基于约束图正则的块稀疏对称非负矩阵分解

刘威, 邓秀勤, 刘冬冬, 刘玉兰

计算机科学 COMPUTER SCIENCE
http://www.jsjkx.com
DOI:10.11896/jsjkx.220500050

基于约束图正则的块稀疏对称非负矩阵分解

刘 威 邓秀勤 刘冬冬 刘玉兰
广东工业大学数学与统计学院 广州 510000
(2534652861@qq.com)

摘 要 现有的基于对称非负矩阵因式分解(Symmetric Nonnegative matrix Factorization,SymNMF)算法大都仅依赖初始数据构造亲和矩阵,并且一定程度上忽视了样本有限的成对约束信息,无法有效区分不同类别的相似样本以及学习样本的几何特征。针对以上问题,提出了基于约束图正则的块稀疏对称非负矩阵分解(Block Sparse Symmetric Nonnegative Matrix Factorization Based on Constrained Graph Regularization,CGBS-SymNMF)。首先,通过先验信息构造约束图矩阵,用于指导类别指示矩阵区分高相似度的不同类别样本;然后,引入PCP-SDP(Pairwise Constraint Propagation by Semi-definite Programming)方法,利用成对约束学习一个新的样本图映射矩阵;最后,利用"勿连"约束构造不相似矩阵,用于引导一个块稀疏正则项,以增强模型抗噪能力。实验结果表明,所提算法具有更高的聚类精确度和稳定性。

关键词:对称非负矩阵因式分解;亲和矩阵;成对约束;图正则;块稀疏

中图法分类号 TP301.6

Block Sparse Symmetric Nonnegative Matrix Factorization Based on Constrained Graph Regularization

LIU Wei,DENG Xiuqin,LIU Dongdong and LIU Yulan
School of Mathematics and Statistics,Guangdong University of Technology,Guangzhou 510000,China

Abstract The existing algorithms based on symmetric nonnegative matrix factorization(SymNMF) are mostly rely on initial data to construct affinity matrices,and neglect the limited pairwise constraints,so these methods are unable to effectively distinguish similar samples of different categories or learn the geometric features of samples. To solve the above problems,this paper proposes a block sparse symmetric nonnegative matrix factorization based on constrained graph regularization(CGBS-SymNMF). Firstly,the constrained graph matrix is constructed by prior information,which is used to guide the clustering indicator matrix to distinguish different clusters of samples with high similarity. Secondly,pairwise constraint propagation by semidefinite programming (PCP-SDP) is introduced to learn a new sample graph mapping matrix by using pairwise constraints. Finally,a dissimilarity matrix is constructed by cannot-link constraints,which is used to guide a block sparse regular term for enhancing the anti-noise capability of the model. Experimental results demonstrate a higher clustering accuracy and stability of the proposed algorithm.

Keywords Symmetric nonnegative matrix factorization,Affinity matric,Pairwise constraint,Graph regularization,Block sparse

1 引言

聚类作为一种处理海量数据以及提取数据信息效益的重要方法,在机器学习和数据挖掘方面获得了广泛的应用。因此,对于如何提升聚类质量、算法效率以及稳定性等问题有着重要的研究意义。

聚类的目标是对原始数据按照一定的相似度进行类别划分,而在现实应用中所获取的数据往往是高维的并且广泛分布于流形空间上,计算样本之间的距离变得十分困难,因此需要预先对数据进行降维处理。在众多的降维方法中,非负矩阵因式分解[1](Nonnegative matrix Factorization,NMF)因其高效且简便的特点得到了广泛的应用,在文件聚类[2-3]、文本挖掘[4-5]、图像分割[6-7]和图像聚类[8-9]等方面取得了卓越的成绩,并且在此之后不断发展。文献[10]在NMF的基础上引入了图正则,提出了GNMF算法,使得重构的矩阵能够学习样本的局部几何特点;文献[11]提出了SNMF-BCS算法,

到稿日期:2022-05-06 返修日期:2022-09-12
基金项目:国家自然科学基金(12101136);广东省研究生教育创新计划项目(2021SFKC030);广州市科技基金(202102020273);广东省区域联合基金(2020A1515110967);重庆师范大学数学学科省部级重点实验室开放课题(CSSXKFKTQ202002)
This work was supported by the National Natural Science Foundation of China(12101136),Postgraduate Educational Innovation program of Guangdong(2021SFKC030),Science and Technology Foundation of Guangzhou(202102020273),Regional Join foundation of Guangdong (2020A1515110967) and Open project of Provincial Key Laboratory of Mathematics,Chongqing Normal University(CSSXKFKTQ202002).
通信作者:邓秀勤(dxq706@gdut.edu.cn)

该方法引入了矩阵的 $L_{2,1}$ 范数对分解项进行约束，提高了解的稀疏性，同时使用余弦相似度来降低数据各特征之间的多重共线性。

上述 NMF 及其相关算法变种解决了实际应用中的诸多问题，但该方法仅用于数据降维，需要借助如 K-means 等传统算法来辅助进行聚类分析，大大降低了效率，并且由于 K-means 等基础算法的缺陷导致聚类结果不稳定。为了弥补这种缺陷，文献[12]提出了 SymNMF 算法，该方法将样本亲和矩阵分解为一个非负的类别指示矩阵和它转置的乘积形式，并且最终通过指示矩阵的特定元素索引获得样本类别；文献[13]在 SymNMF 的基础上提出了 SSymNMF 算法，增加了类别指示矩阵的一范数惩罚项，提高了解的稀疏性和模型的稳定性；文献[14]提出了 GrsymNMF 算法，该算法引入图正则项来增强 SymNMF 的图学习能力，促使隶属相同类簇的样本在聚类域中彼此靠近，进而提高样本相似度。但是上述算法仅依赖于样本图，对数据图结构的学习不彻底。

在现实应用中，需要进行聚类分析的样本往往带有部分监督信息，如成对约束。成对约束分为两种：

(1)“必连”(must-link, ML)约束，隶属于“必连”约束的两个样本在进行聚类划分时必然会被划分到同一类。

(2)“勿连”(cannot-link, CL)约束，隶属于“勿连”约束的两个样本在进行聚类划分时必然不会被划分到同一类。

而传统的基于 SymNMF 算法大都忽略了这部分信息，造成了大量的信息浪费。为了利用样本成对约束信息，大量的基于 SymNMF 的半监督聚类被提出。文献[15]提出了 GSymNMF 算法，该算法在 SymNMF 基础上添加利用“必连”约束构造的约束图正则；文献[16]提出了 SymNMFCC 算法，使用成对约束促使属于“必连”约束的两个样本相似度为 1，“勿连”约束的两个样本相似度为 0；文献[17]提出了 SASymNMF 算法，该算法将学习具有监督信息的约束图和亲和矩阵两个过程相结合，使其相互增强；文献[18]提出了 PCPISymNMF 算法，该算法通过约束传播模型，用少量的成对约束学习一个高质量的亲和矩阵，增强了模型的稳定性。

尽管上述基于 SymNMF 的半监督方法取得了不错的聚类效果，但是仍然存在以下两个问题：

(1)大部分方法都着重使用“必连”约束，忽视了“勿连”约束所发挥的作用，对于隶属不同类簇的高度相似的样本区分力度不大，在聚类过程中比较容易误分类簇。

(2)初始的约束矩阵以及亲和矩阵对模型的影响程度较大，低质量的约束矩阵和亲和矩阵对模型聚类效果有较大的限制。

针对上述问题，提出了基于约束图正则的块稀疏对称非负矩阵分解(CGBS-SymNMF)，该算法综合考虑现有的半监督 SymNMF 算法的优缺点，使亲和矩阵和类别指示矩阵的学习、样本聚类 3 个过程相互结合，相互增强，以得到更加准确和稳定的聚类结果。为了提高模型对样本的区分能力以及提高模型的稳定性，着重使用了“勿连”约束生成不相似约束矩阵，从而引导类别指示矩阵的块稀疏表示；同时该算法为了增强模型对样本的区分能力，构建了“正负”约束图，使得类别指示矩阵充分学习约束图的空间几何结构；最后通过样本图学习一个更加优质的亲和矩阵，并不断通过成对约束对其元素进行修正，使得“必连”约束对应的元素趋向于 1。通过交叉迭代获得最优的类别指示矩阵并进行聚类，在 UCI 数据集和图像数据集上的实验结果表明，提出的算法优于当前经典的基于 SymNMF 的算法。

2 相关工作

2.1 NMF 与 SNMF 算法

对于一个给定的非负数据集 $\boldsymbol{X}=[x_1,x_2,x_3,\cdots,x_n]\in R^{d\times n}$，其中 $x_i\in R^{d\times 1}$ 为数据集 $\boldsymbol{X}$ 的一个样本。NMF 算法将初始数据集分解为两个非负矩阵 $\boldsymbol{U},\boldsymbol{V}$ 的内积，其模型如式(1)所示：

$$\begin{aligned}&\min_{\boldsymbol{U},\boldsymbol{H}}\|\boldsymbol{X}-\boldsymbol{U}\boldsymbol{V}^{\mathrm{T}}\|_F^2\\&\text{s.t.}\ \boldsymbol{U}\geqslant 0,\boldsymbol{V}\geqslant 0\end{aligned}\tag{1}$$

其中，$\boldsymbol{U}\in R^{d\times k}$ 为基础向量矩阵，$\boldsymbol{V}\in R^{n\times k}$ 为数据集 $\boldsymbol{X}$ 的低维表示矩阵，然后对矩阵 $\boldsymbol{V}$ 应用 K-means 等传统聚类算法以获得最终结果，因此 NMF 可用作数据的降维。根据文献[19]，NMF 算法可按照式(2)、式(3)交叉迭代直至收敛。

$$\boldsymbol{V}_{ij}\leftarrow\boldsymbol{V}_{ij}\frac{(\boldsymbol{X}^{\mathrm{T}}\boldsymbol{U})_{ij}}{(\boldsymbol{V}\boldsymbol{U}^{\mathrm{T}}\boldsymbol{U})_{ij}}\tag{2}$$

$$\boldsymbol{U}_{ij}\leftarrow\boldsymbol{U}_{ij}\frac{(XV)_{ij}}{(\boldsymbol{U}\boldsymbol{V}^{\mathrm{T}}V)_{ij}}\tag{3}$$

作为 NMF 的一个变体，SymNMF 的目标是将一个根据初始数据集构造的亲和矩阵 $\boldsymbol{S}$ 分解为一个类别指示矩阵 $\boldsymbol{H}$ 与其转置的乘积，其模型如式(4)所示：

$$\begin{aligned}&\min_{\boldsymbol{H}}\|\boldsymbol{S}-\boldsymbol{H}\boldsymbol{H}^{\mathrm{T}}\|_F^2\\&\text{s.t.}\ \boldsymbol{H}\geqslant 0\end{aligned}\tag{4}$$

其中，$\boldsymbol{H}\in R^{n\times k}$，$\boldsymbol{S}\in R^{n\times n}$，矩阵 $\boldsymbol{S}$ 通常用以下 3 种方法给出。

(1)$\boldsymbol{S}=\boldsymbol{W}$，详见文献[20]，其中 $\boldsymbol{W}$ 为根据数据集 $\boldsymbol{X}$ 构造的基于 P-NN 的对称相似矩阵，其中 $\boldsymbol{W}_{ij}=\boldsymbol{W}_{ji}$，$\boldsymbol{W}_{ij}$ 可用式(5)进行构造。

$$\boldsymbol{W}_{ij}=\begin{cases}\mathrm{e}^{\frac{-\|x_i-x_j\|^2}{2\sigma^2}}, & x_i\in\boldsymbol{N}(x_j)\ \text{or}\ x_j\in\boldsymbol{N}(x_i)\\0, & \text{otherwise}\end{cases}\tag{5}$$

其中，$\boldsymbol{N}(x)$ 表示样本 x 的 $\boldsymbol{P}$ 邻域。

(2)$\boldsymbol{S}=\boldsymbol{D}^{-\frac{1}{2}}\boldsymbol{W}\boldsymbol{D}^{-\frac{1}{2}}$，详见文献[21]，其中矩阵 $\mathbf{D}$ 为对角矩阵，对角元素 $\mathbf{D}_{ii}=\sum_{j=1}^{n}\boldsymbol{W}_{ij}$。

(3)$\boldsymbol{S}=\boldsymbol{K}$，详见文献[22]，此时，矩阵 $\boldsymbol{K}$ 与 $\boldsymbol{H}$ 通过逐步交叉迭代，互相增强，共同优化。

由于 SymNMF 的非负性以及对称性，$\boldsymbol{H}$ 能很好地捕获样本的类簇信息，因此，对于初始数据集的每个样本，其类别可以取 $\boldsymbol{H}$ 矩阵对应行最大元素的列坐标值，根据文献[23]，SymNMF 算法可由式(6)迭代收敛。

$$\boldsymbol{H}_{ij}\leftarrow\frac{1}{2}\left[\boldsymbol{H}_{ij}\left(1+\frac{(\boldsymbol{S}\boldsymbol{H})_{ij}}{(\boldsymbol{H}\boldsymbol{H}^{\mathrm{T}}\boldsymbol{H})_{ij}}\right)\right]\tag{6}$$

SymNMF 算法比 NMF 算法能够更好地适应各种数据的分解，但是同时也忽略了样本中少量的监督信息，造成了信息浪费。

2.2 PCP-SDP 算法

为了使用样本中有限的成对约束揭示更多未标记数据的标签信息，诸多成对约束传播方法[24-26](Pairwise Constraint

Propagation，PCP）被提出，本文主要介绍 PCP-SDP 算法[27]。

根据文献[28]，成对约束具有如下性质：

（1）对称性

$$(x_i,x_j)\in \boldsymbol{ML} \Rightarrow (x_j,x_i)\in \boldsymbol{ML} \\ (x_i,x_j)\in \boldsymbol{CL} \Rightarrow (x_j,x_i)\in \boldsymbol{CL} \tag{7}$$

（2）传递性

$$(x_i,x_j)\in \boldsymbol{ML},(x_j,x_l)\in \boldsymbol{ML} \Rightarrow (x_i,x_l)\in \boldsymbol{ML} \\ (x_i,x_j)\in \boldsymbol{ML},(x_j,x_l)\in \boldsymbol{CL} \Rightarrow (x_i,x_l)\in \boldsymbol{CL} \tag{8}$$

PCP-SDP 算法的主要目标是在样本图上学习一个光滑的映射，使得被 $\boldsymbol{ML}$ 约束的两个样本尽可能接近，而被 $\boldsymbol{CL}$ 约束的两个样本则尽可能远离，设 $\boldsymbol{\Phi}$ 是一个将样本空间 $\mathscr{X}$ 投影到高维空间 $\mathscr{F}$ 的一个映射，则有：

$$x_i \in \mathscr{X} \mapsto \boldsymbol{\Phi}(x_i)\in \mathscr{F} \tag{9}$$

考虑到 PCP-SDP 的投影目标，建立如下优化框架：

$$\begin{aligned} &\min_{\Phi} \mathscr{S}(\boldsymbol{\Phi}) \\ &\text{s.t.}\ \|\boldsymbol{\Phi}(x_i)-\boldsymbol{\Phi}(x_j)\|_{\mathscr{F}}<\varepsilon,\forall (x_i,x_j)\in \boldsymbol{ML} \\ &\quad\ \ \|\boldsymbol{\Phi}(x_i)-\boldsymbol{\Phi}(x_j)\|_{\mathscr{F}}<\sigma,\forall (x_i,x_j)\in \boldsymbol{CL} \end{aligned} \tag{10}$$

其中，ε 是一个远小于 σ 的正数；$\mathscr{S}(\boldsymbol{\Phi})$ 为 $\boldsymbol{\Phi}$ 的光滑度度量函数，$\mathscr{S}(\boldsymbol{\Phi})$ 越小，意味着 $\boldsymbol{\Phi}$ 越光滑；$\|\cdot\|_{\mathscr{F}}$ 为 $\mathscr{F}$ 空间上的距离度量。

为了度量 $\boldsymbol{\Phi}$ 的光滑度，PCP-SDP 算法引入了谱图理论[29]，令 $k_{ij}=\langle \boldsymbol{\Phi}(x_i),\boldsymbol{\Phi}(x_j)\rangle_{\mathscr{F}}$，则 $\boldsymbol{K}=[k_{ij}]_{n\times n}$ 为对称的正定矩阵，记为 $\boldsymbol{K}\geq 0$，构造的度量函数如下：

$$\begin{aligned} \mathscr{S}(\boldsymbol{\Phi}) &= \frac{1}{2}\sum_{i,j=1}^{n} w_{ij}\left\| \frac{\boldsymbol{\Phi}(x_i)}{\sqrt{d_{ii}}}-\frac{\boldsymbol{\Phi}(x_j)}{\sqrt{d_{jj}}}\right\|_{\mathscr{F}}^2 \\ &= \frac{1}{2}\sum_{i,j=1}^{n} w_{ij}\left(\frac{1}{d_{ii}}k_{ii}+\frac{1}{d_{jj}}k_{jj}-2\frac{1}{\sqrt{d_{ii}d_{jj}}}k_{ij}\right) \\ &= \sum_{i=1}^{n}k_{ii}-\sum_{i,j=1}^{n}\frac{w_{ij}}{\sqrt{d_{ii}d_{jj}}}k_{ij} \\ &= \boldsymbol{I}\cdot\boldsymbol{K}-(\boldsymbol{D}^{-1/2}\boldsymbol{W}\boldsymbol{D}^{-1/2})\cdot\boldsymbol{K} \\ &= \bar{\boldsymbol{L}}\cdot\boldsymbol{K} \end{aligned} \tag{11}$$

其中，$\boldsymbol{A}\cdot\boldsymbol{B}$ 表示矩阵点积，如式(12)所示：

$$\boldsymbol{A}\cdot\boldsymbol{B}=\sum_{i=1}^{n}\sum_{j=1}^{n}a_{ij}b_{ij} \tag{12}$$

其中，a_{ij}，b_{ij} 分别为 $\boldsymbol{A}$，$\boldsymbol{B}$ 的元素。

基于以上分析，PCP-SDP 模型的构建过程如下：

$$\begin{aligned} &\min_{K}\bar{\boldsymbol{L}}\cdot\boldsymbol{K} \\ &\text{s.t.}\ k_{ii}=1,i=1,2,\cdots,n \\ &\quad\ \ k_{ij}=1,\forall (x_i,x_j)\in \boldsymbol{ML} \\ &\quad\ \ k_{ij}=0,\forall (x_i,x_j)\in \boldsymbol{CL} \\ &\quad\ \ \boldsymbol{K}\geq 0 \end{aligned} \tag{13}$$

该优化问题为一个半定规划问题[30]（Semidefinite Programing，SDP），可以采用文献[31]提出的 CSDP 方法或者文献[32]提出的 SeDuMi 方法不断迭代直至稳定，最终促使在 Φ 上的映射逐步光滑，使位于同一流形结构上的数据尽可能紧密，结合成对约束的基本性质，最终使得成对约束被逐步传播到整个数据集上，因此 PCP-SDP 算法因其优秀的约束传播效果，被广泛地应用于各种算法。

3　CGBS-SymNMF 算法

3.1　算法框架

对称非负矩阵分解算法将亲和矩阵分解为类指示矩阵 $\boldsymbol{H}\in R^{n\times k}$ 和它转置的乘积。对于给定的数据集 $\boldsymbol{X}$，假设存在 k 个类簇 $\{c_1,c_2,c_3,\cdots,c_k\}$，且有 $c_1+c_2+\cdots+c_k=n$，$\forall x_i\in\boldsymbol{X}$，当 x_i 属于第 $m\in[1,k]$ 个类簇时，对于类指示矩阵有 $\boldsymbol{H}_{im}=1$，反之 $\boldsymbol{H}_{im}=0$。因此一个高质量的类指示矩阵应当具备良好的类别判断能力；同样地，作为被分解的亲和矩阵 $\boldsymbol{S}$，也应当具备类别明显的空间几何结构。基于前文的分析，本文提出了 CGBS-SymNMF 算法，该算法在 SymNMF 的基础上引入了 PCP-SDP 模型，然后构建了两个正则项，用于提高模型的性能，具体构造过程如下。

（1）模型约束图正则项。本文从谱图理论的角度出发，利用成对约束构造一个约束图正则项，使得 $\boldsymbol{H}$ 能够充分学习约束图的空间几何结构。通过成对约束集构造约束图，使用其中的“必连”约束构造正约束图 $\boldsymbol{L}^M=\boldsymbol{D}^M-\boldsymbol{S}^M$，使用“勿连”约束构造负约束图 $\boldsymbol{L}^C=\boldsymbol{D}^C-\boldsymbol{S}^C$，其中 $\boldsymbol{D}^M$ 和 $\boldsymbol{D}^C$ 为对角矩阵，其对角元素 $\boldsymbol{D}_{ii}^M=\sum_{j=1}^{n}\boldsymbol{S}_{ij}^M$，$\boldsymbol{D}_{ii}^C=\sum_{j=1}^{n}\boldsymbol{S}_{ij}^C$，$\boldsymbol{S}^M$，$\boldsymbol{S}^C$ 可由式(14)、式(15)进行构造。

$$\boldsymbol{S}^M=\begin{cases}1, & \text{if}(x_i,x_j)\in \boldsymbol{ML} \\ 0, & \text{otherwise}\end{cases} \tag{14}$$

$$\boldsymbol{S}^C=\begin{cases}1, & \text{if}(x_i,x_j)\in \boldsymbol{CL} \\ 0, & \text{otherwise}\end{cases} \tag{15}$$

基于以上分析，对于给定的 $\boldsymbol{S}^M$，$\boldsymbol{S}^C$ 以及需要优化的 $\boldsymbol{H}$，正、负约束图正则可以分别由式(16)、式(17)进行构造。

$$\begin{aligned} \frac{1}{2}\sum_{i,j\in ML}^{n}\boldsymbol{S}_{i,j}^M\|\boldsymbol{h}_i-\boldsymbol{h}_j\|^2 &= \frac{1}{2}(2\sum_{i}^{n}\boldsymbol{D}_{ii}^M\boldsymbol{h}_i\boldsymbol{h}_i^{\mathrm{T}}-2\sum_{i,j}^{n}\boldsymbol{S}_{ij}^M\boldsymbol{h}_i\boldsymbol{h}_j^{\mathrm{T}}) \\ &= Tr(\boldsymbol{H}^{\mathrm{T}}\boldsymbol{D}^M\boldsymbol{H})-Tr(\boldsymbol{H}^{\mathrm{T}}\boldsymbol{S}^M\boldsymbol{H}) \\ &= Tr(\boldsymbol{H}^{\mathrm{T}}\boldsymbol{L}^M\boldsymbol{H}) \end{aligned} \tag{16}$$

$$\begin{aligned} \frac{1}{2}\sum_{i,j\in CL}^{n}\boldsymbol{S}_{i,j}^C\|\boldsymbol{h}_i-\boldsymbol{h}_j\|^2 &= \frac{1}{2}(2\sum_{i}^{n}\boldsymbol{D}_{ii}^C\boldsymbol{h}_i\boldsymbol{h}_i^{\mathrm{T}}-2\sum_{i,j}^{n}\boldsymbol{S}_{ij}^C\boldsymbol{h}_i\boldsymbol{h}_j^{\mathrm{T}}) \\ &= Tr(\boldsymbol{H}^{\mathrm{T}}\boldsymbol{D}^C\boldsymbol{H})-Tr(\boldsymbol{H}^{\mathrm{T}}\boldsymbol{S}^C\boldsymbol{H}) \\ &= Tr(\boldsymbol{H}^{\mathrm{T}}\boldsymbol{L}^C\boldsymbol{H}) \end{aligned} \tag{17}$$

其中，$\boldsymbol{h}_i$ 为 $\boldsymbol{H}$ 的行向量。因此综合式(16)、式(17)，可以得到本文的约束图正则项，如式(18)所示：

$$Tr(\boldsymbol{H}^{\mathrm{T}}(\boldsymbol{L}^M-\alpha\boldsymbol{L}^C)\boldsymbol{H}) \tag{18}$$

（2）模型块稀疏正则项。在理想情况下，$\boldsymbol{H}\boldsymbol{H}^{\mathrm{T}}$ 应当表现出块对角结构[33]，当 $(x_i,x_j)\in\boldsymbol{ML}$ 时，其对应元素 $(\boldsymbol{H}\boldsymbol{H}^{\mathrm{T}})_{ij}=1$；当 $(x_i,x_j)\in\boldsymbol{CL}$ 时，$(\boldsymbol{H}\boldsymbol{H}^{\mathrm{T}})_{ij}=0$。但是，在实际应用上，$\boldsymbol{H}\boldsymbol{H}^{\mathrm{T}}$ 往往是非块对角的，因此为了促使其具备块对角结构，本文主要使用“勿连”约束构建块对角正则项，同时为了增加该项的稳定性，降低因低质量的初始“勿连”约束对模型效果产生的影响，采用矩阵 L_1 范式进行约束，最终构建的块稀疏正则项如式(19)所示：

$$\|\boldsymbol{S}^C\odot\boldsymbol{H}\boldsymbol{H}^{\mathrm{T}}\|_1 \tag{19}$$

基于以上分析，最终提出的 CGBS-SNMF 算法模型如式(20)所示：

$$\begin{aligned} &\min_{K,H}\|\boldsymbol{K}-\boldsymbol{H}\boldsymbol{H}^{\mathrm{T}}\|_F^2+\lambda_1 Tr(\boldsymbol{H}^{\mathrm{T}}(\boldsymbol{L}^M-\alpha\boldsymbol{L}^C)\boldsymbol{H})+ \\ &\lambda_2 Tr(\boldsymbol{L}\boldsymbol{K})+\lambda_3\|\boldsymbol{S}^C\odot\boldsymbol{H}\boldsymbol{H}^{\mathrm{T}}\| \\ &\text{s.t.}\ \boldsymbol{H}\geqslant 0 \\ &\quad\ \ \boldsymbol{K}\geqslant 0 \\ &\quad\ \ k_{ii}=1,\forall x_i\in\boldsymbol{X} \\ &\quad\ \ k_{ij}=1,\forall (x_i,x_j)\in\boldsymbol{ML} \\ &\quad\ \ k_{ij}=0,\forall (x_i,x_j)\in\boldsymbol{CL} \end{aligned} \tag{20}$$

3.2 模型优化求解

由于对矩阵 $\boldsymbol{K}$ 中元素的等式约束，问题(20)变得难以求解，因此本文采用文献[34]提出的方法对 $\boldsymbol{K}$ 进行松弛，将问题(20)转化为以下问题进行优化：

$$\begin{aligned}&\min_{\boldsymbol{K},\boldsymbol{H}}\|\boldsymbol{K}-\boldsymbol{H}\boldsymbol{H}^{\mathrm{T}}\|_F^2+\lambda_1 Tr(\boldsymbol{H}^{\mathrm{T}}(\boldsymbol{L}^{\boldsymbol{M}}-\alpha\boldsymbol{L}^{\boldsymbol{C}})\boldsymbol{H})+\\&\quad\lambda_2 Tr(\boldsymbol{LK})+\lambda_3\|\boldsymbol{S}^{\boldsymbol{C}}\odot\boldsymbol{H}\boldsymbol{H}^{\mathrm{T}}\|_1+\beta\|\boldsymbol{P}\odot(\boldsymbol{K}-\boldsymbol{Z})\|_F^2\\&\text{s. t. }\boldsymbol{H}\geqslant 0\\&\quad\ \boldsymbol{K}\geqslant 0\end{aligned}\tag{21}$$

其中，$\boldsymbol{P}\in R^{n\times n}$ 为指示矩阵，可根据式(22)进行构造；$\boldsymbol{Z}\in R^{n\times n}$ 为辅助矩阵，可根据式(23)进行构造。

$$p_{ij}=\begin{cases}1, & \text{if}(x_i,x_j)\in \boldsymbol{ML}\ \text{or}\ (x_i,x_j)\in\boldsymbol{CL}\ \text{or}\ i=j\\0, & \text{otherwise}\end{cases}\tag{22}$$

$$z_{ij}=\begin{cases}1, & \text{if}(x_i,x_j)\in \boldsymbol{ML}\ \text{or}\ i=j\\0, & \text{otherwise}\end{cases}\tag{23}$$

考虑到问题(21)仅对单个变量的优化是凸的，因此本文引入拉格朗日乘子法[35] (Lagrangian multiplier method)，采用交叉迭代的方式优化 $\boldsymbol{K}$ 和 $\boldsymbol{H}$ 变量，构造的拉格朗日函数如下：

$$\begin{aligned}\mathscr{L}(\boldsymbol{K},\boldsymbol{H},\boldsymbol{\Psi},\boldsymbol{\gamma})=&\|\boldsymbol{K}-\boldsymbol{H}\boldsymbol{H}^{\mathrm{T}}\|_F^2+\lambda_1 Tr(\boldsymbol{H}^{\mathrm{T}}(\boldsymbol{L}^{\boldsymbol{M}}-\alpha\boldsymbol{L}^{\boldsymbol{C}})\\&\boldsymbol{H})+\lambda_2 Tr(\boldsymbol{LK})+\beta\|\boldsymbol{P}\odot(\boldsymbol{K}-\boldsymbol{Z})\|_F^2+\\&\lambda_3\|\boldsymbol{S}^{\boldsymbol{C}}\odot\boldsymbol{H}\boldsymbol{H}^{\mathrm{T}}\|_1+\langle\boldsymbol{\Psi},\boldsymbol{K}\rangle+\langle\boldsymbol{\gamma},\boldsymbol{H}\rangle\end{aligned}\tag{24}$$

其中，$\boldsymbol{\Psi},\boldsymbol{\gamma}$ 为拉格朗日乘子。

对式(24)的优化变量 $\boldsymbol{K}$ 和 $\boldsymbol{H}$ 分别求一阶偏导，可得：

$$\frac{\partial\mathscr{L}}{\partial K}=2\boldsymbol{K}-2\boldsymbol{H}\boldsymbol{H}^{\mathrm{T}}+\lambda_2\boldsymbol{L}+2\beta\boldsymbol{P}\odot\boldsymbol{K}-2\boldsymbol{P}\odot\boldsymbol{Z}+\boldsymbol{\Psi}\tag{25}$$

$$\frac{\partial\mathscr{L}}{\partial\boldsymbol{H}}=-4\boldsymbol{KH}+4\boldsymbol{H}\boldsymbol{H}^{\mathrm{T}}\boldsymbol{H}+2\lambda_1(\boldsymbol{L}^{\boldsymbol{M}}-\alpha\boldsymbol{L}^{\boldsymbol{C}})\boldsymbol{H}+2\lambda_3\boldsymbol{S}^{\boldsymbol{C}}\boldsymbol{H}+\boldsymbol{\gamma}\tag{26}$$

根据 KKT 条件(Karush-Kuhn-Tucker conditions)有 $\boldsymbol{\Psi}_{ij}\boldsymbol{K}_{ij}=0$ 和 $\gamma_{ij}\boldsymbol{H}_{ij}^4=0$，令式(25)、式(26)两个偏导式等于零并且分别在等式两边乘以 $\boldsymbol{K}_{ij}$ 和 $\boldsymbol{H}_{ij}^4$，可得：

$$\begin{aligned}&2\boldsymbol{K}_{ij}\boldsymbol{K}_{ij}-2(\boldsymbol{H}\boldsymbol{H}^{\mathrm{T}})_{ij}\boldsymbol{K}_{ij}+\lambda_2\boldsymbol{L}_{ij}\boldsymbol{K}_{ij}+2\beta(\boldsymbol{P}\odot\boldsymbol{K})_{ij}\boldsymbol{K}_{ij}-\\&2(\boldsymbol{P}\odot\boldsymbol{Z})_{ij}\boldsymbol{K}_{ij}+\boldsymbol{\Psi}_{ij}\boldsymbol{K}_{ij}=0\end{aligned}\tag{27}$$

$$\begin{aligned}&-4(\boldsymbol{KH})_{ij}\boldsymbol{H}_{ij}^4+4(\boldsymbol{H}\boldsymbol{H}^{\mathrm{T}}\boldsymbol{H})_{ij}\boldsymbol{H}_{ij}^4+2\lambda_1(\boldsymbol{L}^M-\alpha\boldsymbol{L}^C)_{ij}\boldsymbol{H}_{ij}^4+\\&2\lambda_3(\boldsymbol{S}^{\boldsymbol{C}}\boldsymbol{H})_{ij}\boldsymbol{H}_{ij}^4+\gamma_{ij}\boldsymbol{H}_{ij}^4=0\end{aligned}\tag{28}$$

通过式(27)、式(28)，可以获得优化变量 $\boldsymbol{K}$ 和 $\boldsymbol{H}$ 的更新公式，如式(29)、式(30)所示：

$$\boldsymbol{K}_{ij}\leftarrow\boldsymbol{K}_{ij}\left(\frac{2\boldsymbol{H}\boldsymbol{H}^{\mathrm{T}}+\lambda_2\boldsymbol{S}+2\beta\boldsymbol{P}\odot\boldsymbol{Z}}{2\boldsymbol{K}+\lambda_2\boldsymbol{D}+2\beta\boldsymbol{P}\odot\boldsymbol{K}}\right)_{ij}\tag{29}$$

$$\boldsymbol{H}_{ij}\leftarrow\boldsymbol{H}_{ij}\left(\frac{(2\boldsymbol{KH}+\lambda_1\boldsymbol{S}^M\boldsymbol{H}+\alpha\lambda_1\boldsymbol{D}^C\boldsymbol{H})_{ij}}{(2\boldsymbol{H}\boldsymbol{H}^{\mathrm{T}}\boldsymbol{H}+\lambda_1\boldsymbol{D}^M\boldsymbol{H}+\alpha\lambda_1\boldsymbol{S}^C\boldsymbol{H}+\lambda_3\boldsymbol{S}^C\boldsymbol{H})_{ij}}\right)^{\frac{1}{4}}\tag{30}$$

通过不断迭代更新式(29)、式(30)，该算法的收敛精度设置为 0.02，当达到该收敛精度时，就能获得模型(21)的解 $\boldsymbol{K}$ 和 $\boldsymbol{H}$，最终 CGBS-SymNMF 算法如算法 1 所示。

算法 1 CGBS-SymNMF 算法

输入：成对约束集 CL 和 ML，α=0.01，β=10，数据集 **X**，样本类簇数 k

输出：样本标签 $\{l_1,l_2,l_3,\cdots,l_n\}$

1. 根据 CL，ML 集合计算 $\mathbf{L}^{\mathrm{M}}$，$\mathbf{L}^{\mathrm{C}}$，**P**，Z，根据样本集计算图矩阵 $\mathbf{L}=\mathrm{I}-\mathbf{D}^{-1/2}\mathbf{S}\mathbf{D}^{-1/2}$，初始化优化变量 $\mathbf{K}=\mathrm{rand}^+(\mathrm{n},\mathrm{n})$，$\mathbf{H}=\mathrm{rand}^+(\mathrm{n},\mathrm{k})$
2. 通过式(28)更新 **K**
3. 通过式(29)更新 **H**
4. 重复步骤 2、步骤 3 直至算法收敛
5. 计算矩阵 **H** 每一行最大元素的坐标，并将其作为对应样本的标签

3.3 收敛性证明

为了便于后续的证明描述，设定：

$$\begin{aligned}\Omega(\boldsymbol{K},\boldsymbol{H})=&\|\boldsymbol{K}-\boldsymbol{H}\boldsymbol{H}^{\mathrm{T}}\|_F^2+\lambda_1 Tr(\boldsymbol{H}^{\mathrm{T}}(\boldsymbol{L}^M-\alpha\boldsymbol{L}^C)\boldsymbol{H})+\\&\lambda_2 Tr(\boldsymbol{LK})+\beta\|P\odot(\boldsymbol{K}-\boldsymbol{Z})\|_F^2+\\&\lambda_3\|S^C\odot\boldsymbol{H}\boldsymbol{H}^{\mathrm{T}}\|_1\end{aligned}\tag{31}$$

定理 1 算法 1 是收敛的当且仅当 $\Omega(\boldsymbol{K},\boldsymbol{H})$ 是非增的。

若要证明定理 1，则需要证明以下两个问题：

(1)固定 $\boldsymbol{H}$，通过更新式(29)，$\Omega(\boldsymbol{K})$ 是非增的。

(2)固定 $\boldsymbol{K}$，通过更新式(30)，$\Omega(\boldsymbol{H})$ 是非增的。

根据定义 1[16] 和引理 1[17]，只需要找到 $\Omega(\boldsymbol{K})$ 和 $\Omega(\boldsymbol{H})$ 的上界辅助函数，就可直接证明问题(1)、问题(2)。

定义 1 $G(x,x^t)$ 为 $\Omega(x)$ 的一个辅助函数，当且仅当以下条件成立：

$$G(x,x^t)\geqslant\Omega(x),G(x^t,x^t)=\Omega(x^t)\tag{32}$$

引理 1 如果 $G(x,x^t)$ 是一个辅助函数，并且在其变量迭代更新下的 $\Omega(x)$ 非增，则有：

$$x^{t+1}=\arg\min_x G(x,x^t)\tag{33}$$

首先，证明问题(1)：

通过固定变量 $\boldsymbol{H}$，可得 $\Omega(\boldsymbol{K})$ 的表达式，如式(34)所示：

$$\begin{aligned}\Omega(\boldsymbol{K})=&Tr(\boldsymbol{K}\boldsymbol{K}^{\mathrm{T}}-2K\boldsymbol{H}\boldsymbol{H}^{\mathrm{T}})+\lambda_2 Tr(\boldsymbol{LK})+\\&\beta Tr(\boldsymbol{P}\odot(\boldsymbol{K}\boldsymbol{K}^{\mathrm{T}}-2\boldsymbol{K}\boldsymbol{Z}^{\mathrm{T}}))\end{aligned}\tag{34}$$

因此，从元素更新的角度来看，$\Omega(\boldsymbol{K})$ 的一阶导数可表示为式(35)：

$$\begin{aligned}\Omega'(\boldsymbol{K}_{ij})&=\frac{\partial\Omega(\boldsymbol{K}_{ij})}{\partial\boldsymbol{K}_{ij}}\\&=2(\boldsymbol{K}+\beta\boldsymbol{P}\odot\boldsymbol{K})_{ij}+\lambda_2\boldsymbol{L}-2(\boldsymbol{H}\boldsymbol{H}^{\mathrm{T}}+\boldsymbol{P}\odot\boldsymbol{Z})_{ij}\end{aligned}\tag{35}$$

同样地，$\Omega(\boldsymbol{K})$ 的二阶导数可表示为式(36)：

$$\Omega''(\boldsymbol{K}_{ij})=\frac{\partial^2\Omega(\boldsymbol{K}_{ij})}{\partial\boldsymbol{K}_{ij}^2}=2(1+\beta\boldsymbol{P})_{ij}\tag{36}$$

其中，$1\in R^{n\times n}$ 为一个全 1 矩阵，$\Omega(\boldsymbol{K})$ 后续的更高阶导函数如式(37)所示：

$$\Omega^{(3)}(\boldsymbol{K}_{ij})=\Omega^{(4)}(\boldsymbol{K}_{ij})=\cdots=\Omega^{(n)}(\boldsymbol{K}_{ij})=0\tag{37}$$

根据定义 2[36]，可以将 $\Omega(\boldsymbol{K})$ 在 K_{ij}^t 点进行泰勒展开(Taylor expansion)，如式(38)所示：

$$\begin{aligned}\Omega(\boldsymbol{K}_{ij})=&\Omega(\boldsymbol{K}_{ij}^t)+\Omega'(\boldsymbol{K}_{ij}^t)(\boldsymbol{K}_{ij}-\boldsymbol{K}_{ij}^t))+\\&\frac{\Omega''(\boldsymbol{K}_{ij}^t))}{2}(\boldsymbol{K}_{ij}-\boldsymbol{K}_{ij}^t)^2\end{aligned}\tag{38}$$

定义 2 一个函数可以表示为有限项的和，并且这些项是通过函数在单个点的导数值计算出的，表达式如下：

$$\Omega(y)=\sum_{n=0}^{\infty}\frac{\Omega^{(n)}(a)}{n!}(y-a)^n\tag{39}$$

其中，$\Omega^{(n)}(a)$ 表示 $\Omega(\boldsymbol{K}_{ij})$ 在点 a 的 n 阶导。

考虑到 $\Omega(\boldsymbol{K}_{ij})$ 在 $\boldsymbol{K}_{ij}^t$ 点的泰勒展开式以及定义 1，因此构造辅助函数 $G(\boldsymbol{K}_{ij},\boldsymbol{K}_{ij}^t)$ 如下：

$$\begin{aligned}G(\boldsymbol{K}_{ij},\boldsymbol{K}_{ij}^t)=&\Omega(\boldsymbol{K}_{ij}^t)+\Omega'(\boldsymbol{K}_{ij}^t)(\boldsymbol{K}_{ij}-\boldsymbol{K}_{ij}^t)+\\&\Gamma(\boldsymbol{K}_{ij}^t)(\boldsymbol{K}_{ij}-\boldsymbol{K}_{ij}^t)^2\end{aligned}\tag{40}$$

其中，

$$\Gamma(\boldsymbol{K}_{ij}^t)=\frac{(2\boldsymbol{K}^t+\lambda_2\boldsymbol{D}+2\beta\boldsymbol{P}\odot\boldsymbol{K}^t)_{ij}}{2K_{ij}^t} \tag{41}$$

显然,当 $\boldsymbol{K}=\boldsymbol{K}^t$ 时,有 $G(\boldsymbol{K}^t,\boldsymbol{K}^t)=\Omega(\boldsymbol{K}^t)$,因此只需要证明不等式 $G(\boldsymbol{K},\boldsymbol{K}_{ij}^t)\geqslant\Omega(\boldsymbol{K})$,则 $\Omega(\boldsymbol{K})$ 非增。结合式(38)以及式(40),若要上述不等式成立,可等价地证明不等式(42)。

$$\Gamma(\boldsymbol{K}_{ij}^t)\geqslant\frac{\Omega''(\boldsymbol{K}_{ij}^t)}{2} \tag{42}$$

结合式(36)以及式(41)有:

$$\frac{(2\boldsymbol{K}^t+\lambda_2\boldsymbol{D}+2\beta\boldsymbol{P}\odot\boldsymbol{K}^t)_{ij}}{2\boldsymbol{K}_{ij}^t}\geqslant\frac{2(\boldsymbol{K}^t+\beta\boldsymbol{P}\odot\boldsymbol{K}^t)_{ij}}{2\boldsymbol{K}_{ij}^t}=(1+\beta\boldsymbol{P})_{ij} \tag{43}$$

因此不等式(42)满足,则定义1成立,因此矩阵 $\boldsymbol{K}$ 在式(29)的更新下是非增的。

由于 $\Omega(\boldsymbol{K})$ 是非增的,根据引理1,并且结合辅助函数 $G(\boldsymbol{K}_{ij},\boldsymbol{K}_{ij}^t)$,$\boldsymbol{K}$ 的迭代式可以由如下方式给出:

$$\boldsymbol{K}_{ij}^{t+1}=\arg\min_{K} G(\boldsymbol{K}_{ij},\boldsymbol{K}_{ij}^{t+1}) \tag{44}$$

令 $\partial G(\boldsymbol{K}_{ij},\boldsymbol{K}_{ij}^{t+1})/\partial\boldsymbol{K}_{ij}=0$,可得:

$$\frac{\partial(\Omega'(\boldsymbol{K}_{ij}^t)(\boldsymbol{K}_{ij}-\boldsymbol{K}_{ij}^t)+\Gamma(\boldsymbol{K}_{ij}^t)(\boldsymbol{K}_{ij}-\boldsymbol{K}_{ij}^t)^2)}{\partial\boldsymbol{K}_{ij}}=0$$

$$\Rightarrow\Omega'(\boldsymbol{K}_{ij}^t)+2\Gamma(\boldsymbol{K}_{ij}^t)(\boldsymbol{K}_{ij}-\boldsymbol{K}_{ij}^t)=0$$

$$\Rightarrow\boldsymbol{K}_{ij}=\frac{2\Gamma(\boldsymbol{K}_{ij}^t)\boldsymbol{K}_{ij}^t-\Omega'(\boldsymbol{K}_{ij}^t)}{2\Gamma(\boldsymbol{K}_{ij}^t)}$$

$$\Rightarrow\boldsymbol{K}_{ij}=\boldsymbol{K}_{ij}^t\left(\frac{2\boldsymbol{H}\boldsymbol{H}^{\mathrm{T}}+\lambda_2\boldsymbol{S}+2\beta\boldsymbol{P}\odot\boldsymbol{Z}}{\boldsymbol{K}^t+\lambda_2\boldsymbol{D}+2\beta\boldsymbol{P}\odot\boldsymbol{K}^t}\right)_{ij} \tag{45}$$

下面证明问题(2):

通过固定变量 $\boldsymbol{K}$,可得 $\Omega(\boldsymbol{H})$ 的表达式:

$$\Omega(\boldsymbol{H})=Tr(\boldsymbol{H}\boldsymbol{H}^{\mathrm{T}}\boldsymbol{H}\boldsymbol{H}^{\mathrm{T}})-\lambda_1 Tr(\boldsymbol{H}^{\mathrm{T}}\alpha\boldsymbol{L}^C-\boldsymbol{L}^M\boldsymbol{H})-\lambda_3 Tr((\boldsymbol{L}^C-\boldsymbol{D}^C)\boldsymbol{H}\boldsymbol{H}^{\mathrm{T}})-2Tr(\boldsymbol{K}\boldsymbol{H}\boldsymbol{H}^{\mathrm{T}}) \tag{46}$$

基于式(46),本文构建辅助函数 $G(\boldsymbol{H}_{ij},\boldsymbol{H}_{ij}^t)$ 如下:

$$\begin{aligned}G(\boldsymbol{H}_{ij},\boldsymbol{H}_{ij}^t)=&\sum_{i,j=1}^{n}\sum_{m=1}^{k}(\boldsymbol{H}^t\boldsymbol{H}^{t^{\mathrm{T}}})_{ij}\boldsymbol{H}_{im}^t\frac{\boldsymbol{H}_{jm}^4}{\boldsymbol{H}_{jm}^{t^3}}-\lambda_1\sum_{i,j=1}^{n}\sum_{m=1}^{k}(\alpha\boldsymbol{L}_{ij}^C-\\&\boldsymbol{L}_{ij}^M)\boldsymbol{H}_{im}^t\boldsymbol{H}_{jm}^t\left(1+\ln\frac{\boldsymbol{H}_{im}\boldsymbol{H}_{jm}}{\boldsymbol{H}_{im}^t\boldsymbol{H}_{jm}^t}\right)-\\&2\sum_{i,j=1}^{n}\sum_{m=1}^{k}\boldsymbol{K}_{ij}\boldsymbol{H}_{im}^t\boldsymbol{H}_{jm}^t\left(1+\ln\frac{\boldsymbol{H}_{im}\boldsymbol{H}_{jm}}{\boldsymbol{H}_{im}^t\boldsymbol{H}_{jm}^t}\right)-\\&\lambda_3\sum_{i,j=1}^{n}\sum_{m=1}^{k}(\boldsymbol{L}_{ij}^C-\boldsymbol{D}_{ij}^C)\boldsymbol{H}_{im}^t\boldsymbol{H}_{jm}^t\left(1+\ln\frac{\boldsymbol{H}_{im}\boldsymbol{H}_{jm}}{\boldsymbol{H}_{im}^t\boldsymbol{H}_{jm}^t}\right)\end{aligned} \tag{47}$$

当 $\boldsymbol{H}=\boldsymbol{H}^t$ 时,有 $G(\boldsymbol{H}^t,\boldsymbol{H}^t)=\Omega(\boldsymbol{H}^t)$,下面证明 $G(\boldsymbol{H},\boldsymbol{H}_{ij}^t)\geqslant\Omega(\boldsymbol{H})$,则 $\Omega(\boldsymbol{H})$ 非增。为便于后续的证明,令 $\boldsymbol{H}=\eta\boldsymbol{H}^t$,则 $G(\boldsymbol{H}_{ij},\boldsymbol{H}_{ij}^t)$ 的第一项为:

$$\begin{aligned}&\sum_{i,j=1}^{n}\sum_{m=1}^{k}(\boldsymbol{H}^t\boldsymbol{H}^{t^{\mathrm{T}}})_{ij}\boldsymbol{H}_{im}^t\frac{\boldsymbol{H}_{jm}^4}{\boldsymbol{H}_{jm}^{t^3}}\\&=\sum_{i,j=1}^{n}\sum_{m=1}^{k}(\boldsymbol{H}^t\boldsymbol{H}^{t^{\mathrm{T}}})_{ij}\boldsymbol{H}_{im}^t\boldsymbol{H}_{jm}^t\eta_{jm}^4\\&=\sum_{i,j=1}^{n}\sum_{m,r=1}^{k}\boldsymbol{H}_{ir}^t\boldsymbol{H}_{jr}^t\boldsymbol{H}_{im}^t\boldsymbol{H}_{jm}^t\eta_{jm}^4\end{aligned} \tag{48}$$

因为 i,j 和 m,r 都是对称的,并且根据不等式原理,式(49)成立:

$$a^4+b^4+c^4+d^4\geqslant2(a^2b^2+c^2d^2)\geqslant4abcd \tag{49}$$

所以可以得到:

$$\eta_{jm}^4=\frac{\eta_{im}^4+\eta_{jm}^4+\eta_{ir}^4+\eta_{jr}^4}{4}\geqslant\eta_{im}\eta_{jm}\eta_{ir}\eta_{jr} \tag{50}$$

根据式(48)和式(50)可以推出:

$$\begin{aligned}&\sum_{i,j=1}^{n}\sum_{m,r=1}^{k}\boldsymbol{H}_{ir}^t\boldsymbol{H}_{jr}^t\boldsymbol{H}_{im}^t\boldsymbol{H}_{jm}^t\eta_{jm}^4\\&\geqslant\sum_{i,j=1}^{n}\sum_{m,r=1}^{k}\boldsymbol{H}_{ir}^t\boldsymbol{H}_{jr}^t\boldsymbol{H}_{im}^t\boldsymbol{H}_{jm}^t\eta_{im}\eta_{jm}\eta_{ir}\eta_{jr}\\&=\sum_{i,j=1}^{n}\sum_{m,r=1}^{k}\boldsymbol{H}_{ir}\boldsymbol{H}_{jr}\boldsymbol{H}_{im}\boldsymbol{H}_{jm}=Tr(\boldsymbol{H}\boldsymbol{H}^{\mathrm{T}}\boldsymbol{H}\boldsymbol{H}^{\mathrm{T}})\end{aligned} \tag{51}$$

对于 $G(\boldsymbol{H}_{ij},\boldsymbol{H}_{ij}^t)$ 的第三项,根据 $1+\log x\leqslant x$,可以推出:

$$\begin{aligned}&-2\sum_{i,j=1}^{n}\sum_{m=1}^{k}\boldsymbol{K}_{ij}\boldsymbol{H}_{im}^t\boldsymbol{H}_{jm}^t\left(1+\ln\frac{\boldsymbol{H}_{im}\boldsymbol{H}_{jm}}{\boldsymbol{H}_{im}^t\boldsymbol{H}_{jm}^t}\right)\\&\geqslant-2\sum_{i,j=1}^{n}\sum_{m=1}^{k}K_{ij}\boldsymbol{H}_{im}^t\boldsymbol{H}_{jm}^t\frac{\boldsymbol{H}_{im}\boldsymbol{H}_{jm}}{\boldsymbol{H}_{im}^t\boldsymbol{H}_{jm}^t}\\&=-2\sum_{i,j=1}^{n}\sum_{m=1}^{k}K_{ij}\boldsymbol{H}_{im}\boldsymbol{H}_{jm}=-2Tr(\boldsymbol{K}\boldsymbol{H}\boldsymbol{H}^{\mathrm{T}})\end{aligned} \tag{52}$$

使用同样的方法也可以证明 $G(\boldsymbol{H}_{ij},\boldsymbol{H}_{ij}^t)$ 的第二项和第四项。

根据引理1,有:

$$\boldsymbol{H}^{t+1}=\arg\min_{\boldsymbol{H}} G(\boldsymbol{H},\boldsymbol{H}^t) \tag{53}$$

因此,通过对 $G(\boldsymbol{H},\boldsymbol{H}^t)$ 进行求导可以得到:

$$\begin{aligned}\frac{G(\boldsymbol{H},\boldsymbol{H}^t)}{\partial\boldsymbol{H}_{jm}}=&4\sum_{i,j=1}^{n}\sum_{m=1}^{k}(\boldsymbol{H}^t\boldsymbol{H}^{t^{\mathrm{T}}})_{ij}\boldsymbol{H}_{im}^t\frac{\boldsymbol{H}_{jm}^3}{\boldsymbol{H}_{jm}^{t^3}}-2\lambda_1\sum_{i,j=1}^{n}\sum_{m=1}^{k}(\alpha L_{ij}^C-\\&L_{ij}^M)\boldsymbol{H}_{im}^t\frac{\boldsymbol{H}_{jm}^t}{\boldsymbol{H}_{jm}}-4\sum_{i,j=1}^{n}\sum_{m=1}^{k}\boldsymbol{K}_{ij}\boldsymbol{H}_{im}^t\\&\frac{\boldsymbol{H}_{jm}^t}{\boldsymbol{H}_{jm}}-2\lambda_3\sum_{i,j=1}^{n}\sum_{m=1}^{k}(\boldsymbol{L}_{ij}^C-\boldsymbol{D}_{ij}^C)\boldsymbol{H}_{im}^t\frac{\boldsymbol{H}_{jm}^t}{\boldsymbol{H}_{jm}}\end{aligned} \tag{54}$$

通过对式(54)中的项进行替换,可以得到:

$$\begin{aligned}\frac{G(\boldsymbol{H},\boldsymbol{H}^t)}{\partial\boldsymbol{H}_{jm}}=&4(\boldsymbol{H}^t\boldsymbol{H}^{t^{\mathrm{T}}}\boldsymbol{H}^t)_{jm}\frac{\boldsymbol{H}_{jm}^3}{\boldsymbol{H}_{jm}^{t^3}}-4(K\boldsymbol{H}^t)_{jm}\frac{\boldsymbol{H}_{jm}^t}{\boldsymbol{H}_{jm}}-\\&2\lambda_1((\alpha\boldsymbol{L}^C-\boldsymbol{L}^M)\boldsymbol{H}^t)_{jm}\frac{\boldsymbol{H}_{jm}^t}{\boldsymbol{H}_{jm}}-2\lambda_3((\boldsymbol{L}^C-\\&D^C)\boldsymbol{H}^t)_{jm}\frac{\boldsymbol{H}_{jm}^t}{\boldsymbol{H}_{jm}}\end{aligned} \tag{55}$$

根据式(55),可以推出 $\boldsymbol{H}$ 的更新式,如下:

$$\boldsymbol{H}_{ij}=\boldsymbol{H}_{ij}^t\left(\frac{(2\boldsymbol{K}\boldsymbol{H}^t+\lambda_1\boldsymbol{S}^M\boldsymbol{H}^t+\alpha\lambda_1\boldsymbol{D}^C\boldsymbol{H}^t)_{ij}}{(2\boldsymbol{H}^t\boldsymbol{H}^{t^{\mathrm{T}}}\boldsymbol{H}^t+\lambda_1D^M\boldsymbol{H}^t+\alpha\lambda_1\boldsymbol{S}^C\boldsymbol{H}^t+\lambda_3\boldsymbol{S}^C\boldsymbol{H}^t)_{ij}}\right)^{\frac{1}{4}} \tag{56}$$

3.4 算法时间复杂度

CGBS-SymNMF 算法的时间复杂度主要是迭代更新 K 和 H 两个优化变量,通过式(28)更新 $\boldsymbol{K}$,其复杂度为 $O(\max(n^2k,n^3))$;通过式(29)更新 $\boldsymbol{H}$,其复杂度为 $O(n^2k)$;由于提出的算法通常情况下 $n>k$,因此 CGBS-SymNMF 算法的时间复杂度为 $O(n^3)$。

4 实验结果及分析

4.1 实验设置及聚类评价

本文共选取6个数据集进行实验,其基本信息如表1所列,包括3个UCI数据集,即Iris,Zoo,Seed,以及3个图片数据集,即 Mnist,USPS,YaleB 部分样图如图1所示,其中Mnist,USPS为手写数据集,分别包含0-9这10类数字的2000,1 500张手写数字图片,每张图片规格分别为28×28,16×16,Yale为人脸数据集,包含15个人的不同姿势的165张图片,每个人11张图片,每张图片规格为32×32。

为了评价提出的CGBS-SNMF算法的聚类性能,本文选取了6个经典的基于NMF和SymNMF的算法进行比较。

表 1 数据集

Table 1 Datasets

Datasets	instances	features	classes
Iris	150	4	3
Seed	210	7	3
Zoo	101	17	7
Mnist	2000	784	10
USPS	1500	256	10
Yale	165	1024	15

(1)GNMF 算法(无监督)[10]：在 NMF 基础上，加入图正则项，使得嵌入矩阵能够学习数据集的空间几何结构。

(2)SymNMF 算法(无监督)[12]：利用图聚类框架构建亲和矩阵，并将其分解为类别指示矩阵。

(3)GSymNMF 算法(半监督)[15]：利用"必连"约束构建一个约束图正则，增强指示矩阵的样本识别能力。

(4)SymNMFCC 算法(半监督)[16]：利用成对约束构建了一个新的约束聚类框架，促使隶属于同一个类簇的样本的相似度趋向于 1，隶属于不同类簇的样本的相似度趋向于 0。

(5)PCPISymNMF 算法(半监督)[18]：使用 PCP 方法生成更合理的相似图，同时对学习到的相似图使用对称非负矩阵进行分解。

(6)BNMF 算法(半监督)[37]：通过构建一个块对角结构对应标签信息，增强 NMF 的数据表达能力。

(a)手写数字 (b)Yale

图 1 图片数据部分样例图

Fig. 1 Partial examples of figure data

所有算法在 3 个 UCI 数据集上进行实验，实验设置初始成对约束量为[100:100:1000](从 100 到 1000 选取成对约束量，步长为 100)，在图片数据集选取初始成对约束量为[600:600:6000]，样本成对约束均根据样本标签随机产生。对于部分算法需要基于 P-nn 图构造亲和矩阵，设置邻域 $P=40$，σ 为每个数据集的标准差。对于提出的 CGBS-SymNMF 算法，设置 3 个参数 $\lambda_1=\lambda_2=\lambda_3=[10^{-3},10^{-2},10^{-1},1,10^1,10^2,10^3]$，由于选取成对约束的随机性以及部分算法需要依赖 K-means 完成聚类，对实验结果有较大影响，为了平衡这部分因素，每种算法在最优参数下，取相同成对约束并且在其他外部条件一致的情况下运行 20 次并取平均结果。

为了评价各算法的最终聚类效果，根据文献[34]选取聚类精度(Clustering Accuracy, Acc)作为评价指标，$Acc\in[0,1]$ 且越接近 1，模型的聚类效果越好。设样本真实标签 $\boldsymbol{L}=[l_1,l_2,l_3,\cdots,l_n]$，模型对样本的预测标签 $\boldsymbol{Y}=[y_1,y_2,y_3,\cdots,y_n]$，Acc 可按照式(57)进行计算。

$$Acc=\frac{\sum_{i=1}^{n}\delta(l_i,map(y_i))}{n} \tag{57}$$

其中 $map(\cdot)$ 为映射函数，将预测标签按最优匹配进行重排列，对于 $\forall\, i,j$，有：

$$\delta(i,j)=\begin{cases}1, & i=j\\ 0, & i\neq j\end{cases} \tag{58}$$

本次实验使用联想电脑，配置为 windows 10 专业版 64 位操作系统，Intel(R) Core(TM) i3-4150 CPU @ 3.50 GHz 3.50 GHz，8 GB DDR3 1600 MHz 内存，所有算法均在 MATLAB 9.0.0.341360(R2016a)上运行。

4.2 实验结果分析

在其他情况一致的情况下，表 2 列出了所有算法在 6 个数据集上选定最优参数运行 20 次得到的平均 Acc 值，通过比较分析可以得出以下结论：

(1)在 6 个不同的数据集上，CGBS-SymNMF 算法的聚类性能均高于同样带图正则的 GSymNMF 算法，在 USPS，Mnist，Yale 上的聚类精度提升最大，说明模型的块稀疏正则能够有效促使类判别矩阵的块对角化，从而增强模型的图学习能力。

(2)在 6 个数据集上，带正约束图正则的 GSymNMF 算法的聚类精确度超过 SymNMF 算法和 GNMF 算法，说明正约束图正则能更好地捕获非线性流型数据的空间几何结构，有效提高模型的类判别能力。CGBS-SymNMF 算法在其基础上构建了正负约束图正则，进一步增强了模型的聚类性能，其较高的聚类精度证实了这一点。

表 2 各算法在 6 个数据集上的平均 Acc 值

Table 2 Average Acc of different algorithms on six datasets

算法	数据集					
	Iris	Zoo	Seed	Mnist	USPS	Yale
SymNMF	0.944	0.857	0.913	0.242	0.183	0.397
GNMF	0.905	0.863	0.935	0.262	0.271	0.233
BNMF	0.913	0.873	0.941	0.227	0.243	0.584
GSymNMF	0.982	0.917	0.973	0.435	0.319	0.903
SymNMFCC	0.991	0.875	0.964	0.516	0.454	0.993
PCPISymNMF	0.954	0.813	0.941	0.414	0.443	0.981
CGBS-SymNMF	**1.000**	**0.941**	**0.973**	**0.524**	**0.612**	**1.000**

图 2 给出了不同成对约束数量对算法 Acc 值的影响，从图 2(a)—图 2(f)这 6 个子图可以得到如下信息：

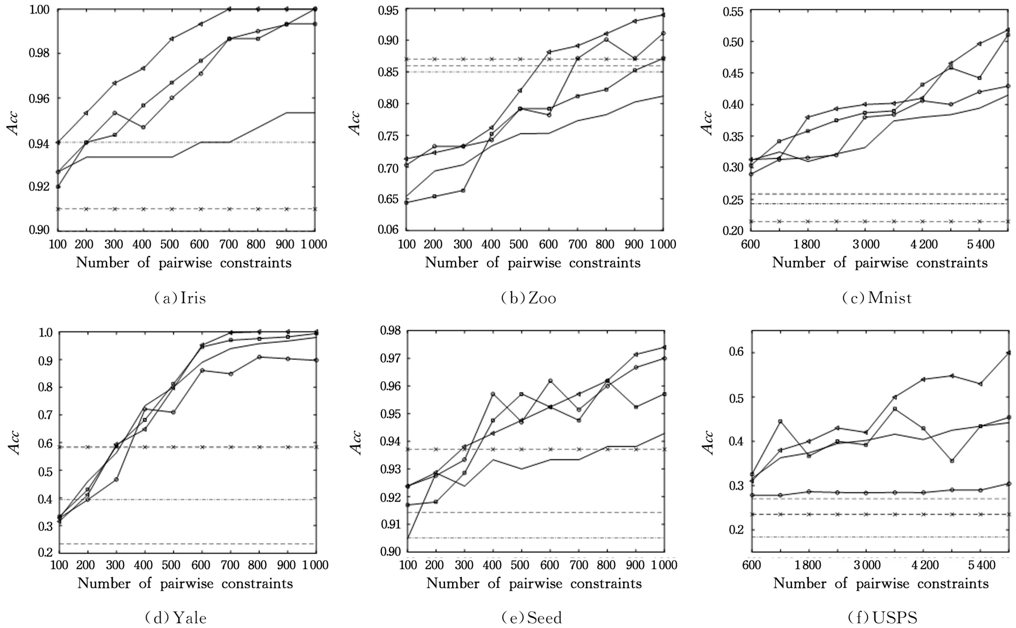
(1)提出的 CGBS-SymNMF 算法在大部分情况下，优于取同等成对约束量的基于 NMF 或 SymNMF 的半监督约束聚类算法，证明了所提算法的优秀的聚类精度。

(2)通过增加成对约束的数量，可以明显看到 CGBS-SymNMF 算法、PCPISymNMF 算法、SymNMFCC 算法和 GSymNMF 算法的聚类精度有较大提升，说明了在样本类别划分过程中，成对约束起到了较大的作用。

(3)通过增加成对约束的数量，能使得约束聚类算法的聚类精度大大超过无监督的 SymNMF 算法和 GNMF 算法，说明成对约束有助于增大隶属不同类簇样本的差异性，能有效提高算法类识别的能力。

(4)对于 PCPISymNMF 算法、SymNMFCC 算法和 GSymNMF 算法，部分成对约束数量大的对应的算法 Acc 值比成对约束数量小的对应的算法 Acc 值要低，说明随机生成的成对约束的质量并不高，而 CGBS-SymNMF 算法对于各成对约束数量的 Acc 值总体上都呈逐步上升趋势，因此提出的

算法对初始成对约束敏感度不高，算法比较稳定，同时也可以看到由“勿连”约束引导的块稀疏正则起到了很好的作用。



(a) Iris　(b) Zoo　(c) Mnist

(d) Yale　(e) Seed　(f) USPS

图 2　各算法取不同成对约束量下的 Acc 值比较

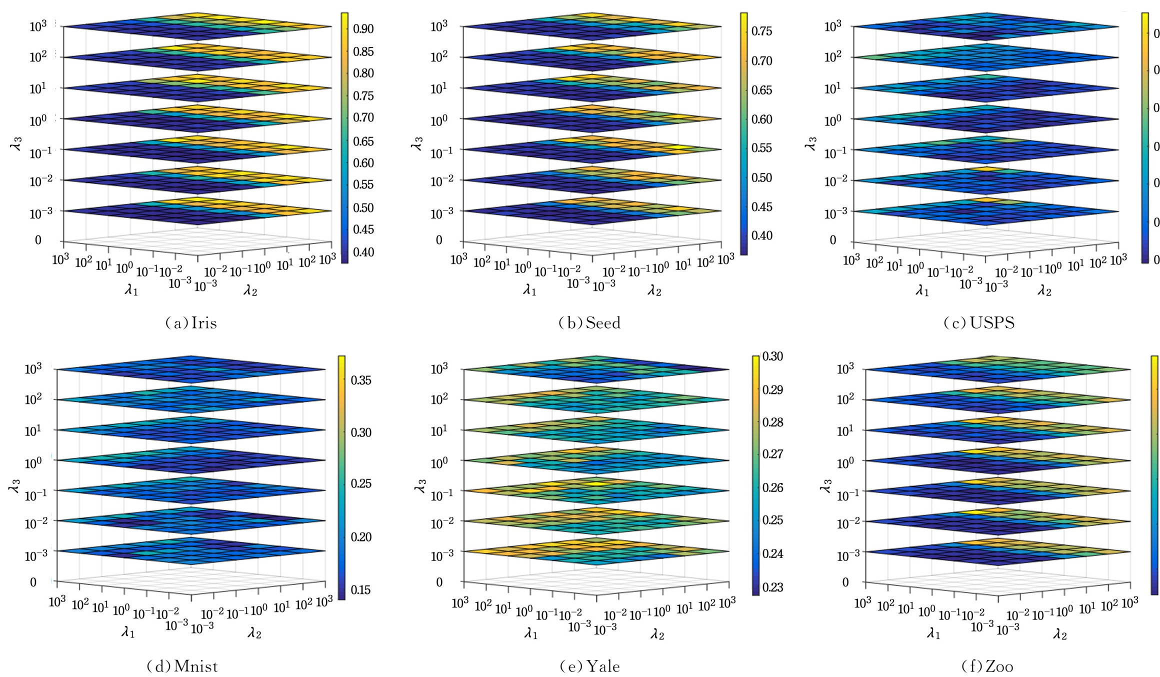
Fig. 2　Acc comparison of different algorithms under different pairwise constraints

4.3　算法参数分析

影响 CGBS-SymNMF 算法聚类精度的参数主要是 λ_1，λ_2，λ_3，其他因素一致的情况下，在 Iris，Seed，Zoo 这 3 个数据集上选取 100 成对约束量，USPS，Mnist，Yale 这 3 个数据集上选取 600 成对约束量，遍历 3 个参数所有取值，计算算法的 Acc 值，最终结果如图 3 所示，从图 3(a)－图 3(f)这 6 个子图可以得到以下信息：

(1)在 Iris，Seed，Zoo 这 3 个 UCI 数据集上，当参数 $\lambda_3 \in \{10, 10^2, 10^3\}$，$\lambda_1$，$\lambda_2 \in \{10^2, 10^3\}$ 时总能得到高的 Acc 值，较大的 λ_1，λ_2，λ_3 使得 CGBS-SymNMF 算法的约束图正则、样本图正则、块稀疏项的学习能力大大提升，增强了算法的稀疏性。

(2)在 Mnist，USPS 以及 Yale 数据集上，参数 $\lambda_3 \in \{10^{-3}, 10^{-2}, 10^{-1}\}$，参数 $\lambda_1 \in \{10^2, 10^3\}$ 时模型效果较好，参数 λ_2 对算法的影响程度较低，主要是因为 Mnist 和 USPS 数据集本身比较稀疏，不需要通过参数提高模型的稀疏性，同时基于 P-nn 构造的样本图对于大规模的样本影响较小，过于稀疏，效果比约束图正则差。



(a) Iris　(b) Seed　(c) USPS

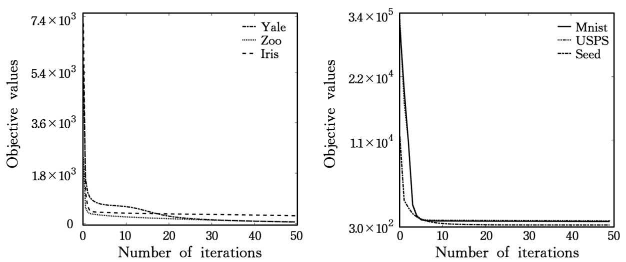
(d) Mnist　(e) Yale　(f) Zoo

图 3　CGBS-SymNMF 算法参数对 Acc 值影响

Fig. 3　Impact of CGBS-SymNMF algorithm parameters on Acc values

4.4 算法的收敛性分析

通过固定 CGBS-SymNMF 算法的 $\lambda_1,\lambda_2,\lambda_3$ 值,并且取在 6 个数据集上选取的最低成对约束量,从图 4 可以看到提出的算法在迭代 10 次之后,目标函数值能够收敛到一个稳定的水平。图 5 给出了各算法在保证收敛精度一致情况下的单次收敛耗时情况,可以看到,提出的算法在保证拥有较高聚类精度的性能前提下,运行耗时相对其他算法而言较低,而运行耗时低的 SymNMF 算法和 GSymNMF 算法却又无法达到一个高的聚类精度。因此,综合考虑聚类表现和耗时,证明了所提算法的高效性。



(a) (b)

图 4 CGBS-SymNMF 算法在 6 个数据集上收敛性能

Fig. 4 Convergence behavior of CGBS-SymNMF algorithm on six datasets

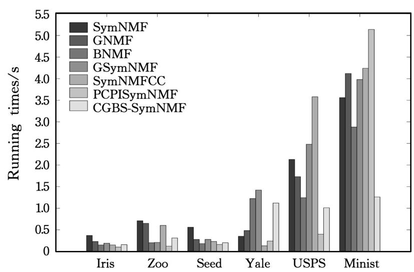


图 5 各算法运行时间比较

Fig. 5 Running times comparison

结束语 本文提出了基于约束图正则的块稀疏对称非负矩阵分解算法,能够利用成对约束有效地区分不同类簇的相近样本以及增强模型图学习能力。该算法建立在 SymNMF 算法基础上,引入了 PCP-SDP 模型,将样本通过图结构映射到一个光滑面上,使得相近样本传播约束标签,加强算法对样本的分辨能力;同时构建了约束图正则,进一步增强算法的图学习能力,最终使用块稀疏正则项提高模型抗噪能力,并且加强所分解矩阵 $\boldsymbol{HH}^{\mathrm{T}}$ 的块对角表示能力,提高模型聚类精确度。但是提出的算法在 Mnist 和 USPS 这类数据集上的聚类精确度不够高,对稀疏数据类的聚类效果有待加强,在这一方面,CGBS-SymNMF 算法还有进一步提升的空间。

参 考 文 献

[1] LEE D D, SEUNG H S. Learning the parts of objects by non-negative matrix factorization[J]. Nature, 1999, 401(6755): 788-791.

[2] SHAHBAZI Z, BYUN Y C. Topic modeling in short-text using non-negative matrix factorization based on deep reinforcement learning[J]. Journal of Intelligent & Fuzzy Systems, 2020, 39(1): 753-770.

[3] LAXMI LYDIA E, KRISHNA KUMAR P, SHANKAR K, et al. Charismatic document clustering through novel K-Means non-negative matrix factorization (KNMF) algorithm using key phrase extraction[J]. International Journal of Parallel Programming, 2020, 48(3): 496-514.

[4] HASSANI A, IRANMANESH A, MANSOURI N. Text mining using nonnegative matrix factorization and latent semantic analysis[J]. Neural Computing and Applications, 2021, 33(20): 13745-13766.

[5] SHAHBAZI Z, BYUN Y C. Topic modeling in short-text using non-negative matrix factorization based on deep reinforcement learning[J]. Journal of Intelligent & Fuzzy Systems, 2020, 39(1): 753-770.

[6] MCGRAW T, KANG J, HERRING D. Sparse non-negative matrix factorization for mesh segmentation[J]. International Journal of Image and Graphics, 2016, 16(1): 1650004.

[7] LIU H, WU J, SUN J Y, et al. A robust segmentation method with triple-factor non-negative matrix factorization for myocardial blood flow quantification from dynamic 82Rb positron emission tomography[J]. Medical Physics, 2019, 46(11): 5002-5013.

[8] TIAN L, DU Q, KOPRIVA I, et al. Orthogonal graph-regularized non-negative matrix factorization for hyperspectral image clustering[C]// IEEE International Geoscience and Remote Sensing Symposium. Piscataway: IEEE, 2019: 795-798.

[9] YANG Z, ZHANG Y, XIANG Y, et al. Non-negative matrix factorization with dual constraints for image clustering[J]. IEEE Transactions on Systems, Man, and Cybernetics: Systems, 2018, 50(7): 2524-2533.

[10] CAI D, HE X, HAN J, et al. Graph regularized nonnegative matrix factorization for data representation[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2010, 33(8): 1548-1560.

[11] ZHOU C, LI X L, LI Q L, et al. Sparse Non-negative Matrix Factorization Algorithm Based on Cosine Similarity[J]. Computer Science, 2020, 47(10): 108-113.

[12] HE Z, XIE S, ZDUNEK R, et al. Symmetric nonnegative matrix factorization: Algorithms and applications to probabilistic clustering[J]. IEEE Transactions on Neural Networks, 2011, 22(12): 2117-2131.

[13] DOBROVOLSKYI H, KEBERLE N, TERNOVYY Y. Sparse symmetric nonnegative matrix factorization applied to face recognition[C]// IEEE International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications (IDAACS). Piscataway: IEEE, 2017: 1042-1045.

[14] GAO Z, GUAN N, SU L. Graph Regularized Symmetric Non-Negative Matrix Factorization for Graph Clustering[C]// 2018 IEEE International Conference on Data Mining Workshops (ICDMW). Piscataway: IEEE, 2018: 379-384.

[15] YANG L, CAO X, JIN D, et al. A unified semi-supervised community detection framework using latent space graph regularization[J]. IEEETransactions on Cybernetics, 2014, 45(11): 2585-2598.

[16] ZHANG X,ZONG L,LIU X,et al. Constrained clustering with nonnegative matrix factorization[J]. IEEETransactions on Neural Networks and Learning Systems,2015,27(7):1514-1526.

[17] JIA Y,LIU H,HOU J,et al. Semisupervised adaptive symmetric non-negative matrix factorization[J]. IEEE Transactions on Cybernetics,2020,51(5):2550-2562.

[18] WU W,JIA Y,KWONG S,et al. Pairwise constraint propagation-induced symmetric nonnegative matrix factorization[J]. IEEE Transactions on Neural Networks and Learning Systems,2018,29(12):6348-6361.

[19] PENG C,ZHANG Z,KANG Z,et al. Nonnegative matrix factorization with local similarity learning[J]. Information Sciences,2021,562(8):325-346.

[20] QIN Y,WU H,FENG G. Structured subspace learning-induced symmetric nonnegative matrix factorization[J]. Signal Processing,2021,186(12):108115.

[21] PARK S,ZHAO H. Spectral clustering based on learning similarity matrix[J]. Bioinformatics,2018,34(12):2069-2076.

[22] QIN Y,WU H,FENG G. Structured subspace learning-induced symmetric nonnegative matrix factorization[J]. Signal Processing,2021,186(12):108115.

[23] WANG S,LI G,HU G,et al. Community detection in dynamic networks using constraint non-negative matrix factorization[J]. Intelligent Data Analysis,2020,24(1):119-139.

[24] JIA Y,LIU H,HOU J,et al. Pairwise constraint propagation with dual adversarial manifold regularization[J]. IEEE Transactions on Neural Networks and Learning Systems,2020,31(12):5575-5587.

[25] JIA Y,WU W,WANG R,et al. Joint optimization for pairwise constraint propagation[J]. IEEE Transactions on Neural Networks and Learning Systems,2020,32(7):3168-3180.

[26] BAI L,LIANG J Y,CAO F. Semi-supervised clustering with constraints of different types from multiple information sources[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence,2020,43(9):3247-3258.

[27] LI Z,LIU J,TANG X. Pairwise constraint propagation by semi-definite programming for semi-supervised classification[C]// Proceedings of the 25th International Conference on Machine Learning. Berlin:Springer,2008:576-583.

[28] QIN Y,DING S F. Survey of Semi-supervised Clustering[J]. Computer Science,2019,46(9):15-21.

[29] SPIELMAN D A. Spectral graph theory and its applications[C]// 48th Annual IEEE Symposium on Foundations of Computer Science(FOCS'07). Piscataway:IEEE,2007:29-38.

[30] LEMON A,SO A M C,YE Y. Low-rank semidefinite programming:Theory and applications[M]. Boston:Now Publishers,2016:6-17.

[31] HANG M, YUAN J L, PANG L P,et al. UV-theory of a Class of Semidefinite Programming and Its Applications[J]. Acta Mathematicae Applicatae Sinica,2021,37(4):717-737.

[32] WONG W K,YIN Y,ZHOU J. Using SeDuMi to find various optimal designs for regression models[J]. Statistical Papers,2019,60(5):1583-1603.

[33] LU C,FENG J,LIN Z,et al. Subspace clustering by block diagonal representation[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence,2018,41(2):487-501.

[34] JIA Y,HOU J,KWONG S. Constrained clustering with dissimilarity propagation-guided graph-Laplacian PCA[J]. IEEE Transactions on Neural Networks and Learning Systems,2020,32(9):3985-3997.

[35] HASSAN S N H B,NIIMI T,YAMASHITA N. Augmented lagrangian method with alternating constraints for nonlinear optimization problems[J]. Journal of Optimization Theory and Applications,2019,181(3):883-904.

[36] JIA Y,KWONG S,HOU J,et al. Semi-supervised non-negative matrix factorization with dissimilarity and similarity regularization[J]. IEEE Transactions on Neural Networks and Learning Systems,2019,31(7):2510-2521.

[37] LI Z,TANG J,HE X. Robust structured nonnegative matrix factorization for image representation[J]. IEEE Transactions on Neural Networks and Learning Systems,2018,29(5):1947-1960.

LIU Wei, born in 1996, postgraduate. His main research interests include machine learning and data mining.

DENG Xiuqin, born in 1966, professor, master supervisor. Her main research interests include machine learning and data mining.

(责任编辑:喻藜)