

基于超图正则化的概念分解及在数据表示中的应用

李 雪, 赵春霞, 舒振球, 郭剑辉

(南京理工大学 计算机科学与工程学院, 南京 210094)

摘要: 针对传统图模型的流形学习无法准确表达数据间多元几何结构信息的问题, 提出一种基于超图正则化的概念分解(HRCF)算法. 该算法用一组具有相似属性的数据子集构建超边, 建立数据间高阶关系的超图模型. 通过在概念分解算法中增加超图正则项, 保持数据间多元几何流形结构, 提高了算法的鉴别性. 在 Yale 库、USPS 库和 TDT2 库上的实验表明, HRCF 算法明显提高了聚类的准确率和归一化互信息, 验证了算法的有效性.

关键词: 概念分解; 流形正则项; 非负矩阵分解; 聚类

中图分类号: TP391

文献标志码: A

Hyper-graph regularized concept factorization algorithm and its application to data representation

LI Xue, ZHAO Chun-xia, SHU Zhen-qiu, GUO Jian-hui

(College of Computer Science and Engineering, Nanjing University of Science and Technology, Nanjing 210094, China. Correspondent: LI Xue, E-mail: lixue_angel@163.com)

Abstract: The manifold learning methods of the simple graph model ignored the high-order relationship between data points. Therefore, an algorithm, called hyper-graph regularized concept factorization(HRCF) is proposed. HRCF considers the high-order relationship of samples by constructing the hyper-edge in hyper-graph with a subset of data points sharing with some attribute. The concept factorization(CF) algorithm can preserve the high-order relationship of the manifold structure, by adding hyper-graph regularization term in clustering. Thus, the algorithm has more discrimination power. The experimental results on Yale, USPS and TDT2 database show that the proposed approach provides a better representation and achieves better clustering results in terms of accuracy and normalized mutual information, and verify the effectiveness of the proposed method.

Keywords: concept factorization; manifold regularization; non-negative matrix factorization; cluster

0 引言

目前, 数据表示已成为信息检索、机器学习、数据挖掘中的热点研究之一^[1]. 在实际应用中, 数据通常是成千上万维的, 传统方法在处理时变得难以实现, 因此, 对高维数据进行低维表示成为数据处理的前提条件. 用于数据表示的矩阵分解算法包括奇异值分解(SVD)^[2]、主成分分析(PCA)、线性鉴别分析(LDA)、非负矩阵分解(NMF)^[3]和概念分解(CF)^[4]等.

Lee 等^[3]提出的 NMF 算法基于部分的数据表示, 在不改变原始高维数据结构的前提下, 将一个高维矩阵分解成两个非负的低秩矩阵的乘积, 由于非负性

的约束, 使得数据包含负值时 NMF 算法受到限制. 鉴于 NMF 算法无法进行核化, Xu 等^[4]提出了 CF 算法, CF 模型的思想是每个聚类可用数据点的线性组合来表示, 而每个数据点又可以用聚类中心的线性组合表示. 近年来, Liu 等^[5]提出了基于局部受限的概念分解(LC-CF)算法, 该算法根据局部坐标系编码思想对传统的 CF 增加一个局部正则化限制. Cai 等^[6]提出了一种局部一致性概念分解(LCCF)算法, 该算法通过构造一个传统图模型, 使其在低维表示空间中保持数据原有的流形结构信息, 但是只考虑两个数据的成对关系, 无法准确表达数据间的高阶关系.

收稿日期: 2014-05-20; 修回日期: 2014-10-14.

基金项目: 国家自然科学基金项目(61272220, 61101197, 90820306); 中国博士后科学基金项目(2014M551599); 江苏省社会安全图像与视频理解重点实验室基金项目(30920130122006); 江苏省普通高校研究生科研创新计划项目(KYLX_0383).

作者简介: 李雪(1989—), 女, 博士生, 从事图像处理、模式识别的研究; 赵春霞(1964—), 女, 教授, 博士生导师, 从事机器人、人工智能、图像处理与模式识别等研究.

为解决传统图模型无法准确表达数据间多元几何结构信息的缺陷,本文提出一种基于超图正则化的概念分解(HRCF)算法.该算法基于无向有权重的拉普拉斯图,使用 K 近邻(KNN)方法构建超边,超边上的任意一点与该超边上的所有点都有交互作用,从而可以有效提取数据间的多元几何结构信息,兼顾数据间几何结构信息的整体性,有效刻画了数据间的高阶关系^[7-12].HRCF算法采用乘性迭代方法来求解目标函数,并证明了算法的收敛性,实验结果表明了算法的有效性和准确性.

1 非负矩阵分解和概念分解算法简介

1.1 非负矩阵分解(NMF)

给定一个非负矩阵

$$\mathbf{X} = [x_1, x_2, \dots, x_n] \in \mathbf{R}^{m \times n},$$

其中每一列表示一个样本.寻求两个非负矩阵 $\mathbf{U} \in \mathbf{R}^{m \times k}$ 和 $\mathbf{V} \in \mathbf{R}^{n \times k}$ 使其满足

$$\mathbf{X}_{m \times n} \approx \mathbf{U}_{m \times k} \mathbf{V}_{n \times k}^T,$$

其中 $k \ll m, k \ll n$.NMF的目标函数表示为

$$O = \|\mathbf{X} - \mathbf{UV}^T\|^2, \quad (1)$$

其中 $\|\cdot\|$ 为Frobenius范数.目标函数 O 同时对于 \mathbf{U} 和 \mathbf{V} 是非凸函数,因此不能找到全局最优解,但是对于单独的 \mathbf{U} 或 \mathbf{V} 是凸函数.

1.2 概念分解(CF)

在CF算法中,要求每个基 u_k 是样本向量 x_i 的非负线性组合,即

$$u_k = \sum_{j=1}^N x_j w_{jk},$$

其中 $w_{jk} > 0$.有 $\mathbf{W} = [w_{jk}] \in \mathbf{R}^{N \times K}$,CF模型是寻求矩阵 \mathbf{W} 和 \mathbf{V} 使得 $\mathbf{X} \approx \mathbf{XWV}^T$,其目标函数为

$$O = \|\mathbf{X} - \mathbf{XWV}^T\|^2. \quad (2)$$

采用乘性迭代^[5]更新算法分别更新 \mathbf{W} 和 \mathbf{V} ,即

$$\begin{aligned} v_{jk}^{t+1} &\leftarrow v_{jk}^t \frac{(\mathbf{KW})_{jk}}{(\mathbf{VW}^T \mathbf{KW})_{jk}}, \\ w_{jk}^{t+1} &\leftarrow w_{jk}^t \frac{(\mathbf{KW})_{jk}}{(\mathbf{KWV}^T \mathbf{V})_{jk}}, \end{aligned} \quad (3)$$

其中 $\mathbf{K} = \mathbf{X}\mathbf{X}^T$.

2 基于超图正则化的概念分解算法

谱图理论和流形学习理论^[13]的研究表明,数据间的几何结构关系可通过数据点的最近邻图近似表示.

2.1 构建超图正则项

超图由 $G_h = (\mathbf{V}, \mathbf{E}, \mathbf{W})$ 组成, \mathbf{V} 为节点集合, \mathbf{E} 为超边集合, \mathbf{W} 为对角矩阵,描述了超边的权重.数据映射到低维空间后,通常定义 \mathfrak{R} 来度量近邻点在低维空间的平滑度,有

$$\begin{aligned} \mathfrak{R} &= \frac{1}{2} \sum_{e \in E} \sum_{(i,j) \in e} \|z_i - z_j\|^2 \frac{w(e)}{\delta(e)} = \\ &= \text{Tr}(\mathbf{V}^T \mathbf{D}_v \mathbf{V}) - \text{Tr}(\mathbf{V}^T \mathbf{S} \mathbf{V}) = \\ &= \text{Tr}(\mathbf{V}^T \mathbf{L}_{\text{hyper}} \mathbf{V}). \end{aligned} \quad (4)$$

其中: \mathbf{D}_v 为对角矩阵,表示顶点的度; $\mathbf{L}_{\text{hyper}} = \mathbf{D}_v - \mathbf{S}$; $\mathbf{S} = \mathbf{H}\mathbf{W}\mathbf{D}_e^{-1}\mathbf{H}^T$.

为了尽可能使数据集在新的表示空间中保持光滑,需要最小化 \mathfrak{R} .

2.2 构建HRCF算法的目标函数

为了保持样本的多元几何结构信息,在CF算法的目标函数中添加基于超图的拉普拉斯正则项约束,HRCF算法的目标函数可以表示为

$$O = \|\mathbf{X} - \mathbf{XWV}^T\|^2 + \alpha \text{Tr}(\mathbf{V}^T \mathbf{L}_{\text{hyper}} \mathbf{V}). \quad (5)$$

其中: $\text{Tr}(\cdot)$ 为矩阵的迹, \mathbf{W} 和 \mathbf{V} 为非负矩阵,正则项参数 α 满足 $\alpha \geq 0$.

2.3 HRCF目标函数求解

HRCF的目标函数 O 同时对于 \mathbf{W} 和 \mathbf{V} 是非凸函数,因此无法得到目标函数的全局最优解,但是对于单独的 \mathbf{W} 和 \mathbf{V} 是凸函数,可以采用乘性迭代算法求解目标函数的局部最优解.目标函数(5)可化简为

$$\begin{aligned} O &= \\ &= \text{Tr}((\mathbf{X} - \mathbf{XWV}^T)^T (\mathbf{X} - \mathbf{XWV}^T)) + \\ &= \alpha \text{Tr}(\mathbf{V}^T \mathbf{L}_{\text{hyper}} \mathbf{V}) = \\ &= \text{Tr}(\mathbf{K}) - 2\text{Tr}(\mathbf{VW}^T \mathbf{K}) + \text{Tr}(\mathbf{VW}^T \mathbf{KWV}^T) + \\ &= \alpha \text{Tr}(\mathbf{V}^T \mathbf{L}_{\text{hyper}} \mathbf{V}), \end{aligned} \quad (6)$$

其中 $\mathbf{K} = \mathbf{X}^T \mathbf{X}$.

分别对 O 求 \mathbf{W} 和 \mathbf{V} 的偏导数,通过KKT条件, $\psi_{jk} w_{jk} = 0, \phi_{jk} v_{jk} = 0$,可得到HRCF算法的更新迭代规则

$$w_{jk} \leftarrow w_{jk} \frac{(\mathbf{KV})_{jk}}{(\mathbf{KWV}^T \mathbf{V})_{jk}}, \quad (7)$$

$$v_{jk} \leftarrow v_{jk} \frac{(\mathbf{KW} + \alpha \mathbf{S})_{jk}}{(\mathbf{VW}^T \mathbf{KW} + \alpha \mathbf{D}_v \mathbf{V})_{jk}}. \quad (8)$$

2.4 收敛性证明

定理1 目标函数在更新迭代规则(7)和(8)下是非增函数,当且仅当 \mathbf{W} 和 \mathbf{V} 为稳定点时,目标函数是不变的.

定理1确保了 \mathbf{W} 和 \mathbf{V} 分别在更新迭代规则(7)和(8)下目标函数(5)是收敛的,并且可以找到局部最优解.为证明定理1,使用期望最大算法^[1]增加辅助函数,定义如下.

定义1 若函数 $G(x, x')$ 满足如下条件:

$$G(x, x') \geq F(x), \quad G(x, x) = F(x),$$

则称 $G(x, x')$ 是 $F(x)$ 的辅助函数.

引理1 如果函数 G 是函数 F 的辅助函数, 则 F 在如下条件下是非增的:

$$x^{(t+1)} = \arg \min_x G(x, x'). \quad (9)$$

证明 函数 G 是函数 F 的辅助函数, 由定义1, 有

$$F(x^{(t+1)}) \leq G(x^{(t+1)}, x^{(t)}) \leq G(x^{(t)}, x^{(t)}) = F(x^{(t)}). \quad \square$$

为证明定理1, 依次证明更新迭代规则(7)和(8)是收敛的. 对于式(8), 定义 v_{ab} 是矩阵 \mathbf{V} 的元素, $F_{v_{ab}}$ 表示目标函数 O 中与变量 v_{ab} 相关的函数, 由于本质上目标函数 O 是逐个元素更新的, 首先证明 $F_{v_{ab}}$ 在迭代式(8)下是非增的.

引理2 函数

$$G(v, v_{ab}^{(t)}) = F_{v_{ab}}(v_{ab}^{(t)}) + F'_{v_{ab}}(v_{ab}^{(t)})(v - v_{ab}^{(t)}) + \frac{(\mathbf{V}\mathbf{W}^T\mathbf{K}\mathbf{W})_{ab} + \alpha(\mathbf{D}_v\mathbf{V})_{ab}}{v_{ab}^{(t)}}(v - v_{ab}^{(t)})^2 \quad (10)$$

是 $F_{v_{ab}}$ 的辅助函数.

证明 容易证明 $G(v, v) = F_{v_{ab}}(v)$, 下面通过比较 $G(v, v_{ab}^{(t)})$ 和 $F_{v_{ab}}(v)$ 泰勒级数展开式的大小来证明 $G(v, v_{ab}^{(t)}) \geq F_{v_{ab}}(v)$, 即证明引理2. $F_{v_{ab}}(v)$ 的泰勒级数展开式为

$$F_{v_{ab}}(v) = F_{v_{ab}}(v_{ab}^{(t)}) + F'_{v_{ab}}(v_{ab}^{(t)})(v - v_{ab}^{(t)}) + \frac{1}{2}[(\mathbf{W}^T\mathbf{K}\mathbf{W})_{ab} + \alpha\mathbf{L}_{\text{hyper}}](v - v_{ab}^{(t)})^2. \quad (11)$$

由式(10)可知, 只需证明

$$(\mathbf{V}\mathbf{W}^T\mathbf{K}\mathbf{W})_{ab} = \sum_{l=1}^k v_{al}^{(t)}(\mathbf{W}^T\mathbf{K}\mathbf{W})_{lb} \geq v_{ab}^{(t)}(\mathbf{W}^T\mathbf{K}\mathbf{W})_{bb}, \quad (12)$$

$$\alpha(\mathbf{D}_v, v)_{ab} \geq \alpha(\mathbf{D}_v - \mathbf{S})_{aa}v_{ab}^{(t)} = \alpha(\mathbf{L}_{\text{hyper}})_{aa}v_{ab}^{(t)}. \quad (13)$$

综上所述, $G(v, v_{ab}^{(t)}) \geq F_{v_{ab}}(v)$. \square

同理, 对于式(7), 定义 w_{ab} 是矩阵 \mathbf{W} 的元素, $F_{w_{ab}}$ 表示目标函数 O 中与变量 w_{ab} 相关的函数.

引理3 函数

$$G(w, w^{(t)}) = F_{w_{ab}}(w_{ab}^{(t)}) + F'_{w_{ab}}(w_{ab}^{(t)})(w - w_{ab}^{(t)}) + \frac{\mathbf{K}\mathbf{W}\mathbf{V}^T\mathbf{V}}{w_{ab}^{(t)}}(w - w_{ab}^{(t)})^2 \quad (14)$$

是 $F_{w_{ab}}$ 的辅助函数.

引理3的证明过程同引理2, 此处略.

下面证明定理1.

证明 由引理2可知, $G(v, v_{ab}^{(t)})$ 是 $F_{v_{ab}}$ 的辅助函数, 将式(10)代入(9)可得

$$v_{ab}^{(t+1)} = v_{ab}^{(t)} - v_{ab}^{(t)} \frac{F'_{v_{ab}}(v_{ab}^{(t)})}{2(\mathbf{V}\mathbf{W}^T\mathbf{K}\mathbf{W})_{ab} + 2\alpha(\mathbf{D}_v\mathbf{V})_{ab}} = v_{ab}^{(t)} \frac{(\mathbf{K}\mathbf{W} + \alpha\mathbf{S}\mathbf{V})_{ab}}{(\mathbf{V}\mathbf{W}^T\mathbf{K} + \alpha\mathbf{D}_v\mathbf{V})_{ab}}. \quad (15)$$

由于式(10)是 $F_{v_{ab}}$ 的辅助函数, 由引理1可得, $F_{v_{ab}}$ 在更新过程中是非增的, $F_{w_{ab}}$ 在更新过程中是非增的, 其证明参见 $F_{v_{ab}}$ 的证明过程. \square

HRCF 目标函数(5)在式(7)和(8)的更新迭代规则下是非增的, 最后趋向于一个局部最优值.

3 实验分析

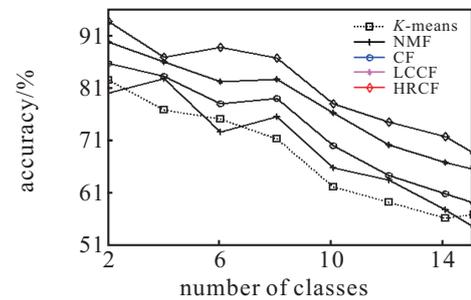
本节重点评估 HRCF 算法在 Yale 库、USPS 库、TDT2 库的实验效果, 并与 K 均值 (K -means) 算法、NMF 算法、CF 算法、LCCF 算法在相同数据库上的结果进行比较. 每次随机选择 p 类样本进行聚类实验, 重复实验 20 次.

3.1 数据集介绍

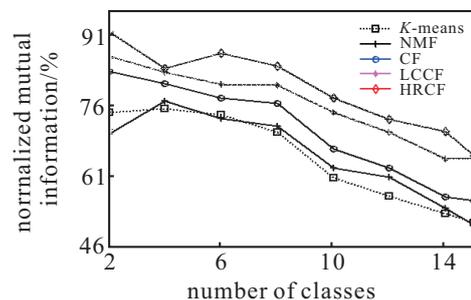
在人脸库 Yale、手写体库 USPS 和文本库 TDT2 上进行实验. 在 Yale 和 USPS 库上进行实验时, 将图像大小归一化为 32×32 ; 在 TDT2 库上进行实验时, 只选取样本数目大于 10 的样本, 共选出用于实验的包括 56 类组成的 10021 个样本.

3.2 实验结果

在聚类中常利用准确率 (AC) 和归一化互信息 (NMI) 作为评价标准. 图 1~图 3 为 5 种算法在 Yale

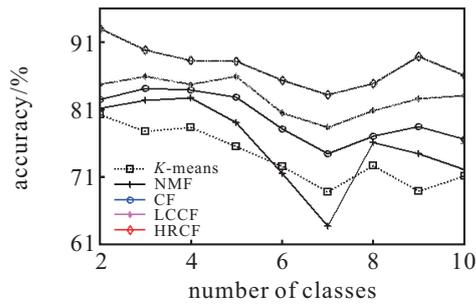


(a) 准确率

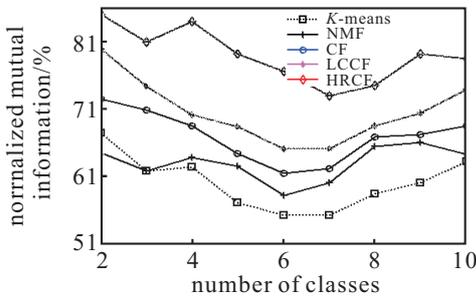


(b) 归一化互信息

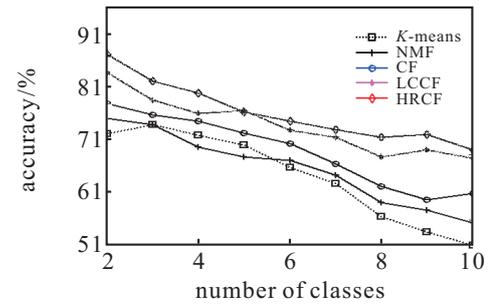
图1 在 Yale 数据库上聚类的准确率和归一化互信息



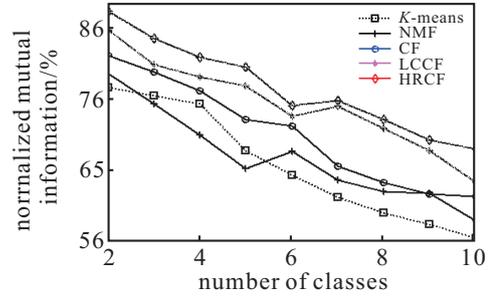
(a) 准确率



(b) 归一化互信息



(a) 准确率



(b) 归一化互信息

图2 在USPS数据库上聚类的准确率和归一化互信息

图3 在TDT2数据库上聚类的准确率和归一化互信息

表1 在Yale数据库上的聚类实验

P	accuracy					normalized mutual information				
	K-means	NMF	CF	LCCF	HRCF	K-means	NMF	CF	LCCF	HRCF
2	82.46	80.01	85.76	89.79	93.72	74.56	70.04	83.26	86.38	91.57
4	76.79	82.78	83.27	86.06	86.92	75.29	76.88	80.58	83.01	83.81
6	75.12	72.56	78.01	82.21	88.79	74.06	73.22	77.54	80.45	87.06
8	71.22	75.57	79.06	82.69	86.72	70.33	71.68	76.47	80.31	84.26
10	62.09	65.71	69.99	76.26	78.06	60.74	62.82	66.74	74.51	77.49
12	59.21	63.42	64.32	70.06	74.49	56.81	60.78	62.79	70.31	73.06
14	56.06	56.06	57.71	60.78	71.78	53.08	54.29	56.62	64.71	70.55
15	56.78	56.78	54.33	58.98	68.56	51.12	50.67	55.78	64.77	65.28
avg	67.47	69.12	71.56	75.85	81.13	64.50	65.05	69.97	75.56	79.14

表2 在USPS数据库上的聚类实验

P	accuracy					normalized mutual information				
	K-means	NMF	CF	LCCF	HRCF	K-means	NMF	CF	LCCF	HRCF
2	80.14	81.06	82.45	84.62	92.75	67.38	64.42	72.34	79.74	84.88
3	77.74	82.31	84.06	85.74	89.67	61.77	61.79	70.78	74.26	80.79
4	78.28	82.68	83.79	84.66	88.14	62.38	63.78	68.42	70.06	83.81
5	75.46	79.01	82.74	85.76	87.96	57.09	62.40	64.32	68.32	79.06
6	72.58	71.47	78.05	80.42	85.15	55.21	58.12	61.39	65.07	76.45
7	68.82	63.79	74.46	78.32	83.06	55.18	59.96	62.07	64.98	72.82
8	72.71	76.07	77.07	80.79	84.76	58.31	65.34	66.76	68.42	74.42
9	68.89	74.42	78.38	82.46	88.71	60.02	65.97	67.12	70.33	79.06
10	71.19	72.06	76.54	82.99	85.89	63.11	64.26	68.44	73.79	78.42
avg	73.98	75.87	79.73	82.86	87.34	60.05	62.90	66.85	70.55	78.86

库、USPS库、TDT2库上聚类的准确率和归一化互信息。

表1~表3为Yale、USPS、TDT2数据库上的聚类实验。由表1可见，HRCF比CF算法的平均AC和

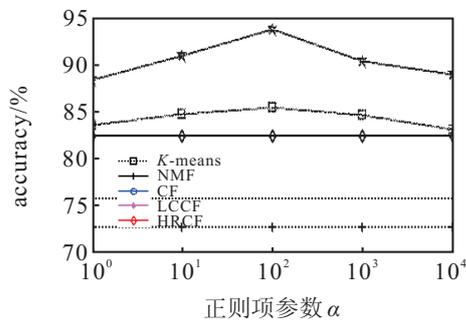
平均NMI分别提高了9.57%和9.17%；由表2可见，HRCF在USPS库上平均AC和NMI高达87.34%和78.86%；由表3可见，HRCF比LCCF的平均AC和NMI分别提高了2.90%和2.44%。

表 3 在 TDT2 数据库上的聚类实验

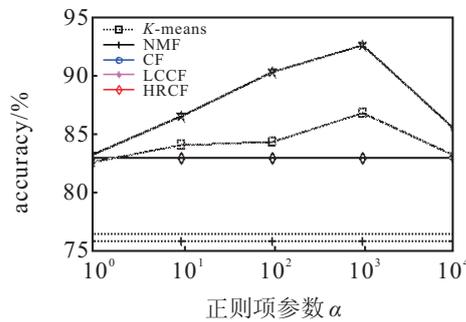
P	accuracy					normalized mutual information				
	K -means	NMF	CF	LCCF	HRCF	K -means	NMF	CF	LCCF	HRCF
2	75.87	79.26	82.79	89.78	93.96	77.88	79.92	82.42	86.12	88.78
3	77.84	77.84	80.06	83.48	87.78	76.71	75.54	80.15	81.12	84.96
4	75.48	72.76	78.74	80.54	85.11	75.59	71.13	77.42	79.38	82.24
5	73.26	70.44	76.00	81.06	80.74	68.92	66.32	73.28	78.13	80.82
6	68.11	69.62	73.49	76.62	78.64	65.42	68.78	72.45	73.82	75.36
7	64.32	66.35	68.81	75.07	76.69	77.88	64.72	66.72	75.22	76.08
8	56.79	60.00	63.77	70.55	75.02	76.71	63.06	64.38	72.06	73.38
9	53.22	58.27	60.67	72.14	75.59	75.59	62.79	62.66	68.89	70.42
10	50.18	55.42	62.12	70.22	72.04	68.92	58.82	62.32	64.37	69.12
avg	66.12	67.77	71.83	77.72	80.62	65.42	67.90	71.31	75.47	77.91

3.3 参数选择

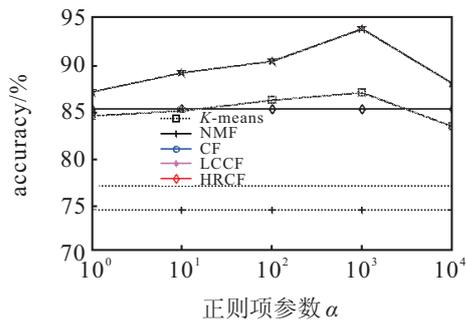
HRCF 模型中需要确定 2 个参数, 分别为创建超边时所选择的 k 个近邻节点和正则项参数 α . 通过搜索不同参数值对实验结果的影响进行评估, 由实验分析可知, k 取值为 2~10、正则项参数 α 在 10~10000 范围内, 实验效果最理想.



(a) 在 Yale 库上

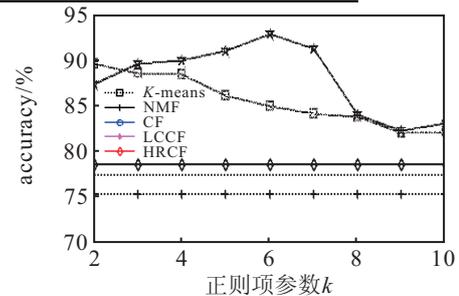


(b) 在 USPS 库上

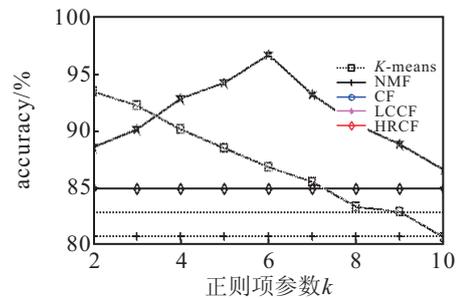


(c) 在 TDT2 库上

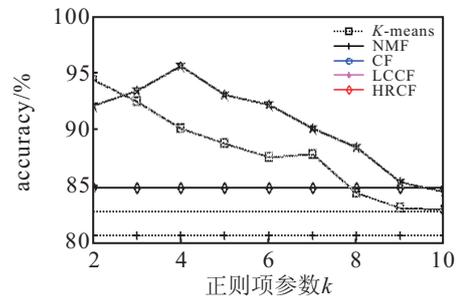
图 4 正则项参数 α 对聚类准确率的影响



(a) 在 Yale 库上



(b) 在 USPS 库上



(c) 在 TDT2 库上

图 5 构建超边的顶点数 k 对聚类准确率影响

当正则项参数 α 变化时, 5 种算法在 3 个数据集上聚类的准确率如图 4 所示. 参数 α 对 K -means、NMF、CF 算法无影响, LCCF 算法对 α 值变化不敏感. 在 Yale 库上 $\alpha = 100$ 、在 USPS 和 TDT2 上 $\alpha = 1000$, HRCF 算法取得最好效果. 图 5 为构建超边的顶点数 k 对聚类准确率影响, 由图 5 可见, 当 $k = 6$ 时, HRCF 算法在 Yale 和 USPS 上效果最好, 在 TDT2 上, 当 $k = 4$ 时效果最好.

3.4 结论分析

由第 3.2 节和第 3.3 节可得如下结论:

1) 由于 K -means 算法、NMF 算法和 CF 算法没有考虑到数据间的几何结构信息, 其聚类 AC 和 NMI 远不如保持了原始空间几何结构信息的 LCCF 算法;

2) 由于 LCCF 算法只能描述数据间成对关系, 不能有效提取分布复杂的数据间的高阶几何结构信息, 聚类效果远不如 HRCF 算法;

3) 当参数 α 较小时, 在目标函数中超图正则化所占比重较少, 即忽略了数据间几何结构的高阶多元关系, 当参数 α 较大时, 降维后重构误差变大, 无法恰当地表示原始数据;

4) 当参数 k 较大时, 不具有相同属性的节点构造同一条超边, 使得相似度降低, 实验效果随之下降, 当参数 k 较小时, 将具有相同潜在结构信息的数据离散, 故曲线先上升到最优值, 如果 k 继续增大, 便会使聚类准确率下降.

4 结 论

本文提出的 HRCF 算法通过对 CF 算法的目标函数增加拉普拉斯超图正则化项, 获得数据间固有的高阶几何结构信息, 比 CF 算法和普通图模型的 LCCF 算法具有更强的鉴别能力. 给出了算法模型的迭代求解方法、收敛性证明和参数选择分析, 并在 Yale 库、USPS 库和 TDT2 库上进行了实验, 由实验结果可知, HRCF 算法的准确率和归一化互信息明显优于 K -means、NMF、CF、LCCF 算法, 表明了 HRCF 算法的有效性. 但是, HRCF 模型中参数 k 和 α 需要通过区间搜索得到最优值, 因此如何有效选择参数是今后研究的重点方向之一.

参考文献(References)

- [1] Xu W, Liu X, Gong Y H. Document clustering based on non-negative matrix factorization[C]. Annual ACM SIGIR Conference. Toronto: Sheffield, 2003: 267-273.
- [2] Duda R O, Hart P E, Stork D G. Pattern classification[M]. The 2nd ed. Hoboken: Wiley-Interscience, 2000: 5-10.
- [3] Lee D D, Seung H S. Algorithms for non-negative matrix factorization[C]. Advances in Neural Information Processing Systems. Columbia: Vancouver, 2001: 556-562.
- [4] Xu W, Gong Y H. Document clustering by concept factorization[C]. Proc of ACM SIGIR. Sheffield, 2004: 202-209.
- [5] Liu H F, Zheng Y, Wu Z H. Locality-constrained concept factorization[C]. The Twenty-Second Int Joint Conf Artificial Intelligence. Barcelona: Morgan Kaufmann, 2011: 1378-1383.
- [6] Cai D, He X F, Han J W. Locally consistent concept factorization for document clustering[J]. IEEE Trans on Knowledge and Data Engineering, 2011, 23(6): 902-913.
- [7] Zhou D Y, Huang J Y, Bernhard S. Learning with hypergraphs: Clustering, classification and embedding[C]. Advances in Neural Information Processing Systems. Cambridge: MIT Press, 2006: 1601-1608.
- [8] Hong C Q, Yu J, Li J, et al. Multi-view hyper-graph learning by patch alignment framework[J]. IEEE Trans on Neurocomputing, 2013, 118(2013): 79-86.
- [9] Yu J, Tao D C, Wang M. Adaptive hyper-graph learning and its application in image classification[J]. IEEE Trans on Image Process, 2012, 21(7): 3262-3272.
- [10] Yu J, Rui Y, Chen B. Exploiting click constraints and multiview features for image reranking[J]. IEEE Trans on Multimed, 2014, 16(1): 159-168.
- [11] Zeng K, Yu J, Li C H, et al. Image clustering by hyper-graph regularized non-negative matrix factorization[J]. IEEE Trans on Neurocomputing, 2014, 138(22): 209-217.
- [12] Cai D, He X F, Wu X Y, et al. Non-negative matrix factorization on manifold[C]. Proc of the 8th IEEE Int Conf on Data Mining. Pisa, 2008: 63-72.
- [13] Wang C, Yu J, Tao D P. High-level attributes modeling for indoor scenes classification[J]. Neurocomputing, 2013, 12(2003): 337-343.

(责任编辑: 郑晓蕾)