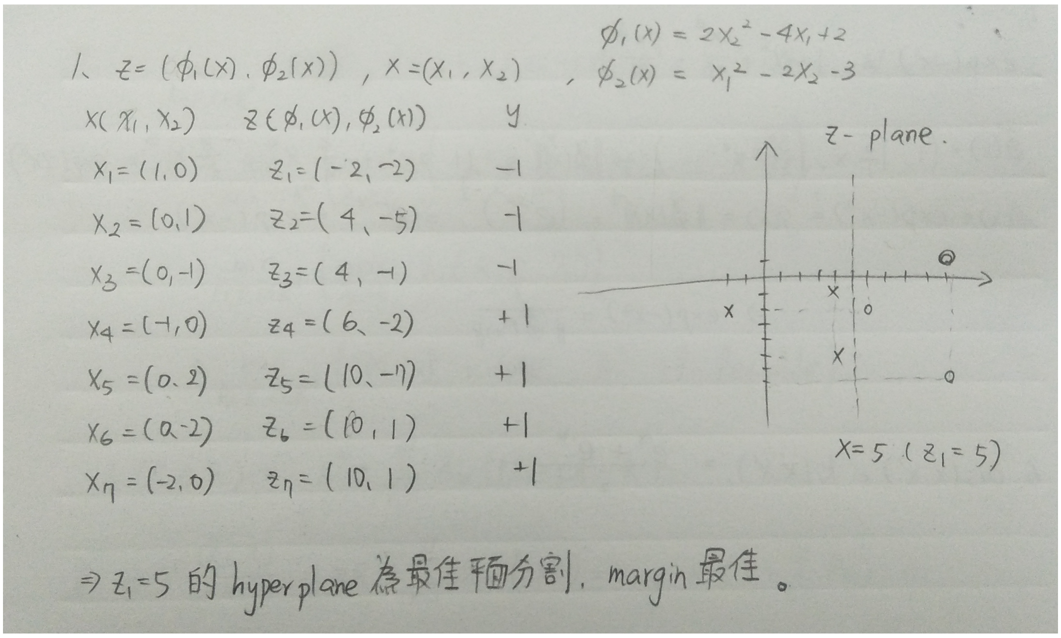
NTUEE, WAN - CYUAN FAN 范萬泉

"no collaborators, with reference"

Homework 1

Problem 1 :



1、 $z = (\phi_1(x), \phi_2(x))$, $x = (x_1, x_2)$, $\phi_1(x) = 2x_2^2 - 4x_1 + 2$, $\phi_2(x) = x_1^2 - 2x_2 - 3$

$X(x_1, x_2)$	$z(\phi_1(x), \phi_2(x))$	y
$x_1 = (1,0)$	$z_1 = (-2,-2)$	-1
$x_2 = (0,1)$	$z_2 = (4,-5)$	-1
$x_3 = (0,-1)$	$z_3 = (4,-1)$	-1
$x_4 = (-1,0)$	$z_4 = (6,-2)$	$+1$
$x_5 = (0,2)$	$z_5 = (10,-7)$	$+1$
$x_6 = (0,-2)$	$z_6 = (10,1)$	$+1$
$x_7 = (-2,0)$	$z_7 = (10,1)$	$+1$

$\Rightarrow z_1 = 5$ 的 hyperplane 為最佳平分割, margin 最佳。

Problem 2 :

這題general polynomial kernal(L3 p.9)上可以看出來， $\zeta = 1$ ，$\gamma = 1$ ，且為2次kernel，所以我們直接帶入SVM即可求出。

- the optimal α :(0 , 0.59647182 , 0.81065085 , 0.8887034 , 0.20566488 , 0.31275439 , 0)
- Support vectors : ((0,1) , (0,-1) , (-1,0) , (0,2) , (0,-2))

```python
import numpy as np
x = np.array([[1,0],[0,1],[0,-1],[-1,0],[0,2],[0,-2],[-2,0]])
y = np.array([-1,-1,-1,1,1,1,1])
from sklearn.svm import SVC
clf = SVC(kernel = 'poly' , degree = 2, coef0 = 1, gamma = 1)
clf.fit(x, y)

print(clf.support_)
print("support vectors : ",clf.support_vectors_)
print("Alpha * y: ",clf.dual_coef_)
```

Problem 3 :

由上提的資訊我們可以知道個別support vectors的 alpha值，接著利用(L3 p.5)的公式，我們可以找出b , w數值，其相加就是我們的軌跡方程式。

- $b \leftarrow \left(y_s - \sum_{SVindicesn} \alpha_n y_n K(x_n, x_s) \right)$ ，計算出 b
- $w \leftarrow \sum_{SVindicesn} \alpha_n y_n K(x_n, x)$，計算出w，再將兩者相加，等於零可以算出方程。

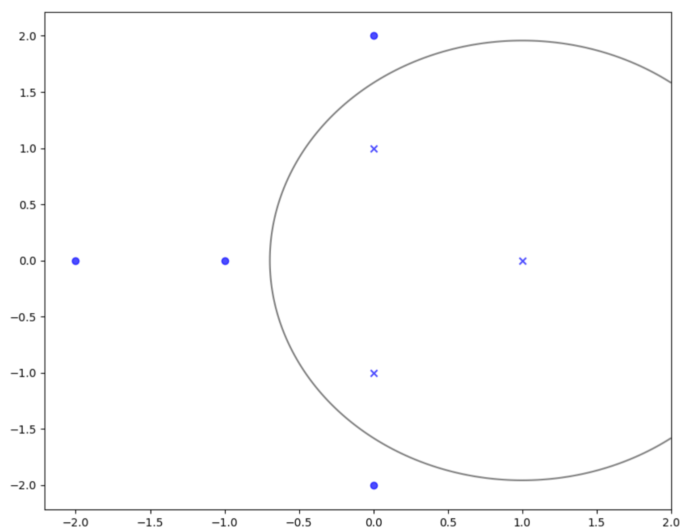
利用python 工具，我們找到方程式如下：

$0.8887x_1^2 + 0.6665544x_2^2 - 1.7774x_1 + 2.22044605 \times 10^{-16}x_2 - 1.66655442 = 0$

w 部分：我們用np.array() 表示 $x_1^2, x_2^2, x_1, x_2, 1$ 等不同級數。

```python
#find b
b = []
for i in clf.support_:
    tmp = 0.
    for j , k in zip(clf.support_,range(0,5)):
        tmp += clf.dual_coef_[0][k] * (1+np.inner(x[j],x[i]))**2.
    b.append(y[i] - tmp)
b = np.array(b)
print("b : ",b)

#find w
w = []
for j , k in zip(clf.support_,range(0,5)):
    ay = clf.dual_coef_[0][k]
    tmp = [ay*x[j][0]**2.,ay*x[j][1]**2.,ay*2*x[j][0],ay*2*x[j][1],ay*1]
    tmp = np.array(tmp)
    w.append(tmp)
w = np.array(w)
w = np.sum(w,axis = 0)
print("w : ",w)
```



Problem 4 :

不會相同，因為在problem 01的轉換 $z_1 = \phi_1(x) = 2x_2^2 - 4x_1 + 2$或是 z_2 與我們在problem 03得到的轉換函式

$0.8887x_1^2 + 0.6665544x_2^2 - 1.7774x_1 + 2.22044605 \times 10^{-16}x_2 - 1.66655442$

是不同的空間轉換，z_1呈現拋物線，而下面方程式橢圓，所以分割圖形不會相等。

我們可以將problem 01中，$z_1 = 5$ 方程式轉回x space，$z_1 = \phi_1(x) = 2x_2^2 - 4x_1 + 2 = 5$ 為一條拋物線。

Problem 5 :

5.

$$\exp(-x^2) = 1 - x^2 + \frac{x^4}{2!} - \frac{x^6}{3!} + \cdots$$

$$\tilde{\phi}(x) = \left(1, \sqrt{\tfrac{2}{1}}\,x, \sqrt{\tfrac{2^2}{2!}}\,x^2 \ldots\right) \rightarrow \|\tilde{\phi}(x)\|^2 = 1 + 2x^2 + \frac{4}{2!}x^4 + \frac{8}{3!}x^6 \ldots = \exp(2x^2)$$

$$f(x) = \exp(-x^2) \overset{?}{=} g(x) = \|\tilde{\phi}(x)\|^{-1} = \left(e^{2x^2}\right)^{-\frac{1}{2}} = e^{-x^2} = \exp(-x^2)$$

$$\Rightarrow \exp(-x^2) = \frac{1}{\|\tilde{\phi}(x)\|}$$

Problem 6 :

6. $\cos(x, x') = k(x, x')$，令 $k(x_i, x_j) = \cos(x_i, x_j) = k_{ij}$

根據 Mercer's condition，若 $\cos(x, x')$為合法 kernel.

滿足 ① symmetric

則 $\cos(x_1, x_2) = \cos(x_2, x_1)$ 成立.

② positive semi-definite.

則 matrix K 滿足

$c^T \cdot K \cdot c \geq 0$ for all $c \in R^n \backslash 0$ (c is all non-zero in R^n)

∵ cos 值 $\in [-1, 1]$

$$[c_1\ c_2 \ldots c_N] \begin{bmatrix} \cos(x_1, x_1) & \cos(x_1, x_2) & \cdots \\ \cos(x_2, x_1) & \vdots & \\ \vdots & \vdots & \\ \cos(x_N, x_1) & & \cos(x_N, x_N) \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_N \end{bmatrix}$$

(反例)

$$[1\ 2\ 3] \begin{bmatrix} 1 & -0.8 & -0.9 \\ -0.8 & 1 & -0.7 \\ -0.9 & -0.7 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} = [-3.3\ \ -0.9\ \ 0.7] \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$$

$$= -3.3 - 1.8 + 2.1 = -3 \leq 0 \text{ 不合}$$

∴ $\cos(x, x')$ 不為一個 valid kernel

Problem 7 :

7. $\min\limits_{R\in\mathbb{R}, c\in\mathbb{R}^d} R^2$, s.t. $\|z_n - c\|^2 \le R^2$ $n = 1, 2, 3 \cdots, N$.

$$\text{SVM} \equiv \min_{R\in\mathbb{R}, c\in\mathbb{R}^d} \left(\max_{\lambda_n \geq 0} \left(R^2 + \sum_{n=1}^{N} \lambda_n \left(\|z_n - c\|^2 - R^2 \right) \right) \right)$$

$$\equiv \min_{R\in\mathbb{R}, c\in\mathbb{R}^d} \left(\max_{\lambda_n \geq 0} \mathcal{L}(R, c, \lambda) \right)$$

$$\equiv \min_{R\in\mathbb{R}, c\in\mathbb{R}^d} \left(\infty \text{ if violate}, R^2 \text{ if feasible} \right)$$

$$\mathcal{L}(R, c, \lambda) = R^2 + \sum_{n=1}^{N} \lambda_n \left(\|z_n - c\|^2 - R^2 \right)$$

Problem 8 :

8. ① $\dfrac{\partial \mathcal{L}(R,c,\lambda)}{\partial R} = \cancel{2}R + \sum_{n=1}^{N} \lambda_n \cdot (-\cancel{2}R) = 0$, $R\left(1 - \sum_{n=1}^{N} \lambda_n\right) = 0$

② $\dfrac{\partial \mathcal{L}(R,c,\lambda)}{\partial c_i} = \sum_{n=1}^{N} \lambda_n \left(-2\|z_{n,i} - c_i\|\right) = 0$, $\sum_{n=1}^{N} \lambda_n \|z_{n,i} - c_i\| = 0$

* KKT condition : (P) (D)

① primal feasible : $\|z_n - c\|^2 \le R^2$

② dual feasible : $\lambda_n \geq 0$

③ dual - inner optimal : $\sum_{n=1}^{N} \lambda_n = 1$ (if $R \neq 0$), $\sum_{n=1}^{N} \lambda_n \|z_{n,i} - c_i\| = 0$

④ primal - inner optimal : $\lambda_n \left(\|z_n - c\|^2 - R^2\right) = 0$

$\left(c = \dfrac{\sum_{n=1}^{N} \lambda_n \cdot z_n}{\sum_{n=1}^{N} \lambda_n} \right)$

<proof>

$$\left\| \sum_{n=1}^{N} \lambda_n \cdot z_{n,i} - \sum_{n=1}^{N} \lambda_n c_i \right\| = 0 \quad , \quad c_i = \frac{\sum_{n=1}^{N} \lambda_n \cdot z_{n,i}}{\sum_{n=1}^{N} \lambda_n}$$

$$c = \frac{\sum_{n=1}^{N} \lambda_n \cdot z_n}{\sum_{n=1}^{N} \lambda_n}$$

Problem 9 :

9. $R(1-\sum_{n=1}^{N}\lambda_n)=0$, $\sum_{n=1}^{N}\lambda_n\|z_n-c\|=0$ $\quad R^2 = \sum_{n=1}^{N}\lambda_n R^2$

$$L(R,c,\lambda) = R^2+\sum_{n=1}^{N}\lambda_n(\|z_n-c\|^2-R^2)$$

$$= \cancel{R^2}+\sum_{n=1}^{N}\lambda_n\|z_n-c\|^2-\cancel{\sum_{n=1}^{N}\lambda_n R^2} = \sum_{n=1}^{N}\lambda_n\|z_n-c\|^2$$

$$\max_{\text{all } \lambda_n\geq 0,\ \sum_{n=1}^{N}\lambda_n=1,\ c=\frac{\sum_{n=1}^{N}\lambda_n\cdot z_n}{\sum_{n=1}^{N}\lambda_n}} \sum_{n=1}^{N}\lambda_n\|z_n-\underbrace{\frac{\sum_{i=1}^{N}\lambda_i\cdot z_i}{\sum_{i=1}^{N}\lambda_i}}_{c}\|^2$$

$$\Rightarrow \sum_{n=1}^{N}\lambda_n\left[\|z_n\|^2+\|c\|^2+2\|z_n\cdot c\|\right] \qquad \left(c\sum_{n=1}^{N}\lambda_n=\sum_{n=1}^{N}\lambda_n\cdot z_n\right)$$

$$= \sum_{n=1}^{N}\lambda_n\|z_n\|^2+\sum_{n=1}^{N}\lambda_n\|c\|^2+2\sum_{n=1}^{N}\lambda_n\|z_n\cdot c\| \qquad \hookrightarrow \sum_{n=1}^{N}\lambda_n\|c\|^2=\sum_{n=1}^{N}\lambda_n\|z_n c\|$$

$$= \sum_{n=1}^{N}\lambda_n\|z_n\|^2+3\sum_{n=1}^{N}\lambda_n\|z_n\cdot c\|$$

$$= \sum_{n=1}^{N}\lambda_n\|z_n\|^2+3\sum_{n=1}^{N}\lambda_n\frac{\sum_{i=1}^{N}\lambda_i z_n^T\cdot z_i}{\sum_{i=1}^{N}\lambda_i} = \sum_{n=1}^{N}\lambda_n\|z_n\|^2+\frac{3}{\underbrace{\boxed{\sum_{i=1}^{N}\lambda_i}}_{1}}\sum_{n=1}^{N}\sum_{m=1}^{N}\lambda_n\lambda_m z_n^T z_m$$

$$\Rightarrow \max_{\lambda_n\geq 0}\ 3\sum_{n=1}^{N}\sum_{m=1}^{N}\lambda_n\lambda_m z_n^T z_m+\sum_{n=1}^{N}\lambda_n\|z_n\|^2,\ \text{s.t.}\ \sum_{n=1}^{N}\lambda_n=1$$

Problem 10 :

10.

$\min$ 加負号

$$\max_{\lambda_n \ge 0} \; 3\sum_{n=1}^{N}\sum_{m=1}^{N} \lambda_n \lambda_m z_n^T z_m + \sum_{n=1}^{N} \lambda_n z_n^T z_n \quad \text{s.t. } \sum_{n=1}^{N} \lambda_n = 1$$

原 QP Solver

using QP(Q, p, A, C)

$$\min_{\alpha} \; \frac{1}{2}\alpha^T Q \alpha + p^T \alpha$$

s.t. $a_i^T \alpha \ge c_i$, $i = 1, 2 \ldots$

- $q_{n,m} = -6 z_n^T z_m$
- $p_n = -z_n^T z_n$
- $a_i = 1$
- $c_{\ge} = 1$, $c_{\le} = 1$, $c_n = 1$

① kernel: $K(x, x') = z^T z'$

$q_{n,m} = -6K(x_n, x_m)$

$p_n = -K(x_n, x_n)$

② $\lambda \leftarrow QP(Q, P, A, C)$

③ $c \leftarrow \sum_{n=1}^{N} \lambda_n z_n$

④ $\because \|z_n - c\|^2 \le R^2$, $\min R^2$

求出 $\|z_n - c\|^2$ 後, R^2 即最小等於 $\|z_n - c\|^2$

Problem 11 :

11. Hard-margin: $\min_{\alpha} \frac{1}{2}\sum_{n=1}^{N}\sum_{m=1}^{N} \alpha_n \alpha_m y_n y_m z_n^T z_m - \sum_{n=1}^{N} \alpha_n$, s.t. $\begin{cases} \sum_{n=1}^{N} y_n \alpha_n = 0 \\ \alpha_n \ge 0, \; n = 1, 2 \ldots N \end{cases}$

$\Rightarrow$ optimal sol. $\Rightarrow \alpha = \alpha^*$

Soft-margin: $\min_{\alpha} \frac{1}{2}\sum_{n=1}^{N}\sum_{m=1}^{N} \alpha_n \alpha_m y_n y_m z_n^T z_m - \sum_{n=1}^{N} \alpha_n$, s.t. $\begin{cases} \sum_{n=1}^{N} y_n \alpha_n = 0 \\ \alpha_n \ge 0, \; n = 1, 2 \ldots N \\ \alpha_n \le C, \; n = 1, 2 \ldots N \end{cases}$

在Hard-margin SVM 和 Soft-margin SVM 中,

差別只在 Soft-margin SVM 多了 $\alpha_n \le C$ 的 upper bound.

因此,若加上條件 $C \ge \max_{1 \le n \le N} \alpha_n^*$,則在Hard-margin SVM 求得的 optimal sol. α^*

必滿足 α^* 其中任何一項 $\le \max_{1 \le n \le N} \alpha_n^* \le C$

使 Soft-margin SVM 亦成立,故 α^* 亦為 Soft-margin SVM 的 optimal sol.

Problem 12 :

12、 $\tilde{K}(x,x') = k(x,x') \times p \quad , p > 0$

$$\begin{cases} g_{SVM}(x) = \text{sign}\left(\sum_{SV\ indices\ n} \alpha_n y_n K(x_n, x) + b\right) \\ b = y_s - \sum_{SV\ in\ n} \alpha_n y_n K(x_n, x_s) \end{cases}$$

$$g'_{SVM}(x) = \text{sign}\left(\sum_{SV\ in\ n} \alpha_n' \cdot y_n \cdot \tilde{K}(x_n, x) + b'\right)$$

$$= \text{sign}\left(\sum_{SV\ in\ n} \alpha_n' y_n p \cdot k(x_n, x) + y_s - \sum_{SV\ in\ n} \alpha_n' y_n p \cdot k(x_n, x_s)\right)$$

$$= \text{sign}\left(\sum_{SV\ in\ n} \alpha_n' y_n p \cdot \left(k(x_n, x) - k(x_n, x_s)\right) + y_s\right) \text{ —— ①}$$

$$g_{SVM}(x) = \text{sign}\left(\sum_{SV\ in\ n} \alpha_n y_n k(x_n, x) + y_s - \sum_{SV\ in\ n} \alpha_n y_n k(x_n, x_s)\right)$$

$$= \text{sign}\left(\sum_{SV\ in\ n} \alpha_n y_n \left(k(x_n, x) - k(x_n, x_s)\right) + y_s\right) \text{ —— ②}$$

在①中，若 $\tilde{C} = \frac{C}{P}$ ，則 $C \geq \alpha_n \geq 0 \rightarrow \tilde{C} \geq \frac{\alpha_n}{P} \geq 0$

$\frac{\alpha_n}{P} = \alpha_n'$

則①

$$g'_{SVM}(x) = \text{sign}\left(\sum_{SV\ in\ n} \alpha_n y_n \left(k(x_n, x) - k(x_n, x_s)\right) + y_s\right) = g_{SVM}(x)$$

*檢查 QP solver

(原本)

$$\min_{\alpha} \ \frac{1}{2} \alpha^T Q_D \alpha - 1^T \alpha$$

$$\text{s.t. } y^T \alpha = 0$$

$$\alpha_n \geq 0 \text{ for } 1, 2, \dots N$$

$\left(\alpha_n' = \frac{\alpha_n}{P} \ , \ \tilde{K}(x,x') = k(x,x') \times P\right)$

$$\min_{\alpha} \ \frac{1}{2} \frac{\alpha^T}{P} (P Q_D) \frac{\alpha}{P} - 1^T \frac{\alpha}{P}$$

$$\text{s.t. } y^T \frac{\alpha}{P} = 0$$

$$\frac{\alpha_n}{P} \geq 0 \text{ for } 1, 2, \dots N$$

$$\Rightarrow \min_{\alpha} \ \frac{1}{P}\left(\frac{1}{2} \alpha^T Q_D \alpha - 1^T \alpha\right)$$

$$\text{s.t. } y^T \alpha = 0$$

$$\alpha_n \geq 0 \text{ for } 1, 2, \dots N$$

(α_n 相同解)正確

Problem 13 :

我使用sklearn module裡面的SVC，其數學套用如下：

- primal problem:

$$\min_{w,b,\zeta} \frac{1}{2} w^T w + C \sum_{i=1}^{n} \zeta_i$$

$$\text{subject to } y_i(w^T \phi(x_i) + b) \geq 1 - \zeta_i,$$
$$\zeta_i \geq 0, i = 1,\ldots,n$$

- dual problem:

$$\min_{\alpha} \frac{1}{2} \alpha^T Q \alpha - e^T \alpha$$

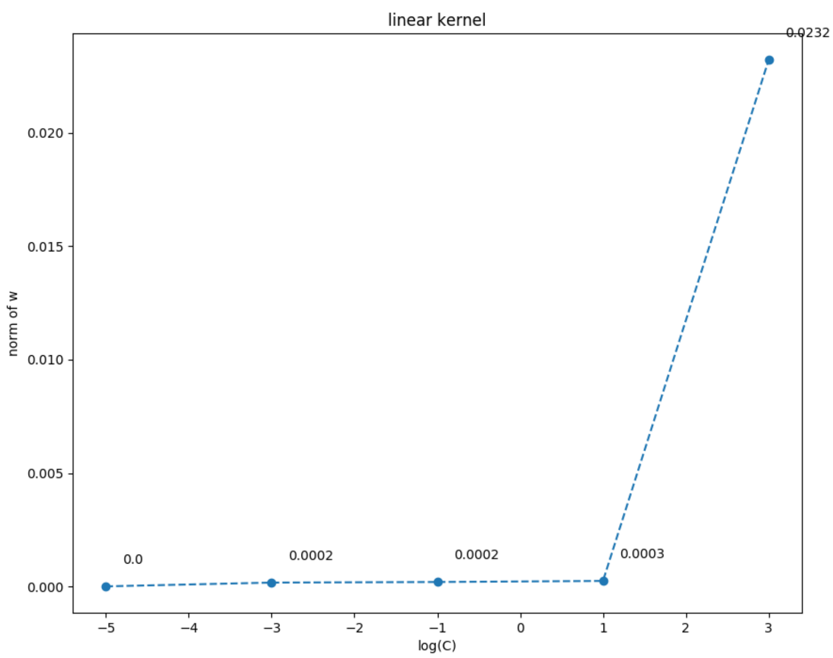
$$\text{subject to } y^T \alpha = 0$$
$$0 \leq \alpha_i \leq C, i = 1,\ldots,n$$

Where e is the vector of all one. C > 0 is the upper bound. Q is an n by n positive semidefinite matrix. $Q_{ij} \equiv y_i y_j K(x_i, x_j)$, and $K(x_i, x_j) = \phi(x_i)^T \phi(x_j)$ is the kernel.

Here training vectors are implicitly mapped into a higher (maybe infinite) dimensional space by the function ϕ.

The decision function is: $\operatorname{sgn}(\sum_{i=1}^{n} y_i \alpha_i K(x_i, x) + \rho)$

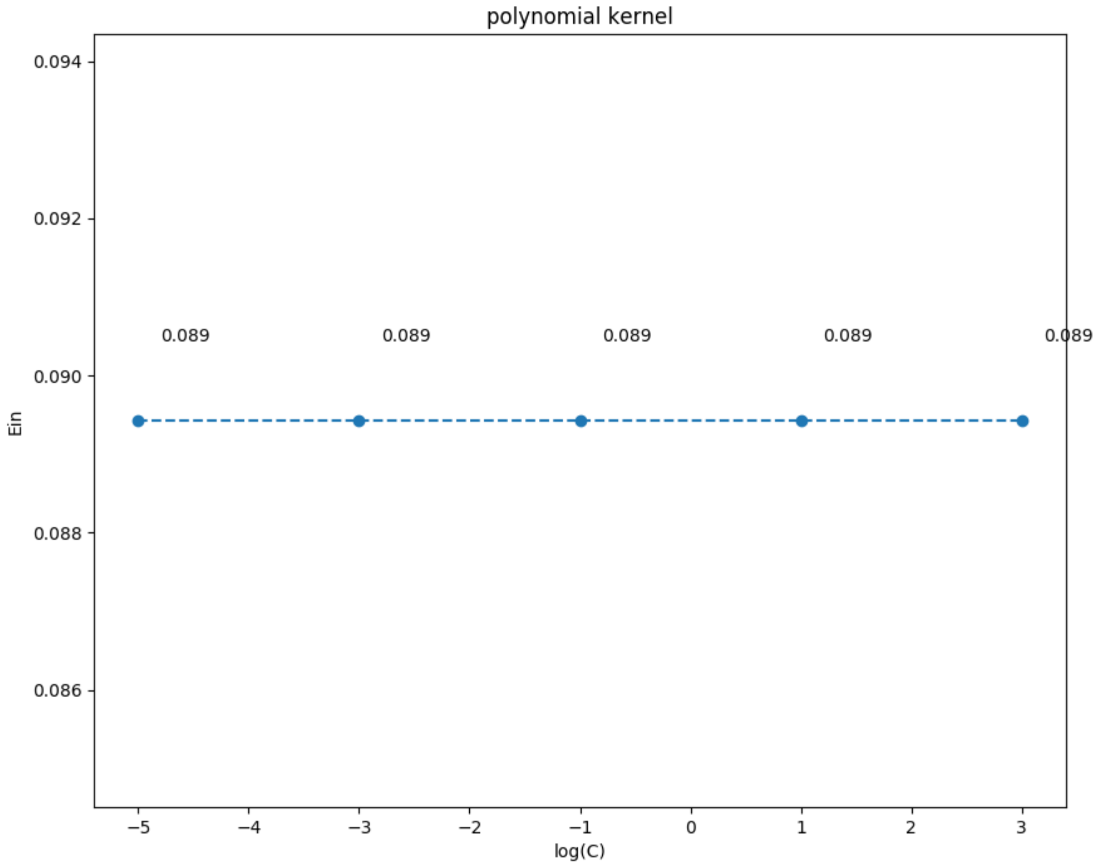
因為數學架構與講義上的一樣，所以我們直接帶入數值做計算，計算到小數點後第四位為止，結果畫成圖片如下：



我們發現在提升c值之後，因為對於錯誤的容忍度下降，也就是說對錯誤點的處罰權重上升，使得w的絕對值也跟著上升。

Problem 14 :

使用與problem 13相同的sklearn module ，只是此時我們kernel使用polynomial kernel 。而Ein計算方式如下： $E_{in} = 1 - clf.score\left(train_{data}[: ,1 :], train_{label}\right)$
再將圖形化出如下：



這部分我們發現，$\log c$ 上升在-5到3的範圍中，Ein並沒有變化，也就是錯誤率是相同的，因此在這段區間裡面，我們調整c對於提升正確率沒有效果。
繼續提升C值有可能可以降低錯誤率。

Problem 15 :

此題的module使用與前幾題一樣的sklearn，kernel改成Gaussian kernel ，將參數帶入後，可以求得參數 dual_coef , support vector等，接下來使用以下方式計算“the distance of any free support vector to the hyperplane in the (infinite-dimensional) Z space”。

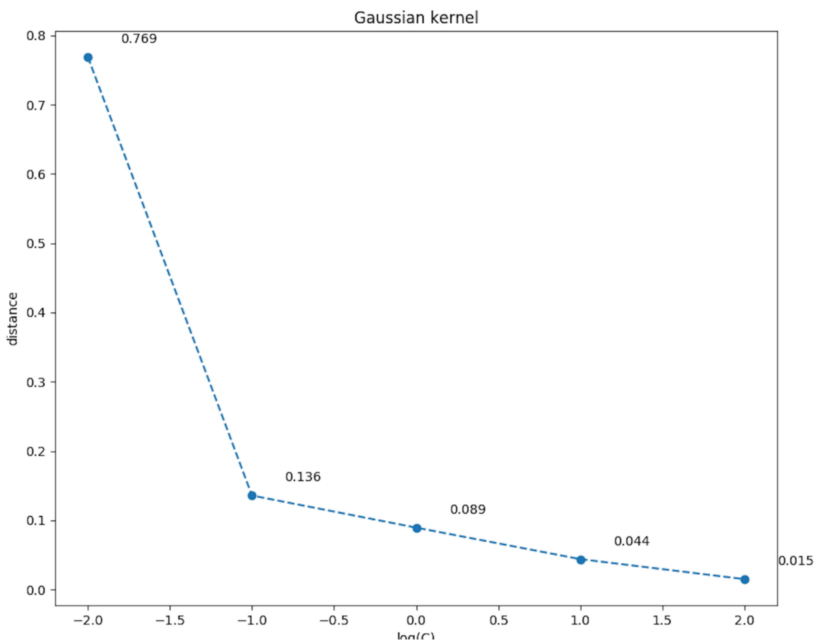
參考講義距離公式：$\frac{1}{\|w\|}\left|w^Tx+b\right|$, in Z Space $x \leftarrow z$。

如果我們挑的是free support vector，則會滿足式子：

$$\alpha_n(1-y_n(w^Tz_n+b))=0 \rightarrow (1-y_n(w^Tz_n+b))=0 \rightarrow |w^Tz_n+b|=1$$

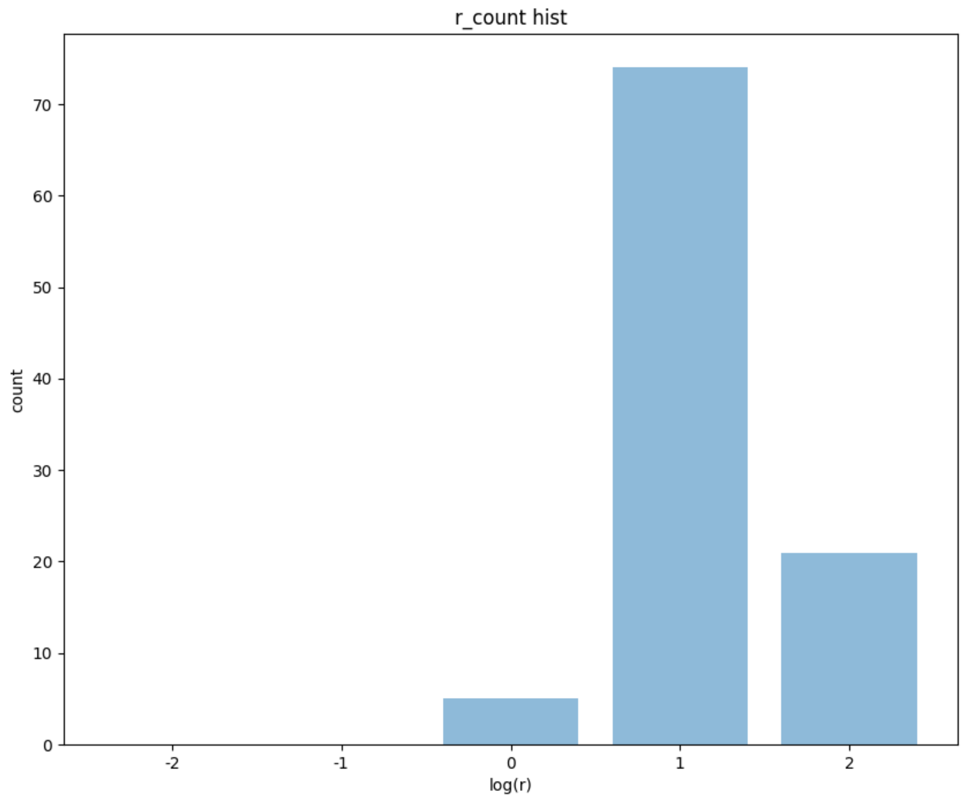
$$\frac{1}{\|w\|}\left|w^Tx+b\right| = \frac{1}{\|w\|}$$

因此distance $= \frac{1}{\|w\|}$。求出後，畫以下圖形：我們會發現所有c值對應的distance會隨c上升而下降，表示margin的大小越來越小，事實上，當我們將c調大時，對於犯錯時的加成會更大，也就是模型會更複雜更貼合support vector。



Problem 16:

在切training data與validation data時，我們使用“from sklearn.model_selection import train_test_split”切割，切割後，帶入與problem 15一樣的module計算SVM數值，重複100次檢驗γ值對應到的錯誤率誰最小，我們得到以下hist ：



我們可以發現，不同的training data分佈，可能會造成不一樣的γ挑選，而大部分的可能分布下，$\gamma = 10$ 的Ein最小，因此如果要選擇的話，我們會選之當作參數。

Problem 17 :

17. 在 Soft-margin SVM 中, $w_i = \sum_{n=1}^{N} \alpha_n y_n \cdot z_{n,i}$ ($\because \frac{\partial L}{\partial w_i} = 0$)

若 z_i = constant 時,

$w_i = z_i \sum_{n=1}^{N} \alpha_n \cdot y_n$. 又因為 $\sum_{n=1}^{N} \alpha_n \cdot y_n = 0$ ($\because \frac{\partial L}{\partial b} = 0$)

$w_i = 0$. 成立。

Problem 18 :

18. $\tilde{g}_{SVM} = \text{sign}\left(\sum_{SV\ in\ n} \tilde{\alpha}_n y_n \tilde{K}(x_n, x') + b\right)$

$= \text{sign}\left(\sum_{SV\ in\ n} \tilde{\alpha}_n y_n (K(x_n, x') + q) + b\right)$

$= \text{sign}\left(\sum_{SV\ in\ n} \tilde{\alpha}_n y_n (K(x_n, x') + \cancel{q}) + y_s - \sum_{SV\ in\ n} \tilde{\alpha}_n y_n (K(x_n, x_s) + \cancel{q})\right)$

$= \text{sign}\left(\sum_{SV\ in\ n} \tilde{\alpha}_n y_n (K(x_n, x') - K(x_n, x_s)) + y_s\right)$

若 $\tilde{C} = C$ $0 \le \alpha_n \le C = \tilde{C}$ 則 $\tilde{\alpha}_n$ upper bound 與 α_n 相同

檢查 QP solver α_n, $\tilde{\alpha}_n$ 值

α_n:

$$\min_{\alpha} \frac{1}{2}\alpha^T Q_D \alpha + 1^T \alpha$$

s.t. $y^T \alpha = 0$

$\alpha_n \ge 0$ for 1, 2, ... N

$\tilde{\alpha}_n$

$$\min_{\tilde{\alpha}} \frac{1}{2}\tilde{\alpha}^T (Q_D + q1)\tilde{\alpha} + 1^T \tilde{\alpha}$$

$$= \min_{\tilde{\alpha}} \frac{1}{2}\tilde{\alpha}^T Q_D \tilde{\alpha} + 1^T \tilde{\alpha} + \frac{1}{2}\tilde{\alpha}^T (q \cdot 1)\tilde{\alpha}$$

s.t. $y^T \cdot \tilde{\alpha} = 0$

$\tilde{\alpha}_n \ge 0$ for 1, 2, ... N

式子不相同, 故所求 α 不一定相同, g_{SVM} 不一定相同.

My Code is available on my Github :

"https://github.com/davidhalladay/Computer-vision-project/tree/master/mini_01"

reference:

1. *SVM , "https://scikit-learn.org/stable/modules/generated/sklearn.svm.SVC.html"*
2. *Matplotlib.contour , "https://matplotlib.org/api/_as_gen/matplotlib.pyplot.contour.html"*
3. *SVM.math , "https://scikit-learn.org/stable/modules/svm.html#svm-kernels"*
4. *SVM parameters details , "http://www.stardustsky.net/index.php/post/53.html"*
5. *Train_test_split , "https://scikit-learn.org/stable/modules/generated/sklearn.model_selection.train_test_split.html"*
6. *Mever's theorem , "https://en.wikipedia.org/wiki/Mercer%27s_theorem"*
7. *Construct kernel func. , "http://mlweb.loria.fr/book/en/constructingkernels.html"*
8. *Standford SVM , "http://cs229.stanford.edu/notes/cs229-notes3.pdf"*