Lecture 37: §4.7 Similarities & Diagonalization

Recall: An $n\times n$ real matrix A is diagonalizable over $\mathbb{C}$ if $A \sim D$ where $D = \begin{pmatrix} d_1 & & 0 \\ & \ddots & \\ 0 & & d_n \end{pmatrix}$ is a diagonal matrix in $\mathrm{Mat}_{n\times n}(\mathbb{C})$

So $D = S^{-1} A S$ for S in $\mathrm{Mat}_{n\times n}(\mathbb{C})$ nonsingular ($\det S \neq 0$)

- $\{d_1, \dots, d_n\}$ are the eigenvalues of A (with possible repetitions)
- If $S = [\vec{v}_1, \dots \vec{v}_n]$ then $A\vec{v}_1 = d_1 \vec{v}_1$, $A\vec{v}_2 = d_2\vec{v}_2$, $\vdots$ $A\vec{v}_n = d_n\vec{v}_n$ & $\{\vec{v}_1, \dots, \vec{v}_n\}$ is a basis for $\mathbb{C}^n$ (basis of eigenvectors!)
- S invertible so $S^{-1} = \dfrac{\mathrm{Cof}(S)^T}{\det(S)}$, also $(S \mid I_n) \overset{GJ}{\sim} (I_n \mid S^{-1})$
  $\mathbb{C}$ row-operations

Conclusion: A is diagonalizable over $\mathbb{C}$ means $\mathbb{C}^n$ has a basis of eigenvectors

Example: $A = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$ $\quad P_A(\lambda) = \det\begin{pmatrix} -\lambda & 1 \\ -1 & -\lambda \end{pmatrix} = \lambda^2 + 1 = (\lambda - i)(\lambda + i) \rightsquigarrow \lambda = \pm i$ eigenvalues

$E_i(A) = \mathcal{N}(A - iI) = \mathcal{N}\begin{pmatrix} -i & 1 \\ -1 & -i \end{pmatrix} = \mathcal{N}\begin{pmatrix} 1 & i \\ 0 & 0 \end{pmatrix} = \mathrm{Sp}\left(\begin{bmatrix} -i \\ 1 \end{bmatrix}\right)$

$\rightsquigarrow E_{-i}(A) = \overline{E_i(A)} = \mathrm{Sp}\left(\begin{bmatrix} +i \\ 1 \end{bmatrix}\right)$ so $D = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}$ $\quad S = \begin{bmatrix} -i & +i \\ 1 & 1 \end{bmatrix} \rightsquigarrow S^{-1} = \frac{1}{-2i}\begin{bmatrix} 1 & -i \\ -1 & -i \end{bmatrix} = \begin{bmatrix} i/2 & 1/2 \\ -i/2 & 1/2 \end{bmatrix}$

Real Symmetric Matrices

Diagonalizable over $\mathbb{R}$!

Theorem: Fix A $n\times n$ real symmetric matrix. Then:

1. A has only real eigenvalues $\lambda_1, \dots, \lambda_n$ (possibly repeated!) L35
2. $\mathbb{R}^n$ has an orthonormal basis B of eigenvectors for A

that is, $B\{\vec{v}_1, \dots, \vec{v}_n\}$ with v_i in $E_{\lambda_i}(A)$. and $v_i \cdot v_j = \begin{cases} 0 & i\neq j \\ 1 & i=j \end{cases}$

3. $\begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{bmatrix} = Q^T A Q$ with $Q = [\vec{v}_1 \cdots \vec{v}_n]$

($Q^{-1} = Q^T$ because B is orthonormal!)

Example: $A = \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} = A^T$

$p_A(\lambda) = \det \begin{pmatrix} 1-\lambda & -1 \\ -1 & 1-\lambda \end{pmatrix} = (1-\lambda)^2 - 1 = \lambda(\lambda-2)$

$E_0 = \mathcal{N}(A) = \mathcal{N}\begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} = \mathrm{Sp}\left(\begin{bmatrix} 1 \\ 1 \end{bmatrix}\right) = \mathrm{Sp}\left(\begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}\right)$

$E_2 = \mathcal{N}(A - 2I_2) = \mathcal{N}\begin{pmatrix} -1 & -1 \\ -1 & -1 \end{pmatrix} = \mathcal{N}\begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix} = \mathrm{Sp}\left(\begin{bmatrix} 1 \\ -1 \end{bmatrix}\right) = \mathrm{Sp}\left(\begin{bmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \end{bmatrix}\right)$

norm $= 1$

Check: $\begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix} \cdot \begin{bmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \end{bmatrix} = 0$

Take $Q = \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & -1/\sqrt{2} \end{bmatrix}$ $C = \begin{bmatrix} 0 & 0 \\ 0 & 2 \end{bmatrix}$.

$\boxed{\text{Why diagonalize?}}$

A: We can perform power operations very fast!

Proposition: ① If $D = S^{-1}AS$ $(A \sim D)$, then $A^k = SD^kS^{-1}$ for all $k = 1, 2, \ldots$

② If $D = \begin{pmatrix} d_1 & & 0 \\ & \ddots & \\ 0 & & d_n \end{pmatrix}$, then $D^k = \begin{pmatrix} d_1^k & & 0 \\ & \ddots & \\ 0 & & d_n^k \end{pmatrix}$ for all $k = 1, 2, \ldots$

If D is invertible, the same formulas work for $k = -1, -2, \ldots$

Why? ① $D = S^{-1}AS \rightsquigarrow SDS^{-1} = A$

- $A^2 = A \cdot A = SD\underbrace{S^{-1}S}_{=I_n}DS = SD^2S$
- $A^3 = A \cdot A^2 = SD\underbrace{S^{-1}S}_{=I_n}D^2S = SD^3S$, etc.

In general

$A^k = SD^kS^{-1}$

② $D^2 = \begin{pmatrix} d_1 & & 0 \\ & \ddots & \\ 0 & & d_n \end{pmatrix} \begin{pmatrix} d_1 & & 0 \\ & \ddots & \\ 0 & & d_n \end{pmatrix} = \begin{pmatrix} d_1^2 & & & \\ & d_2^2 & & \\ & & \ddots & \\ & & & d_n^2 \end{pmatrix}$

$D^3 = DD^2 = \begin{pmatrix} d_1 & & \\ & \ddots & \\ & & d_n \end{pmatrix} \begin{pmatrix} d_1^2 & & \\ & \ddots & \\ & & d_n^2 \end{pmatrix} = \begin{pmatrix} d_1^3 & & \\ & \ddots & \\ & & d_n^3 \end{pmatrix}$ etc.

- If D is invertible, all $d_i \neq 0$ so $D^{-1} = \begin{pmatrix} d_1^{-1} & & \\ & \ddots & \\ & & d_n^{-1} \end{pmatrix}$ is also diagonal.

& $A^{-1} = (SDS^{-1}) = (S^{-1})^{-1}D^{-1}S^{-1} = SD^{-1}S^{-1}$ Same ideas work for $A^{-1} \sim D^{-1}$

Example (last time) $A = \begin{bmatrix} 5 & -6 \\ 3 & -4 \end{bmatrix}$ $\quad p_A(\lambda) = (\lambda-2)(\lambda+1)$ $\quad E_{-1} = Sp\left(\begin{bmatrix} 1 \\ 1 \end{bmatrix}\right)$

$E_2 = Sp\left(\begin{bmatrix} 2 \\ 1 \end{bmatrix}\right)$

$$D = \begin{bmatrix} -1 & 0 \\ 0 & 2 \end{bmatrix}, \quad S = \begin{bmatrix} 1 & 2 \\ 1 & 1 \end{bmatrix}, \quad S^{-1} = \frac{1}{1-2}\begin{bmatrix} 1 & -2 \\ -1 & 1 \end{bmatrix} = \begin{bmatrix} -1 & 2 \\ 1 & -1 \end{bmatrix}$$

basis for E_{-1} $\quad$ basis for E_2

$D = S^{-1}AS$, so $A = SDS^{-1}$

① Compute A^{-1}: $A^{-1} = (SDS^{-1})^{-1} = SD^{-1}S^{-1} = \begin{bmatrix} 1 & 2 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} -1 & 0 \\ 0 & 1/2 \end{bmatrix}\begin{bmatrix} -1 & 2 \\ 1 & -1 \end{bmatrix}$

$$= \begin{bmatrix} 1 & 2 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} 1 & -2 \\ 1/2 & -1/2 \end{bmatrix} = \begin{bmatrix} 1+1 & -2-1 \\ 1+1/2 & -2-1/2 \end{bmatrix} = \begin{bmatrix} 2 & -3 \\ 3/2 & -5/2 \end{bmatrix}$$

Check $\begin{bmatrix} 5 & -6 \\ 3 & -4 \end{bmatrix}\begin{bmatrix} 2 & -3 \\ 3/2 & -5/2 \end{bmatrix} = \begin{bmatrix} 10-9 & -15+15 \\ 6-6 & -9+10 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ ✓

② Compute A^{10}: $A^{10} = SD^{10}S^{-1} = \begin{bmatrix} 1 & 2 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & 2^{10} \end{bmatrix}\begin{bmatrix} -1 & 2 \\ 1 & -1 \end{bmatrix}$

$$= \begin{bmatrix} 1 & 2^{11} \\ 1 & 2^{10} \end{bmatrix}\begin{bmatrix} -1 & 2 \\ 1 & -1 \end{bmatrix} = \begin{bmatrix} -1+2^{11} & 2-2^{11} \\ -1+2^{10} & 2-2^{10} \end{bmatrix} = \begin{bmatrix} 2047 & -2046 \\ 1023 & -1022 \end{bmatrix}$$

③ Compute $(A^{-1})^{10}$: $A^{-10} = S(D^{-1})^{10}S^{-1} = \begin{bmatrix} 1 & 2 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & 1/2^{10} \end{bmatrix}\begin{bmatrix} -1 & 2 \\ 1 & -1 \end{bmatrix}$

$$= \begin{bmatrix} 1 & 1/2^9 \\ 1 & 1/2^{10} \end{bmatrix}\begin{bmatrix} -1 & 2 \\ 1 & -1 \end{bmatrix} = \begin{bmatrix} \frac{-2^9+1}{2^9} & \frac{2^{10}-1}{2^9} \\ \frac{-2^{10}+1}{2^{10}} & \frac{2^{11}-1}{2^{10}} \end{bmatrix} = \begin{bmatrix} \frac{-511}{512} & \frac{1023}{512} \\ \frac{-1023}{1024} & \frac{2047}{1024} \end{bmatrix}$$

Diagonalization over $\mathbb{R}$ vs over $\mathbb{C}$

Fix A an $n\times n$ matrix with entries in $\mathbb{R}$.

Theorem: ① A is diagonalizable over $\mathbb{R}$ if, and only if, all eigenvalues of A are real & $\mathbb{R}^n$ has a basis of eigenvectors for A.

- If $\{\lambda_1, \ldots, \lambda_r\}$ are the eigenvalues, then each $E_{\lambda_i}(A)$ is a subspace of $\mathbb{R}^n$. We have a basis of eigenvectors if, and only if,

$$\dim_{\mathbb{R}} E_{\lambda_1}(A) + \cdots + \dim_{\mathbb{R}} E_{\lambda_r}(A) = n$$

Equivalently: $\dim_{\mathbb{R}} E_{\lambda_i}(A) = \mathrm{mult}(\lambda_i, P_A)$ for all $i = 1, \ldots, r$ (A non-defective)

② A is diagonalizable over $\mathbb{C}$ if, and only if, $\mathbb{C}^n$ has a basis of eigenvectors for A. This happens if and only if

$$\dim_{\mathbb{C}} E_\lambda(A) = \mathrm{mult}(\lambda, P_A) \text{ for each eigenvalue } \lambda.$$

For ②: View $E_\lambda(A)$ as a subspace of $\mathbb{C}^n$ even if λ is real

[If λ is real, we can pick a basis for $E_\lambda(A)$ using vectors in $\mathbb{R}^n$.

- Diagonalization of A over $\mathbb{C}$ $\underset{\textcircled{2}}{\overset{\textcircled{1}}{\rightleftarrows}}$ basis for $\mathbb{C}^n$ with eigenvectors

Why? ① $A = S\,D\,S^{-1}$ $\underset{L36}{\rightsquigarrow}$ Eigenvalues of A are diagonal entries of D

Write $S = [\vec{v}_1, \ldots, \vec{v}_n]$. Then: $S^{-1} v_i = S^{-1}\,\mathrm{Col}_i(S) = \mathrm{Col}_i(\underbrace{S^{-1}S}_{=I_n}) = e_i$

So $A\vec{v}_i = S\,D\,S^{-1}\vec{v}_i = S\,D\,e_i = S(d_i e_i) = d_i \underbrace{S e_i}_{=\vec{v}_i} = d_i \vec{v}_i$

So $\vec{v}_i$ is an eigenvector for A

S is nonsingular means the columns are li, so $\{\vec{v}_1, \ldots, \vec{v}_n\}$ is a basis for $\mathbb{C}^n$ (li + size = dim $\mathbb{C}^n$)

② Pick $B = \{\vec{v}_1, \ldots, \vec{v}_n\}$ basis for $\mathbb{C}^n$ using eigenvectors $[A v_i = \lambda v_i]$

Then $S = [\vec{v}_1, \ldots, \vec{v}_n]$ is nonsingular. Pick $i = 1, \ldots, n$:

$\mathrm{Col}_i(S^{-1}AS) = S^{-1}A\,\mathrm{Col}_i(S) = S^{-1}A\vec{v}_i = S^{-1}\lambda_i\vec{v}_i = \lambda_i S^{-1}\mathrm{Col}_i S = \lambda_i e_i$

So $S^{-1}AS = \begin{pmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_n \end{pmatrix}$ is diagonal!

Conclude: A is diagonalizable

- From $\mathbb{C}$ to $\mathbb{R}$: eigenvalues λ_i in $\mathbb{R}$ & E_{λ_i}: computations can be done in $\mathbb{R}^n$.

$\boxed{\text{Diagonalizable} = \text{non-defective}}$ A in $\mathrm{Mat}_{n\times n}(\mathbb{R})$

[L33]

Proposition: Pick $\{\lambda_1, \ldots, \lambda_p\}$ distinct eigenvalues for A & $S = \{\vec{v}_1, \ldots, \vec{v}_p\}$ with $\vec{v}_i \neq \vec{0}$ & $A\vec{v}_i = \lambda_i \vec{v}_i$. Then S is li

Consequence: If $\{\lambda_1, \ldots, \lambda_r\}$ are all the eigenvalues of A (in $\mathbb{C}$) & $\dim_{\mathbb{C}} E_{\lambda_i}(A) = \mathrm{mult}(\lambda_i, P_A)$ for $i = 1, \ldots, r$, then $\mathbb{C}^n$ has a basis of eigenvectors.
$\leq$ (always)

Why? If $m_i = \dim_{\mathbb{C}} E_{\lambda_i}(A)$ then $m_1 + m_2 + \cdots + m_r = \deg(P_A) = n$.
$\leq$ (always!)

$$\left.\begin{array}{ll} B_1 = \{\vec{v}_1^{(1)}, \ldots, \vec{v}_{m_1}^{(1)}\} & \text{basis for } E_{\lambda_1} \\ \vdots & \\ B_r = \{\vec{v}_1^{(r)}, \ldots, \vec{v}_{m_r}^{(r)}\} & \text{———} \quad E_{\lambda_r} \end{array}\right]$$

Claim $B_1 \cup \cdots \cup B_r$ is a basis for $\mathbb{C}^n$

Show B is li: (*) $$\vec{0} = \underbrace{\alpha_{11}\vec{v}_1^{(1)} + \cdots + \alpha_{1m_1}\vec{v}_{m_1}^{(1)}}_{=\vec{v}_1 \text{ in } E_{\lambda_1}} + \cdots + \underbrace{\alpha_{r1}v_1^{(r)} + \cdots + \alpha_{rm_r}v_{m_r}^{(r)}}_{\vec{v}_r \text{ in } E_{\lambda_r}}$$

- If $\vec{v}_i = \vec{0}$ then $\alpha_{i1} = \cdots = \alpha_{im_i} = 0$ (B_i is li)
- After removing the $\vec{v}_i$'s that are $\vec{0}$ from (*), the remaining ones will be li (their sum is $\vec{0}$!) This contradicts the Proposition, so all $\vec{v}_i = 0$

Example: $A = \begin{bmatrix} 25 & -8 & 30 \\ 24 & -7 & 30 \\ -12 & 4 & -14 \end{bmatrix} \rightsquigarrow P_A(\lambda) = -\lambda^3 + 4\lambda^2 - 5\lambda + 2$
$= -(\lambda-1)^2(\lambda-2)$

Eigenvalues: $\lambda = 1$ with mult 2 & $\lambda = 2$ with mult 1

- We compute E_1 & E_2 to decide if A is diagnalizable or not! (E_1 is key!)

- $E_1 = \mathcal{N}(A - I_3) = \mathcal{N}\begin{pmatrix} 24 & -8 & 30 \\ 24 & -8 & 30 \\ -12 & 4 & -15 \end{pmatrix} = \mathcal{N}\begin{pmatrix} 24 & -8 & 30 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} = \mathcal{N}\begin{pmatrix} 1 & -1/3 & 5/4 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$ REF

x_1 dep, x_2 & x_3 indep

$x_1 = \frac{1}{3}x_2 - \frac{5}{4}x_3$

$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} \frac{1}{3}x_2 - \frac{5}{4}x_3 \\ x_2 \\ x_3 \end{bmatrix} = x_2 \begin{bmatrix} 1/3 \\ 1 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} -5/4 \\ 0 \\ 1 \end{bmatrix}$

$E_1 = Sp\left(\begin{bmatrix} 1/3 \\ 1 \\ 0 \end{bmatrix} \begin{bmatrix} -5/4 \\ 0 \\ 1 \end{bmatrix} \right) = Sp\left(\begin{bmatrix} 1 \\ 3 \\ 0 \end{bmatrix}, \begin{bmatrix} -5 \\ 4 \\ 0 \end{bmatrix} \right)$ $\dim E_1(A) = 2$.

- $E_2 = \mathcal{N}(A - 2I_3) = \mathcal{N}\begin{pmatrix} 23 & -8 & 30 \\ 24 & -9 & 30 \\ -12 & 4 & -16 \end{pmatrix} = \mathcal{N}\begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{pmatrix} = \mathcal{N}\begin{pmatrix} 1 & 0 & 2 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{pmatrix} = Sp\left(\begin{bmatrix} -2 \\ -2 \\ 1 \end{bmatrix}\right)$

$\dim E_1 = 2 = \text{mult}(1, P_A)$ & $\dim E_2 = 1 = \text{mult}(2, P_A)$ so A is diagnalizable

$D = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix}$ $S = \begin{bmatrix} 1 & -5 & -2 \\ 3 & 4 & -2 \\ 0 & 0 & 1 \end{bmatrix}$ $S^{-1} = \begin{bmatrix} 4/19 & 5/19 & 18/19 \\ -3/19 & 1/19 & -4/19 \\ 0 & 0 & 1 \end{bmatrix}$ ($S = \begin{bmatrix} -5 & 1 & -2 \\ 4 & 3 & -2 \\ 0 & 0 & 1 \end{bmatrix}$ would also work!)

$D = S^{-1}AS$ (diagonalizable over $\mathbb{R}$)