

Convex combinations

Tiny Explanations 1

Carl Joshua Quines

April 30, 2020

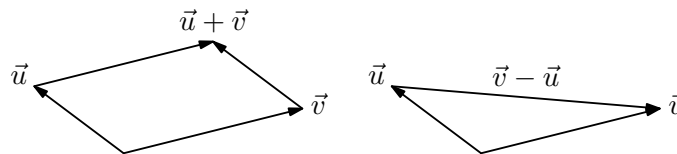
A *convex combination* of vectors is a linear combination, where all the scalars are non-negative and sum to 1. In other words, if your vectors were $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n$, then a convex combination would be of the form

$$a_1\vec{v}_1 + a_2\vec{v}_2 + \dots + a_n\vec{v}_n$$

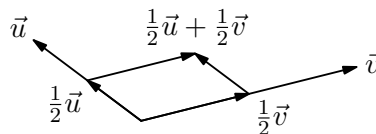
where $a_1, a_2, \dots, a_n \geq 0$, and $a_1 + a_2 + \dots + a_n = 1$. These have some nice geometric properties.

1 Two vectors

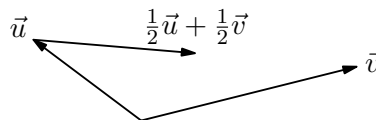
We know what it means to add two vectors, say, $\vec{u} + \vec{v}$. You draw \vec{u} , then you draw \vec{v} starting from the tip of \vec{u} . Similarly, the interpretation of $\vec{v} - \vec{u}$ is a vector going from the tip of \vec{u} to the tip of \vec{v} .



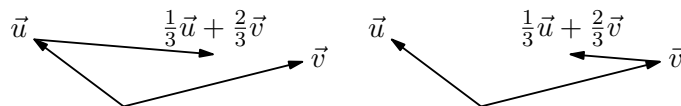
We can similarly define what $\frac{1}{2}\vec{u} + \frac{1}{2}\vec{v}$ would be. We add half of \vec{u} to half of \vec{v} .



But I claim this isn't the right way to think about what $\frac{1}{2}\vec{u} + \frac{1}{2}\vec{v}$ is. Instead, we should think of it as meaning $\vec{u} + \frac{1}{2}(\vec{v} - \vec{u})$. So this means, starting from \vec{u} , we go half the way towards \vec{v} :

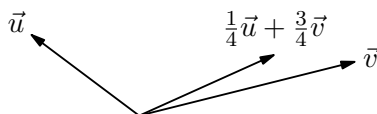


Similarly, the morally correct way to think of $\frac{1}{3}\vec{u} + \frac{2}{3}\vec{v}$ is something like $\vec{u} + \frac{2}{3}(\vec{v} - \vec{u})$. Which means “starting from \vec{u} , go two-thirds of the way to \vec{v} .” Alternatively, we can interpret this as $\vec{v} + \frac{1}{3}(\vec{u} - \vec{v})$, which means “starting from \vec{v} , go one-third of the way to \vec{u} .”



One thing to notice is that $\frac{1}{3}\vec{u} + \frac{2}{3}\vec{v}$ is closer to \vec{v} than it is to \vec{u} . In particular, it's twice as close to \vec{v} than it is to \vec{u} . Which makes sense, because $\frac{2}{3}$ is twice $\frac{1}{3}$.

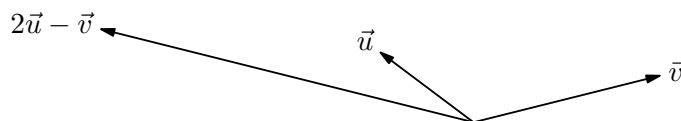
So where in the line connecting \vec{u} and \vec{v} would $\frac{1}{4}\vec{u} + \frac{3}{4}\vec{v}$ be? It would be closer to \vec{v} , because the $\frac{3}{4}$ is larger; in fact, thrice as close to \vec{v} :



This gives us a way to interpret convex combinations of two vectors. A convex combination of two vectors is just a vector that lies somewhere in between them!

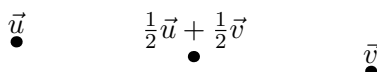
We can also use this to interpret where, say, $2\vec{u} - \vec{v}$ would be. This isn't really a convex combination, because -1 is negative. Although it is true that $2 - 1 = 1$, so we can interpret this similarly.

As 2 is bigger than -1 , it's closer to \vec{u} . We can rewrite it as $\vec{u} - (\vec{v} - \vec{u})$. So it's if we started at \vec{u} , took our distance to \vec{v} , and then went in the opposite direction. It would be outside the segment joining the two vectors, but still on the same line:



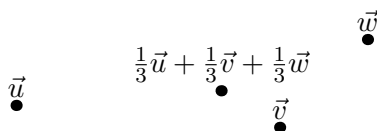
2 Three vectors

We're going to drop drawing the origin now, and instead just where the vectors are in the plane. So for example, we can draw \vec{u} , \vec{v} , and their midpoint $\frac{1}{2}\vec{u} + \frac{1}{2}\vec{v}$ like so:



Does it matter where the origin is? Not really. This is because the vector $\frac{1}{2}\vec{u} + \frac{1}{2}\vec{v}$ would always be at the midpoint of the two vectors, and the midpoint doesn't depend on where the origin is! And this is generally true of any convex combination. (This isn't generally true for other linear combinations.)

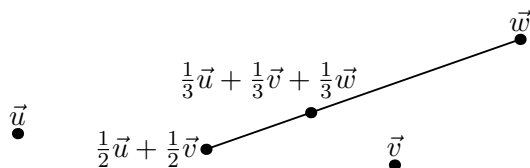
The same intuition carries over when we're dealing with three vectors. Like with two vectors, the center of the mass would be where the average is, which is $\frac{1}{3}\vec{u} + \frac{1}{3}\vec{v} + \frac{1}{3}\vec{w}$:



This has another nice interpretation, which is

$$\frac{1}{2}\vec{u} + \frac{1}{2}\vec{v} + \frac{1}{3}\left(\vec{w} - \left(\frac{1}{2}\vec{u} + \frac{1}{2}\vec{v}\right)\right).$$

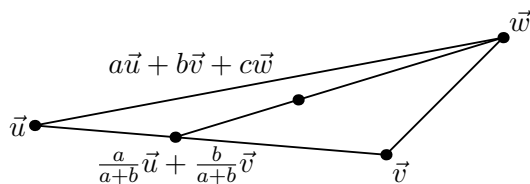
From our previous work, we know that this is the point that's one-third of the way from the midpoint of \vec{u} and \vec{v} to \vec{w} . Which gives us the well-known result about how the centroid divides the medians of the triangle.



This inspires us with the same logic as the two-vector case. If you have $a\vec{u} + b\vec{v} + c\vec{w}$, we can write it as the convex combination first of a vector between \vec{u} and \vec{v} :

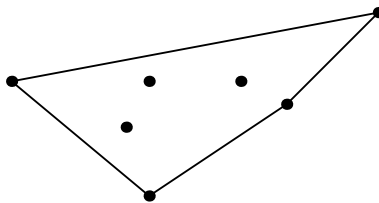
$$\frac{a}{a+b}\vec{u} + \frac{b}{a+b}\vec{v}.$$

And then we can write it as a convex combination of *that* vector and \vec{w} . So any convex combination of these three vectors has to lie inside the triangle they form:



And this is how barycentric coordinates work. The point with coordinates (a, b, c) , corresponds to the vector $a\vec{u} + b\vec{v} + c\vec{w}$.

In general, the convex combination of several vectors is the set of vectors that lie inside the “space” bounded by the vectors. More precisely, it's their *convex hull*, which you would get if you took all your vectors and imagined wrapping a rubber band around them.

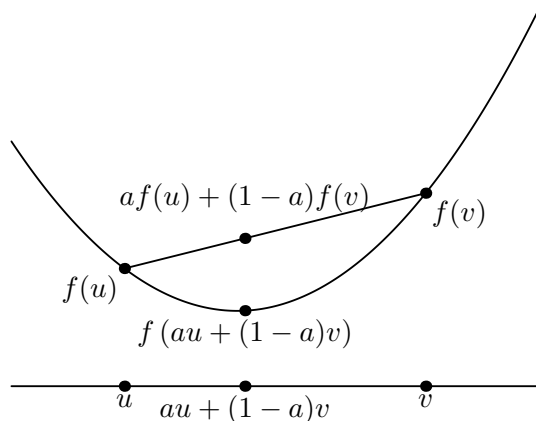


In three dimensions, it would be like wrapping all the points with a rubber sheet, say. You can probably imagine that; I'm not going to illustrate it.

3 Jensen's inequality

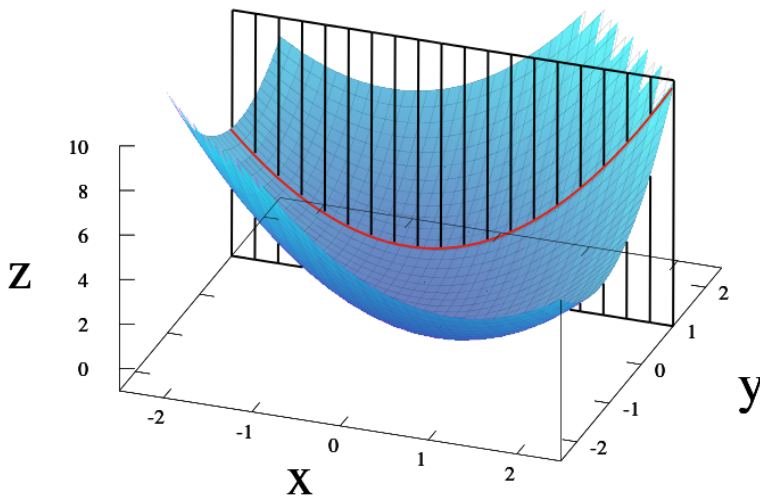
Why are convex functions called convex? Let's consider the one-dimensional case, where we have some function f that takes real numbers to real numbers. Then the convex combination of two points u and v on the real line is just some other point that lies in between them, like $au + (1-a)v$. And a function is convex if the value of the function at this point is lower than convex combination of the function values, $af(u) + (1-a)f(v)$. In other words, it's convex if, for all u and v ,

$$af(u) + (1-a)f(v) \geq f(au + (1-a)v).$$



Geometrically, the segment joining two points lies entirely above the values of the function in between.

We can generalize this to more dimensions, say, if f was a function that took vectors to real numbers. For example, here's what a convex function from \mathbb{R}^2 to \mathbb{R} looks like:



by [IkamuseFan](#) , CC-BY-SA 4.0

And we can also generalize this to more vectors, like we turned convex combinations from two vectors to three vectors, and the proof is similar to how we reduced convex combinations from three to two. The general statement is then Jensen's inequality, where for a convex combination $a_1\vec{x}_1 + a_2\vec{x}_2 + \dots + a_n\vec{x}_n$, we get

$$a_1f(\vec{x}_1) + a_2f(\vec{x}_2) + \dots + a_nf(\vec{x}_n) \geq f(a_1\vec{x}_1 + a_2\vec{x}_2 + \dots + a_n\vec{x}_n).$$