Section 5.2

The Characteristic Equation

We have a couple of new ways of saying "A is invertible" now:

The Invertible Matrix Theorem

Let A be a square $n \times n$ matrix, and let $T\colon \mathbf{R}^n \to \mathbf{R}^n$ be the linear transformation $T(x) = Ax$. The following statements are equivalent.

1. A is invertible.
2. T is invertible.
3. A is row equivalent to I_n.
4. A has n pivots.
5. $Ax = 0$ has only the trivial solution.
6. The columns of A are linearly independent.
7. T is one-to-one.
8. $Ax = b$ is consistent for all b in $\mathbf{R}^n$.
9. The columns of A span $\mathbf{R}^n$.
10. T is onto.
11. A has a left inverse (there exists B such that $BA = I_n$).
12. A has a right inverse (there exists B such that $AB = I_n$).
13. A^T is invertible.
14. The columns of A form a basis for $\mathbf{R}^n$.
15. $\operatorname{Col} A = \mathbf{R}^n$.
16. $\dim \operatorname{Col} A = n$.
17. $\operatorname{rank} A = n$.
18. $\operatorname{Nul} A = \{0\}$.
19. $\dim \operatorname{Nul} A = 0$.

19. The determinant of A is *not* equal to zero.
20. The number 0 is *not* an eigenvalue of A.

The Characteristic Polynomial

Let A be a square matrix.

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

This gives us a way to compute the eigenvalues 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

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

Important

The eigenvalues of A are the roots of the characteristic polynomial $f(\lambda) = \det(A - \lambda I)$.

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:

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

Question: What is the characteristic polynomial of

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

Answer:

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

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

What do you notice about $f(\lambda)$?

- The constant term is $\det(A)$, which is zero if and only if $\lambda = 0$ is a root.
- The linear term $-(a + d)$ is the negative of the sum of the diagonal entries of A.

Definition

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

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:

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

We know from before that one eigenvalue is $\lambda = 2$: indeed, $f(2) = -8 + 6 + 2 = 0$. Doing polynomial long division, we get:

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

Hence $\lambda = -1$ is also an eigenvalue.

Algebraic Multiplicity

Definition

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

This is not a very interesting notion *yet*. It will become interesting when we also define *geometric* multiplicity later.

Example

In the rabbit population matrix, $f(\lambda) = -(\lambda - 2)(\lambda + 1)^2$, so the algebraic multiplicity of the eigenvalue 2 is 1, and the algebraic multiplicity of the eigenvalue -1 is 2.

Example

In the matrix $\begin{pmatrix} 5 & 2 \\ 2 & 1 \end{pmatrix}$, $f(\lambda) = (\lambda - (3 - 2\sqrt{2}))(\lambda - (3 + 2\sqrt{2}))$, so the algebraic multiplicity of $3 + 2\sqrt{2}$ is 1, and the algebraic multiplicity of $3 - 2\sqrt{2}$ is 1.

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

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

turns out to be a polynomial of degree n, and its roots are the eigenvalues of A:

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

Poll

If you count the eigenvalues of A, with their algebraic multiplicities, you will get:

A. Always n.

B. Always at most n, but sometimes less.

C. Always at least n, but sometimes more.

D. None of the above.

The answer depends on whether you allow *complex* eigenvalues. If you only allow real eigenvalues, the answer is B. Otherwise it is A, because any degree-n polynomial has exactly n *complex* roots, counted with multiplicity. Stay tuned.

Recall: If $\{v_1, v_2, \ldots, v_m\}$ is a basis for a subspace V and x is in V, then the $\mathcal{B}$-**coordinates** of x are the (unique) coefficients $c_1, c_2, \ldots, c_m$ such that

$$x = c_1 v_1 + c_2 v_2 + \cdots + c_m v_m.$$

In this case, the $\mathcal{B}$-**coordinate vector** of x is

$$[x]_{\mathcal{B}} = \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_m \end{pmatrix}.$$

Example: The vectors

$$v_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \qquad v_2 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

form a basis for $\mathbf{R}^2$ because they are not collinear. [interactive]

Coordinate Systems on $\mathbf{R}^n$

Recall: A set of n vectors $\{v_1, v_2, \ldots, v_n\}$ form a basis for $\mathbf{R}^n$ if and only if the matrix C with columns $v_1, v_2, \ldots, v_n$ is invertible.

If $x = c_1v_1 + c_2v_2 + \cdots + c_nv_n$ then

$$[x]_{\mathcal{B}} = \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{pmatrix} \implies x = c_1v_1 + c_2v_2 + c_nv_n = C[x]_{\mathcal{B}}.$$

Since $x = C[x]_{\mathcal{B}}$ we have $[x]_{\mathcal{B}} = C^{-1}x$.

Translation: Let $\mathcal{B}$ be the basis of columns of C. Multiplying by C changes from the $\mathcal{B}$-coordinates to the usual coordinates, and multiplying by C^{-1} changes from the usual coordinates to the $\mathcal{B}$-coordinates:

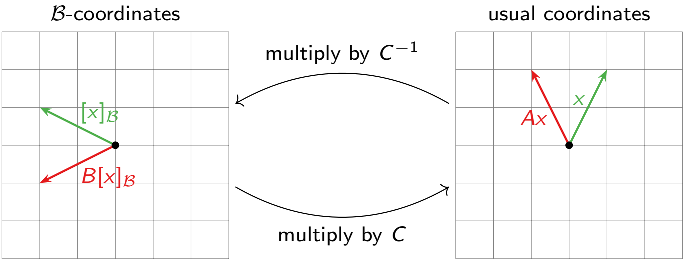
$$[x]_{\mathcal{B}} = C^{-1}x \qquad x = C[x]_{\mathcal{B}}.$$

Definition

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

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

What does this mean? This gives you a different way of thinking about multiplication by A. Let $\mathcal{B}$ be the basis of columns of C.



To compute Ax, you:

1. multiply x by C^{-1} to change to the $\mathcal{B}$-coordinates: $[x]_{\mathcal{B}} = C^{-1}x$
2. multiply this by B: $B[x]_{\mathcal{B}} = BC^{-1}x$
3. multiply this by C to change to usual coordinates: $Ax = CBC^{-1}x = CB[x]_{\mathcal{B}}$.

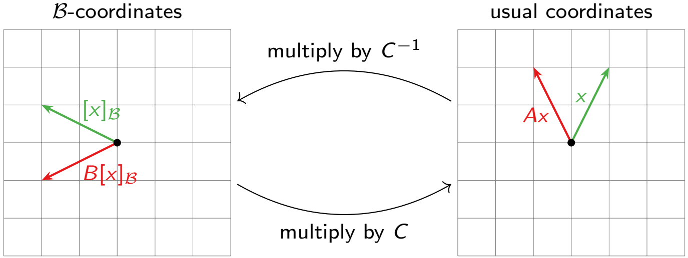
Similarity

Definition

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

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

What does this mean? This gives you a different way of thinking about multiplication by A. Let $\mathcal{B}$ be the basis of columns of C.



If $A = CBC^{-1}$, then A and B do the same thing, but B operates on the $\mathcal{B}$-coordinates, where $\mathcal{B}$ is the basis of columns of C.

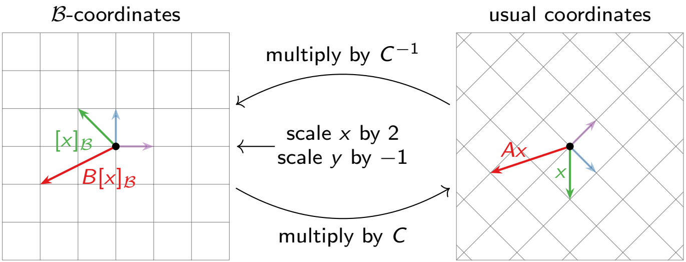
$$A = \begin{pmatrix} 1/2 & 3/2 \\ 3/2 & 1/2 \end{pmatrix} \quad B = \begin{pmatrix} 2 & 0 \\ 0 & -1 \end{pmatrix} \quad C = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \qquad A = CBC^{-1}.$$

What does B do geometrically?

It scales the x-direction by 2 and the y-direction by -1.

To compute Ax, first change to the $\mathcal{B}$ coordinates, then multiply by B, then change back to the usual coordinates, where

$$\mathcal{B} = \left\{ \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \end{pmatrix} \right\} = \{v_1, v_2\} \qquad \text{(the columns of } C\text{).}$$



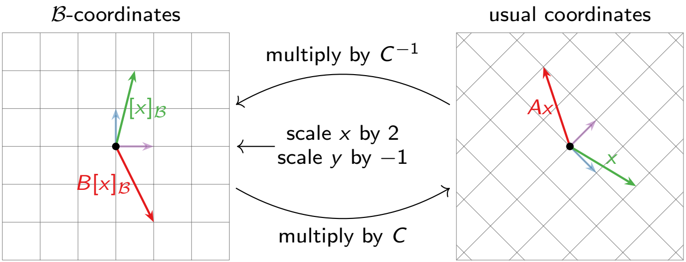
$$A = \begin{pmatrix} 1/2 & 3/2 \\ 3/2 & 1/2 \end{pmatrix} \quad B = \begin{pmatrix} 2 & 0 \\ 0 & -1 \end{pmatrix} \quad C = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \qquad A = CBC^{-1}.$$

What does B do geometrically?

It scales the x-direction by 2 and the y-direction by -1.

To compute Ax, first change to the $\mathcal{B}$ coordinates, then multiply by B, then change back to the usual coordinates, where

$$\mathcal{B} = \left\{ \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \end{pmatrix} \right\} = \{v_1, v_2\} \qquad \text{(the columns of } C\text{)}.$$



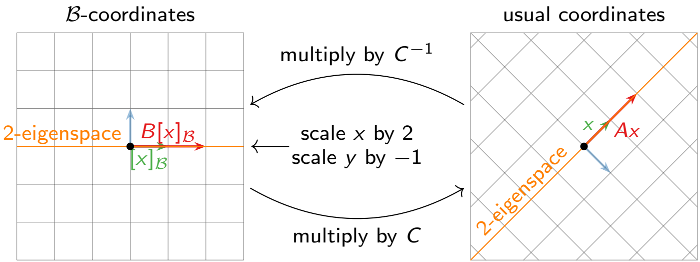
$$A = \begin{pmatrix} 1/2 & 3/2 \\ 3/2 & 1/2 \end{pmatrix} \quad B = \begin{pmatrix} 2 & 0 \\ 0 & -1 \end{pmatrix} \quad C = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \qquad A = CBC^{-1}.$$

What does B do geometrically?

It scales the x-direction by 2 and the y-direction by -1.

To compute Ax, first change to the $\mathcal{B}$ coordinates, then multiply by B, then change back to the usual coordinates, where

$$\mathcal{B} = \left\{ \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \end{pmatrix} \right\} = \{v_1, v_2\} \qquad \text{(the columns of } C\text{)}.$$



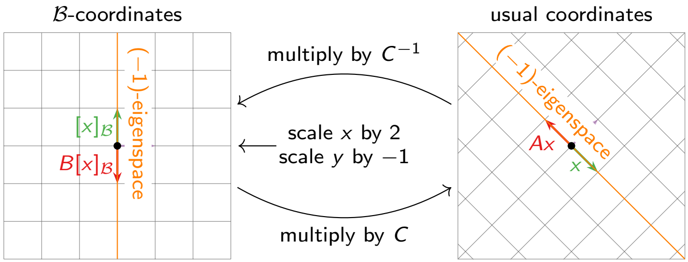
$$A = \begin{pmatrix} 1/2 & 3/2 \\ 3/2 & 1/2 \end{pmatrix} \quad B = \begin{pmatrix} 2 & 0 \\ 0 & -1 \end{pmatrix} \quad C = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \qquad A = CBC^{-1}.$$

What does B do geometrically?

It scales the x-direction by 2 and the y-direction by -1.

To compute Ax, first change to the $\mathcal{B}$ coordinates, then multiply by B, then change back to the usual coordinates, where

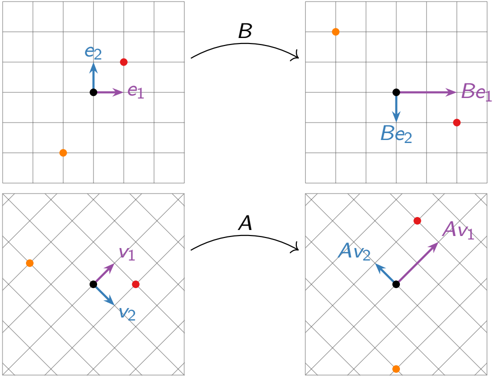
$$\mathcal{B} = \left\{ \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \end{pmatrix} \right\} = \{v_1, v_2\} \qquad \text{(the columns of } C\text{).}$$



What does A do geometrically?

- B scales the e_1-direction by 2 and the e_2-direction by -1.
- A scales the v_1-direction by 2 and the v_2-direction by -1.

columns of C



[interactive]

Since B is simpler than A, this makes it easier to understand A.

Note the relationship between the eigenvalues/eigenvectors of A and B.

Similarity

Example (3 × 3)

$$A = \begin{pmatrix} -3 & -5 & -3 \\ 2 & 4 & 3 \\ -3 & -5 & -2 \end{pmatrix} \quad B = \begin{pmatrix} 2 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad C = \begin{pmatrix} -1 & 1 & 0 \\ 1 & -1 & 1 \\ -1 & 0 & 1 \end{pmatrix}$$

$$\implies \quad A = CBC^{-1}.$$

What do A and B do geometrically?

- B scales the e_1-direction by 2, the e_2-direction by -1, and fixes e_3.
- A scales the v_1-direction by 2, the v_2-direction by -1, and fixes v_3.

Here v_1, v_2, v_3 are the columns of C.

[interactive]

Similar Matrices Have the Same Characteristic Polynomial

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

Why? Suppose $A = CBC^{-1}$.

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

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

Consequence: similar matrices have the same eigenvalues! (But different eigenvectors in general.)

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 has nothing to do with 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.

Summary

We did two different things today.

First we talked about characteristic polynomials:

- We learned to find the eigenvalues of a matrix by computing the roots of the characteristic polynomial $p(\lambda) = \det\bigl(A - \lambda I\bigr)$.
- For a 2×2 matrix A, the characteristic polynomial is just

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

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

Then we talked about similar matrices:

- Two square matrices A, B of the same size are **similar** if there is an invertible matrix C such that $A = CBC^{-1}$.
- Geometrically, similar matrices A and B do the same thing, except B operates on the coordinate system $\mathcal{B}$ defined by the columns of C:

$$B[x]_{\mathcal{B}} = [Ax]_{\mathcal{B}}.$$

- This is useful when we can find a similar matrix B which is *simpler* than A (e.g., a diagonal matrix).