

HW5

Amitesh Badkul

September 29, 2020

Abstract

This document contains my attempt at the homework 5 problems of the course Learning From Data (**CS156**) as taught by Professor *Yaser Abu-Mostafa, Caltech*.

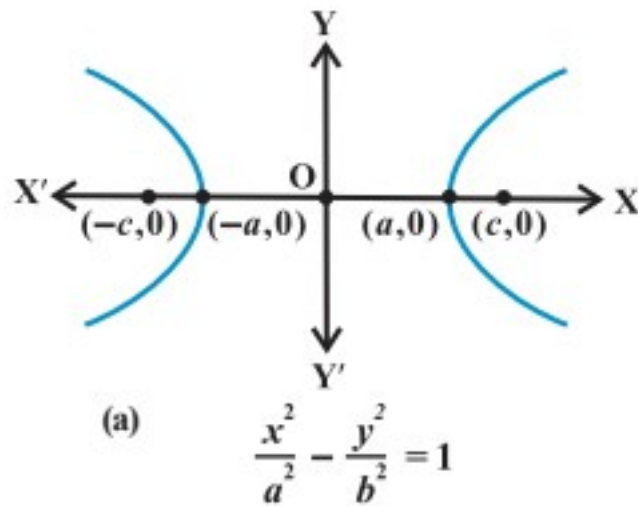
- **Linear Regression**

1.[c], Using the given equation the answer can be obtained

- **Non Linear Transform**

2.[d], the x_1 should be negative and x_2 should be positive to achieve the given graph

Fig. 3



3.[c], The VC Dimensions of the above equation is 13

- **Gradient Descent**

4.[e], self-explanatory

5.[d], a program gives the number of iterations as 10, the program calculates the gradient of the error function and tries to achieve a minima

6.[e], check the output for the written code

7.[a], An inefficient method for reaching the minima

Gradient Descent

Remember that the general form of gradient descent is:

$$\begin{aligned} & \textit{Repeat} \{ \\ & \quad \theta_j := \theta_j - \alpha \frac{\partial}{\partial \theta_j} J(\theta) \\ & \} \end{aligned}$$

We can work out the derivative part using calculus to get:

$$\begin{aligned} & \textit{Repeat} \{ \\ & \quad \theta_j := \theta_j - \frac{\alpha}{m} \sum_{i=1}^m (h_{\theta}(x^{(i)}) - y^{(i)}) x_j^{(i)} \\ & \} \end{aligned}$$

- **Logistic Regression**

8.[c], the value after 100 runs comes out to be 0.153 on average

9.[a], the number of epochs turn out to be less than 350

- **PLA vs SGD**

10.[e]