Guest editorial: special issue on data mining with matrices, graphs and tensors

Tao Li · Chris Ding · Fei Wang

Published online: 2 March 2011
© The Author(s) 2011

1 Data mining with matrices, graphs and tensors

The field of data mining increasingly adapts methods and algorithms from advanced matrix computations, graph theory and optimization. In these methods, the data is described using matrix representations (graphs are represented by their adjacency matrices) and the data mining problem is formulated as an optimization problem with matrix variables. With these, the data mining task becomes a process of minimizing or maximizing a desired objective function of matrix variables. Prominent examples include spectral clustering, non-negative matrix factorization, Principal Component Analysis (PCA) and Singular Value Decomposition (SVD) related clustering and dimension reduction, tensor analysis and L-1 regularization.

These *matrix-formulated optimization-centric* methodologies are rapidly becoming a significant part of data mining, and have evolved into a popular and rapidly-expanding research area for solving challenging data mining problems. First, they are amenable to vigorous analysis and benefit from the well-established knowledge in linear algebra, graph theory and optimization accumulated through centuries. Second, they can be efficiently computed thanks to the mature software tools developed by scientific computing communities. Third, they are simple to implement and easy to

T. Li (✉)

School of Computer Science, Florida International University, Miami, FL, USA
e-mail: taoli@cs.fiu.edu

C. Ding

CSE Department, University of Texas at Arlington, Arlington, TX, USA
e-mail: chqding@uta.edu

F. Wang

Healthcare Transformation Group, IBM T. J. Watson Research Center, Yorktown Heights, NY, USA
e-mail: fwang@us.ibm.com

understand, in comparison to probabilistic, information-theoretic, and other methods. In addition, they are well-suited to parallel and distributed processing for solving large scale challenging problems such as searching and extracting patterns from the entire Web. Last but not the least, these methodologies are quite flexible and they can be used to formulate a large number of data mining tasks (e.g., classification, clustering, semi-supervised learning, feature selection, dimensionality reduction).

2 The special issue

Recently, we started a number of activities to establish a platform for exchanging ideas in the field of data mining using matrices, graphs and tensors. We have successfully organized four workshops in related areas over the last 3 years:

- KDD 2008 workshop on data mining using matrices and tensors (DMMT'08),
- KDD 2009 workshop on data mining using matrices and tensors (DMMT'09),
- ICDM 2009 workshop on optimization-based methods for emerging data mining problems (OEDM'09),
- ICDM 2010 workshop on optimization-based methods for emerging data mining problems (OEDM'10).

These activities culminated in the special issue. This special issue provides a leading forum for timely, in-depth presentation of recent advances in algorithms, theory and applications in data mining using matrices, graphs and tensors. The selected papers underwent a rigorous refereeing and revision process. Here we provide a brief overview of the papers included in the special issue:

1. Dimensionality reduction, i.e., finding a succinct and effective representation for high dimensional data sets, is one of the primary data mining tasks. The paper by Zhou, Tao and Wu proposes the manifold elastic net (MEN) for dimensionality reduction which integrates manifold learning with sparse learning to obtain a sparse low-dimensional projection. MEN directly imposes the elastic net penalty (i.e., the combination of the lasso penalty and the L_2 -norm penalty) over the loss of a discriminative manifold learning algorithm. In particular, MEN can be formulated as a lasso penalized least square problem and the least angle regression (LARS) is applied to obtain the optimal sparse projection matrix.
2. Recently several 2D or higher-order PCA-style dimensionality reduction algorithms have been developed for feature extraction from two-dimensional or higher-order data, such as face images and surveillance videos. However, most of these methods lack probabilistic interpretations and it is also difficult to apply them in real applications. The paper by Yu, Bi and Ye proposes a probabilistic dimensionality reduction framework for 2D and higher-order data and introduces a family of probabilistic dimensionality reduction models called PHOPCA for matrix-variate data, including both 2D data as well as higher-order data. PHOPCA explicitly specifies the generative process of matrix-variate objects, and their proposed optimization framework recovers the optimal solutions of several matrix-variate PCA-style algorithms under mild conditions.

3. Collaborative filtering recommends items based on similarity measures between users and/or items. One of the challenging questions in collaborative filtering is how to measure similarity of users and/or items. The paper by Cao, Yang, Sun and Chen proposes a matrix factorization model for learning user and item similarities simultaneously for collaborative filtering based on adaptive bidirectional asymmetric similarity measurement. Their proposed model automatically learns asymmetric similarities between users and items at the same time using matrix factorization.
4. The paper by Zafeiriou and Petrou studies Nonnegative Tensor Factorization (NTF) based on the Kullback–Leibler (KL) divergence. In particular, they explore the properties of the optimization problem, investigate the convergence properties, and calculate generalized pythagorean rules for KL divergence by formulating the problem as an alternative Csiszar–Tusnady procedure. They also study the relation between the proposed algorithms with probabilistic tensor latent component analysis and provide probabilistic interpretations for the algorithms.
5. The paper by Cui, Liu, Sun, and Yang proposes a hierarchical visual event pattern mining approach and utilize the patterns to address the key problems in video mining and understanding. In their work, they propose a tensor-based video representation and Joint Matrix Factorization (JMF) for unsupervised primitive event categorization.
6. Detecting community structures in complex networks such as social networks and the Word-Wide Web is an important task in mining network data. The paper by Wang, Li, Wang, Zhu and Ding investigates the problem of extending nonnegative matrix factorization (NMF) for community discovery. They propose three concrete algorithms, Symmetric NMF, Asymmetric NMF and Joint NMF to work on undirected networks, directed networks and compound networks.

Acknowledgments We are very indebted to the reviewers who reviewed the papers very carefully. We would also like to thank all the authors who submitted their papers to the special issue. Special thanks to Dr. Geoff Webb for his great help and support in organizing the issue.