Problem Solutions – Chapter 11

Problem 11.1.1 Solution
For this problem, it is easiest to work with the expectation operator. The mean function of the output is

$$E\left[Y(t)\right] = 2 + E\left[X(t)\right] = 2 \tag{1}$$

The autocorrelation of the output is

$$R_Y(t,\tau) = E\left[(2 + X(t))(2 + X(t + \tau))\right] \tag{2}$$
$$= E\left[4 + 2X(t) + 2X(t + \tau) + X(t)X(t + \tau)\right] \tag{3}$$
$$= 4 + 2E\left[X(t)\right] + 2E\left[X(t + \tau)\right] + E\left[X(t)X(t + \tau)\right] \tag{4}$$
$$= 4 + R_X(\tau) \tag{5}$$

We see that $R_Y(t,\tau)$ only depends on the time difference τ. Thus $Y(t)$ is wide sense stationary.

Problem 11.1.2 Solution
By Theorem 11.2, the mean of the output is

$$\mu_Y = \mu_X \int_{-\infty}^{\infty} h(t)\, dt \tag{1}$$

$$= -3 \int_0^{10^{-3}} (1 - 10^6 t^2)\, dt \tag{2}$$

$$= -3 \left(t - (10^6/3)t^3\right)\Big|_0^{10^{-3}} \tag{3}$$

$$= -2 \times 10^{-3} \text{ volts} \tag{4}$$

Problem 11.1.3 Solution
By Theorem 11.2, the mean of the output is

$$\mu_Y = \mu_X \int_{-\infty}^{\infty} h(t)\, dt = 4 \int_0^{\infty} e^{-t/a}\, dt = -4a e^{-t/a}\Big|_0^{\infty} = 4a. \tag{1}$$

Since $\mu_Y = 1 = 4a$, we must have $a = 1/4$.

Problem 11.1.4 Solution
Since $E[Y^2(t)] = R_Y(0)$, we use Theorem 11.2(a) to evaluate $R_Y(\tau)$ at $\tau = 0$. That is,

$$R_Y(0) = \int_{-\infty}^{\infty} h(u) \int_{-\infty}^{\infty} h(v) R_X(u - v)\, dv\, du \tag{1}$$

$$= \int_{-\infty}^{\infty} h(u) \int_{-\infty}^{\infty} h(v) \eta_0 \delta(u - v)\, dv\, du \tag{2}$$

$$= \eta_0 \int_{-\infty}^{\infty} h^2(u)\, du, \tag{3}$$

by the sifting property of the delta function.

Problem 11.2.1 Solution

(a) Note that

$$Y_i = \sum_{n=-\infty}^{\infty} h_n X_{i-n} = \frac{1}{3} X_{i+1} + \frac{1}{3} X_i + \frac{1}{3} X_{i-1} \tag{1}$$

By matching coefficients, we see that

$$h_n = \begin{cases} 1/3 & n = -1, 0, 1 \\ 0 & \text{otherwise} \end{cases} \tag{2}$$

(b) By Theorem 11.5, the output autocorrelation is

$$R_Y[n] = \sum_{i=-\infty}^{\infty} \sum_{j=-\infty}^{\infty} h_i h_j R_X[n + i - j] \tag{3}$$

$$= \frac{1}{9} \sum_{i=-1}^{1} \sum_{j=-1}^{1} R_X[n + i - j] \tag{4}$$

$$= \frac{1}{9} (R_X[n + 2] + 2R_X[n + 1] + 3R_X[n] + 2R_X[n - 1] + R_X[n - 2]) \tag{5}$$

Substituting in $R_X[n]$ yields

$$R_Y[n] = \begin{cases} 1/3 & n = 0 \\ 2/9 & |n| = 1 \\ 1/9 & |n| = 2 \\ 0 & \text{otherwise} \end{cases} \tag{6}$$

Problem 11.2.2 Solution
Applying Theorem 11.4 with sampling period $T_s = 1/4000$ s yields

$$R_X[k] = R_X(kT_s) = 10 \frac{\sin(2000\pi k T_s) + \sin(1000\pi k T_s)}{2000\pi k T_s} \tag{1}$$

$$= 20 \frac{\sin(0.5\pi k) + \sin(0.25\pi k)}{\pi k} \tag{2}$$

$$= 10 \operatorname{sinc}(0.5k) + 5 \operatorname{sinc}(0.25k) \tag{3}$$

Problem 11.2.3 Solution

(a) By Theorem 11.5, the expected value of the output is

$$\mu_W = \mu_Y \sum_{n=-\infty}^{\infty} h_n = 2\mu_Y = 2 \tag{1}$$

(b) Theorem 11.5 also says that the output autocorrelation is

$$R_W[n] = \sum_{i=-\infty}^{\infty} \sum_{j=-\infty}^{\infty} h_i h_j R_Y[n+i-j] \tag{2}$$

$$= \sum_{i=0}^{1} \sum_{j=0}^{1} R_Y[n+i-j] \tag{3}$$

$$= R_Y[n-1] + 2R_Y[n] + R_Y[n+1] \tag{4}$$

For $n = -3$,

$$R_W[-3] = R_Y[-4] + 2R_Y[-3] + R_Y[-2] = R_Y[-2] = 0.5 \tag{5}$$

Following the same procedure, its easy to show that $R_W[n]$ is nonzero for $|n| = 0, 1, 2$. Specifically,

$$R_W[n] = \begin{cases} 0.5 & |n| = 3 \\ 3 & |n| = 2 \\ 7.5 & |n| = 1 \\ 10 & n = 0 \\ 0 & \text{otherwise} \end{cases} \tag{6}$$

(c) The second moment of the output is $E[W_n^2] = R_W[0] = 10$. The variance of W_n is

$$\mathrm{Var}[W_n] = E\left[W_n^2\right] - (E\left[W_n\right])^2 = 10 - 2^2 = 6 \tag{7}$$

(d) This part doesn't require any probability. It just checks your knowledge of linear systems and convolution. There is a bit of confusion because h_n is used to denote both the filter that transforms X_n to Y_n as well as the filter that transforms Y_n to W_n. To avoid confusion, we will use $\hat{h}_n$ to denote the filter that transforms X_n to Y_n. Using Equation (11.25) for discrete-time convolution, we can write

$$W_n = \sum_{j=-\infty}^{\infty} h_j Y_{n-j}, \qquad\qquad Y_{n-j} = \sum_{i=-\infty}^{\infty} \hat{h}_i X_{n-j-i}. \tag{8}$$

Combining these equations yields

$$W_n = \sum_{j=-\infty}^{\infty} h_j \sum_{i=-\infty}^{\infty} \hat{h}_i X_{n-j-i}. \tag{9}$$

For each j, we make the substitution $k = i + j$, and then reverse the order of summation to obtain

$$W_n = \sum_{j=-\infty}^{\infty} h_j \sum_{k=-\infty}^{\infty} \hat{h}_{k-j} X_{n-k} = \sum_{k=-\infty}^{\infty} \left(\sum_{j=-\infty}^{\infty} h_j \hat{h}_{k-j} \right) X_{n-k}. \tag{10}$$

Thus we see that

$$g_k = \sum_{j=-\infty}^{\infty} h_j \hat{h}_{k-j}. \tag{11}$$

That is, the filter g_n is the convolution of the filters $\hat{h}_n$ and h_n.

383

In the context of our particular problem, the filter $\hat{h}_n$ that transforms X_n to Y_n is given in Example 11.5. The two filters are

$$\hat{\mathbf{h}} = \begin{bmatrix} \hat{h}_0 & \hat{h}_1 \end{bmatrix}' = \begin{bmatrix} 1/2 & 1/2 \end{bmatrix}' \qquad\qquad \mathbf{h} = \begin{bmatrix} h_0 & h_1 \end{bmatrix}' = \begin{bmatrix} 1 & 1 \end{bmatrix}' \qquad (12)$$

Keep in mind that $h_n = \hat{h}_n = 0$ for $n < 0$ or $n > 1$. From Equation (11),

$$g_k = \hat{h}_k + \hat{h}_{k-1} = \begin{cases} 1/2 & k = 0 \\ 1 & k = 1 \\ 1/2 & k = 2 \\ 0 & \text{otherwise} \end{cases} \qquad (13)$$

Problem 11.2.4 Solution

(a) By Theorem 11.5, the mean output is

$$\mu_V = \mu_Y \sum_{n=-\infty}^{\infty} h_n = (-1+1)\mu_Y = 0 \qquad (1)$$

(b) Theorem 11.5 also says that the output autocorrelation is

$$R_V[n] = \sum_{i=-\infty}^{\infty} \sum_{j=-\infty}^{\infty} h_i h_j R_Y[n+i-j] \qquad (2)$$

$$= \sum_{i=0}^{1} \sum_{j=0}^{1} h_i h_j R_Y[n+i-j] \qquad (3)$$

$$= -R_Y[n-1] + 2R_Y[n] - R_Y[n+1] \qquad (4)$$

For $n = -3$,

$$R_V[-3] = -R_Y[-4] + 2R_Y[-3] - R_Y[-2] = R_Y[-2] = -0.5 \qquad (5)$$

Following the same procedure, it's easy to show that $R_V[n]$ is nonzero for $|n| = 0, 1, 2$. Specifically,

$$R_V[n] = \begin{cases} -0.5 & |n| = 3 \\ -1 & |n| = 2 \\ 0.5 & |n| = 1 \\ 2 & n = 0 \\ 0 & \text{otherwise} \end{cases} \qquad (6)$$

(c) Since $E[V_n] = 0$, the variance of the output is $E[V_n^2] = R_V[0] = 2$. The variance of W_n is

$$\mathrm{Var}[V_n] = E\left[W_n^2\right] R_V[0] = 2 \qquad (7)$$

(d) This part doesn't require any probability. It just checks your knowledge of linear systems and convolution. There is a bit of confusion because h_n is used to denote both the filter that transforms X_n to Y_n as well as the filter that transforms Y_n to V_n. To avoid confusion, we will

384

use $\hat{h}_n$ to denote the filter that transforms X_n to Y_n. Using Equation (11.25) for discrete-time convolution, we can write

$$V_n = \sum_{j=-\infty}^{\infty} h_j Y_{n-j}, \qquad\qquad Y_{n-j} = \sum_{i=-\infty}^{\infty} \hat{h}_i X_{n-j-i}. \qquad (8)$$

Combining these equations yields

$$V_n = \sum_{j=-\infty}^{\infty} h_j \sum_{i=-\infty}^{\infty} \hat{h}_i X_{n-j-i}. \qquad (9)$$

For the inner sum, we make the substitution $k = i+j$ and then reverse the order of summation to obtain

$$V_n = \sum_{j=-\infty}^{\infty} h_j \sum_{k=-\infty}^{\infty} \hat{h}_{k-j} X_{n-k} = \sum_{k=-\infty}^{\infty} \left(\sum_{j=-\infty}^{\infty} h_j \hat{h}_{k-j} \right) X_{n-k}. \qquad (10)$$

Thus we see that

$$f_k = \sum_{j=-\infty}^{\infty} h_j \hat{h}_{k-j}. \qquad (11)$$

That is, the filter f_n is the convolution of the filters $\hat{h}_n$ and h_n. In the context of our particular problem, the filter $\hat{h}_n$ that transforms X_n to Y_n is given in Example 11.5. The two filters are

$$\hat{\mathbf{h}} = \begin{bmatrix} \hat{h}_0 & \hat{h}_1 \end{bmatrix}' = \begin{bmatrix} 1/2 & 1/2 \end{bmatrix}' \qquad\qquad \mathbf{h} = \begin{bmatrix} h_0 & h_1 \end{bmatrix}' = \begin{bmatrix} 1 & -1 \end{bmatrix}' \qquad (12)$$

Keep in mind that $h_n = \hat{h}_n = 0$ for $n < 0$ or $n > 1$. From Equation (11),

$$f_k = \hat{h}_k - \hat{h}_{k-1} = \begin{cases} 1/2 & k = 0 \\ 0 & k = 1 \\ -1/2 & k = 2 \\ 0 & \text{otherwise} \end{cases} \qquad (13)$$

Problem 11.2.5 Solution
We start with Theorem 11.5:

$$R_Y[n] = \sum_{i=-\infty}^{\infty} \sum_{j=-\infty}^{\infty} h_i h_j R_X[n+i-j] \qquad (1)$$

$$= R_X[n-1] + 2R_X[n] + R_X[n+1] \qquad (2)$$

First we observe that for $n \leq -2$ or $n \geq 2$,

$$R_Y[n] = R_X[n-1] + 2R_X[n] + R_X[n+1] = 0 \qquad (3)$$

This suggests that $R_X[n] = 0$ for $|n| > 1$. In addition, we have the following facts:

$$R_Y[0] = R_X[-1] + 2R_X[0] + R_X[1] = 2 \qquad (4)$$
$$R_Y[-1] = R_X[-2] + 2R_X[-1] + R_X[0] = 1 \qquad (5)$$
$$R_Y[1] = R_X[0] + 2R_X[1] + R_X[2] = 1 \qquad (6)$$

A simple solution to this set of equations is $R_X[0] = 1$ and $R_X[n] = 0$ for $n \neq 0$.

Problem 11.2.6 Solution

The mean of $Y_n = (X_n + Y_{n-1})/2$ can be found by realizing that Y_n is an infinite sum of the X_i's.

$$Y_n = \left[\frac{1}{2}X_n + \frac{1}{4}X_{n-1} + \frac{1}{8}X_{n-2} + \ldots\right] \tag{1}$$

Since the X_i's are each of zero mean, the mean of Y_n is also 0. The variance of Y_n can be expressed as

$$\text{Var}[Y_n] = \left[\frac{1}{4} + \frac{1}{16} + \frac{1}{64} + \ldots\right]\text{Var}[X] = \sum_{i=1}^{\infty}(\frac{1}{4})^i\sigma^2 = (\frac{1}{1-1/4} - 1)\sigma^2 = \sigma^2/3 \tag{2}$$

The above infinite sum converges to $\frac{1}{1-1/4} - 1 = 1/3$, implying

$$\text{Var}\left[Y_n\right] = (1/3)\text{Var}\left[X\right] = 1/3 \tag{3}$$

The covariance of $Y_{i+1}Y_i$ can be found by the same method.

$$\text{Cov}[Y_{i+1}, Y_i] = [\frac{1}{2}X_n + \frac{1}{4}X_{n-1} + \frac{1}{8}X_{n-2} + \ldots][\frac{1}{2}X_{n-1} + \frac{1}{4}X_{n-2} + \frac{1}{8}X_{n-3} + \ldots] \tag{4}$$

Since $E[X_iX_j] = 0$ for all $i \neq j$, the only terms that are left are

$$\text{Cov}[Y_{i+1}, Y_i] = \sum_{i=1}^{\infty}\frac{1}{2^i}\frac{1}{2^{i-1}}E[X_i^2] = \frac{1}{2}\sum_{i=1}^{\infty}\frac{1}{4^i}E[X_i^2] \tag{5}$$

Since $E[X_i^2] = \sigma^2$, we can solve the above equation, yielding

$$\text{Cov}\left[Y_{i+1}, Y_i\right] = \sigma^2/6 \tag{6}$$

Finally the correlation coefficient of Y_{i+1} and Y_i is

$$\rho_{Y_{i+1}Y_i} = \frac{\text{Cov}[Y_{i+1}, Y_i]}{\sqrt{\text{Var}[Y_{i+1}]}\sqrt{\text{Var}[Y_i]}} = \frac{\sigma^2/6}{\sigma^2/3} = \frac{1}{2} \tag{7}$$

Problem 11.2.7 Solution

There is a technical difficulty with this problem since X_n is not defined for $n < 0$. This implies $C_X[n, k]$ is not defined for $k < -n$ and thus $C_X[n, k]$ cannot be completely independent of k. When n is large, corresponding to a process that has been running for a long time, this is a technical issue, and not a practical concern. Instead, we will find $\bar{\sigma}^2$ such that $C_X[n, k] = C_X[k]$ for all n and k for which the covariance function is defined. To do so, we need to express X_n in terms of $Z_0, Z_1, \ldots, Z_{n_1}$. We do this in the following way:

$$X_n = cX_{n-1} + Z_{n-1} \tag{1}$$
$$= c[cX_{n-2} + Z_{n-2}] + Z_{n-1} \tag{2}$$
$$= c^2[cX_{n-3} + Z_{n-3}] + cZ_{n-2} + Z_{n-1} \tag{3}$$
$$\vdots \tag{4}$$
$$= c^nX_0 + c^{n-1}Z_0 + c^{n-2}Z_2 + \cdots + Z_{n-1} \tag{5}$$
$$= c^nX_0 + \sum_{i=0}^{n-1}c^{n-1-i}Z_i \tag{6}$$

Since $E[Z_i] = 0$, the mean function of the X_n process is

$$E[X_n] = c^n E[X_0] + \sum_{i=0}^{n-1} c^{n-1-i} E[Z_i] = E[X_0] \tag{7}$$

Thus, for X_n to be a zero mean process, we require that $E[X_0] = 0$. The autocorrelation function can be written as

$$R_X[n,k] = E[X_n X_{n+k}] = E\left[\left(c^n X_0 + \sum_{i=0}^{n-1} c^{n-1-i} Z_i\right)\left(c^{n+k} X_0 + \sum_{j=0}^{n+k-1} c^{n+k-1-j} Z_j\right)\right] \tag{8}$$

Although it was unstated in the problem, we will assume that X_0 is independent of $Z_0, Z_1, \ldots$ so that $E[X_0 Z_i] = 0$. Since $E[Z_i] = 0$ and $E[Z_i Z_j] = 0$ for $i \neq j$, most of the cross terms will drop out. For $k \geq 0$, autocorrelation simplifies to

$$R_X[n,k] = c^{2n+k} \text{Var}[X_0] + \sum_{i=0}^{n-1} c^{2(n-1)+k-2i} \bar{\sigma}^2 = c^{2n+k} \text{Var}[X_0] + \bar{\sigma}^2 c^k \frac{1 - c^{2n}}{1 - c^2} \tag{9}$$

Since $E[X_n] = 0$, $\text{Var}[X_0] = R_X[n,0] = \sigma^2$ and we can write for $k \geq 0$,

$$R_X[n,k] = \bar{\sigma}^2 \frac{c^k}{1 - c^2} + c^{2n+k}\left(\sigma^2 - \frac{\bar{\sigma}^2}{1 - c^2}\right) \tag{10}$$

For $k < 0$, we have

$$R_X[n,k] = E\left[\left(c^n X_0 + \sum_{i=0}^{n-1} c^{n-1-i} Z_i\right)\left(c^{n+k} X_0 + \sum_{j=0}^{n+k-1} c^{n+k-1-j} Z_j\right)\right] \tag{11}$$

$$= c^{2n+k} \text{Var}[X_0] + c^{-k} \sum_{j=0}^{n+k-1} c^{2(n+k-1-j)} \bar{\sigma}^2 \tag{12}$$

$$= c^{2n+k} \sigma^2 + \bar{\sigma}^2 c^{-k} \frac{1 - c^{2(n+k)}}{1 - c^2} \tag{13}$$

$$= \frac{\bar{\sigma}^2}{1 - c^2} c^{-k} + c^{2n+k}\left(\sigma^2 - \frac{\bar{\sigma}^2}{1 - c^2}\right) \tag{14}$$

We see that $R_X[n,k] = \sigma^2 c^{|k|}$ by choosing

$$\bar{\sigma}^2 = (1 - c^2)\sigma^2 \tag{15}$$

Problem 11.2.8 Solution

We can recusively solve for Y_n as follows.

$$Y_n = aX_n + aY_{n-1} \tag{1}$$

$$= aX_n + a[aX_{n-1} + aY_{n-2}] \tag{2}$$

$$= aX_n + a^2 X_{n-1} + a^2[aX_{n-2} + aY_{n-3}] \tag{3}$$

By continuing the same procedure, we can conclude that

$$Y_n = \sum_{j=0}^{n} a^{j+1} X_{n-j} + a^n Y_0 \qquad (4)$$

Since $Y_0 = 0$, the substitution $i = n - j$ yields

$$Y_n = \sum_{i=0}^{n} a^{n-i+1} X_i \qquad (5)$$

Now we can calculate the mean

$$E[Y_n] = E\left[\sum_{i=0}^{n} a^{n-i+1} X_i\right] = \sum_{i=0}^{n} a^{n-i+1} E[X_i] = 0 \qquad (6)$$

To calculate the autocorrelation $R_Y[m, k]$, we consider first the case when $k \geq 0$.

$$C_Y[m, k] = E\left[\sum_{i=0}^{m} a^{m-i+1} X_i \sum_{j=0}^{m+k} a^{m+k-j+1} X_j\right] = \sum_{i=0}^{m} \sum_{j=0}^{m+k} a^{m-i+1} a^{m+k-j+1} E[X_i X_j] \qquad (7)$$

Since the X_i is a sequence of iid standard normal random variables,

$$E[X_i X_j] = \begin{cases} 1 & i = j \\ 0 & \text{otherwise} \end{cases} \qquad (8)$$

Thus, only the $i = j$ terms make a nonzero contribution. This implies

$$C_Y[m, k] = \sum_{i=0}^{m} a^{m-i+1} a^{m+k-i+1} \qquad (9)$$

$$= a^k \sum_{i=0}^{m} a^{2(m-i+1)} \qquad (10)$$

$$= a^k \left[(a^2)^{m+1} + (a^2)^m + \cdots + a^2\right] \qquad (11)$$

$$= \frac{a^2}{1 - a^2} a^k \left[1 - (a^2)^{m+1}\right] \qquad (12)$$

For $k \leq 0$, we start from

$$C_Y[m, k] = \sum_{i=0}^{m} \sum_{j=0}^{m+k} a^{m-i+1} a^{m+k-j+1} E[X_i X_j] \qquad (13)$$

As in the case of $k \geq 0$, only the $i = j$ terms make a contribution. Also, since $m + k \leq m$,

$$C_Y[m, k] = \sum_{j=0}^{m+k} a^{m-j+1} a^{m+k-j+1} = a^{-k} \sum_{j=0}^{m+k} a^{m+k-j+1} a^{m+k-j+1} \qquad (14)$$

By steps quite similar to those for $k \geq 0$, we can show that

$$C_Y[m, k] = \frac{a^2}{1 - a^2} a^{-k} \left[1 - (a^2)^{m+k+1}\right] \qquad (15)$$

A general expression that is valid for all m and k would be

$$C_Y[m, k] = \frac{a^2}{1 - a^2} a^{|k|} \left[1 - (a^2)^{\min(m, m+k)+1}\right] \qquad (16)$$

Since $C_Y[m, k]$ depends on m, the Y_n process is not wide sense stationary.

Problem 11.3.1 Solution

Since the process X_n has expected value $E[X_n] = 0$, we know that $C_X(k) = R_X(k) = 2^{-|k|}$. Thus $\mathbf{X} = \begin{bmatrix} X_1 & X_2 & X_3 \end{bmatrix}'$ has covariance matrix

$$\mathbf{C_X} = \begin{bmatrix} 2^0 & 2^{-1} & 2^{-2} \\ 2^{-1} & 2^0 & 2^{-1} \\ 2^{-2} & 2^{-1} & 2^0 \end{bmatrix} = \begin{bmatrix} 1 & 1/2 & 1/4 \\ 1/2 & 1 & 1/2 \\ 1/4 & 1/2 & 1 \end{bmatrix}. \tag{1}$$

From Definition 5.17, the PDF of $\mathbf{X}$ is

$$f_{\mathbf{X}}(\mathbf{x}) = \frac{1}{(2\pi)^{n/2}[\det(\mathbf{C_X})]^{1/2}} \exp\left(-\frac{1}{2}\mathbf{x}'\mathbf{C_X^{-1}}\mathbf{x}\right). \tag{2}$$

If we are using MATLAB for calculations, it is best to decalre the problem solved at this point. However, if you like algebra, we can write out the PDF in terms of the variables x_1, x_2 and x_3. To do so we find that the inverse covariance matrix is

$$\mathbf{C_X^{-1}} = \begin{bmatrix} 4/3 & -2/3 & 0 \\ -2/3 & 5/3 & -2/3 \\ 0 & -2/3 & 4/3 \end{bmatrix} \tag{3}$$

A little bit of algebra will show that $\det(\mathbf{C_X}) = 9/16$ and that

$$\frac{1}{2}\mathbf{x}'\mathbf{C_X^{-1}}\mathbf{x} = \frac{2x_1^2}{3} + \frac{5x_2^2}{6} + \frac{2x_3^2}{3} - \frac{2x_1 x_2}{3} - \frac{2x_2 x_3}{3}. \tag{4}$$

It follows that

$$f_{\mathbf{X}}(\mathbf{x}) = \frac{4}{3(2\pi)^{3/2}} \exp\left(-\frac{2x_1^2}{3} - \frac{5x_2^2}{6} - \frac{2x_3^2}{3} + \frac{2x_1 x_2}{3} + \frac{2x_2 x_3}{3}\right). \tag{5}$$

Problem 11.3.2 Solution

The sequence X_n is passed through the filter

$$\mathbf{h} = \begin{bmatrix} h_0 & h_1 & h_2 \end{bmatrix}' = \begin{bmatrix} 1 & -1 & 1 \end{bmatrix}' \tag{1}$$

The output sequence is Y_n.

(a) Following the approach of Equation (11.58), we can write the output $\mathbf{Y}_3 = \begin{bmatrix} Y_1 & Y_2 & Y_3 \end{bmatrix}'$ as

$$\mathbf{Y}_3 = \begin{bmatrix} Y_1 \\ Y_2 \\ Y_3 \end{bmatrix} = \begin{bmatrix} h_1 & h_0 & 0 & 0 \\ h_2 & h_1 & h_0 & 0 \\ 0 & h_2 & h_1 & h_0 \end{bmatrix} \begin{bmatrix} X_0 \\ X_1 \\ X_2 \\ X_3 \end{bmatrix} = \underbrace{\begin{bmatrix} -1 & 1 & 0 & 0 \\ 1 & -1 & 1 & 0 \\ 0 & 1 & -1 & 1 \end{bmatrix}}_{\mathbf{H}} \underbrace{\begin{bmatrix} X_0 \\ X_1 \\ X_2 \\ X_3 \end{bmatrix}}_{\mathbf{X}}. \tag{2}$$

We note that the components of $\mathbf{X}$ are iid Gaussian $(0, 1)$ random variables. Hence $\mathbf{X}$ has covariance matrix $\mathbf{C_X} = \mathbf{I}$, the identity matrix. Since $\mathbf{Y}_3 = \mathbf{HX}$,

$$\mathbf{C_{Y_3}} = \mathbf{HC_XH'} = \mathbf{HH'} = \begin{bmatrix} 2 & -2 & 1 \\ -2 & 3 & -2 \\ 1 & -2 & 3 \end{bmatrix}. \tag{3}$$

Some calculation (by hand or by MATLAB) will show that $\det(\mathbf{C_{Y_3}}) = 3$ and that

$$\mathbf{C_{Y_3}}^{-1} = \frac{1}{3}\begin{bmatrix} 5 & 4 & 1 \\ 4 & 5 & 2 \\ 1 & 2 & 2 \end{bmatrix}. \tag{4}$$

Some algebra will show that

$$\mathbf{y}'\mathbf{C_{Y_3}}^{-1}\mathbf{y} = \frac{5y_1^2 + 5y_2^2 + 2y_3^2 + 8y_1y_2 + 2y_1y_3 + 4y_2y_3}{3}. \tag{5}$$

This implies $\mathbf{Y}_3$ has PDF

$$f_{\mathbf{Y}_3}(\mathbf{y}) = \frac{1}{(2\pi)^{3/2}[\det(\mathbf{C_{Y_3}})]^{1/2}} \exp\left(-\frac{1}{2}\mathbf{y}'\mathbf{C_{Y_3}}^{-1}\mathbf{y}\right) \tag{6}$$

$$= \frac{1}{(2\pi)^{3/2}\sqrt{3}} \exp\left(-\frac{5y_1^2 + 5y_2^2 + 2y_3^2 + 8y_1y_2 + 2y_1y_3 + 4y_2y_3}{6}\right). \tag{7}$$

(b) To find the PDF of $\mathbf{Y}_2 = \begin{bmatrix} Y_1 & Y_2 \end{bmatrix}'$, we start by observing that the covariance matrix of $\mathbf{Y}_2$ is just the upper left 2×2 submatrix of $\mathbf{C_{Y_3}}$. That is,

$$\mathbf{C_{Y_2}} = \begin{bmatrix} 2 & -2 \\ -2 & 3 \end{bmatrix} \quad \text{and} \quad \mathbf{C_{Y_2}}^{-1} = \begin{bmatrix} 3/2 & 1 \\ 1 & 1 \end{bmatrix}. \tag{8}$$

Since $\det(\mathbf{C_{Y_2}}) = 2$, it follows that

$$f_{\mathbf{Y}_2}(\mathbf{y}) = \frac{1}{(2\pi)^{3/2}[\det(\mathbf{C_{Y_2}})]^{1/2}} \exp\left(-\frac{1}{2}\mathbf{y}'\mathbf{C_{Y_2}}^{-1}\mathbf{y}\right) \tag{9}$$

$$= \frac{1}{(2\pi)^{3/2}\sqrt{2}} \exp\left(-\frac{3}{2}y_1^2 - 2y_1y_2 - y_2^2\right). \tag{10}$$

Problem 11.3.3 Solution
The sequence X_n is passed through the filter

$$\mathbf{h} = \begin{bmatrix} h_0 & h_1 & h_2 \end{bmatrix}' = \begin{bmatrix} 1 & -1 & 1 \end{bmatrix}' \tag{1}$$

The output sequence is Y_n. Following the approach of Equation (11.58), we can write the output $\mathbf{Y} = \begin{bmatrix} Y_1 & Y_2 & Y_3 \end{bmatrix}'$ as

$$\mathbf{Y} = \begin{bmatrix} Y_1 \\ Y_2 \\ Y_3 \end{bmatrix} = \begin{bmatrix} h_2 & h_1 & h_0 & 0 & 0 \\ 0 & h_2 & h_1 & h_0 & 0 \\ 0 & 0 & h_2 & h_1 & h_0 \end{bmatrix} \begin{bmatrix} X_{-1} \\ X_0 \\ X_1 \\ X_2 \\ X_3 \end{bmatrix} = \underbrace{\begin{bmatrix} 1 & -1 & 1 & 0 & 0 \\ 0 & 1 & -1 & 1 & 0 \\ 0 & 0 & 1 & -1 & 1 \end{bmatrix}}_{\mathbf{H}} \underbrace{\begin{bmatrix} X_{-1} \\ X_0 \\ X_1 \\ X_2 \\ X_3 \end{bmatrix}}_{\mathbf{X}}. \tag{2}$$

Since X_n has autocovariance function $C_X(k) = 2^{-|k|}$, $\mathbf{X}$ has covariance matrix

$$\mathbf{C_X} = \begin{bmatrix} 1 & 1/2 & 1/4 & 1/8 & 1/16 \\ 1/2 & 1 & 1/2 & 1/4 & 1/8 \\ 1/4 & 1/2 & 1 & 1/2 & 1/4 \\ 1/8 & 1/4 & 1/2 & 1 & 1/2 \\ 1/16 & 1/8 & 1/4 & 1/2 & 1 \end{bmatrix}. \tag{3}$$

Since $\mathbf{Y} = \mathbf{HX}$,

$$\mathbf{C_Y} = \mathbf{HC_X H'} = \begin{bmatrix} 3/2 & -3/8 & 9/16 \\ -3/8 & 3/2 & -3/8 \\ 9/16 & -3/8 & 3/2 \end{bmatrix}. \tag{4}$$

Some calculation (by hand or preferably by MATLAB) will show that $\det(\mathbf{C_Y}) = 675/256$ and that

$$\mathbf{C_Y^{-1}} = \frac{1}{15} \begin{bmatrix} 12 & 2 & -4 \\ 2 & 11 & 2 \\ -4 & 2 & 12 \end{bmatrix}. \tag{5}$$

Some algebra will show that

$$\mathbf{y'C_Y^{-1}y} = \frac{12y_1^2 + 11y_2^2 + 12y_3^2 + 4y_1y_2 + -8y_1y_3 + 4y_2y_3}{15}. \tag{6}$$

This implies $\mathbf{Y}$ has PDF

$$f_{\mathbf{Y}}(\mathbf{y}) = \frac{1}{(2\pi)^{3/2}[\det(\mathbf{C_Y})]^{1/2}} \exp\left(-\frac{1}{2}\mathbf{y'C_Y^{-1}y}\right) \tag{7}$$

$$= \frac{16}{(2\pi)^{3/2}15\sqrt{3}} \exp\left(-\frac{12y_1^2 + 11y_2^2 + 12y_3^2 + 4y_1y_2 + -8y_1y_3 + 4y_2y_3}{30}\right). \tag{8}$$

This solution is another demonstration of why the PDF of a Gaussian random vector should be left in vector form.

Comment: We know from Theorem 11.5 that Y_n is a stationary Gaussian process. As a result, the random variables Y_1, Y_2 and Y_3 are identically distributed and $\mathbf{C_Y}$ is a symmetric Toeplitz matrix. This might make on think that the PDF $f_{\mathbf{Y}}(\mathbf{y})$ should be symmetric in the variables y_1, y_2 and y_3. However, because Y_2 is in the middle of Y_1 and Y_3, the information provided by Y_1 and Y_3 about Y_2 is different than the information Y_1 and Y_2 convey about Y_3. This fact appears as asymmetry in $f_{\mathbf{Y}}(\mathbf{y})$.

Problem 11.3.4 Solution

The sequence X_n is passed through the filter

$$\mathbf{h} = \begin{bmatrix} h_0 & h_1 & h_2 \end{bmatrix}' = \begin{bmatrix} 1 & 0 & -1 \end{bmatrix}' \tag{1}$$

The output sequence is Y_n. Following the approach of Equation (11.58), we can write the output $\mathbf{Y} = \begin{bmatrix} Y_1 & Y_2 & Y_3 \end{bmatrix}'$ as

$$\mathbf{Y} = \begin{bmatrix} Y_1 \\ Y_2 \\ Y_3 \end{bmatrix} = \begin{bmatrix} h_2 & h_1 & h_0 & 0 & 0 \\ 0 & h_2 & h_1 & h_0 & 0 \\ 0 & 0 & h_2 & h_1 & h_0 \end{bmatrix} \begin{bmatrix} X_{-1} \\ X_0 \\ X_1 \\ X_2 \\ X_3 \end{bmatrix} = \underbrace{\begin{bmatrix} 1 & 0 & -1 & 0 & 0 \\ 0 & 1 & 0 & -1 & 0 \\ 0 & 0 & 1 & 0 & -1 \end{bmatrix}}_{\mathbf{H}} \underbrace{\begin{bmatrix} X_{-1} \\ X_0 \\ X_1 \\ X_2 \\ X_3 \end{bmatrix}}_{\mathbf{X}}. \tag{2}$$

Since X_n has autocovariance function $C_X(k) = 2^{-|k|}$, $\mathbf{X}$ has the Toeplitz covariance matrix

$$\mathbf{C_X} = \begin{bmatrix} 1 & 1/2 & 1/4 & 1/8 & 1/16 \\ 1/2 & 1 & 1/2 & 1/4 & 1/8 \\ 1/4 & 1/2 & 1 & 1/2 & 1/4 \\ 1/8 & 1/4 & 1/2 & 1 & 1/2 \\ 1/16 & 1/8 & 1/4 & 1/2 & 1 \end{bmatrix}. \tag{3}$$

Since $\mathbf{Y} = \mathbf{HX}$,

$$\mathbf{C_Y} = \mathbf{HC_XH'} = \begin{bmatrix} 3/2 & 3/8 & -9/16 \\ 3/8 & 3/2 & 3/8 \\ -9/16 & 3/8 & 3/2 \end{bmatrix}. \tag{4}$$

Some calculation (preferably by MATLAB) will show that $\det(\mathbf{C_Y}) = 297/128$ and that

$$\mathbf{C_Y^{-1}} = \begin{bmatrix} 10/11 & -1/3 & 14/33 \\ -1/3 & 5/6 & -1/3 \\ 14/33 & -1/3 & 10/11 \end{bmatrix}. \tag{5}$$

Some algebra will show that

$$\mathbf{y'C_Y^{-1}y} = \frac{10}{11}y_1^2 + \frac{5}{6}y_2^2 + \frac{10}{11}y_3^2 - \frac{2}{3}y_1y_2 + \frac{28}{33}y_1y_3 - \frac{2}{3}y_2y_3. \tag{6}$$

This implies $\mathbf{Y}$ has PDF

$$f_{\mathbf{Y}}(\mathbf{y}) = \frac{1}{(2\pi)^{3/2}[\det(\mathbf{C_Y})]^{1/2}} \exp\left(-\frac{1}{2}\mathbf{y'C_Y^{-1}y}\right) \tag{7}$$

$$= \frac{8\sqrt{2}}{(2\pi)^{3/2}3\sqrt{33}} \exp\left(-\frac{5}{11}y_1^2 - \frac{5}{12}y_2^2 - \frac{5}{11}y_3^2 + \frac{1}{3}y_1y_2 - \frac{14}{33}y_1y_3 + \frac{1}{3}y_2y_3\right). \tag{8}$$

This solution is yet another demonstration of why the PDF of a Gaussian random vector should be left in vector form.

Problem 11.4.1 Solution

This problem is solved using Theorem 11.9 with $M = 2$ and $k = 1$. The optimum linear predictor filter $\mathbf{h} = \begin{bmatrix} h_0 & h_1 \end{bmatrix}'$ of X_{n+1} given $\mathbf{X}_n = \begin{bmatrix} X_{n-1} & X_n \end{bmatrix}'$ is given by

$$\overleftarrow{\mathbf{h}} = \begin{bmatrix} h_1 \\ h_0 \end{bmatrix} = \mathbf{R}_{\mathbf{X}_n}^{-1} \mathbf{R}_{\mathbf{X}_n X_{n+k}}, \tag{1}$$

where

$$\mathbf{R}_{\mathbf{X}_n} = \begin{bmatrix} R_X[0] & R_X[1] \\ R_X[1] & R_X[0] \end{bmatrix} = \begin{bmatrix} 1 & 3/4 \\ 3/4 & 1 \end{bmatrix} \tag{2}$$

and

$$\mathbf{R}_{\mathbf{X}_n X_{n+1}} = E\left[\begin{bmatrix} X_{n-1} \\ X_n \end{bmatrix} X_{n+1}\right] = \begin{bmatrix} R_X[2] \\ R_X[1] \end{bmatrix} = \begin{bmatrix} 1/2 \\ 3/4 \end{bmatrix}. \tag{3}$$

Thus the filter vector $\mathbf{h}$ satisfies

$$\overleftarrow{\mathbf{h}} = \begin{bmatrix} h_1 \\ h_0 \end{bmatrix} = \begin{bmatrix} 1 & 3/4 \\ 3/4 & 1 \end{bmatrix}^{-1} \begin{bmatrix} 1/2 \\ 3/4 \end{bmatrix} = \begin{bmatrix} -1/7 \\ 6/7 \end{bmatrix}. \tag{4}$$

Thus $\mathbf{h} = \begin{bmatrix} 6/7 & -1/7 \end{bmatrix}'$ and the optimum linear predictor of X_{n+1} given X_n and X_{n-1} is

$$\hat{X}_{n+1} = \overleftarrow{\mathbf{h}}'\mathbf{X}_n = \begin{bmatrix} -\frac{1}{7} & \frac{6}{7} \end{bmatrix} \begin{bmatrix} X_{n-1} \\ X_n \end{bmatrix} = -\frac{1}{7}X_{n-1} + \frac{6}{7}X_n. \tag{5}$$

To find the mean square error of this predictor, we can calculate it directly as

$$e_L^* = E\left[(\hat{X}_{n+1} - X_{n+1})^2\right] = E\left[\left(-\frac{1}{7}X_{n-1} + \frac{6}{7}X_n - X_{n+1}\right)^2\right]. \tag{6}$$

By expanding the square and taking the expectation, each cross-term is of the form $E[X_i X_j] = R_X[i-j]$, so that

$$e_L^* = E\left[\left(-\frac{1}{7}X_{n-1} + \frac{6}{7}X_n - X_{n+1}\right)\left(-\frac{1}{7}X_{n-1} + \frac{6}{7}X_n - X_{n+1}\right)^2\right] \tag{7}$$

$$= \frac{1}{49}R_X[0] - \frac{12}{49}R_X[1] + \frac{2}{7}R_X[2] + \frac{36}{49}R_X[0] - \frac{12}{7}R_X[1] + R_X[0] \tag{8}$$

$$= \frac{86}{49}R_X[0] - \frac{96}{49}R_X[1] + \frac{2}{7}R_X[2] = \frac{3}{7}. \tag{9}$$

This direct method is already tedious, even for a simple filter of order $M = 2$. A better way to calculate the mean square error is to recall that Theorem 11.9 is just of Theorem 9.7 expressed in the terminology of filters. Expressing part (c) of Theorem 9.7 in terms of the linear prediction filter $\mathbf{h}$, the mean square error of the predictor is

$$e_L^* = \text{Var}[X_{n+1}] - \overleftarrow{\mathbf{h}}' \mathbf{R}_{\mathbf{X}_n X_{n+k}} \tag{10}$$

$$= R_X[0] - \overleftarrow{\mathbf{h}}' \begin{bmatrix} R_X[2] \\ R_X[1] \end{bmatrix} \tag{11}$$

$$= 1 - \begin{bmatrix} -1/7 & 6/7 \end{bmatrix} \begin{bmatrix} 1/2 \\ 3/4 \end{bmatrix} = \frac{3}{7}. \tag{12}$$

For an arbitrary filter order M, Equation (10) is a much simpler way to compute the mean square error.

Problem 11.4.2 Solution
This problem is solved using Theorem 11.9 with $k = 1$. The optimum linear predictor filter $\mathbf{h} = \begin{bmatrix} h_0 & h_1 \end{bmatrix}'$ of X_{n+1} given $\mathbf{X}_n = \begin{bmatrix} X_{n-1} & X_n \end{bmatrix}'$ is given by

$$\overleftarrow{\mathbf{h}} = \begin{bmatrix} h_1 \\ h_0 \end{bmatrix} = \mathbf{R}_{\mathbf{X}_n}^{-1} \mathbf{R}_{\mathbf{X}_n X_{n+k}}, \tag{1}$$

where

$$\mathbf{R}_{\mathbf{X}_n} = \begin{bmatrix} R_X[0] & R_X[1] \\ R_X[1] & R_X[0] \end{bmatrix} = \begin{bmatrix} 1.1 & 0.75 \\ 0.75 & 1.1 \end{bmatrix} \tag{2}$$

and

$$\mathbf{R}_{\mathbf{X}_n X_{n+1}} = E\left[\begin{bmatrix} X_{n-1} \\ X_n \end{bmatrix} X_{n+1}\right] = \begin{bmatrix} R_X[2] \\ R_X[1] \end{bmatrix} = \begin{bmatrix} 0.5 \\ 0.75 \end{bmatrix}. \tag{3}$$

Thus the filter vector $\mathbf{h}$ satisfies

$$\overleftarrow{\mathbf{h}} = \begin{bmatrix} h_1 \\ h_0 \end{bmatrix} = \begin{bmatrix} 1.1 & 0.75 \\ 0.75 & 1.1 \end{bmatrix}^{-1} \begin{bmatrix} 0.5 \\ 0.75 \end{bmatrix} = \begin{bmatrix} -0.0193 \\ 0.6950 \end{bmatrix}. \tag{4}$$

Thus $\mathbf{h} = \begin{bmatrix} 0.6950 & -0.0193 \end{bmatrix}'$ and the optimum linear predictor of X_{n+1} given X_n and X_{n-1} is

$$\hat{X}_{n+1} = \overleftarrow{\mathbf{h}}' \mathbf{X}_n = \begin{bmatrix} -0.0193 & 0.6950 \end{bmatrix} \begin{bmatrix} X_{n-1} \\ X_n \end{bmatrix} = -0.0193 X_{n-1} + 0.6950 X_n. \tag{5}$$

To find the mean square error of this predictor, we can calculate it directly as

$$e_L^* = E\left[(\hat{X}_{n+1} - X_{n+1})^2\right] = E\left[(-0.0193 X_{n-1} + 0.6950 X_n - X_{n+1})^2\right]. \tag{6}$$

We can expand the square and take the expectation term by term since each cross-term is of the form $E[X_i X_j] = R_X[i-j]$. This approach is followed in the solution of Problem 11.4.1 and it is quite tedious. A better way to calculate the mean square error is to recall that Theorem 11.9 is just Theorem 9.7 expressed in the terminology of filters. Expressing part (c) of Theorem 9.7 in terms of the linear prediction filter $\mathbf{h}$, the mean square error of the predictor is

$$e_L^* = \text{Var}[X_{n+1}] - \overleftarrow{\mathbf{h}}' \mathbf{R}_{\mathbf{X}_n X_{n+k}} \tag{7}$$

$$= R_X[0] - \overleftarrow{\mathbf{h}}' \begin{bmatrix} R_X[2] \\ R_X[1] \end{bmatrix} \tag{8}$$

$$= 1.1 - \begin{bmatrix} -0.0193 & 0.6950 \end{bmatrix} \begin{bmatrix} 1/2 \\ 3/4 \end{bmatrix} = 0.5884. \tag{9}$$

Comment: It is instructive to compare this solution to the solution of Problem 11.4.1 where the random process, denoted $\tilde{X}_n$ here to distinguish it from X_n in this problem, has autocorrelation function

$$R_{\tilde{X}}[k] = \begin{cases} 1 - 0.25\,|k| & |k| \leq 4, \\ 0 & \text{otherwise.} \end{cases} \tag{10}$$

The difference is simply that $R_{\tilde{X}}[0] = 1$, rather than $R_X[0] = 1.1$ as in this problem. This difference corresponds to adding an iid noise sequence to $\tilde{X}_n$ to create X_n. That is,

$$X_n = \tilde{X}_n + Z_n \tag{11}$$

where Z_n is an iid additive noise sequence with autocorrelation function $R_Z[k] = 0.1\delta[k]$ that is independent of the X_n process. Thus X_n in this problem can be viewed as a noisy version of $\tilde{X}_n$ in Problem 11.4.1. Because the $\tilde{X}_n$ process is less noisy, the optimal predictor filter of $\tilde{X}_{n+1}$ given $\tilde{X}_{n-1}$ and $\tilde{X}_n$ is $\tilde{\mathbf{h}} = \begin{bmatrix} 6/7 & -1/7 \end{bmatrix}' = \begin{bmatrix} 0.8571 & -0.1429 \end{bmatrix}'$, which places more emphasis on the current value $\tilde{X}_n$ in predicting the next value.

In addition, the mean squared error of the predictor of $\tilde{X}_{n+1}$ is only $3/7 = 0.4285$, which is less than 0.5884. Not surprisingly, the noise in the X_n process reduces the performance of the predictor.

Problem 11.4.3 Solution

This problem generalizes Example 11.14 in that -0.9 is replaced by the parameter c and the noise variance 0.2 is replaced by η^2. Because we are only finding the first order filter $\mathbf{h} = \begin{bmatrix} h_0 & h_1 \end{bmatrix}'$, it is relatively simple to generalize the solution of Example 11.14 to the parameter values c and η^2.

Based on the observation $\mathbf{Y} = \begin{bmatrix} Y_{n-1} & Y_n \end{bmatrix}'$, Theorem 11.11 states that the linear MMSE estimate of $X = X_n$ is $\overleftarrow{\mathbf{h}}' \mathbf{Y}$ where

$$\overleftarrow{\mathbf{h}} = \mathbf{R}_{\mathbf{Y}}^{-1} \mathbf{R}_{\mathbf{Y} X_n} = (\mathbf{R}_{\mathbf{X}_n} + \mathbf{R}_{\mathbf{W}_n})^{-1} \mathbf{R}_{\mathbf{X}_n X_n}. \tag{1}$$

From Equation (11.82), $\mathbf{R}_{\mathbf{X}_n X_n} = \begin{bmatrix} R_X[1] & R_X[0] \end{bmatrix}' = \begin{bmatrix} c & 1 \end{bmatrix}'$. From the problem statement,

$$\mathbf{R}_{\mathbf{X}_n} + \mathbf{R}_{\mathbf{W}_n} = \begin{bmatrix} 1 & c \\ c & 1 \end{bmatrix} + \begin{bmatrix} \eta^2 & 0 \\ 0 & \eta^2 \end{bmatrix} = \begin{bmatrix} 1+\eta^2 & c \\ c & 1+\eta^2 \end{bmatrix}. \tag{2}$$

This implies

$$\overleftarrow{\mathbf{h}} = \begin{bmatrix} 1+\eta^2 & c \\ c & 1+\eta^2 \end{bmatrix}^{-1} \begin{bmatrix} c \\ 1 \end{bmatrix} \tag{3}$$

$$= \frac{1}{(1+\eta^2)^2 - c^2} \begin{bmatrix} 1+\eta^2 & -c \\ -c & 1+\eta^2 \end{bmatrix} \begin{bmatrix} c \\ 1 \end{bmatrix} \tag{4}$$

$$= \frac{1}{(1+\eta^2)^2 - c^2} \begin{bmatrix} c\eta^2 \\ 1+\eta^2 - c^2 \end{bmatrix} . \tag{5}$$

The optimal filter is

$$\mathbf{h} = \frac{1}{(1+\eta^2)^2 - c^2} \begin{bmatrix} 1+\eta^2 - c^2 \\ c\eta^2 \end{bmatrix} . \tag{6}$$

To find the mean square error of this predictor, we recall that Theorem 11.11 is just Theorem 9.7 expressed in the terminology of filters. Expressing part (c) of Theorem 9.7 in terms of the linear estimation filter $\mathbf{h}$, the mean square error of the estimator is

$$e_L^* = \mathrm{Var}[X_n] - \overleftarrow{\mathbf{h}}' \mathbf{R}_{\mathbf{Y}_n X_n} \tag{7}$$

$$= \mathrm{Var}[X_n] - \overleftarrow{\mathbf{h}}' \mathbf{R}_{\mathbf{X}_n X_n} \tag{8}$$

$$= R_X[0] - \overleftarrow{\mathbf{h}}' \begin{bmatrix} c \\ 1 \end{bmatrix} \tag{9}$$

$$= 1 - \frac{c^2\eta^2 + \eta^2 + 1 - c^2}{(1+\eta^2)^2 - c^2} . \tag{10}$$

Note that we always find that $e_L^* < \mathrm{Var}[X_n] = 1$ simply because the optimal estimator cannot be worse than the blind estimator that ignores the observation $\mathbf{Y}_n$.

Problem 11.4.4 Solution

In this problem, we find the mean square estimation error of the optimal first order filter in Problem 11.4.3. This problem highlights a shortcoming of Theorem 11.11 in that the theorem doesn't explicitly provide the mean square error associated with the optimal filter $\mathbf{h}$. We recall from the discussion at the start of Section 11.4 that Theorem 11.11 is derived from Theorem 9.7 with

$$\overleftarrow{\mathbf{h}} = \mathbf{a} = \mathbf{R}_{\mathbf{Y}}^{-1} \mathbf{R}_{\mathbf{Y}X} = (\mathbf{R}_{\mathbf{X}_n} + \mathbf{R}_{\mathbf{W}_n})^{-1} \mathbf{R}_{\mathbf{X}_n X_n} \tag{1}$$

From Theorem 9.7, the mean square error of the filter output is

$$e_L^* = \mathrm{Var}[X] - \mathbf{a}' \mathbf{R}_{\mathbf{Y}X} \tag{2}$$

$$= R_X[0] - \overleftarrow{\mathbf{h}}' \mathbf{R}_{\mathbf{X}_n X_n} \tag{3}$$

$$= R_X[0] - \mathbf{R}_{\mathbf{X}_n X_n}' (\mathbf{R}_{\mathbf{X}_n} + \mathbf{R}_{\mathbf{W}_n})^{-1} \mathbf{R}_{\mathbf{X}_n X_n} \tag{4}$$

Equations (3) and (4) are general expressions for the means square error of the optimal linear filter that can be applied to any situation described by Theorem 11.11.

To apply this result to the problem at hand, we observe that $R_X[0] = c^0 = 1$ and that

$$\overleftarrow{\mathbf{h}} = \frac{1}{(1+\eta^2)^2 - c^2} \begin{bmatrix} c\eta^2 \\ 1+\eta^2 - c^2 \end{bmatrix}, \qquad \mathbf{R}_{\mathbf{X}_n X_n} = \begin{bmatrix} R_X[1] \\ R_X[0] \end{bmatrix} = \begin{bmatrix} c \\ 1 \end{bmatrix} \tag{5}$$

This implies

$$e_L^* = R_X[0] - \overleftarrow{\mathbf{h}}' \mathbf{R}_{\mathbf{X}_n X_n} \tag{6}$$

$$= 1 - \frac{1}{(1+\eta^2)^2 - c^2} \begin{bmatrix} c\eta^2 & 1+\eta^2 - c^2 \end{bmatrix} \begin{bmatrix} c \\ 1 \end{bmatrix} \tag{7}$$

$$= 1 - \frac{c^2\eta^2 + 1 + \eta^2 - c^2}{(1+\eta^2)^2 - c^2} \tag{8}$$

$$= \eta^2 \left(\frac{1 + \eta^2 - c^2}{(1+\eta^2)^2 - c^2} \right) \tag{9}$$

The remaining question is what value of c minimizes the mean square error e_L^*. The usual approach is to set the derivative $\frac{de_L^*}{dc}$ to zero. This would yield the *incorrect* answer $c = 0$. In fact, evaluating the second derivative at $c = 0$ shows that $\frac{d^2 e_L^*}{dc^2}\Big|_{c=0} < 0$. Thus the mean square error e_L^* is maximum at $c = 0$. For a more careful analysis, we observe that $e_L^* = \eta^2 f(x)$ where

$$f(x) = \frac{a - x}{a^2 - x}, \tag{10}$$

with $x = c^2$, and $a = 1 + \eta^2$. In this case, minimizing $f(x)$ is equivalent to minimizing the mean square error. Note that for $R_X[k]$ to be a respectable autocorrelation function, we must have $|c| \leq 1$. Thus we consider only values of x in the interval $0 \leq x \leq 1$. We observe that

$$\frac{df(x)}{dx} = -\frac{a^2 - a}{(a^2 - x)^2} \tag{11}$$

Since $a > 1$, the derivative is negative for $0 \leq x \leq 1$. This implies the mean square error is minimized by making x as large as possible, i.e., $x = 1$. Thus $c = 1$ minimizes the mean square error. In fact $c = 1$ corresponds to the autocorrelation function $R_X[k] = 1$ for all k. Since each X_n has zero expected value, every pair of sample X_n and X_m has correlation coefficient

$$\rho_{X_n, X_m} = \frac{\mathrm{Cov}[X_n, X_m]}{\sqrt{\mathrm{Var}[X_n]\,\mathrm{Var}[X_m]}} = \frac{R_X[n-m]}{R_X[0]} = 1. \tag{12}$$

That is, $c = 1$ corresponds to a degenerate process in which every pair of samples X_n and X_m are perfectly correlated. Physically, this corresponds to the case where where the random process X_n is generated by generating a sample of a random variable X and setting $X_n = X$ for all n. The observations are then of the form $Y_n = X + Z_n$. That is, each observation is just a noisy observation of the random variable X. For $c = 1$, the optimal filter

$$\mathbf{h} = \frac{1}{2 + \eta^2} \begin{bmatrix} 1 \\ 1 \end{bmatrix}. \tag{13}$$

is just an equally weighted average of the past two samples.

Problem 11.4.5 Solution
This problem involves both linear estimation and prediction because we are using Y_{n-1}, the noisy observation of X_{n_1} to estimate the future value X_n. Thus we can't follow the cookbook recipes of Theorem 11.9 or Theorem 11.11. Instead we go back to Theorem 9.4 to find the minimum mean

square error linear estimator. In that theorem, X_n and Y_{n-1} play the roles of X and Y. That is, our estimate $\hat{X}_n$ of X_n is

$$\hat{X}_n = \hat{X}_L(Y_{n-1}) = \rho_{X_n, Y_{n-1}} \left(\frac{\text{Var}[X_n]}{\text{Var}[Y_{n-1}]} \right)^{1/2} (Y_{n-1} - E[Y_{n-1}]) + E[X_n] \tag{1}$$

By recursive application of $X_n = cX_{n-1} + Z_{n-1}$, we obtain

$$X_n = a^n X_0 + \sum_{j=1}^{n} a^{j-1} Z_{n-j} \tag{2}$$

The expected value of X_n is $E[X_n] = a^n E[X_0] + \sum_{j=1}^{n} a^{j-1} E[Z_{n-j}] = 0$. The variance of X_n is

$$\text{Var}[X_n] = a^{2n} \text{Var}[X_0] + \sum_{j=1}^{n} [a^{j-1}]^2 \text{Var}[Z_{n-j}] = a^{2n} \text{Var}[X_0] + \sigma^2 \sum_{j=1}^{n} [a^2]^{j-1} \tag{3}$$

Since $\text{Var}[X_0] = \sigma^2/(1 - c^2)$, we obtain

$$\text{Var}[X_n] = \frac{c^{2n}\sigma^2}{1 - c^2} + \frac{\sigma^2(1 - c^{2n})}{1 - c^2} = \frac{\sigma^2}{1 - c^2} \tag{4}$$

Note that $E[Y_{n-1}] = dE[X_{n-1}] + E[W_n] = 0$. The variance of Y_{n-1} is

$$\text{Var}[Y_{n-1}] = d^2 \text{Var}[X_{n-1}] + \text{Var}[W_n] = \frac{d^2\sigma^2}{1 - c^2} + \eta^2 \tag{5}$$

Since X_n and Y_{n-1} have zero mean, the covariance of X_n and Y_{n-1} is

$$\text{Cov}[X_n, Y_{n-1}] = E[X_n Y_{n-1}] = E[(cX_{n-1} + Z_{n-1})(dX_{n-1} + W_{n-1})] \tag{6}$$

From the problem statement, we learn that

$$E[X_{n-1}W_{n-1}] = 0 \qquad\qquad E[X_{n-1}]E[W_{n-1}] = 0$$
$$E[Z_{n-1}X_{n-1}] = 0 \qquad\qquad E[Z_{n-1}W_{n-1}] = 0$$

Hence, the covariance of X_n and Y_{n-1} is

$$\text{Cov}[X_n, Y_{n-1}] = cd\,\text{Var}[X_{n-1}] \tag{7}$$

The correlation coefficient of X_n and Y_{n-1} is

$$\rho_{X_n, Y_{n-1}} = \frac{\text{Cov}[X_n, Y_{n-1}]}{\sqrt{\text{Var}[X_n]\text{Var}[Y_{n-1}]}} \tag{8}$$

Since $E[Y_{n-1}]$ and $E[X_n]$ are zero, the linear predictor for X_n becomes

$$\hat{X}_n = \rho_{X_n, Y_{n-1}} \left(\frac{\text{Var}[X_n]}{\text{Var}[Y_{n-1}]} \right)^{1/2} Y_{n-1} = \frac{\text{Cov}[X_n, Y_{n-1}]}{\text{Var}[Y_{n-1}]} Y_{n-1} = \frac{cd\,\text{Var}[X_{n-1}]}{\text{Var}[Y_{n-1}]} Y_{n-1} \tag{9}$$

Substituting the above result for $\text{Var}[X_n]$, we obtain the optimal linear predictor of X_n given Y_{n-1}.

$$\hat{X}_n = \frac{c}{d} \frac{1}{1 + \beta^2(1 - c^2)} Y_{n-1} \tag{10}$$

where $\beta^2 = \eta^2/(d^2\sigma^2)$. From Theorem 9.4, the mean square estimation error at step n

$$e_L^*(n) = E[(X_n - \hat{X}_n)^2] = \text{Var}[X_n](1 - \rho_{X_n, Y_{n-1}}^2) = \sigma^2 \frac{1 + \beta^2}{1 + \beta^2(1 - c^2)} \tag{11}$$

We see that mean square estimation error $e_L^*(n) = e_L^*$, a constant for all n. In addition, e_L^* is an increasing function β.

Problem 11.5.1 Solution

To use Table 11.1, we write $R_X(\tau)$ in terms of the autocorrelation
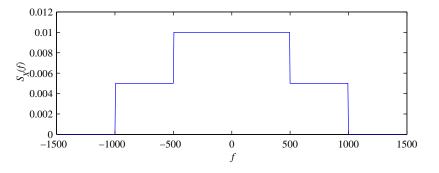
$$\operatorname{sinc}(x) = \frac{\sin(\pi x)}{\pi x}. \tag{1}$$

In terms of the $\operatorname{sinc}(\cdot)$ function, we obtain

$$R_X(\tau) = 10\operatorname{sinc}(2000\tau) + 5\operatorname{sinc}(1000\tau). \tag{2}$$

From Table 11.1,

$$S_X(f) = \frac{10}{2{,}000}\operatorname{rect}\left(\frac{f}{2000}\right) + \frac{5}{1{,}000}\operatorname{rect}\left(\frac{f}{1{,}000}\right) \tag{3}$$

Here is a graph of the PSD.

Problem 11.5.2 Solution

The process $Y(t)$ has expected value $E[Y(t)] = 0$. The autocorrelation of $Y(t)$ is

$$R_Y(t,\tau) = E[Y(t)Y(t+\tau)] = E[X(\alpha t)X(\alpha(t+\tau))] = R_X(\alpha\tau) \tag{1}$$

Thus $Y(t)$ is wide sense stationary. The power spectral density is

$$S_Y(f) = \int_{-\infty}^{\infty} R_X(\alpha\tau)e^{-j2\pi f\tau}\, d\tau. \tag{2}$$

At this point, we consider the cases $\alpha > 0$ and $\alpha < 0$ separately. For $\alpha > 0$, the substitution $\tau' = \alpha\tau$ yields

$$S_Y(f) = \frac{1}{\alpha}\int_{-\infty}^{\infty} R_X(\tau')e^{-j2\pi(f/\alpha)\tau'}\, d\tau' = \frac{S_X(f/\alpha)}{\alpha} \tag{3}$$

When $\alpha < 0$, we start with Equation (2) and make the substitution $\tau' = -\alpha\tau$, yielding

$$S_Y(f) = \frac{1}{-\alpha}\int_{-\infty}^{\infty} R_X(-\tau')e^{-j2\pi\frac{f}{-\alpha}\tau'}\, d\tau'. \tag{4}$$

Since $R_X(-\tau') = R_X(\tau')$,

$$S_Y(f) = \frac{1}{-\alpha}\int_{-\infty}^{\infty} R_X(\tau')e^{-j2\pi\frac{f}{-\alpha}\tau'}\, d\tau'. = \frac{1}{-\alpha}S_X\left(\frac{f}{-\alpha}\right) \tag{5}$$

For $-\alpha = |\alpha|$ for $\alpha < 0$, we can combine the $\alpha > 0$ and $\alpha < 0$ cases in the expression

$$S_Y(f) = \frac{1}{|\alpha|}S_X\left(\frac{f}{|\alpha|}\right). \tag{6}$$

398

Problem 11.6.1 Solution

Since the random sequence X_n has autocorrelation function

$$R_X[k] = \delta_k + (0.1)^{|k|}, \tag{1}$$

We can find the PSD directly from Table 11.2 with $0.1^{|k|}$ corresponding to $a^{|k|}$. The table yields

$$S_X(\phi) = 1 + \frac{1 - (0.1)^2}{1 + (0.1)^2 - 2(0.1)\cos 2\pi\phi} = \frac{2 - 0.2\cos 2\pi\phi}{1.01 - 0.2\cos 2\pi\phi}. \tag{2}$$

Problem 11.7.1 Solution

First we show that $S_{YX}(f) = S_{XY}(-f)$. From the definition of the cross spectral density,

$$S_{YX}(f) = \int_{-\infty}^{\infty} R_{YX}(\tau)e^{-j2\pi f\tau}\,d\tau \tag{1}$$

Making the subsitution $\tau' = -\tau$ yields

$$S_{YX}(f) = \int_{-\infty}^{\infty} R_{YX}(-\tau')e^{j2\pi f\tau'}\,d\tau' \tag{2}$$

By Theorem 10.14, $R_{YX}(-\tau') = R_{XY}(\tau')$. This implies

$$S_{YX}(f) = \int_{-\infty}^{\infty} R_{XY}(\tau')e^{-j2\pi(-f)\tau'}\,d\tau' = S_{XY}(-f) \tag{3}$$

To complete the problem, we need to show that $S_{XY}(-f) = [S_{XY}(f)]^*$. First we note that since $R_{XY}(\tau)$ is real valued, $[R_{XY}(\tau)]^* = R_{XY}(\tau)$. This implies

$$[S_{XY}(f)]^* = \int_{-\infty}^{\infty} [R_{XY}(\tau)]^*[e^{-j2\pi f\tau}]^*\,d\tau \tag{4}$$

$$= \int_{-\infty}^{\infty} R_{XY}(\tau)e^{-j2\pi(-f)\tau}\,d\tau \tag{5}$$

$$= S_{XY}(-f) \tag{6}$$

Problem 11.8.1 Solution

Let $a = 1/RC$. The solution to this problem parallels Example 11.22.

(a) From Table 11.1, we observe that

$$S_X(f) = \frac{2 \cdot 10^4}{(2\pi f)^2 + 10^4} \qquad H(f) = \frac{1}{a + j2\pi f} \tag{1}$$

By Theorem 11.16,

$$S_Y(f) = |H(f)|^2 S_X(f) = \frac{2 \cdot 10^4}{[(2\pi f)^2 + a^2][(2\pi f)^2 + 10^4]} \tag{2}$$

To find $R_Y(\tau)$, we use a form of partial fractions expansion to write

$$S_Y(f) = \frac{A}{(2\pi f)^2 + a^2} + \frac{B}{(2\pi f)^2 + 10^4} \tag{3}$$

399

Note that this method will work only if $a \neq 100$. This same method was also used in Example 11.22. The values of A and B can be found by

$$A = \left.\frac{2 \cdot 10^4}{(2\pi f)^2 + 10^4}\right|_{f=\frac{ja}{2\pi}} = \frac{-2 \cdot 10^4}{a^2 - 10^4} \qquad B = \left.\frac{2 \cdot 10^4}{a^2 + 10^4}\right|_{f=\frac{j100}{2\pi}} = \frac{2 \cdot 10^4}{a^2 - 10^4} \tag{4}$$

This implies the output power spectral density is

$$S_Y(f) = \frac{-10^4/a}{a^2 - 10^4} \frac{2a}{(2\pi f)^2 + a^2} + \frac{1}{a^2 - 10^4} \frac{200}{(2\pi f)^2 + 10^4} \tag{5}$$

Since $e^{-c|\tau|}$ and $2c/((2\pi f)^2 + c^2)$ are Fourier transform pairs for any constant $c > 0$, we see that

$$R_Y(\tau) = \frac{-10^4/a}{a^2 - 10^4} e^{-a|\tau|} + \frac{100}{a^2 - 10^4} e^{-100|\tau|} \tag{6}$$

(b) To find $a = 1/(RC)$, we use the fact that

$$E\left[Y^2(t)\right] = 100 = R_Y(0) = \frac{-10^4/a}{a^2 - 10^4} + \frac{100}{a^2 - 10^4} \tag{7}$$

Rearranging, we find that a must satisfy

$$a^3 - (10^4 + 1)a + 100 = 0 \tag{8}$$

This cubic polynomial has three roots:

$$a = 100 \qquad a = -50 + \sqrt{2501} \qquad a = -50 - \sqrt{2501} \tag{9}$$

Recall that $a = 100$ is not a valid solution because our expansion of $S_Y(f)$ was not valid for $a = 100$. Also, we require $a > 0$ in order to take the inverse transform of $S_Y(f)$. Thus $a = -50 + \sqrt{2501} \approx 0.01$ and $RC \approx 100$.

Problem 11.8.2 Solution

(a) $R_W(\tau) = \delta(\tau)$ is the autocorrelation function whose Fourier transform is $S_W(f) = 1$.

(b) The output $Y(t)$ has power spectral density

$$S_Y(f) = |H(f)|^2 S_W(f) = |H(f)|^2 \tag{1}$$

(c) Since $|H(f)| = 1$ for $f \in [-B, B]$, the average power of $Y(t)$ is

$$E\left[Y^2(t)\right] = \int_{-\infty}^{\infty} S_Y(f) \, df = \int_{-B}^{B} df = 2B \tag{2}$$

(d) Since the white noise $W(t)$ has zero mean, the mean value of the filter output is

$$E\left[Y(t)\right] = E\left[W(t)\right] H(0) = 0 \tag{3}$$

Problem 11.8.3 Solution

Since $S_Y(f) = |H(f)|^2 S_X(f)$, we first find

$$|H(f)|^2 = H(f)H^*(f) \tag{1}$$

$$= \left(a_1 e^{-j2\pi f t_1} + a_2 e^{-j2\pi f t_2}\right)\left(a_1 e^{j2\pi f t_1} + a_2 e^{j2\pi f t_2}\right) \tag{2}$$

$$= a_1^2 + a_2^2 + a_1 a_2 \left(e^{-j2\pi f(t_2 - t_1)} + e^{-j2\pi f(t_1 - t_2)}\right) \tag{3}$$

It follows that the output power spectral density is

$$S_Y(f) = (a_1^2 + a_2^2)S_X(f) + a_1 a_2 S_X(f) e^{-j2\pi f(t_2 - t_1)} + a_1 a_2 S_X(f) e^{-j2\pi f(t_1 - t_2)} \tag{4}$$

Using Table 11.1, the autocorrelation of the output is

$$R_Y(\tau) = (a_1^2 + a_2^2)R_X(\tau) + a_1 a_2 \left(R_X(\tau - (t_1 - t_2)) + R_X(\tau + (t_1 - t_2))\right) \tag{5}$$

Problem 11.8.4 Solution

(a) The average power of the input is

$$E\left[X^2(t)\right] = R_X(0) = 1 \tag{1}$$

(b) From Table 11.1, the input has power spectral density

$$S_X(f) = \frac{1}{2}e^{-\pi f^2/4} \tag{2}$$

The output power spectral density is

$$S_Y(f) = |H(f)|^2 S_X(f) = \begin{cases} \dfrac{1}{2}e^{-\pi f^2/4} & |f| \leq 2 \\ 0 & \text{otherwise} \end{cases} \tag{3}$$

(c) The average output power is

$$E\left[Y^2(t)\right] = \int_{-\infty}^{\infty} S_Y(f)\, df = \frac{1}{2}\int_{-2}^{2} e^{-\pi f^2/4}\, df \tag{4}$$

This integral cannot be expressed in closed form. However, we can express it in the form of the integral of a standardized Gaussian PDF by making the substitution $f = z\sqrt{2/\pi}$. With this subsitution,

$$E\left[Y^2(t)\right] = \frac{1}{\sqrt{2\pi}}\int_{-\sqrt{2\pi}}^{\sqrt{2\pi}} e^{-z^2/2}\, dz \tag{5}$$

$$= \Phi(\sqrt{2\pi}) - \Phi(-\sqrt{2\pi}) \tag{6}$$

$$= 2\Phi(\sqrt{2\pi}) - 1 = 0.9876 \tag{7}$$

The output power almost equals the input power because the filter bandwidth is sufficiently wide to pass through nearly all of the power of the input.

Problem 11.8.5 Solution

(a) From Theorem 11.13(b),

$$E\left[X^2(t)\right] = \int_{-\infty}^{\infty} S_X(f)\, df \int_{-100}^{100} 10^{-4}\, df = 0.02 \tag{1}$$

(b) From Theorem 11.17

$$S_{XY}(f) = H(f)S_X(f) = \begin{cases} 10^{-4}H(f) & |f| \leq 100 \\ 0 & \text{otherwise} \end{cases} \tag{2}$$

(c) From Theorem 10.14,

$$R_{YX}(\tau) = R_{XY}(-\tau) \tag{3}$$

From Table 11.1, if $g(\tau)$ and $G(f)$ are a Fourier transform pair, then $g(-\tau)$ and $G^*(f)$ are a Fourier transform pair. This implies

$$S_{YX}(f) = S_{XY}^*(f) = \begin{cases} 10^{-4}H^*(f) & |f| \leq 100 \\ 0 & \text{otherwise} \end{cases} \tag{4}$$

(d) By Theorem 11.17,

$$S_Y(f) = H^*(f)S_{XY}(f) = |H(f)|^2\, S_X(f) \tag{5}$$

$$= \begin{cases} 10^{-4}/[10^4\pi^2 + (2\pi f)^2] & |f| \leq 100 \\ 0 & \text{otherwise} \end{cases} \tag{6}$$

(e) By Theorem 11.13,

$$E\left[Y^2(t)\right] = \int_{-\infty}^{\infty} S_Y(f)\, df = \int_{-100}^{100} \frac{10^{-4}}{10^4\pi^2 + 4\pi^2 f^2}\, df \tag{7}$$

$$= \frac{2}{10^8\pi^2} \int_0^{100} \frac{df}{1 + (f/50)^2} \tag{8}$$

By making the substitution, $f = 50\tan\theta$, we have $df = 50\sec^2\theta\, d\theta$. Using the identity $1 + \tan^2\theta = \sec^2\theta$, we have

$$E\left[Y^2(t)\right] = \frac{100}{10^8\pi^2} \int_0^{\tan^{-1}(2)} d\theta = \frac{\tan^{-1}(2)}{10^6\pi^2} = 1.12 \times 10^{-7} \tag{9}$$

Problem 11.8.6 Solution

The easy way to do this problem is to use Theorem 11.17 which states

$$S_{XY}(f) = H(f)S_X(f) \tag{1}$$

(a) From Table 11.1, we observe that

$$S_X(f) = \frac{8}{16 + (2\pi f)^2} \qquad H(f) = \frac{1}{7 + j2\pi f} \tag{2}$$

From Theorem 11.17,

$$S_{XY}(f) = H(f)S_X(f) = \frac{8}{[7 + j2\pi f][16 + (2\pi f)^2]} \tag{3}$$

(b) To find the cross correlation, we need to find the inverse Fourier transform of $S_{XY}(f)$. A straightforward way to do this is to use a partial fraction expansion of $S_{XY}(f)$. That is, by defining $s = j2\pi f$, we observe that

$$\frac{8}{(7 + s)(4 + s)(4 - s)} = \frac{-8/33}{7 + s} + \frac{1/3}{4 + s} + \frac{1/11}{4 - s} \tag{4}$$

Hence, we can write the cross spectral density as

$$S_{XY}(f) = \frac{-8/33}{7 + j2\pi f} + \frac{1/3}{4 + j2\pi f} + \frac{1/11}{4 - j\pi f} \tag{5}$$

Unfortunately, terms like $1/(a - j2\pi f)$ do not have an inverse transforms. The solution is to write $S_{XY}(f)$ in the following way:

$$S_{XY}(f) = \frac{-8/33}{7 + j2\pi f} + \frac{8/33}{4 + j2\pi f} + \frac{1/11}{4 + j2\pi f} + \frac{1/11}{4 - j2\pi f} \tag{6}$$

$$= \frac{-8/33}{7 + j2\pi f} + \frac{8/33}{4 + j2\pi f} + \frac{8/11}{16 + (2\pi f)^2} \tag{7}$$

$$\tag{8}$$

Now, we see from Table 11.1 that the inverse transform is

$$R_{XY}(\tau) = -\frac{8}{33}e^{-7\tau}u(\tau) + \frac{8}{33}e^{-4\tau}u(\tau) + \frac{1}{11}e^{-4|\tau|} \tag{9}$$

Problem 11.8.7 Solution

(a) Since $E[N(t)] = \mu_N = 0$, the expected value of the output is $\mu_Y = \mu_N H(0) = 0$.

(b) The output power spectral density is

$$S_Y(f) = |H(f)|^2 S_N(f) = 10^{-3}e^{-2\times10^6|f|} \tag{1}$$

(c) The average power is

$$E\left[Y^2(t)\right] = \int_{-\infty}^{\infty} S_Y(f)\, df = \int_{-\infty}^{\infty} 10^{-3}e^{-2\times10^6|f|}\, df \tag{2}$$

$$= 2 \times 10^{-3} \int_{0}^{\infty} e^{-2\times10^6 f}\, df \tag{3}$$

$$= 10^{-3} \tag{4}$$

(d) Since $N(t)$ is a Gaussian process, Theorem 11.3 says $Y(t)$ is a Gaussian process. Thus the random variable $Y(t)$ is Gaussian with

$$E\left[Y(t)\right] = 0 \qquad \mathrm{Var}[Y(t)] = E\left[Y^2(t)\right] = 10^{-3} \tag{5}$$

Thus we can use Table 3.1 to calculate

$$P\left[Y(t) > 0.01\right] = P\left[\frac{Y(t)}{\sqrt{\mathrm{Var}[Y(t)]}} > \frac{0.01}{\sqrt{\mathrm{Var}[Y(t)]}}\right] \tag{6}$$

$$1 - \Phi\left(\frac{0.01}{\sqrt{0.001}}\right) \tag{7}$$

$$= 1 - \Phi(0.32) = 0.3745 \tag{8}$$

Problem 11.8.8 Solution

Suppose we assume that $N(t)$ and $Y(t)$ are the input and output of a linear time invariant filter $h(u)$. In that case,

$$Y(t) = \int_0^t N(u)\,du = \int_{-\infty}^\infty h(t-u)N(u)\,du \tag{1}$$

For the above two integrals to be the same, we must have

$$h(t-u) = \begin{cases} 1 & 0 \le t - u \le t \\ 0 & \text{otherwise} \end{cases} \tag{2}$$

Making the substitution $v = t - u$, we have

$$h(v) = \begin{cases} 1 & 0 \le v \le t \\ 0 & \text{otherwise} \end{cases} \tag{3}$$

Thus the impulse response $h(v)$ depends on t. That is, the filter response is linear but not time invariant. Since Theorem 11.2 requires that $h(t)$ be time invariant, this example does not violate the theorem.

Problem 11.8.9 Solution

(a) Note that $|H(f)| = 1$. This implies $S_{\hat{M}}(f) = S_M(f)$. Thus the average power of $\hat{M}(t)$ is

$$\hat{q} = \int_{-\infty}^\infty S_{\hat{M}}(f)\,df = \int_{-\infty}^\infty S_M(f)\,df = q \tag{1}$$

(b) The average power of the upper sideband signal is

$$E\left[U^2(t)\right] = E\left[M^2(t)\cos^2(2\pi f_c t + \Theta)\right] \tag{2}$$

$$- E\left[2M(t)\hat{M}(t)\cos(2\pi f_c t + \Theta)\sin(2\pi f_c t + \Theta)\right] \tag{3}$$

$$+ E\left[\hat{M}^2(t)\sin^2(2\pi f_c t + \Theta)\right] \tag{4}$$

To find the expected value of the random phase cosine, for an integer $n \neq 0$, we evaluate

$$E\left[\cos(2\pi f_c t + n\Theta)\right] = \int_0^{2\pi} \cos(2\pi f_c t + n\theta)\frac{1}{2\pi}d\theta \tag{5}$$

$$= \frac{1}{2n\pi}\sin(2\pi f_c t + n\theta)\Big|_0^{2\pi} \tag{6}$$

$$= \frac{1}{2n\pi}\left(\sin(2\pi f_c t + 2n\pi) - \sin(2\pi f_c t)\right) = 0 \tag{7}$$

Similar steps will show that for any integer $n \neq 0$, the random phase sine also has expected value

$$E\left[\sin(2\pi f_c t + n\Theta)\right] = 0 \tag{8}$$

Using the trigonometric identity $\cos^2 \phi = (1 + \cos 2\phi)/2$, we can show

$$E\left[\cos^2(2\pi f_c t + \Theta)\right] = E\left[\frac{1}{2}\left(1 + \cos(2\pi(2f_c)t + 2\Theta)\right)\right] = 1/2 \tag{9}$$

Similarly,

$$E\left[\sin^2(2\pi f_c t + \Theta)\right] = E\left[\frac{1}{2}\left(1 - \cos(2\pi(2f_c)t + 2\Theta)\right)\right] = 1/2 \tag{10}$$

In addition, the identity $2\sin\phi\cos\phi = \sin 2\phi$ implies

$$E\left[2\sin(2\pi f_c t + \Theta)\cos(2\pi f_c t + \Theta)\right] = E\left[\cos(4\pi f_c t + 2\Theta)\right] = 0 \tag{11}$$

Since $M(t)$ and $\hat{M}(t)$ are independent of Θ, the average power of the upper sideband signal is

$$E\left[U^2(t)\right] = E\left[M^2(t)\right]E\left[\cos^2(2\pi f_c t + \Theta)\right] + E\left[\hat{M}^2(t)\right]E\left[\sin^2(2\pi f_c t + \Theta)\right] \tag{12}$$

$$- E\left[M(t)\hat{M}(t)\right]E\left[2\cos(2\pi f_c t + \Theta)\sin(2\pi f_c t + \Theta)\right] \tag{13}$$

$$= q/2 + q/2 + 0 = q \tag{14}$$

Problem 11.8.10 Solution

(a) Since $S_W(f) = 10^{-15}$ for all f, $R_W(\tau) = 10^{-15}\delta(\tau)$.

(b) Since Θ is independent of $W(t)$,

$$E\left[V(t)\right] = E\left[W(t)\cos(2\pi f_c t + \Theta)\right] = E\left[W(t)\right]E\left[\cos(2\pi f_c t + \Theta)\right] = 0 \tag{1}$$

(c) We cannot initially assume $V(t)$ is WSS so we first find

$$\begin{aligned}
R_V(t,\tau) &= E[V(t)V(t+\tau)] & (2)\\
&= E[W(t)\cos(2\pi f_c t + \Theta)W(t+\tau)\cos(2\pi f_c(t+\tau) + \Theta)] & (3)\\
&= E[W(t)W(t+\tau)]E[\cos(2\pi f_c t + \Theta)\cos(2\pi f_c(t+\tau) + \Theta)] & (4)\\
&= 10^{-15}\delta(\tau)E[\cos(2\pi f_c t + \Theta)\cos(2\pi f_c(t+\tau) + \Theta)] & (5)
\end{aligned}$$

We see that for all $\tau \neq 0$, $R_V(t, t + \tau) = 0$. Thus we need to find the expected value of

$$E\left[\cos(2\pi f_c t + \Theta)\cos(2\pi f_c(t + \tau) + \Theta)\right] \tag{6}$$

only at $\tau = 0$. However, its good practice to solve for arbitrary τ:

$$E[\cos(2\pi f_c t + \Theta)\cos(2\pi f_c(t + \tau) + \Theta)] \tag{7}$$

$$= \frac{1}{2}E[\cos(2\pi f_c \tau) + \cos(2\pi f_c(2t + \tau) + 2\Theta)] \tag{8}$$

$$= \frac{1}{2}\cos(2\pi f_c \tau) + \frac{1}{2}\int_0^{2\pi} \cos(2\pi f_c(2t + \tau) + 2\theta)\frac{1}{2\pi}\, d\theta \tag{9}$$

$$= \frac{1}{2}\cos(2\pi f_c \tau) + \frac{1}{2}\sin(2\pi f_c(2t + \tau) + 2\theta)\Big|_0^{2\pi} \tag{10}$$

$$= \frac{1}{2}\cos(2\pi f_c \tau) + \frac{1}{2}\sin(2\pi f_c(2t + \tau) + 4\pi) - \frac{1}{2}\sin(2\pi f_c(2t + \tau)) \tag{11}$$

$$= \frac{1}{2}\cos(2\pi f_c \tau) \tag{12}$$

Consequently,

$$R_V(t, \tau) = \frac{1}{2}10^{-15}\delta(\tau)\cos(2\pi f_c \tau) = \frac{1}{2}10^{-15}\delta(\tau) \tag{13}$$

(d) Since $E[V(t)] = 0$ and since $R_V(t, \tau) = R_V(\tau)$, we see that $V(t)$ is a wide sense stationary process. Since $L(f)$ is a linear time invariant filter, the filter output $Y(t)$ is also a wide sense stationary process.

(e) The filter input $V(t)$ has power spectral density $S_V(f) = \frac{1}{2}10^{-15}$. The filter output has power spectral density

$$S_Y(f) = |L(f)|^2 S_V(f) = \begin{cases} 10^{-15}/2 & |f| \leq B \\ 0 & \text{otherwise} \end{cases} \tag{14}$$

The average power of $Y(t)$ is

$$E\left[Y^2(t)\right] = \int_{-\infty}^{\infty} S_Y(f)\, df = \int_{-B}^{B} \frac{1}{2}10^{-15}\, df = 10^{-15}B \tag{15}$$

Problem 11.9.1 Solution

The system described in this problem corresponds exactly to the system in the text that yielded Equation (11.146).

(a) From Equation (11.146), the optimal linear filter is

$$\hat{H}(f) = \frac{S_X(f)}{S_X(f) + S_N(f)} \tag{1}$$

In this problem, $R_X(\tau) = \text{sinc}(2W\tau)$ so that

$$S_X(f) = \frac{1}{2W}\text{rect}\left(\frac{f}{2W}\right). \tag{2}$$

It follows that the optimal filter is

$$\hat{H}(f) = \frac{\frac{1}{2W} \operatorname{rect}\left(\frac{f}{2W}\right)}{\frac{1}{2W} \operatorname{rect}\left(\frac{f}{2W}\right) + 10^{-5}} = \frac{10^5}{10^5 + 2W} \operatorname{rect}\left(\frac{f}{2W}\right). \tag{3}$$

(b) From Equation (11.147), the minimum mean square error is

$$e_L^* = \int_{-\infty}^{\infty} \frac{S_X(f) S_N(f)}{S_X(f) + S_N(f)} \, df = \int_{-\infty}^{\infty} \hat{H}(f) S_N(f) \, df \tag{4}$$

$$= \frac{10^5}{10^5 + 2W} \int_{-W}^{W} 10^{-5} \, df \tag{5}$$

$$= \frac{2W}{10^5 + 2W}. \tag{6}$$

It follows that the mean square error satisfies $e_L^* \leq 0.04$ if and only if $W \leq 2{,}083.3$ Hz. What is occurring in this problem is the optimal filter is simply an ideal lowpass filter of bandwidth W. Increasing W increases the bandwidth of the signal and the bandwidth of the filter $\hat{H}(f)$. This allows more noise to pass through the filter and decreases the quality of our estimator.

Problem 11.9.2 Solution

The system described in this problem corresponds exactly to the system in the text that yielded Equation (11.146).

(a) From Equation (11.146), the optimal linear filter is

$$\hat{H}(f) = \frac{S_X(f)}{S_X(f) + S_N(f)} \tag{1}$$

In this problem, $R_X(\tau) = e^{-5000|\tau|}$ so that

$$S_X(f) = \frac{10^4}{(5{,}000)^2 + (2\pi f)^2}. \tag{2}$$

It follows that the optimal filter is

$$\hat{H}(f) = \frac{\frac{10^4}{(5{,}000)^2+(2\pi f)^2}}{\frac{10^4}{(5{,}000)^2+(2\pi f)^2} + 10^{-5}} = \frac{10^9}{1.025 \times 10^9 + (2\pi f)^2}. \tag{3}$$

From Table 11.2, we see that the filter $\hat{H}(f)$ has impulse response

$$\hat{h}(\tau) = \frac{10^9}{2\alpha} e^{-\alpha|\tau|} \tag{4}$$

where $\alpha = \sqrt{1.025 \times 10^9} = 3.20 \times 10^4$.

(b) From Equation (11.147), the minimum mean square error is

$$e_L^* = \int_{-\infty}^{\infty} \frac{S_X(f) S_N(f)}{S_X(f) + S_N(f)} \, df = \int_{-\infty}^{\infty} \hat{H}(f) S_N(f) \, df \tag{5}$$

$$= 10^{-5} \int_{-\infty}^{\infty} \hat{H}(f) \, df \tag{6}$$

$$= 10^{-5} \hat{h}(0) = \frac{10^4}{2\alpha} = 0.1562. \tag{7}$$

Problem 11.10.1 Solution

Although it is straightforward to calculate sample paths of Y_n using the filter response $Y_n = \frac{1}{2}Y_{n-1} + \frac{1}{2}X_n$ directly, the necessary loops makes for a slow program. A solution using vectors and matrices tends to run faster. From the filter response, we can write

$$Y_1 = \frac{1}{2}X_1 \tag{1}$$

$$Y_2 = \frac{1}{4}X_1 + \frac{1}{2}X_2 \tag{2}$$

$$Y_3 = \frac{1}{8}X_1 + \frac{1}{4}X_2 + \frac{1}{2}X_3 \tag{3}$$

$$\vdots \tag{4}$$

$$Y_n = \frac{1}{2^n}X_1 + \frac{1}{2^{n-1}}X_2 + \cdots + \frac{1}{2}X_n \tag{5}$$

In vector notation, these equations become

$$\underbrace{\begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix}}_{\mathbf{Y}} = \underbrace{\begin{bmatrix} 1/2 & 0 & \cdots & 0 \\ 1/4 & 1/2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 1/2^n & \cdots & 1/4 & 1/2 \end{bmatrix}}_{\mathbf{H}} \underbrace{\begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_n \end{bmatrix}}_{\mathbf{X}}. \tag{6}$$
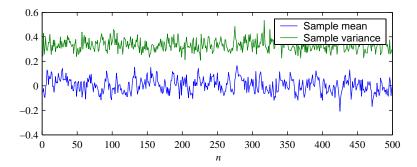
When $\mathbf{X}$ is a column of iid Gaussian $(0, 1)$ random variables, the column vector $\mathbf{Y} = \mathbf{HX}$ is a single sample path of $Y_1, \ldots, Y_n$. When $\mathbf{X}$ is an $n \times m$ matrix of iid Gaussian $(0, 1)$ random variables, each column of $\mathbf{Y} = \mathbf{HX}$ is a sample path of $Y_1, \ldots, Y_n$. In this case, let matrix entry $Y_{i,j}$ denote a sample Y_i of the jth sample path. The samples $Y_{i,1}, Y_{i,2}, \ldots, Y_{i,m}$ are iid samples of Y_i. We can estimate the mean and variance of Y_i using the sample mean $M_n(Y_i)$ and sample variance $V_m(Y_i)$ of Section 7.3. These estimates are

$$M_n(Y_i) = \frac{1}{m}\sum_{j=1}^{m}Y_{i,j}, \qquad\qquad V(Y_i) = \frac{1}{m-1}\sum_{j=1}^{m}(Y_{i,j} - M_n(Y_i))^2 \tag{7}$$

This is the approach of the following program.

```
function ymv=yfilter(m);
%ymv(i) is the mean and var (over m paths) of y(i),
%the filter output of 11.2.6 and 11.10.1
X=randn(500,m);
H=toeplitz([(0.5).^(1:500)],[0.5 zeros(1,499)]);
Y=H*X;
yav=sum(Y,2)/m;
yavmat=yav*ones(1,m);
yvar=sum((Y-yavmat).^2,2)/(m-1);
ymv=[yav yvar];
```

The commands ymv=yfilter(100);plot(ymv) will generate a plot similar to this:

We see that each sample mean is small, on the other order of 0.1. Note that $E[Y_i] = 0$. For $m = 100$ samples, the sample mean has variance $1/m = 0.01$ and standard deviation 0.1. Thus it is to be expected that we observe sample mean values around 0.1.

Also, it can be shown (in the solution to Problem 11.2.6 for example) that as i becomes large, $\mathrm{Var}[Y_i]$ converges to $1/3$. Thus our sample variance results are also not surprising.

Comment: Although within each sample path, Y_i and Y_{i+1} are quite correlated, the sample means of Y_i and Y_{i+1} are not very correlated when a large number of sample paths are averaged. Exact calculation of the covariance of the sample means of Y_i and Y_{i+1} might be an interesting exercise. The same observations apply to the sample variance as well.
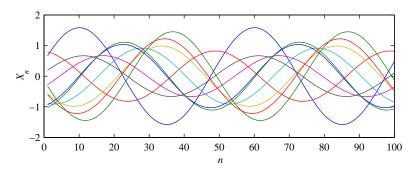
Problem 11.10.2 Solution

This is just a MATLAB question that has nothing to do with probability. In the MATLAB operation R=fft(r,N), the shape of the output R is the same as the shape of the input r. If r is a column vector, then R is a column vector. If r is a row vector, then R is a row vector. For fftc to work the same way, the shape of n must be the same as the shape of R. The instruction n=reshape(0:(N-1),size(R)) does this.
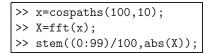
Problem 11.10.3 Solution

The program cospaths.m generates Gaussian sample paths with the desired autocorrelation function $R_X(k) = \cos(0.04 * pi * k)$. Here is the code:

```
function x=cospaths(n,m);
%Generate m  sample paths of length n of a
%Gaussian process with ACF R[k]=cos(0.04*pi*k)
k=0:n-1;
rx=cos(0.04*pi*k)';
x=gaussvector(0,rx,m);
```

The program is simple because if the second input parameter to gaussvector is a length m vector rx, then rx is assumed to be the first row of a symmetric Toeplitz covariance matrix. The commands x=cospaths(100,10);plot(x) will produce a graph like this one:
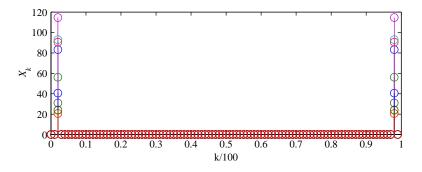

409

We note that every sample path of the process is Gaussian random sequence. However, it would also appear from the graph that every sample path is a perfect sinusoid. this may seem strange if you are used to seeing Gaussian processes simply as noisy processes or fluctuating Brownian motion processes. However, in this case, the amplitude and phase of each sample path is random such that over the ensemble of sinusoidal sample functions, each sample X_n is a Gaussian $(0, 1)$ random variable.

Finally, to confirm that that each sample path is a perfect sinusoid, rather than just resembling a sinusoid, we calculate the DFT of each sample path. The commands

```
>> x=cospaths(100,10);
>> X=fft(x);
>> stem((0:99)/100,abs(X));
```

will produce a plot similar to this:

The above plot consists of ten overlaid 100-point DFT magnitude stem plots, one for each Gaussian sample function. Each plot has exactly two nonzero components at frequencies $k/100 = 0.02$ and $(100 - k)/100 = 0.98$ corresponding to each sample path sinusoid having frequency 0.02. Note that the magnitude of each 0.02 frequency component depends on the magnitude of the corresponding sinusoidal sample path.

Problem 11.10.4 Solution

Searching the MATLAB full product help for `inv` yields this bit of advice:

> In practice, it is seldom necessary to form the explicit inverse of a matrix. A frequent misuse of `inv` arises when solving the system of linear equations . One way to solve this is with $x = \mathtt{inv}(A) * b$. A better way, from both an execution time and numerical accuracy standpoint, is to use the matrix division operator $x = A \backslash b$. This produces the solution using Gaussian elimination, without forming the inverse. See \ and / for further information.

The same discussion goes on to give an example where $x = A \backslash b$ is both faster and more accurate.

Problem 11.10.5 Solution

The function `lmsepredictor.m` is designed so that if the sequence X_n has a finite duration autocorrelation function such that $R_X[k] = 0$ for $|k| \geq m$, but the LMSE filter of order $M - 1$ for $M \geq m$ is supposed to be returned, then the `lmsepredictor` automatically pads the autocorrelation vector `rx` with a sufficient number of zeros so that the output is the order $M - 1$ filter. Conversely, if `rx` specifies more values $R_X[k]$ than are needed, then the operation `rx(1:M)` extracts the M values $R_X[0], \ldots, R_X[M - 1]$ that are needed.

410

However, in this problem $R_X[k] = (-0.9)^{|k|}$ has infinite duration. When we pass the truncated representation `rx` of length $m = 6$ and request `lmsepredictor(rx,M)` for $M \geq 6$, the result is that `rx` is incorrectly padded with zeros. The resulting filter output will be the LMSE filter for the filter response

$$R_X[k] = \begin{cases} (-0.9)^{|k|} & |k| \leq 5, \\ 0 & \text{otherwise,} \end{cases} \tag{1}$$

rather than the LMSE filter for the true autocorrelation function.

Problem 11.10.6 Solution

Applying Theorem 11.4 with sampling period $T_s = 1/4000$ s yields

$$R_X[k] = R_X(kT_s) = 10\operatorname{sinc}(0.5k) + 5\operatorname{sinc}(0.25k). \tag{1}$$

To find the power spectral density $S_X(\phi)$, we need to find the DTFT of $\operatorname{sinc}(\phi_0 k)$ Unfortunately, this was omitted from Table 11.2 so we now take a detour and derive it here. As with any derivation of the transform of a sinc function, we guess the answer and calculate the inverse transform. In this case, suppose

$$S_X(\phi) = \frac{1}{\phi_0}\operatorname{rect}(\phi/\phi_0) = \begin{cases} 1 & |\phi| \leq \phi_0/2, \\ 0 & \text{otherwise.} \end{cases} \tag{2}$$

We find $R_X[k]$ from the inverse DTFT. For $|\phi_0| \leq 1$,

$$R_X[k] = \int_{-1/2}^{1/2} S_X(\phi)\, e^{j2\pi\phi k}\, d\phi = \frac{1}{\phi_0}\int_{-\phi_0/2}^{\phi_0/2} e^{j2\pi\phi k}\, d\phi = \frac{1}{\phi_0}\frac{e^{j\pi\phi_0 k} - e^{-j\pi\phi_0 k}}{j2\pi k} = \operatorname{sinc}(\phi_0 k) \tag{3}$$
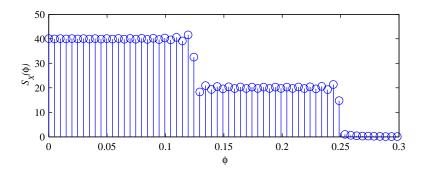
Now we apply this result to take the transform of $R_X[k]$ in Equation (1). This yields

$$S_X(\phi) = \frac{10}{0.5}\operatorname{rect}(\phi/0.5) + \frac{5}{0.25}\operatorname{rect}(\phi/0.25). \tag{4}$$

Ideally, an $2N + 1$-point DFT would yield a sampled version of the DTFT $S_X(\phi)$. However, the truncation of the autocorrelation $R_X[k]$ to 201 points results in a difference. For $N = 100$, the DFT will be a sampled version of the DTFT of $R_X[k]\operatorname{rect}(k/(2N+1))$. Here is a MATLAB program that shows the difference when the autocorrelation is truncated to $2N + 1$ terms.

```
function DFT=twosincsdft(N);
%Usage: SX=twosincsdft(N);
%Returns and plots the 2N+1
%point DFT of R(-N) ... R(0) ... R(N)
%for ACF R[k] in Problem 11.2.2
k=-N:N;
rx=10*sinc(0.5*k) + 5*sinc(0.25*k);
DFT=fftc(rx);
M=ceil(0.6*N);
phi=(0:M)/(2*N+1);
stem(phi,abs(DFT(1:(M+1))));
xlabel('\it \phi');
ylabel('\it S_X(\phi)');
```

Here is the output of `twosincsdft(100)`.

From the stem plot of the DFT, it is easy to see the deviations from the two rectangles that make up the DTFT $S_X(\phi)$. We see that the effects of windowing are particularly pronounced at the break points.

Comment: In `twosincsdft`, DFT must be real-valued since it is the DFT of an autocorrelation function. Hence the command `stem(DFT)` should be sufficient. However, due to numerical precision issues, the actual `DFT` tends to have a tiny imaginary hence we use the `abs` operator.

Problem 11.10.7 Solution
In this problem, we generalize the solution of Problem 11.4.1 using Theorem 11.9 with $k = 1$ for filter order $M > 2$ The optimum linear predictor filter $\mathbf{h} = \begin{bmatrix} h_0 & h_1 & \cdots & h_{M-1} \end{bmatrix}'$ of X_{n+1} given $\mathbf{X}_n = \begin{bmatrix} X_{n-(M-1)} & \cdots & X_{n-1} & X_n \end{bmatrix}'$ is given by

$$\overleftarrow{\mathbf{h}} = \begin{bmatrix} h_{M-1} \\ \vdots \\ h_0 \end{bmatrix} = \mathbf{R}_{\mathbf{X}_n}^{-1} \mathbf{R}_{\mathbf{X}_n X_{n+1}}, \tag{1}$$

where $\mathbf{R}_{\mathbf{X}_n}$ is given by Theorem 11.6 and $\mathbf{R}_{\mathbf{X}_n X_{n+1}}$ is given by Equation (11.66). In this problem,

$$\mathbf{R}_{\mathbf{X}_n X_{n+1}} = E\left[\begin{bmatrix} X_{n-M+1} \\ \vdots \\ X_n \end{bmatrix} X_{n+1} \right] = \begin{bmatrix} R_X[M] \\ \vdots \\ R_X[1] \end{bmatrix}. \tag{2}$$

Going back to Theorem 9.7(a), the mean square error is
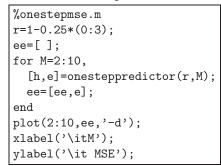
$$e_L^* = E\left[(X_{n+1} - \overleftarrow{h}'\mathbf{X}_n)^2 \right] = \text{Var}[X_{n+1}] - \overleftarrow{\mathbf{h}}'\mathbf{R}_{\mathbf{X}_n X_{n+1}}. \tag{3}$$

For $M > 2$, we use the MATLAB function `onesteppredictor(r,M)` to perform the calculations.

```
function [h,e]=onesteppredictor(r,M);
%usage: h=onesteppredictor(r,M);
%input: r=[R_X(0) R_X(1) .. R_X(m-1)]
%assumes R_X(n)==0 for n >=m
%output=vector h for lmse predictor
% xx=h'[X(n),X(n-1),..,X(n-M+1)] for X(n+1)
m=length(r);
r=[r(:);zeros(M-m+1,1)];%append zeros if needed
RY=toeplitz(r(1:M));
RYX=r(M+1:-1:2);
h=flipud(RY\RYX);
e=r(1)-(flipud(h))'*RYX;
```

412

The code is pretty straightforward. Here are two examples just to show it works.

```
>> [h2,e2]=onesteppredictor(r,2)
h2 =
    0.8571
   -0.1429
e2 =
    0.4286


>> [h4,e4]=onesteppredictor(r,4)
h4 =
    0.8000
    0.0000
   -0.0000
   -0.2000
e4 =
    0.4000
```

The problem also requested that we calculate the mean square error as a function of the filter order M. Here is a script and the resulting plot of the MSE.

```
%onestepmse.m
r=1-0.25*(0:3);
ee=[ ];
for M=2:10,
  [h,e]=onesteppredictor(r,M);
  ee=[ee,e];
end
plot(2:10,ee,'-d');
xlabel('\itM');
ylabel('\it MSE');
```


Problem 11.10.8 Solution

This problem generalizes the solution of Problem 11.10.7. to a k-step predictor. The optimum linear predictor filter $\mathbf{h} = \begin{bmatrix} h_0 & h_1 & \cdots & h_{M-1} \end{bmatrix}'$ of X_{n+1} given $\mathbf{X}_n = \begin{bmatrix} X_{n-(M-1)} & \cdots & X_{n-1} & X_n \end{bmatrix}'$ is given by

$$\overleftarrow{\mathbf{h}} = \begin{bmatrix} h_{M-1} \\ \vdots \\ h_0 \end{bmatrix} = \mathbf{R}_{\mathbf{X}_n}^{-1} \mathbf{R}_{\mathbf{X}_n X_{n+k}}, \tag{1}$$

where $\mathbf{R}_{\mathbf{X}_n}$ is given by Theorem 11.6 and $\mathbf{R}_{\mathbf{X}_n X_{n+k}}$ is given by Equation (11.66). In this problem,

$$\mathbf{R}_{\mathbf{X}_n X_{n+k}} = E\left[\begin{bmatrix} X_{n-M+1} \\ \vdots \\ X_n \end{bmatrix} X_{n+k}\right] = \begin{bmatrix} R_X\left[M+k-1\right] \\ \vdots \\ R_X\left[k\right] \end{bmatrix}. \tag{2}$$

Going back to Theorem 9.7(a), the mean square error is

$$e_L^* = E\left[(X_{n+k} - \overleftarrow{h}'\mathbf{X}_n)^2\right] = \text{Var}[X_{n+k}] - \overleftarrow{\mathbf{h}}'\mathbf{R}_{\mathbf{X}_n X_{n+k}}. \tag{3}$$
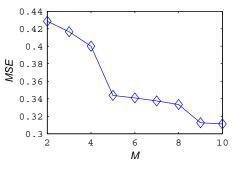
The MATLAB function kpredictor(r,M) implements this solution.

```
function h=kpredictor(r,M,k);
%usage: h=kpredictor(r,M,k);
%input: r=[R_X(0) R_X(1) .. R_X(m-1)]
%assumes R_X(n)==0 for n >=m
%output=vector a
% for lmse predictor xx=h'[X(n),X(n-1),..,X(n-N+1)] for X(n+k)
m=length(r);
r=[r(:);zeros(M-m+1,1)]; %appends zeros if needed
RY=toeplitz(r(1:M));
RYX=r(1+k:M+k);
h=flipud(RY\RYX);
```
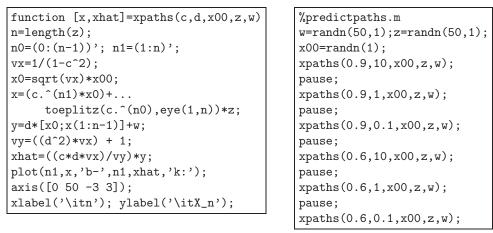
The code is pretty straightforward.
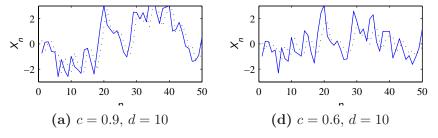
Problem 11.10.9 Solution

To generate the sample paths X_n and Y_n is relatively straightforward. For $\sigma = \eta = 1$, the solution to Problem 11.4.5 showed that the optimal linear predictor $\hat{X}_n(Y_{n-1})$ of X_n given Y_{n-1} is
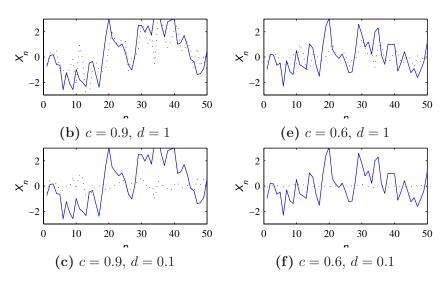
$$\hat{X}_n = \frac{cd}{d^2 + (1 - c^2)} Y_{n-1}. \tag{1}$$

In the following, we plot sample paths of X_n and $\hat{X}_n$. To show how sample paths are similar for different values of c and d, we construct all samples paths using the same sequence of noise samples W_n and Z_n. In addition, given c, we choose $X_0 = X_{00}/\sqrt{1 - c^2}$ where X_{00} is a Gaussian $(0, 1)$ random variable that is the same for all sample paths. To do this, we write a function `xpathplot(c,d,x00,z,w)` which constructs a sample path given the parameters c and d, and the noise vectors `z` and `w`. A simple script `predictpaths.m` calls `xpathplot` to generate the requested sample paths. Here are the two programs:

```
function [x,xhat]=xpaths(c,d,x00,z,w)
n=length(z);
n0=(0:(n-1))'; n1=(1:n)';
vx=1/(1-c^2);
x0=sqrt(vx)*x00;
x=(c.^(n1)*x0)+...
    toeplitz(c.^(n0),eye(1,n))*z;
y=d*[x0;x(1:n-1)]+w;
vy=((d^2)*vx) + 1;
xhat=((c*d*vx)/vy)*y;
plot(n1,x,'b-',n1,xhat,'k:');
axis([0 50 -3 3]);
xlabel('\itn'); ylabel('\itX_n');
```

```
%predictpaths.m
w=randn(50,1);z=randn(50,1);
x00=randn(1);
xpaths(0.9,10,x00,z,w);
pause;
xpaths(0.9,1,x00,z,w);
pause;
xpaths(0.9,0.1,x00,z,w);
pause;
xpaths(0.6,10,x00,z,w);
pause;
xpaths(0.6,1,x00,z,w);
pause;
xpaths(0.6,0.1,x00,z,w);
```

Some sample paths for X_n and $\hat{X}_n$ for the requested parameters are shown below. In each pair, the one-step prediction $\hat{X}_n$ is marked by dots.

(a) $c = 0.9$, $d = 10$

(d) $c = 0.6$, $d = 10$

414

(b) $c = 0.9$, $d = 1$

(e) $c = 0.6$, $d = 1$

(c) $c = 0.9$, $d = 0.1$

(f) $c = 0.6$, $d = 0.1$

The mean square estimation error at step n was found to be

$$e_L^*(n) = e_L^* = \sigma^2 \frac{d^2 + 1}{d^2 + (1 - c^2)} \tag{2}$$

We see that the mean square estimation error is $e_L^*(n) = e_L^*$, a constant for all n. In addition, e_L^* is a decreasing function of d. In graphs (a) through (c), we see that the predictor tracks X_n less well as d decreases because decreasing d corresponds to decreasing the contribution of X_{n-1} to the measurement Y_{n-1}. Effectively, the impact of the measurement noise is increased. As d decreases, the predictor places less emphasis on the measurement Y_n and instead makes predictions closer to $E[X] = 0$. That is, when d is small in graphs (c) and (f), the predictor stays close to zero. With respect to c, the performance of the predictor is less easy to understand. In Equation (11), the mean square error e_L^* is the product of

$$\mathrm{Var}[X_n] = \frac{\sigma^2}{1 - c^2} \qquad 1 - \rho_{X_n, Y_{n-1}}^2 = \frac{(d^2 + 1)(1 - c^2)}{d^2 + (1 - c^2)} \tag{3}$$

As a function of increasing c^2, $\mathrm{Var}[X_n]$ increases while $1 - \rho_{X_n, Y_{n-1}}^2$ decreases. Overall, the mean square error e_L^* is an increasing function of c^2. However, $\mathrm{Var}[X]$ is the mean square error obtained using a blind estimator that always predicts $E[X]$ while $1 - \rho_{X_n, Y_{n-1}}^2$ characterizes the extent to which the optimal linear predictor is better than the blind predictor. When we compare graphs (a)-(c) with $c = 0.9$ to graphs (d)-(f) with $c = 0.6$, we see greater variation in X_n for larger a but in both cases, the predictor worked well when d was large.