UNIVERSITY OF LONDON
IMPERIAL COLLEGE OF SCIENCE, TECHNOLOGY AND MEDICINE

BSc and MSci EXAMINATIONS (MATHEMATICS) MAY-JUNE 2001

This paper is also taken for the relevant examination for the Associateship

M3S8/M4S8 TIME SERIES

DATE: Tuesday, 5th June 2001 TIME: 2 pm — 4 pm

Credit will be given for all questions attempted, but extra credit will be given for complete or nearly complete answers.

Calculators may not be used.

Note: throughout this paper $\{\epsilon_t\}$ is a sequence of uncorrelated random variables having zero mean and variance σ_ϵ^2, unless stated otherwise.

1. What is meant by saying that a stochastic process is second-order stationary?

   Determine whether the following stochastic processes are second-order stationary, giving full justification:

   a)
$$X_t = (3/4)X_{t-1} - (1/8)X_{t-2} + \epsilon_t.$$

   b)
$$Y_t = X_t + C,$$

   where $\{X_t\}$ is a second-order stationary process with zero mean and autocovariance sequence $\{s_{\tau,X}\}$, and C is a random variable with zero mean and variance σ_C^2, uncorrelated with X_t for all t.

   c)
$$X_t = \mu + \sum_{l=1}^{L} D_l \cos(2\pi f_l t + \phi_l),$$

   where $\mu, D_1, ..., D_L, f_1, ..., f_L$ are real-valued constants, and the phases $\{\phi_l\}$ are independent and identically distributed with probability density function

$$g(\phi_l) = \frac{1}{2\pi}(1 + \sin\,\phi_l), \qquad |\phi_l| \le \pi.$$

2. a) Let $\{X_t\}$ be a real-valued zero mean second-order stationary moving-average process, or MA(2),

$$X_t = \epsilon_t - \theta_{1,2}\epsilon_{t-1} - \theta_{2,2}\epsilon_{t-2}.$$

Derive the frequency response function $G(f)$ associated with the LTI filter given by

$$L\{\epsilon_t\} = \epsilon_t - \theta_{1,2}\epsilon_{t-1} - \theta_{2,2}\epsilon_{t-2},$$

and hence show that the spectral density function for the MA(2) process $\{X_t\}$ can be written

$$S_X(f) = \sigma_\epsilon^2[(1+\theta_{1,2}^2+\theta_{2,2}^2) - 2\theta_{1,2}(1-\theta_{2,2})\cos(2\pi f) - 2\theta_{2,2}\cos(4\pi f)].$$

b)
(i) What is meant by saying that a moving-average process is invertible?
(ii) How may we check whether such a process is invertible?
(iii) Consider the second-order stationary moving average processes $\{Y_t\}$ and $\{Z_t\}$ defined by

$$Y_t = \epsilon_t - \epsilon_{t-1} + \tfrac{1}{4}\epsilon_{t-2},$$

and

$$Z_t = \epsilon_t - 4\epsilon_{t-1} + 4\epsilon_{t-2}.$$

Determine whether $\{Y_t\}$ and $\{Z_t\}$ are invertible processes.

c) Find the spectral density function for each of the processes $\{Y_t\}$ and $\{Z_t\}$. Of what is this result an example? What is its practical significance?

3. Let $\{X_t\}$ be a zero mean stationary stochastic process, with spectral density function $S_X(f)$.

a) Specify the three conditions which must be satisfied by a linear time-invariant (LTI) digital filter.
b) If a filter has transfer function $G(f)$, what is the spectrum $S_Y(f)$ of the output in terms of the spectrum $S_X(f)$ of the input?
c) Show that the first-order backward difference process

$$Y_t = X_t - X_{t-1}$$

has the spectral density function $S_Y(f)$ given by

$$S_Y(f) = 4\sin^2(\pi f)\, S_X(f)\,.$$

Does a first-order backward difference filter resemble a low-pass or high-pass filter?
d) If a two coefficient filter is now applied to the process $\{Y_t\}$ to give $\{Z_t\}$ according to

$$Z_t = aY_t + bY_{t-1},$$

where

$$a = \frac{1-\sqrt{3}}{4} \text{ and } b = -\frac{1+\sqrt{3}}{4}$$

show that

$$S_Z(f) = \sin^2(\pi f)[1 + 2\cos^2(\pi f)]\, S_X(f)\,.$$

where $S_Z(f)$ is the spectral density function of $\{Z_t\}$.

4. Give one advantage and one disadvantage accruing from the use of a single data taper in spectrum analysis.

Let $\{h_t\}$ be a real-valued taper, standardized so that $\sum_{t=1}^{N} h_t^2 = 1$.

a) Suppose that $X_1, ..., X_N$ is a sample of length N of a second-order stationary process $\{X_t\}$ with known mean μ. The direct spectral estimator with the mean subtracted before the time series is tapered is defined as

$$\hat{S}^{(d)}(f) = \left| \sum_{t=1}^{N} h_t (X_t - \mu) e^{-i2\pi f t} \right|^2 .$$

Use the spectral representation $X_t - \mu = \int_{-1/2}^{1/2} e^{i2\pi f t} dZ(f)$, to show that the mean of the direct spectral estimator $\hat{S}^{(d)}(f)$ is given by

$$E\{\hat{S}^{(d)}(f)\} = \int_{-1/2}^{1/2} \mathcal{H}(f - f') S(f') df' ,$$

where $\mathcal{H}(f) = \left| \sum_{t=1}^{N} h_t e^{-i2\pi f t} \right|^2$.

b) Now suppose that $X_1, ..., X_N$ is in fact a segment of length N of a *white noise* process (i.e., $X_t = \epsilon_t$) with spectral density function $S(f) = \sigma_\epsilon^2$ and nonzero mean μ. If the ϵ_t's are not centred before computing the direct spectral estimator, show that for $0 < f < 1/2$ the mean of the resulting direct estimator,

$$\tilde{S}^{(d)}(f) = \left| \sum_{t=1}^{N} h_t \epsilon_t e^{-i2\pi f t} \right|^2 ,$$

is given by

$$E\{\tilde{S}^{(d)}(f)\} = \sigma_\epsilon^2 + \mu^2 \mathcal{H}(f).$$

c) What does a comparison of $E\{\hat{S}^{(d)}(f)\}$ and $E\{\tilde{S}^{(d)}(f)\}$ tell us about $\int_{-1/2}^{1/2} \mathcal{H}(f) df$?

5. a) Carefully explain what is meant by multitaper spectrum estimation. What benefits arise from using multitapering rather than single tapering in spectrum estimation?

b) Consider a real-valued sequence $h_1, \ldots, h_N$, chosen to maximize the fraction of energy, $\beta(W)$, concentrated in the frequency band $|f| \leq W < 1/2$:

$$\beta(W) = \frac{\int_{-W}^{W} |H(f)|^2 df}{\int_{-1/2}^{1/2} |H(f)|^2 df},$$

where $H(f) = \sum_{t=1}^{N} h_t e^{-i2\pi ft}$.

(i) Show that

$$\beta(W) = \frac{\sum_{j=1}^{N} \sum_{k=1}^{N} h_k \frac{\sin[2\pi W(j-k)]}{\pi(j-k)} h_j}{\sum_{t=1}^{N} h_t^2}.$$

(ii) Define the matrix $\mathbf{A}$ as the $N \times N$ matrix with (j,k)th element given by $\sin[2\pi W(j-k)]/[\pi(j-k)]$. Using the result that

$$\frac{d}{d\mathbf{x}} \mathbf{x}^T \mathbf{A} \mathbf{x} = 2\mathbf{A}\mathbf{x},$$

show that the sequence $h_1, \ldots, h_N$, that maximizes $\beta(W)$ is the eigenvector corresponding to the largest eigenvalue of $\mathbf{A}$.

(iii) Using this approach, how are the other discrete prolate tapers of length N derived, and why are they orthogonal?

SOLUTIONS

M3S8/M4S8 TIME SERIES

DATE: Someday, 0th June 2001 TIME: x pm — y pm

Soln. 1. $\{X_t\}$ is second-order stationary if $E\{X_t\}$ is a finite constant for all t, var$\{X_t\}$ is a finite constant for all t, and $cov\{X_t, X_{t+\tau}\} = s_\tau$, a finite quantity depending only on τ and not on t.

a) We have

$$X_t - (3/4)X_{t-1} + (1/8)X_{t-2} = \epsilon_t.$$

The characteristic polynomial for this AR(2) process is

$$\Phi(z) = 1 - (3/4)z + (1/8)z^2 = (1 - \frac{1}{4}z)(1 - \frac{1}{2}z).$$

The roots are $z = 4$ and $z = 2$, i.e., both outside the unit circle, so the process is stationary.

b) First, we have $E\{Y_t\} = E\{X_t + C\} = E\{X_t\} + E\{C\} = 0$, so $E\{Y_t\}$ is independent of t. Next, we have

$$\begin{aligned}\text{cov}\,\{Y_t, Y_{t+\tau}\} &= E\{(X_t + C)(X_{t+\tau} + C)\} \\ &= E\{X_t X_{t+\tau}\} + E\{X_t C\} + E\{X_{t+\tau} C\} + E\{C^2\} \\ &= s_{\tau,X} + \sigma_C^2,\end{aligned}$$

(because $E\{(X_t - E\{X_t\})(C - E\{C\})\} = 0$ for all t), which is also independent of t, so $\{Y_t\}$ is a stationary process with acvs given by $s_{\tau,Y} \equiv s_{\tau,X} + \sigma_C^2$ (in the above, $E\{X_t C\} = E\{X_{t+\tau} C\} = 0$ because C is uncorrelated with X_t for all t).

c) For the first moment, we have

$$E\{X_t\} = \mu + \sum_{l=1}^{L} D_l E\{\cos(2\pi f_l t + \phi_l)\}.$$

Now since a cosine integrates to zero over a whole period,

$$\begin{aligned}E\{\cos(2\pi f_l t + \phi_l)\} &= \frac{1}{2\pi}\int_{-\pi}^{\pi} \cos(2\pi f_l t + \phi_l)(1 + \sin(\phi_l))\, d\phi_l \\ &= \frac{1}{2\pi}\int_{-\pi}^{\pi} \cos(2\pi f_l t + \phi_l)\sin(\phi_l)\, d\phi_l.\end{aligned}$$

But since $\cos A \sin B = \{\sin(A+B) - \sin(A-B)\}/2$, we have that $\cos(2\pi f_l t + \phi_l)\sin(\phi_l) = \{\sin(2\pi f_l t + 2\phi_l) - \sin(2\pi f_l t)\}/2$, and hence

$$E\{\cos(2\pi f_l t + \phi_l)\} = \frac{1}{4\pi}\int_{-\pi}^{\pi} \sin(2\pi f_l t + 2\phi_l)d\phi_l - \frac{1}{4\pi}\sin(2\pi f_l t)2\pi$$
$$= 0 - \frac{1}{2}\sin(2\pi f_l t).$$

Hence,

$$E\{X_t\} = \mu - \frac{1}{2}\sum_{l=1}^{L} D_l \sin(2\pi f_l t),$$

which in general is not independent of t, so that the process is non-stationary.

Soln. 2.

a)

$$
\begin{aligned}
L\{\epsilon_t\} &= \epsilon_t - \theta_{1,2}\epsilon_{t-1} - \theta_{2,2}\epsilon_{t-2} \\
\Rightarrow L\{e^{i2\pi ft}\} &= e^{i2\pi ft}(1 - \theta_{1,2}e^{-i2\pi f} - \theta_{2,2}e^{-i4\pi f}) \\
\Rightarrow G(f) &= 1 - \theta_{1,2}e^{-i2\pi f} - \theta_{2,2}e^{-i4\pi f} \\
\Rightarrow |G(f)|^2 &= (1 + \theta_{1,2}^2 + \theta_{2,2}^2) - \theta_{1,2}(e^{-i2\pi f} + e^{i2\pi f}) - \theta_{2,2}(e^{-i4\pi f} + e^{i4\pi f}) \\
&\quad + \theta_{1,2}\theta_{2,2}(e^{-i2\pi f} + e^{i2\pi f}) \\
&= (1 + \theta_{1,2}^2 + \theta_{2,2}^2) - 2\theta_{1,2}(1 - \theta_{2,2})\cos(2\pi f) - 2\theta_{2,2}\cos(4\pi f).
\end{aligned}
$$

Hence,

$$S_X(f) = \sigma_\epsilon^2[(1+\theta_{1,2}^2+\theta_{2,2}^2) - 2\theta_{1,2}(1-\theta_{2,2})\cos(2\pi f) - 2\theta_{2,2}\cos(4\pi f)].$$

b)

(i) An MA process is said to be invertible if it can be written in autoregressive form $\Phi(B)X_t = \epsilon_t$, with the $z-$polynomial $\Phi(z)$ admitting a power series expansion.

(ii) We can check whether an MA process is invertible when written in autoregressive form by determining that the poles of $\Phi(z)$ are all outside the unit circle. For a moving average process this means that the characteristic MA polynomial will have all roots outside the unit circle.

(iii) The characteristic polynomial for the MA(2) process $\{Y_t\}$ is

$$1 - z + (1/4)z^2 = (1 - \frac{1}{2}z)(1 - \frac{1}{2}z).$$

The roots are $z = 2$ (double) and so process is invertible. The characteristic polynomial for the MA(2) process $\{Z_t\}$ is

$$1 - 4z + 4z^2 = (1 - 2z)(1 - 2z).$$

The roots are $z = 1/2$ (double) and so process is non-invertible.

c) For $\{Y_t\}$ we have, w.r.t. (a), $\theta_{1,2} = 1, \theta_{2,2} = -1/4$. Putting these parameter values into the expression for $S_X(f)$ we get

$$S_Y(f) = \frac{\sigma_\epsilon^2}{16}[33 - 40\cos(2\pi f) + 8\cos(4\pi f)].$$

For $\{Z_t\}$ we have, w.r.t. (a), $\theta_{1,2} = 4, \theta_{2,2} = -4$. Putting these parameter values into the expression for $S_X(f)$ we get

$$S_Z(f) = \sigma_\epsilon^2[33 - 40\cos(2\pi f) + 8\cos(4\pi f)].$$

The spectra have the same shape, differing only in a constant of proportionality. This is an example of the non-identifiability of the model from the spectrum (or autocovariance). Inverting the roots leaves the spectral shape unchanged. The practical significance is that more information is required to identify the parameter values than simply the spectrum (or autocovariance).

Soln. 3.

a) A digital filter L that transforms an input sequence $\{X_t\}$ into an output sequence $\{Y_t\}$ is called a linear time-invariant digital filter if it has the following three properties:

[1] Scale preservation:

$$L\{\alpha X_t\} = \alpha L\{X_t\}.$$

[2] Superposition:

$$L\{X_{t,1} + X_{t,2}\} = L\{X_{t,1}\} + L\{X_{t,2}\}.$$

[3] Time invariance:

$$\text{if } L\{X_t\} = \{Y_t\}, \text{ then } L\{X_{t+\tau}\} = \{Y_{t+\tau}\},$$

where τ is integer-valued and the notation $\{X_{t+\tau}\}$ refers to the sequence whose tth element is $X_{t+\tau}$.

b) $S_Y(f) = |G(f)|^2 S_X(f)$.

c) The filter is defined by $L\{X_t\} = X_t - X_{t-1}$. The transfer function is obtained by inputting $X_t = \exp(i2\pi ft)$:

$$L\{e^{i2\pi ft}\} = e^{i2\pi ft} - e^{i2\pi f(t-1)} = e^{i2\pi ft}\left(1 - e^{-i2\pi f}\right) = e^{i2\pi ft}G_1(f),$$

where $G_1(f) \equiv 1 - \exp(-i2\pi f)$ is the transfer function. Now

$$|G_1(f)|^2 = \left|e^{-i\pi f}(e^{i\pi f} - e^{-i\pi f})\right|^2 = \left|e^{-i\pi f}2i\sin(\pi f)\right|^2 = 4\sin^2(\pi f).$$

Hence, $S_Y(f) = 4\sin^2(\pi f)S_X(f)$.

Now $|G_1(f)|^2$ increases from 0 to Nyquist (1/2) so the first difference filter resembles a high-pass filter.

d) The filter is defined by $L\{X_t\} = aX_t + bX_{t-1}$. The transfer function is obtained by inputing $X_t = \exp(i2\pi ft)$:

$$L\{e^{i2\pi ft}\} = ae^{i2\pi ft} + be^{i2\pi f(t-1)} = e^{i2\pi ft}\left(a + be^{-i2\pi f}\right) = e^{i2\pi ft}G_2(f),$$

Now $S_Z(f) = |G_2(f)|^2 S_Y(f) = 4\sin^2(\pi f)|G_2(f)|^2 S_X(f)$. But

$$\begin{aligned}
|G_2(f)|^2 &= \left[\frac{1-\sqrt{3}}{4} - \frac{1+\sqrt{3}}{4} e^{-i2\pi f}\right]\left[\frac{1-\sqrt{3}}{4} - \frac{1+\sqrt{3}}{4} e^{i2\pi f}\right] \\
&= \left[\frac{(1-\sqrt{3})^2}{16} - \frac{(1-\sqrt{3})(1+\sqrt{3})}{16}\left(e^{-i2\pi f} + e^{i2\pi f}\right) + \frac{(1+\sqrt{3})^2}{16}\right] \\
&= \frac{1}{4}(2+\cos(2\pi f)) \\
&= \frac{1}{4}(1+2\cos^2(\pi f)),
\end{aligned}$$

since $\cos(2x) = 2\cos^2(x) - 1$. Hence,

$$S_Z(f) = \sin^2(\pi f)(1+2\cos^2(\pi f))S_X(f).$$

Soln. 4. The main advantage of data tapering in spectrum analysis is reduction in sidelobe leakage by changing the default Fejér blurring kernel into something with smaller sidelobes. Main disadvantages: the main lobe of the blurring kernel is made wider, reducing resolution. [Other correct answers accepted].

a) Let

$$J(f) \equiv \sum_{t=1}^{N} h_t (X_t - \mu) e^{-i2\pi f t}.$$

By the spectral representation theorem

$$X_t - \mu = \int_{-1/2}^{1/2} e^{i2\pi f' t} \, dZ(f'),$$

where $\{Z(\cdot)\}$ is a process with orthogonal increments, and $E\{dZ(f)\} = 0$. Thus

$$\begin{aligned} J(f) &= \sum_{t=1}^{N} h_t \left(\int_{-1/2}^{1/2} e^{i2\pi f' t} \, dZ(f') \right) e^{-i2\pi f t} \\ &= \int_{-1/2}^{1/2} \sum_{t=1}^{N} h_t e^{-i2\pi (f - f') t} \, dZ(f') \\ &= \int_{-1/2}^{1/2} H(f - f') \, dZ(f'), \end{aligned}$$

where $\{h_t\}$ and $H(\cdot)$ form a Fourier transform pair under the assumption that $\{h_t\}$ is an infinite sequence with $h_t = 0$ for $t < 1$ and $t > N$; i.e.,

$$H(f) \equiv \sum_{t=1}^{N} h_t e^{-i2\pi f t}.$$

Now it is given that,

$$\hat{S}^{(d)}(f) \equiv |J(f)|^2 = \left| \sum_{t=1}^{N} h_t (X_t - \mu) e^{-i2\pi f t} \right|^2.$$

Because $\{Z(\cdot)\}$ has orthogonal increments, we therefore have

$$E\{\hat{S}^{(d)}(f)\} = \int_{-1/2}^{1/2} \mathcal{H}(f - f')S(f')\,df',$$

where

$$\mathcal{H}(f) \equiv |H(f)|^2 = \left|\sum_{t=1}^{N} h_t e^{-i2\pi ft}\right|^2.$$

b)

$$\begin{aligned} E\{\tilde{S}^{(d)}(f)\} &= E\left\{\left|\sum_{t=1}^{N} h_t \epsilon_t e^{-i2\pi ft}\right|^2\right\} \\ &= \sum_{j=1}^{N}\sum_{k=1}^{N} h_j h_k E\{\epsilon_j \epsilon_k\} e^{-i2\pi f(k-j)}. \end{aligned}$$

But ϵ_t here has a nonzero mean μ, so that $E\{\epsilon_j^2\} = \sigma_\epsilon^2 + \mu^2 \quad (j = k)$ and $E\{\epsilon_j \epsilon_k\} = \mu^2 \quad (j \neq k)$. So

$$\begin{aligned} E\{\tilde{S}^{(d)}(f)\} &= \left[\sigma_\epsilon^2 \sum_{j=1}^{N} h_j^2 + \mu^2 \sum_{j=1}^{N}\sum_{k=1}^{N} h_j h_k e^{-i2\pi f(k-j)}\right]. \\ &= \sigma_\epsilon^2 + \mu^2 \left|\sum_{t=1}^{N} h_t e^{-i2\pi ft}\right|^2 \\ &= \sigma_\epsilon^2 + \mu^2 \mathcal{H}(f). \end{aligned}$$

c) For white noise, $E\{\hat{S}^{(d)}(f)\} = \sigma_\epsilon^2 \int_{-1/2}^{1/2} \mathcal{H}(f)df$, since $\mathcal{H}$ has unit periodicity. When $\mu = 0$, this must be the same as for $E\{\tilde{S}^{(d)}(f)\}$, and hence $\int_{-1/2}^{1/2} \mathcal{H}(f)df = 1$, (which also follows from Parseval's relation since $\sum_{t=1}^{N} h_t^2 = 1$.)

Soln. 5.

a) Multitaper spectrum estimation is carried out by computing a set of K direct spectrum estimates, each using a taper, where the set of K tapers are orthonormal. The set of direct spectrum estimates are averaged to produce the final spectrum estimate.

A single taper will protect the spectrum estimate against side-lobe leakage, but the estimate will look very ragged; multitapering confers the same protection against leakage, but using several tapers reduces the variance (by a factor of K.)

b) (i) We can write $\beta(W)$ as

$$\begin{aligned}
\beta(W) &= \frac{\int_{-W}^{W} \sum_j h_j e^{-i2\pi fj} \sum_k h_k e^{i2\pi fk} df}{\int_{-1/2}^{1/2} \sum_j h_j e^{-i2\pi fj} \sum_k h_k e^{i2\pi fk} df} \\
&= \frac{\sum_j \sum_k h_j h_k \int_{-W}^{W} e^{-i2\pi f(j-k)} df}{\sum_j \sum_k h_j h_k \int_{-1/2}^{1/2} e^{-i2\pi f(j-k)} df} \\
&= \frac{\sum_j \sum_k h_j h_k \{\sin[2\pi f(j-k)]/[2\pi(j-k)]|_{-W}^{W}\}}{\sum_j \sum_k h_j h_k \delta_{j,k}} \\
&= \frac{\sum_j \sum_k h_k \{\sin[2\pi W(j-k)]/[\pi(j-k)]\} h_j}{\sum_j h_j^2}.
\end{aligned}$$

(ii) Write $\beta(W) = \mathbf{h}^T\mathbf{A}\mathbf{h}/\mathbf{h}^T\mathbf{h}$. Then differentiate:

$$\begin{aligned}
\frac{d}{d\mathbf{h}}\beta(W) &= \frac{\mathbf{h}^T\mathbf{h}\frac{d}{d\mathbf{h}}(\mathbf{h}^T\mathbf{A}\mathbf{h}) - \mathbf{h}^T\mathbf{A}\mathbf{h}\frac{d}{d\mathbf{h}}(\mathbf{h}^T\mathbf{h})}{(\mathbf{h}^T\mathbf{h})^2} \\
&= \frac{2\mathbf{A}\mathbf{h}}{\mathbf{h}^T\mathbf{h}} - \beta(W)\frac{2\mathbf{h}}{\mathbf{h}^T\mathbf{h}}.
\end{aligned}$$

Setting the derivative to zero gives $\mathbf{A}\mathbf{h} = \beta(W)\mathbf{h}$, so that the maximum of $\beta(W)$ is the largest eigenvalue of $\mathbf{A}$.

(iii) The next $K-1$ discrete prolate tapers ($K \leq N$) are the next $K-1$ eigenvectors (in decreasing order of their eigenvalues) and are guaranteed orthogonal as they are just eigenvectors.