Solutions to 4F7, Adaptive Filters and Spectrum Estimation, 2017, Part IIB

Q1a. For $k \geq 0$, noting that $E(x(n-k)w(n)) = 0$,

$$\begin{aligned} E(x(n-k)d(n)) &= \sum_{m=0}^{\infty} a_m E(x(n-k)x(n-m)) \\ &= \sum_{m=0}^{\infty} a_m \beta^{|k-m|}. \end{aligned}$$

This is a convolution of two sequences. We also see that matrix $\mathbf{R}$ is

$$\begin{bmatrix} 1 & \beta & & \cdots & \beta^{M-1} \\ \beta & 1 & \beta & & \\ & & 1 & & \vdots \\ & & & \ddots & \beta \\ \beta^{M-1} & & \cdots & \beta & 1 \end{bmatrix}$$

Q1b. Differentiate $E\left(d(n) - \mathbf{h}^T\mathbf{x}(n)\right)^2$ wrt to vector $\mathbf{h}$ and set to zero to get the answer.

Q1c. Using $\mathbf{Rh} = \mathbf{p}$ for $M = 1$ we get $h_0 = \sum_{m=0}^{\infty} a_m \beta^m$. The solution is not a_0 since the sequence $x(n)$ is not an independent sequence.

The MSE for $h_0 = a_0$ is $E\left(w(n) + \sum_{m=1}^{\infty} a_m x(n-m)\right)^2 = E\left(w(n)^2\right) + E\left(\sum_{m=1}^{\infty} a_m x(n-m)\right)^2$.

Q1d. Bookwork.

Q1e. Looking at the ith row of the relationship $\mathbf{Rh} = \mathbf{p}$ we see that it is

$$\beta^{i-1}h_0 + \ldots + \beta^0 h_{i-1} + \beta^1 h_i + \ldots + \beta^{|M-i|}h_{M-1} = \sum_{m=0}^{\infty} a_m \beta^{|i-1-m|}.$$

In the limit as $M \to \infty$, the left hand side becomes $\sum_{m=0}^{\infty} h_m \beta^{|i-1-m|}$. Thus each row i of $\mathbf{Rh} = \mathbf{p}$ is expressing the relationship

$$\left(\{h_m\}_{m=0}^{\infty} * \left\{\beta^{|m|}\right\}_{m=-\infty}^{m=\infty}\right)_i = \left(\{a_m\}_{m=0}^{\infty} * \left\{\beta^{|m|}\right\}_{m=-\infty}^{m=\infty}\right)_i.$$

Clearly one solution is $h_m = a_m$ for all m. This is a convex problem and thus the solution is unique.

Q1f-i. For iid source symbols $\beta = 0$ since $E(x(n)) = 0$. The matrix $\mathbf{R}$ is the identity matrix. Row i of vector $\mathbf{p}$ is a_{i-1}. Thus Wiener solution is $h_0 = a_0, \ldots, h_{M-1} = a_{M-1}$.

Q1f-ii. The minimum MSE (or MMSE) is $E\left(w(n) + \sum_{m=M}^{\infty} a_m x(n-m)\right)^2$ which evaluates to $\sigma^2 + \sum_{m=M}^{\infty} a_m^2$. Since sequence $\{a_m\}_{m=0}^{\infty}$ is square summable, the MMSE will decrease with M and asymptote towards σ^2 from above.

Examiner: The most popular and straightforward question, well-answered by most candidates. Part (f)-ii was surprisingly difficult for many.

Q2a. The cost function is $E\left\{(\theta_0(c-1)+cv(0))^2\right\}$ which evaluates to $(c-1)^2E(\theta_0^2)+c^2\sigma_v^2$. Minimising wrt c gives

$$c=\frac{E(\theta_0^2)}{E(\theta_0^2)+\sigma_v^2} \qquad \text{or} \qquad 1-c=\frac{\sigma_v^2}{E(\theta_0^2)+\sigma_v^2}.$$

The MSE is obtained by substitution. Call this minimum MSE

$$\sigma_0^2=(c-1)^2E\left\{\theta_0^2\right\}+c^2E\left\{v(0)^2\right\}.$$

Q2b. $\theta_1-\tilde{\theta}_1=(\theta_0-d\hat{\theta}_0)+w(1)$. The MSE is $E\left\{(\theta_0-d\hat{\theta}_0)^2\right\}+\sigma_w^2$. Note that $E\left\{(\theta_0-d\hat{\theta}_0)^2\right\}=E\left\{(\theta_0-dcy_0)^2\right\}$. Since c is optimal for $E\left\{(\theta_0-cy_0)^2\right\}$, the MSE is minimised when $d=1$. Thus $\tilde{\theta}_1=\hat{\theta}_0=cy_0$. The minimum MSE is $\sigma_0^2+\sigma_w^2$.

Q2c-i. $E(\tilde{\theta}_1)=E(cy_0)=cE(\theta_0)$. $E(y_1)=E(\theta_0)$. $E(\theta_1)=E(\theta_0)$. So $Kc+L=1$ will give an unbiased estimate.

Q2c-ii. Write

$$\begin{aligned}
\hat{\theta}_1 &= Kc\frac{\tilde{\theta}_1}{c}+(L-1)\theta_1+\theta_1+Lv(1)\\
\hat{\theta}_1-\theta_1 &= Kc\frac{\tilde{\theta}_1}{c}-Kc\theta_1+Lv(1)\\
&= Kc\left(\frac{\tilde{\theta}_1}{c}-\theta_1\right)+Lv(1)\\
E\left(\hat{\theta}_1-\theta_1\right)^2 &= (1-L)^2E\left(\frac{\tilde{\theta}_1}{c}-\theta_1\right)^2+L^2\sigma_v^2
\end{aligned}$$

Minimising wrt L gives (from part a)

$$L=\frac{E\left(\frac{\tilde{\theta}_1}{c}-\theta_1\right)^2}{E\left(\frac{\tilde{\theta}_1}{c}-\theta_1\right)^2+\sigma_v^2}=\frac{\sigma_v^2+\sigma_w^2}{2\sigma_v^2+\sigma_w^2}$$

using $\tilde{\theta}_1=cy_0$. Solve for K using $Kc+L=1$.

Q2d. To match the observed process to the previous parts, use y_n/b_n as the observation. But there is a discrepancy in the state transition.

Use the optimal c from part a for the observation y_0/b_0. This gives best estimate of θ_0 given y_0/b_0.

step 1: To get the best estimate $\tilde{\theta}_1$ for θ_1 before observing y_1/b_1, re-solve part (b) to get $d=a_1$.

step 2: Use the scheme in part c-ii to update $\tilde{\theta}_1$ to $\hat{\theta}_1$ using y_1/b_1.

Now loop procedure step 1 and step 2.

Examiner: The second most popular question and turned out to be the most difficult question for the candidates. The solution to part (b) should have followed part (a) almost by inspection. Unfortunately most ended up repeating the calculations. Candidates were largely lost in calculations in part (c)-ii. Part (d) poorly answered and candidates failed to properly connect with the solution to previous parts.

Q3a-i. Strict stationarity implies $p(x_0, \ldots, x_k) = p(x_n, \ldots, x_{n+k})$ for all n and k.

Q3a-ii. We need to assume w_n is white Gaussian noise. This implies $p(x_0, \ldots, x_{P-1})$ is a Gaussian density. Need to find its mean vector and covariance matrix. Let $\mathbf{x}_n = (x_n, \ldots, x_{n-P+1})^T$. Write $\mathbf{x}_n$ as

$$\mathbf{x}_n = A\mathbf{x}_{n-1} + \mathbf{b}w_n$$

where $\mathbf{b} = (1, 0, \ldots 0)^T$. $E(\mathbf{x}_n) = AE(\mathbf{x}_{n-1})$. Thus mean is 0.

Compute $E(\mathbf{x}_n\mathbf{x}_n^T) = AE(\mathbf{x}_{n-1}\mathbf{x}_{n-1}^T)A^T + \sigma^2\mathbf{b}\mathbf{b}^T$. Let $S = E(\mathbf{x}_n\mathbf{x}_n^T) = E(\mathbf{x}_{n-1}\mathbf{x}_{n-1}^T)$ and solve this equation for components of matrix S.

Q3b-i. The process is $x_n = a_2 x_{n-2} + w_n$. We see that odd and even time indices define two independent AR(1) processes. That is let $z_k = x_{2k}$ and $y_k = x_{2k+1}$ for $k = 0, 1, \ldots$ Thus $\{z_k\}_{k\geq 0}$ and $\{y_k\}_{k\geq 0}$ are independent of each other, $x_0 = z_0$, $y_0 = x_1$, $z_k = a_2 z_{k-1} + w_{2k}$ and $y_k = a_2 y_{k-1} + w_{2k+1}$.

Using procedure from a-ii, stationary mean of z_k is 0. Stationary variance is $E(z_k^2) = a_2^2 E(z_{k-1}^2) + E(w_{2k}^2)$, solving gives $E(z_k^2) = \sigma^2/(1 - a_2^2)$. Same mean and variance for y_k. So

$$p(x_0, x_1) = p(z_0)p(y_0)$$

gives final answer.

Q3b-ii. Assume $n = 2m$. Using this,

$$p(x_2, \ldots, x_n|x_0, x_1) = p(z_1, \ldots, z_m|z_0)p(y_1, \ldots, y_{m-1}|y_0).$$

Now solve for the MLE.

$$\log p(z_1, \ldots, z_m|z_0) = m \log \frac{1}{\sqrt{2\pi}\sigma} - \frac{1}{2\sigma^2}\sum_{i=1}^{m}(z_i - a_2 z_{i-1})^2$$

$$\log p(y_1, \ldots, y_{m-1}|y_0) = (m-1) \log \frac{1}{\sqrt{2\pi}\sigma} - \frac{1}{2\sigma^2}\sum_{i=1}^{m-1}(y_i - a_2 y_{i-1})^2$$

Fix σ^2, differentiate and solve for a_2 first. Then differentiate and solve for σ^2 using found a_2. This gives

$$a_2 = \left(\sum_{i=1}^{m} z_i z_{i-1} + \sum_{i=1}^{m-1} y_i y_{i-1}\right) \Big/ \left(\sum_{i=1}^{m} z_{i-1}^2 + \sum_{i=1}^{m-1} y_{i-1}^2\right)$$

Let $L = \sum_{i=1}^{m}(z_i - a_2 z_{i-1})^2 + \sum_{i=1}^{m-1}(y_i - a_2 y_{i-1})^2$ at found a_2. Solving for σ^2 yields

$$\sigma^2 = \frac{L}{n+1-2}.$$

Q3b-iii: Bookwork.

Examiner: Very well answered question although only attempted by roughly 50% of candidates. Many did not know the definition of strict stationarity.

Part (b)-ii was not solved correctly, it was the stationary distribution that was needed here. Part (b)-iii was well answered with candidates reverse engineering the Yule-Walker solution from the obtained MLE one as opposed to having to remember it.

Q4a. Follow procedure in lecture notes.

Q4b-i. Since x_n is white Gaussian noise, the periodograms are independent of each other.

$$\mathrm{var}\left(\hat{S}^{(k)}\right) = (1-\gamma_k)^2 \mathrm{var}\left(\hat{S}^{(k-1)}\right) + \gamma_k^2 \mathrm{var}\left(\hat{P}^{(k)}\right)$$

where $\hat{P}^{(k)} = |\cdots|^2/N$ is the periodogram of frame k. When $\gamma_k = 1/k$, we are computing the sample average $\hat{P}^{(1)}, \hat{P}^{(2)}, \ldots, \hat{P}^{(k)}$. Note that $\mathrm{var}\left(\hat{P}^{(k)}\right)$ is not a function of k, call this value p_N where N indicates frame length. So $\mathrm{var}\left(\hat{S}^{(k)}\right) = p_N/k$.

When $\gamma_k = a$, we have a geometric series. Note that the variance becomes $\sum_{n=1}^{k} p_N a^2 \left[(1-a)^2\right]^{k-n} = p_N a^2 \sum_{i=0}^{k-1} \left[(1-a)^2\right]^i$.

Q4b-ii. The sample means limiting variance is (trivially) 0. The geometric series case is $p_N a^2/\left(1-(1-a)^2\right) = p_N a/(2-a)$.

Q4c. First note that $E(x_n^2) = 2\sigma^2$, $E(x_n x_{n+1}) = \sigma^2$ and other $E(x_n x_{n+k}) = 0$ for other values of k. Also,

$$\sum_{n=0}^{N-1}\sum_{m=0}^{N-1} x_n x_m = \sum_{n=0}^{N-1} x_n^2 + \sum_{n=0}^{N-2} x_n x_{n+1} + \sum_{n=1}^{N-1} x_n x_{n-1} + R$$

where $E(R) = 0$ owing to the autocorrelation of x_n. So mean value is

$$\frac{1}{N}\left(N2\sigma^2 + (N-1)2\sigma^2\right).$$

Examiner: Attempted by roughly 50% of candidates. Part (a) was not very well answered. A properly executed sketch of the derivation would have been sufficient. Part (b) was also not well answered and candidates failed to calculate the variance in the most straightforward manner as the crib details. This had a knock on effect to the remaining parts. Part (c) was meant to be an easy mark earner but many did not capitalise on it.

Sumeetpal S. Singh