Indian Institute of Technology Delhi-Abu Dhabi
Course: AMTL 101 (Linear Algebra)
Semester II, 2024-2025
Instructor: Prof. Amit Priyadarshi

Contents

1 Change of Bases Matrix for a Linear Transformation

Let $T : V \to W$ be a linear transformation and β_1, β_2 be two ordered bases for V and β_1', β_2' be two ordered bases for W. Then we have two matrices $[T]_{\beta_1}^{\beta_1'}$ and $[T]_{\beta_2}^{\beta_2'}$. How are they related?

We know that for any $v \in V$,

$$\begin{aligned} [T(v)]_{\beta_1'} &= [T]_{\beta_1}^{\beta_1'}[v]_{\beta_1} \\ \text{and} \quad [T(v)]_{\beta_2'} &= [T]_{\beta_2}^{\beta_2'}[v]_{\beta_2}. \end{aligned}$$

Let P be the change of bases matrix from β_1 to β_2 and let Q be the change of bases matrix from β_1' to β_2'. Then

$$\begin{aligned} [v]_{\beta_2} &= P[v]_{\beta_1} \quad \forall\ v \in V \\ \text{and} \quad [w]_{\beta_2'} &= Q[w]_{\beta_1'} \quad \forall\ w \in W. \end{aligned}$$

Taking $w = T(v)$, we get

$$\begin{aligned} & [T(v)]_{\beta_2'} = Q[T(v)]_{\beta_1'} \\ \implies & [T]_{\beta_2}^{\beta_2'}[v]_{\beta_2} = Q[T]_{\beta_1}^{\beta_1'}[v]_{\beta_1} \\ \implies & [T]_{\beta_2}^{\beta_2'}P[v]_{\beta_1} = Q[T]_{\beta_1}^{\beta_1'}[v]_{\beta_1} \quad \forall\ v \in V \\ \implies & [T]_{\beta_2}^{\beta_2'}P = Q[T]_{\beta_1}^{\beta_1'} \\ \implies & [T]_{\beta_2}^{\beta_2'} = Q[T]_{\beta_1}^{\beta_1'}P^{-1}. \end{aligned}$$

Definition 1. *A linear operator is a linear transformation from a vector space V to itself.*

Let $T : V \to V$ be a linear operator and let β be an ordered basis for V. Then $[T]_\beta^\beta$ will simply be denoted by $[T]_\beta$.

Now suppose β_1 & β_2 be two ordered bases for V and P is the change of bases matrix from β_1 to β_2. Then $Q = P$.

$$\therefore\ [T]_{\beta_1} = P^{-1}[T]_{\beta_2}P.$$

Definition 2. *(**Similar Matrices**) Two square matrices A and B of the same size say $n \times n$ are said to be similar if there exists an invertible matrix P such that $B = P^{-1}AP$.*

Remark 1. *$[T]_{\beta_1}$ and $[T]_{\beta_2}$ are similar matrices.*

Note that similar matrices have the same determinant and the same trace.

Proof. Let $B = P^{-1}AP$ then $\det(B) = \det(P^{-1}AP) = \det(APP^{-1}) = \det(A)$.
Also, $\operatorname{trace}(B) = \operatorname{trace}(P^{-1}(AP)) = \operatorname{trace}((AP)P^{-1}) = \operatorname{trace}(A)$. $\square$

2 Eigenvalues and Eigenvectors

Definition 3. *Let A be an $n \times n$ matrix with real or complex entries. A scalar λ is called an eigenvalue of the matrix A if there exists some non-zero $X \in \mathbb{R}^n$ such that $AX = \lambda X$. All such $X \neq 0$ are called eigenvectors of A corresponding to the eigenvalue λ.*

Definition 4. *The eigenspace of A corresponding to an eigenvalue λ is defined as*

$$E_\lambda = \{X \in \mathbb{R}^n : AX = \lambda X\},$$

i.e. E_λ consists of all eigenvectors corresponding to eigenvalue λ and the zero vector.

Note that if X is an eigenvector of A corresponding to the eigenvalue λ, then cX is also an eigenvector for any $c \neq 0$.
Also, if $AX_1 = \lambda X_1$ & $AX_2 = \lambda X_2$ then $A(X_1 + X_2) = AX_1 + AX_2 = \lambda X_1 + \lambda X_2 = \lambda(X_1 + X_2)$. Therefore, E_λ is a subspace of $\mathbb{R}^n$. In fact, $E_\lambda = \text{Nullspace}(A - \lambda I)$, since $AX = \lambda X \iff (A - \lambda I)X = 0$.

Example 1. Let $A = 0$, the zero matrix. Then for any $0 \neq X \in \mathbb{R}^n$, $AX = 0 = \lambda X \implies \lambda = 0$. $\therefore$ $\lambda = 0$ is the only eigenvalue of A and the corresponding eigenspace is $\mathbb{R}^n$.

Example 2. Let $A = I$, the identity matrix. Then $AX = IX = X \implies \lambda = 1$ is the only eigenvalue.

Example 3. Let $A = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix}$ be a diagonal matrix. Observe that

$$A\begin{pmatrix}1\\0\\\vdots\\0\end{pmatrix}=\begin{pmatrix}\lambda_1\\0\\\vdots\\0\end{pmatrix}=\lambda_1\begin{pmatrix}1\\0\\\vdots\\0\end{pmatrix}$$

$$A\begin{pmatrix}0\\1\\\vdots\\0\end{pmatrix}=\begin{pmatrix}0\\\lambda_2\\\vdots\\0\end{pmatrix}=\lambda_2\begin{pmatrix}0\\1\\\vdots\\0\end{pmatrix}$$

$$\text{similarly, }A\begin{pmatrix}0\\0\\\vdots\\1\end{pmatrix}=\begin{pmatrix}0\\0\\\vdots\\\lambda_n\end{pmatrix}=\lambda_n\begin{pmatrix}0\\0\\\vdots\\1\end{pmatrix}$$

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

Theorem 1. *λ is an eigenvalue of $A \iff \det(A-\lambda I)=0$.*

Proof. λ is an eigenvalue of $A \iff AX = \lambda X$ for some non-zero X $\iff (A-\lambda I)X = 0$ for some non-zero $X \iff A-\lambda I$ is not invertible $\iff \det(A-\lambda I)=0$. □

Definition 5. ***(Characteristic Polynomial)*** *For $A \in M_{n\times n}(\mathbb{R})$, we define the characteristic polynomial of A as*

$$p(x) = \det(xI - A).$$

From the previous theorem, we see that the eigenvalues of A are the roots of the characteristic polynomial, i.e. λ is an eigenvalue of $A \iff p(\lambda)=0$.

Example 4.

1. For $A=0$, $p(x)=\det(xI-A)=\det(xI)=x^n$.

2. For $A=I$, $p(x)=\det(xI-I)=\det((x-1)I)=(x-1)^n$.

3. For $A = \begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{pmatrix}$,

$$p(x) = \begin{vmatrix} x-\lambda_1 & & & \\ & x-\lambda_2 & & \\ & & \ddots & \\ & & & x-\lambda_n \end{vmatrix} = (x-\lambda_1)(x-\lambda_2)\dots(x-\lambda_n).$$

Example 5. Let $A = \begin{pmatrix} 2 & 3 \\ 3 & 2 \end{pmatrix}$, find the eigenvalues and eigenvectors of A.

Solution

$$\begin{aligned} p(x) &= \det(xI - A) \\ &= \det \begin{pmatrix} x-2 & -3 \\ -3 & x-2 \end{pmatrix} \\ &= (x-2)^2 - 9 \\ &= (x+1)(x-5) \end{aligned}$$

$\therefore$ eigenvalues are -1 and 5.

Now, eigenvectors corresponding to the eigenvalue $\lambda = -1$:

$$= \{0 \neq x \in \mathbb{R}^2 : (A - (-1)I)X = 0\}$$

since,

$$\begin{aligned} &(A+I)X = 0 \\ &\Longrightarrow \begin{pmatrix} 3 & 3 \\ 3 & 3 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \\ &\Longleftrightarrow x + y = 0 \end{aligned}$$

$\therefore\ x = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$ is an eigenvector corresponding to the eigenvalue $\lambda = -1$.

Now, eigenvectors corresponding to the eigenvalue $\lambda = 5$:

$$= \{0 \neq x \in \mathbb{R}^2 : (A - 5I)X = 0\}$$

since,

$$
\begin{aligned}
&(A-5I)X = 0 \\
&\Longrightarrow \begin{pmatrix} -3 & 3 \\ 3 & -3 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \\
&\Longleftrightarrow x - y = 0
\end{aligned}
$$

$\therefore\ x = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ is an eigenvector corresponding to the eigenvalue $\lambda = 5$. ■

Example 6. Let $A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$, then

$$
\begin{aligned}
p(x) &= \det(xI - A) \\
&= \det \begin{pmatrix} x & 1 \\ -1 & x \end{pmatrix} \\
&= x^2 + 1.
\end{aligned}
$$

Here, the characteristic polynomial of A has no real roots. So, it has no real eigenvalue. So, no vector in $\mathbb{R}^2$ is an eigenvector.
However, if we allow complex eigenvectors, then $\lambda = \pm\iota$ (the complex roots of x^2+1) are complex eigenvalues, and we can calculate complex eigenvectors as follows:
1. For $\lambda = \iota$:

$$
\begin{aligned}
&\lambda I - A = \begin{pmatrix} \iota & 1 \\ -1 & \iota \end{pmatrix} \\
&\begin{pmatrix} \iota & 1 \\ -1 & \iota \end{pmatrix}\begin{pmatrix} z_1 \\ z_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \\
&\Longrightarrow \iota z_1 + z_2 = 0 \Longleftrightarrow z_2 = -\iota z_1.
\end{aligned}
$$

So, $\begin{pmatrix} 1 \\ -\iota \end{pmatrix}$ is an eigenvector for $\lambda = \iota$.
2. For $\lambda = -\iota$:

$$
\begin{aligned}
&\lambda I - A = \begin{pmatrix} -\iota & 1 \\ -1 & -\iota \end{pmatrix} \\
&\begin{pmatrix} -\iota & 1 \\ -1 & -\iota \end{pmatrix}\begin{pmatrix} z_1 \\ z_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \\
&\Longrightarrow -\iota z_1 + z_2 = 0 \Longleftrightarrow z_2 = \iota z_1.
\end{aligned}
$$

So, $\begin{pmatrix} 1 \\ \iota \end{pmatrix}$ is an eigenvector for $\lambda = -\iota$.

Remark 2. *If $\alpha \pm \iota\beta$ are eigenvalues of a real matrix A, then if $X \in M_{n\times 1}(\mathbb{C})$ such that $AX = (\alpha + \iota\beta)X$ (i.e. X is an eigenvector of A corresponding to eigenvalue $\alpha + \iota\beta$), then*

$$A\bar{X} = (\alpha - \iota\beta)\bar{X}$$

so, eigenvectors for a conjugate pair of eigenvalues can be found by taking the complex conjugate of eigenvectors.

3 Diagonalizability

Definition 6. *A matrix $A \in M_{n\times n}(\mathbb{R})$ or $M_{n\times n}(\mathbb{C})$ is said to be diagonalizable if A is similar to a diagonal matrix. I.e. A is diagonalizable if there exists an invertible matrix P such that $P^{-1}AP = D$, where D is a diagonal matrix.*

Of course, every diagonal matrix is diagonalizable.
Suppose A is diagonalizable, then there exists an invertible matrix P such that

$$P^{-1}AP = D = \begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{pmatrix}$$

$$\implies P^{-1}APe_i = De_i = \lambda_i e_i$$

where e_i is the vector in $\mathbb{R}^n$ with i^{th} coordinate 1 and 0 otherwise.

$$\implies A(Pe_i) = P(\lambda_i e_i) = \lambda_i(Pe_i)$$

$\implies$ Pe_i is an eigenvector of A with eigenvalue λ_i (Note that $Pe_i \neq 0$ because if $Pe_i = 0 \implies P^{-1}Pe_i = 0 \implies e_i = 0$, which is not true). So, we get eigenvectors $Pe_1, Pe_2, \ldots, Pe_n$ of the matrix A. Also, $Pe_1, Pe_2, \ldots, Pe_n$ are linearly independent (because they are columns of an invertible matrix P). Therefore, $\{Pe_1, Pe_2, \ldots, Pe_n\}$ is a basis for $\mathbb{R}^n$ or $\mathbb{C}^n$.
So, if A is diagonalizable then we can find a basis of $\mathbb{R}^n$ or $\mathbb{C}^n$ consisting of

eigenvectors of A.

Conversely, suppose $\{X_1, X_2, \ldots, X_n\}$ is a basis of $\mathbb{R}^n$ or $\mathbb{C}^n$ consisting of eigenvectors of A, i.e. $X_1, X_2, \ldots, X_n$ are n linearly independent eigenvectors of A.

$$\therefore \quad AX_i = \lambda_i X_i \quad \text{for} \quad i = 1, 2, \ldots, n.$$

Note that $\lambda_1, \lambda_2, \ldots, \lambda_n$ need not be distinct.
Let P be the matrix whose columns are $X_1, X_2, \ldots, X_n$.

Claim: $P^{-1}AP = D = \begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{pmatrix}$. We will show that $AP = PD$. Since,

$$\begin{aligned} APe_i &= A(Pe_i) = AX_i = \lambda_i X_i \\ \text{also,} \quad PDe_i &= P(\lambda_i e_i) = \lambda_i Pe_i = \lambda_i X_i \\ \therefore APe_i &= PDe_i \quad \forall\ i = 1, 2, \ldots, n. \\ \Longrightarrow AP &= PD \\ \Longrightarrow P^{-1}AP &= D. \end{aligned}$$

So, we get the following result:

Theorem 2. *A matrix $A \in M_{n\times n}(\mathbb{R})$ or $M_{n\times n}(\mathbb{C})$ is diagonalizable if and only if A has n linearly independent eigenvectors.*

Example 7. Let $A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$. Is A diagonalizable?

Solution The characteristic polynomial of A is

$$\begin{aligned} p(x) &= \det(xI - A) \\ &= \det \begin{pmatrix} x & -1 \\ 0 & x \end{pmatrix} \\ &= x^2 \end{aligned}$$

$\Longrightarrow$ $\lambda = 0$ is the only eigenvalue of A.

Now, $\lambda I - A = \begin{pmatrix} 0 & -1 \\ 0 & 0 \end{pmatrix}$, if $\begin{pmatrix} x \\ y \end{pmatrix}$ is an eigenvector of A, then

$$\begin{pmatrix} 0 & -1 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$
$$\iff y = 0.$$

$\therefore \left\{\begin{pmatrix} x \\ 0 \end{pmatrix} : x \neq 0\right\}$ is the set of eigenvectors of A.

Since, we cannot find two linearly independent eigenvectors of A, A is not diagonalizable. ■

Proposition 1. *If λ_1 and λ_2 are two distinct eigenvalues of A, then the corresponding eigenvectors are linearly independent.*

Proof. Since, λ_1, λ_2 are eigenvalues of A, there exist $X_1 \neq 0, X_2 \neq 0$ such that $AX_1 = \lambda_1 X_1$ and $AX_2 = \lambda_2 X_2$.
To show $\{X_1, X_2\}$ is linearly independent, assume

$$\begin{aligned} & c_1X_1 + c_2X_2 = 0 && (1)\\ \implies & A(c_1X_1 + c_2X_2) = A.0 = 0 \\ \implies & c_1AX_1 + c_2AX_2 = 0 \\ \implies & c_1\lambda_1X_1 + c_2\lambda_2X_2 = 0 && (2) \end{aligned}$$

subtracting the equation 2 from $\lambda_2\times$ eq.1, we get,

$$\begin{aligned} & c_1(\lambda_2 - \lambda_1)X_1 = 0 \\ \implies & c_1 = 0 \quad (\because X_1 \neq 0,\ \lambda_2 \neq \lambda_1). \end{aligned}$$

From equation 1, $c_2X_2 = 0 \implies c_2 = 0$ $(\because X_2 \neq 0)$. □

Exercise 1. *If $\lambda_1, \lambda_2, \ldots, \lambda_k$ are distinct eigenvalues of A with eigenvectors $X_1, X_2, \ldots, X_k$, then $\{X_1, X_2, \ldots, X_k\}$ is linearly independent.*

4 Eigenvalues and Eigenvectors of a Linear Operator

Definition 7. *Let $T : V \to V$ be a linear operator, where V is a vector space over a field $\mathbb{F}$. A scalar $\lambda \in \mathbb{F}$ is called an eigenvalue of T if there exists a* *non-zero vector $v \in V$ such that $T(v) = \lambda v$. Such non-zero v is called an eigenvector of T corresponding to eigenvalue λ.*

Remark 3. *If $A \in M_{n\times n}(\mathbb{F})$, then we can define $T : \mathbb{F}^n \to \mathbb{F}^n$ by $T(X) = AX$, where $X = (x_1, x_2, \ldots, x_n)^t$. So, eigenvalues and eigenvectors of T and A are the same.*

Suppose, V is a finite dimensional vector space and $T : V \to V$ is a linear operator. If $\beta = \{v_1, v_2, \ldots, v_n\}$ is an ordered basis for V, then we get an $n \times n$ matrix $A = [T]_\beta$ such that $[T(v)]_\beta = A[v]_\beta$.
Also, if β_1 and β_2 are two ordered bases for V, then $[T]_{\beta_1}$ and $[T]_{\beta_2}$ are similar matrices.

$\therefore$ $\lambda I - [T]_{\beta_1}$ and $\lambda I - [T]_{\beta_2}$ are similar matrices. Since, similar matrices have the same determinant, we can define the characteristic polynomial of T as

$$p_T(x) = \det(\lambda I - [T]_\beta),$$

where β is any ordered basis for V. This is a polynomial of degree n. Thus, to find the eigenvalues of a linear operator T on a finite dimensional vector space we choose any ordered basis of V and write the matrix of T w.r.t. this basis. We find the characteristic polynomial of this matrix and find the roots to get all eigenvalues of T.

Example 8. Let $T : P_2(\mathbb{R}) \to P_2(\mathbb{R})$ be the operator $T(a_0 + a_1x + a_2x^2) = a_1 + 2a_2x$. Find the eigenvalues and eigenvectors of T.

Solution Let $\beta = \{1, x, x^2\}$, then

$$\begin{aligned} T(1) &= 0 = 0.1 + 0.x + 0.x^2 \\ T(x) &= 1 = 1.1 + 0.x + 0.x^2 \\ T(x^2) &= 2x = 0.1 + 2.x + 0.x^2 \\ \therefore \quad [T]_\beta &= \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{pmatrix} = A. \end{aligned}$$

So, the characteristic polynomial of T will be

$$\begin{aligned}\det(\lambda I - A) &= \det\begin{pmatrix} \lambda & -1 & 0 \\ 0 & \lambda & -2 \\ 0 & 0 & \lambda \end{pmatrix} \\ &= \lambda^3 \\ &= 0 \iff \lambda = 0.\end{aligned}$$

$\therefore$ $\lambda = 0$ is the only eigenvalue of T.

To find the eigenvectors, consider

$$\begin{aligned}T(a_0 + a_1x + a_2x^2) &= 0.(a_0 + a_1x + a_2x^2) \\ \implies a_1 + 2a_2x &= 0 \\ \implies a_1 = 0,\ a_2 &= 0\end{aligned}$$

$\therefore$ $p(x) = a_0$ is an eigenvector of T for any $a_0 \neq 0$.

We can also use the matrix $[T]_\beta$ to find the eigenvectors as follows: first find the eigenvectors of A.

$$\begin{aligned}&\text{Since,}\quad AX = 0X = 0 \\ &\implies \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \\ &\implies x_2 = 0,\ x_3 = 0,\ x_1 = a \in \mathbb{R}.\end{aligned}$$

$\therefore$ $\begin{pmatrix} a \\ 0 \\ 0 \end{pmatrix}$, $a \neq 0$ is an eigenvector of A. Therefore, if v is an eigenvector of T, then $[v]_\beta = \begin{pmatrix} a \\ 0 \\ 0 \end{pmatrix}$.

$$\implies \quad v = a.1 + 0.x + 0.x^2 = a,$$

$\therefore$ $v = a \neq 0$ is an eigenvector of T. ∎

Remark 4. *The eigenvalues and eigenvectors are defined even for infinite dimensional vector spaces. However, we cannot define characteristic polynomial in such cases.*

Example 9. Let $T : V \to V$ be given by $T(p(x)) = p'(x)$, where V is the vector space of all real polynomials.
If $p(x) = a_0 \neq 0$, then

$$T(p(x)) = 0 = 0.p(x)$$

$\therefore$ $\lambda = 0$ is an eigenvalue and any non-zero constant polynomial is an eigenvector. Can you show that these are the only eigenvectors of T ?

Proposition 2. *For any matrix $A \in M_{n\times n}(\mathbb{C})$, the product of complex eigenvalues (counted with multiplicity) equals the determinant of A.*

Proof. Let $p(x) = \det(xI - A)$, we know that the roots of $p(x)$ are the eigenvalues.

$$\therefore\ p(x) = (x - \lambda_1)(x - \lambda_2)\ldots(x - \lambda_n),$$

where $\lambda_i's$ are the complex eigenvalues. They may not be distinct.
Then

$$\begin{aligned} p(0) &= (-1)^n \lambda_1\lambda_2\ldots\lambda_n \\ \implies \det(-A) &= (-1)^n \lambda_1\lambda_2\ldots\lambda_n \\ \implies \det(A) &= \lambda_1\lambda_2\ldots\lambda_n. \end{aligned}$$

□

Proposition 3. *For any matrix $A \in M_{n\times n}(\mathbb{C})$, the sum of the eigenvalues (counted with multiplicity) equals the trace of A.*

Proof. We will proof for 2×2 matrices:

$$\begin{aligned} \text{Let} \quad A &= \begin{pmatrix} a & b \\ c & d \end{pmatrix} \\ p(x) &= (x-a)(x-d) - bc \\ &= x^2 - (a+d)x + ad - bc \end{aligned}$$

$\therefore$ Sum of eigenvalues $= a + d = \text{trace}(A)$. □

5 Necessary and Sufficient Condition for Diagonalizability

Theorem 3. ***(Necessary and sufficient condition for diagonalizability)*** *Let A be any $n \times n$ matrix and suppose $\lambda_1, \lambda_2, \ldots, \lambda_k$ are the distinct* *eigenvalues of A. Let E_i be the eigenspace corresponding to eigenvalue λ_i, for $i = 1, 2, \ldots, k$. Then A is diagonalizable if and only if*

$$\sum_{i=1}^{k} \dim(E_i) = n.$$

Definition 8. *Let λ be an eigenvalue of A. Then the **geometric multiplicity** of λ is the dimension of the eigenspace corresponding to λ.*

Definition 9. *Let λ be an eigenvalue of A. Then the **algebraic multiplicity** of λ is the number of times λ is repeated in the roots of the characteristic polynomial.*

For example, if $p(x) = (x-1)(x-2)^3(x-3)^2$, then the algebraic multiplicity of 1, 2 & 3 are 1, 3 & 2, respectively.

Theorem 4. *The geometric multiplicity of each eigenvalue is less than or equal to the algebraic multiplicity.*

Theorem 5. *A is diagonalizable if and only if the geometric multiplicity equals the algebraic multiplicity for each eigenvalue.*

Example 10. Let $A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 3 \\ 0 & 0 & 2 \end{pmatrix}$. Is A diagonalizable?

Solution The characteristic polynomial of A is

$$\begin{aligned} p(x) &= \det \begin{pmatrix} x-1 & 0 & 0 \\ 0 & x-2 & -3 \\ 0 & 0 & x-2 \end{pmatrix} \\ &= (x-1)(x-2)^2 \end{aligned}$$

$\therefore$ eigenvalues are $1, 2, 2$. So, algebraic multiplicity of 2 is two. Let's find the geometric multiplicity of 2.

Since, $2I - A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & -3 \\ 0 & 0 & 0 \end{pmatrix}$ has rank 2,

$$\therefore \dim(\mathrm{Ker}(2I - A)) = 3 - 2 = 1,$$

$\therefore$ geometric multiplicity(2) $= 1 <$ algebraic multiplicity(2).

$\therefore$ A is not diagonalizable.

■

6 Cayley-Hamilton Theorem

Theorem 6. *Suppose A is any $n \times n$ matrix and let $p(x) = \det(xI - A)$ be the characteristic polynomial of A. Then the matrix $p(A)$ is the zero matrix.*

(If $p(x) = x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0$, then

$$p(A) = A^n + a_{n-1}A^{n-1} + \cdots + a_1 A + a_0 I.)$$

Example 11. If $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, then

$$\begin{aligned} p(x) &= \det(xI - A) \\ &= \begin{vmatrix} x-a & -b \\ -c & x-d \end{vmatrix} \\ &= (x-a)(x-d) - bc \\ &= x^2 - (a+d)x + ad - bc \\ \therefore\ p(A) &= A^2 - (a+d)A + (ad-bc)I. \end{aligned}$$

Verify that $p(A) = 0$.

6.1 Applications of Cayley-Hamilton Theorem

1. Calculating Inverse of a Matrix:

Let A be an $n \times n$ matrix and $p(x) = x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0$ be the characteristic polynomial of A. Note that A is invertible iff $a_0 \neq 0$ (because $a_0 = p(0) = \det(-A)$). By the Cayley-Hamilton theorem,

$$\begin{aligned} p(A) &= A^n + a_{n-1}A^{n-1} + \cdots + a_1 A + a_0 I \\ &= 0 \\ i.e.\ \ a_0 I &= -\left(A^n + a_{n-1}A^{n-1} + \cdots + a_1 A\right) \\ &= -\left(A^{n-1} + a_{n-1}A^{n-2} + \cdots + a_1 I\right) A \\ \Longrightarrow\ I &= -\frac{1}{a_0}\left(A^{n-1} + a_{n-1}A^{n-2} + \cdots + a_1 I\right) A \quad \text{if } a_0 \neq 0, \\ \Longrightarrow\ A^{-1} &= -\frac{1}{a_0}\left(A^{n-1} + a_{n-1}A^{n-2} + \cdots + a_1 I\right). \end{aligned}$$

Example 12. Let $A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 1 & 1 & 0 \end{pmatrix}$. Find A^{-1}.

Solution The characteristic polynomial of A is

$$\begin{aligned}
p(x) &= \begin{vmatrix} x-1 & 0 & 0 \\ 0 & x-1 & -1 \\ -1 & -1 & x \end{vmatrix} \\
&= (x-1)(x^2-x-1) \\
&= x^3-2x^2+1.
\end{aligned}$$

Since $p(0) = 1 \neq 0$, A is invertible. Also, by the Cayley-Hamilton theorem,

$$\begin{aligned}
& A^3 - 2A^2 + I = 0 \\
\implies & I = 2A^2 - A^3 = A(2A - A^2) \\
\implies & A^{-1} = 2A - A^2 \\
& = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 2 \\ 2 & 2 & 0 \end{pmatrix} - \begin{pmatrix} 1 & 0 & 0 \\ 1 & 2 & 1 \\ 1 & 1 & 1 \end{pmatrix} \\
& = \begin{pmatrix} 1 & 0 & 0 \\ -1 & 0 & 1 \\ 1 & 1 & -1 \end{pmatrix}
\end{aligned}$$

■

2. Calculating Powers of A or $g(A)$ for any Polynomial $g(x)$:

By the division algorithm, $g(x) = p(x)q(x) + r(x)$, where $p(x)$ is the characteristic polynomial and either $r(x) = 0$ or $\deg r(x) < \deg p(x) = n$. Therefore, $r(x)$ is a polynomial of degree at most $(n-1)$.
Now, $g(A) = p(A)q(A)+r(A) \implies g(A) = r(A)$ ($\because p(A) = 0$, by the Cayley-Hamilton theorem). To calculate $r(A)$, we only need $A, A^2, \ldots, A^{n-1}$.

Example 13. Let $A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 1 & 1 & 0 \end{pmatrix}$. Compute A^{100}.

Solution We saw in example 12, $p(x) = x^3-2x^2+1$, if we want to compute

A^{100}, let $g(x) = x^{100}$. By the division algorithm,

$$\begin{aligned} x^{100} &= (x^3 - 2x^2 + 1)q(x) + (ax^2 + bx + c) \\ &= (x-1)(x^2 - x - 1)q(x) + (ax^2 + bx + c) \\ &= (x-1)\left(x - \left(\frac{1+\sqrt{5}}{2}\right)\right)\left(x - \left(\frac{1-\sqrt{5}}{2}\right)\right) q(x) + (ax^2 + bx + c) \end{aligned}$$

after putting $x = 1,\ x = \frac{1+\sqrt{5}}{2}\ \frac{1-\sqrt{5}}{2}$ in the above equation, we get

$$\begin{aligned} 1 &= a + b + c \\ \left(\frac{1+\sqrt{5}}{2}\right)^{100} &= a\left(\frac{1+\sqrt{5}}{2}\right)^2 + b\left(\frac{1+\sqrt{5}}{2}\right) + c \\ \left(\frac{1-\sqrt{5}}{2}\right)^{100} &= a\left(\frac{1-\sqrt{5}}{2}\right)^2 + b\left(\frac{1-\sqrt{5}}{2}\right) + c \end{aligned}$$

We get 3 linear equations in a, b, c, which can be solved to get a, b, c. After that

$$A^{100} = r(A) = aA^2 + bA + cI.$$

■

Example 14. Calculate A^{50} for $A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$.

Solution The characteristic polynomial of A is

$$\begin{aligned} p(x) &= \det(xI - A) \\ &= \begin{vmatrix} x & -1 \\ 1 & x \end{vmatrix} \\ &= x^2 + 1. \end{aligned}$$

By the Cayley-Hamilton theorem, $A^2 + I = 0$.

$$\implies A^{50} = (A^2)^{25} = -I.$$

Alternatively,

$$x^{50} = (x^2 + 1)g(x) + ax + b,$$

after putting $x = i$ and $x = -i$, we get

$$-1 = i^{50} = ia + b$$
$$-1 = (-i)^{50} = -ia + b$$

after solving the above two equations, we get $a = 0$ and $b = -1$.

$$\therefore\ x^{50} = (x^2 + 1)g(x) - 1$$
$$\implies A^{50} = 0 - I = -I.$$

$\blacksquare$