Exercises 5.10

Exercise 5.10.1

Recall that $\neg x$ is the negative of the Boolean variable x.

a. Show that a single perceptron can learn the Boolean function $y = x_1 \wedge \neg x_2$, with some x_1 , $x_2 \in \{0,1\}$.

b. The same question as in part a for the Boolean function $y = x_1 \vee \neg x_2$, with some x_1 , $x_2 \in \{0,1\}$.

c. Show that a perceptron with one Boolean input, x, can learn the negation function $y = \neg x$. What about the linear neuron?

d. Show that a perceptron with three Boolean inputs, x_1 , x_2 , x_3, can learn the negation function $y = \neg x$. What about $x_1 \vee x_2 \vee x_3$?

Exercise 5.10.2

Show that two finite linearly separable sets A and B can be separated by a perceptron with rational weights.

Exercise 5.10.3

Assume the inputs to a linear neuron are independent and normaly distributed, $X_i \sim \mathcal{N}(0, \sigma_i^2),\ i = 1, \ldots, n$. Find the optimal weights, $w^\star$

a. A one dimensional random variable with zero mean, Z is learned by a linear neuron with input X. Assume the input, X, and the target Z are independent. Write the cost function and find the optimal parameters, $w^\star$. Provide an interpretation of the result.

b. Use Newton's method to obtain the optimal parametersof a linear neuron.

Exercise 5.10.4

Explain the equivalence between the linear regression algorithm an the learning of a linear neuron.

Exercise 5.10.5

Consider a neuron with a continuum input, whose output is $y = H(\int_0^1 x d\mu(x))$. Find the output in the case when the measure is $\mu = \delta_{x_0}$.

Exercise 5.10.6

Consider n points $P_1, \dots P_n$, included in a half-circle, and denote by $\mathbf{x}_1, \dots, \mathbf{x}_n$ their coordinate vectors.

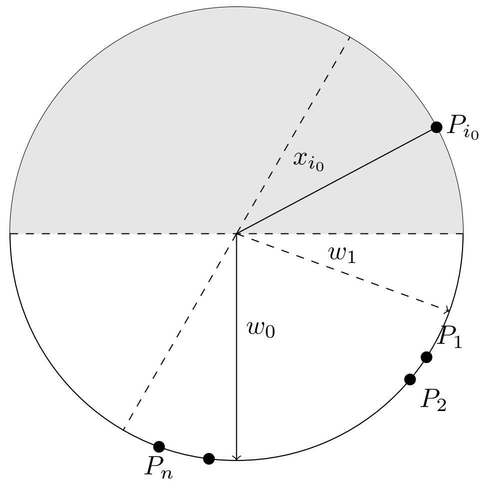

Figure 1: Perceptron algorithm

A perceptron can learn the aforementioned half-circle by the following algorithm:

1. Start from an arbitaryhalf-circle determined by its diameter and a unit normal vector w_0. Then select an incorreclty classified point, P_{i_0}, i.e. , a point for which $\langle w_0, \mathbf{x}_{i_0} \rangle < 0$. See Fig 1.
2. Rotate the diameter that the new normal is $w_1 = w_0 + \mathbf{x}_{i_0}$. Show that the point is now correclty classified.
3. Repeating the previoustwo steps, we constructed inductively the sequence of vectors $(w_m)_m$. Such that $w_{m+1} = w_m + \mathbf{x}_{i_m}$, where P_{i_m} is a point misclassified at step m. Show that the process ends in a finite number of steps, i.e. , there is a $N > 1$ such that $\langle w_0, \mathbf{x}_j \rangle > 0, \ \forall \ 1 \leq j \leq n$. Find an estimate of the number N.

Exercise 5.10.7

Modify the perceptron learning algorithm given by Exercise 5.10.6 for the case when the points $P_1, \dots P_n$ are included in a half-plane.

Exercise 5.10.8

Let $\mathbf{1}_A(x)$ denote the characteristic function of the set A, namely, $\mathbf{1}_A(x) = 1$ if $x \in A$ and $\mathbf{1}_A(x) = 0$ if $x \notin A$

a. Show that the function $\phi(x_1, x_2) = \mathbf{1}_{x_2 > x_1 + 0.5}(x_1, x_2) + \mathbf{1}_{x_2 > x_1 - 0.5}(x_1, x_2)$ implements XOR.

b. Show that the XOR function can be implement by a linear combination of the two perceptrons.

Exercise 5.10.9

Show that if all input vectors x_k have the same length, then the $\alpha - LMS$ algorithm minimizes the mean square error and in this case the updating rule (5.6.6) becomes the gradient descent rule.

Exercise 5.10.10

Find the weights of a Madaline with two Adaline units which implements the $XNOR$ function.

SOLUTIONS