ENGINEERING TRIPOS PART IIA

Thurs 8 May 2008 2.30 to 4

Module 3F3

SIGNAL AND PATTERN PROCESSING

Answer not more than ***three*** *questions.*

All questions carry the same number of marks.

The ***approximate*** *number of marks allocated to each part of a question is indicated in the right margin.*

There are no attachments.

STATIONERY REQUIREMENTS
Single-sided script paper

SPECIAL REQUIREMENTS
Engineering Data Book
CUED approved calculator allowed

You may not start to read the questions printed on the subsequent pages of this question paper until instructed that you may do so by the Invigilator

1 The Discrete Fourier Transform (DFT) for a data sequence $\{x_n\}$ of length N, where N is here assumed to be a power of 2, is defined as

$$X_p = \sum_{n=0}^{N-1} x_n e^{-j\frac{2\pi}{N}np}, \quad p = 0, 1, ..., N-1$$

(a) Show that the DFT values X_p and $X_{p+N/2}$ may be expressed as

$$X_p = A_p + W^p B_p, \quad \text{and} \quad X_{p+N/2} = A_p - W^p B_p$$

where A_p is a series involving only the even numbered data points $(x_0, x_2, ...)$ and B_p is a series involving only the odd numbered data points $(x_1, x_3, ...)$ and W is a constant which should be carefully defined. [30%]

Derive the total computational complexity for evaluating X_p and $X_{p+N/2}$ for $p = 0, 1, ..., N/2-1$ and compare this with a full evaluation of the DFT (assume that complex exponentials are pre-computed and stored). [20%]

(b) Define a vector of data points as $\mathbf{x} = [x_0, x_1, ..., x_{N-1}]^T$ and the corresponding vector of frequency components as $\mathbf{X} = [X_0, X_1, ..., X_{N-1}]^T$.

Show that the DFT may be expressed in matrix-vector form as

$$\mathbf{X} = \mathbf{M}\mathbf{x}$$

where $\mathbf{M}$ is an $(N \times N)$ matrix whose nth column is defined as

$$\mathbf{m}_n = [W^0, W^n, W^{2n}, W^{3n}, \ldots, W^{(N-1)n}]^T$$

and W is a constant which should be defined. [20%]

(c) Hence derive, by direct application of matrix algebra, the inverse DFT in the form:

$$\mathbf{x} = \mathbf{H}\mathbf{X}$$

where the elements of $\mathbf{H}$ should be carefully defined. [30%]

[Hint: consider the products $\mathbf{m}_i^H \mathbf{m}_j$ in the case where $i = j$ and $i \neq j$. Note that $\mathbf{m}_i^H$ is the complex conjugate of the transpose of vector $\mathbf{m}_i$]

2 (a) An autoregressive (AR) model of order P is defined by the following equation:

$$x_n = -\sum_{i=1}^{P} a_i x_{n-i} + e_n$$

where $\{e_n\}$ is a zero mean white noise process having standard deviation σ.

Show that transfer function $H(z)$ from e to x can be expressed as

$$H(z) = \frac{1}{A(z)}$$

where $A(z)$ should be derived. Hence or otherwise write an expression for the power spectrum of this AR process. [30%]

(b) A second autoregressive process $\{y_n\}$ is defined in the same way, driven by a second zero-mean white noise process $\{f_n\}$, which is uncorrelated with $\{e_n\}$, and also having standard deviation σ:

$$y_n = -\sum_{i=1}^{P} b_i y_{n-i} + f_n$$

The two processes are now added, i.e. we take $w_n = x_n + y_n$. Show that, in the z-transform domain, $W(z)$ may be expressed as:

$$W(z) = \frac{E(z)B(z) + F(z)A(z)}{A(z)B(z)}$$

where $E(z)$ is the z-transform of e_n, etc. [10%]

Now, take both AR processes to be first order, i.e. $P = 1$. Consider the numerator term $U(z) = E(z)B(z) + F(z)A(z)$, which corresponds itself to a sum of random processes.

Show that the process corresponding to $E(z)B(z)$ is uncorrelated with that corresponding to $F(z)A(z)$. [20%]

Hence show that the power spectrum of the process $\{u_n\}$ is:

$$S_U(e^{j\Omega}) = \sigma^2 \left(2 + 2(a_1 + b_1)\cos\Omega + a_1^2 + b_1^2\right)$$

[20%]

(c) Hence or otherwise write down the power spectrum for $\{w_n\}$. Is $\{w_n\}$ itself an AR process? Justify your answer. [20%]

3 (a) Discuss the principles of the Wiener filter, including the error function to be minimised, the assumptions made about the random processes involved, the information required to determine the filter, and application scenarios where the filter can be employed. [25%]

(b) In a stock market trading system it is desired to predict the price of a share six months into the future. The raw prices are first adjusted to remove any trends or offsets of the mean away from zero, leading to a random process of adjusted prices $\{x_t\}$. It is then proposed that the price x_t and the price one month ago x_{t-1} might be used to make the prediction via a linear estimate of the form

$$\hat{x}_{t+6} = ax_t + bx_{t-1}$$

where a and b are constants to be determined. You may assume that the adjusted price data are approximately wide-sense stationary over the period of interest. Under an expected mean squared error criterion for the prediction, show that a and b must satisfy the following conditions:

$$ar_{xx}[0] + br_{xx}[1] = r_{xx}[6], \quad br_{xx}[0] + ar_{xx}[1] = r_{xx}[7]$$

where $r_{xx}[k]$ is the autocorrelation function of the data. [40%]

(c) A financial analyst believes that the following formula applies for this data:

$$E[x_t x_{t-k}] = \delta[k] + 0.2/(|k| + 0.2)$$

where $\delta[k]$ is the unit pulse function.

Compute the optimal constants a and b, assuming that the above formula is correct. Compute also the expected squared error for the prediction when these optimal values are used. How much of an improvement is this compared with the simple predictor given by $\hat{x}_{t+6} = x_t$? [35%]

4 Consider a data set of pairs of observations $\mathscr{D} = \{(x_n, y_n)\}$ where $n = 1, \ldots, N$ and N is the total number of data points. Assume we wish to learn a regression model

$$y_n = a x_n + \varepsilon_n$$

where ε_n is independent zero-mean Gaussian noise with variance σ^2.

(a) Write down the log likelihood $\log p(y_1, \ldots, y_N | x_1, \ldots, x_N, a, \sigma^2)$ in terms of $y_1, \ldots, y_N, x_1, \ldots, x_N, a, \sigma^2$. [30%]

(b) Assume the following data set of $N = 4$ pairs of points

$$\mathscr{D} = \{(0,1), (1,2), (2,0), (3,4)\}$$

Solve for the maximum likelihood estimates of a and σ^2. [40%]

(c) Assume the same data set, but instead a regression model that predicts x given y:

$$x_n = b y_n + \varepsilon_n$$

Is the maximum likehood estimate of b equal to $\frac{1}{a}$? Explain why or why not, giving a derivation if necessary. [30%]

END OF PAPER