Name: Yuxuan Zhang (yuxuanz8)

University of Illinois Spring 2020

CS 446/ECE 449 Machine Learning
Homework 1: Linear Regression

Due on Thursday February 6 2020, noon Central Time

1. **[17 points]** Linear Regression

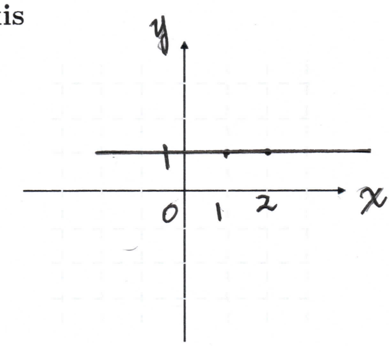
We are given a dataset $\mathcal{D} = \{(1,1),(2,1)\}$ containing two pairs (x,y), where each $x \in \mathbb{R}, y \in \mathbb{R}$ denotes a real-valued number.

We want to find the parameters $w = \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} \in \mathbb{R}^2$ of a linear regression model $\hat{y} = w_1 x + w_2$ using

$$\min_w \frac{1}{2} \sum_{(x,y)\in\mathcal{D}} \left(y - w^\top \begin{bmatrix} x \\ 1 \end{bmatrix} \right)^2. \tag{1}$$

(a) (2 points) Plot the given dataset and find the optimal w^* by inspection.

Your answer: **Mark the axis**



$$w^* = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

(b) (4 points) Using general matrix vector notation, the program in Eq. (1) is equivalent to

$$\min_w \frac{1}{2} \|\mathbf{y} - \mathbf{X}w\|_2^2. \tag{2}$$

Specify the dimensions of the introduced matrix $\mathbf{X}$ and the introduced vector $\mathbf{y}$. Also write down explicitly the matrices and vectors using the values in the given dataset $\mathcal{D}$.

Your answer: $X: 2\times 2$. $y: 2\times 1$ $w = \begin{pmatrix} w_1 \\ w_2 \end{pmatrix}$

$$X = \begin{pmatrix} x_1 & 1 \\ x_2 & 1 \end{pmatrix} \qquad y = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

$$= \begin{pmatrix} 1 & 1 \\ 2 & 1 \end{pmatrix}$$

Name: Yuxuan Zhang (yuxuanz8)

(c) (3 points) **Derive** the general analytical solution for the program given in Eq. (2). Also plug in the values for the given dataset $\mathcal{D}$ and compute the solution numerically.

Your answer: $\min_w \frac{1}{2}\|y - Xw\|_2^2$

$$Q(w) = \frac{1}{2}\|y - Xw\|^2 = \frac{1}{2}(y^T - w^T X^T)(y - Xw)$$
$$= \frac{1}{2}(y^T y - 2y^T X w + w^T X^T X w)$$
$$\frac{\partial Q(w)}{\partial w} = \frac{1}{2}(-2X^T y + 2X^T X w) = 0$$
$$X^T X w = X^T y$$
$$w = (X^T X)^{-1} X^T y$$

$$X^T y = \begin{pmatrix} 1 & 2 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 3 \\ 2 \end{pmatrix}$$
$$X^T X = \begin{pmatrix} 1 & 2 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 2 & 1 \end{pmatrix} = \begin{pmatrix} 5 & 3 \\ 3 & 2 \end{pmatrix}$$
$$(X^T X)^{-1} = \begin{pmatrix} 2 & -3 \\ -3 & 5 \end{pmatrix}$$
$$w^* = (X^T X)^{-1} X^T y = \begin{pmatrix} 2 & -3 \\ -3 & 5 \end{pmatrix}\begin{pmatrix} 3 \\ 2 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

(d) (1 point) Numerous ways exist to compute this solution via PyTorch. Read the docs for the functions 'torch.gels', 'torch.gesv', and 'torch.inverse'. Use all three approaches when completing the file A1_LinearRegression.py and verify your answer. Which solution provides the most accurate value for w_1 for our dataset?

Your answer: The most accurate solution is provided by Solution 2 and 3. I used lstsq ~~to~~ instead gels. ~~This got $\begin{pmatrix} -0.0 \\ 1.0 \end{pmatrix}$ Solution 3 get a float~~ $\begin{pmatrix} 0.0 \\ 1.0 \end{pmatrix}$

it returns ~~float number~~ a $\begin{pmatrix} -4.77e\text{-}07 \\ 1.000e\text{+}00 \end{pmatrix}$

(e) (6 points) We are now given a dataset $\mathcal{D} = \{(0,0),(1,1),(2,1)\}$ of pairs (x,y) with $x, y \in \mathbb{R}$ for which we want to fit a quadratic model $\hat{y} = w_1 x^2 + w_2 x + w_3$ using the program given in Eq. (2). Specify the dimensions of the matrix $\mathbf{X}$ and the vector $\mathbf{y}$. Also write down explicitly the matrix and vector using the values in the given dataset. Find the optimal solution w^* and draw it together with the dataset into a plot.

Your answer: **Mark the axis**

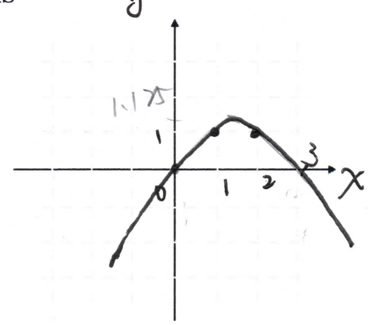
X: 3x3.

y: 3x1.

$$X = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 1 & 1 \\ 4 & 2 & 1 \end{pmatrix}$$

$$y = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}$$

$$w^* = (X^T X)^{-1} X^T y$$



$$X^T y = \begin{pmatrix} 0 & 1 & 4 \\ 0 & 1 & 2 \\ 1 & 1 & 1 \end{pmatrix}\begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 5 \\ 3 \\ 2 \end{pmatrix}$$
$$(X^T X) = \begin{pmatrix} 0 & 1 & 4 \\ 0 & 1 & 2 \\ 1 & 1 & 1 \end{pmatrix}\begin{pmatrix} 0 & 0 & 1 \\ 1 & 1 & 1 \\ 4 & 2 & 1 \end{pmatrix} = \begin{pmatrix} 17 & 9 & 5 \\ 9 & 5 & 3 \\ 5 & 3 & 3 \end{pmatrix}$$

$$w^* = (X^T X)^{-1} X^T y$$
$$= \begin{pmatrix} \frac{3}{2} & -3 & \frac{1}{2} \\ -3 & \frac{13}{2} & -\frac{3}{2} \\ \frac{1}{2} & -\frac{3}{2} & 1 \end{pmatrix}\begin{pmatrix} 5 \\ 3 \\ 2 \end{pmatrix}$$
$$= \begin{pmatrix} -\frac{1}{2} \\ \frac{3}{2} \\ 0 \end{pmatrix}$$

(f) (1 points) Complete A2_LinearRegression2.py and verify your reply for the previous answer. How did you specify the matrix $\mathbf{X}$?

Your answer: X= torch.Tensor([0, 0, 1], [1, 1, 1], [4, 2, 1])

$$X = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 1 & 1 \\ 4 & 2 & 1 \end{pmatrix} \quad w^* = \begin{pmatrix} -5e\text{-}01 \\ 1.5e\text{+}00 \\ -2.91e\text{-}07 \end{pmatrix}$$

Previous answer is right.