SMO 算法的最优化条件计算

Tuesday, November 26, 2024 10:49 PM

$$\psi(\alpha_1,\alpha_2)=\frac{1}{2}(\alpha_1^2K_{11}+2\alpha_1\alpha_2y_1y_2K_{12}+\alpha_2^2K_{22})-(\alpha_1+\alpha_2)+\alpha_1y_1V_1+\alpha_2y_2V_2$$

有： $\alpha_1=(\xi-\alpha_2y_2)y_1$，代入：

$$\psi(\alpha_1,\alpha_2)=\frac{1}{2}\left[(\xi-\alpha_2y_2)^2K_{11}+2(\xi-\alpha_2y_2)\alpha_2y_2K_{12}+\alpha_2^2K_{22}\right]-(\xi-\alpha_2y_2)y_1-\alpha_2+(\xi-\alpha_2y_2)v_1+\alpha_2y_2v_2+C$$

求导：有：

$$\frac{\partial\psi}{\partial\alpha_2}=\frac{1}{2}\left[-2y_2(\xi-\alpha_2y_2)K_{11}+2(\xi-\alpha_2y_2)y_2K_{12}+(-2y_2)\alpha_2y_2K_{12}+2\alpha_2K_{22}\right]+y_1y_2-1-y_2V_1+y_2V_2.$$

整理，有：

$$\frac{\partial\psi}{\partial\alpha_2}=\left[\alpha_2K_{11}-y_2\xi K_{11}+\xi y_2K_{12}-\alpha_2K_{12}-\alpha_2K_{12}+\alpha_2K_{22}\right]+y_1y_2-1-y_2V_1+y_2V_2$$

$$=\alpha_2(K_{11}-2K_{12}+K_{22})+y_2\xi K_{12}-y_2\xi K_{11}+y_1y_2-1-y_2(V_1-V_2) \quad ☆$$

此处重新考虑 SVM 计算方式：对于某个 x，预测值为：

$f(x)=w^Tx+b$，其中 w 为权重向量 $\xrightarrow{\text{有关系}} w=\sum_{i=1}^{n}\alpha_iy_iK(x_i,x)$

即 $f(x)$ 为原先的预测值，对应地，α 为原始权重向量

由 $f(x)^{old}=\sum_{i=1}^{n}\alpha_i^{old}y_iK(x_i,x)+b$ 为迭代前的SVM估计值

由于 $\alpha_3,\alpha_4\cdots\alpha_n$ 是不变的，则可以以 $f(x)$ 表示出 V_1, V_2.

有
$$\begin{cases}V_1=\sum_{i=3}^{n}\alpha_i^{old}y_iK_{i1}=f(x_1)-\alpha_1^{old}y_1K_{11}-\alpha_2^{old}y_2K_{12}-b\\ V_2=\sum_{i=3}^{n}\alpha_i^{old}y_iK_{i2}=f(x_2)-\alpha_1^{old}y_1K_{12}-\alpha_2^{old}y_2K_{22}-b\end{cases}$$

其中. V_1, V_2 ⇒

取 $\frac{\partial\psi}{\partial\alpha_2}=0$，则：

$$\alpha_2(K_{11}-2K_{12}+K_{22})+y_2\xi(K_{12}-K_{11})+y_1y_2-1=y_2(V_1-V_2)$$

$$\alpha_2(K_{11}-2K_{12}+K_{22})+y_2\xi(K_{12}-K_{11})+y_1y_2-1=y_2(v_1-v_2)$$

代入 $v_1-v_2=f(x_1)-f(x_2)-\alpha_1^{old}y_1(K_{11}-K_{12})-\alpha_2^{old}y_2(K_{12}-K_{22})$

$$\xrightarrow{\alpha_1=y_1(\xi-\alpha_2y_2)} = f(x_1)-f(x_2)-\xi(K_{11}-K_{12})+\alpha_2^{old}y_2(K_{11}-K_{12})+\alpha_2^{old}y_2(K_{22}-K_{12})$$

$$=f(x_1)-f(x_2)-\xi(K_{11}-K_{12})+\alpha_2^{old}y_2(K_{11}-2K_{12}+K_{22}) \quad ☆_2$$

将☆$_2$代入☆$_1$,有:

$$\frac{\partial\psi}{\partial\alpha_2}=\alpha_2^{new}(K_{11}-2K_{12}+K_{22})+\cancel{y_2\xi(K_{12}-K_{11})}+y_1y_2-1-y_2(f(x_1)-f(x_2))+\cancel{\xi y_2(K_{11}-K_{12})}$$

$$-\alpha_2^{old}(K_{11}-2K_{12}+K_{22})$$

$$=(\alpha_2^{new}-\alpha_2^{old})(K_{11}-2K_{12}+K_{22})+y_2(y_1-y_2)-y_2(f(x_1)-f(x_2))=0$$

则有: $\alpha_2^{new}-\alpha_2^{old}=\dfrac{y_2(f(x_1)-f(x_2))-y_2(y_1-y_2)}{K_{11}-2K_{12}+K_{22}}=y_2\dfrac{(f(x_1)-y_1)-(f(x_2)-y_2)}{K_{11}-2K_{12}+K_{22}}$

我们取 $E_1=f(x_1)-y_1$, $E_2=f(x_2)-y_2$. $\eta=K_{11}-2K_{12}+K_{22}$, 则

$$\boxed{\alpha_2^{new}=\alpha_2^{old}+y_2\frac{E_1-E_2}{\eta}}$$

☆ 为权重迭代公式。