Solve $p + \frac{1}{k}z = \frac{1}{k}v$

Multi-class SVM Loss Subproblem

Ruohan Zhan

August 21, 2016

Consider problem

$$\min_{z\in\mathbf{R}^n} \frac{1}{2\kappa}\|z-v\|^2 + \sum_{j\neq i}\max(0, 1 - z_i + z_j), \tag{1}$$

which is equivalent to the following problem with constraints:

$$\min_{z,\xi\in\mathbf{R}^n} \frac{1}{2\kappa}\|z-v\|^2 + \mathbf{1}^T\xi, \quad s.t. \xi \succeq 0, \xi \succeq Az + b, \tag{2}$$

where $A = I - \mathbf{1}e_i^T$ and $b = \mathbf{1} - e_i = -Ae_i$.

The Lagrangian is as follows:

$$L(z,\xi,\lambda,\mu) = \frac{1}{2\kappa}\|z-v\|^2 + \mathbf{1}^T\xi - \mu^T\xi + \lambda^T(Az+b-\xi), \quad \mu \succeq 0, \lambda \succeq 0. \tag{3}$$

The dual function is

$$g(\lambda,\mu) = \inf_{z,\xi} L(z,\xi,\lambda,\mu) = \begin{cases} \lambda^T(b+Av) - \dfrac{\kappa}{2}\lambda^T AA^T\lambda, \text{ if } \lambda+\mu = 1 \\ -\infty, \text{ otherwise.} \end{cases} \tag{4}$$

where $z = v - \kappa A^T\lambda$.

Therefore, the dual problem of problem 2 is

$$\min_{\lambda\in\mathbf{R}^n} \quad \frac{\kappa}{2}\lambda^T AA^T\lambda - \lambda^T(b+Av), \text{ s.t. } 1 \succeq \lambda, \lambda \succeq 0. \tag{5}$$

which is equivalent to the following problem:

$$\min_{\lambda\in\mathbf{R}^n} \quad \frac{\kappa}{2}\|A^T\lambda - \frac{1}{\kappa}(v - e_i)\|^2, \text{ s.t. } 1 \succeq \lambda, \lambda \succeq 0. \tag{6}$$

Denote $c = \frac{1}{\kappa}(v - e_i)$, and $x = (\lambda_1, \ldots, \lambda_{i-1}, -\sum_{j\neq i}\lambda_j, \lambda_{i+1}, \ldots, \lambda_n)$, we then wish to solve the following problem:

$$\min_{x\in\mathbf{R}^n} \sum_{j=1}^{n}(x_j - c_j)^2, \quad \sum_{j=1}^{n} x_j = 0, 0 \leq x_j \leq 1, j \neq i. \tag{7}$$

Relax it ,we can get

$$\min_{x\in\mathbf{R}^n} \frac{1}{2}\sum_{j=1}^{n}(x_j - c_j)^2 + \frac{\mu}{2}(\sum_{j=1}^{n} x_j)^2, \quad 0 \leq x_j \leq 1, j \neq i. \tag{8}$$

Note that when $\mu \to \infty$, the solution of 8 converges to the solution of 7. (this needs to be double checked. Is it reasonable that the solution by taking $\mu \to \infty$ is the solution of problem 8 with μ going to ∞?)

There are THREE steps in total to solve this problem 7:

1. solve problem 8 without constraint

The first order optimality condition for the nonconstraint version of problem 8 is as follows:

$$\begin{pmatrix} 1+\mu & \mu & \ldots & \mu \\ \mu & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \mu \\ \mu & \ldots & \mu & 1+\mu \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{pmatrix}, \tag{9}$$

so

$$x_i = c_i - \mu\frac{\sum_{j=1}^{n} c_j}{1 + n\mu}$$

2.Take projection operation onto $[0, 1]$

Denote $P(x)$ as the projection operation onto $[0, 1]$:

$$P(x) = \begin{cases} 0, & x \leq 0 \\ x, & 0 \leq x \leq 1 \\ 1, & 1 \leq x. \end{cases} \tag{10}$$

Then the solution of problem 8 is as follows:

$$\begin{cases} x_k = P(c_k - \mu\frac{\sum_{j=1}^{n} c_j}{1 + n\mu}), & k \neq i \\ x_i = c_i - \mu\frac{\sum_{j=1}^{n} c_j}{1 + n\mu}. \end{cases} \tag{11}$$

3. Take $\mu \to \infty$ and update $x_i = -\sum_{j \neq i} x_j$

$$\begin{cases} x_k^* = P(c_k - \frac{\sum_{j=1}^{n} c_j}{n}), \quad k \neq i \\ x_i^* = -\sum_{k \neq i} x_k. \end{cases} \tag{12}$$

It's obvious $A^T \lambda = x$. From Slater's constraint qualification, the strong duality holds, and the solution of target problem 1 is $z^* = v - \kappa A^T \lambda^* = v - \kappa x^*$.