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

f, f_j are convex functions

(P) $f_i(x) \le 0 \quad i = 1, \dots, m$

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

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

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

(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) $f_i(x) \le 0 \qquad a_j^T x = b_j$

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

<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).

If x^* is a K.K.T. point then it is global minimum of (P)

Wolfe Dual

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

$$\nabla_x \mathcal{L}(x, \lambda, \mu) = 0$$

$$x \geq 0$$

Wolfe Dual

x^*, λ^*, μ^* solves P

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

Hence it is a feasible point of Wolfe Dual.

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

f, f_i 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 $\lambda_i \geq 0$

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

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

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

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

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

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

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

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

Thus for any feasible x, λ, μ.

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

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

Dual of SVM problem

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

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

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

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

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

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

$$\lambda_i \geq 0$$

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

$\Phi(x)$.