**Sums of Squares (FNS 195-S)** **Fall 2014**
**Record of What We Did** **Drew Armstrong**

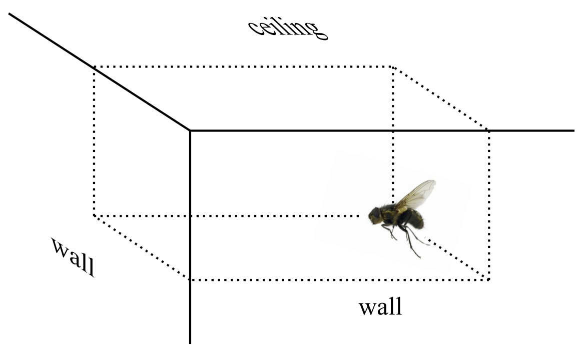
---

## Vectors

When we tried to apply Cartesian coordinates in 3 dimensions we ran into some difficulty tryiing to describe lines and planes. To describe lines and planes properly we will need to introduce the concept of "vectors". You may be familiar with vectors from physics, but how do we describe vectors mathematically?

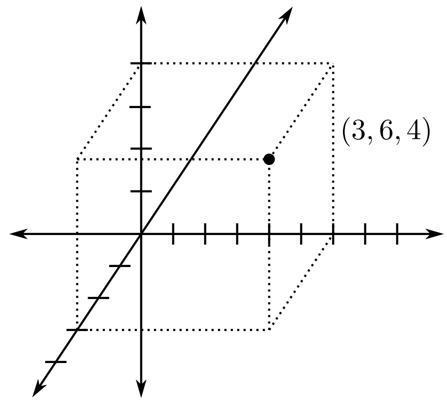
• How are 3-dimensional Cartesian coordinates defined?

Legend says that René Descartes was lying in bed one day, watching a fly buzz around in the corner. He realized that he could uniquely specify the position of the fly by recording its distance from the two walls and the ceiling.




For geometric reasons, the fly and the point on a given surface closest to the fly determine a line **perpendicular** to the surface. So the fly is at one vertex of a **rectangular box** whose opposite vertex is the corner. If the dimensions of the box are $x \times y \times z$ then we say the fly is at position $(x, y, z)$. Does it matter in which order we list these three numbers? Yes and no. The original choice of ordering is completely arbitrary, but after that we have to stick with it. (Just like Ben Franklin and the negative electron.)

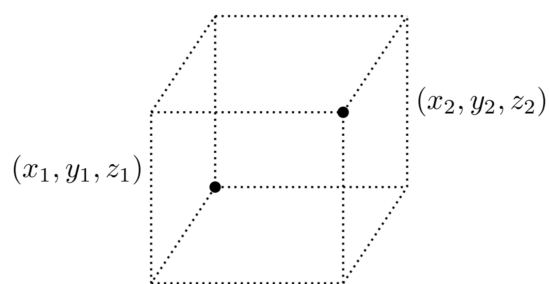
In more modern terms, we begin with three mutually perpendicular "axes" meeting at a point (called the **origin**). Each pair of axes determines a 2D Cartesian plane. There are three pairs of axes and we will say that they determine three "coordinate planes" (think of a ceiling and two walls). Given any point $P$ in space we consider the rectangular box whose opposite vertex is at the origin. If the dimensions of the box are $x \times y \times z$ then we say that $P = (x, y, z)$. (Unlike Descartes, we allow the dimensions of the box to be negative.)



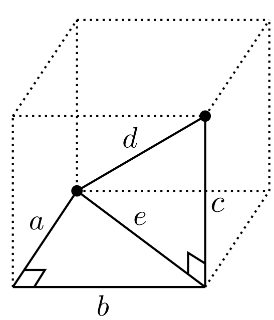

This picture shows a point in the "first octant" of $\mathbb{R}^3$, i.e., in which all three coordinates are positive. To display all the different possibilities of zero, positive, and negative coordinates we would need $3^3 = 27$ different looking pictures!

• What is the distance between points $(x_1, y_1, z_1)$ and $(x_2, y_2, z_2)$ in Cartesian space $\mathbb{R}^3$? Why? (This is a bit harder than you might think.)

When we computed the distance between points $(x_1, y_1)$ and $(x_2, y_2)$ in $\mathbb{R}^2$ we drew a triangle between them and used the Pythagorean Theorem. We'll do the same thing here, except we will need **two** triangles. The points $(x_1, y_1, z_1)$ and $(x_2, y_2, z_2)$ lie at opposite vertices of a rectangular box with faces parallel to the coordinate planes.




We can connect the points with two perpendicular right triangles like so:



The lengths $a, b, c$ are the dimensions of the box, so we have

$$\begin{aligned} a &= |x_2 - x_1| \\ b &= |y_2 - y_1| \\ c &= |z_2 - z_1|. \end{aligned}$$

I used absolute values so that the same formulas will work for all possible configurations of two points, not just the one in the picture. The length $d$ is the distance we want to compute and the length $e$ is just some number that we don't care about. Applying the PT to the first triangle gives

$$a^2 + b^2 = e^2$$

and then applying the PT to the second triangle gives

$$e^2 + c^2 = d^2.$$

Finally, putting the equations together gives

$$d^2 = e^2 + c^2 = a^2 + b^2 + c^2.$$

We conclude that the distance $d$ between the points $(x_1, y_1, z_1)$ and $(x_2, y_2, z_2)$ is satisfies

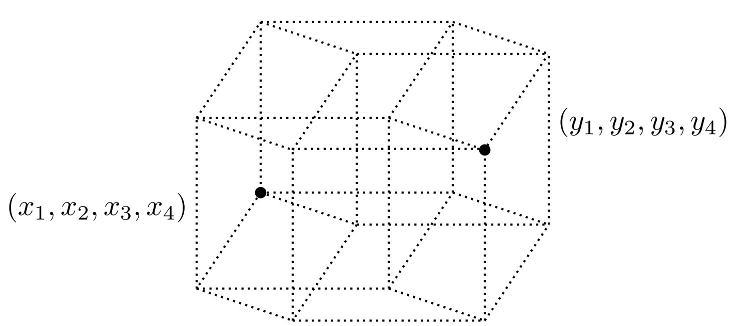
$$\boxed{d^2 = (x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2.}$$

Since we are squaring each quantity the absolute value signs become irrelevant. Once again, algebra is smarter than geometry. Discussion: Do the three coordinate planes in Cartesian space *need* to be perpendicular? They do if we want this nice formula to be true.

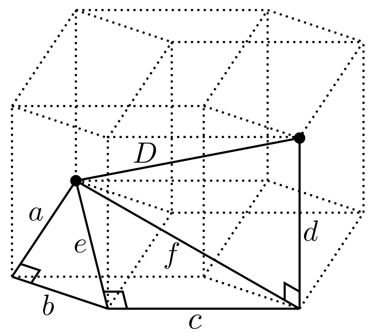
We might call the above formula the 3D Pythagorean Theorem. Here's something to think about: What is the distance between the points $(x_1, x_2, x_3, x_4)$ and $(y_1, y_2, y_3, y_4)$ in "4-dimensional Cartesian space" $\mathbb{R}^4$? There are at least two possible answers: (1) There is no such thing as "4-dimensional space" so we can define it however we want. We **choose** to say that the distance $D$ satisfies

$$\boxed{D^2 = (y_1 - x_1)^2 + (y_2 - x_2)^2 + (y_3 - x_3)^2 + (y_4 - x_4)^2}$$

because this formula has nice properties. (2) There **is** such a thing as "4-dimensional space" but we can't visualize it so we have to argue indirectly by projecting into lower dimensional space. Our two points are the opposite corners of a rectangular hyperbox. If we squash this hyperbox down onto a 2D plane it will look something like this:




Then we can connect the points with **three** perpendicular right triangles like so:




The angles don't look right, but that's just because we squashed everything into two dimensions. (Hey, the angles didn't look right in the 3D picture either and you didn't complain then!) Now we can apply the PT separately to the three right triangles to get

$$\begin{aligned} a^2 + b^2 &= e^2 \\ e^2 + c^2 &= f^2 \\ f^2 + d^2 &= D^2. \end{aligned}$$

Finally, putting these equations together gives

$$\begin{aligned} D^2 &= f^2 + d^2 \\ &= e^2 + c^2 + d^2 \\ &= a^2 + b^2 + c^2 + d^2. \end{aligned}$$

This is the same formula as before.

• What is the equation of the sphere with radius $r$ centered at $(a, b, c)$?

This one is easy now that we have the tools. Note that the point $(x, y, z)$ is on the sphere if and only if it has distance $r$ from the point $(a, b, c)$ (this is just the definition of a sphere), and by the above formula this happens if and only if

$$\boxed{(x - a)^2 + (y - b)^2 + (z - c)^2 = r^2.}$$

• What is the equation of a plane in $\mathbb{R}^3$? What is the equation of a line in $\mathbb{R}^3$. (This is quite a bit harder than you might think.)

Yes this *is* quite a bit harder. We don't have the tools necessary to answer this question right now so I'll just tell you that the equation of a general plane in $\mathbb{R}^3$ is

$$\boxed{ax + by + cz + d = 0,}$$

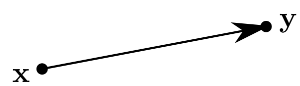
and we'll leave the proof until later. Asking for the equation of a line was a trick question because **there is no equation of a line in** $\mathbb{R}^3$! What I mean by this is that a line in $\mathbb{R}^3$ can not be determined by a single equation; we need at least two. And there is no *best* way to choose these two equations. So what is the best way to describe a line in $\mathbb{R}^3$?

The best way to describe both lines and planes in $\mathbb{R}^3$ is by using vectors.

• What is a "vector"?

In physics they say that a vector is a quantity that has both magnitude and direction. But what to they mean by "quantity", "magnitude", and "direction"? In mathematics we need a more precise definition.

For us, a **vector** is an ordered pair of points in $\mathbb{R}^n$. (Yes, we have now jumped to $n$-dimensional space. Why not?) Given points $\mathbf{x} = (x_1, x_2, \ldots, x_n)$ and $\mathbf{y} = (y_1, y_2, \ldots, y_n)$ in $\mathbb{R}^n$ we will think of the ordered pair $[\mathbf{x}, \mathbf{y}]$ as an "arrow" with "tail" at $\mathbf{x}$ and "head" at $\mathbf{y}$.




Yes, this is happening in $n$-dimensional space but a vector is a 1-dimensional thing so I can still draw it perfectly well.

• Find an algebraic formula for the **length** of the vector $[\mathbf{x}, \mathbf{y}]$.

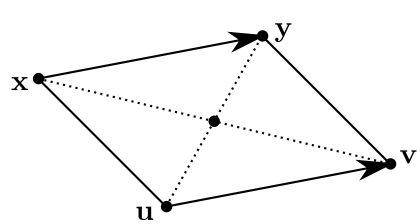
We will use the notation $\|[\mathbf{x}, \mathbf{y}]\|$ for the length of the vector $[\mathbf{x}, \mathbf{y}]$. This is the same thing as the distance between the points $\mathbf{x}$ and $\mathbf{y}$ in $\mathbb{R}^n$ so the above discussion suggests that we have the formula

$$\boxed{\|[\mathbf{x}, \mathbf{y}]\|^2 = (y_1 - x_1)^2 + (y_2 - x_2)^2 + \cdots + (y_n - x_n)^2.}$$

We can either take this as a definition, or we can generalize the argument that we used to compute distance in $\mathbb{R}^4$. We regard this boxed formula as the $n$-dimensional version of the Pythagorean Theorem.

• How does this mathematical definition compare to the physical definition? We will say that two vectors are **equal** if they have the same "magnitude and direction". Find an algebraic formula to determine when the vectors $[\mathbf{x}, \mathbf{y}]$ and $[\mathbf{u}, \mathbf{v}]$ are equal. [Hint: Draw a parallelogram.]

Suppose that the vectors $[\mathbf{x}, \mathbf{y}]$ and $[\mathbf{u}, \mathbf{v}]$ have the same length and suppose that they are parallel. Then by connecting the tails and the heads we obtain 2D parallelogram. This parallelogram lives in $n$-dimensional space, but I can still draw it:




Here I have also drawn the diagonals of the parallelogram. One can show that the diagonals intersect at their common midpoint (using similar triangles, for example), but what are the **coordinates** of this midpoint? There are two ways to compute it. On the one hand it is the midpoint of $\mathbf{x}$ and $\mathbf{v}$,

$$\frac{\mathbf{x} + \mathbf{v}}{2} = \left( \frac{x_1 + v_1}{2}, \frac{x_2 + v_2}{2}, \ldots, \frac{x_n + v_n}{2} \right),$$

and on the other hand it is the midpoint of $\mathbf{u}$ and $\mathbf{y}$,

$$\frac{\mathbf{u} + \mathbf{y}}{2} = \left( \frac{u_1 + y_1}{2}, \frac{u_2 + y_2}{2}, \ldots, \frac{u_n + y_n}{2} \right).$$

[Discussion: Why is the this the correct way to compute midpoints? Look at the cases $n = 1$, 2, and 3. Can you convince yourself that the formula is true in general?] Since these two points are equal we conclude that

$$\mathbf{x} + \mathbf{v} = (x_1 + v_1, x_2 + v_2, \ldots, x_n + v_n) = (u_1 + y_1, u_2 + y_2, \ldots, u_n + y_n) = \mathbf{u} + \mathbf{y}.$$

We can equivalently write this as

$$\mathbf{y} - \mathbf{x} = (y_1 - x_1, y_2 - x_2, \ldots, y_n - x_n) = (v_1 - u_1, v_2 - u_2, \ldots, v_n - u_n) = \mathbf{v} - \mathbf{u}.$$

In summary, we declare that the *vectors* $[\mathbf{x}, \mathbf{y}]$ and $[\mathbf{u}, \mathbf{v}]$ are "equal" whenever the *points* $\mathbf{y} - \mathbf{x}$ and $\mathbf{v} - \mathbf{u}$ are equal:

$$\boxed{[\mathbf{x}, \mathbf{y}] = [\mathbf{u}, \mathbf{v}] \quad \Longleftrightarrow \quad \mathbf{y} - \mathbf{x} = \mathbf{v} - \mathbf{u}.}$$

Hey, this is starting to suggest some kind of "algebra" of points.

• Any vector can be put into **standard form** by moving its tail to the origin $\mathbf{0} = (0, \ldots, 0) \in \mathbb{R}^n$ without chainging its length or direction. If $[\mathbf{x}, \mathbf{y}] = [\mathbf{0}, \mathbf{v}]$, find a formula for $\mathbf{v}$.

Well, this is easy based on the previous problem. Since $[\mathbf{x}, \mathbf{y}] = [\mathbf{0}, \mathbf{v}]$ we have

$$\mathbf{y} - \mathbf{x} = \mathbf{v} - \mathbf{0}$$
$$\mathbf{y} - \mathbf{x} = \mathbf{v}.$$

We conclude that $\mathbf{v} = \mathbf{y} - \mathbf{x} = (y_1 - x_1, y_2 - x_2, \ldots, y_n - x_n)$ and hence

$$\boxed{[\mathbf{x}, \mathbf{y}] = [\mathbf{0}, \mathbf{y} - \mathbf{x}].}$$

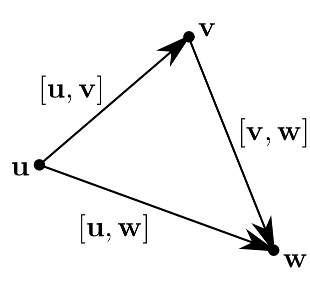
[A point of confusion: Many people and books will use the symbol $\mathbf{v}$ to refer to the vector $[\mathbf{0}, \mathbf{v}]$ with tail at the origin $\mathbf{0}$ and head at the point $\mathbf{v}$. This is certainly an abuse of notation because a vector is a very different thing from a point. An "abuse of notation" is the same thing as a "white lie".]

• Vectors can be "added". Given the three points $\mathbf{u}, \mathbf{v}, \mathbf{w} \in \mathbb{R}^n$ we define the sum

$$[\mathbf{u}, \mathbf{v}] + [\mathbf{v}, \mathbf{w}] := [\mathbf{u}, \mathbf{w}].$$

We say that vectors add "head-to-tail". Draw a picture of this. Why do we use the symbol "+" to refer to this operation? What does it have to do with addition of numbers?

The picture looks like this:




Addition of vectors is a very natural operation if we think of the vector $[\mathbf{u}, \mathbf{v}]$ as a function that says "move from point $\mathbf{u}$ to point $\mathbf{v}$" and the vector $[\mathbf{v}, \mathbf{w}]$ as a function that says "move from point $\mathbf{v}$ to point $\mathbf{w}$". By performing these two rules in sequence we get "move from point $\mathbf{u}$ to point $\mathbf{v}$ and **then** move from point $\mathbf{v}$ to point $\mathbf{w}$". Obviously this is the same as the rule "move from point $\mathbf{u}$ to point $\mathbf{w}$". Thus we can think of addition of vectors as "composition of functions".

OK, it may be a natural operation, but what does it have to do with addition of numbers? Let's postpone this question for a moment.

• Come up with a formula for adding two vectors in standard position:

$$[\mathbf{0},\mathbf{x}]+[\mathbf{0},\mathbf{y}]=?$$

Use your formula to explain why $[\mathbf{0},\mathbf{x}]+[\mathbf{0},\mathbf{y}]=[\mathbf{0},y]+[\mathbf{0},x]$.

To add vectors they must line up head to tail. Thus to compute $[\mathbf{0},\mathbf{x}]+[\mathbf{0},\mathbf{y}]$ we must first move $[\mathbf{0},\mathbf{y}]$ so that its tail is at the point $\mathbf{x}$. Then where will its head be? If $[\mathbf{0},\mathbf{y}]=[\mathbf{x},\mathbf{v}]$ then we have

$$\begin{aligned}\mathbf{y}-\mathbf{0}&=\mathbf{v}-\mathbf{x}\\ \mathbf{y}&=\mathbf{v}-\mathbf{x}\\ \mathbf{y}+\mathbf{x}&=\mathbf{v}.\end{aligned}$$

So if we move the vector $[\mathbf{0},\mathbf{y}]$ so its tail is at $\mathbf{x}$ then its head will be at the point $\mathbf{y}+\mathbf{x}$ (which is the same as $\mathbf{x}+\mathbf{y}$). Finally, we can add the vectors:

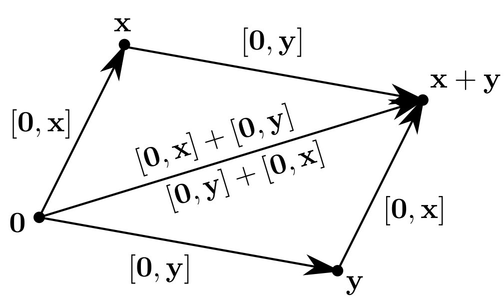
$$[\mathbf{0},\mathbf{x}]+[\mathbf{0},\mathbf{y}]=[\mathbf{0},\mathbf{x}]+[\mathbf{x},\mathbf{x}+\mathbf{y}]=[\mathbf{0},\mathbf{x}+\mathbf{y}].$$

Aha! Now we see why "addition" is a good name for this operation.

Now we also have two reasons why vector addition is commutative. The first reason is that addition of numbers is commutative:

$$[\mathbf{0},\mathbf{x}]+[\mathbf{0},\mathbf{y}]=[\mathbf{0},\mathbf{x}+\mathbf{y}]=[\mathbf{0},\mathbf{y}+\mathbf{x}]=[\mathbf{0},\mathbf{y}]+[\mathbf{0},\mathbf{x}].$$

The second reason is geometric. The vectors $[\mathbf{0},\mathbf{x}]$ and $[\mathbf{0},\mathbf{y}]$ form a parallelogram with corners at the points $\mathbf{0}$, $\mathbf{x}$, $\mathbf{y}$, and $\mathbf{x}+\mathbf{y}$. The sum of the vectors $[\mathbf{0},\mathbf{x}]$ and $[\mathbf{0},\mathbf{y}]$ is the diagonal of the parallelogram. By decomposing the parallelogram into two triangles it now becomes obvious why $[\mathbf{0},\mathbf{x}]+[\mathbf{0},\mathbf{y}]=[\mathbf{0},\mathbf{y}]+[\mathbf{0},\mathbf{x}]$:




• The "zero vector" $[\mathbf{0},\mathbf{0}]$ has a special property. What is this property?

Recall that we defined a "vector" as an ordered pair of points, and we said that two vectors are equal if the have the same length and direction. The zero vector has length 0, but it is a bit strange because it doesn't have any direction. Using the algebraic formula for equality of vectors we find that $[\mathbf{u},\mathbf{u}]=[\mathbf{0},\mathbf{0}]$ for any point $\mathbf{u}\in\mathbb{R}^n$ because

$$\mathbf{u}-\mathbf{u}=\mathbf{0}=\mathbf{0}-\mathbf{0},$$

so we can think of the zero vector as any point repeated twice. [Warning: We do **not** think of the zero vector as a point. It is a point of multiplicity two.]

The most important property of the zero vector is how it behaves with respect to vector addition. For any vector $[\mathbf{u}, \mathbf{v}]$ we have

$$[\mathbf{u}, \mathbf{v}] + [\mathbf{0}, \mathbf{0}] = [\mathbf{0}, \mathbf{0}] + [\mathbf{u}, \mathbf{v}] = [\mathbf{u}, \mathbf{v}].$$

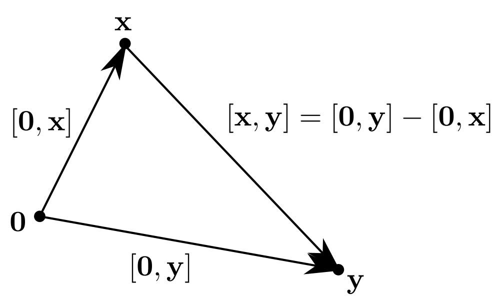
We say that $[\mathbf{0}, \mathbf{0}]$ is the "identity element" for vector addition. So "zero vector" is a good name because it plays an analogous role to the number zero in the arithmetic of numbers. We can also think of the zero vector as the function that tells you to "move from point $\mathbf{0}$ to point $\mathbf{0}$", i.e., it tells you "don't move". Typically the function that tells you "don't do anything" is called the "identity function".

• What does it mean to "subtract" vectors? Draw a picture.

Consider the vectors $[\mathbf{0}, \mathbf{x}]$ and $[\mathbf{0}, \mathbf{y}]$. On one hand, the vector "$[\mathbf{0}, \mathbf{y}] - [\mathbf{0}, \mathbf{x}]$" is just the vector $V$ that solves the equation

$$[\mathbf{0}, \mathbf{x}] + V = [\mathbf{0}, \mathbf{y}].$$

We see that $V = [\mathbf{x}, \mathbf{y}]$ will do the trick. On the other hand, we can think of the vectors $[\mathbf{0}, \mathbf{x}]$, $[\mathbf{0}, \mathbf{y}]$, and $[\mathbf{x}, \mathbf{y}]$ as the three sides of a triangle.




Note that subtraction of vectors is easy to compute by putting the vectors in standard position:

$$\begin{aligned}[\mathbf{0}, \mathbf{y}] - [\mathbf{0}, \mathbf{x}] &= [\mathbf{x}, \mathbf{y}] \\ &= [\mathbf{0}, \mathbf{y} - \mathbf{x}].\end{aligned}$$

From this point on, we will try to save space by denoting the vector $[\mathbf{0}, \mathbf{u}]$ simply as $\mathbf{u}$. This does **not** mean that a vector is a point; it is just a shorthand notation. So when I refer to the "vector" $\mathbf{u} \in \mathbb{R}^n$ I am really referring to the vector $[\mathbf{0}, \mathbf{u}]$. Luckily the notation has been set up so that if we just apply it mindlessly we will get correct results.