Lecture 32: §4.4 Eigenvalues & the characteristic polynomial

§1 **Recall**: The Eigenvalue problem

Input: A of size $n\times n$

Output All values of λ (eigenvalues) for which the system

$$A\begin{bmatrix} x_1 \\ \vdots \\ x_n\end{bmatrix} = \lambda \begin{bmatrix} x_1 \\ \vdots \\ x_n\end{bmatrix} \quad \text{admits a solution } \underbrace{\begin{bmatrix} x_1 \\ \vdots \\ x_n\end{bmatrix}}_{\hookrightarrow \text{ eigenvectors}} \neq \vec{0}$$

Key Fact (last time): All valid λ's can be obtain from $\det(A-\lambda I_n)=0$

Example: $A = \begin{bmatrix} 1 & 0 & -2 \\ 1 & 3 & 1 \\ 1 & 3 & 1\end{bmatrix}$ $\qquad \det(A) = 0$

$$A-\lambda I_3 = \begin{bmatrix} 1 & 0 & -2 \\ 1 & 3 & 1 \\ 1 & 3 & 1\end{bmatrix} - \begin{bmatrix} \lambda & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \lambda\end{bmatrix} = \begin{bmatrix} 1-\lambda & 0 & -2 \\ 1 & 3-\lambda & 1 \\ 1 & 3 & 1-\lambda\end{bmatrix} \leftarrow \text{expand along this row}$$

$$\det(A-\lambda I_3) = (-1)^{1+1}(1-\lambda)\begin{vmatrix} 3-\lambda & 1 \\ 3 & 1-\lambda\end{vmatrix} + 0 + (-1)^{1+3}(-2)\begin{vmatrix} 1 & 3-\lambda \\ 1 & 3\end{vmatrix}$$

$$= (1-\lambda)\left((3-\lambda)(1-\lambda)-3\right) - 2\left(3-(3-\lambda)\right)$$

$$= (1-\lambda)(\lambda^2-4\lambda) - 2\lambda$$

$$= (1-\lambda)\lambda(\lambda-4) - 2\lambda \qquad \text{(common factor)}$$

$$= \lambda\left((1-\lambda)(\lambda-4)-2\right)$$

$$= \lambda(-\lambda^2+5\lambda-6) = -\lambda(\lambda^2-5\lambda+6) = \boxed{-\lambda(\lambda-2)(\lambda-3)}$$

Roots of $(\lambda^2-5\lambda+6)$ are 2 & 3 by quadratic formula

Conclusion: Eigenvalues of A are $\lambda=0$, $\lambda=2$ & $\lambda=3$.

Name $\boxed{\det(A-\lambda I_n) = \text{Characteristic polynomial of } A}$

Why? It's a polynomial in $\mathbb{R}[\lambda]$ (think of λ as a variable!)

More formally: $P_A(t) = \det(A-tI_n)$ is the characteristic polynomial of A in the variable t.

§2. The Characteristic Polynomial:

Q: What can we say about $P_A(t)$?

Theorem 1. The eigenvalues of A are the roots of the polynomial $P_A(t)$.

Q What else? Let's look at some examples

Example (page 1) A has size 3×3, $P_A(t)$ has degree 3 & $P_A(0)=0=\det(A)$

Another example $A=\begin{bmatrix}1&1\\0&2\end{bmatrix}$ has size 2×2 & $P_A(t)=(t-1)(t-2)=t^2-3t+2$ has degree 2

$\det A = 2$ $\qquad P_A(0)=2=\det(A)$

This is a general statement!

Theorem 2: $P_A(t)$ is a polynomial of degree n in t & $P_A(0)=\det(A)$

Why? • $P_A(t)=\det(A-tI_n)$ so specializing at $t=0$ clearly gives $\det(A)$

To check the degree claim, think about the process of computing determinants

• t appears only along the diagonal of $A-tI_n = \begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1n}\\ \vdots & & & \\ b_{n1} & b_{n2} & & b_{nn}\end{pmatrix}$

So $\det(A-tI_n) = (-1)^{1+1}b_{11}\begin{vmatrix} b_{22} & \cdots & b_{2n}\\ b_{n2} & \cdots & b_{nn}\end{vmatrix} + (-1)^{1+2}b_{12}\begin{vmatrix} b_{21} & b_{23} & \cdots & b_{2n}\\ \vdots & & & \\ b_{n1} & b_{n3} & \cdots & b_{nn}\end{vmatrix} +$

$\cdots\cdots + (-1)^{1+n}b_{1n}\begin{vmatrix} b_{21} & \cdots & b_{2(n-1)}\\ \vdots & & \\ b_{n1} & \cdots & b_{n(n-1)}\end{vmatrix}$

replace values for b_{ij}

$$\overset{\text{replace values for } b_{ij}}{=} (-1)^{1+1}\underbrace{(a_{11}-t)}_{\deg 1}\underbrace{\begin{vmatrix}(a_{22}-t) & \cdots & a_{2n}\\ a_{32} & & a_{3n}\\ & & \vdots\\ a_{n2} & \cdots & (a_{nn}-t)\end{vmatrix}}_{\text{degree}=n-1} + (-1)^{1+2}a_{12}\underbrace{\begin{vmatrix} a_{21} & a_{23} & \cdots & a_{2n}\\ & (a_{33}-t) & \cdots & \\ & & \ddots & \\ a_{n1} & \cdots & & (a_{nn}-t)\end{vmatrix}}_{\text{degree}_t\,\le n-2}$$

$$+\cdots + (-1)^{1+n}a_{1n}\underbrace{\begin{vmatrix} a_{21} & (a_{22}-t) & \cdot & & a_{2(n-1)}\\ a_{31} & a_{32} & (a_{33}-t) & & \\ \vdots & & & \ddots & \\ a_{n1} & a_{n2} & \cdots & \cdots & a_{n(n-1)}\end{vmatrix}}_{\deg\le n-2}$$

Now, Look at the summands & try to determine their degrees in t (or at least a bound for the degree)

1st summand: Matrix has t's along the diagonal & size $=(n-1)\times(n-1)$
It is the characteristic polynomial of $\begin{pmatrix} a_{22} & a_{23} & \cdots & a_{2n} \\ \vdots & & & \\ a_{n2} & & \cdots & a_{nn} \end{pmatrix}$ so it's a polynomial of degree $(n-1)$ in t.

The rest of the summands: Matrices have $(n-2)$ entries with a t so the degree cannot be larger than $(n-2)$.

Conclusion: The first summand has degree n in t & the others have degree $\leq n-2$. Overall, we get that $P_A(t)$ has degree n. $\square$

Natural questions:

Q1: How many real roots does $P_A(t)$ have? A: At most $n = \deg P_A(t)$

Q2: How can we find them? A: No unique nor complete list of methods, but we have some heuristic methods.

Our next examples describe some of them:

$\boxed{\text{EXAMPLE 1}}$ $A = \begin{bmatrix} -2 & -1 \\ 1 & -2 \end{bmatrix}$

$P_A(t) = \det \begin{bmatrix} -2-t & -1 \\ 1 & 2-t \end{bmatrix} = (-2-t)^2 + 1 = t^2+4t+5$ has no real roots

Why? Quadratic formula gives roots: $\dfrac{-4 \pm \sqrt{16-20}}{2} = \dfrac{-4 \pm \sqrt{-4}}{2}$ <0

- Another way to see this:

$P_A(t) = \underbrace{(-2-t)^2}_{\geq 0 \text{ for any } t \text{ in } \mathbb{R}} + 1 \geq 1 \rightsquigarrow$ can't get $=0$ using t in $\mathbb{R}$.

(Heuristic: Write $P_A(t)$ as a sum of positive terms (eg squares))
("Sum of squares" expressions)

EXAMPLE 2: $A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -2 & -1 \\ 0 & 1 & -2 \end{bmatrix}$

$$P_A(t) = \det \begin{pmatrix} \boxed{1-t} & 0 & 0 \\ 0 & -2-t & -1 \\ 0 & 1 & -2-t \end{pmatrix} = (1-t)(t^2+4t+5)$$

"block decomposition of the matrix"

no real root

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

$$P_A(t) = \det \begin{pmatrix} \boxed{1-t} & 0 & 0 \\ 0 & \boxed{2-t} & 0 \\ 0 & 0 & \boxed{-3-t} \end{pmatrix} = (1-t)(2-t)(-3-t)$$

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

3 blocks along diagonal

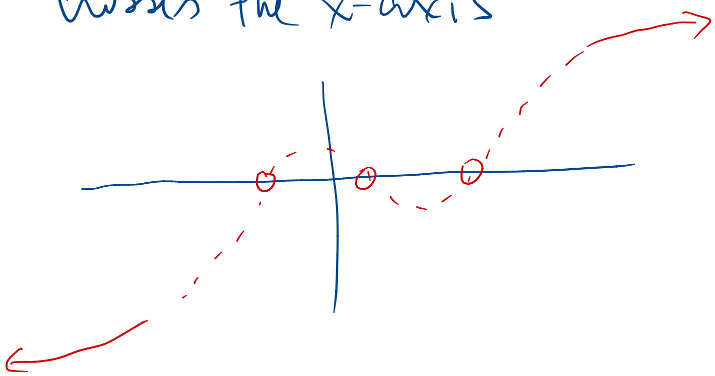
3 roots

- In general If n odd, then $P_A(t)$ has at least one real root

Why? $P_A(t) = a\,t^n + (\text{lower order terms})$

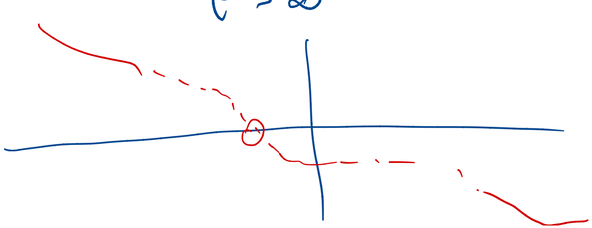
If $a > 0$: $\lim_{t\to+\infty} P_A(t) = +\infty$ & $\lim_{t\to-\infty} P_A(t) = -\infty$ (n odd)

$P_A(t)$ is continuous, so at some point, the graph of $P_A(t)$ crosses the x-axis



In this case: 3 roots but we know at least one root

If $a < 0$ $\lim_{t\to\infty} P_A(t) = -\infty$ & $\lim_{t\to-\infty} P_A(t) = +\infty$



So at some point, we must cross the x-axis

- If n even, $P_A(t)$ may have no real roots.

Note: If we allow complex numbers as roots, then $P_A(t)$ has exactly n roots over $\mathbb{C}$, if we count them with multiplicity, (meaning $(x-1)^2$ has 2 roots = 1 double root $x=1$)

$\boxed{\text{EXAMPLE 1 (revisited)}}$ $\quad P_A(t) = t^2+4t+5$

Roots: $t = \dfrac{-4 \pm \sqrt{-4}}{2}$

Complex numbers provide roots of negative numbers, by adding a root of -1, which we call i.

So $\sqrt{-4} = \sqrt{4(-1)} = \sqrt{4}\sqrt{-1} = 2i$

Then, roots of $P_A(t) = \dfrac{-4 \pm 2i}{2} = -2 \pm i$ $\begin{cases} -2+2i \\ -2-i \end{cases}$

We'll see more about this in §4.6

§2. Properties of Eigenvalues:

- Questions: - How do eigenvalues of A relate to eigenvalues of A^2, A^3, A^4, ...?
  - How do eigenvalues of A, with A invertible, relate to A^{-1}?

Theorem 3: Fix A an $n \times n$ matrix & let λ be an eigenvalue of A.

Then: (1) λ^k is an eigenvalue of A^k for $k = 2, 3, 4, \ldots$

(2) If A is nonsingular, then $\lambda \neq 0$ & $\frac{1}{\lambda}$ is an eigenvalue of A^{-1}.

Moreover, they all share the same eigenvector!

Proof: (1) Let's discuss $k=2$. The rest will follow by iterating the same argument.

By definition: $A\begin{bmatrix}x_1\\ \vdots\\ x_n\end{bmatrix} = \lambda\begin{bmatrix}x_1\\ \vdots\\ x_n\end{bmatrix}$ admits a solution $\begin{bmatrix}x_1\\ \vdots\\ x_n\end{bmatrix} \neq \vec{0}$

(eigenvector)

Multiply by A: & associate

$$A^2\begin{bmatrix}x_1\\ \vdots\\ x_n\end{bmatrix} = A\left(A\begin{bmatrix}x_1\\ \vdots\\ x_n\end{bmatrix}\right) = A\left(\lambda\begin{bmatrix}x_1\\ \vdots\\ x_n\end{bmatrix}\right) = \lambda\left(A\begin{bmatrix}x_1\\ \vdots\\ x_n\end{bmatrix}\right) = \lambda\,\lambda\begin{bmatrix}x_1\\ \vdots\\ x_n\end{bmatrix}$$
$$= \lambda^2\begin{bmatrix}x_1\\ \vdots\\ x_n\end{bmatrix}$$

So $A^2\begin{bmatrix}x_1\\ \vdots\\ x_n\end{bmatrix} = \lambda^2\begin{bmatrix}x_1\\ \vdots\\ x_n\end{bmatrix}$ & $\begin{bmatrix}x_1\\ \vdots\\ x_n\end{bmatrix} \neq \vec{0}$.

By definition, λ^2 is an eigenvalue of A^2 with eigenvector $\begin{bmatrix}x_1\\ \vdots\\ x_n\end{bmatrix}$.

(shared with A)

The proof for $k=3,4,\ldots$ follows the same reasoning.

(2) First, we show $\lambda \neq 0$.

A is invertible, so it's nonsingular. In particular $N(A) = \{\vec{0}\}$,

so $A\begin{bmatrix}x_1\\ x_n\end{bmatrix} = 0\begin{bmatrix}x_1\\ x_n\end{bmatrix}$ has no nontrivial solution. So $\lambda = 0$ is NOT an eigenvalue of A.

Another way to see this:

$P_A(t) = \det(A - tI_n)$ has $P_A(0) = \det(A) \neq 0$ so $t=0$ is not a root of $P_A(t)$. Thus, it can't be an eigenvalue.

- Next, we show λ^{-1} is an eigenvalue of A^{-1}.

Start from $A\begin{bmatrix}x_1\\ \vdots\\ x_n\end{bmatrix} = \lambda\begin{bmatrix}x_1\\ \vdots\\ x_n\end{bmatrix}$ has a solution in $\begin{bmatrix}x_1\\ \vdots\\ x_n\end{bmatrix} \neq \vec{0}$

(eigenvector)

Now, multiply by A^{-1} & associate:

$$\underbrace{(A^{-1}A)}_{I_n}\begin{bmatrix}x_1\\ \vdots\\ x_n\end{bmatrix} = A^{-1}\left(A\begin{bmatrix}x_1\\ \vdots\\ x_n\end{bmatrix}\right) = A^{-1}\left(\lambda\begin{bmatrix}x_1\\ \vdots\\ x_n\end{bmatrix}\right) = \lambda A^{-1}\begin{bmatrix}x_1\\ \vdots\\ x_n\end{bmatrix}$$

So $\begin{bmatrix}x_1\\ \vdots\\ x_n\end{bmatrix} = \lambda A^{-1}\begin{bmatrix}x_1\\ \vdots\\ x_n\end{bmatrix}$

Now, multiply by $\frac{1}{\lambda}$:

$$\frac{1}{\lambda}\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = A^{-1}\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$$ has a solution $\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \neq \vec{0}$

(same eigenvector as the one for A)

Theorem 4: A of size $n \times n$ with eigenvalue λ. Then, for any scalar μ, we have that $(\lambda+\mu)$ is an eigenvalue for $(A+\mu I_n)$. Furthermore, both matrices share the corresponding eigenvectors.

Proof: Again, we use the definitions.

Write $A\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \lambda\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$ for some eigenvector $\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \neq \vec{0}$

Add $\mu\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$ to both sides & regroup

$$A\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} + \mu\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \lambda\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} + \mu\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = (\lambda+\mu)\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$$

Use $\mu\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = (\mu I_n)\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$ to regroup the right-hand side: We get

$$(A+\mu I_n)\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = (\lambda+\mu)\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$$ & $\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \neq \vec{0}$, so it's an eigenvector & $(\lambda+\mu)$ is eigenvalue

Example (page 1, revisited)

$$A = \begin{bmatrix} 1 & 0 & -2 \\ 1 & 3 & 1 \\ 1 & 3 & 1 \end{bmatrix} \qquad P_A(t) = -t\,(t-2)(t-3)$$

- A is singular because $\lambda = 0$ is an eigenvalue
- Eigenvalues of A^2 include $0^2, 2^2, 3^2$. Since we can have at most $n=3$ then they are all the eigenvalues of A^2
- Eigenvalues of A^3 include $0^3, 2^3, 3^3$. Since we can have at most $n=3$, then they are all the eigenvalues of A^3
- Eigenvalues of $A-5I_3$ include $0-5,\ 2-5,\ 3-5 = -5, -3, -2$. By the same reasoning from above, these are <u>all</u> the eigenvalues of $A-5I_3$.