Lecture 18 : Orthogonal Matrix & Diagonalization of real symmetric matrices

We have seen that if a square matrix has distinct eigen values then it is diagonalizable. Here, we will see that a real symmetric matrix is diagonalizable. First we prove the following lemma

Lemma : Let A be an $n \times n$ matrix. Then the following conditions are equivalent.

1. The column vectors are orthonormal.
2. $A^T A = A A^T = I_n$ (so $A^{-1} = A^T$)
3. $\|Ax\| = \|x\| \;\forall\, x \in \mathbb{R}^n$
4. $\langle Ax, Ay \rangle = \langle x, y \rangle \;\forall\, x, y \in \mathbb{R}^n$
5. The row vectors are orthonormal.

Proof:- Let $A = (a_{ij})$ $1 \le i \le n$, $1 \le j \le n$.

Column vectors $v_j = (a_{1j}, a_{2j}, \dots, a_{nj})$

& row vectors $w_i = (a_{i1}, a_{i2}, \dots, a_{in})$

$1 \Rightarrow 2$:- $\langle v_p, v_q \rangle = 0$ $p \neq q$

& $\|v_p\| = 1$

Hence $\sum_{l=1}^{n} a_{lp} a_{lq} = 0$ for $p \neq q$

& $\sum_{l=1}^{n} a_{lp}^2 = 1$

p th row vector of A^T is $(a_{1p}, a_{2p}, \dots, a_{np})$
& q th column vector of A is $(a_{1q}, a_{2q}, \dots, a_{nq})$
pq th entry of $A^T A$ is dot product of $(a_{1p}, a_{2p}, \dots, a_{np})$ with $(a_{1q}, a_{2q}, \dots, a_{nq})$

So, $A^T A = I_n$

$\Rightarrow$ A^T is invertible & $(A^T)^{-1} = A$

& $A A^T = I_n$, So $AA^T = A^T A = I_n$.

(2) $\Rightarrow$ (3) :- Let $e_i = (0, \dots, \underset{i\text{th position}}{1}, \dots 0)$

$\{e_1, e_2, \dots, e_n\}$ is a basis of $\mathbb{R}^n$

Let $x = (x_1, \dots, x_n) \in \mathbb{R}^n$ then

$$x = (x_1, \dots x_n) = \sum_{i=1}^{n} x_i e_i$$

$$A x = \sum_{i=1}^{n} x_i A e_i$$

Ae_i is the ith column vector $(a_{1i}, a_{2i}, \dots, a_{ni})$ of A

$$\| A e_i \| = \| (a_{1i}, a_{2i}, \dots, a_{ni}) \| = \sqrt{a_{1i}^2 + \dots + a_{ni}^2}$$
$= \sqrt{1}$ (as iith element
$= 1$ of $A^T A$ is 1)

$$\|Ax\|^2 = \langle Ax, Ax\rangle$$
$$= \langle \sum_{i=1}^{n} x_i Ae_i, \sum_{i=1}^{n} x_i Ae_i\rangle$$
$$= \sum_{i,j=1}^{n} x_i x_j \langle Ae_i, Ae_j\rangle$$
$$= \sum_{i=1}^{n} x_i^2$$ (Since for $i \neq j$ $\langle Ae_i, Ae_j\rangle$ = ijth entry of A^TA which is 0 as $A^TA = I$)
$$= \|x\|^2$$
$\Rightarrow \|Ax\| = \|x\| \ \forall x \in \mathbb{R}^n$.

(3) $\Rightarrow$ (4) $\|A(x+y)\| = \|x+y\|$

$\Rightarrow \langle A(x+y), A(x+y)\rangle = \langle x+y, x+y\rangle$

$\Rightarrow \langle Ax+Ay, Ax+Ay\rangle = \langle x,x\rangle + \langle x,y\rangle + \langle y,x\rangle + \langle y,y\rangle$

$\Rightarrow \|Ax\|^2 + 2\langle Ax, Ay\rangle + \|Ay\|^2$
$= \|x\|^2 + 2\langle x,y\rangle + \|y\|^2$

$\Rightarrow \langle Ax, Ay\rangle = \langle x,y\rangle$ as $\|Ax\| = \|x\|$ & $\|Ay\| = \|y\|$.

(4) $\Rightarrow$ (5) Suppose $\langle Ax, Ay\rangle = \langle x, y\rangle \ \forall x, y \in \mathbb{R}^n$

consider the orthonormal basis $\{e_1, \dots, e_n\}$ of $\mathbb{R}^n$ then for $i \neq j$ $\langle Ae_i, Ae_j\rangle = \langle e_i, e_j\rangle = 0$

& $\langle Ae_i, Ae_i\rangle = \langle e_i, e_i\rangle = 1$

Let $A = (a_{ij})$ then Ae_i is the ith column vector $\begin{pmatrix} a_{1i} \\ a_{2i} \\ \vdots \\ a_{ni} \end{pmatrix}$ of A.

or ith row vector of A^T.

$\langle Ae_i, Ae_j\rangle = 0$ for $i \neq j$, $\langle Ae_i, Ae_i\rangle = 1$

implies $A^TA = I \Rightarrow AA^T = I$

$\Rightarrow$ row vectors of A are orthonormal

(5) $\Rightarrow$ (1) Suppose row vectors of A are orthonormal. This implies $AA^T = I$

$\Rightarrow$ A is invertible & $A^TA = I$

$A^TA = I \Rightarrow$ column vectors are orthogonal

FACT: Every non-constant polynomial over $\mathbb{C}$ has a root in $\mathbb{C}$ i.e. if $p(x) = a_0 + a_1 x + \cdots + a_n x^n$ is a polynomial of degree n $(n \geq 1)$ then $\exists\, \alpha \in \mathbb{C}$ s.t $p(\alpha) = 0$. Infact, for $p(x)$ $\exists\, \alpha_1, \ldots, \alpha_n$ in $\mathbb{C}$ such that $p(\alpha_1) = \cdots = p(\alpha_n) = 0$

We will use this fact to prove that eigen values of real symmetric matrix is real

Theorem:- Let A be a real symmetric matrix of order $n \times n$. Then all eigen values of A are real.

Proof:- Eigen values are roots of the characteristic polynomial $\det(A - xI)$.

Let $p(x) = \det(A - xI)$. Then by the fact above $\exists\, \lambda \in \mathbb{C}$ such that $p(\lambda) = 0$

i.e $\det(A - \lambda I) = 0$

Let $u \in \mathbb{C}^n$ be a non-zero eigen vector of A corresponding to eigen value λ

Then $Au = \lambda u$.

$$\Rightarrow (Au)^T = \lambda u^T$$

$$\Rightarrow u^T A = \lambda u^T \quad (\text{since } A^T = A)$$

Take complex conjugate both sides, then

$\bar{u}^T \bar{A} = \bar{\lambda} \bar{u}^T$ if $u = (u_1, \ldots, u_n)$ where $u_i \in \mathbb{C}$

then $\bar{u} = (\bar{u}_1, \ldots, \bar{u}_n)$

& if $A = (a_{ij})$,

$\bar{A} = (\bar{a}_{ij})$.

Now, $\bar{u}^T A u = \bar{\lambda} \bar{u}^T u$

$$\Rightarrow \bar{u}^T (\lambda u) = \bar{\lambda} u^T u$$

$$\Rightarrow \lambda \bar{u}^T u = \bar{\lambda} u^T u$$

$$\Rightarrow (\lambda - \bar{\lambda}) \bar{u}^T u = 0$$

$\bar{u}^T u$ is a non-zero number

Hence $\lambda = \bar{\lambda} \Rightarrow \lambda$ is real. $\boxtimes$

Lemma: Let A be a real matrix with order $n \times n$ and suppose there exists a real eigen value λ of A. Then $\exists$ a real non-zero eigen vectors of λ.

<u>Proof:</u> Consider the system of linear equation $(A - \lambda I)x = 0$.

det $(A - \lambda I) = 0$ implies it has a non-zero solution $u \in \mathbb{R}^n$.

$$(A - \lambda I)u = 0 \Rightarrow Au = \lambda u$$

u is an eigen vector corresponding to λ. $\square$

<u>Orthogonal Matrix:</u> A square matrix A is said to be orthogonal if $AA^T = A^TA = I$.

<u>Theorem:</u> Let A be a real symmetric matrix of order $n \times n$. Then there exists a real orthogonal matrix P such that $P^{-1}AP$ is a diagonal matrix.

<u>Proof:-</u> Let λ be an eigen value of A.

Then by previous theorem, λ is real.

By lemma, there exists a non-zero real vector u of A corresponding to λ.

By Gram-Schmidt orthogonalization process, $\exists$ an orthonormal basis $\{u, f_2, \ldots, f_n\}$ of $\mathbb{R}^n$ containing u.

Let P_1 be the matrix whose column vectors are $u, f_2, \ldots, f_n$. As column vectors are orthonormal therefore the matrix $P_1 \equiv (u \; f_2 \; \cdots \; f_n)$ is orthogonal.

Consider the matrix $P_1 A P_1$ and let $e_1 = (1, 0, \ldots 0)$.

The first column of $P_1^T A P_1$ is given by

$$
\begin{aligned}
(P_1^{-1} A P_1) e_1 = (P^{-1} A)(P_1 e_1) &= P_1^{-1} A u \\
&= P_1^{-1} \lambda u \\
&= \lambda P_1^{-1} u \\
&= \lambda e_1 \quad \text{since, } P_1 e_1 = u
\end{aligned}
$$

Note that $P_1^T = P_1^{-1}$ & $P_1^{-1} A P_1$ is also symmetric as

$$
\begin{aligned}
(P_1^{-1} A P_1)^T &= (P_1^T A P_1)^T \\
&= P_1^T A^T (P_1^T)^T \\
&= P_1^T A P_1
\end{aligned}
$$

Thus, first row of $P_1^{-1} A P_1$ is also λe_1

So, $P_1^{-1} A P_1$ is of the form

$$
\begin{pmatrix} \lambda & 0 \\ 0 & A_1 \end{pmatrix}
$$

where A_1 is real symmetric matrix of order $(n-1) \times (n-1)$

The proof of theorem is by induction on n.
If $n = 1$, it is obvious.
Let the statement be true for all real Symmetric matrices of order $(n-1)\times(n-1)$
Then $\exists$ a real orthogonal matrix P_2 of order $(n-1)\times(n-1)$ such that
$P_2^{-1} A_1 P_2$ is a diagonal matrix D.

Let $P = P_1 \begin{pmatrix} 1 & 0 \\ 0 & P_2 \end{pmatrix}$

Check that P is orthogonal and

$$P^{-1} A P = \begin{pmatrix} \lambda & 0 \\ 0 & D \end{pmatrix} \qquad \square$$