# Projective Unsupervised Flexible Embedding with Optimal Graph

Wei Wang[1]
wei.wang@unitn.it
Yan Yan[1]
yan.yan@unitn.it
Feiping Nie[2]
feipingnie@gmail.com
Xavier Alameda Pineda[1]
xavier.alamedapineda@unitn.it
Shuicheng Yan[3]
eleyans@nus.edu.sg
Nicu Sebe[1]
niculae.sebe@unitn.it

[1] DISI,
University of Trento
Trento, Italy

[2] OPTIMAL,
Northwestern Polytechnical University,
Xi'an, China

[3] ECE,
National University of Singapore,
Singapore

---

Graph based dimensionality reduction techniques have been successfully applied to clustering and classification tasks. The fundamental basis of these algorithms is the constructed graph which dominates their performance. Usually, the graph is defined by the input affinity matrix. However, the affinity matrix is sub-optimal for dimension reduction as there is much noise in the data. To address this issue, we propose the projective unsupervised flexible embedding with optimal graph (PUFE-OG) model. We build an optimal graph by adjusting the affinity matrix. To tackle the out-of-sample problem, we employ a linear regression term to learn a projection matrix. The optimal graph and projection matrix are jointly learned by integrating the manifold regularizer and regression residual into a unified model. An efficient algorithm is derived to solve the challenging model. The experimental results on several public benchmark datasets demonstrate that the presented PUFE-OG outperforms other state-of-the-art methods.

When the labels are available, the most popular dimensionality reduction algorithm is linear discriminant analysis (LDA). It has excellent performance as LDA utilizes discriminant information to learn the subspace. In addition, simultaneously performing clustering and subspace learning can yield even better clustering result. [3] proposed an effective discriminating K-Means (DisKmeans) algorithm by integrating LDA and K-Means. However, labeled data are often very costly to obtain.

When the labels are unavailable, unsupervised dimensionality reduction methods become the only choice. For example, PCA is widely used because of its simplicity and efficiency. The unsupervised graph based dimensionality reduction methods usually outperform PCA. This is because these methods take advantage of manifold information. Many graph based dimensionality reduction methods have been explored, such as locally linear embedding (LLE), Laplacian eigenmap (LE), and ISOMAP. However, these methods suffer from out-of-sample problem. They can not map the new data points that are not included in the training set. To tackle this problem, many works [1] integrated the manifold regularizer with the ridge regression loss into the subspace learning framework. Similar to other manifold learning algorithms, their performance is also controlled by the graph constructed by the fixed affinity matrix, which might lead to a sub-optimal result [2]. To address this issue, we propose a projective unsupervised flexible embedding with optimal graph (PUFE-OG) framework. Instead of utilizing the fixed affinity matrix to preserve the manifold structure, we construct an optimal graph by adapting the affinity matrix for subspace learning.

[1] Feiping Nie, Dong Xu, Ivor Wai-Hung Tsang, and Changshui Zhang. Flexible manifold embedding: A framework for semi-supervised and unsupervised dimension reduction. *TIP*, 2010.

[2] Feiping Nie, Xiaoqian Wang, and Heng Huang. Clustering and projected clustering with adaptive neighbors. In *SIGKDD*. ACM, 2014.

[3] Jieping Ye, Zheng Zhao, and Mingrui Wu. Discriminative k-means for clustering. In *NIPS*, 2008.