Review of "Generalized Low-rank plus Sparse Tensor Estimation by Fast Riemannian Optimizaiton"

Jiaxin Hu

May 12, 2021

This paper proposes a generalized framework to estimate a tensor with low-rank plus sparse structure. A fast algorithm based on Riemannian optimization is proposed, and corresponding theorems are provided. As far as I am concerned, the main contributions of the paper include (1) incorporating the Riemannian optimization and gradient prunign techniques to estimate the general low-rank plus sparse tensor, and (2) establishing the non-asymptotic error bounds for the algorithm outputs. Here are my comments.

Introduction

1. Gu et al. (2014) and Zhou and Feng (2017) also tackle the similar robust tensor estimation problems with low-rank plus sparse structure. It would be interesting to discuss the difference and similarity with the purposed method.

Numerical Experiment and real data analysis

1. According to the paper, the advantages for using low-rank plus sparse structure include (1) making the model more robust to model misspecification and outliers, and (2) capturing the useful heterogenous signals. I believe more experiments which compare the purposed model with other models can better emphasize these mentioned advantages. For example, compare the estimation accuracy as well as computing time with other tensor models only assume low-rankness or sparsity. Similar comparisons can be implemented with International Commodity data. For example, compare the PCA clustering results of low-rank model (HOSVD) with the proposed method.

2. Gu et al. (2014) and Zhou and Feng (2017) apply the low-rank plus sparse structure but with different algorithms. It would be interesting to see the comparison with these methods to highlight the benefits to use Riemannian optimization.

3. In the analysis of International Commodity data, Figure 3 implies that the low-rank estimate is sensitive to the sparsity α. The paper implies that the economic similarity gradually dominates the geographical relations as α increases. However, it is hard to understand the relationship between the sparsity α and economic similarity. It would be easier to understand the clustering shifts with more explanation to the meaning of α and its relationship with economic similarity.

4. By the simulation results in section 6, the estimation error is sensitive to the tuning parameter γ. In addition, the main Theorem 4.1 suggests that the stepsize β should be in a proper range. I believe more guidance on the selection of γ and β would be useful for practical applications.

5. In addition, I think more simulations which check the accuracy for the sparse tensor $\mathcal{S}$ would benefit the understanding of the theorems.

References

Gu, Q., Gui, H., and Han, J. (2014). Robust tensor decomposition with gross corruption. *Advances in Neural Information Processing Systems*, 27:1422–1430.

Zhou, P. and Feng, J. (2017). Outlier-robust tensor pca. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 2263–2271.