

Low-Rank Multi-View Subspace Clustering Based on Sparse Regularization

Yan Sun, Fanlong Zhang*

School of Computer Science, Nanjing Audit University, Nanjing, China

Email: *csfzhang@nau.edu.cn

How to cite this paper: Sun, Y. and Zhang, F.L. (2024) Low-Rank Multi-View Subspace Clustering Based on Sparse Regularization. *Journal of Computer and Communications*, 12, 14-30.
<https://doi.org/10.4236/jcc.2024.124002>

Received: March 2, 2024

Accepted: March 30, 2024

Published: April 2, 2024

Copyright © 2024 by author(s) and Scientific Research Publishing Inc. This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

Abstract

Multi-view Subspace Clustering (MVSC) emerges as an advanced clustering method, designed to integrate diverse views to uncover a common subspace, enhancing the accuracy and robustness of clustering results. The significance of low-rank prior in MVSC is emphasized, highlighting its role in capturing the global data structure across views for improved performance. However, it faces challenges with outlier sensitivity due to its reliance on the Frobenius norm for error measurement. Addressing this, our paper proposes a Low-Rank Multi-view Subspace Clustering Based on Sparse Regularization (LMVSC-Sparse) approach. Sparse regularization helps in selecting the most relevant features or views for clustering while ignoring irrelevant or noisy ones. This leads to a more efficient and effective representation of the data, improving the clustering accuracy and robustness, especially in the presence of outliers or noisy data. By incorporating sparse regularization, LMVSC-Sparse can effectively handle outlier sensitivity, which is a common challenge in traditional MVSC methods relying solely on low-rank priors. Then Alternating Direction Method of Multipliers (ADMM) algorithm is employed to solve the proposed optimization problems. Our comprehensive experiments demonstrate the efficiency and effectiveness of LMVSC-Sparse, offering a robust alternative to traditional MVSC methods.

Keywords

Clustering, Multi-View Subspace Clustering, Low-Rank Prior, Sparse Regularization

1. Introduction

Clustering plays a significant role in machine learning and artificial intelligence (AI) for several reasons, acting as a foundational technique that underpins many

of the processes and applications within these fields [1] [2].

Multi-view Subspace Clustering (MVSC) is an advanced clustering technique that is particularly suited for handling data that naturally comes from multiple sources or “views.” This approach is based on the principle that different views of the data can provide complementary information that should be integrated when performing clustering. The goal of MVSC is to find a common subspace that best represents the underlying structure of the data across all views, thereby improving the quality and accuracy of the clustering results. As shown in [3] and [4], by leveraging multiple views of the data, MVSC can achieve higher clustering accuracy than single-view clustering methods, especially when the views are complementary. The works [5] and [6] claim that the integration of multiple views can make the clustering process more robust to noise and redundancy within individual views, as the method can exploit the clean and informative parts of each view.

As a result, the Multi-view Subspace Clustering has been applied in various areas, such as image and video analysis [7], bioinformatics [8] [9], and social network analysis [10]. Among them, low-rank prior is a critical technique used to capture the global structure of data across multiple views while ensuring that the representation is compact and meaningful. The low-rank prior is based on the assumption that the data from all views lie on or near a low-dimensional subspace, and this inherent structure can be exploited to improve clustering performance. The key contributions in the domain of low-rank subspace-based methods encompass several notable works, such as Latent Multi-view Subspace Clustering (LMSC) [11], Multimodal Sparse and Low-rank Subspace Clustering (MLRSSC) [12], Flexible Multi-view Representation Learning for Subspace Clustering (FMR) [13] and Dual Shared-Specific Multi-view Subspace Clustering (DSS-MS) [14]. These methods, grounded in the well-established framework of low-rank representation, have demonstrated competitive clustering performance in empirical studies.

Especially, the recent work [15] introduced an efficient and effective approach termed Facilitated Low-rank Multi-view Subspace Clustering (FLMSC), they factorize the view-specific representation matrix into two small factor matrices, *i.e.*, an orthogonal dictionary and a latent representation, which can fully explore the underlying subspace structure of multiple views. However, this approach still suffers from one issue. They focus on the F-norm to measure the error between observation data and reconstruction data. However, as well known F-norm is sensitive to outliers, as it squares the values before summing them, which can proportionately increase the impact of larger differences.

To address the aforementioned drawback, this paper develops a Low-Rank Multi-view Subspace Clustering Based on Sparse Regularization approach, which is featured by sparse regularization. The main contributions and novelty of this work can be summarized below. 1) By employing sparse regularization, this paper presents a robust low-rank multi-view subspace clustering approach termed,

LMVSC-Sparse. 2) This paper develops an Alternating Direction Method of Multipliers (ADMM) algorithm to solve proposed optimization problems. 3) Comprehensive experiments are conducted on benchmark data sets, which have shown the advantage of our approach in both efficiency and effectiveness.

The rest of the paper is organized as follows: Section 2 reviews the related model. In Section 3, we propose a novel model, low-rank multi-view subspace clustering based on sparse regularization (LMVSC-Sparse) and its corresponding optimization algorithm. Section 4 illustrates the experiment's benchmark data sets.

2. Foundational Model

Denote $X = \{x_1, \dots, x_n\} \in \mathbb{R}^{d \times n}$ a collection of data samples, where d is the feature dimension and n is the number of data samples. Then, the traditional subspace clustering based on the self-expression can be modeled by (1):

$$X = XZ + E. \tag{1}$$

where $Z \in \mathbb{R}^{n \times n}$ denotes the subspace representation at dictionary X and E denotes the error matrix. The corresponding optimization problem based on low rank prior can be given by (2)

$$\min_Z \|XZ - X\|_F^2 + \lambda \|Z\|_*. \tag{2}$$

in which $\|Z\|_*$ means nuclear norm, computed by summing the singular values of $\|Z\|_*$. The nuclear norm is widely used to measure the low rank property. (2) is also called low-rank representation (LRR) [16]. LRR has being successfully applied in dimensionality reduction [17], noise reduction [18], recommendation systems [19], image processing [20], and also clustering and classification [21] [22].

After solving (2), the affinity matrix W can be obtained by

$$W = \frac{1}{2} \left(|Z| + |Z^T| \right). \tag{3}$$

Then one can obtain the final clustering results by conducting spectral clustering algorithm.

The (2) can be extended to a general problem (4):

$$\min_Z \|XZ - X\|_F^2 + \lambda f \|Z\|_*. \tag{4}$$

where function f means general regularization.

Now, considering the multi-view data samples, denoted by $X = \{X^{(1)T}, \dots, X^{(V)T}\}^T$, where $X^{(v)} = [x_1^{(v)}, \dots, x_n^{(v)}] \in \mathbb{R}^{d_v \times n}$ being the v th view. The LRR (2) can be extended to (5):

$$\min_{\{Z^{(v)}\}_{v=1}^V} \sum_{v=1}^V \|X^{(v)} - X^{(v)}Z^{(v)}\|_F^2 + \lambda \sum_{v=1}^V \|Z^{(v)}\|_*. \tag{5}$$

Recently, the [15] furthermore extends (5) to following (6):

$$\begin{aligned} & \min_{\{Z^{(v)}, L^{(v)}, C^{(v)}\}_{v=1}^V} \sum_{v=1}^V \|X^{(v)} - X^{(v)} Z^{(v)}\|_F^2 + \lambda_1 \sum_{v=1}^V \|C^{(v)}\|_* + \lambda_2 \sum_{v=1, v \neq w}^V \|C^{(v)} - C^{(w)}\|_F^2 \\ & \text{s.t. } Z^{(v)} = L^{(v)} C^{(v)}, L^{(v)T} L^{(v)} = I, \forall v \end{aligned} \quad (6)$$

where λ_1 and λ_2 are two positive balancing parameters.

Compared to (5), the (6) factorize the view specific representation $Z^{(v)}$ into two small matrices $L^{(v)}$ and $C^{(v)}$, and employed the property $\|C^{(v)}\|_* = \|Z^{(v)}\|_*$ when $L^{(v)}$ has orthogonal columns, *i.e.* $L^{(v)T} L^{(v)} = I$.

3. Proposed Approach

In this section, we utilize sparse regularization instead of F norm regularization. in (6). Sparse regularization is often preferred over F-norm regularization in various machine learning and signal processing applications due to its unique properties and benefits, especially when dealing with high-dimensional data or models that incorporate many parameters. There exist two benefits for sparse regularization. 1) Sparse regularization can promote sparsity of solution. This means that they encourage the model to use fewer features or parameters by driving the coefficients of less important features to zero. This is particularly useful in feature selection and for models where interpretability is important, as it highlights which features are most relevant to the prediction. 2) Sparse regularization is effective at preventing overfitting, especially in high-dimensional settings where the number of features greatly exceeds the number of observations. By encouraging a model to concentrate on fewer variables, it reduces the model's complexity and enhances its capacity to fit noise.

Thus, we replace the F-norm in (6) by L1-norm, and obtain the Low-Rank Multi-view Subspace Clustering Based on Sparse Regularization (LMVSC-Sparse):

$$\begin{aligned} & \min_{\{Z^{(v)}, L^{(v)}, C^{(v)}\}_{v=1}^V} \sum_{v=1}^V \|X^{(v)} - X^{(v)} Z^{(v)}\|_1 + \lambda_1 \sum_{v=1}^V \|C^{(v)}\|_* + \lambda_2 \sum_{v=1, v \neq w}^V \|C^{(v)} - C^{(w)}\|_F^2 \\ & \text{s.t. } Z^{(v)} = L^{(v)} C^{(v)}, L^{(v)T} L^{(v)} = I, \forall v \end{aligned} \quad (7)$$

Based on the framework of ADMM [23], we propose an efficient optimization algorithm to solve the minimization problem abovementioned. First, the corresponding augmented Lagrange function is formulated as follows:

$$\begin{aligned} & L\left(\{Z^{(v)}, L^{(v)}, C^{(v)}\}_{v=1}^V\right) \\ & = \sum_{v=1}^V \|X^{(v)} - X^{(v)} Z^{(v)}\|_1 + \lambda_1 \sum_{v=1}^V \|C^{(v)}\|_* + \frac{\lambda_2}{V-1} \sum_{v=1, v \neq w}^V \|C^{(v)} - C^{(w)}\|_F^2 \\ & \quad + \sum_{v=1}^V \langle Y^{(v)}, Z^{(v)} - L^{(v)} C^{(v)} \rangle + \frac{\mu}{2} \sum_{v=1}^V \|Z^{(v)} - L^{(v)} C^{(v)}\|_F^2 \\ & \text{s.t. } L^{(v)T} L^{(v)} = I, \forall v \end{aligned} \quad (8)$$

where $\{Y^{(v)}\}_{v=1}^V$ represent the Lagrange multipliers and μ represents the penalty parameter. Apparently, it is not easy to optimize all the variables at the same time. Therefore, we adopt an iterative optimization scheme to update the va-

riables one by one. The corresponding procedure of updating steps as shown in what follows.

1) Update the variables $\{Z^{(v)}\}_{v=1}^V$

When fixing the other variables, we can solve the following minimization sub-problem w.r.t. variable $Z^{(v)}$:

$$\min_{Z^{(v)}} \|X^{(v)} - X^{(v)}Z^{(v)}\|_1 + \langle Y^{(v)}, Z^{(v)} - L^{(v)}C^{(v)} \rangle + \frac{\mu}{2} \|Z^{(v)} - L^{(v)}C^{(v)}\|_F^2. \quad (9)$$

This can be rewritten equivalently as:

$$\min_{Z^{(v)}} \|X^{(v)} - X^{(v)}Z^{(v)}\|_1 + \frac{\mu}{2} \left\| Z^{(v)} - L^{(v)}C^{(v)} + \frac{1}{\mu} Y^{(v)} \right\|_F^2. \quad (10)$$

By introducing the auxiliary variable H , and omitting the superscript for simplicity, we have:

$$\min_{Z,H} \|H\|_1 + \frac{\mu}{2} \left\| Z - LC + \frac{1}{\mu} Y \right\|_F^2 \quad \text{s.t. } X - XZ = H. \quad (11)$$

This is a constrained optimization problem. We adopt half-quadratic splitting (HQS) [24] algorithm for its simplicity and fast convergence. Then it is solved by following minimization

$$\min_{Z,H} \|H\|_1 + \frac{\mu}{2} \left\| Z - LC + \frac{1}{\mu} Y \right\|_F^2 + \frac{\eta}{2} \|X - XZ - H\|_F^2. \quad (12)$$

where η is a penalty parameter that forces $X - XZ$ and H to approach the same fixed point. Subsequently, H and Z can be updated by following two sub-problems.

Sub-problem one:

$$\min_H f_1 = \|H\|_1 + \frac{\eta}{2} \|X - XZ - H\|_F^2. \quad (13)$$

Sub-problem two:

$$\min_Z f_2 = \frac{\mu}{2} \left\| Z - LC + \frac{1}{\mu} Y \right\|_F^2 + \frac{\eta}{2} \|X - XZ - H\|_F^2. \quad (14)$$

For first sub-problem, let $\mathcal{S}_\tau : R \rightarrow R$ denote the shrinkage operator $\mathcal{S}_\tau(x) = \text{sgn}(x) \max(|x| - \tau, 0)$ and extend it to matrices by applying it to each element. It is easy to show that above sub-problem's solution can be given by

$$H = \mathcal{S}_\eta(X - XZ). \quad (15)$$

For second sub-problem, it is equal to following problem

$$\min_Z f_3 = \frac{\mu}{\eta} \left\| Z - LC + \frac{1}{\mu} Y \right\|_F^2 + \|X - XZ - H\|_F^2. \quad (16)$$

Its solution can be given by setting the derivative of above sub-problem to zero and obtain:

$$\begin{aligned}
 f_3 &= \frac{\mu}{\eta} \text{trace} \left[\left(Z - LC + \frac{1}{\mu} Y \right)^T \left(Z - LC + \frac{1}{\mu} Y \right) \right] \\
 &\quad + \text{trace} \left[(X - XZ - H)^T (X - XZ - H) \right] \\
 &= \frac{\mu}{\eta} \text{trace} \left[Z^T Z + 2Z^T \left(\frac{1}{\mu} Y - LC \right) \right] + \text{trace} \left[Z^T X^T XZ + 2Z^T X^T (H - X) \right] \\
 &= \text{trace} \left[\frac{\mu}{\eta} Z^T Z + 2Z^T \left(\frac{1}{\eta} Y - \frac{\mu}{\eta} LC \right) \right] + \text{trace} \left[Z^T X^T XZ + 2Z^T X^T (H - X) \right] \\
 &= \text{trace} \left[Z^T \left(\frac{\mu}{\eta} + X^T X \right) Z + 2Z^T \left(\frac{1}{\eta} Y - \frac{\mu}{\eta} LC + X^T (H - X) \right) \right]
 \end{aligned} \tag{17}$$

Then

$$\frac{\partial f_3}{\partial Z} = 2 \left(\frac{\mu}{\eta} + X^T X \right) Z + 2 \left(\frac{1}{\eta} Y - \frac{\mu}{\eta} LC + X^T (H - X) \right). \tag{18}$$

Let $\frac{\partial f_3}{\partial Z} = 0$, we have

$$Z = \left(\frac{\mu}{\eta} + X^T X \right)^{-1} \left(-\frac{1}{\eta} Y + \frac{\mu}{\eta} LC + X^T (X - H) \right). \tag{19}$$

By Sherman-Morrison-Woodbury equation, according the size of X , Z can also be rewritten as

$$\begin{aligned}
 \left(\frac{\mu}{\eta} + X^T X \right)^{-1} &= \frac{\eta}{\mu} I - \frac{\eta}{\mu} X^T \left(I + X \frac{\eta}{\mu} X^T \right)^{-1} X \frac{\eta}{\mu} \\
 &= \frac{\eta}{\mu} \left(I - \frac{\eta}{\mu} X^T \left(I + \frac{\eta}{\mu} X X^T \right)^{-1} X \right)
 \end{aligned} \tag{20}$$

2) Update rule for the variables $\{L^{(v)}\}_{v=1}^V$

When fixing the other variables, we can solve the following minimization sub-problem w.r.t. variable $L^{(v)}$:

$$\begin{aligned}
 \min_{L^{(v)}} &\left\langle Y^{(v)}, Z^{(v)} - L^{(v)} C^{(v)} \right\rangle + \frac{\mu}{2} \left\| Z^{(v)} - L^{(v)} C^{(v)} \right\|_F^2 \\
 \text{s.t.} &L^{(v)T} L^{(v)} = I
 \end{aligned} \tag{21}$$

This constrained problem could be further reduced into the form as follows:

$$\max_{L^{(v)}} \text{Tr} \left(L^{(v)T} R^{(v)} \right) \quad \text{s.t. } L^{(v)T} L^{(v)} = I. \tag{22}$$

where $R^{(v)} = \left(Z^{(v)} + \frac{1}{\mu} Y^{(v)} \right) C^{(v)T}$. Before solving (22), we need following lemma:

Lemma 1. For any matrices $A \in \mathbb{R}^{m \times n}$, suppose the singular value decomposition (SVD) of matrix A is $U \Lambda V^T$, then we consider the following constrained problem:

$$\max_Y \text{Tr} \left(Y^T A \right) \quad \text{s.t. } Y^T Y = I. \tag{23}$$

has closed form as follows:

$$Y = UV^T. \tag{24}$$

Based on the **Lemma 1**, by performing the SVD decomposition of matrix $R^{(v)}$ as $R^{(v)} = U_R^{(v)} \Lambda_R^{(v)} W_R^{(v)}$, the solution for (22) can be achieved by:

$$L^{(v)} = U_R^{(v)} W_R^{(v)T}. \tag{25}$$

3) Update rule for the variables $\{C^{(v)}\}_{v=1}^V$

When fixing the other variables, we obtain the problem (26):

$$\begin{aligned} \min_{C^{(v)}} \lambda_1 \|C^{(v)}\|_* + \frac{\lambda_2}{V-1} \sum_{v \neq w} \|C^{(v)} - C^{(w)}\|_F^2 \\ + \langle Y^{(v)}, Z^{(v)} - L^{(v)} C^{(v)} \rangle + \frac{\mu}{2} \|Z - L^{(v)} C^{(v)}\|_F^2 \end{aligned} \tag{26}$$

Before solving (26), we need following lemma [25]:

Lemma 2. For a given matrix F and a positive parameter $\tau > 0$, the optimal solution to the following problem

$$\min_D \tau \|D\|_* + \frac{1}{2} \|D - F\|_F^2. \tag{27}$$

is given by

$$D = U_F \Theta_\tau(\Sigma_F) W_F^T. \tag{28}$$

where $U_F \Sigma_F W_F^T$ is the SVD decomposition of matrix F . Meanwhile, $\Theta_\tau(\cdot)$ is defined as follows:

$$\Theta_\tau(\Sigma_F) = \max(0, \Sigma_F - \tau) + \min(0, \Sigma_F + \tau). \tag{29}$$

Based on **Lemma 2**, by setting $\gamma = \frac{\lambda_1}{2(\mu + \lambda_2)}$, the closed-formed solution for variable $C^{(v)}$ is shown as follows:

$$C^{(v)} = U_H^{(v)} \Theta_\gamma(\Sigma_H^{(v)}) W_H^{(v)T}. \tag{30}$$

where, $U_H^{(v)} \Sigma_H^{(v)} W_H^{(v)T}$ represents the SVD decomposition of $\frac{1}{\mu + \lambda_2} H^{(v)}$, and

$$H^{(v)} = \mu L^{(v)T} \left(Z^{(v)} + \frac{1}{\mu} Y^{(v)} \right) + \frac{\lambda_2}{V-1} \sum_{w \neq v} C^{(w)} \tag{31}$$

The final affinity matrix S could be obtained as follows:

$$\tilde{S} = \frac{1}{V} \sum_{v=1}^V L^{(v)} C^{(v)}. \tag{32}$$

and

$$S = \frac{|\tilde{S}| + |\tilde{S}|^T}{2}.$$

In a nutshell, the detailed optimization process for LMVSC-Sparse is summarized in **Algorithm 1**.

Algorithm 1. ADMM for LMVSC-Sparse.

Input: multi-view data the multi-view data samples, denoted by $X = \{X^{(1)}, \dots, X^{(V)}\}$, λ_1 , λ_2 , and the numbers of clusters.

Initialize: $\mu = 10^{-3}$, $\rho = 1.2$, $\mu_{\max} = 10^6$, $\gamma = 100$, $L^{(v)} = 0$, $C^{(v)} = 0$, $\forall v$, stooping tolerance ϵ .

Repeat:

Step 1: Computing $\{Z^{(v)}\}_{v=1}^V$ by (19).

Step 2: Computing $\{L^{(v)}\}_{v=1}^V$ by (25).

Step 3: Computing $\{C^{(v)}\}_{v=1}^V$ by (30).

Step 4: Computing $\{Y^{(v)}\}_{v=1}^V$ by $Y^{(v)} = Y^{(v)} + \mu(Z^{(v)} - L^{(v)}C^{(v)})$.

Step 5: Computing $\mu = \min(\rho\mu, \mu_{\max})$.

Until $\sum_{v=1}^V \|Z^{(v)} - L^{(v)}C^{(v)}\|_{\infty} \leq \epsilon$ or achieved maxim iteration number.

Output $Z^{(v)}$.

4. Experiments

The proposed algorithm is compared with four state-of-the-art cluster algorithms, namely, Facilitated Low-rank Multi-view Subspace Clustering (FLMSC) [15], Scalable Multi-view Subspace Clustering with Unified Anchors (SMVSC) [26], Large-Scale Multi-View Subspace Clustering (LMVSC) [3], Graph-based Multi-view Clustering (GMC) [27].

4.1. Data and Metrics

To verify performance, the BBC [28] is used for clustering. There are 2225 documents over 5 annotated topics in this data set. In the experiments, we use as ample subset of original BBC consisting of 685 documents and four different views, with 4659, 4633, 4665 and 4684 in each view, respectively.

For the evaluation metrics, F-score, Normalized Mutual Information (NMI), Accuracy (ACC), and Adjusted Rand index (AR) are employed. The F-score, also known as the F1-score or F-measure, considers both the precision and the recall to compute the score. Normalized Mutual Information (NMI) is a measure used to evaluate the similarity between two clustering of a dataset. It's a measure of the mutual dependence between two variables, in this case, the clustering assignments obtained from different algorithms or methods. Accuracy (ACC) is a common evaluation metric used in the context of clustering, particularly when the ground truth labels are available. It measures the proportion of data points that are correctly assigned to their true clusters.

It's a popular metric due to its simplicity and ability to handle varying cluster sizes and shapes. AR is a measure that assesses the similarity between two clustering by considering all pairs of samples and counting pairs that are assigned to the same or different clusters in the predicted and true clustering. It then adjusts

the raw Rand Index to account for the expected similarity between clustering due to chance.

4.2. Comparison with State-of-Arts

In the first experiments, 1% elements in BBC dataset is added noise, the noise level vary from 0 to 0.5. Then we compare the F-score, NMI, ACC, and AR for various algorithms. The results are shown in **Figures 1-4**.

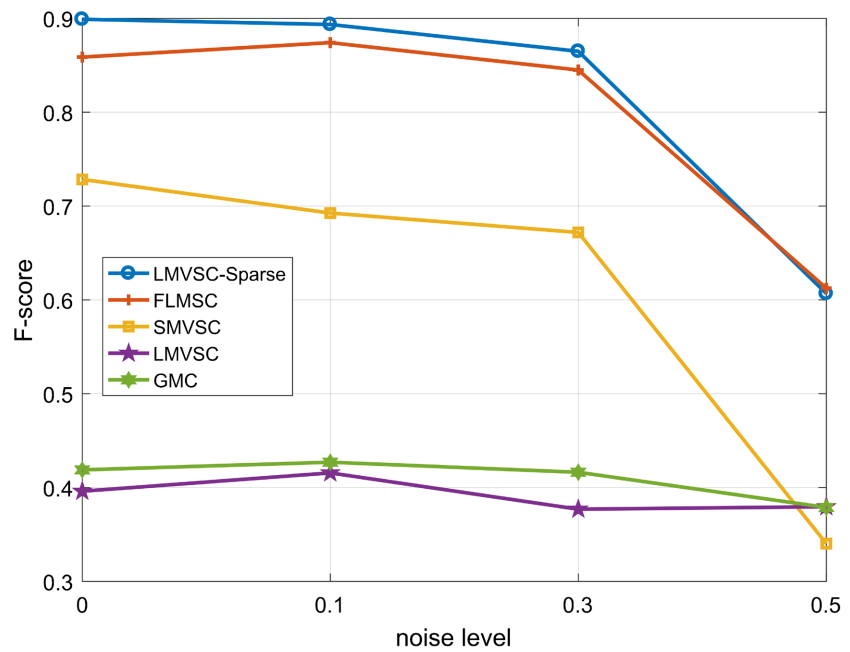


Figure 1. The F-score values at different noise level.

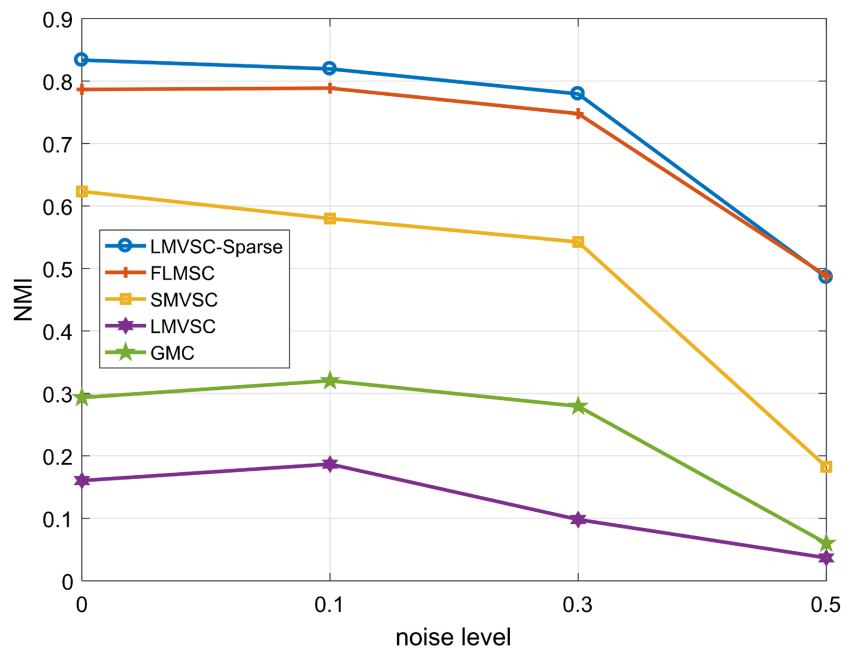


Figure 2. The NMI values at different noise level.

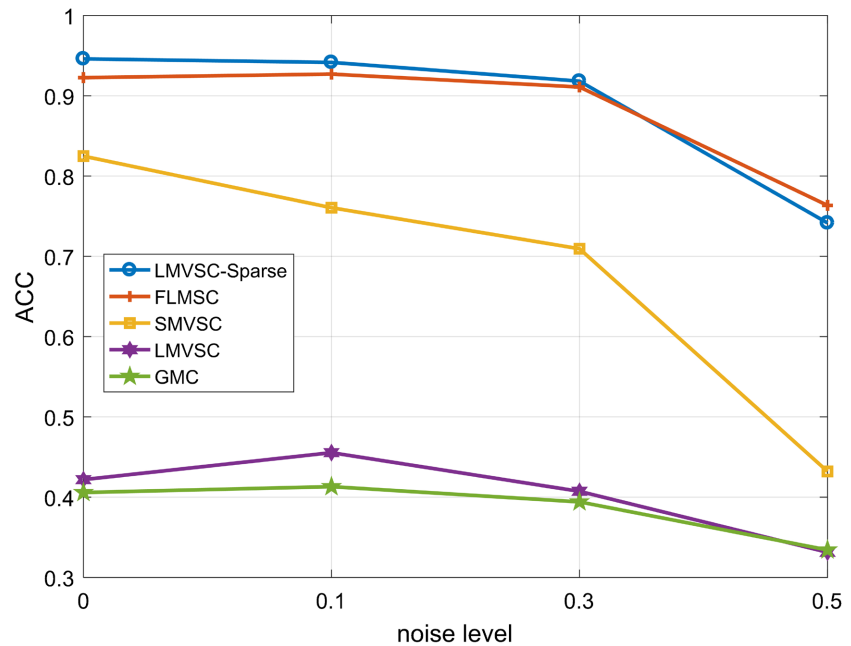


Figure 3. The ACC values at different noise level.

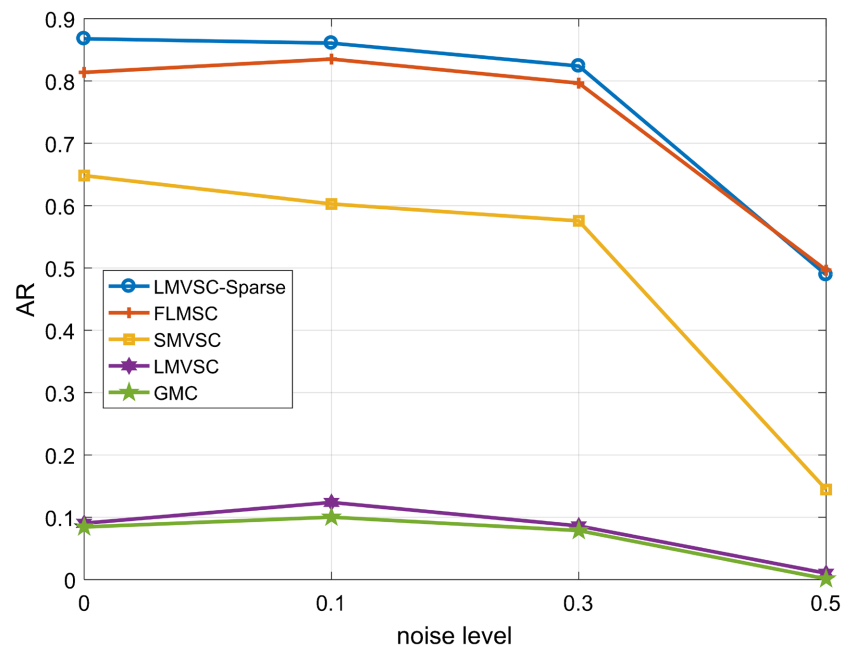


Figure 4. The AR values at different noise level.

From **Figure 1**, one might conclude that LMVSC-Sparse is the most robust method in the presence of noise, maintaining a high F-score across all tested noise levels. In contrast, GMC is the least effective method in terms of F-score, regardless of the noise level. The other methods show varying degrees of decline in their F-scores as the noise level increases, suggesting that they are more sensitive to noise than LMVSC-Sparse. These observations could be useful for selecting a method for applications where data is expected to have a certain level of

noise. LMVSC-Sparse might be preferable in environments where noise is unavoidable or difficult to control.

In **Figure 2**, the overall trend indicates that all methods suffer a decline in clustering performance as noise increases, but to varying degrees. LMVSC-Sparse appears to be the most robust against noise, maintaining a high NMI throughout. GMC is markedly affected by noise, with a significant decrease in NMI as the noise level rises. For applications where maintaining clustering quality in the presence of noise is important, LMVSC-Sparse would likely be the preferred choice based on this data. The other methods may still be considered, but their performance will depend on the acceptable threshold for NMI in the context of the specific application. **Figure 3** and **Figure 4** show the similar trend as **Figure 1** and **Figure 2**.

Also, the mean running times of algorithms are summarized in **Table 1**. From this table, we can conclude that, LMVSC is the fastest algorithm across all noise levels, making it suitable for applications where running time is critical. LMVSC-Sparse is the slowest, which might be a trade-off for its robust performance in terms of F-score, NMI, ACC and AR, as indicated in the previous figures. Choosing the right algorithm would depend on the balance between accuracy (as measured by F-score, NMI, ACC and AR) and efficiency (as measured by running time), alongside the specific requirements of the application or task at hand.

In the second experiments, 1%, 5%, 10%, 20%, 50% elements in BBC dataset is added noise, the noise level is 0.1. Then we compare the F-score, NMI, ACC, and AR for various algorithms. The results are shown in **Figures 5-8**. From these figures, we can conclude that:

- 1) All algorithms show a decline in F-score with increasing sparsity levels, with LMVSC-Sparse and FLMSC being the least affected. GMC's performance drops significantly and remains low across sparsity levels.
- 2) For NMI again, all algorithms show a decline as sparsity increases, with LMVSC-Sparse showing the least impact. GMC performs poorly at higher sparsity levels.
- 3) For ACC, we see a sharp decline for all methods as sparsity increases, with LMVSC-Sparse being the most robust but still affected.
- 4) For AC, all algorithms experience a drop sparsity increase. LMVSC-Sparse and FLMSC tend to have better robustness compared to others.

4.3. Parameter Sensitivity Analysis

Figure 9 shows a series of heatmaps that represent a parameter analysis for different evaluation metrics and rank bounds. Each heatmap corresponds to a combination of two parameters, which are λ_1 and λ_2 . The colors in each heatmap represent different values of the metric being evaluated, with darker or brighter colors typically indicating better performance. Here's a breakdown of the analysis:

Table 1. The running times compare (Time unit: s).

| Noise Level | LMVSC-Sparse | FLMSC | SMVSC | LMVSC | GMC |
|-------------|--------------|-------|-------|-------|------|
| 0 | 44.47 | 19.93 | 21.84 | 2.26 | 4.02 |
| 0.1 | 44.39 | 19.66 | 21.72 | 2.24 | 3.89 |
| 0.3 | 44.76 | 19.85 | 21.22 | 2.18 | 3.70 |
| 0.5 | 45.55 | 20.55 | 20.97 | 2.08 | 3.21 |

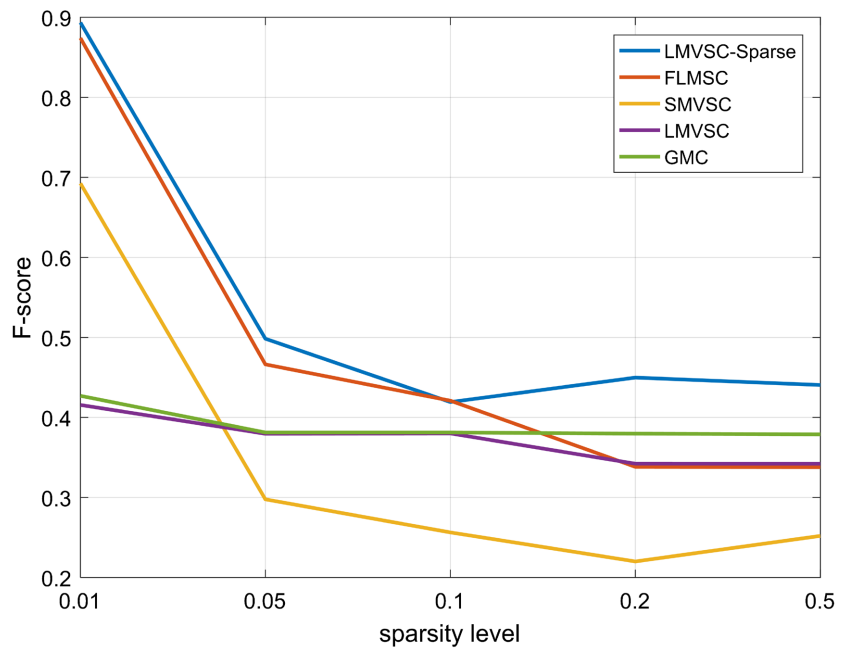


Figure 5. The F-score values at different sparsity level.

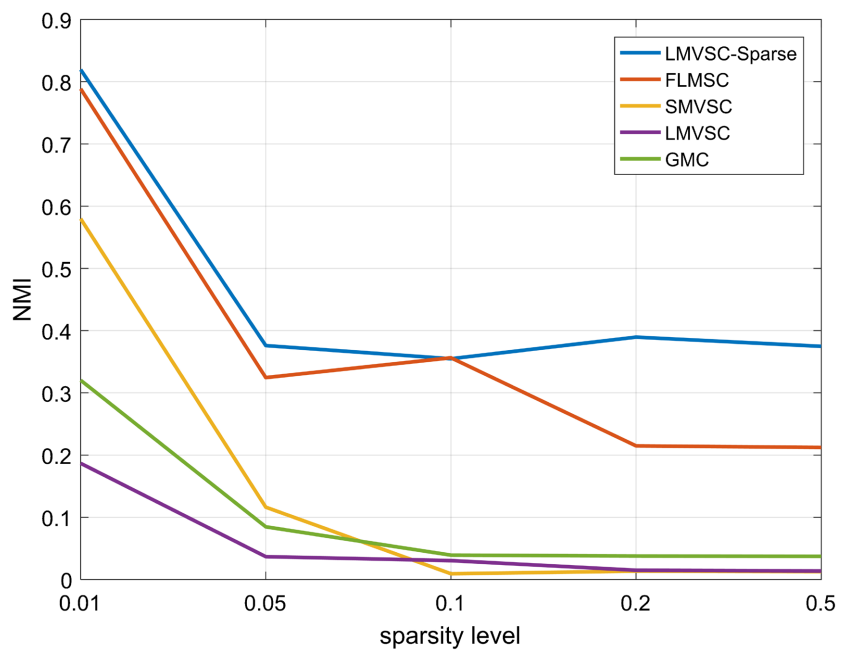


Figure 6. The NMI values at different sparsity level.

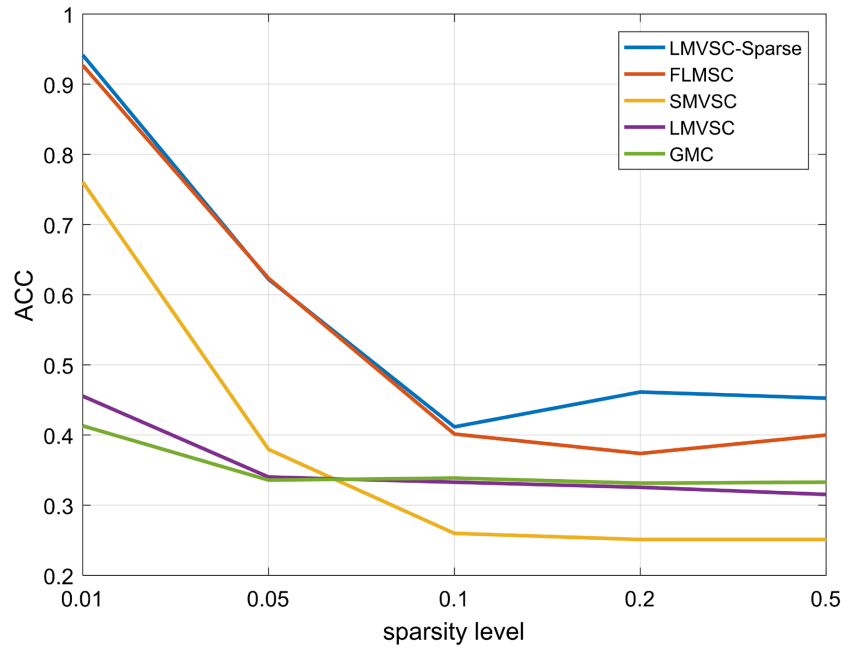


Figure 7. The ACC values at different sparsity level.

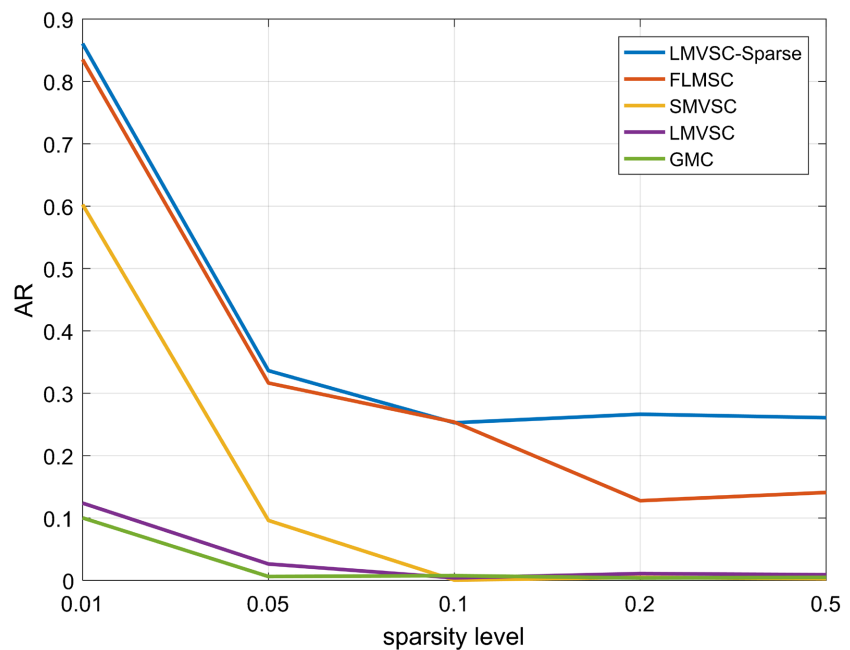


Figure 8. The AR values at different sparsity level.

In summary, the performance of all algorithms degrades with increasing sparsity, which is expected as sparser data tends to have less information for the algorithms to leverage in the clustering process. LMSC-Sparse seems to be the most robust across all evaluated metrics, maintaining higher values than the others as sparsity increases.

1) There is a consistent trend where the F-score, NMI, and AC seem to be more sensitive to the second parameter across rank bounds.

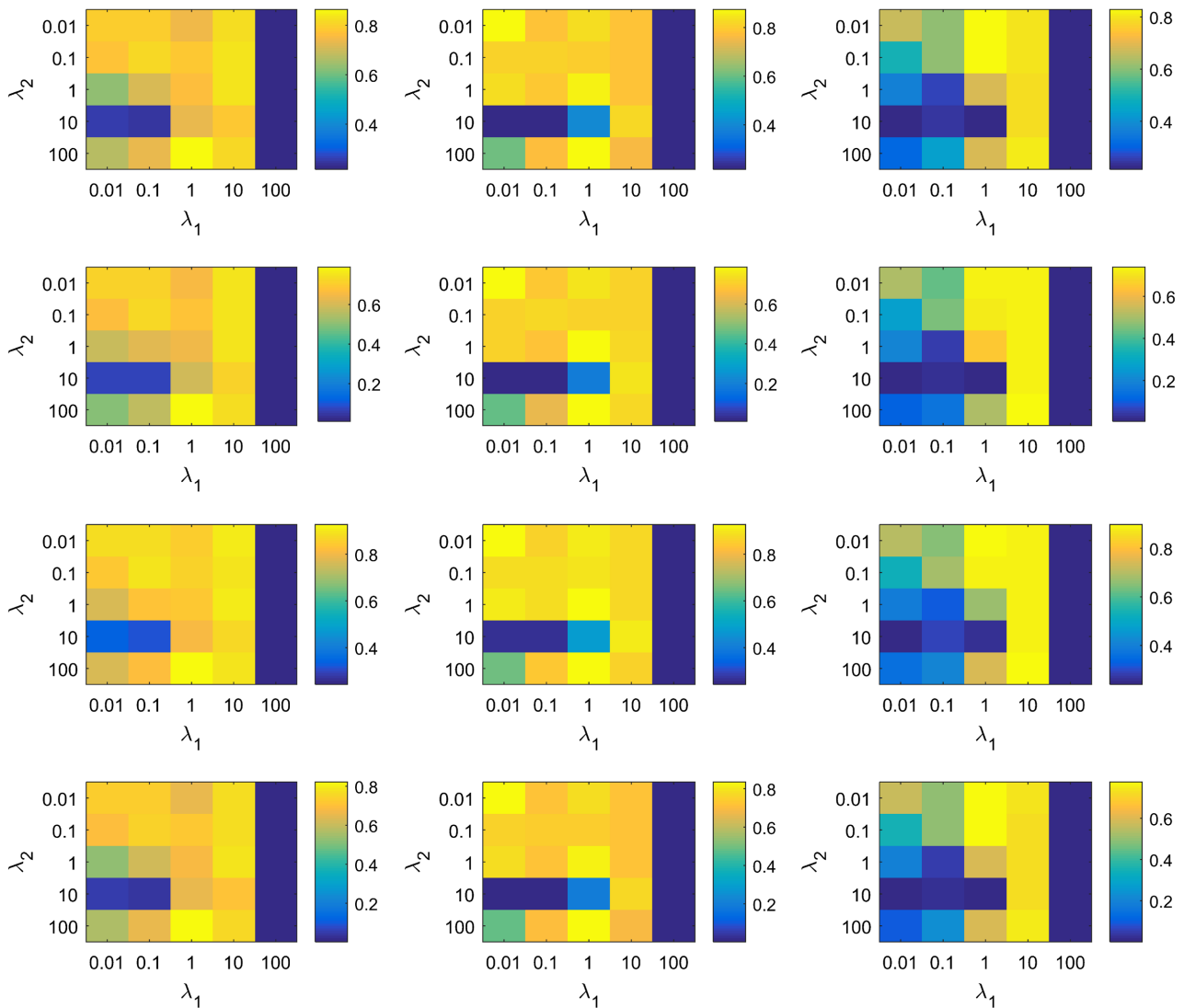


Figure 9. Parameter analysis, the first row means F-score, second row NMI, third row ACC, fourth row AC, the first column means rank bound is 50, second column 100, third column 200.

2) For all metrics, there are specific parameter combinations that yield high values, indicating optimal regions for each rank bound setting.

3) The first parameter appears to have a less significant impact on the F-score and NMI compared to ACC and AC.

4) The optimal regions for high values seem to shift and become less extensive as the rank bound increases, which may indicate that models with higher complexity (larger rank bounds) do not necessarily perform better and can be harder to tune.

In conclusion, these heatmaps can be used to identify the optimal parameter settings for each rank bound and metric. It is important to balance model complexity with the ability to tune the parameters effectively, as overly complex models may not yield better performance and can be more challenging to optimize.

5. Conclusion

This paper introduced an innovative Low-Rank Multi-view Subspace Clustering based on Sparse Regularization (LMVSC-Sparse) method. LMVSC-Sparse incorporated sparse regularization to mitigate the impact of outliers, thus enhancing the robustness of the clustering process. The developed ADMM algorithm efficiently solved the optimization problem, ensuring both effectiveness and efficiency, as evidenced by the experimental results on benchmark datasets. The performance of LMVSC-Sparse, particularly in noisy and sparse conditions, demonstrated its superiority over other state-of-the-art MVSC methods. This robustness is critical for practical applications in fields such as image analysis, bioinformatics, and social network analysis, where data often contain noise and come from diverse sources. The results of this work not only further the understanding of multi-view clustering dynamics but also open avenues for future research in optimizing clustering methods for complex, real-world datasets.

Acknowledgements

This work was supported partly by the National Natural Science Foundation of China under Grant No. 62276137.

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

- [1] Ezugwu, A.E., Ikotun, A.M., Oyelade, O.O., *et al.* (2022) A Comprehensive Survey of Clustering Algorithms: State-of-the-Art Machine Learning Applications, Taxonomy, Challenges, and Future Research Prospects. *Engineering Applications of Artificial Intelligence*, **110**, Article 104743. <https://doi.org/10.1016/j.engappai.2022.104743>
- [2] Tieghi, L., Becker, S., Corsini, A., *et al.* (2023) Machine-Learning Clustering Methods Applied to Detection of Noise Sources in Low-Speed Axial Fan. *Journal of Engineering for Gas Turbines and Power*, **145**, Article 031020. <https://doi.org/10.1115/1.4055417>
- [3] Kang, Z., Zhou, W., Zhao, Z., *et al.* (2020) Large-Scale Multi-View Subspace Clustering in Linear Time. *Proceedings of the AAAI Conference on Artificial Intelligence*, **34**, 4412-4419. <https://doi.org/10.1609/aaai.v34i04.5867>
- [4] Wu, H., Huang, S., Tang, C., *et al.* (2023) Pure Graph-Guided Multi-View Subspace Clustering. *Pattern Recognition*, **136**, Article 109187. <https://doi.org/10.1016/j.patcog.2022.109187>
- [5] Zhang, X., Ren, Z., Sun, H., *et al.* (2021) Multiple Kernel Low-Rank Representation-Based Robust Multi-View Subspace Clustering. *Information Sciences*, **551**, 324-340. <https://doi.org/10.1016/j.ins.2020.10.059>
- [6] Zhao, N. and Bu, J. (2022) Robust Multi-View Subspace Clustering Based on Consensus Representation and Orthogonal Diversity. *Neural Networks*, **150**, 102-111. <https://doi.org/10.1016/j.neunet.2022.03.009>

- [7] Zhu, W., Lu, J. and Zhou, J. (2019) Structured General and Specific Multi-View Subspace Clustering. *Pattern Recognition*, **93**, 392-403. <https://doi.org/10.1016/j.patcog.2019.05.005>
- [8] Shi, Q., Hu, B., Zeng, T., *et al.* (2019) Multi-View Subspace Clustering Analysis for Aggregating Multiple Heterogeneous Omics Data. *Frontiers in Genetics*, **10**, Article 744. <https://doi.org/10.3389/fgene.2019.00744>
- [9] Liu, H., Shang, M., Zhang, H., *et al.* (2021) Cancer Subtype Identification Based on Multi-View Subspace Clustering with Adaptive Local Structure Learning. *IEEE International Conference on Bio-Informatics and Biomedicine (BIBM)*, Houston, TX, 9-12 December 2021, 484-490. <https://doi.org/10.1109/BIBM52615.2021.9669659>
- [10] Zhang, G.Y., Chen, X.W., Zhou, Y.R., *et al.* (2022) Kernelized Multi-View Subspace Clustering via Auto-Weighted Graph Learning. *Applied Intelligence*, **52**, 716-731. <https://doi.org/10.1007/s10489-021-02365-8>
- [11] Tao, H., Hou, C., Qian, Y., *et al.* (2020) Latent Complete Row Space Recovery for Multi-View Subspace Clustering. *IEEE Transactions on Image Processing*, **29**, 8083-8096. <https://doi.org/10.1109/TIP.2020.3010631>
- [12] Abavisani, M. and Patel, V.M. (2018) Multimodal Sparse and Low-Rank Subspace Clustering. *Information Fusion*, **39**, 168-177. <https://doi.org/10.1016/j.inffus.2017.05.002>
- [13] Li, R., Zhang, C., Hu, Q., *et al.* (2019) Flexible Multi-View Representation Learning for Subspace Clustering. *Proceedings of the Twenty-Eighth International Joint Conference on Artificial Intelligence*, Macao, China, 10-16 August 2019, 2916-2922. <https://doi.org/10.24963/ijcai.2019/404>
- [14] Zhou, T., Zhang, C., Peng, X., *et al.* (2019) Dual Shared-Specific Multi-View Subspace Clustering. *IEEE Transactions on Cybernetics*, **50**, 3517-3530. <https://doi.org/10.1109/TCYB.2019.2918495>
- [15] Zhang, G.-Y., Huang, D. and Wang, C.-D. (2023) Facilitated Low-Rank Multi-View Sub-Space Clustering. *Knowledge-Based Systems*, **260**, Article 110141. <https://doi.org/10.1016/j.knosys.2022.110141>
- [16] Liu, G., Lin, Z., Yan, S., Sun, J., Yu, Y. and Ma, Y. (2012) Robust Recovery of Subspace Structures by Low-Rank Representation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **35**, 171-184. <https://doi.org/10.1109/TPAMI.2012.88>
- [17] Liu, Z., Lu, Y., Lai, Z., *et al.* (2021) Robust Sparse Low-Rank Embedding for Image Dimension Reduction. *Applied Soft Computing*, **113**, Article 107907. <https://doi.org/10.1016/j.asoc.2021.107907>
- [18] Du, S., Liu, B., Shan, G., *et al.* (2022) Enhanced Tensor Low-Rank Representation for Clustering and Denoising. *Knowledge-Based Systems*, **243**, Article 108468. <https://doi.org/10.1016/j.knosys.2022.108468>
- [19] Wang, J., Zhu, L., Dai, T., *et al.* (2021) Low-Rank and Sparse Matrix Factorization with Prior Relations for Recommender Systems. *Applied Intelligence*, **51**, 3435-3449. <https://doi.org/10.1007/s10489-020-02023-5>
- [20] Peng, J., Sun, W., Li, H.C., *et al.* (2021) Low-Rank and Sparse Representation for Hyperspectral Image Processing: A Review. *IEEE Geoscience and Remote Sensing Magazine*, **10**, 10-43. <https://doi.org/10.1109/MGRS.2021.3075491>
- [21] Chen, J., Yang, S., Mao, H., *et al.* (2021) Multiview Subspace Clustering Using Low-Rank Representation. *IEEE Transactions on Cybernetics*, **52**, 12364-12378. <https://doi.org/10.1109/TCYB.2021.3087114>
- [22] Hui, K., Shen, X., Abhadiomhen, S.E., *et al.* (2022) Robust Low-Rank Representa-

- tion via Residual Projection for Image Classification. *Knowledge-Based Systems*, **241**, Article 108230. <https://doi.org/10.1016/j.knosys.2022.108230>
- [23] Falsone, A., Notarnicola, I., Notarstefano, G., *et al.* (2020) Tracking-ADMM for Distributed Constraint-Coupled Optimization. *Automatica*, **117**, Article 108962. <https://doi.org/10.1016/j.automatica.2020.108962>
- [24] Sun, Y., Yang, Y., Liu, Q., *et al.* (2020) Learning Non-Locally Regularized Compressed Sensing Network with Half-Quadratic Splitting. *IEEE Transactions on Multimedia*, **22**, 3236-3248. <https://doi.org/10.1109/TMM.2020.2973862>
- [25] Cai, J.F., Candès, E.J. and Shen, Z. (2010) A Singular Value Thresholding Algorithm for Matrix Completion. *SIAM Journal on Optimization*, **20**, 1956-1982. <https://doi.org/10.1137/080738970>
- [26] Sun, M., Zhang, P., Wang, S., Zhou, S., Tu, W. and Liu, X. (2021) Scalable Multi-View Subspace Clustering with Unified Anchors. *Proceedings of the 29th ACM International Conference on Multimedia*, China, 20-24 October 2021, 3528-3536 <https://doi.org/10.1145/3474085.3475516>
- [27] Wang, H., Yang, Y. and Liu, B (2020) GMC: Graph-Based Multi-View Clustering, *IEEE Transactions on Knowledge and Data Engineering*, **32**, 1116-1129. <https://doi.org/10.1109/TKDE.2019.2903810>
- [28] Huang, L., Chao, H.Y. and Wang, C.D. (2019) Multi-View Intact Space Clustering. *Pattern Recognition*, **86**, 344-353. <https://doi.org/10.1016/j.patcog.2018.09.016>