$$\min_{x \in \mathbb{R}^d} \quad f(x)$$

f, f_j are convex functions

(P)

$$f_i(x) \leq 0 \quad i = 1, \ldots, m$$

$$a_i^T x = b_i \quad j = 1, \ldots, n$$

$$f_j : \mathbb{R}^d \to \mathbb{R}$$

$$\mathcal{L}(x, \lambda, \mu) = f(x) + \sum_{i=1}^{m} \lambda_i f_i(x) + \sum_{j=1}^{n} \mu_j (a_j^T x - b_j)$$

1. $\nabla f(x) + \sum_{i=1}^{m} \lambda_i \nabla f_i(x) + \sum_{j=1}^{n} \mu_j a_j = 0$

2. $\lambda_i f_i(x) = 0 \qquad \lambda_i \geq 0$

3. $f_i(x) \leq 0 \qquad a_j^T x = b_j$

   $i = 1, \ldots, m \qquad j = 1, \ldots, n$

<u>K.K.T point</u>

If for any x^* there exists λ^*, μ^* such that (x^*, λ^*, μ^*) satisfy (1)-(3) then x^* is a K.K.T. point of (P).

$\boxed{\text{If } x^* \text{ is a K.K.T. point then it is global minimum of (P)}}$

Wolfe Dual

$$\left.\begin{aligned} \max_{x,\lambda,\mu} \quad & \mathcal{L}(x, \lambda, \mu) \\ & \nabla_x \mathcal{L}(x, \lambda, \mu) = 0 \\ & x \geq 0 \end{aligned}\right\} \text{Wolfe Dual}$$

$\boxed{x^*, \lambda^*, \mu^* \text{ solves Wolfe Dual}}$

$$\boxed{\nabla_x \mathcal{L}(x^*, \lambda^*, \mu^*) = 0 \;, \lambda^* \geq 0}$$

Hence it is a feasible point of Wolfe Dual.

$$\boxed{\mathcal{L}(x^*, \lambda^*, \mu^*) = f(x^*)}$$

f, f_j are convex. | For any $x \in \mathbb{R}^d$

$$f(x^*) \geq f(x) + \nabla f(x)^T (x^* - x)$$

$$f_i(x^*) \geq f_i(x) + \nabla f_i(x)^T (x^* - x)$$

$$a_j^T x^* - b = a_j^T x - b + a_j^T (x - x^*)$$

$$f(x^*) = \mathcal{L}(x^*, \lambda^*, \mu^*)$$

$$\geq f(x^*) + \sum_{i=1}^{m} \lambda_i \boxed{f_i(x^*)} < 0$$

$$+ \sum_{j=1}^{n} \mu_j (a_j^T x^* - b_j)$$

For any $\lambda_i \geq 0$

$$\geq f(x) + \nabla f(x)^T (x^* - x)$$

$$+ \sum_{i=1}^{m} \lambda_i \left(f_i(x) + \nabla f_i(x)^T (x^* - x) \right.$$

$$+ \sum_{j=1}^{n} \mu_j (a_j^T x - b)$$

$$+ \sum_{j=1}^{n} \mu_j a_j^T (x^* - x)$$

$$f(x^*) = \mathcal{L}(x^*, \lambda^*, \mu^*)$$

$$\geq f(x) + \sum_{i=1}^{m} \lambda_i f_i(x) + \sum_{j=1}^{n} \mu_j (a_j^T x - b)$$

$$+ (x^* - x)^T \left[\nabla f(x) + \sum_{i=1}^{m} \lambda_i \nabla f_i(x) + \sum_{j=1}^{n} \mu_j a_j \right)$$

$$= \mathcal{L}(x, \lambda, \mu) + (x^* - x)^T \nabla_x \mathcal{L}(x, \lambda, \mu)$$

Thus for any feasible x, λ, μ.

$$f(x^*) = \mathcal{L}(x^*, \lambda^*, \mu^*)$$

$$\geq \mathcal{L}(x, \lambda, \mu)$$

Dual of SVM problem

$$\min_{\{w,b\}} \frac{1}{2}\|w\|^2$$

$$y_i(w^T x_i + b) \geq 1$$

$$\mathcal{L}(w, b, \lambda)$$

$$= \frac{1}{2}\|w\|^2 - \sum_{i=1}^{N} \lambda_i \{ y^{(i)} (w^T x^{(i)} + b) - 1 \}$$

$$\nabla_w \mathcal{L} = 0 \Rightarrow w - \sum_{i=1}^{N} \lambda_i y^{(i)} x^{(i)} = 0$$

$$\nabla_b \mathcal{L} = 0 \Rightarrow -\sum_{i=1}^{N} \lambda_i y^{(i)} = 0$$

$$\lambda_i \{ y_i (w^T x^{(i)} + b) - 1 \} = 0$$

$$\lambda_i \geq 0$$

Wolfe Dual

$$\max_{w, b, \lambda} \frac{1}{2}\|w\|^2 - \sum_{i=1}^{N} \lambda_i \left\{ y^{(i)} \left(w^T x^{(i)} + b \right) - 1 \right\}$$

$$w = \sum_{i=1}^{N} \lambda_i y^{(i)} x^{(i)}$$

$$\sum_{i=1}^{N} \lambda_i y^{(i)} = 0$$

Eliminate $\underline{\underline{w}}$, b

$$\max_{\lambda \geq 0} \sum_{i=1}^{N} \lambda_i - \left\| \sum_{i=1}^{N} \lambda_i y^{(i)} x^{(i)} \right\|^2$$

$$\sum_{i=1}^{N} \lambda_i y^{(i)} = 0$$

$$\max_{\lambda \geq 0} \sum_{i=1}^{N} \lambda_i - \sum_{i=1}^{N} \sum_{j=1}^{N} \lambda_i \lambda_j y^{(i)} x^{(i)T} x^{(j)}$$

$$\sum_{i=1}^{N} \lambda_i y^{(i)} = 0$$

Solve SVM. in feature space

$\Phi(x)$.