Problem 1 Let $(w, b) \in \mathbb{R}^d \times R$ and $x \in \mathbb{R}^d$. Assume that $\|w\|_2 = 1$. Let us consider the point v defined by $v = x - (w \cdot x + b)w$.

1. (2 points) Show that $w \cdot v + b = 0$

$(w,b) \in \mathbb{R}^d \times \mathbb{R}$, $x \in \mathbb{R}^d$, $\|w\|^2 = 1$ where $v = x - (w \cdot x + b)w$

$w \cdot v + b = w \cdot [x - (w \cdot x + b)w] + b = w \cdot x - \|w\|^2 (w \cdot x + b) + b$ where $\|w\|^2 = 1$

$= w \cdot x - w \cdot x - b + b = 0$

2. (3 points) Using the previous question, prove that the following holds:

$$\min_{u \in \mathbb{R}^d \text{ s.t. } w \cdot u + b = 0} \|x - u\|_2 \le |w \cdot x + b|$$

Since $w \cdot u + b = 0$, $\therefore w \cdot x + b - (w \cdot u + b) = w \cdot x + b$

$w \cdot x - w \cdot u = wx + b \Rightarrow |w \cdot (x - u)| = |w \cdot x + b|$

On the left side: $|w \cdot (x-u)| = |w^T (x-u)|$

Recall Cauchy-Bunyakovsky-Schwarz inequality

$$|x^T y| \le \|x\|_2 \|y\|_2$$

$\therefore |w^T (x-u)| \le \|w\|_2 \|x-u\|_2 \Rightarrow \|x-u\|_2 \geq |w^T (x-u)|$

The minimum value of $\|x-u\|_2$ is 0, when $x = u$

$|w \cdot x + b|$ has minimal value 0, the rest values are all greater than 0

$$\therefore \|x-u\|_2 \le |w \cdot x + b|$$

3. (4 points) Let $u \in \mathbb{R}^d$ such that $w \cdot u + b = 0$. Show that $\|x - u\|_2 \geq \|x - v\|_2$.

Recall: $v = x - (w \cdot x + b)w \Rightarrow x - v = (w \cdot x + b)w \therefore \|x - v\| = |wx + b| \|w\|$ ---- ①

Since $w \cdot u + b = 0$, $\therefore w \cdot x + b - (w \cdot u + b) = w \cdot x + b$

We can get: $w \cdot x - w \cdot u = wx + b$

$w \cdot (x-u) = wx + b \Rightarrow |w^T (x-u)| = |wx + b|$ ---- ②

Recall Cauchy-Bunyakovsky-Schwarz inequality

$$|x^T y| \le \|x\|_2 \|y\|_2$$

Apply the inequality to ②, we get:

$$|w^T(X-u)| \le \|w\|_2 \, \|X-u\|, \quad \text{where } \|w\|_2 = 1$$

$$\therefore |w^T(X-u)| \le \|X-u\| \quad \text{------} ③$$

From ① and ②, we know $|w^T(X-u)| = |wx+b| = \|x-v\|$

Combine ①, ② and ③, we get: $\|X-u\|_2 \geq \|X-v\|_2$

4. (2 points) Use the above results to find the analytical expression for the ℓ_2 distance between a point x and the hyperplane defined by $w \cdot u + b = 0$ for all $u \in \mathbb{R}^d$.

From the results above, we know that: $w \cdot u + b = 0$, $w \cdot v + b = 0$

$$\therefore w \cdot v + b - (w \cdot u + b) = 0 \Rightarrow w(v-u) = 0$$

Thus, w is a orthogonal vector to the plane $w \cdot u + b = 0$

Then, we can get ux is parallel to w

$$\therefore |w \cdot ux| = |w||xu|\cos\theta = |w||xu| = \|w\|_2$$

$$\therefore |w \cdot ux| = |w \cdot (u-x)| = |wu - wx|$$

$$= |-(b+wx)| = |b+wx|$$

The ℓ_2 distance from the point to the plane is:

$$\ell_2 = \frac{|wx+b|}{\|w\|}$$

Problem 2 Recall the perceptron algorithm from class. It can be found in lecture notes posted on the course website.

1. (2 points) Let us modify the update rule to add a learning rate α which multiplies the update applied to weights. The update rule becomes: $w' \leftarrow w + \alpha t^{(i)} x^{(i)}$. Show that the value of the learning rate does not change the perceptron's prediction after an update.

The choice of learning rate α does not matter becaust it just changes the scaling of w. Learing rate is the length of steps the algorithm makes down the gradient on the error curve,

$$w' \leftarrow w + \alpha t^{(i)} x^{(i)}$$

Adding a constant to the update function does not change the step direction, as it does not change the sign of the decision.

As long as $\alpha > 0$, it does not change perceptron decision in any step.