Advanced Digital Signal Processing - Introduction

LECTURE-2

AP9211- ADVANCED DIGITAL SIGNAL PROCESSING

UNIT I DISCRETE RANDOM SIGNAL PROCESSING

UNIT II SPECTRUM ESTIMATION
Estimation of spectra from finite duration signals, Nonparametric methods - Periodogram, Modified periodogram, Bartlett, Welch and Blackman-Tukey methods, Parametric methods – ARMA, AR and MA model based spectral estimation, Solution using Levinson-Durbin algorithm

UNIT III LINEAR ESTIMATION AND PREDICTION
Linear prediction – Forward and Backward prediction, Solution of Prony’s normal equations, Least mean-squared error criterion, Wiener filter for filtering and prediction, FIR and IIR Wiener filters, Discrete Kalman filter
UNIT IV ADAPTIVE FILTERS
FIR adaptive filters – adaptive filter based on steepest descent method-
Widrow-Hopf LMS algorithm, Normalized LMS algorithm, Adaptive
channel equalization, Adaptive echo cancellation, Adaptive noise
cancellation, RLS adaptive algorithm.

UNIT V MULTIRATE DIGITAL SIGNAL PROCESSING
Mathematical description of change of sampling rate – Interpolation and
Decimation, Decimation by an integer factor, Interpolation by an integer
factor, Sampling rate conversion by a rational factor, Poly phase filter
structures, Multistage implementation of multirate system, Application to
sub-band coding – Wavelet transform

REFERENCES:
John Wiley and Sons, Inc, Singapore, 2002
Pearson Education, 2002
The DFT and FFT

• The discrete-time Fourier transform
  – is a very useful tool,
  • but the frequency variable is a continuous variable,
  – and hence not well suited for computation

\[ X(e^{j\omega}) = \sum_{n=-\infty}^{\infty} x(n)e^{-j\omega n} \]

The DFT and FFT

• For finite length sequences
  – there is another representation called the Discrete Fourier Transform (DFT)
    • which is a function of a discrete frequency variable k

\[ X(k) = \sum_{n=0}^{N-1} x(n)e^{-j2\pi kn/N} \]
The DFT and FFT

- If two finite length sequences of length $N_1$ and $N_2$
  - have to be linearly convolved using the DFT,
  - then the DFTs have to be at least of length $N_1+N_2-1$.

- The length of a DFT is normally a power of 2,
  - so that a fast implementation of the DFT
  - can be used the so called fast Fourier transform (FFT)

- While the DTFT has a time complexity of $N^2$,
  - the DFT has a time complexity of $N \log_2(N)$

Linear Algebra

- Much of discrete-time signal processing
  - is done with finite length sequences
- These sequences
  - can be ideally represented by a vector, containing the samples of these sequences
  \[
  \mathbf{x} = \begin{bmatrix}
  x(0) \\
  x(1) \\
  \vdots \\
  x(N-1)
  \end{bmatrix}
  \]
- This
  - allows discrete-time signal processing to access the power of linear algebra to solve real world problems
Vectors

- A vector is an array of real-valued or complex-valued numbers or functions. Normally we assume that the vectors are column vectors

\[
x = \begin{bmatrix}
x_1 \\
x_2 \\
\vdots \\
x_N
\end{bmatrix}
\]

- Hence the transpose of a vector is a row vector

\[
x^T = [x_1, x_2, \ldots, x_N]
\]

- The Hermitian transpose is the complex conjugate of the transpose

\[
x^H = (x^T)^* = [x_1^*, x_2^*, \ldots, x_N^*]
\]

---

Vectors

- A finite length sequence \(x(n)\) that is equal to zero outside the interval \([0,N-1]\) may be represented in vector form

\[
x = \begin{bmatrix}
x(0) \\
x(1) \\
\vdots \\
x(N-1)
\end{bmatrix}
\]

- Also it is convenient to define the vector \(x(n)\) as

\[
x(n) = \begin{bmatrix}
x(n) \\
x(n-1) \\
\vdots \\
x(n-N+1)
\end{bmatrix}
\]
Vectors

- Often the goal will be to minimize the length magnitude of an error vector.

- Hence the magnitude needs to be defined.

- Usually this is the $L_2$ norm

\[
\|x\|_2 = \left( \sum_{i=1}^{N} |x_i|^2 \right)^{1/2}
\]

Vectors

- Other useful norms are the $L_1$ norm

\[
\|x\|_1 = \sum_{i=1}^{N} |x_i|
\]

- And the $L_{\infty}$ norm

\[
\|x\|_{\infty} = \max_i |x_i|
\]

- In this module the $L_2$ will be used almost exclusively and hence $\|x\|$ implies the $L_2$ norm
Vectors

• A vector might be normalized to have unit magnitude by dividing by its norm
  \[ v_x = \frac{x}{|x|} \]

• \( v_x \) is now a unit norm vector pointing in the direction of \( x \)
• For a signal vector the norm squared represents its energy
  \[ |x|^2 = \sum_{n=0}^{N-1} |x(n)|^2 \]

• The norm can also be used to measure the distance between two vectors (RMS of error signal)
  \[ d(x, y) = |x - y| = \left( \sum_{i=1}^{N} |x_i - y_i|^2 \right)^{1/2} \]

Vectors

• Given two complex vectors, the inner product is defined as
  \[ \langle a, b \rangle = a^H b = \sum_{i=1}^{N} a_i^* b_i \]

• For real vectors the inner product becomes
  \[ \langle a, b \rangle = a^T b = \sum_{i=1}^{N} a_i b_i \]

• The inner product defines the geometrical relationship between two vectors, where \( \theta \) is the angle between the two vectors.
  \[ \langle a, b \rangle = |a| |b| \cos \theta \]

• Two nonzero vectors are said to be orthogonal if their inner product is zero
  \[ \langle a, b \rangle = 0 \]
Vectors Example Inner Product

- Consider the two unit norm vectors
\[ \mathbf{v}_1 = \begin{bmatrix} 0 \\ 1 \end{bmatrix} ; \quad \mathbf{v}_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix} \]

- The inner product is
\[ \langle \mathbf{v}_1, \mathbf{v}_2 \rangle = \frac{1}{\sqrt{2}} = \cos \theta \]

- Hence \( \theta = \pi/4 \)
\[ \mathbf{v}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix} ; \quad \mathbf{v}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \]

- If however, The inner product is 0 and the vectors are orthogonal to each other

Vectors

- Since the cosine is bounded by +/- 1 it follows
\[ \langle \mathbf{a}, \mathbf{b} \rangle = \| \mathbf{a} \| \| \mathbf{b} \| \cos \theta \]

- Where the equal sign holds if the vectors are co-linear (a=\( \alpha \mathbf{b} \)). This is known as the Cauchy-Schwarz inequality
\[ |\langle \mathbf{a}, \mathbf{b} \rangle| \leq \| \mathbf{a} \| \| \mathbf{b} \| \]

- Another useful inequality is
\[ 2|\langle \mathbf{a}, \mathbf{b} \rangle| \leq \| \mathbf{a} \|^2 + \| \mathbf{b} \|^2 \]

- This follows from
\[ |\mathbf{a} \pm \mathbf{b}|^2 \geq 0 \]
\[ |\mathbf{a} \pm \mathbf{b}|^2 = |\mathbf{a}|^2 \pm 2\langle \mathbf{a}, \mathbf{b} \rangle + |\mathbf{b}|^2 \geq 0 \]
\[ 2|\langle \mathbf{a}, \mathbf{b} \rangle| \leq |\mathbf{a}|^2 + |\mathbf{b}|^2 \]
Vectors

- The inner product can be used to formulate the output of an LSI filter in a concise way
  \[ y(n) = \sum_{k=0}^{N-1} h(k)x(n - k) \]

- Let
  \[
  h = \begin{bmatrix}
  h(0) \\
  h(1) \\
  \vdots \\
  h(N-1)
  \end{bmatrix},
  x(n) = \begin{bmatrix}
  x(n) \\
  x(n-1) \\
  \vdots \\
  x(n-N+1)
  \end{bmatrix}
  \]

- It follows
  \[ y(n) = h^T x(n) \]

Linear Independence, Vector Spaces, & Basis Vectors

- A set of n vectors is said to be linearly independent if it implies that all alphas are zero
  \[ \alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n = 0 \]

  - If there is a set of nonzero alphas so that this holds, the set is called linearly dependent

  - For a set of linearly dependent vectors there exist therefore a nonzero set of betas such that this holds
    \[ x_1 = \beta_2 v_2 + \beta_3 v_3 + \cdots + \beta_n v_n \]

  - For vectors of dimension N, no more than N vectors can be linearly independent
Linear Independence Example

• Given the following pair of vectors
  \[ \mathbf{v}_1 = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} \]

• If we assume that there are alphas such that
  \[ \alpha_1 \mathbf{v}_1 + \alpha_2 \mathbf{v}_2 = \alpha_1 \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} + \alpha_2 \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \]

• Then it follows that
  \[ \alpha_1 + \alpha_2 = 0 \]
  \[ 2\alpha_1 = 0 \]

• Which means that
  \[ \alpha_1 = \alpha_2 = 0 \]

• Hence the vectors are linearly independent

Linear Independence Example

• However, adding an additional vector
  \[ \mathbf{v}_1 = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \quad \mathbf{v}_3 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \]

• Results in a linearly dependent set since
  \[ \mathbf{v}_1 = \mathbf{v}_2 + 2\mathbf{v}_3 \]
Linear Independence, Vector Spaces, and Basis Vectors

• Given a set of N vectors \( V = \{ v_1, v_2, \ldots, v_N \} \)

• Consider the set of all vectors that may be formed from a linear combination of these vectors
  \[ v = \sum_{i=1}^{N} a_i v_i \]

• This set forms a vector space and the vectors are said to span the space.

• If the set is linearly independent, then they are said to form a basis for the space.

• The number of vectors in the basis are referred to as the dimension of the space.

Linear Independence, Vector Spaces, and Basis Vectors

For example, the set of all real vectors of the form \( x = [x_1, x_2, \ldots, x_N]^T \) forms an \( N \)-dimensional vector space, denoted by \( \mathbb{R}^N \), that is spanned by the basis vectors

\[
\begin{align*}
  u_1 &= [1, 0, 0, \ldots, 0]^T \\
  u_2 &= [0, 1, 0, \ldots, 0]^T \\
  \vdots \\
  u_N &= [0, 0, 0, \ldots, 1]^T
\end{align*}
\]

In terms of this basis, any vector \( v = [v_1, v_2, \ldots, v_N]^T \) in \( \mathbb{R}^N \) may be uniquely decomposed as follows

\[
v = \sum_{i=1}^{N} v_i u_i
\]

It should be pointed out, however, that the basis for a vector space is not unique.
Matrices

• An \( n \times m \) matrix is an array of numbers (real or complex) or functions having \( n \) rows and \( m \) columns

\[
A = \{a_{ij}\} = \begin{bmatrix}
  a_{11} & a_{12} & a_{13} & \cdots & a_{1m} \\
  a_{21} & a_{22} & a_{23} & \cdots & a_{2m} \\
  a_{31} & a_{32} & a_{33} & \cdots & a_{3m} \\
  \vdots & \vdots & \vdots & \ddots & \vdots \\
  a_{n1} & a_{n2} & a_{n3} & \cdots & a_{nm}
\end{bmatrix}
\]

Matrices

• Sometimes it is convenient to let a matrix have an infinite number of rows or columns
• The prime example for this is the convolution sum written as a matrix multiplication.

\[
y(n) = h^T x(n) = x^T(n) h
\]

• Let \( h(n) \) be the unit sample response of a LTI FIR filter, then the output \( y(n) \) can be written as an inner product

\[
h = \begin{bmatrix}
  h(0) \\
  h(1) \\
  \vdots \\
  h(N - 1)
\end{bmatrix} \quad x(n) = \begin{bmatrix}
  x(n) \\
  x(n - 1) \\
  \vdots \\
  x(n - N + 1)
\end{bmatrix}
\]
Matrices

- If \( x(n) = 0 \) for \( n < 0 \), then we can express \( y(n) \) for \( n \geq 0 \) as
  \[
  X_0 h = y
  \]
- Where \( X_0 \) is a convolution matrix

\[
X_0 = \begin{bmatrix}
x(0) & 0 & 0 & \cdots & 0 \\
x(1) & x(0) & 0 & \cdots & 0 \\
x(2) & x(1) & x(0) & \cdots & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
x(N-1) & x(N-2) & x(N-3) & \cdots & x(0)
\end{bmatrix}
\]
- And \( y \)

\[
y = [y(0), y(1), y(2), \ldots]^T.
\]

Matrices

- Sometimes it is convenient to write the matrix as a set of column or row vectors

\[
A = [c_1, c_2, \ldots, c_m]
\]

\[
A = \begin{bmatrix}
x_1^H \\
x_2^H \\
\vdots \\
x_n^H
\end{bmatrix}
\]

- A matrix can also be partitioned into sub-matrices

\[
A = \begin{bmatrix}
A_{11} & A_{12} \\
A_{21} & A_{22}
\end{bmatrix}
\]
Matrices Partitioning Example

• Consider the 3x3 matrix

\[
A = \begin{bmatrix}
1 & 0 & 0 \\
0 & 2 & 1 \\
0 & 1 & 2
\end{bmatrix} = \begin{bmatrix}
1 & 0 & 0 \\
0 & 2 & 1 \\
0 & 1 & 2
\end{bmatrix}
\]

• This matrix may be partitioned as

\[
A = \begin{bmatrix}
1 & 0^T \\
0 & A_{22}
\end{bmatrix}
\]

• Where \(0 = [0, 0]^T\)

• And \(A_{22} = \begin{bmatrix}
2 & 1 \\
1 & 2
\end{bmatrix}\)

Matrices

• If \(A\) is a \(n \times m\) matrix,
  – then its transpose \(A^T\) is the \(m \times n\) matrix that is formed by interchanging the rows and the columns of \(A\)

• For a square matrix if \(A\) is equal to its transpose,
  – it is called a symmetric matrix

\[
A = A^T
\]

• For complex matrices,
  – the Hermitian transpose is the complex conjugate of the transpose

\[
A^H = (A^*)^T = (A^T)^*
\]
Matrices

- For a complex square matrix if \( A \) is equal to its Hermitian transpose it is called a Hermitian matrix

\[ A = A^H \]

- Properties of the Hermitian transpose are

1. \((A + B)^H = A^H + B^H\)
2. \((A^H)^H = A\)
3. \((AB)^H = B^H A^H\)

- Equivalent properties for the transpose may be obtained by replacing the Hermitian transpose with the regular transpose

Matrix Inverse

- Let \( A \) be an \( n \times m \) matrix that is partitioned in terms of its \( m \) column vectors

\[ A = [c_1, c_2, \ldots, c_m] \]

- The rank of \( A \), \( \rho(A) \) is defined to be the number of linearly independent columns in \( A \), i.e., the number of independent vectors in the above set.

\[ \rho(A) = \rho(A^H) \]

- One of the properties of the rank is, that the rank of a matrix is equal to the rank of its Hermitian transpose
Matrix Inverse

- Therefore if \( A \) is partitioned in terms of its \( n \) row vectors, then the rank is equivalently equal to the number of linearly independent row vectors.

\[
A = \begin{bmatrix}
    r_1^T \\
    r_2^T \\
    \vdots \\
    r_n^T
\end{bmatrix}
\]

- A useful property of the rank of a matrix is the following:

Property. The rank of \( A \) is equal to the rank of \( AA^T \) and \( A^T A \).

\[
\rho(A) = \rho(AA^T) = \rho(A^T A)
\]

- Since the rank is equal to the number of independent rows and the number of independent columns it follows that for an \( m \times n \) matrix

\[
\rho(A) \leq \min(m, n)
\]

If \( A \) is an \( m \times n \) matrix and \( \rho(A) = \min(m, n) \) then \( A \) is said to be of full rank.

Matrix Inverse

- If \( A \) is a square matrix of full rank, then there exists a unique matrix \( A^{-1} \) called the inverse of \( A \) such that

\[
A^{-1} A = AA^{-1} = I
\]

- Where \( I \) is the identity matrix

\[
I = \begin{bmatrix}
    1 & 0 & 0 & \cdots & 0 \\
    0 & 1 & 0 & \cdots & 0 \\
    0 & 0 & 1 & \cdots & 0 \\
    \vdots & \vdots & \vdots & \ddots & \vdots \\
    0 & 0 & 0 & \cdots & 1
\end{bmatrix}
\]

- In this case \( A \) is called invertible or nonsingular.
- If \( A \) is not of full rank the it is said noninvertible or singular and \( A \) does not have an inverse.
Matrix Inverse

- Some properties of the matrix inverse are as follows. If $A$ and $B$ are invertible then

$$(AB)^{-1} = B^{-1}A^{-1}$$

- Second

$$(A^H)^{-1} = (A^{-1})^H$$

Matrix Inverse

- Finally a formula that is useful for inverting matrices that arise in adaptive filtering (RLS) algorithms is the Matrix inversion lemma

$$(A + BCD)^{-1} = A^{-1} - A^{-1}B(C^{-1} + DA^{-1}B)^{-1}DA^{-1}$$  \hspace{1cm} (2.28)$$

where it is assumed that $A$ is $n \times n$, $B$ is $n \times m$, $C$ is $m \times m$, and $D$ is $m \times n$ with $A$ and $C$ nonsingular matrices. A special case of this lemma occurs when $C = I$, $B = u$, and $D = v^H$ where $u$ and $v$ are $n$-dimensional vectors. In this case the lemma becomes

$$(A + uv^H)^{-1} = A^{-1} - A^{-1}uv^HA^{-1} \frac{1}{1 + vv^HA^{-1}u}$$  \hspace{1cm} (2.29)$$

which is sometimes referred to as Woodbury's Identity [5]. As a special case, note that for $A = I$, Eq. (2.29) becomes

$$(I + uv^H)^{-1} = I - \frac{1}{1 + vv^H}$$  \hspace{1cm} (2.30)$$
Determinant And Trace

- If $A = a_{11}$ is a 1x1 matrix, then its determinate $\det(A)$ is defined as $\det(A) = a_{11}$

- The determinant of an n x n matrix is defined recursively in terms of the determinants of (n-1) x (n-1) matrices as follows. For any $j$:
  \[
  \det(A) = \sum_{i=1}^{n} (-1)^{i+j} a_{ij} \det(A_{ij})
  \]

- Where $A_{ij}$ is the (n-1) x (n-1) matrix that is formed by deleting the $i^{th}$ row and the $j^{th}$ column of $A$

- The determinant also has an important geometric interpretation as the area of a parallelogram, and more generally as the volume of a higher-dimensional parallelepiped.

Determinant Example

- For a 2 x 2 matrix
  \[
  A = \begin{bmatrix}
  a_{11} & a_{12} \\
  a_{21} & a_{22}
  \end{bmatrix}
  \]
  The determinant is
  \[
  \det(A) = a_{11}a_{22} - a_{12}a_{21}
  \]

- And for the 3 x 3 matrix
  \[
  A = \begin{bmatrix}
  a_{11} & a_{12} & a_{13} \\
  a_{21} & a_{22} & a_{23} \\
  a_{31} & a_{32} & a_{33}
  \end{bmatrix}
  \]
  The determinant is
  \[
  \det(A) = a_{11} \det \begin{bmatrix}
  a_{22} & a_{23} \\
  a_{32} & a_{33}
  \end{bmatrix} - a_{12} \det \begin{bmatrix}
  a_{21} & a_{23} \\
  a_{31} & a_{33}
  \end{bmatrix} + a_{13} \det \begin{bmatrix}
  a_{21} & a_{22} \\
  a_{31} & a_{32}
  \end{bmatrix}
  \]
  \[
  = a_{11}[a_{22}a_{33} - a_{23}a_{32}] - a_{12}[a_{21}a_{33} - a_{23}a_{31}] + a_{13}[a_{21}a_{32} - a_{22}a_{31}]
  \]
Determinant And Trace

• The determinant may be used to determine whether or not a matrix is invertible

\[ \text{Property. An } n \times n \text{ matrix } A \text{ is invertible if and only if its determinant is nonzero, } \det(A) \neq 0 \]

• Additional properties of the determinant are listed below, where it is assumed that \( A \) and \( B \) are \( n \times n \) matrices

1. \( \det(AB) = \det(A) \det(B) \).
2. \( \det(A^T) = \det(A) \).
3. \( \det(\alpha A) = \alpha^n \det(A) \), where \( \alpha \) is a constant.
4. \( \det(A^{-1}) = \frac{1}{\det(A)} \), assuming that \( A \) is invertible.

Determinant And Trace

• Another useful function of a matrix is the trace.

• Often the trace of a covariance matrix needs to minimized.

• Given an \( n \times n \) matrix \( A \) the trace is the sum of the terms along the diagonal

\[ \text{tr}(A) = \sum_{i=1}^{n} a_{ii} \]
Linear Equations

• Signal modeling, Wiener filtering and spectrum estimation require finding the solution to a set of linear equations

Consider the following set of \( n \) linear equations in the \( m \) unknowns \( x_i, i = 1, 2, \ldots, m \),

\[
\begin{align*}
& a_{11}x_1 + a_{12}x_2 + \cdots + a_{1m}x_m = b_1 \\
& a_{21}x_1 + a_{22}x_2 + \cdots + a_{2m}x_m = b_2 \\
& \vdots \\
& a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nm}x_m = b_n
\end{align*}
\]

• This can be efficiently written as \( Ax = b \)

• And interpreted as a linear combination of the column vectors of \( A \)

\[
b = \sum_{i=1}^{n} x_i a_i
\]

Linear Equations-Square Matrix

• For a square matrix \( n=m \)
  – If \( A \) is nonsingular, there exists a unique solution \( x = A^{-1}b \)
  – If \( A \) is singular, there exist no solution (inconsistent equations) or there exist many solutions (at least one degree of freedom is still left)

• If \( A \) is singular, then the columns of \( A \) are linearly dependent and there exists a nonzero solution to the homogeneous equations. In fact there will be \( k=n-r(A) \) linearly independent solution to the homogeneous equations \( Az = 0 \)
Linear Equations-Square Matrix

- If there is at least one vector $x_0$ that solves
  \[ x = A^{-1}b \]
- Then any vector of the form will also be a solution where are linearly independent solutions of
  \[ x = x_0 + \alpha_1 z_1 + \cdots + \alpha_k z_k \]
  \[ z_i, \ i = 1, 2, \ldots, k \]
  \[ Az = 0 \]

Linear Equations-Square Matrix Example

Consider the following pair of equations in two unknowns $x_1$ and $x_2$,
\[ x_1 + x_2 = 1 \]
\[ x_1 + x_2 = 2 \]

In matrix form, these equations are
\[ \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \end{bmatrix} \]

Clearly, the matrix $A$ is singular, $\det(A) = 0$, and no solution exists. However, if the second equation is modified so that
\[ \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \]
then there are many solutions. Specifically, note that for any constant $\alpha$, the vector
\[ x = \begin{bmatrix} 1 \\ 0 \end{bmatrix} + \alpha \begin{bmatrix} 1 \\ -1 \end{bmatrix} \]
will satisfy these equations.
Special Matrix Forms

• Diagonal matrix
\[ A = \begin{bmatrix}
  a_{11} & 0 & 0 & \cdots & 0 \\
  0 & a_{22} & 0 & \cdots & 0 \\
  0 & 0 & a_{33} & \cdots & 0 \\
  \vdots & \vdots & \vdots & \ddots & \vdots \\
  0 & 0 & 0 & \cdots & a_{nn}
\end{bmatrix} \]

• Which can be written as
\[ A = \text{diag} \{a_{11}, a_{22}, \ldots, a_{nn}\} \]

• Identity matrix
\[ I = \text{diag}(1, 1, \ldots, 1) = \begin{bmatrix}
  1 & 0 & 0 & \cdots & 0 \\
  0 & 1 & 0 & \cdots & 0 \\
  0 & 0 & 1 & \cdots & 0 \\
  \vdots & \vdots & \vdots & \ddots & \vdots \\
  0 & 0 & 0 & \cdots & 1
\end{bmatrix} \]

Special Matrix Forms

• If the entries on the diagonal become matrices, then \( A \) is said to be a block diagonal matrix
\[ A = \begin{bmatrix}
  A_{11} & 0 & \cdots & 0 \\
  0 & A_{22} & \cdots & 0 \\
  \vdots & \vdots & \ddots & \vdots \\
  0 & 0 & \cdots & A_{nn}
\end{bmatrix} \]

• The exchange matrix reverses the order of entries
\[ J = \begin{bmatrix}
  0 & \cdots & 0 & 1 \\
  0 & \cdots & 1 & 0 \\
  \vdots & \vdots & \vdots & \vdots \\
  1 & \cdots & 0 & 0
\end{bmatrix} \]
\[ J^T A J = \begin{bmatrix}
  a_{11} & a_{12} & a_{13} \\
  a_{21} & a_{22} & a_{23} \\
  a_{31} & a_{32} & a_{33}
\end{bmatrix} \]
\[ A J = \begin{bmatrix}
  d_{11} & d_{12} & d_{13} \\
  d_{21} & d_{22} & d_{23} \\
  d_{31} & d_{32} & d_{33}
\end{bmatrix} \]
\[ J^T A J = \begin{bmatrix}
  e_{11} & e_{12} & e_{13} \\
  e_{21} & e_{22} & e_{23} \\
  e_{31} & e_{32} & e_{33}
\end{bmatrix} \]
Special Matrix Forms

- An upper triangular matrix is a square matrix in which all of the terms below the diagonal are equal to zero

\[
A = \begin{bmatrix}
    a_{11} & a_{12} & a_{13} & a_{14} \\
    0 & a_{22} & a_{23} & a_{24} \\
    0 & 0 & a_{33} & a_{34} \\
    0 & 0 & 0 & a_{44}
\end{bmatrix}
\]

- An lower triangular matrix is a square matrix in which all of the terms above the diagonal are equal to zero

- Transposing a lower triangular matrix results in an upper triangular matrix and vice versa

1. The determinant of a lower triangular matrix or an upper triangular matrix is equal to the product of the terms along the diagonal

\[
\det(A) = \prod_{i=1}^{n} a_{ii}
\]

2. The inverse of an upper (lower) triangular matrix is upper (lower) triangular.

3. The product of two upper (lower) triangular matrices is upper (lower) triangular.
Special Matrix Forms

• A very important matrix is a Toeplitz matrix, which is completely specified when the first row and the first column is known. Note that a convolution matrix is a Toeplitz matrix.

An \( n \times n \) matrix \( A \) is said to be Toeplitz if all of the elements along each of the diagonals have the same value, i.e.,

\[ a_{ij} = a_{i+1,j+1} \quad \text{for all } i < n \text{ and } j < n \]

An example of a \( 4 \times 4 \) Toeplitz matrix is

\[
A = \begin{bmatrix}
1 & 3 & 5 & 7 \\
2 & 1 & 3 & 5 \\
4 & 2 & 1 & 3 \\
6 & 4 & 2 & 1
\end{bmatrix}
\]

Special Matrix Forms

• If a Toeplitz matrix is symmetric, or Hermitian in the case of a complex matrix, then all the elements are completely determined by either the first row or the first column. Note that an autocorrelation matrix is a Hermitian Toeplitz matrix.

\[
A = \begin{bmatrix}
1 & 3 & 5 & 7 \\
3 & 1 & 3 & 5 \\
5 & 3 & 1 & 3 \\
7 & 5 & 3 & 1
\end{bmatrix}
\]

\[
A = \text{Toep}\{1, j, 1-j\} = \begin{bmatrix} 1 & -j & 1+j \\ j & 1 & -j \\ 1-j & j & 1 \end{bmatrix}
\]
Special Matrix Forms

• A real $n \times n$ matrix is said to be orthogonal if the columns and rows are orthonormal

\[
A = [a_1, a_2, \ldots, a_n]
\]

\[
a_i^T a_j = \begin{cases} 
1 & \text{for } i = j \\
0 & \text{for } i \neq j
\end{cases}
\]

• Hence for orthogonal matrices holds

\[
A^{-1} = A^T \\
A^T A = I
\]

• In the case of a complex matrix, such matrices are called unitary matrices

\[
a^*_i a_j = \begin{cases} 
1 & \text{for } i = j \\
0 & \text{for } i \neq j
\end{cases}
\]

\[
A^* A = I \\
A^{-1} = A^*
\]

Quadratic and Hermitian Forms

• The quadratic form of a real symmetric $n \times n$ matrix $A$ is the scalar defined by, where $x$ is a real vector

\[
Q_A(x) = x^T A x = \sum_{i=1}^{n} \sum_{j=1}^{n} x_i a_{ij} x_j
\]

• For example

\[
A = \begin{bmatrix} 3 & 1 \\ 1 & 2 \end{bmatrix} \\
Q_A(x) = x^T A x = 3x_1^2 + 2x_1 x_2 + 2x_2^2
\]

• For a Hermitian matrix

\[
Q_A(x) = x^H A x = \sum_{i=1}^{n} \sum_{j=1}^{n} x_i^* a_{ij} x_j
\]

• If the quadratic form of a matrix $A$ is positive for all nonzero vectors $x$, then the matrix is called positive definite, written as $A > 0$

\[
Q_A(x) > 0
\]
Quadratic and Hermitian Forms

- For example this matrix is positive definite, since the quadratic form is positive for any non-zero vector \( \mathbf{x} \)

\[
A = \begin{bmatrix}
2 & 0 \\
0 & 3
\end{bmatrix}
\]

\[
Q_A(\mathbf{x}) = [x_1, x_2] \begin{bmatrix}
2 & 0 \\
0 & 3
\end{bmatrix} \begin{bmatrix}
x_1 \\
x_2
\end{bmatrix} = 2x_1^2 + 3x_2^2
\]

Quadratic and Hermitian Forms

- If the quadratic form is non-negative for all nonzero vectors \( \mathbf{x} \), then \( A \) is said to be positive semidefinite

\[
Q_A(\mathbf{x}) \geq 0
\]

- For example

\[
A = \begin{bmatrix}
2 & 0 \\
0 & 0
\end{bmatrix}
\]

since the quadratic form is zero for \( \mathbf{x} = [0, x_2]^T \)

\[
Q_A(\mathbf{x}) = 2x_1^2 \geq 0
\]