Determinants continued

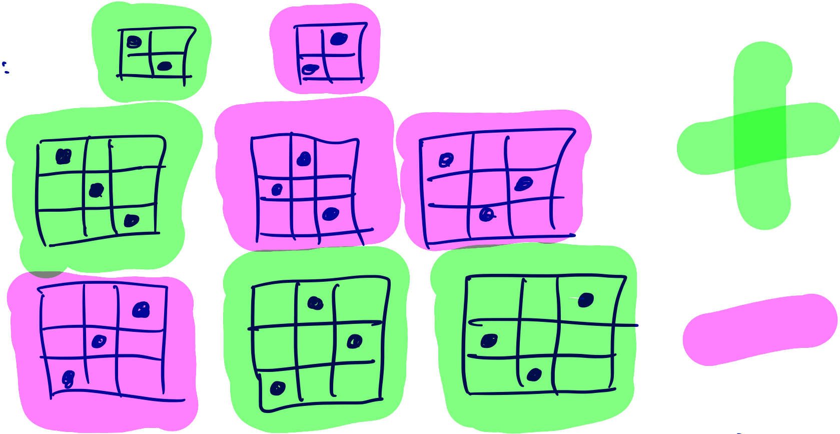
9 Mar

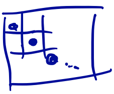
Here's yet another perspective on determinants.

A <u>rook arrangement</u> (my term) is a positioning of n rooks on an $n \times n$ grid so that no pair threaten each other:

2x2:

3x3:



The <u>sign</u> of a rook arrangement is (-1) to the number of column flips ←i.e., switch two columns you need to get there from  ←(rooks on main diagonal). Then given A, we make a term for every rook arrangement: multiply the sign of the rook arrangement by the product of all the matrix entries in the rooks' positions.

Theorem Adding up the aforementioned terms gives det A. [Traditional expression, if you prefer: $\det A = \sum_{\sigma \text{ a permutation of } \{1,\dots,n\}} \operatorname{sgn}(\sigma) a_{1,\sigma(1)} \cdots a_{n,\sigma(n)}$]

What's the point? This is our first concrete and computational definition. The geometric one doesn't give you a way to calculate det A for $n > 2$, and the cofactor expansion is leaning on its uniqueness (that is, doesn't matter what row/column you expand along). In fact, to show uniqueness of the cofactor expansion, one shows by induction that all of give the above rook arrangement expression.

Example Find $\begin{vmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 10 \end{vmatrix}$ using rook arrangements.

Sol. $\det A = (1)(5)(10) - (2)(4)(10) + (3)(4)(8)$
$- (1)(6)(8) + (7)(5)(3) - (7)(2)(6)$
$= -3$

Example Show that when $\begin{vmatrix} a & b & c & d \\ e & f & g & h \\ i & j & k & l \\ m & n & o & p \end{vmatrix}$ is written out and simplified, there are no occurrences of a^2.

Solution Each rook arrangement includes at most one factor of a. ■

Exercise Show from the rook def$^{\underline{n}}$ that $\det A = \det A^T$.

One feature of determinants that is geometrically clear but bears emphasizing is that

$$\boxed{\det A \neq 0 \iff A \text{ is invertible}}$$

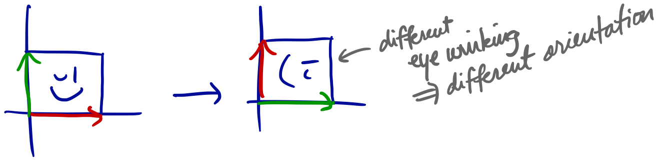
We use the term "singular" to describe an $n \times n$ matrix A with $\det A = 0$.

We still don't have a fast way of calculating $\det A$! The rook def$^{\underline{n}}$ is slow because it has $n!$ terms

Determinants and elementary row operations

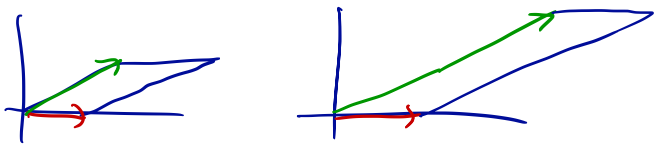
* <u>Row switching</u> flips the orientation and thus multiplies det by -1.

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \longrightarrow \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$



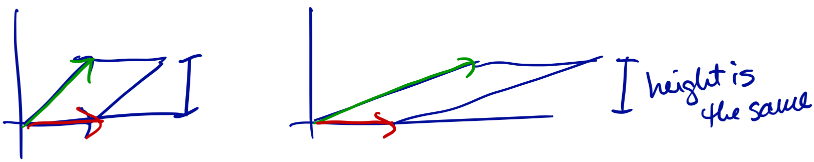
* <u>Row scaling</u> Scales the det:

$$\begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix} \longrightarrow \begin{bmatrix} 2 & 4 \\ 0 & 1 \end{bmatrix}$$



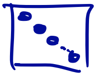
* <u>Row add</u> Doesn't alter det!

$$\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \longrightarrow \begin{bmatrix} 1 & 3 \\ 0 & 1 \end{bmatrix}$$



So we can row reduce A to a diagonal matrix, avoiding row scaling operations, and use the following fact.

<u>Proposition</u> If A is upper triangular (all zeros below the main diagonal), then $\det A$ equals the product of the diagonal entries.

<u>Proof</u> The only rook arrangement without a zero factor is . So that's the only nonzero term in the rook expansion of $\det A$. ∎

Example Find $\begin{vmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 10 \end{vmatrix}$ by row operations.

<u>Solution</u>

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 10 \end{pmatrix} \sim \begin{pmatrix} 1 & 2 & 3 \\ 0 & -3 & -6 \\ 0 & -6 & -11 \end{pmatrix} \sim \begin{pmatrix} 1 & 2 & 3 \\ 0 & -3 & -6 \\ 0 & 0 & 1 \end{pmatrix}$$

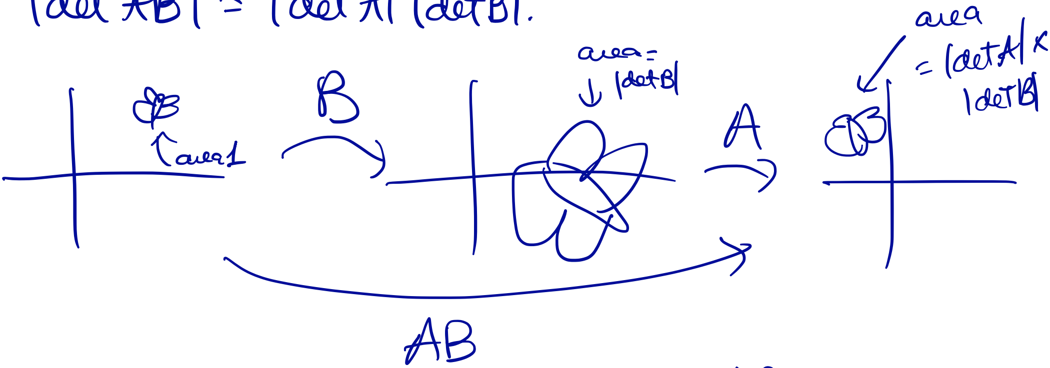
So $\det A = (1)(-3)(1) = -3.$

(Note: if we'd row switched k times, we'd have needed an extra factor of $(-1)^k$ to compensate)

Here are a couple properties of the det which are clear from the geometric definition.

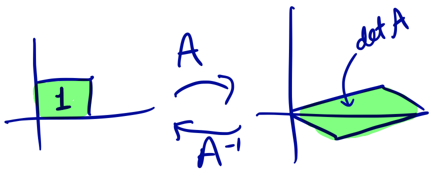
Fact $\det(AB) = (\det A)(\det B)$

Proof If a shape has volume 1, then its image under B has area $|\det B|$ and the image of that has area $|\det B|\,|\det A|$. So $|\det AB| = |\det A|\,|\det B|$.



But AB reverses orientations iff exactly one of A or B does, so the signs of $\det A \det B$ and $\det AB$ match too.

Fact $\det(A^{-1}) = (\det A)^{-1}$



Cramer's Rule

Theorem If A is an invertible $n \times n$ matrix, then the unique solution $\vec{x}$ of $A\vec{x} = \vec{b}$ is given by $\vec{x} = (x_1, \ldots, x_n)$ where for all i from 1 to n,

$$x_i = \frac{\det \overbrace{[\vec{a}_1 \cdots \vec{a}_{i-1}\ \vec{b}\ \vec{a}_{i+1} \cdots \vec{a}_n]}^{\text{subout the } i^{\text{th}} \text{ column of } A \text{ with } \vec{b}}}{\det A}.$$

Example Solve $\begin{cases} x+y = 20 \\ x-y = 8 \end{cases}$ using Cramer's Rule.

Solution

$$x = \frac{\begin{vmatrix} 20 & 1 \\ 8 & -1 \end{vmatrix}}{\begin{vmatrix} 1 & 1 \\ 1 & -1 \end{vmatrix}} = \frac{-28}{-2} = 14$$

$$y = \frac{\begin{vmatrix} 1 & 20 \\ 1 & 8 \end{vmatrix}}{\begin{vmatrix} 1 & 1 \\ 1 & -1 \end{vmatrix}} = \frac{-12}{-2} = 6.$$

So $(x,y) = \boxed{(14,6)}$.

Cramer's rule can be used to derive a formula for matrix inverses, by solving $A\vec{x} = \vec{e}_1$, $A\vec{x} = \vec{e}_2$, etc. This is called the 'adjugate' formula for A^{-1}; see the book if you're interested.

Proof of Cramer's rule

Suppose $A\vec{x} = \vec{b}$ and consider

$$M = \begin{bmatrix} 1 & 0 & & x_1 & & 0 \\ 0 & 1 & & \vdots & & 0 \\ 0 & 0 & \cdots & \vdots & \cdots & \vdots \\ \vdots & \vdots & & \vdots & & \vdots \\ 0 & \vdots & & x_n & & 1 \end{bmatrix}$$

sub out $\vec{x}$ for i^{th} column

Then $AM = [\vec{a}_1 \cdots \vec{b} \cdots \vec{a}_n]$. Also $\det M = x_i$, since the diagonal rook arrangement is the only one without zero factors. So

$$\det A \underbrace{\det M}_{x_i} = \det [\vec{a}_1 \cdots \vec{b} \cdots \vec{a}_n],$$

$$\& \; x_i = \frac{\det[\vec{a}_1 \cdots \vec{b} \cdots \vec{a}_n]}{\det A}.$$

Cramer's theorem is useful for expressing numerical stability:

Example Explain PSet #5, question 2 using Cramer's rule.

Solution We're interested in the solution of $\left[\begin{array}{cc|c} 4.5 & 3.1 & 19.249 \\ 1.6 & 1.1 & 6.843 \end{array}\right]$, which is (let's focus on x_1)

$$x_1 = \frac{\begin{vmatrix} 19.249 & 3.1 \\ 6.843 & 1.1 \end{vmatrix}}{\begin{vmatrix} 4.5 & 3.1 \\ 1.6 & 1.1 \end{vmatrix}} = \frac{-197/5000}{-1/100} = 3.94$$

The numerator here is very sensitive to small changes in $\vec{b} = \begin{bmatrix} 19.249 \\ 6.843 \end{bmatrix}$, because of the following crucial numerical analysis concept: if $d = A - B$ where A, B are much larger than d, then small percent changes in A or B can lead to huge percent changes in d:

$$62.469 - 62.467 = 0.002$$

tiny change

$$62.489 - 62.467 = 0.022$$

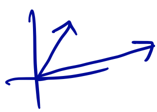
11x!!

the upshot: subtracting two very close numbers can lead to large percent errors in the difference.

If the coefficient matrix had been $\begin{bmatrix} 4.5 & 3.1 \\ 1.6 & 3.0 \end{bmatrix}$, we'd get

$$x_1 = \frac{\det \begin{bmatrix} 19.249 & 3.1 \\ 6.843 & 3.0 \end{bmatrix}}{\det \begin{bmatrix} 4.5 & 3.1 \\ 1.6 & 3.0 \end{bmatrix}} = \frac{36.5337\ldots}{8.54\ldots} = 4.27795\ldots$$

which is now much stabler; no tiny numbers.

Graphically, it's easier to use  than