Tensor LS Regression - STAT676 Project Proposal (Revised)

Zixiang Xu Ben Seiyon Lee

zxu27@gmu.edu slee287@gmu.edu

1. Background and Research Question

Recently, after reading the tensor regression part of the paper: *Tensor in Modern Statistical Learning*, I thought a simple tensor regression maybe a good project topic. So I Googled and looked at the *Tensor Computation for Data Analysis* book and decided to take this LS tensor regression as my final project.

The main purpose of this project is to:

- Generate tensor random variables under different settings;
- Perform simple least square tensor regression. And see which setting(s) would be the best assumption for the simple LS tensor regression.

2. Planned Methods and Models

2.1. Generate tensor random variables under different settings

The simple tensor regression model is defined as:

$$\mathcal{Y} = \mathcal{X} \times_1 B^{(1)} \cdots \times_M B^{(M)} \times_{M+1} \mathbf{I}_N + \mathcal{E},$$

where $\mathcal{X}$ would be generated under different settings (eg. multi-normal, uniform etc.), and $\mathcal{Y}$ would be generated based on the relation:

$$\mathcal{Y} = \mathcal{X} \times_1 A^{(1)} \times_2 A^{(2)} \ldots \times_N A^{(N)} \iff Y_{(n)} = A^{(n)} X_{(n)} (A^{(N)} \otimes \ldots \otimes A^{(n+1)} \otimes A^{(n-1)} \otimes \ldots \otimes A^{(1)})^T.$$

2.2. Perform simple least square tensor regression

After generating data samples $\{\mathcal{X}, \mathcal{Y}\}$'s, the corresponding optimization problem would be:

$$\min_{B^{(1)}, \cdots, B^{(M)}} ||\mathcal{Y} - \mathcal{X} \times_1 B^{(1)} \cdots \times_M B^{(M)} \times_{M+1} \mathbf{I}_N||_F^2.$$

Combining the relation mentioned in section 2.1, the above optimization problem would be reduced to iteratively solving the optimization problem:

$$\min_{B^{(m)}} ||Y_{(m)} - B^{(m)} X_{(m)} (B^{(N)} \otimes \ldots \otimes B^{(m+1)} \otimes B^{(m-1)} \otimes \ldots \otimes B^{(1)})^T||_F^2.$$

3. Expected Result

I am not sure which setting of $\mathcal{X}$ might be the best for this simple LS model. Maybe any multi-dimension distribution with linear relationship would be the choice.