Multiple kernel Learning, version2

Govinda Surampudi, Avinash Sharma, Dipanjan Roy

April 12, 2016

Abstract

In this version, it is hypothesised that

1. all ROIs have same set of diffusion scales.
2. for each scale of diffusion, each ROI diffuses at different amplitude.

1 Mathematical formulation

Given a fixed set of scales for all ROIs, the setup learns a set of α 's. Instead of horizontally stacking all the kernels together in a single matrix and trying to find α matrix, whose many entries are known to be 0, explicit linear combination of kernels weighted with their corresponding $\alpha^{(i)l}$'s is used for approximating **FC** matrices.
Let 'n' be the number of ROIs, and there be 'p' diffusion kernels for each of the 'L' subjects(samples).

$$\mathbf{E} = \| \sum_{l=1}^{L} \left\{ \sum_{i=1}^{p} \mathbf{H}_i^l \alpha^{(i)l} - \mathbf{FC}^l \right\} \|_F^2 \qquad \forall i \in (1, ..., p)$$

$$\mathbf{E} = \| \sum_{l=1}^{L} \left\{ \sum_{j=1}^{n} \left(\sum_{i=1}^{p} h_{ij}^l \alpha_j^{(i)l} - f_j^l \right) \right\} \|_F^2 \quad \forall j \in (1, ..., n)$$

$$\mathbf{E} = \| \sum_{j=1}^{n} \left\{ \sum_{l=1}^{L} \mathbf{X}_j^l \hat{\alpha}_j^l - f_j^l \right\} \|_F^2 \qquad \forall l \in (1, ..., L)$$

$$\mathbf{E} = \| \sum_{j=1}^{n} \mathbf{\Psi}_j \hat{\alpha}_j - \mathbf{\Phi}_j \|_F^2$$

each $\hat{\alpha}_j$ is a least squares solution. Combine $\hat{\alpha}_j$'s into $\hat{\alpha}^i$'s.
where,

$$\alpha^{(i)l} = \begin{bmatrix} \alpha_1^{(i)l} & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \alpha_n^{(i)l} \end{bmatrix}_{n \times n}$$

$$\mathbf{H}_i^l = \begin{bmatrix} h_{i1}^l & \cdots & h_{in}^l \end{bmatrix}_{n \times n}, \mathbf{H}_i^l \alpha^{(i)l} = \begin{bmatrix} h_{i1}^l \alpha_1^{(i)l} & \cdots & h_{in}^l \alpha_n^{(i)l} \end{bmatrix}_{n \times n}$$

$$\mathbf{X}_j^l = \begin{bmatrix} h_{1j}^l & \cdots & h_{pj}^l \end{bmatrix}_{n \times p}$$

$$\mathbf{\Psi} = \begin{bmatrix} \mathbf{X}_j^1 \\ \vdots \\ \mathbf{X}_j^L \end{bmatrix}_{nL \times p}, \mathbf{\Psi} = \begin{bmatrix} f_j^1 \\ \vdots \\ f_j^L \end{bmatrix}_{nL \times 1}, \hat{\alpha}_j^l = \hat{\alpha}_j = \begin{bmatrix} \hat{\alpha}_j^{(1)} \\ \vdots \\ \hat{\alpha}_j^{(p)} \end{bmatrix}_{p \times 1} \quad \forall l \in (1, ..., L)$$