EGT3
ENGINEERING TRIPOS PART IIB

Tuesday 1 May 2018 2.00 to 3.40

Module 4F7

STATISTICAL SIGNAL ANALYSIS

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

All questions carry the same number of marks.

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

*Write your candidate number **not** your name on the cover sheet.*

STATIONERY REQUIREMENTS
Single-sided script paper

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

10 minutes reading time is allowed for this paper at the start of the exam.

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

1 (a) Consider the following gradient descent algorithm

$$\mathbf{h}(n) = \mathbf{h}(n-1) + \mu\big(\mathbf{p} - \mathbf{R}\mathbf{h}(n-1)\big), \quad n > 0,$$

where $\mathbf{R}$ is a positive definite square matrix, $\mathbf{h}(0)$ and $\mathbf{p}$ are vectors, and μ is a scalar.

(i) Assuming $\mathbf{h}(n)$ converges to a limit $\mathbf{h}^*$, find $\mathbf{h}^*$. [10%]

(ii) Deduce conditions on μ that would ensure $\mathbf{h}(n)$ converges regardless of the initial vector $\mathbf{h}(0)$ chosen. Carefully explain how the eigenvalues of the matrix $\mathbf{R}$ influence the convergence speed. (Hint: you may use the fact that $\mathbf{A}^k \to 0$ when all the eigenvalues of the matrix $\mathbf{A}$ have modulus strictly less than one.) [20%]

(b) The gradient descent algorithm is modified to

$$\mathbf{h}(n) = \mathbf{h}(n-1) + \mu\mathbf{S}\big(\mathbf{p} - \mathbf{R}\mathbf{h}(n-1)\big), \quad n > 0,$$

where the step-size is now $\mu\mathbf{S}$ for some matrix $\mathbf{S}$ and scalar μ. Find the optimal matrix $\mathbf{S}$ that reduces the sensitivity of the convergence of $\mathbf{h}(n) \to \mathbf{h}^*$ on the eigenvalue spread of the matrix $\mathbf{R}$. [15%]

(c) Adaptive filters are commonly used for prediction. The aim is to form a linear predictor of the real valued signal $\{x(n)\}$ using only noisy measurements of it:

$$u(n) = x(n) + v(n)$$

where $v(n)$ is zero mean white noise with variance σ_v^2 and $E\{x(n)v(m)\} = 0$ for all integers n, m.

(i) Give the M-tap Least Mean Square (LMS) algorithm, with all quantities carefully defined, for designing a linear predictor for $\{x(n)\}$. (You may use the following standard notation: write the LMS algorithm as $\mathbf{h}(n) = \mathbf{h}(n-1) + \cdots$ where $\mathbf{h}(n)$ is the vector of filter coefficients being updated by the LMS algorithm.) [30%]

(ii) State the range of values of step-size for which the LMS converges in mean. Express this limit point,

$$\lim_{n\to\infty} E\{\mathbf{h}(n)\},$$

in terms of the autocorrelation matrix of $\{x(n)\}$. [25%]

2 We have repeated observations of a random variable X through

$$Y_n = X + V_n \quad \text{for } n = 1, 2, \ldots$$

where $\{V_n\}$ is an independent and identically distributed zero-mean scalar noise sequence, independent of X, with variance $E\left(V_n^2\right) = \sigma_v^2$. Also, $E(X) = 0$ and $E\left(X^2\right) = \sigma_0^2$.

(a) Let $\widehat{X}_{n-1}$ be an estimate of X using $\{Y_1, \ldots, Y_{n-1}\}$. Upon receiving Y_n, the estimate of X is updated to

$$\widehat{X}_n = G_n\left(Y_n - \widehat{X}_{n-1}\right) + \widehat{X}_{n-1}.$$

(i) Find $\widehat{X}_0$ which results in the estimate being unbiased for all n. [10%]

(ii) Find the value of the gain G_n that minimises the mean square error $E\left\{(\widehat{X}_n - X)^2\right\}$. Carefully detail your derivation. [40%]

(b) The Kalman filter for estimating X can be expressed as

$$\widehat{X}_n = \widehat{X}_{n-1} + \frac{\sigma_0^2}{n\sigma_0^2 + \sigma_v^2}\left(Y_n - \widehat{X}_{n-1}\right).$$

(i) Using this equation find an expression for $\widehat{X}_n$ in terms of $\widehat{X}_0$ and $Y_1, \ldots, Y_n$. [20%]

(ii) Compute the variance of the Kalman filter estimate and that of the sample mean estimate,

$$\frac{1}{n}\sum_{i=1}^{n} Y_i,$$

and thus conclude which estimator is better. [30%]

3 Consider the following Markov random process

$$X_{k+1} = X_k + W_{k+1}, \quad k = 0, 1, \ldots$$

where $X_0 = 0$ and $W_1, W_2, \ldots$ is a sequence of independent and identically distributed random variables with common probability mass function (pmf) $f(i)$ such that $\sum_{i=-\infty}^{\infty} f(i) = 1$. That is, W_k (for $k = 1, 2, \ldots$) are integer valued random variables.

(a) Find $P_{i,j} = \Pr(X_{k+1} = j \mid X_k = i)$. [10%]

(b) Let

$$Y_k = X_k + dV_k, \quad k = 1, 2, \ldots$$

where $V_1, V_2, \ldots$ are independent Gaussian random variables with zero mean and unit variance and d is a positive constant.

(i) Find $p(x_1, y_1, \ldots, x_n, y_n)$ which is the joint probability mass and density function of $X_1, Y_1, \ldots, X_n, Y_n$ evaluated at $x_1, y_1, \ldots, x_n, y_n$. [10%]

(ii) Given the conditional pmf $p(x_n \mid y_1, \ldots, y_n)$, find $p(x_{n+1} \mid y_1, \ldots, y_{n+1})$. [20%]

(c) Let $X^1_{1:n}, X^2_{1:n}, \ldots, X^N_{1:n}$ be N independent samples from $p(x_1, \ldots, x_n)$, the joint probability mass function of $X_1, \ldots, X_n$.

(i) Give the values of the weights $w^1_n, w^2_n, \ldots, w^N_n$ so that

$$\frac{1}{N} \sum_{i=1}^{N} w^i_n$$

is an unbiased estimate of $p(y_1, \ldots, y_n)$. [10%]

(ii) Hence construct an importance sampling estimate of

$$\sum_{x_1=-\infty}^{\infty} \cdots \sum_{x_n=-\infty}^{\infty} H(x_1, \ldots, x_n) p(x_1, \ldots, x_n \mid y_1, \ldots, y_n)$$

where $H(x_1, \ldots, x_n)$ is a real-valued function of interest. [10%]

(iii) Using sequential importance sampling, extend this to an importance sampling estimate of

$$\sum_{x_{n+1}=-\infty}^{\infty} h(x_{n+1}) p(x_{n+1} \mid y_1, \ldots, y_{n+1})$$

where $h(x_{n+1})$ is a real-valued function of interest. [20%]

(d) Let $J_1, \ldots, J_M$ be independent and identically distributed random variables with pmf

$$\Pr(J = j) = w_n^j / (\sum_{i=1}^{N}) w_n^i.$$

Show that the following *resampled* estimate

$$\frac{1}{M}\left(H(X_{1:n}^{J_1}) + \ldots + H(X_{1:n}^{J_M})\right)$$

has the same expected value as the importance sampling estimate constructed in part (c)(ii). [20%]

4 Let X be an integer valued random variable with probability mass function (pmf) $p(i) = c/i^2$ for $i > 0$ and $p(i) = 0$ otherwise. Constant c is chosen to ensure that $\sum_{i=1}^{\infty} p(i) = 1$. We have repeated observations of the random variable X through

$$Y_n = X + V_n \quad \text{for } n = 1, 2, \ldots$$

where $V_1, V_2, \ldots$ are independent Gaussian random variables with zero mean and unit variance.

(a) Find $p_k(i) = p(X = i \,|\, Y_1 = y_1, \ldots, Y_k = y_k)$, which is the conditional pmf of X given observations $y_1, \ldots, y_k$. [10%]

(b) Let $X_1, \ldots, X_N$ be N independent samples from a pmf $q(i)$ where $q(i) = 0$ for $i < 0$.

(i) Using $X_1, \ldots, X_N$, give the importance sampling estimate of $\sum_{i=0}^{\infty} p_k(i)h(i)$, where h is some real-valued function of interest. [15%]

(ii) Give the importance sampling estimate of $p(y_1, \ldots, y_k)$ when c is unknown. [15%]

(c) Assume c is known.

(i) Give the importance sampling estimate of $p(y_1, \ldots, y_k)$ and show that it is unbiased. [25%]

(ii) Calculate the variance of the importance sampling estimate of $p(y_1, \ldots, y_k)$ and show that it can be written as

$$\frac{1}{N} p(y_1, \ldots, y_k)^2 \left(-1 + \sum_{i=0}^{\infty} q(i) \frac{p_k(i)}{q(i)} \frac{p_k(i)}{q(i)} \right).$$

[25%]

(iii) Find the optimal choice for $q(i)$. [10%]

END OF PAPER