Locally linear representation for image clustering

Liangli Zhen, Zhang Yi, Xi Peng and Dezhong Peng

The construction of the similarity graph plays an essential role in a spectral clustering (SC) algorithm. There exist two popular schemes to construct a similarity graph, i.e. the pairwise distance-based scheme (PDS) and the linear representation-based scheme (LRS). It is notable that the above schemes suffered from some limitations and drawbacks, respectively. Specifically, the PDS is sensitive to noises and outliers, while the LRS may incorrectly select inter-subspaces points to represent the objective point. These drawbacks degrade the performance of the SC algorithms greatly. To overcome these problems, a novel scheme to construct the similarity graph is proposed, where the similarity computation among different data points depends on both their pairwise distances and the linear representation relationships. This proposed scheme, called locally linear representation (LLR), encodes each data point using a collection of data points that not only produce the minimal reconstruction error but also are close to the objective point, which makes it robust to noises and outliers, and avoids selecting inter-subspaces points to represent the objective point to a large extent.

Introduction: Spectral clustering (SC) is one of the most popular clustering algorithms, whose key is to build a similarity graph to describe the similarities among different data points [1]. In the graph, each vertex denotes a data point, and the edge weight between two vertexes represents the similarity of the corresponding data points. Currently, there are two schemes to calculate the similarity among data points, i.e. the pairwise distance-based scheme (PDS) and the linear representation-based scheme (LRS). The PDS computes the similarity between two points according to the distance between two points, e.g. Laplacian eigenmaps [2]. On the other hand, the LRS assumes that each data point could be denoted as a linear combination of some intra-subspace points [3]. On the basis of this observation, this scheme uses the linear representation coefficients as a measure of similarity. Recently, the LRS has attracted a lot of interest in the field of image clustering, since it captures the real structure of the data set better. Numerous clustering algorithms are developed based on the LRS, such as locally linear embedding (LLE) [4], sparse subspace clustering (SSC) [3] and low rank representation (LRR) [5].

It is notable that the above-mentioned similarity computation schemes suffer from some limitations. Specifically, the PDS is sensitive to noises and outliers, because it only depends on the distance between the two considered data points, and ignores the global structure of the whole data set. Fig. 1*a* illustrates the disadvantages of the PDS. On the other hand, the LRS has the possibility that a data point is represented as a linear combination of the inter-subspace data points. Fig. 1*b* shows the drawbacks of the LRS. SSC [3] and LRR [5] overcome this problem to some extent by bringing a sparsity constraint and a low-rank constraint into linear representation, but both of them are iterative algorithms with high computational complexity.

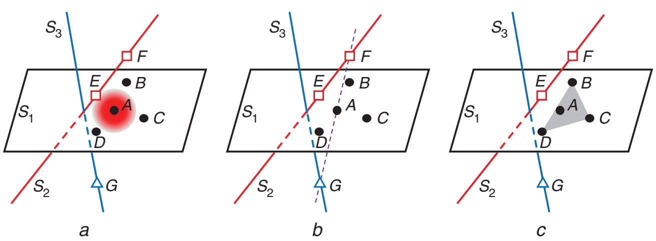

Fig. 1 *Key observation of geometric analysis of three different similarity graph construction strategies*

There are three subspaces S_1, S_2 and S_3 that lie in $\mathbb{R}^3$, where $\dim(S_1) = 2$, $\dim(S_2) = 1$ and $\dim(S_3) = 1$

Points A, B, C and D are drawn from S_1, points E, F from S_2, and point G from S_3

a PDS: Most similar point to A is E in terms of Euclidean distance (kind of PDS), but E is not in same cluster of A

b LDS: Most similar points to A are F and G in terms of linear representation-based similarity (i.e. LRS), because point A lies on line spanned by F and G

c Our method: Our method will select B, C and D as most similar points to A

Points B, C and D not only can represent A with minimal residual but are close to A. They will be divided into same cluster

To overcome the above-mentioned problems, this Letter presents a novel scheme to construct the similarity graph, where the similarity computation among different data points depends not only on their pairwise distances but also on mutually linear representation relationships. The proposed scheme, called locally linear representation (LLR), encodes each data point using a set of data points which produce the minimal error, and are close to the objective point. Our developed scheme is more robust to noises and outliers than the PDS. At the same time, compared with the LRS, it can effectively avoid selecting inter-subspaces points to represent the objective point. Moreover, the new scheme uses an analytic solution to construct the similarity graph, and has lower computational complexity than the iterative methods, such as the SSC and LRR.

Locally linear representation: Our basic idea was derived from a theoretical result in manifold learning that a topological manifold is a topological space which is locally homeomorphic to an Euclidean space [4]. It implies that in a subspace, mutually adjacent points can provide the linear representation for each other. This inspires us to construct the similarity graph by solving the following optimisation problem:

For each point $\boldsymbol{x}_i$, $i = 1, 2, \ldots, n$

$$\min_{\boldsymbol{c}_i} \lambda \,\|\boldsymbol{S}_i\boldsymbol{c}_i\|_2^2 + (1-\lambda)\,\|\boldsymbol{x}_i - \boldsymbol{D}_i\boldsymbol{c}_i\|_2^2 \text{ s.t. } \boldsymbol{1}^{\mathrm{T}}\boldsymbol{c}_i = 1 \quad (1)$$

where $\boldsymbol{D}_i = [\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_{i-1}, 0, \boldsymbol{x}_{i+1}, \ldots, \boldsymbol{x}_n]$ is a dictionary for $\boldsymbol{x}_i$; $\boldsymbol{S}_i$ is a diagonal matrix whose jth diagonal element is the pairwise distance between $\boldsymbol{x}_i$ and the jth data point in $\boldsymbol{D}_i$; $\boldsymbol{1} \in \mathbb{R}^m$ is a vector that consists of ones, $\lambda \in [0, 1)$ is a balance parameter and $\boldsymbol{c}_i \in \mathbb{R}^n$ is the representation coefficient of $\boldsymbol{x}_i$.

In the above problem, the first term makes $\boldsymbol{x}_i$ prefer to choose the nearby points to represent itself; and the second term makes it produce a minimal reconstruction error. Fig. 1*c* is a toy example showing the effectiveness of our approach.

By solving problem (1), it gives that

$$\boldsymbol{c}_i = \frac{\boldsymbol{M}_i^{-1}\boldsymbol{1}}{\boldsymbol{1}^{\mathrm{T}}\boldsymbol{M}_i^{-1}\boldsymbol{1}} \quad (2)$$

where $\boldsymbol{M}_i = \lambda\boldsymbol{S}_i^{\mathrm{T}}\boldsymbol{S}_i + (\boldsymbol{1}-\lambda)(\boldsymbol{x}_i\boldsymbol{1}^{\mathrm{T}} - \boldsymbol{D}_i)^{\mathrm{T}}(\boldsymbol{x}_i\boldsymbol{1}^{\mathrm{T}} - \boldsymbol{D}_i)$.

Note that the above solution is not sparse. It contains many trivial coefficients. This will increase the time cost of SC. By following [6], we obtain a sparse similarity graph by keeping k largest entries in $\boldsymbol{c}_i$ and setting the rest to zeros.

Once the similarity graph is built, we could apply the graph to the image clustering problem under the framework of SC [1, 7, 8]. Algorithm 1 summarises the whole procedure of our algorithm.

Algorithm 1. Learning LLR for SC

Input: A given data set $\boldsymbol{X} \in \mathbb{R}^{m\times n}$, balance parameter $\lambda \in [0, 1)$ and thresholding parameter (k).

1. For each point $\boldsymbol{x}_i \in \mathbb{R}^m$ ($i = 1, 2, \ldots, n$), calculate its representation coefficients $\boldsymbol{c}_i \in \mathbb{R}^n$ by solving
$$\min_{\boldsymbol{c}_i} \lambda \,\|\boldsymbol{S}_i\boldsymbol{c}_i\|_2^2 + (1-\lambda)\,\|\boldsymbol{x}_i - \boldsymbol{D}_i\boldsymbol{c}_i\|_2^2 \text{ s.t. } \boldsymbol{1}^{\mathrm{T}}\boldsymbol{c}_i = 1$$
2. Remove the trivial coefficients from $\boldsymbol{c}_i$ by performing hard thresholding operator, i.e. keeping k largest entries in $\boldsymbol{c}_i$ and zeroing all other elements.
3. Construct an undirected similarity graph via $\boldsymbol{W} = |\boldsymbol{C}| + |\boldsymbol{C}^{\mathrm{T}}|$.
4. Perform SC [8] over $\boldsymbol{W}$ to obtain the clustering membership.

Output: The clustering labels of the input data points.

Baselines and evaluation metrics: We ran the experiments over two widely used facial image data sets, i.e. the Extended Yale Database B [9] and the AR database [10]. The Extended Yale Database B contains 2014 near frontal face images of 38 individuals. The AR database contains 1400 face images without disguises distributed over 100 individuals (14 images for each subject). We downsized the images of the Extended Yale Database B from 192 × 168 to 48 × 42 and the AR images from 165 × 120 to 55 × 40. Moreover, as in [3, 5], principal component analysis is used as a pre-processing step by retaining 98% energy of the cropped images.

We compared LLR with several state-of-the-art algorithms, i.e. LRR [5], SSC [3], LLE-graph based clustering (LLEC) [4] and standard SC [8]. Moreover, we also tested the performance of k-means clustering as a baseline.

Two popular metrics, accuracy (AC) and normalised mutual information (NMI), are used to measure the clustering performance of these algorithms. The method works better, the value of AC or NMI being higher. In addition, the time cost for building similarity graph (t_1) and the whole time cost for clustering (t_2) are recorded to evaluate efficiency.

In each test, we tuned the parameters of all the methods to obtain their best AC. Briefly, LLR needs two user-specified parameters, balance parameter λ and thresholding parameter k. We set $\lambda \in \{0.001, 0.01, 0.1\}$ and $k \in \{3, 4, 5, 6\}$. Moreover, considering the computation efficiency, we only use 300 closest data points as dictionary $\boldsymbol{D}_i$ for each $\boldsymbol{x}_i$ in terms of Euclidean distance. For the other compared methods, we set the parameters by following [3–5, 8].

We report the clustering results of the evaluate algorithms in Tables 1 and 2, from which we have the following observations:

- LLR outperforms the other methods in AC and NMI by a considerable performance margin. LLR is 6.6% and 1.9% higher than the second best method (LRR) over AR in AC and NMI, respectively. The corresponding values are 17.0% and 15.0% over the Extended Yale Database B.
- LRR and SSC are two recently proposed algorithms, which are superior to LLEC and SC. Note that only SC is a pairwise distance-based SC method.
- LLR finds an elegant balance between time cost and clustering quality; it is not the fastest algorithm but achieves the best clustering quality.
- The k-means is the fastest algorithm, but performs the worst in AC and NMI.

Table 1: Performance comparisons of different methods over Extended Yale Database B

Metric	LLR	LRR [5]	SSC [3]	LLEC [4]	SC [8]	k-means
AC	**0.883**	0.713	0.613	0.461	0.426	0.098
NMI	**0.922**	0.772	0.684	0.540	0.539	0.115
t_1	14.628	38.095	159.665	0.678	0.264	–
t_2	102.256	90.8268	231.235	74.309	64.606	**4.543**

t_1 denotes CPU elapsed time (s) for building similarity graph and t_2 is whole time cost

Table 2: Performance comparisons of different methods over AR database

Metric	LLR	LRR [5]	SSC [3]	LLEC [4]	SC [8]	k-means
AC	**0.837**	0.771	0.767	0.396	0.361	0.311
NMI	**0.929**	0.910	0.886	0.682	0.652	0.611
t_1	8.696	30.495	164.327	0.318	0.147	–
t_2	111.618	128.343	286.978	107.779	113.918	**4.460**

Conclusion: Linear representation and pairwise distance are two popular methods to construct a similarity graph for SC. But both of them encountered some problems in practical applications. The pairwise distance-based method is sensitive to noise and outliers, while the linear representation-based method might fail when the data came from a union of dependent subspaces. In this Letter, we propose a new algorithm that represents the objective point $\boldsymbol{x}$ using some data points that not only can reconstruct $\boldsymbol{x}$ better but also are close to $\boldsymbol{x}$ in terms of pairwise distance. The incorporation of pairwise distance and linear representation largely improve the discrimination of the data model, which is beneficial to the clustering problem. Extensive experiments have verified the effectiveness and efficiency of our approach.

Acknowledgments: This work was supported by the National Basic Research Program of China (973 Program) under grant no. 2011CB302201, the Program for New Century Excellent Talents in University of China under grant no. NCET-12-0384 and the National Natural Science Foundation of China under grant no. 61172180.

24 February 2014
doi: 10.1049/el.2014.0666
One or more of the Figures in this Letter are available in colour online.

Liangli Zhen, Zhang Yi, Xi Peng and Dezhong Peng (*College of Computer Science, Sichuan University, No. 24 South Section 1 Yihuan Road, Chengdu 610065, People's Republic of China*)

E-mail: pengdz@scu.edu.cn

References

1 Cheng, B., Yang, J., Yan, S., Fu, Y., and Huang, T.: 'Learning with ℓ^1-graph for image analysis', *IEEE Trans. Image Process.*, 2010, **19**, (4), pp. 858–866
2 Belkin, M., and Niyogi, P.: 'Laplacian eigenmaps for dimensionality reduction and data representation', *Neural Comput.*, 2003, **15**, (6), pp. 1373–1396
3 Elhamifar, E., and Vidal, R.: 'Sparse subspace clustering: algorithm, theory, and applications', *IEEE Trans. Pattern Anal. Mach. Intell.*, 2013, **35**, (11), pp. 2765–2781
4 Roweis, S., and Saul, L.: 'Nonlinear dimensionality reduction by locally linear embedding', *Science*, 2000, **290**, (5500), pp. 2323–2326
5 Liu, G., Lin, Z., Yan, S., Sun, J., Yu, Y., and Ma, Y.: 'Robust recovery of subspace structures by low-rank representation', *IEEE Trans. Pattern Anal. Mach. Intell.*, 2013, **35**, (1), pp. 171–184
6 Peng, X., Zhang, L., and Zhang, Y.: 'Constructing L2-graph for subspace learning and segmentation', preprint arXiv:1209.0841v4, 2012
7 Shi, J., and Malik, J.: 'Normalized cuts and image segmentation', *IEEE Trans. Pattern Anal. Mach. Intell.*, 2000, **22**, (8), pp. 888–905
8 Ng, A.Y., Jordan, M.I., and Weiss, Y.: 'On spectral clustering: analysis and an algorithm', in (Eds): 'Advances in neural information processing systems (NIPS)', (2002), pp. 849–856
9 Georghiades, A., Belhumeur, P., and Kriegman, D.: 'From few to many: illumination cone models for face recognition under variable lighting and pose', *IEEE Trans. Pattern Anal. Mach. Intell.*, 2001, **23**, (6), pp. 643–660
10 Martinez, A., and Benavente, R.: 'The AR face database'. CVC Technical Report No. 24, 1998