

基于冗余紧框架的 l_2/l_1 极小化块稀疏压缩感知*

张枫[†] 王建军[‡]

(西南大学数学与统计学院, 重庆 400715)

摘要: 压缩感知是(近似)稀疏信号处理的研究热点之一, 它突破了Nyquist/Shannon采样率, 实现了信号的高效采集和鲁棒重构. 本文采用 l_2/l_1 极小化方法和Block D-RIP理论研究了在冗余紧框架下的块稀疏信号, 所获结果表明, 当Block D-RIP常数 $\delta_{2k|\tau}$ 满足 $0 < \delta_{2k|\tau} < 0.2$ 时, l_2/l_1 极小化方法能够鲁棒重构原始信号, 同时改进了已有的重构条件和误差上界. 基于离散傅里叶变换(DFT)字典, 我们执行了一系列仿真实验充分地证实了理论结果.

关键词: 压缩感知; l_2/l_1 极小化方法; Block D-RIP; 冗余紧框架; 块稀疏信号

中图分类号: O29 **文献标识码:** A **文章编号:** 0372-2112(20xx)xx-xxxx-xx

DOI: 10.3969/j.issn.0372-2112.20xx.xx.xxx

1 引言

压缩感知^[1-3] (Compressed sensing)作为一种新的采样理论, 它利用信号的稀疏特性, 在远小于Nyquist 采样率的条件下, 用随机采样获取信号的离散样本, 通过非线性重建算法完美的恢复信号. 压缩感知理论一经提出, 就引起了学者广泛关注. 目前已在压缩成像^[4], 医学成像^[5], 模式识别^[6], 图像处理^[7]等领域得到了广泛应用.

在压缩感知中, 我们主要考虑以下模型:

$$y = Ax + w,$$

其中, $A \in \mathbb{R}^{m \times n}$ 是测量矩阵, $y \in \mathbb{R}^m$ 是已知的线性测量, $x \in \mathbb{R}^n$ 是待重构的未知信号, $w \in \mathbb{R}^n$ 是噪声($\|w\|_2 \leq \epsilon$). 压缩感知的核心思想是依赖于信号是否是稀疏的或者近似稀疏的, 即信号 x 的非零元素的个数是否远小于 x 的长度. 然而, 在现实中常见的自然信号不一定都具有稀疏特性, 甚至这类信号在某些正交基上都不能够进行稀疏表示. 自然地, 上

*基金项目:国家自然科学基金资助项目(NO. 61673015, 61273020),中央高校基本业务费专项(NO. XD-JK2015A007), 科学计算与智能信息处理广西高校重点实验室(合同编号: GXSCIP201702).

[†]作者简介: 张枫(1991-), 男, 博士生, 研究方向: 压缩感知、矩阵修补和鲁棒主成分分析研究.

[‡]通信作者:王建军, 教授.

述信号重构过程不能直接应用于自然信号的重构. 研究表明, 一些自然信号在某些紧框架 $D \in \mathbb{R}^{n \times N}$ ($n \leq N, DD^* = I_n$)^[8, 9] 上能够稀疏表示, 即 $x = D\alpha$, 其中 $\alpha \in \mathbb{R}^N$ 是(近似)稀疏的. 从数学的角度讲, 上述问题可以写成以下形式

$$\min_{x \in \mathbb{R}^n} \|D^*x\|_0 \quad s.t. \quad \|Ax - y\|_2 \leq \varepsilon, \quad (1)$$

其中 D^* 表示矩阵 D 的共轭转置, $\|D^*x\|_0$ 表示向量 D^*x 中非零元素的个数, 若 $\|D^*x\|_0 \leq k$, 那么就称向量 D^*x 为 k -稀疏信号. 由于问题(1)是一个NP-hard问题, 即在多项式时间内, 计算机无法有效求解. 所以一种更为实际的和易于处理的凸优化方法被提出

$$\min_{x \in \mathbb{R}^n} \|D^*x\|_1 \quad s.t. \quad \|Ax - y\|_2 \leq \varepsilon, \quad (2)$$

其中 $\|D^*x\|_1 = \sum_{i=1}^N |(D^*x)_i|$. 为了保证信号的重构效果, Candès等人在文献^[10]中提出了以下D-限制性等容条件(D-Restricted Isometry Property, D-RIP):

定义1.1 (D-RIP) 对于任意 k -稀疏信号 v ($\|v\|_0 \leq k$), 若存在 $0 < \delta_k < 1$, 使得

$$(1 - \delta_k) \|Dv\|_2^2 \leq \|ADv\|_2^2 \leq (1 + \delta_k) \|Dv\|_2^2,$$

那么称矩阵 A 满足 k 阶D-限制性等容性质(D-RIP), 最小的 δ_k 称为D-限制性等容常数(D-RIC).

进一步, 他们指出当测量矩阵 A 满足D-限制性等容性质且 $\delta_{2k} < 0.08$ 时通过求解有约束优化问题(2)可以实现信号的鲁棒重构. 之后, 2011年, Li等人^[11]将上述条件作了进一步改善, 得到 $\delta_{2k} < 0.4931$. 我们知道, 问题(2)等价于以下转换为以下无约束优化问题

$$\min_{x \in \mathbb{R}^n} \lambda \|D^*x\|_1 + \frac{1}{2} \|Ax - y\|_2^2. \quad (3)$$

2014年, Zhao Tan等人在文献^[8]中提出, 若 A 满足D-限制性等容性质(D-RIP), D 是一个紧框架, 当D-RIP常数 $\delta_{2k} < 0.1907$ 并且 $\|D^*A^*w\|_\infty \leq \frac{\lambda}{2}$ 时, 问题(3)的解 \hat{x} 满足

$$\|\hat{x} - x\|_2 \leq V_1 \sqrt{k} \lambda + V_2 \frac{\|D^*x - (D^*x)_k\|_1}{\sqrt{k}},$$

其中 $\|D^*x - (D^*x)_k\|_1$ 是最佳 k -项 ℓ_1 逼近误差. C_1, C_2 是两个数值型常数并且 C_1 由D-RIP常数和 $\|D^*D\|_{1,1}$ ($\|D^*D\|_{1,1} := \sup\{\|D^*Dv\|_1 : v \in \mathbb{R}^N, \|v\|_1 \leq 1\}$) 共同量化, C_2 仅取决于D-RIP常数.

然而现实世界中的自然信号其结构千变万化, 一种常见的结构方式是自然信号在某些冗余字典下是块稀疏的, 即其非零元素以块的形式出现, 例如彩色图像处理^[7]和DNA-微阵列^[12]等. 从数学的角度看, 给定分块 $\tau = \{\tau_1, \tau_2, \tau_3, \dots, \tau_d\}$, 对于任意向量 $\alpha \in \mathbb{R}^N$ 都可以被描述为

$$\alpha = \underbrace{[\alpha_1, \dots, \alpha_{\tau_1}]}_{\alpha[1]}, \underbrace{[\alpha_{\tau_1+1}, \dots, \alpha_{\tau_1+\tau_2}]}_{\alpha[2]}, \dots, \underbrace{[\alpha_{N-\tau_d+1}, \dots, \alpha_N]}_{\alpha[d]},$$

其中 $\alpha[i]$ 表示向量 α 的第 i 个子块, 而 α_i 则表示向量 α 的第 i 个分量元素. 如果向量 α 最多有 k 个非零块, 即 $\|\alpha\|_{2,0} \leq k$, 则称向量 α 为块 k -稀疏信号. 特别地, 当 $d = 1$ 时, 等价于传统的带字典的压缩感知问题. 相应地, 测量矩阵 $A \in \mathbb{R}^{m \times n}$ 和冗余字典 $D \in \mathbb{R}^{n \times N}$ 也可以分别被描述为

$$A = \underbrace{[A_1, \dots, A_{\tau_1}]}_{A[1]}, \underbrace{[A_{\tau_1+1}, \dots, A_{\tau_1+\tau_2}]}_{A[2]}, \dots, \underbrace{[A_{N-\tau_d+1}, \dots, A_n]}_{A[d]},$$

$$D = \underbrace{[D_1, \dots, D_{\tau_1}]}_{D[1]}, \underbrace{[D_{\tau_1+1}, \dots, D_{\tau_1+\tau_2}]}_{D[2]}, \dots, \underbrace{[D_{N-\tau_d+1}, \dots, D_N]}_{D[d]}.$$

其中 $A[i]$, $D[i]$ 分别表示矩阵 A 和 D 的第 i 个子块(矩阵), 而 A_i , D_i 则分别表示矩阵 A 和 D 的第 i 个列向量.

然而使用原始的 ℓ_1 极小化方法来恢复此类块稀疏信号不能充分利用信号的结构性特征, 即非零元素是以块的形式出现的这一特性. 为此, 一些学者对传统的压缩感知方法进行了针对性的改进. 在文献^[13]中, Wang等人提出并研究了如下的 ℓ_2/ℓ_1 问题

$$\min_{x \in \mathbb{R}^n} \|D^* x\|_{2,1} \quad s.t. \quad \|Ax - y\|_2 \leq \varepsilon, \quad (4)$$

其中 $\|D^* x\|_{2,1} = \sum_{i=1}^N \|(D[i])^* x[i]\|_2$. 此外借鉴D-RIP概念, 他们引入了Block D-RIP, 定义如下:

定义1.2 (Block D-RIP) 对于任意块 k -稀疏信号 v ($\|v\|_{2,0} \leq k$), 若存在 $0 < \delta_{k|\tau} < 1$, 使得

$$(1 - \delta_{k|\tau}) \|Dv\|_2^2 \leq \|ADv\|_2^2 \leq (1 + \delta_{k|\tau}) \|Dv\|_2^2,$$

那么称矩阵 A 满足 k 阶Block D-限制性等容性质(Block D-RIP), 最小的 $\delta_{k|\tau}$ 称为Block D-限制性等容常数(Block D-RIC).

在文献^[14]中, Eldar和Mishalizi指出当测量矩阵 A 满足Block D-限制性等容性质(Block D-RIP)且 $\delta_{2k|\tau} < 0.4142$ 时, 块稀疏信号能够通过求解有约束优化问题(4)进行鲁棒重构. 其后, 2013年, Lin等人^[15]将上述条件作了进一步改善, 得到了 $\delta_{2k|\tau} < 0.4931$. 类似地, 问题(4)可以转化为以下块无约束优化问题

$$\min_{x \in \mathbb{R}^n} \lambda \|D^* x\|_{2,1} + \frac{1}{2} \|Ax - y\|_2^2. \quad (5)$$

假设 A 满足Block D-限制性等容性质(Block D-RIP), D 是一个紧框架, 本文研究表明, 当Block D-RIP常数 $\delta_{2k|\tau} < 0.2$ 并且 $\|D^* A^* w\|_{2,\infty} \leq \frac{\lambda}{2}$ 时, 问题(5)的解 \hat{x} 满足

$$\|\hat{x} - x\|_2 \leq C_1 \sqrt{k} \lambda + C_2 \frac{\|D^* x - (D^* x)_k\|_{2,1}}{\sqrt{k}},$$

其中 C_1, C_2 是两个数值型常数并且 C_1 由D-RIP常数和 $\|D^* D\|_{(2,1),(2,1)}$ ($\|D^* D\|_{(2,1),(2,1)} := \sup\{\|D^* Dv\|_{2,1} : v \in \mathbb{R}^N, \|v\|_{2,1} \leq 1\}$)共同量化, C_2 仅取决于D-RIP常数.

2 预备知识

为了方便介绍后文,我们首先给出以下记号.

- 给定正整数 d , 记索引集 $T \subset \{1, 2, \dots, d\}$ 且 $|T| = k$, T^c 表示 T 在 $\{1, 2, \dots, d\}$ 中的补集.
- $D_T \in R^{n \times |T|}$ 表示从 D 中取出索引集 T 对应的列所组成的矩阵, 记 $D_T^* = (D_T)^*$.
- 记 T_1 为包含 $D_{T^c}^* h$ 的 k 个最大2范数块的索引集, T_2 为包含 $D_{(T \cup T_1)^c}^* h$ 的 k 个最大2范数块的索引集, 等.

接下来, 为了证明主要定理, 我们需要以下引理.

引理2.1 $\|D_{T^c}^* h\|_2 \leq \|D_{T_1}^* h\|_2 + \frac{\|D_{T^c}^* h\|_{2,1}}{2\sqrt{k}}$.

Proof. 由 T_j 的构造, 我们有

$$\|D_{T_{j+1}}^* [i] h [i]\|_2 \leq \frac{\|D_{T_j}^* h\|_{2,1}}{k},$$

上式两边取2范数, 得

$$\|D_{T_{j+1}}^* h\|_2 \leq \frac{\|D_{T_j}^* h\|_{2,1}}{\sqrt{k}},$$

所以, 我们有

$$\sum_{j \geq 2}^l \|D_{T_j}^* h\|_2 \leq \frac{\sum_{j \geq 1}^l \|D_{T_j}^* h\|_{2,1}}{\sqrt{k}} = \frac{\|D_{T^c}^* h\|_{2,1}}{\sqrt{k}}. \quad (6)$$

我们注意到 $\{\|D_{T_j}^* h\|_2\}_{j=1}^l$ 是一个单调不增序列. 下面的不等式可以通过两边平方得证

$$\|D_{T^c}^* h\|_2 \leq \|D_{T_1}^* h\|_2 + (\sqrt{2} - 1) \sum_{j=2}^l \|D_{T_j}^* h\|_2. \quad (7)$$

又因 $\sqrt{2} - 1 < \frac{1}{2}$, 结合(6)式和(7)式, 引理得证. \square

引理2.2 令测量 $y = Ax + w$, $h = \hat{x} - x$, 若 $\|D^* A^* w\|_{2,\infty} \leq \frac{\lambda}{2}$, 则问题(5)的解 \hat{x} 满足

$$\|D^* A^* A h\|_{2,\infty} \leq \left(\frac{1}{2} + \|D^* D\|_{(2,1),(2,1)}\right) \lambda.$$

Proof. 由实值凸的低阶半连续函数的次微分的定义和 \hat{x} 是问题(5)的解可知, \hat{x} 满足问题(5)的子梯度最优化条件, 即

$$A^*(A\hat{x} - y) + \lambda Dv = 0,$$

其中 $v \in \mathbb{R}^N$, 若 $\|D_i^* x\|_2 \neq 0$, 则 $v_i = \text{sgn}(\|D_i^* x\|_2)$, 否则 $\|v_i\|_2 \leq 1$. 因此存在 $v \in \mathbb{R}^N$ 使得 $\|v\|_{2,\infty} \leq 1$, 进一步, 我们有

$$\|D^* A^*(A\hat{x} - y)\|_{2,\infty} = \lambda \|D^* D v\|_{2,\infty} \leq \lambda \|D^* D\|_{(2,\infty),(2,\infty)} = \lambda \|D^* D\|_{(2,1),(2,1)}.$$

于是

$$\|D^* A^* A h\|_{2,\infty} \leq \|D^* A^*(A\hat{x} - y)\|_{2,\infty} + \|D^* A^*(A x - y)\|_{2,\infty} \leq \left(\frac{1}{2} + \|D^* D\|_{(2,1),(2,1)}\right) \lambda.$$

引理得证. \square

引理2.3 令 $D \in \mathbb{R}^{n \times N}$ 为满足 $DD^* = I_n$ 的矩阵, $A \in \mathbb{R}^{m \times n}$ 为满足Block D-RIP($0 < \delta_{2k|\tau} < \frac{1}{2}$)条件的矩阵. 令索引集 $T \subseteq \{1, 2, \dots, d\}$ 恰有 k 个块, 测量 $y = Ax + w$, $h = \hat{x} - x$, 若 $\|D^* A^* w\|_{2,\infty} \leq \frac{\lambda}{2}$, 则问题(5)的解 \hat{x} 满足

$$\|D_T^* h\|_2 \leq \beta_1 \sqrt{k} \lambda + \beta_2 \frac{\|D_{T^c}^* h\|_{2,1}}{\sqrt{k}}, \quad (8)$$

其中

$$\beta_1 = \frac{1 + 2\|D^* D\|_{(2,1),(2,1)}}{2 - 4\delta_{2k|\tau}}, \quad \beta_2 = \frac{\delta_{2k|\tau}}{1 - 2\delta_{2k|\tau}}.$$

Proof. 令 T_0 为包含 $D^* h$ 的 k 个最大2范数块的索引集, 因为 $\|D_T^* h\|_2 \leq \|D_{T_0}^* h\|_2, \|D_{T_0^c}^* h\|_{2,1} \leq \|D_{T^c}^* h\|_{2,1}$, 所以我们取 $T = T_0$ 就能充分证明该引理.

因为 A 满足Block D-RIP($0 < \delta_{2k|\tau} < \frac{1}{2}$)条件, 不失一般性, 我们假设存在 $u, v \in \sum_k := \{Dc : c \in \mathbb{R}^N, \|c\|_{2,0} \leq k\}$, 使得 $\|u\|_2 = \|v\|_2 = 1$, 因此我们有

$$\begin{aligned} \langle Au, Av \rangle &= \frac{1}{4} (\|A(u+v)\|_2^2 - \|A(u-v)\|_2^2) \\ &\geq \frac{1}{4} (1 - \delta_{2k|\tau}) \|u+v\|_2^2 - \frac{1}{4} (1 + \delta_{2k|\tau}) \|u-v\|_2^2 \\ &= \langle u, v \rangle - \delta_{2k|\tau}. \end{aligned}$$

对任意的 $u, v \in \sum_k$, 我们有

$$\langle Au, Av \rangle \geq \langle u, v \rangle - \delta_{2k|\tau} \|u\|_2 \|v\|_2.$$

于是

$$\begin{aligned} \langle Ah, ADD_{T_0}^* h \rangle &= \langle ADD_{T_0}^* h, ADD_{T_0}^* h \rangle + \sum_{j \geq 1}^l \langle ADD_{T_j}^* h, ADD_{T_0}^* h \rangle \\ &\geq (1 - \delta_{2k|\tau}) \|DD_{T_0}^* h\|_2^2 + \sum_{j \geq 1}^l \langle DD_{T_j}^* h, DD_{T_0}^* h \rangle - \delta_{2k|\tau} \|DD_{T_0}^* h\|_2 \sum_{j \geq 1}^l \|DD_{T_j}^* h\|_2 \\ &= (1 - \delta_{2k|\tau}) \|DD_{T_0}^* h\|_2^2 + \langle h - DD_{T_0}^* h, DD_{T_0}^* h \rangle - \delta_{2k|\tau} \|DD_{T_0}^* h\|_2 \sum_{j \geq 1}^l \|DD_{T_j}^* h\|_2 \\ &= (1 - \delta_{2k|\tau}) \|DD_{T_0}^* h\|_2^2 + \|D_{T_0}^* h\|_2^2 - \|DD_{T_0}^* h\|_2^2 - \delta_{2k|\tau} \|DD_{T_0}^* h\|_2 \sum_{j \geq 1}^l \|DD_{T_j}^* h\|_2, \end{aligned}$$

我们注意到 D 是一个紧框架, 即 $DD^* = I_n$, 故

$$\|DD_{T_0}^* h\|_2 \leq \|D\|_2 \|D_{T_0}^* h\|_2 \leq \|D_{T_0}^* h\|_2,$$

因此

$$\langle Ah, ADD_{T_0}^* h \rangle \geq (1 - \delta_{2k|\tau}) \|D_{T_0}^* h\|_2^2 - \delta_{2k|\tau} \|D_{T_0}^* h\|_2 \sum_{j \geq 1}^l \|D_{T_j}^* h\|_2. \quad (9)$$

另一方面, 由引理2.2, $\|D_{T_0}^* h\|_{2,1} \leq \sqrt{k} \|D_{T_0}^* h\|_2$, 我们有

$$\begin{aligned} \langle Ah, ADD_{T_0}^* h \rangle &= \langle D^* A^* Ah, D_{T_0}^* h \rangle \\ &\leq \|D^* A^* Ah\|_{2,\infty} \|D_{T_0}^* h\|_{2,1} \\ &\leq \left(\frac{1}{2} + \|D^* D\|_{(2,1),(2,1)}\right) \lambda \sqrt{k} \|D_{T_0}^* h\|_2. \end{aligned} \quad (10)$$

结合(6), (9), (10)式, 我们有

$$\begin{aligned} (1 - \delta_{2k|\tau}) \|D_{T_0}^* h\|_2^2 &\leq \left(\frac{1}{2} + \|D^* D\|_{(2,1),(2,1)}\right) \sqrt{k} \lambda \|D_{T_0}^* h\|_2 + \delta_{2k|\tau} \|D_{T_0}^* h\|_2 \sum_{j \geq 1}^l \|D_{T_j}^* h\|_2 \\ &= \|D_{T_0}^* h\|_2 \left(\left(\frac{1}{2} + \|D^* D\|_{(2,1),(2,1)}\right) \sqrt{k} \lambda + \delta_{2k|\tau} \|D_{T_1}^* h\|_2 + \delta_{2k|\tau} \sum_{j \geq 2}^l \|D_{T_j}^* h\|_2 \right) \\ &\leq \|D_{T_0}^* h\|_2 \left(\left(\frac{1}{2} + \|D^* D\|_{(2,1),(2,1)}\right) \sqrt{k} \lambda + \delta_{2k|\tau} \|D_{T_0}^* h\|_2 + \delta_{2k|\tau} \frac{\|D_{T_0}^* h\|_{2,1}}{\sqrt{k}} \right). \end{aligned}$$

当 $\|D_{T_0}^* h\|_2 = 0$ 且 $T = T_0$, 引理2.3显然成立, 当 $\|D_{T_0}^* h\|_2 \neq 0$ 时

$$\|D_T^* h\|_2 \leq \frac{1 + 2\|D^* D\|_{(2,1),(2,1)} \sqrt{k} \lambda}{2 - 4\delta_{2k|\tau}} + \frac{\delta_{2k|\tau}}{1 - 2\delta_{2k|\tau}} \frac{\|D_{T^c}^* x\|_{2,1}}{\sqrt{k}}.$$

引理得证. □

引理2.4 令 $D \in \mathbb{R}^{n \times N}$ 为满足 $DD^* = I_n$ 的矩阵, 测量 $y = Ax + w$, $h = \hat{x} - x$, 若 $\|D^* A^* w\|_{2,\infty} \leq \frac{\lambda}{2}$, 则问题(5)的解 \hat{x} 满足

$$\|D_{T^c}^* h\|_{2,1} \leq 3\|D_T^* h\|_{2,1} + 4\|D_{T^c}^* x\|_{2,1}. \quad (11)$$

Proof. 由于 \hat{x} 是问题(5)的解, 故

$$\lambda \|D^* \hat{x}\|_{2,1} + \frac{1}{2} \|A\hat{x} - y\|_2^2 \leq \lambda \|D^* x\|_{2,1} + \frac{1}{2} \|Ax - y\|_2^2.$$

将 $h = \hat{x} - x$, $y = Ax + w$ 代入上式, 由 D 是一个紧框架可知

$$\begin{aligned} \lambda \|D^* \hat{x}\|_{2,1} + \frac{1}{2} \|Ah\|_2^2 &\leq \lambda \|D^* x\|_{2,1} + \langle Ah, w \rangle \\ &\leq \lambda \|D^* x\|_{2,1} + \langle D^* h, D^* A^* w \rangle \\ &\leq \lambda \|D^* x\|_{2,1} + \|D^* h\|_{2,1} \|D^* A^* w\|_{2,\infty} \\ &\leq \lambda \|D^* x\|_{2,1} + \frac{\lambda}{2} \|D^* h\|_{2,1}, \end{aligned}$$

于是

$$\|D^*x\|_{2,1} \leq \|D^*x\|_{2,1} + \frac{1}{2}\|D^*h\|_{2,1}.$$

进一步, 我们有

$$\|D_T^*h + D_T^*x\|_{2,1} + \|D_{T^c}^*h + D_{T^c}^*x\|_{2,1} \leq \|D_T^*x\|_{2,1} + \|D_{T^c}^*x\|_{2,1} + \frac{1}{2}\|D_T^*h\|_{2,1} + \frac{1}{2}\|D_{T^c}^*h\|_{2,1},$$

由三角不等式, 易知

$$\|D_T^*x\|_{2,1} - \|D_T^*h\|_{2,1} + \|D_{T^c}^*h\|_{2,1} - \|D_{T^c}^*x\|_{2,1} \leq \|D_T^*x\|_{2,1} + \|D_{T^c}^*x\|_{2,1} + \frac{1}{2}\|D_T^*h\|_{2,1} + \frac{1}{2}\|D_{T^c}^*h\|_{2,1},$$

整理后得(11)式, 引理得证. \square

3 主要结果

定理3.1 令 $D \in \mathbb{R}^{n \times N}$ 为满足 $DD^* = I_n$ 的矩阵, $A \in \mathbb{R}^{m \times n}$ 为满足Block D-RIP($0 < \delta_{2k|\tau} < 0.2$)条件的矩阵, $(D^*x)_k$ 为由 D^*x 的 k 个最大2范数块组成的向量, 测量 $y = Ax + w$, $h = \hat{x} - x$, 若 $\|D^*A^*w\|_{2,\infty} \leq \frac{\lambda}{2}$, 则问题(5)的解 \hat{x} 满足

$$\|\hat{x} - x\|_2 \leq C_1\sqrt{k}\lambda + C_2 \frac{\|D^*x - (D^*x)_k\|_{2,1}}{\sqrt{k}}, \quad (12)$$

其中

$$C_1 = \frac{(1 + 2\|D^*D\|_{(2,1),(2,1)})(\frac{7}{4} - 3\delta_{2k|\tau})}{(1 - 2\delta_{2k|\tau})(1 - 5\delta_{2k|\tau})}, \quad C_2 = \frac{2 - 4\delta_{2k|\tau}^2}{(1 - 2\delta_{2k|\tau})(1 - 5\delta_{2k|\tau})}$$

Proof. 由引理2.1和引理2.4, 我们有

$$\begin{aligned} \|\hat{x} - x\|_2 = \|h\|_2 &= \|D^*h\|_2 \\ &\leq \|D_T^*h\|_2 + \|D_{T^c}^*h\|_2 \\ &\leq \|D_T^*h\|_2 + \|D_{T_1}^*h\|_2 + \frac{\|D_{T^c}^*h\|_{2,1}}{2\sqrt{k}} \\ &\leq \|D_T^*h\|_2 + \|D_{T_1}^*h\|_2 + \frac{3\|D_T^*h\|_{2,1} + 4\|D_{T^c}^*x\|_{2,1}}{2\sqrt{k}} \\ &\leq \frac{5}{2}\|D_T^*h\|_2 + \|D_{T_1}^*h\|_2 + \frac{2\|D_{T^c}^*x\|_{2,1}}{\sqrt{k}}. \end{aligned} \quad (13)$$

下面我们分别估计 $\|D_T^*h\|_2$ 和 $\|D_{T_1}^*h\|_2$, 由(8)式知

$$\begin{aligned} \|D_T^*h\|_2 &\leq \beta_1\sqrt{k}\lambda + \beta_2 \frac{\|D_{T^c}^*h\|_{2,1}}{\sqrt{k}} \\ &\leq \beta_1\sqrt{k}\lambda + \beta_2 \frac{3\|D_T^*h\|_{2,1} + 4\|D_{T^c}^*x\|_{2,1}}{\sqrt{k}} \\ &\leq \beta_1\sqrt{k}\lambda + 3\beta_2\|D_T^*h\|_2 + 4\beta_2 \frac{\|D_{T^c}^*x\|_{2,1}}{\sqrt{k}}, \end{aligned}$$

由于 $1 - 3\beta_2 > 0$ ($0 < \delta_{2k|T} < 0.2$), 故

$$\|D_T^* h\|_2 \leq \frac{\beta_1}{1 - 3\beta_2} \sqrt{k} \lambda + \frac{4\beta_2}{1 - 3\beta_2} \frac{\|D_{T^c}^* x\|_{2,1}}{\sqrt{k}}. \quad (14)$$

由(8)式和(14)式, 我们有

$$\begin{aligned} \|D_{T_1}^* h\|_2 &\leq \beta_1 \sqrt{k} \lambda + \beta_2 \frac{\|D_{T_1}^* h\|_{2,1}}{\sqrt{k}} \\ &\leq \beta_1 \sqrt{k} \lambda + \beta_2 \frac{\|D_T^* h\|_{2,1} + \|D_{T^c}^* h\|_{2,1}}{\sqrt{k}} \\ &\leq \beta_1 \sqrt{k} \lambda + \frac{\beta_2}{\sqrt{k}} \|D_T^* h\|_{2,1} + \frac{3\beta_2}{\sqrt{k}} \|D_T^* h\|_{2,1} + 4\beta_2 \frac{\|D_{T^c}^* x\|_{2,1}}{\sqrt{k}} \\ &\leq \beta_1 \sqrt{k} \lambda + 4\beta_2 \|D_T^* h\|_2 + 4\beta_2 \frac{\|D_{T^c}^* x\|_{2,1}}{\sqrt{k}} \\ &\leq \beta_1 \sqrt{k} \lambda + 4\beta_2 \left(\frac{\beta_1}{1 - 3\beta_2} \sqrt{k} \lambda + \frac{4\beta_2}{1 - 3\beta_2} \frac{\|D_{T^c}^* x\|_{2,1}}{\sqrt{k}} \right) + 4\beta_2 \frac{\|D_{T^c}^* x\|_{2,1}}{\sqrt{k}} \\ &= \frac{\beta_1 + \beta_1 \beta_2}{1 - 3\beta_2} \sqrt{k} \lambda + \frac{4\beta_2 + 4\beta_2^2}{1 - 3\beta_2} \frac{\|D_{T^c}^* x\|_{2,1}}{\sqrt{k}}. \end{aligned} \quad (15)$$

由(13), (14), (15)式知, (12)式成立. \square

4 数值实验

为了验证理论结果, 本文分别做了两组实验: (1)设计满足定理3.1条件的算法; (2)理论误差上界对比实验. 实验在CPU为Inter (R)Core (TM)i3, 内存为2GB的台式电脑上进行, 运行软件为MATLAB (R2014a). 实验中, 测量矩阵 $A \in \mathbb{R}^{128 \times 256}$ 服从标准高斯分布, 字典 $D \in \mathbb{R}^{256 \times 1024}$ 由傅里叶变换矩阵和单位阵合并而成, 并且满足 $DD^* = I_n$, 测量误差 $w \in \mathbb{R}^{128}$ 服从正态分布, 取正则化参数 $\lambda = 1e - 3$, 从而满足 $\|D^* A^* w\|_{2,\infty} \leq \frac{\lambda}{2}$, 待重构的信号 $x \in \mathbb{R}^{256}$, 在字典 D 下的块稀疏信号 $\alpha \in \mathbb{R}^{1024}$ 中的非零块位置随机产生. 为了克服实验结果的偶然性, 所有实验将独立重复地进行100次.

在本文中, 我们将冗余字典和Block-IRLS算法^[16-18]相结合提出D-Block-IRLS算法, 参见算法1.

首先针对D-Block-IRLS算法, 我们采用两种分块形式, 均分256块和非均分256块. 图2分别研究了均匀分块和非均匀分块的情况下通过D-Block-IRLS算法得到的误差 $\|\hat{x} - x\|_2$ 以及理论误差上界 $C_1 \sqrt{k} \lambda + C_2 \frac{\|D^* x - (D^* x)_k\|_{2,1}}{\sqrt{k}}$ 与块稀疏度 k 的关系. 由图可知, 无论是均匀分块还是非均匀分块通过D-Block-IRLS算法得到的误差都远小于理论误差上界, 换言之, 利用D-Block-IRLS算法来重构块稀疏信号可以满足实验设计的需要并且从侧面印证了我们理论分析的正确性.

由文献^[8]知, 当 $D \in \mathbb{R}^{n \times N}$ 为满足 $DD^* = I_n$ 的矩阵, $A \in \mathbb{R}^{m \times n}$ 为满足D-RIP($0 < \delta_{2k} < 0.1907$) 条件的矩阵, $(D^* x)_k$ 为由 $D^* x$ 的 k 个最大元素组成的向量, 测量 $y = Ax + w$, $h = \hat{x} - x$,

算法 1 D-Block-IRLS算法

输入: 分块 $\tau = \{\tau_1, \tau_2, \dots, \tau_d\}$, 测量矩阵 A , 字典 D , 观测信号 Y , 块稀疏度估计 k .

输出: 重构信号 x .

- 1: 选择适当的惩罚参数 $\lambda(0 < \lambda < 1)$.
- 2: 初始化迭代向量 $\alpha^{(0)}$, 使其满足 $AD\alpha^{(0)} = y$. 设置 $\epsilon_0 = 1$.
- 3: 开始迭代 $j = 0$.
- 4: 通过 $\alpha^{(j)}$ 解决下面的线性问题

$$\left((AD)^T AD + \text{Diag} \left[\frac{\lambda I_{\tau_l}}{(\epsilon_j^2 + \|\alpha^{(j)}[i]\|_2^2)^{\frac{1}{2}}} \right]_{1 \leq l \leq d} \right) \alpha^{(j+1)} = (AD)^T y.$$

- 5: 当 $\alpha^{(j)}$ 满足停机条件, 将 $D\alpha^{(j)}$ 作为输出赋值给 x , 同时结束算法, 否则执行下一步.
 - 6: 取定一个常数 $c(0 < c < \frac{1}{d})$ 并更新 $\epsilon_{j+1} = \min \{ \epsilon_j, c \cdot r(\alpha^{(j+1)})_{k+1}/d \}$. 其中, $r(\alpha)$ 表示将向量 α 的分块取 ℓ_2 范数后, 再由大到小依次排列形成的向量. 而 $r(\alpha)_{k+1}$ 表示向量 $r(\alpha)$ 的第 $k+1$ 个分量值.
 - 7: $j = j + 1$, 并返回到第4步继续执行.
-

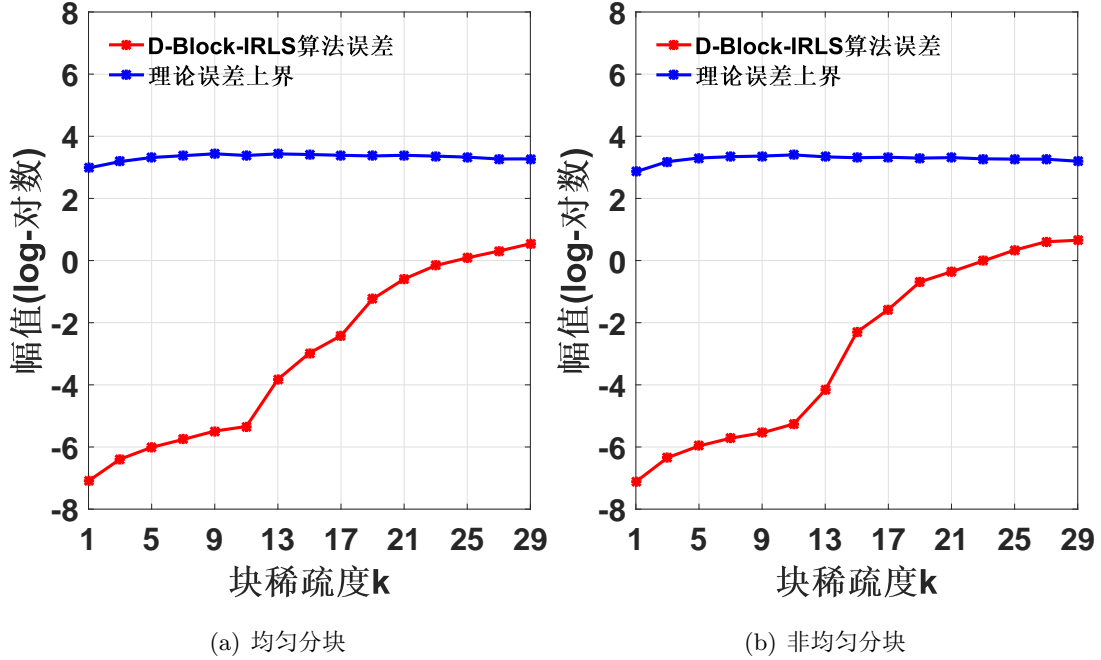


图 1: D-Block-IRLS算法误差与理论误差上界对比.

若 $\|D^*A^*w\|_\infty \leq \frac{\lambda}{2}$, 则问题(3)的解 \hat{x} 满足

$$\|\hat{x} - x\|_2 \leq V_1\sqrt{k}\lambda + V_2\frac{\|D^*x - (D^*x)_k\|_1}{\sqrt{k}}, \quad (16)$$

其中

$$V_1 = \frac{4\sqrt{2}(\frac{1}{2} + \|D^*D\|_{1,1})}{1 - (1 + 3\sqrt{2})\delta_{2k}}, \quad V_2 = \frac{4((\sqrt{2} - 1)\delta_{2k} + 1)}{1 - (1 + 3\sqrt{2})\delta_{2k}}$$

现在分别用块和非块的方式来处理块稀疏信号 α , 取 $\|D^*D\|_{(2,1),(2,1)} = \frac{1+\sqrt{256}}{2}$, $\|D^*D\|_{1,1} = \frac{1+\sqrt{1024}}{2}$ (参见文献[9]), 其他参数保持一致. 图3表明, 无论是均匀分块还是非均匀分块, 我们所获理论误差上界都明显优于(16)式的误差上限.

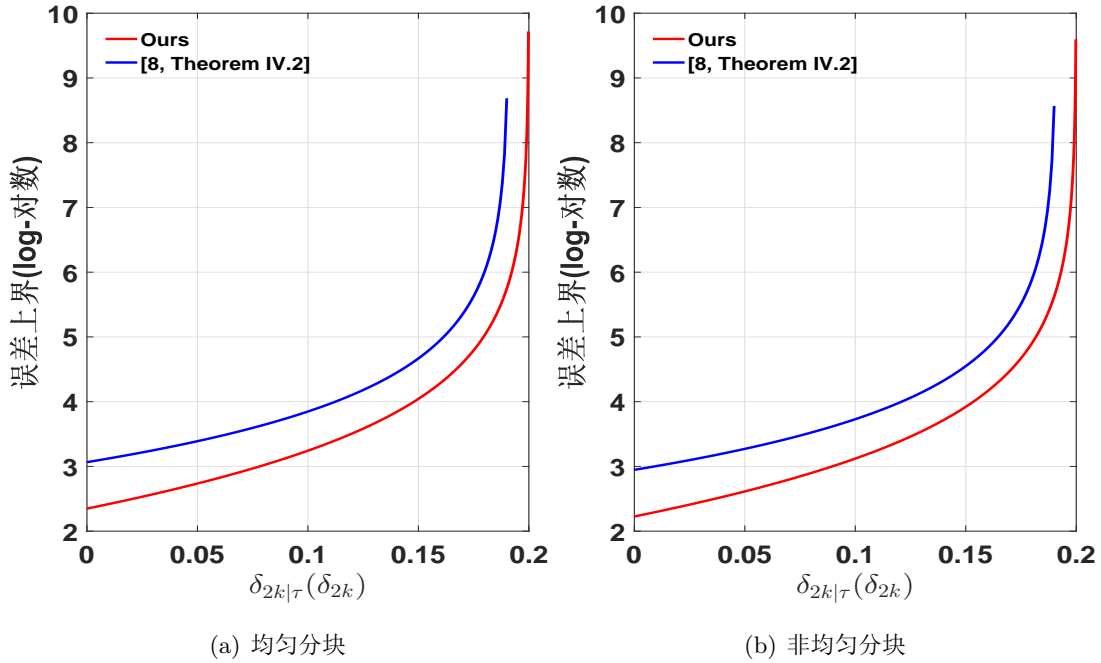


图 2: 理论误差上界对比.

5 小结

本文采用 l_2/l_1 极小化方法研究了基于冗余紧框架下的块稀疏信号的恢复, 获得了该方法鲁棒重构原始信号的充分条件和误差上界估计, 所获结果表明, 误差上界可以通过正则化参数 λ, k -项逼近和块稀疏度来控制. 仿真实验证明了我们理论结果的准确性, 该结果对于推动压缩感知的进一步发展具有一定的理论价值和借鉴作用.

参考文献:

- [1] Nyquist H. Certain topics in telegraph transmission theory[J]. American Institute of Electrical Engineers, 1928, 47 (2): 617-644.
- [2] Candès E, Romberg J, Tao T. Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information[J]. IEEE Transactions on Information Theory, 2006, 52 (2): 489-509.
- [3] Donoho D L. Compressed sensing[J]. IEEE Transactions on Information Theory, 2006, 52 (4): 1289-1306.
- [4] Baraniuk R G. Single-pixel imaging via compressive sampling[J]. IEEE Signal Processing Magazine, 2008, 25 (2): 83-91.
- [5] Lustig M, Donoho D, Pauly J M. Sparse MRI: The application of compressed sensing for rapid MR imaging[J]. Magnetic resonance in medicine, 2007, 58 (6): 1182-1195.
- [6] Wright J, Yang A Y, Ganesh A, et al. Robust face recognition via sparse representation[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2009, 31 (2): 210-227.
- [7] Majumdar A, Ward R. Compressed sensing of color images[J]. Signal Processing, 2010, 90 (12): 3122-3127.
- [8] Z Tan, Y C Eldar, A. Beck, A. Nehorai. Smoothing and decomposition for analysis sparse recovery. IEEE Trans. Signal Process, 2014 62 (7): 1762 - 1774.
- [9] Y Shen, B Han, E Braverman. Stable recovery of analysis based approaches. Appl. Comput. Harmon. Anal, 2015 39 (1): 162-163.
- [10] E J Candès, Y C Eldar, D. Needell, P. Randall. Compressed sensing with coherent and redundant dictionaries. Appl. Comput. Harmon. Anal, 2011 31 (1): 59 - 73.
- [11] Mo Q, Li S. New bounds on the restricted isometry constant δ_{2k} . Appl. Comput. Harmon. Anal, 2011 31 (3): 460 - 468.
- [12] Parvaresh F, Vikalo H, Misra H, et al. Recovering sparse signals using sparse measurement matrices in compressed DNA microarrays. IEEE Journal of Selected Topics in Signal Processing, 2008, 2 (3): 275-285.
- [13] Wang Y, Wang J J, Xu Z B. A Note on Block-sparse Signals Recovery with Coherent Tight Frames. Discrete Dynamics in Nature and Society, 2013 2013 (1): 1-8.
- [14] Eldar Y C, Mishali M. Robust recovery of signals from a structured union of subspaces[J]. IEEE Transactions on Information Theory, 2009, 55 (11): 5302-5316.
- [15] Lin J H, Li S. Block Sparse Recovery via Mixed ℓ_2/ℓ_1 -Minimization[J]. Acta Mathematica Sinica, 2013, 29 (7): 1401-1412.
- [16] 王文东, 王尧, 王建军. 基于迭代重赋权最小二乘算法的块稀疏压缩感知. 电子学报, 2015, 45 (5): 923-928 (Wang W D, Wang Y, Wang J J. Iterative Reweighted Least Squares Algorithm for Block-Sparse Compressed Sensing. Chinese Journal of Electronics, 2015, 45 (5): 923-928.

- [17] Wang J, Zhang J, Wang W, et al. A perturbation analysis of nonconvex block-sparse compressed sensing[J]. Communications in Nonlinear Science and Numerical Simulation, 2015, 29(1 - 3):416-426.
- [18] Liu C, Wang J, Wang W, et al. Non-convex block-sparse compressed sensing with redundant dictionaries[J]. Iet Signal Processing, 2017, 11 (2): 171-180.

Block-sparse Compressed Sensing with Redundant Tight Frames via ℓ_2/ℓ_1 -Minimization

Zhang Feng, Wang Jianjun

(School of Mathematics and Statistic, Southwest University, Chongqing 400715, China)

Abstract: Compressed sensing is one of the hot research theories for (approximately) sparse signal processing which breaks through Nyquist/Shannon sampling theory, and makes it into reality that one can efficiently acquire and exactly reconstruct a signal. This paper mainly investigated the signals which are block-sparse under redundant tight frames based on ℓ_2/ℓ_1 -minimization method and Block D-RIP theory. Under the condition $0 < \delta_{2k|\tau} < 0.2$, the obtained results show that ℓ_2/ℓ_1 -minimization method can robustly reconstruct the original signal, meanwhile, improve the existing reconstruction condition and error upper bound. Using the discrete Fourier transform dictionary, we conducted a series of simulation experiments which sufficiently verified the theoretical results.

Keywords: Compressed sensing; ℓ_2/ℓ_1 -Minimization method; Block D-RIP; Redundant tight frames; Block-sparse signals

2010 MSC: 94A12, 94A15