High Dimensional Functional Model

1 Model

Consider outcome Y_{ij} is the grey matter volume (WVOLUME variable in the data) , let X_{ij} be the τ protein (TAU), and Z_{ij} be the amyloid protein (AMYLOID) measured at $j = 1, \ldots, d$ brain regions of interest (ROI) $i = 1, \ldots, n$ subjects indexed by PIDN. Let $\mathbf{W}_{ij}$ be the q dimensional baseline covariate. Our model is

$$Y_{ij} = \beta_j(Z_{ij})X_{ij} + \boldsymbol{\alpha}^{\mathrm{T}}\mathbf{W}_{ij} + \epsilon_{ij},$$

where ϵ_{ij}, $i = 1, \ldots, n$, $j = 1, \ldots, d$, are independent mean zeros random errors. We assume $\beta_j(Z_{ij})$ is in the K dimensional functional spaces, where $K << d$. That is

$$\beta_j(\cdot) = \sum_{k=1}^{K} \eta_{kl}\kappa_k(\cdot). \tag{1}$$

We estimate parameter β_j and $\boldsymbol{\alpha}$ by using B-spline technique. We replace $\beta_j(\cdot)$ by a B-spline approximation $\mathbf{B}(t)^{\mathrm{T}}\boldsymbol{\gamma}_j$, where $\mathbf{B}(t)$ is a rth order B-spline basis with N knots. To satisfy the relation in (1), we assume $\boldsymbol{\Gamma} \equiv (\boldsymbol{\gamma}_j, j = 1, \ldots, d)^{\mathrm{T}}$ is a low rank matrix. Now we reduce the problem of estimating $\beta_j(\cdot)$ to the problem of recovering $\boldsymbol{\Gamma}$. Let $\mathbf{e}_j$ be the d dimensional unit vector with the jth entry to be one. We obtain the estimators for $\boldsymbol{\Gamma}, \boldsymbol{\alpha}$ through minimizing

$$\sum_{i=1}^{n}\sum_{j=1}^{d}\{Y_{ij} - \mathbf{e}_j^{\mathrm{T}}\boldsymbol{\Gamma}\mathbf{B}(Z_{ij})X_{ij} - \boldsymbol{\alpha}^{\mathrm{T}}\mathbf{W}_{ij}\}^2 + \lambda_L\|\boldsymbol{\Gamma}\|_* + \lambda_P\sum_{j=1}^{d}\mathbf{e}_j^{\mathrm{T}}\boldsymbol{\Gamma}\mathbf{D}\boldsymbol{\Gamma}^{\mathrm{T}}\mathbf{e}_j,$$

where $\mathbf{D}$ is the second order weights matrix for the penalized spline regression and $\|\cdot\|_*$ is the nuclear norm. One choice of $\mathbf{D}$ can be $\int \mathbf{B}''(s)\mathbf{B}''(s)^{\mathrm{T}}ds$, where $\mathbf{B}''$ is the secondary derivative of $\mathbf{B}$.

2 Algorithm

Let's first disregard $\boldsymbol{\alpha}^{\mathrm{T}}\mathbf{W}_{ij}$ and solve

$$
\begin{aligned}
& \min \sum_{i=1}^{n}\sum_{j=1}^{d}\{Y_{ij} - \mathbf{e}_j^{\mathrm{T}}\boldsymbol{\Gamma}\mathbf{B}(Z_{ij})X_{ij}\}^2 + \lambda_L\|\boldsymbol{\Gamma}\|_* + \lambda_P\sum_{j=1}^{d}\mathbf{e}_j^{\mathrm{T}}\boldsymbol{\Gamma}\mathbf{D}\boldsymbol{\Gamma}^{\mathrm{T}}\mathbf{e}_j, \\
\Leftrightarrow \quad & \min \sum_{i=1}^{n}\sum_{j=1}^{d}\{Y_{ij} - \mathbf{e}_j^{\mathrm{T}}\boldsymbol{\Gamma}\mathbf{B}(Z_{ij})X_{ij}\}^2 + \lambda_L\|\boldsymbol{\Gamma}\|_* + \lambda_P \operatorname{trace}(\boldsymbol{\Gamma}\mathbf{D}\boldsymbol{\Gamma}^{\mathrm{T}}).
\end{aligned}
$$

Let $f(\boldsymbol{\Gamma}) = \sum_{i=1}^{n}\sum_{j=1}^{d}\{Y_{ij} - \mathbf{e}_j^{\mathrm{T}}\boldsymbol{\Gamma}\mathbf{B}(Z_{ij})X_{ij}\}^2 + \lambda_P \operatorname{trace}(\boldsymbol{\Gamma}\mathbf{D}\boldsymbol{\Gamma}^{\mathrm{T}})$ and $g(\boldsymbol{\Gamma}) = \lambda_L\|\boldsymbol{\Gamma}\|_*$. Since $f(\cdot)$ is smooth and $g(\cdot)$ is convex, we use proximal gradient method to solve the above minimization. Denote $\mathbf{B}(Z_{ij})X_{ij}$ by $\mathbf{v}_{ij}$, the gradient of $f(\boldsymbol{\Gamma})$ is

$$
\nabla f(\boldsymbol{\Gamma}) = \sum_{i=1}^{n}\sum_{j=1}^{d} 2(\mathbf{e}_j\mathbf{e}_j^{\mathrm{T}}\boldsymbol{\Gamma}\mathbf{v}_{ij}\mathbf{v}_{ij}^{\mathrm{T}} - Y_{ij}\mathbf{e}_j\mathbf{v}_{ij}^{\mathrm{T}}) + \lambda_P(\boldsymbol{\Gamma}\mathbf{D}^{\mathrm{T}} + \boldsymbol{\Gamma}\mathbf{D}).
$$

Then we can estimate $\widehat{\boldsymbol{\Gamma}}$ as follows.

Algorithm 1 Proximal gradient estimation

---

1. Given initial guess $\boldsymbol{\Gamma}^0 = \mathbf{0}$, a threshold ϵ, λ_L, and λ_P.
2. For $t \geq 0$, repeat

$$
\begin{aligned}
\boldsymbol{\Gamma}^{t+1} &= \mathbf{prox}_{\alpha^t}(\boldsymbol{\Gamma}^t - \alpha^t\nabla f(\boldsymbol{\Gamma}^t)) \\
&= \underset{\boldsymbol{\Gamma}}{\operatorname{argmin}} \frac{1}{2\alpha^t}\|\boldsymbol{\Gamma} - (\boldsymbol{\Gamma}^t - \alpha^t\nabla f(\boldsymbol{\Gamma}^t))\|_F^2 + \lambda_L\|\boldsymbol{\Gamma}\|_*. \qquad (2)
\end{aligned}
$$

3. If $\|\boldsymbol{\Gamma}^{t+1} - \boldsymbol{\Gamma}^t\|_F / \|\boldsymbol{\Gamma}^t\|_F \leq \epsilon$, stop; else, go to Step 2.

α^t is the step size, which can be chosen by a constant or by backtracking line search. The minimization in (2) can be solved by soft-thresholding SVD. Let $\mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^{\mathrm{T}}$ be the SVD of the matrix $\boldsymbol{\Gamma}^t - \alpha^t\nabla f(\boldsymbol{\Gamma}^t)$, then the solution of (2) is given by $\boldsymbol{\Gamma}^{t+1} = \mathbf{U}\boldsymbol{\Sigma}_{\alpha^t\lambda_L}\mathbf{V}^{\mathrm{T}}$, where $\boldsymbol{\Sigma}_{\alpha^t\lambda_L} = \operatorname{diag}\{(\Sigma_{i,i} - \alpha^t\lambda_L)_+\}$.