All-at-once Optimization for Mining Higher-order Tensors

In many disciplines, we deal with data that has more than two axes of variation. For instance, multi-channel EEG (electroencephalogram) signals have spatial, spectral and temporal dimensions. We represent such data as higher-order tensors and use tensor factorizations, i.e., higher-order analogues of matrix factorizations, to extract the underlying patterns in the data. However, information extraction is often challenging due to large-scale and incomplete data sets, i.e., data with missing or unknown values. Besides, we are commonly faced with data from multiple sources and need to fuse data sets to capture the underlying structures.

In order to address these issues, we introduce a gradient-based optimization framework for factorizing higher-order tensors that can easily extend to the analysis of incomplete tensors as well as joint analysis of multiple data sets. In particular, we are interested in fitting the tensor model called CANDECOMP/PARAFAC (CP), which expresses a tensor as the sum of rank-one tensors and has been widely used in various disciplines including chemometrics, signal processing, computational neuroscience and social network analysis. The traditional CP algorithm is an alternating least squares algorithm, which is computationally very efficient but not robust in the case of overfactoring, i.e., extracting more factors than the true underlying factors. Alternative algorithms shown to be robust to overfactoring do not scale to large-scale data sets. The proposed gradient-based approach uses all-at-once optimization and solves for all CP factor matrices simultaneously. Using numerical experiments, we demonstrate that it is both accurate and scalable. We also show how the proposed approach extends to fit a CP model to incomplete data and joint analysis of multiple data sets. We further demonstrate the usefulness of the proposed framework on several biomedical applications.