HW8. Support Vector Machines

Name - Sayem Lincoln
PID - A54207835

1) we have $Q = \begin{bmatrix} 0 & 0_d^T \\ 0_d & I_d \end{bmatrix}$ where $\begin{cases} 0_d^T = [0 \dots 0] \in \mathbb{R}^d \\ I_d = d \times d \text{ identity matrix} \end{cases}$

We have to prove that Q is a positive semi-definite matrix

Q positive semideflite $\Leftrightarrow u^T Q u \geq 0$ for any u

for $u \in \mathbb{R}^{d+1} \Rightarrow u = \begin{bmatrix} u_1 \\ \vdots \\ u_{d+1} \end{bmatrix}$

$$u^T Q u = \sum_{i=2}^{d+1} u_i^2 \geq 0$$

$\Rightarrow$ Thus Q is positive semi-definite.

$$Q = \left(\begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \begin{pmatrix} \overbrace{0 \; \cdots \; 0}^{d} \\ 1 \quad\quad 0 \\ \ddots \\ 0 \quad\quad 1 \end{pmatrix} \right)$$

2) a) minimize $\frac{1}{2} \sum_{m=1}^{N} \sum_{n=1}^{N} y_n y_m \alpha_n \alpha_m x_n^T x_m - \sum_{n=1}^{N} \alpha_n$ (i)

subject to $\sum_{n=1}^{N} y_n \alpha_n = 0$; $\alpha_n \geq 0$ $(n = 1, \dots, N)$ (ii)

Problem written as $\Rightarrow$

minimize $\Rightarrow \frac{1}{2} u^T Q u + p^T u$

for any u subject to $\Rightarrow a_m^T u \geq c_m$ $(m = 1, \dots, N)$

we begin by simplifying the expression of (i) $\Rightarrow$

$$\frac{1}{2} \sum_{m=1}^{N} \sum_{n=1}^{N} y_n y_m \alpha_n \alpha_m x_N^T x_M - \sum_{n=1}^{N} \alpha_m =$$

$$= \frac{1}{2} [\alpha_1 \cdots \alpha_N] \underbrace{\begin{bmatrix} y_1 y_1 x^T x_1 & \cdots & y_1 y_N x_1^T x_N \\ y_2 y_1 x_2^T x_1 & \cdots & y_2 y_N x_2^T x_N \\ \vdots & & \vdots \\ y_N y_1 x_N^T x_1 & \cdots & y_N y_N x_N^T x_N \end{bmatrix}}_{D} \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_N \end{bmatrix} -$$

$$[1 \cdots 1] \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_N \end{bmatrix}$$

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

We simplify the expression (ii) $\Rightarrow$

we have $\sum_{m=1}^{N} y_n \alpha_n = 0 \Rightarrow \begin{cases} \sum_{m=1}^{N} y_n \alpha_n \geq 0 \\ \sum_{m=1}^{N} y_m \alpha_n \leq 0 \end{cases}$

$$\Rightarrow \begin{cases} \sum_{m=n}^{N} y_n \alpha_n \geq 0 \\ \sum_{m=n}^{N} -y_n \alpha_n \geq 0 \end{cases}$$

we have also $\alpha_m \geq 0$ $(m = 1, \ldots, N)$

$$\Rightarrow \begin{pmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{pmatrix} \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_N \end{bmatrix} \geq 0_N$$

Thus we can write $\Rightarrow$

$$\begin{bmatrix} [y_1 \cdots y_N] \\ [-y_1 \cdots -y_N] \\ \begin{bmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{bmatrix} \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_N \end{bmatrix} \geq 0_{N+2}$$

$$\Rightarrow A_D \times \alpha \geq 0_{N+2}$$

Finally, we conclude that the problem is a standard QP-Problem $\Rightarrow$

$$\begin{cases} \min_{\alpha \in \mathbb{R}} \frac{1}{2} \alpha^T Q_D \alpha - I_N^T \alpha \end{cases}$$

$$Q_D = \begin{bmatrix} y_1 y_1 x_1^T x_1 & \cdots & y_1 y_N x_1^T x_N \\ \vdots & & \vdots \\ y_N y_1 x_N^T x_1 & \cdots & y_N y_N x_N^T x_N \end{bmatrix} \quad \text{and} \quad A_D = \begin{bmatrix} y^T \\ -y^T \\ I_{N \times N} \end{bmatrix}$$

b) We suppose $\Rightarrow$

$$X_s = \begin{bmatrix} -y_1 x_1^T \\ | \\ -y_N x_N^T \end{bmatrix}$$

$$X_s X_s^T = \begin{bmatrix} -y_1 x_1^T \\ | \\ -y_N x_N^T \end{bmatrix} \begin{bmatrix} -y_1 x_1^T & \text{———} & -y_N x_N^T \end{bmatrix}$$

$$= \underbrace{\begin{bmatrix} y_1 y_1 x_1^T x_1 & \text{———} & y_1 y_N x_1^T x_N \\ | & & | \\ y_N y_1 x_N^T x_N & \text{———} & y_N y_N x_N^T x_N \end{bmatrix}}_{Q_D}$$

$$= Q_D$$

$$\Rightarrow Q_D = X_s X_s^T$$

Q_D is positive semi-definite $\Longleftrightarrow u^T Q_D u \geq 0$ for any u

$$\Longleftrightarrow u^T X_s X_s^T u \geq 0$$

$$\Longleftrightarrow (u^T X_s)(u^T X_s)^T \geq 0$$

$$\Longleftrightarrow \|u^T X_s\|^2 \geq 0$$

Thus, Q_D is positive semi-definite.

3) Assume D (data set) with two data points $(x_+, +1)$ and $(x_-, -1)$.
We compute hyperplane and its margin (b^*, w^*).
The two seperation constraints are $\Rightarrow$

$$\begin{cases} (w^T x_+ + b) \geq 1 & \text{— (i)} \\ \text{and} \\ -(w^T x_- + b) \geq 1 & \text{— (ii)} \end{cases}$$

By adding (i) and (ii) $\Rightarrow w^T(x_+ - x_-) \geq 2$

The Cauchy - Schwarz Inequality gives us $\Rightarrow$

$$\Rightarrow 2 \leq w^T(x_+ - x_-) \leq |w^T(x_+ - x_-)| \leq \|w\| \cdot \|x_+ - x_-\|$$

So we get $\Rightarrow$ $\|w\| \geq \dfrac{2}{\|x_+ - x_-\|}$

Thus $\Rightarrow$ $\|w^*\| = \dfrac{2}{\|x_+ - x_-\|}$ (The minimum)

In this case we want w^* to satisfy both constraints (1) & (4). This means that $\Rightarrow$

$$\left| w^{*T}(x_+ - x_-) \right| = \|w^*\| \cdot \|x_+ - x_-\|$$

This can happen only if $\Rightarrow$

$$w^* = k(x_+ - x_-)$$

but we have $\|w^*\| = \dfrac{2}{\|x_+ - x_+\|}$

$$\Rightarrow k = \frac{2}{\|x_+ - x_-\|^2}$$

Finally we may write $\Rightarrow$ $w^* = \dfrac{2(x_+ - x_-)}{\|x_+ - x_-\|^2}$

It remains to be determined b^* $\Rightarrow$

we have $\Rightarrow$ $2\left(\dfrac{(x_+ - x_-)}{\|x_+ - x_-\|^2}\right)^T x_+ + b^* = 1$

$$\Rightarrow b^* = 1 - 2\,\frac{x_+^T x_+ - x_-^T x_+}{\|x_+ - x_-\|^2}$$

$$= \frac{\|x_-\|^2 - \|x_+\|^2}{\|x_+ - x_-\|^2}$$

Thus, $b^* = \dfrac{\|x_-\|^2 - \|x_+\|^2}{\|x_+ - x_-\|^2}$

Finally, we can say that (w^*, b^*) satisfies both constraints and minimizes $\|w\|$, and therefore gives us the optimal hyperplane.

4) We have a data set of three data points in $\mathbb{R}^2 \Rightarrow$

$$X = \begin{pmatrix} 0 & 0 \\ 0 & -1 \\ -2 & 0 \end{pmatrix} \qquad y = \begin{pmatrix} -1 \\ -1 \\ +1 \end{pmatrix}$$

In this case the constraints are $\Rightarrow$

$$\begin{cases} -b \geq 1 & — \text{(i)} \\ -(w_2 + b) \geq 1 & — \text{(ii)} \\ (-2w_1 + b) \geq 1 & — \text{(iii)} \end{cases}$$

We combine (1) and (11) $\Rightarrow$ $w_1 \leq -1$

if we consider $W = \begin{pmatrix} w_1 \\ w_2 \end{pmatrix}$, the quantity we seek to minimize is $\Rightarrow$ $\frac{1}{2} w^T w = \frac{1}{2}(w_1^2 + w_2^2) \geq \frac{1}{2}(1+0) \geq \frac{1}{2}$

We have equality when $w_1 = -1$ and $w_2 = 0$. So, when $w^* = (-1, 0)$, the constraint (iii) gives us $\Rightarrow$ $b \geq -1$.

So we choose $b^* = -1$.

Thus, we verify (w^*, b^*) satisfies all constraints and minimizes $\|w\|$ and therefore gives us the optimal hyperplane.

NB the margin in this case is $\dfrac{1}{\|w^*\|} = 1$.

Problem 5 Code output
<-----

```
     pcost       dcost       gap    pres   dres
 0: -1.2161e+02 -3.4142e+02  1e+03  3e+01  3e+00
 1: -5.2504e+02 -1.0309e+03  1e+03  2e+01  2e+00
 2: -1.8811e+03 -2.5912e+03  8e+02  1e+01  1e+00
 3: -5.4551e+03 -6.4620e+03  1e+03  1e+01  1e+00
 4: -1.9177e+04 -2.1331e+04  2e+03  1e+01  1e+00
 5: -7.2845e+04 -7.8706e+04  6e+03  1e+01  1e+00
 6: -2.6268e+05 -2.7970e+05  2e+04  1e+01  1e+00
 7: -1.2219e+06 -1.2834e+06  6e+04  1e+01  1e+00
 8: -9.4425e+06 -9.7643e+06  3e+05  1e+01  1e+00
 9: -1.9505e+08 -1.9739e+08  2e+06  1e+01  1e+00
10: -1.5176e+10 -1.5192e+10  2e+07  1e+01  1e+00
11: -1.7305e+10 -1.7324e+10  2e+07  1e+01  1e+00
12: -1.9347e+10 -1.9368e+10  2e+07  1e+01  1e+00
13: -3.4497e+11 -3.4533e+11  4e+08  1e+01  1e+00
14: -4.6438e+11 -4.6487e+11  5e+08  1e+01  1e+00
15: -4.7433e+11 -4.7483e+11  5e+08  1e+01  1e+00
16: -5.8783e+11 -5.8845e+11  6e+08  1e+01  1e+00
17: -1.3542e+12 -1.3556e+12  1e+09  1e+01  1e+00
18: -2.3250e+12 -2.3271e+12  2e+09  1e+01  1e+00
Terminated (singular KKT matrix).
w =  [ 3.05175781e-05 -1.22070312e-04 -1.83105469e-04 -2.89916992e-04
  1.67846680e-04  1.83105469e-04  0.00000000e+00  2.30789185e-04
 -7.62939453e-05  1.37329102e-04 -1.52587891e-05 -2.13623047e-04
 -4.57763672e-05]

b =  [1.00061076]
     pcost       dcost       gap    pres   dres
 0: -1.0917e+02 -1.9548e+02  1e+04  1e+02  2e+00
 1: -2.5975e+01 -8.2113e+00  2e+03  2e+01  3e-01
 2: -5.6936e-01 -4.4739e-01  5e+01  5e-01  8e-03
 3: -1.0950e-01 -1.3442e-01  6e+00  6e-02  1e-03
 4: -3.2196e-02 -5.4630e-02  1e+00  1e-02  2e-04
 5: -1.0622e-02 -3.3207e-02  4e-01  4e-03  6e-05
 6: -4.1912e-03 -2.3472e-02  2e-01  1e-03  2e-05
 7: -4.2660e-03 -1.6118e-02  8e-02  6e-04  9e-06
 8: -5.6269e-03 -8.6678e-03  2e-02  1e-04  2e-06
 9: -5.8634e-03 -5.8987e-03  6e-04  4e-06  6e-08
10: -5.8737e-03 -5.8740e-03  6e-06  4e-08  6e-10
11: -5.8738e-03 -5.8738e-03  6e-08  4e-10  6e-12
Optimal solution found.
w =  [-3.28305305e-11 -2.54024062e-10 -8.16977670e-10 -1.97164389e-09
 -3.29201747e-09 -6.08261872e-09 -4.69466584e-09  1.21107466e-08
 -7.04056011e-09 -2.96726194e-08 -2.63955109e-08 -1.76340118e-08
 -9.87410280e-09 -4.64008554e-09 -1.60132777e-09 -3.83735192e-10
 -1.25173491e-10 -2.58377263e-10 -1.58348754e-09 -4.62551085e-09
 -9.75406465e-09 -1.90203836e-08 -1.70212927e-08  7.35532512e-09
  3.04353769e-09 -2.71300519e-08 -3.66912638e-08 -3.17225374e-08
 -1.95357080e-08 -1.00637619e-08 -4.34128225e-09 -1.18266042e-09
 -1.51314282e-10 -4.45674497e-10 -2.30122505e-09 -7.38592971e-09
 -1.80969293e-08 -2.68025293e-08 -1.25931586e-08  2.09388383e-08
  2.58314862e-08 -2.15324546e-09 -2.26599187e-08 -2.98070182e-08
 -2.17633359e-08 -1.12738795e-08 -5.17528445e-09 -1.61776781e-09
 -2.74491418e-10 -7.59217871e-10 -2.85786622e-09 -1.18881428e-08
 -2.41576519e-08 -2.46852636e-08 -4.06834200e-09  3.10613525e-08
  3.73763558e-08  9.28084413e-09 -1.52068572e-08 -2.56437189e-08
 -2.01730936e-08 -1.05122789e-08 -4.34278099e-09 -1.20950804e-09
 -2.88459682e-10 -1.37648390e-09 -5.35174987e-09 -1.63129310e-08
 -2.66213424e-08 -2.40482455e-08 -2.55917385e-09  3.27186183e-08
  3.79302812e-08  9.12333711e-09 -1.64479498e-08 -2.46565715e-08
 -1.77480495e-08 -8.64003485e-09 -2.95036629e-09 -6.21257918e-10
 -6.00521842e-10 -2.29012908e-09 -6.98918688e-09 -1.77127951e-08
 -2.88140878e-08 -2.74453994e-08 -5.77751460e-09  2.91086954e-08
  3.23561615e-08  1.78056423e-09 -2.03335957e-08 -2.30450044e-08
 -1.43033477e-08 -5.78002745e-09 -1.62414576e-09 -3.07400750e-10
 -7.45570778e-10 -2.61721212e-09 -7.16140296e-09 -1.57899583e-08
 -2.64501042e-08 -2.69208575e-08 -8.06494753e-09  2.41051105e-08
  2.54616466e-08 -4.87318072e-09 -2.46891337e-08 -2.08003966e-08
 -1.03890975e-08 -3.68561191e-09 -1.00788829e-09 -9.94896848e-11
 -4.19213173e-10 -1.97536099e-09 -5.82834774e-09 -1.31961759e-08
```

```
 -2.10471300e-08 -2.15752309e-08 -6.19947587e-09  2.18678614e-08
  2.15174030e-08 -1.09397488e-08 -2.60346883e-08 -1.65996095e-08
 -7.04772221e-09 -2.42165373e-09 -6.11253747e-10  1.93970878e-10
 -1.93100646e-10 -9.69353362e-10 -3.45290076e-09 -8.52922552e-09
 -1.35313805e-08 -1.36890393e-08  3.25798345e-11  2.43729226e-08
  1.84966263e-08 -1.73597709e-08 -2.61493406e-08 -1.29379802e-08
 -5.12279898e-09 -1.71890987e-09 -2.47726171e-10  3.26735388e-10
  5.71877145e-10  1.26067200e-09 -3.60867714e-10 -3.29826202e-09
 -6.02977569e-09 -5.47340323e-09  9.09382950e-09  3.00071981e-08
  1.52574694e-08 -2.26525326e-08 -2.32521334e-08 -9.71815260e-09
 -4.32746517e-09 -1.17270048e-09  1.66201917e-10  4.89500516e-10
  2.71109619e-09  4.61660507e-09  2.88197666e-09  8.18878999e-10
 -1.33043631e-09  1.15829720e-11  1.49510565e-08  3.31739147e-08
  1.22240475e-08 -2.22923387e-08 -1.80973971e-08 -6.67212784e-09
 -2.81953585e-09 -2.77185324e-10  1.41825256e-09  1.63910171e-09
  4.88991611e-09  7.75539061e-09  5.74592414e-09  1.84674183e-09
 -7.67431785e-10  4.29979850e-10  1.72491833e-08  3.20585127e-08
  1.21000749e-08 -1.49555055e-08 -1.11806267e-08 -3.87599430e-09
 -1.29652914e-09  1.59310860e-09  3.40000483e-09  2.17726763e-09
  4.60511024e-09  8.93695821e-09  6.51870243e-09  1.38508473e-09
 -3.00005893e-09 -2.56791412e-09  1.34029759e-08  2.71085877e-08
  1.33474685e-08 -8.33218755e-09 -9.95068927e-09 -3.89125223e-09
  6.88829402e-10  4.07588628e-09  4.98944655e-09  2.98262482e-09
  3.45849079e-09  8.73404309e-09  8.82235546e-09  3.72622894e-09
 -3.58992041e-09 -7.67802903e-09  3.83167493e-09  1.75250548e-08
  9.51404910e-09 -1.07173968e-08 -1.38948331e-08 -4.94793524e-09
  3.08807560e-09  6.46880455e-09  6.58750146e-09  2.92493943e-09
  2.18143683e-09  6.30373775e-09  8.47691698e-09  5.99734523e-09
 -1.09559613e-09 -1.17268762e-08 -1.29158016e-08  5.07253343e-10
 -1.71524037e-09 -1.58025609e-08 -1.37638956e-08 -2.63317566e-09
  3.78391202e-09  5.53866412e-09  5.15985111e-09  2.30117203e-09
  5.56342597e-10  2.11035998e-09  3.83398346e-09  4.45171114e-09
  2.28623347e-09 -5.29758541e-09 -1.74859032e-08 -1.05813591e-08
 -7.33702785e-10 -9.18681335e-09 -5.75484405e-09 -9.94979456e-10
  1.55721452e-09  2.52908535e-09  2.12738196e-09  6.94154941e-10]

 =  []
unning on sample size 50...
unning on sample size 100...
unning on sample size 200...
unning on sample size 500...
unning on sample size 1000...
unning on sample size 2000...
unning on sample size 3000...
unning on sample size 4000...
unning on feature size 5...
unning on feature size 10...
unning on feature size 20...
unning on feature size 50...
unning on feature size 100...
unning on feature size 300...
unning on feature size 500...
```