* 所有python檔對csv檔的讀檔相對路徑都是data/filename.csv
* problem1(a)

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Coefficient | Std. Error | t-statistic | p-value |
| (intercept) | -0.178484 | 0.185403 | -0.962678 | 0.337228 |
|  | 0.085325 | 0.056887 | 1.499920 | 0.135696 |
|  | 0.004536 | 0.011481 | 0.395128 | 0.693298 |
|  | -0.090711 | 0.472928 | -0.191808 | 0.848147 |
|  | -0.057838 | 0.113619 | -0.509051 | 0.611450 |
|  | 0.141110 | 0.239256 | 0.589784 | 0.556205 |
|  | 0.033006 | 0.006104 | 5.407511 | 0.000000 |

RSE = 0.6096428946580033

=

* problem1(b)

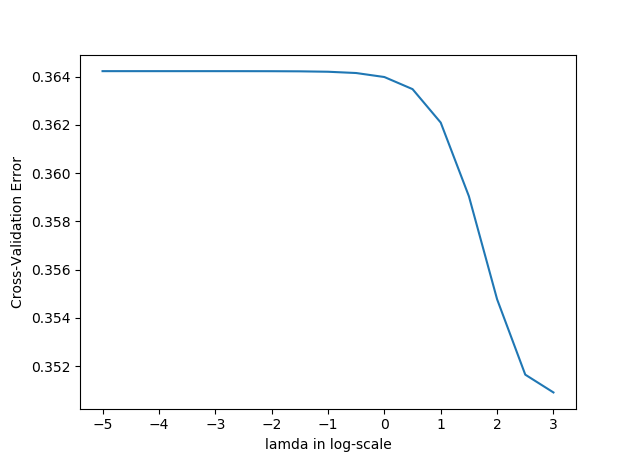
注：function有考慮intercept

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Coefficient | Std. Error | t-statistic | p-value |
| Feature1:(seafood) | 0.084306 | 0.054379 | 1.550341 | 0.123073 |
| Feature2:(obesity) | 0.034448 | 0.005323 | 6.471520 | 0.000000 |

RSE = 0.603186602943349

=

* problem1(c)



透過ridge regression，在各種lamda的搭配下，最小的RSS都是由seafood(X1) & obesity(X6)此兩個features產生。由上圖所示，當lamda = 1000.000000時，Cross-Validation Error最小。

|  |  |
| --- | --- |
|  | Coefficient |
| Feature1:(seafood) | 0.008600 |
| Feature2:(obesity) | 0.030436 |

RSE = 0.6072550960783837

=

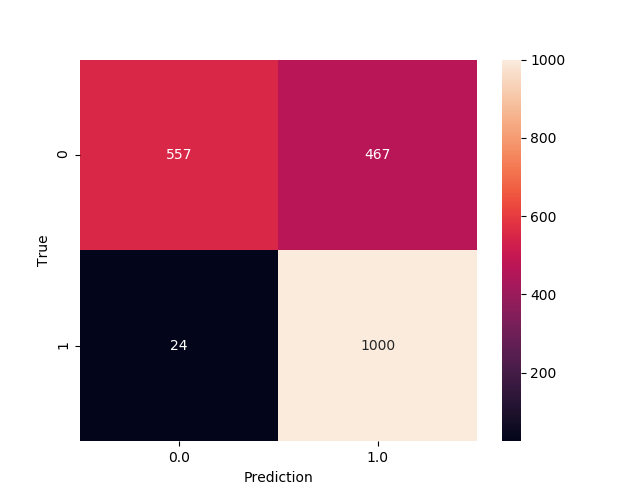
* problem2(a)

注:training data與test data的pixel value都normalize至0到1之間，若沒有經過normalization，計算上會發生overflow。

step size = 0.001

所有beta的初值皆設為0

Beta總共更新10次



* problem2(b)

注:training data與test data的pixel value都normalize至0到1之間，並且都有加上雜訊，若沒有加上雜訊，會無法計算covariance matrix的反矩陣。有加上亂數種子使每次執行結果相同。

