Version Sjg/0

EGT3
ENGINEERING TRIPOS PART IIA

April 2022

Module 3F3

STATISTICAL SIGNAL PROCESSING - WORKED SOLUTIONS

*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 Two machines, machine A and machine B, both produced 1kg bags of sugar. Each bag of sugar produced by machine A either weighs 1kg with probability p or weighs 1.001kg with probability $1 - p$. A bag of sugar produced by machine B will either weigh 1kg with probability q or weigh 0.999kg with probability $1 - q$. Let X_n be the weight of the nth bag of sugar produced by machine A, $n = 1, 2, \ldots$ and let Y_n be the weight of the nth bag of sugar produced by machine B.

A consignment is assembled by taking bags from machines A and B according to the following rule:

$$S_{n+1} = \begin{cases} S_n + X_{n+1} & \text{if } S_n = n \\ S_n + Y_{n+1} & \text{otherwise} \end{cases}$$

where S_n is the total weight of the consignment after selecting n bags. Assume S_0=0.

(a) List the possible values, or range, for the random variables S_1, S_2 and S_n. [10%]

(b) Find $\Pr(S_2 = 2)$ and $\Pr(S_2 = 2.001)$. [10%]

(c) Verify the Markov property $\Pr(S_n = s_n | S_1 = s_1, \ldots, S_{n-1} = s_{n-1}) = \Pr(S_n = s_n | S_{n-1} = s_{n-1})$ is satisfied. [20%]

(d) Find the transition probabilities $\Pr(S_{n+1} = j | S_n = i)$ for all values of i and j in the range of S_n and S_{n+1} respectively. [30%]

(e) Let $\pi_n = \Pr(S_n = n)$. Find π_{n+1} in terms of π_n, p and q. [20%]

(f) For a consignment of K bags, find the expected value of the number of bags that come from machine A. [10%]

SECTION *

Question 1

Part (a)

Range of S_1: since $S_1 = X_1$, $S_1 \in \{1, 1.001\}$.
Range of S_2: if $S_1 = 1$ then $S_2 = S_1 + X_2 \in \{2, 2.001\}$. If $S_1 = 1.001$ then $S_2 = S_1 + Y_2 \in \{2, 2.001\}$. So range of S_2 is $\{2, 2.001\}$.
Range of S_3: can be similarly found using the method for S_2 to be $\{3, 3.001\}$. Extrapolate to S_n.

Part (b)

$$\begin{aligned}\Pr(S_2 = 2) &= \Pr(S_2 = 2|S_1 = 1)\Pr(S_1 = 1) \\ &\quad + \Pr(S_2 = 2|S_1 = 1.001)\Pr(S_1 = 1.001) \\ &= p \times p + (1 - q) \times p\end{aligned}$$

and $\Pr(S_2 = 2.001) = 1 - \Pr(S_2 = 2)$.

Part (c)

First find the joint probability and then find the conditional probability.
Case 1: $s_{n-1} = n - 1$

$$\begin{aligned}&\Pr(S_1 = s_1, \ldots, S_{n-1} = s_{n-1}, S_n = s_n) \\ &= \Pr(S_1 = s_1, \ldots, S_{n-1} = s_{n-1}, X_n + s_{n-1} = s_n) \\ &= \Pr(S_1 = s_1, \ldots, S_{n-1} = s_{n-1})\Pr(X_n + s_{n-1} = s_n)\end{aligned}$$

and $\Pr(S_n = s_n|S_{n-1} = s_{n-1}) = \Pr(X_n + s_{n-1} = s_n)$. Thus, after dividing,

$$\Pr(S_n = s_n|S_1 = s_1, \ldots, S_{n-1} = s_{n-1}) = \Pr(S_n = s_n|S_{n-1} = s_{n-1})$$

Case 2: $s_{n-1} = n - 1 + 0.001$, follow exactly the same steps.

Full marks for either case 1 or case 2.

Part (d)

The range of S_n was previously found to be $\{n, n.001\}$

$$\begin{aligned}
\Pr(S_{n+1} = n+1|S_n = n) &= p \\
\Pr(S_{n+1} = n+1+0.001|S_n = n) &= 1-p \\
\Pr(S_{n+1} = n+1+0.001|S_n = n.001) &= q \\
\Pr(S_{n+1} = n+1|S_n = n.001) &= 1-q
\end{aligned}$$

Part (e)

$$\begin{aligned}
\Pr(S_{n+1} = n+1) &= \Pr(S_{n+1} = n+1|S_n = n)\pi_n + \Pr(S_{n+1} = n+1|S_n = n.001)(1-\pi_n) \\
\pi_{n+1} &= p\pi_n + (1-q)(1-\pi_n)
\end{aligned}$$

Part (f)

Machine A is selected every time $S_k = k$ or number of times machine A is selected is

$$\sum_{k=0}^{K-1} \mathbb{I}_{[S_k=k]}$$

where $\mathbb{I}_{[S_k=k]} = 1$ if $S_k = k$ and is 0 otherwise. Taking the expectations gives the sum $\sum_{k=0}^{K-1} \pi_k$.

2 (a) From measurements, a wide sense stationary random process $\{X_n\}$ is found to have zero mean and its autocorrelation function is estimated to be $\widehat{R}_X(0) = 1, \widehat{R}_X(1) = 0.9, \ldots, \widehat{R}_X(l) = 0.9^l, \ldots$
Find α and the variance of the noise W_n of the first order autoregressive (AR) model

$$X_n = \alpha X_{n-1} + W_n$$

which corresponds to these estimated autocorrelation values and show the corresponding power spectrum estimate is

$$\frac{1-0.9^2}{|1-0.9\exp(-j\omega)|^2}.$$

[30%]

 (cont.

(b) Let $\{Y_n\}$ be another (mean zero) first order autoregressive random process, independent from $\{X_n\}$, but with the same estimated autocorrelation function $\widehat{R}_Y(k) = 0.9^k$ for $k \geq 0$. Data from the two autoregressive random processes have been merged together mistakenly to form a single data sequence $Z_0, Z_1, \ldots$

$$Z_n = \begin{cases} X_k & \text{if } n = 2k \\ Y_k & \text{if } n = 2k+1. \end{cases}$$

For example, $(Z_0, Z_1, Z_2, Z_3, Z_4, Z_5) = (X_0, Y_0, X_1, Y_1, X_2, Y_2)$. Find the autocorrelation function $R_Z(l) = \mathbf{E}(Z_0 Z_l)$ for all l. [30%]

(c) Find the autocorrelation function of the following AR(2) model

$$Z_n = \alpha Z_{n-2} + \beta_0 W_n + \beta_1 W_{n-1}$$

where $\{W_n\}$ is an independent sequence with zero mean and variance 1. [30%]

(d) Find the values of β_0 and β_1 that best matches the autocorrelation you found in part-(b). [10%]

SECTION *

Question 2

Part (a)

Let variance of W_n be σ^2. $\mathbf{E}(X_n^2) = \alpha^2 \mathbf{E}(X_{n-1}^2) + \sigma^2$ so $R_X(0) = \sigma^2/(1-\alpha^2)$. $R_X(1) = \mathbf{E}(X_n X_{n-1}) = \alpha R_X(0)$ and $R_X(2) = \mathbf{E}(X_n X_{n-2}) = \mathbf{E}(\alpha X_{n-1} X_{n-2}) + 0 = \alpha R_X(1)$. Thus $R_X(k) = \alpha R_X(k-1)$.

Matching estimate: clearly $\alpha = 0.9$ and $\sigma^2 = 1 - \alpha^2$.

The power spectrum is

$$
\begin{aligned}
S_X(\omega) &= \sum_{k=-\infty}^{\infty} R_X(k)\exp(-j\omega k) \\
&= \sum_{k=-\infty}^{0} R_X(k)\exp(-j\omega k) + \sum_{k=0}^{\infty} R_X(k)\exp(-j\omega k) - 1 \\
&= \sum_{k=0}^{\infty} \alpha^k \exp(j\omega k) + \sum_{k=0}^{\infty} \alpha^k \exp(-j\omega k) - 1 \\
&= \frac{1}{1-\alpha\exp(j\omega)} + \frac{1}{1-\alpha\exp(-j\omega)} - 1
\end{aligned}
$$

Simplify to get final form.

Part (b)

We can see that

$$(Z_0, Z_1, Z_2, Z_3, Z_4, Z_5) = (X_0, Y_0, X_1, Y_1, X_2, Y_2)$$

and so on.

$$
\begin{aligned}
\mathbf{E}(Z_0^2) &= \mathbf{E}(X_0^2) = R_X(0) = 1 \\
\mathbf{E}(Z_0 Z_1) &= \mathbf{E}(X_0 Y_0) = 0 \\
\mathbf{E}(Z_0 Z_2) &= \mathbf{E}(X_0 X_1) = R_X(1) = \alpha \\
\mathbf{E}(Z_0 Z_3) &= \mathbf{E}(X_0 Y_1) = 0 \\
\mathbf{E}(Z_0 Z_4) &= \mathbf{E}(X_0 X_2) = R_X(2) = \alpha^2 \\
\mathbf{E}(Z_0 Z_l) &= R_X(l/2) = \alpha^{l/2} \qquad \text{if } l \text{ even} \\
\mathbf{E}(Z_0 Z_l) &= 0 \qquad \text{if } l \text{ even}
\end{aligned}
$$

Part (c)

We can see there are no noise terms W in common between Z_n and Z_{n+2}, between Z_n and Z_{n+3}, and between Z_n and Z_{n+k} for $k \geq 2$. Thus for $k \geq 2$,

$$R_Z(k) = \mathbf{E}(Z_n Z_{n+k}) = \alpha \mathbf{E}(Z_n Z_{n+k-2}) = \alpha R_Z(k-2).$$

So now find $R_Z(0)$ and $R_Z(1)$

$$R_Z(0) = \mathbf{E}(Z_n^2) = \alpha^2 R_Z(0) + \beta_0^2 + \beta_1^2$$

$$R_Z(1) = \mathbf{E}(Z_n Z_{n+1}) = \alpha R_Z(1) + \beta_0 \beta_1$$

Part (d)

Setting $\beta_1 = 0$ then $R_Z(0) = \alpha^2 R_Z(0) + \beta_0^2$ and $R_Z(1) = \alpha R_Z(1)$ which implies $R_Z(1) = 0$. So we choose $\beta_0 = 1 - \alpha^2$ to give $R_Z(0) = 1$.

3 (a) It is desired to estimate the value of a parameter λ from a sequence of random data. The estimator is denoted $\hat{\lambda}$. Explain the terms *bias* and *variance* of a statistical estimator. The mean-squared error is defined as $MSE = E[(\hat{\lambda} - \lambda)^2]$. Show that

$$MSE = variance + bias^2$$

and hence explain why it is not always best to design an unbiased estimator. [20%]

Solution: *Bias is* $E[\hat{\lambda}] - \lambda$. *Variance is* $E[(\hat{\lambda} - E[\hat{\lambda}])^2]$.

MSE:

$$E[(\hat{\lambda} - \lambda)^2] = E[\hat{\lambda}^2] + \lambda^2 - 2\lambda E[\hat{\lambda}]$$

$$Variance + bias^2 = E[\hat{\lambda}^2] - E[\hat{\lambda}]^2 + (E[\hat{\lambda}] - \lambda)^2 = E[\hat{\lambda}^2] - E[\hat{\lambda}]^2 + E[\hat{\lambda}]^2 + \lambda^2 - 2\lambda E[\hat{\lambda}]$$

$$= E[\hat{\lambda}^2] + \lambda^2 - 2\lambda E[\hat{\lambda}] = MSE$$

So we can see that it may be possible to get lower MSE by trading off some bias for some lower variance.

(b) In a low illumination imaging experiment it is assumed that the number of photons N_k arriving at a particular pixel in time interval k is independent and Poisson distributed with mean $\mu_k = a\exp(-k)$, $k = 0, ..., K-1$, i.e.

$$p(N_k|a) = \frac{\mu_k^{N_k} e^{-\mu_k}}{N_k!}$$

It is proposed to estimate the underlying constant a using the following formula:

$$\hat{a} = \alpha \sum_{k=0}^{K-1} w_k N_k$$

where w_k are positive weights such that $\sum_{k=0}^{K-1} w_k = 1$, and where α is a constant. Determine the value of α which makes the bias of the estimator zero, for a fixed set of weights w_k. [20%]

Solution:

$$E[\hat{a}] = \alpha \sum w_k E[N_k] = \alpha a \sum w_k \exp(-k)$$

So we need $\alpha = 1/(\sum w_k \exp(-k))$ *for an unbiased estimator.*

Examiner's comment: most candididates did not use this simple method and used a more lengthy but correct (if error-prone) method for directly calculating $E[\hat{a}^2]$.

(c) Show that the MSE of the estimator, under the unbiased condition derived above, can be written as:

$$a\frac{\sum_{k=0}^{K-1} w_k^2 e^{-k}}{(\sum_{k=0}^{K-1} w_k e^{-k})^2}$$

[30%]

Solution: *MSE:*

The estimator is unbiased so MSE=var.

$$var(\hat{a}) = \alpha^2 \sum w_k^2 var(N_k) = \alpha^2 \sum w_k^2 a e^{-k} = a\frac{\sum w_k^2 e^{-k}}{(\sum w_k \exp(-k))^2}$$

since $var(aX + bY) = a^2 var(X) + b^2 var(Y)$ *for independent RVs.*

(d) Determine and sketch the likelihood function for a, clearly marking any salient points. Determine the Maximum likelihood (ML) estimator and discuss how it relates to the weighted estimator above. What is the bias and variance of the ML estimator? [30%]

Solution: *Likelihood is:*

$$\prod_{k=0}^{N-1} \frac{\mu_k^{N_k} e^{-\mu_k}}{N_k!}$$

Sketch... gamma density shape...

ML found by taking log-likelihood and neglecting any additive constants (that don't depend on a):

$$(\sum N_k) \log a - a \sum \exp(-k)$$

Take derivative wrt a and set to zero:

$$a^{ML} = (\sum N_k)/(\sum \exp(-k))$$

Comparing with the above estimator, it is unbiased with uniform weights, $w_k = 1/K$. *Interestingly the ML estimator does not require the individual* N_k*s, only their sum (a 'sufficient' statistic), which reflects the special properties of Poisson processes. Under this condition the bias is zero and the variance=MSE is equal to* $a/\sum e^{-k}$.

Examiner's comment: many candidates wrote down the likelihood for just a single point N_k *and not for all K data points.*

---

4 (a) A random process is defined as:

$$x_n = A\cos(n\omega_0 + \phi) + B$$

where $A \sim \mathcal{N}(\mu_A, \sigma_A^2)$ and $B \sim \mathcal{N}(0, \sigma_B^2)$ are mutually independent Gaussian random variables. ω_0 is a fixed frequency and ϕ is a random variable uniformly distributed between 0 and π, independent from A and B.

Show that the autocorrelation function of $\{x_n\}$ is

$$R_{xx}[n_1, n_2] = \sigma_B^2 + 0.5(\mu_A^2 + \sigma_A^2)\cos((n_2 - n_1)\omega_0)$$

and determine whether the process is wide-sense stationary (WSS). [25%]

Solution:

$$\begin{aligned} R_{xx}[n_1, n_2] &= E[x_{n1}x_{n2}] = E[B^2] + E[A^2]E[\cos(n_1\omega_0 + \phi)\cos(n_2\omega_0 + \phi)] \\ &= \sigma_B^2 + E[A^2]0.5E[\cos((n_1 + n_2)\omega_0 + 2\phi) + \cos((n_2 - n_1)\omega_0)], \textit{(trig identities)} \\ &= \sigma_B^2 + E[A^2]0.5(0 + \cos((n_2 - n_1)\omega_0)), \textit{(expectation of cos over one whole period=0)} \\ &= \sigma_B^2 + (\mu_A^2 + 0.5\sigma_A^2)(0 + \cos((n_2 - n_1)\omega_0)), \textit{(expectation of cos over one whole period=0)} \end{aligned}$$

as required.

However the mean of the process is

$$\begin{aligned} E[x_n] &= \mu_B + \mu_A E\cos(n\omega_0 + \phi) \\ &= \mu_B + \mu_A \frac{1}{\pi}\int_0^{\pi} \cos(n\omega_0 + \phi)d\phi = \mu_B + \mu_A/\pi[\sin(n\omega_0 + \phi)]_0^{\pi} \\ &= \mu_B - 2\mu_A/\pi[\sin(n\omega_0)] \end{aligned}$$

which is time-varying. Hence the process is not WSS.

Examiner's comment: many candidates did not check that the mean was constant, or incorrectly calculated it, and just assumed the process was stationary, which it is not, because the phase is not uniformly distributed in 2π.

(b) A WSS process has aero mean and autocorrelation function

$$R_{xx}[n_1, n_2] = \cos((n_2 - n_1)\omega_0)$$

where ω_0 is a constant. Determine whether the process is mean-ergodic. [25%]

Solution: *For mean ergodicity theorem, we need the autocovariance function:*

$$C_{xx}[n_1, n_2] = R_{xx}[n_1, n_2] - E[x_n]^2 = \cos((n_2 - n_1)\omega_0)$$

Now,

$$
\begin{aligned}
1/N\sum_{n=0}^{N} C_{xx}[n=n_2-n_1] &= 1/NRe[\sum_{n=0}^{N}\exp(in\omega_0)], \textit{(Real part of complex exponential)}\\
&= 1/NRe[(1-\exp(i(N+1)\omega_0)/(1-\exp(i\omega_0))], \textit{(Sum of GP)}\\
&\leq 1/N|[(1-\exp(i(N+1)\omega_0)/(1-\exp(i\omega_0))]|, \textit{Real part <= to magnitude}\\
&\leq 1/N2/|(1-\exp(i\omega_0))| \rightarrow 0, \textit{as } N\rightarrow\infty
\end{aligned}
$$

hence it is ergodic in the mean.

[This is quite involved and arguments involving area under cosine function being zero were acceptable]

Examiner's comment: most candidates did not test the above necessary condition, and those that did were often not able to spot the solution using a GP. One or two candidates spotted that the process is not ergodic when $1-\exp(i\omega_0)=0$ i.e. $\omega_0=2k\pi$, but this was not required to get full marks.

(c) The WSS process in (b) above is observed in independent white Gaussian noise:

$$y_n = x_n + v_n,\ \ v_n \sim \mathcal{N}(0,\sigma_v^2)$$

It is desired to predict the next value x_{n+1} based on two previous observed values, y_{n-1} and y_n, using a 2-tap FIR filter.

Derive the filter coefficients that minimise the mean-squared error (MSE) in this prediction. [30%]

Solution: *MSE is*

$$H = E[\epsilon_n^2] = E[(x_{n+1}-\hat{x}_{n+1})^2$$

where

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

Derivatives are:

$$\frac{dH}{dh_0} = E[2\epsilon d\epsilon/dh_0] = -2E[(x_{n+1}-\hat{x}_{n+1})y_n] = -2(r_{XY}[-1]-h_0 r_{YY}[0]-h_1 r_{YY}[1])$$

and

$$\frac{dH}{dh_1} = E[2\epsilon d\epsilon/dh_1] = -2E[(x_{n+1}-\hat{x}_{n+1})y_{n-1}] = -2(r_{XY}[-2]-h_0 r_{YY}[-1]-h_1 r_{YY}[0])$$

and so the conditions for optimal error are:

$$r_{XY}[-1]-h_0 r_{YY}[0]-h_1 r_{YY}[1]=0,\quad r_{XY}[-2]-h_0 r_{YY}[-1]-h_1 r_{YY}[0]=0$$

Now, since $y_n = x_n + v_n$ *and* v_n *is zero mean and independent of* x_n*, we have*

$$R_{YX}[k] = R_{XY}[k] = R_{XX}[|k|]$$

and

$$R_{YY}[k] = R_{XX}[k] + \sigma_v^2 \delta[k]$$

Hence

$$h_0 r_1 = r_2 - h_1(r_0 + \sigma_v^2))$$

$$h_1 r_1 = r_1 - h_0(r_0 + \sigma_v^2)$$

where $r_k = R_{XX}[k]\cos(k\omega_0)$.
Back-substituting for h_1 *gives:*

$$h_0 r_1 = r_2 - 1/r_1(r_1 - h_0(r_0 + \sigma_v^2))(r_0 + \sigma_v^2)$$

and

$$h_1 r_1 = r_1 - 1/r_1(r_2 - h_1(r_0 + \sigma_v^2))(r_0 + \sigma_v^2)$$

so,

$$h_0(r_1 - (1 + \sigma_v^2)^2/r_1) = r_2 - (1 + \sigma_v^2)$$

and

$$h_1(r_1 - (1 + \sigma_v^2)^2/r_1) = r_1 - r_2/r_1(1 + \sigma_v^2)$$

and finally,

$$h_0 = \frac{r_2 r_1 - r_1(1 + \sigma_v^2)}{r_1^2 - (1 + \sigma_v^2)^2}$$

and

$$h_1 = \frac{r_1^2 - r_2(1 + \sigma_v^2)}{r_1^2 - (1 + \sigma_v^2)^2}$$

Examiner's comment: most people got the right idea here, although many made algebraic slips, used the wrong error term $x_n - \hat{x}_{n+1}$ *or simply failed to solve fully for the coefficients.*

(d) Show that as the noise becomes very small (i.e. $\sigma_v^2 \to 0$) the optimal coefficients are:

$$h_0 = 2\cos\omega_0, \quad h_1 = -1$$

Calculate the MSE for this optimal filter and comment on your result. [20%]

Solution: *Substituting* $\sigma_v^2 = 0$ *gives:*

$$h_0 = \frac{r_2 r_1 - r_1}{r_1^2 - 1} = \frac{\cos(2\omega_0) - \cos(\omega_0)}{-\sin^2(\omega_0)^2} = 2\cos\omega_0$$

and

$$h_1 = \frac{r_1^2 - r_2}{r_1^2 - 1} = -1$$

The MSE is:

$$H = E[(x_{n+1} - \hat{x}_{n+1})^2] = r_{XX}[0] - 2r_{X\hat{X}}[0] + r_{\hat{X}\hat{X}}[0]$$

Now,

$$r_{XX}[0] = 1$$

$$r_{X\hat{X}}[0] = E[X_n(h_0 Y_{n-1} + h_1 Y_{n-2})] = h_0 r_1 + h_1 r_2 = 2\cos^2\omega - \cos(2\omega) = 1$$

$$r_{\hat{X}\hat{X}}[0] = E((h_0 Y_{n-1} + h_1 Y_{n-2})^2] = h_0^2 r_{YY}[0] + h_1^2 r_{YY}[0] + 2h_0 h_1 r_{YY}[1] = h_0^2 + h_1^2 + 2h_1 h_2 \cos\omega_0 = 1$$

(Using the optimal values:

$$h_0 = 2\cos(\omega), \quad h_1 = -1)$$

So, finally,

$$H^{opt} = 1 - 2 + 1 = 0$$

Hence it is possible to perfectly predict X given two previous data points (predictable process). In fact this is stating that sinusoids can be perfectly predicted with zero error, but don't necessarily expect candidates to quote this.

Examiner's comment: A satisfying number made a good attempt at explaining why the error goes to zero in the final part

---

END OF PAPER

THIS PAGE IS BLANK