DOCTORAL THESIS

Low rank tensor decomposition for feature extraction and tensor recovery

Shi, Qiquan

Date of Award:
2018

Abstract

Feature extraction and tensor recovery problems are important yet challenging, particularly for multi-dimensional data with missing values and/or noise. Low-rank tensor decomposition approaches are widely used for solving these problems. This thesis focuses on three common tensor decompositions (CP, Tucker and t-SVD) and develops a set of decomposition-based approaches. The proposed methods aim to extract low-dimensional features from complete/incomplete data and recover tensors given partial and/or grossly corrupted observations.

Based on CP decomposition, semi-orthogonal multilinear principal component analysis (SO-MPCA) seeks a tensor-to-vector projection that maximizes the captured variance with the orthogonality constraint imposed in only one mode, and it further integrates the relaxed start strategy (SO-MPCA-RS) to achieve better feature extraction performance. To directly obtain the features from incomplete data, low-rank CP and Tucker decomposition with feature variance maximization (TDVM-CP and TDVM-Tucker) are proposed. TDVM methods explore the relationship among tensor samples via feature variance maximization, while estimating the missing entries via low-rank CP and Tucker approximation, leading to informative features extracted directly from partial observations. TDVM-CP extracts low-dimensional vector features viewing the weight vectors as features and TDVM-Tucker learns low-dimensional tensor features viewing the core tensors as features. TDVM methods can be generalized to other variants based on other tensor decompositions. On the other hand, this thesis solves the missing data problem by introducing low-rank matrix/tensor completion methods, and also contributes to

automatic rank estimation. Rank-one matrix decomposition coupled with L1-norm regularization (L1MC) addresses the matrix rank estimation problem. With the correct estimated rank, L1MC refines its model without L1-norm regularization (L1MC-RF) and achieve optimal recovery results given enough observations. In addition, CP-based nuclear norm regularized orthogonal CP decomposition (TREL1) solves the challenging CP- and Tucker-rank estimation problems. The estimated rank can improve the tensor completion accuracy of existing decomposition-based methods. Furthermore, tensor singular value decomposition (t-SVD) combined with tensor nuclear norm (TNN) regularization (ARE_{TNN}) provides automatic tubal-rank estimation. With the accurate tubal-rank determination, ARE_{TNN} relaxes its model without the TNN constraint (TC-ARE) and results in optimal tensor completion under mild conditions. In addition, ARE_{TNN} refines its model by explicitly utilizing its determined tubal-rank a priori and then successfully recovers low-rank tensors based on incomplete and/or grossly corrupted observations (RTC-ARE: robust tensor completion/RTPCA-ARE: robust tensor principal component analysis).

Experiments and evaluations are presented and analyzed using synthetic data and real-world images/videos in machine learning, computer vision, and data mining applications. For feature extraction, the experimental results of face and gait recognition show that SO-MPCA-RS achieves the best overall performance compared with competing algorithms, and its relaxed start strategy is also effective for other CP-based PCA methods. In the applications of face recognition, object/action classification, and face/gait clustering, TDVM methods not only stably yield similar good results under various multi-block missing settings and different parameters in general, but also outperform the competing methods with significant improvements. For matrix/tensor rank estimation and recovery, L1MC-RF efficiently estimates the true rank and exactly recovers the incomplete images/videos under mild conditions, and outperforms the state-of-the-art algorithms on the whole. Furthermore, the empirical evaluations show that TREL1 correctly determines the CP-/Tucker- ranks well, given sufficient observed entries, which consistently improves the recovery performance of existing decomposition-based tensor completion. The t-SVD recovery methods TC-ARE, RTPCA-ARE, and RTC-ARE not only inherit the ability of $\mathrm{ARE}_{\mathrm{TNN}}$ to achieve accurate rank estimation, but also achieve good performance in the tasks of (robust) image/video completion, video denoising, and background modeling. This outperforms the state-of-the-art methods in all cases we have tried so far with significant improvements.

Keywords: Low-rank tensor decomposition, CP, Tucker, t-SVD, feature extraction, automatic rank estimation, tensor completion, robust tensor recovery, missing data.

Table of Contents