Homework 1

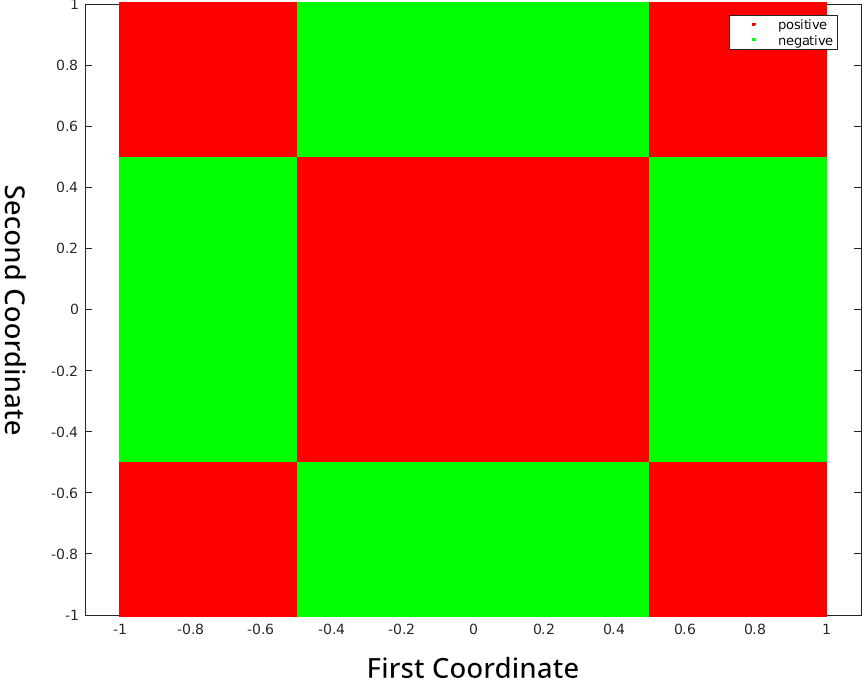
108B (5403) Machine Learning　王傳鈞　0416047

本次作業所使用到的程式碼，都已經上傳到[GitHub](https://github.com/a2468834/Machine_Learning_Grad/tree/master/HW1/MATLAB%20code)當中，若有需要參考其詳細內容，歡迎點擊連結直接前往網頁瀏覽。

檔案名稱為「Q1\_a.m」的MATLAB source code代表是第一題的(a)小題所使用到的程式碼，其餘檔案命名方式依此類推。

第一題

1. 利用MATLAB內建函數fitcknn()，我們可以得到下面這張經過1-nearest neighbor algorithm處理過後的座標圖：



在上圖當中，紅色區塊是被algorithm分類為positive的instances，綠色區塊則是被分類為negative的instances。

1. 經過3-nearest neighbor algorithm運算，classification結果如下：

|  |  |  |  |
| --- | --- | --- | --- |
| Part | Part | | Correctness |
| Predicted Class | Actual Class |
|  | negative | positive | W |
|  | negative | positive | W |
|  | negative | positive | W |
|  | negative | positive | W |
|  | negative | positive | W |
|  | positive | negative | W |
|  | positive | negative | W |
|  | positive | negative | W |
|  | positive | negative | W |

1. 3-nearest neighbor algorithm的confusion matrix為：

|  |  |  |  |
| --- | --- | --- | --- |
|  | | Predicted Class | |
| ＋ | － |
| Actual Class | ＋ | True  Positive | False  Negative |
| － | False  Positive | True  Negative |

第二題

證明： 且 是對稱正定矩陣 的所有特徵值均為正數

* where is and eigenvalue of

and such that

* 接著，我們考慮 的值：

因為 是real symmetric positive definite matrix，所以

綜合和的結果，我們得到：

* 必為一個正實數

第三題

證明還未完成

第四題

1. 20次的隨機實驗結果分別如下：

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| No. | Average | No. | Average | No. | Average | No. | Average |
| 1 | 0.5420 | 6 | 0.4857 | 11 | 0.5412 | 16 | 0.5271 |
| 2 | 0.4037 | 7 | 0.4409 | 12 | 0.6227 | 17 | 0.3614 |
| 3 | 0.4785 | 8 | 0.6096 | 13 | 0.6752 | 18 | 0.5278 |
| 4 | 0.4315 | 9 | 0.4478 | 14 | 0.3314 | 19 | 0.5932 |
| 5 | 0.3910 | 10 | 0.4941 | 15 | 0.5567 | 20 | 0.5819 |

1. 50次的隨機實驗結果分別如下：

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| No. | Average | No. | Average | No. | Average | No. | Average | No. | Average |
| 1 | 0.4946 | 11 | 0.5166 | 21 | 0.4986 | 31 | 0.5062 | 41 | 0.4830 |
| 2 | 0.4925 | 12 | 0.5030 | 22 | 0.5123 | 32 | 0.5010 | 42 | 0.5120 |
| 3 | 0.5035 | 13 | 0.5002 | 23 | 0.5230 | 33 | 0.4925 | 43 | 0.4958 |
| 4 | 0.4998 | 14 | 0.5017 | 24 | 0.4798 | 34 | 0.4907 | 44 | 0.4963 |
| 5 | 0.5031 | 15 | 0.4938 | 25 | 0.5163 | 35 | 0.4997 | 45 | 0.5076 |
| 6 | 0.5153 | 16 | 0.5139 | 26 | 0.5006 | 36 | 0.5086 | 46 | 0.4936 |
| 7 | 0.4895 | 17 | 0.5106 | 27 | 0.5116 | 37 | 0.5114 | 47 | 0.4982 |
| 8 | 0.4936 | 18 | 0.5079 | 28 | 0.4950 | 38 | 0.4889 | 48 | 0.5017 |
| 9 | 0.4879 | 19 | 0.5099 | 29 | 0.4999 | 39 | 0.5065 | 49 | 0.4968 |
| 10 | 0.5032 | 20 | 0.4991 | 30 | 0.5061 | 40 | 0.4942 | 50 | 0.4888 |

第五題

1. 取 、、，則原方程組即為
2. 求解normal equation：

、

利用Gauss-Jordan消去法，我們得到以下結果：

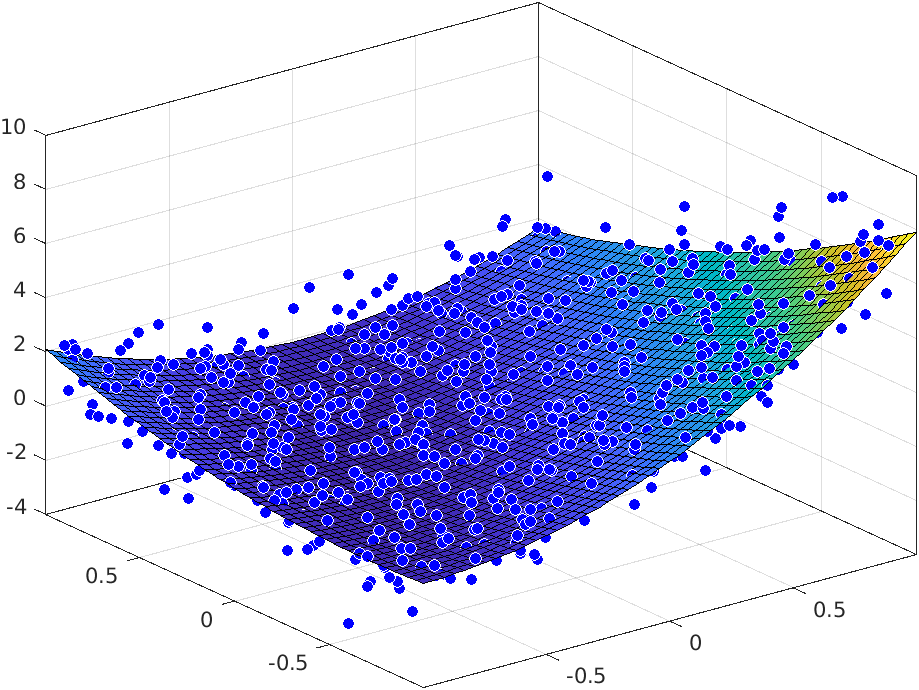
1. 因此，least squares approximation solution

第六題

1. 關於本題使用到的training dataset ，詳細請見[GitHub](https://github.com/a2468834/Machine_Learning_Grad/blob/master/HW1/MATLAB%20code/Q6_dataset.csv)連結。
2. 利用MATLAB內建函數fit()，我們可以得到以下的fitting surface：

1. 先用 計算出input data的預測值 ，其與實際值 的差異即為error；接著將1000筆資料的error取絕對值再平均，即可求得MAE。

，每筆資料的error值詳細請見[GitHub](https://github.com/a2468834/Machine_Learning_Grad/blob/master/HW1/MATLAB%20code/Q6_error_vector.csv)連結。

另外，這個fitting surface的3D圖如下：