Lecture Diagonalizability

- In the previous lecture, we saw that **not every** linear operator or matrix is diagonalizable.
- While we already have a **necessary and sufficient condition** for diagonalizability (**Theorem 5.1**), we haven't yet developed a **practical method** to test whether a given operator (or matrix) *can* be diagonalized —and if so, **how to do it**.

💡 **This lecture fills that gap.**

We'll develop a step-by-step method to:

- ✅ Determine if a linear operator is diagonalizable.

- 🔍 Understand the role of algebraic vs geometric multiplicity.

- 🧮 Actually **construct a basis of eigenvectors** when diagonalizability holds.

By the end of this lecture, you will have a **clear test** (**Theorem 5.8**) and a good understanding of the **diagonalization process**.

Theorem 5.5.

Let T be a linear operator on a vector space, and let $\lambda_1, \lambda_2, \ldots, \lambda_k$ be distinct eigenvalues of T. For each $i = 1, 2, \ldots, k$, let S_i be a finite set of eigenvectors of T corresponding to λ_i. If each $S_i (i = 1, 2, \ldots, k)$, is linearly independent, then $S_1 \cup S_2 \cup \cdots \cup S_k$ is linearly independent.

Proof: We prove this by induction.

- Base step. When $k=1$, it is obvious S_1 is linearly independent. by the assumption.

- Induction step. Assume the theorem is true for $k-1$ distinct eigenvalues. and $k \geq 2$

We will show the theorem is true for k.

For $1 \leq i \leq k$, consider $S_i = \{v_{i1}, v_{i2}, \cdots, v_{in_i}\}$.

be linearly independent set of eigenvectors corresponding to λ_i. We want to show. $S = S_1 \cup S_2 \cup \cdots \cup S_k$ is linearly independent.

Let a_{ij} be any scalar, where $1 \leq i \leq k$, $1 \leq j \leq n_i$.

s.t. $$\sum_{i=1}^{k} \sum_{j=1}^{n_i} a_{ij} \vec{v_{ij}} = \vec{0} \qquad \circledast$$

As $\vec{v_{ij}}$ is an eigenvector of T w.r.t. λ_i, $\underline{T v_{ij} = \lambda_i v_{ij}}$.

or $(T - \lambda_i I_v) v_{ij} = \vec{0}$

Apply $(T - \lambda_k I_v)$ on both sides of $\circledast$, we have.

$$\vec{0} = (T - \lambda_k I_v) \sum_{i=1}^{k} \sum_{j=1}^{n_i} a_{ij} \vec{v_{ij}} = \sum_{i=1}^{k} \sum_{j=1}^{n_i} a_{ij} (T - \lambda_k I_v) \vec{v_{ij}}$$

$$= \sum_{i=1}^{k-1} \sum_{j=1}^{n_i} a_{ij} (T - \lambda_k I_V) \vec{v}_{ij} + \sum_{j=1}^{n_k} a_{kj} \cancel{(T - \lambda_k I_V) \vec{v}_{kj}}^{\vec{0}}$$

$$= \sum_{i=1}^{k-1} \sum_{j=1}^{n_i} a_{ij} (\cancel{T \vec{v}_{ij}}^{\lambda_i v_{ij}} - \lambda_k \vec{v}_{ij})$$

$$= \sum_{i=1}^{k-1} \sum_{j=1}^{n_i} a_{ij} (\lambda_i - \lambda_k) \vec{v}_{ij} = \vec{0}$$

As $S_1 \cup S_2 \cup \cdots \cup S_{k-1}$ is linearly independent by the induction hypothesis, we must have.

$$a_{ij} (\lambda_i - \lambda_k) = 0, \quad \text{for } 1 \le i \le k-1, \ 1 \le j \le n_i$$

As $\lambda_1, \cdots \lambda_k$ are distinct, we know $\lambda_i - \lambda_k \ne 0$ for $1 \le i \le k-1$

Therefore $a_{ij} = 0$ for $1 \le i \le k-1$ and $1 \le j \le n_i$

Then $\circledast$ $\left(\sum_{i=1}^{k} \sum_{j=1}^{n_i} a_{ij} \vec{v}_{ij} = \vec{0} \right)$ reduces to

$$\sum_{j=1}^{n_k} a_{kj} \vec{v}_{kj} = \vec{0}.$$

Since S_k is also linearly independent $a_{kj} = 0$, $1 \le j \le n_k$.

Therefore $a_{ij} = 0$, where $1 \le i \le k$, $1 \le j \le n_i$

So S is linearly independent. $\square$

Corollary.

Let T be a linear operator on an n-dimensional vector space V. If T has n distinct eigenvalues, then T is diagonalizable.

Proof: Exercise.

Remark. The converse of the above corollary is not true. That is, even if T is diagonalizable, it does **not** necessarily have n distinct eigenvalues. For example, the identity operator is diagonalizable but has only one eigenvalue.

Example 1: Let

$$A = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \in M_{2\times 2}(\mathbb{R}).$$

The characteristic polynomial of A (thus L_A) is

$$\det(A - tI) = \det \begin{pmatrix} 1-t & 1 \\ 1 & 1-t \end{pmatrix} = t(t-2).$$

So the eigenvalues of A are 0 and 2. By the previous corollary, we know A (and hence L_A) is diagonalizable.

Definition.

A polynomial $f(t)$ in $\mathrm{P}(F)$ **splits over** F if there are scalars $c, a_1, \ldots, a_n$ (not necessarily distinct) in F such that

$$f(t) = c\,(t-a_1)\,(t-a_2)\cdots(t-a_n)$$

Examples.

1. $t^2 - 1 = (t+1)(t-1)$ splits over R.
2. $\left(t^2+1\right)(t-2)$ does not splits over R since t^2+1 cannot be factored into a product of linear factors over R.

   However, $\left(t^2+1\right)(t-2)$ does split over C because it factors into the product $(t+i)(t-i)(t-2)$.

Theorem 5.6.

The characteristic polynomial of any diagonalizable linear operator on a vector space V over a field F splits over F.

Proof: Let T be a diagonalizable linear operator on V, where $\dim(V) = n$. Let β be such a basis so that $[T]_\beta = D$ is diagonal and

Definition: Determinant and Characteristic Polynomial of an Operator.
Let T be a linear operator on a finite-dimensional vector space V, and let β be any ordered basis for V.

We define:

- The **determinant** of T, denoted $\det(T)$, to be the determinant of the matrix $A = [T]_\beta$.
- The **characteristic polynomial** of T to be the characteristic polynomial of A. That is,

$$f(t) = \det(A - tI_n).$$

These definitions are well-defined because they do not depend on the choice of basis.

$$D = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix}$$

Then by the def of the char. poly. of T is

$$f(t) = \det(D - tI_n) = \begin{vmatrix} \lambda_1 - t & 0 & \cdots & 0 \\ 0 & \lambda_2 - t & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & \lambda_n - t \end{vmatrix}$$

would introduce a -1

$$= (\lambda_1 - t)(\lambda_2 - t)\cdots(\lambda_n - t)$$

$$= (-1)^n (t - \lambda_1)(t - \lambda_2)\cdots(t - \lambda_n)$$

Thus by def of $f(t)$ splits over F (previous page), $f(t)$ split over F.

Remark.

- The **converse is false**:
  - A characteristic polynomial may **split** (i.e., factor completely),
  - but T may still **not** be diagonalizable. (See **Exercise 2** below)

The following concept provides a criterion for determining when an operator with a split characteristic polynomial is diagonalizable.

Definition.

Let λ be an eigenvalue of a linear operator or matrix with characteristic polynomial $f(t)$. The **multiplicity** (sometimes called the **algebraic multiplicity**) of λ is the largest positive integer k for which $(t-\lambda)^k$ is a factor of $f(t)$.

Remark

- If T is a diagonalizable linear operator on a finite-dimensional vector space V, then by **Theorem 5.1**:
  - The eigenvalues of T appear on the diagonal of $[T]_\beta$, each repeated according to its **algebraic multiplicity**.
- Therefore:
  - The number of **linearly independent eigenvectors** associated to an eigenvalue λ—also known as the **geometric multiplicity** of λ—plays a key role in determining whether T is diagonalizable.

    null space of $T-\lambda I$.
- According to **Theorem 5.4**: ($\vec{v}$ is an an eigenvector to $\lambda \Leftrightarrow (T-\lambda I)\vec{v} = \vec{0}$) $\Leftrightarrow \vec{v} \in \ker(T-\lambda I)$
  - Eigenvectors corresponding to an eigenvalue λ lie in $\ker(T-\lambda I_V)$.
  - This motivates the study of $\ker(T-\lambda I_V)$ in analyzing diagonalizability.

☆ Want: $\dim(\ker(T-\lambda I_V))$ = multiplicity of λ to ensure enough elements to form a basis. so T is diagona

---

Definition.

Let T be a linear operator on a vector space V, and let λ be an eigenvalue of T. The **eigenspace** of T corresponding to λ is defined as

$$E_\lambda = \{x \in V : T(x) = \lambda x\} = \ker(T-\lambda I_V).$$

Analogously, for a square matrix A, the eigenspace corresponding to the eigenvalue λ is defined to be the eigenspace of the associated linear operator L_A, that is,

$$E_\lambda = \ker(A-\lambda I).$$

The next theorem relates the **dimension of the eigenspace** E_λ (also called the **geometric multiplicity**) to the **algebraic multiplicity** of the eigenvalue λ.

Theorem 5.7.

Let T be a linear operator on a finite-dimensional vector space V, and let λ be an eigenvalue of T with algebraic multiplicity m. Then

$$1 \leq \dim(E_\lambda) \leq m.$$

Proof: See Page 264.

Exercise 2: Let T be the linear operator on $P_2(\mathbb{R})$ defined by $T(f(x)) = f'(x)$.

In this exercise, you will determine whether T is diagonalizable.

(1) Write down the matrix representation of T with respect to the standard basis $\beta = \{1, x, x^2\}$ for $P_2(\mathbb{R})$.

$T(1) = 1' = 0$, $\quad T(x) = x' = 1 + 0\cdot x + 0\, x^2$, $\quad T(x^2) = 0\cdot 1 + 2\cdot x + 0\cdot x^2$

Thus $\quad [T]_\beta = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{bmatrix}$

(2) Compute the characteristic polynomial of T. Then identify the eigenvalues and their algebraic multiplicities.

Let $\det([T]_\beta - \lambda I_3) = \begin{vmatrix} -\lambda & 1 & 0 \\ 0 & -\lambda & 2 \\ 0 & 0 & -\lambda \end{vmatrix} = -\lambda^3 = 0$

$\Rightarrow \lambda = 0$ with algebraic multiplicity 3.

(3) Solve $T(f(x)) = f'(x) = 0$ to find the eigenspace $E_0 = \ker(T)$. What is the geometric multiplicities of $\lambda = 0$? Recall geometric multiplicity of an eigenvalue λ is $\dim(E_\lambda)$.

Then eigenvector to $\lambda = 0$. satisfies

$$\underbrace{T(f(x))}_{f'(x)} = \underbrace{\lambda}_{0} f(x) \Rightarrow f'(x) = 0.$$

$$f'(x) = 0$$

which are precisely all constants. Thus $\{1\}$. so. $\dim(E_{\lambda=0}) = 1$.

(4) Is T diagonalizable? Why?

No, Becaus there is no basis of $P_2(\mathbb{R})$ consisting of eigenvectors of T.
So T is not diagonalizable.

Example 3 Let T be the linear operator on $\mathbb{R}^3$ defined by

$$T\begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} = \begin{pmatrix} 4a_1 + a_3 \\ 2a_1 + 3a_2 + 2a_3 \\ a_1 + 4a_3 \end{pmatrix}.$$

(1) Find the eigenvalues and corresponding eigenvectors of T.

(2) Decide whether T is diagonalizable.

(1) Let $\beta = \{e_1, e_2, e_3\}$ be the set of standard basis for $\mathbb{R}^3$.
Then

$$T(e_1) = T\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 4 \\ 2 \\ 1 \end{bmatrix}, \quad T(e_2) = T\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 3 \\ 0 \end{bmatrix}, \quad T(e_3) = T\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 4 \end{bmatrix}$$

Thus

$$A = [T]_\beta = \begin{bmatrix} 4 & 0 & 1 \\ 2 & 3 & 2 \\ 1 & 0 & 4 \end{bmatrix}$$

We compute.

$$\det(A-\lambda I_3) = \begin{vmatrix} 4-\lambda & 0 & 1 \\ 2 & 3-\lambda & 2 \\ 1 & 0 & 4-\lambda \end{vmatrix} = (4-\lambda)\begin{vmatrix} 3-\lambda & 2 \\ 0 & 4-\lambda \end{vmatrix}$$

$$+ \begin{vmatrix} 2 & 3-\lambda \\ 1 & 0 \end{vmatrix} = (4-\lambda)^2(3-\lambda) - (3-\lambda) = (3-\lambda)\left[(\lambda-4)^2 - 1\right]$$

$$= -(\lambda-3)(\lambda^2-8\lambda+15) = -(\lambda-3)(\lambda-3)(\lambda-5) = 0$$

Thus $\lambda_1 = 5$, $\lambda_2 = 3$ with multiplicity 2.

- When $\lambda_1 = 5$, we solve.

$$(A-\lambda_1 I)\vec{v} = \vec{0} \Rightarrow \begin{bmatrix} -1 & 0 & 1 \\ 2 & -2 & 2 \\ 1 & 0 & -1 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\left[\begin{array}{ccc|c} -1 & 0 & 1 & 0 \\ 2 & -2 & 2 & 0 \\ 1 & 0 & -1 & 0 \end{array}\right] \sim \left[\begin{array}{ccc|c} 0 & 0 & 0 & 0 \\ 0 & -2 & 4 & 0 \\ 1 & 0 & -1 & 0 \end{array}\right] \sim \left[\begin{array}{ccc|c} 1 & 0 & -1 & 0 \\ 0 & 1 & -2 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

$$\Rightarrow \begin{cases} x_1 = x_3 \\ x_2 = 2x_3 \end{cases} \quad \text{Thus} \quad \vec{v} = \begin{bmatrix} x_3 \\ 2x_3 \\ x_3 \end{bmatrix} = x_3\begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}$$

Thus an eigenvector to $\lambda_1 = 5$ is a scalar of $\vec{v}_1 = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}$

So $E_1 = \text{span}\left\{\begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}\right\}$. $\dim E_1 = 1$

When $\lambda_2 = 3$, we solve

$$(A-\lambda_2 I)\vec{v} = \vec{0} \Rightarrow \begin{bmatrix} 1 & 0 & 1 \\ 2 & 0 & 2 \\ 1 & 0 & 1 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

$\Rightarrow x_1 = -x_3$.

Thus $\vec{v} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} -x_3 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ x_2 \\ 0 \end{bmatrix} + \begin{bmatrix} -x_3 \\ 0 \\ x_3 \end{bmatrix} = x_2 \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$

So $\left\{ \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} \right\}$. is a basis for $E_{\lambda_2} = E_3$.

and $\dim E_3 = 2$ which coincides with the algebraic multiplicity of $\lambda_2 = 3$.

By Thm 5.3. $\beta' = \left\{ \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} \right.$ forms a linearly independent set with 3 vectors thus a basis for $\mathbb{R}^3$ consisting of eigenvectors of T.

So T is diagonalizable.

Exercise 2: $\dim(E_0) = 1 <$ alg. multipicity of 1. T not diagonalizable

🔍 **Question:**

What can you observe by comparing the geometric multiplicity $\dim(E_\lambda)$ and the algebraic multiplicity of each λ in **Exercise 2** and **Example 3**?

Example 3. $\dim(E_{\lambda_1}) =$ alg. mul. of λ_1 and $\dim(E_{\lambda_2}) =$ alg. mul of λ_2
T is diagonalizable

The guess from the observation is in general true. We have the following Theorem:

Theorem 5.8

Let T be a linear operator on a finite-dimensional vector space V, and suppose the characteristic polynomial of T splits. Let $\lambda_1, \lambda_2, \ldots, \lambda_k$ be the distinct eigenvalues of T. Then:

(a) T is diagonalizable if and only if the algebraic multiplicity of each λ_i is equal to $\dim(E_{\lambda_i})$, where E_{λ_i} is the eigenspace corresponding to λ_i.

(b) If T is diagonalizable and β_i is an ordered basis of E_{λ_i} for each i, then

$$\beta = \beta_1 \cup \beta_2 \cup \cdots \cup \beta_k$$

is an ordered basis for V consisting of eigenvectors of T.

Proof. (a). $\forall i$, $i = 1, \cdots, k$, let m_i be the (algebraic) multiplicity of λ_i
$d_i := \dim(E_i)$, $n = \dim(V)$.

"$\Rightarrow$" If T is diagonalizable, β is a basis of eigenvectors of T. Let $\beta_i = \beta \cap E_{\lambda_i}$ and let $n_i := |\beta_i|$, i.e., the number of vectors in β_i

Then

- $n_i \le d_i$ $= \dim(E_i)$, since β_i is linearly independent subset of β and E_i.

- $\dim(E_{\lambda_i}) = d_i \le m_i$ by Thm 5.7.

Then

$$n = \sum_{i=1}^{k} n_i \le \sum_{i=1}^{k} d_i \le \sum_{i=1}^{k} m_i = n$$

(forces all the inequalities to be "=")

since β has n vectors

since the degree of the char. poly of T = sum of the (alg) multiplicity of eigenvalues.

Thus $\sum_{i=1}^{k} d_i = \sum_{i=1}^{k} m_i$

$\Rightarrow \sum_{i=1}^{k} (m_i - d_i) = 0$.

As $(m_i - d_i) \geq 0 \ \forall i$, we know $m_i = d_i$.

Thus the algebraic multiplicity of λ_i (m_i) is the same as $\dim(E_{\lambda_i}) = d_i$.

"$\Leftarrow$" Assume $m_i = d_i$, $\forall i = 1, \cdots, k$, we show T is diagonalizable.

$\forall i$, let β_i be an ordered basis for E_{λ_i}. and

$\beta = \beta_1 \cup \beta_2 \cup \cdots \cup \beta_k$

By Thm 5.5, β is linearly independent.

Also as $d_i = m_i$, $\forall i$, β has

$$\sum_{i=1}^{k} d_i = \sum_{i=1}^{k} m_i = n.$$

vectors.

Thus β is an ordered basis consisting of the eigenvectors

So T is diagonalizable. This also proves (b)

Theorem 5.8 completes our study of the diagonalization problem. We summarize our results.

Test for Diagonalizability:

Let T be a linear operator on an n-dimensional vector space V. Then T is diagonalizable **if and only if** both of the following conditions hold:

1. The characteristic polynomial of T **splits** (i.e., it factors completely into linear terms over the field).
2. For each eigenvalue λ of T, the **algebraic multiplicity** of λ equals the **nullity** of $T - \lambda I$; that is,

$$\text{algebraic multiplicity of } \lambda = \text{geometric multiplicity of } \lambda = \text{nullity}(T - \lambda I) = n - \text{rank}(T - \lambda I).$$

These same conditions can be used to test whether a square matrix A is diagonalizable, since diagonalizability of A is equivalent to the diagonalizability of the associated linear operator L_A.