EN 257: Applied Stochastic Processes
Problem Set 7

Douglas Lanman
dlanman@brown.edu
3 May 2007

Problem 6.25

Consider a wide sense stationary random sequence $X[n]$ input to a linear filter with impulse response

$$h[n] = \begin{cases} 1/2, & n = \{0,1\} \\ 0, & \text{otherwise.} \end{cases} \tag{1}$$

Write the PSD of the output sequence $S_{YY}(\omega)$ in terms of the PSD of the input sequence $S_{XX}(\omega)$.

(a) Show that the PSD is real-valued, even if $X[n]$ is a complex-valued random sequence.

(b) Show that if $X[n]$ is real-valued, then $S_{XX}(\omega) = S_{XX}(-\omega)$.

(c) Show that $S_{XX}(\omega) \geq 0$ for every ω, regardless of whether $X[n]$ is complex-valued or not.

Let's begin by determining the PSD of the output sequence $S_{YY}(\omega)$ in terms of the PSD of the input sequence $S_{XX}(\omega)$. First, we recall that the autocorrelation of the output sequence $R_{YY}[m]$ is given by Equation 6.4-1 on page 350 in [5].

$$R_{YY}[m] = g[m] * R_{XX}[m], \text{ for } g[m] \triangleq h[m] * h^*[-m]$$

Note that $g[m]$, the *autocorrelation impulse response*, is given by

$$g[m] = h[m] * h^*[-m] = \sum_{k=-\infty}^{\infty} h[k]h^*[m-k] = \begin{cases} 1/4, & m = \{0,2\} \\ 1/2, & m = 1 \\ 0, & \text{otherwise.} \end{cases}$$

Also recall that the power spectral density (PSD) of the input sequence $S_{XX}(\omega)$ is defined as the discrete-time Fourier transform (DFT) of the input autocorrelation function $R_{XX}[m]$.

$$S_{XX}(\omega) \triangleq \sum_{m=-\infty}^{\infty} R_{XX}[m]e^{-j\omega m}, \text{ for } -\pi \leq \omega \leq \pi \tag{2}$$

Similarly, the PSD of the output sequence $S_{YY}(\omega)$ is given by the DFT of $R_{YY}(\omega)$.

$$S_{YY}(\omega) \triangleq \sum_{m=-\infty}^{\infty} R_{YY}[m]e^{-j\omega m}, \text{ for } -\pi \leq \omega \leq \pi$$

Substituting the previous expression for $R_{YY}[m]$, we find

$$S_{YY}(\omega) = \sum_{m=-\infty}^{\infty} (h[k] * h^*[-m]) * R_{XX}[m]e^{-j\omega m}, \text{ for } -\pi \leq \omega \leq \pi.$$

From the derivation of Equation 6.4-2b on page 351, we conclude that the general form of the output PSD is given by the following expression.

$$S_{YY}(\omega) = |H(\omega)|^2 S_{XX}(\omega), \text{ for } -\pi \le \omega \le \pi \tag{3}$$

At this point we can evaluate the DFT of the impulse response, as defined in Equation 1.

$$H(\omega) = \sum_{m=-\infty}^{\infty} h[m]e^{-j\omega m} = \frac{1}{2}\left(1 - e^{-j\omega}\right)$$

The magnitude of $H(\omega)$ is then given by

$$|H(\omega)|^2 = \frac{1}{4}\left(1 - e^{-j\omega}\right)\left(1 - e^{j\omega}\right) = \cos^2(\omega/2).$$

Substituting this result in Equation 3 gives the desired expression for the PSD of the output.

$$\boxed{S_{YY}(\omega) = \cos^2(\omega/2)\, S_{XX}(\omega), \text{ for } -\pi \le \omega \le \pi}$$

Part (a)

To prove that the PSD $S_{XX}(\omega)$ is a real-valued function, we begin by proving that the autocorrelation function is conjugate symmetric, such that $R_{XX}[m] = R^*_{XX}[-m]$. From the definition of the autocorrelation function $R_{XX}[m]$, we have

$$R_{XX}[m] = E\{X[k+m]X^*[k]\} = E\{X[k]X^*[k-m]\} = E^*\{X[k-m]X^*[k]\} = R^*_{XX}[-m]. \tag{4}$$

If $S_{XX}(\omega)$ is real-valued, then it must satisfy the following condition.

$$S_{XX}(\omega) = S^*_{XX}(\omega)$$

Substituting Equation 2, we must show that the following equality holds.

$$\begin{aligned}
\sum_{m=-\infty}^{\infty} R_{XX}[m]e^{-j\omega m} &= \left(\sum_{m=-\infty}^{\infty} R_{XX}[m]e^{-j\omega m}\right)^* \\
&= \sum_{m=-\infty}^{\infty} R^*_{XX}[m]e^{j\omega m} \\
&= \sum_{m=-\infty}^{\infty} R_{XX}[-m]e^{j\omega m} \\
&= \sum_{m=-\infty}^{\infty} R_{XX}[m]e^{-j\omega m}
\end{aligned}$$

Note that in the previous expression we have applied the result of Equation 4 to conclude that $R^*_{XX}[m] = R_{XX}[-m]$. In addition, we have substituted $-m$ for m since the order of summation can be reversed. In conclusion, since the left-hand and right-hand sides of the previous expression are equal, we find that $S_{XX}(\omega)$ is a real-valued function, even if $X[n]$ is a complex-valued sequence.

$$\therefore \boxed{S_{XX}(\omega) = S^*_{XX}(\omega) \Rightarrow S_{XX}(\omega) \in \mathbb{R}, \text{ for } -\pi \le \omega \le \pi}$$

Part (b)

Let's begin by using Equation 2 to express $S_{XX}(-\omega)$.

$$S_{XX}(-\omega) = \sum_{m=-\infty}^{\infty} R_{XX}[m]e^{j\omega m} = \sum_{m=-\infty}^{\infty} R_{XX}[-m]e^{-j\omega m}$$

Note that, since the order of summation can be reversed, we have substituted $-m$ for m in the right-hand expression. At this point we can apply the result of Equation 4 to conclude $R_{XX}[-m] = R_{XX}^*[m]$. Now recall that the autocorrelation function $R_{XX}[m]$ will be real-valued if $X[n]$ is real-valued, such that

$$R_{XX}^*[m] = E^*\{X[k+m]X^*[k]\} = E\{X[k+m]X[k]\} = R_{XX}[m], \text{ for } \{X[n]\} \in \mathbb{R}.$$

We conclude that $R_{XX}[-m] = R_{XX}[m]$ if $X[n]$ is real-valued and, substituting into the previous expression, we find that $S_{XX}(\omega)$ is an even function if $X[n]$ is real-valued.

$$\therefore \boxed{S_{XX}(\omega) = S_{XX}(-\omega), \text{ for } \{X[n]\} \in \mathbb{R}}$$

Part (c)

Recall, from page 351 in [5], that the inverse Fourier transform of the PSD $S_{XX}(\omega)$ is equal to the autocorrelation function $R_{XX}[m]$.

$$R_{XX}[m] = \frac{1}{2\pi}\int_{-\pi}^{\pi} S_{XX}(\omega)e^{j\omega m}d\omega$$

Applying Equation 3, we find that the output autocorrelation function is given by

$$R_{YY}[m] = \frac{1}{2\pi}\int_{-\pi}^{\pi} |H(\omega)|^2 S_{XX}(\omega)e^{j\omega m}d\omega.$$

At this point, consider a narrowband low-pass filter $H(\omega)$, with bandwidth 2ϵ, centered at frequency ω_o, where $|\omega_0| < \pi$, and with unity gain in the passband. Substituting into the previous expression and evaluating at $m = 0$, we find

$$R_{YY}[0] = \frac{1}{2\pi}\int_{\omega_o-\epsilon}^{\omega_o+\epsilon} S_{XX}(\omega)d\omega \simeq \frac{\epsilon}{\pi}S_{XX}(\omega_o), \text{ for } -\pi \leq \omega_o \leq \pi.$$

At this point we recall that $R_{YY}[0]$ is a non-negative function, since

$$R_{YY}[0] = E\{Y[k]Y^*[k]\} = E\{|Y[k]|^2\} \geq 0.$$

Substituting this result into the previous expression, we find that $S_{XX}(\omega) \geq 0$ for every ω, regardless of whether $X[n]$ is complex-valued or not.

$$\therefore \boxed{S_{XX}(\omega) \geq 0, \text{ for } -\pi \leq \omega \leq \pi}$$

Problem 6.26

Let the WSS random sequence X have the correlation function

$$R_{XX}[m] = 10e^{-\lambda_1|m|} + 5e^{-\lambda_2|m|}$$

with $\lambda_1 > 0$ and $\lambda_2 > 0$. Find the corresponding PSD $S_{XX}(\omega)$ for $|\omega| \leq \pi$.

Substituting into Equation 2, we find that the PSD $S_{XX}(\omega)$ is given by the following expression.

$$\begin{aligned}
S_{XX}(\omega) &= \sum_{m=-\infty}^{\infty} R_{XX}[m]e^{-j\omega m} \\
&= \sum_{m=-\infty}^{\infty} \left(10e^{-\lambda_1|m|} + 5e^{-\lambda_2|m|}\right) e^{-j\omega m} \\
&= \left(10 \sum_{m=-\infty}^{\infty} e^{-\lambda_1|m|}e^{-j\omega m}\right) + \left(5 \sum_{m=-\infty}^{\infty} e^{-\lambda_2|m|}e^{-j\omega m}\right)
\end{aligned}$$

As an aside, we note that the first term in the previous expression can be further simplified by separating the summation into several components.

$$\begin{aligned}
10 \sum_{m=-\infty}^{\infty} e^{-\lambda_1|m|}e^{-j\omega m} &= 10 + 10 \sum_{m=-\infty}^{-1} e^{-\lambda_1|m|}e^{-j\omega m} + 10 \sum_{m=1}^{\infty} e^{-\lambda_1|m|}e^{-j\omega m} \\
&= 10 + 10 \sum_{m=1}^{\infty} e^{-\lambda_1|m|}e^{j\omega m} + 10 \sum_{m=1}^{\infty} e^{-\lambda_1|m|}e^{-j\omega m} \\
&= 10 + 10 \sum_{m=1}^{\infty} e^{-\lambda_1|m|} \left(e^{j\omega m} + e^{-j\omega m}\right) \\
&= 10 + 20 \sum_{m=1}^{\infty} e^{-\lambda_1 m} \cos(\omega m)
\end{aligned}$$

Substituting this result into the previous expression yields the following solution for the PSD $S_{XX}(\omega)$ for $\omega \leq \pi$.

$$\boxed{S_{XX}(\omega) = 15 + \sum_{m=1}^{\infty} \left(20e^{-\lambda_1 m} + 10e^{-\lambda_2 m}\right) \cos(\omega m), \text{ for } |\omega| \leq \pi}$$

Applying additional trigonometric identities, we find that this expression can be further simplified.

$$S_{XX}(\omega) = \frac{10\sinh(\lambda_1)}{\cosh(\lambda_1) - \cos(\omega)} + \frac{5\sinh(\lambda_2)}{\cosh(\lambda_2) - \cos(\omega)}, \text{ for } |\omega| \leq \pi$$

Problem 6.29

Consider the LTI system show in Figure P6.29 on page 395 in [5]. Let $X[n]$ and $V[n]$ be WSS and mutually uncorrelated with zero-mean and PSD's $S_{XX}(\omega)$ and $S_{VV}(\omega)$, respectively.

(a) Find the PSD of the output $S_{YY}(\omega)$.

(b) Find the cross-power spectral density $S_{XY}(\omega)$ between the input X and output Y.

Part (a)

Recall from Equation 2 that the PSD of the input sequence $S_{XX}(\omega)$ is defined as the discrete-time Fourier transform (DFT) of the input autocorrelation function $R_{XX}[m]$.

$$S_{XX}(\omega) \triangleq \sum_{m=-\infty}^{\infty} R_{XX}[m]e^{-j\omega m}, \text{ for } -\pi \leq \omega \leq \pi$$

Similarly, the PSD of the output sequence $S_{YY}(\omega)$ is given by the DFT of $R_{YY}(\omega)$.

$$S_{YY}(\omega) \triangleq \sum_{m=-\infty}^{\infty} R_{YY}[m]e^{-j\omega m}, \text{ for } -\pi \leq \omega \leq \pi$$

In order to proceed, we must first find an expression for $R_{YY}[m]$ in terms of $R_{XX}[m]$ and $R_{VV}[m]$. Let us define

$$Z[n] \triangleq X[n] + V[n]$$

as the input to the LTI system with impulse response $h[n]$. By the definition of the autocorrelation function, we must have the following condition (see Table 7.5-1 on page 444 in [5]).

$$\begin{aligned}
R_{ZZ}[m] &= E\{Z[k+m]Z^*[k]\} \\
&= E\{(X[k+m]+V[k+m])(X[k]+V[k])^*\} \\
&= E\{X[k+m]X^*[k]\} + E\{X[k+m]V^*[k]\} + E\{V[k+m]X^*[k]\} + E\{V[k+m]V^*[k]\} \\
&= E\{X[k+m]X^*[k]\} + E\{V[k+m]V^*[k]\} \\
&= R_{XX}[m] + R_{VV}[m] \qquad (5)
\end{aligned}$$

Note that, since $X[n]$ and $V[n]$ are mutually uncorrelated, we can conclude that $E\{X[k+m]V^*[k]\} = E\{V[k+m]X^*[k]\} = 0$. Now recall the expression for S_{YY} previously derived in Problem 6.25.

$$S_{YY}(\omega) = \sum_{m=-\infty}^{\infty} (h[k] * h^*[-m]) * R_{ZZ}[m]e^{-j\omega m}, \text{ for } -\pi \leq \omega \leq \pi$$

Substituting Equation 5, we find that

$$S_{YY}(\omega) = \sum_{m=-\infty}^{\infty} (h[k] * h^*[-m]) * R_{XX}[m]e^{-j\omega m} + \sum_{m=-\infty}^{\infty} (h[k] * h^*[-m]) * R_{VV}[m]e^{-j\omega m}.$$

Finally, we recall the following properties of the DFT: (1) the DFT of the convolution of two functions is the product of their discrete-time Fourier transforms, and (2) the DFT of the complex conjugate of a time-reversed function is equal to the complex conjugate of the Fourier transform of the original function [3]. In addition, we remember from Equation 2 that the autocorrelation function and the power spectral density are Fourier transform pairs. Applying these conditions to the previous expression gives the following result for the PSD of the output $S_{YY}(\omega)$.

$$\boxed{S_{YY}(\omega) = |H(\omega)|^2 S_{XX}(\omega) + |H(\omega)|^2 S_{VV}(\omega), \text{ for } -\pi \le \omega \le \pi}$$

Part (b)

Similar to the previous part, we begin by recalling the definition of the cross-power spectral density $S_{XY}(\omega)$ given on page 352 in [5].

$$S_{XY}(\omega) \triangleq \sum_{m=-\infty}^{\infty} R_{XY}[m] e^{-j\omega m}, \text{ for } -\pi \le \omega \le \pi \tag{6}$$

In order to evaluate $S_{XY}(\omega)$, we require a closed-form expression for the cross-correlation function $R_{XY}(\omega)$. Following the derivation on page 349, we find the following result.

$$\begin{aligned}
R_{XY}[m,n] &= E\{X[m]Y^*[n]\} \\
&= \sum_{k=-\infty}^{\infty} h^*[n-k] E\{X[m]Z^*[k]\} \\
&= \sum_{k=-\infty}^{\infty} h^*[n-k] E\{X[m](X^*[k] + V^*[k])\} \\
&= \sum_{k=-\infty}^{\infty} h^*[n-k] E\{X[m]X^*[k]\} + \sum_{k=-\infty}^{\infty} h^*[n-k] E\{X[m]V^*[k]\} \\
&= \sum_{k=-\infty}^{\infty} h^*[n-k] R_{XX}[m-k] \\
&= \sum_{k=-\infty}^{\infty} h^*[-l] R_{XX}[(m-n)-l], \text{ for } l \triangleq k-n
\end{aligned}$$

At this point we note that the cross-correlation $R_{XY}[m,n]$ is shift-invariant. As a result, we can define $R_{XY}[m] \triangleq R_{XY}[m,0]$. Substituting this result into the previous expression yields

$$R_{XY}[m] = \sum_{k=-\infty}^{\infty} h^*[-l] R_{XX}[m-l] = h^*[-m] * R_{XX}[m].$$

Now we can evaluate Equation 6 to find the desired expression for the cross-power spectral density.

$$S_{XY}(\omega) = \sum_{m=-\infty}^{\infty} h^*[-m] * R_{XX}[m] e^{-j\omega m}, \text{ for } -\pi \le \omega \le \pi$$

Applying the previous properties of the DFT used in Part (a), we conclude that $S_{XY}(\omega)$ is given by the following expression.

$$\boxed{S_{XY}(\omega) = H^*(\omega) S_{XX}(\omega), \text{ for } -\pi \le \omega \le \pi}$$

Problem 6.32

Recall that a Markov random sequence $X[n]$ is specified by its first-order pdf $f_X(x;n)$ and its one-step conditional pdf

$$f_X(x_n|x_{n-1};n,n-1) = f_X(x_n|x_{n-1}).$$

(a) Find the two-step pdf $f_X(x_n|x_{n-2})$ for a Markov random sequence in terms of the above functions. For this problem, take $n \geq 2$ for a sequence starting at $n = 0$.

(b) Find the N-step pdf $f_X(x_n|x_{n-N})$ for an arbitrary positive integer N, where $n \geq N$.

Part (a)

Recall from pages 429-430 in [5] that the Chapman-Kolmogorov equations can be used to compute the conditional density of $X[n_3]$ given $X[n_1]$, for $n_3 > n_2 > n_1$. To begin our analysis, we note that the joint pdf can be written as follows.

$$f_X(x_3,x_1;n_3,n_1) = \int_{-\infty}^{\infty} f_X(x_3|x_2,x_1;n_3,n_2,n_1) f_X(x_2,x_1;n_2,n_1)dx_2$$

Dividing both sides of this expression by $f(x_1;n_1)$ yields the following result.

$$f_X(x_3|x_1;n_3,n_1) = \int_{-\infty}^{\infty} f_X(x_3|x_2,x_1;n_3,n_2,n_1) f_X(x_2|x_1;n_2,n_1)dx_2$$

Finally, by applying the Markov property we arrive at the well-known Chapman-Kolmogorov equation for the transition density $f_X(x_3|x_1)$.

$$f_X(x_3|x_1;n_3,n_1) = \int_{-\infty}^{\infty} f_X(x_3|x_2;n_3,n_2) f_X(x_2|x_1;n_2,n_1)dx_2, \text{ for } n_3 > n_2 > n_1 \tag{7}$$

From this condition we can conclude that the two-step pdf $f_X(x_n|x_{n-2})$ for a Markov random sequence is given by the following expression (using the simplified notation in the original problem statement).

$$\boxed{f_X(x_n|x_{n-2}) = \int_{-\infty}^{\infty} f_X(x_n|x_{n-1}) f_X(x_{n-1}|x_{n-2})dx_{n-1}}$$

Part (b)

The N-step pdf $f_X(x_n|x_{n-N})$, for an arbitrary positive integer N, can be found in a similar manner as the previous problem by repeatedly using conditioning (i.e., the chain rule of probability [5]).

$$\boxed{f_X(x_n|x_{n-N}) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f_X(x_n|x_{n-1}) f_X(x_{n-1}|x_{n-2}) \ldots f_X(x_{n-N+1}|x_{n-N}) dx_{n-1} dx_{n-2} \ldots dx_{n-N+1}}$$

Problem 6.33

Consider a Markov random sequence $X[n]$ on $1 \leq n \leq N$ that is statistically described by its first-order pdf $f_X(x;1)$ and its one-step transition (conditional) pdf $f_X(x_n|x_{n-1};n,n-1)$. By the Markov definition and suppressing time variables, we have

$$f_X(x_n|x_{n-1}) = f_X(x_n|x_{n-1},x_{n-2},\ldots,x_1) \text{ for } 2 \leq n \leq N.$$

Show that such a Markov sequence is also Markov in the reverse order, such that

$$f_X(x_n|x_{n+1}) = f_X(x_n|x_{n+1},x_{n+2},\ldots,x_N) \text{ for } 1 \leq n \leq N-1$$

and, as a result, one can alternatively describe a Markov random sequence by its one-step backward pdf $f_X(x_{n-1}|x_n;n-1,n)$ and its first-order pdf $f_X(x;N)$.

Recall, from Equation 2.6-52 on page 104 in [5], that the following expression relates the conditional pdfs for two random variables X and Y.

$$f_{X|Y}(x|y) = \frac{f_{Y|X}(y|x)f_X(x)}{f_Y(y)} \tag{8}$$

Furthermore, we recall that the chain rule of probability and the Markov property can be used to express the joint pdf as follows.

$$f_X(x_1,x_2,\ldots,x_{N-1},x_N) = f_X(x_N|x_{N-1})f_X(x_{N-1}|x_{N-2})\ldots f_X(x_2|x_1)f_X(x_1)$$

From Equation 8 we observe that the following condition can be used to relate the conditional probability density functions.

$$f_X(x_n|x_{n-1}) = \frac{f_X(x_{n-1}|x_n)f_X(x_n)}{f_X(x_{n-1})}$$

Substituting this result into the previous expression yields the following equation.

$$f_X(x_1,x_2,\ldots,x_{N-1},x_N) = \left(\frac{f_X(x_{N-1}|x_N)f_X(x_N)}{f_X(x_{N-1})}\right)\left(\frac{f_X(x_{N-2}|x_{N-1})f_X(x_{N-1})}{f_X(x_{N-2})}\right)\cdots\left(\frac{f_X(x_1|x_2)f_X(x_2)}{f_X(x_1)}\right)f_X(x_1)$$

Simplifying this expression yields the following result.

$$f_X(x_1,x_2,\ldots,x_{N-1},x_N) = f_X(x_1|x_2)f_X(x_2|x_3)\ldots f_X(x_{N-1}|x_N)f_X(x_N)$$

Note that this expression has the general form of a Markov sequence in reverse order. As a result, we conclude that a Markov sequence is also Markov in the reverse order, such that

$$f_X(x_n|x_{n+1}) = f_X(x_n|x_{n+1},x_{n+2},\ldots,x_N) \text{ for } 1 \leq n \leq N-1$$

and, as a result, one can alternatively describe a Markov random sequence by its one-step backward pdf $f_X(x_{n-1}|x_n;n-1,n)$ and its first-order pdf $f_X(x;N)$.

Problem 6.34

Let $X[n]$ be a Markov chain on $n \geq 1$ with transition probabilities given by $P(x[n]|x[n-1])$.

(a) Find an expression for the two-step transition probabilities $P(x[n]|x[n-2])$.

(b) Show that $P(x[n+1]|x[n-1], x[n-2], \ldots, x[1]) = P(x[n+1]|x[n-1])$, for $n \geq 1$.

Part (a)

Similar to Problem 6.32 we can apply the Chapman-Kolmogorov equations to compute the conditional probability of $X[n_3]$ given $X[n_1]$, for $n_3 > n_2 > n_1$. To begin our analysis, we note that the joint probability can be written as

$$P(x[n_3], x[n_1]) = \sum_{x[n_2] \in \mathcal{X}_2} P(x[n_3]|x[n_2], x[n_1]) P(x[n_2], x[n_1]),$$

where $\mathcal{X}_2$ denotes all possible values of the discrete random variable $X[n_2]$. Dividing both sides of this expression by $P(x[n_1])$ yields the following result.

$$P(x[n_3]|x[n_1]) = \sum_{x[n_2] \in \mathcal{X}_2} P(x[n_3]|x[n_2], x[n_1]) P(x[n_2]|x[n_1])$$

Finally, by applying the Markov property we arrive at the well-known Chapman-Kolmogorov equation for the transition probability $P(x[n_3]|x[n_1])$.

$$P(x[n_3]|x[n_1]) = \sum_{x[n_2] \in \mathcal{X}_2} P(x[n_3]|x[n_2]) P(x[n_2]|x[n_1])$$

From this condition we conclude that the two-step transition probabilities $P(x[n]|x[n-2])$ are given by the following expression.

$$\boxed{P(x[n]|x[n-2]) = \sum_{x[n-1] \in \mathcal{X}_{n-1}} P(x[n]|x[n-1]) P(x[n-1]|x[n-2])}$$

Part (b)

From the previous derivation we conclude that the following condition must hold for $n \geq 1$.

$$P(x[n+1]|x[n-1], \ldots, x[1]) = \sum_{x[n] \in \mathcal{X}_n} P(x[n+1]|x[n], \ldots, x[1]) P(x[n]|x[n-1], \ldots, x[1])$$

By the Markov property this expression has the following simplified form.

$$P(x[n+1]|x[n-1], \ldots, x[1]) = \sum_{x[n] \in \mathcal{X}_n} P(x[n+1]|x[n]) P(x[n]|x[n-1])$$

By inspection, we conclude that this result is identical to that found in Part (a). As a result, we conclude that the desired relationship holds.

$$\boxed{P(x[n+1]|x[n-1], x[n-2], \ldots, x[1]) = P(x[n+1]|x[n-1]), \text{ for } n \geq 1}$$

Problem 7.40

Express the answers to the following questions in terms of probability density functions.

(a) State the definition of an independent-increments random process.

(b) State the definition of a Markov random process.

(c) Prove that any random process that has independent increments also has the Markov property.

Part (a)

Recall, from pages 326 and 410 in [5], that a random process has independent increments when the set of n random variables

$$X(t_1), X(t_2) - X(t_1), \ldots, X(t_n) - X(t_{n-1})$$

are jointly independent for all $t_1 < t_2 < \ldots < t_n$ and $n \geq 1$. In terms of probability density functions, we note that a random process $X(t)$ with independent increments must satisfy

$$f_X(x_n - x_{n-1}|x_{n-1}, x_{n-2}, \ldots, x_1; t_n, t_{n-1}, \ldots, t_1) = f_X(x_n - x_{n-1}; t_n, t_{n-1}), \tag{9}$$

for all $x_1, x_2, \ldots, x_n$ and integers $n > 0$ where $t_1 < t_2 < \ldots < t_n$.

Part (b)

From page 422 in [5] we recall that a (first-order) continuous-valued Markov process $X(t)$ satisfies

$$f_X(x_n|x_{n-1}, x_{n-2}, \ldots, x_1; t_n, t_{n-1}, \ldots, t_1) = f_X(x_n|x_{n-1}; t_n, t_{n-1}), \tag{10}$$

for all $x_1, x_2, \ldots, x_n$ and integers $n > 0$ where $t_1 < t_2 < \ldots < t_n$. Note that this expression defines the so-called one-step conditional pdf, however a continuous-valued Markov process must also satisfy the following k-step pdf given by

$$f_X(x_{n+k}|x_n, x_{n-1}, \ldots, x_1; t_{n+k}, t_n, t_{n-1}, \ldots, t_1) = f_X(x_{n+k}|x_n; t_{n+k}, t_n),$$

for all $x_1, x_2, \ldots, x_n, x_{n+k}$ and integers $n > 0$ and $k > 0$ where $t_1 < t_2 < \ldots < t_n < t_{n+k}$.

Part (c)

Following the derivation outlined on page 423 in [5], we note that any random process $X(t)$ with independent increments must have a pdf that can be expressed in the following form.

$$\begin{aligned} f_X(x_n|x_{n-1}, x_{n-2}, \ldots, x_1; t_n, t_{n-1}, \ldots, t_1) &= f_X(x_n - x_{n-1}|x_{n-1}, x_{n-2}, \ldots, x_1; t_n, t_{n-1}, \ldots, t_1) \\ &= f_X(x_n - x_{n-1}; t_n, t_{n-1}) \\ &= f_X(x_n - x_{n-1}|x_{n-1}; t_n, t_{n-1}) \\ &= f_X(x_n|x_{n-1}; t_n, t_{n-1}) \end{aligned}$$

Note that, on the first two lines of this expression, we have applied the definition of independent increments given by Equation 9. Since the last line is identical to the definition of Markov random processes given by Equation 10, we conclude that any random process that has independent increments also has the Markov property.

Problem 7.46

Consider the linear system shown in Figure P7.46 on page 485 in [5], which is excited by two *orthogonal* zero-mean, jointly wide-sense stationary random processes $X(t)$, the signal, and $U(t)$, the noise. Let the input to the system G be

$$Y(t) = h(t) * X(t) + U(t),$$

which models a distorted-signal-in-noise estimation problem. If we pass the received signal $Y(t)$ through the filter G, we obtain an estimate $\hat{X}(t)$. Finally, we define the estimation error $\varepsilon(t)$ such that

$$\varepsilon(t) = \hat{X}(t) - X(t).$$

In the following problems we will evaluate some relevant power and cross-power spectral densities.

(a) Find $S_{YY}(\omega)$.

(b) Find $S_{\hat{X}X}(\omega) = S^*_{X\hat{X}}(\omega)$ in terms of $H(\omega)$, $G(\omega)$, $S_{XX}(\omega)$, and $S_{UU}(\omega)$.

(c) Find $S_{\varepsilon\varepsilon}(\omega)$.

(d) Show that, in order to minimize $S_{\varepsilon\varepsilon}(\omega)$ at frequencies where $S_{XX}(\omega) >> S_{UU}(\omega)$, we should select $G \approx H^{-1}$. Similarly, where $S_{XX}(\omega) << S_{UU}(\omega)$, we should have $G \approx 0$.

Part (a)

To begin our analysis we recall that the power spectral density $S_{XX}(\omega)$, for a continuous-valued wide-sense stationary random process $X(t)$, is given on page 443 in [5].

$$S_{XX}(\omega) \triangleq \int_{-\infty}^{\infty} R_{XX}(\tau) e^{-j\omega\tau} d\tau$$

Similar to Part (a) of Problem 6.29 and as given by Equation7.5-14, we recall that a LTI system with frequency response $H(\omega)$ has the following output PSD $S_{YY}(\omega)$ for the input process $X(t)$.

$$S_{YY}(\omega) = \int_{-\infty}^{\infty} R_{YY}(\tau) e^{-j\omega\tau} d\tau = \int_{-\infty}^{\infty} h(\tau) * R_{XX}(\tau) * h^*(-\tau) e^{-j\omega\tau} d\tau = |H(\omega)|^2 S_{XX}(\omega)$$

Next, we observe that $h(t) * X(t)$ and $U(t)$ are orthogonal since

$$\begin{aligned} E\left\{[h(t) * X(t_1)]\, Y^*(t_2)\right\} &= E\left\{\int_{-\infty}^{\infty} h(\tau) X(t_1 - \tau) Y^*(t_2) d\tau\right\} \\ &= \int_{-\infty}^{\infty} h(\tau) E\left\{X(t_1 - \tau) Y^*(t_2)\right\} d\tau = 0 \end{aligned}$$

and $E\{X(t_1 - \tau)Y^*(t_2)\} = 0$ for the orthogonal random processes $X(t)$ and $Y(t)$ (see page 437). Finally, we recall from Table 7.5-1 that the PSD of two orthogonal random process $X_1(t)$ and $X_2(t)$ is given by $S_{X_1X_1}(\omega) + S_{X_2X_2}(\omega)$. In conclusion $S_{YY}(\omega)$ is given by the following expression.

$$\boxed{S_{YY}(\omega) = |H(\omega)|^2 S_{XX}(\omega) + S_{UU}(\omega)}$$

Part (b)

From the problem statement, we have

$$\hat{X}(t) = g(t) * [h(t) * X(t) + U(t)] = g(t) * h(t) * X(t) + g(t) * U(t). \tag{11}$$

Now we recall that the cross-power spectral density $S_{\hat{X}X}(\omega)$ is given by the following expression.

$$S_{\hat{X}X}(\omega) = \int_{-\infty}^{\infty} R_{\hat{X}X}(\tau) e^{-j\omega\tau} d\tau$$

To proceed we need to obtain a closed-form expression for the cross-correlation $R_{\hat{X}X}(\omega)$. By definition, we have

$$\begin{aligned}
R_{\hat{X}X}(\tau) &= E\left\{\hat{X}(t+\tau)X^*(t)\right\} \\
&= E\left\{[g(t+\tau) * h(t+\tau) * X(t+\tau) + g(t+\tau) * U(t+\tau)]\, X^*(t)\right\} \\
&= E\left\{[g(t+\tau) * h(t+\tau) * X(t+\tau)]\, X^*(t)\right\} + E\left\{[g(t+\tau) * U(t+\tau)]\, X^*(t)\right\} \\
&= g(t+\tau) * h(t+\tau) * E\left\{X(t+\tau)X^*(t)\right\} \\
&= g(\tau) * h(\tau) * R_{XX}(\tau),
\end{aligned} \tag{12}$$

since, by definition, $X(t)$ and $U(t)$ are orthogonal random processes such that $R_{UX}(t_1, t_2) = E\{U(t_1)X^*(t_2)\} = 0$ for all t_1 and t_2. Substituting Equation 12 into Equation 11 yields the following solution for $S_{\hat{X}X}(\omega)$.

$$\boxed{S_{\hat{X}X}(\omega) = H(\omega)G(\omega)S_{XX}(\omega)}$$

Part (c)

From the problem statement and Equation 11, we find

$$\varepsilon(t) = \hat{X}(t) - X(t) = [g(t) * h(t) - \delta(t)] * X(t) + g(t) * U(t).$$

Correspondingly, the power spectral density of the estimation error is given by

$$S_{\varepsilon\varepsilon}(\omega) = \int_{-\infty}^{\infty} R_{\varepsilon\varepsilon}(\tau) e^{-j\omega\tau} d\tau.$$

Since the Fourier transform of the Dirac delta function $\delta(t)$ is equal to unity, we conclude that the power spectral density of the estimation error has the following solution.

$$\boxed{S_{\varepsilon\varepsilon}(\omega) = |G(\omega)H(\omega) - 1|^2 S_{XX}(\omega) + |G(\omega)|^2 S_{UU}(\omega)} \tag{13}$$

Part (d)

From Equation 13 we find that, in order to minimize $S_{\varepsilon\varepsilon}(\omega)$ for $S_{XX}(\omega) >> S_{UU}(\omega)$, we must select $G(\omega) \approx [H(\omega)]^{-1}$. Similarly, $G \approx 0$ minimizes $S_{\varepsilon\varepsilon}(\omega)$ for $S_{XX}(\omega) << S_{UU}(\omega)$. In summary, the following conditions on $G(\omega)$ will minimize the power spectral density of the estimation error.

$$\boxed{\begin{gathered}
S_{\varepsilon\varepsilon}(\omega) = \left|[H(\omega)]^{-1} H(\omega) - 1\right|^2 S_{XX}(\omega) + |H(\omega)|^{-2} S_{UU}(\omega) \approx 0, \text{ for } S_{XX}(\omega) >> S_{UU}(\omega) \\
S_{\varepsilon\varepsilon}(\omega) = S_{XX}(\omega) \approx 0, \text{ for } S_{XX} << S_{UU}(\omega)
\end{gathered}}$$

Problem 7.47

Let $X(t)$, the input to the system in Figure P7.47 on page 486 in [5], be a stationary Gaussian random process. The power spectral density of $Z(t)$ is measured experimentally and found to be

$$S_{ZZ}(\omega) = \pi\delta(\omega) + \frac{2\beta}{(\omega^2+\beta^2)(\omega^2+1)}.$$

(a) Find the correlation function $S_{YY}(\omega)$ in terms of β.

(b) Find the correlation function $S_{XX}(\omega)$.

Part (a)

To begin our analysis we recall that the power spectral density $S_{ZZ}(\omega)$ is given on page 443 in [5].

$$S_{ZZ}(\omega) = \int_{-\infty}^{\infty} R_{ZZ}(\tau)e^{-j\omega\tau}d\tau = \int_{-\infty}^{\infty} h(\tau) * R_{YY}(\tau) * h^*(-\tau)e^{-j\omega\tau}d\tau = |H(\omega)|^2 S_{YY}(\omega)$$

To proceed we need to determine a closed-form expression for the system frequency response $H(\omega)$. By definition the frequency response $H(\omega)$ is the discrete-time Fourier transform of the impulse response $h(t)$. As a result, we have

$$H(\omega) = \int_{-\infty}^{\infty} h(\tau)e^{-j\omega\tau}d\tau = \int_{-\infty}^{\infty} e^{-\tau}u(\tau)e^{-j\omega\tau}d\tau = \int_{0}^{\infty} e^{-(1+j\omega)\tau}d\tau = \frac{1}{1+j\omega}.$$

Evaluating the frequency response magnitude $|H(\omega)|^2$, we find

$$|H(\omega)|^2 = H(\omega)H^*(\omega) = \frac{1}{\omega^2+1}.$$

Substituting into the previous expression for $S_{ZZ}(\omega)$ yields a solution for $S_{YY}(\omega)$ in terms of β.

$$\boxed{S_{YY}(\omega) = \frac{S_{ZZ}(\omega)}{|H(\omega)|^2}, \text{ for } |H(\omega)| \neq 0 \quad \Rightarrow \quad S_{YY}(\omega) = \pi(\omega^2+1)\delta(\omega) + \frac{2\beta}{\omega^2+\beta^2}}$$

Part (b)

Recall from Table 7.5-1 in [5] that the power spectral density of a random process $X^n(t)$, generated from the random process $X(t)$, is given by $\omega^{2n}S_{XX}(\omega)$. As a result, we find that the PSD of $X^2(t)$ should be given by $\omega^4 S_{XX}(\omega)$. Since $Y(t) = X^2(t)$, we conclude that the correlation function $S_{XX}(\omega)$ has the following form.

$$\boxed{S_{XX}(\omega) = \frac{S_{YY}(\omega)}{\omega^4}, \text{ for } \omega \neq 0 \quad \Rightarrow \quad S_{YY}(\omega) = \pi\left(\frac{\omega^2+1}{\omega^4}\right)\delta(\omega) + \frac{2\beta}{\omega^4(\omega^2+\beta^2)}, \text{ for } \omega \neq 0}$$

Problem 9.3

Use the orthogonality principle to show that the minimum mean-square error (MMSE)

$$\varepsilon^2 \triangleq E[(X - E[X|Y])^2], \tag{14}$$

for real-valued random variables, can be expressed as

$$\varepsilon^2 = E[X(X - E[X|Y])]$$

or as

$$\varepsilon^2 = E[X^2] - E[E[X|Y]^2].$$

Generalize to the case where $\mathbf{X}$ and $\mathbf{Y}$ are real-valued random vectors. That is, show that the MMSE matrix is

$$\begin{aligned} \varepsilon^2 &\triangleq E[(\mathbf{X} - E[\mathbf{X}|\mathbf{Y}])(\mathbf{X} - E[\mathbf{X}|\mathbf{Y}])^T] \\ &= E[\mathbf{X}(\mathbf{X} - E[\mathbf{X}|\mathbf{Y}])^T] \\ &= E[\mathbf{X}\mathbf{X}^T] - E[E[\mathbf{X}|\mathbf{Y}]E^T[\mathbf{X}|\mathbf{Y}]]. \end{aligned} \tag{15}$$

Let's begin by expanding the product in Equation 14.

$$\begin{aligned} \varepsilon^2 &= E[(X - E[X|Y])(X - E[X|Y])] \\ &= E[X(X - E[X|Y]) - E[X|Y](X - E[X|Y])] \\ &= E[X(X - E[X|Y])] - E[E[X|Y](X - E[X|Y])] \end{aligned} \tag{16}$$

At this point we recall the *orthogonality principle*, as given by Property 9.1-1 on page 555 in [5] and Theorem 5.4.1 on page 327 in [2]. That is, the MMSE error vector

$$\varepsilon \triangleq \mathbf{X} - E[\mathbf{X}|\mathbf{Y}]$$

is orthogonal to any measurable function $h(\mathbf{Y})$ of the data, such that

$$E[h^*(\mathbf{Y})(\mathbf{X} - E[\mathbf{X}|\mathbf{Y}])^T] = \mathbf{0}. \tag{17}$$

For the random variables X and Y, Equation 17 yields the following condition for $h^*(Y) \triangleq E[X|Y]$.

$$E[E[X|Y](X - E[X|Y])] = 0$$

Substituting this result into Equation 16 yields the desired relation via the orthogonality principle.

$$\varepsilon^2 = E[X(X - E[X|Y])]$$

To complete the scalar-valued derivation, we further expand this product as follows.

$$\varepsilon^2 = E[X(X - E[X|Y])] = E[X^2] - E[XE[X|Y]]$$

Recall, from Equation 4.2-27 in [5], the smoothing property of the conditional expectation ensures

$$E[X] = E[E[X|Y]]$$

for the random variables X and Y. Applying this condition to the previous expression yields the final solution.

$$\therefore \boxed{\varepsilon^2 = E[X(X - E[X|Y])] = E[X^2] - E[E[X|Y]^2]}$$

Now let's generalize to the case where $\mathbf{X}$ and $\mathbf{Y}$ are real-valued random vectors. We begin by expanding the product in Equation 15.

$$\begin{aligned}
\varepsilon^2 &= E[(\mathbf{X} - E[\mathbf{X}|\mathbf{Y}])(\mathbf{X} - E[\mathbf{X}|\mathbf{Y}])^T] \\
&= E[\mathbf{X}(\mathbf{X} - E[\mathbf{X}|\mathbf{Y}])^T - E[\mathbf{X}|\mathbf{Y}](\mathbf{X} - E[\mathbf{X}|\mathbf{Y}])^T] \\
&= E[\mathbf{X}(\mathbf{X} - E[\mathbf{X}|\mathbf{Y}])^T] - E[E[\mathbf{X}|\mathbf{Y}](\mathbf{X} - E[\mathbf{X}|\mathbf{Y}])^T] \qquad (18)
\end{aligned}$$

For the random vectors $\mathbf{X}$ and $\mathbf{Y}$, Equation 17 yields the following condition for $h^*(\mathbf{Y}) \triangleq E[\mathbf{X}|\mathbf{Y}]$.

$$E[E[\mathbf{X}|\mathbf{Y}](\mathbf{X} - E[\mathbf{X}|\mathbf{Y}])^T] = \mathbf{0}$$

Substituting this result into Equation 18 yields the desired relation via the orthogonality principle.

$$\varepsilon^2 = E[\mathbf{X}(\mathbf{X} - E[\mathbf{X}|\mathbf{Y}])^T]$$

To complete the vector-valued derivation, we further expand this product as follows.

$$\varepsilon^2 = E[\mathbf{X}\mathbf{X}^T] - E[\mathbf{X}E^T[\mathbf{X}|\mathbf{Y}]]$$

As in the scalar-valued case, the smoothing property of the conditional expectation ensures

$$E[\mathbf{X}] = E[E[\mathbf{X}|\mathbf{Y}]].$$

Applying this condition to the previous expression yields the desired solution.

$$\therefore \boxed{\varepsilon^2 = E[\mathbf{X}(\mathbf{X} - E[\mathbf{X}|\mathbf{Y}])^T] = E[\mathbf{X}\mathbf{X}^T] - E[E[\mathbf{X}|\mathbf{Y}]E^T[\mathbf{X}|\mathbf{Y}]]}$$

(QED)

Problem 9.6

A Gaussian random sequence $X[n]$, for $n = 0, 1, 2, \ldots$, is defined as

$$X[n] = -\sum_{k=1}^{n} \binom{k+2}{2} X[n-k] + W[n], \tag{19}$$

where $X[0] = W[0]$ and $W[n]$ is a Gaussian white noise sequence with zero mean and unity variance.

(a) Show that $W[n]$ is the innovations sequence for $X[n]$.

(b) Show that $X[n] = W[n] - 3W[n-1] + 3W[n-2] - W[n-3]$, for $W[-1] = W[-2] = W[-3] = 0$.

(c) Use the preceding result to obtain the best two-step predictor of $X[12]$ as a linear combination of $X[0], \ldots, X[10]$. Also calculate the resulting mean-square prediction error.

Part (a)

Recall, from Definition 9.2-1 on page 571 in [5], that the innovations sequence for a random sequence $X[n]$ is defined to be a white random sequence which is a casual and causally-invertible linear transformation of the sequence $X[n]$. From Equation 19 we find that $W[n]$ is a causal linear transformation of $\{X[0], X[1], \ldots, X[n]\}$ such that

$$W[n] = X[n] + \sum_{k=1}^{n} \binom{k+2}{2} X[n-k].$$

In addition, we note that each $X[n]$ is composed of a linear combination of zero-mean Gaussian random variables and, as a result, must also be a white random sequence. In conclusion, we find that $W[n]$ is a white random sequence that is causally equivalent to $X[n]$. Similarly, as we'll show in Part (b), $X[n]$ can be expressed as a causal linear combination of $\{W[n-3], W[n-2], W[n-1], W[n]\}$. As a result, we find that $W[n]$ is the innovations sequence for $X[n]$ since it satisfies Definition 9.2-1. In other words, $W[n]$ contains the new information obtained when we observe $X[n]$ given the past observations $\{X[n-1], X[n-2], \ldots, X[0]\}$.

Part (b)

Let's begin by evaluating $X[1]$ by direct evaluation of Equation 19.

$$\begin{aligned} X[1] &= W[1] - \sum_{k=1}^{1} \binom{k+2}{2} X[1-k] \\ &= W[1] - 3X[0] = W[1] - 3W[0] \end{aligned}$$

Similarly, for $X[2]$ we find the following result.

$$\begin{aligned} X[2] &= W[2] - \sum_{k=1}^{2} \binom{k+2}{2} X[2-k] \\ &= W[2] - 3X[1] - 6X[0] = W[2] - 3W[1] + 3W[0] \end{aligned}$$

Continuing our analysis we find that $X[3]$ has the following solution.

$$\begin{aligned}X[3] &= W[3] - \sum_{k=1}^{3} \binom{k+2}{2} X[3-k] \\ &= W[3] - 3X[2] - 6X[1] - 10X[0] = W[3] - 3W[2] + 3W[1] - W[0]\end{aligned}$$

By induction we conclude that the general solution for $X[n]$, for $n = 0, 1, 2, \ldots$, is given by the following expression.

$$\boxed{X[n] = W[n] - 3[n-1] + 3W[n-2] - W[n-3], \text{ for } W[-1] = W[-2] = W[-3] = 0}$$

Part (c)

Recall that the best two-step predictor $\hat{X}[12]$ of $X[12]$ will be given by the following conditional expectation.

$$\hat{X}[12] = E\{X[12]|X[10], \ldots, X[0]\}$$

Note that $W[n]$, the innovations sequence, is causally equivalent to $X[n]$. As a result, we can also express the two-step predictor as follows.

$$\begin{aligned}\hat{X}[12] &= E\{X[12]|W[10], \ldots, W[0]\} \\ &= E\{W[12] - 3W[11] + 3W[10] - W[9]|W[10], \ldots, W[0]\} \\ &= 3W[10] - W[9]\end{aligned}$$

Note that we substituted for $X[n]$ using the result found in Part (b). Since $W[n]$ is a white random process, we also conclude that $E\{W[12]|W[10], \ldots, W[0]\} = E\{W[11]|W[10], \ldots, W[0]\} = 0$. As a result, the best two-step predictor of $X[12]$ is given by the following expression.

$$\boxed{\hat{X}[12] = 3W[10] - W[9] = 3\left\{X[10] + \sum_{k=1}^{10} \binom{k+2}{2} X[10-k]\right\} - \left\{X[9] + \sum_{k=1}^{9} \binom{k+2}{2} X[9-k]\right\}}$$

Finally, we note that the mean-square prediction error ε^2 is given by the following expression.

$$\begin{aligned}\varepsilon^2 &= E\{(X[12] - \hat{X}[12])^2\} = E\{(W[12] - 3W[11])^2\} \\ &= E\{W[12]^2\} - 6E\{W[12]W[11]\} + 9E\{W[11]^2\} \\ &= E\{W[12]^2\} - 6E\{W[12]\}E\{W[11]\} + 9E\{W[11]^2\} = 10\end{aligned}$$

Since $W[n]$ is a mean-zero white random process we conclude that $E\{W[12]^2\} = E\{W[11]^2\} = 1$ and $E\{W[12]W[11]\} = E\{W[12]\}E\{W[11]\} = 0$. In conclusion, the mean-square prediction error ε^2 for $X[12]$ is given by the following equation.

$$\boxed{\varepsilon^2 = E\{(X[12] - \hat{X}[12])^2\} = 10}$$

Problem 9.8

A random sequence $Y[n]$, for $n = 0, 1, 2, \ldots$, satisfies the second-order linear difference equation

$$2Y[n+2] + Y[n+1] + Y[n] = 2W[n], \text{ for } Y[0] = 0, Y[1] = 1,$$

with $W[n]$ a standard white Gaussian random sequence. Transform this equation into the state-space representation and evaluate the mean function $\mu_{\mathbf{X}}[n]$ and the correlation function $\mathbf{R}_{\mathbf{XX}}[n_1, n_2]$ for at least the first few values of n. (Hint: Define the state vector $\mathbf{X}[n] \triangleq (Y[n+2], Y[n+1])^T$.)

As requested, let's begin by transforming the linear constant coefficient difference equation into the state-space representation. Following the method outlined in Example 6.6-2 on page 374 in [5], we conclude that the state-space representation has the following form.

$$\boxed{\begin{gathered}\mathbf{X}[n] = \mathbf{A}\mathbf{X}[n-1] + \mathbf{b}W[n], \text{ where} \\ \mathbf{X}[n] = \begin{pmatrix} Y[n+2] \\ Y[n+1] \end{pmatrix}, \quad \mathbf{A} = \begin{pmatrix} -1/2 & -1/2 \\ 1 & 0 \end{pmatrix}, \quad \mathbf{b} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad \text{and} \quad \mathbf{X}[-1] = \begin{pmatrix} 1 \\ 0 \end{pmatrix}\end{gathered}}$$

To confirm this expression, we write out the matrix-vector product and compare to the original difference equation.

$$\begin{pmatrix} Y[n+2] \\ Y[n+1] \end{pmatrix} = \begin{pmatrix} -1/2 & -1/2 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} Y[n+1] \\ Y[n] \end{pmatrix} + \begin{pmatrix} 1 \\ 0 \end{pmatrix} W[n]$$

Now recall that the general solution to the resulting vector-valued difference equation is given by Equation 9.2-2 on page 571 in [5].

$$\mathbf{X}[n] = \mathbf{A}^{n+1}\mathbf{X}[-1] + \sum_{m=0}^{n} \mathbf{A}^{n-m}\mathbf{b}W[m]$$

The mean function can be obtained using the standard definition as follows.

$$\begin{aligned}\mu_{\mathbf{X}}[n] = E\{\mathbf{X}[n]\} &= E\left\{\mathbf{A}^{n+1}\mathbf{X}[-1] + \sum_{m=0}^{n} \mathbf{A}^{n-m}\mathbf{b}W[m]\right\} \\ &= \mathbf{A}^{n+1}\mathbf{X}[-1] + \sum_{m=0}^{n} \mathbf{A}^{n-m}\mathbf{b}E\{W[m]\} \\ &= \mathbf{A}^{n+1}\mathbf{X}[-1]\end{aligned}$$

Note that in the previous expression we have exploited the linearity property of the expectation operator and the fact that $E\{W[m]\} = 0$, $\forall n$. As a result, we conclude that the mean function $\mu_{\mathbf{X}}[n]$ is given by the following expression (with the resulting first few values also shown below).

$$\boxed{\begin{gathered}\mu_{\mathbf{X}}[n] = \mathbf{A}^{n+1}\mathbf{X}[-1] \\ \mu_{\mathbf{X}}[0] = \begin{pmatrix} -1/2 \\ 1 \end{pmatrix}, \quad \mu_{\mathbf{X}}[1] = \begin{pmatrix} -1/4 \\ -1/2 \end{pmatrix}, \quad \text{and} \quad \mu_{\mathbf{X}}[2] = \begin{pmatrix} 3/8 \\ -1/4 \end{pmatrix}\end{gathered}}$$

To complete our analysis, we recall that the autocorrelation function $\mathbf{R}_{\mathbf{XX}}[n_1, n_2]$ can also be obtained using the standard definition

$$\mathbf{R}_{\mathbf{XX}}[n_1, n_2] = E\{\mathbf{X}[n_1]\mathbf{X}^{\dagger}[n_2]\},$$

where $\mathbf{X}^\dagger[n]$ denotes the conjugate transpose of $\mathbf{X}[n]$. In this problem $\mathbf{X}$ is real-valued, so we conclude that $\mathbf{R}_{\mathbf{XX}}[n_1,n_2]$ is given by the following expression.

$$\mathbf{R}_{\mathbf{XX}}[n_1,n_2] = E\{\mathbf{X}[n_1]\mathbf{X}^T[n_2]\}$$

Substituting the general solution for $\mathbf{X}[n]$, we find the following result.

$$\begin{aligned}\mathbf{R}_{\mathbf{XX}}[n_1,n_2] &= E\left\{\left(\mathbf{A}^{n_1+1}\mathbf{X}[-1] + \sum_{m_1=0}^{n_1}\mathbf{A}^{n_1-m_1}\mathbf{b}W[m_1]\right)\left(\mathbf{A}^{n_2+1}\mathbf{X}[-1] + \sum_{m_2=0}^{n_2}\mathbf{A}^{n_2-m_2}\mathbf{b}W[m_2]\right)^T\right\}\\ &= E\left\{\left(\mathbf{A}^{n_1+1}\mathbf{X}[-1] + \sum_{m_1=0}^{n_1}\mathbf{A}^{n_1-m_1}\mathbf{b}W[m_1]\right)\left(\mathbf{X}^T[-1](\mathbf{A}^T)^{n_2+1} + \sum_{m_2=0}^{n_2}\mathbf{b}^T(\mathbf{A}^T)^{n_2-m_2}W[m_2]\right)\right\}\end{aligned}$$

Once again we can exploit the linearity property of the expectation operator. In addition, notice that the cross-terms in $W[n]$ will be eliminated since $E\{W[n]\} = 0,\ \forall n$. As a result, the autocorrelation function has the following solution.

$$\mathbf{R}_{\mathbf{XX}}[n_1,n_2] = \mathbf{A}^{n_1+1}\mathbf{X}[-1]\mathbf{X}^T[-1](\mathbf{A}^T)^{n_2+1} + \sum_{m_1=0}^{n_1}\sum_{m_2=0}^{n_2}\mathbf{A}^{n_1-m_1}\mathbf{b}\mathbf{b}^T(\mathbf{A}^T)^{n_2-m_2}E\{W[m_1]W[m_2]\}$$

At this point we recall that $E\{W[m_1]W[m_2]\} = \sigma_W^2\delta[m_1-m_2]$ for $W[n]$ a white Gaussian random sequence. According to the problem state, $\sigma_W^2 = 1$ which leads to the following solution for the autocorrelation function $\mathbf{R}_{\mathbf{XX}}[n_1,n_2]$ (with the resulting first few values also shown below).

$$\boxed{\begin{gathered}\mathbf{R}_{\mathbf{XX}}[n_1,n_2] = \begin{cases}\mathbf{A}^{n_1+1}\mathbf{X}[-1]\mathbf{X}^T[-1](\mathbf{A}^T)^{n_2+1} + \sum_{m=0}^{n_2}\mathbf{A}^{n_1-m}\mathbf{b}\mathbf{b}^T(\mathbf{A}^T)^{n_2-m}, & \text{for } n_1 \geq n_2\\ \mathbf{A}^{n_1+1}\mathbf{X}[-1]\mathbf{X}^T[-1](\mathbf{A}^T)^{n_2+1} + \sum_{m=0}^{n_1}\mathbf{A}^{n_1-m}\mathbf{b}\mathbf{b}^T(\mathbf{A}^T)^{n_2-m}, & \text{for } n_1 < n_2\end{cases}\\ \mathbf{R}_{\mathbf{XX}}[0,0] = \begin{pmatrix}5/4 & -1/2\\ -1/2 & 1\end{pmatrix},\ \mathbf{R}_{\mathbf{XX}}[0,1] = \begin{pmatrix}-3/8 & 5/4\\ -1/4 & -1/2\end{pmatrix},\ \text{and } \mathbf{R}_{\mathbf{XX}}[1,1] = \begin{pmatrix}21/16 & -3/8\\ -3/8 & 5/4\end{pmatrix}\end{gathered}}$$

References

[1] Geoffrey Grimmett and David Stirzaker. *Probability and Random Processes (Third Edition)*. Oxford University Press, 2001.

[2] Harold J. Larson and Bruno O. Shubert. *Probabilistic Models in Engineering Sciences: Random Variables and Stochastic Processes*. John Wiley & Sons, 1979.

[3] Sanjit K. Mitra. *Digital Signal Processing: A Computer-Based Approach*. McGraw-Hill, New York, NY, 2006.

[4] Michael Mitzenmacher and Eli Upfal. *Probability and Computing: Randomized Algorithms and Probabilistic Analysis*. Cambridge University Press, 2005.

[5] Henry Stark and John W. Woods. *Probability and Random Processes with Applications to Signal Processing (Third Edition)*. Prentice-Hall, 2002.