6. Advanced Linear Algebra

[1] D. G. Zill, W. S. Wright, and J. J. Ding, Engineering Mathematics, Metric Edition, Cengage Learning, Taipei, Taiwan, 2019.

[2] A. V. Oppenheim and R. W. Schafer, *Discrete-Time Signal Processing*, London: Prentice-Hall, 3rd ed., 2010.

Linear Algebra

- Linear Algebra
  - Operations
    - Addition, Multiplication, Scalar Multi. Transpose, Hermitian, Permutation, Determinant, Rank, Inverse (Sec. 6-1)
    - Kroneker Product (Sec. 6-2)
    - Element-Wise Product (Sec. 6-2)
    - Functions of Matrices (Sec. 6-4)
    - Generalized Norm (Sec. 6-5)
  - Eigenvectors and Eigenvalues
    - Definition and Computation (Sec. 6-1)
    - Jordan Canonical Form (Sec. 6-3)
    - Singular Value Decomposition (Chap. 8)
  - Applications
    - Discrete Linear System (Sec. 6-1)
    - Markov Model (Sec. 6-6)
    - Discrete Basis Approximate (Chap. 7)
    - Principal Component Analysis (Chap. 8)
    - Independent Component Analysis (Sec. 9-6)
  - Discrete Transforms (Sec. 6-7)

6.1 Review of Linear Algebra

6.1.1 Matrix

Scalar: x_1

Vector: $\begin{bmatrix} x_1 & x_2 & \cdots & \cdots & x_N \end{bmatrix}$

1D

Matrix: $\begin{bmatrix} a_{1,1} & a_{1,2} & \cdots & \cdots & a_{1,N} \\ a_{2,1} & a_{2,2} & \cdots & \cdots & a_{2,N} \\ \vdots & \vdots & \ddots & & \vdots \\ \vdots & \vdots & & \ddots & \vdots \\ a_{M,1} & a_{M,2} & \cdots & \cdots & a_{M,N} \end{bmatrix}$

2D

$a_{m,n}$: entry (also called an element or a scalar)

$$\begin{bmatrix} a_{1,1} & a_{1,2} & \cdots & \cdots & a_{1,N} \\ a_{2,1} & a_{2,2} & \cdots & \cdots & a_{2,N} \\ \vdots & \vdots & \ddots & & \vdots \\ \vdots & \vdots & & \ddots & \vdots \\ a_{M,1} & a_{M,2} & \cdots & \cdots & a_{M,N} \end{bmatrix}$$

$M = N$: square matrix

$M = 1$: row vector, $N = 1$: column vector, $M = N = 1$: scalar

For a square matrix,

if $a_{m,n} = 0$ for $m \neq n$ $\rightarrow$ diagonal matrix

if $a_{m,n} = 0$ for $m \neq n$, $a_{n,n} = 1$ $\rightarrow$ identity matrix (denoted by **I**)

if $a_{m,n} = 0$ for $m > n$ $\rightarrow$ upper triangular matrix

if $a_{m,n} = 0$ for $m < n$ $\rightarrow$ lower triangular matrix

A linear system can be expressed as a matrix operation

$$\begin{cases} 2x+3y=1 \\ x+4y=2 \end{cases} \Longrightarrow \overset{A}{\begin{bmatrix} 2 & 3 \\ 1 & 4 \end{bmatrix}} \overset{x}{\begin{bmatrix} x \\ y \end{bmatrix}} = \overset{y}{\begin{bmatrix} 1 \\ 2 \end{bmatrix}}$$

$$\begin{cases} a_{1,1}x_1 + a_{1,2}x_2 + \cdots + a_{1,N}x_N = b_1 \\ a_{2,1}x_1 + a_{2,2}x_2 + \cdots + a_{2,N}x_N = b_2 \\ \qquad\qquad\qquad \vdots \\ a_{N,1}x_1 + a_{N,2}x_2 + \cdots + a_{N,N}x_N = b_N \end{cases}$$

$$\Longrightarrow \begin{bmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,N} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,N} \\ \vdots & \vdots & \ddots & \vdots \\ a_{N,1} & a_{N,2} & \cdots & a_{N,N} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_N \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_N \end{bmatrix}$$

6.1.2 Matrix Operations

(1) Addition

$$\mathbf{C} = \mathbf{A} + \mathbf{B} \quad \text{where } c_{m,n} = a_{m,n} + b_{m,n}.$$

(2) Multiplication

$$\mathbf{C} = \mathbf{AB} \qquad \text{where} \qquad c_{m,n} = \sum_{v=1}^{N} a_{m,v} b_{v,n}$$

n^{th} column of B

m^{th} row of A

$$\overset{\mathbf{C}}{\begin{bmatrix} c_{1,1} & c_{1,2} & \cdots & c_{1,n} & \cdots & \cdots & c_{1,Q} \\ c_{2,1} & c_{2,2} & \cdots & c_{2,n} & \cdots & \cdots & c_{2,Q} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \vdots & \ddots & \ddots & \vdots \\ c_{m,1} & c_{m,2} & \cdots & c_{m,n} & \cdots & \cdots & c_{m,Q} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \ddots & \vdots \\ c_{M,1} & c_{M,2} & \cdots & c_{M,n} & \cdots & \cdots & c_{M,Q} \end{bmatrix}} = \overset{\mathbf{A}}{\begin{bmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} & \cdots & \cdots & a_{1,N} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} & \cdots & \cdots & a_{2,N} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \vdots & \ddots & \ddots & \vdots \\ a_{m,1} & a_{m,2} & \cdots & a_{m,n} & \cdots & \cdots & a_{m,N} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \ddots & \vdots \\ a_{M,1} & a_{M,2} & \cdots & a_{M,n} & \cdots & \cdots & a_{M,N} \end{bmatrix}} \overset{\mathbf{B}}{\begin{bmatrix} b_{1,1} & b_{1,2} & \cdots & b_{1,n} & \cdots & \cdots & b_{1,Q} \\ b_{2,1} & b_{2,2} & \cdots & b_{2,n} & \cdots & \cdots & b_{2,Q} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \vdots & \ddots & \ddots & \vdots \\ b_{m,1} & b_{m,2} & \cdots & b_{m,n} & \cdots & \cdots & b_{m,Q} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \ddots & \vdots \\ b_{N,1} & b_{N,2} & \cdots & b_{N,n} & \cdots & \cdots & b_{N,Q} \end{bmatrix}}$$

output

inner product

Multiplication (with the Sub-Matrices)

If $$\mathbf{A} = \begin{bmatrix} \mathbf{A}_{1,1} & \mathbf{A}_{1,2} & \cdots & \mathbf{A}_{1,V} \\ \mathbf{A}_{2,1} & \mathbf{A}_{2,2} & \cdots & \mathbf{A}_{2,V} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{A}_{U,1} & \mathbf{A}_{U,2} & \cdots & \mathbf{A}_{U,V} \end{bmatrix} \qquad \mathbf{B} = \begin{bmatrix} \mathbf{B}_{1,1} & \mathbf{B}_{1,2} & \cdots & \mathbf{B}_{1,X} \\ \mathbf{B}_{2,1} & \mathbf{B}_{2,2} & \cdots & \mathbf{B}_{2,X} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{B}_{W,1} & \mathbf{B}_{W,2} & \cdots & \mathbf{B}_{W,X} \end{bmatrix}$$

and the columns of $\mathbf{A}_{m,v}$ is equal to the rows of $\mathbf{B}_{v,x}$

then $$\mathbf{AB} = \begin{bmatrix} \mathbf{C}_{1,1} & \mathbf{C}_{1,2} & \cdots & \mathbf{C}_{1,X} \\ \mathbf{C}_{2,1} & \mathbf{C}_{2,2} & \cdots & \mathbf{C}_{2,X} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{C}_{U,1} & \mathbf{C}_{U,2} & \cdots & \mathbf{C}_{U,X} \end{bmatrix}$$ where $\mathbf{C}_{u,x} = \sum_{v=1}^{V} \mathbf{A}_{u,v}\mathbf{B}_{v,x}$

[Example 1]

A_{11} A_{12} B_{11} B_{12}

$$\left[\begin{array}{cc|c} 1 & 2 & 0 \\ 3 & 1 & 0 \\ \hline 1 & 0 & 1 \end{array}\right]\left[\begin{array}{cc|c} 0 & 1 & 0 \\ 1 & 0 & 2 \\ \hline 0 & 1 & 1 \end{array}\right]=\begin{bmatrix} \mathbf{C}_{1,1} & \mathbf{C}_{1,2} \\ \mathbf{C}_{2,1} & \mathbf{C}_{2,2} \end{bmatrix}$$

A_{21} A_{22} B_{21} B_{22}

where
$$\mathbf{C}_{1,1}=\underbrace{\begin{bmatrix} 1 & 2 \\ 3 & 1 \end{bmatrix}}_{A_{11}}\underbrace{\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}}_{B_{11}}+\underbrace{\begin{bmatrix} 0 \\ 0 \end{bmatrix}}_{A_{12}}\underbrace{\begin{bmatrix} 0 & 1 \end{bmatrix}}_{B_{21}}=\begin{bmatrix} 2 & 1 \\ 1 & 3 \end{bmatrix}$$

$$\mathbf{C}_{1,2}=\begin{bmatrix} 1 & 2 \\ 3 & 1 \end{bmatrix}\begin{bmatrix} 0 \\ 2 \end{bmatrix}+\begin{bmatrix} 0 \\ 0 \end{bmatrix}1=\begin{bmatrix} 4 \\ 2 \end{bmatrix}$$

$$\mathbf{C}_{2,1}=\begin{bmatrix} 1 & 0 \end{bmatrix}\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}+1\begin{bmatrix} 0 & 1 \end{bmatrix}=\begin{bmatrix} 0 & 2 \end{bmatrix} \qquad \mathbf{C}_{2,2}=\begin{bmatrix} 1 & 0 \end{bmatrix}\begin{bmatrix} 0 \\ 2 \end{bmatrix}+1\cdot 1=1$$

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

(3) Scalar Multiplication

$$k\mathbf{A} = k\begin{bmatrix} a_{1,1} & a_{1,2} & \cdots & \cdots & a_{1,N} \\ a_{2,1} & a_{2,2} & \cdots & \cdots & a_{2,N} \\ \vdots & \vdots & \ddots & & \vdots \\ \vdots & \vdots & & \ddots & \vdots \\ a_{M,1} & a_{M,2} & \cdots & \cdots & a_{M,N} \end{bmatrix} = \begin{bmatrix} ka_{1,1} & ka_{1,2} & \cdots & \cdots & ka_{1,N} \\ ka_{2,1} & ka_{2,2} & \cdots & \cdots & ka_{2,N} \\ \vdots & \vdots & \ddots & & \vdots \\ \vdots & \vdots & & \ddots & \vdots \\ ka_{M,1} & ka_{M,2} & \cdots & \cdots & ka_{M,N} \end{bmatrix}$$

(4) Transpose

$$\mathbf{B} = \mathbf{A}^{\mathbf{T}} \quad \mathit{if}\ b_{m,n} = a_{n,m}$$

(5) Hermitian

$$\mathbf{B} = \mathbf{A}^{\mathbf{H}} \quad \mathit{if}\ b_{m,n} = a^{*}_{n,m}$$

* means conjugation

In fact,

$$\mathbf{B} = \left(\mathbf{A}^{T}\right)^{*} = \left(\mathbf{A}^{*}\right)^{T}$$

* Definition related to transpose and Hermitian

If $\mathbf{A} = \mathbf{A}^{\mathbf{T}}$ $\Longrightarrow$ Symmetric Matrix

If $\mathbf{A} = -\mathbf{A}^{\mathbf{T}}$ $\Longrightarrow$ Skew Symmetric Matrix

If $\mathbf{A} = \mathbf{A}^{\mathbf{H}}$ $\Longrightarrow$ Symmetric Hermitian Matrix

If $\mathbf{A} = -\mathbf{A}^{\mathbf{H}}$ $\Longrightarrow$ Skew Symmetric Hermitian Matrix

(6) Permutation: exchanging rows or exchanging columns

For a permutation matrix, only one entry in each row or each column is 1, others are 0.

Examples: Exchange the 2nd and the 4th rows:

row permutation

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 10 & 11 & 12 \\ 13 & 14 & 15 & 16 \end{bmatrix} = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 13 & 14 & 15 & 16 \\ 9 & 10 & 11 & 12 \\ 5 & 6 & 7 & 8 \end{bmatrix}$$

permuting matrix

Exchange the 2nd and the 3rd columns:

column permutation

$$\begin{bmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 10 & 11 & 12 \\ 13 & 14 & 15 & 16 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 3 & 2 & 4 \\ 5 & 7 & 6 & 8 \\ 9 & 11 & 10 & 12 \\ 13 & 15 & 14 & 16 \end{bmatrix}$$

permuting matrix

(7) Determinant

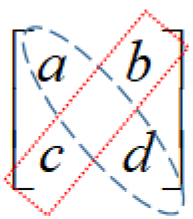
denoted by det(**A**) or |**A**|

$$\det\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right) = \begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc$$

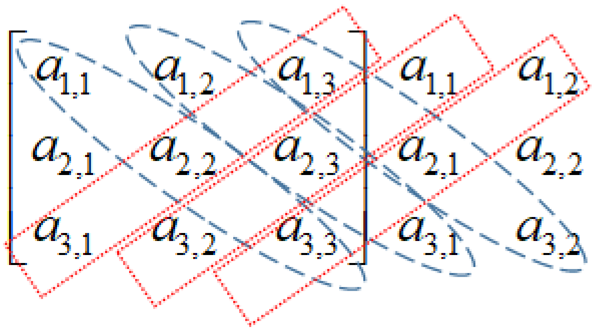
$$\det\left(\begin{bmatrix} a_{1,1} & a_{1,2} & a_{1,3} \\ a_{2,1} & a_{2,2} & a_{2,3} \\ a_{3,1} & a_{3,2} & a_{3,3} \end{bmatrix}\right) = \begin{vmatrix} a_{1,1} & a_{1,2} & a_{1,3} \\ a_{2,1} & a_{2,2} & a_{2,3} \\ a_{3,1} & a_{3,2} & a_{3,3} \end{vmatrix} =$$

$$= a_{1,1}a_{2,2}a_{3,3} + a_{1,2}a_{2,3}a_{3,1} + a_{1,3}a_{2,1}a_{3,2} - a_{1,1}a_{2,3}a_{3,2} - a_{1,2}a_{2,1}a_{3,3} - a_{1,3}a_{2,2}a_{3,1}$$

(a)



(b)



In general, if

$$\mathbf{A}=\begin{bmatrix} a_{1,1} & a_{1,2} & \cdots & \cdots & a_{1,N} \\ a_{2,1} & a_{2,2} & \cdots & \cdots & a_{2,N} \\ \vdots & \vdots & \ddots & & \vdots \\ \vdots & \vdots & & \ddots & \vdots \\ a_{N,1} & a_{N,2} & \cdots & \cdots & a_{N,N} \end{bmatrix}$$

$$\det(\mathbf{A}) = a_{i,1}C_{i,1} + a_{i,2}C_{i,2} + \cdots + a_{i,N}C_{i,N}$$ → choosing any i

$$= a_{k,1}C_{k,1} + a_{k,2}C_{k,2} + \cdots + a_{k,N}C_{k,N}$$ → choosing any k

We call $C_{i,j}$ the cofactor of $a_{i,j}$ and call $M_{i,j}$ a minor determinant.

where $C_{i,j} = (-1)^{i+j} M_{i,j}$

$$M_{i,j} = \det\begin{bmatrix} a_{1,1} & \cdots & a_{1,j-1} & a_{1,j} & a_{1,j+1} & \cdots & a_{1,N} \\ \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{i-1,1} & \cdots & a_{i-1,j-1} & a_{i-1,j} & a_{i-1,j+1} & \cdots & a_{i-1,N} \\ a_{i,1} & \cdots & a_{i,j-1} & a_{i,j} & a_{i,j+1} & \cdots & a_{i,N} \\ a_{i+1,1} & \cdots & a_{i+1,j-1} & a_{i+1,j} & a_{i+1,j+1} & \cdots & a_{i+1,N} \\ \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{N,1} & \cdots & a_{N,j-1} & a_{N,j} & a_{N,j+1} & \cdots & a_{N,N} \end{bmatrix}$$

The determinant is always hard to determine when N is large.

However, when $\mathbf{A}$ is an upper triangular matrix:

$$\mathbf{A} = \begin{bmatrix} a_{1,1} & a_{1,2} & a_{1,3} & \cdots & a_{1,N-1} & a_{1,N} \\ 0 & a_{2,2} & a_{2,3} & \cdots & a_{2,N-1} & a_{2,N} \\ 0 & 0 & a_{3,3} & \cdots & a_{3,N-1} & a_{3,N} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & a_{N-1,N-1} & a_{N-1,N} \\ 0 & 0 & 0 & \cdots & 0 & a_{N,N} \end{bmatrix} \qquad a_{m,n} = 0 \quad if \ m > n$$

its determinant is the product of diagonal entries:

$$\det(\mathbf{A}) = a_{1,1}a_{2,2}a_{3,3}\cdots a_{N-1,N-1}a_{N,N}$$

When **A** is a lower triangular matrix:

$$\mathbf{A} = \begin{bmatrix} a_{1,1} & 0 & 0 & \cdots & 0 & 0 \\ a_{2,1} & a_{2,2} & 0 & \cdots & 0 & 0 \\ a_{3,1} & a_{3,2} & a_{3,3} & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{N-1,1} & a_{N-1,2} & a_{N-1,3} & \cdots & a_{N-1,N-1} & 0 \\ a_{N,1} & a_{N,2} & a_{N,3} & \cdots & a_{N,N-1} & a_{N,N} \end{bmatrix} a_{m,n} = 0 \quad if\ m < n$$

its determinant is also the product of diagonal entries:

$$\det(\mathbf{A}) = a_{1,1}a_{2,2}a_{3,3}\cdots a_{N-1,N-1}a_{N,N}$$

We can use the fact that

$$\det(\mathbf{AB}) = \det(\mathbf{A})\det(\mathbf{B})$$

and the row elimination method to determine the determinant.

(See page 500)

(8) Rank

The number of linearly independent rows in a matrix.
(It is equivalent to the number of linearly independent columns in a matrix).

row3 = row1 + 2 X row2

Example: $$rank\left(\begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 2 & 1 \end{bmatrix}\right) = 2$$

column3 = column1
+0·column2

In a linearly independent set, any vector cannot be expressed by a linearly combination of other vectors.

That is, the solution of

$$c_1\mathbf{v_1} + c_2\mathbf{v_2} + \cdots + c_N\mathbf{v_N} = \mathbf{0}$$

is

$$c_1 = c_2 = \cdots = c_N = \mathbf{0}$$

Note: For an NxN square matrix $\mathbf{A}$, if

$$rank(\mathbf{A}) < N$$

then

$$\det(\mathbf{A}) = 0$$

(9) Inverse

If (i) $\mathbf{A}$ is a square matrix and (ii) $\det(\mathbf{A}) \neq 0$, then the inverse of $\mathbf{A}$ (denoted by $\mathbf{A}^{-1}$) exists and it is a matrix that satisfies

$$\mathbf{A}^{-1}\mathbf{A} = \mathbf{A}\mathbf{A}^{-1} = \mathbf{I}$$

Application: For a linear system that can be expressed by

$$\mathbf{A}\mathbf{x} = \mathbf{y}$$

its solution is:

$$\mathbf{x} = \mathbf{A}^{-1}\mathbf{y}$$

Determining the Inverse (Method 1: by Adjoint Matrix)

$$\mathbf{A}^{-1} = \frac{1}{\det(\mathbf{A})}\operatorname{adj}\mathbf{A}$$

where

$$\operatorname{adj}\mathbf{A} = \begin{bmatrix} C_{1,1} & C_{2,1} & \cdots & \cdots & C_{N,1} \\ C_{1,2} & C_{2,2} & \cdots & \cdots & C_{N,2} \\ \vdots & \vdots & \ddots & & \vdots \\ \vdots & \vdots & & \ddots & \vdots \\ C_{1,N} & C_{2,N} & \cdots & \cdots & C_{N,N} \end{bmatrix} = \begin{bmatrix} C_{1,1} & C_{1,2} & \cdots & \cdots & C_{1,N} \\ C_{2,1} & C_{2,2} & \cdots & \cdots & C_{2,N} \\ \vdots & \vdots & \ddots & & \vdots \\ \vdots & \vdots & & \ddots & \vdots \\ C_{N,1} & C_{N,2} & \cdots & \cdots & C_{N,N} \end{bmatrix}^T$$

$C_{i,j}$ the cofactor of $a_{i,j}$

Determining the Inverse (Method 2: by Row Elimination)

$$\left[\mathbf{A}\middle|\mathbf{I}\right] \xrightarrow[\text{elementary row operations}]{B} \left[\mathbf{I}\middle|\mathbf{A}^{-1}\right]$$

$B[A|I] = [BA|B \cdot I]$

If $BA = I$, $B = A^{-1}$, $B \cdot I = A^{-1}$

(preferred)

[Example 2]: Determine the inverse of the following matrix

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 1 & 2 \\ 2 & 3 & 1 & 1 \\ -1 & -1 & 2 & 3 \\ 1 & 3 & 3 & 4 \end{bmatrix}$$

(Solution):

$$\left[\mathbf{A}\middle|\mathbf{I}\right] = \left[\begin{array}{cccc|cccc} 1 & 2 & 1 & 2 & 1 & 0 & 0 & 0 \\ 2 & 3 & 1 & 1 & 0 & 1 & 0 & 0 \\ -1 & -1 & 2 & 3 & 0 & 0 & 1 & 0 \\ 1 & 3 & 3 & 4 & 0 & 0 & 0 & 1 \end{array}\right]$$

$$\mathbf{E}_1\left[\mathbf{A}\middle|\mathbf{I}\right]=\left[\begin{array}{cccc|cccc} 1 & 2 & 1 & 2 & 1 & 0 & 0 & 0 \\ 0 & -1 & -1 & -3 & -2 & 1 & 0 & 0 \\ 0 & 1 & 3 & 5 & 1 & 0 & 1 & 0 \\ 0 & 1 & 2 & 2 & -1 & 0 & 0 & 1 \end{array}\right]$$

where $\mathbf{E}_1=\begin{bmatrix} 1 & 0 & 0 & 0 \\ -2 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ -1 & 0 & 0 & 1 \end{bmatrix}$

$\det(E_1)=1$

$$\mathbf{E}_2\mathbf{E}_1\left[\mathbf{A}\middle|\mathbf{I}\right]=\left[\begin{array}{cccc|cccc} 1 & 2 & 1 & 2 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 3 & 2 & -1 & 0 & 0 \\ 0 & 1 & 3 & 5 & 1 & 0 & 1 & 0 \\ 0 & 1 & 2 & 2 & -1 & 0 & 0 & 1 \end{array}\right]$$

where $\mathbf{E}_2=\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$

$\det(E_2)=-1$

$$\mathbf{E}_3\mathbf{E}_2\mathbf{E}_1\left[\mathbf{A}\middle|\mathbf{I}\right]=\left[\begin{array}{cccc|cccc} 1 & 2 & 1 & 2 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 3 & 2 & -1 & 0 & 0 \\ 0 & 0 & 2 & 2 & -1 & 1 & 1 & 0 \\ 0 & 0 & 1 & -1 & -3 & 1 & 0 & 1 \end{array}\right]$$

where $\mathbf{E}_3=\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & -1 & 0 & 1 \end{bmatrix}$

$\det(E_3)=1$

$$\mathbf{E}_4\mathbf{E}_3\mathbf{E}_2\mathbf{E}_1\left[\mathbf{A}\middle|\mathbf{I}\right]=\left[\begin{array}{cccc|cccc} 1 & 2 & 1 & 2 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 3 & 2 & -1 & 0 & 0 \\ 0 & 0 & 1 & 1 & -1/2 & 1/2 & 1/2 & 0 \\ 0 & 0 & 1 & -1 & -3 & 1 & 0 & 1 \end{array}\right]$$

where $\mathbf{E}_4=\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1/2 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$

$\det(E_4)=\frac{1}{2}$

$$\mathbf{E}_5\mathbf{E}_4\mathbf{E}_3\mathbf{E}_2\mathbf{E}_1\left[\mathbf{A}\middle|\mathbf{I}\right]=\left[\begin{array}{cccc|cccc}1&2&1&2&1&0&0&0\\0&1&1&3&2&-1&0&0\\0&0&1&1&-1/2&1/2&1/2&0\\0&0&0&-2&-5/2&1/2&-1/2&1\end{array}\right] \quad \text{where } \mathbf{E}_5=\begin{bmatrix}1&0&0&0\\0&1&0&0\\0&0&1&0\\0&0&-1&1\end{bmatrix}$$

$\det(E_5)=1$

$$\mathbf{E}_6\mathbf{E}_5\mathbf{E}_4\mathbf{E}_3\mathbf{E}_2\mathbf{E}_1\left[\mathbf{A}\middle|\mathbf{I}\right]=\left[\begin{array}{cccc|cccc}1&2&1&2&1&0&0&0\\0&1&1&3&2&-1&0&0\\0&0&1&1&-1/2&1/2&1/2&0\\0&0&0&1&5/4&-1/4&1/4&-1/2\end{array}\right] \quad \mathbf{E}_6=\begin{bmatrix}1&0&0&0\\0&1&0&0\\0&0&1&0\\0&0&0&-1/2\end{bmatrix}$$

$\det(E_6)=\frac{-1}{2}$

$$\mathbf{E}_7\mathbf{E}_6\mathbf{E}_5\mathbf{E}_4\mathbf{E}_3\mathbf{E}_2\mathbf{E}_1\left[\mathbf{A}\middle|\mathbf{I}\right]=\left[\begin{array}{cccc|cccc}1&2&1&0&-3/2&1/2&-1/2&1\\0&1&1&0&-7/4&-1/4&-3/4&3/2\\0&0&1&0&-7/4&3/4&1/4&1/2\\0&0&0&1&5/4&-1/4&1/4&-1/2\end{array}\right] \quad \mathbf{E}_7=\begin{bmatrix}1&0&0&-2\\0&1&0&-3\\0&0&1&-1\\0&0&0&1\end{bmatrix}$$

$$E_6E_5E_4E_3E_2E_1A=\begin{bmatrix}1&2&1&2\\0&1&1&3\\0&0&1&1\\0&0&0&1\end{bmatrix} \quad -\frac{1}{2}\cdot 1\cdot\frac{1}{2}\cdot 1\cdot -1\cdot 1\cdot\det(A)=1$$

$\det(A)=4$

$$\underbrace{\mathbf{E}_9\mathbf{E}_8\mathbf{E}_7\mathbf{E}_6\mathbf{E}_5\mathbf{E}_4\mathbf{E}_3\mathbf{E}_2\mathbf{E}_1}_{A^{-1}}\left[\mathbf{A}\middle|\mathbf{I}\right] = \left[\begin{array}{cccc|cccc} 1 & 0 & 0 & 0 & 1/4 & 7/4 & 5/4 & -3/2 \\ 0 & 1 & 0 & 0 & 0 & -1 & -1 & 1 \\ 0 & 0 & 1 & 0 & -7/4 & 3/4 & 1/4 & 1/2 \\ 0 & 0 & 0 & 1 & 5/4 & -1/4 & 1/4 & -1/2 \end{array}\right]$$

where

$$\mathbf{E}_8 = \begin{bmatrix} 1 & 0 & -1 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \qquad \mathbf{E}_9 = \begin{bmatrix} 1 & -2 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Therefore,

$$\mathbf{A}^{-1} = \begin{bmatrix} 1/4 & 7/4 & 5/4 & -3/2 \\ 0 & -1 & -1 & 1 \\ -7/4 & 3/4 & 1/4 & 1/2 \\ 5/4 & -1/4 & 1/4 & -1/2 \end{bmatrix}$$

Specially, for a 2x2 matrix:

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

if det(**A**) = $ad - bc \neq 0$, then $\mathbf{A}^{-1}$ can be determined from

$$\mathbf{A}^{-1} = \frac{1}{ad - bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

(Note): The row elimination method is also helpful for calculating the determinant.

For example, in Example 2, since

$$\mathbf{E}_5\mathbf{E}_4\mathbf{E}_3\mathbf{E}_2\mathbf{E}_1\mathbf{A} = \begin{bmatrix} 1 & 2 & 1 & 2 \\ 0 & 1 & 1 & 2 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & -2 \end{bmatrix}$$

$$\mathbf{E}_1 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ -2 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ -1 & 0 & 0 & 1 \end{bmatrix} \quad \mathbf{E}_2 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad \mathbf{E}_3 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & -1 & 0 & 1 \end{bmatrix}$$

$$\mathbf{E}_4 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1/2 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad \mathbf{E}_5 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & -1 & 1 \end{bmatrix}$$

$$\det(\mathbf{E}_5)\det(\mathbf{E}_4)\det(\mathbf{E}_3)\det(\mathbf{E}_2)\det(\mathbf{E}_1)\det(\mathbf{A}) = \det\left(\begin{bmatrix} 1 & 2 & 1 & 2 \\ 0 & 1 & 1 & 2 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & -2 \end{bmatrix}\right)$$

$$1 \cdot \frac{1}{2} \cdot 1 \cdot (-1) \cdot 1 \cdot \det(\mathbf{A}) = -2$$

$$\det(\mathbf{A}) = 4$$

6.1.3 Eigenvalues and Eigenvectors

[Definitions]

For a square matrix **A**, if

$$\mathbf{A}\mathbf{e} = \lambda\mathbf{e}$$

$Ae = \lambda I e$ $\quad (A-\lambda I)e = 0$

if $\det(A-\lambda I) \neq 0$ $\quad (A-\lambda I)^{-1}$ exists

the only solution of e is

$e = (A-\lambda I)0 = 0$

then λ is called the eigenvalue

e is called the eigenvector corresponding to λ

$\therefore \det(A-\lambda I) = 0$ must be satisfied.

[Determining Eigenvalues]

Eigenvalues are the solutions of

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

characteristic equation

[Determining Eigenvectors]

Solve **e** from

$$\left(\mathbf{A}-\lambda\mathbf{I}\right)\mathbf{e}=\mathbf{0}$$

Note: If **e** is an eigenvector of **A**, then $c\mathbf{e}$ is also an eigenvector of **A** where c is a scalar.

$$\left(\mathbf{A}-\lambda\mathbf{I}\right)c\mathbf{e}=c\mathbf{0}=\mathbf{0}$$

[Eigenspace]

Suppose that $\mathbf{e_1}$, $\mathbf{e_2}$, …, $\mathbf{e_k}$ are linearly independent eigenvectors corresponding to λ. Then the eigenspace corresponding to λ is

$$span\left\{\mathbf{e_1},\mathbf{e_2},\cdots,\mathbf{e_k}\right\}$$

$$=c_1\mathbf{e_1}+c_2\mathbf{e_2}+\cdots+c_k\mathbf{e_k}$$

c_1, c_2, …, c_k can be any constants

$A(c_1e_1+c_2e_2)$

$=c_1Ae_1+c_2Ae_2=c_1\lambda e_1+c_2\lambda e_2$

$=\lambda(c_1e_1+c_2e_2)$

ex: $A=\begin{bmatrix}2&0\\0&2\end{bmatrix}$

$\lambda=2,2$

$e_1=\begin{bmatrix}1\\0\end{bmatrix}$ $e_2=\begin{bmatrix}0\\1\end{bmatrix}$

$Ae_1=2e_1$

$Ae_2=2e_2$

$\det(A) = \det(E)\det(D)\det(E^{-1}) = \det(D)$

$\det(E^{-1}) = 1/\det(E)$

[Eigenvalue-Eigenvector Decomposition]

$$\mathbf{A} = \mathbf{E}\mathbf{D}\mathbf{E}^{-1}$$

$A^2 = ED\cancel{E^{-1}}\cancel{E}DE^{-1} = ED^2E^{-1}$

$A^n = ED^nE^{-1}$

where

$D^n = \begin{bmatrix} \lambda_1^n & & 0 \\ & \lambda_2^n & \\ 0 & & \ddots & \lambda_N^n \end{bmatrix}$

$$\mathbf{E} = \begin{bmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \cdots & \mathbf{e}_{N\text{-}1} & \mathbf{e}_N \end{bmatrix}$$

$$\mathbf{D} = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 & 0 \\ 0 & \lambda_2 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & \lambda_{N-1} & 0 \\ 0 & 0 & \cdots & 0 & \lambda_N \end{bmatrix}$$

$$\mathbf{A}\mathbf{e}_n = \lambda_n \mathbf{e}_n \qquad n = 1, 2, \ldots, N$$

$\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_{N\text{-}1}, \mathbf{e}_N$ are linearly independent

$\det(E) \neq 0$, E^{-1} exists

Since

$$\mathbf{A}\begin{bmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \cdots & \mathbf{e}_{N\text{-}1} & \mathbf{e}_N \end{bmatrix} = \begin{bmatrix} \lambda_1\mathbf{e}_1 & \lambda_2\mathbf{e}_2 & \cdots & \lambda_{N-1}\mathbf{e}_{N\text{-}1} & \lambda_N\mathbf{e}_N \end{bmatrix}$$

$$\mathbf{A}\mathbf{E} = \mathbf{E}\mathbf{D} \qquad \mathbf{A} = \mathbf{E}\mathbf{D}\mathbf{E}^{-1}$$

$\text{If } A = A^H,\ A e_1 = \lambda_1 e_1,\ A e_2 = \lambda_2 e_2,\ \lambda_1 \neq \lambda_2$

Specially, if
$e_2^H A e_1 = \lambda_1 e_2^H e_1$
$e_2^H A^H e_1 = (A e_2)^H e_1 = \lambda_2 e_2^H e_1$
$\lambda_1 e_2^H e_1 = \lambda_2 e_2^H e_1$
$e_2^H e_1 = 0$

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

the eigenvectors form a complete and orthogonal set (after normalization, it becomes a complete and orthonormal set), then

$$\mathbf{E}^{-1} = \mathbf{E}^{\mathbf{H}}$$

$E^H E = I,\ E^{-1} = E^H$

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

$\begin{bmatrix} e_1^H \\ e_2^H \\ \vdots \\ e_N^H \end{bmatrix} [e_1\ e_2 \cdots e_N] = I$

since $e_m^H e_n = 0$ if $m \neq n$

and **A** can be decomposed into

$$\mathbf{A} = \lambda_1 \mathbf{e}_1 \mathbf{e}_1^{\mathbf{H}} + \lambda_2 \mathbf{e}_2 \mathbf{e}_2^{\mathbf{H}} + \cdots + \lambda_{N-1} \mathbf{e}_{\mathbf{N}-1} \mathbf{e}_{\mathbf{N}-1}^{\mathbf{H}} + \lambda_N \mathbf{e}_{\mathbf{N}} \mathbf{e}_{\mathbf{N}}^{\mathbf{H}}$$

$e_n^H e_n = 1$
(after normalization)

$[e_1, e_2, \cdots e_N] \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_N \end{bmatrix} \begin{bmatrix} e_1^H \\ e_2^H \\ \vdots \\ e_N^H \end{bmatrix}$

$= [\lambda_1 e_1, \lambda_2 e_2, \cdots \lambda_N e_N] \begin{bmatrix} e_1^H \\ e_2^H \\ \vdots \\ e_N^H \end{bmatrix}$

[Example 3]:

$$\mathbf{A} = \begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix}$$

$A = A^H$

Eigenvalues of $\mathbf{A}$:

$$\det(\mathbf{A} - \lambda\mathbf{I}) = \det\left(\begin{bmatrix} 3-\lambda & 1 \\ 1 & 3-\lambda \end{bmatrix}\right) = (3-\lambda)^2 - 1$$

$$= \lambda^2 - 6\lambda + 8 = 0$$

$\lambda = 2, 4$

Eigenvectors of $\mathbf{A}$:

For $\lambda = 2$,

$$(\mathbf{A} - 2\mathbf{I})\mathbf{e}_1 = 0 \qquad \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \overbrace{\begin{bmatrix} e_{1,1} \\ e_{1,2} \end{bmatrix}}^{e_1} = 0 \qquad \text{where} \quad \mathbf{e}_1 = \begin{bmatrix} e_{1,1} \\ e_{1,2} \end{bmatrix}$$

$$e_{1,1} + e_{1,2} = 0 \qquad e_{1,1} = 1, \quad e_{1,2} = -1 \qquad \mathbf{e}_1 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

Eigenvectors of **A**:

For $\lambda = 4$,

$$(\mathbf{A} - 4\mathbf{I})\mathbf{e}_1 = 0 \quad \begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix}\begin{bmatrix} e_{2,1} \\ e_{2,2} \end{bmatrix} = 0 \quad \text{where} \quad \mathbf{e}_2 = \begin{bmatrix} e_{2,1} \\ e_{2,2} \end{bmatrix}$$

$$-e_{2,1} + e_{2,2} = 0 \qquad e_{2,1} = 1, \quad e_{2,2} = 1 \qquad \mathbf{e}_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

$\|e_1\| = \sqrt{1^2 + (-1)^2} = \sqrt{2}$

Eigenvector-eigenvalue decomposition of **A**:

$$\mathbf{A} = \mathbf{E}\mathbf{D}\mathbf{E}^{-1} \qquad \mathbf{E} = [\mathbf{e}_1 \quad \mathbf{e}_2] = \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} \qquad \mathbf{D} = \begin{bmatrix} 2 & 0 \\ 0 & 4 \end{bmatrix}$$

$$\mathbf{A} = \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}\begin{bmatrix} 2 & 0 \\ 0 & 4 \end{bmatrix}\begin{bmatrix} 1/2 & -1/2 \\ 1/2 & 1/2 \end{bmatrix}$$

e_1 is orthogonal to e_2

With normalization

$$\mathbf{e_1} = \begin{bmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \end{bmatrix} \qquad \mathbf{e_2} = \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}$$

$$\mathbf{A} = \mathbf{EDE^H} \qquad \mathbf{E} = \begin{bmatrix} \mathbf{e_1} & \mathbf{e_2} \end{bmatrix} = \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ -1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix} \qquad \mathbf{D} = \begin{bmatrix} 2 & 0 \\ 0 & 4 \end{bmatrix}$$

$$\mathbf{A} = \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ -1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix} \begin{bmatrix} 2 & 0 \\ 0 & 4 \end{bmatrix} \begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix}$$

6.1.4 Properties

(A) Geometric Operation Formula

(1) Rotation	Clockwise $\begin{bmatrix} x_1 \\ y_1 \end{bmatrix} = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$ Clockwise rotation with respect to (x_0, y_0) $\begin{bmatrix} x_1 \\ y_1 \end{bmatrix} = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} 1-\cos\theta & -\sin\theta \\ \sin\theta & 1-\cos\theta \end{bmatrix} \begin{bmatrix} x_0 \\ y_0 \end{bmatrix}$
(2) Scaling	$\begin{bmatrix} x_1 - x_0 \\ y_1 - x_0 \end{bmatrix} = \begin{bmatrix} \sigma_x & 0 \\ 0 & \sigma_y \end{bmatrix} \begin{bmatrix} x - x_0 \\ y - x_0 \end{bmatrix}$
(3) Shearing	$\begin{bmatrix} x_1 - x_0 \\ y_1 - x_0 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ c & 1 \end{bmatrix} \begin{bmatrix} x - x_0 \\ y - x_0 \end{bmatrix}$ or $\begin{bmatrix} x_1 - x_0 \\ y_1 - x_0 \end{bmatrix} = \begin{bmatrix} 1 & b \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x - x_0 \\ y - x_0 \end{bmatrix}$

(4) Reflection	with respect to (x_0, y_0) $$\begin{bmatrix} x_1 - x_0 \\ y_1 - y_0 \end{bmatrix} = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} x - x_0 \\ y - y_0 \end{bmatrix}$$ with respect to (x_0, y_0), only reflect on x-axis $$\begin{bmatrix} x_1 - x_0 \\ y_1 - y_0 \end{bmatrix} = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x - x_0 \\ y - y_0 \end{bmatrix}$$
(5) Affine Transformation	$$\begin{bmatrix} x_1 - x_0 \\ y_1 - x_0 \end{bmatrix} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} x - x_0 \\ y - x_0 \end{bmatrix}$$
(6) Projection	on the x-axis $$\begin{bmatrix} x_1 \\ y_1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$ on the axis of $c(\cos\theta, \sin\theta)$: $$\begin{bmatrix} x_1 \\ y_1 \end{bmatrix} = \begin{bmatrix} \cos^2\theta & \cos\theta\sin\theta \\ \cos\theta\sin\theta & \sin^2\theta \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

(B) Properties of Determinants

(1) Transpose	$\det(\mathbf{A}) = \det(\mathbf{A}^{\mathrm{T}})$
(2) Multiplication	$\det(\mathbf{AB}) = \det(\mathbf{BA}) = \det(\mathbf{A})\det(\mathbf{B})$
(3) Zero row / column	$\det(\mathbf{A}) = 0$ if all the entries of a row (or a column) are 0.
(4) Row exchange	If **B** is the same as **A** but two of the rows are exchanged $$\det(\mathbf{B}) = -\det(\mathbf{A})$$
(5) Scaling	If $b_{i,j} = a_{i,j}$ when $i \neq k$, and $b_{k,j} = \tau a_{k,j}$, then $$\det(\mathbf{B}) = \tau \det(\mathbf{A})$$
(6) Row addition	If $b_{i,j} = a_{i,j}$ when $i \neq k$, and $b_{k,j} = a_{k,j} + \tau a_{h,j}$, then $$\det(\mathbf{B}) = \det(\mathbf{A})$$

(C) Properties of Ranks

(i)	rank($\mathbf{0}$) = 1
(ii)	rank($\mathbf{I}$) = N if $\mathbf{I}$ is an $N \times N$ identify matrix
(iii)	rank($\mathbf{P}$) = N if $\mathbf{P}$ is an $N \times N$ permutation matrix
(iv)	rank($\mathbf{D}$) = N_1 if $\mathbf{D}$ is a diagonal matrix and there are N_1 nonzero entries in the diagonal line
(v)	rank($\mathbf{A}$) = N if $\mathbf{A}$ is an $N \times N$ triangular matrix and all of the entries in the diagonal line are nonzero
(vi)	rank($\mathbf{A}$) = rank($\mathbf{A}^{\mathrm{T}}$)
(vii)	rank($\mathbf{A}$) $\leq$ min(M, N) if $\mathbf{A}$ is an $M \times N$ matrix
(viii)	rank($\mathbf{DA}$) = rank($\mathbf{A}$) if $\mathbf{A}$ is an $M \times N$ and $\mathbf{D}$ is an $N \times N$ matrix whose entries in the diagonal line are all nonzero
(ix)	rank($\mathbf{BA}$) $\leq$ min(rank($\mathbf{A}$), rank($\mathbf{B}$))
(x)	rank($\mathbf{BA}$) = rank($\mathbf{A}$) if $\mathbf{A}$ is an $M \times N$ matrix and rank($\mathbf{B}$) = M

D. G. Zill, W. S. Wright, and J. J. Ding, Engineering Mathematics, Metric Edition, Cengage Learning, Taipei, Taiwan, 2019.

(D) Properties of Matrix Inverse

(i)	$(\mathbf{A}^{-1})^{-1} = \mathbf{A}$
(ii)	$(\mathbf{AB})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1}$
(iii)	$(\mathbf{A}_1\mathbf{A}_2 \cdots \mathbf{A}_h)^{-1} = \mathbf{A}_h{}^{-1} \cdots \mathbf{A}_2{}^{-1}\,\mathbf{A}_1{}^{-1}$
(iv)	$(\mathbf{A}^{\mathrm{T}})^{-1} = (\mathbf{A}^{-1})^{\mathrm{T}}, \quad (\mathbf{A}^{\mathrm{H}})^{-1} = (\mathbf{A}^{-1})^{\mathrm{H}}$
(v)	$\mathbf{I}^{-1} = \mathbf{I}$
(vi)	If $\mathbf{D}$ is a diagonal matrix, $\mathbf{D}^{-1} = \mathbf{F}$ where $f_{m,n} = 0$ if $m \neq n$, $f_{n,n} = 1/d_{n,n}$
(vii)	$\mathbf{P}^{-1} = \mathbf{P}^{\mathrm{T}}$ if $\mathbf{P}$ is a permutation matrix
(viii)	$\begin{bmatrix} \mathbf{A}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{A}_2 \end{bmatrix}^{-1} = \begin{bmatrix} \mathbf{A}_1^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{A}_2^{-1} \end{bmatrix}$
(ix)	$\det(\mathbf{A}^{-1}) = 1/\det(\mathbf{A})$

D. G. Zill, W. S. Wright, and J. J. Ding, Engineering Mathematics, Metric Edition, Cengage Learning, Taipei, Taiwan, 2019.

(E) Properties of Eigenvectors and Eigenvalues

(1) sum of eigenvalues	$\sum_{m=1}^{N} \lambda_m = \sum_{n=1}^{N} A[n,n]$
(2) product of eigenvalues	$\prod_{m=1}^{N} \lambda_m = \det(\mathbf{A})$
(3) eigenvectors / eigenvalues for $\mathbf{A} = \mathbf{A}^{\mathbf{H}}$	If $\mathbf{A} = \mathbf{A}^{\mathbf{H}}$, then (i) the eigenvalues are real, (ii) if $\mathbf{Ae_1} = \lambda_1 \mathbf{e_1}, \mathbf{Ae_2} = \lambda_2 \mathbf{e_2}$, $\lambda_1 \neq \lambda_2$, then $\langle \mathbf{e_1}, \mathbf{e_2} \rangle = 0$ Eigenvectors with different eigenvalues are orthogonal.
(4) eigenvectors / eigenvalues for $\mathbf{A}^{-1}$	If $\mathbf{Ae} = \lambda \mathbf{e}$, then $\mathbf{A}^{-1}\mathbf{e} = \lambda^{-1}\mathbf{e}$
(5) similar matrix	$\mathbf{S}(\mathbf{Be}) = \lambda \mathbf{Be}$ if $\mathbf{Ae} = \lambda \mathbf{e}$ and $\mathbf{S} = \mathbf{BAB}^{-1}$

(6) quadratic form	If $f(x,y)=ax^2+bxy+cy^2=\begin{bmatrix}x & y\end{bmatrix}\mathbf{M}\begin{bmatrix}x\\ y\end{bmatrix}$, where $\mathbf{M}=\begin{bmatrix}a & b/2\\ b/2 & c\end{bmatrix}$, then $f(x_1,y_1)=\lambda_1 x_1^2+\lambda_2 y_1^2$ where $\begin{bmatrix}x_1 & y_1\end{bmatrix}=\begin{bmatrix}x & y\end{bmatrix}\begin{bmatrix}\mathbf{e}_1 & \mathbf{e}_2\end{bmatrix}$ $\mathbf{M}=\begin{bmatrix}\mathbf{e}_1 & \mathbf{e}_2\end{bmatrix}\begin{bmatrix}\lambda_1 & 0\\ 0 & \lambda_2\end{bmatrix}\begin{bmatrix}\mathbf{e}_1^{\mathbf{T}}\\ \mathbf{e}_2^{\mathbf{T}}\end{bmatrix}$
(7) shape for the quadratic form	(i) ellipse: $\lambda_1>0$ and $\lambda_2>0$ or $\lambda_1<0$ and $\lambda_2<0$ (ii) hyperbola: $\lambda_1>0$ and $\lambda_2<0$ or $\lambda_1<0$ and $\lambda_2>0$ (iii) parabola: $\lambda_1=0$ or $\lambda_2=0$ (but are not all zero) (iv) line: $\lambda_1=\lambda_2=0$

6.2 Kronecker and Element-Wise Products

6.2.1 Kronecker Product

Suppose that

$$\mathbf{A} = \begin{bmatrix} a_{1,1} & a_{1,2} & \cdots & \cdots & a_{1,N} \\ a_{2,1} & a_{2,2} & \cdots & \cdots & a_{2,N} \\ \vdots & \vdots & \ddots & & \vdots \\ \vdots & \vdots & & \ddots & \vdots \\ a_{M,1} & a_{M,2} & \cdots & \cdots & a_{M,N} \end{bmatrix}$$

then

global information

$$\mathbf{A} \otimes \mathbf{B} = \begin{bmatrix} a_{1,1}\mathbf{B} & a_{1,2}\mathbf{B} & \cdots & \cdots & a_{1,N}\mathbf{B} \\ a_{2,1}\mathbf{B} & a_{2,2}\mathbf{B} & \cdots & \cdots & a_{2,N}\mathbf{B} \\ \vdots & \vdots & \ddots & & \vdots \\ \vdots & \vdots & & \ddots & \vdots \\ a_{M,1}\mathbf{B} & a_{M,2}\mathbf{B} & \cdots & \cdots & a_{M,N}\mathbf{B} \end{bmatrix}$$

local information

$\text{rank}([a_{11}B \ a_{12}B \cdots a_{1N}B])$
$= \text{rank}(B)$

If $[a_{21}, a_{22}, \cdots a_{2N}]$ is independent of $[a_{11} \ a_{12}, \cdots a_{1N}]$

$\text{rank}\left(\begin{bmatrix} a_{11}B & a_{12}B & \cdots & a_{1N}B \\ a_{21}B & a_{22}B & \cdots & a_{2N}B \end{bmatrix}\right)$
$= 2\,\text{rank}(B)$

Kronecker Product

$$\mathbf{A}\otimes\mathbf{B}=\begin{bmatrix} a_{1,1}\mathbf{B} & a_{1,2}\mathbf{B} & \cdots & \cdots & a_{1,N}\mathbf{B} \\ a_{2,1}\mathbf{B} & a_{2,2}\mathbf{B} & \cdots & \cdots & a_{2,N}\mathbf{B} \\ \vdots & \vdots & \ddots & & \vdots \\ \vdots & \vdots & & \ddots & \vdots \\ a_{M,1}\mathbf{B} & a_{M,2}\mathbf{B} & \cdots & \cdots & a_{M,N}\mathbf{B} \end{bmatrix}$$

Notation of the Kronecker product: $\otimes$

A: Global information
B: Local information

If the size of **A** is $M_1 \times N_1$
the size of **B** is $M_2 \times N_2$
then the size of $\mathbf{A}\otimes\mathbf{B}$ is $(M_1 M_2)\times(N_1 N_2)$

[Example 1]

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

$$\mathbf{A} \otimes \mathbf{B} = \begin{array}{c} \qquad 1 \cdot B \qquad\qquad (-1)B \\ \left[\begin{array}{ccc|ccc} 1 & 1 & 0 & -1 & -1 & 0 \\ 0 & 2 & 0 & 0 & -2 & 0 \\ 0 & 1 & 1 & 0 & -1 & -1 \\ \hline -2 & -2 & 0 & 2 & 2 & 0 \\ 0 & -4 & 0 & 0 & 4 & 0 \\ 0 & -2 & -2 & 0 & 2 & 2 \end{array}\right] \\ \qquad (-2)B \qquad\qquad 2B \end{array}$$

[Addition Property]

If size($\mathbf{A}$) = size($\mathbf{C}$)

$$\mathbf{A} \otimes \mathbf{B} + \mathbf{C} \otimes \mathbf{B} = \left(\mathbf{A} + \mathbf{C}\right) \otimes \mathbf{B}$$

If size($\mathbf{B}$) = size($\mathbf{D}$)

$$\mathbf{A} \otimes \mathbf{B} + \mathbf{A} \otimes \mathbf{D} = \mathbf{A} \otimes \left(\mathbf{B} + \mathbf{D}\right)$$

[Multiplication Property]

If size($\mathbf{A}$) = $M_1 \times N_1$, size($\mathbf{B}$) = $M_2 \times N_2$,
size($\mathbf{C}$) = $M_3 \times N_3$, size($\mathbf{D}$) = $M_4 \times N_4$, and

$$N_1 = M_3, \quad N_2 = M_4$$

then

$$\left(\mathbf{A} \otimes \mathbf{B}\right)\left(\mathbf{C} \otimes \mathbf{D}\right) = \left(\mathbf{AC}\right) \otimes \left(\mathbf{BD}\right)$$

(Proof of the Multiplication Property)

$$\mathbf{A}\otimes\mathbf{B}=\begin{bmatrix} a_{1,1}\mathbf{B} & a_{1,2}\mathbf{B} & \cdots & \cdots & a_{1,N}\mathbf{B} \\ a_{2,1}\mathbf{B} & a_{2,2}\mathbf{B} & \cdots & \cdots & a_{2,N}\mathbf{B} \\ \vdots & \vdots & \ddots & & \vdots \\ \vdots & \vdots & & \ddots & \vdots \\ a_{M,1}\mathbf{B} & a_{M,2}\mathbf{B} & \cdots & \cdots & a_{M,N}\mathbf{B} \end{bmatrix} \quad \mathbf{C}\otimes\mathbf{D}=\begin{bmatrix} c_{1,1}\mathbf{D} & c_{1,2}\mathbf{D} & \cdots & \cdots & c_{1,K}\mathbf{D} \\ c_{2,1}\mathbf{D} & c_{2,2}\mathbf{D} & \cdots & \cdots & c_{2,K}\mathbf{D} \\ \vdots & \vdots & \ddots & & \vdots \\ \vdots & \vdots & & \ddots & \vdots \\ c_{N,1}\mathbf{D} & c_{M,2}\mathbf{D} & \cdots & \cdots & c_{N,K}\mathbf{D} \end{bmatrix}$$

If

$$\mathbf{G}=(\mathbf{A}\otimes\mathbf{B})(\mathbf{C}\otimes\mathbf{D})=\begin{bmatrix} \mathbf{G}_{1,1} & \mathbf{G}_{1,2} & \mathbf{G}_{1,3} & \cdots & \mathbf{G}_{1,K} \\ \mathbf{G}_{2,1} & \mathbf{G}_{2,2} & \mathbf{G}_{2,3} & \cdots & \mathbf{G}_{2,K} \\ \mathbf{G}_{3,1} & \mathbf{G}_{3,2} & \mathbf{G}_{3,3} & \cdots & \mathbf{G}_{3,K} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \mathbf{G}_{M,1} & \mathbf{G}_{M,2} & \mathbf{G}_{M,3} & \cdots & \mathbf{G}_{M,K} \end{bmatrix}$$

$$\mathbf{G}_{m,k}=\sum_{n=1}^{N}a_{m,n}\mathbf{B}c_{n,k}\mathbf{D}=\left(\sum_{n=1}^{N}a_{m,n}c_{n,k}\right)\mathbf{BD}=E_{m,k}\mathbf{BD}$$

where $E_{m,k}$ is an entry of $\mathbf{E}=\mathbf{AC}$

[Example 2]

$$\mathbf{A}=\begin{bmatrix}1 & 0\\ 2 & 3\end{bmatrix}\qquad \mathbf{B}=\begin{bmatrix}2 & 1\\ 1 & 0\end{bmatrix}\qquad \mathbf{C}=\begin{bmatrix}0 & 1\\ 2 & 1\end{bmatrix}\qquad \mathbf{D}=\begin{bmatrix}1 & 1\\ 0 & 2\end{bmatrix}$$

$$\begin{array}{c}\begin{array}{cc}B & 0\end{array}\\ \mathbf{A}\otimes\mathbf{B}=\left[\begin{array}{cc|cc}2 & 1 & 0 & 0\\ 1 & 0 & 0 & 0\\ \hline 4 & 2 & 6 & 3\\ 2 & 0 & 3 & 0\end{array}\right]\\ \begin{array}{cc}2B & 3B\end{array}\end{array}\qquad \begin{array}{c}\begin{array}{cc}0 & D\end{array}\\ \mathbf{C}\otimes\mathbf{D}=\left[\begin{array}{cc|cc}0 & 0 & 1 & 1\\ 0 & 0 & 0 & 2\\ \hline 2 & 2 & 1 & 1\\ 0 & 4 & 0 & 2\end{array}\right]\\ \begin{array}{cc}2D & D\end{array}\end{array}$$

$$(\mathbf{A}\otimes\mathbf{B})(\mathbf{C}\otimes\mathbf{D})=\begin{bmatrix}0 & 0 & 2 & 4\\ 0 & 0 & 1 & 1\\ 12 & 24 & 10 & 20\\ 6 & 6 & 5 & 5\end{bmatrix}$$

$$\mathbf{AC}=\begin{bmatrix}0 & 1\\ 6 & 5\end{bmatrix}\qquad \mathbf{BD}=\begin{bmatrix}2 & 4\\ 1 & 1\end{bmatrix}\qquad \begin{array}{c}\begin{array}{cc}0\cdot BD & 1\cdot BD\end{array}\\ \mathbf{AC}\otimes\mathbf{BD}=\left[\begin{array}{cc|cc}0 & 0 & 2 & 4\\ 0 & 0 & 1 & 1\\ \hline 12 & 24 & 10 & 20\\ 6 & 6 & 5 & 5\end{array}\right]\\ \begin{array}{cc}6\,BD & 5\,BD\end{array}\end{array}$$

[Inverse Property]

$$\left(\mathbf{A}\otimes\mathbf{B}\right)^{-1}=\mathbf{A}^{-1}\otimes\mathbf{B}^{-1}$$

(Proof): $\left(\mathbf{A}^{-1}\otimes\mathbf{B}^{-1}\right)\left(\mathbf{A}\otimes\mathbf{B}\right)=\mathbf{A}^{-1}\mathbf{A}\otimes\mathbf{B}^{-1}\mathbf{B}=\mathbf{I}\otimes\mathbf{I}=\mathbf{I}$

page 524

[Eigenvector Property]

If

$$\mathbf{A}\mathbf{e}_1=\lambda_1\mathbf{e}_1,\qquad \mathbf{B}\mathbf{e}_2=\lambda_2\mathbf{e}_2$$

then

$$= (\mathbf{A}\mathbf{e}_1)\otimes(\mathbf{B}\mathbf{e}_2) = \lambda_1\mathbf{e}_1\otimes\lambda_2\mathbf{e}_2$$

$$\left(\mathbf{A}\otimes\mathbf{B}\right)\left(\mathbf{e}_1\otimes\mathbf{e}_2\right)=\lambda_1\lambda_2\left(\mathbf{e}_1\otimes\mathbf{e}_2\right)$$

That is, $\mathbf{e}_1\otimes\mathbf{e}_2$ is an eigenvector of $\mathbf{A}\otimes\mathbf{B}$

and $\lambda_1\lambda_2$ is the corresponding eigenvalue.

[Orthogonal Property]

If

$$\mathbf{AA}^{\mathrm{H}} = \mathbf{I}, \quad \mathbf{BB}^{\mathrm{H}} = \mathbf{I}, \quad \mathbf{C} = (\mathbf{A} \otimes \mathbf{B})$$

then

$$\mathbf{CC}^{\mathrm{H}} = \mathbf{I}$$

$$(A \otimes B)(A^H \otimes B^H) = (AA^H) \otimes (BB^H) = I \otimes I = I$$

[Rank Property]

If rank($\mathbf{A}$) = c_1, rank($\mathbf{B}$) = c_2, then

rank($\mathbf{A} \otimes \mathbf{B}$) = $c_1 c_2$.

[Determinant Property]

Suppose that size(**A**) = $M\times M$, size(**B**) = $N\times N$, then

$$\det(\mathbf{A}\otimes\mathbf{B}) = \det(\mathbf{A})^N \det(\mathbf{B})^M$$

(Proof can be done by eigenvector-eigenvalue decomposition)

If $\lambda_1, \lambda_2, \cdots \lambda_M$ are eigenvalues of A, $\det(A) = \lambda_1\lambda_2\cdots\lambda_M$

$\tau_1, \tau_2 \cdots \tau_N$ are eigenvalues of B $\det(B) = \tau_1\tau_2\cdots\tau_N$

From page 527, $\lambda_m\tau_n$ are eigenvalues of $A\otimes B$ $m = 1, 2, \cdots M$, $n = 1, 2, \cdots N$

$\det(A\otimes B) = (\lambda_1\tau_1 \times \lambda_1\tau_2 \times \cdots \times \lambda_1\tau_N) \times (\lambda_2\tau_1 \times \lambda_2\tau_2 \times \cdots \times \lambda_2\tau_N) \times \cdots$
$\times (\lambda_M\tau_1 \times \lambda_M\tau_2 \times \cdots \times \lambda_M\tau_N)$
$= \lambda_1^N\lambda_2^N\cdots\lambda_M^N\,\tau_1^M\tau_2^M\cdots\tau_N^M = \det(A)^N\det(B)^M$

6.2.2 Element-Wise Product

Suppose that **A** and **B** are two matrices and their sizes are the same ($M \times N$). Then

$$\mathbf{A} \circ \mathbf{B} = \left(\begin{bmatrix} a_{1,1} & a_{1,2} & \cdots & \cdots & a_{1,N} \\ a_{2,1} & a_{2,2} & \cdots & \cdots & a_{2,N} \\ \vdots & \vdots & \ddots & & \vdots \\ \vdots & \vdots & & \ddots & \vdots \\ a_{M,1} & a_{M,2} & \cdots & \cdots & a_{M,N} \end{bmatrix} \circ \begin{bmatrix} b_{1,1} & b_{1,2} & \cdots & \cdots & b_{1,N} \\ b_{2,1} & b_{2,2} & \cdots & \cdots & b_{2,N} \\ \vdots & \vdots & \ddots & & \vdots \\ \vdots & \vdots & & \ddots & \vdots \\ b_{M,1} & b_{M,2} & \cdots & \cdots & b_{M,N} \end{bmatrix} \right)$$

$$= \begin{bmatrix} a_{1,1}b_{1,1} & a_{1,2}b_{1,2} & \cdots & \cdots & a_{1,N}b_{1,N} \\ a_{2,1}b_{2,1} & a_{2,2}b_{2,2} & \cdots & \cdots & a_{2,N}b_{2,N} \\ \vdots & \vdots & \ddots & & \vdots \\ \vdots & \vdots & & \ddots & \vdots \\ a_{M,1}b_{M,1} & a_{M,2}b_{M,2} & \cdots & \cdots & a_{M,N}b_{M,N} \end{bmatrix}$$

[Example 3]

$$\mathbf{A} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 4 \\ 1 & 3 & 5 \end{bmatrix} \qquad \mathbf{B} = \begin{bmatrix} 1 & 2 & 1 \\ 1 & -2 & 1 \\ 2 & -1 & -1 \end{bmatrix}$$

$$\mathbf{A} \circ \mathbf{B} = \begin{bmatrix} 1\cdot 1 & 1\cdot 2 & 1\cdot 1 \\ 1\cdot 1 & 2\cdot(-2) & 4\cdot 1 \\ 1\cdot 2 & 3\cdot(-1) & 5\cdot(-1) \end{bmatrix} = \begin{bmatrix} 1 & 2 & 1 \\ 1 & -4 & 4 \\ 2 & -3 & -5 \end{bmatrix}$$

[Properties of the Element-Wise Product]

Commutative Property

$$\mathbf{A} \circ \mathbf{B} = \mathbf{B} \circ \mathbf{A}$$

Associative Property

$$(\mathbf{A} \circ \mathbf{B}) \circ \mathbf{C} = \mathbf{A} \circ (\mathbf{B} \circ \mathbf{C})$$

Distributive Property

$$\mathbf{A} \circ (\mathbf{B}+\mathbf{C}) = \mathbf{A} \circ \mathbf{B} + \mathbf{A} \circ \mathbf{C}$$

$$(\mathbf{A}+\mathbf{B}) \circ \mathbf{C} = \mathbf{A} \circ \mathbf{C} + \mathbf{B} \circ \mathbf{C}$$

6.3 Jordan-Canonical Form

6.3.1 Generalization for Eigenvector-Eigenvalue Decomposition

[Eigenvalue-Eigenvector Decomposition]

$$\mathbf{A} = \mathbf{E}\mathbf{D}\mathbf{E}^{-1}$$

where

$$\mathbf{E} = \begin{bmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \cdots & \mathbf{e}_{N-1} & \mathbf{e}_N \end{bmatrix} \qquad \mathbf{D} = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 & 0 \\ 0 & \lambda_2 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & \lambda_{N-1} & 0 \\ 0 & 0 & \cdots & 0 & \lambda_N \end{bmatrix}$$

However, not all the matrices have a complete eigenvector set.

[Example 1]

$$\mathbf{A} = \begin{bmatrix} 2 & 1 & 6 \\ 0 & 2 & 5 \\ 0 & 0 & 2 \end{bmatrix}$$

$\det(\mathbf{A} - \lambda\mathbf{I}) = (2-\lambda)^3$ eigenvalues: 2, 2, 2

$$(\mathbf{A} - 2\mathbf{I})\begin{bmatrix} k_1 \\ k_2 \\ k_3 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 6 \\ 0 & 0 & 5 \\ 0 & 0 & 0 \end{bmatrix}\begin{bmatrix} k_1 \\ k_2 \\ k_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

eigenvector: $\begin{bmatrix} k_1 \\ 0 \\ 0 \end{bmatrix} = k_1\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$

$$\underline{5k_3 = 0,} \quad \underline{k_2 + 6k_3 = 0} \implies k_2 = k_3 = 0$$

only one independent solution: $\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$ $\qquad \mathbf{e}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$

How do we perform eigenvector-eigenvalue decomposition for it?

original : $(A-\lambda I)e=0$

We try to solve

replace 0 by e_1

$$\underline{(\mathbf{A}-\lambda\mathbf{I})\mathbf{e_2}} = \underline{\mathbf{e_1}}, \qquad \begin{bmatrix} 0 & 1 & 6 \\ 0 & 0 & 5 \\ 0 & 0 & 0 \end{bmatrix} \overset{e_2}{\begin{bmatrix} k_1 \\ k_2 \\ k_3 \end{bmatrix}} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

$$5k_3 = 0, \quad k_2 + 6k_3 = 1$$

$k_3=0$, $k_2=1$

One of the solution is $\mathbf{e_2} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$ $\begin{bmatrix} k_1 \\ 1 \\ 0 \end{bmatrix}$

for simplification set $k_1=0$

e_2 is independent of e_1

The, we try to solve

$$(\mathbf{A}-\lambda\mathbf{I})\mathbf{e_3} = \mathbf{e_2}, \qquad \begin{bmatrix} 0 & 1 & 6 \\ 0 & 0 & 5 \\ 0 & 0 & 0 \end{bmatrix} \overset{e_3}{\begin{bmatrix} k_1 \\ k_2 \\ k_3 \end{bmatrix}} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$$

$$5k_3 = 1, \quad k_2 + 6k_3 = 0$$

general solution of e_3

$\begin{bmatrix} k_1 \\ -6/5 \\ 1/5 \end{bmatrix}$

One of the solution is $\mathbf{e_3} = \begin{bmatrix} 0 \\ -6/5 \\ 1/5 \end{bmatrix}$

for simplification, set $k_1=0$

e_3 is independent of e_1, e_2

Since

$(A-\lambda I)e_1=0$ $$\mathbf{A}\mathbf{e}_1 = \lambda \mathbf{e}_1$$

$(A-\lambda I)e_2=e_1$ $$\mathbf{A}\mathbf{e}_2 = \lambda \mathbf{e}_2 + \mathbf{e}_1$$

$(A-\lambda I)e_3=e_2$ $$\mathbf{A}\mathbf{e}_3 = \lambda \mathbf{e}_3 + \mathbf{e}_2$$

$$\mathbf{A}\underbrace{\begin{bmatrix}\mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3\end{bmatrix}}_{E} = \underbrace{\begin{bmatrix}\mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3\end{bmatrix}}_{E}\begin{bmatrix}\lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda\end{bmatrix}$$

$$\mathbf{A} = \mathbf{E}\begin{bmatrix}\lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda\end{bmatrix}\mathbf{E}^{-1} \quad \text{where} \quad \mathbf{E} = \begin{bmatrix}\mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3\end{bmatrix}$$

$$\mathbf{A} = \begin{bmatrix}1 & 0 & 0 \\ 0 & 1 & -6/5 \\ 0 & 0 & 1/5\end{bmatrix}\begin{bmatrix}2 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2\end{bmatrix}\begin{bmatrix}1 & 0 & 0 \\ 0 & 1 & -6/5 \\ 0 & 0 & 1/5\end{bmatrix}^{-1}$$

[Jordan-Canonical Form]

It try to decompose a matrix into

$$\mathbf{A} = \mathbf{E}\begin{bmatrix} \mathbf{D_1} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{D_2} & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{D_K} \end{bmatrix}\mathbf{E}^{-1}$$

Any matrix has the Jordan-Canonical form.

where $\mathbf{D_k}$ has the form of

$$\mathbf{D_k} = \lambda_k \mathbf{I} \qquad \text{or} \qquad \mathbf{D_k} = \begin{bmatrix} \lambda_k & 1 & 0 & \cdots & 0 \\ 0 & \lambda_k & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_k & 1 \\ 0 & 0 & \cdots & 0 & \lambda_k \end{bmatrix}$$

[Example 2]

$$\mathbf{A} = \begin{bmatrix} 4 & 0 & 0 & -2 \\ 0 & 4 & -2 & 0 \\ -1 & -1 & 4 & 0 \\ -1 & -1 & 0 & 4 \end{bmatrix}$$

$$\lambda = 2, 4, 4, 6$$

For $\lambda = 2$, $\quad \mathbf{e} = \begin{bmatrix} 1 & 1 & 1 & 1 \end{bmatrix}^T$

For $\lambda = 6$, $\quad \mathbf{e} = \begin{bmatrix} 1 & 1 & -1 & -1 \end{bmatrix}^T$

$$e = \begin{bmatrix} k_1 \\ -k_1 \\ 0 \\ 0 \end{bmatrix} \quad \text{set } k_1 = 1 \quad \begin{bmatrix} 1 \\ -1 \\ 0 \\ 0 \end{bmatrix}$$

For $\lambda = 4$,

$$(\mathbf{A} - 4\mathbf{I})\mathbf{e} = 0, \quad \begin{bmatrix} 0 & 0 & 0 & -2 \\ 0 & 0 & -2 & 0 \\ -1 & -1 & 0 & 0 \\ -1 & -1 & 0 & 0 \end{bmatrix} \underbrace{\begin{bmatrix} k_1 \\ k_2 \\ k_3 \\ k_4 \end{bmatrix}}_{e} = 0$$

$$-2k_4 = 0 \quad k_4 = 0$$

$$-2k_3 = 0 \quad k_3 = 0$$

$$-k_1 - k_2 = 0$$

$$\rightarrow k_2 = -k_1$$

For $\lambda = 4$, there is only one linearly independent eigenvector

$$\mathbf{e_1} = \begin{bmatrix} 1 & -1 & 0 & 0 \end{bmatrix}^T$$

Then, we try to solve

replace $(A-\lambda I)e=0$
by $(A-\lambda I)e_2=e_1$

$$(\mathbf{A}-4\mathbf{I})\mathbf{e_2} = \mathbf{e_1}, \quad \begin{bmatrix} 0 & 0 & 0 & -2 \\ 0 & 0 & -2 & 0 \\ -1 & -1 & 0 & 0 \\ -1 & -1 & 0 & 0 \end{bmatrix} \overset{e_2}{\begin{bmatrix} k_1 \\ k_2 \\ k_3 \\ k_4 \end{bmatrix}} = \begin{bmatrix} 1 \\ -1 \\ 0 \\ 0 \end{bmatrix}$$

$$-2k_4 = 1$$
$$-2k_3 = -1$$
$$-k_1 - k_2 = 0$$

One of the solution is

$$\mathbf{e_2} = \begin{bmatrix} 0 & 0 & 1/2 & -1/2 \end{bmatrix}^T$$

$$\begin{bmatrix} k_1 \\ -k_1 \\ 1/2 \\ -1/2 \end{bmatrix}$$

set $k_1=0$

e_2 is independent of e_1

Therefore,

$$\mathbf{A} = \mathbf{E}\begin{bmatrix} 2 & 0 & 0 & 0 \\ 0 & 4 & 1 & 0 \\ 0 & 0 & 4 & 0 \\ 0 & 0 & 0 & 6 \end{bmatrix}\mathbf{E}^{-1}$$

$$\overbrace{\begin{bmatrix} 1/4 & 0 & 0 & 0 \\ 0 & 1/2 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 1/4 \end{bmatrix} E^T}^{E^{-1}} E = I$$

$$E^T E = \begin{bmatrix} 4 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 1/2 & 0 \\ 0 & 0 & 0 & 4 \end{bmatrix}$$

where

$$\mathbf{E} = \begin{array}{c} \begin{array}{cccc} \lambda=2 & e_1 \quad \lambda=4 & e_2 & \lambda=6 \end{array} \\ \begin{bmatrix} 1 & 1 & 0 & 1 \\ 1 & -1 & 0 & 1 \\ 1 & 0 & 1/2 & -1 \\ 1 & 0 & -1/2 & -1 \end{bmatrix} \end{array}$$

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

[Example 3]

$$\mathbf{A}=\begin{bmatrix}4&0&1&2&1\\0&4&1&-2&1\\0&0&3&0&1\\0&0&0&4&0\\0&0&1&0&3\end{bmatrix}$$

$$\det\left(\begin{bmatrix}4-\lambda&0&1&2&1\\0&4-\lambda&1&-2&1\\0&0&3-\lambda&0&1\\0&0&0&4-\lambda&0\\0&0&1&0&3-\lambda\end{bmatrix}\right)$$

$$=(4-\lambda)^2\det\left(\begin{bmatrix}3-\lambda&0&1\\0&4-\lambda&0\\1&0&3-\lambda\end{bmatrix}\right)$$

$$=(4-\lambda)^2\left((3-\lambda)^2(4-\lambda)-(4-\lambda)\right)$$

$$=(4-\lambda)^3(\lambda^2-6\lambda+8)$$

$$=-(\lambda-4)^3(\lambda-2)(\lambda-4)$$

(Solution): Since

$$\det(\mathbf{A}-\lambda\mathbf{I})=-(\lambda-2)(\lambda-4)^4$$

the eigenvalues are 2, 4, 4, 4, 4.

The eigenvector corresponding to $\lambda=2$ is

$$(\mathbf{A}-2\mathbf{I})\mathbf{e}=\mathbf{0},\qquad \begin{bmatrix}2&0&1&2&1\\0&2&1&-2&1\\0&0&1&0&1\\0&0&0&2&0\\0&0&1&0&1\end{bmatrix}\begin{bmatrix}k_1\\k_2\\k_3\\k_4\\k_5\end{bmatrix}=\begin{bmatrix}0\\0\\0\\0\\0\end{bmatrix}$$

$2k_1+k_3+2k_4+k_5=0$

$2k_2+k_3-2k_4+k_5=0$

$k_3+k_5=0$

$2k_4=0$

One of the solution is $\begin{bmatrix}0\\0\\1\\0\\-1\end{bmatrix}$

set $k_3=1$

$k_1=k_2=0$

$k_3=1$

$k_4=0$

$k_5=-1$

The eigenvectors corresponding to $\lambda = 4$ are

general solution

$$e = \begin{bmatrix} k_1 \\ k_2 \\ 0 \\ 0 \\ 0 \end{bmatrix} = k_1 e_{1,a} + k_2 e_{1,b}$$

$$(\mathbf{A} - 4\mathbf{I})\mathbf{e} = \mathbf{0}, \qquad \begin{bmatrix} 0 & 0 & 1 & 2 & 1 \\ 0 & 0 & 1 & -2 & 1 \\ 0 & 0 & -1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & -1 \end{bmatrix} \begin{bmatrix} k_1 \\ k_2 \\ k_3 \\ k_4 \\ k_5 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

$k_3 + k_5 = 0$, $2k_3 = 0$, $k_3 = 0$

(① − ②)/4

$$\begin{cases} k_3 + 2k_4 + k_5 = 0 & ① \\ k_3 - 2k_4 + k_5 = 0 & ② \\ -k_3 + k_5 = 0 \\ \cancel{k_3 - k_5 = 0} \end{cases} \Rightarrow \begin{cases} k_3 + 2k_4 + k_5 = 0 \\ k_4 = 0 \\ k_3 = k_5 \end{cases} \Rightarrow k_3 = k_4 = k_5 = 0$$

Therefore, there are only two linearly independent solutions:

$$\mathbf{e}_{1,a} = [1 \quad 0 \quad 0 \quad 0 \quad 0]^T \quad \text{and} \quad \mathbf{e}_{1,b} = [0 \quad 1 \quad 0 \quad 0 \quad 0]^T$$

If we set

$$\mathbf{e}_{\mathbf{1,a}} = \begin{bmatrix}1 & 0 & 0 & 0 & 0\end{bmatrix}^T$$

then

$e_{2,a}$

$$(\mathbf{A}-4\mathbf{I})\mathbf{e}_{\mathbf{2,a}} = \mathbf{e}_{\mathbf{1,a}}, \quad \begin{bmatrix} 0 & 0 & 1 & 2 & 1 \\ 0 & 0 & 1 & -2 & 1 \\ 0 & 0 & -1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & -1 \end{bmatrix} \begin{bmatrix} k_1 \\ k_2 \\ k_3 \\ k_4 \\ k_5 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

$k_3 + k_5 = 1/2 \quad 2k_3 = 1/2$

$$\begin{cases} k_3 + 2k_4 + k_5 = 1 \\ k_3 - 2k_4 + k_5 = 0 \\ -k_3 + k_5 = 0 \\ k_3 - k_5 = 0 \end{cases} \Rightarrow \begin{cases} k_3 - 2k_4 + k_5 = 0 \\ 4k_4 = 1 \\ k_3 = k_5 \end{cases} \Rightarrow k_3 = k_4 = k_5 = 1/4$$

One of the solution is $\mathbf{e}_{\mathbf{2,a}} = \begin{bmatrix}0 & 0 & 1/4 & 1/4 & 1/4\end{bmatrix}^T$

(We set $k_1 = k_2 = 0$ for simplification)

If we set

$$\mathbf{e}_{\mathbf{1,b}} = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \end{bmatrix}^T$$

then

e_{2b}

$$\left(\mathbf{A}-4\mathbf{I}\right)\mathbf{e}_{\mathbf{2,b}} = \mathbf{e}_{\mathbf{1,b}}, \quad \begin{bmatrix} 0 & 0 & 1 & 2 & 1 \\ 0 & 0 & 1 & -2 & 1 \\ 0 & 0 & -1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & -1 \end{bmatrix} \begin{bmatrix} k_1 \\ k_2 \\ k_3 \\ k_4 \\ k_5 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\begin{cases} k_3 + 2k_4 + k_5 = 0 \\ k_3 - 2k_4 + k_5 = 1 \\ -k_3 + k_5 = 0 \\ k_3 - k_5 = 0 \end{cases} \Rightarrow \begin{cases} k_3 + 2k_4 + k_5 = 0 \\ 4k_4 = -1 \\ k_3 = k_5 \end{cases} \Rightarrow k_3 = -k_4 = k_5 = 1/4$$

One of the solution is $\mathbf{e}_{\mathbf{2,b}} = \begin{bmatrix} 0 & 0 & 1/4 & -1/4 & 1/4 \end{bmatrix}^T$

k_1 k_2

set $k_1 = k_2 = 0$

Therefore, the Jordan-Canonical form of **A** is

Note! (2 groups for $\lambda = 4$)

for e_{1a}

for e_{1b}

$$\mathbf{A} = \mathbf{E}\begin{bmatrix} 4 & 1 & 0 & 0 & 0 \\ 0 & 4 & 0 & 0 & 0 \\ 0 & 0 & 4 & 1 & 0 \\ 0 & 0 & 0 & 4 & 0 \\ 0 & 0 & 0 & 0 & 2 \end{bmatrix}\mathbf{E}^{-1}$$

where

$$\mathbf{E} = \begin{array}{c} \overbrace{\begin{array}{cccc} e_{1a} & e_{2a} & e_{1b} & e_{2b} \end{array}}^{\lambda = 4} \;\; \lambda = 2 \\ \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 1/4 & 0 & 1/4 & 1 \\ 0 & 1/4 & 0 & -1/4 & 0 \\ 0 & 1/4 & 0 & 1/4 & -1 \end{bmatrix} \end{array} \qquad \mathbf{E}^{-1} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 2 & 1 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & -2 & 1 \\ 0 & 0 & 1/2 & 0 & -1/2 \end{bmatrix}$$

6.3.2 Power of the Jordan Canonical Form

If

$$\mathbf{A} = \mathbf{E}\mathbf{D}\mathbf{E}^{-1}$$

then

$$\mathbf{A}^{\alpha} = \underline{\mathbf{E}\mathbf{D}^{\alpha}\mathbf{E}^{-1}}$$

Problem: How do we determine $\mathbf{D}^{\alpha}$ for the Jordan-canonical form?

$$\mathbf{D} = \begin{bmatrix} \mathbf{D}_1 & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{D}_2 & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{D}_K \end{bmatrix}$$

where $\mathbf{D}_\mathbf{k} = \lambda_k \mathbf{I}$

or

$$\mathbf{D}_\mathbf{k} = \begin{bmatrix} \lambda_k & 1 & 0 & \cdots & 0 \\ 0 & \lambda_k & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_k & 1 \\ 0 & 0 & \cdots & 0 & \lambda_k \end{bmatrix}$$

$$\mathbf{D} = \begin{bmatrix} \mathbf{D}_1 & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{D}_2 & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{D}_\mathbf{K} \end{bmatrix} \qquad \mathbf{D}^\alpha = \begin{bmatrix} \mathbf{D}_1^\alpha & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{D}_2^\alpha & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{D}_\mathbf{K}^\alpha \end{bmatrix}$$

$$\mathbf{D}_\mathbf{k} = \begin{bmatrix} \lambda_k & 1 & 0 & \cdots & 0 \\ 0 & \lambda_k & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_k & 1 \\ 0 & 0 & \cdots & 0 & \lambda_k \end{bmatrix} = \lambda_k \mathbf{I} + \mathbf{U} \quad \text{where} \quad \mathbf{U} = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & 1 \\ 0 & 0 & \cdots & 0 & 0 \end{bmatrix}$$

Therefore,

$$(x+y)^\alpha = \sum_\beta \binom{\alpha}{\beta} x^\beta y^{\alpha-\beta}$$

$$\mathbf{D}_\mathbf{k}^\alpha = \sum_{\beta=0}^{\alpha} \binom{\alpha}{\beta} \lambda_k^{\alpha-\beta} \mathbf{U}^\beta$$

$$\binom{\alpha}{\beta} = \frac{\alpha!}{\beta!(\alpha-\beta)!}$$

$$= 1 \cdot \lambda_k^\alpha I + \alpha \lambda_k^{\alpha-1} U + \frac{\alpha(\alpha-1)}{2} \lambda^{\alpha-2} U^2 + \cdots \binom{\alpha}{l} \lambda^{\alpha-l} U^l$$

$$l = \min(\alpha, M-1)$$

Note:

$U^0 = I$

If $$\mathbf{U} = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$ then

$$\mathbf{U}^2 = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \quad \mathbf{U}^3 = \begin{bmatrix} 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \quad \mathbf{U}^4 = \begin{bmatrix} 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

$$\mathbf{U}^\beta = \mathbf{0} \quad \textit{for } \beta \geq 5$$

In general, if size of $\mathbf{D}_k$ is MxM then

$$\mathbf{U}^{\beta}[m,n]=1 \quad when \quad n-m=\beta$$

$$\mathbf{U}^{\beta}[m,n]=0 \quad otherwise$$

Also,

$$\mathbf{U}^{\beta}=\mathbf{0} \quad when\ \beta \geq M$$

Therefore,

$$\mathbf{D}_{\mathbf{k}}^{\alpha}[n,n]=\lambda_k^{\alpha},$$

$\beta = n-m$

$$\mathbf{D}_{\mathbf{k}}^{\alpha}[m,n]=\binom{\alpha}{n-m}\lambda_k^{\alpha-n+m} \quad if\ 1 \leq n-m \leq \alpha$$

$$\mathbf{D}_{\mathbf{k}}^{\alpha}[m,n]=0 \quad if\ n-m<0\ \ or\ \ n-m>\alpha$$

$\min(\alpha, M-1)$

$$C^{\alpha}_{\beta} = \binom{\alpha}{\beta} = \frac{\alpha!}{\beta!\,(\alpha-\beta)!}$$

$$\mathbf{D}_{\mathbf{k}}^{\alpha} = \begin{bmatrix} \lambda_k^{\alpha} & C_1^{\alpha}\lambda_k^{\alpha-1} & C_2^{\alpha}\lambda_k^{\alpha-2} & \cdots & C_{\alpha}^{\alpha}\lambda_k^{0} & 0 & \cdots & 0 \\ 0 & \lambda_k^{\alpha} & C_1^{\alpha}\lambda_k^{\alpha-1} & C_2^{\alpha}\lambda_k^{\alpha-2} & \ddots & C_{\alpha}^{\alpha}\lambda_k^{0} & \ddots & \vdots \\ 0 & 0 & \lambda_k^{\alpha} & C_1^{\alpha}\lambda_k^{\alpha-1} & \ddots & \ddots & \ddots & 0 \\ 0 & 0 & 0 & \lambda_k^{\alpha} & \ddots & C_2^{\alpha}\lambda_k^{\alpha-2} & \ddots & C_{\alpha}^{\alpha}\lambda_k^{0} \\ 0 & 0 & 0 & 0 & \ddots & C_1^{\alpha}\lambda_k^{\alpha-1} & C_2^{\alpha}\lambda_k^{\alpha-2} & \vdots \\ \vdots & \vdots & \vdots & \vdots & \ddots & \lambda_k^{\alpha} & C_1^{\alpha}\lambda_k^{\alpha-1} & C_2^{\alpha}\lambda_k^{\alpha-2} \\ 0 & 0 & 0 & 0 & \cdots & 0 & \lambda_k^{\alpha} & C_1^{\alpha}\lambda_k^{\alpha-1} \\ 0 & 0 & 0 & 0 & \cdots & 0 & 0 & \lambda_k^{\alpha} \end{bmatrix}$$

$$\mathbf{D}_{\mathbf{k}}^{\alpha}[n,n] = \lambda_k^{\alpha},$$

$$\mathbf{D}_{\mathbf{k}}^{\alpha}[m,n] = \binom{\alpha}{n-m}\lambda_k^{\alpha-n+m} \quad \text{if } 1 \le n-m \le \alpha$$

$$\mathbf{D}_{\mathbf{k}}^{\alpha}[m,n] = 0 \quad \text{if } n-m < 0 \text{ or } n-m > \alpha$$

[Example 4] If

$$\mathbf{A}=\begin{bmatrix}4 & 0 & 0 & -2\\ 0 & 4 & -2 & 0\\ -1 & -1 & 4 & 0\\ -1 & -1 & 0 & 4\end{bmatrix}$$

then determine $\mathbf{A}^5$

(Solution): From Example 2,

$$\mathbf{A}=\mathbf{E}\begin{bmatrix}2 & 0 & 0 & 0\\ 0 & 4 & 1 & 0\\ 0 & 0 & 4 & 0\\ 0 & 0 & 0 & 6\end{bmatrix}\mathbf{E}^{-1}$$

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

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

$$\mathbf{A}^5 = \mathbf{E}\begin{bmatrix} 2^5 & 0 & 0 & 0 \\ 0 & 4^5 & 4^4\binom{5}{1} & 0 \\ 0 & 0 & 4^5 & 0 \\ 0 & 0 & 0 & 6^5 \end{bmatrix}\mathbf{E}^{-1} = \mathbf{E}\begin{bmatrix} 32 & 0 & 0 & 0 \\ 0 & 1024 & 1280 & 0 \\ 0 & 0 & 1024 & 0 \\ 0 & 0 & 0 & 7776 \end{bmatrix}\mathbf{E}^{-1}$$

$$\mathbf{A}^5 = \begin{bmatrix} 2464 & 1440 & -656 & -3216 \\ 1440 & 2464 & -3216 & -656 \\ -1936 & -1936 & 2464 & 1440 \\ -1936 & -1936 & 1440 & 2464 \end{bmatrix}$$

6.4 Functions of Matrices (只教不考)

Functions of Matrices are also called spectral mapping.

6.4.1 Exponential of a Matrix

From the Taylor series

$\sum_{n=0}^{\infty} \frac{x^n}{n!}$

$$\exp(x) = 1 + \frac{x}{1!} + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots\cdots$$

therefore,

$$\exp(\mathbf{A}) = 1 + \frac{\mathbf{A}}{1!} + \frac{\mathbf{A}^2}{2!} + \frac{\mathbf{A}^3}{3!} + \cdots\cdots$$

If

$$\mathbf{A} = \mathbf{E}\mathbf{D}\mathbf{E}^{-1}$$

$A^n = ED^nE^{-1}$

then

$$\exp(\mathbf{A}) = \mathbf{E}\left(\mathbf{I} + \frac{\mathbf{D}}{1!} + \frac{\mathbf{D}^2}{2!} + \frac{\mathbf{D}^3}{3!} + \cdots\cdots\right)\mathbf{E}^{-1} = \mathbf{E}\exp(\mathbf{D})\mathbf{E}^{-1}$$

$$\exp(\mathbf{D}) = \mathbf{I} + \frac{\mathbf{D}}{1!} + \frac{\mathbf{D}^2}{2!} + \frac{\mathbf{D}^3}{3!} + \cdots\cdots$$

(Case 1): If **A** has a complete eigenvector set, then

$$\mathbf{D} = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_N \end{bmatrix}$$

$$\exp(\mathbf{D}) = \mathbf{I} + \frac{\mathbf{D}}{1!} + \frac{\mathbf{D}^2}{2!} + \frac{\mathbf{D}^3}{3!} + \cdots\cdots = \begin{bmatrix} 1+\lambda_1+\frac{\lambda_1^2}{2!}+\cdots & 0 & \cdots & 0 \\ 0 & 1+\lambda_2+\frac{\lambda_2^2}{2!}+\cdots & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1+\lambda_N+\frac{\lambda_N^2}{2!}+\cdots \end{bmatrix}$$

$$= \begin{bmatrix} \exp(\lambda_1) & 0 & \cdots & 0 \\ 0 & \exp(\lambda_2) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \exp(\lambda_N) \end{bmatrix}$$

(Case 2): If **A** does not have a complete eigenvector set, then

$$\mathbf{D}=\begin{bmatrix}\mathbf{D_1} & \mathbf{0} & \cdots & \mathbf{0}\\ \mathbf{0} & \mathbf{D_2} & \cdots & \mathbf{0}\\ \vdots & \vdots & \ddots & \vdots\\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{D_K}\end{bmatrix} \quad \mathbf{D_k}=\lambda_k\mathbf{I} \text{ or } \mathbf{D_k}=\begin{bmatrix}\lambda_k & 1 & 0 & \cdots & 0\\ 0 & \lambda_k & 1 & \cdots & 0\\ \vdots & \vdots & \ddots & \ddots & \vdots\\ 0 & 0 & \cdots & \lambda_k & 1\\ 0 & 0 & \cdots & 0 & \lambda_k\end{bmatrix}$$

$$\exp(D)= \mathbf{I}+\frac{\mathbf{D}}{1!}+\frac{\mathbf{D}^2}{2!}+\frac{\mathbf{D}^3}{3!}+\cdots\cdots=\begin{bmatrix}\mathbf{S_1} & \mathbf{0} & \cdots & \mathbf{0}\\ \mathbf{0} & \mathbf{S_2} & \cdots & \mathbf{0}\\ \vdots & \vdots & \ddots & \vdots\\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{S_K}\end{bmatrix}$$

where $\mathbf{S_k}=\exp(\mathbf{D_k})=\mathbf{I}+\frac{\mathbf{D_k}}{1!}+\frac{\mathbf{D_k}^2}{2!}+\frac{\mathbf{D_k}^3}{3!}+\cdots\cdots$

$$\mathbf{S_k} = \mathbf{I} + \frac{\mathbf{D_k}}{1!} + \frac{\mathbf{D_k}^2}{2!} + \frac{\mathbf{D_k}^3}{3!} + \cdots\cdots \quad \text{If} \quad \mathbf{D_k} = \begin{bmatrix} \lambda_k & 1 & 0 & \cdots & 0 \\ 0 & \lambda_k & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_k & 1 \\ 0 & 0 & \cdots & 0 & \lambda_k \end{bmatrix}$$

$$\underline{S_k = \sum_{\alpha=0}^{\infty} \frac{D_k^\alpha}{\alpha!}}$$

$$D_k^2 = \begin{bmatrix} \lambda_k^2 & 2\lambda_k & 1 & \ddots & 0 \\ & \lambda_k^2 & 2\lambda_k & \ddots & 1 \\ & & \ddots & \ddots & 2\lambda_k \\ & & & & \lambda_k^2 \end{bmatrix}$$

$$D_k^3 = \begin{bmatrix} \lambda_k^3 & 3\lambda_k^2 & 3\lambda_k & 1 & \ddots \; 1 \\ & \lambda_k^3 & \ddots & \ddots & 1 \\ & & & \ddots & 3\lambda_k^2 \\ & & & \ddots & \lambda_k^3 \end{bmatrix}$$

then $\mathbf{S_k}[m,n] = 0 \qquad \text{if } m > n,$

$$\mathbf{S_k}[n,n] = 1 + \frac{\lambda_k}{1!} + \frac{\lambda_k^2}{2!} + \frac{\lambda_k^3}{3!} + \cdots = \exp(\lambda_k)$$

$$\underbrace{\mathbf{S_k}[m,n]}_{n>m} = \sum_{\alpha=n-m}^{\infty} \frac{1}{\alpha!} \underline{\binom{\alpha}{n-m} \lambda_k^{\alpha-n+m}}$$

$D_k^\alpha(m,n)$
page 549

$$= \sum_{\alpha=n-m}^{\infty} \frac{1}{(\alpha-n+m)!} \frac{1}{(n-m)!} \lambda_k^{\alpha-n+m} = \frac{1}{(n-m)!} \sum_{\gamma=0}^{\infty} \frac{1}{\gamma!} \lambda_k^{\gamma}$$

$\gamma = \alpha - n + m$

$$= \frac{1}{(n-m)!} \exp(\lambda_k)$$

Therefore,

$$\exp\left(\mathbf{D_k}\right)=\mathbf{S_k}=\begin{bmatrix} e^{\lambda_k} & \frac{1}{1!}e^{\lambda_k} & \frac{1}{2!}e^{\lambda_k} & \cdots & \frac{1}{(M-1)!}e^{\lambda_k} \\ 0 & e^{\lambda_k} & \frac{1}{1!}e^{\lambda_k} & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & \frac{1}{2!}e^{\lambda_k} \\ 0 & 0 & \cdots & e^{\lambda_k} & \frac{1}{1!}e^{\lambda_k} \\ 0 & 0 & \cdots & 0 & e^{\lambda_k} \end{bmatrix}$$

[Example 1] If

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

determine $\exp(\mathbf{A})$

(Solution): The eigenvalues of **A** are 2, 2, 2.
The Jordan-Canonical form of **A** is

$$\mathbf{A} = \mathbf{EDE}^{-1} \qquad \mathbf{D} = \begin{bmatrix} 2 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{bmatrix}$$

$$\mathbf{E} = \begin{bmatrix} 1 & 0 & 1/2 \\ 0 & 1/2 & 0 \\ 1 & 0 & 0 \end{bmatrix}, \quad \mathbf{E}^{-1} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 2 & 0 \\ 2 & 0 & -2 \end{bmatrix}$$

$$\exp(\mathbf{A}) = \mathbf{E}\exp(\mathbf{D})\mathbf{E}^{-1}$$

where $$\exp(\mathbf{D}) = \begin{bmatrix} e^2 & e^2 & e^2/2 \\ 0 & e^2 & e^2 \\ 0 & 0 & e^2 \end{bmatrix}$$

Therefore,

$$\exp(\mathbf{A}) = \begin{bmatrix} 2e^2 & 2e^2 & -e^2 \\ e^2 & e^2 & -e^2 \\ e^2 & 2e^2 & 0 \end{bmatrix}$$

6.4.2 Functions Using Taylor Series

(這個 subsection 只教不考)

(1) cos(**A**)

To derive $\cos(\mathbf{A})$

we can apply $\cos x = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \cdots\cdots$

$$\cos(\mathbf{A}) = \mathbf{I} - \frac{\mathbf{A}^2}{2!} + \frac{\mathbf{A}^4}{4!} - \frac{\mathbf{A}^6}{6!} + \cdots\cdots$$

$$= \mathbf{E}\left(\mathbf{I} - \frac{\mathbf{D}^2}{2!} + \frac{\mathbf{D}^4}{4!} - \frac{\mathbf{D}^6}{6!} + \cdots\cdots\right)\mathbf{E}^{-1} = \mathbf{E}\cos(\mathbf{D})\mathbf{E}^{-1}$$

where $\cos(\mathbf{D}) = \left(\mathbf{I} - \frac{\mathbf{D}^2}{2!} + \frac{\mathbf{D}^4}{4!} - \frac{\mathbf{D}^6}{6!} + \cdots\cdots\right)$

if $\mathbf{A} = \mathbf{EDE}^{-1}$

(Case 1) When **A** has a complete eigenvector set, then

$$\mathbf{D}=\begin{bmatrix}\lambda_1 & 0 & \cdots & 0\\ 0 & \lambda_2 & \cdots & 0\\ \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & \cdots & \lambda_N\end{bmatrix}$$

$$\mathbf{I}-\frac{\mathbf{D}^2}{2!}+\frac{\mathbf{D}^4}{4!}-\cdots\cdots=\begin{bmatrix}1-\frac{\lambda_1^2}{2!}+\frac{\lambda_1^2}{4!}-\cdots & 0 & \cdots & 0\\ 0 & 1-\frac{\lambda_2^2}{2!}+\frac{\lambda_2^2}{4!}-\cdots & \cdots & 0\\ \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & \cdots & 1-\frac{\lambda_N^2}{2!}+\frac{\lambda_N^2}{4!}-\cdots\end{bmatrix}$$

$$\cos(\mathbf{D})=\begin{bmatrix}\cos(\lambda_1) & 0 & \cdots & 0\\ 0 & \cos(\lambda_2) & \cdots & 0\\ \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & \cdots & \cos(\lambda_N)\end{bmatrix}$$

(Case 2) When **A** does not have a complete eigenvector set, then

$$\mathbf{D} = \begin{bmatrix} \mathbf{D_1} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{D_2} & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{D_K} \end{bmatrix} \qquad \mathbf{D_k} = \begin{bmatrix} \lambda_k & 1 & 0 & \cdots & 0 \\ 0 & \lambda_k & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_k & 1 \\ 0 & 0 & \cdots & 0 & \lambda_k \end{bmatrix}$$

$$\mathbf{I} - \frac{\mathbf{D}^2}{2!} + \frac{\mathbf{D}^4}{4!} - \frac{\mathbf{D}^6}{6!} + \cdots\cdots = \begin{bmatrix} \mathbf{S_1} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{S_2} & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{S_K} \end{bmatrix}$$

where $\quad \mathbf{S_k} = \cos\left(\mathbf{D_k}\right) = \mathbf{I} - \frac{\mathbf{D_k^2}}{2!} + \frac{\mathbf{D_k^4}}{4!} - \frac{\mathbf{D_k^6}}{6!} + \cdots\cdots$

$$\mathbf{S_k}[m,n]=0 \qquad \text{if } m>n,$$

$$\mathbf{S_k}[n,n]=1-\frac{\lambda_k^2}{2!}+\frac{\lambda_k^4}{4!}-\frac{\lambda_k^6}{6!}+\cdots=\cos(\lambda_k)$$

(1) If $n-m$ is positive and even,

$$\mathbf{S_k}[m,n]=\sum_{\alpha=(n-m)/2}^{\infty}\frac{(-1)^{\alpha}}{(2\alpha)!}\binom{2\alpha}{n-m}\lambda_k^{2\alpha-n+m}$$

$$=\sum_{\alpha=(n-m)/2}^{\infty}\frac{(-1)^{\alpha}}{(2\alpha-n+m)!(n-m)!}\lambda_k^{2\alpha-n+m}$$

$$=\sum_{\beta=0}^{\infty}\frac{(-1)^{\beta+(n-m)/2}}{(2\beta)!(n-m)!}\lambda_k^{2\beta} \qquad \beta=\alpha-\frac{n-m}{2}$$

$$=\frac{(-1)^{(n-m)/2}}{(n-m)!}\cos(\lambda_k)$$

(2) If $n - m$ is positive and odd,

$$\mathbf{S_k}[m,n] = \sum_{\alpha=(n-m+1)/2}^{\infty} \frac{(-1)^{\alpha}}{(2\alpha)!} \binom{2\alpha}{n-m} \lambda_k^{2\alpha-n+m}$$

$$= \sum_{\alpha=(n-m+1)/2}^{\infty} \frac{(-1)^{\alpha}}{(2\alpha-n+m)!(n-m)!} \lambda_k^{2\alpha-n+m} \qquad \beta = \alpha - \frac{n-m+1}{2}$$

$$= \sum_{\beta=0}^{\infty} \frac{(-1)^{\beta+(n-m+1)/2}}{(2\beta+1)!(n-m)!} \lambda_k^{2\beta+1} = \frac{(-1)^{(n-m+1)/2}}{(n-m)!} \sin(\lambda_k)$$

$$\cos(\mathbf{D_k}) = \mathbf{S_k} = \begin{bmatrix} \cos\lambda_k & \frac{-1}{1!}\sin\lambda_k & \frac{-1}{2!}\cos\lambda_k & \cdots & \cdots \\ 0 & \cos\lambda_k & \frac{-1}{1!}\sin\lambda_k & \ddots & \cdots \\ \vdots & \vdots & \ddots & \ddots & \frac{-1}{2!}\cos\lambda_k \\ 0 & 0 & \cdots & \cos\lambda_k & \frac{-1}{1!}\sin\lambda_k \\ 0 & 0 & \cdots & 0 & \cos\lambda_k \end{bmatrix}$$

(2) sin(A)

$$\sin x = \frac{x}{1!} - \frac{x^3}{3!} + \frac{x^5}{5!} - \cdots\cdots$$

(Case 1) When **A** has a complete eigenvector set and

$$\mathbf{A} = \mathbf{E}\mathbf{D}\mathbf{E}^{-1} \qquad \mathbf{D} = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_N \end{bmatrix}$$

then

$$S = \frac{A}{1!} - \frac{A^3}{3!} + \frac{A^5}{5!} - \cdots$$

$$\sin(\mathbf{A}) = \mathbf{E}\mathbf{S}\mathbf{E}^{-1} \qquad \mathbf{S} = \sin(\mathbf{D}) = \begin{bmatrix} \sin\lambda_1 & 0 & \cdots & 0 \\ 0 & \sin\lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sin\lambda_N \end{bmatrix}$$

(Case 2) When **A** does not have a complete eigenvector set and

$$\mathbf{A} = \mathbf{E}\mathbf{D}\mathbf{E}^{-1}$$

$$\mathbf{D} = \begin{bmatrix} \mathbf{D_1} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{D_2} & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{D_K} \end{bmatrix} \qquad \mathbf{D_k} = \begin{bmatrix} \lambda_k & 1 & 0 & \cdots & 0 \\ 0 & \lambda_k & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_k & 1 \\ 0 & 0 & \cdots & 0 & \lambda_k \end{bmatrix}$$

then

$$\sin(\mathbf{A}) = \mathbf{E}\mathbf{S}\mathbf{E}^{-1} \qquad \mathbf{S} = \begin{bmatrix} \mathbf{S_1} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{S_2} & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{S_K} \end{bmatrix}$$

$$\sin(\mathbf{A}) = \mathbf{E}\mathbf{S}\mathbf{E}^{-1} \qquad \mathbf{S} = \begin{bmatrix} \mathbf{S_1} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{S_2} & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{S_K} \end{bmatrix}$$

$$\mathbf{S_k}[m,n] = 0 \qquad \text{if } m > n,$$

$$\mathbf{S_k}[n,n] = \sin(\lambda_k)$$

$$\mathbf{S_k}[m,n] = \frac{(-1)^{(n-m-1)/2}}{(n-m)!}\cos(\lambda_k) \qquad \text{if } n-m \text{ is positive and odd,}$$

$$\mathbf{S_k}[m,n] = \frac{(-1)^{(n-m)/2}}{(n-m)!}\sin(\lambda_k) \qquad \text{if } n-m \text{ is positive and even}$$

(3) In general,

if

$$f(x) = f(0) + \frac{f'(0)}{1!}x + \frac{f''(0)}{2!}x^2 + \frac{f'''(0)}{3!}x^3 + \cdots\cdots\cdots$$

and

$$\mathbf{A} = \mathbf{EDE}^{-1}$$

then

$$f(\mathbf{A}) = \mathbf{E}f(\mathbf{D})\mathbf{E}^{-1}$$

where $f(\mathbf{D}) = f(0)\mathbf{I} + \frac{f'(0)}{1!}\mathbf{D} + \frac{f''(0)}{2!}\mathbf{D}^2 + \frac{f'''(0)}{3!}\mathbf{D}^3 + \cdots\cdots\cdots$

$$f(\mathbf{A}) = \mathbf{E} f(\mathbf{D}) \mathbf{E}^{-1}$$

where $$f(\mathbf{D}) = f(0) + \frac{f'(0)}{1!}\mathbf{D} + \frac{f''(0)}{2!}\mathbf{D}^2 + \frac{f'''(0)}{3!}\mathbf{D}^3 + \cdots\cdots\cdots$$

(i) When **A** has a complete eigenvector set and

$$\mathbf{D} = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_N \end{bmatrix}$$

then

$$f(\mathbf{D}) = \begin{bmatrix} f(\lambda_1) & 0 & \cdots & 0 \\ 0 & f(\lambda_2) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & f(\lambda_N) \end{bmatrix}$$

(ii) However, the form of $f(\mathbf{D})$ is rather complicated if **A** does not have a complete eigenvector set.

6.5 Generalized Norm

For a vector

$$\mathbf{x} = \begin{bmatrix} x[1] & x[2] & \cdots & x[N-1] & x[N] \end{bmatrix}$$

(1) Norm (L_α norm): $\|\mathbf{x}\|_\alpha = \sqrt[\alpha]{\sum_{n=0}^{N-1} |x[n]|^\alpha}$

(2) L_2 norm (conventional norm):

$\alpha = 2$

$$\|\mathbf{x}\|_2 = \sqrt{\sum_{n=0}^{N-1} |x[n]|^2}$$

(Physical meaning: Distance)

$$\|\mathbf{x}\|_\alpha = \sqrt[\alpha]{\sum_{n=0}^{N-1} |x[n]|^\alpha}$$

(3) L_1 norm: $$\|\mathbf{x}\|_1 = \sum_{n=0}^{N-1} |x[n]|$$

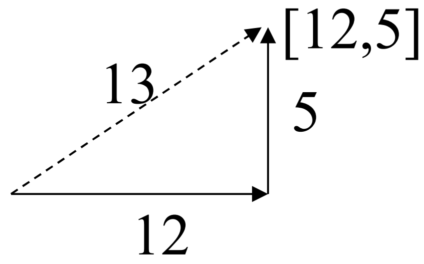
(Physical meaning: Sum of Amplitudes)

[Example 1]

$$\|[12,5]\|_2 = \sqrt{12^2+5^2} = 13$$

$\|[12,5]\|_\infty = 12$

$$\|[12,5]\|_1 = 12+5 = 17$$



$$\|\mathbf{x}\|_\alpha = \sqrt[\alpha]{\sum_{n=0}^{N-1} |x[n]|^\alpha}$$

(3) L_∞ norm: $\|\mathbf{x}\|_\infty = \mathrm{Max}\{|x[n]|\}$

(Physical meaning: The maximal amplitude)

Note: $$\lim_{\alpha\to\infty} \sum_{n=0}^{N-1} |x[n]|^\alpha \cong \left(Max\{|x[n]|\}\right)^\alpha$$

$$\lim_{\alpha\to\infty} \sqrt[\alpha]{\sum_{n=0}^{N-1} |x[n]|^\alpha} \cong Max\{|x[n]|\}$$

If there are M entries of $|x[n]|$ that equals to Max$|x[n]|$

$$\lim_{\alpha\to\infty} \sum_{n=0}^{N-1} |x[n]|^\alpha \cong M\left(Max\{|x[n]|\}\right)^\alpha$$

$$\lim_{\alpha\to\infty} \sqrt[\alpha]{\sum_{n=0}^{N-1} |x[n]|^\alpha} \cong M^{1/\alpha} Max\{|x[n]|\} = Max\{|x[n]|\}$$

$$\|\mathbf{x}\|_\alpha = \sqrt[\alpha]{\sum_{n=0}^{N-1} |x[n]|^\alpha}$$

(5) $\lim_{\alpha \to 0} (L_\alpha \text{ norm})^\alpha$ (Also call as the L_0 norm)

remove $\sqrt[\alpha]{\ }$

Note that $\lim_{\alpha \to 0} \left(\|\mathbf{x}\|_\alpha \right)^\alpha = \lim_{\alpha \to 0} \sum_{n=0}^{N-1} |x[n]|^\alpha$

$$\lim_{\alpha \to 0} \left(\|\mathbf{x}\|_\alpha \right)^\alpha = K$$

where K is the number of points such that $x[n] \neq 0$

(Physical meaning: <u>The number of nonzero points</u>)

The L_2 norm is easier for optimization, but it often happens that using the L_0 or the L_1 norm can obtain even better optimization results.

L_α-norm, $\alpha \geq 1 \Rightarrow$ convex

[Norms of a Matrix: Definition 1]

$$\|\mathbf{A}\|_\alpha = \sqrt[\alpha]{\sum_{m=0}^{M-1}\sum_{n=0}^{N-1}|A[m,n]|^\alpha}$$

L_2 norm: (also call the Frobenius norm):

$$\|\mathbf{A}\|_2 = \sqrt{\sum_{m=0}^{M-1}\sum_{n=0}^{N-1}|A[m,n]|^2}$$

L_1 norm:

$$\|\mathbf{A}\|_1 = \sum_{m=0}^{M-1}\sum_{n=0}^{N-1}|A[m,n]|$$

L_∞ norm: (also call the max norm):

$$\|\mathbf{A}\|_\infty = \underset{m,n}{Max}|A[m,n]|$$

L_0 norm:

$$\lim_{\alpha\to 0}\left(\|\mathbf{A}\|_\alpha\right)^\alpha = K$$

where K is the number of points that satisfy $A[m, n] \neq 0$

[Example 2]

$$\mathbf{A} = \begin{bmatrix} -1 & 2 & 1 \\ 2 & 3 & -1 \\ 0 & -2 & 1 \end{bmatrix}$$

L_2 norm: $\sqrt{4(1^2)+3(2^2)+3^2} = 5$

L_1 norm: $4\cdot 1+3\cdot 2+3=13$

L_∞ norm: 3

L_0 norm: 8

[Norms of a Matrix: Definition 2]

Note: In fact, a more standard definition for the norm of a matrix is:

$$\|\mathbf{A}\|_\alpha = \sup_{\mathbf{x}\neq 0} \frac{\|\mathbf{A}\mathbf{x}\|_\alpha}{\|\mathbf{x}\|_\alpha}$$

$\sup_{x>0} 1 - e^{-x} = 1$

where sup (supremum) means the least upper bound.

上确界

The norm with this definition is called the operator norm. One possible application of the operator norm is to determine the passivity of electrical components.

For image processing and machine learning applications, it is more often to use the same definition of the norms of a vector to define the norms of a matrix, as on page 574.

The norms on page 574 are also called "Entry-wise" matrix norms or vector-based norms.

6.6 Markov Model

6.6.1 Definitions

State Vector:

$$\mathbf{x}(t) = \left[x_1(t), x_2(t), \cdots, x_N(t)\right]^T$$ (varies with t)

In a Markov model,

(1) $x_1(t)$, $x_2(t)$, …, $x_N(t)$ can be expressed as linear combinations of $\{x_1(t-1), x_2(t-1), \ldots, x_N(t-1)\}$

$$x_m(t) = a_{m,1}x_1(t-1) + a_{m,2}x_2(t-1) + \cdots + a_{m,N}x_N(t-1)$$

$m = 1, 2, \ldots, N$:

(2) $0 \le a_{m,n} \le 1$

(3) $\sum_{m=1}^{N} a_{m,n} = 1$ for $n = 1, 2, \ldots, N$:

Note that, if $\sum_{m=1}^{N} a_{m,n} = 1$

then

$$\sum_{m=1}^{N} x_m(t) = \sum_{m=1}^{N} \left(a_{m,1} x_1(t-1) + a_{m,2} x_2(t-1) + \cdots + a_{m,N} x_N(t-1) \right)$$

$$= \sum_{m=1}^{N} a_{m,1} x_1(t-1) + \sum_{m=1}^{N} a_{m,2} x_2(t-1) + \cdots + \sum_{m=1}^{N} a_{m,N} x_N(t-1)$$

$$\sum_{m=1}^{N} x_m(t) = \sum_{m=1}^{N} x_m(t-1)$$

Many problems in physics, environment, and social science can be expressed by the Markov model.

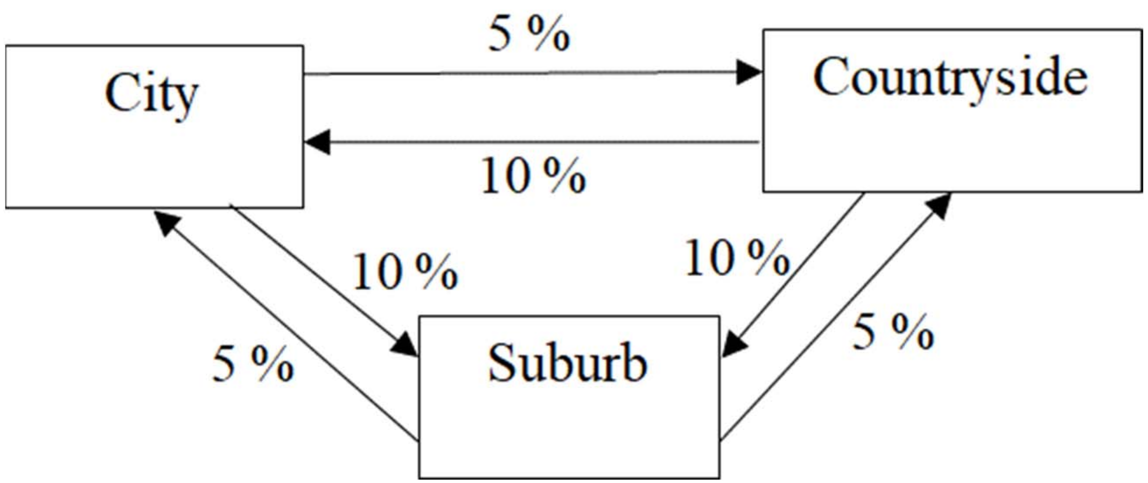
Markov Model (Matrix Form)

$$\mathbf{x}(t) = \mathbf{A}\mathbf{x}(t-1)$$

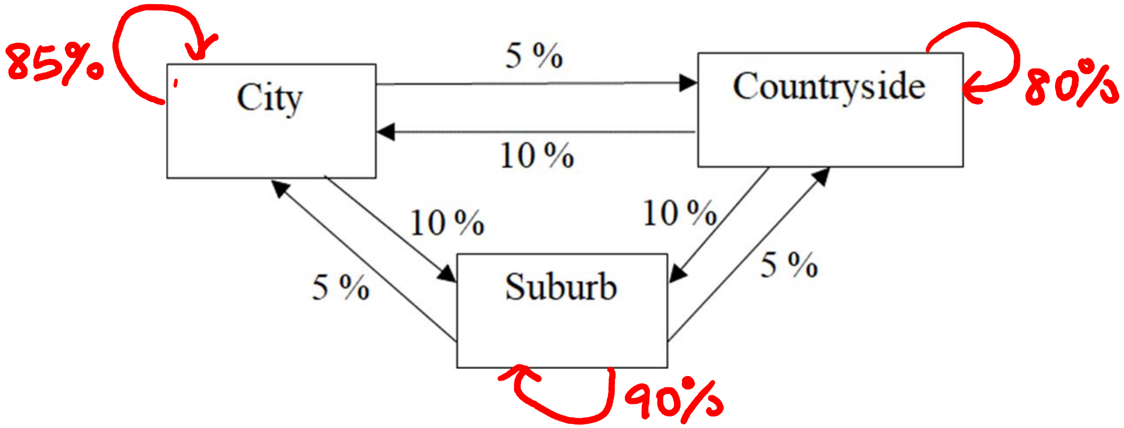
$$\mathbf{A} = \begin{bmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,N-1} & a_{1,N} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,N-1} & a_{2,N} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{N-1,1} & a_{N-1,2} & \cdots & a_{N-1,N-1} & a_{N-1,N} \\ a_{N,1} & a_{N,2} & \cdots & a_{N,N-1} & a_{N,N} \end{bmatrix}$$

$$\mathbf{x}(t) = \begin{bmatrix} x_1(t) \\ x_2(t) \\ \vdots \\ x_{N-1}(t) \\ x_N(t) \end{bmatrix} \qquad \mathbf{x}(t-1) = \begin{bmatrix} x_1(t-1) \\ x_2(t-1) \\ \vdots \\ x_{N-1}(t-1) \\ x_N(t-1) \end{bmatrix}$$

[Example 1] Migration Model



Suppose that, every year,
(1) 10% of the people lived in the city move to the suburb and 5% of the people lived in the city move to the countryside every year
(2) In the suburb, 5% of the people move to the city and 5% of the people move to the countryside every year.
(3) 10% of the people lived in the countryside move to the city and 10% of the people lived in the countryside move to the suburb.

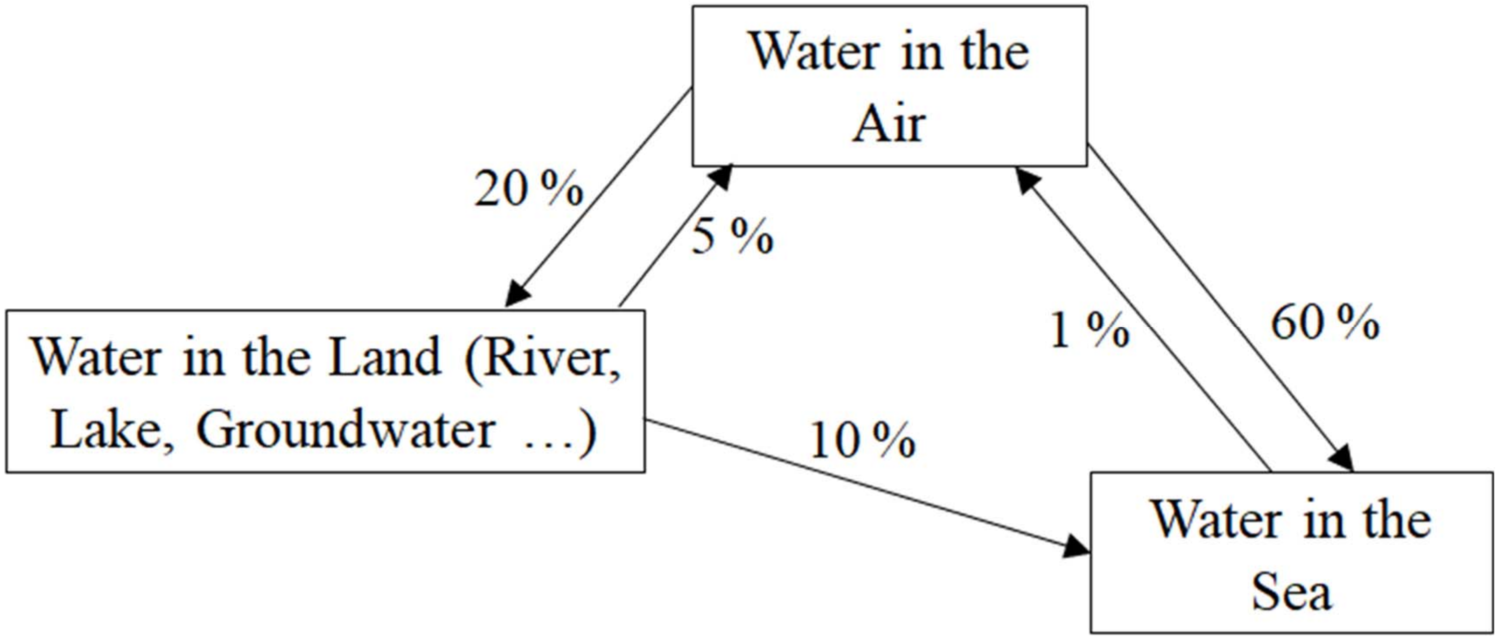


$$\mathbf{X}(t) = \mathbf{A}\mathbf{X}(t-1) \qquad \text{where}$$

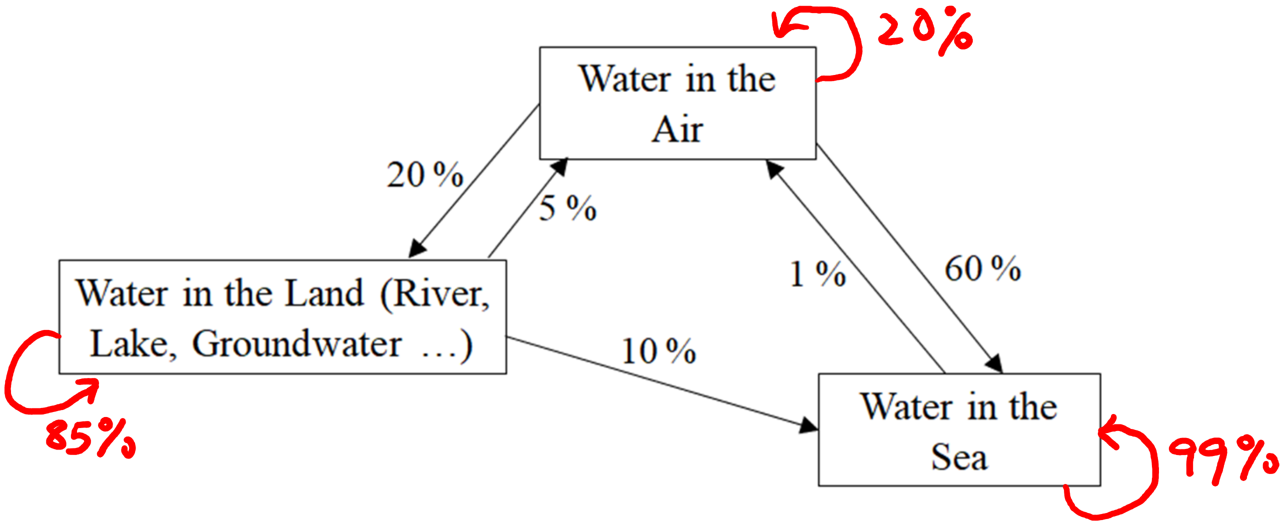
$$\mathbf{A} = \begin{bmatrix} 0.85 & 0.05 & 0.1 \\ 0.1 & 0.9 & 0.1 \\ 0.05 & 0.05 & 0.8 \end{bmatrix} \qquad \mathbf{X}(t) = \begin{array}{c} \begin{bmatrix} x_1(t) \\ x_2(t) \\ x_3(t) \end{bmatrix} \end{array} \begin{array}{l} \text{city} \\ \text{suburb} \\ \text{countryside} \end{array}$$

$$\begin{bmatrix} x_1(t) \\ x_2(t) \\ x_3(t) \end{bmatrix} = \begin{bmatrix} 0.85 & 0.05 & 0.1 \\ 0.1 & 0.9 & 0.1 \\ 0.05 & 0.05 & 0.8 \end{bmatrix} \begin{bmatrix} x_1(t-1) \\ x_2(t-1) \\ x_3(t-1) \end{bmatrix}$$

[Example 2] Water Cycle System



Suppose that, every year,
(1) 20% of the water in the air will precipitate to the land through rain or snow, 60% of the water in the air will precipitate to the sea.
(2) 10% of the water in the land will flow into the sea and 5% of the water will evaporate into the air.
(3) Also, 1% of the water in the see will evaporate to the air every year.



$$\mathbf{X}(t) = \mathbf{A}\mathbf{X}(t-1) \quad \text{where}$$

$$\mathbf{A} = \begin{bmatrix} 0.2 & 0.05 & 0.01 \\ 0.2 & 0.85 & 0 \\ 0.6 & 0.1 & 0.99 \end{bmatrix} \qquad \mathbf{X}(t) = \begin{bmatrix} x_1(t) \\ x_2(t) \\ x_3(t) \end{bmatrix} \begin{array}{l} \text{air} \\ \text{land} \\ \text{sea} \end{array}$$

6.6.2 Analysis

In Example 1, suppose that, initially, the populations of the city, the suburb, and the country are 50,000, 50,000, and 100,000.
(1) Determine the populations of the three places 1 year, 2 years, and 10 years later.
(2) Also, determine what will the populations of the three places converge to.

$t \to \infty$

$$\mathbf{A} = \begin{bmatrix} 0.85 & 0.05 & 0.1 \\ 0.1 & 0.9 & 0.1 \\ 0.05 & 0.05 & 0.8 \end{bmatrix}$$

Eigenvector-eigenvalue decomposition

$$\mathbf{A} = \mathbf{EDE}^{-1} \qquad \text{where}$$

$\lambda = 1$

$$\mathbf{E} = \begin{bmatrix} 3 & 1 & 1 \\ 5 & -1 & 0 \\ 2 & 0 & -1 \end{bmatrix} \qquad \mathbf{D} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0.8 & 0 \\ 0 & 0 & 0.75 \end{bmatrix} \qquad \mathbf{E}^{-1} = \begin{bmatrix} 0.1 & 0.1 & 0.1 \\ 0.5 & -0.5 & 0.5 \\ 0.2 & 0.2 & -0.8 \end{bmatrix}$$

Initial $$\mathbf{X}(0)=\begin{bmatrix}50000\\50000\\100000\end{bmatrix}$$

After one year $$\mathbf{X}(1)=\mathbf{AX}(0)=\begin{bmatrix}55000\\60000\\85000\end{bmatrix}$$

After two years $$\mathbf{X}(2)=\mathbf{AX}(1)=\begin{bmatrix}58250\\68000\\73750\end{bmatrix}$$

After ten years $$\mathbf{X}(10)=\mathbf{A}^{10}\mathbf{X}(0)=\mathbf{ED^{10}E^{-1}}\mathbf{X}(0)=\begin{bmatrix}61990\\94631\\43379\end{bmatrix}$$

$$\mathbf{D}^{10}=\begin{bmatrix}1 & 0 & 0\\0 & (0.8)^{10} & 0\\0 & 0 & (0.75)^{10}\end{bmatrix}$$

To determine what will the populations of the three places converge to, we can apply

$$\lim_{t\to\infty}\mathbf{X}(t)=\lim_{t\to\infty}\mathbf{A}^t\mathbf{X}(0)=\lim_{t\to\infty}\mathbf{E}\mathbf{D}^t\mathbf{E}^{-1}\mathbf{X}(0)$$

Since

$$\lim_{t\to\infty}\mathbf{D}^t=\begin{bmatrix}1^t & 0 & 0\\ 0 & (0.8)^t & 0\\ 0 & 0 & (0.75)^t\end{bmatrix}=\begin{bmatrix}1 & 0 & 0\\ 0 & 0 & 0\\ 0 & 0 & 0\end{bmatrix}$$

$$\lim_{t\to\infty}\mathbf{X}(t)=\begin{bmatrix}3 & 1 & 1\\ 5 & -1 & 0\\ 2 & 0 & -1\end{bmatrix}\begin{bmatrix}1 & 0 & 0\\ 0 & 0 & 0\\ 0 & 0 & 0\end{bmatrix}\begin{bmatrix}0.1 & 0.1 & 0.1\\ 0.5 & -0.5 & 0.5\\ 0.2 & 0.2 & -0.8\end{bmatrix}\begin{bmatrix}50000\\ 50000\\ 100000\end{bmatrix}$$

$$\begin{bmatrix}3 & 0 & 0\\ 5 & 0 & 0\\ 2 & 0 & 0\end{bmatrix}\quad\begin{bmatrix}0.3 & 0.3 & 0.3\\ 0.5 & 0.5 & 0.5\\ 0.2 & 0.2 & 0.2\end{bmatrix}$$

$$\lim_{t\to\infty}\mathbf{X}(t)=\begin{bmatrix}60000\\ 100000\\ 40000\end{bmatrix}$$

In Example 2, suppose that, initially, the amounts water in the air, in the land, and in the sea are $x_1(0)$, $x_2(0)$, and $x_3(0)$, respectively.

(1) Determine the amounts of water in the air, the land, and the sea after 10 years.

(2) Also determine what will the amounts of the water in the three places converge to.

$t \to \infty$

$$\mathbf{A} = \begin{bmatrix} 0.2 & 0.05 & 0.01 \\ 0.2 & 0.85 & 0 \\ 0.6 & 0.1 & 0.99 \end{bmatrix}$$

Eigenvector-eigenvalue decomposition $\mathbf{A} = \mathbf{EDE}^{-1}$

$$\begin{array}{cccc} & \lambda=1 & & \\ \mathbf{E} = & \left[\begin{array}{c} 3 \\ 4 \\ 220 \end{array}\right. & \begin{array}{c} 1 \\ 16.7977 \\ -17.9799 \end{array} & \left.\begin{array}{c} 3.3595 \\ -1 \\ -2.3595 \end{array}\right] \end{array} \quad \mathbf{D} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0.8619 & 0 \\ 0 & 0 & 0.1781 \end{bmatrix}$$

$$\mathbf{E}^{-1} = \begin{bmatrix} 1/227 & 1/227 & 1/227 \\ 0.01615 & 0.05723 & -0.00126 \\ 0.28892 & -0.02097 & -0.00356 \end{bmatrix}$$

After ten years

$$\mathbf{X}(10)=\mathbf{A}^{10}\mathbf{X}(0)=\mathbf{E}\mathbf{D}^{10}\mathbf{E}^{-1}\mathbf{X}(0)=\mathbf{E}\begin{bmatrix}1 & 0 & 0\\ 0 & 0.2263 & 0\\ 0 & 0 & 3.2098\cdot 10^{-8}\end{bmatrix}\mathbf{E}^{-1}\mathbf{X}(0)$$

$$\boxed{\begin{bmatrix}1 & 0 & 0\\ 0 & (0.8619)^{10} & 0\\ 0 & 0 & (0.1781)^{10}\end{bmatrix}}$$

$$\begin{bmatrix}x_1(10)\\ x_2(10)\\ x_3(10)\end{bmatrix}=\begin{bmatrix}0.01687x_1(0)+0.02617x_2(0)+0.01293x_3(0)\\ 0.07900x_1(0)+0.23515x_2(0)+0.01283x_3(0)\\ 0.90413x_1(0)+0.73869x_2(0)+0.97424x_3(0)\end{bmatrix}$$

To determine what will the amounts of the water in the three places converge to,

$$\lim_{t\to\infty}\mathbf{X}(t)=\lim_{t\to\infty}\mathbf{A}^t\mathbf{X}(0)=\lim_{t\to\infty}\mathbf{E}\mathbf{D}^t\mathbf{E}^{-1}\mathbf{X}(0)$$

Since

$$\lim_{t\to\infty}\mathbf{D}^t=\begin{bmatrix}1^t & 0 & 0\\ 0 & (0.8619)^t & 0\\ 0 & 0 & (0.1781)^t\end{bmatrix}=\begin{bmatrix}1&0&0\\0&0&0\\0&0&0\end{bmatrix}$$

$$\lim_{t\to\infty}\mathbf{X}(t)=\mathbf{E}\begin{bmatrix}1&0&0\\0&0&0\\0&0&0\end{bmatrix}\mathbf{E}^{-1}\mathbf{X}(0)$$

$$\lim_{t\to\infty}\begin{bmatrix}x_1(t)\\x_2(t)\\x_3(t)\end{bmatrix}=\begin{bmatrix}\frac{3}{227}\left(x_1(0)+x_2(0)+x_3(0)\right)\\ \frac{4}{227}\left(x_1(0)+x_2(0)+x_3(0)\right)\\ \frac{220}{227}\left(x_1(0)+x_2(0)+x_3(0)\right)\end{bmatrix}$$

[**Theorem 1**] For a Markov model matrix,
(i) At least one of the eigenvalue is $\lambda = 1$.
(ii) Other eigenvalues are no larger than 1.
(iii) Moreover, if the multiplicity of $\lambda = 1$ is 1, then the eigenvector corresponding to $\lambda = 1$ determines the ratio of $x_1(t) : x_2(t) : \ldots : x_N(t)$ in the convergence case ($t \to \infty$).

(Proof): Suppose that **A** is the transfer matrix of a Markov model.

(i) To show 1 must be an eigenvalue of **A**:

If $\mathbf{A}_1 = \mathbf{L}\mathbf{A}\mathbf{L}^{-1}$ where

$\det(A_1) = \cancel{\det(L)}\det(A)\frac{1}{\cancel{\det(L)}}$

If $Ae = \lambda e$
$A_1(Le)$
$= LA\cancel{L^{-1}}\cancel{L}e$
$= \lambda(Le)$

$$\mathbf{L} = \begin{bmatrix} 1 & 1 & \cdots & 1 & 1 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \\ 0 & 0 & \cdots & 0 & 1 \end{bmatrix} \qquad \mathbf{L}^{-1} = \begin{bmatrix} 1 & -1 & \cdots & -1 & -1 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \\ 0 & 0 & \cdots & 0 & 1 \end{bmatrix}$$

$$\mathbf{LA} = \begin{bmatrix} 1 & 1 & \cdots & 1 & 1 \\ a_{2,1} & a_{2,2} & \cdots & a_{2,N-1} & a_{2,N} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{N-1,1} & a_{N-1,2} & \cdots & a_{N-1,N-1} & a_{N-1,N} \\ a_{N,1} & a_{N,2} & \cdots & a_{N,N-1} & a_{N,N} \end{bmatrix}$$

$$\mathbf{A_1} = \mathbf{LAL}^{-1} = \begin{bmatrix} 1 & 0 & \cdots & 0 & 0 \\ a_{2,1} & a_{2,2}-a_{2,1} & \cdots & a_{2,N-1}-a_{2,1} & a_{2,N}-a_{2,1} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{N-1,1} & a_{N-1,2}-a_{N-1,1} & \cdots & a_{N-1,N-1}-a_{N-1,1} & a_{N-1,N}-a_{N-1,1} \\ a_{N,1} & a_{N,2}-a_{N,1} & \cdots & a_{N,N-1}-a_{N,1} & a_{N,N}-a_{N,1} \end{bmatrix}$$

$$\det\left(\mathbf{A_1} - \lambda\mathbf{I}\right) = (1-\lambda)\det\left(\begin{bmatrix} a_{2,2}-a_{2,1}-\lambda & \cdots & a_{2,N-1}-a_{2,1} & a_{2,N}-a_{2,1} \\ \vdots & \ddots & \vdots & \vdots \\ a_{N-1,2}-a_{N-1,1} & \cdots & a_{N-1,N-1}-a_{N-1,1}-\lambda & a_{N-1,N}-a_{N-1,1} \\ a_{N,2}-a_{N,1} & \cdots & a_{N,N-1}-a_{N,1} & a_{N,N}-a_{N,1}-\lambda \end{bmatrix}\right)$$

Therefore, $\det\left(\mathbf{A_1} - \lambda\mathbf{I}\right) = 0 \quad \textit{for } \lambda = 1$

$\lambda = 1$ should be one of the eigenvalues of $\mathbf{A_1}$ and hence $\mathbf{A}$ (note that $\mathbf{A}$ and $\mathbf{A_1}$ are similar) .

(ii) To show 1 is the largest eigenvalue of $\mathbf{A}$:

If

$$\mathbf{A}_2 = \mathbf{A}^{\mathbf{T}}$$

then $\mathbf{A_2}$ and $\mathbf{A}$ have the same eigenvalues since

$$\det(\mathbf{A} - \lambda\mathbf{I}) = \det(\mathbf{A_2} - \lambda\mathbf{I}).$$

If $\mathbf{g} = [g_1, g_2, \ldots, g_N]^T$ is an eigenvector of $\mathbf{A_2}$ and the corresponding eigenvalue is λ. Without the loss of generalization, we suppose that

$$|g_p| \geq |g_n| \text{ where } n = 1, 2, \ldots, N.$$

inner product of the pth column of A and g

Compare the p^{th} entry on the both sides of $\mathbf{A_2 g} = \lambda\mathbf{g}$, we have

$$a_{1,p}g_1 + a_{2,p}g_2 + \cdots + a_{N,p}g_N = \lambda g_p$$

$$a_{1,p}g_1 + a_{2,p}g_2 + \cdots + a_{N,p}g_N = \lambda g_p$$

$$\left|\lambda g_p\right| = \left|a_{1,p}g_1 + a_{2,p}g_2 + \cdots + a_{N,p}g_N\right| \le a_{1,p}\left|g_1\right| + a_{2,p}\left|g_2\right| + \cdots + a_{N,p}\left|g_N\right|$$

$$\le (a_{1,p} + a_{2,p} + \cdots + a_{N,p})\left|g_p\right| = \left|g_p\right|$$

$$\left|\lambda\right| \le 1$$

(iii) To show that, if the multiplicity of $\lambda = 1$ is 1, then its corresponding eigenvector determines the ratio of $x_1(t)$: $x_2(t)$: … : $x_N(t)$ in the convergence case, suppose that

$$\mathbf{E} = \begin{bmatrix} \mathbf{e_1} & \mathbf{e_2} & \cdots & \mathbf{e_N} \end{bmatrix}, \quad \mathbf{E}^{-1} = \begin{bmatrix} \mathbf{f_1} & \mathbf{f_2} & \cdots & \mathbf{f_N} \end{bmatrix}^T$$

then

$$D : \begin{bmatrix} \lambda_1 & & 0 \\ 0 & \lambda_2 & \\ & & \ddots \\ & & & \lambda_N \end{bmatrix}$$

$$\mathbf{A} = \mathbf{EDE^{-1}}$$

$$\mathbf{A} = \lambda_1 \mathbf{e_1 f_1^T} + \lambda_2 \mathbf{e_2 f_2^T} + \cdots + \lambda_{N-1} \mathbf{e_{N-1} f_{N-1}^T} + \lambda_N \mathbf{e_N f_N^T}$$

$$\mathbf{A} = \lambda_1 \mathbf{e_1 f_1^T} + \lambda_2 \mathbf{e_2 f_2^T} + \cdots + \lambda_{N-1} \mathbf{e_{N-1} f_{N-1}^T} + \lambda_N \mathbf{e_N f_N^T}$$

$$\lim_{t\to\infty} \mathbf{A}^t = \lim_{t\to\infty} \lambda_1^t \mathbf{e_1 f_1^T} + \lambda_2^t \mathbf{e_2 f_2^T} + \cdots + \lambda_{N-1}^t \mathbf{e_{N-1} f_{N-1}^T} + \lambda_N^t \mathbf{e_N f_N^T}$$

If $\lambda_1 = 1$ and $|\lambda_m| < 1$ for $m \neq 1$, then

$$\lim_{t\to\infty} \mathbf{A}^t = \lim_{t\to\infty} \mathbf{e_1 f_1^T}$$

$$\lim_{t\to\infty} \mathbf{X}(t) = \lim_{t\to\infty} \mathbf{A}^t \mathbf{X}(0) = c\mathbf{e_1} \qquad \text{where} \qquad c = \mathbf{f_1^T X}(0)$$

Note that c is a constant. In other words, when $t \to \infty$,

$$x_1(t) : x_2(t) : \cdots : x_N(t) = e_1[1] : e_1[2] : \cdots : e_1[N]$$

6.7 Discrete Transforms

(只教不考)

If the rows of A are orthogonal

$A A^H = D$

(1) Discrete Fourier Transform (DFT)

$$G[m] = \sum_{n=0}^{N-1} g[n] e^{-j\frac{2\pi mn}{N}}$$

$n = 0, 1, 2, \ldots, N\text{-}1$

$m = 0, 1, 2, \ldots, N\text{-}1$

inverse:

$$g[n] = \frac{1}{N}\sum_{m=0}^{N-1} G[m] e^{j\frac{2\pi mn}{N}}$$

$A A^H D^{-1} = I$

$A^{-1} = A^H D^{-1}$

When $N = 2$

$$\begin{bmatrix} G[0] \\ G[1] \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} g[0] \\ g[1] \end{bmatrix}$$

When $N = 3$

$$\begin{bmatrix} G[0] \\ G[1] \\ G[2] \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & \frac{-1-j\sqrt{3}}{2} & \frac{-1+j\sqrt{3}}{2} \\ 1 & \frac{-1+j\sqrt{3}}{2} & \frac{-1-j\sqrt{3}}{2} \end{bmatrix} \begin{bmatrix} g[0] \\ g[1] \\ g[2] \end{bmatrix}$$

(2) Discrete Cosine Transform (DCT)

$$\mathbf{y} = \mathbf{C_N}\mathbf{x}$$

where

$$C_N[m,n] = k_m \cos\left(\pi \frac{m(n+1/2)}{N}\right)$$

$m = 0, 1, \ldots, N\text{-}1,\qquad n = 0, 1, \ldots, N\text{-}1,$

$$k_0 = \sqrt{1/N}$$

$$k_m = \sqrt{2/N} \qquad \text{when } m \neq 0$$

Inverse:

$$\mathbf{x} = \mathbf{C_N^T}\mathbf{y} \qquad \mathbf{C_N^{-1}} = \mathbf{C_N^T}$$

Application: Data Compression

When $N = 8$

8-Point DCT

$$\mathbf{C_8} = \begin{bmatrix} 0.3536 & 0.3536 & 0.3536 & 0.3536 & 0.3536 & 0.3536 & 0.3536 & 0.3536 \\ 0.4904 & 0.4157 & 0.2778 & 0.0975 & -0.0975 & -0.2778 & -0.4157 & -0.4904 \\ 0.4619 & 0.1913 & -0.1913 & -0.4619 & -0.4619 & -0.1913 & 0.1913 & 0.4619 \\ 0.4157 & -0.0975 & -0.4904 & -0.2778 & 0.2778 & 0.4904 & 0.0975 & -0.4157 \\ 0.3536 & -0.3536 & -0.3536 & 0.3536 & 0.3536 & -0.3536 & -0.3536 & 0.3536 \\ 0.2778 & -0.4904 & 0.0975 & 0.4157 & -0.4157 & -0.0975 & 0.4904 & -0.2778 \\ 0.1913 & -0.4619 & 0.4619 & -0.1913 & -0.1913 & 0.4619 & -0.4619 & 0.1913 \\ 0.0975 & -0.2778 & 0.4157 & -0.4904 & 0.4904 & -0.4157 & 0.2778 & -0.0975 \end{bmatrix}$$

Zero Crossings of the 8-Point DCT

$$
\begin{bmatrix}
0.3536 & 0.3536 & 0.3536 & 0.3536 & 0.3536 & 0.3536 & 0.3536 & 0.3536 \\
0.4904 & 0.4157 & 0.2778 & 0.0975 & * -0.0975 & -0.2778 & -0.4157 & -0.4904 \\
0.4619 & 0.1913 & * -0.1913 & -0.4619 & -0.4619 & -0.1913 & * 0.1913 & 0.4619 \\
0.4157 & * -0.0975 & -0.4904 & -0.2778 & * 0.2778 & 0.4904 & 0.0975 & * -0.4157 \\
0.3536 & * -0.3536 & -0.3536 & * 0.3536 & 0.3536 & * -0.3536 & -0.3536 & * 0.3536 \\
0.2778 & * -0.4904 & * 0.0975 & 0.4157 & * -0.4157 & -0.0975 & * 0.4904 & * -0.2778 \\
0.1913 & * -0.4619 & * 0.4619 & * -0.1913 & -0.1913 & * 0.4619 & * -0.4619 & * 0.1913 \\
0.0975 & * -0.2778 & * 0.4157 & * -0.4904 & * 0.4904 & * -0.4157 & * 0.2778 & * -0.0975
\end{bmatrix}
$$

more zero crossings $\rightarrow$ high frequency component

(3) Walsh Transform (Hadamard Transform)

$$\mathbf{y} = \mathbf{W_N x}$$

where $\mathbf{W_N}$ is an $N \times N$ matrix and the entries of $\mathbf{W_N}$ is either 1 or -1, N is limited to 2^k.

Inverse: $$(\mathbf{W_N})^{-1} = \frac{1}{N}\mathbf{W_N^T}$$

Applications:

(i) simplification for computation

(ii) analysis step-like functions

(iii) code division multiple access (CDMA)

2-point Walsh transform

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

4-point Walsh transform

$$\mathbf{W}_4 = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \\ 1 & -1 & 1 & -1 \end{bmatrix}$$

To obtain the 2^{k+1}-point Walsh transform from the 2^k-point Walsh transform,

Step 1

$$\mathbf{V}_{2^{k+1}} = \begin{bmatrix} \mathbf{W}_{2^k} & \mathbf{W}_{2^k} \\ \mathbf{W}_{2^k} & -\mathbf{W}_{2^k} \end{bmatrix}$$

Step 2 Reorder according to sign changes

$$\mathbf{V}_{2^{k+1}} \xrightarrow{\textit{permutation}} \mathbf{W}_{2^{k+1}}$$

$$\mathbf{W}_2 = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \qquad \mathbf{V}_4 = \begin{bmatrix} \mathbf{W}_2 & \mathbf{W}_2 \\ \mathbf{W}_2 & -\mathbf{W}_2 \end{bmatrix} = \begin{array}{c} \\ \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \end{bmatrix} \end{array} \begin{array}{c} \text{sign changes} \\ 0 \\ 3 \\ 1 \\ 2 \end{array}$$

$$
\mathbf{W}_4 = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \\ 1 & -1 & 1 & -1 \end{bmatrix}
\quad
\mathbf{V}_8 = \begin{array}{cc}
 & \text{sign changes} \\
\begin{bmatrix}
1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\
1 & 1 & -1 & -1 & 1 & 1 & -1 & -1 \\
1 & -1 & -1 & 1 & 1 & -1 & -1 & 1 \\
1 & -1 & 1 & -1 & 1 & -1 & 1 & -1 \\
1 & 1 & 1 & 1 & -1 & -1 & -1 & -1 \\
1 & 1 & -1 & -1 & -1 & -1 & 1 & 1 \\
1 & -1 & -1 & 1 & -1 & 1 & 1 & -1 \\
1 & -1 & 1 & -1 & -1 & 1 & -1 & 1
\end{bmatrix}
&
\begin{array}{c} 0 \\ 3 \\ 4 \\ 7 \\ 1 \\ 2 \\ 5 \\ 6 \end{array}
\end{array}
$$

$$
\mathbf{W}_8 = \begin{bmatrix}
1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\
1 & 1 & 1 & 1 & -1 & -1 & -1 & -1 \\
1 & 1 & -1 & -1 & -1 & -1 & 1 & 1 \\
1 & 1 & -1 & -1 & 1 & 1 & -1 & -1 \\
1 & -1 & -1 & 1 & 1 & -1 & -1 & 1 \\
1 & -1 & -1 & 1 & -1 & 1 & 1 & -1 \\
1 & -1 & 1 & -1 & -1 & 1 & -1 & 1 \\
1 & -1 & 1 & -1 & 1 & -1 & 1 & -1
\end{bmatrix}
$$

(4) Haar Transform

$$\mathbf{y} = \mathbf{H_N}\mathbf{x}$$

where $\mathbf{H_N}$ is an $N\times N$ matrix and the entries of $\mathbf{H_N}$ is 1 0, or -1, N is limited to 2^k.

Inverse: $(\mathbf{H_N})^{-1} = \mathbf{D_N}\mathbf{H_N^T}$ $\mathbf{D_N}$ is diagonal and

$$\mathbf{D_N}[0,0] = \mathbf{D_N}[1,1] = 1/N$$

$$\mathbf{D_N}[2,2] = \mathbf{D_N}[3,3] = 2/N$$

$$\mathbf{D_N}[4,4] = \mathbf{D_N}[5,5] = \mathbf{D_N}[6,6] = \mathbf{D_N}[7,7] = 4/N$$

$$\mathbf{D_N}[n,n] = 8/N \quad \textit{for}\ 8 \le n \le 15$$

$$\vdots \qquad\qquad \vdots \qquad\qquad \vdots$$

Applications:

(i) simplification for computation

(ii) extract local features

$N = 2$

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

$N = 4$

$$\mathbf{H}_4 = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & -1 \end{bmatrix}$$

$N = 8$

$$\mathbf{H}_8 = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & -1 & -1 & -1 & -1 \\ 1 & 1 & -1 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & -1 & -1 \\ 1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & -1 \end{bmatrix}$$

General way to generate the Haar transform:

$$\mathbf{H}_{\mathbf{2N}} = \begin{bmatrix} \mathbf{H}_{\mathbf{N}} \otimes [1,1] \\ \mathbf{I}_{\mathbf{N}} \otimes [1,-1] \end{bmatrix}$$ where $\otimes$ means the Kronecker product

$$\mathbf{I}_{\mathbf{N}} = \begin{bmatrix} 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \\ 0 & 0 & \cdots & 0 & 1 \end{bmatrix}$$

$N = 16$

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