① for KKT conditions:

$$\lambda_i (y_i - (w\varphi(x) + w_0) + \xi_i) = 0$$

s.t.

$$\xi_i \geq 0$$
$$\lambda_i \geq 0$$
$$y_i - (w\varphi(x) + w_0) \leq \varepsilon$$

for $\xi_i = 0 \Rightarrow \lambda_i (y_i - (w\varphi(x) - w_0)) = 0$

if $\lambda_i > 0$, then $y_i - (w\varphi(x) + w_0) = \varepsilon$
(suppore vectors)

$$w\varphi(x) + w_0 = y_i - \varepsilon$$
$$w_0 = y_i - \varepsilon - w\varphi(x)$$

for suppore vector at the upper edge of ε - table, which is presented by x_j

$$w_0 = y_j - \varepsilon - w\varphi(x_j)$$

~~since~~ $\cancel{w\varphi(x_j) = \sum \lambda_i y_i k(x_i, x_j)}$

$\cancel{\therefore w_0 = y_j - \varepsilon - \sum \lambda_i y_i k(x_i, x_j)}$

SVR: $L(w, w_0, \xi, \lambda^+, \lambda^-) = 0.5\|w\|^2 + c\sum(\xi_i + \xi_i^*) - \sum \lambda_i (y_i - w\varphi(x) + w_0) - \xi_i) - \sum \lambda_i^* (y - w\varphi(x) + w_0) + \xi_i^*)$

$$\frac{\partial L}{\partial w} = 0 \Rightarrow w + \sum \lambda_i^* \varphi(x) - \sum \lambda_i (\varphi(x)) = 0$$

$$w = \sum (\lambda_i - \lambda_i^*) \varphi(x)$$

$$\frac{\partial L}{\partial w_0} = 0 \qquad \sum(\lambda_i - \lambda_i^*) = 0$$

substitute w into the equation with w_0

$$w_0 = y_j - \varepsilon - \Sigma(\lambda_i - \lambda_i^*)\, \varphi(x_i) \cdot \varphi(x_j)$$

$$K(x_i, x_j) = \varphi(x_i)\varphi(x_j)$$

$$\therefore w_0 = y_j - \varepsilon - \Sigma(\lambda_i - \lambda_i^*) K(x_i, x_j)$$

$$= y_j - \varepsilon - \Sigma(\lambda^+ - \lambda^-) K(x_i, x_j)$$

② a. learning rate $a = \eta(i) = 0.5$

2 epochs only

$$\begin{cases} X_1 = C_1 = [1, -1, -1] \\ X_2 = C_2 = [-1, 1, -1] \\ X_3 = C_3 = [-1, -1, 1] \end{cases}$$

Start with $\begin{bmatrix} 0 & 2 \\ 3 & 0 \\ 1 & -1 \end{bmatrix}$

for X_1 target $[1, -1, -1]$

$e_1 = y_1 - \text{sign}(W(0)^T x(1)) = [1\ -1\ -1] - \text{sign}([0\ 3\ 1], [2\ 0\ -1]) * [1\ 1]$

$= [0\ -2\ -2]$

$W(1) = W(0) + \eta(1) * x(1) * e^T = \begin{bmatrix} 0 & 1 \\ 3 & -1 \\ 1 & -2 \end{bmatrix}$

$e_2 = y_2 - \text{sign}(W(1)^T x(2)) = [0\ 2\ 0]$

$$w(2) = w(1) + \eta(2) * x_2 \cdot e_2^T = \begin{bmatrix} 0 & 0 \\ 3 & -2 \\ 1 & -2 \end{bmatrix}$$

$e_3 = [0 \ 0 \ 2]$

$$W(3) = W(2) + \eta \, X(3) e_3^T = \begin{bmatrix} 0 & 0 \\ 3 & -2 \\ 1 & -3 \end{bmatrix}$$

not converges yet, and it is keeping changing

(b) $$W(0) = \begin{bmatrix} 0 & 2 \\ 3 & 0 \\ 1 & -1 \end{bmatrix}$$

$e_1 = y_1 - \text{sign}(W(0)^T \cdot X(1)) = [0 \ -2 \ -2]$

$w_1(1) = w_1(0) + \eta e_1 X(1) = 0.5 \cdot 0 \cdot 1 + 0.5 = 0$

$w_2(1) = 1$

$w_3(1) = -2$

$\therefore w(1) = [0 \ 1 \ -2]$

$e_2 = y_2 - \text{sign}(w(1)^T \cdot X(2)) = [0 \ 0 \ -2]$

$w_1(2) = 0$

$w_2(2) = 1$

$w_3(2) = -1$

$\therefore w(2) = [0 \ 1 \ -1]$

$e_3 = y_3 - \text{sign}(w(2)^T X_3) = [0 \ -1 \ 1]$

$w_1(3) = 0$

$w_2(3) = 1.5$

$w_3(3) = 1.5$

$w(3) = [0 \quad 1.5 \quad -1.5]$

Not converge either

③

x_1	x_3
x_2	x_4

class1 : x_1 & x_2 or x_3 & x_4

class2 : x_1 & x_3 or x_2 & x_4

$D = \text{sigmoid}(\text{sigmoid}(x_1 * x_2) + \text{sigmoid}(x_3 * x_4))$

If $D > 0.5$ class 1

$D < 0.5$ class 2