SMO 算法概述

Yimeng Ren

1 SMO 算法概述

SMO 是由 Platt 在 1998 年提出的、针对软间隔最大化 SVM 对偶问题求解的一个算法，其基本思想很简单：在每一步优化中，挑选出诸多参数（$\alpha_k(k=1,2,\dots,N)$）中的两个参数（α_i,α_j）作为"真正的参数"，其余参数都视为常数，从而问题就变成了类似于二次方程求最大值的问题，从而我们就能求出解析解。

具体而言，SMO 要解决的是如下对偶问题：

$$\max_{\alpha} L(\alpha)=-\frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N}\alpha_i\alpha_j y_i y_j\left(x_i\cdot x_j\right)+\sum_{i=1}^{N}\alpha_i s.t. \sum_{i=1}^{N}\alpha_i y_i=0, 0\le\alpha_i\le C, i=1,2,...,N$$

其大致求解步骤则可以概括如下：

1. 选出 $\alpha_1,\alpha_2,\dots,\alpha_N$ 中"最不好的"两个参数 α_i,α_j
2. 只把 α_i,α_j 视为参数并把其他的 α_k 视为常数。注意到 α_i,α_j 需要满足 $\sum_{i=1}^{N}\alpha_i y_i=0$ 和 $0\le\alpha_i,\alpha_j\le C$，所以求完解后需要检查是否满足约束；如不满足，则进行调整。

SMO 算法

输入：训练数据集 $T=\{(x_1,y_1),(x_2,y_2),\cdots,(x_N,y_N)\}$，其中 $x_i\in\mathcal{X}=\mathbf{R}^n, y_i\in\mathcal{Y}=\{-1,+1\}, i=1,2,\cdots,N$，精度 ϵ；

输出：近似解 $\hat{\alpha}$

（1）取初始值 $\alpha^{(0)}, k=0$；

(2) 选取优化变量 $\alpha_1^{(k)},\alpha_2^{(k)}$，解析求解两个变量的最优化问题（如下），求得最优解 $\alpha_1^{(k+1)},\alpha_2^{(k+1)}$，更新 α 为 $\alpha^{(k+1)}$；

$$
\begin{aligned}
\min_{\alpha_1,\alpha_2} \quad & W(\alpha_1,\alpha_2)=\frac{1}{2}K_{11}\alpha_1^2+\frac{1}{2}K_{22}\alpha_2^2+y_1y_2K_{12}\alpha_1\alpha_2- \\
& (\alpha_1+\alpha_2)+y_1\alpha_1\sum_{i=3}^{N}y_i\alpha_iK_{i1}+y_2\alpha_2\sum_{i=3}^{N}y_i\alpha_iK_{i2} \\
\text{s.t.} \quad & \alpha_1y_1+\alpha_2y_2=-\sum_{i=3}^{N}y_i\alpha_i=\zeta \\
& 0\leqslant\alpha_i\leqslant C, \quad i=1,2
\end{aligned}
$$

其中，$K_{ij}=K\left(x_i,x_j\right),i,j=1,2,\cdots,N,\zeta$ 是常数，第一行目标函数中省略了不含 α_1,α_2 的常数项。

（3）若在精度 ϵ 范围内满足停止条件：

$$
\sum_{i=1}^{N}\alpha_iy_i=0, \quad 0\leqslant\alpha_i\leqslant C, \quad i=1,2,\cdots,N
$$

$$
y_i\cdot g(x_i)\begin{cases}\geqslant 1, & \{x_i \mid \alpha_i=0\} \\ =1, & \{x_i \mid 0<\alpha_i<C\} \\ \leqslant 1, & \{x_i \mid \alpha_i=C\}\end{cases}
$$

其中，

$$
g(x_i)=\sum_{j=1}^{N}\alpha_jy_jK\left(x_j,x_i\right)+b
$$

则进入（4），否则令 $k=k+1$，转回（2）；

（4）取 $\hat{\alpha}=\alpha^{(k+1)}$