3.

```
 Plabo@fuguanjundeMBP  ~/Dropbox/ML/ML_NTU/ML_NTU_HW/Hw5  master ●+  python 4.py
X.shape:(7, 2)
X:[[ 1  0]
 [ 0  1]
 [ 0 -1]
 [-1  0]
 [ 0  2]
 [ 0 -2]
 [-2  0]]

Y.shape:(7,)
Y:[-1 -1 -1  1  1  1  1]

K [[  4.   1.   1.   0.   1.   1.   1.]
 [  1.   4.   0.   1.   9.   1.   1.]
 [  1.   0.   4.   1.   1.   9.   1.]
 [  0.   1.   1.   4.   1.   1.   9.]
 [  1.   9.   1.   1.  25.   9.   1.]
 [  1.   1.   9.   1.   9.  25.   1.]
 [  1.   1.   1.   9.   1.   1.  25.]]
     pcost       dcost       gap    pres   dres
 0: -2.0644e+00 -4.5383e+00  2e+01  3e+00  2e+00
 1: -3.0994e+00 -4.4384e+00  5e+00  1e+00  6e-01
 2: -1.2189e+00 -2.0952e+00  9e-01  7e-16  2e-14
 3: -1.3986e+00 -1.4261e+00  3e-02  4e-16  9e-16
 4: -1.4073e+00 -1.4076e+00  3e-04  1e-16  7e-16
 5: -1.4074e+00 -1.4074e+00  3e-06  2e-16  5e-16
 6: -1.4074e+00 -1.4074e+00  3e-08  2e-16  9e-16
Optimal solution found.
```

svm_type c_svc
kernel_type polynomial
degree 2
gamma 1
coef0 1
nr_class 2
total_sv 5
rho 1.66653
label 1 -1
nr_sv 3 2

SV
0.8887164347987063 1:-1 => x4
0.1502852229699299 2:2 => x5
0.3681704323551298 2:-2 => x6
-0.4857048695492317 2:1 => x2
-0.9214672205745343 2:-1 => x3
對應到SV=x2,x3,x4,x5,x6
因此min (alpha i i從1到7) = alpha 7

4.
如講義第三講第4頁，
alpha_n*y_n*kernel(x_n,x)+b=1/9(8*x1^2–16*x1+6*x2^2–15)=0

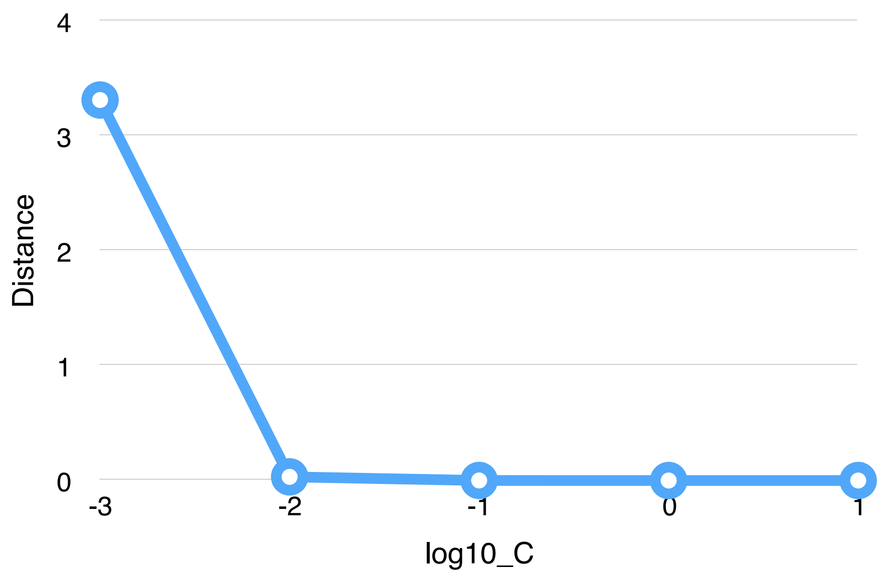
5.
The curves should be different in the x space, because they are learned with respect to different z spaces.
如第三講第7頁，藍色的transition跟綠色的transition一樣，雖然具有同樣的power但是有不同的inner product不同的geometry。
或8頁所說，change selecting K if and only if change of margin definition。

15.
用libsvm，參數為 -t 0 線性kernel，根據不同的C。
log10_C:-6,lWl:1.23092046817e-08
log10_C:-4,lWl:5.97920361981e-05
log10_C:-2,lWl:0.326392895857
log10_C:0,lWl:128.172269053
log10_C:2,lWl:171.301590146
因為margin = 2d（線到SV的距離）= 2/llwll
當C越大的時候越不能容忍錯誤，margin越小，w越大。

16.
log10_C:-6,Ein:0.0743382252092
log10_C:-4,Ein:0.0743382252092
log10_C:-2,Ein:0.0743382252092
log10_C:0,Ein:0.0743382252092
log10_C:2,Ein:0.0743382252092
因為屬於8這類的class樣本很少，非8的樣本很多，就算c跟w怎麼變都不會影響到Ein。

17.
log10_C:-6,alpha:0.001084
log10_C:-4,alpha:0.1084
log10_C:-2,alpha:10.84
log10_C:0,alpha:1084.0
log10_C:2,alpha:108400.0
因為C越大W越大Margin越小alpha越大。
應證第四講第8頁
No loss of optimality if solving with implicit constraint beta = C - alpha
and explicit constraint 0 <= alpha <= C:

18.
參考：http://my.oschina.net/u/1461744/blog/209104
先算出lWl，再運用svm_predict去算出free support vector的p_val。最後 p_val / lwl 得到距離
log10_C:-3, Distance:3.31437425017
log10_C:-2, Distance:0.0320639400527
log10_C:-1, Distance:0.000671724709803
log10_C:0, Distance:0.000257625011818
log10_C:1, Distance:0.000235309159599



19.
Eout:0.107125062282
Eout:0.0991529646238
Eout:0.105132037867
Eout:0.178873941206
Eout:0.178873941206
因為固定C所以在gamma增加時，Eout沒有顯著的改變微微上升。

20.
gamma:0 times:7
gamma:1 times:71
gamma:2 times:22