Lecture XXXIV: §4.7 Similarity Transformations & Diagonalization

Recall: Last time we defined defective / non-defective matrices, by comparing algebraic & geometric multiplicities of eigenvalues (roots of $P_A(t)$)

A $n\times n$ matrix. Then: $\boxed{\lambda \text{ eigenvalue of } A \longleftrightarrow E_\lambda(A) = \mathcal{N}(A-\lambda\,Id) \neq \{\vec{0}\}}$

- Two different multiplicities for eigenvalues:

$\begin{cases} \text{• alg multiplicity of } \lambda = \text{number of times } \lambda \text{ appears as a root of } P_A(t) \\ \text{• geometric multiplicity of } \lambda = \dim E_\lambda(A) \end{cases}$

Key inequalities: $\boxed{1 \leq \text{Geom. mult } \lambda \leq \text{Alg Mult of } \lambda}$ (λ eigenvalue of A)

Def: $\underset{n\times n}{A}$ non defective if Geom mult = alg mult for all eigenvalues of A

Ex: $A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ is non-defective $P_A(t) = (t-1)^2$ & $E_1(A) = \mathbb{R}^2$

Alg mult of 1 = geom mult of 1 = 2

$A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$ is defective $P_A(t) = (t-1)^2$ & $E_1(A) = \mathrm{Sp}\left(\begin{bmatrix} 1 \\ 0 \end{bmatrix}\right)$

Alg mult of 1 is 2 but Geom mult = 1

Definition $\underset{n\times n}{A}$ is diagonalizable if $\mathbb{R}^n$ has a basis B consisting of eigenvectors for A

Equivalently, we can find a diagonal matrix $D = \begin{bmatrix} d_1 & & 0 \\ & \ddots & \\ 0 & & d_n \end{bmatrix}$ and an invertible matrix S with $S^{-1}AS = D$

In this situation, columns of S are the vectors in B & $d_1 \ldots d_n$ are the eigenvalues for each vector $\vec{v}_1, \ldots, \vec{v}_n$ of B (in the same order as ds!)

Def. Two $n\times n$ matrices A, C are called similar if $C = S^{-1}AS$ for some $\underset{n\times n}{S}$

This says A & C represent the same linear transf $T: \mathbb{R}^n \to \mathbb{R}^n$, $\vec{x} \mapsto A\vec{x}$, in different bases

$C = [T]_{BB}$ where B = columns of S & $A = [T]_{EE}$

TODAY: ① A non-defective & with only real eigenvalues is ALWAYS diagonalizable

② Symmetric matrices have only real eigenvalues & are diagonalizable

Furthermore eigenspaces are mutually orthogonal

§1. Linear independence of eigenvectors:

THEOREM 1: Fix A an $n\times n$ matrix and a list of k distinct eigenvalues $\{\lambda_1, \ldots, \lambda_k\}$. Pick one eigenvector $\vec{v_j} \neq \vec{0}$ for each j (ie $A\vec{v_j} = \lambda_j \vec{v_j}$). Then, $S = \{\vec{v_1}, \ldots, \vec{v_k}\}$ is linearly indep in $\mathbb{R}^n$

Why? • $k=1$ $\quad \vec{v_1} \neq \vec{0}$ so $S = \{\vec{v_1}\}$ is lin. indep.

• Assume $k>1$ & argue by contradiction. Suppose $S = \{\vec{v_1}, \ldots, \vec{v_k}\}$ is lin. dep. After reordering the eigenvalues, we can find $2 \leq m \leq k$ with $\{\vec{v_1}, \ldots, \vec{v_{m-1}}\}$ linearly indep.

BUT $\{\vec{v_1}, \ldots, \vec{v_m}\}$ ——— dep

1 vector — k vectors

li — ld

at some point we go from li to ld

Write the dependency relation for $\{\vec{v_1}, \ldots, \vec{v_m}\}$

$$(1) \quad \boxed{a_1\vec{v_1} + \cdots + a_m\vec{v_m} = \vec{0}} \qquad \text{with } a_m \neq 0$$

Multiply by A on the left, on both sides of $=$:

$$A(a_1\vec{v_1} + \cdots + a_m\vec{v_m}) = A\vec{0}$$

(use distribution)

$$a_1\underbrace{A\vec{v_1}}_{=\lambda_1\vec{v_1}} + \cdots + a_m\underbrace{A\vec{v_m}}_{=\lambda_m\vec{v_m}} = \vec{0}$$

$$(2) \quad \boxed{a_1\lambda_1\vec{v_1} + \cdots + a_m\lambda_m\vec{v_m} = \vec{0}}$$

Next: compute $(2) - \lambda_m(1)$:

$$\begin{array}{rl} & a_1\lambda_1\vec{v_1} + a_2\lambda_2\vec{v_2} + \cdots + a_m\lambda_m\vec{v_m} = \vec{0} \\ - & a_1\lambda_m\vec{v_1} + a_2\lambda_m\vec{v_2} + \cdots + a_m\lambda_m\vec{v_m} = \vec{0} \\ \hline & a_1(\lambda_1-\lambda_m)\vec{v_1} + a_2(\lambda_2-\lambda_m)\vec{v_2} + \cdots + \vec{0} = \vec{0} \end{array}$$

(substract column by column)

We get: $\underbrace{a_1(\lambda_1-\lambda_m)}_{=b_1}\vec{v}_1+\underbrace{a_2(\lambda_2-\lambda_m)}_{=b_2}\vec{v}_2+\cdots+\underbrace{a_{m-1}(\lambda_{m-1}-\lambda_m)}_{=b_{m-1}}\vec{v}_{m-1}=\vec{0}$

But $\{\vec{v}_1,\ldots,\vec{v}_{m-1}\}$ is li, so $b_1=b_2=\cdots=b_{m-1}=0$

Now $\lambda_1-\lambda_m\neq 0$, $\lambda_2-\lambda_m\neq 0$, ..., $\lambda_{m-1}-\lambda_m\neq 0$, so

$a_1=a_2=\cdots=a_{m-1}=0$

Looking back at (1), we get $\vec{0}+a_m\vec{v}_m=\vec{0}$ & $\vec{v}_m\neq\vec{0}$,

so $a_m=0$. This contradicts our assumption on $a_m\neq 0$.

Conclusion: S must be linearly independent.

Consequence: Assume $\underset{n\times n}{A}$ is non-defective & all its eigenvalues are real. Then A is diagonalizable.

Why? Collect all the eigenvalues of A $\{\lambda_1,\ldots,\lambda_k\}$

& write their algebraic multiplicities $m_1,\ldots,m_k$

Note: $m_1+\cdots+m_k=$ degree of $P_A=n$

Pick $B_1=$ basis for $E_{\lambda_1}(A)$ (m_1 elements since $m_1=\dim E_{\lambda_1}(A)$)

$B_2=$ ——— $E_{\lambda_2}(A)$ ($m_2=\dim E_{\lambda_2}(A)$ elements)

$\vdots$ $\vdots$

$B_k=$ ——— $E_{\lambda_k}(A)$ ($m_k=\dim E_{\lambda_k}(A)$ ———)

Write $B=B_1\cup B_2\cup\cdots\cup B_k$ (vectors in $\mathbb{R}^n$)

- size of $B=m_1+m_2+\cdots+m_k=n$
- Claim(*): B is linearly independent

Conclude: B is a basis for $\mathbb{R}^n$. All its elements are eigenvectors

(*) Proof of Claim: Write $B_1 = \{\vec{v}_1^{(1)}, \dots, \vec{v}_{m_1}^{(1)}\}$

$B_2 = \{\vec{v}_1^{(2)}, \dots, \vec{v}_{m_2}^{(2)}\}$

$\vdots$

$B_k = \{\vec{v}_1^{(1)}, \dots, \vec{v}_{m_k}^{(k)}\}$

& consider a dependency relation for these n vectors

$$\vec{0} = \underbrace{(a_{11}\vec{v}_1^{(1)} + \dots + a_{1m_1}\vec{v}_{m_1}^{(1)})}_{=\vec{v}_1 \text{ in } Sp(B_1) = E_{\lambda_1}(A)} + \underbrace{(a_{21}\vec{v}_1^{(2)} + \dots + a_{2m_2}\vec{v}_{m_2}^{(2)})}_{=\vec{v}_2 \text{ in } Sp(B_2) = E_{\lambda_2}(A)} + \dots$$
$$\dots + \underbrace{(a_{k1}\vec{v}_1^{(k)} + \dots + a_{km_k}\vec{v}_{m_k}^{(k)})}_{=\vec{v}_k \text{ in } Sp(B_k) = E_{\lambda_k}(A)}$$

So we get $\vec{0} = \vec{v}_1 + \vec{v}_2 + \dots + \vec{v}_k$ and $\vec{v}_1, \dots, \vec{v}_k$ are eigenvectors for different eigenvalues. By our Theorem above, the only way this relation holds is if $\vec{v}_1 = \vec{v}_2 = \dots = \vec{v}_k = \vec{0}$

But now we get k dependency relations, one for each basis $B_1, \dots, B_k$

$\vec{0} = \vec{v}_1 = a_{11}\vec{v}_1^{(1)} + \dots + a_{1m_1}\vec{v}_{m_1}^{(1)}$ relation for B_1

$\vdots$

$\vec{0} = \vec{v}_k = a_{k1}\vec{v}_1^{(k)} + \dots + a_{km_k}\vec{v}_{m_k}^{(k)}$ ——— B_k

Since $B_1, \dots, B_k$ are each l.i., we conclude all our coefficients $a_{11}, \dots, a_{1m_1}, a_{21}, \dots, a_{k1}, \dots, a_{km_k}$ are 0, so B is l.i.

Example: (last time) $A = \begin{bmatrix} 1 & 0 & -2 \\ 1 & 3 & 1 \\ 1 & 3 & 1 \end{bmatrix}$ $E_0 = Sp\left(\begin{bmatrix} 2 \\ -1 \\ 1 \end{bmatrix}\right)$, $E_2 = Sp\left(\begin{bmatrix} -2 \\ 1 \\ 1 \end{bmatrix}\right)$, $E_3 = Sp\left(\begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix}\right)$

non-defective & $\lambda = 0, 2, 3$ eigenvalues

Then $B = \left\{\begin{bmatrix} 2 \\ -1 \\ 1 \end{bmatrix}, \begin{bmatrix} -2 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix}\right\}$ is a basis (of eigenvectors) for $\mathbb{R}^3$.

(we don't need to check it, eg via a determinant computation)

Special case: If $P_A(t)$ has distinct roots & they are all real, then A is diagonalizable (over $\mathbb{R}$)

- We'll have a similar result over the complex numbers, but this requires talking about $\mathrm{Mat}_{n\times n}(\mathbb{C})$ & $\mathbb{C}^n$, which will feature in the next 2 lectures

§2. (Real) Symmetric Matrices:

THEOREM 2: Fix a symmetric matrix A $(A^T = A)$ of size $n\times n$. Then:

→ ① A has only real eigenvalues $\lambda_1, \ldots, \lambda_n$ (possibly repeated!)
HARD

② $\mathbb{R}^n$ has an orthonormal basis B of eigenvectors for A, that is, $B = \{\vec{v}_1, \ldots, \vec{v}_n\}$ with $v_i \perp v_j$ if $i\neq j$ & $\|v_1\| = \cdots = \|v_n\| = 1$

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

(note: $Q^{-1} = Q^T$ because B is orthonormal!)
(Q is called an orthogonal matrix)

Example: $A = \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} = A^T$ $\quad P_A(t) = \det\left(\begin{bmatrix} 1-t & -1 \\ -1 & 1-t \end{bmatrix}\right) = (1-t)^2 - 1 = t(t-2)$

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

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

$B = \left\{\begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}, \begin{bmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \end{bmatrix}\right\}$ is an orthonormal basis. $Q = \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & -1/\sqrt{2} \end{bmatrix}$, $D = \begin{bmatrix} 0 & 0 \\ 0 & 2 \end{bmatrix}$

Q: Why is the Theorem True?

- To discuss ① we need complex numbers & complex vector spaces $\mathbb{C}^n$

This will be the topic of the next 2 lectures.

- Next: discuss 2 & 3

Idea: Write the non-repeated eigenvalues : $\lambda_1, \ldots, \lambda_k$

- Claim: $E_{\lambda_i}(A) \perp E_{\lambda_j}(A)$ for $i \neq j$ (This is the main thing to show)
- Then, picking orthonormal basis $B_1, \ldots, B_k$ for each $E_{\lambda_1}(A), \ldots, E_{\lambda_k}(A)$ will produce the basis B for 2

Proof of claim:

Note: $(A\vec{x})\cdot\vec{y} = \vec{x}\cdot A\vec{y}$ if A is symmetric

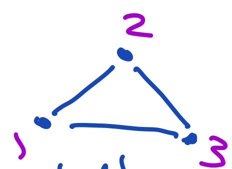
$$\left[\ (A\vec{x})\cdot\vec{y} = (A\vec{x})^T\vec{y} = (\vec{x}^TA^T)\vec{y} = \vec{x}\cdot(A^T\vec{y}) = \vec{x}\cdot A\vec{y}\ \right]$$

Now pick $\vec{v}$ in $E_{\lambda_1}(A)$ & $\vec{u}$ in $E_{\lambda_2}(A)$. for $\lambda_1 \neq \lambda_2$ eigenvalues

Want to show $\vec{v}\cdot\vec{u} = 0$

$$\lambda_1(\vec{v}\cdot\vec{u}) = \lambda_1\vec{v}\cdot\vec{u} = A\vec{v}\cdot\vec{u} \underset{\text{Note}}{=} \vec{v}\cdot A\vec{u} = \vec{v}\cdot\lambda_2\vec{u} = \lambda_2\vec{v}\cdot\vec{u}$$

so $\underbrace{(\lambda_1-\lambda_2)}_{\neq 0}(\vec{v}\cdot\vec{u}) = 0$ forces $\vec{v}\cdot\vec{u} = 0$.

Example: $A = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}$ adjacency matrix of G 
(symmetric if G is not oriented)

$$P_A(t) = \det\left(\begin{bmatrix} -t & 1 & 1 \\ 1 & -t & 1 \\ 1 & 1 & -t \end{bmatrix}\right) = -t(t^2-1) - 1(-t-1) + 1(1+t)$$

$$= -t^3 + t + 2t + 2 = -(t^3 - 3t - 2) = -(t-2)(t^2+2t+1) = -(t-2)(t+1)^2$$

Eigenvalues of A: 2 with mult 1 are real
-1 ——— 2

$$E_2 = \mathcal{N}(A - 2I_3) = \mathcal{N}\left(\begin{bmatrix} -2 & 1 & 1 \\ 1 & -2 & 1 \\ 1 & 1 & -2 \end{bmatrix}\right) = \mathrm{Sp}\left(\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}\right) \quad \dim = 1 \checkmark$$

$$E_{-1} = \mathcal{N}(A + I_3) = \mathcal{N}\left(\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}\right) = \mathrm{Sp}\left(\begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}\right) \quad \dim = 2 \checkmark$$

A is non-defective & all eigenvalues are real, so it's diagonalizable

- In particular $B = \left\{ \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} \right\}$ is a basis of eigenvectors

eigenvalues $\quad 2 \quad -1 \quad -1$

So $$S^{-1} A S = \begin{bmatrix} 2 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix} \quad \text{with} \quad S = \begin{bmatrix} 1 & -1 & -1 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} \qquad (*)$$

- Note: $E_2 \perp E_{-1}$ but B is not orthogonal because $\begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} \not\perp \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$. To get an orthonormal basis we can use Gram-Schmidt on our basis $\left\{ \underset{\vec{u}_1}{\begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}}, \underset{\vec{u}_2}{\begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}} \right\}$ of E_{-1}.

Gram Schmidt: $\vec{w}_1 = \vec{u}_1$, $\quad \vec{w}_2 = \vec{u}_2 - \dfrac{\vec{u}_2 \cdot \vec{w}_1}{\vec{w}_1 \cdot \vec{w}_1} \vec{w}_1 = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} - \dfrac{(-1)^2}{2} \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}$

$\rightsquigarrow$ Orthogonal basis for $E_{-1} = \{\vec{w}_1, \vec{w}_2\}$ $\qquad = \begin{bmatrix} -1/2 \\ -1/2 \\ 1 \end{bmatrix} = 2 \begin{bmatrix} -1 \\ -1 \\ 2 \end{bmatrix}$

$\rightsquigarrow$ Our orthonormal basis for $\mathbb{R}^3$ of eigenvectors of A is:

$$B = \left\{ \begin{bmatrix} 1/\sqrt{3} \\ 1/\sqrt{3} \\ 1/\sqrt{3} \end{bmatrix}, \begin{bmatrix} -1/\sqrt{2} \\ 1/\sqrt{2} \\ 0 \end{bmatrix}, \begin{bmatrix} -1/\sqrt{6} \\ -1/\sqrt{6} \\ 2/\sqrt{6} \end{bmatrix} \right\}$$

eigenvalues $\quad 2 \quad -1 \quad -1$

Now: $$\begin{bmatrix} 2 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix} = Q^T A Q \quad \text{with} \quad Q = \begin{bmatrix} 1/\sqrt{3} & -1/\sqrt{2} & -1/\sqrt{6} \\ 1/\sqrt{3} & 1/\sqrt{2} & -1/\sqrt{6} \\ 1/\sqrt{3} & 0 & 2/\sqrt{6} \end{bmatrix}$$

Compare with $(*)$: $Q^{-1} = Q^T$ comes for free, but S^{-1} needs to be computed!

Observations:
- For each $l > 0$, (i,j)-entry of A^l is the number of paths of length l between node i & node j.
- A^l is easy to compute because A is diagonalizable ($A^l = Q \begin{bmatrix} 2^l & & 0 \\ & (-1)^l & \\ 0 & & (-1)^l \end{bmatrix} Q^T$)