Probability Mass Estimation in Tensors with Hidden Elements Using Structure

Martin Haardt
Ilmenau University of Technology
Communications Research Laboratory
P.O. Box 10 05 65, D-98684 Ilmenau, Germany
Homepage: http://tu-ilmenau.de/crl
E-Mail: Martin.Haardt@tu-ilmenau.de

Estimating the joint probability mass function (PMF) of a set of random variables from partially observed data is a crucial part of statistical learning and data analysis, with applications in areas such as recommender systems and data classification. Since the PMF takes the form of a multi-way tensor, under certain model assumptions the problem becomes closely associated with tensor factorization. In recent studies, it was shown that a low-rank PMF tensor can be fully recovered (under some mild conditions) by applying a low-rank (approximate) joint factorization to all estimated joint PMFs of subsets of fixed cardinality larger than two (e.g., triplets). The joint factorization is based on a Least Squares (LS) fit to the estimated lower-order sub-tensors. In this plenary talk, we take a different estimation approach by fitting the partial factorization directly to the observed partial data in the sense of Kullback-Leibler divergence (KLD). Consequently, we avoid the need for a particular selection and direct estimation of sub-tensors of a particular order, as we inherently apply proper weighting to all the available partial data. We show that our approach essentially attains the Maximum Likelihood estimate of the full PMF tensor (under the low-rank model) and therefore enjoys its well-known properties of consistency and asymptotic efficiency.

Furthermore, based on the Bayesian model interpretation of the low-rank model, we propose an Estimation-Maximization (EM) based approach, which is computationally cheap per iteration. To this end, we have developed a hybrid alternating-directions expectation-maximization (AD-EM) algorithm to solve the ML optimization problem, consisting of computationally more expensive AD iterations followed by an EM refinement stage. It is well known that the convergence rate of EM decreases as the fraction of missing data increases. In this talk, we also address the slow convergence of the EM algorithm. By adapting the squared iterative methods (SQUAREM) acceleration scheme to the context of PMF estimation, we propose the SQUAREM-PMF algorithm to speed up the convergence of the EM algorithm. Moreover, we demonstrate that running the computationally cheaper EM algorithm alone is sufficient after an appropriate initialization. Numerical results on both synthetic and real data in the context of movie recommendation show that our algorithm outperforms state-of-the-art PMF estimation algorithms.