AI2 - Ex 08

Q1 (a) → Std. QP form is:

$$\min \frac{1}{2} x^T Q x + c^T x \Rightarrow \frac{1}{2}\sum_{i=1}^{m}\sum_{j=1}^{m} x_i Q_{ij} x_j + \sum_{i=1}^{m} c_i x_i$$

subject to $Ax \le b$

$A_{eq}\, x = b_{eq}$

~~lb ≤ x ≤ ub~~

→ Given QP is:

$$\max -\frac{1}{2}\sum_{i=1}^{m}\sum_{j=1}^{m}(\alpha_i-\alpha_i')(\alpha_j-\alpha_j')K(x_i,x_j)$$

$$-\sum_{i=1}^{m}\varepsilon(\alpha_i+\alpha_i') + \sum_{i=1}^{m} y_i(\alpha_i-\alpha_i')$$

Sub. to: $\sum_{i=1}^{m}(\alpha_i-\alpha_i') = 0$

$0 \le \alpha, \alpha' \le C$

→ $\max -\frac{1}{2}(\alpha-\alpha')^T K(\alpha-\alpha') + \sum_{i=1}^{m}\alpha_i(y_i-\epsilon)$

$+ \sum_{i=1}^{m}\alpha_i'(-y_i-\epsilon)$

$\Rightarrow \min \frac{1}{2}\left[\alpha^T K\alpha + \alpha'^T K\alpha - \alpha^T K\alpha' - \alpha'^T K\alpha\right]$

$+\sum_{i=1}^{m}\alpha_i(\epsilon-y_i) + \sum_{i=1}^{m}\alpha_i'(\epsilon+y_i)$

~~⇒ min ½ α^T K~~

~~⇒ min~~

Block matrix $(2m \times 2m)$

$$\Rightarrow \min \quad \frac{1}{2}\left(\begin{bmatrix}\alpha \\ \alpha'\end{bmatrix}^T \begin{bmatrix}K & -K \\ -K & K\end{bmatrix} \begin{bmatrix}\alpha \\ \alpha'\end{bmatrix}\right) + c\begin{bmatrix}\alpha \\ \alpha'\end{bmatrix}$$

where c $(\in \mathbb{C}^{2m\times 1})$ $= \begin{cases} \epsilon - y_i, & i \in [1, m] \\ \epsilon + y_{i-m}, & i \in [m+1, 2m] \end{cases}$

For converting this to standard QP.

Let $x = \begin{bmatrix}\alpha \\ \alpha'\end{bmatrix} \in \mathbb{C}^{2m\times 1}$

$c \in \mathbb{C}^{2m\times 1} = \begin{cases} \epsilon - y_i, & i \in [1, m] \\ \epsilon + y_{i-m}, & i \in [m+1, 2m] \end{cases}$

$Q = \begin{bmatrix} K & -K \\ -K & K \end{bmatrix} \in \mathbb{C}^{2m\times 2m}$

To derive the constraints.

$A_{eq} = \begin{cases} 1, & i \in [1, m] \\ -1, & i \in [m+1, 2m] \end{cases} \in \mathbb{C}^{1\times 2m}$

$b_{eq} = 0$

~~A = 0 b = 0~~

~~$l_b = 0$~~, u_b

~~$u_b = \{ \epsilon, \; i \in [1, 2m]$~~

~~Ae~~ For A & b

$0 \le x_i \le C \Rightarrow \begin{bmatrix} -I_{2m \times 2m} \\ I_{2m \times 2m} \end{bmatrix}^T x \le \begin{bmatrix} 0 \\ C \end{bmatrix}^T$

$A = \begin{bmatrix} -I_{2m \times 2m} \\ I_{2m \times 2m} \end{bmatrix}^T$

$b = \begin{cases} 0, & i \in [1, 2m] \\ C, & i \in [2m+1, 4m] \end{cases} \in \mathbb{C}^{4m \times 1}$

(d) From the optimization of C & ϵ, it was observed that when ϵ was kept too small, the regression curve ~~graph~~ tends to be less smooth & if ϵ is kept too large the regression curve tends to be not fit the data properly.

Also when C is large, the ~~graph~~ reg. curve tends to be not very smooth, while when C is small, the reg. curve doesn't fit the data properly.

In the above context, reg. curve tends to be not very smooth means that it fits the data well but becomes highly complex. ~~to be generalised for the on making ...~~