



Machine Learning Foundations

HOMEWORK 3

ReadMe

R05942060 蘇宛琳



Problem7

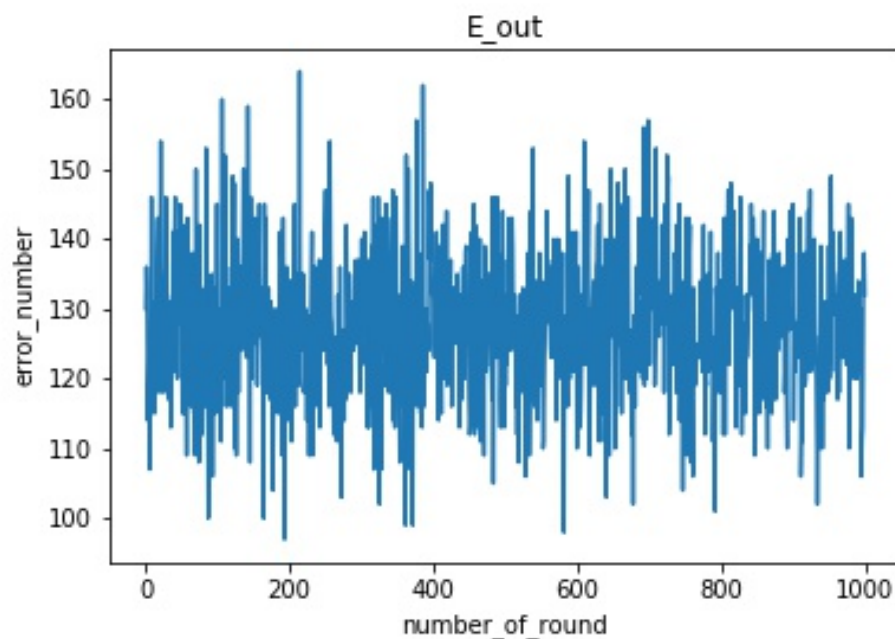
執行 linearRegression7.py 的檔案

`def generate_random_data(size)` random 產生 1000 筆 x 資料與對應的 Y，並加上 10% 的 noise。

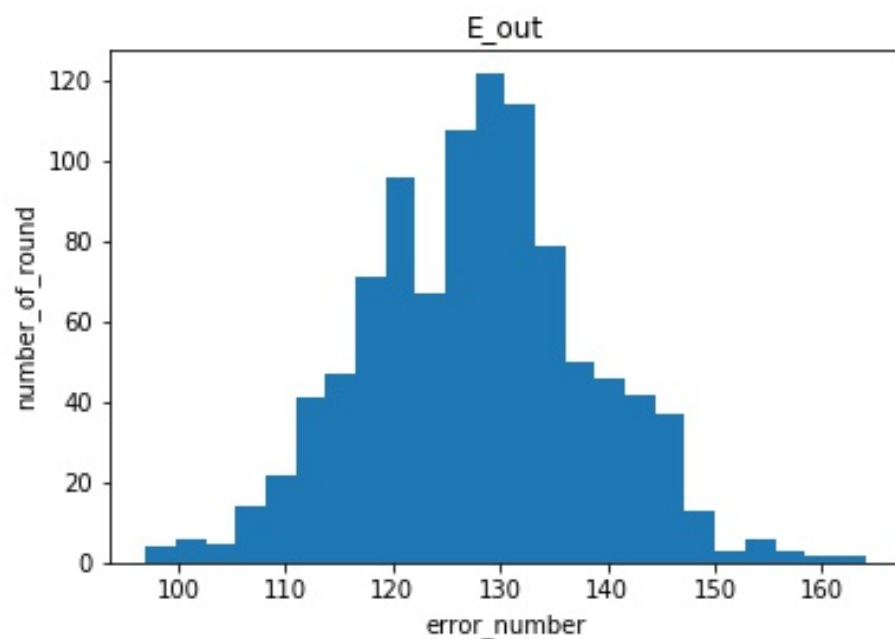
`def Z_transform(X)` 接下來將 x 轉換至 Z 平面。

使用 `np.linalg.pinv` 計算 pseudo inverse，找出 `Wlin`。

使用 `Wlin` 跑 1000 次不同的 1000 筆資料計算 `Eout`，結果如下圖：



下圖是做 Histogram 統計，從圖中可看出 `E_out` 集中在 0.12-0.14 這個區間，符合原本設定的 10% 的 noise：



Problem8,9

執行 `logisticRegression8.py` 的檔案

利用 `def load_data(file_path)` 讀取 'hw3_train_data.txt' 及 'hw3_train_test.txt' 的檔案，存進 `x_tr, y_tr, x_te, y_te` 的變數中。

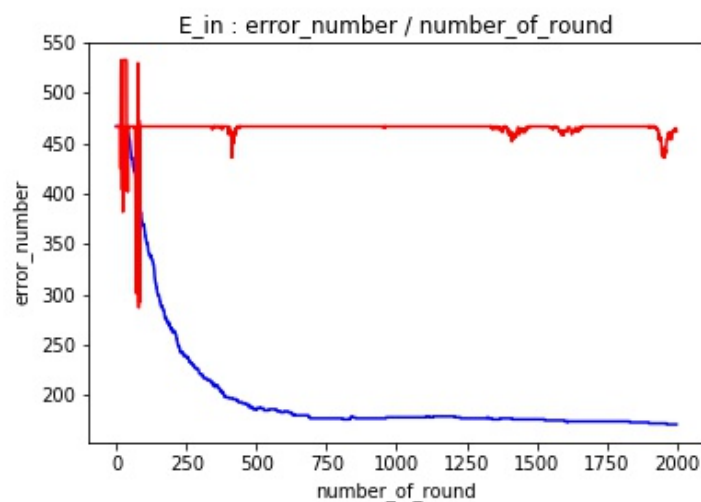
接下來進入迴圈(size : 2000)，每次迴圈都同時利用

`def logisticRegression_FixedLearningRate(x, y, eta, size)`, 及

`def logisticRegression_StochasticGradientDescent(x, y, eta, size)`

算出 Fixed learning rate 的 `w` 以及 Stochastic 的 `w`，並使用更新後的 `w` count error，最後將每次的 `E_in` `E_out` 記錄起來，連成 curve。

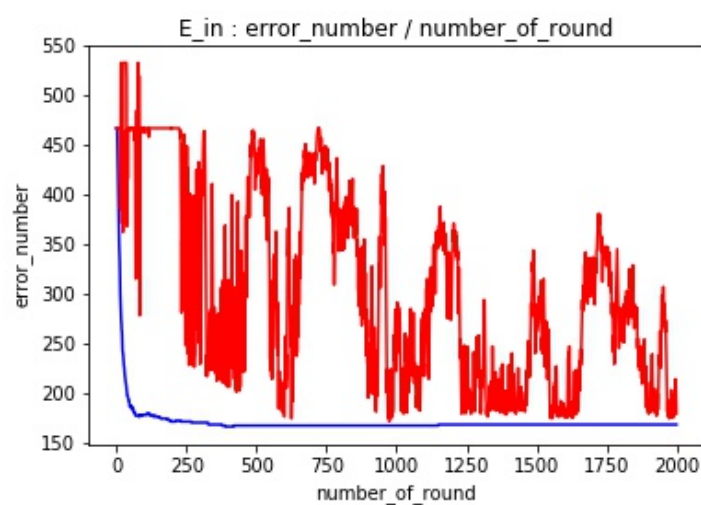
Problem8 的 `E_in` 圖形：



Blue line : Fixed learning rate

Red line : Stochastic Gradient Descent

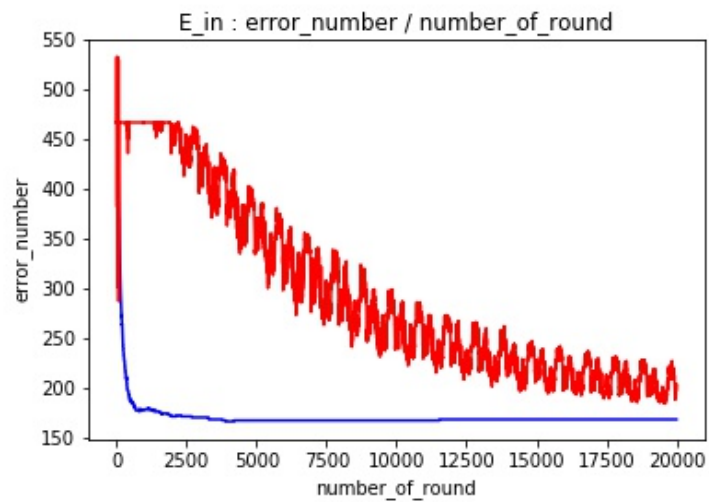
eta 設定為 0.001 / size : 2000



Blue line : Fixed learning rate

Red line : Stochastic Gradient Descent

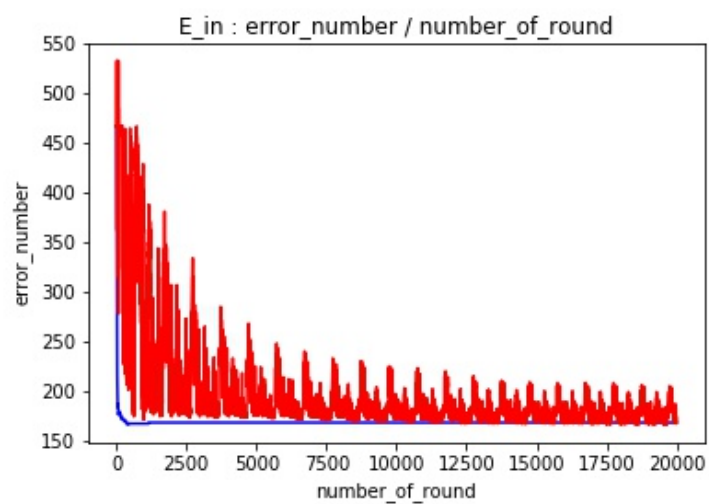
eta 設定為 0.01 / size : 2000



Blue line : Fixed learning rate

Red line : Stochastic Gradient Descent

eta 設定為 0.001 / size : 20000

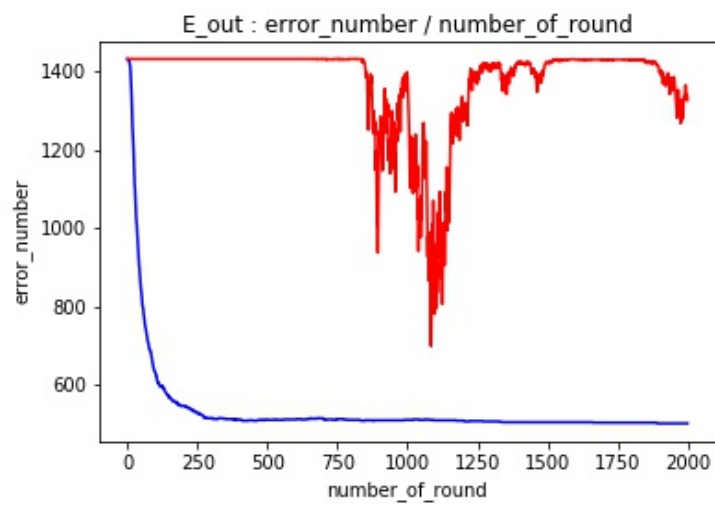


Blue line : Fixed learning rate

Red line : Stochastic Gradient Descent

eta 設定為 0.01 / size : 20000

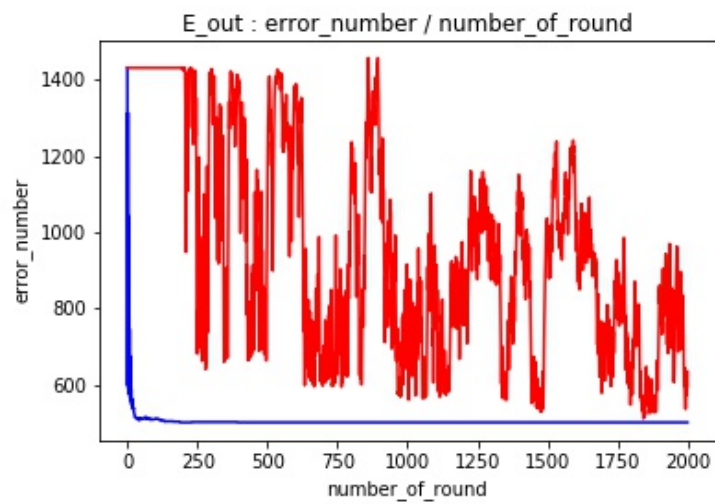
Problem9 的 E_out 圖形：



Blue line : Fixed learning rate

Red line : Stochastic Gradient Descent

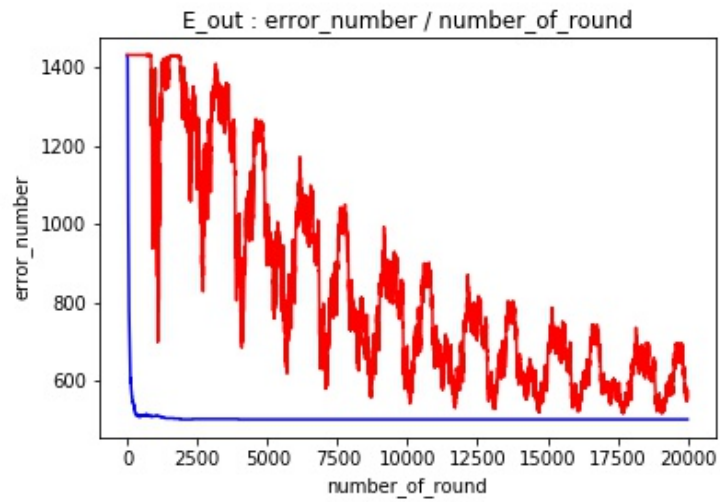
eta 設定為 0.001 / size : 2000



Blue line : Fixed learning rate

Red line : Stochastic Gradient Descent

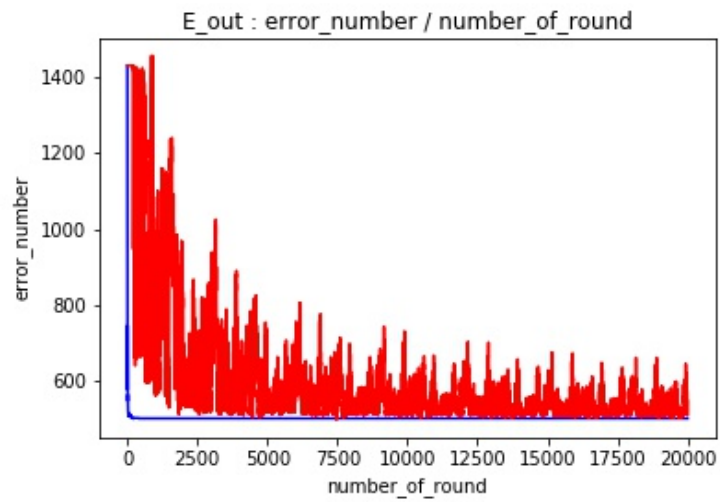
eta 設定為 0.01 / size : 2000



Blue line : Fixed learning rate

Red line : Stochastic Gradient Descent

eta 設定為 0.001 / size : 20000



Blue line : Fixed learning rate

Red line : Stochastic Gradient Descent

eta 設定為 0.01 / size : 20000