Caio Brighenti
COSC 480 - Learning From Data
Fall 2019
Problem Set 4

1. Show $||w(t)||^2 \leq ||w(t-1)||^2 + ||x(t-1)||^2$ where x is the misclassified observation.

   We start by defining more clearly the terms we are considering. w^{new} in the PLA algorithm can be defined as $w^{new} = w + y \cdot x$, where w is the current weight, x the misclassified observation, and y the correct label. Given that x is misclassified, we can know for certain the quantity $y \cdot x$ is negative, as x and y must necessarily have different signs. Thus, we can rewrite this definition as an inequality as follows: $w^{new} \leq w + abs(x)$. y is omitted as simply takes the value 1.

   We thus know that w^{new} is smaller than or equal to w plus the magnitude of x. Instead of taking the absolute value, we could have used any function that took the magnitude of x. Another such example is the norm operation, which squares each component of x and thus cancels out the negatives. Thus, we can rewrite the previous inequality as $||w^{new}||^2 \leq ||w||^2 + ||x||^2$, as taking the norm of each component will still preserve the inequality. Thus we have shown the claim to be true.

2. Use induction to show $||w(t)||^2 \leq tR^2$.

   As we are proceeding by induction, we first show a base case and establish an inductive hypothesis.

   Base case: Our base case is that $t = 0$. This base is trivial, as when $t = 0$ we have:

   $$||w(0)||^2 \leq 0 \cdot R \tag{1}$$
   $$0 \leq 0 \tag{2}$$

   Thus the base case holds.

   Inductive hypothesis: Assume the claim holds for $t-1$. We will show it then holds for t.

   $$\begin{aligned} ||w(t)||^2 &\leq ||w||^2 + ||x||^2 && \text{shown in 1} && (3)\\ &\leq (t-1)R^2 + ||x||^2 && \text{inductive hypothesis} && (4)\\ &\leq (t-1)R^2 + R^2 && ||x|| \text{ must be at most } R && (5)\\ &\leq tR^2 && && (6) \end{aligned}$$

   Therefore the claim must be true.

3. Use the properties found in 1 and 2 to show $\frac{w(t) \cdot w^*}{||w(t)||} \geq \sqrt{t}\frac{M}{R}$.

   We start by first slightly modifying the relationship found in 2 in order to remove the squared terms. We simply take the square root of both sides, resulting in the property $||w(t)|| \leq \sqrt{t}R$. We now observe that given an inequality $x \geq y$, we can obtain the inequality $\frac{x}{a} \geq \frac{y}{b}$, on the condition that $a \leq b$. Using this property, we thus have:

   $$w(t) \cdot w^* \geq tM \tag{7}$$
   $$\frac{w(t) \cdot w^*}{||w(t)||} \geq \frac{t}{\sqrt{t}}\frac{M}{R} \tag{8}$$
   $$\frac{w(t) \cdot w^*}{||w(t)||} \geq \sqrt{t}\frac{M}{R} \qquad \frac{x}{\sqrt{x}} = \sqrt{x} \tag{9}$$

   Thus the claim is true.

4. Show that $t \leq \frac{R^2||w^*||^2}{M^2}$

   We can arrive at this claim by simply modifying a few terms from the previous problem as follows.

$$\sqrt{t}\frac{M}{R} \leq \frac{w(t) \cdot w^*}{||w(t)||} \tag{10}$$

$$t\frac{M^2}{R^2} \leq \frac{w(t)^2 \cdot (w^*)^2}{||w(t)||^2} \qquad \text{square both sides} \tag{11}$$

$$t\frac{M^2}{R^2} \leq (w^*)^2 \qquad w(t)^2 \text{is the dot product of } t \text{ with } t\text{, which is the same as } ||w(t)||^2 \tag{12}$$

$$t \leq \frac{R^2||w^*||^2}{M^2} \qquad \text{rearrange} \tag{13}$$

   Thus the claim is true.