ECE 595: Homework 2

Yi Qiao, Class ID 187

(Spring 2019)

Exercise 1

(a)

Refer to code in the back.

(b)

The data was successfully read, shown by the screen below.

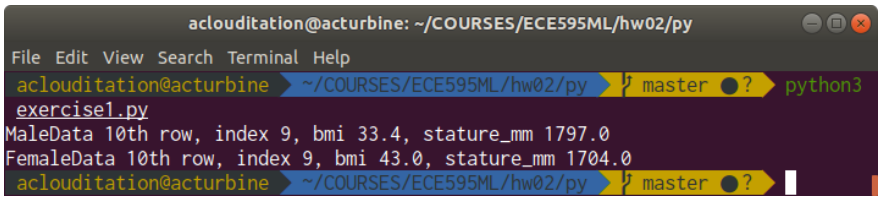


Figure 1: screenshot for reading data

Exercise 2

(a)

$$\begin{bmatrix} \boldsymbol{\omega}^* \\ \omega_0^* \end{bmatrix} = \underset{\boldsymbol{\omega},\omega_0}{argmin} \sum_{j=1}^{N} (\boldsymbol{\omega}^T \boldsymbol{x}_j + \omega_0 - y_j)^2$$

$$set\ \boldsymbol{\theta} = \begin{bmatrix} \boldsymbol{\omega}^* \\ \omega_0^* \end{bmatrix} \tag{1}$$

$$\boldsymbol{\theta}^* = \underset{\boldsymbol{\theta}}{argmin} \sum_{j=1}^{N} (\begin{bmatrix} \boldsymbol{x}_j^T & 1 \end{bmatrix} \boldsymbol{\theta} - y_j)^2$$

thus,

$$A = \begin{bmatrix} -\boldsymbol{x}_1- & 1 \\ -\boldsymbol{x}_2- & 1 \\ \dots & \dots \\ -\boldsymbol{x}_N- & 1 \end{bmatrix} \quad \boldsymbol{b} = \begin{bmatrix} y_1 \\ y_2 \\ \dots \\ y_N \end{bmatrix}$$

(b)

by least square

$$\begin{aligned} \boldsymbol{A}^T \boldsymbol{A} \boldsymbol{\theta}^* &= \boldsymbol{A}^T \boldsymbol{b} \\ \boldsymbol{\theta}^* &= (\boldsymbol{A}^T \boldsymbol{A})^{-1} \boldsymbol{A}^T \boldsymbol{b} \end{aligned} \tag{2}$$

if $A^T A$ is invertible, A needs to be full column rank (or $null(A) = 0$). TODO: find out how to avoid this issue

(c)

Refer to the code in the back.
The optimal weight θ^* is:

$$\boldsymbol{\theta}^* = \begin{bmatrix} -1.23396767e-2 \\ 6.67486843e-3 \\ -1.07017505e+1 \end{bmatrix}$$

(d)

Refer to the code in the back again.
The weight vector computed using CVXPY is the same as the one computed in the previous step.

$$\boldsymbol{\theta}^* = \begin{bmatrix} -1.23396767e-2 \\ 6.67486843e-3 \\ -1.07017505e+1 \end{bmatrix}$$

Exercise 3

(a)

(i)

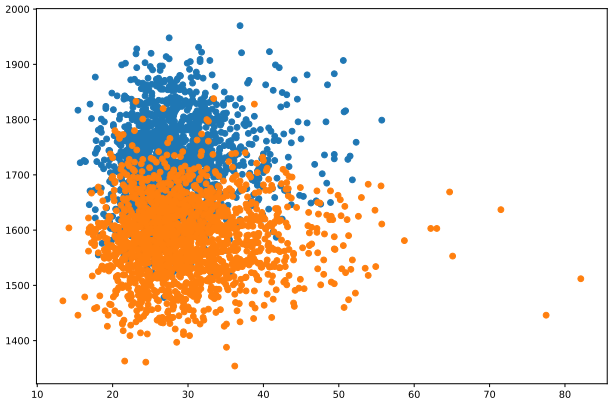


Figure 2: Training data

(ii)

$$\begin{aligned} \boldsymbol{\omega}^{*T}\boldsymbol{x} + \omega_0^* &= 0 \\ \omega_1^* x_1 + \omega_2^* x_2 + \omega_0^* &= 0 \\ x_2 &= -\frac{\omega_1^* x_1 + \omega_0^*}{\omega_2^*} \end{aligned} \tag{3}$$

(iii)

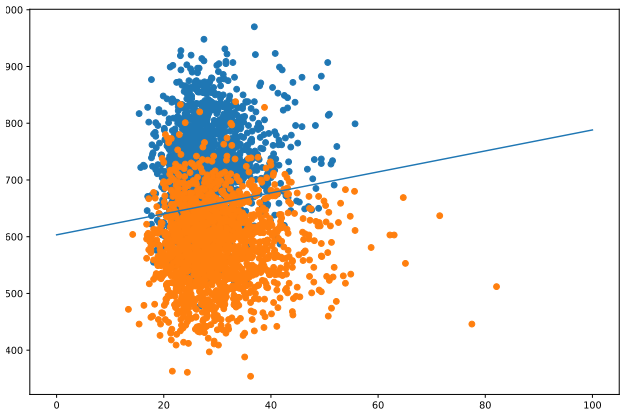


Figure 3: Training data with decision boundary

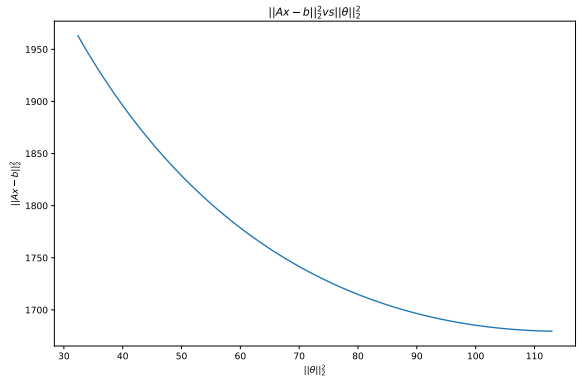
(b)

Refer to the code in the back, the success rate is:

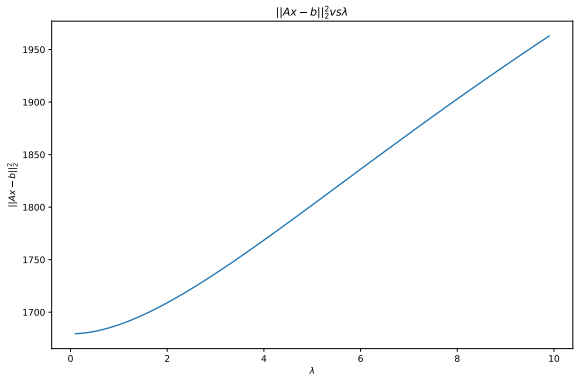
$$success\ rate = 83.93\%$$

Exercise 4

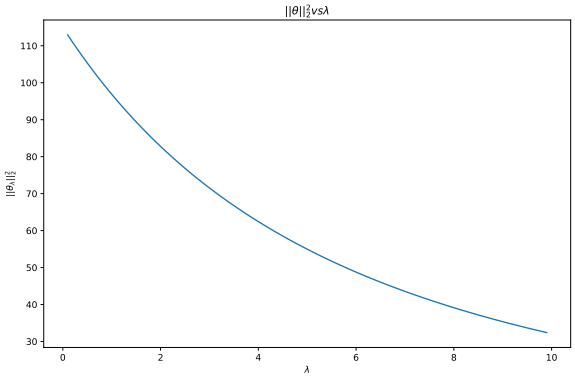
(a)



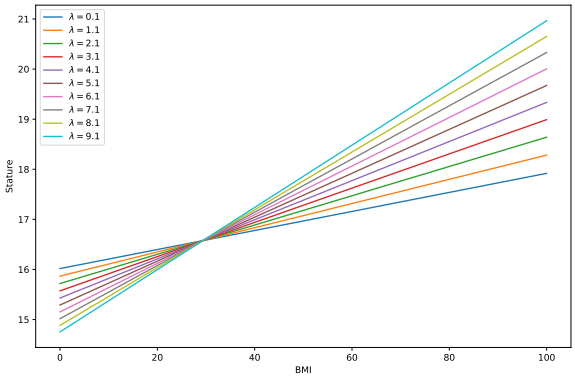
(a) 1



(b) 2



(c) 3



(d) 4

(b)

for equation (8) idk
for equation (9), $\boldsymbol{\theta}_\alpha = \underset{\boldsymbol{\theta}}{argmin} ||\boldsymbol{A\theta} - \boldsymbol{b}||_2^2 \; subject \; to \; ||\boldsymbol{\theta}||_2^2 \leq \alpha$:

$$\begin{aligned} \mathcal{L}(\boldsymbol{\theta}, \lambda) &= ||\boldsymbol{A\theta} - \boldsymbol{b}||_2^2 - \lambda(||\boldsymbol{\theta}||_2^2 - \alpha) \\ let \; \nabla\mathcal{L}(\boldsymbol{\theta}, \lambda) &= 0 \\ \nabla||\boldsymbol{A\theta} - \boldsymbol{b}||_2^2 - \lambda\nabla(||\boldsymbol{\theta}||_2^2 - \alpha) &= 0 \end{aligned} \tag{4}$$

KKT conditions:

$$\nabla_\theta||\boldsymbol{A\theta} - \boldsymbol{b}||_2^2 + \boldsymbol{\lambda}\nabla_\theta(||\boldsymbol{\theta}||_2^2 - \alpha) = 0 \tag{5}$$

for equation (10), $\boldsymbol{\theta}_\epsilon = \underset{\boldsymbol{\theta}}{argmin} ||\boldsymbol{\theta}||_2^2 \; subject \; to \; ||\boldsymbol{A\theta} - \boldsymbol{b}||_2^2 \leq \epsilon$