Worksheet 6: determinants

Introducing the concept of **determinant** in a first linear algebra class is always a challenge, because the determinant of a matrix is some initially very abstruse and complicated *magic formula* that then turns out to have all sort of good properties. However this is also an opportunity to understand an important idea in modern mathematics: it is important to shift the focus from the question " what **is** the determinant?" to the question " what **does the determinant do** for us?". Mathematicians wanted some gadget that did a particular job, and then it just turned out that the gadget that does that job is the ugly weird formula that makes our eyes cross. Ok, so given this premise, let us dig in.

1 What should the determinant do?

The determinant is a gadget that should allow us to solve the following problems:

1. Decide if a linear function is invertible.

2. Decide if a list of vectors is linearly independent.

3. Determine the dimension of the range of a linear function.

The determinant is a function that takes as input square matrices and returns real numbers:

$$\begin{array}{rccc} det: & Mat(n \times n) & \to & \mathbb{R} \\ \\ & M & \mapsto & det(M) \end{array}$$

The way it accomplishes the above goals is the following:

1. A linear function L is invertible if and only if the determinant of ANY matrix that represents it is non-zero.

2. To determine if a list of n vectors is linearly independent, one does the following. First, choose a basis for your vector space and express your vectors as column vectors. Arrange them as the consecutive columns of a matrix. Then they are linearly independent if and only if you can find an $n \times n$ submatrix that has non-zero determinant.

3. To determine the dimension of the range of a linear function $L : V \to W$, one does the following. Let M_L be a matrix that represents L for some choice of bases. Then dimension of the range of L is equal to the size of the larges square submatrix of M_L that has non-zero determinant.

Question 1 *What is the determinant in the* 1×1 *and* 2×2 *cases?*

Question 2 *Consider the vectors* $(1, 2, 3, 4, 5, 6)$ *and* $(2, 4, 5, 8, 10, 12) \in \mathbb{R}^6$. *Use the determinant to decide if they are linearly independent.*

2 How do we accomplish that?

The key idea is that to obtain a gadget that does the things we mentioned in the previous section, we ask this gadget to satisfy a bunch of properties. First off, there is another change of point of view that is necessary and important. We think of the inputs of the determinant function not as being square matrices, but instead the ordered set of the column vectors of a square matrix. This may seem like a completely silly thing, but we will soon see how important this shift in point of view is. So now the determinant wants to be a function:

$$\begin{array}{rccc} det: & (\mathbb{R}^n)^n & \to & \mathbb{R} \\ & v_1, \ldots, v_n & \mapsto & det(v_1, \ldots, v_n) \end{array}$$

We wish det to be a function that satisfies these three magic properties:

Normalization: Denote by $e_1, \ldots, e_n$ the standard basis of $\mathbb{R}^n$. Then

$$det(e_1, \ldots, e_n) = 1.$$

(In the matrix perspective, this is saying that the determinant of the identity matrix is required to be equal to 1).

Linearity in each entry: Hold $n - 1$ of the vectors fixed. Without loss of generality let us assume $v_2, \ldots v_n$ are fixed. Then

$$det(\alpha v_1 + \beta w_1, v_2, \ldots, v_n) = \alpha \cdot det(v_1, v_2, \ldots, v_n) + \beta \cdot det(w_1, v_2, \ldots, v_n)$$

This is saying that if you hold $n - 1$ entries of det fixed and think of one entry as variable inputs, then you obtain a linear function from $\mathbb{R}^n \to \mathbb{R}$. Notice, this does NOT mean that det is a linear function, when you allow ALL vectors to vary as inputs!

Sign reversing property: The sign changes when two vectors are switched. Here we switch the first two, but it doesn't matter which ones you switch:

$$det(v_1, v_2, \ldots, v_n) = -det(v_2, v_1, \ldots, v_n)$$

Let us now see how, if we can find such a function, we indeed obtain the results from Section 1. First we observe the following additional properties of the determinant.

(1) If one of the entries is the zero vector, the determinant is zero.

$$det(0, v_2, \dots, v_n) = 0.$$

This follows from (L).

(2) If two entries are the same, the determinant is zero.

$$det(v, v, \dots, v_n) = 0.$$

This follows from (SRP).

(3) If one of the entries is a linear combination of the others, the determinant is zero.

$$det(v_1, v_2, \dots, v_{n-1}, \sum_{i=1}^{n-1} \alpha_i v_i) = 0.$$

This follows from (L) and (SRP).

The punchline is the following: $det(v_1, v_2, \dots, v_n) \neq 0$ if and only if $v_1, \dots, v_n$ are linearly independent!

Hence:

A linear function $L : \mathbb{R}^n \to \mathbb{R}^n$ is invertible $\iff$ $L(e_1), \dots, L(e_n)$ are linearly independent

$$\iff det(L(e_1), \dots, L(e_n)) \neq 0 \iff det(M_L) \neq 0.$$

To establish the second and third result from Section 1 we make use of the following idea: assume $m < n$.

m vectors in $\mathbb{R}^n$ are linearly independent if and only if there exist m coordinates such that looking ONLY at those coordinates, you have m linearly independent vectors in $\mathbb{R}^m$.

Notice that in the previous sentence I said "there exist", not "for every". Look at Question 2 to convince yourself of why this is an important distinction.

Question 3 *Use this idea to show that m column vectors (of arbitrary length) are linearly independent if and only if you can find an $m \times m$ submatrix that has non-zero determinant.*

3 These properties determine the determinant (pun intended)

I hopefully convinced you that if there existed a function *det* with the three magic properties from Section 2, then that would be the "linear independence detector” that we have dreamed of in Section 1. The upshot is that the three properties (N), (L), (SRP) in fact determine the determinant in a unique way.

Let me illustrate how this works for $n = 2$. Let M be an arbitrary 2×2 matrix:

$$M = \begin{bmatrix} a & c \\ b & d \end{bmatrix}$$

Then, recalling that we denote by e_1, e_2 the standard basis of $\mathbb{R}^2$, we have:

$$det(M) = det(ae_1 + be_2, ce_1 + de_2).$$

Now we apply property (L) in the first entry, to have:

$$det(ae_1 + be_2, ce_1 + de_2) = a \cdot det(e_1, ce_1 + de_2) + b \cdot det(e_2, ce_1 + de_2).$$

Now we apply property (L) in the second entry, to have:

$$a \cdot det(e_1, ce_1 + de_2) + b \cdot det(e_2, ce_1 + de_2) = ac \cdot det(e_1, e_1) + ad \cdot det(e_1, e_2) + bc \cdot det(e_2, e_1) + bd \cdot det(e_2, e_2).$$

By (SRP) the first and fourth summands are zero, because they have two repeated entries. We can also flip the sign of the third summand and switch the order of the entries.

$$ac \cdot det(e_1, e_1) + ad \cdot det(e_1, e_2) + bc \cdot det(e_2, e_1) + bd \cdot det(e_2, e_2) = ad \cdot det(e_1, e_2) - bc \cdot det(e_1, e_2).$$

Finally by (N), $det(e_1, e_2) = 1$, thus obtaining:

$$det(M) = ad - bc.$$

Problem 1 *Do the same in the $n = 3$ case. I.e. take an arbitrary 3×3 matrix*

$$M = \begin{bmatrix} a & d & g \\ b & e & h \\ c & f & i \end{bmatrix}$$

Think of M as three vectors in $\mathbb{R}^3$ that you can express as linear combinations of the three standard basis vectors. Use properties (N), (L) and (SRP) to manipulate the determinant until you have a bunch of summands some of which are zero, and some of which are multiples of $det(e_1, e_2, e_3)$.