Orthogonal Projection

The Punch Line: Inner products make it quite easy to compute the component of vectors that lie in interesting subspaces—in particular, components in the direction of any other vector.

Warm-Up What is the closest vector on the x-axis to the following vectors?

(a) $\begin{bmatrix} 4 \\ 0 \end{bmatrix}$ (b) $\begin{bmatrix} 4 \\ 9 \end{bmatrix}$ (c) $\begin{bmatrix} x \\ y \end{bmatrix}$

What is the closest point on the y-axis to these vectors? On the xy-plane?

(d) $\begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}$ (e) $\begin{bmatrix} 9 \\ 1 \\ 2 \end{bmatrix}$ (f) $\begin{bmatrix} x \\ y \\ z \end{bmatrix}$

(a) This vector is on the x-axis, so it is itself the closest vector on that axis.

(b) Here we have a y-component, but $\begin{bmatrix} 4 \\ 0 \end{bmatrix}$ is still the closest vector on the x-axis (any other vector would have a longer distance, as you can see by drawing out a triangle with the x-component, y-component, and distance between the given vector and another vector on the x-axis as legs).

(c) In general, the closest vector on the x-axis will be the vector $\begin{bmatrix} x \\ 0 \end{bmatrix}$.

(d) The closest on the y-axis is $\begin{bmatrix} 0 \\ 2 \\ 0 \end{bmatrix}$. The vector itself lives in the xy-plane.

(e) The closest vector on the y-axis is $\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$ (which is not particularly close, but changing y won't get you closer in the x or z directions). The closest vector in the xy-plane is $\begin{bmatrix} 9 \\ 1 \\ 0 \end{bmatrix}$, which is much closer (it's "right below" the real vector).

(f) In general, the closest vector in the y-axis is $\begin{bmatrix} 0 \\ y \\ 0 \end{bmatrix}$, and on the xy-plane is $\begin{bmatrix} x \\ y \\ 0 \end{bmatrix}$. This is the right kind of picture to have in mind for orthogonal projection—just keeping the components in the directions we specify. If they aren't already our axes, we have to do some computation (often with inner products), but geometrically we're doing the same thing.

Orthogonal Projection: If we have some vector $\vec{u}$ that we're interested in, we can compute the *orthogonal projection* of any other vector $\vec{v}$ onto $\vec{u}$ as $\frac{\vec{v}\cdot\vec{u}}{\vec{u}\cdot\vec{u}}\vec{u}$. That is, the ratio of the inner product of $\vec{v}$ and $\vec{u}$ to the inner product of $\vec{u}$ with itself is the coefficient on $\vec{u}$ giving the closest vector in Span $\{\vec{u}\}$ to $\vec{v}$. This coefficient can be thought of as "the amount of $\vec{v}$ in the direction of $\vec{u}$", and the projection (which is a vector) as "the component of $\vec{v}$ in the direction of $\vec{u}$."

1 Compute the projection of $\vec{v}$ onto $\vec{u}$.

(a) $\vec{v} = \begin{bmatrix}1\\2\\3\end{bmatrix}$ onto $\vec{u} = \begin{bmatrix}1\\1\\1\end{bmatrix}$

(b) $\vec{v} = \begin{bmatrix}2\\-1\end{bmatrix}$ onto $\vec{u} = \begin{bmatrix}1\\3\end{bmatrix}$

(c) $\vec{v} = \begin{bmatrix}2\\-1\end{bmatrix}$ onto $\vec{u} = \begin{bmatrix}1\\2\end{bmatrix}$

(d) $\vec{v} = \begin{bmatrix}1\\0\end{bmatrix}$ onto $\vec{u} = \begin{bmatrix}x\\y\end{bmatrix}$

(a) Here we get $\vec{v}\cdot\vec{u} = (1)(1)+(2)(1)+(3)(1) = 6$ and $\vec{u}\cdot\vec{u} = (1)^2+(1)^2+(1)^2 = 3$, so the projection is $\frac{6}{3}\begin{bmatrix}1\\1\\1\end{bmatrix} = \begin{bmatrix}2\\2\\2\end{bmatrix}$.

(b) Here we get $\vec{v}\cdot\vec{u} = -1$ and $\vec{u}\cdot\vec{u} = 10$, so the projection is $-\frac{1}{10}\begin{bmatrix}1\\3\end{bmatrix}$. This vector is fairly short (its length is $\frac{1}{\sqrt{10}} \approx 0.32$) compared to $\vec{v}$ (which has length $\sqrt{5} \approx 2.24$, about seven times larger), which means that most of $\vec{v}$ is in a direction orthogonal to $\vec{u}$. Also, the negative coefficient means that the part of $\vec{v}$ that lies "in the direction of $\vec{u}$" is in fact going away from $\vec{u}$—by in the direction, we mean along that same line, which includes going backwards.

(c) Here we get $\vec{v}\cdot\vec{u} = 0$. This means that $\vec{v}$ is orthogonal to $\vec{u}$. Geometrically, there is no component of $\vec{v}$ that is parallel to $\vec{u}$—they in a sense have "nothing to do with each other."

(d) Here we get $\vec{v}\cdot\vec{u} = x$ and $\vec{u}\cdot\vec{u} = x^2+y^2$. Thus, the projection is $\frac{x}{x^2+y^2}\begin{bmatrix}x\\y\end{bmatrix}$. It's worth noting that the length of this vector is less than (or equal to, if $y = 0$) 1—the length of the projection of a vector onto some other vector is at most the length of the original vector, and less unless they are parallel.

Projection Onto Subspaces: If W is a subspace of $\mathbb{R}^n$, we can compute the projection of a vector onto W. This is found by taking all and only the component of a vector which lie in W, which is most easily done if we have an orthogonal (or orthonormal) basis for W. Then we can simply compute the relevant inner products to project onto each basis vector, then add up all the results. (Note that this won't work if the basis isn't orthogonal.)

2 Project the vector $\vec{v}$ onto the subspace spanned by the given vectors.

(a) $\vec{v} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$, $W = \text{Span}\left\{\begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}\right\}$

(b) $\vec{v} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$, $W = \text{Span}\left\{\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}\right\}$

(a) Here we can check that the basis given for W is orthogonal, so we compute the inner product of $\vec{v}$ with the given vectors—0 and -2, respectively—and their inner products with themselves—6 and 2, respectively, to see that the projection of $\vec{v}$ onto W is $\frac{0}{6}\begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix} + \frac{-2}{2}\begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$.

(b) Here our basis is not orthogonal. We first remove the component of the second basis vector in the direction of the first to get an orthogonal vector—the projection of $\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$ onto $\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$ is $\frac{1}{2}\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$, so our new basis vector (the second minus its component in the direction of the first) is $\begin{bmatrix} 1/2 \\ -1/2 \\ 1 \end{bmatrix}$. You can check that this is indeed orthogonal to the first and in W.

Then we can project $\vec{v}$ onto $\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$ and $\begin{bmatrix} 1/2 \\ -1/2 \\ 1 \end{bmatrix}$ to get the projection as $\frac{3}{2}\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} + \frac{5/2}{3/2}\begin{bmatrix} 1/2 \\ -1/2 \\ 1 \end{bmatrix} = \begin{bmatrix} 7/3 \\ 2/3 \\ 5/3 \end{bmatrix}$.

Under the Hood: Why are orthogonal bases so much easier to project onto (we don't even have a good way to project onto the span of non-orthogonal vectors other than finding an orthogonal basis for that same subspace)? Heuristically, each vector in an orthogonal set is giving "independent information" about a vector in their span. Travelling in the direction of one of them doesn't move at all in the direction of the others, while for non-orthogonal vectors, increasing in one direction also moves in some of the others, and it's hard to separate the effects.

So, a basis gives enough information to describe any vector (it spans the space) and doesn't have redundant information (it's linearly independent), while an *orthogonal* basis also has the property that pieces of that description don't interfere with each other. An orthonormal basis is even nicer, in that the information requires less processing to get information about lengths—the coefficient on each component is the length in that direction (in other bases, the length of the basis vector changes this).