Section 5.2

The Characteristic Equation

The Characteristic Polynomial

Last section we learn that for a square matrix A:

$$\begin{aligned}\lambda \text{ is an eigenvalue of } A &\iff Ax = \lambda x \text{ has a nontrivial solution} \\ &\iff (A - \lambda I)x = 0 \text{ has a nontrivial solution} \\ &\iff A - \lambda I \text{ is not invertible} \\ &\iff \det(A - \lambda I) = 0.\end{aligned}$$

Compute Eigenvalues

The *eigenvalues* of A are **the roots** of $\det(A - \lambda I)$, which is the characteristic polynomial of A.

Definition

Let A be a square matrix. The **characteristic polynomial** of A is

$$f(\lambda) = \det(A - \lambda I).$$

The **characteristic equation** of A is the equation

$$f(\lambda) = \det(A - \lambda I) = 0.$$

The Characteristic Polynomial

Example

Question: What are the eigenvalues of

$$A = \begin{pmatrix} 5 & 2 \\ 2 & 1 \end{pmatrix}?$$

Answer: First we find the *characteristic polynomial*:

$$\begin{aligned} f(\lambda) &= \det(A - \lambda I) = \det\left[\begin{pmatrix} 5 & 2 \\ 2 & 1 \end{pmatrix} - \begin{pmatrix} \lambda & 0 \\ 0 & \lambda \end{pmatrix}\right] = \det\begin{pmatrix} 5-\lambda & 2 \\ 2 & 1-\lambda \end{pmatrix} \\ &= (5-\lambda)(1-\lambda) - 2\cdot 2 \\ &= \lambda^2 - 6\lambda + 1. \end{aligned}$$

The eigenvalues are the roots of the characteristic polynomial, which we can find *using the quadratic formula*:

$$\lambda = \frac{6 \pm \sqrt{36-4}}{2} = 3 \pm 2\sqrt{2}.$$

Definition

The **trace** of a square matrix A is $\mathrm{Tr}(A) =$ sum of the diagonal entries of A.

What do you notice about: the characteristic polynomial of $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$?

Answer:

$$\det(A - \lambda I) = \det \begin{pmatrix} a-\lambda & b \\ c & d-\lambda \end{pmatrix} = (a-\lambda)(d-\lambda) - bc$$
$$= \lambda^2 - (a+d)\lambda + (ad - bc)$$

- The coefficient of λ is the trace of A and the constant term is $\det(A)$.
- Recall that A is not invertible if and only if $\lambda = 0$ is a root.

Shortcut

The characteristic polynomial of a 2×2 matrix A is

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

Question: What are the eigenvalues of the *rabbit population matrix*

$$A = \begin{pmatrix} 0 & 6 & 8 \\ \frac{1}{2} & 0 & 0 \\ 0 & \frac{1}{2} & 0 \end{pmatrix}?$$

Answer: First we find the characteristic polynomial:

$$\begin{aligned} f(\lambda) &= \det(A - \lambda I) = \det \begin{pmatrix} -\lambda & 6 & 8 \\ \frac{1}{2} & -\lambda & 0 \\ 0 & \frac{1}{2} & -\lambda \end{pmatrix} \\ &= 8\left(\frac{1}{4} - 0 \cdot -\lambda\right) - \lambda\left(\lambda^2 - 6 \cdot \frac{1}{2}\right) \\ &= -\lambda^3 + 3\lambda + 2. \end{aligned}$$

Already know one *eigenvalue is* $\lambda = 2$, check : $f(2) = -8 + 6 + 2 = 0$.

Doing polynomial long division, we get:

$$\frac{-\lambda^3 + 3\lambda + 2}{\lambda - 2} = -\lambda^2 - 2\lambda - 1 = -(\lambda + 1)^2.$$

Hence $f(\lambda) = -(\lambda + 1)^2(\lambda - 2)$ and so $\lambda = -1$ *is also* an eigenvalue.

Algebraic Multiplicity

Definition

The **algebraic multiplicity** of an eigenvalue λ is its *multiplicity as a root* of the characteristic polynomial.

There is a geometric multiplicity notion, but this one is easier to work with.

Example

In the rabbit population matrix, $f(\lambda) = -(\lambda - 2)(\lambda + 1)^2$. The algebraic multiplicities are

$$\lambda = \begin{cases} 2 & \text{multiplicity } 1, \\ -1 & \text{multiplicity } 2 \end{cases}$$

Example

In the matrix $\begin{pmatrix} 5 & 2 \\ 2 & 1 \end{pmatrix}$, $f(\lambda) = (\lambda - (3 + 2\sqrt{2}))(\lambda - (3 - 2\sqrt{2}))$. The algebraic multiplicities are $\lambda = \begin{cases} 3 + 2\sqrt{2} & \text{alg. multiplicity } 1, \\ 3 - 2\sqrt{2} & \text{alg. multiplicity } 1 \end{cases}$

Multiplicities

Theorem

If A is an $n \times n$ matrix, the **characteristic polynomial**

$$f(\lambda) = \det(A - \lambda I)$$

is a polynomial of degree n, and its roots are the eigenvalues of A:

$$f(\lambda) = (-1)^n \lambda^n + a_{n-1}\lambda^{n-1} + a_{n-2}\lambda^{n-2} + \cdots + a_1\lambda + a_0.$$

Complex numbers

If you **count the eigenvalues** of A, with their algebraic multiplicities, depending on *whether you allow complex eigenvalues*, you will get :

- **Do allow** complex numbers: *Always n*.
- Only real numbers: Always at most n, but *sometimes less*.

This is because any degree-n polynomial has exactly n *complex roots*, counted with multiplicity. Stay tuned!

Similarity

Definition

Two $n \times n$ matrices A and B are **similar** if there is an invertible $n \times n$ matrix C such that

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

The intuition

C keeps record of a basis $\mathcal{C} = \{v_1, \ldots, v_n\}$ of $\mathbf{R}^n$.

B transforms the $\mathcal{C}$-coordinates of x: $B[x]_{\mathcal{C}} = [Ax]_{\mathcal{C}}$ in *the same way that* A transforms the standard coordinates of x

Why does it work?

- First, $\mathcal{C} = \{v_1, v_2, \ldots, v_n\}$ is a basis for $\mathbf{R}^n$ (C is invertible), so

$$w = c_1 v_1 + c_2 v_2 + c_n v_n = C[w]_{\mathcal{C}}$$

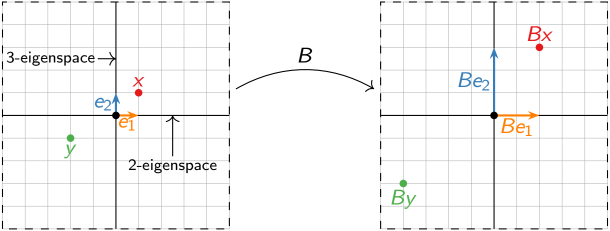
- *Using $\mathcal{C}$-coordinates* for any vector w, is $[w]_{\mathcal{C}} = C^{-1} w$.
- Then $A = CBC^{-1}$ implies $C^{-1}A = BC^{-1}$. Using $\mathcal{C}$-coordinates:

$$[Ax]_{\mathcal{C}} = C^{-1}(Ax) = B(C^{-1}x) = B([x]_{\mathcal{C}}).$$

$$A = \begin{pmatrix} 1 & 2 \\ -1 & 4 \end{pmatrix} \quad B = \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix} \quad C = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix} \quad \implies \quad A = CBC^{-1}.$$

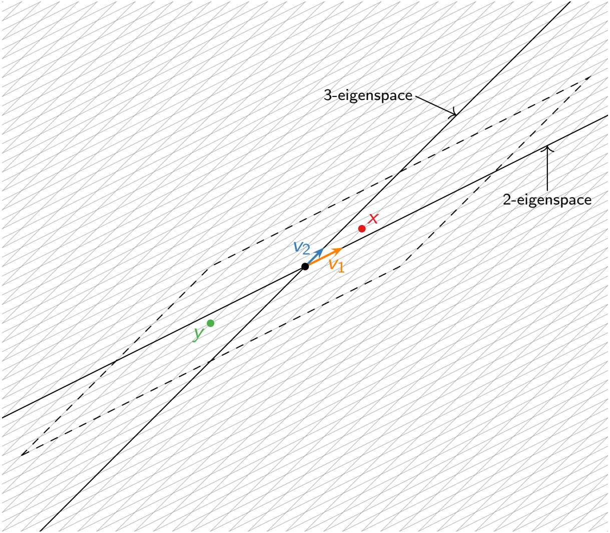
What does B do geometrically? **Scaling:** x-direction by 2 and y-direction by 3.

B acting on the usual coordinates



Now A will do to the standard coordinates what B does to the $\mathcal{C}$-coordinates, where $\mathcal{C} = \left\{ \begin{pmatrix} 2 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right\}$.

From $\mathcal{C}$-coordinates to standard coordinates



$$\left.\begin{aligned} v_1 &= \begin{pmatrix} 2 \\ 1 \end{pmatrix} \\ v_2 &= \begin{pmatrix} 1 \\ 1 \end{pmatrix} \end{aligned}\right\} \text{vectors in } \mathcal{C}$$

$$[x]_{\mathcal{C}} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

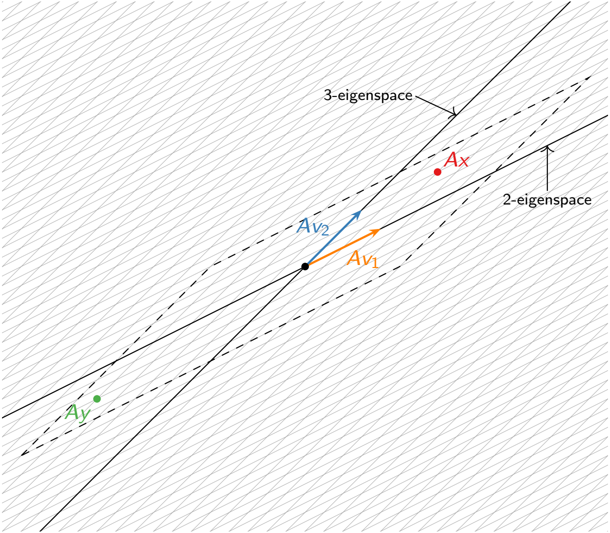
$$x = v_1 + v_2 = \begin{pmatrix} 3 \\ 2 \end{pmatrix}$$

$$[y]_{\mathcal{C}} = \begin{pmatrix} -2 \\ -1 \end{pmatrix}$$

$$y = -2v_1 - v_2$$

$$= \begin{pmatrix} -5 \\ -3 \end{pmatrix}$$

A does to the usual coordinates what B does to the $\mathcal{C}$-coordinates



$$Av_1 = 2v_1 = \begin{pmatrix}4\\2\end{pmatrix}$$

$$Av_2 = 3v_2 = \begin{pmatrix}3\\3\end{pmatrix}$$

$$B[x]_{\mathcal{C}} = \begin{pmatrix}2\\3\end{pmatrix} = [Ax]_{\mathcal{C}}$$

$$Ax = 2v_1 + 3v_2 = \begin{pmatrix}7\\5\end{pmatrix}$$

$$B[y]_{\mathcal{C}} = \begin{pmatrix}-4\\-3\end{pmatrix} = [Ay]_{\mathcal{C}}$$

$$Ay = -4v_1 - 3v_2 = \begin{pmatrix}-11\\-7\end{pmatrix}$$

Check: $Ax = \begin{pmatrix}1 & 2\\-1 & 4\end{pmatrix}\begin{pmatrix}3\\2\end{pmatrix} = \begin{pmatrix}7\\5\end{pmatrix}$ $Ay = \begin{pmatrix}1 & 2\\-1 & 4\end{pmatrix}\begin{pmatrix}-5\\-3\end{pmatrix} = \begin{pmatrix}-11\\-7\end{pmatrix}$ ✔

Similar Matrices Have the Same Characteristic Polynomial

Fact

If A and B are **similar**,
then they have the *same characteristic polynomial*.

Consequence:
similar matrices have the *same eigenvalues*! Though different eigenvectors in general.

Why? Suppose $A = CBC^{-1}$. We can show that $\det(A - \lambda I) = \det(B - \lambda I)$.

$$\begin{aligned} A - \lambda I &= CBC^{-1} - \lambda I \\ &= CBC^{-1} - C(\lambda I)C^{-1} \\ &= C(B - \lambda I)C^{-1}. \end{aligned}$$

Therefore,

$$\begin{aligned} \det(A - \lambda I) &= \det\big(C(B - \lambda I)C^{-1}\big) \\ &= \det(C)\det(B - \lambda I)\det(C^{-1}) \\ &= \det(B - \lambda I), \end{aligned}$$

because $\det(C^{-1}) = \det(C)^{-1}$.

Similarity Caveats

Warning

1. Matrices with the *same eigenvalues* need not be similar.
   For instance,

$$\begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$$

   both only have the eigenvalue 2, but they are *not similar*.

2. Similarity is *lost in row equivalence*.
   For instance,

$$\begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

   are row equivalent, but they have *different eigenvalues*.