Determinants: A Summary

1 Introduction

So far we have only defined determinants of 2×2 matrices. Given $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in \mathbb{R}^{2\times 2}$, we defined

$$\det(A) \;=\; ad - bc.$$

It can be shown that for every $A \in \mathbb{R}^{2\times 2}$, $|\det A|$ is the area of the parallelogram determined by the column vectors of A (see Figure 1 below). You will prove this on Worksheet 21.5. In particular, this means that $\det A = 0$ if and only if the columns of A determine a "degenerate" parallelogram with zero area. Put differently, the determinant of A is zero if and only if the columns of A are linearly dependent. It follows that A is invertible if and only if $\det A \neq 0$, so the determinant offers a convenient test for invertibility.

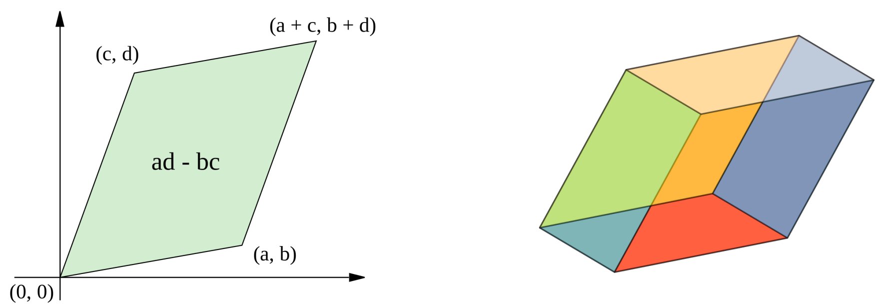

Figure 1: A parallelogram and a parallelepiped.

You may recall from Calc 3 that the 3-dimensional version of a parallelogram is called a *parallelepiped* (see Figure 1 above). And the 3-dimensional version of area is usually called *volume* (the 1-dimensional version is called *length*). More generally, the n-dimensional analogue of a parallelogram is called an *n-parallelepiped* and the n-dimensional analogue of area is called *n-volume*. What we would like to do now is define determinants of square matrices in general, so that for each $A \in \mathbb{R}^{n\times n}$, $|\det A|$ is the "n-dimensional volume" of the n-parallelepiped determined by the columns of A; in particular, this again means that for all $A \in \mathbb{R}^{n\times n}$ we will have $\det A = 0$ if and only if the columns of A are linearly dependent, so that A is invertible if and only if $\det(A) \neq 0$.

We recommend the 3-Blue-1-Brown animation to help you visualize the determinant. In fact, we recommend the entire 3-Blue-1-Brown series on linear algebra (link from Canvas Handouts page) to develop your intuition and appreciation of the subject (and other subjects!).

There are several ways to define the determinant, but in Math 217 this semester, we will not focus on the definition. Instead, we focus on how to *compute* the determinant, as well as its useful properties presented in Section 6.2 and 6.3 in the textbook, and summarized in the write-up. One definition is presented in Section 6.1 of the textbook, but reading 6.1 is optional— you will not be asked to know the definition in 6.1 or need it to do homework or exams. Instead, we will describe a convenient procedure for calculating determinants called *Laplace expansion.*

2 Laplace Expansions

It is easiest to describe Laplace expansions by example, so let's start with the 3×3 matrix

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}.$$

The first step is to pick a row or column of A. It doesn't matter which one – all choices will give the same determinant in the end. So let's say we pick the first column, $\begin{bmatrix} 1 \\ 4 \\ 7 \end{bmatrix}$. We will get one term in our calculation for each of the entries 1, 4, and 7. For each of these three entries, we cross out the row and column that entry belongs to and calculate the determinant of the 2×2 matrix that is left over. We multiply this 2×2 determinant by the corresponding entry, and add up the three results, making sure to flip the sign of the second one. So the calculation looks like this:

$$\begin{aligned} \det \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix} &= 1 \cdot \det \begin{bmatrix} 5 & 6 \\ 8 & 9 \end{bmatrix} - 4 \cdot \det \begin{bmatrix} 2 & 3 \\ 8 & 9 \end{bmatrix} + 7 \cdot \det \begin{bmatrix} 2 & 3 \\ 5 & 6 \end{bmatrix} \\ &= (45 - 48) - 4(18 - 34) + 7(12 - 15 = -3 + 24 - 21 = 0. \end{aligned}$$

But wait – how did we know to flip the sign on the second term? In general, you use a checker-board pattern of $+$'s and $-$'s, like this:

$$\begin{bmatrix} + & - & + \\ - & + & - \\ + & - & + \end{bmatrix}$$

Now you can do the same thing along any row or column, and you will always get the same answer. For instance, taking a Laplace expansion along the second row gives

$$\begin{aligned} \det \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix} &= -4 \cdot \det \begin{bmatrix} 2 & 3 \\ 8 & 9 \end{bmatrix} + 5 \cdot \det \begin{bmatrix} 1 & 3 \\ 7 & 9 \end{bmatrix} - 6 \cdot \det \begin{bmatrix} 1 & 2 \\ 7 & 8 \end{bmatrix} \\ &= -4(18 - 24) + 5(9 - 21) - 6(8 - 14) = 24 - 60 + 36 = 0. \end{aligned}$$

Notice that this calculation is telling us that A is not invertible, a fact which you can check by row-reducing.

How about larger matrices? Well, you can do the same thing, as long as you remember to use a full checker-board pattern starting with a $+$ in the upper-left corner. For instance, using the checker-board pattern

$$\begin{bmatrix} + & - & + & - \\ - & + & - & + \\ + & - & + & - \\ - & + & - & + \end{bmatrix}$$

and a pair of bottom-row Laplace expansions, we get

$$\det \begin{bmatrix} 1 & -2 & 3 & -4 \\ 0 & 5 & -6 & 7 \\ 0 & 0 & -8 & 9 \\ 0 & 0 & 0 & -10 \end{bmatrix} = -10 \cdot \det \begin{bmatrix} 1 & -2 & 3 \\ 0 & 5 & -6 \\ 0 & 0 & -8 \end{bmatrix} = (-10)(-8) \cdot \det \begin{bmatrix} 1 & -2 \\ 0 & 5 \end{bmatrix} = 400.$$

From this example we can draw several important observations:

(1) Computing determinants using Laplace expansions is a recursive, or iterative, procedure: in calculating an $n \times n$ determinant, we will obtain n different determinants of size $(n-1) \times (n-1)$, and we just have to keep going until we get down to a bunch of 2×2 determinants.

(2) When calculating a determinant using a Laplace expansion, you should choose a row or column having lots of zeros, if possible, since this will shorten the calculation.

(3) The determinant of an upper (or lower) triangular matrix is simply the product of its diagonal entries. (Recall that the matrix A is *upper triangular* if the (i,j)-entry of A is zero whenever $i > j$, and *lower triangular* if the (i,j)-entry of A is zero whenever $i < j$.)

Unfortunately, (1) means that calculating determinants using Laplace expansions is extremely inefficient unless the matrix happens to be "sparse," i.e., filled with lots of zeros. But fortunately, (3) will provide a faster way of calculating determinants of large matrices once we know how elementary row operations affect determinants.

2.1 Caution!

The skeptical math student reading this will wonder whether our "definition" of the determinant makes sense: how do we know that the above calculations of the determinant all give the same number? Maybe expanding along the first column gives a different determinant than expanding along the last row? How can we be sure we get the same result no matter what row or column we pick?

This is a serious problem, and the reason for the complicated definition in 6.1 or other complicated definitions that may be more popular (and more abstract) such as saying that the "determinant function on $\mathbb{R}^{n\times n}$ is the unique multi-linear function in the columns of the matrix which takes the value 1 on the identity matrix." Don't worry that this makes no sense to you right now—we'll explain it later. For now, you'll have to take our word for it that expanding along any row or column gives the same determinant. You can at least check it yourself in the 2×2 case:

Exercise: Verify that for a 2×2 matrix $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$, Laplace expansion along each of the two rows and each of the four columns yields the same determinant.

3 Elementary Row Operations and Determinants

Recall that there are three types of elementary row operations:

(i) *Row scaling:* scale some row by a nonzero scalar.

(ii) *Row swap:* interchange two rows.

(iii) *Row addition:* add a scalar multiple of one row to another row.

Here is how the elementary row operations affect the determinant:

- If the matrix A' is obtained from the square matrix A by scaling one row of A by the nonzero scalar c, then $\det A' = c \cdot \det A$.
- If the matrix A' is obtained from the square matrix A by interchanging two rows of A, then $\det A' = -\det A$.
- If the matrix A' is obtained from the square matrix A by adding a scalar multiple of some row of A to another row, then $\det A' = \det A$.

All of these facts can be shown to follow by picking appropriate rows/columns to expand along.

These facts have important theoretical consequences, but they also give us a more efficient way of calculating determinants: just use elementary row operations to reduce a matrix to upper-triangular form, making sure to keep track of which row-scalings and how many row-swaps you performed. To illustrate, consider the 4×4 matrix

$$A = \begin{bmatrix} \frac{1}{2} & -\frac{3}{2} & -\frac{1}{2} & \frac{5}{2} \\ 2 & -4 & -2 & 8 \\ -1 & 3 & 6 & -1 \\ 1 & -3 & -1 & 2 \end{bmatrix}.$$

Scaling the first row of A by 2 and then applying three successive row addition operations reduces A to the upper triangular matrix

$$R = \begin{bmatrix} 1 & -3 & -1 & 5 \\ 0 & 2 & 0 & -2 \\ 0 & 0 & 5 & 4 \\ 0 & 0 & 0 & -3 \end{bmatrix}.$$

Thus $-30 = \det R = 2 \cdot \det A$, so $\det A = -15$.

4 The Multiplicative Property of Determinants

Another important property of determinants is the **multiplicative property** which states that

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

for any two square matrices A and B of the same size. You should check that this implies that that the determinant of an invertible matrix is non-zero, since if $AB = I_n$, then

$$\det A\ \det B = \det(A\ B) = \det I_n = 1.$$

This says that neither A, nor its inverse B, can have determinant zero. It also tells us that for an invertible matrix A,

$$\det A^{-1} = \frac{1}{\det A}.$$

You will prove the formula $\det(AB) = (\det A)(\det B)$ on Worksheet 21.4.

5 The Determinant of a Linear Transformation

Now, let V be an finite dimensional vector space and let $T : V \to V$ be a linear transformation. We would like to define the determinant of T. We can model T by matrix multiplication, of course, by choosing a basis. Let $\mathcal{B} = (b_1, b_2, \ldots, b_n)$ be an ordered basis for V so that $[T]_\mathcal{B}$ is the $n \times n$ matrix modeling T in $\mathcal{B}$-coordinates. Suppose we define the determinant of T as

$$\det T = \det[T]_\mathcal{B}.$$

Is this well-defined? That is, is it independent of our choice of basis $\mathcal{B}$? It is possible that a different choice of ordered basis, say $\mathcal{C}$, would have a different determinant for $[T]_\mathcal{C}$?

Fortunately, the answer is *yes*, this determinant *is well-defined:* all bases produce the same determinant for T. To see this, recall that for each pair of ordered bases $\mathcal{B}$ and $\mathcal{C}$ of V, the $\mathcal{B}$-matrix and $\mathcal{C}$-matrix of T are similar to each other:

$$[T]_\mathcal{C} = S\ [T]_\mathcal{B}\ S^{-1},$$

where S is the change of basis matrix $S_{\mathcal{B}\to\mathcal{C}}$. Thus using the multiplicative property of determinants,

$$\det[T]_{\mathcal{C}} = \det\big(S\ [T]_{\mathcal{B}}\ S^{-1}\big) = \big(\det S\big)\big(\det[T]_{\mathcal{B}}\big)\big(\det S^{-1}\big) = \big(\det S\big)\big(\det[T]_{\mathcal{B}}\big)\big(\det S\big)^{-1} = \det[T]_{\mathcal{B}}.$$

Therefore we **can define** the determinant of T to be $\det[T]_{\mathcal{B}}$, where $\mathcal{B}$ is any ordered basis of V. The calculation shows that it does not matter which basis we pick. Determinants of linear transformations of the form $T : V \to V$ will be used extensively in Chapter 7.

If we have a linear transformation $T : \mathbb{R}^n \to \mathbb{R}^n$ with standard matrix A, then the determinant of T is just the determinant A. This is because $A = [T]_{\mathcal{E}}$, where $\mathcal{E}$ is the standard basis.

5.1 The Determinant as a scaling factor for volume.

Geometrically, the determinant of T (or more accurately, $|\det T|$) tells us the *scale factor* for the n-volume for the transformation T. For example, in dimension two, the transformation $T : \mathbb{R}^2 \to \mathbb{R}^2$ will take any plane figure Ω to the transformation plane figure $T[\Omega]$ where

$$\text{Area } T[\Omega] = |\det T| \text{ Area } \Omega.$$

You should try this for a few simple 2×2 matrices, such as a diagonal matrix or the matrix of a shear, to make sure you believe it.

We recommend the 3-Blue-1-Brown animation to help you visualize the determinant of a transformation.

This geometric intuition can help explain why $\det(AB) = (\det A)(\det B)$. If we think of A and B as the standard matrices of two linear transformations which we compose:

$$\mathbb{R}^n \xrightarrow{T_B} \mathbb{R}^n \xrightarrow{T_A} \mathbb{R}^n,$$

then we know that AB is the standard matrix of the composition. Since the first transformation scales volumes by $|\det B|$ and the second scales values by $|\det A|$, their composition scales volumes by $|\det A|\ |\det B|$.

6 Multilinearity and Alternating Properties

There are two very important properties of determinants: multi-linearity and the alternating property. Let us fix a size n, and consider $n\times n$ matrices. So far we have been thinking of the determinant as a function

$$\det : \mathbb{R}^{n\times n} \to \mathbb{R}.$$

Now, $\mathbb{R}^{n\times n}$ and $\mathbb{R} = \mathbb{R}^1$ are both vector spaces, so it makes sense to ask whether this function is a linear transformation; that is, do we have $\det(A + B) = \det A + \det B$ and $\det(cA) = c \cdot \det A$? It should not take you long to convince yourself that both of these properties fail, so det is *not* a linear transformation when viewed as a map from $\mathbb{R}^{n\times n}$ to $\mathbb{R}$. However, it does have a different sort of linearity property that we now describe.

It turns out that det *is* linear *as a function of individual columns*, while other columns are left fixed. For example, given vectors $\vec{b}, \vec{a}_1, \ldots, \vec{a}_n \in \mathbb{R}^n$ and $c \in \mathbb{R}$, we have

$$\det\begin{bmatrix} | & | & & | \\ \vec{a}_1+\vec{b} & \vec{a}_2 & \cdots & \vec{a}_n \\ | & | & & | \end{bmatrix} = \det\begin{bmatrix} | & | & & | \\ \vec{a}_1 & \vec{a}_2 & \cdots & \vec{a}_n \\ | & | & & | \end{bmatrix} + \det\begin{bmatrix} | & | & & | \\ \vec{b} & \vec{a}_2 & \cdots & \vec{a}_n \\ | & | & & | \end{bmatrix}$$

and

$$\det\begin{bmatrix} | & | & & | \\ c\vec{a}_1 & \vec{a}_2 & \cdots & \vec{a}_n \\ | & | & & | \end{bmatrix} = c\cdot\det\begin{bmatrix} | & | & & | \\ \vec{a}_1 & \vec{a}_2 & \cdots & \vec{a}_n \\ | & | & & | \end{bmatrix}.$$

We describe this by saying that the determinant function is *linear in the first column.* Similarly, the determinant function is linear in every other column as well, and also in every individual row. This means that given $A = [\vec{a}_1\ \vec{a}_2] \in \mathbb{R}^{2\times 2}$, $B = [\vec{b}_1\ \vec{b}_2] \in \mathbb{R}^{2\times 2}$, and $c \in \mathbb{R}$, we have

$$\begin{aligned} \det(A+B) &= \det\begin{bmatrix}\vec{a}_1+\vec{b}_1 & \vec{a}_2+\vec{b}_2\end{bmatrix} \\ &= \det\begin{bmatrix}\vec{a}_1 & \vec{a}_2+\vec{b}_2\end{bmatrix} + \det\begin{bmatrix}\vec{b}_1 & \vec{a}_2+\vec{b}_2\end{bmatrix} \\ &= \det\begin{bmatrix}\vec{a}_1 & \vec{a}_2\end{bmatrix} + \det\begin{bmatrix}\vec{a}_1 & \vec{b}_2\end{bmatrix} + \det\begin{bmatrix}\vec{b}_1 & \vec{a}_2\end{bmatrix} + \det\begin{bmatrix}\vec{b}_1 & \vec{b}_2\end{bmatrix} \end{aligned}$$

and

$$\det(cA) = \det\begin{bmatrix}c\vec{a}_1 & c\vec{a}_2\end{bmatrix} = c^2\cdot\det\begin{bmatrix}\vec{a}_1 & \vec{a}_2\end{bmatrix} = c^2\cdot\det(A).$$

More generally, for $c \in \mathbb{R}$ and $A \in \mathbb{R}^{n\times n}$, we see that $\det(cA) = c^n \cdot \det(A)$. Any function $f : \mathbb{R}^{n\times n} \to \mathbb{R}$ that is "linear in every column" is said to be *multilinear.* Thus for each n, the determinant function $\det : \mathbb{R}^{n\times n} \to \mathbb{R}$ is multilinear even though it is *not* linear. The determinant function $\mathbb{R}^{2\times 2} \to \mathbb{R}$ is an **alternating** multi-linear function. This means that if B is obtained from A by swapping any two columns, then $\det B = -\det A$.

An approach to determinants often taken in more advanced linear algebra course is to study multi-linear functions in general, and then show that there is a *unique* alternating n-multi-linear function $\mathbb{R}^{n\times n} \to \mathbb{R}$ such that $I_n \in \mathbb{R}^{n\times n}$ is sent to $1 \in \mathbb{R}$. This unique alternating n-multi-linear function is then defined to be the **determinant**. It assigns to each matrix $A \in \mathbb{R}^{n\times n}$, some scalar $\det A \in \mathbb{R}$.

* * *

We conclude this handout with a summary of properties of determinants that you will be expected to know. Not all of these properties were discussed above, so please read carefully through the entire list!

Properties of Determinants

- For any square matrix A, $\det A = 0$ if and only if A is *not* invertible.
- The determinant of a triangular matrix (upper or lower) is just the product of its diagonal entries.
- $\det(A^\top) = \det(A)$ for any square matrix A.
- For any column vectors $\vec{a}_1, \ldots, \vec{a}_n, \vec{b}_i \in \mathbb{R}^m$

  $$\det \begin{bmatrix} \vec{a}_1 & \cdots & \vec{a}_{i-1} & \vec{a}_i + \vec{b}_i & \vec{a}_{i+1} & \cdots & \vec{a}_n \end{bmatrix} = \det \begin{bmatrix} \vec{a}_1 & \cdots & \vec{a}_n \end{bmatrix} + \det \begin{bmatrix} \vec{a}_1 & \cdots & \vec{a}_{i-1} & \vec{b}_i & \vec{a}_{i+1} & \cdots & \vec{a}_n \end{bmatrix}$$

- For any column vectors $\vec{a}_1, \ldots, \vec{a}_n \in \mathbb{R}^m$ and $k \in \mathbb{R}$,

  $$\det \begin{bmatrix} \vec{a}_1 & \cdots & \vec{a}_{i-1} & k\vec{a}_i & \vec{a}_{i+1} & \cdots & \vec{a}_n \end{bmatrix} = k \cdot \det \begin{bmatrix} \vec{a}_1 & \cdots & \vec{a}_n \end{bmatrix}$$

- The fancy word for the previous two properties is "multilinear." So the determinant function is "multilinear" but it is *not* linear as a mapping from $\mathbb{R}^{n\times n}$ to $\mathbb{R}$, since in general

  $$\det(A+B) \neq \det(A) + \det(B) \qquad \text{and} \qquad \det(kA) \neq k \cdot \det(A).$$

- If A is an $n \times n$ matrix and $k \in \mathbb{R}$, then $\det(kA) = k^n \cdot \det(A)$.
- The determinant function is alternating in the columns and in the rows: swapping two columns (or two rows) changes the sign of the determinant.
- The determinant function is "multiplicative," i.e.,

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

  for all square matrices A and B of the same size.
- If A is invertible, then $\det(A^{-1}) = \dfrac{1}{\det A}$.
- If $[\vec{a} \;\; \vec{b}]$ is 2×2, then $|\det A|$ is the area of the parallelogram determined by $\vec{a}$ and $\vec{b}$, and if $A = [\vec{a} \;\; \vec{b} \;\; \vec{c}]$ is 3×3, then $|\det A|$ is the volume of the parallelepiped determined by $\vec{a}$, $\vec{b}$, and $\vec{c}$. (This generalizes to higher dimensions, once you define "n-dimensional volume.")
- Similar matrices have the same determinant, which allows us to define the determinant of a linear transformation $T : \mathbb{R}^n \to \mathbb{R}^n$ to be $\det[T]_{\mathcal{B}}$ where $\mathcal{B}$ is any basis of $\mathbb{R}^n$.
- The determinant of an orthogonal matrix is ± 1.

Techniques for computing determinants

- Know how to compute determinants using Laplace expansions. Remember that all Laplace expansions give the same value, so you should choose rows or columns with lots of zeros.
- Know how elementary row operations affect determinants, and how to simplify computing determinants using elementary row operations.