**機器學習理論 HW2**

108064535 陳文遠

|  |
| --- |
| **Step 1. 使用 Maximum Likelihood & Least Squares 法來訓練 Training\_set.csv 檔中的資料** |

* 假設有一線性方程式，而題目給定此自變數為一個 feature vector 如下：

其中，

最後，

|  |
| --- |
| **Step 2. 將 Testing\_set.csv 中的測試資料送進 Maximum Likelihood & Least Squares 的 Model 中來預測 chance of admit 以及計算其 squared error** |

* 假設線性函數為( 為 Added Gaussian Noise)：

其機率分布可以寫成：

其中， 是我們輸入的資料(分數)， 為 Targets (chance of admit)，則我們可以獲得下方的 likelihood function：

再將 likelihood function 取對數：

最後將上述式子對 微分令為 0 來求極值：

* 接下來，將 Testing\_csv 檔中的測試資料代入 來構成 testing data 的 feature vector ()：

同樣地，

* 最後求出 predict 的值：

Squared Error 為：

|  |
| --- |
| **Step 3. 使用 Bayesian Linear Regression 法來訓練 Training\_csv 檔中的資料** |

* 根據貝式定理我們得知

假設事前機率 (priori probability) 為

假設後驗機率 (posteriori probability) 為

經推導後可得 以及 ，其中 與 **Step 1.** 的求法一樣

* 在此實驗中，我僅先將 以及 設為 1

|  |
| --- |
| **Step 4. 將 Testing\_set.csv 中的測試資料送進 Bayesian Linear Regression 的 Model 中來預測 w, chance of admit 以及計算其 squared error** |

* 權重值為 ，此題計算出的 分別為：

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| -0.18731298 | 0.07518969 | 0.03911943 | 0.49466673 | 0.35914929 |

* 接下來，將 Testing\_csv 檔中的測試資料代入 來構成 testing data 的 feature vector ()：

同樣地，

* 最後求出 predict 的值：

Squared Error 為：

|  |
| --- |
| **Step 5. 討論** |

1. **Maximum likelihood and least squares 以及 Bayesian linear regression 的差異：**

已知貝式定理如下，其中 為欲估計的數， 為輸入的資料集

Maximum likelihood and least squares 的目的是為了找出一個 來使得 likelihood 。換句話說在此方法中，他將 視為一個常數而非隨機變數，並且不需要借助 prior probability 來估計

而在 Bayesian linear regression 中會完全計算出 posterior probability ，並且在此方法中會將 視為隨機變數，倘若 的變異數足夠小，就將他的期望值視為估計值。Bayesian 法與 Maximum likelihood 法的關鍵區別就是他允許採用 prior information

1. **對結果的影響**

為了比較其影響，我分別將 Maximum likelihood 法以及 Bayesian 法對 100 的樣本估計的 squared error 進行加總，接著修改 的值來觀察何時會得到最小 squared error

從觀察的結果中會發現 對 Bayesian 法的結果影響不大，對 Maximum likelihood 法的影響較大，而當 時我可以得到最小的總 squared error

當 時，Maximum likelihood 法之 100 個樣本的 squared error 加總為 0.4363359203046397，而 Bayesian 法之 100 個樣本的 squared error 加總為 0.44922167547056285