Lecture 13: Diagonalization, Part II

Non-distinct eigenvalues

As usual let A be an $n \times n$ matrix with real entries.

Recall that if you have a basis $\{\vec{v}_i\}_{i=1}^n$ of $\mathbb{R}^n$ consisting of eigenvectors (with eigenvalues $\{\lambda_i\}_{i=1}^n$), then we have

$$A = PDP^{-1} = \begin{pmatrix} \uparrow & & \uparrow \\ \vec{v}_1 & \cdots & \vec{v}_n \\ \downarrow & & \downarrow \end{pmatrix} \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix} \begin{pmatrix} \uparrow & & \uparrow \\ \vec{v}_1 & \cdots & \vec{v}_n \\ \downarrow & & \downarrow \end{pmatrix}^{-1}$$

— i.e. A is diagonalizable. We also saw that this may not be the case — we may not have an A-eigenbasis of $\mathbb{R}^n$ — if the λ_i (appearing as solutions of $\det(A - \lambda I_n) = 0$) are not all different. We now examine this case more closely. Assume for this part that $\det(A - \lambda I_n) = \prod_{i=1}^n (\lambda - \lambda_i)$ with all λ_i real.

Suppose that for some $\lambda_0 \in \mathbb{R}$ we have $\dim E_{\lambda_0} = k$, i.e. A has exactly k independent eigenvectors $\vec{v}_1, \dots, \vec{v}_k$ with the same eigenvalue λ_0. Extend this to a basis $\vec{v}_1, \dots, \vec{v}_n$ of $\mathbb{R}^n$, and put $P = \begin{pmatrix} \uparrow & & \uparrow \\ \vec{v}_1 & \cdots & \vec{v}_n \\ \downarrow & & \downarrow \end{pmatrix}$. For $1 \le j \le k$,

$$P^{-1}AP\vec{e}_j = P^{-1}A\vec{v}_j = P^{-1}\lambda_0\vec{v}_j = \lambda_0 P^{-1}\vec{v}_j = \lambda_0\vec{e}_j.$$

So we know the first k columns of

$$P^{-1}AP = \underbrace{\left(\begin{array}{ccc|c} \lambda_0 & & & \\ & \ddots & & C \\ & & \lambda_0 & \\ \hline & O & & B \end{array}\right)}_{1^{\text{st}}\ k \text{ columns}} \Longrightarrow \underset{P^{-1}(A-\lambda I_n)P}{\overset{}{P^{-1}AP - \lambda I_n}} = \left(\begin{array}{ccc|c} \lambda_0-\lambda & & O & \\ & \ddots & & C \\ O & & \lambda_0-\lambda & \\ \hline & O & & B-\lambda I_{n-k} \end{array}\right)$$

$$\Longrightarrow \det(A-\lambda I_n) = \det\left(P^{-1}(A-\lambda I_n)P\right) = (\lambda_0-\lambda)^k \cdot \det(B-\lambda I_{n-k})$$

$\Longrightarrow$ the multiplicity of λ_0 (as an eigenvalue of A) is <u>at least</u> k, proving

<u>Theorem 1</u>: For any eigenvalue λ_0 of A, $1 \le \dim E_{\lambda_0} \le \text{mult}(\lambda_0)$.

Since our only hope for an A-eigenbasis of $\mathbb{R}^n$ is to have the eigenspace dimensions sum to n, and these are bounded by the eigenvalue multiplicities (which do sum to n), we really need all the eigenspaces to be as large as possible:

<u>Theorem 2</u>: The following are equivalent:

(i) An A-eigenspace of $\mathbb{R}^n$ exists.

(ii) A is diagonalizable (over $\mathbb{R}$).

(iii) $\underbrace{\boxed{\dim E_{\lambda_0} = \text{mult}(\lambda_0)}}_{(*)}$ for each eigenvalue λ_0 of A.

<u>Remark</u>: $\dim E_{\lambda_0} = \text{mult}(\lambda_0)$ is automatic when $\text{mult}(\lambda_0) = 1$, by Theorem 1.

Ex 1/ Is $A = \begin{bmatrix} \boxed{5} & -3 & 0 & 9 \\ 0 & \boxed{3} & 1 & -2 \\ 0 & 0 & \boxed{2} & 0 \\ 0 & 0 & 0 & \boxed{2} \end{bmatrix}$ diagonalizable?

Eigenvalues are 5, 3, and 2 (with mult. 2). We have to check that $\dim E_2 = 2$, i.e. that $\text{nullity}(A - 2I_4) = 2$.

$$A - 2I_4 = \begin{array}{c} \begin{bmatrix} 3 & -3 & 0 & 9 \\ 0 & 1 & 1 & -2 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \\ \uparrow \ \ \uparrow \\ \text{free} \end{array} \Longrightarrow \dim E_2 = 2. \text{ So } (*) \text{ holds.}$$

$\Downarrow$

Yes.

Ex 2/ Is $A = \begin{bmatrix} 5 & -3 & 0 & 9 \\ 0 & 3 & 1 & 2 \\ 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 2 \end{bmatrix}$ diagonalizable?

Same eigenvalues/mults., but $\dim E_2 =$

$$\text{nullity}(A - 2I_4) = \text{nullity}\begin{array}{c} \begin{bmatrix} 3 & -3 & 0 & 9 \\ 0 & 1 & 1 & -2 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} \\ \uparrow \\ \text{free} \end{array} = 1. \text{ So } (*) \text{ fails.}$$

$\Downarrow$

No.

$\boxed{\text{Complex Eigenvalues}}$

Ex 3/ Let $A = \begin{pmatrix} a & -b \\ b & a \end{pmatrix}$, $a, b \in \mathbb{R}$. Can you diagonalize this?

$$A - \lambda I_2 = \begin{pmatrix} a-\lambda & -b \\ b & a-\lambda \end{pmatrix} \Rightarrow \det(A - \lambda I_2) = (a-\lambda)^2 + b^2$$
$$= \lambda^2 - 2a\lambda + (a^2+b^2)$$

$$\Rightarrow \det(A - \lambda I_2) = 0 \text{ iff } \lambda = \frac{2a \pm \sqrt{4a^2 - 4(a^2+b^2)}}{2} = a \pm ib,$$

where $i = \sqrt{-1}$. Needless to say, we are not in the situation examined in the last part! We have distinct λ_1 & λ_2, but they are <u>not real numbers</u>. However, <u>Theorems 1 & 2 remain true provided you replace $\mathbb{R}$ and $\mathbb{R}^n$ by $\mathbb{C}$ and $\mathbb{C}^n$</u>. That is, while you <u>cannot</u> diagonalize A "over $\mathbb{R}$", you <u>can</u> diagonalize it "over $\mathbb{C}$", like this:

$$E_{a+ib} = \text{Nul}(A - (a+ib)I_2) = \text{Nul}\begin{pmatrix} -bi & -b \\ b & -bi \end{pmatrix} = \text{Nul}\begin{pmatrix} 1 & -i \\ 0 & 0 \end{pmatrix} = \text{span}\left\{\begin{pmatrix} i \\ 1 \end{pmatrix}\right\}$$

$$E_{a-ib} = \text{Nul}(A - (a-ib)I_2) = \text{Nul}\begin{pmatrix} bi & -b \\ b & bi \end{pmatrix} = \text{Nul}\begin{pmatrix} 1 & i \\ 0 & 0 \end{pmatrix} = \text{span}\left\{\begin{pmatrix} -i \\ 1 \end{pmatrix}\right\}.$$

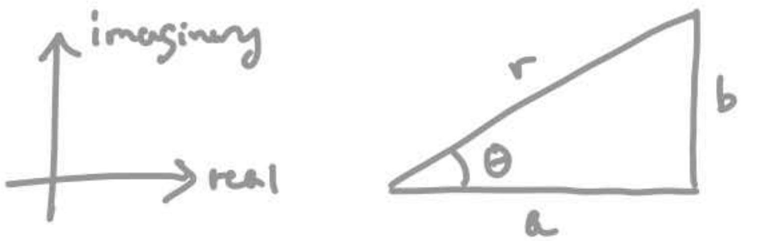
Writing $D = \begin{pmatrix} a+ib & 0 \\ 0 & a-ib \end{pmatrix}$ and $P = \begin{pmatrix} i & -i \\ 1 & 1 \end{pmatrix}$, we get

$$A = PDP^{-1}.$$

//

This suggests that <u>complex eigenvalues and eigenvectors of real matrices</u> occur in complex-conjugate <u>pairs</u>, which is in fact true in general.

In the above Example, A is a rotation-dilation matrix: taking $r = \sqrt{a^2+b^2}$, we have the right triangle



$$A = r \cdot \begin{pmatrix} a/r & -b/r \\ b/r & a/r \end{pmatrix} = \underbrace{\begin{pmatrix} r & 0 \\ 0 & r \end{pmatrix}}_{\text{dilation}} \underbrace{\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}}_{\substack{\text{rotation} \\ =: R_\theta}}$$

(Note here that $a \pm ib = re^{\pm i\theta}$.)

More generally, if $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, then the characteristic polynomial is $\det(A - \lambda I_2) = (a-\lambda)(d-\lambda) - bc$

$$= \lambda^2 - \underbrace{(a+d)}_{\substack{\mathrm{tr}(A) = \\ \text{trace of } A}}\lambda + \underbrace{(ad-bc)}_{\det(A)}$$

There are 3 cases, according to the quadratic formula:

(i) $(\mathrm{tr}\, A)^2 > 4 \det A$: 2 distinct real roots $\Rightarrow$ A diagonalizable

(ii) $(\mathrm{tr}\, A)^2 = 4 \det A$: real root with multiplicity 2 $\Rightarrow$ may or may not diagonalize

(iii) $(\mathrm{tr}\, A)^2 < 4 \det A$: 2 distinct complex roots $\Rightarrow$ A diagonalizable (but with P & D complex matrices).

While case (i) was what played a rôle in our wolf-sheep "dynamical system", in case (iii) repeatedly applying A to a vector turns out to "rotate" it elliptically about the origin.

<u>Remark</u>: Note that since $a+bi$ was an eigenvalue above, we know that the 2 rows of $(A-(a+bi)\mathbb{I}_2)$ are multiples of one another. So you can just cross out the second w/o thinking.

Ex 4/ Find the eigenvalues & eigenvectors of

$$A = \begin{pmatrix} 5 & 1 \\ -8 & 1 \end{pmatrix}.$$

$$\det(A-\lambda\mathbb{I}_2) = \det\begin{pmatrix} 5-\lambda & 1 \\ -8 & 1-\lambda \end{pmatrix} = \lambda^2 - 6\lambda + 13$$
$$= (\lambda-(3+2i))(\lambda-(3-2i))$$

$$\Rightarrow A-(3+2i)\mathbb{I}_2 = \begin{pmatrix} 2-2i & 1 \\ -8 & -2-2i \end{pmatrix} \xrightarrow[\text{Remark}]{\text{use}} \begin{pmatrix} 2-2i & 1 \\ 0 & 0 \end{pmatrix}$$

gives $E_{3+2i} = \operatorname{span}\left\{\begin{pmatrix} 1 \\ -2+2i \end{pmatrix}\right\}$, and (taking complex conjugates)

$$E_{3-2i} = \operatorname{span}\left\{\begin{pmatrix} 1 \\ -2-2i \end{pmatrix}\right\}.$$

//

$\boxed{\text{Trace}}$ The *trace* of an $n \times n$ matrix A is the sum of its diagonal entries,

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

Write $p_A(\lambda) := \det(\lambda I - A) = \prod_{i=1}^{n} (\lambda - \lambda_i)$ for the characteristic polynomial. (This does not change when you conjugate A by an invertible matrix P, because $\det(\lambda I - PAP^{-1}) = \det(P(\lambda I - A)P^{-1}) = \det(\lambda I - A)$.)

Recall that we showed in Lect. 12 that $\prod_{i=1}^{n} \lambda_i = \det(A)$.

__Claim:__ $\sum_{i=1}^{n} \lambda_i = \mathrm{tr}(A)$.

__Proof:__ Since $p_A(\lambda) = \prod_i (\lambda - \lambda_i)$, the coefficient of λ^{n-1} is $-\sum_{i=1}^{n} \lambda_i$. But also

$$p_A(\lambda) = \det(\lambda I - A) = \sum_{\sigma \text{ permutation}} \mathrm{sgn}(\sigma) \prod_{i=1}^{n} (\lambda \delta_{i\sigma_i} - a_{i\sigma_i})$$

$$= \underbrace{\prod_{i=1}^{n} (\lambda - a_{ii})}_{\substack{\sigma = \text{identity} \\ (\text{trivial permutation} \\ \sigma_i = i) \text{ term}}} + \sum_{\substack{\sigma \text{ nontrivial} \\ \text{permutation}}} \mathrm{sgn}(\sigma) \prod_{i=1}^{n} (\lambda \delta_{i\sigma_i} - a_{i\sigma_i})$$

$\Rightarrow$ at least two σ_i's are different from i

$\Rightarrow$ product has degree $\leq n-2$

$$= \lambda^n - \left(\sum_{i=1}^{n} a_{ii}\right) \lambda^{n-1} + \text{lower-degree terms.} \qquad \square$$

__Consequence:__ $\mathrm{tr}(A)$ doesn't change when you conjugate A by an invertible matrix!