18.06.23: Determinants & Permutations

Lecturer: Barwick

Friday, 8 April 2016

So here's an $n \times n$ matrix

$$A = \begin{pmatrix} \vec{v}_1 & \cdots & \vec{v}_n \end{pmatrix},$$

and we have this number

$$\det(A) = \det(\vec{v}_1, \ldots, \vec{v}_n) \in \mathbf{R}$$

that measures the *signed n-dimensional volume* of the parallelopiped spanned by $\vec{v}_1, \ldots, \vec{v}_n$.

Let's list the things we know about $\det(A)$.

(1) *Normalization.* The identity matrix has determinant 1:

$$\det(I) = \det(\hat{e}_1, \cdots, \hat{e}_n) = 1.$$

(2) *Multilinearity.* For any real numbers $r, s \in \mathbf{R}$,

$$\begin{aligned}\det(\vec{v}_1, \cdots, r\vec{x}_i + s\vec{y}_i, \cdots, \vec{v}_n) &= r\det(\vec{v}_1, \cdots, \vec{x}_i, \cdots, \vec{v}_n) \\ &\quad + s\det(\vec{v}_1, \cdots, \vec{y}_i, \cdots, \vec{v}_n).\end{aligned}$$

(3) *Alternation.* The determinant

$$\det(\vec{v}_1, \cdots, \cdots, \vec{v}_n) = 0$$

if any two of the $\vec{v}_i$s are equal.

These are the core, defining properties of det.

Here are more properties, which we *deduced* from the three core properties above:

(4) Multiplying a row or column by a number $r \in \mathbf{R}$ multiplies the determinant by that r.

(5) Swapping two rows or columns in A multiplies the determinant by a -1.

(6) Adding a multiple of a row or column onto another row or column doesn't change the determinant.

And here are some general computational facts we extract from these properties.

(8) $\det(A) \neq 0$ *if and only if* A is invertible.

(9) $\det(MN) = \det(M)\det(N)$.

(10) $\det(A^{\intercal}) = \det(A)$. (Why?)

(11) $\det \operatorname{diag}(\lambda_1, \ldots, \lambda_n) = \prod_{i=1}^{n} \lambda_i$.

(12) More generally, if A is triangular, then $\det(A)$ is the product of the entries along the diagonal. (Why?)

Let's compute the determinant of the following matrices:

$$\begin{pmatrix} 1 & 2 & 4 \\ 1 & 3 & 9 \\ 1 & 4 & 16 \end{pmatrix}, \quad \begin{pmatrix} 1 & 1 & 1 & 2 \\ 1 & 1 & 2 & 2 \\ 1 & 2 & 2 & 3 \\ 2 & 2 & 3 & 4 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 1 & 2 & 5 \\ 1 & 2 & 5 & 14 \\ 2 & 5 & 14 & 42 \\ 5 & 14 & 42 & 132 \end{pmatrix}, \quad \begin{pmatrix} 1 & 2 & 5 & 14 \\ 2 & 5 & 5 & 14 \\ 5 & 14 & 42 & 132 \\ 14 & 42 & 132 & 429 \end{pmatrix}$$

I was half-joking about the last two; this is actually a neat little piece of mathematics: there's only one sequence of integers $c_0, c_1, c_2, \ldots$ such that for any $n \geq 1$,

$$\det(c_{i+j-2}) = \det(c_{i+j-1}) = 1.$$

These are called the *Catalan numbers*, and your mathematical life isn't complete until you've read about them! (I was going to put a problem about these on the homework, but I thought that might be too much.)

Suppose $\sigma\colon \{1,\ldots,n\} \longrightarrow \{1,\ldots,n\}$ a *permutation* of $\{1,\ldots,n\}$ – i.e., a bijection from that set to itself.

One can express a permutation very compactly, by writing down the matrix

$$P_\sigma = \begin{pmatrix} \hat{e}_{\sigma(1)} & \cdots & \hat{e}_{\sigma(n)} \end{pmatrix}$$

called the *permutation matrix* corresponding to σ.

This is great, because matrix multiplication corresponds to composition of permutations:

$$P_{\sigma\circ\tau} = P_\sigma P_\tau \quad \text{and} \quad P_{id} = I.$$

Also, these matrices are orthogonal; in fact,

$$P_\sigma^\intercal = P_\sigma^{-1} = P_{\sigma^{-1}}.$$

Here's a permutation matrix for $n = 5$:

$$\begin{pmatrix} \hat{e}_2 \\ \hat{e}_4 \\ \hat{e}_1 \\ \hat{e}_3 \\ \hat{e}_5 \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

What's its determinant?

This is a general pattern: the determinant of a permutation matrix P_σ is called the *sign* of the permutation σ:

$$\operatorname{sgn}(\sigma) \coloneqq \det(P_\sigma)$$

In effect, it's

$$(-1)^{\text{number of swaps in } \sigma}.$$

What's weird about this is that you can imagine performing more or fewer swaps to get sigma. The magic of determinants is telling you that the *parity* of the number of swaps stays the same!

Note that $\operatorname{sgn}(\sigma) = \operatorname{sgn}(\sigma^{-1})$. (Why?)

It turns out that permutations give you a formula for the determinant of any matrix $A = \begin{pmatrix} \vec{v}_1 & \cdots & \vec{v}_n \end{pmatrix}$. Let's see why.

First, the j-th column $\vec{v}_j$ can be written as

$$\vec{v}_j = \sum_{k=1}^{n} a_{k,j}\hat{e}_k.$$

The multilinearity of det can then be deployed:

$$\begin{aligned}\det(A) &= \det\left(\sum_{k(1)=1}^{n} a_{k(1),j}\hat{e}_{k(1)}, \ldots, \sum_{k(n)=1}^{n} a_{k(n),j}\hat{e}_{k(n)}\right) \\ &= \sum_{k(1)=1}^{n} \cdots \sum_{k(n)=1}^{n} \left(\prod_{i=1}^{n} a_{k(i),i}\right) \det(\hat{e}_{k(1)}, \ldots, \hat{e}_{k(n)})\end{aligned}$$

Now all those sums can be combined into one sum. You're summing over the *set* E_n of all maps $k\colon \{1, \ldots, n\} \longrightarrow \{1, \ldots, n\}$:

$$\det(A) = \sum_{k \in E_n} \left(\prod_{i=1}^{n} a_{k(i),i}\right) \det(\hat{e}_{k(1)}, \ldots, \hat{e}_{k(n)}).$$

$$\det(A) = \sum_{k \in E_n} \left(\prod_{i=1}^{n} a_{k(i),i} \right) \det(\hat{e}_{k(1)}, \ldots, \hat{e}_{k(n)}).$$

Now we use the alternatingness: if any two columns are equal, then the determinant is zero. So any summand in which $k\colon \{1, \ldots, n\} \longrightarrow \{1, \ldots, n\}$ is not injective doesn't appear:

$$\det(A) = \sum_{\sigma \in \Sigma_n} \left(\prod_{i=1}^{n} a_{\sigma(i),i} \right) \det(\hat{e}_{\sigma(1)}, \ldots, \hat{e}_{\sigma(n)}),$$

where Σ_n is the set of permutations of $\{1, \ldots, n\}$.

Now we have:

$$\begin{aligned}
\det(A) &= \sum_{\sigma\in\Sigma_n}\left(\prod_{i=1}^{n} a_{\sigma(i),i}\right)\det(\hat{e}_{\sigma(1)},\ldots,\hat{e}_{\sigma(n)}) \\
&= \sum_{\sigma\in\Sigma_n}\left(\prod_{i=1}^{n} a_{\sigma(i),i}\right)\det(P_\sigma) \\
&= \sum_{\sigma\in\Sigma_n}\operatorname{sgn}(\sigma)\left(\prod_{i=1}^{n} a_{\sigma(i),i}\right).
\end{aligned}$$

$$\det(A) = \sum_{\sigma \in \Sigma_n} \operatorname{sgn}(\sigma) \left(\prod_{i=1}^{n} a_{\sigma(i),i} \right)$$

There it is – *the Leibniz formula* for the determinant.

Do you care? Well, if you're trying to program a computer to compute determinants, no. Evaluating this formula involves $\Omega(n!\,n)$ operations. Gaussian elimination uses $O(n^3)$ operations. We have a winner.

On the other hand, the fact that there *is* a formula is vaguely reassuring. But there's another advantage …

Think of the determinant as a function from $\mathbf{R}^{n^2} \longrightarrow \mathbf{R}$; this formula expresses that function as a *polynomial* in n^2 variables. That means that it's continuous and infinitely differentiable. So this leads us to the following result:

Proposition. *Suppose $A = (a_{i,j})$ an invertible $n \times n$ matrix. Then there exists an $\varepsilon > 0$ such that if $A' = (a'_{i,j})$ is an $n \times n$ matrix such that $|a'_{i,j} - a_{i,j}| < \varepsilon$, then A' is invertible too.*

That is, invertible matrices are stable under small perturbations.

Here's another wacky-sounding consequence. Suppose $L \subseteq \mathbf{R}^{n^2}$ is a line. Then if there exists one point on L that corresponds to an invertible matrix, then all but finitely many points on L correspond to invertible matrices.

Question. Suppose A an $n \times n$ matrix. For how many real numbers $t \in \mathbf{R}$ is $A + tI$ is invertible (none, finitely many, infinitely many, all)?