

DOI: 10.11992/tis.201904058

网络出版地址: <http://kns.cnki.net/kcms/detail/23.1538.TP.20190829.1513.006.html>

低秩分块矩阵的核近似

王中元, 刘惊雷

(烟台大学 计算机与控制工程学院, 山东 烟台 264005)

摘要: 为了探讨结构受限下的矩阵分解问题, 通过最小化块外对角线来增强类与类之间数据表示的不相关性, 从而实现分块约束, 即数据来源于不同的聚类结构, 是一种局部结构的约束; 同时通过增强样本的自表达属性并缩小样本之间的差距来增强类内数据表示的相关性, 从而实现低秩约束, 即数据行出现冗余, 是一种全局结构的约束。随后设计了一个低秩分块矩阵的核近似算法, 通过交替方向乘子法迭代求解。最后将该方法分别在人脸识别和字符识别上进行测试。实验结果表明, 所提出的低秩分块矩阵分解算法在收敛速度和近似精度上都具有一定的优势。

关键词: 低秩近似; 块对角矩阵; 稀疏矩阵; 核近似; 矩阵分解; 交替向量乘子法; 子空间聚类; 图像识别

中图分类号: TP181 **文献标志码:** A **文章编号:** 1673-4785(2019)06-1209-08

中文引用格式: 王中元, 刘惊雷. 低秩分块矩阵的核近似 [J]. 智能系统学报, 2019, 14(6): 1209-1216.

英文引用格式: WANG Zhongyuan, LIU Jinglei. Kernel approximation of a low-rank block matrix[J]. CAAI transactions on intelligent systems, 2019, 14(6): 1209-1216.

Kernel approximation of a low-rank block matrix

WANG Zhongyuan, LIU Jinglei

(School of Computer and Control Engineering, Yantai University, Yantai 264005, China)

Abstract: In order to explore the matrix decomposition problem under structural constraints, irrelevance of data representation between classes was enhanced in this paper by minimizing the diagonal outside the block, thus realizing the block constraint, i.e., the data is derived from different cluster structures. It is a local structure constraint. At the same time, by enhancing the self-expressing property of the sample and narrowing the gap between samples, the correlation of the data representation in the class was enhanced, thereby realizing the low-rank constraint, i.e., the redundancy of the data row was a constraint of the global structure, thereby realizing the low-rank constraint. A kernel approximation algorithm for low-rank block matrix was then designed and solved iteratively by alternating the direction method of multipliers (ADMM). Finally, the method was tested on face recognition and character recognition. Experimental results showed that the proposed low-rank block matrix decomposition algorithm has certain advantages in solving speed and approximate accuracy.

Keywords: low-rank approximation; block diagonal matrix; sparse matrix; kernel approximation; matrix factorization; alternating direction method of multipliers (ADMM); subspace clustering; image identification

近年来, 随着互联网的快速发展, 人们获取数据以及存储数据的能力都迅速提高。不论在科学技术还是在日常生活的各个领域都累积了大量数据。怎样才能快速有效地对这些数据进行分析处

理并发掘其中蕴含的有效信息引起人们的关注。大部分机器学习都能够通过矩阵的形式表示, 但是在现实生活中经常会使用到数以百万计的样本, 通过对矩阵进行处理的机器学习技术的复杂度会随着应用规模的增加呈二次方增长, 这会使很多问题无法解决。

目前通常通过低秩矩阵分解、核方法和聚类等方法来解决矩阵分解的问题^[1]。聚类是将输入

收稿日期: 2019-04-24. 网络出版日期: 2019-08-30.

基金项目: 国家自然科学基金项目 (61572419, 61773331, 61703360);
山东省高等学校科技计划 (J17KA091).

通信作者: 刘惊雷. E-mail: jinglei_liu@sina.com.

数据按照某种相似性规则划分为若干类。低秩矩阵分解是将数据矩阵分解成噪声和稀疏低秩两部分。在众多方法之中,通过核方法近似低秩分块矩阵由于其模型合理直观、实施简单、效果显著而受到关注。

本文考虑结构受限下的矩阵分解问题^[2],其中结构受限主要是低秩结构受限和分块结构受限。本文通过消除噪声并最小化块外对角线来增强类与类之间数据表示的不相关性,从而实现分块约束^[3];同时通过增强训练样本的自表达属性并缩小样本之间的差距来增强类内数据表示的相关性,从而实现低秩约束^[4],最后迭代优化求解一系列子问题来实现矩阵分解的目的。随后设计了一个低秩分块矩阵的核近似算法 (low-rank block kernel approximation, LBKA)。LBKA 不仅能够显著提高算法速度,而且由于低秩分块约束极大的消除了噪声的影响,使得算法近似精度也有了提高。

相较于传统的矩阵分解问题,本文设计了低秩分块矩阵的核近似算法。本文的特点和贡献如下:

1) 引入了一种低秩分块矩阵的框架,通过核方法将非线性问题转化为高维空间中的线性问题来解决,这样不仅可以提高算法的精度,还可以通过低秩约束来降低算法复杂度,提高了计算的收敛速度与计算精度。

2) 设计了一种低秩分块矩阵的核近似算法。该算法基于增广拉格朗日乘子法和交替方向乘子法构造迭代公式,依次更新系数矩阵 \mathbf{Z} 和辅助变量 \mathbf{J} 、 \mathbf{Q} 、 \mathbf{S} ,直到收敛到稳定的特征矩阵。

3) 在人脸识别数据集和字符识别数据集上进行了实验验证。实验结果表明,相较于传统的识别算法, LBKA 在收敛速度和近似精度上有所提高。同时通过核矩阵近似处理,极大地降低了噪声对实验的影响。

1 相关工作

本文主要在低秩、稀疏、对角矩阵的分解领域探讨有关低秩分块矩阵的近似问题。

随着非负矩阵分解^[4]等矩阵分解算法的提出,给数据处理领域带来了新的思路。低秩可以看作稀疏的特殊形式,低秩就是特征值稀疏;块对角也是,块对角矩阵自身就具有低秩和稀疏的特性。

现阶段,稀疏表示被广泛应用于信号处理^[3]、机器学习和计算机视觉^[5]等方面。随着基于稀疏表示的分类 (SRC)^[6] 在人脸识别中的成功应用,已经提出了许多基于 SRC 的优化算法。例如,

Nie 等人通过对损失函数和正则化项进行 l_{21} 范数约束,提出了一种有效的特征选择方法^[7]。

然而稀疏约束只能实现局部约束,而低秩约束却能够实现全局约束。鲁棒主成分分析 (robust principal component analysis, RPCA) 是最典型的低秩方法,最初是用来将损坏的观测数据恢复为具有低秩结构的数据。后来发现 RPCA 可以将一个子空间的数据分解成低秩和稀疏噪声两部分。但 RPCA 无法处理多个子空间的并集。为此提出了潜在的低秩表示 (LatLRR)^[8],通过利用子空间分割的低秩表示的方法解决多个子空间问题。

本节主要介绍一些基本符号并简单说明了两种典型的基于低秩表示的方法:鲁棒主成分分析 (RPCA)^[9] 和低秩表示 (LRR)^[10]。

1.1 基本符号

本文使用粗体大写字母表示矩阵,例如 \mathbf{A} 。用粗体小写字母表示矢量,例如 \mathbf{a} 。对于 \mathbf{A} ,其第 i 行被表示为行向量 $[\mathbf{A}]_i$,其第 j 列被表示为列向量 $[\mathbf{A}]_{:,j}$ 。将 a_{ij} 或 A_{ij} 表示为 \mathbf{A} 的第 (i,j) 项。将 $\text{Diag}(\mathbf{A})$ 作为对角矩阵,其对角线上的第 i 项是 a_{ii} 。将单位矩阵表示为 \mathbf{I} 。如果 \mathbf{A} 中的所有项都是非负的,则表示为 $\mathbf{A} \geq 0$ 。文中将使用一些范数,包括 l_0 范数 $\|\mathbf{A}\|_0$, l_1 范数 $\|\mathbf{A}\|_1 = \sum_{ij} |a_{ij}|$, Frobenius 范数 $\|\mathbf{A}\|_F = \sqrt{\sum_{ij} a_{ij}^2}$, l_{21} 范数 $\|\mathbf{A}\|_{21} = \sum_j \|[\mathbf{A}]_{:,j}\|$ 和核范数 $\|\mathbf{A}\|_*$ 。

1.2 低秩分块矩阵相关定义

定义 1 假设矩阵 $\mathbf{A} = [\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n] \in \mathbf{R}^{m \times n}$ 是一个包含 n 个数据点的列矩阵。其核矩阵定义为: $\mathbf{B}_{ij} = \mathbf{b}(\mathbf{a}_i, \mathbf{a}_j) = \langle \phi(\mathbf{a}_i), \phi(\mathbf{a}_j) \rangle$, $\forall i, j = 1, 2, \dots, n$, 其中 $\phi: \mathbf{a} \mapsto \phi(\mathbf{a})$ 是内核诱导的特征图。 n 个映射数据点的所有成对内积存储在核矩阵 $\mathbf{K} \in \mathbf{R}^{n \times n}$ 中。

由于核矩阵的内存为 $O(n^2)$,所以在实际应用中是使用数据数量的二次方来计算和存储的。例如核主成分分析算法需要计算特征值分解 (SVD),它的复杂度为 $O(n^3)$,并且要多次访问核矩阵 \mathbf{K} 。因此要想计算大规模的数据,必须考虑复杂度的影响。

图 1 是低秩矩阵分块的一个示例,对于任意一个矩阵 $\mathbf{A} \in \mathbf{R}^{m \times n}$,有

图 1 低秩矩阵分块示例

Fig. 1 An example of low-rank matrix partitioning

$$\mathbf{A} = \mathbf{H}\mathbf{H}^T$$

为了用低秩分块矩阵逼近该相似度矩阵并同时得到聚类结构, 可以分别近似这些对角块, 每个对角块表示一个聚类。

通过求矩阵秩的最小化来求解低秩矩阵。然而这很难, 可以用核近似方法将核范数最小化为

$$\mathbf{X}^* = \arg \min_{\mathbf{X}} f(\mathbf{X}) + \Gamma \|\mathbf{X}\|_*$$

式中: $\mathbf{X} \in \mathbf{R}^{m \times n}$, $\Gamma > 0$ 是正则化参数。

定义 2 (奇异值阈值 (SVT)) 要想求解核范数最小化问题 (NNM), 例如:

$$\mathbf{X}^* = \arg \min_{\mathbf{X}} \Gamma \|\mathbf{X}\|_* + \frac{1}{2} \|\mathbf{X} - \mathbf{A}\|_F^2$$

需要使用奇异值阈值处理算子 $S_\Gamma(\cdot)$ 给出的闭式解:

$$\mathbf{X}^* = S_\Gamma(\mathbf{A}) = \mathbf{U}_A S_\Gamma(\Sigma_A) \mathbf{V}_A^T$$

式中: $S_\Gamma(x) = \text{sgn}(x) \cdot \max(|x| - \Gamma, 0)$ 是软收缩算子; $\mathbf{U}_A \Sigma_A \mathbf{V}_A^T$ 是 \mathbf{A} 的 SVD 分解^[11]。

1.2.1 鲁棒的主成分分析 (RPCA)

假设 $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n] \in \mathbf{R}^{d \times n}$ 是由 n 个样本组成的数据矩阵, RPCA 的目标是从损坏的矩阵 \mathbf{X} 中确定低秩矩阵 \mathbf{X}_0 , 同时消除稀疏噪声 \mathbf{E} , 即 $\mathbf{X} = \mathbf{X}_0 + \mathbf{E}$ 。因此, RPCA 的目标函数为

$$\begin{aligned} \min_{\mathbf{X}_0, \mathbf{E}} & \text{rank}(\mathbf{X}_0) + \lambda \|\mathbf{E}\|_0 \\ \text{s.t. } & \mathbf{X} = \mathbf{X}_0 + \mathbf{E} \end{aligned} \quad (1)$$

式中 λ 是平衡参数; $\|\cdot\|_0$ 表示 l_0 范数。由于低秩函数的离散性质和最小化 l_0 范数十分困难, 目前通常分别使用核范数和 l_1 范数正则化代替低秩约束和 l_0 范数正则化。因此, 式 (1) 可以重新表述为

$$\begin{aligned} \min_{\mathbf{X}_0, \mathbf{E}} & \|\mathbf{X}_0\|_* + \lambda \|\mathbf{E}\|_1 \\ \text{s.t. } & \mathbf{X} = \mathbf{X}_0 + \mathbf{E} \end{aligned} \quad (2)$$

其中 $\|\cdot\|_*$ 和 $\|\cdot\|_1$ 表示核范数和 l_1 范数。式 (2) 可以使用增广拉格朗日乘子法 (ALM)^[12] 来计算。

1.2.2 低秩表示 (LRR)

RPCA 是基于先验假设, 即数据大多是从低秩子空间中提取的, 可以用单个子空间来描述。然而真实数据集很难用单个子空间来描述。为此, LRR 假设每个数据可以由多个线性低秩子空间的并集近似表示。由此可以得到 LRR 的目标函数为

$$\begin{aligned} \min_{\mathbf{Z}, \mathbf{E}} & \text{rank}(\mathbf{Z}) + \lambda \|\mathbf{E}\|_1 \\ \text{s.t. } & \mathbf{X} = \mathbf{D}\mathbf{Z} + \mathbf{E} \end{aligned} \quad (3)$$

其中 \mathbf{D} 和 λ 分别是字典和平衡参数。 $\|\cdot\|_1$ 表示不同范数的约束。与 RPCA 类似, 式 (3) 可以重新表述为

$$\begin{aligned} \min_{\mathbf{Z}, \mathbf{E}} & \|\mathbf{Z}\|_* + \lambda \|\mathbf{E}\|_1 \\ \text{s.t. } & \mathbf{X} = \mathbf{D}\mathbf{Z} + \mathbf{E} \end{aligned} \quad (4)$$

1.3 块对角矩阵的核近似

定义 3 (自表达属性) 来自多个子空间的并

集的每个数据实例可以通过其他数据实例的线性组合来表示, 这种性质被称为自表达属性^[13]。即

$$\mathbf{X} = \mathbf{X}\mathbf{Z}$$

在存在自表达属性的情况下, 数据集中的每个数据点可以由来自其子空间的几个点的线性组合来表示^[14]。所以, 自表达属性应该是块对角的。

本文令类与类之间块对角线分量尽可能小, 同时提高相关的类内表示, 将其结构化形式表示为

$$\begin{aligned} \min_{\mathbf{Z}} & \lambda_1 \|\mathbf{A} \odot \mathbf{Z}\|_F^2 + \lambda_2 \|\mathbf{D} \odot \mathbf{Z}\|_0 \\ \text{s.t. } & \mathbf{X} = \mathbf{X}\mathbf{Z} \end{aligned} \quad (5)$$

其中 λ_1 和 λ_2 是用于衡量相应的项的正数, \odot 表示 Hadamard 乘积, 并且 $\mathbf{X} \in \mathbf{R}^{d \times n}$ 。具体地说, 第 1 项是为了最小化类与类之间块对角线分量, 并且 $\mathbf{A} = \mathbf{1}_n \mathbf{1}_n^T - \mathbf{Y}$ 。第 2 项是构造的子空间度量, 以提高相关的类内表示。 d_{ij} 是度量 x_i 和 x_j 之间的距离。本文将两个样本之间的距离定义为欧几里德距离的平方, 即 $\|x_i - x_j\|_2^2$ 。由于 l_0 范数最小化问题是 NP 难问题, 所以第 2 项可以被放宽表示为 $\|\mathbf{D} \odot \mathbf{Z}\|_1$ 。因此, 式 (5) 可以重新表述为

$$\begin{aligned} \min_{\mathbf{Z}} & \lambda_1 \|\mathbf{A} \odot \mathbf{Z}\|_F^2 + \lambda_2 \|\mathbf{D} \odot \mathbf{Z}\|_1 \\ \text{s.t. } & \mathbf{X} = \mathbf{X}\mathbf{Z} \end{aligned} \quad (6)$$

通过整合公式 (4) 和 (6), 可以得到:

$$\begin{aligned} \min_{\mathbf{Z}} & \|\mathbf{Z}\|_* + \lambda_1 \|\tilde{\mathbf{A}} \odot \mathbf{Z}\|_F^2 + \lambda_2 \|\mathbf{D} \odot \mathbf{Z}\|_1 + \lambda_3 \|\mathbf{E}\|_{21} \\ \text{s.t. } & \mathbf{X} = \mathbf{X}\mathbf{Z} + \mathbf{E} \end{aligned} \quad (7)$$

其中 λ_1 、 λ_2 和 λ_3 为式 (7) 的相应项加权。

下面本文对式 (7) 进行核化。

设 $\phi: \mathbf{R}^d \rightarrow H$ 是从输入空间到 Hilbert 空间的映射, $\mathbf{K} \in \mathbf{R}^{n \times n}$ 为半正定核矩阵:

$$\mathbf{K}_{ij} = (\phi(\mathbf{X})^T \phi(\mathbf{X}))_{ij} = \phi(\mathbf{x}_i)^T \phi(\mathbf{x}_j) = \text{ker}(\mathbf{x}_i, \mathbf{x}_j)$$

其中 $\text{ker}: \mathbf{R}^d \times \mathbf{R}^d \rightarrow \mathbf{R}$ 是核函数。式 (7) 中的 \mathbf{E} 可以表示为 $\mathbf{E} = \mathbf{X} - \mathbf{X}\mathbf{Z}$, 所以有

$$\min_{\mathbf{Z}} \|\mathbf{Z}\|_* + \lambda_1 \|\tilde{\mathbf{A}} \odot \mathbf{Z}\|_F^2 + \lambda_2 \|\mathbf{D} \odot \mathbf{Z}\|_1 + \lambda_3 \|\mathbf{X} - \mathbf{X}\mathbf{Z}\|_{21} \quad (8)$$

然后定义一个变量 $\mathbf{S} = \mathbf{I} - \mathbf{Z} \in \mathbf{R}^{n \times n}$, 可以得到

$$\|\mathbf{X} - \mathbf{X}\mathbf{Z}\|_{21} = \|\mathbf{X}\mathbf{S}\|_{21} = \sum_{i=1}^n \|\mathbf{X}\mathbf{s}_i\|_2 = \sum_{i=1}^n (\mathbf{s}_i^T \mathbf{X}^T \mathbf{X} \mathbf{s}_i)^{1/2}, \text{ 其中 } \mathbf{S} = [\mathbf{s}_1, \mathbf{s}_2, \dots, \mathbf{s}_n], \mathbf{s}_i \in \mathbf{R}^n \text{ 表示 } \mathbf{S} \text{ 的第 } i \text{ 列。将其中的 } \mathbf{X} \text{ 替换为 } \phi(\mathbf{X}), \text{ 可以得到}$$

$$\begin{aligned} \min_{\mathbf{Z}} & \|\mathbf{Z}\|_* + \lambda_1 \|\tilde{\mathbf{A}} \odot \mathbf{Z}\|_F^2 + \\ & \lambda_2 \|\mathbf{D} \odot \mathbf{Z}\|_1 + \lambda_3 \sum_{i=1}^n \sqrt{\mathbf{s}_i^T \phi(\mathbf{X})^T \phi(\mathbf{X}) \mathbf{s}_i} \end{aligned} \quad (9)$$

$$\text{s.t. } \mathbf{S} = \mathbf{I} - \mathbf{Z}$$

因为 $\mathbf{K} = \phi(\mathbf{x})^T \phi(\mathbf{x})$, 所以可以通过将函数 $g(\mathbf{S})$

定义为 $g(\mathbf{S}) \triangleq \sum_{i=1}^n \sqrt{\mathbf{s}_i^T \mathbf{K} \mathbf{s}_i}$, 将式 (9) 重写为以下问题:

$$\begin{aligned} \min_{\mathbf{Z}} & \|\mathbf{Z}\|_* + \lambda_1 \|\tilde{\mathbf{A}} \odot \mathbf{Z}\|_F^2 + \lambda_2 \|\mathbf{D} \odot \mathbf{Z}\|_1 + \lambda_3 g(\mathbf{S}) \\ \text{s.t. } & \mathbf{S} = \mathbf{I} - \mathbf{Z} \end{aligned} \quad (10)$$

其中核矩阵 \mathbf{K} 包含在 $g(\mathbf{S})$ 中。

定义4 (拉格朗日乘子法、增广拉格朗日乘子法、交替方向乘子法)

1) 假定要求解 $f(\mathbf{X})$ 的最优化问题, 满足 $h(\mathbf{X})=0$, 其中 $f(\mathbf{X}): \mathbf{R}^n \rightarrow \mathbf{R}$, $h(\mathbf{X}): \mathbf{R}^n \rightarrow \mathbf{R}^m$ 。基于拉格朗日乘子法^[13]可以得到以下目标函数:

$$L_c(x, \lambda) = f(\mathbf{X}) + \mu h(\mathbf{X})$$

2) 在计算等式约束问题时, 使用增广拉格朗日乘子法 (ALM), ALM 增加了二次惩罚项, 基于 ALM 可以得到以下目标函数:

$$L_c(x, \lambda) = f(\mathbf{X}) + \mu h(\mathbf{X}) + \frac{c}{2} \|h(\mathbf{X})\|^2$$

3) 要求解 $f(\mathbf{X})+g(\mathbf{Z})$ 的最优化问题, 满足约束条件 $\mathbf{A}\mathbf{X}+\mathbf{B}\mathbf{Z}=\mathbf{C}$ 。这类问题需要用到交替方向乘子法 (ADMM)^[13], ADMM 是 ALM 的进一步推广, 它将对偶上升法的可分解性与 ALM 的收敛性相结合, 基于 ADMM 方法可以得到以下目标函数:

$$L_c(\mathbf{x}, \mathbf{z}, \lambda) = f(\mathbf{X}) + g(\mathbf{Z}) + \lambda^T (\mathbf{A}\mathbf{X} + \mathbf{B}\mathbf{Z} - \mathbf{C}) + \frac{c}{2} \|\mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{z} - \mathbf{C}\|^2$$

2 低秩分块矩阵的核近似算法

算法1 低秩分块矩阵的核近似算法

输入 特征矩阵 \mathbf{X} ; 参数 $\lambda_1, \lambda_2, \lambda_3$; 距离测量矩阵 \mathbf{D} ;

输出 系数矩阵 \mathbf{Z} 。

初始化: $\mathbf{J}=0, \mathbf{Z}=0, \mathbf{Q}=0, \mathbf{S}=0, \lambda_1, \lambda_2, \lambda_3 > 0$,

$\mathbf{C}_1=0, \mathbf{C}_2=0, \mathbf{C}_3=0$,

while

$$\max(\|\mathbf{I} - \mathbf{Z}^{t+1} - \mathbf{S}^{t+1}\|_\infty, \|\mathbf{J}^{t+1} - \mathbf{Z}^{t+1}\|_\infty, \|\mathbf{Q}^{t+1} - \mathbf{Z}^{t+1}\|_\infty) > \text{tol}$$

do

1) 通过式 (14) 更新系数矩阵 \mathbf{Z} ;

2) 通过式 (15) 更新辅助变量 \mathbf{J} ;

3) 通过式 (17) 更新辅助变量 \mathbf{Q} ;

4) 通过式 (18) 更新辅助变量 \mathbf{S} ;

5) 更新拉格朗日乘子 $\mathbf{C}_1, \mathbf{C}_2$ 和 \mathbf{C}_3 ;

$$\begin{cases} \mathbf{C}_1^{t+1} = \mathbf{C}_1^t + \mu^t (\mathbf{I} - \mathbf{Z}^{t+1} - \mathbf{S}^{t+1}) \\ \mathbf{C}_2^{t+1} = \mathbf{C}_2^t + \mu^t (\mathbf{J}^{t+1} - \mathbf{Z}^{t+1}) \\ \mathbf{C}_3^{t+1} = \mathbf{C}_3^t + \mu^t (\mathbf{Q}^{t+1} - \mathbf{Z}^{t+1}) \end{cases}$$

6) 更新 μ

$$\mu^{t+1} = \min(\mu_{\max}, \rho \mu^t)$$

end

为了优化式 (10), 首先引入两个辅助变量 \mathbf{J} 和 \mathbf{Q} 来使问题可分离, 将式 (10) 重写为

$$\min_{\mathbf{J}, \mathbf{Z}, \mathbf{Q}, \mathbf{S}} \|\mathbf{J}\|_* + \frac{\lambda_1}{2} \|\tilde{\mathbf{A}} \odot \mathbf{Z}\|_F^2 + \lambda_2 \|\mathbf{D} \odot \mathbf{Q}\|_1 + \lambda_3 g(\mathbf{S}) \quad (11)$$

$$\text{s.t. } \mathbf{S} = \mathbf{I} - \mathbf{Z}, \mathbf{J} = \mathbf{Z}, \mathbf{Q} = \mathbf{Z}, \mathbf{Z} \geq 0$$

然后, 可以通过 ALM 得到式 (11) 的增广拉格朗日函数

$$L(\mathbf{J}, \mathbf{Z}, \mathbf{Q}, \mathbf{S}, \mathbf{C}_1, \mathbf{C}_2, \mathbf{C}_3) =$$

$$\begin{aligned} & \|\mathbf{J}\|_* + \frac{\lambda_1}{2} \|\tilde{\mathbf{A}} \odot \mathbf{Z}\|_F^2 + \lambda_2 \|\mathbf{D} \odot \mathbf{Q}\|_1 + \lambda_3 g(\mathbf{S}) + \\ & \langle \mathbf{C}_1, \mathbf{I} - \mathbf{Z} - \mathbf{S} \rangle + \langle \mathbf{C}_2, \mathbf{J} - \mathbf{Z} \rangle + \langle \mathbf{C}_3, \mathbf{Q} - \mathbf{Z} \rangle + \\ & \frac{\mu}{2} (\|\mathbf{J} - \mathbf{Z}\|_F^2 + \|\mathbf{I} - \mathbf{Z} - \mathbf{S}\|_F^2 + \|\mathbf{Q} - \mathbf{Z}\|_F^2) \end{aligned} \quad (12)$$

其中 $\langle \mathbf{J}, \mathbf{Q} \rangle = \text{tr}(\mathbf{J}^T \mathbf{Q})$ 。 \mathbf{C}_1 、 \mathbf{C}_2 和 \mathbf{C}_3 是拉格朗日乘子, $\mu > 0$ 是惩罚参数。下面介绍详细迭代步骤:

更新 \mathbf{Z} 固定其他变量并通过下述步骤更新 \mathbf{Z}

$$\begin{aligned} L = \min_{\mathbf{Z}} & \frac{\lambda_1}{2} \|\tilde{\mathbf{A}} \odot \mathbf{Z}\|_F^2 + \langle \mathbf{C}_1^t, \mathbf{I} - \mathbf{Z} - \mathbf{S}^t \rangle + \\ & \langle \mathbf{C}_2^t, \mathbf{J}^{t+1} - \mathbf{Z} \rangle + \langle \mathbf{C}_3^t, \mathbf{Q}^t - \mathbf{Z} \rangle + \frac{\mu^t}{2} (\|\mathbf{J}^{t+1} - \mathbf{Z}\|_F^2 + \\ & \|\mathbf{I} - \mathbf{Z} - \mathbf{S}^t\|_F^2 + \|\mathbf{Q}^t - \mathbf{Z}\|_F^2) \end{aligned}$$

这相当于

$$\begin{aligned} L = \min_{\mathbf{Z}} & \frac{\lambda_1}{2} \|\mathbf{Z} - \mathbf{R}\|_F^2 + \frac{\mu^t}{2} \left(\left\| \mathbf{I} - \mathbf{Z} - \mathbf{S}^t + \frac{\mathbf{C}_1^t}{\mu^t} \right\|_F^2 + \right. \\ & \left. \left\| \mathbf{J}^{t+1} - \mathbf{Z} + \frac{\mathbf{C}_2^t}{\mu^t} \right\|_F^2 + \left\| \mathbf{Q}^t - \mathbf{Z} + \frac{\mathbf{C}_3^t}{\mu^t} \right\|_F^2 \right) \end{aligned} \quad (13)$$

其中 $\mathbf{R} = [\mathbf{Y}, \mathbf{0}_{n(N-n)}] \odot \mathbf{Z}^t$ 。通过推导 $\partial L / \partial \mathbf{Z} = 0$, 可以得到 \mathbf{Z} 的最优解, 式 (13) 的闭式解为

$$\mathbf{Z}^{t+1} = [(2 + \frac{\lambda_1}{\mu^t})\mathbf{I} + \mathbf{K}]^{(-1)} (\frac{\lambda_1}{\mu^t} \mathbf{R} + \mathbf{U}_1 + \mathbf{U}_2 + \mathbf{U}_3) \quad (14)$$

其中 $\mathbf{U}_1 = \mathbf{I} - \mathbf{S}^t + \frac{\mathbf{C}_1^t}{\mu^t}$, $\mathbf{U}_2 = \mathbf{J}^{t+1} + \frac{\mathbf{C}_2^t}{\mu^t}$, $\mathbf{U}_3 = \mathbf{Q}^t + \frac{\mathbf{C}_3^t}{\mu^t}$ 。

更新 \mathbf{J} 当固定其他变量时, 式 (12) 的目标函数可以表示为 \mathbf{J} 的函数, 即

$$\begin{aligned} \mathbf{J}^{t+1} = \arg \min_{\mathbf{J}} & \|\mathbf{J}\|_* + \langle \mathbf{C}_2^t, \mathbf{J} - \mathbf{Z}^t \rangle + \frac{\mu^t}{2} \|\mathbf{J} - \mathbf{Z}^t\|_F^2 = \\ & \|\mathbf{J}\|_* + \frac{\mu^t}{2} \left\| \mathbf{J} - \left(\mathbf{Z}^t - \frac{\mathbf{C}_2^t}{\mu^t} \right) \right\|_F^2 \end{aligned}$$

通过使用奇异值阈值算子 (见定义 3), 可以得到一个封闭形式的解决方案, 即

$$\mathbf{J}^{t+1} = \Gamma_{1/\mu^t}(\mathbf{Z}^t - \frac{\mathbf{C}_2^t}{\mu^t}) = \mathbf{U} \mathbf{S}_{1/\mu^t}(\Sigma) \mathbf{V}^T \quad (15)$$

式中: $\mathbf{U} \Sigma \mathbf{V}^T$ 是 $(\mathbf{Z}^t - \frac{\mathbf{C}_2^t}{\mu^t})$ 的奇异值分解; $\mathbf{S}_{1/\mu^t}(\cdot)$ 是软阈值算子^[7], 定义为

$$S_\lambda(x) = \begin{cases} x - \lambda, & x > \lambda \\ x + \lambda, & x < -\lambda \\ 0, & -\lambda < x < \lambda \end{cases}$$

更新 \mathbf{Q} 当固定其他变量时, 式 (12) 的目标函数可以表示为 \mathbf{Q} 的函数, 即

$$L = \min_{\mathbf{Q}} \lambda_2 \|\mathbf{D} \odot \mathbf{Q}\|_1 + \frac{\mu^t}{2} \left\| \mathbf{Q} - \left(\mathbf{Z}^{t+1} - \frac{\mathbf{C}_3^t}{\mu^t} \right) \right\|_F^2 \quad (16)$$

可以通过逐元素的方法进行更新。显然, 式 (16) 可以等效为解决 $n \times N$ 个子问题。对于第 i 行第 j 列元素 \mathbf{K}_{ij} , 式 (16) 的最优解是

$$\mathbf{Q}_{ij}^{t+1} = \arg \min_{\mathbf{Q}_{ij}} \lambda_2 \mathbf{D}_{ij} |\mathbf{Q}_{ij}| + \frac{\mu^t}{2} (\mathbf{Q}_{ij} - \mathbf{M}_{ij})^2 = \mathbf{S}_{\frac{\lambda_2 \mathbf{D}_{ij}}{\mu^t}}(\mathbf{M}_{ij}) \quad (17)$$

更新 \mathbf{S} 当固定其他变量时, 式 (12) 的目标函数可以表示为 \mathbf{S} 的函数, 即

$$L = \min_{\mathbf{S}} \lambda_3 g(\mathbf{S}) + \frac{\mu^t}{2} \left\| \mathbf{S} - \left(\mathbf{I} - \mathbf{Z}^{t+1} - \frac{\mathbf{C}_1^t}{\mu^t} \right) \right\|_F^2$$

然后令 $\mathbf{\Gamma} = \mathbf{I} - \mathbf{Z}^{t+1} - \frac{\mathbf{C}_1^t}{\mu^t}$, 则 \mathbf{S}^{t+1} 的第 i 行为

$$\mathbf{S}^{t+1}(i, :) = \begin{cases} \frac{\|\mathbf{\Gamma}^i\|_2 - \frac{\lambda_3}{\mu^t}}{\|\mathbf{\Gamma}^i\|_2} \mathbf{\Gamma}^i & \text{if } \|\mathbf{\Gamma}^i\|_2 > \frac{\lambda_3}{\mu^t} \\ 0 & \text{if } \|\mathbf{\Gamma}^i\|_2 \leq \frac{\lambda_3}{\mu^t} \end{cases} \quad (18)$$

其中 $\mathbf{\Gamma}^i$ 是矩阵 $\mathbf{\Gamma}$ 的第 i 行。

在优化变量 \mathbf{J} 、 \mathbf{Z} 、 \mathbf{Q} 和 \mathbf{S} 之后, ADMM 算法还需要更新拉格朗日乘子 \mathbf{C}_1 、 \mathbf{C}_2 、 \mathbf{C}_3 以及参数 μ , 以便更快地收敛。

最后, 本文通过谱聚类算法进行聚类, 即先通过构造亲和矩阵来找到数据的低维嵌入, 然后使用 k 均值聚类来实现最后的聚类。

3 算法的性质分析

3.1 算法的收敛性分析

为了解决目标函数, 即式 (10), 本文使用了迭代更新 ADMM 算法, 如第 2 章所示。下面证明 LBKA 的收敛性。

定理 1 LBKA 是收敛的。

证明 经典的 ADMM 算法主要解决下述问题:

$$\begin{aligned} \min_{\mathbf{z} \in \mathbf{R}^n, \mathbf{w} \in \mathbf{R}^m} & f(\mathbf{z}) + h(\mathbf{w}) \\ \text{s.t. } & \mathbf{R}\mathbf{z} + \mathbf{T}\mathbf{w} = \mu \end{aligned} \quad (19)$$

其中 $\mathbf{R} \in \mathbf{R}^{p \times n}$, $\mathbf{T} \in \mathbf{R}^{p \times m}$, $\mu \in \mathbf{R}^p$, f 和 h 是凸函数。

可以看出, 式 (11) 是式 (19) 的一种特殊情况。经变换, 式 (12) 可以被转换为式 (19)。然后, 就可以使用 ADMM 算法以交替方式更新两个原始变量, 并迭代地解决式 (19):

$$\begin{aligned} \mathbf{Z}^{t+1} &= \arg \min_{\mathbf{Z} \in \mathbf{R}^{n \times N}} L_{\mu}(\mathbf{Z}, \mathbf{W}^t, \mathbf{C}^t) \\ \mathbf{W}^{t+1} &= \arg \min_{\mathbf{W} \in \mathbf{R}^{m \times N}} L_{\mu}(\mathbf{Z}^{t+1}, \mathbf{W}, \mathbf{C}^t) \\ \mathbf{C}^{t+1} &= \mathbf{C}^t + \mu(\mathbf{R}\mathbf{Z}^{t+1} + \mathbf{T}\mathbf{W}^{t+1} - \mathbf{U}) \end{aligned} \quad (20)$$

它与第 2 章中的算法 1 有相同的更新步骤。因此, 式 (11) 是经典 ADMM 问题的一个特殊情况。算法 1 中所提出的优化算法相当于两块 ADMM, 它的收敛性在理论上得以证明。

3.2 算法的复杂度分析

定理 2 根据第 2 章可以看出 LBKA 的总时间复杂度为 $O_k(2n^2N + n^2d + nN)$, 其中 n 为训练样本数, N 为样本总数, d 为样本维数, k 为迭代次数。

证明 算法 1 的主要过程在第二节算法迭代的前 3 步给出, 因为需要进行奇异值分解 (SVD) 和矩阵运算。因此, 当训练样本数 n 和样本总数

N 非常大时, LBKA 的时间复杂度会很高。特别是计算矩阵 $\mathbf{J} \in \mathbf{R}^{n \times N}$ 的 SVD 分解需要 $O(n^2N)(N > n)$ 的复杂度。在这里需要注意的是, 由于要计算矩阵的逆, 迭代更新 \mathbf{Z} 时需要 $O(n^2N + n^2d)$ 时间, 其中 d 是样本维数。步骤 3 的时间复杂度是 $O(nN)$ 。因此, LBKA 的总时间复杂度为 $O_k(2n^2N + n^2d + nN)$, 其中 k 是迭代次数。

相比之下, 基于稀疏表示的分类方法 (如 SRC 和 LatLRR) 的时间复杂度是 $O(n^2(N-n)d)$, 这要比 LBKA 慢很多。LRLR^[15] 和 LRRR^[16] 等回归方法的计算复杂度为 $O(nd + n^2d)$, 比 LBKA 快一点, 但计算精度比 LBKA 低。基于低秩和稀疏表示的方法 (如 RPCA) 需要同时计算特征矩阵的 SVD 并求解软阈值问题。LBKA 的总体时间复杂度与低秩稀疏表示方法的总时间大致相同。

4 实验结果与分析

4.1 实验环境

本文在两个数据集上分别测试了人脸识别和字符识别两个识别任务, 然后与一些最先进的识别方法进行比较, 包括基于稀疏表示的方法, 如 LRC^[17], 基于低秩标准的方法, 如 RPCA, 和传统的分类方法, 如支持向量机 (SVM)^[18], 以及块对角低秩表示方法 (block-diagonal low-rank representation, BDLRR)^[5]。

为了保证 LBKA 和对比算法在实验中的参数相同, 本文使用交叉验证方法重新实现了所有算法。因此, 本文使用的所有算法都是在相同条件下进行测试的, 所以实验结果是真实可靠的。

4.2 评估标准

本文使用 2 个常用的度量指标对实验结果进行评估, 分别是准确度 (Accuracy) 和兰德指数 (Rand index):

1) Accuracy (Acc):

$$\text{Acc} = \max_{\pi} \frac{1}{n} \sum_{ij} \mathbf{X}_{\pi(i)j}^t \mathbf{X}_{ij} \quad (21)$$

其中 π 是 n 个组的排列, \mathbf{X}^t 和 \mathbf{X} 分别是分类准确的样本和所有测试样本, 如果点 j 属于簇 i 则它们的第 i 个条目等于 1, 否则为 0。

2) Rand index (RI):

$$\text{RI} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \quad (22)$$

其中, TN 表示不同类样本的被分到不同个集合, FN 表示同一类样本的被分到不同个集合。

4.3 人脸识别

在本节中, 本文使用扩展 YaleB^[5] 面部图像数据集进行实验。

扩展的 YaleB 数据库: 由 38 个人的 2 414 个人脸图像组成, 每个人有 59~64 个亮度不同的图像。在实验过程中, 随机选择其中的 20、25、30、35 个图像作为训练集, 其余的作为测试集进行实验。

图 2 为 LBKA 获得的扩展 YaleB 数据库的数据表示, 从中可以看出获得的矩阵为对角分块结构。表 1 和表 2 分别是扩展的 YaleB 数据库中不同算法的识别精度和兰德指数。从中可以看出, 在使用不同数量的训练集进行实验时, LBKA 都能得到最好的识别结果。而且随着样本数量的增加, LBKA 的识别精度也随之逐渐提高。

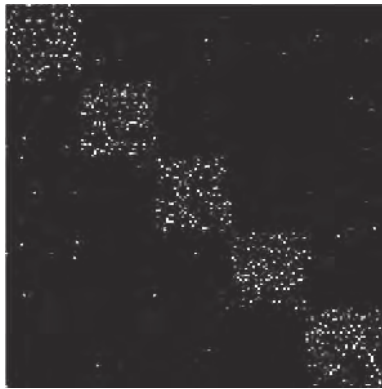


图 2 LBKA 获得的扩展 YaleB 数据库的数据表示

Fig. 2 Data representation of the Extended YaleB database obtained by LBKA

表 1 扩展 YaleB 数据库中不同数量训练集中不同方法的准确度 (Acc±std)

Table 1 Recognition accuracies of different methods of different numbers of training samples on the Extended YaleB database (Acc±std) %

算法	训练集数量			
	20	25	30	35
LRC	92.15±0.95	93.55±0.65	94.55±0.68	95.49±0.55
RPCA	93.58±0.61	95.51±0.36	96.70±0.46	96.96±0.49
SVM	92.81±0.68	95.20±0.44	96.11±0.41	96.70±0.69
BDLRR	94.67±0.31	95.79±0.29	96.46±0.15	96.81±0.14
LBKA	97.46±0.67	98.22±0.42	98.67±0.46	99.26±0.29

4.4 字符识别

本文选取 Char74K 场景人物图像数据集^[5]来进行实验。在以往的模式识别任务中, 由于场景图像十分复杂, 很难将想要识别的文本分离出来, 而 LBKA 中对类内类外的处理有助于对场景中字符进行提取。本文将 LBKA 与其他先进的字符识别算法进行对比, 包括 CoHOG^[19]、RTPD^[20]、PHOG^[21]、MLFP^[22]和 BDLRR^[5]。

Char74K 数据库: 该数据库总共包含 7 万 4 千幅图像, 所以叫 Chars74K。本文主要关注其中英语字符和数字的识别。本文在实验中只使用了原

数据库的一个小子集 Char74K-15, 其中包含 15 个训练样本和 15 个测试样本。

表 2 扩展 YaleB 数据库中不同数量训练集中不同方法的兰德指数 (RI)

Table 2 Rand index of different methods of different numbers of training samples on the Extended YaleB database(RI) %

算法	训练集数量			
	20	25	30	35
LRC	91.25	92.63	93.35	93.76
RPCA	94.61	94.84	95.75	95.96
SVM	93.28	95.51	92.65	96.96
BDLRR	96.34	97.11	96.37	97.03
LBKA	97.82	98.53	97.89	99.11

表 3 和表 4 列出了 LBKA 和其他字符识别方法的识别准确度和兰德指数, 从中可以看出, LBKA 识别的准确度和兰德指数都是最高的。

表 3 场景角色数据库上不同方法的准确度

Table 3 Recognition accuracies of different methods on the scene character database. %

算法	准确度
CoHOG	63
PHOG	65
MLFP	64
RTPD	67
BDLRR	70
LBKA	73

表 4 场景角色数据库上不同方法的兰德指数

Table 4 Rand index of different methods on the scene character database. %

算法	兰德指数
CoHOG	60
PHOG	63
MLFP	61
RTPD	68
BDLRR	72
LBKA	74

4.5 实验分析

首先, 与两种数据集上的对比算法相比, LBKA 进行模式识别时更加准确。

其次, LBKA 比现有的识别算法更准确, 这说明扩大了对角块和非对角块之间的差异的同时增强了相关数据表示, 能够获得更好的识别效果。

第三, LBKA 在识别数量有限的样本时比其它对比算法都好。主要原因是 LBKA 具有稀疏, 低秩和分块三大特性。低秩可以有效地挖掘数据

相关的基础结构,并且揭示数据矩阵的全局潜在结构;稀疏能够寻找最近的数据子空间;块对角表示能够发掘数据内在结构并阐明数据点的最近子空间。

随着迭代次数的增加 LBKA 的相对误差在减小,并且总值在图 3 所示的 60 次迭代后基本没有变化,这验证了 LBKA 具有收敛性。图 4 表示 LBKA 关于参数和的性能变化。可以看出, LBKA 对和的变化值不敏感。更具体的说,当参数处于一个合理范围内时, LBKA 准确度较高,这说明增加类内类外约束是十分有必要的。

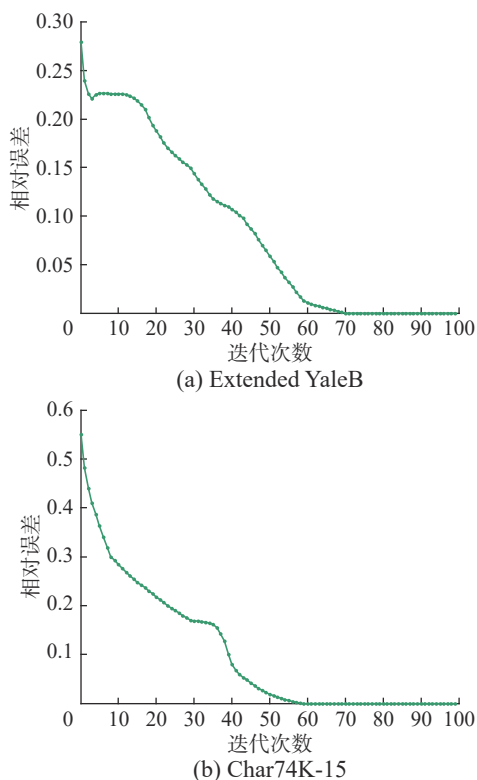


图 3 LBKA 在不同数据库上的收敛曲线

Fig. 3 Convergence curves of LBKA on different databases

5 结束语

本文给出了低秩分块矩阵的相关定义,说明了矩阵分解的应用。分析了核近似的优点,并提出了低秩分块矩阵核近似的方法。最后将该方法在人脸识别和字符识别中进行比较。结果表明,所提出的低秩分块矩阵分解算法在收敛速度和近似精度上都具有一定的优势。未来工作包括:

1) 从矩阵的核近似出发,对原有的子空间聚类算法进行改进,提高分块的速度和准确性,从而提升算法找到全局最优解的能力。

2) 通过与多种矩阵分解算法的比较,观察提出的低秩分块矩阵的核近似算法的表现,进一步分析算法的可行性。

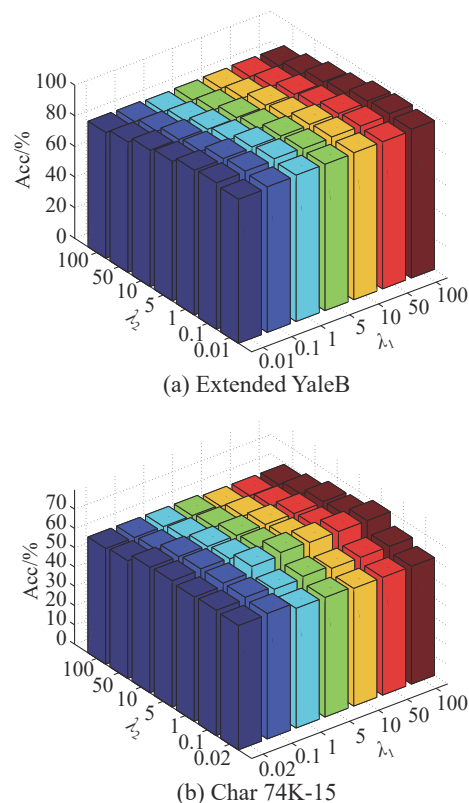


图 4 λ_1, λ_2 对两种数据集的影响

Fig. 4 The impact of λ_1 and λ_2 on the two data sets

3) 本文通过实验证明了 LBKA 在人脸识别和字符识别等领域的效果比传统识别算法更好,接下来可以研究其在回归、聚类、排序等方面的应用。

参考文献:

- [1] 雷恒鑫, 刘惊雷. 基于行列联合选择矩阵分解的偏好特征提取[J]. 模式识别与人工智能, 2017, 30(3): 279-288. LEI Hengxin, LIU Jinglei. Preference feature extraction based on column union row matrix decomposition[J]. Pattern recognition and artificial intelligence, 2017, 30(3): 279-288.
- [2] 张恒敏, 杨健, 郑玮. 低秩矩阵近似与优化问题研究进展[J]. 模式识别与人工智能, 2018, 31(1): 23-36. ZHANG Hengmin, YANG Jian, ZHENG Wei. Research progress of low-rank matrix approximation and optimization problem[J]. Pattern recognition and artificial intelligence, 2018, 31(1): 23-36.
- [3] CHIANG K Y, DHILLON I S, HSIEH C J. Using side information to reliably learn low-rank matrices from missing and corrupted observations[J]. Journal of machine learning research, 2018, 19(76): 1-35.
- [4] LU Canyi, FENG Jiashi, LIN Zhouchen. Subspace clustering by block diagonal representation[J]. IEEE transactions on pattern analysis and machine intelligence, 2019, 41(2): 487-501.
- [5] ZHANG Zheng, XU Yong, SHAO Ling. Discriminative

- block-diagonal representation learning for image recognition[J]. *IEEE transactions on neural networks and learning systems*, 2018, 29(7): 3111–3125.
- [6] WEI C P, CHEN C F, WANG Y C F. Robust face recognition with structurally incoherent low-rank matrix decomposition[J]. *IEEE transactions on image processing*, 2014, 23(8): 3294–3307.
- [7] NI Yuzhao, SUN Ju, YUAN Xiaotong, et al. Robust low-rank subspace segmentation with semidefinite guarantees [C]//Proceedings of 2010 IEEE International Conference on Data Mining Workshops. Sydney, NSW, Australia, 2010: 1179–1188.
- [8] LEE K C, HO J, KRIEGMAN D J. Acquiring linear subspaces for face recognition under variable lighting[J]. *IEEE transactions on pattern analysis and machine intelligence*, 2005, 27(5): 684–698.
- [9] CANDÈS E J, LI Xiaodong, MA Yi. Robust principal component analysis?[J]. *Journal of the ACM*, 2011, 58(3): 11.
- [10] LIU Guangcan, LIN Zhouchen, YAN Shuicheng. Robust recovery of subspace structures by low-rank representation[J]. *IEEE transactions on pattern analysis and machine intelligence*, 2013, 35(1): 171–184.
- [11] CHEN Yudong, JALALI A, SANGHAVI S, et al. Clustering partially observed graphs via convex optimization [J]. *The journal of machine learning research*, 2014, 15(1): 2213–2238.
- [12] LIN Zhouchen, CHEN Minming, MA Yi. The augmented lagrange multiplier method for exact recovery of corrupted low-rank matrices[J]. *arXiv preprint arXiv: 1009.5055*, 2010.
- [13] LU Canyi, FENG Jiashi, YAN Shuicheng. A unified alternating direction method of multipliers by majorization minimization[J]. *IEEE transactions on pattern analysis and machine intelligence*, 2018, 40(3): 527–541.
- [14] 刘松华, 张军英, 丁彩英. 核矩阵列相关低秩近似分解算法 [J]. *模式识别与人工智能*, 2011, 24(6): 776–782.
- LIU Songhua, ZHANG Junying, DING Caiying. Low-Rank approximation and decomposition for kernel matrix based on column correlation[J]. *Pattern recognition and artificial intelligence*, 2011, 24(6): 776–782.
- [15] YIN Ming, GAO Junbin, LIN Zhouchen. Laplacian regularized low-rank representation and its applications[J]. *IEEE transactions on pattern analysis and machine intelligence*, 2016, 38(3): 504–517.
- [16] HE Xiaofei, CAI Deng, SHAO Yuanlong, et al. Laplacian regularized Gaussian mixture model for data clustering[J]. *IEEE transactions on knowledge and data engineering*, 2011, 23(9): 1406–1418.
- [17] NASEEM I, TOGNERI R, BENNAMOUN M. Linear regression for face recognition[J]. *IEEE transactions on pattern analysis and machine intelligence*, 2010, 32(11): 2106–2112.
- [18] 丁立中, 廖士中. KMA- α : 一个支持向量机核矩阵的近似计算算法 [J]. *计算机研究与发展*, 2012, 49(4): 746–753.
- DING Lizhong, LIAO Shizhong. KMA- α : a kernel matrix approximation algorithm for support vector machines[J]. *Journal of computer research and development*, 2012, 49(4): 746–753.
- [19] TIAN Shangxuan, LU Shijian, SU Bolan. Scene text recognition using co-occurrence of histogram of oriented gradients[C]//Proceedings of the 12th International Conference on Document Analysis and Recognition. Washington, DC, USA, 2013: 912–916.
- [20] PHAN T Q, SHIVAKUMARA P, TIAN S X. Recognizing text with perspective distortion in natural scenes[C]//Proceedings of 2013 IEEE International Conference on Computer Vision. Sydney, NSW, Australia, 2013: 569–576.
- [21] TAN Zhirong, TIAN Shangxuan, TAN C L. Using pyramid of histogram of oriented gradients on natural scene text recognition[C]//Proceedings of 2014 IEEE International Conference on Image Processing. Paris, France, 2014: 2629–2633.
- [22] LEE C Y, BHARDWAJ A, DI Wei, et al. Region-based discriminative feature pooling for scene text recognition [C]//Proceedings of 2014 IEEE Conference on Computer Vision and Pattern Recognition. Columbus, OH, USA, 2014: 4050–4057.

作者简介:



王中元,男,1996年生,硕士研究生,主要研究方向为核方法与矩阵分解。



刘惊雷,男,1970年生,教授,博士,主要研究方向为人工智能和理论计算机科学。主持国家自然科学基金面上项目、山东省自然科学基金面上项目各1项。发表学术论文40余篇。