

Fall 2007

# 線性系統 Linear Systems

## Chapter 03 Linear Algebra

Feng-Li Lian  
NTU-EE  
Sep07 – Jan08

Materials used in these lecture notes are adopted from  
“Linear System Theory & Design,” 3rd. Ed., by C.-T. Chen (1999)

### Outline

- Introduction
- Basis, Representation, & Orthonormalization (3.2)
- Linear Algebraic Equations (3.3)
- Similarity Transformation (3.4)
- Diagonal Form and Jordan Form (3.5)
- Functions of a Square Matrix (3.6)
- Lyapunov Equation (3.7)
- Some Useful Formulas (3.8)
- Quadratic Form and Positive Definiteness (3.9)
- Singular-Value Decomposition (3.10)
- Norms of Matrices (3.11)

- Matrix partition and block matrix

$$\mathbf{A} = [\mathbf{A}_{ij}] = \begin{bmatrix} & & \\ & \text{---} & \text{---} \\ & | & | \\ & | & | \end{bmatrix}$$

$$\mathbf{B} = [\mathbf{B}_{ij}] = \begin{bmatrix} & & & \\ & \text{---} & \text{---} & \text{---} \\ & | & | & | \\ & | & | & | \end{bmatrix}$$

- Matrix multiplication with compatible partitions

$$\mathbf{C} = \mathbf{A} \mathbf{B} \quad [\mathbf{C}_{ik}] = \left[ \sum_j \mathbf{A}_{ij} \mathbf{B}_{jk} \right]$$

$$= \begin{bmatrix} & & & \\ & \text{---} & \text{---} & \text{---} \\ & | & | & | \\ & | & | & | \end{bmatrix}$$

Compatible Partitions:

column partition of the 1st matrix = row partition of the 2nd matrix

$$\mathbf{AB} = [\mathbf{a}_1 \ \mathbf{a}_2 \ \cdots \ \mathbf{a}_m] \begin{bmatrix} \mathbf{b}_1 \\ \mathbf{b}_2 \\ \vdots \\ \mathbf{b}_m \end{bmatrix} = \mathbf{a}_1\mathbf{b}_1 + \mathbf{a}_2\mathbf{b}_2 + \cdots + \mathbf{a}_m\mathbf{b}_m$$

Linear combination of  $\mathbf{A}$ 's columns

$$\mathbf{CA} = \mathbf{C}[\mathbf{a}_1 \ \mathbf{a}_2 \ \cdots \ \mathbf{a}_m] = [\mathbf{Ca}_1 \ \mathbf{Ca}_2 \ \cdots \ \mathbf{Ca}_m]$$

$$\mathbf{BD} = \begin{bmatrix} \mathbf{b}_1 \\ \mathbf{b}_2 \\ \vdots \\ \mathbf{b}_m \end{bmatrix} \mathbf{D} = \begin{bmatrix} \mathbf{b}_1\mathbf{D} \\ \mathbf{b}_2\mathbf{D} \\ \vdots \\ \mathbf{b}_m\mathbf{D} \end{bmatrix}$$

### Linear Dependence (L.D.) vs. Linear Independence (L.I.) – 1 (3.2)

- A set  $\{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_m\}$  in  $\mathbb{R}^n$  is L.D. (linearly dependent)

IF  $\exists \alpha = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_m \end{bmatrix} \neq \mathbf{0}$ , such that  $[\mathbf{x}_1 \ \mathbf{x}_2 \ \cdots \ \mathbf{x}_m] \cdot \alpha = \mathbf{0}$

$\iff$  at least one  $\mathbf{x}_i$  is a linear combination of other vectors

- A set of vectors in  $\mathbb{R}^n$  is a basis
  - if 1) The set of vectors is L.I. (linearly independent)
  - 2) All vectors in  $\mathbb{R}^n$  can be linearly combined by those in the set
- Every basis in  $\mathbb{R}^n$  has  $n$  vectors. Thus  $\dim(\mathbb{R}^n) = n$
- Let  $\{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n\}$  be a basis and  $\mathbf{x}$  be a vector in  $\mathbb{R}^n$ , then
 
$$\mathbf{x} = [\mathbf{x}_1 \mathbf{x}_2 \cdots \mathbf{x}_n] \boldsymbol{\alpha}$$

unique representation  
of  $\mathbf{x}$  w.r.t. the basis
- Any set of more than  $n$  vectors in  $\mathbb{R}^n$  is L.D.

- Standard orthonormal basis

$$\mathbf{i}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 0 \end{bmatrix}, \quad \mathbf{i}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \\ 0 \end{bmatrix}, \quad \dots, \quad \mathbf{i}_{n-1} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 1 \\ 0 \end{bmatrix}, \quad \mathbf{i}_n = \begin{bmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix}$$

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} =$$

- Norms of vectors: real-valued function  $\|\mathbf{x}\|$

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \quad \|\bullet\| :$$

- satisfying

1.  $\|\mathbf{x}\| \geq 0$  for every  $\mathbf{x}$  and  $\|\mathbf{x}\| = 0$  if and only if  $\mathbf{x} = \mathbf{0}$ .

2.  $\|\alpha\mathbf{x}\| = |\alpha|\|\mathbf{x}\|$ , for any real  $\alpha$ .

3.  $\|\mathbf{x}_1 + \mathbf{x}_2\| \leq \|\mathbf{x}_1\| + \|\mathbf{x}_2\|$  for every  $\mathbf{x}_1$  and  $\mathbf{x}_2$ . (triangle inequality)

- Common norm: 1-norm, 2-norm (Euclidean),  $\infty$ -norm, p-norm

$$\|\mathbf{x}\|_1 := |x_1| + |x_2| + \cdots + |x_n|$$

$$:= \sum_{i=1}^n |x_i|$$

$$\|\mathbf{x}\|_2 := \sqrt{|x_1|^2 + |x_2|^2 + \cdots + |x_n|^2}$$

$$:= \sqrt{\mathbf{x}^\top \mathbf{x}} := \left( \sum_{i=1}^n |x_i|^2 \right)^{1/2}$$

$$\|\mathbf{x}\|_\infty := \max \{ |x_1|, |x_2|, \dots, |x_n| \}$$

$$:= \max_i |x_i|$$

$$\|\mathbf{x}\| := \left( \sum_{i=1}^n |x_i| \right)^{1/p}$$

- A vector  $\mathbf{x}$  is **normalized** if
- Two vectors  $\mathbf{x}_i, \mathbf{x}_j$  are **orthogonal** if
- A set of vectors  $\{\mathbf{x}_1, \dots, \mathbf{x}_m\}$  is **orthonormal** if
- Given a set of **L.I.** vectors  $\{\mathbf{e}_1, \dots, \mathbf{e}_m\}$ ,  
use Schmidt orthonormalization procedure  
to obtain an **orthonormal** set

- Schmidt orthonormalization procedure for  $\{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_m\}$ :

$$\mathbf{u}_1 := \mathbf{e}_1$$

$$\mathbf{u}_2 := \mathbf{e}_2 - (\mathbf{q}_1^\top \mathbf{e}_2) \mathbf{q}_1$$

$$\mathbf{u}_3 := \mathbf{e}_3 - (\mathbf{q}_1^\top \mathbf{e}_3) \mathbf{q}_1 - (\mathbf{q}_2^\top \mathbf{e}_3) \mathbf{q}_2$$

⋮

$$\mathbf{u}_m := \mathbf{e}_m - \sum_{k=1}^{m-1} (\mathbf{q}_k^\top \mathbf{e}_m) \mathbf{q}_k$$

$$\mathbf{q}_1 := \frac{\mathbf{u}_1}{\|\mathbf{u}_1\|}$$

$$\mathbf{q}_2 := \frac{\mathbf{u}_2}{\|\mathbf{u}_2\|}$$

$$\mathbf{q}_3 := \frac{\mathbf{u}_3}{\|\mathbf{u}_3\|}$$

⋮

$$\mathbf{q}_m := \frac{\mathbf{u}_m}{\|\mathbf{u}_m\|}$$

- Representation by Orthonormal vectors  $\{\mathbf{q}_1, \mathbf{q}_2, \dots, \mathbf{q}_m\}$ :

Let  $\{\mathbf{q}_1, \mathbf{q}_2, \dots, \mathbf{q}_n\}$  be an orthonormal basis

$$\mathbf{Q} = [\mathbf{q}_1 \ \mathbf{q}_2 \ \cdots \ \mathbf{q}_m] \rightarrow \mathbf{Q}'\mathbf{Q} = \mathbf{I}_m$$

$$\rightarrow \mathbf{Q}' = \mathbf{Q}^{-1}$$

and  $\mathbf{x} = [\mathbf{q}_1 \ \mathbf{q}_2 \ \cdots \ \mathbf{q}_n] \ \boldsymbol{\alpha}$ ,

then  $\boldsymbol{\alpha} = [\mathbf{q}_1 \ \mathbf{q}_2 \ \cdots \ \mathbf{q}_n]' \mathbf{x}$

### Linear Algebraic Equations – 1 (3.3)

$$\mathbf{A} \mathbf{x} = \mathbf{y}, \quad \mathbf{A}: m \times n, \quad \mathbf{x}: n \times 1, \quad \mathbf{y}: m \times 1$$

- range space of  $\mathbf{A} := \{\beta \mid \beta = \mathbf{A}\alpha\}$
- $\text{rank}(\mathbf{A}) := \rho(\mathbf{A}) := \text{dim. of } \mathbf{A}'\text{s range space} \leq \min\{m, n\}$
- null space of  $\mathbf{A} := \{\alpha \mid \mathbf{A}\alpha = \mathbf{0}\}$
- $\text{nullity}(\mathbf{A}) := \text{dim. of } \mathbf{A}'\text{s null space} = n - \rho(\mathbf{A})$

- $\mathbf{Ax} = \mathbf{y}$  has at least a solution

$$\Leftrightarrow \rho(\mathbf{A}) = \rho([\mathbf{A} \ \mathbf{y}])$$

- $\mathbf{Ax} = \mathbf{y}$  has at least a solution for every  $\mathbf{y}$   $\Leftrightarrow \rho(\mathbf{A}) = m$

### Theorem 3.1

- Given an  $m \times n$  matrix  $\mathbf{A}$ , & an  $m \times 1$  vector  $\mathbf{y}$ ,  
an  $n \times 1$  solution  $\mathbf{x}$  exists in  $\mathbf{Ax} = \mathbf{y}$   
if and only if  
 $\mathbf{y}$  lies in the **range space** of  $\mathbf{A}$ ,  
or, equivalently,

$$\rho(\mathbf{A}) = \rho([\mathbf{A} \ \mathbf{y}])$$

where  $[\mathbf{A} \ \mathbf{y}]$  is an  $m \times (n + 1)$  matrix  
with  $\mathbf{y}$  appended to  $\mathbf{A}$  as an additional column.

- Given an  $m \times n$  matrix  $\mathbf{A}$ ,  
a solution  $\mathbf{x}$  exists in  $\mathbf{Ax} = \mathbf{y}$ , for every  $\mathbf{y}$ ,  
if and only if  
 $\mathbf{A}$  has **rank  $m$**  (**full row rank**).

- Suppose

$$\text{nullity}(\mathbf{A}) = k,$$

$\{\mathbf{n}_1, \mathbf{n}_2, \dots, \mathbf{n}_k\}$  is a basis of the null space of  $\mathbf{A}$ ,

and  $\mathbf{Ax} = \mathbf{y}$  has a solution  $\mathbf{x}_p$

- Then

every vector  $\mathbf{x} = \mathbf{x}_p + \alpha_1 \mathbf{n}_1 + \alpha_2 \mathbf{n}_2 + \dots + \alpha_k \mathbf{n}_k$  is also a solution

### Theorem 3.2

- Given an  $m \times n$  matrix  $\mathbf{A}$ , & an  $m \times 1$  vector  $\mathbf{y}$ ,  
let  $\mathbf{x}_p$  be a solution of  $\mathbf{Ax} = \mathbf{y}$   
and let  $k = n - \rho(\mathbf{A})$  be the nullity of  $\mathbf{A}$ .

If  $\mathbf{A}$  has rank  $n$  (full column rank) or  $k = 0$ ,  
then the solution  $\mathbf{x}_p$  is unique.

If  $k > 0$ , then for every real  $\alpha_i, i = 1, 2, \dots, k$ ,  
the vector

$$\mathbf{x} = \mathbf{x}_p + \alpha_1 \mathbf{n}_1 + \alpha_2 \mathbf{n}_2 + \dots + \alpha_k \mathbf{n}_k$$

is a solution of  $\mathbf{Ax} = \mathbf{y}$ , where the set  $\{\mathbf{n}_1, \dots, \mathbf{n}_k\}$   
is a basis of the null space of  $\mathbf{A}$ .

- Given an  $m \times n$  matrix  $\mathbf{A}$ ,  
a solution  $\mathbf{x}$  exists in  $\mathbf{x}\mathbf{A} = \mathbf{y}$ , for any  $\mathbf{y}$ ,  
if and only if  
 $\mathbf{A}$  has full column rank.
- Given an  $m \times n$  matrix  $\mathbf{A}$ , & an  $1 \times n$  vector  $\mathbf{y}$ ,  
let  $\mathbf{x}_p$  be a solution of  $\mathbf{x}\mathbf{A} = \mathbf{y}$   
and let  $k = m - \rho(\mathbf{A})$ .

If  $k = 0$ , the solution  $\mathbf{x}_p$  is unique.

If  $k > 0$ , then for any  $\alpha_i, i = 1, 2, \dots, k$ ,  
the vector

$$\mathbf{x} = \mathbf{x}_p + \alpha_1 \mathbf{n}_1 + \alpha_2 \mathbf{n}_2 + \dots + \alpha_k \mathbf{n}_k$$

is a solution of  $\mathbf{x}\mathbf{A} = \mathbf{y}$ , where  $\mathbf{n}_i \mathbf{A} = \mathbf{0}$   
and the set  $\{\mathbf{n}_1, \dots, \mathbf{n}_k\}$  is linearly independent.

## Determinant – 1

- Determinant of  $\mathbf{A} = [a_{ij}] \in \mathbb{R}^{n \times n}$

$$\det \mathbf{A} = \sum_i^n a_{ij} c_{ij}$$

$$\text{cofactor } c_{ij} = (-1)^{i+j} \det \mathbf{M}_{ij}$$

$$\mathbf{M}_{ij} = (n-1) \times (n-1) \text{ submatrix of } \mathbf{A}$$

by deleting its  $i$ th row and  $j$ th column

- A minor of order  $r$  = the determinant of any  $r \times r$  submatrix of  $\mathbf{A}$
- Rank:  $\rho(\mathbf{A})$  = the largest order of all nonzero minors of  $\mathbf{A}$

- $\mathbf{A}$  is nonsingular  $\Leftrightarrow \det \mathbf{A} \neq 0 \Leftrightarrow \rho(\mathbf{A}) = n$

- If  $\mathbf{A}$  is nonsingular, then the inverse of  $\mathbf{A}$  is

$$\mathbf{A}^{-1} = \frac{\text{Adj } \mathbf{A}}{\det \mathbf{A}} = \frac{1}{\det \mathbf{A}} [c_{ij}]'$$

and satisfies  $\mathbf{A}\mathbf{A}^{-1} = \mathbf{A}^{-1}\mathbf{A} = \mathbf{I}_n$

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

- Consider  $\mathbf{Ax} = \mathbf{y}$  with  $\mathbf{A}$  square
- If  $\mathbf{A}$  is nonsingular,  
then the equation has a unique solution  
for every  $\mathbf{y}$   
and the solution equals  $\mathbf{A}^{-1}\mathbf{y}$ .

In particular,  
the only solution of  $\mathbf{Ax} = \mathbf{0}$  is  $\mathbf{x} = \mathbf{0}$

- $\mathbf{Ax} = \mathbf{0}$  has nonzero solutions  
if and only if  
 $\mathbf{A}$  is singular.

The number of linearly independent solutions  
equals the nullity of  $\mathbf{A}$

### Similarity Transformation – 1 (3.4)

- Linear transform w.r.t. different bases

$$\mathbf{Ax} = \mathbf{y}, \quad \mathbf{A}: n \times n, \quad \mathbf{x}: n \times 1, \quad \mathbf{y}: n \times 1$$

-- a linear transform from  $\mathbb{R}^n$  to  $\mathbb{R}^n$

- If  $\mathbf{x}$  &  $\mathbf{y}$  are represented w.r.t. a new basis  $\{\mathbf{q}_1, \mathbf{q}_2, \dots, \mathbf{q}_n\}$  as

$$\mathbf{x} = \mathbf{Q} \bar{\mathbf{x}}$$

$$\mathbf{y} = \mathbf{Q} \bar{\mathbf{y}}$$

where  $\mathbf{Q} = [\mathbf{q}_1 \ \mathbf{q}_2 \ \cdots \ \mathbf{q}_n]$ ,

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

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

then  $\mathbf{A} \mathbf{Q} \bar{\mathbf{x}} = \mathbf{Q} \bar{\mathbf{y}} \implies \mathbf{Q}^{-1} \mathbf{A} \mathbf{Q} \bar{\mathbf{x}} = \bar{\mathbf{y}}$

**Similar Transformation:**  $\bar{\mathbf{A}} = \mathbf{Q}^{-1} \mathbf{A} \mathbf{Q}$      $\mathbf{A} = \mathbf{Q} \bar{\mathbf{A}} \mathbf{Q}^{-1}$

- Equivalently,  $\mathbf{A}\mathbf{Q} = \mathbf{Q}\bar{\mathbf{A}}$  or

$$\mathbf{A}[\mathbf{q}_1 \ \mathbf{q}_2 \ \cdots \ \mathbf{q}_n] = [\mathbf{A}\mathbf{q}_1 \ \mathbf{A}\mathbf{q}_2 \ \cdots \ \mathbf{A}\mathbf{q}_n] = [\mathbf{q}_1 \ \mathbf{q}_2 \ \cdots \ \mathbf{q}_n]\bar{\mathbf{A}}$$

### Example 3.4: a special basis

- Suppose  $\{\mathbf{b}, \mathbf{Ab}, \dots, \mathbf{A}^{n-1}\mathbf{b}\}$  is a basis, and

$$\mathbf{A}^n \mathbf{b} = \beta_1 \mathbf{b} + \beta_2 \mathbf{Ab} + \cdots + \beta_n \mathbf{A}^{n-1} \mathbf{b}$$

- Then

$$\mathbf{A} [\mathbf{b} \ \mathbf{Ab} \ \cdots \ \mathbf{A}^{n-2}\mathbf{b} \ \mathbf{A}^{n-1}\mathbf{b}] = [\mathbf{Ab} \ \mathbf{A}^2\mathbf{b} \ \cdots \ \mathbf{A}^{n-1}\mathbf{b} \ \mathbf{A}^n\mathbf{b}]$$

$$= [\mathbf{b} \ \mathbf{Ab} \ \mathbf{A}^2\mathbf{b} \ \cdots \ \mathbf{A}^{n-1}\mathbf{b}]$$

$$\begin{bmatrix} 0 & 0 & \cdots & 0 & \beta_1 \\ 1 & 0 & \cdots & 0 & \beta_2 \\ 0 & 1 & \cdots & 0 & \beta_3 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & \beta_{n-1} \\ 0 & 0 & \cdots & 1 & \beta_n \end{bmatrix}$$

$\bar{\mathbf{A}}$  is in a companion form

## Diagonal & Jordan Forms – 1 (3.5)

- An eigenvalue/eigenvector pair  $\{\lambda, \mathbf{x}\}$  of an  $n \times n$  real  $\mathbf{A}$ :

$$(\mathbf{A} - \lambda \mathbf{I}_n) \mathbf{x} = \mathbf{0}, \quad \mathbf{x} \neq \mathbf{0}$$

- $\mathbf{A}$  has distinct eigenvalues  $\{\lambda_1, \lambda_2, \dots, \lambda_n\}$ ,

then corresponding eigenvectors  $\{\mathbf{q}_1, \mathbf{q}_2, \dots, \mathbf{q}_n\}$  form a basis

- Characteristic polynomial of  $\mathbf{A}$ : (monic, deg. =  $n$ )

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

$$= \lambda^n + \alpha_1 \lambda^{n-1} + \cdots + \alpha_{n-1} \lambda + \alpha_n$$

## Diagonal & Jordan Forms – 2

Feng-Li Lian © 2007  
NTUEE-LS1-Matrix-28

$$\begin{bmatrix} 0 & 0 & 0 & -\alpha_4 \\ 1 & 0 & 0 & -\alpha_3 \\ 0 & 1 & 0 & -\alpha_2 \\ 0 & 0 & 1 & -\alpha_1 \end{bmatrix} \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -\alpha_4 & -\alpha_3 & -\alpha_2 & -\alpha_1 \end{bmatrix} \begin{bmatrix} -\alpha_1 & 1 & 0 & 0 \\ -\alpha_2 & 0 & 1 & 0 \\ -\alpha_3 & 0 & 0 & 1 \\ -\alpha_4 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} -\alpha_1 & -\alpha_2 & -\alpha_3 & -\alpha_4 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

$$\det \begin{bmatrix} \lambda & 0 & 0 & \alpha_4 \\ -1 & \lambda & 0 & \alpha_3 \\ 0 & -1 & \lambda & \alpha_2 \\ 0 & 0 & -1 & \lambda + \alpha_1 \end{bmatrix} = \lambda \det \begin{bmatrix} \lambda & 0 & \alpha_3 \\ -1 & \lambda & \alpha_2 \\ 0 & -1 & \lambda + \alpha_1 \end{bmatrix} - \alpha_4 \det \begin{bmatrix} -1 & \lambda & 0 \\ 0 & -1 & \lambda \\ 0 & 0 & -1 \end{bmatrix}$$

→  $\Delta(\lambda) = \lambda^4 + \alpha_1 \lambda^3 + \alpha_2 \lambda^2 + \alpha_3 \lambda + \alpha_4$

### Example 3.5

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

$$\Delta(\lambda) = \det \begin{bmatrix} \lambda & 0 & 0 \\ -1 & \lambda & -2 \\ 0 & -1 & \lambda - 1 \end{bmatrix} = (\lambda - 2)(\lambda + 1)\lambda$$

$$\mathbf{Q} = [\mathbf{q}_1 \ \mathbf{q}_2 \ \mathbf{q}_3] = \begin{bmatrix} 0 & 0 & 2 \\ 1 & -2 & 1 \\ 1 & 1 & -1 \end{bmatrix}$$

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

In general,  $\hat{\mathbf{A}} = \text{diag}[\lambda_1, \lambda_2, \dots, \lambda_n]$

### Example 3.6

$$\mathbf{A} = \begin{bmatrix} -1 & 1 & 1 \\ 0 & 4 & -13 \\ 0 & 1 & 0 \end{bmatrix} \quad \text{has eigenvalues } -1, 2 \pm 3j.$$

$$\mathbf{Q} = \begin{bmatrix} 1 & j & -j \\ 0 & -3+2j & -3-2j \\ 0 & j & j \end{bmatrix}$$

Real matrices may have

complex eigenvalues & eigenvectors (in conjugate pairs).

- $\mathbf{A}$  has distinct eigenvalues  $\{\lambda_1, \lambda_2, \dots, \lambda_m\}$ ,  $m < n$ ,  
then some eigenvalues are repeated

- Suppose  $\lambda_i$  is repeated  $n_i$  times,

then  $\sum_i n_i = n$ , and

there are  $l_i$  eigenvectors corresponding to  $\lambda_i$ ,

where  $1 \leq l_i \leq n_i$

- There exists a set of eigenvectors and  
generalized eigenvectors forming a nonsingular matrix  $\mathbf{Q}$

such that  $\hat{\mathbf{A}} = \mathbf{Q}^{-1}\mathbf{A}\mathbf{Q}$  has the Jordan form

$$\begin{bmatrix} \lambda_1 & 1 & & 0 \\ \lambda_1 & 1 & & \\ \lambda_1 & & \ddots & \\ 0 & & & \lambda_1 \end{bmatrix} \xrightarrow{\text{Transformation}} \begin{bmatrix} \mathbf{J}_1^{(1)} & & & \\ & \ddots & & \\ & & \mathbf{J}_{l_1}^{(1)} & \\ & & & n_1 \times n_1 \\ & & & \\ & & & \mathbf{0} \\ & & & \\ & & & \\ & & & \mathbf{J}_1^{(2)} & \\ & & & & n_2 \times n_2 \\ & & & & \\ & & & & \mathbf{0} \\ & & & & \\ & & & & \mathbf{J}_1^{(m)} & \\ & & & & & n_m \times n_m \\ & & & & & \\ & & & & & \mathbf{J}_{l_m}^{(m)} \end{bmatrix}$$

- Consider the equality

$$\mathbf{AQ} = \mathbf{A} [\mathbf{q}_1 \mathbf{q}_2 \cdots \mathbf{q}_k \hat{\mathbf{Q}}] = [\mathbf{q}_1 \mathbf{q}_2 \cdots \mathbf{q}_k \hat{\mathbf{Q}}]$$

$$\begin{bmatrix} \lambda & 1 & & \\ & \ddots & & \\ & & \lambda & 1 \\ & & & \ddots \\ & & & & \lambda \end{bmatrix} \hat{\mathbf{J}}$$

- We have  $\mathbf{Aq}_1 = \lambda \mathbf{q}_1$ ,

$$\mathbf{Aq}_2 = \lambda \mathbf{q}_2 + \mathbf{q}_1,$$

...,

$$\mathbf{Aq}_k = \lambda \mathbf{q}_k + \mathbf{q}_{k-1}$$

and

$$(\mathbf{A} - \lambda \mathbf{I}) \mathbf{q}_k = \mathbf{q}_{k-1}$$

$$(\mathbf{A} - \lambda \mathbf{I}) \mathbf{q}_{k-1} = \mathbf{q}_{k-2}$$

$$(\mathbf{A} - \lambda \mathbf{I})^2 \mathbf{q}_k = \mathbf{q}_{k-2}$$

...

$$(\mathbf{A} - \lambda \mathbf{I})^{k-1} \mathbf{q}_k = \mathbf{q}_1$$

$$(\mathbf{A} - \lambda \mathbf{I})^k \mathbf{q}_k = \mathbf{0}$$

$\{\mathbf{q}_1, \mathbf{q}_2, \dots, \mathbf{q}_k\}$ : chain of generalized eigenvectors of length  $k$

$\mathbf{q}_k$ : a generalized eigenvector of grade  $k$

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

$$\mathbf{A} = \mathbf{Q} \hat{\mathbf{A}} \mathbf{Q}^{-1}$$

$$\det \mathbf{Q} \det \mathbf{Q}^{-1} = \det \mathbf{I} = 1$$

$$\det \mathbf{A} = \det \mathbf{Q} \det \hat{\mathbf{A}} \det \mathbf{Q}^{-1} = \det \hat{\mathbf{A}}$$

= product of all eigenvalues of  $\mathbf{A}$

$\mathbf{A}$  is nonsingular if and only if  $\mathbf{A}$  has no zero eigenvalue.

### Nilpotent

$$\mathbf{J} := \begin{bmatrix} \lambda & 1 & 0 & 0 \\ 0 & \lambda & 1 & 0 \\ 0 & 0 & \lambda & 1 \\ 0 & 0 & 0 & \lambda \end{bmatrix} \quad (\mathbf{J} - \lambda \mathbf{I}) = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

$$(\mathbf{J} - \lambda \mathbf{I})^2 = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

$$(\mathbf{J} - \lambda \mathbf{I})^3 = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

$$(\mathbf{J} - \lambda \mathbf{I})^k = \mathbf{0} \text{ for } k \geq 4.$$

a nilpotent

## Functions of a square matrix A – 1 (3.6)

integer power  $\mathbf{A}^k$  = 
$$\begin{cases} \mathbf{AA} \cdots \mathbf{A} & \text{for } k > 0 \\ \mathbf{I} & \text{for } k = 0 \\ \mathbf{A}^{-1}\mathbf{A}^{-1} \cdots \mathbf{A}^{-1} & \text{for } k < 0 \end{cases}$$

Note that:  $\mathbf{A}^{21} = \mathbf{A}^{16}\mathbf{A}^4\mathbf{A}$  for faster computation

## Functions of a square matrix A – 2

Feng-Li Lian © 2007  
NTUEE-LS1-Matrix-38

If  $f(\lambda) = \lambda^m + c_1\lambda^{m-1} + \cdots + c_{m-1}\lambda + c_m$   
 $= (\lambda - r_1)(\lambda - r_2) \cdots (\lambda - r_m)$

Polynomial  $f(\mathbf{A}) = \mathbf{A}^m + c_1\mathbf{A}^{m-1} + \cdots + c_{m-1}\mathbf{A} + c_m\mathbf{I}$   
 $= (\mathbf{A} - r_1\mathbf{I})(\mathbf{A} - r_2\mathbf{I}) \cdots (\mathbf{A} - r_m\mathbf{I})$

$$\mathbf{A} = \begin{bmatrix} \mathbf{A}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{A}_2 \end{bmatrix} \quad \rightarrow \quad \mathbf{A}^k = \begin{bmatrix} \mathbf{A}_1^k & \mathbf{0} \\ \mathbf{0} & \mathbf{A}_2^k \end{bmatrix}$$

$$f(\mathbf{A}) = \begin{bmatrix} f(\mathbf{A}_1) & \mathbf{0} \\ \mathbf{0} & f(\mathbf{A}_2) \end{bmatrix}$$

$$\mathbf{A} = \mathbf{Q}\hat{\mathbf{A}}\mathbf{Q}^{-1} \quad \rightarrow \quad \mathbf{A}^k = (\mathbf{Q}\hat{\mathbf{A}}\mathbf{Q}^{-1})(\mathbf{Q}\hat{\mathbf{A}}\mathbf{Q}^{-1}) \cdots (\mathbf{Q}\hat{\mathbf{A}}\mathbf{Q}^{-1}) = \mathbf{Q}\hat{\mathbf{A}}^k\mathbf{Q}^{-1}$$

$$f(\mathbf{A}) = \mathbf{Q}f(\hat{\mathbf{A}})\mathbf{Q}^{-1}$$

### Minimal Polynomial of A

- Minimal polynomial of  $\mathbf{A}$ :  $\psi(\lambda) = \prod(\lambda - \lambda_i)^{\bar{n}_i}$

where  $\bar{n}_i$  is the dim. of the largest Jordan block associated with  $\lambda_i$ , also called the index of  $\lambda_i$ ,

- $\psi(\lambda)$  divides  $\Delta(\lambda)$ , the char. poly. of  $\mathbf{A}$ , and  $\psi(\mathbf{A}) = \mathbf{0}$

## Jordan Form

$$A = QDQ^T$$

$D$	$\begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_1 & 0 \\ 0 & 0 & \lambda_1 \end{bmatrix}$	$\begin{bmatrix} \lambda_1 & 1 & 0 \\ 0 & \lambda_1 & 0 \\ 0 & 0 & \lambda_1 \end{bmatrix}$	$\begin{bmatrix} \lambda_1 & 1 & 0 \\ 0 & \lambda_1 & 1 \\ 0 & 0 & \lambda_1 \end{bmatrix}$
$\lambda I - D$	$\begin{bmatrix} \lambda - \lambda_1 & 0 & 0 \\ 0 & \lambda - \lambda_1 & 0 \\ 0 & 0 & \lambda - \lambda_1 \end{bmatrix}$	$\begin{bmatrix} \lambda - \lambda_1 & -1 & 0 \\ 0 & \lambda - \lambda_1 & 0 \\ 0 & 0 & \lambda - \lambda_1 \end{bmatrix}$	$\begin{bmatrix} \lambda - \lambda_1 & -1 & 0 \\ 0 & \lambda - \lambda_1 & -1 \\ 0 & 0 & \lambda - \lambda_1 \end{bmatrix}$
$\Delta(\lambda)$	$(\lambda - \lambda_1)(\lambda - \lambda_1)(\lambda - \lambda_1)$		
$D - \lambda_1 I$	$\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$	$\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$	$\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$
	$(D - \lambda_1 I) = 0$	$(D - \lambda_1 I)^2 = 0$	$(D - \lambda_1 I)^3 = 0$
$\Psi(\lambda)$	$(\lambda - \lambda_1)$	$(\lambda - \lambda_1)^2$	$(\lambda - \lambda_1)^3$

## Generalized Eigenvectors – 1

$$A = QDQ^T$$

$D$	$\begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_1 & 0 \\ 0 & 0 & \lambda_1 \end{bmatrix}$	$\begin{bmatrix} \lambda_1 & 1 & 0 \\ 0 & \lambda_1 & 0 \\ 0 & 0 & \lambda_1 \end{bmatrix}$	$\begin{bmatrix} \lambda_1 & 1 & 0 \\ 0 & \lambda_1 & 1 \\ 0 & 0 & \lambda_1 \end{bmatrix}$
$\Delta(\lambda)$	$(\lambda - \lambda_1)(\lambda - \lambda_1)(\lambda - \lambda_1)$		
$\lambda_1 I - D$	$\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$	$\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$	$\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$
$nullity$	3	2	1
$\mathbf{v}_i$	$\begin{bmatrix} * \\ * \\ * \end{bmatrix}$	$\begin{bmatrix} * \\ 0 \\ * \end{bmatrix}$	$\begin{bmatrix} * \\ 0 \\ 0 \end{bmatrix}$
$(\lambda_1 I - D)\mathbf{v}_i = 0$			

## Generalized Eigenvectors – 2

$$A = QDQ^T$$

$D$

$$\mathbf{v}_3 = \mathbf{v}$$

$$\begin{bmatrix} \lambda_1 & 1 & 0 \\ 0 & \lambda_1 & 1 \\ 0 & 0 & \lambda_1 \end{bmatrix}$$

$$\mathbf{v}_2 = (D - \lambda_1 I)\mathbf{v}$$

$$\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

$$\mathbf{v}_1 = (D - \lambda_1 I)^2\mathbf{v}$$

$$\begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

$$\Rightarrow (D - \lambda_1 I)^3\mathbf{v} = 0 \text{ but } (D - \lambda_1 I)^2\mathbf{v} \neq 0$$

$$\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

1

$\Rightarrow \mathbf{v}$  : a generalized eigenvector of grade 3

$$\begin{bmatrix} * \\ 0 \\ 0 \end{bmatrix}$$

## Generalized Eigenvectors – 3

$$A = QDQ^T \quad \mathbf{v}_3 = \mathbf{v}$$

$D$

$$\mathbf{v}_2 = (D - \lambda_1 I)\mathbf{v}$$

$$\mathbf{v}_1 = (D - \lambda_1 I)^2\mathbf{v}$$

$$\Rightarrow (D - \lambda_1 I)^3\mathbf{v} = 0 \text{ but } (D - \lambda_1 I)^2\mathbf{v} \neq 0$$

$$\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

$$\mathbf{v}_1 = (D - \lambda_1 I)^3 \mathbf{v} = 0 \Rightarrow \mathbf{v}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

$$\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

$$\mathbf{v}_2 = \mathbf{v}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \Rightarrow \mathbf{v}_2 = \begin{bmatrix} * \\ 1 \\ 0 \end{bmatrix}$$

$$\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

$$\mathbf{v}_3 = \mathbf{v}_2 = \begin{bmatrix} * \\ 1 \\ 0 \end{bmatrix} \Rightarrow \mathbf{v}_3 = \begin{bmatrix} * \\ * \\ 1 \end{bmatrix}$$

**Theorem 3.4 (Cayley-Hamilton theorem)**

Let

$$\Delta(\lambda) = \det(\lambda\mathbf{I} - \mathbf{A}) = \lambda^n + \alpha_1\lambda^{n-1} + \cdots + \alpha_{n-1}\lambda + \alpha_n$$

be the characteristic polynomial of  $\mathbf{A}$ . Then

$$\Delta(\mathbf{A}) = \mathbf{A}^n + \alpha_1\mathbf{A}^{n-1} + \cdots + \alpha_{n-1}\mathbf{A} + \alpha_n\mathbf{I} = \mathbf{0}$$

- The C-H Theorem implies that

$\mathbf{A}^k$  can be linearly combined

by  $\{\mathbf{I}, \mathbf{A}, \dots, \mathbf{A}^{n-1}\}$  for any  $k \geq 0$ :

$$\mathbf{A}^{n+1} + \alpha_1\mathbf{A}^n + \cdots + \alpha_{n-1}\mathbf{A}^2 + \alpha_n\mathbf{A} = \mathbf{0} \cdot \mathbf{A} = \mathbf{0}$$

- Given any polynomial  $f(\lambda)$ , there exist  $\beta_0, \beta_1, \dots, \beta_{n-1}$ , s.t.

$$f(\lambda) = \alpha_0 + \alpha_1 \lambda + \cdots + \alpha_{n-1} \lambda^{n-1} + \alpha_n \lambda^n + \alpha_{n+1} \lambda^{n+1} + \cdots$$

$$f(\mathbf{A}) = \alpha_0 + \alpha_1 \mathbf{A} + \cdots + \alpha_{n-1} \mathbf{A}^{n-1} + \alpha_n \mathbf{A}^n + \alpha_{n+1} \mathbf{A}^{n+1} + \cdots$$

$$= \beta_0 + \beta_1 \mathbf{A} + \cdots + \beta_{n-1} \mathbf{A}^{n-1}$$

Let  $h(\lambda) := \beta_0 + \beta_1 \lambda + \cdots + \beta_{n-1} \lambda^{n-1}$

- To determine  $\beta_0, \beta_1, \dots, \beta_{n-1}$  may use long division:

$$f(\lambda) = q(\lambda) \Delta(\lambda) + h(\lambda)$$

$$f(\mathbf{A}) = q(\mathbf{A}) \Delta(\mathbf{A}) + h(\mathbf{A}) = q(\mathbf{A}) \mathbf{0} + h(\mathbf{A}) = h(\mathbf{A})$$

$$f^{(l)}(\lambda_i) = h^{(l)}(\lambda_i) \quad \text{for } l = 0, 1, \dots, n_i - 1 \quad \text{and } i = 1, 2, \dots, m$$

### Theorem 3.5: Polynomial Representation – 1

- Given  $f(\lambda)$   
and an  $n \times n$  matrix  $\mathbf{A}$  with charac. poly.

$$\Delta(\lambda) = \prod_{i=1}^m (\lambda - \lambda_i)^{n_i}$$

where  $n = \sum_{i=1}^m n_i$

- Define  $h(\lambda) = \beta_0 + \beta_1 \lambda + \cdots + \beta_{n-1} \lambda^{n-1}$   
These  $n$  unknowns are solved by:

$$f^{(l)}(\lambda_i) = h^{(l)}(\lambda_i)$$

for  $l = 0, 1, \dots, n_i - 1$ , and  $i = 1, 2, \dots, m$

where

$$f^{(l)}(\lambda_i) := \frac{d^l f(\lambda)}{d\lambda^l} \Big|_{\lambda=\lambda_i}, \quad h^{(l)}(\lambda_i) := \frac{d^l h(\lambda)}{d\lambda^l} \Big|_{\lambda=\lambda_i}$$

- Then we have

$$f(\mathbf{A}) = h(\mathbf{A})$$

and  $h(\lambda)$  is said to equal  $f(\lambda)$   
on the spectrum of  $\mathbf{A}$

## Example 3.7 – 1

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

$$\Rightarrow \mathbf{A}^{100} = \begin{bmatrix} 0 & 1 \\ -1 & -2 \end{bmatrix}^{100}$$

$$\Rightarrow f(\lambda) = \lambda^{100}$$

$$\Rightarrow h(\lambda) = \beta_0 + \beta_1 \lambda$$

$$\Rightarrow f(\quad) = h(\quad)$$

⇒

### Example 3.7 – 2

Feng-Li Lian © 2007  
NTUEE-LS1-Matrix-51

$$\Delta(\lambda) = \lambda^2 + 2\lambda + 1 = (\lambda + 1)^2$$

$$h(\lambda) = \beta_0 + \beta_1 \lambda$$

$$f(-1) = h(-1) : \quad (-1)^{100} = \beta_0 - \beta_1$$

$$f'(-1) = h'(-1) : \quad 100 \cdot (-1)^{99} = \beta_1$$

$$\beta_1 = -100, \beta_0 = 1 + \beta_1 = -99$$

$$\mathbf{A}^{100} = \beta_0 \mathbf{I} + \beta_1 \mathbf{A} = -99 \mathbf{I} - 100 \mathbf{A}$$

$$= -99 \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} - 100 \begin{bmatrix} 0 & 1 \\ -1 & -2 \end{bmatrix} = \begin{bmatrix} -199 & -100 \\ 100 & 101 \end{bmatrix}$$

### Example 3.8 – 1

Feng-Li Lian © 2007  
NTUEE-LS1-Matrix-52

$$\mathbf{A}_1 = \begin{bmatrix} 0 & 0 & -2 \\ 0 & 1 & 0 \\ 1 & 0 & 3 \end{bmatrix}$$

$$\Rightarrow e^{\mathbf{A}_1 t} = e^{\begin{bmatrix} 0 & 0 & -2 \\ 0 & 1 & 0 \\ 1 & 0 & 3 \end{bmatrix} t} \text{ or } \exp\left(\begin{bmatrix} 0 & 0 & -2 \\ 0 & 1 & 0 \\ 1 & 0 & 3 \end{bmatrix} t\right)$$

$$\Rightarrow \mathbf{f}(\lambda) = e^{\lambda t}$$

$$\Rightarrow \mathbf{h}(\lambda) = \beta_0 + \beta_1 \lambda + \beta_2 \lambda^2$$

$$\Rightarrow \mathbf{f}( ) = \mathbf{h}( )$$

⇒

### Example 3.8 – 2

Feng-Li Lian © 2007  
NTUEE-LS1-Matrix-53

$$\Delta(\lambda) = (\lambda - 1)^2(\lambda - 2)$$

$$h(\lambda) = \beta_0 + \beta_1\lambda + \beta_2\lambda^2$$

$$f(1) = h(1) : \quad e^t = \beta_0 + \beta_1 + \beta_2$$

$$f'(1) = h'(1) : \quad te^t = \beta_1 + 2\beta_2$$

$$f(2) = h(2) : \quad e^{2t} = \beta_0 + 2\beta_1 + 4\beta_2$$

$$\beta_0 = -2te^t + e^{2t}, \beta_1 = 3te^t + 2e^t - 2e^{2t}, \text{ and } \beta_2 = e^{2t} - e^t - te^t$$

$$e^{\mathbf{A}_1 t} = h(\mathbf{A}_1) = \begin{bmatrix} 2e^t - e^{2t} & 0 & 2e^t - 2e^{2t} \\ 0 & e^t & 0 \\ e^{2t} - e^t & 0 & 2e^{2t} - e^t \end{bmatrix}$$

### Example 3.10 – 1

Feng-Li Lian © 2007  
NTUEE-LS1-Matrix-54

$$\hat{\mathbf{A}} = \begin{bmatrix} \lambda_1 & 1 & 0 & 0 \\ 0 & \lambda_1 & 1 & 0 \\ 0 & 0 & \lambda_1 & 1 \\ 0 & 0 & 0 & \lambda_1 \end{bmatrix}$$

⇒ arbitrary  $\mathbf{f}(\lambda)$

⇒ Select a convenient form of  $h(\lambda)$  (i.e., different set of  $\beta_i$ 's)

$$\Rightarrow h(\lambda) = \beta_0 + \beta_1(\lambda - \lambda_1) + \beta_2(\lambda - \lambda_1)^2 + \beta_3(\lambda - \lambda_1)^3$$

Then the “[interpolation](#)” conditions yield

$$\beta_0 = f(\lambda_1), \quad \beta_1 = f'(\lambda_1), \quad \beta_2 = \frac{f''(\lambda_1)}{2!}, \quad \beta_3 = \frac{f^{(3)}(\lambda_1)}{3!}$$

Therefore

$$f(\hat{\mathbf{A}}) = f(\lambda_1)\mathbf{I} + \frac{f'(\lambda_1)}{1!}(\hat{\mathbf{A}} - \lambda_1\mathbf{I}) + \frac{f''(\lambda_1)}{2!}(\hat{\mathbf{A}} - \lambda_1\mathbf{I})^2 + \frac{f^{(3)}(\lambda_1)}{3!}(\hat{\mathbf{A}} - \lambda_1\mathbf{I})^3$$

$$= \begin{bmatrix} f(\lambda_1) & f'(\lambda_1)/1! & f''(\lambda_1)/2! & f^{(3)}(\lambda_1)/3! \\ 0 & f(\lambda_1) & f'(\lambda_1)/1! & f''(\lambda_1)/2! \\ 0 & 0 & f(\lambda_1) & f'(\lambda_1)/1! \\ 0 & 0 & 0 & f(\lambda_1) \end{bmatrix}$$

$$\hat{\mathbf{A}} = \begin{bmatrix} \lambda_1 & 1 & 0 & 0 \\ 0 & \lambda_1 & 1 & 0 \\ 0 & 0 & \lambda_1 & 1 \\ 0 & 0 & 0 & \lambda_1 \end{bmatrix}$$

$$e^{\hat{\mathbf{A}}t} = \begin{bmatrix} e^{\lambda_1 t} & te^{\lambda_1 t} & t^2 e^{\lambda_1 t}/2! & t^3 e^{\lambda_1 t}/3! \\ 0 & e^{\lambda_1 t} & te^{\lambda_1 t} & t^2 e^{\lambda_1 t}/2! \\ 0 & 0 & e^{\lambda_1 t} & te^{\lambda_1 t} \\ 0 & 0 & 0 & e^{\lambda_1 t} \end{bmatrix}$$

### Example 3.11

$$\mathbf{A} = \begin{bmatrix} \lambda_1 & 1 & 0 & 0 & 0 \\ 0 & \lambda_1 & 1 & 0 & 0 \\ 0 & 0 & \lambda_1 & 0 & 0 \\ 0 & 0 & 0 & \lambda_2 & 1 \\ 0 & 0 & 0 & 0 & \lambda_2 \end{bmatrix}$$

$$f(\lambda) = e^{\lambda t}$$

$$e^{\mathbf{At}} = \begin{bmatrix} e^{\lambda_1 t} & te^{\lambda_1 t} & t^2 e^{\lambda_1 t}/2! & 0 & 0 \\ 0 & e^{\lambda_1 t} & te^{\lambda_1 t} & 0 & 0 \\ 0 & 0 & e^{\lambda_1 t} & 0 & 0 \\ 0 & 0 & 0 & e^{\lambda_2 t} & te^{\lambda_2 t} \\ 0 & 0 & 0 & 0 & e^{\lambda_2 t} \end{bmatrix}$$

$$f(\lambda) = (s - \lambda)^{-1}$$

$$(s\mathbf{I} - \mathbf{A})^{-1} = \begin{bmatrix} \frac{1}{(s - \lambda_1)} & \frac{1}{(s - \lambda_1)^2} & \frac{1}{(s - \lambda_1)^3} & 0 & 0 \\ 0 & \frac{1}{(s - \lambda_1)} & \frac{1}{(s - \lambda_1)^2} & 0 & 0 \\ 0 & 0 & \frac{1}{(s - \lambda_1)} & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{(s - \lambda_2)} & \frac{1}{(s - \lambda_2)^2} \\ 0 & 0 & 0 & 0 & \frac{1}{(s - \lambda_2)} \end{bmatrix}$$

### Using Power Series

- Defining functions of  $\mathbf{A}$  with power series

$$f(\lambda) = \sum_{i=0}^{\infty} \beta_i \lambda^i \text{ with the radius of convergence } \rho$$

$$f(\mathbf{A}) = \sum_{i=0}^{\infty} \beta_i \mathbf{A}^i \quad \text{if } |\lambda_i(\mathbf{A})| < \rho \text{ for all } i$$

### Example 3.12

$$\hat{\mathbf{A}} = \begin{bmatrix} \lambda_1 & 1 & 0 & 0 \\ 0 & \lambda_1 & 1 & 0 \\ 0 & 0 & \lambda_1 & 1 \\ 0 & 0 & 0 & \lambda_1 \end{bmatrix} \Rightarrow \text{arbitrary } \mathbf{f}(\lambda)$$

$$f(\lambda) = f(\lambda_1) + f'(\lambda_1)(\lambda - \lambda_1) + \frac{f''(\lambda_1)}{2!}(\lambda - \lambda_1)^2 + \dots$$

$$f(\hat{\mathbf{A}}) = f(\lambda_1)\mathbf{I} + f'(\lambda_1)(\hat{\mathbf{A}} - \lambda_1\mathbf{I}) + \dots + \frac{f^{(n-1)}(\lambda_1)}{(n-1)!}(\hat{\mathbf{A}} - \lambda_1\mathbf{I})^{n-1} + \dots$$

$$\begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

in this case,  $(\hat{\mathbf{A}} - \lambda_1\mathbf{I})^k = \mathbf{0}$  for  $k \geq n = 4$

### Matrix Exponential Function – 1

$$e^{\lambda t} = 1 + \lambda t + \frac{\lambda^2 t^2}{2!} + \dots + \frac{\lambda^n t^n}{n!} + \dots \quad \forall \lambda, t$$

$$e^{\mathbf{At}} = \mathbf{I} + t\mathbf{A} + \frac{t^2}{2!}\mathbf{A}^2 + \dots = \sum_{k=0}^{\infty} \frac{1}{k!} t^k \mathbf{A}^k \quad \forall \mathbf{A}, t$$

$$e^{\mathbf{0}} = \mathbf{I}$$

$$e^{\mathbf{A}(t_1+t_2)} = e^{\mathbf{At}_1} e^{\mathbf{At}_2} \quad \text{because } e^{\lambda(t_1+t_2)} = e^{\lambda t_1} e^{\lambda t_2} \text{ and } \mathbf{A}^k \mathbf{A}^l = \mathbf{A}^l \mathbf{A}^k$$

$$[e^{\mathbf{At}}]^{-1} = e^{-\mathbf{At}} \quad \text{because } e^{\mathbf{At}} e^{-\mathbf{At}} = e^{\mathbf{A}(t-t)} = e^{\mathbf{0}} = \mathbf{I}$$

$$e^{(\mathbf{A}+\mathbf{B})t} \neq e^{\mathbf{A}t} e^{\mathbf{B}t} \quad \text{because } \mathbf{A}\mathbf{B} \neq \mathbf{B}\mathbf{A} \text{ in general}$$

$$e^{\mathbf{A}t} = \mathbf{I} + t\mathbf{A} + \frac{t^2}{2!}\mathbf{A}^2 + \dots$$

$$e^{\mathbf{B}t} = \mathbf{I} + t\mathbf{B} + \frac{t^2}{2!}\mathbf{B}^2 + \dots$$

$$e^{\mathbf{A}t} e^{\mathbf{B}t} = \left( \mathbf{I} + t\mathbf{A} + \frac{t^2}{2!}\mathbf{A}^2 + \dots \right) \left( \mathbf{I} + t\mathbf{B} + \frac{t^2}{2!}\mathbf{B}^2 + \dots \right)$$

$$= \mathbf{I} + t\mathbf{A} + t\mathbf{B} + \frac{t^2}{2!}\mathbf{A}^2 + \frac{t^2}{2!}\mathbf{B}^2 + t^2\mathbf{AB} + \frac{t^3}{2!}\mathbf{A}^2\mathbf{B} + \frac{t^3}{2!}\mathbf{A}\mathbf{B}^2 + \dots$$

$$e^{(\mathbf{A}+\mathbf{B})t} = \mathbf{I} + t(\mathbf{A}+\mathbf{B}) + \frac{t^2}{2!}(\mathbf{A}+\mathbf{B})^2 + \dots$$

$$\mathbf{A}^2 + \mathbf{B}^2 + 2\mathbf{AB} = (\mathbf{A} + \mathbf{B})^2 = \mathbf{A}^2 + \mathbf{AB} + \mathbf{BA} + \mathbf{B}^2$$

$$e^{\mathbf{A}t} = \mathbf{I} + t\mathbf{A} + \frac{t^2}{2!}\mathbf{A}^2 + \dots$$

$$\frac{d}{dt} e^{\mathbf{A}t} = \sum_{k=1}^{\infty} \frac{1}{(k-1)!} t^{k-1} \mathbf{A}^k$$

$$= \mathbf{A} \left( \sum_{k=0}^{\infty} \frac{1}{k!} t^k \mathbf{A}^k \right) = \left( \sum_{k=0}^{\infty} \frac{1}{k!} t^k \mathbf{A}^k \right) \mathbf{A} = \mathbf{A} e^{\mathbf{A}t} = e^{\mathbf{A}t} \mathbf{A}$$

$$e^{\mathbf{A}t} = \mathbf{I} + t\mathbf{A} + \frac{t^2}{2!}\mathbf{A}^2 + \dots \Rightarrow \mathcal{L}[e^{\mathbf{A}t}] = ?$$

$$\mathcal{L}[1] =$$

$$\mathcal{L}[t] =$$

$$\mathcal{L}\left[\frac{t^2}{2!}\right] =$$

$$1 + s^{-1}\lambda + s^{-2}\lambda^2 + \dots$$

$$= (1 - s^{-1}\lambda)^{-1}$$

converges for  $|s^{-1}\lambda| < 1$

$$\mathcal{L}[e^{\mathbf{A}t}] = s^{-1}\mathbf{I} + s^{-2}\mathbf{A} + s^{-3}\mathbf{A}^2 + \dots$$

$$= s^{-1}(\mathbf{I} + s^{-1}\mathbf{A} + s^{-2}\mathbf{A}^2 + \dots)$$

$$= s^{-1}(\mathbf{I} - s^{-1}\mathbf{A})^{-1} \quad \text{for } |s^{-1}\lambda_i(\mathbf{A})| < 1, \forall i$$

$$= [s(\mathbf{I} - s^{-1}\mathbf{A})]^{-1}$$

$$= (s\mathbf{I} - \mathbf{A})^{-1}$$

### Lyapunov Equation $\mathbf{AM} + \mathbf{MB} = \mathbf{C}$ (3.7)

- Lyapunov equation:  $\mathbf{AM} + \mathbf{MB} = \mathbf{C}$

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \begin{bmatrix} m_{11} & m_{12} \\ m_{21} & m_{22} \\ m_{31} & m_{32} \end{bmatrix} + \begin{bmatrix} m_{11} & m_{12} \\ m_{21} & m_{22} \\ m_{31} & m_{32} \end{bmatrix} \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} = \begin{bmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \\ c_{31} & c_{32} \end{bmatrix}$$

$$\begin{bmatrix} a_{11} + b_{11} & a_{12} & a_{13} & b_{21} & 0 & 0 \\ a_{21} & a_{22} + b_{11} & a_{23} & 0 & b_{21} & 0 \\ a_{31} & a_{32} & a_{33} + b_{11} & 0 & 0 & b_{21} \\ b_{12} & 0 & 0 & a_{11} + b_{22} & a_{12} & a_{13} \\ 0 & b_{12} & 0 & a_{21} & a_{22} + b_{22} & a_{23} \\ 0 & 0 & b_{12} & a_{31} & a_{32} & a_{33} + b_{22} \end{bmatrix} \begin{bmatrix} m_{11} \\ m_{21} \\ m_{31} \\ m_{12} \\ m_{22} \\ m_{32} \end{bmatrix} = \begin{bmatrix} c_{11} \\ c_{21} \\ c_{31} \\ c_{12} \\ c_{22} \\ c_{32} \end{bmatrix}$$

matrix representation of  $\mathcal{A}$  w.r.t. a certain basis

$$\mathcal{A}(\mathbf{M}) = \mathbf{C} \quad \text{where } \mathcal{A}(\mathbf{M}) := \mathbf{AM} + \mathbf{MB}$$

$\mathbf{A}\mathbf{u} = \lambda_i \mathbf{u}$  : right eigenvalue-eigenvector pair

$\mathbf{v}\mathbf{B} = \mathbf{v}\mu_j$  : left eigenvalue-eigenvector pair

$$\mathcal{A}(\mathbf{u}\mathbf{v}) = \mathbf{A}\mathbf{u}\mathbf{v} + \mathbf{u}\mathbf{v}\mathbf{B} = \lambda_i \mathbf{u}\mathbf{v} + \mathbf{u}\mathbf{v}\mu_j = (\lambda_i + \mu_j)\mathbf{u}\mathbf{v}$$

→  $(\lambda_i + \mu_j)$  is an eigenvalue of  $\mathcal{A}$

→  $\mathcal{A}$  is nonsingular if  $\lambda_i + \mu_j \neq 0$  for all  $i, j$

### Some Useful Formulas – 1 (3.8)

- $\mathbf{A}: m \times n, \mathbf{B}: n \times p$

$$\rho(\mathbf{A}) + \rho(\mathbf{B}) - n \leq \rho(\mathbf{AB}) \leq \min(\rho(\mathbf{A}), \rho(\mathbf{B}))$$

$$\therefore \rho(\mathbf{A}) + \rho(\mathbf{B}) - n = \rho(\mathbf{B}) - \text{nullity}(\mathbf{A})$$

$$\boxed{\rho(\mathbf{AC}) = \rho(\mathbf{A}) = \rho(\mathbf{DA})} \quad \text{if } \mathbf{C} \text{ and } \mathbf{D} \text{ are square and nonsingular}$$

$$\therefore \rho(\mathbf{AC}) \leq \rho(\mathbf{A}) \text{ and } \rho(\mathbf{A}) = \rho(\mathbf{ACC}^{-1}) \leq \rho(\mathbf{AC})$$

- Elementary Operation Matrices:

$$\mathbf{E}_1 = \begin{bmatrix} 1 & & & 0 \\ & 1 & & \neq 0 \\ & & c & \\ & 1 & & \\ 0 & & & 1 \end{bmatrix}, \quad \mathbf{E}_2 = \begin{bmatrix} 1 & & & 0 \\ & 0 & & 1 \\ & & 1 & \\ & 1 & & 0 \\ 0 & & & 1 \end{bmatrix}, \quad \mathbf{E}_3 = \begin{bmatrix} 1 & & & 0 \\ & 1 & & \\ & & 1 & \\ & & & c \\ 0 & & & 1 \end{bmatrix}$$

elementary column (row) operations:

post- (pre-) multiply the elementary operation matrices

$$\det(\mathbf{I}_m + \mathbf{AB}) = \det(\mathbf{I}_n + \mathbf{BA}) \quad \mathbf{A}: m \times n, \mathbf{B}: n \times m$$

$$\mathbf{N} = \begin{bmatrix} \mathbf{I}_m & \mathbf{A} \\ \mathbf{0} & \mathbf{I}_n \end{bmatrix} \quad \mathbf{Q} = \begin{bmatrix} \mathbf{I}_m & \mathbf{0} \\ -\mathbf{B} & \mathbf{I}_n \end{bmatrix} \quad \mathbf{P} = \begin{bmatrix} \mathbf{I}_m & -\mathbf{A} \\ \mathbf{B} & \mathbf{I}_n \end{bmatrix}$$

$$\mathbf{NP} = \begin{bmatrix} \mathbf{I}_m + \mathbf{AB} & \mathbf{0} \\ \mathbf{B} & \mathbf{I}_n \end{bmatrix} \quad \mathbf{QP} = \begin{bmatrix} \mathbf{I}_m & -\mathbf{A} \\ \mathbf{0} & \mathbf{I}_n + \mathbf{BA} \end{bmatrix}$$

$$\det(\mathbf{NP}) = \det(\mathbf{I}_m + \mathbf{AB})$$

$$= \det \mathbf{N} \det \mathbf{P} = \det \mathbf{P}$$

$$= \det \mathbf{Q} \det \mathbf{P} = \det(\mathbf{QP}) = \det(\mathbf{I}_n + \mathbf{BA})$$

$$s^n \det(s\mathbf{I}_m - \mathbf{AB}) = s^m \det(s\mathbf{I}_n - \mathbf{BA})$$

$$\det[\mathbf{I}_m - (\mathbf{A}/\sqrt{s})(\mathbf{B}/\sqrt{s})] = \det[\mathbf{I}_n - (\mathbf{B}/\sqrt{s})(\mathbf{A}/\sqrt{s})]$$

$$\Leftrightarrow \frac{1}{s^m} \det(s\mathbf{I}_m - \mathbf{AB}) = \frac{1}{s^n} \det(s\mathbf{I}_n - \mathbf{BA})$$

$$\det(s\mathbf{I}_n - \mathbf{AB}) = \det(s\mathbf{I}_n - \mathbf{BA}) \quad \text{if } n = m$$

### Quadratic Form – 1 (3.9)

$$x_1^2 + 2x_2^2 - 3x_3^2 + 4x_1x_2 - 5x_2x_3 - 7x_1x_3$$

: real coefficients and variables  $\Rightarrow$  real values

$$= [x_1 \ x_2 \ x_3] \begin{bmatrix} 1 & 4 & -7 \\ 0 & 2 & -5 \\ 0 & 0 & -3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

$$= [x_1 \ x_2 \ x_3] \begin{bmatrix} 1 & 3 & 0 \\ 1 & 2 & -5 \\ -7 & 0 & -3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

$$= [x_1 \ x_2 \ x_3] \begin{bmatrix} 1 & 2 & -7/2 \\ 2 & 2 & -5/2 \\ -7/2 & -5/2 & -3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

$$= \mathbf{x}' \ \mathbf{M} \ \mathbf{x} \quad \text{real symmetric } \mathbf{M} \ (= \mathbf{M}')$$

To accommodate **complex** variables,  
 consider **real symmetric** **M** and the **quadratic form**  
 (\* means **complex conjugate transpose**)

$$(\mathbf{x}^* \mathbf{M} \mathbf{x})^* = \mathbf{x}^* \mathbf{M}^* \mathbf{x} = \mathbf{x}^* \mathbf{M}' \mathbf{x} = \mathbf{x}^* \mathbf{M} \mathbf{x}$$

For  $\mathbf{Mv} = \lambda \mathbf{v}$ , where  $\mathbf{v} \neq \mathbf{0}$ ,  $\mathbf{v}^* \mathbf{Mv} = \mathbf{v}^* \lambda \mathbf{v} = \lambda (\mathbf{v}^* \mathbf{v})$

Thus, **eigenvalues** of real symmetric matrices are **real**

## Diagonalizable Symmetric Matrix

Suppose  $\mathbf{x}$  is a generalized eigenvector of grade 2 or

$$(\mathbf{M} - \lambda \mathbf{I})^2 \mathbf{x} = \mathbf{0}$$

$$(\mathbf{M} - \lambda \mathbf{I})\mathbf{x} \neq \mathbf{0}$$

then

$$[(\mathbf{M} - \lambda \mathbf{I})\mathbf{x}]' (\mathbf{M} - \lambda \mathbf{I})\mathbf{x} = \mathbf{x}' (\mathbf{M}' - \lambda \mathbf{I}') (\mathbf{M} - \lambda \mathbf{I})\mathbf{x} = \mathbf{x}' (\mathbf{M} - \lambda \mathbf{I})^2 \mathbf{x}$$

$\underbrace{\quad}_{\neq 0} \qquad \qquad \qquad \underbrace{\quad}_{= 0}$

i.e., all **Jordan blocks** of **M** are  $1 \times 1$ , and **M** is **diagonalizable**

Orthogonal matrix:  $\mathbf{Q}' \mathbf{Q} = \mathbf{Q} \mathbf{Q}' = \mathbf{I}$

**Theorem 3.6**

For every real symmetric matrix  $\mathbf{M}$ , there exists an orthogonal matrix  $\mathbf{Q}$  such that

$$\mathbf{M} = \mathbf{Q}\mathbf{D}\mathbf{Q}' \quad \text{or} \quad \mathbf{D} = \mathbf{Q}'\mathbf{M}\mathbf{Q}$$

where  $\mathbf{D}$  is a diagonal matrix with the eigenvalues of  $\mathbf{M}$ , which are all real, on the diagonal.

- Positive definiteness of a real symmetric matrix  $\mathbf{M}$ :

$$\mathbf{x}' \mathbf{M} \mathbf{x} > 0 \quad \text{for all } \mathbf{x} \neq \mathbf{0}.$$

- Positive semi-definiteness of a real symmetric matrix  $\mathbf{M}$ :

$$\mathbf{x}' \mathbf{M} \mathbf{x} \geq 0 \quad \text{for all } \mathbf{x} \neq \mathbf{0}.$$

**Theorem 3.7**

A symmetric  $n \times n$  matrix  $\mathbf{M}$  is positive definite (positive semidefinite) if and only if any one of the following conditions holds.

- 1. Every eigenvalue of  $\mathbf{M}$  is positive (zero or positive).**

$$\mathbf{x}'\mathbf{M}\mathbf{x} = \mathbf{x}'\mathbf{Q}'\mathbf{D}\mathbf{Q}\mathbf{x} = \tilde{\mathbf{x}}'\mathbf{D}\tilde{\mathbf{x}} = \lambda_1\tilde{x}_1^2 + \lambda_2\tilde{x}_2^2 + \cdots + \lambda_n\tilde{x}_n^2$$

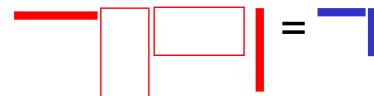
- 2. All the *leading* principal minors of  $\mathbf{M}$  are positive (all the principal minors of  $\mathbf{M}$  are zero or positive).**

$$\det[m_{11}] > 0, \quad \det \begin{bmatrix} m_{11} & m_{12} \\ m_{12} & m_{22} \end{bmatrix} > 0, \quad \det \begin{bmatrix} m_{11} & m_{12} & m_{13} \\ m_{21} & m_{22} & m_{23} \\ m_{31} & m_{32} & m_{33} \end{bmatrix} > 0, \quad \dots, \quad \det \mathbf{M} > 0$$

$$\det[m_{11}] \geq 0, \quad \det \begin{bmatrix} m_{11} & m_{12} \\ m_{12} & m_{22} \end{bmatrix} \geq 0, \quad \det \begin{bmatrix} m_{11} & m_{12} & m_{13} \\ m_{21} & m_{22} & m_{23} \\ m_{31} & m_{32} & m_{33} \end{bmatrix} \geq 0, \quad \dots, \quad \det \mathbf{M} \geq 0$$

- 3.** There exists an  $n \times n$  nonsingular matrix  $\mathbf{N}$  (an  $n \times n$  singular matrix  $\mathbf{N}$  or an  $m \times n$  matrix  $\mathbf{N}$  with  $m < n$ ) such that  $\mathbf{M} = \mathbf{N}'\mathbf{N}$ .

$$\mathbf{x}'\mathbf{M}\mathbf{x} = \mathbf{x}'\mathbf{N}'\mathbf{N}\mathbf{x} = (\mathbf{N}\mathbf{x})'(\mathbf{N}\mathbf{x}) = \|\mathbf{N}\mathbf{x}\|_2^2 \geq 0$$



If  $\mathbf{N}$  is nonsingular, the only  $\mathbf{x}$  to make  $\mathbf{N}\mathbf{x} = \mathbf{0}$  is  $\mathbf{x} = \mathbf{0}$ .

**Theorem 3.8**

- 1.** An  $m \times n$  matrix  $\mathbf{H}$ , with  $m \geq n$ , has rank  $n$ , if and only if the  $n \times n$  matrix  $\mathbf{H}'\mathbf{H}$  has rank  $n$  or  $\det(\mathbf{H}'\mathbf{H}) \neq 0$ .

**Proof :** $(\Leftarrow)$  $\mathbf{H}'\mathbf{H}$  is positive semidefinite ( $\mathbf{x}'\mathbf{H}'\mathbf{H}\mathbf{x} = \|\mathbf{H}\mathbf{x}\|_2^2 \geq 0$ ) and has no zero eigenvalues $\Rightarrow \mathbf{H}'\mathbf{H}$  is positive definite, i.e.,  $\mathbf{x}'\mathbf{H}'\mathbf{H}\mathbf{x} > 0 \quad \forall \mathbf{x} \neq \mathbf{0}$  $\Rightarrow \mathbf{H}\mathbf{x} \neq \mathbf{0} \quad \forall \mathbf{x} \neq \mathbf{0}$ , i.e.,  $\mathbf{H}$  has rank  $n$ .

$$\begin{array}{c} \text{---} \\ \boxed{\mathbf{H}'} \\ \boxed{\mathbf{H}} \\ \text{---} \end{array} \mid = \begin{array}{c} \text{---} \\ \text{---} \end{array}$$

 $(\Rightarrow)$  $\mathbf{H}$  has rank  $n$ , i.e.,  $\mathbf{H}\mathbf{x} \neq \mathbf{0} \quad \forall \mathbf{x} \neq \mathbf{0}$  $\Rightarrow \mathbf{x}'\mathbf{H}'\mathbf{H}\mathbf{x} > 0 \quad \forall \mathbf{x} \neq \mathbf{0}$ , i.e.,  $\mathbf{H}'\mathbf{H}$  is positive definite $\Rightarrow \mathbf{H}'\mathbf{H}$  has only positive eigenvalues, i.e.,  $\det(\mathbf{H}'\mathbf{H}) > 0$ .

2. An  $m \times n$  matrix  $\mathbf{H}$ , with  $m \leq n$ , has rank  $m$ , if and only if the  $m \times m$  matrix  $\mathbf{H}\mathbf{H}'$  has rank  $m$  or  $\det(\mathbf{H}\mathbf{H}') \neq 0$ .

$$\begin{array}{c} \text{---} \\ \boxed{\mathbf{H}} \\ \boxed{\mathbf{H}'} \\ \text{---} \end{array} \mid = \begin{array}{c} \text{---} \\ \text{---} \end{array}$$

- For a real  $m \times n$  matrix  $\mathbf{H}$ ,  
 $\mathbf{HH}'$  has  $m$  eigenvalues,  $\mathbf{H}'\mathbf{H}$  has  $n$  eigenvalues, and

$$\det(s\mathbf{I}_m - \mathbf{HH}') = s^{m-n} \det(s\mathbf{I}_n - \mathbf{H}'\mathbf{H})$$

- $\mathbf{HH}'$  and  $\mathbf{H}'\mathbf{H}$  have the same nonzero eigenvalues, and at most  $\min\{m,n\}$  nonzero eigenvalues.

- If  $\mathbf{HH}'$  and  $\mathbf{H}'\mathbf{H}$  are positive semidefinite

Then  $\mathbf{HH}'$  and  $\mathbf{H}'\mathbf{H}$  have all nonnegative eigenvalues, and at most  $\min\{m,n\}$  positive eigenvalues.

### Singular-Value Decomposition – 1 (3.10)

$$\mathbf{H} = \begin{bmatrix} -4 & -1 & 2 \\ 2 & 0.5 & -1 \end{bmatrix}$$

$$\mathbf{M}_1 = \mathbf{H}^\top \mathbf{H} = \begin{bmatrix} 20 & 5 & -10 \\ 5 & 1.25 & -2.5 \\ -10 & -2.5 & 5 \end{bmatrix}$$

$$\det(\lambda\mathbf{I} - \mathbf{M}_1) = \lambda^2(\lambda - 26.25)$$

$$\lambda_i = 26.25, 0, 0$$

$$\mathbf{M}_2 = \mathbf{H}\mathbf{H}^\top = \begin{bmatrix} 21 & -10.5 \\ -10.5 & 5.25 \end{bmatrix}$$

$$\det(\lambda\mathbf{I} - \mathbf{M}_2) = \lambda(\lambda - 26.25)$$

$$\lambda_i = 26.25, 0$$

- For a **real symmetric** matrix  $\mathbf{M} = \mathbf{H}^T \mathbf{H}$ ,  
there exist an **orthogonal** matrix  $\mathbf{Q}$  and a **diagonal** matrix  $\mathbf{D}$ ,  
such that

$$\mathbf{Q}^T \mathbf{M} \mathbf{Q} = \mathbf{D}$$

$$\mathbf{D} = \begin{bmatrix} \lambda_1^2 & & & \\ & \lambda_2^2 & & \\ & & \ddots & \\ & & & \lambda_n^2 \end{bmatrix} = \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{bmatrix} \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{bmatrix} =: \mathbf{S}^T \mathbf{S}$$

$$\Rightarrow \mathbf{Q}^T \mathbf{H}^T \mathbf{H} \mathbf{Q} = \mathbf{S}^T \mathbf{S}$$

$$\mathbf{M} = \mathbf{Q} \mathbf{D} \mathbf{Q}^T \quad \mathbf{H} = \mathbf{R} \mathbf{S} \mathbf{Q}^T$$

- The **eigenvalues** of  $\mathbf{H}' \mathbf{H}$  can be arranged as

$$\lambda_1^2 \geq \lambda_2^2 \geq \cdots \lambda_r^2 > 0 = \lambda_{r+1}^2 = \cdots = \lambda_n^2$$

Let  $\bar{n} = \min\{m, n\}$ ,

then the **singular values** of  $\mathbf{H}$  are defined as

$$\lambda_1 \geq \lambda_2 \geq \cdots \lambda_r > 0 = \lambda_{r+1} = \cdots = \lambda_{\bar{n}}$$

- There exists an **orthogonal matrix  $\mathbf{Q}$**  such that

$$\mathbf{Q}'\mathbf{H}'\mathbf{H}\mathbf{Q} = \mathbf{D} =: \mathbf{S}'\mathbf{S}$$

**D:**  $n \times n$  diagonal, with  $\lambda_i^2$  on the diagonal

**S:**  $m \times n$  diagonal, with singular values  $\lambda_i$  on the diagonal

### Theorem 3.9: SVD – 1

#### Theorem 3.9 (Singular-value decomposition)

Every  $m \times n$  matrix  $\mathbf{H}$  can be transformed into the form

$$\mathbf{H} = \mathbf{RSQ}'$$

with  $\mathbf{R}'\mathbf{R} = \mathbf{RR}' = \mathbf{I}_m$ ,  $\mathbf{Q}'\mathbf{Q} = \mathbf{QQ}' = \mathbf{I}_n$ , and  $\mathbf{S}$  being  $m \times n$  with the singular values of  $\mathbf{H}$  on the diagonal.

Proof:

$$\mathbf{H}'\mathbf{H} = \mathbf{Q}\mathbf{S}'\mathbf{R}'\mathbf{R}\mathbf{S}\mathbf{Q}' = \mathbf{Q}\mathbf{S}'\mathbf{S}\mathbf{Q}',$$

with  $\mathbf{R}$  satisfying  $\mathbf{H}\mathbf{H}' = \mathbf{R}\mathbf{S}\mathbf{Q}'\mathbf{Q}\mathbf{S}'\mathbf{R}' = \mathbf{R}\mathbf{S}\mathbf{S}'\mathbf{R}'$

$$\mathbf{H}'\mathbf{H} = \mathbf{Q}\mathbf{S}'\mathbf{S}\mathbf{Q}' \text{ or } \mathbf{Q}'\mathbf{H}'\mathbf{H}\mathbf{Q} = \mathbf{S}'\mathbf{S}$$

$$\mathbf{H}\mathbf{H}' = \mathbf{R}\mathbf{S}\mathbf{S}'\mathbf{R}' \text{ or } \mathbf{R}'\mathbf{H}'\mathbf{H}\mathbf{R} = \mathbf{S}\mathbf{S}'$$

### Example 3.14

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

$$\mathbf{AQ} = \mathbf{RS}$$

$$\begin{aligned} \mathbf{a} &= [0 \ 1 \ 1 \ 2; 1 \ 2 \ 3 \ 4; 2 \ 0 \ 2 \ 0]; \\ [\mathbf{r}, \mathbf{s}, \mathbf{q}] &= \text{svd}(\mathbf{a}) \end{aligned}$$

orthonormal basis of  $\mathbf{A}$ 's range space      rank = 2, nullity = 2

$$\mathbf{r} = \begin{bmatrix} 0.3782 & -0.3084 & 0.8729 \\ 0.8877 & -0.1468 & -0.4364 \\ 0.2627 & 0.9399 & 0.2182 \end{bmatrix} \quad \mathbf{s} = \begin{bmatrix} 6.1568 & 0 & 0 & 0 \\ 0 & 2.4686 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

$$\mathbf{q} = \begin{bmatrix} 0.2295 & 0.7020 & 0.3434 & -0.5802 \\ 0.3498 & -0.2439 & 0.8384 & 0.3395 \\ 0.5793 & 0.4581 & -0.3434 & 0.5802 \\ 0.6996 & -0.4877 & -0.2475 & -0.4598 \end{bmatrix}$$

orthonormal basis of  $\mathbf{A}$ 's null space

## (Induced) Norms of Matrices – 1 (3.11)

- (Induced) norms of matrices ( $\mathbf{A}$ : real,  $m \times n$ )

$$\|\mathbf{x}\|_1 := \sum_{i=1}^n |x_i| \quad \|\mathbf{x}\|_2 := \sqrt{\mathbf{x}'\mathbf{x}} = \left( \sum_{i=1}^n |x_i|^2 \right)^{1/2} \quad \|\mathbf{x}\|_\infty := \max_i |x_i|$$

$$\|\mathbf{A}\| = \sup_{\mathbf{x} \neq \mathbf{0}} \frac{\|\mathbf{Ax}\|}{\|\mathbf{x}\|} = \sup_{\|\mathbf{x}\|=1} \|\mathbf{Ax}\|$$

## (Induced) Norms of Matrices – 2

Feng-Li Lian © 2007  
NTUEE-LS1-Matrix-90

$$\|\mathbf{A}\|_1 = \max_j \left( \sum_{i=1}^m |a_{ij}| \right) = \text{largest column absolute sum}$$

$$\|\mathbf{A}\|_2 = \text{largest singular value of } \mathbf{A}$$

$$= (\text{largest eigenvalue of } \mathbf{A}'\mathbf{A})^{1/2}$$

$$\|\mathbf{A}\|_\infty = \max_i \left( \sum_{j=1}^n |a_{ij}| \right) = \text{largest row absolute sum}$$

$$\|\mathbf{Ax}\| \leq \|\mathbf{A}\| \|\mathbf{x}\|$$

$$\|\mathbf{A} + \mathbf{B}\| \leq \|\mathbf{A}\| + \|\mathbf{B}\|$$

$$\|\mathbf{AB}\| \leq \|\mathbf{A}\| \|\mathbf{B}\|$$

- Example:

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

$$\|A\|_1 = 3 + |-1|$$

$$\|A\|_2 = 3.7$$

$$\|A\|_\infty = 3 + 2$$