Barycentric Coordinates

http://palitri.com

Contents

1 Quick Solution

For a barycentric coordinate system, defined by initial point A and axis vectors $\vec{a}$ and $\vec{b}$, the barycentric coordinates (u, v) of a given point S are as follows:

$$u = \frac{(\vec{a} \cdot \vec{c})\left(\vec{b} \cdot \vec{b}\right) - \left(\vec{a} \cdot \vec{b}\right)\left(\vec{b} \cdot \vec{c}\right)}{(\vec{a} \cdot \vec{a})\left(\vec{b} \cdot \vec{b}\right) - \left(\vec{a} \cdot \vec{b}\right)^2}$$

$$v = \frac{\left(\vec{b} \cdot \vec{c}\right)(\vec{a} \cdot \vec{a}) - \left(\vec{a} \cdot \vec{b}\right)(\vec{a} \cdot \vec{c})}{(\vec{a} \cdot \vec{a})\left(\vec{b} \cdot \vec{b}\right) - \left(\vec{a} \cdot \vec{b}\right)^2}$$

Where u is the coordinate along $\vec{a}$, v is the coordinate along $\vec{b}$, $\vec{c} = \vec{AS}$.
For consistency, the following conditions should be met

$$\vec{a} \cdot \vec{a} \neq 0$$

$$\vec{b} \cdot \vec{b} \neq 0$$

$$\left(\vec{a} \cdot \vec{b}\right)^2 \neq (\vec{a} \cdot \vec{a})\left(\vec{b} \cdot \vec{b}\right)$$

The solution and the approach to barycentric coordinates are thoroughly covered in the rest of the document.

2 Introduction

Strictly, a barycentric coordinate system is defined by the vertices of a triangle. The coordinates of a point in such system, are defined by the point's location between a triangle's vertex and opposite side, varying from 1 to 0. Thus, a point has three barycentric coordinates - one for each triangle vertex. Also, the sum of the three barycentric coordinates of a point equals 1.

Let's take another approach, by which defining a point in barycentric coordinates can be looked upon as defining it in a coordinate system with a center and two axes, much like Cartesian.
A point S with barycentric coordinates (u, v, w) is expressed in Cartesian space as $S = wA + uB + vC$, where points A, B and C are the vertices of the triangle, which define the barycentric coordinate system.
Because $u + v + w = 1$, therefore $w = 1 - u - v$. This way $S = (1 - u - v)A + uB + vC$, which is $S = A + u(B - A) + v(C - A)$. Thus, S can be expressed via point A, two vectors - $\vec{AB}$ and $\vec{AC}$ and the coordinates (u, v).
I prefer this view on the matter, because it uses an already familiar concept and removes the complementary third coordinate.

Now, let's have a plain example of how a point is traditionally expressed in a two dimensional Cartesian coordinate system.

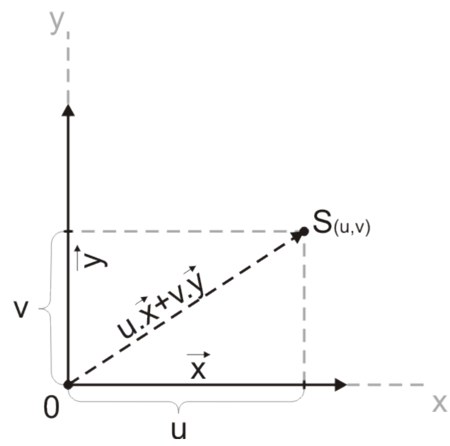

Figure 1: Cartesian coordinates

Point S is fully described by its Cartesian coordinates (u, v). They mean that we find the point by starting from the center of the coordinate system 0 then shifting u units along the x-axis direction and v units along the y-axis direction.
If we have the x-axis expressed by the vector $\vec{x} = (1, 0)$, and the y-axis expressed by the vector $\vec{y} = (0, 1)$, then the above statement can be expressed as $S = 0 + u\vec{x} + v\vec{y}$.

Well, it's just the same case with barycentric coordinates, only that the axes of a barycentric coordinate system can be any non-collinear vectors and the center can be any point in space.

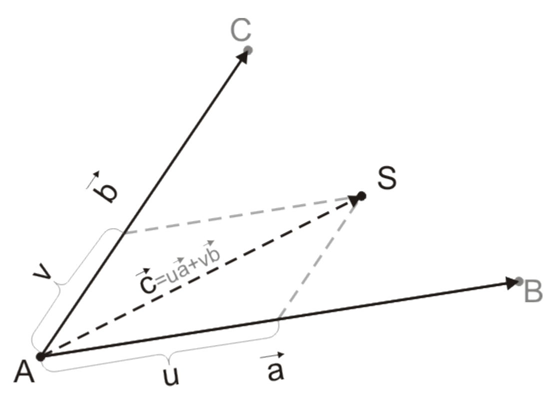

Figure 2: Setting

Let's have a point A and two non-collinear vectors $\vec{a}$ and $\vec{b}$. They fully describe a space with coordinate system that has A as its center and the vectors as its axes.
Any point can be described by them, for example the point S can be found by starting from A and shifting u units along $\vec{a}$ direction, then v units along $\vec{b}$ direction (one unit being the respective vector's length). And that is $S = A + u\vec{a} + v\vec{b}$ - quite the same as in Cartesian system.
Note that the length of the vector axes doesn't have to equal 1, nor do both vectors have to be of equal length.
Now, let's calculate the barycentric coordinates of a point

3 Setting

We have

- Point A
- Vectors $\vec{a}$ and $\vec{b}$
- Point S

We'll calculate the (u, v) coordinates of point S in terms of the barycentric coordinate system defined by A, $\vec{a}$ and $\vec{b}$.

4 Solution

As discussed, $S = A + u\vec{a} + v\vec{b}$.
Let's have vector $\vec{c} = \vec{AS}$. This way $S = A + \vec{c}$ and

$$\vec{c} = u\vec{a} + v\vec{b}$$

Now we have a vector-equation with two variables. The quickest way to solve it may seem to be to pick a vector component (x, y, z, etc.) to create one scalar equation out of it, pick another to create a second and solve them for the two variables. This approach however would rely on just one vector component to express u or v, which is not the whole relevant information, describing the setting, and that would lead to incompleteness in the result, see the Notes section for details.

For our approach, we'll create two equations out of the upper one, by dot multiplying it by $\vec{a}$ to create one scalar equation, and by $\vec{b}$ to create the other. This will give us the following equation set

$$\vec{a} \cdot \vec{a}u + \vec{a} \cdot \vec{b}v = \vec{a} \cdot \vec{c}$$
$$\vec{a} \cdot \vec{b}u + \vec{b} \cdot \vec{b}v = \vec{b} \cdot \vec{c}$$

Simple as that, it's only a matter of solving this equations set. All the following steps are just its thorough solution (perhaps too thorough).
So, let's substitute the dot products to simplify further calculations' appearance

$$p = \vec{a} \cdot \vec{b}$$
$$q = \vec{a} \cdot \vec{c}$$
$$r = \vec{b} \cdot \vec{c}$$
$$s = \vec{a} \cdot \vec{a}$$
$$t = \vec{b} \cdot \vec{b}$$

To skip the thorough solving process and get to the solution, go straight to Results section
So, now we have

$$su + pv = q$$
$$pu + tv = r$$

Express u from the first equation and v from the other

$$u = \frac{q - pv}{s}, s \neq 0$$
$$v = \frac{r - pu}{t}, t \neq 0$$

Replace v in the first equation by its expression from the second and replace u in the second equation by its expression from the first.

$$u = \frac{q}{s} - \frac{pr - p^2 u}{st}$$
$$v = \frac{r}{t} - \frac{pq - p^2 v}{st}$$

And having two equations containing only one of the variables each, work out u and v.

$$u\left(1 - \frac{p^2}{st}\right) = \frac{q}{s} - \frac{pr}{st}, p^2 \neq st$$
$$v\left(1 - \frac{p^2}{st}\right) = \frac{r}{t} - \frac{pq}{st}$$

$$u = \frac{q}{s - \frac{p^2}{t}} - \frac{pr}{st - p^2}$$
$$v = \frac{r}{t - \frac{p^2}{s}} - \frac{pq}{st - p^2}$$

$$u = \frac{q}{\frac{st-p^2}{t}} - \frac{pr}{st-p^2}$$
$$v = \frac{r}{\frac{st-p^2}{s}} - \frac{pq}{st-p^2}$$

$$u = \frac{qt}{st-p^2} - \frac{pr}{st-p^2}$$
$$v = \frac{rs}{st-p^2} - \frac{pq}{st-p^2}$$

Finally, we have

$$u = \frac{qt-pr}{st-p^2}$$
$$v = \frac{rs-pq}{st-p^2}$$

With the restrictions

$$s \neq 0$$
$$t \neq 0$$
$$p^2 \neq st$$

5 Notes

5.1 Restrictions

Looking back at the substitutions we did, let's take a brief to see what Math is saying with the restrictions that popped out while solving the equation set.

- $s \neq 0$, or $\vec{a} \cdot \vec{a} \neq 0$. That is $|\vec{a}| |\vec{a}| \cos\angle(\vec{a}, \vec{a}) \neq 0$. With the cosine being equal to 1, this leads to $|\vec{a}| \neq 0$. And of course, if the length of $\vec{a}$ is zero, it would be a point rather than a vector and our very setting would be inconsistent.
- $t \neq 0$ is analogous to $s \neq 0$ and means that $\left|\vec{b}\right| \neq 0$
- $p^2 \neq st$, or $\left(\vec{a} \cdot \vec{b}\right)^2 \neq (\vec{a} \cdot \vec{a}) \left(\vec{b} \cdot \vec{b}\right)$. That is $|\vec{a}| \left|\vec{b}\right| \cos\angle\left(\vec{a}, \vec{b}\right) |\vec{a}| \left|\vec{b}\right| \cos\angle\left(\vec{a}, \vec{b}\right) \neq |\vec{a}| |\vec{a}| 1 \left|\vec{b}\right| \left|\vec{b}\right| 1$, which leaves $\cos\angle\left(\vec{a}, \vec{b}\right) \neq 1$. In other words, it says $\vec{a}$ and $\vec{b}$ should not be collinear.

5.2 Point inside a triangle

Often, barycentric coordinates are used to determine whether a point is inside the boundaries of a triangle.

If $u \geq 0$, then the point is on the same side of $\vec{b}$ as the triangle.
If $v \geq 0$, then the point is on the same side of $\vec{a}$ as the triangle.
With that, we only need one such condition for the third side.
The line equation for that side would be $v = 1-u$, or $u+v = 1$. So if $u+v \leq 1$, then the point is on the same side as the triangle.

With these three put together, we can say that a point is inside the triangle if and only if

$$\begin{aligned} u &\geq 0 \\ v &\geq 0 \\ u+v &\leq 1 \end{aligned}$$

It's the same method as in the common check for a point in an axis aligned rectangle - we have one condition for each side and if all conditions are met, then the point is inside the boundaries of the figure. The difference here is that the sides are not four, but three, and that one of the sides is not axis aligned.

For the same thing, but using the classical approach to barycentric coordinates, it's quite easy to see, that the point is inside if all three coordinates are between 0 and 1. That is

$$\begin{aligned} 0 \leq &u \leq 1 \\ 0 \leq &v \leq 1 \\ 0 \leq &w \leq 1 \end{aligned}$$

And with $w = 1 - u - v$, the last line becomes $0 \leq 1 - u - v \leq 1$, which is $0 \leq u + v \leq 1$. But both u and v are already ensured to be greater than zero, so it's safe to reduce it to $u + v \leq 1$. This, on the other hand ensures $u \leq 1$ and $v \leq 1$, so we finally get to

$$\begin{aligned} u &\geq 0 \\ v &\geq 0 \\ u+v &\leq 1 \end{aligned}$$

5.3 Other means for determining if a point is inside a triangle

Another way to check in which half plane a point lies, may be to use cross products instead of barycentric coordinates. For example, if

$$\left(\vec{AB} \times \vec{AS}\right) \text{ and } \left(\vec{AB} \times \vec{AC}\right)$$

both have the same orientation, that is if

$$\left(\vec{AB} \times \vec{AS}\right) \cdot \left(\vec{AB} \times \vec{AC}\right) > 0$$

then both points S and C lie in the same half plane.
If this is carried out for each side of the triangle, and S is in the same half plane as the respective triangle vertex, the point has to be nowhere else, but inside the boundaries of the triangle.

A different approach would be to form three triangles - $\triangle ABS$, $\triangle BCS$ and $\triangle CAS$ and check if sum of the surfaces of these triangles is equal or bigger than the surface of the original. If bigger, then the point is outside.
The surface of a triangle can easily be computed as half the length of the cross product between two triangle sides, taken as vectors.
For example, $S_{\triangle ABC} = \frac{\left|\vec{AB} \times \vec{AC}\right|}{2}$.
And so, the point would be inside the triangle, if

$$\left|\vec{AB} \times \vec{AC}\right| = \left|\vec{AB} \times \vec{AS}\right| + \left|\vec{AC} \times \vec{AS}\right| + \left|\vec{BC} \times \vec{BS}\right|$$

Some floating point biasing may have to be considered.

The barycentric approach is preferable for several reasons - it's computationally most inexpensive, robust and gives additional information, that in most cases would be furtherly needed anyway. It has only advantages over other methods.

5.4 Interpolation with barycentric coordinates

Interpolation (as well as extrapolation) of values between the vertices of a triangle is a very common task. It's frequently used in graphics when vertex values like color, texture coordinates and so on, have to be calculated (interpolated) for a certain point on the triangle.
In fact, we've already used interpolation when formulating the position of the point S - in a sense, knowing the triangle has a flat surface, we interpolated its position bi-linearly between the position values of the triangle's vertices.
In the same way other values can be interpolated. Suppose V_a, V_b and V_c are values at A, B and C respectively. We can calculate the value at S as

$$V_s = V_a + uV_b + vV_c$$

5.5 Barycenter

Practically the center, or initial point, of a barycentric coordinate system is not a barycenter. It's just an origin to relate the other points to.
If talking in terms of astronomy, a barycentric coordinate system really is one that has the center of mass as its origin, but in general that's not the case.
The origin of a barycentric coordinate system is not its center of mass.

5.6 Component approach to solving the equation set

We can try to solve the equation $\vec{a}u+\vec{b}v=\vec{c}$ for u and v by splitting it into two of its components, let that be x and y. We'd have

$$\vec{a}_x u + \vec{b}_x v = \vec{c}_x$$
$$\vec{a}_y u + \vec{b}_y v = \vec{c}_y$$

But, if we express u from the first equation and v from the second, we'd have to divide by $\vec{a}_x$ and $\vec{b}_y$ respectively, which would lead to the limitation $\vec{a}_x \neq 0$ and $\vec{b}_y \neq 0$, and that is not quite pleasant, because we obviously CAN have such vectors forming a perfectly valid coordinate system.
For that, we can have another case, by expressing u from the second equation and v from the first, this time with the conditions $\vec{a}_y \neq 0$ and $\vec{b}_x \neq 0$.
Later on, in both cases we'd have the limitation $\frac{\vec{a}_x}{\vec{a}_y} \neq \frac{\vec{b}_x}{\vec{b}_y}$, which has some sense but only on the XY plane, where $\vec{a}$ and $\vec{b}$ would project into collinear vectors. However in three dimensional space they may not be collinear at all, which would call for more cases of the solution, including the z component. In more dimensions, we'll need to consider more vector components in order to have a fully consistent result, which will multiply the number of cases.

All this is due to the partial information that we plug in on the first place. That is, we express u and v with just one vector component each, omitting much information which may actually hold the relevant data. In case of three dimensional space, that's omitting two thirds of the information.
Luckily, math is faithful and whispering, we only need to listen.