Tutorial 9

CS5487 Lecture Notes (2020)
Dr. Antoni B. Chan
Dept of Computer Science
City University of Hong Kong

Problem 9.4 - Soft Margin SVM

$$\min_{w,b,\xi} \underbrace{\frac{1}{2}\|w\|^2}_{\text{max-margin}} + \underbrace{C \sum_{i=1}^n \xi_i}_{\text{penalty on too many violators.}}$$

hyperparameter

s.t. ① $y_i(w^T x_i + b) \geq 1 - \xi_i, \ \forall i$ — "soft" margin constraint

② $\xi_i \geq 0, \ \forall i$ — slack variable

a) Lagrangian: Lagrange multipliers: α_i for ①, r_i for ②

① $y_i(w^T x_i + b) \geq 1 - \xi_i$

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

Lagrangian:

$$L(w,b,\xi,\alpha,r) = \frac{1}{2}\|w\|^2 + C\sum_i \xi_i - \sum_{i=1}^n \alpha_i\left(y_i(w^T x_i + b) - 1 + \xi_i\right) - \sum_{i=1}^n r_i \xi_i$$

b) $\frac{\partial L}{\partial w} = w - \sum_i \alpha_i y_i x_i = 0 \Rightarrow \boxed{w^* = \sum_{i=1}^n \alpha_i y_i x_i}$

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

$\frac{\partial L}{\partial \xi_i} = \underbrace{C - \alpha_i - r_i = 0}$ $\Rightarrow \boxed{r_i = C - \alpha_i}$

c) $L(\alpha) = \min_{w,b,\xi} L(w,b,\xi,\alpha,r)$

subst. $w = \sum_i \alpha_i y_i x_i$ into L:

$$L(\alpha) = \frac{1}{2}\Big\|\sum_i \alpha_i y_i x_i\Big\|^2 + C\sum_i \xi_i - \sum_i \alpha_i\Big(y_i\Big(\big(\sum_j \alpha_j y_j x_j\big)^T x_i + b\Big) - 1 + \xi_i\Big) - \sum_i r_i \xi_i$$

$\sum_i \sum_j \alpha_i \alpha_j y_i y_j x_i^T x_j$ ✓

$\sum_i \alpha_i y_i b = b \underbrace{\sum_i \alpha_i y_i}_{0}$

$\frac{1}{2}\|\sum_i \alpha_i y_i x_i\|^2 = \frac{1}{2}(\sum_i \alpha_i y_i x_i)^T(\sum_j \alpha_j y_j x_j) = \frac{1}{2}\sum_i\sum_j \alpha_i \alpha_j y_i y_j x_i^T x_j$ ✓

$$= -\frac{1}{2}\sum_i\sum_j \alpha_i\alpha_j y_i y_j x_i^T x_j + C\sum_i \xi_i + \sum_i \alpha_i - \sum_i \alpha_i \xi_i - \sum_i r_i \xi_i$$

$$\sum_i (C\xi_i - \alpha_i \xi_i - r_i \xi_i) = \sum_i \xi_i \underbrace{(C - \alpha_i - r_i)}_{0} = 0$$

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

d) $\begin{cases} r_i = C - \alpha_i \\ \alpha_i \geq 0 \\ r_i \geq 0 \end{cases} \Rightarrow \alpha_i + r_i = C$

$\alpha_i \leq C$

$0 \leq \alpha_i \leq C$

active $\Rightarrow$ "equality is active"

inactive $\Rightarrow$ "equality is inactive" i.e. inequality.

Dual SVM problem

$$\boxed{\begin{aligned} \max_{\alpha} L(\alpha) = \max_{\alpha} \; & \sum_i \alpha_i - \frac{1}{2}\sum_i \sum_j \alpha_i \alpha_j y_i y_j x_i^T x_j \\ \text{s.t.} \; & 0 \leq \alpha_i \leq C \;, \forall i \\ & \sum_i \alpha_i y_i = 0 \end{aligned}}$$

KKT conditions

1) $r_i > 0$ (active), $\alpha_i = 0$ (inactive)

$\Downarrow$ $\xi_i = 0$

$\Downarrow$ $y_i(w^T x_i + b) - 1 + \underbrace{\xi_i}_{0} > 0$

$\Downarrow$ $y_i(w^T x_i + b) > 1 \rightarrow$ correctly classified (beyond margin)

2) $r_i > 0$ (active), $\alpha_i > 0$ (active)

$\Downarrow$ $\xi_i = 0$

$\Downarrow$ $y_i(w^T x_i + b) - 1 + \underbrace{\xi_i}_{0} = 0$

$\Downarrow$ $y_i(w^T x_i + b) = 1 \rightarrow$ correctly classified on margin.

3) $r_i = 0$ (inactive) $\Rightarrow \underbrace{\cancel{r_i}}_{0} = C - \alpha_i \Rightarrow \alpha_i = C$ (active)

$\Downarrow$ $y_i(w^T x_i + b) - 1 + \xi_i = 0$

$\hookrightarrow y_i(w^T x_i + b) = 1 - \xi_i$

$\Downarrow$ $\xi_i > 0 \Rightarrow$ violates margin (slack variable is > 0)