- M9227 - Week 1: Linear Algebra -

1 Vectors

Linear algebra is the art of manipulating vectors and matrices. A **vector** v is an object that we can write as a list of n numbers (for some integer n). The number n is called the **dimension** of the vector. We can add vectors by just adding their entries individually. For instance, $(1,2)+(3,4) = (4,6)$ or $(1,1,1)+(-1,.2,\pi) = (0,1.2,1+\pi)$. We can also multiply a number a by a vector v by multiplying each entry in v by a. For instance, if $v = (1,2,3)$ and $a = 3$, then $av = (3,6,9)$. Vectors are really supposed to represent points in space, as shown in the example below.

Example 1.1. *If we have a particle in* 3*d space, we can specify its position in terms of* 3 *numbers by specifying an* x*-axis, a* y*-axis and a* z*-axis. So we can specify the position as a vector* (a,b,c) *with* a *the* x*-position,* b *the* y*-position and* c *the* z*-position.*

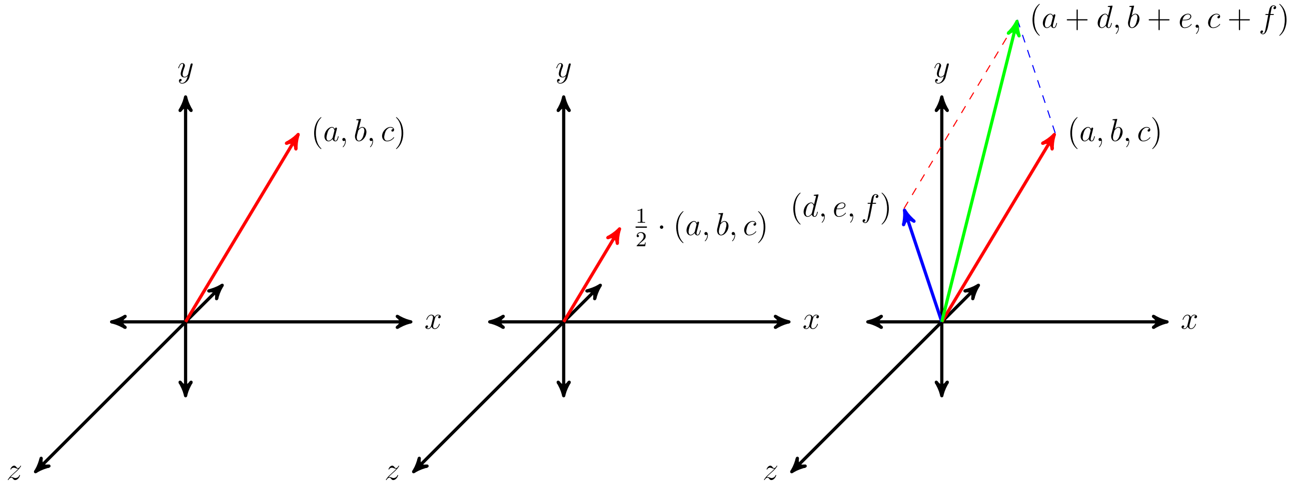

Figure 1: A 3-d particle as a vector. The position of the particle, (a,b,c), is at the end of the arrow.

We call the set of all the n-dimensional vectors $\mathbb{R}^n$.

2 Linear Map & Matrices

Linear map is maps on vectors which preserves the algebra. Maybe this sentence looks scary because you don't know what we mean by "map" or "preserve the algebra". So let's break it down:

- By **map**, we mean a rule for taking one vector and getting a new vector from the old one.
- For example, the rule that vector (a,b,c) goes to (a^2,b^2,c^2) is a map. If you give this map the vector $(1,2,3)$, then the map gives you $(1,4,9)$. Another example is the map $(a,b,c) \mapsto (a+1,b+2,c+3)$ (here $\mapsto$ means "goes to"). What does $(0,0,0)$ go to under that map? Yet another example is the map $(a,b) \mapsto (b,a)$. What does $(3,2)$ map to under that map?

- When we define the map $(a,b,c) \mapsto$[something], it is usually convenient to give the map a name instead of repeatedly saying "that map". We can name the map A, for instance. Then if we have a vector v, the map A takes v and gives me something new, v'. We can denote v' as Av.

- For example, suppose we define the map $(a,b) \mapsto (b+a, a)$ and we call the map A. Then if we define $v = (1,2)$ we can see that $Av = (2+1, 1) = (3,1)$ by just applying the rule. Now you try. Suppose $w = (3,4)$. What is Aw?

- Now that we have a notion of what a map is, we can say what **preserves the algebra** means. This is **very important**, so read it a few times if it doesn't make sense. As we stated in the first section, vectors can be added and scaled (multiplied by a number). This is the algebra of vectors, linear algebra. A map A **preserves** this algebra when:

$$A(v+w) = Av + Aw \text{ and } A(\lambda v) = \lambda Av$$

  In other words, sums and scaling just pass right through these maps. We call these linear maps. These maps are the central objects of study in linear algebra, and are ubiquitous in mathematics.

- Here's an example. The map B defined as $(a,b) \mapsto (b,a)$ is linear. We can just prove it directly:

$$B[(a,b)+(c,d)] = B[(a+c, b+d) = (b+d, a+c) = (b,a)+(d,c) = B(a,b) + B(c,d)$$

$$B(\lambda(a,b)) = B(\lambda a, \lambda b) = (\lambda b, \lambda a) = \lambda(b,a) = \lambda B(a,b)$$

  Do you see how we just checked that this map obeyed the rules of a linear map? Checking things is important! Now check that the map C defined by $(a,b,c) \mapsto (2a, -b, a+c)$ is linear. What about the map $(a,b,c) \mapsto (a,b,c+1)$? What about the map $(a,b,c) \mapsto (a^2, b^2, c^2)$?

A linear map can always be written as a matrix. A matrix is a square table like so:

$$\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}; \begin{bmatrix} 0 & -1 & 2 \\ 1 & 0 & 4 \\ 5 & -\pi & 4.3 \end{bmatrix}; \begin{bmatrix} 10^4 & .2 & 8 \\ 7.\bar{8} & \hbar & 20 \end{bmatrix}$$

They don't need to be square. When they're not square, they represent linear maps from a space of n-dimensions to a space of m-dimensions where $m \neq n$. We won't talk too much about those here. Given a vector $v = (v_1, v_2, \ldots, v_n)$ and a matrix A we can apply a matrix to it as so:

$$Av = \begin{bmatrix} a_{1,1} & a_{2,1} & \ldots & a_{n-1,1} & a_{n,1} \\ a_{1,2} & a_{2,2} & \ldots & a_{n-1,2} & a_{n,2} \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ a_{1,n-1} & a_{2,n-1} & \ldots & a_{n-1,n-1} & a_{n,n-1} \\ a_{1,n} & a_{2,n} & \ldots & a_{n-1,n} & a_{n,n} \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ \ldots \\ v_{n-1} \\ v_n \end{bmatrix} = \begin{bmatrix} \sum_i a_{1,i} v_i \\ \sum_i a_{2,i} v_i \\ \ldots \\ \sum_i a_{n-1,i} v_i \\ \sum_i a_{n,i} v_i \end{bmatrix}$$

How can we prove that any linear map can be written this way? Well, take a linear map T. Then it takes the unit vector e_i (which has 1 in the ith spot and 0 elsewhere, for example, $e_1 = (1,0,0,\ldots)$) to some vector Te_i. Then the matrix for T can be written as the matrix with the ith column equal to the vector $Te_i = \sum_i a_{j,i} e_j$. Then since every vector v can be expressed as a sum of e_i's, maybe $v = \sum_i v_i e_i$, then:

$$Tv = \sum_i v_i T e_i = \sum_i v_i a_{j,i} e_j = \begin{bmatrix} a_{1,1} & a_{2,1} & \dots & a_{n-1,1} & a_{n,1} \\ a_{1,2} & a_{2,2} & \dots & a_{n-1,2} & a_{n,2} \\ \dots & \dots & \dots & \dots & \dots \\ a_{1,n-1} & a_{2,n-1} & \dots & a_{n-1,n-1} & a_{n,n-1} \\ a_{1,n} & a_{2,n} & \dots & a_{n-1,n} & a_{n,n} \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ \dots \\ v_{n-1} \\ v_n \end{bmatrix} = Av$$

So every linear map can be expressed in terms of a matrix when we express a vector v as a sum of the unit vectors e_i. The important thing about e_i here is the following; every vector v can be expressed uniquely as a sum of the vectors e_i. Any other set of vectors b_i for which this is true is called a **basis**. You can learn more about bases in the homework. They're very important.

3 The Determinant

There are many important quantities associated to a matrix M. Perhaps one of the most important is the **determinant** of M, written $\det(M)$. The determinant is simple to explain, with one caveat which we will discuss a little later.

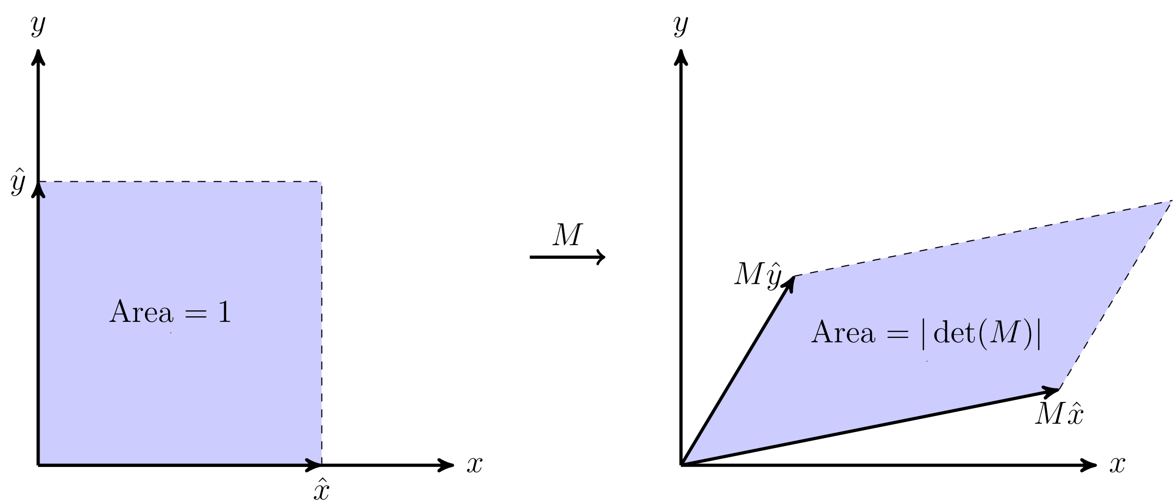

Figure 2: A unit square transforms as so under M.

In short, the determinant measures the deformation in *area* caused by the matrix. To see what this means, consider the figure above. On the left, we depict a unit[1] with sides given by vectors $\hat{x} = (1, 0)$ and $\hat{y} = (0, 1)$. On the right, we depict the parallelogram which the square with these sides maps to under the linear transformation given by the matrix M. In other words, when we apply the matrix M to the sides $\hat{x}$ and $\hat{y}$ of the square, we get new vectors $M\hat{x}$ and $M\hat{y}$. These vectors are the sides of the parallelogram on the right.

The area of this parallelogram is equal to the *absolute value* of the determinant (thus the bars in the diagram above). This is the caveat that we were talking about, and we'll explain how we know what sign the determinant will be below. First, let's find a formula for the determinant.

[1] A **unit** square is defined as a square that has area 1.

Proposition 3.1. *Suppose we have a matrix M, given by:*

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

Then $|\det(M)| = |ad - bc|$.

Proof. We do a proof by picture, which is the best kind. Before we start, we observe that:

$$M\hat{x} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} a \\ c \end{pmatrix}; M\hat{y} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} b \\ d \end{pmatrix}$$

So the vector $M\hat{x}$ is just the first column of the matrix and $M\hat{y}$ is just the second column.

Now we're going to prove a special case of this proposition; we're going to assume that $M\hat{x}$ and $M\hat{y}$ are in the first quadrant of the plane, and that $\tan^{-1}(c/a) < \tan^{-1}(b/d)$. This lets us draw the following picture:

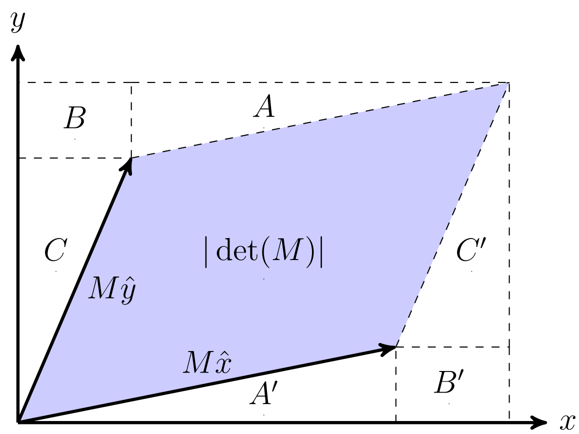

Figure 3: The picture.

With this picture, we can just use geometry that Greek's knew about to compute the area of the shaded parallelogram. The area of the big rectangle (made out of all of the shapes) is just the product of its side lengths, $a+b$ and $c+d$. The area of A is the same as that of A', and likewise with B and B', and C and C'. The area of A is $\frac{1}{2}ac$ (the side lengths are just the x and y components of $M\hat{x}$). The area of C is $\frac{1}{2}bd$. Finally, the area of B is bc. So the area of the inner parallelogram is:

$$|\det(M)| = \text{Area(Big Rectangle)} - 2\text{Area}(A) - 2\text{Area}(B) - 2\text{Area}(C)$$

$$= (a+c)(b+d) - ab - cd - 2bc = ad - bc$$

This proves the statement for our special case! Now we reduce every other case to the one we just proved. If we just take an arbitrary matrix $M = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, we can do some rotations and reflections to get new matrix $M' = \begin{pmatrix} a' & b' \\ c' & d' \end{pmatrix}$. This just amounts to multiplying M by reflection matrices and rotation matrices (rotations by 90 degrees) which can be expressed as:

$$R = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}; U = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}; V = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

R is rotation around 90 degrees counter-clockwise, U is reflection across the x-axis and V is reflection across the y-axis. Notice that when we multiply a matrix M by R, U or V on the right, it doesn't change $|ad - bc|$. In other words, it only changes the determinant formula that we listed above by a sign. Furthermore, enough of these rotations and reflections will cause the assumptions of our special case to be satisfied. This you should check. □

Now a brief aside. You might ask "So now we know how cubes transform. But cubes are boring. What about a general shape!?!". Well, first we would ask you to clarify what you meant by "general shape." Then we would stop being pedantic and answer your question.

Proposition 3.2. *If S is a solid shape sitting in n-dimensional space, then the volume of S under a linear transformation (under operation by a matrix M) is just multiplied by $|\det(M)|$.*

Proof. Here we will give an intuitive proof that has a rigorous version. Since the rigorous version requires limits (calculus!), we leave it to the more advanced students to fill in the gaps if they wish. Suppose that we have such a solid shape. Then we can approximate the shape by a bunch of cubes of some small volume ϵ.

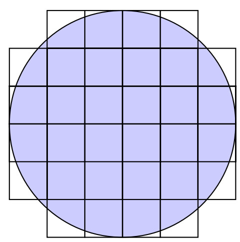

Figure 4: A lil' cube approximation of S. In this case S is a ball

When we apply the linear transformation M to the n-dimensional space that S sits inside of (getting a new shape, MS), each of the cubes in the approximation transforms to a higher dimensional parallelogram of volume $\epsilon \det(M)$. Furthermore, the set of all the little parallelograms forms a close approximation of the new shape MS. So we can argue that the volume of the shape MS must be pretty close to $\det(M) \cdot \text{volume}(S)$, with the closeness determined by how small ϵ is (i.e how tiny our original little cubes were).

If we take our cubes smaller and smaller (taking the "limit as ϵ goes to 0"), we get a better and better approximation of the volume of MS, and in fact we see that this approximation approaches the exact number $\det(M) \cdot \text{volume}(S)$. So, in fact, the volume of MS is exactly $\det(M) \cdot \text{volume}(S)$. □

4 Properties Of The Determinant

The determinant has some important properties that we will use through out this course (and that you will certainly use later in your life if you do more math). The first one is a multiplication rule.

Proposition 4.1. *Let A and B be matrices. Then* $\det(AB) = \det(A)\det(B)$.

In other words, the determinant of a product of matrices is the product of their determinants.

Proof. We only know the determinant in $2d$-case, so we'll prove it in that case. It's a little ugly, but not that ugly.

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}; B = \begin{pmatrix} e & f \\ g & h \end{pmatrix}; AB = \begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} e & f \\ g & h \end{pmatrix} = \begin{pmatrix} ae+bg & ce+dg \\ af+bh & cf+dh \end{pmatrix}$$

$$\det(AB) = (ae+bg)(cf+dh) - (af+bh)(ce+dg) =$$

$$acef + adeh + bcfg + bdgh - acef - adfg - bceh - bdgh = adeh + bcfg - adfg - bceh =$$

$$(ad-bc)(eh-fg) = \det(A)\det(B)$$

This multiplication formula also can also basically be proven with Proposition 3.2 (and in higher dimensions!). To see this, let S be the unit cube, and let A, B be two matrices. Then Proposition 3.1 implies that the volume of AS is $|\det(A)|$. Then Proposition 3.2 implies that the shape BAS (the image of the parallelogram AS under B) has volume $|\det(B)|\text{volume}(AS) = |\det(B)||\det(A)|$. But this is also just the volume of the image of the unit square under the matrix BA, which is $|\det(BA)|$. So $|\det(BA)| = |\det(B)||\det(A)|$. $\square$

The multiplication rule has a bunch of very useful corollaries.

Corollary 4.1. *The following properties of the determinant hold.*

1. $\det(Id) = 1$
2. $\det(A)\det(A^{-1}) = 1$
3. $\det(BAB^{-1}) = \det(A)$

Proof. (1) is obvious geometrically speaking: the area and orientation of the unit cube are left unchanged by the identity map. (2) then follows from the product formula, since $\det(A)\det(A^{-1}) = \det(AA^{-1}) = \det(Id) = 1$. (3) is similar. Again, we just use the fact that $\det(AB) = \det(A)\det(B)$ to say that:

$$\det(BAB^{-1}) = \det(B)\det(AB^{-1}) = \det(B)\det(A)\det(B^{-1}) = \det(B)\det(B^{-1})\det(A) = \det(A)$$

$\square$

These are pretty handy formulae. Furthermore, it shows that there's a relationship between invertibility and the determinant.

Proposition 4.2. *Let M be a square matrix. Then if* $\det(M) \neq 0$ *then M has an inverse. Likewise, if M has an inverse then* $\det(M) \neq 0$.

This means that $\det(M) \neq 0$ and M being invertible are equivalent conditions. In this situation, we say that $\det(M) \neq 0$ **if and only if** M is invertible.

Proof. First we prove that if M is invertible then $\det(M) \neq 0$. We do this with a **proof by contradiction**. A proof like this always starts with "suppose our statement were false" and the derives something impossible. In this case our impossible statement will be $0 = 1$. Very impossible indeed!

So here we go. Suppose that our statement were false, so we found an invertible matrix M with determinant 0. Then $\det(M)\det(M^{-1}) = 0\det(M^{-1} = 0$. But by Corollary (4.1), wwe also know that $\det(M)\det(M^{-1}) = 1$. So $0 = 1$. That's impossible, so an invertible matrix must have non-zero determinant.

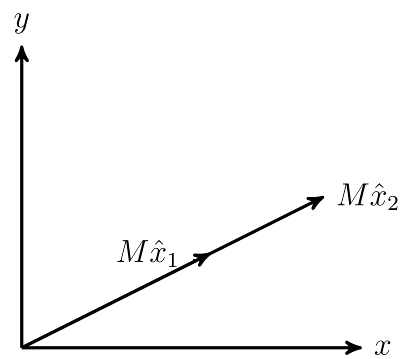

Figure 5: When the volume is 0 (like in this picture), $M\hat{x}_1$ and $M\hat{x}_2$ are linearly dependent.

That was easy! Now let's go the other way. Suppose a matrix M has $\det(M) \neq 0$. Then the volume of the parallelogram with sides given by the vectors $M\hat{x}_1, M\hat{x}_2, \ldots, M\hat{x}_n$ is non-zero. This means that these vectors cannot be linear combinations of each other, since if they were then they would be coplanar and then the columns of the matrix would be linearly dependent! This is illustrated pictorially above. $\square$

The determinant is even involved in the most well-known formula for the inverse of matrix M^{-1} in terms of M's entries. In the 2-dimensional case, we can figure this out with a little guessing. Suppose we had a matrix $M = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ for which $\det(M) \neq 0$. We want to mix the entries up to get something close to the identity. One thing to note is that $(a, c) \cdot (c, -a) = 0$. These are 90 degree rotations of each other. With this in mind we can try $M' = \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$. Notice how the second row of M' is orthogonal to the first column of M and the first row of M' is orthogonal to the second column of M. So MM' is diagonal, which is a good start. If we actually compute MM', we see that:

$$MM' = \begin{pmatrix} ad - bc & 0 \\ 0 & ad - bc \end{pmatrix}$$

Hmm, that thing in on the diagonal looks familiar. If we divide the top formula It's the determinant! In other words, we have:

$$M^{-1} = \det(M)^{-1}M' = \det(M)^{-1}\text{cofactor}(M)$$

$M' = \text{cofactor}(M)$ is called the **cofactor matrix**. Google its general definition if you're interested.