methods are based on transforming the sample autocovariance function. An alternative approach is to smooth the periodogram by simple grouping the periodogram ordinates in a set and finding their average value. This approach is based on a suggestion by Daniell in 1946. It can be shown that for large n the periodogram ordinates at fixed frequencies ($I_n(\omega_j)$, where $\omega_j = 2\pi j/n$) are approximately independent with variances changing only slightly over small frequency interval. So we might hope to construct a consistent estimator of $h(\omega)$ by averaging the periodogram estimates in a small frequency interval containing ω (just as we obtain a consistent estimator of a population mean by averaging the observed values in a random sample of size n). For example, consider a simple moving average filter by smoothing the series $\{I_n(\omega_j)\}$, i.e., $(1/2\pi)\sum_{|k|\leqslant M}\{1/(2M+1)\}I_n(\omega_{j+k})$, or more generally,

$$\widetilde{h}(\omega_j) = \frac{1}{2\pi}\sum_{|k|\leqslant M} W_n(k) I_n(\omega_{j+k}), \tag{9}$$

where in order this estimate to be consistent, we must impose the following conditions on $\{W_n(\cdot)\}$.

i) $W_n(k) = W_n(-k), \qquad W_n(k) \geqslant 0, \quad \forall k,$

ii) $\sum_{|k|\leqslant M} W_n(k) = 1,$

iii) $\sum_{|k|\leqslant M} W_n^2(k) \to 0, \qquad \text{as } n \to \infty.$

The conditions on $\{W_n(\cdot)\}$ ensure that the mean and variance of $\widetilde{h}(\omega)$ converge to $h(\omega)$ and 0, respectively, as $n \to \infty$. The estimator $\widetilde{h}(\omega)$ given by (9) is called a discrete spectral average estimator of $h(\omega)$.

The above methods were proposed originally for estimating the spectral density function of a stationary process whose spectrum is absolutely continuous. However, in recent years spectral estimation has been used with the objective of estimating the line spectrum of a signal observed in the presence of white noise. These considerations have led engineers to look for spectral estimates with high resolution, since spectral peaks with narrow bandwidths may correspond to the presence of periodicities in the signal. Mewes and Dermühl (2001) consider the time series

$$X_t = 0.1 \times \sin(2\pi \times 100t) + \sin(2\pi \times 200t) + \sin(2\pi \times 210t) + \varepsilon_t,$$

where $\{\varepsilon_t\}$ is a sequence of uncorrelated random variables, each with zero mean and variance 1. Such a sequence is referred to as white noise and is denoted by $\varepsilon_t \sim \mathrm{WN}(0,1)$. They compared the true spectral density function of X_t with four estimates: periodogram with a sample size $n = 64$, $n = 2048$, and a discrete spectral average estimator using Hamming and rectangular lag windows with a sample size $n = 1000$. They observed that the classical methods (using periodogram) for estimating the spectrum (specially when the sample size is small) have not enough precision and they are not able to distinguish two strong peaks from each other. Larson et al. (2003) have also presented numerical examples showing these difficulties.

Two particular estimates for solving this problem are given by Capon (1969) and Pisarenko (1972), and these are widely used in signal processing problems. The Capon method is a classical spectral analysis method that has also, somewhat incorrectly, been referred to as a maximum likelihood spectral estimator, for more details, see Capon (1983), Stoica and Moses (2005), Marple (1987) and Kay (1988). These estimates, though frequently cited in the engineering literature, are less familiar to time series analysts. The fastest available technique for the computation of the Capon spectra is that of Larson and Stoica (2002). In Li et al. (1998) the bias and variance of Capon and several estimators have been compared. For more information refer to Jakobsson (2000), Jakobsson et al. (2000) and Ekman et al. (2000). Subba Rao and Gabr (1989) discussed the motivation for the Pisarenko estimator, using the properties of circular symmetric matrices, and considered its relationship with Capon's high resolution estimator. They considered the estimator,

$$h_n(\omega) = \frac{1}{2\pi n} \sum_{t=1}^{n} \sum_{s=1}^{n} R(t-s) \exp\{-i(t-s)\omega\}, \tag{10}$$

which is asymptotically equivalent to estimating $h(\omega)$, (and called $h_n(\omega)$ the "truncated spectral density function"). They showed that $h_n(\omega)$ can be written

as

$$h_n(\omega_l) = \frac{1}{4\pi}\left(\lambda_{n,0}A_{n,0}(\omega_l) + \sum_{j=1}^{(n-1)/2}(\lambda_{n,2j}+\lambda_{n,2j-1})A_{n,2j}(\omega_l)\right), \qquad (11)$$

where $A_{n,0}(\omega_l) = (2/n)2\pi F_{n-1}(\omega_l)$, $A_{n,2j}(\omega_l) = A_{n,2j-1}(\omega_l) = (1/n)\{2\pi F_{n-1}(\omega_j+\omega_l) + 2\pi F_{n-1}(\omega_j - \omega_l)\}$, and $\lambda_{n,0}, \lambda_{n,1}, \ldots, \lambda_{n,n}$ are the eigenvalues of the Toeplitz matrix $\mathbf{R}_n$ given by

$$\mathbf{R}_n = \begin{pmatrix} R(0) & R(1) & \cdots & R(n-1) \\ R(-1) & R(0) & \cdots & R(n-2) \\ \vdots & \vdots & & \vdots \\ R(-(n-1)) & R(-(n-2)) & \cdots & R(0) \end{pmatrix}.$$

Note that $F_n(\theta)$ is the Fejér kernel (e.g. Priestley, 1989) and is given by

$$F_n(\theta) = \frac{1}{2\pi n}\cdot\frac{\sin^2\left(\frac{n\theta}{2}\right)}{\sin^2\left(\frac{\theta}{2}\right)}.$$

In finding the estimator (11), they used the asymptotic equivalence of $\mathbf{R}_n$ with a related circular symmetric matrix which is well-known (e.g. Gray, 1972). Note that $h_n(\omega_l)$ is a smoothed version of the eigenvalues of $\mathbf{R}_n$, and the smoothing function (lag window) is the Fejér kernel, it is linear in the eigenvalues $\lambda_{n,j}$ and that $A_{n,j}(\omega_l)$ does not depend on the process. This suggests that these eigenvalues can be replaced by any non-linear function of $\lambda_{n,j}$ and by suitably defining an inverse function we can, in the limiting form, recover the original spectrum. To be more precise, consider a strictly monotone continuous function $G(\cdot)$ over the interval $(0,\infty)$ and let $g(\cdot)$ be an inverse function, i.e. $g(G(x)) = x$. Then we can consider the function

$$h_{n,\mathrm{P}}(\omega_l) = g\left(\sum_{j=0}^{n-1} G(\lambda_{n,j})A_{n,j}(\omega_l)\right), \qquad (12)$$

as an approximation to $h_n(\omega)$. In fact, this is the way that Pisarenko (1972) derived his estimate, and (12) is the theoretical form of the general Pisarenko estimate. Note that (12) must be multiplied by an appropriate scale factor to recover $h_n(\omega)$, and that this factor is dependent to the form of $G(\cdot)$. In its general form (12) also includes the theoretical form of the high resolution estimator of Capon (1969). To obtain this, substitute $G(x) = x^{-1}$ in (12). Then we obtain

$$h_{n,\mathrm{P}}(\omega_l) = \frac{1}{\pi}\left(\sum_{j=0}^{n-1} \lambda_{n,j}^{-1} A_{n,j}(\omega_l)\right)^{-1}. \tag{13}$$

The theoretical form of Capon's estimator is given by

$$\begin{aligned} h_{n,\mathrm{Cap}}(\omega_l) &= \frac{1}{\pi} h_{n,\mathrm{P}}(\omega_l) \\ &= \frac{1}{\pi}\left(\sum_{j=0}^{n-1} \lambda_{n,j}^{-1} A_{n,j}(\omega_l)\right)^{-1}. \end{aligned} \tag{14}$$

Note that Capon (1969) defined the theoretical form of his minimum variance spectral estimator as $(2\pi/n)h_{n,\mathrm{Cap}}(\omega_l)$. So, it is clear that the estimator defined by Capon (1969) does not strictly qualify as a power spectral density. Further remarks and details are given in Larson et al. (2003).

4 Parametric Methods

The principal difference between the spectral estimation methods of Section 3 and those in this section, is that in Section 3 we made no assumption on the studied time series (except for its stationarity). The parametric or model-based methods of spectral estimation assume that the time series satisfies a generating model with known functional form, and then proceed by estimation the parameters in the assumed model. The spectral density are then derived from the estimated model. In those cases where the assumed model is near to the reality, the parametric methods provide more accurate spectral estimates than the nonparametric techniques. The nonparametric approach remains useful, where there is little or no information about the time series model in question.

In parametric methods, the spectral density estimator is usually obtained by fitting an ARMA model to the data and then the spectral density of the fitted model is computed. Provided there is an ARMA model that fits the data satisfactorily, this procedure has the advantage that it can be made systematic by selecting the model according, for example, to a bias-corrected version of the AIC (the information criterion of Akaike) known as the AICC (Akaike, 1973a).

In this way, the pth order autoregressive estimator $\widehat{h}_p(\omega)$ of the spectral density of a stationary time series $\{X_t\}$ is the spectral density of the autoregressive process $\{Y_t\}$ defined by

$$Y_t - \widehat{\phi}_{p1} Y_{t-1} - \cdots - \widehat{\phi}_{pp} Y_{t-p} = Z_t, \qquad \{Z_t\} \sim WN(0, \widehat{v}_p) \tag{15}$$

where $\widehat{\phi}_p = (\widehat{\phi}_{p1}, \ldots, \widehat{\phi}_{pp})'$ and $\widehat{v}_p$ are the well-known Yule-Walker estimators of the coefficient vector $\phi_p = (\phi_{p1}, \ldots, \phi_{pp})'$ and white noise variance σ^2 from AR(p) model: $Y_t - \phi_{p1}Y_{t-1} - \cdots - \phi_{pp}Y_{t-p} = Z_t$, respectively. Then

$$\widehat{h}_p(\omega) = \frac{\widehat{v}_p}{2\pi} \cdot \left|1 - \widehat{\phi}_{p1}\exp\{-i\omega\} - \cdots - \widehat{\phi}_{pp}\exp\{-ip\omega\}\right|^{-2}. \qquad (16)$$

The choice of p for which the approximating AR(p) process "best" represents the data can be made by minimizing AICC. The pth order autoregressive estimator $\widehat{h}_p(\omega)$ defined by (16) is the same as the maximum entropy estimator, i.e. the spectral density $\widehat{h}$ which maximizes the entropy,

$$E = \int_{-\pi}^{\pi} \ln g(\omega) d\omega,$$

over the class of all densities g which satisfy the constraints,

$$\int_{-\pi}^{\pi} \exp\{-i\omega\tau\} g(\omega) d\omega = \widehat{R}(\tau), \qquad \tau = 0, \pm 1, \ldots, \pm p$$

(see Brockwell and Davis, 1991).

In the definition (16) it is natural to consider replacing the Yule-Walker estimates $\widehat{\phi}_p$ and $\widehat{v}_p$ by the corresponding maximum likelihood estimates. Also, there is no need to restrict attention to autoregressive models, and there are processes for which autoregressive spectral estimation performs poorly. Spectra with both sharp peaks and deep nulls cannot be modeled by either AR or MA equations of reasonably small orders. It is in these cases that the more general ARMA model is valuable. The practical ARMA estimators are computationally simple and often quite reliable, but their statistical accuracy is in some cases poor. See Byrnes et al. (2000, 2001) for some results on ARMA parameter estimation. To deal with cases of this kind we can use the estimate suggested by Akaike (1973b), i.e.

$$\widehat{h}(\omega) = \frac{\widehat{\sigma}^2}{2\pi} \cdot \frac{\left|1 + \widehat{\theta}_1\exp\{-i\omega\} + \cdots + \widehat{\theta}_q\exp\{-iq\omega\}\right|^2}{\left|1 - \widehat{\phi}_1\exp\{-i\omega\} - \cdots - \widehat{\phi}_p\exp\{-ip\omega\}\right|^2}, \qquad (17)$$

where $\widehat{\phi} = (\widehat{\phi}_1, \ldots, \widehat{\phi}_p)'$ and $\widehat{\theta} = (\widehat{\theta}_1, \ldots, \widehat{\theta}_q)'$ and $\widehat{\sigma}^2$ are maximum likelihood estimates of an ARMA(p, q) process fitted to the data, with p and q chosen by AICC. The estimate $\widehat{h}(\omega)$ is called the maximum likelihood ARMA (or MLARMA) spectral density estimate. A simple but less efficient estimator than (17) which is particularly useful for processes whose MA(∞) representation has

rapidly decaying coefficients is the moving average estimator (Brockwell and Davis, 1991) given by

$$\widehat{h}_q(\omega) = \frac{\widehat{v}_q^2}{2\pi} \cdot \left|1 + \widehat{\theta}_{q1} \exp\{-i\omega\} + \cdots + \widehat{\theta}_{qq} \exp\{-iq\omega\}\right|^2, \qquad (18)$$

where $\widehat{\theta}_q = (\widehat{\theta}_{q1}, \ldots, \widehat{\theta}_{qq})'$, and $\widehat{v}_q^2$ are the innovation estimates. The advantage of both estimators (16) and (18) over the MLARMA estimator is the substantial reduction in computation time. Moreover, under specified conditions, the asymptotic distributions of the autoregressive and moving average spectral density estimators can be determined for a large class of linear processes (Brockwell and Davis, 1991).

An indirect but computationally efficient method with numerical results is given by Stoica et al. (2000) and Dumitrescu et al. (2001).

5 Multivariate Case

In this section we consider the estimation of the spectral density function of a vector-valued series. The approach in multivariate case is based on the properties of the eigenvalues of block-Toeplitz matrices, and is similar to the approach given by Subba Rao and Gabr (1989) and heavily based on the results of Hannan and Wahlberg (1989). Let $\{\mathbf{X}_t,\ t \in \mathbb{Z}\}$ be zero mean vector, discrete-parameter, and second order T-dimensional stationary series with $X_t(j)$, $j = 0, \ldots, T-1$, as its jth element, and $\mathbf{R}(\tau) = E(\mathbf{X}_{t+\tau}\mathbf{X}_t')$, $\tau \in \mathbb{Z}$, be the autocovariance matrix of $\mathbf{X}_t$. We assume that the series has an absolutely continuous spectrum and let $\mathbf{h}(\omega) = [h_{jk}(\omega)]_{j,k=0,1,\ldots,T-1}$, denote its spectral density matrix, i.e., let

$$h_{jk}(\omega) = \frac{1}{2\pi} \sum_{\tau=-\infty}^{\infty} R_{jk}(\tau) \exp\{-i\omega\tau\}$$

where $R_{jk}(\tau) = EX_{t+\tau}(j)X_t(k)$, $(j, k = 0, 1, \ldots, T-1)$ is the (j,k)-th element of $\mathbf{R}(\tau)$. Equivalently, we may write

$$\mathbf{h}(\omega) = \frac{1}{2\pi} \sum_{\tau=-\infty}^{\infty} \mathbf{R}(\tau) \exp\{-i\omega\tau\}, \qquad 0 \leqslant \omega \leqslant 2\pi$$

Let $\{\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_n\}$ be a sample of size n from $\{\mathbf{X}_t\}$. A natural estimator of the cross-covariance $R_{jk}(\tau)$ is

$$\widehat{R}_{jk}(\tau) = \frac{1}{n} \sum_t X_{t+\tau}(j)X_t(k), \qquad \tau = 0, \pm 1, \ldots, \pm(n-1)$$

where the summation is from $t = 1$ to $(n - \tau)$, $\tau \geqslant 0$, and from $t = (1 - \tau)$ to n, $\tau < 0$. It can be shown that (Priestley, 1989)

$$E\{\widehat{R}_{jk}(\tau)\} = R_{jk}(\tau) + O\left(\frac{1}{n}\right),$$

where "$f(x) = O(g(x))$" denote the existence of a constant C, such that $|f(x)| \leqslant Cg(x)$ for all x.

Let $\mathbf{I}_n(\omega) = [I_{n,jk}(\omega)]_{j,k=0,\ldots,T-1}$ be the periodogram matrix with (j,k)-th element

$$I_{n,jk}(\omega) = \xi_{\mathbf{X}_j}(\omega)\xi^*_{\mathbf{X}_k}(\omega),$$

where

$$\xi_{\mathbf{X}_j}(\omega) = \frac{1}{\sqrt{2\pi n}} \sum_{t=1}^{n} X_t(j) \exp\{-i\omega t\}, \qquad j = 0, \ldots, T-1$$

denotes the finite Fourier transform of $X_t(j)$, (for $j = k$, $I_{n,jj}$ is simply the periodogram of $X_t(j)$, and for $j \neq k$, $I_{n,jk}$ is the cross-periodogram between $X_t(j)$ and $X_t(k)$). Then

$$\begin{aligned} E\{I_{n,jk}(\omega)\} &= \frac{1}{2\pi n} \sum_{t=1}^{n} \sum_{s=1}^{n} R_{jk}(t-s) \exp\{-i(t-s)\omega\} \\ &= h_{n,jk}(\omega), \text{ say.} \end{aligned} \tag{19}$$

In matrix form, under some regularity conditions on $\mathbf{h}(\omega)$, (Priestley, 1989), we can write

$$E\{\mathbf{I}_n(\omega)\} = \frac{1}{2\pi n} \sum_{t=1}^{n} \sum_{s=1}^{n} \mathbf{R}(t-s) \exp\{-i(t-s)\omega\} \tag{20}$$

$$= \mathbf{h}_n(\omega), \text{ say,} \tag{21}$$

$$= \mathbf{h}(\omega) + \mathbf{O}\left(\frac{\log n}{n}\right), \tag{22}$$

where $\mathbf{O}(\cdot) = O(\cdot)\mathbf{1}_n\mathbf{1}_n'$, with $\mathbf{1}_n = (1, 1, \ldots, 1)'$. We shall call $\mathbf{h}_n(\omega)$ the *truncated spectral density matrix.*

Thus $\mathbf{I}_n(\omega)$ is an asymptotically unbiased estimator of $\mathbf{h}(\omega)$, but of course it is not consistent (Brillinger, 1981). In order to find a consistent estimator of $\mathbf{h}(\omega)$, the usual procedure is to smooth the periodogram by a suitable kernel

(Brillinger, 1981). Nematollahi and Subba Rao (2005) showed that the theoretical spectral density matrix $\mathbf{h}(\omega)$ can be written in terms of the eigenvalues of the variance-covariance matrix, then one can estimate $\mathbf{h}(\omega)$ using the eigenvalue decomposition of the sample variance-covariance matrix and show that it intrinsically makes use of Fejér kernel type of weight functions.

More precisely, define a $nT \times 1$ vector $\mathbb{X}_n = (\mathbf{X}_n', \mathbf{X}_{n-1}', \ldots, \mathbf{X}_1')'$ and let $\mathbf{\Gamma}_n = E\mathbb{X}_n\mathbb{X}_n'$ be its variance-covariance matrix. We have

$$\mathbf{\Gamma}_n = \begin{pmatrix} \mathbf{R}(0) & \mathbf{R}(1) & \cdots & \mathbf{R}(n-1) \\ \mathbf{R}(-1) & \mathbf{R}(0) & \cdots & \mathbf{R}(n-2) \\ \vdots & \vdots & & \vdots \\ \mathbf{R}(-(n-1)) & \mathbf{R}(-(n-2)) & \cdots & \mathbf{R}(0) \end{pmatrix}.$$

We note that $\mathbf{\Gamma}_n$ is a block-Toeplitz matrix. Individual matrix elements are not, in general, symmetric ($\mathbf{R}(\tau) \neq \mathbf{R}'(\tau)$), although $\mathbf{R}(-\tau) = \mathbf{R}'(\tau)$.

Nematollahi and Subba Rao (2005) derived the following expression for estimation of spectral density matrix $\mathbf{h}(\omega)$.

$$\mathbf{h}_n(\omega_l) = \frac{1}{4\pi} \sum_{j=0}^{n-1} \mathbf{A}_n(\omega_j, \omega_l), \tag{23}$$

with

$$\mathbf{A}_n(\omega_j, \omega_l) = \frac{2}{n} \sum_{t=1}^{n} \sum_{s=1}^{n} \mathbf{W}_n^{t*}(\omega_j) \mathbf{\Lambda}_n(\omega_j) \mathbf{W}_n^{s}(\omega_j) \exp\{-i(t-s)\omega_l\}, \tag{24}$$

where $\mathbf{\Lambda}_n(\omega_j)$ is an "eigenvalue-matrix" of $\mathbf{\Gamma}_n$, and $\mathbf{W}_n(\omega_j)$ is an "eigenvector-matrix" associated with $\mathbf{\Lambda}_n(\omega_j)$, (clearly, they are not the eigenvalue and the eigenvector in the usual sense). In finding the estimator (23), they used the asymptotic equivalence of $\mathbf{R}_n$ with a related circular symmetric matrix which have been proved in the multivariate case by Nematollahi and Shishebor (2005) and is a multivariate generalization of the well-known result of Gray (1972).

Nematollahi and Subba Rao (2005) showed that an equivalence form for $\mathbf{h}_n(\omega_l)$ given by (23) is

$$\mathbf{h}_n(\omega_l) = \frac{1}{4\pi} \Bigg(\frac{2}{n} \times 2\pi F_{n-1}(\omega_l) \mathbf{\Lambda}_n(\omega_0) + 2 \sum_{j=1}^{(n-1)/2} \frac{1}{n} \mathbf{\Lambda}_n(\omega_{2j}) \{ 2\pi F_{n-1}(\omega_j + \omega_l) + 2\pi F_{n-1}(\omega_j - \omega_l) \} \Bigg), \tag{25}$$

where

$$F_n(\theta) = \frac{1}{2\pi n} \cdot \frac{\sin^2\left(\frac{n\theta}{2}\right)}{\sin^2\left(\frac{\theta}{2}\right)},$$

which is the Fejér kernel. The approximate relation (25) tells us that the spectral density function $\mathbf{h}_n(\omega_l)$ is in fact a smooth function of $\mathbf{\Lambda}_n(\omega_j)$ and the smoothing function is the well-known Fejér kernel. We observe that $\mathbf{h}_n(\omega_l)$ is linear in $\mathbf{\Lambda}_n(\omega_j)$ and that $\mathbf{A}_n(\omega_j, \omega_l)$ does not depend on the time series $\{\mathbf{X}_t\}$. Similar to the univariate case, this suggests that these eigenvalue-matrices can be replaced by any nonlinear function of $\mathbf{\Lambda}_n(\omega_j)$ and by suitably defining an inverse function, we can, in the limiting form, recover the original spectrum. Consider a strictly monotonic continuous function $G(\cdot)$ and $g(\cdot)$ be an inverse function, i.e. $g(G(x)) = x$. Then we can introduce the function

$$\mathbf{h}_{n,\mathrm{P}}(\omega_l) = g\left(\sum_{j=0}^{n-1} \frac{2}{n} \mathbf{B}_n^*(\omega_j, \omega_l) G(\mathbf{\Lambda}_n(\omega_j)) \mathbf{B}_n(\omega_j, \omega_l)\right), \tag{26}$$

as an approximation to $\mathbf{h}_n(\omega_l)$, where $\mathbf{B}_n(\omega_j, \omega_l) = \sum_{s=1}^{n} \mathbf{W}_n^s(\omega_j) \exp\{is\omega_l\}$, and $\mathbf{W}_n^s(\omega_j)$ is the sth block of $\mathbf{W}_n(\omega_j)$. In fact, this is a generalization of the way that Pisarenko (1972) derived his estimate in the univariate case, and (26) is the generalization of the theoretical form of the Pisarenko's estimator given by (12), to the multivariate case. Also, note that (26) must be multiplied by an appropriate scale factor to recover $\mathbf{h}_n(\omega_l)$, and this factor is independent of the form of $G(\cdot)$.

As we mentioned in Section 3, the high resolution estimation of the spectral density function given by (14) was introduced by Capon (1969) in the univariate case. A multivariate generalization and also an explicit expression for the high resolution spectral density matrix (a generalization of Capon's estimator) of the vector series $\mathbf{X}_t$ are also given by Nematollahi and Subba Rao (2005). They derived an appropriate minimum variance (MV) spectral estimator with form

$$\mathbf{h}_{n,\mathrm{Cap}}(\omega) = \frac{1}{\pi}\left(\frac{2}{n}\mathbf{L}_n^*(\omega)\Gamma_n^{-1}\mathbf{L}_n(\omega)\right)^{-1}, \tag{27}$$

where $\mathbf{L}_n(\omega_l) = (\mathbf{I}, \exp\{i\omega_l\}\mathbf{I}, \ldots, \exp\{ni\omega_l\}\mathbf{I})'$. It can be shown that another theoretical form of generalized Capon's estimator is given by

$$\mathbf{h}_{n,\mathrm{Cap}}(\omega_l) = \frac{1}{\pi}\left(\sum_{j=0}^{n-1} \widetilde{\mathbf{A}}_n(\omega_j, \omega_l)\right)^{-1}, \tag{28}$$

where

$$\widetilde{\mathbf{A}}_n(\omega_j,\omega_l)=\frac{2}{n}\sum_{t=1}^{n}\sum_{s=1}^{n}\mathbf{W}_n^{t*}(\omega_j)\mathbf{\Lambda}_n^{-1}(\omega_j)\mathbf{W}_n^{s}(\omega_j)\exp\{-i(t-s)\omega_l\}. \qquad (29)$$

This is a multivariate generalization of the *minimum variance spectral* (MVS) *estimator* due to Capon (1969) or relation (14).

The relationship between the AR and Capon spectra and the extension of the result to two-dimensional time series can be found in Jakobsson et al. (2000). The extension of the nonparametric spectral analysis methods to two-dimensional (2D) time series can be found in Stoica and Moses (2005). They also provide new interpretations for some of these methods, which are particularly useful when we want very simple (although somewhat heuristic) derivations of the methods in question. The 2D spectral analysis finds applications in image processing, synthetic aperture radar imagery, etc. See Larson et al. (2003) and the references therein for a review that covers the well-known 2D methods and their application to synthetic aperture radar. The 2D extensions of some parametric methods are also discussed in Stoica and Moses (2005).

6 Conclusions

This paper has summarized some main results in the area of parametric and nonparametric analysis of the spectral estimation problem. We have also discussed various extensions of the methods, with the extension to the multivariate time series. Spectral estimation via methods mentioned above is a theme that can be extended and developed in many ways. The intention is to summarize some extension and variations of the mentioned methods that have appeared in the literature. Instead of providing a full treatment of all techniques, we have referred the reader to the original articles for more details. Although there exist an abundance of research papers in this area, there are still topics that should form the subject of future investigations. We have not given an exhaustive list of open problem, but it seems the most interesting and relevant ones to solve include finding a fast time-recursive implementation of Capon estimator and deriving an optimal choice of the truncation point (bandwidth) and windows.

Acknowledgement

I wish to thank the referees and editor for many useful suggestions.

References

Akaike, H. (1973a). Information theory and an extension of maximum likelihood principle. 2nd International Symposium on Information Theory, B. N. Petrov and F. Csaki, eds. akademiai Kiado, Budapest, pp. 267-281.

Akaike, H. (1973b). Maximum likelihood identification of Gaussian autoregressive moving average models. *Biometrika* **60**, 255-265.

Bartlett, M.S. (1963). Statistical estimation of density functions. *Sankhyā Ser. A* **25**, 245-254.

Blackman, R.B.; Tukey, J.W. (1959). Particular pairs of windows, *The Measurement of Power Spectra, From the Point of View of Communications Engineering.* Dover, New York.

Bloomfield, P. (1976). *Fourier Analysis of Time Series: An Introduction.* Wiley, New York.

Bracewell, R.N. (1986). *The Fourier Transform and Its Applications*, 2nd ed. McGraw-Hill, New York.

Brillinger, D.R. (1981). *Time Series: Data Analysis and Theory*, Expanded ed. Holt, Rinehart and Winston, New York.

Brockwell, P.J.; Davis, R.A. (1991). *Time Series: Theory and Methods.* Springer, New York.

Byrnes, C.L.; Georgious, T.T.; Lindquist, A. (2000). A new approach to spectral estimation: A tunable high-resolution spectral estimator. *IEEE Trans. signal Process.* **48**, 3189-3205.

Byrnes, C.L.; Georgious, T.T.; Lindquist, A. (2001). A generalized entropy criterion for nevanlinna-pick interpolation with degree constraint. *IEEE Trans. Automat. Control.* **46**, 822-839.

Capon, J. (1969). High resolution frequency-wave number spectrum analysis. *Proc. IEEE* **57**, 1408-1418.

Capon, J. (1983). Maximum-likelihood spectral estimation. In *Nonlinear Methods of Spectral Analysis*, S. Haykin, ed. Springer, New York.

Chatfield, C. (1975). *The Analysis of Time Series: An Introduction,* 2nd ed. Chapman and Hall, London.

Daniell, P.J. (1946). Discussion on symposium on autocorrelation in time series. *J. Roy. Statist. Soc. Suppl.* **8**, 88-90.

DeGraaf, S.R. (1998). Imaging via modern 2-D spectral estimation methods, *IEEE Trans. Image Process.* **7**, 729-761.

Dumitrescu, B.; Tabus, I.; Stoica, P. (2001). On the parametrization of positive real sequences and MA parameter estimation. *IEEE Trans. Signal Process.* **49**, 2630-2639.

Ekman, T.; Jakobsson, A.; Stoica, P. (2000). On efficient implementation of the CAPON algorithm. In *Proc. European Signal Processing Conference (EUSIPCO)*, Tampere, Finland.

Epanechnikov, V.A. (1969). Non-parametric estimation of a multivariate probability density. *Theory Probab. Appl.* **14**, 153-158.

Gini, F.; Lombardini, F. (2002). Multilook APES for multibaseline SAR interferometry. *IEEE Trans. Signal Process.* **50**, 1800-1803.

Gray, R.M. (1972). On the asymptotic eigenvalue distribution of Toeplitz matrices. *IEEE Trans. Inform. Theory* **18**, 725-30.

Hannan, E.J. (1970). *Multiple Time Series.* Wiley, New York.

Hannan, E.J.; Wahlberg, B. (1989). Convergence rates for inverse Toeplitz matrix forms. *J. Multivariate Anal.* **31**, 127-135.

Haykin, S. (ed.) (1991). *Advances in Spectrum Analysis and Array Processing*, Volumes 1 and 2. Prentice-Hall. Englewood Cliffs, NJ.

Haykin, S. (ed.) (1995). *Advances in Spectrum Analysis and Array Processing*, Volume 3. Prentice-Hall. Englewood Cliffs, NJ.

Jakobsson, A. (2000). Model-Based and Matched-Filterbank Signal Analysis. PhD thesis, Uppsala University, Uppsala, Sweden.

Jakobsson, A.; Ekman, T.; Stoica, P. (2000). Computationally efficient 2-D spectral estimation. *Proc. Intl. Conf. Acoust., Speech, Signal Processing.*

Jakobsson, A.; Marple, L.; Stoica, P. (2000). Computaionally efficient two-dimensional capon spectrum analysis. *IEEE Trans. Signal Process.* **48**, 2651-2661.

Jakowatz, C.V.; Wahl, D.E.Jr.; Eichel, P.H.; Ghiglia, D.C.; Thompson, P.A. (1996). *Spotlight-Mode Synthetic Aperture Radar: A Signal Processing Approach.* Kluwer, Norwell.

Kay, S.M. (1988). *Modern Spectral Estimation: Theory and Application.* Prentice-Hall. Englewood Cliffs, NJ.

Koopmans, L.H. (1974). *The Spectral Analysis of Time Series.* Academic Press, New York.

Larsson, E.G.; Stoica, P. (2002). Fast implementation of two-dimensional APES and CAPON spectral estimators. *Multidimens. Systems Signal Process.* **13**, 35-53.

Larson, E.G.; Li, J.; Stoica, P. (2003) High-resolution nonparametric spectral analysis: theory and applications. In *High-Resolution Signal Processing*, Y. Hua, Q. Cheng, and A. B. Gershman, eds. Marcel Dekker, New York.

Li, H.; Li, J.; Stoica, P. (1998). Performance analysis of forward-backward matched-filterbank spectral estimators. *IEEE Trans. Signal Process.* **46**, 1954-1966.

Marple, S.; Lawrence, J. (1987). *Digital Spectral Analysis with Applications.* Prentice-Hall. Englewood Cliffs, NJ.

Mewes Dermühl (2001). Improvement on high resolution spectral estimation techniques in sensor array signal processing. A available in *http://wwwgssr.uah.es/ftp/felipecatedra/ficheros%20pdf%20DOA/*).

Nematollahi, A.R; Shishebor, Z. (2005). Asymptotic equivalence of the covaiance matrix of a multivariate stationary time series with the related circular symmetric matrix. *J. Stat. Theory Appl.* **4**, 209-217.

Nematollahi, A.R.; Subba Rao, T. (2005). On the spectral density estimation of periodically correlated (cyclostationary) time series, *Sankhyā Ser. A.* **67**, 568-589.

Percival, D.B.; Walden, A.T. (1993). *Spectral Analysis for Physical Applications: Multitaper and Conventional Univariate Techniques.* Cambridge University Press. Cambridge, UK.

Pisarenko, V.F. (1972). On the estimation of spectra by means of nonlinear functions of the covariance matrix. *Geophys., J. R. Astr. Soc.* **28**, 511-531.

Porat, B. (1994). *A Course in Digital Signal Processing.* Wiley, New York.

Priestley, M.B. (1962). Basic considerations in the estimation of spectra. *Technometrics* **4**, 551-564.

Priestley, M.B. (1989). *Spectral Analysis and Time Series.* Academic Press. London.

Proakis, J.G.; Rader, C.M.; Ling, F.; Nikias, C.L. (1992). *Advanced Digital Signal Processing.* Macmillan, New York.

Rahnemaee, M.; Zare, M.; Nematollahi, A.R.; Sedghi, H. (2005). Application of spectral analysis of daily water level and spring discharge hydrographs data for comparing physical characteristics of karstic aquifers. *J. Hydrology* **311**, 106-116.

Scargle, J.D. (1997). Astronomical Time Series Analysis. *Astronomical Time Series,* (eds. D. Maoz et al.) Kluwer.

Scharf, L.L. (1991). *Statistical Signal Processing: Detection, Estimation, and Time Seires Analysis.* Addison-Wesley. Reading, MA.

Schuster, A. (1898). On the investigation of hidden periodicities with application to a supposed 26-day period of meteorological phenomena. *Terr. Mag. Atmos. Elect.* **3**, 13-41.

Stoica, P.; McKelvey, T.; Mari, J. (2000). MA estimation in polynomial time. *IEEE Trans. Signal Process.* **48**, 1999-2012.

Stoica, P.; Moses, R.L. (1997). *Introduction to Spectral Analysis.* Prentice-Hall. NJ.

Stoica, P.; Moses, R.L. (2005). *Spectral Analysis of Signals*, Prentice-Hall. NJ.

Subba Rao, T.; Gabr, M.M. (1989). The estimation of spectrum, inverse spectrum and inverse autocovariances of a stationary time series. *J. Time Ser. Anal.* **10**, 183-202.

Therrien, C.W. (1992). *Discrete Random Signals and Statistical Signal Processing.* Prentice-Hall. Englwood Cliff, NJ.

Ali-reza Nematollahi
Department of Statistics,
College of Science,
Shiraz University,
Shiraz 71454, Iran.
e-mail: *nematollahi@susc.ac.ir*