Mathematics for Robotics (ROB-GY 6013 Section A)

- **Week 5:**
  - Eigenvalues and Eigenvectors
  - Similar Matrices
- **Homework 3 will be posted by tomorrow**

$$Av = v$$

A is a square ($n \times n$) matrix

$$Av = v$$

A is the *identity matrix*

- …if we take v to be all possible real n-tuples

$$Av = v$$

Find the v

$$\begin{bmatrix} 4 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} v = v$$

$$R_{\hat{k}}(\theta) v = v$$

Find the v

$$\begin{bmatrix} 4 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} v = v \qquad v = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}$$

$$R_{\hat{k}}(\theta) v = v \qquad v = \hat{k}$$

Identity-ish

- For some scalar λ $$Av = \lambda v$$

$$\begin{bmatrix} 4 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} v = \lambda v$$

Identity-ish

- For some scalar λ $$Av = \lambda v$$

$$\begin{bmatrix} 4 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} v = \lambda v \qquad \begin{aligned} \lambda_1 &= 4 \\ \lambda_2 &= 3 \\ \lambda_3 &= 2 \\ \lambda_4 &= 1 \end{aligned} \qquad v_1 = \begin{bmatrix} 100 \\ 0 \\ 0 \\ 0 \end{bmatrix}, v_2 = \begin{bmatrix} 0 \\ 3.1 \\ 0 \\ 0 \end{bmatrix}, v_3 = \begin{bmatrix} 0 \\ 0 \\ 42 \\ 0 \end{bmatrix}, v_4 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}$$

Identity-ish

- For some scalar λ $$Av = \lambda v$$

$$\begin{bmatrix} 4 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} v = \lambda v \qquad \begin{matrix} \lambda_1 = 4 \\ \lambda_2 = 3 \\ \lambda_3 = 2 \\ \lambda_4 = 1 \end{matrix} \qquad v_1 = \begin{bmatrix} 100 \\ 0 \\ 0 \\ 0 \end{bmatrix}, v_2 = \begin{bmatrix} 0 \\ 3.1 \\ 0 \\ 0 \end{bmatrix}, v_3 = \begin{bmatrix} 0 \\ 0 \\ 42 \\ 0 \end{bmatrix}, v_4 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}$$

- The scalars are the **eigenvalues**.
- The vectors corresponding to the scalars are **eigenvectors**.

Definition: Eigenvalues and Eigenvectors

- Let A be an $n \times n$ matrix with **complex** coefficients.

  A scalar $\lambda \in \mathbb{C}$ is an **eigenvalue (e-value)** of A, if there exists ***a*** **non-zero vector** $v \in \mathbb{C}^n$ such that $Av = \lambda v$. Any such vector v is called an **eigenvector (e-vector)** associated with λ.

Definition: Eigenvalues and Eigenvectors

- Let A be an $n \times n$ matrix with **complex** coefficients.

  A scalar $\lambda \in \mathbb{C}$ is an **eigenvalue (e-value)** of A, if there exists **a non-zero vector** $v \in \mathbb{C}^n$ such that $Av = \lambda v$. Any such vector v is called an **eigenvector (e-vector)** associated with λ.

- Eigenvectors are **not** unique
- Is 0 a permissible eigenvalue?
- Why can't the zero vector be an eigenvector?

Definition: Eigenvalues and Eigenvectors

- Let A be an $n \times n$ matrix with **complex** coefficients.

  A scalar $\lambda \in \mathbb{C}$ is an **eigenvalue (e-value)** of A, if there exists **a non-zero vector** $v \in \mathbb{C}^n$ such that $Av = \lambda v$. Any such vector v is called an **eigenvector (e-vector)** associated with λ.

- Eigenvectors are **not** unique
- Is 0 a permissible eigenvalue? **Yes! Why wouldn't it?**
- Why can't the zero vector be an eigenvector? **Then all possible scalars would be eigenvalues because** $A\,0 = 0$.

Finding Eigenvalues and Eigenvalues

$$\boxed{\exists v \neq 0 \quad \text{s.t} \quad Av = \lambda v}$$

$$\exists v \neq 0 \quad \text{s.t} \quad (\lambda I - A)v = 0 \Leftrightarrow \det(\lambda I - A) = 0$$

Finding Eigenvalues and Eigenvalues

$$\boxed{\exists v \neq 0 \ \ \text{s.t} \ \ Av = \lambda v}$$

$$\exists v \neq 0 \ \ \text{s.t} \ \ (\lambda I - A)v = 0 \Leftrightarrow \det(\lambda I - A) = 0$$

- **Characteristic polynomial** $\det(\lambda I - A)$
- **Characteristic equation** $\det(\lambda I - A) = 0$

Finding Eigenvalues and Eigenvalues

$$\boxed{\exists v \neq 0 \ \ \text{s.t} \ \ Av = \lambda v}$$

$$\exists v \neq 0 \ \ \text{s.t} \ \ (\lambda I - A)v = 0 \Leftrightarrow \det(\lambda I - A) = 0$$

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

Finding Eigenvalues and Eigenvalues

$$\boxed{\exists v \neq 0 \quad \text{s.t} \quad Av = \lambda v}$$

$$\exists v \neq 0 \quad \text{s.t} \quad (\lambda I - A)v = 0 \Leftrightarrow \det(\lambda I - A) = 0$$

$$A = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \qquad \lambda I - A = \begin{bmatrix} \lambda & -1 \\ 1 & \lambda \end{bmatrix} \qquad \det(\lambda I - A) = \lambda^2 + 1 = 0$$

Finding Eigenvalues and Eigenvalues

$$A = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \qquad \lambda I - A = \begin{bmatrix} \lambda & -1 \\ 1 & \lambda \end{bmatrix} \qquad \det(\lambda I - A) = \lambda^2 + 1 = 0$$

- **Eigenvalues**

$$\lambda_1 = j$$

$$\lambda_2 = -j$$

Finding Eigenvalues and Eigenvalues

$$A = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \qquad \lambda I - A = \begin{bmatrix} \lambda & -1 \\ 1 & \lambda \end{bmatrix} \qquad \det(\lambda I - A) = \lambda^2 + 1 = 0$$

- **Eigenvalues**

$$\lambda_1 = j \qquad Av^i = \lambda_i v^i \qquad \begin{bmatrix} -\lambda_i & 1 \\ -1 & -\lambda_i \end{bmatrix} \begin{bmatrix} v_1^i \\ v_2^i \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$\lambda_2 = -j \qquad (A - \lambda_i I)v^i = 0$$

$$\begin{bmatrix} -j & 1 \\ -1 & -j \end{bmatrix} \begin{bmatrix} v_1^1 \\ v_2^1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \qquad \begin{matrix} -jv_1^1 + v_2^1 = 0 \\ -1v_1^1 - jv_2^1 = 0 \end{matrix} \qquad \begin{matrix} jv_1^1 = v_2^1 \\ v_1^1 = -jv_2^1 \end{matrix}$$

$$\begin{bmatrix} j & 1 \\ -1 & j \end{bmatrix} \begin{bmatrix} v_1^2 \\ v_2^2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \qquad \begin{matrix} jv_1^2 + v_2^2 = 0 \\ -v_1^2 + jv_2^2 = 0 \end{matrix} \qquad \begin{matrix} jv_1^2 = -v_2^2 \\ v_1^2 = jv_2^2 \end{matrix}$$

Non-unique so pick
$v_1{}^1 = v_2{}^1 = 1$

Note that some equations are redundant, as you would expect when there are non-unique solutions!

- **Eigenvectors**

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

Finding Eigenvalues and Eigenvalues

$$A = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \qquad \lambda I - A = \begin{bmatrix} \lambda & -1 \\ 1 & \lambda \end{bmatrix} \qquad \det(\lambda I - A) = \lambda^2 + 1 = 0$$

- **Eigenvalues**

$$\lambda_1 = \boxed{j}$$
$$\lambda_2 = \boxed{-j}$$

$$Av^i = \lambda_i v^i$$
$$(A - \lambda_i I)v^i = 0$$

$$\begin{bmatrix} -\lambda_i & 1 \\ -1 & -\lambda_i \end{bmatrix} \begin{bmatrix} v_1^i \\ v_2^i \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

- **Eigenvectors**

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

- Eigenvalues and eigenvectors of real matrices are not always real, but must be **complex conjugate pairs**

Definition: Characteristic Polynomial/Equation

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

  $\Delta(\lambda) = 0$ is called the **characteristic equation**.

  By the **Fundamental Theorem of Algebra**, Δ(λ) can be factored as

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

  where $\lambda_1, \ldots, \lambda_p$ are the distinct **eigenvalues (roots)**,
  m_i is the **algebraic multiplicity** of λ_i, and $m_1 + m_2 + \ldots + m_p = n$.

  The **geometric multiplicity** of λ_i is defined as $\eta_i := \dim \operatorname{null}(A - \lambda_i I)$.

Definition: Characteristic Polynomial/Equation

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

  $\Delta(\lambda) = 0$ is called the **characteristic equation**.

  By the **Fundamental Theorem of Algebra**, Δ(λ) can be factored as

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

  where $\lambda_1, \ldots, \lambda_p$ are the distinct **eigenvalues (roots)**,
  m_i is the **algebraic multiplicity** of λ_i, and $m_1 + m_2 + \ldots + m_p = n$.

  The **geometric multiplicity** of λ_i is defined as $\eta_i := \dim \mathrm{null}(A - \lambda_i I)$.

$\mathrm{null}(B)$ denotes the null-space of matrix B, the space of all x such that $Bx = 0$

New Theorem: Distinct Eigenvalues

- Let A be an $n \times n$ matrix with **complex** or **real** coefficients.

  If the **eigenvalues** $\{\lambda_1, \ldots, \lambda_n\}$ are distinct, that is $\lambda_i \neq \lambda_j$ for all $1 \leq i \neq j \leq n$, then the **eigenvectors** $\{v^1, \ldots, v^n\}$ are **linearly independent** in $(\mathbb{C}^n, \mathbb{C})$

New Theorem: Distinct Eigenvalues

- Let A be an $n \times n$ matrix with **complex** or **real** coefficients.

  If the **eigenvalues** $\{\lambda_1, \ldots, \lambda_n\}$ are distinct, that is $\lambda_i \neq \lambda_j$ for all $1 \leq i \neq j \leq n$, then the **eigenvectors** $\{v^1, \ldots, v^n\}$ are **linearly independent** in $(\mathbb{C}^n, \mathbb{C})$

  Restatement of the theorem:
  If the **eigenvalues** $\{\lambda_1, \ldots, \lambda_n\}$ are distinct then the **eigenvectors** $\{v^1, \ldots, v^n\}$ are form a **basis** in $(\mathbb{C}^n, \mathbb{C})$

New Theorem: Distinct Eigenvalues

- Let A be an $n \times n$ matrix with **complex** or **real** coefficients.

  If the **eigenvalues** $\{\lambda_1,\dots,\lambda_n\}$ are distinct, that is $\lambda_i \neq \lambda_j$ for all $1 \leq i \neq j \leq n$, then the **eigenvectors** $\{v^1,\dots,v^n\}$ are **linearly independent** in $(\mathbb{C}^n,\mathbb{C})$

  Restatement of the theorem:
  If the **eigenvalues** $\{\lambda_1,\dots,\lambda_n\}$ are distinct then the **eigenvectors** $\{v^1,\dots,v^n\}$ are form a **basis** in $(\mathbb{C}^n,\mathbb{C})$

$$\begin{bmatrix} 4 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} v = v \qquad \begin{aligned} \lambda_1 &= 4 \\ \lambda_2 &= 3 \\ \lambda_3 &= 2 \\ \lambda_4 &= 1 \end{aligned} \qquad v_1 = \begin{bmatrix} 100 \\ 0 \\ 0 \\ 0 \end{bmatrix}, v_2 = \begin{bmatrix} 0 \\ 3.1 \\ 0 \\ 0 \end{bmatrix}, v_3 = \begin{bmatrix} 0 \\ 0 \\ 42 \\ 0 \end{bmatrix}, v_4 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}$$

New Theorem: Distinct Eigenvalues

- Let A be an $n \times n$ matrix with **complex** or **real** coefficients.

  If the **eigenvalues** $\{\lambda_1, \ldots, \lambda_n\}$ are distinct, that is $\lambda_i \neq \lambda_j$ for all $1 \leq i \neq j \leq n$, then the **eigenvectors** $\{v^1, \ldots, v^n\}$ are **linearly independent** in $(\mathbb{C}^n, \mathbb{C})$

  Restatement of the theorem:
  If the **eigenvalues** $\{\lambda_1, \ldots, \lambda_n\}$ are distinct then the **eigenvectors** $\{v^1, \ldots, v^n\}$ are form a **basis** in $(\mathbb{C}^n, \mathbb{C})$

$$\begin{bmatrix} 4 & 0 & 0 & 0 \\ & & & \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \qquad \begin{array}{l} \lambda_1 = 4 \\ \\ \lambda_3 = 2 \\ \lambda_4 = 1 \end{array} \qquad \begin{bmatrix} 100 \\ \\ 0 \\ 0 \end{bmatrix} \quad \begin{bmatrix} 0 \\ \\ 0 \\ 0 \end{bmatrix} \quad \begin{bmatrix} 0 \\ \\ 42 \\ 0 \end{bmatrix} \quad \begin{bmatrix} 0 \\ \\ 0 \\ 1 \end{bmatrix}$$

Converse is not true. Check identity matrix .

New Theorem: Distinct Eigenvalues

- Let A be an $n \times n$ matrix with **complex** or **real** coefficients.

  If the **eigenvalues** $\{\lambda_1, \ldots, \lambda_n\}$ are distinct, that is $\lambda_i \neq \lambda_j$ for all $1 \leq i \neq j \leq n$, then the **eigenvectors** $\{v^1, \ldots, v^n\}$ are **linearly independent** in $(\mathbb{C}^n, \mathbb{C})$

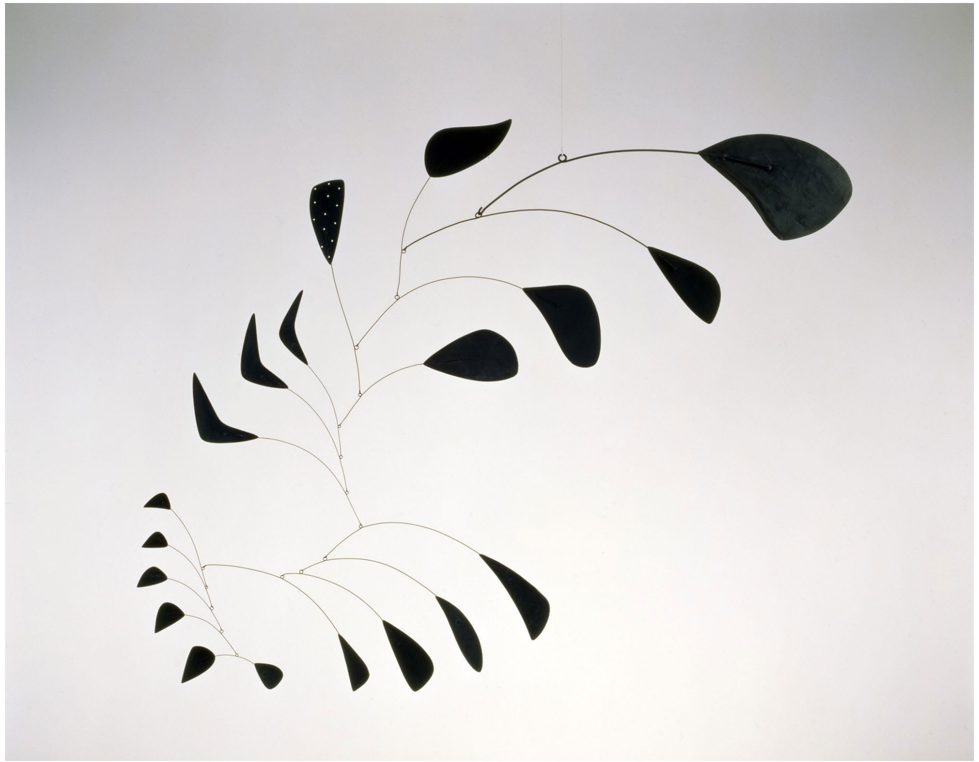
  Restatement of the theorem:
  If the **eigenvalues** $\{\lambda_1, \ldots, \lambda_n\}$ are distinct then the **eigenvectors** $\{v^1, \ldots, v^n\}$ are form a **basis** in $(\mathbb{C}^n, \mathbb{C})$

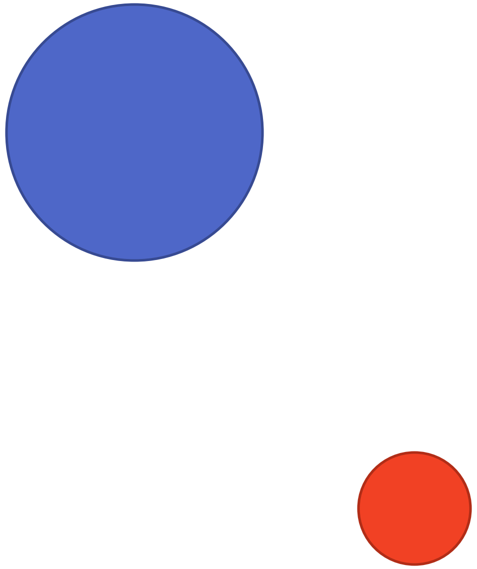
$$\begin{bmatrix} 4 & 0 & 0 & 0 \\ & & & \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad \begin{matrix} \lambda_1 = 4 \\ \\ \lambda_3 = 2 \\ \lambda_4 = 1 \end{matrix} \quad \begin{bmatrix} 100 \\ \\ 0 \\ 0 \end{bmatrix} \quad \begin{bmatrix} 0 \\ \\ 0 \\ 0 \end{bmatrix} \quad \begin{bmatrix} 0 \\ \\ 42 \\ 0 \end{bmatrix} \quad \begin{bmatrix} 0 \\ \\ 0 \\ 1 \end{bmatrix}$$

See Proof of Theorem 2.53 in the main text .

Similarity

- What does it mean for two things to be similar?
- Same underlying "essence"?





Similarity

- Suppose two square $n \times n$ matrices correspond to the same linear operator in different bases u and v

$$\begin{tikzcd} \mathcal{X} \arrow[r, "\mathcal{L}"] \arrow[d, "{[x]_u}"'] & \mathcal{Y} \arrow[d, "{[y]_u}"] \\ \mathcal{F}^n \arrow[r, "A"] & \mathcal{F}^n \end{tikzcd} \qquad \begin{tikzcd} \mathcal{X} \arrow[r, "\mathcal{L}"] \arrow[d, "{[x]_v}"'] & \mathcal{Y} \arrow[d, "{[y]_v}"] \\ \mathcal{F}^n \arrow[r, "B"] & \mathcal{F}^n \end{tikzcd}$$

Similarity

- Suppose two $n \times n$ matrices correspond to the same linear operator in different bases u and v

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

$$A[x]_u = [y]_u$$

$$B[x]_v = [y]_v$$

Similarity

- Suppose two $n \times n$ matrices correspond to the same linear operator in different bases u and v

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

$$A[x]_u = [y]_u$$

$$B[x]_v = [y]_v$$

$$P[\ \ ]_u = [\ \ ]_v$$

$$\bar{P}[\ \ ]_v = [\ \ ]_u$$

$$P^{-1} = \bar{P}$$

Similarity

- Suppose two $n \times n$ matrices correspond to the same linear operator in different bases u and v

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

$$A[x]_u = [y]_u$$

$$B[x]_v = [y]_v$$

$$BP[x]_u = P[y]_u$$

$$P^{-1}BP[x]_u = [y]_u$$

$$P[\ \ ]_u = [\ \ ]_v$$

$$\bar{P}[\ \ ]_v = [\ \ ]_u$$

$$P^{-1} = \bar{P}$$

$$\boxed{P^{-1}BP = A} \quad \textbf{OR} \quad \boxed{B = PAP^{-1}}$$

Definition (Not a proof): Similarity

- Two $n \times n$ matrices A and B are **similar** if there exists an invertible $n \times n$ matrix P such that $B = P \cdot \mathrm{A} \cdot P^{-1}$.

  P is called a **similarity matrix**.

Definitions: Full Set of Eigenvectors

- Two $n \times n$ matrices A and B are **similar** if there exists an invertible $n \times n$ matrix P such that $B = P \cdot \mathrm{A} \cdot P^{-1}$.

  P is called a **similarity matrix**.

- An $n \times n$ matrix A is said to have a **full set of eigenvectors** if there exists a **basis** $\{v^1, \ldots, v^n\}$ of $(\mathbb{C}^n, \mathbb{C})$ such that

$$Av^i = \lambda_i v^i, 1 \le i \le n$$

Theorem: Diagonalization

- Two $n \times n$ matrices A and B are **similar** if there exists an invertible $n \times n$ matrix P such that $B = P \cdot \mathrm{A} \cdot P^{-1}$.

  P is called a **similarity matrix**.

- An $n \times n$ matrix A is said to have a **full set of eigenvectors** if there exists a **basis** $\{v^1, \ldots, v^n\}$ of $(\mathbb{C}^n, \mathbb{C})$ such that

$$Av^i = \lambda_i v^i, 1 \leq i \leq n$$

- An $n \times n$ matrix A has a **full set of eigenvectors** if, and only if, it is **similar** to a **diagonal matrix**. And when this happens, the entries on the diagonal matrix are **eigenvalues** of A.

Proof: Diagonalization

Proof: We assume that $\{v^1, \ldots, v^n\}$ is a basis for $(\mathbb{C}^n, \mathbb{C})$ and that for $1 \leq i \leq n$, $Av^i = \lambda_i v^i$. Define two $n \times n$ matrices

$$M := \begin{bmatrix} v^1 & v^2 & \cdots & v^n \end{bmatrix}$$
$$\Lambda := \mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n).$$

Then

$$\begin{aligned} A \cdot M &:= \begin{bmatrix} Av^1 & Av^2 & \cdots & Av^n \end{bmatrix} \\ &= \begin{bmatrix} \lambda_1 v^1 & \lambda_2 v^2 & \cdots & \lambda_n v^n \end{bmatrix} \\ &= M \cdot \Lambda. \end{aligned}$$

We'll leave as an Exercise,

$$M\alpha = \begin{bmatrix} v^1 & v^2 & \cdots & v^n \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{bmatrix} = \alpha_1 v^2 + \alpha_2 v^2 + \cdots + \alpha_n v^n,$$

and hence M is invertible if, and only if, $\{v^1, \ldots, v^n\}$ is linearly independent. Therefore we have

$$A = M \Lambda M^{-1} \text{ and } \Lambda = M^{-1} A M,$$

proving that $\{v^1, \ldots, v^n\}$ is a basis implies A is similar to a diagonal matrix.

The other direction is straightforward and left to the reader. You need to recognize the columns of the "similarity matrix" as being e-vectors of A. ■

Similarity/Diagonalization

- Is the diagonalization unique?

Similarity/Diagonalization

- Is the diagonalization unique? **No. Take a look at how the diagonalization was made. The eigenvectors are not unique, and the eigenvalues don't have a preferred order from** 1 **to** n.
- If A and B are similar, they have the same eigenvalues. Moreover, their eigenvalues have the same **algebraic** and **geometric multiplicities**.

Miscellaneous Things about A

- For an $n \times m$ matrix A with coefficients in field $\mathcal{F}$.
  - **Definition: rank** is the number of linearly independent columns of A.
  - **Theorem:** $\mathrm{rank}(A) = \mathrm{rank}(A^T) = \mathrm{rank}(AA^T) = \mathrm{rank}(A^TA)$
    - Column rank = row rank, which is less than or equal to $\min(n,m)$
- For an $n \times n$ matrix A with coefficients in field $\mathcal{F}$.
  - **Definition:** A is symmetric if $A = A^T$
    - **Definition:** If its eigenvalues are also all positive, A is also **positive definite**
  - **Definition:** Trace of matrix A or $\mathrm{tr}(A)$ is the sum of the diagonal entries

$$\mathrm{tr}(A) := \sum_{i=1}^{n} A_{ii}$$

  - See **2.8 A Few Additional Properties of Matrices** in the book for the rest

After the exam

Norms and Normed Spaces

- We finally arrive at a notion of **distance**.
- We are used to Pythagorean distance and dot product in 3-D space.
- **Goals:**
    - Generalize **distance** $d(x,y)$ for any vector space (can be complex)
        - This will be very useful when talking about **errors** (think least squares)
    - Generalize dot product to the **inner product** $\langle x,y \rangle$ for any vector space (can be complex)

Distance

- We can define distance by equipping a vector space with a special function, the **norm**

$$d(x,y) := \|x-y\|$$

Definition 3.1 *Let* $(\mathcal{X}, \mathcal{F})$ *be a vector space where the field* $\mathcal{F}$ *is either* $\mathbb{R}$ *or* $\mathbb{C}$*. A function* $\|\cdot\|: \mathcal{X} \to \mathbb{R}$ *is a* ***norm*** *if it satisfies*

(a) $\|x\| \geq 0, \forall x \in \mathcal{X}$ *and* $\|x\| = 0 \iff x = 0.$

(b) Triangle inequality: $\|x+y\| \leq \|x\| + \|y\|, \forall x, y \in \mathcal{X}$

(c) $\|\alpha x\| = |\alpha| \cdot \|x\|, \forall x \in \mathcal{X}, \alpha \in \mathcal{F}, \begin{cases} \text{If } \alpha \in \mathbb{R}, |\alpha| \text{ means the absolute value} \\ \text{If } \alpha \in \mathbb{C}, |\alpha| \text{ means the magnitude} \end{cases}.$

Norms

(a) $\mathcal{F} := \mathbb{R}$ *or* $\mathbb{C}$, $\mathcal{X} := \mathcal{F}^n$.

- $\|x\|_2 := \left(\sum_{i=1}^{n} |x_i|^2\right)^{\frac{1}{2}}$, *Euclidean norm or 2-norm*
- $\|x\|_p := \left(\sum_{i=1}^{n} |x_i|^p\right)^{\frac{1}{p}}$, $1 \leq p < \infty$, *p-norm*
- $\|x\|_\infty := \max\limits_{1 \leq i \leq n} |x_i|$, *max-norm, sup-norm, ∞-norm*

(b) $\mathcal{F} := \mathbb{R}$, $\mathcal{D} \subset \mathbb{R}$, $\mathcal{D} := [a, b]$, $a < b < \infty$, *and* $\mathcal{X} := \{f : \mathcal{D} \to \mathbb{R} \mid f \text{ is continuous}\}$.

- $\|f\|_2 := (\int_a^b |f(t)|^2 \,\mathrm{d}t)^{\frac{1}{2}}$
- $\|f\|_p := (\int_a^b |f(t)|^p \,\mathrm{d}t)^{\frac{1}{p}}$, $1 \leq p < \infty$
- $\|f\|_\infty := \max\limits_{a \leq t \leq b} |f(t)|$, *which is also written* $\|f\|_\infty := \sup\limits_{a \leq t \leq b} |f(t)|$

Norms

(a) $\mathcal{F} := \mathbb{R}$ *or* $\mathbb{C}$, $\mathcal{X} := \mathcal{F}^n$.

- $\boxed{\|x\|_2 := \left(\sum_{i=1}^{n} |x_i|^2\right)^{\frac{1}{2}}, \text{ Euclidean norm or 2-norm}}$
- $\|x\|_p := \left(\sum_{i=1}^{n} |x_i|^p\right)^{\frac{1}{p}}$, $1 \le p < \infty$, *p-norm*
- $\|x\|_\infty := \max\limits_{1\le i\le n} |x_i|$, *max-norm, sup-norm,* ∞*-norm*

(b) $\mathcal{F} := \mathbb{R}$, $\mathcal{D} \subset \mathbb{R}$, $\mathcal{D} := [a, b]$, $a < b < \infty$, *and* $\mathcal{X} := \{f : \mathcal{D} \to \mathbb{R} \mid f \text{ is continuous}\}$.

- $\|f\|_2 := (\int_a^b |f(t)|^2 \,\mathrm{d}t)^{\frac{1}{2}}$
- $\|f\|_p := (\int_a^b |f(t)|^p \,\mathrm{d}t)^{\frac{1}{p}}$, $1 \le p < \infty$
- $\|f\|_\infty := \max\limits_{a\le t\le b} |f(t)|$, *which is also written* $\|f\|_\infty := \sup\limits_{a\le t\le b} |f(t)|$

Inner Product

Definition 3.11 *Let* $(\mathcal{X}, \mathbb{C})$ *be a vector space. A function* $\langle \cdot\,, \cdot \rangle : \mathcal{X} \times \mathcal{X} \to \mathbb{C}$ *is an inner product if*

(a) $\langle a, b \rangle = \overline{\langle b, a \rangle}$.

(b) $\langle \alpha_1 x_1 + \alpha_2 x_2, y \rangle = \alpha_1 \langle x_1, y \rangle + \alpha_2 \langle x_2, y \rangle$, *linear in the left argument. Some books place the linearity on the right side.*

(c) $\langle x, x \rangle \geq 0$ *for any* $x \in \mathcal{X}$, *and* $\langle x, x \rangle = 0 \iff x = 0$. *(See below:* $\langle x, x \rangle$ *is a real number and therefore it can be compared to 0.)*

Inner Product

$$(a)\ \ (\mathbb{C}^n, \mathbb{C}),\ \langle x, y\rangle := x^\top \overline{y} = \sum_{i=1}^{n} x_i \overline{y_i}.$$

$$(b)\ \ (\mathbb{R}^n, \mathbb{R}),\ \langle x, y\rangle := x^\top y = \sum_{i=1}^{n} x_i y_i.$$

$$(c)\ \ \mathcal{F} = \mathbb{R},\ \mathcal{X} = \{A \,|\, n \times m \textit{ real matrices}\},\ \langle A, B\rangle := \mathrm{tr}(AB^\top) = \mathrm{tr}(A^\top B).$$

$$(d)\ \ \mathcal{X} = \{f : [a, b] \to \mathbb{R},\ f \textit{ continuous}\},\ \mathcal{F} = \mathbb{R},\ \langle f, g\rangle := \int_a^b f(t)g(t)\,\mathrm{d}t.$$

Inner Product

$$(a)\ \ (\mathbb{C}^n, \mathbb{C}),\ \langle x, y\rangle := x^\top \overline{y} = \sum_{i=1}^{n} x_i \overline{y_i}.$$

$$\boxed{(b)\ \ (\mathbb{R}^n, \mathbb{R}),\ \langle x, y\rangle := x^\top y = \sum_{i=1}^{n} x_i y_i.}$$

$$(c)\ \ \mathcal{F} = \mathbb{R},\ \mathcal{X} = \{A \,|\, n \times m \textit{ real matrices}\},\ \langle A, B\rangle := \operatorname{tr}(AB^\top) = \operatorname{tr}(A^\top B).$$

$$(d)\ \ \mathcal{X} = \{f : [a, b] \to \mathbb{R},\ f \textit{ continuous}\},\ \mathcal{F} = \mathbb{R},\ \langle f, g\rangle := \int_a^b f(t) g(t)\, \mathrm{d}t.$$

Supremum and Infimum

- Supremum is the least upper bound
- Infimum is the greatest lower bound
- Makes sense for any **partially ordered set**
- Real numbers are **totally ordered** and complex numbers are **not ordered**
- Principle of Mathematical Induction is somehow connected to the **well-ordering principle**: *every non-empty set of positive integers contains a least element*