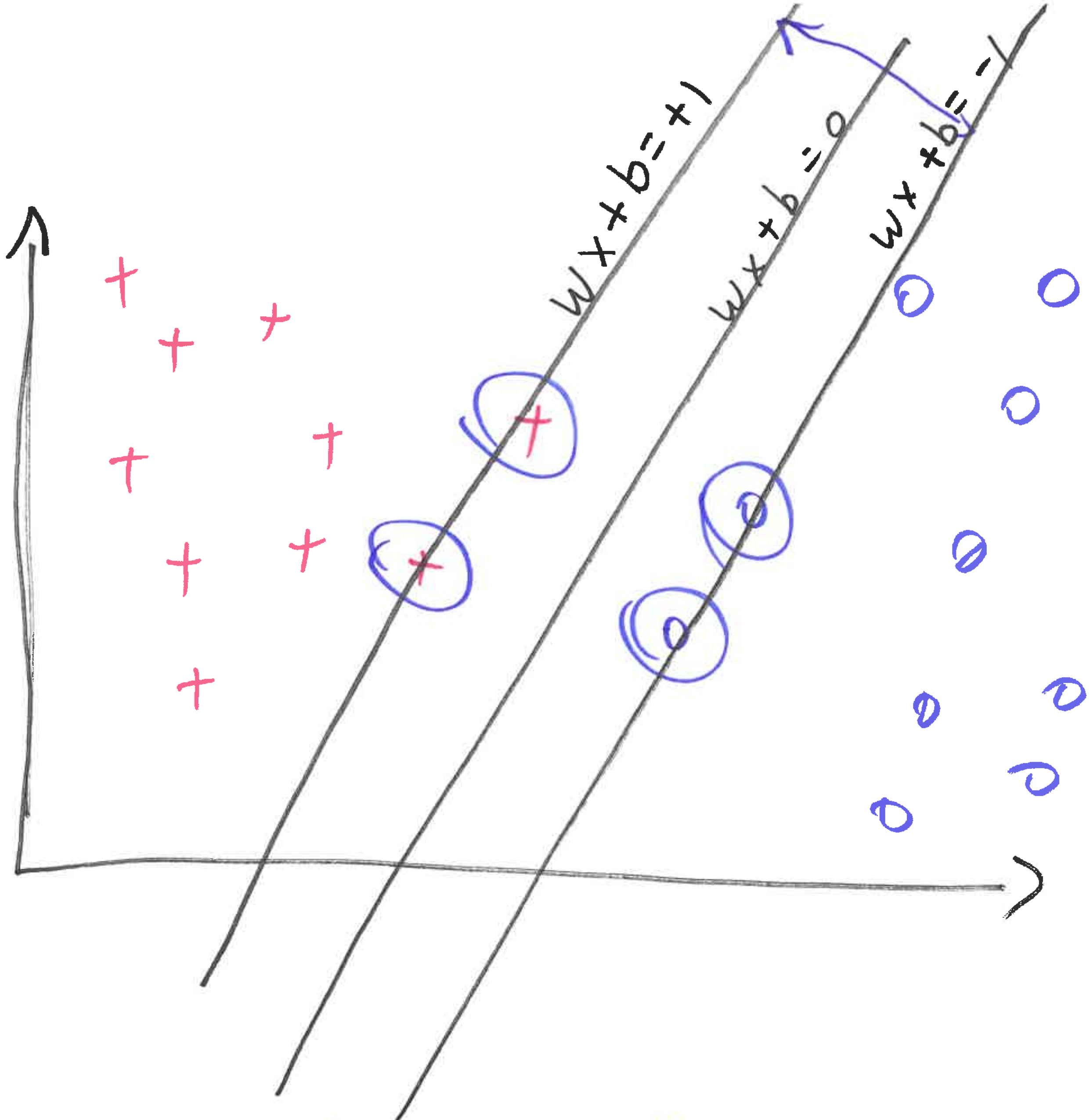
Support Vector Machines

1. Optimal Margin Classifier : Hard Margin
2. Kernel Trick
3. Soft Margin Case
4. SMO Algorithm
5. SVM Regression



$y \in \{+1 \quad -1\}$

$wx + b \geq +1$ for all +ve points

$wx + b \leq -1$ for all -ve points

$$\left.\begin{aligned} wx + b &= +1 \\ wx + b &= -1 \end{aligned}\right\} \text{parallel Lines}$$

$$wx + b - 1 = 0$$

$$wx + b + 1 = 0$$

Distance between $\parallel^{le}$ lines $= d = \dfrac{|\cancel{b} - 1 - \cancel{b} - 1|}{\sqrt{w_1^2 + w_2^2 + w_3^2 + \cdots w_n^2}}$

$$d = \frac{2}{\|w\|}$$

Objective

$$\text{Max} \quad \frac{2}{\|w\|} \quad \text{s.t}$$

$wx + b \geq +1$ for +ve points $y = +1$

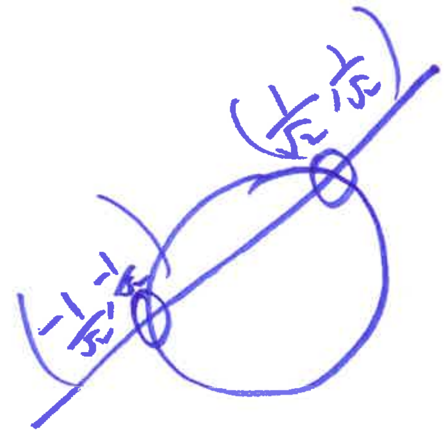
$wx + b \leq -1$ for -ve point $y = -1$

$$\boxed{\text{Max} \quad \frac{2}{\|w\|} \quad \text{s.t} \qquad y(wx + b) \geq +1 \quad \text{for all points}}$$

Lagrangian Multiplier

$$\min \; f(x, y) = x + y \quad s.t$$

$$g(x, y) = x^2 + y^2 = 1$$



$$\min \; f(x) \quad s.t \quad g(x) = 0$$

$$\mathcal{L}(x, \lambda) = f(x) + \lambda\, g(x)$$

$$f(x, y) = x + y \qquad g(x, y) = x^2 + y^2 - 1 = 0$$

$$\mathcal{L}(x, y, \lambda) =$$

$$f(x, y) + \lambda g(x, y)$$

$$L = x + y + \lambda(x^2 + y^2 - 1)$$

$$\frac{\delta L}{\delta x} = 0 \qquad \frac{\delta L}{\delta y} = 0 \qquad \frac{\delta L}{\delta \lambda} = 0$$

$$\frac{\delta L}{\delta x} = 1 + 2\lambda x = 0$$

$$x = \frac{-1}{2\lambda}$$

$$\frac{\delta L}{\delta y} \Rightarrow \qquad y = \frac{-1}{2\lambda}$$

$$\frac{\delta L}{\delta x} \Rightarrow \quad x = \frac{-1}{2\lambda}$$

$$\frac{\delta L}{\delta y} \Rightarrow \quad y = \frac{-1}{2\lambda}$$

$$\frac{\delta L}{\delta \lambda} \Rightarrow \quad x^2 + y^2 - 1 = 0$$

$$\frac{1}{4\lambda^2} + \frac{1}{4\lambda^2} - 1 = 0$$

$$\lambda^2 = \frac{1}{2}$$

$$\lambda = \pm \frac{1}{\sqrt{2}}$$

$$x = \cancel{\mp \frac{\lambda}{\sqrt{\cdot}}} \quad \pm \frac{1}{\sqrt{2}} \qquad y = \pm \frac{1}{\sqrt{2}}$$

$$f(x,y) = \frac{1}{\sqrt{2}} + \frac{1}{\sqrt{2}} = \sqrt{2} \ ; \quad \frac{-1}{\sqrt{2}} - \frac{1}{\sqrt{2}} = -\sqrt{2}$$

Duality Optimization

Primal Problem

$$\boxed{\begin{aligned} &\min \quad f(w) \qquad \text{s.t} \\ &h_i(w) = 0 \qquad i = 1 \ldots\ldots l \\ &g_i(w) \leq 0 \qquad i = 1 \ldots k \end{aligned}}$$

$$\mathcal{L}(w, \alpha, \beta) =$$

$$f(w) + \sum_{i=1}^{l} \beta_i h_i(w) + \sum_{i=1}^{k} \alpha_i g_i(w)$$

$$\theta_p(w) = \max_{\substack{\alpha, \beta \\ \alpha \geq 0}} \mathcal{L}(w, \alpha, \beta)$$

$$\theta_p(w) = \max_{\substack{\alpha\, \beta \\ \alpha \geq 0}} \mathcal{L}(w, \alpha, \beta)$$

Primal Problem $\Rightarrow \min_{w} \theta_p(w)$

Primal Problem

$$\theta_{\underline{p}}(w) = \begin{cases} f(w) & \text{if all contrain satisfied} \\ \infty & \end{cases}$$

Optimal solution for primal Problem p^*

$$\theta_D(\alpha, \beta) = \min_{w} \mathcal{L}(\alpha, \beta, w)$$

~~Original~~ Problem $= \max_{\alpha, \beta} \theta_p(\alpha, \beta)$

$$\min \max f(x) \geq \max \min f(x)$$

$$p^* \geq d^*$$

$p^* = d^*$ under certain conditions

if $f(w)$ & $g(w)$ are convex

& $h(w)$ is affine // linear

$g(w)$ is feasible

$$\frac{\delta L}{\delta w} = 0$$

$$\frac{\delta L}{\delta B_i} = 0$$

$$\frac{\delta L}{\delta \alpha_i} = 0$$

Karush-kuhn-Tucker

K.K.T dual

Complimentary

Conditions

$$\alpha_i \; g_i(w) = 0$$

$$\alpha_i \geq 0$$

$$g_i(w) \leq 0$$

$$L(w, \alpha)$$

$$L = \frac{1}{2} \|\underline{w}\|^2 - \sum_{i=1}^{m} \underline{\alpha_i} \left(\underline{y^{(i)}} \left(\vec{w} \cdot \vec{x}^{(i)} + \underline{b} \right) - 1 \right)$$

$$\alpha_i \geq 0$$

$$\frac{\delta L}{\delta w} = \frac{2}{2} \|w\| - \sum_{i=1}^{m} \alpha_i \, y^{(i)} \vec{x}^{(i)} = 0$$

$$\|w\| = \sum_{i=1}^{m} \alpha_i \, y^{(i)} x^{(i)}$$

$$\frac{\partial L}{\partial b} \Rightarrow \boxed{\sum_{i=1}^{m} \alpha_i y^{(i)} = 0}$$

$$\mathcal{L} = \frac{1}{2}\|w\|^2 - \sum \alpha_i \left(y^{(i)} \left(\underline{w} \cdot x^{(i)} + b \right) - 1 \right)$$

$$= \frac{1}{2} \sum \alpha_i y^i x^i \sum_j \alpha_j y^j x^j - \cancel{\sum \alpha_i \left(y^i \left(\alpha_j y^j x^i \right.\right.}$$

$$- \sum \alpha_i \left(y^{(i)} \left(\alpha_j y^{(j)} x^{(j)} \cdot x^{(i)} + b - 1 \right) \right)$$

$$= - \sum \alpha_i y^{(i)} x^{(i)} \sum \alpha_j y^{(j)} x^{(j)} + \boxed{\underline{\alpha_i} y^i b}^{=0}$$

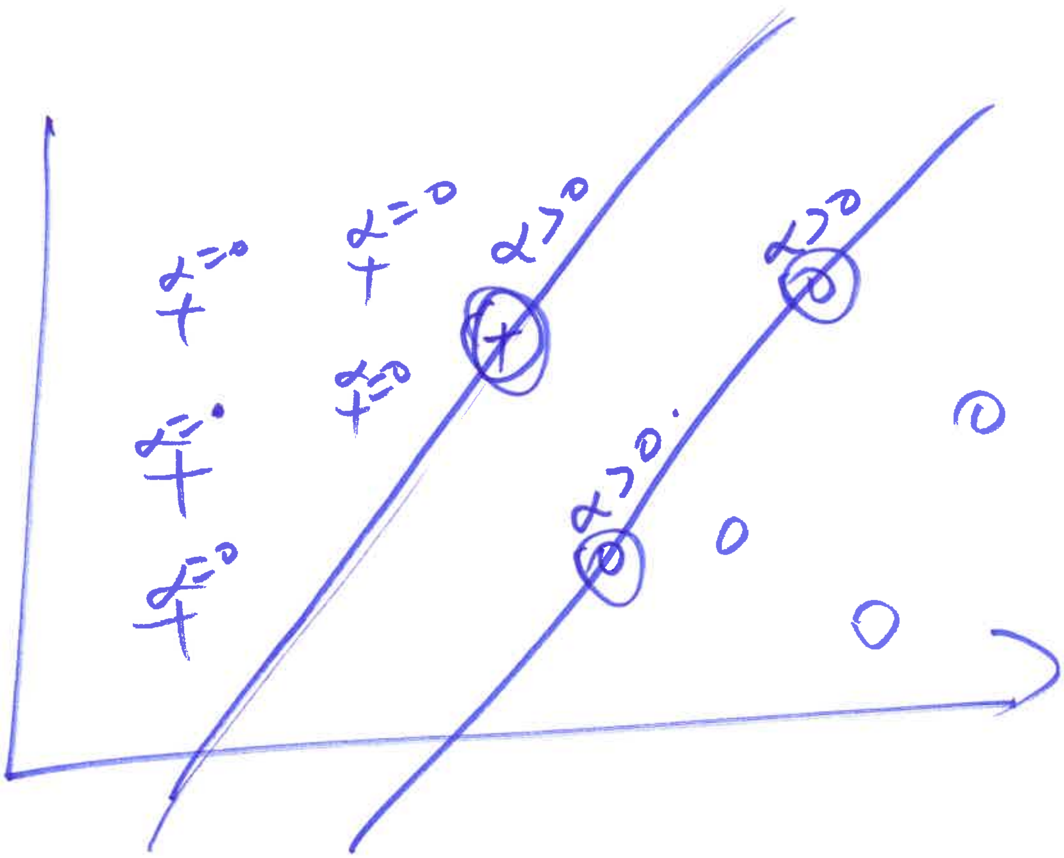
$$+ \sum \alpha_i$$

Max

$$\sum_{i=1}^{m} \alpha_i - \frac{1}{2} \sum_{i=1}^{m} \sum_{j=1}^{m} \alpha_i \alpha_j y^{(i)} y^{(j)} x^{(i)} \cdot x^{(j)}$$

$k(x^i, x^j)$

$\alpha_i \geq 0$ & $\sum \alpha_i y^{(i)} = 0$

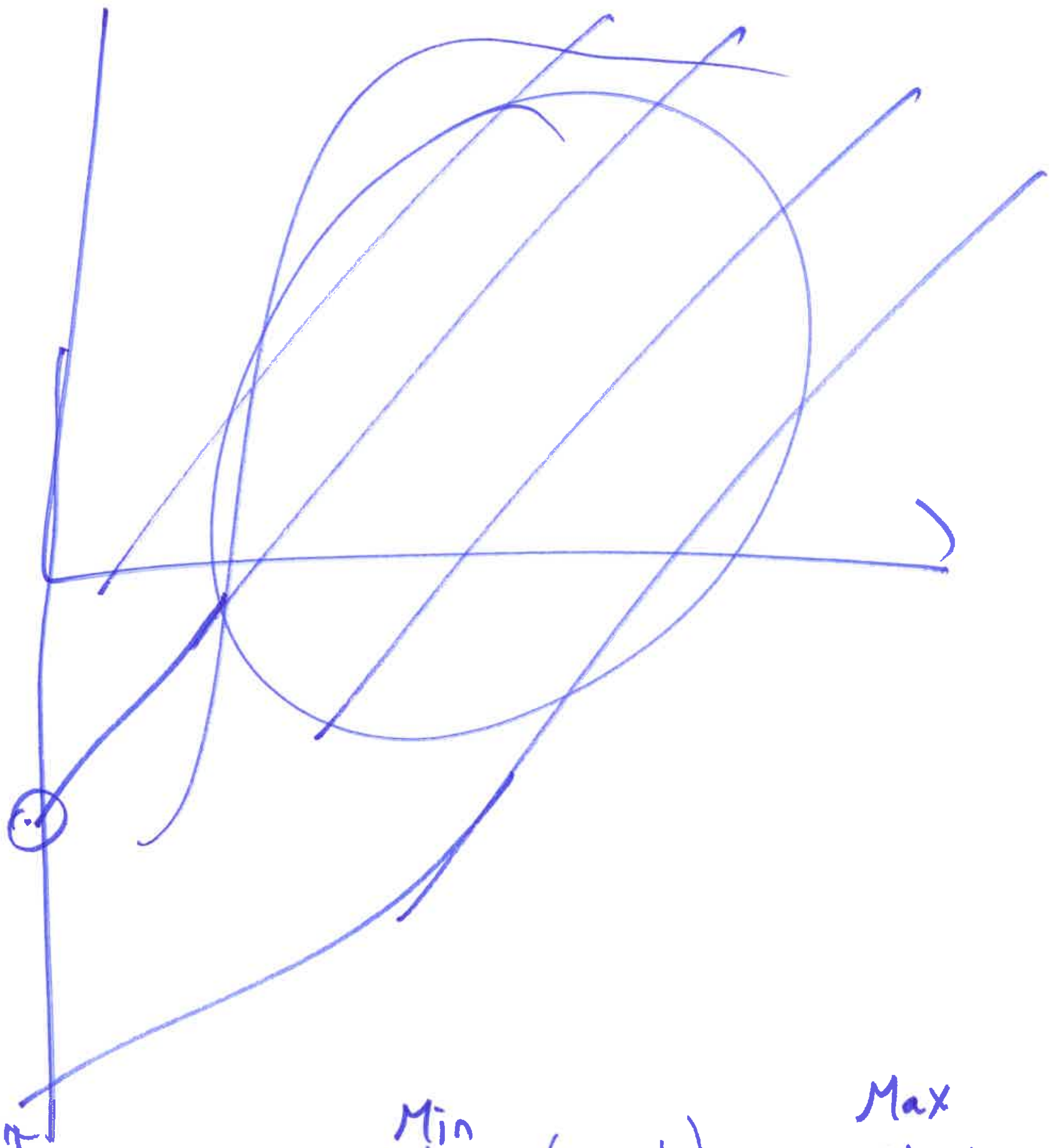


$$w = \sum \alpha_i y^i x^{(i)}$$

$x^t \to$ test case

$$w x^{(t)} + \underbrace{b}_{\gamma} = \begin{cases} > 0 & +ve \\ \leq 0 & -ve \end{cases}$$

$$\sum \alpha_i y^i \underline{\langle x^i \cdot x^t \rangle} + \underline{b}$$



$$b = -\frac{\underset{+ve\ point}{\overset{Min}{\cancel{Max}}}(wx+b) + \underset{-ve}{\overset{Max}{\cancel{Min}}}(wx+b)}{2}$$