E4.13
AS2
SO15
ISE4.31

IMPERIAL COLLEGE LONDON

DEPARTMENT OF ELECTRICAL AND ELECTRONIC ENGINEERING
EXAMINATIONS 2008

MSc and EEE/ISE PART IV: MEng and ACGI

Corrected copy

SPECTRAL ESTIMATION AND ADAPTIVE SIGNAL PROCESSING

Tuesday, 6 May 10:00 am

Time allowed: 3:00 hours

There are FIVE questions on this paper.

Answer ONE of questions 1,2 and TWO of questions 3,4,5.

All questions carry equal marks

Any special instructions for invigilators and information for candidates are on page 1.

Examiners responsible First Marker(s) : D.P. Mandic

Second Marker(s) : M.K. Gurcan

1) Consider the problem of periodogram based spectrum estimation.

a) Show how the periodogram can be computed from the discrete Fourier transform of a signal. [1]

i) Compute the bias of the periodogram and show whether the periodogram is biased. [2]

ii) Comment on the performance of the periodogram in terms of the frequency resolution and variance. What is the variance of the periodogram proportional to? [3]

b) Suppose that we compute the periodogram $\hat{P}_{per}(e^{\jmath\omega})$ using N samples of $x(n)$ (Bartlett method). Find and prepare a carefully labelled sketch of the expected value of this spectrum estimate. Is this estimate biased? Is it consistent? What is the role of averaging of periodograms? [6]

c) Given $N = 10,000$ samples of a process $x(n)$, you are asked to compute the periodogram. However, with only a finite amount of memory resources, you are unable to compute a DFT any longer than 1024. Using these $10,000$ samples, describe how you would be able to compute a periodogram that has a resolution of

$$\Delta\omega = 0.89\frac{2\pi}{10000}$$

Hint: the resolution of the periodogram is $\Delta\omega = 0.89\frac{2\pi}{N}$. [8]

2) Consider the problem of parametric spectrum estimation.

a) Explain the limitations of classical periodogram based methods and the need for parametric spectrum estimation techniques. [3]

i) Define the autoregressive (AR) spectrum estimation, and state the expression of a general AR power spectrum. State the expression for a general autoregressive moving average (ARMA) power spectrum. Explain how the positions of poles and zeros in the transfer function affect the shape of the power spectrum. [5]

ii) Figure 2.1 shows the power spectrum of a random process. From the shape of the power spectrum, explain whether it is best modelled by AR or MA spectrum estimation. What is the order of the model which best describes the power spectrum of this process? Explain the relationship between the positions of the poles/zeros within the model of this power spectrum and the frequencies and magnitudes of the peaks in the spectrum. [6]

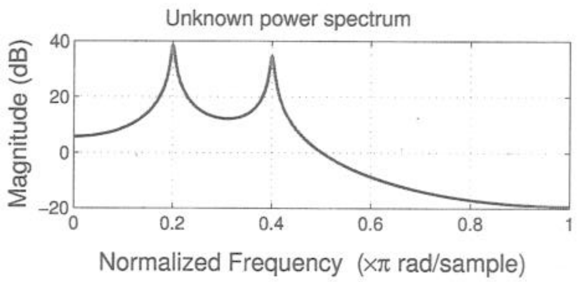


Figure 2.1: Power spectrum of an unknown process

b) In the MUSIC algorithm, finding the peaks of the frequency estimation function (for p frequencies of interest)

$$\hat{P}_{MU}(e^{\jmath\omega}) = \frac{1}{\sum_{i=p+1}^{M} \left|\mathbf{e}^H \mathbf{v}_i\right|^2}$$

is equivalent to finding the minima of the denominator, where $\mathbf{v}_i, i = 1, \ldots, M$ are the eigenvectors of the correlation matrix, $(\cdot)^H$ is the Hermitian transpose operator, and $\mathbf{e} = \left[1, e^{\jmath\omega}, \ldots, e^{\jmath(M-1)\omega}\right]^T$. Show that finding the minima of the denominator is equivalent to finding the maxima of

$$\sum_{i=1}^{p} \left|\mathbf{e}^H \mathbf{v}_i\right|^2$$

Hint: Use the fact that $\mathbf{I} = \sum_{i=1}^{M} \mathbf{v}_i \mathbf{v}_i^H$ and that $\mathbf{e}^H \mathbf{e} = M$. [6]

3) Consider a nonlinear adaptive finite impulse response (FIR) filter (dynamical perceptron), shown in Figure 3.1.

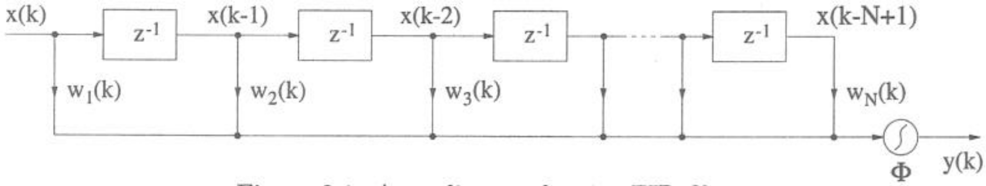


Figure 3.1: A nonlinear adaptive FIR filter.

The nonlinear function Φ is of saturation type.

a) Derive a least mean square (LMS)–type learning algorithm for such a filter. The cost function is $E = \frac{1}{2}e^2(k)$, and the nonlinear function at the output of the filter is $\Phi(net(k)) = tanh(net(k))$, where $net(k) = \mathbf{x}^T(k)\mathbf{w}(k)$. Comment on the operation of such a filter. What is the difference in the operation between the quasi–linear middle region and the operation nearer the tails of the nonlinearity? [6]

b) To suit the online adaptive mode of operation and an unknown range of inputs, the so called dynamical range transformation is often performed. This way the range of the input signals is reduced so as to suit the range of the nonlinear activation function Φ. One technique for dynamical range transformation is based on differentiation, whereby the transformed input $u(k)$ is given by $u(k) = x(k) - x(k-1)$.
Explain how this transformation provides dynamical range reduction. Suggest one more such technique. [4]

c) To suit the unknown range of the input we may perform adaptation of the amplitude $\lambda(k)$ of the nonlinear function Φ, based on cost function $E = \frac{1}{2}e^2(k)$, that is

$$\begin{aligned} \Phi(net(k)) &= \lambda(k)\bar{\Phi}(net(k)) \\ \lambda(k+1) &= \lambda(k) - \rho\nabla_\rho E(k) \end{aligned}$$

where $\bar{\Phi}$ denotes the nonlinear function with a unit amplitude and ρ is a learning rate. Derive the update for the amplitude of the nonlinearity $\lambda(k)$. [6]

d) The so called bilinear model is given by

$$y(k) = \sum_{j=1}^{N-1} c_j y(k-j) + \sum_{i=0}^{N-1}\sum_{j=1}^{N-1} b_{i,j} y(k-j)x(k-i) + \sum_{i=0}^{N-1} a_i x(k-i)$$

Can this model be approximated with the nonlinear filter from Figure 3.1? [4]

4) Consider the problem of finite impulse response (FIR) adaptive filtering.

a) Explain the notions of parametric, nonparametric and semi-parametric modelling. [3]

b) A digital filter can operate in the filtering, smoothing and prediction mode. Explain these modes of operation. [3]

i) What operation is performed by the filter given by

$$y(k) = 0.2x(k-2) + 0.8x(k-1) + 0.8x(k+1) + 0.2x(k+2)$$

[2]

ii) What operation is performed by the filter given by

$$y(k) = x(k) + 0.9x(k-1) + 0.81x(k-2) + 0.729x(k-3)$$

[2]

c) The recursive least squares (RLS) algorithm is employed for the training of an FIR filter.

i) Write down and explain the cost function for the RLS algorithm. [2]

ii) State the general recursive least squares filtering problem. [2]

iii) Explain the difference between the least mean square (LMS) and RLS algorithms. Compare the respective error surfaces. [2]

iv) Show the way to calculate recursively the update of the correlation matrix, that is $\mathbf{R}(k+1) = f\big(\mathbf{R}(k), \mathbf{x}(k)\big)$. [2]

v) A forgetting factor λ is often used within the cost function of the RLS algorithm. Explain the role of the forgetting factor. What are typical values of the forgetting factor? [2]

5) Consider an adaptive finite impulse response (FIR) filter.

a) Show in block diagram how an adaptive noise cancellation configuration can be used to suppress the effects of noise. Discuss and compare with the adaptive line enhancement configuration. [4]

b) The coefficient update equation for the normalised least mean square (LMS) algorithm is given by

$$\mathbf{w}(k+1) = \mathbf{w}(k) + \frac{\beta e(k)\mathbf{x}(k)}{\varepsilon + \parallel \mathbf{x}(k) \parallel_2^2}$$

where $\mathbf{w}(k)$ is the weight (coefficient) vector at time instant k, $\mathbf{x}(k)$ is the tap input vector, β is the learning rate, $e(k) = d(k) - \mathbf{x}^T(k)\mathbf{w}(k)$ is the instantaneous output error, $d(k)$ is the desired response, and ε is a regularisation factor.
Explain in your own words the merits of this algorithm and the roles of the terms in the denominator of the normalised LMS (NLMS) update. [4]

c) Verify that the update equation

$$\mathbf{w}(k+1) = \mathbf{w}(k) + \mu e_p(k)\mathbf{x}(k)$$

based upon the a posteriori error

$$e_p(k) = d(k) - \mathbf{x}^T(k)\mathbf{w}(k+1)$$

is equivalent to the NLMS algorithm. Comment upon the relationship between ε, β, and μ. [6]
Hint: Show that the weight update $\Delta\mathbf{w}(k) = \mathbf{w}(k+1) - \mathbf{w}(k)$ based on the a posteriori error $e_p(k)$ can be expressed in terms of the a priori error $e(k)$ as

$$\Delta\mathbf{w}(k) = \mu e(k)\left[1 - \mu\frac{\parallel \mathbf{x}(k) \parallel_2^2}{1 + \mu \parallel \mathbf{x}(k) \parallel_2^2}\right]\mathbf{x}(k) = \frac{1}{\frac{1}{\mu} + \parallel \mathbf{x}(k) \parallel_2^2} e(k)\mathbf{x}(k)$$

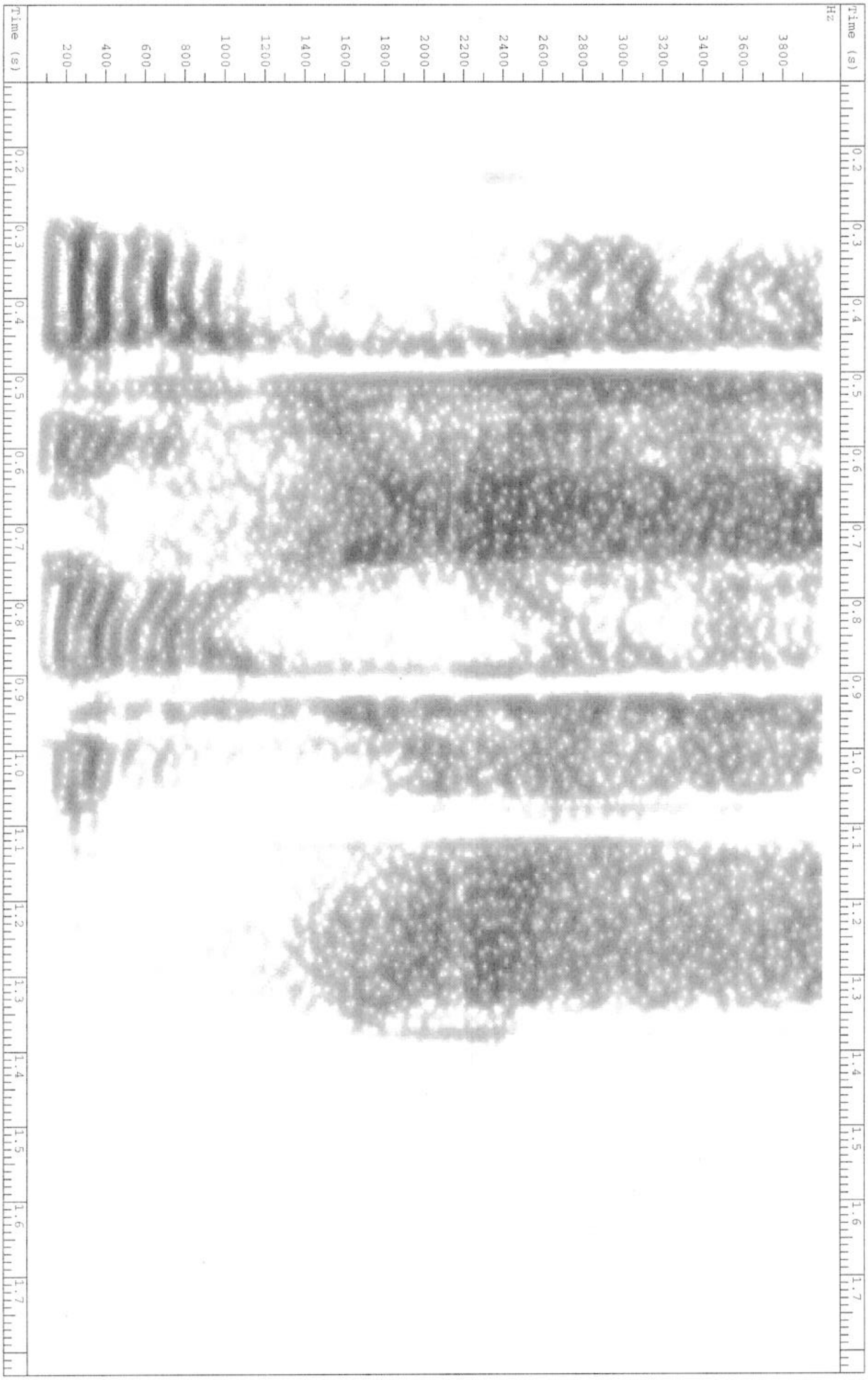
d) The affine projection (AP) algorithm is a generalisation of NLMS based upon forcing the following L a posteriori errors to be zero

$$\begin{bmatrix} e_p(k) \\ e_p(k-1) \\ \vdots \\ e_p(k-L) \end{bmatrix} = \begin{bmatrix} d(k) \\ d(k-1) \\ \vdots \\ d(k-L) \end{bmatrix} - \begin{bmatrix} \mathbf{x}^T(k)\mathbf{w}(k+1) \\ \mathbf{x}^T(k-1)\mathbf{w}(k+1) \\ \vdots \\ \mathbf{x}^T(k-L)\mathbf{w}(k+1) \end{bmatrix} = \mathbf{0}$$

i) Explain in your own words the operation of the AP algorithm. [2]

ii) Derive the update for the AP algorithm and compare with the NLMS update. [4]

Figure 1


Time (s)
Hz
200
400
600
800
1000
1200
1400
1600
1800
2000
2200
2400
2600
2800
3000
3200
3400
3600
3800
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1.0
1.1
1.2
1.3
1.4
1.5
1.6
1.7