Assignment 3

107062208 邱靖豪

(1) Linear regression model for Ozone:

```
[[ 5.31731200e+01  4.43153600e+01  9.11152960e+02 -8.50679421e+07
   1.03467200e+02  0.00000000e+00]
 [ 5.67000000e+01  4.56086400e+01  9.25795760e+02 -8.49814425e+07
   1.01300000e+02  0.00000000e+00]
 [ 5.41120000e+01  4.49200000e+01  9.31605000e+02 -8.49468550e+07
   9.20768000e+01  0.00000000e+00]
 ...
 [ 5.75434800e+01  4.28377600e+01  8.11225000e+02 -8.56399430e+07
   1.13513600e+02  0.00000000e+00]
 [ 5.53630000e+01  4.43153600e+01  7.08683440e+02 -8.62138605e+07
   1.22904800e+02  5.00000000e-01]
 [ 5.13676800e+01  6.56720700e+01  8.26640000e+02 -8.55531530e+07
   1.16849600e+02  0.00000000e+00]]
```

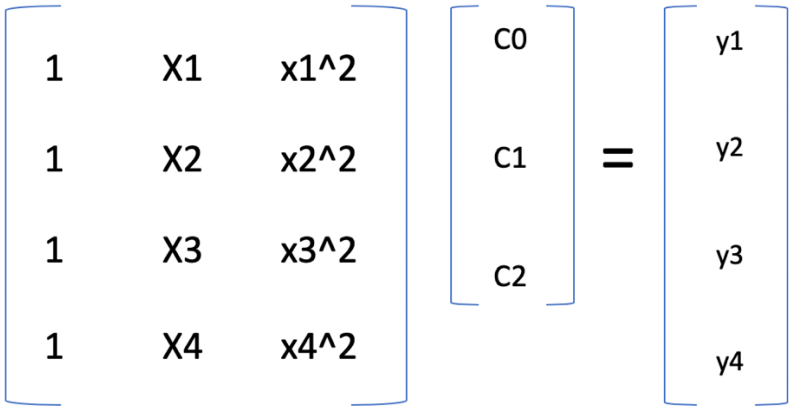
(2) Linear regression model for PM10:

```
[[-1.00806041e+02  2.93157895e+01  1.89220000e+02 -8.50679421e+07
   3.34444444e+01  2.00000000e+02]
 [-1.24387755e+02  2.91052632e+01  2.16620000e+02 -8.49814425e+07
   3.38888889e+01  2.00000000e+02]
 [-1.07400000e+02  2.92105263e+01  2.28000000e+02 -8.49468550e+07
   3.54444444e+01  2.00000000e+02]
 ...
 [-1.29624163e+02  2.98421053e+01  4.60000000e+01 -8.56399430e+07
   3.07777778e+01  2.00000000e+02]
 [-1.15804082e+02  2.93157895e+01 -3.18800000e+01 -8.62138605e+07
   2.76666667e+01  5.00000000e-01]
 [ 4.65970612e+01  2.61052632e+01  6.35000000e+01 -8.55531530e+07
   2.96666667e+01  2.00000000e+02]]
```

(3) Linear regression model for PM2.5:

```
[[ 9.52480000e+00  2.01666667e+01  2.12320000e+01 -8.50679421e+07
   2.43333333e+01  3.00000000e+01]
 [ 1.20000000e+01  1.98333333e+01  2.26970000e+01 -8.49814425e+07
   2.50000000e+01  3.00000000e+01]
 [ 9.88000000e+00  2.00000000e+01  2.33250000e+01 -8.49468550e+07
   2.73333333e+01  3.00000000e+01]
 ...
 [ 1.29792000e+01  2.10000000e+01  1.53250000e+01 -8.56399430e+07
   2.03333333e+01  3.00000000e+01]
 [ 1.07200000e+01  2.01666667e+01  1.59520000e+01 -8.62138605e+07
   1.56666667e+01  3.00000000e+00]
 [ 2.32333333e+01  1.50833333e+01  1.58000000e+01 -8.55531530e+07
   1.86666667e+01  3.00000000e+01]]
```

至於模型的實作，我是利用上課教過的 normal equation 來求，比方說，我假設方程式為二次曲線，則可依課本做法，先假設 y = ax^2 + bx + c ，再做出矩陣 Ax = B，並將每點帶入公式，比如說 x 為 WindDirec 的值，y 為 O3 的值即可求出 matrix A 和 vector B，再利用 normal equation 求出 x = (AtA)-1AtB，即為二次方程式的係數



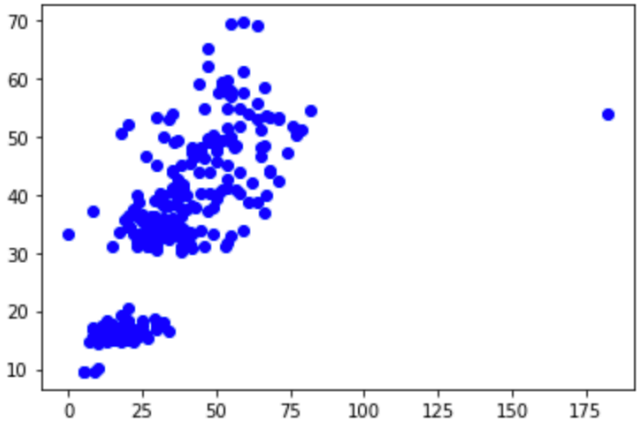
(參考課本 p247 頁做法)

(4)The validation errors of your models:

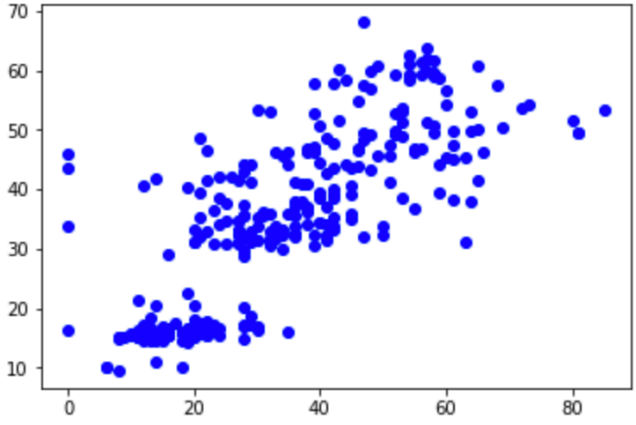
在 PM10 會是比較變動的部分，因為是 random 取值，所以有時候在那欄會超過 1.5，甚至在 2.多左右。

1. O3 多界在 1.1~1.4 間：

validation errors of O3 : 1.2590458894436705

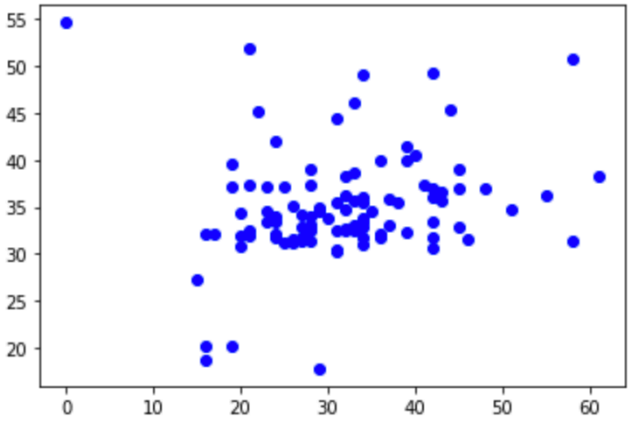


validation errors of O3 : 1.19451751451674

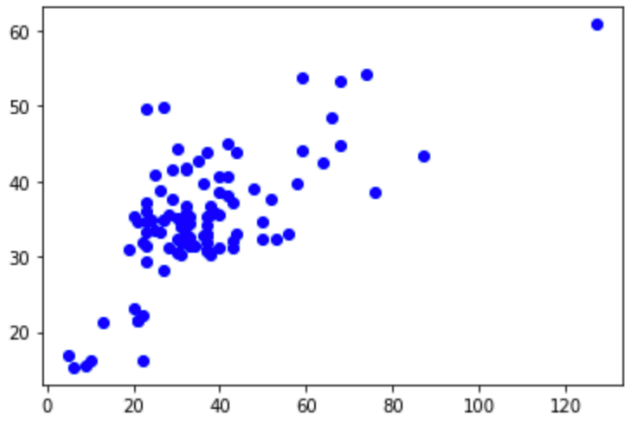


2. PM10 多界在 1.1~1.4 間

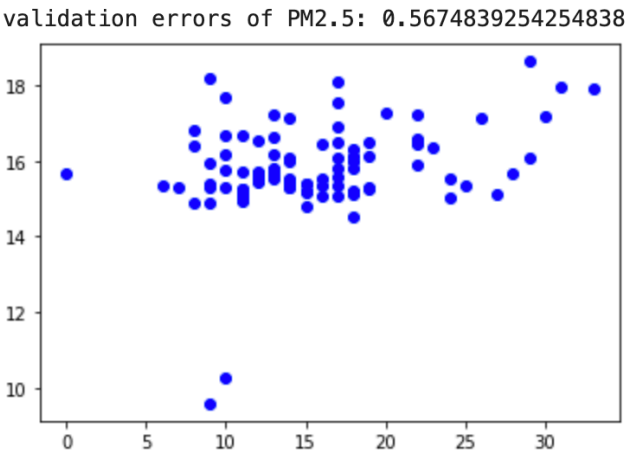
validation errors of PM10 : 1.1329715194076653

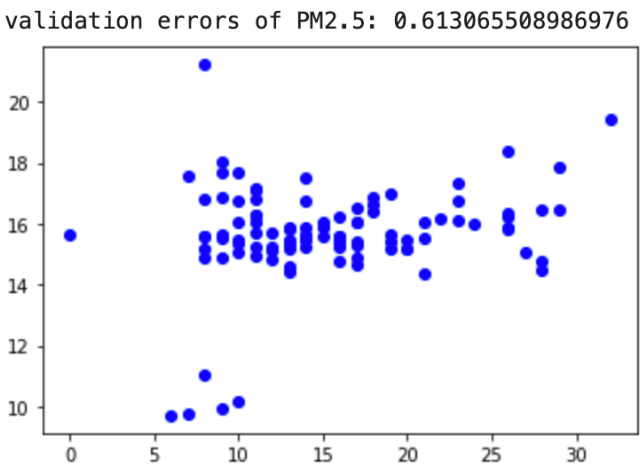


validation errors of PM10 : 1.3234940787810046



3.PM2.5 多界在 0.5~0.7





(4)The most significant function:

根據 linear least square 的解，我們只要找到掛絕對值後最大的係數，就是

每個 model 中影響最大的 function

1. O3 : WindSpeed（function 2）

```
[ 7.46553039e-01  9.12483457e-01 -2.33421386e-02  1.82371995e-08
 -1.52256397e-01  1.55900080e-01]
```

2. PM10 : WindSpeed （function 2）

```
[ 7.23424213e-03  3.52659525e-01  8.01327279e-02 -1.76241434e-07
  2.29875698e-01  1.57336980e-02]
```

3. PM2.5 : Temperature （function 3）

```
[ 8.38899711e-02  1.02673837e-01  3.03879554e-01 -5.39991950e-08
  1.08833066e-01  2.05644677e-02]
```