ISBN: 978-981-94-4198-3
2025 the 15th International Workshop on Computer Science and Engineering (WCSE 2025)
doi: 10.18178/wcse.2025.06.062

Group-weighted Subspace Clustering for Multivariate Compositional Data

Zhengyan Liu [1,3], Huiwen Wang [1,2], Qing Zhao [4 +] and Qiying Wu[1,3]

[1] School of Economics and Management, Beihang University, Beijing, 100191, China

[2] Key Laboratory of Complex System Analysis, Management and Decision (Beihang University), Ministry of Education, Beijing, 100191, China

[3] Beijing Key Laboratory of Emergency Support Simulation Technologies for City Operations, Beijing, 100191, China

[4] School of Mathematics and Physics, University of Science and Technology Beijing, Beijing, 100083, China

Abstract. Compositional data which contain relative information occur frequently in many practical scenarios. Promising achievements have been made in clustering for single compositional data variable, yet few works are available for multivariate compositional data clustering, let alone the importance measurement of compositional data variables. In this paper, we develop a novel clustering method for multivariate compositional data which integrates the self-expressive property and the group-weighted strategy. The self-expressive property is generalized to multivariate compositional data analysis, upon which the coefficient matrix is obtained to reflect the global structure of data. Moreover, considering the importance of each compositional data variable is different during the process of clustering, we propose a group-weighted strategy to adaptively learn appropriate weights for them. Experimental results on a practical dataset demonstrate the effectiveness of the proposed method.

Keywords: multivariate compositional data, self-expressive property, variable importance.

1. Introduction

Compositional data describe the relative information of components on a whole by proportions or percentages, occur frequently in practice [1]. For example, industry-level gross domestic product (GDP) and emotion structures in social media. Compositional data are constrained to be positive and constant-sum (e.g. 1, 100% or other constants), which contributes to the closure of compositional data.

Owing to the constraints, traditional statistical methods, such as clustering, can not be directly applied to the raw data [2]. A common approach is to eliminate these constraints by some transformation methods before modeling. Along this line, many compositional data clustering methods have been proposed. Xu et al. [3] discussed four hierarchical clustering methods utilizing two types of log-ritio transformations, pivot coordinates [4] and weighted symmetric pivot coordinates [5]. Godichon-Baggioni et al. [6] excuted K-means clustering utilizing the centered log-ratio (clr) [7] transformation. Wang et al. [8] proposed the theoretical framework of the convex clustering for compositional data with respect to the Aitchison geometry using the isometric log-ratio (ilr) transformation. Although these methods have made great progress in the clustering of compositional data, there still exists some limitations. These approaches are limited to single compositional data variable, lacking the ability to deal with multivariate scenario, and the performances of them are heavily influenced by the selection of the metric functions, such as Cosine, Euclidean distance and Gaussian function. Moreover, how to distinguish the importance of different compositional data variables is also a crucial issue for clustering.

To address these problems, we propose a group-weighted subspace clustering for multivariate compositional data termed GWSCMC, which integrates the subspace clustering with the group-weighted strategy. Specifically, subspace clustering [9] is generalized to multivariate compositional data clustering since it is widely used and does not depend upon the metric functions. The key to the success of subspace

[+] Corresponding author.
E-mail address: qing.zhao@ustb.edu.cn.

clustering is the self-expressive property which assumes that all samples belong to different subspace and each sample can be linearly represented by other samples within the same subspace [10]. We first construct the self-expressive property for multivariate compositional data. Then, the coefficient matrix is obtained under the self-expressive property and the scaled simplex constraints are imposed on it to make it physically meaningful and more discriminative. Meanwhile, considering the varying importance of compositional data variables, we propose a group-weighted strategy to learn appropriate weights for these variables, reflecting the importance explicitly. Finally, the optimization problem is solved by the alternating direction method of multipliers (ADMM). Experimental results on the practical dataset demonstrate the effectiveness of our proposed method in comparison with other clustering methods.

2. Preliminaries

2.1. Aitchison Geometry and ilr Transformation

The D-part simplex space S^D which is constituted by all the D-part compositional data is defined as

$$S^D = \{\mathbf{u} = (u_1, u_2, \cdots, u_D)' \mid u_j > 0, \sum_{j=1}^{D} u_j = 1\}. \tag{1}$$

In the Aitchison geometry, the perturbation $\oplus$ and powering $\otimes$ are two important operations for compositional data. For any two compositions $u = (u_1, u_2, \cdots, u_D)'$ and $v = (v_1, v_2, \cdots, v_D)'$ in S^D, and $\alpha \in \mathbb{R}$,

$$\begin{aligned} \mathbf{u} \oplus \mathbf{v} &= \mathcal{C}(u_1 v_1, u_2 v_2, \cdots, u_D v_D)', \\ \alpha \otimes \mathbf{u} &= \mathcal{C}(u_1^{\alpha}, u_2^{\alpha}, \cdots, u_D^{\alpha})', \end{aligned} \tag{2}$$

where $\mathcal{C}(\cdot)$ denotes the closure operation, $\mathcal{C}(\mathbf{z}) = (\frac{z_1}{\mathrm{sum}z}, \frac{z_2}{sumz}, \cdots, \frac{z_D}{sumz})'$, $\mathbf{z} = (z_1, z_2, \cdots, z_D)' \in S^D$ and $sumz = \sum_{i=1}^{D} z_i$. Apparently, the perturbation and powering operations construct a linear space on S^D. The zero element corresponding to the perturbation is $\mathcal{C}(\mathbf{1}_D)$, where $\mathbf{1}_D \in \mathbb{R}^D$ is a vector with all elements equal to 1.

Moreover, the ilr transformation has been widely utilized to remove the constrains of compositional data, which is conducted via the orthonormal basis in S^D. Specifically, composition $\mathbf{u}$ can be transformed to a vector $ilr(\mathbf{u}) = \mathbf{u}^* = (u_1^*, u_2^*, \cdots, u_{D-1}^*)'$. For $i = 1, 2, \cdots, D-1$,

$$u_i^* = \frac{1}{\sqrt{(D-i)(D-i+1)}} \left(\sum_{j=1}^{D-i} \log u_j - \sqrt{\frac{D-i}{D-i+1}} \log u_{D-i+1} \right). \tag{3}$$

Owing to the isometry property of ilr transformation, the perturbation and powering in Aitchison geometry can be transformed into the addition and multiplication operations in real space. For any $\alpha, \beta \in \mathbb{R}$ and $\mathbf{u}, \mathbf{v} \in S^D$, we can obtain

$$ilr(\alpha \otimes \mathbf{u} \oplus \beta \otimes \mathbf{v}) = \alpha \times ilr(\mathbf{u}) + \beta \times ilr(\mathbf{v}). \tag{4}$$

2.2. Multivariate Compositional Data

In analogy with the multivariate scalar data in the Euclidean space, we can denote the p-dimensional D-part compositional data vector in S_p^D as $\mathbf{U}_i = (\mathbf{u}_{i1}', \mathbf{u}_{i2}', \cdots, \mathbf{u}_{ip}')'$, where $\mathbf{u}_{ij} = (u_{ij}^1, u_{ij}^2, \cdots, u_{ij}^D)' \in S^D$ for $i = 1, 2, \cdots, n$ and $j = 1, 2, \cdots, p$.

The Aitchison geometry and ilr transformation in S_p^D can be derived from univariate compositional data in S^D [11]. For any two multivariate compositional samples $\mathbf{U}_r = (\mathbf{u}_{r1}', \mathbf{u}_{r2}', \cdots, \mathbf{u}_{rp}')'$ and $\mathbf{U}_l = (\mathbf{u}_{l1}', \mathbf{u}_{l2}', \cdots, \mathbf{u}_{lp}')'$ in S_p^D, and $\alpha \in \mathbb{R}$, the perturbation $\oplus$ and powering $\otimes$ in S_p^D can be defined as

$$\begin{aligned} \mathbf{U}_r \oplus \mathbf{U}_l &= ((\mathbf{u}_{r1} \oplus \mathbf{u}_{l1})', (\mathbf{u}_{r2} \oplus \mathbf{u}_{l2})', \cdots, (\mathbf{u}_{rp} \oplus \mathbf{u}_{lp})')', \\ \alpha \otimes \mathbf{U}_r &= ((\alpha \otimes \mathbf{u}_{r1})', (\alpha \otimes \mathbf{u}_{r2})', \cdots, (\alpha \otimes \mathbf{u}_{rp})')'. \end{aligned} \tag{5}$$

The definition of ilr transformation for multivariate compositional data can be formulated as

$$ilr(\mathbf{U}_r) = \mathbf{U}_r^* = (ilr(\mathbf{u}_{r1})', ilr(\mathbf{u}_{r2})', \cdots, ilr(\mathbf{u}_{rp})')'. \tag{6}$$

3. Group-weighted Subspace Clustering for Multivariate Compositional Data

In this section, we first define the self-expressive property for multivariate compositional data. Then, based on this, we propose our model and develop an optimization method to solve it.

3.1. The Self-expressive Property for Multivariate Compositional Data

We first define the self-expressive property for multivariate compositional data. That is, all samples of multivariate compositional data are in a union of different subspaces which can be seen as clusters, and each sample can be represented as a linear combination of the other samples in the same subspace. This can be formulated as follows:

$$\mathbf{U}_i = b_{1i} \otimes \mathbf{U}_1 \oplus \cdots \oplus b_{ni} \otimes \mathbf{U}_n = \bigoplus_{j \neq i} b_{ji} \otimes \mathbf{U}_j, \tag{7}$$

where $\mathbf{U}_i = (\mathbf{u}_{i1}', \mathbf{u}_{i2}', \cdots, \mathbf{u}_{ip}')'$, $i = 1, 2, \cdots, n$ and $j = 1, 2, \cdots, n$. $\mathbf{b}_i = (b_{1i}, b_{2i}, \cdots, b_{ni})' \in \mathbb{R}^{n \times 1}$ is the coefficient vector corresponding to the sample $\mathbf{U}_i$ and $b_{ii} = 0$.

3.2. Model Formulation

To effectively solve the multiple linear regression problem in Eq.(7), we first simplify it by a proposition on account of ilr transformation.

Proposition 1: For any two multivariate compositional samples $\mathbf{U}_r$ and $\mathbf{U}_l$ in S_p^D, and $\alpha, \beta \in \mathbb{R}$,

$$ilr(\alpha \otimes \mathbf{U}_r \oplus \beta \otimes \mathbf{U}_l) = \alpha \times ilr(\mathbf{U}_r) + \beta \times ilr(\mathbf{U}_l). \tag{8}$$

By means of Proposition 1, if we perform ilr transformation on both sides of Eq.(7). The equation can be rewritten as follows:

$$ilr(\mathbf{U}_i) = b_{1i} ilr(\mathbf{U}_1) + \cdots + b_{ni} ilr(\mathbf{U}_n) = \sum_{j \neq i} b_{ji} ilr(\mathbf{U}_j), \tag{9}$$

To further simplify the notations, we can rewrite the above equation in the matrix form as follows:

$$\min_{\mathbf{B}} \| \mathbf{H} - \mathbf{HB} \|_F^2 + \lambda \| \mathbf{B} \|_F^2, \tag{10}$$

where $\mathbf{H} = ilr(\mathbf{U}) = (ilr(\mathbf{U}_1), ilr(\mathbf{U}_2), \cdots, ilr(\mathbf{U}_n)) \in \mathbb{R}^{d \times n}$ is the multivariate compositional data after ilr transformation and d is the number of dimensions. $\mathbf{B} = (\mathbf{b}_1, \cdots, \mathbf{b}_n) \in \mathbb{R}^{n \times n}$ denotes the coefficient matrix and the constraint $b_{ii} = 0$ is removed since the Frobenius norm imposed on $\mathbf{B}$ can avoid the trivial solution [12]. λ is a trade-off parameter.

Inspired by [12], To make $\mathbf{B}$ more physically meaningful and discriminative, the scaled simplex constraints are imposed on it, which make each coefficient vector non-negative and the vector sum equal to a constant s where $0 < s \leq 1$. As a result, the objective function can be formulated as:

$$\min_{\mathbf{B}} \| \mathbf{H} - \mathbf{HB} \|_F^2 + \lambda \| \mathbf{B} \|_F^2, s.t. \mathbf{B} \geq 0, \mathbf{B}' \mathbf{1}_n = s \mathbf{1}_n, \tag{11}$$

where $\mathbf{B} \geq 0$ denotes that all entries of $\mathbf{B}$ are non-negative and $\mathbf{1}_n \in \mathbb{R}^n$ represents a n-dimensional vector with all entries equal to 1.

Considering the fact that the importance of compositional data variables is different during the process of clustering, we introduce a weight matrix $\mathbf{W} = diag(\mathbf{W}^{(1)}, \mathbf{W}^{(2)}, \cdots, \mathbf{W}^{(p)})$ to explicitly reflect the importance of different compositional data variables, where $\mathbf{W}^{(k)}$ denotes the block diagonal weight matrix corresponding to the k-th compositional data variable, $k = 1, 2, \cdots, p$. Let d_k represents the number of parts of the k-th compositional data variable after ilr transformation, we have $\sum_{k=1}^{p} d_k = d$ and $diag(\mathbf{W}^{(k)}) = w_k \mathbf{1}_{d_k} \in \mathbb{R}^{d_k}$, where w_k is the weight of the k-th compositional data variable. To learn weights adaptively, we integrate the constraint $w_k \geq 0, \sum_{k=1}^{p} w_k = 1$ into Eq.(11). The objective function of GWSCMC can be written as

$$\min_{\mathbf{B}, \mathbf{W}} \| \mathbf{WH} - \mathbf{WHB} \|_F^2 + \lambda \| \mathbf{B} \|_F^2, s.t. \mathbf{B} \geq 0, \mathbf{B}' \mathbf{1}_n = s \mathbf{1}_n, w_k \geq 0, \sum_{k=1}^{p} w_k = 1. \tag{12}$$

3.3. Solving the Optimization with ADMM

The optimization problem shown in Eq.(12) can not obtain the analytical solution directly, so we solve it with ADMM. Specifically, we first introduce a new auxiliary variable $\mathbf{C}$, which serves to separate the objective function in Eq.(12), and then rewrite Eq.(12) in the following form:

$$\min_{\mathbf{B},\mathbf{W}} \| \mathbf{WH} - \mathbf{WHC} \|_F^2 + \lambda \| \mathbf{B} \|_F^2, s.t. \mathbf{C} = \mathbf{B}, \mathbf{C} \geq 0, \mathbf{C}'\mathbf{1}_n = s\mathbf{1}_n, w_k \geq 0, \sum_{k=1}^{p} w_k = 1. \tag{13}$$

The corresponding augmented Lagrangian function is formulated as

$$\mathcal{L}(\mathbf{B}, \mathbf{C}, \mathbf{W}) = \| \mathbf{WH} - \mathbf{WHB} \|_F^2 + \lambda \| \mathbf{C} \|_F^2 + tr(\mathbf{Y}^T(\mathbf{C} - \mathbf{B})) + \frac{\mu}{2} \| \mathbf{C} - \mathbf{B} \|_F^2, \tag{14}$$

where $\mathbf{Y}$ is the Lagrange multiplier and μ denotes an adaptive balance parameter. According to ADMM, we can iteratively update each of $\mathbf{B}$, $\mathbf{C}$ and $\mathbf{W}$ while keeping the other two variables fixed.

1) Updating $\mathbf{B}$ while fixing $\mathbf{C}$ and $\mathbf{W}$

By taking the derivative of Eq.(14) with respect to $\mathbf{B}$ and setting it to be zero, $\mathbf{B}$ is given by

$$\mathbf{B} = (2\mathbf{M}^T\mathbf{M} + \mu\mathbf{I})^{-1}(2\mathbf{M}^T\mathbf{M} + \mu\mathbf{C} + \mathbf{Y}), \tag{15}$$

where $\mathbf{M} = \mathbf{WH}$ and $\mathbf{I}$ is the identity matrix.

2) Updating $\mathbf{C}$ while fixing $\mathbf{B}$ and $\mathbf{W}$

$$\mathbf{C} = \arg\min_{\mathbf{C}} \frac{1}{2} \| \mathbf{C} - \frac{\mu\mathbf{B} - \mathbf{Y}}{\mu + 2\lambda} \|_F^2, s.t. \mathbf{C} \geq 0, \mathbf{C}'\mathbf{1}_n = s\mathbf{1}_n. \tag{16}$$

This is a quadratic problem and can be effectively solved by the method based on projection.

3) Updating $\mathbf{W}$ while fixing $\mathbf{B}$ and $\mathbf{C}$

We divide the data matrix $\mathbf{H}$ according to the variables and have $\mathbf{H} = (\mathbf{H}_{(1)}', \mathbf{H}_{(2)}', \cdots, \mathbf{H}_{(p)}')' \in \mathbb{R}^{d \times n}$ where $\mathbf{H}_{(k)} \in \mathbb{R}^{d_k \times n}$ is the k-th compositional data variable, for $k = 1, 2, \cdots, p$. Then, the optimization problem according to $\mathbf{W}$ is

$$w_k = \arg\min_{w_k} \sum_{k=1}^{p} w_k^2 \| \mathbf{H}_{(k)} - \mathbf{H}_{(k)}\mathbf{B} \|_F^2, s.t. w_k \geq 0, \sum_{k=1}^{p} w_k = 1. \tag{17}$$

Eq.(17) can be solved by the Lagrange multiplier method. Suppose $g_k = \| \mathbf{H}_{(k)} - \mathbf{H}_{(k)}\mathbf{B} \|_F^2$, the solution is

$$w_k = 1 \Big/ \sum_{l=1}^{p} \frac{g_k}{g_l}. \tag{18}$$

4) Updating $\mathbf{Y}$ and μ

We update the lagrangian multiplier $\mathbf{Y}$ and penalty parameter μ by the following formulations:

$$\begin{aligned} \mathbf{Y} &= \mathbf{Y} + \mu(\mathbf{C} - \mathbf{B}), \\ \mu &= \min(\eta\mu, \mu_{max}), \end{aligned} \tag{19}$$

where η and μ_{max} are two predefined positive constants, and η controls the speed of convergence.

We repeat the above updates until the convergence criteria is satisfied. Then the affinity matrix $\mathbf{A}$ can be constructed by $\mathbf{A} = (\mathbf{B} + \mathbf{B}')/2$, which is utilized for spectral clustering to obtain the final clustering result.

4. Real Data Analysis

In this section, we conduct comparative experiments on practical datasets to validate the effectiveness of GWSCMC. The subsequent subsections provide the dataset description, comparison methods and the experimental results.

4.1. Dataset

The Beer Reviews from Beer Advocate dataset is utilized to verify our method. This dataset contains 1.5 million beer reviews and each review includes rating in terms of appearance, aroma, palate, taste, and overall impression. Moreover, reviews also include product information, such as the style of beer and the alcohol by volume of the beer. In our experiment, we integrate all the beer reviews for different styles of beers. As a result, the dataset is simplified to contain 104 beer styles with five compositional data variables termed appearance, aroma, palate, taste and abv (alcohol by volume). The first four variables all include nine parts and the last variable contains ten parts. To test the performance of clustering intuitively, we compute the mean value of the overall impression rating for each style of beer. Then, we divide the mean value into three intervals which are below 3.6, (3.6, 3.8], and exceed 3.8. Furthermore, each interval can be seen as a group

so the true cluster labels are known. The labels learned from the algorithms will be compared to the true labels to evaluate the clustering results. Three commonly used metrics are employed to evaluate the clustering performance, i.e. the normalized mutual information (NMI), the adjust rand index (ARI), and the accuracy (ACC). Note that larger values of these metrics means better clustering performance.

4.2. Comparison Methods

We compare the proposed GWSCMC with several state-of-the-art clustering methods including K-means, Spectral Clustering, Hierarchical Clustering, least square regression (LSR) subspace clustering [10], and scaled simplex representation (SSR) subspace clustering [12]. The first three methods are traditional clustering methods which mainly utilize the distance to measure the similarity among samples. LSR and SSR are two effective subspace clustering methods which are constructed based on the self-expressive property. It is worth noting that these methods are always utilized to deal with single compositional data variable, but not designed for multivariate compositional data. As a result, these compositional data variables after ilr transformation are concatenated as a whole before using these methods. For GWSCMC, the optimal parameters λ and s are determined by a grid search strategy. In other words, these two parameters are selected from candidate sets. The candidate set for λ is $[10^{-4},10^{-3},10^{-2},10^{-1},1,10,10^{2}]$, while s takes values from [0.1, 0.2, …, 0.9, 1].

4.3. Results

To evaluate the performance of different clustering algorithms, each method is executed for 10 repetitions on the dataset, and then the mean value and standard deviation of NMI, ARI, and ACC are calculated respectively. All experimental results are presented in Table 1. From Table 1, we can see that the clustering results of K-means and spectral clustering are inferior to the self-expressive property based methods LSR, SSR and GWSCMC. Hierarchical clustering method is better than LSR but not as good as SSR and GWSCMC. Comparing to LSR, SSR can mitigate the noise effect to obtain a sparse and accurate subspace representation, resulting in a better performance. Overall, GWSCMC outperforms all the other clustering methods. The performance gain of GWSCMC over SSR demonstrates the power of the group-weighted strategy which distinguishes the importance of different compositional data variables.

Table 1: Clustering results of the compared methods

Methods	NMI (std)	ARI (std)	ACC (std)
K-means	0.1711 (0.0311)	0.1119(0.0263)	0.4865 (0.0458)
Spectral Clustering	0.1496 (0.0000)	0.0887 (0.0000)	0.4519 (0.0000)
Hierarchical	0.3185 (0.0000)	0.2050 (0.0000)	0.6154 (0.0000)
LSR	0.2283 (0.0051)	0.1649 (0.0022)	0.4962 (0.0050)
SSR	0.3382 (0.0191)	0.4124 (0.0212)	0.7298 (0.0153)
GWSCMC	0.5668 (0.0028)	0.5614 (0.0036)	0.8462 (0.0000)

In addition, the weights learned by the proposed GWSCMC on the dataset are illustrated in Fig.1. Clearly, the taste variable gains the largest weight with around 0.30, followed by the palate and aroma variables with around 0.24 and 0.21. The weights of appearance and abv variables are relatively small with around 0.16 and 0.09. In fact, this is not against our instincts. The taste, palate and aroma are the key to the overall impression of beers. By contrast, the appearance and abv are relatively not critical.

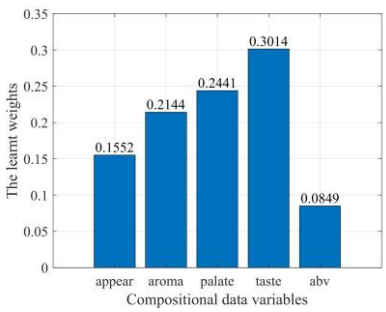

Fig. 1: The weights learned by GWSCMC for compositional data variables.

5. Conclusion

This paper focus on the clustering of multivariate compositional data. We propose an effective clustering method which utilizes the self-expressive property and the group-weighted strategy. Specifically, different from the metric based methods which suffer from the selection of the metric functions, the proposed method explores the self-expressive property for multivariate compositional data to find the relationships among samples. Moreover, to intuitively reflect the importance of each compositional data variable with different numbers of parts, we propose a group-weighted strategy to adaptively learn appropriate weights for different variables. From the practical case study, it is shown that our proposed method can achieve good performance and when the importance of variables are distinguished, the clustering results are significantly improved.

6. References

[1] Alenazi A. A review of compositional data analysis and recent advances[J]. Communications in Statistics-Theory and Methods, 2023, 52(16): 5535-5567.

[2] Lu S, Wang W, Guan R. Kent feature embedding for classification of compositional data with zeros[J]. Statistics and Computing, 2024, 34(2): 69.

[3] Xu N, Xu C, Finkelman R B, et al. Coal elemental (compositional) data analysis with hierarchical clustering algorithms[J]. International Journal of Coal Geology, 2022, 249: 103892.

[4] Egozcue J J, Pawlowsky-Glahn V, Mateu-Figueras G, et al. Isometric logratio transformations for compositional data analysis[J]. Mathematical geology, 2003, 35(3): 279-300.

[5] Hron K, Engle M, Filzmoser P, et al. Weighted symmetric pivot coordinates for compositional data with geochemical applications[J]. Mathematical Geosciences, 2021, 53: 655-674.

[6] Godichon-Baggioni A, Maugis-Rabusseau C, Rau A. Clustering transformed compositional data using K-means, with applications in gene expression and bicycle sharing system data[J]. Journal of Applied Statistics, 2019, 46(1): 47-65.

[7] Aitchison J. The statistical analysis of compositional data[J]. Journal of the Royal Statistical Society: Series B (Methodological), 1982, 44(2): 139-160.

[8] Wang X, Wang H, Wang S, et al. Convex clustering method for compositional data via sparse group lasso[J]. Neurocomputing, 2021, 425: 23-36.

[9] Qu W, Xiu X, Chen H, et al. A survey on high-dimensional subspace clustering[J]. Mathematics, 2023, 11(2): 436.

[10] Lu C Y, Min H, Zhao Z Q, et al. Robust and efficient subspace segmentation via least squares regression[C]//Computer Vision–ECCV 2012: 12th European Conference on Computer Vision, Florence, Italy, October 7-13, 2012, Proceedings, Part VII 12. Springer Berlin Heidelberg, 2012: 347-360.

[11] Wang H, Shangguan L, Wu J, et al. Multiple linear regression modeling for compositional data[J]. Neurocomputing, 2013, 122: 490-500.

[12] Xu J, Yu M, Shao L, et al. Scaled simplex representation for subspace clustering[J]. IEEE Transactions on Cybernetics, 2019, 51(3): 1493-1505.