

# 稀疏—平坦矩阵信号的重构条件

田园, 黄鑫, 李明楚

大连理工大学软件学院, 辽宁 大连 116620

基金项目: 国家自然科学基金(61877007)

通信作者: 田园, tianyuan\_ca@dlut.edu.cn 收稿/录用/修回: 2021-08-24/2021-11-01/2021-12-13

## 摘要

建立基于凸优化方法重构矩阵信号  $\mathbf{X} = [\mathbf{x}_1, \dots, \mathbf{x}_n]$  的一组充分条件,  $\mathbf{X}$  具有列稀疏性和平坦性的结构特征, 即每个列向量  $\mathbf{x}_j$  至多具有  $s$  个非零分量、同时所有列向量的  $l_1$  范数具有相同数值。所采用的矩阵范数是  $\|\|\mathbf{X}\|\|_1 \equiv \max_j \|\mathbf{x}_j\|_1$ 。工作分两部分, 第一部分分别对无观测误差和有观测误差的情况, 针对求解  $\min\|\|\cdot\|\|_1$  型凸优化问题重构以上类型矩阵信号的方法, 建立保障稳定性和鲁棒性的充分条件; 第二部分工作建立随机观测情况下观测空间维数的下界, 用以保证信号以高概率被正确重构。所得结果优于将重构向量信号的  $\min-l_1$  方法直接推广到针对矩阵信号的  $\min-l_1$  方法所得到的结果, 并给出数值仿真验证。所针对的信号模型出现在具有定常或缓变包络波形的多输入/多输出雷达及合成孔径雷达等新应用领域, 本文工作针对这类应用提供一组实用的信号重构条件。

## 关键词

压缩感知  
矩阵信号  
凸优化  
逐列稀疏  
 $l_1$  列平坦  
中图法分类号: TP18  
文献标识码: A

# Conditions for Reconstructing Sparse-Flat Matrix Signals

TIAN Yuan, HUANG Xin, LI Mingchu

Software School, Dalian University of Technology, Dalian 116620, China

## Abstract

In this study, a group of sufficient conditions is established for reconstructing the structured matrix signal  $\mathbf{X} = [\mathbf{x}_1, \dots, \mathbf{x}_n]$  with the column-wise sparsity and flatness features via convex programming, where each column  $\mathbf{x}_j$  is a vector of  $s$ -sparsity and all columns have the same value of  $l_1$  norms. The regularizer in use is a matrix norm  $\|\|\mathbf{X}\|\|_1 \equiv \max_j \|\mathbf{x}_j\|_1$ . The contributions have two parts: First, sufficient conditions are established for stability and robustness in signal reconstruction by solving the  $\min\|\|\cdot\|\|_1$  convex program from noise-free or noisy measurements. Second, fundamental lower bounds on the dimensions of the measurement space are established for correct reconstruction with high probability. The results are superior to those obtained by directly extending the vector signal's  $\min-l_1$  approach to the matrix signal's  $\min-l_1$  approach. Numerical simulations are presented. The investigated signal model appears in some emerging application fields, e.g., multiple-input and multiple-output and synthetic aperture radars with constant or slowly varying envelope waveforms. The work in this paper provides practical conditions to reconstruct such signals.

## Keywords

compressive sensing;  
matrix signal;  
convex programming;  
column-wise sparsity;  
 $l_1$ -column-flatness

## 0 引言

信号的重构是数字通信、计算机成像、自动控制等领域的关键问题之一。该问题常可表述为如何基于未知信号的部分先验知识(如某种结构特征), 基于尽可能少的观测数据来重构未知信号。压缩感知理论的概念、方法和技术, 为解决这类问题提供了不同于经典方法的新途径, 并取得很多创新成果。目前这方面的大多数工作是针对向量信号<sup>[1-3]</sup>。从代数的观点看, 向量属于一阶张量。在越来越多的实际应用中, 也很自然地出现所谓高阶张量信号, 例如在先进雷达信号处理和阵列信号处理领域<sup>[4]</sup>, 核心问题是根据接收数据  $y_{kl}$  和观测模型  $y_{kl} = \sum_{ij} \Psi_{kl, ij} X_{ij} + e_{kl}$  重构矩阵信号  $[X_{ij}]$ 。矩阵属于所谓二阶张量, 这里的  $X_{ij}$  具体表示在距离—方向单元或距离—速度单元  $(i, j)$  上的目标信号 ( $X_{ij}$  为 0 则表示目标不存在)。更复杂的雷达数据处理问题需要重构距离—速度—方向单元  $(i, j, k)$  上的目标信号  $X_{ijk}$ , 这是重构三阶张量信号的例子。

虽然对元素作简单重排后, 任何  $(n_1, \dots, n_p)$  型  $p$  阶张量总可被视为  $n_1, \dots, n_p$  维向量, 但高阶张量信号的代数结构与最简单的一阶张量(向量)有很大差异, 两者所具有的结构化特征也有很大不同<sup>[5]</sup>。以二阶张量(矩阵)信号为例, 如果从点集在平面和直线上的分布模式来直观地类比, 会看出矩阵信号能表现出较向量信号更为丰富的结构特征, 前者的稀疏性模式也比后者更为多样。

在应用压缩感知方法来解决信号重构问题的工作中, 已开始有越来越多的工作涉及高阶张量信号, 例如在数字雷达和分布式传感等领域。然而, 总体而言, 大部分工作是将高阶张量信号作为向量信号来看待, 在此基础上采用较为成熟的  $\min-l_1$  理论和算法来加以处理, 而针对高阶张量信号的固有特点来进行处理的工作相对较少。针对高阶张量信号重构问题的理论研究还远没有针对向量信号的研究广泛和深入。在有代表性的工作中, 文[6]针对稀疏矩阵信号的重构问题, 论述了 Kronecker 乘积分解形式的观测算子的重构条件和观测空间维数的界估计。该工作的研究途径是采用矩阵的  $l_1$  范数  $\sum_{ij} |X_{ij}|$ , 通过求解  $\min-l_1$  凸优化问题来重构稀疏矩阵信号。该方法的优点是能够直接移植基于  $\min-l_1$  方法重构向量信号的诸多成果, 缺点是  $l_1$  范数只适合表达矩阵信号一般性的稀疏特征, 或者说

$l_1$  范数只是表达矩阵信号众多结构特征的凸范数之一, 对表达某些更精细的结构特征并非总是适合。文[7]将文[6]的处理方法推广到处理视频压缩问题中的一类高阶张量信号, 针对具有完全分解结构的特殊类型的观测算子, 建立了基于该算子的相干系数(coherence)的重构条件, 并讨论了对视频重构问题的应用。文[8]面向雷达应用论述了重构距离—速度矩阵信号的一种基本理论, 建立了相应的重构条件。该工作通过求解矩阵的  $\min-l_1$  凸优化问题定出含有目标速度即多普勒频移的一组中间观测量, 再应用谱估计方法来完全确定矩阵信号的元素。文[9]进一步论述了针对多输入/多输出雷达中的距离—方向—速度信号(三阶张量)重构条件, 文[10]针对短时观测序列的情况研究了重构条件的优化, 文[11]针对合成孔径雷达成像问题, 基于上述类似的思路建立求解模型并讨论了相应的近似算法, 在这里所涉及的信号是目标的距离—方向—速度—反射率构成的高阶张量。上述工作都是将高阶张量信号直接视为向量信号, 采用(基于某种特殊函数系的) $l_1$  范数来表达信号的稀疏性质, 再通过仔细利用特定领域中观测算子的特殊性质来建立重构条件。许多其他工作也采取类似途径。

采用非  $\min-l_1$  途径研究矩阵信号重构的工作则较少, 其中有代表性和成果较多的是矩阵完备化(matrix completion)问题和低秩矩阵的重构等问题<sup>[1, 12-13]</sup>。以低秩矩阵重构问题为例, 通过求解矩阵的核范数(矩阵奇异值之和)最小化问题, 目前已建立较完整的理论和多种有效算法<sup>[12-13]</sup>。本文所涉及的矩阵信号的结构特征与已研究过的这些问题不同, 需采用完全不同的矩阵范数来表达, 所得结果也很不相同, 然而这些工作正表明, 因矩阵信号较向量信号具有更为丰富的结构特征, 不同的结构特征需要不同的矩阵范数来表达, 对高阶张量信号就更是如此, 而简单移植对向量信号行之有效的  $\min-l_1$  型的理论方法与结果未必总是各种应用的最优选择。

总之, 矩阵信号和高阶张量信号的重构问题在应用领域越来越多, 特别是传感器阵列、多天线雷达及合成孔径雷达(SAR)成像等领域, 但相应的基础理论尚不完整。尽管算法方面工作相对而言更多一些, 但因理论尚在发展, 相关算法的很多实际问题如误差稳定性、对结构特征的敏感性等还有待深入<sup>[14]</sup>。本文研讨重构一类矩阵信号的理论问题和数值验证, 故不在此详述算法方面的工作。

这里将针对具有稀疏—平坦结构特征的矩阵信号,研究其重构条件。称矩阵  $\mathbf{X} = [\mathbf{x}_1, \dots, \mathbf{x}_n]$  具有  $s$  列稀疏性和  $l_1$  列平坦性是指其每个列向量  $\mathbf{x}_j$  至多具有  $s$  个非零分量、同时所有列向量的  $l_1$  范数具有相同的数值。这类信号模型出现在带特殊波形的合成孔径雷达及高维超声成像等新应用领域<sup>[4, 15]</sup>, 信号的列稀疏性反映散射体在波束范围内的稀疏性, 平坦性反映波形包络定常或缓变的特点。

本文工作由两部分组成, 第一部分将凸优化理论与范数的次微分  $\partial \|\mathbf{X}\|_1$  的结构特征相结合来建立一组重构条件, 第二部分基于  $\partial \|\mathbf{X}\|_1$  的结构特征与高维概率理论, 来建立重构这类矩阵信号的观测空间维数须满足的下界, 最后讨论数值验证。

### 1 基本问题、相关概念和理论基础

首先说明符号及约定: 为行文简洁, 本文仅讨论实数域上的方阵信号, 但所有结果对复数域上的矩阵信号同样成立。文中的向量  $\mathbf{x}$  指列向量,  $\mathbf{x}^T$  表示行向量。 $\langle \mathbf{x}, \mathbf{y} \rangle$  表示两个向量  $\mathbf{x}$  和  $\mathbf{y}$  的 Euclid 标量积,  $\langle \mathbf{X}, \mathbf{Y} \rangle$  表示两个矩阵  $\mathbf{X}$  和  $\mathbf{Y}$  的 Euclid 标量积, 即迹形式  $\text{tr}(\mathbf{X}^T \mathbf{Y})$ , 相应的 Frobenius 范数  $\langle \mathbf{X}, \mathbf{X} \rangle^{1/2}$  记为  $\|\mathbf{X}\|_F$ 。

一个向量或矩阵的支集 (support), 是指其非零元素下标的集合, 以  $\text{supp}(\cdot)$  表示。一个线性算子  $\Psi$  的核空间  $\ker \Psi$  是指被该算子映射为零元素的那些元素的集合, 即  $\ker \Psi \equiv \{\mathbf{X}; \Psi(\mathbf{X}) = \mathbf{O}\}$ 。

对实向量  $\mathbf{x} = [x_1, \dots, x_n]^T$ ,  $\text{sgn } \mathbf{x}$  表示相应的向量  $(\text{sgn } x_1, \dots, \text{sgn } x_n)^T$ , 其中的分量是符号函数: 若  $x_j > 0$  则  $\text{sgn } x_j = 1$ ; 若  $x_j < 0$  则  $\text{sgn } x_j = -1$ ;  $\text{sgn } 0 = 0$ 。

给定正整数  $s$ , 用记号  $\sum_s^{n \times n}$  表示逐列  $s$  稀疏的  $n \times n$  矩阵  $\mathbf{X}$  的集合, 即这些矩阵的每一列至多只有  $s$  个非零分量, 以下称之为  $s$  列稀疏信号。

如果矩阵  $\mathbf{X} = [\mathbf{x}_1, \dots, \mathbf{x}_n]$  所有列向量的  $l_1$  范数  $\|\mathbf{x}_j\|_1$  数值相同, 则称之为  $l_1$  列平坦矩阵。

设  $S \equiv S_1 \cup \dots \cup S_n$  是矩阵元素的下标集合, 其中每个  $S_j$  是第  $j$  列元素的下标集合  $\{(i, j): i = 1, \dots, n\}$  的子集且基数  $|S_j| \leq s$ , 以下称这样的集合  $S$  为矩阵信号的  $s$  稀疏模式。记号  $\sum_s^{n \times n}(S)$  表示以  $S$  为支集的  $n \times n$  矩阵的集合, 即集合  $\{\mathbf{M}; \text{若}(i, j) \text{ 不属于 } S \text{ 则 } M_{ij} = 0\}$ 。显然, 对给定的稀疏模式  $S$ ,  $\sum_s^{n \times n}(S)$  是一个维数不超过  $ns$  的线性空间且记

$$\sum_s^{n \times n} = \cup_s \sum_s^{n \times n}(S)$$

若  $S$  是一个  $s$  稀疏模式,  $\mathbf{M}$  是一个矩阵, 则  $\mathbf{M}_S$  表示这样的矩阵: 若  $(i, j)$  属于  $S$  则其  $(i, j)$  元素 =  $M_{ij}$ , 否则为 0。对  $S$  的列下标集  $S_j$  和向量  $\mathbf{m}$ ,  $\mathbf{m}|_{S_j}$  的含义类似。

如果  $X_k$  是一组随机变量,  $p(x)$  是某个概率密度函数, 则记号  $X_k \sim \text{iid } p(x)$  表示该组随机变量  $X_k$  彼此概率独立并遵循相同的概率分布  $p(x)$ 。

#### 1.1 基本问题

本文研究基于凸优化方法求解具有  $s$  列稀疏且  $l_1$  列平坦特征的矩阵信号  $\mathbf{X} = [\mathbf{x}_1, \dots, \mathbf{x}_n]$  的重构问题, 目标函数是  $\|\mathbf{X}\|_1 \equiv \max_j \|\mathbf{x}_j\|_1$ 。列平坦性意味着对所有的  $j$  有  $\|\mathbf{X}\|_1 = \|\mathbf{x}_j\|_1$ , 拟处理的凸优化问题为以下形式, 记为  $\text{MP}_{y, \Psi, \eta}^{(\alpha)}$ :

$$\min \|\mathbf{Z}\|_1 \quad \text{s.t. } \mathbf{Z} \in \mathbb{R}^{n \times n}, \|\mathbf{y} - \Psi(\mathbf{Z})\|_\alpha \leq \eta \quad (1)$$

其中, 观测数据  $\mathbf{y}$  是空间  $\mathbb{R}^m$  中的向量,  $\|\cdot\|_\alpha$  是给定在该空间上的某个范数, 例如  $l_2$ -范数;  $\Psi: \mathbb{R}^{n \times n} \rightarrow \mathbb{R}^m$  是线性算子。这时存在某个(待重构的真实矩阵信号)  $\mathbf{X}$  满足  $\mathbf{y} = \Psi(\mathbf{X}) + \mathbf{e}$ ,  $\mathbf{e}$  是观测误差向量且满足  $\|\mathbf{e}\|_\alpha \leq \eta$ 。该观测模型等价的分量形式为  $y_i = \langle \Psi_i, \mathbf{X} \rangle + e_i$ , 对每个  $i = 1, \dots, m$  有  $\Psi_i \in \mathbb{R}^{n \times n}$ 。

注意对问题  $\text{MP}_{y, \Psi, \eta}^{(\alpha)}$  仅须考虑误差界  $\eta: 0 \leq \eta < \|\mathbf{y}\|_\alpha$  的情况, 若超出该界则显然问题  $\text{MP}_{y, \Psi, \eta}^{(\alpha)}$  仅有平凡的最优解  $\mathbf{X}^* = \mathbf{O}$ 。

#### 1.2 相关概念

在此汇集对本文工作至关重要的凸几何理论中的某些概念与关系<sup>[16]</sup>。为行文简洁, 均以向量形式来表述, 但都不难直接推广为针对矩阵的相应形式。

空间  $\mathbb{R}^n$  中的锥  $C$  是这样—个区域: 对任何正数  $t > 0$ ,  $tC$  恒为  $C$  的子域。对空间  $\mathbb{R}^n$  中的某个区域  $K$ , 其对偶域  $K^* \equiv \{\mathbf{v}; \text{对 } K \text{ 中所有的 } \mathbf{x} \text{ 有 } \langle \mathbf{x}, \mathbf{v} \rangle \leq 0\}$ 。不难验证  $K^*$  总是一个凸锥。

对一个凸函数  $\varphi(\mathbf{x})$ , 有以下的相关集合:

$$\begin{aligned} & \text{切锥 } D(\varphi, \mathbf{x}) \\ & \equiv \{\mathbf{u}; \text{存在 } t > 0 \text{ 使 } \varphi(\mathbf{x} + t\mathbf{u}) \\ & \leq \varphi(\mathbf{x})\} \\ & \text{次微分 } \partial\varphi(\mathbf{x}) \\ & \equiv \{\mathbf{u}; \text{对所有的 } \mathbf{v} \text{ 有 } \varphi(\mathbf{v}) \geq \varphi(\mathbf{x}) + \langle \mathbf{v} - \mathbf{x}, \mathbf{u} \rangle\} \end{aligned}$$

且两者间成立重要的对偶关系  $D(\varphi, \mathbf{x})^* = \cup_{t > 0} t\partial\varphi(\mathbf{x})$ 。

对任何范数  $\|\cdot\|$ , 其对偶范数  $\|\cdot\|^*$  定义为  $\|\mathbf{v}\|^* \equiv \max \{\langle \mathbf{x}, \mathbf{v} \rangle; \|\mathbf{x}\| \leq 1\}$ , 例如  $\|\mathbf{X}\|_1^* = \sum_j \|\mathbf{x}_j\|_\infty$ 。不难验证任何范数  $\|\cdot\|$  的次微分总成立

恒等式:

$$\partial|\mathbf{x}| = \{\mathbf{v}: |\mathbf{x}| = \langle \mathbf{x}, \mathbf{v} \rangle \text{ 且 } |\mathbf{v}|^* \leq 1\} \quad (2)$$

设  $K$  是线性空间  $L$  中的锥,  $\Psi$  是定义在该空间上的某个线性算子,  $|\cdot|_\beta$  是  $L$  上的某个范数,  $|\cdot|_\alpha$  是  $\Psi$  的像空间上的某个范数,  $\Psi$  相对于这两个范数的、 $K$  限定性的最小奇异值定义为

$$\lambda_{\min, \alpha, \beta}(\Psi; K) \equiv \min\{|\Psi\mathbf{v}|_\alpha: \mathbf{v} \text{ 属于 } K \text{ 且 } |\mathbf{v}|_\beta = 1\} \quad (3)$$

如果范数  $|\cdot|_\beta$  和  $|\cdot|_\alpha$  均为  $l_2$ -范数或均为 Frobenius 范数, 则  $\lambda_{\min, \alpha, \beta}(\Psi; K)$  简记为  $\lambda_{\min}(\Psi; K)$ 。

设  $K$  是赋范空间  $(L, |\cdot|_\beta)$  中的一个锥 (未必为凸), 其高斯宽度定义为

$$w_\beta(K) \equiv E_g(\max\{\langle \mathbf{g}, \mathbf{v} \rangle: \mathbf{v} \text{ 属于 } K \text{ 且 } |\mathbf{v}|_\beta = 1\}) \quad (4)$$

其中,  $\mathbf{g}$  是  $L$  上具有标准高斯分布的随机向量。若  $|\cdot|_\beta$  为向量的  $l_2$  或矩阵的 Frobenius 范数, 则  $w_\beta(K)$  简记为  $w(K)$ 。

### 1.3 理论基础

本文需用到凸几何及高维概率理论某些较新近的结果, 在此以几个定理的形式做集中表述。这些结果的原始表述都是向量而非矩阵形式, 为适合本文处理矩阵信号的需要, 以下在形式上作了重新表述。定理 1 的原始证明参见文[17], 定理 2 和定理 3 的证明参见文[18], 定理 4 的证明以及对这些结果中最优常数的讨论参见文[19]。

**定理 1** 1) 设矩阵信号  $\mathbf{X} \in \mathbb{R}^{n \times n}$  满足观测方程  $\mathbf{y} = \Psi(\mathbf{X})$ ,  $\mathbf{X}^*$  是问题  $\text{MP}_{\mathbf{y}, \Psi, \eta}^{(\alpha)}$ :  $\eta = 0$  (即无测量噪声情况下的优化问题) 的最优解, 则  $\mathbf{X}^* = \mathbf{X}$  当且仅当  $\ker \Psi \cap D(\|\cdot\|_1, \mathbf{X}) = \{\mathbf{0}\}$ 。2) 设矩阵信号  $\mathbf{X} \in \mathbb{R}^{n \times n}$  满足观测方程  $\mathbf{y} = \Psi(\mathbf{X}) + \mathbf{e}$ , 其中观测噪声有上界  $|\mathbf{e}|_\alpha \leq \eta$  ( $\eta > 0$ ),  $\mathbf{X}^*$  是问题  $\text{MP}_{\mathbf{y}, \Psi, \eta}^{(\alpha)}$  的最优解,  $|\cdot|_\beta$  是某种给定的矩阵范数, 则

$$\|\mathbf{X}^* - \mathbf{X}\|_\beta \leq 2\eta / \lambda_{\min, \alpha, \beta}(\Psi; D(\|\cdot\|_1, \mathbf{X})) \quad (5)$$

**定理 2** 设  $\mathbf{X}$  和  $\mathbf{X}^*$  分别表示真实矩阵信号和问题  $\text{MP}_{\mathbf{y}, \Psi, \eta}^{(2)}$  的最优解,  $\Psi_{kij} \sim \text{iid} N(0, 1)$ , 则对任何  $t > 0$  有:

$$P(\|\mathbf{X}^* - \mathbf{X}\|_F \leq 2\eta / (m - w(D(\|\cdot\|_1, \mathbf{X})) - t) +) \geq 1 - \exp(-t^2/2)$$

其中,  $(u)_+ \equiv \max(u, 0)$ 。

**定理 3** 设  $F$  是凸函数且在  $\mathbf{X}$  处的次微分  $\partial F(\mathbf{X})$  不含零矩阵, 则:

$$w_\beta^2(D(F, \mathbf{X})) \leq E_G(\min\{\|\mathbf{G} - t\mathbf{V}\|_{\beta^*}^2: t > 0, \mathbf{V} \text{ 属于 } \partial F(\mathbf{X})\})$$

其中,  $|\cdot|_{\beta^*}$  是  $|\cdot|_\beta$  的对偶范数,  $\mathbf{G}$  是矩阵元

$\mathbf{G}_{ij} \sim \text{iid} N(0, 1)$  的随机矩阵。特别地, 若  $|\cdot|_\beta$  是矩阵的 Frobenius 范数  $|\cdot|_F$  则:  $w^2(D(F, \mathbf{X})) \leq E_G(\min\{\|\mathbf{G} - t\mathbf{V}\|_F^2: t > 0, \mathbf{V} \text{ 属于 } \partial F(\mathbf{X})\})$ 。

以上定理将用以估计高斯宽度  $w_\beta(D(F, \mathbf{X}))$  的上界。

**定理 4** 设  $\mathbf{X}$  和  $\mathbf{X}^*$  分别表示真实矩阵信号和问题  $\text{MP}_{\mathbf{y}, \Psi, \eta}^{(2)}$  的最优解, 观测数据的分量  $y_k = \langle \Psi_k, \mathbf{X} \rangle + e_k$ ,  $\Psi_k \sim \text{iid} \Psi$  且随机矩阵  $\Psi$  满足以下条件: 1)  $E(\Psi) = 0$ ; 2) 存在常数  $\alpha > 0$ , 使所有的  $\mathbf{U}: \|\mathbf{U}\|_F = 1$  满足  $\alpha \leq E(\langle \Psi, \mathbf{U} \rangle)$ ; 3) 存在常数  $\sigma > 0$  使  $P(|\langle \Psi, \mathbf{U} \rangle| \geq t) \leq 2\exp(-t^2/2\sigma^2)$ 。记  $\rho \equiv \sigma/\alpha$ , 则对任何  $t > 0$  有:

$$P(\|\mathbf{X} - \mathbf{X}^*\|_F \leq 2\eta / (c_1 \alpha \rho^{-2} m - c_2 \sigma w(D(\|\cdot\|_1, \mathbf{X})) - \alpha t)) \geq 1 - \exp(-c_3 t^2) \quad (6)$$

其中,  $c_1, c_2, c_3$  均为绝对常数。

## 2 主要结果和一个基本工具

### 2.1 主要结果

第 3 节分别在无观测误差和有观测误差的情况下, 针对保证重构的稀疏—平坦矩阵信号  $\mathbf{X}$  具有误差稳定性, 建立观测算子  $\Psi$  需满足的充分条件。该组结果的核心是定理 7 关于观测算子的近似等距性质, 通过适当的参数组合, 该性质能衍生出针对各种情况的重构条件, 因此和重构向量信号的  $\text{min-}l_1$  理论所建立的近似等距 (RIP) 性质类似, 具有广泛的实用性。

第 4 节建立随机观测情况下观测空间维数  $m$  (观测数据的数量) 需满足的下界及数值验证。该界用以保证稀疏—平坦矩阵信号以高概率被正确重构。所得结果优于将向量形式的普遍理论直接推广为矩阵信号的  $\text{min-}l_1$  理论所能得到的结果, 这种优点一方面缘于范数  $\|\cdot\|_1$  较之  $l_1$  范数更能准确表达矩阵信号的稀疏—平坦特性, 另一方面缘于本文具体针对矩阵信号 (而非简单将其视为向量) 的分析方法。

### 2.2 基本工具: 矩阵范数 $\|\mathbf{Z}\|_1$ 的次微分结构

凸范数的次微分 (参见 1.2 节) 是做相应凸分析的关键工具之一。以下引理完整刻画了次微分  $\partial\|\mathbf{X}\|_1$  的结构。

**引理 1** 矩阵范数  $\|\mathbf{Z}\|_1 \equiv \max_j |z_j|_1$  在矩阵  $\mathbf{X} = [\mathbf{x}_1, \dots, \mathbf{x}_n]$  上的次微分  $\partial\|\mathbf{X}\|_1$  是矩阵集合:  $\partial\|\mathbf{X}\|_1 = \{[\lambda_1 \xi_1, \dots, \lambda_n \xi_n]: \text{对所有的 } j \text{ 有 } \xi_j \text{ 属于}$

$\partial|x_j|_1, \lambda_j \geq 0, \lambda_1 + \dots + \lambda_n = 1$  且对  $j: |x_j|_1 < \max_k |x_k|_1$  有  $\lambda_j = 0$

**证明** 首先容易验证上述集合中的元素确实属于  $\partial\|X\|_1$ , 这是因为对具有上述性质的矩阵  $W = [\lambda_1 \xi_1, \dots, \lambda_n \xi_n]$  有

$$\begin{aligned} \langle W, X \rangle &= \sum_j \lambda_j \langle \xi_j, x_j \rangle \\ &= \sum_j \lambda_j |x_j|_1 \\ &= \|X\|_1 \sum_j \lambda_j \\ &= \|X\|_1 \end{aligned}$$

且  $\|\cdot\|_1$  的对偶范数  $\|W\|_1^* = \sum_j \lambda_j \|\xi_j\|_\infty \leq \sum_j \lambda_j = 1$ , 因此按照次微分的性质(参见 1.2 节),  $W$  属于  $\partial\|X\|_1$ .

接下来证明若  $W$  属于  $\partial\|X\|_1$ , 则  $W$  确实具有该引理对集合元素所描述的性质. 设  $W$  有列向量的形式  $W = [\mu_1, \dots, \mu_n]$ , 于是对所有的  $Y = [y_1, \dots, y_n]$  有  $\|Y\|_1 \geq \|X\|_1 + \langle Y - X, W \rangle$ , 即:

$$\max_j |y_j|_1 \geq \max_j |x_j|_1 + \sum_j \langle y_j - x_j, \mu_j \rangle \quad (7)$$

令  $\mu_j = |\mu_j|_\infty \xi_j$  (于是  $|\xi_j|_\infty = 1$  若  $\mu_j \neq 0$ ) 则式(7)表为  $\max_j |y_j|_1 \geq \max_j |x_j|_1 + \sum_j |\mu_j|_\infty \langle y_j - x_j, \xi_j \rangle$ . 注意到对每个非零的第  $j$  列总可选取  $i_j$  使  $|\xi_j(i_j)| = 1$  并设置相应的向量  $e_j^*$ , 其分量  $e_j^*(i_j) = \text{sgn } \xi_j(i_j)$  而其他分量皆为零, 取  $y_j = x_j + e_j^*, j = 1, \dots, n$ , 代入式(7)得:

$$\begin{aligned} &1 + \max_j |x_j|_1 \\ &\geq \max_j |y_j|_1 \\ &\geq \max_j |x_j|_1 + \sum_j |\mu_j|_\infty \langle e_j^*, \xi_j \rangle \\ &= \max_j |x_j|_1 + \sum_j |\mu_j|_\infty |\xi_j|_\infty \\ &= \max_j |x_j|_1 + \sum_j |\mu_j|_\infty \end{aligned}$$

即  $1 \geq \sum_j |\mu_j|_\infty$ .

现在固定某个列指标  $i$ , 对所有的  $j \neq i$  取  $y_j = x_j$  并取  $y_i$  是满足  $|y_i|_1 \leq |x_i|_1$  的向量, 将这样的  $y_1, \dots, y_n$  代入式(7)得

$$\begin{aligned} &\max_j |x_j|_1 \\ &\geq \max_j |y_j|_1 \\ &\geq \max_j |x_j|_1 + \sum_j \langle y_j - x_j, \mu_j \rangle \\ &= \max_j |x_j|_1 + |\mu_j|_\infty \langle y_i - x_i, \xi_i \rangle \end{aligned}$$

即  $\langle y_i - x_i, \xi_i \rangle \leq 0$ , 亦即对所有的  $y_i: |y_i|_1 \leq |x_i|_1$  有  $\langle x_i, \xi_i \rangle \geq \langle y_i, \xi_i \rangle$  进而有  $\langle x_i, \xi_i \rangle \geq \sup\{\langle y_i, \xi_i \rangle: |y_i|_1 \leq |x_i|_1\} = |x_i|_1 |\xi_i|_\infty = |x_i|_1$ , 从而  $\langle x_i, \xi_i \rangle = |x_i|_1$ .

该结论连同  $|\xi_i|_\infty = 1$  意味着  $\xi_i$  属于  $\partial|x_i|_1$ .

至此已证明  $\partial\|X\|_1$  中的  $W$  恒有形式  $[\lambda_1 \xi_1, \dots, \lambda_n \xi_n]$ , 其中每个  $\xi_j$  属于  $\partial|x_j|_1$ 、对所有的  $j$  有  $\lambda_j \geq 0$  且  $\lambda_1 + \dots + \lambda_n \leq 1$ . 由于  $\|X\|_1 = \langle W, X \rangle = \sum_j \lambda_j \langle \xi_j, x_j \rangle = \sum_j \lambda_j |x_j|_1 \leq \max_j |x_j|_1 \sum_j \lambda_j \leq \|X\|_1$ , 因此必有  $\lambda_1 + \dots + \lambda_n = 1$  且对  $j: |x_j|_1 < \max_k |x_k|_1$  必有  $\lambda_j = 0$ . 证毕.

如果显式地结合  $\partial|x_j|_1$  的形式, 就得到次微分更完备的表达形式

$$\partial\|X\|_1 = \{[\lambda_1 \xi_1, \dots, \lambda_n \xi_n]: \text{对所有的 } j \text{ 有 } |\xi_j|_\infty \leq 1, \xi_j(i) = \text{sgn } X_{ij} \text{ 若 } X_{ij} \neq 0; \lambda_j \geq 0, \lambda_1 + \dots + \lambda_n = 1, \text{ 对 } j: |x_j|_1 < \max_k |x_k|_1 \text{ 有 } \lambda_j = 0\} \quad (8)$$

### 3 基于求解 $\min\|\cdot\|_1$ 凸优化问题重构矩阵信号的充分条件

本节分别针对观测数据无误差、观测数据有误差等不同情形, 建立基于求解凸优化问题  $MP_{y, \Psi, \eta}^{(\alpha)}$  来重构稀疏—平坦矩阵信号的各种充分条件. 有关结果的证明与经典理论相应结果的证明<sup>[1]</sup> 在逻辑上类似, 但需仔细处理矩阵范数  $\|\cdot\|_1$  带来的一些复杂性.

#### 3.1 无观测误差的情形

本节研究当线性观测算子  $\Psi$  满足什么条件时, 矩阵信号  $X$  能通过求解以下问题被精确地重构:

$$MP_{y, \Psi, 0}: \min \|Z\|_1 \quad \text{s.t. } Z \in \mathbb{R}^{n \times n}, y = \Psi(Z)$$

在此将问题适当放宽, 研究算子  $\Psi$  满足什么条件时, 对那些并非严格具有列稀疏和列平坦特征的矩阵信号  $X$ ,  $MP_{y, \Psi, 0}$  的最优解  $X^*$  能逼近  $X$  到什么程度. 然后, 作为特殊情况引出结论. 逼近误差以逐列  $l_1$  误差的最大值来度量. 以下的引理 2 建立逼近误差的上界及相应的充分条件, 引理 3 给出保障最优解具有  $l_1$  列平坦性的一个充分条件.

**引理 2**  $s$  是给定的正整数,  $\Psi$  是给定的线性算子, 满足参量为  $\rho$  的  $s\text{-}\|\cdot\|_1$  稳定性条件, 即存在一个常数  $0 < \rho < 1$ , 使得任何  $M = [m_1, \dots, m_n] \in \ker \Psi$  在任何  $s$  稀疏模式  $S = S_1 S_n$  ( $|S_j| \leq s$ ) 上恒满足

$$\|M_S\|_1 \leq \rho \|M_{\sim S}\|_1 \quad (9)$$

在这条件下, 问题  $MP_{y, \Psi, 0}$  (其中  $y = \Psi(X)$ ) 的任何可行解  $V = [v_1, \dots, v_n]$  均满足不等式

$$\|V - X\|_1 \leq (1 + \rho) (\max_j (|v_j|_1 - |x_j|_1) + 2 \max_j \sigma_s(x_j)_1) / (1 - \rho) \quad (10)$$

其中,  $\sigma_s(x)_1 \equiv |x|_1 - (|x(i_1)| + \dots + |x(i_s)|)$ ,  $x(i_1), \dots, x(i_s)$  是向量  $x$  的前  $s$  个绝对值最大的分量.

引理 3 涉及观测算子  $\Psi$  的对偶算子。如果  $\Psi$  的分量表达式为  $\Psi(V)_k = \text{tr}(\Psi_k^T V)$ , 其中  $\Psi_k$  是一组矩阵,  $k=1, \dots, m$ , 则  $\Psi$  的对偶算子有表达式  $\Psi^T(z) = \Psi_k: \mathbb{R}^m \rightarrow \mathbb{R}^{n \times n}$ 。此外, 引理 3 对存在观测观测误差的情况 ( $\eta > 0$ ) 也成立。

**引理 3** 给定正整数  $s$  和  $s$  稀疏模式  $S = S_1 \cup \dots \cup S_n$  (其中每个  $|S_j| \leq s$ ) 以及线性算子  $\Psi$ ,  $\Psi^T$  表示其对偶算子,  $\Psi_S$  表示  $\Psi$  限定在子空间  $\sum_s^{n \times n}(S)$  上的映射 (因而  $\Psi_S(X) = \Psi(X_S)$ )。如果对任何  $v \neq 0$ , 矩阵  $\Psi_S^T(v)$  不含列向量  $\mathbf{0}$ , 则  $\text{MP}_{y, \Psi, \eta}$  的支集含于  $S$  的最优解  $X^*$  必是  $l_1$  列平坦的。

在证明之前, 先综合以上结果做一个重要推论。当观测算子满足引理 3 的条件时, 在式(10)中取  $Z$  为最优解  $X^*$ , 就得到  $X^*$  逼近实际信号  $X$  的误差上界, 其中  $\max_j \sigma_s(x_j)_1$  和  $\max_j (|x_j^*|_1 - |x_j|_1)$  分别度量信号偏离  $s$  列稀疏性和  $l_1$  列平坦性的程度。如果  $X$  是  $s$  列稀疏的, 则  $\max_j \sigma_s(x_j)_1 = 0$ ; 如果是  $l_1$  列平坦的, 则每列均有  $|x_j|_1 = \|X\|_1$  故  $\max_j (|x_j^*|_1 - |x_j|_1) = \|X^*\|_1 - \|X\|_1 \leq 0$ , 由此得到结论:

**定理 5** 若观测算子  $\Psi$  满足引理 1 和引理 2 中的条件, 则问题  $\text{MP}_{y, \Psi, 0}$  的最优解  $X^*$  能精确重构任何  $s$ -列稀疏且  $l_1$  列平坦的矩阵信号  $X$ , 即  $X^* = X$ 。

引理 2 对任意信号的逼近误差的上界估计本身也具有独立的意义, 这是因为在实际应用中更多存在的是近似、而非理想的列稀疏且列平坦信号, 例如少数分量承载绝大部分能量、波形包络存在小幅变化的信号。式(10)为这类信号的重构误差提供实用的上界估计, 表达重构问题的解在实际信号偏离稀疏—平坦特征时的稳定程度。对重构误差的度量与表达信号结构特征采用相同的范数  $\|\cdot\|_1$ , 符合模型一致性<sup>[20]</sup>, 并且式(10)所表达的是对每列误差的上界控制, 优于采用  $l_1$  矩阵范数  $\|X\|_1 = \sum_j |x_j|_1$  求解  $\min \|X\|_1$  问题的逼近效果, 后者得出的是相对较粗的误差度量  $\|X^* - X\|_1$  受控于更高的界  $\sigma_{ms}(X)_1$  (参见文[1])。

以下证明上述引理, 为此需要一个简单的工具, 即对下标的任何子集  $\Delta$  和任何向量  $u$  与  $v$  有<sup>[1]</sup>:

$$\begin{aligned} & |(u-v)_{-\Delta}|_1 \\ & \leq |v|_1 - |u|_1 + |(u-v)_{\Delta}|_1 + 2|u_{-\Delta}|_1 \end{aligned} \quad (11)$$

引理 2 的证明对  $y = \Psi(X)$ ,  $\text{MP}_{y, \Psi, 0}$  的可行解  $V = [v_1, \dots, v_n]$  有  $M = [m_1, \dots, m_n]$  属于  $\ker \Psi$  使  $V = M + X$ , 对每个列向量  $v_j$  和  $x_j$  应用(11)得

$|m_{j_1} \sim s_j|_1 \leq |v_j|_1 - |x_j|_1 + |m_{j_1 \sim s_j}|_1 + 2|x_{j_1 \sim s_j}|_1$  进而结合条件(9)得到

$$\begin{aligned} & \|M_{-s}\|_1 \equiv \max_j |m_{j_1 \sim s_j}|_1 \\ & \leq \max_j (|v_j|_1 - |x_j|_1) + \|M_S\|_1 + 2\max_j |x_{j_1 \sim s_j}|_1 \\ & \leq \max_j (|v_j|_1 - |x_j|_1) + \rho \|M_{-s}\|_1 + 2\max_j |x_{j_1 \sim s_j}|_1 \\ & \text{即 } \|M_{-s}\|_1 \leq (2\max_j |x_{j_1 \sim s_j}|_1 + \max_j (|v_j|_1 - |x_j|_1)) / \\ & (1 - \rho), \text{ 从而对任何 } s \text{ 稀疏模式 } S \text{ 有} \end{aligned}$$

$$\begin{aligned} & \|M\|_1 = \|M_S\|_1 + \|M_{-s}\|_1 \\ & \leq (1 + \rho) \|M_{-s}\|_1 \\ & \leq (1 + \rho) (2\max_j |x_{j_1 \sim s_j}|_1 + \\ & \max_j (|v_j|_1 - |x_j|_1)) / (1 - \rho) \end{aligned}$$

由于  $\min_s \max_j |x_{j_1 \sim s_j}|_1 = \max_j \sigma_s(x_j)_1$  故该式蕴含结论(10)。

证毕。

引理 3 的证明首先考虑有观测误差的情形, 即问题  $\text{MP}_{y, \Psi, \eta}$ :

$$\min \|V\|_1 \quad \text{s.t. } V, |y - \Psi_S(V)|_2 \leq \eta$$

其中,  $\eta > 0$ 。

由于  $\|\cdot\|_1$  和约束式  $|y - \Psi_S(\cdot)|_2$  均为自变量的凸函数, 根据凸优化的普遍理论, 存在乘子  $\gamma^* > 0$  和矩阵  $Z^*$  属于  $\partial \|X^*\|_1$  满足一阶极值条件:

$$Z^* + \gamma^* \Psi_S^T(\Psi_S(X^*) - y) = \mathbf{0} \quad (12)$$

$$|y - \Psi_S(X^*)|_2 = \eta$$

由于  $y - \Psi_S(X^*) \neq 0$ , 因而在定理的条件下矩阵  $Z^* = \gamma^* \Psi_S^T(y - \Psi_S(X^*))$  不包含列向量  $\mathbf{0}$ , 由引理 1 这意味着  $X$  的列向量均满足  $|x_j^*|_1 = \max_k |x_k^*|_1$ , 即  $X$  是  $l_1$  列平坦的。

再考虑无观测误差的情形, 即问题  $\text{MP}_{y, \Psi, 0}$ :  $\min \|V\|_1$  s.t.  $V, y = \Psi_S(V)$ 。根据凸优化的普遍理论, 存在乘子向量  $u$  和矩阵  $Z^*$  属于  $\partial \|X^*\|_1$  满足一阶极值条件  $Z^* + \Psi_S^T(u) = \mathbf{0}$ 。若  $u \neq 0$  则在定理的条件下  $Z^*$  不含  $\mathbf{0}$  列, 由前面同样的理由得  $X$  是  $l_1$  列平坦的。若  $u = 0$  则极值条件蕴含  $Z^* = \mathbf{0}$ , 但根据引理 1 这是不可能的, 除非  $X^* = \mathbf{0}$ 。证毕。

### 3.2 有观测误差的情形

当存在观测误差时, 问题的最优解不再能精确等于实际信号, 这时更恰当的要求是两者的数值误差能被合理地控制, 即满足鲁棒性(Robustness)。

**定理 6** 给定正整数  $s$  和线性算子  $\Psi$ , 满足  $s$ - $\|\cdot\|_1$  鲁棒性条件, 即存在常数  $0 < \rho < 1$  和  $\beta > 0$  使得任何  $n \times n$  矩阵  $M$  在任何  $s$  稀疏模式  $S = S_1 \cup \dots \cup S_n$  上满足不等式:

$$\|M_S\|_1 \leq \rho \|M_{-s}\|_1 + \beta |\Psi(M)|_2 \quad (13)$$

在这条件下,若  $V = [v_1, \dots, v_n]$  是问题  $MP_{y, \Psi, \eta}$  的某个可行解,其中  $y = \Psi(X) + e$ ,  $X = [x_1, \dots, x_n]$  是实际信号,观测误差满足  $\|e\|_2 \leq \eta$ , 则必有不等式

$$\|V - X\|_1 \leq C(2\max_j \sigma_s(x_j)_1 + 2\beta\eta + \max_j (|v_j|_1 - |x_j|_1)) \quad (14)$$

其中,  $C \equiv (1 + \rho)/(1 - \rho)$ 。

若  $\Psi$  还进一步满足引理 3 的条件,设实际信号  $X$  是  $s$  列稀疏且  $l_1$  列平坦的,  $X^*$  是问题  $MP_{y, \Psi, \eta}$  的最优解,则有误差上界:

$$\|X^* - X\|_1 \leq 2\beta\eta(1 + \rho)/(1 - \rho) \quad (15)$$

**证明** 令  $V - X \equiv M = [m_1, \dots, m_n]$ , 将不等式(11)应用到列向量  $v_j$  和  $x_j$  得到  $|m_{j1-sj}|_1 \leq |v_j|_1 - |x_j|_1 + |m_{j1-sj}|_1 + 2|x_{j1-sj}|_1$ , 从而  $\|M_{-s}\|_1 \equiv \max_j |m_{j1-sj}|_1 \leq \max_j (|v_j|_1 - |x_j|_1) + \|M_s\|_1 + 2\max_j |x_{j1-sj}|_1 \leq \max_j (|v_j|_1 - |x_j|_1) + \rho \|M_{-s}\|_1 + 2\max_j |x_{j1-sj}|_1 + \beta \|\Psi(M)\|_2$  (这里已应用条件(13)), 即:

$$\|M_{-s}\|_1 \leq (2\max_j |x_{j1-sj}|_1 + \max_j (|v_j|_1 - |x_j|_1) + \beta \|\Psi(M)\|_2)/(1 - \rho)$$

因此,

$$\begin{aligned} \|M\|_1 &\leq \|M_s\|_1 + \|M_{-s}\|_1 \\ &\leq (1 + \rho) \|M_{-s}\|_1 + \beta \|\Psi(M)\|_2 \\ &\leq (1 + \rho) (2\max_j |x_{j1-sj}|_1 + \max_j (|v_j|_1 - |x_j|_1) + 2\beta \|\Psi(X)\|_2)/(1 - \rho) \end{aligned}$$

对每个  $s$  稀疏模式  $S$  均成立。由于  $\min_j \max_j |x_{j1-sj}|_1 = \max_j \sigma_s(x_j)_1$  因而取使得误差界  $\max_j |x_{j1-sj}|_1$  达到最小的那个稀疏模式  $S$  就得到式(14)。

若  $\Psi$  还满足引理 3 的条件、实际信号  $X$  是  $s$  列稀疏且  $l_1$  列平坦的,取  $V$  为最优解  $X^*$ , 这时  $X^*$  也是  $l_1$  列平坦的,从而式(14)蕴含式(15)。证毕。式(15)表明当存在观测误差时,矩阵信号的重构误差仍可以被逐列控制(而非如  $\min$ - $l_1$  理论那样仅被平均地控制),符合模型一致性概念<sup>[20]</sup>,特别在实际信号为稀疏—平坦的情况下,重构误差  $\|X^* - X\|_1 \leq 2\beta\eta(1 + \rho)/(1 - \rho) = O(\eta)$  具有线性收敛速率。

### 3.3 观测算子 $\Psi$ 的近似等距条件

在重构向量信号的  $\min$ - $l_1$  理论中,观测算子作用在稀疏信号集合上的近似等距性质(RIP)是一类重要条件,由此可派生出诸多其他条件<sup>[11]</sup>。本节给出针对矩阵信号的  $\min$ - $\|\cdot\|_1$  理论的类似性质,并证明在适当的参数组合之下该性质衍生出上节所建立的鲁棒性条件。

**定理 7** 给定正整数  $s$  和线性算子  $\Psi: \mathbb{R}^{n \times n} \rightarrow \mathbb{R}^m$ , 设存在正数  $0 < \delta_s < 1$  和  $\Delta_s > 0$  满足以下条件:

1)  $(1 - \delta_s) \|V\|_F^2 \leq \|\Psi(V)\|_2^2 \leq (1 + \delta_s) \|V\|_F^2$  对所有  $s$  列稀疏的矩阵  $V$  成立;

2)  $|\langle \Psi(V), \Psi(W) \rangle| \leq (\Delta_s/n) \sum_{j=1 \dots n} |v_j|_2 |w_j|_2$  对所有  $s$  列稀疏且支集不相交的矩阵  $V = [v_1, \dots, v_n]$  和矩阵  $W = [w_1, \dots, w_n]$  ( $\text{supp}(V) \cap \text{supp}(W) = \emptyset$ ) 成立。

在上述条件下,存在正值的常数  $\rho$  和  $\beta$  使得不等式

$$\|M_s\|_1 \leq \rho \|M_{-s}\|_1 + \beta \|\Psi(M)\|_2 \quad (16)$$

对任何  $n \times n$  矩阵  $M$  和任何  $s$  稀疏模式  $S$  均成立,且常数  $\rho$  和  $\beta$  可如下选择:

$$\rho \leq \Delta_s / (1 - \delta_s - \Delta_s/4) \quad (17)$$

$$\beta \leq s^{1/2} (1 + \delta_s)^{1/2} / (1 - \delta_s - \Delta_s/4)$$

特别地,若参数  $\delta_s$  和  $\Delta_s$  满足  $\delta_s + 5\Delta_s/4 < 1$ , 则  $\Psi$  满足鲁棒性条件:  $\rho < 1$ 。

**注 1** 条件(1)是标准的 RIP 性质,不难检验它对任何  $s$  稀疏且支集不相交的矩阵  $V$  和  $W$  蕴含不等式  $|\langle \Psi(V), \Psi(W) \rangle| \leq \delta_s \|V\|_F \|W\|_F$ , 但这还不足以得到结论(16),需要稍强的条件(2)来加强,该性质在直观上等价于算子  $\Psi$  对支集不相交(从而相互垂直)的矩阵信号具有接近正交变换的效果,“接近”的程度由新参数  $\Delta_s$  度量,  $\Delta_s$  数值越小则接近程度越高。因此,这里的近似等距性质稍强于  $\min$ - $l_1$  理论中观测算子的 RIP 性质。

**证明** 对任何一个  $n \times n$  矩阵  $M = [m_1, \dots, m_n]$  及每个  $j$ , 设  $|m_j(i_1)| \geq |m_j(i_2)| \geq \dots \geq |m_j(i_n)|$  是第  $j$  列  $m_j$  的分量按其绝对值的降序排序,令  $S_0(j) \equiv \{(i_1, j), \dots, (i_s, j)\}$  是列  $m_j$  的绝对值最大的前  $s$  个分量的下标,令  $S_1(j) \equiv \{(i_{1+s}, j), \dots, (i_{2s}, j)\}$  是列  $m_j$  的绝对值次大的  $s$  个分量的下标,如此等等。对每个  $k=0, 1, 2, \dots$  令  $S_k = \cup_{j=1}^n S_k(j)$ , 显然  $M = \sum_{k=0} M_{S_k}$ 。首先注意到,要证明式(17)对所有的  $S$  均成立,只要证明其对  $S_0$  成立,因此以下对  $S_0$  证明该结论。从定理的条件(1)出发:

$$\begin{aligned} (1 - \delta_s) \|M_{S_0}\|_F^2 &\leq \|\Psi(M_{S_0})\|_2^2 \\ &= \langle \Psi(M_{S_0}), \Psi(M) - \sum_{k \geq 1} \Psi(M_{S_k}) \rangle \\ &= \langle \Psi(M_{S_0}), \Psi(M) \rangle - \sum_{k \geq 1} \langle \Psi(M_{S_0}), \Psi(M_{S_k}) \rangle \\ &\leq \|\Psi(M_{S_0})\|_2 \|\Psi(M)\|_2 + \\ &\quad (\Delta_s/n) \sum_{1 \leq j \leq n} \sum_{k \geq 1} |m_{j1S_0(j)}|_2 |m_{j1S_k(j)}|_2 \end{aligned}$$

(应用柯西不等式及定理 7 条件(2))

$$\begin{aligned}
&\leq (1 + \delta_s)^{1/2} \|\mathbf{M}_{S_0}\|_F \|\Psi(\mathbf{M})\|_2 + \\
&\quad (\Delta_s/n) \|\mathbf{M}_{S_0}\|_F \sum_{1 \leq j \leq n} \sum_{k \geq 1} |\mathbf{m}_{j1Sk(j)}|_2 \\
&\quad (\text{应用定理 1 条件 1}) \text{ 和 } |\mathbf{m}_{j1S_0(j)}|_2 \leq \|\mathbf{M}_{S_0}\|_F \\
&\leq (1 + \delta_s)^{1/2} \|\mathbf{M}_{S_0}\|_F \|\Psi(\mathbf{M})\|_2 + (\Delta_s/n) \|\mathbf{M}_{S_0}\|_F \\
&\quad \sum_{1 \leq j \leq n} (s^{-1/2} |\mathbf{m}_{j1 \sim S_0(j)}|_1 + (1/4) |\mathbf{m}_{j1S_0(j)}|_2) \\
&\quad (\text{应用对任何 } 0 \leq a_s \leq \dots \leq a_2 \leq a_1 \text{ 成立 } (\sum_{1 \leq k \leq s} a_k^2)^{1/2} \\
&\leq s^{-1/2} \sum_{1 \leq k \leq s} a_k + (s^{1/2}/4) (a_1 - a_s)^{[1]} \text{ 和对任} \\
&\text{何 } j \text{ 成立} \\
&\quad \min_{1 \leq i \leq s} |\mathbf{m}_{j1Sk(j)}(i)| \geq \max_{1 \leq i \leq s} |\mathbf{m}_{j1Sk+1(j)}(i)| \\
&\text{这一事实)} \\
&= \|\mathbf{M}_{S_0}\|_F ((1 + \delta_s)^{1/2} \|\Psi(\mathbf{M})\|_2 + (s^{-1/2} \Delta_s/n) \\
&\quad \sum_{1 \leq j \leq n} |\mathbf{m}_{j1 \sim S_0(j)}|_1 + (\Delta_s/4n) \sum_{1 \leq j \leq n} |\mathbf{m}_{j1S_0(j)}|_2) \\
&\leq \|\mathbf{M}_{S_0}\|_F ((1 + \delta_s)^{1/2} \|\Psi(\mathbf{M})\|_2 + s^{-1/2} \Delta_s \max_j \\
&\quad |\mathbf{m}_{j1 \sim S_0(j)}|_1 + (\Delta_s/4n) n^{1/2} (\sum_{1 \leq j \leq n} |\mathbf{m}_{j1S_0(j)}|_2^2)^{1/2}) \\
&\quad (\text{对最后一项中的 } \sum_{n \geq j \geq 1} |\mathbf{m}_{j1S_0(j)}|_2 \text{ 应用柯西不} \\
&\text{等式}) \\
&= \|\mathbf{M}_{S_0}\|_F ((\delta_s + 1)^{1/2} \|\Psi(\mathbf{M})\|_2 + s^{-1/2} \Delta_s \|\mathbf{M}_{\sim S_0}\|_1 + \\
&\quad (\Delta_s/4n^{1/2}) \|\mathbf{M}_{S_0}\|_F) \\
&\quad \text{从以上不等式两端消去 } \|\mathbf{M}_{S_0}\|_F \text{ 后得到} \\
&(1 - \delta_s) \|\mathbf{M}_{S_0}\|_F \leq s^{-1/2} \Delta_s \|\mathbf{M}_{\sim S_0}\|_1 + (\Delta_s/4n^{1/2}) \|\mathbf{M}_{S_0}\|_F + \\
&\quad (\delta_s + 1)^{1/2} \|\Psi(\mathbf{M})\|_2 \\
&\text{即 } \|\mathbf{M}_{S_0}\|_F \leq (s^{-1/2} \Delta_s \|\mathbf{M}_{\sim S_0}\|_1 + (\delta_s + 1)^{1/2} \|\Psi(\mathbf{M})\|_2) / \\
&(1 - \delta_s - \Delta_s/4n^{1/2}). \text{ 最后注意到 } \|\mathbf{M}_{S_0}\|_1 = \max_j |\mathbf{m}_{j1S_0(j)}|_1 \\
&\leq s^{1/2} \max_j |\mathbf{m}_{j1S_0(j)}|_2 \leq s^{1/2} \|\mathbf{M}_{S_0}\|_F \text{ 并应用上述不等式,} \\
&\text{就得到对 } S_0 \text{ 成立的结论式(16)和式(17), 进而两者} \\
&\text{对任何 } S \text{ 均成立。证毕。}
\end{aligned}$$

#### 4 保障鲁棒性重构的观测空间维数

在建立了针对各种情况的重构条件后, 接下来要解决的问题就是寻求能满足各种重构条件的观测算子  $\Psi$ , 特别是寻求这样的算子, 在能够满足相应重构条件的同时, 其像空间的维数尽可能地低, 也就是在实际应用中所需观测数据尽可能地少。本节证明  $\min \|\mathbf{Z}\|_1$  凸优化问题在  $\Psi$  遵循某些典型的概率分布的情况下, 只要观测空间维数超过某个下界, 问题的最优解就能以任意接近于 1 的概率相对于 Frobenius 范数任意逼近实际的信号。

为阅读方便, 在此简述所考虑的问题  $\text{MP}_{y, \Psi, \eta}^{(2)}$ :  $\min \|\mathbf{Z}\|_1$  s.t.  $\mathbf{Z}, \|\mathbf{y} - \Psi(\mathbf{Z})\|_2 \leq \eta$ , 其中  $\mathbf{y} = \Psi(\mathbf{X}) + \mathbf{e}$ ,  $\mathbf{X}$  是某个未知的矩阵信号、误差满足  $\|\mathbf{e}\|_2 \leq \eta$ ,  $\Psi: \mathbb{R}^{n \times n} \rightarrow \mathbb{R}^m$ . 该问题的分量形式为  $y_i =$

$\langle \Psi_i, \mathbf{X} \rangle + e_i, i = 1, \dots, m$ , 每个  $\Psi_i$  是彼此独立的随机矩阵。以下分别论述其每个矩阵元遵循彼此独立的高斯分布  $N(0, 1)$  和亚高斯分布(sub-Gaussian)两种情形。

##### 4.1 宽度估计

根据 1.3 节的基础理论, 相关宽度指标  $w(D(\|\mathbf{X}\|_1, \mathbf{X}))$  的上界估计是分析随机观测情况下信号重构问题的关键步骤之一。引理 4 是针对矩阵信号和范数  $\|\cdot\|_1$  的基本结果, 明确表达出该类信号的列稀疏度参数  $s$  和列平坦度参数  $r$  如何影响宽度, 即此类信号的复杂度。

**引理 4** 给定  $n \times n$  矩阵  $\mathbf{X} = [\mathbf{x}_1, \dots, \mathbf{x}_n]$ , 其每列  $\mathbf{x}_j$  至多  $s$  个非零分量,  $r$  是集合  $\{j: |\mathbf{x}_j|_1 = \max_k |\mathbf{x}_k|_1\}$  的基数, 即具有最大  $l_1$  范数值的列的数目, 于是:

$$w^2(D(\|\cdot\|_1, \mathbf{X})) \leq 1 + n^2 - r(n - s \lg(Cn^4 r^2)) \quad (18)$$

特别地, 对  $s$  列稀疏且  $l_1$  列平坦信号( $r = n$ )有:

$$w^2(D(\|\cdot\|_1, \mathbf{X})) \leq 1 + n s \lg(Cn^6) \quad (19)$$

其中,  $C$  是绝对常数(即与问题的任何参量无关)。

**注 2** 以下称参量  $r$  为矩阵信号的  $l_1$  列平坦度、 $s$  为列稀疏度。从式(18)和式(19)看到, 宽度  $w$  的数值随  $s$  的减少或者随  $r$  的增大而下降, 当列稀疏度  $s$  固定时, 宽度对平坦信号具有最小数值。

**证明** 证明从 1.3 节定理 3 中的不等式  $w^2(D(\|\cdot\|_1, \mathbf{X})) \leq E_C(\min\{\|\mathbf{G} - t\mathbf{V}\|_F^2: t > 0, \mathbf{V} \text{ 属于 } \partial\|\mathbf{X}\|_1\})$  出发, 其中  $\mathbf{G}$  是矩阵元遵循高斯分布  $G_{ij} \sim \text{iid } N(0, 1)$  的随机矩阵。

令  $\mathbf{G} = [\mathbf{g}_1, \dots, \mathbf{g}_n]$ , 其中每列是随机向量  $\mathbf{g}_j \sim \text{iid } N(0, \mathbf{I}_n)$ 。由引理 1,  $\partial\|\mathbf{X}\|_1$  中的  $\mathbf{V} = [\lambda_1 \xi_1, \dots, \lambda_n \xi_n]$ , 其中(对列经必要的重编号后)对  $j = 1, \dots, r$  有  $\lambda_j \geq 0$  和  $|\mathbf{x}_j|_1 = \max_k |\mathbf{x}_k|_1$ , 对  $j \geq r + 1$  有  $\lambda_j = 0$  和  $|\mathbf{x}_j|_1 < \max_k |\mathbf{x}_k|_1$ ,  $\lambda_1 + \dots + \lambda_r = 1$ , 对  $X_{ij} \neq 0$  有  $\xi_j(i) = \text{sgn } X_{ij}$ , 对任何  $i$  和  $j$  有  $|\xi_j(i)| \leq 1$ 。逐步展开计算得:

$$\begin{aligned}
&w^2(D(\|\cdot\|_1, \mathbf{X})) \\
&\leq E_C(\min_{t > 0, \lambda_j, \xi_j} \sum_{j=1}^r |\mathbf{g}_j - t\lambda_j \xi_j|_2^2 + \sum_{j=r+1}^n |\mathbf{g}_j|_2^2) \\
&\leq \min_{t > 0, \lambda_j} E_C(\sum_{j=r+1}^n |\mathbf{g}_j|_2^2 + \min_{\xi_j} \sum_{j=1}^r \\
&\quad |\mathbf{g}_j - t\lambda_j \xi_j|_2^2) \\
&= \min_{t > 0, \lambda_j} E_C(\sum_{j=1}^n |\mathbf{g}_j - t\lambda_j \xi_j|_2^2) + \\
&\quad \min_{\xi_j} \sum_{j=r+1}^n E_C(|\mathbf{g}_j|_2^2) \\
&= (n - r)n + \min_{t > 0, \lambda_j} E_C(\sum_{r=j+1}^n \min_{\xi_j} |\mathbf{g}_j - t\lambda_j \xi_j|_2^2) \\
&\quad (\text{各变量 } \xi_j \text{ 彼此无关且 } E_C[|\mathbf{g}_j|_2^2] = n)
\end{aligned}$$

$$= (n-r)n + \min_{t>0, \lambda} \sum_{j=1}^r E_{g_j}(\min_{g_j} |g_j - t\lambda_j \xi_j|_2^2)$$

对每个  $j=1, \dots, r$ , 令  $S(j) = \text{supp}(\mathbf{x}_j)$  (因而  $|S(j)| \leq s$ ),  $\sim S(j)$  表示其补集, 显然有  $|g_j - t\lambda_j \xi_j|_2^2 = |g_{j|S(j)} - t\lambda_j \xi_{j|S(j)}|_2^2 + |g_{j|\sim S(j)} - t\lambda_j \xi_{j|\sim S(j)}|_2^2$ . 注意  $\xi_{j|S(j)}$  的所有分量取值为 1 而  $\xi_{j|\sim S(j)}$  的所有分量取值于区间  $[-1, +1]$ . 选择  $\lambda_1 = \dots = \lambda_r = 1/r$ , 令  $\varepsilon > 0$  是充分小的正数,  $t = t(\varepsilon)$  是满足  $P(|g| > t/r) \leq \varepsilon$  的正数, 其中标量  $g \sim N(0, 1)$ , 由高斯积分的 Chernoff 界知  $t$  可这样选取使  $\varepsilon = \exp(-t^2/2r^2)$ . 对每个  $j$  和不在第  $j$  列的支集  $S(j)$  中的每个行标号  $i$ , 若  $|g_j(i)| \leq t(\varepsilon)/r$  则取  $\xi_j$  的相应分量  $\xi_j(i) = rg_j(i)/t(\varepsilon)$  (这时  $|g_j(i) - t\lambda_j \xi_j(i)| = 0$ ), 否则取  $\xi_j(i) = \text{sgn}(g_j(i))$  (这时  $|g_j(i) - t\lambda_j \xi_j(i)| = |g_j(i)| - t(\varepsilon)/r$ ), 于是当  $|g_{j|\sim S(j)}|_\infty < t(\varepsilon)/r$  时有  $|g_{j|\sim S(j)} - t\lambda_j \xi_{j|\sim S(j)}|_2^2 = 0$ , 进而有:

$$\begin{aligned} & E(|g_{j|\sim S(j)} - t\lambda_j \xi_{j|\sim S(j)}|_2^2) \\ &= \int_0^\infty P(|g_{j|\sim S(j)} - t\lambda_j \xi_{j|\sim S(j)}|_2^2 > u) du \\ &= 2 \int_0^\infty uP(|g_{j|\sim S(j)} - t\lambda_j \xi_{j|\sim S(j)}|_2 > u) du \\ &\leq 2 \int_0^\infty uP(|g_{j|\sim S(j)} - t\lambda_j \xi_{j|\sim S(j)}| \text{ 至少存在一个分量的绝对值} > (n-s)^{-1/2}u) du \\ &\leq 2(n-s) \int_0^\infty uP[|g| - t(\varepsilon)/r > (n-s)^{-1/2}u] du \\ &\leq 2(n-s) \int_0^\infty u \exp(-((t(\varepsilon)/r) + (n-s)^{-1/2}u)^2/2) du \\ &\leq C_0(n-s)^2 \exp(-t(\varepsilon)^2/2r^2) \leq C_0(n-s)^2 \varepsilon \end{aligned}$$

这里  $C_0$  为某个绝对常数. 另一方面:

$$\begin{aligned} & E_{g_j}(|g_{j|S(j)} - t\lambda_j \xi_{j|S(j)}|_2^2) = (t(\varepsilon)^2/r^2) |\xi_{j|S(j)}|_2^2 + \\ & E_{g_j}(|g_{j|S(j)}|_2^2) \\ &= (1 + t(\varepsilon)^2/r^2)s = (1 + 2\lg(1/\varepsilon))s \end{aligned}$$

因此

$$\begin{aligned} & w^2(D(\|\cdot\|_1, \mathbf{X})) \leq (n-r)n + r(n-s)^2 + \\ & \varepsilon(1 + 2\lg(1/\varepsilon))rs \\ & \leq n^2 + C_0 n^2 r \varepsilon - r(n-s)\lg(e/\varepsilon^2) \end{aligned}$$

上式对任意  $\varepsilon > 0$  均成立, 取  $\varepsilon = 1/C_0 n^2 r$  就得到  $w^2(D(\|\cdot\|_1, \mathbf{X})) \leq n^2 - r(n-s)\lg(Cn^4 r^2) + 1$ . 证毕.

## 4.2 主要结果

### 情形 1 高斯型随机算子 $\Psi$

将引理 4 同 1.3 节的定理 2 相结合, 通过简单而直接的计算就得到以下结论:

**定理 8** 设  $\Psi_k$  的随机矩阵元  $\Psi_{kij} \sim \text{iid } N(0, 1)$ ,

矩阵信号  $\mathbf{X} \in \sum_s^{n \times n}$  且  $l_1$  列平坦,  $R^m \ni \mathbf{y} = \Psi(\mathbf{X}) + \mathbf{e}$ , 误差界  $|\mathbf{e}|_2 \leq \eta$ ,  $\mathbf{X}^*$  是问题  $\text{MP}_{\mathbf{y}, \Psi, \eta}^{(2)}$  的最优解. 对任何  $t > 0$ , 如果

$$m \geq t + 2\eta/\delta + (n \text{slg}(Cn^6))^{1/2} \quad (20)$$

其中  $C$  是绝对常数, 那么  $P(|\mathbf{X}^* - \mathbf{X}|_F \leq \delta) \geq 1 - \exp(-t^2/2)$ , 即  $s$  列稀疏且  $l_1$  列平坦的矩阵信号  $\mathbf{X}$  能由问题  $\text{MP}_{\mathbf{y}, \Psi, \eta}^{(2)}$  的最优解  $\mathbf{X}^*$  在逼近误差  $|\mathbf{X}^* - \mathbf{X}|_F$  以高概率充分小的意义上被重构.

### 情形 2 亚高斯型算子 $\Psi$

在这种情形, 将引理 4、定理 1 结论 2) 和定理 4 相结合, 通过简单的计算就得到以下结论:

**定理 9** 设  $\mathbf{X}$  和  $\mathbf{X}^*$  分别是实际的矩阵信号和问题  $\text{MP}_{\mathbf{y}, \Psi, \eta}^{(2)}$  的最优解,  $\mathbf{X} \in \sum_s^{n \times n}$  且  $l_1$  列平坦, 观测向量  $\mathbf{y}$  的分量  $y_k = \langle \Psi_k, \mathbf{X} \rangle + e_k$ , 每个观测矩阵  $\Psi_k \sim \text{iid } \Psi$ ,  $k=1, \dots, m$ ,  $\Psi$  的矩阵元是亚高斯随机变量且满足定理 4 中的条件 1)、2)、3),  $\alpha, \rho, \sigma$  是与该组条件有关的常数.

对任何  $t > 0$ , 若

$$m \geq (C_1 \rho^4 / \alpha) (\alpha t + 2\eta/\delta + \sigma C_2 (\rho^6 n \text{slg}(C_3 n^6))^{1/2}) \quad (21)$$

其中,  $C_i$  均为绝对常数, 那么  $P[|\mathbf{X}^* - \mathbf{X}|_F \leq \delta] \geq 1 - \exp(-C_4 t^2)$ , 即  $s$  列稀疏且  $l_1$  列平坦的信号  $\mathbf{X}$  能通过问题  $\text{MP}_{\mathbf{y}, \Psi, \eta}^{(2)}$  的最优解  $\mathbf{X}^*$  在逼近误差  $|\mathbf{X}^* - \mathbf{X}|_F$  以高概率充分小的意义上被重构.

**注 3** 注意由柯西不等式  $|z|_1 \leq n^{1/2} |z|_2$  有  $|Z|_F^2 = \sum_{i,j} |z_{ij}|^2 \geq (1/n) \sum_i |z_i|_1^2 \geq \max_j |z_j|_1 = \|\mathbf{Z}\|_1$ , 因此式(20)和式(21)显然也分别使  $P(\|\mathbf{X}^* - \mathbf{X}\|_1 \leq \delta) \geq 1 - \exp(-t^2/2)$  和  $P(\|\mathbf{X}^* - \mathbf{X}\|_1 \leq \delta) \geq 1 - \exp(-C_4 t^2)$ .

**注 4** 在略去次要项之后, 上述结果均具有渐近形式  $m \geq C_1 (n \text{slg } n)^{1/2}$ . 该界优于将向量形式的普遍理论直接推广为矩阵信号的  $\min$ - $l_1$  理论所能得到的结果, 后者的渐近形式是  $m \geq C_2 n \text{slg } n$ . 这一方面缘于范数  $\|\cdot\|_1$  较之  $l_1$  范数更能够准确表达矩阵信号的稀疏—平坦特性, 另一方面是缘于这里专门针对矩阵信号、而非简单将其视为向量的分析方法.

## 5 数值仿真

存在很多方式的数值仿真实验对以上得到的观测维数  $m$ 、列稀疏度  $s$ 、矩阵信号的行/列数  $n$  的理论关系进行验证. 在此采用的基本方式是首先固定一个随机生成的稀疏—平坦信号, 通过统计在不同

观测维数上求解  $\min-\|\cdot\|_1$  凸优化问题得以成功重构原始信号的经验型概率(百分比)来进行。

用以测试的信号选取 3232 方阵信号 ( $n=32$ ), 每列按均匀分布随机选取 2 个非零元素, 每个元素的数值按均值为 1 的指数型分布生成, 全部元素彼此概率独立, 最后将每个元素的绝对值按其在列的  $l_1$  范数归一化, 由此得到具有列稀疏度  $s=2$ 、 $\|\cdot\|_1$  范数等于 1 的稀疏—平坦矩阵信号  $\mathbf{X}$ 。为限制计算量, 其他实验信号也按 3232 方阵形式如上类似构造, 列稀疏度分别取  $s=3$  和 4。

观测算子  $\Psi$  采用分量形式  $\Psi(\mathbf{Z})_k = \text{tr}(\Psi_k^T \mathbf{Z})$ ,  $k=1, \dots, m$ ,  $m$  个矩阵  $\Psi_k$  是  $n$  行  $n$  列随机方阵, 全部  $\Psi_k$  的元素按  $N(0, 1)$  正态分布独立生成, 观测维数  $m$  是在数值实验中的可变参数。

对亚高斯型观测算子的数值实验,  $\Psi$  的构造方式同上, 差别在于全部  $\Psi_k$  的元素按亚高斯分布独立生成, 具体如 Radmather 型(数值为  $\pm 1$  的均匀分布)和指数型等分布。

存在很多凸优化算法求解  $\min-\|\cdot\|_1$  问题, 例如较为高效并适合于非光滑目标函数的空间分割型算法求解<sup>[21-22]</sup>。在没有观测误差的情形, 最简捷的方法是(通过补充一组线性约束式)将该问题转化为等价的线性规划问题求解; 在有观测误差且以 Frobenius 范数度量误差的情形, 则可以转化为等价的二次型优化问题<sup>[16]</sup>求解。仿真实验应用该方法进行。重构信号和原始信号间的误差以 Frobenius 范数度量,  $\delta$  是基本控制参数, 当误差小于  $\delta$  时作为正确重构事件来统计。取  $\delta=0.05\|\mathbf{X}\|_1$ , 实验在每个观测维数  $m$  上完整执行 400 次, 数值结果如图 1 和图 2 所示。

图 1 是对以上几种稀疏—平坦矩阵信号通过求解  $\min-\|\cdot\|_1$  凸优化问题被成功重构的概率随观测维数  $m$  变化的结果, 实验分别对高斯型和指数型(典型的亚高斯型分布之一)随机观测算子来进行。作为比较, 图 2 是该实验在相同的参数上对相同的矩阵信号完整执行相同次数求解  $\min-l_1$  问题的数值结果。

从图 1 和图 2 可看到, 两者的概率曲线随观测维数  $m$  的变化形状相似, 这同定理 8 和定理 9 所得到的  $\min-\|\cdot\|_1$  理论的成功概率的下界与  $m$  的关系和  $\min-l_1$  理论结果的形式相同(除常数不同之外)是一致的。同时从图 1 和图 2 也看到, 在每种对应的情况下,  $\min-\|\cdot\|_1$  理论的成功概率曲线都比  $\min-l_1$  理论的曲线在明显更低的维数  $m$  处开始

上升。须指出的是, 本文的证明方法不足以确定定理 8 和定理 9 的结论中那些常数的大小, 因此上述实验给出的是趋势性验证, 更完整的验证有赖于更细致的理论结果。

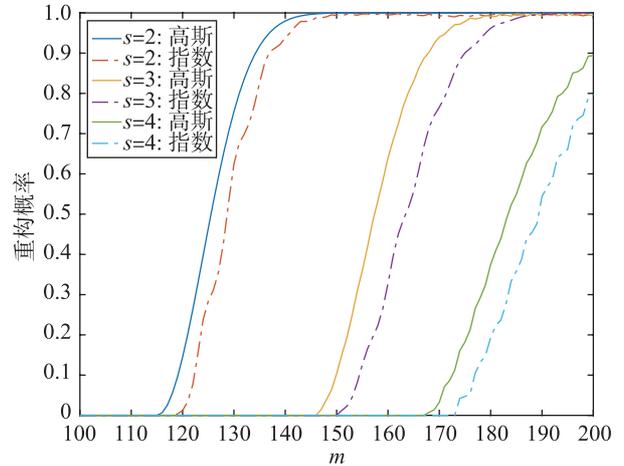


图 1  $\min-\|\cdot\|_1$  成功重构信号的概率与观测维数  $m$  的关系  
Fig.1 Relationship between measurement dimensions  $m$  and success probability of reconstruction via  $\min-\|\cdot\|_1$

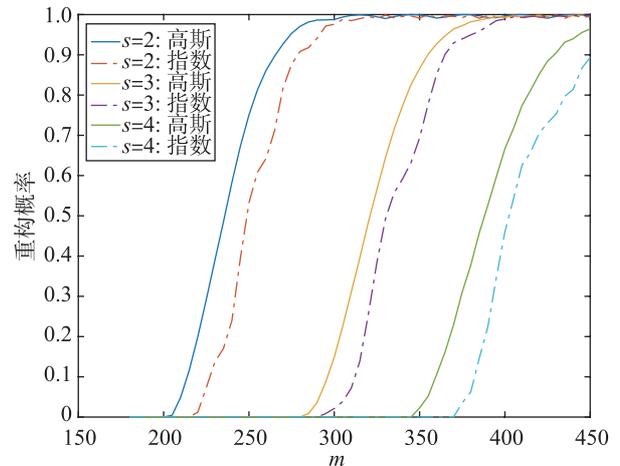


图 2  $\min-l_1$  成功重构信号的概率与观测维数  $m$  的关系  
Fig.2 Relationship between measurement dimensions  $m$  and success probability of reconstruction via  $\min-l_1$

以上实验是针对稀疏—平坦矩阵信号。在实际应用中更多类型的矩阵信号仅具有近似的稀疏—平坦特性。采用引理 4 注记中的记号, 用  $r$  表示矩阵信号的  $l_1$  列平坦度, 于是  $r=n$  意味着信号平坦,  $n-r(>0)$  数值越大则约不平坦。图 3 是通过求解  $\min-\|\cdot\|_1$  来重构非平坦信号的数值仿真结果, 原始信号  $\mathbf{X}$  仍然采用前述形式, 同时随机选取  $n-r$  个列, 每列对其中某元素减去一个幅度不超过  $0.1\|\mathbf{X}\|_1$  的随机正实数以破坏信号的平坦特性。以列稀疏度  $s=2$  且高斯随机观测算子的情形为例, 对  $n-r=1, 2,$

3、4、5 的数值仿真结果如图 3，概率曲线族具有宽度估计式所预期的近似等间距性质(本文的证明方法不足以确定这些理论估计式中常数的大小，因此所得并非完整的验证)。对其他列稀疏度和亚高斯随机观测算子(如指数型)也具有相似的结果。

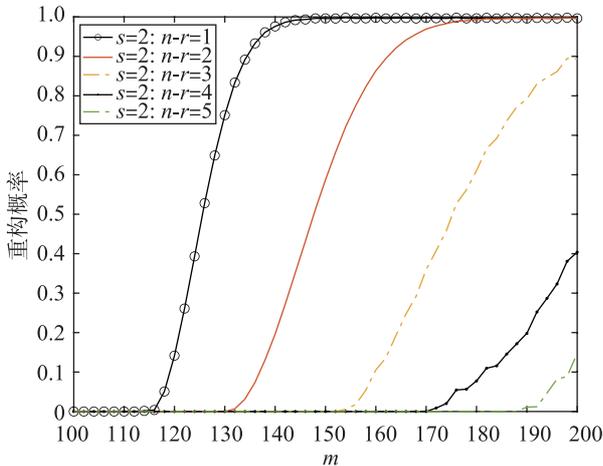


图3  $\min\|\cdot\|_1$  重构非平坦信号的概率与观测维数  $m$  的关系

Fig.3 Relationship between measurement dimensions  $m$  and success probability of reconstructing non-flat signals via  $\min\|\cdot\|_1$

## 6 结论

本文研究基于  $\min\|\cdot\|_1$  凸优化方法重构稀疏—平坦矩阵信号的一组充分条件。所得结果优于将针对向量信号的  $\min-l_1$  方法直接移植为针对矩阵

信号的  $\min-l_1$  处理的结果，主要表现在第 3 节的条件形式较之后者更为具体，同时第 4 节的界条件也优于后者的界条件。这一方面是因为所采用的矩阵范数  $\|\cdot\|_1$  较之  $L_1$  范数更能准确表达稀疏—平坦矩阵信号的结构特征，另一方面是因为这里采用专门针对矩阵信号、而非简单将其视为向量信号的分析方法。

本文的分析方法缩小了结论对矩阵信号的适用范围，仅限于一类特殊结构的矩阵信号。本文认为对不同结构特征的矩阵信号，采用特殊类型的矩阵范数(例如对重构低秩矩阵采用核范数、对重构稀疏—平坦矩阵信号采用  $\|\cdot\|_1$  范数等)和分析途径来建立相应的结果、而非总是采用  $l_1$  范数，将具有更好的针对性和实用性。这也意味着矩阵信号和高阶张量信号重构问题的研究较之向量信号会更为复杂和丰富。

在进一步待研究的问题中，一类是结合应用领域中观测算子所具有的结构特征，来进一步建立更为具体和实用的重构条件，这将特别有益于优化第 3 节中的那些条件，同时也利于定出第 4 节结果中某些重要常数的范围；另一类工作是结合矩阵数值算法的丰富成果，发展相应的求解算法并应用于稀疏—平坦矩阵信号所出现的典型问题。最后，由于高阶张量在多源成像等应用领域中越来越多地出现，类似问题针对高阶张量信号的情形也值得深入研究。

## 参考文献

- [1] Foucart F, Rauhut H. A mathematical introduction to compressive sensing[M]. 2nd ed. Berlin, Germany: Birkhaeuser, 2019.
- [2] Eldar Y C. Sampling theory: Beyond bandlimited systems[M]. Cambridge, UK: Cambridge University Press, 2015.
- [3] 丛爽, 张娇娇. 压缩传感理论、优化算法及其在系统状态重构中的应用[J]. 信息与控制, 2017, 46(3): 267-274.  
Cong S, Zhang J J. Compressive sensing theory, optimization algorithm and application in system state reconstruction[J]. Information and Control, 2017, 46(3): 267-274.
- [4] Cohen D, Eldar Y C. Sub-Nyquist radar systems: Temporal, spectral and spatial compression[J]. IEEE Signal Processing Magazine, 2018, 11(2): 35-57.
- [5] Kannan R, Shake N, Berman A. Properties of strong-tensors and general-tensors[J]. Linear Algebra and Its Applications, 2017, 455(1): 133-166.
- [6] Duarte M F, Baraniuk R G. Kronecker compressive sensing[J]. IEEE Transactions on Image Processing, 2012, 21(2): 494-504.
- [7] Dai W, Li Y, Zou J. et al. Fully decomposable compressive sampling with optimization for multidimensional sparse representation[J]. IEEE Transactions on Signal Processing, 2018, 66(3): 603-616.
- [8] Heckel R, Morgenshtern V, Soltanolkotabi M. Super-resolution radar[J]. Information and Inference, 2016, 5(1): 22-75.
- [9] Dorsch D, Rauhut H. Refined analysis of sparse MIMO radar[J]. Journal of Fourier Analysis and Applications, 2017, 23(6): 1120-1152.

- 
- [10] Aubel C, Stotz D, Boelcskei H. Super-resolution from short-time Fourier transform measurements[J]. *Journal of Fourier Analysis and Applications*, 2018, 24(1): 22–35.
- [11] Khwaja A S, Onhon N O, Cetin M. Compressed sensing methods for radar imaging in the presence of phase errors and moving objects[M]. Cambridge, UK: Cambridge University Press, 2020.
- [12] Davenport M A, Romberg J. An overview of low-rank matrix recovery from incomplete observations[J]. *IEEE Journal of Selected Topics in Signal Processing*, 2016, 10(4): 592–625.
- [13] Tropp J A, Yurtsever A, Udell S. Practical sketching algorithms for low rank matrix approximation[J]. *SIAM Journal of Matrix Analysis and Applications*, 2017, 38(4): 1454–1485.
- [14] Aldroubi P, Cabrelli R, Jaffard S. New trends in applied harmonic analysis: Sparse representation, compressed sensing and multifractal analysis[M]. Archen, Germany: SciTech Publishing, 2019.
- [15] Cohen R, Eldar Y C. Sparse Doppler sensing based on nested arrays[J]. *IEEE Transaction on Ultrasonics, Ferroelectrics and Frequency Control*, 2018, 65(12): 2349–2364.
- [16] Agrawal A, Barrat S, Boyd S. Learning convex optimization models[J]. *IEEE Journal of Automatica Sinica*, 2021, 8(8): 1355–1364.
- [17] Chandrasekaran V, Recht B, Parrio P A, et al. The convex geometry of linear inverse problems[J]. *Foundations of Computing Mathematics*, 2012, 12: 805–849.
- [18] Tropp J A. Convex recovery of a structured signal from independent random linear measurements[M]//*Sampling theory: A Renaissance in Compressive Sampling and Other Developments*. Basel, Germany: Birkhaeuser, 2016.
- [19] Mendelson S. Learning without concentration for general loss functions[J]. *Probability Theory and Related Fields*, 2018, 71(2): 459–502.
- [20] Vatié S, Peyre G, Dadili J. Model consistency of partly smooth regularizers[J]. *IEEE Transactions on Information Theory*, 2018, 64(3): 1725–1747.
- [21] Narasimhan F, Boyd S. CVXR: A package for disciplined convex optimization[J]. *Journal of Statistical Software*, 2020, 94(1): 1–34.

## 作者简介

田 园(1966–), 男, 博士, 教授。研究领域为信号处理, 信息论和非线性控制。

黄 鑫(1999–), 男, 硕士生。研究领域为信号处理和大数据分析。

李明楚(1962–), 男, 博士, 教授。研究领域为信息安全及智能大数据分析。