Orientation

The story so far:

- Solve $Ax=b$ ✓
  (elimination, LU, PA=LU, ...)
- Approximately solve $Ax=b$ ✓
  (subspaces, orthogonality, projections, QR, ...)

Next up:

- Solve $Ax=\lambda x$

This is the eigenvalue problem used to solve difference equations (rabbit population, PageRanks, ...). It will also play a central role in the SVD.

In $Ax=\lambda x$, the "λ" is a scalar. Hence $Ax=\lambda x$ only makes sense when A is square. We will be dealing exclusively with square matrices for the next 5 weeks.

The determinant of a square matrix is a number with many magical properties. Importantly: it gives you a formula that tells you whether a matrix is invertible.

Today we'll do determinants using row operations.
Next time: other ways to compute them.

Determinants

Def: The determinant of a square matrix A is a number $\det(A)$ satisfying:

(1) If $A \xrightarrow{R_i += c \cdot R_j} B$ then $\det(B) = \det(A)$

(2) If $A \xrightarrow{R_i \times = c} B$ then $\det(B) = c \cdot \det(A)$

(3) If $A \xrightarrow{R_i \leftrightarrow R_j} B$ then $\det(B) = -\det(A)$

(4) $\det(I_n) = 1$.

Consequence: If A has a zero row then $\det(A) = 0$

Eg:

$$\det\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 0 & 0 & 0 \end{pmatrix} \xrightarrow[(2)]{R_3 \times = -1} -\det\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ -0 & -0 & -0 \end{pmatrix}$$

$$= -\det\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 0 & 0 & 0 \end{pmatrix}$$

This says $\det(A) = -\det(A)$, so $\det(A) = 0$.

In fact, we'll see that if the rows are LD then $\det = 0$.

Consequence: If A is (upper/lower) triangular then $\det(A) =$ product of diagonal entries.

$$\boxed{\det\left(\begin{matrix}\text{triangular} \\ \text{matrix}\end{matrix}\right) = \begin{matrix}\text{product of the} \\ \text{diagonal entries}\end{matrix}}$$

eg. a matrix in REF

Eg: First note:

$$A = \begin{pmatrix} a & * & * \\ 0 & b & * \\ 0 & 0 & c \end{pmatrix} \overset{R_1 x = \frac{1}{a}}{\rightsquigarrow} \begin{pmatrix} 1 & * & * \\ 0 & b & * \\ 0 & 0 & c \end{pmatrix} = B$$

$$\overset{(2)}{\Longrightarrow} \det(B) = \frac{1}{a}\det(A)$$

$$\Longrightarrow \det(A) = a \cdot \det(B)$$

i.e. $\det\begin{pmatrix} a & * & * \\ 0 & b & * \\ 0 & 0 & c \end{pmatrix} = a \cdot \det\begin{pmatrix} 1 & * & * \\ 0 & b & * \\ 0 & 0 & c \end{pmatrix}$

If $abc \neq 0$, we can do this 3 times:

$$\det\begin{pmatrix} a & * & * \\ 0 & b & * \\ 0 & 0 & c \end{pmatrix} \underset{(2)\ R_3 x = \frac{1}{c}}{\overset{R_1 x = \frac{1}{a}\ R_2 x = \frac{1}{b}}{=\!=\!=}} abc \det\begin{pmatrix} 1 & * & * \\ 0 & 1 & * \\ 0 & 0 & 1 \end{pmatrix}$$

$$\underset{\text{replacements}}{\overset{\text{row (1)}}{=\!=\!=}} abc \det\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \overset{(4)}{=\!=} abc$$

What if $abc = 0$ though? For example, if $b = 0, c \neq 0$ then

$$\det\begin{pmatrix} a & * & * \\ 0 & 0 & * \\ 0 & 0 & c \end{pmatrix} \underset{\text{replacements}}{\overset{\text{row (1)}}{=\!=\!=}} \det\begin{pmatrix} a & * & 0 \\ 0 & 0 & 0 \\ 0 & 0 & c \end{pmatrix} = 0 = abc$$

Something like this will happen every time if there's a zero on the diagonal.

Since a REF matrix is upper-triangular, you can compute $\det(A)$ using Gaussian elimination!

Eg: $\det\begin{pmatrix}0&1&3\\1&2&1\\1&1&0\end{pmatrix} \xrightarrow[(3)]{R_1\leftrightarrow R_2} -\det\begin{pmatrix}1&2&1\\0&1&3\\1&1&0\end{pmatrix}$

$\xrightarrow[(1)]{R_3 -= R_1} -\det\begin{pmatrix}1&2&1\\0&1&3\\0&-1&-1\end{pmatrix} \xrightarrow[(1)]{R_3 += R_2} -\det\begin{pmatrix}1&2&1\\0&1&3\\0&0&2\end{pmatrix}$

$= -2$

triangular

Magically, you get the same number for det(A) no matter which row operations you do!

Eg: $\det\begin{pmatrix}0&1&3\\1&2&1\\1&1&0\end{pmatrix} \xrightarrow[(3)]{R_1\leftrightarrow R_3} -\det\begin{pmatrix}1&1&0\\1&2&1\\0&1&3\end{pmatrix}$

different REF

$\xrightarrow[(1)]{R_2 -= R_1} -\det\begin{pmatrix}1&1&0\\0&1&1\\0&1&3\end{pmatrix} \xrightarrow[(1)]{R_3 -= R_2} -\det\begin{pmatrix}1&1&0\\0&1&1\\0&0&2\end{pmatrix}$

$= -2$ same det

Gaussian elimination is the fastest general algorithm for computing det(A) (if A has "random", known entries).

Procedure: To compute det(A), run elimination:

$A \underset{\text{ops}}{\overset{\text{row}}{\rightsquigarrow}} U = \text{REF}$. Then

$$\boxed{\det(A) = (-1)^{\#\text{row swaps}} \cdot \frac{1}{\prod(\text{row scaling})} \prod\left(\begin{matrix}\text{diagonal}\\ \text{entries of } U\end{matrix}\right)}$$

"product"

NB: You don't need to do row scaling operations in Gaussian elimination, in which case the (row scaling) term doesn't appear.

We'll have a much better way to remember this formula using PALU decompositions a little later on.

NB: Row operations multiply the determinant by a nonzero scalar:

$$A \underset{\text{ops}}{\overset{\text{row}}{\rightsquigarrow}} B \implies \det(B) = \begin{pmatrix}\text{nonzero} \\ \text{number}\end{pmatrix} \cdot \det(A).$$

We can compute the determinant of an arbitrary 2×2 matrix. (Next time: 3×3)

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

→ If $a \neq 0$:

$$\det\begin{pmatrix} a & b \\ c & d \end{pmatrix} \overset{R_2 -= \frac{c}{a}R_1}{=\!=\!=} \det\begin{pmatrix} a & b \\ 0 & d - \frac{c}{a}\cdot b \end{pmatrix} = a\left(d - \frac{c}{a}b\right)$$

$$= ad - bc$$

→ If $a = 0$:

$$\det\begin{pmatrix} 0 & b \\ c & d \end{pmatrix} \overset{R_1 \leftrightarrow R_2}{=\!=\!=} -\det\begin{pmatrix} c & d \\ 0 & b \end{pmatrix} = -bc$$

$$= ad - bc$$

In both cases: $\boxed{\det\begin{pmatrix} a & b \\ c & d \end{pmatrix} = ad - bc}$

Determinants are so much fun because of the

Magical Properties of the Determinant:

(1) Existence: There exists a number $\det(A)$ satisfying (1)–(4).

(2) Invertibility: $\boxed{A \text{ is invertible} \iff \det(A) \neq 0}$

(3) Multiplicativity:

$$\det(AB) = \det(A)\det(B)$$

and $\det(A^{-1}) = \dfrac{1}{\det(A)}$ if $\det(A) \neq 0$

(4) Transposes: $\det(A^T) = \det(A)$

We'll prove (2) in class. See ILA for the rest.

What do these properties mean?

(1) Existence: This says you get the same number no matter which row ops you do!

This is actually the only magic property that is difficult to prove.

(2) Invertibility:

Proof: Suppose $A \underset{\text{ops}}{\overset{\text{row}}{\rightsquigarrow}} U = \text{REF}$.

Then $\det(U) = \begin{pmatrix}\text{nonzero}\\ \text{number}\end{pmatrix} \cdot \det(A)$.

But $\det(U)$ = product of diagonal entries, so

$\det(A) \neq 0 \Leftrightarrow$ $\det(U) \neq 0 \Leftrightarrow$ all diagonal entries are nonzero

$\Leftrightarrow$ A has n pivots

$\Leftrightarrow$ A is invertible //

Eg: Is $\begin{pmatrix} 1 & 4 & 7 \\ 2 & 5 & 8 \\ 3 & 6 & 9 \end{pmatrix}$ invertible?

$$\det\begin{pmatrix} 1 & 4 & 7 \\ 2 & 5 & 8 \\ 3 & 6 & 9 \end{pmatrix} \underset{R_3 -= 3R_1}{\overset{R_2 -= 2R_1}{=\!=}} \det\begin{pmatrix} 1 & 4 & 7 \\ 0 & -3 & -6 \\ 0 & -6 & -12 \end{pmatrix}$$

$$\overset{R_3 -= 2R_2}{=\!=} \det\begin{pmatrix} 1 & 4 & 7 \\ 0 & -3 & -6 \\ 0 & 0 & 0 \end{pmatrix} = 0$$ so not invertible.

NB: If the columns (rows) of A are linearly dependent then A does not have FCR (FRR) $\Rightarrow$ not invertible $\Rightarrow \det(A) = 0$.

$$\boxed{\text{The rows or cols of } A \text{ are LD} \Leftrightarrow \det(A) = 0}$$

(3) Multiplicativity:

Eg: $\left[\det\begin{pmatrix} 0 & 1 & 3 \\ 1 & 2 & 1 \\ 1 & 1 & 0 \end{pmatrix}^{100}\right] = \det\left[\begin{pmatrix} 0 & 1 & 3 \\ 1 & 2 & 1 \\ 1 & 1 & 0 \end{pmatrix}^{99}\begin{pmatrix} 0 & 1 & 3 \\ 1 & 2 & 1 \\ 1 & 1 & 0 \end{pmatrix}\right]$

$$= \det\begin{pmatrix} 0 & 1 & 3 \\ 1 & 2 & 1 \\ 1 & 1 & 0 \end{pmatrix} \cdot \det\left[\begin{pmatrix} 0 & 1 & 3 \\ 1 & 2 & 1 \\ 1 & 1 & 0 \end{pmatrix}^{99}\right] = \cdots = \left[\det\begin{pmatrix} 0 & 1 & 3 \\ 1 & 2 & 1 \\ 1 & 1 & 0 \end{pmatrix}\right]^{100}$$

$$= (-2)^{100}$$

This is way faster than multiplying a 3×3 matrix by itself 100 times! (Constant time vs. $\geq O(n^3)$)

More generally: $\boxed{\det(A^n) = \det(A)^n \quad \text{for all } n \geq 0 \text{ (and } n<0 \text{ if } \det(A)\neq 0)}$

Eg: Suppose A has a $PA=LU$ decomposition.

$$\det(L) = \det\begin{pmatrix} 1 & 0 & 0 \\ * & 1 & 0 \\ * & * & 1 \end{pmatrix} = 1$$

You form P by doing row swaps on I_n. Each row swap negates the determinant, so

$$\det(P) = (-1)^{\#\text{row swaps}}$$

Taking det of both sides of $PA=LU$ gives:

$$(-1)^{\#\text{row swaps}}\det(A) = \det(PA)$$
$$= \det(LU) = \det(U)$$

This recovers the formula from before (we did no row scaling operations to compute $PA=LU$).

(4) Transposes:

One fun consequence is that det satisfies properties (1)–(3) (of the definition of det) for column operations too: they're just row operations on A^T.

Eg:

$$\begin{tikzcd}
\det\begin{pmatrix} 2 & 7 & 4 \\ 3 & 1 & 3 \\ 4 & 0 & 1 \end{pmatrix} \arrow[r, equal, "C_1 -= 4C_3"] \arrow[d, equal, "\text{transpose}"] & \det\begin{pmatrix} -14 & 7 & 4 \\ -9 & 1 & 3 \\ 0 & 0 & 1 \end{pmatrix} \arrow[d, equal, "\text{transpose}"] \\
\det\begin{pmatrix} 2 & 3 & 4 \\ 7 & 1 & 0 \\ 4 & 3 & 1 \end{pmatrix} \arrow[r, equal, "R_1 -= 4R_3"] & \det\begin{pmatrix} -14 & -9 & 0 \\ 7 & 1 & 0 \\ 4 & 3 & 1 \end{pmatrix}
\end{tikzcd}$$

So we can compute det using column operations:

$$\det\begin{pmatrix} 2 & 7 & 4 \\ 3 & 1 & 3 \\ 4 & 0 & 1 \end{pmatrix} \xlongequal{C_1 -= 4C_3} \det\begin{pmatrix} -14 & 7 & 4 \\ -9 & 1 & 3 \\ 0 & 0 & 1 \end{pmatrix}$$

$$\xlongequal{C_1 += 9C_2} \det\begin{pmatrix} 49 & 7 & 4 \\ 0 & 1 & 3 \\ 0 & 0 & 1 \end{pmatrix} = 49$$

I don't know how useful this is, but it's fun!

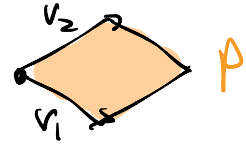
NB: Given (1) Existence, it's straightforward to prove (3) Multiplicativity and (4) Transposes using elementary matrices.

Determinants and Volumes

This is a geometric interpretation of determinants that helps to explain the behavior of det with respect to row operations.

It also explains why determinants come up in multivariable calculus (see below).

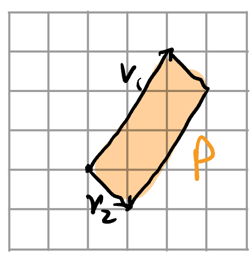
Two vectors $v_1, v_2 \in \mathbb{R}^2$ determine ("span") a parallelogram:



$$P = \{x_1 v_1 + x_2 v_2 : x_1, x_2 \in [0,1]\}$$

Fact: $\operatorname{area}(P) = \left|\det\begin{pmatrix} - & v_1^T & - \\ - & v_2^T & - \end{pmatrix}\right| = \left|\det\begin{pmatrix} | & | \\ v_1 & v_2 \\ | & | \end{pmatrix}\right|$

Eg: $v_1 = \begin{pmatrix} 2 \\ 3 \end{pmatrix}$ $\quad v_2 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$



$$\operatorname{area}(P) = \left|\det\begin{pmatrix} 2 & 3 \\ 1 & -1 \end{pmatrix}\right|$$

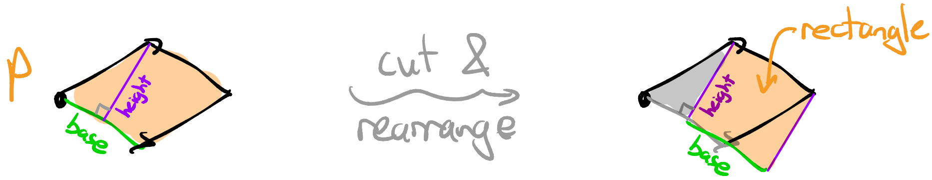
$$= |2(-1) - 3(1)| = 5$$

Why is the Fact true?

Let's check that you can compute areas using row operations in the same way as determinants.

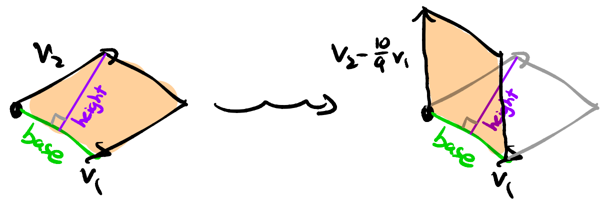
(Same computation $\Rightarrow$ same number!)

NB: area(P) = base × height



(1) Row replacement:

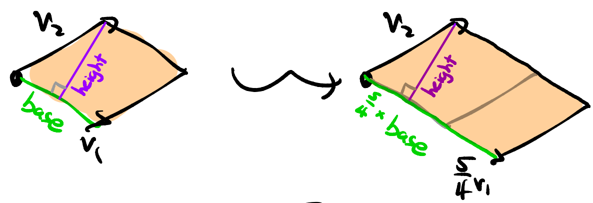
$v_2 \rightsquigarrow v_2 - \frac{10}{9} v_1$



base: unchanged
height: unchanged
$\Longrightarrow$ area: unchanged
and |det|: unchanged ✓

(2) Row Scaling:

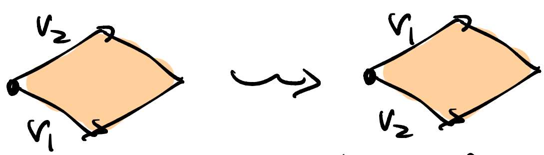
$v_1 \rightsquigarrow \frac{5}{4} v_1$



base $\rightsquigarrow \frac{5}{4}$ × base
height: unchanged
$\Longrightarrow$ area $\rightsquigarrow \frac{5}{4}$ × area
and |det| $\rightsquigarrow \frac{5}{4}$ × |det| ✓

(3) Row Swap:

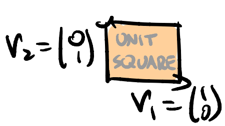
$v_1 \longleftrightarrow v_2$



area: unchanged
and |det|: unchanged ✓

Q: where is the −1?
A: on the HW

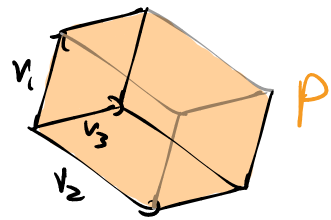
(4) Identity:



|det| = 1
area = 1 ✓

The same fact holds true in higher dimensions, with the same reasoning.

Def: The *parallelepiped* determined ("spanned") by n vectors in $\mathbb{R}^n$



$$v_1, v_2, \ldots, v_n \in \mathbb{R}^n \text{ is}$$

$$P = \{ x_1 v_1 + x_2 v_2 + \cdots + x_n v_n : x_1, x_2, \ldots, x_n \in \mathbb{R} \}$$

Thm (Determinants and Volumes)

$$\text{volume}(P) = \left| \det \begin{pmatrix} - v_1^T - \\ \vdots \\ - v_n^T - \end{pmatrix} \right| = \left| \det \begin{pmatrix} | & & | \\ v_1 & \cdots & v_n \\ | & & | \end{pmatrix} \right|$$

NB: When $n=1$, "volume" = "length"

$$\text{length}(\overset{0}{\bullet}\!\!-\!\!-\!\!-\!\!\overset{a}{\bullet}) = |a| = |\det(a)|$$

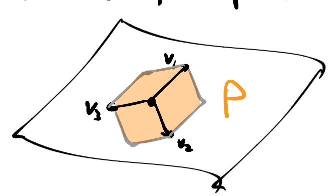
NB: When $n=2$, "volume" = "area"

Q: When is $\text{volume}(P) = 0$?

When P is squashed flat.

This means $\{v_1, \ldots, v_n\}$ are LD, so $\det = 0$ ✓

$\{v_1, v_2, v_3\}$ coplanar



$\Rightarrow$ P is flat!

Multivariable Calculus

To do integration, you approximate shapes by tiny cubes, which turn into tiny parallellepipeds after applying a function. This is why determinants appear in the change of variables formula for integration.

$$(y_1, y_2) = f(x_1, x_2)$$

$$\Rightarrow dy_1 dy_2 = \det\begin{pmatrix} \partial y_1/\partial x_1 & \partial y_1/\partial x_2 \\ \partial y_2/\partial x_1 & \partial y_2/\partial x_2 \end{pmatrix} dx_1 dx_2$$