The Characteristic Polynomial, cont'd

Recall from last time:

- An eigenvector of A is a vector v such that
  $$Av = \lambda v \qquad \lambda = \text{eigenvalue}$$
- The λ-eigenspace is
  $$\mathrm{Nul}(A - \lambda I_n) = \{\text{all } \lambda\text{-eigenvectors and } 0\}$$
- The characteristic polynomial of A is
  $$p(\lambda) = \det(A - \lambda I_n)$$
  The eigenvalues are the solns of $p(\lambda) = 0$.
- We like eigenvectors because
  $$Av = \lambda v \implies A^k v = \lambda^k v$$
  So we can use these to solve the difference eq^n
  $$v_{k+1} = A v_k \rightsquigarrow v_k = A^k v_0$$

What kind of function is $p(\lambda)$? What does it look like?

2×2 case: $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$

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

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

This is a polynomial of degree 2 (quadratic).

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 2×2 Matrix A

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

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

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

We know how to factor quadratic polynomials: the quadratic formula!

Eg: Find all eigenvalues of $A = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$

$Tr(A) = 4 \quad det(A) = 3$

$$p(\lambda) = \lambda^2 - 4\lambda + 3 = 0$$

$$\Rightarrow \lambda = \frac{1}{2}\left(4 \pm \sqrt{16-12}\right) = \frac{1}{2}(4 \pm 2) = 2 \pm 1$$

so the eigenvalues are 1 and 3.

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

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

→ This is a degree-n polynomial

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

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

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

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

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

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 won't ask you to factor any polynomials of degree ≥ 3 by hand.

NB: This is not a Gaussian elimination problem!

Diagonalization

Solving a difference equation $v_{k+1} = Av_k$ is easy when v_0 is an eigenvector:

$$Av_0 = \lambda v_0 \implies v_k = A^k v_0 = \lambda^k v_0.$$

It is also easy if v_0 is a linear combination of eigenvectors: suppose

$$v_0 = x_1 w_1 + \cdots + x_n w_n \quad \text{where } Aw_i = \lambda_i w_i.$$

Then

$$v_k = A^k v_0 = A^k(x_1 w_1 + \cdots + x_n w_n)$$

no matrix multiplication!

$$= x_1 A^k w_1 + \cdots + x_n A^k w_n = x_1 \lambda_1^k w_1 + \cdots + x_n \lambda_n^k w_n.$$

$$\boxed{\begin{array}{l}\text{If } Aw_1 = \lambda_1 w_1,\ Aw_2 = \lambda_2 w_2,\ \ldots,\ Aw_n = \lambda_n w_n, \text{ then} \\ A^k(x_1 w_1 + \cdots + x_n w_n) = \lambda_1^k x_1 w_1 + \cdots + \lambda_n^k x_n w_n.\end{array}}$$

Rabbit Example Cont'd: We computed the matrix

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

Compute eigenspaces (bases for $\mathrm{Nul}(A - \lambda I_3)$):

$$2: \mathrm{Span}\left\{\begin{pmatrix} 32 \\ 4 \\ 1 \end{pmatrix}\right\} \qquad -\tfrac{1}{2}: \mathrm{Span}\left\{\begin{pmatrix} 2 \\ -1 \\ 1 \end{pmatrix}\right\} \qquad -\tfrac{3}{2}: \mathrm{Span}\left\{\begin{pmatrix} 18 \\ -3 \\ 1 \end{pmatrix}\right\}$$

Let's give names to some eigenvectors:

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

Can we write our initial state $v_0 = (16, 6, 1)$ as a LC of w_1, w_2, w_3? Need to solve

$$\begin{pmatrix} 16 \\ 6 \\ 1 \end{pmatrix} = x_1 \begin{pmatrix} 32 \\ 4 \\ 1 \end{pmatrix} + x_2 \begin{pmatrix} 2 \\ -1 \\ 1 \end{pmatrix} + x_3 \begin{pmatrix} 18 \\ -3 \\ 1 \end{pmatrix}$$

Augmented matrix:
$$\left(\begin{array}{ccc|c} 32 & 2 & 18 & 16 \\ 4 & -1 & -3 & 6 \\ 1 & 1 & 1 & 1 \end{array}\right) \xrightarrow{\text{solve}} \begin{array}{l} x_1 = 1 \\ x_2 = 1 \\ x_3 = -1 \end{array}$$

So $v_0 = w_1 + w_2 - w_3$

$$\Longrightarrow v_k = A^k v_0 = 2^k w_1 + \left(-\tfrac{1}{2}\right)^k w_2 - \left(-\tfrac{3}{2}\right)^k w_3$$

$$= \begin{pmatrix} 32 \cdot 2^k + 2 \cdot (-1/2)^k - 18 \cdot (-3/2)^k \\ 4 \cdot 2^k - (-1/2)^k + 3 \cdot (-3/2)^k \\ 2^k + (-1/2)^k - (-3/2)^k \end{pmatrix}$$

closed form: no matrix multiplication

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

So $A^k v_0 \sim 2^k w_1$ (most significant digits)

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)

$\xrightarrow[\text{thm}]{\text{basis}}$ $\{w_1, w_2, w_3\}$ is a basis for $\mathbb{R}^3$

$\Rightarrow$ 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 initial state $v_0 \in \mathbb{R}^3$.

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 $w_1,\dots,w_n$. In this case, $\{w_1,\dots,w_n\}$ is called an **eigenbasis**.

In this case, any vector in $\mathbb{R}^n$ is a linear combination of eigenvectors. Writing a vector as a LC of eigenvectors is called expanding in an eigenbasis.

→ This means solving the vector equation

$$v = x_1 w_1 + x_2 w_2 + \cdots + x_n w_n.$$

Important! When working with a diagonalizable matrix, everything is much easier if you expand your vectors in an eigenbasis!

Procedure for Solving a Difference Equation:

Consider a difference equation

$$v_{k+1} = Av_k \quad \text{with initial state } v_0.$$

(1) Diagonalize A to get an eigenbasis $\{w_1, \dots, w_n\}$ with eigenvalues $\lambda_1, \dots, \lambda_n$.

Stop if the matrix is not diagonalizable: this procedure fails.

(2) Expand v_0 in the eigenbasis: i.e., solve

$$v_0 = x_1 w_1 + \dots + x_n w_n$$

Solution: $v_k = A^k v_0 = \lambda_1^k x_1 w_1 + \dots + \lambda^k x_n w_n$

Of course, this only works if A is diagonalizable.

Procedure for Diagonalizing a Matrix:

Let A be an $n \times n$ matrix.

(1) Compute the characteristic polynomial

$$p(\lambda) = \det(A - \lambda I_n)$$

(2) Factor $p(\lambda)$ to find the eigenvalues of A.

(3) Find a basis for each eigenspace.

(4) Combine your bases in (3).

- If you have n vectors, they form an eigenbasis.
- Otherwise, A is not diagonalizable.

Eg: We ran this procedure on $A = \begin{pmatrix} 0 & 13 & 12 \\ 1/4 & 0 & 0 \\ 0 & 1/2 & 0 \end{pmatrix}$ above.

Eg: Solve the difference equation

$$v_{k+1} = Av_k \quad \text{for} \quad A = \begin{pmatrix} 14 & -18 & -33 \\ -12 & 20 & 33 \\ 12 & -18 & -31 \end{pmatrix} \quad v_0 = \begin{pmatrix} 6 \\ 0 \\ 2 \end{pmatrix}$$

First we diagonalize A.

$$p(\lambda) = \det\begin{pmatrix} 14-\lambda & -18 & -33 \\ -12 & 20-\lambda & 33 \\ 12 & -18 & -31-\lambda \end{pmatrix}$$

$$= (14-\lambda)\det\begin{pmatrix} 20-\lambda & 33 \\ -18 & -31-\lambda \end{pmatrix} - 12(-1)\det\begin{pmatrix} -18 & -33 \\ -18 & -31-\lambda \end{pmatrix}$$
$$+ 12\det\begin{pmatrix} -18 & -33 \\ 20-\lambda & 33 \end{pmatrix}$$

$$= \cdots = -\lambda^3 + 3\lambda^2 - 4$$

Ask a computer for the roots:

$$p(\lambda) = -(\lambda-2)^2(\lambda+1)$$

twice?

So the eigenvalues are $\lambda = 2$ and $\lambda = -1$.

Let's find bases for eigenspaces:

$\lambda = 2$: $A - 2I_3 = \begin{pmatrix} 12 & -18 & -33 \\ -12 & 18 & 33 \\ 12 & -18 & -33 \end{pmatrix} \xrightarrow{\text{rref}} \begin{pmatrix} 1 & -3/2 & -11/4 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$

$$\xrightarrow{\text{PVF}} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = x_2\begin{pmatrix} 3/2 \\ 1 \\ 0 \end{pmatrix} + x_3\begin{pmatrix} 11/4 \\ 0 \\ 1 \end{pmatrix}$$

$$\xrightarrow{\text{basis}} \left\{ \begin{pmatrix} 3/2 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 11/4 \\ 0 \\ 1 \end{pmatrix} \right\}$$

Let's clear denominators to make our lives easier:

$$w_1 = \begin{pmatrix} 3 \\ 2 \\ 0 \end{pmatrix} \qquad w_2 = \begin{pmatrix} 11 \\ 0 \\ 4 \end{pmatrix}$$

$\lambda = -1$: $A + I_3 = \begin{pmatrix} 15 & -18 & -33 \\ -12 & 21 & 33 \\ 12 & -18 & -30 \end{pmatrix} \overset{\text{rref}}{\rightsquigarrow} \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{pmatrix}$

$$\overset{\text{PVF}}{\rightsquigarrow} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = x_3 \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix} \xrightarrow{\text{basis}} w_3 = \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}$$

We have 3 eigenvectors $w_1, w_2, w_3 \Rightarrow A$ is diagonalizable with eigenbasis

$$\{w_1, w_2, w_3\} = \left\{ \begin{pmatrix} 3 \\ 2 \\ 0 \end{pmatrix}, \begin{pmatrix} 11 \\ 0 \\ 4 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix} \right\}$$

Now we expand $v_0 = \begin{pmatrix} 6 \\ 0 \\ 2 \end{pmatrix}$ in our eigenbasis:

$$\begin{pmatrix} 6 \\ 0 \\ 2 \end{pmatrix} = x_1 \begin{pmatrix} 3 \\ 2 \\ 0 \end{pmatrix} + x_2 \begin{pmatrix} 11 \\ 0 \\ 4 \end{pmatrix} + x_3 \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}$$

$$\overset{\text{aug}}{\underset{\text{matrix}}{\rightsquigarrow}} \left(\begin{array}{ccc|c} 3 & 11 & 1 & 6 \\ 2 & 0 & -1 & 0 \\ 0 & 4 & 1 & 2 \end{array}\right) \overset{\text{rref}}{\rightsquigarrow} \left(\begin{array}{ccc|c} 1 & 0 & 0 & -1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & -2 \end{array}\right)$$

$$\rightsquigarrow v_0 = -w_1 + w_2 - 2w_3$$

Now we're done:

$$v_k = A^k v_0 = -2^k w_1 + 2^k w_2 - 2(-1)^k w_3$$

$$= -2^k \begin{pmatrix} 3 \\ 2 \\ 0 \end{pmatrix} + 2^k \begin{pmatrix} 11 \\ 0 \\ 4 \end{pmatrix} - 2(-1)^k \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}$$

$$= \begin{pmatrix} 8 \cdot 2^k - 2 \cdot (-1)^k \\ -2 \cdot 2^k + 2 \cdot (-1)^k \\ 4 \cdot 2^k - 2(-1)^k \end{pmatrix} \longleftarrow \text{closed form:}$$

no matrix multiplication required!

NB: $2^k \gg (-1)^k$, so $v_k \sim 2^k (w_2 - w_1) = 2^k \begin{pmatrix} 8 \\ -2 \\ 4 \end{pmatrix}$

as $k \to \infty$

Eg: $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ shear

$$p(\lambda) = \lambda^2 - Tr(A)\lambda + \det(A)$$
$$= \lambda^2 - 2\lambda + 1 = (\lambda-1)^2$$

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

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

We only got one eigenvector $(1,0)$ and not two:

$\Longrightarrow$ not diagonalizable! (all eigenvectors lie on the x-axis.)

So we can't use diagonalization to solve

$$v_{k+1} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} v_k.$$

NB: The shear should be your favorite example of a non-diagonalizable matrix.

Fact: A matrix with "random entries" will be diagonalizable.

In the diagonalization procedure, how did we know that when we combined our eigenbases we would get a linearly independent set of vectors?

Fact: If $w_1, \ldots, w_p$ are eigenvectors of A with different eigenvalues then $\{w_1, \ldots, w_p\}$ is LI.

More generally, say

- $\{w_1, w_2\}$ is a basis for the λ_1-eigenspace
- $\{w_3\}$ is a basis for the λ_2-eigenspace.

Suppose $x_1 w_1 + x_2 w_2 + x_3 w_3 = 0$.

- $x_1 w_1 + x_2 w_2$ is in the λ_1-eigenspace
- Since $(x_1 w_1 + x_2 w_2) + x_3 w_3 = 0$, the Fact implies $x_1 w_1 + x_2 w_2 = 0$ and $x_3 w_3 = 0$ (so $x_3 = 0$)
- Since $\{w_1, w_2\}$ is LI, this implies $x_1 = x_2 = 0$.

This shows $\{w_1, w_2, w_3\}$ is LI ✓

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

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

$$Aw_1 = \lambda_1 w_1, \ldots, Aw_n = \lambda_n w_n$$

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

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 $w_{i+1} \in \operatorname{Span}\{w_1, \dots, w_i\}$, so

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

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

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

If $\lambda_{i+1} = 0$ then $\lambda_1 x_1 w_1 + \cdots + \lambda_i x_i w_i = 0 \xRightarrow[\text{LI}]{\{w_1, \dots, w_i\}}$ $x_1 = \cdots = x_i = 0$ (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

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

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

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

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

$$\Rightarrow x_1 = \cdots = x_i = 0$$

which is impossible, as before.