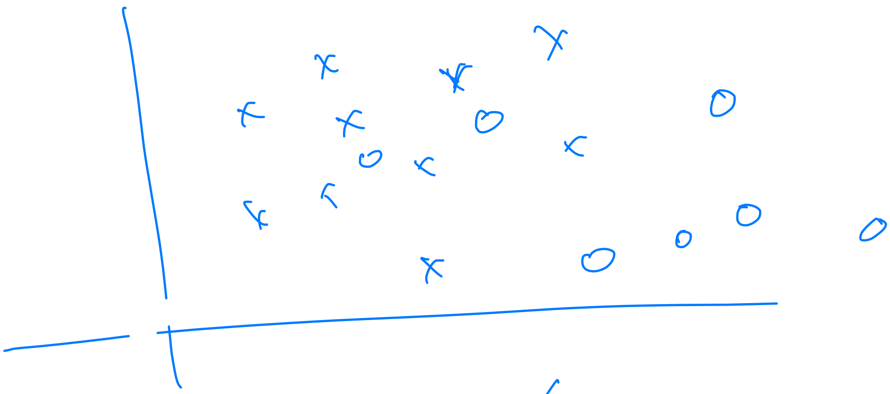
2021-12-08

Μοντέλο Support Vector Classifier για μη διαχωρίσιμες κλάσεις



$$\max_{\|b\|=1} M$$

$$y_i (x_i^T b + b_0) \geq M \quad \forall i$$

F εφικτή περιοχή

$F = \emptyset$

ανέφικτο.

Χαλάρωση

$$\max_{b, b_0, \xi_1, \xi_2, \dots, \xi_N} M$$

$$\|b\| = 1$$

$$y_i (x_i^T b + b_0) \geq M - \xi_i \quad i = 1, \dots, N$$

$$\xi_i \geq 0$$

(α)

$$\min \sum \xi_i$$

$$\|b\| = 1$$

$$y_i (x_i^T b + b_0) \geq M$$

$$\xi_i \geq 0 .$$

(β)

$$\max M$$

$$\|b\| = 1$$

$$y_i (x_i^T b + b_0) \geq M - \xi_i \ , i = 1, \dots, N$$

$$\xi_i \geq 0$$

$$\sum_{i=1}^{N} \xi_i \leq K$$

Μη κυρτός περιορ.

$\|b\| = 1$

$$\max M$$

$$y_i(x_i^T b + b_0) \geq M\|b\| - \underline{\xi_i}\|b\|$$

Εναλλακτικό μοντέλο

$$\max M$$

$$\|b\|=1$$

$$y_i(x_i^T b + b_0) \geq M - M\xi_i = M(1-\xi_i)$$

$$\xi_i \geq 0$$

$$\sum_{i=1}^{N} \xi_i \leq K$$

$$\Downarrow$$

$$\max M \qquad \text{Θέτουμε } \|b\| = \frac{1}{M}$$

$$y_i(x_i^T b + b_0) \geq M\|b\|(1-\xi_i)$$

$$\xi_i \geq 0$$

$$\sum \xi_i \leq K$$

$$\Downarrow$$

$$\max \frac{1}{\|b\|}$$

$$y_i(x_i^T b + b_0) \geq 1-\xi_i$$

$$\xi_i \geq 0$$

$$\sum \xi_i \leq K$$

$$\Longrightarrow$$

SVC

Πρόβλημα Τετραγωνικού Προγραμματισμού

$$\boxed{\begin{aligned} &\min \frac{1}{2}\|b\|^2 \\ &y_i(x_i^T b + b_0) \geq 1-\xi_i \\ &\xi_i \geq 0 \\ &\sum_{i=1}^{N} \xi_i \leq K \end{aligned}}$$

Lagrangean primal-dual formulation

Έστω πρωτεύον πρόβλημα : $Z_p = \min f(x)$

$$g_i(x) \ge 0 \quad , i = 1, \dots, M$$

$F_p = \{ x : g_i(x) \ge 0 , i = 1, .., M \}$ εφικτή περιοχή του P

Primal Lagrangean:

$$L_p(x,\mu) = f(x) - \sum_{i=1}^{M} \mu_i g_i(x) \qquad (\mu_i \ge 0)$$

Έστω $L_D(\mu) = \boxed{\min_{x} L_p(x,\mu)}$

$$\min_{x} L_p(x,\mu) \le \min_{x \in F_p} L_p(x,\mu)$$

$$\left.\begin{array}{l} \forall x \in F_p \Rightarrow g_i(x) \ge 0 \quad \forall i \\ \qquad\qquad \mu_i \ge 0 \quad \forall i \end{array}\right\} \mu_i g_i(x) \ge 0 \quad \forall i$$

$$\Rightarrow L_p(x,\mu) = f(x) - \underbrace{\sum \mu_i g_i(x)}_{\ge 0} \le f(x) \qquad \forall x \in F_p$$

$$\Rightarrow \min_{x \in F_p} L_p(x,\mu) \le \min_{x \in F_p} f(x) = Z_p$$

$$\Longrightarrow L_D(\mu) \le Z_p \qquad \forall \mu \ge 0 .$$

$$\Rightarrow \boxed{\underbrace{\max_{\mu \geq 0} L_D(\mu)}_{\text{δυϊκό πρόβλημα.}} \overset{z_P}{\leq} z_P} \Rightarrow \boxed{z_D \leq z_P}$$

↑ ασθενής δυϊκότητα.

ισχυρή δυϊκότητα $\boxed{z_D = z_P}$

Συμπληρωματικότητα

Στη βέλτιστη λύση, x^*, μ^*:

$$\mu_i^* \, g_i(x^*) = 0 \quad \forall i = 1, \ldots, M.$$

(αν $g_i(x^*) > 0 \Rightarrow \mu_i^* = 0$,

$\underline{\mu_i^* > 0} \Rightarrow g_i(x^*) = 0$)

$$\min \frac{1}{2}\|\beta\|^2$$

$$y_i(x_i^T\beta+\beta_0)\ge 1-\xi_i$$

$$\xi_i\ge 0$$

$$\sum \xi_i \le K.$$

$$\longrightarrow \boxed{\begin{array}{l} \min\limits_{\beta,\beta_0,\xi_i} \frac{1}{2}\|\beta\|^2 + C\cdot\sum_{i=1}^{N}\xi_i \\ y_i(x_i^T\beta+\beta_0)\ge 1-\xi_i \\ \xi_i \ge 0 \qquad i=1,\dots,N \end{array}}$$

πολ/στές Lagrange

$$\begin{cases} \alpha_i & y_i(x_i^T\beta+\beta_0)-(1-\xi_i)\ge 0 \quad i=1,\dots,N \\ \mu_i & \xi_i \ge 0, \quad i=1,\dots,N \end{cases}$$

$$\nabla_\beta\, x_i^T\beta = x_i$$

$$L_P(\beta,\beta_0,\underline{\xi_i},\alpha,\mu)=$$

$$\frac{1}{2}\beta^T\beta + C\sum_{i=1}^{N}\xi_i - \sum_{i=1}^{N}\alpha_i\left[y_i(x_i^T\beta+\beta_0)-(1-\xi_i)\right] - \sum_{i=1}^{N}\mu_i\xi_i$$

$$\underline{L_D(\alpha,\mu)} = \min_{\beta,\beta_0,\xi} \boxed{L_P(\beta,\beta_0,\xi,\alpha,\mu)}$$

$$\nabla_\beta L_P = \beta - \sum_{i=1}^{N}\alpha_i y_i x_i = 0 \quad\Rightarrow\quad \beta = \sum_{i=1}^{N}\alpha_i y_i x_i$$

$$\frac{\partial L_P}{\partial \beta_0} = \sum_{i=1}^{N}\alpha_i y_i = 0$$

$$\frac{\partial L}{\partial \xi_i} = C - \alpha_i - \mu_i = 0 \Rightarrow \alpha_i = C - \mu_i \quad \forall i = 1, \dots, N$$

$$\max_{\alpha, \mu \geq 0} L_D(\alpha, \mu)$$

Αντικαθιστούμε

$$b = \sum_i \alpha_i y_i x_i$$

$$\mu_i = C - \alpha_i$$

$$b^T \cdot b = \left(\sum_i \alpha_i y_i x_i\right)^T \cdot \left(\sum_{i'} \alpha_{i'} y_{i'} x_{i'}\right)$$

$$= \sum_{i,i'=1}^{N} \alpha_i \alpha_{i'} y_i y_{i'} x_i^T x_{i'}$$

Επίσης $\left[\sum_{i=1}^{N} \alpha_i y_i x_i^T\right] b = b^T b$

$$\left(\sum_{i=1}^{N} \alpha_i\right)\left(\sum_{j=1}^{N} b_j\right) \neq \sum_{i=1}^{N} \alpha_i b_i$$

$$= \sum_{i,j} \alpha_i b_j$$

$$L_D = -\frac{1}{2}\sum_{i,i'} \alpha_i \alpha_{i'} y_i y_{i'} x_i^T x_{i'}$$

$$+ C\sum_{i=1}^{N} \xi_i + \sum_i \alpha_i (1-\xi_i) - \sum \mu_i \xi_i$$

$$= -\frac{1}{2}\sum_{i,i'} \cdots\cdots$$

$$+ \sum_{i=1}^{N} \alpha_i + \sum_{i=1}^{N} \cancelto{0}{(C-\alpha_i-\mu_i)}\,\xi_i$$

$$\boxed{L_D(\alpha) = \sum_{i=1}^{N} \alpha_i - \frac{1}{2}\sum_{i,i'} \alpha_i \alpha_{i'} y_i y_{i'} x_i^T x_{i'}}$$

complementarity

$$\cdots \Longrightarrow \beta^*$$