# 图像表示的约束非负矩阵分解

Constrained Nonnegative Matrix Factorization for Image Representation

Haifeng Liu, Member, IEEE, Zhaohui Wu, Senior Member, IEEE, Xuelong Li, Senior Member, IEEE,Deng Cai, Member, IEEE, and Thomas S. Huang, Life Fellow, IEEE

## 摘要

非负矩阵分解(NMF)是一种流行的技术，用于寻找基于零件的非负数据的线性表示。它已成功地应用于模式识别、信息检索和计算机视觉等广泛的应用领域。然而，NMF本质上是一种非监督方法，不能利用标签信息。在本文中，我们提出了一种新的半监督矩阵分解方法，称为约束非负矩阵分解(CNMF)，该方法将标签信息作为附加约束。具体来说，我们将展示如何显式地组合标签信息来提高所得到的矩阵分解的鉴别能力。对提出的CNMF方法进行了研究，并给出了优化问题的更新方案。通过一组基于真实世界应用的评估，经验实验证明了我们的新算法的有效性，与最先进的方法相比。

关键词：非负矩阵分解，半监督学习，降维，聚类。

## 1介绍

在许多数据分析任务中，一个基本问题是找到合适的数据[1]、[2]、[3]、[4]、[5]、[6]、[7]、[8]的表示。一个有用的表示通常使数据中的潜在结构显式，从而可以应用进一步的处理。矩阵分解技术作为此类数据表示的基本工具，已经受到越来越多的关注。通过使用不同的标准，已经发展出若干不同的方法来做到这一点。最流行的技术包括主成分分析(PCA)[9]、奇异值分解(SVD)[10]和矢量量化[11]。矩阵分解的核心是找到两个或两个以上的矩阵因子，它们的乘积是原始矩阵的一个很好的近似。在实际应用中，分解后的矩阵因子的维数通常比原矩阵的维数小得多。这使得数据点的表示更加紧凑，这可以促进其他学习任务，如聚类和分类。

在矩阵分解方法中，非负矩阵分解(Nonnegative matrix factorization，NMF)[2]，[3]的特点是它强制了因子矩阵必须是非负的约束，即所有元素必须等于或大于零。这种非负性约束将NMF引入到基于部件的对象表示中，因为它只允许对原始数据进行加法(而不是减法)组合。因此，对于图像处理、人脸识别[2]、[12]和文档聚类[13]、[14]来说，它是一种理想的降维算法。在这些降维算法中，很自然地会将物体看作是部分的组合而形成一个整体。

NMF是一种无监督学习算法。也就是说，NMF不适用于许多现实世界的问题，在这些问题上，来自领域专家的知识有限。然而，许多机器学习研究人员发现，当未标记数据与少量标记数据结合使用时，可以在学习准确度[15]，[16]，[17]上产生相当大的提高。与标记过程相关的成本可能使一个完全标记的训练集不可行，而获取一个小的标记数据集是相对便宜的。在这种情况下，半监督学习可以有很大的实用价值。因此，将NMF的使用扩展到半监督的方式将会有很大的好处。

最近，Cai[1]提出了一种图形正则化NMF(GNMF)方法来编码数据空间的几何信息。GNMF构造一个最近邻图来建模局部流形结构。当标签信息可用时，可以很自然地将其合并到图结构中。具体来说，如果两个数据点共享相同的标签，则可以为连接它们的边分配一个较大的权重。如果两个数据点有不同的标签，则相应的权值设为0。这就产生了半监督GNMF。这种方法的主要缺点是，在理论上无法保证来自同一类的数据点会在新的表示空间中映射到一起，并且仍然不清楚如何以一种有原则的方式选择权重。

本文提出了一种新的矩阵分解方法——约束非负矩阵分解(Constrainednon-negative matrix Factorization，CNMF)，该方法将标签信息作为附加的硬约束。我们的方法的中心思想是，来自同一个类的数据点应该合并到新的表示空间中。这样，得到的基于零件的表示与原始数据具有一致的标签，从而具有更强的鉴别能力。我们的方法的另一个优点是它是无参数的，这避免了调优参数以获得最佳结果的成本。它使我们的算法适用于许多现实世界的应用容易和有效。讨论了如何有效地解决相应的优化问题。并给出了优化方案的收敛性证明。本文的贡献有：

1、标准的NMF是一种不能包含标签信息的无监督学习算法。在本文中，我们将其推广到半监督学习算法中。此外，我们的方法将标签信息作为硬约束；因此，共享相同标签的数据点在新的表示空间中具有相同的坐标。通过这种方式，学习到的表征可以有更强的辨别能力。

2、早期的研究[18]表明，NMF和PLSA都是多项PCA的实例。特别是，PLSA解决了具有KL散度[19]，[20]的NMF问题。为了进一步探讨这方面的问题，我们将CNMF应用于KL发散公式，并提供了求解优化问题的更新规则。

3、与半监督GNMF不同，我们的方法的一个优点是它是无参数的。因此，为了获得最佳结果，不需要对参数进行调优。因此，CNMF可以轻松高效地应用于许多实际应用。实验结果表明，该算法能显著提高聚类性能。

4、 据我们所知，目前还没有直接得到NMF问题解的方法；最先进的算法是使用更新规则迭代地得到目标函数的最优值。因此，算法的效率对于实际应用非常重要。在本文中，我们定性地分析了算法的计算复杂性，并对算法的收敛速度进行了实验测试，以定量地证明算法的有效性。

本文的结构如下：在第二部分，我们简要回顾了NMF的背景和相关工作。第3节介绍了约束NMF的思想。第4节和第5节给出了两种公式中算法的详细算法和算法收敛性的理论证明。第6节讨论算法的计算复杂度。最后，第7节给出了实验结果，第8节总结了本文。

## 2相关工作

矩阵的因式分解通常是非唯一的，通过合并不同的约束，已经发展出许多不同的方法来实现这一点。PCA[9]和SVD[10]将矩阵分解为主成分的线性组合。Hoyer[21]和Dueck等人[22]计算稀疏矩阵分解。

NMF与这些方法的不同之处在于，它强制约束因子矩阵的元素必须是非负的。假设有n个数据点。每个数据点是m维的，用一个向量表示。向量放在列中，整个数据集由矩阵表示。NMF的目标是找出两个非负矩阵因子U和V，其中这两个因子的乘积是原始矩阵的近似，表示为。

该近似由一个成本函数来量化，该函数可以由一些距离测度来构造。一个简单的度量是两个矩阵[23]之间的欧几里德距离(也称为弗罗本尼乌斯范数)的平方。然后，NMF的目标可以重述如下：将X分解为矩阵U和矩阵VT，从而最小化以下目标函数：。[24]中描述的另一个度量称为X与Y的“散度”：



这里。这个度量不是对称的。X到Y的距离不一定等于Y到X的距离。当和时，它减小为Kullback-Leibler散度，即相对熵。

这两个目标函数在变量U和v上都不是凸的，因此，很难找到或的全局极小值。Lee和Seung提出了一种迭代更新算法[24]来寻找上述优化问题的局部最优解。

NMF分解，每列向量U，ui，可以被视为一个基础和每个数据点ξ是近似的线性组合这些k基地，加权的组件v。换句话说，NMF习每个数据映射从m维空间k维空间。新的表示空间由k个基的ui生成。在真实的应用程序中，如图像处理[2]，人脸识别[12]，[25]，和文档聚类[13]，[14]，我们通常设置k<<m和k<<n。然后，高维数据可以表示为一组低维向量，希望基向量可以发现潜在语义结构的数据集。与PCA、LDA、LPP[4]等降维算法不同的是，U和V上的非负约束只允许基向量的可加性组合，这也是NMF被认为是基于零件的表示的原因。

Ding等人[26]提出了一种半非负矩阵分解算法，该算法只限制一个矩阵因子包含非负项，同时松弛了对基向量的约束。

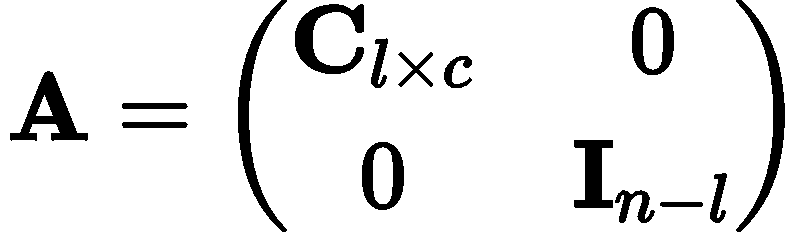
除了用乘法更新方法寻找矩阵因子外，还有人提出了梯度下降算法[28]来解决优化问题。该算法也被称为交替非负最小二乘或“投影梯度”。Lin[29]表明项目梯度法比乘法更新法收敛更快。Heiler等人[30]推导了基于序列二次和二阶规划的NMF优化方案。这种方法的一个关键优点是，NMF可以通过在相同的优化框架中以额外约束的形式合并先验知识来扩展。附加的约束是，对于每个类以及属于该类的每个标记点，将其系数限制为围绕类中心的圆锥。这一思想与我们的论文不同，我们的论文将标签信息作为附加的硬约束，要求来自同一类的数据点在新的表示空间中合并在一起。

据我们所知，大多数现有的NMF变体和扩展集中在矩阵分解上，而没有将标签信息作为硬约束考虑在内。本文提出了一种新的半监督矩阵分解方法，该方法将标签信息作为附加约束。因此，数据点的新表示形式具有更强的鉴别能力。

## 3半监督NMF

NMF是一种无监督学习算法。它不能直接应用于标签信息可用的情况。在本节中，我们介绍了一种新的矩阵分解方法，称为约束非负矩阵分解，它将标签信息作为附加的约束。该方法可以保证共享相同标签的数据点在低维空间中映射到相同的类中。

考虑一个由n个数据点组成的数据集，其中前l个数据点x1的标签信息可用和其余n-l数据点没有标记。假设有c类。每个数据点从被标记为一个类。我们首先构建指标矩阵c，其中；如果xi被标记为第j类，则。利用指标矩阵C，定义一个标签约束矩阵a，如下所示：



其中是一个单位矩阵。以n个数据点为例，其中x1、x2标记为I类，x3、x4标记为II类，x5标记为III类，另外n-5个数据点未标记。基于本例的标签约束矩阵A表示为：



回想一下，NMF将每个数据点xi映射到vi，从m维空间映射到k维空间。为了合并标签信息，我们可以通过引入一个辅助矩阵Z来施加标签约束：V=AZ。

从上面的方程，很容易检查，如果xi和xj有相同的标签，那么vi=vj。在标签约束下，通过寻找两个非负矩阵因子U和Z，将原始的NMF扩展为半监督学习算法(CNMF)，其中因子U、Z和a的乘积是原始矩阵的近似表示为。

## 4最小化F范数代价的算法

### 4.1更新算法

使用Frobenius范数作为代价函数，我们的带有标签约束的CNMF算法减少，以最小化以下目标函数：。约束为ui，j和zi，j是非负的。

(1)中CNMF的目标函数在变量U和Z上都不是凸的，因此寻找的全局极小值是不现实的。在下文中，我们描述了一种迭代更新算法来获得的局部最优值。

使用矩阵属性，可以将目标函数重写为：



让和是约束的拉格朗日乘子uij>0和zij>0，分别和。拉格朗日函数是。

要求L对U和Z的导数为零，我们得到：



利用Kuhn-Tucker条件和，我们得到以下关于和的方程：



这些方程得到以下更新规则：



对于上述迭代更新规则，有如下定理：

定理1：在(2)和(3)的更新规则下，(1)中的目标函数是不增长的。当且仅当U和Z处于平稳点时，目标函数在这些更新下是不变的。

定理1保证(2)和(3)的迭代收敛，因此最终的解将是一个局部最优解。下面，我们将给出定理1的证明。

### 4.2收敛性的证明

为了证明定理1，我们使用了一个辅助函数的如下性质，正如期望最大化算法[31]，[32]所使用的。

引理2。如果存在辅助函数G，满足条件和，则F不增加在更新。

等式仅当是。的局部最小值时成立。通过迭代(4)中的更新，估计序列将收敛到局部最小。我们将通过为(1)中的目标函数定义一个适当的辅助函数来说明这一点。

首先，我们证明(3)中更新规则的收敛性。对于Z中的任意元素zab，设Fzab表示中与zab相关的部分。由于更新本质上是元素的，因此可以充分证明在(3)的更新步骤下，每个Fzab都是不递增的。我们通过定义关于zab的辅助函数来证明这一点：

引理3。设f0表示对应于z的一阶导数，函数



是Fzab的辅助函数，是中仅与zab相关的部分。

证明。显然，。根据辅助函数的定义，我们只需证明。为了做到这一点，我们比较在(5)中的Fzab的泰勒级数展开：。

是对z的二阶导数，很容易检验



将(7)代入(6)并与(5)比较，我们可以看到，不是显示，就等于证明。

为了证明上述不等式，我们有



接下来，我们为(2)中的更新规则定义一个辅助函数。同样，让Fuab表示中与uab相关的部分。那么，关于uab的辅助函数定义如下：

引理4。函数是Fuab的辅助函数，是中仅与uab相关的部分。

引理4的证明与引理3的证明在本质上是相似的，由于篇幅的限制，此处略去。利用上面的引理，现在我们给出定理1的证明。

定理1的证明。将(5)变成(4)，我们得到



根据引理3，由于(5)是一个辅助函数，因此在此更新规则下Fzab是非递增的。

虽然对于每一个更新步骤，KL发散公式中CNMF的代价比NMF略高一些，但由于收敛的迭代次数不同，CNMFKL的整体算法复杂度可能并不会更慢。我们将在下一节实验研究这些算法的收敛速度。假设NMF、CNMF和CNMFKL的相乘更新分别在t1、t2和t3迭代后停止，则这些算法的总体计算复杂度为，和。

## 7实验结果

在本节中，我们将研究如何使用我们提出的CNMF算法进行数据聚类。通过实验验证了该算法在图像聚类中的有效性。

## 7.1评价指标

我们使用两个指标来评估聚类性能[1]，[13]。通过比较每个样本的聚类标签与数据集提供的标签来评估结果。一个度量是准确度(AC)，它用于测量获得的正确标签的百分比。给定一个包含n幅图像的数据集，对于每幅样本图像，设li为我们通过不同算法得到的聚类标签，ri为数据集提供的标签。精度定义为。

在是如果x=y等于1，否则等于0的delta函数，map(li)是映射函数，将每个集群标签li映射到数据集中的等价标签。利用Kuhn-Munkres算法[34]可以找到最佳映射。

第二个度量是标准化互信息(MIc)。在聚类应用程序中，互信息用于度量两组聚类的相似程度。给定两组图像聚类C和C’，它们的互信息度量定义为。

式中，表示任意从数据集中选取的图像分别属于和簇的概率，表示任意选取的图像同时属于和簇的概率。取值范围为0~，为C的熵值，为C0的熵值。当两组图像簇相同时，达到最大，当两组图像簇完全独立时，则为零。的一个重要特征；是该值对于所有类型的排列保持相同。在我们的实验中，我们使用归一化度量，取值范围在0到1之间：



## 7.2评估与比较

为了展示数据聚类的性能，我们在四个数据集上比较了我们的算法和其他相关方法。我们评估的算法如下：

我们提出的约束非负矩阵分解算法最小化f范数代价。我们提出的约束非负矩阵分解算法最小化KL发散代价(CNMFKL)。基于非负矩阵分解的聚类。我们实现了一个标准化的NMF削减加权版本，如[13]所建议的。非负张量分解(NTF)NTF是NMF对张量数据的扩展。在NTF中，每个人脸图像被表示为一个二阶张量，而不是一个矢量。图正则化非负矩阵分解[1]，将数据空间的几何信息编码为矩阵分解。流形上的半监督图正则化非负矩阵分解(SemiGNMF)[1]。该方法通过修改权值矩阵，将标签信息整合到图结构中。

分解聚类(CF)概念[36]。我们评估了在四个图像数据集上的聚类性能。这些数据集包含许多类别的图像。表3总结了这些数据集的重要统计数据。稍后我们将分别描述数据集的细节。对于每个数据集，用2到10个不同数量的聚类进行评价。对于固定的聚类数k，我们进行如下实验：

1。我们从数据集中随机选取k个类别，将这k个类别的图像混合作为集合X进行聚类。对于半监督算法(CNMF、CNMFKL和SemiGNMF)，我们随机从X中的每个类别中选取10%的图像，并使用它们的类别号作为可用的标签信息。例外情况是ORL数据库。ORL中每个类别只有10张图片，10%的图片只有一张。对于CNMF来说，一个标签是没有意义的，因为该算法将具有相同标签的图像映射到同一点上。因此，对于ORL，我们从每个类别中随机选择两幅图像来提供标签信息。

2。在聚类集合上，我们应用上述不同的矩阵分解算法来获得新的数据表示v。我们将新空间的维数设置为与聚类数量相同(光谱聚类中也使用了同样的技术)。因此，这一步将数据从原始空间映射到低维(k维)空间。

3。然后将K-means应用到新的数据表示V中进行图像聚类。用不同的初始点重复K-means20次，得到K-means代价函数的最佳结果。

4。我们将得到的聚类与原始图像分类进行比较，计算精度和归一化互信息。

以上过程重复10次，记录平均的聚类性能作为最终结果。正如我们前面提到的，在我们的方法中没有参数。对于其他算法，参数设置为每个算法能够达到其最佳结果的值。

### 7.2.1ORL数据库

AT&TORL数据库1包含40个不同的受试者的10张不同的图像，因此总共有400张图像。对于一些受试者，图像是在不同的时间拍摄的，光线变化，面部表情(睁眼/闭眼，微笑/不微笑)，面部细节(戴眼镜/不戴眼镜)。所有的图像都是在深色的均匀背景下拍摄的，受试者处于直立的正面位置。

在所有的实验中，图像都进行了预处理，以确定人脸的位置。原始图像首先在尺度和方向上进行归一化，使两只眼睛在同一位置对齐。然后，将面部区域裁剪成最终的图像进行聚类。每个图像是32x32像素，每个像素256个灰度级别。

图1显示了精度和归一化互信息与簇数的关系。我们提出的CNMF和CNMFKL算法始终优于所有其他算法。CF、NMF、GNMF和SemiGNMF的性能相当。在该数据集中，CNMF的性能最好，CNMFKL的性能次之。

我们还从ORL数据库中随机选择了25名受试者，每个受试者有10张图片，并对总共250张图片运行算法。图2显示了效果。第一幅图像包含了25个主题，另外三幅是由NMF、CNMF和CNMFKL获得的基向量。

### 7.2.2Yale数据库

耶鲁数据库2包含15个个体的165张灰度图像。每个受试者有11张图片，每种不同的面部表情或配置都有一张：中心光、带眼镜、快乐、左光、不带眼镜、正常、右光、悲伤、困倦、惊讶和眨眼。我们对这个数据集进行与ORL数据集相同的预处理。因此，每个图像也由图像空间中的1，024维向量表示。

聚类结果如表6所示。图示如图4所示。我们可以看到，NMF、GNMF和SemiGNMF有相似的结果。虽然SemiGNMF利用了标签信息，但与NMF和GNMF相比，SemiGNMF并没有表现出多大的优势。相比之下，CNMF和CNMFKL在所有情况下都大大优于上述三种方法，其中CNMFKL的性能最好。从表6可以看出，与我们提出的CNMF算法之外的最佳算法SemiGNMF相比，CNMFKL提高了7.46%，CNMF提高了4.41%的准确率。对于规范化的互信息，CNMFKL和CNMF分别提高了8.38%和4.81%。

### 7.2.3Caltech-101数据库

Caltech-101是由加州理工大学创建的数字图像数据集，包含101个对象类别。每个类别包含大约40到800张图片。每个图像的大小大约是300x200像素。在我们的实验中，我们选择了10个最大的类别，除了BACKGROUND\_GOOGLE类别。总的来说，我们测试的子集包含3，044张图像。我们对每幅图像进行与Corel数据集相同的预处理。我们首先提取SIFT描述符，然后生成码字作为每幅图像的特征。对于Caltech-101数据集，SIFT描述符的数量为555，292，我们还生成了500个码字。最后，Caltech-101中的每幅图像都由一个500维频率直方图表示。

聚类结果如表7所示，图5所示。在这个数据集上，CNMFKL仍然显示出最好的性能。CNMF在归一化互信息测量和大多数精度测量情况下都是第二好的。总的来说，我们的方法在聚类中显示了更好的有效性。如表7所示，与SemiGNMF相比，CNMFKL的准确度提高了7.2%。对于标准化互信息，CNMFKL提高了9.01%，CNMF提高了2.87%。

## 7.3收敛性研究

我们使用迭代更新规则来获得CNMF目标函数的局部最优值，无论成本度量是在Frobenius范数还是KL发散。在前几节中，我们已经证明了更新规则的收敛性，并分析了计算复杂性。本文通过实验证明了算法的收敛速度。

我们比较了原始NMF算法、我们的CNMF最小化f范数代价(CNMF)和CNMF最小化KL发散代价(CNMFKL)的收敛速度。图6为三种算法在四个图像数据库上的收敛速度。对于每个图，x轴是迭代次数，y轴是目标函数的值。我们可以看到，CNMF和CNMFKL算法收敛速度都非常快。对于ORL、Corel和Caltech-101数据库，所有三种算法在20次迭代中收敛。对于耶鲁的数据库，它们在100次迭代中收敛。特别是，我们注意到，对于Yale数据库和Corel数据库，CNMFKL在不到10次迭代的时间内收敛，展示了非凡的性能。这也验证了我们在第6节中所做的陈述，即尽管CNMFKL在每个更新步骤中需要更多的操作，但整个过程的总体成本并不高，因为它可能比NMF收敛得更快。

## 8结论

在本文中，我们提出了一种新的矩阵分解方法，称为约束非负矩阵分解，它利用有标记和无标记的数据点。CNMF将标签信息作为硬约束强加于目标函数。这样，数据点的新表示形式可以具有更强的鉴别能力。我们在两个公式中展示了CNMF方法，并提出了两个优化问题的更新算法。在四个标准图像数据库上的实验结果证明了该方法的有效性。CNMF和CNMFKL算法都优于其他现有算法。特别是在所有情况下，kl-散度形式的CNMF算法都比其他算法具有明显的优势。此外，我们的算法是无参数的。因此，我们提出的CNMF可以很容易地应用于广泛的实际问题。