"Tensor Regression with Applications in Neuroimaging Data Analysis"

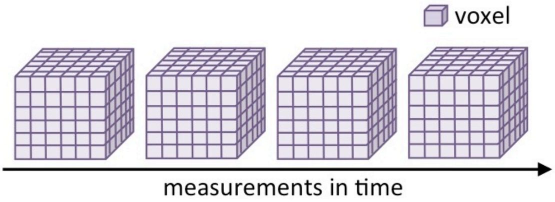
Hua Zhou, Lexin Li, & Hongtu Zhu

Jennifer Kirk

May 7, 2013

Review

- Analysis of neuroimage data is important to mental health
- fMRI data: 4-D array (tensor) with spatial and temporal correlation
- Current methods reduce the dimensions of the data, ignoring the correlation
- **New Method**: Extend GLM to use **fMRI image as one covariate observation** in regression model



One fMRI Observation from One Subject

Special Case: Matrix Covariates

- Outcome $Y_i \sim$ univariate exponential family
- Vector covariate: $\mathbf{z}_i$
- Matrix covariate: $\mathbf{X_i}$ (p $\times$ q)
- Link function:

$$g(\mu_i) = \alpha + \gamma^{\mathsf{T}} \mathbf{z_i} + \underbrace{\begin{bmatrix} \beta_{11} \\ \beta_{12} \\ \vdots \\ \beta_{1p} \end{bmatrix}^{\mathsf{T}}}_{1 \times p} \underbrace{\begin{bmatrix} x_{i11} & x_{i12} & \cdots & x_{i1q} \\ x_{i2,1} & x_{i22} & \cdots & x_{i2q} \\ \vdots & \vdots & \ddots & \vdots \\ x_{ip1} & x_{ip2} & \cdots & x_{ipq} \end{bmatrix}}_{p \times q} \underbrace{\begin{bmatrix} \beta_{21} \\ \beta_{22} \\ \vdots \\ \beta_{2q} \end{bmatrix}}_{q \times 1}$$

Tensor Math

- Order: the number of indices need to describe the tensor
- Kronecker Product: A is $m \times p$, B is $n \times q$:

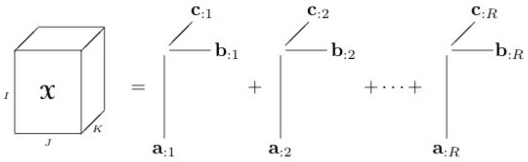
$$\mathbf{A} \otimes \mathbf{B}_{mn \times pq} \equiv \begin{bmatrix} a_{1,1}\mathbf{B} & a_{1,2}\mathbf{B} & \cdots & a_{1,p}\mathbf{B} \\ a_{2,1}\mathbf{B} & a_{2,2}\mathbf{B} & \cdots & a_{2,p}\mathbf{B} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m,1}\mathbf{B} & a_{m,2}\mathbf{B} & \cdots & a_{m,p}\mathbf{B} \end{bmatrix}$$

- Khatri-Rao Product: A is $m \times p$, B is $n \times p$:

$$\mathbf{A} \odot \mathbf{B}_{mn \times p} \equiv [\mathbf{a}_{\cdot 1} \otimes \mathbf{b}_{\cdot 1} \cdots \mathbf{a}_{\cdot p} \otimes \mathbf{b}_{\cdot p}]$$

Rank-R Decomposition

- If $\mathbf{X}$ is an $I \times J \times K$ (order 3) tensor and $\mathbf{A}_{I \times R}$, $\mathbf{B}_{J \times R}$, $\mathbf{C}_{K \times R}$ are matrices then $\mathbf{X} = [\![\mathbf{A}, \mathbf{B}, \mathbf{C}]\!]$ means

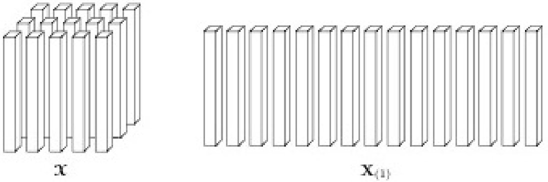


- If $\mathbf{X}$ is an $I_1 \times \ldots \times I_D$ (order D) tensor, then the rank-R decomposition is

$$\mathbf{X} = [\![\mathbf{A}_1, \mathbf{A}_2, \ldots, \mathbf{A}_D]\!] = \sum_{r=1}^{R} \mathbf{a}_1^{(r)} \circ \ldots \circ \mathbf{a}_D^{(r)}$$

Mode-d Matricization

- Denoted $\mathbf{X}_{(d)}$
- $\mathbf{X}$ is an $I \times J \times K$ (order 3) tensor then $\mathbf{X}_{(1)}$:



- In general, we "spread out" the tensor, keeping the d^{th} dimension, to get a matrix

Special Case: Matrix Covariates

Recall:

- $\mathbf{X_i}$ is a $p \times q$ matrix
- β_1^T is a $1 \times p$ vector
- β_2 is a $q \times 1$ vector

$$
\begin{aligned}
g(\mu_i) &= \alpha + \gamma^\mathsf{T}\mathbf{Z_i} + \beta_1^\mathsf{T}\mathbf{X_i}\beta_2 \\
&= \alpha + \gamma^\mathsf{T}\mathbf{Z_i} + (\beta_2 \odot \beta_1)^\mathsf{T}\mathrm{vec}(\mathbf{X_i}) \\
&= \alpha + \gamma^\mathsf{T}\mathbf{Z_i} + \langle(\beta_2 \odot \beta_1), \mathrm{vec}(\mathbf{X_i})\rangle
\end{aligned}
$$

where $(\beta_2 \odot \beta_1)$ is a $pq \times 1$ vector, $\langle\cdot\rangle$ is the inner product, and $\mathrm{vec}(\mathbf{X_i})$ is a vector made from $\mathbf{X_i}$

Rank-R Generalized Linear Tensor Regression

- Outcome: $Y_i \sim$ univariate exponential family
- Vector covariate: $\mathbf{z_i}$
- Tensor covariate: $\mathbf{X_i}$ (Order D: $I_1 \times \ldots \times I_D$)
- Assume tensor $\mathbf{B}$ has a rank-R decomposition

$$[\![\mathbf{B}_1, \ldots, \mathbf{B}_D]\!]$$

  where $\mathbf{B}_d$ is $I_d \times R$ matrix
- Link function:

$$g(\mu_i) = \alpha + \gamma^{\mathsf{T}} \mathbf{z}_i + \langle (\mathbf{B}_D \odot \ldots \odot \mathbf{B}_1) \mathbf{1}_R, \mathrm{vec}(\mathbf{X_i}) \rangle$$

Rank-R Generalized Linear Tensor Regression

- $\mathbf{X_i}$ Order D: $I_1 \times \ldots \times I_D$
- $\mathbf{B} = [\![\mathbf{B}_1, \ldots, \mathbf{B}_D]\!]$
- $\mathbf{B}_d$ is $I_d \times R$ matrix

$$
\begin{aligned}
g(\mu_i) &= \alpha + \gamma^{\mathsf{T}} \mathbf{z}_i + \langle (\mathbf{B}_D \odot \ldots \odot \mathbf{B}_1) \mathbf{1}_R, \mathrm{vec}(\mathbf{X_i}) \rangle \\
&= \alpha + \gamma^{\mathsf{T}} \mathbf{z}_i + \langle \mathbf{B}_d, \mathbf{X}_{\mathbf{i}(d)} (\mathbf{B}_D \odot \ldots \odot \mathbf{B}_{d+1} \odot \mathbf{B}_{d-1} \odot \ldots \odot \mathbf{B}_1) \rangle
\end{aligned}
$$

Parameter Estimation

- Maximum Likelihood Estimation
- Estimation Algorithm
  1. Set $\mathbf{B}^{(0)} = 0$ & estimate $\hat{\alpha}^{(0)}$, $\hat{\gamma}^{(0)}$
  2. Set $\alpha = \hat{\alpha}^{(n-1)}$, $\gamma = \hat{\gamma}^{(n-1)}$ & for each $\mathbf{B}_d$:
     - Set $\mathbf{B}_k = \hat{\mathbf{B}}_k^{(n)}$, $k < d$
     - Set $\mathbf{B}_k = \hat{\mathbf{B}}_k^{(n-1)}$, $k > d$
     - Estimate $\hat{\mathbf{B}}_d$
  3. Estimate $\hat{\alpha}^{(n)}$ and $\hat{\gamma}^{(n)}$, assuming $\mathbf{B}_d = \hat{\mathbf{B}}_d^{(n)}$ for all d
  4. Iterate 2–3 until the likelihood converges

$$g(\mu_i) = \alpha + \gamma^{\mathsf{T}}\mathbf{z}_i + \langle \mathbf{B}_d, \mathbf{X}_{\mathbf{i}(d)}(\mathbf{B}_D \odot \ldots \odot \mathbf{B}_{d+1} \odot \mathbf{B}_{d-1} \odot \ldots \odot \mathbf{B}_1) \rangle$$

Summary

- Extend GLM framework to tensor covariates
- Assume tensor parameter has a simple form
- Exploited algebraic properties of this simple form to estimate the parameter easily

Completed Work & Next Steps

- Completed Work
  - Understand model & estimation algorithm
  - Implemented simple model
  - Ran small scale simulation
- Next Steps
  - Speed up estimation algorithm
  - Generalize estimation algorithm
  - Run Simulation # 1: simple matrix data
  - Run Simulation # 2: real data, simulated outcomes
  - Analyze real data