Information Sciences 424 (2018) 27–38

Contents lists available at ScienceDirect

Information Sciences

journal homepage: www.elsevier.com/locate/ins

Subspace clustering using a low-rank constrained autoencoder

Yuanyuan Chen, Lei Zhang, Zhang Yi*

Machine Intelligence Laboratory, College of Computer Science, Sichuan University, Chengdu 610065, PR China

ARTICLE INFO

Article history:
Received 28 March 2017
Revised 14 September 2017
Accepted 16 September 2017
Available online 18 September 2017

Keywords:
Deep neural networks
Subspace clustering
Autoencoder
Low-rank representation

ABSTRACT

The performance of subspace clustering is affected by data representation. Data representation for subspace clustering maps data from the original space into another space with the property of better separability. Many data representation methods have been developed in recent years. Typical among them are low-rank representation (LRR) and an autoencoder. LRR is a linear representation method that captures the global structure of data with low-rank constraint. Alternatively, an autoencoder nonlinearly maps data into a latent space using a neural network by minimizing the difference between the reconstruction and input. To combine the advantages of an LRR (globality) and autoencoder (self-supervision based locality), we propose a novel data representation method for subspace clustering. The proposed method, called low-rank constrained autoencoder (LRAE), forces the latent representation of the neural network to be of low rank, and the low-rank constraint is computed as a prior from the input space. One major advantage of the LRAE is that the learned data representation not only maintains the local features of the data, but also preserves the underlying low-rank global structure. Extensive experiments on several datasets for subspace clustering were conducted. They demonstrated that the proposed LRAE substantially outperformed state-of-the-art subspace clustering methods.

1. Introduction

Subspace clustering seeks to determine clusters in different subspaces within a dataset, and has been widely used in many scientific and industrial applications, including the clustering of images containing different objects, segmentation of different objects in a video sequence, and separation of a mixture of speech. Formally, let $X = [x_1, x_2, \cdots, x_N]$ be a $d \times N$ data matrix that consists of N vectors that are assumed to belong to K clusters. Subspace clustering determines to which cluster each datum is assigned. Among different subspace clustering methods, spectral clustering is an important branch for managing high-dimensional data.

One key step in the spectral clustering algorithm is to construct an affinity matrix $C \in R^{N \times N}$, where C_{jk} quantizes the similarity between data points j and k. A popular measurement of similarity is a function of some distance. For example, in [21], the similarity between two points is defined by $C_{jk} = \exp\left\{-s_{jk}^2\right\}$, where s_{jk} is a value of the distance between j and k. After obtaining the affinity matrix C, the segmentation of the data is performed by applying the k-means algorithm to the largest k eigenvectors of a normalized Laplacian matrix.

Different from [21], ℓ_1-graph [4], L2-graph [25], and sparse subspace clustering (SSC) [6], low-rank representation (LRR) [12,13,16,17] learns a representation by constructing the affinity matrix using the linear reconstruction coefficient. The major

* Corresponding author.
E-mail address: zhangyi@scu.edu.cn (Z. Yi).

http://dx.doi.org/10.1016/j.ins.2017.09.047
0020-0255/

difference between the approaches is the constraint on the coefficient, which describes the prior on the data distribution. Although these methods are state of the art in the area of subspace clustering, we observe that they suffer from the following limitations: First, these methods have to use the entire dataset as a dictionary; thus, they cannot manage large scale and incremental data [26]. Second, these methods assume that the data can be linearly represented with respect to each other. In this case, subspace clustering could be solved properly using convex methods. However, in reality, the relations between data are usually highly nonlinear or "nonlinear in linear" [11,14,23,27]. The aforementioned methods might achieve degraded performance. Third, spectral clustering based approaches inherit the disadvantages of manifold learning because they can be the same as performing manifold learning to obtain a data representation and conducting k-means clustering on the representation to achieve clustering membership. The disadvantages of manifold learning, such as poor robustness and a smooth sampling requirement, further limit applications of subspace clustering methods in the scenario of big data.

As another popular data representation method, a deep autoencoder [9] overcomes the disadvantages of manifold learning. A deep autoencoder progressively maps input data into a latent space by minimizing the difference between the reconstruction and input using a parametric neural network. Unlike manifold learning, an autoencoder performs in parallel and has a fast inference speed, which enables it to manage large scale and incremental data. It is a data-driven approach and does not depend on the linear assumption. Furthermore, the deep architecture [22] can capture the nonlinear complex structure of big data because of the nonlinear activation function and multi-layered structure.

Motivated by the substantial success of deep learning [9,20,33,37], we propose a novel subspace clustering method that combines the advantages of traditional LRR models and the popular deep model. The proposed method, called a low-rank constrained autoencoder (LRAE), is a variant of an autoencoder. Unlike the standard autoencoder and LRR, the LRAE not only enforces the latent representation that can reconstruct input well, but also requires the set of latent representations to be of low rank, where the low-rank constraint is computed from the input space. To summarize, the LRAE integrates the locality of an autoencoder and globality of LRR to learn a good data representation, and thus achieve better clustering results. To the best of our knowledge, this is the first work to incorporate low-rank constraint into an autoencoder. The most closely related work to our proposed method PARTY which was proposed in [24]. However, the LRAE and PARTY are quite different. In details, PARTY assumes that the latent representation can be sparsely reconstructed, whereas the LRAE adopts a low-rank prior. In practice, a low-rank prior is better than a sparsity prior because the former can guarantee the globality of the entire dataset, whereas the latter cannot. Moreover, the objective functions of PARTY and the LRAE are different. PARTY aims to solve an ℓ_1-minimization problem to achieve sparsity, whereas our method achieves a low-rank with a nuclear-norm based constraint. Benefiting from the intrinsic property of our objective function, our method is much more efficient than PARTY.

The remainder of this paper is organized as follows: In Section 2, we briefly discuss related work. In Section 3, we present the LRAE in detail, including the network architecture, objective function, and learning algorithm. We report the results of experiments and make some discussion in Section 4. We present our conclusions in Section 5.

2. Related work

In this section, we briefly discuss existing work on LRR and autoencoders.

2.1. Low-rank representation

In [13], Liu et al. proposed an LRR to determine the coefficient matrix C of data X with the lowest rank:

$$\min_C rank(C) \quad \text{s.t.} \quad X = AC, \tag{1}$$

where A is a dictionary of the data. If the data in X are arranged to satisfy the true segmentation of the data, the optimal solution to (1), C^*, is block diagonal.

$$C^* = \begin{bmatrix} C_1^* & 0 & \cdots & 0 \\ 0 & C_2^* & \cdots & 0 \\ 0 & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & C_K^* \end{bmatrix}, \tag{2}$$

which captures both dense within-class affinities and zero between-class affinities. Compared with sparse representation, low-rank representation of data captures the global structure of the data. In [14,15], Liu established a deeper theoretical analysis of LRR. Based on Liu's main idea, a great number of researchers have attempted to search a low-rank matrix to reveal the structure of the data or improve the efficiency of the original algorithm [3,7,18,32].

By solving a nuclear norm minimization problem, Favaro [7] proposed a closed-form solution to the low-rank minimization problem:

$$\min_C \|C\|_* + \frac{\lambda}{2}\|X - XC\|_F^2. \tag{3}$$