Automatic estimation of multivariate spectra via smoothing splines

BY ORI ROSEN

Department of Mathematical Sciences, University of Texas at El Paso, El Paso, Texas 79968, U.S.A.
ori@math.utep.edu

AND DAVID S. STOFFER

Department of Statistics, University of Pittsburgh, Pittsburgh, Pennsylvania 15260, U.S.A.
stoffer@pitt.edu

SUMMARY

The classical method for estimating the spectral density of a multivariate time series is first to calculate the periodogram, and then to smooth it to obtain a consistent estimator. Typically, to ensure the estimate is positive definite, all the elements of the periodogram are smoothed the same way. There are, however, many situations for which different components of the spectral matrix have different degrees of smoothness. We propose a Bayesian approach that uses Markov chain Monte Carlo techniques to fit smoothing splines to each component, real and imaginary, of the Cholesky decomposition of the periodogram matrix. The spectral estimator is then obtained by reconstructing the spectral estimator from the smoothed Cholesky decomposition components. Our technique produces an automatically smoothed spectral matrix estimator along with samples from the posterior distributions of the parameters to facilitate inference.

Some key words: Coherency; Cholesky decomposition; DNA nucleotide sequence; Markov chain Monte Carlo; Multivariate spectral density; Smoothing spline; Spectral analysis; Spectral envelope.

1. INTRODUCTION

The classical method for estimating the spectral density of a multivariate time series is first to calculate the periodogram and then to smooth it to obtain a consistent estimator (Brillinger, 2001, Ch. 5; Shumway & Stoffer, 2006, Ch. 4). A major difficulty is to guarantee that the final estimate is positive definite while allowing optimal smoothing for each element of the spectral matrix. Typically, to ensure the estimate is positive definite, all the elements of the periodogram are smoothed the same way. Pawitan (1996) proposed a penalized likelihood estimator for the cross-spectrum of a bivariate time series. The smoothing parameters for the real and imaginary parts can be chosen objectively from the data. Thus the real and imaginary parts can have different smoothness. With an implicit restriction on the estimation procedure to make the coherence less than or equal to one, the estimator is positive semidefinite. Extension of this method beyond bivariate time series is difficult because estimating the cross-spectra one at a time cannot guarantee that the final multivariate spectral estimate is positive semidefinite. Another difficulty is in constructing

confidence intervals for the spectrum. In the univariate setting, Franke & Haïdle (1992) proposed a bootstrap procedure for constructing confidence intervals. However, it is difficult to generalize this method to the multivariate setting. As pointed out in Dai & Guo (2004), there are many situations for which different components of the spectral matrix have different degrees of smoothness, and hence require different smoothing parameters in order to obtain optimal estimators.

Dai & Guo (2004) overcame these problems by smoothing the Cholesky decomposition of a multitaper spectral estimator and then reconstructing the spectral estimator from the smoothed Cholesky components. Given data from a stationary, vector-valued time series, their estimation and inference procedure consisted of the following steps: compute the periodogram of the data; smooth the periodogram using a multitaper spectral estimator (Thomson, 1982); perform the Cholesky decomposition on the multitaper estimator; smooth each of the Cholesky decomposition components with its own smoothing parameter; reconstruct the spectral estimator from the smoothed Cholesky decomposition components; and finally use a bootstrap method to obtain pointwise confidence intervals.

In this paper, we take a Bayesian approach that uses Markov chain Monte Carlo techniques to fit smoothing splines to each component, real and imaginary, of the Cholesky decomposition of the periodogram matrix. The spectral estimator is then obtained by reconstructing the spectral estimator from the smoothed Cholesky decomposition components. The advantage of our technique is that it allows for automatic smoothing of the components but avoids having to pre-smooth the periodogram by calculating a multitaper estimate. In addition, because our procedure produces a sample from the posterior distribution of all the parameters, credible intervals are easily obtained.

Throughout, we assume that we have a sufficiently large number, n, of observations from a p-dimensional stationary time series, x_t, whose $p \times p$ autocovariance matrix, $\Gamma(h) = \{\gamma_{jk}(h)\}$, satisfies $\sum_{h=-\infty}^{\infty} |\gamma_{jk}(h)| < \infty$ for all $j, k = 1, \ldots, p$. The $p \times p$ spectral density matrix is given by

$$f(\nu) = \sum_{h=-\infty}^{\infty} \Gamma(h) e^{-2\pi i \nu h}, \quad -1/2 \leqslant \nu \leqslant 1/2,$$

where $f(\nu) = \{f_{jk}(\nu)\}$, for $j, k = 1, \ldots, p$, and frequency, ν, is measured in cycles per unit time; note that $f(\nu) = f^*(\nu)$, where $*$ denotes the conjugate transpose. Finally, we assume that $f(\nu)$ is positive definite.

2. THE MODEL AND PRIOR SPECIFICATION

Given a realization $x_1, \ldots, x_n$ from a multivariate stationary time series, the discrete Fourier transform of the data is given by

$$y(\nu_k) = n^{-1/2} \sum_{t=1}^{n} x_t e^{-2\pi i \nu_k t},$$

for $k = 0, 1, \ldots, n - 1$, where $\nu_k = k/n$ are the Fourier frequencies. The discrete Fourier transforms, $y(\nu_k)$, $k = 0, \ldots, n - 1$, of a zero-mean stationary multivariate time series are approximately independent complex multivariate normal random variables. Let $y_k \equiv y(\nu_k)$ and $f_k \equiv f(\nu_k)$, $k = 0, \ldots, n - 1$. The approximate likelihood is given by

$$L(y_0, \ldots, y_{n-1}; f_0, \ldots, f_{n-1}) \simeq \prod_{k=0}^{n-1} \det(f_k)^{-1} \exp(-y_k^* f_k^{-1} y_k), \qquad (1)$$

where det(·) denotes determinant. Equation (1) is an extension of the Whittle likelihood (Whittle, 1957) to the multivariate case. Note that, since the spectral matrix and the discrete Fourier transform are even functions of ν, there are only $[n/2]$ distinct observations.

Our goal is to obtain smooth estimators of the elements of f as a function of ν while satisfying the constraint that f is positive definite. To this end, we express the inverse of the spectral matrix at frequency ν_k as the modified complex Cholesky factorization

$$f_k^{-1} = T_k^* D_k^{-1} T_k, \tag{2}$$

where T_k is a complex unit lower triangular matrix, and D_k is a diagonal matrix. To be more specific,

$$T_k = \begin{pmatrix} 1 & & & & \\ -\theta_{21}^{(k)} & 1 & & & \\ -\theta_{31}^{(k)} & -\theta_{32}^{(k)} & 1 & & \\ \vdots & \vdots & & \ddots & \\ -\theta_{p1}^{(k)} & -\theta_{p2}^{(k)} & \cdots & -\theta_{p,p-1}^{(k)} & 1 \end{pmatrix}$$

and $D_k = \operatorname{diag}(\delta_{1k}^2, \ldots, \delta_{pk}^2)$. Note that, in general, the $\theta_{il}^{(k)}$'s are complex-valued. The real modified Cholesky decomposition has been used to model covariance matrices; see for example Pourahmadi (1999), Pourahmadi (2000), Daniels & Pourahmadi (2002), Pourahmadi & Daniels (2002) and Wu & Pourahmadi (2003).

We note that Dai & Guo (2004) worked directly with the complex Cholesky decomposition of f_k, $f_k = L_k L_k^*$ say, where L_k is lower triangular. In particular, they obtained smoothed estimators of the elements of L_k. In our case, we use the modified decomposition of f_k^{-1} given in (2) because it appears naturally in the likelihood, (1). As previously stated, it is difficult to model the elements of the spectral matrix directly because of the constraint that the spectral matrix must be positive definite at each frequency. In the factorization (2), the $\theta_{il}^{(k)}$'s are unconstrained and the δ_{jk}^2's are positive. Thus, it is much easier to model these parameters rather than the elements of the spectral matrix. Once T_k and D_k have been estimated, the resulting estimator of f_k is automatically positive definite.

To facilitate the estimation of the $\theta_{il}^{(k)}$'s and the δ_{jk}^2's and thereby the estimation of the spectral matrix, we use the likelihood (1) in combination with the factorization (2). We first rewrite the likelihood (1) as a function of the $\theta_{il}^{(k)}$'s and the δ_{jk}^2's. Let $N = [n/2]$, θ_k be the $p(p-1)/2$-dimensional vector $(\theta_{21}^{(k)}, \theta_{31}^{(k)}, \theta_{32}^{(k)}, \ldots, \theta_{p,p-1}^{(k)})'$, $\Theta = (\theta_1, \ldots, \theta_N)$, $\Delta = \{\delta_{1k}^2, \ldots, \delta_{pk}^2\}_{k=1}^N$ and $Y = (y_1, \ldots, y_N)$. From (1) and (2) it follows that the likelihood can be expressed as

$$L(Y; \Delta, \Theta) \propto \prod_{k=1}^N \prod_{j=1}^p \delta_{jk}^{-2} \exp\{(y_k - Z_k \theta_k)^* D_k^{-1} (y_k - Z_k \theta_k)\}, \tag{3}$$

where Z_k is a $p \times p(p-1)/2$ design matrix given by

$$Z_k = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & & \cdots & & 0 \\ y_{1k} & 0 & 0 & 0 & 0 & 0 & 0 & & \cdots & & 0 \\ 0 & y_{1k} & y_{2k} & 0 & 0 & 0 & 0 & & \cdots & & 0 \\ 0 & 0 & 0 & y_{1k} & y_{2k} & y_{3k} & 0 & & \cdots & & 0 \\ \vdots & & & & & & & & & & \\ 0 & & & \cdots & & & \cdots & 0 & y_{1k} & y_{2k} & \cdots & y_{p-1,k} \end{pmatrix}$$

with y_{ik} denoting the ith entry of y_k. Note that in (3) we have ignored the endpoint involving y_0. Next, we place linear smoothing spline priors on the $\theta_{il}^{(k)}$'s and the δ_{jk}^2's. Dai & Guo, 2004) used cubic smoothing splines. In our experience, linear smoothing splines are better suited to estimating the spectral matrix, as they can better accommodate narrowband peaks. In particular, each of the $\log \delta_{jk}^2$'s and the real and imaginary parts of each of the negative $\theta_{il}^{(k)}$'s are expressed as

$$\alpha_0 + \alpha_1 \nu_k + \sum_{s=1}^{N} \psi_s(\nu_k)\beta_s , \tag{4}$$

where $\psi_s(\nu_k) = \sqrt{2}\cos\{(s-1)\pi\nu_k\}$. The $\psi_s(\cdot)$'s are the Demmler–Reinsch basis functions for linear smoothing splines (Eubank, 1999). Let X_β be the matrix whose columns are the basis functions $\psi_s(\cdot)$ evaluated at $\nu_1, \ldots, \nu_N$, and let X_α be a matrix whose columns are the vector of ones and $(\nu_1, \ldots, \nu_N)'$. Let $X = (X_\alpha \mid X_\beta)$ be the matrix formed by binding X_α and X_β columnwise, $\gamma_j = (\alpha_j', \beta_j')'$, $\Delta_j = (\delta_{j1}^2, \ldots, \delta_{jN}^2)'$ and $\theta_{il} = (\theta_{il}^{(1)}, \ldots, \theta_{il}^{(N)})'$. Then

$$\log \Delta_j = X\gamma_j, \quad -\Re(\theta_{il}) = X\gamma_{il(re)}, \quad -\Im(\theta_{il}) = X\gamma_{il(im)} , \tag{5}$$

for $j = 1, \ldots, p$, $i = 2, \ldots, p$, and $l = 1, \ldots, i-1$, where $\Re(\cdot)$ and $\Im(\cdot)$ denote the real part and the imaginary part, respectively. Corresponding to (5), the priors on α_j, $\alpha_{il(re)}$ and $\alpha_{il(im)}$ are taken to be $N(0, \sigma_\alpha^2 I_2)$, and those on β_j, $\beta_{il(re)}$ and $\beta_{il(im)}$ are taken to be $N(0, \tau_j^2 I_N)$, $N(0, \tau_{il(re)}^2 I_N)$ and $N(0, \tau_{il(im)}^2 I_N)$, respectively. With the $\theta_{il}^{(k)}$'s and the δ_{jk}^2's viewed as functions of ν, the parameters τ_j^2, $\tau_{il(re)}^2$ and $\tau_{il(im)}^2$ are smoothing parameters, governing the amount of smoothing of each of these functions. A zero value of a smoothing parameter corresponds to a linear fit, while a value tending to infinity results in an interpolating linear spline. The priors on the smoothing parameters are $p(\tau_j^2) \propto 1/\tau_j^2$, $p(\tau_{il(re)}^2) \propto 1/\tau_{il(re)}^2$ and $p(\tau_{il(im)}^2) \propto 1/\tau_{il(im)}^2$. We estimate the spectral matrix by its posterior mean using Markov chain Monte Carlo methods to perform the required multidimensional integration.

3. THE SAMPLING SCHEME

Let $u_k = y_k - Z_k\theta_k$. If we plug the expression for $\log \Delta_j$ in (5) into the likelihood (3) and incorporate the priors, the conditional distribution of γ_j, $j = 1, \ldots, p$, is given by

$$p(\gamma_j \mid \Theta, Y) \propto \exp\left[-\sum_{k=1}^{N}\{x_k'\gamma_j + |u_{jk}|^2 \exp(-x_k'\gamma_j)\} - \frac{1}{2\sigma_\alpha^2}\alpha_j'\alpha_j - \frac{1}{2\tau_j^2}\beta_j'\beta_j\right], \tag{6}$$

where u_{jk} is the jth element of u_k, $|\cdot|$ denotes the complex modulus and x_k is the kth row of X. Since this is not a standard distribution, we use a Metropolis–Hastings step to sample from it. The conditional distributions of the corresponding smoothing parameters are $\mathrm{IG}(N/2, \frac{1}{2}\beta_j'\beta_j)$, for $j = 1, \ldots, p$. The conditional distributions of $\gamma_{il(re)}$ and $\gamma_{il(im)}$, for $i = 2, \ldots, p$ and $j = 1, \ldots, i-1$, are multivariate normal, and the conditional distributions of the corresponding smoothing parameters are $\mathrm{IG}(N/2, \frac{1}{2}\beta_{il(re)}'\beta_{il(re)})$ and $\mathrm{IG}(N/2, \frac{1}{2}\beta_{il(im)}'\beta_{il(im)})$, where IG denotes the inverse gamma distribution. In principle, for smoothing splines, the knots are at the abscissa values of the data points. Thus, in our case, there are N basis functions. In practice, however, not all the N basis functions are necessary; in fact, only about $N/10$ are usually sufficient. More details of the sampling scheme are given in the Appendix.

4. EXAMPLES

4·1. *Simulated data*

We simulated $n = 1024$ observations from the bivariate times series, $x_t = (x_{1t}, x_{2t})'$,

$$x_t = \Phi_1 x_{t-1} + \Phi_2 x_{t-2} + z_t,$$

in which

$$\Phi_1 = \begin{pmatrix} 0·5 & 0 \\ 0 & -0·3 \end{pmatrix}, \Phi_2 = \begin{pmatrix} 0 & 0 \\ 0 & -0·5 \end{pmatrix} \quad \text{and} \quad z_t \sim N \left\{ 0, \Sigma = \begin{pmatrix} 1 & 0·9 \\ 0·9 & 1 \end{pmatrix} \right\}.$$

The z_t were generated independently. The 2×2 spectral matrix of the process is (Shumway & Stoffer, 2006, Ch. 4)

$$f(\nu) = \Phi^{-1}(\nu) \, \Sigma \, \Phi^{*-1}(\nu),$$

where

$$\Phi(\nu) = I - \Phi_1 \exp(-2\pi i \nu) - \Phi_2 \exp(-4\pi i \nu).$$

From this fact, we may calculate the elements of the spectral matrix of this process:

$$f_{11}(\nu) = \{1·25 - \cos(2\pi \nu)\}^{-1},$$

$$f_{22}(\nu) = \{1·34 + 0·9 \cos(2\pi \nu) + \cos(4\pi \nu)\}^{-1},$$

$$f_{12}(\nu) = 0·9[0·85 - 0·45 \cos(2\pi \nu) + 0·5 \cos(4\pi \nu) + i\{0·5 \sin(4\pi \nu) + 0·55 \sin(2\pi \nu)\}]^{-1}$$

and $f_{21}(\nu) = f_{12}^*(\nu)$. In addition, the squared coherency between the component processes is constant:

$$\rho_{12}^2(\nu) = \frac{|f_{12}(\nu)|^2}{f_{11}(\nu) f_{22}(\nu)} = 0·81.$$

Figure 1 displays the periodogram ordinates of the simulated series, the true spectral components, $f_{11}(\nu)$ and $f_{22}(\nu)$, and the corresponding estimated spectra along with 99% pointwise credible intervals, all on a log scale. These intervals are obtained as the 0·005 and 0·995 empirical percentiles of the Markov chain Monte Carlo iterates, after the burn-in period, of $X\gamma_1$ and $X\gamma_2$ at each ν_k. For each example, the Markov chain Monte Carlo procedure was run for a total of 2000 iterations with a burn-in period of 1000 iterations. In addition, the estimated squared coherency, not shown, is approximately a straight line around the true squared coherency, $\rho_{12}^2(\nu) = 0·81$. This example demonstrates that our technique is a viable way of obtaining smoothed spectral estimates of multivariate processes.

Some comments are in order. First, note that the spectral density of the second component, $f_{22}(\nu)$, has a fairly sharp peak around $\nu = 0·28$. This fact is obscured in Fig. 1 because, to improve the visual impact, the spectra are plotted on a log scale, and in landscape mode; spectra are typically plotted this way. Secondly, because changing the order of the time series will change the Cholesky decomposition of the spectral matrix, we investigated the possibility that the ordering matters. In particular, we repeated this simulation with the components reversed, that is, with $x_t = (x_{2t}, x_{1t})'$, instead of $x_t = (x_{1t}, x_{2t})'$. Although the components of the Cholesky decomposition were changed, the final estimates of the spectra and cross-spectra did not change. Finally, to monitor convergence, we ran four chains with dispersed random starting points. We then looked at iteration plots, i.e. the iterates from the four chains for a given parameter as a function
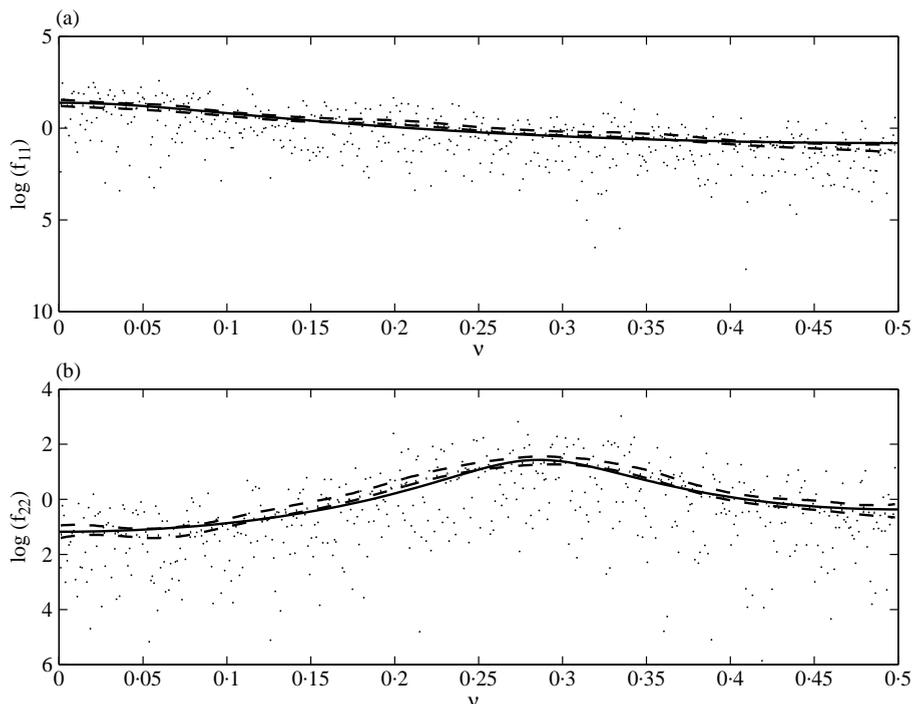
Fig. 1. Simulated data. The true spectra (solid line), the periodogram ordinates (dots) and 99% credible intervals (dashed lines), all on a log scale, for (a) f_{11} and (b) f_{22}.

of the iteration number. In all the plots, following the burn-in period, the within-chain variation was in close agreement with the between-chain variation, i.e. the four chains were almost indistinguishable.

4·2. *The El Niño cycle*

Throughout Shumway & Stoffer (2006), two simultaneously recorded series are used to explore the El Niño cycle. Figure 2 shows these series, which are monthly values of the Southern Oscillation Index and associated Recruitment, i.e. number of new fish, for a period of 453 months ranging over the years 1950–1987. The Southern Oscillation Index measures changes in air pressure related to sea surface temperatures in the central Pacific Ocean. The central Pacific Ocean warms every 3 to 7 years because of the El Niño effect. Both series exhibit regularly repeating cycles, and the two series are related because the fish spawn in colder waters.

Figure 3 presents the estimated spectra, i.e. the diagonal elements of the estimated spectral matrix, as a function of frequency. Both spectra have peaks at about the same frequencies. One is at $\nu = 1/12$ cycles per month, which is the obvious yearly cycle. The other peak at about $\nu = 1/48$ represents a possible El Niño effect. The yearly cycle has more power in the Southern Oscillation Index series. In addition, both spectra show small peaks at the harmonic frequencies of the yearly cycle.

To examine how the two series are related, we also plotted the estimated squared coherency in Fig. 4. This plot shows that the El Niño and yearly cycles are strongly coherent. Other frequencies are also coherent but this is less meaningful, because they are harmonics of the yearly cycle and the power spectra at the harmonics, see Fig. 3, are small.
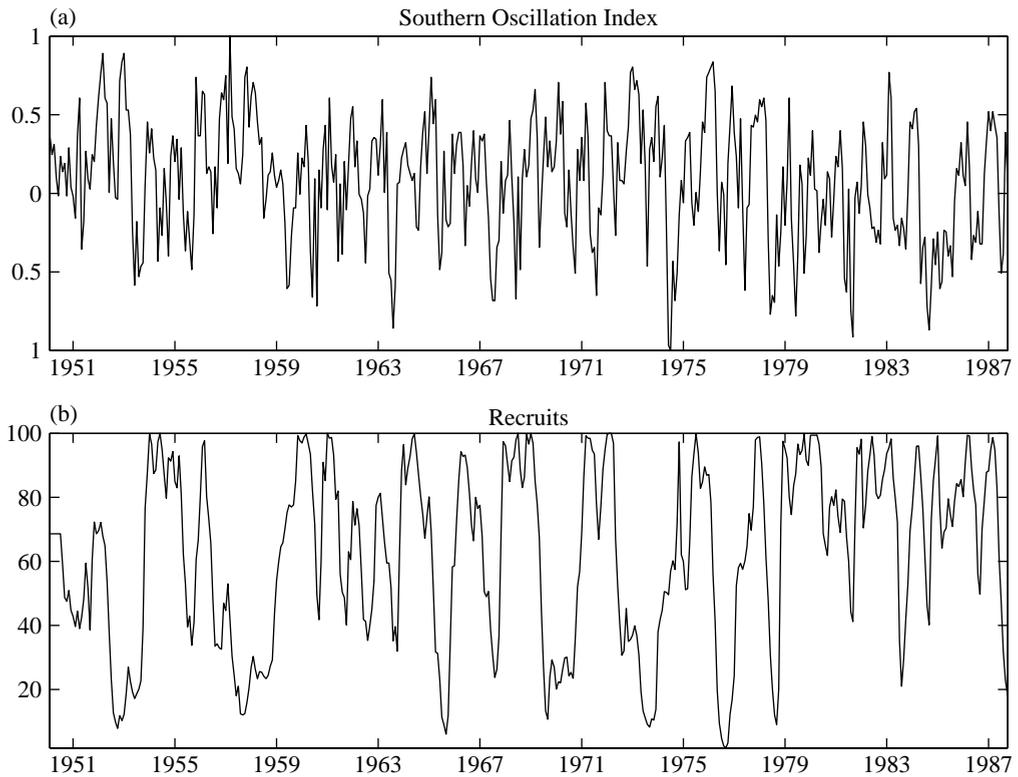
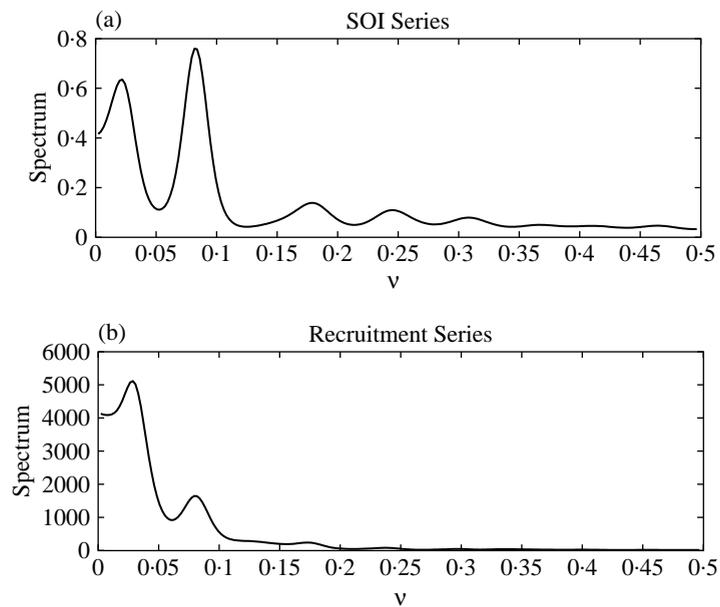Fig. 2. Monthly (a) Southern Oscillation Index and (b) Recruitment series for the years 1950–1987.

Fig. 3. Individual estimated spectra of the (a) Southern Oscillation Index and (b) recruitment series.
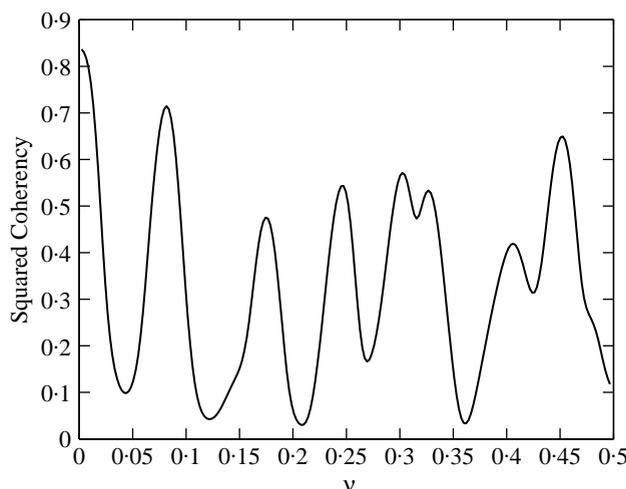
Fig. 4. Squared coherency function relating Southern Oscilla-
tion Index to recruitment.

4·3. *Spectral envelope and* DNA *sequences*

In some applications, an investigator may be more interested in a function of the components of a spectral matrix, such as coherency, described in the previous subsections. We now investigate the case of estimating eigenvalues and eigenvectors of the spectral matrix. In particular, we focus on estimating the spectral envelope, i.e. an eigenvalue, and the corresponding scaling, i.e. an eigenvector, of a DNA nucleotide sequence. In this case, the data are a sequence of multivariate 3×1 indicator vectors that correspond to the sequence of nucleotides. For brevity, we refer the reader to Stoffer et al. (1993) or Shumway & Stoffer (2006, §7·9) for a discussion of the spectral envelope.

In this example, we analyze part of the BNRF1 gene in Herpesvirus saimiri; the data are taken from GenBank. This particular coding sequence occurs from bp 6821 to bp 10 561, where 'bp' means 'base-pair'. As in Stoffer (2002), we estimate the spectral envelope for 1000 bps starting at bp 8820 of the coding sequence. Figure 5 displays the estimated spectral envelope corresponding to this subset of the sequence. The spectral envelope picks up a signal at one cycle every three bps, which occurs often in coding sequences we have analyzed. There is another peak in the spectral envelope indicating a signal at one cycle every 10 bps. This signal is particularly interesting because, while the double helix makes one turn about every 10 base-pairs, the 10 bps signal is rarely seen and the importance of this twisting is not clear.

Finally, it is worthwhile to look at the scalings corresponding to each peak. For $\nu = 1/10$, the scalings for the bases are A = 2·12, C = 1·93, G = 0·17 and T = 0. This suggests that the signal is attributed to the M-K alphabet, where M = A or C and K = G or T is the complement of M. This structure itself is of interest. The relationship between A and C is that both have aMino, hence the M, groups at the ring position most distant from the point of attachment to the sugar, and the relationship between G and T is that both have Keto, hence the K, groups at the corresponding position. For $\nu = 1/3$, the scalings are A = 2·28, C = 0·89, G = 2·16 and T = 0. This suggests that the signal may be attributed to the common RYY alphabet, where R denotes a puRine, A or G, and Y denotes a pYrimidine, C or T.
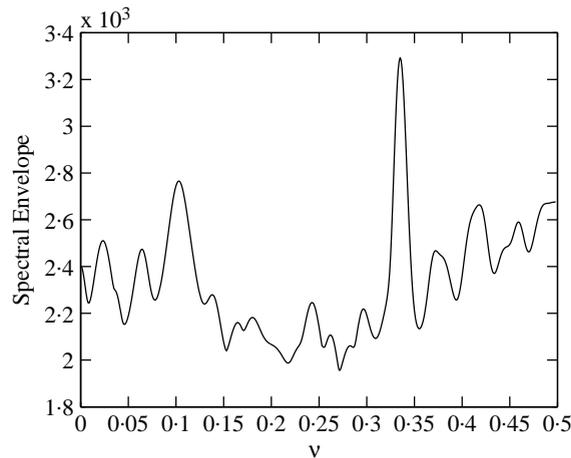
Fig. 5. Spectral envelope for part of a coding sequence
in Herpesvirus saimiri.

APPENDIX

Technical Details

Details of the Sampling Scheme. In this appendix we give more details about the conditional distributions of $\gamma_{il(re)}$ and $\gamma_{il(im)}$, $i = 2, \ldots, p$, $l = 1, \ldots, i-1$, used for performing the Gibbs sampler. We use the notation established in §§2 and 3.

Bivariate time series. Let

$$v_{21(re)} = \sum_{k=1}^{N} \delta_{2k}^{-2}(y_{2k}^* y_{1k} + y_{1k}^* y_{2k})x_k,$$

$$v_{21(im)} = \sum_{k=1}^{N} \delta_{2k}^{-2}i(y_{2k}^* y_{1k} - y_{1k}^* y_{2k})x_k,$$

$$A_{21(\cdot)} = \mathrm{diag}\left\{\frac{1}{2}(\sigma_\alpha^{-2}, \sigma_\alpha^{-2}, \tau_{21(\cdot)}^{-2}, \ldots, \tau_{21(\cdot)}^{-2})\right\},$$

for $\cdot = re, im$, and

$$B_{21} = \sum_{k=1}^{N} \delta_{2k}^{-2}|y_{1k}|^2 x_k x_k',$$

where y_{jk}^* denotes the complex conjugate of y_{jk} and $i = \sqrt{(-1)}$. Then

$$\gamma_{21(re)}\big|\gamma_2, \tau_{21(re)}^2, Y \sim N(\mu_{21(re)}, \Sigma_{21(re)}), \tag{A1}$$

$$\gamma_{21(im)}\big|\gamma_2, \tau_{21(im)}^2, Y \sim N(\mu_{21(im)}, \Sigma_{21(im)}), \tag{A2}$$

where $\Sigma_{21(\cdot)} = \frac{1}{2}(A_{21(\cdot)} + B_{21})^{-1}$ and $\mu_{21(\cdot)} = \Sigma_{21(\cdot)}v_{21(\cdot)}$, for $\cdot = re, im$.

The conditional distribution for drawing γ_j, $j = 1, 2$, is given in (6). To draw from it, we use a Metropolis–Hastings step with a multivariate normal proposal distribution. The mean and variance-covariance matrix of this multivariate normal distribution are the maximizer of (6) with respect to γ_j, and the inverse of the negative hessian of (6) evaluated at this maximizer, respectively. Thus, the sampling scheme consists of the following steps.

Step 1. Draw γ_j from (6), for $j = 1, 2$.

Step 2. Draw $\gamma_{21(re)}$ and $\gamma_{21(im)}$ from (A1) and (A2), respectively.

Step 3. Draw τ_j^2, $j = 1, 2$, $\tau_{21(re)}^2$ and $\tau_{21(im)}^2$ from the inverse Gamma distributions described in §3.

Trivariate time series. The vectors $\gamma_{21(re)}$ and $\gamma_{21(im)}$ are drawn according to (A1) and (A2), respectively. Let

$$v_{31(re)} = \sum_{k=1}^{N} \delta_{3k}^{-2} \{y_{3k}^* y_{1k} + y_{3k} y_{1k}^* - x_k' \gamma_{32(re)}(y_{1k}^* y_{2k} + y_{2k}^* y_{1k})\} x_k,$$

$$v_{31(im)} = \sum_{k=1}^{N} \delta_{3k}^{-2} \{i(y_{3k}^* y_{1k} - y_{3k} y_{1k}^*) - x_k' \gamma_{32(im)}(y_{1k}^* y_{2k} + y_{2k}^* y_{1k})\} x_k,$$

$$B_{31} = \sum_{k=1}^{N} \delta_{3k}^{-2} |y_{1k}|^2 x_k x_k',$$

$$v_{32(re)} = \sum_{k=1}^{N} \delta_{3k}^{-2} \{y_{3k}^* y_{2k} + y_{3k} y_{2k}^* - x_k' \gamma_{31(re)}(y_{1k}^* y_{2k} + y_{2k}^* y_{1k})\} x_k,$$

$$v_{32(im)} = \sum_{k=1}^{N} \delta_{3k}^{-2} \{i(y_{3k}^* y_{2k} - y_{3k} y_{2k}^*) - x_k' \gamma_{31(im)}(y_{1k}^* y_{2k} + y_{2k}^* y_{1k})\} x_k,$$

$$B_{32} = \sum_{k=1}^{N} \delta_{3k}^{-2} |y_{2k}|^2 x_k x_k'.$$

For $i = 3, l = 1, 2$ and $\cdot = re, im$, let $A_{il(\cdot)} = \text{diag}\left\{\frac{1}{2}(\sigma_\alpha^{-2}, \sigma_\alpha^{-2}, \tau_{il(\cdot)}^{-2}, \ldots, \tau_{il(\cdot)}^{-2})\right\}$. Then

$$\gamma_{31(re)} \mid \gamma_{32(re)}, \gamma_3, \tau_{31(re)}^2, Y \sim N(\mu_{31(re)}, \Sigma_{31(re)}),$$

$$\gamma_{31(im)} \mid \gamma_{32(im)}, \gamma_3, \tau_{31(im)}^2, Y \sim N(\mu_{31(im)}, \Sigma_{31(im)}),$$

$$\gamma_{32(re)} \mid \gamma_{31(re)}, \gamma_3, \tau_{32(re)}^2, Y \sim N(\mu_{32(re)}, \Sigma_{32(re)})$$

$$\gamma_{32(im)} \mid \gamma_{31(im)}, \gamma_3, \tau_{32(im)}^2, Y \sim N(\mu_{32(im)}, \Sigma_{32(im)}),$$

where $\Sigma_{il(\cdot)} = \frac{1}{2}(A_{il(\cdot)} + B_{il})^{-1}$ and $\mu_{il(\cdot)} = \Sigma_{il(\cdot)} v_{il(\cdot)}$, for $i = 3, l = 1, 2$ and $\cdot = re, im$. The different parameters are drawn in an analogous fashion to that described in the bivariate case.

REFERENCES

BRILLINGER, D. R. (2001). *Time Series: Data Analysis and Theory*, 2nd ed. Philadelphia, PA: SIAM.

DAI, M. & GUO, W. (2004). Multivariate spectral analysis using Cholesky decomposition. *Biometrika* **91**, 629–43.

DANIELS, M. J. & POURAHMADI, M. (2002). Bayesian analysis of covariance matrices and dynamic models for longitudinal data. *Biometrika* **89**, 553–66.

EUBANK, R. L. (1999). *Nonparametric Regression and Spline Smoothing*, 2nd ed. New York: Marcel Dekker.

FRANKE, J. & HARDLE, W. (1992). On bootstrapping kernel spectral estimates. *Ann. Statist.* **20**, 121–45.

PAWITAN, Y. (1996). Automatic estimation of the cross-spectrum of a bivariate time series. *Biometrika* **83**, 419–32.

POURAHMADI, M. (1999). Joint mean-covariance models with applications to longitudinal data: unconstrained parameterisation. *Biometrika* **86**, 677–90.

POURAHMADI, M. (2000). Maximum likelihood estimation of generalised linear models for multivariate normal covariance matrix. *Biometrika* **87**, 425–35.

POURAHMADI, M. & DANIELS, M. (2002). Dynamic conditionally linear mixed models for longitudinal data. *Biometrics* **58**, 82–8.

SHUMWAY, R. H. & STOFFER, D. S. (2006). *Time Series Analysis and its Applications: With R Examples*, 2nd ed. New York: Springer.

STOFFER, D. S., TYLER, D. E. & MCDOUGALL, A. J. (1993). Spectral analysis for categorical time series: scaling and the spectral envelope. *Biometrika* **80**, 611–22.

STOFFER, D. S. (2002). Nonparametric frequency detection and optimal coding in molecular biology. In *Modeling Uncertainty: An Examination of Stochastic Theory, Methods, and Applications*, Ed. M. Dror, P. L'Ecuyer & F. Szidarovszky, pp. 129–54, Boston, MA: Kluwer Academic Publishers.

THOMSON, D. J. (1982). Spectrum estimation and harmonic analysis. *Proc. IEEE* **70**, 1055–96.

WHITTLE, P. (1957). Curve and periodogram smoothing. *J. R. Statist. Soc.* B **19**, 38–47.

WU, W. B. & POURAHMADI, M. (2003). Nonparametric estimation of large covariance matrices of longitudinal data. *Biometrika* **90**, 831–44.