Linear Operators

ROB 501

Necmiye Ozay

- **Linear operators**
  - **Matrix representation of linear operators**
  - **Quick proof of representation theorem for linear operators**
- **Eigenvalues and eigenvectors (if time)**

Announcements

- No office hours today after lecture (I can stick around for 10-15 mins but need to leave by 10:35am)
  - feel free to email for an appointment at a later time if you have questions

Linear operator

Let $(\mathcal{X}, \mathcal{F})$ and $(\mathcal{Y}, \mathcal{F})$ be vector spaces.

$$L : \underbrace{\mathcal{X}}_{\text{domain of } L} \to \underbrace{\mathcal{Y}}_{\text{co-domain of } L}$$

is a **linear operator (mapping, transformation)** if $\forall x, \bar{x} \in \mathcal{X}$ and $\forall \alpha, \bar{\alpha} \in \mathcal{F}$

- 1. $L(x + \bar{x}) = L(x) + L(\bar{x})$ (additivity)
- 2. $L(\alpha x) = \alpha L(x)$ (homogeneity)

Alternatively (or equivalently) $\forall x, \bar{x} \in \mathcal{X}, \forall \alpha, \bar{\alpha} \in \mathcal{F}$

$$(*) \quad L(\alpha x + \bar{\alpha} \bar{x}) = \alpha L(x) + \bar{\alpha} L(\bar{x})$$

Examples:

1. Let $\mathcal{X} = \mathcal{F}^n$, $\mathcal{Y} = \mathcal{F}^m$, $\mathcal{F} = \mathcal{F}$, then any matrix $A \in \mathcal{F}^{m \times n}$, defines a linear operator

$$L : \mathcal{X} \to \mathcal{Y} \quad \text{s.t.} \quad L(x) = Ax.$$

(Easy to prove that 1. and 2. hold)

2. Let $\mathcal{X} = \mathbb{P}_3(t)$, $\mathcal{Y} = \mathbb{P}_3(t)$, $\mathcal{F} = \mathbb{R}$. Consider

$$L : \mathcal{X} \to \mathcal{Y}, \quad \forall p \in \mathcal{X} \quad L(p) = \frac{d}{dt} p \quad \text{(derivative)}.$$

L is a linear operator.

take $p_1, p_2 \in \mathcal{X}$, $\alpha_1, \alpha_2 \in \mathbb{R}$

$$L(\alpha_1 p_1 + \alpha_2 p_2) = \frac{d}{dt}(\alpha_1 p_1 + \alpha_2 p_2) = \frac{d}{dt}(\alpha_1 p_1) + \frac{d}{dt}(\alpha_2 p_2)$$

$$= \alpha_1 \frac{d}{dt} p_1 + \alpha_2 \frac{d}{dt} p_2$$

$$= \alpha_1 L(p_1) + \alpha_2 L(p_2) \checkmark$$

$(*)$ holds $\longrightarrow$ L is a linear operator.

3\. Let $\mathcal{X} = \mathbb{P}_3(t)$, $\mathcal{Y} = \mathbb{R}$, $\mathcal{F} = \mathbb{R}$. Consider

$L : \mathcal{X} \longrightarrow \mathcal{Y}$ s.t. $\forall p \in \mathcal{X}$ $\quad L(p) = \underbrace{p(5)}$.

evaluate p at $t=5$
or "$p(t)\big|_{t=5}$"

$p_1, p_2 \in \mathbb{P}_3(t) \qquad p = p_1 + p_2$

$$
\begin{aligned}
L(p_1 + p_2) = L(p) &= p(5) \\
&= p_1(5) + p_2(5) \\
&= L(p_1) + L(p_2)
\end{aligned}
$$

1\. ✓

$\alpha \in \mathcal{F} \quad p_1 \in \mathbb{P}_3(t)$

$\bar{p} = \alpha p_1$

$$
\begin{aligned}
L(\alpha p_1) &= L(\bar{p}) \\
&= \bar{p}(5) \\
&= \alpha\, p_1(5) \\
&= \alpha\, L(p_1)
\end{aligned}
$$

2\. ✓

Matrix representations of linear operators

$A \in \mathcal{F}^{n \times m}$

Let $(\mathcal{X}, \mathcal{F})$ have basis $\{u^1, \ldots, u^m\}$, let $(\mathcal{Y}, \mathcal{F})$ have basis $\{v^1, \ldots, v^n\}$ and let $L : \mathcal{X} \to \mathcal{Y}$ be a linear transformation. Then an $n \times m$ matrix A with entries in $\mathcal{F}$ is a **matrix representation** of L if:

$$\forall x \in \mathcal{X},\ \overset{\in \mathcal{F}^n}{[L(x)]_v} = A\, \overset{\in \mathcal{F}^m}{[x]_u}.$$

Representation Theorem for Linear Operators:

Thm: Given the definition above, define $A = [A_1 | \ldots | A_m]$ with $A_i = [L(u^i)]_v$. Then A is a matrix representation of L.

Ex: Let $\mathcal{X} = \mathbb{P}_3(t)$, $\mathcal{Y} = \mathbb{P}_3(t)$, $\mathcal{F} = \mathbb{R}$,

$L : \mathcal{X} \to \mathcal{Y}$, $\forall p \in \mathcal{X}$, $L(p) = \frac{d}{dt} p$

Take $u = \{\overset{u^1}{1}, \overset{u^2}{t}, \overset{u^3}{t^2}, \overset{u^4}{t^3}\}$ as a basis for $\mathcal{X}$ and $v = u$ as a basis for $\mathcal{Y}$. Find a matrix representation of L with respect to u and v.

$$A_i = \left[\frac{d}{dt} u^i\right]_v$$

$$A_1 = [L(u^1)]_v = \left[\frac{d}{dt} 1\right]_v = [0]_v = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

$$A_2 = [L(u^2)]_v = \left[\frac{d}{dt} t\right]_v = [1]_v = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

$$A_3 = [L(u^3)]_V = \left[\frac{d}{dt} t^2\right]_V = [2t]_V = \begin{bmatrix} 0 \\ 2 \\ 0 \\ 0 \end{bmatrix}$$

$$A_4 = [L(u^4)]_V = \left[\frac{d}{dt} t^3\right]_V = [3t^2]_V = \begin{bmatrix} 0 \\ 0 \\ 3 \\ 0 \end{bmatrix}$$

$$A = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Sanity check:

$p \in \mathbb{P}_3(t)$ $\quad p(t) = a_0 + a_1 t + a_2 t^2 + a_3 t^3$

$$[p]_U = \begin{bmatrix} a_0 \\ a_1 \\ a_2 \\ a_3 \end{bmatrix}$$

Claim:

$$\left[\frac{d}{dt} p\right]_V = A[p]_U$$

Check:

$$A \cdot \begin{bmatrix} a_0 \\ a_1 \\ a_2 \\ a_3 \end{bmatrix} = \begin{bmatrix} a_1 \\ 2a_2 \\ 3a_3 \\ 0 \end{bmatrix} \leftrightarrow a_1 + 2a_2 t + 3a_3 t^2$$

$$= \frac{d}{dt} p \ \checkmark$$

$$\begin{tikzcd} X \arrow[r, "L"] \arrow[d, "{[\cdot]_U}"'] & Y \arrow[d, "{[\cdot]_V}"] \\ \mathbb{F}^m \arrow[r, "A"] & \mathbb{F}^n \end{tikzcd}$$

"commutation diagram"

<u>Remark</u>: (very special case of commutation diagram).

Consider $(X, \mathbb{F})$ w/ basis $\{u^1, \dots, u^m\}$

and $(\mathcal{X}, \mathcal{F})$ w/ basis $\{\bar{u}^1, \dots, \bar{u}^m\}$

Define the <u>identity operator</u> $L : \mathcal{X} \to \mathcal{X}$ s.t. $\forall x \in \mathcal{X}$ $L(x) = x$. Then, the matrix representation of L is the change of basis matrix!

$$A_i = [L(u^i)]_{\bar{u}} = [u^i]_{\bar{u}} = P_i$$

We denote $L(x) = x$ as $Id(x) = x$.

$$\begin{tikzcd}
\mathcal{X} \arrow[r, "Id"] \arrow[d, "{[\cdot]_u}"'] & \mathcal{X} \arrow[d, "{[\cdot]_{\bar{u}}}"] \\
\mathcal{F}^m \arrow[r, "P"] & \mathcal{F}^m
\end{tikzcd}$$

$x = t + 1$ $\quad x = 1 \cdot u^1 + 1 \cdot u^2$

$u = \{1, t\}$ $\quad [x]_u = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$

$\tilde{u} = \{1, t+1\}$.

$x = 0 \cdot \tilde{u}^1 + 1 \cdot \tilde{u}^2$

$[x]_{\tilde{u}} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$

Proof of representation Theorem:

$x \in \mathcal{X}$, $\quad x = \alpha_1 u^1 + \dots + \alpha_m u^m \quad [x]_u = \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_m \end{bmatrix}$

$L : \mathcal{X} \to \mathcal{Y}$

$$L(x) = L(\alpha_1 u^1 + \dots + \alpha_m u^m)$$

} L is a linear operator

$$L(x) = \alpha_1 L(u^1) + \dots + \alpha_m L(u^m)$$

$$\mathcal{F}^n \ni [L(x)]_V = [\alpha_1 L(u^1) + \dots + \alpha_m L(u^m)]_V$$

} repres. of vectors preserve addition and scalar mult.

$$= \alpha_1 [L(u^1)]_V + \dots + \alpha_m [L(u^m)]_V$$

$$= \left[[L(u^1)]_V \,\middle|\, \dots \,\middle|\, [L(u^m)]_V \right] \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_m \end{bmatrix}$$

$$[L(x)]_v = A[x]_u \quad \forall x \in D.$$

$A_1 \quad A_m \quad [x]_u$

Eigenvalues and Eigenvectors

Def: Let A be an $n \times n$ matrix w/ complex coefficients ($A \in \mathbb{C}^{n \times n}$). $\lambda \in \mathbb{C}$ is an eigenvalue of A if $\exists\, v \in \mathbb{C}^n$, $v \neq 0$ s.t. $A \cdot v = \lambda v$. The vector v is called an eigenvector corresponding to λ.

not unique!

identity matrix

Facts: λ is an e-value $\Longleftrightarrow$ $\det(\lambda I - A) = 0$

(b.c. $0 = \underbrace{\lambda v - Av}_{v \neq 0} = \lambda \cdot I \cdot v - Av = (\lambda I - A) v$)

Ex:
$$A=\begin{bmatrix}0 & 1\\ -1 & 0\end{bmatrix} \qquad \lambda I - A = \begin{bmatrix}\lambda & -1\\ 1 & \lambda\end{bmatrix}$$

$$\det(\lambda I - A) = \lambda^2 + 1 = 0 \longrightarrow \lambda_1 = j \quad \lambda_2 = -j$$

complex conjugate pair

Moreover, $v^1 = \begin{bmatrix}1\\ j\end{bmatrix}$, $v^2 = \begin{bmatrix}1\\ -j\end{bmatrix}$

Def. $\Delta(\lambda) = \det(\lambda I - A)$ is called the characteristic polynomial of A.

$\Delta(\lambda) = \det(\lambda I - A) = 0$ is called the characteristic equation of A.

* Fundamental theorem of algebra says

$$\Delta(\lambda) = (\lambda - \lambda_1)^{m_1} (\lambda - \lambda_2)^{m_2} \cdots (\lambda - \lambda_p)^{m_p}$$

where $\lambda_1, \ldots, \lambda_p$ are distinct and

$m_1 + m_2 + \ldots + m_p = n$.

($m_i \geq 1$ is called the multiplicity of λ_i)

<u>Theorem</u>: Let A be $n \times n$ w/ complex coefficients. Suppose $m_1 = m_2 = \ldots = m_n = 1$ (all e-values are distinct, $\lambda_i \neq \lambda_j$ $i \neq j$). Then, the corresponding e-vectors form a <u>basis</u> for $(\mathbb{C}^n, \mathbb{C})$.

That is, $\{v^1, \ldots, v^n\}$ are linearly independent.

<u>Note</u>: (a) Converse is false (i.e. repeated e-values can still have lin. indep. e-vectors) (example, $A = I$).)

(b) Proof will be posted w/ the lecture notes (it is a nice read!)

<u>Def</u>: Two square matrices A and B are <u>similar</u> if $\exists$ an invertible matrix P s.t. $B = \underbrace{PAP^{-1}}_{\text{similarity transformation}}$

EECS 560 Handout
Eigenvectors corresponding to distinct eigenvalues

Theorem: Let A be an $n \times n$ matrix with complex coefficients. Let $\lambda_1, \ldots, \lambda_n$ be the eigenvalues of A and let $v^1, \ldots, v^n$ be the eigenvectors corresponding to these eigenvalues. If $\lambda_1, \ldots, \lambda_n$ are distinct (i.e., $\lambda_i \neq \lambda_j$ for $i \neq j$), then $\{v^1, \ldots, v^n\}$ is linearly independent over $\mathbb{C}$.

Before proving this theorem we will prove or recall some facts that will be useful in the proof.

Fact 1: $(A - \lambda_i I)\, v^i = 0$, $v^i \neq 0$ for $i = 1, 2, \ldots, n$.
This is by definition of the eigenvalues and eigenvectors.

Fact 2: $(A - \lambda_i I)(A - \lambda_k I) = (A - \lambda_k I)(A - \lambda_i I)$ for $1 \leq k,\, i \leq n$.
In general matrix multiplication does not commute but it commutes for this specific case which can be verified by performing the multiplications on both sides. In particular, both sides are equal to $A^2 - \lambda_i A - \lambda_k A + \lambda_k \lambda_i I$.

Fact 3: $(A - \lambda_2 I)(A - \lambda_3 I) \cdots (A - \lambda_n I)\, v^i = 0$ for $2 \leq i \leq n$.
From Fact 2, we can deduce that for any $2 \leq i \leq n$,

$$(A - \lambda_2 I)(A - \lambda_3 I) \cdots (A - \lambda_n I) = \left[\prod_{k=2, k \neq i}^{n} (A - \lambda_k I) \right] [A - \lambda_i I].$$

Multiplying both sides with v^i, *we get*

$$(A - \lambda_2 I)(A - \lambda_3 I) \cdots (A - \lambda_n I)\, v^i = \left[\prod_{k=2, k \neq i}^{n} (A - \lambda_k I) \right] \underbrace{[A - \lambda_i I]\, v^i}_{=0 \text{(by Fact 1)}} = 0.$$

Proof of the Theorem: Let eigenvalues be distinct and let's assume by contradiction that $\{v^1, \ldots, v^n\}$ is linearly dependent. Then, there exists $\alpha_1, \ldots, \alpha_n \in \mathbb{C}$, not all zero, such that

$$\alpha_1 v^1 + \alpha_2 v^2 + \ldots + \alpha_n v^n = 0. \tag{1}$$

Without loss of generality, we suppose that $\alpha_1 \neq 0$. Multiplying Eq. (1) from left with the matrix $(A - \lambda_2 I)(A - \lambda_3 I) \cdots (A - \lambda_n I)$ we get

$$(A - \lambda_2 I)(A - \lambda_3 I) \cdots (A - \lambda_n I)\left[\alpha_1 v^1 + \alpha_2 v^2 + \ldots + \alpha_n v^n\right] = 0 \qquad (2)$$

By using Fact 3, Eq. (2) reduces to

$$(A - \lambda_2 I)(A - \lambda_3 I) \cdots (A - \lambda_n I)\,\alpha_1 v^1 = 0 \qquad (3)$$

Noting that $(A - \lambda_k I) v^i = A v^i - \lambda_k v^i = \lambda_i v^i - \lambda_k v^i = (\lambda_i - \lambda_k) v^i$, we get

$$\alpha_1 \underbrace{(A - \lambda_2 I) \cdots \underbrace{(A - \lambda_{n-1} I) \underbrace{(A - \lambda_n I) v^1}_{(\lambda_1 - \lambda_n) v^1}}_{(\lambda_1 - \lambda_{n-1})(\lambda_1 - \lambda_n) v^1}}_{\vdots} = \underbrace{\alpha_1}_{\neq 0} (\lambda_1 - \lambda_2) \cdots (\lambda_1 - \lambda_{n-1})(\lambda_1 - \lambda_n) \underbrace{v^1}_{\neq 0} = 0$$

$$\Longrightarrow (\lambda_1 - \lambda_2) \cdots (\lambda_1 - \lambda_{n-1})(\lambda_1 - \lambda_n) = 0,$$

which implies at least one of the λ_i's for $i > 1$ should be the same as λ_1; that is, eigenvalues are not distinct, which is a contradiction.

Therefore, $\left\{v^1, \ldots, v^n\right\}$ is linearly independent (since we have n linearly independent vectors, they form a basis for $(\mathbb{C}^n, \mathbb{C})$).