Keertana . Chidambaram (PS3)

Problem 1

Dual Optimization Problem:

$$\max_{\alpha} \left\{ \sum_i \alpha_i - \frac{1}{2} \sum_{i,j} \alpha_i \alpha_j y_i y_j \phi(x_i).\phi(x_j) \right\}$$

$$\text{s.t.} \quad \sum_i \alpha_i y_i = 0$$

$$0 \leq \alpha_i \leq C \quad \forall i$$

canonical form of a QP is:

$$\underset{\alpha}{\operatorname{argmin}} \quad \frac{1}{2} \alpha^T H \alpha + f^T \alpha$$

$$\text{s.t.} \quad A\alpha \leq a$$

$$B\alpha = b$$

Dual opt. problem can be rewritten as:

$$\min_{\alpha} \left\{ -\sum_i \alpha_i + \frac{1}{2} \sum_{i,j} \alpha_i \alpha_j y_i y_j \phi(x_i).\phi(x_j) \right\} \quad -①$$

$$\text{s.t} \quad \sum y_i \alpha_i = 0, \quad -②$$

$$1.\alpha_i \leq C \quad -③ \quad \forall i$$

$$-1.\alpha_i \leq 0 \quad -④.$$

$$② \Rightarrow [y_1 \; y_2 \; .. \; y_n] \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{bmatrix} = \sum_i y_i \alpha_i = 0$$

So, $\boxed{\begin{array}{l} B = [y_1 \; y_2 \; .. \; y_n] = y^T \\ b = 0 \text{ (1x1 scalar number)} \end{array}}$

$$③ \Rightarrow I\alpha_i \le \begin{bmatrix} c \\ \vdots \\ c \end{bmatrix}$$

$$④ \Rightarrow -I\alpha_i \le \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}$$

$$\boxed{\text{so } A = \begin{bmatrix} I \\ -I \end{bmatrix}_{2n \times n} \qquad a = \begin{bmatrix} c \\ \vdots \\ c \\ 0 \\ \vdots \\ 0 \end{bmatrix}_{2n \times 1} \begin{matrix} \}n \\ \}n \end{matrix}}$$

$$① \Rightarrow [-1 \cdots -1] \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{bmatrix}$$

$$+ [\alpha_1 \cdots \alpha_n] \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} [y_1 \cdots y_n] \begin{bmatrix} \phi(x_1)\phi(x_1) & \cdots & \phi(x_1)\phi(x_n) \\ \vdots & & \\ \phi(x_1)\phi(x_n) & \cdots & \phi(x_n)\phi(x_n) \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{bmatrix}$$

So $\boxed{f = \begin{bmatrix} -1 \\ \vdots \\ -1 \end{bmatrix} \qquad H = yy^T K}$ [where K is the Kernel matrix NOT the function]

Problem 2

[refer to PRML]

α_is have all been found.

Assumption: at least 1 support vector lies on the margin. AND the optimization solves for α_is with an upper bound, i.e. $0 \leq \alpha_i \leq C$.

Therefore, support vectors with $0 < \alpha_i < C$ will be such that they lie on the margin.

[$\alpha_i = 0$ means they are not support vectors & $\alpha_i = C$ means margin < 1].

given that
$$h(x) = \text{sign}\left(\sum_{i:\alpha_i>0} \alpha_i y_i K(x_i, x) + b \right)$$

so for the support vectors with margin 1,
Lets call these $\{(x_m, y_m)\}$ as set N_M.

Since margin is 1,
$$y_m h(x_m) = 1 \quad \forall (x_m, y_m) \in N_M$$

$$\Rightarrow \quad y_m \left[\sum_{i:\alpha_i>0} \alpha_i y_i K(x_i, x_m) + b \right] = 1 \quad \forall (x_m, y_m) \in N_M$$

Solving for b and taking the average over the set N_M,

$$\boxed{b = \frac{1}{|N_M|} \sum_{m \in M} \left(y_m - \sum_{i:\alpha_i>0} \alpha_i y_i K(x_i, x_m) \right)}$$