11-10-2021

Determinants - Axioms

Determinants are functions which take matrices as inputs and produce numbers. They are of enormous importance in linear algebra, but perhaps you've also seen them in other courses. They're used to define the **cross product** of two 3-dimensional vectors. They appear in **Jacobians** which occur in the **change-of-variables formula for multiple integrals**.

Determinants take as inputs $n \times n$ (square) matrices with entries in R, where R is a commutative ring with identity. The set of such matrices is denoted $M(n, R)$. (I will be careful to prove everything for a commutative ring with identity — but for many things, you can pretend that R is just the real numbers $\mathbb{R}$ if it helps you to understand.)

In this section, I'll define determinants as functions satisfying three axioms. Mathematicians often proceed in this way: Define an object by identifying *properties* which characterize it, rather than simply writing down a formula.

Definition. A **determinant function** is a function $D : M(n, R) \to R$ which satisfies the following axioms:

1. D is a linear function in each row. That is, if $a \in R$ and $x, y \in R^n$,

$$D \begin{bmatrix} \leftarrow & r_1 & \rightarrow \\ & \vdots & \\ \leftarrow & ax + y & \rightarrow \\ & \vdots & \\ \leftarrow & r_n & \rightarrow \end{bmatrix} = a \cdot D \begin{bmatrix} \leftarrow & r_1 & \rightarrow \\ & \vdots & \\ \leftarrow & x & \rightarrow \\ & \vdots & \\ \leftarrow & r_n & \rightarrow \end{bmatrix} + D \begin{bmatrix} \leftarrow & r_1 & \rightarrow \\ & \vdots & \\ \leftarrow & y & \rightarrow \\ & \vdots & \\ \leftarrow & r_n & \rightarrow \end{bmatrix}.$$

2. A matrix with two equal rows has determinant 0:

$$D \begin{bmatrix} \leftarrow & r_1 & \rightarrow \\ & \vdots & \\ \leftarrow & x & \rightarrow \\ & \vdots & \\ \leftarrow & x & \rightarrow \\ & \vdots & \\ \leftarrow & r_n & \rightarrow \end{bmatrix} = 0.$$

3. $D(I) = 1$, where I is the $n \times n$ identity matrix.

Note: Later on, you'll see the following standard notations instead of "D" for determinants. You can either write "det" in place of D, or put vertical bars around the matrix:

$$\det A \quad \text{or} \quad |A|.$$

For now, I'll use "D".

If you've never seen something defined this way, you might be a bit uneasy. Am I going to give you a formula? Not yet.

First, a formula may be fine for computing things, but it doesn't really tell you *what the thing is.* In that respect, it's superficial, like judging someone by their appearance. Giving axioms for a thing can give a deeper view of what the thing is, and does — its *essence*

Second, defining determinants using axioms makes it a lot easier to prove many of the important properties of determinants — for example, that the determinant of a product of matrices is the product of the determinants.

You still want a formula or a recipe for computing determinants — right? Well, in some of the examples below, we'll see how you can compute determinants just using the axioms, or using row reduction. Be patient, and you'll feel better shortly!

Let's see how the axioms look like in particular cases.

The first axiom (linearity) is probably the hardest to understand. It allows you to add or subtract, or move constants in and out, *in a single row, assuming that all the other rows stay the same.* It is easier to show you what this means than to describe it in words.

For example, here are two determinants being combined into one. The two third rows are added, and the other two rows (which must be the same in both matrices) are unchanged:

$$D\begin{bmatrix} a & b & c \\ d & e & f \\ x_1 & x_2 & x_3 \end{bmatrix} + D\begin{bmatrix} a & b & c \\ d & e & f \\ y_1 & y_2 & y_3 \end{bmatrix} = D\begin{bmatrix} a & b & c \\ d & e & f \\ x_1+y_1 & x_2+y_2 & x_3+y_3 \end{bmatrix}.$$

In row 3, I used the fact that

$$(x_1, x_2, x_3) + (y_1, y_2, y_3) = (x_1 + y_1, x_2 + y_2, x_3 + y_3).$$

You can also take a single determinant apart into two determinants. In this example, we have subtraction instead of addition. All the action takes place in row 2; the first and third rows are the same in all of the matrices.

$$D\begin{bmatrix} a & b & c \\ x_1-y_1 & x_2-y_2 & x_3-y_3 \\ d & e & f \end{bmatrix} = D\begin{bmatrix} a & b & c \\ x_1 & x_2 & x_3 \\ d & e & f \end{bmatrix} - D\begin{bmatrix} a & b & c \\ y_1 & y_2 & y_3 \\ d & e & f \end{bmatrix}.$$

Linearity also allows you to factor a constant out of a *single row*, in this case row 1:

$$D\begin{bmatrix} kx_1 & kx_2 & kx_3 \\ a & b & c \\ d & e & f \end{bmatrix} = k \cdot D\begin{bmatrix} x_1 & x_2 & x_3 \\ a & b & c \\ d & e & f \end{bmatrix}.$$

You can do the opposite: Multiply a constant outside the determinant into a *single row*:

$$5 \cdot D\begin{bmatrix} a & b \\ c & d \end{bmatrix} = D\begin{bmatrix} 5a & 5b \\ c & d \end{bmatrix}.$$

You could do this as well:

$$5 \cdot D\begin{bmatrix} a & b \\ c & d \end{bmatrix} = D\begin{bmatrix} a & b \\ 5c & 5d \end{bmatrix}.$$

Here's an example where you "take apart" a determinant using linearity. Notice that the first two rows are the same in all the matrices; all the action takes place in the third row:

$$D\begin{bmatrix} a & b & c \\ d & e & f \\ x_1+ky_1 & x_2+ky_2 & x_3+ky_3 \end{bmatrix} = D\begin{bmatrix} a & b & c \\ d & e & f \\ x_1 & x_2 & x_3 \end{bmatrix} + D\begin{bmatrix} a & b & c \\ d & e & f \\ ky_1 & ky_2 & ky_3 \end{bmatrix} =$$

$$D\begin{bmatrix} a & b & c \\ d & e & f \\ x_1 & x_2 & x_3 \end{bmatrix} + k \cdot D\begin{bmatrix} a & b & c \\ d & e & f \\ y_1 & y_2 & y_3 \end{bmatrix}.$$

First, I used linearity to break the given determinant up into two determinants. Then I factored k out of the third row of the second determinant.

Perhaps you feel that I haven't *really* told you what a determinant *is*, because I haven't given you a formula or recipe for *computing* a determinant. Those axioms seem pretty abstract. What you'd like is to start with a matrix and produce a *number*. In fact, the three axioms above are enough to be able to compute determinants (though not very efficiently). Here's an example.

Suppose I have a determinant function D for 2×2 real matrices — so D satisfies the three axioms above. *Using only the axioms*, I'll compute

$$D \begin{bmatrix} 1 & -1 \\ 3 & 2 \end{bmatrix}.$$

First, I'll break up the second row into a sum of a multiple of the first row and another vector:

$$(3, 2) = (3, -3) + (0, 5) = 3 \cdot (1, -1) + (0, 5).$$

Then I can use linearity to break the determinant up into two pieces.

$$D \begin{bmatrix} 1 & -1 \\ 3 & 2 \end{bmatrix} = D \begin{bmatrix} 1 & -1 \\ 3+0 & -3+5 \end{bmatrix} = D \begin{bmatrix} 1 & -1 \\ 3 & -3 \end{bmatrix} + D \begin{bmatrix} 1 & -1 \\ 0 & 5 \end{bmatrix} = 3 \cdot D \begin{bmatrix} 1 & -1 \\ 1 & -1 \end{bmatrix} + 5 \cdot D \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix} =$$

$$3 \cdot D \begin{bmatrix} 1 & -1 \\ 1 & -1 \end{bmatrix} + 5 \cdot D \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix} = 3 \cdot 0 + 5 \cdot D \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix} = 5 \cdot D \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix}.$$

Notice that in the second equality the first row stays the same, while the new second rows are $(3, -3)$ and $(0, 5)$.

Notice that linearity also allows me to factor 3 and 5 out of the second rows for the third equality.

The alternating axiom says that a matrix with two equal rows has determinant 0, and that gave me the fifth equality.

Now I do a similar trick with the first row:

$$5 \cdot D \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix} = 5 \cdot D \begin{bmatrix} 1+0 & 0+-1 \\ 0 & 1 \end{bmatrix} = 5 \cdot D \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + 5 \cdot D \begin{bmatrix} 0 & -1 \\ 0 & 1 \end{bmatrix} =$$

$$5 \cdot 1 + (-1) \cdot 5 \cdot D \begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix} = 5 + (-5) \cdot 0 = 5.$$

The second equality used linearity applied to the first row: $(1+0, 0+-1) = (1, 0) + (0, -1)$. The third equality used the fact that the determinant of the identity matrix is 1, and used linearity to factor -1 out of the second row. The fourth equality used the fact that the determinant of a matrix with two equal rows is 0.

Thus,

$$D \begin{bmatrix} 1 & -1 \\ 3 & 2 \end{bmatrix} = 5.$$

Notice that we computed a determinant *using only the axioms for a determinant*. We don't have a formula at the moment (though you may have seen a formula for 2×2 determinants before). It's true that the computation took a lot of steps, and this is not the best way to do this — but this example gives some evidence that our axioms actually tell what determinants are.

The next two results are often useful in computations.

First, you might suspect that *a matrix with an all-zero row has determinant* 0, and it's easy to prove using linearity. Rather than give a formal proof, I'll illustrate the idea with a particular example.

$$D \begin{bmatrix} 0 & 0 & 0 \\ 2 & 0 & 17 \\ 7 & -6 & 1 \end{bmatrix} = D \begin{bmatrix} 0+0 & 0+0 & 0+0 \\ 2 & 0 & 17 \\ 7 & -6 & 1 \end{bmatrix} = D \begin{bmatrix} 0 & 0 & 0 \\ 2 & 0 & 17 \\ 7 & -6 & 1 \end{bmatrix} + D \begin{bmatrix} 0 & 0 & 0 \\ 2 & 0 & 17 \\ 7 & -6 & 1 \end{bmatrix},$$

The last equation says "(stuff) = (stuff) + (stuff)". This means that (stuff) = 0. So

$$D \begin{bmatrix} 0 & 0 & 0 \\ 2 & 0 & 17 \\ 7 & -6 & 1 \end{bmatrix} = 0.$$

The same idea can be used to prove the result in general.

The next result tells us what happens to a determinant when we swap two rows of the matrix.

Lemma. If $D : M(n, r) \to R)$ is a function which is linear in the rows (Axiom 1) and is 0 when a matrix has equal rows (Axiom 2), then swapping two rows multiplies the value of D by -1:

$$D \begin{bmatrix} & \vdots & \\ \leftarrow & r_i & \rightarrow \\ & \vdots & \\ \leftarrow & r_j & \rightarrow \\ & \vdots & \end{bmatrix} = -D \begin{bmatrix} & \vdots & \\ \leftarrow & r_j & \rightarrow \\ & \vdots & \\ \leftarrow & r_i & \rightarrow \\ & \vdots & \end{bmatrix}.$$

Proof. The proof will use the first and second axioms repeatedly. The idea is to swap rows i and j by adding or subtracting rows.

In the diagrams below, the gray rectangles represent rows which are unchanged and the same in all of the matrices. All the "action" takes place in row i and row j.

$$\mathrm{D} \begin{bmatrix} \\ \longleftarrow r_i \longrightarrow \\ \\ \longleftarrow r_j \longrightarrow \\ \\ \end{bmatrix} \underset{\text{(by Axiom 2)}}{=} \mathrm{D} \begin{bmatrix} \\ \longleftarrow r_i \longrightarrow \\ \\ \longleftarrow r_j \longrightarrow \\ \\ \end{bmatrix} + \mathrm{D} \begin{bmatrix} \\ \longleftarrow r_i \longrightarrow \\ \\ \longleftarrow r_i \longrightarrow \\ \\ \end{bmatrix} \underset{\text{(by Axiom 1)}}{=} \mathrm{D} \begin{bmatrix} \\ \longleftarrow r_i \longrightarrow \\ \\ \leftarrow r_i + r_j \rightarrow \\ \\ \end{bmatrix} \underset{\text{(by Axiom 2)}}{=} \mathrm{D} \begin{bmatrix} \\ \longleftarrow r_i \longrightarrow \\ \\ \leftarrow r_i + r_j \rightarrow \\ \\ \end{bmatrix} - \mathrm{D} \begin{bmatrix} \\ \leftarrow r_i + r_j \rightarrow \\ \\ \leftarrow r_i + r_j \rightarrow \\ \\ \end{bmatrix} \underset{\text{(by Axiom 1)}}{=}$$

$$\mathrm{D} \begin{bmatrix} \\ \longleftarrow -r_j \longrightarrow \\ \\ \leftarrow r_i + r_j \rightarrow \\ \\ \end{bmatrix} \underset{\text{(by Axiom 2)}}{=} \mathrm{D} \begin{bmatrix} \\ \longleftarrow -r_j \longrightarrow \\ \\ \leftarrow r_i + r_j \rightarrow \\ \\ \end{bmatrix} + \mathrm{D} \begin{bmatrix} \\ \longleftarrow -r_j \longrightarrow \\ \\ \longleftarrow -r_j \longrightarrow \\ \\ \end{bmatrix} \underset{\text{(by Axiom 1)}}{=} \mathrm{D} \begin{bmatrix} \\ \longleftarrow -r_j \longrightarrow \\ \\ \longleftarrow r_i \longrightarrow \\ \\ \end{bmatrix} \underset{\text{(by Axiom 1)}}{=} -\mathrm{D} \begin{bmatrix} \\ \longleftarrow r_j \longrightarrow \\ \\ \longleftarrow r_i \longrightarrow \\ \\ \end{bmatrix}$$

Notice that in each addition or subtraction step (the steps that use Axiom 1), only one of row i or row j changes at a time. $\square$

Remarks. (a) I'll show later that it's enough to assume (instead of Axiom 2) that $D(A) = 0$ vanishes whenever two *adjacent* rows of A are equal. (This is a technical point which you can forget about until we need it.)

(b) Suppose that $D : M(n, r) \to R$ is a function satisfying Axioms 1 and 3, and suppose that swapping two rows multiplies the value of D by -1. Must D satisfy Axiom 2? In other words, is "swapping multiplies the value by -1" *equivalent* to "equal rows means determinant 0"?

Assuming that swapping two rows multiplies the value of D by -1, I have

$$D(A) = D \begin{bmatrix} \leftarrow & r_1 & \rightarrow \\ & \vdots & \\ \leftarrow & x & \rightarrow \\ & \vdots & \\ \leftarrow & x & \rightarrow \\ & \vdots & \\ \leftarrow & r_n & \rightarrow \end{bmatrix} = -D \begin{bmatrix} \leftarrow & r_1 & \rightarrow \\ & \vdots & \\ \leftarrow & x & \rightarrow \\ & \vdots & \\ \leftarrow & x & \rightarrow \\ & \vdots & \\ \leftarrow & r_n & \rightarrow \end{bmatrix} = -D(A).$$

(I swapped the two equal x-rows, which is why the matrix didn't change. But by assumption, this useless swap multiplies D by -1.)

Hence, $2 \cdot D(A) = 0$.

If R is $\mathbb{R}$, $\mathbb{Q}$, $\mathbb{C}$, or $\mathbb{Z}_n$ for n prime and not equal to 2, then $2 \cdot D(A) = 0$ implies $D(A) = 0$. However, if $R = \mathbb{Z}_2$, then $2x = 0$ for all x. Hence, $2 \cdot D(A) = 0$, no matter what $D(A)$ is. I can't conclude that $D(A) = 0$ in this case. Therefore, Axiom 2 need not hold. You can see, however, that it will hold if R is a field of characteristic other than 2.

Fortunately, since I took "equal rows means determinant 0" as an *axiom* for determinants, and since the lemma shows that this *implies* that "swapping rows multiplies the determinant by -1", I know that *both* of these properties will hold for determinant functions. $\square$

Example. (**Computing determinants using the axioms**) Suppose that $D : M(3, \mathbb{R}) \to \mathbb{R}$ is a determinant function and

$$D \begin{bmatrix} \leftarrow & a_1 & \rightarrow \\ \leftarrow & b & \rightarrow \\ \leftarrow & c & \rightarrow \end{bmatrix} = 5 \quad \text{and} \quad D \begin{bmatrix} \leftarrow & a_2 & \rightarrow \\ \leftarrow & b & \rightarrow \\ \leftarrow & c & \rightarrow \end{bmatrix} = -3.$$

Compute

$$D \begin{bmatrix} \leftarrow & c & \rightarrow \\ \leftarrow & b & \rightarrow \\ \leftarrow & a_1 + 4a_2 & \rightarrow \end{bmatrix}.$$

$$D \begin{bmatrix} \leftarrow & c & \rightarrow \\ \leftarrow & b & \rightarrow \\ \leftarrow & a_1 + 4a_2 & \rightarrow \end{bmatrix} = -D \begin{bmatrix} \leftarrow & a_1 + 4a_2 & \rightarrow \\ \leftarrow & b & \rightarrow \\ \leftarrow & c & \rightarrow \end{bmatrix} =$$

$$-\left(D \begin{bmatrix} \leftarrow & a_1 & \rightarrow \\ \leftarrow & b & \rightarrow \\ \leftarrow & c & \rightarrow \end{bmatrix} + 4D \begin{bmatrix} \leftarrow & a_2 & \rightarrow \\ \leftarrow & b & \rightarrow \\ \leftarrow & c & \rightarrow \end{bmatrix} \right) = -(5 + 4 \cdot (-3)) = 7. \quad \square$$

Determinants and elementary row operations.

Elementary row operations are used to reduce a matrix to row reduced echelon form, and as a consequence, to solve systems of linear equations. We can use them to compute determinants with more ease than using the axioms directly — and, even when we have some better algorithms (like expansion by cofactors), row operations will be useful in simplifying computations. How are determinants affected by elementary row operations?

Adding a multiple of a row to another row does not change the determinant. Suppose, for example, I'm performing the operation $r_i \to r_i + a \cdot r_j$. Let

$$A = \begin{bmatrix} & \vdots & \\ \leftarrow & r_i & \rightarrow \\ & \vdots & \\ \leftarrow & r_j & \rightarrow \\ & \vdots & \end{bmatrix}.$$

Then

$$D \begin{bmatrix} & \vdots & \\ \leftarrow & r_i + a \cdot r_j & \rightarrow \\ & \vdots & \\ \leftarrow & r_j & \rightarrow \\ & \vdots & \end{bmatrix} = D \begin{bmatrix} & \vdots & \\ \leftarrow & r_i & \rightarrow \\ & \vdots & \\ \leftarrow & r_j & \rightarrow \\ & \vdots & \end{bmatrix} + a \cdot D \begin{bmatrix} & \vdots & \\ \leftarrow & r_j & \rightarrow \\ & \vdots & \\ \leftarrow & r_j & \rightarrow \\ & \vdots & \end{bmatrix} = D \begin{bmatrix} & \vdots & \\ \leftarrow & r_i & \rightarrow \\ & \vdots & \\ \leftarrow & r_j & \rightarrow \\ & \vdots & \end{bmatrix} = D(A).$$

Therefore, this kind of row operation leaves the determinant unchanged.

The alternating property implies that swapping two rows multiplies the determinant by -1. For example,

$$D\begin{bmatrix} a & b \\ c & d \end{bmatrix} = -D\begin{bmatrix} c & d \\ a & b \end{bmatrix}.$$

Our third kind of row operation involves multiplying a row by a number (which must be invertible in the ring from which the entries of the matrix come). So if I wanted to multiply the second row of a real matrix by 19, I could do this:

$$D\begin{bmatrix} a & b \\ c & d \end{bmatrix} = \frac{1}{19} \cdot 19 \cdot D\begin{bmatrix} a & b \\ c & d \end{bmatrix} = \frac{1}{19} D\begin{bmatrix} a & b \\ 19c & 19d \end{bmatrix}.$$

Thus, multiplying the second row by 19 leaves a factor of $\dfrac{1}{19}$ outside.

However, when you're using row operations *to compute a determinant*, you usually want to factor a number out of a row, which you can do using the linearity axiom. Thus:

$$D\begin{bmatrix} 6a & 6b \\ c & d \end{bmatrix} = 6 \cdot D\begin{bmatrix} a & b \\ c & d \end{bmatrix}.$$

Example. (**Computing a determinant using row operations**) Suppose D is a determinant function on 2×2 real matrices. Use row operations to compute the following determinant:

$$D\begin{bmatrix} 1 & 5 \\ -3 & 4 \end{bmatrix}.$$

$$D\begin{bmatrix} 1 & 5 \\ -3 & 4 \end{bmatrix} \underset{r_2 \to r_2 + 3r_1}{=} D\begin{bmatrix} 1 & 5 \\ 0 & 19 \end{bmatrix} \underset{\text{(Linearity)}}{=} 19 \cdot D\begin{bmatrix} 1 & 5 \\ 0 & 1 \end{bmatrix} \underset{r_1 \to r_1 - 5r_2}{=}$$

$$19 \cdot D\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = 19 \cdot 1 = 19. \quad \square$$

Example. (**Computing a determinant using row operations**) Suppose D is a determinant function on 2×2 matrices with entries in $\mathbb{Z}_5$. Use row operations to compute the following determinant:

$$D\begin{bmatrix} 2 & 1 \\ 3 & 4 \end{bmatrix}.$$

In $\mathbb{Z}_5$, I have $1 = 2 \cdot 3$. So

$$D\begin{bmatrix} 2 & 1 \\ 3 & 4 \end{bmatrix} = D\begin{bmatrix} 2 \cdot 1 & 2 \cdot 3 \\ 3 & 4 \end{bmatrix} \underset{\text{(Linearity)}}{=} 2 \cdot D\begin{bmatrix} 1 & 3 \\ 3 & 4 \end{bmatrix} \underset{r_2 \to r_2 + 2r_1}{=}$$

$$2 \cdot D\begin{bmatrix} 1 & 3 \\ 0 & 0 \end{bmatrix} = 2 \cdot 0 = 0.$$

I used the fact that a matrix with an all-zero row has determinant 0. $\square$

Your experience with row reducing matrices tells you that either the row reduced echelon form will be the identity, or it will have an all-zero row at the bottom. In the second case, we've seen that the determinant is 0. In the first case, there may be constants multiplying the determinant, and the determinant of the identity is 1 — and so, you know the value of the determinant by multiplying everything together.

Row reduction gives you a way of computing determinants that is a little more practical than applying the axioms directly. It should also convince you that, starting with the three determinant axioms, we now have something which takes a square matrix and produces a number.

There are still some questions we need to address. We need to *prove* that there really are functions which satisfy the three axioms. (Just being able to compute in particular cases is not a proof.) This is called an **existence** question. We will answer this question by producing an algorithm which gives a determinant function for square matrices. It is called **expansion by cofactors**.

Could there be multiple determinant functions? Could more than one function on square matrices satisfy the axioms? This seems unlikely given that we were able to start with numerical matrices and compute specific numbers — but maybe a different approach might produce a different answer. This is called a **uniqueness** question.

We will show that, in fact, there is only one determinant function — a function satisfying the three axioms — on square matrices. It can be computed in various ways, but you'll get the same answer in all cases.

Along the way, we'll find another approach which uses **permutations** to compute determinants. We'll also prove some important properties of determinants, such as the rule for products.