

# 基于低秩约束的熵加权多视角模糊聚类算法

张嘉旭<sup>1</sup> 王骏<sup>1,2</sup> 张春香<sup>1</sup> 林得富<sup>1</sup> 周塔<sup>3</sup> 王士同<sup>1</sup>

**摘要** 如何有效挖掘多视角数据内部的一致性以及差异性,是构建多视角模糊聚类算法的两个重要问题. 本文在 Co-FKM 算法框架上,提出了基于低秩约束的熵加权多视角模糊聚类算法 (Entropy-weighting multi-view fuzzy C-means with low rank constraint, LR-MVEWFCM). 一方面,从视角之间的一致性出发,引入核范数对多个视角之间的模糊隶属度矩阵进行低秩约束;另一方面,基于香农熵理论引入视角权重自适应调整策略,使算法根据各视角的重要程度来处理视角间的差异性. 本文使用交替方向乘子法 (Alternating direction method of multipliers, ADMM) 进行目标函数的优化. 最后,人工模拟数据集和 UCI (University of California Irvine) 数据集上进行的实验结果验证了该方法的有效性.

**关键词** 多视角模糊聚类, 香农熵, 低秩约束, 核范数, 交替方向乘子法

**引用格式** 张嘉旭, 王骏, 张春香, 林得富, 周塔, 王士同. 基于低秩约束的熵加权多视角模糊聚类算法. 自动化学报, 2022, 48(7): 1760-1770

**DOI** 10.16383/j.aas.c190350

## Entropy-weighting Multi-view Fuzzy C-means With Low Rank Constraint

ZHANG Jia-Xu<sup>1</sup> WANG Jun<sup>1,2</sup> ZHANG Chun-Xiang<sup>1</sup> LIN De-Fu<sup>1</sup> ZHOU Ta<sup>3</sup> WANG Shi-Tong<sup>1</sup>

**Abstract** Effective mining both internal consistency and diversity of multi-view data is important to develop multi-view fuzzy clustering algorithms. In this paper, we propose a novel multi-view fuzzy clustering algorithm called entropy-weighting multi-view fuzzy c-means with low-rank constraint (LR-MVEWFCM). On the one hand, we introduce the nuclear norm as the low-rank constraint of the fuzzy membership matrix. On the other hand, the adaptive adjustment strategy of view weight is introduced to control the differences among views according to the importance of each view. The learning criterion can be optimized by the alternating direction method of multipliers (ADMM). Experimental results on both artificial and UCI (University of California Irvine) datasets show the effectiveness of the proposed method.

**Key words** Multi-view fuzzy clustering, Shannon entropy, low-rank constraint, nuclear norm, alternating direction method of multipliers (ADMM)

**Citation** Zhang Jia-Xu, Wang Jun, Zhang Chun-Xiang, Lin De-Fu, Zhou Ta, Wang Shi-Tong. Entropy-weighting multi-view fuzzy C-means with low rank constraint. *Acta Automatica Sinica*, 2022, 48(7): 1760-1770

随着多样化信息获取技术的发展,人们可以从不同途径或不同角度来获取对象的特征数据,即多视角数据. 多视角数据包含了同一对象不同角度的信息. 例如: 网页数据中既包含网页内容又包含网页链接信息; 视频内容中既包含视频信息又包含音频信息; 图像数据中既涉及颜色直方图特征、纹理

特征等图像特征,又涉及描述该图像内容的文本. 多视角学习能有效地对多视角数据进行融合,避免了单视角数据信息单一的问题<sup>[1-4]</sup>.

多视角模糊聚类是一种有效的无监督多视角学习方法<sup>[5-7]</sup>. 它通过在多视角聚类过程中引入各样本对不同类别的模糊隶属度来描述各视角下样本属于该类别的不确定性程度. 经典的工作有: 文献 [8] 以经典的单视角模糊 C 均值 (Fuzzy C-means, FCM) 算法作为基础模型,利用不同视角间的互补信息确定协同聚类的准则,提出了 Co-FC (Collaborative fuzzy clustering) 算法; 文献 [9] 参考文献 [8] 的协同思想提出 Co-FKM (Multiview fuzzy clustering algorithm collaborative fuzzy K-means) 算法,引入双视角隶属度惩罚项,构造了一种新型的无监督多视角协同学习方法; 文献 [10] 借鉴了 Co-FKM 和 Co-FC 所使用的双视角约束思想,通过引入视角权重,并采用集成策略来融合多视角的模糊隶属

收稿日期 2019-05-09 录用日期 2019-07-17

Manuscript received May 9, 2019; accepted July 17, 2019

国家自然科学基金 (61772239), 江苏省自然科学基金 (BK20181339) 资助

Supported by National Natural Science Foundation of China (61772239) and Natural Science Foundation of Jiangsu Province (BK20181339)

本文责任编辑 刘艳军

Recommended by Associate Editor LIU Yan-Jun

1. 江南大学数字媒体学院 无锡 214122 2. 上海大学通信与信息工程学院 上海 200444 3. 江苏科技大学电子信息学院 镇江 212100

1. School of Digital Media, Jiangnan University, Wuxi 214122

2. School of Communication and Information Engineering, Shanghai University, Shanghai 200444 3. School of Electronic Information, Jiangsu University of Science and Technology, Zhenjiang 212100

度矩阵, 提出了 WV-Co-FCM (Weighted view collaborative fuzzy C-means) 算法; 文献 [11] 通过最小化双视角下样本与聚类中心的欧氏距离来减小不同视角间的差异性, 基于 K-means 聚类框架提出了 Co-K-means (Collaborative multi-view K-means clustering) 算法; 在此基础上, 文献 [12] 提出了基于模糊划分的 TW-Co-K-means (Two-level weighted collaborative K-means for multi-view clustering) 算法, 对 Co-K-means 算法中的双视角欧氏距离加入一致性权重, 获得了比 Co-K-means 更好的多视角聚类结果. 以上多视角聚类方法都基于成对视角来构造不同的正则化项来挖掘视角之间的一致性和差异性信息, 缺乏对多个视角的整体考虑.

一致性和差异性设计多视角聚类算法需要考虑的两个重要原则<sup>[10-14]</sup>. 一致性是指在多视角聚类过程中, 各视角的聚类结果应该尽可能保持一致. 在设计多视角聚类算法时, 往往通过协同、集成等手段来构建全局划分矩阵, 从而得到最终的聚类结果<sup>[14-16]</sup>. 差异性是指多视角数据中的每个视角均反映了对象在不同方面的信息, 这些信息互为补充<sup>[10]</sup>, 在设计多视角聚类算法时需要对这些信息进行充分融合. 综合考虑这两方面的因素, 本文拟提出新型的低秩约束熵加权多视角模糊聚类算法 (Entropy-weighting multi-view fuzzy C-means with low rank constraint, LR-MVEWFCM), 其主要创新点可以概括为以下 3 个方面:

1) 在模糊聚类框架下提出了面向视角一致性的低秩约束准则. 已有的多视角模糊聚类算法大多基于成对视角之间的两两关系来构造正则化项, 忽视了多个视角的整体一致性信息. 本文在模糊聚类框架下从视角全局一致性出发引入低秩约束正则化项, 从而得到新型的低秩约束多视角模糊聚类算法.

2) 在模糊聚类框架下同时考虑多视角聚类的一致性和差异性, 在引入低秩约束的同时进一步使用面向视角差异性的多视角香农熵加权策略; 在迭代优化的过程中, 通过动态调节视角权重系数来突出具有更好分离性的视角的权重, 从而提高聚类性能.

3) 在模糊聚类框架下首次使用交替方向乘法 (Alternating direction method of multipliers, ADMM)<sup>[15]</sup> 对 LR-MVEWFCM 算法进行优化求解.

在本文中, 令  $N$  为样本总量,  $D$  为样本维度,  $K$  为视角数目,  $C$  为聚类数目,  $m$  为模糊指数. 设  $\mathbf{x}_{j,k}$  表示多视角场景中第  $j$  个样本第  $k$  个视角的特征向量,  $j = 1, \dots, N$ ,  $k = 1, \dots, K$ ;  $\mathbf{v}_{i,k}$  表示第  $k$  个视角下, 第  $i$  个聚类中心,  $i = 1, \dots, C$ ;  $U_k = [\mu_{ij,k}]$  表示第  $k$  个视角下的模糊隶属度矩阵, 其中  $\mu_{ij,k}$  是第  $k$  个视角下第  $j$  个样本属于第  $i$  个聚类中心的模

糊隶属度,  $i = 1, \dots, C$ ,  $j = 1, \dots, N$ .

本文第 1 节在相关工作中回顾已有的经典模糊 C 均值聚类算法 FCM 模型<sup>[17]</sup> 和多视角模糊聚类 Co-FKM 模型<sup>[9]</sup>; 第 2 节将低秩理论与多视角香农熵理论相结合, 提出本文的新方法; 第 3 节基于模拟数据集和 UCI (University of California Irvine) 数据集验证本文算法的有效性, 并给出实验分析; 第 4 节给出实验结论.

## 1 相关工作

### 1.1 模糊 C 均值聚类算法 FCM

设单视角环境下样本  $\mathbf{x}_1, \dots, \mathbf{x}_N \in \mathbf{R}^D$ ,  $U = [\mu_{i,j}]$  是模糊划分矩阵,  $V = [\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_C]$  是样本的聚类中心. FCM 算法的目标函数可表示为

$$J_{\text{FCM}} = \sum_{i=1}^C \sum_{j=1}^N \mu_{ij}^m \|\mathbf{x}_j - \mathbf{v}_i\|^2$$

$$\sum_{i=1}^C \mu_{ij} = 1, \quad 0 < \mu_{ij} \leq 1, \quad m > 1 \quad (1)$$

可得到  $J_{\text{FCM}}$  取得局部极小值的必要条件为

$$\mathbf{v}_i = \frac{\sum_{j=1}^N (\mu_{ij})^m \mathbf{x}_j}{\sum_{j=1}^N (\mu_{ij})^m} \quad (2)$$

$$\mu_{ij} = \frac{1}{\sum_{s=1}^c \left( \frac{\|\mathbf{x}_j - \mathbf{v}_i\|^2}{\|\mathbf{x}_j - \mathbf{v}_s\|^2} \right)^{\frac{1}{m-1}}} \quad (3)$$

根据式 (2) 和式 (3) 进行迭代优化, 使目标函数收敛于局部极小点, 从而得到样本属于各聚类中心的模糊划分矩阵  $U$ .

### 1.2 多视角模糊聚类 Co-FKM 模型

在经典 FCM 算法的基础上, 文献 [9] 通过引入视角协同约束正则项, 对视角间的一致性信息加以约束, 提出了多视角模糊聚类 Co-FKM 模型.

多视角模糊聚类 Co-FKM 模型需要满足如下条件:

$$\sum_{i=1}^C \mu_{ij,k} = 1,$$

$$\mu_{ij,k} \leq 1, \quad i = 1, 2, \dots, C; \quad j = 1, 2, \dots, N \quad (4)$$

多视角模糊聚类 Co-FKM 模型的目标函数  $J_{\text{Co-FKM}}$  定义为

$$J_{\text{Co-FKM}} = \sum_{k=1}^K \sum_{i=1}^C \sum_{j=1}^N \mu_{ij,k}^m \|\mathbf{x}_{j,k} - \mathbf{v}_{i,k}\|^2 + \eta \Delta_k \quad (5)$$

$$\Delta_k = \frac{1}{K-1} \sum_{k=1}^K \sum_{k \neq \bar{k}}^C \sum_{j=1}^N (\mu_{ij, \bar{k}}^m - \mu_{ij, k}^m) \|\mathbf{x}_{j, k} - \mathbf{v}_{i, k}\|^2 \quad (6)$$

式(5)中,  $\eta$ 表示协同划分参数;  $\Delta$ 表示视角一致项, 由式(6)可知, 当各视角趋于一致时,  $\Delta$ 将趋于0.

迭代得到各视角的模糊隶属度  $\mu_{ij, k}$  后, 为了最终得到一个具有全局性的模糊隶属度划分矩阵, Co-FKM 算法对各视角下的模糊隶属度采用几何平均的方法, 得到数据集的整体划分, 具体形式为

$$\hat{\mu}_{ij} = \sqrt[K]{\prod_{k \in K} \mu_{ij, k}} \quad (7)$$

其中,  $\hat{\mu}_{ij}$  为全局模糊划分结果.

## 2 基于低秩约束的熵加权多视角模糊聚类算法

针对当前多视角模糊聚类算法研究中存在的不足, 本文提出一种基于低秩约束的熵加权多视角模糊聚类新方法 LR-MVEWFCM. 一方面通过向多视角模糊聚类算法的目标学习准则中引入低秩约束项, 在整体上控制聚类过程中各视角的一致性; 另一方面基于香农熵理论, 通过熵加权机制来控制各视角之间的差异性. 同时使用交替方向乘法对模型进行优化求解.

设多视角隶属度  $U_1, \dots, U_K$  融合为一个整体的隶属度矩阵  $U$ , 将矩阵  $U$  的秩函数凸松弛为核范数, 通过对矩阵  $U$  进行低秩约束, 可以将多视角数据之间的一致性转化为核范数最小化问题进行求解, 具体定义为

$$\Gamma(U) = \|U\|_* \quad (8)$$

其中,  $U = [U_1 \ \dots \ U_K]^T$  表示全局划分矩阵,  $\|\cdot\|_*$  表示核范数. 式(8)的优化过程保证了全局划分矩阵的低秩约束. 低秩约束的引入, 可以弥补当前大多数多视角聚类算法仅能基于成对视角构建约束的缺陷, 从而更好地挖掘多视角数据中包含的全局一致性信息.

目前已有的多视角的聚类算法在处理多视角数据时, 通常默认每个视角平等共享聚类结果<sup>[11]</sup>, 但实际上某些视角的数据往往因空间分布重叠而导致可分性较差. 为避免此类视角的数据过多影响聚类效果, 本文拟对各视角进行加权处理, 并构建香农熵正则项从而在聚类过程中有效地调节各视角之间的权重, 使得具有较好可分离性的视角的权重系数尽可能大, 以达到更好的聚类效果.

令视角权重系数  $\sum_{k=1}^K w_k = 1$  且  $w_k \geq 0$ , 则香农熵正则项表示为

$$f(w_k) = - \sum_{k=1}^K w_k \ln w_k \quad (9)$$

综上所述, 本文作如下改进: 首先, 用本文提出的低秩约束全局模糊隶属度矩阵  $U$ ; 其次, 计算损失函数时考虑视角权重  $w_k$ , 并加入视角权重系数的香农熵正则项. 设  $U = [U_1 \ \dots \ U_K]^T$ ;  $\mathbf{w} = [w_1, \dots, w_k, \dots, w_K]$  表示  $K$  个视角下的视角权重. 本文所构建 LR-MVEWFCM 的目标函数为

$$J(\mathbf{w}, U, V) = \sum_{k=1}^K w_k \sum_{i=1}^C \sum_{j=1}^N (\mu_{ij, k})^m \|\mathbf{x}_{j, k} - \mathbf{v}_{i, k}\|^2 + \theta \|U\|_* + \lambda \sum_{k=1}^K w_k \ln w_k \quad (10)$$

其中, 约束条件为

$$\begin{cases} \sum_{i=1}^C \mu_{ij, k} = 1, & \mu_{ij, k} \in [0, 1] \\ \sum_{k=1}^K w_k = 1, & w_k \in [0, 1] \end{cases} \quad (11)$$

本文取模糊指数  $m = 2$ .

### 2.1 基于 ADMM 的求解算法

在本节中, 我们将使用 ADMM 方法, 通过交替方向迭代的策略来实现目标函数(11)的最小化.

$$f(\mathbf{w}, U, V) = \sum_{k=1}^K w_k \sum_{i=1}^C \sum_{j=1}^N (\mu_{ij, k})^2 \times \|\mathbf{x}_{j, k} - \mathbf{v}_{i, k}\|^2 + \lambda \sum_{k=1}^K w_k \ln w_k \quad (12)$$

$$g(Z) = \theta \|Z\|_* \quad (13)$$

最小化式(10)可改写为如下约束优化问题:

$$\min L(\mathbf{w}, U, V, Z) = f(\mathbf{w}, U, V) + g(Z)$$

$$\begin{cases} \sum_{i=1}^C \mu_{ij, k} = 1, & \mu_{ij, k} \in [0, 1] \\ \sum_{k=1}^K w_k = 1, & w_k \in [0, 1] \end{cases} \quad (14)$$

$$Z = U$$

其求解过程可分解为如下几个子问题:

1)  $V$ -子问题. 固定  $\mathbf{w}$  和  $U$ , 更新  $V$  为

$$\mathbf{v}_{i,k}^{(t+1)} = \arg \min_{\mathbf{v}_{i,k}^t} \sum_{k=1}^K w_k^{(t)} \times \sum_{i=1}^C \sum_{j=1}^N \left( \mu_{ij,k}^{(t)} \right)^2 \left\| \mathbf{x}_{j,k} - \mathbf{v}_{i,k}^{(t)} \right\|^2 \quad (15)$$

通过最小化式 (15), 可得到  $\mathbf{v}_{i,k}^{(t+1)}$  的闭合解为

$$\mathbf{v}_{i,k}^{(t+1)} = \frac{\sum_{j=1}^N \left( \mu_{ij,k}^{(t)} \right)^2 \mathbf{x}_{j,k}}{\sum_{j=1}^N \left( \mu_{ij,k}^{(t)} \right)^2} \quad (16)$$

2)  $U$ -子问题. 固定  $\mathbf{w}$ ,  $Q$  和  $Z$ , 更新  $U$  为

$$U^{(t+1)} = \arg \min_{U^{(t)}} f(\mathbf{w}, U, V) + \alpha \left( \sum_{k=1}^K w_k^{(t)} - 1 \right) + \sum_{j=1}^N \sum_{k=1}^K \beta_{j,k} \left( \sum_{i=1}^C \mu_{ij,k}^{(t)} - 1 \right) + \frac{\rho}{2} \left\| U^{(t)} - Z^{(t)} + Q^{(t)} \right\|_F^2 \quad (17)$$

通过最小化式 (17), 可得到  $U^{(t+1)}$  的封闭解为

$$\mu_{ij,k}^{(t+1)} = \frac{\rho z_{ij,k}^{(t)} - \rho q_{ij,k}^{(t)}}{2w_k^{(t)} \left\| \mathbf{x}_{j,k} - \mathbf{v}_{i,k}^{(t+1)} \right\|^2 + \rho} - \left( \sum_{i=1}^C \frac{\rho z_{ij,k}^{(t)} - \rho q_{ij,k}^{(t)}}{2w_k^{(t)} \left\| \mathbf{x}_{j,k} - \mathbf{v}_{i,k}^{(t+1)} \right\|^2 + \rho} - 1 \right) \times \left( \frac{1}{\sum_{i=1}^C \frac{1}{2w_k^{(t)} \left\| \mathbf{x}_{j,k} - \mathbf{v}_{i,k}^{(t+1)} \right\|^2 + \rho}} \right) \times \left( \frac{1}{2w_k^{(t)} \left\| \mathbf{x}_{j,k} - \mathbf{v}_{i,k}^{(t+1)} \right\|^2 + \rho} \right) \quad (18)$$

3)  $w$ -子问题. 固定  $V$  和  $U$ , 更新  $\mathbf{w}$  为

$$w_k^{(t+1)} = \frac{\exp \left( \frac{-\sum_{i=1}^C \sum_{j=1}^N \left( \mu_{ij,k}^{(t+1)} \right)^2 \left\| \mathbf{x}_{j,k} - \mathbf{v}_{i,k}^{(t+1)} \right\|^2}{\lambda} - 1 \right)}{\sum_{k=1}^K \exp \left( \frac{-\sum_{i=1}^C \sum_{j=1}^N \left( \mu_{ij,k}^{(t+1)} \right)^2 \left\| \mathbf{x}_{j,k} - \mathbf{v}_{i,k}^{(t+1)} \right\|^2}{\lambda} - 1 \right)} \quad (19)$$

4)  $Z$ -子问题. 固定  $Q$  和  $U$ , 更新  $Z$  为

$$Z^{(t+1)} = \arg \min_{Z^{(t)}} Z^{(t)} + \frac{\rho}{2} \left\| U^{(t+1)} - Z^{(t)} + Q^{(t)} \right\|_F^2 = \theta \left\| Z^{(t)} \right\|_* + \frac{\rho}{2} \left\| U^{(t+1)} - Z^{(t)} + Q^{(t)} \right\|_F^2 \quad (20)$$

通过引入软阈值算子, 可得式 (20) 的解为

$$Z^{(t+1)} = D_{\frac{\theta}{\rho}} \left( U^{(t+1)} + Q^{(t)} \right) = AS_{\frac{\theta}{\rho}}(\Sigma)B^T \quad (21)$$

其中,  $U^{(t+1)} + Q^{(t)} = A\Sigma B^T$  为矩阵  $U^{(t+1)} + Q^{(t)}$  的奇异值分解, 核范数的近邻算子可由软阈值算子  $S_{\theta/\rho}(\Sigma) = \text{diag}(\{\max(0, \sigma_i - \theta/\rho)\})$  ( $i = 1, 2, \dots, N$ ) 给出.

5)  $Q$ -子问题. 固定  $Z$  和  $U$ , 更新  $Q$  为

$$Q^{(t+1)} = Q^{(t)} + \frac{\rho}{2} \left( U^{(t+1)} - Z^{(t+1)} \right) \quad (22)$$

经过上述迭代过程, 目标函数收敛于局部极值, 同时得到不同视角下的模糊隶属度矩阵. 本文借鉴文献 [10] 的集成策略, 使用视角权重系数  $\mathbf{w} = [w_1, \dots, w_k, \dots, w_K]$  和模糊隶属度矩阵  $U$  来构建具有全局特性的模糊空间划分矩阵  $\tilde{U}$ :

$$\tilde{U} = \sum_{k=1}^K w_k U_k \quad (23)$$

其中,  $w_k$ ,  $U_k$  分别表示第  $k$  个视角的视角权重系数和相应的模糊隶属度矩阵.

LR-MVEWFCM 算法描述如下:

**输入.** 包含  $K$  ( $1 \leq k \leq K$ ) 个视角的多视角样本集, 其中任意一个视角对应样本集  $X_k = \{\mathbf{x}_{1,k}, \dots, \mathbf{x}_{N,k}\}$ , 聚类中心  $C$ , 迭代阈值  $\epsilon$ , 最大迭代次数  $T$ ;

**输出.** 各视角聚类中心  $\mathbf{v}_{i,k}^{(t)}$ , 模糊空间划分矩阵  $\tilde{U}$  和各视角权重  $w_k$ ;

**步骤 1.** 随机初始化  $V^{(t)}$ , 归一化  $U^{(t)}$  及  $\mathbf{w}^{(t)}$ ,  $t = 0$ ;

**步骤 2.** 根据式 (21) 更新  $\mathbf{v}_{i,k}^{(t+1)}$ ;

**步骤 3.** 根据式 (23) 更新  $U^{(t+1)}$ ;

**步骤 4.** 根据式 (24) 更新  $w_k^{(t+1)}$ ;

**步骤 5.** 根据式 (26) 更新  $Z^{(t+1)}$ ;

**步骤 6.** 根据式 (27) 更新  $Q^{(t+1)}$ ;

**步骤 7.** 如果  $\|L^{(t+1)} - L^{(t)}\| < \epsilon$  或者  $t > T$ , 则算法结束并跳出循环, 否则, 返回步骤 2;

**步骤 8.** 根据步骤 7 所获取的各视角权重  $w_k$  及各视角下的模糊隶属度  $U_k$ , 使用式 (23) 计算  $\tilde{U}$ .

## 2.2 讨论

### 2.2.1 与低秩约束算法比较

近年来, 基于低秩约束的机器学习模型得到了广泛的研究. 经典工作包括文献 [16] 中提出 LRR (Low rank representation) 模型, 将矩阵的秩函数凸松弛为核范数, 通过求解核范数最小化问题, 求得基于低秩表示的亲和矩阵; 文献 [14] 提出低秩张量多视角子空间聚类算法 (Low-rank tensor constrained multiview subspace clustering, LT-MS),



在各视角间求出带有低秩约束的子空间表示矩阵; 文献 [18] 则进一步将低秩约束引入多模型子空间聚类算法中, 使算法模型取得了较好的性能. 本文将低秩约束与多视角模糊聚类框架相结合, 提出了 LR-MVEWFCM 算法, 用低秩约束来实现多视角数据间的一致性. 本文方法可作为低秩模型在多视角模糊聚类领域的重要拓展.

### 2.2.2 与多视角 Co-FKM 算法比较

图 1 和图 2 分别给出了多视角 Co-FKM 算法和本文 LR-MVEWFCM 算法的工作流程.

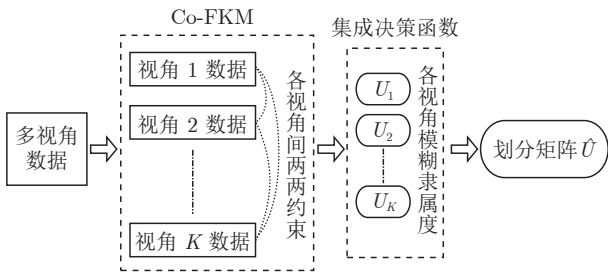


图 1 Co-FKM 算法处理多视角聚类任务工作流程  
Fig.1 Co-FKM algorithm for multi-view clustering task

本文算法与经典的多视角 Co-FKM 算法在多视角信息的一致性约束和多视角聚类结果的集成策略上均有所不同. 在多视角信息的一致性约束方面, 本文将 Co-FKM 算法中的视角间两两约束进一步扩展到多视角全局一致性约束; 在多视角聚类结果的集成策略上, 本文不同于 Co-FKM 算法对隶属度矩阵简单地求几何平均值的方式, 而是将各视角隶属度与视角权重相结合, 构建具有视角差异性的集成决策函数.

## 3 实验与分析

### 3.1 实验设置

本文采用模拟数据集和 UCI 中的真实数据集

表 1 参数定义和设置  
Table 1 Parameter setting in the experiments

算法	算法说明	参数设置
FCM	经典的单视角模糊聚类算法	模糊指数 $m = \frac{\min(N, D-1)}{\min(N, D-1)-2}$ , 其中, $N$ 表示样本数, $D$ 表示样本维数
CombKM	组合 K-means 算法	—
Co-FKM	多视角协同划分的模糊聚类算法	模糊指数 $m = \frac{\min(N, D-1)}{\min(N, D-1)-2}$ , 协同学习系数 $\eta \in \frac{K-1}{K}$ , 其中, $K$ 为视角数, 步长 $\rho = 0.01$
Co-Clustering	基于样本与特征空间的协同聚类算法	正则化系数 $\lambda \in \{10^{-3}, 10^{-2}, \dots, 10^3\}$ , 正则化系数 $\mu \in \{10^{-3}, 10^{-2}, \dots, 10^3\}$
LR-MVEWFCM	基于低秩约束的熵加权多视角模糊聚类算法	视角权重平衡因子 $\lambda \in \{10^{-5}, 10^{-4}, \dots, 10^5\}$ , 低秩约束正则项系数 $\theta \in \{10^{-3}, 10^{-2}, \dots, 10^3\}$ , 模糊指数 $m = 2$
MVEWFCM	LR-MVEWFCM 算法中低秩约束正则项系数 $\theta = 0$	视角权重平衡因子 $\lambda \in \{10^{-5}, 10^{-4}, \dots, 10^5\}$ , 模糊指数 $m = 2$

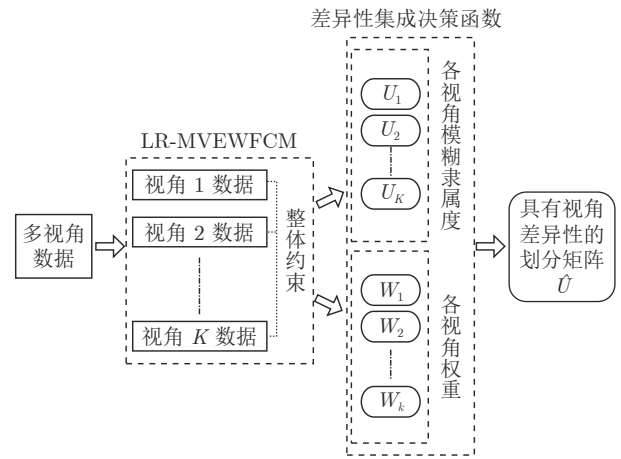


图 2 LR-MVEWFCM 算法处理多视角聚类任务工作流程  
Fig.2 LR-MVEWFCM algorithm for multi-view clustering task

进行实验验证, 选取 FCM<sup>[17]</sup>、CombKM<sup>[19]</sup>、Co-FKM<sup>[9]</sup> 和 Co-Clustering<sup>[20]</sup> 这 4 个聚类算法作为对比算法, 参数设置如表 1 所示. 实验环境为: Intel Core i5-7400 CPU, 其主频为 2.3 GHz, 内存为 8 GB. 编程环境为 MATLAB 2015b.

本文采用如下两个性能指标对各算法所得结果进行评估.

1) 归一化互信息 (Normalized mutual information, NMI)<sup>[10]</sup>

$$NMI = \frac{\sum_{i=1}^C \sum_{j=1}^C N_{i,j} \log_2 N \times \frac{N_{i,j}}{N_i} \times N_j}{\sqrt{\sum_{i=1}^C N_i \log_2 \frac{N_i}{N} \times \sum_{j=1}^C N_j \log_2 \frac{N_j}{N}}} \quad (24)$$

其中,  $N_{i,j}$  表示第  $i$  类与第  $j$  类的契合程度,  $N_i$  表示第  $i$  类中所属样本量,  $N_j$  表示第  $j$  类中所属样本量, 而  $N$  表示数据的样本总量;

2) 芮氏指标 (Rand index, RI)<sup>[10]</sup>

$$RI = \frac{f_{00} + f_{11}}{\frac{N(N-1)}{2}} \quad (25)$$

其中,  $f_{00}$  表示具有不同类标签且属于不同类的数据配对点数目,  $f_{11}$  则表示具有相同类标签且属于同一类的数据配对点数目,  $N$  表示数据的样本总量. 以上两个指标的取值范围介于  $[0, 1]$  之间, 数值越接近 1, 说明算法的聚类性能越好. 为了验证算法的鲁棒性, 各表中统计的性能指标值均为算法 10 次运行结果的平均值.

### 3.2 模拟数据集实验

为了评估本文算法在多视角数据集上的聚类效果, 使用文献 [10] 的方法来构造具有三维特性的模拟数据集  $A(x, y, z)$ , 其具体生成过程为: 首先在 MATLAB 环境下采用正态分布随机函数 `normrnd` 构建数据子集  $A_1(x, y, z)$ ,  $A_2(x, y, z)$  和  $A_3(x, y, z)$ , 每组对应一个类簇, 数据均包含 200 个样本.

其中第 1 组与第 2 组数据集在特征  $z$  上数值较为接近, 第 2 组与第 3 组数据集在特征  $x$  上较为接近; 然后将 3 组数据合并得到集合  $A(x, y, z)$ , 共计 600 个样本; 最后对数据集内的样本进行归一化处理. 我们进一步将特征  $x, y, z$  按表 2 的方式两两组合, 从而得到多视角数据.

表 2 模拟数据集特征组成

Table 2 Characteristic composition of simulated dataset

视角	包含特征
视角 1	$x, y$
视角 2	$y, z$
视角 3	$x, z$

将各视角下的样本可视化, 如图 3 所示.

通过观察图 3 可以发现, 视角 1 中的数据集在空间分布上具有良好的可分性, 而视角 2 和视角 3 的数据在空间分布上均存在着一定的重叠, 从而影

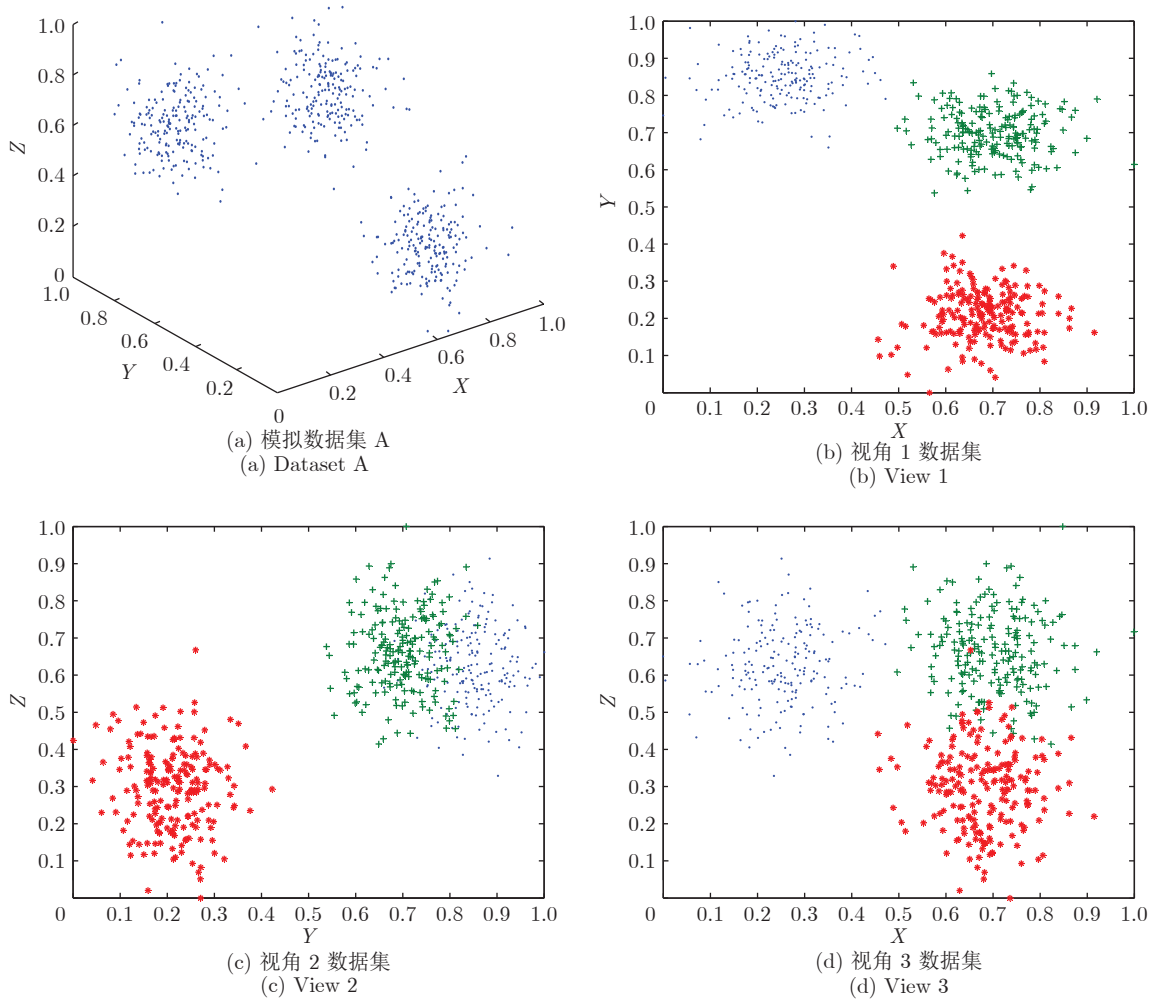


图 3 模拟数据集及各视角数据集

Fig. 3 Simulated data under multiple views

响了所在视角下的聚类性能. 通过组合不同视角生成若干新的数据集, 如表 3 所示, 并给出了 LR-MVEWFCM 重复运行 10 次后的平均结果和方差.

表 3 模拟数据实验算法性能对比  
Table 3 Performance comparison of the proposed algorithms on simulated dataset

编号	包含特征	NMI	RI
1	视角1	1.0000 ± 0.0000	1.0000 ± 0.0000
2	视角2	0.7453 ± 0.0075	0.8796 ± 0.0081
3	视角3	0.8750 ± 0.0081	0.9555 ± 0.0006
4	视角1, 视角2	1.0000 ± 0.0000	1.0000 ± 0.0000
5	视角1, 视角3	1.0000 ± 0.0000	1.0000 ± 0.0000
6	视角2, 视角3	0.9104 ± 0.0396	0.9634 ± 0.0192
7	视角2, 视角3	1.0000 ± 0.0000	1.0000 ± 0.0000

对比 LR-MVEWFCM 在数据集 1~3 上的性能, 我们发现本文算法在视角 1 上取得了最为理想的效果, 在视角 3 上的性能要优于视角 2, 这与图 3 中各视角数据的空间可分性是一致的. 此外, 将各视角数据两两组合构成新数据集 4~6 后, LR-MVEWFCM 算法都得到了比单一视角更好的聚类效果, 这都说明了本文采用低秩约束来挖掘多视角数据中一致性的方法, 能够有效提高聚类性能.

基于多视角数据集 7, 我们进一步给出本文算法与其他经典聚类算法的比较结果.

从表 4 中可以发现, 由于模拟数据集在某些特征空间下具有良好的空间可分性, 所以无论是本文的算法还是 Co-Clustering 算法、FCM 算法等算法均取得了很好的聚类效果, 而 CombKM 算法的性能较之以上算法则略有不足, 分析其原因在于 CombKM 算法侧重于挖掘样本之间的信息, 却忽视了多视角之间的协作, 而本文算法通过使用低秩约束进一步挖掘了多视角之间的全局一致性, 因而得到了比 CombKM 算法更好的聚类效果.

### 3.3 真实数据集实验

本节采用 5 个 UCI 数据集: 1) Iris 数据集; 2) Image Segmentation (IS) 数据集; 3) Balance 数据集; 4) Ionosphere 数据集; 5) Wine 数据集来进行

实验. 由于这几个数据集均包含了不同类型的特征, 所以可以将这些特征进行重新分组从而构造相应的多视角数据集. 表 5 给出了分组后的相关信息.

我们在多视角数据集上运行各多视角聚类算法; 同时原数据集上运行 FCM 算法. 相关结果统计见表 6 和表 7.

通过观察表 6 和表 7 中的 NMI 和 RI 指标值可知, Co-FKM 算法的聚类性能明显优于其他几种经典聚类算法, 而相比于 Co-FKM 算法, 由于 LR-MVEWFCM 采用了低秩正则项来挖掘多视角数据之间的一致性关系, 并引入多视角自适应熵加权策略, 从而有效控制各视角之间的差异性. 很明显, 这种聚类性能更为优异和稳定, 且收敛性的效果更好. 表 6 和表 7 中的结果也展示了在 IS、Balance、Iris、Ionosphere 和 Wine 数据集上, 其 NMI 和 RI 指标均提升 3~5 个百分点, 这也说明了本文算法在多视角聚类过程中的有效性.

为进一步说明本文低秩约束发挥的积极作用, 将 LR-MVEWFCM 算法和 MVEWFCM 算法共同进行实验, 算法的性能对比如图 4 所示.

从图 4 中不难发现, 无论在模拟数据集上还是 UCI 真实数据集上, 相比较 MVEWFCM 算法, LR-MVEWFCM 算法均可以取得更好的聚类效果. 由此可见, LR-MVEWFCM 目标学习准则中的低秩约束能够有效利用多视角数据的一致性来提高算法的聚类性能.

为研究本文算法的收敛性, 同样选取 8 个数据集进行收敛性实验, 其目标函数变化如图 5 所示.

从图 5 中可以看出, 本文算法在真实数据集上仅需迭代 15 次左右就可以趋于稳定, 这说明本文算法在速度要求较高的场景下具有较好的实用性.

综合以上实验结果, 我们不难发现, 在具有多视角特性的数据集上进行模糊聚类分析时, 多视角模糊聚类算法通常比传统单视角模糊聚类算法能够得到更优的聚类效果; 在本文中, 通过多视角模糊聚类学习中引入低秩约束来增强不同视角之间的一致性关系, 并引入香农熵调节视角权重关系, 控制不同视角之间的差异性, 从而得到了比其他多视角聚类算法更好的聚类效果.

表 4 模拟数据集 7 上各算法的性能比较  
Table 4 Performance comparison of the proposed algorithms on simulated dataset 7

数据集	指标	Co-Clustering	CombKM	FCM	Co-FKM	LR-MVEWFCM
A	NMI-mean	1.0000	0.9305	1.0000	1.0000	1.0000
	NMI-std	0.0000	0.1464	0.0000	0.0000	0.0000
	RI-mean	1.0000	0.9445	1.0000	1.0000	1.0000
	RI-std	0.0000	0.1171	0.0000	0.0000	0.0000

表 5 基于 UCI 数据集构造的多视角数据  
Table 5 Multi-view data constructed based on UCI dataset

编号	原数据集	说明	视角特征	样本	视角	类别
8	IS	Shape	9	2 310	2	7
		RGB	9			
9	Iris	Sepal长度	2	150	2	3
		Sepal宽度				
		Petal长度	2			
		Petal宽度				
10	Balance	天平左臂重量	2	625	2	3
		天平左臂长度				
		天平右臂重量	2			
		天平右臂长度				
11	Iris	Sepal长度	1	150	4	3
		Sepal宽度	1			
		Petal长度	1			
		Petal宽度	1			
12	Balance	天平左臂重量	1	625	4	3
		天平左臂长度	1			
		天平右臂重量	1			
		天平右臂长度	1			
13	Ionosphere	每个特征单独作为一个视角	1	351	34	2
14	Wine	每个特征单独作为一个视角	1	178	13	3

### 3.4 参数敏感性实验

LR-MVEWFCM 算法包含两个正则项系数,

即视角权重平衡因子  $\lambda$  和低秩约束正则项系数  $\theta$ , 图 6 以 LR-MVEWFCM 算法在模拟数据集 7 上的实验为例, 给出了系数从 0 到 1000 过程中, 算法性能的变化情况, 当低秩正则项系数  $\theta = 0$  时, 即不添加此正则项, 算法的性能最差, 验证了本文加入的低秩正则项的有效性, 当  $\theta$  值变化过程中, 算法的性能相对变化较小, 说明本文算法在此数据集上对于  $\theta$  值变化不敏感, 具有一定的鲁棒性; 而当香农熵正则项系数  $\lambda = 0$  时, 同样算法性能较差, 也说明引入此正则项的合理性. 当  $\lambda$  值变大时, 发现算法的性能也呈现变好趋势, 说明在此数据集上, 此正则项相对效果比较明显.

## 4 结束语

本文从多视角聚类学习过程中的一致性和差异性两方面出发, 提出了基于低秩约束的熵加权多视角模糊聚类算法. 该算法采用低秩正则项来挖掘多视角数据之间的一致性关系, 并引入多视角自适应熵加权策略从而有效控制各视角之间的差异性, 从而提高了算法的性能. 在模拟数据集和真实数据集上的实验均表明, 本文算法的聚类性能优于其他多视角聚类算法. 同时本文算法还具有迭代次数少、收敛速度快的优点, 具有良好的实用性. 由于本文采用经典的 FCM 框架, 使用欧氏距离来衡量数据对象之间的差异, 这使得本文算法不适用于某些高维数据场景. 如何针对高维数据设计多视角聚类算法, 这也将是我们今后的研究重点.

表 6 5 种聚类方法的 NMI 值比较结果  
Table 6 Comparison of NMI performance of five clustering methods

编号	Co-Clustering		CombKM		FCM		Co-FKM		LR-MVEWFCM
	均值	P-value	均值	P-value	均值	P-value	均值	P-value	均值
8	0.5771 ±	0.0019	0.5259 ±	0.2056	0.5567 ±	0.0044	0.5881 ±	3.76 × 10 <sup>-4</sup>	0.5828 ±
	0.0023		0.0551		0.0184		0.0109		0.0044
9	0.7582 ±	2.03 × 10 <sup>-24</sup>	0.7251 ±	2.32 × 10 <sup>-7</sup>	0.7578 ±	1.93 × 10 <sup>-24</sup>	0.8317 ±	8.88 × 10 <sup>-16</sup>	0.9029 ±
	7.4015 × 10 <sup>-17</sup>		0.0698		0.0698		0.0064		0.0057
10	0.2455 ±	0.0165	0.1562 ±	3.47 × 10 <sup>-5</sup>	0.1813 ±	0.0061	0.2756 ±	0.1037	0.3030 ±
	0.0559		0.0749		0.1172		0.0309		0.0402
11	0.7582 ±	2.28 × 10 <sup>-16</sup>	0.7468 ±	5.12 × 10 <sup>-16</sup>	0.7578 ±	5.04 × 10 <sup>-16</sup>	0.8244 ±	2.16 × 10 <sup>-16</sup>	0.8768 ±
	1.1703 × 10 <sup>-16</sup>		0.0079		1.1703 × 10 <sup>-16</sup>		1.1102 × 10 <sup>-16</sup>		0.0097
12	0.2603 ±	0.3825	0.1543 ±	4.61 × 10 <sup>-4</sup>	0.2264 ±	0.1573	0.2283 ±	0.0146	0.2863 ±
	0.0685		0.0763		0.1127		0.0294		0.0611
13	0.1385 ±	2.51 × 10 <sup>-9</sup>	0.1349 ±	2.35 × 10 <sup>-13</sup>	0.1299 ±	2.60 × 10 <sup>-10</sup>	0.2097 ±	0.0483	0.2608 ±
	0.0085		2.9257 × 10 <sup>-17</sup>		0.0984		0.0329		0.0251
14	0.4288 ±	1.26 × 10 <sup>-08</sup>	0.4215 ±	7.97 × 10 <sup>-09</sup>	0.4334 ±	2.39 × 10 <sup>-08</sup>	0.5295 ±	0.4376	0.5413 ±
	1.1703 × 10 <sup>-16</sup>		0.0095		5.8514 × 10 <sup>-17</sup>		0.0301		0.0364



表 7 5 种聚类方法的 RI 值比较结果  
Table 7 Comparison of RI performance of five clustering methods

编号	Co-Clustering		CombKM		FCM		Co-FKM		LR-MVEWFCM
	均值	P-value	均值	P-value	均值	P-value	均值	P-value	均值
8	0.8392 ± 0.0010	1.3475 × 10 <sup>-14</sup>	0.8112 ± 0.0369	1.95 × 10 <sup>-7</sup>	0.8390 ± 0.0115	0.0032	0.8571 ± 0.0019	0.0048	0.8508 ± 0.0013
9	0.8797 ± 0.0014	1.72 × 10 <sup>-26</sup>	0.8481 ± 0.0667	2.56 × 10 <sup>-5</sup>	0.8859 ± 1.1703 × 10 <sup>-16</sup>	6.49 × 10 <sup>-26</sup>	0.9358 ± 0.0037	3.29 × 10 <sup>-14</sup>	0.9665 ± 0.0026
10	0.6515 ± 0.0231	3.13 × 10 <sup>-1</sup>	0.6059 ± 0.0340	1.37 × 10 <sup>-6</sup>	0.6186 ± 0.0624	0.0016	0.6772 ± 0.0227	0.0761	0.6958 ± 0.0215
11	0.8797 ± 0.0014	1.25 × 10 <sup>-18</sup>	0.8755 ± 0.0029	5.99 × 10 <sup>-12</sup>	0.8859 ± 0.0243	2.33 × 10 <sup>-18</sup>	0.9267 ± 2.3406 × 10 <sup>-16</sup>	5.19 × 10 <sup>-18</sup>	0.9527 ± 0.0041
12	0.6511 ± 0.0279	0.0156	0.6024 ± 0.0322	2.24 × 10 <sup>-5</sup>	0.6509 ± 0.0652	0.1139	0.6511 ± 0.0189	0.008	0.6902 ± 0.0370
13	0.5877 ± 0.0030	1.35 × 10 <sup>-12</sup>	0.5888 ± 0.0292	2.10 × 10 <sup>-14</sup>	0.5818 ± 1.1703 × 10 <sup>-16</sup>	4.6351 × 10 <sup>-13</sup>	0.6508 ± 0.0147	0.0358	0.6855 ± 0.0115
14	0.7187 ± 1.1703 × 10 <sup>-16</sup>	3.82 × 10 <sup>-6</sup>	0.7056 ± 0.0168	1.69 × 10 <sup>-6</sup>	0.7099 ± 1.1703 × 10 <sup>-16</sup>	8.45 × 10 <sup>-7</sup>	0.7850 ± 0.0162	0.5905	0.7917 ± 0.0353

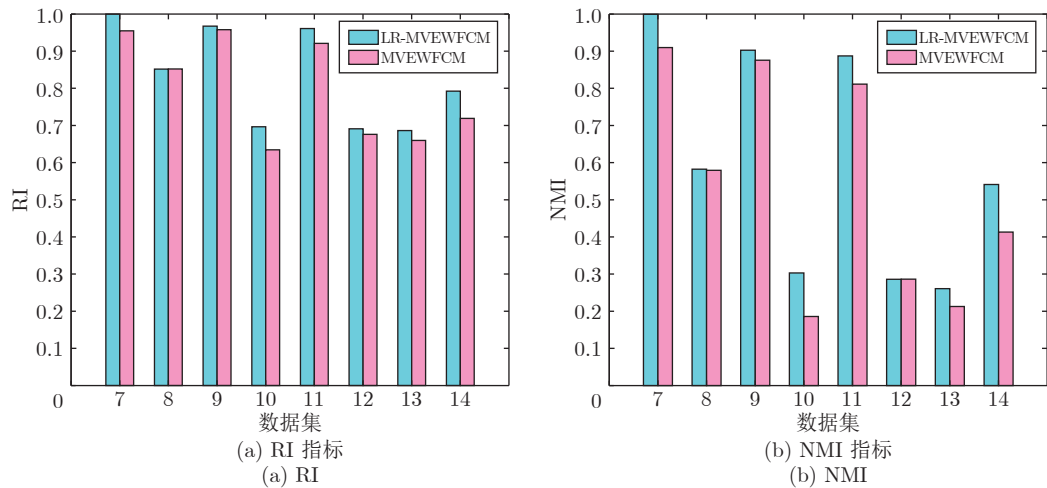
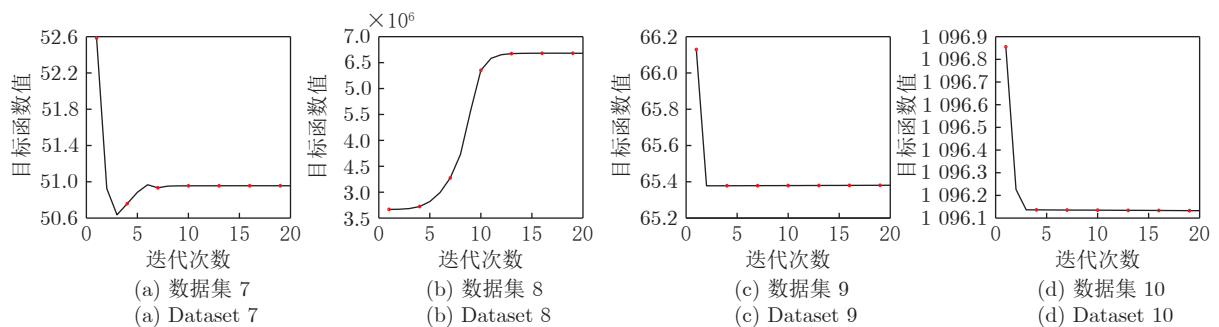


图 4 低秩约束对算法性能的影响 (横坐标为数据集编号, 纵坐标为聚类性能指标)

Fig.4 The influence of low rank constraints on the performance of the algorithm (the X-coordinate is the data set number and the Y-coordinate is the clustering performance index)



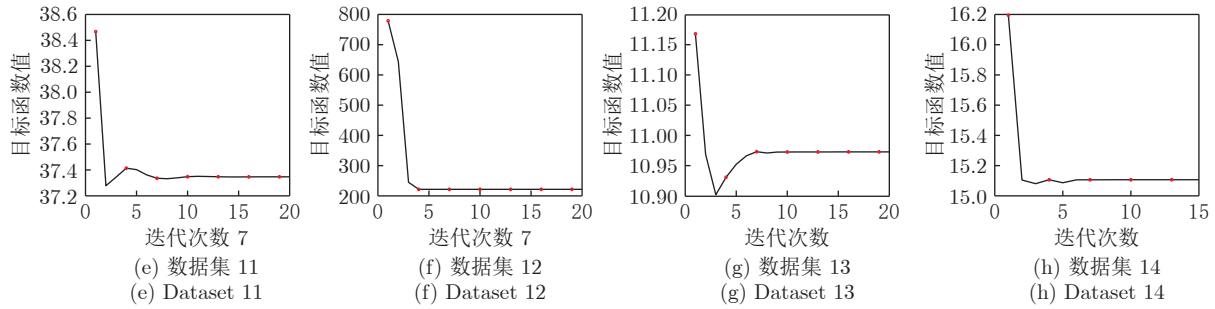


图 5 LR-MVEWFCM 算法的收敛曲线

Fig. 5 Convergence curve of LR-MVEWFCM algorithm

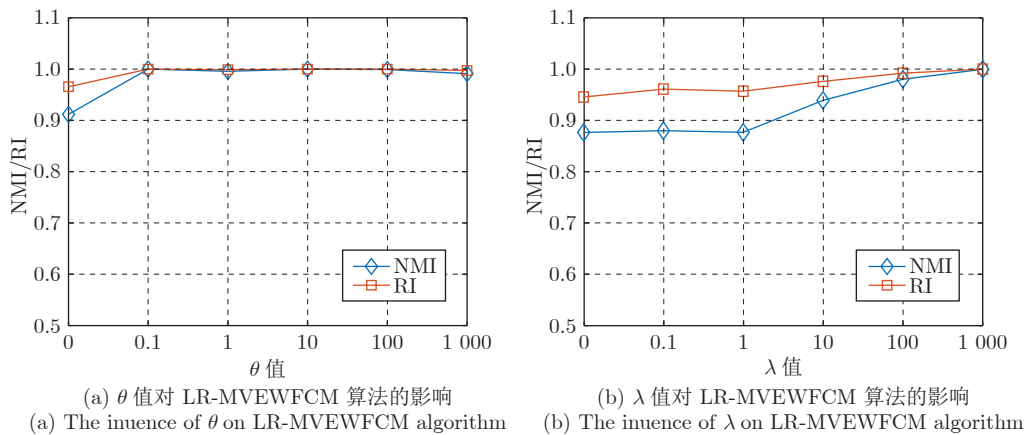
(a)  $\theta$  值对 LR-MVEWFCM 算法的影响  
(a) The influence of  $\theta$  on LR-MVEWFCM algorithm(b)  $\lambda$  值对 LR-MVEWFCM 算法的影响  
(b) The influence of  $\lambda$  on LR-MVEWFCM algorithm

图 6 模拟数据集 7 上参数敏感性分析

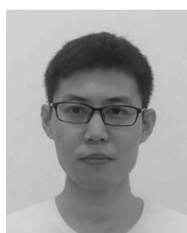
Fig. 6 Sensitivity analysis of parameters on simulated dataset 7

## References

- Xu C, Tao D C, Xu C. Multi-view learning with incomplete views. *IEEE Transactions on Image Processing*, 2015, **24**(12): 5812–5825
- Brefeld U. Multi-view learning with dependent views. In: Proceedings of the 30th Annual ACM Symposium on Applied Computing, Salamanca, Spain: ACM, 2015. 865–870
- Muslea I, Minton S, Knoblock C A. Active learning with multiple views. *Journal of Artificial Intelligence Research*, 2006, **27**(1): 203–233
- Zhang C Q, Adeli E, Wu Z W, Li G, Lin W L, Shen D G. Infant brain development prediction with latent partial multi-view representation learning. *IEEE Transactions on Medical Imaging*, 2018, **38**(4): 909–918
- Bickel S, Scheffer T. Multi-view clustering. In: Proceedings of the 4th IEEE International Conference on Data Mining (ICDM'04), Brighton, UK: IEEE, 2004. 19–26
- Wang Y T, Chen L H. Multi-view fuzzy clustering with minimax optimization for effective clustering of data from multiple sources. *Expert Systems with Applications*, 2017, **72**: 457–466
- Wang Jun, Wang Shi-Tong, Deng Zhao-Hong. Survey on challenges in clustering analysis research. *Control and Decision*, 2012, **27**(3): 321–328  
(王骏, 王士同, 邓赵红. 聚类分析研究中的若干问题. *控制与决策*, 2012, **27**(3): 321–328)
- Pedrycz W. Collaborative fuzzy clustering. *Pattern Recognition Letters*, 2002, **23**(14): 1675–1686
- Cleuziou G, Exbrayat M, Martin L, Sublemontier J H. CoFKM: A centralized method for multiple-view clustering. In: Proceedings of the 9th IEEE International Conference on Data Mining, Miami, FL, USA: IEEE, 2009. 752–757
- Jiang Y Z, Chung F L, Wang S T, Deng Z H, Wang J, Qian P J. Collaborative fuzzy clustering from multiple weighted views. *IEEE Transactions on Cybernetics*, 2015, **45**(4): 688–701
- Bettoumi S, Jlassi C, Arous N. Collaborative multi-view K-means clustering. *Soft Computing*, 2019, **23**(3): 937–945
- Zhang G Y, Wang C D, Huang D, Zheng W S, Zhou Y R. TW-Co-K-means: Two-level weighted collaborative K-means for multi-view clustering. *Knowledge-Based Systems*, 2018, **150**: 127–138
- Cao X C, Zhang C Q, Fu H Z, Liu S, Zhang H. Diversity-induced multi-view subspace clustering. In: Proceedings of the 2015 IEEE Conference on Computer Vision and Pattern Recognition, Boston, MA, USA: IEEE, 2015. 586–594
- Zhang C Q, Fu H Z, Liu S, Liu G C, Cao X C. Low-rank tensor constrained multiview subspace clustering. In: Proceedings of the 2015 IEEE International Conference on Computer Vision, Santiago, Chile: IEEE, 2015. 1582–1590
- Boyd S, Parikh N, Chu E, Peleato B, Eckstein J. Distributed optimization and statistical learning via the alternating direction method of multipliers. *Foundations and Trends in Machine Learning*, 2011, **3**(1): 1–122
- Liu G C, Lin Z C, Yan S C, Sun J, Yu Y, Ma Y. Robust recovery of subspace structures by low-rank representation. *IEEE*

*Transactions on Pattern Analysis and Machine Intelligence*, 2013, **35**(1): 171–184

- 17 Bezdek J C, Ehrlich R, Full W. FCM: The fuzzy C-means clustering algorithm. *Computers Geosciences*, 1984, **10**(2): 191–203
- 18 Abavisani M, Patel V M. Multimodal sparse and low-rank subspace clustering. *Information Fusion*, 2018, **39**: 168–177
- 19 Gu Q Q, Zhou J. Learning the shared subspace for multi-task clustering and transductive transfer classification. In: Proceedings of the 9th IEEE International Conference on Data Mining, Miami beach, FL, USA: IEEE, 2009. 159–168
- 20 Gu Q Q, Zhou J. Co-clustering on manifolds. In: Proceedings of the 15th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Paris, France: ACM, 2009. 359–368



**张嘉旭** 江南大学数字媒体学院硕士研究生. 主要研究方向为人工智能和模式识别.

E-mail: zhangjiayu@hl.chinamobile.com

(**ZHANG Jia-Xu** Master student at the School of Digital Media, Jiangnan University. His research interest covers artificial intelligence and data mining.)



**王 骏** 上海大学通信与信息工程学院副教授. 主要研究方向为人工智能, 模糊聚类 and 医学图像分类. 本文通信作者.

E-mail: wangjun\_syty@hotmail.com

(**WANG Jun** Associate professor at the School of Communication and Information Engineering, Shanghai University. His research interest covers artificial intelligence, fuzzy clustering, and medical image classification. Corresponding author of this paper.)



**张春香** 江南大学数字媒体学院硕士研究生. 主要研究方向为人工智能和模式识别.

E-mail: 17851308360@163.com

(**ZHANG Chun-Xiang** Master student at the School of Digital Media, Jiangnan University. Her research interest covers artificial intelligence and data mining.)



**林得富** 江南大学数字媒体学院硕士研究生. 主要研究方向为人工智能和模式识别.

E-mail: jiangnandaxu\_2022@yeah.net

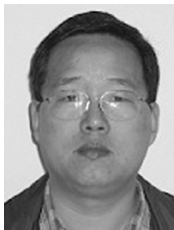
(**LIN De-Fu** Master student at the School of Digital Media, Jiangnan University. His research interest covers artificial intelligence and data mining.)



**周 塔** 江苏科技大学电子信息学院副教授. 主要研究方向为人工智能, 模式识别与智能系统.

E-mail: jkdzhout@just.edu.cn

(**ZHOU Ta** Associate professor at the School of Electronic Information, Jiangsu University of Science and Technology. His research interest covers artificial intelligence, pattern recognition, and intelligent systems.)



**王士同** 江南大学数字媒体学院教授. 主要研究方向为人工智能和模式识别. E-mail: wxwangst@aliyun.com

(**WANG Shi-Tong** Professor at the School of Digital Media, Jiangnan University. His research interest covers artificial intelligence and

data mining.)