Journal of Computer Science and Frontier Technologies
ISSN Print: 3104-4204
ISSN Online: 3104-4212
CD ISSN: 3104-4220
https://doi.org/10.63313/JCSFT.9008

# Graph Regularized Double Log-determinant Minimization for Subspace Clustering

**Yuang Deng**
Qingdao University, Qingdao 266000, China
Email: dya0718@163.com

**How to cite this paper:** Deng, Y. (2025). Graph Regularized Double Log-determinant Minimization for Subspace Clustering. Journal of Computer Science and Frontier Technologies, 1(1), 87-96. ISSN Print: 3104-4204, ISSN Online: 3104-4212.
https://doi.org/10.63313/JCSFT.9008
**Published: 2025-09-18**

Copyright © 2025 by author(s) and Erytis Publishing Limited.
This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).
http://creativecommons.org/licenses/by/4.0/

## Abstract

Subspace clustering methods have been successful in various applications such as face clustering. Existing methods usually adopt the l_1 or l_2-based norms to minimize the fitting error. However, the fitting residual constructed with such norms usually has structural information since these norms are computed in an element-wisely independent way and all elements of the residual are treated equally. To maximumly minimize structural information in residual so as to maximumly retain such information in recovered data, in this paper we propose to minimize fitting error of each example with log-determinant rank approximation. Extensive experiments verify the effectiveness of the proposed method.

## Keywords

Subspace Clustering; Machine learning; Log-determinant

## 1. Introduction

It has been increasingly demanding to handle high-dimensional data in various real world applications such as computer vision, image processing, pattern recognition, etc.Often times, high-dimensional data lie in a union of subspaces, which implies that they have low-dimensional structures.For example, face images of several individuals are high-dimensional, but they lie in several subspaces corresponding to each individual. This encourages us to recover low-dimensional structures of high-dimensional data for efficient representation.To recover such structures, it usually requires clustering data points into different groups such that each group is fitted with a latent subspace,which is known as subspace clustering[1]–[4].

During the last decades, a large number of subspace clustering methods have been developed, among which the spectral clustering-based ones have drawn significant attentions[1]– [5]. The spectral clustering-based subspace clustering methods usually assume that the data are self-expressed.That is, each example of the data can be represented by a low-dimensional vector with respect to the data itself as diction-

ary.Low-rank representation (LRR)[1],[6] and sparse subspace clustering (SSC) [4] are two typical ones, based on which various variants have been developed.LRR requires that the sought low-dimensional matrix to be low-rank while SSC enforces sparseness on the representation matrix by minimizing the nuclear or$\ell_1$norm of the representation matrix, respectively. Such constraints allow the representation matrix to have clear group structures.Recently, a number of subspace clustering methods have been developed based on LRR or SSC.For example, [7], [8] are variants of LRR and SSC, respectively, where the former recovers the low-rank representation with the most representative features while the latter seeks the low-dimensional structure in the latent feature space. Othern than minimizing the nuclear norm, SCLA [3]seeks a low-rank representation matrix of the data by minimizing log-determinant rank approximation (LDRA).The LRDA is a more accurate approximation of the rank than the nuclear norm, thus[3] seeks representation with more structural information than the original LRR with improved clustering performance.To enhance the capability of counting nonlinear structures,[9] seeks the new representation with manifold regularization.

The above methods, as well as other subspace clustering methods, adopt the $\ell_1$ or$\ell_2$ norm-based loss functions to minimize the reconstruction error. It is known that these norms are based on element-wise operations and treats all elements equally. This might be problematic when dealing with 2-dimensional data in applications such as face clustering. Indeed, existing works show that there is still structural information existing in residual images by using $\ell_1$or Frobenius norms.This implies that the recovered data do not maximumly retain structural information, which may potentially degrade the clustering performance.Moreover,[10] shows that it is inaccurate to measure the true distances of examples for face image data with$\ell_1$or$\ell_2$norms while spatial-aware norm such as the nuclear norm well reveals the true distances. These inspires us to adopt spatial-aware norms to measure the fitting residual in subspace clustering models.

To maximumally recover structural information from fitting residual and maximumally retain such information in the recovered data,in this paper we propose to adopt the LRDA to minimize the residual in a low-rank based model scheme. Moreover, to enhance the capability of recovering nonlinear relationships of the data, we recover the representation with manifold regularization.The optimization algorithm is guaranteed to converge.Extensive experiments confirm the effectiveness of the new method.

## 2. RELATED WORK

### 2.1. Low-rank Representation

Given the data matrix $X \in \mathcal{R}^{d\times n}$with$d$and$n$being the number of features and examples, respectively, the LRR seeks a low-rank representation $Z$of the data with

$$\min_{X,E} \| X - XZ \|_{2,1} + \lambda \| Z \|_*, \tag{1}$$

where $\| \cdot \|_{2,1}$ is the $\ell_{2,1}$ norm that adds $\ell_2$norms of all columns of a matrix, $\| \cdot \|_*$ is the nuclear norm that adds all singular values of a matrix, and $\lambda > 0$ is a balancing parameter. Then an affinity matrix is constructed with the representation matrix $Z$, on which the final clustering is performed.

### 2.2. Graph Laplacian

Graph Laplacian [11] is widely used to incorporate the intrinsic geometrical structure of the data, which enforces the smoothness of the data in linear and nonlinear spaces by minimizing

$$\begin{aligned} &\frac{1}{2}\sum_{i,j=1}^{n} \| v_i - v_j \|_2^2 w_{ij} = \sum_{j=1}^{n} d_{jj}\, v_j^T v_j - \sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij}\, v_i^T v_j, \\ &= \mathbf{Tr}(VDV^T) - \mathbf{Tr}(VWV^T) = \mathbf{Tr}(VLV^T), \end{aligned} \tag{2}$$

Where $\mathbf{Tr}(\cdot)$ is the trace operator, $W = [w_{ij}]$ is the weight matrix that measures the pair-wise similarities of original data points, $D = [d_{ij}]$ is a diagonal matrix with $d_{ii} = \sum_j w_{ij}$, and$L = D - W$ .

## 3. METHODOLOGY

Existing subspace clustering methods usually adopt the $\ell_2$or $\ell_1$-norm based loss function to minimize the reconstruction error.However, $\ell_2$and $\ell_1$ norms are based on element-wise operations, which rarely consider the structural information in the fitting error.As shown in original LRR and SSC works, rich structural information still exists in the fitting residual. This inspires us to develop alternative measurement to capture such information.To minimize the structural information in the residual, we propose to minimize the rank of residual matrix of each example, which leads to the following model:

$$\min_{Z} \sum_{i=1}^{n} \operatorname{rank}(X_i - \sum_{j=1}^{n} X_j Z_{ji}) + \lambda f(Z), \tag{3}$$

where $\lambda \geq 0$is a balancing parameter and$f(\cdot)$ is a proper function to restrict the structure of the representation matrix $Z$. The rank function is hard to solve and is usually replaced with some approximations for ease of optimization.Recent researches have shown that nonconvex approaches such as the log-determinant rank approximation (LDRA) are effective in approximating the rank function.For a matrix $M$, the log-determinant rank approximation is defined as $\log\det(M) = \log\det(I + M^T M) = \sum_i \log(1 + \sigma_i^2(M))$, where$\sigma_i(M)$ is the$i-$th singular value of $M$.The effectiveness of LDRA inspires us to replace the rank function in formula（3） with

LDRA,which leads to the following model:

$$\min_{Z} \sum_{i=1}^{n} \text{logdet}\,(X_i - \sum_{j=1}^{n} X_j Z_{ji}) + \lambda f(Z). \tag{4}$$

Without loss of generality, we follow the LRR and require low-rank structure for $Z$. We minimize the LRDA of $Z$ to guarantee the low-rank struture of $Z$, which leads to the following model:

$$\min_{Z} \sum_{i=1}^{n} \text{logdet}\,(X_i - \sum_{j=1}^{n} X_j Z_{ji}) + \lambda\, \text{logdet}(Z). \tag{5}$$

It is seen that (5) only considers linear relationships of the data.To further account nonlinear relationships of the data, we seek the low-rank representation matrix $Z$ on manifold, which leads to following model for subspace clustering:

$$\min_{Z} \sum_{i=1}^{n} \text{logdet}\,(X_i - \sum_{j=1}^{n} X_j Z_{ji}) + \lambda\, \text{logdet}(Z) + \gamma \mathbf{Tr}(Z^T L Z) \tag{6}$$

where $\gamma \geq 0$ is a balancing parameter and $L$ is the graph Laplacian matrix. We name (6) as graph regularized double log-determinant minimization (GLoD). We will develop an efficient optimization algorithm in the following section.

## 4. OPTIMIZATION

In this section, we will develop an optimization scheme based on method of augmented Lagrangian multiplier (ALM).For ease of notation, we denote the LDRA as $\|\cdot\|_{ld}$ .We first introduce auxilary variables $D_i$ with $i = 1, \cdots, n$ and $W$ to (6), which leads to

$$\begin{aligned} &\min_{Z} \sum_{i=1}^{n} \| D_i\|_{ld} + \lambda \|W\|_{ld} + \gamma \mathbf{Tr}(ZLZ^T), \\ &s.t. \quad D_i = X_i - \sum_{j=1}^{n} X_j Z_{ji}, W = Z. \end{aligned} \tag{7}$$

The corresponding augmented Lagrangian function is:

$$\begin{aligned} \mathcal{L} = &\sum_{i=1}^{n} \| D_i\|_{ld} + \lambda \|W\|_{ld} + \frac{\rho}{2} \|W - Z + \frac{1}{\rho}\Lambda\|_F^2 \\ &+ \gamma \mathbf{Tr}(ZLZ^T) + \frac{\rho}{2} \|D_i - X_i + \sum_{j=1}^{n} X_j Z_{ji} + \frac{1}{\rho}\Theta_i\|_F^2. \end{aligned} \tag{8}$$

Next, we will alternatively solve the above problem with respect to each variable while fixing the others. The detailed optimization strategy is discussed in the following.

### 4.1. $D_i$-minimization

The subproblem with respect to $D_i$ is

$$\min_{D_i} \sum_{i=1}^{n} \| D_i \|_{ld} + \frac{\rho}{2} \| D_i - X_i + \sum_{j=1}^{n} X_j Z_{ji} + \frac{1}{\rho} \Theta_i \|_F^2. \tag{9}$$

The above problem is in the standard form of [3] and can be solved with closed-form solution.For a matrix$M$ , we define an operator $\mathcal{T}_\tau(M)$ as

$$\mathcal{T}_\tau(M) = \mathcal{P}(M) \operatorname{diag}\{\sigma_i^*\}\mathcal{Q}(M),$$

with $\mathcal{P}(M)$ and $\mathcal{Q}(M)$ being the left and right singular vectors of $M$, and $\sigma_i^*$ being the solution to

$$arg \min_{\sigma_i \geq 0} (\sigma_i - \sigma_i^M)^2 + \tau \log(1 + \sigma_i^2),$$

where $\sigma_i^M$ is the i-th largest singular value of $M$. The detailed solution and analysis of the constrained minimization problem can be found in [3]. Due to space limit, we do not fully expand it here.According to [3], the $D_i$-associated subproblem can be solved with closed-form solution as follows:

$$D_i = \mathcal{T}_{\frac{1}{\rho}}(X_i - \sum_{j=1}^{n} X_j Z_{ji} - \frac{1}{\rho}\Theta_i). \tag{10}$$

### 4.2. $W$-minimization

The subproblem associated with $W$-minimization is

$$\min_W \lambda \|W\|_{ld} + \frac{\rho}{2} \|W - Z + \frac{1}{\rho}\Lambda\|_F^2, \tag{11}$$

which, similar to $D_i$-minimization, can be solved with the following operator:

$$W = \mathcal{T}_{\frac{\lambda}{\rho}}(Z - \frac{1}{\rho}\Lambda). \tag{12}$$

### 4.3. $Z$-minimization

The sub-optimization problem associated with $Z$ is

$$\min_{Z} \gamma \, \mathbf{Tr}(ZLZ^T) + \frac{\rho}{2}\|D_i - X_i + \sum_{j=1}^{n} X_j Z_{ji} + \frac{1}{\rho}\Theta_i\|_F^2 + \frac{\rho}{2}\|W - Z + \frac{1}{\rho}\Lambda\|_F^2 \tag{13}$$

The above problem is convex and quadratic in $Z$ , which admits its solution by the first-order optimality condition.We take the derivative of (13) and set it to zero, then we have

$$\gamma Z(L + L^T) + \rho(Z - W - \frac{1}{\rho}\Lambda) + \rho HZ = Z(2\gamma L + \rho I) + \rho HZ - \rho(W + \frac{1}{\rho}\Lambda) = 0, \tag{14}$$

where

$$H_{st} = \mathbf{Tr} X_s^T (D_t - X_t + \frac{1}{\rho}\Theta_t). \tag{15}$$

Thus, $Z$ can be obtained by solving the above sylvester equation, which is denoted as

$$Z = sylvester(2\gamma L + \rho I, \rho H, \rho W + \Lambda). \tag{16}$$

### 4.4. $\Lambda, \Theta_i$ – updating

The updating of $\Lambda$, $\Theta_i$, and $\rho$ are standard as follows:

$$\Lambda = \Lambda + \rho(W - Z) \tag{17}$$

$$\Theta_i = \Theta_i + \rho(D_i - X_i + \sum_{j=1}^{n} X_j Z_{ji}) \tag{18}$$

$$\rho = \rho\kappa \tag{19}$$

where$\kappa > 1$ is a parameter that keeps $\rho$ increasing at each iteraton.

## 5. SUBSPACE CLUSTERING VIA KDRR

After we solve (6), we follow a common post-processing step to construct an affinity matrix $\mathbf{A}$ from $Z$ for final clustering [3][6].Detailed steps to construct $\mathbf{A}$ is listed as follows:

1)Let $Z = U\Sigma V^T$ be the skinny SVD of $Z$. Define $\bar{Z} = U\Sigma^{1/2}$ to be the weighted column space of $Z$.

2)Obtain $\bar{U}$by normalizing each row of $\bar{Z}$.

3)Construct the affinity matrix $\mathbf{A}$ as $[\mathbf{A}]_{ij} = \left(|[\bar{U}\bar{U}^T]_{ij}|\right)^{\phi}$,where follow[1] and set $\phi = 4$ throughout this paper for fair comparison.

Subsequently, we perform Normalized Cut (NCut) [12]on $\mathbf{A}$ in a way similar to [3],[6],[13].

## 6. EXPERIMENT

In this section, we conduct extensive experiments to testify the effectiveness of the proposed method.In particular, we compare the GLoD with several state-of-the-art subspace clustering methods, including LRR[6], LapLRR [9], and SCLA[3]. We compare the methods on 3 widely used benchmark data sets including Yale, Jaffe, and Alphadigit.Clustering accuracy and purity are used as the metrics to evaluate the clustering performance, whose definition can be found in [14],[15].

To better evaluate the performance of all methods in comparison, we conduct experiments in a way similar to [3],[15].In particular, for each data set we consider the overall data as well as its subsets of different sizes in the experiment.For example, the Yale data set has 15 clusters.In the experiment, beside the overall data, we conduct experiments using its subsets of $N$clusters, where $N$ takes values of 3, 6, 9, and 12, respectively.For each $N$ value, there are multiple choices of subsets, among which we randomly select 10 and report the average performance.For each method, the best performance is reported by searching all possible combinations of parameters. In particular, all parameters are tuned in a grid-search scheme within the set {0.0001,0.001,0.01,0.1,1,10,100,1000,10000}. For LapLRR and GLoD, binary kernel is used to construct the graph Laplacian.We present the detailed clustering performance of all methods in Table. 1-3.

It is observed that the GLoD has the best clustering performance in both clustering accuracy and purity.

**Table 1.** C lustering Performance on Yale

| N | Accuracy | | | |
|---|---|---|---|---|
| | LRR | LapLRR | SCLA | GLoD |
| 3 | 0.5485±0.0942 | 0.6758±0.1221 | 0.5424±0.0692 | **0.7121±0.1118** |
| 6 | 0.4258±0.0723 | 0.5394±0.0458 | 0.4197±0.0423 | **0.5591±0.0564** |
| 9 | 0.3576±0.0344 | 0.5131±0.0497 | 0.3667±0.0393 | **0.5212±0.0524** |
| 12 | 0.3303±0.0243 | 0.4924±0.0277 | 0.3303±0.0224 | **0.5015±0.0247** |
| 15 | 0.3394 | **0.4848** | 0.3515 | 0.4727 |
| Average | 0.4003 | 0.5411 | 0.4021 | **0.5533** |
| N | Purity | | | |
| | LRR | LapLRR | SCLA | GLoD |
| 3 | 0.5515±0.0901 | 0.6879±0.1098 | 0.5455±0.0623 | **0.7152±0.1061** |
| 6 | 0.4409±0.0694 | 0.5500±0.0429 | 0.4333±0.0469 | **0.5682±0.0573** |
| 9 | 0.3798±0.0337 | 0.5212±0.0518 | 0.3889±0.0302 | **0.5364±0.0529** |
| 12 | 0.3530±0.0268 | 0.4955±0.0288 | 0.3492±0.0268 | **0.5136±0.0269** |
| 15 | 0.3697 | **0.4848** | 0.3758 | 0.4788 |
| Average | 0.4190 | 0.5479 | 0.4185 | **0.5624** |

**Table 2.** C lustering Performance on Jaffe

| N | Accuracy | | | |
|---|---|---|---|---|
| | LRR | LapLRR | SCLA | GLoD |
| 2 | 0.8444±0.1557 | **1.0000±0.0000** | 0.6885±0.1910 | **1.0000±0.0000** |
| 3 | 0.5987±0.1503 | 0.9984±0.0049 | 0.4882±0.0680 | **1.0000±0.0000** |
| 4 | 0.5150±0.0711 | 0.9604±0.0948 | 0.4377±0.0453 | **0.9988±0.0037** |

| N | Accuracy | | | |
|---|---|---|---|---|
| | LRR | LapLRR | SCLA | GLoD |
| 5 | 0.4809±0.0834 | 0.9113±0.1121 | 0.3927±0.0637 | **0.9990±0.0031** |
| 6 | 0.4313±0.0809 | 0.9153±0.0750 | 0.3729±0.0386 | **0.9976±0.0038** |
| 7 | 0.4126±0.0575 | 0.9620±0.0305 | 0.3457±0.0417 | **0.9987±0.0028** |
| 8 | 0.4039±0.0401 | 0.9417±0.0453 | 0.3420±0.0396 | **1.0000±0.0000** |
| 9 | 0.3701±0.0177 | 0.9671±0.0099 | 0.3231±0.0257 | **1.0000±0.0000** |
| 10 | 0.3568 | 0.9671 | 0.3286 | **1.0000** |
| Average | 0.4904 | 0.9582 | 0.4133 | **0.9994** |
| N | Accuracy | | | |
| | LRR | LapLRR | SCLA | GLoD |
| 2 | 0.8444±0.1557 | **1.0000±0.0000** | 0.6885±0.1910 | **1.0000±0.0000** |
| 3 | 0.5987±0.1503 | 0.9984±0.0049 | 0.4882±0.0680 | **1.0000±0.0000** |
| 4 | 0.5150±0.0711 | 0.9604±0.0948 | 0.4377±0.0453 | **0.9988±0.0037** |
| 5 | 0.4809±0.0834 | 0.9113±0.1121 | 0.3927±0.0637 | **0.9990±0.0031** |
| 6 | 0.4313±0.0809 | 0.9153±0.0750 | 0.3729±0.0386 | **0.9976±0.0038** |
| 7 | 0.4126±0.0575 | 0.9620±0.0305 | 0.3457±0.0417 | **0.9987±0.0028** |
| 8 | 0.4039±0.0401 | 0.9417±0.0453 | 0.3420±0.0396 | **1.0000±0.0000** |
| 9 | 0.3701±0.0177 | 0.9671±0.0099 | 0.3231±0.0257 | **1.0000±0.0000** |
| 10 | 0.3568 | 0.9671 | 0.3286 | **1.0000** |
| Average | 0.4904 | 0.9582 | 0.4133 | **0.9994** |

**Table 3.** C lustering Performance on Alphadigit

| N | Accuracy | | | |
|---|---|---|---|---|
| | LRR | LapLRR | SCLA | GLoD |
| 5 | 0.8697±0.0660 | 0.8318±0.0911 | 0.8723±0.0604 | **0.8754±0.0640** |
| 10 | 0.6751±0.0744 | 0.7105±0.0830 | 0.6982±0.0936 | **0.7121±0.0923** |
| 15 | 0.6140±0.0245 | 0.6306±0.0403 | 0.6289±0.0493 | **0.6381±0.0579** |
| 20 | 0.5953±0.0222 | **0.6137±0.0422** | 0.6050±0.0224 | 0.6021±0.0265 |
| 25 | 0.5447±0.0229 | 0.5562±0.0272 | 0.5439±0.0170 | 0.5631±0.0258 |
| 30 | 0.5213±0.0112 | 0.5170±0.0121 | 0.5091±0.0240 | **0.5263±0.0134** |
| 35 | 0.4933±0.0111 | **0.4970±0.0148** | 0.4931±0.0104 | **0.4960±0.0150** |
| 36 | 0.4922 | **0.5107** | 0.4936 | 0.4929 |
| Average | 0.6007 | 0.6084 | 0.6055 | **0.6132** |
| N | Purity | | | |
| | LRR | LapLRR | SCLA | GLoD |
| 5 | 0.8697±0.0660 | 0.8421±0.0759 | 0.8723±0.0604 | **0.8754±0.0640** |
| 10 | 0.6928±0.0662 | **0.7297±0.0685** | 0.7133±0.0876 | 0.7264±0.0829 |
| 15 | 0.6400±0.0213 | **0.6540±0.0377** | 0.6489±0.0441 | 0.6535±0.0524 |
| 20 | 0.6297±0.0214 | 0.6367±0.0348 | 0.6287±0.0208 | **0.6317±0.0284** |
| 25 | 0.5793±0.0228 | 0.5831±0.0248 | 0.5742±0.0163 | **0.5914±0.0203** |
| 30 | 0.5541±0.0071 | 0.5415±0.0120 | 0.5406±0.0209 | **0.5563±0.0126** |
| 35 | 0.5306±0.0099 | 0.5181±0.0108 | 0.5264±0.0077 | **0.5311±0.0153** |
| 36 | **0.5306** | 0.5292 | 0.5178 | 0.5228 |
| Average | 0.6284 | 0.6293 | 0.6278 | 0.6361 |

Specifically, LRR and SCLA have similar performance on Yale and Alphadigit data sets. On Jaffe data set, LRR shows better performance than SCLA. GLoD and LapLRR have the best performances on all three data sets.GLoD improves the clustering performance significantly on Yale and Jaffe data sets.On Jaffe data set, GLoG has almost 100% accuracy in all cases.It appears challenging to cluster face images of Yale data set since all methods have relatively low clustering accuracy.In this case, GLoD has the top performance in almost all cases and the improvement is promising for such a challenging task.These observations confirm the effectiveness of GLoD.

## 7. CONCLUSIONS

In this paper, we propose a novel subspace clustering method with double log-determinant rank minimization.Unlike existing subspace clustering methods that measure the fitting residual with $\ell_1$or $\ell_2$ norms,in this paper we propose to minimize the residual of each example with the LRDA.This renders the new model to maximumly minimize structural information in the residual such that structural information can be maximumly retained in the reconstructed data.Extensive experiments are conducted to evaluate the proposed method, from which we observe superior performance in clustering.

## References


[1] Liu, G., Lin, Z. and Yu, Y. (2010) Robust Subspace Segmentation by Low-Rank Representation. Proceedings of the 27th International Conference on Machine Learning (ICML-10), 663–670.

[2] Liu, G. and Yan, S. (2011) Latent Low-Rank Representation for Subspace Segmentation and Feature Extraction. Proceedings of the IEEE International Conference on Computer Vision (ICCV), Barcelona, Spain, 6-13 November 2011, 1615-1622.

[3] Peng, C., Kang, Z., Li, H. and Cheng, Q. (2015) Subspace Clustering Using Log-Determinant Rank Approximation. Proceedings of the 21st ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD), Sydney, Australia, 10-13 August 2015, 925-934.

[4] Elhamifar, E. and Vidal, R. (2013) Sparse Subspace Clustering: Algorithm, Theory, and Applications. IEEE Transactions on Pattern Analysis and Machine Intelligence, 35(11), 2765–2781. https://doi.org/10.1109/TPAMI.2013.57

[5] Lu, C.-Y., Min, H., Zhao, Z.-Q., Zhu, L., Huang, D.-S. and Yan, S. (2012) Robust and Efficient Subspace Segmentation via Least Squares Regression. Computer Vision – ECCV, Florence, Italy, 7-13 October 2012, 347-360.

[6] Liu, G., Lin, Z., Yan, S., Sun, J., Yu, Y. and Ma, Y. (2013) Robust Recovery of Subspace Structures by Low-Rank Representation. IEEE Transactions on Pattern Analysis and Machine Intelligence, 35(1), 171–184. https://doi.org/10.1109/TPAMI.2012.88

[7] Peng, C., Kang, Z., Yang, M. and Cheng, Q. (2016) Feature Selection Embedded Subspace Clustering. IEEE Signal Processing Letters, 23(7), 1018–1022. https://doi.org/10.1109/LSP.2016.2573159

[8] Li, C.-G. and Vidal, R. (2015) Structured Sparse Subspace Clustering: A Unified Optimization Framework. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Boston, MA, USA, 7-12 June 2015, 277-286.

[9] Liu, J., Chen, Y., Zhang, J. and Xu, Z. (2014) Enhancing Low-Rank Subspace Clustering by Manifold Regularization. IEEE Transactions on Image Processing, 23(9), 4022–4030. https://doi.org/10.1109/TIP.2014.2343458

[10] Zhang, F., Qian, J. and Jian, Y. (2013) Nuclear Norm Based 2DPCA. Proceedings of the 2013 2nd IAPR Asian Conference on Pattern Recognition (ACPR), Naha, Japan, 5-8 November 2013, 74-78.

[11] Chung, F.R. (1997) Spectral Graph Theory. American Mathematical Society, Providence, Vol. 92.

[12] Shi, J. and Malik, J. (2000) Normalized Cuts and Image Segmentation. IEEE Transactions on Pattern Analysis and Machine Intelligence, 22(8), 888–905. https://doi.org/10.1109/34.868688

[13] Agarwal, P.K. and Mustafa, N.H. (2004) k-Means Projective Clustering. Proceedings of the Twenty-Third ACM SIGMOD-SIGACT-SIGART Symposium on Principles of Database

Systems (PODS), Paris, France, 14-16 June 2004, 155-165.

[14] Peng, C., Kang, Z., Hu, Y., Cheng, J. and Cheng, Q. (2017) Nonnegative Matrix Factorization with Integrated Graph and Feature Learning. ACM Transactions on Intelligent Systems and Technology, 8(3), 42:1–42:29. https://doi.org/10.1145/2987378

[15] Peng, C., Kang, Z., Cai, S. and Cheng, Q. (2018) Integrate and Conquer: Double-Sided Two-Dimensional k-Means via Integrating of Projection and Manifold Construction. ACM Transactions on Intelligent Systems and Technology, 9(5). https://doi.org/10.1145/3200488