Problem set #3 ZIYAO YANG 9490544195

P1 $L = v_{M+1}^T X^T X v_{M+1} - \alpha_0 (v_{M+1}^T v_{M+1} - 1) - \sum_{i=1}^{M} \alpha_i v_{M+1}^T v_i$,

$\alpha_0 \lessgtr 0, \ \alpha_i \lessgtr 0, \ i=1 \dots M$

$$\frac{\partial L}{\partial v_{M+1}} = 2X^T X v_{M+1} - 2\alpha_0 v_{M+1} - \sum_{i=1}^{M} \alpha_i v_i = 0$$

$\xrightarrow[v_1^T \text{ both side}]{\text{multiply by}}$ $2v_1^T X^T X v_{M+1} - 2\alpha_0 v_1^T v_{M+1} - \sum_{i=2}^{M} \alpha_i v_i - \alpha_1 v_1^T v_1 = 0$

since $v^T v_i = 0$, $\|v\|_2 = 1$

$2 v_1^T X^T X v_{M+1} = \alpha_1$

$\Rightarrow 2(X^T X v_1)^T \cdot v_{M+1} = \alpha_1$ $(X^T X v_1 = \lambda_1 v_1)$

$\Rightarrow 2\lambda_1 v_1^T v_{M+1} = \alpha_1$ $(v_1^T v_{M+1} = 0)$

$\Rightarrow \alpha_1 = 0$

similarly $\alpha_1 = \alpha_2 = \cdots \alpha_M = 0$

so $\frac{\partial L}{\partial v_{M+1}} = 2X^T X v_{M+1} - 2\alpha_0 v_{M+1} = 0$

$X^T X v_{M+1} = \alpha_0 v_{M+1}$

Therefore $\alpha_0 = \lambda_{M+1}$ and v_{M+1} is eign vector with eign value λ_{M+1}

$\max v^T X^T X v \Rightarrow \max (v_{M+1}^T X^T X v_{M+1})$

$= \max (\lambda_{M+1} v_{M+1}^T \cdot v_{M+1})$

$= \lambda_{M+1}$

P2

1) By introducing slack variables

$$\min_{\vec{w}, b} \frac{1}{2}\|\vec{w}\|_2^2 + C\sum_{i=1}^{N}(\xi_i + \xi_i')$$

w.t

$$f(x_i) - y_i \le \epsilon + \xi_i$$
$$y_i - f(x_i) \le \epsilon + \xi_i'$$
$$\xi_i \ge 0, \quad \xi_i' \ge 0, \quad i = 1, 2 \dots N$$

2) $L(\vec{w}, b, \vec{\alpha}, \vec{\alpha}', \vec{\xi}, \vec{\xi}', \vec{\mu}, \vec{\mu}')$

$$= \frac{1}{2}\|\vec{w}\|^2 + C\sum_{n=1}^{N}(\xi_n + \xi_n') + \sum_{n=1}^{N}\alpha_n(f(x_n) - y_n - \epsilon - \xi_n)$$
$$+ \sum_{n=1}^{N}\alpha_n'(y_n - f(x_n) - \epsilon - \xi_n') - \sum_{i=1}^{N}\mu_i\xi_i - \sum_{i=1}^{N}\mu_i'\xi_i'$$

$$= \frac{1}{2}\|\vec{w}\|^2 + C\sum_{n=1}^{N}(\xi_n + \xi_n') + \sum_{n=1}^{N}\alpha_n(w^T\vec{x}_n + b - y_n - \epsilon - \xi_n)$$
$$+ \sum_{n=1}^{N}\alpha_n'(y_n - \vec{w}^T\vec{x}_n - b - \epsilon - \xi_n') - \sum_{i=1}^{N}\mu_i\xi_i - \sum_{i=1}^{N}\mu_i'\xi_i'$$

$$\frac{\partial L}{\partial w} = w + \sum_{n=1}^{N}\alpha_n x_n - \sum_{n=1}^{N}\alpha_n' x_n = 0$$

$$\frac{\partial L}{\partial b} = \sum_{n=1}^{N}\alpha_n - \sum_{n=1}^{N}\alpha_n' = 0$$

$$\frac{\partial L}{\partial \xi_n} = C - \mu_n - \alpha_n = 0$$

$$\frac{\partial L}{\partial \xi_n'} = C - \mu_n' - \alpha_n' = 0$$

plug into L

$$L = \sum_{i=1}^{N} y_i(\alpha_i' - \alpha_i) - \epsilon\sum(\alpha_i' + \alpha_i) + C\cancel{\sum_{i=1}^{N}(\xi_i + \xi_i')}$$

$$+ \frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N}(\alpha_i' - \alpha_i)(\alpha_j' - \alpha_j)\vec{x_i}^T\vec{x_j}$$

$$- C\cancel{\sum_{i=1}^{N}\xi_i} + \cancel{\sum_{i=1}^{N}\alpha_i\xi_i} - C\cancel{\sum_{i=1}^{N}\xi_i'} + \cancel{\sum_{i=1}^{N}\alpha_i'\xi_i'}$$

$$- \sum_{i=1}^{N}(\alpha_i' - \alpha)w^T x_n - \cancel{\sum_{i=1}^{N}\alpha_i\xi_i} - \cancel{\sum_{i=1}^{N}\alpha_i'\xi_i'}$$

$$= \sum_{i=1}^{N} y_i(\alpha_i' - \alpha_i) - \epsilon\sum(\alpha_i' + \alpha_i) + \frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N}(\alpha_i' - \alpha_i)(\alpha_j' - \alpha_j)\vec{x_i}^T\vec{x_j}$$

$$- \sum_{i=1}^{N}\sum_{j=1}^{N}(\alpha_i' - \alpha_i)(\alpha_j' - \alpha_j)\cdot\vec{x_i}\cdot\vec{x_j}$$

$$= \sum_{i=1}^{N} y_i(\alpha_i' - \alpha_i) - \epsilon\sum(\alpha_i' + \alpha_i) - \frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N}(\alpha_i' - \alpha_i)(\alpha_j' - \alpha_j)\vec{x_i}^T\vec{x_j}$$

$$\text{s.t.} \quad \sum_{i=1}^{N}(\alpha_i' - \alpha_i) = 0$$

$$0 \le \alpha_i, \alpha_i' \le C, \quad i = 1, 2 \cdots N.$$

P3.

1) primal: $\min_{\vec{w},b} \frac{1}{2}\|\vec{w}\|^2$

s.t. $y_n[\vec{w}^T\phi(\vec{x_n})+b] \geq 1, \forall n.$

To be specific. $\vec{w} = (w_1, w_2)$

$$\min_{\vec{w},b} \frac{1}{2}(w_1^2+w_2^2)$$

plug points into,
$y_n(w_1\cos x_1 + w_2 \sin x_1 + b) \geq 1$ $\Rightarrow$
we get constraints

s.t. $w_1 + b \geq 1$
$-(w_2+b) \geq 1$
$-w_1 + b \geq 1$

$\Rightarrow \begin{cases} b \geq 1 \\ w_1 \geq 1 \\ w_2 + b \leq -1 \end{cases}$

dual $\max_{\alpha} \sum_n \alpha_n - \frac{1}{2}\sum_m\sum_n y_m y_n \alpha_m \alpha_n k(x_m, x_n)$

s.t. $\alpha_n \geq 0, \forall n$

$\sum_n \alpha_n y_n = 0$

To be specific,

$k(x,x)$	x_1	x_2	x_3
x_1	1	0	-1
x_2	0	1	0
x_3	-1	0	1

$y_m y_n$	y_1 (1)	y_2 (-1)	y_3 (1)
y_1	1	-1	1
y_2	-1	1	-1
y_3	1	-1	1

$$\Rightarrow \max_{\alpha_1,\alpha_2,\alpha_3 \geq 0} \alpha_1+\alpha_2+\alpha_3 - \frac{1}{2}(\alpha_1^2+\alpha_2^2+\alpha_3^2 - 2\alpha_1\alpha_3)$$

s.t. $\alpha_1 - \alpha_2 + \alpha_3 = 0$

2) using constraints $\alpha_2 = \alpha_1 + \alpha_3$ to elimate α_2

objective become

$$\max_{\alpha_1, \alpha_3 \geq 0} 2\alpha_1 + 2\alpha_3 - \frac{1}{2}(\alpha_1^2 + \alpha_3^2 + \alpha_1^2 + 2\alpha_1\alpha_3 + \alpha_3^2 - 2\alpha_1\alpha_3)$$

$$= \max_{\alpha_1, \alpha_3 \geq 0} 2\alpha_1 + 2\alpha_3 - \alpha_1^2 - \alpha_3^2$$

$$= \max_{\alpha_1, \alpha_3 \geq 0} -(\alpha_1 - 1)^2 - (\alpha_3 - 1)^2 + 2$$

so when $\alpha_1 = 1$ and $\alpha_3 = 1$, obj is maximum.

and thus $\alpha_2 = 2$

The primal solution

$$(w_1^*, w_2^*)^T = \sum_{n=1}^{3} y_n \alpha_n^* \phi(x_n)$$

$$= 1 \times 1 \times \begin{pmatrix} 1 \\ 0 \end{pmatrix} + (-1) \times 2 \times \begin{pmatrix} 0 \\ 1 \end{pmatrix} + 1 \times 1 \times \begin{pmatrix} -1 \\ 0 \end{pmatrix}$$

$$= \begin{pmatrix} 0 \\ -2 \end{pmatrix}$$

$$b^* = y_1 - w^{*T} \phi(x_1)$$

$$= 1 - \begin{pmatrix} 0 \\ -2 \end{pmatrix}^T \begin{pmatrix} 1 \\ 0 \end{pmatrix} = 1$$

P4

1) $L = \varepsilon_t (e^{\beta_t} - e^{-\beta_t}) + e^{-\beta_t}$

$$\frac{\partial L}{\partial \beta_t} = \varepsilon_t (e^{\beta_t} + e^{-\beta_t}) - e^{-\beta_t} = 0$$

$$\Rightarrow \quad \varepsilon_t e^{2\beta_t} + \varepsilon_t - 1 = 0$$

$$\Rightarrow \quad \beta_t = \frac{1}{2} \ln\left(\frac{1-\varepsilon_t}{\varepsilon_t}\right)$$

2) ignore denomtnator of $D_{t+1}(n)$

$$\sum_{n: h_t(\vec{x}_n) \neq y_n} D_{t+1}(n) \propto \sum_{n: h_t(\vec{x}_n) \neq y_n} D_t(n) e^{\beta_t} = \varepsilon_t e^{\beta_t} = \sqrt{\varepsilon_t (1-\varepsilon_t)}$$

$$\sum_{n: h_t(\vec{x}_n) = y_n} D_{t+1}(n) \propto \sum_{n: h_t(\vec{x}_n) = y_n} D_t(n) e^{-\beta_t} = \varepsilon_t e^{-\beta_t} = \sqrt{(1-\varepsilon_t)\varepsilon_t}$$

since $\displaystyle\sum_{n: h_t(\vec{x}_n) = y_n} D_{t+1}(n) + \sum_{n: h_t(\vec{x}_n) \neq y_n} D_{t+1}(n) = 1$

$$\sum_{n: h_t(\vec{x}_n) = y_n} + \sum_{n: h_t(\vec{x}_n) \neq y_n} = \frac{1}{2}$$