The Margin of a Separating Hyperplane

We are given the following setting:

- An inner product space \( \mathcal{V} \) with norm \( \|v\| = \sqrt{\langle v, v \rangle} \)

- Labelled vectors \((y_1, x_1), \ldots, (y_m, x_1)\)

- A separating hyperplane defined by the vector \( w \) where \( w \) is normal to the hyperplane with norm \( \|w\| = 1 \). This hyperplane separates all of the labelled vectors by their label. That is

\[
\forall i \; y_i \langle w, x_i \rangle > 0
\]

In this setting, the **margin** of the hyperplane is defined to be the closest vector to the separating hyperplane. This margin is given as follows:

\[
\text{margin} = \min_i \|\langle w, x_i \rangle\|
\]

As we will soon prove, \( |\langle w, x_i \rangle| \) is the distance of \( x_i \) to the hyperplane. Thus, we see that the margin is defined to be the minimal distance of any point to the hyperplane.

Distance of a Vector to the Hyperplane

The distance of a point \( x \) to the hyperplane is defined as

\[
\min \{ \|x - v\| \mid \langle w, v \rangle = 0 \}
\]

That is, the distance to hyperplane is the smallest distance between \( x \) and some vector \( v \) that lies on the hyperplane (i.e. \( v \) lies on the hyperplane because it is orthogonal to \( w \) as evident by the fact \( \langle w, v \rangle = 0 \)). In fact, this distance is given by

\[
|\langle w, x \rangle + b| = \min \{ \|x - v\| \mid \langle w, v \rangle = 0 \}
\]

**Proof:**

Let the vector \( v \) be defined to be

\[
v := x - \langle w, x \rangle w
\]

We see that this vector lies on the hyperplane because it’s inner product with \( w \) is 0 as shown here:
\[
\langle w, v \rangle = \langle w, x - \langle w, x \rangle w \rangle
\]

\[
= \langle w, x \rangle - \langle w, \langle w, x \rangle w \rangle
\]

\[
= \langle w, x \rangle - \langle w, x \rangle \langle w, w \rangle
\]

\[
= \langle w, x \rangle - \langle w, x \rangle \|w\|^2
\]

\[
= \langle w, x \rangle
\]

\[
= 0
\]

Next we see that no vector on the hyperplane is closer to \( x \) as \( v \).

\[
\|x - u\|^2 = \|x - v + v - u\|^2
\]

\[
= \|x - v\|^2 + \|v - u\|^2 + 2\langle x - v, v - u \rangle \quad \text{See Lemma 1.}
\]

\[
\geq \|x - v\|^2 + 2\langle x - v, v - u \rangle 
\]

\[
= \|x - v\|^2 + 2\langle x - (x - \langle w, x \rangle w), v - u \rangle
\]

\[
= \|x - v\|^2 + 2\langle x - x + \langle w, x \rangle w, v - u \rangle
\]

\[
= \|x - v\|^2 + 2\langle \langle w, x \rangle w, v - u \rangle
\]

\[
= \|x - v\|^2 + 2\langle w, x \rangle \langle w, v - u \rangle
\]

\[
= \|x - v\|^2 
\]

\( v - u \) is on the hyperplane, thus \( \langle w, v - u \rangle = 0 \).

Lastly, we see that the norm of \( x - v \) is \( \langle w, x \rangle \):
\[\|x - v\| = \|x - (x - \langle w, x \rangle w)\|\]

\[= \|\langle w, x \rangle w\|\]

\[= \langle w, x \rangle \|w\|\]

\[= \langle w, x \rangle\]

\[\square\]

**Lemma 1:**

\[\|v + u\|^2 = \|v\|^2 + \|u\|^2 + 2\langle v, u \rangle\]

**Proof of Lemma 1:**

\[\|v + u\|^2 = \langle v + u, v + u \rangle\]

\[= \langle v, v + u \rangle + \langle u, v + u \rangle\]

\[= \langle v, v \rangle + \langle v, u \rangle + \langle u, v + u \rangle\]

\[= \langle v, v \rangle + \langle v, u \rangle + \langle u, v \rangle + \langle u, u \rangle\]

\[= \|v\|^2 + \langle v, u \rangle + \langle v, u \rangle + \|u\|^2\]

\[= \|v\|^2 + \|u\|^2 + 2\langle v, u \rangle\]

\[\square\]