The Characteristic Polynomial

Last time:

- An eigenvector of a square matrix A is a nonzero vector v such that
$$Av = \lambda v$$
for a scalar λ.
- The associated eigenvalue is λ.

We like eigenvectors because $A^k v = \lambda^k v$ for all k.

Given an eigenvalue λ, we know how to compute all λ-eigenvectors: the λ-eigenspace is
$$\mathrm{Nul}(A - \lambda I_n)$$

How do you find the eigenvalues of A?

λ is an eigenvalue of A

$\iff A - \lambda I_n$ is not invertible

$\iff \det(A - \lambda I_n) = 0$

Def: The characteristic polynomial of an $n \times n$ matrix A is $p(\lambda) = \det(A - \lambda I_n)$

$$\boxed{\lambda \text{ is an eigenvalue of } A \iff p(\lambda)=0}$$

Eg: Find all eigenvalues of $A = \begin{pmatrix} 0 & 13 & 12 \\ 1/4 & 0 & 0 \\ 0 & 1/2 & 0 \end{pmatrix}$

$$\det(A - \lambda I_3) = \det\begin{pmatrix} -\lambda & 13 & 12 \\ 1/4 & -\lambda & 0 \\ 0 & 1/2 & -\lambda \end{pmatrix}$$

$$\xrightarrow[\text{cofactors}]{\text{expand}} \cdots = -\lambda^3 + \frac{13}{4}\lambda + \frac{3}{2} = p(\lambda)$$

How do we find the roots of a degree-n polynomial?

- In real life: ask a computer

  NB the computer will turn this back into an eigenvalue problem and will use a different (faster) eigenvalue-finding algorithm

- By hand: I'll give you one root λ_0.

  Compute the quadratic polynomial $p(\lambda)/(\lambda-\lambda_0)$ using synthetic division, then use the quadratic formula.

NB: This is not a Gaussian elimination problem!

Eg cont'd: We know 2 is an eigenvalue of the rabbit matrix. Check:

$$p(2) = -8 + \frac{13}{4}\cdot 2 + \frac{3}{2} = -\frac{16}{2} + \frac{13}{2} + \frac{3}{2} = 0 \checkmark$$

This means $(\lambda-2)$ divides $p(\lambda)$. What's $p(\lambda)/(\lambda-2)$?

Synthetic division (long division of polynomials):

$$\begin{array}{r l}
 & -\lambda^2 - 2\lambda - \frac{3}{4} \\
\lambda - 2 \, \big) & -\lambda^3 + \frac{13}{4}\lambda + \frac{3}{2} \\
 & -\left[-\lambda^2(\lambda-2)\right] \\
\hline
 & -2\lambda^2 + \frac{13}{4}\lambda + \frac{3}{2} \\
 & -\left[-2\lambda(\lambda-2)\right] \\
\hline
 & -4\lambda + \frac{13}{4}\lambda + \frac{3}{2} = -\frac{3}{4}\lambda + \frac{3}{2} \\
 & -\left[-\frac{3}{4}(\lambda-2)\right] = 0 \checkmark
\end{array}$$

So $-\lambda^3 + \frac{13}{4}\lambda + \frac{3}{2} = (\lambda-2)\left(-\lambda^2 - 2\lambda - \frac{3}{4}\right)$

Quadratic formula: roots of $-\lambda^2 - 2\lambda - \frac{3}{4}$ are

$$-\frac{1}{2}\left(2 \pm \sqrt{4-3}\right) = -\frac{1}{2}, -\frac{3}{2}$$

$$\Rightarrow -\lambda^3 + \frac{13}{4}\lambda + \frac{3}{2} = -(\lambda-2)\left(\lambda+\frac{1}{2}\right)\left(\lambda+\frac{3}{2}\right)$$

So the eigenvalues are $2, -\frac{1}{2}, -\frac{3}{2}$.

To check these are eigenvalues, let's find some eigenvectors!

$$2 \rightsquigarrow \mathrm{Nul}(A - 2I_3) = \mathrm{Nul}\begin{pmatrix} -2 & 13 & 12 \\ 1/4 & 0 & 0 \\ 0 & 1/2 & 0 \end{pmatrix} \overset{\text{PrF}}{=} \mathrm{Span}\left\{\begin{pmatrix} 32 \\ 4 \\ 1 \end{pmatrix}\right\}$$

$$-\frac{1}{2} \rightsquigarrow \mathrm{Nul}(A + \tfrac{1}{2}I_3) = \mathrm{Nul}\begin{pmatrix} 1/2 & 13 & 12 \\ 1/4 & 1/2 & 0 \\ 0 & 1/2 & 1/2 \end{pmatrix} \overset{\text{PrF}}{=} \mathrm{Span}\left\{\begin{pmatrix} 2 \\ -1 \\ 1 \end{pmatrix}\right\}$$

$$-\frac{3}{2} \rightsquigarrow \mathrm{Nul}(A + \tfrac{3}{2}I_3) = \mathrm{Nul}\begin{pmatrix} 3/2 & 13 & 12 \\ 1/4 & 3/2 & 0 \\ 0 & 1/2 & 3/2 \end{pmatrix} \overset{\text{PrF}}{=} \mathrm{Span}\left\{\begin{pmatrix} 18 \\ -3 \\ 1 \end{pmatrix}\right\}$$

In this example, $p(\lambda) = -\lambda^3 + \frac{13}{4}\lambda + \frac{3}{2}$ is a degree-3 polynomial. What does it look like in general? Let's try 2×2 matrices first.

Eg: $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$

$$\det(A - \lambda I_2) = \det\begin{pmatrix} a-\lambda & b \\ c & d-\lambda \end{pmatrix} = (a-\lambda)(d-\lambda) - bc$$

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

This is a polynomial of degree 2.

Def: The trace of a matrix A is

$\mathrm{Tr}(A) =$ the sum of the diagonal entries of A.

Eg: $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ $\quad \mathrm{Tr}(A) = a+d$

Characteristic Polynomial of a 2x2 Matrix A

$$p(\lambda) = \lambda^2 - \mathrm{Tr}(A)\lambda + \det(A)$$

NB: $p(0) = \det(A - 0I_n) = \det(A)$

so the constant term is always $\det(A)$.

General Form: If A is an $n\times n$ matrix, then

$$p(\lambda) = (-1)^n \lambda^n + (-1)^{n-1}\mathrm{Tr}(A)\lambda^{n-1} + (\text{other terms}) + \det(A)$$

→ This is a degree-n polynomial

→ You only get the λ^{n-1} and constant coeffs "for free" — the rest are more complicated.

Eg: $A = \begin{pmatrix} 0 & 13 & 12 \\ 1/4 & 0 & 0 \\ 0 & 1/2 & 0 \end{pmatrix} \rightsquigarrow p(\lambda) = -\lambda^3 + 0\lambda^2 + \frac{13}{4}\lambda + \frac{3}{2}$

$\mathrm{Tr}(A) = 0 + 0 + 0 = 0$ ✓ $\quad \det(A) = -\frac{1}{4}\cdot\left(-\frac{12}{2}\right) = \frac{3}{2}$ ✓

Fact: A polynomial of degree n has at most n roots ("roots" = "zeros")

Consequence: An $n\times n$ matrix has at most n eigenvalues.

Diagonalization

Rabbit Example Cont'd: We computed the matrix

$$A = \begin{pmatrix} 0 & 13 & 12 \\ 1/4 & 0 & 0 \\ 0 & 1/2 & 0 \end{pmatrix}$$ has eigenvalues $2, -\frac{1}{2}, -\frac{3}{2}$

& eigenspaces

$$2 : \text{Span}\left\{\begin{pmatrix} 32 \\ 4 \\ 1 \end{pmatrix}\right\} \quad -\frac{1}{2} : \text{Span}\left\{\begin{pmatrix} 2 \\ -1 \\ 1 \end{pmatrix}\right\} \quad -\frac{3}{2} : \text{Span}\left\{\begin{pmatrix} 18 \\ -3 \\ 1 \end{pmatrix}\right\}$$

Let's give names to some eigenvectors:

$$w_1 = \begin{pmatrix} 32 \\ 4 \\ 1 \end{pmatrix} \qquad w_2 = \begin{pmatrix} 2 \\ -1 \\ 1 \end{pmatrix} \qquad w_3 = \begin{pmatrix} 18 \\ -3 \\ 1 \end{pmatrix}$$

We know what happens if we start with $\begin{pmatrix} 32 \\ 4 \\ 1 \end{pmatrix}$ rabbits: $A^k\begin{pmatrix} 32 \\ 4 \\ 1 \end{pmatrix} = 2^k\begin{pmatrix} 32 \\ 4 \\ 1 \end{pmatrix} \rightsquigarrow$ doubles each time.

What if we start with $v_0 = \begin{pmatrix} 16 \\ 6 \\ 1 \end{pmatrix}$?

Fact: $v_0 = \begin{pmatrix} 16 \\ 6 \\ 1 \end{pmatrix}$ can be written as a linear combination of eigenvectors:

$$\begin{pmatrix} 16 \\ 6 \\ 1 \end{pmatrix} = w_1 + w_2 - w_3$$

Now it's easy to compute $A^k v_0$!

$$A^k v_0 = A^k(w_1 + w_2 - w_3) = A^k w_1 + A^k w_2 - A^k w_3$$

$$= 2^k w_1 + \left(-\tfrac{1}{2}\right)^k w_2 + \left(-\tfrac{3}{2}\right)^k w_3$$

Observation 1: $2^k \gg |(-\frac{1}{2})^k|$ and $|(-\frac{3}{2})^k|$ for large k

so $A^k v_0 \sim 2^k w_1$

This explains why eventually,

- ratios converge to $(32:4:1)$
- population roughly doubles each year

Observation 2: $\{w_1, w_2, w_3\}$ is linearly independent

(this is automatic — more later)

$\overset{\text{basis}}{\underset{\text{thm}}{\Longrightarrow}}$ $\{w_1, w_2, w_3\}$ is a basis for $\mathbb{R}^3$

$\Longrightarrow$ any vector in $\mathbb{R}^3$ is a linear combination of w_1, w_2, w_3

So if $v_0 = x_1 w_1 + x_2 w_2 + x_3 w_3$ then

$$A^k v_0 = x_1 A^k w_1 + x_2 A^k w_2 + x_3 A^k w_3$$
$$= 2^k x_1 w_1 + (-\tfrac{1}{2})^k x_2 w_2 + (-\tfrac{3}{2})^k x_3 w_3$$

So observation 1 holds for any starting vector $v_0 \in \mathbb{R}^3$.

Q: What if $x_1 = 0$?

The fact that A has 3 LI eigenvectors means we can understand how A acts on $\mathbb{R}^3$ entirely in terms of its eigenvectors & eigenvalues.

Def: Let A be an $n \times n$ matrix. A is **diagonalizable** if it has n linearly independent eigenvectors $w_1, \ldots, w_n$. In this case, $\{w_1, \ldots, w_n\}$ is called an **eigenbasis**.

In this case, to compute $A^k v$ for $v \in \mathbb{R}^n$:

(1) solve $v = x_1 w_1 + \cdots + x_n w_n$

(2) $\boxed{A^k v = \lambda_1^k x_1 w_1 + \cdots + \lambda^k x_n w_n}$ ← vector form

$(\lambda_i = \text{eigenvalue for } w_i : A w_i = \lambda_i w_i)$

So if A is diagonalizable, then we can understand how A acts on $\mathbb{R}^n$ entirely in terms of its eigenvectors & eigenvalues. Work in an eigenbasis!

Matrix Form of Diagonalization

A is diagonalizable $\iff$ there exists an invertible matrix C and a diagonal matrix D such that

$$A = CDC^{-1}$$

In this case the columns of C form an eigenbasis & the diagonal entries of D are the corresponding eigenvalues.

$$C = \begin{pmatrix} | & & | \\ w_1 & \cdots & w_n \\ | & & | \end{pmatrix} \quad D = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix} \quad Aw_i = \lambda_i w_i$$

same order: $w_i \longleftrightarrow \lambda_i$

Eg: $A = \begin{pmatrix} 0 & 13 & 12 \\ 1/4 & 0 & 0 \\ 0 & 1/2 & 0 \end{pmatrix} \Rightarrow A = CDC^{-1}$ for

$$C = \begin{array}{cccc} & w_1 & w_2 & w_3 \\ & \left(\begin{matrix} 32 \\ 4 \\ 1 \end{matrix}\right. & \begin{matrix} 2 \\ -1 \\ 1 \end{matrix} & \left.\begin{matrix} 18 \\ -3 \\ 1 \end{matrix}\right) \end{array} \quad D = \begin{pmatrix} 2 & 0 & 0 \\ 0 & -1/2 & 0 \\ 0 & 0 & -3/2 \end{pmatrix}$$

Proof: $C\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = x_1 w_1 + \cdots + x_n w_n$

$$\Rightarrow C^{-1}(x_1 w_1 + \cdots + x_n w_n) = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

Any vector has the form $v = x_1 w_1 + \cdots + x_n w_n$, and

$$CDC^{-1}v = CDC^{-1}(x_1 w_1 + \cdots + x_n w_n)$$

$$= C\begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

$$= \begin{pmatrix} | & & | \\ w_1 & \cdots & w_n \\ | & & | \end{pmatrix}\begin{pmatrix} \lambda_1 x_1 \\ \vdots \\ \lambda_n x_n \end{pmatrix} = x_1 \lambda_1 w_1 + \cdots + x_n \lambda_n w_n$$

$$= A(x_1 w_1 + \cdots + x_n w_n) = Av$$

✓

NB: If $A = CDC^{-1}$ then

$$A^k = (CDC^{-1})^k = (CD\cancel{C^{-1}})(\cancel{C}D\cancel{C^{-1}})\cdots(\cancel{C}DC^{-1})$$
$$= CD^kC^{-1} = C\begin{pmatrix} \lambda_1^k & & 0 \\ & \ddots & \\ 0 & & \lambda_n^k \end{pmatrix}C^{-1}$$

This is a closed form expression for A^k in terms of k: much easier to compute!

$$\boxed{A^k = CD^kC^{-1}}$$

← matrix form: compare p.8

↑ this matrix has n^2 entries that are functions of k

Eg: $A = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 4 \end{pmatrix}$ is diagonal:

$$Ae_1 = 2e_1 \quad Ae_2 = 3e_2 \quad Ae_3 = 3e_3$$

So $\{e_1, e_2, e_3\}$ is an eigenbasis → can take $C = I_3$, so the diagonalization is

$$A = I_3 A I_3$$

Q: What if we take e_2 to be our first eigenvector?

Eg: $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ $\quad p(\lambda) = \lambda^2 - Tr(A)\lambda + \det(A)$
$$= \lambda^2 - 2\lambda + 1 = (\lambda-1)^2$$

The only eigenvalue is 1, and the 1-eigenspc is

$$\text{Nul}(A - I_2) = \text{Nul}\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = \text{Span}\left\{\begin{pmatrix} 1 \\ 0 \end{pmatrix}\right\}$$

So all eigenvectors lie on the x-axis

$\Longrightarrow$ not diagonalizable!

(cf. p.5 of W10L1 notes)

Procedure to Diagonalize a Matrix:

(1) Compute the characteristic polynomial $p(\lambda)$

(2) Find the roots of $p(\lambda)$ = eigenvalues of A

(3) Find a basis for each eigenspace $= \text{Nul}(A - \lambda I_n)$ (using PVF)

Combine your bases from (3). If you end up with n vectors, they form an eigenbasis. Otherwise, A is not diagonalizable.

Fact: If $w_1, \dots, w_p$ are eigenvectors with different eigenvalues then $\{w_1, \dots, w_p\}$ is linearly independent.

So in the procedure, you never have to check if bases of different eigenspaces are LI when you combine them.

Proof of the Fact: Say $Aw_i = \lambda_i w_i$ and all of the $\lambda_1, \dots, \lambda_p$ are distinct. Suppose $\{w_1, \dots, w_p\}$ is LD. Then for some i, $\{w_1, \dots, w_i\}$ is LI but $w_{i+1} \in \mathrm{Span}\{w_1, \dots, w_i\}$, so

$$w_{i+1} = x_1 w_1 + \dots + x_i w_i$$

$$\Rightarrow Aw_{i+1} = A(x_1 w_1 + \dots + x_i w_i)$$

$$\Rightarrow \lambda_{i+1} w_{i+1} = \lambda_1 x_1 w_1 + \dots + \lambda_i x_i w_i$$

If $\lambda_{i+1} = 0$ then $\lambda_1 x_1 w_1 + \dots + \lambda_i x_i w_i = 0 \xRightarrow[\text{LI}]{\{w_1, \dots, w_i\}}$ $x_1 = \dots = x_i = 0$ (because $\lambda_1, \dots, \lambda_i \neq 0$), so $w_{i+1} = 0$, which can't happen because w_{i+1} is an eigenvector.

If $\lambda_{i+1} \neq 0$ then

$$w_{i+1} = \frac{\lambda_1}{\lambda_{i+1}} x_1 w_1 + \dots + \frac{\lambda_i}{\lambda_{i+1}} x_i w_i$$

Subtract $\quad w_{i+1} = x_1 w_1 + \dots + x_i w_i$

$$\rightsquigarrow 0 = \left(\frac{\lambda_1}{\lambda_{i+1}} - 1\right) x_1 w_1 + \dots + \left(\frac{\lambda_i}{\lambda_{i+1}} - 1\right) x_i w_i$$

But $\lambda_j \neq \lambda_{i+1}$ for $j \leq i$, so $\frac{\lambda_j}{\lambda_{i+1}} - 1 \neq 0$

$$\Rightarrow x_1 = \dots = x_i = 0$$

which is impossible, as before. ✓

Consequence: If A has n (different) eigenvalues then A is diagonalizable.

Indeed, if $\lambda_1, \dots, \lambda_n$ are eigenvalues and

$$A w_1 = \lambda_1 w_1, \dots, A w_n = \lambda_n w_n$$

then $\{w_1, \dots, w_n\}$ is an eigenbasis by the Fact.