

基于多样性和谱嵌入的张量 多视图子空间聚类

张沙沙, 王长鹏

(长安大学 理学院, 西安 710064)

摘要: 针对如何有效利用多视图的多样性信息和高阶信息, 并建立系数矩阵的学习过程与谱聚类之间联系的问题, 提出一种基于多样性和谱嵌入的张量多视图子空间聚类算法. 首先, 在自表示张量部分采用张量自适应对数行列式正则化, 从而能根据奇异值的大小自适应地选择逼近函数. 其次, 采用 Hilbert-Schmidt 独立准则衡量多样性, 以确保不同视图的系数表示矩阵具有足够的多样性. 再次, 为避免谱聚类过程的独立进行, 将其引入模型中联合学习, 使低秩张量学习、多样性学习和谱嵌入学习在一个统一的框架内进行. 最后, 通过在 5 个真实数据集上与 10 种优秀算法进行比较, 验证了该算法在提升聚类性能方面的有效性.

关键词: 多视图子空间聚类; 张量自适应对数行列式; 多样性; 谱嵌入; Hilbert-Schmidt 独立准则

中图分类号: TP391.4 **文献标志码:** A **文章编号:** 1671-5489(2025)02-0499-14

Tensor Multi-view Subspace Clustering Based on Diversity and Spectral Embedding

ZHANG Shasha, WANG Changpeng

(School of Sciences, Chang'an University, Xi'an 710064, China)

Abstract: Aiming at how to effectively utilize the diversity information and higher-order information of multi-views, and establish the connection between the learning process of coefficient matrices and spectral clustering, we proposed a tensor multi-view subspace clustering algorithm based on diversity and spectral embedding. Firstly, the algorithm used tensor adaptive log-determinant regularization in the self-representation tensor part, which could adaptively select the approximation function according to the size of the singular values. Secondly, the Hilbert-Schmidt independence criterion was used to measure diversity to ensure that the coefficient representation matrices from different views exhibited sufficient diversity. Thirdly, in order to avoid the spectral clustering process to be performed independently, it was introduced into the model for joint learning, so that the low-rank tensor learning, diversity learning and spectral embedding learning were performed in a unified framework. Finally, the effectiveness of the algorithm in improving the clustering performance was verified by comparing it with ten excellent algorithms on five real datasets.

收稿日期: 2024-04-17.

第一作者简介: 张沙沙(2000—), 女, 汉族, 硕士研究生, 从事机器学习的研究, E-mail: 3212146414@qq.com. **通信作者简介:** 王长鹏(1985—), 男, 汉族, 博士, 副教授, 从事机器学习的研究, E-mail: cpwang@chd.edu.cn.

基金项目: 国家自然科学基金(批准号: 12471480)和长安大学中央高校基本科研业务费专项基金(批准号: 300102122101).

Keywords: multi-view subspace clustering; tensor adaptive log-determinant; diversity; spectral embedding; Hilbert-Schmidt independence criterion

0 引言

聚类广泛应用于数据挖掘、机器学习等多领域,旨在将数据集分成几个由类似数据对象构成的组.多视图数据是对同一事物由不同方式进行描述得到的数据,这些数据在本质上有一定的相关性和多样性.多视图数据在机器学习中应用广泛,采用传统的单视图聚类算法处理这些数据易忽略视图之间的内部联系,导致聚类性能难以提高,所以关于多视图聚类算法的研究备受关注.

近年来,已提出了许多多视图聚类算法^[1-4],这些算法主要分为 4 类:基于协同训练的多视图聚类算法、基于核学习的多视图聚类算法、基于图学习的多视图聚类算法以及基于子空间学习的多视图聚类算法.本文考虑基于子空间学习的多视图聚类算法,其原理是将原始的特征数据映射到其子空间中学习一个鲁棒系数矩阵,然后构造低秩的相似矩阵用于谱聚类.该方法保留了原始数据的某些相关信息和互补信息,能更好揭示数据的子空间结构,从而达到满意的聚类效果.

低秩表示子空间聚类(low-rank representation, LRR)^[5]和稀疏子空间聚类(sparse subspace clustering, SSC)^[6]是早期经典的两种子空间聚类算法,之后受此启发提出了一些基于自表示子空间学习的多视图聚类算法.例如:Luo 等^[7]提出了一种将一致性和特异性共同用于子空间表示学习的方法,很好地拟合了真实世界的数据集;Zhang 等^[8]通过一种潜在表示探索多视图的互补信息,以确保获得更准确的子空间表示;Chen 等^[9]充分利用了低秩表示技术和多样性正则化项全面探索多视图的多样性和一致性,将其应用于子空间聚类;张绎凡等^[10]考虑到不同视图之间的多样性信息,采用 Hilbert-Schmidt 独立准则(Hilbert-Schmidt independence criterion, HSIC)约束特定于视图的投影矩阵.上述方法都取得了很好的聚类效果,但都仅通过矩阵层面挖掘不同视图之间的内在联系,而基于张量可探索视图间更深层的互补信息.

基于张量的多视图子空间聚类算法先将多个视图的系数表示矩阵叠加成一个三阶张量,然后通过不同的张量约束方法恢复一个低秩张量,进一步采用线性融合方法从低秩张量获得一致相似矩阵进行谱聚类.例如,Zhang 等^[11]在子空间学习中引入一种独特的低秩张量约束,巧妙地捕捉了隐藏在多视图数据中的高阶相关性.但张量展开矩阵的核范数并不是 Tucker 低秩和 L_1 范数的凸松弛.因此,Lu 等^[12]提出了基于张量奇异值分解(tensor singular value decomposition, t-SVD)的张量核范数,这个新范数解决了 L_1 范数存在的凸松弛问题.受此启发,Xie 等^[13]提出了基于 t-SVD 的多视图子空间聚类算法,该算法能很好地表征多视图数据中嵌入的高阶相关信息.进一步,Gao 等^[14]考虑到奇异值之间的差异性,提出了基于加权张量核范数最小化的多视图子空间聚类算法.Sun 等^[15]将视图表示分解为低秩部分和多样化部分,分别采用不同的正则化项,综合考虑视图之间的一致性和多样性.这些方法在聚类性能方面都有很大提升,但大多数方法仍存在问题:1)视图的相似矩阵是多样化的,他们之间的内在差异未得到充分考虑,可能导致很难全面探索多视图数据中蕴含的信息;2)低秩张量和聚类指标矩阵的学习过程相互独立,忽略了他们之间的依赖关系,可能导致很难获得后者的最优解.

基于以上分析,为更好地利用多视图数据中嵌入的互补信息和高阶信息,并避免谱聚类独立进行,本文提出一种基于多样性和谱嵌入的张量多视图子空间聚类算法(tensor multiview subspace clustering based on diversity and spectral embedding, TMSCDSE).该算法通过重组所有视图的自表示系数矩阵构造一个三阶张量,引入一种张量自适应对数行列式正则化更好地逼近张量的秩函数,并借助低秩张量实现多视图之间的内在一致性.此外,通过 Hilbert-Schmidt 独立准则有效地利用多视图之间的互补信息,加强子空间表示的多样性.最后,将低秩张量学习、多样性学习和谱嵌入学习在一个统一的框架中,得到一致相似矩阵对应的聚类指标矩阵,然后使用 k -means 算法获取最终的聚类结果.图 1 为 TMSCDSE 算法的整体框架.

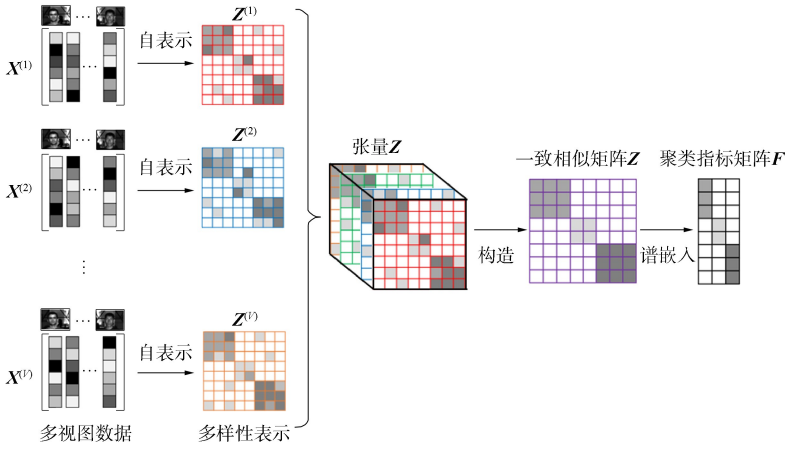


图 1 TMSCDSE 算法的整体框架

Fig. 1 Overall framework of TMSCDSE algorithm

1 预备知识

1.1 符号和定义

为方便, 先介绍本文使用的符号和定义. 对于一个三阶张量 $\mathbf{A} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$, 其第 i 个正面切片用矩阵 $\mathbf{A}^{(i)}$ 表示, 并且用张量 $\bar{\mathbf{A}} = \text{fft}(\mathbf{A}, [], 3)$ 表示沿着第三维的快速 Fourier 变换(FFT), 所以张量 \mathbf{A} 也可以通过 $\bar{\mathbf{A}}$ 的逆进行一系列变换获得(即 $\mathbf{A} = \text{ifft}(\bar{\mathbf{A}}, [], 3)$).

文献[13]提出了使用 t-SVD^[16]描述跨多视图的共同信息, 其定义如下.

定义 1(t-SVD)^[16] 若给定一个张量 $\mathbf{A} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$, 则 \mathbf{A} 可被分解为如下形式:

$$\mathbf{A} = \mathbf{USV}^T, \tag{1}$$

其中: $\mathbf{U} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$ 和 $\mathbf{V} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$ 是正交张量; $\mathbf{S} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$ 是一个 f 对角张量, 其每个切片都是一个对角矩阵.

通常采用张量核范数(t-SVD based tensor nuclear norm, t-TNN)^[17]松弛张量的秩函数. 而基于 t-TNN 的模型中采用 Fourier 域中奇异值的范数近似秩函数, 这种方法对较大奇异值收缩较大, 但较大奇异值通常表征嵌入在矩阵中的重要结构信息, 所以导致聚类效果不理想.

定义 2(t-TNN)^[17] 若给定一个张量 $\mathbf{A} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$, 则 \mathbf{A} 的张量核范数可被定义为如下形式:

$$\|\mathbf{A}\|_* = \frac{1}{n_3} \sum_{j=1}^{n_3} \text{tr} \sqrt{\mathbf{A}^{(j)\text{T}} \mathbf{A}^{(j)}} = \frac{1}{n_3} \sum_{i=1}^{\min\{n_1, n_2\}} \sum_{j=1}^{n_3} |\tilde{\mathbf{D}}(i, i, j)|, \tag{2}$$

其中矩阵 $\mathbf{A}^{(j)}$ 表示张量 \mathbf{A} 的第 j 个正面切片, $\text{tr}(\cdot)$ 表示矩阵的迹, $\tilde{\mathbf{D}} = \text{fft}(\mathbf{D}, [], 3)$ 通过在 Fourier 域中张量 \mathbf{A} 的奇异值分解获得.

文献[15]提出了一个张量自适应对数行列式正则化器(tensor adaptive log-determinant regularization, TALR)用于低秩张量部分, 该方法可以根据奇异值的差异自动选择合适的逼近函数, 其定义如下.

定义 3(TALR)^[15] 若给定一个张量 $\mathbf{A} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$, 则 \mathbf{A} 的张量自适应对数行列式正则化定义为

$$\|\mathbf{A}\|_{\text{TALR}} = \frac{1}{n_3} \sum_{j=1}^{n_3} \log \det(\Upsilon \sqrt{\mathbf{A}^{(j)\text{T}} \mathbf{A}^{(j)}} + \mathbf{I}) = \frac{1}{n_3} \sum_{i=1}^{\min\{n_1, n_2\}} \sum_{j=1}^{n_3} \log(\tilde{\mathbf{M}}(i, i, j) + 1), \tag{3}$$

其中: $\log \det(\cdot)$ 为对数行列式函数; $\tilde{\mathbf{D}}$ 可以通过 Fourier 域中 \mathbf{A} 的奇异值分解得到; $\Upsilon > 0$ 为自适应收缩因子, 用来保证正则化的稳定性.

1.2 多视图子空间聚类

多视图子空间聚类是能有效处理高维数据的方法之一, 其旨在从原始的多视图数据中学习一个低维的嵌入子空间. 而基于自表示的模型是最常用的一种多视图子空间聚类模型, 一般可表述为

$$\begin{aligned} \min_{\mathbf{Z}^{(v)}, \mathbf{E}^{(v)}} \left\{ \sum_{v=1}^V \Omega(\mathbf{Z}^{(v)}) + \lambda \sum_{v=1}^V \Phi(\mathbf{E}^{(v)}) \right\}, \\ \text{s. t. } \mathbf{X}^{(v)} = \mathbf{X}^{(v)} \mathbf{Z}^{(v)} + \mathbf{E}^{(v)}, \quad v = 1, 2, \dots, V, \end{aligned} \quad (4)$$

其中 $\mathbf{X}^{(v)}$ 表示第 v 个视图的特征矩阵, $\mathbf{Z}^{(v)}$ 和 $\mathbf{E}^{(v)}$ 分别表示第 v 个视图的系数矩阵和误差矩阵, $\lambda > 0$ 表示平衡参数, $\Omega(\cdot)$ 为正则化项, $\Phi(\cdot)$ 为设计的损失函数. 所以不同的正则化和损失函数会构成不同的多视图子空间聚类方法. Tang 等^[18] 采用低秩表示模型学习共享的系数矩阵, 然后生成亲和图, 同时采用多样性正则化为每个视图学习最优的权重值, 以减少冗余并增强特定视图的特征. Zhang 等^[19] 提出了鲁棒的低秩核子空间聚类算法, 该算法可有效处理多视图数据中的复杂噪声和非线性结构问题. Chen 等^[20] 提出了一种广义的非凸低秩张量逼近方法, 明确地考虑到不同奇异值之间的显著差异性, 能更有效地探索不同视图之间的高阶相关性. 此外, 赵晓佳等^[21] 提出了一步张量学习的多视图子空间聚类方法, 联合学习表示张量和亲和矩阵.

1.3 多样性表示

不同视图的数据之间有一定差异, 这种差异构成了多视图数据的多样性信息. 而多视图子空间聚类的主要挑战之一就是充分利用数据之间的多样性和互补性信息, 获得全面的特征表示. 考虑到该问题, 张绎凡等^[10] 将 HSIC 作为多样性的衡量引入多视图子空间聚类模型, 这种策略尽可能地保留了视图的多样性信息, 并有效提高了聚类性能. HSIC 具有以下特征: 1) 将变量映射到再现 Hilbert 空间, 以测量变量之间的相关性, 达到可解决变量之间更复杂相关性的效果, 如可用于处理非线性相关的情况; 2) 在不估计随机变量联合分布的情况下, 也可以估计变量之间的相关性, 具有明显的计算优势; 3) 经验 HSIC 的结果等价于矩阵乘积的迹, 使问题更容易求解.

定义 4(HSIC)^[10] 给定张量 $\mathbf{Z} := \{(\mathbf{x}_1, \mathbf{y}_1), \dots, (\mathbf{x}_n, \mathbf{y}_n)\} \subseteq \mathbf{X} \times \mathbf{Y}$, 考虑从中得到一系列 n 个联合分布 p_{xy} 的独立观测数据, 经验 HSIC 记作 $\text{HSIC}(\mathbf{Z}, \mathbf{F}, \mathbf{G})$, 可表示为

$$\text{HSIC}(\mathbf{Z}, \mathbf{F}, \mathbf{G}) = (n-1)^{-2} \text{tr}(\mathbf{K}_1 \mathbf{H} \mathbf{K}_2 \mathbf{H}), \quad (5)$$

其中: \mathbf{F} 和 \mathbf{G} 是两个可分离的再生 Hilbert 空间; \mathbf{K}_1 和 \mathbf{K}_2 为 Gram 矩阵, 且 $k_{1,ij} = k_1(\mathbf{x}_i, \mathbf{x}_j)$, $k_{2,ij} = k_2(\mathbf{y}_i, \mathbf{y}_j)$, $k_1(\mathbf{x}_i, \mathbf{x}_j)$ 和 $k_2(\mathbf{y}_i, \mathbf{y}_j)$ 分别为 \mathbf{X} 和 \mathbf{Y} 上的核函数; $h_{ij} = \delta_{ij} - 1/n$ 使 Gram 矩阵中心化, 使其在特征空间中的均值为零.

1.4 谱嵌入学习

给定一个多视图数据集 $\mathbf{X} = (\mathbf{X}^1, \mathbf{X}^2, \dots, \mathbf{X}^V) \in \mathbb{R}^{d_n \times v}$, 其包含 V 个视图和 n 个样本. 一般地, 多视图谱聚类方法的目的是从数据空间中学习相应的 Laplace 矩阵. 因此, 谱嵌入学习的目标函数可表示为

$$\begin{aligned} \min_{\mathbf{F}} \text{tr}(\mathbf{F}^T \mathbf{L} \mathbf{F}), \\ \text{s. t. } \mathbf{F}^T \mathbf{F} = \mathbf{I}, \end{aligned} \quad (6)$$

其中: $\mathbf{L} = \mathbf{D} - \mathbf{W}$ 为 Laplace 矩阵; \mathbf{D} 为对角矩阵, 其对角元素为 $d_i = \sum_j \mathbf{W}_{ij}$, \mathbf{W} 是用于表示数据点之间相似性的相似矩阵; \mathbf{F} 为谱嵌入矩阵, 由 Laplace 矩阵 \mathbf{L} 的前 c 个最小特征值对应的特征向量组成, c 表示簇的个数.

2 TMSCDSE 算法

2.1 算法模型

给定多视图数据集 $\mathbf{X}^{(v)} = (\mathbf{X}_1^{(v)}, \mathbf{X}_2^{(v)}, \dots, \mathbf{X}_n^{(v)}) \in \mathbb{R}^{d_v \times n}$, $v = 1, 2, \dots, V$, 其中 d_v 和 n 分别表示第 v 个视图的特征维数和样本数量, v 表示视图数量. 受 LRR^[5] 的启发, 通过低秩表示在第 v 个视图中学习到的系数矩阵 $\mathbf{Z}^{(v)}$ 具有低秩结构特征, 所以从 V 个视图学习到的所有系数矩阵都具有相似的低秩结构. 因此, 由这些系数矩阵 $\mathbf{Z}^{(1)}, \mathbf{Z}^{(2)}, \dots, \mathbf{Z}^{(V)}$ 可构造一个具有低秩结构的张量 \mathbf{Z} . 为更好地利用低秩结构和嵌入在张量 \mathbf{Z} 中的高阶信息, 目前提出的大多数张量多视图聚类算法都是基于 t-SVD 实现的, 并取得了满意的聚类效果^[22-23].

现有的低秩张量约束方法通常平等地对待每个奇异值的贡献, 使用相同参数收缩所有的奇异值. 但在实际应用中, 矩阵的这些非零奇异值存在显著差异, 通常前几个较大的奇异值表征矩阵中的重要结构信息. 这种显著差异称为先验信息, 其对图像去噪、矩阵补全等都有重要意义. 为解决该问题, 受 TALR^[15] 的启发, 本文在低秩张量部分采用一种张量自适应对数行列式正则化器, 使奇异值的贡献被区别对待, 实现根据奇异值大小自适应选择合适的逼近函数. 最初的目标函数可表示为

$$\begin{aligned} & \min_{\mathbf{Z}^{(v)}, \mathbf{E}^{(v)}} \{ \|\mathbf{Z}\|_{\text{TALR}} + \lambda \|\mathbf{E}\|_{2,1} \}, \\ & \text{s. t. } \mathbf{X}^{(v)} = \mathbf{X}^{(v)} \mathbf{Z}^{(v)} + \mathbf{E}^{(v)}, \quad v = 1, 2, \dots, V, \\ & \quad \mathbf{E} = [\mathbf{E}^{(1)}; \mathbf{E}^{(2)}; \dots; \mathbf{E}^{(V)}], \\ & \quad \mathbf{Z} = \Phi(\mathbf{Z}^{(1)}, \mathbf{Z}^{(2)}, \dots, \mathbf{Z}^{(V)}), \end{aligned} \tag{7}$$

其中 $\Phi(\cdot)$ 表示将多视图的所有系数矩阵 $\mathbf{Z}^{(v)}$ 合并成一个三阶张量 \mathbf{Z} , $\|\mathbf{E}\|_{2,1}$ 表示对误差矩阵的稀疏约束, $\lambda > 0$ 用于平衡这两项. 为更好地捕捉视图之间的低秩特性, 降低计算复杂度, 需对构造的张量 \mathbf{Z} 进行旋转操作. 通过式(7)获得最优的系数矩阵, 再构造出相似矩阵用于谱聚类得出聚类结果.

通常不同视图包含固有的多样化信息, 促成了多视图数据的多样性. 因此, 为探索嵌入在多视图数据中的互补信息, 本文引入 HSIC 作为一种相关性的度量方式, 通过 HSIC 对成对的系数矩阵进行约束. 目标函数进一步表示为

$$\begin{aligned} & \min_{\mathbf{Z}^{(v)}, \mathbf{E}^{(v)}} \left\{ \|\mathbf{Z}\|_{\text{TALR}} + \lambda \|\mathbf{E}\|_{2,1} + \frac{\beta}{2} \sum_{v \neq w} \text{HSIC}(\mathbf{Z}^{(v)}, \mathbf{Z}^{(w)}) \right\}, \\ & \text{s. t. } \mathbf{X}^{(v)} = \mathbf{X}^{(v)} \mathbf{Z}^{(v)} + \mathbf{E}^{(v)}, \quad v = 1, 2, \dots, V, \\ & \quad \mathbf{E} = [\mathbf{E}^{(1)}; \mathbf{E}^{(2)}; \dots; \mathbf{E}^{(V)}], \\ & \quad \mathbf{Z} = \Phi(\mathbf{Z}^{(1)}, \mathbf{Z}^{(2)}, \dots, \mathbf{Z}^{(V)}). \end{aligned} \tag{8}$$

此外, 现有的大多数聚类方法都将低秩张量学习和谱聚类分两步进行, 而聚类指标矩阵依赖于相似矩阵, 如果分开求解不利于得到最优的聚类指标矩阵, 导致其很难较好地表征聚类结构. 因此, 本文提出将低秩张量学习和谱聚类在一个统一的框架中进行联合优化, 得到 TMSCDSE 算法的目标函数为

$$\begin{aligned} & \min_{\mathbf{Z}^{(v)}, \mathbf{E}^{(v)}, \mathbf{F}} \left\{ \|\mathbf{Z}\|_{\text{TALR}} + \lambda \|\mathbf{E}\|_{2,1} + \frac{\beta}{2} \sum_{v \neq w} \text{HSIC}(\mathbf{Z}^{(v)}, \mathbf{Z}^{(w)}) + 2\alpha \text{tr}(\mathbf{F}^T \mathbf{L}_{\tilde{\mathbf{Z}}} \mathbf{F}) \right\}, \\ & \text{s. t. } \mathbf{X}^{(v)} = \mathbf{X}^{(v)} \mathbf{Z}^{(v)} + \mathbf{E}^{(v)}, \quad v = 1, 2, \dots, V, \quad \mathbf{F}^T \mathbf{F} = \mathbf{I}, \\ & \quad \mathbf{E} = [\mathbf{E}^{(1)}; \mathbf{E}^{(2)}; \dots; \mathbf{E}^{(V)}], \\ & \quad \mathbf{Z} = \Phi(\mathbf{Z}^{(1)}, \mathbf{Z}^{(2)}, \dots, \mathbf{Z}^{(V)}), \end{aligned} \tag{9}$$

其中: $\mathbf{L}_{\tilde{\mathbf{Z}}} = \mathbf{D}_{\tilde{\mathbf{Z}}} - \tilde{\mathbf{Z}}$ 为 Laplace 矩阵, $\tilde{\mathbf{Z}} = \frac{1}{V} \sum_{v=1}^V [(|\mathbf{Z}^{(v)}| + |\mathbf{Z}^{(v)}|^T)/2]$ 为构造的一致相似矩阵, $\mathbf{D}_{\tilde{\mathbf{Z}}}$ 是一个对角矩阵, 其对角项为 $\mathbf{D}_{\tilde{\mathbf{Z}}}(i, i) = \sum_j (\tilde{\mathbf{Z}}_{ij} + \tilde{\mathbf{Z}}_{ji})$; $\mathbf{F} \in \mathbb{R}^{c \times N}$ 表示聚类指标矩阵, c 表示聚类个数; α 和 β 为两个平衡参数. 式(9)不仅充分利用了嵌入在多视图中的高阶信息和多样性信息, 而且将低秩张量学习和谱嵌入学习同时进行, 使获得的聚类指标矩阵能更有效地表征聚类结构, 达到提高多视图聚类性能的目的.

2.2 算法优化

采用增广 Lagrange 乘子算法解决上述优化问题, 即更新某个变量时固定其他变量. 在优化前, 先引入一个辅助张量变量 $\mathbf{J} \in \mathbb{R}^{n \times n \times V}$, 并且令 $\mathbf{J} = \mathbf{Z}$, 则式(9)的增广 Lagrange 函数为

$$\begin{aligned} & \mathbf{L} = (\mathbf{Z}^{(1)}, \mathbf{Z}^{(2)}, \dots, \mathbf{Z}^{(V)}; \mathbf{E}^{(1)}, \mathbf{E}^{(2)}, \dots, \mathbf{E}^{(V)}; \mathbf{J}; \mathbf{F}) = \\ & \quad \|\mathbf{J}\|_{\text{TALR}} + \lambda \|\mathbf{E}\|_{2,1} + \frac{\beta}{2} \sum_{v \neq w} \text{HSIC}(\mathbf{Z}^{(v)}, \mathbf{Z}^{(w)}) + 2\alpha \text{tr}(\mathbf{F}^T \mathbf{L}_{\tilde{\mathbf{Z}}} \mathbf{F}) + \\ & \quad \sum_{v=1}^V (\langle \mathbf{Y}^{(v)}, \mathbf{X}^{(v)} - \mathbf{X}^{(v)} \mathbf{Z}^{(v)} - \mathbf{E}^{(v)} \rangle + \frac{\mu}{2} \|\mathbf{X}^{(v)} - \mathbf{X}^{(v)} \mathbf{Z}^{(v)} - \mathbf{E}^{(v)}\|_{\text{F}}^2) + \end{aligned}$$

$$\langle \mathbf{W}, \mathbf{Z} - \mathbf{J} \rangle + \frac{\rho}{2} \|\mathbf{Z} - \mathbf{J}\|_{\mathbb{F}}^2, \quad (10)$$

其中矩阵 $\mathbf{Y}^{(v)}$ 和张量 \mathbf{W} 为两个 Lagrange 乘子, μ 和 ρ 为惩罚参数.

2.2.1 更新 $\mathbf{Z}^{(v)}$

固定变量 \mathbf{E}, \mathbf{F} 和 \mathbf{J} , 由于 $\varphi_{(v)}^{-1}(\mathbf{J}) = \mathbf{J}^{(v)}$, $\varphi_{(v)}^{-1}(\mathbf{W}) = \mathbf{W}^{(v)}$, 因此模型(10)可进一步变为

$$\begin{aligned} & \operatorname{argmin}_{\mathbf{Z}^{(v)}} \left\{ \frac{\beta}{2} \sum_{v \neq w} \text{HSIC}(\mathbf{Z}^{(v)}, \mathbf{Z}^{(w)}) + 2\alpha \operatorname{tr}(\mathbf{F}^T \mathbf{L}_2 \mathbf{F}) + \langle \mathbf{W}, \mathbf{Z} - \mathbf{J} \rangle + \frac{\rho}{2} \|\mathbf{Z} - \mathbf{J}\|_{\mathbb{F}}^2 + \right. \\ & \left. \sum_{v=1}^V (\langle \mathbf{Y}^{(v)}, \mathbf{X}^{(v)} - \mathbf{X}^{(v)} \mathbf{Z}^{(v)} - \mathbf{E}^{(v)} \rangle + \frac{\mu}{2} \|\mathbf{X}^{(v)} - \mathbf{X}^{(v)} \mathbf{Z}^{(v)} - \mathbf{E}^{(v)}\|_{\mathbb{F}}^2) \right\} = \\ & \operatorname{argmin}_{\mathbf{Z}^{(v)}} \left\{ \sum_{v=1}^V \left(\frac{\mu}{2} \|\mathbf{X}^{(v)} - \mathbf{X}^{(v)} \mathbf{Z}^{(v)} - \mathbf{E}^{(v)} + \frac{\mathbf{Y}^{(v)}}{\mu}\|_{\mathbb{F}}^2 + \frac{\rho}{2} \|\mathbf{Z}^{(v)} - \mathbf{J}^{(v)} + \frac{\mathbf{W}^{(v)}}{\rho}\|_{\mathbb{F}}^2 \right) + \right. \\ & \left. \frac{\beta}{2} \sum_{v \neq w} \text{HSIC}(\mathbf{Z}^{(v)}, \mathbf{Z}^{(w)}) + 2\alpha \operatorname{tr}(\mathbf{F}^T \mathbf{L}_2 \mathbf{F}) \right\}. \quad (11) \end{aligned}$$

定义 $\mathbf{P} = [\mathbf{P}_1, \dots, \mathbf{P}_j, \dots, \mathbf{P}_N]$, 其中 $\mathbf{P}_j = [\|\mathbf{F}_1 - \mathbf{F}_j\|_2^2; \dots; \|\mathbf{F}_N - \mathbf{F}_j\|_2^2]$, \mathbf{F}_j 是 \mathbf{F} 的第 j 行, 则有如下变换:

$$\begin{aligned} 2\operatorname{tr}(\mathbf{F}^T \mathbf{L}_2 \mathbf{F}) &= \operatorname{tr}(\mathbf{P}^T \tilde{\mathbf{Z}}) = \operatorname{tr}\left(\mathbf{P}^T \left(\frac{1}{V} \sum_{v=1}^V \frac{|\mathbf{Z}^{(v)}| + |\mathbf{Z}^{(v)}|^T}{2}\right)\right) = \\ & \frac{1}{2V} \sum_{v=1}^V \operatorname{tr}(\mathbf{P}^T |\mathbf{Z}^{(v)}| + \mathbf{P}^T |\mathbf{Z}^{(v)}|^T). \quad (12) \end{aligned}$$

在计算 $\mathbf{Z}^{(v)}$ 时, 矩阵 $\mathbf{Z}^{(w)}$ ($v \neq w$) 是固定的. 为方便求解, 采用 HSIC 的内积核, 即 $\mathbf{K}^{(v)} = \mathbf{Z}^{(v)T} \mathbf{Z}^{(v)}$, 因此 HSIC 可以写成:

$$\sum_{v=1, v \neq w}^V \text{HSIC}(\mathbf{Z}^{(v)}, \mathbf{Z}^{(w)}) = \sum_{v=1, v \neq w}^V \operatorname{tr}(\mathbf{H} \mathbf{K}^{(v)} \mathbf{H} \mathbf{K}^{(w)}) = \sum_{v=1, v \neq w}^V \operatorname{tr}(\mathbf{Z}^{(v)} \mathbf{H} \mathbf{K}^{(w)} \mathbf{H} \mathbf{Z}^{(v)T}) = \operatorname{tr}(\mathbf{Z}^{(v)} \mathbf{K} \mathbf{Z}^{(v)T}), \quad (13)$$

其中 $\mathbf{K} = \sum_{v=1, v \neq w}^V \mathbf{H} \mathbf{K}^{(v)} \mathbf{H}$.

结合式(12)和式(13), 可进一步将模型(11)变为

$$\begin{aligned} & \operatorname{argmin}_{\mathbf{Z}^{(v)}} \left\{ \frac{\beta}{2} \operatorname{tr}(\mathbf{Z}^{(v)} \mathbf{K} \mathbf{Z}^{(v)T}) + \frac{\alpha}{2V} \operatorname{tr}(\mathbf{P}^T |\mathbf{Z}^{(v)}| + \mathbf{P}^T |\mathbf{Z}^{(v)}|^T) + \right. \\ & \left. \frac{\mu}{2} \|\mathbf{X}^{(v)} - \mathbf{X}^{(v)} \mathbf{Z}^{(v)} - \mathbf{E}^{(v)} + \frac{\mathbf{Y}^{(v)}}{\mu}\|_{\mathbb{F}}^2 + \frac{\rho}{2} \|\mathbf{Z}^{(v)} - \mathbf{J}^{(v)} + \frac{\mathbf{W}^{(v)}}{\rho}\|_{\mathbb{F}}^2 \right\}, \quad (14) \end{aligned}$$

再关于 $\mathbf{Z}^{(v)}$ 求偏导, 并使求导结果为 0, 则得到的最优解满足下式:

$$\mathbf{L}^{(v)} \mathbf{Z}^{(v)} + \mathbf{Z}^{(v)} \mathbf{R}^{(v)} = \mathbf{C}^{(v)}, \quad (15)$$

其中

$$\begin{aligned} \mathbf{L}^{(v)} &= \rho \mathbf{I} + \mu \mathbf{X}^{(v)T} \mathbf{X}^{(v)}, \quad \mathbf{R}^{(v)} = \beta \mathbf{K}, \\ \mathbf{C}^{(v)} &= \mu \mathbf{X}^{(v)T} \mathbf{X}^{(v)} + \mathbf{X}^{(v)T} \mathbf{Y}^{(v)} + \rho \mathbf{J}^{(v)} - \mu \mathbf{X}^{(v)T} \mathbf{E}^{(v)} - \mathbf{W}^{(v)} - \\ & \frac{\alpha}{2V} (\mathbf{P} \cdot \operatorname{sign}(\mathbf{Z}_{i-1}^{(v)}) + \mathbf{P}^T \cdot \operatorname{sign}(\mathbf{Z}_{i-1}^{(v)}))^T. \quad (16) \end{aligned}$$

上述方程是具有唯一解的标准 Sylvester 方程^[24], 所以可以有效求解.

2.2.2 更新 $\mathbf{E}^{(v)}$

固定变量 $\mathbf{Z}^{(v)}, \mathbf{F}$ 和 \mathbf{J} , 则有

$$\begin{aligned} & \operatorname{argmin}_{\mathbf{E}} \left\{ \lambda \|\mathbf{E}\|_{2,1} + \sum_{v=1}^V \langle \mathbf{Y}^{(v)}, \mathbf{X}^{(v)} - \mathbf{X}^{(v)} \mathbf{Z}^{(v)} - \mathbf{E}^{(v)} \rangle + \sum_{v=1}^V \frac{\mu}{2} \|\mathbf{X}^{(v)} - \mathbf{X}^{(v)} \mathbf{Z}^{(v)} - \mathbf{E}^{(v)}\|_{\mathbb{F}}^2 \right\} = \\ & \operatorname{argmin}_{\mathbf{E}} \left\{ \frac{\lambda}{\mu} \|\mathbf{E}\|_{2,1} + \frac{1}{2} \|\mathbf{E} - \mathbf{D}\|_{\mathbb{F}}^2 \right\}, \quad (17) \end{aligned}$$

其最优解为

$$\mathbf{E}_{:,i}^* = \begin{cases} \frac{\|\mathbf{D}_{:,i}\|_2 - \lambda/\mu}{\|\mathbf{D}_{:,i}\|_2} \mathbf{D}_{:,i}, & \|\mathbf{D}_{:,i}\|_2 > \frac{\lambda}{\mu}, \\ 0, & \text{其他,} \end{cases} \quad (18)$$

其中 $\mathbf{D}_{:,i}$ 表示 \mathbf{D} 的第 i 列, $\mathbf{D} = (\mathbf{D}^1, \dots, \mathbf{D}^V)$, $\mathbf{D}^j = \mathbf{X}^{(j)} - \mathbf{X}^{(j)} \mathbf{Z}^{(j)} + \frac{1}{\mu} \mathbf{Y}^{(j)}$, $j = 1, 2, \dots, V$.

2.2.3 更新 \mathbf{J}

固定变量 $\mathbf{Z}^{(v)}$, $\mathbf{E}^{(v)}$ 和 \mathbf{F} , 可通过以下模型更新 \mathbf{J} :

$$\begin{aligned} \mathbf{J}^* &= \operatorname{argmin}_{\mathbf{J}} \left\{ \|\mathbf{J}\|_{\text{TALR}} + \langle \mathbf{W}, \mathbf{Z} - \mathbf{J} \rangle + \frac{\rho}{2} \|\mathbf{Z} - \mathbf{J}\|_{\text{F}}^2 \right\} = \\ &\operatorname{argmin}_{\mathbf{J}} \left\{ \frac{1}{\rho} \|\mathbf{J}\|_{\text{TALR}} + \frac{1}{2} \left\| \mathbf{J} - \left(\mathbf{Z} + \frac{1}{\rho} \mathbf{W} \right) \right\|_{\text{F}}^2 \right\}, \end{aligned} \quad (19)$$

为解决上述最小化问题, 首先引入下列定理, 并在本文算法中总结变量 \mathbf{J} 的更新过程.

定理 1^[15] 假设 $\tau > 0$, 张量 $\mathbf{X}, \mathbf{Y} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$, 并且 \mathbf{Y} 的 t-SVD 分解为 $\mathbf{Y} = \mathbf{U} \mathbf{D} \mathbf{V}^T$. 则对于模型

$$\operatorname{argmin}_{\mathbf{X}} \left\{ \tau \|\mathbf{X}\|_{\text{TALR}} + \frac{1}{2} \|\mathbf{X} - \mathbf{Y}\|_{\text{F}}^2 \right\}, \quad (20)$$

其最优解为 \mathbf{X}^* , 并且 $\mathbf{X}^* = \mathbf{U} \mathbf{D}_x \mathbf{V}^T$. 其中 $\mathbf{U} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$ 和 $\mathbf{V} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$ 是正交张量, $\mathbf{D}_x \in \mathbb{R}^{n_1 \times n_2 \times n_3}$ 是一个 f 对角张量. 当 $1 \leq i \leq \min\{n_1, n_2\}$, $1 \leq j \leq n_3$ 时, 使 $\tilde{\mathbf{D}}_y = \text{fft}(\mathbf{D}_y, [\], 3)$, $\tilde{\mathbf{D}}_x = \text{fft}(\mathbf{D}_x, [\], 3)$, 则 $\tilde{\mathbf{D}}_x$ 的对角元素可通过以下方程得到:

$$(\tilde{\mathbf{D}}_x(i, i, j))^2 + \left(\frac{1}{\gamma} - \tilde{\mathbf{D}}_y(i, i, j) \right) \tilde{\mathbf{D}}_x(i, i, j) + \frac{1}{\gamma} (\tau - \tilde{\mathbf{D}}_y(i, i, j)) = 0. \quad (21)$$

通过求解式(21)可得其解为

$$\tilde{\mathbf{D}}_x(i, i, j) = \begin{cases} \frac{(\tilde{\mathbf{D}}_y(i, i, j) - 1/\gamma) + \sqrt{\Delta}}{2}, & \Delta \geq 0, \\ 0, & \Delta < 0, \end{cases} \quad (22)$$

其中 $\Delta = \left(\frac{1}{\gamma} - \tilde{\mathbf{D}}_y(i, i, j) \right)^2 - \frac{4}{\gamma} (\tau - \tilde{\mathbf{D}}_y(i, i, j))$.

算法 1 基于定理 1 更新 \mathbf{J} .

输入: $\mathbf{Z}, \mathbf{W}, \rho$ 和 γ ;

步骤 1) $\tau = \frac{1}{\rho}$, $\mathbf{M} = \mathbf{Z} + \frac{1}{\rho} \mathbf{W}$;

步骤 2) $\tilde{\mathbf{M}} = \text{fft}(\mathbf{M}, [\], 3)$;

步骤 3) for $j = 1 : n_3$ do

$$[\tilde{\mathbf{U}}^{(j)}, \tilde{\mathbf{D}}_m^{(j)}, \tilde{\mathbf{V}}^{(j)}] = \text{SVD}(\tilde{\mathbf{M}}^{(j)});$$

由定理 1 得 $\tilde{\mathbf{D}}_g(i, i, j)$;

$$\tilde{\mathbf{J}}^{(j)} = \tilde{\mathbf{U}}^{(j)} \tilde{\mathbf{D}}_m^{(j)} \tilde{\mathbf{V}}^{(j)T};$$

步骤 4) end for

步骤 5) $\mathbf{J} = \text{ifft}(\tilde{\mathbf{J}}, [\], 3)$;

输出: 张量 \mathbf{J} .

2.2.4 更新 Lagrange 乘子 $\mathbf{Y}^{(v)}$ 和 \mathbf{W}

$\mathbf{Y}^{(v)}$ 和 \mathbf{W} 的更新公式分别为

$$\mathbf{Y}^{(v)*} = \mathbf{Y}^{(v)} + \mu (\mathbf{X}^{(v)} - \mathbf{X}^{(v)} \mathbf{Z}^{(v)} - \mathbf{E}^{(v)}), \quad (23)$$

$$\mathbf{W}^* = \mathbf{W} + \rho (\mathbf{Z} - \mathbf{J}). \quad (24)$$

2.2.5 更新 \mathbf{F}

固定变量 $\mathbf{Z}^{(v)}$, $\mathbf{E}^{(v)}$ 和 \mathbf{J} , 可通过求解以下模型更新 \mathbf{F} :

$$\mathbf{F} = \operatorname{argmin}_{\mathbf{F}} \operatorname{tr}(\mathbf{F}^T \mathbf{L}_Z \mathbf{F}), \quad (25)$$

$$\text{s. t. } \mathbf{F}^T \mathbf{F} = \mathbf{I}, \quad \mathbf{F} \in \mathbb{R}^{N \times c}.$$

最优解 F 由 Laplace 矩阵 L_z 的 c 个最小特征值对应的特征向量组成. 下列算法总结了 TMSCDSE 算法的整体优化过程.

算法 2 TMSCDSE 算法.

输入: 给定多视图数据矩阵 $X^{(1)}, X^{(2)}, \dots, X^{(V)}$, λ, β, α 和聚类数量 K ;

初始化 $Z^{(v)}=0, E^{(v)}=0, Y^{(v)}=0, J=W=0, \mu=10^{-5}, \rho=10^{-4}, \eta=2, \mu_{\max}=\rho_{\max}=10^{10}, \epsilon=10^{-7}$;

while 不收敛 do

步骤 1) 利用式(15)更新 $Z^{(v)}$;

步骤 2) 利用式(18)更新 E ;

步骤 3) 利用式(23)更新 $Y^{(v)}$;

步骤 4) 得到 $Z=\Phi(Z^{(1)}, Z^{(2)}, \dots, Z^{(V)})$;

步骤 5) 由算法 1 得到 J ;

步骤 6) 利用式(24)更新 W ;

步骤 7) 利用式(25)更新 F ;

步骤 8) 更新参数 μ 和 ρ : $\mu=\min\{\eta\mu, \mu_{\max}\}, \rho=\min\{\eta\rho, \rho_{\max}\}$;

步骤 9) 得到 $(J^{(1)}, J^{(2)}, \dots, J^{(V)})=\Phi^{-1}(J)$;

步骤 10) 根据收敛条件检查是否收敛: $\|X^{(v)}-X^{(v)}Z^{(v)}-E^{(v)}\|_{\infty}<\epsilon, \|Z^{(v)}-J^{(v)}\|_{\infty}<\epsilon$;

end

步骤 11) 得到聚类指标矩阵 F , 并对其执行 k -means 算法;

输出: 聚类结果 C .

2.3 复杂度分析

算法 2 的计算复杂度主要取决于 4 个变量 $Z^{(v)}, J, E$ 和 F 的更新. 变量 $Z^{(v)}$ 的优化过程涉及 Sylvester 方程的求解, 所以计算复杂度为 $O(VN^3)$. 在更新变量 J 时, 需要提前计算一个维度为 $N \times V \times N$ 的张量的 3D fft 和 ifft 值, 以及在 Fourier 域中 $N \times V (N \geq V)$ 矩阵的 N 个奇异值, 因此计算复杂度为 $O(VN^2 \log(N) + V^2 N^2)$. 对于更新 E 的每次迭代, V 个视图的总复杂度为 $O(VN^2)$. 进一步考虑谱嵌入部分, 变量 F 的计算复杂度为 $O(N^3)$. 其中, V 和 N 分别表示视图数量和样本数量. 最后, 考虑到迭代次数 t 和 $V \leq N$, 所以 TMSCDSE 算法的总计算复杂度为 $O(t(VN^3 + VN^2 \log(N)))$.

3 实验与结果分析

3.1 实验数据

为验证 TMSCDSE 算法的有效性, 在下列 5 个数据集上对其进行评估, 各数据集的信息列于表 1.

表 1 多视图数据集信息

Table 1 Information of multi-view datasets

数据集	样本数	视图数	聚类数	视图维数
MSRCv1	210	5	7	24,576,512,256,254
COIL20	1 440	3	20	1 024,3 304,6 750
BBCsport	544	2	5	3 183,3 203
yaleB	650	3	10	2 500,3 304,6 750
ORL	400	3	40	4 096,3 304,6 750

MSRCv1: 该数据集包含树、汽车、自行车、建筑、飞机、牛和人脸 7 个类别的 210 张图像, 每个类别有 30 张图像. 在进行聚类实验时, 从这些图像中选取 5 个特征组成该数据集的不同视图. COIL20: 该数据集来自哥伦比亚对象图像库, 其包含了 20 个对象的 1 440 张图像, 每个类别由 72 张图片构成; 并且该数据集使用了 3 种类型的特征, 分别是 Intensity, LBP 和 Gabor. BBCsport: 该数据集包含来自 BBC 体育网站的 544 个文档, 其对应 5 个不同的主题, 在该数据集上提取两种类型的特征进行实验. YaleB: 该数据集是由美国耶鲁大学发布的来自 10 个人的 650 张人脸图像, 其每个类包含 65 张不同图像, 实验选取了 Intensity, LBP 和 Gabor 3 种类型的特征, 从而构成 3 种多视图. ORL: 该

数据集是一个人脸图像数据集, 其包含来自 40 人的 400 张人脸图像, 在拍摄时间、光线、面部表情等不同的情况下获取每人的 10 张人脸图像构成样本, 然后再提取 Intensity, LBP 和 Gabor 这 3 种特征构成该多视图数据的不同视图。

3.2 评价指标

实验采用 6 个常用的评估指标验证 TMSCDSE 算法的性能, 包括准确性 (ACC)、归一化互信息 (NMI)、准确率 (Precision)、 F 值 (F -score)、召回率 (Recall) 和调整 Rand 系数 (ARI)。不同指标侧重于聚类的不同属性, 但都满足其值越大表示聚类性能越好的特点。

3.3 对比方法

基于上述 5 个数据集和 6 个评价指标, 本文将提出的 TMSCDSE 算法与现有的多种聚类算法进行比较。对比算法有多样性诱导的多视图子空间聚类 (diversity-induced multi-view subspace clustering, DiMSC)^[25]、低秩张量约束的多视图子空间聚类方法 (low-rank tensor constrained multiview subspace clustering, LT-MSC)^[11]、多视图张量多秩最小化方法 (on unifying multi-view self-representations for clustering by tensor multi-rank minimization, t-SVD-MSC)^[13]、多视图光谱聚类的增强张量学习 (essential tensor learning for multi-view spectral clustering, ETLMSC)^[22]、自适应加权 Procrustes 方法 (multiview clustering via adaptively weighted Procrustes, AWP)^[26]、多视图一致图聚类方法 (multiview consensus graph clustering, MCGC)^[27]、基于图的多视图聚类系统方法 (a study of graph-based system for multi-view clustering, GBS)^[28]、多视图聚类的相似矩阵低秩逼近与不一致分离融合方法 (a similarity matrix low-rank approximation and inconsistency separation fusion approach for multi-view clustering, FAMvC)^[29]、基于低秩矩阵分解的秩一致性诱导多视图子空间聚类方法 (rank consistency induced multiview subspace clustering via low-rank matrix factorization, RC-MSC)^[30] 和统一一步多视图光谱聚类方法 (unified one-step multi-view spectral clustering, UOMvSC)^[31]。

在对所有对比算法进行实验时都采用原文献中所给的最优参数, 以实现最佳结果。此外, 由于所有算法最终都是基于 k -means 实现的, 所以为减少随机初始化带来的误差, 对所有对比算法以及 TMSCDSE 算法都重复运行 20 次, 取其平均值作为最终聚类结果进行记录。

3.4 实验结果

表 2~表 6 分别列出了各算法在上述 5 个数据集上的聚类结果。由表 2~表 6 可见, 大多数情况下 TMSCDSE 算法都优于其他聚类算法。这可能是因为该算法能根据奇异值的不同大小, 自适应地选择合适的逼近函数。此外, TMSCDSE 算法将低秩张量学习、多样性学习和谱嵌入学习在一个统一的框架中进行, 使学习到的系数矩阵能更好地表征聚类结构, 并有效结合了多视图数据的一致性信息和多样性信息。例如, 在数据集 BBCsport 上与次优方法相比, TMSCDSE 算法在 ACC, NMI, F -score, Precision, Recall 和 ARI 6 个指标上分别实现了 1.1, 3.5, 1.4, 1.9, 2.5, 2.6 个百分点的改进。在数据集 yaleB 上这 6 个指标分别提高了 33.4, 27.7, 41.4, 38.8, 36.0, 43.4 个百分点, 提升效果更明显。

表 2 不同算法在数据集 MSRCv1 上的聚类性能对比

Table 2 Comparison of clustering performance of different algorithms on MSRCv1 dataset

算法	ACC	NMI	Precision	F -score	Recall	ARI
DiMSC	0.722	0.627	0.606	0.613	0.621	0.550
LT-MSC	0.820	0.750	0.694	0.717	0.741	0.670
t-SVD-MSC	0.991	0.982	0.980	0.981	0.982	0.978
ETLMSC	0.849	0.801	0.753	0.770	0.789	0.732
AWP	0.748	0.716	0.605	0.682	0.783	0.623
MCGC	0.762	0.712	0.647	0.691	0.741	0.637
GBS	0.833	0.790	0.703	0.749	0.803	0.706
FAMvC	0.767	0.724	0.658	0.697	0.741	0.645
RC-MSC	0.859	0.758	0.759	0.754	0.749	0.714
UOMvSC	0.905	0.819	0.814	0.819	0.826	0.790
TMSCDSE	0.995	0.989	0.990	0.990	0.990	0.989

表 3 不同算法在数据集 COIL20 上的聚类性能对比

Table 3 Comparison of clustering performance of different algorithms on COIL20 dataset

算法	ACC	NMI	Precision	F -score	Recall	ARI
DiMSC	0.764	0.834	0.714	0.724	0.734	0.709
LT-MSC	0.703	0.784	0.613	0.637	0.663	0.617
t-SVD-MSC	0.809	0.884	0.753	0.772	0.793	0.760
ETLMSC	0.882	0.955	0.849	0.884	0.923	0.878
AWP	0.679	0.886	0.525	0.663	0.898	0.641
MCGC	0.748	0.859	0.547	0.647	0.792	0.625
GBS	0.791	0.941	0.694	0.794	0.929	0.782
FAMvC	0.843	0.949	0.743	0.828	0.934	0.818
RC-MSC	0.759	0.824	0.724	0.715	0.706	0.699
UOMvSC	0.824	0.941	0.795	0.819	0.926	0.809
TMSCDSE	0.910	0.950	0.890	0.899	0.908	0.893

表 4 不同算法在数据集 BBCsport 上的聚类性能对比

Table 4 Comparison of clustering performance of different algorithms on BBCsport dataset

算法	ACC	NMI	Precision	F -score	Recall	ARI
DiMSC	0.829	0.654	0.759	0.750	0.741	0.673
LT-MSC	0.926	0.803	0.845	0.858	0.871	0.813
t-SVD-MSC	0.987	0.958	0.984	0.978	0.972	0.971
ETLMSC	0.928	0.925	0.931	0.925	0.919	0.902
AWP	0.634	0.501	0.454	0.564	0.744	0.380
MCGC	0.570	0.330	0.324	0.456	0.770	0.181
GBS	0.857	0.865	0.802	0.877	0.968	0.834
FAMvC	0.559	0.304	0.392	0.489	0.649	0.272
RC-MSC	0.840	0.725	0.796	0.793	0.791	0.728
UOMvSC	0.526	0.314	0.379	0.437	0.935	0.209
TMSCDSE	0.998	0.993	0.998	0.997	0.997	0.997

表 5 不同算法在数据集 yaleB 上的聚类性能对比

Table 5 Comparison of clustering performance of different algorithms on yaleB dataset

算法	ACC	NMI	Precision	F -score	Recall	ARI
DiMSC	0.620	0.644	0.497	0.523	0.552	0.468
LT-MSC	0.468	0.473	0.323	0.346	0.373	0.269
t-SVD-MSC	0.554	0.583	0.423	0.448	0.476	0.384
ETLMSC	0.244	0.173	0.164	0.167	0.171	0.074
AWP	0.585	0.579	0.478	0.505	0.529	0.444
MCGC	0.414	0.411	0.282	0.282	0.346	0.186
GBS	0.434	0.449	0.204	0.265	0.378	0.157
FAMvC	0.605	0.574	0.318	0.386	0.492	0.303
RC-MSC	0.511	0.501	0.410	0.341	0.292	0.255
UOMvSC	0.400	0.409	0.234	0.273	0.328	0.179
TMSCDSE	0.954	0.921	0.911	0.911	0.912	0.902

在这些对比算法中,基于张量的方法包括 LT-MSC,t-SVD-MSC 和 ETLMSC,从聚类结果可观察到 TMSCDSE 算法的性能基本都优于这些算法,原因可能是这些算法将低秩张量的学习过程与谱聚类过程分两步进行,忽略了其中的依赖关系,使学习聚类指标矩阵的准确性降低,从而进一步影响了最终的聚类结果.针对该问题,TMSCDSE 算法将低秩张量学习和谱嵌入学习集成在一个统一的模型中同时进行优化,经过迭代后可直接获得聚类指标矩阵.

基于图的方法 MCGC,GBS 和 FAMvC 在 5 个数据集上的结果都不如张量低秩方法,这是因为数

据集中通常包含一些噪声和冗余信息, 对构造出的图质量会产生一定影响. DiMSC 算法与 TMSCDSE 算法都采用了 HSIC 作为多样性的衡量, 表明充分利用视图间的多样性信息对提升聚类性能有较大作用. 此外, TMSCDSE 算法能更全面地探索嵌入在所有多视图中的高阶相关信息和互补信息.

表 6 不同算法在数据集 ORL 上的聚类性能对比

Table 6 Comparison of clustering performance of different algorithms on ORL dataset

算法	ACC	NMI	Precision	F-score	Recall	ARI
DiMSC	0.798	0.909	0.694	0.738	0.789	0.732
LT-MSC	0.806	0.914	0.708	0.752	0.802	0.746
t-SVD-MSC	0.963	0.992	0.945	0.964	0.985	0.964
ETLMSC	0.880	0.969	0.826	0.879	0.939	0.876
AWP	0.683	0.873	0.567	0.652	0.769	0.643
MCGC	0.712	0.873	0.462	0.567	0.734	0.555
GBS	0.633	0.857	0.232	0.360	0.801	0.337
FAMvC	0.780	0.898	0.643	0.706	0.783	0.699
RC-MSC	0.777	0.865	0.704	0.669	0.637	0.652
UOMvSC	0.858	0.926	0.768	0.798	0.832	0.794
TMSCDSE	0.973	0.995	0.960	0.974	0.989	0.974

3.5 参数分析

TMSCDSE 算法中含有 3 个重要参数, 分别为 α, β 和 λ . 其中, 参数 α 表示谱嵌入项的重要性, 参数 β 表示视图之间的差异, 参数 λ 表示与多视图数据的损坏程度密切相关. 即在 $[1 \times 10^{-9}, 1 \times 10^{-8}, 1 \times 10^{-7}, 1 \times 10^{-6}, 1 \times 10^{-5}]$ 内调整 α , 在 $[1 \times 10^{-4}, 5 \times 10^{-4}, 1 \times 10^{-3}, 5 \times 10^{-3}, 0.01, 0.05, 0.1, 0.5]$ 内调整 λ , 在 $[1 \times 10^{-5}, 5 \times 10^{-5}, 1 \times 10^{-4}, 5 \times 10^{-4}, 1 \times 10^{-3}, 5 \times 10^{-3}, 0.01, 0.05]$ 内调整 β . 通过在上述范围内进行参数调优, 找到每个数据集对应的最优参数, 然后获取最优的聚类结果.

以两个数据集 MSRCv1 和 COIL20 为例, 图 2 为固定参数 α 时, 不同参数 β, λ 对评价指标 ACC 的影响. 由图 2(A) 可见, 数据集 MSRCv1 对 β 和 λ 有一定的敏感性, 不同参数取值使算法性能出现波动. 由图 2(B) 可见, 当参数 β 在 $[5 \times 10^{-5}, 1 \times 10^{-4}, 5 \times 10^{-4}, 1 \times 10^{-3}, 5 \times 10^{-3}, 0.01]$ 内调整和参数 λ 在 $[1 \times 10^{-4}, 5 \times 10^{-4}, 1 \times 10^{-3}, 5 \times 10^{-3}, 0.01, 0.05, 0.1]$ 内调整时, 算法性能的波动较小, 有一定的鲁棒性. 说明低秩张量学习和多样性学习之间相互联系, 只有同时考虑这两项的不同重要性才能获得最佳的聚类结果.

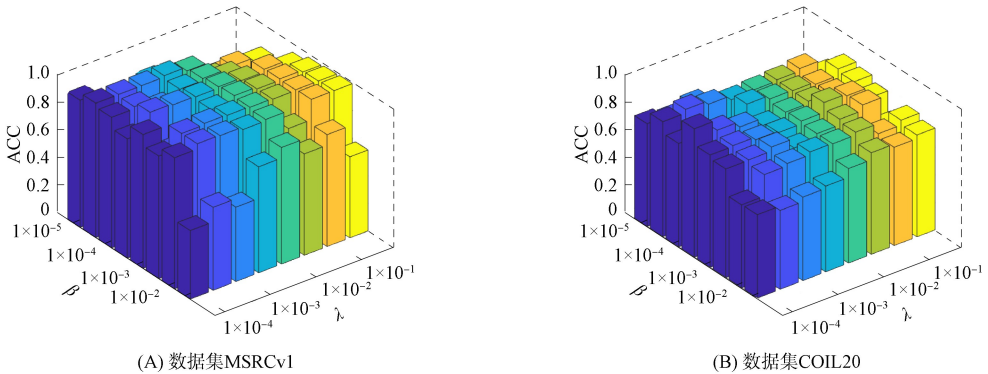


图 2 在数据集 MSRCv1 和 COIL20 上对参数 β 和 λ 的敏感性分析

Fig. 2 Sensitivity analysis of parameters β and λ on MSRCv1 and COIL20 datasets

图 3 为固定参数 β 和 λ 时, 参数 α 在数据集 MSRCv1 和 COIL20 上的取值变化对算法聚类性能的影响. 由图 3 可见, 随着 α 的变化, ACC 和 NMI 对应的值发生明显波动. 当 $\alpha = 1 \times 10^{-8}$ 时, 两个数据集均可获得最佳性能. 而当 $\alpha > 1 \times 10^{-8}$ 时, TMSCDSE 算法的聚类性能显著下降. 这可能是因为低秩张量学习捕获了嵌入在不同视图之间的高阶相关信息和互补信息, 且多样性项充分考虑到多视图的多样性信息, 使学习到的系数矩阵已经很好地表征了数据之间的关系. 谱嵌入项是为了使学习到的系数

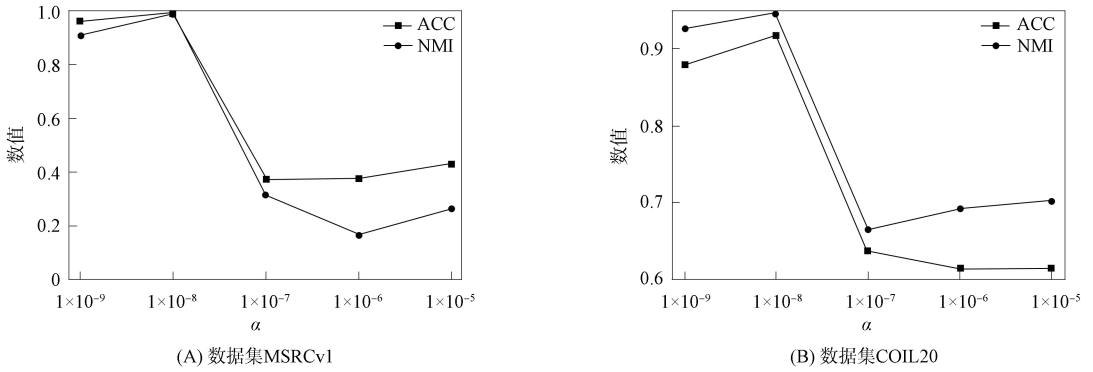


图 3 在数据集 MSRCv1 和 COIL20 上对参数 α 的敏感性分析

Fig. 3 Sensitivity analysis of parameters α on MSRCv1 and COIL20 datasets

矩阵进一步利用聚类结构,为减少其对前两部分已经取得效果的影响,所以应给参数 α 赋予一个较小值.

3.6 收敛性分析

当存在 3 个或更多变量时,非精确的 ALM 的收敛性无法确定^[5].由算法 2 可见, TMSCDSE 算法存在 4 个块变量 $\mathbf{Z}, \mathbf{E}, \mathbf{J}$ 和 \mathbf{F} , 所以很难证明该算法的收敛性. 因此,为验证优化算法的收敛性,本文采用重构误差(RE)和匹配误差(ME)绘制收敛曲线,并在图 4 中给出 TMSCDSE 算法在数据集 BBCsport 和 yaleB 上的收敛曲线. 其中,重构误差和匹配误差分别定义为

$$RE = \frac{1}{V} \sum_{v=1}^V \|\mathbf{X}^{(v)} - \mathbf{X}^{(v)} \mathbf{Z}^{(v)} - \mathbf{E}^{(v)}\|_{\infty}, \quad (26)$$

$$ME = \frac{1}{V} \sum_{v=1}^V \|\mathbf{Z}^{(v)} - \mathbf{J}^{(v)}\|_{\infty}. \quad (27)$$

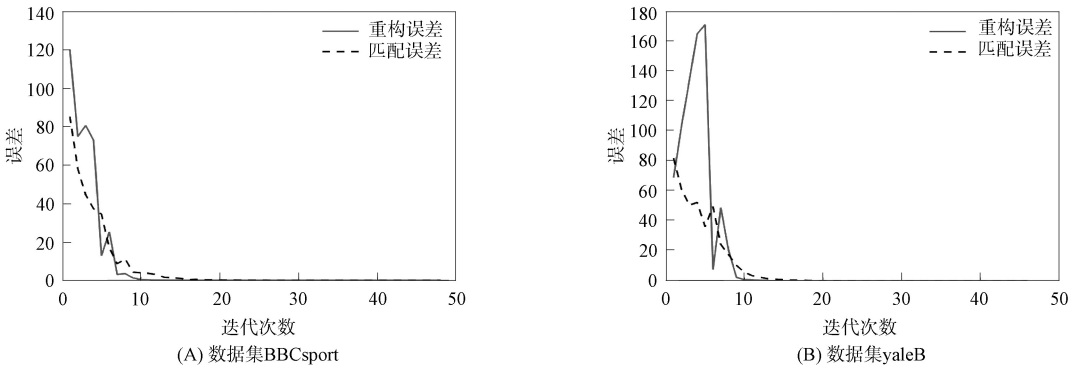


图 4 在数据集 BBCsport 和 yaleB 上的收敛曲线

Fig. 4 Convergence curves on BBCsport and yaleB datasets

由图 4 可见,随着迭代次数的增加,停止准则的值趋于 0,而且在迭代约 45 次时已经达到稳定水平,从而证明了 TMSCDSE 算法的收敛性.

3.7 消融实验

为进一步说明 TMSCDSE 算法的有效性,对其进行消融实验分别考察谱嵌入部分和多样性部分的作用. 第一个实验:去掉其中的多样性项,仅考虑谱嵌入项的作用,将其表示为 TMSCDSE-S. 第二个实验:仅考虑加入多样性项,去除最后的谱嵌入项,将其表示为 TMSCDSE-D. 这两个实验以及 TMSCDSE 算法在 ACC 和 NMI 指标上的聚类结果如图 5 所示. 由图 5 可见, TMSCDSE-D 和 TMSCDSE-S 算法的聚类效果在 5 个数据集上都不如 TMSCDSE 算法. 因此,只有同时考虑到多视图数据的多样性,以及将系数矩阵和谱聚类的学习过程在一个统一框架中同时进行,才能实现最佳的聚类效果.

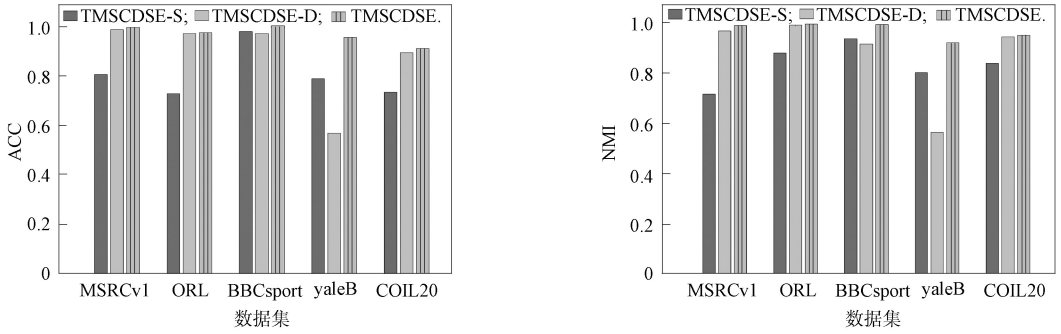


图 5 TMSCDSE 算法性能的消融实验结果

Fig. 5 Results of ablation experiment of TMSCDSE algorithm performance

综上所述, 针对如何有效利用多视图的多样性信息和高阶信息, 并建立系数矩阵的学习过程与谱聚类之间联系的问题, 本文提出了一种融合低秩张量学习、多样性学习和谱嵌入学习的多视图子空间聚类统一框架, 称为基于多样性和谱嵌入的张量多视图子空间聚类算法(TMSCDSE). 首先, 利用所有视图的系数矩阵构造一个三阶张量, 以充分挖掘多视图数据中的高阶相关性, 并采用对数行列式正则化项恢复低秩张量. 其次, 考虑到不同视图的多样性, 采用 Hilbert-Schmidt 独立准则对视图的系数矩阵进行约束. 最后, 通过联合谱嵌入构建统一的模型框架, 学习更有效的聚类指标矩阵, 再利用 k -means 算法得到最终的聚类结果. 在 5 个真实数据集上进行大量对比实验的结果表明, TMSCDSE 算法优越性显著.

参 考 文 献

- [1] TAN Y, OU W H, LONG F, et al. Multi-view Clustering via Co-regularized Nonnegative Matrix Factorization with Correlation Constraint [C]//2016 7th International Conference on Cloud Computing and Big Data (CCBD). Piscataway, NJ: IEEE, 2016: 1-6.
- [2] WANG C D, CHEN M S, HUANG L, et al. Smoothness Regularized Multiview Subspace Clustering with Kernel Learning [J]. IEEE Transactions on Neural Networks and Learning Systems, 2021, 32(11): 5047-5060.
- [3] 赵伟豪, 林浩申, 曹传杰, 等. 基于相似图投影学习的多视图聚类 [J]. 计算机应用研究, 2024, 41(1): 102-107. (ZHAO W H, LIN H S, CAO C J, et al. Multi-view Clustering Based on Similarity Graph Projection Learning [J]. Application Research of Computers, 2024, 41(1): 102-107.)
- [4] 王丽娟, 丁世飞, 夏菁. 基于多样性的多视图低秩稀疏子空间聚类算法 [J]. 智能系统学报, 2023, 18(2): 399-407. (WANG L J, DING S F, XIA J. Multiview Low-Rank Sparse Subspace Clustering Based on Diversity [J]. CAAI Transactions on Intelligent Systems, 2023, 18(2): 399-407.)
- [5] LUI G C, LIN Z C, YAN S H, et al. Robust Recovery of Subspace Structures by Low-Rank Representation [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2013, 35(1): 171-184.
- [6] EIHAMIFAR E, VIDAL R. Sparse Subspace Clustering: Algorithm, Theory, and Applications [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2013, 35(11): 2765-2781.
- [7] LUO S R, ZHANG C Q, ZHANG W, et al. Consistent and Specific Multi-view Subspace Clustering [C]//Proceedings of the AAAI Conference on Artificial Intelligence. Palo Alto: AAAI Press, 2018: 3730-3736.
- [8] ZHANG C Q, FU H Z, HU Q H, et al. Generalized Latent Multi-view Subspace Clustering [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2020, 42(1): 86-99.
- [9] CHEN J, YANG S X, MAO H, et al. Multiview Subspace Clustering Using Low-Rank Representation [J]. IEEE Transactions on Cybernetics, 2022, 52(11): 12364-12378.
- [10] 张绎凡, 李婷, 葛洪伟. 多样性诱导的潜在嵌入多视图聚类 [J]. 南京大学学报(自然科学), 2023, 59(3): 388-397. (ZHANG Y F, LI T, GE H W. Diversity-Induced Multi-view Clustering in Latent Embedded Space [J]. Journal of Nanjing University (Natural Science), 2023, 59(3): 388-397.)
- [11] ZHANG C Q, FU H Z, LIN S, et al. Low-Rank Tensor Constrained Multiview Subspace Clustering [C]//2015 IEEE International Conference on Computer Vision (ICCV). Piscataway, NJ: IEEE, 2015: 1582-1590.

- [12] LU C Y, FENG J S, CHEN Y D, et al. Tensor Robust Principal Component Analysis: Exact Recovery of Corrupted Low-Rank Tensors via Convex Optimization [C]//2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway, NJ: IEEE, 2016: 5249-5257.
- [13] XIE Y, TAO D C, ZHANG W S, et al. On Unifying Multi-view Self-representations for Clustering by Tensor Multi-rank Minimization [J]. International Journal of Computer Vision, 2016, 126(4): 1157-1179.
- [14] GAO Q X, XIA W, WAN Z Z, et al. Tensor-SVD Based Graph Learning for Multi-view Subspace Clustering [C]//Proceedings of the AAAI Conference on Artificial Intelligence. Palo Alto: AAAI Press, 2020: 3930-3937.
- [15] SUN X L, HAI Y, ZHANG X J, et al. Adaptive Tensor Rank Approximation for Multi-view Subspace Clustering [J]. Chinese Journal of Electronics, 2023, 32(4): 840-853.
- [16] KILMER M E, MARTIN C D. Factorization Strategies for Third-Order Tensors [J]. Linear Algebra and Its Applications, 2011, 435(3): 641-658.
- [17] LU C Y, FENG J S, CHEN Y D, et al. Tensor Robust Principal Component Analysis with a New Tensor Nuclear Norm [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2020, 42(4): 925-938.
- [18] TANG C, ZHU X Z, LIU X W, et al. Learning a Joint Affinity Graph for Multiview Subspace Clustering [J]. IEEE Transactions on Multimedia, 2019, 21(7): 1724-1736.
- [19] ZHANG X Q, CHEN B J, SUN H J, et al. Robust Low-Rank Kernel Subspace Clustering Based on the Schatten p -Norm and Correntropy [J]. IEEE Transactions on Knowledge and Data Engineering, 2020, 32(12): 2426-2437.
- [20] CHEN Y Y, WANG S Q, PENG C, et al. Generalized Nonconvex Low-Rank Tensor Approximation for Multi-view Subspace Clustering [J]. IEEE Transactions on Image Processing, 2021, 30(1): 4022-4035.
- [21] 赵晓佳, 徐婷婷, 陈勇勇, 等. 基于一步张量学习的多视图子空间聚类 [J]. 自动化学报, 2023, 49(1): 40-53. (ZHAO X J, XU T T, CHEN Y Y, et al. One-Step Tensor Learning for Multi-view Subspace Clustering [J]. Acta Automatica Sinica, 2023, 49(1): 40-53.)
- [22] WU J L, LIN Z C, ZHA H B. Essential Tensor Learning for Multi-view Spectral Clustering [J]. IEEE Transactions on Image Processing, 2019, 28(12): 5910-5922.
- [23] XIE Y, ZHANG W S, QU Y Y, et al. Hyper-Laplacian Regularized Multilinear Multiview Self-representations for Clustering and Semisupervised Learning [J]. IEEE Transactions on Cybernetics, 2020, 50(2): 572-586.
- [24] BARTELS R H, STEWART G W. Algorithm 432 [C2]: Solution of the Matrix Equation $AX, XB = C$ [J]. Communications of the ACM, 1972, 15(9): 820-826.
- [25] CAO X C, ZHANG C Q, FU H Z, et al. Diversity-Induced Multi-view Subspace Clustering [C]//2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway, NJ: IEEE, 2015: 586-594.
- [26] NIE F P, TIAN L, LI X L. Multiview Clustering via Adaptively Weighted Procrustes [C]//Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining. [S. l.]: Association for Computing Machinery, 2018: 2022-2030.
- [27] ZHAN K, NIE F P, WANG J, et al. Multiview Consensus Graph Clustering [J]. IEEE Transactions on Image Processing, 2019, 28(3): 1261-1270.
- [28] WANG H, YANG Y, LIU B, et al. A Study of Graph-Based System for Multi-view Clustering [J]. Knowledge-Based Systems, 2019, 163(1): 1009-1019.
- [29] HE Z Q, WAN S H, ZAPPATORE M, et al. A Similarity Matrix Low-Rank Approximation and Inconsistency Separation Fusion Approach for Multi-view Clustering [J]. IEEE Transactions on Artificial Intelligence, 2024, 5(2): 868-881.
- [30] GUO J P, SUN Y F, GAO J B, et al. Rank Consistency Induced Multiview Subspace Clustering via Low-Rank Matrix Factorization [J]. IEEE Transactions on Neural Networks and Learning Systems, 2022, 33(7): 3157-3170.
- [31] TANG C, LI Z L, WANG J, et al. Unified One-Step Multi-view Spectral Clustering [J]. IEEE Transactions on Knowledge and Data Engineering, 2023, 35(6): 6449-6460.