SOLUTIONS TO CHAPTER 2
Background

2.1 The DFT of a sequence $x(n)$ of length N may be expressed in matrix form as follows

$$\mathbf{X} = \mathbf{W}\mathbf{x}$$

where $\mathbf{x} = [x(0),\ x(1),\ \ldots, x(N-1)]^{T}$ is a vector containing the signal values and $\mathbf{X}$ is a vector containing the DFT coefficients $X(k)$,

  (a) Find the matrix $\mathbf{W}$.

  (b) What properties does the matrix $\mathbf{W}$ have?

  (c) What is the inverse of $\mathbf{W}$?

Solution _____

  (a) The DFT of a sequence $x(n)$ of length N is

$$X(k) = \sum_{n=0}^{N-1} x(n) e^{-j\frac{2\pi}{N}nk} = \sum_{n=0}^{N-1} x(n) W_N^{nk}$$

where $W_N \equiv e^{-j\frac{2\pi}{N}}$. If we define

$$\mathbf{w}_k^H = \left[1, W_N^k, W_N^{2k}, \ldots, W_N^{k(N-1)}\right]$$

then $X(k)$ is the inner product

$$X(k) = \mathbf{w}_k^H \cdot \mathbf{x}$$

Arranging the DFT coefficients in a vector we have,

$$\mathbf{X} = \begin{bmatrix} X(0) \\ X(1) \\ \vdots \\ X(N-1) \end{bmatrix} = \begin{bmatrix} \mathbf{w}_0^H \mathbf{x} \\ \mathbf{w}_1^H \mathbf{x} \\ \vdots \\ \mathbf{w}_{N-1}^H \mathbf{x} \end{bmatrix} = \mathbf{W}\mathbf{x}$$

where

$$\mathbf{W} = \begin{bmatrix} \mathbf{w}_0^H \\ \mathbf{w}_1^H \\ \vdots \\ \mathbf{w}_{N-1}^H \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & \cdots & 1 \\ 1 & W_N & W_N^2 & \cdots & W_N^{N-1} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & W_N^{N-1} & W_N^{2(N-1)} & \cdots & W_N^{(N-1)^2} \end{bmatrix}$$

  (b) The matrix $\mathbf{W}$ is *symmetric* and nonsingular. In addition, due to the orthogonality of the complex exponentials,

$$\mathbf{w}_k^H \cdot \mathbf{w}_l = \sum_{n=0}^{N-1} e^{-j\frac{2\pi}{N}(k-l)} = \begin{cases} N & ; & k = l \\ 0 & ; & k \neq l \end{cases}$$

it follows that $\mathbf{W}$ is *orthogonal*.

  (c) Due to the orthogonality of $\mathbf{W}$, the inverse is

$$\mathbf{W}^{-1} = \tfrac{1}{N}\mathbf{W}^H$$

2.2 Prove or disprove each of the following statements

    (a) The product of two upper triangular matrices is upper triangular.

    (b) The product of two Toeplitz matrices is Toeplitz.

    (c) The product of two centrosymmetric matrices is centrosymmetric.

Solution _____

    (a) With

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}$$

it follows that if $\mathbf{A}$ is upper triangular then $a_{ij} = 0$ for all $i < j$. If $\mathbf{B}$ is also upper triangular, then the (i,j)th element of the product $\mathbf{C} = \mathbf{AB}$ is

$$c_{ij} = \sum_{k=1}^{n} a_{ik} \cdot b_{kj}$$

For $i < j$ we have

$$c_{ij} = \sum_{k=1}^{j-1} a_{ik} \cdot b_{kj} + \sum_{k=j}^{n} a_{ik} \cdot b_{kj}$$

The first summation is equal to zero since $b_{kj} = 0$ for $k = 1, \ldots, j-1$, and the second term is equal to zero since $a_{ik} = 0$ for $k = j, \ldots, n$. Therefore, $c_{ij} = 0$ for $i < j$ and $\mathbf{C}$ is upper triangular.

    (b) The product of two Toeplitz matrices is *not* necessarily Toeplitz. This may be easily demonstrated by example. Let $\mathbf{A}$ be the following 3×3 Toeplitz matrix,

$$\mathbf{A} = \begin{bmatrix} a_0 & a_{-1} & a_{-2} \\ a_1 & a_0 & a_{-1} \\ a_2 & a_1 & a_0 \end{bmatrix}$$

and let $\mathbf{B}$ be the Toeplitz matrix

$$\mathbf{B} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix}$$

The product, $\mathbf{AB}$, is

$$\mathbf{AB} = \begin{bmatrix} a_{-2} & 0 & a_0 \\ a_{-1} & 0 & a_1 \\ a_0 & 0 & a_2 \end{bmatrix}$$

which is *not* Toeplitz.

    (c) If $\mathbf{A}$ and $\mathbf{B}$ are centrosymmetric matrices, then

$$\mathbf{A} = \mathbf{J}^H \mathbf{A} \mathbf{J} \quad ; \quad \mathbf{B} = \mathbf{J}^H \mathbf{B} \mathbf{J}$$

and

$$\mathbf{AB} = (\mathbf{J}^H \mathbf{A} \mathbf{J})(\mathbf{J}^H \mathbf{B} \mathbf{J})$$

Since $\mathbf{J}\mathbf{J}^H = \mathbf{I}$, then

$$\mathbf{AB} = \mathbf{J}^H\mathbf{AB}\mathbf{J}$$

which means that $\mathbf{AB}$ is centrosymmetric.

2.3 Find the minimum norm solution to the following set of underdetermined linear equations,

$$
\begin{bmatrix} 1 & 0 & 2 & -1 \\ -1 & 1 & 0 & 1 \end{bmatrix}
\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}
=
\begin{bmatrix} 1 \\ 1 \end{bmatrix}
$$

Solution

With

$$
\mathbf{A} = \begin{bmatrix} 1 & 0 & 2 & -1 \\ -1 & 1 & 0 & 1 \end{bmatrix}
$$

since the rows of $\mathbf{A}$ are linearly independent, then the minimum norm solution is unique and given by

$$
\mathbf{x}_0 = \mathbf{A}^H (\mathbf{A}\mathbf{A}^H)^{-1} \mathbf{b}
$$

With

$$
\mathbf{A}\mathbf{A}^H = \begin{bmatrix} 6 & -2 \\ -2 & 3 \end{bmatrix}
$$

and

$$
(\mathbf{A}\mathbf{A}^H)^{-1} = \tfrac{1}{14} \begin{bmatrix} 3 & 2 \\ 2 & 6 \end{bmatrix}
$$

it follows that the minimum norm solution is

$$
\mathbf{x} = \tfrac{1}{14}
\begin{bmatrix} 1 & -1 \\ 0 & 1 \\ 2 & 0 \\ -1 & 1 \end{bmatrix}
\begin{bmatrix} 3 & 2 \\ 2 & 6 \end{bmatrix}
\begin{bmatrix} 1 \\ 1 \end{bmatrix}
= \tfrac{1}{14}
\begin{bmatrix} -3 \\ 8 \\ 10 \\ 3 \end{bmatrix}
$$

2.4 Consider the set of inconsistent linear equations $\mathbf{Ax} = \mathbf{b}$ given by

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$$

(a) Find the least squares solution to these equations.

(b) Find the projection matrix $\mathbf{P}_A$.

(c) Find the best approximation $\widehat{\mathbf{b}} = \mathbf{P}_A \mathbf{b}$ to $\mathbf{b}$.

(d) Consider the matrix

$$\mathbf{P}_A^{\perp} = \mathbf{I} - \mathbf{P}_A$$

Find the vector $\mathbf{b}^{\perp} = \mathbf{P}_A^{\perp}\mathbf{b}$ and show that it is orthogonal to $\widehat{\mathbf{b}}$. What does the matrix $\mathbf{P}_A^{\perp}$ represent?

Solution _____

(a) Since the columns of $\mathbf{A}$ are linearly independent, the least squares solution is unique and given by

$$\mathbf{x}_0 = (\mathbf{A}^H \mathbf{A})^{-1} \mathbf{A}^H \mathbf{b}$$

With

$$\mathbf{A}^H \mathbf{A} = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$$

it follows that

$$(\mathbf{A}^H \mathbf{A})^{-1} = \tfrac{1}{3} \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix}$$

and, therefore,

$$\begin{aligned} \mathbf{x}_0 &= \tfrac{1}{3} \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix} \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} \\ &= \tfrac{1}{3} \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} \\ &= \tfrac{1}{3} \begin{bmatrix} 1 \\ 1 \end{bmatrix} \end{aligned}$$

(b) The projection matrix is

$$\begin{aligned} \mathbf{P}_A &= \mathbf{A}(\mathbf{A}^H \mathbf{A})^{-1} \mathbf{A}^H = \tfrac{1}{3} \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix} \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix} \\ &= \tfrac{1}{3} \begin{bmatrix} 2 & -1 \\ -1 & 2 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix} \\ &= \tfrac{1}{3} \begin{bmatrix} 2 & -1 & 1 \\ -1 & 2 & 1 \\ 1 & 1 & 2 \end{bmatrix} \end{aligned}$$

(c) The best approximation to **b** is

$$\widehat{\mathbf{b}} = \mathbf{P}_A \mathbf{b} = \frac{1}{3} \begin{bmatrix} 2 & -1 & 1 \\ -1 & 2 & 1 \\ 1 & 1 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} = \frac{1}{3} \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}$$

(d) The matrix $\mathbf{P}_\mathbf{A}^\perp$ is

$$\mathbf{P}_\mathbf{A}^\perp = \mathbf{I} - \mathbf{P}_A = \frac{1}{3} \begin{bmatrix} 1 & 1 & -1 \\ 1 & 1 & -1 \\ -1 & -1 & 1 \end{bmatrix}$$

and

$$\mathbf{b}^\perp = \mathbf{P}_\mathbf{A}^\perp \mathbf{b} = \frac{1}{3} \begin{bmatrix} 2 \\ 2 \\ -2 \end{bmatrix}$$

The inner product between $\widehat{\mathbf{b}}$ and $\mathbf{b}^\perp$ is

$$< \widehat{\mathbf{b}}, \ \mathbf{b}^\perp > = \frac{1}{9} \begin{bmatrix} 1 & 1 & 2 \end{bmatrix} \begin{bmatrix} 2 \\ 2 \\ -2 \end{bmatrix} = 0$$

Therefore, $\widehat{\mathbf{b}}$ is *orthogonal* to $\mathbf{b}^\perp$. The matrix $\mathbf{P}_\mathbf{A}^\perp$ is a projection matrix that projects a vector onto the space that is orthogonal to the space spanned by the columns of $\mathbf{A}$.

2.5 Consider the problem of trying to model a sequence $x(n)$ as the sum of a constant plus a complex exponential of frequency ω_0,

$$\widehat{x}(n) = c + ae^{jn\omega_0} \quad ; \quad n = 0, 1, \ldots, N-1$$

where c and a are unknown. We may express the problem of finding the values for c and a as one of solving a set of overdetermined linear equations

$$\begin{bmatrix} 1 & 1 \\ 1 & e^{j\omega_0} \\ \vdots & \vdots \\ 1 & e^{j(N-1)\omega_0} \end{bmatrix} \begin{bmatrix} c \\ a \end{bmatrix} = \begin{bmatrix} x(0) \\ x(1) \\ \vdots \\ x(N-1) \end{bmatrix}$$

(a) Find the least squares solution for c and a.

(b) If N is even and $\omega_0 = 2\pi k/N$ for some integer k, find the least squares solution for c and a.

Solution _____

(a) Assuming that $\omega_0 \neq 0, 2\pi, \ldots$, the columns of the matrix

$$\mathbf{A} = \begin{bmatrix} 1 & 1 \\ 1 & e^{j\omega_0} \\ \vdots & \vdots \\ 1 & e^{j(N-1)\omega_0} \end{bmatrix}$$

are linearly independent, and the least squares solution for c and a is given by

$$\begin{bmatrix} c \\ a \end{bmatrix} = (\mathbf{A}^H\mathbf{A})^{-1}\mathbf{A}^H\mathbf{x}$$

Since

$$\mathbf{A}^H\mathbf{A} = \begin{bmatrix} N & \displaystyle\sum_{n=0}^{N-1} e^{jn\omega_0} \\ \displaystyle\sum_{n=0}^{N-1} e^{-jn\omega_0} & N \end{bmatrix} = \begin{bmatrix} N & \dfrac{1-e^{jN\omega_0}}{1-e^{j\omega_0}} \\ \dfrac{1-e^{-jN\omega_0}}{1-e^{-j\omega_0}} & N \end{bmatrix}$$

Therefore, the inverse of $(\mathbf{A}^H\mathbf{A})$ is

$$(\mathbf{A}^H\mathbf{A})^{-1} = \dfrac{1}{N^2 - \dfrac{1-\cos N\omega_0}{1-\cos\omega_0}} \begin{bmatrix} N & -\dfrac{1-e^{jN\omega_0}}{1-e^{j\omega_0}} \\ -\dfrac{1-e^{-jN\omega_0}}{1-e^{-j\omega_0}} & N \end{bmatrix}$$

and we have

$$\begin{bmatrix} c \\ a \end{bmatrix} = \dfrac{1}{N^2 - \dfrac{1-\cos N\omega_0}{1-\cos\omega_0}} \begin{bmatrix} N & -\dfrac{1-e^{jN\omega_0}}{1-e^{j\omega_0}} \\ -\dfrac{1-e^{-jN\omega_0}}{1-e^{-j\omega_0}} & N \end{bmatrix} \begin{bmatrix} \displaystyle\sum_{n=0}^{N-1} x(n) \\ \displaystyle\sum_{n=0}^{N-1} x(n)e^{-jn\omega_0} \end{bmatrix}$$

which becomes

$$
\begin{bmatrix} c \\ a \end{bmatrix} = \cfrac{1}{N^2 - \cfrac{1 - \cos N\omega_0}{1 - \cos \omega_0}} \begin{bmatrix} N \sum_{n=0}^{N-1} x(n) - \dfrac{1 - e^{jN\omega_0}}{1 - e^{j\omega_0}} \sum_{n=0}^{N-1} x(n)e^{-jn\omega_0} \\[4mm] N \sum_{n=0}^{N-1} x(n)e^{-jn\omega_0} - \dfrac{1 - e^{-jN\omega_0}}{1 - e^{-j\omega_0}} \sum_{n=0}^{N-1} x(n) \end{bmatrix}
$$

(b) If $\omega_0 = 2\pi k/N$ and $k \neq 0$, then

$$
\frac{1 - e^{jN\omega_0}}{1 - e^{j\omega_0}} = \frac{1 - e^{-jN\omega_0}}{1 - e^{-j\omega_0}} = 0
$$

and

$$
\frac{1 - \cos N\omega_0}{1 - \cos \omega_0} = 0
$$

Therefore, we have

$$
\begin{bmatrix} c \\ a \end{bmatrix} = \begin{bmatrix} \dfrac{1}{N} \sum_{n=0}^{N-1} x(n) \\[4mm] \dfrac{1}{N} \sum_{n=0}^{N-1} x(n)e^{-jn\omega_0} \end{bmatrix}
$$

2.6 It is known that the sum of the squares of n from $n = 1$ to $N - 1$ has a closed form expression of the following form

$$\sum_{n=0}^{N-1} n^2 = a_0 + a_1 N + a_2 N^2 + a_3 N^3$$

Given that a third-order polynomial is uniquely determined in terms of the values of the polynomial at four distinct points, derive a closed form expression for this sum by setting up a set of linear equations and solving these equations for a_0, a_1, a_2, a_3. Compare your solution to that given in Table 2.3.

Solution

Assuming that

$$\sum_{n=0}^{N-1} n^2 = a_0 + a_1 N + a_2 N^2 + a_3 N^3$$

we may evaluate this sum for $N = 1, 2, 3, 4$ and write down the following set of four equations in four unknowns

$$\begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 4 & 8 \\ 1 & 3 & 9 & 27 \\ 1 & 4 & 16 & 64 \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ a_2 \\ a_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 5 \\ 14 \end{bmatrix}$$

Solving these equations for a_0, a_1, and a_2, we find

$$\begin{bmatrix} a_0 \\ a_1 \\ a_2 \\ a_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 1/6 \\ -1/2 \\ 1/3 \end{bmatrix}$$

which gives the following closed-form expression for the sum,

$$\sum_{n=0}^{N-1} n^2 = \frac{1}{6} N - \frac{1}{2} N^2 + \frac{1}{3} N^3 = \frac{1}{6} N(N-1)(2N-1)$$

2.7 Show that a projection matrix $\mathbf{P}_A$ has the following two properties,

1. It is *idempotent*, $\mathbf{P}_A^2 = \mathbf{P}_A$.

2. It is Hermitian.

Solution ───

Given a matrix $\mathbf{A}$, the projection matrix $\mathbf{P}_A$ is

$$\mathbf{P}_A = \mathbf{A}(\mathbf{A}^H \mathbf{A})^{-1} \mathbf{A}^H$$

Therefore,

$$\begin{aligned} \mathbf{P}_A^2 &= \mathbf{A}(\mathbf{A}^H \mathbf{A})^{-1} \mathbf{A}^H \mathbf{A}(\mathbf{A}^H \mathbf{A})^{-1} \mathbf{A}^H \\ &= \mathbf{A}(\mathbf{A}^H \mathbf{A})^{-1} \mathbf{A}^H = \mathbf{P}_A \end{aligned}$$

and it follows that $\mathbf{P}_A$ is idempotent. Also,

$$\mathbf{P}_A^H = \left[\mathbf{A}(\mathbf{A}^H \mathbf{A})^{-1} \mathbf{A}^H \right]^H = \mathbf{A}\left[(\mathbf{A}^H \mathbf{A})^{-1} \right]^H \mathbf{A}^H$$

Since $\mathbf{A}\mathbf{A}^H$ is Hermitian, then so is its inverse,

$$\left[(\mathbf{A}^H \mathbf{A})^{-1} \right]^H = (\mathbf{A}^H \mathbf{A})^{-1}$$

and

$$\mathbf{P}_A^H = \mathbf{A}(\mathbf{A}^H \mathbf{A})^{-1} \mathbf{A}^H$$

Thus, $\mathbf{P}_A$ is Hermitian.

───

2.8 Let $\mathbf{A} > 0$ and $\mathbf{B} > 0$ be positive definite matrices. Prove or disprove the following statements.

(a) $\mathbf{A}^2 > 0$.

(b) $\mathbf{A}^{-1} > 0$.

(c) $\mathbf{A} + \mathbf{B} > 0$.

Solution _____

(a) Let $\mathbf{v}_k$ be an eigenvector and λ_k the corresponding eigenvalue of $\mathbf{A}$. Since

$$\mathbf{A}^2\mathbf{v}_k = \mathbf{A}(\mathbf{A}\mathbf{v}_k) = \lambda_k\mathbf{A}\mathbf{v}_k = \lambda_k^2\mathbf{v}_k$$

then $\mathbf{v}_k$ is an eigenvector of $\mathbf{A}^2$ and λ_k^2 is the corresponding eigenvalue. If $\mathbf{A} > 0$, then $\lambda_k > 0$. Therefore, $\lambda_k^2 > 0$, and it follows that $\mathbf{A}^2 > 0$.

(b) If $\mathbf{A} > 0$, then the eigenvalues of $\mathbf{A}$ are positive, $\lambda_k > 0$. In addition, $\mathbf{A}^{-1}$ exists and the eigenvalues of $\mathbf{A}^{-1}$ are λ_k^{-1}. Since $\lambda_k > 0$, it follows that $\lambda_k^{-1} > 0$ and, therefore, $\mathbf{A}^{-1} > 0$.

(c) Let $\mathbf{v} \neq 0$ be an arbitrary vector. Then

$$\mathbf{v}^H(\mathbf{A} + \mathbf{B})\mathbf{v} = \mathbf{v}^H\mathbf{A}\mathbf{v} + \mathbf{v}^H\mathbf{B}\mathbf{v}$$

If $\mathbf{A} > 0$ and $\mathbf{B} > 0$, then

$$\mathbf{v}^H\mathbf{A}\mathbf{v} > 0 \quad ; \quad \mathbf{v}^H\mathbf{B}\mathbf{v} > 0$$

Therefore,

$$\mathbf{v}^H(\mathbf{A} + \mathbf{B})\mathbf{v} > 0$$

and it follows that $(\mathbf{A} + \mathbf{B}) > 0$.

2.9 (a) Prove that each eigenvector of a symmetric Toeplitz matrix is either symmetric or anti-symmetric, i.e., $\mathbf{v}_k = \pm\mathbf{J}\mathbf{v}_k$.

(b) What property can you state about the eigenvalues of a Hermitian Toeplitz matrix?

Solution _____

(a) If $\mathbf{A}$ is a symmetric Toeplitz matrix, then

$$\mathbf{J}^T\mathbf{A}\mathbf{J} = \mathbf{A}$$

where $\mathbf{J}$ is the exchange matrix. If $\mathbf{v}_k$ is an eigenvector of $\mathbf{A}$ with eigenvalue λ_k, then

$$\mathbf{A}\mathbf{v}_k = \lambda_k\mathbf{v}_k$$

and, using the identity above, we have

$$\mathbf{J}^T\mathbf{A}\mathbf{J}\mathbf{v}_k = \lambda_k\mathbf{v}_k$$

Since $\mathbf{J}$ is unitary, $\mathbf{J}^T\mathbf{J} = \mathbf{I}$, if we multiply both sides of this equation on the left by $\mathbf{J}$, it follows that

$$\mathbf{A}\mathbf{J}\mathbf{v}_k = \lambda_k\mathbf{J}\mathbf{v}_k$$

Therefore, if $\mathbf{v}_k$ is an eigenvector with eigenvalue λ_k, then $\mathbf{J}\mathbf{v}_k$ is also an eigenvector with the same eigenvalue. Consequently, if the eigenvalue λ_k is *distinct*, then $\mathbf{v}_k$ and $\mathbf{J}\mathbf{v}_k$ must be equal to within a constant,

$$\mathbf{v}_k = c\mathbf{J}\mathbf{v}_k$$

However, since the exchange matrix reverses the order of the elements of the vector $\mathbf{v}_k$, the only possible values for this constant are $c = \pm 1$. Therefore,

$$\mathbf{v}_k = \pm\mathbf{J}\mathbf{v}_k$$

and the eigenvector $\mathbf{v}_k$ is either symmetric or anti-symmetric.

Now let us consider the case in which the eigenvalue λ_k is *not distinct*. We will assume that the multiplicity is two. The following discussion may be easily generalized to higher multiplicities. In this case, $\mathbf{v}_k$ and $\mathbf{J}\mathbf{v}_k$ span a two-dimensional space, and any two linearly independent vectors in this space may be selected as the eigenvectors. Therefore, we may choose

$$\widetilde{\mathbf{v}}_{k_1} = \mathbf{v}_k + \mathbf{J}\mathbf{v}_k$$

and

$$\widetilde{\mathbf{v}}_{k_2} = \mathbf{v}_k - \mathbf{J}\mathbf{v}_k$$

as the two eigenvectors. Note that $\widetilde{\mathbf{v}}_{k_1}$ is symmetric and $\widetilde{\mathbf{v}}_{k_2}$ is anti-symmetric. This completes the proof.

(b) In the case of Hermitian Toeplitz matrices, the eigenvectors are either Hermitian or anti-Hermitian, i.e.,

$$\mathbf{v}_k = \pm\mathbf{v}_k^*$$

2.10 (a) Find the eigenvalues and eigenvectors of the real 2×2 symmetric Toeplitz matrix

$$\mathbf{A} = \begin{bmatrix} a & b \\ b & a \end{bmatrix}$$

(b) Find the eigenvalues and eigenvectors of the 2×2 Hermitian matrix

$$\mathbf{A} = \begin{bmatrix} a & b^* \\ b & a \end{bmatrix}$$

Solution _____

(a) The eigenvalues are the roots of the characteristic equation

$$\det(\mathbf{A} - \lambda \mathbf{I}) = (a - \lambda)^2 - b^2 = 0$$

Expanding the quadratic in λ we have

$$\lambda^2 - 2a\lambda + (a^2 - b^2) = \left[\lambda - (a + b)\right]\left[\lambda - (a - b)\right] = 0$$

Therefore, the eigenvalues are $\lambda_1 = a + b$ and $\lambda_2 = a - b$. The eigenvectors, on the other hand, are solutions to the equation

$$\mathbf{A}\mathbf{v}_k = \lambda_k \mathbf{v}_k$$

For the first eigenvector, $\mathbf{v}_1$, we have

$$\begin{bmatrix} a & b \\ b & a \end{bmatrix} \begin{bmatrix} v_{11} \\ v_{12} \end{bmatrix} = (a + b) \begin{bmatrix} v_{11} \\ v_{12} \end{bmatrix}$$

which gives $v_{11} = v_{12}$, or

$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

Similarly, the eigenvector $\mathbf{v}_2$ is found to be

$$\mathbf{v}_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

(b) With

$$\mathbf{A} = \begin{bmatrix} a & b^* \\ b & a \end{bmatrix}$$

the eigenvalues are the roots of the characteristic equation

$$\det(\mathbf{A} - \lambda \mathbf{I}) = (a - \lambda)^2 - |b|^2 = 0$$

or,

$$\lambda^2 - 2a\lambda + a^2 + |b|^2 = \left[\lambda - (a + |b|)\right]\left[\lambda - (a - |b|)\right] = 0$$

Thus, $\lambda_1 = a + |b|$ and $\lambda_2 = a - |b|$.

The eigenvector that has eigenvalue λ_1 is the solution to

$$\begin{bmatrix} a & b^* \\ b & a \end{bmatrix} \begin{bmatrix} v_{11} \\ v_{12} \end{bmatrix} = (a + |b|) \begin{bmatrix} v_{11} \\ v_{12} \end{bmatrix}$$

which gives $v_{12} = \dfrac{b}{|b|} v_{11}$, or

$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ b/|b| \end{bmatrix}$$

Similarly, for $\mathbf{v}_2$ we have

$$\mathbf{v}_2 = \begin{bmatrix} 1 \\ -b/|b| \end{bmatrix}$$

2.11 Establish Property 5 on p. 45.

Solution ──

Let $\mathbf{B}$ be an $n \times n$ matrix with eigenvalues λ_k and eigenvectors $\mathbf{v}_k$. With

$$\mathbf{A} = \mathbf{B} + \alpha\mathbf{I}$$

note that

$$\begin{aligned}
\mathbf{A}\mathbf{v}_k &= \mathbf{B}\mathbf{v}_k + \alpha\mathbf{v}_k \\
&= \lambda_k\mathbf{v}_k + \alpha\mathbf{v}_k = (\lambda_k + \alpha)\mathbf{v}_k
\end{aligned}$$

Therefore, $\mathbf{A}$ and $\mathbf{B}$ have the same eigenvectors, and the eigenvalues of $\mathbf{A}$ are $\lambda_i + \alpha$.

──

2.12 A necessary and sufficient condition for a Hermitian matrix $\mathbf{A}$ to be positive definite is that there exists a non-singular matrix $\mathbf{W}$ such that

$$\mathbf{A} = \mathbf{W}^H \mathbf{W}$$

(a) Prove this result.

(b) Find a factorization of the form $\mathbf{A} = \mathbf{W}^H \mathbf{W}$ for the matrix

$$\mathbf{A} = \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix}$$

Solution _____

(a) If $\mathbf{A} > 0$, then $\mathbf{A}$ may be factored as follows

$$\mathbf{A} = \mathbf{V}\mathbf{\Lambda}\mathbf{V}^H$$

where $\mathbf{\Lambda} = \text{diag}\{\lambda_1, \ldots, \lambda_N\}$ with $\lambda_i > 0$. Therefore, $\mathbf{\Lambda}$ may be factored as follows

$$\mathbf{\Lambda} = \mathbf{\Lambda}^{1/2}\mathbf{\Lambda}^{1/2}$$

where $\mathbf{\Lambda}^{1/2} = \text{diag}\{\lambda_1^{1/2}, \ldots, \lambda_N^{1/2}\} > 0$. Thus, we may write

$$\mathbf{A} = \left(\mathbf{V}\mathbf{\Lambda}^{1/2}\right)\left(\mathbf{\Lambda}^{1/2}\mathbf{V}^H\right) = \left(\mathbf{\Lambda}^{1/2}\mathbf{V}^H\right)^H \left(\mathbf{\Lambda}^{1/2}\mathbf{V}^H\right) = \mathbf{W}^H \mathbf{W}$$

where $\mathbf{W} = \mathbf{\Lambda}^{1/2}\mathbf{V}^H > 0$ is nonsingular.

Conversely, suppose that $\mathbf{A}$ may be factored as

$$\mathbf{A} = \mathbf{W}^H \mathbf{W}$$

where $\mathbf{W}$ is a nonsingular matrix. Then $\mathbf{W}$ may be factored as follows

$$\mathbf{W} = \mathbf{V}\mathbf{\Lambda}\mathbf{V}^H$$

where $\mathbf{\Lambda}$ is a diagonal matrix and $\mathbf{V}$ is a unitary matrix. Thus,

$$\mathbf{A} = \mathbf{W}^H \mathbf{W} = \left(\mathbf{V}\mathbf{\Lambda}\mathbf{V}^H\right)^H \left(\mathbf{V}\mathbf{\Lambda}\mathbf{V}^H\right) = \mathbf{V}\mathbf{\Lambda}^2\mathbf{V}^H$$

Since the diagonal terms of $\mathbf{\Lambda}^2$ are positive, then $\mathbf{A} > 0$.

(b) The eigenvalues of $\mathbf{A}$ are $\lambda_1 = 3$ and $\lambda_2 = 1$, and the normalized eigenvectors are

$$\mathbf{v}_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \end{bmatrix} \quad ; \quad \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

Therefore,

$$\mathbf{W}^H = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & \sqrt{3} \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & \sqrt{3} \\ 1 & -\sqrt{3} \end{bmatrix}$$

2.13 Consider the 2×2 matrix

$$\mathbf{A} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$$

(a) Find the eigenvalues and eigenvectors of $\mathbf{A}$.

(b) Are the eigenvectors unique? Are they linearly independent? Are they orthogonal?

(c) Diagonalize $\mathbf{A}$, i.e., find $\mathbf{V}$ and $\mathbf{D}$ such that

$$\mathbf{V}^H \mathbf{A} \mathbf{V} = \mathbf{D}$$

where $\mathbf{D}$ is a diagonal matrix.

Solution _____

(a) The eigenvalues are the roots of the characteristic equation

$$\det(\mathbf{A} - \lambda \mathbf{I}) = \lambda^2 + 1 = 0$$

which are $\lambda = \pm j$. The eigenvector corresponding to the eigenvalue $\lambda_1 = j$ satisfies the equation

$$\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = j \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$$

which implies that $v_2 = j v_1$. Therefore, the normalized eigenvector is

$$\mathbf{v}_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ j \end{bmatrix}$$

Similarly for the eigenvector corresponding to the eigenvalue $\lambda_2 = -j$ we have

$$\mathbf{v}_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -j \end{bmatrix}$$

(b) The eigenvectors are unique, linearly independent, and orthogonal,

$$\langle \mathbf{v}_1, \ \mathbf{v}_2 \rangle = \mathbf{v}_1^H \mathbf{v}_2 = 0$$

(c) With $\mathbf{V}$ the matrix of normalized eigenvectors,

$$\mathbf{V} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ j & -j \end{bmatrix}$$

we have

$$\mathbf{V}^H \mathbf{A} \mathbf{V} = \mathbf{D}$$

where

$$\mathbf{D} = \begin{bmatrix} j & 0 \\ 0 & -j \end{bmatrix}$$

2.14 Find the eigenvalues and eigenvectors of the matrix

$$\mathbf{A} = \begin{bmatrix} 1 & -1 \\ 2 & 4 \end{bmatrix}$$

Solution _____

The eigenvalues of a matrix $\mathbf{A}$ are the roots of the characteristic equation

$$\det(\mathbf{A} - \lambda\mathbf{I}) = 0$$

For the given matrix, we have

$$\begin{aligned}
\det(\mathbf{A} - \lambda\mathbf{I}) &= \det\begin{pmatrix} 1-\lambda & -1 \\ 2 & 4-\lambda \end{pmatrix} \\
&= (1-\lambda)(4-\lambda) + 2 = \lambda^2 - 5\lambda + 6 = (\lambda - 3)(\lambda - 2)
\end{aligned}$$

Therefore, the eigenvalues are $\lambda_1 = 3$ and $\lambda_2 = 2$. The eigenvectors are found by solving the equations

$$\mathbf{A}\mathbf{v}_i = \lambda_i\mathbf{v}_i \quad ; \quad i = 1, 2$$

For $\lambda_1 = 3$ we have

$$\begin{bmatrix} 1 & -1 \\ 2 & 4 \end{bmatrix} \begin{bmatrix} v_{11} \\ v_{12} \end{bmatrix} = 3 \begin{bmatrix} v_{11} \\ v_{12} \end{bmatrix}$$

The first equation is

$$v_{11} - v_{12} = 3v_{11}$$

or

$$v_{12} = -2v_{11}$$

Therefore, the eigenvector is

$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ -2 \end{bmatrix}$$

Repeating for $\lambda_2 = 2$ we find

$$\mathbf{v}_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

2.15 Consider the following 3×3 symmetric matrix

$$\mathbf{A} = \begin{bmatrix} 1 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 1 \end{bmatrix}$$

(a) Find the eigenvalues and eigenvectors of $\mathbf{A}$.

(b) Find the determinant of $\mathbf{A}$.

(c) Find the spectral decomposition of $\mathbf{A}$.

(d) What are the eigenvalues of $\mathbf{A} + \mathbf{I}$ and how are the eigenvectors related to those of $\mathbf{A}$?

Solution ——————————————————————————

(a) The eigenvalues are found from the roots of the characteristic equation,

$$\det(\mathbf{A} - \lambda \mathbf{I}) = 0$$

The roots are $\lambda = 3, 1, 0$. Given the eigenvalues, the eigenvectors are found by solving the equations $\mathbf{A}\mathbf{v}_i = \lambda_i \mathbf{v}_i$ for $i = 1, 2, 3$. The eigenvectors (unnormalized) are

$$\mathbf{V} = \begin{bmatrix} \mathbf{v}_1; \mathbf{v}_2; \mathbf{v}_3 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 \\ -2 & 0 & 1 \\ 1 & -1 & 1 \end{bmatrix}$$

(b) The determinant is equal to the product of the eigenvalues,

$$\det \mathbf{A} = \prod_{i=1}^{3} \lambda_i = 0$$

(c) The spectral decomposition for $\mathbf{A}$ is

$$\mathbf{A} = \sum_{i=1}^{3} \lambda_i \mathbf{v}_i \mathbf{v}_i^H$$

where $\mathbf{v}_i$ are the *normalized* eigenvectors of $\mathbf{A}$. Since $\lambda_3 = 0$, this decomposition becomes

$$
\begin{aligned}
\mathbf{A} &= 3 \cdot \left(\tfrac{1}{6}\right) \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix} \begin{bmatrix} 1 & -2 & 1 \end{bmatrix} + \tfrac{1}{2} \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} \begin{bmatrix} 1 & 0 & -1 \end{bmatrix} \\
&= \tfrac{1}{2} \begin{bmatrix} 1 & -2 & 1 \\ -2 & 4 & -2 \\ 1 & -2 & 1 \end{bmatrix} + \tfrac{1}{2} \begin{bmatrix} 1 & 0 & -1 \\ 0 & 0 & 0 \\ -1 & 0 & 1 \end{bmatrix}
\end{aligned}
$$

(d) If the eigenvalues of $\mathbf{A}$ are λ_i, then the eigenvalues of $\mathbf{A} + \mathbf{I}$ are $\lambda_i + 1$, and the eigenvectors are the same. Therefore, the eigenvalues of $\mathbf{A} + \mathbf{I}$ are $\lambda = 4, 2, 1$.

2.16 Suppose that an $n \times n$ matrix $\mathbf{A}$ has eigenvalues $\lambda_1, \ldots, \lambda_n$ and eigenvectors $\mathbf{v}_1, \ldots, \mathbf{v}_n$.

(a) What are the eigenvalues and eigenvectors of $\mathbf{A}^2$?

(b) What are the eigenvalues and eigenvectors of $\mathbf{A}^{-1}$?

Solution _____

(a) With $\mathbf{v}_i$ an eigenvector of $\mathbf{A}$ with eigenvalue λ_i, note that

$$\mathbf{A}^2\mathbf{v}_i = \mathbf{A}(\mathbf{A}\mathbf{v}_i) = \lambda_i(\mathbf{A}\mathbf{v}_i) = \lambda_i^2\mathbf{v}_i$$

Therefore, the eigenvectors of $\mathbf{A}^2$ are the same as those for $\mathbf{A}$, and the eigenvalues are λ_i^2.

(b) Since

$$\mathbf{A}\mathbf{v}_i = \lambda_i\mathbf{v}_i$$

then, assuming that $\mathbf{A}^{-1}$ exists,

$$\mathbf{v}_i = \lambda_i\mathbf{A}^{-1}\mathbf{v}_i$$

or

$$\mathbf{A}^{-1}\mathbf{v}_i = \frac{1}{\lambda_i}\mathbf{v}_i$$

Therefore, $\mathbf{A}^{-1}$ has the same eigenvectors as $\mathbf{A}$, and the eigenvalues are $1/\lambda_i$.

2.17 Find a matrix whose eigenvalues are $\lambda_1 = 1$ and $\lambda_2 = 4$ with eigenvectors $\mathbf{v}_1 = [3,1]^T$ and $\mathbf{v}_2 = [2,1]^T$.

Solution _____

From the given information, we have

$$\mathbf{A} \begin{bmatrix} 3 \\ 1 \end{bmatrix} = \begin{bmatrix} 3 \\ 1 \end{bmatrix} \quad ; \quad \mathbf{A} \begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 8 \\ 4 \end{bmatrix}$$

Let

$$\mathbf{A} = \begin{bmatrix} \mathbf{a}_1, & \mathbf{a}_2 \end{bmatrix}$$

Then we have

$$3\mathbf{a}_1 + \mathbf{a}_2 = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$$

and

$$2\mathbf{a}_1 + \mathbf{a}_2 = \begin{bmatrix} 8 \\ 4 \end{bmatrix}$$

Subtracting these two equations gives

$$\mathbf{a}_1 = \begin{bmatrix} -5 \\ -3 \end{bmatrix}$$

Also, we have

$$\mathbf{a}_2 = \begin{bmatrix} 8 \\ 4 \end{bmatrix} - 2\mathbf{a}_1 = \begin{bmatrix} 18 \\ 10 \end{bmatrix}$$

Therefore,

$$\mathbf{A} = \begin{bmatrix} -5 & 18 \\ -3 & 10 \end{bmatrix}$$

2.18 Gerschgorin's circle theorem states that every eigenvalue of a matrix $\mathbf{A}$ lies in at least one of the circles $C_1, \ldots, C_N$ in the complex plane where C_i has center at the diagonal entry a_{ii} and its radius is $r_i = \sum_{j \neq i} |a_{ij}|$.

1. Prove this theorem by using the eigenvalue equation $\mathbf{A}\mathbf{x} = \lambda \mathbf{x}$ to write

$$(\lambda - a_{ii})x_i = \sum_{j \neq i} a_{ij}x_j$$

   and then use the triangle inequality,

$$\left| \sum_{j \neq i} a_{ij}x_j \right| \leq \sum_{j \neq i} |a_{ij}x_j|$$

2. Use this theorem to establish the bound on λ_{max} given in Property 7.

3. The matrix

$$\begin{bmatrix} 4 & 1 & 2 \\ 2 & 3 & 0 \\ 3 & 2 & 6 \end{bmatrix}$$

   is said to be *diagonally dominant* since $|a_{ii}| > r_i$. Use Gerschgorin's circle theorem to show that this matrix is nonsingular.

Solution _____

1. Let $\mathbf{x} = \begin{bmatrix} x_1, \ldots, x_N \end{bmatrix}^T$ be an eigenvector, and λ the corresponding eigenvalue for the matrix $\mathbf{A}$. Assume that x_i is the largest component of $\mathbf{x}$, i.e, $|x_i| \geq |x_j|$ for all $j \neq i$. With $\mathbf{A}\mathbf{x} = \lambda \mathbf{x}$, it follows that

$$\sum_{j=1}^{N} a_{ij}x_j = \lambda x_i$$

   or,

$$(\lambda - a_{ii})x_i = \sum_{j \neq i} a_{ij}x_j$$

   Therefore,

$$|\lambda - a_{ii}| = \left| \sum_{j \neq i} a_{ij}\frac{x_j}{x_i} \right| \leq \sum_{j \neq i} |a_{ij}| \left| \frac{x_j}{x_i} \right|$$

   Since $|x_i| \geq |x_j|$ for all $j \neq i$, then the ratios $|x_j/x_i|$ are less than or equal to one, and λ lies in the ith circle defined by

$$|\lambda_i - a_{ii}| \leq r_i$$

   where

$$r_i = \sum_{j \neq i} |a_{ij}|$$

2. From Gerschgorin's circle theorem, for each eigenvalue, λ, there is an i such that

$$|\lambda - a_{ii}| \leq \sum_{j \neq i} |a_{ij}|$$

Since

$$|\lambda| - |a_{ii}| \leq |\lambda - a_{ii}|$$

then

$$|\lambda| \leq \sum_{j=1}^{n} |a_{ij}|$$

Therefore,

$$|\lambda_{\max}| \leq \max_{i} \sum_{j=1}^{n} |a_{ij}|$$

3. Let $\mathbf{A}$ be a matrix that is diagonally dominant,

$$|a_{ii}| > r_i$$

Assume that one of the eigenvalues is zero ($\mathbf{A}$ is singular). From Gerschgorin's circle theorem, we know that, for each eigenvalue,

$$|\lambda - a_{ii}| \leq r_i$$

However, if $\lambda_k = 0$, then

$$|\lambda_k - a_{ii}| = |a_{ii}| \leq r_i$$

for some i. Therefore, $\mathbf{A}$ is not diagonally dominant, which contradicts the hypothesis. Thus, if $\mathbf{A}$ is diagonally dominant, then it cannot have any zero eigenvalues and must, therefore, be nonsingular.

2.19 Consider the following quadratic function of two variables z_1 and z_2,

$$f(z_1, z_2) = 3z_1^2 + 3z_2^2 + 4z_1z_2 + 8$$

Find the values of z_1 and z_2 that minimize $f(z_1, z_2)$ subject to the constraint that $z_1 + z_2 = 1$ and determine the minimum value of $f(z_1, z_2)$.

Solution _____

To minimize the function

$$f(z_1, z_2) = 3z_1^2 + 3z_2^2 + 4z_1z_2 + 8$$

subject to the constraint

$$z_1 + z_2 = 1$$

we may use Lagrange multipliers as follows. If we define the objective function $Q(z_1, z_2)$ as follows

$$Q(z_1, z_2) = 3z_1^2 + 3z_2^2 + 4z_1z_2 + 8 + \lambda(1 - z_1 - z_2)$$

then the values for z_1 and z_2 that minimize $f(z_1, z_2)$ may be found by solving the equations

$$\frac{\partial}{\partial z_1} Q(z_1, z_2) = 6z_1 + 4z_2 - \lambda = 0$$

$$\frac{\partial}{\partial z_2} Q(z_1, z_2) = 6z_2 + 4z_1 - \lambda = 0$$

$$\frac{\partial}{\partial \lambda} Q(z_1, z_2) = 1 - z_1 - z_2 = 0$$

Writing the first two equations in matrix form we have

$$\begin{bmatrix} 6 & 4 \\ 4 & 6 \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} = \lambda \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

Solving for z_1 and z_2 we find

$$\begin{bmatrix} z_1 \\ z_2 \end{bmatrix} = \frac{\lambda}{20} \begin{bmatrix} 6 & -4 \\ -4 & 6 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \frac{\lambda}{10} \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

Plugging these values into the third equation above, we may solve for the Lagrange multiplier, λ, as follows,

$$1 - z_1 - z_2 = 1 - \tfrac{\lambda}{10} - \tfrac{\lambda}{10} = 1 - \tfrac{\lambda}{5} = 0$$

or

$$\lambda = 5$$

Given λ we may explicitly evaluate z_1 and z_2,

$$z_1 = 1/2 \quad ; \quad z_2 = 1/2$$

Substituting these values into $f(z_1, z_2)$ we find that the minimum value of the function is

$$\min\big[f(z_1, z_2)\big] = 10.5$$

SOLUTIONS TO CHAPTER 3
Discrete Time Random Processes

3.1 Let x be a random variable with mean m_x and variance σ_x^2. Let x_i for $i = 1, 2, \ldots, N$ be N independent measurements of the random variable x.

(a) With $\widehat{m}_x$ the sample mean defined by

$$\widehat{m}_x = \frac{1}{N} \sum_{i=1}^{N} x_i$$

determine whether or not the sample variance

$$\widehat{\sigma}_x^2 = \frac{1}{N} \sum_{i=1}^{N} (x_i - \widehat{m}_x)^2$$

is unbiased, i.e., is $E\{\widehat{\sigma}_x^2\} = \sigma_x^2$?

(b) If x is a Gaussian random variable, find the variance of the sample variance, $E\{(\widehat{\sigma}_x^2 - E\{\widehat{\sigma}_x^2\})^2\}$.

Solution _____

(a) The expected value of the sample variance is

$$E\{\widehat{\sigma}_x^2\} = E\left\{ \frac{1}{N} \sum_{i=1}^{N} \left(x_i - \frac{1}{N} \sum_{j=1}^{N} x_j \right)^2 \right\} = \frac{1}{N} \sum_{i=1}^{N} E\left\{ \left[(x_i - m_x) - \frac{1}{N} \sum_{j=1}^{N} (x_j - m_x) \right]^2 \right\}$$

Expanding the square we have

$$E\{\widehat{\sigma}_x^2\} = \frac{1}{N} \sum_{i=1}^{N} E\left\{ (x_i - m_x)^2 - \frac{2}{N} \sum_{j=1}^{N} (x_i - m_x)(x_j - m_x) + \frac{1}{N^2} \sum_{i=1}^{N} \sum_{j=1}^{N} (x_i - m_x)(x_j - m_x) \right\}$$

Since the measurements are assumed to be independent, then

$$E\{(x_i - m_x)(x_j - m_x)\} = \left\{ \begin{array}{ll} \sigma_x^2 & ; \quad i = j \\ 0 & ; \quad i \neq j \end{array} \right.$$

and the expression for $\widehat{\sigma}_x^2$ becomes

$$E\{\widehat{\sigma}_x^2\} = \frac{1}{N} \sum_{i=1}^{N} \left\{ \sigma_x^2 - \frac{2}{N} \sigma_x^2 + \frac{1}{N^2} N \sigma_x^2 \right\} = \sigma_x^2 \left(1 - \frac{1}{N} \right) = \sigma_x^2 \frac{N-1}{N}$$

Therefore, although the sample variance is biased, it is asymptotically unbiased.

(b) Finding the variance of the sample variance directly is very tedious. A simpler way is as follows. With

$$\widehat{\sigma}_x^2 = \frac{1}{N} \sum_{i=1}^{N} \left(x_i - \widehat{m}_x \right)^2$$

it is well-known that

$$\frac{N\widehat{\sigma}_x^2}{\sigma_x^2} = \sum_{i=1}^{N} \left(\frac{x_i - \widehat{m}_x}{\sigma_x} \right)^2$$

is a Chi-square random variable with $n - 1$ degrees of freedom, which has a variance of $2(n - 1)$. Therefore,

$$\text{Var}\left(\frac{N\widehat{\sigma}_x^2}{\sigma_x^2} \right) = 2(N - 1)$$

and, consequently, we have

$$\text{Var}\left(\widehat{\sigma_x^2} \right) = 2\frac{\sigma_x^4}{N^2}(N - 1)$$

3.2 Let $x(n)$ be a stationary random process with zero mean and autocorrelation $r_x(k)$. We form the process, $y(n)$, as follows

$$y(n) = x(n) + f(n)$$

where $f(n)$ is a known deterministic sequence. Find the mean $m_y(n)$ and the autocorrelation $r_y(k, l)$ of the process $y(n)$.

Solution _____

The mean of the process is

$$m_y(n) = E\{y(n)\} = E\{x(n)\} + f(n) = f(n)$$

and the autocorrelation is

$$
\begin{aligned}
r_y(k, l) &= E\{y(k)y(l)\} = E\{[x(k) + f(k)][x(l) + f(l)]\} \\
&= E\{x(k)x(l)\} + f(k)f(l) = r_x(k, l) + f(k)f(l)
\end{aligned}
$$

3.3 A discrete-time random process $x(n)$ is generated as follows,

$$x(n) = \sum_{k=1}^{p} a(k)x(n-k) + w(n)$$

where $w(n)$ is a white noise process with variance σ_w^2. Another process, $z(n)$, is formed by adding noise to $x(n)$,

$$z(n) = x(n) + v(n)$$

where $v(n)$ is white noise with a variance of σ_v^2 that is uncorrelated with $w(n)$.

(a) Find the power spectrum of $x(n)$.

(b) Find the power spectrum of $z(n)$.

Solution _____

(a) Since $x(n)$ is the output of an all-pole filter driven by white noise, $x(n)$ is an AR(p) process with a power spectrum

$$P_x(e^{j\omega}) = \frac{\sigma_w^2}{|A(e^{j\omega})|^2}$$

where

$$A(e^{j\omega}) = 1 - \sum_{k=1}^{p} a(k)e^{-jk\omega}$$

(b) The process $z(n)$ is a sum of two random processes

$$z(n) = x(n) + v(n)$$

Since $x(n)$ is a linear combination of values of $w(n)$,

$$x(n) = \sum_{k=-\infty}^{n} h(k)w(n-k)$$

where $h(n)$ is the unit sample response of the filter generating $x(n)$, and since $v(n)$ is uncorrelated with $w(n)$, then $v(n)$ is uncorrelated with $x(n)$, and we have

$$r_z(k) = r_x(k) + r_v(k)$$

Therefore,

$$P_z(e^{j\omega}) = P_x(e^{j\omega}) + P_v(e^{j\omega})$$

and

$$P_z(e^{j\omega}) = \frac{\sigma_w^2}{|A(e^{j\omega})|^2} + \sigma_v^2 = \frac{\sigma_w^2 + \sigma_v^2|A(e^{j\omega})|^2}{|A(e^{j\omega})|^2}$$

3.4 Suppose we are given a linear shift-invariant system having a system function

$$H(z) = \frac{1 - \frac{1}{2}z^{-1}}{1 - \frac{1}{3}z^{-1}}$$

that is excited by zero mean exponentially correlated noise $x(n)$ with an autocorrelation sequence

$$r_x(k) = \left(\frac{1}{2}\right)^{|k|}$$

Let $y(n)$ be the output process, $y(n) = x(n) * h(n)$.

(a) Find the power spectrum, $P_y(z)$, of $y(n)$.

(b) Find the autocorrelation sequence, $r_y(k)$, of $y(n)$.

(c) Find the cross-correlation, $r_{xy}(k)$, between $x(n)$ and $y(n)$.

(d) Find the *cross-power spectral density*, $P_{xy}(z)$, which is the z transform of the cross-correlation $r_{xy}(k)$.

Solution _____

(a) The power spectrum of $x(n)$ is

$$P_x(z) = \frac{3/4}{(1 - \frac{1}{2}z^{-1})(1 - \frac{1}{2}z)}$$

and the power spectrum of $y(n)$ is

$$P_y(z) = H(z)H(z^{-1})P_x(z) = \frac{3/4}{(1 - \frac{1}{3}z^{-1})(1 - \frac{1}{3}z)}$$

(b) The autocorrelation sequence for $y(n)$ may be easily found using the z-transform pair

$$\alpha^{|k|} \longleftrightarrow \frac{1 - \alpha^2}{(1 - \alpha z^{-1})(1 - \alpha z)}$$

Since

$$\left(\tfrac{1}{3}\right)^{|k|} \longleftrightarrow \frac{8/9}{(1 - \frac{1}{3}z^{-1})(1 - \frac{1}{3}z)}$$

then

$$r_y(k) = \tfrac{27}{32}\left(\tfrac{1}{3}\right)^{|k|}$$

(c) The cross-correlation $r_{xy}(k)$ between $x(n)$ and $y(n)$ is

$$r_{xy}(k) = r_x(k) * h(-k)$$

This may be easily computed using z-transforms as follows,

$$
\begin{aligned}
P_{xy}(z) &= P_x(z)H(z^{-1}) = \frac{3/4}{(1 - \frac{1}{2}z^{-1})(1 - \frac{1}{2}z)} \cdot \frac{1 - \frac{1}{2}z}{1 - \frac{1}{3}z} \\
&= \frac{3/4}{(1 - \frac{1}{2}z^{-1})(1 - \frac{1}{3}z)}
\end{aligned}
$$

Writing this in terms of z^{-1} and performing a partial fraction expansion gives

$$P_{xy}(z) = \frac{3}{4}\frac{z^{-1}}{(1-\frac{1}{2}z^{-1})(z^{-1}-\frac{1}{3})} = \frac{9/10}{1-\frac{1}{2}z^{-1}} + \frac{3/10}{z^{-1}-\frac{1}{3}}$$

Inverse z-transforming gives

$$r_{xy}(k) = \frac{9}{10}(\tfrac{1}{2})^k u(k) + \frac{9}{10}(3)^{-k}u(-k-1)$$

(d) The cross-power spectral density, $P_{xy}(z)$, as computed in part (a), is

$$P_{xy}(z) = \frac{3/4}{(1-\frac{1}{2}z^{-1})(1-\frac{1}{3}z)}$$

(e) The cross-correlation, $r_{xy}(k)$, between $x(n)$ and $y(n)$ may found by computing the inverse z-transform of the cross-power spectral density,

$$P_{xy}(z) = \frac{3}{4}\frac{z^{-1}}{(1-\frac{1}{2}z^{-1})(z^{-1}-\frac{1}{3})} = \frac{\frac{9}{10}}{1-\frac{1}{2}z^{-1}} + \frac{\frac{3}{10}}{z^{-1}-\frac{1}{3}}$$

Inverse transforming gives

3.5 Find the power spectrum for each of the following wide-sense stationary random processes that have the given autocorrelation sequences.

(a) $r_x(k) = 2\delta(k) + j\delta(k-1) - j\delta(k+1)$.

(b) $r_x(k) = \delta(k) + 2(0.5)^{|k|}$.

(c) $r_x(k) = 2\delta(k) + \cos(\pi k/4)$.

(d) $r_x(k) = \begin{cases} 10 - |k| & ; \quad |k| < 10 \\ 0 & ; \quad \text{otherwise} \end{cases}$

Solution _____

(a) This autocorrelation sequence is finite in length, and the power spectrum is simply

$$P_x(e^{j\omega}) = 2 + je^{-j\omega} - je^{j\omega} = 2 + 2\sin\omega$$

Note that, as required, $P_x(e^{j\omega})$ is real and non-negative.

(b) With α real, using the DTFT pair

$$\alpha^{|k|} \longrightarrow \frac{1 - \alpha^2}{|1 - \alpha e^{-j\omega}|^2}$$

we have

$$P_x(e^{j\omega}) = 1 + 2\frac{1 - \frac{1}{4}}{|1 - \frac{1}{2}e^{-j\omega}|^2} = 1 + \frac{\frac{3}{2}}{\frac{5}{4} - \cos\omega} = \frac{11 - 4\cos\omega}{5 - 4\cos\omega}$$

(c) Since the DTFT of a complex exponential is an impulse,

$$e^{jn\omega_0} \longleftrightarrow 2\pi\delta(\omega - \omega_0)$$

it follows that the power spectrum of $r_x(k) = 2\delta(k) + \cos(\pi k/4)$ is

$$P_x(e^{j\omega}) = 2 + \pi\delta(\omega - \tfrac{\pi}{4}) + \pi\delta(\omega + \tfrac{\pi}{4})$$

(d) Observe that $r_x(k)$ is a triangle that is symmetric about $k = 0$ and extends from $k = -9$ to $k = 9$. Therefore, $r_x(k)$ may be written as the convolution of two pulses,

$$r_x(k) = p(k) * p(-k)$$

where

$$p(k) = \begin{cases} 1 & ; \quad 0 \le k \le 9 \\ 0 & ; \quad \text{else} \end{cases}$$

Since the DTFT of $p(k)$ is

$$P(e^{j\omega}) = \sum_{k=0}^{9} e^{-jk\omega} = \frac{1 - e^{-j10\omega}}{1 - e^{-j\omega}} = e^{-j\frac{9}{2}\omega}\frac{\sin 5\omega}{\sin\omega/2}$$

then the power spectrum is

$$P_x(e^{j\omega}) = |P(e^{j\omega})|^2 = \frac{\sin^2 5\omega}{\sin^2 \omega/2}$$

3.6 Find the autocorrelation sequence corresponding to each of the following power spectral densities.

(a) $P_x(e^{j\omega}) = 3 + 2\cos\omega$.

(b) $P_x(e^{j\omega}) = \dfrac{1}{5 + 3\cos\omega}$.

(c) $P_x(z) = \dfrac{-2z^2 + 5z - 2}{3z^2 + 10z + 3}$.

Solution _____

(a) Expanding $P_x(e^{j\omega})$ in terms of complex exponentials,

$$P_x(e^{j\omega}) = 3 + 2\cos\omega = 3 + e^{-j\omega} + e^{j\omega}$$

it follows that $r_x(0) = 3$ and $r_x(1) = r_x(-1) = 1$.

(b) Recall the DTFT pair

$$\alpha^{|k|} \longleftrightarrow \frac{1 - \alpha^2}{(1 - \alpha e^{-j\omega})(1 - \alpha e^{j\omega})} = \frac{1 - \alpha^2}{(1 + \alpha^2) - 2\alpha\cos\omega}$$

Since

$$P_x(e^{j\omega}) = \frac{1}{5 + 3\cos\omega} = \frac{1/5}{1 + \frac{3}{5}\cos\omega}$$

it follows that

$$r_x(k) = \tfrac{1}{4}\left(-\tfrac{1}{3}\right)^{|k|}$$

(c) With

$$P_x(z) = \frac{-2z^2 + 5z - 2}{3z^2 + 10z + 3} = \frac{-2z + 5 - 2z^{-1}}{(3 + z)(3 + z^{-1})} = \frac{1}{9}\frac{-2z + 5 - 2z^{-1}}{(1 + \frac{1}{3}z)(1 + \frac{1}{3}z^{-1})}$$

using the pair

$$\left(-\tfrac{1}{3}\right)^{|k|} \longleftrightarrow \frac{\frac{8}{9}}{(1 + \frac{1}{3}z)(1 + \frac{1}{3}z^{-1})}$$

it follows that

$$r_x(k) = \tfrac{5}{8}\left(-\tfrac{1}{3}\right)^{|k|} - \tfrac{2}{8}\left(-\tfrac{1}{3}\right)^{|k-1|} - \tfrac{2}{8}\left(-\tfrac{1}{3}\right)^{|k+1|}$$

3.7 Let $x(n)$ be a zero mean WSS process with an $N \times N$ autocorrelation matrix $\mathbf{R}_x$. Determine whether each of the following statements are *True* or *False*.

(a) If the eigenvalues of $\mathbf{R}_x$ are equal, $\lambda_1 = \lambda_2 = \cdots = \lambda_N$, then $r_x(k) = 0$ for $k = 1, 2, \ldots, N$.

(b) If $\lambda_1 > 0$ and $\lambda_k = 0$ for $k = 2, 3, \ldots, N$, then $r_x(k) = Ae^{jk\omega_0}$.

Solution _____

(a) This statement is *True*. To show this, the first step is to recall the *Bordering Theorem* which, for Toeplitz matrices, states that if $\mathbf{R}_{p-1}$ is a $p \times p$ Toeplitz matrix with ordered eigenvalues $\lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_p$, and if $\mathbf{R}_p$ is the $(p+1) \times (p+1)$ Toeplitz matrix that is formed by adding one row and column to $\mathbf{R}_{p-1}$, then the ordered eigenvalues of $\mathbf{R}_p$, denoted by $\tilde{\lambda}_1 \leq \tilde{\lambda}_2 \leq \cdots \leq \tilde{\lambda}_{p+1}$ are interlaced with those of $\mathbf{R}_{p-1}$ as follows

$$\tilde{\lambda}_1 \leq \lambda_1 \leq \tilde{\lambda}_2 \leq \lambda_2 \leq \cdots \leq \tilde{\lambda}_p \leq \lambda_p \leq \tilde{\lambda}_{p+1}$$

What this implies is that if the eigenvalues of $\mathbf{R}_p$ are equal, $\tilde{\lambda}_1 = \tilde{\lambda}_2 = \cdots = \tilde{\lambda}_{p+1}$, then the eigenvalues of each of the lower-order Toeplitz matrices must also be equal.

The next step is to note that for any 2×2 Toeplitz matrix

$$\mathbf{R}_1 = \begin{bmatrix} r_x(0) & r_x(1) \\ r_x(1) & r_x(0) \end{bmatrix}$$

if the eigenvalues are equal then $r_x(1) = 0$. This follows easily from the fact that the eigenvalues of $\mathbf{R}_1$ are $\lambda_1 = r_x(0) + r_x(1)$ and $\lambda_2 = r_x(0) - r_x(1)$. Therefore, λ_1 will be equal to λ_2 if and only if $r_x(1) = 0$. This property may also be established for a 3×3 Toeplitz matrix by equating the coefficients of the powers of λ in the equation

$$\det[\mathbf{R}_2 - \lambda\mathbf{I}] = (\lambda - \lambda_0)^3$$

where the three eigenvalues of $\mathbf{R}_2$ are equal to λ_0.

We may now establish the result by induction. Assume that $\mathbf{R}_{k-1}$ is a $k \times k$ Toeplitz matrix with equal eigenvalues, and that

$$\mathbf{R}_{k-1} = \text{Toep}\{r_x(0), 0, 0, \ldots, 0\}$$

We will now show that if $\mathbf{R}_k$ is a $(k+1) \times (k+1)$ Toeplitz matrix

$$\mathbf{R}_k = \text{Toep}\{r_x(0), 0, \ldots, 0, r_x(k)\}$$

with equal eigenvalues, then $r_x(k) = 0$. The eigenvalues of $\mathbf{R}_k$ are the roots of the polynomial

$$\det(\mathbf{R}_k - \lambda\mathbf{I}) = 0$$

If the eigenvalues are to be equal, then

$$\det(\mathbf{R}_k - \lambda\mathbf{I}) = (\lambda - \lambda_0)^k$$

However, note that

$$\det(\mathbf{R}_k - \lambda\mathbf{I}) = [\lambda - r_x(0)]^k \pm r_x^2(p)[\lambda - r_x(0)]^{k-2}$$

where the sign depends on whether k is even or odd. Therefore,

$$\left[\lambda - r_x(0)\right]^k \pm r_x^2(p)\left[\lambda - r_x(0)\right]^{k-2} = \left(\lambda - \lambda_0\right)^k$$

if and only if $\lambda_0 = r_x(0)$ and $r_x(k) = 0$ as was to be shown.

(b) This statement is *True*. To show this, note that if $\lambda_1 > 0$ and $\lambda_k = 0$ for $k = 2, 3, \ldots, N$, then the autocorrelation matrix has the form

$$\mathbf{R}_x = \lambda_1 \mathbf{v}_1 \mathbf{v}_1^H$$

where $\mathbf{v}_1$ is the eigenvector associated with the nonzero eigenvalue λ_1. Let $v_1(k)$ be the coefficients of the eigenvector $\mathbf{v}_1$,

$$\mathbf{v}_1 = \left[v_1(1), \; v_1(2), \; \ldots, \; v_1(N)\right]^T$$

Then

$$\mathbf{R}_x = \lambda_1 \begin{bmatrix} v_1(1) \\ v_1(2) \\ \vdots \\ v_1(N) \end{bmatrix} \begin{bmatrix} v_1^*(1) & v_1^*(2) & \cdots & v_1^*(N) \end{bmatrix} = \begin{bmatrix} |v_1(1)|^2 & v_1(1)v_1^*(2) & \cdots & v_1(1)v_1^*(N) \\ v_1(2)v_1^*(1) & |v_1(2)|^2 & \cdots & v_1(2)v_1^*(N) \\ \vdots & \vdots & & \vdots \\ v_1(N)v^*(1) & v_1(N)v_1^*(2) & \cdots & |v_1(N)|^2 \end{bmatrix}$$

Since $\mathbf{R}_x$ is Toeplitz, then the terms along the main diagonal must be equal,

$$|v_1(1)|^2 = |v_1(2)|^2 = \cdots = |v_1(N)|^2$$

Therefore, the coefficients $v_1(k)$ must have the form,

$$v_1(k) = \sqrt{A}\; e^{j\phi_k}$$

In addition, the Toeplitz structure of $\mathbf{R}_x$ implies that the terms along the diagonal below the main diagonal must be equal,

$$v_1(2)v_1^*(1) = v_1(3)v_1^*(2) = \cdots = v_1(N)v_1^*(N-1)$$

Therefore, $(\phi_k - \phi_{k-1})$ must be a constant,

$$\phi_k - \phi_{k-1} = \omega_0$$

and, with $v_1(1) = \sqrt{A}e^{j\phi_1}$, it follows that

$$v_1(k) = \sqrt{A}\; e^{j\left[\phi_1 + (k-1)\omega_0\right]}$$

Finally, note that since the first column of $\mathbf{R}_x$ contains the autocorrelations $r_x(k)$ for $k = 0, 1, \ldots, N-1$, then

$$r_x(k) = v_1(k+1)v_1^*(1) = \sqrt{A}\; e^{j\left[\phi_1 + k\omega_0\right]} \cdot \sqrt{A}\; e^{-j\phi_1} = A\; e^{jk\omega_0}$$

as was to be shown.

3.8 Consider the random process

$$x(n) = A\cos(n\omega + \phi) + w(n)$$

where $w(n)$ is zero mean white Gaussian noise with a variance σ_w^2. For each of the following cases, find the autocorrelation sequence and if the process is WSS, find the power spectrum.

(a) A is a Gaussian random variable with zero mean and variance σ_A^2 and both ω and ϕ are constants.

(b) ϕ is uniformly distributed over the interval $[-\pi, \pi]$ and both A and ω are constants.

(c) ω is a random variable that is uniformly distributed over the interval $[\omega_0 - \Delta, \omega_0 + \Delta]$ and both A and ϕ are constants.

Solution _____

(a) When ω and ϕ are constants, then

$$r_x(k,l) = E\{x(k)x(l)\} = E\{A\cos(k\omega + \phi)A\cos(l\omega + \phi)\} + \sigma_w^2\delta(k-l)$$

Thus,

$$\begin{aligned}
r_x(k,l) &= E\{A^2\}\cos(k\omega + \phi)\cos(k\omega + \phi) + \sigma_w^2\delta(k-l) \\
&= \sigma_A^2\cos(k\omega + \phi)\cos(l\omega + \phi) + \sigma_w^2\delta(k-l)
\end{aligned}$$

Note that since $r_x(k,l)$ does not depend on the difference $k - l$, then $x(n)$ is not wide-sense stationary, and the power spectrum is not defined for this process.

(b) When A and ω are constants and ϕ is a random variable that is uniformly distributed over the interval $[-\pi, \pi]$, then the autocorrelation is

$$\begin{aligned}
r_x(k,l) &= E\{A^2\cos(k\omega + \phi)\cos(l\omega + \phi)\} + \sigma_w^2\delta(k-l) \\
&= \tfrac{1}{2}A^2 E\{\cos[(k+l)\omega + 2\phi]\} + \tfrac{1}{2}A^2 E\{\cos(k-l)\omega\} + \sigma_w^2\delta(k-l)
\end{aligned}$$

However, since $E\{\cos[(k+l)\omega + 2\phi]\} = 0$, then the autocorrelation is

$$r_x(k,l) = \frac{1}{2}A^2\cos(k-l)\omega + \sigma_w^2\delta(k-l)$$

Therefore, $r_x(k,l)$ depends on the difference $(k-l)$, and the process is WSS. The power spectrum is

$$P_x(e^{j\omega}) = \frac{\pi A^2}{2}\delta(\omega - \omega) + \frac{\pi A^2}{2}\delta(\omega + \omega) + \sigma_w^2$$

(c) As in parts (a) and (b), the autocorrelation of the process $x(n)$ is

$$r_x(k,l) = E\{x(k)x(l)\} = E\{A\cos(k\omega + \phi)A\cos(l\omega + \phi)\} + \sigma_w^2\delta(k-l)$$

In this case, however, ω is a random variable, and the expectation of the product of the cosines is

$$E\{A\cos(k\omega + \phi)A\cos(l\omega + \phi)\} = A^2 E\left\{\tfrac{1}{2}\cos[(k-l)\omega] + \tfrac{1}{2}\cos[(k+l)\omega + 2\phi]\right\}$$

Since ω is uniformly distributed over the interval $[\omega_0 - \Delta, \omega_0 + \Delta]$, the expectation of the first term is

$$
\begin{aligned}
E\big\{\cos\big[(k-l)\omega\big]\big\} &= \frac{1}{2\Delta}\int_{\omega_0-\Delta}^{\omega+\Delta}\cos\big[(k-l)\omega\big]\,d\omega \\
&= \frac{1}{2\Delta(k-l)}\Big[\sin\big[(k-l)(\omega_0+\Delta)\big] - \sin\big[(k-l)(\omega_0-\Delta)\big]\Big]
\end{aligned}
$$

With ϕ a constant, the expectation of the second term is

$$
E\big\{\cos\big[(k+l)\omega + 2\phi\big]\big\} = \frac{1}{2\Delta(k+l)}\Big[\sin\big[(k+l)(\omega_0+\Delta+2\phi)\big] - \sin\big[(k+l)(\omega_0-\Delta+2\phi)\big]\Big]
$$

Therefore, $x(n)$ is *not* WSS. However, if ϕ is a random variable that is uniformly distributed over the interval $[-\pi, \pi]$, then this second expectation is zero, and the autocorrelation becomes

$$
\begin{aligned}
r_x(k,l) &= \frac{A^2}{4\Delta(k-l)}\Big\{\sin\big[(k-l)(\omega_0+\Delta)\big] - \sin\big[(k-l)(\omega_0-\Delta)\big]\Big\} \\
&= \frac{A^2}{2\Delta(k-l)}\sin\big[(k-l)\Delta\big]\cos\big[(k-l)\omega_0\big]
\end{aligned}
$$

and the process is WSS. With an autocorrelation sequence given by

$$
r_x(k) = \frac{\pi A^2}{2\Delta}\frac{\sin k\Delta}{\pi k}\cos k\omega_0
$$

using the DTFT pair

$$
r_x(k) = \frac{\sin k\Delta}{\pi k} \Longleftrightarrow P_x(e^{j\omega}) = \begin{cases} 1 & ; \quad |\omega| \le \Delta \\ 0 & ; \quad \text{else} \end{cases}
$$

it follows that the power spectrum of $x(n)$ is

$$
P_x(e^{j\omega}) = \begin{cases} \dfrac{\pi A^2}{4\Delta} & ; \quad \omega_0 - \Delta \le |\omega| \le \omega_0 + \Delta \\ 0 & ; \quad \text{else} \end{cases}
$$

3.9 Determine whether or not each of the following are valid autocorrelation matrices. If they are not, explain why not.

(a) $\mathbf{R}_1 = \begin{bmatrix} 4 & 1 & 1 \\ -1 & 4 & 1 \\ -1 & -1 & 4 \end{bmatrix}$

(b) $\mathbf{R}_2 = \begin{bmatrix} 2 & 2 \\ 2 & 2 \end{bmatrix}$

(c) $\mathbf{R}_3 = \begin{bmatrix} 1 & 1+j \\ 1-j & 1 \end{bmatrix}$

(d) $\mathbf{R}_4 = \begin{bmatrix} 3 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 3 \end{bmatrix}$

(e) $\mathbf{R}_5 = \begin{bmatrix} 2 & j & 1 \\ -j & 4j & -j \\ 1 & j & 2 \end{bmatrix}$

Solution _____

(a) Since $\mathbf{R}_1$ is not symmetric, it is *not* a valid autocorrelation matrix.

(b) Since $\mathbf{R}_2$ is symmetric and non-negative definite, this *is* a valid autocorrelation matrix.

(c) Although $\mathbf{R}_3$ is Hermitian, note that its determinant is negative,

$$\det \mathbf{R}_3 = 1 - (1+j)(1-j) = -1$$

Therefore, $\mathbf{R}_3$ is not non-negative definite and, therefore, it is *not* a valid autocorrelation matrix.

(d) $\mathbf{R}_4$ *is* a valid autocorrelation matrix since it is symmetric and non-negative definite.

(e) The entries along the diagonal of an autocorrelation matrix must be real-valued (this follows from the Hermitian property, and the fact that the ith entry along the diagonal is equal to $E\{|x(i)|^2\}$ which is real). Since the middle element is imaginary, this is *not* a valid autocorrelation matrix.

3.10 The input to a linear shift-invariant filter with unit sample response

$$h(n) = \delta(n) + \tfrac{1}{2}\delta(n-1) + \tfrac{1}{4}\delta(n-2)$$

is a zero mean wide-sense stationary process with autocorrelation

$$r_x(k) = \left(\tfrac{1}{2}\right)^{|k|}$$

(a) What is the variance of the output process?

(b) Find the autocorrelation of the output process, $r_y(k)$, for all k.

Solution _____

Before we find the variance, let's find the autocorrelation. With

$$r_y(k) = r_x(k) * h(k) * h(-k) = r_x(k) * \left[\tfrac{21}{16}\delta(k) + \tfrac{5}{8}\delta(k-1) + \tfrac{5}{8}\delta(k+1) + \tfrac{1}{4}\delta(k-2) + \tfrac{1}{4}\delta(k-2)\right]$$

it follows that

$$r_y(k) = \tfrac{21}{16}(\tfrac{1}{2})^{|k|} + \tfrac{5}{8}(\tfrac{1}{2})^{|k-1|} + \tfrac{5}{8}(\tfrac{1}{2})^{|k+1|} + \tfrac{1}{4}(\tfrac{1}{2})^{|k-2|} + \tfrac{1}{4}(\tfrac{1}{2})^{|k+2|}$$

Finally, since $x(n)$ has zero mean, the variance is

$$\sigma_y^2 = r_y(0) = \left(\tfrac{21}{16} + \tfrac{5}{16} + \tfrac{5}{16} + \tfrac{1}{16} + \tfrac{1}{16}\right) = \tfrac{33}{16}$$

3.11 Consider a first-order AR process that is generated by the difference equation

$$y(n) = ay(n-1) + w(n)$$

where $|a| < 1$ and $w(n)$ is a zero mean white noise random process with variance σ_w^2.

(a) Find the unit sample response of the filter that generates $y(n)$ from $w(n)$.

(b) Find the autocorrelation of $y(n)$.

(c) Find the power spectrum of $y(n)$.

Solution _____

(a) The process $y(n)$ is generated by filtering white noise with a first-order filter that has a system function given by

$$H(z) = \frac{1}{1 - az^{-1}}$$

Thus,

$$h(n) = a^n u(n)$$

(b) Since the autocorrelation sequence for $w(n)$ is $r_w(k) = \sigma_w^2 \delta(k)$, then

$$r_y(k) = \sigma_w^2 \delta(k) * h(k) * h(-k) = \sigma_w^2 h(k) * h(-k) = \frac{\sigma_w^2}{1 - a^2} a^{|k|}$$

(c) The power spectrum of $y(n)$ is

$$P_y(e^{j\omega}) = \sigma_w^2 \, |H(e^{j\omega})|^2 = \frac{\sigma_w^2}{1 + a^2 - 2a \cos \omega}$$

3.12 Consider an MA(q) process that is generated by the difference equation

$$y(n) = \sum_{k=0}^{q} b(k)w(n-k)$$

where $w(n)$ is zero mean white noise with variance σ_w^2.

(a) Find the unit sample response of the filter that generates $y(n)$ from $w(n)$.

(b) Find the autocorrelation of $y(n)$.

(c) Find the power spectrum of $y(n)$.

Solution _____

(a) The process $y(n)$ is generated by filtering white noise with an FIR filter that has a system function given by

$$H(z) = \sum_{k=0}^{q} b(k)z^{-k}$$

Thus, the unit sample response is

$$h(n) = \sum_{k=0}^{q} b(k)\delta(n-k)$$

(b) The autocorrelation sequence for $y(n)$ is

$$r_y(k) = \sigma_w^2 \delta(k) * h(k) * h(-k) = \sigma_w^2 \sum_{l=0}^{q-|k|} b(l)b(|k|+l)$$

(c) The power spectrum of $y(n)$ is

$$P_y(e^{j\omega}) = \sigma_w^2 B(e^{j\omega})B(e^{-j\omega})$$

3.13 Suppose we are given a zero-mean process $x(n)$ with autocorrelation

$$r_x(k) = 10 \left(\tfrac{1}{2}\right)^{|k|} + 3 \left(\tfrac{1}{2}\right)^{|k-1|} + 3 \left(\tfrac{1}{2}\right)^{|k+1|}$$

(a) Find a filter which, when driven by unit variance white noise, will yield a random process with this autocorrelation.

(b) Find a stable and causal filter which, when excited by $x(n)$, will produce zero mean, unit variance, white noise.

Solution _____

(a) The power spectrum of $x(n)$ is

$$P_x(z) = \frac{3/4}{(1 - \tfrac{1}{2}z^{-1})(1 - \tfrac{1}{2}z)} \left[10 + 3z^{-1} + 3z\right] = \tfrac{3}{4} \frac{(1 + 3z^{-1})(1 + 3z)}{(1 - \tfrac{1}{2}z^{-1})(1 - \tfrac{1}{2}z)}$$

Therefore, if

$$H(z) = \frac{\sqrt{3}}{2} \frac{1 + 3z^{-1}}{1 - \tfrac{1}{2}z^{-1}} \quad ; \quad |z| > \tfrac{1}{2}$$

then the response of this filter to unit variance white noise will be a random process with the given autocorrelations.

(b) Consider the filter having a system function

$$G(z) = \frac{2}{3\sqrt{3}} \frac{1 - \tfrac{1}{2}z^{-1}}{1 + \tfrac{1}{3}z^{-1}} \quad ; \quad |z| > \tfrac{1}{3}$$

Clearly this filter is stable and causal. Furthermore, if we filter $x(n)$ with $g(n)$ then the power spectrum of the filtered signal will be

$$P_y(z) = G(z)G(z^{-1})P_x(z) = 1$$

Therefore, $g(n)$ is the *whitening filter* that will produce unit variance white noise from $x(n)$.

Chapter 3

3.14 For each of the following, determine whether or not the random process is

1. Wide-sense stationary.
2. Mean ergodic.

(a) $x(n) = A$ where A is a random variable with probability density function $f_A(\alpha)$.

(b) $x(n) = A\cos(n\omega_0)$ where A is a Gaussian random variable with mean m_A and variance σ_A^2.

(c) $x(n) = A\cos(n\omega_0+\phi)$ where ϕ is a random variable that is uniformly distributed between $-\pi$ and π.

(d) $x(n) = A\cos(n\omega_0) + B\sin(n\omega_0)$ where A and B are uncorrelated zero mean random variables with variance σ^2.

(e) A Bernoulli process with $\Pr\{x(n) = 1\} = p$ and $\Pr\{x(n) = -1\} = 1 - p$.

(f) $y(n) = x(n) - x(n-1)$ where $x(n)$ is the Bernoulli process defined in part (e).

Solution _____

(a) We are given a process $x(n) = A$, where A is a random variable with probability density function $f_A(\alpha)$. To check for wide-sense stationarity we need to compute the mean and autocorrelation of $x(n)$. The mean of this process is

$$m_x(n) = E\{x(n)\} = E\{A\}$$

which is a constant. The autocorrelation is

$$r_x(k,l) = E\{x(k)x(l)\} = E\{A^2\}$$

which is also a constant. Therefore, $x(n)$ is WSS.
To check for ergodicity in the mean, note that

$$c_x(k) = r_x(k) - m_x^2 = E\{A^2\} - E^2\{A\}$$

which is a constant. Therefore, $x(n)$ is ergodic in the mean only if the variance of A is zero, $c_x(k) = 0$.

(b) With $x(n) = A\cos(n\omega_0)$, note that the mean of the process is

$$m_x(n) = E\{x(n)\} = E\{A\cos(n\omega_0)\} = E\{A\}\cos(n\omega_0) = m_A\cos(n\omega_0)$$

which depends on n. Thus, $x(n)$ is not WSS and, therefore, not ergodic in the mean.

(c) For $x(n) = A\cos(n\omega_0 + \phi)$ with ϕ a random variable that is uniformly distributed between $-\pi$ and π, the mean of $x(n)$ is

$$m_x(n) = E\{x(n)\} = E\{A\cos(n\omega_0 + \phi)\} = \frac{A}{2\pi}\int_{-\pi}^{\pi}\cos(n\omega_0 + \phi)d\phi = 0$$

which is a constant. For the autocorrelation we have

$$
\begin{aligned}
r_x(k,l) &= E\big\{A\cos(k\omega_0+\phi)A\cos(l\omega_0+\phi)\big\} \\
&= \tfrac{1}{2}A^2 E\Big\{\cos([k-l]\omega_0)+\cos([k+l]\omega_0+2\phi)\Big\} \\
&= \tfrac{1}{2}A^2\cos(k-l)\omega_0
\end{aligned}
$$

which is a function of $(k-l)$. Therefore, $x(n)$ is WSS.
To check for ergodicity in the mean, note that

$$
\begin{aligned}
\frac{1}{N}\sum_{k=0}^{N-1}c_x(k) &= \frac{1}{2}\frac{A^2}{N}\sum_{k=0}^{N-1}\cos k\omega_0 \\
&= \frac{1}{2}\frac{A^2}{N}\sum_{k=0}^{N-1}\big\{\tfrac{1}{2}e^{jk\omega_0}+\tfrac{1}{2}e^{-jk\omega_0}\big\} \\
&= \frac{1}{4}\frac{A^2}{N}\Big\{\frac{1-e^{jN\omega_0}}{1-e^{j\omega_0}}+\frac{1-e^{-jN\omega_0}}{1-e^{-j\omega_0}}\Big\} \\
&= \frac{A^2}{2N}\frac{\sin(N\omega_0/2)}{\sin(\omega_0/2)}\cos\Big(\frac{N-1}{2}\omega_0\Big)
\end{aligned}
$$

which goes to zero as $N\to\infty$, provided that $\omega_0=0$. If $\omega_0=0$, then $x(n)=A\cos\phi$ and $c_x(k)=A^2/2$. In this case, $x(n)$ is not ergodic in the mean. Therefore, $x(n)$ is ergodic in the mean only if $\omega_0\neq 0$.

(d) The mean of this process is

$$
E\big\{x(n)\big\}=E\big\{A\big\}\cos(n\omega_0)+E\big\{B\big\}\sin(n\omega_0)
$$

Since $E\{A\}=E\{B\}=0$, then $E\{x(n)\}=0$. For the autocorrelation we have

$$
\begin{aligned}
r_x(k,l) &= E\big\{x(k)x(l)\big\}=E\big\{[A\cos(\omega_0 k)+B\sin(\omega_0 k)][A\cos(\omega_0 l)+B\sin(\omega_0 l)]\big\} \\
&= E\big\{A^2\big\}\cos(k\omega_0)\cos(l\omega_0)+E\big\{B^2\big\}\sin(k\omega_0)\sin(l\omega_0) \\
&\quad +E\big\{AB\big\}[\cos(k\omega_0)\sin(l\omega_0)+\sin(k\omega_0)\cos(l\omega_0)]
\end{aligned}
$$

Since A and B are uncorrelated and have zero mean, then $E\{AB\}=0$ and $E\{A^2\}=E\{B^2\}=\sigma^2$. Therefore, we have

$$
r_x(k,l)=\sigma^2[\cos(k\omega_0)\cos(l\omega_0)+\sin(k\omega_0)\sin(l\omega_0)]=\sigma^2\cos(k-l)\omega_0
$$

Since the mean is a constant and the correlation function $r_x(k,l)$ depends only the difference, $k-l$, then $x(n)$ is a wide-sense stationary process.
As in part (c), $x(n)$ is ergodic in the mean only if $\omega_0=0$.

(e) With $x(n)$ a Bernoulli process with $\Pr\{x(n)=1\}=p$ and $\Pr\{x(n)=-1\}=1-p$, the mean of $x(n)$ is

$$
m_x(n)=E\big\{x(n)\big\}=p-(1-p)=2p-1
$$

and the autocorrelation is

$$
r_x(k,l)=E\big\{x(k)x(l)\big\}=\begin{cases}E\big\{x^2(k)\big\} & ; \quad k=l \\ E\big\{x(k)x(l)\big\} & ; \quad k\neq l\end{cases}
$$

Therefore,

$$r_x(k,l) = \begin{cases} p + (1-p) = 1 & ; \quad k = l \\ (1-2p)^2 & ; \quad k \neq l \end{cases}$$

or,

$$r_x(k,l) = 4p(1-p)\delta(k-l) + (1-2p)^2$$

which is a function of $k - l$. Therefore, $x(n)$ is WSS. With

$$c_x(k) = r_x(k) - m_x^2 = 4p(1-p)\delta(k)$$

it follows that

$$\frac{1}{N}\sum_{k=0}^{N-1} c_x(k) = \frac{1}{N}4p(1-p) \longrightarrow 0 \text{ as } N \to \infty$$

Thus, $x(n)$ is ergodic in the mean.

(f) With $y(n) = x(n) - x(n-1)$ where $x(n)$ is a Bernoulli process, wide-sense stationarity may be easily checked using the direct approach taken in parts (a)-(e) of this problem. However, it is easier to note that since $y(n)$ is the response of a linear shift-invariant system to an input that is a WSS process, then $y(n)$ will be WSS.

For ergodicity in the mean, note that since

$$c_x(k) = 4p(1-p)\delta(k)$$

then

$$c_y(k) = c_x(k) * h(k) * h(-k) = 4p(1-p)\big[2\delta(k) - \delta(k-1) - \delta(k+1)\big]$$

and, clearly,

$$\frac{1}{N}\sum_{k=0}^{N-1} c_x(k) \to 0 \text{ as } N \to \infty$$

3.15 Determine which of the following correspond to a valid autocorrelation sequence for a WSS random process. For those that are not valid, state why not. For those that are valid, describe a way for generating a process with the given autocorrelation.

(a) $r_x(k) = \delta(k - 1) + \delta(k + 1)$

(b) $r_x(k) = 3\delta(k) + 2\delta(k - 1) + 2\delta(k + 1)$

(c) $r_x(k) = \exp(jk\pi/4)$

(d) $r_x(k) = \begin{cases} 1 & ; \quad |k| < N \\ 0 & ; \quad \text{else} \end{cases}$

(e) $r_x(k) = \begin{cases} \dfrac{N - |k|}{N} & ; \quad |k| < N \\ 0 & ; \quad \text{else} \end{cases}$

(f) $r_x(k) = 2^{-k^2}$

Solution

(a) This autocorrelation sequence is not valid since we must have $r_x(0) \geq |r_x(1)|$.

(b) This autocorrelation sequence is not valid since the power spectrum is not non-negative

$$P_x(e^{j\omega}) = 3 + 2e^{-j\omega} + 2e^{j\omega} = 3 + 4\cos\omega$$

(c) This autocorrelation sequence is valid, and corresponds to a harmonic process. Given a random variable ϕ that is uniformly distributed between $-\pi$ and π, this process may be generated as follows,

$$x(n) = e^{j(\frac{n\pi}{4} + \phi)}$$

(d) This autocorrelation sequence is not valid since the power spectrum

$$P_x(e^{j\omega}) = \frac{\sin(N - \frac{1}{2})\omega}{\sin\omega/2}$$

is not non-negative.

(e) This autocorrelation sequence is valid since $r_x(k)$ is symmetric, and its discrete-time Fourier transform is non-negative for all ω. This process may be generated by filtering unit variance white noise with the FIR filter that has a unit sample response given by

$$h(n) = \begin{cases} 1 & ; \quad 0 \leq n < N \\ 0 & ; \quad \text{else} \end{cases}$$

(f) The sequence

$$r_x(k) = 2^{-k^2} = e^{-(\ln 2)k^2}$$

is a sampled Gaussian pulse. The DTFT of $r_x(k)$ is an aliased Gaussian, which is positive for all ω. Since $r_x(k)$ is symmetric and $P_x(e^{j\omega}) \geq 0$, this represents a valid autocorrelation sequence. This process may be generated by filtering white noise with a linear shift-invariant system that has a Gaussian shaped unit sample response.

3.16 Show that the cross-correlation, $r_{xy}(k)$, between two jointly wide-sense stationary processes $x(n)$ and $y(n)$ satisfies the following inequalities,

(a) $|r_{xy}(k)| \leq \left[r_x(0) r_y(0) \right]^{1/2}$

(b) $|r_{xy}(k)| \leq \frac{1}{2} \left[r_x(0) + r_y(0) \right]$

Solution _____

(a) Note that for any constant a,

$$E\left\{ \left[ax(n+k) - y(n) \right]^2 \right\} \geq 0$$

Expanding the square we have

$$a^2 E\{x^2(n+k)\} - 2aE\{x(n+k)y(n)\} + E\{y^2(n)\} \geq 0$$

This is a quadratic equation in a, and is non-negative. Therefore, its discriminant must be non-positive,

$$4E^2\{x(n+k)y(n)\} \leq 4E\{x^2(n+k)\}E\{y^2(n)\}$$

or,

$$r_{xy}^2(k) \leq r_x(0)r_y(0)$$

Taking the square root, the result follows.

(b) To establish this inequality, note that

$$E\{|x(n+k) \pm y(n)|^2\} \geq 0$$

Expanding the square it follows that

$$r_x(0) \pm 2r_{xy}(k) + r_y(0) \geq 0$$

Therefore,

$$\pm r_{xy}(k) \leq \frac{1}{2}\left[r_x(0) + r_y(0)\right]$$

and

$$|r_{xy}(k)| \leq \frac{1}{2}\left[r_x(0) + r_y(0)\right]$$

3.17 Given a wide-sense stationary random process $x(n)$, we would like to design a "linear predictor" that will predict the value of $x(n+1)$ using a linear combination of $x(n)$ and $x(n-1)$. Thus, our predictor for $x(n+1)$ is of the form

$$\widehat{x}(n+1) = ax(n) + bx(n-1)$$

where a and b are constants. Assume that the process has zero mean

$$E\{x(n)\} = 0$$

and that we want to minimize the mean-square error

$$\xi = E\left\{[x(n+1) - \widehat{x}(n+1)]^2\right\}$$

   (a) With $r_x(k)$ the autocorrelation of $x(n)$, determine the optimum predictor of $x(n)$ by finding the values of a and b that minimize the mean-square error.

   (b) What is the minimum mean-square error of the predictor? Express your answer in terms of the autocorrelation $r_x(k)$.

   (c) If $x(n+1)$ is uncorrelated with $x(n)$, what form does your predictor take?

   (d) If $x(n+1)$ is uncorrelated with both $x(n)$ and $x(n-1)$, what form does your predictor take?

Solution _____

   (a) The mean-square error that we want to minimize is

$$\xi = E\left\{[x(n+1) - \widehat{x}(n+1)]^2\right\} = E\left\{x^2(n+1) - 2x(n+1)\widehat{x}(n+1) + \widehat{x}^2(n+1)\right\}$$

Since the estimate of $x(n+1)$ is

$$\widehat{x}(n+1) = ax(n) + bx(n-1)$$

then setting the derivative of ξ with respect to a and b equal to zero we have

$$\frac{\partial \xi}{\partial a} = -2E\{x(n+1)x(n)\} + E\{2\widehat{x}(n+1)x(n)\} = 0$$

$$\frac{\partial \xi}{\partial b} = -2E\{x(n+1)x(n-1)\} + E\{2\widehat{x}(n+1)x(n-1)\} = 0$$

Dividing by 2 and substituting for $\widehat{x}(n+1)$ gives

$$aE\{x^2(n)\} + bE\{x(n)x(n-1)\} = E\{x(n+1)x(n)\}$$

$$aE\{x(n)x(n-1)\} + bE\{x^2(n-1)\} = E\{x(n+1)x(n-1)\}$$

Putting these equations in matrix form we have

$$\begin{bmatrix} r_x(0) & r_x(1) \\ r_x(1) & r_x(0) \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} r_x(1) \\ r_x(2) \end{bmatrix}$$

Solving for a and b we find

$$\begin{bmatrix} a \\ b \end{bmatrix} = \frac{1}{r_x^2(0) - r_x^2(1)} \begin{bmatrix} r_x(0)r_x(1) - r_x(1)r_x(2) \\ r_x(0)r_x(2) - r_x^2(1) \end{bmatrix}$$

(b) For the minimum mean-square error we have

$$
\begin{aligned}
\xi &= E\Big\{[x(n+1) - \widehat{x}(n+1)]^2\Big\} = E\big\{[x(n+1) - \widehat{x}(n+1)]\,[x(n+1) - \widehat{x}(n+1)]\big\} \\
&= E\big\{[x(n+1) - \widehat{x}(n+1)]\,x(n+1)\big\} - E\big\{[x(n+1) - \widehat{x}(n+1)]\,\widehat{x}(n+1)\big\}
\end{aligned}
$$

Note that for values of a and b that minimize the mean-square error, the derivatives of ξ with respect to a and b are equal to zero, which implies that the second term in the equation above is equal to zero. Therefore, the minimum mean-square error is

$$
\xi = E\big\{[x(n+1) - \widehat{x}(n+1)]\,x(n+1)\big\} = E\big\{[x(n+1) - ax(n) - bx(n-1)]\,x(n+1)\big\}
$$

or,

$$
\begin{aligned}
\xi &= r_x(0) - a r_x(1) - b r_x(2) \\
&= r_x(0) - r_x(1)\frac{r_x(0)r_x(1) - r_x(1)r_x(2)}{r_x^2(0) - r_x^2(1)} - r_x(2)\frac{r_x(0)r_x(2) - r_x^2(1)}{r_x^2(0) - r_x^2(1)}
\end{aligned}
$$

(c) If $x(n+1)$ and $x(n)$ are uncorrelated, then $r_x(1) = 0$, and the values for a and b become

$$
a = 0 \quad ; \quad b = r_x(2)/r_x(0)
$$

In this case, the linear predictor is

$$
\widehat{x}(n+1) = \frac{r_x(2)}{r_x(0)}x(n-1)
$$

(d) If $x(n+1)$ is uncorrelated with both $x(n)$ and $x(n-1)$, then the values for a and b are

$$
a = b = 0
$$

and the linear predictor is

$$
\widehat{x}(n+1) = 0
$$

which is equal to the expected value of $x(n+1)$,

$$
\widehat{x}(n+1) = E\left\{x(n+1)\right\}
$$

3.18 *True* or *False:* If $x(n)$ is a WSS process and $y(n)$ is the process that is formed by filtering $x(n)$ with a stable, linear shift-invariant filter $h(n)$, then

$$\sigma_y^2 = \sigma_x^2 \sum_{n=-\infty}^{\infty} |h(n)|^2$$

where σ_x^2 and σ_y^2 are the variances of the processes $x(n)$ and $y(n)$, respectively.

Solution _____

If a WSS process $x(n)$ is filtered with a filter that has a unit sample response $h(n)$, then the autocorrelation of the output process is

$$r_y(k) = r_x(k) * h(k) * h^*(-k)$$

Assuming that $x(n)$ has zero-mean, then

$$\sigma_y^2 = r_y(0) \qquad \text{and} \qquad \sigma_x^2 = r_x(0)$$

Therefore, the question is whether or not the following relationship is true:

$$r_y(0) = r_x(0) \sum_{n=-\infty}^{\infty} |h(n)|^2$$

If this is true, then σ_y^2 depends *only* on $r_x(0)$, and not on $r_x(k)$ for $k \neq 0$. Clearly, this is not true unless $r_x(k) = 0$ for $|k| > 0$, i.e., if $x(n)$ is white noise. Therefore, the statement is *False*.

3.19 Show that a sufficient condition for a wide-sense stationary process to be ergodic in the mean is that the autocovariance be absolutely summable,

$$\sum_{k=-\infty}^{\infty} |c_x(k)| < \infty$$

Solution ───

Beginning with the variance of the sample mean,

$$\text{Var}\{\widehat{m}_x(N)\} = \frac{1}{N} \sum_{k=-N+1}^{N-1} \left(1 - \frac{|k|}{N}\right) c_x(k) \leq \frac{1}{N} \sum_{k=-N+1}^{N-1} |c_x(k)|$$

note that if

$$\sum_{k=-\infty}^{\infty} |c_x(k)| < \infty$$

then

$$\lim_{N\to\infty} \text{Var}\{\widehat{m}_x(N)\} = \lim_{N\to\infty} \frac{1}{N} \sum_{k=-N+1}^{N-1} |c_x(k)| = 0$$

and the process is ergodic in the mean.

───

3.20 For each of the following, determine whether the statements are *True* or *False*.

    (a) All wide-sense stationary moving average processes are ergodic in the mean.

    (b) All wide-sense stationary autoregressive processes are ergodic in the mean.

Solution _____

    (a) This statement is true. Recall that a WSS moving average process has an autocovariance that is finite in length, $c_x(k) = 0$ for all $|k| > k_0$ for some k_0. Therefore, if we let

$$C = \sum_{k=0}^{k_0} c_x(k)$$

then,

$$\lim_{N \to \infty} \frac{1}{N} \sum_{k=0}^{N-1} c_x(k) = \lim_{N \to \infty} \frac{C}{N} = 0$$

and ergodicity is established.

    (b) This statement is true and may be shown as follows. An autoregressive process is formed by filtering finite variance white noise $w(n)$ with a stable, causal all-pole filter $h(n)$,

$$x(n) = h(n) * w(n)$$

Assuming, without any loss in generality, that $w(n)$ has zero mean, the covariance of $x(n)$ is

$$c_x(k) = \sigma_w^2 \, h(k) * h(-k)$$

where $\sigma_w^2 < \infty$ is the variance of $w(n)$. The condition that $h(n)$ is stable guarantees that $c_x(k)$ is absolutely summable and, therefore, $x(n)$ is ergodic in the mean.

3.21 Let $x(n)$ be a real WSS Gaussian random process with autocovariance function $c_x(k)$. Show that $x(n)$ will be *correlation ergodic* if and only if

$$\lim_{N \to \infty} \frac{1}{N} \sum_{k=0}^{N-1} c_x^2(k) = 0$$

Hint: Use the *moment factoring theorem* for real Gaussian random variables which states that

$$E\{x_1 x_2 x_3 x_4\} = E\{x_1 x_2\} E\{x_3 x_4\} + E\{x_1 x_3\} E\{x_2 x_4\} + E\{x_1 x_4\} E\{x_2 x_3\}$$

Solution _____

We are given a WSS Gaussian random process, $x(n)$, that is ergodic in the mean. For convenience, let us assume that the process has zero mean. In this problem we will be using the moment theorem for Gaussian random variables. This theorem states that if $x(m)$, $x(n)$, $x(k)$, and $x(l)$, are zero mean jointly distributed Gaussian random variables then the fourth-order moment $E\{x(m)x(n)x(k)x(l)\}$ is

$$
\begin{aligned}
E\{x(m)x(n)x(k)x(l)\} &= E\{x(m)x(n)\}E\{x(k)x(l)\} + E\{x(m)x(k)\}E\{x(n)x(l)\} \\
&\quad + E\{x(m)x(l)\}E\{x(n)x(k)\}
\end{aligned}
$$

Now, for a fixed value of k, let

$$y(n) = x(n+k)x(n)$$

Since

$$E\{y(n)\} = r_x(k)$$

it follows that $x(n)$ will be correlation ergodic

$$\lim_{N \to \infty} \frac{1}{N} \sum_{n=0}^{N-1} x(n+k)x(n) = r_x(k)$$

if and only if[1]

$$\lim_{N \to \infty} \frac{1}{2N+1} \sum_{l=-N+1}^{N-1} c_y(l) = 0 \tag{P3.21-1}$$

where

$$
\begin{aligned}
c_y(l) &= E\{y(m+l)y(m)\} - E\{y(m+l)\}E\{y(m)\} \\
&= E\{x(m+l+k)x(m+l)x(m+k)x(m)\} - r_x^2(k)
\end{aligned}
$$

Using the moment theorem, we have

$$
\begin{aligned}
c_y(l) &= E\{x(m+l+k)x(m+l)\}E\{x(m+k)x(m)\} + \\
&\quad E\{x(m+l+k)x(m+k)\}E\{x(m+l)x(m)\} + \\
&\quad E\{x(m+l+k)x(m)\}E\{x(m+k)x(m+l)\} - r_x^2(k) \\
&= r_x^2(l) + r_x(k+l)r_x(k-l)
\end{aligned}
$$

_____

[1] Note that we are assuming that $y(n)$ is wide-sense stationary. However, it may be shown that since $x(n)$ is a WSS Gaussian process then it is also strict sense stationary and, therefore, $y(n)$ is WSS.

Therefore, $x(n)$ will be correlation ergodic if and only if

$$\lim_{N \to \infty} \frac{1}{N} \sum_{l=0}^{N-1} \left[r_x^2(l) + r_x(k+l) r_x(k-l) \right] = 0 \qquad \text{(P3.21-1)}$$

What we would like to establish is the equivalence of Eq. (P3.21-1) and Eq. (P3.21-1). It is clear that if Eq. (P3.21-1) is true, then Eq. (P3.21-1) holds. We may see this by setting $k = 0$ in Eq. (P3.21-1). To establish the converse, we use the inequality

$$2|r_x(k+l) r_x(k-l)| \le r_x^2(k+l) + r_x^2(k-l)$$

Therefore,

$$\left| \frac{1}{N} \sum_{l=0}^{N-1} r_x(k+l) r_x(k-l) \right| \le \frac{1}{2N} \left| \sum_{l=0}^{N-1} r_x^2(k+l) + r_x^2(k-l) \right|$$

$$= \frac{1}{2N} \left| \sum_{l=k}^{N-1+k} r_x^2(k) + \sum_{l=-k}^{N-1-k} r_x^2(k) \right|$$

$$\le \frac{1}{2N} \left| \sum_{l=0}^{N-1} r_x^2(k) + \frac{k r_x^2(0)}{N} \right|$$

where, in the last inequality, we used the fact that

$$|r_x(k)| \le r_x(0)$$

for all k, Thus, as $N \to \infty$, the right-side goes to zero and we have established the equivalence of Eq. (P3.21-1) and Eq. (P3.21-1).

3.22 Ergodicity in the mean depends on the asymptotic behavior of the autocovariance of a process, $c_x(k)$. The asymptotic behavior of $c_x(k)$, however, is related to the behavior of the power spectrum $P_x(e^{j\omega})$ at $\omega = 0$. Show that $x(n)$ is ergodic in the mean if and only if $P_x(e^{j\omega})$ is continuous at the origin, $\omega = 0$. *Hint:* Express Eq. (3.65) as a limit, as $N \to \infty$, of the convolution of $c_x(k)$ with a pulse

$$p_N(n) = \begin{cases} 1/N & ; \quad 0 \le n < N \\ 0 & ; \quad \text{otherwise} \end{cases}$$

with the convolution being evaluated at $k = N$.

Solution _____

A necessary and sufficient condition for a process WSS process $x(n)$ with autocovariance $c_x(k)$ to be ergodic in the mean is

$$\lim_{N \to \infty} \frac{1}{N} \sum_{k=0}^{N-1} c_x(k) = 0$$

To show that $x(n)$ is ergodic in the mean if and only if $P_x(e^{j\omega})$ is continuous at the origin, let

$$s_N(k) = c_x(k) * p_N(k)$$

where

$$p_N(n) = \begin{cases} 1/N & ; \quad 0 \le n < N \\ 0 & ; \quad \text{otherwise} \end{cases}$$

Note that

$$s_N(N) = \frac{1}{N} \sum_{k=0}^{N-1} c_x(k)$$

Since $s_N(k)$ is the convolution of $c_x(k)$ with $p_N(k)$, then

$$S_N(e^{j\omega}) = P_x(e^{j\omega})P_N(e^{j\omega})$$

where

$$P_N(e^{j\omega}) = e^{-j(N-1)/2} \frac{\sin N\omega/2}{N \sin \omega/2}$$

Therefore,

$$s_N(N) = \frac{1}{2\pi} \int_{-\pi}^{\pi} S_N(e^{j\omega})e^{jN\omega}d\omega = \frac{1}{2\pi} \int_{-\pi}^{\pi} P_x(e^{j\omega})e^{-j(N-1)/2} \frac{\sin N\omega/2}{N \sin \omega/2} \, e^{jN\omega}d\omega$$

Note that the term multiplying the power spectrum $P_x(e^{j\omega})$ inside the integral is bounded by one in magnitude and, as $N \to \infty$, this term goes to zero for all $\omega \ne 0$, and for $\omega = 0$ this term is equal to one. Therefore,

$$\lim_{N \to \infty} s_N(N) = \frac{1}{2\pi} \left[P_x(e^{j\omega})|_{\omega=0^+} - P_x(e^{j\omega})|_{\omega=0^-} \right]$$

and it follows that $s_N(N)$ goes to zero as $N \to \infty$ if and only if $P_x(e^{j\omega})|_{\omega=0^+} = P_x(e^{j\omega})|_{\omega=0^-}$, i.e., $P_x(e^{j\omega})$ is continuous at the origin.

3.23 In Section 3.2.4 it was shown that for real-valued zero mean random variables the autocorrelation is bounded by one in magnitude

$$|\rho_{xy}| \le 1$$

Show that this bound also applies when x and y are complex random variables with nonzero mean. Determine what relationship must hold between x and y if

$$|\rho_{xy}| = 1$$

Solution _____

Without any loss of generality, we may assume zero mean for both x and y. For non-zero mean random variables, the following derivation is modified by replacing x with $x - m_x$ and y with $y - m_y$. What we want to show is that

$$\left| E\left\{ xy^* \right\} \right|^2 \le E\{|x|^2\} E\{|y|^2\}$$

Note that for any constant a,

$$E\{|x - ay|^2\} \ge 0$$

with equality if and only if $x = ay$ with probability one. Expanding the square we have

$$E\{|x - ay|^2\} = E\{|x|^2\} - aE\{x^*y\} - a^*E\{xy^*\} + |a|^2 E\{|y|^2\} \ge 0$$

Now let

$$a = \frac{E\{xy^*\}}{E\{|y|^2\}}$$

Then

$$E\{|x|^2\} - E\{x^*y\}\frac{E\{xy^*\}}{E\{|y|^2\}} - E\{xy^*\}\frac{E\{x^*y\}}{E\{|y|^2\}} + E\{|y|^2\}\frac{E\{x^*y\}E\{xy^*\}}{E^2\{|y|^2\}} \ge 0$$

Cancelling terms and simplifying this becomes

$$E\{|x|^2\} - \frac{|E\{xy^*\}|^2}{E\{|y|^2\}} \ge 0$$

and the inequality follows.

3.24 Let $P_x(e^{j\omega})$ be the power spectrum of a wide-sense stationary process $x(n)$ and let λ_k be the eigenvalues of the $M \times M$ autocorrelation matrix $\mathbf{R}_x$. *Szegö's theorem* states that if $g(\cdot)$ is a continuous function then

$$\lim_{M \to \infty} \frac{g(\lambda_1) + g(\lambda_2) + \cdots + g(\lambda_M)}{M} = \frac{1}{2\pi} \int_{-\pi}^{\pi} g[P_x(e^{j\omega})] d\omega$$

Using this theorem, show that

$$\lim_{M \to \infty} \left[\det \mathbf{R}_x \right]^{1/M} = \exp\left\{ \frac{1}{2\pi} \int_{-\pi}^{\pi} \ln\left[P_x(e^{j\omega}) \right] d\omega \right\}$$

Solution
The determinant of a matrix is equal to the product of its eigenvalues,

$$\det \mathbf{R}_x = \prod_{k=1}^{M} \lambda_k$$

Therefore

$$(\det \mathbf{R}_x)^{1/M} = \left(\prod_{k=1}^{M} \lambda_k \right)^{1/M}$$

Taking logarithms, we have

$$\ln (\det \mathbf{R}_x)^{1/M} = \frac{1}{M} \sum_{k=1}^{M} \ln \lambda_k$$

Using Szegö's theorem yields

$$\lim_{M \to \infty} \ln (\det \mathbf{R}_x)^{1/M} = \lim_{M \to \infty} \frac{1}{M} \sum_{k=1}^{M} \ln \lambda_k = \frac{1}{2\pi} \int_{-\pi}^{\pi} \ln\left[P_x(e^{j\omega}) \right] d\omega$$

and the result follows.

3.25 In some applications, the data collection process may be flawed so that there are either missing data values or outliers that should be discarded. Suppose that we are given N samples of a WSS process $x(n)$ with one value, $x(n_0)$, missing. Let $\mathbf{x}$ be the vector containing the given sample values,

$$\mathbf{x} = [x(0),\ x(1),\ \ldots, x(n_0 - 1),\ x(n_0 + 1),\ \ldots, x(N)]^T$$

(a) Let $\mathbf{R}_x$ be the autocorrelation matrix for the vector $\mathbf{x}$,

$$\mathbf{R}_x = E\{\mathbf{x}\mathbf{x}^H\}$$

Which of the following statements are true:

1. $\mathbf{R}_x$ is Toeplitz.
2. $\mathbf{R}_x$ is Hermitian.
3. $\mathbf{R}_x$ is positive semidefinite.

(b) Given the autocorrelation matrix for $\mathbf{x}$, is it possible to find the autocorrelation matrix for the vector

$$\mathbf{x} = \Big[x(0),\ x(1),\ \ldots, x(N)\Big]^T$$

that does not have $x(n_0)$ missing? If so, how would you find it? If not, explain why not.

Solution _____

(a) The matrix is *not Toeplitz*. This may be shown easily by example. If $\mathbf{x} = \big[x(0), x(2), x(3)\big]$, then

$$\mathbf{R}_x = E\{\mathbf{x}\mathbf{x}^H\} = E\left[\begin{array}{ccc} |x(0)|^2 & x(0)x^*(2) & x(0)x^*(3) \\ x(2)x^*(0) & |x(2)|^2 & x(2)x^*(3) \\ x(3)x^*(0) & x(3)x^*(2) & |x(3)|^2 \end{array}\right] = \left[\begin{array}{ccc} r_x(0) & r_x(2) & r_x(3) \\ r_x(2) & r_x(0) & r_x(1) \\ r_x(3) & r_x(1) & r_x(0) \end{array}\right]$$

which is clearly not Toeplitz. However, by definition, $\mathbf{R}_x$ is Hermitian,

$$\mathbf{R}_x^H = E\{\mathbf{x}\mathbf{x}^H\}^H = E\{\mathbf{x}\mathbf{x}^H\} = \mathbf{R}_x$$

Finally, $\mathbf{R}_x$ is non-negative definite, which may be shown as follows. Let $\mathbf{v}$ be any non-zero vector. Then,

$$\mathbf{v}^H \mathbf{R}_x \mathbf{v} = \mathbf{v}^H E\{\mathbf{x}\mathbf{x}^H\}\mathbf{v} = E\{\mathbf{v}^H \mathbf{x}\mathbf{x}^H \mathbf{v}\}$$

Therefore,

$$\mathbf{v}^H \mathbf{R}_x \mathbf{v} = E\{|\mathbf{v}^H \mathbf{x}|^2\} \geq 0$$

and $\mathbf{R}_x \geq 0$.

(b) There are several ways to find the autocorrelation matrix for the vector

$$\mathbf{x} = \Big[x(0),\ x(1),\ \ldots, x(N)\Big]^T$$

that does not have $x(n_0)$ missing. One way is as follows. Note that the first column of $\mathbf{R}_x$ that is formed from the vector that has $x(n_0)$ missing is as follows,

$$\Big[r_x(0), r_x(1), \cdots, r_x(n_0 - 1), r_x(n_0 + 1), \cdots, r_x(N)\Big]$$

Therefore, all that we need is the missing correlation $r_x(n_0)$. Note, however, that this term is found in the second column of row $(n_0 + 2)$ (see example in part (a) above). Thus, given $r_x(n_0)$ the Toeplitz matrix may then be formed.

3.26 The power spectrum of a wide-sense stationary process $x(n)$ is

$$P_x(e^{j\omega}) = \frac{25 - 24\cos\omega}{26 - 10\cos\omega}$$

Find the whitening filter $H(z)$ that produces unit variance white noise when the input is $x(n)$.

Solution _____

Expanding the cosines in the expression for the power spectrum we have

$$P_x(e^{j\omega}) = \frac{25 - 12e^{j\omega} - 12e^{-j\omega}}{26 - 5e^{j\omega} - 5e^{-j\omega}}$$

or, in terms of z, the power spectrum becomes

$$P_x(z) = \frac{25 - 12z - 12z^{-1}}{26 - 5z - 5z^{-1}} = \frac{(4 - 3z)(4 - 3z^{-1})}{(5 - z)(5 - z^{-1})} = G(z)G(z^{-1})$$

where

$$G(z) = \frac{4 - 3z}{5 - z} = \frac{3 - 4z^{-1}}{1 - 5z^{-1}}$$

Therefore, if

$$H(z) = \frac{1}{G(z)} = \frac{1 - 5z^{-1}}{3 - 4z^{-1}}$$

then

$$y(n) = h(n) * x(n)$$

will be unit variance white noise.

_____

3.27 We have seen that the autocorrelation matrix of a WSS process is positive semidefinite,

$$\mathbf{R}_x \geq 0$$

The spectral factorization theorem states that if $P_x(e^{j\omega})$ is continuous then the power spectrum may be factored as

$$P_x(e^{j\omega}) = \sigma_0^2 \, |Q(e^{j\omega})|^2$$

where $Q(e^{j\omega})$ corresponds to a causal and stable filter.

(a) Assuming that $\sigma_0^2 \neq 0$ and $Q(e^{j\omega})$ is nonzero for all ω, show that the autocorrelation matrix is positive definite.

(b) Given an example of a nontrivial process for which $\mathbf{R}_x$ is *not* positive definite.

Solution

(a) The positive definite property of $\mathbf{R}_x$ may be easily established with the help of the *Eigenvalue Extremal Property* which states that the eigenvalues of the autocorrelation matrix of a zero mean WSS random process are upper and lower bounded by the maximum and minimum values, respectively, of the power spectrum,

$$\min_\omega P_x(e^{j\omega}) \leq \lambda_i \leq \max_\omega P_x(e^{j\omega})$$

If $\sigma_0^2 \neq 0$ and $Q(e^{j\omega})$ is nonzero for all ω, then

$$P_x(e^{j\omega}) = \sigma_0^2 \, |Q(e^{j\omega})|^2 > 0$$

and, therefore,

$$0 < \min_\omega P_x(e^{j\omega}) \leq \lambda_i \quad ; \quad i = 1, 2, \ldots, n$$

Thus, it follows that $\mathbf{R}_x$ is positive definite.

(b) A non-trivial process for which $\mathbf{R}_x$ is *not* positive definite is the harmonic process

$$x(n) = e^{j(n\omega_0 + \phi)}$$

where ϕ is a random variable that is uniformly distributed over the interval $[-\pi, \pi]$. For this process, the autocorrelation matrix $\mathbf{R}_x$ has a rank of one.

SOLUTIONS TO CHAPTER 4
Signal Modeling

4.1 Find the Padé approximation of second order to a signal $x(n)$ that is given by

$$\mathbf{x} = [2,\ 1,\ 0,\ -1,\ 0,\ 1,\ 0,\ -1,\ 0,\ 1,\ \ldots]^T$$

i.e., $x(0) = 2$, $x(1) = 1$, $x(2) = 0$, and so on. In other words, using an approximation of the form

$$H(z) = \frac{b(0) + b(1)z^{-1} + b(2)z^{-2}}{1 + a(1)z^{-1} + a(2)z^{-2}}$$

find the coefficients $b(0)$, $b(1)$, $b(2)$, $a(1)$, and $a(2)$.

Solution

The Padé equations that must be solved are

$$\begin{bmatrix} 2 & 0 & 0 \\ 1 & 2 & 0 \\ 0 & 1 & 2 \\ -1 & 0 & 1 \\ 0 & -1 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ a(1) \\ a(2) \end{bmatrix} = \begin{bmatrix} b(0) \\ b(1) \\ b(2) \\ 0 \\ 0 \end{bmatrix}$$

The last two equations in this set are

$$\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \begin{bmatrix} a(1) \\ a(2) \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

Solving for $a(1)$ and $a(2)$ we have

$$a(1) = 0 \qquad \text{and} \qquad a(2) = 1$$

Using the first three equations, we may solve for $b(0), b(1)$, and $b(2)$ as follows:

$$\begin{bmatrix} b(0) \\ b(1) \\ b(2) \end{bmatrix} = \begin{bmatrix} 2 & 0 & 0 \\ 1 & 2 & 0 \\ 0 & 1 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \\ 2 \end{bmatrix}$$

Therefore, the model is

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

4.2 A third-order all-pole Padé approximation to a signal $x(n)$ has been determined to be

$$H(z) = \frac{1}{1 + 2z^{-1} + z^{-2} + 3z^{-3}}$$

What information about $x(n)$ can be determined from this model?

Solution _____

The Padé approximation using p poles and q zeros matches the first $p + q + 1$ values of $x(n)$ exactly (assuming that the Padé equations are nonsingular). Therefore, all that can be determined from the Padé model

$$H(z) = \frac{1}{1 + 2z^{-1} + z^{-2} + 3z^{-3}}$$

which has $p = 3$ poles and $q = 0$ zeros, are the first 4 values of $x(n)$. These may be determined by finding the inverse z-transform of $H(z)$ or, alternatively, from the Padé equations,

$$\begin{bmatrix} x(0) & 0 & 0 & 0 \\ x(1) & x(0) & 0 & 0 \\ x(2) & x(1) & x(0) & 0 \\ x(3) & x(2) & x(1) & x(0) \\ x(4) & x(3) & x(2) & x(1) \end{bmatrix} \begin{bmatrix} 1 \\ a(1) \\ a(2) \\ a(3) \end{bmatrix} = \begin{bmatrix} b(0) \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

Substituting the given values for $a(k)$ and $b(k)$ these become

$$\begin{bmatrix} x(0) & 0 & 0 & 0 \\ x(1) & x(0) & 0 & 0 \\ x(2) & x(1) & x(0) & 0 \\ x(3) & x(2) & x(1) & x(0) \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ 1 \\ 3 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

which may be solved by back substitution as follows. From the first equation, it follows that $x(0) = 1$. Next, given $x(0)$, we see from the second equation that

$$x(1) + 2x(0) = 0$$

or, $x(1) = -2$. Then, from the third equation we find

$$x(2) + 2x(1) + x(0) = 0$$

or $x(2) = 3$. Finally, from the last equation we have

$$x(3) + 2x(2) + x(1) + 3x(0) = 0$$

or

$$x(3) = -7$$

Therefore,

$$\mathbf{x} = \begin{bmatrix} 1, & -2, & 3, & -7, & \ldots \end{bmatrix}$$

4.3 Suppose that a signal $x(n)$ is known to be of the form

$$x(n) = \sum_{k=1}^{L} c_k (\lambda_k)^n u(n)$$

where the λ_k's are distinct complex numbers.

(a) Show that the Padé approximation method can be used to determine the parameters c_k and λ_k for $k = 1, 2, \ldots, L$. Is the answer unique?

(b) The first eight values of a signal $x(n)$, which is known to be of the form given above with $L = 3$, are

$$\mathbf{x} = [32, \ 16, \ 8, \ 12, \ 18, \ 33, \ 64.5, \ 128.25]^T$$

Determine c_k and λ_k for $k = 1, 2, 3$.

Solution _____

(a) With

$$x(n) = \sum_{k=1}^{L} c_k (\lambda_k)^n u(n)$$

the z-transform is

$$X(z) = \sum_{k=1}^{L} \frac{c_k}{1 - \lambda_k z^{-1}} = \frac{b(0) + b(1)z^{-1} + \cdots + b(L-1)z^{-(L-1)}}{1 + a(1)z^{-1} + \cdots + a(L)z^{-L}} = \frac{B(z)}{A(z)}$$

which is a rational function of z of order $(L-1)$ in the numerator and order L in the denominator. Therefore, the Padé approximation may be used to find the polynomials $A(z)$ and $B(z)$ provided $p \geq L$ and $q \geq (L-1)$. The coefficients λ_k are the roots of the polynomial $A(z)$ and the coefficients c_k may be found with a partial fraction expansion of $X(z)$.

(b) The Padé equations for the denominator coefficients are

$$\begin{bmatrix} x(q) & x(q-1) & \cdots & x(q-p+1) \\ x(q+1) & x(q) & \cdots & x(q-p+2) \\ \vdots & \vdots & & \vdots \\ x(q+p-1) & x(q+p-2) & \cdots & x(q) \end{bmatrix} \begin{bmatrix} a(1) \\ a(2) \\ \vdots \\ a(p) \end{bmatrix} = - \begin{bmatrix} x(q+1) \\ x(q+2) \\ \vdots \\ x(q+p) \end{bmatrix}$$

If $L = 3$, then we set $p = 3$ and $q = 2$. With

$$\mathbf{x} = \begin{bmatrix} 32, & 16, & 8, & 12, & 18, & 33, & 64.5, & 128.25 \end{bmatrix}^T$$

the Padé equations become

$$\begin{bmatrix} 8 & 16 & 32 \\ 12 & 8 & 16 \\ 18 & 12 & 8 \end{bmatrix} \begin{bmatrix} a(1) \\ a(2) \\ a(3) \end{bmatrix} = - \begin{bmatrix} 12 \\ 18 \\ 33 \end{bmatrix}$$

The solution for $\mathbf{a} = [1, \ a(1), \ a(2), \ a(3)]^T$ is

$$\mathbf{a} = \begin{bmatrix} 1 \\ a(1) \\ a(2) \\ a(3) \end{bmatrix} = \begin{bmatrix} 1 \\ -1.5 \\ -0.75 \\ 0.375 \end{bmatrix}$$

For the numerator coefficients, the Padé equations are

$$\begin{bmatrix} x(0) & 0 & 0 & \cdots & 0 \\ x(1) & x(0) & 0 & \cdots & 0 \\ x(2) & x(1) & x(0) & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ x(q) & x(q-1) & x(q-2) & \cdots & x(q-p) \end{bmatrix} \begin{bmatrix} 1 \\ a(1) \\ a(2) \\ \vdots \\ a(p) \end{bmatrix} = \begin{bmatrix} b(0) \\ b(1) \\ b(2) \\ \vdots \\ b(q) \end{bmatrix}$$

Thus,

$$\begin{bmatrix} b(0) \\ b(1) \\ b(2) \end{bmatrix} = \begin{bmatrix} 32 & 0 & 0 & 0 \\ 16 & 32 & 0 & 0 \\ 8 & 16 & 32 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ a(1) \\ a(2) \\ a(3) \end{bmatrix} = \begin{bmatrix} 32 \\ -32 \\ -40 \end{bmatrix}$$

Thus, the model is

$$H(z) = \frac{32 - 32z^{-1} - 40z^{-2}}{1 - 1.5z^{-1} - 0.75z^{-2} + 0.375z^{-3}}$$

4.4 A consumer electronics device includes a DSP chip that contains a linear shift-invariant digital filter that is implemented in ROM. In order to perform some reverse engineering on the product, it is necessary to determine the system function of the filter. Therefore, the unit sample response is measured and it is determined that the first eight values are as listed in the following table.

Unit sample response
n	0	1	2	3	4	5	6	7
$h(n)$	-1	2	3	2	1	2	0	1

Having no knowlege of the order of the filter, it is assumed that $H(z)$ contains two poles and two zeros.

(a) Based on this assumption, determine a candidate system function, $H(z)$, for the filter.

(b) Based on the solution found in (a) and the given values for $h(n)$, is it possible to determine whether or not the hypothesis about the order of the system is correct? Explain.

Solution _____

(a) The Padé approximation may be used to find the system function of the filter. With $p = q = 2$ the Padé equations are

$$
\begin{bmatrix} -1 & 0 & 0 \\ 2 & -1 & 0 \\ 3 & 2 & -1 \\ 2 & 3 & 2 \\ 1 & 2 & 3 \end{bmatrix}
\begin{bmatrix} 1 \\ a(1) \\ a(2) \end{bmatrix} =
\begin{bmatrix} b(0) \\ b(1) \\ b(2) \\ 0 \\ 0 \end{bmatrix}
$$

Using the last two equations we have

$$
\begin{bmatrix} 3 & 2 \\ 2 & 3 \end{bmatrix}
\begin{bmatrix} a(1) \\ a(2) \end{bmatrix} = -
\begin{bmatrix} 2 \\ 1 \end{bmatrix}
$$

Solving for $a(1)$ and $a(2)$ we find

$$
\begin{bmatrix} a(1) \\ a(2) \end{bmatrix} = \frac{1}{5}
\begin{bmatrix} 3 & -2 \\ -2 & 3 \end{bmatrix}
\begin{bmatrix} -2 \\ -1 \end{bmatrix} =
\begin{bmatrix} -4/5 \\ 1/5 \end{bmatrix}
$$

Next, solving for $b(0)$, $b(1)$, and $b(2)$ using the first three equations we have

$$
\begin{bmatrix} b(0) \\ b(1) \\ b(2) \end{bmatrix} =
\begin{bmatrix} -1 & 0 & 0 \\ 2 & -1 & 0 \\ 3 & 2 & -1 \end{bmatrix}
\begin{bmatrix} 1 \\ -4/5 \\ 1/5 \end{bmatrix} =
\begin{bmatrix} -1 \\ 14/5 \\ 6/5 \end{bmatrix}
$$

Therefore, the system function is

$$
H(z) = \frac{-1 + \frac{14}{5}z^{-1} + \frac{6}{5}z^{-2}}{1 - \frac{4}{5}z^{-1} + \frac{1}{5}z^{-2}}
$$

(b) We may check to see if the inverse z-transform of $H(z)$ matches all of the given values of $h(n)$. Alternatively, if the system function is correct, then $e(n)$ should be equal to zero for $n \geq 5$,

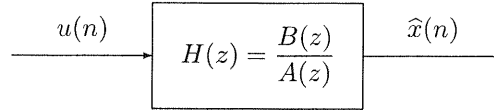$$e(n) = a(n) * h(n) = 0 \quad ; \quad n \geq 5$$

Since

$$e(5) = h(5) + a(1)h(4) + a(2)h(3) = 8/5 \neq 0$$

then the hypothesis about the model order is not correct. There must be more poles and/or zeros.

4.5 The Padé approximation models a signal as the response of a filter to a unit sample input, $\delta(n)$. Suppose, however, that we would like to model a signal $x(n)$ as the *step response* of a filter as shown in the following figure

In the following, assume that $H(z)$ is a second-order filter having a system function of the form

$$H(z) = \frac{b(0) + b(1)z^{-1} + b(2)z^{-2}}{1 + a(1)z^{-1} + a(2)z^{-2}}$$

(a) Using the Padé approximation method with a *unit step* input, derive the set of equations that must be solved so that

$$\widehat{x}(n) = x(n) \quad \text{for} \quad n = 0, 1, \dots, 4$$

(b) If the first eight values of the signal, $x(n)$, are

$$\mathbf{x} = [1, \ 0, \ 2, \ -1, \ 2, \ 0, \ 1, \ 2]^T$$

find $b(0)$, $b(1)$, $b(2)$, $a(1)$, and $a(2)$.

Solution
_____

(a) What we would like to find are polynomials $A(z)$ and $B(z)$ so that

$$U(z) \cdot \frac{B(z)}{A(z)} = X(z)$$

or

$$\frac{1}{1 - z^{-1}} B(z) = A(z)X(z)$$

Multiplying both sides of the equation by $1 - z^{-1}$, we have

$$(1 - z^{-1})A(z)X(z) = B(z)$$

Expressing this equation in the time domain by combining $(1 - z^{-1})$ with $X(z)$ leads to the following set of linear equations that may be easily solved for the filter coefficients $a(k)$ and $b(k)$,

$$\begin{bmatrix} x'(0) & 0 & 0 \\ x'(1) & x'(0) & 0 \\ x'(2) & x'(1) & x'(0) \\ x'(3) & x'(2) & x'(1) \\ x'(4) & x'(3) & x'(2) \end{bmatrix} \begin{bmatrix} 1 \\ a(1) \\ a(2) \end{bmatrix} = \begin{bmatrix} b(0) \\ b(1) \\ b(2) \\ 0 \\ 0 \end{bmatrix}$$

where the sequence $x'(n)$ is defined as follows

$$x'(n) = x(n) - x(n-1)$$

with $x(n) = 0$ for $n < 0$. Alternatively, combining the term $(1 - z^{-1})$ with $A(z)$ leads to the following set of equations

$$\begin{bmatrix} x(0) & 0 & 0 & 0 \\ x(1) & x(0) & 0 & 0 \\ x(2) & x(1) & x(0) & 0 \\ x(3) & x(2) & x(1) & x(0) \\ x(4) & x(3) & x(2) & x(1) \end{bmatrix} \begin{bmatrix} 1 \\ a'(1) \\ a'(2) \\ a'(3) \end{bmatrix} = \begin{bmatrix} b(0) \\ b(1) \\ b(2) \\ 0 \\ 0 \end{bmatrix}$$

where the coefficients $a'(k)$ are defined by

$$a'(k) = a(k) - a(k-1)$$

with $a(0) = 1$ and $a(k) = 0$ for $k < 0$ and $k > 2$. Yet another possibility is to express the following equation in the time domain

$$A(z)X(z) = \frac{1}{1 - z^{-1}} B(z)$$

Since the term on the right is the convolution of a step with the coefficients $b(k)$ we have

$$\begin{bmatrix} x(0) & 0 & 0 \\ x(1) & x(0) & 0 \\ x(2) & x(1) & x(0) \\ x(3) & x(2) & x(1) \\ x(4) & x(3) & x(2) \end{bmatrix} \begin{bmatrix} 1 \\ a(1) \\ a(2) \end{bmatrix} = \begin{bmatrix} b(0) \\ b(0) + b(1) \\ b(0) + b(1) + b(2) \\ b(0) + b(1) + b(2) \\ b(0) + b(1) + b(2) \end{bmatrix}$$

Any one of these sets of equations may be used to solve for the coefficients $a(k)$ and $b(k)$.

(b) Using the first approach derived in part (a), we first form the sequence $x'(n)$,

$$\mathbf{x}' = \begin{bmatrix} 1, & -1, & 2, & -3, & 3, & -2, & 1, & 1 \end{bmatrix}^T$$

The Padé equations are then

$$\begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 2 & -1 & 1 \\ -3 & 2 & -1 \\ 3 & -3 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ a(1) \\ a(2) \end{bmatrix} = \begin{bmatrix} b(0) \\ b(1) \\ b(2) \\ 0 \\ 0 \end{bmatrix}$$

From the last two equations we have

$$\begin{bmatrix} 2 & -1 \\ -3 & 2 \end{bmatrix} \begin{bmatrix} a(1) \\ a(2) \end{bmatrix} = \begin{bmatrix} 3 \\ -3 \end{bmatrix}$$

Solving for $a(1)$ and $a(2)$ we find

$$\begin{bmatrix} a(1) \\ a(2) \end{bmatrix} = \begin{bmatrix} 2 & 1 \\ 3 & 2 \end{bmatrix} \begin{bmatrix} 3 \\ -3 \end{bmatrix} = \begin{bmatrix} 3 \\ 3 \end{bmatrix}$$

Finally, solving for $b(0)$, $b(1)$, and $b(2)$ we have

$$\left[\begin{array}{c} b(0) \\ b(1) \\ b(2) \end{array}\right] = \left[\begin{array}{ccc} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 2 & -1 & 1 \end{array}\right] \left[\begin{array}{c} 1 \\ a(1) \\ a(2) \end{array}\right] = \left[\begin{array}{c} 1 \\ 2 \\ 2 \end{array}\right]$$

Therefore, the model is

$$H(z) = \frac{1 + 2z^{-1} + 2z^{-2}}{1 + 3z^{-1} + 3z^{-2}}$$

4.6 With a real-valued signal $x(n)$ known only for $n = 0, 1, \ldots, N$, the *backwards covariance method* finds the coefficients of an all-pole model that minimize the *backward prediction error*

$$\mathcal{E}_p^- = \sum_{n=p}^{N} [e_p^-(n)]^2$$

where

$$e_p^-(n) = x(n-p) + \sum_{k=1}^{p} a_p(k) x(n+k-p)$$

(a) Show that the coefficients $a_p(k)$ that minimize $\mathcal{E}_p^-$ satisfy a set of normal equations of the form

$$\mathbf{R}_x \bar{\mathbf{a}}_p = -\mathbf{r}_x$$

where

$$\bar{\mathbf{a}}_p = [a_p(1), \ a_p(2), \ \ldots \ , a_p(p)]^T$$

and find explicit expressions for the entries in $\mathbf{R}_x$ and $\mathbf{r}_x$.

(b) Is the solution to the backwards covariance method the same as the solution to the covariance method? Why or why not?

(c) Consider a new error that is the sum of the forward and backward prediction errors

$$\mathcal{E}_p^B = \sum_{n=p}^{N} \left\{ [e_p^+(n)]^2 + [e_p^-(n)]^2 \right\}$$

where $e_p^-(n)$ is the error defined above and $e_p^+(n)$ is the forward prediction error used in the covariance method,

$$e_p^+(n) = x(n) + \sum_{k=1}^{P} a_p(k) x(n-k)$$

Derive the normal equations for the coefficients that minimize this error. (This approach is known as the *Modified Covariance Method.*)

(d) Consider the signal

$$x(n) = \beta^n \quad ; \quad n = 0, 1, \ldots, N$$

With $p = 1$ find the first-order all-pole model that minimizes $\mathcal{E}_p^B$ and determine the minimum value of $\mathcal{E}_p^B$. For what values of β is the model stable? What happens to the model and the modeling error as $N \to \infty$?

Solution _____

(a) With $\mathcal{E}_p^- = \sum_{n=p}^{N}[e_p^-(n)]^2$, setting the partial derivative of $\mathcal{E}_p^-$ with respect to $a_p(l)$ equal to zero we have

$$\frac{\partial \mathcal{E}_p^-}{\partial a_p(l)} = \sum_{n=p}^{N} 2e_p^-(n)\frac{\partial e_p^-(n)}{\partial a_p(l)} = 2\sum_{n=p}^{N} e_p^-(n)x(n+l-p) = 0$$

Dividing by two, and substituting for $e_p^-(n)$ yields

$$\sum_{n=p}^{N} x(n-p)x(n+l-p) + \sum_{k=1}^{p} a_p(k)\left[\sum_{n=p}^{N} x(n+k-p)x(n+l-p)\right] = 0$$

If we define

$$r_x(k,l) = \sum_{n=p}^{N} x(n+k-p)x(n+l-p) = \sum_{n=0}^{N-p} x(n+k)x(n+l)$$

then the normal equations become

$$\sum_{k=1}^{p} a_p(k)r_x(k,l) = -r_x(0,l) \quad ; \quad l = 1,2,\ldots,p$$

(b) No, the backwards covariance method does not give the same solution as the covariance method since the definitions for $r_x(k,l)$ are different in the two methods. Specifically, for the covariance method

$$r_x(k,l) = \sum_{n=p}^{N} x(n-l)x(n-k)$$

(c) As before, we differentiate $\mathcal{E}_p^B$ with respect to $a_p(l)$, and set the result equal to zero,

$$\frac{\partial \mathcal{E}_p^B}{\partial a_p(l)} = \sum_{n=p}^{N} 2\left[e_p^+(n)x(n-l) + e_p^-(n)x(n+l-p)\right] = 0$$

Dividing by two, and substituting for $e_p^-(n)$ and $e_p^+(n)$ we have

$$0 = \sum_{n=p}^{N}[x(n)x(n-l) + x(n-p)x(n+l-p)$$
$$+ \sum_{k=1}^{p} a_p(k)\sum_{n=p}^{N}\left[x(n-k)x(n-l) + x(n+k-p)x(n+l-p)\right]$$

Defining

$$r_x(k,l) = \sum_{n=p}^{N} x(n-l)x(n-k)$$

we have

$$r_x(l,0) + r_x(p-l,p) + \sum_{k=1}^{p} a_p(k)\left[r(l,k) + r_x(p-l,p-k)\right] = 0$$

Therefore, the normal equations are

$$\sum_{k=1}^{p} a_p(k)\left[r_x(l,k) + r_x(p-l,p-k)\right] = -\left[r_x(l,0) + r_x(p-l,p)\right] \quad ; \quad l = 1,\ldots,p$$

(d) With $p = 1$ we have

$$a(1) = -\frac{r_x(1,0) + r_x(0,1)}{r_x(1,1) + r_x(0,0)} = -\frac{2r_x(1,0)}{r_x(1,1) + r_x(0,0)}$$

Since $x(n) = \beta^n$ for $n = 0, 1, \ldots, N$, then

$$r_x(0,0) = \sum_{n=1}^{N} x^2(n) = \beta^2 \frac{1 - \beta^{2N}}{1 - \beta^2}$$

$$r_x(1,1) = \sum_{n=1}^{N} x^2(n-1) = \frac{1 - \beta^{2N}}{1 - \beta^2}$$

$$r_x(1,0) = \sum_{n=1}^{N} x(n)x(n-1) = \beta \frac{1 - \beta^{2N}}{1 - \beta^2}$$

Therefore,

$$a(1) = -\frac{2\beta}{1 + \beta^2}$$

Note that for any value of β,

$$(1 + \beta)^2 = 1 + \beta^2 + 2\beta \geq 0$$

Thus,

$$1 + \beta^2 \geq -2\beta$$

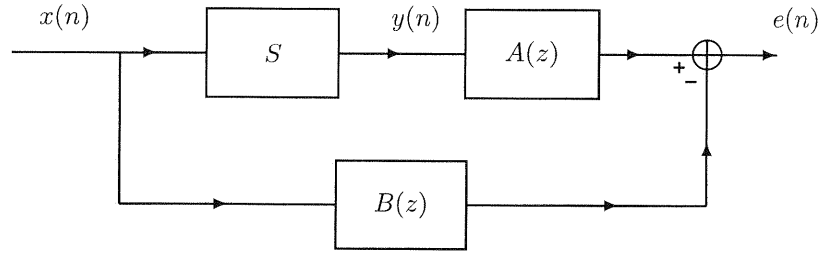and the model is *stable* for all values of β.
For the modeling error, we have

$$\mathcal{E}_p^B = r_x(0,0) + r_x(p,p) + \sum_{k=1}^{p} a_p(k)\big[r_x(0,k) + r_x(p, p-k)\big]$$

Therefore,

$$\begin{aligned}
\mathcal{E}_1^B &= r_x(0,0) + r_x(1,1) + a(1)\big[r_x(0,1) + r_x(1,0)\big] \\
&= \frac{1 - \beta^{2N}}{1 - \beta^2}\Big[\beta^2 + 1 - \frac{4\beta^2}{1 + \beta^2}\Big] = \frac{1 - \beta^2}{1 + \beta^2}(1 - \beta^{2N})
\end{aligned}$$

and the modelling error *does not* go to zero as $N \to \infty$.

4.7 Suppose that we would like to derive a rational model for an unknown system S using the approach shown in the following figure

For a given input $x(n)$ the output of the system, $y(n)$, is observed. The coefficients of the two FIR filters $A(z)$ and $B(z)$ that minimize the sum of the squares of the error signal $e(n)$ are then to be determined. Assume that the sum is for all $n \geq 0$ as in Eq. (4.73).

(a) Derive the normal equations that define the optimal solution for the coefficients of $A(z)$ and $B(z)$.

(b) The philosophy of this method is that if the error is small then $B(z)/A(z)$ is a reasonable model for S. Suppose that S is a linear shift-invariant system with a rational system function. Show that this method will identify the parameters of S exactly assuming that the orders of the filters $A(z)$ and $B(z)$ are chosen appropriately.

Solution _____

(a) Note that $E(z) = Y(z)A(z) - X(z)B(z)$, so

$$e(n) = \sum_{k=0}^{p} a(k)y(n-k) - \sum_{k=0}^{q} b(k)x(n-k)$$

With

$$\mathcal{E} = \sum_{n=0}^{\infty} e^2(n)$$

the *Normal Equations* are found by setting the derivatives of $\mathcal{E}$ with respect to $a(k)$ and $b(k)$ equal to zero,

$$\frac{\partial \mathcal{E}}{\partial a(k)} = 0 \quad ; \quad \frac{\partial \mathcal{E}}{\partial b(k)} = 0$$

Thus,

$$\frac{\partial \mathcal{E}}{\partial a(k)} = \sum_{n=0}^{\infty} 2e(n)y(n-k) = 2\sum_{n=0}^{\infty} \left\{ \sum_{l=0}^{p} a(l)y(n-l) - \sum_{l=0}^{q} b(l)x(n-l) \right\} y(n-k) = 0$$

and

$$\frac{\partial \mathcal{E}}{\partial b(k)} = -\sum_{n=0}^{\infty} 2e(n)x(n-k) = -2\sum_{n=0}^{\infty} \left\{ \sum_{l=0}^{p} a(l)y(n-l) - \sum_{l=0}^{q} b(l)x(n-l) \right\} x(n-k) = 0$$

Dividing by two, and rearranging the sums, we have

$$\sum_{l=0}^{p} a(l) \left\{ \sum_{n=0}^{\infty} y(n-l)y(n-k) \right\} - \sum_{l=0}^{q} b(l) \left\{ \sum_{n=0}^{\infty} x(n-l)y(n-k) \right\} = 0 \quad ; \quad k = 1, \ldots, p$$

and

$$-\sum_{l=0}^{p} a(l) \left\{ \sum_{n=0}^{\infty} y(n-l)x(n-k) \right\} + \sum_{l=0}^{q} b(l) \left\{ \sum_{n=0}^{\infty} x(n-l)x(n-k) \right\} = 0 \quad ; \quad k = 0, \ldots, q$$

If we define

$$r_{xy}(k,l) \quad = \quad \sum_{n=0}^{\infty} x(n-l)y(n-k)$$

$$r_y(k,l) \quad = \quad \sum_{n=0}^{\infty} y(n-l)y(n-k)$$

$$r_x(k,l) \quad = \quad \sum_{n=0}^{\infty} x(n-l)x(n-k)$$

then these equations become

$$\sum_{l=0}^{p} a(l)r_y(k,l) - \sum_{l=0}^{q} b(l)r_{xy}(k,l) \quad = \quad 0 \quad ; \quad k = 1, 2, \ldots, p$$

$$-\sum_{l=0}^{p} a(l)r_{yx}(k,l) + \sum_{l=0}^{q} b(l)r_x(k,l) \quad = \quad 0 \quad ; \quad k = 0, 1, \ldots, q$$

Assuming that the coefficients have been normalized so that $a(0) = 1$, we have

$$\sum_{l=1}^{p} a(l)r_y(k,l) - \sum_{l=0}^{q} b(l)r_{xy}(k,l) \quad = \quad -r_y(k,0) \quad ; \quad k = 1, 2, \ldots, p$$

$$-\sum_{l=1}^{p} a(l)r_{yx}(k,l) + \sum_{l=0}^{q} b(l)r_x(k,l) \quad = \quad r_{yx}(k,0) \quad ; \quad k = 0, 1, \ldots, q$$

Writing these in matrix form we obtain

$$\begin{bmatrix} \mathbf{R}_y & -\mathbf{R}_{xy} \\ -\mathbf{R}_{yx} & \mathbf{R}_x \end{bmatrix} \begin{bmatrix} \mathbf{a} \\ \mathbf{b} \end{bmatrix} = \begin{bmatrix} -\mathbf{r}_x \\ \mathbf{r}_{xy} \end{bmatrix}$$

where $\mathbf{a}^T = [a(1), a(2), \ldots, a(p)]$, $\mathbf{b}^T = [b(0), b(1), \ldots, b(q)]$, $\mathbf{r}_x^T = [r_x(1,0), r_x(2,0), \ldots, r_x(p,0)]$, and $\mathbf{r}_{xy}^T = [r_{xy}(1,0), r_{xy}(2,0), \ldots, r_{xy}(q,0)]$. Also, $\mathbf{R}_x$ is a $p \times p$ matrix with entries $r_x(k,l)$, $\mathbf{R}_y$ is a $(q+1) \times (q+1)$ matrix with entries $r_q(k,l)$, and $\mathbf{R}_{xy}$ is a $p \times (q+1)$ matrix with entries $r_x(k,l)$,

(b) Suppose $S(z) = C(z)/D(z)$. Then

$$E(z) = B(z)X(z) - \frac{C(z)}{D(z)}A(z)X(z)$$

and the error can be made equal to zero if

$$B(z) = \frac{C(z)}{D(z)}A(z) \quad \text{or} \quad \frac{B(z)}{A(z)} = \frac{C(z)}{D(z)}$$

4.8 Consider a signal, $x(n)$, which is the unit sample response of a causal all-pole filter with system function

$$H(z) = \frac{1}{(1 + 0.5z^{-1})(1 + 0.75z^{-1})(1 + 2z^{-1})}$$

We observe $x(n)$ over the interval $0 \leq n \leq N$ where $N \gg 1$.

(a) Using the covariance method, we determine a third-order all-pole model for $x(n)$. What, if anything, can you say about the location of the poles in the model? Do the pole locations depend on N? If so, where do the poles go in the limit as $N \to \infty$?

(b) Repeat part (a) for the case in which you use the autocorrelation method.

Solution ⸻⸻⸻⸻⸻⸻⸻⸻⸻⸻⸻⸻⸻⸻⸻⸻

Note that the sequence that we are trying to model is the unit sample response of a causal filter that has poles at $z = -0.5, -0.75, -2$.

(a) Since $x(n)$ is the unit sample response of an all-pole filter, the covariance method will produce an exact model for the signal, independent of the value of N (assuming that $N \geq 6$).

(b) For the autocorrelation method, the roots will always lie inside the unit circle and will vary with N. However, as N gets large, the roots will move towards the minimum phase solution, with poles at $z = -0.5, -0.75, -0.5$.

4.9 Equation (4.129) may be used to reduce the amount of computation required to set-up the covariance normal equations.

(a) Show that the elements along the main diagonal may be computed recursively beginning with $r_x(1,1)$.

(b) Show how the elements along the lower diagonals may be computed recursively beginning with $r_x(k,1)$. How may the terms along the upper diagonals be obtained?

(c) Determine how many multiplies and adds are necessary to set-up the covariance normal equations (do not forget the evaluation of the vector on the right-hand side).

Solution _____

Using the relationship between $r_x(k+1, l+1)$ and $r_x(k,l)$,

$$r_x(k+1, l+1) = r_x(k,l) - x(N-l)x^*(N-k) + x(p-1-l)x^*(p-1-k)$$

we may evaluate the terms in the covariance normal equations recursively.

(a) Beginning with $r_x(1,1)$, the elements along the main diagonal of the covariance normal equations may be found recursively as follows

$$r_x(k+1, k+1) = r_x(k,k) - |x(N-k)|^2 + |x(p-1-k)|^2$$

This requires two multiplications and two additions for each term.

(b) Beginning with $r_x(k,1)$, the elements along the lower diagonals may be computed recursively as follows

$$r_x(k+1, 2) = r_x(k,1) - x(N-1)x^*(N-k) + x(p-2)x^*(p-k)$$

and

$$r_x(k+2, 3) = r_x(k+1, 2) - x(N-2)x^*(N-k-1) + x(p-3)x^*(p-k-1)$$

or, in general,

$$r_x(k+l+1, l+2) = r_x(k+l, l+1) - x(N-l-1)x^*(N-k-l) + x(p-l-2)x^*(p-k-l-1)$$

As with the terms along the main diagonal, each term on the lower diagonals requires two multiplications and two additions. Note that the upper diagonals may be found using the conjugate symmetry of the covariance normal equations.

(c) The covariance normal equations require finding the elements of a $p \times p$ Hermitian matrix. As shown in parts (a) and (b), given the first column of the matrix, the remaining entries may be computed recursively. Given the elements in the first column, $r_x(k,1)$, each of the $(p-k)$ terms along the diagonals, i.e, $r_x(k+l, l)$ for $l = 1, \ldots, p-k-1$, require 2 multiplications and 2 additions. This requires

$$\sum_{k=1}^{p-1} 2(p-k) = 2p(p-1) - p(p-1) = p(p-1)$$

multiplications, and the same number of additions. In addition, it is necessary to evaluate the p terms in the first column,

$$r_x(k,1) = \sum_{n=p}^{N} x(n-l)x^*(n-k) \quad ; \quad k = 1, 2, \ldots, p$$

This requires $p(N-p+1)$ multiplications and $p(N-p)$ additions. Therefore, the total number of multiplications is .

$$\# \text{ mults. } = p(N - p + 1) + p(p - 1) = Np$$

and the total number of additions is

$$\# \text{ adds. } = p(N - p) + p(p - 1) = Np - p$$

4.10 We want to model a signal $x(n)$ using an all-pole model of the form

$$H(z) = \frac{b(0)}{1 + z^{-N}\left[\sum\limits_{k=1}^{p} a_p(k)z^{-k}\right]}$$

For example, with $p = 2$ the model is

$$H(z) = \frac{b(0)}{1 + a(1)z^{-N-1} + a(2)z^{-N-2}}$$

Derive the normal equations that define the coefficients $a_p(k)$ that minimize the Prony error

$$\mathcal{E}_p = \sum_{n=0}^{\infty} |e(n)|^2$$

where

$$e(n) = x(n) + \sum_{l=1}^{p} a_p(l)x(n-l-N)$$

and derive an expression for the minimum error.

Solution _____

The equations for the coefficients $a_p(k)$, $k = 1, \ldots p$, that minimize the error $\mathcal{E}_p$ are found by setting the derivatives of $\mathcal{E}_p$ with respect to $a_p(k)$ equal to zero. Thus, assuming that $x(n)$ is real, we have

$$\frac{\partial \mathcal{E}_p}{\partial a_p(k)} = \sum_{n=0}^{\infty} 2e(n)x(n-k-N) = 0$$

Dividing by two, and substituting for $e(n)$, we have

$$\sum_{n=0}^{\infty} \left[x(n) + \sum_{l=1}^{p} a_p(l)x(n-l-N) \right] x(n-k-N) = 0$$

or

$$\sum_{l=1}^{p} a_p(l) \left[\sum_{n=0}^{\infty} x(n-l-N)x(n-k-N) \right] = -\sum_{n=0}^{\infty} x(n)x(n-k-N)$$

If we define

$$r_x(k,l) = \sum_{n=0}^{\infty} x(n-l)x(n-k)$$

then it is easily shown that $r_x(k,l)$ depends only on the difference, $k - l$, and we may write

$$r_x(k) = \sum_{n=0}^{\infty} x(n)x(n-k)$$

Thus, the normal equations become

$$\sum_{l=1}^{p} a_p(l)r_x(k-l) = -r_x(k+N)$$

Finally, using the orthogonality condition

$$\sum_{n=0}^{\infty} e(n)x(n-k-N) = 0$$
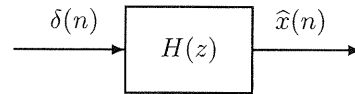
we have, for the minimum error,

$$\{\mathcal{E}_p\}_{\min} = \sum_{n=0}^{\infty} e(n)\left[x(n) + \sum_{l=1}^{p} a_p(l)x(n-l-N)\right] = \sum_{n=0}^{\infty} e(n)x(n)$$

Therefore,

$$\begin{aligned}
\{\mathcal{E}_p\}_{\min} &= \sum_{n=0}^{\infty}\left[x(n) + \sum_{l=1}^{p} a_p(l)x(n-l-N)\right]x(n) \\
&= r_x(0) + \sum_{l=1}^{p} a_p(l)r_x(l+N)
\end{aligned}$$

4.11 Suppose that we would like to model a signal $x(n)$ as shown in the following figure.

where $h(n)$ is an all-pole filter that has a system function of the form

$$H(z) = \frac{b(0)}{1 + \displaystyle\sum_{k=1}^{p} a_p(k)z^{-2k}}$$

Modify the Prony normal equations so that one can determine the coefficients $a_p(k)$ in $H(z)$ from a sequence of signal values, $x(n)$.

Solution _____

To minimize the Prony error

$$\mathcal{E}_p = \sum_{n=0}^{\infty} |e(n)|^2$$

where

$$e(n) = x(n) + \sum_{l=1}^{p} a_p(l)x(n-2l)$$

we set the derivative of $\mathcal{E}_p$ with respect to $a_p(k)$ equal to zero,

$$\frac{\partial \mathcal{E}_p}{\partial a_p(k)} = 2\sum_{n=0}^{\infty} e(n)x(n-2k) = 0$$

which gives

$$\sum_{n=0}^{\infty} \left[x(n) + \sum_{l=1}^{p} a_p(l)x(n-2l) \right] x(n-2k) = 0$$

Therefore,

$$\sum_{l=1}^{p} a_p(l)r_x(2k-2l) = -r_x(2k)$$

where

$$r_x(2k-2l) = \sum_{n=0}^{\infty} x(n-2l)x(n-2k)$$

For example, with $p = 2$ the equations have the form

$$\begin{bmatrix} r_x(0) & r_x(2) \\ r_x(2) & r_x(0) \end{bmatrix} \begin{bmatrix} a(1) \\ a(2) \end{bmatrix} = \begin{bmatrix} r_x(2) \\ r_x(4) \end{bmatrix}$$

4.12 Suppose that we would like to model a signal $x(n)$ that we believe to be quasiperiodic. Based on our observations of $x(n)$ we estimate the autocorrelations through lag $k = 10$ to be

$$r_x(k) = [1.0, \ 0.4, \ 0.4, \ 0.3, \ 0.2, \ 0.9, \ 0.4, \ 0.4, \ 0.2, \ 0.1, \ 0.7]^T$$

(a) In formulating an all-pole model to take into account the suspected periodicity let us consider a two-coefficient model of the form

$$H(z) = \frac{b(0)}{1 + a(5)z^{-5} + a(10)z^{-10}}$$

Find the values for the coefficients $a(5)$ and $a(10)$ that minimize the all-pole modeling error.

(b) Compare the error obtained with the model found in (a) to the error that is obtained with a model of the form

$$H(z) = \frac{b(0)}{1 + a(1)z^{-1} + a(2)z^{-2}}$$

(c) Now consider an all-pole model of the form

$$H(z) = \frac{b(0)}{1 + a(N)z^{-N}}$$

where both $a(N)$ and N are considered to be model parameters. Find the value for $a(N)$ and N that minimize the all-pole modeling error and evaluate the modeling error.

Solution _____

(a) With an all-pole model of the form

$$H(z) = \frac{b(0)}{1 + a(5)z^{-5} + a(10)z^{-10}}$$

we begin by defining the error that we want to minimize. Let

$$\mathcal{E} = \sum_{n=0}^{\infty} e^2(n)$$

with

$$e(n) = x(n) + a(5)x(n - k) + a(10)x(n - 10)$$

To find the coefficients $a(5)$ and $a(10)$ that minimize $\mathcal{E}$, we set the partial derivatives of $\mathcal{E}$ with respect to $a(5)$ and $a(10)$ equal to zero,

$$\frac{\partial \mathcal{E}}{\partial a(5)} = 2 \sum_{n=0}^{\infty} e(n)x(n-5) = 2 \sum_{n=0}^{\infty} \big[x(n) + a(5)x(n-5) + a(10)x(n-10) \big] x(n-5) = 0$$

and

$$\frac{\partial \mathcal{E}}{\partial a(10)} = 2 \sum_{n=0}^{\infty} e(n)x(n-10) = 2 \sum_{n=0}^{\infty} \left[x(n) + a(5)x(n-5) + a(10)x(n-10) \right] x(n-10) = 0$$

Dividing by two and rearranging we have

$$a(5) \sum_{n=0}^{\infty} x^2(n-5) + a(10) \sum_{n=0}^{\infty} x(n-10)x(n-5) = - \sum_{n=0}^{\infty} x(n)x(n-5)$$

and

$$a(5) \sum_{n=0}^{\infty} x(n-5)x(n-10) + a(10) \sum_{n=0}^{\infty} x^2(n-10) = - \sum_{n=0}^{\infty} x(n)x(n-10)$$

These equations may be written as

$$\begin{bmatrix} r_x(0) & r_x(5) \\ r_x(5) & r_x(0) \end{bmatrix} \begin{bmatrix} a(5) \\ a(10) \end{bmatrix} = - \begin{bmatrix} r_x(5) \\ r_x(10) \end{bmatrix}$$

where

$$r_x(k) = \sum_{n=0}^{\infty} x(n)x(n-k)$$

Using the given autocorrelations, thse become

$$\begin{bmatrix} 1 & .9 \\ .9 & 1 \end{bmatrix} \begin{bmatrix} a(5) \\ a(10) \end{bmatrix} = - \begin{bmatrix} .9 \\ .7 \end{bmatrix}$$

Solving for $a(5)$ and $a(10)$ we find

$$\begin{bmatrix} a(5) \\ a(10) \end{bmatrix} = \begin{bmatrix} -1.4211 \\ 0.5789 \end{bmatrix}$$

Finally, for the modeling error, we have

$$\mathcal{E} = r_x(0) + a(5)r_x(5) + a(10)r_x(10) = 0.1263$$

(b) With a model of the form

$$H(z) = \frac{b(0)}{1 + a(1)z^{-1} + a(2)z^{-2}}$$

the normal equations are

$$\begin{bmatrix} r_x(0) & r_x(1) \\ r_x(1) & r_x(0) \end{bmatrix} \begin{bmatrix} a(1) \\ a(2) \end{bmatrix} = - \begin{bmatrix} r_x(1) \\ r_x(2) \end{bmatrix}$$

and the filter coefficients are solutions to the equations

$$\begin{bmatrix} 1 & 0.4 \\ 0.4 & 1 \end{bmatrix} \begin{bmatrix} a(1) \\ a(2) \end{bmatrix} = - \begin{bmatrix} 0.4 \\ 0.4 \end{bmatrix}$$

Thus,

$$\begin{bmatrix} a(1) \\ a(2) \end{bmatrix} = \begin{bmatrix} 2/7 \\ 2/7 \end{bmatrix}$$

Finally, the modeling error is

$$\mathcal{E} = r_x(0) + a(1)r_x(1) + a(2)r_x(2) = 1.2286$$

(c) Using a model of the form

$$H(z) = \frac{b(0)}{1 + a(N)z^{-N}}$$

the value for the coefficient $a(N)$ that minimizes the mean-square error is

$$a(N) = -\frac{r_x(N)}{r_x(0)}$$

and the minimum mean-square error is given by

$$\{\xi_1\}_{\min} = \frac{r_x^2(0) - r_x^2(N)}{r_x(0)}$$

Therefore, the mean-square error is the smallest when $N = 5$, which is the value of k for which $|r_x(k)|$ is the largest. In other words, to minimize the error we want to find the value of N for which $x(n)$ and $x(n + N)$ have the *highest* correlation.

4.13 We would like to build a predictor of digital waveforms. Such a system would form an estimate of a later sample (say n_0 samples later) by observing p consecutive data samples. Thus we would set

$$\widehat{x}(n + n_0) = \sum_{k=1}^{p} a_p(k)x(n - k)$$

The predictor coefficients $a_p(k)$ are to be chosen to minimize

$$\mathcal{E}_p = \sum_{n=0}^{\infty} [x(n + n_0) - \widehat{x}(n + n_0)]^2$$

(a) Derive the equations that define the optimum set of coefficients $a_p(k)$.

(b) If $n_0 = 0$, how is your formulation of this problem different from Prony's method?

Solution _____

(a) We want to find the predictor coefficients $a_p(k)$ that minimize the linear prediction error

$$\mathcal{E}_p = \sum_{n=0}^{\infty} [e(n)]^2$$

where

$$e(n) = x(n + n_0) - \widehat{x}(n + n_0)$$

To find these coefficients, differentiate $\mathcal{E}_p$ with respect to $a_p(k)$, and set the derivatives equal to zero as follows

$$\frac{\partial \mathcal{E}_p}{\partial a_p(k)} = -\sum_{n=0}^{\infty} 2e(n)\frac{\partial \widehat{x}(n + n_0)}{\partial a_p(k)} = 0$$

Since

$$\widehat{x}(n + n_0) = \sum_{k=1}^{p} a_p(k)x(n - k)$$

then

$$\frac{\partial \widehat{x}(n + n_0)}{\partial a_p(k)} = x(n - k)$$

and we have

$$\frac{\partial \mathcal{E}_p}{\partial a_p(k)} = -2\sum_{n=0}^{\infty} e(n)x(n - k) = 0$$

Dividing by two, and substituting for $e(n)$, we have

$$\sum_{n=0}^{\infty} \left\{ x(n + n_0) - \sum_{l=1}^{p} a_p(l)x(n - l) \right\} x(n - k) = 0 \quad ; \quad k = 1, 2, \ldots, p$$

Therefore, the *normal equations* are

$$\sum_{l=1}^{p} a_p(l)r_x(k, l) = r_x(k, -n_0)$$

where

$$r_x(k,l) = \sum_{n=0}^{\infty} x(n-l)x(n-k)$$

(b) With $n_0 = 0$, these equations are the same as the all-pole normal equations using Prony's method, except that the right-hand side does not have a minus sign. Therefore, the solution differs in sign.

4.14 You are told that it is always possible to determine whether or not a causal all-pole filter is stable from a finite number of values of its unit sample response. For example, if $H(z)$ is a pth-order all-pole filter, given $h(n)$ for $n = 0, 1, \ldots, N$, then the stability of $H(z)$ may be determined. If this is true, explain the procedure and list any conditions that must be placed on p or N. If false, explain why it cannot be done.

Solution _____

It is true that the stability of a causal all-pole filter can be determined from a finite number of values of its unit sample response. Given a pth-order all-pole filter,

$$H(z) = \frac{b(0)}{1 + \sum_{k=1}^{p} a_p(k) z^{-k}}$$

the coefficients $a(k)$ may be found using the Padé approximation for an all-pole model,

$$\begin{bmatrix} h(0) & 0 & \cdots & 0 \\ h(1) & h(0) & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ h(p-1) & h(p-2) & \cdots & h(0) \end{bmatrix} \begin{bmatrix} a_p(1) \\ a_p(2) \\ \vdots \\ a_p(p) \end{bmatrix} = - \begin{bmatrix} h(1) \\ h(2) \\ \vdots \\ h(p) \end{bmatrix}$$

Thus, given $h(n)$ for $n = 0, 1, \ldots, p$, the coefficients may be found, and the roots of the polymomial

$$A(z) = 1 + \sum_{k=1}^{p} a(k) z^{-k}$$

checked to see if they lie inside the unit circle.

4.15 Let $H(z)$ be a first-order model for a real-valued signal $x(n) = \delta(n) + \delta(n - 1)$,

$$H(z) = \frac{b(0)}{1 - a(1)z^{-1}}$$

and let

$$\mathcal{E}_{LS} = \sum_{n=0}^{\infty} [x(n) - h(n)]^2$$

be the error that is to be minimized. By setting the derivatives of $\mathcal{E}_{LS}$ with respect to $b(0)$ and $a(1)$ equal to zero, try to find an analytic solution for the values of $b(0)$ and $a(1)$ that minimize $\mathcal{E}_{LS}$. (This problem illustrates how difficult the direct method of signal modeling may be, even for a first-order model.)

Solution _____

We are given a signal $x(n) = \delta(n) + \delta(n - 1)$ that we would like to model as the unit sample response of the all-pole filter

$$H(z) = \frac{b(0)}{1 - a(1)z^{-1}}$$

The goal is to minimize the error

$$\mathcal{E}_{LS} = \sum_{n=0}^{\infty} [h(n) - x(n)]^2$$

Thus, the model for $x(n)$ will be

$$\hat{x}(n) = b(0)[a(1)]^n u(n)$$

To find the values for $a(1)$ and $b(0)$ that minimize $\mathcal{E}_{LS}$, we begin by setting the derivative of $\mathcal{E}_{LS}$ with respect to $a(1)$ and $b(0)$ equal to zero. Thus, we have

$$\frac{\partial \mathcal{E}_{LS}}{\partial a(1)} = -\sum_{n=0}^{\infty} 2 \left[x(n) - \hat{x}(n) \right] \, nb(0)a^{n-1}(1) = 0$$

Dividing by two, and substituting for $x(n)$ and $\hat{x}(n)$ we have

$$
\begin{aligned}
\sum_{n=0}^{\infty} [x(n) - \hat{x}(n)] \, nb(0)a^{n-1}(1) &= b(0)(1 - a(1)b(0)) - \sum_{n=2}^{\infty} (b(0)a^n(1) \, nb(0)a^{n-1}(1) \\
&= b(0)(1 - a(1)b(0)) - b(0)^2 a^{-1}(1) \sum_{n=2}^{\infty} na^{2n}(1) \\
&= b(0)(1 - a(1)b(0)) - b(0)^2 a^{-1}(1) \sum_{n=0}^{\infty} na^{2n}(1) + b(0)^2 a(1) \\
&= b(0) - b(0)^2 a^{-1}(1) \sum_{n=0}^{\infty} na^{2n}(1)] \\
&= b(0) - \frac{a(1)b(0)^2}{(1 - a^2(1))^2} = 0
\end{aligned}
$$

Simplifying we have

$$a(1)b(0) = \left[1 - a^2(1) \right]^2 \tag{P4.15-1}$$

Differentiating $\mathcal{E}_{LS}$ with respect to $b(0)$ we have

$$\frac{\partial \mathcal{E}_{LS}}{\partial b(0)} = - \sum_{n=0}^{\infty} 2 \left[x(n) - \hat{x}(n) \right] a^n(1) = 0$$

Again, dividing by two, and substituting for $x(n)$ and $\hat{x}(n)$ we have

$$\begin{aligned}
\sum_{n=0}^{\infty} \left[x(n) - \hat{x}(n) \right] a^n(1) &= \left[1 - b(0) \right) + a(1)(1 - a(1)b(0)) \right] - \sum_{n=2}^{\infty} b(0) a^{2n}(1) \\
&= \left[1 - b(0) \right) + a(1)(1 - a(1)b(0)) \right] - b(0) a^4(1) \sum_{n=0}^{\infty} a^{2n}(1) \\
&= \left[1 - b(0) \right) + a(1)(1 - a(1)b(0)) \right] - \frac{b(0) a^4(1)}{1 - a^2(1)} = 0
\end{aligned}$$

which may be simplified to

$$b(0) = 1 + a(1) - a^2(1) - a^3(1)$$

Multiplying both sides of this equation by $a(1)$, and using Eq. (P4.15-1), we have

$$\left[1 - a^2(1) \right]^2 = a(1) \left[1 + a(1) - a^2(1) - a^3(1) \right]$$

or

$$2a^4(1) + a^3(1) - 3a^2(1) - a(1) + 1 = 0$$

which may be factored as follows

$$(a(1) + 1)^2 (a(1) - 1)(a(1) - 0.5) = 0$$

Of these roots, clearly we want $a(1) = 0.5$. The value for $b(0)$ is, therefore,

$$b(0) = 1 + a(1) - a^2(1) - a^3(1) = \tfrac{9}{8}$$

Thus, our model for $x(n)$ becomes

$$\hat{x}(n) = \tfrac{9}{8}(0.5)^n u(n)$$

with a squared error of

$$\mathcal{E}_{LS} = (1 - \tfrac{9}{8})^2 + (1 - \tfrac{9}{16})^2 + \sum_{n=2}^{\infty} (0.5)^{2n} = .259$$

4.16 We have a signal $x(n)$ for which we would like to obtain an all-pole model of the form

$$H(z) = \frac{b(0)}{1 + a(1)z^{-1} + a(2)z^{-2}}$$

Using the autocorrelation method, find explicit formulas for $b(0)$, $a(1)$, and $a(2)$ in terms of $r_x(0)$, $r_x(1)$, and $r_x(2)$.

Solution

The autocorrelation normal equations are

$$\begin{bmatrix} r_x(0) & r_x(1) \\ r_x(1) & r_x(0) \end{bmatrix} \begin{bmatrix} a(1) \\ a(2) \end{bmatrix} = - \begin{bmatrix} r_x(1) \\ r_x(2) \end{bmatrix}$$

Solving for the coefficients we have

$$\begin{bmatrix} a(1) \\ a(2) \end{bmatrix} = -\frac{1}{r_x^2(0) - r_x^2(1)} \begin{bmatrix} r_x(0) & -r_x(1) \\ -r_x(1) & r_x(0) \end{bmatrix} \begin{bmatrix} r_x(1) \\ r_x(2) \end{bmatrix} = -\frac{1}{r_x^2(0) - r_x^2(1)} \begin{bmatrix} r_x(0)r_x(1) - r_x(1)r_x(2) \\ -r_x^2(1) + r_x(0)r_x(2) \end{bmatrix}$$

and

$$b^2(0) = r_x(0) + a(1)r_x(1) + a(2)r_x(2)$$

4.17 If one is modeling a signal $x(n)$ whose transform, $X(z)$, contains zeros, then an all-pole model may be used to effectively model a zero with an infinite number of poles. In this problem, we look at how a zero is modeled with the autocorrelation method. Let $x(n) = \delta(n) - \alpha\delta(n-1)$ where $|\alpha| < 1$ and α is real.

(a) Determine the *pth*-order all-pole model $A_p(z)$ for $x(n)$ where p is an arbitrary positive integer, and find the value for the squared error $\mathcal{E}_p$.

(b) For the all-pole model determined in part (a), what is the limit of $A_p(z)$ as $p \to \infty$? What does $\mathcal{E}_p$ converge to as $p \to \infty$? Justify your answers.

(c) Repeat parts (a) and (b) for $|\alpha| > 1$.

Solution _____

(a) With $x(n) = \delta(n) - \alpha\delta(n-1)$, note that the autocorrelation sequence is

$$r_x(k) = (1 + \alpha^2)\delta(k) - \alpha\Big[\delta(k-1) + \delta(k+1)\Big]$$

Therefore, the autocorrelation normal equations for a *pth*-order all-pole model are

$$\begin{bmatrix} 1+\alpha^2 & -\alpha & 0 & \cdots & 0 \\ -\alpha & 1+\alpha^2 & -\alpha & \cdots & 0 \\ 0 & -\alpha & 1+\alpha^2 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & 1+\alpha^2 \end{bmatrix} \begin{bmatrix} 1 \\ a_p(1) \\ a_p(2) \\ \vdots \\ a_p(p) \end{bmatrix} = \epsilon_p \begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

or, in matrix notation,

$$\mathbf{R}_p \mathbf{a}_p = \epsilon_p \mathbf{u}_1$$

Thus, the *pth*-order model is

$$\mathbf{a}_p = \epsilon_p \mathbf{R}_p^{-1} \mathbf{u}_1$$

where $\mathbf{R}_p^{-1}\mathbf{u}_1$ is the first column of the inverse of the $(p+1) \times (p+1)$ autocorrelation matrix, $\mathbf{R}_p$. With $\Delta_j = \det \mathbf{R}_j$, note that $\mathbf{a}_p$ may be written as

$$\mathbf{a}_p = \epsilon_p \mathbf{R}_p^{-1} \mathbf{u}_1 = \epsilon_p \frac{1}{\Delta_p} \begin{bmatrix} \Delta_{p-1} \\ \alpha\Delta_{p-2} \\ \alpha^2\Delta_{p-3} \\ \vdots \\ \alpha^p \end{bmatrix}$$

Furthermore, since the first coefficient of $\mathbf{a}_p$ is equal to one, then we must have

$$\epsilon_p = \Delta_p / \Delta_{p-1}$$

Thus,

$$\mathbf{a}_p = \frac{1}{\Delta_{p-1}} \begin{bmatrix} \Delta_{p-1} \\ \alpha\Delta_{p-2} \\ \alpha^2\Delta_{p-3} \\ \vdots \\ \alpha^p \end{bmatrix}$$

and, for the kth coefficient,

$$a_p(k) = \alpha^k \frac{\Delta_{p-k-1}}{\Delta_{p-1}}$$

Due to the tri-diagonal Toeplitz structure of $\mathbf{R}_p$ we may find a closed form expression for Δ_j as follows. First, note that for $j = 0$ and $j = 1$ we have

$$\begin{aligned} \Delta_0 &= 1 + \alpha^2 \\ \Delta_1 &= \left(1 + \alpha^2\right)^2 - \alpha^2 = 1 + \alpha^2 + \alpha^4 \end{aligned}$$

We may show, by induction, that

$$\Delta_j = 1 + \alpha^2 + \alpha^4 + \cdots + \alpha^{2(j+1)} = \sum_{k=0}^{j+1} \alpha^{2k} = \frac{1 - \alpha^{2(j+2)}}{1 - \alpha^2}$$

Specifically, assume that this relation holds for Δ_{j-1}, and let us verify that it must also hold for Δ_j. Note that the determinant of $\mathbf{R}_j$ is

$$\Delta_j = (1 + \alpha^2)\Delta_{j-1} + \alpha \det \mathbf{A}_{j-1}$$

where

$$\mathbf{A}_{j-1} = \begin{bmatrix} -\alpha & -\alpha & 0 & \cdots & 0 \\ 0 & 1+\alpha^2 & -\alpha & \cdots & 0 \\ 0 & -\alpha & 1+\alpha^2 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & 1+\alpha^2 \end{bmatrix}$$

is the $(j-1) \times (j-1)$ submatrix formed by deleting the second column and the first row of $\mathbf{R}_j$. Since

$$\det \mathbf{A}_{j-1} = -\alpha \Delta_{j-2}$$

then we have the following expression for Δ_j,

$$\Delta_j = (1 + \alpha^2)\Delta_{j-1} - \alpha^2 \Delta_{j-2}$$

Therefore,

$$\Delta_j = (1 + \alpha^2) \sum_{k=0}^{j} \alpha^{2k} - \alpha^2 \sum_{k=0}^{j-1} \alpha^{2k} = \sum_{k=0}^{j+1} \alpha^{2k}$$

as was to be shown. Thus, for the coefficient $a_p(k)$ we have

$$a_p(k) = \alpha^k \frac{\Delta_{p-k-1}}{\Delta_{p-1}} = \alpha^k \frac{1 - \alpha^{2(p-k+1)}}{1 - \alpha^{2(p+1)}} \qquad \text{(P4.17-1)}$$

(b) If we assume that $|\alpha| < 1$, then as $p \to \infty$, the term multiplying α^k in $a_p(k)$ goes to zero, and

$$a_\infty(k) = \alpha^k$$

Therefore, in the limit as $p \to \infty$, the all-pole model is

$$\frac{1}{A_\infty(z)} = \frac{1}{\displaystyle\sum_{k=0}^{\infty} \alpha^k z^{-k}} = 1 - \alpha z^{-1}$$

(c) Now, let us consider what happens when $|\alpha| > 1$. The expression for $a_p(k)$ given in Eq. (P4.17-1) still holds. However, as $p \to \infty$ we have

$$\lim_{p \to \infty} a_p(k) = \lim_{p \to \infty} \alpha^k \frac{1 - \alpha^{2(p-k+1)}}{1 - \alpha^{2(p+1)}} = \lim_{p \to \infty} \alpha^k \frac{\alpha^{2(p-k+1)}}{\alpha^{2(p+1)}} = \alpha^{-k}$$

and, in the limit as $p \to \infty$, the all-pole model is

$$\frac{1}{A_\infty(z)} = \frac{1}{\displaystyle\sum_{k=0}^{\infty} \alpha^{-k} z^{-k}} = 1 - \alpha^{-1} z^{-1}$$

For the squared error, we have

$$\epsilon_p = \frac{\Delta_p}{\Delta_{p-1}}$$

Thus, for $|\alpha| < 1$ we have

$$\epsilon_\infty = \lim_{p \to \infty} \frac{\Delta_p}{\Delta_{p-1}} = 1$$

and, for $|\alpha| > 1$

$$\epsilon_\infty = \lim_{p \to \infty} \frac{\Delta_p}{\Delta_{p-1}} = \alpha^2$$

4.18 Find a closed-form expression for the FIR least squares inverse of length N for each of the following systems.

(a) $G(z) = \dfrac{1}{1 - \alpha z^{-1}}$; $|\alpha| < 1$.

(b) $G(z) = 1 - z^{-1}$.

(c) $G(z) = \dfrac{\alpha - z^{-1}}{1 - \alpha z^{-1}}$; $|\alpha| < 1$.

Solution ───

(a) Since

$$G(z) = \frac{1}{1 - \alpha z^{-1}}$$

is an all-pole filter, the FIR least squares inverse is simply the denominator of $G(z)$,

$$H_N(z) = 1 - \alpha z^{-1}$$

or

$$h_N(n) = \delta(n) - \alpha\delta(n - 1)$$

(b) To find the least squares inverse of

$$G(z) = 1 - z^{-1}$$

we must solve the linear equations

$$\mathbf{R}_g \mathbf{h}_N = g(0)\mathbf{u}_1$$

Since

$$r_g(k) = \begin{cases} 2 & ; \quad k = 0 \\ -1 & ; \quad |k| = 1 \\ 0 & ; \quad |k| > 1 \end{cases}$$

and $g(0) = 1$, then these equations become

$$\begin{bmatrix} 2 & -1 & 0 & \dots \\ -1 & 2 & -1 & \dots \\ 0 & -1 & 2 & \dots \\ \vdots & \vdots & \vdots & \vdots \end{bmatrix} \begin{bmatrix} h_N(0) \\ h_N(1) \\ h_N(2) \\ \vdots \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \end{bmatrix}$$

The solution to these equations (see Example 4.4.5) is of the form

$$h_N(n) = c_1 + n c_2$$

where c_1 and c_2 are constants that are determined by the boundary conditions at $n = 0$ and $n = N - 1$,

$$2h_N(0) - h_N(1) = 1 \quad ; \quad -h_N(N - 2) + 2h_N(N - 1) = 0$$

Using the given form for $h_N(n)$, these boundary conditions become

$$2(c_1) - (c_1 + c_2) = 1 \quad ; \quad -\big[c_1 + (N - 2)c_2\big] + 2\big[c_1 + (N - 1)c_2\big] = 0$$

Solving for c_1 and c_2 we find

$$c_1 = \frac{N}{N+1} \quad ; \quad c_2 = -\frac{1}{N+1}$$

Therefore, for $n = 0, 1, \ldots, N-1$, we have

$$h_N(n) = \frac{N}{N+1} - \frac{n}{N+1} = \frac{N-n}{N+1}$$

and $h_N(n) = 0$ for all other values of n.

(c) Again, to find the least squares inverse, we must solve the linear equations

$$\mathbf{R}_g \mathbf{h}_N = g(0)\mathbf{u_1}$$

Note, however, that $G(e^{j\omega})$ is an all-pass filter,

$$|G(e^{j\omega})| = 1$$

Therefore,

$$r_g(k) = g(k) * g(-k) = \delta(k)$$

and the least-squares inverse is

$$h_N(n) = \delta(n)$$

Note that the least squares inverse is the same for all systems that are related by an all-pass filter.

4.19 An important application of least squares inverse filtering is deconvolution, which is concerned with the recovery of a signal $d(n)$ that has been convolved with a filter $g(n)$

$$x(n) = d(n) * g(n)$$

The problem is to design a filter $h_N(n)$ that may be used to produce an estimate of $d(n)$ from $x(n)$,

$$\widehat{d}(n) = x(n) * h_N(n)$$

One of the difficulties, however, is that noise in the observed signal may be amplified by the filter. For example, if we observe

$$y(n) = d(n) * g(n) + v(n)$$

then the filtered observations become

$$y(n) * h_N(n) = \widehat{d}(n) + v(n) * h_N(n) = \widehat{d}(n) + u(n)$$

where

$$u(n) = v(n) * h_N(n)$$

is the filtered noise. One way to reduce this noise is to design a least squares inverse filter that minimizes

$$\mathcal{E} = \sum_{n=0}^{\infty} |e(n)|^2 + \lambda E\{|u(n)|^2\}$$

where

$$e(n) = \delta(n - n_0) - h_N(n) * g(n)$$

and $\lambda > 0$ is a parameter that is to be selected. Note that for large values of λ, minimizing $\mathcal{E}$ will force a large reduction in the filtered noise at the expense of a decrease in resolution, i.e., larger $e(n)$, whereas smaller values of λ lead to higher resolution and larger noise.

(a) Assume that $v(n)$ is zero-mean white noise with a variance σ_v^2. Show that

$$E\{|u(n)|^2\} = \sigma_v^2 \, \mathbf{h}_N^H \mathbf{h}_N$$

where $\mathbf{h}_N$ is a vector containing the coefficients of the filter $h_N(n)$.

(b) Derive the normal equations that result from minimizing the error

$$\mathcal{E} = \mathbf{e}^H \mathbf{e} + \lambda \, \sigma_v^2 \, \mathbf{h}_N^H \mathbf{h}_N$$

where $\mathbf{e} = [e(0), \ e(1), \ \ldots]^T$, and show that they may be written in the form

$$(\mathbf{R}_g + \alpha \mathbf{I})\mathbf{h}_N = \mathbf{g}_{n_0}^*$$

where $\alpha > 0$ is a *prewhitening parameter* that depends upon the values of λ, and $\mathbf{g}_{n_0}^*$ is the vector on the right-side of Eq. (4.101).

Solution _____

(a) From Eq. (3.90) on p. 101, we have

$$E\{|u(n)|^2\} = \mathbf{h}_N^H \mathbf{R}_v \mathbf{h}_N$$

Since

$$\mathbf{R}_v = \sigma_v^2 \mathbf{I}$$

then

$$E\{|u(n)|^2\} = \sigma_v^2 \, \mathbf{h}_N^H \mathbf{h}_N$$

as was to be shown.

(b) The error that we want to minimize is

$$\mathcal{E} = \sum_{n=0}^{\infty} |e(n)|^2 + \lambda E\{|u(n)|^2\}$$

where

$$e(n) = \delta(n - n_0) - h_N(n) * g(n)$$

and

$$E\{|u(n)|^2\} = \sigma_v^2 \, \mathbf{h}_N^H \mathbf{h}_N = \sigma_v^2 \sum_{l=0}^{N-1} |h_N(l)|^2$$

To minimize the error, we set the derivative with respect to $h_N^*(k)$ equal to zero for $k = 0, 1, \dots, N$,

$$\frac{\partial \mathcal{E}}{\partial h_N^*(k)} = -\sum_{n=0}^{\infty} e(n) g^*(n - k) + \lambda \, \sigma_v^2 \, h_N(k) = 0$$

Substituting for $e(n)$ we have

$$-\sum_{n=0}^{\infty} \Big[\delta(n - n_0) - \sum_{l=0}^{N-1} h_N(l) g(n - l) \Big] g^*(n - k) + \lambda \, \sigma_v^2 \, h_N(k) = 0$$

or,

$$-g^*(n_0 - k) + \sum_{n=0}^{\infty} \Big[\sum_{l=0}^{N-1} h_N(l) g(n - l) \Big] g^*(n - k) + \lambda \, \sigma_v^2 \, h_N(k) = 0$$

Interchanging the order of summation yields

$$-g^*(n_0 - k) + \sum_{l=0}^{N-1} h_N(l) \Big[\sum_{n=0}^{\infty} g(n - l) g^*(n - k) \Big] + \lambda \, \sigma_v^2 \, h_N(k) = 0$$

With

$$r_g(k - l) = \sum_{n=0}^{\infty} g(n - l) g^*(n - k)$$

it follows that

$$\sum_{l=0}^{N-1} h_N(l) r_g(k - l) + \lambda \, \sigma_v^2 \, h_N(k) = g^*(n_0 - k)$$

Written in matrix form, this becomes

$$\left(\mathbf{R}_g + \alpha \mathbf{I}\right)\mathbf{h}_N = \mathbf{g}_{n_0}^*$$

where $\alpha = \lambda \sigma_v^2 > 0$.

4.20 We are given a signal, $x(n)$, that we want to model as the unit sample response of an all-pole filter. We have reason to believe that the signal is periodic and, consequently, the poles of the model should lie on the unit circle. Thus, assuming a second-order model for the signal, the system function is constrained to have the form

$$H(z) = \frac{b(0)}{1 + a(1)z^{-1} + z^{-2}}$$

With $|a(1)| < 2$ this model produces a pair of poles on the unit circle at an angle of θ defined by

$$2\cos\theta = -a(1)$$

(a) Using the autocorrelation method, derive the normal equations that define the value of $a(1)$ that minimizes the error

$$\mathcal{E}_p = \sum_{n=0}^{\infty} e^2(n)$$

(b) Find an expression for the minimum error, $\{\mathcal{E}_p\}_{\min}$.

Solution _____

(a) The error that we want to minimize is

$$\mathcal{E}_p = \sum_{n=0}^{\infty} e^2(n)$$

where

$$e(n)x(n) + a(1)x(n-1) + x(n-2)$$

To find the value of $a(1)$ that minimizes $\mathcal{E}_p$, we set the derivative of $\mathcal{E}$ with respect to $a(1)$ equal to zero,

$$\frac{\partial \mathcal{E}_p}{\partial a(1)} = \sum_{n=0}^{\infty} 2e(n) \frac{\partial e(n)}{\partial a(1)} = 0$$

Since the partial of $e(n)$ with respect to $a(1)$ is $x(n-1)$, the normal equations are

$$\sum_{n=0}^{\infty} e(n)x(n-1) = 0$$

or

$$\sum_{n=0}^{\infty} \Big[x(n) + a(1)x(n-1) + x(n-2) \Big] x(n-1) = 0$$

With

$$r_x(k) = \sum_{n=0}^{\infty} x(n)x(n-k)$$

these become

$$a(1)r_x(0) = -2r_x(1)$$

Therefore,

$$a(1) = -2r_x(1)/r_x(0)$$

(Note that $|a(1)| \le 2$).

(b) To find the minimum error, we have

$$
\begin{aligned}
\{\mathcal{E}_p\}_{\min} &= \sum_{n=0}^{\infty}\Big[x(n)+a(1)x(n-1)+x(n-2)\Big]e(n) = \sum_{n=0}^{\infty}\Big[x(n)+x(n-2)\Big]e(n) \\
&= \sum_{n=0}^{\infty}\Big[x(n)+x(n-2)\Big]\Big[x(n)+a(1)x(n-1)+x(n-2)\Big] \\
&= 2\Big[r_x(0)+a(1)r_x(1)+r_x(2)\Big]
\end{aligned}
$$

4.21 Voiced speech may be modeled as the output of an all-pole filter driven by an impulse train

$$p_{n_0}(n) = \sum_{k=1}^{K} \delta(n - kn_0)$$

where the time between pulses, n_0, is known as the *pitch period*. Suppose that we have a segment of voiced speech, and that we know the pitch period, n_0. We extract a subsequence, $x(n)$, of length $N = 2n_0$ and model this signal as shown in the following figure
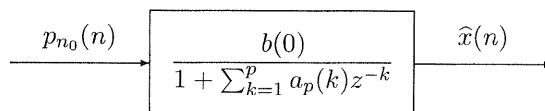
where the input, $p_{n_0}(n)$, consists of two pulses,

$$p_{n_0}(n) = \delta(n) + \delta(n - n_0)$$

Find the normal equations that define coefficients $a_p(k)$ that minimize the error

$$\mathcal{E}_p = \sum_{n=0}^{N-1} e^2(n)$$

where

$$e(n) = a_p(n) * x(n) - b(n) * p_{n_0}(n)$$

and $b(n) = b(0)\delta(n)$.

Solution _____

If we define $a_p(0) = 1$, then the error $e(n)$ is

$$e(n) = a_p(n) * x(n) - b(n) * p_{n_0}(n) = \sum_{l=0}^{p} a_p(l)x(n-l) - b(0)\big[\delta(n) + \delta(n - n_0)\big]$$

and the mean-square error that we want to minimize is

$$\mathcal{E}_p = \sum_{n=0}^{2n_0-1} e^2(n) = \sum_{n=0}^{2n_0-1} \left[\sum_{l=0}^{p} a_p(l)x(n-l) - b(0)\delta(n) - b(0)\delta(n - n_0) \right]^2$$

Setting the derivative with respect to $a_p(k)$ equal to zero, we have

$$\frac{\partial \mathcal{E}_p}{\partial a_p(k)} = \sum_{n=0}^{2n_0-1} 2 \left[\sum_{l=0}^{p} a_p(l)x(n-l) - b(0)\delta(n) - b(0)\delta(n - n_0) \right] x(n-k) = 0$$

If we define

$$r_x(k,l) = \sum_{n=0}^{2n_0-1} x(n-l)x(n-k)$$

then the normal equations become (recall that $a_p(0) = 1$)

$$\sum_{l=1}^{p} a_p(l) r_x(k, l) - b(0) x(-k) - b(0) x(n_0 - k) = -r_x(k, 0) \quad ; \quad k = 1, 2, \ldots, p$$

Assuming that $x(n) = 0$ for $n < 0$, with $\mathbf{x} = [x(n_0 - 1), x(n_0 - 2), \ldots, x(n_0 - p)]^T$, the normal equations may be written in matrix form as follows

$$\mathbf{R}_x \mathbf{a} - b(0) \mathbf{x} = -\mathbf{r}_x$$

Finally, differentiating with respect to $b(0)$ we have

$$\frac{\partial \mathcal{E}}{\partial b(0)} = -\sum_{n=0}^{\infty} 2 \left[\sum_{l=0}^{p} a_p(l) x(n-l) - b(0) \delta(n) - b(0) \delta(n - n_0) \right] \left[\delta(n) + \delta(n - n_0) \right]$$

Thus,

$$x(0) - b(0) + \sum_{l=1}^{p} a_p(l) x(n_0 - l) - b(0) = -x(n_0)$$

or, in vector form, we have

$$\mathbf{x}^T \mathbf{a} - 2b(0) = -x(0) - x(n_0)$$

Putting all of these together in matrix form yields

$$\begin{bmatrix} \mathbf{R}_x & \mathbf{x} \\ \mathbf{x}^T & 1 \end{bmatrix} \begin{bmatrix} \mathbf{a} \\ -2b(0) \end{bmatrix} = - \begin{bmatrix} \mathbf{r}_x \\ x(0) + x(n_0) \end{bmatrix}$$

4.22 You would like to design a linear predictor of a signal $x(n)$ but, due to hardware constraints, are limited to a two-tap predictor. However, since delays can be tolerated in the predictor, you decide to design a predictor of the form

$$\widehat{x}(n) = a(1)x(n - N_1) + a(2)x(n - N_2)$$

where N_1 and N_2 are positive integers. The goal is to find the values of $a(1)$, $a(2)$, N_1, and N_2 that minimize the mean-square error $E\{e^2(n)\}$ where

$$e(n) = x(n) - \widehat{x}(n)$$

Assuming that $x(n)$ is a zero mean wide-sense stationary process, you estimate the autocorrelation of $x(n)$ and find that the values of $r_x(k)$ for $k = 0, 1, \ldots, 7$ are as given in the following table. For $k > 7$, it is determined that the autocorrelation is approximately zero.

Autocorrelation values							
k	0	1	2	3	4	5	6	7
$r_x(k)$	1.0	−0.1	0.0	−0.5	−0.2	0.6	0.2	0.2

   (a) If you were to design an optimum predictor of the form $\widehat{x}(n) = a(1)x(n-1)$, what would be the mean-square error in your prediction of $x(n)$? What about for the predictor $\widehat{x}(n) = a(1)x(n - 3)$?

   (b) Derive a general expression for the minimum mean-square error for a predictor of the form $\widehat{x}(n) = a(1)x(n - N_1)$ with your expression given only in terms of autocorrelations $r_x(k)$. What value of N_1 minimizes the mean-square error?

   (c) Find the values of $a(1)$, $a(2)$, N_1, and N_2 in the two-coefficient linear predictor defined above that minimize the mean-square error $E\{e^2(n)\}$.

   (d) Find the value for the mean-square prediction error for your predictor designed in part (c).

Solution   ───

   (a) With a predictor of the form $\hat{x}(n) = a(1)x(n - 1)$, the value of $a(1)$ that minimizes the mean-square error is

$$a(1) = -\frac{r_x(1)}{r_x(0)}$$

and the minimum error is

$$\xi_{MS} = r_x(0) + a(1)r_x(1) = r_x(0) - \frac{r_x^2(1)}{r_x(0)} = \frac{r_x^2(0) - r_x^2(1)}{r_x(0)}$$

For the given autocorrelations, this becomes

$$\xi_{MS} = 0.99$$

On the other hand, with a predictor of the form $\hat{x}(n) = a(1)x(n-3)$, the value of $a(3)$ that minimizes the mean-square error is

$$a(1) = -\frac{r_x(3)}{r_x(0)}$$

and the minimum mean-square error is

$$\xi_{MS} = r_x(0) + a(1)r_x(3) = r_x(0) - \frac{r_x^2(3)}{r_x(0)} = \frac{r_x^2(0) - r_x^2(3)}{r_x(0)}$$

For the given autocorrelations, this becomes

$$\xi_{MS} = 0.75$$

(b) For a predictor of the form $\hat{x}(n) = a(1)x(n - N_1)$, the optimum value for $a(1)$ is

$$a(1) = -\frac{r_x(N_1)}{r_x(0)}$$

and the minimum mean-square error is

$$\xi_{MS} = r_x(0) + a(1)r_x(N_1) = r_x(0) - \frac{r_x^2(N_1)}{r_x(0)} = \frac{r_x^2(0) - r_x^2(N_1)}{r_x(0)}$$

Therefore, we want to pick the value of N_1 that maximizes the absolute value of $r_x(N_1)$. For the given set of autocorrelations, the value we want to select is $N_1 = 5$.

(c) With a linear predictor of the form

$$\hat{x}(n) = a(1)x(n - N_1) + a(2)x(n - N_2)$$

the normal equations that define the optimum set of predictor coefficients are found by differentiating the mean-square error with respect to $a(1)$ and $a(2)$ and setting the result equal to zero. Thus, with

$$\frac{\partial \xi}{\partial a(1)} = 2E\left\{e(n)x(n - N_1)\right\} = 0$$

$$\frac{\partial \xi}{\partial a(2)} = 2E\left\{e(n)x(n - N_2)\right\} = 0$$

we have

$$E\left\{\left[x(n) - a(1)x(n - N_1) - a(2)x(n - N_2)\right]x(n - N_1)\right\} = 0$$

$$E\left\{\left[x(n) - a(1)x(n - N_1) - a(2)x(n - N_2)\right]x(n - N_2)\right\} = 0$$

Therefore,

$$a(1)r_x(0) + a(2)r_x(N_1 - N_2) = r_x(N_1)$$
$$a(1)r_x(N_2 - N_1) + a(2)r_x(0) = r_x(N_2)$$

or, in matrix form,

$$\begin{bmatrix} r_x(0) & r_x(N_1 - N_2) \\ r_x(N_1 - N_2) & r_x(0) \end{bmatrix} \begin{bmatrix} a(1) \\ a(2) \end{bmatrix} = \begin{bmatrix} r_x(N_1) \\ r_x(N_2) \end{bmatrix}$$

Based on the given set of autocorrelations, the values for N_1 and N_2 that minimize

$$\xi_{MS} = r_x(0) + a(1)r_x(N_1) - a(2)r_x(N_2)$$

are

$$N_1 = 3 \qquad N_2 = 5$$

and the coefficient values are

$$a(1) = -0.5 \qquad a(2) = 0.6$$

(d) The minimum mean-square error is

$$\xi_{MS} = E\{e(n)x(n)\} = r_x(0) - a(1)r_x(N_1) - a(2)r_x(N_2)$$

which, for the given values of N_1, N_2, $a(1)$, and $a(2)$, is

$$\xi = 0.39$$

4.23 If $r_x(0) = 1$, $r_x(1) = 0.5$, and $r_x(2) = 0.75$, find the values of $a(1)$, $a(2)$, and $b(0)$ in the following AR(2) model for $x(n)$,

$$x(n) + a(1)x(n-1) + a(2)x(n-2) = b(0)w(n)$$

where $w(n)$ is unit variance white noise.

Solution

The equations we want to solve, $\mathbf{R}_x \mathbf{a} = -\mathbf{r}_x$, are as follows

$$\begin{bmatrix} 1 & \frac{1}{2} \\ \frac{1}{2} & 1 \end{bmatrix} \begin{bmatrix} a(1) \\ a(2) \end{bmatrix} = - \begin{bmatrix} \frac{1}{2} \\ \frac{3}{4} \end{bmatrix}$$

Therefore,

$$\begin{bmatrix} a(1) \\ a(2) \end{bmatrix} = -\frac{1}{1 - \frac{1}{4}} \begin{bmatrix} 1 & -\frac{1}{2} \\ -\frac{1}{2} & 1 \end{bmatrix} \begin{bmatrix} \frac{1}{2} \\ \frac{3}{4} \end{bmatrix} = -\frac{4}{3} \begin{bmatrix} \frac{1}{8} \\ \frac{1}{2} \end{bmatrix}$$

and

$$b(0) = r_x(0) + a(1)r_x(1) + a(2)r_x(2) = 1 - (\tfrac{1}{6})(\tfrac{1}{2}) - (\tfrac{2}{3})(\tfrac{3}{4}) = \tfrac{5}{12}$$

4.24 Use the method of spectral factorization to find a moving average model of order 2 for a process whose autocorrelation sequence is

$$\mathbf{r}_x = [3, \ 1.5, \ 1]^T$$

Solution _____

Given $\mathbf{r}_x = \begin{bmatrix} 3, & 1.5, & 1 \end{bmatrix}^T$, the power spectrum of $x(n)$ is

$$P_x(z) = 3 + 1.5\big[z^{-1} + z\big] + \big[z^{-2} + z^2\big]$$

which is a fourth-order polynomial. Using the method of spectral factorization, we must factor $P_x(z)$ in the form

$$P_x(z) = \sigma_x^2\big[1 + b(1)z^{-1} + b(2)z^{-2}\big]\big[1 + b(1)z + b(2)z^2\big]$$

Either by trial and error, or using a polynomial factoring program or MATLAB, we find that the factorization is

$$P_x(z) = \frac{1}{2}\big[1 + z^{-1} + 2z^{-2}\big]\big[1 + z + 2z^2\big]$$

Therefore, a moving average model for $x(n)$ is

$$B(z) = \frac{1}{\sqrt{2}}\big(1 + z^{-1} + 2z^{-2}\big)$$

What should be learned from this problem is that spectral factorization is a computationally difficult solution to the MA modeling problem.

4.25 Suppose that the first five values in the autocorrelation sequence for the process $x(n)$ are

$$\mathbf{r}_x = [3,\ 9/4,\ 9/8,\ 9/16,\ 9/32\ \ldots]^T$$

(a) Use the modified Yule-Walker equation method to find an ARMA(1,1) model for $x(n)$.

(b) Are the given values in the autocorrelation sequence consistent with the model that you found in part (a)?

Solution _____

(a) As we did in Example 4.7.1, we would like to find an ARMA(1,1) model for $x(n)$ that has the given autocorrelation values. Since the Yule-Walker equations are

$$\begin{bmatrix} r_x(0) & r_x(1) \\ r_x(1) & r_x(0) \\ r_x(2) & r_x(1) \end{bmatrix} \begin{bmatrix} 1 \\ a_1(1) \end{bmatrix} = \begin{bmatrix} c_1(0) \\ c_1(1) \\ 0 \end{bmatrix}$$

then the modified Yule-Walker equations for $a(1)$ are

$$r_x(1)a(1) = -r_x(2)$$

which gives $a_1(1) = -r_x(2)/r_x(1) = -1/2$.

For the moving average coefficients, we begin by computing $c_1(0)$ and $c_1(1)$ using the Yule-Walker equations as follows

$$\begin{bmatrix} r_x(0) & r_x(1) \\ r_x(1) & r_x(0) \end{bmatrix} \begin{bmatrix} 1 \\ a_1(1) \end{bmatrix} = \begin{bmatrix} c_1(0) \\ c_1(1) \end{bmatrix}$$

With the given values for $r_x(k)$, using $a_1(1) = -1/2$, we find

$$\begin{bmatrix} c_1(0) \\ c_1(1) \end{bmatrix} = \begin{bmatrix} 3 & \frac{9}{4} \\ \frac{9}{4} & 3 \end{bmatrix} \begin{bmatrix} 1 \\ -1/2 \end{bmatrix} = \begin{bmatrix} 15/8 \\ 3/4 \end{bmatrix}$$

and

$$\left[C_1(z)\right]_+ = \tfrac{15}{8} + \tfrac{3}{4}z^{-1}$$

Multiplying by $A_1^*(1/z^*) = 1 - \tfrac{1}{2}z$ we have

$$\left[C_1(z)\right]_+ A_1^*(1/z^*) = \left(\tfrac{15}{8} + \tfrac{3}{4}z^{-1}\right)\left(1 - \tfrac{1}{2}z\right) = -\tfrac{15}{16}z + \tfrac{3}{2} + \tfrac{3}{4}z^{-1}$$

Therefore, the causal part of $P_y(z)$ is

$$\left[P_y(z)\right]_+ = \left[\left[C_1(z)\right]_+ A_1^*(1/z^*)\right]_+ = \tfrac{3}{2} + \tfrac{3}{4}z^{-1}$$

Using the symmetry of $P_y(z)$, we have

$$C_1(z)A_1^*(1/z^*) = B_1(z)B_1^*(1/z^*) = \tfrac{3}{4}z + \tfrac{3}{2} + \tfrac{3}{4}z^{-1}$$

Performing a spectral factorization gives

$$P_y(z) = B(z)B^*(1/z^*) = \frac{3}{4}(1 + z^{-1})(1 + z)$$

so the ARMA(1,1) model is

$$H(z) = \frac{\sqrt{3}}{2}\frac{1 + z^{-1}}{1 - \tfrac{1}{2}z^{-1}}$$

(b) Yes. The model matches $r_x(k)$ for $k = 0, 1, 2$, and for $k > 2$ note that

$$r_x(k) = \tfrac{1}{2} r_x(k-1)$$

which they do.

SOLUTIONS TO CHAPTER 5
The Levinson Recursion

[1] Given the sequence of autocorrelation values,

$$r_x(0) = 1, \quad r_x(1) = 0.8, \quad r_x(2) = 0.5, \quad r_x(3) = 0.1$$

Find the reflection coefficients, Γ_j, the model parameters, $a_j(k)$, and the modeling errors, ϵ_j, for $j = 1, 2, 3$.

Solution _____

This problem is a straightforward application of the Levinson-Durbin recursion. For the sequence of autocorrelations

$$r_x(0) = 1, \quad r_x(1) = 0.8, \quad r_x(2) = 0.5, \quad r_x(3) = 0.1$$

the reflection coefficient sequence is

$$\mathbf{\Gamma} = \begin{bmatrix} -0.8, & 0.3889, & 0.4727 \end{bmatrix}^T$$

The model parameters $a_j(k)$, therefore, are

$$\mathbf{a}_1 = \begin{bmatrix} 1 \\ \Gamma_1 \end{bmatrix} = \begin{bmatrix} 1 \\ -0.8 \end{bmatrix}$$

$$\mathbf{a}_2 = \begin{bmatrix} 1 \\ -0.8 \\ 0 \end{bmatrix} + \Gamma_2 \begin{bmatrix} 0 \\ -0.8 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ -1.1111 \\ 0.3889 \end{bmatrix}$$

and

$$\mathbf{a}_3 = \begin{bmatrix} 1 \\ -1.1111 \\ 0.3889 \\ 0 \end{bmatrix} + \Gamma_3 \begin{bmatrix} 0 \\ 0.3889 \\ -1.1111 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ -0.9273 \\ -0.1364 \\ 0.4727 \end{bmatrix}$$

Finally, for the model errors, we have

$$
\begin{aligned}
\epsilon_0 &= r_x(0) = 1 \\
\epsilon_1 &= \epsilon_0 \left(1 - \Gamma_1^2\right) = 0.36 \\
\epsilon_2 &= \epsilon_1 \left(1 - \Gamma_2^2\right) = 0.3056 \\
\epsilon_3 &= \epsilon_2 \left(1 - \Gamma_3^2\right) = 0.2373
\end{aligned}
$$

[2] Let $A_{p-1}(z)$ be a polynomial of order $(p-1)$ of the form

$$A_{p-1}(z) = 1 + \sum_{k=1}^{p-1} a_{p-1}(k) z^{-k} \qquad \text{(P5.2-1)}$$

and let $\Gamma_1, \Gamma_2, \ldots, \Gamma_{p-1}$ be the reflection coefficients that are generated by the Levinson-Durbin recursion. A pth-order polynomial is formed via the Levinson update equation as follows

$$A_p(z) = A_{p-1}(z) + \Gamma_p z^{-p} A_{p-1}^*(1/z^*)$$

(a) If $|\Gamma_j| < 1$ for $j = 1, \ldots, p-1$ and if $|\Gamma_p| = 1$, what can be said about the location of the zeros of $A_p(z)$?

(b) Suppose $A_p(z)$ may be factored as

$$A_p(z) = \prod_{k=1}^{p} (1 - \alpha_k z^{-1})$$

i.e., the zeros of $A_p(z)$ are at $\alpha_1, \alpha_2, \ldots, \alpha_p$. If $\widetilde{A}_p(z)$ is a polynomial with reflection coefficients $\widetilde{\Gamma}_j$ where

$$\begin{aligned} \widetilde{\Gamma}_j &= \Gamma_j \quad \text{for } j = 1, 2, \ldots, p-1 \\ \widetilde{\Gamma}_p &= 1/\Gamma_p^* \end{aligned}$$

How are the zeros of $\widetilde{A}_p(z)$ related to those of $A_p(z)$?

(c) If we consider $A_p(z)$ in (P5.2-1) to be a continuous function of the reflection coefficient Γ_p, using your results derived in parts (a) and (b), describe how the zeros move in the z-plane as Γ_p is varied in a continuous fashion from some number, say ϵ, to its reciprocal, $1/\epsilon$.

Solution _____

(a) If $|\Gamma_j| < 1$ for $j = 1, 2, \ldots, p-1$, then the zeros of $A_{p-1}(z)$ are *inside* the unit circle. With $|\Gamma_p| = 1$, let $\Gamma_p = e^{j\theta}$. We then have

$$A_p(z) = A_{p-1}(z) + e^{j\theta} z^{-p} A_{p-1}^*(1/z^*)$$

Conjugating and replacing z with $1/z^*$,

$$A_p^*(1/z^*) = A_{p-1}^*(1/z^*) + e^{-j\theta} z^p A_{p-1}(z)$$

Multiplying both sides of this equation by $e^{j\theta} z^{-p}$ gives

$$e^{j\theta} z^{-p} A_p^*(1/z^*) = A_{p-1}(z) + e^{j\theta} z^{-p} A_{p-1}^*(1/z^*)$$

which is equal to $A_p(z)$. Thus, with

$$A_p(z) = e^{j\theta} z^{-p} A_p^*(1/z^*)$$

it follows that if $A_p(z)$ has a zero at $z = z_0$, then $A_p(z)$ must also have a zero at $z = 1/z_0^*$. Now, recall that if $|\Gamma_j| < 1$ for $j = 1, 2, \ldots, p$ then the zeros of $A_p(z)$ must lie *inside* the unit circle, no matter how close $|\Gamma_p|$ may be to one. Since the zeros of $A_p(z)$ move continuously as Γ_p is varied, the zeros of $A_p(z)$, which lie in reciprocal pairs when $\Gamma_p = 1$, must all be on the unit circle.

(b) By definition, we have

$$
\begin{aligned}
A_p(z) &= A_{p-1}(z) + \Gamma_p z^{-p} A_{p-1}^*(1/z^*) \\
\widetilde{A}_p(z) &= A_{p-1}(z) + (1/\Gamma_p^*) z^{-p} A_{p-1}^*(1/z^*)
\end{aligned}
$$

Multiplying both sides of the expression for $\widetilde{A}_p(z)$ by Γ_p^* yields

$$
\Gamma_p^* \widetilde{A}_p(z) = \Gamma_p^* A_{p-1}(z) + z^{-p} A_{p-1}^*(1/z^*)
$$

Conjugating and replacing z with $1/z^*$ this becomes

$$
\Gamma_p \widetilde{A}_p^*(1/z^*) = \Gamma_p A_{p-1}^*(1/z^*) + z^p A_{p-1}(z)
$$

Multiplying both sides by z^{-p} gives

$$
z^{-p} \Gamma_p \widetilde{A}_p^*(1/z^*) = z^{-p} \Gamma_p A_{p-1}^*(1/z^*) + A_{p-1}(z)
$$

and we see that the right-hand side is equal to $A_p(z)$. Therefore,

$$
A_p(z) = \Gamma_p z^{-p} \widetilde{A}_p^*(1/z^*)
$$

and it follows that $\widetilde{A}_p^*(1/z^*)$ is equal to zero when $A_p(z)$ is equal to zero. In other words, the zeros of $A_p(z)$ are *reflected* about the unit circle, so that a zero at $z = z_0$ in $A_p(z)$ becomes a zero at $z = 1/z_0^*$ in $\widetilde{A}_p(z)$.

(c) As Γ_p increases from $\Gamma_p = \epsilon$ to $\Gamma_p = 1$, the zeros of $A_p(z)$ move *towards* the unit circle. When $\Gamma_p = 1$ all of the zeros lie on the unit circle. As Γ_p increases beyond 1, the zeros move outside the unit circle and approach their mirror image location as $\Gamma_p \to 1/\epsilon$.

[3] Let $a_p(k)$ be the set of filter coefficients corresponding to the reflection coefficients Γ_k for $k = 1, 2, \ldots, p$.

(a) Prove that if the reflection coefficients are modulated by $(-1)^k$

$$\widehat{\Gamma}_k = (-1)^k \Gamma_k$$

then the new set of filter coefficients $\widehat{a}_p(k)$ are

$$\widehat{a}_p(k) = (-1)^k a_p(k)$$

(b) Can you generalize the result in part (a) to the case in which

$$\widehat{\Gamma}_k = \alpha^k \Gamma_k$$

where α is a complex number with $|\alpha| = 1$? What about if $|\alpha| < 1$?

Solution _____

(a) We want to show that, if $\widehat{\Gamma}_k = (-1)^k \Gamma_k$, then

$$\widehat{a}_p(k) = (-1)^k a_p(k)$$

or, if we let $\widehat{A}_p(z)$ and $A_p(z)$ be the z-transforms of $\widehat{a}_p(k)$ and $a_p(k)$, respectively, we want to show that

$$\widehat{A}_p(z) = A_p(-z)$$

We begin by noting that, for $p = 1$, we have

$$A_1(z) = 1 + \Gamma_1 z^{-1}$$

and

$$\widehat{A}_1(z) = 1 - \Gamma_1 z^{-1} = A_1(-z)$$

Therefore, let us assume that $\widehat{A}_{p-1}(z) = A_{p-1}(-z)$, and show that $\widehat{A}_p(z) = A_p(-z)$. From the Levinson order update equation we have

$$A_p(z) = A_{p-1}(z) + \Gamma_p z^{-p} A_{p-1}^*(1/z^*)$$

and

$$\begin{aligned}
\widehat{A}_p(z) &= \widehat{A}_{p-1}(z) + \widehat{\Gamma}_p z^{-p} \widehat{A}_{p-1}^*(1/z^*) \\
&= \widehat{A}_{p-1}(z) + (-1)^p \Gamma_p z^{-p} \widehat{A}_{p-1}^*(1/z^*)
\end{aligned}$$

Thus,

$$\widehat{A}_p(z) = A_{p-1}(-z) + \Gamma_p (-z)^{-p} A_{p-1}^*(-1/z^*) = A_p(-z)$$

and we have the desired result.

(b) If $\widehat{\Gamma}_k = \alpha^k \Gamma_k$ with $|\alpha| = 1$, then we may write

$$\widehat{\Gamma}_k = e^{jk\theta} \Gamma_k$$

for some real number θ. As in part (a), for $p = 1$ we have

$$A_1(z) = 1 + \Gamma_1 z^{-1}$$

and

$$\widehat{A}_1(z) = 1 + e^{j\theta} \Gamma_1 z^{-1} = A_1(e^{-j\theta} z)$$

Therefore, let us assume that $\widehat{A}_{p-1}(z) = A_{p-1}(e^{-j\theta} z)$. From the Levinson order update equation we have

$$A_p(z) = A_{p-1}(z) + \Gamma_p z^{-p} A_{p-1}^*(1/z^*)$$

and

$$\begin{aligned} \widehat{A}_p(z) &= \widehat{A}_{p-1}(z) + \widehat{\Gamma}_p z^{-p} \widehat{A}_{p-1}^*(1/z^*) \\ &= \widehat{A}_{p-1}(z) + e^{jp\theta} \Gamma_p z^{-p} \widehat{A}_{p-1}^*(1/z^*) \end{aligned}$$

Thus,

$$\widehat{A}_p(z) = A_{p-1}(e^{-j\theta} z) + e^{jp\theta} \Gamma_p z^{-p} A_{p-1}^*(e^{-j\theta}/z^*) = A_p(e^{-j\theta} z)$$

and we have the desired result,

$$\widehat{A}_p(z) = A_p(e^{-j\theta} z)$$

If $|\alpha| < 1$, then the coefficients change in no predictable manner. Consider, for example, the case of a second-order model,

$$\mathbf{a}_2 = \begin{bmatrix} 1 \\ \Gamma_1(1 + \Gamma_2) \\ \Gamma_2 \end{bmatrix}$$

If $\widehat{\Gamma}_k = \alpha^k \Gamma_k$ and α is real, then

$$\widehat{\mathbf{a}}_2 = \begin{bmatrix} 1 \\ \alpha \Gamma_1(1 + \alpha^2 \Gamma_2) \\ \alpha^2 \Gamma_2 \end{bmatrix}$$

[4] For the reflection coefficient sequence

$$\Gamma_k = \alpha^k \quad ; \quad k = 1, 2, \ldots$$

with $|\alpha| < 1$, prove that the zeros of the polynomials $A_p(k)$ lie on a circle of radius α for every $p \geq 1$. **Solution** _____

We will show this by induction. For a first-order model, we have

$$A_1(z) = 1 + \alpha z^{-1}$$

which has a root at $z = -\alpha$. Now, assume that $\Gamma_k = \alpha^k$ and that $A_{p-1}(z)$ has all of its roots on a circle of radius α. If we multiply the sequence $a_{p-1}(k)$ by α^{-k}, then the z-transform of the sequence is

$$\widetilde{A}_{p-1}(z) = A_{p-1}(\alpha z)$$

and it has all of its zeros *on* the unit circle. Therefore, $\widetilde{A}_{p-1}(z)$ is symmetric,

$$\widetilde{A}_{p-1}(z) = z^{-(p-1)} \widetilde{A}_{p-1}(z^{-1})$$

or

$$A_{p-1}(\alpha z) = z^{-(p-1)} A_{p-1}(\alpha z^{-1})$$

Now, for $A_p(z)$, we have

$$A_p(z) = A_{p-1}(z) + \Gamma_p z^{-p} A_{p-1}(z^{-1}) = A_{p-1}(z) + \alpha^p z^{-p} A_{p-1}(z^{-1})$$

If we multiply the coefficients $a_p(k)$ by α^{-k}, then the z-transform of this sequence becomes

$$\widetilde{A}_p(z) = A_p(\alpha z) = A_{p-1}(\alpha z) + \alpha^p (\alpha z)^{-p} A_{p-1}(\alpha/z) = A_{p-1}(\alpha z) + z^{-p} A_{p-1}(\alpha/z)$$

Since $\widetilde{A}_{-1}p(z)$ is symmetric, then so is $A_p(z)$, and the roots of $A_p(\alpha z)$ either lie on the unit circle or in reciprocal pairs, i.e., if $z = a$ is a root, then so is $z = 1/a$. However, since $A_p(z)$ has all of its roots *inside* the unit circle ($|\Gamma_k| < 1$), then the roots of $\widetilde{A}_p(z)$ must lie on the unit circle. Thus, it follows that the roots of $A_p(z)$ must lie on a circle of radius α.

_____

[5] Without factoring any polynomials, determine whether or not a linear shift-invariant filter with system function

$$H(z) = \frac{1 + 0.8z^{-1} - 0.9z^{-2} + 0.3z^{-3}}{1 - 0.9z^{-1} + 0.8z^{-2} - 0.5z^{-3}}$$

is minimum phase, i.e., all the poles and zeros are inside the unit circle.

Solution

The reflection coefficients corresponding to the numerator coefficients, $\mathbf{b} = \begin{bmatrix} 1, 0.8, -0.9, 0.3 \end{bmatrix}^T$ is

$$\mathbf{\Gamma} = \begin{bmatrix} -4.6522, & -1.2527, & 0.3 \end{bmatrix}^T$$

Since $|\Gamma_1| > 1$ and $|\Gamma_2 > 1$, then the numerator polynomial is not minimum phase (there is at least one root outside the unit circle). However, since the reflection coefficients of the denominator polynomial are $\mathbf{\Gamma} = \begin{bmatrix} -0.4545, & 0.4667, & -0.5 \end{bmatrix}^T$. then the filter is stable.

[**6**] Consider the signal

$$x(n) = \delta(n) + b\delta(n - 1)$$

Suppose that we observe $x(n)$ for $n = 0, 1, \ldots, N$.

(a) Using the autocorrelation method, find the 2nd-order all-pole model for $x(n)$.

(b) Suppose we want to find a p-pole model for $x(n)$. Let Γ_j denote the *jth* reflection coefficient of the lattice filter implementation of the all-pole model. Find a recursion for the reflection coefficients that expresses Γ_j in terms of Γ_{j-1} and Γ_{j-2}.

Solution _____

(a) For the signal $x(n) = \delta(n) + b\delta(n - 1)$, the autocorrelation sequence is

$$r_x(k) = (1 + b^2)\delta(k) + b\delta(k - 1) + b\delta(k + 1)$$

Therefore, using the autocorrelation method, the second-order all-pole model for $x(n)$ is found by solving the normal equations

$$\left[\begin{array}{cc} r_x(0) & r_x(1) \\ r_x(1) & r_x(0) \end{array} \right] \left[\begin{array}{c} a_2(1) \\ a_2(2) \end{array} \right] = - \left[\begin{array}{c} r_x(1) \\ r_x(2) \end{array} \right]$$

Thus,

$$\left[\begin{array}{cc} 1 + b^2 & b \\ b & 1 + b^2 \end{array} \right] \left[\begin{array}{c} a_2(1) \\ a_2(2) \end{array} \right] = - \left[\begin{array}{c} b \\ 0 \end{array} \right]$$

and we have

$$\left[\begin{array}{c} a_2(1) \\ a_2(2) \end{array} \right] = \frac{1}{1 + b^2 + b^4} \left[\begin{array}{c} -b(1 + b^2) \\ b^2 \end{array} \right]$$

(b) The reflection coefficients are given by

$$\Gamma_{j+1} = -\frac{\gamma_j}{\epsilon_j}$$

where

$$\gamma_j = r_x(j + 1) + \sum_{i=1}^{j} a_j(i) r_x(j - i + 1)$$

However, since $r_x(k) = 0$ for $|k| > 1$ then

$$\gamma_j = a_j(j) r_x(1) = b\,\Gamma_j$$

Therefore,

$$\frac{\Gamma_{j+1}}{\Gamma_j} = \frac{\gamma_j}{\gamma_{j-1}} \frac{\epsilon_{j-1}}{\epsilon_j} = \frac{\Gamma_j}{\Gamma_{j-1}} \frac{\epsilon_{j-1}}{\epsilon_j}$$

and, since

$$\frac{\epsilon_j}{\epsilon_{j-1}} = 1 - \Gamma_j^2$$

then

$$\frac{\Gamma_{j+1}}{\Gamma_j} = \frac{\Gamma_j}{\Gamma_{j-1}} \frac{1}{1 - \Gamma_j^2}$$

Thus, a recursion for Γ_j is given by

$$\Gamma_{j+1} = \frac{\Gamma_j^2}{\Gamma_{j-1}(1 - \Gamma_j^2)}$$

[**7**] Suppose we have a data sequence whose z-transform is of the form:

$$X(z) = \frac{G}{1 + \displaystyle\sum_{k=1}^{p} a_p(k) z^{-k}}$$

although the value of p is unknown. The coefficients of the model are computed using Levinson's recursion. How can the value of p be determined by looking at the sequence of reflection coefficients Γ_j for $j = 1, 2, \ldots$?

Solution

Since $x(n)$ may be modeled exactly with an all-pole model of order p, if the autocorrelation sequence for $x(n)$ is known exactly, then Prony's method should yield an exact model. Looking at the reflection coefficients, we would discover that $\Gamma_j = 0$ for all $j \geq p$.

[8] Let $r_x(k)$ be a complex autocorrelation sequence given by

$$\mathbf{r}_x = [2, \ 0.5(1+j), \ 0.5j]^T$$

Use the Levinson-Durbin recursion to solve the autocorrelation normal equations for a second-order all-pole model.

Solution _____

1. First-order model:

$$\Gamma_1 = -\frac{r_x(1)}{r_x(0)} = -\frac{1+j}{4} \quad ; \quad \epsilon_1 = r_x(0)\left[1 - |\Gamma_1|^2\right] = 2\left[1 - \tfrac{1}{8}\right] = \tfrac{7}{4}$$

and

$$\mathbf{a}_1 = \begin{bmatrix} 1 \\ \Gamma_1 \end{bmatrix} = \begin{bmatrix} 1 \\ -\dfrac{1+j}{4} \end{bmatrix}$$

2. Second-order model:

$$\gamma_1 = r_x(2) + a_1(1)r_x(1) = \frac{j}{2} - \frac{1+j}{4}\frac{1+j}{2} = \frac{j}{4}$$

$$\Gamma_2 = -\frac{\gamma_1}{\epsilon_1} = -\frac{j}{4}\frac{4}{7} = -\frac{j}{7}$$

and

$$\mathbf{a}_2 = \begin{bmatrix} 1 \\ -\dfrac{1+j}{4} \\ 0 \end{bmatrix} - \frac{j}{7}\begin{bmatrix} 0 \\ -\dfrac{1-j}{4} \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ -\dfrac{3(1+j)}{14} \\ -\dfrac{j}{7} \end{bmatrix}$$

[9] Determine whether the following statements are *True* or *False*.

(a) If $r_x(k)$ is an autocorrelation sequence with $r_x(k) = 0$ for $|k| > p$, then $\Gamma_k = 0$ for $|k| > p$.

(b) Given an autocorrelation sequence, $r_x(k)$ for $k = 0, \ldots, p$, if the $(p+1) \times (p+1)$ Toeplitz matrix

$$\mathbf{R}_p = \text{Toep}\{r_x(0), r_x(1), \ldots, r_p(p)\}$$

is positive definite, then

$$\mathbf{r}_x = [r_x(0), \ldots, r_x(p), 0, 0, \ldots]^T$$

will *always* be a valid autocorrelation sequence, i.e., extending $r_x(k)$ with zeros is a valid autocorrelation extension.

(c) If $r_x(k)$ is periodic, then Γ_j will be periodic with the same period.

Solution

(a) **False**: If $\Gamma_k = 0$ for $k > p$, then this corresponds to an all-pole model which does *not* have a finite length autocorrelation sequence.

(b) **False**: The condition required for $r_x(p + 1)$ to be a valid extension is given by Eq. (5.98) and is illustrated in Fig. 5.15. Clearly, unless the point $r_x(p + 1) = 0$ lies within the shaded circle, then this is not a valid extension of the partial autocorrelation sequence. For example, suppose $r_x(0) = r_x(1) = 1$. In this case, there is only one valid extension ($r_x(k) = 1$ for all $k > 1$), and the discrete-time Fourier transform of the sequence that is formed by extending this partial autocorrelation sequence with zeros is not a valid power spectrum.

(c) **False**: If $r_x(k)$ is periodic, then $x(n)$ is periodic (perfectly predictable) and the model will have all of its poles on the unit circle ($\Gamma_p = \pm 1$ for some value of p). For example, if $r_x(k) = 1$ for all k, then $\mathbf{\Gamma} = \begin{bmatrix} 1, & 0, & 0, & \ldots \end{bmatrix}^T$.

[**10**] In our discussions of the Levinson-Durbin recursion, we demonstrated the equivalence between the following three sets of parameters:

- $r_x(0), r_x(1), \ldots, r_x(p)$
- $a_p(1), a_p(2), \ldots, a_p(p), b(0)$
- $\Gamma_1, \Gamma_2, \ldots, \Gamma_p, \epsilon_p$

For each of the following signal transformations, determine which of these parameters change and, if possible, describe how.

(a) The signal $x(n)$ is scaled by a constant C,

$$x'(n) = Cx(n)$$

(b) The signal $x(n)$ is modulated by $(-1)^n$,

$$x'(n) = (-1)^n x(n)$$

Solution

(a) With $x'(n) = Cx(n)$ the autocorrelation sequence is scaled by C^2, i.e., $r'_x(k) = C^2 r_x(k)$. For the all-pole model, the numerator $b(0)$ will change by a factor of C, but the location of the poles will not change and, therefore, the reflection coefficients will not change. This may be shown in a number of ways. For example, note that the solution to the normal equations

$$\mathbf{R}_x \bar{\mathbf{a}} = -\mathbf{r}_x$$

will not be affected if all of the correlations are scaled by C^2.

(b) If $x(n)$ is modulated by $(-1)^n$, then $r_x(k)$ is modulated by $(-1)^k$. This results in a modulation of the all-pole parameters by $(-1)^k$ (this may also be seen by looking at the normal equations) and, since $\Gamma_j = a_j(j)$, then the reflection coefficients will be modulated by $(-1)^j$. Since $\epsilon_j = \epsilon_{j-1}(1 - \Gamma_j^2)$, then ϵ_p will be unchanged and, therefore, $b(0)$ will not be affected.

[**11**] The autocorrelation of a signal consisting of a random phase sinusoid in noise is

$$r_x(k) = P\cos(k\omega_0) + \sigma_w^2\delta(k)$$

where ω_0 is the frequency of the sinusoid, P is the power, and σ_w^2 is the variance of the noise. Suppose that we fit an AR(2) model to the data.

(a) Find the coefficients $\mathbf{a}_2 = [1, \ a_2(1), \ a_2(2)]^T$ of the AR(2) model as a function of ω_0, σ_w^2, and P.

(b) Find the reflection coefficients, Γ_1 and Γ_2, corresponding to the AR(2) model.

(c) What are the limiting values of the AR(2) parameters and the reflection coefficients as $\sigma_w^2 \to 0$?

Solution _____

(a) The coefficients of a second-order model for $x(n)$ are found by solving the normal equations

$$\begin{bmatrix} r_x(0) & r_x(1) \\ r_x(1) & r_x(0) \end{bmatrix} \begin{bmatrix} a_2(1) \\ a_2(2) \end{bmatrix} = \begin{bmatrix} r_x(1) \\ r_x(2) \end{bmatrix}$$

Given that the autocorrelation sequence of a random phase sinusoid is

$$r_x(k) = P\cos(k\omega_0) + \sigma_w^2\delta(k)$$

then these equations become

$$\begin{bmatrix} P + \sigma_w^2 & P\cos\omega_0 \\ P\cos\omega_0 & P + \sigma_w^2 \end{bmatrix} \begin{bmatrix} a_2(1) \\ a_2(2) \end{bmatrix} = \begin{bmatrix} P\cos\omega_0 \\ P\cos 2\omega_0 \end{bmatrix}$$

Solving these equations for $a_2(1)$ and $a_2(2)$ we find

$$\begin{bmatrix} a_2(1) \\ a_2(2) \end{bmatrix} = -\frac{1}{(P + \sigma_w^2)^2 - P^2\cos^2\omega_0} \begin{bmatrix} P\cos\omega_0(P + \sigma_w^2) - P^2\cos\omega_0\cos 2\omega_0 \\ -P^2\cos^2\omega_0 + P\cos 2\omega_0(P + \sigma_w^2) \end{bmatrix}$$

(b) For the reflection coefficients, we have

$$\Gamma_1 = -\frac{r_x(1)}{r_x(0)} = -\frac{P\cos\omega_0}{P + \sigma_w^2}$$

and

$$\Gamma_2 = a_2(2) = \frac{P^2\cos^2\omega_0 - P\cos 2\omega_0(P + \sigma_w^2)}{(P + \sigma_w^2)^2 - P^2\cos^2\omega_0}$$

(c) If we let $\sigma_w^2 \to 0$ then

$$\begin{bmatrix} a_2(1) \\ a_2(2) \end{bmatrix} \to \begin{bmatrix} -2\cos\omega_0 \\ 1 \end{bmatrix}$$

and, for the reflection coefficients,

$$\Gamma_1 \to -\cos\omega_0 \quad \text{and} \quad \Gamma_1 \to 1$$

[12] Given that $r_x(0) = 1$ and that the first three reflection coefficients are $\Gamma_1 = 0.5$, $\Gamma_2 = 0.5$, and $\Gamma_3 = 0.25$,

    (a) Find the associated autocorrelation sequence $r_x(1), r_x(2), r_x(3)$.

    (b) Find the autocorrelation sequence $r_x(1), r_x(2), r_x(3)$ for the case in which the reflection coefficient Γ_3 is a free variable, i.e., solve for the autocorrelation values as a function of Γ_3.

    (c) Repeat part (b) when both Γ_2 and Γ_3 are free parameters.

Solution

    (a) From the given reflection coefficient sequence we first find the vectors $\mathbf{a}_1$, $\mathbf{a}_2$, and $\mathbf{a}_3$. For $\mathbf{a}_1$ we have

$$\mathbf{a}_1 = \begin{bmatrix} 1 \\ \Gamma_1 \end{bmatrix} = \begin{bmatrix} 1 \\ 1/2 \end{bmatrix}$$

and for $\mathbf{a}_2$

$$\mathbf{a}_2 = \begin{bmatrix} 1 \\ a_1(1) \\ 0 \end{bmatrix} + \Gamma_2 \begin{bmatrix} 0 \\ a_1(1) \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 3/4 \\ 1/2 \end{bmatrix}$$

and for $\mathbf{a}_3$,

$$\mathbf{a}_3 = \begin{bmatrix} 1 \\ a_2(1) \\ a_2(2) \\ 0 \end{bmatrix} + \Gamma_3 \begin{bmatrix} 0 \\ a_2(2) \\ a_2(1) \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 7/8 \\ 11/16 \\ 1/4 \end{bmatrix}$$

Using the inverse Levinson-Durbin recursion, we may find the autocorrelation sequence. First, we have

$$r_x(1) = -\Gamma_1 r_x(0) = -1/2$$

and

$$\epsilon_1 = r_x(0)(1 - \Gamma_1^2) = 3/4$$

Now, with

$$\gamma_1 = -\Gamma_2 \epsilon_1 = -3/8$$

and

$$\gamma_1 = r_x(2) + a_1(1)r_x(1)$$

solving for $r_x(2)$ we have

$$r_x(2) = \gamma_1 - a_1(1)r_x(1) = -1/8$$

and ϵ_2 is

$$\epsilon_2 = \epsilon_1(1 - \Gamma_2^2) = 9/16$$

Finally, with

$$\gamma_2 = -\Gamma_3 \epsilon_2 = -9/64$$

and

$$\gamma_2 = r_x(3) + a_2(1)r_x(2) + a_2(2)r_x(1)$$

solving for $r_x(3)$ we have

$$r_x(3) = \gamma_2 - a_2(1)r_x(2) - a_2(2)r_x(1) = 13/64$$

Thus, the autocorrelation sequence is

$$\mathbf{r} = [1, \ -1/2, \ -1/8, \ 13/64]^T$$

(b) From part (a) we have

$$r_x(3) = -\epsilon_2 \Gamma_3 - a_2(1)r_x(2) - a_2(2)r_x(1)$$

Since $\epsilon_2 = 9/16$ then

$$r_x(3) = -\tfrac{9}{16}\Gamma_3 - a_2(1)r_x(2) - a_2(2)r_x(1) = -\tfrac{9}{16}\Gamma_3 + \tfrac{11}{32}$$

(c) With Γ_2 a free parameter, the vector $\mathbf{a}_2$ becomes

$$\mathbf{a}_2 = \left[\begin{array}{c} 1 \\ \tfrac{1}{2}(1 + \Gamma_2) \\ \Gamma_2 \end{array} \right]$$

Also, $\epsilon_2 = \epsilon_1(1 - \Gamma_2^2) = \tfrac{3}{4}(1 - \Gamma_2^2)$. Therefore,

$$r_x(3) = -\tfrac{3}{4}\Gamma_3(1 - \Gamma_2^2) - \tfrac{1}{2}(1 + \Gamma_2)r_x(2) - \Gamma_2 r_x(1)$$

However, note that

$$r_x(2) = -\epsilon_1 \Gamma_2 - a_1(1)r_x(1) = -\tfrac{3}{4}\Gamma_2 + \tfrac{1}{4}$$

Thus, with $r_x(1) = -1/2$, we have

$$r_x(3) = -\tfrac{3}{4}\Gamma_3(1 - \Gamma_2^2) - \tfrac{1}{2}(1 + \Gamma_2)(-\tfrac{3}{4}\Gamma_2 + \tfrac{1}{4}) + \tfrac{1}{2}\Gamma_2$$

[**13**] Using the autocorrelation method, an all-pole model of the form

$$H(z) = \frac{b(0)}{1 + a(1)z^{-1} + a(2)z^{-2} + a(3)z^{-3}}$$

has been found for a signal $x(n)$. The constant in the numerator has been chosen so that the autocorrelation of the signal matches the autocorrelation of the model, i.e.,

$$r_x(k) = r_h(k)$$

where $r_h(k)$ is the autocorrelation of $h(n)$. In addition, you know the following about the signal and the model:

1. $r_x(0) = 4$
2. $\Gamma_3 = 0.5$
3. $\Gamma_1 > 0$ and $\Gamma_2 > 0$
4. $x(0) = 1$
5. $\epsilon_3 = 11/16$, where ϵ_3 is the third-order modeling error.
6. $\det(\mathbf{R}_2) = 11$, where $\mathbf{R}_2 = \text{Toep}\{r_x(0), r_x(1), r_x(2)\}$.

Find the values for the model parameters

$$a(1), \quad a(2), \quad a(3), \quad \text{and} \quad b(0)$$

Solution _____

We are given

$$\epsilon_3 = \epsilon_2(1 - \Gamma_3^2) = 11/16$$

Since $\Gamma_3 = 1/2$, we may solve for ϵ_2 as follows,

$$\epsilon_2 = \frac{11/16}{1 - \Gamma_3^2} = 11/12$$

Now, with

$$\det(\mathbf{R}_2) = \epsilon_0 \epsilon_1 \epsilon_2 = 11$$

and $\epsilon_0 = r_x(0) = 4$ we may solve for ϵ_1,

$$\epsilon_1 = \frac{11}{\epsilon_0 \epsilon_2} = \frac{11}{4 \cdot (11/12)} = 3$$

We may now solve for the reflection coefficients. With

$$\epsilon_1 = \epsilon_0(1 - \Gamma_1^2)$$

we have

$$\Gamma_1^2 = 1 - \frac{\epsilon_1}{\epsilon_0} = 1 - \tfrac{3}{4} = \tfrac{1}{4}$$

and, since $\Gamma_1 > 0$, then

$$\Gamma_1 = \tfrac{1}{2}$$

For the second reflection coefficient, we have

$$\epsilon_2 = \epsilon_1 (1 - \Gamma_2^2)$$

Thus,

$$\Gamma_2^2 = 1 - \frac{\epsilon_2}{\epsilon_1} = 1 - \frac{11/12}{3} = \frac{25}{36}$$

and, since $\Gamma_2 > 0$, then

$$\Gamma_2 = \frac{5}{6}$$

Finally, we may generate the coefficients $a_3(k)$ using the step-up recursion,

$$\mathbf{a}_2 = \begin{bmatrix} 1 \\ \Gamma_1 \\ 0 \end{bmatrix} + \Gamma_2 \begin{bmatrix} 0 \\ \Gamma_1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 1/2 \\ 0 \end{bmatrix} + 5/6 \begin{bmatrix} 0 \\ 1/2 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 11/12 \\ 5/6 \end{bmatrix}$$

and

$$\mathbf{a}_3 = \begin{bmatrix} 1 \\ 11/12 \\ 5/6 \\ 0 \end{bmatrix} + 1/2 \begin{bmatrix} 0 \\ 5/6 \\ 11/12 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 4/3 \\ 31/24 \\ 1/2 \end{bmatrix}$$

Therefore,

$$A_3(z) = 1 + \frac{4}{3}z^{-1} + \frac{31}{24}z^{-2} + \frac{1}{2}z^{-3}$$

and

$$b(0) = \sqrt{\epsilon_3} = \sqrt{11/12}$$

[14] The first seven values of the unit sample response, $h(n)$, of a 3rd-order linear shift-invariant filter

$$H(z) = A \frac{1 + b(1)z^{-1} + b(2)z^{-2} + b(3)z^{-3}}{1 + a(1)z^{-1} + a(2)z^{-2} + a(3)z^{-3}}$$

are given by

$$\mathbf{h} = [1,\ 1/4,\ 1/2,\ 1,\ 0,\ 0,\ 7/8]^T$$

Determine whether or not the filter is stable. If more information is needed, state what is required and make the necessary assumptions.

Solution
Since $h(n)$ is the unit sample response of a linear shift-invariant filter having 3 poles and 3 zeros, given the first 7 values of $h(n)$ the system function may be found using the Padé approximation. For the denominator polynomial coefficients we have

$$\begin{bmatrix} h(q) & h(q-1) & \cdots & h(q-p+1) \\ h(q+1) & h(q) & \cdots & h(q-p+2) \\ \vdots & \vdots & & \vdots \\ h(q+p-1) & h(q+p-2) & \cdots & h(q) \end{bmatrix} \begin{bmatrix} a_p(1) \\ a_p(2) \\ \vdots \\ a_p(p) \end{bmatrix} = - \begin{bmatrix} h(q+1) \\ h(q+2) \\ \vdots \\ h(q+p) \end{bmatrix}$$

With $p = q = 3$ and using the given values for $h(n)$ we have

$$\begin{bmatrix} 1 & 1/2 & 1/4 \\ 0 & 1 & 1/2 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} a_3(1) \\ a_3(2) \\ a_3(3) \end{bmatrix} = - \begin{bmatrix} 0 \\ 0 \\ 7/8 \end{bmatrix}$$

Solving for $a_3(k)$ we find

$$a_3(1) = 0 \quad ; \quad a_3(2) = 7/16 \quad ; \quad a_3(3) = -7/8$$

and, therefore,

$$A(z) = 1 + \tfrac{7}{16}z^{-2} - \tfrac{7}{8}z^{-3}$$

To test for stability, we use the Schur-Cohn stability test. With

$$\Gamma_3 = -\tfrac{7}{8}$$

we have, for $\mathbf{a}_2$

$$\begin{bmatrix} a_2(1) \\ a_2(2) \end{bmatrix} = \frac{1}{1 - \Gamma_3^2} \left\{ \begin{bmatrix} a_3(1) \\ a_3(2) \end{bmatrix} - \Gamma_3 \begin{bmatrix} a_3(2) \\ a_3(1) \end{bmatrix} \right\} = \begin{bmatrix} 49/30 \\ 28/15 \end{bmatrix}$$

Therefore,

$$\Gamma_2 = a_2(2) = 28/15 > 1$$

and the filter is unstable.

[**15**] The extendibility problem in power spectrum estimation concerns the issue of whether or not a finite length sequence,

$$r_x(1), r_x(2), r_x(3), \ldots, r_x(p)$$

may be extended (extrapolated) into a legitimate autocorrelation sequence in such a way that

$$P_x(e^{j\omega}) = \sum_{k=-\infty}^{\infty} r_x(k)e^{-jk\omega} \qquad (P5.15\text{-}1)$$

is a valid power spectrum. In other words, is it possible to find values of $r_x(k)$ for $|k| > p$ such that $P_x(e^{j\omega})$ in (P5.15-1) is a non-negative real function of ω?

(a) Develop a procedure that uses Levinson's recursion to test the extendibility of a sequence.

(b) Use your procedure developed in (a) to determine the constraints on a and b that are necessary and sufficient in order for the sequence

$$r_x(0) = 1, \quad r_x(1) = a, \quad r_x(2) = b$$

to be an extendible sequence.

(c) Assuming that the sequence in (b) is extendible, find two different legitimate extensions.

Solution

(a) In order for a sequence $r_x(0), r_x(1), \ldots, r_x(p)$ to be *extendible*, it is necessary and sufficient for the corresponding reflection coefficient sequence to be bounded by one in magnitude. This may be justified as follows. Suppose that $r_x(0), r_x(1), \ldots, r_x(p)$ is extendible. Then the autocorrelation matrix $\mathbf{R}_x$ must be positive definite and, as we know, this implies that $|\Gamma_j| < 1$. Conversely, suppose that $|\Gamma_j| < 1$ for $j = 1, \ldots, p$. Then the sequence

$$\Gamma_j^e = \begin{cases} \Gamma_j & ; \quad j = 1, \ldots, p \\ 0 & ; \quad j > p \end{cases}$$

which corresponds to an all-pole power spectrum, represents a valid extension of the reflection coefficient sequence and a valid autocorrelation sequence extension.

(b) With $r_x(0) = 1$, $r_x(1) = a$, and $r_x(2) = b$, we require that

$$|\Gamma_1| = a < 1$$

and

$$|\Gamma_2| = \left| \frac{r_x(2) + a_1(1)r_x(1)}{\epsilon_1} \right| = \left| \frac{r_x(2) + \Gamma_1 r_x(1)}{r_x(0)\left(1 - \Gamma_1^2\right)} \right| = \left| \frac{b - a^2}{1 - a^2} \right| < 1$$

or

$$2a^2 - 1 < b < 1$$

(c) Two different extensions may be constructed by defining two different extensions of the reflection coefficient sequence. For example, either of the following would be appropriate

$$\mathbf{\Gamma} = \begin{bmatrix} \Gamma_1, \Gamma_2, 0, 0, 0, \ldots \end{bmatrix}^T$$

and

$$\mathbf{\Gamma} = \begin{bmatrix} \Gamma_1, \Gamma_2, 0.5, 0, 0, \ldots \end{bmatrix}^T$$

[**16**] Which of the following autocorrelation sequences are extendible? For those that are extendible, find an extension and determine whether or not the extension is unique.

(a) $\mathbf{r}_x = [1.0, \; 0.6, \; 0.6]^T$

(b) $\mathbf{r}_x = [1.0, \; 0.6, \; -0.6]^T$

(c) $\mathbf{r}_x = [1.0, \; 0.0, \; 1.0]^T$

Solution _____

(a) To check for extendibility, we need to find the reflection coefficients that correspond to the autocorrelations $\mathbf{r}_x = [1.0, \; 0.6, \; 0.6]^T$. With $\Gamma_1 = -0.6$, and

$$\Gamma_2 = -\frac{\gamma_1}{\epsilon_1} = -\frac{0.6 + (-0.6)(0.6)}{1 - (0.6)^2} = -\frac{0.24}{0.64}$$

we see that since $|\Gamma_k| < 1$, then this sequence *is extendible*. Furthermore, the extension is *not* unique. One possible extension is the all-pole autocorrelation that corresponds to the reflection coefficient sequence $\Gamma_3 = \Gamma_4 = \cdots = 0$.

(b) For this sequence, since

$$\Gamma_1 = -\frac{\gamma_1}{\epsilon_1} = -\frac{-0.6 + (-0.6)(0.6)}{1 - (0.6)^2} = -\frac{0.24}{0.64} = 1.5 > 1$$

this sequence *is not extendible*.

(c) The sequence $\mathbf{r}_x = [1.0, \; 0.0, \; 1.0]^T$ *is extendible*, and the extension is unique and given by

$$\mathbf{r}_x = [1, 0, 1, 0, 1, \ldots]^T$$

[**17**] Let $\mathbf{R}_3$ be the symmetric Toeplitz matrix formed from the autocorrelation sequence $r_x(0)$, $r_x(1)$, $r_x(2)$, and $r_x(3)$. If the reflection coefficients that result from applying the Levinson-Durbin recursion to $\mathbf{R}_3$ are

$$\Gamma_1 = \tfrac{1}{2} \qquad \Gamma_2 = \tfrac{1}{3} \qquad \Gamma_3 = \tfrac{1}{4}$$

and if $r_x(0) = 1$, find the determinant of $\mathbf{R}_3$.

Solution _____

The determinant is equal to the product of the errors ϵ_k,

$$\det(\mathbf{R}_p) = \prod_{k=0}^{p} \epsilon_k$$

Since

$$\epsilon_0 = r_x(0) = 1$$

and

$$
\begin{aligned}
\epsilon_1 &= \epsilon_0\left(1 - \Gamma_1^2\right) = \tfrac{3}{4} \\
\epsilon_2 &= \epsilon_1\left(1 - \Gamma_2^2\right) = \tfrac{2}{3} \\
\epsilon_3 &= \epsilon_2\left(1 - \Gamma_3^2\right) = \tfrac{5}{8}
\end{aligned}
$$

then

$$\det \mathbf{R}_3 = \left(\tfrac{3}{4}\right)\left(\tfrac{2}{3}\right)\left(\tfrac{5}{8}\right) = \tfrac{5}{16}$$

[18] Let $r_x(k) = \sigma_x^2 \delta(k) + 1$. Find the reflection coefficients Γ_k for all $k \geq 0$ and find the all-pole models $A_p(k)$ for all $p \geq 1$.

Solution _____

With an autocorrelation sequence

$$r_x(k) = \sigma_x^2 \delta(k) + 1$$

for the first reflection coefficient we have,

$$\Gamma_1 = -\frac{r_x(1)}{r_x(0)} = -\frac{1}{1 + \sigma_x^2}$$

Note that since $r_x(k) = 1$ for $k \geq 1$ then

$$\gamma_j = \begin{bmatrix} r_x(j+1), r_x(j), \ldots, r_x(1) \end{bmatrix} \begin{bmatrix} 1 \\ a_j(1) \\ \vdots \\ a_j(j) \end{bmatrix} = 1 + \sum_{k=1}^{j} a_j(k)$$

From the Levinson-Durbin coefficient update equation we have

$$\begin{bmatrix} 1 \\ a_j(1) \\ \vdots \\ a_j(j-1) \\ a_j(j) \end{bmatrix} = \begin{bmatrix} 1 \\ a_{j-1}(1) \\ \vdots \\ a_{j-1}(j-1) \\ 0 \end{bmatrix} + \Gamma_j \begin{bmatrix} 0 \\ a_{j-1}(j-1) \\ \vdots \\ a_{j-1}(1) \\ 1 \end{bmatrix}$$

Since γ_j is the sum of the coefficients $a_j(k)$, and γ_{j-1} is the sum of the coefficients $a_{j-1}(k)$, then

$$\gamma_j = \gamma_{j-1}(1 + \Gamma_j)$$

With

$$\gamma_j = -\epsilon_j \Gamma_{j+1}$$

it follows that

$$-\epsilon_j \Gamma_{j+1} = -\epsilon_{j-1} \Gamma_j (1 + \Gamma_j)$$

Therefore,

$$\Gamma_{j+1} = \frac{\epsilon_{j-1}}{\epsilon_j} \Gamma_j (1 + \Gamma_j) = \Gamma_j \frac{1 + \Gamma_j}{1 - \Gamma_j^2} = \frac{\Gamma_j}{1 - \Gamma_j} \tag{P5.18-1}$$

which is a recursion for the reflection coefficients. With $\Gamma_1 = -1/(1 + \sigma_x^2)$ it follows that

$$\Gamma_2 = -\frac{1}{2 + \sigma_x^2}$$

and, in general, we have by induction,

$$\Gamma_k = -\frac{1}{k + \sigma_x^2}$$

For the model, note that

$$\mathbf{a}_1 = \begin{bmatrix} 1 \\ \Gamma_1 \end{bmatrix}$$

and

$$\mathbf{a}_2 = \begin{bmatrix} 1 \\ \Gamma_1 \\ 0 \end{bmatrix} + \Gamma_2 \begin{bmatrix} 0 \\ \Gamma_1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ \Gamma_1(1 + \Gamma_2) \\ \Gamma_2 \end{bmatrix}$$

However, from Eq. (P5.18-1) we see that, in general,

$$\Gamma_{j+1} = \frac{\Gamma_j}{1 - \Gamma_j}$$

or

$$\Gamma_{j+1} - \Gamma_j \Gamma_{j+1} = \Gamma_j$$

and

$$\Gamma_{j+1} = \Gamma_j(1 + \Gamma_{j+1}) \tag{P5.18-1}$$

Therefore,

$$\mathbf{a}_2 = \begin{bmatrix} 1 \\ \Gamma_2 \\ \Gamma_2 \end{bmatrix}$$

Our claim, then, is that

$$\mathbf{a}_j = \begin{bmatrix} 1 \\ \Gamma_j \\ \vdots \\ \Gamma_j \end{bmatrix}$$

for all j. To show this, assume that this holds for $\mathbf{a}_{j-1}$, i.e.,

$$\mathbf{a}_{j-1} = \begin{bmatrix} 1, \Gamma_{j-1}, \Gamma_{j-1}, \dots, \Gamma_{j-1} \end{bmatrix}^T$$

Then

$$\mathbf{a}_j = \begin{bmatrix} 1 \\ \Gamma_{j-1} \\ \vdots \\ \Gamma_{j-1} \\ 0 \end{bmatrix} + \Gamma_j \begin{bmatrix} 0 \\ \Gamma_{j-1} \\ \vdots \\ \Gamma_{j-1} \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ \Gamma_{j-1}(1 + \Gamma_j) \\ \vdots \\ \Gamma_{j-1}(1 + \Gamma_j) \\ \Gamma_j \end{bmatrix}$$

and, using Eq. (P5.18-1), the result follows.

[**19**] A pth-order all-pole model for a signal $x(n)$ is parameterized by the $p+1$ parameters ϵ_p and $a_p(k)$. Since the reflection coefficients Γ_j for a stable model are bounded by one in magnitude, they are automatically scaled. In speech processing, therefore, there has been an interest in coding a speech waveform in terms of its reflection coefficient sequence. The relationship between a reflection coefficient sequence and the spectrum, however, is not easily discernable. Consider, for example, the following three reflection coefficient sequences

1. $\Gamma_k = \dfrac{1}{k+1}$

2. $\Gamma_k = -\dfrac{1}{k+1}$

3. $\Gamma_k = \dfrac{(-1)^k}{k+1}$

Although the only difference between these reflection coefficient sequences is in terms of the sign of Γ_k, the power spectra are quite different. For each of these sequences, find the corresponding autocorrelation sequence, $r_x(k)$, and power spectrum, $P_x(e^{j\omega})$, in closed form.

Solution _____

(a) Since the reflection coefficients only determine the autocorrelation sequence to within a scale factor, let us normalize $r_x(k)$ so that $r_x(0) = 1$. Then

$$r_x(1) = -\Gamma_1 = -0.5$$

Note that since

$$\Gamma_k = \frac{1}{k+1} = \frac{k}{k+1}\Gamma_{k-1}$$

then

$$\epsilon_k = \epsilon_{k-1}(1 - \Gamma_k^2) = \epsilon_{k-1}\left[1 - \frac{1}{(k+1)^2}\right] = \epsilon_{k-1}\frac{k}{k+1}\frac{k+2}{k+1}$$

Since $\epsilon_0 = r_x(0) = 1$, solving this recursion for ϵ_k we have

$$\epsilon_k = \frac{1}{2}\frac{k+2}{k+1}$$

Now, let us find $r_x(2)$,

$$r_x(2) = \gamma_1 - a_1(1)r_x(1) = -\epsilon_1\Gamma_2 + \tfrac{1}{4} = -(\tfrac{3}{4})(\tfrac{1}{3}) + \tfrac{1}{4} = 0$$

We will now show by induction that $r_x(k) = 0$ for all $k > 2$. Assume that $r_x(2) = r_x(3) = \cdots = r_x(k) = 0$. Then

$$r_x(k+1) = \gamma_k - a_k(k)r_x(1) = -\epsilon_k\Gamma_{k+1} + \tfrac{1}{2}\Gamma_k = -\frac{1}{2}\frac{k+2}{k+1}\frac{1}{k+2} + \frac{1}{2}\frac{1}{k+1} = 0$$

as was to be shown. Thus, we have

$$r_x(k) = \delta(k) - 0.5\delta(k-1) - 0.5\delta(k+1)$$

which has a power spectrum

$$P_x(e^{j\omega}) = 1 - \cos\omega$$

(b) Note that this reflection coefficient sequence corresponds to the one given in Problem 5.18 with $\sigma_x^2 = 1$. Therefore,
$$r_x(k) = \delta(k) + 1$$
which corresponds to a power spectrum
$$P_x(e^{j\omega}) = 1 + 2\pi\delta(\omega)$$

(c) Since these reflection coefficients are obtained from those in part (a) by multiplying by $(-1)^k$, then the autocorrelation sequence is $(-1)^k$ times the autocorrelation sequence found in part (a), i.e.,
$$r_x(k) = \delta(k) + 0.5\delta(k-1) + 0.5\delta(k+1)$$
which has a power spectrum
$$P_x(e^{j\omega}) = 1 + \cos\omega$$

[20] Let $x(n)$ be a random process with autocorrelation sequence

$$r_x(k) = (0.2)^{|k|}$$

(a) Find the reflection coefficients Γ_1 and Γ_2 for a second-order predictor and draw the lattice filter network.

(b) Suppose that uncorrelated white noise with a variance of $\sigma_x^2 = 0.1$ is added to $x(n)$,

$$y(n) = x(n) + w(n)$$

How do the reflection coefficients change?

(c) Can you make any general statements about the effect on the reflection coefficients when white noise is added to a process?

Solution _____

(a) Since the autocorrelation sequence corresponds to an AR(1) process, the second-order prediction error filter is $A(z) = 1 - 0.2z^{-1}$, and the reflection coefficients are $\Gamma_1 = -0.2$ and $\Gamma_2 = 0$. The corresponding lattice filter is shown below,
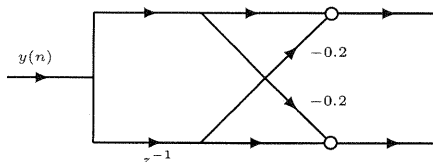
(b) If white noise added to $x(n)$, then the autocorrelation sequence is

$$r_y(k) = r_x(k) + 0.1\delta(k)$$

and the equations for the second-order predictor are

$$\begin{bmatrix} 1.01 & 0.2 \\ 0.2 & 1.01 \end{bmatrix} \begin{bmatrix} a(1) \\ a(2) \end{bmatrix} = \begin{bmatrix} 0.2 \\ 0.04 \end{bmatrix}$$

The predictor, therefore, is

$$\mathbf{a} = \begin{bmatrix} 1, & a(1), & a(2) \end{bmatrix}^T = \begin{bmatrix} 1, & -0.1979, & -0.0004 \end{bmatrix}^T$$

Thus, the first reflection coefficient is smaller than in the no-noise case, and the second reflection coefficient is non-zero.

(c) Generally, the addition of white tends to moves the poles closer to the origin.

[21] The reflection coefficients corresponding to a third-order all-pole model of a signal $x(n)$ are

$$\Gamma_1 = 0.25 \qquad \Gamma_2 = 0.50 \qquad \Gamma_3 = 0.25$$

and the modeling error is given by

$$\epsilon_3 = (15/16)^2$$

(a) Find the direct form filter coefficients, $a_3(k)$, for this third-order model.

(b) Find the autocorrelation values $r_x(1)$, $r_x(2)$, and $r_x(3)$ that led to this model.

(c) If a fourth-order model were found for $x(n)$, what value or values for $r_x(4)$ result in the minimum model error, ϵ_4?

(d) Repeat part (c) and find the value or values for $r_x(4)$ that produce the maximum error, ϵ_4.

Solution _____

(a) Using the step-up recursion, we find that the transversal filter coefficients are

$$\mathbf{a}_1 = \left[\begin{array}{c} 1 \\ \Gamma_1 \end{array} \right] = \left[\begin{array}{c} 1 \\ 1/4 \end{array} \right]$$

$$\mathbf{a}_2 = \left[\begin{array}{c} 1 \\ 1/4 \\ 0 \end{array} \right] + 1/2 \left[\begin{array}{c} 0 \\ 1/4 \\ 1 \end{array} \right] = \left[\begin{array}{c} 1 \\ 3/8 \\ 1/2 \end{array} \right]$$

$$\mathbf{a}_3 = \left[\begin{array}{c} 1 \\ 3/8 \\ 1/2 \\ 0 \end{array} \right] + 1/4 \left[\begin{array}{c} 0 \\ 1/2 \\ 3/8 \\ 1 \end{array} \right] = \left[\begin{array}{c} 1 \\ 1/2 \\ 19/32 \\ 1/4 \end{array} \right]$$

(b) Using the inverse Levinson-Durbin recursion, we have

$$\epsilon_3 = \left(\tfrac{15}{16} \right)^2 = r_x(0)(1 - \tfrac{1}{16})(1 - \tfrac{1}{4})(1 - \tfrac{1}{16}) = r_x(0)(\tfrac{15}{16})^2(\tfrac{3}{4})^2$$

Therefore, $r_x(0) = 4/3$ and

$$r_x(1) = -r_x(0)\Gamma_1 = - \left(\tfrac{4}{3} \right)\left(\tfrac{1}{4} \right) = -\tfrac{1}{3}$$

$$r_x(2) = -\sum_{i=1}^{2} a_2(i) r_x(2-i) = -a_2(1)r_x(1) - a_2(2)r_x(0) = -\tfrac{13}{24}$$

$$r_x(3) = -\sum_{i=1}^{3} a_3(i) r_x(3-i) = -a_3(1)r_x(2) - a_3(2)r_x(1) - a_3(3)r_x(0) = \tfrac{13}{96}$$

(c) From the Levinson-Durbin recursion we have

$$-\epsilon_j \Gamma_{j+1} = r_x(j+1) + \sum_{i=1}^{j} a_j(i) r_x(j-i+1)$$

so,

$$r_x(j+1) = -\epsilon_j \Gamma_{j+1} - \sum_{i=1}^{j} a_j(i) r_x(j-i+1)$$

With $j = 3$, we have for $r_x(4)$

$$r_x(4) = -\epsilon_3 \Gamma_4 - \sum_{i=1}^{3} a_3(i) r_x(4-i) = -\left(\tfrac{15}{16}\right)^2 \Gamma_4 + \left(\tfrac{259}{768}\right) = -(0.8789)\Gamma_4 + 0.3372$$

Now, recall that if $|\Gamma_4| = 1$ then $\epsilon_4 = 0$. Therefore, the error is minimum if

$$r_x(4) = \pm 0.8789 + 0.3372$$

(d) Using the results of part (c), we see that if $\Gamma_4 = 0$ then ϵ_4 takes on the maximum value. Therefore, we have

$$r_x(4) = 0.3372$$

[22] The reflection coefficients for a two-pole model are $\Gamma_1 = 0.25$ and $\Gamma_2 = 0.25$ and the "modeling error" is $\epsilon_2 = 9$.

    (a) If $r_x(3) = 1$ find the modeling error, ϵ_3, for a three-pole model.

    (b) If the signal values, $x(n)$, are multiplied by $1/2$, i.e., $y(n) = 0.5x(n)$, find the reflection coefficients and the modeling error for a two-pole model of $y(n)$.

Solution   _____

    (a) To find the error, ϵ_3, we use

$$\epsilon_3 = r_x(0) \prod_{i=1}^{3}(1 - \Gamma_i^2)$$

We begin by finding $r_x(0)$ as follows

$$\epsilon_2 = r_x(0)(1 - \Gamma_1^2)(1 - \Gamma_2^2) = r_x(0)\left(\tfrac{15}{16}\right)^2 = 9$$

Therefore,

$$r_x(0) = \left(\tfrac{16}{5}\right)^2$$

We now need to find the value for the third reflection coefficient, $\Gamma_3 = \gamma_2/\epsilon_2$, where

$$\gamma_2 = r_x(3) + a_2(1)r_x(2) + a_2(2)r_x(1)$$

First, we must find γ_3 as follows. The first and second order models are

$$\mathbf{a}_1 = \left[\begin{array}{c} 1 \\ \Gamma_1 \end{array}\right] = \left[\begin{array}{c} 1 \\ 1/4 \end{array}\right]$$

$$\mathbf{a}_2 = \left[\begin{array}{c} 1 \\ 1/4 \\ 0 \end{array}\right] + \Gamma_2 \left[\begin{array}{c} 0 \\ 1/4 \\ 1 \end{array}\right] = \left[\begin{array}{c} 1 \\ 5/16 \\ 1/4 \end{array}\right]$$

Computing the autocorrelation values, we find

$$r_x(1) = -\Gamma_1 r_x(0) = -\tfrac{64}{25}$$

$$r_x(2) = -a_2(1)r_x(1) - a_2(2)r_x(0) = -\tfrac{44}{25}$$

Finally, evaluating γ_2 we have

$$\gamma_2 = r_x(3) + a_2(1)r_x(2) + a_2(2)r_x(1) = 1 - \tfrac{11}{20} - \tfrac{16}{25} = -\tfrac{19}{100}$$

Therefore,

$$\Gamma_3 = -\gamma_2/\epsilon_2 = \frac{0.19}{9} = 0.021$$

and

$$\epsilon_3 = \epsilon_2(1 - \Gamma_3^2) = 8.996$$

(b) If the signal values $x(n)$ are scaled by any number, then the all-pole model will not change. Specifically, if $y(n) = 0.5x(n)$ then $r_y(k) = 0.25r_x(k)$. Therefore, since the normal equations are

$$\mathbf{R}_y \mathbf{a} = -\mathbf{r}_y$$

then they are the same as the normal equations for $x(n)$

$$\mathbf{R}_x \mathbf{a} = -\mathbf{r}_x$$

Thus the model coefficients for $x(n)$ and $y(n)$ must be the same and, therefore, $\Gamma_1 = \Gamma_2 = 0.25$ in both cases. However, since

$$\epsilon_2 = r_y(0) \prod_{i=1}^{2} (1 - \Gamma_i^2)$$

then the error will be reduced by $1/4$, i.e.,

$$\epsilon_2 = 9/4 = 2.25$$

[23] You are given the following sequence of autocorrelation values

$$\mathbf{r}_x = [10, \ -1, \ 0.1, \ -1]^T$$

(a) Use the Schur recursion to find the reflection coefficient sequence $\Gamma_1, \Gamma_2, \Gamma_3$.

(b) What is the final generator matrix, $\mathbf{G}_3$, equal to?

(c) Find the modeling error, ϵ_3.

Solution _____

With the autocorrelation sequence $\mathbf{r}_x = \begin{bmatrix} 10, -1, 0.1, -1 \end{bmatrix}^T$, the Schur recursion proceeds as follows:

1. **Step 1**: Define the generator matrix $\mathbf{G}_0$

$$\mathbf{G}_0 = \begin{bmatrix} 0 & r_x(1) & r_x(2) & r_x(3) \\ r_x(0) & r_x(1) & r_x(2) & r_x(3) \end{bmatrix} = \begin{bmatrix} 0 & -1 & 0.1 & -1 \\ 10 & -1 & 0.1 & -1 \end{bmatrix}$$

2. **Step 2**: From the shifted matrix

$$\tilde{\mathbf{G}}_0 = \begin{bmatrix} 0 & -1 & 0.1 & -1 \\ 0 & 10 & -1 & 0.1 \end{bmatrix}$$

it follows that $\Gamma_1 = 0.1$ and

$$\Theta_1 = \begin{bmatrix} 1 & 0.1 \\ 0.1 & 1 \end{bmatrix}$$

3. **Step 3**: Forming the product $\Theta_1 \tilde{\mathbf{G}}_1$ we find

$$\mathbf{G}_1 = \Theta_1 \tilde{\mathbf{G}}_0 = \begin{bmatrix} 1 & 0.1 \\ 0.1 & 1 \end{bmatrix} \begin{bmatrix} 0 & -1 & 0.1 & -1 \\ 0 & 10 & -1 & 0.1 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & -0.99 \\ 0 & 9.9 & -0.99 & 0 \end{bmatrix}$$

4. **Step 4**: From the shifted matrix

$$\tilde{\mathbf{G}}_1 = \begin{bmatrix} 0 & 0 & 0 & -0.99 \\ 0 & 0 & 9.9 & -0.99 \end{bmatrix}$$

we see that $\Gamma_2 = 0$ and

$$\Theta_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

5. **Step 5**: Forming the product $\Theta_2 \tilde{\mathbf{G}}_1$ we obtain

$$\mathbf{G}_2 = \Theta_2 \tilde{\mathbf{G}}_1 = \begin{bmatrix} 0 & 0 & 0 & -0.99 \\ 0 & 0 & 9.9 & -0.99 \end{bmatrix}$$

6. **Step 6**: Finally, forming the shifted matrix $\tilde{\mathbf{G}}_2$,

$$\tilde{\mathbf{G}}_2 = \begin{bmatrix} 0 & 0 & 0 & -0.99 \\ 0 & 0 & 0 & 9.9 \end{bmatrix}$$

we find that $\Gamma_3 = 0.1$ and

$$\Theta_3 = \begin{bmatrix} 1 & 0.1 \\ 0.1 & 1 \end{bmatrix}$$

Therefore, with

$$\mathbf{G}_3 = \Theta_3 \tilde{\mathbf{G}}_2 = \begin{bmatrix} 1 & 0.1 \\ 0.1 & 1 \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 & -0.99 \\ 0 & 0 & 0 & 9.9 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 9.801 \end{bmatrix}$$

we find, that the error is

$$\epsilon_3 = 0.891$$

[24] Let $r_x(0), r_x(1), \ldots, r_x(p)$ be a set of autocorrelation values and let $\mathbf{R}_p$ be the corresponding $(p+1) \times (p+1)$ autocorrelation matrix. Show that

$$1 - \Gamma_p^2 = \frac{\det \mathbf{R}_p \det \mathbf{R}_{p-2}}{\left[\det \mathbf{R}_{p-1}\right]^2}$$

where Γ_p is the pth reflection coefficient.

Solution
Recall that

$$\epsilon_p = \epsilon_{p-1}\left(1 - \Gamma_p^2\right)$$

Therefore,

$$1 - \Gamma_p^2 = \frac{\epsilon_p}{\epsilon_{p-1}}$$

In addition, since

$$\det \mathbf{R}_p = \prod_{k=0}^{p} \epsilon_k$$

then

$$\epsilon_p = \frac{\det \mathbf{R}_p}{\det \mathbf{R}_{p-1}} \quad ; \quad \epsilon_{p-1} = \frac{\det \mathbf{R}_{p-1}}{\det \mathbf{R}_{p-2}}$$

Thus,

$$1 - \Gamma_p^2 = \frac{\det \mathbf{R}_p \det \mathbf{R}_{p-2}}{\left[\det \mathbf{R}_{p-1}\right]^2}$$

[25] If $|\Gamma_j| < 1$, derive a bound for the coefficients δ_j in the split Levinson recursion.

Solution _____

The definition of δ_j, given in terms of the reflection coefficients in Eq. (5.149), is

$$\delta_j = (1 - \Gamma_j)(1 + \Gamma_{j-1})$$

Therefore, if $|\Gamma_j| < 1$ for each j, then

$$0 < 1 - \Gamma_j < 2 \quad ; \quad 0 < 1 - \Gamma_{j-1} < 2$$

and

$$0 < \delta_j < 4$$

[26] Let $h_{n_0}(n)$ be the FIR least squares inverse filter of length N with delay n_0 for a sequence $g(n)$, i.e.,

$$h_{n_0}(n) * g(n) \approx \delta(n - n_0)$$

The coefficients $h_{n_0}(n)$ are the solution to the Toeplitz equations (see p. 174 in Chapter 4)

$$\mathbf{R}_g \mathbf{h}_{n_0} = \mathbf{g}_{n_0} \qquad\qquad\qquad (P5.26\text{-}1)$$

which may be solved efficiently using the Levinson recursion. Since the value for the delay n_0 that produces the smallest least squares error is typically unknown, to find the optimum value for n_0 these equations must be solved for each value of n_0, beginning with $n_0 = 0$. Instead of using the Levinson recursion to solve these equations repeatedly, it is possible to take advantage of the relationship between $\mathbf{g}_{n_0}$ and $\mathbf{g}_{n_0+1}$,

$$\mathbf{g}_{n_0} = \begin{bmatrix} g^*(n_0) \\ g^*(n_0 - 1) \\ \vdots \\ g^*(0) \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \quad ; \quad \mathbf{g}_{n_0+1} = \begin{bmatrix} g^*(n_0 + 1) \\ g^*(n_0) \\ \vdots \\ g^*(1) \\ g^*(0) \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

to derive a recursion for $h_{n_0}(n)$. In this problem we derive this recursion which is known as the *Simpson Sideways Recursion*.

(a) The solution to the normal equations $\mathbf{R}_g \mathbf{h}_{n_0} = \mathbf{g}_{n_0}$ for $n_0 = 0$ may be found using the Levinson-Durbin recursion. Show how to generate the solution for $n_0 = 1$ from the solution for $n_0 = 0$ in less than $4N$ multiplications and divisions where N is the length of the inverse filter $\mathbf{h}_{n_0}$. Note that any information generated in the Levinson-Durbin recursion (for $n_0 = 0$) can be used to construct the new solution.

(b) Generalize the result of part (a) to obtain a recursion that will successively construct the solution for all $n_0 > 0$. Again your method should have less than $4N$ multiplications and divisions at each step.

(c) Write an expression for the error $\mathcal{E}_{n_0}$ at the n_0th step of the recursion in terms of the coefficients $g(n)$ and the coefficients of the least squares inverse filter $h_{n_0}(n)$.

(d) Write a MATLAB program that implements the Simpson sideways recursion.

(e) How can this recursion may be used to find the inverse of a Toeplitz matrix?

Solution _____

(a) If we let $a_\alpha(k)$ be the filter coefficients using a lag α then

$$\widehat{x}(n + 1) = \sum_{k=0}^{p} a_\alpha(k) x(n - \alpha - k)$$

and the optimum filter coefficients are found as follows. With

$$\frac{\partial}{\partial a_\alpha(k)} E\left\{e^2(n)\right\} = 2E\{e(n)x(n-\alpha-k)\} = 0$$

then

$$E\left\{\left[x(n+1) - \sum_{l=0}^{p} a_\alpha(l)x(n-\alpha-l)\right]x(n-\alpha-k)\right\} = 0$$

or

$$\sum_{l=0}^{p} a_\alpha(l)r_x(k-l) = r_x(k+\alpha+1)$$

In matrix form, these equations are

$$\begin{bmatrix} r_x(0) & r_x(1) & \cdots & r_x(p) \\ r_x(1) & r_x(0) & \cdots & r_x(p-1) \\ \vdots & \vdots & & \vdots \\ r_x(p) & r_x(p-1) & \cdots & r_x(0) \end{bmatrix} \begin{bmatrix} a_\alpha(0) \\ a_\alpha(1) \\ \vdots \\ a_\alpha(p) \end{bmatrix} = \begin{bmatrix} r_x(\alpha+1) \\ r_x(\alpha+2) \\ \vdots \\ r_x(\alpha+p+1) \end{bmatrix} \qquad \text{(P5.26-1)}$$

or,

$$\mathbf{R}_p \mathbf{a}_\alpha = \mathbf{r}_x(\alpha)$$

Finally, since the mean square error is $\epsilon_p = E\{e(n)x(n+1)\}$ then

$$\begin{aligned} \epsilon_p &= E\left\{\left[x(n+1) - \sum_{k=0}^{p} a_\alpha(k)x(n-\alpha-k)\right]x(n+1)\right\} \\ &= r_x(0) - \sum_{k=0}^{p} a_\alpha(k)r_x(\alpha+k+1) \end{aligned}$$

(b) Let $\mathbf{f} = [f(0), f(1), \ldots, f(p)]^T$ be the solution to

$$\begin{bmatrix} r_x(0) & r_x(1) & \cdots & r_x(p) \\ r_x(1) & r_x(0) & \cdots & r_x(p-1) \\ \vdots & \vdots & & \vdots \\ r_x(p) & r_x(p-1) & \cdots & r_x(0) \end{bmatrix} \begin{bmatrix} f(0) \\ f(1) \\ \vdots \\ f(p) \end{bmatrix} = - \begin{bmatrix} r_x(1) \\ r_x(2) \\ \vdots \\ r_x(p+1) \end{bmatrix}$$

i.e., $\mathbf{f} = -\mathbf{a}_0$. Note that this equation may be equivalently written as

$$\begin{bmatrix} r_x(0) & r_x(1) & \cdots & r_x(p+1) \\ r_x(1) & r_x(0) & \cdots & r_x(p) \\ r_x(2) & r_x(1) & \cdots & r_x(p-1) \\ \vdots & \vdots & & \vdots \\ r_x(p+1) & r_x(p) & \cdots & r_x(0) \end{bmatrix} \begin{bmatrix} 1 \\ f(0) \\ f(1) \\ \vdots \\ f(p) \end{bmatrix} = \begin{bmatrix} \rho_p \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \qquad \text{(P5.26-2)}$$

where

$$\rho_p = r_x(0) + \sum_{l=0}^{p} f(l)r_x(l+1)$$

is a constant that is related to the modeling error ϵ_p. Due to the Toeplitz structure of $\mathbf{R}_{p+1}$ we may rewrite (P5.26-2) as follows

$$
\begin{bmatrix}
r_x(0) & r_x(1) & \cdots & r_x(p+1) \\
r_x(1) & r_x(0) & \cdots & r_x(p) \\
\vdots & \vdots & & \vdots \\
r_x(p) & r_x(p-1) & \cdots & r_x(1) \\
r_x(p+1) & r_x(p) & \cdots & r_x(0)
\end{bmatrix}
\begin{bmatrix}
f(p) \\
f(p-1) \\
\vdots \\
f(0) \\
1
\end{bmatrix}
=
\begin{bmatrix}
0 \\
0 \\
\vdots \\
0 \\
\rho_p
\end{bmatrix}
\qquad \text{(P5.26-3)}
$$

Now, we begin the derivation of the recursion. Assume that we have the solution to

$$
\begin{bmatrix}
r_x(0) & r_x(1) & \cdots & r_x(p) \\
r_x(1) & r_x(0) & \cdots & r_x(p-1) \\
\vdots & \vdots & & \vdots \\
r_x(p) & r_x(p-1) & \cdots & r_x(0)
\end{bmatrix}
\begin{bmatrix}
a_0(0) \\
a_0(1) \\
\vdots \\
a_0(p)
\end{bmatrix}
=
\begin{bmatrix}
r_x(1) \\
r_x(2) \\
\vdots \\
r_x(p+1)
\end{bmatrix}
\qquad \text{(P5.26-4)}
$$

($\alpha = 0$) and that we want to derive the solution to

$$
\begin{bmatrix}
r_x(0) & r_x(1) & \cdots & r_x(p) \\
r_x(1) & r_x(0) & \cdots & r_x(p-1) \\
\vdots & \vdots & & \vdots \\
r_x(p) & r_x(p-1) & \cdots & r_x(0)
\end{bmatrix}
\begin{bmatrix}
a_1(0) \\
a_1(1) \\
\vdots \\
a_1(p)
\end{bmatrix}
=
\begin{bmatrix}
r_x(2) \\
r_x(3) \\
\vdots \\
r_x(p+2)
\end{bmatrix}
\qquad \text{(P5.26-5)}
$$

($\alpha = 1$). If we augment the vector $\mathbf{a}_0$ with a zero and multiply by the Toeplitz matrix $\mathbf{R}_{p+1}$ then

$$
\begin{bmatrix}
r_x(0) & r_x(1) & \cdots & r_x(p) & r_x(p+1) \\
r_x(1) & r_x(0) & \cdots & r_x(p-1) & r_x(p) \\
\vdots & \vdots & & \vdots & \vdots \\
r_x(p) & r_x(p-1) & \cdots & r_x(0) & r_x(1) \\
r_x(p+1) & r_x(p) & \cdots & r_x(1) & r_x(0)
\end{bmatrix}
\begin{bmatrix}
a_0(0) \\
a_0(1) \\
\vdots \\
a_0(p) \\
0
\end{bmatrix}
=
\begin{bmatrix}
r_x(1) \\
r_x(2) \\
\vdots \\
r_x(p+1) \\
\xi_1
\end{bmatrix}
\qquad \text{(P5.26-6)}
$$

where

$$
\xi_1 = \sum_{k=0}^{p} r_x(p+1-k)a_0(k)
$$

Combining (P5.26-3) and (P5.26-6) we have

$$
\mathbf{R}_{p+1}
\left\{
\begin{bmatrix}
a_0(0) \\
a_0(1) \\
\vdots \\
a_0(p) \\
0
\end{bmatrix}
+ \gamma_1
\begin{bmatrix}
f(p) \\
f(p-1) \\
\vdots \\
f(0) \\
1
\end{bmatrix}
\right\}
=
\begin{bmatrix}
r_x(1) \\
r_x(2) \\
\vdots \\
r_x(p+1) \\
\xi_1 + \gamma_1 \rho_p
\end{bmatrix}
$$

Therefore, if we set

$$
\gamma_1 = [r_x(p+2) - \xi_1]/\rho_p
$$

then

$$
\mathbf{R}_{p+1}
\left\{
\begin{bmatrix}
a_0(0) \\
a_0(1) \\
\vdots \\
a_0(p) \\
0
\end{bmatrix}
+ \gamma_1
\begin{bmatrix}
f(p) \\
f(p-1) \\
\vdots \\
f(0) \\
1
\end{bmatrix}
\right\}
=
\begin{bmatrix}
r_x(1) \\
r_x(2) \\
\vdots \\
r_x(p+1) \\
r_x(p+2)
\end{bmatrix}
$$

Finally, using (P5.26-2) we have

$$
\mathbf{R}_{p+1}\left\{\begin{bmatrix} a_0(0) \\ a_0(1) \\ \vdots \\ a_0(p) \\ 0 \end{bmatrix} + \gamma_1 \begin{bmatrix} f(p) \\ f(p-1) \\ \vdots \\ f(0) \\ 1 \end{bmatrix} + \mu_1 \begin{bmatrix} 1 \\ f(0) \\ \vdots \\ f(p-1) \\ f(p) \end{bmatrix}\right\} = \begin{bmatrix} r_x(1)+c \\ r_x(2) \\ \vdots \\ r_x(p+1) \\ r_x(p+2) \end{bmatrix} \qquad \text{(P5.26-7)}
$$

where c is some constant. Therefore, let

$$
\mu_1 = -a_0(0) - \gamma_1 f(p)
$$

then the first coefficient in the sum of the three vectors above is zero and we may write

$$
\begin{bmatrix} r_x(0) & r_x(1) & \cdots & r_x(p) & r_x(p+1) \\ r_x(1) & r_x(0) & \cdots & r_x(p-1) & r_x(p) \\ \vdots & \vdots & & \vdots & \vdots \\ r_x(p) & r_x(p-1) & \cdots & r_x(0) & r_x(1) \\ r_x(p+1) & r_x(p) & \cdots & r_x(1) & r_x(0) \end{bmatrix}\begin{bmatrix} 0 \\ a_1(0) \\ \vdots \\ a_1(p-1) \\ a_1(p) \end{bmatrix} = \begin{bmatrix} r_x(1)+c \\ r_x(2) \\ \vdots \\ r_x(p+1) \\ r_x(p+2) \end{bmatrix}
$$

where we have defined

$$
\mathbf{a}_1 = \begin{bmatrix} a_1(0) \\ \vdots \\ a_1(p-1) \\ a_1(p) \end{bmatrix} = \begin{bmatrix} a_0(1) \\ \vdots \\ a_0(p) \\ 0 \end{bmatrix} + \gamma_1 \begin{bmatrix} f(p-1) \\ \vdots \\ f(0) \\ 1 \end{bmatrix} + \mu_1 \begin{bmatrix} f(0) \\ \vdots \\ f(p-1) \\ f(p) \end{bmatrix}
$$

If we eliminate the first equation from (P5.26-7) and use the fact that the leading coefficient in the vector multiplied by $\mathbf{R}_{p+1}$ is zero we see that $\mathbf{a}_1$ is the solution to (P5.26-5).

The extension of this approach to solve for $\mathbf{a}_{\alpha+1}$ given $\mathbf{a}_\alpha$ is straight-forward. In particular, for $i = 0, 1, \ldots, p-1$ set

$$
a_{\alpha+1}(i) = a_\alpha(i+1) + \gamma_{\alpha+1} f(p-i-1) + \mu_{\alpha+1} f(i)
$$

and

$$
a_{\alpha+1}(p) = \gamma_{\alpha+1} + \mu_{\alpha+1} f(p)
$$

where

$$
\gamma_{\alpha+1} = [r_x(p+\alpha+2) - \xi_{\alpha+1}]/\rho_p
$$
$$
\mu_{\alpha+1} = -a_\alpha(0) - \gamma_1 f(p)
$$

are coefficients that need to be evaluated for each value of α and where

$$
\rho_p = r_x(0) + \sum_{l=0}^{p} f(l) r_x(l+1)
$$

$$
\xi_1 = \sum_{k=0}^{p} r_x(p+1-k) a_0(k)
$$

are fixed constants that are independent of α.

SOLUTIONS TO CHAPTER 6
Lattice Filters

6.1 Design a two-pole lattice filter that has poles at $re^{j\theta}$ and $re^{-j\theta}$ and draw a carefully labeled flowgraph of your filter.

Solution _____

A second-order filter with roots at $z = re^{\pm\theta}$ has a system function given by

$$A(z) = 1 - 2r\cos\theta z^{-1} + r^2 z^{-2}$$

A first-order lattice filter, on the other hand, has a system function given by $A_1(z) = 1 + \Gamma_1 z^{-1}$, and a second-order filter has a system function

$$\begin{bmatrix} 1 \\ a_2(1) \\ a_2(2) \end{bmatrix} = \begin{bmatrix} 1 \\ \Gamma_1 \\ 0 \end{bmatrix} + \Gamma_2 \begin{bmatrix} 0 \\ \Gamma_1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ \Gamma_1(1 + \Gamma_2) \\ \Gamma_2 \end{bmatrix}$$
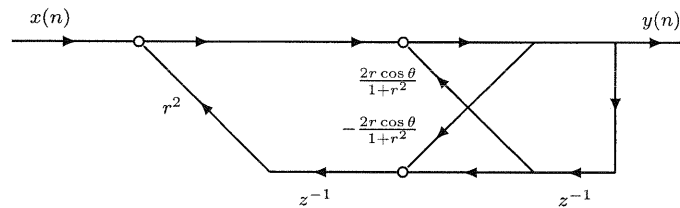
Therefore, $A_2(z) = 1 + \Gamma_1(1 + \Gamma_2)z^{-1} + \Gamma_2 z^{-2}$. It follows that we want to pick

$$\Gamma_2 = r^2$$

and

$$\Gamma_1 = -\frac{2r\cos\theta}{1 + r^2}$$

A flowgraph for this filter is shown below

6.2 Consider the all-pole filter

$$H(z) = \frac{1}{1 - 0.2z^{-1} + 0.4z^{-2} + 0.6z^{-3}}$$

Draw the flowgraph for a lattice filter implementation of $H(z)$ using

(a) A Kelly-Lochbaum lattice filter.

(b) A normalized lattice filter.

(c) A one-muliplier lattice filter.

For each structure, determine the number of multiplies, adds, and delays required to implement the filter and compare them to a direct-form realization of $H(z)$. Based solely on computational considerations, which structure is the most efficient?
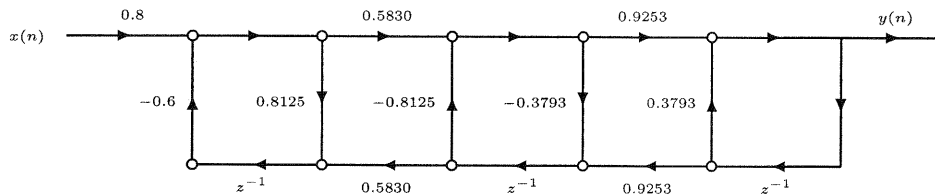
Solution _____

The first step in the implementation of this all-pole filter as a lattice filter is to compute the reflection coefficients. Using the Levinson-Durbin recursion, we find

$$\mathbf{\Gamma} = \begin{bmatrix} -0.3793, & 0.8125, & 0.6000 \end{bmatrix}^T$$

The structures may now be drawn as follows. For the Kelly-Lochbaum lattice filter we have
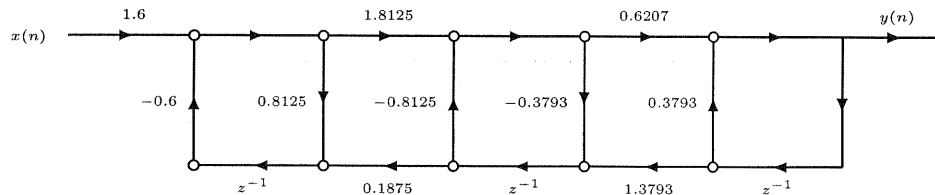
Note that this filter requires 10 multiplications and 5 adds per output point, and it has 3 delays. For the normalized lattice lattice filter we have


which has the same number of multiplies, adds, and delays as the Kelly-Lochbaum filter. Finally, for the one-multiplier lattice filter we have
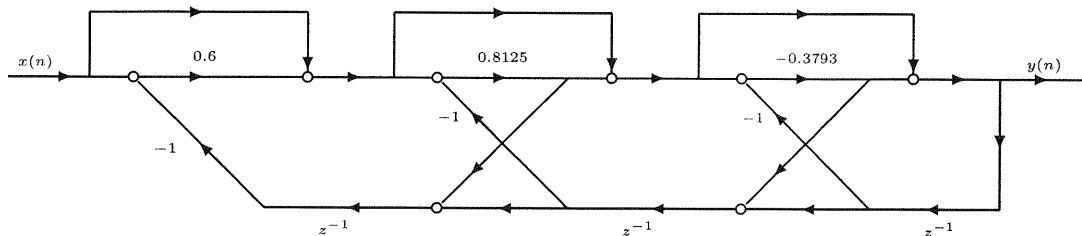
Unlike the previous two filters, the one-multiplier lattice filter only requires 3 multiplications, but 8 adds per output point. As with the other filters, it has 3 delays.

6.3 Find the system function $H(z)$ for the lattice filter given in the figure below.
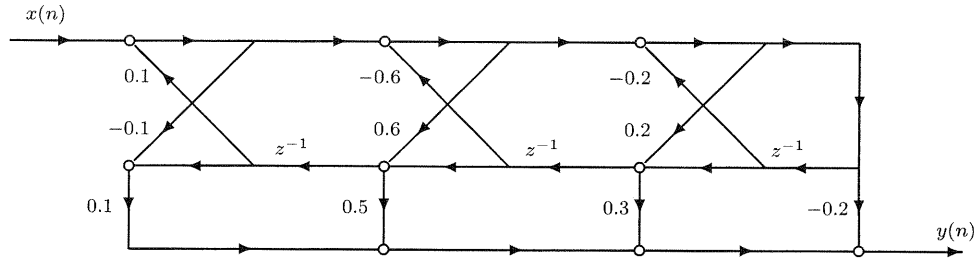
Solution _____

The first step is to use the step-up recursion on the reflection coefficient sequence $\mathbf{\Gamma} = \begin{bmatrix} 0.2, & 0.6, & -0.1 \end{bmatrix}^T$ to find the all-pole models $A_1(z)$, $A_2(z)$, and $A_3(z)$. Beginning with the first-order model, we have

$$\mathbf{a}_1 = \begin{bmatrix} 1 \\ \Gamma_1 \end{bmatrix} = \begin{bmatrix} 1 \\ 0.2 \end{bmatrix}$$

Next, for $\mathbf{a}_2$ we have

$$\mathbf{a}_2 = \begin{bmatrix} 1 \\ 0.2 \\ 0 \end{bmatrix} + \Gamma_2 \begin{bmatrix} 0 \\ 0.2 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 0.32 \\ 0.6 \end{bmatrix}$$

Finally, for $\mathbf{a}_3$ we have

$$\mathbf{a}_3 = \begin{bmatrix} 1 \\ 0.32 \\ 0.6 \\ 0 \end{bmatrix} + \Gamma_3 \begin{bmatrix} 0 \\ 0.6 \\ 0.32 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 0.26 \\ 0.568 \\ -0.1 \end{bmatrix}$$

Therefore, the denominator polynomial is

$$A_3(z) = 1 + 0.26z^{-1} + 0.568z^{-2} - 0.1z^{-3}$$

To find the numerator $B_3(z)$, we use Eq. (6.50),

$$b_q(k) = \sum_{j=k}^{q} c_q(j)a_j^*(j-k)$$

Thus, we have

$$
\begin{aligned}
b_3(3) &= c_3(3)a_3(0) = 0.1 \\
b_3(2) &= c_3(2)a_2(0) + c_3(3)a_3(1) = 0.5260 \\
b_3(1) &= c_3(1)a_1(0) + c_3(2)a_2(1) + c_3(3)a_3(2) = 0.5168 \\
b_3(0) &= c_3(0)a_0(0) + c_3(1)a_1(1) + c_3(2)a_2(2) + c_3(3)a_2(2) = 0.15
\end{aligned}
$$

Therefore, the system function is

$$H(z) = \frac{0.15 + 0.5168z^{-1} + 0.5260z^{-2} + 0.1z^{-3}}{1.0000 + 0.26z^{-1} + 0.568z^{-2} - 0.1z^{-3}}$$

6.4 Sketch a lattice filter structure for each of the following system functions.

(a) $H(z) = \dfrac{2 - z^{-1}}{1 + 0.7z^{-1} + 0.49z^{-2}}$

(b) $H(z) = \dfrac{1 + 1.3125z^{-1} + 0.75z^{-2}}{1 + 0.875z^{-1} + 0.75z^{-2}}$

(c) $H(z) = \dfrac{0.75 + 0.875z^{-1} + z^{-2}}{1 + 0.875z^{-1} + 0.75z^{-2}}$

Solution _____

(a) To implement the filter

$$H(z) = \frac{2 - z^{-1}}{1 + 0.7z^{-1} + 0.49z^{-2}}$$

using a lattice filter structure, we must first find the reflection coefficients corresponding to the denominator polynomial. Using the step-down recursion, we have

$$\Gamma = \begin{bmatrix} 0.4698, & 0.49 \end{bmatrix}^T$$

Next, we must find the coefficients $c_1(k)$ that will produce the numerator polynomial

$$B(z) = 2 - z^{-1}$$

Using the recursion

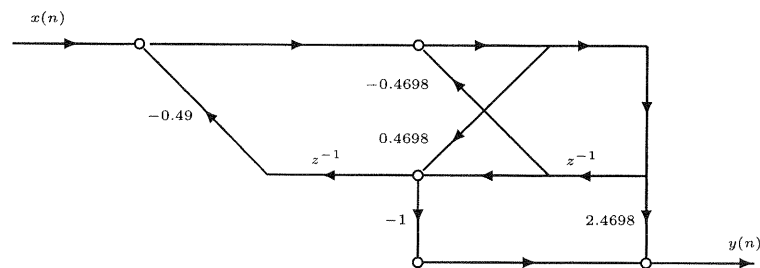$$c_q(k) = b_q(k) - \sum_{j=k+1}^{q} c_q(j)a_j^*(j - k)$$

we have

$$c_1(1) = b_1(1) = -0.1$$

and

$$c_1(0) = b_1(0) - c_1(1)a_1(1) = 2.4698$$

(note that $a_1(1) = \Gamma_1$). Therefore, the lattice filter structure for this system is as shown in the figure below

(b) To find the lattice filter structure for

$$H(z) = \frac{1 + 1.3125z^{-1} + 0.75z^{-2}}{1 + 0.875z^{-1} + 0.75z^{-2}}$$

we first use the step-down recursion to find the reflection coefficients, which are

$$\mathbf{\Gamma} = \begin{bmatrix} 0.5, & 0.75 \end{bmatrix}^T$$

Next, we use the recursion

$$c_q(k) = b_q(k) - \sum_{j=k+1}^{q} c_q(j)a_j^*(j - k)$$

to find the coefficients $c_3(k)$. This recursion requires the first and second-order all-pole models, which are

$$
\begin{aligned}
\mathbf{a}_1 &= \begin{bmatrix} 1, & a_1(1) \end{bmatrix}^T = \begin{bmatrix} 1, & 0.5 \end{bmatrix}^T \\
\mathbf{a}_2 &= \begin{bmatrix} 1, & a_2(1), & a_2(2) \end{bmatrix}^T = \begin{bmatrix} 1, & 0.875, & 0.75 \end{bmatrix}^T
\end{aligned}
$$

Thus, for the coefficients $c_2(k)$, we have

$$
\begin{aligned}
c_2(2) &= b_2(2) = 0.75 \\
c_2(1) &= b_2(1) - c_2(2)a_2(1) = 0.65625 \\
c_2(0) &= b_2(0) - c_2(1)a_1(1) - c_2(2)a_2(2) = 0.109375
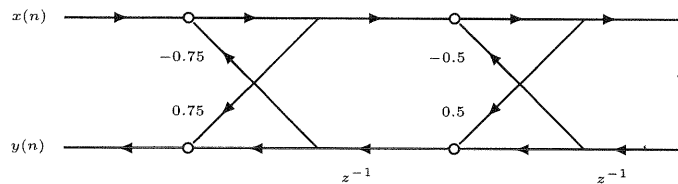\end{aligned}
$$

The lattice filter structure for this system is shown in the figure below
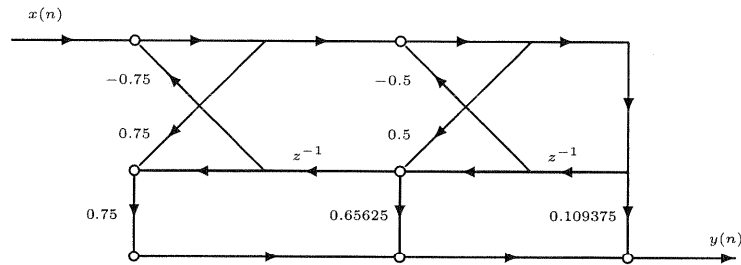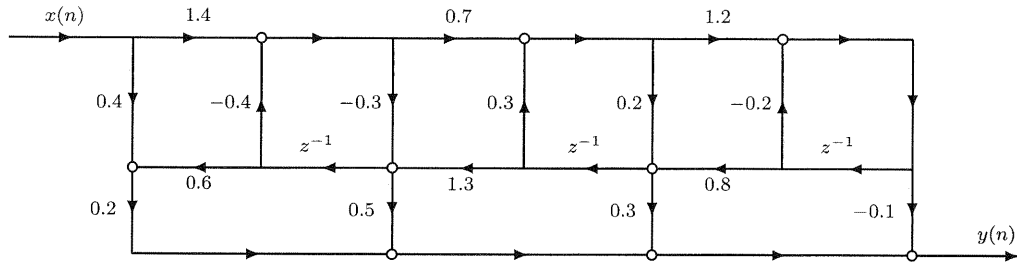

(c) Note that the system function

$$H(z) = \frac{0.75 + 0.875z^{-1} + z^{-2}}{1 + 0.875z^{-1} + 0.75z^{-2}}$$

is an allpass filter. Since the denominator is the same as the system function in part (b), then the reflection coefficients are $\mathbf{\Gamma} = \begin{bmatrix} 0.5, & 0.75 \end{bmatrix}^T$, and the lattice filter is as shown in the figure below

6.5 Determine the system function of the lattice filter shown in the figure below.

Solution _____

This is a third-order lattice filter with 3 poles and 3 zeros. Note that since the coefficients in the upper branch are equal to one plus the coefficient in the adjacent down-going branch, and the coefficients in the lower branch are one minus the coefficient in the adjacent down-going branch, then the all-pole part of the lattice filter uses the Kelly-Lochbaum lattice. Thus, the reflection coefficients are

$$\mathbf{\Gamma} = [-0.2, \; -0.3, \; 0.4]^T$$

Using the step-up recursion, we find

$$\mathbf{a}_1 = \begin{bmatrix} 1 \\ \Gamma_1 \end{bmatrix} = \begin{bmatrix} 1 \\ -0.2 \end{bmatrix}$$

and,

$$\mathbf{a}_2 = \begin{bmatrix} 1 \\ -0.2 \\ 0 \end{bmatrix} + \Gamma_2 \begin{bmatrix} 0 \\ -0.2 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ -0.14 \\ -0.3 \end{bmatrix}$$

and, finally,

$$\mathbf{a}_3 = \begin{bmatrix} 1 \\ -0.14 \\ -0.3 \\ 0 \end{bmatrix} + \Gamma_3 \begin{bmatrix} 0 \\ -0.3 \\ -0.14 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ -0.26 \\ -0.356 \\ 0.4 \end{bmatrix}$$

Therefore, the denominator polynomial is

$$A_3(z) = 1 - 0.26z^{-1} - 0.356z^{-2} + 0.4z^{-3}$$

To find the numerator, $B_3(z)$, we use Eq. (6.50),

$$b_q(k) = \sum_{j=k}^{q} c_q(j) a_j^*(j - k)$$

Note that this requires that we have $A_j(z)$ for $j = 1, 2, 3$. Thus, we have

$$
\begin{aligned}
b_3(3) &= c_3(3)a_3(0) = 0.2 \\
b_3(2) &= c_3(2)a_2(0) + c_3(3)a_3(1) = 0.448 \\
b_3(1) &= c_3(1)a_1(0) + c_3(2)a_2(1) + c_3(3)a_3(2) = 0.1588 \\
b_3(0) &= c_3(0)a_0(0) + c_3(1)a_1(1) + c_3(2)a_2(2) + c_3(3)a_3(3) = 0.23
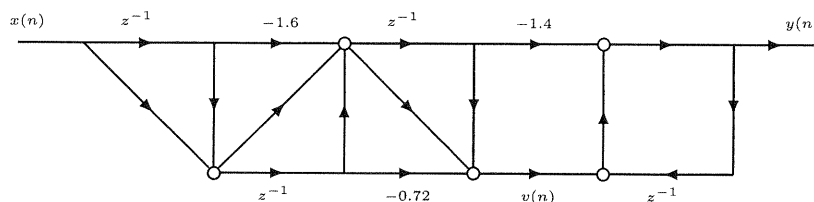\end{aligned}
$$

Therefore, the system function is

$$H(z) = \frac{0.23 + 0.1588z^{-1} + 0.448z^{-2} + 0.2z^{-3}}{1 - 0.26z^{-1} - 0.356z^{-2} + 0.4z^{-3}}$$

6.6 Shown in the figure below is a split lattice filter.

(a) What is the order of the filter (number of poles and zeros)?

(b) Does this filter have minimum phase?

(c) Find the system function $H(z) = Y(z)/X(z)$.

(d) What is the transfer function $V(z)/X(z)$ between $x(n)$ and $v(n)$?

(e) How would you modify this structure to add a zero in $H(z)$ at $z = -1$?

Solution _____

(a) This filter is a secon-order FIR filter. Therefore, it has two zeros and no poles.

(b) The filter parameters for this structure are as follows. First of all, since $2\delta_1 = 1.6$, then

$$\delta_1 = 0.8 = 1 - \Gamma_1$$

which gives

$$\Gamma_1 = 0.2$$

Secondly, since

$$\delta_2 = (1 - \Gamma_2)(1 + \Gamma_1) = 0.72$$

then

$$\Gamma_2 = 0.4$$

Therefore, since $|\Gamma_1| < 1$ and $\Gamma_2| < 1$, then the poles of the filter are inside the unit circle, i.e., it has minimum phase.

(c) With $\Gamma_1 = 0.2$ and $\Gamma_2 = 0.4$, we have

$$\mathbf{a}_2 = \begin{bmatrix} 1 \\ 0.2 \\ 0 \end{bmatrix} + 0.4 \begin{bmatrix} 0 \\ 0.2 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 0.28 \\ 0.4 \end{bmatrix}$$

Therefore, system function $H(z) = Y(z)/X(z)$ is

$$H(z) = 1 + 0.28z^{-1} + 0.4z^{-2}$$

(d) The sequence $v(n)$ is equal to

$$v(n) = \varepsilon_3(n) = e_2^+(n) + e_2^-(n-1)$$

(e) Since
$$y(n) = y(n-1) + \varepsilon_2(n-1) - 1.5\varepsilon_3(n)$$

if we were to add a zero to $H(z)$, then the output would be

$$z(n) = y(n) - y(n-1) = \varepsilon_2(n-1) - 1.5\varepsilon_3(n)$$

Therefore, all that we need to do is remove the feedback path at the end of the structure.

6.7 *True* or *False*: Let $H(z)$ be the system function of a linear shift-invariant filter with coefficients $a_p(k)$ and $b_q(k)$, and let Γ_k and $c_q(k)$ be the coefficients in the lattice filter realization of $H(z)$. If $a_p(k)$ and $b_q(k)$ are modified as follows

$$\widetilde{a}_p(k) = (-1)^k a_p(k) \quad ; \quad \widetilde{b}_p(k) = (-1)^k b_p(k)$$

then the coefficients in the lattice filter are modified in the same way, i.e.,

$$\widetilde{\Gamma}_k = (-1)^k \Gamma_k \quad ; \quad \widetilde{c}_q(k) = (-1)^k c_q(k)$$

Solution _____

This statement is true. As shown in Problem 5.3, given $a_p(k)$ and Γ_k, the reflection coefficients corresponding to the coefficients $\widetilde{a}_p(k) = (-1)^k a_p(k)$ are $\widetilde{\Gamma}_k = (-1)^k \Gamma_k$. Now, for the coefficients $c_q(k)$, note that they are defined recursively in terms of the coefficients $b_q(k)$ and $a_p(k)$ as follows,

$$c_q(k) = b_q(k) - \sum_{j=k+1}^{q} c_q(j) a_j^*(j-k)$$

with

$$c_q q = b_q(q)$$

Therefore, if $\widetilde{b}_q(q) = (-1)^q b_q(q)$, then

$$\widetilde{c}_q(q) = \widetilde{b}_q(q) = (-1)^q b_q(q) = (-1)^q c_q(q)$$

Furthermore, for each k, from $k = q-1, q-2, \ldots, 0$ we have

$$
\begin{aligned}
\widetilde{c}_q(k) &= \widetilde{b}_q(k) - \sum_{j=k+1}^{q} \widetilde{c}_q(j) \widetilde{a}_j^*(j-k) \\
&= (-1)^k b_q(k) - \sum_{j=k+1}^{q} (-1)^j c_q(j)(-1)^{j-k} a_j^*(j-k) \\
&= (-1)^k b_q(k) - \sum_{j=k+1}^{q} (-1)^{2j}(-1)^k c_q(j) a_j^*(j-k) \\
&= (-1)^k \left[b_q(k) - \sum_{j=k+1}^{q} c_q(j) a_j^*(j-k) \right] \\
&= (-1)^k c_q(k)
\end{aligned}
$$

6.8 As shown in Fig. 6.2b, a lattice filter may be used to generate the forward and backward prediction errors, $e_p^+(n)$ and $e_p^-(n)$, respectively.

(a) What is the relationship between the magnitudes of the discrete-time Fourier transforms of $e_p^+(n)$ and $e_p^-(n)$?

(b) Is it possible to design a realizable filter (causal and stable) that will produce the response $e_p^-(n)$ to the input $e_p^+(n)$? If so, describe how and, if not, state why not.

(c) Is it possible to design a realizable filter that will produce the response $e_p^+(n)$ to the input $e_p^-(n)$? If so, describe how and, if not, state why not.

Solution _____

(a) Recall that the backward prediction error, $e_p^-(n)$ is related to the forward prediction error $e_p^+(n)$ by an allpass filter,

$$E_p^-(z) = \frac{A_p^R(z)}{A_p(z)} E_p^+(z) = \prod_{i=1}^{p} \left[\frac{z^{-1} - \alpha_i^*}{1 - \alpha_i z^{-1}} \right] E_p^+(z) = H_{ap}(z) E_p^+(z)$$

Therefore, the DTFT magnitudes of $e_p^+(n)$ and $e_p^-(n)$ are equal,

$$|E_p^-(e^{j\omega})|^2 = |E_p^+(e^{j\omega})|^2$$

(b) If the roots of $A_p(z)$ are inside the unit circle, $|\alpha_k| < 1$, then $e_p^-(n)$ may be generated from $e_p^-(n)$ using the stable and causal allpass filter shown in the figure below,
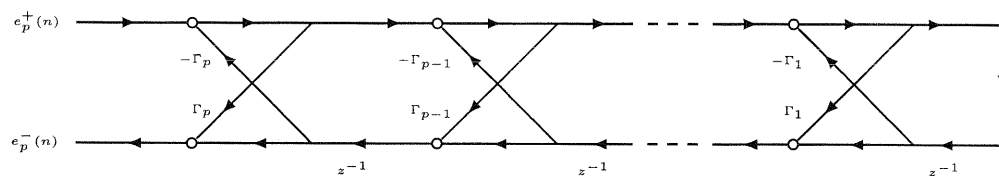
(c) It is not possible, in general, to find a causal and stable filter to produce a response, $e_p^+(n)$, to the input $e_p^-(n)$. The reason is that, if the roots of $A_p(z)$ are inside the unit circle, then $e_p^-(n)$ will be related to $e_p^+(n)$ by an allpass filter that has p poles inside the unit circle and p zeros outside the unit circle. Therefore, the inverse of this allpass has p poles outside the unit circle and p zeros inside, and therefore cannot be causal and stable.

6.9 The all-pole lattice filter in Fig. 6.7b may be used to generate the all-pole approximation $\hat{x}(n)$ to a signal $x(n)$. In this problem, we investigate another use for this filter. Suppose that $e_0^+(n)$ is initialized to one at time $n = 0$, and all of the remaining states are set equal to zero. Determine the output of the all-pole filter for $n = 1, 2, \ldots, N$. Hint: Consider the expression for γ_j given in Eq. (5.10) of Chapter 5.

Solution _____

Before we begin, let is see if we can gain some insight into this problem by computing a few output values. The system that we are considering, for a third-order lattice, is shown in the figure below.
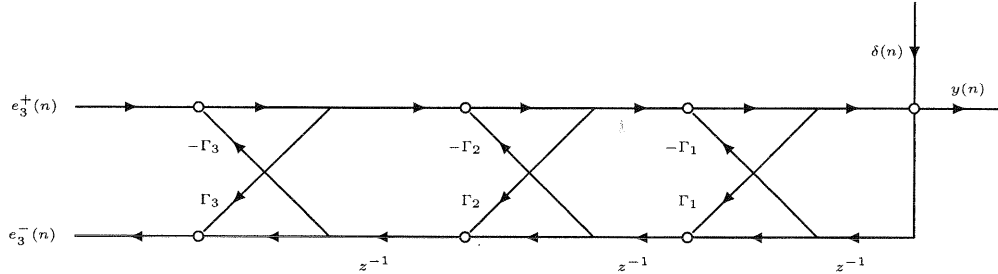
Beginning with time $n = 0$, we can easily generate the following sequence of outputs,

$$
\begin{aligned}
y(0) &= 1 \\
y(1) &= -\Gamma_1 = r_x(1)/r_x(0) \\
y(2) &= \Gamma_1 - \Gamma_2(1 - \Gamma_1^2)
\end{aligned}
$$

Now note that

$$
\begin{aligned}
y(2) &= \Gamma_1^2 - \Gamma_2(1 - \Gamma_1^2) = -\frac{r_x^2(1)}{r_x^2(0)} - \frac{\gamma_1}{\epsilon_1}(1 - \Gamma_1^2) \\
&= -\frac{r_x^2(1)}{r_x^2(0)} - \frac{\gamma_1}{r_x(0)(1 - \Gamma_1^2)}(1 - \Gamma_1^2) = -\frac{r_x^2(1)}{r_x^2(0)} - \frac{\gamma_1}{r_x(0)} \\
&= -\frac{r_x^2(1)}{r_x^2(0)} - \frac{r_x(2) + \Gamma_1 r_x(1)}{r_x(0)} = \frac{r_x(2)}{r_x(0)}
\end{aligned}
$$

Therefore, let us hypothesize that

$$
y(n) = \frac{r_x(n)}{r_x(0)}
$$

We will show this by induction. First, let us make a few observations.

1. The input acts as an initial condition in the delay line that propagates to the left.

2. The last non-zero value in the delay line (at time $n = j$) is

$$
\frac{\epsilon_j}{r_x(0)} = \prod_{k=1}^{j}(1 - \Gamma_j^2)
$$

This is clear from the lattice filter structure.

3. Looking *into* the all-pole lattice filter, at time j, there is effectively an input, at time $j+1$, equal to $-\Gamma_{j+1}\epsilon_j/r_x(0)$ as illustrated in the figure below.
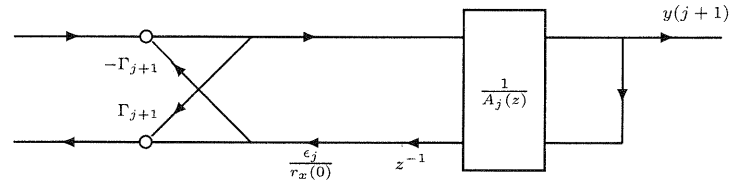
Now let us assume that $y(j) = r_x(j)/r_x(0)$. Since $y(j)$ is equal to a linear combination of outputs, $r_x(j-i)$ multiplied by the filter coefficients, $a_j(i)$, then

$$y(j) = \frac{r_x(j)}{r_x(0)} = \sum_{i=0}^{j} a_j(i)\frac{r_x(j-i)}{r_x(0)}$$

Therefore, it follows from the third observation above that

$$
\begin{aligned}
y(j+1) &= -\frac{\Gamma_{j+1}\epsilon_j}{r_x(0)} + \sum_{i=0}^{j} a_j(i)\frac{r_x(j-i-1)}{r_x(0)} \\
&= \frac{\gamma_j}{r_x(0)} + \sum_{i=0}^{j} a_j(i)\frac{r_x(j-i-1)}{r_x(0)} \\
&= \frac{r_x(j+1)}{r_x(0)}
\end{aligned}
$$

6.10 Consider the following modification to the Burg error

$$\mathcal{E}_j^w = \sum_{n=j}^N w_j(n) \left\{ \left| e_j^+(n) \right|^2 + \left| e_j^-(n) \right|^2 \right\}$$

where $w_j(n)$ is a window that is applied to the forward and backward prediction errors.

(a) Derive an expression that defines the value for the reflection coefficient Γ_j^w that minimizes the modified Burg error $\mathcal{E}_j^w$.

(b) What conditions, if any, are necessary in order to guarantee that the reflection coefficients are bounded by one in magnitude?

Solution _____

(a) Reproducing the derivation of the Burg recursion, we begin by setting the derivative of the error

$$\mathcal{E}_j^w = \sum_{n=j}^N w_j(n) \left\{ \left| e_j^+(n) \right|^2 + \left| e_j^-(n) \right|^2 \right\}$$

with respect to Γ_j^* equal to zero as follows

$$\begin{aligned}
\frac{\partial \mathcal{E}_j^w}{\partial \Gamma_j^*} &= \sum_{n=j}^N w_j(n) \frac{\partial}{\partial \Gamma_j^*} \left\{ \left| e_j^+(n) \right|^2 + \left| e_j^-(n) \right|^2 \right\} \\
&= \sum_{n=j}^N w_j(n) \left\{ e_j^+(n) \left[e_{j-1}^-(n-1) \right]^* + e_j^-(n) \left[e_{j-1}^+(n) \right]^* \right\} = 0
\end{aligned}$$

Substituting for $e_j^+(n)$ and $e_j^-(n)$ and solving for Γ_j gives

$$\Gamma_j = - \frac{2 \displaystyle\sum_{n=j}^N w_j(n) e_{j-1}^+(n) \left[e_{j-1}^-(n-1) \right]^*}{\displaystyle\sum_{n=j}^N w_j(n) \left\{ \left| e_{j-1}^+(n) \right|^2 + \left| e_{j-1}^-(n-1) \right|^2 \right\}}$$

(b) For the Burg recursion, the property that $|\Gamma_j| < 1$ is based on the inequality

$$2 |\langle \mathbf{a}, \mathbf{b} \rangle| \leq \|\mathbf{a}\|^2 + \|\mathbf{b}\|^2$$

We may use this inequality to place constraints on the window, $w_j(n)$, so that the reflection coefficients are bounded by one in magnitude. We do this as follows. If $a = e_{j-1}^+(n)$ and $b = e_{j-1}^+(n-1)$, then

$$2 \left| e_{j-1}^+(n) \left[e_{j-1}^+(n-1) \right]^* \right| \leq \left| e_{j-1}^+(n) \right|^2 + \left| e_{j-1}^+(n-1) \right|^2$$

If $w_j(n) \geq 0$, then we may multiply both sides of this inequality by $w_j(n)$,

$$2 w_j(n) \left| e_{j-1}^+(n) \left[e_{j-1}^+(n-1) \right]^* \right| \leq w_j(n) \left\{ \left| e_{j-1}^+(n) \right|^2 + \left| e_{j-1}^+(n-1) \right|^2 \right\}$$

Summing both sides from $n = j$ to $n + N$ we have

$$2 \sum_{n=j}^{N} w_j(n) \left| e_{j-1}^+(n) \left[e_{j-1}^+(n-1) \right]^* \right| \leq \sum_{n=j}^{N} w_j(n) \left\{ \left| e_{j-1}^+(n) \right|^2 + \left| e_{j-1}^+(n-1) \right|^2 \right\}$$

Since

$$\left| \sum_{n=j}^{N} w_j(n) e_{j-1}^+(n) \left[e_{j-1}^+(n-1) \right]^* \right| \leq \sum_{n=j}^{N} w_j(n) \left| e_{j-1}^+(n) \left[e_{j-1}^+(n-1) \right]^* \right|$$

(recall that $w_j(n) \geq 0$), then

$$2 \left| \sum_{n=j}^{N} w_j(n) e_{j-1}^+(n) \left[e_{j-1}^+(n-1) \right]^* \right| \leq \sum_{n=j}^{N} w_j(n) \left\{ \left| e_{j-1}^+(n) \right|^2 + \left| e_{j-1}^+(n-1) \right|^2 \right\}$$

and it follows that $|\Gamma_j| \leq 1$. Therefore, the constraint on the window is that it be non-negative,

$$w_j(n) \geq 0$$

6.11 Consider the sequence of reflection coefficients Γ_j^B, Γ_j^I, and $\Gamma_j^{\min}$ where Γ_j^I is the reflection coefficient as proposed by Itakura and $\Gamma_j^{\min}$ is the reflection coefficient that is formed using the minimum method (see Section 6.5.3).

(a) Establish the following relationship between these reflection coefficients:

$$|\Gamma_j^{\min}| \le |\Gamma_j^B| \le |\Gamma_j^I|$$

(b) Are there any conditions under which all three reflection coefficients will be the same for all j?

(c) Let Γ_j^M be the set of reflection coefficients corresponding to the all-pole model that is derived from the modified covariance method. Is it possible to upper or lower bound these coefficients in terms of Γ_j^B, Γ_j^I, or $\Gamma_j^{\min}$?

Solution _____

(a) To establish this inequality, we begin by showing that the sign of Γ_j^+ and Γ_j^- is the same. With

$$\Gamma_j^+ = -\frac{\displaystyle\sum_{n=j}^{N} e_{j-1}^+(n)\left[e_{j-1}^-(n-1)\right]^*}{\displaystyle\sum_{n=j}^{N}\left|e_{j-1}^-(n-1)\right|^2}$$

and

$$\Gamma_j^- = -\frac{\displaystyle\sum_{n=j}^{N} e_{j-1}^+(n)\left[e_{j-1}^-(n-1)\right]^*}{\displaystyle\sum_{n=j}^{N}\left|e_{j-1}^+(n)\right|^2}$$

note that the numerators are the same, and the denominators are positive. Therefore, Γ_j^+ and Γ_j^- have the same sign. Now note that since Γ_j^B is the harmonic mean of Γ_j^+ and Γ_j^-, then it is always true that

$$|\Gamma_j^{\min}| \le |\Gamma_j^B|$$

Therefore, all that we need to establish is that

$$|\Gamma_j^B| \le |\Gamma_j^I|$$

We may do this from the definitions of Γ_j^B and Γ_j^I. Since

$$\Gamma_j^B = \frac{2\Gamma_j^+\Gamma_j^-}{\Gamma_j^+ + \Gamma_j^-}$$

and

$$\Gamma_j^I = \sqrt{|\Gamma_j^+|\,|\Gamma_j^-|}$$

then we want to show that

$$\left| \frac{2\Gamma_j^+ \Gamma_j^-}{\Gamma_j^+ + \Gamma_j^-} \right| \leq \sqrt{|\Gamma_j^+| \, |\Gamma_j^-|}$$

or

$$\left| \frac{2\sqrt{|\Gamma_j^+| \, |\Gamma_j^-|}}{\Gamma_j^+ + \Gamma_j^-} \right| \leq 1$$

Squaring both sides we have

$$4|\Gamma_j^+| \, |\Gamma_j^-| \leq |\Gamma_j^+ + \Gamma_j^-|^2$$

Let us assume that $\Gamma_j^+ > 0$ and $\Gamma_j^- > 0$ (the proof is similar if both are negative). Then the inequality that we want to establish becomes

$$0 \leq (\Gamma_j^+ + \Gamma_j^-)^2 - 4\Gamma_j^+ \, \Gamma_j^-$$

However, the right-hand side is equal to $(\Gamma_j^+ - \Gamma_j^-)^2$, which is clearly non-negative, and the desired result follows.

(b) Since Γ_j^B is the harmonic mean of Γ_j^+ and Γ_j^-, then Γ_j^B will be equal to $\Gamma_j^{\min}$ (and also equal to $\Gamma_j^{\max}$) if and only if $\Gamma_j^+ = \Gamma_j^-$. In this case, it will also be true that $\Gamma_j^B = \Gamma_j^I$. In general, however, Γ_j^+ is not equal to Γ_j^-. A special case in which they are equal is when modeling a random process. With

$$\Gamma_j^+ = -\frac{E\left\{ e_{j-1}^+(n)\left[e_{j-1}^-(n-1)\right]^* \right\}}{E\left\{ |e_{j-1}^-(n-1)|^2 \right\}}$$

and

$$\Gamma_j^- = -\frac{E\left\{ e_{j-1}^+(n)\left[e_{j-1}^-(n-1)\right]^* \right\}}{E\left\{ |e_{j-1}^+(n)|^2 \right\}}$$

these reflection coefficients will be equal if the process is wide-sense stationary.

(c) As we saw in Example 6.5.4, the all-pole model that is derived from the modified covariance method is not guaranteed to be stable. Therefore, the reflection coefficients Γ_j^M may be greater than one in magnitude, and the only possible bound that may be valid is

$$|\Gamma_j^I| \leq |\Gamma_j^{MC}|$$

6.12 In Sect. 6.6 it was shown that the backward prediction errors are orthogonal, i.e.,

$$E\left\{e_i^-(n)[e_j^-(n)]^*\right\} = \begin{cases} \epsilon_j & ; \quad i = j \\ 0 & ; \quad i \neq j \end{cases}$$

Establish the following orthogonality conditions:

(a) $E\{e_i^+(n)x^*(n-k)\} = 0 \quad ; \quad 1 \leq k \leq i$

(b) $E\{e_i^-(n)x^*(n-k)\} = 0 \quad ; \quad 0 \leq k \leq i-1$

(c) $E\{e_i^+(n)x^*(n)\} = E\{e_i^-(n)x^*(n-i)\} = \epsilon_i$

Solution

(a) The fact that the forward prediction error $e_i^+(n)$ is orthogonal to the input $x^*(n-k)$ is a restatement of the orthogonality principle. In particular, the forward prediction error $e_i^+(n)$ is, by definition, equal to the difference between $x(n)$ and the prediction of $x(n)$ that is based on the inputs $x(n-1), x(n-2), \ldots, x(n-i)$. By the principle of orthogonality, this error must be orthogonal to $x(n-k)$, for $k = 1, \ldots, i$, and the property is established.

(b) As in part (a), this property also follows from the orthogonality principle. The backward prediction error $e_i^-(n)$ is equal to the difference between $x(n-i)$ and the prediction of $x(n-i)$ given the inputs $x(n), x(n-1), \ldots, x(n-i+1)$. By the orthogonality principle, this error must be orthogonal to $x(n-k)$ for $k = 0, 1, \ldots, i-1$.

(c) First, we show that

$$E\left\{e_i^+(n)x^*(n)\right\} = \epsilon_i$$

Note that $x(n)$ is equal to the forward prediction error plus the prediction of $x(n)$,

$$x(n) = e_i^+(n) + \widehat{x}(n)$$

Since the prediction is orthogonal to the error $e_i^+(n)$, it follows that

$$E\left\{e_i^+(n)x^*(n)\right\} = E\left\{e_i^+(n)[e_i^+(n)]^*\right\} = \epsilon_i$$

To prove the second part of the property,

$$E\left\{e_i^-(n)x^*(n-i)\right\} = \epsilon_i$$

note that $x(n-i)$ is equal to the backward prediction error, $e_i^-(n)$ plus the prediction of $x(n-i)$,

$$x(n-i) = e_i^-(n) + \widehat{x}(n-i)$$

Since this prediction is orthogonal to the error $e_i^-(n)$, it follows that

$$E\left\{e_i^-(n)x^*(n-i)\right\} = E\left\{e_i^-(n)[e_i^-(n)]^*\right\} = \epsilon_i$$

6.13 In this problem it will be shown that the prediction error filters, $A_n(z)$, are *orthogonal* on the unit circle. Specifically, let $P_x(e^{j\omega})$ be the power spectrum of a zero mean random process $x(n)$ and let $A_n(z)$ be the system function of the nth-order prediction error filter. Define the polynomials

$$\phi_n(z) = z^n + a_n(1)z^{n-1} + \cdots + a_n(n-1)z + a_n(n)$$

Show that these polynomials satisfy the orthogonality property

$$\int_{-\pi}^{\pi} P_x(e^{j\omega})\phi_n(e^{j\omega})\phi_m^*(e^{j\omega})d\omega = \lambda_n \delta_{mn}$$

and find an expression for the constant λ_n. Polynomials that satisfy this orthogonality condition are called Szegö polynomials.

Solution _____

Without any loss in generality, assume that $m \leq n$, and let

$$\mathbf{e}_n = \begin{bmatrix} 1 \\ e^{j\omega} \\ e^{j2\omega} \\ \vdots \\ e^{jn\omega} \end{bmatrix}$$

be a vector of complex exponentials. With

$$r_x(k) = \frac{1}{2\pi}\int_{-\pi}^{\pi} P_x(e^{j\omega})e^{jk\omega}d\omega$$

and

$$\mathbf{R}_x = \text{Toep}\{r_x(0),\ r_x(1),\ \ldots, r_x(n)\}$$

it follows that the autocorrelation matrix may be expressed in terms of $P_x(e^{j\omega})$ and $\mathbf{e}_n$ as follows

$$\mathbf{R}_x = \frac{1}{2\pi}\int_{-\pi}^{\pi} P_x(e^{j\omega})\,\mathbf{e}\,\mathbf{e}^H d\omega$$

Let $\mathbf{a}_n = [1, a_n(1), \ldots, a_n(n)]^T$ be the nth-order prediction error filter, which is the solution to the normal equations

$$\mathbf{R}_x \mathbf{a}_n = \epsilon_n \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

Since

$$\mathbf{e}^H \mathbf{a}_n = 1 + a_n(1)e^{j\omega} + \cdots + a_n(n)e^{jn\omega} = e^{-jn\omega}\phi_n(e^{j\omega})$$

then

$$\mathbf{R}_x \mathbf{a}_n = \frac{1}{2\pi}\int_{-\pi}^{\pi} P_x(e^{j\omega})\phi_n(e^{j\omega})e^{-jn\omega}\,\mathbf{e}_n d\omega = \frac{1}{2\pi}\int_{-\pi}^{\pi} P_x(e^{j\omega})\phi_n(e^{j\omega})\begin{bmatrix} e^{-jn\omega} \\ e^{-j(n-1)\omega} \\ \vdots \\ e^{-j\omega} \\ 1 \end{bmatrix} d\omega = \epsilon_n \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \\ 0 \end{bmatrix}$$

Therefore, with

$$\phi_m^*(e^{j\omega}) = e^{-jm\omega} + a_m^*(1)e^{-j(m-1)\omega} + \cdots + a_m^*(m) = \mathbf{e}_n^H \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ a_m^*(1) \\ \vdots \\ a_m^*(m) \end{bmatrix}$$

it follows that

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} P_x(e^{j\omega})\phi_n(e^{j\omega})\phi_m^*(e^{j\omega})d\omega = \begin{cases} \epsilon_n & ; \quad n = m \\ 0 & ; \quad n \neq m \end{cases}$$

SOLUTIONS TO CHAPTER 7
Optimum Filters

7.1 A random process $x(n)$ is generated as follows

$$x(n) = \alpha x(n-1) + v(n) + \beta v(n-1)$$

where $v(n)$ is white noise with mean m_v and variance σ_v^2.

(a) Design a first-order linear predictor

$$\widehat{x}(n+1) = w(0)x(n) + w(1)x(n-1)$$

that minimizes the mean-square error in the prediction of $x(n+1)$, and find the minimum mean-square error.

(b) Now consider a predictor of the form

$$\widehat{x}(n+1) = c + w(0)x(n) + w(1)x(n-1)$$

Find the values for c, $w(0)$, and $w(1)$ that minimize the mean-square error, and compare the mean-square error of this predictor with that found in part (a).

Solution _____

(a) As we have seen in Section 7.2.2, the solution to this problem is found by solving the Wiener-Hopf equations

$$\left[\begin{array}{cc} r_x(0) & r_x(1) \\ r_x(1) & r_x(0) \end{array} \right] \left[\begin{array}{c} w(0) \\ w(1) \end{array} \right] = \left[\begin{array}{c} r_x(1) \\ r_x(2) \end{array} \right]$$

and the minimum mean-square error is

$$\xi_{\min} = r_x(0) - w(0)r_x(1) - w(1)r_x(2)$$

All that is required is to find the autocorrelation sequence

$$r_x(k) = E\big\{x(n+k)x(n)\big\}$$

Since $x(n)$ is formed by filtering $v(n)$ with a filter that has a system function given by

$$H(z) = \frac{1 + \beta z^{-1}}{1 - \alpha z^{-1}}$$

then the power spectrum of $x(n)$, which is the DTFT of the *autocovariance*, $c_x(k)$, is

$$P_x(e^{j\omega}) = \sigma_v^2 \frac{\left|1 + \beta e^{-j\omega}\right|^2}{\left|1 - \alpha e^{-j\omega}\right|^2}$$

Therefore,

$$c_x(k) = \frac{\sigma_v^2}{1 - \alpha^2}\left[(1 + \beta^2)\alpha^{|k|} + \beta\alpha^{|k-1|} + \beta\alpha^{|k+1|}\right]$$

Since the mean of $v(n)$ is nonzero, then the mean of $x(n)$ is nonzero, and the autocorrelation $r_x(k)$ is related to the autocovariance $c_x(k)$ as follows,

$$r_x(k) = c_x(k) + m_x^2$$

With

$$m_x = m_v H(e^{j\omega})|_{\omega=0} = m_v \frac{1+\beta}{1-\alpha}$$

we have

$$r_x(k) = \frac{\sigma_v^2}{1-\alpha^2} \left[(1+\beta^2)\alpha^{|k|} + \beta\alpha^{|k-1|} + \beta\alpha^{|k+1|} \right] + m_v^2 \frac{(1+\beta)^2}{(1-\alpha)^2}$$

Given values of β, α, σ_v^2, and m_v, we could then solve the Wiener-Hopf equations.

(b) With a predictor of the form,

$$\widehat{x}(n+1) = c + w(0)x(n) + w(1)x(n-1)$$

in order to minimize the mean-square error, we must have

$$\frac{\partial \xi}{\partial c} = -2E\{e(n)\} = 0$$

This, however, implies that

$$E\{x(n+1)\} - \left[c + w(0)E\{x(n)\} + w(1)E\{x(n-1)\} \right] = 0$$

Therefore, with $m_x = E\{x(n)\}$, we have

$$c = m_x\left[1 - w(0) - w(1) \right]$$

Now note that since

$$\begin{aligned} \xi &= E\left\{ x(n+1) - \left[c + w(0)x(n) + w(1)x(n-1) \right] \right\} \\ &= E\left\{ [x(n+1) - m_x] - w(0)[x(n) - m_x] + w(1)[x(n-1) - m_x] \right\} \end{aligned}$$

then minimizing ξ is equivalent to finding the optimum linear predictor

$$\widehat{x}(n+1) = w(0)x(n) + w(1)x(n-1)$$

for a zero mean process. The Wiener-Hopf equations are

$$\begin{bmatrix} r_x(0) & r_x(1) \\ r_x(1) & r_x(0) \end{bmatrix} \begin{bmatrix} w(0) \\ w(1) \end{bmatrix} = \begin{bmatrix} r_x(1) \\ r_x(2) \end{bmatrix}$$

with

$$r_x(k) = \frac{\sigma_v^2}{1-\alpha^2} \left[(1+\beta^2)\alpha^{|k|} - \beta\alpha^{|k-1|} - \beta\alpha^{|k+1|} \right]$$

and, again, given values for β, α, and σ_v^2, the Wiener filter may be found explicitly.

7.2 In this problem we consider the design of a three-step predictor using a first-order filter

$$W(z) = w(0) + w(1)z^{-1}$$

In other words, with $x(n)$ the input to the predictor $W(z)$, the output

$$\hat{x}(n+3) = w(0)x(n) + w(1)x(n-1)$$

is the minimum mean-square estimate of $x(n+3)$.

(a) What are the Wiener-Hopf equations for the Wiener three-step predictor?

(b) If the values of $r_x(k)$ for lags $k = 0$ to $k = 4$ are

$$\mathbf{r}_x = [1.0,\ 0,\ 0.1,\ -0.2,\ -0.9]^T$$

solve the Wiener-Hopf equations and find the optimum three-step predictor.

(c) Does the prediction error filter

$$F(z) = 1 + w(0)z^{-3} + w(1)z^{-4}$$

have minimum phase, i.e., are the zeros of $F(z)$ inside the unit circle? How does this compare to what you know about the prediction error filter for a one-step predictor?

Solution _____

(a) The Wiener-Hopf equations are

$$\mathbf{R}_x \mathbf{w} = \mathbf{r}_{dx}$$

where

$$\mathbf{r}_{dx} = E\{d(n)\mathbf{x}(n)\}$$

With $d(n) = x(n+3)$ this becomes

$$\mathbf{r}_{dx} = \begin{bmatrix} r_x(3) \\ r_x(4) \end{bmatrix}$$

Thus,

$$\begin{bmatrix} r_x(0) & r_x(1) \\ r_x(1) & r_x(0) \end{bmatrix} \begin{bmatrix} w(0) \\ w(1) \end{bmatrix} = \begin{bmatrix} r_x(3) \\ r_x(4) \end{bmatrix}$$

(b) With $\mathbf{r}_x = [1.0,\ 0,\ 0.1,\ -0.2,\ -0.9]^T$, it follows that $\mathbf{R}_x = \mathbf{I}$, and the solution to the Wiener-Hopf equations is

$$\mathbf{w} = \begin{bmatrix} w(0) \\ w(1) \end{bmatrix} = \begin{bmatrix} -0.2 \\ -0.9 \end{bmatrix}$$

(c) The reflection coefficient sequence corresponding to the prediction error filter,

$$F(z) = 1 - 0.2z^{-3} - 0.9z^{-4}$$

is

$$\mathbf{\Gamma} = [0.8571,\ 9.2308,\ -1.0526,\ -0.9000]^T$$

Since the reflection coefficients are not bounded by one in magnitude, $F(z)$ is not minimim phase. This is not the case for a one-step predictor, which is guaranteed to have minimum phase.

7.3 Repeat Example 7.2.1 using a second-order Wiener filter, and compare the mean-square error for the second-order filter with the mean-square error of the first-order filter.

Solution _____

For a second-order Wiener filter, the Wiener-Hopf equations are

$$\begin{bmatrix} r_x(0) & r_x(1) & r_x(2) \\ r_x(1) & r_x(0) & r_x(1) \\ r_x(2) & r_x(1) & r_x(0) \end{bmatrix} \begin{bmatrix} w(0) \\ w(1) \\ w(2) \end{bmatrix} = \begin{bmatrix} r_{dx}(0) \\ r_{dx}(1) \\ r_{dx}(2) \end{bmatrix}$$

Since $d(n)$ and $v(n)$ are assumed to be uncorrelated, then

$$r_{dx}(k) = r_d(k) = \alpha^{|k|}$$

and

$$r_x(k) = r_d(k) + r_v(k) = \alpha^{|k|} + \sigma_v^2 \delta(k)$$

Thus, the Wiener-Hopf equations become

$$\begin{bmatrix} 1 + \sigma_v^2 & \alpha & \alpha^2 \\ \alpha & 1 + \sigma_v^2 & \alpha \\ \alpha^2 & \alpha & 1 + \sigma_v^2 \end{bmatrix} \begin{bmatrix} w(0) \\ w(1) \\ w(2) \end{bmatrix} = \begin{bmatrix} 1 \\ \alpha \\ \alpha^2 \end{bmatrix}$$

With $\alpha = 0.8$ and $\sigma_v^2 = 1$, these become

$$\begin{bmatrix} 2 & 0.8 & 0.64 \\ 0.8 & 2 & 0.8 \\ 0.64 & 0.8 & 2 \end{bmatrix} \begin{bmatrix} w(0) \\ w(1) \\ w(2) \end{bmatrix} = \begin{bmatrix} 1 \\ 0.8 \\ 0.64 \end{bmatrix}$$

Solving for $w(0)$ and $w(1)$ we have

$$\begin{bmatrix} w(0) \\ w(1) \\ w(2) \end{bmatrix} = \begin{bmatrix} 0.3824 \\ 0.2000 \\ 0.1176 \end{bmatrix}$$

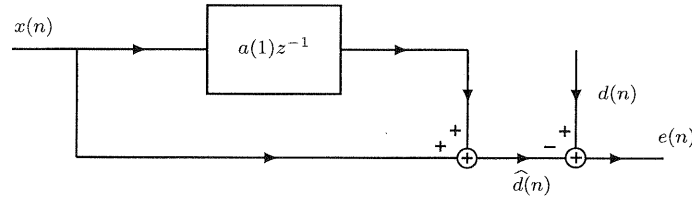Therefore, the Wiener filter is

$$W(z) = 0.3824 + 0.2z^{-1} + 0.1176z^{-2}$$

For the mean-square error, we have

$$\xi_{\min} = r_d(0) - w(0)r_{dx}(0) - w(1)r_{dx}(1) - w(2)r_{dx}(2) = 0.3824$$

Note that this is smaller than the mean-square error for a first-order Wiener, which is $\xi_{\min} = 0.4048$.

7.4 Consider the system shown in the figure below for estimating a process $d(n)$ from $x(n)$.

If $\sigma_d^2 = 4$ and

$$\mathbf{r}_x = [1.0, \ 0.5, \ 0.25]^T \quad ; \quad \mathbf{r}_{dx} = [-1.0, \ 1.0]^T$$

find the value of $a(1)$ that minimizes the mean-square error $\xi = E\{|e(n)|^2\}$, and find the minimum mean-square error.

Solution _____

The error is given by

$$e(n) = d(n) - \widehat{d}(n) = d(n) - x(n) - a(1)x(n-1)$$

The value for the coefficient $a(1)$ that minimizes the mean-square error

$$\xi = E\{|e(n)|^2\}$$

is found by setting the derivative of ξ with respect to $a^*(1)$ equal to zero and solving for $a(1)$. With

$$\frac{\partial \xi}{\partial a^*(1)} = -E\{e(n)x^*(n-1)\} = 0$$

substituting for $e(n)$ we have

$$E\{d(n)x^*(n-1) - x(n)x^*(n-1) - a(1)|x(n-1)|^2\} = 0$$

or

$$r_{dx}(1) - r_x(1) - a(1)r_x(0) = 0$$

Solving for $a(1)$ we have

$$a(1) = \frac{r_{dx}(1) - r_x(1)}{r_x(0)} = 1/2$$

The minimum mean-square error is

$$\xi = E\{e(n)[d(n) - x(n) - a(1)x(n-1)]^*\}$$

Using the orthogonality condition

$$E\{e(n)x^*(n-1)\} = 0$$

this becomes

$$
\begin{aligned}
\xi &= E\{e(n)[d(n) - x(n)]^*\} \\
&= r_d(0) - r_{dx}(0) - a(1)r_{dx}(1) - r_{dx}(0) + r_x(0) + a(1)r_x(1) \\
&= 6.75
\end{aligned}
$$

7.5 In this problem we consider linear prediction in a noisy environment. Suppose that a signal $d(n)$ is corrupted by noise,

$$x(n) = d(n) + w(n)$$

where $r_w(k) = 0.5\delta(k)$ and $r_{dw}(k) = 0$. The signal $d(n)$ is an AR(1) process that satisfies the difference equation

$$d(n) = 0.5d(n-1) + v(n)$$

where $v(n)$ is white noise with variance $\sigma_v^2 = 1$. Assume that $w(n)$ and $v(n)$ are uncorrelated.

(a) Design a first-order FIR linear predictor $W(z) = w(0) + w(1)z^{-1}$ for $d(n)$, and find the mean-square prediction error $\xi = E\{[d(n+1) - \hat{d}(n+1)]^2\}$.

(b) Design a causal Wiener predictor and compare the mean-square error with that found in part (a).

Solution _____

(a) The Wiener-Hopf equations for the FIR Wiener filter are

$$\mathbf{R}_x \mathbf{w} = \mathbf{r}_{dx}$$

Since the noise, $w(n)$, is uncorrelated with the signal, $d(n)$, the cross-correlation $r_{dx}(k)$ is

$$r_{dx}(k) = E\Big\{d(n+1)\big[d(n-k) + w(n-k)\big]\Big\} = r_d(k+1)$$

and $P_{dx}(z) = zP_d(z)$. Furthermore, the autocorrelation of $x(n)$ is

$$r_x(k) = E\big\{x(n)x(n-k)\big\} = r_d(k) + r_w(k)$$

Since the power spectrum of $d(n)$ is

$$P_d(z) = \frac{1}{(1 - \frac{1}{2}z^{-1})(1 - \frac{1}{2}z)}$$

then the autocorrelation sequence is

$$r_d(k) = \tfrac{4}{3}(\tfrac{1}{2})^{|k|}$$

With

$$r_w(k) = \tfrac{1}{2}\delta(k)$$

the Wiener-Hopf equations for the second-order predictor are

$$\begin{bmatrix} 11/6 & 2/3 \\ 2/3 & 11/6 \end{bmatrix} \begin{bmatrix} w(0) \\ w(1) \end{bmatrix} = \begin{bmatrix} 2/3 \\ 1/3 \end{bmatrix}$$

Solving for $w(0)$ and $w(1)$, we find

$$\begin{bmatrix} w(0) \\ w(1) \end{bmatrix} = \begin{bmatrix} 12/35 \\ 2/35 \end{bmatrix}$$

with a minimum mean-square error of

$$\xi_{\min} = r_d(0) - w(0)r_d(1) - w(1)r_d(2) = \tfrac{4}{3} - \tfrac{12}{35}\tfrac{2}{3} - \tfrac{2}{35}\tfrac{1}{3} = \tfrac{38}{35} = 1.086$$

(b) For the causal Wiener filter, the system function is

$$H(z) = \frac{1}{\sigma_0^2 Q(z)} \left[\frac{P_{dx}(z)}{Q(z^{-1})} \right]_+$$

Since

$$P_x(z) = P_d(z) + P_w(z) = \frac{1}{(1 - \frac{1}{2}z^{-1})(1 - \frac{1}{2}z)} + \frac{1}{2} = \frac{\frac{13}{8} - \frac{1}{4}z^{-1} - \frac{1}{4}z}{(1 - \frac{1}{2}z^{-1})(1 - \frac{1}{2}z)}$$

then we have, after factoring the numerator polynomial,

$$P_x(z) = 1.5856 \frac{(1 - 0.1577z^{-1})(1 - 0.1577z)}{(1 - \frac{1}{2}z^{-1})(1 - \frac{1}{2}z)}$$

Therefore,

$$Q(z) = \frac{1 - 0.1577z^{-1}}{1 - \frac{1}{2}z^{-1}}$$

and, with $P_{dx}(z) = zP_d(z)$, it follows that

$$
\begin{aligned}
H(z) &= \frac{1}{1.5856} \frac{(1 - \frac{1}{2}z^{-1})}{(1 - 0.1577z^{-1})} \left[\frac{z}{(1 - \frac{1}{2}z^{-1})(1 - \frac{1}{2}z)} \times \frac{(1 - \frac{1}{2}z)}{(1 - 0.1577z)} \right]_+ \\
&= \frac{1}{1.5856} \frac{(1 - \frac{1}{2}z^{-1})}{(1 - 0.1577z^{-1})} \left[\frac{z}{(1 - \frac{1}{2}z^{-1})(1 - 0.1577z)} \right]_+
\end{aligned}
$$

or

$$H(z) = 0.6307 \frac{(1 - \frac{1}{2}z^{-1})}{(1 - 0.1577z^{-1})} \left[\frac{1}{(1 - \frac{1}{2}z^{-1})(z^{-1} - 0.1577)} \right]_+$$

With

$$\frac{1}{(1 - \frac{1}{2}z^{-1})(z^{-1} - 0.1577)} = \frac{0.5428}{1 - \frac{1}{2}z^{-1}} + \frac{1.0856}{z^{-1} - 0.1577}$$

it follows that

$$\left[\frac{1}{(1 - \frac{1}{2}z^{-1})(z^{-1} - 0.1577)} \right]_+ = \frac{0.5428}{1 - \frac{1}{2}z^{-1}}$$

Therefore,

$$H(z) = 0.6307 \frac{(1 - \frac{1}{2}z^{-1})}{(1 - 0.1577z^{-1})} \cdot \frac{0.5428}{(1 - \frac{1}{2}z^{-1})} = \frac{0.3423}{1 - 0.1577z^{-1}}$$

Thus, the unit sample response is

$$h(n) = 0.3423(0.1577)^n u(n)$$

For the minimum mean-square error, we have

$$
\begin{aligned}
\xi_{\min} &= r_d(0) - \sum_{l=0}^{\infty} h(l) r_{dx}(l) = \frac{4}{3} - \sum_{l=0}^{\infty} 0.3423(0.1577)^l \cdot \frac{4}{3}\left(\frac{1}{2}\right)^{l+1} \\
&= \frac{4}{3} - 0.2282 \sum_{l=0}^{\infty} (0.07885)^l = 1.0856
\end{aligned}
$$

What we observe is that the minimum mean-square error is approximately the same as that for the second-order FIR Wiener filter. This is because $h(n)$ for the causal Wiener filter is approximately zero for $n > 1$ and, for $n = 0, 1$, the filter coefficients are approximately the same as those found in part (a).

7.6 Suppose that a process $x(n)$ has been recorded but there is a missing gap over the interval $[N_1, N_2]$, i.e., $x(n)$ is unknown over this interval.

   (a) Derive the optimum estimate of $x(N_1)$ using the data in the semi-infinite interval $(-\infty, N_1 - 1]$.

   (b) Derive the optimum estimate of $x(N_1)$ using the data in the semi-infinite interval $[N_2 + 1, \infty)$.

   (c) Derive the optimum estimate of $x(N_1)$ that is formed by combining together the two estimates found in parts (a) and (b).

   (d) Generalize your result in part (c) to find the optimum estimate of $x(n)$ at an arbitrary point n in the interval $[N_1, N_2]$.

Solution _____

   (a) This is the one-step causal linear prediction problem. The optimum estimate of $x(N_1)$ is

$$\widehat{x}_1(N_1) = \sum_{k=0}^{\infty} h_1(k) x(N_1 - 1 - k)$$

where

$$H_1(z) = \frac{1}{Q(z)} \Big[zQ(z) \Big]_+$$

and $Q(z)$ is the minimum phase factor of $P_x(z)$,

$$P_x(z) = \sigma_0^2 Q(z) Q(z^{-1})$$

   (b) Here we want to design the optimum filter for predicting *backwards* $N_2 - N_1 + 1$ samples. The system function for the optimum causal $(N_2 - N_1 + 1)$-step *forward* predictor is

$$H_2(z) = \frac{1}{Q(z)} \Big[z^{N_2 - N_1 + 1} Q(z) \Big]_+$$

Therefore, the optimum estimate of $x(N_1)$ is

$$\widehat{x}_2(N_1) = \sum_{k=0}^{\infty} h_2(k) x(N_2 + 1 + k)$$

   (c) We are given two estimates of $x(N_1)$, which we have called $\widehat{x}_1(N_1)$ and $\widehat{x}_2(N_1)$. Let the variance of these estimates be denoted by σ_1^2 and σ_2^2, respectively. We would now like to find the optimum estimate of $x(N_1)$ using an estimate of the form

$$\widehat{x}(N_1) = K\widehat{x}_1(N_1) + (1 - K)\widehat{x}_2(N_1)$$

where K is a constant that is to be determined. Note that this form for the estimate guarantees that $\widehat{x}(N_1)$ will be unbiased if $\widehat{x}_1(N_1)$ and $\widehat{x}_2(N_1)$ are unbiased. To find the value of K that minimizes the mean-square error,

$$\xi = E\Big\{ \big[x(N_1) - \widehat{x}(N_1) \big]^2 \Big\}$$

we set the derivative of ξ with respect to K equal to zero and solve. With

$$e(N_1) = x(N_1) - \widehat{x}(N_1) = K\big[x(N_1) - \widehat{x}_1(N_1)\big] + (1 - K)\big[x(N_1) - \widehat{x}_2(N_1)\big]$$

we have

$$\frac{\partial \xi}{\partial K} = 2E\left\{ e(N_1)\Big(\big[x(N_1) - \widehat{x}_1(N_1)\big] - \big[x(N_1) - \widehat{x}_2(N_1)\big] \Big) \right\} = 0$$

Substituting for $e(N_1)$ and taking the expected value we have, assuming that the estimation errors $e_1(N_1) = x(N_1) - \widehat{x}_1(N_1)$ and $e_2(N_1) = x(N_1) - \widehat{x}_2(N_1)$ are uncorrelated,

$$K\sigma_1^2 - (1 - K)\sigma_2^2 = 0$$

Solving for K we have

$$K = \frac{\sigma_2^2}{\sigma_1^2 + \sigma_2^2}$$

which leads to the following estimate for $x(N_1)$,

$$\widehat{x}(N_1) = \frac{\sigma_2^2}{\sigma_1^2 + \sigma_2^2}\widehat{x}_1(N_1) + \frac{\sigma_1^2}{\sigma_1^2 + \sigma_2^2}\widehat{x}_2(N_1)$$

(d) For an arbitrary point N_0 in the interval $[N_1, N_2]$, the result derived in part (c) produces the optimum estimate of $x(N_0)$ provided $H_1(z)$ is replaced with

$$H_1(z) = \frac{1}{Q(z)}\Big[z^{N_0 - N_1 + 1}Q(z)\Big]_+$$

and $H_2(z)$ is replaced with

$$H_2(z) = \frac{1}{Q(z)}\Big[z^{N_2 - N_0 + 1}Q(z)\Big]_+$$

7.7 In this problem we consider the design of a causal IIR Wiener filter for p-step prediction,

$$\widehat{x}(n+p) = \sum_{k=0}^{\infty} h(k)x(n-k)$$

(a) If $x(n)$ is a real-valued random process with power spectral density

$$P_x(z) = \sigma_0^2 Q(z)Q(z^{-1})$$

find the system function of the causal Wiener filter that minimizes the mean-square error

$$\xi = E\{|x(n+p) - \widehat{x}(n+p)|^2\}$$

(b) If $x(n)$ is an AR(1) process with power spectrum

$$P_x(z) = \frac{1 - a^2}{(1 - az^{-1})(1 - az)}$$

find the causal p-step linear predictor and evaluate the mean-square error.

(c) If $x(n)$ is an MA(2) process that is generated by the difference equation

$$x(n) = 4v(n) - 2v(n-1) + v(n-2)$$

where $v(n)$ is zero mean unit variance white noise, find the system function of the two-step ($p = 2$) predictor and evaluate the mean-square error.

(d) Repeat part (c) for a three-step predictor.

Solution _____

(a) The optimum *causal* Wiener filter is

$$H(z) = \frac{1}{\sigma_0^2 Q(z)} \left[\frac{P_{dx}(z)}{Q(z^{-1})} \right]_+$$

With $d(n) = x(n+p)$,

$$r_{dx}(k) = r_x(k+p)$$

and

$$P_{dx}(z) = z^p P_x(z) = z^p \sigma_0^2 Q(z)Q(z^{-1})$$

Therefore,

$$H(z) = \frac{1}{Q(z)} \left[z^p Q(z) \right]_+$$

(b) With the power spectrum of $x(n)$ given by

$$P_x(z) = \frac{1 - a^2}{(1 - az^{-1})(1 - az)}$$

we see that $\sigma_0^2 = 1 - a^2$, and

$$Q(z) = \frac{1}{1 - az^{-1}}$$

Therefore,

$$
\begin{aligned}
\left[z^p Q(z)\right]_+ &= \left[\frac{z^p}{1 - az^{-1}}\right]_+ \\
&= \left[z^p\left(1 + az^{-1} + a^2 z^{-2} + \cdots\right)\right]_+ \\
&= a^p + a^{p+1} z^{-1} + a^{p+2} z^{-2} + \cdots \\
&= \frac{a^p}{1 - az^{-1}}
\end{aligned}
$$

and the optimum causal Wiener filter for estimating $x(n+p)$ is

$$H(z) = \frac{1}{Q(z)}\left[z^p Q(z)\right]_+ = (1 - az^{-1}) \cdot \frac{a^p}{1 - az^{-1}} = a^p$$

Thus, the estimate of $x(n+p)$ is given by

$$\widehat{x}(n+p) = a^p x(n)$$

Note that the predictor only uses the most recent value of the sequence, $x(n)$, to predict $x(n+p)$. Since $x(n)$ is an autoregressive process, this value carries all of the information about the past history of $x(n)$.

The minimum error is equal to

$$\xi_{\min} = r_d(0) - \sum_{l=0}^{\infty} h(l) r_{dx}^*(l)$$

With $r_d(k) = r_x(k) = a^{|k|}$, $r_{dx}(k) = r_x(k+p)$, and $h(n) = a^p \delta(n)$, this becomes

$$\xi_{\min} = 1 - a^{2p}$$

(c) If $x(n)$ is a moving average process that is generated by the difference equation

$$x(n) = 4v(n) - 2v(n-1) + v(n-2)$$

then the power spectrum of $x(n)$ is

$$P_x(z) = (4 - 2z^{-1} + z^{-2})(4 - 2z + z) = 16(1 - \tfrac{1}{2}z^{-1} + \tfrac{1}{4}z^{-2})(1 - \tfrac{1}{2}z + \tfrac{1}{4}z)$$

With

$$\left[z^p Q(z)\right]_+ = \left[z^p(1 - \tfrac{1}{2}z^{-1} + \tfrac{1}{4}z^{-2})\right]_+$$

if $p = 2$, then this becomes

$$\left[z^2 Q(z)\right]_+ = \tfrac{1}{4}$$

Therefore, the optimum predictor is

$$H(z) = \frac{1}{Q(z)}\left[z^2 Q(z)\right]_+ = \frac{\tfrac{1}{4}}{1 - \tfrac{1}{2}z^{-1} + \tfrac{1}{4}z^{-2}}$$

or

$$\widehat{x}(n+1) = \tfrac{1}{2}\widehat{x}(n) - \tfrac{1}{4}\widehat{x}(n-1) + \tfrac{1}{4}x(n)$$

For the minimum error, we have

$$\xi_{\min} = r_d(0) - \frac{1}{2\pi} \int_{-\pi}^{\pi} H(e^{j\omega}) P_{dx}^*(e^{j\omega}) d\omega$$

With $r_d(0) = r_x(0) = 21$ and

$$P_{dx}^*(e^{j\omega}) = e^{-jp\omega} P_x(e^{j\omega}) = e^{-jp\omega} 16\big|1 - \tfrac{1}{2}e^{-j\omega} + \tfrac{1}{4}e^{-j2\omega}\big|^2$$

we have, with $p = 2$,

$$\xi_{\min} = 21 - \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{1/4}{1 - \tfrac{1}{2}e^{-j\omega} + \tfrac{1}{4}e^{-j2\omega}} \cdot 16 e^{-j2\omega} \big|1 - \tfrac{1}{2}e^{-j\omega} + \tfrac{1}{4}e^{-j2\omega}\big|^2 d\omega$$

Therefore,

$$\xi_{\min} = 21 - \frac{1}{2\pi} \int_{-\pi}^{\pi} 4 e^{-j2\omega}\big[1 - \tfrac{1}{2}e^{j\omega} + \tfrac{1}{4}e^{j2\omega}\big] d\omega = 20$$

(d) With $p = 3$, note that

$$\Big[z^3 Q(z)\Big]_+ = 0$$
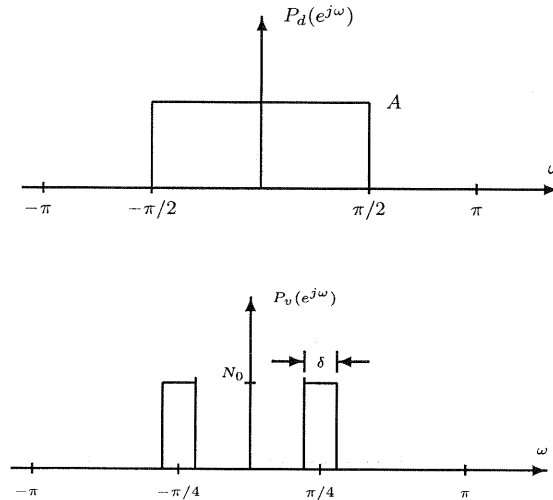
Therefore, the optimum predictor is

$$\widehat{x}(n+3) = 0$$

7.8 Suppose that we would like to estimate a signal $d(n)$ from the noisy observations

$$x(n) = d(n) + v(n)$$

where the noise, $v(n)$, is uncorrelated with $d(n)$. The power spectral densities of $d(n)$ and $v(n)$ are shown in the following figure.

(a) Design a *noncausal* Wiener smoothing filter for estimating $d(n)$ from $x(n)$,

$$\widehat{d}(n) = \sum_{k=-\infty}^{\infty} h(k)x(n-k)$$

(b) Compute the mean-square error $E\{|d(n) - \widehat{d}(n)|^2\}$ and compare it to the mean-square error that results when $h(n) = \delta(n)$, i.e., with no filtering of $x(n)$.
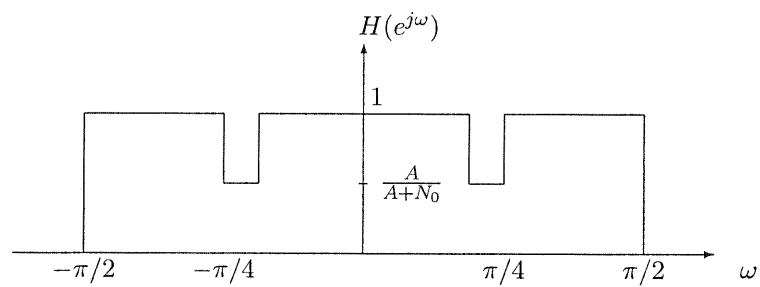
Solution _____

(a) The frequency response of the optimum noncausal Wiener filter is

$$H(e^{j\omega}) = \frac{P_{dx}(e^{j\omega})}{P_x(e^{j\omega})} = \frac{P_d(e^{j\omega})}{P_d(e^{j\omega}) + P_w(e^{j\omega})}$$

Using the given power spectral densities for $d(n)$ and $w(n)$, it follows that the frequency response of the Wiener filter is as shown in the figure below.

(b) The minimum mean-square error is

$$
\begin{aligned}
\xi_{min} &= r_d(0) - \sum_{l=-\infty}^{\infty} h(l)r_{dx}(l) \\
&= \frac{1}{2\pi} \int_{-\pi}^{\pi} \left[P_d(e^{j\omega}) - H(e^{j\omega})P_d(e^{-j\omega}) \right] d\omega
\end{aligned}
$$

$$H(e^{j\omega})$$

Evaluating the integral we find

$$\xi_{min} = \frac{\delta A}{\pi} \frac{N_0}{N_0 + A}$$

The mean-square error using $h(n) = \delta(n)$ is

$$\xi = E\left[v^2(n)\right] = \frac{1}{2\pi} \int_{-\pi}^{\pi} P_v(e^{j\omega}) d\omega = \frac{\delta N_0}{\pi}$$

7.9 We would like to estimate a process $d(n)$ from noisy observations,

$$x(n) = d(n) + v(n)$$

where $v(n)$ is white noise with variance $\sigma_v^2 = 1$ and $d(n)$ is a wide-sense stationary random process with the first four values of the autocorrelation sequence given by

$$\mathbf{r}_d = [1.5, \ 0, \ 1.0, \ 0]^T$$

Assume that $d(n)$ and $v(n)$ are uncorrelated. Our goal is to design an FIR filter to reduce the noise in $d(n)$. Hardware constraints, however, limit the filter to only three nonzero coefficients in $W(z)$.

(a) Derive the optimal three-multiplier causal filter

$$W(z) = w(0) + w(1)z^{-1} + w(2)z^{-2}$$

for estimating $d(n)$ and evaluate the mean-square error $E\{|d(n) - \widehat{d}(n)|^2\}$.

(b) Repeat (a) for the noncausal FIR filter

$$W(z) = w(-1)z + w(0) + w(1)z^{-1}$$

(c) Can you suggest a way to reduce the mean-square error below that obtained for the filters designed in parts (a) and (b) without using any more than three filter coefficients?

Solution _____

(a) The Wiener-Hopf equations for the optimal three-multiplier filter

$$W(z) = w(0) + w(1)z^{-1} + w(2)z^{-2}$$

for estimating $d(n)$ are

$$\begin{bmatrix} r_x(0) & r_x(1) & r_x(2) \\ r_x(1) & r_x(0) & r_x(1) \\ r_x(2) & r_x(1) & r_x(0) \end{bmatrix} \begin{bmatrix} w(0) \\ w(1) \\ w(2) \end{bmatrix} = \begin{bmatrix} r_{dx}(0) \\ r_{dx}(1) \\ r_{dx}(2) \end{bmatrix}$$

With

$$r_x(k) = r_d(k) + r_v(k)$$

and

$$r_{dx}(k) = r_d(k)$$

using the given values for the autocorrelation $r_d(k)$, these equations become

$$\begin{bmatrix} 2.5 & 0 & 1.0 \\ 0 & 2.5 & 0 \\ 1.0 & 0 & 2.5 \end{bmatrix} \begin{bmatrix} w(0) \\ w(1) \\ w(2) \end{bmatrix} = \begin{bmatrix} 1.5 \\ 0 \\ 1.0 \end{bmatrix}$$

The solution for the coefficients $w(k)$ is

$$\begin{bmatrix} w(0) \\ w(1) \\ w(2) \end{bmatrix} = \begin{bmatrix} 0.5238 \\ 0 \\ 0.1905 \end{bmatrix}$$

The mean-square error is

$$E\{|d(n) - \widehat{d}(n)|^2\} = r_d(0) - \sum_{k=0}^{2} w(k) r_{dx}(k) = 1.5 - 0.5238 \cdot (1.5) - 0.1905 = 0.5238$$

(b) For the noncausal FIR filter

$$W(z) = w(-1)z + w(0) + w(1)z^{-1}$$

the Wiener-Hopf equations for the optimum coefficients may be derived as follows. With ξ the mean-square error

$$\xi = E\{e^2(n)\}$$

we set the derivative of ξ with respect to $w(k)$ equal to zero as follows

$$\frac{\partial \xi}{\partial w(k)} = E\{-2e(n)x(n-k)\} = 0 \quad ; \quad k = 0, \pm 1$$

Substituting for $e(n)$ and simplifying we have

$$\sum_{l=-1}^{1} w(l) r_x(k-l) = r_{dx}(k) \quad ; \quad k = 0, \pm 1$$

Since $r_{dx}(k) = r_d(k)$, then these equations are

$$\begin{bmatrix} r_x(0) & r_x(1) & r_x(2) \\ r_x(1) & r_x(0) & r_x(1) \\ r_x(2) & r_x(1) & r_x(0) \end{bmatrix} \begin{bmatrix} w(-1) \\ w(0) \\ w(1) \end{bmatrix} = \begin{bmatrix} r_d(-1) \\ r_d(0) \\ r_d(1) \end{bmatrix}$$

Using the given autocorrelations, these become

$$\begin{bmatrix} 2.5 & 0 & 1.0 \\ 0 & 2.5 & 0 \\ 1.0 & 0 & 2.5 \end{bmatrix} \begin{bmatrix} w(-1) \\ w(0) \\ w(1) \end{bmatrix} = \begin{bmatrix} 0 \\ 1.5 \\ 0 \end{bmatrix}$$

Therefore, we have

$$\begin{bmatrix} w(-1) \\ w(0) \\ w(1) \end{bmatrix} = \begin{bmatrix} 0 \\ 0.6 \\ 0 \end{bmatrix}$$

with a mean-square error

$$E\{|d(n) - \widehat{d}(n)|^2\} = r_d(0) - \sum_{k=-1}^{1} w(k) r_{dx}(k) = 0.6$$

(c) Since $r_d(1) = 0$ and the additive noise, $v(n)$, is white, then $x(n \pm 1)$ is of no use in estimating $d(n)$. Therefore, a better estimator to use that has only three coefficients is the following

$$W(z) = w(0) + w(2)z^{-2} + w(-2)z^2$$

7.10 Suppose that a signal $x(n)$ is recorded and that, due to measurement errors, there are *outliers* in the data, i.e., for some values of n there is a large error in the measured value of $x(n)$. Instead of eliminating these data values, suppose that we perform a minimum mean-square interpolation as follows. Given a "bad" data value at time $n = n_0$, consider an estimate for $x(n_0)$ of the form

$$\widehat{x}(n_0) = ax(n_0 - 1) + bx(n_0 + 1)$$

(a) Assuming that $x(n)$ is a wide-sense stationary random process with autocorrelation sequence $r_x(k)$, find the values for a and b that minimize the mean-square error

$$\xi = E\{|x(n_0) - \widehat{x}(n_0)|^2\}$$

(b) If $r_x(k) = (0.5)^{|k|}$, evaluate the mean-square error for the interpolator found in part (a).

(c) Discuss when it may be better to use an estimator of the form

$$\widehat{x}(n_0) = ax(n_0 - 1) + bx(n_0 - 2)$$

or explain why such an estimator should not be used.

(d) Given an autocorrelation sequence $r_x(k)$, derive the Wiener-Hopf equations that define the optimum filter for interpolating $x(n)$ to produce the best estimate of $x(n_0)$ in terms of the $2p$ data values

$$x(n_0 - 1), x(n_0 - 2), \ldots, x(n_0 - p) \quad \text{and} \quad x(n_0 + 1), x(n_0 + 2), \ldots, x(n_0 + p)$$

(e) Find an expression for the mean-square error for your estimate in part (d).

Solution _____

(a) With an interpolator of the form

$$\widehat{x}(n_0) = ax(n_0 - 1) + bx(n_0 + 1)$$

the values for a and b that minimize the mean-square error

$$\xi = E\{|x(n_0) - \widehat{x}(n_0)|^2\}$$

are found by setting the derivative of ξ with respect to a and b equal to zero. Assuming that $x(n)$ is a real-valued process, we have

$$\frac{\partial \xi}{\partial a} = E\left\{-2[x(n_0) - \widehat{x}(n_0)]x(n_0 - 1)\right\} = 0$$

and

$$\frac{\partial \xi}{\partial b} = E\left\{-2[x(n_0) - \widehat{x}(n_0)]x(n_0 + 1)\right\} = 0$$

Dividing by 2 and substituting for $\widehat{x}(n)$ gives

$$E\left\{[x(n_0) - ax(n_0 - 1) - bx(n_0 + 1)]x(n_0 - 1)\right\} = 0$$

$$E\Big\{\big[x(n_0) - ax(n_0 - 1) - bx(n_0 + 1)\big]x(n_0 + 1)\Big\} = 0$$

Writing these equations in matrix form in terms of the autocorrelation sequence $r_x(k)$, we have

$$\begin{bmatrix} r_x(0) & r_x(2) \\ r_x(2) & r_x(0) \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} r_x(1) \\ r_x(1) \end{bmatrix}$$

Solving for a and b we find

$$\begin{aligned}
\begin{bmatrix} a \\ b \end{bmatrix} &= \frac{1}{r_x^2(0) - r_x^2(2)} \begin{bmatrix} r_x(0) & -r_x(2) \\ -r_x(2) & r_x(0) \end{bmatrix} \begin{bmatrix} r_x(1) \\ r_x(1) \end{bmatrix} \\
&= \frac{1}{r_x^2(0) - r_x^2(2)} \begin{bmatrix} r_x(1)\big[r_x(0) - r_x(2)\big] \\ r_x(1)\big[r_x(0) - r_x(2)\big] \end{bmatrix} \\
&= \frac{1}{r_x(0) + r_x(2)} \begin{bmatrix} 1 \\ 1 \end{bmatrix}
\end{aligned}$$

(b) The mean-square error for the interpolator found in part (a) is

$$\xi_{\min} = E\{e(n_0)x(n_0)\} = E\Big\{\big[x(n_0) - ax(n_0 - 1) - bx(n_0 + 1)\big]x(n_0)\Big\} = r_x(0) - ar_x(1) - br_x(1)$$

With $r_x(k) = (0.5)^{|k|}$,

$$a = b = 0.5$$

and

$$\xi_{\min} = 0.5$$

(c) Whether we use the interpolator

$$\widehat{x}(n_0) = ax(n_0 - 1) + bx(n_0 - 2)$$

or

$$\widehat{x}(n_0) = ax(n_0 - 1) + bx(n_0 + 1)$$

depends on the autocorrelation sequence for $x(n)$. If $|r_x(2)| > |r_x(1)|$, then the first estimate would be better. Otherwise, the second one will produce a smaller mean-square error.

(d) With an estimate of $x(n_0)$ of the form

$$\widehat{x}(n_0) = \sum_{\substack{l=-p \\ l \neq 0}}^{p} a(l)x(n_0 - l)$$

the filter coefficients that minimize the mean-square error are found by setting the derivatives of $E\{e^2(n_0)\}$ with respect to $a(k)$ equal to zero as follows,

$$\frac{\partial}{\partial a(k)} E\{e^2(n_0)\} = -2E\{e(n_0)x(n_0 - k)\} \quad ; \quad k = \pm 1, \pm 2, \ldots, \pm p$$

Dividing by two, and substituting for the error $e(n)$, this becomes

$$E\Big\{\big[x(n_0) - \sum_{\substack{l=-p \\ l \neq 0}}^{p} a(l)x(n_0 - l)\big]x(n_0 - k)\Big\} = 0$$

or

$$r_x(k) - \sum_{\substack{l=-p \\ l \neq 0}}^{p} a(l) r_x(k-l) = 0 \quad ; \quad k = \pm 1, \pm 2, \ldots, \pm p$$

where $r_x(k)$ is the autocorrelation of $x(n)$.

(e) The minimum mean-square error is

$$\xi_{\min} = E\{e(n_0)x(n_0)\} = E\left\{ \left[x(n_0) - \sum_{\substack{l=-p \\ l \neq 0}}^{p} a(l) x(n_0 - l) \right] x(n_0) \right\}$$

or,

$$\xi_{\min} = r_x(0) - \sum_{\substack{l=-p \\ l \neq 0}}^{p} a(l) r_x(l)$$

7.11 In this problem we consider the design of an optimum smoothing filter for estimating a process $d(n)$ from the measurements

$$x(n) = d(n) + v(n)$$

Our goal is to use a noncausal FIR filter that has a system function of the form:

$$W(z) = \sum_{k=-p}^{p} w(k)z^{-k}$$

In other words, we want to produce an estimate of $d(n)$ as follows

$$\widehat{d}(n) = \sum_{k=-p}^{p} w(k)x(n-k)$$

(a) Derive the Wiener-Hopf equations that define the set of coefficients that minimize the mean-square error

$$\xi = E\{|d(n) - \widehat{d}(n)|^2\}$$

(b) How would the Wiener-Hopf equations derived in part (a) change if we used a causal filter with the same number of coefficients? In other words, if the system function was of the form

$$W(z) = \sum_{k=0}^{2p} w(k)z^{-k}$$

how would you modify your equations in (a)?

(c) State qualitatively when you might prefer the noncausal filter over the causal filter and vice versa. For example, for what types of signals and for what types of noise would you expect a causal filter to be superior to the noncausal filter?

(d) FIR digital filters with linear phase (or zero phase) are important in many signal processing applications where frequency dispersion due to nonlinear phase is harmful. An FIR filter with zero phase is characterized by the property that

$$w(n) = w(-n)$$

Thus, the system function may be written as

$$W(z) = w(0) + \sum_{k=1}^{p} w(k)[z^{-k} + z^{k}]$$

Derive the Wiener-Hopf equations that define the optimum zero phase smoothing filter.

(e) With $r_d(k) = 4(0.5)^{|k|}$ and $r_v(k) = \delta(k)$, find the optimum values for the filter coefficients $w(0)$ and $w(1)$ in the zero phase filter

$$W(z) = w(0) + w(1)[z + z^{-1}]$$

Solution ───

(a) Using a noncausal filter of the form

$$W(z) = \sum_{k=-p}^{p} w(k)z^{-k}$$

the estimate of $d(n)$ is given by

$$\widehat{d}(n) = \sum_{k=-p}^{p} w(k)x(n-k)$$

With

$$\xi = E\big\{|e(n)|^2\big\} = E\big\{|d(n) - \widehat{d}(n)|^2\big\}$$

the coefficients that minimize the mean-square error are solutions to the following equations

$$\frac{\partial \xi}{\partial w^*(k)} = E\left\{e(n)\frac{\partial e^*(n)}{\partial a^*(k)}\right\} = 0 \quad ; \quad k = -p, \ldots, p$$

Since

$$e(n) = d(n) - \sum_{l=-p}^{p} w(l)x(n-l)$$

then

$$\frac{\partial e^*(n)}{\partial w^*(k)} = -x^*(n-k)$$

and

$$\frac{\partial \xi}{\partial w^*(k)} = -E\big\{e(n)x^*(n-k)\big\} = 0 \quad ; \quad k = -p, \ldots, p$$

Using the definition of $e(n)$ this becomes

$$E\big\{d(n)x^*(n-k)\big\} - \sum_{l=-p}^{p} w(l)E\big\{x(n-l)x^*(n-k)\big\} = 0$$

and, with $r_x(k) = E\{x(n)x^*(n-k)\}$ and $r_{dx}(k) = E\{d(n)x^*(n-k)\}$ we have

$$\sum_{l=-p}^{p} w(l)r_x(k-l) = r_{dx}(k) \quad ; \quad k = -p, \ldots, p$$

which is a set of $(2p+1)$ linear equations in the $(2p+1)$ unknowns $w(k)$, $k = -p, \ldots, p$. These equations may be written in matrix form as follows

$$\begin{bmatrix} r_x(0) & r_x(1) & \cdots & r_x(2p) \\ r_x(1) & r_x(0) & \cdots & r_x(2p-1) \\ \vdots & \vdots & & \vdots \\ r_x(2p) & r_x(2p-1) & \cdots & r_x(0) \end{bmatrix} \begin{bmatrix} w(-p) \\ w(1-p) \\ \vdots \\ w(p) \end{bmatrix} = \begin{bmatrix} r_{dx}(-p) \\ r_{dx}(1-p) \\ \vdots \\ r_{dx}(p) \end{bmatrix}$$

We may now evaluate the mean-square error,

$$
\begin{aligned}
\xi &= E\{|e(n)|^2\} = E\left\{e(n)\left[d(n) - \sum_{l=-p}^{p} w(l)x(n-l)\right]^*\right\} \\
&= E\{e(n)d^*(n)\} - \sum_{l=-p}^{p} w(l)E\{e(n)x^*(n-l)\}
\end{aligned}
$$

However, from the orthogonality condition, $E\{e(n)x^*(n-k)\} = 0$, it follows that the second term is zero, and we have

$$
\xi_{\min} = E\{e(n)d^*(n)\} = E\left\{\left[d(n) - \sum_{l=-p}^{p} w(l)x(n-l)\right]d^*(n)\right\}
$$

Finally, taking expected values we find

$$
\xi_{\min} = r_d(0) - \sum_{l=-p}^{p} w(l)r_{dx}^*(l)
$$

(b) If we change the filter in (a) to a causal smoothing filter of the form

$$
W(z) = \sum_{k=0}^{2p} w(k)z^{-k}
$$

then the only thing that changes in the derivations above is the limits on the summations. In particular, the Wiener-Hopf equations would take the form

$$
\sum_{l=0}^{2p} w(l)r_x(k-l) = r_{dx}^*(k) \quad ; \quad k = 0, \ldots, 2p
$$

and the minimum error would be

$$
\xi_{\min} = r_d(0) - \sum_{l=0}^{2p} w(l)r_{dx}^*(l)
$$

(c) If the autocorrelation sequence for $d(n)$ is generally larger over the interval $[-p, \ldots, p]$ than it is over the interval $[0, \ldots, 2p]$, then the zero phase smoothing filter would be better to use. Otherwise, the causal filter would be better. For example, with a signal that has an autocorrelation sequence $\mathbf{r}_d = [1, 0, 0, 0, .9, \ldots]$, you may wish to compare the performance of the optimum causal and noncausal filters when $p = 2$.

(d) Using a zero phase smoothing filter

$$
W(z) = w(0) + \sum_{k=1}^{p} w(k)\left[z^{-k} + z^{k}\right]
$$

the estimate of $d(n)$ is now given by

$$
\widehat{d}(n) = w(0)x(n) + \sum_{l=1}^{p} w(l)\left[x(n-l) + x(n+l)\right]
$$

Again, with

$$\xi = E\{|e(n)|^2\} = E\{|d(n) - \widehat{d}(n)|^2\}$$

we differentiate ξ with respect to $w^*(k)$ as follows

$$\frac{\partial \xi}{\partial w^*(k)} = E\left\{e(n)\frac{\partial e^*(n)}{\partial w^*(k)}\right\} = 0 \quad ; \quad k = 0, \ldots, p$$

However, since

$$\frac{\partial e^*(n)}{\partial w^*(0)} = -x^*(n)$$

and

$$\frac{\partial e^*(n)}{\partial w^*(k)} = -x^*(n - k) - x^*(n + k) \quad ; \quad k = 1, \ldots, p$$

then we have

$$E\{e(n)x^*(n)\} = 0$$

and

$$E\{e(n)[x(n - k) + x(n + k)]^*\} = 0 \quad ; \quad k = 1, \ldots, p$$

If we now substitute for $e(n)$ we find

$$r_{dx}(0) - \left\{w(0)r_x(0) + \sum_{l=1}^{p} w(l)\big[r_x(l) + r_x(-l)\big]\right\} = 0$$

and

$$2r_{dx}(k) - \left\{2w(0)r_x(k) + \sum_{l=1}^{p} 2w(l)\big[r_x(k - l) + r_x(k + l)\big]\right\} = 0 \quad ; \quad k = 1, \ldots, p$$

Thus,

$$w(0)r_x(0) + \sum_{k=1}^{p} 2w(k)r_x(k) = r_{dx}(0)$$

and

$$w(0)r_x(k) + \sum_{l=1}^{p} w(l)\big[r_x(k - l) + r_x(k + l)\big] = r_{dx}(k) \quad ; \quad k = 1, \ldots, p$$

(e) The coefficients for the third-order zero phase filter are the solution to the equations,

$$\begin{bmatrix} r_x(0) & 2r_x(1) \\ r_x(1) & r_x(0) + r_x(2) \end{bmatrix} \begin{bmatrix} w(0) \\ w(1) \end{bmatrix} = \begin{bmatrix} r_{dx}(0) \\ r_{dx}(1) \end{bmatrix}$$

Incorporating the given values for the autocorrelations, we have

$$\begin{bmatrix} 5 & 4 \\ 2 & 6 \end{bmatrix} \begin{bmatrix} w(0) \\ w(1) \end{bmatrix} = \begin{bmatrix} 4 \\ 2 \end{bmatrix}$$

and the solution is

$$W(z) = \frac{8}{11} + \frac{1}{11}\left[z^{-1} + z\right]$$

7.12 We observe a signal, $x(n)$, in a noisy and reverberant environment,

$$y(n) = x(n) + 0.8x(n-1) + v(n)$$

where $v(n)$ is white noise with variance $\sigma_w^2 = 1$ that is uncorrelated with $x(n)$. We know that $x(n)$ is a wide-sense stationary AR(1) random process with autocorrelations

$$\mathbf{r}_x = [4,\ 2,\ 1,\ 0.5]^T$$

(a) Find the non-causal IIR Wiener filter, $H(z)$, that produces the minimum mean-square estimate of $x(n)$.

(b) Design a causal IIR Wiener filter, $H(z)$, that produces the minimum mean-square estimate of $x(n)$.

Solution _____

Before we begin this problem, we need to determine the power spectrum for $x(n)$. Since $x(n)$ is an AR(1) process, we can determine $r_x(k)$ and $P_x(z)$ from $r_x(0)$ and $r_x(1)$, which are given. Specifically, we have

$$P_x(z) = \frac{b^2(0)}{\left[1 + a(1)z^{-1}\right]\left[1 + a(1)z\right]}$$

where

$$\begin{bmatrix} r_x(0) & r_x(1) \\ r_x(1) & r_x(0) \end{bmatrix} \begin{bmatrix} 1 \\ a(1) \end{bmatrix} = \epsilon_1 \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

and

$$b^2(0) = \epsilon_1$$

Solving for $a(1)$ and $b(0)$ we find

$$a(1) = -\frac{r_x(1)}{r_x(0)} = 1/2$$

and

$$b^2(0) = \frac{r_x^2(0) - r_x^2(1)}{r_x(0)} = 3$$

Thus,

$$P_x(z) = \frac{3}{(1 - 0.5z^{-1})(1 - 0.5z)}$$

(a) For the noncausal filter, we know that

$$H(z) = \frac{P_{xy}(z)}{P_y(z)}$$

Since

$$y(n) = x(n) + 0.8x(n-1) + v(n)$$

then

$$P_y(z) = (1 + 0.8z^{-1})(1 + 0.8z)P_x(z) + P_v(z)$$

In addition,

$$r_{xy}(k) = E\left\{x(n+k)\left[x(n) + 0.8x(n-1) + v(n)\right]\right\} = r_x(k) + 0.8r_x(k+1)$$

and

$$P_{xy}(z) = (1 + 0.8z)P_x(z)$$

Therefore, the noncausal Wiener filter is

$$H(z) = \frac{(1 + 0.8z)P_x(z)}{(1 + 0.8z^{-1})(1 + 0.8z)P_x(z) + 1}$$

or

$$H(z) = \frac{3(1 + 0.8z)}{3(1 + 0.8z^{-1})(1 + 0.8z) + (1 - 0.5z^{-1})(1 - 0.5z)}$$

which may be simplified to

$$H(z) = \frac{3(1 + 0.8z)}{6.17 + 1.9z^{-1} + 1.9z}$$

(b) For the causal Wiener filter,

$$H(z) = \frac{1}{\sigma_y^2 Q(z)} \left[\frac{P_{xy}(z)}{Q(z^{-1})} \right]_+$$

where σ_y^2 and $Q(z)$ are found from the spectral factorization of $y(n)$,

$$P_y(z) = \sigma_y^2 Q(z)Q(z^{-1})$$

With

$$
\begin{aligned}
P_y(z) &= \frac{3(1 + 0.8z^{-1})(1 + 0.8z)}{(1 - 0.5z^{-1})(1 - 0.5z)} + 1 \\
&= \frac{6.17 + 1.9z^{-1} + 1.9z}{(1 - 0.5z^{-1})(1 - 0.5z)} = 5.5153 \cdot \frac{(1 + 0.3445z^{-1})(1 + 0.3445z)}{(1 - 0.5z^{-1})(1 - 0.5z)}
\end{aligned}
$$

it follows that

$$Q(z) = \frac{1 + 0.3445z^{-1}}{1 - 0.5z^{-1}}$$

and

$$
\begin{aligned}
H(z) &= 0.1813\frac{1 - 0.5z^{-1}}{1 + 0.3445z^{-1}} \left[\frac{3(1 + 0.8z)}{(1 - 0.5z^{-1})(1 - 0.5z)} \times \frac{1 - 0.5z}{1 + 0.3445z} \right]_+ \\
&= 0.1813\frac{1 - 0.5z^{-1}}{1 + 0.3445z^{-1}} \left[\frac{3(1 + 0.8z)}{1 - 0.5z^{-1}} \times \frac{1}{1 + 0.3445z} \right]_+
\end{aligned}
$$

Since

$$\left[\frac{3(1 + 0.8z)}{1 - 0.5z^{-1}} \times \frac{1}{1 + 0.3445z} \right]_+ = \frac{3.5829}{1 - 0.5z^{-1}}$$

then

$$H(z) = \frac{.649}{1 + .3445z^{-1}}$$

7.13 A wide-sense stationary random process has an autocorrelation sequence of the form

$$r_x(k) = \sigma_x^2 \alpha^{|k|}$$

where $|\alpha| < 1$. Over a given time interval, $[n_A, n_B]$, the process $x(n)$ is only known at the end points, i.e., the only given data is $x(n_A)$ and $x(n_B)$. Based on these two observations, determine the optimum estimate

$$\widehat{x}(n) = a(n)x(n_A) + b(n)x(n_B)$$

of $x(n)$ over each of the following intervals

(a) $n > n_B$.

(b) $n < n_A$.

(c) $n_A < n < n_B$.

Solution _____

The error in the estimate of $x(n)$ is

$$e(n) = x(n) - \big[a(n)x(n_A) + b(n)x(n_B)\big]$$

and our goal is to minimize

$$\xi = E\big\{e^2(n)\big\}$$

The coefficients $a(n)$ and $b(n)$ are found by setting the derivative of ξ with respect to $a(n)$ and $b(n)$ equal to zero as follows:

$$\frac{\partial \xi}{\partial a(n)} = -2E\big\{e(n)x(n_A)\big\} = 0$$

and

$$\frac{\partial \xi}{\partial b(n)} = -2E\big\{e(n)x(n_B)\big\} = 0$$

Dividing by two, and substituting for $e(n)$ we find

$$r_x(n - n_A) = a(n)r_x(0) + b(n)r_x(n_B - n_A)$$

and

$$r_x(n - n_B) = a(n)r_x(n_B - n_A) + b(n)r_x(0)$$

or, in matrix form,

$$\begin{bmatrix} r_x(0) & r_x(n_B - n_A) \\ r_x(n_B - n_A) & r_x(0) \end{bmatrix} \begin{bmatrix} a(n) \\ b(n) \end{bmatrix} = \begin{bmatrix} r_x(n - n_A) \\ r_x(n - n_B) \end{bmatrix}$$

Solving for $a(n)$ and $b(n)$ we find

$$\begin{bmatrix} a(n) \\ b(n) \end{bmatrix} = \frac{1}{r_x^2(0) - r_x^2(n_B - n_A)} \begin{bmatrix} r_x(0) & -r_x(n_B - n_A) \\ -r_x(n_B - n_A) & r_x(0) \end{bmatrix} \begin{bmatrix} r_x(n - n_A) \\ r_x(n - n_B) \end{bmatrix}$$

$$= \frac{1}{r_x^2(0) - r_x^2(n_B - n_A)} \begin{bmatrix} r_x(0)r_x(n_A) - r_x(n_B - n_A)r_x(n - n_B) \\ -r_x(n - n_A)r_x(n_B - n_A) + r_x(0)r_x(n - n_B) \end{bmatrix}$$

With $r_x(k) = \sigma_x^2 \alpha^{|k|}$, this becomes

$$\begin{bmatrix} a(n) \\ b(n) \end{bmatrix} = \frac{1}{1 - \alpha^{2|n_B - n_A|}} \begin{bmatrix} \alpha^{|n - n_A|} - \alpha^{|n_B - n_A| + |n - n_A|} \\ \alpha^{|n - n_A| + |n_B - n_A|} - \alpha^{|n - n_B|} \end{bmatrix}$$

Note that this is the <u>general</u> solution which is valid for all n_A and n_B.

(a) For the special case in which $n > n_B$ we have (recall that $n_B > n_A$)

$$\begin{bmatrix} a(n) \\ b(n) \end{bmatrix} = \begin{bmatrix} 0 \\ \alpha^{|n - n_B|} \end{bmatrix}$$

(b) For the special case in which $n < n_A$ we have

$$\begin{bmatrix} a(n) \\ b(n) \end{bmatrix} = \begin{bmatrix} \alpha^{|n - n_A|} \\ 0 \end{bmatrix}$$

(c) For $n_A < n < n_B$,

$$\begin{bmatrix} a(n) \\ b(n) \end{bmatrix} = \frac{1}{1 - \alpha^{2|n_B - n_A|}} \begin{bmatrix} \alpha^{(n - n_A)} - \alpha^{(2n_B - n_A - n)} \\ \alpha^{(n + n_B - 2n_A)} - \alpha^{(n_B - n)} \end{bmatrix}$$

7.14 As shown in Figure 7.12, the Wiener filter may be viewed as a cascade of a whitening filter with a causal filter that produces the minimum mean-square estimate of $d(n)$ from $\epsilon(n)$. For real processes, the system function of the cascade is

$$H(z) = F(z)G(z) = \frac{1}{\sigma_0^2 Q(z)} \left[\frac{P_{dx}(z)}{Q(z^{-1})} \right]_+$$

and the mean-square error is

$$\xi_{\min} = r_d(0) - \sum_{l=0}^{\infty} h(l)r_{dx}(l)$$

(a) If $r_{d\epsilon}(k) = \delta(k)$ and

$$P_x(z) = \frac{4}{(1 - 0.5z^{-1})(1 - 0.5z)}$$

find the unit sample response, $h(n)$, of the causal Wiener filter.

(b) Derive an expression for the mean-square error that expresses $\xi_{\min}$ in terms of the cross correlation, $r_{d\epsilon}(k)$, and evaluate the mean-square error when

$$r_{d\epsilon}(k) = \left(\tfrac{1}{2}\right)^k u(k) + \left(\tfrac{1}{3}\right)^{-k} u(-k - 1)$$

and

$$E\{d^2(n)\} = 4$$

Solution _____

(a) If $r_{d\epsilon}(k) = \delta(k)$, then $P_{d\epsilon}(z) = 1$ and the causal Wiener filter for estimating $d(n)$ from $\epsilon(n)$ is

$$G(z) = \left[P_{d\epsilon}(z) \right]_+ = 1$$

Therefore, the causal Wiener filter is the whitening filter for $x(n)$,

$$H(z) = \frac{1}{\sigma_0 Q(z)} = \tfrac{1}{2}\left(1 - 0.5z^{-1}\right)$$

so the unit sample response is

$$h(n) = \tfrac{1}{2}\delta(n) + \tfrac{1}{4}\delta(n - 1)$$

(b) The mean-square error is

$$\xi_{\min} = r_d(0) - \sum_{k=0}^{\infty} h(k)r_{dx}(k) = r_d(0) - \frac{1}{2\pi j} \oint_C H(z)P_{dx}^*(1/z^*)z^{-1}dz$$

Note that

$$H(z) = \frac{1}{\sigma_0^2 Q(z)} \left[\frac{P_{dx}(z)}{Q^*(1/z^*)} \right]_+$$

where

$$P_{d\epsilon}(z) = \frac{P_{dx}(z)}{\sigma_0 Q^*(1/z^*)}$$

Therefore,

$$P_{dx}(z) = \sigma_0 Q^*(1/z^*) P_{d\epsilon}(z)$$

and we may write the term inside the integral as follows

$$H(z)P_{dx}^*(1/z^*) = \frac{1}{\sigma_0^2 Q(z)} \left[\frac{\sigma_0 Q^*(1/z^*) P_{d\epsilon}(z)}{Q^*(1/z^*)} \right]_+ \sigma_0 Q(z) P_{d\epsilon}^*(1/z^*) = \left[P_{d\epsilon}(z) \right]_+ P_{d\epsilon}^*(1/z^*)$$

Thus, the minimum mean-square error is

$$\xi_{\min} = r_d(0) - \frac{1}{2\pi j} \oint_C [P_{d\epsilon}(z)]_+ P_{d\epsilon}^*(1/z^*) z^{-1} dz = r_d(0) \quad \sum_{k=0}^{\infty} |r_{d\epsilon}(k)|^2$$

For the given cross-correlation sequence, $r_{d\epsilon}(k)$, the minimum error is

$$\xi_{\min} = 4 - \sum_{k=0}^{\infty} (\tfrac{1}{2})^{2k} = \tfrac{8}{3}$$

7.15 Let $x(n)$ be an AR(1) process of the following form

$$x(n) = a(1)x(n-1) + b(0)w(n)$$

where $w(n)$ is unit variance white noise, and let $y(n)$ be noisy measurements

$$y(n) = x(n) + v(n)$$

where $v(n)$ is unit variance white noise that is uncorrelated with $w(n)$. We have seen that the causal Wiener filter for estimating $x(n)$ from $y(n)$ has the form

$$\widehat{x}(n) = a(1)\widehat{x}(n-1) + K\Big[y(n) - a(1)\widehat{x}(n-1)\Big]$$

Find the value of K in terms of $a(1)$ and $b(0)$ that minimizes the mean-square error

$$E\{[x(n) - \widehat{x}(n)]^2\}$$

Solution _____

With an estimate of the form

$$\widehat{x}(n) = a(1)\widehat{x}(n-1) + K\Big[y(n) - a(1)\widehat{x}(n-1)\Big]$$

we want to find the value of K that minimizes the mean-square error

$$\xi = E\Big\{[x(n) - \widehat{x}(n)]^2\Big\}$$

This problem may be solved by differentiating ξ with respect to K, and set the result equal to zero. After a fair amount of work, we find that

$$K = \frac{b^2(0) + a^2(1)\xi_{\min}}{1 + b^2(0) + a^2(1)\xi_{\min}}$$

Unfortunately, however, $\xi_{\min}$ depends upon K. Using the expression

$$\xi_{\min} = r_x(0) - \sum_{l=0}^{\infty} h(l)r_{xy}^*(l)$$

with

$$h(n) = K\Big([1 - K]a(1)\Big)^n u(n)$$

and

$$r_{xy}(k) = r_x(k) = \frac{b^2(0)}{1 - a^2(1)}a(1)^{|k|}$$

we may easily derive the following expression for $\xi_{\min}$

$$\xi_{\min} = b^2(0)\frac{1 - K}{1 - (1 - K)a^2(1)}$$

Solving these two equations for K leads to the following quadratic equation,

$$a^2(1)K^2 + [1 + b^2(0) - a^2(1)]K - b^2(0) = 0$$

and the desired solution is the positive real root of this quadratic. Note that if we substitute $a(1) = 0.8$ and $b(0) = 0.6$ we arrive at the values for K and $\xi_{\min}$ derived in Example 7.3.2.

7.16 The derivation of the Kalman filtering equations for real-valued signals make use of the following matrix differentiation formulas

$$\frac{d}{d\mathbf{K}}\text{tr}(\mathbf{KA}) = \mathbf{A}^T$$

and

$$\frac{d}{d\mathbf{K}}\text{tr}(\mathbf{KAK}^T) = 2\mathbf{KA}$$

where $\mathbf{A}$ is a *symmetric* matrix.

(a) Show that these matrix differentiation formulas are valid.

(b) Derive the equivalent expression for complex data.

(c) Use these matrix differentiation formulas to derive the expression for the Kalman gain given in Eq. (7.113).

Solution _____

(a) With $\mathbf{K}$ a $p \times q$ matrix and $\mathbf{A}$ a $q \times p$ matrix, the (i, i)th element of the matrix product $\mathbf{KA}$ is

$$\left\{ \mathbf{KA} \right\}_{(i,i)} = \sum_{m=1}^{q} k(i, m)a(m, i)$$

Therefore, the trace is

$$\text{Tr}(\mathbf{KA}) = \sum_{i=1}^{p} \sum_{m=1}^{q} k(i, m)a(m, i)$$

and the derivative, with respect to $k(i, j)$, is

$$\frac{\partial}{\partial k(i, j)}\text{Tr}(\mathbf{KA}) = a(j, i)$$

Thus,

$$\frac{d}{d\mathbf{K}}\text{tr}(\mathbf{KA}) = \mathbf{A}^T$$

For the matrix $\mathbf{KAK}^T$, the (i, i)th element is

$$\left\{ \mathbf{KAK}^T \right\}_{(i,i)} = \sum_{m=1}^{q} \sum_{n=1}^{q} k(i, n)a(n, m)k(i, m)$$

Therefore, the trace is

$$\text{Tr}(\mathbf{KAK}^T) = \sum_{i=1}^{p} \sum_{m=1}^{q} \sum_{n=1}^{q} k(i, n)a(n, m)k(i, m)$$

and the derivative with respect to $k(i, j)$ is

$$\frac{\partial}{\partial k(i, j)}\text{Tr}(\mathbf{KAK}^T) = \sum_{m=1}^{q} a(j, m)k(i, m) + \sum_{n=1}^{q} k(i, n)a(n, j)$$

or, since $\mathbf{A}$ is symmetric,

$$\frac{\partial}{\partial k(i,j)} \text{Tr}(\mathbf{K}\mathbf{A}\mathbf{K}^T) = 2\sum_{n=1}^{q} k(i,n)a(n,j) = 2\mathbf{K}\mathbf{A}$$

(b) For complex data, we treat $\mathbf{K}$ and $\mathbf{K}^*$ as independent variables. Differentiating with respect to $\mathbf{K}^*$ we have

$$\frac{d}{d\mathbf{K}^*}\text{tr}\left(\mathbf{A}\mathbf{K}^H\right) = \mathbf{A}$$

$$\frac{d}{d\mathbf{K}^*}\text{tr}\left(\mathbf{K}\mathbf{A}\right) = \mathbf{0}$$

and

$$\frac{d}{d\mathbf{K}^*}\text{tr}\left(\mathbf{K}\mathbf{A}\mathbf{K}^H\right) = \mathbf{K}\mathbf{A}$$

(c) Given these matrix differential formulas, the derivation of the expression for the Kalman gain is straightforward. First, with the error covariance matrix given by

$$\mathbf{P}(n|n) = \left[\mathbf{I} - \mathbf{K}(n)\mathbf{C}(n)\right]\mathbf{P}(n|n-1)\left[\mathbf{I} - \mathbf{K}(n)\mathbf{C}(n)\right]^H + \mathbf{K}(n)\mathbf{Q}_v(n)\mathbf{K}^H(n)$$

we expand it as follows

$$\begin{aligned}\mathbf{P}(n|n) \;=\; & \mathbf{P}(n|n-1) - \mathbf{K}(n)\mathbf{C}(n)\mathbf{P}(n|n-1) + \mathbf{P}(n|n-1)\mathbf{C}^H(n)\mathbf{K}^H(n) \\ & + \mathbf{K}(n)\mathbf{C}(n)\mathbf{P}(n|n-1)\mathbf{C}^H(n)\mathbf{K}^H(n) + \mathbf{K}(n)\mathbf{Q}_v^H\mathbf{K}^H(n)\end{aligned}$$

Since the trace of a sum of matrices is the sum of the traces, using the matrix differentiation formulas above we have

$$\frac{d}{d\mathbf{K}}\text{tr}\{\mathbf{P}(n|n)\} = -\left[\mathbf{I} - \mathbf{K}(n)\mathbf{C}(n)\right]\mathbf{P}(n|n-1)\mathbf{C}^H(n) + \mathbf{K}(n)\mathbf{Q}_v(n) = 0$$

Finally, solving for $\mathbf{K}(n)$ gives the desired expression for the Kalman gain,

$$\mathbf{K}(n) = \mathbf{P}(n|n-1)\mathbf{C}^H(n)\left[\mathbf{C}(n)\mathbf{P}(n|n-1)\mathbf{C}^H(n) + \mathbf{Q}_v(n)\right]^{-1}$$

7.17 Consider a system consisting of two sensors, each making a single measurement of an unknown constant x. Each measurement is noisy and may be modeled as follows

$$
\begin{aligned}
y(1) &= x + v(1) \\
y(2) &= x + v(2)
\end{aligned}
$$

where $v(1)$ and $v(2)$ are zero mean uncorrelated random variables with variance σ_1^2 and σ_2^2, respectively.

(a) In the absence of any other information, we seek the best linear estimate of x of the form

$$
\widehat{x} = k_1 y(1) + k_2 y(2)
$$

Find the values for k_1 and k_2 that yield an unbiased estimate of x that minimizes the mean-square error, $E\{[x - \widehat{x}]^2\}$.

(b) Repeat part (a) for the case where the measurement errors are correlated,

$$
E\{v(1)v(2)\} = \rho \sigma_1 \sigma_2
$$

(c) Repeat part (a) within the framework of Kalman filtering, treating the measurements $y(1)$ and $y(2)$ sequentially.

Solution _____

(a) For an unbiased estimate, we want

$$
E\{x - \widehat{x}\} = E\{x\} - k_1 E\{y(1)\} - k_2 E\{y(2)\} = 0
$$

Since $v(1)$ and $v(2)$ are zero mean, and $E\{x\} = x$ (x is an unknown constant), then

$$
E\{x - \widehat{x}\} = x - (k_1 + k_2)x = 0
$$

Therefore, we want

$$
k_1 + k_2 = 1
$$

or

$$
k_2 = 1 - k_1
$$

The mean-square error that we want to minimize is

$$
\begin{aligned}
E\left\{[x - \widehat{x}]^2\right\} &= E\left\{[x - k_1 y(1) - (1 - k_1)y(2)]^2\right\} \\
&= E\left\{[-k_1 v(1) - (1 - k_1)v(2)]^2\right\} \\
&= k_1^2 \sigma_1^2 + (1 - k_1)^2 \sigma_2^2
\end{aligned}
$$

where the last equality follows from the fact that $v(1)$ and $v(2)$ are uncorrelated. To find the value for k_1 that minimizes the mean-square error, we set the derivative with respect to k_1 equal to zero,

$$
\frac{\partial}{\partial k_1} E\{[x - \widehat{x}]^2\} = 2k_1 \sigma_1^2 - 2(1 - k_1)\sigma_2^2 = 0
$$

Solving for k_1 we find

$$k_1 = \frac{\sigma_2^2}{\sigma_1^2 + \sigma_2^2}$$

and, therefore,

$$k_2 = 1 - k_1 = \frac{\sigma_1^2}{\sigma_1^2 + \sigma_2^2}$$

Thus, the estimate for x is

$$\widehat{x} = \frac{\sigma_2^2}{\sigma_1^2 + \sigma_2^2} \cdot y(1) + \frac{\sigma_1^2}{\sigma_1^2 + \sigma_2^2} \cdot y(2)$$

(b) If $v(1)$ and $v(2)$ are correlated, the only change required in part (a) comes when we evaluate the mean-square error, which is

$$
\begin{aligned}
E\left\{[x - \widehat{x}]^2\right\} &= E\left\{\left[-k_1 v(1) - (1 - k_1)v(2)\right]^2\right\} \\
&= k_1^2 \sigma_1^2 + 2k_1(1 - k_1)\rho\sigma_1\sigma_2 + (1 - k_1)^2\sigma_2^2
\end{aligned}
$$

Setting the derivative of the mean-square error with respect to k_1 equal to zero we have

$$\frac{\partial}{\partial k_1} E\left\{[x - \widehat{x}]^2\right\} = 2k_1\sigma_1^2 + 2(1 - 2k_1)\rho\sigma_1\sigma_2 - 2(1 - k_1)\sigma_2^2 = 0$$

Solving for k_1 we find

$$k_1 = \frac{\sigma_2^2 - \rho\sigma_1\sigma_2}{\sigma_1^2 - 2\rho\sigma_1\sigma_2 + \sigma_2^2}$$

Finally, for k_2 we have

$$k_2 = 1 - k_1 = \frac{\sigma_1^2 - \rho\sigma_1\sigma_2}{\sigma_1^2 - 2\rho\sigma_1\sigma_2 + \sigma_2^2}$$

(c) In the context of a Kalman filter, this problem is the same as the problem worked in Example 7.4.1, except that the variance of $v(n)$ is not a constant, since it changes from σ_1^2 to σ_2^2. For the state equation we have

$$x(n) = x(n - 1)$$

and the measurement equation is

$$y(n) = x(n) + v(n)$$

Therefore, $\mathbf{A}(n) = 1$, $\mathbf{C}(n) = 1$, $\mathbf{Q}_w(n) = 0$, and $\mathbf{Q}_v(n) = \sigma_n^2$. Since $\mathbf{A}(n) = 1$ and $\mathbf{Q}_w(n) = 0$, it follows that $P(n|n - 1)$ and $P(n - 1|n - 1)$ are equal,

$$P(n|n - 1) = P(n - 1|n - 1)$$

and, as in Example 7.4.1, we will simplify the notation and write $P(n)$ instead of $P(n|n)$. For the Kalman gain we have

$$K(n) = P(n - 1)\left[P(n - 1) + \sigma_n^2\right]^{-1}$$

and the update for $P(n|n)$ is

$$
\begin{aligned}
P(n) &= \left[1 - K(n)\right]P(n - 1) \\
&= \left[1 - \frac{P(n - 1)}{P(n - 1) + \sigma_n^2}\right]^{-1} P(n - 1) \\
&= \frac{P(n - 1)\sigma_n^2}{P(n - 1) + \sigma_n^2}
\end{aligned}
$$

Let us now find the first two values for the Kalman gain. With $P(0) = \infty$ we have

$$K(1) = \frac{P(0)}{P(0) + \sigma_1^2} = 1$$

and with

$$P(1) = \frac{P(0)\sigma_1^2}{P(0) + \sigma_1^2} = \sigma_1^2$$

we have

$$K(2) = \frac{P(1)}{P(1) + \sigma_2^2} = \frac{\sigma_1^2}{\sigma_1^2 + \sigma_2^2}$$

Now we may recursively estimate x as follows. With $\widehat{x}(0) = 0$, after the first observation we have

$$\widehat{x}(1) = \widehat{x}(0) + K(1)\big[y(1) - \widehat{x}(0)\big] = y(1)$$

Then, with the second measurement we have

$$
\begin{aligned}
\widehat{x}(2) &= \widehat{x}(1) + K(2)\big[y(2) - \widehat{x}(1)\big] \\
&= y(1) + \frac{\sigma_1^2}{\sigma_1^2 + \sigma_2^2}\big[y(2) - \widehat{x}(1)\big] \\
&= \frac{\sigma_2^2}{\sigma_1^2 + \sigma_2^2}y(1) + \frac{\sigma_1^2}{\sigma_1^2 + \sigma_2^2}y(2)
\end{aligned}
$$

which, of course, is the same as we derived in part (a).

7.18 An autoregressive process of order 1 is described by the difference equation

$$x(n) = 0.5x(n-1) + w(n)$$

where $w(n)$ is zero-mean white noise with a variance $\sigma_w^2 = 0.64$. The observed process $y(n)$ is described by

$$y(n) = x(n) + v(n)$$

where $v(n)$ is zero-mean white noise with a variance $\sigma_v^2 = 1$.

(a) Write the Kalman filtering equations to find the minimum mean-square estimate, $\hat{x}(n|n)$, of $x(n)$ given the observations $y(i)$, $i = 1, \ldots, n$. The initial conditions are $\hat{x}(0|0) = 0$ and $E\{\epsilon^2(0|0)\} = 1$ where $\epsilon(0|0) = x(0) - \hat{x}(0|0)$.

(b) Assuming that the filter reaches a steady state solution, find the steady state Kalman gain and the limiting form of the estimation equation for $\hat{x}(n|n)$.

Solution _____

(a) With the state and observation equation given by

$$
\begin{aligned}
x(n+1) &= 0.5x(n) + w(n) \\
y(n) &= x(n) + v(n)
\end{aligned}
$$

we see that $A = 0.5$ and $C = 1$. Therefore, the Kalman filtering algorithm is

1. $\hat{x}(0) = 0$, $P(0|0) = 1$.
2. For $n = 1, 2, \ldots$
   (a) $P(n|n-1) = (0.5)^2 P(n-1|n-1) + 0.64$
   (b) $K(n) = P(n|n-1) \left[P(n|n-1) + 1 \right]^{-1}$
   (c) $P(n|n) = [1 - K(n)] P(n|n-1)$
   (d) $\hat{x}(n) = 0.5\hat{x}(n-1) + K(n)\left[y(n) - 0.5\hat{x}(n-1) \right]$

(b) From the Kalman filtering equations we have for $P(n|n)$ we have

$$
\begin{aligned}
P(n|n) &= \left[1 - K(n)\right] P(n|n-1) = \left[1 - \frac{P(n|n-1)}{P(n|n-1)+1}\right] P(n|n-1) \\
&= \frac{P(n|n-1)}{P(n|n-1)+1}
\end{aligned}
$$

Thus,

$$P(n|n) = \frac{(0.5)^2 P(n-1|n-1) + 0.64}{1 + (0.5)^2 P(n-1|n-1) + 0.64}$$

In the steady state, $P(n+1|n+1) = P(n|n)$, and we have

$$P(n|n) = \frac{0.25P(n|n) + 0.64}{1.64 + 0.25P(n|n)}$$

Simplifying we have

$$0.25P^2(n|n) + 1.39P(n|n) - 0.64P(n|n) = 0$$

Solving for $P(n|n)$ we find that there is only one positive root, which is

$$P(n|n) = 0.4559$$

Since $K(n) = P(n|n)$, then the steady state estimation equation is

$$\widehat{x}(n) = 0.5\widehat{x}(n-1) + 0.4559\big[y(n) - 0.5\widehat{x}(n-1)\big]$$

7.19 In many cases, the error covariance matrix $\mathbf{P}(n|n-1)$ will converge to a steady-state value $\mathbf{P}$ as $n \to \infty$. Assume that $\mathbf{C}$, $\mathbf{Q}_w$, and $\mathbf{Q}_v$ are the limiting values of $\mathbf{C}(n)$, $\mathbf{Q}_w(n)$, and $\mathbf{Q}_v(n)$, respectively.

(a) For $\mathbf{A}(n) = \mathbf{I}$, show that if $\mathbf{P}(n|n-1)$ converges to a steady state value $\mathbf{P}$, then the limiting value satisfies the *algebraic Ricatti equation*

$$\mathbf{P}\mathbf{C}^H(\mathbf{C}\mathbf{P}\mathbf{C}^H + \mathbf{Q}_v)^{-1}\mathbf{C}\mathbf{P} - \mathbf{Q}_w = 0$$

(b) Derive the Ricatti equation for a general state transition matrix $\mathbf{A}(n)$ that has a limiting value of $\mathbf{A}$.

Solution _____

(a) From the Kalman filtering equations, we have

$$\mathbf{P}(n|n-1) = \mathbf{A}(n-1)\mathbf{P}(n-1|n-1)\mathbf{A}^H(n-1) + \mathbf{Q}_w(n)$$

With $\mathbf{A}(n) = \mathbf{I}$, this becomes

$$\mathbf{P}(n|n-1) = \mathbf{P}(n-1|n-1) + \mathbf{Q}_w(n)$$

For the error covariance matrix, $\mathbf{P}(n|n)$, we have

$$\mathbf{P}(n|n) = \left[\mathbf{I} - \mathbf{K}(n)\mathbf{C}(n)\right]\mathbf{P}(n|n-1)$$

Substituting this into the expression above for $\mathbf{P}(n|n-1)$ gives

$$\mathbf{P}(n|n-1) = \left[\mathbf{I} - \mathbf{K}(n-1)\mathbf{C}(n-1)\right]\mathbf{P}(n-1|n-2) + \mathbf{Q}_w(n)$$

In the steady state, $\mathbf{P}(n|n-1) = \mathbf{P}$, $\mathbf{C}(n) = \mathbf{C}$, and $\mathbf{K}(n) = \mathbf{K}$. Incorporating the steady state conditions into the equation above for $\mathbf{P}(n|n-1)$ we have

$$\mathbf{P} = \left[\mathbf{I} - \mathbf{K}\mathbf{C}\right]\mathbf{P} + \mathbf{Q}_w$$

Cancelling $\mathbf{P}$ on both sides of the equation and rearranging yields,

$$\mathbf{K}\mathbf{C}\mathbf{P} - \mathbf{Q}_w = 0 \qquad\qquad\qquad (\text{P7.19-1})$$

Now, for the Kalman gain, we have

$$\mathbf{K}(n) = \mathbf{P}(n|n-1)\mathbf{C}^H(n)\left[\mathbf{C}(n)\mathbf{P}(n|n-1)\mathbf{C}^H(n) + \mathbf{Q}_v(n)\right]^{-1}$$

which, in steady state, becomes

$$\mathbf{K} = \mathbf{P}\mathbf{C}^H\left[\mathbf{C}\mathbf{P}\mathbf{C}^H + \mathbf{Q}_v\right]^{-1}$$

Substituting this into Eq. (P7.19-1) gives the desired equation

$$\mathbf{P}\mathbf{C}^H\left[\mathbf{C}\mathbf{P}\mathbf{C}^H + \mathbf{Q}_v\right]^{-1}\mathbf{C}\mathbf{P} - \mathbf{Q}_w = 0$$

(b) For the more general case in which there is a general state transition matrix $\mathbf{A}$, the only equation that contains the matrix $\mathbf{A}$ is the update equation for $\mathbf{P}(n|n-1)$ which becomes, in steady state,

$$\mathbf{P} = \mathbf{A}\Big[\mathbf{I} - \mathbf{KC}\Big]\mathbf{PA}^H + \mathbf{Q}_w$$

Substituting in the steady state value for $\mathbf{K}$ we have

$$\mathbf{P} = \mathbf{A}\Big[\mathbf{I} - \mathbf{PC}^H\big(\mathbf{CPC}^H + \mathbf{Q}_v\big)^{-1}\mathbf{C}\Big]\mathbf{PA}^H + \mathbf{Q}_w$$

which is the desired result.

7.20 In Example 7.4.1 we derived the Kalman filter for estimating an unknown constant from noisy measurements. The estimate at time n was shown to be

$$\widehat{x}(n) = \widehat{x}(n-1) + \frac{P(0)}{nP(0) + \sigma_v^2}\Big[y(n) - \widehat{x}(n-1)\Big]$$

(a) Solve this difference equation and find a closed-form expression for $\widehat{x}(n)$ in terms of $\widehat{x}(0)$ and the measurements $y(0), y(1), \ldots, y(n)$.

(b) What does $\widehat{x}(n)$ converge to as $n \to \infty$?

Solution _____

(a) Writing the estimate $\widehat{x}(n)$ in the form

$$\widehat{x}(n) = \widehat{x}(n-1) + \frac{\sigma_v^{-2}P(0)}{1 + n\sigma_v^2 P(0)}\Big[y(n) - \widehat{x}(n-1)\Big]$$

we may derive a closed-form expression for $\widehat{x}(n)$ as follows, First, note that for $\widehat{x}(1)$ we have

$$\widehat{x}(1) = \widehat{x}(0) + \frac{\sigma_v^{-2}P(0)}{1 + \sigma_v^{-2}P(0)}\Big[y(1) - \widehat{x}(0)\Big] = \frac{\widehat{x}(0) + \sigma_v^{-2}P(0)y(1)}{1 + \sigma_v^{-2}P(0)}$$

Then, for $\widehat{x}(2)$ we have

$$
\begin{aligned}
\widehat{x}(2) &= \widehat{x}(1) + \frac{\sigma_v^{-2}P(0)}{1 + 2\sigma_v^{-2}P(0)}\big[y(2) - \widehat{x}(1)\big] \\[2mm]
&= \frac{[1 + \sigma_v^{-2}P(0)]\widehat{x}(1) + \sigma_v^{-2}P(0)y(2)}{1 + 2\sigma_v^{-2}P(0)} \\[2mm]
&= \frac{\widehat{x}(0) + \sigma_v^{-2}P(0)y(1) + \sigma_v^{-2}P(0)y(2)}{1 + 2\sigma_v^{-2}P(0)y(2)}
\end{aligned}
$$

and it follows by induction, that

$$\widehat{x}(n) = \frac{1}{1 + n\sigma_v^{-2}P(0)}\Big[\widehat{x}(0) + \sigma_v^{-2}P(0)\sum_{k=0}^{n} y(k)\Big]$$

(b) As $n \to \infty$

$$\widehat{x}(n) \to \frac{1}{n}\sum_{k=0}^{n} y(k)$$

7.21 In this problem we will derive the following expression for the Kalman gain,

$$\mathbf{K}(n) = \mathbf{P}(n|n)\mathbf{C}^H(n)\mathbf{Q}_v^{-1}(n) \tag{P7.21-1}$$

(a) By substituting Eq. (7.113) for the Kalman gain into Eq. (7.114), show that the error covariance matrix $\mathbf{P}(n|n)$ may be written as

$$\mathbf{P}(n|n) = \mathbf{P}(n|n-1) - \mathbf{P}(n|n-1)\mathbf{C}^H(n)[\mathbf{C}(n)\mathbf{P}(n|n-1)\mathbf{C}^H(n) + \mathbf{Q}_v(n)]^{-1}\mathbf{C}(n)\mathbf{P}(n|n-1)$$

(b) Using the matrix inversion lemma given in Eq. (2.28) on p. 29, show that the inverse covariance matrix may be written as

$$\mathbf{P}^{-1}(n|n) = \mathbf{P}^{-1}(n|n-1) + \mathbf{C}^H(n)\mathbf{Q}_v^{-1}(n)\mathbf{C}(n)$$

(c) By multiplying the expression for the Kalman gain given in Eq. (7.113) on the left by $\mathbf{P}(n|n)\mathbf{P}^{-1}(n|n)$, use your results in part (b) to derive the expression for $\mathbf{K}(n)$ given in Eq. (P7.21-1).

Solution _____

(a) This part follows by inspection. With

$$\mathbf{P}(n|n) = \Big[\mathbf{I} - \mathbf{K}(n)\mathbf{C}(n)\Big]\mathbf{P}(n|n-1)$$

and a Kalman gain given by

$$\mathbf{K}(n) = \mathbf{P}(n|n-1)\mathbf{C}^H(n)\Big[\mathbf{C}(n)\mathbf{P}(n|n-1)\mathbf{C}^H(n) + \mathbf{Q}_v(n)\Big]^{-1}$$

we have

$$\begin{aligned}
\mathbf{P}(n|n) &= \Big[\mathbf{I} - \mathbf{P}(n|n-1)\mathbf{C}^H(n)\Big[\mathbf{C}(n)\mathbf{P}(n|n-1)\mathbf{C}^H(n) + \mathbf{Q}_v(n)\Big]^{-1}\mathbf{C}(n)\Big]\mathbf{P}(n|n-1) \\
&= \mathbf{P}(n|n-1) - \mathbf{P}(n|n-1)\mathbf{C}^H(n)\Big[\mathbf{C}(n)\mathbf{P}(n|n-1)\mathbf{C}^H(n) + \mathbf{Q}_v(n)\Big]^{-1}\mathbf{C}(n)\mathbf{P}(n|n-1)
\end{aligned}$$

(b) The matrix that we want to invert is

$$\mathbf{P}(n|n) = \mathbf{P}(n|n-1) - \mathbf{P}(n|n-1)\mathbf{C}^H(n)\Big[\mathbf{C}(n)\mathbf{P}(n|n-1)\mathbf{C}^H(n) + \mathbf{Q}_v(n)\Big]^{-1}\mathbf{C}(n)\mathbf{P}(n|n-1)$$

The matrix inversion lemma states that

$$(\mathbf{A} + \mathbf{BCD})^{-1} = \mathbf{A}^{-1} - \mathbf{A}^{-1}\mathbf{B}(\mathbf{C}^{-1} + \mathbf{DA}^{-1}\mathbf{B})^{-1}\mathbf{DA}^{-1}$$

Therefore, we define

$$\begin{aligned}
\mathbf{A} &= \mathbf{P}(n|n-1) \\
\mathbf{B} &= -\mathbf{P}(n|n-1)\mathbf{C}^H(n) \\
\mathbf{C} &= \Big[\mathbf{C}(n)\mathbf{P}(n|n-1)\mathbf{C}^H(n) + \mathbf{Q}_v(n)\Big]^{-1} \\
\mathbf{D} &= \mathbf{C}(n)\mathbf{P}(n|n-1)
\end{aligned}$$

Note that

$$(\mathbf{C}^{-1} + \mathbf{D}\mathbf{A}^{-1}\mathbf{B})^{-1} = \left[\mathbf{C}(n)\mathbf{P}(n|n-1)\mathbf{C}^H(n) + \mathbf{Q}_v(n) - \mathbf{C}(n)\mathbf{P}(n|n-1)\mathbf{P}^{-1}(n|n-1)\mathbf{P}(n|n-1)\mathbf{C}^H(n)\right.$$
$$= \mathbf{Q}_v^{-1}(n)$$

Therefore, $\mathbf{P}^{-1}(n|n)$ is

$$\mathbf{P}^{-1}(n|n) = \mathbf{P}^{-1}(n|n-1) + \mathbf{P}^{-1}(n|n-1)\mathbf{P}(n|n-1)\mathbf{C}^H(n)\mathbf{Q}_v^{-1}(n)\mathbf{C}(n)\mathbf{P}(n|n-1)\mathbf{P}^{-1}(n|n-1)$$
$$= \mathbf{P}^{-1}(n|n-1) + \mathbf{C}^H(n)\mathbf{Q}_v^{-1}(n)\mathbf{C}(n)$$

which is what we wanted to show.

(c) The Kalman gain is

$$\mathbf{K}(n) = \mathbf{P}(n|n-1)\mathbf{C}^H(n)\left[\mathbf{C}(n)\mathbf{P}(n|n-1)\mathbf{C}^H(n) + \mathbf{Q}_v(n)\right]^{-1}$$

Multiplying on the left by $\mathbf{P}(n|n)\mathbf{P}^{-1}(n|n)$ gives

$$\mathbf{K}(n) = \mathbf{P}(n|n)\mathbf{P}^{-1}(n|n)\mathbf{P}(n|n-1)\mathbf{C}^H(n)\left[\mathbf{C}(n)\mathbf{P}(n|n-1)\mathbf{C}^H(n) + \mathbf{Q}_v(n)\right]^{-1}$$

Substituting the expression for $\mathbf{P}^{-1}(n|n)$ derived in part (b) yields

$$\mathbf{K}(n) = \mathbf{P}(n|n)\left[\mathbf{P}^{-1}(n|n-1) + \mathbf{C}^H(n)\mathbf{Q}_v^{-1}(n)\mathbf{C}(n)\right]\mathbf{P}(n|n-1)\mathbf{C}^H(n)\left[\mathbf{C}(n)\mathbf{P}(n|n-1)\mathbf{C}^H(n) + \mathbf{Q}_v(n)\right]^{-1}$$
$$= \mathbf{P}(n|n)\left[\mathbf{C}^H(n) + \mathbf{C}^H(n)\mathbf{Q}_v^{-1}(n)\mathbf{C}(n)\mathbf{P}(n|n-1)\mathbf{C}^H(n)\right]\left[\mathbf{C}(n)\mathbf{P}(n|n-1)\mathbf{C}^H(n) + \mathbf{Q}_v(n)\right]^{-1}$$
$$= \mathbf{P}(n|n)\mathbf{C}^H(n)\mathbf{Q}_v^{-1}(n)\left[\mathbf{Q}_v(n) + \mathbf{C}(n)\mathbf{P}(n|n-1)\mathbf{C}^H(n)\right]\left[\mathbf{C}(n)\mathbf{P}(n|n-1)\mathbf{C}^H(n) + \mathbf{Q}_v(n)\right]^{-1}$$
$$= \mathbf{P}(n|n)\mathbf{C}^H(n)\mathbf{Q}_v^{-1}(n)$$

which is the desired relationship.

7.22 Consider the ARMA(1,1) process $y(n)$ given by

$$y(n) + ay(n-1) = w(n) + bw(n-1)$$

where $w(n)$ is a zero-mean white-noise process with a variance σ_w^2.

(a) Show that the state-space representation for this process may be written as

$$\mathbf{x}(n) = \begin{bmatrix} -a & 1 \\ 0 & 0 \end{bmatrix} \mathbf{x}(n-1) + \begin{bmatrix} 1 \\ b \end{bmatrix} w(n)$$

$$y(n) = \begin{bmatrix} 1 & 0 \end{bmatrix} \mathbf{x}(n)$$

where $\mathbf{x}(n)$ is a two-dimensional state vector.

(b) Assuming that the error covariance $\mathbf{P}(n|n)$ converges to a steady state value of $\mathbf{P}$ and is a solution of the Ricatti equation given in Prob. 7.19, show that

$$\mathbf{P} = \sigma_v^2 \begin{bmatrix} 1+c & b \\ b & b^2 \end{bmatrix}$$

where c is a scalar that satisfies the second-order equation

$$c = (b-a)^2 + a^2 c - \frac{(b-a-ac)^2}{1+c}$$

and find the two values of c that satisfy this equation. For each of these values, find the corresponding values for $\mathbf{P}$.

(c) Find the steady-state Kalman gain and determine the values for $\mathbf{K}$ that correspond to the solutions for c found in part (b).

Solution _____

(a) Let

$$\mathbf{x}(n) = \begin{bmatrix} x_1(n) \\ x_2(n) \end{bmatrix}$$

Then we have

$$\begin{bmatrix} x_1(n) \\ x_2(n) \end{bmatrix} = \begin{bmatrix} -a & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x_1(n) \\ x_2(n) \end{bmatrix} + \begin{bmatrix} 1 \\ b \end{bmatrix} w(n)$$

Thus,

$$\begin{aligned} x_1(n) &= -ax_1(n-1) + x_2(n-1) + w(n) \\ x_2(n) &= bw(n) \end{aligned}$$

and

$$x_1(n) = -ax_1(n-1) + bw(n-1) + w(n)$$

which, with

$$y(n) = \begin{bmatrix} 1 & 0 \end{bmatrix} \mathbf{x}(n)$$

is the equation we have for the ARMA(1,1) process.

(b) For the state variable representation of the ARMA process, we have

$$\mathbf{A} = \begin{bmatrix} -a & 1 \\ 0 & 0 \end{bmatrix}$$

$$\mathbf{Q}_w = E\left\{ \begin{bmatrix} v(n) \\ bv(n) \end{bmatrix} \begin{bmatrix} v(n) & bv(n) \end{bmatrix}^2 \right\} = \sigma_w^2 \begin{bmatrix} 1 & b \\ b & b^2 \end{bmatrix}$$

$$\mathbf{Q}_v = \mathbf{0}$$

and

$$\mathbf{C} = \begin{bmatrix} 1 & 0 \end{bmatrix}$$

The algebraic Ricatti equation is

$$\mathbf{P} = \mathbf{A}\left[\mathbf{I} - \mathbf{P}\mathbf{C}^H\left(\mathbf{C}\mathbf{P}\mathbf{C}^H + \mathbf{Q}_v \right)^{-1}\mathbf{C} \right]\mathbf{P}\mathbf{A}^H + \mathbf{Q}_w$$

or

$$\mathbf{P} - \mathbf{A}\mathbf{P}\mathbf{A}^H + \mathbf{A}\mathbf{P}\mathbf{C}^H\left(\mathbf{C}\mathbf{P}\mathbf{C}^H + \mathbf{Q}_v \right)^{-1}\mathbf{C}\mathbf{P}\mathbf{A}^H - \mathbf{Q}_w = 0$$

From the Kalman filtering equations, we know that

$$\mathbf{P} = \mathbf{A}\mathbf{P}\mathbf{A}^H + \mathbf{Q}_w$$

Therefore, let us denote $\mathbf{P}$ by

$$\mathbf{P} = \begin{bmatrix} p(1) & p(2) \\ p(3) & p(4) \end{bmatrix}$$

Incorporating this into the equation above, we have

$$\begin{bmatrix} p(1) & p(2) \\ p(3) & p(4) \end{bmatrix} = \begin{bmatrix} -a & 1 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} p(1) & p(2) \\ p(3) & p(4) \end{bmatrix}\begin{bmatrix} -a & 0 \\ 1 & 0 \end{bmatrix} + \sigma_w^2 \begin{bmatrix} 1 & b \\ b & b^2 \end{bmatrix}$$

Multiplying this out we find

$$\begin{bmatrix} p(1) & p(2) \\ p(3) & p(4) \end{bmatrix} = \begin{bmatrix} a^2 p(1) - 2ap(2) + p(4) & 0 \\ 0 & 0 \end{bmatrix} + \sigma_w^2 \begin{bmatrix} 1 & b \\ b & b^2 \end{bmatrix}$$

Therefore, $\mathbf{P}$ has the form

$$\mathbf{P} = \sigma_w^2 \begin{bmatrix} 1+c & b \\ b & b^2 \end{bmatrix}$$

Now, note that

$$\mathbf{C}\mathbf{P}\mathbf{C}^H + \mathbf{Q}_v = \sigma_w^2 \begin{bmatrix} 1 & 0 \end{bmatrix}\begin{bmatrix} 1+c & b \\ b & b^2 \end{bmatrix}\begin{bmatrix} 1 \\ 0 \end{bmatrix} + 0 = \sigma_w^2(1+c)$$

Therefore,

$$\left(\mathbf{C}\mathbf{P}\mathbf{C}^H + \mathbf{Q}_v \right)^{-1} = \frac{1}{\sigma_w^2(1+c)}$$

and the algebraic Ricatti equation becomes

$$\sigma_w^2 \begin{bmatrix} 1+c & b \\ b & b^2 \end{bmatrix} - \sigma_w^2 \begin{bmatrix} -a & 1 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} 1+c & b \\ b & b^2 \end{bmatrix}\begin{bmatrix} -a & 0 \\ 1 & 0 \end{bmatrix}$$

$$+ \sigma_w^2 \frac{1}{1+c}\begin{bmatrix} -a & 1 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} 1+c & b \\ b & b^2 \end{bmatrix}\begin{bmatrix} 1 \\ 0 \end{bmatrix}\begin{bmatrix} 1 & 0 \end{bmatrix}\begin{bmatrix} 1+c & b \\ b & b^2 \end{bmatrix}\begin{bmatrix} -a & 0 \\ 1 & 0 \end{bmatrix} - \sigma_w^2 \begin{bmatrix} 1 & b \\ b & b^2 \end{bmatrix}$$

$$= \sigma_w^2 \begin{bmatrix} c & 0 \\ 0 & 0 \end{bmatrix} - \sigma_w^2 \begin{bmatrix} a^2 c + (a-b)^2 & 0 \\ 0 & 0 \end{bmatrix} + \frac{\sigma_w^2}{1+c}\begin{bmatrix} (b-a-ac)^2 & 0 \\ 0 & 0 \end{bmatrix}$$

However, in order for this to be equal to zero, as dictated by the algebraic Ricatti equation, the constant c must satisfy the equation

$$c = a^2 c + (a - b)^2 - \frac{(b - a - ac)^2}{1 + c}$$

Multiplying this out and simplifying this may be written as

$$c(1 - b^2 + c) = 0$$

Therefore, there are two possible values for c,

$$c = 0 \quad ; \quad c = b^2 - 1$$

For $c = 0$, the matrix $\mathbf{P}$ is

$$\mathbf{P} = \sigma_w^2 \begin{bmatrix} 1 & b \\ b & b^2 \end{bmatrix}$$

which is non-negative definite for all values of b. For $c = b^2 - 1$, we have

$$\mathbf{P} = \sigma_w^2 \begin{bmatrix} b^2 & b \\ b & b^2 \end{bmatrix}$$

which is non-negative definite provided $b \geq 1$.

(c) Using the following expression for the Kalman gain (see Problem 7.21)

$$\mathbf{K} = \mathbf{P}\mathbf{C}^H \left(\mathbf{C}\mathbf{P}\mathbf{C}^H + \mathbf{Q}_v \right)^{-1}$$

we have

$$\mathbf{K} = \frac{1}{1 + c} \begin{bmatrix} 1 & b \\ b & b^2 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ b/(1 + c) \end{bmatrix}$$

Thus, for $c = 0$

$$\mathbf{K} = \begin{bmatrix} 1 \\ b \end{bmatrix}$$

or for $c = b^2 - 1$,

$$\mathbf{K} = \begin{bmatrix} 1 \\ b^{-1} \end{bmatrix}$$

SOLUTIONS TO CHAPTER 8
Power Spectrum Estimation

8.1 Given $N = 10,000$ samples of a process $x(n)$, you are asked to compute the periodogram. However, with only a finite amount of memory resources, you are unable to compute a DFT any longer than 1024. Using these 10,000 samples, describe how you would be able to compute a periodogram that has a resolution of

$$\Delta\omega = 0.89\frac{2\pi}{10000}$$

Hint: Consider how the decimation-in-time FFT algorithm works.

Solution _____

To get the maximum resolution from $N = 10000$ data values, we want to compute the periodogram of $x(n)$ (segmenting $x(n)$ into subsequences reduces the resolution). The question, therefore, is how to compute the periodogram of $x(n)$ using 1024-point DFT's. Recalling how the FFT works, note that

$$X(e^{j\omega}) = \sum_{n=0}^{9999} x(n)e^{-jn\omega} = \sum_{n=0}^{9999}\sum_{l=0}^{9} x(10n + l)e^{-j(10n+l)\omega} = \sum_{l=0}^{9} e^{-jl\omega}\sum_{n=0}^{9999} x(10n + l)e^{-jn\omega}$$

Therefore, the procedure is to pad $x(n)$ to form a sequence of length $N = 10240$, and then decimate $x(n)$ into 10 sequences $x_l(n)$ of length $M = 1024$,

$$x_l(n) = x(10n + l) \quad ; \quad n = 0, 1, \ldots, 1023$$

Next, the 1024-point DFT's of these sequences, $X_l(k)$, are computed, and combined using the "twiddle factors" $\exp\left(-jl\frac{2\pi k}{10240}\right)$ as follows

$$X(k) = \sum_{l=0}^{9} e^{-jl\frac{2\pi k}{10240}} X_l(k) \quad ; \quad k = 0, 1, \ldots, 10239$$

Finally, squaring the magnitude of $X(k)$ and dividing by $N = 10240$, we have the periodogram with a resolution $\Delta\omega = 0.89(2\pi/10000)$.

_____

8.2 A continuous-time signal $x_a(t)$ is bandlimited to 5 kHz, i.e., $x_a(t)$ has a spectrum $X_a(f)$ that is zero for $|f| > 5$ kHz. Only 10 seconds of the signal has been recorded and is available for processing. We would like to estimate the power spectrum of $x_a(t)$ using the available data in a radix-2 FFT algorithm, and it is required that the estimate have a resolution of at least 10 Hz. Suppose that we use Bartlett's method of periodogram averaging.

(a) If the data is sampled at the Nyquist rate, what is the minimum section length that you may use to get the desired resolution?

(b) Using the minimum section length determined in part (a), with 10 seconds of data, how many sections are available for averaging?

(c) How does your choice of the sampling rate affect the resolution and variance of your estimate? Are there any benefits to sampling above the Nyquist rate?

Solution _____

(a) If we sample at the Nyquist rate, $f_s = 10$kHz, then a resolution of $\Delta f = 10$Hz (in analog frequency) implies that we want a resolution (in radians) of

$$\Delta\omega = 2\pi\frac{\Delta f}{f_s} = 2\pi \times 10^{-3}$$

Since the resolution of the periodogram using an L-point data record is

$$\mathrm{Res}[\widehat{P}_{PER}(\omega)] = \Delta\omega = 0.89\frac{2\pi}{L}$$

then for Bartlett's method we want to use a section length of

$$L \geq 0.89\frac{2\pi}{\Delta\omega} = 890 \quad \text{samples}$$

(b) Sampling at 10 kHz, 10 seconds of data corresponds to $N = (10)(10 \times 10^3) = \times 10^5$ samples. Therefore, with a 1024-point DFT the number of sections we may have in Bartlett's method is

$$K = [N/1024] = 98$$

(c) If the sampling rate is increased then $\Delta\omega$ decreases which, in turn, requires a longer section length for a given resolution. However, an increase in the sampling rate produces a corresponding increase in the total number of samples within a T second interval. Therefore, since the variance (normalized) is
$$V = L/N$$
increasing the sampling rate has no effect. Thus, provided that the sampling rate is not less than the Nyquist frequency, the resolution and the variance do not depend on the sampling rate.

8.3 Bartlett's method is used to estimate the power spectrum of a process from a sequence of $N = 2000$ samples.

   (a) What is the minimum length L that may be used for each sequence if we are to have a resolution of $\Delta f = 0.005$?

   (b) Explain why it would not be advantageous to increase L beyond the value found in (a).

   (c) The *quality factor* of a spectrum estimate is defined to be the inverse of the variability,

$$Q = 1/\mathcal{V}$$

   Using Bartlett's method, what is the minimum number of data samples, N, that are necessary to achieve a resolution of $\Delta f = 0.005$, and a quality factor that is five times larger than that of the periodogram?

Solution ──

   (a) Since $\Delta f = 0.9/L$ then

$$L = \frac{0.89}{\Delta f} = \frac{0.9}{0.005} = 180$$

   (b) Increasing L will increase the resolution, but it will also result in a decrease in the number of segments that may be averaged. This, in turn, will increase the variance of the spectrum estimate.

   (c) For the periodogram, the quality factor is $Q_{per} = 1/\mathcal{V}_{per} = 1$. The quality factor for Bartlett's method is $Q_B = 1/\mathcal{V}_B = K$. Therefore, if we want $Q_{per}/Q_B \geq 5$, then we must have $K \geq 5$. With $M = 180$ (for $\Delta f = 0.005$), then we must have

$$N = KM \geq 5 \times 180 = 900$$

8.4 A random process $x(n)$ is generated by filtering unit variance white noise as shown in the figure below
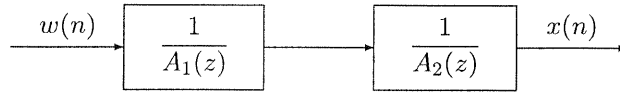
where

$$A_1(z) = 1 + az^{-1} + 0.99z^{-2} \quad ; \quad A_2(z) = 1 - az^{-1} + 0.98z^{-2}$$

(a) Prepare a carefully labeled sketch of the power spectrum of $x(n)$ assuming that a is small, e.g., $0 < a < 0.1$. Pay careful attention to the location and amplitude of the two spectral peaks and the value of $P_x(e^{j\omega})$ at $\omega = \pi/2$.

(b) If $a = 0.1$, determine the section length L required to resolve the spectral peaks of $P_x(e^{j\omega})$ using Bartlett's method. For this value of L, find an approximate value for the bias of the estimate at the peaks of the spectrum. How is the bias related to the area of the spectral peaks?

(c) Consider the method of periodogram smoothing. How many lags of the autocorrelation must be used to obtain a resolution that is comparable to that of Bartlett's estimate considered in part (b)? How much data must be available if the variance of the estimate is to be comparable to that of a four-section Bartlett estimate?

Solution _____

(a) The system function of the cascade is

$$H(z) = \frac{1}{1 + az^{-1} + 0.99z^{-2}} \cdot \frac{1}{1 - az^{-1} + 0.98z^{-2}}$$

and, since the input to this filter is white noise, then the power spectrum of the output process, $x(n)$, is

$$P_x(z) = \frac{1}{1 + az^{-1} + 0.99z^{-2}} \cdot \frac{1}{1 - az^{-1} + 0.98z^{-2}} \cdot \frac{1}{1 + az + 0.99z^2} \cdot \frac{1}{1 - az + 0.98z^2}$$

Thus, $P_x(z)$ has a total of 8 poles, 4 inside the unit circle and 4 outside. Since each pole is close to the unit circle, then the peaks of the power spectrum $P_x(e^{j\omega})$ are at frequencies that correspond, approximately, to the angles of the poles, which are

$$\omega_1 = \cos^{-1} \frac{a}{2\sqrt{0.98}} \quad ; \quad \omega_2 = \cos^{-1} \frac{-a}{2\sqrt{0.99}}$$

For example, if $a = -0.1$, then

$$\omega_1 = 0.51600\pi \quad ; \quad \omega_2 = 0.48392\pi$$

Therefore,

$$\Delta\omega = \omega_2 - \omega_1 \approx 0.03208\pi$$

or
$$\Delta f = \Delta\omega/2\pi = 0.01604$$

Note that the midpoint between the two spectral peaks falls at approximately $\omega_0 = \pi/2$. Since

$$P_x(e^{j\omega}) = \frac{1}{1 + a^2 + (0.98)^2 - 3.96a\cos\omega + 1.96\cos 2\omega} \cdot \frac{1}{1 + a^2 + (0.99)^2 - 3.98a\cos\omega + 1.98\cos 2\omega}$$

The power spectrum at ω_1, ω_2, and ω_0 is

$$P_x(e^{j\omega_1}) = \frac{1}{(4.0 \cdot 10^{-4} - 1.0 \cdot 10^{-4}a^2)(1.0 \cdot 10^{-4} + 4.0204a^2)} \approx \frac{2.5 \cdot 10^3}{1.0 \cdot 10^{-4} + 4.0204a^2}$$

$$P_x(e^{j\omega_2}) = \frac{1}{(4.0 \cdot 10^{-4} + 3.97987a^2)(1.0 \cdot 10^{-4} - 3.0 \cdot 10^{-5})} \approx \frac{10^4}{4.0 \cdot 10^{-4} + 3.97987a^2}$$

$$P_x(e^{j\omega_0}) = \frac{1}{(4.0 \cdot 10^{-4} + a^2)(1.0 \cdot 10^{-4} + a^2)}$$

If $a = 0.1$ then these values become

$$P_x(e^{j\omega_1}) \approx 6.203 \cdot 10^4$$
$$P_x(e^{j\omega_2}) \approx 2.4876 \cdot 10^5$$
$$P_x(e^{j\omega_0}) \approx 9.5202 \cdot 10^3$$

Thus, the peak in $P_x(e^{j\omega})$ at ω_1 is down 6dB from the peak at ω_2, and the "valley" at ω_0 is about 14dB down from the peak at ω_2.

(b) For resolving two peaks that are separated by Δf, we require a section length

$$L \geq \frac{0.89}{\Delta f}$$

With $\Delta f = 0.01604$, we must have

$$L \geq 56$$

Now,

$$E\{\widehat{P}_B(e^{j\omega})\} = E\left[\frac{1}{K}\sum_{k=1}^{K}\widehat{P}_{\text{per}}^{(k)}(e^{j\omega})\right] = E\left[\widehat{P}_{\text{per}}^{(k)}(e^{j\omega})\right]$$

Thus,

$$E\{\widehat{P}_B(e^{j\omega})\} = \frac{1}{2\pi}\int_{-\pi}^{\pi} P_x(e^{j\theta})W_B(e^{j(\omega-\theta)})d\theta$$

where

$$W_B(e^{j\omega}) = \frac{1}{L}\left[\frac{\sin(\omega L/2)}{\sin(\omega/2)}\right]^2$$

Since L has been selected so that the two peaks can be resolved, let us assume that $W_B(e^{j\omega})$ is non-zero only over the interval $-\Delta\omega/2 \leq \omega \leq \Delta\omega/2$. Furthermore, since the width of the main lobe of the window $W_B(e^{j\omega})$ is much wider than the width of the spectral peaks, if we assume that $W_B(e^{j\omega}) \approx L$ over the interval $-\Delta\omega/2 \leq \omega \leq \Delta\omega/2$, then we can form the following approximation

$$E\{\widehat{P}_B(e^{j\omega})\} \approx \frac{L}{2\pi}\int_{-\Delta\omega/2}^{\Delta\omega/2} P_x(e^{j(\omega-\theta)})d\theta$$

Therefore, $E\{\widehat{P}_B(e^{j\omega})\}$ at $\omega = \omega_1$ and $\omega = \omega_2$ is proportional to the *area* under the spectral peaks.

(c) For periodogram smoothing, with a Bartlett window, the resolution is

$$\Delta\omega = 0.64\frac{2\pi}{M}$$

The resolution of Bartlett's method with a section length $L = 56$ is

$$\Delta\omega = 0.89\frac{2\pi}{56}$$

For periodogram smoothing to have the same resolution, we require

$$M = 0.64 \cdot \frac{2\pi}{\Delta\omega} = 0.64 \cdot \frac{56}{0.89} = 40$$

The normalized variance of the Bartlett estimate is

$$\mathcal{V}_B = \frac{1}{K}$$

and for periodogram smoothing it is

$$\mathcal{V}_{BT} = \frac{2M}{3N}$$

In order for periodogram smoothing to have the same variability as Bartlett's method with $K = 4$ sections, we require

$$\frac{2M}{3N} = \frac{1}{4}$$

With $M = 40$, so that the resolutions are the same, this requires a data record length of

$$N = \frac{8}{3}M = 107$$

8.5 Many commercial *Fourier analyzers* continuously update the estimate of the power spectrum of a process $x(n)$ by exponential averaging periodograms as follows,

$$\widehat{P}_i(e^{j\omega}) = \alpha \widehat{P}_{i-1}(e^{j\omega}) + \frac{1-\alpha}{N} \left| \sum_{n=0}^{N-1} x_i(n) e^{-jn\omega} \right|^2$$

where $x_i(n) = x(n + Ni)$ is the ith sequence of N data values. This update equation is initialized with $\widehat{P}_{-1}(e^{j\omega}) = 0$.

(a) Qualitatively describe the philosophy behind this method, and discuss how the value for the weighting factor α should be selected.

(b) Assuming that successive periodograms are uncorrelated and that $0 < \alpha < 1$, find the mean and variance of $\widehat{P}_i(e^{j\omega})$ for a Gaussian random process.

(c) Repeat the analysis in part (b) if the periodograms are replaced with modified periodograms.

Solution _____

(a) As data is being read by a spectrum analyzer, the goal is to continuously update the estimate. As each data record of length N is collected, the periodogram is computed, and *averaged* with the previous spectrum estimate. Although a running average could be formed, this would assume that the process is stationary. Selecting a value of $0 < \alpha < 1$ allows the estimate to *forget* $\widehat{P}_i(e^{j\omega})$ as more data is collected. In the extreme case in which $\alpha = 0$, $\widehat{P}_i(e^{j\omega})$ is the periodogram of the most recent N data values. As we will see in part (b), $\widehat{P}_i(e^{j\omega})$ is an exponentially weighted average of the previous periodograms.

(b) If we define

$$Q_i(e^{j\omega}) = \frac{1}{N} \left| \sum_{n=0}^{N-1} x_i(n) e^{-jn\omega} \right|^2$$

then the expression for the ith spectrum estimate, $P_i(e^{j\omega})$ is

$$\widehat{P}_i(e^{j\omega}) = \alpha \widehat{P}_{i-1}(e^{j\omega}) + \frac{1-\alpha}{N} Q_i(e^{j\omega})$$

which is a difference equation for $\widehat{P}_i(e^{j\omega})$. Since the initial conditions are zero, $\widehat{P}_{-1}(e^{j\omega}) = 0$, then the solution for $\widehat{P}_i(e^{j\omega})$ is

$$\widehat{P}_i(e^{j\omega}) = \sum_{k=0}^{i} (1-\alpha)\alpha^k Q_k(e^{j\omega})$$

Taking the expected value we have

$$E\{\widehat{P}_i(e^{j\omega})\} = \sum_{k=0}^{i} (1-\alpha)\alpha^k E\{Q_k(e^{j\omega})\}$$

Since $Q_k(e^{j\omega})$ is the periodogram of $x_k(n)$, then

$$E\{Q_k(e^{j\omega})\} = \frac{1}{2\pi} P_x(e^{j\omega}) * W_B(e^{j\omega})$$

and

$$
\begin{aligned}
E\{\widehat{P}_i(e^{j\omega})\} &= \frac{1}{2\pi} \sum_{k=0}^{i} (1-\alpha)\alpha^k \big[P_x(e^{j\omega}) * W_B(e^{j\omega}) \big] \\
&= \frac{1}{2\pi} \big[P_x(e^{j\omega}) * W_B(e^{j\omega}) \big] (1-\alpha) \sum_{k=0}^{p} \alpha^k \\
&= \frac{1}{2\pi} \big[P_x(e^{j\omega}) * W_B(e^{j\omega}) \big] (1-\alpha) \frac{1-\alpha^{i+1}}{1-\alpha} \\
&= (1-\alpha^{i+1}) \frac{1}{2\pi} \big[P_x(e^{j\omega}) * W_B(e^{j\omega}) \big]
\end{aligned}
$$

For the variance, we proceed in the same way, using the fact that the variance of the periodogram is

$$
\mathrm{var}\{\widehat{P}_{\mathrm{per}}(e^{j\omega})\} \approx P_x^2(e^{j\omega})
$$

Therefore, we have

$$
\begin{aligned}
E\{\widehat{P}_i(e^{j\omega})\} &= \sum_{k=0}^{i} (1-\alpha)^2 \alpha^{2k} P_x^2(e^{j\omega}) = (1-\alpha)^2 P_x(e^{j\omega}) \sum_{k=0}^{i} \alpha^{2k} \\
&= (1-\alpha)^2 \frac{1-\alpha^{2(i+1)}}{1-\alpha^2} P_x(e^{j\omega}) = \frac{1-\alpha}{1+\alpha}(1-\alpha^{2(i+1)}) P_x(e^{j\omega})
\end{aligned}
$$

(c) For modified periodograms, the only change that is necessary is to use

$$
E\{Q_i(e^{j\omega})\} = \frac{1}{NU} \big| P_x(e^{j\omega}) * W_B(e^{j\omega}) \big|^2
$$

where

$$
U = \frac{1}{N} \sum_{n=0}^{N-1} |w(n)|^2
$$

Substituting this into the expression in part (b), we have for the expected value,

$$
E\{\widehat{P}_i(e^{j\omega})\} = (1-\alpha^{i+1}) \frac{1}{2\pi NU} P_x(e^{j\omega}) * |W_B(e^{j\omega})|^2
$$

and the variance is the same.

8.6 The minimum variance method of spectrum estimation constrains the FIR bandpass filter $G_i(e^{j\omega})$ to have a gain of one at frequency $\omega = \omega_i$,

$$G_i(e^{j\omega_i}) = 1$$

Another approach is to constrain the filter to have unit energy over a frequency band that is centered at $\omega = \omega_i$ and has a bandwidth of Δ,

$$\frac{1}{\Delta} \int_{\omega_i - \Delta/2}^{\omega_i + \Delta/2} \left| G_i(e^{j\omega}) \right|^2 d\omega = 1$$

With this constraint, the filter coefficients $\mathbf{g}_i = [g_i(0), g_i(1), \ldots, g_i(p)]^T$ that minimize the power in the filtered process,

$$E\{|y_i(n)|^2\} = \mathbf{g}_i^H \mathbf{R}_x \mathbf{g}_i$$

may be shown to be the solution to a generalized eigenvalue problem,

$$\mathbf{R}_x \mathbf{g}_i = \lambda(\omega_i, \Delta) \mathbf{T}_i \mathbf{g}_i$$

where $\mathbf{T}_i$ is a matrix whose elements depend upon ω_i and Δ. The spectrum estimate, referred to as the DASE estimate, is

$$\widehat{P}_{DASE}(e^{j\omega}) = \lambda_{\min}(\omega_i, \Delta)$$

where $\lambda_{\min}(\omega_i, \Delta)$ is the minimum eigenvalue of the generalized eigenvalue problem.

(a) Perform the minimization of $E\{|y_i(n)|^2\}$ and determine the form of the matrix $\mathbf{T}_i$.

(b) What happens to the matrix $\mathbf{T}_i$ in the limit as $\Delta \to 0$? What does the power spectrum estimate correspond to in this case?

(c) Repeat part (b) for $\Delta = 2\pi$.

(d) Find the DASE estimate for white noise.

Solution _____

(a) In this problem, we want to minimize

$$E\{|y_i(n)|^2\} = \mathbf{g}_i^H \mathbf{R}_x \mathbf{g}_i$$

subject to the constraint

$$\frac{1}{\Delta} \int_{\omega_i - \Delta/2}^{\omega_i + \Delta/2} |G_i(e^{j\omega})|^2 d\omega = 1$$

where

$$G_i(e^{j\omega}) = \sum_{n=0}^{p} g_i(n) e^{-jn\omega}$$

This constraint may be written as follows

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} G_i(e^{j\omega}) H_i(e^{j\omega}) G_i^*(e^{j\omega}) d\omega = 1$$

where

$$H_i(e^{j\omega}) = \begin{cases} 2\pi/\Delta & ; & |\omega - \omega_i| < \Delta/2 \\ 0 & ; & \text{else} \end{cases}$$

With

$$
\begin{aligned}
h_i(n) &= \frac{1}{2\pi} \int_{-\pi}^{\pi} H_i(e^{j\omega}) e^{jn\omega} d\omega = \frac{1}{\Delta} \int_{\omega_i - \Delta/2}^{\omega_i + \Delta/2} e^{jn\omega} d\omega \\
&= \frac{1}{jn\Delta} \left[e^{jn(\omega_i + \Delta/2)} - e^{jn(\omega_i - \Delta/2)} \right] \\
&= \frac{1}{jn\Delta} e^{jn\omega_i} \cdot 2j \sin\frac{n\Delta}{2} = e^{jn\omega_i} \text{sinc}\left(\frac{n\Delta}{2}\right)
\end{aligned}
$$

where

$$\text{sinc}(x) \equiv \frac{\sin x}{x}$$

we have

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} G_i(e^{j\omega}) H_i(e^{j\omega}) G_i^*(e^{j\omega}) d\omega = \sum_{n=0}^{p} \left[\sum_{k=0}^{p} g_i(k) h_i(n-k) \right] g_i^*(n) = \mathbf{g}_i^H \mathbf{T}_i \mathbf{g}_i$$

where $\mathbf{T}_i$ is a Toeplitz matrix with

$$\left\{ \mathbf{T}_i \right\}_{k,l} = e^{j(k-l)\omega_i} \text{sinc}\left[(k-l)\frac{\Delta}{2} \right]$$

Thus, our problem is

$$\min_{\mathbf{g}_i} \mathbf{g}_i^H \mathbf{R}_x \mathbf{g}_i \quad \text{subject to} \quad \mathbf{g}_i^H \mathbf{T}_i \mathbf{g}_i = 1$$

Introducing a Lagrange multiplier, λ, this is equivalent to minimizing the function

$$Q_R(\mathbf{g}_i, \lambda) = \mathbf{g}_i^H \mathbf{R}_x \mathbf{g}_i + \lambda (1 - \mathbf{g}_i^H \mathbf{T}_i \mathbf{g}_i)$$

Differentiating with respect to $\mathbf{g}_i^*$, and settting the result equal to zero, we have

$$\frac{\partial Q_R(\mathbf{g}_i, \lambda)}{\partial \mathbf{g}_i^*} = \mathbf{R}_x \mathbf{g}_i - \lambda \mathbf{T}_i \mathbf{g}_i = 0$$

or

$$\mathbf{R}_x \mathbf{g}_i = \lambda \mathbf{T}_i \mathbf{g}_i$$

which is a *generalized eigenvalue problem*. Differentiating with respect to λ gives

$$1 - \mathbf{g}_i^H \mathbf{T}_i \mathbf{g}_i = 0$$

or

$$\mathbf{g}_i^H \mathbf{T}_i \mathbf{g}_i = 1$$

Recall that we want to minimize $\mathbf{g}_i^H \mathbf{R}_x \mathbf{g}_i$. Since

$$\mathbf{g}_i^H \mathbf{R}_x \mathbf{g}_i = \lambda \, \mathbf{g}_i^H \mathbf{T}_i \mathbf{g}_i = \lambda$$

then minimizing $\mathbf{g}_i^H \mathbf{R}_x \mathbf{g}_i$ is equivalent to minimizing λ. Since $\mathbf{R}_x$ is nonsingular, the generalized eigenvalue problem may be written as follows

$$\mathbf{R}_x^{-1} \mathbf{T}_i \mathbf{g}_i = \frac{1}{\lambda} \mathbf{g}_i$$

Therefore, $\mathbf{g}_i$ corresponds to the eigenvector of $\mathbf{R}_x^{-1} \mathbf{T}_i$ that has the *maximum* eigenvalue, and $\lambda_{\max}^{-1}$ is the estimate of the power.

(b) As $\Delta \to 0$, the matrix $\mathbf{T}_i$ approaches the limit

$$\lim_{\Delta \to 0} \left\{ \mathbf{T}_i \right\}_{k,l} = \lim_{\Delta \to 0} e^{j(k-l)\omega_i} \operatorname{sinc}\left[(k-l)\frac{\Delta}{2} \right] = e^{j(k-l)\omega_i}$$

Therefore,

$$\mathbf{T}_i = \mathbf{e}_i \mathbf{e}_i^H$$

where

$$\mathbf{e}_i = \left[1, e^{j\omega_i}, \dots, e^{jn\omega_i} \right]^T$$

and the generalized eigenvalue problem becomes

$$\mathbf{R}_x \mathbf{g}_i = \lambda \mathbf{T}_i \mathbf{g}_i = \lambda \mathbf{e}_i \mathbf{e}_i^H \mathbf{g}_i$$

or,

$$\mathbf{R}_x^{-1} \mathbf{e}_i \mathbf{e}_i^H \mathbf{g}_i = \frac{1}{\lambda} \mathbf{g}_i$$

Multiplying both sides on the left by $\mathbf{e}_i^H$ gives

$$\left(\mathbf{e}_i^H \mathbf{R}_x^{-1} \mathbf{e}_i \right) \mathbf{e}_i^H \mathbf{g}_i = \frac{1}{\lambda} \mathbf{e}_i^H \mathbf{g}_i$$

Since $\mathbf{e}_i^H \mathbf{g}_i$ is a scalar, we may divide both sides by $\mathbf{e}_i^H \mathbf{g}_i$, which gives

$$\lambda = \frac{1}{\mathbf{e}_i^H \mathbf{R}_x^{-1} \mathbf{e}_i}$$

which is equivalent to the minimum variance power estimate.

(c) When $\Delta = 2\pi$,

$$\mathbf{T}_i = \mathbf{I}$$

and the generalized eigenvalue problem becomes

$$\mathbf{R}_x \mathbf{g}_i = \lambda \mathbf{g}_i$$

In this case, λ is the minimum eigenvalue of $\mathbf{R}_x$ as in the Pisarenko harmonic decomposition, and the power estimate corresponds to the white noise power under the assumption that the process consists of a sum of complex exponentials in white noise.

(d) In the case of white noise, $\mathbf{R}_x = \sigma_x^2 \mathbf{I}$, and the generalized eigenvalue equation becomes

$$\sigma_x^2 \mathbf{g}_i = \lambda \mathbf{T}_i \mathbf{g}_i$$

or

$$\sigma_x^{-2} \mathbf{T}_i \mathbf{g}_i = \frac{1}{\lambda} \mathbf{g}_i$$

Therefore, the power estimate at frequency ω_i is equal to the maximum eigenvalue of the matrix $\sigma_x^{-2} \mathbf{T}_i$. This value, however, depends on ω_i and Δ.

8.7 Let $x(n)$ be a random process consisting of a single complex exponential in white noise,

$$r_x(k) = P e^{jk\omega_0} + \sigma_w^2 \delta(k)$$

and let $\mathbf{g}_i$ be the minimum variance bandpass filter

$$\mathbf{g}_i = \frac{\mathbf{R}_x^{-1} \mathbf{e}_i}{\mathbf{e}_i^H \mathbf{R}_x^{-1} \mathbf{e}_i}$$

that has a center frequency ω_i with $G(e^{j\omega_i}) = 1$. Assuming that $\omega_i \neq \omega_0$, prove that $G_i(z)$ has a zero that approaches $z = e^{j\omega_0}$ as $\sigma_w^2/P \longrightarrow 0$.

Solution _____

For the given process, the $N \times N$ autocorrelation matrix is

$$\mathbf{R}_x = P\mathbf{e}_0\mathbf{e}_0^H + \sigma_w^2 \mathbf{I} = \sigma_w^2 \left[\mathbf{I} + \frac{P}{\sigma_w^2} \mathbf{e}_0\mathbf{e}_0^H \right]$$

Using Woodbury's identity, the matrix inverse is

$$\mathbf{R}_x^{-1} = \frac{1}{\sigma_w^2} \left[\mathbf{I} - \frac{\frac{P}{\sigma_w^2}\mathbf{e}_0\mathbf{e}_0^H}{1 + \frac{NP}{\sigma_w^2}} \right] = \frac{1}{\sigma_w^2} \left[\mathbf{I} - \frac{P\mathbf{e}_0\mathbf{e}_0^H}{\sigma_w^2 + NP} \right]$$

The minimum variance filter is

$$\mathbf{g}_i = \lambda \mathbf{R}_x^{-1}\mathbf{e}_i = \frac{\lambda}{\sigma_w^2} \left[\mathbf{e}_i - \frac{P\mathbf{e}_0\mathbf{e}_0^H}{\sigma_w^2 + NP}\mathbf{e}_i \right]$$

and the frequency response is

$$G_i(e^{j\omega}) = \mathbf{e}^H \mathbf{g}_i = \frac{\lambda}{\sigma_w^2} \left[\mathbf{e}^H \mathbf{e}_i - \frac{P}{\sigma_w^2 + NP}\mathbf{e}^H \mathbf{e}_0\mathbf{e}_0^H \mathbf{e}_i \right]$$

where

$$\lambda = \frac{1}{\mathbf{e}_i^H \mathbf{R}_x^{-1} \mathbf{e}_i}$$

is simply a scalar. Evaluating $G_i(e^{j\omega})$ at $\omega = \omega_0$, we set $\mathbf{e} = \mathbf{e}_0$ as follows

$$G_i(e^{j\omega})|_{\omega=\omega_0} = \mathbf{e}_0^H \mathbf{g}_i = \frac{\lambda}{\sigma_w^2} \left[\mathbf{e}_0^H \mathbf{e}_i - \frac{P}{\sigma_w^2 + NP}\mathbf{e}_0^H \mathbf{e}_0\mathbf{e}_0^H \mathbf{e}_i \right] = \frac{\lambda}{\sigma_w^2} \left[1 - \frac{NP}{\sigma_w^2 + NP} \right] \mathbf{e}_0^H \mathbf{e}_i$$

Thus,

$$G_i(e^{j\omega})|_{\omega=\omega_0} = \frac{\lambda}{\sigma_w^2} \left[1 - \frac{N}{\frac{\sigma_w^2}{P} + N} \right] \mathbf{e}_0^H \mathbf{e}_i$$

and it follows that

$$\lim_{\sigma_w^2/P \to 0} G_i(e^{j\omega})|_{\omega=\omega_0} = 0$$

i.e., as $\sigma_w^2/P \longrightarrow 0$, a zero in $G(z)$ approaches the unit circle at the point $z = e^{j\omega_0}$.

_____

8.8 A random process is known to consist of a single sinusoid in white noise,

$$x(n) = A\cos(n\omega_0 + \phi) + w(n)$$

Thus, the autocorrelation sequence for $x(n)$ is

$$r_x(k) = \frac{1}{2}A^2\cos(k\omega_0) + \sigma_w^2\delta(k)$$

(a) If $\omega_0 = \pi/4$, $A = \sqrt{2}$, and $\sigma_w^2 = 1$, find the second-order MEM spectrum, $\widehat{P}_{mem}(e^{j\omega})$.

(b) Determine the location of the poles of $\widehat{P}_{mem}(z)$.

(c) Does the peak of $\widehat{P}_{mem}(e^{j\omega})$ provide an accurate estimate of ω_0? How does this estimate of ω_0 compare to that obtained using the Pisarenko Harmonic decomposition?

Solution _____

(a) The second-order MEM spectrum is found by solving the normal equations

$$\begin{bmatrix} r_x(0) & r_x(1) \\ r_x(1) & r_x(0) \end{bmatrix}\begin{bmatrix} a(1) \\ a(2) \end{bmatrix} = -\begin{bmatrix} r_x(1) \\ r_x(2) \end{bmatrix}$$

The first three values of the autocorrelation sequence are

$$\begin{aligned} r_x(0) &= \tfrac{1}{2}A^2 + \sigma_w^2 = 2 \\ r_x(1) &= \tfrac{1}{2}A^2\cos(\tfrac{\pi}{4}) = \tfrac{1}{2}\sqrt{2} \\ r_x(2) &= \tfrac{1}{2}A^2\cos(\tfrac{\pi}{2}) = 0 \end{aligned}$$

Thus, the normal equations are

$$\begin{bmatrix} 2 & \tfrac{1}{2}\sqrt{2} \\ \tfrac{1}{2}\sqrt{2} & 2 \end{bmatrix}\begin{bmatrix} a(1) \\ a(2) \end{bmatrix} = -\begin{bmatrix} \tfrac{1}{2}\sqrt{2} \\ 0 \end{bmatrix}$$

Solving for $a(1)$ and $a(2)$ we find

$$\begin{bmatrix} a(1) \\ a(2) \end{bmatrix} = \frac{1}{7}\begin{bmatrix} -2\sqrt{2} \\ 1 \end{bmatrix}$$

Therefore, the MEM spectrum is

$$\widehat{P}_x(e^{j\omega}) = \frac{\epsilon_2}{1 - \tfrac{2}{7}\sqrt{2}z^{-1} + \tfrac{1}{7}z^{-2}}$$

where

$$\epsilon_2 = r_x(0) + a(1)r_x(1) + a(2)r_x(2) = \tfrac{12}{7}$$

(b) The roots of the denominator polynomial are at

$$z = \frac{\sqrt{2}}{7} \pm j\frac{\sqrt{5}}{7}$$

which, in polar form, are

$$z = 0.378e^{\pm j0.3205\pi}$$

(c) The angles of the poles correspond, approximately, to the locations of the peaks in the maximum entropy spectrum. Since the poles are at an angle of $\pm 0.3205\pi$, and the sinusoid has a frequency of $\omega_0 = 0.25\pi$, then the MEM spectrum does not produce a very accurate estimate of the sinusoid frequency. If we were to use the Pisarenko harmonic decomposition, on the other hand, the frequency would have been determined exactly.

8.9 Suppose that we have determined the following values for the autocorrelation sequence of a real-valued random process $x(n)$:

$$r_x(0) = 1 \quad ; \quad r_x(1) = \alpha \quad ; \quad r_x(2) = 0$$

(a) Using the Blackman-Tukey method with a rectangular window, find and make a carefully labeled sketch of the estimated power spectrum, $\widehat{P}_{BT}(e^{j\omega})$.

(b) Repeat part (a) for a second-order MEM spectrum estimate, $\widehat{P}_{mem}(e^{j\omega})$.

(c) Repeat part (a) for a MV spectrum estimate, $\widehat{P}_{MV}(e^{j\omega})$.

(d) What can you say about the autocorrelation sequences that correspond to the spectrum estimates $\widehat{P}_{BT}(e^{j\omega})$, $\widehat{P}_{mem}(e^{j\omega})$, and $\widehat{P}_{MV}(e^{j\omega})$ found in parts (a)-(c)?

Solution _____

(a) For the Blackman-Tukey estimate with a rectangular window, the spectrum estimate is

$$\widehat{P}_{BT}(e^{j\omega}) = 1 + 2\alpha \cos \omega$$

(b) For the second-order MEM spectrum, the coefficients $a(1)$ and $a(2)$ are solutions to the normal equations

$$\begin{bmatrix} 1 & \alpha \\ \alpha & 1 \end{bmatrix} \begin{bmatrix} a(1) \\ a(2) \end{bmatrix} = - \begin{bmatrix} \alpha \\ 0 \end{bmatrix}$$

Thus,

$$\begin{bmatrix} a(1) \\ a(2) \end{bmatrix} = -\frac{1}{1-\alpha^2} \begin{bmatrix} 1 & -\alpha \\ -\alpha & 1 \end{bmatrix} \begin{bmatrix} \alpha \\ 0 \end{bmatrix} = -\frac{\alpha}{1-\alpha^2} \begin{bmatrix} 1 \\ -\alpha \end{bmatrix}$$

With

$$\epsilon_2 = r_x(0) + a(1)r_x(1) + a(2)r_x(2) = \frac{1-2\alpha^2}{1-\alpha^2}$$

then

$$\begin{aligned} \widehat{P}_{mem}(e^{j\omega}) &= \frac{\epsilon_2}{|1 + a(1)e^{-j\omega} + a(2)e^{-2j\omega}|^2} \\ &= \frac{(1-2\alpha^2)(1-\alpha^2)}{1-\alpha^2 + 2\alpha^4 - 2\alpha(1-\alpha^2)^2 \cos \omega + 2\alpha^2(1-\alpha^2) \cos 2\omega} \end{aligned}$$

(c) The second-order minimum variance estimate is

$$\widehat{P}_{MV}(e^{j\omega}) = \frac{p+1}{\mathbf{e}^H \mathbf{R}_x^{-1} \mathbf{e}}$$

Therefore, with $p = 3$,

$$\mathbf{R}_x = \begin{bmatrix} 1 & \alpha & 0 \\ \alpha & 1 & \alpha \\ 0 & \alpha & 1 \end{bmatrix}$$

and

$$\mathbf{R}_x^{-1} = \frac{1}{1-2\alpha^2} \begin{bmatrix} 1-\alpha^2 & -\alpha & \alpha^2 \\ -\alpha & 1 & -\alpha \\ \alpha^2 & -\alpha & 1-\alpha^2 \end{bmatrix}$$

it follows that the minimum variance estimate is

$$\widehat{P}_{MV}(e^{j\omega}) = \frac{3(1-2\alpha^2)}{(3-2\alpha^2) - 4\alpha\cos\omega + 2\alpha^2\cos 2\omega}$$

(d) For the Blackman-Tukey estimate,

$$\mathcal{F}^{-1}\left[\widehat{P}_{BT}(e^{j\omega})\right] = \widehat{r}_x(k) = \begin{cases} 1 & ; & k = 0 \\ \alpha & ; & k = \pm 1 \\ 0 & ; & \text{else} \end{cases}$$

and for the MEM estimate,

$$\mathcal{F}^{-1}\left[\widehat{P}_{MEM}(e^{j\omega})\right] = \widehat{r}_x(k) = \begin{cases} r_x(k) & ; & |k| \leq 2 \\ -\sum_{l=1}^{2} a(l)\widehat{r}_x(k-l) & ; & k > 2 \end{cases}$$

For the minimum variance estimate,

$$\mathcal{F}^{-1}\left[\widehat{P}_{BT}(e^{j\omega})\right] = \widehat{r}_x(k)$$

which, as with the MEM spectrum estimate, is infinite in length. However, unlike either the Blackman-Tukey or the MEM estimates, the autocorrelation matching property does not hold, i.e., $\widehat{r}_x(k) \neq r_x(k)$ for $|k| \leq 2$.

8.10 Given that the sixth-order minimum variance spectrum estimate of a process $x(n)$ is

$$\widehat{P}_{MV}(e^{j\omega}) = \frac{1}{1 + a\cos 4\omega + 4a\cos 6\omega}$$

and the seventh-order estimate is

$$\widehat{P}_{MV}(e^{j\omega}) = \frac{1}{1 - 2a\cos 2\omega - a\cos 7\omega}$$

find the seventh-order maximum entropy spectrum, $\widehat{P}_{mem}(e^{j\omega})$.

Solution _____

This problem may be easily solved using the following relationship between the MEM and MV spectrum estimates,

$$\frac{1}{\widehat{P}_{MV}^{(p)}(e^{j\omega})} = \frac{1}{p+1}\sum_{k=0}^{p}\frac{1}{\widehat{P}_{mem}^{(k)}(e^{j\omega})}$$

Specifically, note that

$$\frac{8}{\widehat{P}_{MV}^{(7)}(e^{j\omega})} = \sum_{k=0}^{7}\frac{1}{\widehat{P}_{mem}^{(k)}(e^{j\omega})}$$

and

$$\frac{7}{\widehat{P}_{MV}^{(6)}(e^{j\omega})} = \sum_{k=0}^{6}\frac{1}{\widehat{P}_{mem}^{(k)}(e^{j\omega})}$$

Taking the difference yields

$$\frac{1}{\widehat{P}_{mem}^{(7)}(e^{j\omega})} = \frac{8}{\widehat{P}_{MV}^{(7)}(e^{j\omega})} - \frac{7}{\widehat{P}_{MV}^{(7)}(e^{j\omega})}$$

Thus,

$$\frac{1}{\widehat{P}_{mem}^{(7)}(e^{j\omega})} = 8\big[1 - 2a\cos 2\omega - a\cos 7\omega\big] - 7\big[1 + a\cos 4\omega - 4a\cos 6\omega\big]$$

or,

$$\frac{1}{\widehat{P}_{mem}^{(7)}(e^{j\omega})} = 1 - 16a\cos 2\omega - 7a\cos 4\omega - 8a\cos 7\omega - 28a\cos 6\omega$$

8.11 The first-order ($p = 1$) minimum variance spectrum estimate of a random process is

$$\widehat{P}_{MV}(e^{j\omega}) = \frac{8}{3 - \cos \omega}$$

(a) Find the autocorrelations, $r_x(0)$ and $r_x(1)$, that produced this spectrum estimate.

(b) In general, given the pth-order minimum variance estimate $\widehat{P}_{MV}(e^{j\omega})$, is it possible to recover the values of the autocorrelation sequence that produce this estimate?

Solution

(a) The first-order MV spectrum estimate is

$$\widehat{P}_{MV}(e^{j\omega}) = \frac{2}{\mathbf{e}^H \mathbf{R}_x^{-1} \mathbf{e}} = \frac{2}{q(0) + 2q(1) \cos \omega}$$

where

$$\mathbf{R}_x = \begin{bmatrix} r_x(0) & r_x(1) \\ r_x(1) & r_x(0) \end{bmatrix}$$

and $q(k)$ is the sum along the kth diagonal of $\mathbf{R}_x^{-1}$. Since

$$\mathbf{R}_x^{-1} = \frac{1}{r_x^2(0) - r_x^2(1)} \begin{bmatrix} r_x(0) & -r_x(1) \\ -r_x(1) & r_x(0) \end{bmatrix}$$

then

$$q(0) = \frac{2r_x(0)}{r_x^2(0) - r_x^2(1)} \quad ; \quad q(1) = \frac{-r_x(1)}{r_x^2(0) - r_x^2(1)}$$

and

$$\widehat{P}_{MV}(e^{j\omega}) = \frac{2\left[r_x^2(0) - r_x^2(1)\right]}{2r_x(0) - 2r_x(1) \cos \omega}$$

If we express $\widehat{P}_{MV}(e^{j\omega})$ as follows

$$\widehat{P}_{MV}(e^{j\omega}) = \frac{1}{r_x(1)} \frac{\left[r_x^2(0) - r_x^2(1)\right]}{\dfrac{r_x(0)}{r_x(1)} - \cos \omega}$$

then, comparing this to the given minimum variance estimate, we have

$$\frac{r_x(0)}{r_x(1)} = 3 \quad ; \quad \frac{r_x^2(0) - r_x^2(1)}{r_x(1)} = 8$$

Solving these equations for $r_x(0)$ and $r_x(1)$ we find

$$r_x(0) = 3 \quad ; \quad r_x(1) = 1$$

(b) In general, the pth-order minimum variance estimate the autocorrelation sequence $r_x(k)$ may be recovered from the minimum variance spectrum estimate, $\widehat{P}_{MV}(e^{j\omega})$. However, the procedure is not easy. Since $\mathbf{R}_x$ is Toeplitz, if $\mathbf{R}_x$ is a $p \times p$ matrix, then only p entries in $\mathbf{R}_x$ need to be determined — either the first row or the first column. The inverse of $\mathbf{R}_x$ is centrosymmetric and, given $\widehat{P}_{MV}(e^{j\omega})$, we know the sums along the p diagonals of $\mathbf{R}_x^{-1}$. Expressing the inverse of $\mathbf{R}_x^{-1}$ in terms of $r_x(k)$, and summing along the diagonals, we obtain p nonlinear equations in p unknowns which, in theory, can be solved for $r_x(k)$.

8.12 The second-order maximum entropy spectrum of a process $x(n)$ is

$$\widehat{P}_{mem}(e^{j\omega}) = \frac{2}{|1 - 0.5e^{-j\omega} + 0.25e^{-2j\omega}|^2}$$

(a) What is the first-order maximum entropy spectrum?

(b) Find the second-order minimum variance spectrum estimate.

Solution _____

(a) Given the second-order MEM spectrum of $x(n)$, it follows that the second-order all-pole model is

$$A(z) = 1 - 0.5z^{-1} + 0.25z^{-2}$$

The inverse Levinson-Durbin recursion yields the following autocorrelation sequence

$$\mathbf{r}_x = \begin{bmatrix} 2.5397, & 1.0159, & -0.1270 \end{bmatrix}^T$$

Therefore, the first-order MEM spectrum is

$$\widehat{P}_{mem}(e^{j\omega}) = \frac{\epsilon}{\left|1 - \dfrac{r_x(1)}{r_x(0)}e^{-j\omega}\right|^2} = \frac{\epsilon}{|1 - 0.4e^{-j\omega}|^2}$$

where

$$\epsilon = r_x(0) + a(1)r_x(1) = r_x(0) - \frac{r_x^2(1)}{r_x(0)} = 2.1333$$

(b) Having found the autocorrelations $r_x(0)$, $r_x(1)$, and $r_x(2)$, we may now find the second-order MV estimate as follows. With

$$\mathbf{R}_x = \begin{bmatrix} r_x(0) & r_x(1) & r_x(2) \\ r_x(1) & r_x(0) & r_x(1) \\ r_x(2) & r_x(1) & r_x(0) \end{bmatrix} = \begin{bmatrix} 2.5397 & 1.0159 & -0.1270 \\ 1.0159 & 2.5397 & 1.0159 \\ -0.1270 & 1.0159 & 2.5397 \end{bmatrix}$$

for the inverse we have

$$\mathbf{R}_x^{-1} = \begin{bmatrix} 0.50 & -0.25 & 0.125 \\ -0.25 & 0.5937 & -0.25 \\ 0.125 & -0.25 & 0.50 \end{bmatrix}$$

Therefore,

$$\begin{aligned} q(0) &= 1.5937 \\ q(1) &= -0.5 \\ q(2) &= 0.125 \end{aligned}$$

and the MV spectrum becomes,

$$\widehat{P}_{MV}(e^{j\omega}) = \frac{3}{q(0) + 2q(1)\cos\omega + 2q(2)\cos 2\omega} = \frac{3}{1.5937 - \cos\omega + 0.25\cos 2\omega}$$

8.13 From measurements of a process $x(n)$, we estimate the following values for the autocorrelation sequence:

$$r_x(k) = \alpha^{|k|} \quad ; \quad |k| \le M$$

where $|\alpha| < 1$. Estimate the power spectrum using

(a) The Blackman-Tukey method with a rectangular window.

(b) The minimum variance method.

(c) The maximum entropy method.

Solution _____

(a) Using the Blackman-Tukey method with a rectangular window we have

$$
\begin{aligned}
P_{BT}(e^{j\omega}) &= \sum_{k=-M}^{M} \alpha^{|k|} e^{-jk\omega} \\
&= \sum_{k=0}^{M} \alpha^k e^{-jk\omega} + \sum_{k=0}^{M} \alpha^k e^{jk\omega} - 1 \\
&= 2\alpha^{M+1} \frac{1 - \alpha^2 + \cos M\omega - \cos(M+1)\omega}{1 + \alpha^2 - 2\alpha\cos\omega}
\end{aligned}
$$

(Note: the last step above requires a bit of algebra).

(b) For the minimum variance method,

$$\widehat{P}_{MV}(e^{j\omega}) = \frac{M+1}{\mathbf{e}^H \mathbf{R}_M^{-1} \mathbf{e}} = \frac{M+1}{q(0) + 2\sum_{k=1}^{M} q(k)\cos k\omega}$$

where

$$\mathbf{R}_M = \text{Toep}\{r_x(0), r_x(1), \ldots, r_x(M)\}$$

and the coefficients $q(k)$ are the sums along the diagonals of $\mathbf{R}_M^{-1}$. As we saw in Example 5.2.11 (p. 258), the inverse of this autocorrelation matrix is

$$
\mathbf{R}_M^{-1} = \frac{1}{1 - \alpha^2}
\begin{bmatrix}
1 & -\alpha & 0 & \cdots & 0 & 0 \\
-\alpha & 1+\alpha^2 & -\alpha & \cdots & 0 & 0 \\
0 & -\alpha & 1+\alpha^2 & \cdots & 0 & 0 \\
\vdots & \vdots & \vdots & & \vdots & \vdots \\
0 & 0 & 0 & \cdots & 1+\alpha^2 & -\alpha \\
0 & 0 & 0 & \cdots & -\alpha & 1
\end{bmatrix}
$$

Therefore,

$$\widehat{P}_{MV}(e^{j\omega}) = \frac{(M+1)(1-\alpha^2)}{(M+1) + (M-1)\alpha^2 - 2M\alpha\cos\omega}$$

(c) The maximum entropy spectrum is given by

$$\widehat{P}_{mem}(e^{j\omega}) = \frac{\epsilon_p}{\left| 1 + \sum_{k=1}^{p} a_p(k) e^{jk\omega} \right|^2}$$

Solving the autocorrelation normal equations for the coefficients $a_p(k)$, we find

$$\mathbf{a}_p = \begin{bmatrix} 1, -\alpha, 0, \ldots, 0 \end{bmatrix}^T$$

and

$$\epsilon_p = 1 - \alpha^2$$

Therefore, the MEM spectrum is

$$\widehat{P}_{mem}(e^{j\omega}) = \frac{1 - \alpha^2}{\left| 1 - \alpha e^{-j\omega} \right|^2} = \frac{1 - \alpha^2}{1 + \alpha^2 - 2\alpha \cos \omega}$$

8.14 In Eq. (8.97), the entropy of a Gaussian random process was given as

$$H(x) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \ln P_x(e^{j\omega}) d\omega$$

In this problem, we derive another expression for the entropy. Let $x(n)$ be a real-valued zero mean Gaussian random process, and let $\mathbf{x} = [x(0), x(1), \ldots, x(N-1)]^T$ be an N-dimensional Gaussian random vector that is formed from samples of this process. The probability density function for this random vector is

$$f_x(\mathbf{x}) = \frac{1}{(2\pi)^{N/2}(\det \mathbf{R}_x)^{1/2}} \exp\{-\tfrac{1}{2}\mathbf{x}^T \mathbf{R}_x^{-1} \mathbf{x}\}$$

where $\mathbf{R}_x$ is the $N \times N$ autocorrelation matrix of the vector $\mathbf{x}$.

(a) The *average entropy* of a random vector $\mathbf{x}$ is defined as

$$H_N(\mathbf{x}) = -\frac{1}{N} \int f_x(\mathbf{x}) \ln f_x(\mathbf{x}) d\mathbf{x}$$

Show that the average entropy of a zero mean Gaussian random vector is

$$H_N(\mathbf{x}) = \frac{1}{2} \ln(2\pi e) + \frac{1}{2N} \ln(\det \mathbf{R}_x)$$

(b) Show that the average entropy of a Gaussian random vector may be written as

$$H_N(\mathbf{x}) = \frac{1}{2} \ln(2\pi e) + \frac{1}{N} \sum_{k=0}^{N-1} \ln \epsilon_k$$

where ϵ_k is the prediction error sequence that is generated with the Levinson-Durbin recursion from the autocorrelation sequence $r_x(k)$.

(c) The *entropy rate* of a process $x(n)$ is the limit, as $N \to \infty$, of the average entropy,

$$\bar{H}(\mathbf{x}) = \lim_{N \to \infty} H_N(x)$$

Given a partial autocorrelation sequence, $r_x(k)$, for $k = 0, 1, \ldots, N-1$, find the spectrum $P_x(e^{j\omega})$ that maximizes $\bar{H}(x)$ subject to the constraint that the spectrum is consistent with the given autocorrelations.

(d) Let $P_x(e^{j\omega})$ be the power spectrum of a wide-sense stationary process with an $N \times N$ autocorrelation matrix $\mathbf{R}_x$, and let $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_N \geq 0$ be the eigenvalues. *Szegö's theorem* states that if $g(\cdot)$ is a continuous real-valued function then

$$\lim_{N \to \infty} \frac{1}{N} \sum_{k=1}^{N} g(\lambda_k) = \frac{1}{2\pi} \int_{-\pi}^{\pi} g[P_x(e^{j\omega})] d\omega$$

Use Szegö's theorem to show that

$$H(x) = 2\bar{H}(x) + \ln(2\pi e)$$

i.e., that the two entropy expressions are proportional to each other.

Solution _____

(a) The logarithm of the probability density function is

$$\ln f_x(\mathbf{x}) = -\tfrac{N}{2}\ln(2\pi) - \tfrac{1}{2}\ln(\det \mathbf{R}_x) - \tfrac{1}{2}\mathbf{x}^T\mathbf{R}_x^{-1}\mathbf{x}$$

Therefore,

$$
\begin{aligned}
H_N(\mathbf{x}) &= -\frac{1}{N}\int f_x(\mathbf{x})\ln f_x(\mathbf{x})d\mathbf{x} \\[2mm]
&= -\frac{1}{N}\left[-\frac{N}{2}\ln(2\pi) - \frac{1}{2}\ln(\det \mathbf{R}_x) - \frac{N}{2}\right] \\[2mm]
&= \frac{1}{2}\ln(2\pi) + \frac{1}{2N}\ln(\det \mathbf{R}_x) + \frac{1}{2} \\[2mm]
&= \frac{1}{2}\ln(2\pi e) + \frac{1}{2N}\ln(\det \mathbf{R}_x)
\end{aligned}
$$

which is the desired result.

(b) Recall that the determinant of an $N \times N$ autocorrelation matrix $\mathbf{R}_x$ is

$$\det \mathbf{R}_x = \prod_{k=0}^{N-1} \epsilon_k$$

Therefore,

$$\frac{1}{2N}\ln(\det \mathbf{R}_x) = \frac{1}{2N}\sum_{k=0}^{N-1}\ln \epsilon_k$$

and the desired result follows.

(c) As was seen in Chapter 5, there is an equivalence between a set of autocorrelations, $r_x(k)$ for $k = 0, 1, \ldots, p$, and a sequence of reflection coefficients, Γ_j. Therefore, finding a power spectrum that maximizes the entropy rate, subject to the constraint that the first N autocorrelations are equal to $r_x(k)$, is equivalent to finding a set of reflection coefficients Γ_j that maximize the entropy rate. Since the entropy rate is

$$H_N(\mathbf{x}) = \frac{1}{2}\ln(2\pi e) + \frac{1}{N}\sum_{k=0}^{N-1}\ln \epsilon_k$$

then maximizing $\bar{H}(\mathbf{x}) = \lim_{N\to\infty} H_N(\mathbf{x})$ is equivalent to maximizing the prediction errors ϵ_k for $k \geq N$. However, the prediction errors are related to the reflection coefficients as follows,

$$\epsilon_k = \epsilon_{k-1}(1 - |\Gamma_k|^2)$$

Therefore, maximizing the entropy rate is accomplished by setting $\Gamma_j = 0$ for $j \geq N$. Thus, $P_x(e^{j\omega})$ is the all-pole spectrum that corresponds to the given partial autocorrelation sequence

(d) If we let the function $g(\cdot)$ be the logarithm, using Szëgo's theorem we have

$$
\begin{aligned}
H(x) &= \frac{1}{2\pi}\int_{-\pi}^{\pi}\ln P_x(e^{j\omega})d\omega = \lim_{N\to\infty}\frac{1}{N}\sum_{k=1}^{N}\ln(\lambda_k) \\[2mm]
&= \lim_{N\to\infty}\frac{1}{N}\ln(\det\mathbf{R}_x) = 2\bar{H}(x) + \ln(2\pi e)
\end{aligned}
$$

8.15 In this problem, we examine how the entropy of a process changes with the addition of a harmonic process. Let $y(n)$ be a random process with power spectrum

$$P_y(e^{j\omega}) = P_x(e^{j\omega}) + P_\epsilon(e^{j\omega})$$

where

$$P_\epsilon(e^{j\omega}) = \begin{cases} 1/\epsilon & ; \quad |\omega - \omega_0| < \epsilon \\ 0 & ; \quad \text{otherwise} \end{cases}$$

(a) Find the entropy of $y(n)$.

(b) What is the entropy of this process in the limit as $\epsilon \to 0$?

Solution _____

(a) For a Gaussian process, the entropy is

$$H(x) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \ln P_x(e^{j\omega}) d\omega$$

With

$$P_y(e^{j\omega}) = P_x(e^{j\omega}) + P_\epsilon(e^{j\omega})$$

it follows that the entropy of $y(n)$ is

$$H(y) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \ln\left[P_x(e^{j\omega}) + P_\epsilon(e^{j\omega}) \right] d\omega$$

which may be rewritten as follows

$$
\begin{aligned}
H(y) &= \frac{1}{2\pi} \int_{-\pi}^{\pi} \ln\left\{ P_x(e^{j\omega}) \left[1 + \frac{P_\epsilon(e^{j\omega})}{P_x(e^{j\omega})} \right] \right\} d\omega \\
&= \frac{1}{2\pi} \int_{-\pi}^{\pi} \ln P_x(e^{j\omega}) d\omega + \frac{1}{2\pi} \int_{-\pi}^{\pi} \ln\left[1 + \frac{P_\epsilon(e^{j\omega})}{P_x(e^{j\omega})} \right] d\omega \\
&= H(x) + \frac{1}{2\pi} \int_{-\pi}^{\pi} \ln\left[1 + \frac{P_\epsilon(e^{j\omega})}{P_x(e^{j\omega})} \right] d\omega
\end{aligned}
$$

Using the definition of $P_\epsilon(e^{j\omega})$ it follows that $H(y)$ is

$$H(y) = H(x) + \frac{1}{2\pi} \int_{\omega_0-\epsilon}^{\omega_0+\epsilon} \ln\left[1 + \frac{\epsilon}{P_x(e^{j\omega})} \right] d\omega$$

If ϵ is small, then

$$\ln\left[1 + \frac{\epsilon}{P_x(e^{j\omega})} \right] \approx \frac{\epsilon}{P_x(e^{j\omega})}$$

and we have

$$H(y) \approx H(x) + \frac{1}{2\pi} \int_{\omega-\epsilon}^{\omega+\epsilon} \ln \frac{\epsilon}{P_x(e^{j\omega})} d\omega \approx H(x) + \frac{\epsilon}{\pi} \ln \frac{\epsilon}{P_x(e^{j\omega_0})}$$

(b) Using the result from part (a), it follows that in the limit, as $\epsilon \to 0$, the entropy of $y(n)$ approaches the entropy of $x(n)$,

$$\lim_{\epsilon \to 0} H(y) = H(x)$$

i.e., the addition of a harmonic process does not increase the entropy of a process.

8.16 Given an autocorrelation sequence $r_x(k)$ for $k = 0, 1, \ldots, p$, the maximum entropy spectrum is

$$\widehat{P}_{mem}(e^{j\omega}) = \frac{\epsilon_p}{\left| 1 + \displaystyle\sum_{k=1}^{p} a_p(k) e^{-jk\omega} \right|^2}$$

where the coefficients $a_p(k)$ are the solution to the normal equations $\mathbf{R}_x \mathbf{a}_p = \epsilon_p \mathbf{u}_1$. If Γ_k are the reflection coefficients produced by the Levinson-Durbin recursion, show that the MEM spectrum may be upper and lower bounded in terms of Γ_k as follows,

$$r_x(0) \prod_{k=1}^{p} \frac{1 - |\Gamma_k|}{1 + |\Gamma_k|} \leq \widehat{P}_{mem}(e^{j\omega}) \leq r_x(0) \prod_{k=1}^{p} \frac{1 + |\Gamma_k|}{1 - |\Gamma_k|}$$

Hint: Begin with the frequency domain version of the Levinson order-update equation and use the inequality,

$$\big||a| - |b|\big| \leq |a + b| \leq |a| + |b|$$

Solution _____

From the Levinson-Durbin recursion we have

$$A_p(e^{j\omega}) = A_{p-1}(e^{j\omega}) + \Gamma_p e^{-jp\omega} A_{p-1}^*(e^{j\omega})$$

Therefore,

$$\left| A_p(e^{j\omega}) \right| \leq \left| A_{p-1}(e^{j\omega}) \right| + |\Gamma_p| \left| A_{p-1}^*(e^{j\omega}) \right| = \big(1 + |\Gamma_p|\big) \left| A_{p-1}(e^{j\omega}) \right|$$

By induction, it follows that

$$\left| A_p(e^{j\omega}) \right| \leq |A_0(e^{j\omega})| \prod_{k=1}^{p} \big(1 + |\Gamma_k|\big)$$

and, since $A_0(e^{j\omega}) = 1$, then

$$\left| A_p(e^{j\omega}) \right| \leq \prod_{k=1}^{p} \big(1 + |\Gamma_k|\big)$$

In a similar fashion,

$$\left| A_p(e^{j\omega}) \right| \geq \left| \left| A_{p-1}(e^{j\omega}) \right| - |\Gamma_p| \left| A_{p-1}^*(e^{j\omega}) \right| \right| = \big(1 - |\Gamma_p|\big) \left| A_{p-1}(e^{j\omega}) \right|$$

Again, by induction, we find

$$\left| A_p(e^{j\omega}) \right| \geq \prod_{k=1}^{p} \big(1 - |\Gamma_k|\big)$$

Since

$$\widehat{P}_{mem}(e^{j\omega}) = \frac{\epsilon_p}{|A_p(e^{j\omega})|^2}$$

with

$$\epsilon_p = r_x(0) \prod_{k=1}^{p} \big(1 - |\Gamma_k|^2\big) = r_x(0) \prod_{k=1}^{p} \big(1 - |\Gamma_k|\big)\big(1 + |\Gamma_k|\big)$$

we have

$$\frac{r_x(0) \displaystyle\prod_{k=1}^{p} \left(1 - |\Gamma_k|\right)\left(1 + |\Gamma_k|\right)}{\displaystyle\prod_{k=1}^{p} \left(1 + |\Gamma_k|\right)^2} \leq \widehat{P}_{mem}(e^{j\omega}) \leq \frac{r_x(0) \displaystyle\prod_{k=1}^{p} \left(1 - |\Gamma_k|\right)\left(1 + |\Gamma_k|\right)}{\displaystyle\prod_{k=1}^{p} \left(1 - |\Gamma_k|\right)^2}$$

or, performing the division, we have

$$r_x(0) \prod_{k=1}^{p} \frac{1 - |\Gamma_k|}{1 + |\Gamma_k|} \leq \widehat{P}_{mem}(e^{j\omega}) \leq r_x(0) \prod_{k=1}^{p} \frac{1 + |\Gamma_k|}{1 - |\Gamma_k|}$$

as was to be shown.

8.17 Let $x(n)$ be a first-order Gaussian autoregressive process with power spectrum

$$P_x(z) = \frac{c}{(1 - az^{-1})(1 - az)}$$

where a and c are real numbers.

(a) With the constraint that the total power in the signal is equal to one, find the value or values of a and c that maximize the entropy of $x(n)$.

(b) Repeat part (a) and find the value or values of a and c that minimize the entropy.

Solution ───

(a) The autocorrelation sequence corresponding to the power spectrum

$$P_x(z) = \frac{c}{(1 - az^{-1})(1 - az)}$$

is

$$r_x(k) = \frac{c}{1 - |a|^2} a^{|k|}$$

Therefore, the unit power constraint, $r_x(0) = 1$, requires that

$$c = 1 - |a|^2$$

Thus, the power spectrum is

$$P_x(z) = \frac{1 - |a|^2}{(1 - az^{-1})(1 - az)} = \sigma_0^2 Q(z) Q(z^{-1})$$

where

$$\sigma_0^2 = 1 - |a|^2$$

Now recall from the spectral factorization theorem (p. 105) that

$$\sigma_0^2 = \exp\left\{ \frac{1}{2\pi} \int_{-\pi}^{\pi} \ln P_x(e^{j\omega}) d\omega \right\} = e^{H(x)}$$

Therefore, since maximizing the entropy is equivalent to maximizing $\sigma_0^2 = 1 - |a|^2$, then the maximum entropy spectrum is formed when $a = 0$, i.e., $x(n)$ is white noise.

(b) The minimum entropy spectrum is formed in the limit as $|a| \to 1$, which corresponds to a harmonic process.

8.18 The estimated autocorrelation sequence of a random process $x(n)$ for lags $k = 0, 1, 2, 3, 4$ are

$$r_x(0) = 2 \quad ; \quad r_x(1) = 1 \quad ; \quad r_x(2) = 1 \quad ; \quad r_x(3) = 0.5 \quad ; \quad r_x(4) = 0$$

Estimate the power spectrum of $x(n)$ for each of the following cases.

(a) $x(n)$ is an AR(2) process.

(b) $x(n)$ is an MA(2) process.

(c) $x(n)$ is an ARMA(1,1) process.

(d) $x(n)$ contains a single sinusoid in white noise.

Solution _____

(a) For an AR(2) process, we want to find a second-order AR model. This is done by solving the normal equations

$$\begin{bmatrix} r_x(0) & r_x(1) \\ r_x(1) & r_x(0) \end{bmatrix} \begin{bmatrix} a(1) \\ a(2) \end{bmatrix} = - \begin{bmatrix} r_x(1) \\ r_x(2) \end{bmatrix}$$

For the given autocorrelation sequence, these become

$$\begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} a(1) \\ a(2) \end{bmatrix} = - \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

Thus, the coefficients are

$$\begin{bmatrix} a(1) \\ a(2) \end{bmatrix} = -\frac{1}{3} \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

with a modeling error

$$\epsilon_2 = r_x(0) + a(1)r_x(1) + a(2)r_x(2) = \tfrac{4}{3}$$

Therefore,

$$A(e^{j\omega}) = 1 - \tfrac{1}{3}e^{-j\omega} - \tfrac{1}{3}e^{-2j\omega}$$

and the power spectrum is

$$\widehat{P}_x(e^{j\omega}) = \frac{\frac{4}{3}}{\left| 1 - \tfrac{1}{3}e^{-j\omega} - \tfrac{1}{3}e^{-j\omega} \right|^2}$$

(b) For an MA(2) process,

$$\widehat{P}_{MA}(e^{j\omega}) = \sum_{k=-2}^{2} r_x(k)e^{-jk\omega} = 2 + 2\cos\omega + 2\cos 2\omega$$

(c) For an ARMA(1,1) process, we must solve the Yule-Walker equations

$$\begin{bmatrix} r_x(0) & r_x(1) \\ r_x(1) & r_x(0) \\ r_x(2) & r_x(1) \end{bmatrix} \begin{bmatrix} 1 \\ a(1) \end{bmatrix} = \begin{bmatrix} c(0) \\ c(1) \\ 0 \end{bmatrix}$$

The coefficient $a(1)$ is found from the last equation as follows,

$$a(1) = -r_x(2)/r_x(1) = -1$$

Solving for $c(0)$ and $c(1)$ we have

$$\begin{bmatrix} c(0) \\ c(1) \end{bmatrix} = \begin{bmatrix} r_x(0) & r_x(1) \\ r_x(1) & r_x(0) \end{bmatrix} \begin{bmatrix} 1 \\ a(1) \end{bmatrix} = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ -1 \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

From Eq. (4.140) we have

$$\begin{aligned} B(z)B(z^{-1}) &= A(z^{-1})C(z) = \left[1 + a(1)z\right]\left[\sum_{k=-\infty}^{\infty} c(k)z^{-k}\right] \\ &= \cdots + [c(-1) - c(0)z + [c(0) - c(1)] + [c(1) - c(2)]z^{-1} + \cdots \end{aligned}$$

Since $B(z)B(z^{-1})$ is symmetric and of order one in z and z^{-1}, with $c(0) = 1$ and $c(1) = -1$, it follows that

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

Therefore, the power spectrum estimate is

$$\widehat{P}_x(e^{j\omega}) = \frac{|B(e^{j\omega})|^2}{|A(e^{j\omega})|^2} = \frac{2 - 2\cos\omega}{2 - 2\cos\omega} = 1$$

(d) If the signal consists of a single sinusoid, assuming that the autocorrelations are exact, we may use the Pisarenko harmonic decomposition. With one sinusoid, we use $p = 3$, and find the minimum eigenvalue and eigenvector for the 3×3 autocorrelation matrix

$$\mathbf{R} = \begin{bmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{bmatrix}$$

To find the eigenvalues, we have

$$\det(\mathbf{R} - \lambda\mathbf{I}) = \det\begin{bmatrix} 2-\lambda & 1 & 1 \\ 1 & 2-\lambda & 1 \\ 1 & 1 & 2-\lambda \end{bmatrix} = (2-\lambda)\left[(2-\lambda)^2 - 1\right] - \left[(2-\lambda) - 1\right] + \left[1 - (2-\lambda)\right] = (1-\lambda)^2(4-\lambda)$$

Therefore, the minimum eigenvalue is repeated and has a value

$$\lambda_{\min} = \sigma_w^2 = 1$$

Thus, we must reduce the order of the Pisarenko estimate to $p = 2$, and find the eigenvector corresponding to the minimum eigenvalue of the matrix

$$\mathbf{R} = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$$

For this matrix, $\lambda_{\min} = 1$ and the corresponding eigenvector is

$$\mathbf{v}_{\min} = \begin{bmatrix} 1, & -1 \end{bmatrix}^T$$

Therefore, the eigenfilter is

$$V_{\min}(z) = 1 - z^{-1}$$

so the sinusoid frequency is $\omega_0 = \pi$. Finally, from the autocorrelation sequence we see, by inspection, that the sinusoid power is $P = A^2 = 2$.

8.19 The first three values of the autocorrelation sequence for a process $x(n)$ are:

$$r_x(0) = 1 \quad ; \quad r_x(1) = 0 \quad ; \quad r_x(2) = -\alpha$$

where $0 < \alpha < 1$. The eigenvalues of the 3×3 autocorrelation matrix that is formed from these autocorrelations are $\lambda_1 = 1 + \alpha$, $\lambda_2 = 1$, and $\lambda_3 = 1 - \alpha$, and the corresponding eigenvectors are

$$\mathbf{v}_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} \quad ; \quad \mathbf{v}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \quad ; \quad \mathbf{v}_3 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$$

(a) Use the Blackman-Tukey method with a rectangular window to estimate the power spectrum of $x(n)$, and make a carefully labeled sketch of your estimate.

(b) Suppose that $x(n)$ is known to consist of two complex exponentials in white noise. Estimate the power spectrum of $x(n)$ and make a carefully labeled sketch of your estimate.

Solution _____

(a) The Blackman-Tukey estimate with a rectangular window that extends over the interval $|k| \leq 2$ is

$$\widehat{P}_{BT}(e^{j\omega}) = \sum_{k=-2}^{2} r_x(k)e^{-jk\omega} = r_x(0) + 2r_x(1)\cos\omega + 2r_x(2)\cos 2\omega$$

Therefore,

$$\widehat{P}_{BT}(e^{j\omega}) = 1 - 2\alpha\cos 2\omega$$

(b) If $x(n)$ is known to consist of two complex exponentials, then there are several different options. Using principal components with the Blackman-Tukey method we have

$$\widehat{P}_{PC-BT}(e^{j\omega}) = \frac{1}{M}\sum_{i=1}^{p} \lambda_i |\mathbf{e}^H \mathbf{v}_i|^2$$

With $M = 2$ and $p = 2$ this becomes

$$\widehat{P}_{PC-BT}(e^{j\omega}) = \frac{1}{2}\Big[\lambda_1 |\mathbf{e}^H \mathbf{v}_1|^2 + \lambda_2 |\mathbf{e}^H \mathbf{v}_2|^2\Big]$$

From the given eigenvectors and eigenvalues we have

$$\mathbf{e}^H \mathbf{v}_1 = \frac{1}{\sqrt{2}}\begin{bmatrix} 1, & e^{-j\omega}, & e^{-2j\omega} \end{bmatrix}\begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} = \frac{1}{\sqrt{2}}\big[1 - e^{-j2\omega}\big]$$

and

$$\mathbf{e}^H \mathbf{v}_2 = \begin{bmatrix} 1, & e^{-j\omega}, & e^{-2j\omega} \end{bmatrix}\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} = e^{-j\omega}$$

Therefore,

$$\widehat{P}_{PC-BT}(e^{j\omega}) = \frac{1}{2}\left[(1+\alpha)\frac{1}{2}\left|1 - e^{-j2\omega}\right|^2 + \left|e^{-j\omega}\right|^2\right] = \frac{2+\alpha}{2} - \frac{1+\alpha}{2}\cos 2\omega$$

Other methods we could use include the principal components minimum variance and autoregressive frequency estimation methods.

8.20 Suppose that we would like to estimate the power spectrum of an AR(2) process

$$x(n) = a(1)x(n-1) + a(2)x(n-2) + w(n)$$

where $w(n)$ is unit variance white noise. However, our measurements of $x(n)$ are noisy, and what we observe is the process

$$y(n) = x(n) + v(n)$$

where the measurement noise, $v(n)$, is uncorrelated with $x(n)$. It is known that $v(n)$ is a first-order moving average process,

$$v(n) = b(0)q(n) + b(1)q(n-1)$$

where $q(n)$ is white noise. Based on measurements of $v(n)$, the power spectrum of $v(n)$ is estimated to be

$$\widehat{P}_v(e^{j\omega}) = 3 + 2\cos\omega$$

From $y(n)$ we estimate the following values of the autocorrelation sequence $r_y(k)$,

$$\widehat{r}_y(0) = 5 \quad ; \quad \widehat{r}_y(1) = 2 \quad ; \quad \widehat{r}_y(2) = 0 \quad ; \quad \widehat{r}_y(3) = -1 \quad ; \quad \widehat{r}_y(4) = 0.5$$

Using all of the given information, estimate the power spectrum of $x(n)$ using the maximum entropy method.

Solution _____

Since $x(n)$ and $v(n)$ are uncorrelated,

$$r_y(k) = r_x(k) + r_v(k)$$

Therefore, given the autocorrelation $r_y(k)$ and $r_v(k)$, we may find $r_x(k)$,

$$r_x(k) = r_y(k) - r_v(k)$$

We are given an estimate of the power spectrum of $v(n)$,

$$\widehat{P}_v(e^{j\omega}) = 3 + 2\cos\omega$$

which implies that the estimated autocorrelation sequence is

$$r_v(k) = 3\delta(k) + \delta(k-1) + \delta(k+1)$$

(note that since $v(n)$ is an MA(1) process, then $r_v(k) = 0$ for $|k| > 1$). From the given estimated autocorrelations for $y(n)$, we have for $\widehat{r}_x(k)$,

$$\widehat{r}_x(0) = 2 \quad ; \quad \widehat{r}_x(1) = 1 \quad ; \quad \widehat{r}_x(2) = 0 \quad ; \quad \widehat{r}_x(3) = -1 \quad ; \quad \widehat{r}_x(4) = 0.5$$

Now, we may estimate the power spectrum of $x(n)$ using the maximum entropy method. The normal equations that we must solve are,

$$\begin{bmatrix} r_x(0) & r_x(1) \\ r_x(1) & r_x(0) \end{bmatrix} \begin{bmatrix} a(1) \\ a(2) \end{bmatrix} = - \begin{bmatrix} r_1(1) \\ r_x(2) \end{bmatrix}$$

Using the given autocorrelations, these become

$$\begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} a(1) \\ a(2) \end{bmatrix} = - \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

and the solution is $a(1) = -2/3$ and $a(2) = 1/3$, or

$$A(e^{j\omega}) = 1 - \tfrac{2}{3}e^{-j\omega} + \tfrac{1}{3}e^{-j\omega}$$

Finally, for the modeling error,

$$\epsilon = r_x(0) + a(1)r_x(1) + a(2)r_x(2) = 4/3$$

and the MEM spectrum estimate is

$$\widehat{P}_{mem}(e^{j\omega}) = \frac{\tfrac{4}{3}}{\left| 1 - \tfrac{2}{3}e^{-j\omega} + \tfrac{1}{3}e^{-j2\omega} \right|^2}$$

8.21 Show that for $N \gg 1$, estimating the order of an autoregressive process by minimizing $\text{FPE}(p)$ is equivalent to minimizing $\text{AIC}(p)$. Hint: Show that for large N,

$$N \ln \text{FPE}(p) \approx \text{AIC}(p)$$

and use the fact that, if x is small, then $\ln(1 + x) \approx x$.

Solution _____

The Akaiki FPE is

$$\text{FPE}(p) = \mathcal{E}_p \frac{N + p + 1}{N - p - 1} = \mathcal{E}_p \frac{1 + \frac{p+1}{N}}{1 - \frac{p+1}{N}}$$

Therefore,

$$\ln \text{FPE}(p) = \ln \mathcal{E}_p + \ln\left(1 + \frac{p+1}{N}\right) - \ln\left(1 - \frac{p+1}{N}\right)$$

If x is small, then

$$\ln(1 + x) \approx x$$

Using this approximation we have, for large N,

$$\ln \text{FPE}(p) \approx \ln \mathcal{E}_p + 2\frac{p+1}{N}$$

and

$$N \ln \text{FPE}(p) \approx N \ln \mathcal{E}_p + 2(p + 1)$$

Since

$$\text{AIC}(p) = N \ln \mathcal{E}_p + 2p$$

then

$$N \ln \text{FPE}(p) \approx \text{AIC}(p)$$

Therefore, since the logarithm is a monotonically increasing function, then, for large N, minimizing $\text{FPE}(p)$ is equivalent to minimizing $\text{AIC}(p)$.

$$\text{AIC}(p) = N \ln \mathcal{E}_p + 2p$$

8.22 You are given the following values for the autocorrelation sequence of a wide-sense stationary process $x(n)$,

$$r_x(0) = 2 \quad ; \quad r_x(1) = \sqrt{3}/2 \quad ; \quad r_x(2) = 0.5$$

The eigenvalues of the 3×3 Toeplitz autocorrelation matrix are $\lambda_1 = 3.5$, $\lambda_2 = 1.5$, and $\lambda_3 = 1.0$ and the corresponding normalized eigenvectors are

$$\mathbf{v}_1 = \sqrt{2/5} \begin{bmatrix} \sqrt{3}/2 \\ 1 \\ \sqrt{3}/2 \end{bmatrix} \quad ; \quad \mathbf{v}_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} \quad ; \quad \mathbf{v}_3 = \frac{1}{\sqrt{5}} \begin{bmatrix} 1 \\ -\sqrt{3} \\ 1 \end{bmatrix}$$

It is known that $x(n)$ consists of a single sinusoid in white noise.

(a) Estimate the frequency of the sinusoid using the Blackman-Tukey method of frequency estimation.

(b) Use the MUSIC algorithm to estimate the frequency of the sinusoid.

(c) Repeat part (b) using the minimum norm algorithm.

Solution _____

(a) The Blackman-Tukey frequency estimate is given by

$$\widehat{P}_{PC-BT}(e^{j\omega}) = \frac{1}{M} \sum_{i=1}^{p} \lambda_i |\mathbf{e}^H \mathbf{v}_i|^2$$

where M is the number of autocorrelations, and p is the number of complex exponentials. Therefore, with $M = 3$ and $p = 2$, using the given eigenvalues and eigenvectors, we have

$$\widehat{P}_{PC-BT}(e^{j\omega}) = \frac{1}{3} \left[3.5(\tfrac{2}{5}) \left| \tfrac{\sqrt{3}}{2} + e^{-j\omega} + \tfrac{\sqrt{3}}{2} e^{-j2\omega} \right|^2 + 1.5(\tfrac{1}{2}) \left| -1 + e^{-j2\omega} \right|^2 \right]$$

$$= \tfrac{5}{3} + \tfrac{14}{15}\sqrt{3}\cos\omega - \tfrac{6}{5}\cos 2\omega$$

(b) The MUSIC estimate is given by

$$\widehat{P}_{MU}(e^{j\omega}) = \frac{1}{\displaystyle\sum_{i=p+1}^{M} \left| \mathbf{e}^H \mathbf{v}_i \right|^2}$$

With $M = 3$ and $p = 2$ this becomes

$$\widehat{P}_{MU}(e^{j\omega}) = \frac{1}{|\mathbf{e}^H \mathbf{v}_3|^2} = \frac{1}{\tfrac{1}{5}|1 - \sqrt{3}e^{-j\omega} + e^{-j2\omega}|^2} = \frac{5}{5 - 4\sqrt{3}\cos\omega + 2\cos 2\omega}$$

The peak of the MUSIC estimate occurs at the frequency where the denominator is a minimum, which is the frequency where the derivative is equal to zero,

$$\frac{\partial}{\partial\omega}\left[5 - 4\sqrt{3}\cos\omega + 2\cos 2\omega \right] = 4\sqrt{3}\sin\omega - 4\sin 2\omega = 0$$

or

$$\sqrt{3}\sin\omega = \sin 2\omega$$

which, using a trigonometric identity, becomes

$$\sqrt{3}\sin\omega = 2\sin\omega\cos\omega$$

Thus, the peak occurs at the point where

$$\cos\omega = \frac{\sqrt{3}}{2}$$

or

$$\omega_0 = \pi/6$$

(c) For the minimum norm algorithm, we first form the projection matrix

$$\mathbf{P}_n = \mathbf{V}_n\mathbf{V}_n^H = \frac{1}{5}\begin{bmatrix} 1 \\ -\sqrt{3} \\ 1 \end{bmatrix}\begin{bmatrix} 1 & -\sqrt{3} & 1 \end{bmatrix} = \frac{1}{5}\begin{bmatrix} 1 & -\sqrt{3} & 1 \\ -\sqrt{3} & 3 & -\sqrt{3} \\ 1 & -\sqrt{3} & 1 \end{bmatrix}$$

With

$$\lambda = \frac{1}{\mathbf{u}_1^T\mathbf{P}_n\mathbf{u}_1} = 5$$

and

$$\mathbf{a} = \lambda\mathbf{P}_n\mathbf{u}_1 = \begin{bmatrix} 1 \\ -\sqrt{3} \\ 1 \end{bmatrix}$$

we see that the minimum norm estimate

$$\widehat{P}_{MV}(e^{j\omega}) = \frac{1}{|\mathbf{e}^H\mathbf{a}|^2}$$

is the same as the MUSIC algorithm.

8.23 The Pisarenko harmonic decomposition provides a way to estimate the frequencies of a sum of complex exponentials in white noise. As described in Sect. 8.6.2, the powers of the complex exponentials may be found by solving the set of linear equations given in Eq. (8.160). Another method that may be used is based on the orthogonality of the trigonometric sine and cosine functions. This orthogonality condition implies that

$$\det \begin{bmatrix} \sin\omega_1 & \sin\omega_2 & \cdots & \sin\omega_p \\ \sin 2\omega_1 & \sin 2\omega_2 & \cdots & \sin 2\omega_p \\ \vdots & \vdots & & \vdots \\ \sin p\omega_1 & \sin p\omega_2 & \cdots & \sin p\omega_p \end{bmatrix} \neq 0$$

provided $0 < \omega_i < \pi$ and $\omega_i \neq \omega_j$, and

$$\det \begin{bmatrix} 1 & 1 & \cdots & 1 \\ \cos\omega_1 & \cos\omega_2 & \cdots & \cos\omega_p \\ \vdots & \vdots & & \vdots \\ \cos(p-1)\omega_1 & \cos(p-1)\omega_2 & \cdots & \cos(p-1)\omega_p \end{bmatrix} \neq 0$$

provided $0 \leq \omega_i \leq \pi$ and $\omega_i \neq \omega_j$.

(a) Given the autocorrelation sequence of a pth-order harmonic process,

$$r_x(k) = \sum_{i=1}^{p} P_i e^{jk\omega_i} + \sigma_w^2 \delta(k)$$

evaluate the imaginary part of $r_x(k)$ and use the orthogonality of the sine functions to derive a set of linear equations that may be solved to find the signal powers P_i.

(b) How would you modify this algorithm if some of the frequencies were equal to zero or π?

(c) How would you modify this approach for a sum of sinusoids in white noise?

Solution _____

(a) Given the autocorrelation sequence

$$r_x(k) = \sum_{i=1}^{p} P_i e^{jk\omega_i} + \sigma_w^2 \delta(k)$$

we see that the imaginary part of $r_x(k)$ is

$$\text{Im}\{r_x(k)\} = \sum_{i=1}^{p} P_i \sin k\omega_i$$

Therefore, given $r_x(k)$ for $k = 1, 2, \ldots, p$, we may form the following set of p linear equations in the p unknowns, P_i,

$$
\begin{bmatrix}
\sin \omega_1 & \sin \omega_2 & \cdots & \sin \omega_p \\
\sin 2\omega_1 & \sin 2\omega_2 & \cdots & \sin 2\omega_p \\
\vdots & \vdots & & \vdots \\
\sin p\omega_1 & \sin p\omega_2 & \cdots & \sin p\omega_p
\end{bmatrix}
\begin{bmatrix}
P_1 \\
P_2 \\
\vdots \\
P_p
\end{bmatrix}
=
\begin{bmatrix}
\operatorname{Im}\{r_x(1)\} \\
\operatorname{Im}\{r_x(2)\} \\
\vdots \\
\operatorname{Im}\{r_x(p)\}
\end{bmatrix}
$$

Due to the orthogonality of the sine functions, if the frequencies ω_i are not equal to either zero or π, and if $\omega_i \neq \omega_k$, then the matrix is invertible, and we may solve uniquely for the powers P_i.

(b) If one of the frequencies is equal to zero or π, then we may solve for the signal powers as follows. Taking the real part of the autocorrelation sequence we have

$$
\operatorname{Re}\{r_x(k)\} = \sum_{i=1}^{p} P_i \cos k\omega_i + \sigma_w^2 \delta(k)
$$

Therefore, given $r_x(k)$ for $k = 1, 2, \ldots, p$, we may form the following set of p linear equations in the p unknowns, P_i,

$$
\begin{bmatrix}
1 & 1 & \cdots & 1 \\
\cos \omega_1 & \cos \omega_2 & \cdots & \cos \omega_p \\
\vdots & \vdots & & \vdots \\
\cos(p-1)\omega_1 & \cos(p-1)\omega_2 & \cdots & \cos(p-1)\omega_p
\end{bmatrix}
\begin{bmatrix}
P_1 \\
P_2 \\
\vdots \\
P_p
\end{bmatrix}
=
\begin{bmatrix}
\operatorname{Re}\{r_x(0)\} + \sigma_w^2 \\
\operatorname{Re}\{r_x(1)\} \\
\vdots \\
\operatorname{Re}\{r_x(p-1)\}
\end{bmatrix}
$$

Due to the orthogonality of the cosines, this the matrix is invertible, and we may solve uniquely for the powers P_i.

(c) If $x(n)$ is a sum of sinusoids in white noise, then the autocorrelation sequence is a sum of cosines,

$$
r_x(k) = \sum_{i=1}^{p} P_i \cos k\omega_i + \sigma_w^2 \delta(k)
$$

In this case, the powers may be found as in part (b) by solving the set of linear equations

$$
\begin{bmatrix}
1 & 1 & \cdots & 1 \\
\cos \omega_1 & \cos \omega_2 & \cdots & \cos \omega_p \\
\vdots & \vdots & & \vdots \\
\cos(p-1)\omega_1 & \cos(p-1)\omega_2 & \cdots & \cos(p-1)\omega_p
\end{bmatrix}
\begin{bmatrix}
P_1 \\
P_2 \\
\vdots \\
P_p
\end{bmatrix}
=
\begin{bmatrix}
r_x(0) + \sigma_w^2 \\
r_x(1) \\
\vdots \\
r_x(p-1)
\end{bmatrix}
$$

8.24 The Pisarenko harmonic decomposition was derived for a process that consists of a sum of complex exponentials in white noise. In this problem we generalize the decomposition to nonwhite noise. To accomplish this, we begin with an alternate derivation of the Pisarenko decomposition for white noise. Let

$$x(n) = \sum_{k=1}^{p} A_k e^{jn\omega_k} + w(n)$$

where $w(n)$ is noise that is uncorrelated with the complex exponentials.

(a) If $w(n)$ is white noise then, as we saw in Eq. (8.149), the autocorrelation matrix for $x(n)$ may be written as

$$\mathbf{R}_x = \mathbf{EPE}^H + \sigma_w^2 \mathbf{I}$$

where $\mathbf{E}$ is a matrix of complex exponentials and $\mathbf{P}$ is a diagonal matrix of signal powers. If $x(n)$ is filtered with a pth-order FIR filter $\mathbf{a} = [a(0),\ a(1),\ \ldots,\ a(p)]^T$, then the power in the output process is

$$\xi = E\{|y(n)|^2\} = \mathbf{a}^H \mathbf{R}_x \mathbf{a}$$

If $\mathbf{a}$ is constrained to have unit norm, $\mathbf{a}^H \mathbf{a} = 1$, show that the filter that minimizes ξ has p zeros on the unit circle at the frequencies ω_k of the complex exponentials, and show that the minimum value of ξ is equal to σ_w^2.

(b) Now assume that $w(n)$ has an arbitrary power spectrum, $P_w(e^{j\omega})$. If the autocorrelation matrix of $w(n)$ is $\sigma_w^2 \mathbf{Q}$, then the autocorrelation matrix for $x(n)$ becomes

$$\mathbf{R}_x = \mathbf{EPE}^H + \sigma_w^2 \mathbf{Q}$$

Suppose that $x(n)$ is filtered with a pth-order FIR filter $\mathbf{a} = [a(0),\ a(1),\ \ldots,\ a(p)]^T$ that is normalized so that

$$\mathbf{a}^H \mathbf{Q} \mathbf{a} = 1$$

Show that the filter that minimizes the power in the filtered process has p zeros on the unit circle at the frequencies ω_k of the complex exponentials, and that the minimum value is equal to σ_w^2.

(c) Show that minimizing $\xi = \mathbf{a}^H \mathbf{R}_x \mathbf{a}$ subject to the constraint $\mathbf{a}^H \mathbf{Q} \mathbf{a} = 1$ is equivalent to solving the generalized eigenvalue problem

$$\mathbf{R}_x \mathbf{a} = \lambda \mathbf{Q} \mathbf{a}$$

for the minimum eigenvalue and eigenvector. Thus, the frequencies of the complex exponentials correspond to the roots of the polynomial that is formed from the minimum eigenvector

$$V_{\min}(z) = \sum_{k=0}^{p} v_{\min}(k) z^{-k}$$

and σ_w^2 corresponds to the minimum eigenvalue.

(d) A random process consists of single sinusoid in nonwhite noise,

$$x(n) = A\sin(n\omega_0 + \phi) + w(n)$$

The first three values of the autocorrelation sequence for $x(n)$ are

$$\mathbf{r}_x = [9.515, 7.758, 6.472]^T$$

It is known that the additive noise $w(n)$ is a moving average process that is generated by filtering white noise $v(n)$

$$w(n) = v(n) + 0.1v(n-1)$$

However, the variance of $v(n)$ is unknown. Find the frequency ω_0 and the power, $P = \frac{1}{2}A^2$, of the sinusoid.

Solution _____

(a) We want to minimize the power in $y(n)$,

$$\xi = E\{y^2(n)\} = \mathbf{a}^H \mathbf{R}_x \mathbf{a}$$

subject to the constraint that
$$\mathbf{a}^H \mathbf{a} = 1$$

With
$$\mathbf{R}_x = \mathbf{E}\mathbf{P}\mathbf{E}^H + \sigma_w^2 \mathbf{I}$$

where

$$\mathbf{E} = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ e^{j\omega_1} & e^{j\omega_2} & \cdots & e^{j\omega_p} \\ \vdots & \vdots & & \vdots \\ e^{jp\omega_1} & e^{jp\omega_2} & \cdots & e^{jp\omega_p} \end{bmatrix}$$

we have

$$\xi = \mathbf{a}^H \mathbf{E}\mathbf{P}\mathbf{E}^H \mathbf{a} + \sigma_w^2 \mathbf{a}^H \mathbf{a}$$

With the constraint $\mathbf{a}^H \mathbf{a} = 1$, this becomes

$$\xi = \sigma_w^2 + \sum_{i=1}^{p} P_i \left| A(e^{j\omega_i}) \right|^2$$

where

$$A(e^{j\omega_i}) = \mathbf{e}_i^H \mathbf{a} = \sum_{k=0}^{p} a(k) e^{-jk\omega_i}$$

Therefore, ξ will be minimized when

$$A(e^{j\omega_i}) = 0 \quad ; \quad i = 1, 2, \ldots, p$$

i.e., when $A(z)$ has p zeros on the unit circle at the frequencies ω_i corresponding to the complex exponentials in $x(n)$, and the minimum value of ξ is equal to σ_w^2.

(b) If the autocorrelation matrix of the noise is $\sigma_w^2 \mathbf{Q}$, then

$$\mathbf{R}_x = \mathbf{EPE}^H + \sigma_w^2 \mathbf{Q}$$

If $x(n)$ is filtered with a pth-order FIR filter $\mathbf{a} = \begin{bmatrix} a(0), & a(1), & \ldots, & a(p) \end{bmatrix}^T$, then the output power is

$$\xi = \mathbf{a}^H \mathbf{R}_x \mathbf{a} = \mathbf{a}^H \mathbf{EPE}^H \mathbf{a} + \sigma_w^2 \mathbf{a}^H \mathbf{Q} \mathbf{a}$$

If $\mathbf{a}$ is normalized so that

$$\mathbf{a}^H \mathbf{Q} \mathbf{a} = 1$$

then the power becomes

$$\xi = \mathbf{a}^H \mathbf{EPE}^H \mathbf{a} + \sigma_w^2$$

As in part (a), this is minimized when

$$A(e^{j\omega_i}) = 0 \quad ; \quad i = 1, 2, \ldots, p$$

i.e., $A(z)$ has p zeros on the unit circle at the frequencies of the complex exponentials.

(c) Note that

$$\mathbf{R}_x \mathbf{a} = \mathbf{EPE}^H \mathbf{a} + \sigma_w^2 \mathbf{Q} \mathbf{a}$$

For the vector $\mathbf{a}$ that minimizes ξ,

$$\mathbf{E}^H \mathbf{a} = \mathbf{0}$$

Therefore, for this vector,

$$\mathbf{R}_x \mathbf{a} = \sigma_w^2 \mathbf{Q} \mathbf{a}$$

Thus, $\mathbf{a}$ is an eigenvector of $\mathbf{R}_x$, and σ_w^2 is the eigenvalue. Furthermore, σ_w^2 is the *minimum* eigenvalue of $\mathbf{R}_x$. To show this, note that

$$\xi = \mathbf{a}^H \mathbf{R}_x \mathbf{a} = \mathbf{a}^H \mathbf{EPE}^H \mathbf{a} + \sigma_w^2 \mathbf{a}^H \mathbf{Q} \mathbf{a} = \sigma_w^2$$

Therefore, to minimize ξ, we want to minimize σ_w^2, and it follows that σ_w^2 will be the minimum eigenvalue of $\mathbf{R}_x$.

(d) We are given

$$x(n) = A \sin(n\omega_0 + \phi) + w(n)$$

where

$$w(n) = v(n) + 0.1v(n-1)$$

with $v(n)$ white noise that has a variance σ_v^2. Therefore,

$$P_w(e^{j\omega}) = \sigma_w^2 |1 + 0.1e^{-j\omega}|^2 = \sigma_w^2 (1.01 + 0.2 \cos \omega)$$

and

$$r_w(k) = 1.01\sigma_w^2 \delta(k) + 0.1\sigma_w^2 \delta(k-1) + 0.1\sigma_w^2 \delta(k+1)$$

Thus,

$$\mathbf{R}_w = \sigma_w^2 \mathbf{Q}$$

where

$$\mathbf{Q} = \begin{bmatrix} 1.01 & 0.2 & 0 \\ 0.2 & 1.01 & 0.2 \\ 0 & 0.2 & 1.01 \end{bmatrix}$$

The frequency of the sinusoid may then be found by solving the generalized eigenvalue problem

$$\mathbf{R}_x \mathbf{a} = \lambda \mathbf{Q} \mathbf{a}$$

or

$$\begin{bmatrix} 9.575 & 7.758 & 6.472 \\ 7.758 & 9.575 & 7.758 \\ 6.472 & 7.758 & 9.575 \end{bmatrix} \begin{bmatrix} a(0) \\ a(1) \\ a(2) \end{bmatrix} = \begin{bmatrix} 1.01 & 0.1 & 0 \\ 0.1 & 1.01 & 0.1 \\ 0 & 0.1 & 1.01 \end{bmatrix} \begin{bmatrix} a(0) \\ a(1) \\ a(2) \end{bmatrix}$$

Multiplying both sides of the equation by the inverse of $\mathbf{Q}$ we have

$$\begin{bmatrix} 8.8633 & 6.8774 & 5.7911 \\ 6.8774 & 8.8633 & 6.8774 \\ 5.7911 & 6.8774 & 8.8633 \end{bmatrix} \begin{bmatrix} a(0) \\ a(1) \\ a(2) \end{bmatrix} = \begin{bmatrix} a(0) \\ a(1) \\ a(2) \end{bmatrix}$$

The minimum eigenvalue of the matrix $\mathbf{Q}^{-1}\mathbf{R}_x$ is

$$\lambda_{\min} = 1.5692$$

and the corresonding eigenvector is

$$\mathbf{v}_{\min} = \begin{bmatrix} 1, & -1.9026, & 1 \end{bmatrix}^T$$

The roots of the eigenfilter

$$A_{\min}(z) = 1 - 1.9026z^{-1} + z^{-2}$$

are

$$z = e^{\pm j0.1\pi}$$

Therefore, $\omega_0 = 0.1\pi$ and $P = \frac{1}{2}A^2 = 1.5692$.

8.25 A random process is known to consist of a single sinusoid in white noise

$$x(n) = A\sin(n\omega_0 + \phi) + w(n)$$

where the variance of $w(n)$ is σ_w^2.

(a) Suppose that the first three values of the autocorrelation sequence are estimated and found to be

$$r_x(0) = 1 \quad ; \quad r_x(1) = \beta \quad ; \quad r_x(2) = 0$$

Find and prepare a carefully labeled sketch of the spectrum estimate that is formed using the Blackman-Tukey method with a rectangular window.

(b) With the autocorrelation given in part (a), use the Pisarenko harmonic decomposition to estimate the variance of the white noise, σ_w^2, the frequency of the sinusoid, ω_0, and the sinusoid power, $P = \frac{1}{2}A^2$. How does your estimate of the white noise power and the sinusoid frequency depend upon β? Does the sinusoid power depend upon β?

(c) Suppose that we compute the periodogram $\widehat{P}_{per}(e^{j\omega})$ using N samples of $x(n)$. Find and prepare a carefully labeled sketch of the expected value of this spectral estimate. Is this estimate biased? Is it consistent?

(d) Using the autocorrelation estimates given in part (a), find the second-order MEM power spectrum.

Solution _____

(a) Using the Blakman-Tukey method with a rectangular window, the spectrum estimate is

$$\widehat{P}_x(e^{j\omega}) = \sum_{k=-M}^{M} r_x(k)e^{-j\omega} = 1 + 2\beta\cos\omega$$

(b) With a single sinusoid, the white noise power estimate using the Pisarenko harmonic decomposition is equal to the minimum eigenvalue of the 3×3 Toeplitz autocorrelation matrix

$$\mathbf{R}_x = \text{Toep}\{1, \beta, 0\}$$

The eigenvalues are the roots of the polyomial

$$
\det(\mathbf{R}_x) = \det\begin{bmatrix} 1-\lambda & \beta & 0 \\ \beta & 1-\lambda & \beta \\ 0 & \beta & 1-\lambda \end{bmatrix} = (1-\lambda)\left[(1-\lambda)^2 - \beta^2\right] - \beta\left[\beta(1-\lambda)\right]
$$

$$
= (1-\lambda)\left[\lambda^2 - 2\lambda + 1 - 2\beta^2\right]
$$

Therefore, the eigenvalues are

$$\lambda_1 = 1 \quad ; \quad \lambda_2 = 1 + \beta\sqrt{2} \quad ; \quad \lambda_2 = 1 - \beta\sqrt{2}$$

and the minimum eigenvalues is

$$\lambda_{\min} = 1 - |\beta|\sqrt{2}$$

The frequency of the sinusoid is found from the eigenvector corresponding to the minimum eigenvalue which, for $\beta > 0$, is

$$\mathbf{v}_{\min} = \tfrac{1}{2}\big[1, \ -\sqrt{2}, \ 1\big]^T$$

Thus,

$$V_{\min}(z) = \tfrac{1}{2}\big(1 - \sqrt{2}z^{-1} + z^{-2}\big)$$

which has roots at

$$z = \tfrac{\sqrt{2}}{2}\big(1 \pm j\big)$$

which corrsponds to a frequency of

$$\omega_0 = \pi/4$$

If, on the other hand, $\beta < 0$, then the eigenfilter is

$$V_{\min}(z) = \tfrac{1}{2}\big(1 + \sqrt{2}z^{-1} + z^{-2}\big)$$

which has roots at

$$z = \tfrac{\sqrt{2}}{2}\big(-1 \pm j\big)$$

and the frequency of the sinusoid is

$$\omega_0 = 3\pi/4$$

To determine the sinusoid power, note that the autocorrelation sequence is

$$r_x(k) = P\cos k\omega_0 + \sigma_w^2$$

With $\beta > 0$ this becomes

$$r_x(k) = P\cos\frac{k\pi}{4} + (1 - \beta\sqrt{2})$$

Since $r_x(0) = 1$, it follows that

$$P = \beta\sqrt{2}$$

If, on the other hand, $\beta < 0$, then we have

$$P = -\beta\sqrt{2}$$

Therefore, we may write

$$P = |\beta|\sqrt{2}$$

(c) If we compute the periodogram $\widehat{P}_{per}(e^{j\omega})$ using N samples of $x(n)$, then the expected value of this estimate is

$$E\left\{\widehat{P}_{per}(e^{j\omega})\right\} = \frac{1}{2\pi}P_x(e^{j\omega}) * W_B(e^{j\omega})$$

where $W_B(e^{j\omega})$ is the Fourier transform of the Bartlett window, $w_B(k)$,

$$W_B(e^{j\omega}) = \frac{1}{N}\left[\frac{\sin(N\omega/2)}{\sin(\omega/2)}\right]^2$$

and $P_x(e^{j\omega})$ is the power spectrum of $x(n)$. Since $x(n)$ is a random phase sinusoid in white noise, then

$$P_x(e^{j\omega}) = \frac{1}{2}\pi A^2\Big[u_0(\omega - \omega_0) + u_0(\omega + \omega_0)\Big] + \sigma_w^2$$

Therefore, the expected value of the periodogram is

$$
\begin{aligned}
E\left\{\widehat{P}_{per}(e^{j\omega})\right\} &= \frac{1}{2\pi}P_x(e^{j\omega}) * W_B(e^{j\omega}) \\
&= \sigma_w^2 + \frac{1}{4}A^2\left[W_B(e^{j(\omega-\omega_0)}) + W_B(e^{j(\omega+\omega_0)})\right]
\end{aligned}
$$

where $\sigma_w^2 = 1 - |\beta|\sqrt{2}$, and, depending upon the sign of β, the frequency ω_0 is either equal to $\pi/4$ or $3\pi/4$. As we have seen in Chapter 8, the periodogram is a biased estimate of the periodogram, and, since the variance does not go to zero as N goes to infinity, it is not a consistent estimate of the power spectrum.

(d) For a second-order MEM power spectrum, we first must solve the normal equations

$$
\begin{bmatrix} 1 & \beta \\ \beta & 1 \end{bmatrix}\begin{bmatrix} a(1) \\ a(2) \end{bmatrix} = -\begin{bmatrix} \beta \\ 0 \end{bmatrix}
$$

Thus, for the second-order all-pole model we have

$$
\begin{bmatrix} a(1) \\ a(2) \end{bmatrix} = -\frac{1}{1-\beta^2}\begin{bmatrix} 1 \\ -\beta \end{bmatrix}
$$

and

$$
\epsilon_2 = r_x(0) + a(1)r_x(1) + a(2)r_x(2) = \frac{1}{1-\beta^2}
$$

Therefore, the MEM spectrum is

$$
\widehat{P}_{mem}(e^{j\omega}) = \frac{\epsilon_2}{\left|1 + a(1)e^{-j\omega} + a(2)e^{-2j\omega}\right|^2}
$$

Incorporating the values for $a(1)$, $a(2)$, and ϵ_2 and simplifying this becomes

$$
\widehat{P}_{mem}(e^{j\omega}) = \frac{1-\beta^2}{(1-\beta^2+2\beta^4) + \beta^2(1-\beta^2)\cos\omega - 2\beta\cos 2\omega}
$$

8.26 A random process may be classified in terms of the properties of the prediction error sequence ϵ_k that is produced when fitting an all-pole model to the process. Listed below are five different classifications for the error sequence:

1. $\epsilon_k = c > 0$ for all $k \geq 0$.

2. $\epsilon_k = c > 0$ for all $k \geq k_0$ for some $k_0 > 0$.

3. $\epsilon_k \rightarrow c$ as $k \rightarrow \infty$ where $c > 0$.

4. $\epsilon_k \rightarrow 0$ as $k \rightarrow \infty$.

5. $\epsilon_k = 0$ for all $k \geq k_0$ for some $k_0 > 0$.

For each of these classifications, describe as completely as possible the characteristics that may be attributed to the process and its power spectrum.

Solution _____

1. If $\epsilon_k = c > 0$ for all $k \geq 0$, then the reflection coefficients are zero, $\Gamma_k = 0$ for all k. Therefore, the process is not predictable, which is consistent with a white noise process with a power spectrum that is constant, $P_x(e^{j\omega}) = c$.

2. If $\epsilon_k = c > 0$ for all $k \geq k_0$ for some $k_0 > 0$, then the reflection coefficients Γ_k are equal to zero for all $k \geq k_0$. This is consistent with a finite-order autoregressive process of order k_0. Only a finite past history is used to model the process.

3. If $\epsilon_k \rightarrow c$ as $k \rightarrow \infty$ where $c > 0$, then the entire past history of the process is useful in predicting (modeling) the process. Since ϵ_k approaches a limit that is not zero, this behavior is consistent with a moving average or autoregressive moving average process.

4. If $\epsilon_k \rightarrow 0$ as $k \rightarrow \infty$, the process is perfectly predictable given the infinite past history of the process. Recall that the prediction error, expressed in terms of the power spectrum, is (see Eq. (7.72), the Kolmogorov-Szegö formula)

$$\lim_{k \rightarrow \infty} \epsilon_k = \exp\left\{ \frac{1}{2\pi} \int_{-\pi}^{\pi} \ln P_x(e^{j\omega}) d\omega \right\}$$

This will be zero only if $P_x(e^{j\omega}) = 0$ over some finite interval. This behavior is consistent with a lowpass, bandpass, or highpass process in which $P_x(e^{j\omega}) = 0$ for $\omega \in [\omega_1, \omega_2]$.

5. If $\epsilon_k = 0$ for all $k \geq k_0$ for some $k_0 > 0$, then the process is perfectly predictable from the past history of $x(n)$ that is finite in length. This corresponds to the case in which $\Gamma_{k_0} = \pm 1$ and, therefore, represents a harmonic process (line spectrum).

8.27 In the MUSIC algorithm, finding the peaks of the frequency estimation function

$$\widehat{P}_{MU}(e^{j\omega}) = \frac{1}{\displaystyle\sum_{i=p+1}^{M}\left|\mathbf{e}^{H}\mathbf{v}_{i}\right|^{2}}$$

is equivalent to finding the minima of the denominator. Show that finding the *minima* of the denominator is equivalent to finding the *maxima* of

$$\sum_{i=1}^{p}|\mathbf{e}^{H}\mathbf{v}_{i}|^{2}$$

Hint: Use the fact that

$$\mathbf{I} = \sum_{i=1}^{M}\mathbf{v}_{i}\mathbf{v}_{i}^{H}$$

Solution ───

Finding the peaks of the MUSIC frequency estimation function is equivalent to finding the minima of

$$\sum_{i=p+1}^{M}\left|\mathbf{e}^{H}\mathbf{v}_{i}\right|^{2}$$

Since the eigenvectors $\mathbf{v}_{i}$ are orthogonal, if we assume that they have been normalized, then the identity matrix may be expanded in terms of these eigenvectors as follows,

$$\mathbf{I} = \sum_{i=1}^{M}\mathbf{v}_{i}\mathbf{v}_{i}^{H}$$

Mutliplying on the left by $\mathbf{e}^{H}$ and on the right by $\mathbf{e}$, we have

$$\mathbf{e}^{H}\mathbf{e} = \sum_{i=1}^{M}(\mathbf{e}^{H}\mathbf{v}_{i})(\mathbf{v}_{i}^{H}\mathbf{e}) = \sum_{i=1}^{M}|\mathbf{e}^{H}\mathbf{v}_{i}|^{2}$$

Since $\mathbf{e}^{H}\mathbf{e} = M$, then

$$M = \sum_{i=1}^{p}|\mathbf{e}^{H}\mathbf{v}_{i}|^{2} + \sum_{i=p+1}^{M}|\mathbf{e}^{H}\mathbf{v}_{i}|^{2}$$

or

$$\sum_{i=p+1}^{M}|\mathbf{e}^{H}\mathbf{v}_{i}|^{2} = M - \sum_{i=1}^{p}|\mathbf{e}^{H}\mathbf{v}_{i}|^{2}$$

Thus, minimizing the left-hand side is equivalent to maximizing the sum on the right as was to be shown.

8.28 The 3×3 autocorrelation matrix of a harmonic process is

$$\mathbf{R}_x = \begin{bmatrix} 3 & -j & -1 \\ j & 3 & -j \\ -1 & j & 3 \end{bmatrix}$$

(a) Using the Pisarenko harmonic decomposition, find the complex exponential frequencies and the variance of the white noise.

(b) Repeat part (a) using the MUSIC algorithm, the eigenvector method, and the minimum norm method.

Solution _____

(a) To find the complex exponential frequencies and white noise variance using the Pisarenko decomposition, we first must find the eigenvalues of the autocorrelation matrix $\mathbf{R}_x$, which are the roots of the characteristic equation

$$\det(\mathbf{R}_x - \lambda \mathbf{I}) = 0$$

Note that if we subtract two from the terms along the diagonal, then the matrix is singular. Therefore, $\lambda = 0$ is an eigenvalue of the matrix

$$\mathbf{Q}_x = \mathbf{R}_x - 2\mathbf{I}$$

The remaining two eigenvalues of $\mathbf{Q}_x$ are the roots of the the characteristic equation for $\mathbf{Q}_x$,

$$\det(\mathbf{Q}_x - \lambda \mathbf{I}) = \det \begin{bmatrix} 1-\lambda & -j & -1 \\ j & 1-\lambda & -j \\ -1 & j & 1-\lambda \end{bmatrix} = 0$$

or,

$$(1-\lambda)\big[(1-\lambda)^2 - 1\big] + j\big[j(1-\lambda) - j\big] - \big[-1 + (1-\lambda)\big] = 0$$

Multiplying out the terms and simplifying, we have

$$\lambda^2(3 - \lambda) = 0$$

Therefore, the eigenvalues of $\mathbf{Q}_x$ are

$$\widetilde{\lambda_1} = 3 \quad ; \quad \widetilde{\lambda_2} = 0 \quad ; \quad \widetilde{\lambda_3} = 0$$

and the eigenvalues of $\mathbf{R}_x$ are $\lambda_i = \widetilde{\lambda}_i + 2$, or

$$\lambda_1 = 5 \quad ; \quad \lambda_2 = 2 \quad ; \quad \lambda_3 = 2$$

The eigenvectors of $\mathbf{R}_x$ are

$$\mathbf{v}_1 = \frac{1}{\sqrt{3}} \begin{bmatrix} -j \\ 1 \\ j \end{bmatrix} \quad ; \quad \mathbf{v}_2 = \frac{1}{\sqrt{6}} \begin{bmatrix} 2 \\ -j \\ 1 \end{bmatrix} \quad ; \quad \mathbf{v}_3 = \frac{1}{\sqrt{2}} \begin{bmatrix} 0 \\ j \\ 1 \end{bmatrix}$$

Note that since the dimension of the noise subspace is equal to two, then there is only one complex exponential, and the frequency of the complex exponential may be found from the eigenvector corresponding to the zero eigenvalue of the 2×2 matrix

$$\mathbf{R}_x - 2\mathbf{I} = \begin{bmatrix} 1 & -j \\ j & 1 \end{bmatrix}$$

This eigenvector is

$$\mathbf{v}_{\min} = \begin{bmatrix} 1 \\ -j \end{bmatrix}$$

Thus, the eigenfilter is

$$A(z) = 1 - jz^{-1}$$

and the complex exponential frequency is $\omega_0 = -\pi/2$.

(b) For the MUSIC algorithm,

$$\widehat{P}_{MU}(e^{j\omega}) = \frac{1}{\displaystyle\sum_{i=p+1}^{M} \left| \mathbf{e}^H \mathbf{v}_i \right|^2}$$

So, with $M = 3$ and $p = 1$ this becomes

$$\widehat{P}_{MU}^{-1}(e^{j\omega}) = \left| \mathbf{e}^H \mathbf{v}_2 \right|^2 + \left| \mathbf{e}^H \mathbf{v}_3 \right|^2$$

Using the eigenvectors found in part (a) we have

$$\begin{aligned}
\widehat{P}_{MU}^{-1}(e^{j\omega}) &= \frac{1}{6} \left| 2 - je^{-j\omega} + e^{-2j\omega} \right|^2 + \frac{1}{2} \left| je^{-j\omega} + e^{-2j\omega} \right|^2 \\
&= \frac{1}{6} \left[6 - 2\sin\omega + 4\cos 2\omega \right] + \frac{1}{2} \left[2 - 2\sin\omega \right] \\
&= 2 - \frac{4}{3} \sin\omega + \frac{2}{3} \cos 2\omega
\end{aligned}$$

and, therefore,

$$\widehat{P}_{MU}(e^{j\omega}) = \frac{1}{2 - \frac{4}{3}\sin\omega + \frac{2}{3}\cos 2\omega}$$

For the eigenvector method, the frequency estimate is

$$\widehat{P}_{EV}(e^{j\omega}) = \frac{1}{\displaystyle\sum_{i=p+1}^{M} \frac{1}{\lambda_i} \left| \mathbf{e}^H \mathbf{v}_i \right|^2}$$

So, with $p = 1$, $M = 3$, and $\lambda_2 = \lambda_3 = 2$, this becomes

$$\widehat{P}_{EV}^{-1}(e^{j\omega}) = \frac{1}{2} \left| \mathbf{e}^H \mathbf{v}_2 \right|^2 + \frac{1}{2} \left| \mathbf{e}^H \mathbf{v}_3 \right|^2$$

which is the same as the MUSIC estimate to within a scale factor of two.
Finally, for the minimum norm method, we have

$$\widehat{P}_{MN}(e^{j\omega}) = \frac{1}{|\mathbf{e}^H \mathbf{a}|^2}$$

where
$$\mathbf{a} = \lambda \mathbf{P}_n \mathbf{u}_1$$

Forming the projection matrix $\mathbf{P}_n$ we have

$$\mathbf{P}_n = \begin{bmatrix} \mathbf{v}_2, \mathbf{v}_3 \end{bmatrix} \begin{bmatrix} \mathbf{v}_2^{\mathbf{H}} \\ \mathbf{v}_3^{H} \end{bmatrix} = \begin{bmatrix} 2/\sqrt{6} & 0 \\ -j/\sqrt{6} & j/\sqrt{2} \\ 1/\sqrt{6} & 1/\sqrt{2} \end{bmatrix} \begin{bmatrix} 2/\sqrt{6} & -j/\sqrt{6} & 1/\sqrt{6} \\ 0 & j/\sqrt{2} & 1/\sqrt{2} \end{bmatrix}$$

$$= \begin{bmatrix} 2/3 & -j/3 & 1/3 \\ -j/3 & -2/3 & j/3 \\ 1/3 & j/3 & 2/3 \end{bmatrix}$$

With

$$\lambda = \frac{1}{\mathbf{u}_1^T \mathbf{P}_n \mathbf{u}_1} = 3/2$$

and

$$\mathbf{a} = \lambda \mathbf{P}_n \mathbf{u}_1 = \frac{3}{2} \begin{bmatrix} 2/3 \\ -j/3 \\ 1/3 \end{bmatrix}$$

we see that the minimum norm estimate is

$$\widehat{P}_{MV}(e^{j\omega}) = \frac{1}{\frac{3}{2} - \frac{1}{2}\sin\omega + \cos 2\omega}$$

8.29 In this problem we prove that the spurious roots in minimum norm method lie inside the unit circle. Let $x(n)$ be a random process that is a sum of p complex exponentials in white noise, and let $\mathbf{a}$ be an M-dimensional vector that lies in the noise subspace. The z-transform of $\mathbf{a}$ may be factored as follows

$$A(z) = A_0(z)A_1(z)$$

where

$$A_0(z) = \prod_{k=1}^{p}(1 - e^{j\omega_k}z^{-1})$$

is a monic polynomial that has p roots on the unit circle at the frequencies of the complex exponentials in $x(n)$, and $A_1(z)$ is a polynomial that contains the $M - p - 1$ spurious roots.

(a) Show that minimizing $\|\mathbf{a}\|^2$ is equivalent to minimizing

$$\frac{1}{2\pi}\int_{-\pi}^{\pi}|A(e^{j\omega})|^2 d\omega = \frac{1}{2\pi}\int_{-\pi}^{\pi}|A_0(e^{j\omega})|^2|A_1(e^{j\omega})|^2 d\omega$$

where $A_0(e^{j\omega})$ is fixed and $A_1(e^{j\omega})$ is monic.

(b) From the results of part (a), show that minimizing $\|\mathbf{a}\|^2$ is thus equivalent to using the autocorrelation method to find the prediction error filter $A_1(z)$ for the signal whose z-transform is $A_0(z)$.

(c) From (b), argue that $A_1(z)$ must therefore have all of its roots inside the unit circle.

Solution _____

(a) If $\mathbf{a}$ is a vector that lies in the noise subspace, and

$$A(z) = 1 + \sum_{k=1}^{M-1}a(k)z^{-1}$$

then $A(z)$ has p zeros on the unit circle at the frequencies of the complex exponentials in $x(n)$. Therefore, $A(z)$ may be factored as

$$A(z) = A_0(z)A_1(z)$$

where $A_0(z)$ is a fixed polynomial of the form

$$A_0(z) = \prod_{k=1}^{p}(1 - e^{j\omega_k}z^{-1})$$

By Parseval's theorem,

$$\|\mathbf{a}\|^2 = 1 + \sum_{k=1}^{M-1}|a(k)|^2 = \frac{1}{2\pi}\int_{-\pi}^{\pi}|A(e^{j\omega})|^2 d\omega$$

Thus, it follows that minimizing $\|\mathbf{a}\|^2$ is equivalent to minimizing

$$\frac{1}{2\pi}\int_{-\pi}^{\pi}|A(e^{j\omega})|^2 d\omega = \frac{1}{2\pi}\int_{-\pi}^{\pi}|A_0(e^{j\omega})|^2|A_1(e^{j\omega})|^2 d\omega \qquad \text{(P8.29-1)}$$

(b) Recall from Chapter 4 that the autocorrelation method minimizes the error

$$\mathcal{E} = \sum_{n=0}^{\infty} e^2(n)$$

where

$$e(n) = x(n) + \sum_{k=1}^{p} a_p(k)x(n-k) = x(n) * a_p(n)$$

From Parseval's theorem, $\mathcal{E}$ may be written as

$$\mathcal{E} = \frac{1}{2\pi} \int_{-\pi}^{\pi} |E(e^{j\omega})|^2 d\omega = \frac{1}{2\pi} \int_{-\pi}^{\pi} |X(e^{j\omega})A_p(e^{j\omega})|^2 d\omega$$

Therefore, the autocorrelation method involves the minimization of

$$\mathcal{E} = \frac{1}{2\pi} \int_{-\pi}^{\pi} |X(e^{j\omega})A_p(e^{j\omega})|^2 d\omega$$

where $X(e^{j\omega})$ is fixed. If we replace $X(e^{j\omega})$ and $\mathbf{A}_p$ with $A_0(e^{j\omega})$ and $A_1(e^{j\omega})$, respectively, we see that this is equivalent to the problem of minimizing the integral in Eq. (P8.29-1).

(c) Since the autocorrelation method produces a filter, $A_p(z)$ with all of its zeros inside the unit circle, then minimizing Eq. (P8.29-1) will produce a polynomial $A(z)$ with all of its roots inside the unit circle.

8.30 In the minimum norm method, the spurious zeros in the polynomial $A(z)$ are separated from those that lie on the unit circle by forcing the spurious roots to lie inside the unit circle. In some applications of eigenvector methods, such as system identification, some of the desired zeros may lie inside the unit circle. In this case, the desired roots cannot be distinguished from the spurious roots. The minimum norm method may be modified, however, to force the spurious zeros to lie *outside* the unit circle. This is done by constraining the *last* element of the vector $\mathbf{a}$ to have a value of one, i.e.,

$$\mathbf{a}^H \mathbf{u}_M = 1$$

where $\mathbf{u}_M = [0, 0, \ldots, 0, 1]^T$ is a unit vector with the last element equal to one.

(a) Derive the *modified minimum norm* algorithm that uses the constraint that $\mathbf{a}^H \mathbf{u}_M = 1$ instead of $\mathbf{a}^H \mathbf{u}_1 = 1$ as in the minimum norm algorithm.

(b) The 3×3 autocorrelation matrix for a single complex exponential in white noise is

$$\mathbf{R}_x = \begin{bmatrix} 2 & 1-j & -j\sqrt{2} \\ 1+j & 2 & 1-j \\ j\sqrt{2} & 1+j & 2 \end{bmatrix}$$

Find the frequency of the complex exponential and the locations of the spurious roots in the minimum norm frequency estimation function.

(c) Repeat part (b) for the modified minimum norm method.

Solution _____

(a) The derivation of the *modified minimum norm* algorithm is straightforward. With the constraint that $\mathbf{a}^H \mathbf{u}_M = 1$ in place of $\mathbf{a}^H \mathbf{u}_1 = 1$, all that we need to do is replace $\mathbf{u}_1$ in the minimum norm algorithm with $\mathbf{u}_M$. Thus, the modified minimum norm algorithm is

$$\mathbf{a} = \lambda \mathbf{P}_n \mathbf{u}_M$$

where

$$\lambda = \frac{1}{\mathbf{u}_M^H \mathbf{P}_n \mathbf{u}_M}$$

(b) The eigenvalues of the autocorrelation matrix

$$\mathbf{R}_x = \begin{bmatrix} 2 & 1-j & -j\sqrt{2} \\ 1+j & 2 & 1-j \\ j\sqrt{2} & 1+j & 2 \end{bmatrix}$$

are

$$\lambda_1 = \lambda_2 = 2 - \sqrt{2} \quad ; \quad \lambda_3 = 2 + 2\sqrt{2}$$

The normalized noise eigenvectors corresponding to the eigenvalues λ_1 and λ_2 are

$$\mathbf{v}_1 = \frac{1}{2} \begin{bmatrix} 1 \\ -\sqrt{2} \\ 1 \end{bmatrix} \quad ; \quad \mathbf{v}_2 = \frac{1}{2} \begin{bmatrix} 1 \\ j\sqrt{2} \\ -1 \end{bmatrix}$$

Therefore, the projection matrix is

$$\mathbf{P}_n = \frac{1}{4} \begin{bmatrix} 1 & 1 \\ -\sqrt{2} & j\sqrt{2} \\ 1 & -1 \end{bmatrix} \begin{bmatrix} 1 & -j\sqrt{2} & 1 \\ 1 & -j\sqrt{2} & -1 \end{bmatrix} = \frac{1}{4} \begin{bmatrix} 2 & \sqrt{2}(-1-j) & 0 \\ \sqrt{2}(-1+j) & 4 & \sqrt{2}(-1-j) \\ 0 & \sqrt{2}(-1+j) & 2 \end{bmatrix}$$

Thus,

$$\lambda = \frac{1}{\mathbf{u}_M^H \mathbf{P}_n \mathbf{u}_M} = \frac{1}{(\frac{1}{4})2} = 2$$

and

$$\mathbf{a} = \lambda \mathbf{P}_n \mathbf{u}_M = 2(\tfrac{1}{4}) \begin{bmatrix} 2 \\ \sqrt{2}(-1-j) \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ -\frac{1}{2}\sqrt{2}(1+j) \\ 0 \end{bmatrix}$$

Finally, we have

$$\widehat{P}_{MN}(e^{j\omega}) = \frac{1}{|\mathbf{e}^H \mathbf{a}|^2} = \frac{1}{2 + \sqrt{2}\sin\omega - \sqrt{2}\cos\omega}$$

(c) For the modified minimum norm method, the vector **a** is reversed. Thus,

$$\mathbf{a} = \lambda \mathbf{P}_n \mathbf{u}_M = 2(\tfrac{1}{4}) \begin{bmatrix} 0 \\ \sqrt{2}(-1-j) \\ 2 \end{bmatrix}$$

Clearly, with $A(z)$ in part (b) having its roots inside the unit circle, by reversing the vector, with the modified minimum norm algorithm the roots will be outside the unit circle.

SOLUTIONS TO CHAPTER 9
Adaptive Filtering

9.1 In order for the steepest descent algorithm to converge, the step size must be in the range

$$0 < \mu < 2/\lambda_{\max}$$

However, in some cases it may be of interest to find the value of μ that gives the fastest rate of convergence. For a given step size μ, the rate of convergence for the weight vector $\mathbf{w}_n$ is dominated by the slowest converging mode in the expansion

$$\mathbf{w}_n = \mathbf{w} + \sum_{k=0}^{p}(1 - \mu\lambda_k)^n u_0(k)\mathbf{v}_k$$

(a) In terms of the eigenvalues λ_k, find the value for μ that maximizes the rate of convergence. In other words, find the value for μ that maximizes the rate of convergence of the slowest decaying mode.

(b) At what rate does the slowest mode decrease for the step size found in part (a)?

Solution _____

(a) We have seen that the dynamic behavior of the weight vector, $\mathbf{w}_n$, in the steepest descent algorithm behaves as

$$\mathbf{w}_n = \mathbf{w} + \sum_{k=1}^{p}(1 - \mu\lambda_k)^n u_0(k)\mathbf{v}_k$$

To maximize the rate of convergence, we want to choose μ so that we minimize the slowest decaying mode, i.e.,

$$\min_{\mu}\left[\max_{k}|1 - \mu\lambda_k|\right]$$

Shown in the figure below is a plot of $|1 - \mu\lambda_k|$ as a function of μ for $\lambda_{\min}$ and $\lambda_{\max}$.
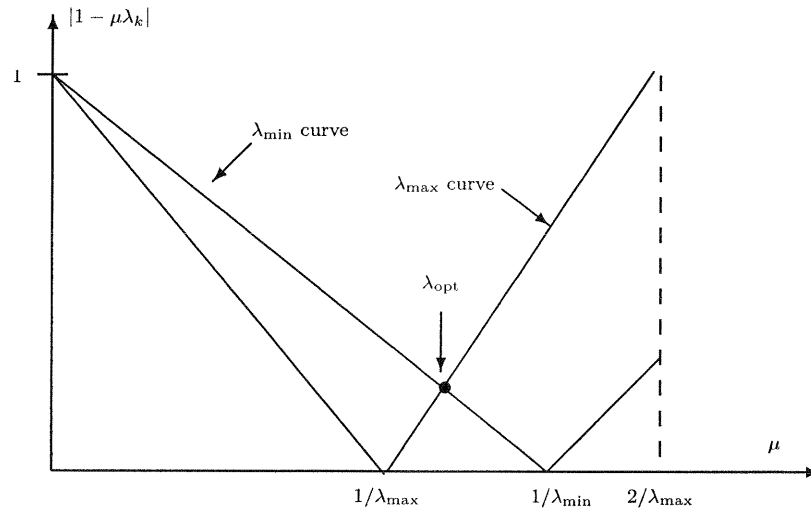
Clearly, for a given value of μ, the points $|1 - \mu\lambda_k|$ will lie between the two curves, $|1 - \mu\lambda_{\max}|$ and $|1 - \mu\lambda_{\min}|$. Therefore, to minimize the largest value of $|1 - \mu\lambda_k|$ we choose $\mu = \mu_{\mathrm{opt}}$ as shown, which is the solution to

$$1 - \mu\lambda_{\max} = -(1 - \mu\lambda_{\min})$$

or,

$$\mu_{\mathrm{opt}} = \frac{2}{\lambda_{\max} + \lambda_{\min}}$$

(b) The slowest decaying mode with μ_{opt} is the rate at which both the minimum and maximum eigenvalues converge, i.e.,

$$\left[1 - \frac{2\lambda_{\max}}{\lambda_{\max} + \lambda_{\min}}\right]^n = \left(\frac{\lambda_{\min} - \lambda_{\max}}{\lambda_{\max} + \lambda_{\min}}\right)^n$$

9.2 Suppose that the input to an adaptive linear predictor is white noise with an autocorrelation sequence $r_x(k) = \sigma_x^2 \delta(k)$.

(a) Solve the normal equations and find the optimum pth-order one-step linear predictor, $\mathbf{w}$.

(b) Minimize the mean-square prediction error using the method of steepest descent with a step size $\mu = 1/(5\sigma_x^2)$ and an initial weight vector $\mathbf{w}_0 = [1, 1, \ldots, 1]^T$. Does the method of steepest descent converge to the solution found in part (a)?

Solution ───

(a) With $\mathbf{R}_x = \sigma_x^2 \mathbf{I}$, and $\mathbf{r}_{dx} = \mathbf{0}$, the solution to the Wiener-Hopf equations is

$$\mathbf{w} = \mathbf{0}$$

(b) The steepest descent algorithm is

$$\mathbf{w}_{n+1} = \mathbf{w}_n - \mu\left[\mathbf{R}_x \mathbf{w}_n - \mathbf{r}_{dx}\right]$$

Since $\mathbf{R}_x = \sigma_x^2 \mathbf{I}$ and $\mathbf{r}_{dx} = \mathbf{0}$, then

$$\mathbf{w}_{n+1} = \left(1 - \mu\sigma_x^2\right)\mathbf{w}_n$$

With $\mu = 1/(5\sigma_x^2)$ the time-evolution of $\mathbf{w}_n$ becomes

$$\mathbf{w}_n = \left(1 - \tfrac{1}{5}\right)^n \mathbf{w}_0$$

which goes to zero as $n \to \infty$.

9.3 Newton's method is an iterative algorithm that may be used to find the minimum of a nonlinear function. Applied to the minimization of the mean-square error

$$\xi(n) = E\{e^2(n)\}$$

where $e(n) = d(n) - \mathbf{w}^T\mathbf{x}(n)$, Newton's method is

$$\mathbf{w}_{n+1} = \mathbf{w}_n - \tfrac{1}{2}\mathbf{R}_x^{-1}\nabla\xi(n)$$

where $\mathbf{R}_x$ is the autocorrelation matrix of $x(n)$. Introducing a step-size parameter μ, Newton's method becomes

$$\mathbf{w}_{n+1} = \mathbf{w}_n - \tfrac{1}{2}\mu\mathbf{R}_x^{-1}\nabla\xi(n)$$

Comparing this to the steepest descent algorithm, we see that the step size μ is replaced with a matrix, $\mu\mathbf{R}_x^{-1}$, which alters the descent direction.

(a) For what values of μ is Newton's method stable, i.e., for what values of μ will $\mathbf{w}_n$ converge?

(b) What is the optimum value of μ, i.e., for what value of μ is the convergence the fastest?

(c) Suppose that we form an LMS version of Newton's method by replacing the gradient with a gradient estimate

$$\widehat{\nabla}\xi(n) = \nabla e^2(n)$$

Derive the coefficient update equation that results from using this gradient estimate and describe how it differs from the LMS algorithm.

(d) Derive an expression that describes the time evolution of $E\{\mathbf{w}_n\}$ using the LMS Newton algorithm derived in part (c).

Solution _____

(a) Evaluating the gradient vector we have

$$\nabla\xi(n) = 2E\{e(n)\nabla e(n)\} = -2E\{e(n)\mathbf{x}(n)\} = -2\mathbf{r}_{dx} + 2\mathbf{R}_x\mathbf{w}_n$$

Thus,

$$\mathbf{w}_{n+1} = \mathbf{w}_n - \tfrac{1}{2}\mu\mathbf{R}_x^{-1}\big[2\mathbf{R}_x\mathbf{w}_n - 2\mathbf{r}_{dx}\big]$$

and we have

$$\mathbf{w}_{n+1} = \mathbf{w}_n - \mu\mathbf{w}_n + \mu\mathbf{w} = (1-\mu)\mathbf{w}_n + \mu\mathbf{w}$$

Thus, the Newton algorithm is stable for $0 < \mu < 2$.

(b) The convergence is the fastest when $\mu = 1$. Note, in fact, that when $\mu = 1$, the Newton iteration converges in one step to $\mathbf{w}$.

(c) The gradient approximation is

$$\nabla e^2(n) = -2e(n)\mathbf{x}(n)$$

Therefore the LMS-type algorithm is

$$\mathbf{w}_{n+1} = \mathbf{w}_n + \mu e(n)\mathbf{R}_x^{-1}\mathbf{x}(n)$$

Comparing this to the LMS algorithm we see that the step direction is changed from $\mathbf{x}(n)$ to $\mathbf{R}^{-1}\mathbf{x}(n)$.

(d) From (c) we see that

$$\mathbf{w}_{n+1} = \mathbf{w}_n + \mu \mathbf{R}_x^{-1} \mathbf{x}(n) d(n) - \mu \mathbf{R}_x^{-1} \mathbf{x}(n) \mathbf{x}^T(n) \mathbf{w}_n$$

Assuming that $\mathbf{x}(n)$ is uncorrelated with the filter tap weight vector, $\mathbf{w}_n$, then

$$E\{\mathbf{w}_{n+1}\} = E\{\mathbf{w}_n\} + \mu \mathbf{R}_x^{-1} \mathbf{r}_{dx} - \mu \mathbf{R}_x^{-1} \mathbf{R}_x E\{\mathbf{w}_n\}$$

which becomes

$$E\{\mathbf{w}_{n+1}\} = (1 - \mu) E\{\mathbf{w}_n\} + \mu \mathbf{w}$$

9.4 One way to derive the steepest descent algorithm for solving the normal equations $\mathbf{R}_x \mathbf{w} = \mathbf{r}_{dx}$ is to use a power series expansion for the inverse of $\mathbf{R}_x$. This expansion is

$$\mathbf{R}_x^{-1} = \mu \sum_{k=0}^{\infty} (\mathbf{I} - \mu \mathbf{R}_x)^k$$

where $\mathbf{I}$ is the identity matrix and μ is a positive constant. In order for this expansion to converge, $\mathbf{R}_x$ must be positive definite and the constant μ must lie in the range

$$0 < \mu < 2/\lambda_{\max}$$

where $\lambda_{\max}$ is the largest eigenvalue of $\mathbf{R}_x$.

(a) Let

$$\mathbf{R}_x^{-1}(n) = \mu \sum_{k=0}^{n} (\mathbf{I} - \mu \mathbf{R}_x)^k$$

be the nth-order approximation to $\mathbf{R}_x^{-1}$, and let

$$\mathbf{w}_n = \mathbf{R}_x^{-1}(n) \mathbf{r}_{dx}$$

be the nth-order approximation to the desired solution $\mathbf{w} = \mathbf{R}_x^{-1} \mathbf{r}_{dx}$. Express $\mathbf{R}_x^{-1}(n+1)$ in terms of $\mathbf{R}_x^{-1}(n)$, and show how this may be used to derive the steepest descent algorithm

$$\mathbf{w}_{n+1} = \mathbf{w}_n - \mu [\mathbf{R}_x \mathbf{w}_n - \mathbf{r}_{dx}]$$

(b) If the statistics of $x(n)$ are unknown, then $\mathbf{R}_x$ is unknown and the expansion for $\mathbf{R}_x^{-1}$ in part (a) cannot be evaluated. However, suppose that we approximate $\mathbf{R}_x = E\{\mathbf{x}(n)\mathbf{x}^T(n)\}$ at time n as follows

$$\widehat{\mathbf{R}}_x(n) = \mathbf{x}(n)\mathbf{x}^T(n)$$

and use, as the nth-order approximation to $\mathbf{R}_x^{-1}$,

$$\widehat{\mathbf{R}}_x^{-1}(n) = \mu \sum_{k=0}^{n} \left[\mathbf{I} - \mu \mathbf{x}(k)\mathbf{x}^T(k)\right]^k$$

Express $\widehat{\mathbf{R}}_x^{-1}(n+1)$ in terms of $\widehat{\mathbf{R}}_x^{-1}(n)$ and use this expression to derive a recursion for $\mathbf{w}_n$.

(c) Compare your recursion derived in part (b) to the LMS algorithm.

Solution _____

(a) Using the nth-order approximation to $\mathbf{R}_x^{-1}$,

$$\mathbf{R}_x^{-1}(n) = \mu \sum_{k=0}^{n} \left(\mathbf{I} - \mu\mathbf{R}_x\right)^k$$

we have

$$\mathbf{R}_x^{-1}(n+1) = \mu \sum_{k=0}^{n+1} \left(\mathbf{I} - \mu\mathbf{R}_x\right)^k = \mu\left(\mathbf{I} - \mu\mathbf{R}_x\right) \sum_{k=0}^{n} \left(\mathbf{I} - \mu\mathbf{R}_x\right)^k + \mu\mathbf{I}$$

Therefore,

$$\mathbf{R}_x^{-1}(n+1) = \left(\mathbf{I} - \mu\mathbf{R}_x\right)\mathbf{R}_x^{-1}(n) + \mu\mathbf{I}$$

Multiplying both sides of the equation by $\mathbf{r}_{dx}$ on the right we have

$$\mathbf{w}_{n+1} = \left(\mathbf{I} - \mu\mathbf{R}_x\right)\mathbf{w}_n + \mu\mathbf{r}_{dx}$$

which is the steepest descent algorithm.

(b) Using the approximation $\widehat{\mathbf{R}}_x = \mathbf{x}(n)\mathbf{x}^T(n)$ for $\mathbf{R}_x$ we have

$$\widehat{\mathbf{R}}_x^{-1}(n+1) = \left[\mathbf{I} - \mu\mathbf{x}(n)\mathbf{x}^T(n)\right]\widehat{\mathbf{R}}_x^{-1}(n) + \mu\mathbf{I}$$

Multiplying both sides of the equation by $\mathbf{r}_{dx}$ on the right, we have

$$\mathbf{w}_{n+1} = \left[\mathbf{I} - \mu\mathbf{x}(n)\mathbf{x}^T(n)\right]\mathbf{w}_n + \mu\mathbf{r}_{dx}$$

(c) The recursion in (b) is the same as the p-vector algorithm (see Problem 14). However, if we use the approximation

$$\widehat{\mathbf{r}}_{dx} = d(n)\mathbf{x}(n)$$

then the recursion becomes equivalent to the LMS algorithm.

9.5 The convergence of a pth-order LMS adaptive filter depends on the eigenvalues of the autocorrelation matrix, $\mathbf{R}_x$, of the input process $x(n)$. These eigenvalues, in turn, depend upon the size, p, of $\mathbf{R}_x$. For example, it follows from the Bordering Theorem (see p. 48) that the maximum eigenvalue is a monotonically nondecreasing function of p, and the minimum eigenvalue is a monotonically nonincreasing function of p. In addition, it follows from the eigenvalue extremal property (see p. 97) that the maximum and minimum eigenvalues approach the maximum and minimum values of the power spectrum, $P_x(e^{j\omega})$, as $p \to \infty$,

$$\lambda_{\max} \to \max_{\omega} P_x(e^{j\omega}) \quad ; \quad \lambda_{\min} \to \min_{\omega} P_x(e^{j\omega})$$

Suppose that the input to an adaptive filter has an autocorrelation

$$r_x(k) = \alpha^{|k|} \quad ; \quad |\alpha| < 1$$

(a) Find the eigenvectors and eigenvalues of the 2×2 autocorrelation matrix $\mathbf{R}_x = \text{Toep}\{1, \alpha\}$. (Your answer will be given in terms of α).

(b) Find the asymptotic values for the maximum and minimum eigenvalues of the $p \times p$ autocorrelation matrix $\mathbf{R}_x$ as $p \to \infty$.

(c) Find, as a function of α, the largest step size μ for convergence in the mean of the LMS algorithm, and find the slowest converging mode (assume that p is large).

Solution _____

(a) The eigenvalues of the 2×2 autocorrelation matrix are $\lambda_1 = 1 + \alpha$ and $\lambda_2 = 1 - \alpha$, and the eigenvectors are $\mathbf{v}_1 = [1, 1]^T$ and $\mathbf{v}_2 = [1, -1]^T$.

(b) Asymptotically, the maximum and minimum eigenvalues are equal to $\max_{\omega}\big[P_x(e^{j\omega})\big]$ and $\min_{\omega}\big[P_x(e^{j\omega})\big]$. Since

$$P_x(e^{j\omega}) = \sum_{k=-\infty}^{\infty} \alpha^{|k|} e^{-jk\omega} = \frac{1 - \alpha^2}{1 + \alpha^2 - 2\alpha \cos(\omega)}$$

then, assuming that $0 < \alpha < 1$, we have

$$\lambda_{\max} = P_x(e^{j\omega})\big|_{\omega=0} = \frac{1 + \alpha}{1 - \alpha}$$

$$\lambda_{\min} = P_x(e^{j\omega})\big|_{\omega=\pi} = \frac{1 - \alpha}{1 + \alpha}$$

(c) Asymptotically, the largest step size for convergence in the mean is

$$\mu = \frac{1}{2\lambda_{\max}} = 0.5 \frac{1 - \alpha}{1 + \alpha}$$

If the step size is greater than $(\lambda_{\max} + \lambda_{\min})/2$, then the slowest decaying mode is the one corresponding to the minimum eigenvalue, which decays as

$$(1 - \mu\lambda_{\min})^n = \left(1 - \mu\frac{1 - \alpha}{1 + \alpha}\right)^n$$

9.6 The *condition number*, χ, of an autocorrelation matrix $\mathbf{R}_x$ may be bounded in terms of the power spectrum of the process $P_x(e^{j\omega})$ as follows,

$$\chi = \frac{\lambda_{\max}}{\lambda_{\min}} \leq \frac{\max\limits_{\omega} P_x(e^{j\omega})}{\min\limits_{\omega} P_x(e^{j\omega})}$$

(a) Use this inequality to bound the condition number of the autocorrelation matrix for the moving average process

$$x(n) = w(n) + \alpha w(n-1)$$

where $w(n)$ is unit variance white noise.

(b) Repeat part (a) for the autoregressive process

$$x(n) = \alpha x(n-1) + w(n)$$

where $|\alpha| < 1$ and $w(n)$ is unit variance white noise.

(c) What does this bound imply about the performance of an adaptive filter when the input to the filter is a lowpass process with a power spectrum of the form

$$P_x(e^{j\omega}) = \begin{cases} 1 & ; \quad |\omega| < \omega_0 \\ 0 & ; \quad \omega_0 \leq |\omega| \leq \pi \end{cases}$$

Solution _____

(a) With

$$x(n) = w(n) + \alpha w(n-1)$$

the power spectrum is

$$P_x(e^{j\omega}) = \left| 1 + \alpha e^{-j\omega} \right|^2 = (1 + \alpha^2) + 2\alpha \cos\omega$$

If we assume that $\alpha > 0$, then

$$\max_{\omega} P_x(e^{j\omega}) = (1 + \alpha)^2$$

and

$$\min_{\omega} P_x(e^{j\omega}) = (1 - \alpha)^2$$

This implies that the condition number of the autocorrelation matrix is bounded by

$$\chi \leq \left(\frac{1+\alpha}{1-\alpha} \right)^2$$

(b) With a power spectrum of

$$P_x(e^{j\omega}) = \frac{1}{(1 + \alpha^2) + 2\alpha \cos\omega}$$

assuming that $\alpha > 0$, the condition number of the autocorrelation matrix for an AR(1) process is bounded by

$$\chi \leq \left(\frac{1+\alpha}{1-\alpha} \right)^2$$

(c) From the Bordering Theorem, we see that for a lowpass process, as the order of the adaptive filter increases, the condition number increases and, in the limit, approaches infinity. Therefore, the time constant for convergence becomes very large as the order increases.

9.7 Suppose that the input to an FIR LMS adaptive filter is a first-order autoregressive process with an autocorrelation

$$r_x(k) = c\alpha^{|k|}$$

where $c > 0$ and $0 < \alpha < 1$. Suppose that the step size μ is

$$\mu = \frac{1}{5\lambda_{\max}}$$

(a) How does the rate of convergence of the LMS algorithm depend upon the value of α?

(b) What effect does the value of c have on the rate of convergence?

(c) How does the rate of convergence of the LMS algorithm depend upon the desired signal $d(n)$?

Solution _____

(a) Recall that $\lambda_{\max}$ and $\lambda_{\min}$ are bounded by the power spectrum as follows,

$$\lambda_{\max} \leq \max_{\omega}\big[P_x(e^{j\omega})\big] = c\frac{1+\alpha}{1-\alpha}$$

$$\lambda_{\min} \geq \min_{\omega}\big[P_x(e^{j\omega})\big] = c\frac{1-\alpha}{1+\alpha}$$

Since the time constant for convergence is proportional to the condition number and, for large p,

$$\chi = \frac{\lambda_{\max}}{\lambda_{\min}} \approx \left(\frac{1+\alpha}{1-\alpha}\right)^2$$

then, as α increases, τ increases, and the convergence is slower.

(b) As the constant c changes, the eigenvalues are scaled by c. However, the condition number χ is unaffected. Therefore, c does not affect the time constant for convergence.

(c) The desired signal $d(n)$ has no effect on the rate of convergence.

9.9 The first three autocorrelations of a process $x(n)$ are

$$r_x(0) = 1, \quad r_x(1) = 0.5, \quad r_x(2) = 0.5$$

Design a two-coefficient LMS adaptive linear predictor for $x(n)$ that has a misadjustment

$$\mathcal{M} = 0.05$$

and find the steady-state mean-square error.

Solution _____

The misadjustment is

$$\mathcal{M} = \frac{\alpha \ \text{tr}(\mathbf{R}_x)}{2 - \alpha \ \text{tr}(\mathbf{R}_x)}$$

Therefore,

$$2\mathcal{M} - \alpha\mathcal{M}\text{tr}(\mathbf{R}_x) = \alpha \ \text{tr}(\mathbf{R}_x)$$

and

$$\alpha = \frac{2\mathcal{M}}{(1 + \mathcal{M})\text{tr}(\mathbf{R}_x)}$$

With $\mathcal{M} = 0.05$ and $\text{tr}(\mathbf{R}_x) = 2r_x(0) = 2$ we have

$$\alpha = 1/21$$

and the LMS update equation becomes

$$\mathbf{w}_{n+1} = \mathbf{w}_n + \tfrac{1}{21}e(n)\mathbf{x}(n)$$

The steady state mean-square error is

$$
\begin{aligned}
\xi(\infty) &= \xi_{\min} + \xi_{\text{ex}}(\infty) = \xi_{\min}\left[1 + \frac{\xi_{\text{ex}}(\infty)}{\xi_{\min}}\right] \\
&= (1 + \mathcal{M})\xi_{\min} = 1.05\xi_{\min}
\end{aligned}
$$

Since

$$\xi_{\min} = r_x(0) - w(1)r_x(1) - w(2)r_x(2)$$

where

$$\begin{bmatrix} 1 & 0.5 \\ 0.5 & 1 \end{bmatrix} \begin{bmatrix} w(1) \\ w(2) \end{bmatrix} = \begin{bmatrix} 0.5 \\ 0.5 \end{bmatrix}$$
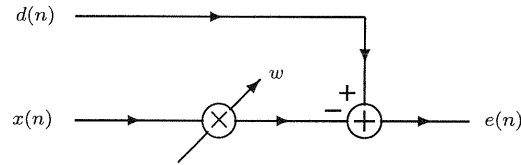
then, with $w(1) = w(2) = 1/3$, and

$$\xi_{\min} = 1 - \left(\tfrac{1}{3}\right)\left(\tfrac{1}{2}\right) - \left(\tfrac{1}{3}\right)\left(\tfrac{1}{2}\right) = 2/3$$

we have

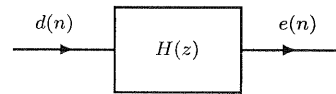$$\xi(\infty) = \left(\tfrac{2}{3}\right)1.05 = 0.7$$

9.10 Consider the single-weight adaptive filter shown in the figure below

(a) Write down the LMS algorithm to update the weight w.

(b) Suppose that $x(n)$ is a constant:

$$x(n) = \begin{cases} K & ; & n \geq 0 \\ 0 & ; & \text{otherwise} \end{cases}$$

Find the system function relating $d(n)$ to $e(n)$ using the LMS algorithm, i.e., find $H(z)$ in the figure below.

(c) Determine the range of values for μ for which $H(z)$ is stable.

Solution _____

(a) The LMS weight update equation is

$$w_{n+1} = w_n + \mu e(n) x(n)$$

(b) With $x(n)$ equal to a constant, the error $e(n)$ is

$$e(n) = d(n) - \widehat{d}(n) = d(n) - K w_n$$

where

$$w_n = w_{n-1} + \mu K e(n-1)$$

Therefore,

$$\begin{aligned} e(n) - e(n-1) &= d(n) - d(n-1) - K[w_n - w_{n-1}] \\ &= d(n) - d(n-1) - \mu K^2 e(n-1) \end{aligned}$$

For the system function we thus have

$$H(z) = \frac{D(z)}{E(z)} = \frac{1 - z^{-1}}{1 - (1 - \mu K^2) z^{-1}}$$

(c) The filter is stable if

$$|1 - \mu K^2| < 1 \qquad \longrightarrow \qquad 0 < \mu < 2/K^2$$

9.11 The LMS adaptive filter minimizes the instantaneous squared error

$$\xi(n) = |e(n)|^2$$

Consider the modified functional

$$\xi'(n) = |e(n)|^2 + \beta \mathbf{w}_n^H \mathbf{w}_n$$

where $\beta > 0$.

(a) Derive the LMS coefficient update equation for $\mathbf{w}_n$ that minimizes $\xi'(n)$.

(b) Determine the condition on the step size μ that will ensure that $\mathbf{w}_n$ converges in the mean.

(c) If μ is small enough so that $\mathbf{w}_n$ converges in the mean, what does $\mathbf{w}_n$ converge to?

Solution _____

(a) First we evaluate the gradient of $\xi'(n)$,

$$\nabla \xi'(n) = \nabla |e(n)|^2 + \beta \nabla \left[\mathbf{w}_n^H \mathbf{w}_n \right]$$

The first term is the same as we have for the LMS algorithm,

$$\nabla |e(n)|^2 = -e(n) x^*(n)$$

whereas the second term is equal to

$$\nabla \left[\mathbf{w}_n^H \mathbf{w}_n \right] = \mathbf{w}_n$$

Therefore, the LMS update equation is

$$\mathbf{w}_{n+1} = \mathbf{w}_n + \mu e(n) \mathbf{x}^*(n) - \mu \beta \mathbf{w}_n = (1 - \mu \beta) \mathbf{w}_n + \mu e(n) \mathbf{x}^*(n)$$

(b) Repeating the analysis that was done for the LMS adaptive filter we have

$$E\{\mathbf{w}_{n+1}\} = \left[(1 - \mu \beta) \mathbf{I} - \mu \mathbf{R}_x \right] E\{\mathbf{w}_n\} + \mu \mathbf{r}_{dx}$$

Therefore, for stability (convergence in the mean) we require

$$\left| (1 - \beta \mu) - \mu \lambda_k \right| < 1 \quad ; \quad k = 0, 1, 2, \ldots, p$$

which implies that

$$0 < \mu < \frac{2}{\beta + \lambda_{\max}}$$

(c) If $\mathbf{w}_n$ converges in the mean to, say $\mathbf{w}_\infty$, then

$$\mathbf{w}_\infty = (1 - \beta \mu) \mathbf{w}_\infty - \mu \mathbf{R}_x \mathbf{w}_\infty + \mu \mathbf{r}_{dx}$$

This implies that

$$\beta \mu \mathbf{w}_\infty + \mu \mathbf{R}_x \mathbf{w}_\infty = \mu \mathbf{r}_{dx}$$

or

$$\left(\mathbf{R}_x + \beta \mathbf{I} \right) \mathbf{w}_\infty = \mathbf{r}_{dx}$$

Therefore,

$$\mathbf{w}_\infty = \left(\mathbf{R}_x + \beta \mathbf{I} \right)^{-1} \mathbf{r}_{dx}$$

9.12 Show that the normalized LMS algorithm is equivalent to using the update equation

$$\mathbf{w}_{n+1} = \mathbf{w}_n + \mu e'(n)\mathbf{x}^*(n)$$

where $e'(n)$ is the error at time n that is based on the new filter coefficients $\mathbf{w}_{n+1}$,

$$e'(n) = d(n) - \mathbf{w}_{n+1}^T\mathbf{x}(n)$$

Discuss the relationship between μ and the parameter ϵ in the normalized LMS algorithm.

Solution _____

Substituting the expression for $e'(n)$ into the update equation gives

$$\begin{aligned}
\mathbf{w}_{n+1} &= \mathbf{w}_n + \mu\big[d(n) - \mathbf{w}_{n+1}^T\mathbf{x}(n)\big]\mathbf{x}^*(n) \\
&= \mathbf{w}_n + \mu d(n)\mathbf{x}^*(n) - \mu\big[\mathbf{x}^*(n)\mathbf{x}^T(n)\big]\mathbf{w}_{n+1}
\end{aligned}$$

Bringing the term with $\mathbf{w}_{n+1}$ on the right to the left-side we have

$$\begin{aligned}
\big[\mathbf{I} + \mu\mathbf{x}^*(n)\mathbf{x}^T(n)\big]\mathbf{w}_{n+1} &= \mathbf{w}_n + \mu d(n)\mathbf{x}^*(n) \\
&= \mathbf{w}_n + \mu\big[d(n) - \mathbf{w}_n^T\mathbf{x}(n)\big]\mathbf{x}^*(n) + \mu\mathbf{w}_n^T\mathbf{x}(n)\mathbf{x}^*(n) \\
&= \big[\mathbf{I} + \mu\mathbf{x}^*(n)\mathbf{x}^T(n)\big]\mathbf{w}_n + \mu e(n)\mathbf{x}^*(n)
\end{aligned}$$

Multiplying both sides by the inverse of $\big[\mathbf{I} + \mu\mathbf{x}^*(n)\mathbf{x}^T(n)\big]$ gives

$$\mathbf{w}_{n+1} = \mathbf{w}_n + \mu e(n)\big[\mathbf{I} + \mu\mathbf{x}^*(n)\mathbf{x}^T(n)\big]^{-1}\mathbf{x}^*(n)$$

Using Woodbury's identity to evaluate the inverse (see Eq. (2.30) on p. 29),

$$\big[\mathbf{I} + \mu\mathbf{x}^*(n)\mathbf{x}^T(n)\big]^{-1} = \mathbf{I} - \mu\frac{\mathbf{x}^*(n)\mathbf{x}^T(n)}{1 + \mu\mathbf{x}^T(n)\mathbf{x}^*(n)} = \mathbf{I} - \mu\frac{\mathbf{x}^*(n)\mathbf{x}^T(n)}{1 + \mu\|\mathbf{x}(n)\|^2}$$

and incorporating this into the update for $\mathbf{w}_n$, we have

$$\mathbf{w}_{n+1} = \mathbf{w}_n + \mu e(n)\mathbf{x}^*(n) - \mu^2 e(n)\frac{\mathbf{x}^*(n)\mathbf{x}^T(n)}{1 + \mu\|\mathbf{x}(n)\|^2}\mathbf{x}^*(n)$$

Since

$$\big[\mathbf{x}^*(n)\mathbf{x}^T(n)\big]\mathbf{x}^*(n) = \mathbf{x}^*(n)\big[\mathbf{x}^T(n)\mathbf{x}^*(n)\big] = \|\mathbf{x}(n)\|^2\mathbf{x}^*(n)$$

then the update may be written as

$$\begin{aligned}
\mathbf{w}_{n+1} &= \mathbf{w}_n + \mu e(n)\mathbf{x}^*(n) - \mu^2 e(n)\frac{\|\mathbf{x}(n)\|^2}{1 + \mu\|\mathbf{x}(n)\|^2}\mathbf{x}^*(n) \\
&= \mathbf{w}_n + \mu e(n)\Big[1 - \frac{\mu\|\mathbf{x}(n)\|^2}{1 + \mu\|\mathbf{x}(n)\|^2}\Big]\mathbf{x}^*(n) \\
&= \mathbf{w}_n + \frac{1}{\dfrac{1}{\mu} + \|\mathbf{x}(n)\|^2}e(n)\mathbf{x}^*(n)
\end{aligned}$$

which is the same as the normalized LMS algorithm with $\beta = 1$ and $\epsilon = \dfrac{1}{\mu}$.

9.13 A process $x(n)$ is formed by passing white noise $w(n)$ through a filter that has a system function

$$H(z) = \frac{1}{1 - 0.08z^{-1} - 0.9z^{-2}}$$

The variance of the white noise is $\sigma_w^2 = (0.19)(0.18)$. The LMS algorithm with two coefficients is used to estimate a process $d(n)$ from $x(n)$.

(a) What is the maximum value for the step size, μ, in order for the LMS algorithm to converge in the mean? Hint: Use the inverse Levinson-Durbin recursion to find the autocorrelation sequence of $x(n)$.

(b) What is the time constant for convergence?

(c) What value for the step size would you use to maximize the rate of convergence of the weights?

(d) If the cross-correlation between $x(n)$ and $d(n)$ is zero,

$$E\{d(n)\mathbf{x}^*(n)\} = 0$$

what are the optimum filter coefficients $\mathbf{w} = [w(0), w(1)]^T$?

Solution _____

(a) The first step is to find the autocorrelation sequence for $x(n)$. This may be done by finding the inverse discrete-time Fourier transform of the power spectrum

$$P_x(e^{j\omega}) = \frac{(0.19)(0.18)}{\left|1 - 0.08e^{-j\omega} - 0.9e^{-j2\omega}\right|^2}$$

or by using the inverse Levinson-Durbin recursion. Here, we will use the inverse Levinson-Durbin recursion.

To begin, we need to find the reflection coefficients, which are

$$\mathbf{\Gamma} = \begin{bmatrix} -0.8, & -0.9 \end{bmatrix}^T$$

and the modeling error, ϵ_2, which is equal to the white noise variance,

$$\epsilon_2 = (0.19)(0.18)$$

To begin the recursion, we set

$$r_x(0) = \frac{\epsilon_2}{\prod_{i=1}^{2}(1 - \Gamma_i^2)} = 0.5$$

We then compute the first-order model,

$$\mathbf{a}_1 = \begin{bmatrix} 1 \\ \Gamma_1 \end{bmatrix} = \begin{bmatrix} 1 \\ -0.8 \end{bmatrix}$$

and then evaluate $r_x(1)$ as follows

$$r_x(1) = -r_x(0)\Gamma_1 = 0.4$$

Then, using the second-order coefficients,

$$\mathbf{a}_2 = \begin{bmatrix} 1, & -0.08, & -0.9 \end{bmatrix}^T$$

we find $r_x(2)$ as follows

$$r_x(2) = -a_2(1)r_x(1) - a_2(2)r_x(0) = 0.482$$

Thus, the autocorrelation sequence is

$$\mathbf{r}_x = \begin{bmatrix} 0.5, & 0.4, & 0.4820 \end{bmatrix}^T$$

Finally, to determine the maximum value for the step size for convergence in the mean, we need the eigenvalues of the autocorrelation matrix. With two coefficients, we must find the eigenvalues of the matrix

$$\mathbf{R}_x = \begin{bmatrix} r_x(0) & r_x(1) \\ r_x(1) & r_x(0) \end{bmatrix}$$

which are

$$\lambda_1 = r_x(0) + r_x(1) = 0.9 \quad ; \quad \lambda_1 = r_x(0) - r_x(1) = 0.1$$

Thus, for convergence in the mean, we must have

$$0 < \mu < \frac{2}{\lambda_{\max}} = 2.22\bar{2}$$

(b) If μ is small in the sense that $\mu\lambda_k \ll 1$, then the time constant for convergence is equal to

$$\tau = \frac{1}{\mu\lambda_{\min}} = \frac{10}{\mu}$$

For example, if we assume that $\mu \cdot \lambda_{\max} = 0.1$, or $\mu = 1/9$, then

$$\tau = \frac{10}{1/9} = 90$$

(c) The expected value of the weights evolves in the same way as the weights in the steepest descent algorithm,

$$E\{\mathbf{w}_n\} = \mathbf{w} + \sum_{k=1}^{p}(1 - \mu\lambda_k)^n u_0(k)\mathbf{v}_k$$

For two weights, we have

$$E\{\mathbf{w}_n\} = \mathbf{w} + (1 - \mu\lambda_1)^n u_0(1)\mathbf{v}_1 + (1 - \mu\lambda_2)^n u_0(2)\mathbf{v}_2$$

In order to maximize the rate at which these terms decay to zero, we want to choose μ so that we maximize the rate at which the slowest decaying mode goes to zero,

$$\min_{\mu} \left[\max_k |1 - \mu\lambda_k| \right]$$

This occurs when we set

$$\mu = \frac{2}{\lambda_1 + \lambda_2} = 2$$

(d) If the cross-correlation between $x(n)$ and $d(n)$ is zero,

$$E\{d(n)\mathbf{x}^*(n)\} = 0$$

then $d(n)$ cannot be estimated from $x(n)$ using a linear estimator. Therefore, the optimum coefficients are zero,

$$\mathbf{w} = \mathbf{0}$$

This may also be seen from the Wiener-Hopf equations, which are

$$\mathbf{R}_x\mathbf{w} = \mathbf{r}_{dx}$$

Thus, since $\mathbf{R}_x$ is nonsingular, if $\mathbf{r}_{dx} = \mathbf{0}$, then $\mathbf{w} = \mathbf{0}$.

9.14 Griffiths developed an algorithm for adaptive beamforming referred to as the p-vector algorithm that eliminates the need for a reference signal $d(n)$. This algorithm may be derived as follows. Recall that the filter coefficient update equation for the LMS algorithm is

$$\mathbf{w}_{n+1} = \mathbf{w}_n + \mu e(n)\mathbf{x}^*(n) = \mathbf{w}_n + \mu d(n)\mathbf{x}^*(n) - \mu[\mathbf{w}_n^T\mathbf{x}(n)]\mathbf{x}^*(n)$$

Note that $d(n)$ is not explicitly needed in this update equation. Instead, what is required is the product $d(n)\mathbf{x}^*(n)$. Therefore, if $d(n)\mathbf{x}^*(n)$ is replaced with its expected value, $\mathbf{r}_{dx} = E\{d(n)\mathbf{x}^*(n)\}$, or by an estimate of this ensemble average, then we have an update equation that does not require knowledge of $d(n)$,

$$\mathbf{w}_{n+1} = \mathbf{w}_n + \mu\mathbf{r}_{dx} - \mu[\mathbf{w}_n^T\mathbf{x}(n)]\mathbf{x}^*(n)$$

This is the p-vector algorithm proposed by Griffiths.

(a) What constraints must be placed on the step size μ in order for $\mathbf{w}_n$ to converge in the mean?

(b) Develop a "leaky" p-vector algorithm and determine the range of values for μ for convergence in the mean. Assuming that μ is selected so that $\mathbf{w}_n$ converges in the mean, find $\lim_{n\to\infty} E\{\mathbf{w}_n\}$.

Solution _____

(a) Taking the expectation of the filter coefficient update equation, using the independence assumption, we have

$$E\{\mathbf{w}_{n+1}\} = E\{\mathbf{w}_n\} + \mu\mathbf{r}_{dx} - \mu\mathbf{R}_x E\{\mathbf{w}_n\} = (\mathbf{I} - \mu\mathbf{R}_x)E\{\mathbf{w}_n\} + \mu\mathbf{r}_{dx}$$

This equation is exactly the same as the one for the LMS algorithm. Therefore, the p-vector algorithm will converge in the mean if

$$0 < \mu < \frac{2}{\lambda_{\max}}$$

(b) For the leaky LMS algorithm, the update equation is

$$
\begin{aligned}
\mathbf{w}_{n+1} &= (1 - \mu\gamma)\mathbf{w}_n + \mu e(n)\mathbf{x}^*(n) \\
&= (1 - \mu\gamma)\mathbf{w}_n + \mu\big[d(n) - \mathbf{w}_n^T\mathbf{x}(n)\big]\mathbf{x}^*(n) \\
&= (1 - \mu\gamma)\mathbf{w}_n + \mu d(n)\mathbf{x}^*(n) - \mu\mathbf{x}^*(n)\mathbf{x}^T(n)\mathbf{w}_n
\end{aligned}
$$

Replacing $d(n)\mathbf{x}^*(n)$ with $\mathbf{r}_{dx}$ we have the update equation for a *leaky* p-vector algorithm,

$$\mathbf{w}_{n+1} = (1 - \mu\gamma)\mathbf{w}_n + \mu\mathbf{r}_{dx} - \mu\mathbf{x}^*(n)\mathbf{x}^T(n)\mathbf{w}_n$$

Taking the expectation of this equation we have

$$E\{\mathbf{w}_{n+1}\} = (1 - \mu\gamma)E\{\mathbf{w}_n\} + \mu\mathbf{r}_{dx} - \mu\mathbf{R}_x E\{\mathbf{w}_n\} = \big[(1 - \mu\gamma)\mathbf{I} - \mu\mathbf{R}_x\big]E\{\mathbf{w}_n\} + \mu\mathbf{r}_{dx}$$

Therefore, $\mathbf{w}_n$ converges in the mean if

$$0 < \mu < \frac{2}{\lambda_{\max} + \gamma}$$

If $E\{\mathbf{w}_n\}$ converges in the mean,

$$\lim_{n \to \infty} E\{\mathbf{w}_n\} = \mathbf{w}_\infty$$

then

$$\mathbf{w}_\infty = \left[(1 - \mu\gamma)\mathbf{I} - \mu\mathbf{R}_x\right]\mathbf{w}_\infty + \mu\mathbf{r}_{dx}$$

and we have

$$\mathbf{w}_\infty = (\gamma\mathbf{I} + \mathbf{R}_x)^{-1}\mathbf{r}_{dx}$$

9.15 An FIR filter with system function $H(z)$ may be implemented using a frequency sampling structure as follows

$$H(z) = \frac{1 - z^{-M}}{M} \sum_{k=0}^{M-1} \frac{H(k)}{1 - e^{j2\pi k/M} z^{-1}} = H_1(z) H_2(z)$$

where $H_1(z)$ is a comb filter that has M zeroes equally spaced around the unit circle and $H_2(z)$ is a filterbank of resonators where the coefficients $H(k)$ are the DFT coefficients of $h(n)$, i.e.,

$$H(k) = \sum_{n=0}^{M-1} h(n) e^{-j2\pi kn/M}$$

Suppose that this structure is implemented as an adaptive filter using the LMS algorithm to adjust the filter (DFT) coefficients, $H(k)$. Derive the time-update equation for these coefficients and sketch the adaptive filter structure.

Solution _____

The frequency sampling structure is shown in the figure below, with the DFT coefficients, $H(k)$, made to be adaptive.
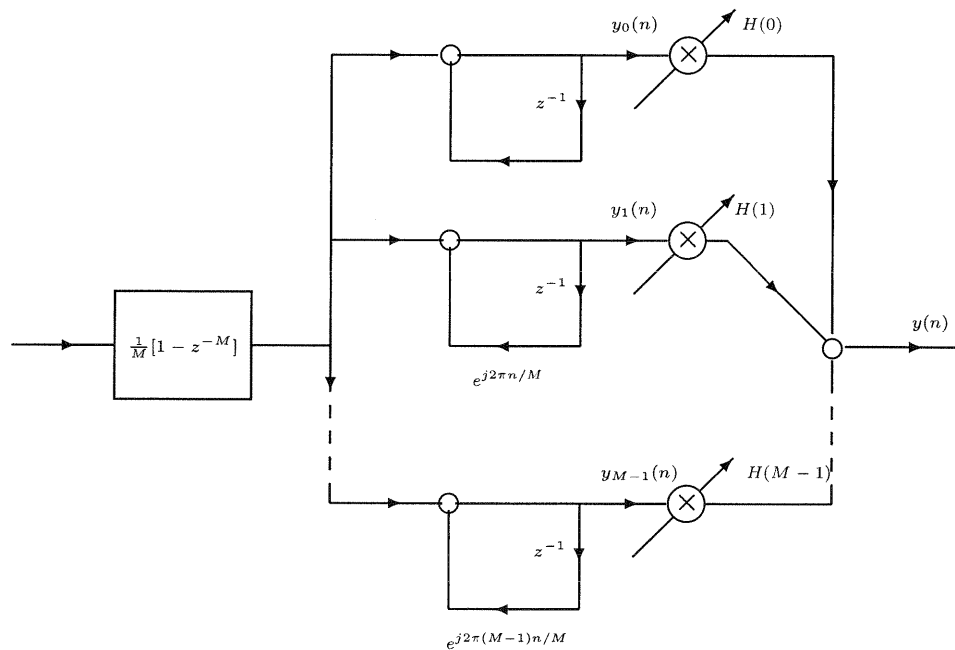
With

$$e(n) = d(n) - y(n) = d(n) - \sum_{k=0}^{M-1} H(k) y_k(n)$$

where

$$y_k(n) = e^{j2\pi kn/M} y_k(n-1) + \frac{1}{M} \big[x(n) - x(n-M) \big]$$

the LMS update of the coefficients $H(k)$ is

$$H_{n+1}(k) = H_n(k) - \mu \nabla |e(n)|^2$$

Since the coefficients $H(k)$ are complex, the gradient is

$$\nabla |e(n)|^2 = -e(n) y_k^*(n)$$

Therefore, the update equation becomes

$$H_{n+1}(k) = H_n(k) + \mu e(n) y_k^*(n)$$

9.16 In many signal processing applications, it is important for a filter to have *linear phase*. This is particularly true in speech and image processing applications, where phase distortion produced by a filter may severely degrade the signal. Therefore, suppose that we would like to design an adaptive linear phase filter whose weights at time n satisfy the following symmetry constraint

$$w_n(k) = w_n(p - k) \quad ; \quad k = 0, 1, \ldots, p$$

For example, consider the two-coefficient linear phase adaptive filter shown in the figure below.
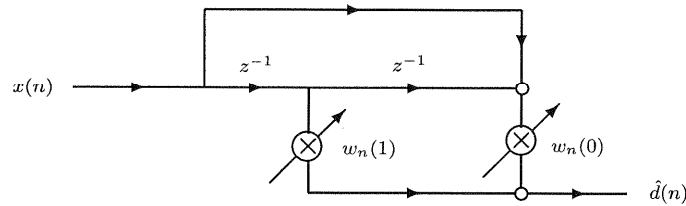
Note that this filter may be viewed as an FIR adaptive filter that is constrained to have the first coefficient equal to the third. As before, define the weight vector, $\mathbf{w}_n$, by

$$\mathbf{w}_n = [w_n(0), w_n(1)]^T$$

and the error sequence, $e(n)$, by

$$e(n) = d(n) - \widehat{d}(n)$$

and assume that

$$r_x(0) = 1, \quad r_x(1) = 0.5, \quad r_x(2) = 0.1$$

(a) Derive the normal equations for the filter that minimizes the mean-square error

$$\xi = E\{[d(n) - \widehat{d}(n)]^2\}$$

(b) Derive the LMS filter coefficient update equations for this constrained transversal filter.

(c) What values for the step size μ may be used if the weights are to converge in the *mean-square* sense?

(d) If we drop the linear phase constraint and use a three-coefficient LMS adaptive filter with $\mathbf{w} = [w(0), w(1), w(2)]^T$, what values of the step size μ may be used if the adaptive filter is to converge in the mean-square sense?

Solution _____

(a) If we define

$$
\begin{aligned}
z_0(n) &= x(n-1) + x(n-3) \\
z_1(n) &= x(n-2)
\end{aligned}
$$

then the estimate of $x(n)$ may be written as

$$\widehat{x}(n) = w(0)z_0(n) + w(1)z_1(n) = \mathbf{w}^T \mathbf{z}(n)$$

and the normal equations are

$$\mathbf{R}_z \mathbf{w} = \mathbf{r}_{dz}$$

where

$$\mathbf{R}_z = E\{\mathbf{z}(n)\mathbf{z}^T(n)\} = \begin{bmatrix} 2r_x(0) + 2r_x(2) & 2r_x(1) \\ 2r_x(1) & r_x(0) \end{bmatrix}$$

and

$$\mathbf{r}_{dz} = E\{d(n)\mathbf{z}^T(n)\} = [r_{dx}(1) + r_{dx}(3), \; r_{dx}(2)]^T$$

(b) The LMS algorithm is

$$\mathbf{w}_{n+1} = \mathbf{w}_n + \mu e(n)\mathbf{z}(n)$$

(c) For mean-square convergence

$$0 < \mu < \frac{2}{\text{Tr}(\mathbf{R}_z)} = \frac{2}{3r_x(0) + 2r_x(2)} = \frac{2}{3.2} \approx 0.625$$

(d) For the unconstrained filter we require

$$0 < \mu < \frac{2}{\text{Tr}(\mathbf{R}_x)} = \frac{2}{3}$$

9.17 In some applications it is known that a given complex-valued process has a constant envelope, e.g. phase modulated signals. The *Constant Modulus Algorithm* (CMA) is an adaptive filtering technique that adjusts the filter coefficients in order to minimize the envelop variation. Given a complex signal $x(n)$ and a set of complex weights $w_n(k)$ at time n, the output of the adaptive filter is

$$y(n) = \mathbf{w}_n^H \mathbf{x}(n)$$

With the constant modulus algorithm, the weights are to be found that minimize the error

$$\xi(n) = \tfrac{1}{4} E\{(|y(n)|^2 - 1)^2\}$$

which is a non-negative measure of the average amount that the envelope of the filter output $y(n)$ deviates from a nominal level (unity in this case). Using Widrow's approach of estimating ensemble averages with one point sample averages, derive the CMA algorithm, which is an LMS version of a steepest descent algorithm to minimize the error $\xi(n)$ defined above.

Solution _____

The constant modulus algorithm for updating the coefficients $\mathbf{w}_n$ has the form

$$\mathbf{w}_{n+1} = \mathbf{w}_n - \mu \nabla e^2(n)$$

where

$$e(n) = \tfrac{1}{2}\Big(|y(n)|^2 - 1\Big)$$

Computing the gradient of $e^2(n)$ we have

$$\nabla e^2(n) = 2e(n)\nabla |y(n)|^2$$

Since

$$|y(n)|^2 = \mathbf{w}_n^H \mathbf{x}(n)\mathbf{x}^H(n)\mathbf{w}_n$$

then

$$\nabla |y(n)|^2 = \big[\mathbf{x}(n)\mathbf{x}^H(n)\big]\mathbf{w}_n$$

and we have

$$\nabla e^2(n) = 2e(n)\mathbf{x}(n)\mathbf{x}^H(n)\mathbf{w}_n = \big(|y(n)|^2 - 1\big)\mathbf{x}(n)\mathbf{x}^H(n)\mathbf{w}_n$$
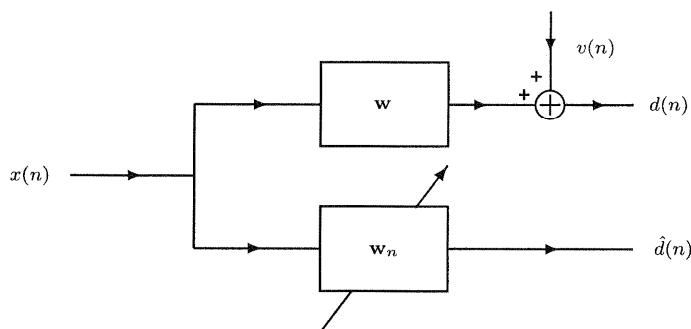
Thus, the CMA algorithm is

$$\mathbf{w}_{n+1} = \mathbf{w}_n + \mu\big(1 - |y(n)|^2\big)y^*(n)\mathbf{x}(n)$$

9.18 The output $d(n)$ of an unknown system is given by

$$d(n) = \sum_{k=0}^{p} w(k)x(n-k) + v(n) = \mathbf{w}^T\mathbf{x}(n) + v(n)$$

where $w(k)$ are the unknown system parameters, $x(n)$ is the system input, and $v(n)$ is zero mean white noise with a variance of σ_v^2. The block diagram below shows an adaptive filter that is used to model the unknown system.

Assume that $x(n)$ is real-valued and suppose that we would like to find the weight vector $\mathbf{w}_n$ that minimizes the error

$$\xi(k) = E\{e^{2k}(n)\}$$

for some positive integer k where $e(n) = d(n) - \widehat{d}(n)$.

(a) As in the LMS algorithm, use the instantaneous gradient vector and derive an LMS update equation for $\mathbf{w}_n$.

(b) Assuming that $v(n)$ is independent of $\mathbf{x}(n)$, and that the weight-error vector

$$\mathbf{c}_n = \mathbf{w}_n - \mathbf{w}$$

is close to zero, and that $\mathbf{c}_n$ is independent of $\mathbf{x}(n)$, show that

$$E\{\mathbf{c}_{n+1}\} = \left[\mathbf{I} - \mu k(2k-1)E\{v^{2k-2}(n)\}\mathbf{R}_x\right]E\{\mathbf{c}_n\}$$

where $\mathbf{R}_x$ is the autocorrelation matrix of $x(n)$.

(c) Show that the modified LMS algorithm derived in part (a) will converge in the mean if the step size μ satisfies the condition

$$0 < \mu < \frac{1}{k(2k-1)E\{v^{2(k-1)}(n)\}\lambda_{\max}}$$

where $\lambda_{\max}$ is the largest eigenvalue of the matrix $\mathbf{R}_x$.

(d) For $k = 1$, show that the results derived in parts (a), (b), and (c) reduce to those in the conventional LMS algorithm.

Solution _____

(a) The instantaneous gradient vector at time n is

$$\nabla \xi(k, n) = \frac{\partial e^{2k}(n)}{\partial \mathbf{w}} = 2ke^{2k-1}(n)\frac{\partial e(n)}{\partial \mathbf{w}} = -2ke^{2k-1}(n)\mathbf{x}(n)$$

Therefore, the LMS update equation for $\mathbf{w}_n$ is

$$\mathbf{w}_{n+1} = \mathbf{w}_n - \mu\nabla\xi(k, n) = \mathbf{w}_n + 2\mu k e^{2k-1}(n)\mathbf{x}(n)$$

(b) With $\mathbf{c}_n = \mathbf{w}_n - \mathbf{w}$, if we subtract $\mathbf{w}$ from both sides of the expression for $\mathbf{w}_{n+1}$ in part (a) we have

$$\mathbf{c}_{n+1} = \mathbf{c}_n + 2\mu k e^{2k-1}(n)\mathbf{x}(n)$$

Evaluating the term $e^{2k-1}(n)$ we have

$$
\begin{aligned}
e^{2k-1}(n) &= \left[\mathbf{w}^T\mathbf{x}(n) + v(n) - \mathbf{w}_n^T\mathbf{x}(n)\right]^{2k-1} = \left[-\mathbf{c}_n^T\mathbf{x}(n) + v(n)\right]^{2k-1} \\
&= v^{2k-1}(n)\left[1 - \frac{\mathbf{c}_n^T\mathbf{x}(n)}{v(n)}\right]^{2k-1}
\end{aligned}
$$

If $\mathbf{c}_n$ is small, then

$$e^{2k-1}(n) \approx v^{2k-1}(n)\left[1 - (2k-1)\frac{\mathbf{c}_n^T\mathbf{x}(n)}{v(n)}\right] = v^{2k-1}(n) - (2k-1)\mathbf{c}_n^T\mathbf{x}(n)v^{2(k-1)}(n)$$

Substituting this into the expression for $\mathbf{c}_{n+1}$ gives

$$
\begin{aligned}
\mathbf{c}_{n+1} &= \mathbf{c}_n + 2\mu k\left[v^{2k-1}(n) - (2k-1)\mathbf{c}_n^T\mathbf{x}(n)v^{2(k-1)}(n)\right]\mathbf{x}(n) \\
&= \left[\mathbf{I} - 2\mu k(2k-1)v^{2(k-1)}(n)\mathbf{x}(n)\mathbf{x}^T(n)\right]\mathbf{c}_n + 2\mu k v^{2k-1}(n)\mathbf{x}(n)
\end{aligned}
$$

Taking the expectation of both sides of this equation, using the assumption that $\mathbf{c}_n$ is independent of $\mathbf{x}(n)$, and that $\mathbf{x}(n)$ is independent of $v(n)$, we have

$$E\{\mathbf{c}_{n+1}\} = \left[\mathbf{I} - 2\mu k(2k-1)E\{v^{2(k-1)}(n)\}\mathbf{R}_x\right]E\{\mathbf{c}_n\} + 2\mu E\{v^{(2k-1)}(n)\}E\{\mathbf{x}(n)\}$$

where $\mathbf{R}_x$ is the autocorrelation matrix for $x(n)$.

(c) This follows immediately as with the LMS algorithm by introducing the factorization $\mathbf{R}_x = \mathbf{Q}\mathbf{\Lambda}\mathbf{Q}^T$ and defining a new weight error vector in a rotated coordinate system.

(d) For $k = 1$, the results in parts (a), (b), and (c) become, respectively,

$$\mathbf{w}_{n+1} = \mathbf{w}_n + 2\mu e(n)\mathbf{x}(n)$$

$$E\{\mathbf{c}_{n+1}\} = \left[\mathbf{I} - 2\mu\mathbf{R}_x\right]E\{\mathbf{c}_n\} + E\{v(n)\mathbf{x}(n)\}$$

and

$$0 < \mu < \frac{1}{\lambda_{\max}}$$

which are the same as for the LMS algorithm.

9.19 Adaptive filters are commonly used for linear prediction. Although harmonic signals such as sinusoids are perfectly predictable, measurement noise will degrade the performance of the predictor and add a bias to the coefficients, $\mathbf{w}$. For example, suppose that we want to design an adaptive linear predictor for a real-valued process $x(n)$ using the noisy measurements

$$y(n) = x(n) + v(n)$$

where $v(n)$ is zero mean white noise that is uncorrelated with $x(n)$. Assume that the variance of $v(n)$ is σ_v^2.

(a) Using the LMS algorithm

$$\mathbf{w}_{n+1} = \mathbf{w}_n + \mu e(n)\mathbf{y}(n)$$

find the range of values for μ for which the LMS algorithm converges in the mean and find

$$\lim_{n \to \infty} E\{\mathbf{w}_n\}$$

(b) The γ-LMS algorithm has been proposed as an adaptive filtering algorithm to combat the effect of measurement noise. Using the noisy observations, $y(n)$, this algorithm is

$$\mathbf{w}_{n+1} = \gamma \mathbf{w}_n + \mu e(n)\mathbf{y}(n)$$

where γ is a constant. Explain how the γ-LMS algorithm can be used to remove the bias in the steady-state solution of the LMS algorithm. Specifically, how would you select values for μ and γ?

Solution _____

(a) The LMS algorithm is

$$\mathbf{w}_{n+1} = \mathbf{w}_n + \mu e(n)\mathbf{y}(n)$$

Taking expected values we find

$$E\{\mathbf{w}_{n+1}\} = \left[\mathbf{I} - \mu \mathbf{R}_y\right] E\{\mathbf{w}_n\} + \mu \mathbf{r}_{dy}$$

where

$$\mathbf{R}_y = \mathbf{R}_x + \sigma_v^2 \mathbf{I}$$

is the autocorrelation matrix of $\mathbf{y}(n)$. Since the eigenvalues of $\mathbf{R}_y$ are

$$\tilde{\lambda}_k = \lambda_k + \sigma_v^2$$

where λ_k are the eigenvalues of $\mathbf{R}_x$, then the LMS algorithm converges in the mean if

$$0 < \mu < \frac{2}{\lambda_{\max} + \sigma_v^2}$$

where $\lambda_{\max}$ is the maximum eigenvalue of $\mathbf{R}_x$. Furthermore, if the LMS algorithm converges in the mean, then

$$\lim_{n \to \infty} E\{\mathbf{w}_n\} = (\mathbf{R}_x + \sigma_v^2 \mathbf{I})^{-1} \mathbf{r}_{dx}$$

(b) The γ-LMS algorithm is

$$
\begin{aligned}
\mathbf{w}_{n+1} &= \gamma \mathbf{w}_n + \mu e(n) \mathbf{y}(n) \\
&= \gamma \mathbf{w}_n + \mu \big[d(n) - \mathbf{w}_n^T \mathbf{y}(n) \big] \mathbf{y}(n) \\
&= \big[\gamma \mathbf{I} - \mu \mathbf{y}(n) \mathbf{y}^T(n) \big] \mathbf{w}_n + \mu d(n) \mathbf{y}(n)
\end{aligned}
$$

Taking expected values, and using the independence assumption, we find

$$
\begin{aligned}
E\{\mathbf{w}_{n+1}\} &= \big[\gamma \mathbf{I} - \mu (\mathbf{R}_x + \sigma_v^2 \mathbf{I}) \big] E\{\mathbf{w}_n\} + \mu \mathbf{r}_{dy} \\
&= \big[(\gamma - \mu \sigma_v^2) \mathbf{I} - \mu \mathbf{R}_x \big] E\{\mathbf{w}_n\} + \mu \mathbf{r}_{dy}
\end{aligned}
$$

The conditions that we would like to impose on μ and γ are

$$
\gamma = 1 + \mu \sigma_v^2
$$

$$
0 < \mu < 2/\lambda_{\max}
$$

where $\lambda_{\max}$ is the maximum eigenvalue of the autocorrelation matrix for $x(n)$.

9.20 In some applications, it may be necessary to delay the update of the filter coefficients for a short period of time. For example, in decision-directed feedback equalization, if a sophisticated algorithm such a Viterbi decoder is used to improve the decisions, then the desired signal and thus the error is not available until a number of samples later. Therefore, assume that $x(n)$ is real-valued, and consider the delayed LMS algorithm that has a filter coefficient update equation given by

$$\mathbf{w}_{n+1} = \mathbf{w}_n + \mu e(n - n_0)\mathbf{x}(n - n_0)$$

where

$$e(n - n_0) = d(n - n_0) - y(n - n_0)$$

Note that if the delay, n_0, is equal to zero then we have the conventional LMS algorithm.

(a) For $n_0 = 1$, determine the values of μ for which the delayed LMS algorithm converges in the mean.

(b) If $\lambda_k = 1$, for $k = 1, \ldots, p$ and if the step size $\mu = 0.1$, find the time constant, τ_0 for the LMS adaptive filter $(n_0 = 0)$ and the time constant τ_1 for the delayed LMS adaptive filter with $n_0 = 1$.

Solution

(a) With $n_0 = 1$ the delayed LMS adaptive filter update equation is

$$
\begin{aligned}
\mathbf{w}_{n+1} &= \mathbf{w}_n + \mu e(n-1)\mathbf{x}(n-1) \\
&= \mathbf{w}_n + \mu \left[d(n-1) - \mathbf{w}_{n-1}^T \mathbf{x}(n-1) \right] \mathbf{x}(n-1)
\end{aligned}
$$

Taking the expected value, assuming that the weight vector $\mathbf{w}_n$ is uncorrelated with the data vector $\mathbf{x}(n)$, we have

$$E\left\{\mathbf{w}_{n+1}\right\} = E\left\{\mathbf{w}_n\right\} - \mu \mathbf{R}_x E\left\{\mathbf{w}_{n-1}\right\} + \mu \mathbf{r}_{dx}$$

Thus, the expected value of the weight vector satisfies a second-order difference equation. Diagonalizing the autocorrelation matrix and expressing this equation in terms of the rotated coefficient vector, $\mathbf{v}(n)$, we have, for the kth coefficient,

$$E\left\{v_{n+1}(k)\right\} = E\left\{v_n(k)\right\} - \mu \lambda_k E\left\{v_{n-1}(k)\right\} + \mu r_{dx}(k)$$

where λ_k for $k = 1, \ldots, N$ are the eigenvalues of $\mathbf{R}_x$. Since the characteristic equation for $E\left\{v_n(k)\right\}$ is

$$1 - z^{-1} + \mu \lambda_k z^{-2} = 0$$

in order for $E\left\{v_n(k)\right\}$ to converge in the mean, the roots of the characteristic equation must lie inside the unit circle. Since the roots are at

$$z_k = \tfrac{1}{2}\left\{ 1 \pm \sqrt{1 - 4\mu\lambda_k} \right\}$$

then the delayed LMS algorithm converges in the mean provided

$$0 < \mu < 1/\lambda_{\max}$$

(b) With $\lambda_k = 1$ and $\mu = 0.1$, the modes of the LMS algorithm behave as

$$(1 - \mu\lambda)^n = (0.9)^n$$

and the time constant is

$$\tau = \frac{1}{\mu\lambda_k} = 10$$

For the delayed LMS, the roots of the characteristic equation are

$$z_k = \tfrac{1}{2}\left\{1 \pm \sqrt{1 - .4}\right\} = 0.887, \ 0.1127$$

Therefore, the slowest decaying mode behaves as $(0.887)^k$ which is approximately the same as the LMS algorithm. Thus, the time constants for delayed LMS with $n_0 = 1$ is about the same as LMS.

9.21 In recent years, there has been an increasing interest in nonlinear digital filters. This interest has included the design of adaptive nonlinear filters. Volterra systems are an important class of nonlinear filters. Assuming that $x(n)$ is real-valued, a second-order Volterra digital filter has the form

$$y(n) = \sum_{k=0}^{K} a(k)x(n-k) + \sum_{k_1=0}^{K_1} \sum_{k_2=k_1}^{K_2} b(k_1, k_2)x(n-k_1)x(n-k_2)$$

Note that the output, $y(n)$, is formed from a linear combination of first-order (linear) terms $x(n-k)$, and a linear combination of second-order (nonlinear) terms $x(n-k_1)x(n-k_2)$. As a specific example, consider the following second-order digital Volterra filter with time-varying coefficients,

$$y(n) = a_n(0)x(n) + a_n(1)x(n-1) + b_n(0)x^2(n) + b_n(1)x(n)x(n-1)$$

Let $\mathbf{\Theta}_n$ be the coefficient vector

$$\mathbf{\Theta}_n = [a_n(0),\ a_n(1),\ b_n(0), b_n(1)]^T$$

and let $\mathbf{x}(n)$ be the data vector

$$\mathbf{x}(n) = [x(n),\ x(n-1),\ x^2(n),\ x(n)x(n-1)]^T$$

(a) Using the LMS update equation

$$\mathbf{\Theta}_{n+1} = \mathbf{\Theta}_n - \frac{1}{2}\mu\nabla e^2(n)$$

where $e(n) = d(n) - y(n)$, derive the coefficient update equations for $a_n(0), a_n(1), b_n(0)$, and $b_n(1)$.

(b) What condition must be placed on μ in order for the coefficient vector $\mathbf{\Theta}_n$ to converge in the mean?

(c) Describe what happens if the third-order statistics of $x(n)$ are zero, i.e.,

$$\begin{aligned}
E\{x^3(n)\} &= 0 \\
E\{x^2(n)x(n-1)\} &= 0 \\
E\{x(n)x^2(n-1)\} &= 0
\end{aligned}$$

Discuss how you might improve the performance of the adaptive Volterra filter by having two step size parameters, μ_1 and μ_2, one for the linear terms and one for the nonlinear terms, and discuss how these parameters must be restricted in order for the filter to converge in the mean.

Solution _____

(a) The update equations are

$$
\begin{aligned}
a_{n+1}(0) &= a_n(0) + \mu e(n)x(n) \\
a_{n+1}(1) &= a_n(1) + \mu e(n)x(n-1) \\
b_{n+1}(0) &= b_n(0) + \mu e(n)x^2(n) \\
b_{n+1}(1) &= b_n(1) + \mu e(n)x(n)x(n-1)
\end{aligned}
$$

where

$$
e(n) = d(n) - y(n)
$$

(b) The step size must satisfy the bound

$$
0 < \mu < \frac{2}{\lambda_{\max}}
$$

where $\lambda_{\max}$ is the maximum eigenvalue of the autocorrelation matrix

$$
\mathbf{R}_x = E\left\{
\begin{bmatrix}
x(n) \\
x(n-1) \\
x^2(n) \\
x(n)x(n-1)
\end{bmatrix}
\begin{bmatrix}
x(n) & x(n-1) & x^2(n) & x(n)x(n-1)
\end{bmatrix}
\right\}
$$

(c) In this case, the autocorrelation matrix $\mathbf{R}_x$ defined above has the form

$$
\mathbf{R}_x =
\begin{bmatrix}
E\{x^2(n)\} & E\{x(n)x(n-1)\} & 0 & 0 \\
E\{x(n)x(n-1)\}E\{x^2(n-1)\} & 0 & 0 & 0 \\
0 & 0 & E\{x^4(n)\} & E\{x^3(n)x(n-1)\} \\
0 & 0 & E\{x^3(n)x(n-1)\} & E\{x^2(n)x^2(n-1)\}
\end{bmatrix}
$$

Since there is a decoupling of the adaptive filter coefficients, we may use *two* step sizes. One for the first two coefficients, and the second for the last two coefficients.

9.22 Consider the system identification problem shown in the figure below.
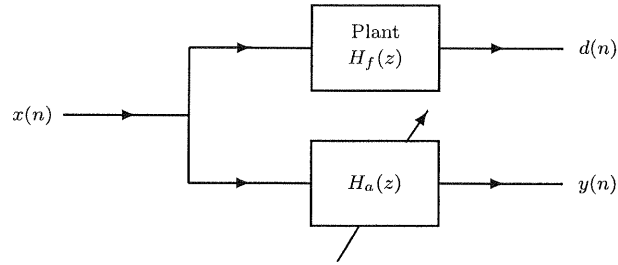
The plant has a rational system function of the form

$$H_f(z) = \frac{0.05 - 0.4z^{-1}}{1 - 1.1314z^{-1} + 0.25z^{-2}}$$

and the adaptive filter that is used to model $H_f(z)$ has two free parameters, a and b, as follows

$$H_a(z) = \frac{b}{1 - az^{-1}}$$

The input, $x(n)$, to both systems is unit variance white noise, and the goal is to find the values of a and b that minimize the mean-square error, $\xi = E\{|e(n)|^2\}$, where $e(n) = d(n) - y(n)$. This mean-square error is a bimodal function of a and b, having a global minimum at $(b, a) = (-0.311, 0.906)$ and a local minimum at $(b, a) = (0.114, -0.519)$.

(a) Write down the equations for the simplified IIR LMS adaptive filter.

(b) Repeat part (a) using the filtered signal approach.

(c) A simplification to the filtered signal approach that has been proposed by Feintuch is to ignore the feedback terms in the equation for the gradient estimates $\psi_k^a(n)$ and $\psi_k^b(n)$. For real-valued signals, this simplification is

$$\psi_k^a(n) = y(n-k) + \sum_{l=1}^{p} a_n(l)\psi_k^a(n-l) \approx y(n-k)$$

$$\psi_k^b(n) = x(n-k) + \sum_{l=1}^{p} a_n(l)\psi_k^b(n-l) \approx x(n-k)$$

Although more efficient than the filtered signal approach, this algorithm may converge to a false minimum, even when the simplified IIR LMS algorithm and the filtered signal approach converge to a minimum. Write down the adaptive filtering equations for $H_a(z)$ using Feintuch's algorithm.

(d) In order for the filter coefficients a_n and b_n in $H_a(z)$ to converge in the mean using Feintuch's algorithm, it is necessary that

$$\lim_{n \to \infty} E\{e(n)x(n)\} = 0$$

and

$$\lim_{n \to \infty} E\{e(n)y(n-1)\} = 0$$

Find the values of $E\{e(n)x(n)\}$ and $E\{e(n)y(n-1)\}$ at the global minimum of ξ. What does this imply about Feintuch's algorithm?

(e) Find the stationary point of the Feintuch adaptive filter, i.e., the value or values of a and b for which $E\{e(n)x(n)\} = 0$ and $E\{e(n)y(n-1)\} = 0$.

Solution _____

(a) The equations for the IIR LMS algorithm are as follows
1. *Coefficient Update Equations*

$$
\begin{aligned}
a_{n+1} &= a_n + \mu e(n)f(n) \\
b_{n+1} &= b_n + \mu e(n)g(n)
\end{aligned}
$$

2. *Gradient Update Equations*

$$
\begin{aligned}
f(n) &= y(n-1) + a_n f(n-1) \\
g(n) &= x(n) + a_n g(n-1)
\end{aligned}
$$

3. *Output Equation*

$$y(n) = a_n y(n-1) + b_n x(n)$$

(b) For the Feintuch algorithm we have the same equations as above except that

$$f(n) = y(n-1) \quad \text{and} \quad g(n) = x(n)$$

Thus,

$$
\begin{aligned}
a_{n+1} &= a_n + \mu e(n)y(n-1) \\
b_{n+1} &= b_n + \mu e(n)x(n)
\end{aligned}
$$

(c) In order for $E\{\theta(n)\} \to \theta^o$ we require

$$E\{e(n)y(n-1)\} = 0 \quad \text{and} \quad E\{e(n)x(n)\} = 0$$

(d) With $H_a(z)$ and $H_f(z)$ denoting the adaptive and fixed filter, respectively, the expected value of the correction term is

$$\mu E \begin{bmatrix} e(n)x(n) \\ e(n)y(n-1) \end{bmatrix}$$

Since $e(n) = d(n) - y(n)$, with

$$d(n) = \sum_{k=0}^{\infty} h_f(k)x(n-k), \qquad y(n) = \sum_{k=0}^{\infty} h_a(k)x(n-k)$$

then, using the assumption that $x(n)$ is unit variance white noise

$$E\{e(n)x(n)\} = E\{d(n)x(n)\} - E\{y(n)x(n)\} = h_f(0) - h_a(0) = 0.05 - b$$

and

$$E\left\{e(n)y(n-1)\right\} = E\left\{d(n)y(n-1)\right\} - E\left\{y(n)y(n-1)\right\}$$

Evaluating the second term we find

$$E\left\{y(n)y(n-1)\right\} = \sum_{k=0}^{\infty} h_a(k)h_a(k-1) = \sum_{k=0}^{\infty} a^k a^{k-1} = \frac{ab^2}{1-a^2}$$

For the first term,

$$E\left\{d(n)y(n-1)\right\} = \sum_{k=0}^{\infty} h_a(k)h_f(k-1)$$

To evaluate this sum, note that may be rewritten as

$$E\{d(n)y(n-1)\} = h_a(-n) * h_f(n)|_{n=1} = \frac{1}{2\pi j} \oint H_a(z^{-1})H_f(z)dz$$

Since

$$
\begin{aligned}
H_a(z^{-1})H_f(z) &= \frac{b}{1-az}\frac{0.05z^2 - 0.4z}{z^2 - 1.1314z + 0.25} \\
&= -\frac{b}{a}\frac{1}{z-\frac{1}{a}}\frac{z(0.05z - 0.4)}{(z - 0.83036)(z - 0.301094)}
\end{aligned}
$$

using residues we find

$$E\left\{d(n)y(n-1)\right\} = \frac{b}{a}\left[\frac{0.5624}{0.83036 - a^{-1}} - \frac{0.21899}{0.301094 - a^{-1}}\right]$$

Finally, at the global minimum we have

$$
\begin{aligned}
E\left\{d(n)y(n-1)\right\} &= 0.6124 \\
E\left\{y(n)y(n-1)\right\} &= 0.4891
\end{aligned}
$$

Therefore, the average value of the correction terms at the global minimum are

$$
\begin{aligned}
E\{e(n)x(n)\} &= 0.05 - b = -0.261 \\
E\{e(n)y(n-1)\} &= 0.123
\end{aligned}
$$

Thus, if the coefficients converge, then they will converge to some other point.

(e) The stationary point is the solution to the equations

$$E\left[\begin{array}{c} e(n)x(n) \\ e(n)y(n-1) \end{array}\right] = 0$$

Solving for a and b from part (d) we find that the stationary point is

$$\left[\begin{array}{c} 0.05 \\ -0.852 \end{array}\right]$$

9.23 The Hyperstable Adaptive Recursive Filter (HARF) is an IIR adaptive filtering algorithm with known convergence properties. Due to its computational complexity, a simplified version of the HARF algorithm, known as SHARF, is often used. Although the convergence properties of HARF are not preserved in the SHARF algorithm, both algorithms are similar when the filters are adapting slowly (small μ). For real-valued signals, the coefficient update equations for the SHARF algorithm are

$$a_{n+1}(k) = a_n(k) + \mu_a y(n-k)v(n)$$
$$b_{n+1}(k) = b_n(k) + \mu_b x(n-k)v(n)$$

where

$$v(n) = e(n) + \sum_{k=1}^{K} c(k)e(n-k)$$

is the error signal that has been filtered with an FIR filter $C(z)$. Suppose that the coefficients of a second-order adaptive filter

$$H(z) = \frac{b(0) + b(1)z^{-1}}{1 + a(1)z^{-1}}$$

are updated using the SHARF algorithm with

$$C(z) = 1 + c(1)z^{-1} + c(2)z^{-2}$$

(a) Write down the SHARF adaptive filter equations for $H(z)$.

(b) If the SHARF algorithm converges in the mean, then $E\{v(n)\mathbf{x}(n)\} = 0$ converges to zero where

$$\mathbf{x}(n) = [x(n),\ x(n-1),\ y(n-1)]^T$$

What does this imply about the relationship between the filter coefficients $c(1)$ and $c(2)$ and $E\{e(n)\mathbf{x}(n)\}$?

(c) What does the SHARF algorithm correspond to when $C(z) = 1$?

Solution _____

(a) The SHARF adaptive filter is

$$\begin{bmatrix} b_{n+1}(0) \\ b_{n+1}(1) \\ a_{n+1}(1) \end{bmatrix} = \begin{bmatrix} b_n(0) \\ b_n(1) \\ a_n(1) \end{bmatrix} + \mu v(n) \begin{bmatrix} x(n) \\ x(n-1) \\ y(n-1) \end{bmatrix}$$

where

$$v(n) = e(n) + c(1)e(n-1) + c(2)e(n-2)$$

(b) We require that $E\{v(n)\mathbf{x}(n)\} = 0$. Thus,

$$E \begin{bmatrix} e(n)x(n) & e(n-1)x(n) & e(n-2)x(n) \\ e(n)x(n-1) & e(n-1)x(n-1) & e(n-2)x(n-1) \\ e(n)y(n-1) & e(n-1)y(n-1) & e(n-2)y(n-1) \end{bmatrix} \begin{bmatrix} 1 \\ c(1) \\ c(2) \end{bmatrix} = 0$$

(c) Same as the Feintuch algorithm.

9.24 Modify the RLS algorithm so that the coefficients $w(k)$ satisfy the linear phase constraint, $w(k) = w(p - k)$. For example, with a five-coefficient filter, the coefficient vector $\mathbf{w}_n$ is of the form

$$\mathbf{w}_n = [w_n(0), w_n(1), w_n(2), w_n(1), w_n(0)]^T$$

Solution _____

In order to ensure that the RLS adaptive filter has linear phase, all that one needs to do is to define a new set of input signals. For a five-coefficient filter, one would use

$$\mathbf{z}(n) = \begin{bmatrix} x(n) + x(n+4) \\ x(n-1) + x(n-3) \\ x(n-2) \end{bmatrix}$$

The RLS equations are unchanged except for the replacement of $\mathbf{z}(n)$ for $\mathbf{x}(n)$.

9.25 There are many different ways that one may compare the performance of adaptive filtering algorithms. Suppose that we are interested in adaptive linear prediction and our measure of performance is the number of arithmetic operations required for the adaptive filter to converge. Let the time constant τ be used as the convergence time of the LMS algorithm. For the RLS algorithm, it is often stated that the rate of convergence is an order of magnitude faster than the LMS algorithm. Therefore, assume that the time constant for the RLS algorithm is one tenth that of the LMS algorithm.

(a) If the eigenvalues of the $p \times p$ autocorrelation matrix for $x(n)$ are

$$\lambda_1 = 1.0 \quad \text{and} \quad \lambda_2 = \cdots = \lambda_p = 0.01$$

and if we use a step size $\mu = 0.1$ for the LMS algorithm, for what order filter, p, are the RLS and LMS adaptive filters equal in terms of their computational requirements to reach convergence?

(b) For high order filters, $p \gg 1$, the computational requirements of the RLS filter become large, and the LMS algorithm becomes an attractive alternative. For what reasons might you prefer to use the RLS algorithm in spite of its increased computational cost?

Solution _____

(a) For the LMS algorithm to converge in the mean, we require

$$0 < \mu < \frac{2}{\lambda_{\max}} = 2$$

With $\mu = \mu_{\max}/10 = 1/5$, the convergence time is

$$\tau_{\text{lms}} = \frac{1}{\mu \lambda_{\min}} = \frac{5}{\lambda_{\min}} = 500$$

For an adaptive filter with p tap weights, the number of multiplications required per update is $2p + 1$ so, for convergence, we require $1000p + 500$ multiplications.

If we accept the premise that the RLS algorithm converges an order of magnitude faster than the LMS algorithm, then the time constant will be $\tau_{\text{rls}} = 50$. Since the number of multiplications required per iteration for the RLS algorithm (using $\lambda = 1$) is $2p^2 + 4p$, if

$$1000p + 500 = 50(2p^2 + 4p)$$

then the RLS and LMS adaptive filters are approximately equal in terms of the computational requirements necessary to reach convergence. Ignoring the constant term and solving for p we find that the two are approximately equivalent when $p = 8$.

(b) Reasons why one may want to consider using RLS instead of LMS are

1. Convergence of RLS is not dependent on the eigenvalues of the autocorrelation matrix of $x(n)$.
2. The misadjustment is zero if the growing window RLS algorithm is used.
3. RLS minimizes a least squares error as compared to a mean-square error.

9.26 Let $\alpha(n)$ be the the a priori error and $e(n)$ the a posteriori error in the RLS algorithm, and let
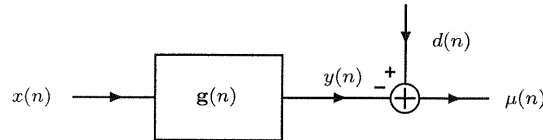
$$\mu(n) = \frac{1}{1 + \mathbf{x}^T(n)\mathbf{R}_x^{-1}(n-1)\mathbf{x}^*(n)}$$

be the scalar that is used in the calculation of the gain vector $\mathbf{g}(n)$.

(a) Show that $e(n)$ may be written in terms of $\alpha(n)$ and $\mu(n)$ by finding an explicit relation between $e(n)$, $\alpha(n)$, and $\mu(n)$. Hint: Begin with the RLS update equation for $\mathbf{w}_n$ and form the product $\mathbf{x}^T(n)\mathbf{w}_n$.

(b) Let $\mathbf{g}(n) = \mu(n)\mathbf{R}_x^{-1}(n-1)\mathbf{x}^*(n)$ be the gain vector in the RLS algorithm. Consider the time-varying filter that has coefficients $\mathbf{g}(n)$ and an input, $x(n)$, that is the same as that used in the RLS algorithm to compute the gain $\mathbf{g}(n)$, i.e.,

$$y(n) = \sum_{k=0}^{p-1} g_n(k)x(n-k)$$

For what signal, $d(n)$, will the difference between $d(n)$ and $y(n)$ be equal to $\mu(n)$ as illustrated in the figure below?

Solution _____

(a) Beginning with the RLS coefficient update equation

$$\mathbf{w}_n = \mathbf{w}_{n-1} + \alpha(n)\mathbf{g}(n)$$

we have

$$\mathbf{x}^T(n)\mathbf{w}_n = \mathbf{x}^T(n)\mathbf{w}_{n-1} + \alpha(n)\mathbf{x}^T(n)\mathbf{g}(n)$$

Therefore,

$$
\begin{aligned}
e(n) &= d(n) - \mathbf{x}^T(n)\mathbf{w}_n \\
&= d(n) - \mathbf{x}^T(n)\mathbf{w}_{n-1} - \alpha(n)\mathbf{x}^T(n)\mathbf{g}(n) \\
&= \alpha(n) - \alpha(n)\mathbf{x}^T(n)\mathbf{g}(n) \\
&= \alpha(n)\left[1 - \mu(n)\mathbf{x}^T(n)\mathbf{R}_x^{-1}(n-1)\mathbf{x}^*(n)\right] \\
&= \alpha(n)\mu(n)
\end{aligned}
$$

(b) Since

$$
\begin{aligned}
\mu(n) &= d(n) - \mathbf{x}^T(n)\mathbf{g}(n) \\
&= d(n) - \mu(n)\mathbf{x}^T(n)\mathbf{R}_x^{-1}(n-1)\mathbf{x}^*(n) \\
&= d(n) - \left[1 - \mu(n)\right]
\end{aligned}
$$

then

$$d(n) = 1$$

9.27 Suppose that the least-squares error used in the RLS algorithm is modified as follows

$$\mathcal{E}_a(n) = \sum_{i=0}^{n} \lambda^{n-i} |e(i)|^2 + \lambda^n \mathbf{w}_n^H \mathbf{w}_n$$

where $\mathbf{w}_n$ is the vector of filter coefficients at time n for a pth-order FIR adaptive filter and

$$e(i) = d(i) - \mathbf{w}_n^T \mathbf{x}(i)$$

Derive the equations for the optimal least squares filter $\mathbf{w}_n$ that minimizes $\mathcal{E}_a(n)$ for each n.

Solution _____

This error differs from the least squares error used in the RLS algorithm in the additive term $\lambda^n \mathbf{w}_n^H \mathbf{w}_n$. To derive the equations for the optimal least squares filter as we did in Eq. (9.91), we differentiate with respect to $w_n^*(k)$ for $k = 0, 1, \ldots, p$, and set the result equal to zero,

$$\frac{\partial \mathcal{E}_a(n)}{\partial w_n^*(k)} = \sum_{i=0}^{n} \lambda^{n-i} e(i) \frac{\partial e^*(i)}{\partial w_n^*(k)} + \lambda^n w_n(k)$$

$$= -\sum_{i=0}^{n} \lambda^{n-i} e(i) x^*(i-k) + \lambda^n w_n(k) = 0 \quad ; \quad k = 0, 1, \ldots, p$$

Substituting for the error we have

$$-\sum_{i=0}^{n} \lambda^{n-i} \left\{ d(i) - \sum_{l=0}^{p} w_n(l) x(i-l) \right\} x^*(i-k) + \lambda^n w_n(k) = 0$$

Interchanging the order of summation and rearranging terms yields

$$\sum_{l=0}^{p} w_n(l) \left[\sum_{i=0}^{n} \lambda^{n-i} x(i-l) x^*(i-k) \right] + \lambda^n w_n(k) = \sum_{i=0}^{n} \lambda^{n-i} d(i) x^*(i-k) \quad ; \quad k = 0, 1, \ldots, p$$

Expressing these in matrix form as in Eq. (9.93) we have

$$\mathbf{R}_x(n) \mathbf{w}_n + \lambda^n \mathbf{w}_n = \mathbf{r}_{dx}(n)$$

where $\mathbf{R}_x(n)$ is a $(p+1) \times (p+1)$ exponentially-weighted deterministic autocorrelation matrix for $x(n)$

$$\mathbf{R}_x(n) = \sum_{i=0}^{n} \lambda^{n-i} \mathbf{x}^*(i) \mathbf{x}^T(i)$$

and $\mathbf{r}_{dx}(n)$ is the deterministic cross-correlation between $d(n)$ and $x(n)$,

$$\mathbf{r}_{dx}(n) = \sum_{i=0}^{n} \lambda^{n-i} d(i) \mathbf{x}^*(i)$$

Alternatively, this matrix equation may be written as

$$\left[\mathbf{R}_x(n) + \lambda^n \mathbf{I} \right] \mathbf{w}_n = \mathbf{r}_{dx}(n)$$