Lecture – 14

Review of Matrix Theory – II

Dr. Radhakant Padhi
Asst. Professor
Dept. of Aerospace Engineering
Indian Institute of Science - Bangalore

Matrix Transformations

- Question:

  Can a matrix be transformed into a “simplified” form, without losing its properties?

- Answer:

  Yes!

- Options:
  - Similarity transformation
  - Equivalence transformation

Similarity Transformation

Definition: If $A_{n\times n}$ and $B_{n\times n}$ are nonsingular matrices and $P_{n\times n}$ is a non-singular matrix such that $B = P^{-1}AP$, then A and B are "similar".

Simplest forms possible:

- **Diagonal form**

  (if there are *n* linearly independent eigenvectors)
- **Jordan form**

  (if the number of linearly independent eigenvectors are less then *n*)

Similarity Transformation

<u>Steps for finding *P*</u>

- Find the eigenvalues of $A_{n\times n}$
- Find *n* linearly independent eigenvectors V_n
  (include generalized eigenvectors, if necessary)
- Construct the *P* matrix by putting the eigenvectors and generalized vectors as column vectors

Similarity Transformation: Some useful results

If $A = A^T$, then $\exists$ an orthogonal matrix P $(i.e.\ PP^T = P^T P = I)$, whose columns are normalized eigenvectors of A, such that

$$B = P^T A P = diag\left(\lambda_1, \lambda_2, \ldots, \lambda_n\right)$$

where $\lambda_1, \lambda_2, \ldots, \lambda_n$ are eigenvalues of A.

A is similar to a diagonal matrix **if and only if** A has n linearly independent eigenvectors.

If A has n distinct eigenvalues, then A has n linearly independent eigenvectors, and hence, it is similar to a diagonal matrix consisting of its eigenvalues in the diagonal elements.

Similarity Transformation

Having *n* distinct eigenvalues is only a sufficient condition for *A* to be similar to a diagonal matrix.

The necessary condition is to have *n* linearly independent eigenvectors. Having repeated eigenvalues does not guarantee that the eigenvectors are linearly independent (e.g. identity matrix).

If the matrix is of full rank, however, it guarantees that it has *n* linearly independent eigenvectors.

Equivalence Transformation

For a $m \times n$ (non-square) matrix A, a transformation of the form $B = PAQ$, where $P_{m\times m}$ and $Q_{n\times n}$ are non-singular matrices is called an "equivalence transformation".

If $A_{m\times n}$ has rank r, then $\exists\, P, Q:$

$$PAQ = \begin{bmatrix} I_{r,r} & 0_{r,n-r} \\ 0_{m-r,r} & 0_{m-r,n-r} \end{bmatrix}_{m\times n}$$

where $I_{r,r}$ is the $r \times r$ identity matrix.

Singular Value Decomposition (SVD)

- Singular Value Decomposition (SVD) is a special class of equivalence transformation, where *P* and *Q* matrices are restricted to be orthogonal.
- Under an orthogonal equivalence transformation (i.e. in SVD), we can achieve a diagonal matrix.
- Under an orthogonal similarity transformation, however, we can achieve only a triangular matrix (Schur's theorem).

Singular Values

$$\sigma_A \triangleq \sqrt{\lambda\left(A^T A\right)}, \quad \text{if } A \text{ is real}$$

$$\triangleq \sqrt{\lambda\left(A^* A\right)}, \quad \text{if } A \text{ is complex}$$

Both $A^T A$ and $A^* A$ are positive semidefinite, and hence, their eigenvalues are always non-negative.

For singular value computation, only positive square roots need to be found out.

Singular Values

For any real (complex) $A_{m\times n}$ of rank r, $\exists$ orthogonal (unitary) matrices $P_{m\times m}$ and $Q_{n\times n}$:

$$A = PDQ\,, \quad D_{m\times n} = \begin{bmatrix} \sigma_1 & \cdots & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & \sigma_r & 0 \\ 0 & \cdots & 0 & 0 \end{bmatrix}$$

where $\sigma_1,\ldots,\sigma_r$ are singular values of A.

How to find matrices *P* and *Q* ?

- Let $\lambda_1,\ldots,\lambda_n$ and $X_1,\ldots,X_n$ be the eigenvalues and corresponding "orthonormal eigenvectors" of $A^T A$
- Construct $\sigma_i\left(A\right)=\sqrt{\lambda_i}\,, \qquad \left(i=1,\ldots,n\right)$
- Order $\sigma_i\,'s$ such that

$$\sigma_i > 0, \quad i=1,\ldots,r$$

$$\sigma_i = 0, \quad i=r+1,\ldots,n$$

How to find matrices *P* and *Q* ?

- Construct $Y_i = \frac{1}{\sigma_i}\left(AX_i\right), \quad i = 1,\ldots,r$
- Extend $Y_1,\ldots,Y_r$ to an orthonormal basis

$$Y_1,\ldots,Y_r,\ Y_{r+1},\ldots,Y_m$$

- Then

$$P = \begin{bmatrix} Y_1 & \cdots & Y_m \\ \downarrow & \downarrow & \downarrow \end{bmatrix} \qquad Q = \begin{bmatrix} X_1 & \cdots & X_n \\ \downarrow & \downarrow & \downarrow \end{bmatrix}^T$$

Example

$$A = \begin{bmatrix} 1 & 0 & -1 \\ -1 & 0 & 1 \end{bmatrix}, \quad A^T A = \begin{bmatrix} 2 & 0 & -2 \\ 0 & 0 & 0 \\ -2 & 0 & 2 \end{bmatrix}$$

$$\lambda_{1,2,3}\left(A^T A\right) = 4, 0, 0$$

$$X_1 = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}, \quad X_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \quad X_3 = \frac{1}{\sqrt{2}}\begin{bmatrix} -1 \\ 0 \\ -1 \end{bmatrix}$$

$$\left(for \quad \lambda = 4 \right) \quad \left(for \quad \lambda = 0 \right) \quad \left(for \quad \lambda = 0 \right)$$

Example

$$\begin{bmatrix} \sigma_1 \\ \sigma_2 \\ \sigma_3 \end{bmatrix} = \begin{bmatrix} \sqrt{\lambda_1} \\ \sqrt{\lambda_2} \\ \sqrt{\lambda_3} \end{bmatrix} = \begin{bmatrix} 2 \\ 0 \\ 0 \end{bmatrix} \quad \left(\text{Note: The values are ordered}\right)$$

$$Y_1 = \frac{1}{\sigma_1}\left(AX_1\right) = \frac{1}{2}\begin{bmatrix} 1 & 0 & -1 \\ -1 & 0 & 1 \end{bmatrix} \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

Next, find $Y_2 = \begin{bmatrix} a & b \end{bmatrix}^{\mathrm{T}}$ such that Y_1 and Y_2 will be orthonormal

i.e. $\left\langle Y_1, Y_2 \right\rangle = Y_1^T Y_2 = a - b = 0$ and $\left\langle Y_2, Y_2 \right\rangle = Y_2^T Y_2 = a^2 + b^2 = 1$

Example

Solution: $a = b = 1/\sqrt{2}.$ Hence, $Y_2 = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ 1 \end{bmatrix}$

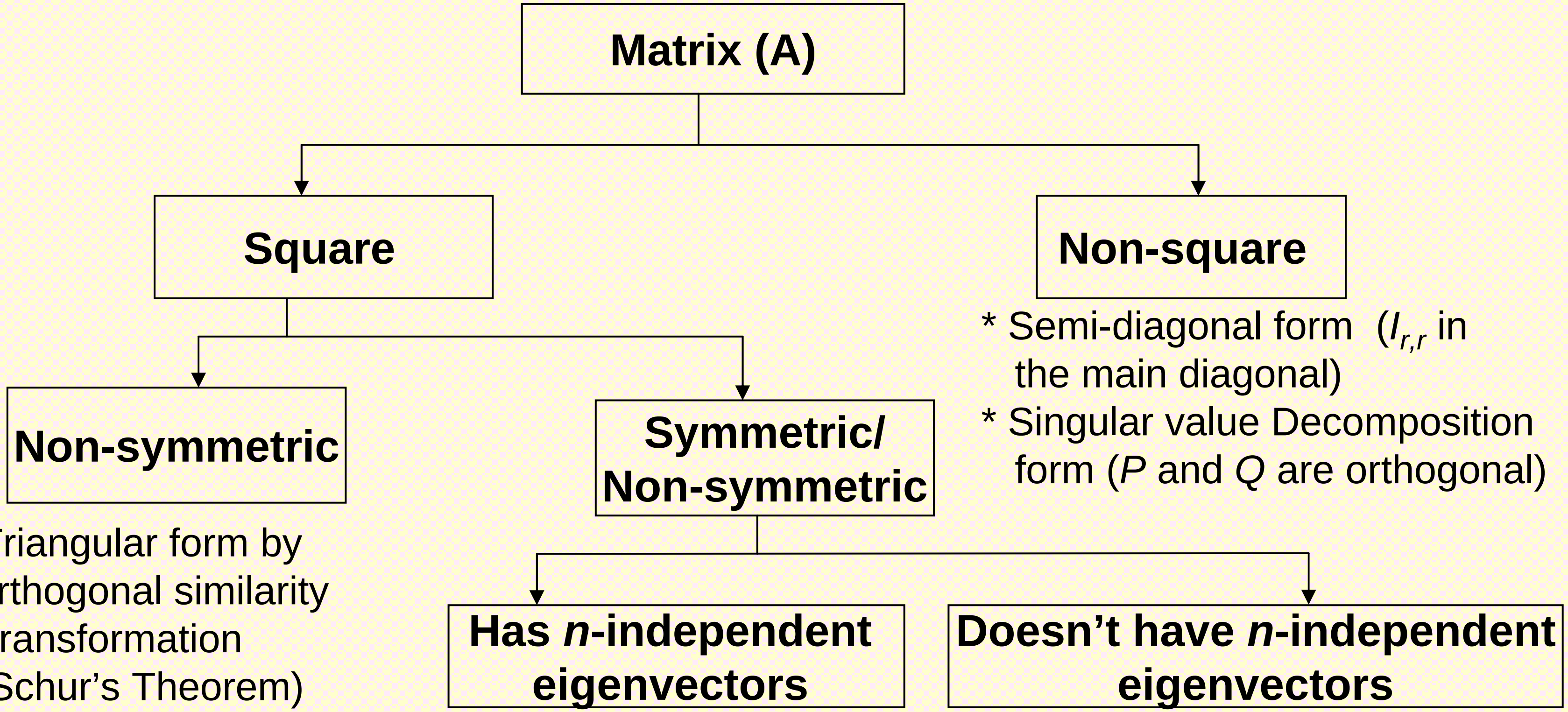
$$P = \begin{bmatrix} Y_1 & Y_2 \\ \downarrow & \downarrow \end{bmatrix} = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}$$

$$Q = \begin{bmatrix} X_1 & X_2 & X_3 \\ \downarrow & \downarrow & \downarrow \end{bmatrix}^T = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 0 & -1 \\ 0 & \sqrt{2} & 0 \\ -1 & 0 & -1 \end{bmatrix}$$

$$D = \begin{bmatrix} \sigma_1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

$$Verify: \quad PDQ = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}\begin{bmatrix} 2 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}\frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 0 & -1 \\ 0 & \sqrt{2} & 0 \\ -1 & 0 & -1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & -1 \\ -1 & 0 & 1 \end{bmatrix} = A$$

Summary of Transformations



Non-square:

* Semi-diagonal form ($I_{r,r}$ in the main diagonal)
* Singular value Decomposition form (*P* and *Q* are orthogonal)

Non-symmetric:

* Triangular form by orthogonal similarity Transformation (Schur's Theorem)
* Diagonal form by orthogonal equivalence Transformation

Has *n*-independent eigenvectors:

Diagonal form by similarity Transformation

Doesn't have *n*-independent eigenvectors:

Jordan form by similarity Tranformation

Vector/Matrix Calculus: Definitions

$$X(t) \triangleq \begin{bmatrix} x_1(t) & x_2(t) & \cdots & x_n(t) \end{bmatrix}^T$$

$$\dot{X}(t) \triangleq \begin{bmatrix} \dot{x}_1(t) & \dot{x}_2(t) & \cdots & \dot{x}_n(t) \end{bmatrix}^T$$

$$\int_0^t X(\tau)\,d\tau \triangleq \begin{bmatrix} \int_0^t x_1(\tau)\,d\tau & \int_0^t x_2(\tau)\,d\tau & \cdots & \int_0^t x_n(\tau)\,d\tau \end{bmatrix}^T$$

Vector/Matrix Calculus: Definitions

$$A(t) = \begin{bmatrix} a_{11}(t) & \cdots & a_{1n}(t) \\ \vdots & \ddots & \vdots \\ a_{m1}(t) & \cdots & a_{mn}(t) \end{bmatrix} \qquad \dot{A}(t) \triangleq \begin{bmatrix} \dot{a}_{11}(t) & \cdots & \dot{a}_{1n}(t) \\ \vdots & \ddots & \vdots \\ \dot{a}_{m1}(t) & \cdots & \dot{a}_{mn}(t) \end{bmatrix}$$

$$\int_0^t A(\tau)d\tau \triangleq \begin{bmatrix} \int_0^t a_{11}(\tau)d\tau & \cdots & \int_0^t a_{1n}(\tau)d\tau \\ \vdots & \ddots & \vdots \\ \int_0^t a_{m1}(\tau)d\tau & \cdots & \int_0^t a_{mn}(\tau)d\tau \end{bmatrix}$$

Vector/Matrix Calculus: Some Useful Results

$$\frac{d}{dt}\left(A(t)+B(t)\right)=\dot{A}(t)+\dot{B}(t)$$

$$\frac{d}{dt}\left(A(t)B(t)\right)=\dot{A}(t)B(t)+A(t)\dot{B}(t)$$

$$\frac{d}{dt}\left(A^{-1}(t)\right)=-A^{-1}(t)\frac{dA(t)}{dt}A^{-1}(t)$$

Vector/Matrix Calculus: Definitions

If $f(X) \in \mathbb{R}$, then $(\partial f / \partial X) \triangleq [\partial f / \partial x_1 \quad \cdots \quad \partial f / \partial x_n]^T$

is called the "gradient" of $f(X)$.

If $f(X) \triangleq [f_1(X) \quad \cdots \quad f_m(X)]^T \in \mathbb{R}^m$, then

$$\frac{\partial f}{\partial X} \triangleq \begin{bmatrix} \partial f_1 / \partial x_1 & \cdots & \partial f_1 / \partial x_n \\ \vdots & \ddots & \vdots \\ \partial f_m / \partial x_1 & \cdots & \partial f_m / \partial x_n \end{bmatrix}$$

is called the "Jacobian matrix" of $f(X)$ with respect to X.

Vector/Matrix Calculus: Definitions

If $f(X) \in \mathbb{R}$, then

$$\frac{\partial^2 f}{\partial X^2} \triangleq \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n} \\ \frac{\partial^2 f}{\partial x_2 \partial x_1} & \ddots & & \frac{\partial^2 f}{\partial x_2 \partial x_n} \\ \vdots & & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1} & \frac{\partial^2 f}{\partial x_n \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_n^2} \end{bmatrix}$$

is called the "Hessian matrix" of $f(X)$.

Vector/Matrix Calculus: Derivative Rules

$$\frac{\partial}{\partial X}\left(b^T X\right) = \frac{\partial}{\partial X}\left(X^T b\right) = b$$

$$\frac{\partial}{\partial X}\left(AX\right) = A$$

$$\frac{\partial}{\partial X}\left(X^T A X\right) = \left(A + A^T\right) X$$

If $A = A^T$, $\quad \frac{\partial}{\partial X}\left(\frac{1}{2} X^T A X\right) = AX$

Vector/Matrix Calculus: Derivative Rules

$$\frac{\partial}{\partial X}\left(f^T(X)\, g(X)\right) = \left[\frac{\partial f}{\partial X}\right]^T g(X) + \left[\frac{\partial g}{\partial X}\right]^T f(X)$$

Corollary :

$$\frac{\partial}{\partial X}\left(C^T g(X)\right) = \left[\frac{\partial g}{\partial X}\right]^T C, \quad \frac{\partial}{\partial X}\left(f^T(X)\, C\right) = \left[\frac{\partial f}{\partial X}\right]^T C$$

$$\frac{\partial}{\partial X}\left(f(X)\, Q\, g(X)\right) = \left[\frac{\partial f}{\partial X}\right]^T Q\, g(X) + \left[\frac{\partial g}{\partial X}\right]^T Q^T f(X)$$

Vector/Matrix Calculus: Derivative Rules

If $G(X) \in \mathbb{R}^{p \times m}, \quad X \in \mathbb{R}^n, \quad U \in \mathbb{R}^m$

$$\frac{\partial}{\partial X}\big(G(X)\,U\big) = \left[\frac{\partial G_1}{\partial X}\right]u_1 + \left[\frac{\partial G_2}{\partial X}\right]u_2 + \cdots + \left[\frac{\partial G_m}{\partial X}\right]u_m$$

where $G \triangleq \begin{bmatrix} G_1 & G_2 & \cdots & G_m \\ \downarrow & \downarrow & & \downarrow \end{bmatrix}$

If $f(X) \in \mathbb{R}, \quad g(X) \in \mathbb{R}^{1 \times m}, \quad X \in \mathbb{R}^n, \quad U \in \mathbb{R}^m$

$$\frac{\partial}{\partial X}\left[f(X)g(X)\,U\right] = \left[f\left[\frac{\partial g}{\partial X}\right] + \left[\frac{\partial f}{\partial X}\right]g\right]U$$

Vector/Matrix Calculus: Chain Rules

If $F\left(f(X)\right)\in\mathbb{R},\quad f(X)\in\mathbb{R},\quad X\in\mathbb{R}^n$

$$\left[\frac{\partial F}{\partial X}\right]_{n\times 1}=\left[\frac{\partial f}{\partial X}\right]_{n\times 1}\left[\frac{\partial F}{\partial f}\right]_{1\times 1}$$

If $F\left(f(X)\right)\in\mathbb{R},\quad f(X)\in\mathbb{R}^m,\quad X\in\mathbb{R}^n$

$$\left[\frac{\partial F}{\partial X}\right]_{n\times 1}=\left[\frac{\partial f}{\partial X}\right]_{n\times m}^T\left[\frac{\partial F}{\partial f}\right]_{m\times 1}$$

If $F\left(f(X)\right)\in\mathbb{R}^p,\quad f(X)\in\mathbb{R}^m,\quad X\in\mathbb{R}^n$

$$\left[\frac{\partial F}{\partial X}\right]_{p\times n}=\left[\frac{\partial F}{\partial f}\right]_{p\times m}^T\left[\frac{\partial f}{\partial X}\right]_{m\times n}$$

Thanks for the Attention….!