Lecture 1n

Projections and Perpendiculars

(pages 40-42)

A common thing to look for is the distance from a point to a line. That is, to look for the *shortest* distance from our starting point to a point on the line.

In this course, we assume that you already know that we find the minimum distance by looking at the connecting segment that is perpendicular to our line.

As such, our goal will be to look at a way to calculate this value. We will, of course, start with a definition.

Definition The part of $\vec{y}$ that is in the direction of $\vec{x}$ will be called the **projection of $\vec{y}$ onto $\vec{x}$**, and is denoted by $\text{proj}_{\vec{x}}\vec{y}$.

Note that $\text{proj}_{\vec{x}}\vec{y}$ will be in the same direction as $\vec{x}$, and thus will be a scalar multiple of $\vec{x}$. The trick, therefore, is to find the correct scalar. So, let k be that scalar, so that $\text{proj}_{\vec{x}}\vec{y} = k\vec{x}$. Now, let $\vec{z}$ be such that $\vec{y} + \vec{z} = \text{proj}_{\vec{x}}\vec{y}$. Then $\vec{z}$ is orthogonal to $\vec{x}$, since that's the main point of taking the projection.

Putting these facts together, we have that $k\vec{x} = \vec{y} + \vec{z}$. And while we can't "divide by $\vec{x}$" to solve for k, we can take the dot product of both sides, getting that $\vec{x} \cdot (k\vec{x}) = \vec{x} \cdot (\vec{y} + \vec{z}) \Rightarrow k(\vec{x} \cdot \vec{x}) = \vec{x} \cdot \vec{y} + \vec{x} \cdot \vec{z}$. But since $\vec{z}$ is orthogonal to $\vec{x}$, we have $\vec{x} \cdot \vec{z} = 0$. And therefore we get that $k = (\vec{x} \cdot \vec{y})/\vec{x} \cdot \vec{x}$. This gives us that

$$\text{proj}_{\vec{x}}\vec{y} = \frac{\vec{y} \cdot \vec{x}}{||\vec{x}||^2}\vec{x}$$

Note that the textbook takes the above formula to be the definition of a projection, but I feel it is important to remember the geometrical origin.

Example: Let $\vec{x} = \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix}$ and let $\vec{y} = \begin{bmatrix} 0 \\ 3 \\ 4 \end{bmatrix}$. Then $\vec{x} \cdot \vec{y} = \vec{y} \cdot \vec{x} = (1)(0) + (2)(3) + (-1)(4) = 2$, $||\vec{x}||^2 = 1^2 + 2^2 + (-1)^2 = 6$, and $||\vec{y}||^2 = 0^2 + 3^2 + 4^2 = 25$. Then $\text{proj}_{\vec{x}}\vec{y} = (2/6)\vec{x} = (1/3)\vec{x} = \begin{bmatrix} 1/3 \\ 2/3 \\ -1/3 \end{bmatrix}$, and $\text{proj}_{\vec{y}}\vec{x} = (2/25)\vec{y} = \begin{bmatrix} 0 \\ 6/25 \\ 8/25 \end{bmatrix}$.

Now, we created and then quickly discarded our orthogonal vector $\vec{z}$ when finding the formula for $\text{proj}_{\vec{x}}\vec{y}$, but it is actually the distance vector we were originally looking for.

Definition For any vectors $\vec{x}, \vec{y} \in \mathbb{R}^n$, with $\vec{x} \neq \vec{0}$, we define the **projection of $\vec{y}$ perpendicular to $\vec{x}$** to be

$$\text{perp}_{\vec{x}}\vec{y} = \vec{y} - \text{proj}_{\vec{x}}\vec{y}$$

Note that $\text{perp}_{\vec{x}}\vec{y}$ travels from $\vec{x}$ to $\vec{y}$, which is the opposite direction of the vector $\vec{z}$ used to find the formula for $\text{proj}_{\vec{x}}\vec{y}$. And so we can split a vector $\vec{y}$ into its portion in the direction of $\vec{x}$ and its portion perpendicular to $\vec{x}$, since $\vec{y} = \text{perp}_{\vec{x}}\vec{y} + \text{proj}_{\vec{x}}\vec{y}$.

Example: Let $\vec{x} = \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix}$ and let $\vec{y} = \begin{bmatrix} 0 \\ 3 \\ 4 \end{bmatrix}$. Then in the previous example, we calculated that $\text{proj}_{\vec{x}}\vec{y} = \begin{bmatrix} 1/3 \\ 2/3 \\ -1/3 \end{bmatrix}$ and $\text{proj}_{\vec{y}}\vec{x} = \begin{bmatrix} 0 \\ 6/25 \\ 8/25 \end{bmatrix}$. So now we get that

$$\text{perp}_{\vec{x}}\vec{y} = \vec{y} - \text{proj}_{\vec{x}}\vec{y} = \begin{bmatrix} 0 \\ 3 \\ 4 \end{bmatrix} - \begin{bmatrix} 1/3 \\ 2/3 \\ -1/3 \end{bmatrix} = \begin{bmatrix} -1/3 \\ 7/3 \\ 11/3 \end{bmatrix}$$

and

$$\text{perp}_{\vec{y}}\vec{x} = \vec{x} - \text{proj}_{\vec{y}}\vec{x} = \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix} - \begin{bmatrix} 0 \\ 6/25 \\ 8/25 \end{bmatrix} = \begin{bmatrix} 1 \\ 44/25 \\ -33/25 \end{bmatrix}$$

Now, before I have you do some of these calculations yourself, we should pause to look at the linearity properties of projections, which I tend to put into the category of "useful properties", so pay attention!

L1 $\text{proj}_{\vec{x}}(\vec{y} + \vec{z}) = \text{proj}_{\vec{x}}\vec{y} + \text{proj}_{\vec{x}}\vec{z}$ and $\text{perp}_{\vec{x}}(\vec{y} + \vec{z}) = \text{perp}_{\vec{x}}\vec{y} + \text{perp}_{\vec{x}}\vec{z}$ for all $\vec{x}, \vec{y}, \vec{z} \in \mathbb{R}^n$

L2 $\text{proj}_{\vec{x}}(t\vec{y}) = t\ \text{proj}_{\vec{x}}\vec{y}$ and $\text{perp}_{\vec{x}}(t\vec{y}) = t\ \text{perp}_{\vec{x}}\vec{y}$ for all $\vec{x}, \vec{y} \in \mathbb{R}^n$ and $t \in \mathbb{R}$

Proof of L2 Let $\vec{x}, \vec{y} \in \mathbb{R}^n$ and $t \in \mathbb{R}$. Then $\text{proj}_{\vec{x}}(t\vec{y}) = \frac{(t\vec{y})\cdot\vec{x}}{||\vec{x}||^2}\vec{x} = \frac{t(\vec{y}\cdot\vec{x})}{||\vec{x}||^2}\vec{x} = t\frac{\vec{y}\cdot\vec{x}}{||\vec{x}||^2}\vec{x} = t\ \text{proj}_{\vec{x}}(\vec{y})$, as desired. And this means that $\text{perp}_{\vec{x}}(t\vec{y}) = t\vec{y} - \text{proj}_{\vec{x}}(t\vec{y}) = t\vec{y} - t\ \text{proj}_{\vec{x}}(\vec{y}) = t(\vec{y} - \text{proj}_{\vec{x}}(\vec{y})) = t\ \text{perp}_{\vec{x}}\vec{y}$, as desired.

The proof of L1 is similar. There are also two other facts worth noting.

(1) $\text{proj}_{\vec{x}}(\text{proj}_{\vec{x}}(\vec{y})) = \text{proj}_{\vec{x}}(\vec{y})$ for all $\vec{x}, \vec{y} \in \mathbb{R}^n$

(2) $\text{perp}_{\vec{x}}(\text{perp}_{\vec{x}}(\vec{y})) = \text{perp}_{\vec{x}}(\vec{y})$ for all $\vec{x}, \vec{y} \in \mathbb{R}^n$

These properties simply state that once you have found the projection/perpendicular part, there is no further action to be taken. In fact, we can make note of the more generic fact that if $\vec{y}$ is a scalar multiple of $\vec{x}$, then $\text{proj}_{\vec{x}}(\vec{y}) = \vec{y}$ (and thus $\text{perp}_{\vec{x}}(\vec{y}) = \vec{0}$), and that if $\vec{y}$ is orthogonal to $\vec{x}$, then $\text{perp}_{\vec{x}}(\vec{y}) = \vec{y}$ (and thus $\text{proj}_{\vec{x}}(\vec{y}) = \vec{0}$).