MTH 100 : Lecture 37

Diagonalization of Symmetric Matrices

Definition: A matrix A is said to be symmetric if $A = A^T$.

A symmetric matrix is necessarily square.

- For the time being we restrict ourselves to matrices and vectors with real entries.

Proposition: If A is symmetric, then any two eigenvectors from different eigenspaces (i.e. eigenvectors corresponding to different eigenvalues) are orthogonal.

Note: Earlier, we have shown that for any square matrix, eigenvectors from different eigenspaces (i.e. corresponding to different eigenvalues) are linearly independent.

For symmetric matrices, we have the stronger result above.

Proof: Let u_1 and u_2 be eigenvectors corresponding to different eigenvalues λ_1 and λ_2 respectively (for matrix A)

$$\lambda_1 \langle u_1, u_2 \rangle = \langle \lambda_1 u_1, u_2 \rangle$$

$$= \langle A u_1, u_2 \rangle = (A u_1)^T u_2$$

$$= (u_1^T A^T) u_2$$

$$= (u_1^T A) u_2 \quad \text{(since } A \text{ is symmetric)}$$

$$= u_1^T (A u_2)$$

$$= u_1^T (\lambda_2 u_2)$$

$$= \langle u_1, \lambda_2 u_2 \rangle$$

$$= \lambda_2 \langle u_1, u_2 \rangle$$

$$\Rightarrow (\lambda_1 - \lambda_2) \langle u_1, u_2 \rangle = 0 \Rightarrow \langle u_1, u_2 \rangle = 0 \quad (\text{since } \lambda_1 - \lambda_2 \neq 0)$$

So, $u_1 \perp u_2$

Definition: A square matrix P is said to be $\boxed{\text{orthogonal}}$ if its columns are orthonormal.

(Please note this slight inconsistency in terminology)

Proposition: An orthogonal matrix P is necessarily invertible and $P^{-1} = P^T$

Proof: Let $P = [v_1 \; v_2 \; \dots \; v_n]$ where v_i's are the orthonormal column vectors.

$$P^T P = \begin{bmatrix} v_1^T \\ v_2^T \\ \vdots \\ v_n^T \end{bmatrix} [v_1 \; v_2 \; \dots \; v_n]$$

$$= \begin{bmatrix} v_1^T v_1 & v_1^T v_2 & \dots & v_1^T v_n \\ v_2^T v_1 & v_2^T v_2 & \dots & v_2^T v_n \\ \dots & \dots & \dots & \dots \\ v_n^T v_1 & v_n^T v_2 & \dots & v_n^T v_n \end{bmatrix}$$

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

(since $v_i^T v_j = \delta_{ij}$ (kronecker delta)
$\delta_{ij} = 1$ if $i = j$
$= 0$ if $i \neq j$)

Therefore $P^T P = I$

$\boxed{\text{QED}}$

Definition: A square matrix A is said to be orthogonally diagonalizable if there is an orthogonal matrix P and a diagonal matrix D, such that $A = PDP^{-1} = PDP^T$
(ie. $AP = PD$)

Note: For an $n \times n$ matrix to be orthogonally diagonalizable, it should have n linearly independent and orthonormal eigenvectors. That happens only in the following case.

Proposition: If an $n \times n$ matrix A is orthogonally diagonalizable, then A is symmetric.

Proof: If A is orthogonally diagonalizable, then $A = PDP^{-1} = PDP^T$ where P is an orthogonal matrix.

Now $A^T = (PDP^T)^T = (P^T)^T D^T P^T = PDP^T = A$
(Since D is a diagonal matrix)

$\Rightarrow A^T = A$ and so A is symmetric.

Definition: The set of eigenvalues of a matrix A is called the spectrum of A.

Theorem (Spectral Theorem for Symmetric Matrices):

An $n \times n$ symmetric matrix A has the following properties.

(a) The eigenspaces are mutually orthogonal.
(i.e. eigenvectors corresponding to different eigenvalues are orthogonal)

(b) A has n real eigenvalues, counting algebraic multiplicities.

(c) A is orthogonally diagonalizable.

(d) The dimension of the eigenspace for each eigenvalue λ equals the algebraic multiplicity of λ (as a root of the characteristic equation),
i.e. the geometric multiplicity is equal to the algebraic multiplicity.

Remarks:

(1) The proof of (a) is given in a previous proposition.

(2) The proof of (b) is an exercise.

(3) The proof of (c) is nontrivial and will be omitted.

(4) The statement (d) follows from the statement (c) using the Diagonalization Theorem.

Corollary: Taking statement (c) and previous proposition we have:

A is orthogonally diagonalizable if and only if A is symmetric.

The Spectral Theorem in Practice:

- In numerical examples, we first factorize the characteristic polynomial. We will always get as many real roots (counting multiplicities) as the dimension of the matrix, i.e. complex roots will not occur.

- While row reducing the matrix $(A-\lambda I)$ for any eigenvalue λ to solve the associated homogeneous system, we get as many free variables as the algebraic multiplicity of λ. Thus we get the desired number of basis vectors.
- For each eigenspace of dimension greater than one, we obtain an orthogonal basis by using the Gram-Schmidt process.
- Finally we normalize all the basis vectors.

<u>Ex:</u> <u>Diagonalization of a symmetric Matrix</u>

Given $A = \begin{bmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{bmatrix}$ we have to diagonalize A.

The characteristic Polynomial $= \det(A-\lambda I)$

$$= \begin{vmatrix} 2-\lambda & 1 & 1 \\ 1 & 2-\lambda & 1 \\ 1 & 1 & 2-\lambda \end{vmatrix}$$

$$= (2-\lambda)\{(2-\lambda)^2-1\} + 1\{1\times 1 - 1(2-\lambda)\} + 1\{1\times 1 - 1(2-\lambda)\}$$

$$= (2-\lambda)(\lambda^2-4\lambda+3) + (\lambda-1) + (\lambda-1)$$

$$= (2-\lambda)(\lambda-1)(\lambda-3) + (\lambda-1) + (\lambda-1)$$

$$= (\lambda-1)\left[(2\lambda-6-\lambda^2+3\lambda) + 1 + 1\right]$$

$$= (\lambda-1)\left(-\lambda^2+5\lambda-4\right) = -(\lambda-1)(\lambda^2-5\lambda+4)$$
$$= -(\lambda-1)(\lambda-4)(\lambda-1) = -(\lambda-1)^2(\lambda-4)$$

So, the eigenvalues are $\lambda_1 = 4$ and $\lambda_2 = 1$

Now for $\lambda_1 = 4$:

$$A - \lambda_1 I = \begin{bmatrix} 2-4 & 1 & 1 \\ 1 & 2-4 & 1 \\ 1 & 1 & 2-4 \end{bmatrix} = \begin{bmatrix} -2 & 1 & 1 \\ 1 & -2 & 1 \\ 1 & 1 & -2 \end{bmatrix} \xrightarrow{R_1 \leftrightarrow R_3} \begin{bmatrix} 1 & 1 & -2 \\ 1 & -2 & 1 \\ -2 & 1 & 1 \end{bmatrix}$$

$$\xrightarrow[R_3 \to R_3 + 2R_1]{R_2 \to R_2 - R_1} \begin{bmatrix} 1 & 1 & -2 \\ 0 & -3 & 3 \\ 0 & 3 & -3 \end{bmatrix} \xrightarrow{R_3 \to R_3 + R_2} \begin{bmatrix} 1 & 1 & -2 \\ 0 & -3 & 3 \\ 0 & 0 & 0 \end{bmatrix}$$

$$\xrightarrow{R_2 \to -\frac{1}{3}R_2} \begin{bmatrix} 1 & 1 & -2 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{bmatrix} \xrightarrow{R_1 \to R_1 - R_2} \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{bmatrix} = \text{RREF matrix}$$

So, the system of equation reduces to

$$\left.\begin{array}{l} x_1 - x_3 = 0 \\ x_2 - x_3 = 0 \end{array}\right\} \Rightarrow \left.\begin{array}{l} x_1 = x_3 \\ x_2 = x_3 \\ x_3 = x_3 \end{array}\right\} \Rightarrow \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = x_3 \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

By taking $x_3 = 1$, we get $u_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ corresponding to $\lambda_1 = 4$

Normalising, we get $v_1 = \frac{1}{\sqrt{3}}\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$

For $\lambda_2 = 1$: $(A - \lambda_2 I) = \begin{bmatrix} 2-1 & 1 & 1 \\ 1 & 2-1 & 1 \\ 1 & 1 & 2-1 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$

$$\text{RREF matrix} = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \longleftarrow \begin{matrix} R_2 \to R_2 - R_1 \\ R_3 \to R_3 - R_1 \end{matrix}$$

So, the system of equation reduces to

$$\left.\begin{aligned} x_1 + x_2 + x_3 = 0 \\ x_2 = x_2 \\ x_3 = x_3 \end{aligned}\right\} \Rightarrow \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = x_2 \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$$

Taking $x_2 = 1, x_3 = 0$ and $x_2 = 0, x_3 = 1$ we get

$$u_2 = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} \quad \text{and} \quad u_3 = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$$

Note that $\langle u_1, u_2 \rangle = 0$, $\langle u_1, u_3 \rangle = 0$

but $\langle u_2, u_3 \rangle = (-1)(-1) = 1 \neq 0$

Therefore we need to apply Gram-Schmidt process to W = eigenspace corresponding to $\lambda_2 = 1$

Let $w_2 = u_2 = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}$

Then $w_3 = u_3 - \dfrac{\langle u_3, u_2 \rangle}{\langle w_2, w_2 \rangle} w_2$

$$= \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} - \frac{1}{2} \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} -\frac{1}{2} \\ -\frac{1}{2} \\ 1 \end{bmatrix}$$

Note that $\langle w_2, w_3 \rangle = (-1)\left(-\frac{1}{2}\right) + 1\left(-\frac{1}{2}\right) = \frac{1}{2} - \frac{1}{2} = 0$

Now we normalise w_2 and w_3 to get

$$v_2 = \frac{1}{\sqrt{2}}\begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} \quad \text{and} \quad v_3 = \frac{1}{\sqrt{\left(\frac{1}{2}\right)^2+\left(\frac{1}{2}\right)^2+1^2}}\begin{bmatrix} -\frac{1}{2} \\ -\frac{1}{2} \\ 1 \end{bmatrix}$$

$$= \frac{2}{\sqrt{6}}\begin{bmatrix} -\frac{1}{2} \\ -\frac{1}{2} \\ 1 \end{bmatrix} = \frac{1}{\sqrt{6}}\begin{bmatrix} -1 \\ -1 \\ 2 \end{bmatrix}$$

So desired P and D will be

$$P = \begin{bmatrix} \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & 0 & \frac{2}{\sqrt{6}} \end{bmatrix} \quad \text{and} \quad D = \begin{bmatrix} 4 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Now we calculate AP and PD:

$$AP = \begin{bmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{bmatrix}\begin{bmatrix} \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & 0 & \frac{2}{\sqrt{6}} \end{bmatrix} = \begin{bmatrix} \frac{4}{\sqrt{3}} & -\frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} \\ \frac{4}{\sqrt{3}} & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} \\ \frac{4}{\sqrt{3}} & 0 & \frac{2}{\sqrt{6}} \end{bmatrix}$$

$$\text{and} \quad PD = \begin{bmatrix} \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & 0 & \frac{2}{\sqrt{6}} \end{bmatrix}\begin{bmatrix} 4 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} \frac{4}{\sqrt{3}} & -\frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} \\ \frac{4}{\sqrt{3}} & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} \\ \frac{4}{\sqrt{3}} & 0 & \frac{2}{\sqrt{6}} \end{bmatrix}$$

Therefore $AP = PD \Rightarrow \boxed{A = PDP^{-1} = PDP^{T}}$