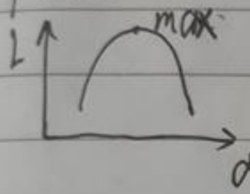
HW3

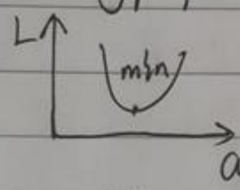
I. DHS p276-Problem 33 (p225-33 in the translated book)

SVM 吴瑞林-电计1701-201785072

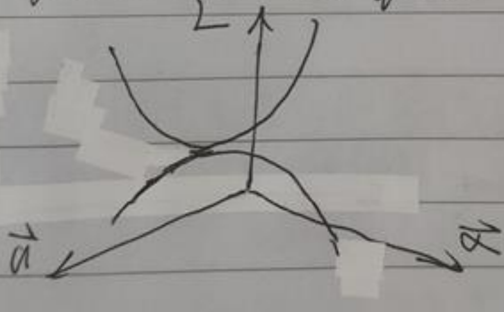
(a) If we want to get optimal solution, we should max L with $\vec{\alpha}$ and min L with a,

max means all patterns are decided correctly, min means the distance between patterns to hyperplane is the max.





to get a saddle point, must make the gap of quesehon and due-questhon be zero.



(b) $\frac{\partial L(a,\alpha)}{\partial a} = 0$, ~~$\partial L/\partial a$~~ = $\partial L/\partial a = \|a\| - \sum_{k=1}^{n} \alpha_k [y_k z_k] = 0$

make $\vec{a} = b\,\vec{w}$, $\partial L/\partial a = \frac{1}{2}\|w\|^2 - \sum_{k=1}^{n} \alpha_k [z_k \vec{w} y_k + z_k b - 1]$

$\partial L/\partial a = 0 \Rightarrow \partial L/\vec{w} = 0$, $\partial L/\partial b = 0$,

$\frac{\partial L}{\partial b} = -\sum_{k=1}^{n} \alpha^*_k z_k = 0$

No.

Date

(c) $\frac{\partial L}{\partial a} = a - \sum_{k=1}^{n} \alpha_k [z_k y_k] = 0$ (matrix_cookbook (69))

$\therefore \quad a^* = \sum_{k=1}^{n} \alpha_k^* z_k y_k$

(d) for suppore vector, $z_k a^t y_k = 1$,

$z_k a^t y_k - 1 = 0$,

for others, because of KKT,

$z_k a^t y_k \neq 1$, $\alpha_k^* = 0$

$\therefore \quad \alpha_k^* [z_k a^t y_k - 1] = 0$

(e) $\tilde{L}(a, \alpha) = \frac{1}{2}\|a\|^2 - \sum_{k=1}^{n} (\alpha_k z_k a^t y_k - \alpha_k)$

$= \frac{1}{2}\|a\|^2 - \sum_{k=1}^{n} \alpha_k z_k a^t y_k + \sum_{k=1}^{n} \alpha_k$

(f) $\tilde{L}(\alpha) = \frac{1}{2}\|\sum_{k=1}^{n} \alpha_k z_k \vec{y}_k\|^2 - \sum_{k=1}^{n} \alpha_k z_k \vec{a}^t \vec{y}_k + \sum_{k=1}^{n} \alpha_k$

$= \frac{1}{2}\left(\sum_{i=1}^{n} \sum_{j=1}^{n} \alpha_i \alpha_j z_i z_j (y_i^t y_j)\right) - \sum_{k=1}^{n} \alpha_k z_k \sum_{i=1}^{n} \alpha_i z_i \vec{y}_i \vec{y}_k$

$= ANS - \sum_{i=1}^{n} \sum_{j=1}^{n} \alpha_i \alpha_j z_i z_j (y_i^t y_j) + \sum_{k=1}^{n} \alpha_k$

$= -\frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} \alpha_i \alpha_j z_i z_j (y_i^t y_j) + \sum_{k=1}^{n} \alpha_k$

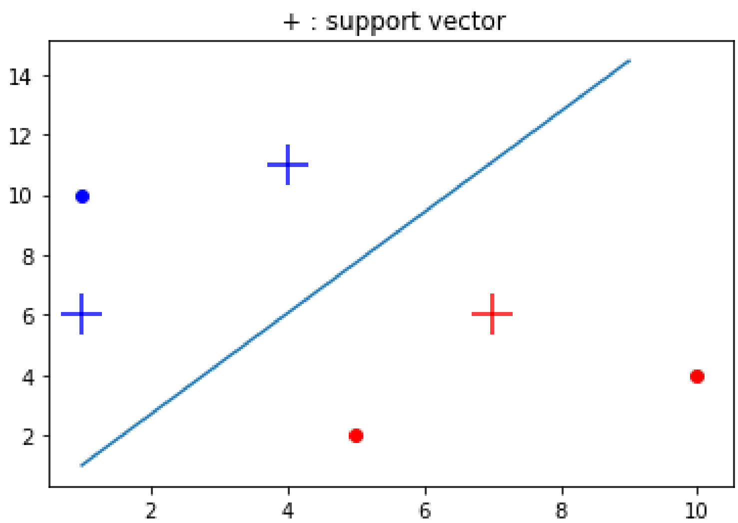
$\therefore \quad \tilde{L}(\alpha) = -\frac{1}{2} \sum_{j,k=1}^{n} \alpha_j \alpha_k z_j z_k (y_j^t y_k) + \sum_{j=1}^{n} \alpha_j$

II Support Vector Machine (SVM)

In this problem, we will write a program to implement the SVM algorithm. Let us start with a toy example (which can be found at SVM matlab Prof olga Veksler.pdf) and then work on more complicated cases. The toy example (credit goes to Prof. Olga Veksler, University of Western Ontario) provides detailed implementation of SVM using Matlab. It is noted that this example works in the original feature space, rather than the augmented one

(a) Try the toy example, and plot the separating hyperplane g(x) = wtx+ w0 and the support vectors

```
alpha
 [[0.036]
 [0.   ]
 [0.04 ]
 [0.   ]
 [0.076]
 [0.   ]]
w [[-0.336  0.2  ]]
w0 [0.136]
```



(b) Train a SVM classifier with TrainSet1.txt, and plot the separating hyperplane and the support vectors.

alpha

[[1.68089774e-10] [1.58970349e-10] [1.27553695e-10] [5.50583818e-11] [8.71972020e-11] [1.07107633e-10] [1.09417379e-10] [1.05125769e-10] [1.03217359e-10] [6.02546155e-10] [6.20334721e-11] [6.72733473e-11] [1.90915524e-10] [4.29591356e-11] [4.91728582e-11] [4.24599334e-11] [1.61220481e-11]

[2.61429378e-11] [3.89468609e-03] [1.44791076e-03] [1.02260988e-10] [-2.49633312e-11] [-2.84975447e-11] [-1.77829149e-11]

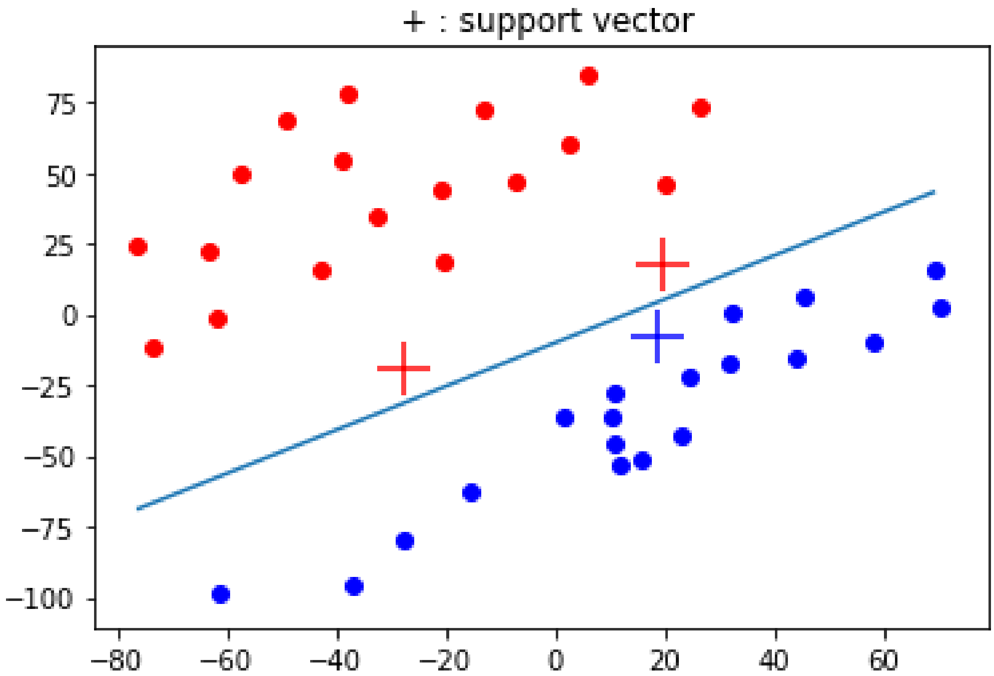
[8.11319178e-11] [8.95448188e-11] [-2.13457366e-11] [-1.45194235e-11] [-1.64119674e-11] [-2.71824843e-11] [2.43045281e-11]

[5.23836093e-10] [5.34259160e-03] [5.62120742e-09] [7.88387657e-10] [3.58266519e-10] [-2.67521656e-11] [-1.63957496e-11]

[1.12594662e-11] [-3.65843179e-11]]

w [[0.0631346   -0.08184873]]

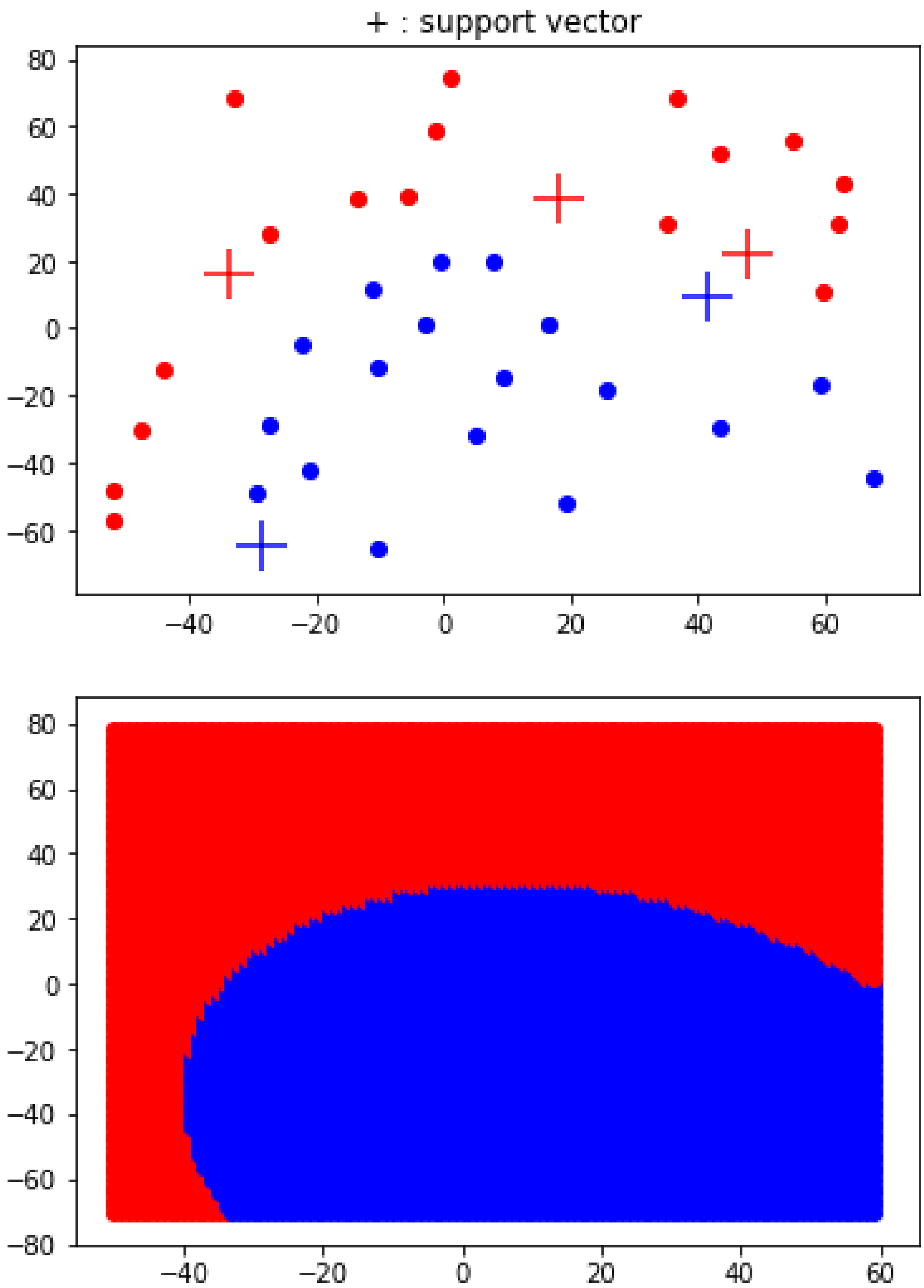
w0 [-0.82449267]



(c) Design a quadratic kernel and train a SVM classifier with TrainSet2.txt. Plot the separating boundary and support vectors in the original feature space.

The kernel make x from 2D to 5D, and do SVM in 5D space

Newx=[x[:,0]**2,x[:,1]**2,x[:,0]*x[:,1],x[:,0],x[:,1]]


+ : support vector
80
60
40
20
0
−20
−40
−60
−40
−20
0
20
40
60
80
60
40
20
0
−20
−40
−60
−80
−40
−20
0
20
40
60