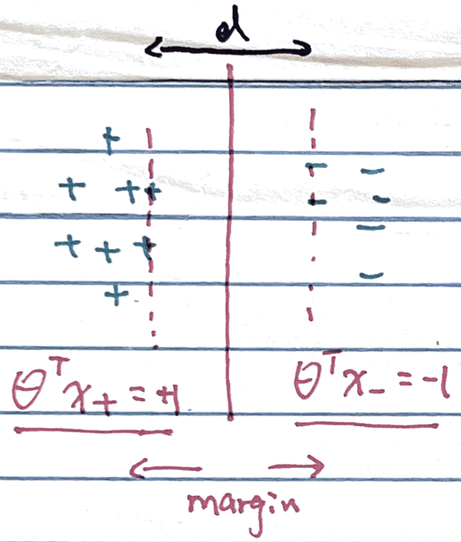
Slide 21

SVM objective

$\theta^T x_+ = +1$

$\theta^T x_- = -1$



$\hat{n} = \frac{\theta}{\|\theta\|_2}$

$\|\theta\| = \|\theta\|_2^2 = \|\theta\|_2 = \|\theta\|^2$

$\Rightarrow \underbrace{\sqrt{\theta_0^2 + \theta_1^2 + \cdots + \theta_{d-1}^2}}_{d \text{ items/scalars}}$

$\hat{n} \cdot \Delta(x_+, x_-) = \frac{\theta}{\|\theta\|_2} \cdot \Delta(x_+, x_-) = \frac{\boxed{\theta \cdot \Delta(x_+, x_-)}^{\,2}}{\|\theta\|_2} = \frac{2}{\|\theta\|_2}$

$(\theta^T x_+ = +1) - (\theta^T x_- = -1) \Rightarrow \theta^T (x_+ - x_-) = 2 = \theta \Delta(x_+, x_-)$

$\min_\theta J = -\overbrace{\sum_{i=1}^{N}}^{\text{cost f}} y_i \cdot \log(h_\theta(x_i)) + (1-y_i) \cdot (\log(1 - h_\theta(x_i))) + \sum \boxed{\frac{\lambda}{2}} \cdot \theta_j^2$

$\min_\theta J = C \overbrace{\sum_{i=1}^{N}}^{\text{cost f}} y_i \, \text{cost}_1(\theta^T x_i) + (1-y_i) \, \text{cost}_0(\theta^T x_i) + \frac{1}{2} \sum \theta_j^2$

$L_{hinge} = \max(0, 1 - y \cdot h_\theta(x))$. $\boxed{\text{Slides 19, 22}}$

What we want: $y = 1$, want $\theta^T x \geq 1$; $y = 0$, want $\theta^T x \leq -1$

$h(x) = \theta^T x$ $\quad n \to n^2$ $\quad \boxed{\text{Slides 23} \sim 25}$

$u = [u_1, u_2]$, $\|u\| = \sqrt{u_1^2 + u_2^2}$

$v = [v_1, v_2]$ $\quad \|v\| = \sqrt{v_1^2 + v_2^2}$

$u^T v = v^T u = u_1 v_1 + u_2 v_2$

$= \|u\| \cdot \underbrace{\|v\| \cdot \cos\theta}_{\text{Projection } P}$

$= \|u\|_2 \cdot P$

$\frac{1}{2}(y-t)^2$, $\nabla_w f \Rightarrow \nabla_w J = \frac{\partial J}{\partial w} = \left(\frac{\partial J}{\partial w_0}, \frac{\partial J}{\partial w_1}, \dots, \frac{\partial J}{\partial w_D}\right)$

$\boxed{\text{Slide 25}}$

$\theta^T x = \|\theta\|_2 \underbrace{\|x\|_2 \cdot \cos\theta}_{P \text{ from } x \text{ on } \theta} = \|\theta\|_2 \cdot P$

→ small θ; angle

θ ↓ weight matrix

$\boxed{\text{slides 26} \sim 27}$

if we ignore the first term by C ↓, $\Rightarrow \min_\theta J = \frac{1}{2} \cdot \sum_{j=1}^{d} \theta_j^2$

if P ↓; $\|\theta\|_2 \cdot P\downarrow$. $\underline{\underline{\geq 1}}$; ≤ -1

$\min_\theta \frac{1}{2} \sum_{j=1}^{N} \theta_j^2$, constrained by ① pos: $y_i = +1$, $\theta^T x_i = +1$.

$\underline{y_i(\theta^T x_i + b) \geq 1}$ ② neg: $y_i = -1$, $\theta^T x_i = -1$

$y_i(\theta^T x_i + b) - 1 \geq 0$

$\boxed{\text{slides 28} \sim 29}$

$J = \min_\theta \underbrace{\frac{1}{2} \sum_{j=1}^{d/N} \underline{\underline{\theta_j^2}} - \sum_i a_i (y_i(\theta^T x_i + \underline{\underline{b}}) - 1)}_{\text{function}}$, goal = $\theta \downarrow$; $a_0, \dots a_{N-1}$

$\frac{\partial J}{\partial w} = 0$; $\frac{\partial J}{\partial b} = 0$.

$J = \frac{1}{2} \sum_{j=1} \theta_j^2 - \left(\sum_i a_i (y_i(\theta^T x_i + b) - \underline{\sum_i a_i}\right)$

$= \boxed{\frac{1}{2} \sum_{j=1} \theta_j^2} - \left(\sum_i a_i y_i \theta^T x_i + b \cdot \sum_{i=1}^{N} a_i y_i - \sum_{i=1}^{N} a_i\right)$

$\frac{\partial J}{\partial \theta} = \theta - \sum_i a_i y_i x_i = 0$ (critical pt).

$\boxed{\theta = \sum_i a_i y_i x_i}$, the val that gives you the critical pt for J

$\frac{\partial J}{\partial b} = 0 - \left(0 + \sum_{i=1}^{N} a_i y_i - 0\right) = -\sum_{i=1}^{N} a_i y_i = 0.$

$\boxed{\sum_{i=1}^{N} a_i y_i = 0.}$

$\boxed{\text{slides 30-31.}}$

$$\theta = \sum_{i=1}^{d} a_i y_i \underline{x_i} \; ; \quad \sum_{i=1}^{N} a_i y_i = 0$$

$$\frac{1}{2}\sum_{j=1}^{d} \theta_j^2 - \sum_{i=1}^{n} a_i \left(y_i (\theta^T x_i + b) - 1 \right)$$

$$= \frac{1}{2}\sum_{j=1}^{d} \theta_j^2 - \left(\sum_{i=1}^{n} \underline{a_i y_i} (\theta^T x_i + \underline{b}) - \sum_{i=1}^{n} a_i \right)$$

$$= \frac{1}{2}\sum_{j=1}^{d} \underline{\theta_j^2} - \left(\underline{\sum_{i=1}^{n} a_i y_i \underline{\theta}^T x_i} + \overset{0}{\cancel{b \cdot \sum_{i=1}^{n} a_i y_i}} - \sum_{i=1}^{n} a_i \right)$$

$$= \frac{1}{2}\sum_{j=1}^{d}\sum_{i=1}^{d} a_i y_i x_i \cdot a_j y_j x_j$$

$$- \left(\sum_{i=1}^{n}\sum_{j=1} a_i y_i \cdot a_j \cdot y_j \cdot x_i \cdot x_j - \sum_{i=1}^{n} a_i \right).$$

$$= \underline{\frac{1}{2}} \sum_i \sum_j a_i a_j y_i y_j x_i x_j - \sum_i \sum_j a_i a_j y_i y_j x_i \cdot x_j + \sum_{i=1}^{n} a_i$$

$$= -\frac{1}{2} \sum_i \sum_j a_i a_j y_i y_j x_i x_j + \sum a_i$$