

信号与数据处理中的低秩模型*

林宙辰¹ 马 毅²

¹北京大学

²上海科技大学

关键词：低秩 稀疏 聚类 去噪

引言

我们正处于大数据时代。大数据除了具有4个V (volume, variety, velocity 和 value)¹的特点外,还常常是高维的。例如,用千万像素相机拍一张照片,相当于在千万维空间里采集了一个样本。因此,如何鲁棒高效地处理高维数据是我们面临的一个重大挑战。幸运的是,高维数据并不是毫无结构的。一个显著的特点是它们经常分布在低维流形附近,这实际上是流形学习的基本假设。比如对数据做主元分析(Principal Component Analysis, PCA)时,往往95%以上的能量都集中在少数几个主方向上。因此,充分利用数据分布的低维特性是解决上述挑战的一条重要途

径。不过,由于传统的流形学习对方法的正确性缺乏深入的理论分析,因此其性能没有保障,尤其是当数据伴随着强噪声、存在损毁或缺失甚至可能失效时。而这样的数据在现实中是大量存在的。

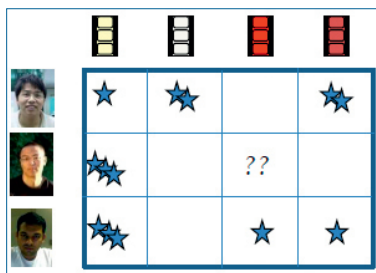


图1 Netflix挑战²(需要预测用户对其未评价过的视频的喜好程度)

基于低秩的模型是近年来处理高维数据的新工具。它的兴起受到稀疏表示和压缩感知理论的

推动,由此系统地发展出新的理论和应用。在此背景下,秩被阐释为二阶(即矩阵)稀疏性³的度量,而不仅仅是一个数学概念。以图像或视频压缩为例,要实现有效的压缩,必须充分利用图像或视频的时间或空间相关性;又如Netflix挑战(见图1),要推断未知的用户评价,需要充分考虑用户喜好的相关性和视频类别的相关性。由于矩阵行列间的相关性和矩阵的秩天然地关联在一起,因此可以把秩定义为二阶稀疏性度量。

数学模型

数学模型可分为线性模型和非线性模型,前者可进一步分为单子空间模型和多子空间

* 本文得到国家自然科学基金(61272341, 61231002)资助。

¹ 体量、多样性、速度和价值。

² Netflix是一家视频租赁公司,拥有很多用户对视频的评价,但这个用户/视频评价矩阵非常稀疏。该公司提供100万美元奖金希望能够把预测用户对视频的评价的准确率提高10%,以便有针对性地推荐,从而提高营收。见<http://www.netflixprize.com/>。

³ 一阶稀疏性是指向量的稀疏性,其度量为非零元的个数,即 l_0 范数 $\|\cdot\|_0$ 。

模型。虽然线性模型看上去比较简单,但是理论分析表明它对强噪声和缺失数据有很强的鲁棒性,在应用中也具有足够的数据表达能力。

单子空间模型

低秩模型的兴起始于坎迪斯(E. Candès)于2008年提出的矩阵填充(Matrix Completion, MC)问题:已知某矩阵 D 在某些位置的值,能否恢复出该矩阵?这是一个广泛的数学模型,如上面提到的Netflix挑战、基因微阵列测量等,其答案显然是不确定的。由于需要考虑矩阵行列间的相关性,因此在数据缺失时,为了恢复低秩结构,坎迪斯建议选秩最小的解。

如何在数据有强噪声时恢复低秩结构?这个问题看似可以用传统的主元分析解决,但实际上传统主元分析只在高斯噪声时可以准确恢复潜在的低秩结构。对于非高斯噪声,如果噪声很强,即使是极少的噪声,也会使传统的主元分析失败。为了提高主元分析的鲁棒性,学者们付出了很多努力,提出了许多“鲁棒”的主元分析方法,但是没有有一个方法在理论上被严格证明是一定能在一定条件下精确恢复出低秩结构的。2009年,钱德拉塞卡兰(V. Chandrasekaran)等人和赖特(J. Wright)等人同时提出了鲁棒主元分析法(Robust PCA, RPCA)。在数据中有稀疏大噪声的情况时,他们采用的方法是把给定观

测矩阵 D 分解成一个低秩矩阵与一个稀疏矩阵的和。后来坎迪斯加入到赖特的工作中,获得了更好的结果,允许只知道观测矩阵 D 部分位置的值^[1]。陈(Y. Chen)等人考虑了噪声在若干列集中的情况,提出了离群追踪(outlier pursuit)模型,把给定观测矩阵分解成一个低秩矩阵与一个列稀疏矩阵的和。

刘(J. Liu)等人把矩阵填充推广到张量填充。张量有基于CP(CANDECOMP-PARAFAC)分解定义的秩,由于它不可计算,因此提出了一种折衷的张量秩,即张量按不同模式展开后得到不同矩阵的秩的和。他们提出的张量填充模型是:已知张量中某些位置的值,通过对折衷的张量秩进行极小化来恢复缺失的值。同样,基于折衷的张量秩,谭(H. Tan)等人把鲁棒主元分析推广到了张量恢复,即把给定张量分解为两个张量之和,一个具有较低的折衷的张量秩,另一个是稀疏张量。

多子空间模型

鲁棒主元分析只能从数据中提取一个子空间,无法刻画数据在此子空间中的精细结构。精细结构的最简单形式是多子空间模型,即数据分布在若干子空间附近,需要找到这些子空间。马毅等人称之为广义主元分析(Generalized PCA, GPCA)。之前已有很多算法,如代数法、随机抽样一致法(RANdom SAMple

Consensus, RANSAC)等,但都没有理论保障。稀疏表示的出现为其提供了新的思路。埃尔赫米法(E. Elhamifar)和维达尔(R. Vidal)于2009年利用样本间相互表达,以表达系数矩阵 Z 稀疏为目标,提出了稀疏子空间聚类(Sparse Subspace Clustering, SSC)模型。受此启发,刘光灿等人提出了低秩表示(Low-Rank Representation, LRR)模型^[4],使表达系数矩阵 Z 低秩,目的是为了增强 Z 各列之间的相关性,以提高对噪声的抵抗能力。稀疏子空间聚类和低秩表示的最优表达系数矩阵可以作为样本间的相似性度量,再通过谱聚类,获得数据所在的若干线性子空间。庄连生等人进一步要求系数矩阵稀疏、非负,以用于半监督学习。

低秩表示模型需要假定数据充足。在样本不足的情况下,刘光灿和颜水成通过引入未观测数据来表达已观测数据,推导出潜在低秩表示(Latent LRR)模型;而刘日升等人提出了固定秩表示(Fixed Rank Representation, FRR)模型,利用不超过给定秩的表达系数矩阵进行聚类。为了进一步提高子空间聚类的精度,卢参义等人提出使用迹套索方法(trace lasso)来约束表达系数。该方法具有自适应性,可随数据相关性而变化,因此称为相关性自适应子空间分割(Correlation Adaptive Subspace Segmentation, CASS)。为了更有效地对张量数据进行聚

类,傅逸凡等人提出了张量低秩表示 (Tensor LRR) 模型,可综合张量在各模式下的信息。

非线性模型

目前用于分割非线性流形的低秩模型比较少。王 (J. Wang) 等人提出了核 (Kernel) 技巧,用于推广低秩表示模型。但对不是高斯型的噪声该方法并不适用。

优化算法

文中所列的离散低秩模型一般都是 NP 问题,只能实现近似求解。通常采用的方法是将其转化成连续优化问题,主要分为两种:一种是凸优化法,如上面提到的,把 l_0 范数 $\|\cdot\|_0$ 换成 l_1 范数 $\|\cdot\|_1$,秩换成核范数 $\|\cdot\|_*$;另一种是非凸优化法,用非凸的连续函数来近似 l_0 范数 $\|\cdot\|_0$ (比如用 l_p 范数 $\|\cdot\|_p$ ($0 < p < 1$)) 和秩 (比如用 Schatten- p 范数 (奇异值的 l_p 范数))。凸优化法的好处是能够得到修改后模型的全局最优解,不过解有可能不够低秩或稀疏;非凸优化法采用的有效算法是迭代重加权最小二乘法 (iteratively reweighted least squares)^[5],好处是能得到更低秩和更稀疏的解,但是不能得到修改后模型的全局最优解,解的质量可能依赖于初值。所以凸优化和非凸优化两种方法互补。此外,还有一种做法是把需要约束为低秩的矩阵直接

表示为两个矩阵的乘积,第一个矩阵的列数和第二个矩阵的行数均为期望的秩,然后交替更新两个矩阵,直至不变为止。目前面向大规模计算的凸优化算法都是一阶算法。代表性的算法包括加速近邻梯度法 (Accelerated Proximal Gradient, APG) 和交错方向法 (Alternating Direction Method, ADM)^[2,3]。针对问题的特点,还可设计随机算法,以大幅降低求解的复杂度。

代表性应用

低秩模型在信号处理和机器

最简单情况是从固定摄像机拍摄的视频中分离背景和前景。背景是基本不变的,如果把背景的每一帧作为矩阵的一列,则该矩阵低秩。同时由于前景是移动的物体,占据像素比例较低,因此前景是视频中的稀疏“噪声”部分。由此得到基于背景建模的鲁棒主元分析模型,其中数据矩阵 D 的每一列是视频的每一帧拉直后得到的向量,低秩矩阵 A 的每一列是背景的每一帧拉直后得到的向量,稀疏矩阵 E 的每一列是前景的每一帧拉直后得到的向量。部分结果如图 2 所示。

图像批量对齐^[6] 背景建



图2 背景建模

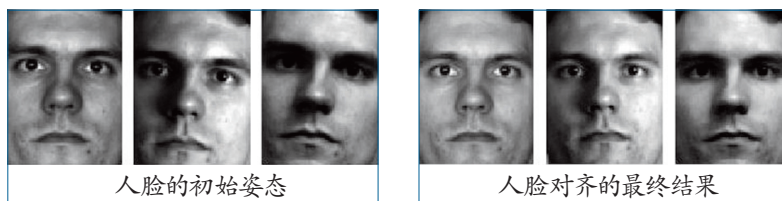


图3 人脸对齐 (这里只显示了部分人脸图像)

学习等领域里已经获得了广泛的应用, NIPS⁴ 2011 上曾出现大量讨论低秩模型的论文。下面仅介绍笔者及合作者在图像处理和计算机视觉领域里的工作。

背景建模^[1] 背景建模的

模需要假定背景已经对齐,而在没有对齐的情况下,则可以把每一帧/每一幅图像做适当的几何变形使其对齐。彭义刚等人^[6]提出了把几何变形的参数进行局部线性化的迭代算法。在仿射变换

⁴ Neural information processing systems, 神经信息处理系统。



图4 利用TILT模型进行图像校正的例子(第一行表示原始的图像块(长方形框)及相应校正变换(四边形框,校正变换就是四边形框相对于长方形框的几何变换),第二行是校正后的图像块)

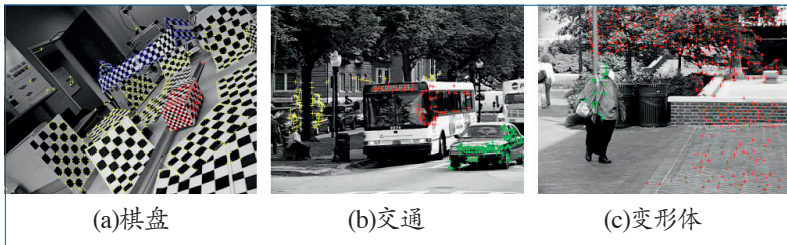


图5 运动分割的例子(不同颜色特征点表示它们属于不同物体)



图6 图像分割的例子(第一行是原图,第二行是分割结果)

Image	F-Map	C-Map	LC-Map	MTSP

图7 图像显著区域检测的例子(第一列是输入图像,第二至五列是不同方法的检测结果,最后一列是基于低秩表示模型的检测方法)

模型下,人脸图像对齐的部分结果如图3所示。

变换不变低秩纹理^[8] 变换不变低秩纹理(Transform Invariant Low-rank Textures, TILT)

的数学模型与图像批量对齐法在表现形式和求解过程方面相同,不过前者的数据矩阵是图像的一个长方形图像块。变换不变低秩纹理法的思想是通过几何变换把

数据矩阵所代表的图像区域校正成正则的区域,如具有横平竖直、对称等特性。这些特性可通过低秩性来刻画。变换不变低秩纹理法原则上对任何参数变换都适用。张铮东等人还考虑了基于广义柱体变换的变换不变低秩纹理法,用于人造建筑物表面纹理的提取。该方法被广泛应用于建筑物几何建模^[8]、相机自动标定和镜头畸变自动校正、文字校正与检测等。任翔和林宙辰^[7]特别研究了该方法基于LADM(linearized alternating direction method, 线性化交替方向法)的快速算法,把求解速度提高了5倍以上。图4是射影变换下图像校正的例子。

运动分割^[4] 低秩表示模型是做刚体运动分割最好的算法之一。所谓刚体运动分割是把视频里做刚体运动物体上的特征点进行聚类,使得每一类与一个独立运动的物体相对应,从而得到物体运动的轨迹。实例见图5。

图像分割 图像分割是特殊的聚类问题。首先把图像过分割(oversegment)成超像素,然后在超像素上提取适当的特征,利用改进的低秩表示模型来综合多种特征(基本上每一种特征对应于一个低秩表示模型),求出整体表示矩阵 Z^* ,然后对用 $(|Z^*|+l(Z^*)^T)/2$ 表示的相似性矩阵进行正则化分割(normalized cut),得出超像素的聚类关系,每一类对应到一个图像区域。实例见图6。

图像显著区域检测 运动

分割和图像分割都是利用了低秩表示模型的表达系数矩阵 Z , 而图像显著区域检测则是利用低秩表示模型里的稀疏“噪声” E 。图像显著区域是图像中“与众不同”的区域, 如果用其他区域进行“预测”, 则会产生较大的误差。因此, 如果把图像分解成小块, 在其上提取一些特征, 则图像显著区域对应于低秩表示模型里的稀疏“噪声” E 较大的部分。实例见图7。

其他应用 低秩模型的其他应用还包括: 部分重复图像检索、人脸识别、结构化纹理修复、街景全景图拼接、人造物体定向 (upright orientation)、基于光度学的立体视觉、图像标签的改进、视觉域适应 (visual domain adaption)、鲁棒视觉跟踪、三维人脸特征提取、计算机断层扫描 (Computed Tomography, CT) 重建、图像半监督分类、图像集的协同分割 (co-segmentation)、视点无关步态识别、文档关键词提取、音频分析、蛋白质-基因相关性分析、网络流量异常检测、鲁棒滤波与系统辨识等等, 限于篇幅就不再介绍了。

结语

近年来, 低秩模型在信号处理、机器学习、计算机视觉等领域获得了广泛的应用, 在短短几年内从理论、算法到应用等各方面得到了快速的发展。本文只是基于我们的工作进行了简单介

绍。在许多实例中, 如果能够结合问题的特性, 适当地引入低秩性约束, 则大多都能得到更好的结果。有些问题的数据本身可能不具有低秩性, 此时, 可通过引入适当的变换, 增强其低秩性 (如用图像批量对齐法和变换不变低秩纹理法对鲁棒主元分析的改进)。有些学者没有检查数据是否具有低秩性或对数据进行适当的预处理就声称使用低秩约束效果不好, 这是不太严谨的。此外, 低秩模型在矩阵张量的推广及用于非线性流形聚类、近线性复杂度的随机算法等方面也有待进一步研究。■



林宙辰

CCF会员。北京大学教授。主要研究方向为计算机视觉、图像处理、机器学习等。zlin@pku.edu.cn



马毅

CCF会员、杰出演讲者。上海科技大学教授。主要研究方向为计算机视觉、高维数据分析。mayi@shanghaitech.edu.cn

参考文献

- [1] Candès E J, Li X, Ma Y, Wright J: Robust principal component analysis? *Journal of the ACM*, 2011, 58(1): 1~37.
- [2] Lin Zhouchen, Liu Risheng, and Li Huan. Linearized alternating direction method with parallel splitting and adaptive penalty

for separable convex programs in machine learning, *Machine Learning*, 2015.

- [3] Lin Z, Liu R, Su Z. Linearized alternating direction method with adaptive penalty for low-rank representation. In: *Proceedings of Advances in Neural Information Processing Systems (NIPS)*, 2011, 612~620.
- [4] Liu G, Lin Z, Yan S, Sun J, Yu Y, Ma Y. Robust recovery of subspace structures by low-rank representation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2013, 35(1): 171~184.
- [5] Lu Canyi, Lin Zhouchen, and Yan Shuicheng. Smoothed low rank and sparse matrix recovery by iteratively reweighted least squared minimization, *IEEE Trans. Image Processing*, 2015.
- [6] Peng Y, Ganesh A, Wright J, Xu W, Ma Y. RASL: Robust alignment by sparse and low-rank decomposition for linearly correlated images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2012, 34(11): 2233~2246.
- [7] Ren X, Lin Z. Linearized alternating direction method with adaptive penalty and warm starts for fast solving transform invariant low-rank textures. *International J. Computer Vision*, 2013, Vol. 104, 1~14.
- [8] Zhang Z, Ganesh A, Liang X, Ma Y. TILT: Transform invariant low-rank textures. *International Journal of Computer Vision*, 2012, 99(1): 1~24.