Announcements

October 17

- WeBWorK assignments 3.1 and 3.2 are due Friday at 6am.
- Quiz on Friday: 3.1 and 3.2.
- Midterm 2 will take place in recitation on Friday, 10/28.
  - This is the day before the withdrawal deadline.
- Office hours: Wednesday 1–2pm, Thursday 3:30–4:30pm, and by appointment, in Skiles 221.
  - As always, TAs' office hours are posted on the website.
  - Math Lab is also a good place to visit.

Section 3.2

Properties of Determinants

Plan for Today

Last time, we gave a recursive formula for determinants in terms of cofactor expansions.

Plan for today:

- An abstract definition of the determinant in terms of its properties.
- Computing determinants using row operations.
- Determinants and products: $\det(AB) = \det(A)\det(B)$.
- Determinants and volumes.
- Determinants and linear transformations.

The determinant is one of the most amazing functions ever devised. Today is about beginning to understand why.

The Determinant is a Function

We can think of the determinant as a function of the entries of a matrix:

$$\det \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} = \begin{aligned} &a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} \\ &\quad - a_{13}a_{22}a_{31} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33}. \end{aligned}$$

The formula for the determinant of an $n \times n$ matrix has $n!$ terms. So the determinant of a 10×10 matrix has 3,628,800 terms!

When mathematicians encounter a function whose formula is too difficult to write down, we try to *characterize* it in terms of its properties.

The determinant function is characterized by how it is changed by row operations.

Defining the Determinant in Terms of its Properties

Definition

The **determinant** is a function

$$\det\colon \{\text{square matrices}\} \longrightarrow \mathbf{R}$$

with the following **defining properties**:

1. $\det(I_n) = 1$
2. If we do a row replacement on a matrix (add a multiple of one row to another), the determinant does not change.
3. If we swap two rows of a matrix, the determinant scales by -1.
4. If we scale a row of a matrix by k, the determinant scales by k.

Why would we think of these properties? This is how volumes work!

1. The volume of the unit cube is 1.
2. Volumes don't change under a shear.
3. Volume of a mirror image is negative of the volume?
4. If you scale one coordinate by k, the volume is multiplied by k.

Properties of the Determinant

2×2 matrix

$$\det\begin{pmatrix} 1 & -2 \\ 0 & 3 \end{pmatrix} = 3$$

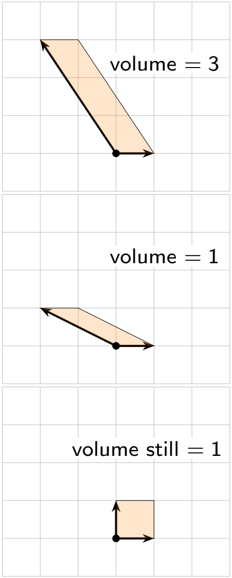
Scale: $R_2 = \frac{1}{3}R_2$

$$\det\begin{pmatrix} 1 & -2 \\ 0 & 1 \end{pmatrix} = 1$$

Row replacement: $R_1 = R_1 + 2R_2$

$$\det\begin{pmatrix} 1 & -2 \\ 0 & 1 \end{pmatrix} = 1$$

(This is a shear by the elementary matrix $\left[\begin{smallmatrix} 1 & 2 \\ 0 & 1 \end{smallmatrix}\right]$.)



Since an elementary matrix differs from the identity matrix by one row operation, and since $\det(I_n) = 1$, it is easy to calulate the determinant of an elementary matrix:

$$\det \begin{pmatrix} 1 & 0 & 8 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \det(I_n) = 1 \qquad \text{(properties 1 and 2)}$$

$$\det \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix} = -\det(I_n) = -1 \qquad \text{(properties 1 and 3)}$$

$$\det \begin{pmatrix} 1 & 0 & 0 \\ 0 & 17 & 0 \\ 0 & 0 & 1 \end{pmatrix} = 17\det(I_n) = 17 \qquad \text{(properties 1 and 4)}$$

Computing the Determinant by Row Reduction

We can use the properties of the determinant and row reduction to compute the determinant of any matrix! This means that det is completely characterized by its defining properties.

$$
\begin{aligned}
\det\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 5 & 7 & -4 \end{pmatrix} &= -\det\begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 5 & 7 & -4 \end{pmatrix} && \text{(property 3)} \\
&= -\det\begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 7 & -9 \end{pmatrix} && \text{(property 2)} \\
&= -\det\begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & -9 \end{pmatrix} && \text{(property 2)} \\
&= (-1)\cdot(-9)\det\begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} && \text{(property 4)} \\
&= (-1)\cdot(-9)\det\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} && \text{(property 2)} \\
&= 9 && \text{(property 1)}
\end{aligned}
$$

Computing the Determinant by Row Reduction

Saving some work

The determinant of an upper (or lower) triangular matrix is the product of the diagonal entries, so we can stop row reducing when we get to row echelon form.

$$\det\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 5 & 7 & -4 \end{pmatrix} = \cdots = -\det\begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & -9 \end{pmatrix} = 9.$$

This is almost always the easiest way to compute the determinant of a large, complicated matrix, either by hand or by computer. (Cofactor expansion is $O(n!) \sim O(n^n\sqrt{n})$, row reduction is $O(n^3)$.)

Poll

Suppose that A is a 4×4 matrix satisfying

$$Ae_1 = e_2 \quad Ae_2 = e_3 \quad Ae_3 = e_4 \quad Ae_4 = e_1.$$

What is $\det(A)$?

A. -1 B. 0 C. 1

These equations tell us the columns of A:

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

You need 3 row swaps to transform this to the identity matrix. So $\det(A) = (-1)^3 = -1$.

A Mathematical IOU

The characterization of the determinant function in terms of its properties is very useful. It gives us a fast way to compute determinants, and prove other properties (later). But...

The disadvantage of defining a function by its properties instead of a formula is: how do you know such a function exists? and if it exists, why is there only one function satisfying those properties?

In our case, we can compute the determinant of a matrix from its defining properties, so if it exists, it is unique. But how do we know that two different row reductions won't give two different answers for the determinant?

Here is a summary of the magical properties of the determinant. Prof. Margalit's notes (on the website) have very understandable proofs.

Magical Properties of the Determinant

you really have to know these

1. There is one and only one function $\det\colon \{\text{square matrices}\} \to \mathbf{R}$ satisfying the defining properties (1)–(4).
2. A is invertible if and only if $\det(A) \neq 0$.
3. If we row reduce A without row scaling, then
$$\det(A) = (-1)^{\#\text{swaps}}\big(\text{product of diagonal entries in REF}\big)$$
4. The determinant can be computed using any of the $2n$ cofactor expansions.
5. $\det(AB) = \det(A)\det(B)$
6. $\det(A) = \det(A^T)$
7. $|\det(A)|$ is the volume of the parallelepiped defined by the columns of A.
8. If A is an $n \times n$ matrix with transformation $T(x) = Ax$, and S is a subset of $\mathbf{R}^n$, then the volume of $T(S)$ is $|\det(A)|$ times the volume of S. (Even for curvy shapes S.)
9. The determinant is multi-linear (we'll talk about this in a few slides).

Property 1 (existence of a function det satisfying the defining properties (1)–(4)) is the hardest one to prove.

We've already discussed property 3: recall that row replacement doesn't change the determinant, and a row swap changes the determinant by -1.

Property 4: One has to show that any cofactor expansion *also* satisfies the defining properties of the determinant (1)–(4). See the notes.

Property 2: An invertible matrix A row reduces to the identity matrix. Since $\det(I_n) \neq 0$, this means $\det(A) \neq 0$. A non-invertible matrix A row reduces to a matrix with a zero row. We know such a matrix has zero determinant by cofactor expansion.

Property 6 is proved using induction and cofactor expansions. It implies that determinants scale the same way under *column* operations as row operations.

Property 7: First you define a *signed* volume vol (i.e. a way to decide whether the volume of a parallelepiped is negative). Then it's easy to show vol satisfies the defining properties (1)–(4), so $\mathrm{vol} = \det$ by Property 1.

Multiplicativity of the Determinant

Why is Property 5 true? In Lay, there's a proof using elementary matrices. Here's a better one.

Let B be an $n \times n$ matrix. There are two cases:

1. If $\det(B) = 0$, then B is not inverible. So for any matrix A, AB is not invertible. (Otherwise $B^{-1} = (AB)^{-1}A$.) So

$$\det(AB) = 0 = \det(A) \cdot 0 = \det(A)\det(B).$$

2. If B is invertible, define another function

$$f\colon \{n \times n \text{ matrices}\} \longrightarrow \mathbf{R} \quad \text{by} \quad f(A) = \frac{\det(AB)}{\det(B)}.$$

   Let's check the defining properties:

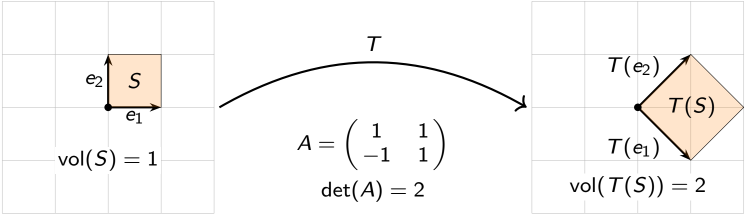
   1. $f(I_n) = \det(I_n B)/\det(B) = 1$.

   2–4. Doing a row operation on A and then multiplying by B, does the *same row operation* on AB. This is because a row operation is left-multiplication by an elementary matrix E, and $(EA)B = E(AB)$. Hence f scales like det with respect to row operations.

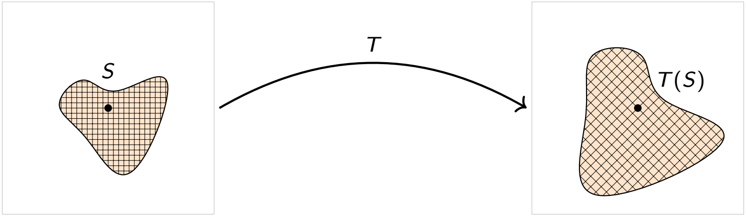
   By uniqueness, $f = \det$, i.e.,

$$\det(A) = f(A) = \frac{\det(AB)}{\det(B)} \quad \text{so} \quad \det(A)\det(B) = \det(AB).$$

Why is Property 8 true? For instance, if S is the unit cube, then $T(S)$ is the parallelepiped defined by the columns of A, since the columns of A are $T(e_1), T(e_2), \ldots, T(e_n)$. In this case, Property 8 is the same as Property 7.



For curvy shapes, you break S up into a bunch of tiny cubes. Each one is scaled by $|\det(A)|$; then you use *calculus* to reduce to the previous situation!



Multi-Linearity of the Determinant

We can also think of det as a function of the columns (or the rows) of an $n \times n$ matrix:

$$\det\colon \underbrace{\mathbf{R}^n \times \mathbf{R}^n \times \cdots \times \mathbf{R}^n}_{n \text{ times}} \longrightarrow \mathbf{R}$$

$$\det(v_1, v_2, \ldots, v_n) = \det\begin{pmatrix} | & | & & | \\ v_1 & v_2 & \cdots & v_n \\ | & | & & | \end{pmatrix}.$$

Property 9 says that for any i and any vectors $v_1, v_2, \ldots, v_n$ and v_i' and any scalar c,

$$\begin{aligned} \det(v_1, \ldots, v_i + v_i', \ldots, v_n) &= \det(v_1, \ldots, v_i, \ldots, v_n) + \det(v_1, \ldots, v_i', \ldots, v_n) \\ \det(v_1, \ldots, cv_i, \ldots, v_n) &= c\det(v_1, \ldots, v_i, \ldots, v_n). \end{aligned}$$

In other words, scaling one column (or row) by c scales det by c (which we already knew), and if column i is a sum of two vectors v_i, v_i', then the determinant is the sum of two determinants, one with v_i in column i, and one with v_i' in column i. *This only works one column at a time.*

Proof: just expand cofactors along column i.