Optimum Signal Processing

Second Edition

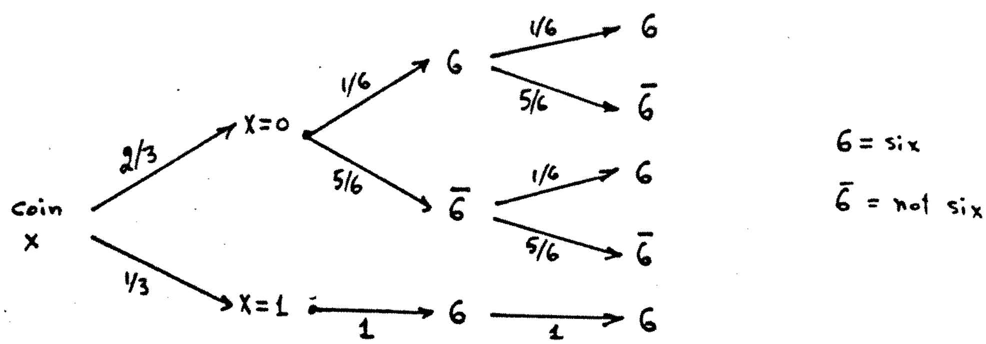
Solutions Manual

Sophocles J. Orfanidis

Department of Electrical Engineering
Rutgers University

Chapter 1

Problem 1.1:
The experiment is described by the following transition probability diagram:



a. Given that the fair die has been selected ($x = 0$), the corresponding conditional probabilities of getting $y = 0, 1, 2$ sixes are:

$$p(y=0 \mid x=0) = \left(\frac{5}{6}\right) = \frac{25}{36}, \quad p(y=1 \mid x=0) = 2 \cdot \frac{5}{6} \cdot \frac{1}{6} = \frac{10}{36}, \quad p(y=2 \mid x=0) = \left(\frac{1}{6}\right)^2 = \frac{1}{36}$$

Similarly, the conditional probabilities of getting y sixes given that the biased die was selected ($x = 1$) are:

$$p(y=0 \mid x=1) = 0, \quad p(y=1 \mid x=1) = 0, \quad p(y=2 \mid x=1) = 1$$

b. The probability $p(y)$ of getting y sixes regardless of which die was selected can be computed using Bayes' rule:

$$p(y) = \sum_{x=0,1} p(y,x) = \sum_{x=0,1} p(y \mid x) p(x) = p(y \mid x=0) p(x=0) + p(y \mid x=1) p(x=1)$$

where the a priori probabilities of x are $p(x=0) = 2/3$ and $p(x=1) = 1/3$. Thus,

$$p(y=0) = \frac{25}{36} \cdot \frac{2}{3} + 0 \cdot \frac{1}{3} = \frac{50}{108}$$

$$p(y=1) = \frac{10}{36} \cdot \frac{2}{3} + 0 \cdot \frac{1}{3} = \frac{20}{108}$$

$$p(y=2) = \frac{1}{36}\cdot\frac{2}{3} + 1\cdot\frac{1}{3} = \frac{38}{108}$$

c. The mean number of sixes regardless of die is computed by:

$$E[y] = \sum_{y=0,1,2} y\,p(y) = 0\cdot\frac{50}{108} + 1\cdot\frac{20}{108} + 2\cdot\frac{38}{108} = 0.89$$

d. Apply Bayes' rule $p(x \mid y) = p(y \mid x)p(x)/p(y)$. Given that we observed either 0 or 1 six, it is certain that we must have started with the fair die; given that we observed 2 sixes there is a small chance that we started with the fair die; indeed,

$$p(x=0 \mid y=0) = \frac{p(y=0 \mid x=0)p(x=0)}{p(y=0)} = \frac{\frac{25}{36}\cdot\frac{2}{3}}{\frac{50}{108}} = 1$$

$$p(x=1 \mid y=0) = \frac{p(y=0 \mid x=1)p(x=1)}{p(y=0)} = \frac{0\cdot\frac{1}{3}}{\frac{50}{108}} = 0$$

$$p(x=0 \mid y=1) = \frac{p(y=1 \mid x=0)p(x=0)}{p(y=1)} = \frac{\frac{10}{36}\cdot\frac{2}{3}}{\frac{20}{108}} = 1$$

$$p(x=1 \mid y=1) = \frac{p(y=1 \mid x=1)p(x=1)}{p(y=1)} = \frac{0\cdot\frac{1}{3}}{\frac{20}{108}} = 0$$

$$p(x=0 \mid y=2) = \frac{p(y=2 \mid x=1)p(x=1)}{p(y=2)} = \frac{\frac{1}{36}\cdot\frac{2}{3}}{\frac{38}{108}} = \frac{1}{19} = 5.3\,\%$$

$$p(x=1 \mid y=2) = \frac{p(y=2 \mid x=1)p(x=1)}{p(y=2)} = \frac{1\cdot\frac{1}{3}}{\frac{38}{108}} = \frac{18}{19} = 94.7\,\%$$

Problem 1.2:
By construction $u = F(x)$ is in the interval $0 \le u \le 1$. Its probability density is obtained by identifying elemental probabilities under the transformation from x to u; that is, $p_u(u)du = p(x)dx$ or, $p_u(u) = p(x)dx/du$. But $du = F'(x)dx = p(x)dx$. Therefore, $p_u(u) = 1$. In conclusion, a uniform u generates x according to $p(x)$.

Problem 1.3:

It follows from the definite integrals $\int_0^\infty xe^{-x}dx = 1$ and $\int_0^\infty x^2e^{-x}dx = 2$ that $E[x] = \mu$ and $E[x^2] = 2\mu^2$. Thus, $\sigma^2 = E[x^2] - E[x]^2 = \mu^2$. The cumulative distribution is $u = \int_0^x e^{-y/\mu}dy/\mu = 1 - e^{-x/\mu}$. Inverting, we obtain $x = -\mu\ln(1-u)$ which can be used to generate x from a uniform u.

Problem 1.4:

The cumulative distribution is $u = \int_0^r xe^{-x^2/2\sigma^2}dx/\sigma^2 = 1 - e^{-r^2/2\sigma^2}$. Inverting, we obtain $r = \sigma\sqrt{-2\ln(1-u)}$.

Problem 1.6:
Multiplying out the matrix factors, we obtain

$$\begin{bmatrix} I_N & -H \\ 0 & I_M \end{bmatrix}\begin{bmatrix} R_{xx} & R_{xy} \\ R_{yx} & R_{yy} \end{bmatrix}\begin{bmatrix} I_N & -H \\ 0 & I_M \end{bmatrix}^T = \begin{bmatrix} R_{xx} - HR_{yx} & R_{xy} - HR_{yy} \\ R_{yx} & R_{yy} \end{bmatrix}\begin{bmatrix} I_N & 0 \\ -H^T & I_M \end{bmatrix} =$$

$$= \begin{bmatrix} R_{xx} - R_{xy}R_{yy}^{-1}R_{yx} & R_{xy} - R_{xy}R_{yy}^{-1}R_{yy} \\ R_{yx} & R_{yy} \end{bmatrix}\begin{bmatrix} I_N & 0 \\ -R_{yy}^{-1}R_{yx} & I_M \end{bmatrix} =$$

$$\begin{bmatrix} R_{xx} - R_{xy}R_{yy}^{-1}R_{yx} & 0 \\ R_{yx} - R_{yy}R_{yy}^{-1}R_{yx} & R_{yy} \end{bmatrix} = \begin{bmatrix} R_{xx} - R_{xy}R_{yy}^{-1}R_{yx} & 0 \\ 0 & R_{yy} \end{bmatrix}$$

The correlation canceling results follow from this identity by recognizing that the above matrix product is the covariance matrix of the transformed vector

$$\begin{bmatrix} \mathbf{e} \\ \mathbf{y} \end{bmatrix} = \begin{bmatrix} \mathbf{x} - H\mathbf{y} \\ \mathbf{y} \end{bmatrix} = \begin{bmatrix} I_N & -H \\ 0 & I_M \end{bmatrix}\begin{bmatrix} \mathbf{x} \\ \mathbf{y} \end{bmatrix}$$

thus,

$$E\left[\begin{bmatrix} \mathbf{e} \\ \mathbf{y} \end{bmatrix}[\mathbf{e}^T, \mathbf{y}^T]\right] = \begin{bmatrix} I_N & -H \\ 0 & I_M \end{bmatrix} E\left[\begin{bmatrix} \mathbf{x} \\ \mathbf{y} \end{bmatrix}[\mathbf{x}^T, \mathbf{y}^T]\right]\begin{bmatrix} I_N & -H \\ 0 & I_M \end{bmatrix}^T$$

or,

$$\begin{bmatrix} R_{ee} & R_{ey} \\ R_{ye} & R_{yy} \end{bmatrix} = \begin{bmatrix} I_N & -H \\ 0 & I_M \end{bmatrix}\begin{bmatrix} R_{xx} & R_{xy} \\ R_{yx} & R_{yy} \end{bmatrix}\begin{bmatrix} I_N & -H \\ 0 & I_M \end{bmatrix}^T = \begin{bmatrix} R_{xx} - R_{xy}R_{yy}^{-1}R_{yx} & 0 \\ 0 & R_{yy} \end{bmatrix}$$

which implies that

$$R_{ey} = 0\,, \quad R_{ee} = R_{xx} - R_{xy}R_{yy}^{-1}R_{yx}$$

Problem 1.7
The matrix identity of Problem 1.6 can also be written as

$$\begin{bmatrix} R_{xx} & R_{xy} \\ R_{yx} & R_{yy} \end{bmatrix} = \begin{bmatrix} I_N & H \\ 0 & I_M \end{bmatrix} \begin{bmatrix} R_{ee} & 0 \\ 0 & R_{yy} \end{bmatrix} \begin{bmatrix} I_N & H \\ 0 & I_M \end{bmatrix}^T$$

Taking determinants of both sides and realizing that $\begin{bmatrix} I_N & H \\ 0 & I_M \end{bmatrix}$ has unit determinant, we find

$$\det\begin{bmatrix} R_{xx} & R_{xy} \\ R_{yx} & R_{yy} \end{bmatrix} = \det\begin{bmatrix} R_{ee} & 0 \\ 0 & R_{yy} \end{bmatrix} = (\det R_{ee})(\det R_{yy})$$

Next, we derive an expression for the conditional density $p\,(\mathbf{x} \mid \mathbf{y})$ using Bayes' rule:

$$p\,(\mathbf{x} \mid \mathbf{y}) = \frac{p\,(\mathbf{x}, \mathbf{y})}{p\,(\mathbf{y})}$$

The joint density may be factored as follows:

$$p\,(\mathbf{x}, \mathbf{y}) = \frac{1}{(2\pi)^{(N+M)/2}\left(\det\begin{bmatrix} R_{xx} & R_{xy} \\ R_{yx} & R_{yy} \end{bmatrix}\right)^{1/2}} \exp\left(-\frac{1}{2}\,[\mathbf{x}^T, \mathbf{y}^T] \begin{bmatrix} R_{xx} & R_{xy} \\ R_{yx} & R_{yy} \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{x} \\ \mathbf{y} \end{bmatrix} \right) =$$

$$= \frac{1}{(2\pi)^{(N+M)/2}(\det R_{ee})^{1/2}(\det R_{yy})^{1/2}} \exp\left(-\frac{1}{2}\,[\mathbf{x}^T, \mathbf{y}^T] \begin{bmatrix} I_N & -H \\ 0 & I_M \end{bmatrix}^T \begin{bmatrix} R_{ee}^{-1} & 0 \\ 0 & R_{yy}^{-1} \end{bmatrix} \begin{bmatrix} I_N & -H \\ 0 & I_M \end{bmatrix} \begin{bmatrix} \mathbf{x} \\ \mathbf{y} \end{bmatrix} \right) =$$

$$= \frac{1}{(2\pi)^{(N+M)/2}(\det R_{ee})^{1/2}(\det R_{yy})^{1/2}} \exp\left(-\frac{1}{2} \begin{bmatrix} \mathbf{x} - H\mathbf{y} \\ \mathbf{y} \end{bmatrix}^T \begin{bmatrix} R_{ee}^{-1} & 0 \\ 0 & R_{yy}^{-1} \end{bmatrix} \begin{bmatrix} \mathbf{x} - H\mathbf{y} \\ \mathbf{y} \end{bmatrix} \right) =$$

$$= \frac{\exp\left[-\frac{1}{2}\mathbf{y}^T R_{yy}^{-1}\mathbf{y} \right]}{(2\pi)^{M/2}(\det R_{yy})^{1/2}} \cdot \frac{\exp\left[-\frac{1}{2}(\mathbf{x} - H\mathbf{y})^T R_{ee}^{-1}(\mathbf{x} - H\mathbf{y}) \right]}{(2\pi)^{N/2}(\det R_{ee})^{1/2}} = p\,(\mathbf{y}) \cdot \frac{\exp\left[-\frac{1}{2}(\mathbf{x} - H\mathbf{y})^T R_{ee}^{-1}(\mathbf{x} - H\mathbf{y}) \right]}{(2\pi)^{N/2}(\det R_{ee})^{1/2}}$$

Therefore,

$$p\,(\mathbf{x} \mid \mathbf{y}) = \frac{p\,(\mathbf{x}, \mathbf{y})}{p\,(\mathbf{y})} = \frac{\exp\left[-\frac{1}{2}(\mathbf{x} - H\mathbf{y})^T R_{ee}^{-1}(\mathbf{x} - H\mathbf{y}) \right]}{(2\pi)^{N/2}(\det R_{ee})^{1/2}}$$

And we recognize this as a gaussian density in $\mathbf{x}$ with mean $H\mathbf{y}$ and covariance R_{ee}. Since the conditional mean is the mean with respect to this conditional density, it follows that

$$E[\mathbf{x} \mid \mathbf{y}] = H\mathbf{y}$$

Problem 1.8:
Using $\hat{\mathbf{x}} = R_{xy}R_{yy}^{-1}\mathbf{y}$ we find

$$\widehat{A_1\mathbf{x}_1 + A_2\mathbf{x}_2} = R_{(A_1x_1 + A_2x_2)y}R_{yy}^{-1}\mathbf{y} = (A_1R_{x_1y} + A_2R_{x_2y})R_{yy}^{-1}\mathbf{y} =$$

$$= A_1R_{x_1y}R_{yy}^{-1}\mathbf{y} + A_2R_{x_2y}R_{yy}^{-1}\mathbf{y} = A_1\hat{\mathbf{x}}_1 + A_2\hat{\mathbf{x}}_2$$

Problem 1.9:
Using $x = s + n_1$, $y = n_2$, and $n_1 = Fn_2$, we find $R_{yy} = R_{n_2n_2}$ and

$$R_{xy} = R_{(s + n_1)y} = R_{(s + Fn_2)n_2} = R_{sn_2} + FR_{n_2n_2} = FR_{n_2n_2}$$

Thus, $H = R_{xy}R_{yy}^{-1} = F$ and the estimate of x becomes $\hat{x} = Hy = Fy$. The estimation error is

$$e = x - \hat{x} = (s + Fn_2) - Fn_2 = s$$

Thus, the noise component n_1 is cancelled completely. If $n_1 = Fn_2 + v$, then we still have $R_{xy} = FR_{n_2n_2}$ and $H = F$. The estimation error is now

$$e = x - \hat{x} = s + Fn_2 + v - Fn_2 = s + v$$

Problem 1.10:
First, we determine H. Noting that $y = n_2 + \varepsilon s = \frac{1}{F}n_1 + \varepsilon s$ and using the given definition for the gain factor G, we find

$$R_{yy} = \frac{1}{F^2}E[n_1^2] + \varepsilon^2E[s^2] = \left(\frac{1}{F^2} + G^2\right)E[n_1^2]$$

Similarly,

$$R_{xy} = \frac{1}{F}E[n_1^2] + \varepsilon E[s^2] = \left(\frac{1}{F} + \varepsilon G\right)E[n_1^2]$$

Therefore,

$$H = R_{xy}R_{yy}^{-1} = \frac{\frac{1}{F} + \varepsilon G}{\frac{1}{F^2} + \varepsilon^2 G} = \frac{F(1 + \varepsilon FG)}{1 + \varepsilon^2 F^2 G}$$

The output of the canceler is:

$$e = x - \hat{x} = x - Hy = (s + n_1) - H(\frac{1}{F}n_1 + \varepsilon s) = (1 - \varepsilon H)s + (1 - \frac{H}{F})n_1$$

Thus, we identify the coefficients a and b as follows:

$$a = 1 - \varepsilon H = 1 - \frac{F\varepsilon(1 + \varepsilon FG)}{1 + \varepsilon^2 F^2 G}$$

$$b = 1 - \frac{H}{F} = 1 - \frac{1 + \varepsilon FG}{1 + \varepsilon^2 F^2 G} = -\varepsilon FG \frac{1 - \varepsilon F}{1 + \varepsilon^2 F^2 G} = -\varepsilon FGa$$

Problem 1.11:
Since $c_n = 1$, it follows that $\mathbf{c}^T\mathbf{c} = M$. Thus, from Example (1.4.3), we find for the optimal estimate:

$$\hat{x} = \frac{1}{1 + \mathbf{c}^T\mathbf{c}}\mathbf{c}^T\mathbf{y} = \frac{1}{M+1}\sum_{n=1}^{M} y_n$$

and the corresponding mean square estimation error:

$$E[e^2]_{\min} = \frac{1}{1 + \mathbf{c}^T\mathbf{c}} = \frac{1}{M+1}$$

The ordinary average $\hat{x}_{av} = \frac{1}{M}\sum_{n=1}^{M} y_n$ will have an estimation error:

$$e = x - \hat{x}_{av} = x - \frac{1}{M}\sum_{n=1}^{M}(x + v_n) = -\frac{1}{M}\sum_{n=1}^{M} v_n$$

and

$$E[e^2] = \frac{1}{M^2}E[(\sum_{n=1}^{M} v_n)^2] = \frac{1}{M^2}\sum_{n,m=1}^{M} E[v_n v_m] = \frac{1}{M^2}\sum_{n,m=1}^{M}\delta_{nm} = \frac{1}{M}$$

which greater than the optimal error $\frac{1}{M+1}$.

Problem 1.12:

a.

$$\hat{x}_n = \sum_{i=1}^{n} E[x\epsilon_i]E[\epsilon_i\epsilon_i]^{-1}\epsilon_i = \sum_{i=1}^{n-1} E[x\epsilon_i]E[\epsilon_i\epsilon_i]^{-1}\epsilon_i + E[x\epsilon_n]E[\epsilon_n\epsilon_n]^{-1}\epsilon_n \quad \text{or,}$$

$$\hat{x}_n = \hat{x}_{n-1} + G_n\epsilon_n$$

but note that $\epsilon_n = y_n - \hat{y}_{n/n-1}$.

b. Since $\hat{x}_{n-1}$ is a linear combination of $\{\epsilon_1, \epsilon_2, \cdots, \epsilon_{n-1}\}$ and each of these is orthogonal to ϵ_n by construction, it follows that $\hat{x}_{n-1}$ will be orthogonal to ϵ_n.

c. Let $\boldsymbol{\epsilon}_n = [\epsilon_1, \epsilon_2, \cdots, \epsilon_n]^T$. Its covariance matrix is diagonal; i.e.,

$$R_{\epsilon_n\epsilon_n} = E[\boldsymbol{\epsilon}_n\boldsymbol{\epsilon}_n^T] = \text{diag}\{E[\epsilon_1^2], E[\epsilon_2^2], \cdots, E[\epsilon_n^2]\}$$

The estimate $\hat{x}_n$ can be expressed as follows:

$$\hat{x}_n = E[x\boldsymbol{\epsilon}_n^T]E[\boldsymbol{\epsilon}_n\boldsymbol{\epsilon}_n^T]^{-1}\boldsymbol{\epsilon}_n = R_{x\epsilon_n}R^{-1}_{\epsilon_n\epsilon_n}\boldsymbol{\epsilon}_n$$

Using Eq. (1.4.4), we find for the corresponding estimation error

$$R_{e_ne_n} = R_{xx} - E[x\boldsymbol{\epsilon}_n^T]E[\boldsymbol{\epsilon}_n\boldsymbol{\epsilon}_n^T]^{-1}E[\boldsymbol{\epsilon}_n x^T] = R_{xx} - \sum_{i=1}^{n} E[x\epsilon_i]E[\epsilon_i\epsilon_i]^{-1}E[\epsilon_i x]$$

where we used the fact that $E[\boldsymbol{\epsilon}_n\boldsymbol{\epsilon}_n^T]$ is diagonal.

d. Separating out the first n -1 terms in the above sum, we find

$$R_{e_ne_n} = R_{e_{n-1}e_{n-1}} - E[x\epsilon_n]E[\epsilon_n\epsilon_n]^{-1}E[\epsilon_n x^T] \leq R_{e_{n-1}e_{n-1}}$$

Problem 1.13:

a. Same as Eq. (1.4.9) but with $\mathbf{c}$ replaced by $\mathbf{c}_n$.

b. Defining $\mathbf{v}_n = [v_1, v_2, \cdots, v_n]^T$, we may write compactly

$$\mathbf{y}_n = \mathbf{c}_n x + \mathbf{v}_n, \qquad \mathbf{y}_{n-1} = \mathbf{c}_{n-1} x + \mathbf{v}_{n-1}$$

It follows that

$$E[y_n y_{n-1}^T] = E[(c_n x + v_n)(\mathbf{c}_{n-1}^T x + \mathbf{v}_{n-1}^T)] = c_n \mathbf{c}_{n-1}^T$$

where we used the fact that $E[x^2] = 1$. Similarly,

$$E[\mathbf{y}_{n-1}\mathbf{y}_{n-1}^T] = E[(\mathbf{c}_{n-1}x + \mathbf{v}_{n-1})(\mathbf{c}_{n-1}^T x + \mathbf{v}_{n-1}^T)] = \mathbf{c}_{n-1}\mathbf{c}_{n-1}^T + I_{n-1}$$

where I_{n-1} is the $(n-1)\times(n-1)$ unit matrix, and we used the fact that $E[\mathbf{v}_{n-1}\mathbf{v}_{n-1}^T] = I_{n-1}$. Then,

$$E[y_n y_{n-1}^T]^{-1} = c_n c_{n-1}^T (I_{n-1} + c_{n-1} c_{n-1}^T)^{-1} = c_n (1 + c_{n-1}^T c_{n-1})^{-1} c_{n-1}^T$$

where we used the matrix identity $c^T(I + cc^T)^{-1} = (1 + c^T c)^{-1} c^T$. Thus,

$$\hat{y}_{n/n-1} = E[y_n y_{n-1}^T] E[y_{n-1} y_{n-1}^T]^{-1} y_{n-1} = c_n (1 + c_{n-1}^T c_{n-1})^{-1} c_{n-1}^T y_{n-1}$$

but note that $\hat{x}_{n-1} = (1 + c_{n-1}^T c_{n-1})^{-1} c_{n-1}^T y_{n-1}$. Therefore,

$$\hat{y}_{n/n-1} = c_n \hat{x}_{n-1}$$

c. Starting with $\varepsilon_n = c_n e_{n-1} + v_n$ and using the fact that e_{n-1} is uncorrelated with v_n, we find

$$E[\varepsilon_n^2] = c_n^2 E[e_{n-1}^2] + E[v_n^2] = \frac{c_n^2}{1 + c_{n-1}^T c_{n-1}} + 1 = \frac{1 + c_n^T c_n}{1 + c_{n-1}^T c_{n-1}}$$

Similarly,

$$E[x\varepsilon_n] = c_n E[x e_{n-1}] = c_n E[e_{n-1}^2] = \frac{c_n}{1 + c_{n-1}^T c_{n-1}}$$

d. Finally, using the results of Problem 1.12 we find

$$\hat{x}_n = \hat{x}_{n-1} + G_n (y_n - \hat{y}_{n/n-1})$$

where the Kalman gain is computed by

$$G_n = E[x\varepsilon_n] E[\varepsilon_n \varepsilon_n]^{-1} = \frac{c_n}{1 + c_n^T c_n}$$

In this problem the dynamics of x was trivial; namely, x was a constant in time. The more general case of non trivial dynamics is discussed in Section 4.9.

Problem 1.14:

Using the above expression for G_n and $\hat{x}_{n-1} = \dfrac{c_{n-1}^T y_{n-1}}{1 + c_{n-1}^T c_{n-1}}$, we find

$$\hat{x}_{n-1} + G_n (y_n - \hat{y}_{n/n-1}) = \hat{x}_{n-1} + G_n (y_n - c_n \hat{x}_{n-1}) = (1 - c_n G_n) \hat{x}_{n-1} + G_n y_n =$$

$$= \left(1 - \frac{c_n^2}{1 + c_n^T c_n}\right) \hat{x}_{n-1} + \frac{c_n}{1 + c_n^T c_n} y_n = \frac{1 + c_n^T c_n - c_n^2}{1 + c_n^T c_n} \hat{x}_{n-1} + \frac{c_n}{1 + c_n^T c_n} y_n =$$

$$= \frac{(1 + c_{n-1}^T c_{n-1}) \hat{x}_{n-1} + c_n y_n}{1 + c_n^T c_n} = \frac{c_{n-1}^T y_{n-1} + c_n y_n}{1 + c_n^T c_n} = \frac{c_n^T y_n}{1 + c_n^T c_n}$$

Problem 1.15:
The Cholesky factors of R are

$$R = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 6 & 14 \\ 3 & 14 & 42 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 3 & 4 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 & 3 \\ 0 & 1 & 4 \\ 0 & 0 & 1 \end{bmatrix} = BR_{ee}B^T$$

Problem 1.16:
The first three iterations of the backward Gram-Schmidt procedure lead to:

$$\begin{aligned} \eta_M &= y_M \\ \eta_{M-1} &= y_{M-1} - E\,[y_{M-1}\eta_M]E\,[\eta_M\eta_M]^{-1}\eta_M \\ \eta_{M-2} &= y_{M-2} - E\,[y_{M-2}\eta_{M-1}]E\,[\eta_{M-1}\eta_{M-1}]^{-1}\eta_{M-1} - E\,[y_{M-2}\eta_M]E\,[\eta_M\eta_M]^{-1}\eta_M \end{aligned}$$

Expressing the ys in terms of the ηs, we find

$$\begin{aligned} y_{M-2} &= \eta_{M-2} + a\eta_{M-1} + b\eta_M \\ y_{M-1} &= \eta_{M-1} + c\eta_M \\ y_M &= \eta_M \end{aligned}$$

with obvious definitions for the coefficients a, b, c. In matrix form:

$$\begin{bmatrix} y_{M-2} \\ y_{M-1} \\ y_M \end{bmatrix} = \begin{bmatrix} 1 & a & b \\ 0 & 1 & c \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \eta_{M-2} \\ \eta_{M-1} \\ \eta_M \end{bmatrix} \quad \text{or,} \quad \mathbf{y} = U\boldsymbol{\eta}$$

The UL factorization follows from

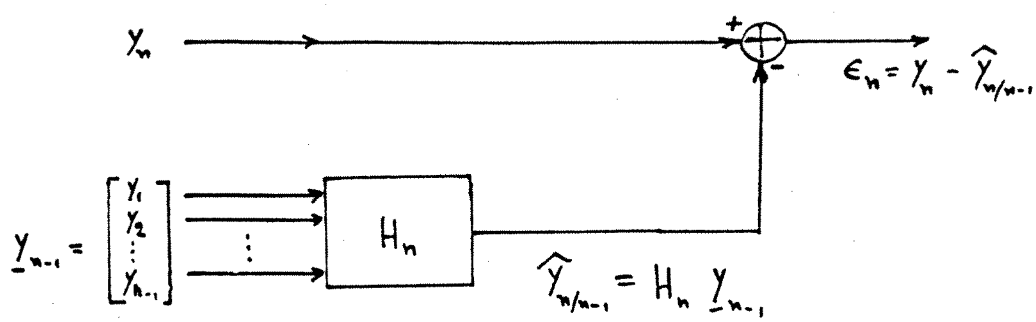
$$R_{yy} = E\,[\mathbf{y}\mathbf{y}^T] = E\,[U\eta\eta^T U^T] = UE\,[\eta\eta^T]U^T = UR_{\eta\eta}U^T$$

where $R_{\eta\eta} = E\,[\eta\eta^T]$ is diagonal by construction.

Problem 1.17:
The correspondences are:

$$\begin{aligned} H = E\,[\mathbf{x}\mathbf{y}^T]E\,[\mathbf{y}\mathbf{y}^T]^{-1} \quad &\rightarrow \quad H_n = E\,[y_n\mathbf{y}_{n-1}^T]E\,[\mathbf{y}_{n-1}\mathbf{y}_{n-1}^T]^{-1} \\ \mathbf{x} \quad &\rightarrow \quad y_n \\ \mathbf{y} \quad &\rightarrow \quad \mathbf{y}_{n-1} \\ \mathbf{e} \quad &\rightarrow \quad \varepsilon_n = y_n - \hat{y}_{n/n-1} \end{aligned}$$

and in diagram form



Problem 1.18:
Equation (1.6.11) may be written as

$$\hat{y}_{n/n-1} = -[a_{n1}, a_{n2}, \cdots, a_{nn}]\begin{bmatrix} y_{n-1} \\ y_{n-2} \\ \vdots \\ y_0 \end{bmatrix} = -[a_{n1}, a_{n2}, \cdots, a_{nn}]\mathbf{y}_{n-1}$$

Comparing with Eq. (1.5.19), we identify the vector of as with

$$-[a_{n1}, a_{n2}, \cdots, a_{nn}] = E[y_n\mathbf{y}_{n-1}^T]E[\mathbf{y}_{n-1}\mathbf{y}_{n-1}^T]^{-1}$$

Problem 1.19:
Write $\varepsilon_n = y_n - \hat{y}_{n/n-1} = y_n - H_n\mathbf{y}_{n-1}$, where $H_n = E[y_n\mathbf{y}_{n-1}^T]E[\mathbf{y}_{n-1}\mathbf{y}_{n-1}^T]^{-1}$. The correlation cancellation conditions are equivalent to the Gram-Schmidt orthogonality conditions; that is, by construction $E[\varepsilon_n\mathbf{y}_{n-1}] = 0$. It follows that $E[\varepsilon_n\hat{y}_{n/n-1}] = 0$ and

$$E[\varepsilon_n^2] = E[\varepsilon_n(y_n - \hat{y}_{n/n-1})] = E[\varepsilon_n y_n] = E[(y_n - H_n\mathbf{y}_{n-1})y_n] \quad \text{or,}$$

$$E[\varepsilon_n^2] = E[y_n^2] - H_n E[\mathbf{y}_{n-1}y_n] = E[y_n^2] - E[y_n\mathbf{y}_{n-1}^T]E[\mathbf{y}_{n-1}\mathbf{y}_{n-1}^T]^{-1}E[\mathbf{y}_{n-1}y_n]$$

Problem 1.20:
First, note that $E[\mathbf{y}_n\mathbf{y}_n^T]$ can be expressed as

$$E[\mathbf{y}_n\mathbf{y}_n^T] = E\left[\begin{bmatrix} y_n \\ \mathbf{y}_{n-1} \end{bmatrix}[y_n, \mathbf{y}_{n-1}^T]\right] = \begin{bmatrix} E[y_n^2] & E[y_n\mathbf{y}_{n-1}^T] \\ E[y_n\mathbf{y}_{n-1}] & E[\mathbf{y}_{n-1}\mathbf{y}_{n-1}^T] \end{bmatrix}$$

Let $\alpha_n = [a_{n1}, a_{n2}, \cdots, a_{nn}]^T$. Then, the results of Problems 1.18 and 1.19 can be written in the compact form:

$$
\begin{aligned}
E[\varepsilon_n^2] &= E[y_n^2] - E[y_n \mathbf{y}_{n-1}^T] E[\mathbf{y}_{n-1}\mathbf{y}_{n-1}^T]^{-1} E[\mathbf{y}_{n-1} y_n] \\
\alpha_n &= -E[\mathbf{y}_{n-1}\mathbf{y}_{n-1}^T]^{-1} E[\mathbf{y}_{n-1} y_n]
\end{aligned}
$$

or, rearranging terms

$$
\begin{aligned}
E[y_n \mathbf{y}_{n-1}] + E[\mathbf{y}_{n-1}\mathbf{y}_{n-1}^T]\alpha_n &= 0 \\
E[y_n^2] + E[y_n \mathbf{y}_{n-1}^T]\alpha_n &= E[\varepsilon_n^2]
\end{aligned}
$$

We can combine these into the matrix equation:

$$
\begin{bmatrix} E[y_n^2] & E[y_n \mathbf{y}_{n-1}^T] \\ E[y_n \mathbf{y}_{n-1}] & E[\mathbf{y}_{n-1}\mathbf{y}_{n-1}^T] \end{bmatrix} \begin{bmatrix} 1 \\ \alpha_n \end{bmatrix} = \begin{bmatrix} E[\varepsilon_n^2] \\ 0 \end{bmatrix} = E[\varepsilon_n^2] \begin{bmatrix} 1 \\ 0 \end{bmatrix}
$$

Problem 1.21:
Let $J = \mathbf{a}_n^T E[\mathbf{y}_n \mathbf{y}_n^T]\mathbf{a}_n + \lambda(1 - \mathbf{a}_n^T \mathbf{u}_n)$ be the extended performance index. Setting its gradient to zero,

$$
\frac{\partial J}{\partial \mathbf{a}_n} = 2E[\mathbf{y}_n \mathbf{y}_n^T]\mathbf{a}_n - \lambda \mathbf{u}_n = 0
$$

we obtain

$$
E[\mathbf{y}_n \mathbf{y}_n^T] = \frac{\lambda}{2}\mathbf{u}_n
$$

The Lagrange multiplier λ is fixed by imposing the constraint $\mathbf{a}_n^T \mathbf{u}_n = 1$. Multiplying both sides of the above equation by $\mathbf{a}_n^T$, we get

$$
\mathbf{a}_n^T E[\mathbf{y}_n \mathbf{y}_n^T]\mathbf{a}_n = \frac{\lambda}{2}\mathbf{a}_n^T \mathbf{u}_n = \frac{\lambda}{2}
$$

But

$$
\mathbf{a}_n^T E[\mathbf{y}_n \mathbf{y}_n^T]\mathbf{a}_n = E[(\mathbf{a}_n^T \mathbf{y}_n)(\mathbf{y}_n^T \mathbf{a}_n)] = E[\varepsilon_n^2]
$$

Thus, the determining equation for $\mathbf{a}_n$ becomes $E[\mathbf{y}_n \mathbf{y}_n^T]\mathbf{a}_n = E[\varepsilon_n^2]\mathbf{u}_n$.

Problem 1.22:
The effective performance index, with the constraint built in, is

$$
J = \mathbf{a}^T R \mathbf{a} + \lambda(1 - \mathbf{u}^T \mathbf{a}) = \min
$$

Setting the gradient with respect to a to zero gives $\dfrac{\partial J}{\partial \mathbf{a}} = 2R\mathbf{a} - \lambda = 0$, or, $\mathbf{a} = \lambda R^{-1}\mathbf{u}/2$. The Lagrange multiplier is fixed by requiring the constraint $\mathbf{a}^T\mathbf{u} = 1$.

Problem 1.23:
Taking determinants of both sides of Eq. (1.7.16) and using the fact that L has unit determinant (being unit lower triangular), it follows that $\det R = (\det \bar{R})E_b$.

Problem 1.24:
Using (1.7.13) into (1.7.19), we find

$$R^{-1} = L^T D_b^{-1} L = \begin{bmatrix} \bar{L} & 0 \\ \boldsymbol{\beta}^T & 1 \end{bmatrix}^T \begin{bmatrix} \bar{D}_b^{-1} & 0 \\ 0^T & E_b^{-1} \end{bmatrix} \begin{bmatrix} L & 0 \\ \boldsymbol{\beta}^T & 1 \end{bmatrix} = \begin{bmatrix} \bar{L}^T \bar{D}_b^{-1} \bar{L} + E_b^{-1} \boldsymbol{\beta}\boldsymbol{\beta}^T & E_b^{-1}\boldsymbol{\beta} \\ E_b^{-1}\boldsymbol{\beta}^T & E_b^{-1} \end{bmatrix}$$

or,

$$R^{-1} = \begin{bmatrix} \bar{R}^{-1} & 0 \\ 0^T & 0 \end{bmatrix} + \frac{1}{E_b} \begin{bmatrix} \boldsymbol{\beta} \\ 1 \end{bmatrix} [\boldsymbol{\beta}^T, 1]$$

Problem 1.25:

a. Since a and b are independent and have zero mean, we have $E[ab] = 0$. Therefore,

$$E[x_n^2] = E[(a + bn)^2] = E[a^2] + n^2 E[b^2] = \sigma_a^2 + n^2 \sigma_b^2$$

b. x_n is not stationary, because $E[x_n^2]$ depends on the absolute time n; similarly, it cannot be ergodic because ergodicity requires stationarity.

c. Being the linear combination of two gaussians, x_n is itself gaussian. It has zero mean and variance $\sigma_n^2 = E[x_n^2] = \sigma_a^2 + n^2\sigma_b^2$. Thus, its density is

$$p(x_n) = \frac{1}{(2\pi)^{1/2}\sigma_n} \exp\left(-\frac{x_n^2}{2\sigma_n^2} \right)$$

d. Since $x_n = x_m + b(n - m)$, it follows that if x_m is given, then the randomness in x_n will arise only from b; that is,

$$p(x_n \mid x_m) dx_n = p_b(b) db$$

Using $dx_n = (n - m)db$ and replacing $b = (x_n - x_m)/(n - m)$, we find

$$p(x_n \mid x_m) = \frac{p_b(b)}{\mid dx_n/db \mid} = \frac{e^{-b^2/2\sigma_b^2}}{(2\pi)^{1/2}\sigma_b \mid n - m \mid} = \frac{e^{-(x_n - x_m)^2/2\sigma_b^2(n - m)^2}}{(2\pi)^{1/2}\sigma_b \mid n - m \mid}$$

Problem 1.26:

a. Working in the time domain, we find for $0 \le k \le 10$

$$\hat{R}(k) = \frac{1}{11}\sum_{n=0}^{10-k} y_{n+k}y_n = \frac{1}{11}\sum_{n=0}^{10-k} 1\cdot 1 = \frac{11-k}{11}$$

For negative ks, we have $\hat{R}(k) = \hat{R}(-k)$. In the z-domain, we have

$$\hat{S}(z) = \frac{1}{11}Y(z)Y(z^{-1}) = \frac{1}{11}(1 + z^{-1} + z^{-2} + \cdots + z^{-10})(1 + z + z^2 + \cdots + z^{10}) =$$
$$= \frac{1}{11}\left(11 + 10(z + z^{-1}) + 9(z^2 + z^{-2}) + \cdots + 2(z^9 + z^{-9}) + (z^{10} + z^{-10})\right)$$

The coefficients of z^{-k} are the $\hat{R}(k)$s.

b. For $0 \le k \le 10$, we find

$$\hat{R}(k) = \frac{1}{11}\sum_{n=0}^{10-k} (-1)^{n+k}(-1)^n = \frac{1}{11}(-1)^k \sum_{n=0}^{10-k} 1 = (-1)^k\frac{11-k}{11}$$

Problem 1.27:

a. Using the time-domain definition, Eq. (1.10.1), we find

$$[R(0), R(\pm 1), R(\pm 2), R(\pm 3)] = \frac{1}{4}[10, 8, 4, 1]$$

b. The 4-point FFT of the sequence $\mathbf{y} = [1, 2, 2, 1]$ is

$$[6, -1 + j, 0, -1 - j]$$

Taking the magnitude squared of each entry, we find: [36, 2, 0, 2]. Taking the inverse 4-point FFT and dividing by 4, we find $\frac{1}{4}[10, 9, 8, 9]$, which does not agree with the correct answer. This is of course, the modulo-4 wrapped version of the correct answer; i.e.

$$[R(0), R(1) + R(-3), R(2) + R(-2), R(3) + R(-1)] = \frac{1}{4}[10, 9, 8, 9]$$

c. The 8-point FFT of $\mathbf{y} = [1, 2, 2, 1, 0, 0, 0, 0]$ is

$$\left[6,\ 1+\frac{1}{\sqrt{2}}-j\,(2+\frac{3}{\sqrt{2}}),\ -1-j,\ 1-\frac{1}{\sqrt{2}}+j\,(2-\frac{3}{\sqrt{2}}),\ 0,\ 1-\frac{1}{\sqrt{2}}-j\,(2-\frac{3}{\sqrt{2}}),\ -1+j,\ 1+\frac{1}{\sqrt{2}}+j\,(2+\frac{3}{\sqrt{2}})\right]$$

The magnitude squared of each entry is

$$[36, 10 + 7\sqrt{2}, 2, 10 - 7\sqrt{2}, 0, 10 - 7\sqrt{2}, 2, 10 + 7\sqrt{2}]$$

Taking the inverse 8-point FFT and dividing by 4 gives

$$\hat{R}(k) = \frac{1}{11}\sum_{n=0}^{10-k} y_{n+k}y_n = \frac{1}{11}\sum_{n=0}^{10-k} 1\cdot 1 = \frac{11-k}{11}$$

For negative ks, we have $\hat{R}(k) = \hat{R}(-k)$. In the z-domain, we have

$$\hat{S}(z) = \frac{1}{11}Y(z)Y(z^{-1}) = \frac{1}{11}(1 + z^{-1} + z^{-2} + \cdots + z^{-10})(1 + z + z^2 + \cdots + z^{10}) =$$
$$= \frac{1}{11}\left[11 + 10(z + z^{-1}) + 9(z^2 + z^{-2}) + \cdots + 2(z^9 + z^{-9}) + (z^{10} + z^{-10})\right]$$

The coefficients of z^{-k} are the $\hat{R}(k)$s.

b. For $0 \le k \le 10$, we find

$$\hat{R}(k) = \frac{1}{11}\sum_{n=0}^{10-k} (-1)^{n+k}(-1)^n = \frac{1}{11}(-1)^k\sum_{n=0}^{10-k} 1 = (-1)^k\frac{11-k}{11}$$

Problem 1.27:

a. Using the time-domain definition, Eq. (1.10.1), we find

$$[R(0), R(\pm 1), R(\pm 2), R(\pm 3)] = \frac{1}{4}[10, 8, 4, 1]$$

b. The 4-point FFT of the sequence $\mathbf{y} = [1, 2, 2, 1]$ is

$$[6, -1 + j, 0, -1 - j]$$

Taking the magnitude squared of each entry, we find: [36, 2, 0, 2]. Taking the inverse 4-point FFT and dividing by 4, we find $\frac{1}{4}[10, 9, 8, 9]$, which does not agree with the correct answer. This is of course, the modulo-4 wrapped version of the correct answer; i.e.

$$[R(0), R(1) + R(-3), R(2) + R(-2), R(3) + R(-1)] = \frac{1}{4}[10, 9, 8, 9]$$

c. The 8-point FFT of $\mathbf{y} = [1, 2, 2, 1, 0, 0, 0, 0]$ is

$$\left[1 + \frac{1}{\sqrt{2}} - j(2 + \frac{3}{\sqrt{2}}), -1-j, 1-\frac{1}{\sqrt{2}} + j(2-\frac{3}{\sqrt{2}}), 0, 1-\frac{1}{\sqrt{2}} - j(2-\frac{3}{\sqrt{2}}), -1 + j, 1 + \frac{1}{\sqrt{2}} + j(2 + \frac{3}{\sqrt{2}})\right]$$

The magnitude squared of each entry is

$$[36, 10 + 7\sqrt{2}, 2, 10 - 7\sqrt{2}, 0, 10 - 7\sqrt{2}, 2, 10 + 7\sqrt{2}]$$

Taking the inverse 8-point FFT and dividing by 4 gives

$$\frac{1}{4}[10, 8, 4, 1, 0, 1, 4, 8]$$

Note, the central entry 0 corresponds to time index $k = 4$ and is the correct answer if we think of the data sequence y as being padded with zeros beyond $k = 4$. The last three entries, corresponding to inverse FFT indices $k = 5,6,7$, are equivalent modulo-8 to the autocorrelation indices $k = -3,-2,-1$. In other words, the answer comes out of the inverse FFT in the order

$$[R(0), R(1), R(2), R(3), 0, R(-3), R(-2), R(-1)]$$

Problem 1.29:
Let $T_R = 500$ msec and $f_s = 2$ kHz be the duration of the data record and the sampling rate, and let $T = 1/f_s$ be the interval between samples. The number of samples within T_R is

$$N = \frac{T_R}{T} = f_s T_R = 1000 \text{ samples}$$

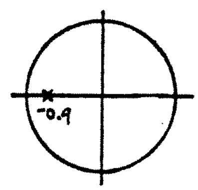
The resolution afforded by a length-M data window is $\Delta f = f_s/M$, from which we find the minimum acceptable M

$$M = \frac{f_s}{\Delta f} = \frac{2\text{ kHz}}{2\text{ Hz}} = 100 \text{ samples}$$

Thus, the maximum number of segments is $K = N/M = 10$.

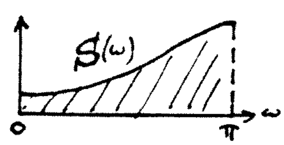
Problem 1.30:

1. $B(z) = \dfrac{1}{1 + 0.9z^{-1}}$ = AR model

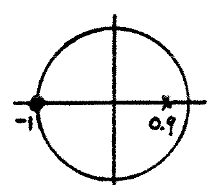


Since the pole of $B(z)$ lies in the high-frequency part of the unit circle, the spectrum $S(\omega) = |B(\omega)|^2$ will look as follows

$$S(\omega) = \left|\frac{1}{1 + 0.9e^{-j\omega}}\right|^2$$

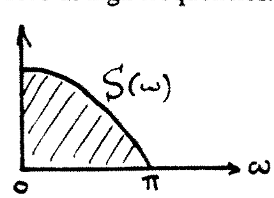


2. $B(z) = \dfrac{1 + z^{-1}}{1 - 0.9z^{-1}}$ = ARMA model



There is a pole at low frequencies and a zero at high frequencies. Thus,

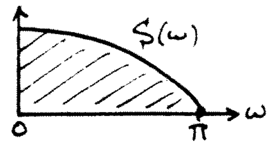
$$S(\omega) = \left|\frac{1 + e^{-j\omega}}{1 - 0.9e^{-j\omega}}\right|^2$$



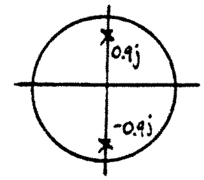
3. $B(z) = 1 + 2z^{-1} + z^{-2} = (1 + z^{-1})^2$ = MA model

It has a double zero at high frequencies:

$$S(\omega) = \left|1 + e^{-j\omega}\right|^4$$

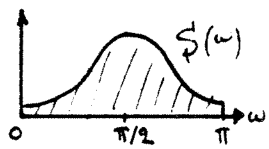


4. $B(z) = \dfrac{1}{1 + 0.81z^{-2}} = \dfrac{1}{(1 - 0.9jz^{-1})(1 + 0.9jz^{-1})}$ = AR model

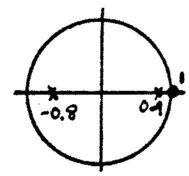


It has two poles at medium frequencies:

$$S(\omega) = \left|\frac{1}{1 + 0.81e^{-j\omega}}\right|^2$$

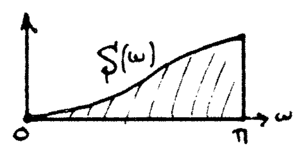


5. $B(z) = \dfrac{1 - 2z^{-1} + z^{-2}}{1 - 0.1z^{-1} - 0.72z^{-2}} = \dfrac{(1 - z^{-1})^2}{(1 + 0.8z^{-1})(1 - 0.9z^{-1})}$



It is ARMA, with an exact zero and a nearby pole at zero frequency, and a high frequency pole.

$$S(\omega) = \left|\frac{1 - 2e^{-j\omega} + e^{-2j\omega}}{1 - 0.1e^{-j\omega} - 0.72e^{-2j\omega}}\right|^2$$



Problem 1.33:
Expanding $\Delta S(\omega) = \dfrac{\partial S(\omega)}{\partial \sigma_\epsilon^2}\Delta\sigma_\epsilon^2 + \dfrac{\partial S(\omega)}{\partial a}\Delta a$, we find

$$E[(\Delta S(\omega))^2] = \left(\frac{\partial S(\omega)}{\partial \sigma_\epsilon^2}\right)^2 E[(\Delta\sigma_\epsilon^2)^2] + \left(\frac{\partial S(\omega)}{\partial a}\right)^2 E[(\Delta a)^2]$$

where we used $E[\Delta a \Delta\sigma_\epsilon^2] = 0$. The indicated partial derivatives are easily computed using the expression

$$S(\omega) = \frac{\sigma_\epsilon^2}{1 - 2a\cos\omega + a^2}$$

This gives,

$$\frac{\partial S(\omega)}{\partial \sigma_\epsilon^2} = \frac{1}{1 - 2a\cos\omega + a^2} = \frac{S(\omega)}{\sigma_\epsilon^2}$$

$$\frac{\partial S(\omega)}{\partial a} = \frac{\sigma_\epsilon^2 2(\cos\omega - a)}{(1 - 2a\cos\omega + a^2)^2} = S(\omega)\frac{2(\cos\omega - a)}{1 - 2a\cos\omega + a^2}$$

It follows that

$$E[(\Delta S(\omega))^2] = \frac{2S(\omega)^2}{N}\left[1 + \frac{2(1 - a^2)(\cos\omega - a)^2}{(1 - 2a\cos\omega + a^2)^2}\right]$$

Problem 1.34:
In terms of the eigenvalues λ_i of B, we have

$$\mathrm{tr}(B - I - \ln B) = \sum_i(\lambda_i - 1 - \ln\lambda_i)$$

The required result follows from the fact that $\lambda_i \geq 0$ and the inequality $\lambda - 1 - \ln\lambda \geq 0$, which is valid for all $\lambda \geq 0$ (with equality attained at $\lambda = 1$). Noting that $f(\hat{R}) = \mathrm{tr}(\ln\hat{R} + I)$, it follows

$$f(R) - f(\hat{R}) = \mathrm{tr}(R^{-1}\hat{R} + \ln R) - \mathrm{tr}(I + \ln\hat{R}) = \mathrm{tr}(B - I - \ln B) \geq 0$$

where we set $B = R^{-1}\hat{R}$. Note B is non negative definite, being the product of two such matrices. Also, we made use of the fact that for any two matrices $\mathrm{tr}(\ln R) - \mathrm{tr}(\ln\hat{R}) = \mathrm{tr}[\ln(R^{-1}\hat{R})]$, which follows easily from the first part of the next problem.

Problem 1.35:
The first relationship follows from

$$\ln(\det R) = \ln\left(\prod_i\lambda_i\right) = \sum_i\ln\lambda_i = \mathrm{tr}(\ln R)$$

where λ_i are the eigenvalues of R. The third follows by differentiating both sides of $RR^{-1} = I$ to get

$$Rd(R^{-1}) + (dR)R^{-1} = 0 \quad\Longrightarrow\quad d(R^{-1}) = -R^{-1}(dR)R^{-1}$$

The second identity follows by making use of the eigendecomposition of R; namely, $R = E\Lambda E^{-1}$, where E and Λ are the orthogonal matrix of eigenvectors and the diagonal matrix of eigenvalues. Taking differentials, we have

$$dR = dE\,\Lambda E^{-1} + E\,d\Lambda E^{-1} - E\Lambda E^{-1}\,dE\,E^{-1}$$

Multiplying by $R^{-1} = E\Lambda^{-1}E^{-1}$, we find

$$R^{-1}dR = E(\Lambda^{-1}d\Lambda)E^{-1} + (E\Lambda^{-1})(E^{-1}dE)(E\Lambda^{-1})^{-1} - dE\,E^{-1}$$

Taking traces and using the fact that the trace is invariant under similarity transformations, it follows that

$$\mathrm{tr}(R^{-1}dR) = \mathrm{tr}(\Lambda^{-1}d\Lambda) + \mathrm{tr}(E^{-1}dE) - \mathrm{tr}(dE\,E^{-1}) = \mathrm{tr}(\Lambda^{-1}d\Lambda) = d\,\mathrm{tr}(\ln\Lambda) = d\,\mathrm{tr}(\ln R)$$

where the last equality follows from $\ln R = E\,(\ln\Lambda)E^{-1}$.

Chapter 2

Problem 2.1:
In all cases, we find the transfer function $H(z)$ and use the relationship

$$S_{yy}(z) = H(z)H(z^{-1})S_{xx}(z) = H(z)H(z^{-1})$$

where we set $S_{xx}(z) = \sigma_x^2 = 1$.

1. The transfer function is $H(z) = 1 - z^{-1}$. Thus,

$$S_{yy}(z) = (1 - z^{-1})(1 - z) = 2 - (z + z^{-1})$$

Picking out the coefficients of z^{-k}, we find

$$R_{yy}(0) = 2\,, \quad R_{yy}(\pm 1) = -1\,, \quad R_{yy}(\pm k) = 0, \ \text{for} \ \ k \geq 2$$

The last result is of course related to the fact that the filter has memory of 1 sampling instant and therefore it can only introduce sequential correlations of lag at most 1.

2. Here, $H(z) = 1 - 2z^{-1} + z^{-2}$. Since the filter is of order 2, we expect to have nontrivial correlations only up to lag 2; indeed,

$$S_{yy}(z) = (1 - 2z^{-1} + z^{-2})(1 - 2z + z^2) = 6 - 4(z + z^{-1}) + (z^2 + z^{-2})$$

Inverting this z-transform, we find the nonzero autocorrelation lags:

$$R_{yy}(0) = 6\,, \quad R_{yy}(\pm 1) = -4\,, \quad R_{yy}(\pm 2) = 1$$

3. Because the filter $H(z) = 1/(1 - 0.5z^{-1})$ is recursive, we must use the contour inversion formula:

$$R_{yy}(k) = \int_{u.c.} S_{yy}(z) z^k \frac{dz}{2\pi jz} = \int_{u.c.} \frac{z^k}{(z - 0.5)(1 - 0.5z)} \frac{dz}{2\pi j}$$

Since $R_{yy}(k) = R_{yy}(-k)$, we only need to compute the above integral for $k \geq 0$. In this case, the integrand has only one pole enclosed by the unit circle; namely, $z = 0.5$. Therefore, we find

$$R_{yy}(k) = \frac{(0.5)^k}{1 - (0.5)^2} \quad \text{for} \quad k \geq 0$$

4. In this case we have:

$$H(z) = \frac{1}{1 - 0.25z^{-2}} = \frac{1}{(1 - 0.5z^{-1})(1 + 0.5z^{-1})}$$

The power spectral density is given by

$$S_{yy}(z) = \frac{1}{(1 - 0.25z^{-2})(1 - 0.25z^2)} = \frac{1}{(1 - 0.5z^{-1})(1 + 0.5z^{-1})(1 - 0.25z^2)}$$

For lags $k \geq 0$, the integrand of the contour inversion formula for $R_{yy}(k)$ has two poles inside the unit circle; namely, $z = \pm 0.5$. Evaluating the residues at these poles we find:

$$R_{yy}(k) = \int_{u.c.} S_{yy}(z) z^k \frac{dz}{2\pi jz} = \int_{u.c.} \frac{z^{k+1}}{(z - 0.5)(z + 0.5)(1 - 0.25z^2)} \frac{dz}{2\pi j} =$$

$$= \operatorname*{Res}_{z=0.5} + \operatorname*{Res}_{z=-0.5} = \frac{1}{1.875}[(0.5)^k + (-0.5)^k]$$

Problem 2.2:
This property is clearly seen for cross-periodograms:

$$\hat{S}_{yw}(z) = Y(z)W(z^{-1}) = H(z)X(z)W(z^{-1}) = H(z)\hat{S}_{xw}(z)$$

where we used the filtering equation $Y(z) = H(z)X(z)$. For the statistical quantities, the proof is best carried out in the time domain working with autocorrelations and using stationarity:

$$R_{yw}(k) = E[y_n w_{n-k}] = E[(\sum_m h_m x_{n-m}) w_{n-k}] = \sum_m h_m E[x_{n-m} w_{n-k}] =$$

$$= \sum_m h_m R_{xw}(n - m - n + k) = \sum_m h_m R_{xw}(k - m)$$

Taking z-transforms of both sides, gives the desired result. Part (b) follows from part (a) and the symmetry property $S_{wy}(z) = S_{yw}(z^{-1})$. Indeed,

$$S_{wy}(z) = S_{yw}(z^{-1}) = H(z^{-1})S_{xw}(z^{-1}) = S_{wx}(z)H(z^{-1})$$

Problem 2.3:
Write e_n in vector form, as follows:

$$e_n = \sum_{m=0}^{M} a_m y_{n-m} = [a_0, a_1, \cdots, a_M] \begin{bmatrix} y_n \\ y_{n-1} \\ \vdots \\ y_{n-M} \end{bmatrix} = \mathbf{a}^T \mathbf{y}(n)$$

Its mean square becomes:

$$E[e_n^2] = E[(\mathbf{a}^T \mathbf{y}(n))(\mathbf{y}(n)^T \mathbf{a})] = \mathbf{a}^T E[\mathbf{y}(n)\mathbf{y}(n)^T]\mathbf{a} = \mathbf{a}^T R_{yy} \mathbf{a}$$

where $R_{yy} = E[\mathbf{y}(n)\mathbf{y}(n)^T]$. Its matrix elements are expressed in terms of the autocorrelation lags, as follows:

$$(R_{yy})_{ij} = E[y_{n-i}y_{n-j}] = R_{yy}(n-i-n+j) = R_{yy}(j-i) = R_{yy}(i-j)$$

The second expression for $E[e_n^2]$ is obtained by noting that it is expressible as the zero-lag autocorrelation of e_n:

$$E[e_n^2] = R_{ee}(0) = \int_{-\pi}^{\pi} S_{ee}(\omega)\frac{d\omega}{2\pi} = \int_{-\pi}^{\pi} |A(\omega)|^2 S_{yy}(\omega)\frac{d\omega}{2\pi}$$

where we used Eq. (1.9.5) and $S_{ee}(\omega) = |A(\omega)|^2 S_{yy}(\omega)$, which follows from the fact that e_n is the output of the linear filter $A(z)$ when the input is y_n.

Problem 2.4:
Because $B(z) = 1/A(z)$ is the signal model of y_n, we must have:

$$S_{yy}(\omega) = \sigma_\epsilon^2 |B(\omega)|^2 = \frac{\sigma_\epsilon^2}{|A(\omega)|^2}$$

Using the results of Problem 2.3, applied to the filter $A'(z)$, we find

$$\sigma_e^2 = E[e_n^2] = \int_{-\pi}^{\pi} |A'(\omega)|^2 S_{yy}(\omega)\frac{d\omega}{2\pi} = \mathbf{a}'^T R_{yy}\mathbf{a}' \qquad \text{or,}$$

$$\sigma_e^2 = \int_{-\pi}^{\pi} |A'(\omega)|^2 \frac{\sigma_\epsilon^2}{|A(\omega)|^2}\frac{d\omega}{2\pi} = \mathbf{a}'^T R_{yy}\mathbf{a}'$$

Applying Problem 2.3 to ϵ_n itself, we get $\sigma_\epsilon^2 = \mathbf{a}^T R_{yy}\mathbf{a}$. We finally find

$$\frac{\sigma_e^2}{\sigma_\epsilon^2} = \int_{-\pi}^{\pi} \left|\frac{A'(\omega)}{A(\omega)}\right|^2 \frac{d\omega}{2\pi} = \frac{\mathbf{a}'^T R_{yy}\mathbf{a}'}{\mathbf{a}^T R_{yy}\mathbf{a}}$$

Part (b) is obtained by interchanging the roles of $A'(z)$ and $A(z)$.

Problem 2.5:
Using stationarity, we obtain

$$R_{yy}(k)^* = (E[y_{n+k}y_n^*])^* = E[y_n y_{n+k}^*] = R_{yy}(-k)$$

In the notation of Problem 2.3, we may write $e_n = \mathbf{a}^T \mathbf{y}(n)$. Its mean square value is

$$E[e^*_n e_n] = \mathbf{a}^\dagger E[\mathbf{y}(n)^* \mathbf{y}(n)^T]\mathbf{a} = \mathbf{a}^\dagger R_{yy}\mathbf{a}$$

The matrix elements of R_{yy} are expressed in terms of the autocorrelation lags:

$$(R_{yy})_{ij} = E[y_{n-i}^* y_{n-j}] = E[y_{n-j}y_{n-i}^*] = R_{yy}(n-j-n+i) = R_{yy}(i-j)$$

Problem 2.6:
Inserting $y_n = A_1 e^{j\phi_1} e^{j\omega_1 n}$ into the definition of Problem 2.5, we find:

$$R_{yy}(k) = E[y_{n+k} y_n^*] = E[A_1 e^{j\phi_1} e^{j\omega_1(n+k)} A_1^* e^{-j\phi_1} e^{-j\omega_1 n}] = |A_1|^2 e^{j\omega_1 k}$$

Next, let y_n be the sum

$$y_n = A_1 e^{j(\omega_1 n + \phi_1)} + A_2 e^{j(\omega_2 n + \phi_2)}$$

Then, we obtain the four terms

$$\begin{aligned} E[y_{n+k} y_n]^* = {} & |A_1|^2 e^{j\omega_1 k} + |A_2|^2 e^{j\omega_2 k} + \\ & + A_1 A_2^* e^{j\omega_1(n+k) - j\omega_2 n} E[e^{j\phi_1} e^{-j\phi_2}] + A_2 A_1^* e^{j\omega_2(n+k) - j\omega_1 n} E[e^{j\phi_2} e^{-j\phi_1}] \end{aligned}$$

Because ϕ_1 and ϕ_2 are independent, we have

$$E[e^{j\phi_1} e^{-j\phi_2}] = E[e^{j\phi_1}] E[e^{-j\phi_2}] = 0 \cdot 0 = 0$$

where we used the fact that for uniformly distributed ϕ, the expectation value $E[e^{j\phi}]$ is zero. Indeed,

$$E[e^{j\phi}] = \int_0^{2\pi} e^{j\phi} p(\phi) d\phi = \int_0^{2\pi} e^{j\phi} \frac{d\phi}{2\pi} = 0$$

It follows that the cross terms in the above autocorrelation are zero. Thus,

$$R_{yy}(k) = E[y_{n+k} y_n^*] = |A_1|^2 e^{j\omega_1 k} + |A_2|^2 e^{j\omega_2 k}$$

Problem 2.7:
We just showed that for mutually independent and uniformly distributed phase angles, $E[e^{j\phi_i} e^{-j\phi_l}] = 0$ if $\phi_i \neq \phi_l$. But if $\phi_i = \phi_l$, then $E[e^{j\phi_i} e^{-j\phi_l}] = E[1] = 1$. To summarize, we have

$$E[e^{j\phi_i} e^{-j\phi_l}] = \delta_{il}$$

Using this result and $E[v_n^* e^{j\phi_i}] = 0$ which follows from the independence of v_n and ϕ_i, we get

$$\begin{aligned} R_{yy}(k) = E[y_{n+k} y_n^*] & = E[v_{n+k} v_n^*] + \sum_{i=1}^{L} A_i e^{j\omega_i(n+k)} E[v_n^* e^{j\phi_i}] + \\ & + \sum_{l=1}^{L} A_l^* e^{-j\omega_l n} E[v_{n+k} e^{-j\phi_l}] + \sum_{i,j=1}^{L} A_i A_l^* e^{j\omega_i(n+k) - j\omega_l n} E[e^{j\phi_i} e^{-j\phi_l}] = \\ & = \sigma_v^2 \delta(k) + \sum_{i,l=1}^{L} A_i A_l^* e^{j\omega_i(n+k) - j\omega_l n} \delta_{il} = \end{aligned}$$

$$= \sigma_v^2 \delta(k) + \sum_{i=1}^{L} |A_i|^2 e^{j\omega_i k}$$

Using $E[\,|e_n|^2] = \mathbf{a}^\dagger R_{yy}\mathbf{a} = \sum_{l,m=0}^{M} a_l^* R_{yy}(l-m) a_m$ and $A(\omega_i) = \sum_{m=0}^{M} a_m e^{-j\omega_i m}$, we find

$$E[\,|e_n|^2] = \sum_{l,m=0}^{M} a_l^* \left[\sigma_v^2 \delta(l-m) + \sum_{i=1}^{L} |A_i|^2 e^{j\omega_i(l-m)} \right] a_m =$$

$$= \sigma_v^2 \sum_{m=0}^{M} |a_m|^2 + \sum_{i=1}^{L} |A_i|^2 \left[\sum_{l=0}^{M} a_l^* e^{j\omega_i l} \right] \left[\sum_{m=0}^{M} a_m e^{-j\omega_i m} \right] = \sigma_v^2 \mathbf{a}^\dagger \mathbf{a} + \sum_{i=1}^{L} |A_i|^2 |A(\omega_i)|^2$$

This last result forms the basis of Pisarenko's method of harmonic retrieval, as discussed in Section 6.2. Part (d) follows the same steps with the replacement of $\delta(l-m)$ by $Q(l-m)$ and recognizing that

$$\sum_{i,l=0}^{M} a_l^* Q(l-m) a_m = \mathbf{a}^\dagger Q \mathbf{a}$$

Problem 2.9:
Consider a more general problem of the form

$$y_n = -R^2 y_{n-2} + (1-R^2) x_n$$

with transfer function

$$H(z) = \frac{1-R^2}{1+R^2 z^{-2}} = \frac{1-R^2}{(1-jRz^{-1})(1+jRz^{-1})}$$

The power spectral density is

$$S_{yy}(z) = \sigma_x^2 H(z) H(z^{-1}) = \frac{(1-R^2)^2}{(1+R^2 z^{-2})(1+R^2 z^2)}$$

where $\sigma_x^2 = 1$. Setting $z = e^{j\omega}$

$$S_{yy}(\omega) = \frac{(1-R^2)^2}{1 + 2R^2\cos(2\omega) + R^4}$$

The autocorrelation function is obtained from

$$R_{yy}(k) = \oint_{u.c.} S_{yy}(z) z^k \frac{dz}{2\pi jz} = \oint_{u.c.} \frac{(1-R^2)^2 z^{k+1}}{(z-jR)(z+jR)(1+R^2z^2)} \frac{dz}{2\pi j} =$$

$$= \operatorname*{Res}_{z=jR} + \operatorname*{Res}_{z=-jR} = \frac{1-R^2}{1+R^2} R^k \cos\left(\frac{\pi k}{2}\right)$$

where we assumed $k \geq 0$. Since $\sigma_y^2 = R_{yy}(0)$, we find for the noise reduction ratio

$$\frac{\sigma_y^2}{\sigma_x^2} = \frac{R_{yy}(0)}{1} = \frac{1-R^2}{1+R^2}$$

Since the filter has unity gain at the frequencies $\omega = \pm\pi/2$, $H(\omega) = 1$, it follows that any linear combination of

$$\sin(\frac{\pi n}{2}) \quad \text{and} \quad \cos(\frac{\pi n}{2})$$

will go through the filter completely unchanged (in the steady state). For values of R close to unity, the steady state is reached more slowly, but the noise reduction ratio is smaller. This is the basic tradeoff between speed of response and effective noise reduction.

Problem 2.11:
In correlation canceler notation, we have $\mathbf{x} = [y_n]$, and $\mathbf{y} = [y_{n-1}]$. Then,

$$R_{xy} = E[y_n y_{n-1}] = R(1)\,, \quad R_{yy} = E[y_{n-1}^2] = R(0)\,, \quad H = R_{xy}R_{yy}^{-1} = \frac{R(1)}{R(0)}$$

The estimate $\hat{\mathbf{x}} = H\mathbf{y}$ becomes,

$$\hat{y}_n = Hy_{n-1} = -a_1 y_{n-1}$$

where we set $a_1 = -R(1)/R(0)$. The minimized estimation error is computed by

$$E[e^2] = R_{xx} - R_{xy}R_{yy}^{-1}R_{yx} = R(0) - \frac{R(1)^2}{R(0)} = R(0) + a_1 R(1)$$

Problem 2.12:
The sample autocorrelation is

$$\hat{R}(k) = \sum_{n=0}^{3-k} y_{n+k}y_n = \sum_{n=0}^{3-k} 1\cdot 1 = 4-k\,, \quad \text{for} \quad k = 0, 1, 2, 3$$

The resulting first order predictor and prediction error are

$$a_1 = -\frac{\hat{R}(1)}{\hat{R}(0)} = -\frac{3}{4}\,, \quad E = \hat{R}(0) + a_1\hat{R}(1) = 4 - \frac{3^2}{4} = \frac{7}{4} = 1.75$$

The estimate is:

$$\hat{y}_n = \frac{3}{4}y_{n-1} \quad \Longrightarrow \quad \hat{y}_4 = \frac{3}{4}\cdot 1 = 0.75$$

The gapped function is defined by

$$g(k) = \sum_{m=0}^{1} a_m \hat{R}(k-m) = \hat{R}(k) - \frac{3}{4}\hat{R}(k-1)$$

It has a gap of length one:

$$g(1) = \hat{R}(1) - \frac{3}{4}\hat{R}(0) = 3 - \frac{3}{4}\cdot 4 = 0$$

Next compute $g(0)$ and $g(2)$:

$$g(0) = \hat{R}(0) - \frac{3}{4}\hat{R}(-1) = 3 - \frac{3}{4}\cdot 3 = \frac{7}{4}, \quad g(2) = \hat{R}(2) - \frac{3}{4}\hat{R}(1) = 2 - \frac{3}{4}\cdot 3 = -\frac{1}{4}$$

Next define the second order gapped function $g'(k)$:

$$g'(k) = g(k) - \gamma_2 g(2-k)$$

where $\gamma_2 = g(2)/g(0) = -1/7$. Using the Levinson recursion, we construct the second order predictor by:

$$\begin{bmatrix} 1 \\ a_1' \\ a_2' \end{bmatrix} = \begin{bmatrix} 1 \\ a_1 \\ 0 \end{bmatrix} - \gamma_2 \begin{bmatrix} 0 \\ a_1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ -3/4 \\ 0 \end{bmatrix} + \frac{1}{7}\begin{bmatrix} 0 \\ -3/4 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ -6/7 \\ 1/7 \end{bmatrix}$$

The 2nd order prediction error is

$$E' = g'(0) - (1-\gamma_2^2)g(0) = (1 - \frac{1}{49})\cdot\frac{7}{4} = 1.714$$

As expected, it is smaller than the error of the 1st order predictor. The estimate is given by

$$\hat{y}_n = \frac{6}{7}y_{n-1} - \frac{1}{7}y_{n-2} \quad \Longrightarrow \quad \hat{y}_4 = \frac{6}{7} - \frac{1}{7} = 0.714$$

Even though the 2nd order predictor is better than the 1st order one in the mean square sense, the actual "predicted" value of the 5th sample is "worse" than that predicted by the 1st order predictor (assuming of course that the most "obvious" value should be $y_4 = 1$).

Finally, the zeros of the 2nd order predictor are found by solving the quadratic equation

$$1 - \frac{6}{7}z^{-1} + \frac{1}{7}z^{-2} = 0$$

It has roots $z = 0.227$ and $z = 0.631$.

Problem 2.13:
We have $y_n = -(-1)^n$ for $n = 0,1,2,3$. The sample autocorrelation is

$$\hat{R}(k) = \sum_{n=0}^{3-k} y_{n+k}y_n = (-1)^k(4-k)\,, \quad \text{for} \quad k = 0,\,1,\,2,\,3$$

Then,

$$a_1 = -\frac{\hat{R}(1)}{\hat{R}(0)} = \frac{3}{4}\,, \quad \hat{y}_n = -a_1 y_{n-1}\,, \quad \hat{y}_4 = -\frac{3}{4}\cdot 1 = -0.75$$

Similarly, we find for the autocorrelation of $\mathbf{y} = [1, 2, 3, 4]$:

$$[\hat{R}(0), \hat{R}(1), \hat{R}(2), \hat{R}(3)] = [30, 20, 11, 4]$$

and $a_1 = -\hat{R}(1)/\hat{R}(0) = -2/3$.

Problem 2.14:
The equation at the upper adder is

$$e_n + \gamma_1 y_{n-1} = y_n \quad \Longrightarrow \quad e_n = y_n - \gamma_1 y_{n-1} = y_n + a_1 y_{n-1}$$

Thus, we find

$$\frac{Y(z)}{e(z)} = \frac{1}{1 + a_1 z^{-1}} = \frac{1}{A(z)}$$

Similarly, at the lower adder we have $r_n = y_{n-1} - \gamma_1 y_n$, which results in the transfer function

$$r(z) = z^{-1}Y(z) - \gamma_1 Y(z) = (a_1 + z^{-1})Y(z) = A^R(z)Y(z)$$

Problem 2.15:
In the z-domain, the equation at the second upper adder is $e(z) = e'(z) + \gamma_2 z^{-1} r(z)$, or

$$e'(z) = e(z) - \gamma_2 z^{-1} r(z) = [A(z) - \gamma_2 z^{-1} A^R(z)]Y(z) = A'(z)Y(z)$$

where we used the Levinson recursion $A'(z) = A(z) - \gamma_2 z^{-1} A^R(z)$.

Problem 2.16:
Let p_1 and p_2 be the two zeros of the prediction-error filter, so that

$$A'(z) = 1 + a_1' z^{-1} + a_2' z^{-2} = (1 - p_1 z^{-1})(1 - p_2 z^{-1})$$

Because of the Levinson recursion, we have

$$p_1 + p_2 = -a_1' = \gamma_1(1 - \gamma_2)\,, \quad p_1 p_2 = a_2' = -\gamma_2$$

The required integral is

$$I = \int_{u.c.} \frac{1}{A'(z)A'(z^{-1})} \frac{dz}{2\pi jz} = \int_{u.c} \frac{z}{(z-p_1)(z-p_2)(1-p_1z)(1-p_2z)} \frac{dz}{2\pi j}$$

Picking out the residues at p_1 and p_2, we find

$$I = \operatorname*{Res}_{z=p_1} + \operatorname*{Res}_{z=p_2} = \frac{1+p_1p_2}{(1-p_1p_2)(1-p_1^2)(1-p_2^2)}$$

Using the identity

$$(1-p_1^2)(1-p_2^2) = (1+p_1p_2)^2 - (p_1+p_2)^2$$

and expressing the ps in terms of the γs, we find finally:

$$I = \frac{1}{(1-\gamma_1^2)(1-\gamma_2^2)}$$

Chapter 3

Problem 3.1:
Starting with the right hand side, we have

$$A(z)\bar{A}(z) = \left(\sum_{i=0}^{M} a_i z^{-i}\right)\left(\sum_{n=0}^{M} a_n^* z^{n}\right) = \sum_{i,n=0}^{M} a_i a_n^* z^{-(i-n)} = \sum_{k,n} a_{n+k} a_n^* z^{-k}$$

where we defined $k = i - n$ and changed summation variables from the pair i,n to the pair k,n. We must now determine the proper range of summation over k and n - doing the n summation first and the k summation last. Since both i and n are in the range $[0, M]$, it follows that k will range in $-M \le k \le M$. Since $i = n + k$, it follows that $0 \le n + k \le M$, or, $-k \le n \le M - k$. But also, $0 \le n \le M$. Therefore, n ranges in the intersection of these two intervals; namely,

$$\max(0, -k) \le n \le \min(M, M-k)$$

Thus,

$$A(z)\bar{A}(z) = \sum_{k=-M}^{M}\left(\sum_{n=\max(0,-k)}^{\min(M,M-k)} a_{n+k} a_n^*\right) z^{-k} = \sum_{k=-M}^{M} r(k) z^{-k}$$

The quantity $r(k)$ is identical to $R_{aa}(k)$. Indeed, if $0 \le k \le M$, then

$$r(k) = \sum_{n=\max(0,-k)}^{\min(M,M-k)} a_{n+k} a_n^* = \sum_{n=0}^{M-k} a_{n+k} a_n^* = R_{aa}(k)$$

Similarly, if $-M \le k \le -1$, the range in the summation of $r(k)$ is $-k \le n \le M$, and we find

$$r(k) = \sum_{n=-k}^{M} a_{n+k} a_n^* = \sum_{i=0}^{M+k} a_{i-k}^* a_i = R_{aa}(-k)^* = R_{aa}(k)$$

Problem 3.2:
We start with Eqs. (3.4.2) which are the time-domain versions of Eqs. (3.4.1). We have

$$|a_n|^2 - |b_n|^2 = |f_n - z_1 f_{n-1}|^2 - |-z_1^* f_n + f_{n-1}|^2 = (|f_n|^2 - |f_{n-1}|^2)(1 - |z_1|^2)$$

Summing over n, we find

$$\sum_{m=0}^{n} |a_m|^2 - \sum_{m=0}^{n} |b_m|^2 = (1 - |z_1|^2) \sum_{m=0}^{n} (|f_m|^2 - |f_{m-1}|^2) = (1 - |z_1|^2)|f_n|^2$$

where we used the fact that $f_{-1} = 0$.

Problem 3.3:
Using the following standard z-transform

$$a^{|k|} \quad \rightarrow \quad \frac{1 - a^2}{(1 - az^{-1})(1 - az)}$$

we find

$$S_{yy}(z) = \sum_{k=-\infty}^{\infty} (0.5)^{|k|} z^{-k} = \frac{1 - (0.5)^2}{(1 - 0.5z^{-1})(1 - 0.5z)}$$

Thus, the model is

$$B(z) = \frac{1}{1 - 0.5z^{-1}}, \quad \text{and} \quad \sigma_\epsilon^2 = 1 - (0.5)^2 = 0.75$$

The difference equation is $y_n = 0.5y_{n-1} + \epsilon_n$.

Problem 3.4:
As in the previous problem, we find

$$S_{yy}(z) = \frac{1 - (0.5)^2}{(1 - 0.5z^{-1})(1 - 0.5z)} + \frac{1 - (-0.5)^2}{(1 + 0.5z^{-1})(1 + 0.5z)} = 1.875 \cdot \frac{1}{1 - 0.25z^{-2}} \cdot \frac{1}{1 - 0.25z^2}$$

Therefore, $\sigma_\epsilon^2 = 1.875$, and

$$B(z) = \frac{1}{1 - 0.25z^{-2}}, \quad y_n = 0.25y_{n-2} + \epsilon_n$$

Problem 3.5:
See solution of Problem 2.9.

Problem 3.6:
Factor the numerator in the form:

$$2.18 - 0.6(z + z^{-1}) = \sigma^2(1 - fz^{-1})(1 - fz) = \sigma^2(1 + f^2) - \sigma^2 f(z + z^{-1})$$

This requires $2.18 = \sigma^2(1 + f^2)$ and $0.6 = \sigma^2 f$. Eliminating σ^2, we find

$$\frac{1 + f^2}{f} = \frac{2.18}{0.6} \quad \Longrightarrow \quad f = 0.3$$

where we chose the solution for f of magnitude less than one. Then, we have $\sigma^2 = 0.6/f = 2$. Thus,

$$2.18 - 0.6(z + z^{-1}) = 2(1 - 0.3z^{-1})(1 - 0.3z)$$

The denominator is factored in a similar fashion to give

$$1.25 - 0.5(z + z^{-1}) = (1 - 0.5z^{-1})(1 - 0.5z)$$

Therefore, $S_{yy}(z)$ is factored as follows

$$S_{yy}(z) = 2\cdot\frac{1 - 0.3z^{-1}}{1 - 0.5z^{-1}}\cdot\frac{1 - 0.3z}{1 - 0.5z}$$

We identify, $\sigma_\epsilon^2 = 2$, and

$$B(z) = \frac{1 - 0.3z^{-1}}{1 - 0.5z^{-1}} \quad \text{and} \quad y_n = 0.5y_{n-1} + \epsilon_n - 0.3\epsilon_{n-1}$$

Problem 3.7:
Using $y_n = cx_n + v_n$ and the fact that x_n and v_n are uncorrelated, we find

$$S_{yy}(z) = c^2 S_{xx}(z) + S_{vv}(z) = \frac{c^2 Q}{(1 - az^{-1})(1 - az)} + R$$

The required signal model $B(z)$ is found by completing the fraction and factoring the numerator in the form (with $|f| < 1$):

$$S_{yy}(z) = \frac{c^2 Q + R(1 - az^{-1})(1 - az)}{(1 - az^{-1})(1 - az)} = \frac{\sigma_\epsilon^2(1 - fz^{-1})(1 - fz)}{(1 - az^{-1})(1 - az)}$$

Therefore, we must have the identity in z:

$$c^2 Q + R(1 - az^{-1})(1 - az) = \sigma_\epsilon^2(1 - fz^{-1})(1 - fz)$$

which gives rise to

$$\sigma_\epsilon^2(1 + f^2) = c^2 Q + R(1 + a^2)\,, \qquad \sigma_\epsilon^2 f = Ra$$

Solving the second, $\sigma_\epsilon^2 = Ra/f$, and inserting in the first, we find the quadratic equation for f:

$$aR(1 + f^2) = f\,[c^2 Q + R(1 + a^2)]$$

This equation remains invariant under the substitution $f \to 1/f$. Therefore, if one solution has magnitude less than one, the other will have magnitude greater than one. Substituting the expression

$$f = \frac{Ra}{R + c^2 P}$$

into the quadratic equation for f, we find

$$aR\left[1 + \left(\frac{Ra}{R + c^2P}\right)^2\right] = \frac{Ra}{R + c^2P}\left[c^2Q + R(1 + a^2)\right]$$

or,

$$\frac{(R + c^2P)^2 + (Ra)^2}{R + c^2P} = c^2Q + R(1 + a^2)$$

which gives after some algebra

$$Q = P - \frac{Ra^2P}{R + c^2P}$$

If P is positive, then f has magnitude less than one. Indeed, since by assumption $|a| \leq 1$, we have

$$|f| = \frac{|a|}{1 + x} \leq \frac{1}{1 + x} < 1, \quad \text{where} \quad x \equiv \frac{c^2P}{R}$$

To show that one solution for P is positive and the other negative, we work with the variable x defined above. Then, the quadratic Riccati equation reads

$$\frac{x^2 + (1 - a^2)x}{1 + x} = \beta, \quad \text{where} \quad \beta \equiv \frac{c^2Q}{R}$$

It can be written in the form

$$x^2 - \alpha x - \beta = 0$$

where $\alpha = \beta - (1 - a^2)$. The two solutions are

$$x_1 = \frac{\alpha + \sqrt{\alpha^2 + 4\beta}}{2}, \quad x_2 = \frac{\alpha - \sqrt{\alpha^2 + 4\beta}}{2}$$

Because $\beta > 0$, it follows that regardless of the value of α, x_1 will be positive and x_2 negative.

Finally, inserting the expression of f in terms of P into σ_ϵ^2, we find

$$\sigma_\epsilon^2 = \frac{Ra}{f} = R + c^2P$$

Problem 3.8:

For $n < i$ we have $B_{ni} = b_{n-i} = 0$ because of the causality of b_n. Also, $(B^T)_{ni} = B_{in} = b_{i-n} = b_{-(n-i)}$. Therefore, it is the matrix that corresponds to b_{-n}. If $C = AB$, then $C_{ni} = (AB)_{ni} = \sum_k A_{nk}B_{ki}$ $= \sum_k a_{n-k}b_{k-i}$. Set $m = k - i$ to get

Chapter 4

Problem 4.1:
The ML estimate is that $\mathbf{x}$ which maximizes the conditional density $p(\mathbf{y} \mid \mathbf{x})$. This density is easily found by recognizing that if $\mathbf{x}$ is given, then the only randomness left in $\mathbf{y}$ arises from the gaussian noise term $\mathbf{v}$; i.e.,

$$p(\mathbf{y} \mid \mathbf{x}) d^N\mathbf{y} = p_v(\mathbf{v}) d^N\mathbf{v}$$

where $N = n_b - n_a + 1$ is the dimension of $\mathbf{y}$. The Jacobian of the transformation from $\mathbf{v}$ to $\mathbf{y}$ is unity, that is, $d^N\mathbf{y} = d^N\mathbf{v}$. Thus,

$$p(\mathbf{y} \mid \mathbf{x}) = p_v(\mathbf{v}) = \frac{e^{-|\mathbf{v}|^2/\sigma_v^2}}{(2\pi)^{N/2}\sigma_v^N} = \frac{e^{-|\mathbf{y}-C\mathbf{x}|^2/\sigma_v^2}}{(2\pi)^{N/2}\sigma_v^2}$$

Maximizing this density with respect to $\mathbf{x}$ is equivalent to minimizing the exponent $\mathcal{E} = |\mathbf{y} - C\mathbf{x}|^2$ with respect to $\mathbf{x}$. Setting the corresponding gradient to zero gives

$$\frac{\partial \mathcal{E}}{\partial \mathbf{x}} = -2C^T(\mathbf{y} - C\mathbf{x}) = 0$$

or, equivalently,

$$C^T C\mathbf{x} = C^T\mathbf{y} \quad \Longrightarrow \quad \mathbf{x} = (C^T C)^{-1} C^T \mathbf{y}$$

Problem 4.2:
The optimal H was found to be:

$$H_{\text{opt}} = E[\mathbf{x}\mathbf{y}^T]E[\mathbf{y}\mathbf{y}^T]^{-1} = R_{xy}R_{yy}^{-1}$$

a. The above is equivalent to the correlation canceler discussed in Section 1.4. There, it was shown that $H = H_{\text{opt}}$ minimizes the estimation error covariance matrix $R_{ee} = E[\mathbf{e}\mathbf{e}^T]$. In fact, we found there that for a deviation $H = H_{\text{opt}} + \Delta H$ from the optimal value, the error covariance matrix was $R_{ee} = R_{ee}^{\text{opt}} + \Delta H R_{yy} \Delta H^T$. And, it is minimized when $\Delta H = 0$.

b. Using this result and the matrix identity $\mathbf{e}^T Q \mathbf{e} = \text{tr}(\mathbf{e}\mathbf{e}^T Q)$, we find

$$E[\mathbf{e}^T Q \mathbf{e}] = \text{tr}\big(E[\mathbf{e}\mathbf{e}^T]Q\big) = \text{tr}(R_{ee}Q) = \text{tr}(R_{ee}^{\text{opt}}Q) + \text{tr}(\Delta H R_{yy} \Delta H^T Q)$$

Since Q is positive semi-definite, the second term will be non-negative, and again the minimum value will be attained at $\Delta H = 0$.

c. For any given n, let Q_n be the matrix that has 1 at the nth slot of its diagonal and zeros everywhere else, that is, its matrix elements are

$$(Q_n)_{ij} = \delta_{in}\,\delta_{jn}$$

Then, it is shown easily that

$$E[\mathbf{e}^T Q\mathbf{e}] = E[e_n^2]$$

Problem 4.3:
The smoothing estimate is $\hat{\mathbf{x}} = H\mathbf{y}$, where $H = R_{xy}R_{yy}^{-1}$. We find

$$R_{xy} = E[\mathbf{x}\mathbf{y}^T] = E[\mathbf{x}(\mathbf{x}^T C^T + \mathbf{v}^T)] = R_{xx}C^T$$
$$R_{yy} = E[\mathbf{y}\mathbf{y}^T] = E[(C\mathbf{x} + \mathbf{v})(\mathbf{x}^T C^T + \mathbf{v}^T)] = CR_{xx}C^T + R_{vv}$$

where we used the fact that $R_{xv} = 0$. It follows that

$$H = R_{xy}R_{yy}^{-1} = R_{xx}C^T[CR_{xx}C^T + R_{vv}]^{-1}$$

Problem 4.4:
In all Wiener filtering problems the optimal estimate is given by

$$\hat{x}_n = E[x_n\mathbf{y}^T]E[\mathbf{y}\mathbf{y}^T]^{-1}\mathbf{y} = H\mathbf{y}$$

In each case, we must identify the vector of observations $\mathbf{y}$ on which the estimate is based. For a first order Wiener filter, we have $\mathbf{y} = \begin{bmatrix} y_n \\ y_{n-1} \end{bmatrix}$. Thus,

$$E[x_n\mathbf{y}^T] = E[x_n[y_n, y_{n-1}]] = [R_{xy}(0), R_{xy}(1)]$$
$$E[\mathbf{y}\mathbf{y}^T] = E[\begin{bmatrix} y_n \\ y_{n-1} \end{bmatrix}[y_n, y_{n-1}]] = \begin{bmatrix} R_{yy}(0) & R_{yy}(1) \\ R_{yy}(1) & R_{yy}(0) \end{bmatrix}$$

Using $y_n = x_n + v_n$ and the fact that x_n is uncorrelated with v_n, we find

$$R_{xy}(k) = R_{xx}(k) + R_{xv}(k) = R_{xx}(k) = \sigma_x^2 a^{|k|}$$
$$R_{yy}(k) = R_{xx}(k) + R_{vv}(k) = \sigma_x^2 a^{|k|} + \sigma_v^2\delta(k)$$

It follows that

$$E[x_n\mathbf{y}^T] = \sigma_x^2[1, a]\,, \quad E[\mathbf{y}\mathbf{y}^T] = \sigma_x^2\begin{bmatrix} 1+\rho & a \\ a & 1+\rho \end{bmatrix}$$

where $\rho = \dfrac{\sigma_v^2}{\sigma_x^2}$ is a noise to signal ratio. Then,

$$H = E[x_n y^T]E[yy^T]^{-1} = [1, a]\begin{bmatrix} 1+\rho & a \\ a & 1+\rho \end{bmatrix}^{-1} = \left[\frac{1+\rho-a^2}{(1+\rho)^2-a^2}, \frac{a\rho}{(1+\rho)^2-a^2}\right]$$

and the estimate becomes

$$\hat{x}_n = Hy = H\begin{bmatrix} y_n \\ y_{n-1} \end{bmatrix} = \frac{1+\rho-a^2}{(1+\rho)^2-a^2}y_n + \frac{a\rho}{(1+\rho)^2-a^2}y_{n-1}$$

Note that if the noise term is absent, $\sigma_v^2 = 0$, then $\rho = 0$ and we have $\hat{x}_n = y_n$, as it should because then $x_n = y_n$. The 2nd order predictor is handled in a similar manner. Now, $y = \begin{bmatrix} y_{n-1} \\ y_{n-2} \end{bmatrix}$ and

$$E[x_n y^T] = E[x_n[y_{n-1}, y_{n-2}]] = [R_{xy}(1), R_{xy}(2)] = a[1, a]\sigma_x^2$$

This differs by a factor a from part (a); also, because of stationarity, the vector y will have the same autocorrelation matrix as that of part (a). Thus, we find

$$H_{\text{pred}} = E[x_n y^T]E[yy^T]^{-1} = a[1, a]\begin{bmatrix} 1+\rho & a \\ a & 1+\rho \end{bmatrix}^{-1} = aH$$

where H was given in part (a). Therefore,

$$\hat{x}_n = H_{\text{pred}}y = H\begin{bmatrix} y_{n-1} \\ y_{n-1} \end{bmatrix} = a\frac{1+\rho-a^2}{(1+\rho)^2-a^2}y_{n-1} + a\frac{a\rho}{(1+\rho)^2-a^2}y_{n-1}$$

If we denote by $x_{n/n}$ the optimal filtered estimate based on $\{y_n, y_{n-1}\}$ and by $x_{n/n-1}$ the optimal predicted estimate based on $\{y_{n-1}, y_{n-2}\}$, then we have found the interesting relationship

$$\hat{x}_{n/n-1} = a\hat{x}_{n-1/n-1} \quad \Longrightarrow \quad \hat{x}_{n+1/n} = a\hat{x}_{n/n}$$

The given expression for $R_{xx}(k)$ implies the signal model for x_n:

$$x_n = a x_{n-1} + \epsilon_n, \quad \sigma_\epsilon^2 = (1-a^2)\sigma_x^2$$

Thus, the above relationship between the predicted and filtered estimates is very reasonable. It states that to get the optimal prediction of x_n on the basis of the past two values of y_n, first find the optimal filtered estimate of x_{n-1} based on the same two past values, and then boost it ahead in time according to the system's transition matrix.

Problem 4.5:
Writing the estimate in vector form, we have

$$\hat{x}_n = h(n,n_a)y(n_a) + h(n,n_b)y(n_b) = [h(n,n_a), h(n,n_b)]\begin{bmatrix} y(n_a) \\ y(n_b) \end{bmatrix} = Hy$$

It is given that $y(n_a) = x_{n_a}$ and $y(n_b) = x_{n_b}$. The required correlations are computed as follows:

$$E[x_n y] = E[x_n[x_{n_a}, x_{n_b}]] = [R_{xx}(n-n_a), R_{xx}(n-n_b)] = \sigma_x^2[a^{|n-n_a|}, a^{|n-n_b|}]$$

Similarly,

$$E[yy^T] = E\left[\begin{bmatrix} x_{n_a} \\ x_{n_b} \end{bmatrix}[x_{n_a}, x_{n_b}]\right] = \begin{bmatrix} R_{xx}(0) & R_{xx}(n_a - n_b) \\ R_{xx}(n_b - n_a) & R_{xx}(0) \end{bmatrix} = \sigma_x^2 \begin{bmatrix} 1 & a^{|n_a - n_b|} \\ a^{|n_b - n_a|} & 1 \end{bmatrix}$$

Solving for $H = E[x_n y^T]E[yy^T]^{-1}$, we find

$$H = \frac{1}{1 - a^{2(n_b - n_a)}}\left[a^{|n-n_a|} - a^{|n-n_b| + n_b - n_a}, a^{|n-n_b|} - a^{|n-n_a| + n_b - n_a}\right]$$

This expression is valid for all values of n.

Problem 4.6:
Since x_n and v_n are uncorrelated, we have

$$S_{yy}(z) = S_{xx}(z) + S_{vv}(z) = \frac{1}{(1 - 0.5z^{-1})(1 - 0.5z)} + 5 = 6.25 \cdot \frac{1 - 0.4z^{-1}}{1 - 0.5z^{-1}} \cdot \frac{1 - 0.4z}{1 - 0.5z}$$

Therefore, we identify the signal model for y_n:

$$B(z) = \frac{1 - 0.4z^{-1}}{1 - 0.5z^{-1}}, \quad \sigma_\epsilon^2 = 6.25$$

Similarly, we find $S_{xy}(z) = S_{xx}(z) = \dfrac{1}{(1 - 0.5z^{-1})(1 - 0.5z)}$. And,

$$\left[\frac{S_{xy}(z)}{B(z^{-1})}\right]_+ = \left[\frac{\dfrac{1}{(1 - 0.5z^{-1})(1 - 0.5z)}}{\dfrac{1 - 0.4z}{1 - 0.5z}}\right]_+ = \left[\frac{1}{(1 - 0.5z^{-1})(1 - 0.4z)}\right]_+ = \frac{1.25}{1 - 0.5z^{-1}}$$

The corresponding Wiener filter is

$$H(z) = \frac{1}{\sigma_\epsilon^2 B(z)}\left[\frac{S_{xy}(z)}{B(z^{-1})}\right]_+ = \frac{\dfrac{1.25}{1 - 0.5z^{-1}}}{6.25 \cdot \dfrac{1 - 0.4z^{-1}}{1 - 0.5z^{-1}}} = \frac{0.2}{1 - 0.4z^{-1}}$$

with a difference equation $\hat{x}_{n/n} = 0.4\hat{x}_{n-1/n-1} + 0.2y_n$. The estimation error is computed by the contour integral:

$$\mathcal{E} = E[e^2_{n/n}] = \int_{u.c.} [S_{xx}(z) - H(z)S_{yx}(z)]\frac{dz}{2\pi jz} = \int_{u.c.} \frac{0.8}{(z-0.4)(1-0.5z)}\frac{dz}{2\pi j} = \operatorname*{Res}_{z=0.4} = 1$$

The improvement over not filtering at all and using y_n as the estimate of x_n is $\frac{E[e^2_{n/n}]}{\sigma_v^2} = 1/5$, or 6.9 dB.

The prediction part is handled by defining $x_1(n) = x_{n+1}$. The problem of predicting x_{n+1} on the basis of Y_n is equivalent to the problem of estimating $x_1(n)$ on the basis of Y_n. Using the filtering equation $X_1(z) = zX(z)$, we find

$$S_{x_1y}(z) = zS_{xy}(z) = \frac{z}{(1-0.5z^{-1})(1-0.5z)}$$

The optimal prediction filter is

$$H_1(z) = \frac{[G(z)]_+}{\sigma_e^2 B(z)}, \quad \text{where} \quad G(z) = \frac{S_{x_1y}(z)}{B(z^{-1})} = \frac{z}{(1-0.5z^{-1})(1-0.4z)}$$

The causal part of G(z)is found by first computing the inverse z-transform g_n and then summing it up for $n \geq 0$:

$$g_n = \int_{u.c.} G(z)z^n \frac{dz}{2\pi jz} = \int_{u.c.} \frac{z^{n+1}}{(z-0.5)(1-0.4z)}\frac{dz}{2\pi j} = 0.4\,(0.5)^n$$

where $n \geq 0$. Therefore,

$$[G(z)]_+ = \sum_{n=0}^{\infty} g_n z^{-n} = \frac{0.4}{1-0.5z^{-1}}$$

The prediction filter is then

$$H_1(z) = \frac{0.1}{1-0.4z^{-1}} \quad \text{or,} \quad \hat{x}_{n+1/n} = 0.4\hat{x}_{n/n-1} + 0.1y_n$$

where we denoted $\hat{x}_{n+1/n} = \hat{x}_1(n)$. We note that the prediction filter $H_1(z)$ is related to the estimation filter $H(z)$ by $H_1(z) = 0.5H(z)$. Since both filters have the same input, that is, y_n, it follows that

$$\hat{x}_{n+1/n} = 0.5\hat{x}_{n/n}$$

Noting that $S_{x_1x_1}(z) = zS_{xx}(z)z^{-1} = S_{xx}(z)$, and $S_{yx_1}(z) = S_{yx}(z)z^{-1}$, we find for the prediction error

$$\mathcal{E}_1 = E[e^2_{n+1/n}] = \int_{u.c.} [S_{x_1x_1}(z) - H_1(z)S_{yx_1}(z)]\frac{dz}{2\pi jz} = \int_{u.c.} \frac{1}{(z-0.4)(1-0.5z)}\frac{dz}{2\pi j} = 1.25$$

It is slightly worse than the estimation error $\mathcal{E}$ because the estimation filter uses one more observation than the prediction filter, and thus, makes a better estimate.

Next, we derive the above results using a Kalman filter formulation. The state and measurement equations are

$$x_{n+1} = 0.5x_n + w_n \quad \text{and} \quad y_n = x_n + v_n$$

with parameters

$$a = 0.5\,, \quad c = 1\,, \quad Q = 1\,, \quad R = 5$$

The corresponding algebraic Riccati equation is:

$$P - \frac{PRa^2}{R + c^2P} = Q \quad \Longrightarrow \quad P - \frac{1.25P}{5 + P} = 1$$

with positive solution $P = 1.25$. The Kalman gains are

$$G = \frac{cP}{R + c^2P} = 0.2\,, \quad K = aG = 0.1$$

We also find

$$f = a - cK = 0.4\,, \quad \sigma_\varepsilon^2 = R + c^2P = 6.25$$

Therefore, the prediction and filtering equations will be

$$\hat{x}_{n+1/n} = f\hat{x}_{n/n-1} + Ky_n = 0.4\hat{x}_{n/n-1} + 0.1y_n$$
$$\hat{x}_{n/n} = f\hat{x}_{n-1/n-1} + Gy_n = 0.4\hat{x}_{n-1/n-1} + 0.2y_n$$

Finally, note

$$\mathcal{E}_1 = E[e_{n+1/n}^2] = P = 1.25\,, \quad \mathcal{E} = E[e_{n/n}^2] = \frac{RP}{R + c^2P} = 1$$

Problem 4.7:
Solving the Riccati equation

$$P - \frac{PRa^2}{R + c^2P} = Q$$

with $Q = 0.5$, $R = 1$, $a = 1$, and $c = 1$, we find $P = 1$. Thus,

$$G = \frac{cP}{R + c^2P} = 0.5\,, \quad K = aG = 0.5\,, \quad f = a - cK = 0.5\,, \quad \sigma_\varepsilon^2 = R + c^2P = 2$$

The prediction and filtering equations become

$$\hat{x}_{n+1/n} = 0.5\hat{x}_{n/n-1} + 0.5y_n$$
$$\hat{x}_{n/n} = 0.5x_{n-1/n-1} + 0.5y_n$$

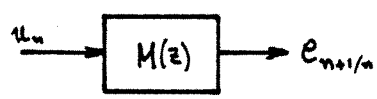
Note, $\hat{x}_{n+1/n} = a\hat{x}_{n/n} = \hat{x}_{n/n}$. Also, because the signal model is marginally stable (rather than strictly stable), the signal x_n is not truly stationary. Therefore, it is not entirely correct to apply the above Kalman filter methods which were derived by assuming strict stationarity. However, as shown in Example 4.9.1, the time-varying Kalman filter converges asymptotically to the above stationary one.

Problem 4.8:
First, we derive a difference equation for $e_{n/n-1} = x_n - \hat{x}_{n/n-1}$.

$$e_{n+1/n} = x_{n+1} - \hat{x}_{n+1/n} = (ax_n + w_n) - (f\hat{x}_{n/n-1} + Ky_n) = fe_{n/n-1} + w_n - Kv_n$$

where we used $y_n = cx_n + v_n$ and $a = f + cK$. We may think of $e_{n+1/n}$ as the output of the filter

$$M(z) = \frac{1}{1 - fz^{-1}}$$



driven by the white noise input $u_n = w_n - Kv_n$. Because w_n and v_n are uncorrelated, the variance of u_n will be

$$\sigma_u^2 = E[u_n^2] = E[w_n^2] + K^2E[v_n^2] = Q + K^2R$$

The power spectral density of $e_{n+1/n}$ will be

$$S_{ee}(z) = \sigma_u^2 M(z)M(z^{-1}) = \frac{Q + K^2R}{(1 - fz^{-1})(1 - fz)}$$

Integrating $S_{ee}(z)$ over the unit circle we find the variance of $e_{n+1/n}$:

$$P = E[e_{n+1/n}^2] = \int_{\text{u.c.}} S_{ee}(z)\frac{dz}{2\pi jz} = \int_{\text{u.c.}} \frac{Q + K^2R}{(z - f)(1 - fz)}\frac{dz}{2\pi j} = \operatorname*{Res}_{z=f} = \frac{Q + K^2R}{1 - f^2}$$

The derivative of P with respect to the Kalman gain K is

$$\frac{dP}{dK} = \frac{2KR - 2fcP}{1 - f^2}$$

Setting this to zero, we find

$$KR = fcP = (a - cK)cP \quad \Longrightarrow \quad K = \frac{acP}{R + c^2P}$$

We can use this result to express f in terms of P:

$$f = a - cK = a - c\,\frac{acP}{R + c^2P} = \frac{aR}{R + c^2P}$$

Inserting the above expressions for f and K into $P = \dfrac{Q + K^2R}{1 - f^2}$, we find

$$P = \frac{Q + \left(\dfrac{acp}{R + c^2P}\right)^2 R}{1 - \left(\dfrac{aR}{R + c^2P}\right)^2} \quad \Longrightarrow \quad P - Q = \frac{PRa^2}{R + c^2P}$$

Problem 4.9:
In the previous problem we derived the difference equation

$$e_{n+1/n} = f\,e_{n/n-1} + u_n\,, \quad \text{where} \quad u_n = w_n - Kv_n$$

Because u_n is white, it follows from the causality of this difference equation that $E\,[e_{n/n-1}\,u_n] = 0$. Using this and stationarity we find

$$P = E\,[e^2_{n+1/n}] = f^2E\,[e^2_{n/n-1}] + E\,[u^2_n] = f^2P + (Q + K^2R)$$

which may be solved for $P = \dfrac{Q + K^2R}{1 - f^2}$. Next, we relate the variance of $e_{n/n}$ to the variance of $e_{n+1/n}$. We have

$$e_{n+1/n} = x_{n+1} - \hat{x}_{n+1/n} = (ax_n + w_n) - a\hat{x}_{n/n} = ae_{n/n} + w_n$$

w_n is uncorrelated with $e_{n/n}$ because it is uncorrelated with x_n, y_n, and all the past xs and ys. It follows that

$$E\,[e^2_{n+1/n}] = a^2E\,[e^2_{n/n}] + E\,[w^2_n] \quad \Longrightarrow \quad P = a^2E\,[e^2_{n/n}] + Q$$

Solving for $E\,[e^2_{n/n}]$ and using the Riccati equation, we find

$$E\,[e^2_{n/n}] = \frac{P - Q}{a^2} = \frac{PR}{R + c^2P}$$

Problem 4.10:
Using the measurement equation $y_n = cx_n + v_n$, we find

$$e_n = y_n - c\hat{x}_{n/n-1} = (cx_n + v_n) - c\hat{x}_{n/n-1} = ce_{n/n-1} + v_n$$

Now, v_n is uncorrelated with x_n and all past xs. Therefore, it is uncorrelated with y_n and all past ys. It follows that it is uncorrelated with $e_{n/n-1}$. Thus,

$$\sigma_\epsilon^2 = E[\epsilon_n^2] = c^2 E[e_{n/n-1}^2] + E[v_n^2] = c^2 P + R$$

Problem 4.13:
First, note that the nth diagonal entry of this performance index incorporates the constraints required for the nth estimate:

$$J_n = E[e_n^2] + (\Lambda H^T)_{nn} + (H\Lambda^T)_{nn} = E[e_n^2] + 2\sum_{i>n} \Lambda_{ni} H_{ni}$$

Using $R_{ee} = R_{xx} - HR_{yx} - R_{xy}H^T + HR_{yy}H^T$, derived in Section 1.4, we find

$$J = R_{ee} = R_{xx} - HR_{yx} - R_{xy}H^T + HR_{yy}H^T + \Lambda H^T + H\Lambda^T$$

The first order variation with respect to H is

$$\delta J = (HR_{yy} - R_{xy} + \Lambda)\,\delta H^T + \delta H\,(HR_{yy} - R_{xy} + \Lambda)^T$$

The unconstrained minimization of J requires $\delta J = 0$ for all δH. Thus, we obtain the conditions

$$HR_{yy} - R_{xy} + \Lambda \quad \Longrightarrow \quad R_{xy} - HR_{yy} = \Lambda = \text{ strictly upper triangular}$$

which is the same as Eq. (4.8.3) subject to Eq. (4.8.2).

Chapter 5

Problem 5.1:
Let $x_n = y_{n+D}$, so that $X(z) = z^D Y(z)$. Then, the cross power spectral density will be

$$S_{xy}(z) = z^D S_{yy}(z) = \sigma_\epsilon^2 z^D B(z) B(z^{-1}) \quad \text{where} \quad S_{yy}(z) = \sigma_\epsilon^2 B(z) B(z^{-1})$$

The Wiener filter for estimating x_n from y_n is then,

$$H(z) = \frac{1}{\sigma_\epsilon^2 B(z)} \left[\frac{S_{xy}(z)}{B(z^{-1})} \right]_+ = \frac{1}{B(z)} [z^D B(z)]_+$$

The causal instruction is removed as follows:

$$\begin{aligned}
[z^D B(z)]_+ &= [z^D + b_1 z^{D-1} + b_2 z^{D-2} + \cdots + b_D + b_{D+1} z^{-1} + \cdots]_+ = \\
&= b_D + b_{D+1} z^{-1} + \cdots = z^D \left(b_D z^{-D} + b_{D+1} z^{-D-1} + \cdots \right) \\
&= z^D \left(B(z) - \sum_{m=0}^{D-1} b_m z^{-m} \right)
\end{aligned}$$

Therefore, the prediction filter becomes

$$H(z) = z^D \left(1 - \frac{1}{B(z)} \sum_{m=0}^{D-1} b_m z^{-m} \right)$$

Note that its input is y_n and its output is the estimate $\hat{y}_{n+D/n}$ of y_{n+D}.
The first few terms in the power series expansion of Example 5.1.1 are

$$B(z) = \frac{1}{1 - 0.9z^{-1} + 0.2z^{-2}} = 1 + 0.9z^{-1} + 0.61z^{-2} + \cdots$$

Therefore, the $D = 2$ predictor will be:

$$H(z) = z^2 \left(1 - (1 - 0.9z^{-1} + 0.2z^{-2})(1 + 0.9z^{-1}) \right) = 0.61 - 0.18z^{-1}$$

with an I/O equation $\hat{y}_{n+2/n} = 0.61 y_n - 0.18 y_{n-1}$. For $D = 3$, we find

$$H(z) = z^2 \left(1 - (1 - 0.9z^{-1} + 0.2z^{-2})(1 + 0.9z^{-1} + 0.61z^{-2}) \right) = -0.729 - 0.122z^{-1}$$

with I/O equation $\hat{y}_{n+3/n} = -0.729 y_n - 0.122 y_{n-1}$. For Example 5.1.2, we have

$$B(z) = \frac{1 - 0.25z^{-2}}{1 - 0.8z^{1}} = 1 + 0.8z^{-1} + 0.39z^{-2} + \cdots$$

The 2-step predictor becomes

$$H(z) = z^2\left(1 - \frac{(1 - 0.8z^{-1})(1 + 0.8z^{-1})}{1 - 0.25z^{-2}}\right) = \frac{0.39}{1 - 0.25z^{-2}}$$

with difference equation $\hat{y}_{n+2/n} = 0.25y_{n/n-2} + 0.39y_n$. The 3-step predictor is

$$H(z) = z^3\left(1 - \frac{(1 - 0.8z^{-1})(1 + 0.8z^{-1} + 0.39z^{-2})}{1 - 0.25z^{-2}}\right) = \frac{0.312}{1 - 0.25z^{-2}}$$

with difference equation $\hat{y}_{n+3/n} = 0.25y_{n+1/n-2} + 0.312y_n$.

Problem 5.2:
The prediction-error filter $A(z) = 1/B(z)$ determines the projection onto the infinite past. Working with z-transforms we have

$$\hat{Y}(z) = Y(z) - \epsilon(z) = [1 - A(z)]Y(z) = -[a_1z^{-1} + a_2z^{-2} + \cdots + a_pz^{-p}]Y(z)$$

or, in the time-domain

$$\hat{y}_n = -[a_1y_{n-1} + a_2y_{n-2} + \cdots + a_py_{n-p}]$$

The prediction coefficients must satisfy Eqs. (5.2.5) which are identical to Eqs. (5.3.7) that determine the projection onto the past p samples.

Problem 5.3:
Using $e_n = \sum_{i=0}^{p} a_iy_{n-i}$, we find

$$\sigma_e^2 = E[e_n^2] = \sum_{i,j=0}^{p} a_iE[y_{n-i}y_{n-j}]a_j = \sum_{i,j=0}^{p} a_iR(j-i)a_j = \mathbf{a}^TR\mathbf{a}$$

The constraint $a_0 = 1$ can be written in vector form $\mathbf{a}^T\mathbf{u} = 1$, where $\mathbf{u}$ is the unit vector $\mathbf{u} = [1, 0, \cdots, 0]^T$. The extended performance index incorporating this constraint with a Lagrange multiplier will be

$$J = \mathbf{a}^TR\mathbf{a} + 2\lambda(1 - \mathbf{a}^T\mathbf{u})$$

Its first order variation with respect to $\mathbf{a}$ is

$$\delta J = 2\delta\mathbf{a}^TR\mathbf{a} - 2\lambda\delta\mathbf{a}^T\mathbf{u} = 2\delta\mathbf{a}^T(R\mathbf{a} - \lambda\mathbf{u})$$

The minimization condition $\delta J = 0$ requires

$$R\mathbf{a} = \lambda\mathbf{u}$$

The Lagrange multiplier is fixed by imposing the constraint. Multiplying by $\mathbf{a}^T$ we find

$$\sigma_e^2 = \mathbf{a}^T R\mathbf{a} = \lambda\mathbf{a}^T\mathbf{u} = \lambda$$

Thus, we obtain Eq. (5.3.7), $R\mathbf{a} = \sigma_e^2\mathbf{u}$.

Problem 5.4:
Noting the $A^R(z)$ are the reverse polynomials, we find for the inverse z-transform of Eq. (5.3.17):

$$\begin{bmatrix} a_{p+1,p+1} \\ a_{p+1,p} \\ \vdots \\ a_{p+1,1} \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ a_{pp} \\ \vdots \\ a_{p1} \\ 1 \end{bmatrix} - \gamma_{p+1} \begin{bmatrix} 1 \\ a_{p1} \\ \vdots \\ a_{pp} \\ 0 \end{bmatrix}$$

It is recognized as the upside down (i.e., reversed) version of Eq. (5.3.15).

Problem 5.5:
For arbitrary γ, we have the matrix identity

$$\frac{1}{1-\gamma^2}\begin{bmatrix} 1 & \gamma \\ \gamma z & z \end{bmatrix}\begin{bmatrix} 1 & -\gamma z^{-1} \\ -\gamma & z^{-1} \end{bmatrix} = \frac{1}{1-\gamma^2}\begin{bmatrix} 1-\gamma^2 & -\gamma z^{-1} + \gamma z^{-1} \\ \gamma z - \gamma z & 1-\gamma^2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

Problem 5.6:
Sending the reflection coefficients through the routine **frwlev**, gives all the prediction-error filters up to order four, with their coefficients arranged in reverse order into the matrix L:

$$L = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ -0.5 & 1 & 0 & 0 & 0 \\ 0.5 & -0.75 & 1 & 0 & 0 \\ -0.5 & 0.875 & -1 & 1 & 0 \\ 0.5 & -1 & 1.3125 & -1.25 & 1 \end{bmatrix}$$

The 4th order polynomial extracted from the last row of L is

$$A_4(z) = 1 - 1.25z^{-1} + 1.3125z^{-2} - z^{-3} + 0.5z^{-4}$$

Sending the coefficients of $A_4(z)$ and $E_4 = 40.5$ through the routine **rlev**, gives the autocorrelation lags:

$$\{R(0), R(1), R(2), R(3), R(4)\} = \{128, 64, -16, -8, 11\}$$

Problem 5.7:
Initialize by $R(0)\cdot 1 = E_0$, which gives $E_0 = 256$. Enlarge the 0th order normal equations by padding a zero:

$$\begin{bmatrix} 256 & 128 \\ 128 & 256 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 256 \\ \Delta_0 \end{bmatrix} \implies \Delta_0 = 128$$

Thus,

$$\gamma_1 = \frac{\Delta_0}{E_0} = \frac{128}{256} = 0.5\,, \quad \text{and} \quad E_1 = (1 - \gamma_1^2)E_0 = 192$$

The 1st order prediction filter will be

$$\begin{bmatrix} 1 \\ a_{11} \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} - 0.5 \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ -0.5 \end{bmatrix}$$

Next, enlarge to next size by padding a zero:

$$\begin{bmatrix} 256 & 128 & -32 \\ 128 & 256 & 128 \\ -32 & 128 & 256 \end{bmatrix} \begin{bmatrix} 1 \\ -0.5 \\ 0 \end{bmatrix} = \begin{bmatrix} 192 \\ 0 \\ \Delta_1 \end{bmatrix} \implies \Delta_1 = -96$$

Therefore,

$$\gamma_2 = \frac{\Delta_1}{E_1} = \frac{-96}{192} = -0.5\,, \quad \text{and} \quad E_2 = (1 - \gamma_2^2)E_1 = 144$$

and the 2nd order prediction filter will be

$$\begin{bmatrix} 1 \\ a_{21} \\ a_{22} \end{bmatrix} = \begin{bmatrix} 1 \\ -0.5 \\ 0 \end{bmatrix} - (-0.5) \begin{bmatrix} 0 \\ -0.5 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ -0.75 \\ 0.5 \end{bmatrix}$$

Enlarging to the next size, we get

$$\begin{bmatrix} 256 & 128 & -32 & -16 \\ 128 & 256 & 128 & -32 \\ -32 & 128 & 256 & 128 \\ -16 & -32 & 128 & 256 \end{bmatrix} \begin{bmatrix} 1 \\ -0.75 \\ 0.5 \\ 0 \end{bmatrix} = \begin{bmatrix} 144 \\ 0 \\ 0 \\ \Delta_2 \end{bmatrix} \implies \Delta_2 = 72$$

Thus,

$$\gamma_3 = \frac{\Delta_2}{E_2} = \frac{72}{144} = 0.5\,, \quad \text{and} \quad E_3 = (1-\gamma_3^2)E_2 = 108$$

and the 3d order prediction filter will be

$$\begin{bmatrix} 1 \\ a_{31} \\ a_{32} \\ a_{33} \end{bmatrix} = \begin{bmatrix} 1 \\ -0.75 \\ 0.5 \\ 0 \end{bmatrix} - 0.5 \begin{bmatrix} 0 \\ 0.5 \\ -0.75 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \\ 0.875 \\ -0.5 \end{bmatrix}$$

Enlarging to order 4 by padding a zero, we get

$$\begin{bmatrix} 256 & 128 & -32 & -16 & 22 \\ 128 & 256 & 128 & -32 & -16 \\ -32 & 128 & 256 & 128 & -32 \\ -16 & -32 & 128 & 256 & 128 \\ 22 & -16 & -32 & 128 & 256 \end{bmatrix} \begin{bmatrix} 1 \\ -1 \\ 0.875 \\ -0.5 \\ 0 \end{bmatrix} = \begin{bmatrix} 108 \\ 0 \\ 0 \\ 0 \\ \Delta_3 \end{bmatrix} \quad \Longrightarrow \quad \Delta_3 = -54$$

Thus,

$$\gamma_4 = \frac{\Delta_3}{E_3} = \frac{-54}{108} = -0.5\,, \quad \text{and} \quad E_4 = (1-\gamma_4^2)E_3 = 81$$

The 4th order filter will be

$$\begin{bmatrix} 1 \\ a_{41} \\ a_{42} \\ a_{43} \\ a_{44} \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \\ 0.875 \\ -0.5 \\ 0 \end{bmatrix} - (-0.5) \begin{bmatrix} 0 \\ -0.5 \\ 0.875 \\ -1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ -1.25 \\ 1.3125 \\ -1 \\ 0.5 \end{bmatrix}$$

Problem 5.8:
Sending the coefficients of $A_4(z)$ and $E_4 = 0.81$ through the routine **rlev**, gives the same matrix L as that of Problem 5.6, and the following autocorrelation lags:

$$\{R(0), R(1), R(2), R(3), R(4)\} = \{2.56,\ 1.28,\ -0.32,\ -0.16,\ 0.22\}$$

Problem 5.10:
The sample autocorrelation of the first sequence is computed by

$$R(k) = \sum_{n=0}^{4-k} (-1)^{n+k}(-1)^n = (-1)^k(5-k)\,, \quad \text{for} \quad k = 0, 1, 2, 3, 4$$

or,

$$\{R(0), R(1), R(2), R(3), R(4)\} = \{5, -4, 3, -2, 1\}$$

Sending these through the routine **lev**, gives all the prediction filters up to order four, arranged into the matrix L:

$$L = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0.8 & 1 & 0 & 0 & 0 \\ -0.1111 & 0.8889 & 1 & 0 & 0 \\ -0.125 & 0 & 0.875 & 1 & 0 \\ 0.1429 & 0 & 0 & 0.8571 & 1 \end{bmatrix}$$

and the vector of prediction errors

$$\{E_0, E_1, E_2, E_3, E_4\} = \{5, 1.8, 1.778, 1.75, 1.7143\}$$

The I/O equation of the 4th order predictor will be

$$\hat{y}_n = -0.8571 y_{n-1} - 0.149 y_{n-4} \quad \Longrightarrow \quad \hat{y}_4 = -0.8571 + 0.1429 = -0.7142$$

The second sequence has autocorrelation lags

$$\{R(0), R(1), R(2), R(3), R(4)\} = \{55, 40, 26, 14, 5\}$$

Sending these through **lev**, gives

$$L = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ -0.7273 & 1 & 0 & 0 & 0 \\ 0.1193 & -0.814 & 1 & 0 & 0 \\ 0.0937 & 0.043 & -0.8029 & 1 & 0 \\ 0.0543 & 0.0501 & 0.0454 & -0.7978 & 1 \end{bmatrix}$$

with 4th order prediction

$$\hat{y}_n = 0.7978 y_{n-1} - 0.0454 y_{n-2} - 0.0501 y_{n-3} - 0.0543 y_{n-4}$$

Problem 5.11:
For the first sequence, we have $\gamma_1 = R(1)/R(0) = 0.5$ and $E_1 = (1 - \gamma_1^2)R(0) = 0.75$. The maximum entropy extension keeps all the prediction error filters equal to the first order filter, that is, $A_1(z) = 1 - \gamma_1 z^{-1} = 1 - 0.5z^{-1}$. The maximum entropy spectral density will be, then

$$S_{yy}(z) = \frac{E_1}{A_1(z)A_1(z^{-1})} = \frac{0.75}{(1 - 0.5z^{-1})(1 - 0.5z)}$$

The inverse z-transform of this leads to $R(k) = (0.5)^{|k|}$. For the second sequence, we determine the first two reflection coefficients $\{\gamma_1, \gamma_2\} = \{0, 0.25\}$, and the corresponding 2nd order prediction filter

$A_2(z) = 1 - 0.25z^{-2}$ and prediction error $E_2 = 3.75$. Therefore, the maximum entropy spectrum will be

$$S_w(z) = \frac{E_2}{A_2(z)A_2(z^{-1})} = \frac{3.75}{(1 - 0.25z^{-2})(1 - 0.25z^2)}$$

Taking the inverse z-transform of this spectral density leads to

$$R(k) = 2(0.5)^{|k|} + 2(-0.5)^{|k|}$$

Problem 5.12:
We have from (5.3.24):

$$R(p+1) = -\sum_{i=1}^{p+1} a_{p+1,i} R(p+1-i) = -\sum_{i=1}^{p} a_{p+1,i} R(p+1-i) - a_{p+1,p+1} R(0)$$

From the Levinson recursion, we have $a_{p+1,p+1} = -\gamma_{p+1}$ and $a_{p+1,i} = a_{pi} - \gamma_{p+1} a_{p,p+1-i}$, for $i = 1, \cdots, p$. Therefore,

$$R(p+1) = -\sum_{i=1}^{p} (a_{pi} - \gamma_{p+1} a_{p,p+1-i}) R(p+1-i) + \gamma_{p+1} R(0) = \gamma_{p+1} [R(0) + \sum_{i=1}^{p} a_{p,p+1-i} R(p+1-i)] - \sum_{i=1}^{p} a_{pi} R(p+1-i)$$

But, we recognize

$$R(0) + \sum_{i=1}^{p} a_{p,p+1-i} R(p+1-i) = \sum_{i=0}^{p} a_{pi} R(i) = E_p$$

Problem 5.13:
Initialize the split Levinson recursion by

$$\mathbf{f}_0 = [2], \quad \mathbf{f}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad \tau_0 = R(0) = 256, \quad \gamma_0 = 0$$

Next, compute $\tau_1 = [R(0), R(1)]\mathbf{f}_1 = [256, 128]\begin{bmatrix} 1 \\ 1 \end{bmatrix} = 384$, and $\alpha_1 = \tau_1/\tau_0 = 384/256 = 1.5$, and then, $\gamma_1 = -1 + \alpha_1/(1 - \gamma_0) = -1 + 1.5 = 0.5$. Then, update the symmetric polynomial

$$\mathbf{f}_2 = \begin{bmatrix} \mathbf{f}_1 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ \mathbf{f}_1 \end{bmatrix} - \alpha_1 \begin{bmatrix} 0 \\ \mathbf{f}_0 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} - 1.5 \begin{bmatrix} 0 \\ 2 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}$$

Next, compute

$$\tau_2 = [R(0), R(1), R(2)]\mathbf{f}_2 = [256,\ 128,\ -32]\begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix} = 96 \quad \Longrightarrow \quad \alpha_2 = \frac{\tau_2}{\tau_1} = \frac{96}{384} = 0.25$$

and find $\gamma_2 = -1 + \alpha_2/(1 - \gamma_1) = -0.5$. And then, update to $\mathbf{f}_3$:

$$\mathbf{f}_3 = \begin{bmatrix} \mathbf{f}_2 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ \mathbf{f}_2 \end{bmatrix} - \alpha_2 \begin{bmatrix} 0 \\ \mathbf{f}_1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \\ 1 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \\ -1 \\ 1 \end{bmatrix} - 0.25 \begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ -0.25 \\ -0.25 \\ 1 \end{bmatrix}$$

Next, compute

$$\tau_3 = [R(0), R(1), R(2), R(3)]\mathbf{f}_3 = [256,\ 128,\ -32,\ -16]\begin{bmatrix} 1 \\ -0.25 \\ -0.25 \\ 1 \end{bmatrix} = 216 \quad \Longrightarrow \quad \alpha_3 = \frac{\tau_3}{\tau_2} = \frac{216}{96} = 2.25$$

and find $\gamma_3 = -1 + \alpha_3/(1 - \gamma_2) = 0.5$. And then, construct $\mathbf{f}_4$:

$$\mathbf{f}_4 = \begin{bmatrix} \mathbf{f}_3 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ \mathbf{f}_3 \end{bmatrix} - \alpha_3 \begin{bmatrix} 0 \\ \mathbf{f}_2 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ -0.25 \\ -0.25 \\ 1 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \\ -0.25 \\ -0.25 \\ 1 \end{bmatrix} - 2.25 \begin{bmatrix} 0 \\ 1 \\ -1 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ -1.5 \\ 1.75 \\ -1.5 \\ 1 \end{bmatrix}$$

Next, compute

$$\tau_4 = [R(0), R(1), R(2), R(3), R(4)]\mathbf{f}_4 = [256,\ 128,\ -32,\ -16,\ 22]\begin{bmatrix} 1 \\ -1.5 \\ 1.75 \\ -1.5 \\ 1 \end{bmatrix} = 54 \quad \Longrightarrow \quad \alpha_4 = \frac{\tau_4}{\tau_3} = \frac{54}{216} = 0.25$$

and find $\gamma_4 = -1 + \alpha_4/(1 - \gamma_3) = -0.5$.

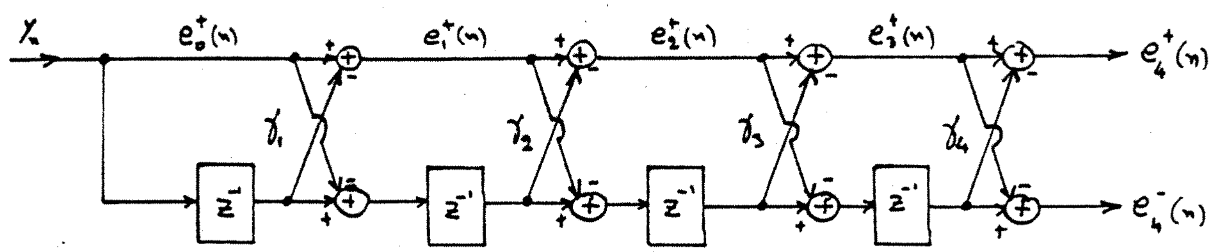
Problem 5.14:
The prediction-error filters of Problems 5.6-5.8 are all the same. The fourth order polynomial is

$$A_4(z) = 1 - 1.25z^{-1} + 1.3125z^{-2} - z^{-3} + 0.5z^{-4}$$

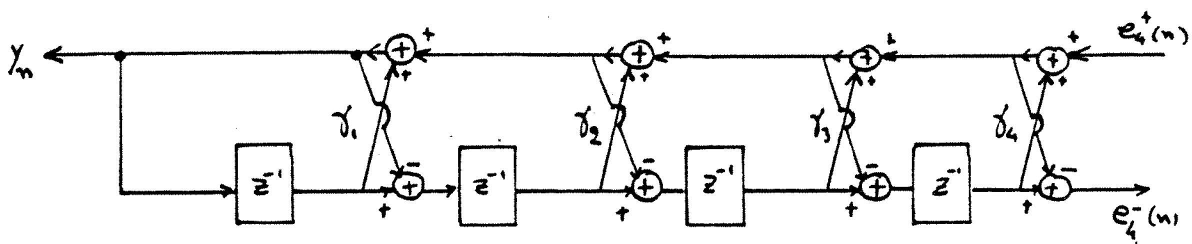
The corresponding reflection coefficients are

$$\{\gamma_1, \gamma_2, \gamma_3, \gamma_4\} = \{0.5,\ -0.5,\ 0.5,\ -0.5\}$$

The analysis lattice filter is



The synthesis filter is



Problem 5.15:
Sending the coefficients of the first polynomial through the routine **bkwlev**, gives the reflection coefficients:

$$\{\gamma_1, \gamma_2, \gamma_3, \gamma_4\} = \{2, 0.3, 0.4, 0.5\}$$

Because one of the reflection coefficients has magnitude greater than one, the polynomial will not be minimal phase. The reflection coefficients of the second polynomial are:

$$\{\gamma_1, \gamma_2, \gamma_3, \gamma_4\} = \{0.2, 0.3, 0.4, 0.5\}$$

Since all of them have magnitude less than one, the polynomial will be minimal phase.

Problem 5.16:
For a gaussian density, we have

$$p(\mathrm{y}) = \frac{\exp\left(-\frac{1}{2}\mathrm{y}^T R^{-1}\mathrm{y}\right)}{(2\pi)^{M/2}(\det R)^{1/2}} \quad \Longrightarrow \quad -\ln p(\mathrm{y}) = \frac{1}{2}\ln(\det R) + \ln\left((2\pi)^{M/2}\right) + \frac{1}{2}\mathrm{y}R^{-1}\mathrm{y}$$

The entropy S is the expectation value $S = E[-\ln p]$. Noting that $\mathrm{y}^T R^{-1}\mathrm{y} = \mathrm{tr}(\mathrm{y}\mathrm{y}^T R^{-1})$, we have $E[\mathrm{y}^T R^{-1}\mathrm{y}] = \mathrm{tr}(E[\mathrm{y}\mathrm{y}^T]R^{-1}) = \mathrm{tr}(RR^{-1}) = \mathrm{tr}(I) = M$. It follows that

$$S = E[-\ln p] = \frac{1}{2}\ln(\det R) + c$$

where $c = \ln((2\pi)^{M/2}) + M/2$. Using the LU decomposition of R, that is, $LRL^T = D$, we have $(\det L)^2 \det R = \det D = \prod_{i=0}^{M} E_i$. But, $\det L = 1$ because L is unit lower triangular. Therefore,

$$S_M = \frac{1}{2}\ln(\det R_M) = \frac{1}{2}\ln(\det D_M) = \frac{1}{2}\ln\left(\prod_{i=0}^{M} E_i\right) = \frac{1}{2}\sum_{i=0}^{M}\ln E_i$$

It follows that the entropies of the order M and order p cases will differ by

$$S_M - S_p = \frac{1}{2}\ln\left(\frac{\det R_M}{\det R_p}\right) = \frac{1}{2}\ln\left(\frac{\det D_M}{\det D_p}\right) = \frac{1}{2}\sum_{i=p+1}^{M}\ln E_i$$

The errors E_i, $i = p+1, \cdots, M$ are given in terms of the reflection coefficients $\gamma_{p+1}, \cdots, \gamma_M$, by

$$E_i = (1-\gamma_i^2)\cdots(1-\gamma_{p+1}^2)E_p \le E_p\,, \quad i = p+1, \cdots, M$$

with the maximum value attained when all the new γs are zero, that is, $\gamma_i^2 = 0$, for $i = p+1, \cdots, M$. It is evident, then, that S_M (with S_p fixed) will be maximized.

Problem 5.17:

a. By definition, we have $L^T = [\mathbf{b}_0, \mathbf{b}_1, \cdots, \mathbf{b}_M]$. It follows that

$$\sum_{p=0}^{M} E_p^{-1}\mathbf{b}_p\mathbf{b}_p^T = [\mathbf{b}_0, \mathbf{b}_1, \cdots, \mathbf{b}_M]\begin{bmatrix} E_0^{-1}\mathbf{b}_0^T \\ E_1^{-1}\mathbf{b}_1^T \\ \vdots \\ E_M^{-1}\mathbf{b}_M^T \end{bmatrix} = L^T D^{-1} L = R^{-1}$$

b.

$$A(z) = a_0 + a_1 z^{-1} + \cdots + a_M z^{-M} = [1, z^{-1}, \cdots, z^{-M}] \begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ a_M \end{bmatrix} = \mathbf{s}(z)^T \mathbf{a}$$

The inverse z-transform can be written vectorially as

$$a_m = \oint_{u.c.} A(z) z^m \frac{dz}{2\pi j z} \quad \Longrightarrow \quad \mathbf{a} = \oint_{u.c.} A(z)\mathbf{s}(z^{-1}) \frac{dz}{2\pi j z}$$

where, $\mathbf{s}(z^{-1}) = [1, z, \cdots, z^M]^T$.

c. Using the results of part (b), we have

$$K(z,w) = \mathbf{s}(z)^T \mathbf{k}(w) = \mathbf{s}(z)^T R^{-1} \mathbf{s}(w) = \mathbf{s}(z)^T R^{-1} R R^{-1} \mathbf{s}(w) = \mathbf{k}(z)^T R \mathbf{k}(w)$$

d. First, note

$$J\mathbf{s}(z) = [z^{-M}, \cdots, z^{-1}, 1]^T = z^{-M}[1, z, \cdots, z^M]^T = z^{-M}\mathbf{s}(z^{-1})$$

Using the fact that R and R^{-1} are invariant under J, that is, $JR^{-1}J = R^{-1}$, we obtain

$$K(z,w) = \mathbf{s}(z)^T J R^{-1} J \mathbf{s}(w) = z^{-M} w^{-M} \mathbf{s}(z^{-1})^T R^{-1} \mathbf{s}(w^{-1}) = z^{-M} w^{-M} K(z^{-1}, w^{-1})$$

e. From parts (a-c), it follows

$$K(z,w) = \mathbf{s}(z)^T R^{-1} \mathbf{s}(w) = \sum_{p=0}^{M} \mathbf{s}(z)^T \mathbf{b}_p \frac{1}{E_p} \mathbf{b}_p^T \mathbf{s}(w) = \sum_{p=0}^{M} \frac{1}{E_p} B_p(z) B_p(w)$$

where, $B_p(z) = \mathbf{s}(z)^T \mathbf{b}_p$. Using part (d) and the fact that polynomial $B_p(z)$ is the reverse of $A_p(z)$, we find

$$K(z,w) = z^{-M} w^{-M} K(z^{-1}, w^{-1}) = z^{-M} w^{-M} \sum_{p=0}^{M} \frac{1}{E_p} B_p(z^{-1}) B_p(w^{-1}) = z^{-M} w^{-M} \sum_{p=0}^{M} z^p A_p(z) w^p A_p(w)$$

or,

$$K(z,w) = \sum_{p=0}^{M} \frac{1}{E_p} A_p(z) A_p(w) z^{-(M-p)} w^{-(M-p)}$$

Problem 5.18:
By definition, we have $R_{ij} = R(i-j)$. Thus,

$$R(i-j) = \int_{u.c.} S_{yy}(z) z^{i-j} \frac{dz}{2\pi jz} \quad \Longrightarrow \quad R = \int_{u.c.} S_{yy}(z) \mathbf{s}(z^{-1}) \mathbf{s}(z)^T \frac{dz}{2\pi jz}$$

Using the above result and the definition $\mathbf{k}(z) = R^{-1}\mathbf{s}(z)$, we find

$$R^{-1} = R^{-1}RR^{-1} = \int_{u.c.} S_{yy}(z) R^{-1}\mathbf{s}(z^{-1})\mathbf{s}(z)^T R^{-1} \frac{dz}{2\pi jz} = \int_{u.c.} S_{yy}(z)\mathbf{k}(z^{-1})\mathbf{k}(z)^T \frac{dz}{2\pi jz}$$

Using the above result and the definition of $K(z,w)$, we find

$$K(z,w) = \mathbf{s}(z)^T R^{-1}\mathbf{s}(w) = \int_{u.c.} S_{yy}(u)\mathbf{s}(z)^T\mathbf{k}(u^{-1})\mathbf{k}(u)^T\mathbf{s}(w)\frac{du}{2\pi ju} = \int_{u.c} S_{yy}(u) K(z,u^{-1}) K(u,w) \frac{du}{2\pi ju}$$

Problem 5.19:

(a) First, note that $\mathbf{s}_p(z)$ admits the decompositions

$$\mathbf{s}_p(z) = \begin{bmatrix} 1 \\ z^{-1} \\ \vdots \\ z^{-(p-1)} \\ z^{-p} \end{bmatrix} = \begin{bmatrix} \mathbf{s}_{p-1}(z) \\ z^{-p} \end{bmatrix} = \begin{bmatrix} 1 \\ z^{-1}\mathbf{s}_{p-1}(z) \end{bmatrix}$$

Using the order updating property (5.9.11) (see also (1.7.28)), we have

$$R_p^{-1} = \begin{bmatrix} R_{p-1}^{-1} & \mathbf{0} \\ \mathbf{0}^T & 0 \end{bmatrix} + \frac{1}{E_p}\mathbf{b}_p\mathbf{b}_p^T$$

It follows that $\mathbf{k}_p(w) = R_p^{-1}\mathbf{s}_p(w)$ can be written as

$$\mathbf{k}^p(w) = \begin{bmatrix} R_{p-1}^{-1} & \mathbf{0} \\ \mathbf{0}^T & 0 \end{bmatrix} \begin{bmatrix} \mathbf{s}_{p-1}(w) \\ w^{-p} \end{bmatrix} + \frac{1}{E_p}\mathbf{b}_p\mathbf{b}_p^T\mathbf{s}_p(w) = \begin{bmatrix} \mathbf{k}_{p-1}(w) \\ 0 \end{bmatrix} + \frac{1}{E_p}\mathbf{b}_p B_p(w)$$

The second decomposition follows from the order updating formula (5.9.16), or (1.7.35), that is,

$$R_p^{-1} = \begin{bmatrix} 0 & \mathbf{0}^T \\ \mathbf{0} & R_{p-1}^{-1} \end{bmatrix} + \frac{1}{E_p}\mathbf{a}_p\mathbf{a}_p^T$$

Then, we have

$$\mathbf{k}_p(w) = \begin{bmatrix} 0 & \mathbf{0}^T \\ \mathbf{0} & R_{p-1}^{-1} \end{bmatrix} \begin{bmatrix} 1 \\ w^{-1}\mathbf{s}_{p-1}(w) \end{bmatrix} + \frac{1}{E_p}\mathbf{a}_p\mathbf{a}_p^T\mathbf{s}_p(z) = \begin{bmatrix} 0 \\ w^{-1}\mathbf{k}_{p-1}(w) \end{bmatrix} + \frac{1}{E_p}\mathbf{a}_p A_p(w)$$

The above equation can also be obtained by reversing the first one. To see this, note first that the reverse of $\mathbf{k}_p(w)$ is $J_p\mathbf{k}_p(w) = R_p^{-1}J_p\mathbf{s}_p(w) = R_p^{-1}w^{-p}\mathbf{s}_p(w^{-1}) = w^{-p}\mathbf{k}_p(w^{-1})$, where we used $J_pR_p^{-1} = R_p^{-1}J_p$. Then, using the fact that $\mathbf{a}_p = J_p\mathbf{b}_p$, and $A_p(w) = w^{-p}B_p(w^{-1})$, we get $\mathbf{k}_p(w) = J_pJ_p\mathbf{k}_p(w) = w^{-p}J_p\mathbf{k}_p(w^{-1})$, or,

$$\mathbf{k}_p(w) = w^{-p}J_p\begin{bmatrix}\mathbf{k}_{p-1}(w^{-1})\\ 0\end{bmatrix} + \frac{1}{E_p}w^{-p}J_p\mathbf{b}_pB_p(w^{-1}) = \begin{bmatrix}0\\ w^{-p}J_{p-1}\mathbf{k}_{p-1}(w^{-1})\end{bmatrix} + \frac{1}{E_p}\mathbf{a}_pA_p(w)$$

and use the property $w^{-p}J_{p-1}\mathbf{k}_{p-1}(w^{-1}) = w^{-1}J_{p-1}J_{p-1}\mathbf{k}_{p-1}(w) = w^{-1}\mathbf{k}_{p-1}(w)$.

(b) These recursions can be proved using the results of Problem 5.17(c), or they can be proved directly as follows

$$K_p(z,w) = \mathbf{s}_p(z)^T\mathbf{k}_p(w) = [\mathbf{s}_{p-1}(z)^T,\, 1]\begin{bmatrix}\mathbf{k}_{p-1}(w)\\ 0\end{bmatrix} + \frac{1}{E_p}\mathbf{s}_p(z)^T\mathbf{b}_pB_p(w)$$

or,

$$K_p(z,w) = K_{p-1}(z,w) + \frac{1}{E_p}B_p(z)B_p(w)$$

Similarly,

$$K_p(z) = \mathbf{s}_p(z)^T\mathbf{k}_p(w) = [1,\, z^{-1}\mathbf{s}_{p-1}(z)^T]\begin{bmatrix}0\\ w^{-1}\mathbf{k}_{p-1}(w)\end{bmatrix} + \frac{1}{E_p}\mathbf{s}_p(z)^T\mathbf{a}_pA_p(w)$$

or,

$$K_p(z,w) = z^{-1}w^{-1}K_{p-1}(z,w) + \frac{1}{E_p}A_p(z)A_p(w)$$

(c) Multiplying the first of the above recursions by $z^{-1}w^{-1}$ and subtracting to cancel the $z^{-1}w^{-1}K_{p-1}(z,w)$ term, we obtain the required result. Similarly, if we subtract them to cancel the $K_p(z,w)$ term and solve for $K_{p-1}(z,w)$, we obtain the result of part (e).

(d) Because $A_p(z)$ has real coefficients, we have $A_p(z^*) = A_p(z)^*$. Then, applying part (c) with $w = z^*$, we obtain

$$K_p(z,z^*) = \mathbf{s}_p(z)^TR_p^{-1}\mathbf{s}_p(z)^* = \frac{1}{E_p}\,\frac{|z|^2\,|A_p(z)|^2 - |B_p(z)|^2}{|z|^2 - 1}$$

If $z = z_i$ is a zero of $A_p(z)$, that is, $A_p(z_i) = 0$, then, we have

$$K_p(z_i, z_i^*) = \mathbf{s}(z_i)^TR_p^{-1}\mathbf{s}(z_i)^* = \frac{1}{E_p}\,\frac{|B_p(z_i)|^2}{1 - |z_i|^2}$$

Because R_p^{-1} is positive definite, the left hand side will be positive. Thus, all the factors in the above equation are positive, and we obtain $1 - |z_i|^2 > 0$ or, $|z_i|^2 < 1$. Similar methods can be used to reach the same conclusion by working with the expression in part (e).

Problem 5.20:
Initialize the Schur recursion by

$$\mathbf{g}_0^+ = \mathbf{g}_0^- = \begin{bmatrix} 256 \\ 128 \\ -32 \\ -16 \\ 22 \end{bmatrix}$$

Compute $\gamma_1 = g_0^+(1)/g_0^-(0) = 128/256 = 0.5$. Then, using the Schur recursions, construct the order-1 gapped functions for $1 \le k \le 4$. This can be done conveniently in vector form:

$$\begin{bmatrix} g_1^+(1) \\ g_1^+(2) \\ g_1^+(3) \\ g_1^+(4) \end{bmatrix} = \begin{bmatrix} 128 \\ -32 \\ -16 \\ 22 \end{bmatrix} - 0.5 \begin{bmatrix} 256 \\ 128 \\ -32 \\ -16 \end{bmatrix} = \begin{bmatrix} 0 \\ -96 \\ 0 \\ 30 \end{bmatrix}$$

$$\begin{bmatrix} g_1^-(1) \\ g_1^-(2) \\ g_1^-(3) \\ g_1^-(4) \end{bmatrix} = \begin{bmatrix} 256 \\ 128 \\ -32 \\ -16 \end{bmatrix} - 0.5 \begin{bmatrix} 128 \\ -32 \\ -16 \\ 22 \end{bmatrix} = \begin{bmatrix} 192 \\ 144 \\ -24 \\ -27 \end{bmatrix}$$

Then, $\gamma_2 = g_1^+(2)/g_1^-(1) = -96/192 = -0.5$. The order-2 gapped functions are for $2 \le k \le 4$:

$$\begin{bmatrix} g_2^+(2) \\ g_2^+(3) \\ g_2^+(4) \end{bmatrix} = \begin{bmatrix} -96 \\ 0 \\ 30 \end{bmatrix} - (-0.5) \begin{bmatrix} 192 \\ 144 \\ -24 \end{bmatrix} = \begin{bmatrix} 0 \\ 72 \\ 18 \end{bmatrix}$$

$$\begin{bmatrix} g_2^-(2) \\ g_2^-(3) \\ g_2^-(4) \end{bmatrix} = \begin{bmatrix} 192 \\ 144 \\ -24 \end{bmatrix} - (-0.5) \begin{bmatrix} -96 \\ 0 \\ 30 \end{bmatrix} = \begin{bmatrix} 144 \\ 144 \\ -9 \end{bmatrix}$$

Then, $\gamma_3 = g_2^+(3)/g_2^-(2) = 72/144 = 0.5$. And, the order-3 three gapped functions are for $3 \le k \le 4$:

$$\begin{bmatrix} g_3^+(3) \\ g_3^+(4) \end{bmatrix} = \begin{bmatrix} 72 \\ 18 \end{bmatrix} - 0.5 \begin{bmatrix} 144 \\ 144 \end{bmatrix} = \begin{bmatrix} 0 \\ -54 \end{bmatrix}$$

$$\begin{bmatrix} g_3^-(3) \\ g_3^-(4) \end{bmatrix} = \begin{bmatrix} 144 \\ 144 \end{bmatrix} - 0.5 \begin{bmatrix} 72 \\ 18 \end{bmatrix} = \begin{bmatrix} 108 \\ 135 \end{bmatrix}$$

And we find $\gamma_4 = g_3^+(4)/g_3^-(3) = -54/108 = -0.5$. The final order-4 gapped functions are for $k = 4$:

$$g_4^+(4) = -54 - (-0.5)\,108 = 0\,, \quad g_4^-(4) = 108 - (-0.5)\,{-54} = 81$$

The computed backward gapped function values are the lower triangular part of G; therefore,

$$G = \begin{bmatrix} 256 & 0 & 0 & 0 & 0 \\ 128 & 192 & 0 & 0 & 0 \\ -32 & 144 & 144 & 0 & 0 \\ -16 & -24 & 144 & 108 & 0 \\ 22 & -27 & -9 & 135 & 81 \end{bmatrix}$$

Problem 5.21:
Let $\mathbf{x} = [256, 128, -32, -16, 22]^T$ be the input to the forward and backward prediction filters. For each order $p = 0, 1, 2, 3, 4$, we compute the convolutions $\mathbf{a}_p * \mathbf{x}$ and $\mathbf{a}_p^R * \mathbf{x}$, and from each keep only the first $M+1 = 5$ outputs. These are the columns of the matrices $Y^{\pm}$. We find

$$Y^+ = \begin{bmatrix} 256 & 256 & 256 & 256 & 256 \\ 128 & 0 & -64 & -128 & -192 \\ -32 & -96 & 0 & 64 & 144 \\ -16 & 0 & 72 & 0 & -64 \\ 22 & 30 & 18 & -54 & 0 \end{bmatrix}, \quad Y^- = \begin{bmatrix} 256 & -128 & 128 & -128 & 128 \\ 128 & 192 & -128 & 160 & -192 \\ -32 & 144 & 144 & -128 & 192 \\ -16 & -24 & 144 & 108 & -128 \\ 22 & -27 & -9 & 135 & 81 \end{bmatrix}$$

The lower triangular parts of these matrices agree with the gapped function values computed in the previous problem.

Problem 5.22:
Initialize the split Schur functions by $g_0(0) = R(0) = 256$, $\gamma_0 = 0$, and $g_0(k) = 2R(k)$, $g_1(k) = R(k) + R(k-1)$, for $k = 1, 2, 3, 4$. This gives,

$$\begin{bmatrix} g_0(0) \\ g_0(1) \\ g_0(2) \\ g_0(3) \\ g_0(4) \end{bmatrix} = \begin{bmatrix} 256 \\ 256 \\ -64 \\ -32 \\ 44 \end{bmatrix}, \quad \begin{bmatrix} g_1(1) \\ g_1(2) \\ g_1(3) \\ g_1(4) \end{bmatrix} = \begin{bmatrix} 384 \\ 96 \\ -48 \\ 6 \end{bmatrix}$$

Then, compute $\alpha_1 = g_1(1)/g_0(0) = 384/256 = 1.5$ and $\gamma_1 = -1 + \alpha_1/(1 - \gamma_0) = 0.5$. Then, using Eq. (5.10.20) written in vector form, construct gapped function $g_2(k)$ for $2 \le k \le 4$:

$$\begin{bmatrix} g_2(2) \\ g_2(3) \\ g_2(4) \end{bmatrix} = \begin{bmatrix} 96 \\ -48 \\ 6 \end{bmatrix} + \begin{bmatrix} 384 \\ 96 \\ -48 \end{bmatrix} - 1.5 \begin{bmatrix} 256 \\ -64 \\ -32 \end{bmatrix} = \begin{bmatrix} 96 \\ 144 \\ 6 \end{bmatrix}$$

and find $\alpha_2 = g_2(2)/g_1(1) = 96/384 = 0.25$, and $\gamma_2 = -1 + \alpha_2/(1-\gamma_1) = -0.5$. Next, construct the gapped function $g_3(k)$ for $3 \le k \le 4$:

$$\begin{bmatrix} g_3(3) \\ g_3(4) \end{bmatrix} = \begin{bmatrix} 144 \\ 6 \end{bmatrix} + \begin{bmatrix} 96 \\ 144 \end{bmatrix} - 0.25 \begin{bmatrix} 96 \\ -48 \end{bmatrix} = \begin{bmatrix} 216 \\ 162 \end{bmatrix}$$

and find $\alpha_3 = g_3(3)/g_2(2) = 216/96 = 2.25$, and $\gamma_3 = -1 + \alpha_3/(1-\gamma_2) = 0.5$. Finally, construct the gapped function $g_4(k)$ for $k = 4$:

$$g_4(4) = g_3(4) + g_3(3) - \alpha_3 g_2(3) = 162 + 216 - 2.25 \times 144 = 54$$

Thus, $\alpha_4 = g_4(4)/g_3(3) = 54/216 = 0.25$, and $\gamma_4 = -1 + \alpha_4/(1-\gamma_3) = -0.5$. The order-4 prediction error is computed by $E_4 = g_4(4)(1-\gamma_4) = 54 \times (1 + .5) = 81$.

We illustrate the computation of the Cholesky factors by computing column 1 of G, that is, $g_1^-(k)$ for $k = 1, 2, \cdots, 4$. The first value is $g_1^-(1) = E_1 = g_1(1)(1-\gamma_1) = 384 \times (1 - 0.5) = 192$. Then, Eq. (5.10.21) written in vector form reads

$$\begin{bmatrix} g_1^-(2) \\ g_1^-(3) \\ g_1^-(4) \end{bmatrix} = \begin{bmatrix} 192 \\ g_1^-(2) \\ g_1^-(3) \end{bmatrix} + (1 - 0.5) \begin{bmatrix} 96 \\ -48 \\ 6 \end{bmatrix} - \begin{bmatrix} 96 \\ 144 \\ 6 \end{bmatrix} = \begin{bmatrix} 192 \\ g_1^-(2) \\ g_1^-(3) \end{bmatrix} + \begin{bmatrix} -48 \\ -168 \\ -3 \end{bmatrix}$$

with solution $[g_1^-(2), g_1^-(3), g_1^-(4)] = [144, -24, -27]$. It agrees with Problem 5.20.

Problem 5.23:

(a) The required result follows from the identity

$$|1 - az^{-1}|^2 - |-a^* + z^{-1}|^2 = 1 + |a|^2 |z|^{-2} - 2\mathrm{Re}(az^{-1}) - |a|^2 - |z|^{-2} + 2\mathrm{Re}(az^{-1}) =$$

$$= (1 - |a|^2)(1 - |z|^{-2})$$

If $|a| < 1$, then part (a) implies that the left hand side is $< 1, = 1, > 1$ according as $|z| > 1, = 1, < 1$. Because $A_p(z)$ is a minimum phase polynomial, it can be factored in the form

$$A_p(z) = \prod_{i=1}^{p} (1 - a_i z^{-1}) \quad \Longrightarrow \quad A_p^R(z) = \prod_{i=1}^{p} (-a_i^* + z^{-1})$$

with $|a_i| < 1$. Thus, $|S_p(z)|^2$ is written as a product of factors as in part (a), and each factor has the required properties.

Problem 5.24:
First, find $\gamma_3 = -S_3(\infty) = -0.125$. Then, apply the backward Schur recursion to find

$$S_2(z) = z\,\frac{S_3(z) + \gamma_3}{1 + \gamma_3 S_3(z)} = \frac{0.1111 - 0.8889z^{-1} + z^{-2}}{1 - 0.8889z^{-1} + 0.1111z^{-2}}$$

This gives, $\gamma_2 = -0.1111$, and

$$S_1(z) = z\,\frac{S_2(z) + \gamma_2}{1 + \gamma_2 S_2(z)} = \frac{-0.8 + z^{-1}}{1 - 0.8z^{-1}}$$

Thus, $\gamma_1 = 0.8$ and $S_0(z) = 1$.

Problem 5.25:
The filtering equations for $v_1(n)$ and $v_2(n)$ are in the z-domain

$$V_1(z) = H_1(z)V(z) \quad \text{and} \quad V_2(z) = H_2(z)V(z)$$

where

$$H_1(z) = \frac{1}{1 - a_1 z^{-1}} \quad \text{and} \quad H_2(z) = \frac{1}{1 - a_2 z^{-1}}$$

Since $s(n)$ is uncorrelated with $v_2(n)$, it follows that

$$S_{xy}(z) = S_{v_1v_2}(z) = H_1(z)H_2(z^{-1})S_{vv}(z) = \frac{1}{(1 - a_1 z^{-1})(1 - a_2 z)}$$

where we used $S_{vv}(z) = \sigma_v^2 = 1$. Similarly,

$$S_{yy}(z) = S_{v_2v_2}(z) = H_2(z)H_2(z^{-1})S_{vv}(z) = \frac{1}{(1 - a_2 z^{-1})(1 - a_2 z)}$$

The cross-correlation is computed for $k \geq 0$ by

$$R_{xy}(k) = \int_{u.c.} S_{xy}(z)z^k \frac{dz}{2\pi jz} = \int_{u.c.} \frac{z^k}{(z - a_1)(1 - a_2 z)} \frac{dz}{2\pi j} = \frac{a_1^k}{1 - a_1 a_2}$$

Similarly,

$$R_{yy}(k) = \int_{u.c.} S_{yy}(z)z^k \frac{dz}{2\pi jz} = \int_{u.c.} \frac{z^k}{(z - a_2)(1 - a_2 z)} \frac{dz}{2\pi j} = \frac{a_2^k}{1 - a_2^2}$$

The infinite-order Wiener filter is obtained by

$$H(z) = \frac{1}{\sigma_\epsilon^2 B(z)} \left[\frac{S_{xy}(z)}{B(z^{-1})} \right]_+$$

with $B(z) = H_2(z)$ and $\sigma_\epsilon^2 = \sigma_v^2 = 1$. We find

$$\frac{S_{xy}(z)}{B(z^{-1})} = \frac{H_1(z)H_2(z^{-1})}{H_2(z^{-1})} = H_1(z)$$

which is already causal. Therefore,

$$H(z) = \frac{1}{H_2(z)} H_1(z) = \frac{1 - a_2 z^{-1}}{1 - a_1 z^{-1}} = 1 + (a_1 - a_2)\frac{z^{-1}}{1 - a_1 z^{-1}}$$

having causal impulse response

$$h_n = \delta(n) + (a_1 - a_2)a_1^{n-1} u(n-1)\,, \quad \text{where} \quad u(n) = \text{unit-step}$$

Note, that the infinite-order Wiener filter causes exact cancellation of the $v_1(n)$ component of $x(n)$. Indeed, working with z-transforms, we find for the estimate of $x(n)$

$$\hat{X}(z) = H(z)Y(z) = \frac{H_1(z)}{H_2(z)} H_2(z)V(z) = H_1(z)V(z) = V_1(z)$$

or, $\hat{x}_n = v_1(n)$, and the estimation error becomes

$$e(n) = x(n) - \hat{x}(n) = s(n) + v_1(n) - v_1(n) = s(n)$$

With the choice of parameters $M = 4$, $a_1 = -0.5$, $a_2 = 0.8$, we find for the first $M+1 = 5$ values of the infinite-order impulse response

$$\mathbf{h} = [1, -1.3, 0.65, -0.325, 0.162]^T$$

Using the routine **firw**, we find for the 4th order Wiener filter

$$\mathbf{h} = [1, -1.3, 0.65, -0.325, 0.116]^T$$

The two **h** differ only in their last entry. The corresponding lattice weights are also produced by **firw**:

$$\mathbf{g} = [0.257, -0.929, 0.464, -0.232, 0.116]^T$$

Problem 5.26:
The matrix inverse is easily verified by direct multiplication. For example, for $M = 2$, the autocorrelation matrix is

$$R_{yy} = \frac{1}{1-a_2^2}\begin{bmatrix} 1 & a_2 & a_2^2 \\ a_2 & 1 & a_2 \\ a_2^2 & a_2 & 1 \end{bmatrix}$$

and we can verify explicitly

$$\frac{1}{1-a_2^2}\begin{bmatrix} 1 & a_2 & a_2^2 \\ a_2 & 1 & a_2 \\ a_2^2 & a_2 & 1 \end{bmatrix}\begin{bmatrix} 1 & -a_2 & 0 \\ -a_2 & b & -a_2 \\ 0 & -a_2 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

The optimum Mth order Wiener filter is given by $\mathbf{h} = R_{yy}^{-1}\mathbf{r}$, where the cross-correlation vector $\mathbf{r}$ has entries

$$r_k = R_{xy}(k) = \frac{a_1^k}{1-a_1a_2} \quad \text{for} \quad k = 0, 1, \cdots, M$$

Using the expression for R_{yy}^{-1}, we find for the first and last entries of $\mathbf{h}$

$$h_0 = r_0 - a_2 r_1 = 1$$

$$h_M = r_M - a_2 r_{M-1} = \frac{(a_1 - a_2)a_2^{M-1}}{1-a_1a_2}$$

Note how h_M differs from the infinite-order case. For all the other values of k in the range $1 \le k \le M-1$, the expression for h_k agrees with the infinite-order case. We have:

$$h_k = -a_2 r_{k-1} + b\, r_k - a_2 r_{k+1} = \frac{-a_2 a_1^{k-1} + (1 + a_2^2)a_1^k - a_2 a_1^{k+1}}{1-a_1a_2} = (a_1 - a_2)a_1^{k-1}$$

It is easily verified that the matrix L appearing in the Cholesky factorization of R_{yy}

$$LR_{yy}L^T = D = \text{diag}\{E_0, E_1, \cdots, E_M\}$$

is the lower triangular matrix (shown here for $M = 3$):

$$L = \begin{bmatrix} 1 & 0 & 0 & 0 \\ -a_2 & 1 & 0 & 0 \\ 0 & -a_2 & 1 & 0 \\ 0 & 0 & -a_2 & 1 \end{bmatrix}$$

with the prediction errors

$$E_0 = 1, \quad E_1 = 1 - a_2^2, \quad E_p = E_1, \quad \text{for } p \geq 2$$

Since $\mathbf{h} = L^T \mathbf{g}$, it follows that $h_M = g_M$ and

$$h_m = g_m - a_2 g_{m+1} \quad \text{for} \quad 0 \leq m \leq M-1$$

Problem 5.28:
Initialize the algorithm by defining

$$e_0^{\pm}(n) = y_n = [4.684, 7.247, 8.423, 8.650, 8.640, 8.392] \quad \text{for} \quad n = 0, 1, 2, 3, 4, 5$$

and computing the 0th order error $E_0 = \frac{1}{6}\sum_{n=0}^{5} y_n^2 = 60.884$. Next, compute the first reflection coefficient using

$$\gamma_1 = \frac{2\sum_{n=1}^{5} e_0^+(n) e_0^-(n-1)}{\sum_{n=1}^{5} [e_0^+(n)^2 + e_0^-(n-1)^2]} = \frac{630.177}{638.242} = 0.987$$

Then, $E_1 = (1 - \gamma_1^2)E_0 = 1.573$, and the 1st order prediction error filter will be

$$\begin{bmatrix} 1 \\ a_{11} \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} - 0.987 \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ -0.987 \end{bmatrix}$$

Next, we update the forward and backward error signals. For $n = 1$ to 5 we have

$$e_1^+(1) = e_0^+(1) - \gamma_1 e_0^-(0) = 7.246 - 0.987 \times 4.684 = 2.622$$

$$e_1^-(1) = e_0^-(0) - \gamma_1 e_0^+(1) = 4.684 - 0.987 \times 8.423 = -2.469$$

Similarly,

$$e_1^+(2) = 8.423 - 0.987 \times 7.247 = 1.270, \quad e_1^-(2) = 7.247 - 0.987 \times 8.423 = -1.067$$

$$e_1^+(3) = 8.650 - 0.987 \times 8.423 = 0.336, \quad e_1^-(3) = 8.423 - 0.987 \times 8.650 = -0.115$$

$$e_1^+(4) = 8.640 - 0.987 \times 8.650 = 0.102, \quad e_1^-(4) = 8.650 - 0.987 \times 8.640 = 0.122$$

$$e_1^+(5) = 8.392 - 0.987 \times 8.640 = -0.136, \quad e_1^-(5) = 8.640 - 0.987 \times 8.392 = 0.357$$

The next reflection coefficient is computed from

$$\gamma_2 = \frac{2\sum_{n=2}^{5} e_1^+(n)e_1^-(n-1)}{\sum_{n=2}^{5}[e_1^+(n)^2 + e_1^-(n-1)^2]} = -\frac{7.044}{9.017} = -0.781$$

with prediction error $E_2 = (1-\gamma_2^2)E_1 = 0.614$. The 2nd order prediction filter will be

$$\begin{bmatrix} 1 \\ a_{21} \\ a_{22} \end{bmatrix} = \begin{bmatrix} 1 \\ -0.987 \\ 0 \end{bmatrix} - (-0.781)\begin{bmatrix} 0 \\ -0.987 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ -1.758 \\ 0.781 \end{bmatrix}$$

Problem 5.29:
See solution of Problem 2.4.

Problem 5.31:
The boundary conditions at the interface are

$$E_+ + E_- = E_+' + E_-' \,, \quad E_+ - E_- = \frac{Z}{Z'}(E_+' - E_-')$$

Adding and subtracting, we find

$$E_+ = \frac{1}{2}\left(1 + \frac{Z}{Z'}\right)E_+' + \frac{1}{2}\left(1 - \frac{Z}{Z'}\right)E_-' \,, \quad E_- = \frac{1}{2}\left(1 - \frac{Z}{Z'}\right)E_+' + \frac{1}{2}\left(1 + \frac{Z}{Z'}\right)E_-'$$

But, note

$$\frac{1}{2}\left(1 + \frac{Z}{Z'}\right) = \frac{Z' + Z}{2Z'} = \frac{1}{\tau} \,, \quad \frac{1}{2}\left(1 - \frac{Z}{Z'}\right) = \left(\frac{Z' + Z}{2Z'}\right)\left(\frac{Z' - Z}{Z' + Z}\right) = \frac{\rho}{\tau}$$

Using the above boundary conditions, we have

$$\frac{1}{2}\mathrm{Re}\left[(E_+ + E_-)^* \frac{1}{Z}(E_+ - E_-)\right] = \frac{1}{2}\mathrm{Re}\left[(E_+' + E_-')^* \frac{1}{Z'}(E_+' - E_-')\right]$$

The conservation of the Poynting vector follows from this.

Problem 5.32:
Writing Eq. (5.13.2) in terms of the normalized fields, changes the scale factor $\frac{1}{\tau}$ to

$$\frac{1}{\tau}\frac{\sqrt{Z'}}{\sqrt{Z}} = \frac{Z + Z'}{2\sqrt{ZZ'}} = \frac{1}{t}$$

where

$$t = (1-\rho^2)^{1/2} = \left(1 - \left[\frac{Z - Z'}{Z + Z'}\right]^2\right)^{1/2} = \frac{2\sqrt{ZZ'}}{Z + Z'}$$

Problem 5.33:
Figure 5.14 defines the reflection and transmission responses of a slab structure. From the first figure, it follows that the incoming fields $\begin{bmatrix}1\\0\end{bmatrix}$ will cause the outgoing fields $\begin{bmatrix}T\\R\end{bmatrix}$. From the second figure, the incoming fields $\begin{bmatrix}0\\1\end{bmatrix}$ will cause the outgoing fields $\begin{bmatrix}R'\\T'\end{bmatrix}$. By definition, the outgoing and incoming fields are related by the scattering matrix S. Thus, we must have

$$\begin{bmatrix}T\\R\end{bmatrix} = S\begin{bmatrix}1\\0\end{bmatrix}, \quad \begin{bmatrix}R'\\T'\end{bmatrix} = S\begin{bmatrix}0\\1\end{bmatrix}$$

These must, therefore, be the columns of S; i.e.,

$$S = \begin{bmatrix}T & R'\\R & T'\end{bmatrix}$$

In general, we may write the incoming fields

$$\begin{bmatrix}E_+\\E_-'\end{bmatrix} = \begin{bmatrix}1\\0\end{bmatrix}E_+ + \begin{bmatrix}0\\1\end{bmatrix}E_-'$$

By linear superposition, these will cause the outgoing fields

$$\begin{bmatrix}E_+'\\E_-\end{bmatrix} = \begin{bmatrix}T\\R\end{bmatrix}E_+ + \begin{bmatrix}R'\\T'\end{bmatrix}E_-' = \begin{bmatrix}T & R'\\R & T'\end{bmatrix}\begin{bmatrix}E_+\\E_-'\end{bmatrix}$$

Problem 5.34:
By direct matrix multiplication, we have

$$\psi(z^{-1})^T J_3 \psi(z) = \frac{z^{-1/2}}{t}\begin{bmatrix}1 & \rho\\ \rho z & z\end{bmatrix}\begin{bmatrix}1 & 0\\0 & -1\end{bmatrix}\frac{z^{1/2}}{t}\begin{bmatrix}1 & \rho z^{-1}\\ \rho & z^{-1}\end{bmatrix} = \frac{1}{t^2}\begin{bmatrix}1 & -\rho\\ \rho z & -z\end{bmatrix}\begin{bmatrix}1 & \rho z^{-1}\\ \rho & z^{-1}\end{bmatrix} = \frac{1}{1-\rho^2}\begin{bmatrix}1-\rho^2 & 0\\ 0 & 1-\rho^2\end{bmatrix} = J_3$$

Similarly,

$$J_1\psi(z)J_1 = \frac{z^{1/2}}{t}\begin{bmatrix}0 & 1\\1 & 0\end{bmatrix}\begin{bmatrix}1 & \rho z^{-1}\\ \rho & z^{-1}\end{bmatrix}\begin{bmatrix}0 & 1\\1 & 0\end{bmatrix} = \frac{z^{-1/2}}{t}\begin{bmatrix}1 & \rho z\\ \rho & z\end{bmatrix} = \psi(z^{-1}) = \bar{\psi}(z)$$

Problem 5.35:

The proof is by induction. The required properties are trivially valid for $m = 0$. Assuming they are valid for order m -1, that is,

$$A_{m-1}(z) = 1 + \cdots + \rho_0\rho_{m-1}z^{-(m-1)}\,, \quad B_{m-1}(z) = \rho_{m-1} + \cdots + \rho_0 z^{-(m-1)}$$

and using the lattice recursions we find that the properties are satisfied for order m. In fact,

$$A_m(z) = [1 + \cdots + \rho_0\rho_{m-1}z^{-(m-1)}] + \rho_m z^{-1}[\rho_{m-1} + \cdots + \rho_0 z^{-(m-1)}] = 1 + \cdots + \rho_0\rho_m z^{-m}$$

$$B_m(z) = \rho_m[1 + \cdots + \rho_0\rho_{m-1}z^{-(m-1)}] + z^{-1}[\rho_{m-1} + \cdots + \rho_0 z^{-(m-1)}] = \rho_m + \cdots + \rho_0 z^{-m}$$

Problem 5.36:
From the given expression for $R(z)$, we identify the polynomials $A_4(z)$ and $B_4(z)$:

$$A_4(z) = 1 - 0.125z^{-1} + 0.664z^{-3} - 0.0625z^{-4}$$
$$B_4(z) = -0.25 + 0.0313z^{-1} + 0.2344z^{-2} - 0.2656z^{-3} + 0.25z^{-4}$$

From the lowest coefficient of $B_4(z)$, we find $\rho_4 = -0.25$. Inserting this into the backward recursions, Eqs. (5.13.36), we find

$$A_3(z) = 1 - 0.125z^{-1} + 0.0625z^{-2}$$
$$B_3(z) = 0.25z^{-1} - 0.2656z^{-2} + 0.25z^{-3}$$

From the lowest coefficient of $B_3(z)$, we find $\rho_3 = 0$. Then, the backward recursion gives

$$A_2(z) = 1 - 0.125z^{-1} + 0.0625z^{-2}$$
$$B_2(z) = 0.25 - 0.2656z^{-1} + 0.25z^{-2}$$

This implies $\rho_2 = 0.25$. The backward recursion gives, then

$$A_1(z) = 1 - 0.0625z^{-1}$$
$$B_1(z) = -0.25 + 0.25z^{-1}$$

therefore, $\rho_1 = -0.25$. The last backward recursion gives

$$A_0(z) = 1\,, \quad B_0(z) = \rho_0 = 0.25$$

Problem 5.37:
Use the recursion

$$\sigma_m = t_m t_{m-1}\cdots t_1 t_0 = t_m\sigma_{m-1}\,, \quad \text{where} \quad t_m = \sqrt{1-\rho_m^2}$$

Then, we have

$$T_m(z) = \frac{\sigma_m z^{-m/2}}{A_m(z)} = \frac{t_m z^{-1/2} \sigma_{m-1} z^{-(m-1)/2}}{A_{m-1}(z) + \rho_m z^{-1} B_{m-1}(z)} = \frac{t_m z^{-1/2}}{1 + \rho_m z^{-1} \dfrac{B_{m-1}(z)}{A_{m-1}(z)}} \frac{\sigma_{m-1} z^{-(m-1)/2}}{A_{m-1}(z)}$$

or,

$$T_m(z) = \frac{t_m z^{-1/2} T_{m-1}(z)}{1 + \rho_m z^{-1} R_{m-1}(z)}$$

Problem 5.41:
Using the property $\mathrm{tr}[AB] = \mathrm{tr}[BA]$ which is valid even for non-square matrices (as long as the product AB is square), we find

$$\mathrm{tr}[P] = \mathrm{tr}[Y(Y^T Y)^{-1} Y^T] = \mathrm{tr}[(Y^T Y)^{-1} Y^T Y] = \mathrm{tr}[I_{M+1}] = M+1$$

where I_{M+1} is the $(M+1)\times(M+1)$ identity matrix.

Chapter 6

Problem 6.2:
Starting with Eq. (6.2.7), we have

$$s_\omega^\dagger R s_\omega = \sum_{m,l=0}^{M} R(m-l)e^{-j\omega(m-l)} = \sum_{k,l} R(k)e^{-j\omega k}$$

where we changed the summation indices from the pair $\{m,l\}$ to the pair $\{k,l\}$ where $k = m - l$. Since m and l are in the range $[0, M]$, it follows that k will be in the range $[-M, M]$. Since the l summation is done first, its range may be further restricted by the value of k. Indeed, l must satisfy both of the following inequalities

$$0 \le l \le M\,, \quad 0 \le m = l + k \le M \quad \Longrightarrow \quad -k \le l \le M - k$$

The two combine into one

$$\max(0, -k) \le l \le \min(M, M-k)$$

Thus, the above sum becomes

$$\sum_{k=-M}^{M} \sum_{l=\max(0,-k)}^{\min(M,M-k)} R(k)e^{-j\omega k} = \sum_{k=-M}^{M} [\min(M, M-k) - \max(0,-k) + 1]R(k)e^{-j\omega k}$$

It is now easily verified that

$$\min(M, M-k) - \max(0, -k) + 1 = M + 1 - |k|$$

Problem 6.3:
The peak of the first sidelobe occurs approximately halfway between the first two zeros of $W(\omega)$, that is, at $\omega_1 = 3\pi/(M+1)$. The relative drop with respect to the main lobe will be

$$\left|\frac{W(0)}{W(\omega_1)}\right| = \frac{3\pi}{2}\frac{\sin\left(\dfrac{3\pi}{2(M+1)}\right)}{\left(\dfrac{3\pi}{2(M+1)}\right)}$$

Using the limiting property $\dfrac{\sin x}{x} \to 1$ for small x, we find for large M:

$$\left|\frac{W(0)}{W(\omega_1)}\right| \to \frac{3\pi}{2} \equiv 13.46 \text{ dB}$$

Problem 6.5:
Multiply $R - \sigma_v^2 I = SPS^\dagger$ from the left by $S^\dagger$ and from the right by S to get

$$S^\dagger(R - \sigma_v^2 I)S = (S^\dagger S)P(S^\dagger S) \quad \Rightarrow \quad (S^\dagger S)^{-1}S^\dagger(R - \sigma_v^2 I)S(S^\dagger S)^{-1} = P$$

Problem 6.6
R itself may be expressed recursively as follows. Let

$$R_k = \sigma_v^2 I + \sum_{i=1}^{k} P_i \mathbf{s}_{\omega_k} \mathbf{s}_{\omega_k}^\dagger$$

Then, $R = R_L$, and we have the recursion

$$R_k = R_{k-1} + P_k \mathbf{s}_{\omega_k} \mathbf{s}_{\omega_k}^\dagger$$

It may be initialized by $R_0 = \sigma_v^2 I$. The required expression for R_k^{-1} follows from the matrix inversion lemma, that is,

$$(R + APB)^{-1} = R^{-1} - R^{-1}A(P^{-1} + BR^{-1}A)^{-1}BR^{-1}$$

with the substitutions:

$$R \to R_{k-1}\,, \quad A \to \mathbf{s}_{\omega_k}\,, \quad B \to \mathbf{s}_{\omega_k}^\dagger\,, \quad P \to P_k$$

Problem 6.7:
First, note that for an arbitrary vector $\mathbf{a}$ and corresponding polynomial $A(z)$,

$$\mathbf{a} = [a_0, a_1, \cdots, a_M]^T\,, \quad A(z) = a_0 + a_1 z^{-1} + \cdots + a_M z^{-M}$$

we can write

$$\mathbf{s}_{\omega_1}^\dagger \mathbf{a} = A(e^{j\omega_1})$$

Next, note that the polynomials corresponding to the given vectors $\mathbf{e}_i$ are

$$E_i(z) = z^{-i} E_0(z)\,, \quad i = 0, 1, \cdots, M-1, \quad \text{where} \quad E_0(z) = 1 - e^{j\omega_1} z^{-1}$$

Since $E_0(z)$ has a zero at $e^{j\omega_1}$, so will all of the above. Thus, we find

$$\mathbf{s}_{\omega_1}^\dagger \mathbf{e}_i = E_i(e^{j\omega_1}) = 0\,, \quad i = 0, 1, \cdots, M-1$$

But this implies that they are all eigenvectors of R belonging to the minimum eigenvalue $\lambda = \sigma_v^2$. Indeed,

$$R\mathbf{e}_i = \sigma_v^2\mathbf{e}_i + P_1\mathbf{s}_{\omega_1}\mathbf{s}_{\omega_1}^\dagger\mathbf{e}_i = \sigma_v^2\mathbf{e}_i$$

The polynomial $A(z)$ corresponding to an arbitrary linear combination of $\mathbf{e}_i$ is

$$A(z) = \sum_{i=0}^{M-1} c_i E_i(z) = \sum_{i=0}^{M-1} c_i z^{-i} E_0(z) = E_0(z)\left(\sum_{i=0}^{M-1} c_i z^{-i}\right)$$

or,

$$A(z) = (1 - e^{j\omega_1}z^{-1})\left(c_0 + c_1 z^{-1} + \cdots + c_{M-1} z^{-(M-1)}\right)$$

Problem 6.8:
The first order minimization condition of the extended performance index is

$$\delta\mathcal{E} = \delta\mathbf{a}^\dagger R\mathbf{a} + \mathbf{a}^\dagger R\delta\mathbf{a} - \lambda\delta\mathbf{a}^\dagger\mathbf{a} - \lambda\mathbf{a}^\dagger\delta\mathbf{a} = \delta\mathbf{a}^\dagger(R\mathbf{a} - \lambda\mathbf{a}) + (R\mathbf{a} - \lambda\mathbf{a})^\dagger\delta\mathbf{a} = 0$$

Because $\delta\mathbf{a}$ are varied as though they were unconstrained, the above condition requires

$$R\mathbf{a} = \lambda\mathbf{a}$$

Multiplying by $\mathbf{a}^\dagger$ and imposing the constraint $\mathbf{a}^\dagger\mathbf{a} = 1$, we obtain

$$\mathcal{E} = \mathbf{a}^\dagger R\mathbf{a} = \lambda\mathbf{a}^\dagger\mathbf{a} = \lambda$$

That is, the extremal value of the performance index is equal to the corresponding eigenvalue. Thus, to minimize it we must choose the minimum eigenvalue and corresponding eigenvector.

Problem 6.9:
The conditions given in Eqs (6.3.9) can be written in vector form

$$E[\mathbf{v}(n)^*\mathbf{v}(n)^T] = \sigma_v^2 I\,, \quad E[A_i(n)\mathbf{v}(n)^*] = 0$$

Using these, we obtain

$$R = E[\mathbf{y}(n)^*\mathbf{y}(n)^T] = E[\mathbf{v}(n)^*\mathbf{v}(n)^T] + \sum_{i,j=1}^{L} \mathbf{s}_{k_i} E[A_i(n)^* A_j(n)]\mathbf{s}_{k_j}^\dagger = \sigma_v^2 I + \sum_{i,j=1}^{L} \mathbf{s}_{k_i} P_{ij}\mathbf{s}_{k_j}^\dagger$$

Problem 6.10:
From the Levinson recursion, we have (depending on whether $\gamma_M = -1$ or $\gamma_M = 1$)

$$A_M(z) = A_{M-1} \pm z^{-1}A_{M-1}^R(z) = A_{M-1}(z) \pm z^{-M}A_{M-1}(z^{-1})$$

It follows that $A_M(z)$ will be symmetric or antisymmetric, and therefore its zeros will come in pairs z_i and $1/z_i$. Now, because all the zeros must satisfy the minimum phase conditions $|z_i| \le 1$, it follows that $|1/z_i| \leq 1$ and therefore $|z_i| = 1$. The vector of coefficients satisfies the normal equations $R\mathbf{a}_M = E_M\mathbf{u} = 0$, because $E_M = (1-\gamma_M^2)E_{M-1} = 0$. Therefore, $\mathbf{a}_M$ is an eigenvector belonging to zero eigenvalue. Normalize $\mathbf{a}_M$ to unit norm by defining $\mathbf{e}_0 = \mathbf{a}_M/\|\mathbf{a}_M\|$. Let $E = [\mathbf{e}_1, \mathbf{e}_2, \cdots, \mathbf{e}_M]$ be the rest of the eigenvectors belonging to the remaining M nonzero (positive) eigenvalues of R. They satisfy the eigenvalue equations $RE = E\Lambda$, where Λ is the $M\times M$ diagonal matrix of the nonzero eigenvalues. The completeness of the $M+1$ eigenvectors implies $EE^\dagger + \mathbf{e}_0\mathbf{e}_0^\dagger = I_{M+1}$. Therefore, R is representable only in terms of the nonzero eigenvectors (i.e., the signal subspace eigenvectors)

$$R = REE^\dagger + R\mathbf{e}_0\mathbf{e}_0^\dagger = REE^\dagger = E\Lambda E^\dagger$$

Let $z_i = e^{j\omega_i}$, $i = 1, \cdots, M$, be the M zeros of the polynomial $\mathbf{e}_0$, and define the corresponding M phasing vectors, arranged as the columns of the matrix S:

$$S = [\mathbf{s}_1, \mathbf{s}_2, \cdots, \mathbf{s}_M], \quad \mathbf{s}_i \equiv \mathbf{s}_{\omega_i}, \quad i = 1, 2, \cdots, M$$

Each $\mathbf{s}_i$ is orthogonal to $\mathbf{e}_0$, because $\mathbf{s}_{\omega_i}^\dagger \mathbf{e}_0 = E(e^{j\omega_i}) = 0$. Therefore, they lie in the space spanned by E. Conversely, if the columns of S are linearly independent (which we assume) then S is a non-orthogonal basis of the same space. Thus, the two bases E and S must be linearly related, that is, $E = SC$, where C is an $M\times M$ invertible matrix. It follows that we may express R in terms of the signal basis S, as follows

$$R = E\Lambda E^\dagger = SC\Lambda C^\dagger S^\dagger = SPS^\dagger, \quad \text{where} \quad P \equiv C\Lambda C^\dagger$$

Next, we argue that P must be diagonal. Each phasing vector admits the decomposition (see Problem 5.19)

$$\mathbf{s}_i = \begin{bmatrix} \bar{\mathbf{s}}_i \\ e^{i\omega_i} \end{bmatrix} = \begin{bmatrix} 1 \\ e^{j\omega_i}\bar{\mathbf{s}}_i \end{bmatrix}$$

where $\bar{\mathbf{s}}_i$ is a phasing vector of order $M-1$. It follows that S will admit the decompositions

$$S = \begin{bmatrix} \bar{S} \\ * \end{bmatrix} = \begin{bmatrix} * \\ \bar{S}D \end{bmatrix}, \quad \text{where} \quad D = \text{diag}\{e^{j\omega_1}, \cdots, e^{j\omega_M}\}$$

and the "don't care" entries have been denoted by *. It follows that R admits the decompositions

$$R = SPS^\dagger = \begin{bmatrix} \bar{S}P\bar{S} & * \\ * & * \end{bmatrix} = \begin{bmatrix} * & * \\ * & \bar{S}DPD^\dagger\bar{S}^\dagger \end{bmatrix}$$

If R is Toeplitz, then its upper left corner must equal its lower right corner (see Section 1.7, p.49). Thus, we must have

$$\bar{S}P\bar{S} = \bar{S}DPD^\dagger\bar{S}^\dagger$$

But $\bar{S}$ is a square ($M\times M$) Vandermonde matrix with distinct columns. Thus, it must be invertible. Using $D^\dagger = D^{-1}$, it follows that

$$P = DPD^\dagger \quad \Longrightarrow \quad PD = DP \quad \Longrightarrow \quad P = \text{diagonal}$$

Problem 6.12:

(a) Using the normalization condition $\mathbf{a}^\dagger Q\mathbf{a} = 1$, we find for any generalized eigenvector

$$\lambda = \lambda\mathbf{a}^\dagger Q\mathbf{a} = \mathbf{a}^\dagger R\mathbf{a} = \sigma_v^2\mathbf{a}^\dagger Q\mathbf{a} + \sum_{i=1}^{L} P_i(\mathbf{a}^\dagger \mathbf{s}_{k_i})(\mathbf{s}_{k_i}^\dagger\mathbf{a}) = \sigma_v^2 + \sum_{i=1}^{L} P_i\,|A(e^{jk_i})|^2$$

The positivity of the second term implies that the minimum eigenvalue σ_v^2 will be attained whenever the eigenpolynomial $A(z)$ has roots at the plane waves $z_i = e^{jk_i}$, that is,

$$\mathbf{s}_{k_i}^\dagger\mathbf{a} = A(z_i) = 0\,, \quad i = 1, 2, \cdots, L$$

Conversely, any eigenvector belonging to the minimum eigenvalue must satisfy these L orthogonality conditions. Thus, the dimensionality of the subspace spanned by such eigenvectors (the noise subspace) will be $M+1-L$. It follows from the above expression for λ that the remaining L eigenvectors will have eigenvalues strictly greater than σ_v^2.

(b) The orthogonality condition follows from the reality of eigenvalues and the hermiticity of R and Q. We have

$$0 = \mathbf{a}_1^\dagger R\mathbf{a}_2 - \mathbf{a}_2^\dagger R\mathbf{a}_1 = \lambda_2\mathbf{a}_1^\dagger Q\mathbf{a}_2 - \lambda_1\mathbf{a}_2^\dagger Q\mathbf{a}_1 = (\lambda_2 - \lambda_1)\mathbf{a}_1^\dagger Q\mathbf{a}_2$$

which implies $\mathbf{a}_1^\dagger Q\mathbf{a}_2 = 0$ if $\lambda_1 \neq \lambda_2$.

(c) The L (linearly independent) direction vectors $Q^{-1}\mathbf{s}_{k_i}$ lie in the signal subspace, because they are Q-orthogonal to the noise subspace. Indeed, if $\mathbf{a}$ is any noise subspace eigenvector, we have

$$\mathbf{a}^\dagger Q(Q^{-1}\mathbf{s}_{k_i}) = \mathbf{a}^\dagger\mathbf{s}_{k_i} = A(z_i)^* = 0$$

(d) We have already seen that any noise subspace eigenpolynomial will have at least L zeros at the desired locations. The remaining $M-L$ zeros could lie anywhere.

Problem 6.13:

It is easily checked that $\mathbf{e}_1$ and $\mathbf{e}_2$ are orthogonal to $\mathbf{s}_{k_2}$. For example,

$$\mathbf{s}_{k_2}^\dagger\mathbf{e}_2 = [1, e^{-jk_2}, e^{-2jk_2}]\begin{bmatrix} 0 \\ 1 \\ -e^{jk_2} \end{bmatrix} = 0$$

It follows that $\mathbf{e}_i$ are minimum eigenvectors; i.e.,

$$R\mathbf{e}_i = \sigma_v^2 Q\mathbf{e}_i + P_2\mathbf{s}_{k_2}(\mathbf{s}_{k_2}^\dagger\mathbf{e}_i) = \sigma_v^2 Q\mathbf{e}_i$$

The z-transforms of these eigenvectors are

$$E_1(z) = 1 - e^{jk_2}z^{-1}, \quad E_2(z) = z^{-1}E_1(z)$$

Therefore, the polynomial corresponding to an arbitrary linear combination $\mathbf{a} = \mathbf{e}_1 + \rho_1\mathbf{e}_2$ will be

$$A(z) = E_1(z) + \rho_1 z^{-1}E_1(z) = (1 - e^{jk_2}z^{-1})(1 + \rho_1 z^{-1})$$

exhibiting the desired planewave zero and an arbitrary spurious zero. Next, we verify that $\mathbf{e}_3 = Q^{-1}\mathbf{s}_{k_2}$ is the signal subspace generalized eigenvector

$$R\mathbf{e}_3 = (\sigma_v^2 Q + P_2\mathbf{s}_{k_2}\mathbf{s}_{k_2}^\dagger)Q^{-1}\mathbf{s}_{k_2} = [\sigma_v^2 + P_2(\mathbf{s}_{k_2}^\dagger Q^{-1}\mathbf{s}_{k_2})]\mathbf{s}_{k_2} = \lambda Q\mathbf{e}_3$$

where $\lambda = \sigma_v^2 + P_2(\mathbf{s}_{k_2}^\dagger Q^{-1}\mathbf{s}_{k_2})$. The Q-orthogonality conditions follow from

$$\mathbf{e}_3^\dagger Q\mathbf{e}_i = \mathbf{s}_{k_2}^\dagger Q^{-1}Q\mathbf{e}_i = \mathbf{s}_{k_2}^\dagger\mathbf{e}_i = 0, \quad \text{for} \quad i = 1, 2$$

In the $M = 3$ case, the z-transforms of the three given vectors are

$$E_1(z) = 1 - e^{jk_2}z^{-1}, \quad E_2(z) = z^{-1}E_1(z), \quad E_3(z) = z^{-2}E_1(z)$$

They all have a zero at $z_2 = e^{jk_2}$. Therefore, they satisfy

$$\mathbf{s}_{k_2}^\dagger\mathbf{e}_i = E_i(z_2) = 0, \quad \text{for} \quad i = 1, 2, 3$$

which implies that they are minimum eigenvectors belonging to $\lambda = \sigma_v^2$. An arbitrary linear combination $\mathbf{a} = \mathbf{e}_1 + c_1\mathbf{e}_2 + c_2\mathbf{e}_3$ will have z-transform $A(z) = E_1(z)(1 + c_1 z^{-1} + c_2 z^{-2})$ exhibiting one desired zero and two spurious ones with arbitrary locations. The generalization to arbitrary M is straightforward.

For the final example, we note that the given vector $\mathbf{e}_1$ is orthogonal to both $\mathbf{s}_{k_2}$ and $\mathbf{s}_{k_3}$. For example,

$$\mathbf{s}_{k_2}^\dagger\mathbf{e}_1 = [1, e^{-jk_2}, e^{-2jk_2}]\begin{bmatrix} 1 \\ -(e^{jk_2} + e^{jk_3}) \\ e^{jk_2}e^{jk_3} \end{bmatrix} = 1 - 1 - e^{jk_3 - jk_2} + e^{jk_3 - jk_2} = 0$$

and similarly, $\mathbf{s}_{k_3}^\dagger\mathbf{e}_1 = 0$. These orthogonality conditions imply that $\mathbf{e}_1$ is a minimum eigenvector:

$$R\mathbf{e}_1 = \sigma_v^2 Q\mathbf{e}_1 + P_2\mathbf{s}_{k_2}(\mathbf{s}_{k_2}^\dagger\mathbf{e}_1) + P_3\mathbf{s}_{k_3}(\mathbf{s}_{k_3}^\dagger\mathbf{e}_1) = \sigma_v^2 Q\mathbf{e}_1$$

The corresponding z-transform of $\mathbf{e}_1$ is

$$E_1(z) = 1 - (e^{jk_2} + e^{jk_3})z^{-1} + e^{jk_2}e^{jk_3}z^{-2} = (1 - e^{jk_2}z^{-1})(1 - e^{jk_3}z^{-1})$$

Problem 6.15:
We have $d\mathbf{r}_a = -\bar{R}d\alpha$ and $d\rho_a = dE + 2\alpha^T\bar{R}d\alpha$. Using Eq. (1.7.35), we find

$$R^{-1}dR = \begin{bmatrix} E^{-1} & E^{-1}\alpha^T \\ E^{-1}\alpha & \bar{R}^{-1} + E^{-1}\alpha\alpha^T \end{bmatrix} \begin{bmatrix} dE + 2\alpha^T\bar{R}d\alpha & -d\alpha^T\bar{R} \\ -\bar{R}d\alpha & 0 \end{bmatrix}$$

The product of these two matrices is given in (6.11.9). Writing a similar equation for $R^{-1}\delta R$, that is,

$$R^{-1}\delta R = E^{-1}\begin{bmatrix} \delta E + \alpha^T\bar{R}\delta\alpha & -\delta\alpha^T\bar{R} \\ (\delta E + \alpha^T\bar{R}\delta\alpha)\alpha - E\delta\alpha & -\alpha\delta\alpha^T\bar{R} \end{bmatrix}$$

and multiplying it with (6.11.9), we obtain

$$R^{-1}dRR^{-1}\delta R = E^{-2}\begin{bmatrix} dE\delta E + dE\alpha^T\bar{R}\delta\alpha + Ed\alpha^T\bar{R}\delta\alpha & * \\ * & Ed\alpha\delta\alpha^T\bar{R} - dE\alpha\delta\alpha^T\bar{R} \end{bmatrix}$$

where the "don't care" entries have been denoted by *. The trace is equal to the sum of the upper left scalar plus the trace of the lower right matrix. Thus,

$$\begin{aligned} \mathrm{tr}[R^{-1}dRR^{-1}\delta R] &= E^{-2}\left(dE\delta E + dE\alpha^T\bar{R}\alpha + Ed\alpha^T\bar{R}\delta\alpha + \mathrm{tr}[Ed\alpha\delta\alpha^T\bar{R} - dE\alpha\delta\alpha^T\bar{R}]\right) = \\ &= E^{-2}\left(dE\delta E + dE\alpha^T\bar{R}\alpha + Ed\alpha^T\bar{R}\delta\alpha + Ed\alpha^T\bar{R}\delta\alpha - dE\alpha^T\bar{R}\delta\alpha\right) = \\ &= E^{-2}dE\delta E + 2E^{-1}d\alpha^T\bar{R}\delta\alpha \end{aligned}$$

Problem 6.16:
(a) The required equivalence follows from

$$E[\mathbf{y}\mathbf{y}^T] = E[(\xi + j\eta)(\xi + j\eta)^T] = \left(E[\xi\xi^T] - E[\eta\eta^T]\right) + j\left(E[\xi\eta^T] + E[\eta\xi^T]\right)$$

Using $A = E[\xi\xi^T] = E[\eta\eta^T]$ and $B = E[\xi\eta^T] = -E[\eta\xi^T]$, we obtain

$$R = E[\mathbf{y}^*\mathbf{y}^T] = E[(\xi - j\eta)(\xi + j\eta)^T)] = E[\xi\xi^T] + E[\eta\eta^T] + jE[\xi\eta^T] - jE[\eta\xi^T] = 2(A + jB)$$

similarly,

$$\bar{R} = E\left[\begin{bmatrix}\xi \\ \eta\end{bmatrix}[\xi^T, \eta^T]\right] = \begin{bmatrix} E[\xi\xi^T] & E[\xi\eta^T] \\ E[\eta\xi^T] & E[\eta\eta^T]\end{bmatrix} = \begin{bmatrix} A & B \\ -B & A\end{bmatrix}$$

To show the equality of quadratic forms, let F and G be the subblocks of $\bar{R}^{-1} = \begin{bmatrix} F & G \\ -G & F\end{bmatrix}$. Then, F must be symmetric and G antisymmetric and must satisfy

$$\begin{bmatrix} A & B \\ -B & A \end{bmatrix}\begin{bmatrix} F & G \\ -G & F \end{bmatrix} = \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix} \quad \Rightarrow \quad AF - BG = I\,, \quad AG + BF = 0$$

These two conditions imply now that the inverse of R is of the form $R^{-1} = \frac{1}{2}(F + jG)$. From the symmetry of F, we have $\xi^T F\eta = \eta^T F\xi$, and from the antisymmetry of G we have $\xi^T G\xi = 0$ and $\eta^T G\eta = 0$. It follows that

$$\mathbf{y}^T R^{-1}\mathbf{y}^* = (\xi + j\eta)^{T)}\frac{1}{2}(F + jG)(\xi - j\eta) = \frac{1}{2}\left(\xi^T F\xi + \xi^T G\eta - \eta^T G\xi + \eta^T F\eta\right) =$$

$$= \frac{1}{2}[\xi^T, \eta^T]\begin{bmatrix} F & G \\ -G & F \end{bmatrix}\begin{bmatrix} \xi \\ \eta \end{bmatrix} = \frac{1}{2}\bar{\mathbf{y}}^T\bar{R}^{-1}\bar{\mathbf{y}}$$

To show the relationship between the determinants, we note the identity

$$\begin{bmatrix} I & -BA^{-1} \\ 0 & I \end{bmatrix}\begin{bmatrix} A & B \\ -B & A \end{bmatrix}\begin{bmatrix} I & 0 \\ A^{-1}B & I \end{bmatrix} = \begin{bmatrix} A + BA^{-1}B & 0 \\ 0 & A \end{bmatrix}$$

The first and third matrix factors are unit upper and lower triangular, and thus, they have unit determinants. Using the identity given in the hint, it follows that

$$\det\begin{bmatrix} A & B \\ -B & A \end{bmatrix} = \det(A + BA^{-1}B)\det A = \det(A + jB)(\det A)^{-1}\det(A - jB)\det A = |\det(A + jB)|^2$$

therefore,

$$2^{2M}\det\bar{R} = 2^{2M}|\det(A + jB)|^2 = |\det[2(A + jB)]|^2 = (\det R)^2$$

To find the complex gaussian probability density, note that the volume elements are the same, that is,

$$d^{2M}\mathbf{y} = d^{2M}\bar{\mathbf{y}} = d^M\xi d^M\eta$$

Using the results of part (a), we have

$$p(\mathbf{y}) = p(\bar{\mathbf{y}}) = \frac{\exp(-\bar{\mathbf{y}}^T\bar{R}^{-1}\bar{\mathbf{y}}/2)}{(2\pi)^{2M/2}(\det\bar{R})^{1/2}} = \frac{\exp(-\mathbf{y}^T R^{-1}\mathbf{y}^*)}{\pi^M\det R}$$

There are many ways to prove part (c). One is to use characteristic functions. Another is to use the standard identity for gaussian variables

$$E[y_1y_2y_3y_4] = E[y_1y_2]E[y_3y_4] + E[y_1y_3]E[y_2y_4] + E[y_1y_4]E[y_2y_3]$$

In the complex valued case one can prove it by splitting each y into its real and imaginary parts and using the real-valued version of this identity on the real parts. Applying this identity to the components of $\mathbf{y}$ and using the fact that $E[\mathbf{y}\mathbf{y}^T] = 0$, we obtain

$$E[y_i{}^*y_jy_k{}^*y_l] = E[y_i{}^*y_i]E[y_k{}^*y_l] + E[y_i{}^*y_k{}^*]E[y_jy_l] + E[y_i{}^*y_l]E[y_k{}^*y_j] = R_{ij}R_{kl} + 0\cdot 0 + R_{il}R_{kj}$$

Perhaps, the simplest proof is to consider the Cholesky factorization $R = L^*L^T$ where L is lower triangular and the associated innovations representation $\mathbf{y} = L\boldsymbol{\varepsilon}$, where $E[\boldsymbol{\varepsilon}^*\boldsymbol{\varepsilon}^T] = I$. Then prove the result for uncorrelated complex gaussian $\boldsymbol{\varepsilon}$ of unit variance, namely,

$$E[\varepsilon_i{}^*\varepsilon_j\varepsilon_k{}^*\varepsilon_l] = \delta_{ij}\delta_{kl} + \delta_{il}\delta_{kj}$$

The proof of this is easy. Then, use the linear transformation $\mathbf{y} = L\boldsymbol{\varepsilon}$ to show

$$E[y_i{}^*y_jy_k{}^*y_l] = \sum_{a,b,c,d} L_{ia}{}^*L_{jb}L_{kc}{}^*L_{ld}E[\varepsilon_a{}^*\varepsilon_b\varepsilon_c{}^*\varepsilon_d] = \sum_{a,b,c,d} L_{ia}{}^*L_{jb}L_{kc}{}^*L_{jd}(\delta_{ab}\delta_{cd} + \delta_{ad}\delta_{cb}) =$$

$$= \Big(\sum_a L_{ia}{}^*L_{ja}\Big)\Big(\sum_c L_{kc}{}^*L_{lc}\Big) + \Big(\sum_a L_{ia}{}^*L_{la}\Big)\Big(\sum_c L_{kc}{}^*L_{jc}\Big) = (L^*L^T)_{ij}(L^*L^T)_{kl} + (L^*L^T)_{il}(L^*L^T)_{kj}$$

To show (6.11.12), let y_{ni} denote the ith component of the nth snapshot $\mathbf{y}(n)$. Then, a small generalization of the above result is

$$E[y_{ni}{}^*y_{nj}y_{mk}{}^*y_{ml}] = R_{ij}R_{kl} + \delta_{nm}R_{il}R_{kj}$$

where the indices i,j are associated with the nth snapshot $\mathbf{y}(n)$ and the indices k,l with the mth snapshot $\mathbf{y}(m)$, and we assumed independent snapshots. Then, we find for the correlation of the matrix elements of $\hat{R}$

$$E[\hat{R}_{ij}\hat{R}_{kl}] = \frac{1}{N^2}\sum_{n,m=0}^{N-1} E[y_{ni}{}^*y_{nj}y_{mk}{}^*y_{ml}] = \frac{1}{N^2}\sum_{n,m=0}^{N-1}(R_{ij}R_{kl} + \delta_{nm}R_{il}R_{kj}) = R_{ij}R_{kl} + \frac{1}{N}R_{il}R_{kj}$$

Using the fact that $E[\hat{R}_{ij}] = R_{ij}$, we find

$$E[\Delta R_{ij}\Delta R_{kl}] = E[(\hat{R}_{ij} - R_{ij})(\hat{R}_{kl} - R_{kl})] = E[\hat{R}_{ij}\hat{R}_{kl}] - R_{ij}R_{kl} = \frac{1}{N}R_{il}R_{kj}$$

Problem 6.17:
The joint density of N independent complex gaussian snapshots is

$$p(\mathbf{y}_0, \mathbf{y}_1, \cdots, \mathbf{y}_{N-1}) = (2\pi)^{-MN}(\det R)^{-N}\exp\Big(-\sum_{n=0}^{N-1}\mathbf{y}(n)^TR^{-1}\mathbf{y}(n)^*\Big)$$

Using $\mathbf{y}(n)^TR^{-1}\mathbf{y}(n)^* = \mathrm{tr}(R^{-1}\mathbf{y}(n)^*\mathbf{y}(n)^T)$ and $N\hat{R} = \sum_{n=0}^{N-1}\mathbf{y}(n)^*\mathbf{y}(n)^T$, we obtain

$$-\ln p = c + N\ln(\det R) + \sum_{n=0}^{N-1}\mathbf{y}(n)^*R^{-1}\mathbf{y}(n)^* = c + N\ln(\det R) + \mathrm{tr}\Big[R^{-1}\sum_{n=0}^{N-1}\mathbf{y}(n)^*\mathbf{y}(n)^T\Big]$$

where $c = \ln(2\pi)^{MN}$. Using the property $\ln(\det R) = \mathrm{tr}(\ln R)$, we find

$$-\ln p = c + N\mathrm{tr}(\ln R + R^{-1}\hat{R})$$

Problem 6.18:
Taking first order variations of $R\mathbf{a} = E\mathbf{u}$, we obtain

$$\Delta\mathbf{a} = E^{-1}\Delta E\mathbf{a} - R^{-1}\delta\mathbf{a}\,, \qquad \Delta E = \mathbf{a}^\dagger\Delta R\mathbf{a} = \mathbf{a}^\dagger\delta\mathbf{a}\,, \qquad \delta\mathbf{a} \equiv \Delta R\mathbf{a}$$

Using (6.11.12), we find $E[\delta\mathbf{a}\delta\mathbf{a}^\dagger] = \dfrac{E}{N}R$. It follows that

$$E[(\Delta E)^2] = \mathbf{a}^\dagger E[\delta\mathbf{a}\delta\mathbf{a}^\dagger]\mathbf{a} = \frac{E}{N}\mathbf{a}^\dagger R\mathbf{a} = \frac{E}{N}E = \frac{E^2}{N}$$

We also find $E[\Delta E\delta\mathbf{a}] = E[\delta\mathbf{a}\delta\mathbf{a}^\dagger]\mathbf{a} = \dfrac{E}{N}R\mathbf{a}$. This implies

$$E[\Delta E\Delta\mathbf{a}] = E[\Delta E(E^{-1}\Delta E\mathbf{a} - R^{-1}\delta\mathbf{a})] = E^{-1}\frac{E^2}{N}\mathbf{a} - R^{-1}\frac{E}{N}R\mathbf{a} = 0$$

Using this result, we obtain

$$E[\Delta\mathbf{a}\Delta\mathbf{a}^\dagger] = E[\Delta\mathbf{a}(E^{-1}\Delta E\mathbf{a}^\dagger - \delta\mathbf{a}^\dagger R^{-1})] = -E[\Delta\mathbf{a}\delta\mathbf{a}^\dagger]R^{-1} = -E[(E^{-1}\Delta E\mathbf{a} - R^{-1}\delta\mathbf{a})\delta\mathbf{a}^\dagger]R^{-1} =$$

$$= -E^{-1}\mathbf{a}\left(\frac{E}{N}\mathbf{a}^\dagger R\right)R^{-1} + R^{-1}\left(\frac{E}{N}R\right)R^{-1} = \frac{E}{N}\left(R - \frac{1}{E}\mathbf{a}\mathbf{a}^\dagger\right)$$

Problem 6.19:
Eq. (6.11.12) implies for any four vectors

$$E[(\mathbf{a}^\dagger\Delta R\mathbf{b})(\mathbf{c}^\dagger\Delta R\mathbf{d})] = \frac{1}{N}(\mathbf{a}^\dagger R\mathbf{d})(\mathbf{c}^\dagger R\mathbf{b})$$

Applying this property to $\Delta S(k) = \mathbf{s}_k^\dagger\Delta R\mathbf{s}_k$, we obtain

$$E[(\Delta S(k))^2] = E[(\mathbf{s}_k^\dagger\Delta R\mathbf{s}_k)(\mathbf{s}_k^\dagger\Delta R\mathbf{s}_k)] = \frac{1}{N}(\mathbf{s}_k^\dagger R\mathbf{s}_k)(\mathbf{s}_k^\dagger R\mathbf{s}_k) = \frac{1}{N}S(k)^2$$

For the ML spectrum, we have $S(k)^{-1} = \mathbf{s}_k^\dagger R^{-1}\mathbf{s}_k$. Thus,

$$-S(k)^{-2}\Delta S(k) = -\mathbf{s}_k^\dagger R^{-1}\Delta R R^{-1}\mathbf{s}_k \quad\Longrightarrow\quad \Delta S(k) = S(k)^2(\mathbf{s}_k^\dagger R^{-1}\Delta R R^{-1}\mathbf{s}_k)$$

It follows that

$$E[(\Delta S(k))^2] = S(k)^4 E[(s_k^\dagger R^{-1}\Delta R R^{-1} s_k)(s_k^\dagger R^{-1}\Delta R R^{-1} s_k)] = S(k)^4 \frac{1}{N}(s_k^\dagger R^{-1} R R^{-1} s_k)^2$$

or,

$$E[(\Delta S(k))^2] = S(k)^4 \frac{1}{N}(s_k^\dagger R^{-1} s_k)^2 = S(k)^4 \frac{1}{N} S(k)^{-2} = \frac{1}{N} S(k)^2$$

Problem 6.20:
We have $\Delta A(k) = s_k^\dagger \Delta a$. Thus,

$$E[\Delta A(k)\Delta A(k)^*] = s_k^\dagger E[\Delta a \Delta a^\dagger] s_k = s_k^\dagger \frac{E}{N}\left[R^{-1} - \frac{1}{E} a a^\dagger\right] s_k = \frac{E}{N}\left[s_k^\dagger R^{-1} s_k - \frac{1}{E} A(k) A(k)^*\right]$$

In Problem 5.17, the reproducing kernel was defined by $K(z,w) = s(z)^T R^{-1} s(w)$. Setting $z = w^* = e^{jk}$, we recognize that $s_k^\dagger R^{-1} s_k = K(z,z^*)$. Using the order recursion of Problem 5.19(b), we find in the order-M case

$$s_k^\dagger R^{-1} s_k - \frac{1}{E} |A_M(k)|^2 = K_M(z,z^*) - \frac{1}{E_M} |A_M(z)|^2 = K_{M-1}(z,z^*)$$

which is positive definite. Alternatively, we may use the Schwarz inequality with respect to the positive definite inner product $x^\dagger R^{-1} y$. Then,

$$(s_k^\dagger R^{-1} s_k)(u^\dagger R^{-1} u) \geq |(s_k^\dagger R^{-1} u)|^2$$

where u is the usual unit vector. The LP solution is $a = E R^{-1} u$, and $E = (u^\dagger R^{-1} u)^{-1}$. It follows that

$$s_k^\dagger R^{-1} s_k - \frac{1}{E} |s_k^\dagger a|^2 = s_k^\dagger R^{-1} s_k - E\,|(s_k^\dagger R^{-1} u)|^2 = (s_k^\dagger R^{-1} s_k) - \frac{|(s_k^\dagger R^{-1} u)|^2}{(u^\dagger R^{1} u)}$$

which is non negative by the above inequality.

As above, we can show that $E[\Delta a \Delta a^T] = -E[\Delta a \delta a^T] R^{-1}$. And,

$$E[\Delta a \delta a^T] = E^{-1} a E[\Delta E \delta a^T] - R^{-1} E[\delta a \delta a^T] = \frac{1}{E} a \frac{E}{N} a^T R^T - R^{-1} E[\delta a \delta a^T]$$

We may also find that $E[\delta a \delta a^T] = \frac{1}{N} R a a^T R^T$. This implies

$$E[\Delta a \delta a^T] = \frac{1}{N} a a^T R^T - \frac{1}{N} R^{-1} R a a^T R^T = 0$$

The first order differential of $S(k) = E / |A(k)|^2$ is

$$\Delta S(k) = S(k)\left[\frac{\Delta E}{E} - \frac{\Delta A(k)}{A(k)} - \frac{\Delta A(k)^*}{A(k)^*}\right]$$

We have $E[\Delta A(k)\Delta A(k)] = \mathbf{s}_k^\dagger E[\Delta\mathbf{a}\Delta\mathbf{a}^T]\mathbf{s}_k = 0$, and also $E[\Delta E\Delta A(k)] = \mathbf{s}_k^\dagger E[\Delta E\Delta\mathbf{a}] = 0$. It follows that

$$E[(\Delta S(k))^2] = S(k)^2\left(\frac{1}{E^2}E[(\Delta E)^2] + \frac{2}{|A(k)|^2}E[|\Delta A(k)|^2]\right) =$$

$$= \frac{1}{N}S(k)^2\left(1 + \frac{2E}{|A(k)|^2}(\mathbf{s}_k^\dagger R^{-1}\mathbf{s}_k - \frac{1}{E}|A(k)|^2)\right) = \frac{1}{N}S(k)^2\left(1 + 2S(k)(\mathbf{s}_k^\dagger R^{-1}\mathbf{s}_k) - 2\right) =$$

$$= \frac{1}{N}S(k)^2\left(2S(k)(\mathbf{s}_k^\dagger R^{-1}\mathbf{s}_k) - 1\right)$$

Using the Schwarz inequality, one can easily show $S(k)(\mathbf{s}_k^\dagger R^{-1}\mathbf{s}_k) - 1 \geq 0$ which implies $2S(k)(\mathbf{s}_k^\dagger R^{-1}\mathbf{s}_k) - 1 \geq 1$.

Chapter 7

Problem 7.2:

(a) Incorporate the constraint by means of a Lagrange multiplier defining the "unconstrained" performance index

$$J = E[e_n^2] + \lambda^T(\mathbf{f} - C^T\mathbf{h}) = E[x_n^2] + \mathbf{h}^T R\mathbf{h} - 2\mathbf{h}^T\mathbf{r} + \lambda^T(\mathbf{f} - C^T\mathbf{h})$$

The minimization conditions are

$$\frac{\partial J}{\partial \mathbf{h}} = 2R\mathbf{h} - 2\mathbf{r} - C\lambda = 0 \quad \Rightarrow \quad \mathbf{h} = \mathbf{h}_u + \frac{1}{2}R^{-1}C\lambda$$

where $\mathbf{h}_u \equiv R^{-1}\mathbf{r}$. The Lagrange multiplier is fixed by imposing the constraint:

$$\mathbf{f} = C^T\mathbf{h} = C^T\mathbf{h}_u + \frac{1}{2}C^T R^{-1}C\lambda \quad \Rightarrow \quad \frac{1}{2}\lambda = (C^T R^{-1}C)^{-1}(\mathbf{f} - C^T\mathbf{h}_u)$$

Substituting into the solution for $\mathbf{h}$, we find

$$\mathbf{h} = \mathbf{h}_u + R^{-1}C(C^T R^{-1}C)^{-1}(\mathbf{f} - C^T\mathbf{h}_u)$$

(b) The condition $C^T\Delta\mathbf{h} = 0$ follows from $C^T P = C^T(I - C(C^TC)^{-1}C^T) = 0$.

(c) The above gradient is $\Delta\mathbf{h} = -2\mu P(R\mathbf{h} - \mathbf{r})$. Therefore,

$$\mathbf{h}(n+1) = \mathbf{h}(n) + \Delta\mathbf{h}(n) + \mathbf{h}(n) - 2\mu P(R\mathbf{h}(n) - \mathbf{r}) = P[\mathbf{h}(n) - 2\mu R\mathbf{h}(n) + 2\mu\mathbf{r}] + (I - P)\mathbf{h}(n)$$

But because $\mathbf{h}(n)$ must satisfy the constraints $C^T\mathbf{h}(n) = \mathbf{f}$ at each iteration, we have $(I - P)\mathbf{h}(n) = C(C^TC)^{-1}C^T\mathbf{h}(n) = C(C^TC)^{-1}\mathbf{f} = \mathbf{h}_{LS}$. Part (d) is straightforward.

Problem 7.3:

We have already computed the required gradient in problem 7.2. Then, the ordinary gradient descent update with respect to the extended performance index will be

$$\mathbf{h}(n+1) = \mathbf{h}(n) - \mu\frac{\partial J}{\partial \mathbf{h}(n)} = \mathbf{h}(n) - 2\mu R\mathbf{h}(n) + 2\mu\mathbf{r} + \mu C\lambda(n)$$

If we denote $\mathbf{g}(n) \equiv (I - 2\mu R)\mathbf{h}(n) + 2\mu\mathbf{r}$, the above difference equation can be written as $\mathbf{h}(n+1) = \mathbf{g}(n) + \mu C\lambda(n)$. It follows that

$$\mathbf{f} = C^T\mathbf{h}(n+1) = C^T\mathbf{g}(n) + \mu(C^TC)\lambda(n) \quad \Rightarrow \quad \mu\lambda(n) = (C^TC)^{-1}\mathbf{f} - (C^TC)^{-1}C^T\mathbf{g}(n)$$

from which we obtain, using $\mathbf{h}_{LS} = C(C^TC)^{-1}\mathbf{f}$,

$$\mu C\lambda(n) = \mathbf{h}_{LS} - C(C^TC)^{-1}C^T\mathbf{g}(n) = \mathbf{h}_{LS} - (I - P)\mathbf{g}(n)$$

The original difference equation now becomes

$$\mathbf{h}(n+1) = \mathbf{g}(n) + \mu C\lambda(n) = \mathbf{g}(n) - (I - P)\mathbf{g}(n) + \mathbf{h}_{LS} = P\mathbf{g}(n) + \mathbf{h}_{LS} = P[(I - 2\mu R)\mathbf{h}(n) + 2\mu\mathbf{r}] + \mathbf{h}_{LS}$$

Problem 7.4:
First note that $\mathbf{h} = R^{-1}\mathbf{r}$ is a particular solution of the inhomogeneous difference equation

$$\mathbf{h} = (I - 2MR)\mathbf{h} + 2M\mathbf{r} \quad\Rightarrow\quad R^{-1}\mathbf{r} = (I - 2MR)R^{-1}\mathbf{r} + 2M\mathbf{r}$$

The homogeneous equation has the general solution

$$\mathbf{h}_{\text{homog}}(n+1) = (I - 2MR)\mathbf{h}_{\text{homog}}(n) \quad\Rightarrow\quad \mathbf{h}_{\text{homog}}(n) = (I - 2MR)^n\mathbf{c}$$

Thus, the general solution of the inhomogeneous equation will be

$$\mathbf{h}(n) = \mathbf{h} + \mathbf{h}_{\text{homog}}(n) = \mathbf{h} + (I - 2MR)^n\mathbf{c}$$

The constant $\mathbf{c}$ is fixed by the initial conditions

$$\mathbf{h}(0) = \mathbf{h} + \mathbf{c} \quad\Rightarrow\quad \mathbf{c} = \mathbf{h}(0) - \mathbf{h}$$

Problem 7.5:
The optimal weights are

$$\mathbf{h} = R^{-1}\mathbf{r} = \begin{bmatrix} 3 & 2 \\ 2 & 3 \end{bmatrix}^{-1} \begin{bmatrix} 10 \\ 5 \end{bmatrix} = \begin{bmatrix} 4 \\ -1 \end{bmatrix}$$

The eigenvalues of R are $\lambda = 1, 5$. Thus, the adaptation parameter must be restricted in the range

$$0 < \mu < \frac{1}{\lambda_{\max}} \quad\Rightarrow\quad 0 < \mu < \frac{1}{5}$$

The choice $\mu = 1/6$ is within this range. The matrix $I - 2\mu R$ can be written in the following diagonalized form

$$I - 2\mu R = \begin{bmatrix} 0 & -\frac{2}{3} \\ -\frac{2}{3} & 0 \end{bmatrix} = U \begin{bmatrix} \frac{2}{3} & 0 \\ 0 & -\frac{2}{3} \end{bmatrix} U^{-1}, \quad \text{where} \quad U = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}$$

Therefore,

$$(I - 2\mu R)^n = U\begin{bmatrix} (2/3)^n & 0 \\ 0 & (-2/3)^n \end{bmatrix}U^{-1} = \frac{1}{2}\begin{bmatrix} (2/3)^n + (-2/3)^n & (-2/3)^n - (2/3)^n \\ (-2/3)^n - (2/3)^n & (2/3)^n + (-2/3)^n \end{bmatrix}$$

The transient weights $\mathbf{h}(n) = \mathbf{h} - (I - 2\mu R)^n\mathbf{h}$ are then

$$\mathbf{h}(n) = \begin{bmatrix} 4 \\ -1 \end{bmatrix} - \frac{1}{2}\begin{bmatrix} 5(2/3)^n + 3(-2/3)^n \\ 5(-2/3)^n - 5(2/3)^n \end{bmatrix}$$

Problem 7.6:
The CCL equations are:

$$e_n = x_n - h_n y_n\,, \quad h_{n+1} = h_n + 2\mu e_n y_n$$

Setting $y_n = 1$, we find

$$e_n = x_n - h_n\,, \quad h_{n+1} = h_n + 2\mu e_n$$

and in the z-domain

$$E(z) = X(z) - H(z)\,, \quad zH(z) = H(z) + 2\mu E(z)$$

Eliminating $H(z)$, we find

$$\frac{E(z)}{X(z)} = \frac{1 - z^{-1}}{1 - (1 - 2\mu)z^{-1}}$$

The filter is causal and stable when $|1 - 2\mu| < 1$ or, equivalently, $0 < \mu < 1$. It may be recognized as a special case of the notch filter of Section 7.9.

Problem 7.7:
With $y_n = (-1)^n$, the CCL equations become

$$e_n = x_n - (-1)^n h_n\,, \quad h_{n+1} = h_n + 2\mu e_n(-1)^n$$

Setting $g_n = h_n(-1)^n$, we obtain

$$e_n = x_n - g_n\,, \quad -g_{n+1} = g_n + 2\mu e_n \quad \Longrightarrow \quad E(z) = X(z) - G(z)\,, \quad -zG(z) = G(z) + 2\mu E(z)$$

Eliminating $G(z)$, we find

$$\frac{E(z)}{X(z)} = \frac{1 + z^{-1}}{1 + (1 - 2\mu)z^{-1}}$$

It is another special case of the notch filter of Section 7.9.

Problem 7.8:
Consider the one-dimensional case first. Using the change of variables

$$h_R = \frac{1}{2}(h + h^*), \quad h_I = -j\frac{1}{2}(h - h^*)$$

we have $\partial h_R/\partial h^* = 1/2$ and $\partial h_I/\partial h^* = j/2$. Therefore,

$$\frac{\partial E}{\partial h^*} = \frac{\partial E}{\partial h_R}\frac{\partial h_R}{\partial h^*} + \frac{\partial E}{\partial h_I}\frac{\partial h_I}{\partial h^*} = \frac{1}{2}\frac{\partial E}{\partial h_R} + \frac{1}{2}j\frac{\partial E}{\partial h_I}$$

Similarly, we find

$$\frac{\partial E}{\partial h} = \frac{1}{2}\left[\frac{\partial E}{\partial h_R} - j\frac{\partial E}{\partial h_I}\right]$$

Gradient descent with respect to h_R and h_I is equivalent to

$$\Delta h = \Delta h_R + j\Delta h_I = -\mu\left[\frac{\partial E}{\partial h_R} + j\frac{\partial E}{\partial h_I}\right] = -2\mu\frac{\partial E}{\partial h^*}$$

In the multidimensional case, the equations apply for each component of $\mathbf{h}$.

Problem 7.9:
The magnitude response of the transfer function of Eq. (7.9.3) is

$$|H(\omega)|^2 = \frac{2(1 - \cos(\omega - \omega_0))}{1 - 2(1 - a\mu)\cos(\omega - \omega_0) + (1 - a\mu)^2}$$

where $a = 2(M+1)|A|^2$. Assuming a narrow notch, we set $\omega = \omega_0 + \Delta\omega$ and expand to lowest order in $\Delta\omega$. The numerator becomes proportional to $(\Delta\omega)^2$ and therefore the denominator may be evaluated at $\Delta\omega = 0$. Thus,

$$|H(\omega_0 + \Delta\omega)|^2 = \frac{(\Delta\omega)^2}{1 - 2(1 - a\mu) + (1 - a\mu)^2} = \frac{(\Delta\omega)^2}{(a\mu)^2}$$

Setting this equal to one half the value at some reference frequency ω_{ref} chosen to be very far from ω_0, we obtain

$$(\Delta\omega)^2 = (a\mu)^2\frac{1}{2}|H(\omega_{ref})|^2 = \frac{(a\mu)^2(1 - \cos(\omega_{ref} - \omega_0))}{1 - 2(1 - a\mu)\cos(\omega_{ref} - \omega_0) + (1 - a\mu)^2}$$

Since the numerator is already of order $O(\mu^2)$, this can be simplified if we set $\mu = 0$ in the denominator. Then,

$$(\Delta\omega)^2 = \frac{1}{2}(a\mu)^2 \quad \Longrightarrow \quad (\Delta\omega)_{3\,dB} = 2(\Delta\omega) = \sqrt{2}\,a\mu$$

Problem 7.13:
The autocorrelation function of x_n is $R_{xx}(k) = \sigma_v^2\delta(k) + P_1\mathbf{s}_1\mathbf{s}_1^\dagger$, where we set $P_1 = |A_1|^2$, and $\mathbf{s}_1 = \mathbf{s}_{\omega_1}$. If $\Delta > 0$, then the vector $\mathbf{r}$ will be

$$r_i = R_{xx}(i + \Delta) = P_1 e^{j\omega_1\Delta} e^{j\omega_1 i} \quad \Longrightarrow \quad \mathbf{r} = P_1 e^{j\omega_1\Delta}\mathbf{s}_1$$

Similarly the autocorrelation matrix can be expressed as

$$R = \sigma_v^2 I + P_1\mathbf{s}_1\mathbf{s}_1^\dagger$$

Using the matrix inversion lemma, we find for R^{-1}

$$R^{-1} = \frac{1}{\sigma_v^2}\left[I + d_1\mathbf{s}_1\mathbf{s}_1^\dagger\right], \quad \text{where} \quad d_1 \equiv -\frac{1}{M + 1 + \dfrac{\sigma_v^2}{P_1}}$$

The ALE weights are computed by

$$\mathbf{h} = R^{-1}\mathbf{r} = \frac{1}{\sigma_v^2}[I + d_1\mathbf{s}_1\mathbf{s}_1^\dagger]P_1 e^{j\omega_1\Delta}\mathbf{s}_1 = \frac{P_1}{\sigma_v^2}[1 + (M+1)d_1]e^{j\omega_1\Delta}\mathbf{s}_1$$

and note that $\dfrac{P_1}{\sigma_v^2}[1 + (M+1)d_1] = 1/[M + 1 + \dfrac{\sigma_v^2}{P_1}]$. It follows that

$$\mathbf{h}^\dagger R\mathbf{h} = \mathbf{h}^\dagger(\sigma_v^2 I + P_1\mathbf{s}_1\mathbf{s}_1^\dagger)\mathbf{h} = \sigma_v^2\mathbf{h}^\dagger\mathbf{h} + P_1|\mathbf{h}^\dagger\mathbf{s}_1|^2$$

and we find

$$P_1|\mathbf{h}^\dagger\mathbf{s}_1|^2 = \frac{P_1(M + 1)^2}{\left(M + 1 + \dfrac{\sigma_v^2}{P_1}\right)^2}, \qquad \sigma_v^2\mathbf{h}^\dagger\mathbf{h} = \frac{\sigma_v^2(M + 1)}{\left(M + 1 + \dfrac{\sigma_v^2}{P_1}\right)^2}$$

Thus,

$$\frac{P_1|\mathbf{h}^\dagger\mathbf{s}_1|^2}{\sigma_v^2\mathbf{h}^\dagger\mathbf{h}} = \frac{P_1}{\sigma_v^2}(M + 1)$$

Since there is only one sinewave present, the signal subspace will be one dimensional and the degenerate noise subspace will be M-dimensional. The single signal subspace eigenvector is easily found to be $\mathbf{s}_1$. Its eigenvalue is the maximum eigenvalue: $\lambda_{max} = \sigma_v^2 + P_1(M + 1)$. On the other hand $\lambda_{min} = \sigma_v^2$. Thus,

$$\frac{\lambda_{\max}}{\lambda_{\min}} = 1 + \frac{P_1}{\sigma_v^2}(M + 1)$$

If the delay is zero, $\Delta = 0$, then the vector $\mathbf{r}$ becomes

$$r_i = R_{xx}(i) = \sigma_v^2 \delta(i) + P_1 e^{j\omega_1 i} \quad \Longrightarrow \quad \mathbf{r} = \sigma_v^2 \mathbf{u} + P_1 \mathbf{s}_1$$

where $\mathbf{u} = [1, 0, \cdots, 0]^T$. It follows that

$$\mathbf{h} = R^{-1}\mathbf{r} = \frac{1}{\sigma_v^2}[I + d_1 \mathbf{s}_1 \mathbf{s}_1^\dagger][\sigma_v^2 \mathbf{u} + P_1 \mathbf{s}_1] = \mathbf{u} + [d_1 + \frac{P_1}{\sigma_v^2}(1 + (M + 1)d_1)]\mathbf{s}_1 = \mathbf{u}$$

Problem 7.15:
The extended performance index is $J = \mathbf{a}^\dagger R \mathbf{a} + \lambda(1 - \mathbf{u}^T \mathbf{a})$, with gradient $\partial J/\partial \mathbf{a} = 2R\mathbf{a} - \lambda \mathbf{u}$. Therefore, gradient descent with respect to J will be

$$\mathbf{a}(n+1) = \mathbf{a}(n) - \mu \frac{\partial J}{\partial \mathbf{a}(n)} = \mathbf{a}(n) - 2\mu R \mathbf{a}(n) + \mu\lambda \mathbf{u} = \mathbf{g}(n) + \mu\lambda \mathbf{u}$$

where, as in Problem 7.3, we set $\mathbf{g} = \mathbf{a} - 2\mu R \mathbf{a}(n)$. Imposing the constraint $\mathbf{u}^T \mathbf{a}(n+1) = 1$, we find

$$1 = \mathbf{u}^T \mathbf{a}(n+1) = \mathbf{u}^T \mathbf{g}(n) + \mu\lambda \quad \Longrightarrow \quad \mu\lambda \mathbf{u} = \mathbf{u} - \mathbf{u}\mathbf{u}^T \mathbf{g}(n)$$

which leads to

$$\mathbf{a}(n+1) = \mathbf{g}(n) - \mathbf{u}\mathbf{u}^T \mathbf{g}(n) + \mathbf{u} = (I - \mathbf{u}\mathbf{u}^T)(\mathbf{a}(n) - 2\mu R \mathbf{a}(n)) + \mathbf{u}$$

Note that $\mathbf{u}$ is the least squares solution of $\mathbf{u}^T \mathbf{a} = 1$. Indeed, $\mathbf{a}_{LS} = \mathbf{u}(\mathbf{u}^T \mathbf{u}) \cdot 1 = \mathbf{u}$. Dropping the expectation values in $R\mathbf{a}(n)$, that is, replacing $R\mathbf{a}(n) = E[\mathbf{y}(n)\mathbf{y}(n)^T]\mathbf{a}(n) = E[\mathbf{y}e_n]$ by $\mathbf{y}e_n$, we obtain

$$\mathbf{a}(n+1) = (I - \mathbf{u}\mathbf{u}^T)(\mathbf{a}(n) - 2\mu e_n \mathbf{y}(n)) + \mathbf{u}$$

This is, of course, equivalent to

$$a_0(n+1) = 1 \quad \text{and} \quad a_i(n+1) = a_i(n) - 2\mu e_n y_{n-i}\,, \quad i = 1, 2, \cdots, M$$

Problem 7.18:
The proof is the same as that in Section 7.13.

Problem 7.20:
Let M be the dimension of the vectors. Counting only operations that grow with M, we note that step 1 requires M^2 operations, steps 2, 3, 5, and 6 require M, and step 4 requires $M(M+1)/2$ operations, assuming that only the lower triangular parts need be updated because of the symmetry of the matrices.