

CSE 404: Introduction to Machine Learning (Fall 2019)
Homework #2

1. (10 points) Problem 1.2 in LFD (page 33).

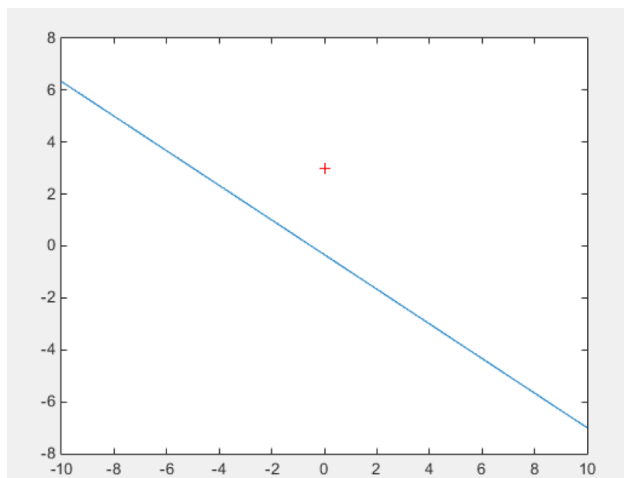
Solution:

(a) From $h(x) = +1$, we can get $w_1x_1 + w_2x_2 + w_0 > 0$, which means $x_2 \geq -(w_1x_1 + w_0)/w_2$ (if $w_2 \neq 0$)

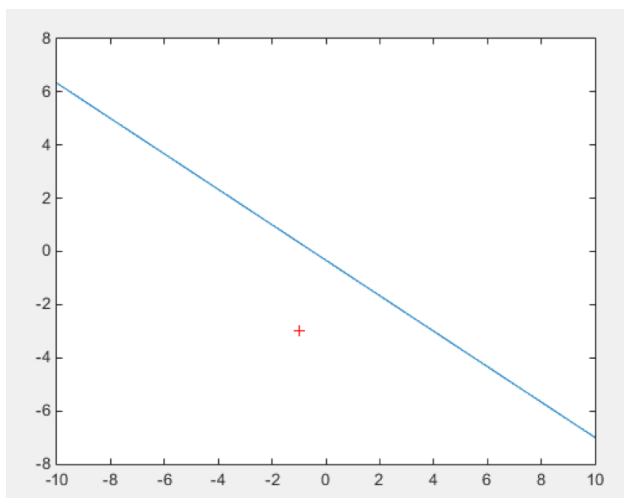
The boundary is the line given by $x_2 = -\frac{w_0}{w_2} - \frac{w_1}{w_2}x_1$

So the slope $a = -w_1/w_2$, $b = -w_0/w_2$ (if $w_2 \neq 0$)

(b) If $w = [1, 2, 3]^T$, we can get $x_2 = -2/3x_1 - 1/3$



If $w = -[1, 2, 3]^T$, we can get $x_2 = -2/3x_1 - 1/3$

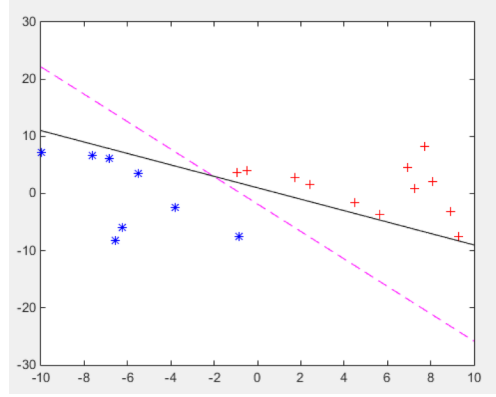


2. (20 points) Problem 1.4 (a-e) in LFD (page 34).

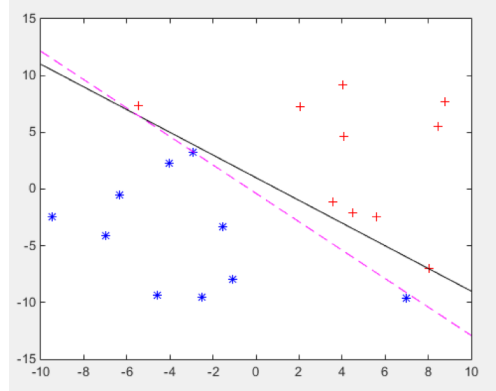
Solution:

Assume target function f is $y = -x + 1$.

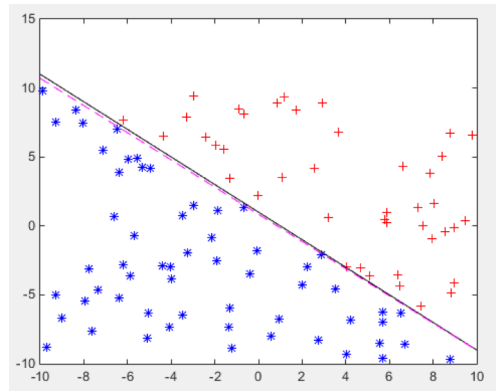
(a)(b) Target function f is the solid line, the final hypothesis g is the magenta dashed line. The PLA starting from $w(0) = 0$ took 12 updates to converge. The g is close to but not identical to f . (c) Using a different random data set: As one can observe, the number of



updates can depend heavily on data set. Update number is 33. (d) As one can observe, the



number of updates increases, typically around 100–200. (e) As one can observe, the number



of updates increases, typically around 500–1000. We conclude that the number of iterations is increasing with N . This is because as N increases, the separation between the data decreases. As the separation between the data decreases, the PLA will generally take more iterations.

