Dot Products and Vector Projections

Jonathan Mui

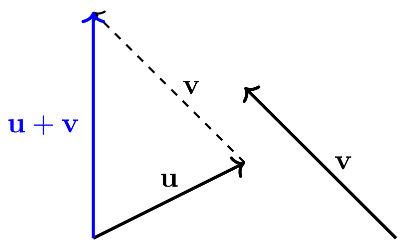
January 2020

1 Review of vector operations

Let's review some basic operations we can do with vectors.

(i) Addition

To add vectors geometrically, we use the **tip-to-tail rule**. Consider the vectors $\mathbf{u}$ and $\mathbf{v}$ below. To construct $\mathbf{u}+\mathbf{v}$, we move the tail of $\mathbf{v}$ to the tip of $\mathbf{u}$, and then connect the tail of $\mathbf{u}$ to the new tip:



To add vectors algebraically, we simply add their coordinates. For example, if $\mathbf{u} = (1, 2, 3)$ and $\mathbf{v} = (-3, 0, 1)$, then

$$\mathbf{u} + \mathbf{v} = (1, 2, 3) + (-3, 0, 1) = (-2, 2, 4)$$

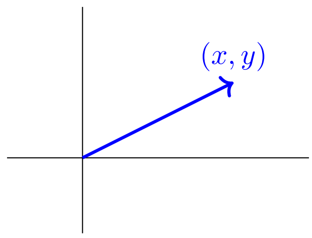
It is easy to show that $\mathbf{u}+\mathbf{v} = \mathbf{v}+\mathbf{u}$ both geometrically and algebraically (try it!). We say that vector addition is **commutative**, just like the addition of real numbers. Moreover, if we have 3 vectors $\mathbf{u}, \mathbf{v}$ and $\mathbf{w}$, see if you can show that

$$(\mathbf{u} + \mathbf{v}) + \mathbf{w} = \mathbf{u} + (\mathbf{v} + \mathbf{w})$$

This property is known as **associativity**, and we say that vector addition is **associative**. When this is the case, we can drop brackets and write $\mathbf{u} + \mathbf{v} + \mathbf{w}$ without ambiguity (just like addition or real numbers). This is in contrast to division of real numbers, for example, where an expression such as $24 \div 2 \div 4$ does not make any sense.

(ii) Magnitude

We may think of a vector concretely as an arrow based at the origin of our coordinate system, and ending at the point described by its coordinates. Here is a picture of the 2D vector $\mathbf{u} = (x, y)$.



Thus the length of the arrow is the same as the distance from the point (x, y) to the origin. By Pythagoras' theorem, this means that the length or **magnitude** of the vector $\mathbf{u}$ is

$$\|\mathbf{u}\| = \sqrt{x^2 + y^2}$$

(In these notes, we consistently denote the length of a vector with double-bars $\|\cdot\|$. This is standard notation in most mathematical literature, although NESA does not use it). This definition extends without any issues to higher dimensions, so for a 3D vector $\mathbf{u} = (x, y, z)$, we have

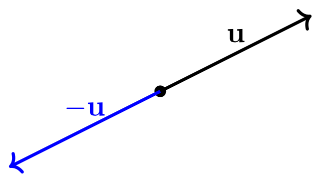
$$\|\mathbf{u}\| = \sqrt{x^2 + y^2 + z^2}$$

(iii) Scaling

Intuitively, scaling is the "stretching" or "squashing" of a vector. However, we can scale vectors by any real number $\lambda \in \mathbb{R}$. For a vector $\mathbf{u} = (x, y, z)$, we define scaling algebraically by

$$\lambda\mathbf{u} = (\lambda x, \lambda y, \lambda z)$$

Notice that $\|\lambda\mathbf{u}\| = |\lambda|\|\mathbf{u}\|$ – you can use the formula for the magnitude of a vector to prove this statement. Consequently, the length of the rescaled vector is $|\lambda|$ times the original one. When $\lambda < 0$, the "direction" of the arrow is flipped:



Thus when we "subtract" two vectors $\mathbf{u} - \mathbf{v}$, what we are actually doing is adding $-\mathbf{v}$ to $\mathbf{u}$. This is why there is no need to define subtraction for vectors separately!

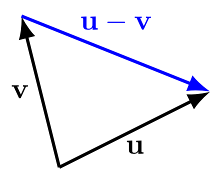
A couple more important definitions:

- A **unit vector** is a vector with length equal to 1. Any non-zero vector $\mathbf{u}$ can be turned into a unit vector $\widehat{\mathbf{u}}$ via rescaling by the length:

$$\widehat{\mathbf{u}} = \frac{\mathbf{u}}{\|\mathbf{u}\|}$$

- Two vectors are said to be **parallel** if one is a scalar multiple of the other.

We conclude this section by reminding ourselves of the geometric significance of the difference between two vectors. Given two vectors $\mathbf{u}$ and $\mathbf{v}$, the expression $\mathbf{u} - \mathbf{v}$ represents the vector that *points from the tip of* $\mathbf{v}$ *to the tip of* $\mathbf{u}$.



You may like to remember this rule as "target minus source": $\mathbf{u}$ is the target, $\mathbf{v}$ is the source, so $\mathbf{u} - \mathbf{v}$ points from the tip of $\mathbf{v}$ to the tip of $\mathbf{u}$.

2 Algebraic definition of the dot product

There is no natural notion of a multiplication for vectors. However, we can define the **dot product**. The product of two real numbers is another real number, but the dot product of two vectors is in fact also a *real number*, and *not* a vector. This explains why this type of product is also called a **scalar product**.

Let $\mathbf{u} = (u_1, u_2, u_3)$ and $\mathbf{v} = (v_1, v_2, v_3)$ be any two 3D vectors. Then

$$\boxed{\mathbf{u} \cdot \mathbf{v} = u_1 v_1 + u_2 v_2 + u_3 v_3}$$

This definition can be easily extended to any number of dimensions, but for the HSC, 3 dimensions will suffice. Notice the dot $\cdot$ in between the vector symbols. It is important to realise that the expression $\mathbf{uv}$ does not make sense!

We collect some properties of the dot product below. We state these without proof, although you are encouraged to construct the proofs yourself (they are very easy).

Properties of the dot product

(i) $\mathbf{u} \cdot \mathbf{v} = \mathbf{v} \cdot \mathbf{u}$ (i.e. the dot product is **commutative**)

(ii) $\mathbf{u} \cdot \mathbf{u} = \|\mathbf{u}\|^2$

(iii) $\lambda(\mathbf{u} \cdot \mathbf{v}) = (\lambda \mathbf{u}) \cdot \mathbf{v} = \mathbf{u} \cdot (\lambda \mathbf{v})$ for all $\lambda \in \mathbb{R}$

(iv) $\mathbf{u} \cdot (\mathbf{v} + \mathbf{w}) = \mathbf{u} \cdot \mathbf{v} + \mathbf{u} \cdot \mathbf{w}$ (distributive law)

Exercise 1. Construct 3 pairs of 3D vectors, such that their dot products are respectively

(i) positive (> 0)

(ii) equal to 0, and

(iii) negative (< 0)

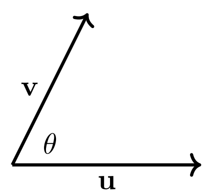
Although the dot product is very easy to calculate, you may wonder what is the motivation behind its definition. How do we know that such an operation is reasonable, let alone useful? To answer this question, we need to study vector projections.

3 Vector projections

3.1 Motivation for the projection formula

In mathematics, we are often concerned with the relationship between different elements in some family of objects. The dot product arises naturally when we ask: given a pair of vectors, how can we describe the relationship between them? This description should contain some geometric information, as well as being easy to use algebraically. For now, let's forget temporarily the definition of the dot product we have just encountered, and look at things from the geometric perspective.

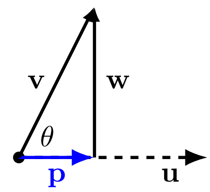
Given two non-zero vectors $\mathbf{u}$ and $\mathbf{v}$, it may have already occurred to you to describe their relationship by considering the angle between them:



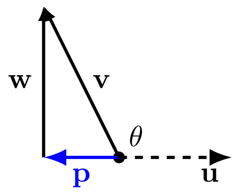
Therefore we should try to get an expression that involves both vectors and the angle. A promising idea is to write $\mathbf{v}$ in terms of $\mathbf{u}$, that is, we want

$$\mathbf{v} = k\mathbf{u} + \mathbf{w}$$

for some constant $k \in \mathbb{R}$ and vector $\mathbf{w}$, both of which are unknown for now. Looking at the picture above, it seems reasonable to consider the *projection of* $\mathbf{v}$ *onto* $\mathbf{u}$. In the following picture, the projection is labelled $\mathbf{p}$:



If the angle between the vectors is obtuse, the situation looks like this:



In either case, $\mathbf{p}$ is a vector parallel to $\mathbf{u}$, i.e. it is a scalar multiple of $\mathbf{u}$. In fact, $\mathbf{p}$ is a scalar multiple of any non-zero vector that is parallel to $\mathbf{u}$. We now make the special choice of $\widehat{\mathbf{u}}$ – the unit vector in the direction of $\mathbf{u}$ – and will therefore write

$$\mathbf{p} = k\widehat{\mathbf{u}} = k\frac{\mathbf{u}}{\|\mathbf{u}\|}$$

Notice that $\|\mathbf{p}\| = \|k\widehat{\mathbf{u}}\| = |k|$. Now we analyse two cases.

Case 1: θ acute

In this case $k > 0$, so $\|\mathbf{p}\| = |k| = k$. On the other hand, trigonometry tells us that

$$\|\mathbf{p}\| = \|\mathbf{v}\| \cos\theta$$

Therefore $k = \|\mathbf{v}\| \cos\theta$ and

$$\mathbf{p} = \frac{\|\mathbf{v}\| \cos\theta}{\|\mathbf{u}\|}\mathbf{u}$$

Case 2: θ obtuse

In this case, $k < 0$, and hence $\|\mathbf{p}\| = |k| = -k$. On the other hand,

$$\|\mathbf{p}\| = \|\mathbf{v}\| \cos(\pi - \theta) = -\|\mathbf{v}\| \cos\theta$$

Therefore $-k = -\|\mathbf{v}\| \cos\theta$, but this gives exactly the same conclusion as Case 1.

Thus in both cases, we have obtained the expression

$$\boxed{\mathbf{p} = \frac{\|\mathbf{v}\| \cos\theta}{\|\mathbf{u}\|}\mathbf{u}}$$

From now on, we denote the vector $\mathbf{p}$ by

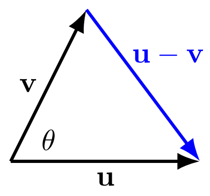
$$\mathrm{Proj}_{\mathbf{u}}(\mathbf{v})$$

This symbol is read as "the projection of $\mathbf{v}$ onto $\mathbf{u}$".

3.2 Definition of the geometric dot product

We have just developed a formula for $\text{Proj}_{\mathbf{u}}(\mathbf{v})$. However, it still involves the quantity $\cos\theta$, which is not so convenient. Perhaps there is a way to express $\cos\theta$ in terms of $\mathbf{u}$ and $\mathbf{v}$?

Once again, trigonometry helps us out. Consider the triangle formed by the vectors $\mathbf{u}, \mathbf{v}$ and $\mathbf{u} - \mathbf{v}$.



By the cosine rule, we obtain

$$\|\mathbf{u} - \mathbf{v}\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2 - 2\|\mathbf{u}\|\|\mathbf{v}\|\cos\theta \qquad (\star)$$

Notice that the term $\|\mathbf{u}\|\|\mathbf{v}\|\cos\theta$ involves the product of the lengths of $\mathbf{u}$ and $\mathbf{v}$ and the angle between them. We therefore define the **geometric dot product** to be

$$\boxed{\mathbf{u} \cdot \mathbf{v} = \|\mathbf{u}\|\|\mathbf{v}\|\cos\theta}$$

We have used the same symbol as the algebraic dot product, but we have not proved that these definitions are the same thing! Let us address this problem now.

Theorem 1. *The algebraic and geometric dot products coincide, i.e.*

$$\boxed{\mathbf{u} \cdot \mathbf{v} = \|\mathbf{u}\|\|\mathbf{v}\|\cos\theta = u_1v_1 + u_2v_2 + u_3v_3}$$

for all vectors $\mathbf{u} = (u_1, u_2, u_3)$ *and* $\mathbf{v} = (v_1, v_2, v_3)$.

Proof. By rearranging equation $(\star)$, we obtain

$$\|\mathbf{u}\|\|\mathbf{v}\|\cos\theta = \frac{1}{2}(\|\mathbf{u}\|^2 + \|\mathbf{v}\|^2 - \|\mathbf{u} - \mathbf{v}\|^2)$$

Using the length formula for vectors (a.k.a Pythagoras' theorem), the right hand side of the above equation becomes

$$\begin{aligned}
&\frac{1}{2}[(u_1^2 + u_2^2 + u_3^2) + (v_1^2 + v_2^2 + v_3^2) - ((u_1 - v_1)^2 + (u_2 - v_2)^2 + (u_3 - v_3)^2)] \\
&\qquad = \frac{1}{2}[2u_1v_1 + 2u_2v_2 + 2u_3v_3] \\
&\qquad = u_1v_1 + u_2v_2 + u_3v_3
\end{aligned}$$

as required. □

Although we have stated the theorem for 3D vectors, it generalises easily to n-dimensional vectors (you may want to try!).

We can now go back to the projection formula and rewrite it in a more practical way. Recall that $\text{Proj}_{\mathbf{u}}(\mathbf{v}) = \frac{\|\mathbf{v}\|\cos\theta}{\|\mathbf{u}\|}\mathbf{u}$, but using what we have just proved, we can write

$$\frac{\|\mathbf{v}\|\cos\theta}{\|\mathbf{u}\|}\mathbf{u} = \frac{\|\mathbf{u}\|\|\mathbf{v}\|\cos\theta}{\|\mathbf{u}\|^2}\mathbf{u} = \frac{\mathbf{u} \cdot \mathbf{v}}{\|\mathbf{u}\|^2}\mathbf{u}$$

Therefore

$$\boxed{\text{Proj}_{\mathbf{u}}(\mathbf{v}) = \frac{\mathbf{u} \cdot \mathbf{v}}{\|\mathbf{u}\|^2}\mathbf{u}}$$

Observe that the projection is another *vector*, not a scalar.

Exercise 2. Prove that

$$\text{Proj}_{\lambda \mathbf{u}}(\mathbf{v}) = \text{Proj}_{\mathbf{u}}(\mathbf{v}) \qquad \text{for all } \lambda \neq 0$$

In other words, the length of the vector onto which we are projecting does not change the projection.

3.3 Orthogonal decomposition

From the formula $\mathbf{u} \cdot \mathbf{v} = \|\mathbf{u}\|\|\mathbf{v}\| \cos\theta$, we immediately deduce the following important fact:

The vectors $\mathbf{u}$ and $\mathbf{v}$ are orthogonal if and only if $\mathbf{u} \cdot \mathbf{v} = 0$

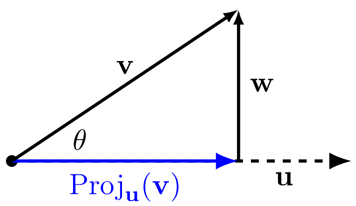
We now have all the tools to complete the task of expressing a vector $\mathbf{v}$ in terms of $\mathbf{u}$ and another component:

$$\mathbf{v} = k\mathbf{u} + \mathbf{w}$$

Using the results of Sections 3.1 and 3.2, we have determined that the scalar k is

$$k = \frac{\mathbf{u} \cdot \mathbf{v}}{\|\mathbf{u}\|^2}$$

so that $k\mathbf{u} = \text{Proj}_{\mathbf{u}}(\mathbf{v})$, the projection of $\mathbf{v}$ onto $\mathbf{u}$. If we revisit the following picture, it is now clear what $\mathbf{w}$ should be.



If we define

$$\boxed{\mathbf{w} = \mathbf{v} - \text{Proj}_{\mathbf{u}}(\mathbf{v})}$$

then by construction, $\mathbf{v} = \text{Proj}_{\mathbf{u}}(\mathbf{v}) + \mathbf{w} = k\mathbf{u} + \mathbf{w}$ as desired. We obtain the following important definition.

The vector $\mathbf{v} - \text{Proj}_{\mathbf{u}}(\mathbf{v})$ is called the **orthogonal complement** of the projection

Exercise 3. Verify directly using the dot product that the orthogonal complement is orthogonal to the projection.