Name:

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.

   Solution: line through points $(x,y) = (1,1)$ and $(x,y) = (2,1)$

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

   (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 $\mathcal{D}$.

   Solution: $\mathbf{X} \in \mathbb{R}^{2\times 2}$, $\mathbf{y} \in \mathbb{R}^2$

   $\mathbf{y} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$, $\mathbf{X} = \begin{bmatrix} 1 & 1 \\ 2 & 1 \end{bmatrix}$

   (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.

   Solution: Compute the derivative and set to zero: $\mathbf{X}^T\mathbf{X}w - \mathbf{X}^T y = 0$

   Solution: $w^* = \left(\mathbf{X}^T\mathbf{X}\right)^{-1}\mathbf{X}^T y = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$

   (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?

   Solution:
   torch.gesv

   (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.

Name:

Solution:
$\mathbf{X} \in \mathbb{R}^{3\times 3}$, $\mathbf{y} \in \mathbb{R}^3$

$$\mathbf{y} = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}, \mathbf{X} = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 1 & 1 \\ 4 & 2 & 1 \end{bmatrix}$$

$$w^* = \begin{bmatrix} -0.5 \\ 1.5 \\ 0 \end{bmatrix}$$

Parabola through all three points in the dataset $\mathcal{D}$

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

Solution:

```
X = torch.Tensor([[0,0,1],[1,1,1],[4,2,1]])
```