Version SJG/2

EGT2
ENGINEERING TRIPOS PART IIA

Wednesday 3 May 2023 9.30 to 11.10

Module 3F3

STATISTICAL SIGNAL PROCESSING

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.

You may not remove any stationery from the Examination Room.

1 Consider the following Markov sequence $X_0, X_1, \ldots$ which describes the population count of prey, where $0 \leq X_k \leq M$,

$$P\left(X_{k+1} = j | X_k = i\right) = \begin{cases} i\theta_1 & \text{if } j = i+1 \text{ and } i < M \\ i\theta_2 & \text{if } j = i-1 \text{ and } i > 0 \\ ? & \text{if } j = i \\ 0 & \text{otherwise} \end{cases}$$

and M denotes the maximum population count. (The value in the $j = i$ row is intentionally left as '?'.)

(a) Find the transition probability matrix for this Markov chain. [20%]

(b) Find the range of values for θ_1 and θ_2 that ensures the transition probability matrix is valid. [10%]

(c) For $X_0 = M$, find the probability mass functions (pmfs) $P(X_1 \mid X_0 = M)$ and $P(X_2 \mid X_0 = M)$. [20%]

(d) For $X_0 = M$, find the values of θ_2 and θ_1 that maximise the probability of observing the following population count: $X_1 = M-1, X_2 = M-2, \ldots, X_{M-1} = 1$. [20%]

(e) Find $P(X_k > X_{k-1})$ in terms of the marginals $P(X_k = i)$. Hence, in a data set $X_1, X_2, \ldots, X_T$ of length T, find the expected number of transitions where the population increases. [30%]

2 Let $\{X_n\}$ be a wide sense stationary random process generated by the AR(2) model

$$X_n = 0.1X_{n-1} + 0.7X_{n-2} + W_n$$

where $\{W_n\}$ is an independent sequence with zero mean and variance 1. Let $R_X(k) = \mathbf{E}(X_n X_{n+k})$, which is the autocorrelation function of $\{X_n\}$.

(a) Find $R_X(k)$ in terms of $R_X(k-1)$ and $R_X(k-2)$ for $k \geq 1$. [30%]

(b) Find $R_X(k)/R_X(0)$ for $k = 1, 2, 3$ and hence sketch the autocorrelation function. [30%]

(c) A modeller suggests that the model

$$Z_n = \alpha Z_{n-2} + W_n$$

for $|\alpha| < 1$ can approximate $R_X(k)$ well.

(i) Find $R_Z(k)/R_Z(0)$ and comment on the appropriateness of this model for data $X_1, X_2, \ldots.$ [30%]

(ii) For an appropriately chosen α, justify why the model $Z_n = \alpha Z_{n-1} + W_n$ may be better suited to explain this data. [10%]

3 It is desired to fit a single sinusoid to noisy oscillating data measured from a distant galaxy. The model is posed as

$$x_n = A\sin(\omega n) + B\cos(\omega n) + \epsilon_n$$

where A and B are unknown coefficients, ω is a *known* frequency and ϵ_n is a random zero-mean noise process. n is an integer time index.

(a) Define the term *joint probability density function* $f_{A,B}(a,b)$ for the random variables (A,B) evaluated at $A = a$ and $B = b$, in terms of an event $(a_l < A \leq a_u,\ b_l < B \leq b_u)$. [10%]

(b) Suppose the joint probability density function of (A,B) is *circularly symmetric*, which means that it may be expressed as

$$f_{A,B}(a,b) = f(a^2 + b^2)$$

where $f()$ is a scalar function of $a^2 + b^2$. Show that the cumulative distribution function (CDF) for $R = \sqrt{A^2 + B^2}$ is given by

$$F_R(r) = 2\pi \int_0^r \rho f(\rho^2) d\rho$$

and hence show that the probability density function for R is

$$f_R(r) = 2\pi r f(r^2)$$

What is the distribution of the phase, $tan^{-1}(B/A)$? [30%]

(c) Suppose that A and B are independent, zero mean Gaussian random variables with variance equal to 1. Determine and sketch the probability density functions for both $R = \sqrt{A^2 + B^2}$ and $R^2 = A^2 + B^2$. [30%]

(d) Suppose there are N data measurements, $\mathbf{x} = [x_0, x_1, ..., x_{N-1}]$. An unbiased estimator $\hat{\theta}(\mathbf{x}) = [\hat{A}, \hat{B}]$ is constructed for $\theta = [A, B]$. An estimate for R^2 is now constructed as $\hat{R^2} = |\hat{\theta}(\mathbf{x})|^2 = \hat{A}^2 + \hat{B}^2$. Show that this estimator for R^2 is biased since

$$E[\hat{A}^2] = A^2 + \mathrm{var}(\hat{A})$$

Explain why the OLS estimator would give the smallest bias in R^2 of any linear unbiased estimator under this scheme. What do you think would happen to this bias as $N \to \infty$? [30%]

4 A smoothing filter is to be used in enhancement of a signal pulse in additive noise. The filter has the following recursive form:

$$x_n = \alpha x_{n-1} + u_n$$

where $0 < \alpha < 1$ is a parameter of the filter, u_n is its input signal and x_n its output. Suppose that the input signal is a noisy rectangular pulse of the form:

$$u_n = s_n + v_n$$

where $\{v_n\}$ is zero-mean white noise having variance σ_v^2 and, for $n = -\infty, \ldots, -1, 0, 1, \ldots, \infty$,

$$s_n = \begin{cases} 1, & n = 0, \ldots, N-1 \\ 0, & \text{otherwise} \end{cases}$$

(a) Show that the impulse (unit pulse) response of the filter is

$$h_p = \alpha^p, \ \ p \geq 0$$

and sketch the output of the filter in response to the pulse s_n alone, i.e. setting $v_n = 0$. [20%]

(b) Show that the power Signal-to-noise ratio (SNR) at the output of the filter at time index $n = N - 1$ can be written as

$$\frac{(1-\alpha^N)^2(1+\alpha)}{(1-\alpha)\sigma_v^2}$$

Determine the value of α that gives the best output SNR when $N = 2$ and comment on the result compared with the SNR of the unfiltered signal u_n at $n = N - 1$. [30%]

(c) Now, set $N = 10$. Determine, using approximate numerical optimisation if necessary, the value of α which gives best output SNR at $n = N - 1$. [20%]

(d) Design the optimal matched filter for this problem and comment on its performance compared with the smoothing filter above for $N = 2$ and $N = 10$, using the optimal values of α computed in (b) and (c) above. [30%]

END OF PAPER

THIS PAGE IS BLANK