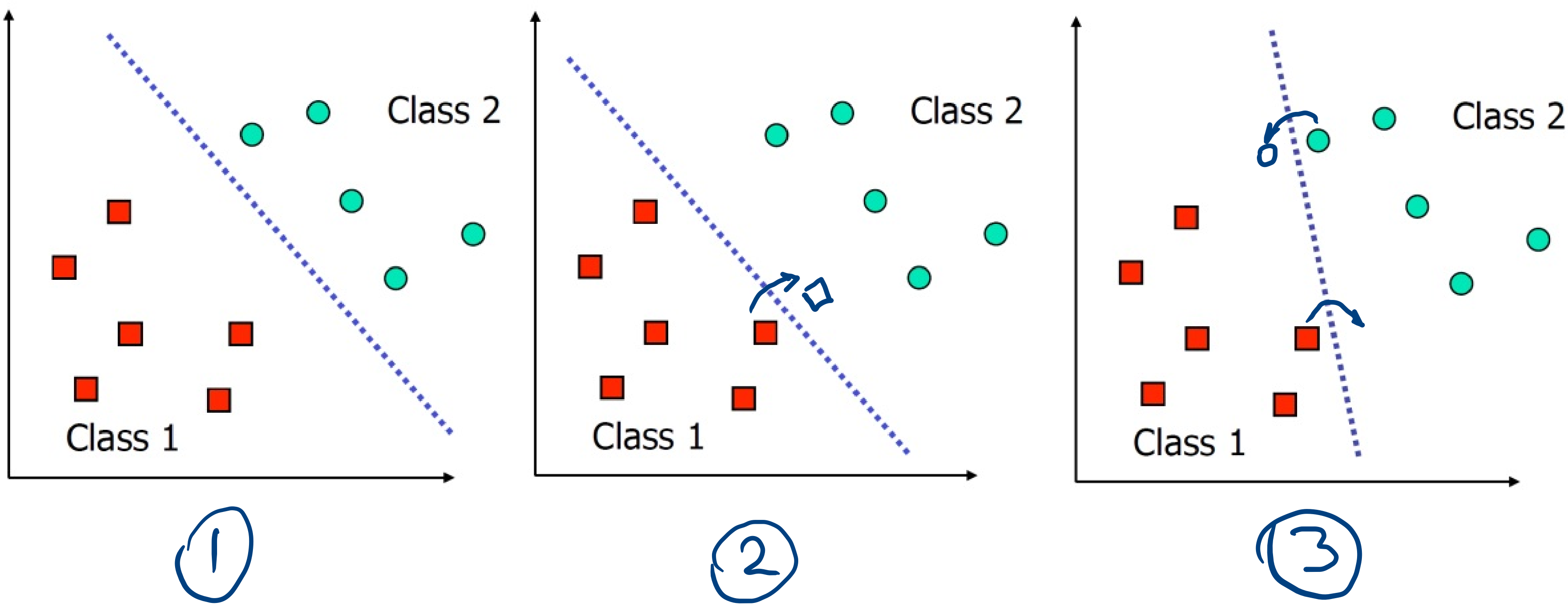
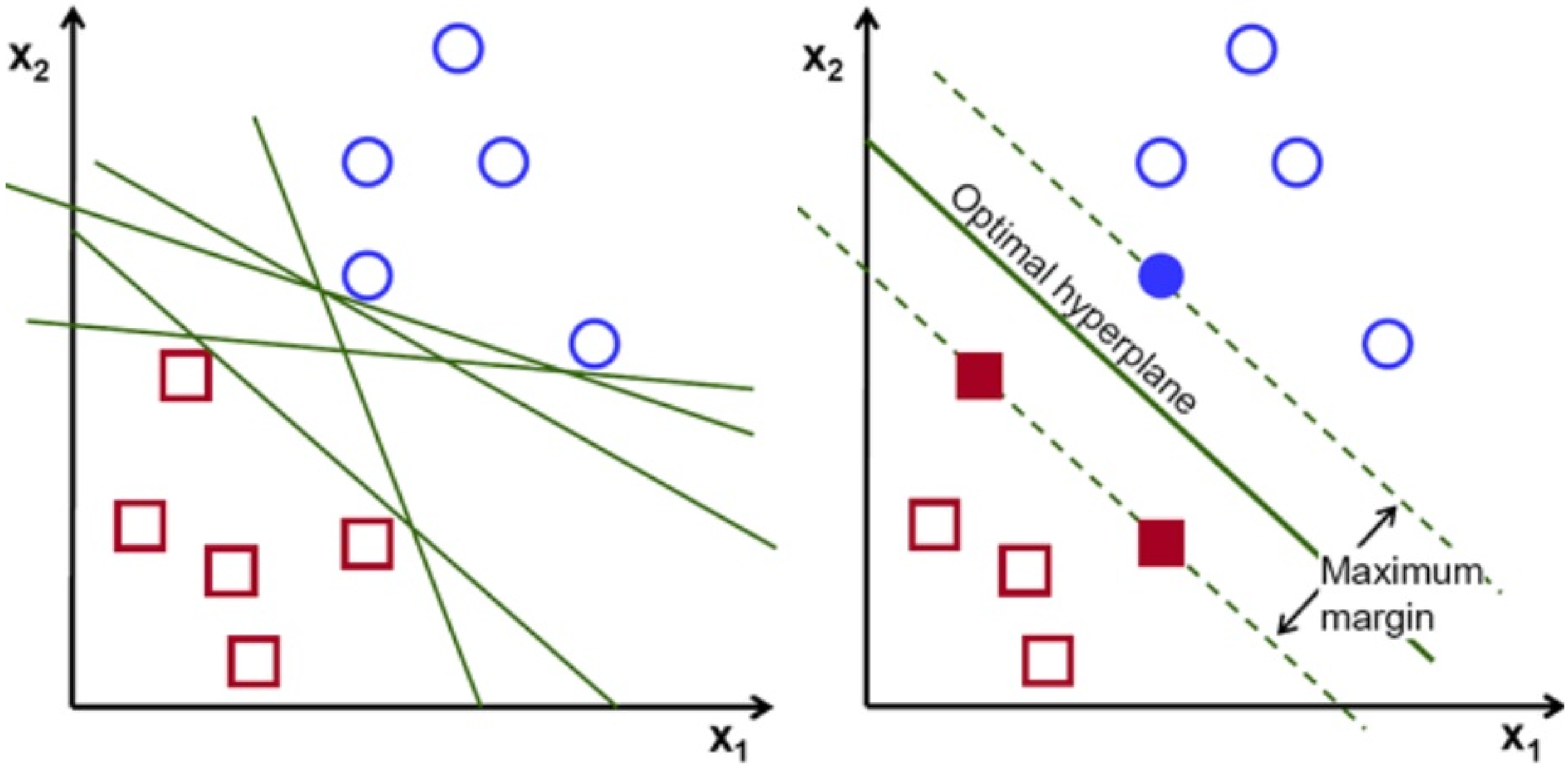
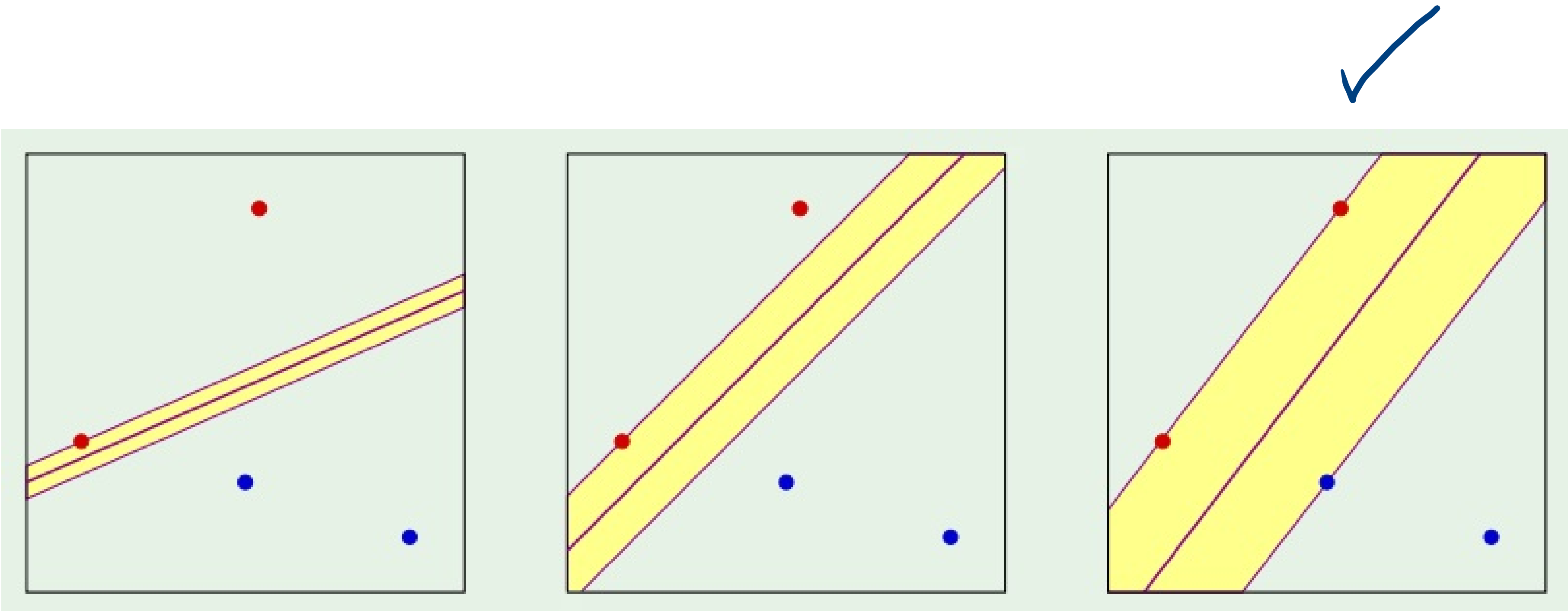
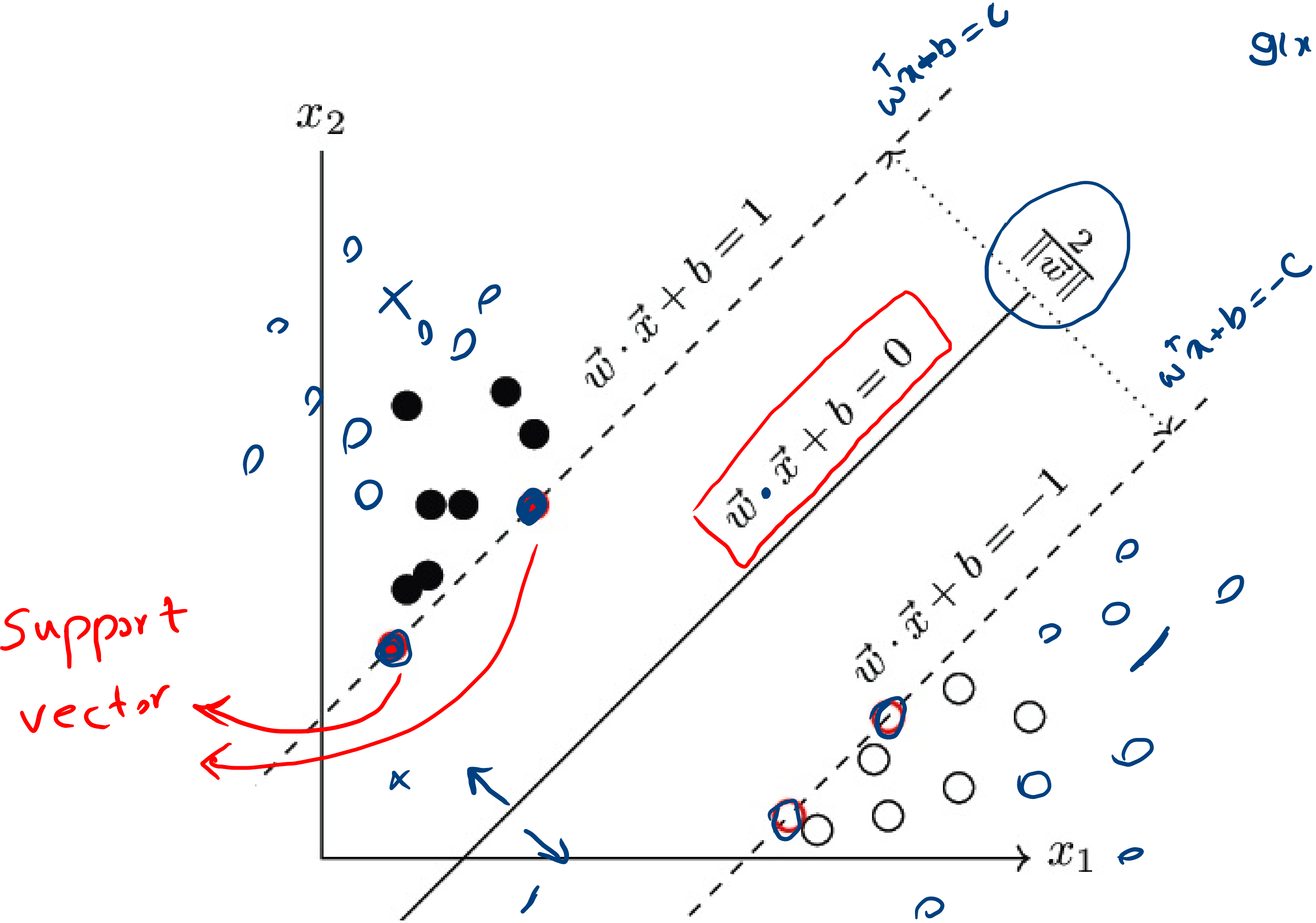
Support vector Machine (SVM)



Support Vector Machine (SVM)







$g(x) = \underline{\underline{w}}^T x + \underline{\underline{b}} = 0$

$g(x) \overset{+}{\underset{-}{\gtrless}} 0$

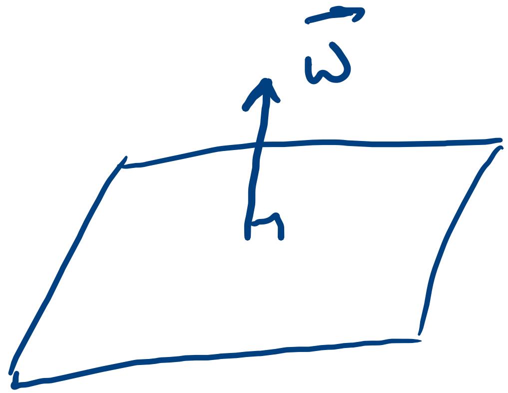
$w \cdot x = w^T x$

$\langle w, x \rangle$

$w^T x + b = c$

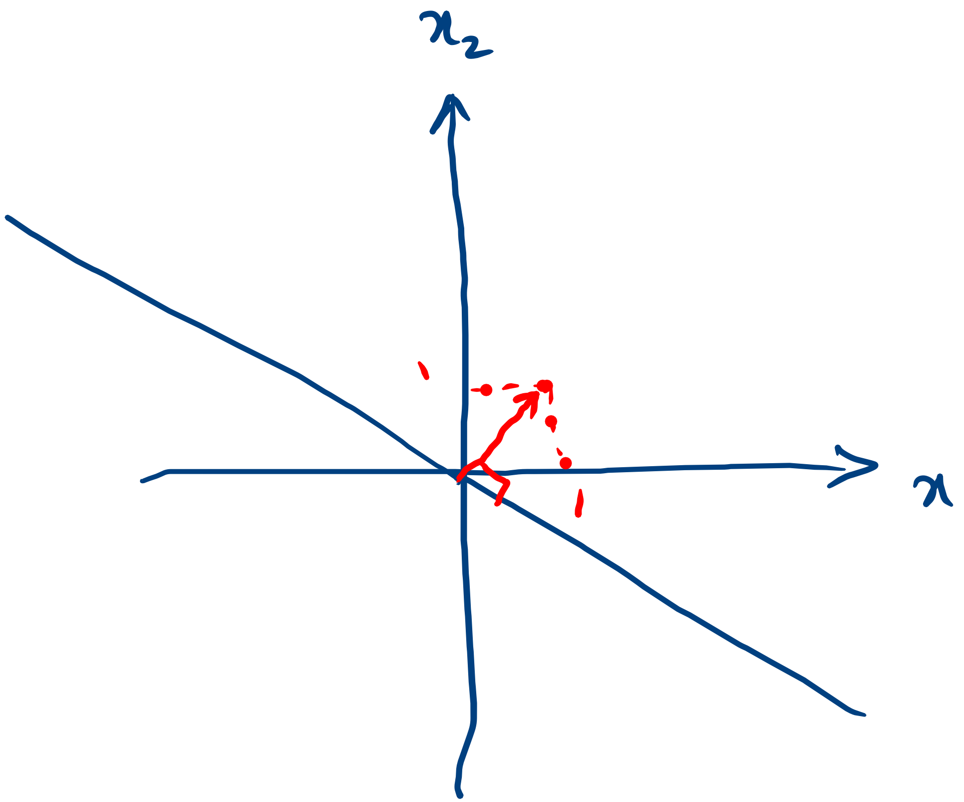
$\underset{w}{\frac{w^T}{c}} x + \underset{b}{\frac{b}{c}} = 1$

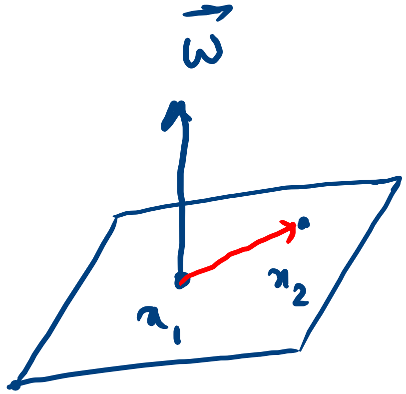
$$w^T x + b$$



$$x_1 + x_2 = 0$$

$$w = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$





$$\boxed{w^T x + b = 0}$$

$$\langle \vec{w}, \overrightarrow{x_1 x_2} \rangle = 0$$

$$w^T(x_2 - x_1) = w^T x_2 - w^T x_1$$

$$= \underbrace{w^T x_2 + b}_{0} - \underbrace{w^T x_1 - b}_{0} = 0$$

$$w^*, b^* = \arg\max_{w,b} \ \text{margin}(w,b)$$

$$D = \{(x_1, y_1), \dots, (x_n, y_n)\}$$

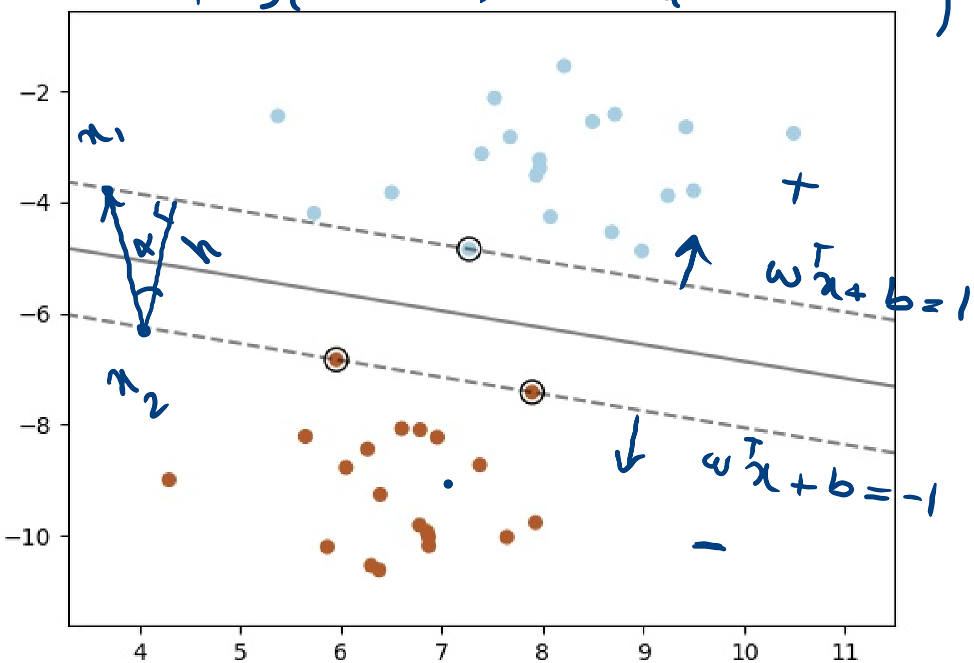
$$y_i \in \{\pm 1\}$$

$$\text{s.t.} \quad \text{if } y_i = +1 \Rightarrow w^T x_i + b \geq 1$$

$$\text{else if } y_i = -1 \Rightarrow w^T x_i + b \leq -1$$

$$\Rightarrow \boxed{y_i (w^T x_i + b) \geq 1}$$

$$i = 1, \dots, n$$



$$h = \left| \, \|x_1 x_2\|_2 \cos\alpha \, \right|$$

$$= \left| \frac{\|w\|_2}{\|w\|_2} \|x_1 x_2\| \cos\alpha \right|$$

$$= \frac{1}{\|w\|_2} \left| \langle w, \overrightarrow{x_1 x_2} \rangle \right|$$

$$\Rightarrow \quad h = \frac{1}{\|w\|_2} \left| w^T (x_1 - x_2) \right| = \frac{1}{\|w\|_2} \left| \underbrace{w^T x_1 + b}_{1} \underbrace{- w^T x_2 - b}_{1} \right| = \frac{2}{\|w\|_2}$$

$$\max_{\omega, b} \frac{2}{\|\omega\|_2}$$

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

$$\equiv \boxed{\begin{array}{l} \min\limits_{\omega, b} \|\omega\|_2^2 \\ \\ \text{s.t.} \quad y_i(x_i^T \omega + b) \geq 1 \end{array}}$$

$\Uparrow$

hard margin SVM

$\omega^T x + b$

$2\omega^T x + b$

$$\min_{x} \quad f(x)$$

s.t.

$$g_i(x) \leq 0 \qquad i=1,\dots,m$$

$$h_i(x) = 0 \qquad i=1,\dots,p$$

$\boxed{\text{KKT conditions}}$

Lagrangian: $L(\underline{x}, \underline{\lambda}, \underline{\mu}) = f(x) + \sum_{i=1}^{m} \lambda_i g_i(x) + \sum_{i=1}^{p} \mu_i h_i(x)$

$$\lambda_i \geq 0 \qquad i=1,\dots,m$$

KKT:

① $\nabla_x L(x, \lambda, \mu) = 0$

② $h_i(x) = 0$

$g_i(x) \leq 0$

③ $\lambda_i \geq 0$

④ $\underline{\lambda_i g_i(x) = 0}$ $\longrightarrow$ Complementary Slackness

x^*, λ^*, μ^*

$$\left.\begin{aligned} &\min_x \quad f(x) \\ &\text{s.t.} \quad g_i(x) \leq 0 \qquad i=1,\dots,m \\ &\qquad\quad h_i(x) = 0 \qquad i=1,\dots,p \end{aligned}\right\} \Longrightarrow x^*$$

$$L(x,\mu,\lambda) = f(x) + \sum_{i=1}^{m} \lambda_i g_i(x) + \sum_{i=1}^{p} \mu_i h_i(x)$$

$$g(\mu,\lambda) = \min_x L(x,\mu,\lambda)$$

$$g(\mu,\lambda) \leq f(x^*)$$

$$g(\mu,\lambda) = \min_x L(x,\mu,\lambda) \leq L(x^*,\mu,\lambda)$$

$$= f(x^*) + \underbrace{\sum_{i=1}^{m} \underbrace{\lambda_i}_{\geq 0} \underbrace{g_i(x^*)}_{\leq 0}}_{\leq 0} + \underbrace{\sum_{i=1}^{p} \mu_i h_i(x^*)}_{0}$$

$$\leq f(x^*)$$

$g(\lambda, \mu) \leq f(x^*)$ primal problem

$\forall \; \lambda_i \geq 0 \; , \; \mu$

$$\boxed{\max_{\lambda, \mu} \; g(\lambda, \mu) \qquad \lambda_i \geq 0}$$

→ Dual problem

Dual var.

Lagrange multipliers

KKT

Dual solution $\le$ primal solution

$$\boxed{g(\lambda^*, \mu^*) \le f(x^*)}$$

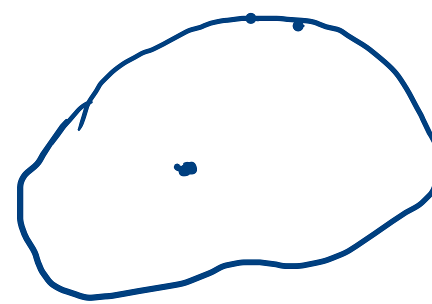
weak dulatily

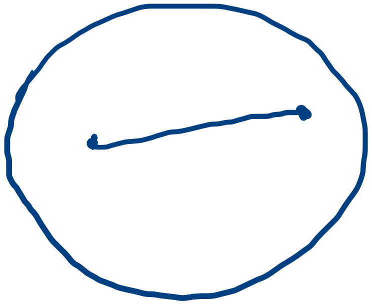
Strong Duality:

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

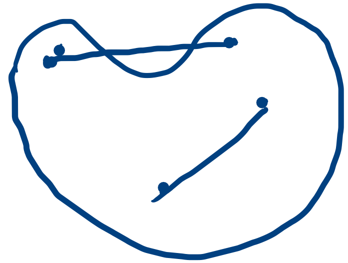
1. convex

Slater's Condition ← 2. $\boxed{\exists x \quad g_i(x) < 0}$



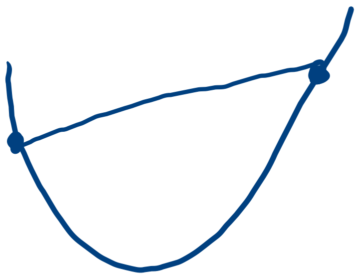


convex



non-convex

$f(x)$



$$w^T \underline{\underline{x}} + 2 < 0$$

$$\begin{cases} f(x) \longrightarrow \text{convex} \\ g_i(x) \longrightarrow \text{〃} \\ h_i(x) \longrightarrow \text{خطی} \end{cases}$$

$$\min_{\omega, b} \ \frac{1}{2} \|\omega\|_2^2$$

$$1 - y_i(\omega^T x_i + b) \leq 0$$

$$\text{s.t.} \quad y_i(\omega^T x_i + b) \geq 1 \qquad i = 1, \dots, n$$

$$L(\omega, b, \lambda) = \frac{1}{2} \|\omega\|_2^2 + \sum_{i=1}^{n} \lambda_i \left(1 - y_i(\omega^T x_i + b)\right)$$

$$\frac{\partial L}{\partial \omega} = \omega + \sum_{i=1}^{n} -\lambda_i y_i x_i = 0 \Rightarrow \boxed{\omega = \sum_{i=1}^{n} \lambda_i y_i x_i}$$

$$\frac{\partial L}{\partial b} = -\sum_{i=1}^{n} \lambda_i y_i = 0 \Rightarrow \boxed{\sum_{i=1}^{n} \lambda_i y_i = 0}$$

$$g(\lambda) = \frac{1}{2}\left(\sum_i \lambda_i y_i x_i\right)^T \left(\sum_i \lambda_i y_i x_i\right) + \sum_{i=1}^{n} \lambda_i$$

$$- \sum_i \lambda_i y_i \left(\sum_j \lambda_j y_j x_j\right)^T x_i - \underbrace{\sum_i \lambda_i y_i b}_{0}$$

$$= \frac{1}{2} \sum_i \sum_j \lambda_i \lambda_j y_i y_j x_i^T x_j - \sum_i \sum_j \lambda_i \lambda_j y_i y_j x_i^T x_j + \sum_i \lambda_i$$

$$\Rightarrow \boxed{g(\lambda) = -\frac{1}{2} \sum_i \sum_j \lambda_i \lambda_j y_i y_j x_i^T x_j + \sum_i \lambda_i} \qquad \lambda_i \geq 0$$

$$\boxed{\begin{aligned} &\max_{\lambda} \ g(\lambda) \\ &\text{s.t.} \quad \lambda_i \geq 0 \end{aligned}} \Rightarrow \lambda^* = \begin{bmatrix} \lambda_1^* \\ \vdots \\ \lambda_n^* \end{bmatrix}$$

$$\boxed{\omega^* = \sum_{i=1}^{n} \lambda_i^* y_i x_i}$$

Complementary Slackness:

$$\lambda_i^* g_i(x^*) = 0$$

$$\boxed{\underbrace{\lambda_k^*}_{\neq 0} \left(y_k (w^{*T} x_k + b) \right) = 0}$$