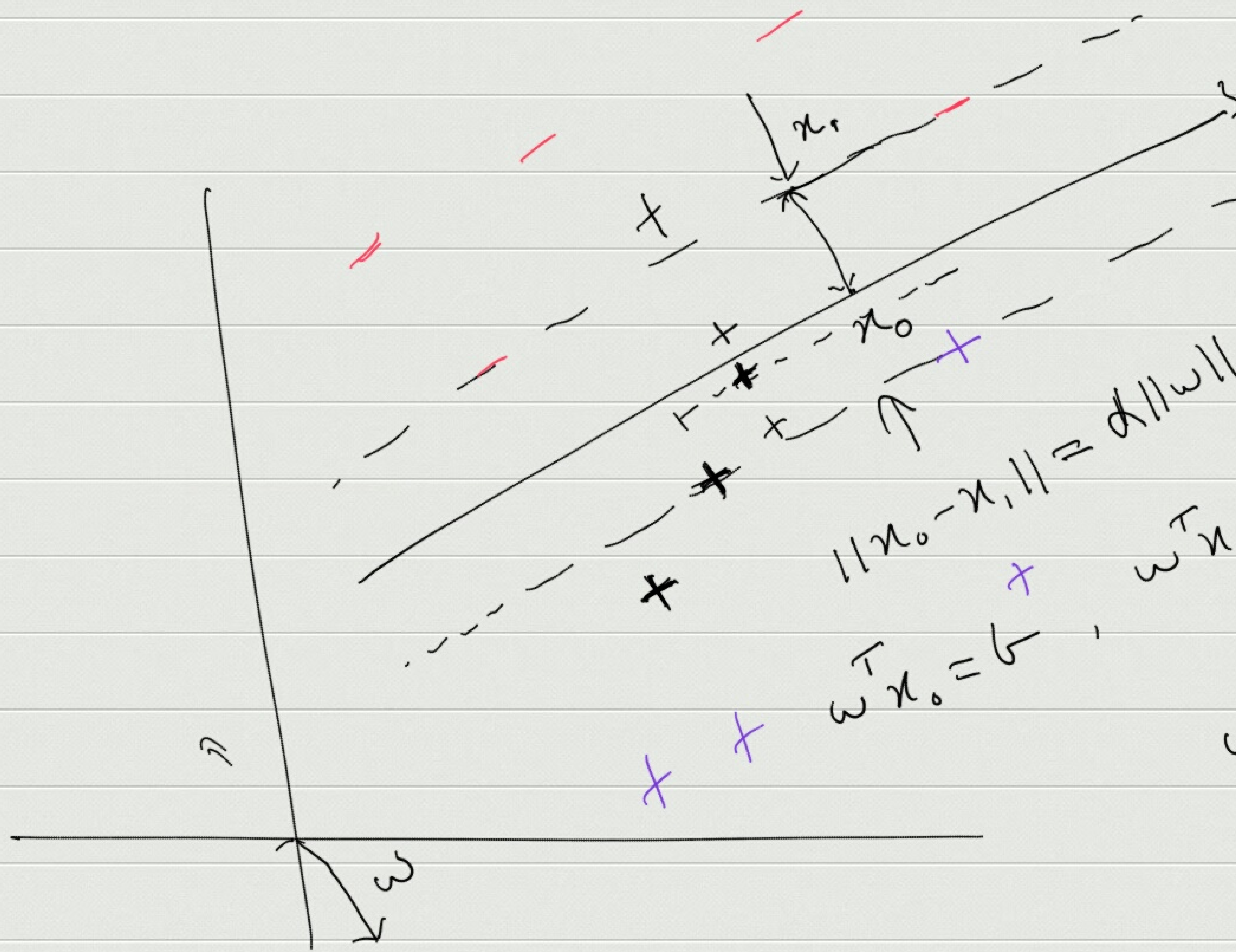


$\underline{\underline{\alpha > 0}}$

$x_0 - x_1 = \alpha w$

$\|x_0 - x_1\| = \alpha \|w\| = \rho / \|w\|$

$w^T x_0 = b \quad , \quad w^T x_1 = b - \rho$

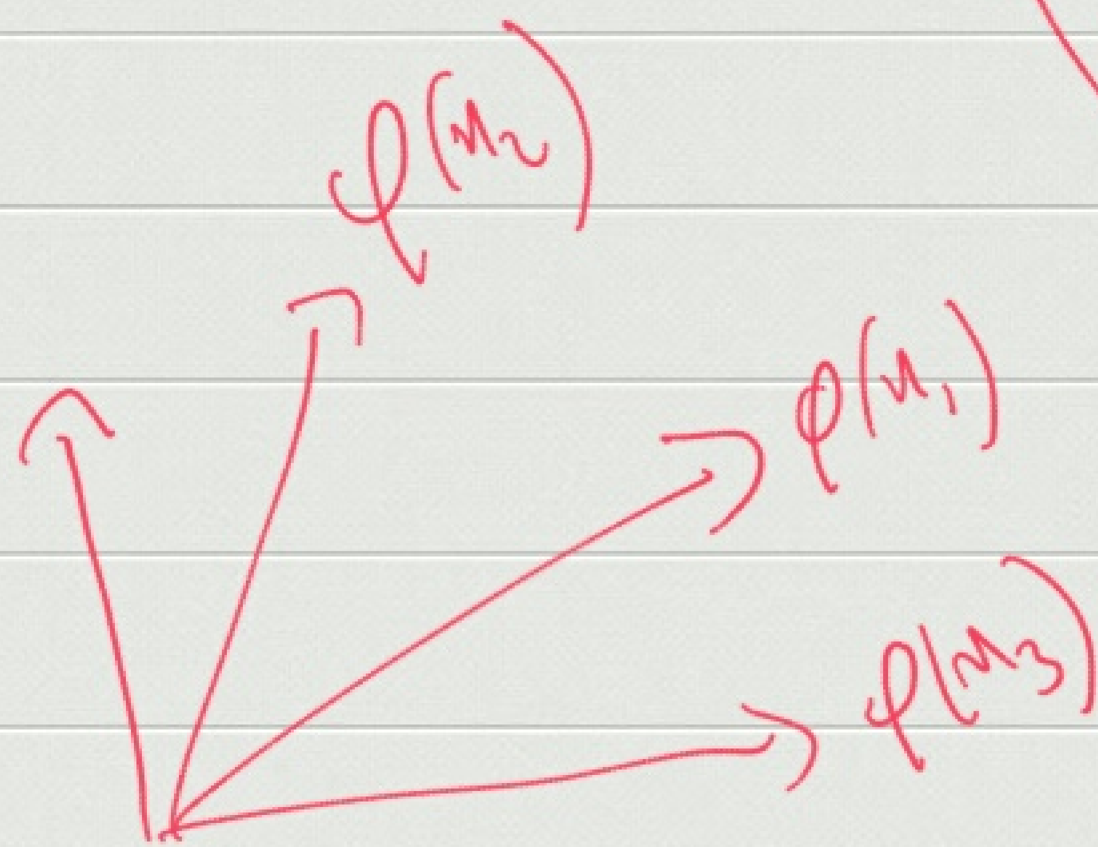
$w^T x_0 - w^T x_1 = \alpha \|w\|^2$

$b - (b - \rho) = \alpha \|w\|^2$

$$\min_{w,b} \frac{1}{2}\|w\|^2 + \frac{C}{m}\sum_i l\left(y_i\left(w^T\varphi(x_i) - b\right)\right)$$

$$w^T\varphi(x) = \sum_{i=1}^{m}\alpha_i\, y_i\, \varphi(x_i)^T\varphi(x) + w_\perp^T\varphi(x)$$

$$w_\perp^T\varphi(x_i) \;\forall i$$

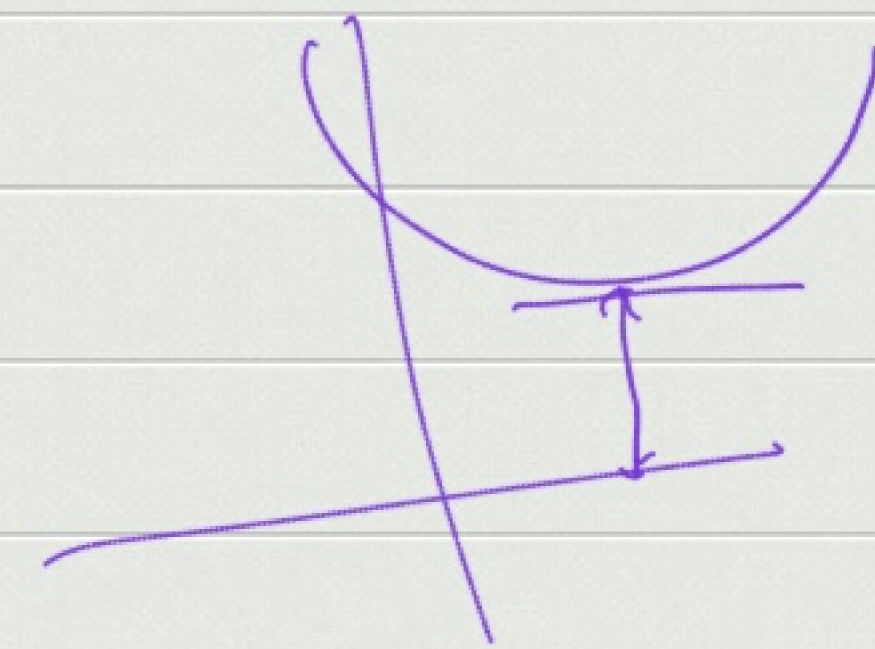


$$\min_{\alpha_i,\, w_\perp} \overbrace{\frac{1}{2}\Big\|\sum_i \alpha_i y_i \varphi(x_i)\Big\|^2 + \frac{1}{2}\|w_\perp\|^2}^{\frac{1}{2}\|w\|^2} + \frac{C}{m}\sum_i l\left(y_i\left((w-w_\perp)^T\varphi(x_i) - b\right)\right)$$

no $w_\perp$

$$w_\perp = 0$$

$$\min_{x \in \mathbb{R}} f(x) \qquad \frac{df(x^*)}{dx} = 0$$



$$\min_{x \in \mathbb{R}^n} f(x) \qquad \nabla f(x^*) = 0$$

$$\downarrow \begin{bmatrix} \frac{\partial f(x^*)}{\partial x_1} \\ \vdots \\ \frac{\partial f(x^*)}{\partial x_n} \end{bmatrix} = 0$$

$$\min_{w \in \mathbb{R}^n,\, b \in \mathbb{R}} \underbrace{\frac{1}{2}\|w\|^2 + \frac{c}{m}\sum_{i=1}^{m} l\left(y_i\left(w^T \phi(x_i) - b\right)\right)}_{f(w,b)} \qquad \frac{\partial f(w^*, b^*)}{\partial w_j^*} = 0$$

$$\downarrow$$

$$w_j^* + \frac{c}{m}\sum_{i=1}^{m} l'\left(y_i\left(w^{*T}\phi(x_i) - b^*\right)\right) \phi_j(x_i)\, y_i = 0 \quad \forall j$$

$$\phi(x_i)^T\phi(x_1) \;\dots\; \phi(x_i)^T\phi(x_n)$$

$$\min_{\substack{w\in\mathbb{R}^n,\, b\in\mathbb{R}\\ \gamma\in\mathbb{R}^m}} \frac{1}{2}\|w\|^2 + \frac{c}{m}\sum_{i=1}^{m} l_i(\gamma_i) + \max_{\alpha\in\mathbb{R}^m}\sum_i \alpha_i\left(\underbrace{\gamma_i - w^T\phi(x_i) + b}_{\neq 0}\right)$$

$$-\frac{1}{2}\sum_i\sum_j \alpha_i\alpha_j\phi(x_i)^T\phi(x_j) = -\frac{1}{2}\alpha^T \boxed{K\alpha}$$

$$\alpha = \begin{bmatrix}\alpha_1\\ \vdots\\ \alpha_m\end{bmatrix}$$

$$(K)_{ij} = \phi(x_i)^T\phi(x_j)$$

$$= \max_{\alpha\in\mathbb{R}^m}\left(\min_w \frac{1}{2}\|w\|^2 - \sum\alpha_i w^T\phi(x_i)\right) + \left(\min_b \sum_i \alpha_i b\right)$$

$$+ \min_{\gamma\in\mathbb{R}^m}\left(\sum_i\alpha_i\gamma_i + \frac{c}{m}\sum_i l_i(\gamma_i)\right)$$

$$w = \sum_i \alpha_i\phi(x_i)$$

$$\sum_i \alpha_i = 0$$

$$-\frac{c}{m}\max_{\gamma_i\in\mathbb{R}}\left(-\frac{m\alpha_i}{c}\right)\gamma_i - l_i(\gamma_i)$$

$$-\frac{c}{m} l_i'(\gamma_i) = \alpha_i \quad \forall i$$

$$= -\frac{c}{m} l_i^*\left(\frac{-m\alpha_i}{c}\right)$$

$$f^*(y) \equiv \max_{x \in dom(f)} x^T y - f(x)$$

f convex then $(f^*)^* = f$

f^* is called as $\longrightarrow$ Conjugate

$\longrightarrow$ Fenchel dual

$\longrightarrow$ Legendre transform.

$f : \mathbb{R}^n \rightarrow \mathbb{R}$

$$\boxed{\begin{aligned} &\max_{\alpha \in \mathbb{R}^m} -\frac{1}{2}\alpha^T K \alpha - \frac{c}{m}\sum_i l_i^*\left(\frac{-m\alpha_i}{c}\right) \\ &\underline{s.t.} \quad \sum_i \alpha_i = 0 \end{aligned}}$$

$$\min_{x \in \mathcal{X}} \quad f(x)$$

$$\underline{\underline{\text{s.t.}}} \qquad g_1(x) \le 0$$

$$\vdots$$

$$g_m(x) \le 0$$

$$= \min_{x \in \mathcal{X}} \boxed{\max_{\lambda \geq 0} \quad L(x, \lambda)} \longrightarrow \begin{cases} 0 & g_i(x) \le 0 \ \forall i \\ \infty & \text{otherwise} \end{cases} \equiv \bar{L}(x)$$

$$L(x, \lambda) = f(x) + \sum_i \lambda_i \, g_i(x)$$

$$\min_{x \in \mathcal{X}} \max_{\lambda \geq 0} L(x, \lambda) \qquad \max_{\lambda \geq 0} \min_{x \in \mathcal{X}} L(x, \lambda)$$

$$\min_{\underline{x} \in \mathcal{X}} \max_{\lambda \geq 0} L(\underline{x}, \lambda) \geq \max_{\lambda \geq 0} \min_{\underline{x} \in \mathcal{X}} L(\underline{x}, \underline{\lambda})$$

$\neq$

Theorem: If All f, g_i's are convex & mild conditions

I is bounded
is feasible → slater's cond.
is solvable

$\exists \, x \in \mathcal{X} \ni g_i(x) < 0 \quad \forall i$

$$\mathrm{I} = \min_{x \in \mathcal{X}} \underbrace{\boxed{\max_{\lambda \geq 0} L(x, \lambda)}}_{\overline{L}(x)} \geq \max_{\lambda \geq 0} \underbrace{\boxed{\min_{x \in \mathcal{X}} L(x, \lambda)}}_{\underline{L}(\lambda)} = \mathrm{II}$$

$$\mathrm{I} \to \overline{L}(x) \to \min_{x \in \mathcal{X}} \to x^*$$

$$\mathrm{II} \to \max_{\lambda \geq 0} \underline{L}(\lambda) \to \lambda^*$$

$$\overline{L}(x^*) \geq \min_{x \in \mathcal{X}} \max_{\lambda \geq 0} L(x, \lambda) \geq \max_{\lambda \geq 0} \min_{x \in \mathcal{X}} L(x, \lambda)$$

$$\geq \underline{L}(\lambda^*)$$

$$=$$

$$=$$

$$\min_{x \in \mathcal{X}} L(x, \lambda^*) = \overline{L}(x^*)$$