

2 Linear Regression

1. Feature selection

Size of Training set: 75% of dataset

Size of Validation set: 25% of dataset

M = 1

M = 2

* Weight越多, Model越複雜, 越容易overfitting

|  |  |  |  |
| --- | --- | --- | --- |
| M | 1 | 2 | 3 |
| Training RMS | 3.8366 | 3.0300 | 1.7722 |
| Validation RMS | 5.7634 | 6.5501 | 9.4487 |

Validation

RMS

Training

M

1. 每次作訓練時, 選一個feature不加入訓練, 若RMS error變大, 則該feature對於資料較為重要。

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Feature | AMB\_TEMP | CH4 | CO | NMHC | NO | NO2 | NOx | O3 | PM10 | RAINFALL |
| TrainRMS | 3.843 | 3.84 | 3.94 | 3.84 | 3.84 | 3.84 | 3.83 | 3.85 | 5.97 | 3.85 |
| ValidRMS | 5.70 | 5.79 | 6.03 | 5.80 | 5.75 | 5.76 | 5.76 | 5.73 | 6.39 | 5.82 |
| Feature | RH | SO2 | THC | WD\_HR | | WIND\_DIREC | | WIND\_SPEED | | WS\_HR |
| TrainRMS | 3.85 | 3.85 | 3.84 | 3.87 | | 3.86 | | 3.84 | | 3.84 |
| ValidRMS | 5.75 | 5.75 | 5.77 | 5.77 | | 5.77 | | 5.77 | | 5.80 |

可以發現去掉PM10做訓練的話, RMS error最大，可以推測PM10對於資料較為有重要性。

1. Maximum likelihood approach

(a) Features: PM10, CO, RAINFALL

Basis function: 先polynomial 再 sigmoid

M = 1

M = 2

(b) 4 folds

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | 1 ~ 274 | 275 ~ 548 | 549 ~ 822 | 823 ~ 1096 |
| F1 | valid | train |  |  |
| F2 |  |  |  |  |
| F3 |  |  |  |  |
| F4 |  |  |  |  |

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | F1 | | F2 | | F3 | | F4 | |
|  | Train | Valid | Train | Valid | Train | Valid | Train | Valid |
| M = 1 | 6.743 | 8.211 | 7.013 | 7.44 | 7.250 | 6.732 | 7.271 | 6.748 |
| M = 2 | 6.695 | 8.018 | 6.875 | 7.455 | 7.104 | 6.705 | 7.132 | 6.876 |
| M = 3 | 6.091 | 11.575 | 6.322 | 7.739 | 6.695 | 5.940 | 6.617 | 6.308 |
| M = 4 | 5.558 | 1851.157 | 5.453 | 6.517 | 5.718 | 7.999 | 5.513 | 5.543 |
| M = 5 | 5.024 | 2755.871 | 4.058 | 12.135 | 3.964 | 54.219 | 4.101 | 4.651 |
| M = 6 | 4.840 | 444.567 | 2.871 | 29394.821 | 3.830 | 10.808 | 2.021 | 137.833 |

從上表可以看出模型越複雜參數越多, 越容易overfitting

1. Maximum a posteriori approach

(a)

🡪

令

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | F1 | | F2 | | F3 | | F4 | |
|  | Train | Valid | Train | Valid | Train | Valid | Train | Valid |
| M = 1 | 8.618 | 11.672 | 9.004 | 11.179 | 8.957 | 11.508 | 9.309 | 13.136 |
| M = 2 | 8.237 | 11.125 | 8.548 | 12.148 | 8.562 | 10.107 | 8.831 | 15.051 |
| M = 3 | 7.321 | 11.645 | 7.670 | 21.294 | 7.849 | 13.504 | 8.001 | 19.412 |
| M = 4 | 6.368 | 7.993 | 6.646 | 12.841 | 6.935 | 8.912 | 6.861 | 13.751 |
| M = 5 | 5.935 | 7.696 | 6.141 | 10.566 | 6.471 | 7.383 | 6.334 | 11.623 |
| M = 6 | 5.775 | 7.825 | 5.899 | 7.440 | 6.276 | 6.402 | 6.122 | 9.499 |

(b)因為MAP的error function是將sum-of-squares error再做regularization, 所以有對weight加上懲罰, 比較不會overfitting, 從上面兩張表也可發現MAP沒有overfitting.