Version GTC/1

Worked Solutions

EGT2

ENGINEERING TRIPOS PART IIA

Wednesday 6 May 2026   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 Let X_t be a random process on the integers. Starting at $X_0 = 0$, at each time the process takes a step forward with probability $\alpha > 0.5$ or a step backwards with probability $1 - \alpha$,

$$X_{t+1} = \begin{cases} X_t + 1 & \text{with probability } \alpha \\ X_t - 1 & \text{with probability } 1 - \alpha \end{cases}$$

Let $q_k(t)$ be the probability that the process reaches the point k for the first time in t steps. So, for example, $q_1(1) = \alpha$ because the process reaches 1 in the first step with probability α.

(a) Show that

(i) $q_1(2) = 0$

(ii) $q_1(3) = (1 - \alpha)\alpha^2$

Further, show that in general

$$q_1(t) = \alpha\delta_{t,1} + (1 - \alpha)q_2(t - 1)$$

[20%]

Solution:

(i) After two steps you can only have X_2 be -2, 0, or 2. So, you cannot reach 1 for the first time at $t = 2$ and $q_1(2) = 0$.

(ii) To reach 1 for the first time in three steps the first step must be backwards and the next two forward. The sequence is $X = (0, -1, 0, +1)$ and probability of this exact sequence is $q_1(3) = (1 - \alpha)\alpha^2$.

In general, the first time to reach 1 will be $t = 1$ if the first step is forward, which occurs with probability α. Otherwise, for $t > 1$ to be at 1 for the first time at t you must initially take a step back, with probability $1 - \alpha$, and then reach two steps forward (to move from -1 to $+1$) for the first time in $(t - 1)$. So, $q_1(t) = \alpha\delta_{t,1} + (1 - \alpha)q_2(t - 1)$.

(b) Let

$$G_k(z) = \sum_{t=1}^{\infty} z^t q_k(t)$$

be the generating function for $q_k(t)$. State the formula relating $G_k(z)$ to $G_1(z)$ and justify your answer. [20%]

Solution:

$G_k(z) = G_1(z)^k$

This is because the time to reach k for the first time is the same as the time it takes to move forward by 1 step k different times. Each of these times are independent, so the generating function for $G_k(z)$ is just $G_1(z)^k$.

(c) Show that $G_1(z)$ obeys

$$G_1(z) = \alpha z + (1-\alpha) z G_1(z)^2$$

and verify that at $z = 1$, $G_1(1) = 1$ is a solution. [30%]

Solution:

$$\begin{aligned}\sum_{t=1}^{\infty} q_1(t) z^t &= \alpha z + (1-\alpha) z \sum_{t=1}^{\infty} q_2(t-1) z^{t-1} \\ &= \alpha z + (1-\alpha) z G_2(z) \\ &= \alpha z + (1-\alpha) z G_1(z)^2\end{aligned}$$

Note, at $z = 1$ we get $G_1(1) = \alpha + (1-\alpha) G_1(1)^2$, which has $G_1(1) = 1$ as a solution, as it should.

(d) Using the generating functions, find the expected time it takes to progress one step forward, $\sum_{t=1}^{\infty} t q_1(t)$, and the expected time to reach k for the first time, $\sum_{t=1}^{\infty} t q_k(t)$. What happens as $\alpha \to 0.5$? [30%]

Solution:

Differentiating $G_1(z) = \sum_t z^t q_1(t)$ term by term gives $G_1'(z) = \sum_t t z^{t-1} q_1(t)$, so $G_1'(1) = \sum_t t q_1(t)$ is the expected time to first reach +1.

Differentiating the relation from part (c):

$$G_1'(z) = \alpha + (1-\alpha) G_1(z)^2 + 2z(1-\alpha) G_1(z) G_1'(z)$$

Setting $z = 1$ and using $G_1'(1) = \mu$, we get

$$\mu = 1 + 2(1-\alpha)\mu = \frac{1}{2\alpha - 1}$$

Also,

$$G_k'(z) = \frac{d}{dz} G_1(z)^k = k G_1(z)^{k-1} G_1'(z)$$

so, using $G_1(1) = 1$ and $G_1'(1) = 1/(2\alpha - 1)$, $G_k'(1) = k/(2\alpha - 1)$.
As $\alpha \to 1/2$, $\mu \to \infty$, and so in expectation it takes an arbitrarily long time to even reach $+1$.

---

2 Consider a Markov chain on N states, labelled $1, 2, \ldots, N$. Let the matrix Q_{ij} denote the transition probability from state i to state j, and assume that the chain has a stationary distribution, $\boldsymbol{\pi} = (\pi_1, \pi_2, \ldots, \pi_N)$.

At time $t = 0$ the chain is initialised in state i with probability $p_i(0)$, and let $p_i(t)$ denote the probabity of being in state i at time t.

(a) State the:

(i) equation that relates $p_i(t+1)$ to the values $p_j(t)$;

(ii) equation that π_i must satisfy because $\boldsymbol{\pi}$ is a stationary distribution.

[20%]

Solution:

(i) $p_i(t+1) = \sum_j p_j(t) Q_{ji}$

(ii) $\pi_i = \sum_j \pi_j Q_{ji}$

(b) Let

$$d(t) = \sum_{i=1}^{N} \left| p_i(t) - \pi_i \right|$$

Show that $d(t+1) \leq d(t)$. [30%]

Solution:

$$\begin{aligned}
d(t+1) &= \sum_i |p_i(t+1) - \pi_i| \\
&= \sum_i \Big| \sum_j Q_{ji}(p_j(t) - \pi_j) \Big| && \text{(by parts (a)i and (a)ii)} \\
&\leq \sum_i \sum_j |Q_{ji}| \left| p_j(t) - \pi_j \right| && \text{(by triangle ineq)} \\
&= \sum_j \left| p_j(t) - \pi_j \right| && \text{(because } \sum_i Q_{ji} = 1) \\
&= d(t)
\end{aligned}$$

(c) Let f_i be any function of the states. At time t the expected value is

$$E_t[f] = \sum_{i=1}^{N} p_i(t) f_i$$

while in the stationary distribution $\boldsymbol{\pi}$ the expected value is $\mu = \sum_{i=1}^{N} \pi_i f_i$. Show that

$$\left|E_t[f] - \mu\right| \leq M d(t)$$

where M is the maximum value of $|f_i|$. [30%]

Solution:

$$\begin{aligned}
\left|E_t[f] - \mu\right| &= \Big|\sum_i p_i(t) f_i - \sum_i \pi_i f_i\Big| \\
&= \Big|\sum_i f_i(p_i(t) - \pi_i)\Big| \\
&\leq \sum_i \big|f_i(p_i(t) - \pi_i)\big| \qquad \text{(by triangle ineq.)} \\
&= \sum_i |f_i| \big|(p_i(t) - \pi_i)\big| \\
&\leq M \sum_i \big|(p_i(t) - \pi_i)\big| = M d(t)
\end{aligned}$$

(d) Finite irreducible aperiodic Markov chains are ergodic. Comment on this fact, and contrast and compare it to the results of parts (b) and (c). [20%]

Solution:

For such an ergodic chain, there is a unique stationary distribution and $\lim_{t\to\infty} p_i(t) = \pi_i$ for all i, regardless of the initial distribution. By (b) this implies $d(t) \to 0$, and by (c) we then have $E_t[f] \to \mu$.

However, in (b) and (c) we did not assume the chain was irreducible or aperiodic so it is possible that $d(t) \to$ const. > 0. For example, the deterministic period-2 chain on two states with

$$Q = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

has stationary $\boldsymbol{\pi} = (1/2, 1/2)$, but starting from $p(0) = (1, 0)$ the chain alternates between $(1, 0)$ and $(0, 1)$, giving $d(t) = 1$ for all t.

---

3 A wide-sense stationary discrete-time real-valued random process $\{X_n\}$ is output through a zero-order hold circuit with sampling interval T seconds so that the continuous-time output is:

$$X(t) = X_n, \ \ nT \leq t < (n+1)T\,,\ n = -\infty, ..., -1, 0, 1, ..., \infty.$$

(a) If $\{X_n\}$ has power spectrum $\mathcal{S}_X(\exp(j\Omega))$, show that the average power of $X(t)$ is

$$\frac{1}{\pi T}\int_0^{\pi} \mathcal{S}_X(\exp(j\Omega))d\Omega\,.$$

[20%]

Solution:

We know from definition of power spectrum that

$$E[X_n^2] = R_{XX}[0] = \frac{1}{2\pi}\int_{-\pi}^{\pi} \mathcal{S}_X(\exp(j\Omega))e^{j0}d\Omega = \frac{1}{\pi}\int_0^{\pi} \mathcal{S}_X(\exp(j\Omega))d\Omega$$

Noting that $\mathcal{S}_X$ has even symmetry for real signals.

But the continuous time signal is a step function with constant level values of X_n between sampling points. So, average power is $E[X_n^2]/T$,

$$= \frac{1}{\pi T}\int_0^{\pi} \mathcal{S}_X(\exp(j\Omega))d\Omega$$

(b) A *causal* discrete time filter with unit pulse response $\{h_n\}$, $n = 0, ..., \infty$ is applied to $\{X_n\}$. Show that the output $\{Y_n\}$ is wide-sense stationary if

$$\sum_{n=0}^{\infty} |h_n| < \infty\,.$$

[30%]

Solution:

Convolution equation in causal case:

$$Y_n = \sum_{m=0}^{\infty} h_m X_{n-m}$$

$$E[Y_n] = \sum_{m=0}^{\infty} h_m E[X_{n-m}] = \mu_X \sum_{m=0}^{\infty} h_m$$

so the mean is stationary (and finite if $\sum_{n=0}^{\infty} |h_n| < \infty$), Condition 1.

$$
\begin{aligned}
R_{YY}[k] = E[Y_n Y_{n+k}] &= E[\sum_{m_1=0}^{\infty} h_{m_1} X_{n-m_1} \sum_{m_2=0}^{\infty} h_{m_2} X_{n+k-m_2}] \\
&= \sum_{m_1=0}^{\infty} \sum_{m_2=0}^{\infty} h_{m_1} h_{m_2} E[X_{n-m_1} X_{n+k-m_2}] \\
&= \sum_{m_1=0}^{\infty} \sum_{m_2=0}^{\infty} h_{m_1} h_{m_2} R_{XX}[m_2 - k - m_1]
\end{aligned}
$$

Hence only depends on k - condition 2.

Now,

$$
\begin{aligned}
|R_{YY}[0]| &= \Big| \sum_{m_1=0}^{\infty} \sum_{m_2=0}^{\infty} h_{m_1} h_{m_2} R_{XX}[m_2 - m_1] \Big| \\
&\leq R_{XX}[0] \Big| \sum_{m_1=0}^{\infty} \sum_{m_2=0}^{\infty} h_{m_1} h_{m_2} \Big| \\
&= R_{XX}[0] \Big| \sum_{m=0}^{\infty} h_m \Big|^2 \\
&\leq R_{XX}[0] \Big(\sum_{m=0}^{\infty} |h_m| \Big)^2
\end{aligned}
$$

which is finite if $\sum_{m=0}^{\infty} |h_m| < \infty$. Under this condition, $\sigma_X^2 = R_{XX}[0] - \mu_Y^2$ is also finite, as required for Condition 3.

(c) Suppose that $\{X_n\}$ is passed through a digital filter satisfying the condition from part (b) above to give a new process $\{Y_n\}$. Show that the power spectrum of $\{Y_n\}$ is

$$
\mathcal{S}_Y(\exp(j\Omega)) = \mathcal{S}_X(\exp(j\Omega)) |G(\exp(j\Omega))|^2
$$

where $G(\exp(j\Omega))$ is a function that should be defined in terms of $\{h_n\}$. [30%]

Solution:

Following lecture notes:

$$
R_{YY}[k] = h_p * h_{-p} * R_{XX}[k]
$$

So, in frequency domain:

$$
\mathcal{S}_Y = \mathcal{S}_X H H^* = \mathcal{S}_X |H|^2
$$

where H is the frequency response of the filter,

$$G(e^{j\Omega}) = \sum_{p=0}^{\infty} h_p e^{-jp\Omega}$$

(d) The function $G(\exp(j\Omega))$ from part (c) above is now defined for $\Omega \in [-\pi, \pi]$ to be:

$$G(\exp(j\Omega)) = \begin{cases} 1, & |\Omega| \in [\Omega_l, \Omega_u) \\ 0, & \text{Otherwise} \end{cases}, \ \Omega_l < \Omega_u < \pi$$

By considering the Power Spectrum of $\{Y_n\}$ in this case, give an interpretation of the Power Spectrum as a *Power spectral-density* for discrete-time random processes. [20%]

Solution:

With this form for G,

$$E[Y_n]^2 = \frac{1}{\pi} \int_{\Omega_l}^{\Omega_u} \mathcal{S}_X(e^{j\Omega}) d\Omega$$

Or for an infinitessimal range Ω to $\Omega + d\Omega$:

$$E[Y_n]^2 = \frac{d\Omega}{\pi} \mathcal{S}_X(e^{j\Omega})$$

So $\mathcal{S}_X$ can be thought of as a *power density* function.

---

4 A set of measurements $\{x_i\}$, $i = 1, ..., N$, is assumed to be drawn independently from a normal distribution, $\mathcal{N}(\mu, \sigma^2)$, but the mean, μ and variance, σ^2, are unknown and believed to take the form

$$\mu = a + bV, \quad \sigma^2 = V$$

where a and b are known constants and V is a random variable.

(a) Show that the likelihood function for V may be expressed as

$$p(\{x_i\}|V = v) = Av^p e^{-(Bv+C/v)/2}$$

and give expressions for A, B, C and p in terms of the $\{x_i\}$, N, a and b. [20%]

Solution:

Likelihood is independent product of normals:

$$\begin{aligned}
p(\{x_i\}|V = v) &= \prod_i \mathcal{N}(x_i|a + bV, V) \\
&= \frac{1}{(2\pi V)^{N/2}} \exp(-\frac{1}{2V} \sum_i (x_i - a - bV)^2) \\
&= \frac{1}{(2\pi V)^{N/2}} \exp(-\frac{1}{2}(\sum_i (x_i - a)^2/V + Nb^2V - 2b \sum_i (x_i - a))) \\
&= \frac{1}{(2\pi)^{N/2}} \exp(\frac{1}{2}(2b \sum_i (x_i - a)))V^{-N/2} \exp(-\frac{1}{2}(\sum_i (x_i - a)^2/V + Nb^2V))
\end{aligned}$$

from which:

$$A = \frac{1}{(2\pi)^{N/2}} \exp(\frac{1}{2}(2b \sum_i (x_i - a)))$$

$$p = -N/2$$

$$B = b^2N$$

$$C = \sum_i (x_i - a)^2$$

(b) Determine the Maximum Likelihood (ML) estimator for V. Sketch the shape of the likelihood function as a function of v, paying attention to any stationary points and the behaviour at $v = 0$. [30%]

Solution:

Need to maximise wrt v:

$$dp(\{x_i\}|v)/dv = pv^{p-1}e^{-(Bv+C/v)/2} + v^p e^{-(Bv+C/v)/2}(-B + C/v^2)/2 = 0$$

i.e.

$$v^2B/2 - vp - C/2 = 0$$

$$v = \frac{p \pm \sqrt{p^2 + CB}}{B}$$

Neglect the negative solution (CB is positive) since this cannot be the variance of the normal :

$$v^{ML} = \frac{p + \sqrt{p^2 + CB}}{B}$$

Hence there is a maximum at v^{ML}. The term $exp(-Bv/2)$ ensures that the curve comes back to zero as $v \to \infty$, and the term $exp(-C/(2v))$ ensures that it starts at zero when $v = 0$, whatever the value of the exponent p (expontial beats polynomial). The derivative at zero is also zero, by a similar exponential vs polynomial argument, so the curve starts flat at zero.

Sketch....

(c) A prior distribution is assigned to V with the form

$$p(v) = \sqrt{\frac{1}{2\lambda\pi v^3}} \exp\left[-\frac{(v-c)^2}{2\lambda c^2 v}\right]$$

where c and $\lambda > 0$ are constants.

Using this prior and the likelihood function of part (a), determine the posterior density $p(v|\{x_i\})$ and find the Maximum *a posteriori* (MAP) estimator for V. You may neglect the normalising constant for the posterior density. [20%]

Solution:

First notice that this prior has the same form as the likeiihood in (a):

$$p(v) \propto v^{-3/2} \exp\left[-\frac{v + c^2/v}{2\lambda c^2}\right]$$

i.e. $B' = 1/(\lambda c^2)$ and $C' = 1/\lambda$, $p' = -3/2$.

Thus the posterior is immediately obained as

$$p(v|\{x_i\}) \propto p(\{x_i\}|v)p(v) = v^{p+p'} e^{-((B+B')v+(C+C')/v)/2}$$

and the MAP solution is

$$v^{MAP} = \frac{p + p' + \sqrt{(p+p')^2 + (C+C')(B+B')}}{(B+B')}$$

(d) In a new experiment, samples $\{v_i\}$, $i = 1, ..., M$, are drawn independently from a prior probability distribution as in part (c). Determine the ML estimators for c and λ, using only the new data $\{v_i\}$ and give an interpretation of c. [30%]

Solution:

Write down the likelihood function:

$$p(\{v_i\}|c,\lambda) = \sqrt{\frac{1}{(2\lambda\pi)^M}} \prod_i v_i^{-3/2} \exp\left[-\frac{\sum_i (v_i - c)^2/v_i}{2\lambda c^2}\right]$$
$$= \sqrt{\frac{1}{(2\lambda\pi)^M}} \prod_i v_i^{-3/2} \exp\left[-\frac{\sum_i (v_i - 2c + c^2/v_i)}{2\lambda c^2}\right]$$
$$= \sqrt{\frac{1}{(2\lambda\pi)^M}} \prod_i v_i^{-3/2} \exp\left[-\frac{1}{\lambda}\left(\frac{\sum_i v_i}{2c^2} - \frac{M}{c} + \frac{\sum_i (1/v_i)}{2}\right)\right]$$

First observe that only the exponent:

$$f(c) := \frac{\sum_i v_i}{2c^2} - \frac{M}{c} + \frac{\sum_i (1/v_i)}{2}$$

depends on c, so this can be optimised wrt c:

$$df(c)/dc = -\frac{\sum_i v_i}{c^3} + \frac{M}{c^2}$$

i.e.

$$c^{ML} = \frac{\sum_i v_i}{M},$$

the sample mean. Thus the interpretation of c is likely to be the mean of the distribution. λ may be obtained from the whole expression, at $c = c^{ML}$:

$$p(\{v_i\}|c,\lambda) \propto \sqrt{\frac{1}{(\lambda)^M}} \exp\left[-\frac{1}{\lambda}\left(-\frac{M^2}{2\sum_i v_i} + \frac{\sum_i (1/v_i)}{2}\right)\right]$$

Differentiating wrt λ gives:

$$-M/2\lambda^{-M/2-1} \exp(...) + \exp(...)\lambda^{-M/2-2}\left(-\frac{M^2}{2\sum_i v_i} + \frac{\sum_i (1/v_i)}{2}\right) = 0$$

i.e.

$$M/2(\lambda) = \left(-\frac{M^2}{2\sum_i v_i} + \frac{\sum_i (1/v_i)}{2}\right)$$

and

$$\lambda^{ML} = \frac{\sum_i (1/v_i)}{M} - \frac{M}{\sum_i v_i} = \frac{\sum_i (1/v_i)}{M} - \frac{1}{c^{ML}}$$

END OF PAPER