NCTU Pattern Recognition, Homework 1

Deadline: April 3, 23:55

Part. 1, Coding (60%):

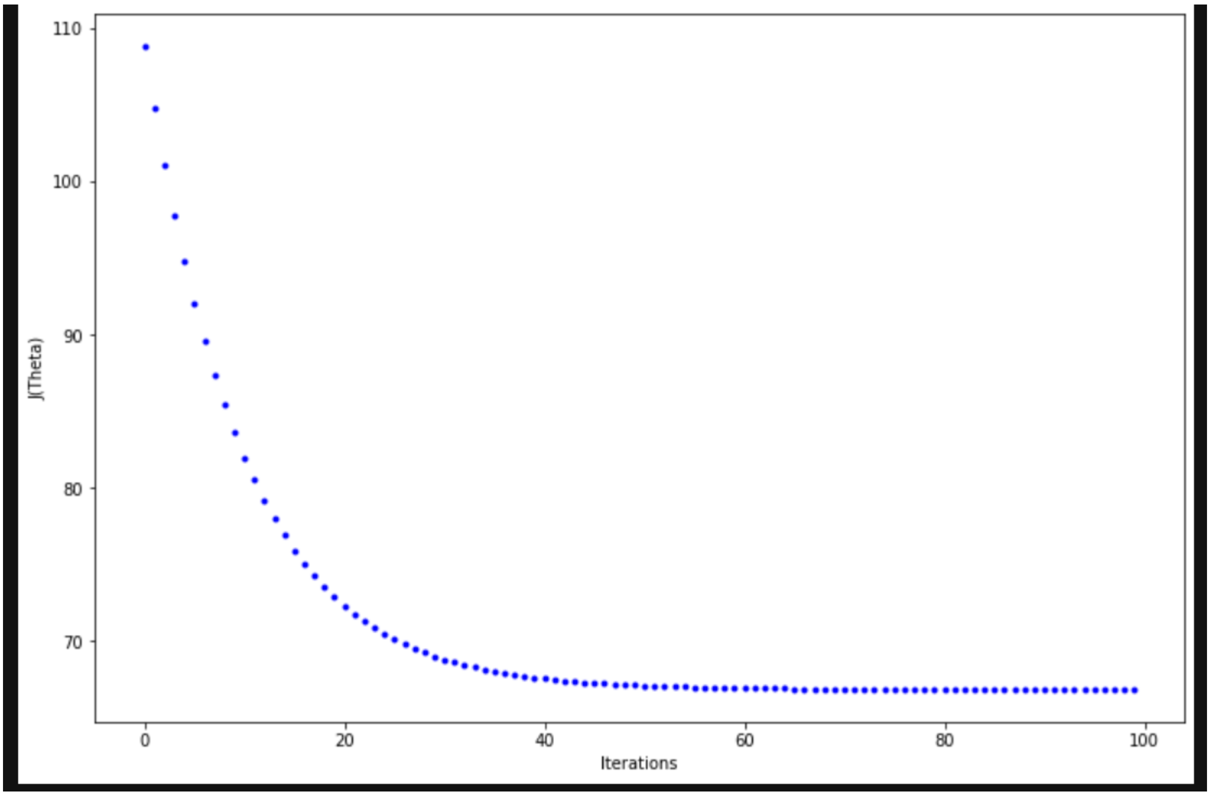
In this coding assignment, you need to implement linear regression by using only NumPy, then train your implemented model using **Gradient Descent** by the provided dataset and test the performance with testing data. Find the sample code and data on the GitHub page https://github.com/NCTU-VRDL/CS_DCP3121/tree/master/HW1

We suggest using the hyper-parameters below:

- Loss function: Mean Square Error
- Learning rate: 1e-4
- Number of training iteration: 100

Please note that only NumPy can be used to implement your model, you will get no points by simply calling sklearn.linear_model.LinearRegression. Moreover, please train your regression model using Gradient Descent, not the closed-form solution.

1. (15%) Plot the learning curve of the training, you should find that loss decreases after a few iterations (x-axis=iteration, y-axis=loss, Matplotlib or other plot tools is available to use)



2. (15%) What's the Mean Square Error of your prediction and ground truth (prediction=model(x_test), ground truth=y_test)
   A:
   MSE 是 prediction 跟 ground 的差的平方和平均，所以是 3.449*2/100=**0.06898**

(每一次跑會有點不一樣，因為一開始 theta 是隨機)

3. (15%) What're the weights and intercepts of your linear model?
   A:
   Weights:0.784
   Intercepts:0.815
   (每一次跑會有點不一樣，因為一開始 theta 是隨機)
4. (10%) What's the difference between Gradient Descent, Mini-Batch Gradient Descent, and Stochastic Gradient Descent?

   A:三種方法主要的不同是在更新的頻率，gradient descent 是一個 iteration 更新一次，mini-batch 則是一個 batch 就更新一次，所以一個 iteration 總共會更新(n/batch size)次，Stochastic 是一筆 data 就更新一次，所以一個 iteration 會更新 n 次。(n 代表 train_data 的 size)

5. (5%) All your codes should follow the PEP8 coding style and with clear comments

Part. 2, Questions (40%):

1. (20%) Suppose that we have three colored boxes R (red), B (blue), and G (green). Box R contains 3 apples, 4 oranges, and 3 guavas, box B contains 2 apples, 0 orange, and 2 guavas, and box G contains 12 apples, 4 oranges, and 4 guavas. If a box is chosen at random with probabilities p(R)=0.2, p(B)=0.4, p(G)=0.4, and a piece of fruit is removed from the box (with equal probability of selecting any of the items in the box), then what is the probability of selecting guava? If we observe that the selected fruit is in fact an apple, what is the probability that it came from the blue box?
2. (20%) Using the definition □□□[□] = □[(□(□) − □[□(□)])^2] show that □□□[□(□)] satisfies □□□[□] = □[□(□)^2] − □[□(□)]^2
   (這轉成 word 後就變這樣了)

1. $p(\text{guava}) = 0.2 \times \frac{3}{10} + 0.4 \times \frac{1}{2} + 0.4 \times \frac{4}{20} = 0.34 = 34\%$

$$p(B \mid A_p) = \frac{0.4 \times \frac{1}{2}}{0.2 \times \frac{3}{10} + 0.4 \times \frac{1}{2} + 0.4 \times \frac{12}{20}} = 0.4 = 40\%$$

2.
$$\begin{aligned}
\text{var}(f) &= E\left[\left(f(x) - E[f(x)]\right)^2 \right] \\
&= E\left[(f(x) - \mu)^2 \right] \\
&= E\left[f(x)^2 + \mu^2 - 2 \cdot \mu f(x) \right] \\
&= E\left[f(x)^2 \right] + \mu^2 - 2\mu \, E[f(x)] \\
&= E\left[f(x)^2 \right] - \mu^2 \\
&= E\left[f(x)^2 \right] - E[f(x)]^2
\end{aligned}$$