MSc- Spectrum Estimation Assignment

This assignment requires access to a UNIX workstation or PC running MATLAB together with hardcopy facility.

Results should be recorded in a laboratory notebook.

One of the assignments should be written up as a formal report.

Assignment 1

Title: **Spectrum Estimation**

Aim:

- To compare the properties of conventional Fourier transform-based techniques with modern model-based methods for power spectral density estimation.

Background:

Estimation of the Power Spectral Density (PSD) of a finite length, temporal or spatial, random signal, is a basic preprocessing operation in many applications, for example target analysis in passive sonar, speech recognition, seismic prospecting, measurement of multiphase flow and performance monitoring. Classical spectrum estimation is based upon the use of the Fast Fourier Transform (FFT) and makes no model assumption for the measurement signal. Some windowed portion of the measurement is applied directly to the FFT, the output of which is used to estimate the PSD. Modern spectrum estimation, however, is based upon the assumption that the measurement signal is generated by a prescribed model which is characterised by a number of parameters. Modern spectrum estimation, therefore, involves algorithms to estimate from the measurement the number and values of these model parameters.

The properties of classical and modern spectrum estimation techniques govern their suitability for a particular application. Therefore, these will be investigated in this assignment. Further background information pertaining to this assignment can be found in section 4.3, 4.4, 7.3 and 7.4 of [Kay, 1988].

Theoretically, the PSD, denoted $P_{xx}(f)$, for a real, ergodic, random process $\{x(k)\}$ is given by

$$P_{xx}(f) = \sum_{m=-\infty}^{\infty} r_{xx}(m) e^{-j2pf} \quad (1)$$

where $f \in (-\frac{1}{2}, \frac{1}{2}]$ is the normalised frequency, note the sampling interval is assumed to be unity, and

$$r_{xx}(m) = E\{x(k)x(k+m)\} \quad (2)$$

is the autocorrelation function, which for an ergodic process [Kay, 1988] can be calculated from

$$r_{xx}(m) = \lim_{N\to\infty} \frac{1}{2N+1} \sum_{k=-N}^{N} x(k)x(k+m) \quad (3)$$

Equation (3) shows that it is necessary to have knowledge of $\{x(k)\}$ for all k to calculate exactly the PSD, whereas, in practice, only N samples of $\{x(k)\}$, k=0,…,N-1 will be available. This is the basis of spectrum estimation, namely, to make the best estimate of the true PSD from a finite length sample. The major difference between classical and modern spectrum estimation techniques is the underlying assumption for the samples of $\{x(k)\}$ outside of the N observed values. Classical methods generally assume the unknown samples are zero, whilst the model-based techniques extrapolate the autocorrelation function with the assumption that the missing values maximise the entropy (randomness) of the underlying process [Kay, 1988].

The assumption that the N measurement samples are from an ergodic process (as in this assignment) also has major practical importance. Ergodicity implies statistical stationarity. Therefore, if the PSD is to be estimated for a real signal, such as a sonar measurement, with statistical properties which are slowly time-varying, the N samples must be drawn from an interval over which the signal can be assumed to be locally stationary. For further details on the analysis of non- stationary signals see [Bendat, 1986].

Project 1 Classical Methods

The simplest classical PSD estimator is called the periodogram which is calculated from

$$\hat{P}_{xx}(f) = \frac{1}{N}\left|\sum_{n=0}^{N-1} x(n)e^{-j2pfn}\right|^2 \quad (4)$$

in which it is assumed that the underlying, infinite length, process $\{x(k)\}$ is multiplied by a length N rectangular window, i.e. $w(k)$=1, k=0,…,N-1 and $w(k)$=0 for all other k. Interestingly, equation (4) can be rewritten as

$$\hat{P}_{xx}(f) = \frac{1}{N} \sum_{m=-(N-1)}^{N-1} \hat{r}_{xx}(m)e^{-j2pfn} \quad (5)$$

where $$\hat{r}_{xx}(m) = \sum_{n=0}^{N-1-|m|} x(n)x(n+|m|) \quad (6)$$

which is the sample autocorrelation function. The similarity between equations (5) and (6), and equations (1) and (3) is the reason that the periodogram is termed the natural estimate of the PSD.

The bias of the periodogram equation (4), namely the difference between the average of the periodogram estimate and the true PSD given by equation (1), as a function of frequency becomes

$$BIAS(f) = \frac{1}{N} P_{xx}(f) * |W(f)|^2 - P_{xx}(f) \quad (7)$$

where $|W(f)|^2 = \left| \sum_{n=0}^{N-1} w(n) e^{-j2pfn} \right|^2$, which for a rectangular window equals $\left(\frac{\sin pfN}{\sin pf} \right)^2$, and the operation * denotes convolution. Inspection of equation (7) shows that in the convolution operation the window, with main lobe width 1/*N*, smears the true PSD which introduces bias into the periodogram estimate. However, in the limit as $N \to \infty$ the main lobe width tends to zero and *W*(*f*) tends to a delta function, hence the bias tends to zero. The periodogram is, as such, termed *asymptotically unbiased.*

The variance of the periodogram as calculated in equation (4) is given approximately by [Kay, 1988]

$$VAR\{\hat{P}_{xx}(f)\} = P_{xx}^2(f) \left[1 + \left(\frac{\sin 2pfN}{N \sin 2pf} \right)^2 \right] \quad (8)$$

for the examples in the following exercises. As $N \to \infty$, equation (8) simplifies to $2P_{xx}^2(f)$ for f=0,1/2 and $P_{xx}^2(f)$ elsewhere. This demonstrates the limitation of the periodogram, namely that even for large *N* its standard deviation is as large as its mean.

The averaged periodogram reduces the variance by dividing the available length *N* data sequence into *K*, possibly overlapping, subsequences of length *M*. The PSD is estimated for each length *M* subsequence with the conventional periodogram and the averaged periodogram PSD estimate is simply the arithmetic mean of these *K* values

$$\hat{P}_{xx}(f) = \frac{1}{K} \sum_{r=0}^{K-1} \hat{P}_{xx}^{(r)}(f) \quad (9)$$

where

$$\hat{P}_{xx}^{(r)}(f) = \frac{1}{M} \left| \sum_{n=0}^{M-1} x(n + r(M - M_o)) e^{-j2pfn} \right|^2 \quad (10)$$

and M_o is the size of the overlap. The effect of averaging a number of smaller periodograms *trades* variance reduction with bias increase. This trade-off is common in estimation theory. To be specific, if there is no overlap between the subsections and the data within each subsection are uncorrelated with the data in adjacent subsections, the variance of the averaged periodogram PSD is K times less than that of a conventional length N periodogram. In practice, there is some correlation between adjacent subsections so the effective reduction in variance is slightly less than K. On the other hand, the reduction in the lengths of the periodograms within the averaged periodogram leads to an increase in bias. It is assumed that N is fixed.

Exercise 1. Write a MATLAB function called *avpgm* which calculates the averaged periodogram PSD estimate given by equations (9) and (10). The function inputs should be the measurement column vector of dimension $N \times 1$, the number of subsequences K, the length of each subsequence M (power of two), the size of the overlap M_o and the number of points in the output PSD, N_{FFT}, default 512. Note, the *fft* routine can be used to calculate the PSD at discrete frequencies $f = 0:1/N_{FFT}:(N_{FFT}-1)/N_{FFT}$. Zero padding should be used to increase the subsequence length M to $N_{FFT.}$ The routine should ensure that $(N- M_o)/(M- M_o)$ is an integer so that each subsequence is full. If this is not the case the length of the input sequence should be reduced accordingly.

a) Use the MATLAB *filter* command to generate a narrowband fourth-order AutoRegressive (AR) random signal, y, with poles at $\{0.99e^{\pm j2p(88/360)}, 0.99e^{\pm j2p(92/360)}\}$. The pole positions can be used to calculate the denominator coefficient vector a, set the numerator coefficient vector b to be [1]. Employ the *randn* function to generate 10,000 points of a white Gaussian input x with zero mean and unity variance in column vector format. (Remember, the MATLAB help facility and documentation to obtain more information about the routines, hint see the *exp*, *conj* and *poly* functions). Use the *plot* routine to examine the AR signal. How long is the initial transient Lt? The *axis* command should be used to examine the initial portion of y. Remove this portion from the AR signal, e.g. $y = y(Lt+1:10000)$. What is the nature of the signal in the time domain beyond the initial transient? Apply the *avpgm* routine to calculate the periodogram estimate of the input Gaussian noise signal x with window lengths 32, 64, 128 and 256 (N_{FFT} should be 512 in each case). Comment on the bias and variance of the results to the equations (7) and (8).

b) Utilise the *freqz* function to plot the magnitude squared function $\left|H(e^{j2pf})\right|^2$ over the frequency interval $f \in [0,1/2]$ which will correspond to the true PSD for y. Due to the even symmetry of the PSD it is only necessary to plot the interval between zero frequency, namely, dc and the Nyquist frequency, i.e. the normalised frequency $f \in [0,1/2]$. This can be achieved with [h w]=freqz(b,a,256), fz=w/(2*pi) and plot(fz,abs(h).^2). Examine the separation of the two peaks. Calculate the periodogram estimate, *pgest*, with a window length of 512 points. Plot the true and estimated PSD

together, comment on the result. Examine the effect of reducing the window length, e.g. use 64 and 128, and observe whether it is possible to resolve the two sinusoids.

c) Experiment with averaging the periodogram. Use K=4, 8, 16 and 32 with a subsequence length M of 256. Is variance a problem with estimating the narrowband process? Does more averaging improve the quality of the estimate?

d) Zoom in on the PSD estimate about the peaks with the *axis* command and comment on whether the estimated PSD has the correct shape around the peak. Is there a bias? How does this relate to the length of the window and the mainlobe bandwidth of the window?

e) Use the *load* command to apply the periodogram to some real speech data contained with the s1.mat file. Examine the s1 data and select an interval of the speech, 256 points, over which the speech is approximately statistically stationary. Estimate the PSD with the periodogram and comment on the shape of the spectrum. Estimate the PSD over the next 256 points and observe the change in the PSD estimate. Do this for a number of intervals and comment upon the nature of the speech spectrum.

f) Advanced challenge, modify the *avpgm* routine to accept a different window, e.g. a Hamming window [Kay, 1988], and repeat exercise d), what is the effect?

Project 2: Modern Spectrum Estimation

A fundamental limitation of classical spectrum estimation techniques based on Fourier analysis is that to resolve two sinusoids separated by Δf , where f is normalised temporal or spatial frequency, a data length of at least $1/(f\Delta)$ is required. In many applications, for example in processing the measurements from a multi sensor, finite aperture, array, only very short data lengths are available. To overcome this problem, over the last 20 years modern spectrum estimation techniques have been proposed. One of these methods is based upon pole-zero modelling of the signal measurement.

If the signal measurement is assumed to be from a Gaussian process, the Maximum Entropy Method (MEM) yields a spectrum estimate which takes the form of an all-pole spectrum

$$\hat{P}(f) = \frac{G^2}{\left|A\left(e^{-j2pf}\right)\right|^2} = \frac{G^2}{\left|1 + a_1 e^{-j2p(2f)} + a_2 e^{-j2p(2f)} + \ldots + a_{NP} e^{-j2p(N_P f)}\right|^2} \quad (11)$$

where N_P is the model order. This equation contains the key difference between classical and modern spectrum estimation techniques. The measurement is assumed to have a power spectrum of the form of equation (11) and, therefore, spectrum estimation becomes estimation of the parameters $\{G, a_1, a_2, \ldots, a_{N_P}, N_P\}$. Estimation of the parameters $\{G, a_1, a_2, \ldots, a_{N_P}\}$ is based upon the solution of a set of $(N_P+1)\times(N_P+1)$ linear equations, name the *normal equations* [Kay, 1987]

$$\begin{bmatrix} r(0) & r(1) & \cdots & r(N_P) \\ r(1) & r(0) & \cdots & r(N_P-1) \\ \vdots & \vdots & \vdots & \vdots \\ r(N_P) & r(N_P-1) & \ldots & r(0) \end{bmatrix} \begin{bmatrix} 1 \\ a_1 \\ \vdots \\ a_{N_P} \end{bmatrix} = \begin{bmatrix} G^2 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \quad (12)$$

with the assumption that N_P, the model order, is known. Order selection is a fundamental issue in modern spectrum estimation and more details can be found in [Kay, 1987]. In this project, however, the model order will be assumed known. The $r(m)$, $m=0,\ldots,N_P$, parameters in equation (12) are the autocorrelation function values for the process $\{x(k)\}$. Again, as in equation (3), to calculate these exactly requires knowledge of the entire process. The issue is to make the best estimates of the autocorrelation function values given the available data $x(k)$, $k=0,1,\ldots,N-1$. In the MEM method, the autocorrelation functions are estimated from

$$\hat{r}(l) = \frac{1}{N}\sum_{m=0}^{N-1} x(m)x(m+l) \quad l = 0,1,\ldots,N_P \quad (13)$$

with the assumption that $x(m+l)=0$ whenever $m+l$ lies outside the interval $0,\ldots,N-1$. The model parameters, $\{G,a_1,a_2,\ldots,a_{N_P}\}$ are then found by solving equation (12). This method, for obvious reasons, is also known as the *autocorrelation* method [Kay, 1987]. A second method, named the *covariance* method, estimates the $r(m)$, $m=0,\ldots,N_P$ parameters with no assumption for the data outside of the interval $0,..,N-1$, that is the (i,j)th entry in the matrix of the left-hand side of equation (12) becomes

$$\hat{r}(i,j) = \frac{1}{N}\sum_{m=N_P}^{N-1} x(m-i)x(m-j) \quad i,j = 0,1,\ldots,N_P \quad (14)$$

In both methods, inspection of equation (12) shows that the autocorrelation function satisfies

$$r(l) = -\sum_{m=1}^{N_P} a_m r(l+m) \quad (15)$$

for $l = 0,\ldots,N_P-1$. The essence of these techniques is that implicitly equation (15) is satisfied for all l, that is, the autocorrelation function is extrapolated by the underlying AutoRegressive (AR) model. This extrapolation means that although only a finite measurement is available, an infinite extent autocorrelation model is assumed which corresponds to a maximally random Gaussian process $\{x'(k)\}$ constrained to have $x'(k)=x(k)$ for $k=0,\ldots,N-1$; hence the name, Maximum Entropy Method.

Exercise 2. Write two MATLAB functions called *autocorr* and *covar* which accept as input the $N \times 1$ measurement vector $x(k)$ and the scalar model order N_P; and output the model parameters $\{G, a_1, a_2, ..., a_{N_P}\}$ as a scalar gain, G, parameter and $N_P \times 1$ a-vector. Use equations (12), (13) and (14). See discussions and FORTRAN code in [Burrus 94, and Appendices of Chapter 7, Kay 88].

a) Apply the *autocorr* and *covar* functions to the AR signal, *y*, generated in Exercise la. Select appropriate values for N_P and the minimum number of points, N, to obtain two distinct peaks in the estimated PSD. This is calculated with equation (11) (This is simple to achieve with the *freqz* function). Explain in detail why the necessary value for N is so much different for the two methods. Once values for N_P and N have been selected apply the best method to ten independent portions of the signal *y*. Produce a plot of the ten estimated PSDs and the true PSD for the best method. Comment on the variability of the estimated PSD. Employ the *roots* and *zplane* functions to show the variation in the estimated pole positions, include on the plots the ideal positions [Burrus 94]. What is the effect on the position of the roots of choosing N_P to equal 6 and 8?

b) Examine the estimated PSD around the position of the peaks similar to Exercise 1d). Is the estimated PSD biased? What is the effect of increasing N?

c) Repeat Exercise 1e) for *autocorr* and *covar*.

References

Kay, S.M., "Modern Spectral Estimation: Theory and Applications", Prentice-Hall, 1988. *Comprehensive coverage of the area.*

Marple, S.L., "Digital Spectral Analysis with Applications", Prentice-Hall, 1987. *Companion for Kay's book, second-edition is about to appear, strong on applications.*

Burrus, C.S., et. al., "Computer-Based Exercises for Signal Processing Using Matlab", Prentice-Hall, 1994. *Further exercises on spectral estimation can be found in Chapter* 6.

Bendat, J.S., and A.G. Piersol, "Random Data: Analysis and Measurement Procedures", Second Edition, John Wiley, 1986. *Useful, practically-orientated discussion of* real signal analysis.

Jonathon Chambers, July 1994, November 1999, ver. 2