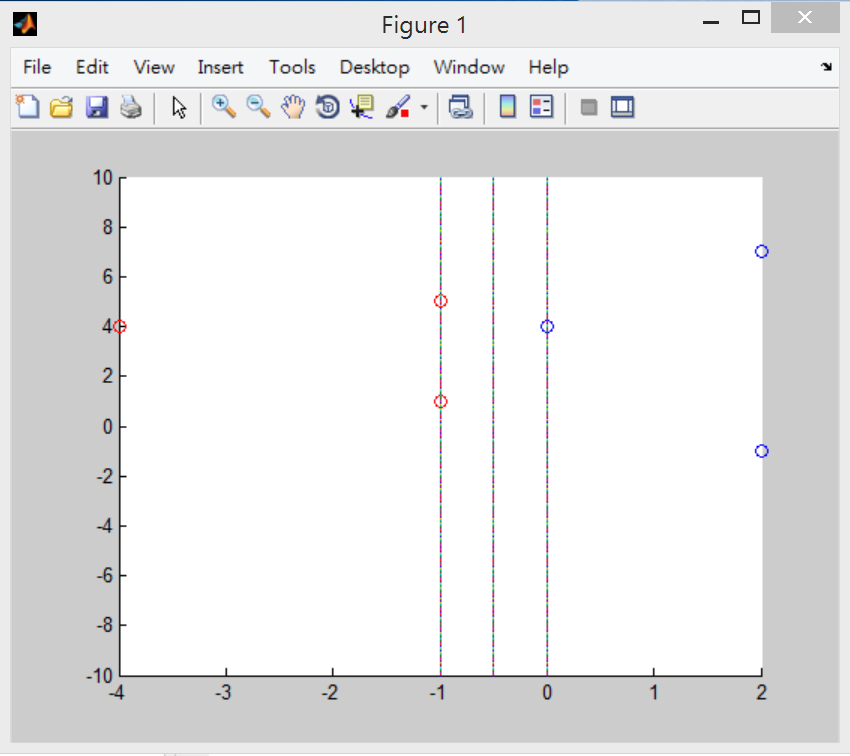
Machine Learning HW5 r01922164資工碩一 李揚

1. 第一題解出來的圖，式子在手寫部分補充

中間為分隔線，左右為margin線

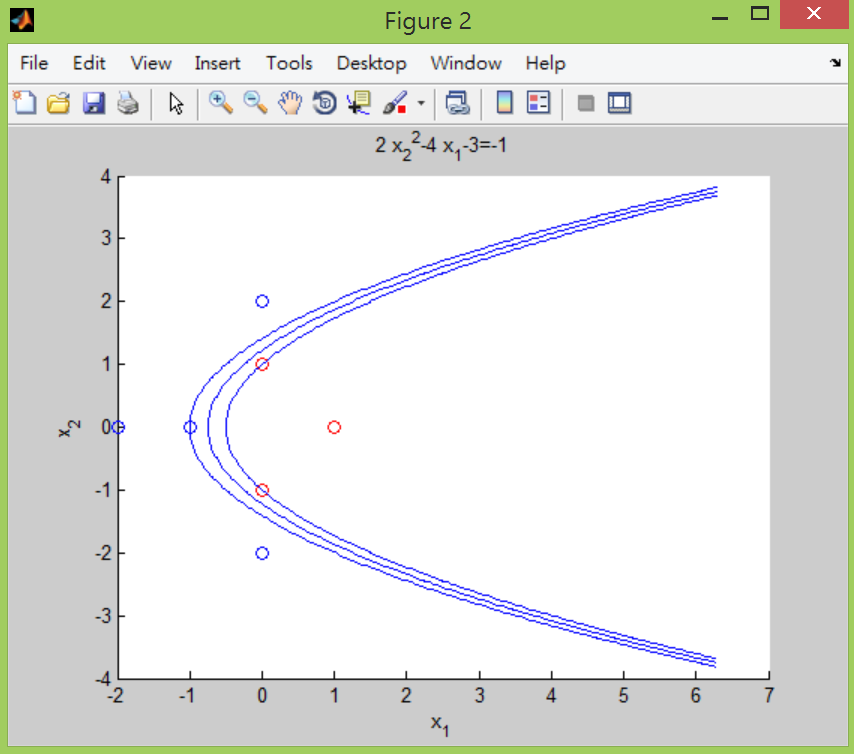


Support vectors

Z2,Z3,Z4

equation: \*Z1 + 1 = 0

1. 第二題畫出來的圖



Support vectors:

X2,X3,X4

Equation:

2X[2]^2 - 4\*X[1] – 3 = 0

第三題的答案

alpha(1) 0.0000

alpha(2) 0.0667

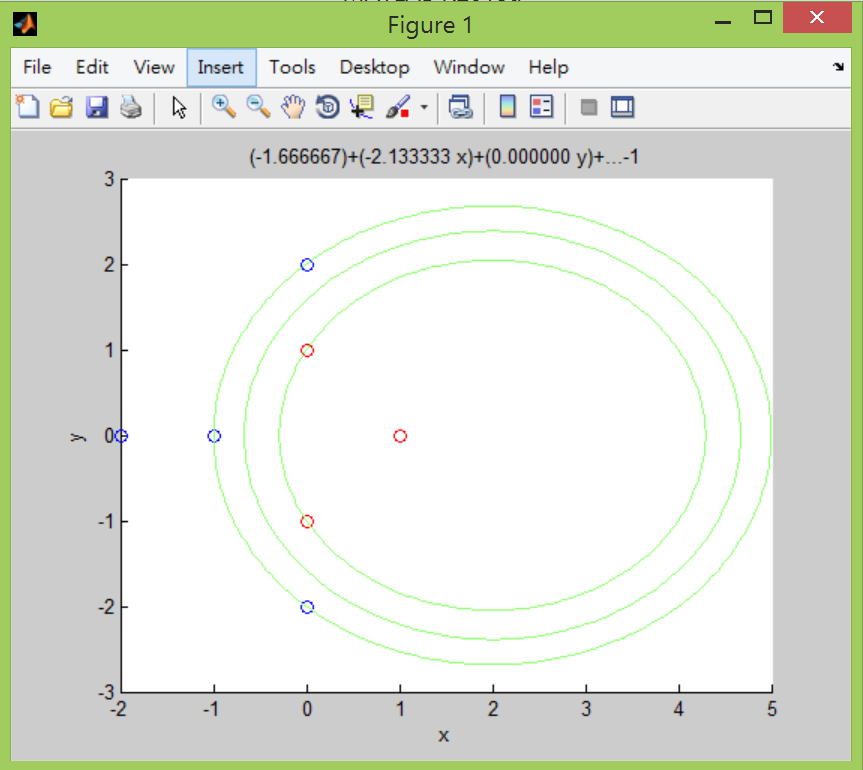
alpha(3) 0.8667

alpha(4) 0.5333

alpha(5) 0.0000

alpha(6) 0.4000

alpha(7) 0.0000

第四題畫出來的圖

Support vectors

Z2, Z3, Z4, Z6

Equation:

(-1.666667)+(-2.133333\*x)+(0.533333\*x1^2)+(0.666667\*x2^2)=0

5. 他們部會是同樣的答案，因為他們轉換到不同的geometry。

第二題是轉換到

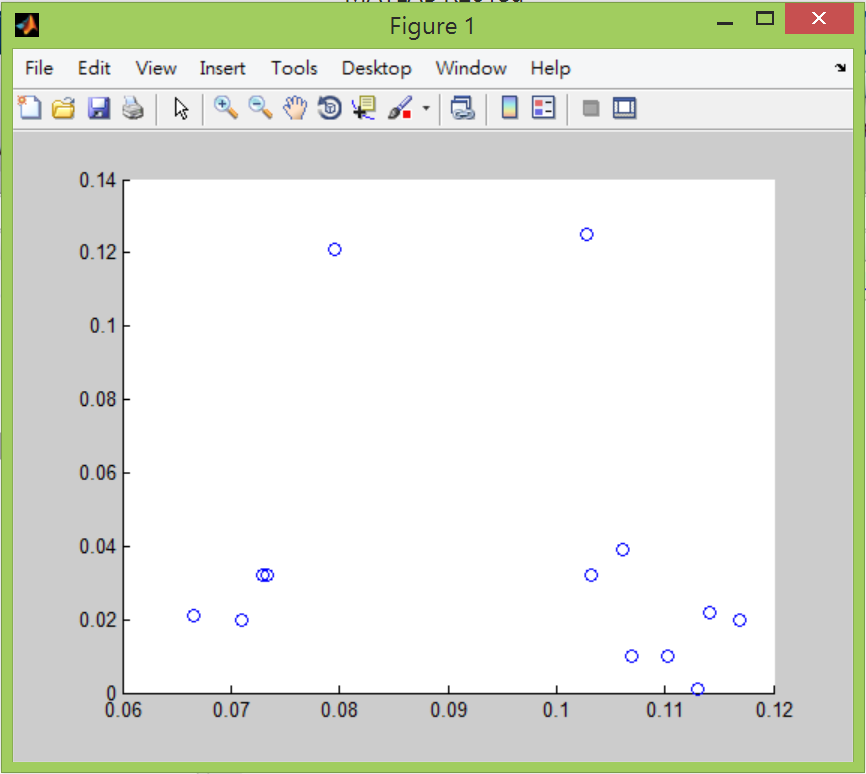
[x2^2-2x1-2 x1^2-2x2+3]

第四題是轉到

[2 2\*x1 2\*x2 sqrt(2) \*x1\*x2 x1^2 x2^2]

所以得到的分隔線不同

13.



橫軸為margin寬度，y軸為Eout。

理論上margin越大Eout 越小，但以這張圖來看，似乎沒有。實際上我跑很多次都差不多是這個結果，所以我想在這組data set裡面margin 越大，Eout越小的現象並不顯著。

14. #SV/N

|  |  |  |  |
| --- | --- | --- | --- |
|  | C=0.001 | C = 1 | C=1000 |
| Sigma=0.125 | 1 | 1 | 1 |
| Sigma= 0.5 | 0.84 | 1 | 1 |
| Sigma= 2 | 0.844 | 0.95 | 0.934 |

我發現隨著C越大或sigma越小support vectors就越多

Ein

|  |  |  |  |
| --- | --- | --- | --- |
|  | C=0.001 | C=1 | C=1000 |
| Sigma=0.125 | 0.416 | 0 | 0 |
| Sigma=0.5 | 0.416 | 0 | 0 |
| Sigma=2 | 0.416 | 0.126 | 0 |

Ecv

|  |  |  |  |
| --- | --- | --- | --- |
|  | C=0.001 | C=1 | C=1000 |
| Sigma=0.125 | 0.416 | 0.416 | 0.416 |
| Sigma=0.5 | 0.416 | 0.416 | 0.416 |
| Sigma=2 | 0.416 | 0.382 | 0.468 |

我發現隨著C上升可以把training set fit得很好。隨著sigma下降也是會把training set fit得很好。不過C過大或sigma過小容易overfitting(如Ein,Ecv表格的右上角就可以看出overfitting的趨勢。

15.

#SV/N

|  |  |  |  |
| --- | --- | --- | --- |
|  | C=0.001 | C=1 | C=1000 |
| Sigma=0.125 | 1 | 1 | 1 |
| Sigma=0.5 | 0.84 | 1 | 1 |
| Sigma=2 | 0.844 | 0.96 | 0.976 |

Ein

|  |  |  |  |
| --- | --- | --- | --- |
|  | C=0.001 | C=1 | C=1000 |
| Sigma=0.125 | 0.416 | 0 | 0 |
| Sigma=0.5 | 0.416 | 0 | 0 |
| Sigma=2 | 0.416 | 0.12 | 0 |

Ecv

|  |  |  |  |
| --- | --- | --- | --- |
|  | C=0.001 | C=1 | C=1000 |
| Sigma=0.125 | 0.416 | 0.416 | 0.416 |
| Sigma=0.5 | 0.416 | 0. 416 | 0. 416 |
| Sigma=2 | 0. 416 | 0.382 | 0.468 |

在這題我們可以看到Ein和support vectors的表格跟第14題非常像。表示說在training set裡面，SVM和SVR for classification表現相差不遠。

隨著sigma越小，C越大也是一樣越有overfitting的趨勢。

16.

#SV/N

|  |  |  |  |
| --- | --- | --- | --- |
|  | lambda=0.001 | lambda=1 | lambda=1000 |
| Sigma=0.125 | 1 | 1 | 1 |
| Sigma=0.5 | 1 | 1 | 1 |
| Sigma=2 | 1 | 1 | 1 |

不論在任何情形，LSSVR都會用到所有的點當support vectors。

Ein

|  |  |  |  |
| --- | --- | --- | --- |
|  | lambda=0.001 | lambda=1 | lambda=1000 |
| Sigma=0.125 | 0 | 0 | 0 |
| Sigma=0.5 | 0 | 0 | 0 |
| Sigma=2 | 0 | 0.048 | 0.246 |

Ecv

|  |  |  |  |
| --- | --- | --- | --- |
|  | lambda=0.001 | lambda=1 | lambda=1000 |
| Sigma=0.125 | 0.474 | 0.474 | 0.474 |
| Sigma=0.5 | 0.472 | 0.47 | 0.47 |
| Sigma=2 | 0.466 | 0.406 | 0.356 |

普遍而言，LSSVR在Ein表現很好(剩餘SVM和SVR)，不過這部分原因也是 因為他用了全部的support vectors去求一個相對比較複雜的模型。

在Ecv上就可以看到其實很多情形都是overfit。