CS534 Xiao Tan Exercise 1

1. We know $w_{i+1} = w_i + y_i x_i$ if $\vec{w}_0 \neq \vec{0}$ and learning rate not 1

let $w_0 = w$ and learning rate c

Part One

$$w_{i+1} = w_i + c y_i x_i$$
$$= w + c \cdot \sum_{k=0}^{i} y_k x_k \quad \because y_i (w \cdot x_i) \geq \delta \text{ for all}$$
$$w \cdot w_{i+1} \geq w w + c \cdot i\delta$$

$$\|w_{i+1}\| = \|w\| \|w_{i+1}\| \geq w \cdot w_{i+1} \geq ww + c \cdot i\delta \quad \leftarrow \text{lower bound}$$

Part two

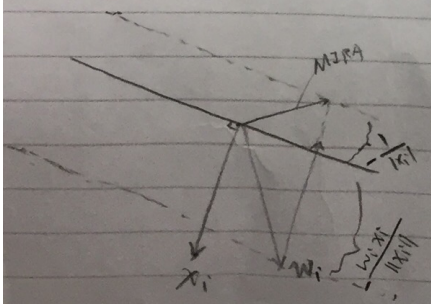
$$w_{i+1} = w_i + c y_i x_i$$
$$\|w_{i+1}\|^2 = \|w_i + c y_i x_i\|^2$$
$$= \|w_i\|^2 + \|c x_i\|^2 + 2 w_i \cdot c y_i x_i$$
$$\leq \|w_i\|^2 + c^2 R^2$$
$$\leq \|w\|^2 + i c^2 R^2$$

Combine Part One and Two, we can prove perception coverges regardless initial weight and learning rate.

2. 对 x_i is a negative example, MIRA make sure $w \cdot x_i = 1$

margin is $\frac{1}{\|x_i\|}$

so the distance is $\frac{-1 - w_i x_i}{\|x_i\|}$



because $y_i = -1$,

$$w_{i+1} = w_i + \frac{y_i - w_i x_i}{\|x_i\|^2} \cdot \frac{x_i}{\|x_i\|}$$
$$= w_i + \frac{y_i - w_i x_i}{\|x_i\|^2} x_i$$

4.
After doing zero-mean and unit-variance, it will make data gradient descent converges much faster. Will make the graph easier to read.

5.
According to the naive implementations, it will do w'<-w' + w every time, so that the running time will be DT times. However, if we do smart average perceptron, we only need to run the branch when it need to be updated. So, it will be run U times. U<<TD, so that smart will be less than naive.

DEBRIEF SECTION (required)

1. Did you work alone, or did you discuss with other students?
   If the latter please write down their names.
   Note: in general, only high-level discussions are allowed.

Discuss with Kaiwen Zheng, Haoyu Zhang

2. How many hours did you spend on this assignment?

8 hours

3. Would you rate it as easy, moderate, or difficult?

Difficult

4. Are the lectures too fast, too slow, or just in the right pace?

Too fast

5. Any other comments?

Notations in PowerPoint are not very clear. If it can be detailed will be better.