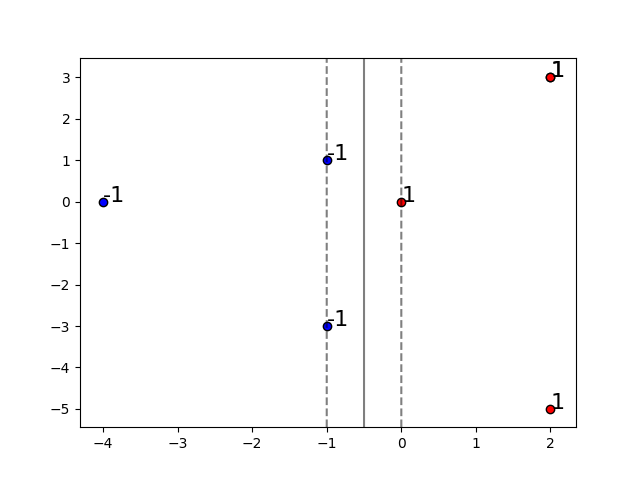
1.

x = [[1,0],[0,1],[0,-1],[-1,0],[0,2],[0,-2],[-2,0]]

y = [-1,-1,-1,1,1,1,1]

z = [[-4,0],[-1,-3],[-1,1],[0,0],[2,-5],[2,3],[2,3]]

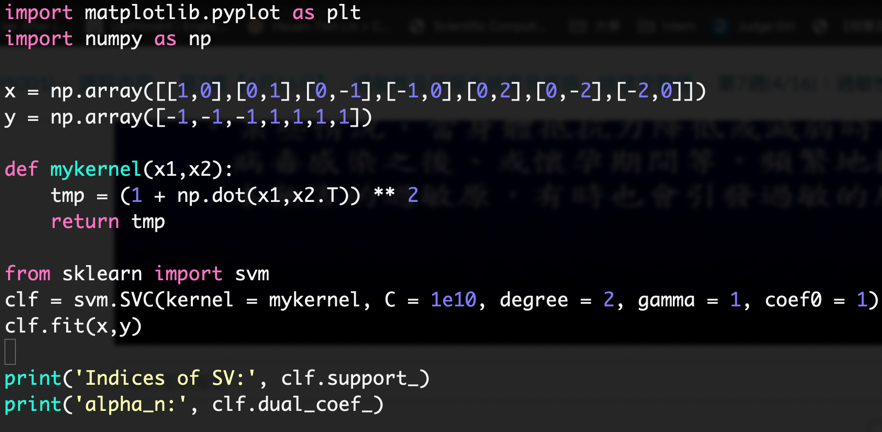
plot of z:



It’s easy to observe that the optimal separating “hyperplane” in Z space is

2.

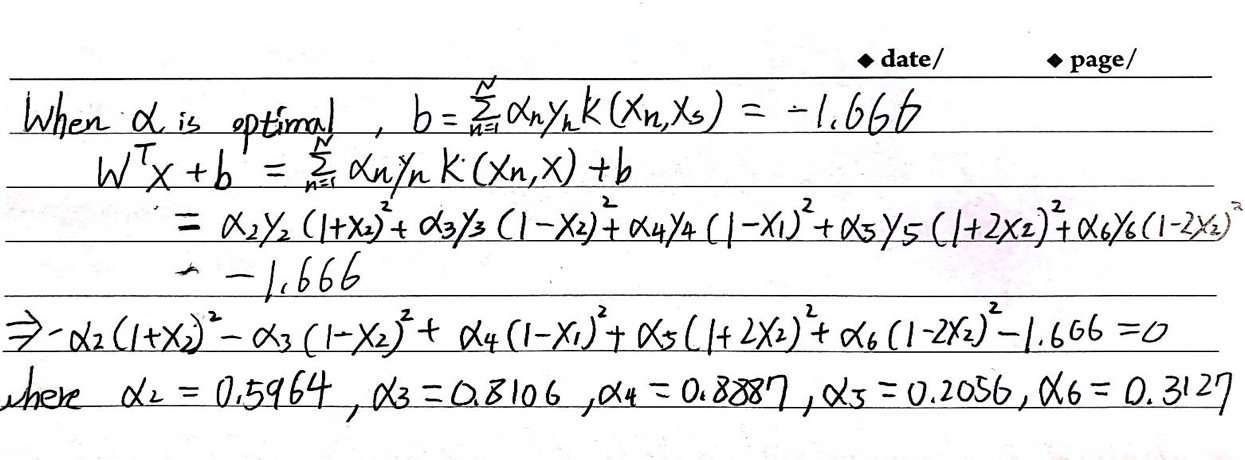
I used sklearn package. Below is my code:



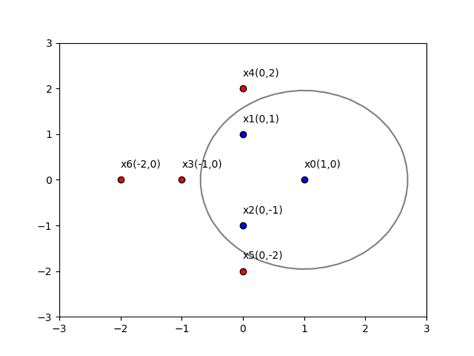
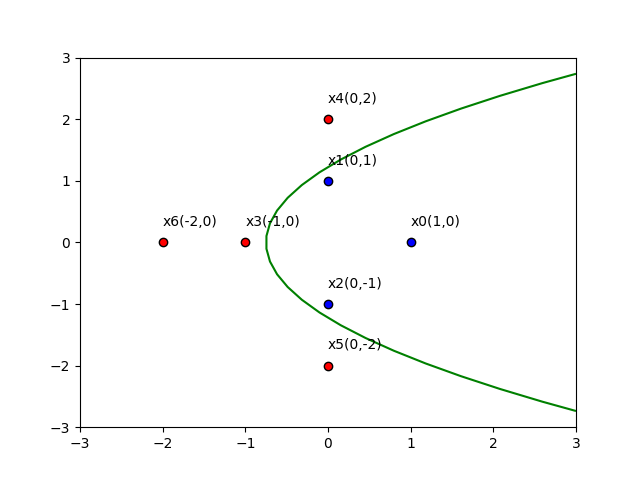
The optimal are

and the support vectors are (0,1), (0,-1), (-1,0), (0,2), (0,-2)

3.

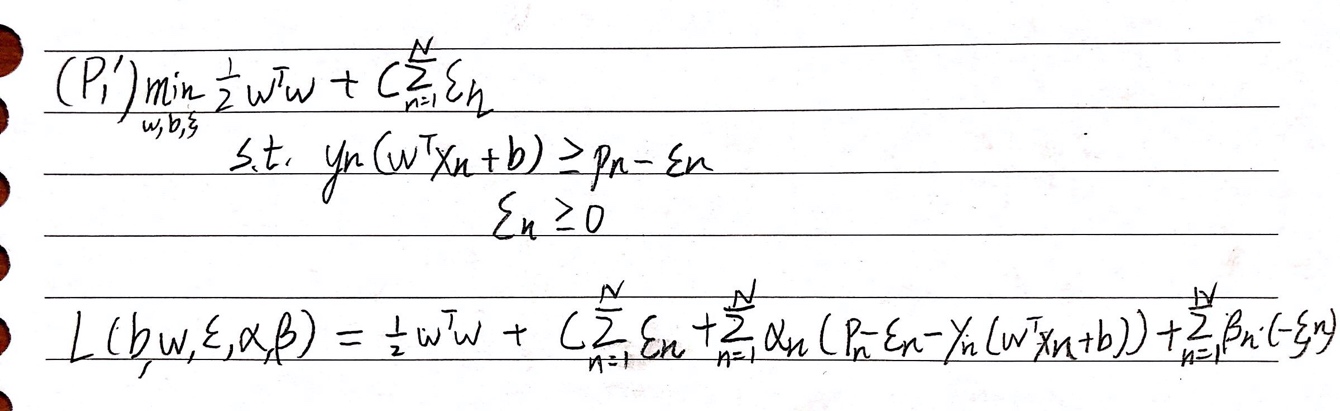


4.

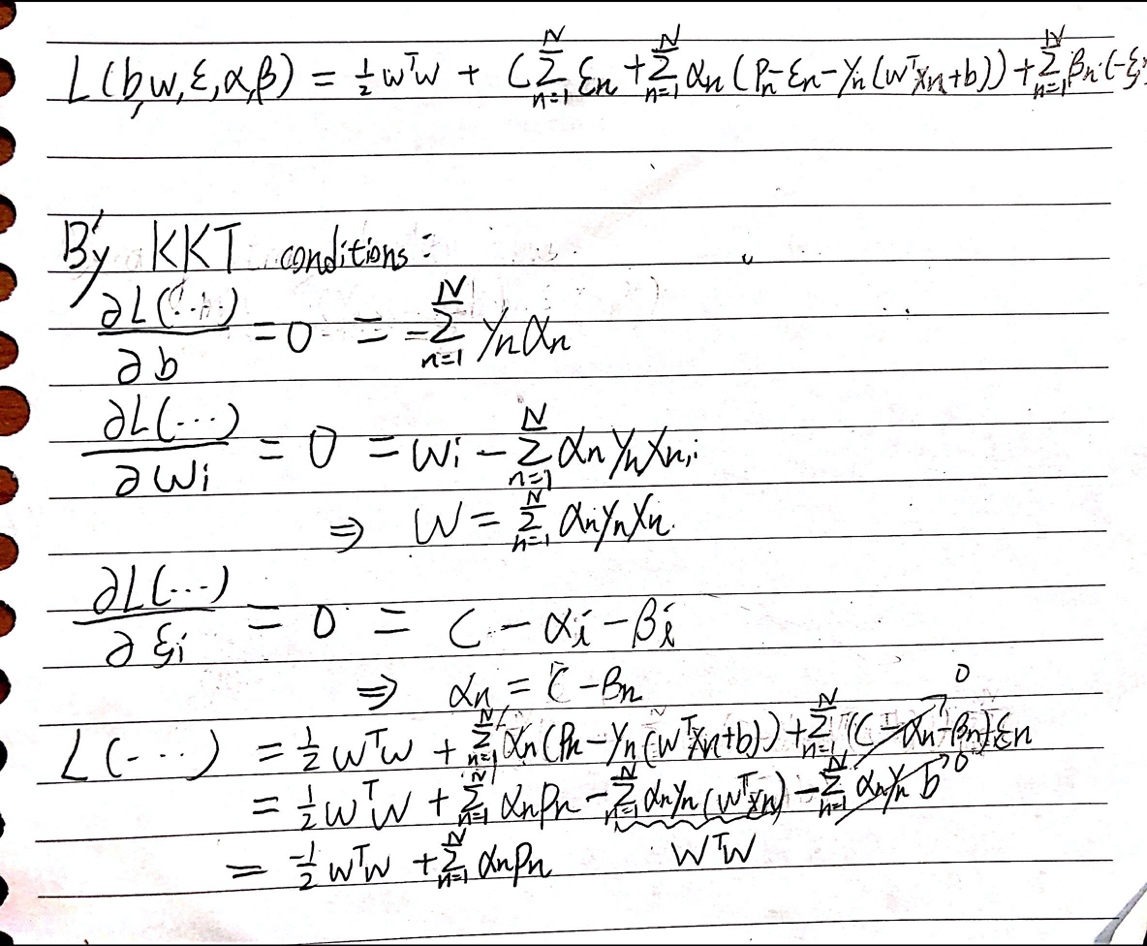


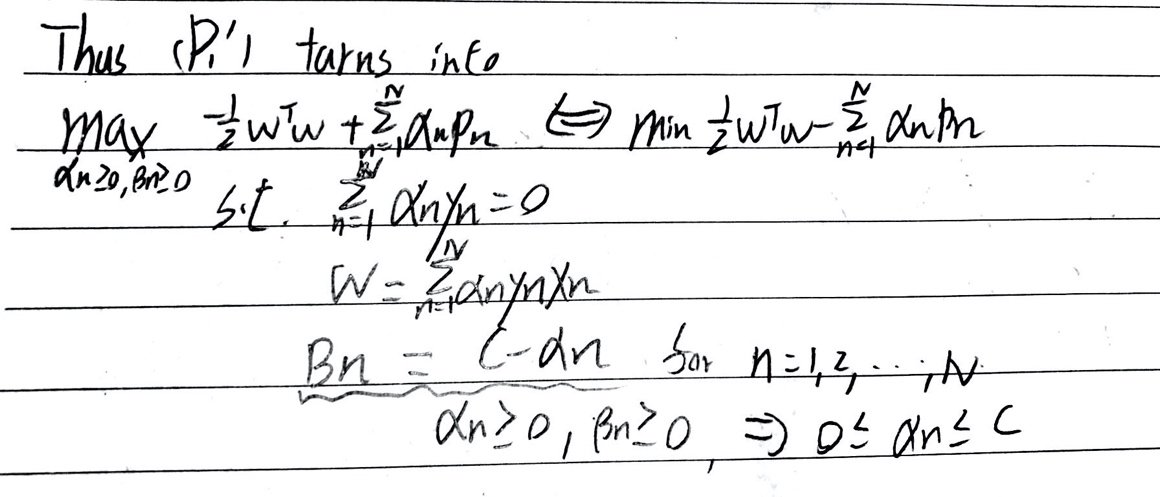
They dont have to be the same, because different kernel function (different nonlinear transformation) will lead to different optimal separating nonlinear curve

5.

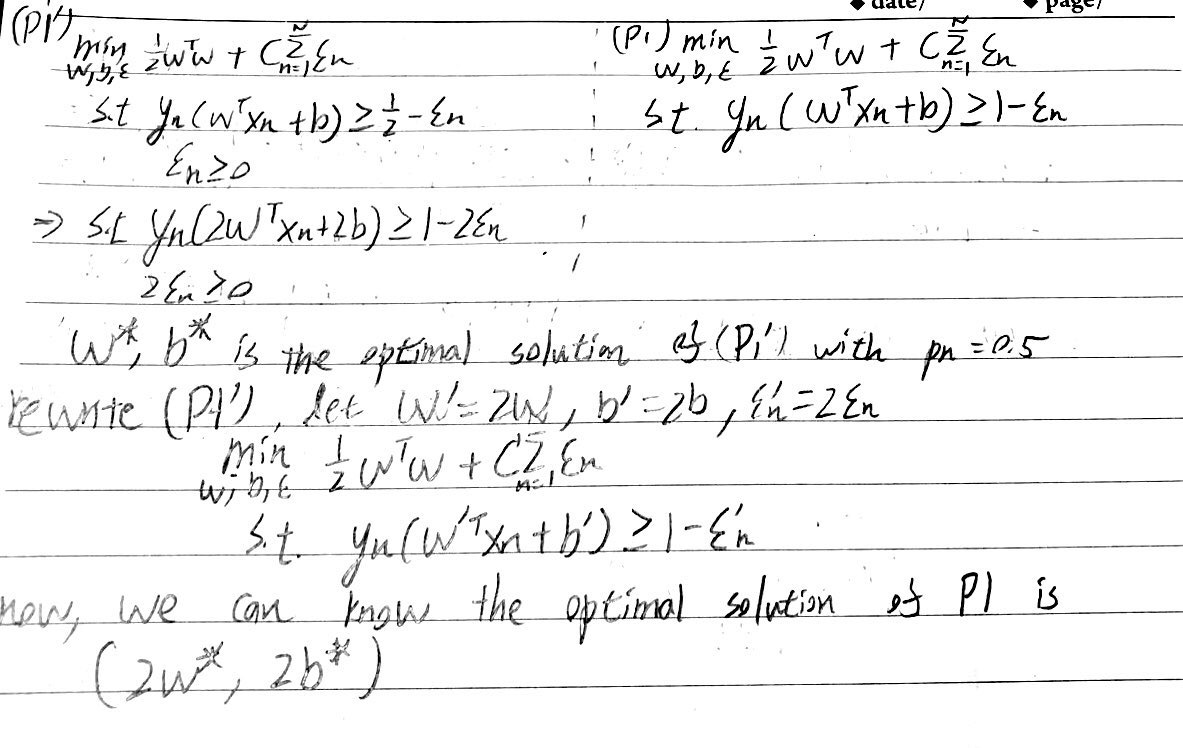


**6.**





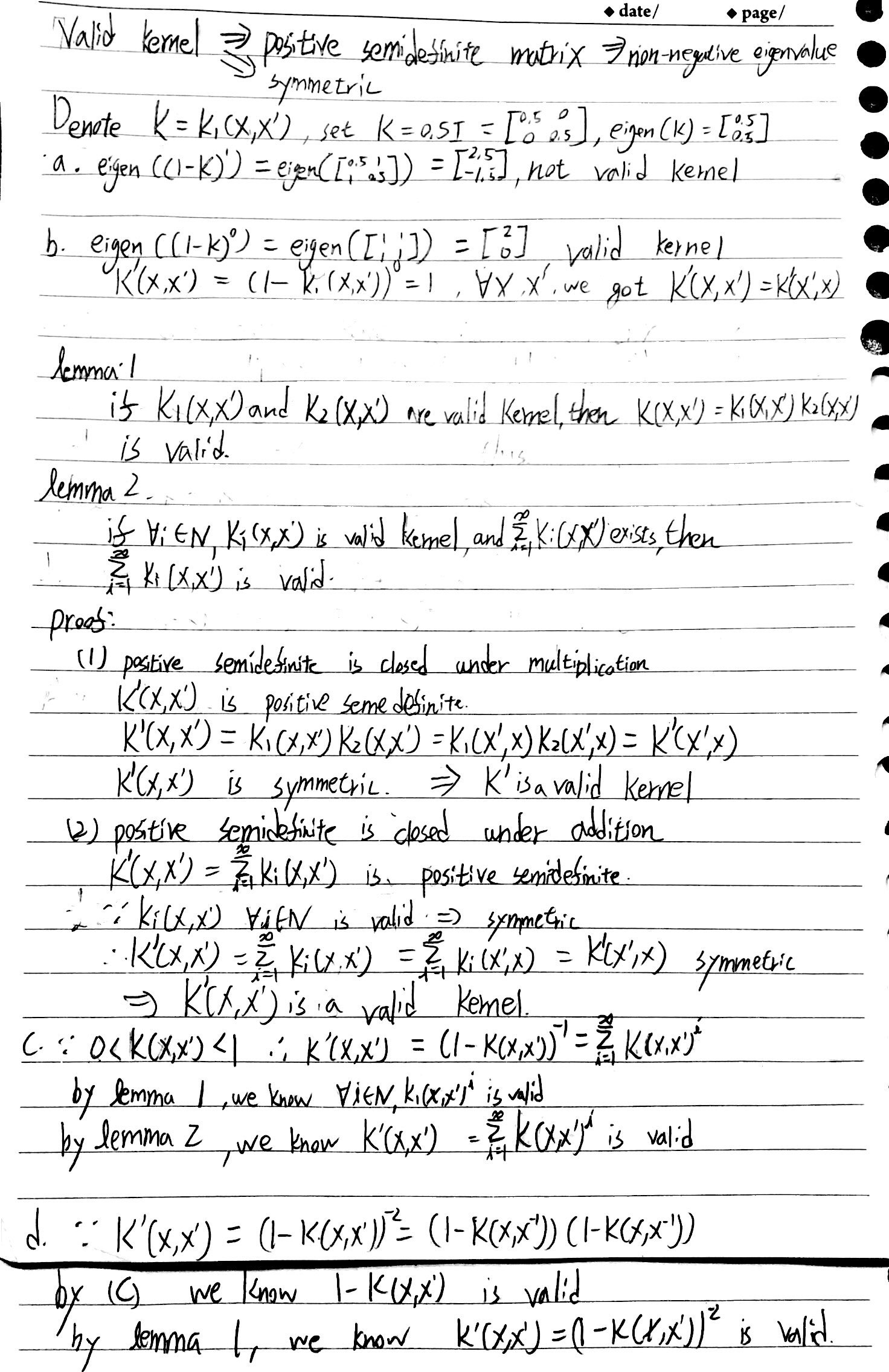
7.



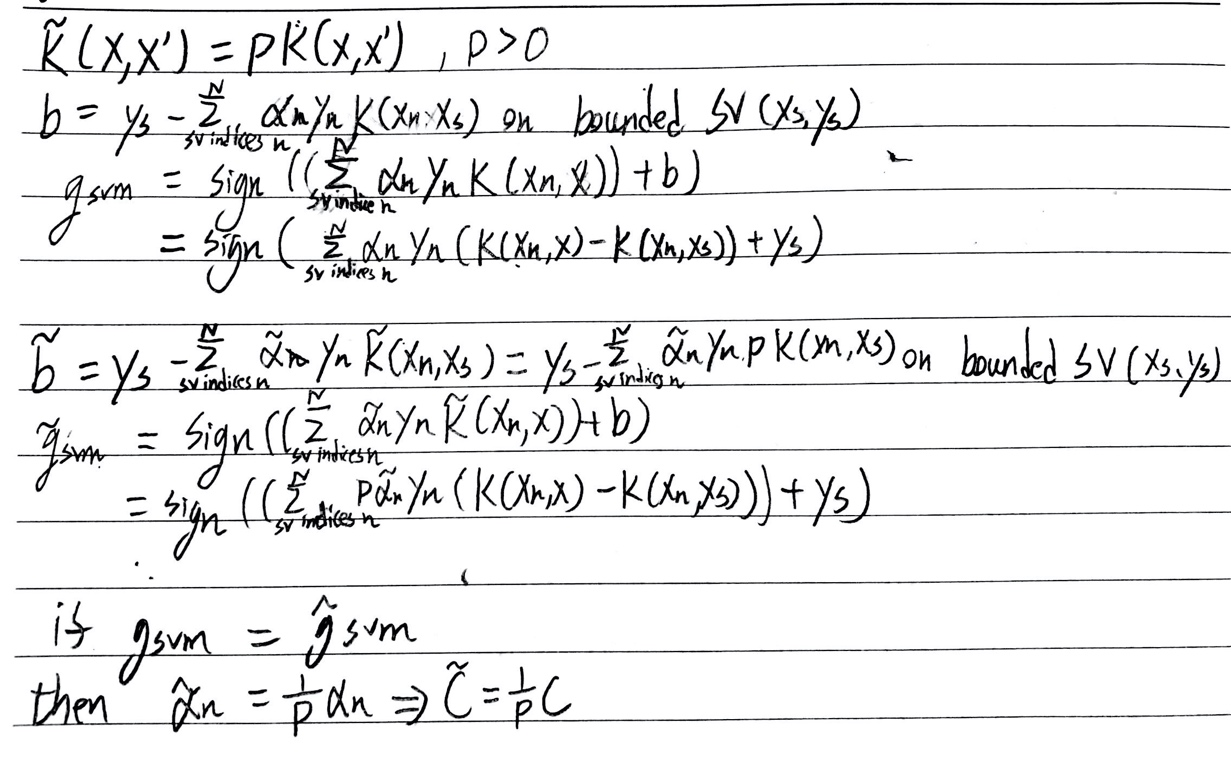
8.

The difference between hard-margin SVM and soft-margin SVM is that in soft-margin SVM there is an upper bound that constraints . So, if C is greater the optimal (which means that in both problems will stay in the range), i.e. , than the two problems have the same solution.

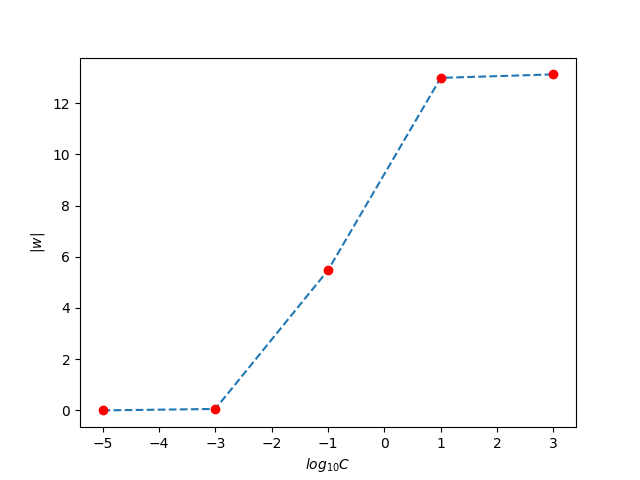
9.



10.

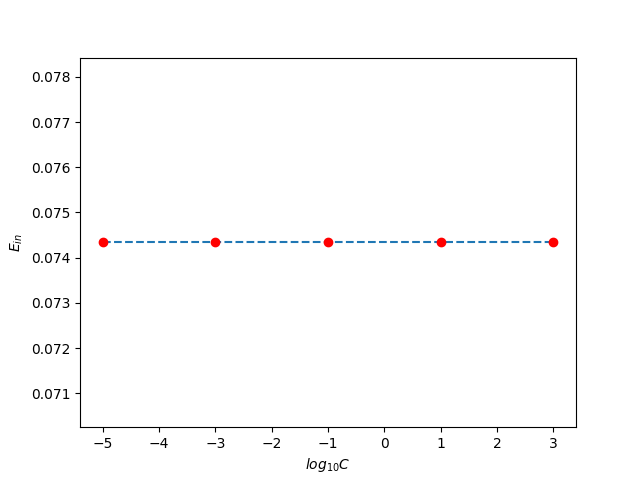


11.



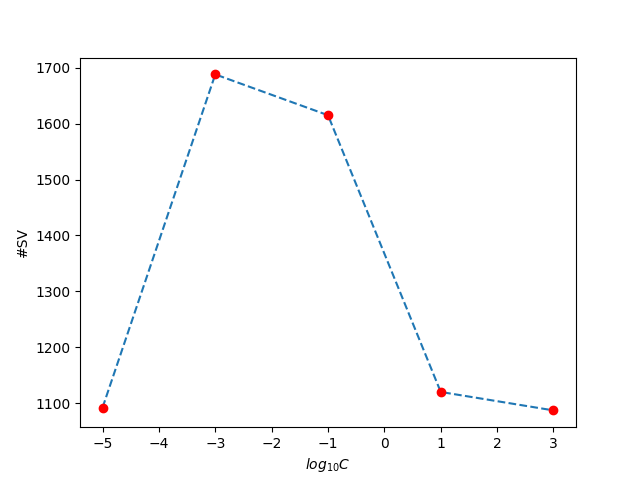
C越大表示:在邊界周圍能容忍的錯誤較少,同時圖形為了滿足這個條件,也會變複雜,複雜的曲線其|W|也會變大。

12.



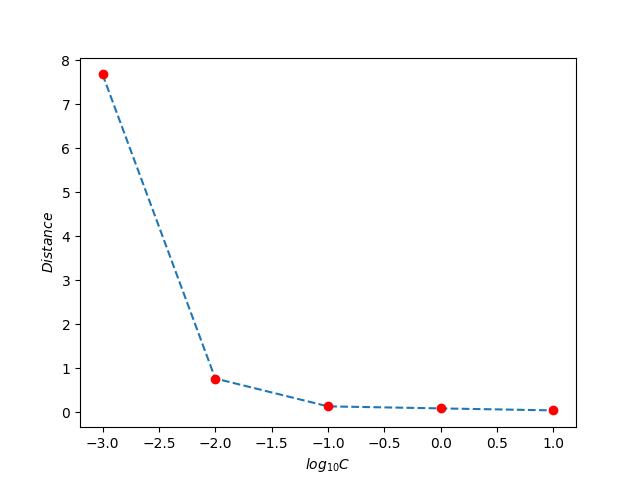
理論上,C越大表示:在邊界周圍,能容忍的錯誤越少。但這筆 Ein 沒什麼變,而且原本 8 所佔的比例剛好就是0.074,所以 C 的影響很小。

13.



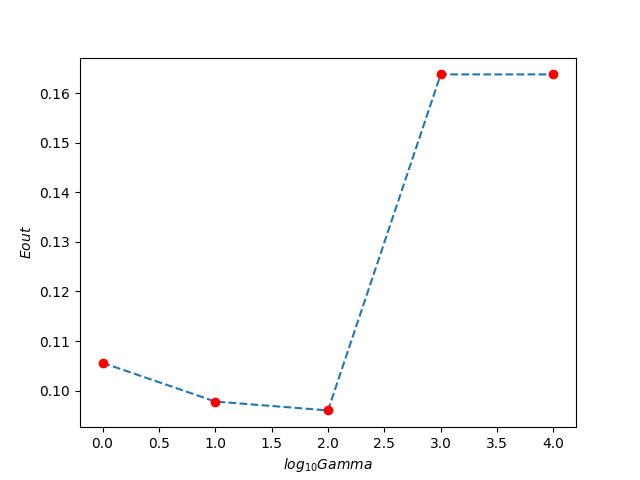
C較小的時候,因為圖形會偏簡單,所以Support Vector 會變少;而C太大的時候,他會盡量避免分類錯誤,讓unbounded Support Vector 不要太多,所以Support Vector 也會變少。

14.



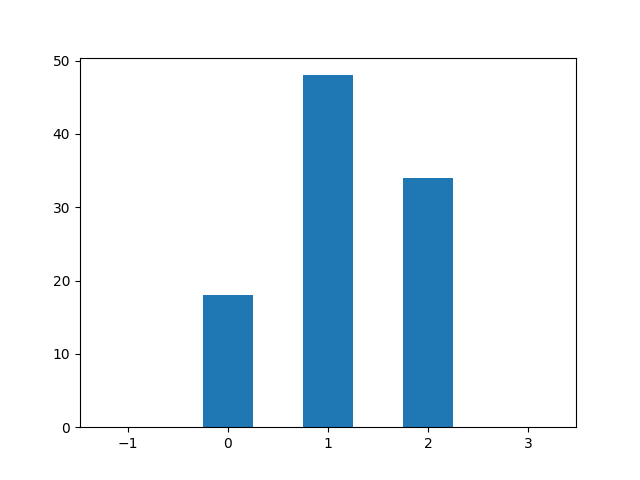
當 log10C 越大時,Z space中 free support vector 到 optimal separating hyperplane的距離越小。

15.



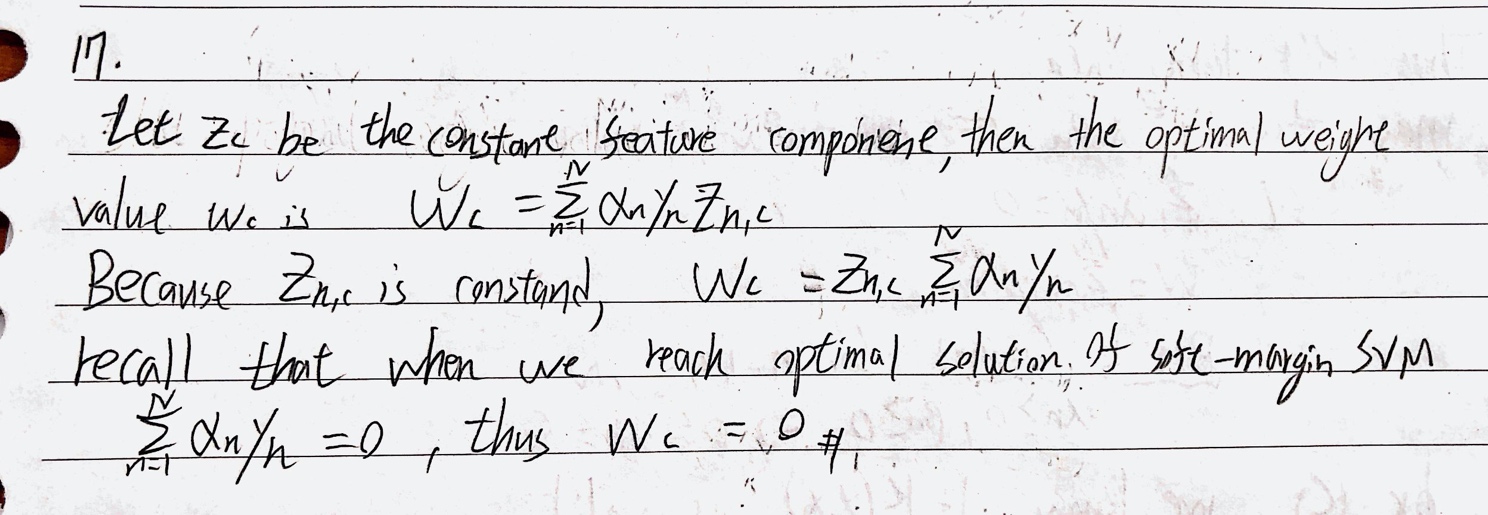
當 log10γ 從 0 增加到 1 時,Eout會下降到最小值,而當 log10γ 從 1 遞增到 4 時,Eout則會遞增。我想是因為γ太大會導致 overfitting,所以最後Eout會變大。

16.



當log10γ為1時,被選中的次數最多，而當log10γ為-1或3時，則皆沒被選中。

17.



18.