HW4 ECS M146
Gawun Kim : 305-186-572

① $y_k(x) = w_k^T x + w_{k0}$, $k = 1, 2, \cdots K$

ⓐ $y(x) = \tilde{W}^T \tilde{x}$

$J(\tilde{W}) = \frac{1}{2} Tr\{(\tilde{X}\tilde{W} - T)^T(\tilde{X}\tilde{W} - T)\}$

$= \frac{1}{2} Tr\{ \tilde{W}^T\tilde{X}^T\tilde{X}\tilde{W} - \tilde{W}^T\tilde{X}^T T - T^T\tilde{X}\tilde{W} + T^T T \}$

$= \frac{1}{2}\left(Tr(\tilde{W}^T\tilde{X}^T\tilde{X}\tilde{W}) - Tr(\tilde{W}^T\tilde{X}^T T) - Tr(T^T\tilde{X}\tilde{W}) + \cancel{Tr(T^T T)} \right)$

$\frac{dJ}{d\tilde{W}} = \frac{1}{2}\left(\left((\tilde{X}^T\tilde{X})^T + (\tilde{X}^T\tilde{X})\right)\tilde{W} \quad -2(T^T\tilde{X})^T \right)$

$= \frac{1}{2}\left(2(\tilde{X}^T\tilde{X})\tilde{W} \quad - 2\tilde{X}^T T \right)$

$= (\tilde{X}^T\tilde{X})\tilde{W} - (\tilde{X}^T T)$

Minimize $\rightarrow \frac{dJ}{d\tilde{W}} = 0$

$(\cancel{\tilde{X}^T}\tilde{X})\tilde{W} - \cancel{\tilde{X}^T} T = 0$ $\qquad \tilde{W} = \tilde{X}^T$

$\tilde{W} = \tilde{X}^{-1} T$

ⓑ From 1-a, $\frac{\partial J}{\partial \tilde{W}}$ was already solved.

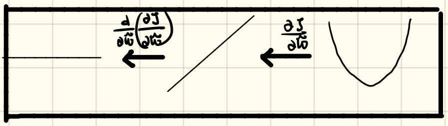
$\frac{\partial J}{\partial \tilde{W}} = (\tilde{X}^T\tilde{X})\tilde{W} - (\tilde{X}^T T)$

$\frac{\partial}{\partial \tilde{W}}\left(\frac{\partial J}{\partial \tilde{W}}\right) = (\tilde{X}^T\tilde{X}) \rightarrow \tilde{X}^2$

so it has unique value.

and it means that $J(\tilde{W})$ has a unique minimum value.

as some specific example of it visually



unique $\qquad$ has Min value

②

$K(x_1, x_2)^2 \leq K(x_1, x_1) K(x_2, x_2)$

$* \; K(x_1, x_1) = \Phi(x_1)^T\Phi(x_1) = (\Phi(x_1))^2$ ①

$K(x_2, x_2) = \Phi(x_2)^T\Phi(x_2) = (\Phi(x_2))^2$ ②

$K(x_1, x_2)^2 = (\Phi(x_1)^T\Phi(x_2))^2$

$\leq (\Phi(x_1))^2 \cdot (\Phi(x_2))^2$

By Cauchy-Schwartz, ① and ②

$\leq K(x_1, x_1) \cdot K(x_2, x_2)$

$\therefore K(x_1, x_2)^2 \leq K(x_1, x_1) \cdot K(x_2, x_2)$

③ $K_1(x, x')$ & $K_2(x, x')$

ⓐ By Gram matrix.

A : Gram Matrix for $K_1(x, x')$

B : Gram Matrix for $K_2(x, x')$

$\forall \; \alpha^T(A+B)\alpha = (\alpha^T A + \alpha^T B)\alpha$

$= \alpha^T A\alpha + \alpha^T B\alpha \geq 0$

$* \; \alpha^T A\alpha \geq 0$

$\alpha^T B\alpha \geq 0$ $\leftarrow$ positive semi~

$\rightarrow$ valid Kernels

ⓑ $K(x, x') = K_1(x, x') K_2(x, x') \rightarrow K = K_1 \cdot K_2$

$K_1 \rightarrow \sum_{i=1}^{n} \lambda_i u_i u_i^T$

$K_2 \rightarrow \sum_{j=1}^{n} \mu_j v_j v_j^T$ $\qquad$ Let they have positive eigenvalue. $\lambda_i \geq 0 \quad \mu_i \geq 0$

$K = \sum_{i=1}^{n}\sum_{j=1}^{n} \lambda_i \mu_j (u_i u_i^T) \circ (v_j v_j^T)$

$= \sum_{i=1}^{n}\sum_{j=1}^{n} \lambda_i \mu_j (u_i \cdot v_j)(u_i \cdot v_j)^T$

$= \sum_{k=1}^{n^2} \gamma_k w_k w_k^T$

$\forall \alpha \in R^n, \; \alpha^T K \alpha = \sum_{k=1}^{n^2} \gamma_k a^T w_k w_k^T a = \sum_{k=1}^{n^2} \gamma_k (w_k^T a)^2 \geq 0$

$\rightarrow$ Valid Kernels

ⓒ $K(x,x') = \exp(K_1(x,x'))$

Let $K_1(x,x') \to K_1$

$$= \exp(K_1) = 1 + K_1 + \frac{1}{2}K_1^2 + \frac{1}{6}K_1^3 + \cdots$$

$$i = 1,\ 2,\ 3,\ 4$$

when $i=1$: pos

$i=2$: pos from question (b)

$i=3$:

$$\to \sum_{t=1}^{n}\sum_{i=1}^{n}\sum_{j=1}^{n} r_t \lambda_i \mu_j (g_t g_t^T)(u_i u_i^T)\cdot(v_j v_j^T)$$

$$= \sum_{t=1}^{n}\sum_{i=1}^{n}\sum_{j=1}^{n} r_t \lambda_i \mu_j (g_t \cdot u_i \cdot v_j)(g_t \cdot u_i \cdot v_j)^T$$

$$= \sum_{K=1}^{n^3} \gamma_K w_K w_K^T$$

$$\forall \alpha \in R^n,\ \alpha^T K \alpha = \sum_{K=1}^{n^3} \gamma_K a^T w_K w_K^T a = \sum_{K=1}^{n^3} \gamma_K (w_K^T a)^2 \geq 0$$

(an infinite series)

Like this way, each of them are not negative. and summation of them are non-negative as well

$\to$ Valid Kernel

④

$$W_i = \sum_{K=1}^{m} a_K y_K x_K$$

$$W^T x_i + b = \sum_{K=1}^{m} a_K y_K \langle x_K, x_i \rangle + b$$

$$y_i(w^T x^i + b) - 1 + \varepsilon_i \geq 0$$

$$y_i\left(\sum_{K=1}^{M} a_K y_K \langle x_K, x_i \rangle + b\right) - 1 + \varepsilon_i \geq 0$$

$$L = \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{m}\varepsilon_i - \sum_{i=1}^{m}\alpha_i\left(y_i\left(\sum_{K=1}^{M} a_K y_K \langle x_K, x_i\rangle + b\right) - 1 + \varepsilon_i\right) - \sum_{j=1}^{n} \zeta_i \varepsilon_i$$

$$\frac{\partial L}{\partial b} \Rightarrow \sum_{i=1}^{n} \alpha_i y_i = 0$$

$$\frac{\partial L}{\partial \varepsilon_i} \Rightarrow C - \alpha_i - \zeta_i = 0$$

From given condition $0 \leq \alpha_i \leq C$

so that $\zeta_i \geq 0$ and $\zeta_i \varepsilon_i = 0$

$\varepsilon_i = 0$

so that

$$y_i\left(\sum_{i=1}^{M} a_K y_K \langle x_K, x_i \rangle + b\right) - 1 = 0$$

$$y_i\left(\sum_{i=1}^{M} a_K y_K \langle x_K, x_i \rangle + b\right) = 1$$

y_i can be divided in each side since it is -1 or 1 and $\frac{1}{y_i}$ is also -1 or 1 // $y_i = \frac{1}{y_i}$

$$\left(\sum_{K=1}^{M} a_K y_K \langle x_K, x_i \rangle + b\right) = y_i$$

$$\sum_{K=1}^{M} a_K y_K \langle x_K, x_i \rangle + mb = y_i$$

$$mb = y_i - \sum_{K=1}^{M} a_K y_K \langle x_K, x_i \rangle$$

$$mb = \frac{1}{m}\sum_{i=1}^{m}\left(y_i - \sum_{K=1}^{m} a_K y_K \langle x_K, x_i \rangle\right)$$

for gaining the mean of the set of mb

$$b_i = \frac{1}{m^2}\sum_{i=1}^{m}\left(y_i - \sum_{K=1}^{M} a_K y_K \langle x_K, x_i \rangle\right)$$

$$b = \frac{1}{N_M}\sum_{n\in M}\left(y^{(n)} - \sum_{m\in S} a_m y^m \langle x^n, x^m \rangle\right)$$

⑤

ⓐ Classify 5 points and not consider for one point. It will be 3.

$-1, 0, 1$ as negative.

$-3, -2$ as positive

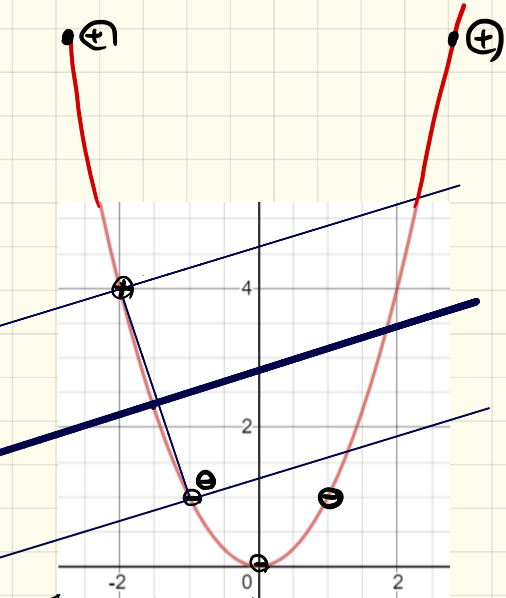
So that The boundary with max margin will be the middle point between -2 & -1

so $\frac{-2-1}{2} = -\frac{3}{2}$

$\mathrm{sgn}(-2x-3)$

(b) $\phi(x) = (x, x^2)$

(c) (d)



The point are linearly separable.

$K_1 = (-2,4)$
$K_2 = (-1,1)$ > support vectors were on these two dots.

and between them, there is a vector to separate two class.

The middle between K_1 and K_2

The linear for K_1 and K_2 is

$$y = -3(x+1)+1$$

and its perpendicular line is

$$y = +\frac{1}{3}\left(x+\frac{3}{2}\right)+\frac{5}{2}$$

$$6y = 2x+3+15$$

$$0 = 2x-6y+18 \Rightarrow \boxed{x-3y+9=0}$$

$x-3y+9=0$

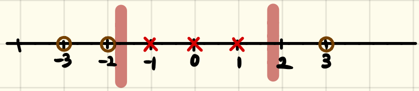
target 1 : $(-3, 9)$
target 2 : $(0, 0)$

target 1 : distance $= \dfrac{|1\cdot(-3)+(-3)\cdot 9+9|}{\sqrt{1+9}} = \dfrac{21}{\sqrt{10}}$

target 2 : distance $= \dfrac{|9|}{\sqrt{10}} = \dfrac{9}{\sqrt{10}}$

∴ When point $(-3, 9)$, it is max which is $\dfrac{21}{\sqrt{10}}$

(e)



(f) there are only two sup vector $K_1(-2,4)$ $K_2(-1,1)$ so a_1, a_2 are non zero.

$$L(a) = \sum_{n=1}^{N} a_n - \frac{1}{2}\sum_{n=1}^{N}\sum_{m=1}^{N} a_n a_m t_n t_m K(x_n, x_m)$$

$$= a_1 + a_2 - \frac{1}{2}\left(a_1^2(K_1\cdot K_1) - 2a_1a_2(K_1\cdot K_2) + a_2^2(K_2\cdot K_2)\right)$$

and since the constraint on the quadratic,

$$a_1 = a_2 = a$$

$$= 2a - 5a^2$$

$$\frac{dL}{da} = 2 - 10a = 0$$

$$a = \frac{1}{5}$$

$$1 = \sum_{n=1}^{S} a_n y_n K(x, K_n) + b$$

$$b = -\frac{9}{5}$$

$$\alpha_i = \frac{1}{5}$$

$$w = \sum_{n\in S} a_n y_n x_n = a_1 y_1 x_1 + a_2 y_2 x_2 = \frac{1}{5}(x_1 - x_2) = \frac{1}{5}(-1\ 3)$$

W will be the slope from (d)

and b is the constant of the line from the equation (d)