smo

2020年8月11日 星期二 上午8:29

$$\min_{\alpha} \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}\alpha_i\alpha_j y_i y_j x_i^T x_j - \sum_{i=1}^{n}\alpha_i$$
$$s.t. \sum_{i=1}^{n}\alpha_i y_i = 0$$
$$0 \le \alpha_i \le C, \quad i=1,2,\dots,n$$

$$\min_{\alpha}\psi(\vec{\alpha}) = \min_{\alpha} \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}\alpha_i\alpha_j y_i y_j x_i^T x_j - \sum_{i=1}^{n}\alpha_i$$
$$s.t. \sum_{i=1}^{n}\alpha_i y_i = 0$$
$$0 \le \alpha_i \le C$$

$$\psi(\alpha_1,\alpha_2) = \frac{1}{2}\alpha_1\alpha_1 y_1 y_1 k_{11} + \frac{1}{2}\alpha_2\alpha_2 y_2 y_2 k_{22} + \frac{1}{2}\alpha_1\alpha_2 y_1 y_2 k_{12} + \frac{1}{2}\alpha_2\alpha_1 y_2 y_1 k_{21} + 2\times\frac{1}{2}\sum_{j=3}^{N} y_1 y_j k_{1j}\alpha_1\alpha_j + 2\times\frac{1}{2}\sum_{j=3}^{N} y_2 y_j k_{2j}\alpha_2\alpha_j - (\alpha_1+\alpha_2) + Constant$$
$$= \frac{1}{2}k_{11}\alpha_1^2 + \frac{1}{2}k_{22}\alpha_2^2 + y_1 y_2 k_{12}\alpha_1\alpha_2 + y_1\alpha_1\underbrace{\sum_{j=3}^{N}\alpha_j y_j k_{1j}}_{v_1} + y_2\alpha_2\underbrace{\sum_{j=3}^{N}\alpha_j y_j k_{2j}}_{v_2} - (\alpha_1+\alpha_2) + Constant$$

$$s.t. \ \alpha_1 y_1 + \alpha_2 y_2 = -\sum_{i=3}^{n} y_i\alpha_i = \tau$$
$$0 \le \alpha_i \le C$$

$$v_i = \sum_{j=3}^{n}\alpha_j y_j k_{ij} = f(x_i) - \sum_{j=1}^{2}\alpha_j y_j k_{ij} - b$$

由约束$\sum_{i=1}^{n}\alpha_i y_i = 0$得,$\alpha_1 y_1 + \alpha_2 y_2 = -\sum_{j=3}^{n}\alpha_j y_j = \tau$ (τ是定值)　　$\alpha_1^{old} y_1 + \alpha_2^{old} y_2 = -\sum_{j=3}^{n}\alpha_j y_j = \tau = \alpha_1^{new} y_1 + \alpha_2^{new} y_2$

两边同乘y_1得,$\alpha_1 y_1^2 + \alpha_2 y_2 y_1 = \tau y_1$

$$\alpha_1 y_1^2 = (\tau - \alpha_2 y_2) y_1$$
$$\alpha_1 = (\tau - \alpha_2 y_2) y_1 \quad (y_1^2 = 1)$$

$$\therefore \quad \psi(\alpha_2) = \frac{1}{2}k_{11}(\tau-\alpha_2 y_2)^2 + \frac{1}{2}k_{22}\alpha_2^2 + y_2 k_{12}(\tau-\alpha_2 y_2)\alpha_2 + (\tau-\alpha_2 y_2)v_1 + y_2\alpha_2 v_2 - (\tau-\alpha_2 y_2)y_1 - \alpha_2 + Constant$$

$$\frac{\partial\psi(\alpha_2)}{\partial\alpha_2} = k_{11}(\tau-\alpha_2 y_2)(-y_2) + k_{22}\alpha_2 + y_2 k_{12}\tau - 2y_2^2 k_{12}\alpha_2 - y_2 v_1 + y_2 v_2 + y_1 y_2 - 1$$
$$= -y_2 k_{11}\tau + k_{11}\alpha_2 y_2^2 + k_{22}\alpha_2 + y_2 k_{12}\tau - 2y_2^2 k_{12}\alpha_2 - y_2 v_1 + y_2 v_2 + y_1 y_2 - 1$$
$$= (k_{11}+k_{22}-2k_{12})\alpha_2 - k_{11}\tau y_2 + k_{12}\tau y_2 + y_1 y_2 - 1 - v_1 y_2 + v_2 y_2 = 0$$

$$(k_{11}+k_{22}-2k_{12})\alpha_2 = y_2(y_2 - y_1 + \tau k_{11} - \tau k_{12} + v_1 - v_2)$$
$$= y_2[y_2 - y_1 + \tau k_{11} - \tau k_{12} + (f(x_1) - \sum_{j=1}^{2} y_j\alpha_j k_{1j} - b) - (f(x_2) - \sum_{j=1}^{2} y_j\alpha_j k_{2j} - b)]$$

将$\tau = \alpha_1^{old} y_1 + \alpha_2^{old} y_2$代入,得:

$$(k_{11}+k_{22}-2k_{12})\alpha_2^{new,unc} = y_2[y_2 - y_1 + f(x_1) - f(x_2) + \underline{\tau k_{11} - \tau k_{12} - \sum_{j=1}^{2} y_j\alpha_j k_{1j} + \sum_{j=1}^{2} y_j\alpha_j k_{2j}}]$$

$$\tau k_{11} - \tau k_{12} - y_1\alpha_1 k_{11} - y_2\alpha_2 k_{12} + y_1\alpha_1 k_{21} + y_2\alpha_2 k_{22}$$
$$= (\alpha_1^{old} y_1 + \alpha_2^{old} y_2)k_{11} - (\alpha_1^{old} y_1 + \alpha_2^{old} y_2)k_{12} - \alpha_1^{old} y_1 k_{11} - \alpha_2^{old} y_2 k_{12} + \alpha_1^{old} y_1 k_{12} + \alpha_2^{old} y_2 k_{22}$$
$$= \alpha_2^{old} y_2 k_{11} - 2\alpha_2^{old} y_2 k_{12} + \alpha_2^{old} y_2 k_{22}$$
$$= (k_{11}+k_{22}-2k_{12})\alpha_2^{old} y_2$$

$$(k_{11}+k_{22}-2k_{12})\alpha_2^{new,unc} = (k_{11}+k_{22}-2k_{12})\alpha_2^{old} + y_2[y_2 - y_1 + f(x_1) - f(x_2)]$$

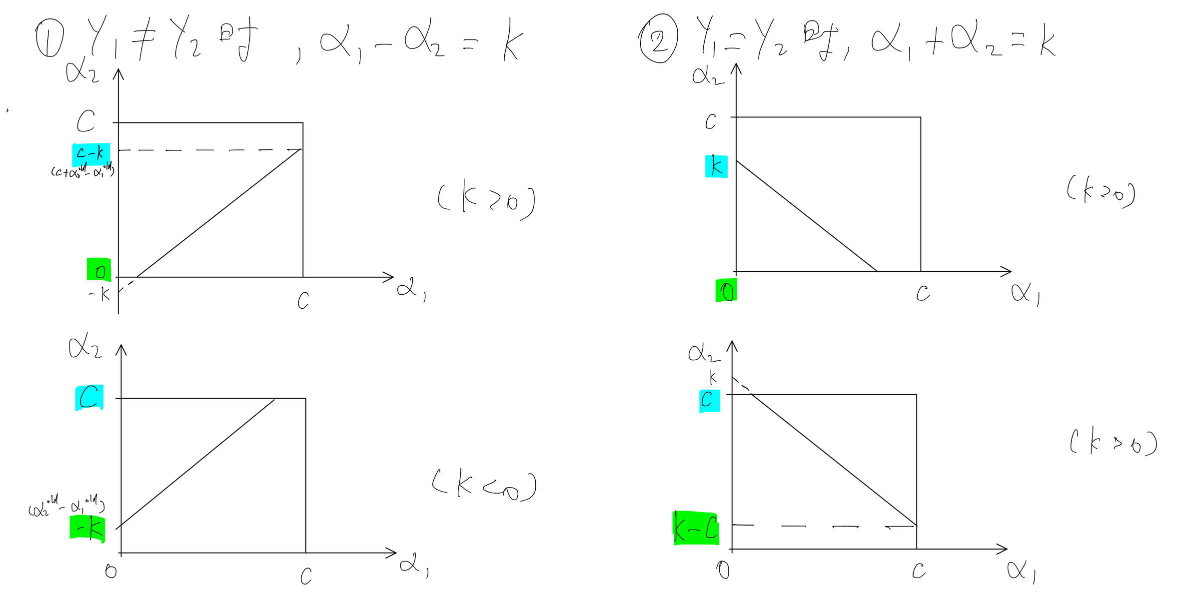
$\hat{w} = \sum_{i\in SV}\hat{\alpha}_i y_i x_i$,预测值$f(x) = \sum_{i=1}^{n}\alpha_i y_i k(x_i, x) + b$,预测值与真实值之差 $E_i = f(x_i) - y_i$

令 $\eta = k_{11} + k_{22} - 2k_{12}$

得 $\alpha_2^{new,unc} = \alpha_2^{old} + \dfrac{y_2(E_1 - E_2)}{\eta}$

由 $\begin{cases} 0 \le \alpha_i \le C \\ \alpha_1 y_1 + \alpha_2 y_2 = \tau \end{cases}$　又: ① $y_1 \ne y_2$时, $\alpha_1 - \alpha_2 = k$　② $y_1 = y_2$时, $\alpha_1 + \alpha_2 = k$

$$L \le \alpha_2^{new} \le H$$



$L = \max(0, \alpha_2^{old} - \alpha_1^{old})$　　$L = \max(0, \alpha_1^{old} + \alpha_2^{old} - C)$

$H = \min(C, C + \alpha_2^{old} - \alpha_1^{old})$　　$H = \min(C, \alpha_2^{old} + \alpha_1^{old})$

$$\therefore \alpha_2^{new} = \begin{cases} H, & \alpha_2^{new,unc} > H \\ \alpha_2^{new,unc}, & L \le \alpha_2^{new,unc} \le H \\ L, & \alpha_2^{new,unc} < L \end{cases}, \quad \alpha_1^{new} = (\tau - \alpha_2^{new} y_2) y_1$$
$$= (\alpha_1^{old} y_1 + \alpha_2^{old} y_2 - \alpha_2^{new} y_2) y_1$$
$$= \alpha_1^{old} + y_1 y_2(\alpha_2^{old} - \alpha_2^{new})$$

$$b_j^{new} = y_j - \sum_{i=1}^{n}\hat{\alpha}_i y_i x_j^T x_i$$
$$b_1^{new} = y_1 - \sum_{i=3}^{n}\alpha_i y_i k_{i1} - \alpha_1^{new} y_1 k_{11} - \alpha_2^{new} y_2 k_{21}$$
$$E_1 = f(x_1) - y_1 = \sum_{i=1}^{n}\alpha_i y_i k_{i1} + b^{old} - y_1$$
$$= \sum_{i=3}^{n}\alpha_i y_i k_{i1} + \alpha_1^{old} y_1 k_{11} + \alpha_2^{old} y_2 k_{21} + b^{old} - y_1$$

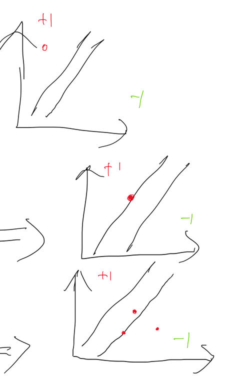
将E_1代入b_1^{new}得: $b_1^{new} = -E_1 + \alpha_1^{old} y_1 k_{11} + \alpha_2^{old} y_2 k_{21} + b^{old} - \alpha_1^{new} y_1 k_{11} - \alpha_2^{new} y_2 k_{21}$

$$= -E_1 - y_1 k_{11}(\alpha_1^{new} - \alpha_1^{old}) - y_2 k_{21}(\alpha_2^{new} - \alpha_2^{old}) + b^{old}$$

同理, $b_2^{new} = -E_2 - y_1 k_{12}(\alpha_1^{new} - \alpha_1^{old}) - y_2 k_{22}(\alpha_2^{new} - \alpha_2^{old}) + b^{old}$

$$\begin{cases} 0 < \alpha_1^{new} < C,\ 0 < \alpha_2^{new} < C: & b_1^{new} = b_2^{new} = b^{new} \\ 0 < \alpha_1^{new} < C,\ \alpha_2^{new} = 0或C: & b^{new} = b_1^{new} \\ \alpha_1^{new} = 0或C,\ 0 < \alpha_2^{new} < C: & b^{new} = b_2^{new} \\ \alpha_1^{new} = 0或C,\ \alpha_2^{new} = 0或C: & b^{new} = \frac{1}{2}(b_1^{new} + b_2^{new}) \end{cases}$$

KKT条件 $\begin{cases} 乘子非负: \alpha_i \ge 0, \ \beta_i \ge 0 \\ 约束条件: y_i(x_i^T w + b) - 1 + \varepsilon_i \ge 0, \ \varepsilon_i \ge 0 \\ 互补条件: \alpha_i(y_i(x_i^T w + b) - 1 + \varepsilon_i) = 0, \ \beta_i\varepsilon_i = 0 \end{cases}$, 即 $\begin{cases} \alpha_i = 0 \Longleftrightarrow y_i f(x_i) > 1 \Longleftrightarrow \\ 0 < \alpha_i < C \Longleftrightarrow y_i f(x_i) = 1 \Longleftrightarrow \\ \alpha_i = C \Longleftrightarrow y_i f(x_i) < 1 \Longleftrightarrow \end{cases}$



当所有变量的解都满足KKT条件时,这个最优化问题的解就得到了.

步骤:

① 选出样本点 x_1, x_2　→ 选出违反KKT条件最严重的样本点x_i

违反KKT: $\begin{cases} y_i f(x_i) - 1 \le 0, & \alpha_i = 0 \\ y_i f(x_i) - 1 \ne 0, & 0 < \alpha_i < C \\ y_i f(x_i) - 1 \ge 0, & \alpha_i = C \end{cases} \Longrightarrow \begin{cases} y_i E_i < -toler, & \alpha_i < C \\ y_i E_i > toler, & \alpha_i > 0 \end{cases}$

$y_i E_i = y_i[f(x_i) - y_i] = y_i f(x_i) - 1$　　toler: 违反KKT条件的程度

② $E_1 = \sum_{i=1}^{n}\alpha_i y_i k_{1i} + b^{old} - y_1$, $E_2 = \underbrace{\sum_{i=1}^{n}\alpha_i y_i k_{2i}}_{} + b^{old} - y_2$,

$\eta = k_{11} + k_{22} - 2k_{12}$

np.multiply($\vec{\alpha}$, $\vec{y}$).T * K

$$\alpha_1 y_1 k_{21} + \alpha_2 y_2 k_{22} + \alpha_3 y_3 k_{23} + \alpha_4 y_4 k_{24}$$
$$= (\alpha_1 y_1 \ \ \alpha_2 y_2 \ \ \alpha_3 y_3 \ \ \alpha_4 y_4)\cdot\begin{pmatrix} k_{21} \\ k_{22} \\ k_{23} \\ k_{24} \end{pmatrix}$$

③ $\alpha_2^{new,unc} = \alpha_2^{old} + \dfrac{y_2(E_1 - E_2)}{\eta}$

④ 若 $y_1 \ne y_2$, $L = \max(0, \alpha_2^{old} - \alpha_1^{old})$

$H = \min(C, C + \alpha_2^{old} - \alpha_1^{old})$

若 $y_1 = y_2$, $L = \max(0, \alpha_1^{old} + \alpha_2^{old} - C)$

$H = \min(C, \alpha_2^{old} + \alpha_1^{old})$

⑤ $\alpha_2^{new} = \begin{cases} H, & \alpha_2^{new,unc} > H \\ \alpha_2^{new,unc}, & L \le \alpha_2^{new,unc} \le H \\ L, & \alpha_2^{new,unc} < L \end{cases}$, $\alpha_1^{new} = \alpha_1^{old} + y_1 y_2(\alpha_2^{old} - \alpha_2^{new})$

⑥ $b_1^{new} = -E_1 - y_1 k_{11}(\alpha_1^{new} - \alpha_1^{old}) - y_2 k_{21}(\alpha_2^{new} - \alpha_2^{old}) + b^{old}$

$b_2^{new} = -E_2 - y_1 k_{12}(\alpha_1^{new} - \alpha_1^{old}) - y_2 k_{22}(\alpha_2^{new} - \alpha_2^{old}) + b^{old}$

$$\begin{cases} 0 < \alpha_1^{new} < C,\ 0 < \alpha_2^{new} < C: & b_1^{new} = b_2^{new} = b^{new} \\ 0 < \alpha_1^{new} < C,\ \alpha_2^{new} = 0或C: & b^{new} = b_1^{new} \\ \alpha_1^{new} = 0或C,\ 0 < \alpha_2^{new} < C: & b^{new} = b_2^{new} \\ \alpha_1^{new} = 0或C,\ \alpha_2^{new} = 0或C: & b^{new} = \frac{1}{2}(b_1^{new} + b_2^{new}) \end{cases}$$

$$\hat{w} = \sum_{i=1}^{n}\hat{\alpha}_i y_i x_i = \alpha_1 y_1\vec{x}_1 + \alpha_2 y_2\vec{x}_2 + \alpha_3 y_3\vec{x}_3 + \alpha_4 y_4\vec{x}_4$$
$$= (\alpha_1 y_1 \ \ \alpha_2 y_2 \ \ \alpha_3 y_3 \ \ \alpha_4 y_4)\cdot\begin{pmatrix} \vec{x}_1 \\ \vec{x}_2 \\ \vec{x}_3 \\ \vec{x}_4 \end{pmatrix}$$
$$= np.multiply(\vec{\alpha}, \vec{y}) * X$$