

# 1 Why the vector that represents the shortest distance from a point to a (hyper)plane must be normal to the plane

Hi everyone. This is a tedious proof of the fact above. Some of you may find it interesting, others may not. Don't worry, this is NOT mandatory reading. Regardless, please remember one thing: If you have a problem where you must relate lengths of vectors to facts about dot products and orthogonality, ALWAYS, ALWAYS, ALWAYS, try to work with the quantity  $|\vec{v}|^2$  instead of  $|\vec{v}|$  because  $|\vec{v}|^2 = \vec{v} \cdot \vec{v}$ .

Just to be clear, in mathematical circles the term hyperplane is used for a subspace of  $\mathbb{R}^m$  of dimension of  $m - 1$ . This generalizes the situation of an ordinary two dimensional plane sitting inside of  $\mathbb{R}^3$ . I tend to get lazy and forget the write the "hyper" part though. OK.

Let  $P$  a hyperplane in  $\mathbb{R}^n$  and  $q \in \mathbb{R}^n$  be a point which does not lie in  $P$ .

To be completely precise we will show two things:

- 1) That there IS a vector,  $\vec{y}$ , from  $q$  to  $\mathbb{R}^n$  that is normal to  $P$ .
- 2) Then it suffices to show that for any other vector  $\vec{x}$  stretching from  $q$  to a point in  $P$  we have  $|\vec{x}| \geq |\vec{y}|$  with equality if and only if  $\vec{x} = \vec{y}$ .

Recall that a hyperplane is given by some equation

$$n_1(v_1 - p_1) + \dots + n_m(v_m - p_m) = 0$$

for constants  $n_1, \dots, n_m, p_1, \dots, p_m \in \mathbb{R}^m$ . Of course now that we're big boys and girls we'd rather just write

$$\vec{n} \cdot (\vec{v} - \vec{p}) = 0.$$

Again, this equation says that all vectors normal to  $\vec{n}$  lie in  $P$  and  $P$  also contains the point  $p$ , (aka the vector  $\vec{p}$  stretching from the origin to the point  $p$ ). Now lets show 1).

Since  $p \in P$ , the vector  $\vec{a} := \begin{bmatrix} q_1 - p_1 \\ \vdots \\ q_m - p_m \end{bmatrix}$  stretches from  $p$  (in the plane) to

$q$ . Then consider  $\text{Proj}_{\vec{n}}(\vec{a})$ . It is normal to  $P$  because by its definition it is a multiple of  $\vec{n}$ . It also represents a line segment from  $q$  to a point in  $P$  for the following reason. Translate  $\vec{a}$  and  $\text{Proj}_{\vec{n}}(\vec{a})$  so their heads point to  $q$ . Then the vector  $\vec{a} - \text{Proj}_{\vec{n}}(\vec{a})$  is a vector whose head lies at the tail of  $\text{Proj}_{\vec{n}}(\vec{a})$  and whose tail is at  $p$ . But notice that

$$\begin{aligned} (\vec{a} - \text{Proj}_{\vec{n}}(\vec{a})) \cdot \vec{n} &= \vec{a} \cdot \vec{n} - \text{Proj}_{\vec{n}}(\vec{a}) \cdot \vec{n} = \vec{a} \cdot \vec{n} - \left( \frac{\vec{a} \cdot \vec{n}}{|\vec{n}|^2} \vec{n} \right) \cdot \vec{n} \\ &= \vec{a} \cdot \vec{n} - \frac{\vec{a} \cdot \vec{n}}{|\vec{n}|^2} |\vec{n}|^2 = \vec{a} \cdot \vec{n} - \vec{a} \cdot \vec{n} = 0. \end{aligned}$$

So the vector  $\vec{a} - \text{Proj}_{\vec{n}}(\vec{a})$  is normal to  $\vec{n}$ , thus it must lie in  $P$ , meaning that if its tail is at a point in  $P$ , then so is its head. Indeed, its tail is  $p \in P$ , so its head is in  $P$ . Remember though that its head is at the same point as the tail of  $\text{Proj}_{\vec{n}}(\vec{a})$ . So  $\text{Proj}_{\vec{n}}(\vec{a})$  represents a line stretching from  $P$  to  $q$ . We're done showing 1). Now we show 2).

Lets allow  $\vec{y}$  to stand in for  $\text{Proj}_{\vec{n}}(\vec{a})$ . Again, by similar reasoning, if  $\vec{x}$  is any other vector which represents a line segment from  $P$  to  $q$ , then the vector  $\vec{x} - \vec{y}$  lies in  $P$ . Now

$$|\vec{x}|^2 = |(\vec{x} - \vec{y}) + \vec{y}|^2 = |\vec{x} - \vec{y}|^2 + 2(\vec{x} - \vec{y}) \cdot \vec{y} + |\vec{y}|^2.$$

Now recall that as  $\vec{x} - \vec{y}$  lies in  $P$ , it must be normal to  $\vec{n}$ . As  $\vec{y}$  is a multiple of  $\vec{n}$ ,  $(\vec{x} - \vec{y}) \cdot \vec{y} = 0$ . So the expression above reduces to

$$|\vec{x} - \vec{y}|^2 + |\vec{y}|^2 \geq |\vec{y}|^2.$$

When do we have equality? Well, precisely when  $|\vec{x} - \vec{y}| = 0$ , which happens precisely when  $\vec{x} - \vec{y} = 0 \Leftrightarrow \vec{x} = \vec{y}$ .

Alright, I hope this wasn't overkill for you.