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 $\lambda$.

* The associated eigenvalue is $\lambda$.

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

Given an eigenvalue $\lambda$, we know how to compute all $\lambda$-eigenvectors: the $\lambda$-eigenspace is

\[ \text{Null}(A - \lambda I_n) \]

How do you find the eigenvalues of $A$?

$\lambda$ is an eigenvalue of $A$ if and only if $A - \lambda I_n$ is not invertible

\[ \iff \det(A - \lambda I_n) = 0 \]

Def: The characteristic polynomial of an $n \times n$ matrix $A$ is $p(\lambda) = \det(A - \lambda I_n)$
\[ \lambda \text{ is an eigenvalue of } A \iff p(\lambda) = 0 \]

**Eg:** Find all eigenvalues of
\[
A = \begin{pmatrix}
\tfrac{1}{4} & 13 & 12 \\
0 & \tfrac{1}{2} & 0 \\
0 & 0 & 0
\end{pmatrix}
\]

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

\[
\text{expand factors} \quad \det(A - \lambda I_3) = -\lambda^3 + \frac{13}{4} \lambda^2 + \frac{3}{2} = p(\lambda)
\]

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 \(\lambda_0\).
  
  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!
Ex continued: We know 2 is an eigenvalue of the rabbit matrix. Check:
\[
p(2) = -8 + \frac{13}{4} \cdot 2 + \frac{3}{2} = -\frac{16}{2} + \frac{13}{2} + \frac{3}{2} = 0
\]
This means \((\lambda-2)\) divides \(p(\lambda)\). What's \(p(\lambda)/(\lambda-2)\)?

**Synthetic division** (long division of polynomials):

\[
\begin{array}{c|ccccc}
\lambda-2 & -\lambda^2 & -2\lambda & -\frac{3}{4} \\
\hline
& -\lambda^3 & +\frac{13}{4}\lambda & +\frac{3}{2} \\
\hline
& -2\lambda^2 & +\frac{13}{4}\lambda & +\frac{3}{2} \\
& & -2\lambda(\lambda-2) \\
& & -4\lambda & +\frac{13}{4}\lambda & +\frac{3}{2} \\
& & & -\frac{3}{4}(\lambda-2) \\
& & & & 0
\end{array}
\]

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

**Quadratic formula:** roots of \(-\lambda^2 - 2\lambda - \frac{3}{4}\) are
\[
-\frac{1}{2} \left(2 \pm \sqrt{4 - 3^2} \right) = -\frac{1}{2}, -\frac{3}{2}
\]

\[\Rightarrow -\lambda^3 + \frac{13}{4}\lambda + \frac{3}{2} = -(\lambda - 2)(\lambda + \frac{1}{2})(\lambda + \frac{3}{2})\]
So the eigenvalues are 2, \(-\frac{1}{2}\), \(-\frac{3}{2}\).
To check these are eigenvalues, let's find some eigenvectors!

\[ 2 \rightarrow \text{Nul}(A - 2I_3) = \text{Nul} \begin{pmatrix} -2 & 1 & 3 \\ 1 & 0 & 2 \\ 1 & 0 & 2 \end{pmatrix} = \text{Span} \left\{ \begin{pmatrix} 1 \\ 3 \\ 1 \end{pmatrix} \right\} \]

\[ -\frac{1}{2} \rightarrow \text{Nul}(A + \frac{1}{2}I_3) = \text{Nul} \begin{pmatrix} 1 & 1 & 3 \\ 1 & 0 & 2 \\ 1 & 0 & 2 \end{pmatrix} = \text{Span} \left\{ \begin{pmatrix} 2 \\ -1 \\ 1 \end{pmatrix} \right\} \]

\[ -\frac{3}{2} \rightarrow \text{Nul}(A + \frac{3}{2}I_3) = \text{Nul} \begin{pmatrix} 1 & 1 & 3 \\ 1 & 0 & 2 \\ 1 & 0 & 2 \end{pmatrix} = \text{Span} \left\{ \begin{pmatrix} 2 \\ -1 \\ 1 \end{pmatrix} \right\} \]

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 2x2 matrices first.

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

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

\[ = \lambda^2 - (a + d)\lambda + (ad - bc) \]

This is a polynomial of degree 2.

Def: The trace of a matrix \( A \) is \( \text{Tr}(A) = \) the sum of the diagonal entries of \( A \).

Eg: \( A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \quad \text{Tr}(A) = a + d \)
Characteristic Polynomial of a 2x2 Matrix $A$

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

**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

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

$$+ \text{(other terms)} + \det(A)$$

→ This is a degree-$n$ polynomial

→ You only get the $\lambda^{n-1}$ and constant coeffs “for free” — the rest are more complicated.

**Eg:** $A = \begin{pmatrix} 6 & 13 & 12 \\ 14 & 0 & 0 \\ 12 & 0 & 0 \end{pmatrix} \rightarrow p(\lambda) = -\lambda^3 + 0\lambda^2 + \frac{12}{4} \lambda + \frac{3}{2}$

$\text{Tr}(A) = 0 + 0 + 0 = 0 \checkmark$ \hspace{1cm} $\det(A) = -\frac{1}{4} \cdot \left( -\frac{12}{2} \right) = \frac{3}{2} \checkmark$

**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

\[
A = \begin{pmatrix}
1 & 1 & 3 \\
0 & 1 & 2 \\
0 & 1 & 0
\end{pmatrix}
\]

has eigenvalues 2, \(-\frac{1}{3}\), \(-\frac{3}{2}\)

& eigenspaces

\[
2: \text{Span}\left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \right\} \quad -\frac{1}{3}: \text{Span}\left\{ \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \right\} \quad -\frac{3}{2}: \text{Span}\left\{ \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\}
\]

Let’s give names to some eigenvectors:

\[
\omega_1 = \begin{pmatrix} 3 \\ 2 \\ 1 \end{pmatrix}, \quad \omega_2 = \begin{pmatrix} 2 \\ -1 \\ 0 \end{pmatrix}, \quad \omega_3 = \begin{pmatrix} 1 \\ 8 \\ -3 \end{pmatrix}
\]

We know what happens if we start with \(\begin{pmatrix} 3 \\ 2 \\ 1 \end{pmatrix}\) rabbits: \(A^k\begin{pmatrix} 3 \\ 2 \\ 1 \end{pmatrix} = 2^k\begin{pmatrix} 3 \\ 2 \\ 1 \end{pmatrix}\) \(\Rightarrow\) doubles each time.

What if we start with \(v_0 = \begin{pmatrix} 1 \\ 6 \\ 9 \end{pmatrix}\)?

Fact: \(v_0 = \begin{pmatrix} 1 \\ 6 \\ 9 \end{pmatrix}\) can be written as a linear combination of eigenvectors:

\[
\begin{pmatrix} 1 \\ 6 \\ 9 \end{pmatrix} = \omega_1 + \omega_2 - \omega_3
\]

Now it’s easy to compute \(A^k v_0\):

\[
A^k v_0 = A^k(\omega_1 + \omega_2 - \omega_3) = A^k \omega_1 + A^k \omega_2 - A^k \omega_3
\]

\[
= 2^k \omega_1 + (-\frac{1}{3})^k \omega_2 + (-\frac{3}{2})^k \omega_3
\]
Observation 1: $2^k \gg \left| \left( -\frac{1}{2} \right)^k \right|$ and $\left| \left( -\frac{3}{2} \right)^k \right|$ for large $k$

So $A^k v_0 \approx 2^k w_1$

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)

$$\Rightarrow \{w_1, w_2, w_3\} \text{ is a basis for } \mathbb{R}^3$$

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

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

$$A^k v_0 = x_1 A^k w_1 + x_2 A^k w_2 + x_3 A^k w_3$$

$$= 2^k x_1 w_1 + \left( -\frac{3}{2} \right)^k x_2 w_2 + \left( -\frac{3}{2} \right)^k x_3 w_3$$

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 $v_1, \ldots, v_n$. In this case, $\{v_1, \ldots, v_n\}$ is called an eigenbasis.

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

1) Solve $v = x_1 v_1 + \ldots + x_n v_n$

2) $A^k v = x_1^k v_1 + \ldots + x_n^k v_n$ ← vector form

($\lambda_i$ = eigenvalue for $v_i$: $Av_i = \lambda_i v_i$)

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 $\implies$ 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.
\[ C = \begin{pmatrix} w_1 & \cdots & w_n \end{pmatrix} \quad D = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix} \quad A_w = \lambda_i w_i \]

**Example:**

\[ A = \begin{pmatrix} 0 & 13 & 12 \\ 14 & 0 & 0 \\ 0 & 12 & 0 \end{pmatrix} \Rightarrow A = CDC^{-1} \text{ for } C = \begin{pmatrix} 32 & 2 & 18 \\ 4 & -1 & -3 \\ 1 & 1 & 1 \end{pmatrix} \quad D = \begin{pmatrix} 2 & 0 & 0 \\ 0 & -\frac{1}{2} & 0 \\ 0 & 0 & -3/2 \end{pmatrix} \]

**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

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

\[ = C \begin{pmatrix} x_1 \\ 0 \\ \vdots \\ 0 \\ x_n \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ \vdots \\ \vdots \\ x_n \end{pmatrix} \]

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

\[ = A (x_1 w_1 + \cdots + x_n w_n) = Av \]
NB: If $A = CDC^{-1}$ then

$$A^k = (CDC^{-1})^k = (CDC^{-1})(CDC^{-1})... (CDC^{-1}) = CD^kC^{-1} = C \left( \begin{array}{ccc} \lambda_1^k & 0 & 0 \\ 0 & \lambda_2^k & 0 \\ 0 & 0 & \lambda_n^k \end{array} \right) C^{-1}$$

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

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

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 -- we can take $C = I_3$, so the diagonalization is

$$A = I_3, \quad A = I_3$$

Q: What if we take $e_2$ to be our first eigenvector?
The only eigenvalue is 1, and the 1-eigenspace is 
\[ \text{Nul}(A - I_2) = \text{Nul}([0 \ 0 ; 0 \ 0]) = \text{Span} \{ (1) \} \]
So all eigenvectors lie on the x-axis
\( \Rightarrow \text{not diagonalizable!} \)
(cf. p.5 of W10L1 notes)

Procedure to Diagonalize a Matrix:
1. Compute the characteristic polynomial \( p(\lambda) \)
2. Find the roots of \( p(\lambda) = \text{eigenvalues of } A \)
3. Find a basis for each eigenspace = \( \text{Nul}(A - \lambda I_n) \)
   (using DVF)

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

Fact: If \( \omega_1, \ldots, \omega_p \) are eigenvectors with different eigenvalues then \( \{\omega_1, \ldots, \omega_p\} \) 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 $A w_i = \lambda_i w_i$ and all of the $\lambda_i, \ldots, \lambda_n$ are distinct. Suppose $\{w_1, \ldots, w_n\}$ is LD. Then for some $j \neq i$, $\{w_1, \ldots, w_j\}$ is LI but $w_i \in \text{Span}\{w_1, \ldots, w_j\}$, so

$$w_i = x_1 w_1 + \cdots + x_j w_j$$

$$\Rightarrow A w_i = A (x_1 w_1 + \cdots + x_j w_j)$$

$$\Rightarrow \lambda_i w_i = \lambda_1 x_1 w_1 + \cdots + \lambda_j x_j w_j$$

If $\lambda_i = 0$ then $\lambda_1 x_1 w_1 + \cdots + \lambda_j x_j w_j = 0 \Rightarrow x_1 = \cdots = x_j = 0$ (because $\lambda_1, \ldots, \lambda_j \neq 0$), so $w_i = 0$, which can't happen because $w_i$ is an eigenvector.

If $\lambda_i \neq 0$ then

$$w_i = x_1 w_1 + \cdots + x_j w_j$$

Subtract $w_i = x_1 w_1 + \cdots + x_j w_j$

$$\Rightarrow 0 = (\frac{\lambda_1}{\lambda_i} - 1) x_1 w_1 + \cdots + (\frac{\lambda_j}{\lambda_i} - 1) x_j w_j$$

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

$$\Rightarrow x_1 = \cdots = x_j = 0$$

which is impossible, as before.
Consequence: If $A$ has $n$ (different) eigenvalues then $A$ is diagonalizable.

Indeed, if $\lambda_1, \ldots, \lambda_n$ are eigenvalues and $A\omega_1 = \lambda_1 \omega_1, \ldots, A\omega_n = \lambda_n \omega_n$ then $\{\omega_1, \ldots, \omega_n\}$ is an eigenbasis by the Fact.