Lecture 15 (Diagonalizability)

Definition 1. Let $A \in M_n(\mathbb{R})$ with the characteristic polynomial $f(x)$. Let λ be an eigenvalue of A then the largest power k such that $(x-\lambda)^k$ is a factor of $f(x)$ is called the algebraic multiplicity of λ (A.M.(λ)).

Theorem 2. Let λ be an eigenvalue of a matrix A. Then the set $E_\lambda = \{x \in \mathbb{F}^n \mid Ax = \lambda x\}$ forms a subspace of $\mathbb{F}^n$ and it is called eigenspace corresponding to the eigenvalue λ. Observe that E_λ is the set of all eigenvectors associated to λ including the zero vector.

Definition 3. The dimension of the eigenspace (E_λ) of eigenvalue λ is called the geometric multiplicity of λ (G.M.(λ)). Thus the geometric multiplicity of λ, G.M.(λ) = Nullity $(A-\lambda I) = n-$ Rank $(A-\lambda I)$.

Remark 4. *1. Thus the geometric multiplicity of λ, G.M.(λ) = Nullity $(A-\lambda I) = n-$ Rank $(A-\lambda I)$.*
2. G.M.(λ) ≥ 1.

Theorem 5. G.M. (λ) $\leq$ A.M. (λ), for an eigenvalue λ of A.

Proof: Let $dim(E_\lambda) = p$ and let $S = \{X_1, X_2, \ldots, X_p\}$ be a basis of E_λ. Then S can be extended to a basis S' of $\mathbb{F}^n$. Let $S' = \{X_1, X_2, \ldots, X_p, X_{p+1}, \ldots, X_n\}$. Then

$$
\begin{aligned}
&AX_1 = \lambda X_1\\
&AX_2 = \lambda X_2\\
&\vdots\\
&AX_p = \lambda X_p\\
&AX_{p+1} = a_{(p+1)1}X_1 + a_{(p+1)2}X_2 + \ldots + a_{(p+n)n}X_n\\
&\vdots\\
&AX_n = a_{n1}X_1 + a_{n2}X_2 + \ldots + a_{nn}X_n.
\end{aligned}
$$

The matrix representation of the above system of equations is

$$A = \begin{bmatrix} \lambda I_p & B \\ 0 & C \end{bmatrix},$$

where I_p is the identity matrix of order p. Thus, the characteristic polynomial of A is $f(x) = det(xI - A) = (x-\lambda)^p g(x)$, where $g(x)$ is a polynomial. Hence, the algebraic multiplicity of λ is at least p.

Definition 6. Let $A \in M_n(\mathbb{F})$. Then A is called diagonalizable if it has n linearly independent eigenvectors.

Lemma 7. Let $\lambda_1, \lambda_2, \ldots, \lambda_k$ be distinct eigenvalues of A and $v_1, v_2, \ldots, v_k$ be the corresponding eigenvectors respectively. Then $v_1, v_2, \ldots, v_k$ are linearly independent.

Proof. The proof is by induction. Let $k = 2$ and v_1, v_2 are linearly dependent. Then $v_1 = \alpha v_2 \Rightarrow$ for some $0 \neq \alpha \in \mathbb{F}$. Thus $Av_1 = \alpha A v_2 \Rightarrow \lambda_1 v_1 = \alpha\lambda_2 v_2 \Rightarrow \alpha(\lambda_1 - \lambda_2)v_2 \Rightarrow \lambda_1 = \lambda_2$, which is a contradiction. Suppose the result is true for $k-1$, that is, $v_1, v_2, \ldots, v_{k-1}$ are linearly independent. Let $\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_k v_k = 0$. Then $A(\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_k v_k) = 0 \Rightarrow \alpha_1\lambda_1 v_1 + \alpha_2\lambda_2 v_2 + \cdots + \alpha_k\lambda_k v_k = 0 \Rightarrow$ $\alpha_1(\lambda_1 - \lambda_k)v_1 + \alpha_2(\lambda_2 - \lambda_k)v_2 + \cdots + \alpha_{k-1}(\lambda_{k-1} - \lambda_1)v_{k-1} = 0$ (since $\lambda_k(\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_k v_k) = 0$). By induction hypothesis, $v_1, v_2, \ldots, v_{k-1}$ are linearly independent, hence $\alpha_i = 0$ for $1 \leq i \leq k-1$ as $\lambda_i \neq \lambda_k$. Thus, $\alpha_k v_k = 0$ so that $\alpha_k = 0$. □

Theorem 8. *Let $A \in M_n(\mathbb{F})$. The following statements are equivalent.*

1. A is diagonalizable.

2. There exists an invertible matrix P such that $P^{-1}AP = D$, where D is a diagonal matrix.

3. $A.M.(\lambda) = G.M.(\lambda)$ for each eigenvalue λ of A.

Proof. Let $X_1, X_2, \ldots, X_n$ be n independent eigenvectors of A. Construct a matrix P having X_i as its i-th column. Then $P^{-1}AP = D$, where D is a diagonal matrix and its i-th diagonal entry is the eigenvalue corresponding to X_i. Thus, $1 \Rightarrow 2$. For $2 \Rightarrow 1$, note that the columns of P are L.I. as P is invertible and each column of P is an eigenvector of A. By Lemma 7, $3 \Leftrightarrow 1$. □

Example 9. *Check diagonalizability of the matrix $A = \begin{bmatrix} 1 & 2 \\ 3 & 2 \end{bmatrix}$. If diagonalizable, find a matrix P such that $P^{-1}AP$ is a diagonal matrix.*

Solution: The characteristic polynomial of A is $(x+1)(x-4)$. Hence $A.M.(\lambda) = 1 = G.M.(\lambda)$ for $\lambda = 4, -1$. Hence, A is diagonalizable. For finding P such that $P^{-1}AP$ is diagonal matrix, we find eigenvectors of A. Eigenvectors corresponding to $\lambda = -1$ and 4 are respectively $v_{-1} = (1, -1)$ and $v_4 = (2, 3)$. Since eigenvectors corresponding to distinct eigenvalues are LI, $\{(1,-1), (2,3)\}$ is LI. Construct

$$P = \begin{bmatrix} 1 & 2 \\ -1 & 3 \end{bmatrix}.$$

One can see easily $D = P^{-1}AP$, where $D = \begin{bmatrix} -1 & 0 \\ 0 & 4 \end{bmatrix}$.

Definition 10. Let $T : V \to V$ be a linear transformation, where V is an n dimensional vector space over $\mathbb{F}$. Then T is called diagonalizable if V has a basis in which each vector is an eigenvector of T, that is, T has n independent eigenvectors.

Remark 11. *Let $V(\mathbb{F})$ is n-dimensional vector space and $T : V \to V$ be a linear operator. Then*

1. if T is diagonalizable and B is a basis of V consisting of eigenvectors, then $[T]_B = D$, where D is a diagonal matrix.

2. if T has n distinct eigenvalues, then T is diagonalizable.

3. if $\lambda_1, \lambda_2, \ldots, \lambda_k$ are distinct eigenvalues of T and E_{λ_i} are the associated eigenspaces, then T is diagonalizable if and only if $\dim V = \dim E_{\lambda_1} + \dim E_{\lambda_2} + \cdots + \dim E_{\lambda_k}$.

Example: The operator $T : \mathbb{R}^3 \to \mathbb{R}^3$ defined by $T(x, y, z) = (2x, x + 2y, 4x + 3z)$ is not diagonalizable. To see this, we consider the standard basis B of $\mathbb{R}^3$ and $[T]_B = \begin{pmatrix} 2 & 0 & 0 \\ 1 & 2 & 0 \\ 4 & 0 & 3 \end{pmatrix}$. The characteristic polynomial is $(x-2)^2(x-3)$. Thus $AM(2) = 2$ and $AM(3) = 1$. $E_2 = \{(0, x, 0) : x \in \mathbb{R}\}$ with $\dim E_2 = 1$ and $E_3 = \{(0, 0, x) : x \in \mathbb{R}\}$ with $\dim E_3 = 1$. Here, we get $\dim \mathbb{R}^3 \neq \dim E_2 + \dim E_3$. Hence, T is not diagonalizable.

Example: Let $T : \mathbb{R}^3 \to \mathbb{R}^3$ defined by $T(x, y, z) = (-x + 2y + 4z, -2x + 4y + 2z, -4x + 2y + 7z)$ is diagonalizable. To see this, we consider the standard basis B of $\mathbb{R}^3$ and $[T]_B = \begin{pmatrix} -1 & 2 & 4 \\ -2 & 4 & 2 \\ -4 & 2 & 7 \end{pmatrix}$. The characteristic polynomial is $\det(xI - [T]_B) = -x^3 + 10x^2 - 33x + 36 = (x-3)^2(x-4)$. Thus $AM(3) = 2$ and $AM(4) = 1$. Solving $([T]_B - 3I)X = 0$, we get $(1, 0, 1), (1, 2, 0)$ are independent solutions. Hence, $\dim E_3 = 2$ and $\dim E_4 = 1$. Here, we get $\dim \mathbb{R}^3 = \dim E_3 + \dim E_4$. Hence, T is diagonalizable.

Further, if we want to find a matrix P such that $P^{-1}[T]_B P = D$ for some diagonal matrix D. We need to compute a basis of eigen vectors. We have already found eigen vectors corresponding to $\lambda = 3$. Now let $\lambda = 4$, solving the system $([T]_B - 4I)X = 0$, we get $(2, 1, 2)$ is an eigen vector. The eigen vectors corresponding to distinct eigen values are linearly independent. Hence, $\{(1, 0, 1), (1, 2, 0), (2, 1, 2)\}$ is a basis consisting of eigen vectors. To find P, we will place basis vectors in the column, i.e., $P = \begin{pmatrix} 1 & 1 & 2 \\ 0 & 2 & 1 \\ 1 & 0 & 2 \end{pmatrix}$. The diagonal matrix D is obtained by placing the eigen values on the diagonal in the same order as eigen vectors in P, that is, if the first column of P is corresponding to eigen vector of λ_1, the first diagonal entry is going to be λ_1 and so on. Here, $D = \begin{pmatrix} 3 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 4 \end{pmatrix}$. Verify yourself that $D = P^{-1}[T]_B P$.