SVM

$C: w^T x + b = 0$ maximum margin

$$\text{Minimum } \frac{1}{2} w^T w$$

$$\text{Subject to, } y_n w^T x_n + y_n b \geq 1$$

Quadratic Program

Dual Optimization Problem (Convex Optimization)

minimize $f(x)$

Lagrangian Function

$$L(w, b, \alpha) = \frac{1}{2} w^T w - \sum_{n=1}^{N} \alpha_n (y_n w^T x_n + y_n b - 1)$$

$\alpha_n \equiv$ dual variables, $\alpha_n \geq 0$

$\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_n)$

Using Lagrangian, We can recover original problem

If

1) maximize $L(\ )$ wrt α
2) minimize $L(\ )$ wrt w, b

maximize

$\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_n)$

$\alpha_i \geq 0$

minimize

w, b

$L(w, b, \alpha)$

$$L(w,b,\alpha) = \frac{1}{2} w^T w - \sum_{n=1}^{N} \alpha_n (y_n w^T x_n + y_n b - 1)$$

Suppose for some (w, b)

$y_n w^T x_n + y_n b < 1$ violate a constraint

We argue that (w, b) cannot be a solution to inner minimization, we can set $\alpha_n \longrightarrow \infty$, but this is a contradiction as $L^* < \infty$

Suppose $y_n w^T x_n + y_n b \geq 1$,

$$\underbrace{-\alpha_n (y_n w^T x_n + y_n b - 1)}_{\geq 0}$$

$\longrightarrow \alpha_n = 0$

$$\boxed{\begin{array}{l} \text{Thus } \forall n = 1, 2, \ldots, N \\ \alpha_n (y_n w^T x_n + y_n b - 1) = 0 \\ \downarrow \\ \text{KKT Condition} \end{array}}$$

$y_n w^T x_n + y_n b > 1 \longrightarrow \alpha_n = 0$

Fix $\underline{\alpha} = (\alpha_1, \alpha_2, \cdots, \alpha_n)$ & solve

$$\min_{\underline{w},\, b} \mathcal{L}(\underline{w}, b, \underline{\alpha})$$

$$\mathcal{L}(\underline{w}, b, \underline{\alpha}) = \frac{1}{2}\underline{w^T w} - \sum_{n=1}^{N} \alpha_n \underline{(y_n w^T x_n + y_n b - 1)}$$

$$\nabla_{\underline{w}} \mathcal{L} = 0 \quad \Big| \quad \nabla_{\underline{w}} \frac{1}{2}\underline{w^T w} - \sum_{n=1}^{N} \alpha_n \underline{(y_n w^T x_n + y_n b - 1)}$$

$$\frac{\partial \mathcal{L}}{\partial b} = 0 \quad \Big| \quad = \underline{w} - \sum_{n=1}^{N} \alpha_n y_n \underline{x_n} = 0$$

$$\underline{w}^* = \sum_{n=1}^{N} \alpha_n y_n \underline{x_n}$$

$$\mathcal{L}(\underline{w}, b, \underline{\alpha}) = \frac{1}{2}\underline{w^T w} - \sum_{n=1}^{N} \alpha_n \underline{(y_n w^T x_n + y_n b - 1)}$$

$$= \frac{1}{2}\underline{w^T w} - \sum_{n=1}^{N} \alpha_n y_n w^T x_n - \sum_{n=1}^{N} \alpha_n y_n b + \sum_{n=1}^{N} \alpha_n$$

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

$$\sum_{n=1}^{N} \alpha_n y_n = 0$$

$$\mathcal{L}(\cdot) = \frac{1}{2}\underline{w^T w} - w^T \sum_{n=1}^{N} \alpha_n y_n x_n + \sum_{n=1}^{N} \alpha_n$$

$$= -\frac{1}{2} w^T w + \sum_{n=1}^{N} \alpha_n$$

$$= -\frac{1}{2} \left(\sum_{n=1}^{N} \alpha_n y_n x_n \right)^T \left(\sum_{n=1}^{N} \alpha_n y_n x_n \right)$$

$$+ \sum_{n=1}^{N} \alpha_n$$

$$= -\frac{1}{2} \sum_{n=1}^{N} \sum_{m=1}^{N} \alpha_n \alpha_m y_n y_m x_n^T x_m + \sum_{n=1}^{N} \alpha_n$$

Dual SVM

Maximize $L(\alpha)$

Subject To $\sum_{n=1}^{N} \alpha_n y_n = 0$

$$L(\alpha) = -\frac{1}{2} \sum_{n=1}^{N} \sum_{m=1}^{N} \alpha_n \alpha_m y_n y_m x_n^T x_m + \sum_{n=1}^{N} \alpha_n$$

$\alpha_i \geq 0$

Kernel Trick

Dual SVM is a Quadratic Program

$$L(\alpha) = \frac{1}{2} \alpha^T M \alpha + p^T \alpha$$

$$\alpha = \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{bmatrix} \quad M = \begin{bmatrix} y_1 y_1 x_1^T x_1 & \cdots & y_1 y_n x_1^T x_n \\ \vdots & & \vdots \\ y_n y_1 x_n^T x_1 & \cdots & y_n y_n x_n^T x_n \end{bmatrix} \quad p = \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}$$