



基于本质自表示的多视角子空间聚类

于晓^{1,2}, 刘慧^{1,2,3*}, 吴彦³, 张彩明^{2,4}

1. 山东财经大学计算机科学与技术学院, 济南 250014, 中国

2. 山东省数字媒体技术重点实验室, 济南 250014, 中国

3. Medical Center, Stanford University, Palo Alto 94305, USA

4. 山东大学软件学院, 济南 250010, 中国

* 通信作者. E-mail: liuh_lh@sdufe.edu.cn

收稿日期: 2020-05-08; 修回日期: 2020-07-22; 接受日期: 2020-10-09; 网络出版日期: 2021-10-12

国家自然科学基金(批准号: 62072274)和山东省重点研发计划(批准号: 2019LYXZ030)资助项目

摘要 大数据时代, 针对高维数据的聚类分析已成为数据挖掘的热点和难点问题, 与此同时, 近年来数据获取方式的重大变革推动了多视角研究快速兴起. 多视角子空间聚类能够针对高维数据进行聚类, 已成为目前聚类研究的一个重要分支. 然而, 现有部分方法直接从原始数据学习相似度矩阵, 极易受到数据中噪声的影响. 因此, 本文提出了一种基于本质自表示的多视角子空间聚类方法 (intrinsic self-representation for multi-view subspace clustering, ISMSC), 从数据的潜在表示中学习相似度矩阵, 并利用数据的自表达特性将所有视角的特征融合成本质表示. 然后, 本文使用交替方向乘子法 (alternating direction method of multipliers, ADMM) 设计了有效的求解策略对模型进行优化. 最后, 将 ISMSC 在人造数据集和 8 个真实基准数据集上的实验结果与 9 种基准方法进行了对比, 结果表明了模型的鲁棒性和有效性.

关键词 多视, 聚类, 子空间, 谱聚类, 稀疏

1 引言

聚类分析一直是数据挖掘中的热点和难点问题^[1,2], 在大数据时代亦不例外. 如今大量的数据以高维的形式存在, 而处理高维数据时, 计算量会随数据维度的增加呈现指数级增长, 出现维度灾难问题. 因此, 在大数据时代背景下研究高维数据的聚类问题具有重要的理论意义和现实意义. 目前, 子空间聚类能够针对高维数据, 基于相同子空间中的数据相似、不同子空间中的数据存在差异性的原理, 将数据映射到不同的子空间中, 已广泛应用于图像分割^[3]、人脸聚类^[4]等众多领域. 与此同时, 大数据、云计算、物联网等技术的发展带来了数据获取方式的变革, 人们可以通过不同的数据源、模态或

引用格式: 于晓, 刘慧, 吴彦, 等. 基于本质自表示的多视角子空间聚类. 中国科学: 信息科学, 2021, 51: 1625–1639, doi: 10.1360/SSI-2020-0274

Yu X, Liu H, Wu Y, et al. Intrinsic self-representation for multi-view subspace clustering (in Chinese). Sci Sin Inform, 2021, 51: 1625–1639, doi: 10.1360/SSI-2020-0274

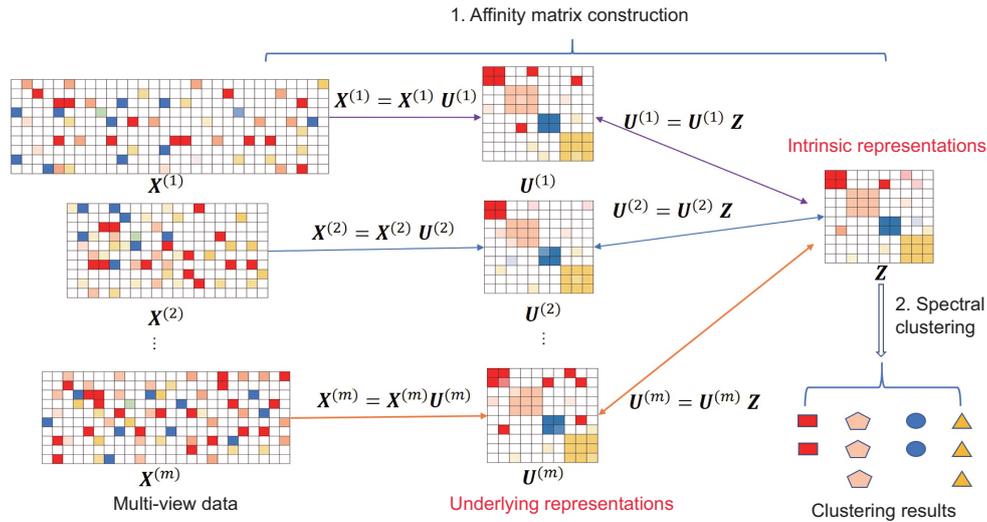


图 1 (网络版彩图) ISMSC 的整体框架
 Figure 1 (Color online) The framework of ISMSC

数据特征获得同一数据的相关信息^[5,6]. 相比单一的数据表示形式, 多视角数据包含的信息更加全面, 故而能够提升数据分析的效果, 因此多视角学习具有重要的研究价值.

多视角子空间聚类在子空间聚类的基础上, 利用各个视角数据的信息, 将高维数据映射到不同子空间中, 可以达到更好的聚类效果, 具有更广阔的应用前景. 目前, 研究者已提出大量多视角子空间聚类算法, 但仍然存在不足. 许多方法直接从原始数据学习建立邻接矩阵 (相似度矩阵)^[7,8]. 但是, 原始数据通常包含噪声和冗余信息, 数据质量不高, 进而影响了其使用效果. 为此, 很多方法^[9,10]使用高斯 (Gauss) 核函数生成相似度矩阵, 但高斯标准差参数的设置会影响聚类的效果. 另一类常见做法是将所有的特征进行连接, 学习数据的潜在表示^[11,12], 但此类方式忽视了各视角数据间的差异, 无法充分利用不同视角数据间的互补信息.

为了解决上述问题, 本文提出了一种基于本质自表示的多视角子空间聚类方法 (intrinsic self-representation for multi-view subspace clustering, ISMSC). 从图 1 中可以看出, ISMSC 首先使用自表示方法学习单视角数据的潜在表示, 然后将多个视角数据的潜在表示融合为本质表示. 最终, 所有视角内包含的数据点之间的相似度信息都在本质表示中得以体现.

本文贡献主要体现在以下 3 个方面.

- (1) 为了减弱原始数据中噪声的影响, ISMSC 从潜在表示中间接学习相似度矩阵.
- (2) 在多视角数据具有本质相似度这一假设下, ISMSC 将多个视角数据的潜在表示融合为本质表示, 充分利用了各个视角数据之间的一致性信息和互补信息.
- (3) 基于交替方向乘法为 ISMSC 设计优化算法. 本文在人造数据集上的实验结果证明了算法对噪声的鲁棒性, 并在多类基准数据集上验证了算法的有效性.

本文在第 2 节对相关工作进行了说明. 第 3 节介绍了基于本质自表示的多视角子空间聚类方法和相应的优化算法. 第 4 节在人造数据集的实验上证明了算法对噪声具有鲁棒性, 并在多种类型基准数据集上与 9 种方法进行比较, 验证了算法的有效性. 最后, 第 5 节对全文进行了总结.

2 相关工作

多视角子空间聚类是多视角聚类的一个分支,由单视角子空间聚类发展而来.本节首先简要介绍最具影响力的单视角子空间聚类方法,然后对多视角子空间聚类方法进行重点说明.

2.1 单视角子空间聚类

目前,子空间聚类^[13]主要有迭代法、代数法、统计法,以及基于谱聚类的方法.近年来,出现了多种基于谱聚类的子空间聚类方法,最具代表性的方法是 Elhamifar 等^[4]提出的稀疏子空间聚类法 (sparse subspace clustering, SSC).它提出了数据的自表达特性,即每个数据点都可以用其他点的线性或仿射组合形式来表示,并根据这一特性对子空间进行聚类.此后,在 SSC 的影响下,出现了多种基于数据自表达特性的算法. Liu 等^[14]专注于数据的低秩表示,设计了低秩表示聚类.结合数据的稀疏性和低秩表示, Luo 等^[15]提出了多空间表示 (multi-space representation, MSR) 来解决子空间聚类的问题. Hu 等^[16]提出了平滑表示模型 (smooth representation, SMR) 中的强制分组效果条件,来判断表示模型中是否存在分组效果.

2.2 多视角子空间聚类

多视角子空间聚类大致可以分为 3 类:基于谱聚类的方法、基于非负矩阵分解的方法^[17]和基于深度学习的方法^[18].其中,基于谱聚类的方法由于其易实现性和有效性而最为流行,通常包含两个阶段.第 1 阶段构造邻接矩阵,第 2 阶段执行谱聚类算法获得聚类结果.具体而言,在获得相似度矩阵之后,使用谱(特征值)进行降维,再调用 k -均值算法将数据划分为维数较低的簇.构造相似度矩阵主要有 4 种方法: ϵ -邻域图、 k -最近邻图、完全连通图^[19]和自表示方法.自表示方法是其中应用最广的,利用了 SSC^[20]中提出的数据自表达特性构造相似度矩阵. Xia 等^[21]基于马尔科夫 (Markov) 链重建了共享转移概率矩阵.为了充分利用不同视角数据之间的相关性, Wang 等^[22]在每次迭代中使用一个视角的低秩表示来对另一个视角的表示进行约束. Zhang 等^[11]利用不同视角之间的潜在互补信息对数据点进行聚类.在低秩表示的基础上, Tang 等^[23]设计了多样性正则项,为各个视角学习权重,得到了统一的邻接图.通过使用张量空间中的 t -积, Yin 等^[24]提出了 t 线性组合的低秩多视角聚类方法. Xie 等^[25]将多个视角的子空间表示矩阵堆叠起来构造张量,探索了多视角数据之间的高阶相关性.通过使用完整空间学习技术, Wang 等^[26]假设相似度矩阵最大程度地依赖于其潜在完整空间,从而构造相似图矩阵. Zhang 等^[27]捕捉了数据间的非线性结构,在单步框架里学习共同的相似度矩阵. Li 等^[28]使用神经网络构建多层的子空间表示,分层学习聚类结构. Kang 等^[29]结合了单视角的图学习、单视角的划分以及多视角的融合,提出了一种多视角子空间聚类方法. Weng 等^[30]提取了多个视角的全局一致性信息,为所有视角学习共同的子空间.

尽管多视角子空间聚类研究已取得较理想的进展,但仍然存在很多问题.首先,许多方法直接从原始数据中学习相似度矩阵,易受到噪声和冗余信息的影响,无法获得理想的邻接矩阵.其次,很多方法将各个视角的所有特征连接起来,忽视了不同视角数据之间的差异,无法探索不同视角之间的相关关系,难以有效利用各视角数据之间的互补信息.相比之下,本文提出的 ISMSC 从潜在表示中学习相似度矩阵,削弱了原始数据中噪声对其的影响.此外,通过将潜在表示融合成本质表示, ISMSC 能够充分利用不同视角数据之间的互补信息.

表 1 符号及含义

Table 1 The meaning of notations in the paper

Notation	Meaning
a	A scalar
\mathbf{a}	A vector
\mathbf{A}	A matrix
$\ \mathbf{A}\ _1$	The l_1 norm of matrix \mathbf{A} , $\ \mathbf{A}\ _1 = \sum_{i,j} a_{i,j}$
$\ \mathbf{A}\ _F$	The Frobenius norm of matrix \mathbf{A} , $\ \mathbf{A}\ _F = \sqrt{\sum_{i,j} a_{i,j} ^2}$
$\text{diag}(\mathbf{A})$	The vector of the diagonal elements of \mathbf{A}

3 基于本质自表示的多视角子空间聚类

本节首先介绍相关的概念, 然后详细说明 ISMSC 的过程, 并提出了解决 ISMSC 的优化方法. 最后, 对 ISMSC 的计算复杂度和收敛性进行了分析.

3.1 相关概念

为了方便描述, 表 1 对本文中所使用的符号进行了说明.

本小节主要介绍数据的自表达特性、潜在表示、本质表示和本质自表示的定义.

定义1 (数据的自表达特性) 子空间并集中的每个数据点都可以通过数据集中其他点的组合进行有效重建. 给定 n 个数据点的集合 \mathbf{X} , $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_i, \dots, \mathbf{x}_n] \in \mathbb{R}^{d \times n}$, 每一列代表一个数据点, 利用数据的自表达特性, \mathbf{X} 可以表示成 $\mathbf{X}\mathbf{Z}$, 其中 $\mathbf{Z} = [z_1, z_2, \dots, z_i, \dots, z_n] \in \mathbb{R}^{n \times n}$, \mathbf{Z} 称为系数矩阵, 其中的每一项 z_i 是原始数据点 \mathbf{x}_i 的表示. 为避免 \mathbf{Z} 是单位矩阵, 将 $\text{diag}(\mathbf{Z})$ 设为 0.

定义2 (自表示) 自表示是利用数据的自表达特性, 将数据点的集合 \mathbf{X} 用 $\mathbf{X}\mathbf{Z}$ 来表示. \mathbf{Z} 也称为自表示矩阵.

定义3 (潜在表示) 对于多视角数据集 \mathbf{X} 而言, 第 v 个视角的数据 $\mathbf{X}^{(v)}$ 的潜在表示捕捉了 $\mathbf{X}^{(v)}$ 的特征, 可以用来表示 $\mathbf{X}^{(v)}$. 本文使用数据的自表达特性来获得 \mathbf{X} 的潜在表示.

定义4 (本质表示) 多视角数据集 \mathbf{X} 中的每个视角都可以看成是对 \mathbf{X} 的一种描述, 存在一种表示, 能够代表 \mathbf{X} 中各数据点之间的相似度. 本质表示就是用来表示多视角数据中实例之间的本质相似度.

定义5 (本质自表示) 本质自表示是使用数据的自表达特性来获得多视角数据的本质相似度表示.

3.2 基于本质自表示的多视角子空间聚类

如图 1 所示, 基于本质自表示的多视角子空间聚类主要由两阶段构成: 构建相似度矩阵和执行谱聚类算法. 本小节主要介绍第 1 阶段. 给定多视角数据 \mathbf{X} , $\mathbf{X}^{(v)} \in \mathbb{R}^{d_v \times n}$ 表示 \mathbf{X} 的第 v 个视角 ($v = 1, 2, \dots, m$), d_v 代表 $\mathbf{X}^{(v)}$ 的数据维度, n 表示样本的数量. 由于原始数据中通常包含噪声, 为了减轻噪声对其的影响, ISMSC 间接学习相似度矩阵. 常见的做法是使用高斯核函数计算相似度矩阵^[9, 10], 但该方式生成的相似度矩阵的质量易受到高斯标准方差参数的影响. 鉴于单视角稀疏子空间聚类法 SSC 的自表示特性能很好地对数据进行重建, 故使用自表示来获得数据的潜在表示:

$$\mathbf{X}^{(v)} = \mathbf{X}^{(v)}\mathbf{U}^{(v)}, \quad (1)$$

其中 $\mathbf{X}^{(v)} \in \mathbb{R}^{d_v \times n}$ 代表着 \mathbf{X} 的第 v 个视角中的数据, $\mathbf{U}^{(v)}$ 是 $\mathbf{X}^{(v)}$ 的潜在表示. 将重建错误作为损

失函数, 得到

$$\min_{\mathbf{U}^{(v)}} \sum_{v=1}^m \left\| \mathbf{X}^{(v)} - \mathbf{X}^{(v)} \mathbf{U}^{(v)} \right\|_F^2. \quad (2)$$

由于 \mathbf{X} 中的每个视角都是对同一数据集的描述, 那么应存在一种表示, 可以代表所有视角中数据点的相似度关系. 在这种假设下, 把 \mathbf{Z} 定义为 \mathbf{X} 的本质表示. 在文献 [31] 的启发下, 将潜在表示融合为本质表示. 基于数据的自表示特性, $\mathbf{U}^{(v)}$ 可以表示为

$$\mathbf{U}^{(v)} = \mathbf{U}^{(v)} \mathbf{Z}. \quad (3)$$

将潜在表示和本质表示的过程结合起来, 得到损失函数:

$$\min_{\mathbf{U}^{(v)}, \mathbf{Z}} \sum_{v=1}^m \left\| \mathbf{X}^{(v)} - \mathbf{X}^{(v)} \mathbf{U}^{(v)} \right\|_F^2 + \lambda_1 \sum_{v=1}^m \left\| \mathbf{U}^{(v)} - \mathbf{U}^{(v)} \mathbf{Z} \right\|_F^2, \quad (4)$$

其中 λ_1 在两个损失函数项之间进行权衡. 此外, 为了使相似度矩阵尽可能稀疏, 对 \mathbf{Z} 使用了 l_1 范数进行限制. 最终目标函数表达式如下所示:

$$\begin{aligned} \min_{\mathbf{U}^{(v)}, \mathbf{Z}} \sum_{v=1}^m \left\| \mathbf{X}^{(v)} - \mathbf{X}^{(v)} \mathbf{U}^{(v)} \right\|_F^2 + \lambda_1 \sum_{v=1}^m \left\| \mathbf{U}^{(v)} - \mathbf{U}^{(v)} \mathbf{Z} \right\|_F^2 + \lambda_2 \|\mathbf{Z}\|_1, \\ \text{s.t. } \text{diag}(\mathbf{Z}) = 0, \end{aligned} \quad (5)$$

其中 λ_2 用来权衡损失函数项和正则项. 为了避免 \mathbf{Z} 是单位矩阵, 将 $\text{diag}(\mathbf{Z})$ 设为零.

3.3 优化

基于 ADMM (alternating direction method of multipliers), 提出了一种优化算法来解决式 (5) 的问题, 如算法 1 所示. 引入 \mathbf{J} 作为辅助变量, 代替 \mathbf{Z} 进行稀疏约束, 目标函数变为

$$\begin{aligned} \min_{\mathbf{U}^{(v)}, \mathbf{Z}} \sum_{v=1}^m \left\| \mathbf{X}^{(v)} - \mathbf{X}^{(v)} \mathbf{U}^{(v)} \right\|_F^2 + \lambda_1 \sum_{v=1}^m \left\| \mathbf{U}^{(v)} - \mathbf{U}^{(v)} \mathbf{Z} \right\|_F^2 + \lambda_2 \|\mathbf{Z}\|_1, \\ \text{s.t. } \mathbf{J} = \mathbf{Z}, \quad \text{diag}(\mathbf{J}) = 0, \quad \text{diag}(\mathbf{Z}) = 0. \end{aligned} \quad (6)$$

算法 1 ISMSC 的优化算法

- 1: Input: Multi-view data matrix \mathbf{X} , superparameter λ and α ;
 - 2: Initialize: $\mathbf{J} = 0, \mathbf{Z} = 0, \mathbf{U}^{(v)} = 0, \mu = 10^{-4}; \mu_{\max} = 10^{10}; \rho = 2; \text{maxiter} = 200, \text{iter} = 0; \text{tol} < 10^{-8}$;
 - 3: **while** iter < maxiter and not converged **do**
 - 4: iter = iter + 1;
 - 5: Update \mathbf{Z} via (9);
 - 6: Update \mathbf{J} via (11);
 - 7: **for** v from 1 to m **do**
 - 8: Update $\mathbf{U}^{(v)}$ via (13);
 - 9: **end for**
 - 10: Update multipliers \mathbf{Y} and penalty parameter μ via (14);
 - 11: **end while**
 - 12: Perform spectral clustering algorithm with \mathbf{Z} as the similarity matrix.
 - 13: Clustering results.
-

通过使用增广拉格朗日乘子法 (augmented Lagrange multiplier, ALM), 上述问题等价于

$$\begin{aligned} \min_{\mathbf{U}^{(v)}, \mathbf{J}} \sum_{v=1}^m \left\| \mathbf{X}^{(v)} - \mathbf{X}^{(v)} \mathbf{U}^{(v)} \right\|_F^2 + \lambda_1 \sum_{v=1}^m \left\| \mathbf{U}^{(v)} - \mathbf{U}^{(v)} \mathbf{Z} \right\|_F^2 + \lambda_2 \|\mathbf{J}\|_1, \\ + \langle \mathbf{Y}, \mathbf{J} - \mathbf{Z} \rangle + \frac{\mu}{2} \|\mathbf{J} - \mathbf{Z}\|_F^2, \end{aligned} \quad (7)$$

其中, $\langle \cdot, \cdot \rangle$ 表示两矩阵的欧几里得 (Euclid) 内积, \mathbf{Y} 是拉格朗日乘子, $\|\cdot\|_F$ 是 Forbenius 范数, $\mu \in \mathbb{R}^+$ 是惩罚项的参数. 式 (5) 可以分解为 4 个子问题, 每个子问题均有闭合解. 在每次迭代过程中, 根据以下步骤对变量进行更新.

(1) 求解 \mathbf{Z} . 保持其他变量不变, 目标函数变为

$$\mathbf{Z} = \operatorname{argmin} \lambda_1 \sum_{v=1}^m \left\| \mathbf{U}^{(v)} - \mathbf{U}^{(v)} \mathbf{Z} \right\|_F^2 + \langle \mathbf{Y}, \mathbf{J} - \mathbf{Z} \rangle + \frac{\mu}{2} \|\mathbf{J} - \mathbf{Z}\|_F^2. \quad (8)$$

使用上述函数对 \mathbf{Z} 求导, 并将导数值设为零, 得到

$$\mathbf{Z} = \left(\lambda_1 \sum_{v=1}^m (\mathbf{U}^{(v)})^T \mathbf{U}^{(v)} + \mu \mathbf{I} \right)^{-1} \left(\mu \left(\mathbf{J} + \frac{\mathbf{Y}}{\mu} \right) + \lambda_1 \sum_{v=1}^m (\mathbf{U}^{(v)})^T \mathbf{U}^{(v)} \right). \quad (9)$$

(2) 求解 \mathbf{J} . 将其他变量保持不变, 更新 \mathbf{J} :

$$\begin{aligned} \mathbf{J} &= \operatorname{argmin} \lambda_2 \|\mathbf{J}\|_1 + \langle \mathbf{Y}, \mathbf{J} - \mathbf{Z} \rangle + \frac{\mu}{2} \|\mathbf{J} - \mathbf{Z}\|_F^2, \\ &= \operatorname{argmin} \lambda_2 \|\mathbf{J}\|_1 + \frac{\mu}{2} \left\| \mathbf{J} - \mathbf{Z} + \frac{\mathbf{Y}}{\mu} \right\|_F^2. \end{aligned} \quad (10)$$

对 \mathbf{J} 求导, 并将导数值设为 0, 得到

$$\mathbf{J} = \mathbf{S}_{\frac{\lambda_2}{\mu}} \left(\mathbf{Z} - \frac{\mathbf{Y}}{\mu} \right), \quad (11)$$

其中 \mathbf{S} 是软阈值操作符^[32].

(3) 求解 $\mathbf{U}^{(v)}$. 保持其他变量不变, 关于 $\mathbf{U}^{(v)}$ 的目标函数变为

$$\mathbf{U}^{(v)} = \operatorname{argmin} \sum_{v=1}^m \left\| \mathbf{X}^{(v)} - \mathbf{X}^{(v)} \mathbf{U}^{(v)} \right\|_F^2 + \lambda_1 \sum_{v=1}^m \left\| \mathbf{U}^{(v)} - \mathbf{U}^{(v)} \mathbf{Z} \right\|_F^2. \quad (12)$$

将上述函数对 $\mathbf{U}^{(v)}$ 求导, 并将导数值设为 0, 求得

$$(\mathbf{X}^{(v)})^T \mathbf{X}^{(v)} \mathbf{U}^{(v)} + \lambda_1 \mathbf{U}^{(v)} (\mathbf{I} - \mathbf{Z}^T - \mathbf{Z} + \mathbf{Z} \mathbf{Z}^T) = (\mathbf{X}^{(v)})^T \mathbf{X}^{(v)}. \quad (13)$$

这是一个标准的西尔维斯特 (Sylvester) 方程, 可以使用 Bartels-Stewart 算法^[33] 进行求解.

(4) 更新乘子 \mathbf{Y} 和惩罚项参数 μ . 对 \mathbf{Y} 和 μ 的更新过程如下:

$$\mathbf{Y} = \mathbf{Y} + \mu(\mathbf{J} - \mathbf{Z}), \quad \mu = \min(\mu_{\max}, \rho \times \mu), \quad (14)$$

其中 $\rho > 1$.

表 2 数据集的相关信息

Table 2 The statistic information of the datasets

Dataset	Instances	Clusters	Views	Dimensions	Domain
MSRC	210	7	6	1320, 48, 512, 100, 256, 210	Object image
Politics-uk	419	5	9	419, 419, 419, 419, 419, 419, 2879, 3614, 19868	Social network
ORL	400	40	3	4096, 3304, 6750	Face image
Still-DB	467	6	3	200, 200, 200	Action image
UCI	2000	10	3	216, 76, 64	Handwritten digit
COIL-20	1440	20	3	1024, 3304, 6750	Object image
Extended-YaleB	650	10	3	2500, 3304, 6750	Face image
Caltech-101	8677	101	4	2048, 4800, 3540, 1240	Object image

3.4 计算复杂度和收敛性

本文 ISMSC 的求解涉及到 4 个子问题. 为表述规范, 用 n 代表样本数量, 在求解 Z 时使用了矩阵的逆运算, 故复杂度为 $O(n^3)$; 更新 J 使用了 Hadamard 乘积, 复杂度为 $O(n^2)$; 求解 $U^{(v)}$ 涉及了矩阵乘法和西尔维斯特方程, 故复杂度为 $O(n^3)$; 更新乘子的复杂度为 $O(n^2)$, 所以每次迭代的复杂度为 $O(n^3 + n^2)$. 关于收敛性, 难以从理论上证明, 与文献 [12, 26, 34] 类似, 我们在实验环节对收敛性进行讨论. 在真实世界数据集上进行实验验证, 证明了 ISMSC 是收敛的.

4 实验

4.1 聚类数据集和评估指标

实验部分使用 ISMSC 在多个数据集上进行了聚类, 包括微软逐像素标记的图像数据集 v1 (MSRC)¹⁾、英国政客推特数据集 (Politics-uk)²⁾、ORL 人脸数据集 (ORL)³⁾、静态影像行为数据集 (Still-DB)^[35]、UCI 手写体识别数据集 (UCI)⁴⁾、纽约市哥伦比亚大学 (Columbia University in the City of New York) 图像数据集 (COIL-20)⁵⁾、扩展版耶鲁人脸数据库 (Extended-YaleB)⁶⁾、加利福尼亚理工学院 (California Institute of Technology) 101 类图像数据库 (Caltech-101)^[36]. 所用数据集的统计性信息参见表 2, 主要包括实例数目、所属领域、视角数量、维度和类别数量. 实验中所用的数据集均是预处理过的数据集, 其中 ORL, COIL-20, MSRC, Politics-uk 是由文献 [18] 提供的, UCI 是由文献 [9] 提供的, Caltech-101 则是由文献 [34] 提供的. 关于特征提取的详细信息, 可以参考上述文献.

本文使用 6 个主流的指标来评估性能: 标准化互信息 (normalized mutual Information, NMI)、准确性 (accuracy, ACC)、F 分数 (F-score, F)、调整后的兰德指数 (adjusted rand index, ARI)、精度 (precision, P) 和召回率 (recall, R). 在这些指标中, ACC 和 NMI 是使用最广的指标. 调整后的兰德指数、F 分数、精度和召回率也被广泛使用. 这些度量标准的值用来表示聚类的质量: 值越高, 效果越好. 这些指数的定义参见文献 [37].

1) <http://research.microsoft.com/en-us/projects/objectclassrecognition/>.

2) <http://mlg.ucd.ie/aggregation/index.html>.

3) <http://www.uk.research.att.com/facedatabase.html>.

4) <http://archive.ics.uci.edu/ml/datasets/Multiple+Features>.

5) <http://www.cs.columbia.edu/CAVE/software/softlib/>.

6) <http://cvc.cs.yale.edu/cvc/projects/yalefacesB/yalefacesB.html>.

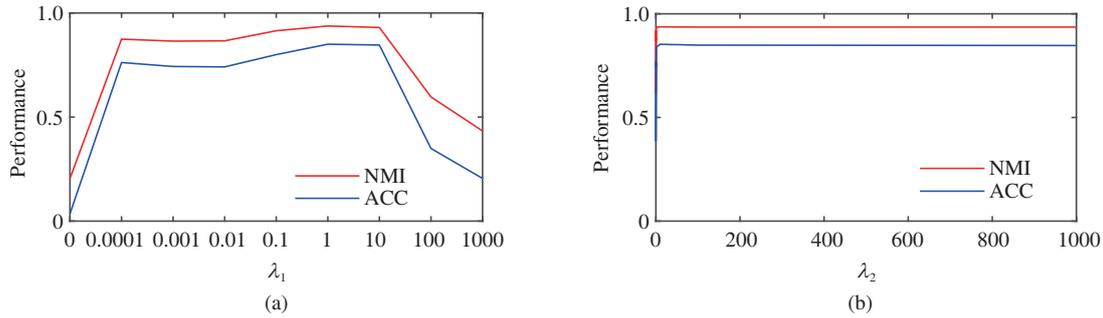


图 2 (网络版彩图) 参数的敏感度分析. (a) λ_1 ; (b) λ_2
 Figure 2 (Color online) The sensitivity analysis of parameters. (a) λ_1 ; (b) λ_2

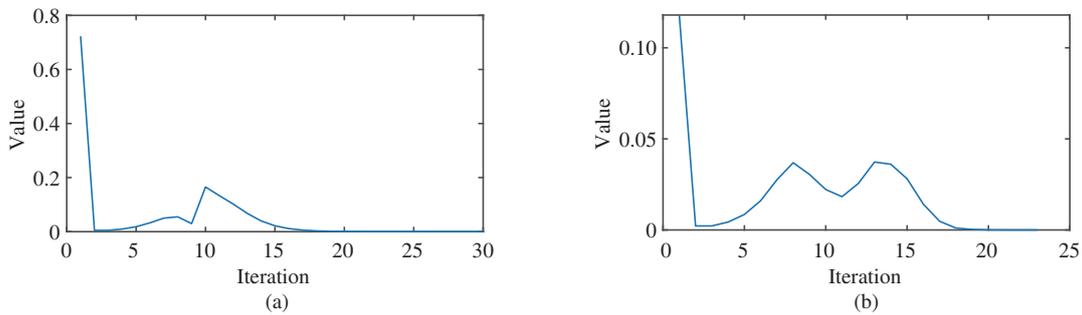


图 3 (网络版彩图) 收敛性分析. (a) Extended-YaleB; (b) MSRC
 Figure 3 (Color online) The convergence analysis. (a) Extended-YaleB; (b) MSRC

4.2 参数的敏感性分析和收敛性分析

ISMSC 算法中共有两个超参, λ_1 和 λ_2 , 它们的取值都设在了 0.0001 和 1000 之间. 本文在 ORL 数据集上对它们进行了敏感性测验. 实验时, 其中一个参数的值不变, 另一个参数的值在值域 [0.001 0.01 0.1 1 10 100 1000] 中变化. 对于每个参数组合, 进行 5 次实验, 记录 NMI 和 ACC 结果的平均值. λ_1 的敏感度测试如图 2(a) 所示, 该参数是用来权衡潜在表示到本质表示的重建错误, 测试结果表明当 λ_1 小于 0.0001 时, 随着 λ_1 的增加, 性能变好; 当 λ_1 在 0.0001 和 10 之间变化时性能稳定, 这表明从潜在表示中学习相似度矩阵能产生较好的结果; 此外, 当 λ_1 大于 10 时, 性能降低, 原因在于当 λ_1 太大时对原始数据的使用过少. λ_2 的敏感度测试如图 2(b) 所示, 该参数用来权衡正则项, 当值大于 1 时性能相对稳定; 随着 λ_2 的增加, \mathbf{Z} 变得尽可能稀疏; 但是, 当 \mathbf{Z} 接近零时, $\lambda_1 \sum_{v=1}^m \|\mathbf{U}^{(v)} - \mathbf{U}^{(v)} \mathbf{Z}\|_F^2$ 这项的值会变大, 由于目标函数的值越小越好, $\lambda_1 \sum_{v=1}^m \|\mathbf{U}^{(v)} - \mathbf{U}^{(v)} \mathbf{Z}\|_F^2$ 和 $\lambda_2 \|\mathbf{Z}\|_1$ 这两项会互相制约, 当它们的和达到最小时停止变化.

本文在 Extended-YaleB 和 MSRC 数据集上对 ISMSC 的收敛性进行了分析, 结果如图 3 所示. 由于矩阵的初始值均设为 0, 所以曲线开始时有些波动. 最终目标值都趋于 0, 这表明 ISMSC 是收敛的.

4.3 对比方法

为了验证 ISMSC 的有效性, 将 ISMSC 和 9 种基准方法进行了比较. 这 9 种方法是: 谱聚类 (spectral clustering, SC) [38]、稀疏子空间聚类 (sparse subspace clustering, SSC) [4]、低秩表示子空间聚类 (low-rank representation, LRR) [14]、基于最小化分歧的两视角聚类 (minimizing-disagreement

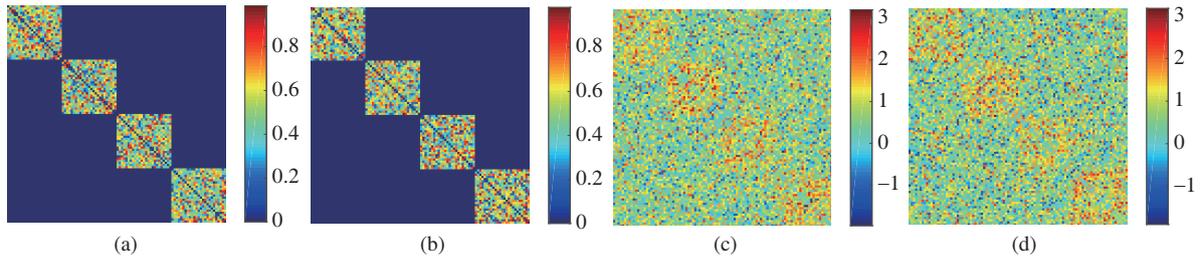


图 4 (网络版彩图) 人造数据集的可视化. (a) 未添加噪声时的视角 1; (b) 未添加噪声时的视角 2; (c) 添加高斯噪声后的视角 1; (d) 添加高斯噪声后的视角 2

Figure 4 (Color online) Visualization of the synthetic dataset. (a) Visualization for view 1 without noise; (b) visualization for view 2 without noise; (c) visualization for view 1 with Gaussian noise; (d) visualization for view 2 with Gaussian noise

表 3 人造数据集上的对比结果 (均值)

Table 3 Results (mean) on Toy dataset

Method	NMI	ACC	F	ARI	P	R
SC	0.633	0.715	0.664	0.554	0.655	0.674
SSC	0.072	0.387	0.268	0.030	0.265	0.271
LRR	0.628	0.830	0.698	0.597	0.678	0.718
sM-D	0.925	0.957	0.942	0.922	0.935	0.951
Co-Reg	0.858	0.950	0.899	0.867	0.896	0.902
DIMSC	0.939	0.980	0.959	0.946	0.958	0.960
LMSC	0.909	0.970	0.939	0.920	0.938	0.940
MSSC	0.633	0.715	0.664	0.554	0.655	0.674
MSC_IAS	1.000	1.000	1.000	1.000	1.000	1.000
ISMSC	0.970	0.990	0.973	0.990	0.979	0.980

spectral clustering, sM-D)^[39]、基于共同正则化的聚类 (co-regularized spectral clustering, Co-Reg)^[10]、多样性诱导的多视角子空间聚类 (diversity-induced multi-view subspace clustering, DIMSC) (CVPR 2015)^[40]、潜在多视角子空间聚类 (latent multi-view subspace clustering, LMSC) (CVPR 2017)^[11]、基于稀疏表示的多模态子空间聚类 (multi-modal sparse representation-based subspace clustering, MSSC) (Information Fusion 2018)^[31]、具有完整感知相似性的多视角子空间聚类 (MSC_IAS) (Pattern Recognition 2019)^[26]。这 9 种方法中, 前 3 个是单视角方法, 中间两个是经典的多视角聚类基准方法, 其余 4 个是最新的多视角子空间聚类基准方法。

4.4 验证性实验

4.4.1 关于减弱噪声影响的验证性实验

为了测试 ISMSC 是否可以减弱原始数据中噪声的影响, 我们设计了 100 个数据点的两视角数据集. 该数据集由对角排列的四个 25×25 块矩阵组成. 块中的值在 0 到 1 范围内随机生成, 块外的值设置为零. 我们在数据集中添加了平均值为 0.5、方差值为 1 的高斯噪声. 无噪声的原始矩阵和添加高斯噪声后的矩阵如图 4 所示. 我们在此数据集上对所有方法进行了实验, 结果如表 3 所示. ISMSC 的 NMI 结果仅次于 MSC_IAS, 比 SC, SSC, LRR, sM-D, Co-Reg, DIMSC, LMSC, MSSC 分别高 33.7%, 89.8%, 34.2%, 34.2%, 4.5%, 8.4%, 11.2%, 3.1%, 33.7%. 结果表明, 与其他方法相比, ISMSC 可以更好地应对原始特征中的噪声.

为了分析 ISMSC 对噪声具有鲁棒性的原因, 我们设计了一种名为 S-MS (sparse multi-view subspace clustering) 的算法, 测试 ISMSC 是否是通过间接学习邻接矩阵应对噪声的. S-MS 的目标函

表 4 针对减弱噪声影响的验证性实验结果 (均值)

Table 4 Verification test results (mean) on the cause of noise reduction

Method	NMI	ACC	F	ARI	P	R
S-MSD	0.692	0.750	0.703	0.608	0.701	0.706
ISMSD	0.970	0.990	0.973	0.990	0.979	0.980

表 5 在 ORL 数据集上不同表示的验证性实验结果 (均值)^{a)}

Table 5 Verification test results (mean) about different representations on ORL dataset

Representation	NMI	ACC	F	ARI	P	R
$\mathbf{U}^{(1)}$	0.812	0.660	0.543	0.532	0.512	0.579
$\mathbf{U}^{(2)}$	0.916	0.802	0.751	0.745	0.707	0.801
$\mathbf{U}^{(3)}$	0.865	0.729	0.650	0.641	0.612	0.693
\mathbf{Z}	0.943	0.861	0.823	0.819	0.789	0.861

a) $\mathbf{U}^{(v)}$ denotes the underlying representation of the v th view and \mathbf{Z} is the intrinsic representation.

数为

$$\min_{\mathbf{Z}} \sum_{v=1}^m \frac{1}{2} \|\mathbf{X}^{(v)} - \mathbf{X}^{(v)} \mathbf{Z}\|_F^2 + \lambda \|\mathbf{Z}\|_1 \quad (15)$$

s.t. $\text{diag}(\mathbf{Z}) = 0.$

式 (15) 可以通过 ADMM 策略求解.

将式 (15) 和 ISMSD 的目标函数式 (5) 对比可以发现, S-MSD 和 ISMSD 之间的主要区别在于 S-MSD 是从原始数据中学习相似度矩阵, 而 ISMSD 则从潜在表示中去学习. 实验结果如表 4 所示, 就 NMI 而言, ISMSD 的结果比 S-MSD 高 27.8%. 由于这两种方法的区别在于是否直接从原始特征中学习相似度矩阵, 因此可以得出结论: ISMSD 通过从数据的潜在表示中学习相似度矩阵, 减弱了原始数据中噪声对它的影响.

4.4.2 关于本质表示的验证性实验

为了分析学习到的本质表示是否比潜在表示能更好地表达聚类结构, 表 5 对每个潜在表示的相似度矩阵以及本质表示的相似度矩阵进行了比较, 结果表明本质表示比潜在表示效果好. 就聚类而言, 本质相似度矩阵优于潜在相似度矩阵, 因为本质相似度矩阵的学习过程, 充分利用了不同视角的一致性信息和互补知识.

4.5 基准数据集上的实验

我们进一步在 8 个基准数据集上进行了实验, 包括 3 个对象聚类数据集、两个人脸聚类数据集、一个动作聚类数据集、一个文档聚类数据集、一个手写数字聚类数据集和一个社交网络数据集. 对于对象聚类, 使用了 MSRC, COIL-20 和 Caltech-101 数据集; 对于人脸聚类, 在两个基准数据集 ORL 和 Extended-YaleB 上进行实验; 对于动作聚类, 使用 Still-DB 数据集; 对于手写数字聚类, 在 UCI 上进行了实验; 对于社交网络聚类, 使用 Politics-uk 数据集. 结果显示在表 6~9 中. 其中, 每列值最好的两个结果加粗表示.

我们根据上述方法所在文献中的设置进行调参. 但是, 在使用 DIMSD 处理 Politics-uk 这个数据集时, 在迭代过程中出现“此西尔维斯特方程的解不存在或不唯一”的错误, 在表 6 中使用 ‘-’ 表示此问题. 此外, 当使用 LMSD 在 Caltech-101 数据集上进行实验时, 一些矩阵的值在经过几次迭代后变为 ‘Inf’, 在求解西尔维斯特方程问题上也出现了问题, 我们在表 9 中使用 ‘-’ 表示此错误.

表 6 在 MSRC 和 Politics-uk 上的对比结果 (均值)

Table 6 Results (mean) on MSRC and Politics-uk

Method	MSRC						Politics-uk					
	NMI	ACC	F	AR	P	R	NMI	ACC	F	AR	P	R
SC	0.609	0.695	0.578	0.508	0.57	0.586	0.611	0.585	0.598	0.444	0.821	0.47
SSC	0.617	0.709	0.567	0.495	0.553	0.58	0.574	0.709	0.694	0.563	0.881	0.572
LRR	0.562	0.65	0.526	0.448	0.517	0.536	0.566	0.606	0.624	0.48	0.855	0.492
sM-D	0.635	0.747	0.619	0.556	0.608	0.631	0.522	0.66	0.617	0.46	0.807	0.499
Co-Reg	0.688	0.773	0.666	0.61	0.648	0.686	0.59	0.647	0.622	0.468	0.818	0.502
DIMSC	0.666	0.743	0.645	0.586	0.627	0.663	-	-	-	-	-	-
LMSC	0.632	0.717	0.605	0.541	0.597	0.613	0.323	0.613	0.524	0.308	0.636	0.445
MSSC	0.768	0.869	0.766	0.728	0.759	0.759	0.579	0.789	0.762	0.646	0.883	0.67
MSC_IAS	0.69	0.73	0.629	0.565	0.599	0.662	0.773	0.815	0.817	0.731	0.962	0.711
ISMSC	0.824	0.91	0.825	0.797	0.821	0.83	0.885	0.964	0.957	0.932	0.98	0.935

表 7 在 ORL 和 Still-DB 上的对比结果 (均值)

Table 7 Results (mean) on ORL and Still-DB

Method	ORL						Still-DB					
	NMI	ACC	F	AR	P	R	NMI	ACC	F	AR	P	R
SC	0.901	0.76	0.707	0.699	0.655	0.768	0.109	0.3	0.224	0.064	0.225	0.223
SSC	0.919	0.79	0.745	0.738	0.689	0.811	0.118	0.336	0.251	0.084	0.236	0.268
LRR	0.851	0.731	0.62	0.61	0.582	0.663	0.115	0.315	0.234	0.058	0.215	0.257
sM-D	0.907	0.765	0.714	0.707	0.655	0.785	0.112	0.332	0.238	0.07	0.227	0.245
Co-Reg	0.866	0.717	0.636	0.627	0.59	0.691	0.081	0.312	0.225	0.054	0.213	0.239
DIMSC	0.933	0.833	0.793	0.788	0.751	0.84	0.128	0.315	0.25	0.081	0.234	0.268
LMSC	0.919	0.801	0.75	0.744	0.697	0.813	0.136	0.322	0.249	0.087	0.241	0.257
MSSC	0.925	0.79	0.75	0.744	0.68	0.838	0.143	0.339	0.246	0.087	0.242	0.25
MSC_IAS	0.93	0.849	0.801	0.796	0.763	0.843	0.137	0.357	0.270	0.090	0.234	0.319
Proposed	0.943	0.861	0.823	0.819	0.789	0.861	0.144	0.362	0.276	0.107	0.251	0.306

表 8 在 UCI 和 COIL-20 上的对比结果 (均值)

Table 8 Results (mean) on UCI and COIL-20

Method	UCI						COIL-20					
	NMI	ACC	F	AR	P	R	NMI	ACC	F	AR	P	R
SC	0.661	0.717	0.616	0.573	0.609	0.627	0.802	0.672	0.637	0.617	0.594	0.688
SSC	0.826	0.796	0.767	0.739	0.716	0.827	0.957	0.835	0.862	0.855	0.794	0.945
LRR	0.548	0.579	0.484	0.426	0.478	0.491	0.774	0.672	0.578	0.553	0.515	0.658
sM-D	0.752	0.774	0.708	0.674	0.682	0.737	0.793	0.665	0.636	0.616	0.602	0.674
Co-Reg	0.733	0.743	0.685	0.649	0.662	0.71	0.788	0.644	0.616	0.595	0.574	0.667
DIMSC	0.455	0.56	0.413	0.347	0.408	0.417	0.847	0.779	0.742	0.728	0.736	0.748
LMSC	0.791	0.861	0.778	0.753	0.769	0.788	0.839	0.749	0.701	0.685	0.664	0.743
MSSC	0.837	0.912	0.837	0.819	0.834	0.84	0.854	0.77	0.748	0.735	0.73	0.768
MSC_IAS	0.941	0.974	0.948	0.943	0.948	0.949	0.953	0.866	0.856	0.848	0.794	0.932
ISMSC	0.913	0.959	0.920	0.911	0.919	0.921	0.957	0.852	0.867	0.860	0.796	0.952

实验结果表明,在这 10 种方法中, MAS_IAS 和本文方法 ISMSC 表现最为突出. 从平均性能的角度而言, ISMSC 是最好的, 它在 MSRC, Politics-uk, ORL, Still-DB 和 Extend-YaleB 数据集上的所有度量中均达到最佳性能. 特别是在 Politics-uk 数据集上, 就 NMI, ACC, F 和 R 而言, ISMSC 分别以 11.2%, 14.9%, 14.0% 和 8.8% 的优势超过了第 2 名. 针对结果, 我们有如下的思考.

(1) 多视角聚类方法在大部分数据集上结果均优于单视角聚类方法. 多视角数据比单视角数据包含更多的信息, 通过利用多个视角数据间的互补知识, 可以更好地对数据进行聚类. 但是, SSC 的性能在某些数据集上仍然出色, 这证明了数据自表达特性的优势.

(2) ISMSC 作为一种多视角子空间聚类方法, 适合处理高维数据. 在 Politics-uk, Extended-YaleB,

表 9 在 Extended-YaleB 和 Caltech-101 上的对比结果 (均值)

Table 9 Results (mean) on Extended-YaleB and Caltech-101

Method	Extended-YaleB						Caltech-101					
	NMI	ACC	F	AR	P	R	NMI	ACC	F	AR	P	R
SC	0.354	0.382	0.297	0.218	0.289	0.306	0.821	0.566	0.387	0.377	0.664	0.273
SSC	0.479	0.474	0.308	0.217	0.262	0.376	0.814	0.558	0.374	0.364	0.641	0.264
LRR	0.775	0.757	0.645	0.603	0.606	0.69	0.82	0.585	0.4	0.39	0.682	0.283
sM-D	0.192	0.261	0.176	0.084	0.172	0.181	0.779	0.526	0.371	0.36	0.628	0.263
Co-Reg	0.138	0.202	0.1548	0.06	0.152	0.157	0.82	0.567	0.391	0.381	0.668	0.277
DIMSC	0.629	0.603	0.499	0.441	0.47	0.532	0.831	0.57	0.4	0.39	0.706	0.279
LMSC	0.529	0.534	0.36	0.279	0.315	0.421	-	-	-	-	-	-
MSSC	0.444	0.443	0.279	0.182	0.233	0.349	0.835	0.606	0.423	0.413	0.715	0.3
MSC-IAS	0.806	0.811	0.741	0.712	0.720	0.765	0.814	0.597	0.459	0.447	0.608	0.370
ISMSC	0.857	0.822	0.793	0.769	0.764	0.823	0.840	0.605	0.451	0.440	0.669	0.334

表 10 与最新多视角子空间聚类方法运行时间比较 (s)

Table 10 The running time comparison with state-of-the-art multi-view subspace clustering methods (s)

Method	Time complexity	MSRC	Politics-uk	Still-DB	ORL	Extended-YaleB
DIMSC	$O(n^3 + n^2)$	4.5	-	15.3	11.1	42.2
LMSC	$O(d^3 + n^3)$	5.4	86.3	15.7	42.6	83.6
MSSC	$O(n^3 + n^2)$	4.4	79.8	3.1	80.6	52.0
MSC-IAS	$O(n^2)$	9.8	7.5	22.3	23.9	26.4
ISMSC	$O(n^3 + n^2)$	4.3	66.1	18.8	11.1	56.6

ORL 等多个数据集上的结果表明了 ISMSC 的有效性.

(3) 与其他 9 种方法相比, ISMSC 在图像聚类数据集、动作聚类数据集、手写数字聚类数据集、社交网络数据集和人脸聚类数据集上取得了更好的结果. 与对比方法相比, 它对不同类型的数据更具有鲁棒性.

(4) ISMSC 的性能远远优于 MSSC. MSSC 从原始数据中学习相似度矩阵, 而 ISMSC 是从潜在表示中学习的, 这一事实进一步证明了间接学习相似度矩阵的优势.

(5) ISMSC 的性能远超过 DIMSC. 与 DIMSC 相比, ISMSC 通过将多个视角的潜在表示融合到本质表示中, 可以更好地利用不同视角之间的一致信息和补充信息, 这证明了本质表示的有效性.

为了测试 ISMSC 的运行效率, 我们在配置为 Intel(R) Core(TM) i7-8565U CPU 和 8 G RAM 的机器上对最新的 5 种方法进行了实验. 在每个数据集上, 都使用达到最佳性能时的参数进行实验, 每个实验重复 3 次, 计算运行时间的平均值, 结果记录在表 10 中, 其中 n 表示样本数, d 是所有视角数据的维度之和, 运行时间最短的两个结果以粗体表示. 在这 5 种方法中, MAS-IAS 的时间复杂度最小, ISMSC 与 DIMSC 和 MSSC 相同. 但是, 实际的运行时间受许多因素影响, 例如迭代次数、样本数量、数据维度、视角数量等. 在 MSRC, Politics-uk, Still-DB, ORL 和 Extended-YaleB 这 5 个数据集的实验结果中, ISMSC 在 MSRC, Politics-uk 和 ORL 这 3 个数据集上运行时间较短.

5 结论

本文提出了一种基于本质自表示的 ISMSC 模型用于多视角子空间聚类. ISMSC 从相应视角的潜在表示而不是从原始特征中建立相似度矩阵, 充分利用了各视角之间的一致信息和互补知识; 将潜在表示和本质表示的学习统一到一个框架中, 对潜在表示和本质表示进行了联合优化. 与其他方法相比, ISMSC 更易于处理高维数据, 可以减弱原始数据中噪声的影响, 具有更强的鲁棒性. 未来, 我们将持续改进模型, 提升对不完整数据集的处理能力, 并探索其在多模态场景中的应用.

参考文献

- 1 Xie J Y, Gao H C, Xie W X. K-nearest neighbors optimized clustering algorithm by fast search and finding the density peaks of a dataset. *Sci Sin Inform*, 2016, 46: 258–280 [谢娟英, 高红超, 谢维信. K 近邻优化的密度峰值快速搜索聚类算法. *中国科学: 信息科学*, 2016, 46: 258–280]
- 2 Li F J, Qian Y H, Wang J T, et al. Clustering method based on sample's stability. *Sci Sin Inform*, 2020, 50: 1239–1254 [李飞江, 钱宇华, 王婕婷, 等. 基于样本稳定性的聚类方法. *中国科学: 信息科学*, 2020, 50: 1239–1254]
- 3 Yang A Y, Wright J, Ma Y, et al. Unsupervised segmentation of natural images via lossy data compression. *Comput Vision Image Underst*, 2008, 110: 212–225
- 4 Elhamifar E, Vidal R. Sparse subspace clustering. In: *Proceedings of Conference on Computer Vision and Pattern Recognition*, Miami, 2009. 2790–2797
- 5 Qu W, Wang D L, Feng S, et al. A novel cross-modal hashing algorithm based on multimodal deep learning. *Sci China Inf Sci*, 2017, 60: 092104
- 6 Wang W, Hu L H, Hu Z Y. Energy-based multi-view piecewise planar stereo. *Sci China Inf Sci*, 2017, 60: 032101
- 7 Zhu X, Huang Z, Yang Y, et al. Self-taught dimensionality reduction on the high-dimensional small-sized data. *Pattern Recogn*, 2013, 46: 215–229
- 8 Hu R, Zhu X, Cheng D, et al. Graph self-representation method for unsupervised feature selection. *Neurocomputing*, 2017, 220: 130–137
- 9 Brbić M, Kopriva I. Multi-view low-rank sparse subspace clustering. *Pattern Recogn*, 2018, 73: 247–258
- 10 Kumar A, Rai P, Daume H. Co-regularized multi-view spectral clustering. In: *Proceedings of the 24th International Conference on Neural Information Processing Systems*, 2011. 1413–1421
- 11 Zhang C, Hu Q, Fu H, et al. Latent multi-view subspace clustering. In: *Proceedings of Conference on Computer Vision and Pattern Recognition*, Honolulu, 2017. 4279–4287
- 12 Zheng Q, Zhu J, Li Z, et al. Feature concatenation multi-view subspace clustering. *Neurocomputing*, 2020, 379: 89–102
- 13 Vidal R. Subspace clustering. *IEEE Signal Process Mag*, 2011, 28: 52–68
- 14 Liu G, Lin Z, Yan S, et al. Robust recovery of subspace structures by low-rank representation. *IEEE Trans Pattern Anal Mach Intell*, 2013, 35: 171–184
- 15 Luo D, Nie F, Ding C, et al. Multi-subspace representation and discovery. In: *Machine Learning and Knowledge Discovery in Databases*. Berlin: Springer, 2011. 405–420
- 16 Hu H, Lin Z, Feng J, et al. Smooth representation clustering. In: *Proceedings of Conference on Computer Vision and Pattern Recognition*, Columbus, 2014. 3834–3841
- 17 Ou W, Yu S, Li G, et al. Multi-view non-negative matrix factorization by patch alignment framework with view consistency. *Neurocomputing*, 2016, 204: 116–124
- 18 Zhang C, Fu H, Hu Q, et al. Generalized latent multi-view subspace clustering. *IEEE Trans Pattern Anal Mach Intell*, 2020, 42: 86–99
- 19 von Luxburg U. A tutorial on spectral clustering. *Stat Comput*, 2007, 17: 395–416
- 20 Elhamifar E, Vidal R. Sparse subspace clustering: algorithm, theory, and applications. *IEEE Trans Pattern Anal Mach Intell*, 2013, 35: 2765–2781
- 21 Xia R, Pan Y, Du L, et al. Robust multi-view spectral clustering via low-rank and sparse decomposition. In: *Proceedings of the 28th AAAI Conference on Artificial Intelligence*, Québec, 2014. 2149–2155
- 22 Wang Y, Wu L, Lin X, et al. Multiview spectral clustering via structured low-rank matrix factorization. *IEEE Trans Neural Netw Learn Syst*, 2018, 29: 4833–4843
- 23 Tang C, Zhu X, Liu X, et al. Learning a joint affinity graph for multiview subspace clustering. *IEEE Trans Multimedia*, 2019, 21: 1724–1736
- 24 Yin M, Gao J, Xie S, et al. Multiview subspace clustering via tensorial t-product representation. *IEEE Trans Neural Netw Learning Syst*, 2019, 30: 851–864
- 25 Xie Y, Tao D, Zhang W, et al. On unifying multi-view self-representations for clustering by tensor multi-rank minimization. *Int J Comput Vis*, 2018, 126: 1157–1179
- 26 Wang X, Lei Z, Guo X, et al. Multi-view subspace clustering with intactness-aware similarity. *Pattern Recogn*, 2019, 88: 50–63
- 27 Zhang G Y, Zhou Y R, He X Y, et al. One-step kernel multi-view subspace clustering. *Knowledge-Based Syst*, 2020,

- 189: 105126
- 28 Li R, Zhang C, Fu H, et al. Reciprocal multi-layer subspace learning for multi-view clustering. In: Proceedings of International Conference on Computer Vision, Seoul, 2019. 8172–8180
- 29 Kang Z, Shi G, Huang S, et al. Multi-graph fusion for multi-view spectral clustering. Knowledge-Based Syst, 2020, 189: 105102
- 30 Weng W, Zhou W, Chen J, et al. Enhancing multi-view clustering through common subspace integration by considering both global similarities and local structures. Neurocomputing, 2020, 378: 375–386
- 31 Abavisani M, Patel V M. Multimodal sparse and low-rank subspace clustering. Inf Fusion, 2018, 39: 168–177
- 32 Cai J F, Candés E J, Shen Z. A singular value thresholding algorithm for matrix completion. SIAM J Optim, 2010, 20: 1956–1982
- 33 Bartels R H, Stewart G W. Solution of the matrix equation $AX + XB = C$ [F4]. Commun ACM, 1972, 15: 820–826
- 34 Wu J, Lin Z, Zha H. Essential tensor learning for multi-view spectral clustering. IEEE Trans Image Process, 2019, 28: 5910–5922
- 35 Ikizler N, Cinbis R G, Pehlivan S, et al. Recognizing actions from still images. In: Proceedings of the 19th International Conference on Pattern Recognition, Tampa, 2008. 1–4
- 36 Li F-F, Fergus R, Perona P. Learning generative visual models from few training examples: an incremental Bayesian approach tested on 101 object categories. In: Proceedings of Conference on Computer Vision and Pattern Recognition Workshop, Washington, 2004. 178–178
- 37 Zhan K, Nie F, Wang J, et al. Multiview consensus graph clustering. IEEE Trans Image Process, 2019, 28: 1261–1270
- 38 Ng A Y, Jordan M I, Weiss Y. On spectral clustering. In: Proceedings of the 14th International Conference on Neural Information Processing Systems: Natural and Synthetic, 2002. 849–856
- 39 de Sa V R. Spectral clustering with two views. In: Proceedings of the 22nd ICML Workshop on Learning with Multiple Views, Bonn, 2005. 20–27
- 40 Cao X, Zhang C, Fu H, et al. Diversity-induced multi-view subspace clustering. In: Proceedings of Conference on Computer Vision and Pattern Recognition, Boston, 2015. 586–594

Intrinsic self-representation for multi-view subspace clustering

Xiao YU^{1,2}, Hui LIU^{1,2,3*}, Yan WU³ & Caiming ZHANG^{2,4}

1. School of Computer Science and Technology, Shandong University of Finance and Economics, Jinan 250014, China;

2. Digital Media Technology Key Lab of Shandong Province, Jinan 250014, China;

3. Medical Center, Stanford University, Palo Alto 94305, USA;

4. Software College, Shandong University, Jinan 250010, China

* Corresponding author. E-mail: liuh_lh@sdufe.edu.cn

Abstract Clustering of high-dimensional data has been a hotspot in the era of big data. Meanwhile, changes in data acquisition have promoted the rapid rise of multi-view learning. Under these circumstances, multi-view subspace clustering has attracted much attention. However, many studies learn the affinity matrix from the original features and therefore, cannot obtain a satisfying affinity matrix due to the noise in the raw data. In this paper, we propose a novel model called intrinsic self-representation for multi-view subspace clustering (ISMSC), which learns the affinity matrix from the underlying representations of multiple views. By fully exploiting the self-representation property of data, ISMSC fuses the features of all views into an intrinsic representation. Moreover, an algorithm is designed to optimize the model by using the alternating direction method of multipliers (ADMM) strategy. Finally, experiments on a synthetic dataset show the robustness to noise, and experiments on different kinds of real-world baseline datasets show the effectiveness of our model.

Keywords multi-view, clustering, subspace, spectral clustering, sparse



Xiao YU was born in 1981. She received her M.S. degree from School of Computer Science and Technology, Shandong University, China, in 2006. She is a Ph.D. candidate as well as a teacher at School of Computer Science and Technology, Shandong University of Finance and Economics. Her research interests include machine learning and data mining.



Hui LIU was born in 1978. She received her M.S. and Ph.D. degrees at School of Computer Science and Technology in Shandong University in 2004 and 2008 respectively. She is now a professor and Ph.D supervisor at Shandong University of Finance and Economics. Her research interests include data mining and visualization.



Yan WU was born in 1976. She received her M.S. degree at Department of Electronic Engineering, University of Illinois, Urbana-Champaign in 2003, and Ph.D. degree at Department of Electronic Engineering, University of Wisconsin-Madison in 2008. Her research interests include MRI and artificial intelligence. She is an associate professor at the Medical Center of Stanford University.



Caiming ZHANG was born in 1955. He received his B.S. and M.E. degrees in computer science from Shandong University, in 1982 and 1984 respectively, and his Ph.D. degree in computer science from the Tokyo Institute of Technology, Japan, in 1994. His research interests include CAGD, CG, information visualization, and medical image processing. He is currently a professor and Ph.D. supervisor at School of Computer Science and Technology, Shandong University.