MULTI-TASK FMRI DATA FUSION USING IVA AND PARAFAC2

ISABELL LEHMANN

The interest in data fusion, i.e., the joint analysis of multiple related datasets, has grown in recent years in various research areas, in particular, in biomedicine. Data-driven methods, especially methods based on joint matrix/tensor factorizations, have shown to be effective for data fusion [1]. Two of them are Independent Vector Analysis (IVA) and PARAFAC2. IVA [2] is an extension of Independent Component Analysis (ICA) to multiple datasets, and a good candidate for data fusion because it makes use of the dependence across datasets. The PARAFAC2 model [3] also has proved useful for jointly analyzing datasets as a more flexible version of the well-known CANDECOMP/PARAFAC tensor method.

With the goal of identifying novel biomarkers for complex neurological disorders, fusion of medical imaging data has received particular attention. Especially important is multi-task functional Magnetic Resonance Imaging (fMRI) data, i.e., data collected from the same subjects while they are performing different tasks. Since different tasks provide complementary information about the brain, analyzing the joint information between tasks may help to better understand these disorders.

In this talk, we study IVA and PARAFAC2 for data fusion [4], first through simulations, where multiple datasets in the form of *subjects* by *voxels* matrices correspond to different tasks. Our simulations reveal that both methods can accurately capture the underlying latent components, albeit with certain differences in capturing the corresponding subject scores. We then apply both methods for the analysis of 13 fMRI datasets from the MCIC collection [5], collected from 271 subjects that perform 3 different tasks with well-defined relationship among them. Both methods are able to achieve two important goals at once, namely capturing group differences between patients with schizophrenia and healthy controls with interpretable components, as well as understanding the relationship across the tasks.

This is joint work with Evrim Acar (Simulamet, Norway), Tanuj Hasija (Paderborn University), M.A.B.S. Akhonda (UMBC), Vince D. Calhoun (TReNDS Center), Peter J. Schreier (Paderborn University), and Tülay Adali (UMBC).

This work was supported in part by NSF grants CCF 1618551, NCS 1631838 and NIH grants R01MH123610 and R01MH118695, the RCN project 300489, and the DFG grant SCHR 1384/3-2. The used hardware is part of the UMBC HPCF.

References

[1] D. Lahat, T. Adali, and C. Jutten, "Multimodal data fusion: an overview of methods, challenges, and prospects," *Proceedings of the IEEE*, vol. 103, no. 9, pp. 1449–1477, 2015.

[2] T. Adali, M. Anderson, and G. S. Fu, "Diversity in independent component and vector analyses: Identifiability, algorithms, and applications in medical imaging," *IEEE Signal Processing Magazine*, vol. 31, no. 3, pp. 18–33, 2014.

[3] R. A. Harshman, "PARAFAC2: Mathematical and technical notes," *UCLA working papers in phonetics*, vol. 22, no. 10, pp. 30–44, 1972.

[4] I. Lehmann, E. Acar, et al., "Multi-task fMRI data fusion using IVA and PARAFAC2," in *2022 IEEE International Conference on Acoustics, Speech and Signal Processing*, 2022.

[5] R. L. Gollub, J. M. Shoemaker, et al., "The MCIC collection: a shared repository of multi-modal, multi-site brain image data from a clinical investigation of schizophrenia," *Neuroinformatics*, vol. 11, no. 3, pp. 367–388, 2013.

Paderborn University, Germany
Email address: isabell.lehmann@sst.upb.de