Information Sciences 340–341 (2016) 144–158

Contents lists available at ScienceDirect

Information Sciences

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

Robust subspace segmentation via nonconvex low rank representation

Wei Jiang [a,c,*], Jing Liu [a], Heng Qi [b], Qionghai Dai [c]

[a] School of Mathematics, Liaoning Normal University, Dalian 116029, China
[b] Electronic Information and Electrical Engineering Department, Dalian University of Technology, Dalian 116024, China
[c] Department of Automation, Tsinghua National Laboratory for Information Science and Technology (TNList), Tsinghua University, Beijing 100084, China

ARTICLE INFO

Article history:
Received 3 March 2015
Revised 14 December 2015
Accepted 25 December 2015
Available online 7 January 2016

Keywords:
Low rank representation
Subspace segmentation
Nuclear norm minimization
Ky Fan p-k-norm
$\ell_{2,q}$-norm

ABSTRACT

Recently, low rank representation (LRR) has been successfully applied to explore subspace segmentation of data. In this paper, we propose a nonconvex formulation to determine the LRR from contaminated data. Unlike in traditional methods, which directly utilize the nuclear norm to approximate the rank function and penalize noise using the $\ell_{2,1}$-norm, our method introduces the Ky Fan p-k-norm and the $\ell_{2,q}$-norm, to better approximate the rank minimization problem and enhance the robustness against noise. An efficient algorithm is derived for solving the novel objective function, and this is followed by a rigorous theoretical proof of the convergence. Extensive experiments on face datasets clearly demonstrate that the proposed methods are more robust to illumination variations, corruptions, and occlusions.

1. Introduction

In recent years, the subspace model has become an important research topic, with widespread applications in computer vision and machine learning, such as face recognition [41,44], image segmentation [22,31], and motion segmentation [35,47]. Traditional principal component analysis (PCA) [18] and the recently developed robust principal component analysis (RPCA) [3], matrix completion [19], and recovery [42] methods are reliant on the hypothesis that data samples are drawn from only one single low rank subspace. However, many real world datasets reside on multiple subspaces, and the subspaces that data samples belong to can be unknown. For example, for handwritten digits, each digit forms its own subspace in the feature space. For human faces, the face of the same person under different conditions lies on the same subspace, and those of different people are associated with different subspaces. This naturally leads to a more reasonable model, in which data are approximately drawn from a union of subspaces. This model presents a more challenging problem concerning subspace segmentation, with an aim of segmenting data samples into clusters, where each cluster corresponds to a subspace.

Over the past two decades, a number of subspace segmentation methods have been developed. According to the survey in [39], the existing methods for subspace segmentation may be roughly categorized into four main families: algebraic methods [7,39], iterative methods [1,40], statistical methods [38,50], and spectral clustering-based methods [10,30,31]. Among these, spectral clustering-based methods are particularly effective, and achieve competitive results. A fundamental problem

* Corresponding author at: School of Mathematics, Liaoning Normal University, Dalian 116029, China. Tel.: +86 41184258356.
E-mail address: swxxjw@aliyun.com (W. Jiang).

http://dx.doi.org/10.1016/j.ins.2015.12.038
0020-0255/

regarding spectral clustering-based methods is that of how to determine an effective affinity graph from the observed data, which represents the underlying data structures. According to Wright et al. [43], an informative affinity graph should exhibit three characteristics: high discriminating power, low sparsity, and an adaptive neighborhood. Based on this insight, many sparse representation (SR) based graph construction methods have been proposed [5,15,37,48]. However, these methods usually compute the sparsest representation of each point individually, and there is no global structure constraint of data. Therefore, such methods may be ineffective in capturing the global structure of data. This characteristic can result in the degradation of subspace segmentation when data are grossly corrupted.

Alongside recent progress in compressive sensing, a new concept of nuclear norm optimization has emerged in the field of rank minimization, and has resulted in a number of interesting applications, such as low rank representation (LRR) [23,24] from corruptions. LRR aims to decompose the original mixed data matrix X into $XZ + E$, where Z is the affinity matrix that describes the correlations between different pairs of data, and E is the associated sparse noise. More precisely, to derive the LRR of the input data matrix X LRR minimizes the rank of the matrix Z while reducing the $\ell_{2,0}$-norm of E. Because LRR involves the rank of the data matrix and the $\ell_{2,0}$-norm minimization problem, it is NP-hard, and thus it is difficult to solve. To address this issue, LRR employs the convex nuclear norm as a surrogate of the rank function, and the convex $\ell_{2,1}$-norm as a surrogate of the $\ell_{2,0}$-norm. A fast implementation for the LRR method is presented in [21], using iteratively linearizing methods.

The concept of manifolds [34] has been widely applied in statistical learning and pattern recognition. Related methods have been applied to solve a variety of machine learning and computer vision problems, such as face synthesis [45], dimensionality reduction [34], motion trajectory [51], and sample selection [46]. The goal of LRR is to determine the lowest rank representation of the data with an approximate dictionary. Motivated by the success of LRR and manifold learning, the graph regularized LRR scheme [29] has been proposed, which involves imposing a local geometrical structure on the traditional LRR. The smooth representation (SMR) model [16] has been proposed to explicitly explore grouping effects. It has been demonstrated that these methods deliver a good general performance for a wide variety of problems, such as face recognition, motion segmentation, and hyperspectral image destriping.

It is a well-known fact that the nuclear norm is the sum of all singular values of a matrix, while the matrix rank is the number of nonzero singular values of a matrix in which each singular value contributes equally. The nuclear norm minimization over-penalizes large singular values, resulting in a biased result. Hence, a solution obtained using the nuclear norm may be suboptimal, because it is not a perfect approximation of the matrix rank. A similar phenomenon has been observed for the convex ℓ_1-norm and nonconvex ℓ_0-norm for feature selection [2]. Thus, the convex method may not achieve effective performances in real applications. In order to achieve a better approximation for the ℓ_0-norm, many nonconvex surrogate functions of ℓ_0-norm have been proposed, including the ℓ_p-norm [12], smoothly clipped absolute deviation [11], logarithm [13], and minimax concave penalty [49]. Some of these nonconvex penalties have been extended to approximate the rank function, including the Schatten p-norm ($0 < p < 1$) used by Nie et al. [32], the truncated nuclear norm proposed by Hu et al. [17], the log-sum heuristic used by Deng et al. [8], the Ky Fan 2-k-norm proposed by Doan and Vavasis [9], and the joint Schatten p-norm and $\ell_{2,q}$-norm regularized LRR problem proposed by Lu et al. [26].

In this paper, we propose a novel LRR algorithm, described as robust subspace segmentation via nonconvex LRR (NLRR), which is designed to replace the rank of a matrix by the Ky Fan p-k-norm and penalize sparse noise via the $\ell_{2,q}$-norm. We investigate two widely employed nonconvex terms, and provide a proximal iteratively reweighted algorithm (PIRA) to solve the nonconvex model. Furthermore, theoretical justifications are provided to prove that the proposed algorithm converges to a stationary point. To the best of our knowledge, this is the first time that the nonconvex model has been generalized to LRR for subspace segmentation, the goal of which is to recover the underlying low rank structure of subspaces in the presence of illumination variations, noisy corruptions, and occlusions. Experimental results demonstrate that the nonconvex model achieves a higher clustering accuracy than other state-of-the-art algorithms, and the objective function can achieve convergence.

2. Related work

2.1. Low rank representation

We assume that $X = [x_1, x_2, \dots, x_n] \in \mathbb{R}^{m\times n}$ is an observed data matrix, whose columns consist of n data samples drawn from a union of k subspaces $\{S_i\}_{i=1}^k$. The LRR [24] can be formulated as the following minimization problem:

$$\begin{aligned}&\min_{Z,E} rank(Z) + \lambda\|E\|_{2,0},\\ &s.t. \quad X = AZ + E,\end{aligned} \tag{1}$$

where λ is a parameter balancing the two components, and A is the dictionary that linearly spans the union of subspaces $\bigcup_{i=1}^k S_i$. $Z \in \mathbb{R}^{m\times n}$ has a low rank structure, which is assumed to be the authentic structure of the observed data, and $E \in \mathbb{R}^{m\times n}$ is the noise matrix in the original data. $rank(Z)$ denotes the rank of the matrix Z, and the $\ell_{2,0}$-norm of matrix E is defined as $\|E\|_{2,0} = \sum_{i=1}^m \| \sum_{j=1}^n e_{ij}^2\|_0$, where for a scalar a, $\|a\|_0 = 1$ if $a \neq 0$ and $\|a\|_0 = 0$ if $\|a\|_0 = 0$.

Unfortunately, the above minimization problem is NP-hard in general, owing to the non-convexity and discontinuous nature of $rank(Z)$ and $\|E\|_{2,0}$. An effective strategy for approaching this problem is to relax the rank function to the convex