

Notation & Setup

April 28, 2015

Hey, Let's use the following notation in our writeups and code ?

The first primal objective is,

$$F_1(\alpha) = \frac{1}{2n} \|Y - \sum_{j=1}^M K^{(j)} \alpha^{(j)}\|_2^2 + \frac{\lambda}{M} \sum_{j=1}^M \sqrt{\alpha^{(j)\top} K^{(j)} \alpha^{(j)}}$$

Let the Cholesky decomposition of $K^{(j)}$ be, $K^{(j)} = L^{(j)} L^{(j)\top}$ where $L^{(j)}$ is lower triangular¹. Let $\beta_j = L^{(j)\top} \alpha^{(j)}$. Then, we can write the above objective as,

$$F_2(\beta) = \frac{1}{2n} \|Y - \sum_{j=1}^M L^{(j)} \beta^{(j)}\|_2^2 + \frac{\lambda}{M} \sum_{j=1}^M \|\beta^{(j)}\|_2$$

Lets call this the second primal objective.

Coding

Lets use the following standards, its basically what we are using right now.

- Maintain α as an $\mathbb{R}^{n \times M}$ matrix. Each column corresponds to an $\alpha^{(j)}$.
- Maintain all the Kernel matrices in a $\mathbb{R}^{n \times n \times M}$ tensor K , where $K(:, :, j)$ is $K^{(j)}$.
- Same for all the $L^{(j)}$'s.
- Lets use n to denote the number of data points and upper case M to denote the number of groups – even in the code.

¹Matlab's `chol` gives the upper triangular matrix, but my wrapper `stableCholesky` takes care of it. It also, takes care of numerical issues in PSD matrices but be careful when using a non-PSD matrix with it.