

DS5220: Problem Set 2

Linear Regression and Gradient Descent

Due: 12 Oct 2022

1. Write a python code whose input is a training dataset $(x_1, y_1), \dots, (x_N, y_N)$ and its output is the weight vector θ in the model $y = \theta^T \phi(x)$ for a non-linear mapping $\phi()$. Consider mean squared error as loss function. Implement two cases i) closed-form solution, and ii) using stochastic gradient descent on mini-batches of size m . [25 Points]
 - (a) Download the train data and apply closed-form solution on train data and with learned weights, plot regression line over training data. [2 Points]
 - (b) Apply gradient descent with mini-batch of size 10. Plot iterations vs. θ . Change the mini-batch size and report its impact on learned weights. [3 Points]
 - (c) Write a predict function that takes a new test sample (x_{tst}, y_{tst}) , and learned weights θ , and outputs prediction \hat{y} . Report train and test error. [2 Points]
2. Consider n degree of polynomials, $\phi(x) = [1, x, x^2, \dots, x^n]$.
 - (a) Run the code again on training data to compute theta for $n \in \{2, 3, 5\}$, and report regression error (mean squared error) on train and test data for all n . [15 Points]
 - (b) Which value of n is the best fit for this data? Explain why. [4 Points]