Review of "Efficient Multidimensional Functional Data Analysis Using Marginal Product Basis Systems"

Jiaxin Hu

January 24, 2022

The proposed method aims to study the functional representation of tensors under the marginal product basis assumption. The main theory indicates the equivalence between the representation study with truncated basis and the CP tensor decomposition. An alternating algorithm to solve the corresponding regularized tensor decomposition is proposed. The asymptotic results for the global estimator are also provided.

Comments:

1. The simulation results in Figure 1 and 2 are inconsistent with the asymptotic results in Section 3. Theorem 3.2 indicates the convergence rate of the optimal estimator is of $\mathcal{O}(N^{-1/2})$. Then, we will expect an improvement of order $\mathcal{O}(1/3)$ in the error with sample size from $N = 5$ to $N = 50$. However, MARGARITA does not indicate such improvement under the cases close to the true settings; i.e., $K_t = 20, m_d = 11$ with degree of freedom $K_t \sum_{d=1}^{3} m_d = 660$ (if I understand correctly; $D = 3$ is not clearly stated in the first experiment setup). Though the variances over replications significantly decrease with $N = 50$, MARGARITA shows the improvement in error only under the overparameterization cases with degree of freedom around 2000 or $K_{fit} = 25, m_d = 25$. Therefore, the algorithm output may be sub-optimal and far away from the global estimator solution $\check{\zeta}^*_{N,m,k}$ in Theorem 2.3. The algorithm output may not enjoy the good asymptotic results in Section 3. More discussions for the algorithm accuracy will strengthen the paper.

2. Further, the asymptotic results in Section 3.2 focus only on properties of the global estimator without regularization. The effects of the regularization to the accuracy are still unknown. Discussions about the global estimator with regularization may be helpful to understand the algorithm performance.

3. General guidance for the choices of basis ϕ, basis m_d, tensor rank K, and tuning parameter λ_d should be provided. The simulation results imply that the choices of parameters, especially m_d and K, have significant effects to the estimation and generalization.

4. In Figure 1, TPB is always fitted with larger degree of freedom than MARGARITA; i.e., $m_d^3 \geq K_{fit} \sum_{d=1}^{3} m_d$. However, the gray and white boxes do not always appear in pairs. For example, there are two consecutive white boxes with degrees of freedom 600 and 675. More explanations will be helpful.