EGT3
ENGINEERING TRIPOS PART IIA

Thursday 23rd April 2015 9.30-11

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 TO BE SUPPLIED FOR THIS EXAM
Engineering Data Book
CUED approved calculator allowed

10 minutes reading time is allowed for this paper.

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

1 (a) The Discrete-time Fourier Transform (DTFT) for a dataset $\{x_n\}$ is denoted $X(e^{j\Omega})$, where Ω is the normalised frequency such that $\Omega = 2\pi$ corresponds to the sampling frequency of the data.

A length-N Discrete Fourier Transform (DFT) is now calculated for the data,

$$X_p = \sum_{n=0}^{N-1} x_n e^{-jnp2\pi/N}, \; p = 0, 1, \ldots, N-1.$$

Show that the DFT coefficients are related to the DTFT through the expression:

$$X_p = \frac{1}{2\pi} \int_0^{2\pi} W_r(e^{j\Theta}) X(e^{j(2p\pi/N - \Theta)}) \, d\Theta$$

where $W_r(e^{j\Theta})$ is a function which should be carefully calculated and defined. [40%]

(b) Sketch $|W_r(e^{j\Theta})|$ for a small range of frequencies Θ, showing the central lobe and several sidelobes. Calculate and clearly mark on the sketch the height of the central lobe and the positions of the first few sidelobes and nulls (i.e. frequencies where $|W_r(e^{j\Theta})| = 0$). [20%]

(c) Suppose that $\{x_n\}$ is composed purely of complex frequency components of the form $a_i \exp(jn\Omega_i)$, where Ω_i is the fixed frequency of the ith component with magnitude a_i. Explain the effect of using the DFT to estimate the spectrum compared with the full DTFT of such a signal. Your explanation should include the effect of N, spectral smearing, leakage and the use of window functions. [20%]

(d) Consider a signal which contains frequencies which lie exactly on one of the DFT bins, i.e. $\Omega_i = k2\pi/N$ where k is an integer. How, if at all, would your comments in the previous section be modified? If it is known in advance that the frequencies satisfy this constraint, what would be the minimum gap between two frequency components in order to detect their presence individually (i.e. 'resolve' them). Consider only the unwindowed case. [20%]

2 A digital filter equaliser is to be designed based on an all-pole analogue prototype filter of the form:

$$H(s) = \prod_{i=1}^{P} \frac{1}{s^2 - 2r_i\cos(\theta_i)s + r_i^2}$$

where $\pi/2 < \theta_i \leq \pi$, $0 < r_i < \infty$, all θ_i are assumed distinct from one another and all r_i are assumed distinct from one another.

The change of variables $s = \frac{1-z^{-1}}{1+z^{-1}}$ is used to design a digital filter $\hat{H}(z)$ based on this prototype.

(a) Show, by keeping $\hat{H}(z)$ factorised in a manner corresponding to $H(s)$, that the new digital filter has poles at

$$\hat{p}_i = \frac{1+p_i}{1-p_i} \text{ and } \hat{p}_i^* = \frac{1+p_i^*}{1-p_i^*}$$

where $p_i = r_i \exp(j\theta_i)$, and show that the digital filter also has $2P$ zeros, all co-located at $z = -1$. [30%]

(b) By considering the poles and/or zeros of the new digital filter as necessary, show whether $\hat{H}(z)$ is guaranteed to be stable or not. [20%]

(c) Sketch a cascade realisation of Direct form II biquadratic sections, using as few multiplications as possible, which would be suitable for implementation of $\hat{H}(z)$. [20%]

(d) A zero-mean white noise sequence with variance $\sigma_w^2 = 1$ is input to a filter $\hat{H}(z)$ as above with $P = 1$, $r_1 = 0.2$ and $\theta_1 = 5\pi/8$. Determine the power spectrum of the output from the filter, sketching the result between normalised frequency $\Omega = 0$ and π, paying particular attention to the DC gain, the gain at $\Omega = \pi$, and any maxima/minima of the gain. [30%]

3 (a) A discrete random process is generated as a sequence of independent time points, assigned a value of +1 or -1 with probabilities p_1 and $p_{-1} = 1 - p_1$, respectively. A possible sequence of bits generated from the process is, for example:

$$\{x_n\} = \{..., +1, +1, -1, +1, -1, +1, ...\}.$$

Determine the mean and autocorrelation function for the process. Hence determine whether the process is wide-sense stationary. [30%]

(b) A noisy communications channel is applied to the random bit stream $\{x_n\}$ such that the received sampled data can be modelled as

$$y_n = 0.9x_n + 0.5x_{n-1} + v_n$$

where $\{v_n\}$ is zero mean white noise having variance $\sigma_v^2 = 0.1$, and which is independent of $\{x_n\}$. It is known that $p_1 = p_{-1} = 0.5$.
Determine the autocorrelation function of the received process $\{y_n\}$ and also the cross-correlation function between $\{x_n\}$ and $\{y_n\}$. [30%]

(c) Hence determine the coefficients of an optimal filter for estimation of $\{x_n\}$ from $\{y_n\}$ according to the formula

$$\hat{x}_n = h_0 y_n + h_1 y_{n-1},$$

where h_0 and h_1 are to be determined according to the Wiener criterion. [30%]

(d) Determine the mean-squared error corresponding to this Wiener filter and comment on the result. [10%]

4 Consider a binary classification problem where the data $\mathscr{D}$ consists of N data points, $\mathscr{D} = \{(x_1, y_1), \ldots, (x_N, y_N)\}$, x_n is a real scalar and $y_n \in \{0, 1\}$, and the goal is to predict class labels y for new x.

Assume a very simple logistic classification model in which the class labels were produced independently and identically from the following model:

$$P(y_n = 1 | x_n, a, b) = \sigma(ax_n + b)$$

where σ is the logistic function, $\sigma(z) = \frac{1}{1+\exp(-z)}$, and a and b are the parameters of the classifier.

(a) Write down the likelihood of a and b for the data $\mathscr{D}$ and describe an algorithm to optimise this likelihood as a function of a and b. [40%]

(b) Consider a data set consisting of only two data points, $\mathscr{D} = \{(-2, 0), (3, 1)\}$. For this data set, describe the set of parameters which classify both data points correctly with probability greater than 0.5. Furthermore, what is the maximum achievable likelihood value? Describe the set of parameters which achieve this maximum. [40%]

(c) Explain how Bayesian learning of the parameters might give more reasonable inferences about a and b from the data set in part (b) than maximum likelihood (ML) and how the Bayesian predictions about future labels differ from the ML predictions. [20%]

END OF PAPER

THIS PAGE IS BLANK