

基于非凸复合函数的稀疏信号恢复算法

周洁容^{1,2} 李海洋^{1,2} 凌军^{1,2} 陈浩^{1,2} 彭济根^{1,2}

摘要 基于泛函深度作用的思想, 通过将两种非凸稀疏泛函进行复合, 构造了一种新的稀疏信号重构模型, 实现了对 0 范数的深度逼近. 综合运用 MM (Majorize minimization) 技术、外点罚函数法和共轭梯度法, 提出一种求解该模型的算法, 称为 NCCS (Non-convex composite sparse) 算法. 为降低重构信号陷入局部极值的可能性, 提出在算法的每步迭代中以 BP (Basis pursuit) 模型的解作为初始迭代值. 为验证所建模型和所提算法的有效性, 进行了多项数值实验. 实验结果表明, 相较于 SL0 (Smoothed L_0) 算法、IRLS (Iterative reweighed least squares) 算法、SCSA (Successive concave sparsity approximation) 算法以及 BP 算法等经典算法, 提出的算法在重构误差、信噪比、归一化均方差、支撑集恢复成功率等方面都有更优的表现.

关键词 压缩感知, 稀疏信号重构, MM 技术, 外点罚函数法, 共轭梯度法

引用格式 周洁容, 李海洋, 凌军, 陈浩, 彭济根. 基于非凸复合函数的稀疏信号恢复算法. 自动化学报, 2022, 48(7): 1782–1793

DOI 10.16383/j.aas.c200666

Sparse Signal Reconstruction Algorithm Based on Non-convex Composite Function

ZHOU Jie-Rong^{1,2} LI Hai-Yang^{1,2} LING Jun^{1,2} CHEN Hao^{1,2} PENG Ji-Gen^{1,2}

Abstract Based on the idea of deep composition of functionals, this paper constructs a new sparse signal reconstruction model by combining two non-convex sparse functionals to achieve the depth of 0 norm approaching. Using the MM (Majorize minimization) technology, the exterior penalty function method and the conjugate gradient method, an algorithm for solving the model is proposed, which is called the NCCS (Non-convex composite sparse) algorithm. In order to reduce the possibility of the reconstructed signal falling into a local extreme value, it is proposed to use the solution of the BP (Basis pursuit) model as the initial iterative value in each iteration of the algorithm. To verify the effectiveness of the model and the proposed algorithm, a number of numerical experiments have been carried out in this paper. The experimental results show that compared with classic algorithms such as the SL0 (Smoothed L_0) algorithm, IRLS (Iterative reweighed least squares) algorithm, SCSA (Successive concave sparsity approximation) algorithm and BP algorithm, the algorithm proposed in this paper has better performance in reconstruction error, signal-to-noise ratio, normalized mean square error, and support set recovery success rate.

Key words Compressed sensing (CS), sparse signal reconstruction, MM (Majorize minimization) technology, exterior penalty function method, conjugate gradient method

Citation Zhou Jie-Rong, Li Hai-Yang, Ling Jun, Chen Hao, Peng Ji-Gen. Sparse signal reconstruction algorithm based on non-convex composite function. *Acta Automatica Sinica*, 2022, 48(7): 1782–1793

根据奈奎斯特 (Nyquist) 采样定律, 想要实现信号的无失真输出, 采样频率必须在信号带宽的两倍以上. 然而, 在大多数情形下, 按此定律采样所获得的信息是冗余的, 这不仅造成采样的浪费, 而且

处理较大带宽的信号时会给硬件系统带来巨大压力. 压缩感知 (Compressed sensing, CS)^[1] 的出现使该问题的解决成为可能. 压缩感知是一种新型的信号采样及重构理论, 利用少量的测量值就可以实现稀疏或可压缩信号的精确重构. 而稀疏信号的重构在众多科学研究和工程应用中十分重要, 如物理学中的量子态层析成像^[2]、天体物理学成像^[3]、磁共振成像^[4]、信号处理^[5]、雷达成像^[6] 等, CS 理论的引入加快了上述应用的研究和发展.

数学上, 稀疏信号重构是指从欠定线性系统 $\mathbf{b} = \mathbf{A}\mathbf{x}$ 中恢复原始信号 \mathbf{x} , 其中 $\mathbf{b} \in \mathbf{R}^n$ 是测量向量, $\mathbf{A} \in \mathbf{R}^{n \times m}$ ($n < m$) 为感知矩阵, $\mathbf{x} \in \mathbf{R}^m$ 是未知的稀疏向量. 最稀疏信号的重构模型是:

收稿日期 2020-08-18 录用日期 2020-12-01

Manuscript received August 18, 2020; accepted December 1, 2020

国家自然科学基金 (11771347, 12031003) 资助

Supported by National Natural Science Foundation of China (11771347, 12031003)

本文责任编辑 王卓

Recommended by Associate Editor WANG Zhuo

1. 广州大学数学与信息科学学院 广州 510006 2. 广州大学机器生命与智能研究中心 广州 510006

1. Mathematics and Information Science College, Guangzhou University, Guangzhou 510006 2. Research Center for Machine Life and Intelligence, Guangzhou University, Guangzhou 510006

$$(P_0) \quad \min_{\mathbf{x}} \|\mathbf{x}\|_0 \quad \text{s. t. } \mathbf{A}\mathbf{x} = \mathbf{b}$$

其中, $\|\mathbf{x}\|_0$ 表示向量 \mathbf{x} 中非零元素的个数^[7].

已经证明, (P_0) 模型的直接求解是 NP (Non-deterministic polynomial) 难的^[8], 其计算量会随着稀疏向量维数的增加而增大, 模型的抗噪能力也很差. 为此人们提出了多种方法对 (P_0) 模型进行求解, 主要方法可归类于启发式方法、凸松弛方法、非凸松弛方法三种类型. 典型的启发式算法有正交匹配追踪^[9]、阈值算法^[10]、子空间追踪算法^[11]、分级正交匹配追踪^[12]等, 这些算法重构理论简单、速度较快, 但往往需要更多观测, 算法的重构精度较低、收敛速度较慢, 并且在有噪声情况下信号的恢复精度较低, 因此其应用范围有限. 典型的凸松弛算法有梯度投影算法^[13]、BP (Basis pursuit) 算法^[14]、BPDN (Basis pursuit denoising) 算法^[15]、IRLS (Iterative reweighted least squares) 算法^[16]、Bregma 迭代法^[17]等, 其中最经典的是 BP 算法, 文献 [18] 证明了在测量矩阵满足有限等距性质^[19]的条件下 BP 模型与 (P_0) 模型等价, 但在多种情形下 BP 模型中的 L_1 范数不能充分反映信号的稀疏性特征, 往往难以获得稀疏解. 为此, 人们提出了用非凸泛函替代 L_0 范数的方法, 称之为非凸松弛方法.

相较于 L_1 范数, 许多非凸泛函能更好地近似 L_0 范数, 从而更好地反映信号的稀疏性特征. 典型的非凸松弛算法有 NSL0 (Newt-on smooth L_0 norm) 算法^[20]、SL0 (Smoothed L_0) 算法^[21]、CT-NRAL0 (Composite trigono-metric function null-space reweighted approximate L_0 norm) 算法^[22]、SCSA (Successive concave sparsity approximation) 算法^[23]等. 易见, 松弛方法的核心是寻找 L_0 范数的逼近函数, 通过极小化该逼近函数寻得最稀疏解. 显然这种近似模型对 (P_0) 模型的逼近性能取决于所选取的近似函数对 L_0 范数的逼近程度. 因此, 如何构造具有更优逼近性能的近似函数已成为稀疏信号重构问题研究中的重要问题.

最近, 文献 [23–24] 分别构造了两种近似函数 $g_\sigma(x) = 1 - e^{-|x|/\sigma}$ 、 $h_\sigma(x) = |x|/(|x| + \sigma)$ 以实现 L_0 范数的逼近, 取得了较好的重构效果. 一个有趣的想法是, 如果将上述每个泛函对信号的作用看成是一次前向 (Forward) 处理过程, 那么我们是否可以通过泛函的复合实现对信号的深度作用, 从而提高信号重构的性能? 基于这种想法, 本文将上述两个泛函进行复合, 构建一种新的非凸松弛模型, 给出该模型的理论分析, 对该模型提出一种新的近似算法, 并通过数值实验验证算法的有效性以及相较于 SL0 算法、IRLS 算法、BP 算法、SCSA 算法等经典算法的优越性.

1 一种用于逼近 L_0 范数的非凸复合指数泛函

设 δ 为 Kronecker 函数, 即

$$\delta(x) = \begin{cases} 1, & x = 0 \\ 0, & x \neq 0 \end{cases} \quad (1)$$

则对任意的 $\mathbf{x} \in \mathbf{R}^m$, $\|\mathbf{x}\|_0$ 可表示为 $\sum_{i=1}^m [1 - \delta(x_i)]$. 由于函数 $\delta(x_i)$ 是不连续的, 人们总希望找到连续函数 $f_\sigma: \mathbf{R} \rightarrow \mathbf{R}$, 通过控制参数 σ ($0 < \sigma < 1$) 来逼近 $1 - \delta(x_i)$, 即

$$\lim_{\sigma \rightarrow 0} f_\sigma(x_i) = 1 - \delta(x_i) = \begin{cases} 0, & x_i = 0 \\ 1, & x_i \neq 0 \end{cases} \quad (2)$$

为方便计, 本文称这样的函数为 Delta 逼近函数 (Delta approximating, DA). 若给定这样的逼近函数 f_σ , 并对任意 $\mathbf{x} \in \mathbf{R}^m$ 定义 $F_\sigma(\mathbf{x}) = \sum_{i=1}^m f_\sigma(x_i)$, 则容易看出

$$\lim_{\sigma \rightarrow 0} F_\sigma(\mathbf{x}) = \|\mathbf{x}\|_0$$

这样, 下面的优化问题可看成是 (P_0) 问题的近似模型

$$\min_{\mathbf{x}} F_\sigma(\mathbf{x}) \quad \text{s. t. } \mathbf{A}\mathbf{x} = \mathbf{b} \quad (3)$$

文献 [21, 23–24] 分别采用以下三种 DA 函数来实现对 L_0 范数的逼近

$$p_\sigma(x) = 1 - e^{-\frac{x^2}{2\sigma^2}} \quad (4)$$

$$g_\sigma(x) = 1 - e^{-\frac{|x|}{\sigma}} \quad (5)$$

$$h_\sigma(x) = \frac{|x|}{|x| + \sigma} \quad (6)$$

并通过对相应近似模型的求解实现了信号的精确重构.

显而易见, 近似式 (3) 对 (P_0) 模型的逼近性能取决于所选取的 DA 函数对 L_0 范数的逼近程度. 因此, 如何构造具有更优逼近性能的 DA 函数已成为稀疏信号重构问题研究中的重要课题. 注意到, 上述 DA 函数 $h_\sigma(x)$ 和 $g_\sigma(x)$ 都能很好地实现对 L_0 范数的逼近. 一个自然的问题是, 是否可以将这两个函数进行复合以实现 L_0 范数的深度逼近? 针对这个问题, 本文对上述两个函数 $h_\sigma(x)$ 和 $g_\sigma(x)$ 的复合进行考察

$$f_\sigma(x) = (g_\sigma \circ h_\sigma)(x) = 1 - e^{-\frac{|x|}{\sigma(|x| + \sigma)}} \quad (7)$$

其中, “ \circ ” 表示函数复合运算, 显然这是一个 DA 函数. 不仅如此, 可以从几何图像和理论分析上说明, 在固定参数 σ 后该 DA 函数相对于 $h_\sigma(x)$ 、

$g_\sigma(x)$ 、 $p_\sigma(x)$ 具有更好的对 L_0 范数的逼近性能.

1) 几何图像分析

图 1 展示了 $p_\sigma(x)$ 、 $h_\sigma(x)$ 、 $g_\sigma(x)$ 以及复合函数 $f_\sigma(x)$ 这 4 种 DA 函数的几何图像, 其中参数 $\sigma = 0.1$. 可以看出, 在 $x = 0$ 附近, $p_\sigma(x)$ 函数曲线最平坦, 而本文所提出的复合函数 $f_\sigma(x)$ 的图像曲线相对其他三种函数最为陡峭, 这表明函数 $f_\sigma(x)$ 具有对 L_0 范数最好的逼近效果.

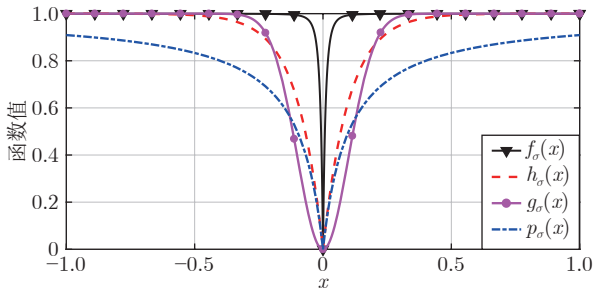


图 1 4 种函数在 $\sigma = 0.1$ 时的一元函数分布
Fig.1 The unary distribution of the four functions at $\sigma = 0.1$

图 2 展示了 $p_\sigma(x)$ 、 $h_\sigma(x)$ 、 $g_\sigma(x)$ 和 $f_\sigma(x)$ 4 种 DA 函数在二维情形的等高几何图像, 即 $p_\sigma(x) = h_\sigma(x) = g_\sigma(x) = f_\sigma(x) = 1$ 的图像, 其中参数 $\sigma = 0.1$. 可以看出, 在 4 种函数与直线 $x_1 = x_2$ 的交点中, 函数 $f_\sigma(x)$ 与 $x_1 = x_2$ 所表示的直线得到的交点的分量绝对值最小, 这表明复合指数函数 $f_\sigma(x)$ 能更好地逼近 L_0 范数.

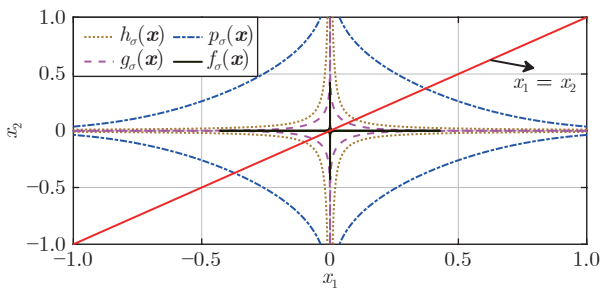


图 2 $p_\sigma(x)$ 、 $h_\sigma(x)$ 、 $g_\sigma(x)$ 和函数 $f_\sigma(x)$ 在 $\sigma = 0.1$ 时的二元函数分布

Fig.2 The bivariate distribution of $p_\sigma(x)$, $h_\sigma(x)$, $g_\sigma(x)$ and the function $f_\sigma(x)$ at $\sigma = 0.1$

2) 理论分析

上述从几何直观上展示了复合指数函数相对于其他 DA 函数的更好逼近性能. 不仅如此, 也可以从理论上证明, 对区间 $(0, 0.9)$ 内任意非零参数 σ , 复合指数函数相对于 $p_\sigma(x)$ 、 $h_\sigma(x)$ 、 $g_\sigma(x)$ 都具有更好的逼近性能. 不失一般性, 仅在第一象限内进行

讨论, 即假定 $x \geq 0$. 对区间内任意 σ ($\sigma \neq 0$), 因为

$$h_\sigma(x) = \frac{x}{x + \sigma} \tag{8}$$

$$g_\sigma(x) = 1 - e^{-\frac{x}{\sigma}} \tag{9}$$

$$p_\sigma(x) = 1 - e^{-\frac{x^2}{2\sigma^2}} \tag{10}$$

$$f_\sigma(x) = 1 - e^{-\frac{x}{\sigma(x + \sigma)}} \tag{11}$$

首先, 固定函数值比较对应变量的大小.

假设 $h_\sigma(x) = g_\sigma(x) = r$, 其中 $0 < r < 1$, $0 < \sigma < 0.9$, 则由式 (8) 以及式 (9) 可得对应变量为

$$\begin{cases} x_h = \frac{\sigma r}{1 - r} \\ x_g = -\sigma \ln(1 - r) \end{cases}$$

因为

$$x_h - x_g = \sigma \left(\frac{r}{1 - r} + \ln(1 - r) \right)$$

令 $R(r) = \frac{r}{1 - r} + \ln(1 - r)$, 有

$$R'(r) = \frac{1}{(1 - r)^2} - \frac{1}{1 - r} > 0$$

所以有 $R(r)$ 为增函数, 即对任意 r 有 $R(r) > R(0) = 0$, 得 $x_h > x_g$. 这说明相较于 $h_\sigma(x)$, 函数 $g_\sigma(x)$ 更贴近坐标轴.

下面固定变量 x , 比较函数值 $p_\sigma(x)$ 、 $g_\sigma(x)$ 、 $f_\sigma(x)$ 的大小. 因为

$$g_\sigma(x) - p_\sigma(x) = e^{-\frac{x^2}{2\sigma^2}} - e^{-\frac{x}{\sigma}}$$

所以, 为讨论 $g_\sigma(x)$ 、 $p_\sigma(x)$ 的大小关系, 只需要比较指数的大小. 易得 x 趋于 0 (至少在 $x \leq 2\sigma$ 时), 对任意非零 σ 恒有

$$-\frac{x^2}{2\sigma^2} \geq -\frac{x}{\sigma}$$

即 $g_\sigma(x) \geq p_\sigma(x)$.

同理, 对

$$f_\sigma(x) - g_\sigma(x) = e^{-\frac{x}{\sigma}} - e^{-\frac{x}{\sigma(x + \sigma)}}$$

进行讨论, 由于

$$\frac{x}{\sigma} - \left(-\frac{x}{\sigma(x + \sigma)} \right) = \frac{x}{\sigma} \left(\frac{1}{x + \sigma} - 1 \right)$$

所以, 在 $x = 0$ 附近 (至少在 $|x| \leq 0.1$ 范围内) 有 $f_\sigma(x) \geq g_\sigma(x)$. 因此有

$$f_\sigma(x) \geq g_\sigma(x) \geq p_\sigma(x)$$

说明 $f_\sigma(x)$ 的函数曲线更陡峭.

利用函数的对称性, 其他象限也可类似讨论.

所以由上述函数的几何图像和理论分析也能得出 $f_\sigma(x)$ 相对于 $h_\sigma(x)$ 、 $g_\sigma(x)$ 、 $p_\sigma(x)$ 具有更强的稀疏性.

2 基于复合指数函数的稀疏模型

以上分析表明, 由式 (7) 定义的复合函数 $f_\sigma(x)$ 具有对 L_0 范数更优的逼近性能. 为此, 考虑由该函数所诱导的优化模型

$$(CE) \quad \min_{\mathbf{x}} F_\sigma(\mathbf{x}) \quad \text{s. t. } A\mathbf{x} = \mathbf{b} \quad (12)$$

其中, $F_\sigma(\mathbf{x}) = \sum_{i=1}^m f_\sigma(x_i)$, 参数 σ 控制着函数 $F_\sigma(\mathbf{x})$ 对 L_0 范数的逼近程度.

下面对优化模型 (CE) 与 (P_0) 之间的关系进行研究, 为此首先证明以下引理.

引理 1. 设模型 (CE) 的最优解为 \mathbf{x}^* , 其稀疏度为 k , 则矩阵 A 中对应解 \mathbf{x}^* 支撑集的子矩阵 A_k 是列满秩的, 其中 \mathbf{x}^* 的支撑集是指集合 $\{i | x_i^* \neq 0, i = 1, \dots, m\}$.

证明. 利用反证法, 假设矩阵 A_k 不是列满秩的, 即矩阵的列向量线性相关, 则存在非零向量 \mathbf{v} , 其支撑集包含在向量 \mathbf{x}^* 的支撑集内, 并且使得 $A_k \mathbf{v} = 0$. 显然, 向量 $\mathbf{x}^* \pm \mathbf{v}$ 满足约束条件 $A\mathbf{x} = \mathbf{b}$. 不失一般性, 假设向量 \mathbf{v} 中最大的分量绝对值不超过向量 \mathbf{x}^* 中最小的分量绝对值.

根据函数 $f_\sigma(x)$ 的严格凹性, 有

$$\left[1 - e^{-\frac{|x+v|}{\sigma(|x+v|+\sigma)}} \right] + \left[1 - e^{-\frac{|x-v|}{\sigma(|x-v|+\sigma)}} \right] < 2 \left[1 - e^{-\frac{|x|}{\sigma(|x|+\sigma)}} \right]$$

于是, 有

$$\sum_{i=1}^m \left[1 - e^{-\frac{|x_i^*+v|}{\sigma(|x_i^*+v|+\sigma)}} \right] + \sum_{i=1}^m \left[1 - e^{-\frac{|x_i^*-v|}{\sigma(|x_i^*-v|+\sigma)}} \right] < 2 \sum_{i=1}^m \left[1 - e^{-\frac{|x_i^*|}{\sigma(|x_i^*|+\sigma)}} \right]$$

即

$$F_\sigma(\mathbf{x}^* + \mathbf{v}) + F_\sigma(\mathbf{x}^* - \mathbf{v}) < 2F_\sigma(\mathbf{x}^*)$$

这表明 $F_\sigma(\mathbf{x}^* + \mathbf{v}) < F_\sigma(\mathbf{x}^*)$ 或 $F_\sigma(\mathbf{x}^* - \mathbf{v}) < F_\sigma(\mathbf{x}^*)$. 这与 \mathbf{x}^* 为模型 (CE) 的最优解相矛盾, 所以假设不成立, 即矩阵 A_k 是列满秩的. \square

上述引理 1 表明, 模型 (CE) 的最优解 \mathbf{x}^* 满足 $\|\mathbf{x}^*\|_0 = k \leq n$, 即模型的最优解包含在有限集 $\Gamma = \{\mathbf{x} | A\mathbf{x} = \mathbf{b}, A_{\text{support}(\mathbf{x})} \text{ 是列满秩矩阵}\}$ 中.

定理 1. 对任意感知矩阵 A 和测量向量 \mathbf{b} , 存在一个依赖于 A 和 \mathbf{b} 的常数 $\sigma(A, \mathbf{b})$, 当 $0 < \sigma < \sigma(A, \mathbf{b})$ 时, 模型 (CE) 与模型 (P_0) 等价.

证明. 要证明模型 (CE) 与模型 (P_0) 等价, 则只需证明模型 (CE) 的最优解为模型 (P_0) 的最优解即可.

根据引理 1, 可以得到以下等价关系

$$\min_{\mathbf{x}} F_\sigma(\mathbf{x}) \quad \text{s. t. } A\mathbf{x} = \mathbf{b} \Leftrightarrow \min_{\mathbf{x} \in \Gamma} F_\sigma(\mathbf{x})$$

说明对任意 σ , (CE) 模型的最优解 $\tilde{\mathbf{x}} \in \Gamma$, 其中 Γ 为有限集.

下面证明存在一个常值 $\sigma(A, \mathbf{b})$, 当 $0 < \sigma < \sigma(A, \mathbf{b})$ 时, 存在唯一一个点 $\tilde{\mathbf{x}} \in \Gamma$ 均是 (CE) 最小化模型的解, 下面利用反证法进行求证.

取一个固定序列 $\{\sigma_i^{(1)}\} \rightarrow 0$, 由于对每个 σ , 集合 Γ 中均存在一个点为模型对应的解, 再根据集合 Γ 为有限集, 所以必然存在一个点能反复解决对应的 (CE) 模型. 不妨假设该序列为 $\{\sigma_i^{(1)}\}$ 本身, 对应的解为 $\tilde{\mathbf{x}}_1$.

又假设存在另一个序列 $\{\sigma_i^{(2)}\} \rightarrow 0$ 和一个异于 $\tilde{\mathbf{x}}_1$ 的集合 Γ 中的点 $\tilde{\mathbf{x}}_2$, 使得当 $\sigma \in \{\sigma_i^{(2)}\}$ 时, $\tilde{\mathbf{x}}_2$ 均为 (CE) 模型的解.

结合序列 $\{\sigma_i^{(1)}\}$ 和 $\{\sigma_i^{(2)}\}$, 构造如下新序列 $\{\sigma_i^{(3)}\}$

$$\sigma_i^{(3)} = \begin{cases} \sigma_i^{(1)}, & i \text{ 为奇数} \\ \sigma_i^{(2)}, & i \text{ 为偶数} \end{cases} \quad (13)$$

显然地, 新序列 $\{\sigma_i^{(3)}\} \rightarrow 0$, 因此必然存在集合 Γ 中的一个点为对应模型 (CE) 的解. 由于序列 $\{\sigma_i^{(3)}\}$ 由 $\{\sigma_i^{(1)}\}$ 和 $\{\sigma_i^{(2)}\}$ 构成, 则这个点一定为 $\tilde{\mathbf{x}}_1$ 和 $\tilde{\mathbf{x}}_2$ 中的一个. 但是无论取哪一个点, 均与 $\tilde{\mathbf{x}}_1$ 为 (CE) 模型对应序列 $\{\sigma_i^{(1)}\}$ 的解和 $\tilde{\mathbf{x}}_2$ 为 (CE) 模型对应序列 $\{\sigma_i^{(2)}\}$ 的解中的其中一种假设相矛盾.

因此存在一个常值 $\sigma(A, \mathbf{b})$, 当 $0 < \sigma < \sigma(A, \mathbf{b})$ 时, 存在唯一一个点 $\tilde{\mathbf{x}} \in \Gamma$, 均是 (CE) 最小化模型的解. 即 $F_\sigma(\tilde{\mathbf{x}}) = \min_{A\mathbf{x}=\mathbf{b}} F_\sigma(\mathbf{x}) \leq \|\mathbf{x}\|_0$, 其中 \mathbf{x} 为可行域中的任意一点. 令 σ 趋于零, 有 $\lim_{\sigma \rightarrow 0} F_\sigma(\tilde{\mathbf{x}}) = \|\tilde{\mathbf{x}}\|_0 \leq \|\mathbf{x}\|_0$, 即 $\|\tilde{\mathbf{x}}\|_0 \leq \|\mathbf{x}\|_0$, 则有 $\tilde{\mathbf{x}}$ 为模型 (P_0) 的最优解, 即 $\tilde{\mathbf{x}}$ 同为模型 (CE)、模型 (P_0) 的最优解, 所以定理 1 成立. \square

3 模型转换

为克服函数 $F_\sigma(\mathbf{x})$ 在原点不可微带来的不便, 可以将稀疏优化模型 (CE) 转换为以下优化问题. 令未知量 $\mathbf{x} = \mathbf{u} - \mathbf{v}$, 其中 $\mathbf{u}, \mathbf{v} \in \mathbf{R}^m$, \mathbf{u} 拥有 \mathbf{x} 中所有正元, 其余元素为零; \mathbf{v} 拥有 \mathbf{x} 中所有负元的

绝对值, 其余元素也为零. 用 $\mathbf{z} = [\mathbf{u}^T, \mathbf{v}^T]^T \in \mathbf{R}^{2m}$ 表示拼接向量. 经过替换, 易得

$$F_\sigma(\mathbf{z}) = F_\sigma(\mathbf{x}) = \sum_{i=1}^{2m} [1 - e^{-\frac{z_i}{\sigma(z_i + \sigma)}}]$$

此时, 约束条件 $\mathbf{A}\mathbf{x} = \mathbf{b}$ 转换为 $[\mathbf{A}, -\mathbf{A}]\mathbf{z} = \mathbf{b}$. 那么无噪模型 (CE) 转换为

$$\min_{\mathbf{z}} F_\sigma(\mathbf{z}) \quad \text{s. t. } [\mathbf{A}, -\mathbf{A}]\mathbf{z} = \mathbf{b}, \mathbf{z} \geq 0 \quad (14)$$

为了便于求解式 (14), 利用文献 [25] 中介绍的 MM 优化方法. MM 优化方法指当目标函数较难实现优化时, 通常可以选择更容易优化的替代目标函数, 当替代函数满足一定的条件时, 其最优解能够无限逼近原目标函数的最优解. 通过利用凹函数 $F_\sigma(\mathbf{z})$ 的一阶判别条件以及 MM 优化方法对目标函数 $F_\sigma(\mathbf{z})$ 进行放缩, 可得

$$\begin{aligned} F_\sigma(\mathbf{z}) &\leq F_\sigma(\tilde{\mathbf{z}}) + \langle \mathbf{z} - \tilde{\mathbf{z}}, \nabla F_\sigma(\tilde{\mathbf{z}}) \rangle \leq \\ &F_\sigma(\tilde{\mathbf{z}}) + \langle \mathbf{z} - \tilde{\mathbf{z}}, \nabla F_\sigma(\tilde{\mathbf{z}}) \rangle + \\ &\lambda(\|\mathbf{z} - \tilde{\mathbf{z}}\|_2^2 + \|\mathbf{z} - \tilde{\mathbf{z}}\|_1) \end{aligned}$$

记

$$\begin{aligned} H_\sigma(\mathbf{z}, \tilde{\mathbf{z}}) &= F_\sigma(\tilde{\mathbf{z}}) + \langle \mathbf{z} - \tilde{\mathbf{z}}, \nabla F_\sigma(\tilde{\mathbf{z}}) \rangle + \\ &\lambda(\|\mathbf{z} - \tilde{\mathbf{z}}\|_2^2 + \|\mathbf{z} - \tilde{\mathbf{z}}\|_1) \end{aligned}$$

其中, $\mathbf{z}, \tilde{\mathbf{z}}$ 是可行域中的点, $\lambda (\lambda > 1/\sigma^2)$ 为正常数, $H_\sigma(\mathbf{z}, \tilde{\mathbf{z}})$ 为 $F_\sigma(\mathbf{z})$ 的一个上界函数. 忽略常数后, 式 (14) 的解可由下式迭代求得

$$\begin{aligned} \mathbf{z}_{k+1}^\sigma &= \arg \min_{\mathbf{z}} \left\{ \langle \nabla F_\sigma(\mathbf{z}_k^\sigma), \mathbf{z} \rangle + \lambda(\|\mathbf{z} - \tilde{\mathbf{z}}\|_2^2 + \right. \\ &\left. \|\mathbf{z} - \tilde{\mathbf{z}}\|_1) \mid [\mathbf{A}, -\mathbf{A}]\mathbf{z} = \mathbf{b}, \mathbf{z} \geq 0 \right\} \end{aligned} \quad (15)$$

其中, $\mathbf{z}_k^\sigma (k = 1, 2, 3, \dots)$ 为可行域中的点, 取值与 σ 相关.

接下来讨论由式 (15) 产生的序列 $\{\mathbf{z}_k^\sigma\}$ 的收敛性.

命题 1. 式 (15) 产生的序列 $\{\mathbf{z}_k^\sigma\}$ 收敛到式 (14) 的局部极小值.

证明. 1) 首先证明式 (15) 迭代产生的函数序列 $\{F_\sigma(\mathbf{z}_k^\sigma)\}$ 的有界性和收敛性.

根据凹函数的一阶判别条件有

$$F_\sigma(\mathbf{z}) \leq F_\sigma(\mathbf{z}_k^\sigma) + \langle \mathbf{z} - \mathbf{z}_k^\sigma, \nabla F_\sigma(\mathbf{z}_k^\sigma) \rangle \quad (16)$$

再由式 (15) 的定义有

$$\begin{aligned} \langle \mathbf{z}_k^\sigma, \nabla F_\sigma(\mathbf{z}_k^\sigma) \rangle &\geq \langle \mathbf{z}_{k+1}^\sigma, \nabla F_\sigma(\mathbf{z}_k^\sigma) \rangle + \\ &\lambda(\|\mathbf{z}_{k+1}^\sigma - \mathbf{z}_k^\sigma\|_2^2 + \|\mathbf{z}_{k+1}^\sigma - \mathbf{z}_k^\sigma\|_1) \end{aligned} \quad (17)$$

在式 (16) 中令 $\mathbf{z} = \mathbf{z}_{k+1}^\sigma$, 结合式 (17) 得

$$F_\sigma(\mathbf{z}_{k+1}^\sigma) - F_\sigma(\mathbf{z}_k^\sigma) \leq \langle \mathbf{z}_{k+1}^\sigma - \mathbf{z}_k^\sigma, \nabla F_\sigma(\mathbf{z}_k^\sigma) \rangle \leq 0$$

即

$$F_\sigma(\mathbf{z}_{k+1}^\sigma) \leq F_\sigma(\mathbf{z}_k^\sigma) \quad (18)$$

其次, 因为

$$F_\sigma(\mathbf{z}) = \sum_{i=1}^{2m} [1 - e^{-\frac{z_i}{\sigma(z_i + \sigma)}}]$$

有

$$0 \leq F_\sigma(\mathbf{z}) \leq 2m \quad (19)$$

所以由式 (18)、(19) 可知, 序列 $\{F_\sigma(\mathbf{z}_k^\sigma)\}$ 是单调有界收敛的.

2) 再证明式 (15) 产生的序列 $\{\mathbf{z}_k^\sigma\}$ 的收敛性.

根据式 (16)、(17) 得

$$\begin{aligned} 0 &\leq \lambda(\|\mathbf{z}_{k+1}^\sigma - \mathbf{z}_k^\sigma\|_2^2 + \|\mathbf{z}_{k+1}^\sigma - \mathbf{z}_k^\sigma\|_1) \leq \\ &\langle \mathbf{z}_k^\sigma - \mathbf{z}_{k+1}^\sigma, \nabla F_\sigma(\mathbf{z}_k^\sigma) \rangle \leq \\ &F_\sigma(\mathbf{z}_k^\sigma) - F_\sigma(\mathbf{z}_{k+1}^\sigma) \end{aligned}$$

所以有

$$\begin{aligned} \lambda \sum_{i=1}^N (\|z_{i+1}^\sigma - z_i^\sigma\|_2^2 + \|z_{i+1}^\sigma - z_i^\sigma\|_1) &\leq \\ F_\sigma(\mathbf{z}_1^\sigma) - F_\sigma(\mathbf{z}_N^\sigma) \end{aligned}$$

当 $N \rightarrow \infty$, 利用式 (19) 得

$$\lambda \sum_{i=1}^N (\|z_{i+1}^\sigma - z_i^\sigma\|_2^2 + \|z_{i+1}^\sigma - z_i^\sigma\|_1)$$

是有界的, 所以序列 $\{\mathbf{z}_k^\sigma\}$ 收敛.

3) 最后证明式 (15) 产生的序列 $\{\mathbf{z}_k^\sigma\}$ 收敛到式 (14) 的局部极小值. 假设式 (15) 迭代产生的序列 $\{\mathbf{z}_k^\sigma\}$ 收敛到 $\tilde{\mathbf{z}}^\sigma$, 即

$$\lim_{k \rightarrow \infty} \mathbf{z}_k^\sigma = \tilde{\mathbf{z}}^\sigma \quad (20)$$

下面证明 $\tilde{\mathbf{z}}^\sigma$ 是

$$\begin{aligned} \min_{\mathbf{z}} \langle \nabla F_\sigma(\tilde{\mathbf{z}}^\sigma), \mathbf{z} \rangle + \lambda(\|\mathbf{z} - \tilde{\mathbf{z}}^\sigma\|_2^2 + \|\mathbf{z} - \tilde{\mathbf{z}}^\sigma\|_1) \\ \text{s. t. } [\mathbf{A}, -\mathbf{A}]\mathbf{z} = \mathbf{b}, \mathbf{z} \geq 0 \end{aligned} \quad (21)$$

的最优解.

记

$$\begin{cases} G_k(\mathbf{z}) = \langle \nabla F_\sigma(\mathbf{z}_k^\sigma), \mathbf{z} \rangle + \lambda(\|\mathbf{z} - \mathbf{z}_k^\sigma\|_2^2 + \|\mathbf{z} - \mathbf{z}_k^\sigma\|_1) \\ G_0(\mathbf{z}) = \langle \nabla F_\sigma(\tilde{\mathbf{z}}^\sigma), \mathbf{z} \rangle + \lambda(\|\mathbf{z} - \tilde{\mathbf{z}}^\sigma\|_2^2 + \|\mathbf{z} - \tilde{\mathbf{z}}^\sigma\|_1) \\ \mathbf{z}_0^\sigma = \arg \min_{\mathbf{z}} \{ G_0(\mathbf{z}) \mid [\mathbf{A}, -\mathbf{A}]\mathbf{z} = \mathbf{b}, \mathbf{z} \geq 0 \} \end{cases}$$

接下来讨论是否有 $\mathbf{z}_0^\sigma = \tilde{\mathbf{z}}^\sigma$ 成立. 由于

$$\begin{aligned}
& |G_k(z) - G_0(z)| = \\
& |\langle \nabla F_\sigma(z_k^\sigma) - \nabla F_\sigma(\tilde{z}^\sigma), z \rangle + \lambda(\|z - z_k^\sigma\|_2^2 - \\
& \|z - \tilde{z}^\sigma\|_2^2) + \lambda(\|z - z_k^\sigma\|_1 - \|z - \tilde{z}^\sigma\|_1)| \leq \\
& \|\nabla F_\sigma(z_k^\sigma) - \nabla F_\sigma(\tilde{z}^\sigma)\|_2 \cdot \|z\|_2 + \lambda\|\tilde{z}^\sigma - z_k^\sigma\|_2^2 + \\
& \lambda\|z - \tilde{z}^\sigma\|_1 + \|\tilde{z}^\sigma - z_k^\sigma\|_1 - \|z - \tilde{z}^\sigma\|_1 = \\
& \|\nabla F_\sigma(z_k^\sigma) - \nabla F_\sigma(\tilde{z}^\sigma)\|_2 \cdot \|z\|_2 + \lambda\|\tilde{z}^\sigma - z_k^\sigma\|_2^2 + \\
& \lambda\|\tilde{z}^\sigma - z_k^\sigma\|_1
\end{aligned}$$

又注意到 $\aleph = \{z_k^\sigma\}_{k=1}^\infty \cup \{\tilde{z}^\sigma\} \cup \{z_0^\sigma\}$ 有界, 所以存在 $M > 0$, 使得对任意 $z \in \aleph$, 有 $\|z\|_2 < M$. 所以有

$$\begin{aligned}
|G_k(z) - G_0(z)| \leq M\|\nabla F_\sigma(z_k^\sigma) - \nabla F_\sigma(\tilde{z}^\sigma)\|_2 + \\
\lambda\|\tilde{z}^\sigma - z_k^\sigma\|_2^2 + \lambda\|\tilde{z}^\sigma - z_k^\sigma\|_1 \quad (22)
\end{aligned}$$

再利用假设式 (20) 有

$$\lim_{k \rightarrow \infty} \|z_k^\sigma - \tilde{z}^\sigma\|_2^2 = 0, \quad \lim_{k \rightarrow \infty} \|z_k^\sigma - \tilde{z}^\sigma\|_1 = 0 \quad (23)$$

即对任意 $\varepsilon_1 > 0$, 存在 $K_1 > 0$, 对任意 $k > K_1$, 有

$$\lim_{k \rightarrow \infty} \|z_k^\sigma - \tilde{z}^\sigma\|_2^2 < \varepsilon_1, \quad \lim_{k \rightarrow \infty} \|z_k^\sigma - \tilde{z}^\sigma\|_1 < \varepsilon_1 \quad (24)$$

又根据初等函数 $\nabla F_\sigma(z)$ 在定义区间内关于变量 z 连续, 有

$$\lim_{k \rightarrow \infty} \nabla F_\sigma(z_k^\sigma) = \nabla F_\sigma(\tilde{z}^\sigma) \quad (25)$$

由此可知, 对任意 $\varepsilon_2 > 0$, 存在 $K_2 > 0$, 对任意 $k > K_2$, 有

$$\|\nabla F_\sigma(z_k^\sigma) - \nabla F_\sigma(\tilde{z}^\sigma)\|_2 < \varepsilon_2 \quad (26)$$

故结合式 (23)、式 (25) 可知, 对任意 $z \in \aleph$, 对于式 (22) 可得

$$\lim_{k \rightarrow \infty} |G_k(z) - G_0(z)| = 0 \quad (27)$$

即对任意 $z \in \aleph$, 有任意 $\varepsilon_3 > 0$, 存在 $K_3 > 0$, 对任意 $k > K_3$, 有

$$-\varepsilon_3 < G_k(z) - G_0(z) < \varepsilon_3 \quad (28)$$

根据 z_0^σ 和 $G_0(z)$ 的定义以及不等式 (28) 有

$$\begin{cases} G_0(z_0^\sigma) \leq G_0(z_k^\sigma) \\ G_0(z_k^\sigma) - G_k(z_k^\sigma) < \varepsilon_3 \end{cases}$$

所以有

$$G_0(z_0^\sigma) - G_k(z_k^\sigma) < \varepsilon_3$$

则由此可得

$$\begin{aligned}
& \langle \nabla F_\sigma(\tilde{z}^\sigma), z_0^\sigma \rangle - \langle \nabla F_\sigma(z_k^\sigma), z_k^\sigma \rangle + \\
& \lambda(\|z_0^\sigma - \tilde{z}^\sigma\|_2^2 + \|z_0^\sigma - \tilde{z}^\sigma\|_1) < \varepsilon_3 \quad (29)
\end{aligned}$$

令 $\varepsilon = \min(\varepsilon_1, \varepsilon_2, \varepsilon_3)$, $K = \max(K_1, K_2, K_3)$, 当 $k > K$ 时, 对不等式 (29) 有

$$\begin{aligned}
& \langle \nabla F_\sigma(z_k^\sigma), z_0^\sigma - z_k^\sigma \rangle + \lambda(\|z_0^\sigma - \tilde{z}^\sigma\|_2^2 + \|z_0^\sigma - \tilde{z}^\sigma\|_1) < \\
& \varepsilon + \langle \nabla F_\sigma(z_k^\sigma) - \nabla F_\sigma(\tilde{z}^\sigma), z_0^\sigma \rangle < \\
& \varepsilon + \|\nabla F_\sigma(z_k^\sigma) - \nabla F_\sigma(\tilde{z}^\sigma)\|_2 \cdot \|z_0^\sigma\|_2 < \\
& \varepsilon + M\varepsilon \quad (30)
\end{aligned}$$

由式 (30) 可知

$$\sum_{i=1}^{2m} \left\{ e^{-\frac{1}{\sigma} \frac{1}{(z_k^\sigma)_i + \sigma}} \frac{(z_0^\sigma)_i - (z_k^\sigma)_i}{((z_k^\sigma)_i + \sigma)^2} + \lambda((z_0^\sigma)_i - \tilde{z}_i^\sigma)^2 + \right. \\
\left. |(z_0^\sigma)_i - \tilde{z}_i^\sigma| \right\} < (M+1)\varepsilon$$

对左端加减同一项, 有

$$\sum_{i=1}^{2m} \left\{ e^{-\frac{1}{\sigma} \frac{1}{(z_k^\sigma)_i + \sigma}} \frac{(z_0^\sigma)_i - \tilde{z}_i^\sigma + \tilde{z}_i^\sigma - (z_k^\sigma)_i}{((z_k^\sigma)_i + \sigma)^2} + \right. \\
\left. \lambda((z_0^\sigma)_i - \tilde{z}_i^\sigma)^2 + |(z_0^\sigma)_i - \tilde{z}_i^\sigma| \right\} < (M+1)\varepsilon \quad (31)$$

对式 (31) 移项, 有

$$\begin{aligned}
& \sum_{i=1}^{2m} |(z_0^\sigma)_i - \tilde{z}_i^\sigma| \left[e^{-\frac{1}{\sigma} \frac{1}{(z_k^\sigma)_i + \sigma}} \frac{\text{sgn}((z_0^\sigma)_i - \tilde{z}_i^\sigma)}{((z_k^\sigma)_i + \sigma)^2} + \right. \\
& \left. \lambda(|(z_0^\sigma)_i - \tilde{z}_i^\sigma| + 1) \right] < \\
& (M+1)\varepsilon + \sum_{i=1}^{2m} \left[e^{-\frac{1}{\sigma} \frac{1}{(z_k^\sigma)_i + \sigma}} \frac{|(z_k^\sigma)_i - \tilde{z}_i^\sigma|}{((z_k^\sigma)_i + \sigma)^2} \right] < \\
& (M+1 + \frac{1}{\sigma^2})\varepsilon \quad (32)
\end{aligned}$$

再对式 (32) 不等号左端进行放缩, 得

$$\begin{aligned}
& \sum_{i=1}^{2m} |(z_0^\sigma)_i - \tilde{z}_i^\sigma| \left[e^{-\frac{1}{\sigma} \frac{1}{(z_k^\sigma)_i + \sigma}} \frac{\text{sgn}((z_0^\sigma)_i - \tilde{z}_i^\sigma)}{((z_k^\sigma)_i + \sigma)^2} + \right. \\
& \left. \lambda(|(z_0^\sigma)_i - \tilde{z}_i^\sigma| + 1) \right] > \\
& \sum_{i=1}^{2m} |(z_0^\sigma)_i - \tilde{z}_i^\sigma| \left[-\frac{1}{\sigma^2} + \lambda(|(z_0^\sigma)_i - \tilde{z}_i^\sigma| + 1) \right] > \\
& \sum_{i=1}^{2m} |(z_0^\sigma)_i - \tilde{z}_i^\sigma| \left(-\frac{1}{\sigma^2} + \lambda \right) \geq 0 \quad (33)
\end{aligned}$$

最后结合式 (32)、(33) 得

$$0 \leq \sum_{i=1}^{2m} |(z_0^\sigma)_i - \tilde{z}_i^\sigma| \left(-\frac{1}{\sigma^2} + \lambda \right) < (M+1 + \frac{1}{\sigma^2})\varepsilon$$

即

$$0 \leq (\lambda - \frac{1}{\sigma^2})\|z_0^\sigma - \tilde{z}^\sigma\|_1 < (M+1 + \frac{1}{\sigma^2})\varepsilon$$

从而由 ε 的任意性, 有 $\|z_0^\sigma - \tilde{z}^\sigma\|_1 = 0$, 即有 $z_0^\sigma = \tilde{z}^\sigma$, 由此可知 \tilde{z}^σ 是式 (21) 的最优解. 令

$$\begin{cases} \langle \nabla F_\sigma(\tilde{z}^\sigma), z \rangle + \lambda[\|z - \tilde{z}^\sigma\|_2^2 + \|z - \tilde{z}^\sigma\|_1] = G(z) \\ [A, -A]z - \mathbf{b} = H(z) = (h_1(z), \dots, h_n(z))^T \\ z = \mathbf{W}(z) = (w_1(z), \dots, w_{2m}(z))^T \end{cases}$$

优化模型 (21) 简记为

$$\min_z G(z) \quad \text{s. t. } H(z) = \mathbf{0}, \mathbf{W}(z) \geq \mathbf{0} \quad (34)$$

约束优化问题式 (34) 的解为 \tilde{z}^σ , 说明在极小值点 \tilde{z}^σ 处必不存在可行域内的下降方向, 即不存在非零向量 $\mathbf{p} \in \mathbf{R}^{2m}$, 使得

$$\begin{cases} \nabla G(\tilde{z}^\sigma)^T \mathbf{p} < 0 \\ \nabla h_i(\tilde{z}^\sigma)^T \mathbf{p} = 0 \\ \nabla w_j(\tilde{z}^\sigma)^T \mathbf{p} \geq 0 \end{cases}$$

其中, $i, j \in \mathbf{I}(\tilde{z}^\sigma)$, $\mathbf{I}(\tilde{z}^\sigma)$ 为点 \tilde{z}^σ 的起作用下标集.

再由

$$\nabla G(z) = \nabla F_\sigma(\tilde{z}^\sigma) + \lambda[2|z - \tilde{z}^\sigma| + \text{sgn}(z - \tilde{z}^\sigma)]$$

易得

$$\nabla G(\tilde{z}^\sigma) = \nabla F_\sigma(\tilde{z}^\sigma)$$

所以对于稀疏优化模型 (14) 在点 \tilde{z}^σ 处也不存在可行域内的下降方向, 即 \tilde{z}^σ 为模型 (14) 的局部极小值. \square

4 算法设计

为给出模型 (14) 的有效求解算法, 需要对算法初始值的选择进行分析, 首先给出复合函数的如下性质.

引理 2. $\lim_{\sigma \rightarrow \infty} \sigma^2 F_\sigma(\mathbf{x}) = \|\mathbf{x}\|_1$.

证明. 已知 $f_\sigma(x) = 1 - e^{-\frac{|x|}{\sigma(|x| + \sigma)}}$, 将 $f_\sigma(x)$ 有关 $|x|$ 的麦克劳林公式记为

$$\begin{aligned} f_\sigma(x) &= f_\sigma(0) + \frac{|x|}{1!} f'_\sigma(0) + \frac{|x|^2}{2!} f''_\sigma(0) + \dots = \\ &= 0 + \frac{|x|}{1!} \cdot \frac{1}{\sigma^2} + g_\sigma(x) = \\ &= \frac{|x|}{\sigma^2} + g_\sigma(x) \end{aligned}$$

其中, $g_\sigma(x)$ 为麦克劳林公式的余项, $f_\sigma^{(n)}(0)$ 表示 $f_\sigma(x)$ 在 0 点处的 n 阶导数.

所以有

$$F_\sigma(\mathbf{x}) = \sum_{i=1}^m f_\sigma(x_i) = \sum_{i=1}^m \left[\frac{|x_i|}{\sigma^2} + g_\sigma(x_i) \right] \quad (35)$$

由于

$$\begin{aligned} f_\sigma^{(2)}(x) &= -\frac{1}{(|x| + \sigma)^3} e^{-\frac{|x|}{\sigma(|x| + \sigma)}} - \\ &= \frac{1}{(|x| + \sigma)^4} e^{-\frac{|x|}{\sigma(|x| + \sigma)}} \end{aligned}$$

其中, 关于 $e^{-|x|/\sigma(|x| + \sigma)}/(|x| + \sigma)^3$ 、 $e^{-|x|/\sigma(|x| + \sigma)}/(|x| + \sigma)^4$ 的任意阶导数满足有关 $(|x| + \sigma)^{-1}$ 的次方数不低于 3 次, 则 $f_\sigma^{(n)}(0)$ ($n \geq 2$) 关于参数 σ^{-1} 的次方数不低于 3 次. 由此可得

$$\lim_{\sigma \rightarrow \infty} \sigma^2 g_\sigma(x) = 0 \quad (36)$$

利用式 (35)、(36), 有

$$\begin{aligned} \lim_{\sigma \rightarrow \infty} \sigma^2 F_\sigma(\mathbf{x}) &= \lim_{\sigma \rightarrow \infty} \sum_{i=1}^m [|x_i| + \sigma^2 g_\sigma(x_i)] = \\ &= \lim_{\sigma \rightarrow \infty} (|x_i|) = \|\mathbf{x}\|_1 \quad \square \end{aligned}$$

引理 2 表明, 在 $\sigma \rightarrow \infty$ 时, 最小化 $F_\sigma(\mathbf{x})$ 近似于最小化 $\|\mathbf{x}\|_1$. $\|\mathbf{x}\|_1$ 的凸性使得算法陷入局部极值的可能性减少, 从而提高算法重构精度. 鉴于此, 本文在设计算法时取初始值为以下问题

$$\min_{\mathbf{x}} \|\mathbf{x}\|_1 \quad \text{s. t. } A\mathbf{x} = \mathbf{b}$$

的解.

在此基础上, 提出如下算法.

步骤 1. 输入: $A, \mathbf{b}, \varepsilon_1, \varepsilon_2, \varepsilon_3$.

步骤 2. 初始化:

- 1) 选取一组下降序列 $\sigma, \sigma_{k+1} = c \cdot \sigma_k, c \in (0, 0.5)$.
- 2) 内、外层循环的终止阈值 $\varepsilon_1, \varepsilon_2, \varepsilon_3 > 0$.
- 3) 设初始值

$$\mathbf{x}_0 = \arg \min_{\mathbf{x}} \{\|\mathbf{x}\|_1 \mid A\mathbf{x} = \mathbf{b}\}$$

$$\mathbf{z}_0 = [(\mathbf{x}_0^+)^T, (\mathbf{x}_0^-)^T]^T$$

步骤 3. 算法迭代:

- 1) $i = 0, \sigma = \sigma_0$.
- 2) while $d_1 > \varepsilon_1$ do
 - a) $\{\hat{\mathbf{z}}_0 = \mathbf{z}_k, l = 0, k = k + 1$
 - b) while $d_2 > \varepsilon_2$ do
 - c) $\{l = l + 1$
 - d) $\hat{\mathbf{z}}_l = \arg \min_{\mathbf{z}} \{G_{l-1}(\mathbf{z}) \mid [A, -A]\mathbf{z} = \mathbf{b}, \mathbf{z} \geq \mathbf{0}\}$
 - e) $d_2 = \frac{\|\hat{\mathbf{z}}_l - \hat{\mathbf{z}}_{l-1}\|_2}{\|\hat{\mathbf{z}}_{l-1}\|_2}$
 - f) $\mathbf{z}_k = \hat{\mathbf{z}}_l$
 - g) $d_1 = \frac{\|\mathbf{z}_k - \mathbf{z}_{k-1}\|_2}{\|\mathbf{z}_{k-1}\|_2}$
 - h) $\sigma = c \cdot \sigma$
 - i) $\mathbf{x}_k = \mathbf{z}_k(1:m) - \mathbf{z}_k(m+1:\text{end})$
 - j) $U = \text{diag} \left\{ \frac{1}{(x_{k-1})_1 + \varepsilon_3}, \dots, \frac{1}{(x_{k-1})_m + \varepsilon_3} \right\}$
 - k) $\mathbf{x}_k = \arg \min_{\mathbf{x}} \{\|U\mathbf{x}\|_1 \mid A\mathbf{x} = \mathbf{b}\}$

步骤 4. 输出稀疏向量: $\mathbf{x} = \mathbf{x}_k$.

其中, 步骤 3 算法迭代环节由内、外两部分循环构成. 外循环为步骤 a) ~ b) 以及 f) ~ h), 循环利用参数 σ 实现复合函数对 L_0 范数的逐次逼近. 内循环即步骤 c) ~ e), 结合外点罚函数法^[26] 和共轭梯

度法迭代求解模型 (14), 具体为利用外点罚函数法引入正则参数实现模型 (14) 的无约束转换, 再通过共轭梯度法求解无约束优化问题. 最后在步骤 j) ~ k), 针对共轭梯度法求得的非稀疏解, 利用加权 L_1 范数最小化^[27] 进行稀疏化处理. 其中, d_1 和 d_2 分别表示外循环和内循环连续迭代的解之间的相对误差, 并用于判断循环是否停止. 值得关注的是, 整个算法的核心环节即步骤 d), 是有关本文模型 (14) 的凸优化求解过程, 具体步骤如下.

模型利用外点罚函数法引入正数 M , 本文取 $M = 1$, 其放大系数取为 5, 将式 (15) 转化为以下无约束问题

$$z_{k+1}^\sigma = \arg \min_z L_{z_k^\sigma}(z, M) \quad (37)$$

其中

$$L_{z_k^\sigma}(z, M) = \langle \nabla F_\sigma(z_k^\sigma), z \rangle + \lambda[\|z - z_k^\sigma\|_2^2 + \|z - z_k^\sigma\|_1] + M[\mathbf{I}^\top \cdot U(z) + \mathbf{I}^\top \cdot H(z)] \quad (38)$$

式 (38) 中, \mathbf{I} 是大小为 $2m \times 1$ 的单位向量.

$$U(z) = (\min(0, z_1^2), \dots, \min(0, z_{2m}^2))^\top$$

$$H(z) = \text{diag}\{([A, -A]z - \mathbf{b}) \cdot ([A, -A]z - \mathbf{b})\}$$

实验中, 函数 $L_{z_k^\sigma}(z, M)$ 有关 z 的一阶、二阶次梯度^[28] 分别取为

$$\nabla L_{z_k^\sigma}(z, M) = \nabla F_\sigma(z_k^\sigma) + \lambda[2(z - z_k^\sigma) + \text{sgn}(z - z_k^\sigma)] + 2M\{V(z) + [A^\top, -A^\top]^\top \cdot H(z)\} \quad (39)$$

$$\nabla^2 L_{z_k^\sigma}(z, M) = 2\lambda\mathbf{I} + [A^\top, -A^\top]^\top [A, -A] \mathbf{I} \quad (40)$$

其中, $V(z) = (\min(0, z_1), \dots, \min(0, z_{2m}))^\top$.

再利用共轭梯度法迭代求解式 (37) 的 z_{k+1}^σ , 其迭代更新格式为

$$z_{t+1} = z_t + \alpha_t \cdot p_t \quad (41)$$

其中, 步长因子

$$\alpha_t = \frac{\|\nabla L_{z_k^\sigma}(z_t, M)\|_2^2}{p_t^\top \cdot \nabla^2 L_{z_k^\sigma}(z_t, M) \cdot p_t} \quad (42)$$

下降方向

$$p_t = -\nabla L_{z_k^\sigma}(z_t, M) + p_{t-1} \cdot \rho_{t-1} \quad (43)$$

其中, 令 $z_k^\sigma = z_0$

$$p_0 = \|\nabla L_{z_k^\sigma}(z_0, M)\|, \rho_{t-1} = \frac{\|\nabla L_{z_k^\sigma}(z_t, M)\|_2^2}{\|\nabla L_{z_k^\sigma}(z_{t-1}, M)\|_2^2} \quad (44)$$

综上, 本文提出的基于非凸复合函数的稀疏信号恢复算法 (Non-convex composite sparse, NCCS) 参考了文献 [23] 中的 SCSA 算法, 主要包括三个部分. 一是算法的初始值选择, 取值为 L_1 范数最小化问题的解, 参考文献 [29], 其计算量为 $O(m^3)$. 二是无约束问题的求解, 通过外点罚函数法和共轭梯度法迭代求解最优值. 由式 (39) ~ 式 (44) 可以看出, 第二部分每次迭代的主要操作是内循环的两部分矩阵乘法, 即

$$2M\{V(z) + [A^\top, -A^\top]^\top \cdot H(z)\}$$

$$2M[A^\top, -A^\top]^\top [A, -A] \mathbf{I}$$

计算复杂度分别为 $O(4m + 2mn)$ 和 $O(4m^2 + 4m^2n)$, 则第二部分的计算量为 $O(l(4m + 2mn + 4m^2 + 4m^2n))$, 其中 l 表示内循环的迭代次数. 三是迭代结果的稀疏化, 利用加权 L_1 范数最小化^[27] 对凸优化结果进行稀疏化处理, 其计算量为 $O(m^3)$.

5 实验仿真与结果分析

为验证本文提出的算法在重构性能上的优越性, 本节设计了几组有关 SLO 算法^[2]、BP 算法^[15]、IRLS 算法^[16]、SCSA 算法^[23] 和本文介绍的 NCCS 这 5 种算法在重构性能上的对照实验. 在仿真实验中, 感知矩阵 A 和稀疏原信号 \mathbf{x} 的具体取值为: 矩阵 A 的大小取为 250×500 , 其中矩阵元素服从零均值、单位方差的高斯分布, 矩阵的列具有单位 L_2 范数; 稀疏原信号 \mathbf{x} 的维数取为 500, 非零分项服从正态分布. 实验将讨论信号稀疏度在区间 [20, 110] 范围内各算法的重构情况. 对于等式约束 $A\mathbf{x} = \mathbf{b}$, 在已知感知矩阵 A 和测量向量 \mathbf{b} 的情况下, 分别利用上述 5 种算法对稀疏原信号进行 100 次仿真实验, 讨论算法的平均重构性能. 为展示本文所提出的模型相对于最新文献 [23] 所提算法的先进性, 本文对 5 种算法采用了与文献 [23] 相同的参数选择, 即

1) 对 SLO 算法: 令 $mu = 2$, $\sigma_{\min} = 10^{-4}$, $c = 0.8$, $L = 8$.

2) 对 IRLS 算法: 令 $p = 0.5$.

3) 对 BP 算法: 令 $l_1 = 1$.

4) 对 SCSA 算法: 令 $\varepsilon_1 = 10^{-3}$, $\varepsilon_2 = 10^{-2}$, $c = 0.1$.

5) 本文 NCCS 算法: 令 $\varepsilon_1 = 10^{-3}$, $\varepsilon_2 = 10^{-2}$, $\varepsilon_3 = 10^{-1}$, $c = 0.1$, $\sigma_0 = \min\{2\|\mathbf{x}_0\|_0, \alpha\}$, 其中 α 是待定数. 仿真实验一将给出模拟实验, 以获得最佳实验数值.

实验中的重构性能包括:

1) 重构信噪比 (Signal noise ratio):

$$SNR = 10 \cdot \log_{10} \left(\frac{\|\tilde{\mathbf{x}}\|_2^2}{\|\tilde{\mathbf{x}} - \hat{\mathbf{x}}\|_2^2} \right)$$

2) 重构误差 (Mean square error):

$$MSE = \|\tilde{\mathbf{x}} - \hat{\mathbf{x}}\|_2^2$$

3) 归一化均方差 (Normalized mean square error):

$$NMSE = \frac{\|\tilde{\mathbf{x}} - \hat{\mathbf{x}}\|_2^2}{\|\tilde{\mathbf{x}}\|_2^2}$$

4) 支撑集恢复成功率 (Recovery success rate of the support set):

$$RSS = \frac{T_2}{T_1}$$

其中, $\tilde{\mathbf{x}}$ 表示稀疏原信号, $\hat{\mathbf{x}}$ 表示算法的恢复信号; T_1 为稀疏原信号的支撑集, T_2 为恢复信号在对应支撑集 T_1 上的支撑恢复集. 具体取值分别为

$$T_1 = \{i | \tilde{x}_i \neq 0, i = 1, \dots, m\}$$

$$T_2 = \{j | |\text{sgn}(\hat{x}_j)| = 1, j \in T_1\}$$

仿真实验一. 待定数 α 的最佳取值.

首先对 NCCS 算法中待定数 α 进行模拟实验, 观察不同大小的数值 α 对算法运行时间的影响, 以获得更优的实验参数. 为节省计算时长, 考虑 α 大小为 0.1 到 0.9, 间隔为 0.1, 分别进行 100 次仿真实验. 从图 3 可以看出, 各 α 值对算法运行时间的影响差距不大, 总体趋势是算法运行时间随 α 的增加而减少; 在 $\alpha = 0.8$ 时, 算法在稀疏度 $k \leq 100$ 时皆能保证运行时间最短. 信号的稀疏度 k 是指信号向量中非零元素的个数, 为便于编程, 在后续实验中待定数 α 全部选定为 0.8.

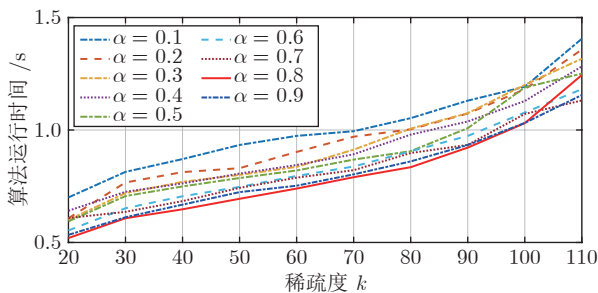


图 3 待定数 α 对 NCCS 算法运行时间的影响

Fig.3 The influence of undetermined number α on the running time of NCCS algorithm

仿真实验二. NCCS 算法对稀疏原信号重构的实验.

图 4 是稀疏原信号以及由 NCCS 算法获得的恢复信号的仿真结果. 图 4 的稀疏度取值为 $k = 65$.

由图 4 可见, 恢复信号和稀疏原信号基本吻合, 实验结果显示算法的重构误差为 2.6933×10^{-17} . 由此可以看出, 本文提出的算法恢复出的稀疏信号很接近稀疏原信号.

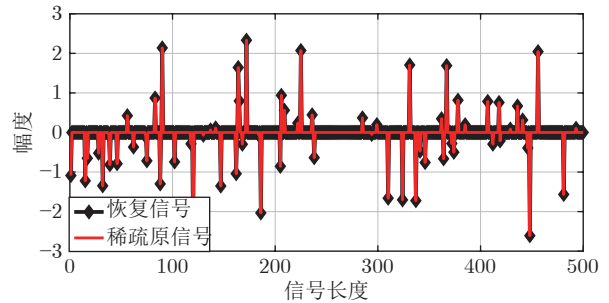


图 4 NCCS 算法的一维信号重构仿真图, 信号大小为 500×1 , 稀疏度为 65

Fig.4 One-dimensional signal reconstruction simulation diagram of NCCS algorithm, the signal size is 500×1 , the sparsity is 65

仿真实验三. 各算法的重构性能和稀疏度的变化关系.

图 5 为 5 组实验算法的重构误差 (MSE) 和稀疏度 k 的变化关系. 其中 IRLS 算法的重构误差最高; 由于最速下降法在迭代过程中存在“锯齿效应”^[30], 实验中 SL0 算法的重构误差也很大, 且 SL0 与 BP、SCSA 这 3 种算法都存在重构误差随着稀疏度的增加而增大的趋势; 而 NCCS 算法的重构误差相对最小, 整体变化较稳定.

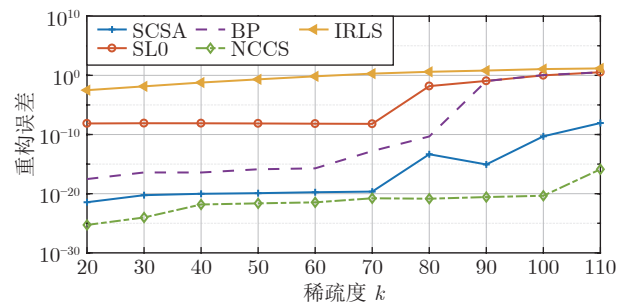


图 5 SL0、IRLS、BP、SCSA、NCCS 5 种算法的重构误差和稀疏度 k 的变化关系

Fig.5 The relationship between the reconstruction error and sparsity of the five algorithms of SL0, IRLS, BP, SCSA, and NCCS

图 6 表示随着稀疏度在区间 $[20, 110]$ 的等量增加, 5 种算法的重构信噪比 (SNR) 的变化情况. 其中 SL0、BP 和 IRLS 3 种算法的稀疏度越高, 重构信噪比越低; NCCS 算法与 SCSA 算法的重构信噪比在稀疏度区间 $[20, 110]$ 上表现较稳定. 在实验的

各组稀疏度下, NCCS 算法的重构信噪比远高于 SL0、BP 和 IRLS 这 3 种算法, 略高于 SCSSA 算法. 所以使用 NCCS 算法对信号的恢复精确度有一定的提高.

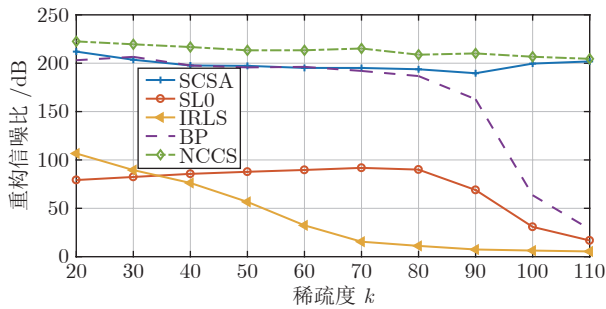


图 6 SL0、IRLS、BP、SCSSA、NCCS 5 种算法的重构信噪比和稀疏度的变化关系

Fig.6 The relationship between the reconstructed signal-to-noise ratio and sparsity of the five algorithms of SL0, IRLS, BP, SCSSA, and NCCS

图 7 展示 5 种算法的重构运行时间和稀疏度 k 的变化关系. 实验结果显示各算法的运行时间随稀疏度的增大而增加, 稀疏度区间在 $[20, 90]$ 时, 各算法的实验平均运行时间能保持在 2 s 以内. 在稀疏度 $k > 90$ 后, SCSSA 算法重构运行时间迅速增加; 当稀疏度 $k > 120$ 后, 该算法的平均重构运行时间不低于 20 s. 本文提出的 NCCS 算法由于参数 σ 值的多次衰减导致算法进行多次迭代, 因此也造成了运行时间相较于剩下的 3 种算法要高, 但对比最新文献 [23] 的 SCSSA 算法, NCCS 算法有更好的实验效果.

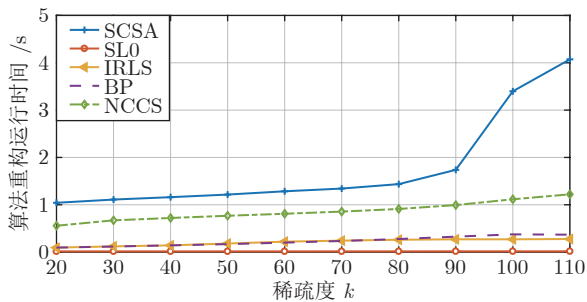


图 7 SL0、IRLS、BP、SCSSA、NCCS 5 种算法的运行时间和稀疏度的变化关系

Fig.7 The relationship between the running time and sparsity of the five algorithms of SL0, IRLS, BP, SCSSA, and NCCS

图 8 是各算法的支撑集恢复成功率 (RSS) 和稀疏度 k 的变化关系. 实验结果显示 SL0、BP 和 IRLS 这 3 种算法的稀疏向量支撑集恢复成功率随稀疏度的增加而下降, SCSSA 算法在稀疏度 $k = 100$ 、 $k = 110$ 时, 其支撑集恢复成功率分别为 $RSS = 0.9980$ 、 $RSS =$

0.9886 , 考虑为实验次数的不足而出现的偏差. 相比之下, 在 NCCS 算法的重构仿真实验中, RSS 值保持稳定, 信号的支撑集在整个实验区间内实现完全恢复.

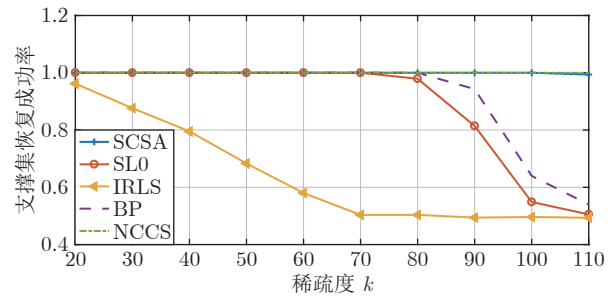


图 8 SL0、IRLS、BP、SCSSA、NCCS 5 种算法的支撑集恢复成功率和稀疏度的变化关系

Fig.8 The relationship between the recovery success rate of the support set and sparsity of the five algorithms of SL0, IRLS, BP, SCSSA, and NCCS

图 9 以及表 1 分别为实验中 5 组算法恢复向量的归一化均方差 ($NMSE$) 和稀疏度 k 的变化关系图像曲线和具体数值记录. 由图 9 和表 1 可见, NCCS 算法和 SCSSA 算法的实验结果远小于 SL0 算法、BP 算法和 IRLS 算法的实验值.

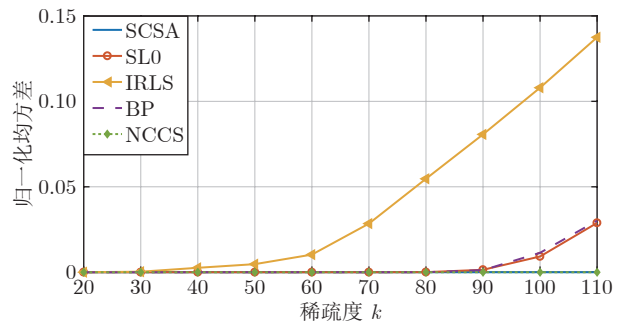


图 9 SL0、IRLS、BP、SCSSA、NCCS 5 种算法的归一化均方差和稀疏度的变化关系

Fig.9 The relationship between the normalized mean square error and sparsity of the five algorithms of SL0, IRLS, BP, SCSSA, and NCCS

综上所述, 本文提出的 NCCS 算法的综合重构性能要比实验的对照算法好, 所以利用 NCCS 算法能更好地恢复稀疏信号向量.

6 结论

本文运用 MM 优化、外点罚函数法、共轭梯度法等方法, 借鉴文献 [23] 提出的 SCSSA 算法思想, 提出了一种新的稀疏信号重构算法, 称为 NCCS 算法. 该算法利用逼近性能更优的非凸复合指数函数实现对 L_0 范数的逼近. 仿真实验表明, 本文所提出

表 1 5 种算法的归一化均方差的数值记录
Table 1 Numerical records of the normalized mean square error of five algorithms

算法	归一化均方差 (NMSE)									
	$k = 20$	$k = 30$	$k = 40$	$k = 50$	$k = 60$	$k = 70$	$k = 80$	$k = 90$	$k = 100$	$k = 110$
SL0	4.11×10^{-10}	2.86×10^{-10}	2.06×10^{-10}	1.55×10^{-10}	1.16×10^{-10}	9.38×10^{-11}	1.46×10^{-4}	1.28×10^{-3}	9.22×10^{-3}	3.13×10^{-2}
IRLS	2.02×10^{-4}	4.08×10^{-4}	1.15×10^{-3}	4.23×10^{-3}	1.07×10^{-2}	2.79×10^{-2}	5.04×10^{-2}	7.20×10^{-2}	1.14×10^{-1}	1.38×10^{-1}
BP	1.65×10^{-22}	4.49×10^{-22}	5.68×10^{-22}	4.95×10^{-22}	4.19×10^{-22}	4.38×10^{-15}	1.18×10^{-5}	9.52×10^{-4}	1.03×10^{-2}	3.57×10^{-2}
SCSA	2.05×10^{-23}	1.94×10^{-22}	1.94×10^{-22}	2.45×10^{-22}	2.50×10^{-22}	2.96×10^{-22}	3.18×10^{-22}	1.27×10^{-16}	1.27×10^{-13}	3.57×10^{-12}
NCCS	5.36×10^{-27}	2.10×10^{-26}	4.81×10^{-24}	4.98×10^{-24}	5.87×10^{-15}	1.85×10^{-17}	3.35×10^{-16}	2.62×10^{-17}	3.58×10^{-16}	1.00×10^{-16}

的 NCCS 算法不仅有效可行, 而且在无噪声情况下, 对比 SL0、IRLS、BP、SCSA 这 4 种算法, NCCS 算法在稀疏信号恢复实验中表现出更优越的重构性能。

References

- Donoho D L. Compressed sensing. *IEEE Transactions on Information Theory*, 2006, **52**(4): 1289–1306
- Gross D, Liu R K, Flammia R T, Becker R, Eisert R. Quantum state tomography via compressed sensing. *Physical Review Letters*, 2010, **105**(15): Article No. 150401
- Wiaux Y, Jacques L, Puy G, Scaife A M M, Vandergheynst P. Compressed sensing imaging techniques for radio interferometry. *Monthly Notices of the Royal Astronomical Society*, 2010, **395**(3): 1733–1742
- Lustig M, Donoho D, Pauly J M. Sparse MRI: The application of compressed sensing for rapid MR imaging. *Magnetic Resonance in Medicine*, 2010, **58**(6): 1182–1195
- Shi B S, Lian Q S, Chen S Z. Compressed sensing magnetic resonance imaging based on dictionary updating and block matching and three dimensional filtering regularization. *Image Processing Letters*, 2016, **10**(1): 68–79
- Herman M, Strohmer T. High-resolution radar via compressed sensing. *IEEE Transactions on Signal Processing*, 2009, **57**(6): 2275–2284
- Chartrand R. Exact reconstruction of sparse signals via nonconvex minimization. *IEEE Signal Processing Letters*, 2007, **14**(10): 707–710
- Peng J G, Yue S G, Li H Y. NP/CMP equivalence: A phenomenon hidden among sparsity models L_0 minimization and L_p minimization for information processing. *IEEE Transactions on Information Theory*, 2015, **61**(7): 4028–4033
- Tropp J A, Gilbert A C. Signal recovery from random measurements via orthogonal matching pursuit. *IEEE Transactions on Information Theory*, 2007, **53**(12): 4655–4666
- Beck A, Teboulle M. A fast iterative shrinkage thresholding algorithm for linear inverse problems. *Siam J Imaging Sciences*, 2009, **2**(1): 183–202
- Dai W, Milenkovic O. Subspace pursuit for compressive sensing signal reconstruction. *IEEE Transactions on Information Theory*, 2009, **55**(5): 2230–2249
- Saadat S A, Safari A, Needell D. Sparse reconstruction of regional gravity signal based on stabilized orthogonal matching pursuit (SOMP). *Pure and Applied Geophysics*, 2016, **173**(6): 2087–2099
- Figueiredo M A T, Nowak R D, Wright S J. Gradient projection for sparse reconstruction: Application to compressed sensing and other inverse problems. *IEEE Journal of Selected Topics in Signal Processing*, 2008, **1**(4): 586–597
- Foucart S, Rauhut H. A mathematical introduction to compressive sensing. *Springer New York*, 2013, **5**: 65–75
- Emmanuel J C, Michael B W, Stephen P B. Enhancing sparsity by reweighted L_1 minimization. *Journal of Fourier Analysis and Applications*, 2008, **14**(5): 877–905
- Chartrand R, Yin W. Iteratively reweighted algorithms for compressive sensing. In: Proceedings of the 2008 IEEE International Conference on Acoustics. Las Vegas, NV, USA: IEEE, 2008. 3869–3872
- Zhang Z, Xu Y, Yang J, Li X, Zhang D. A survey of sparse representation: Algorithms and applications. *IEEE Access*, 2017, **3**: 490–530
- Chen S S, Donoho D L, Saunders M A. Atomic decomposition by basis pursuit. *SIAM Review*, 2001, **43**(1): 129–159
- Chartrand R, Staneva V. Restricted isometry properties and nonconvex compressive sensing. *Inverse Problems*, 2010, **24**(3): 657–682
- Zhao Rui-Zhen, Lin Wan-Juan, Li Hao, Hu Shao-Hai. Reconstruction algorithm for compressive sensing based on smoothed L_0 norm and revised newton method. *Journal of Computer-Aided Design & Computer Graphics*, 2012, **24**(4): 478–484 (赵瑞珍, 林婉娟, 李浩, 胡绍海. 基于光滑 L_0 范数和修正牛顿法的压缩感知重建算法. 计算机辅助设计与图形学学报, 2012, **24**(4): 478–484)
- Mohimani H, Babaie Z M, Jutten C. A fast approach for overcomplete sparse decomposition based on smoothed L_0 norm. *IEEE Transactions on Signal Processing*, 2009, **57**(1): 289–301
- Zhang C J, Hao D B, Hou C B, Yin X J. A new approach for sparse signal recovery in compressed sensing based on minimizing composite trigonometric function. *IEEE Access*, 2018, **6**: 44894–44904
- Malek M M, Babaie M, Koochakzadeh A, Jansson M, Rojas C R. Successive concave sparsity approximation for compressed sensing. *IEEE Transactions on Signal Processing*, 2016, **64**(21): 5657–5671
- Li H Y, Zhang Q, Cui A G, Peng J G. Minimization of fraction function penalty in compressed Sensing. *IEEE Transactions on Neural Networks*, 2020, **31**(5): 1626–1637
- Hunter D R, Lange K. A tutorial on MM algorithms. *The American Statistician*, 2004, **58**(1): 30–37
- Yuan Ya-Xiang, Sun Wen-Yu. *Optimization Theory and Method*. Beijing: Science Press, 1997. 455–482

(袁亚湘, 孙文瑜. 最优化理论与方法. 北京: 科学出版社, 1997. 455-482)

- 27 Emmanuel J, Michael B W, Stephen P B. Enhancing sparsity by reweighted L_1 minimization. *Journal of Fourier Analysis & Applications*, 2007, **14**(5): 877-905
- 28 Clarke F H. *Optimization and Non-smooth Analysis*. New York: Classics in Applied Mathematics, SIAM, 1983. 37-187
- 29 Al-Naffouri T Y, Masood M. Distribution agnostic structured sparsity recovery algorithms. In: Proceedings of the 8th International Workshop on Systems, Signal Processing and Their Applications (WoSSPA). Algiers, Algeria: IEEE, 2013. 283-290
- 30 Chen Jin-Li, Li Wei, Zhu Xiao-Rong, Chen Xuan, Li Jia-Qiang. Smooth L_0 norm calculation based on modified approximate hyperbolic tangent function. *Computer Engineering and Design*, 2018, **39**(12): 3717-3721, 3754
(陈金立, 李伟, 朱筱嵘, 陈宣, 李家强. 基于修正近似双曲正切函数的平滑 L_0 范数计算. 计算机工程与设计, 2018, **39**(12): 3717-3721, 3754)



周洁容 广州大学机器生命与智能研究中心硕士研究生. 广州大学数学与信息科学学院硕士研究生. 2014 年获得岭南师范学院数学学士学位. 主要研究方向为稀疏信息处理.

E-mail: jrzhouzoe@163.com

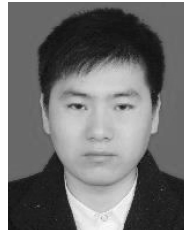
(ZHOU Jie-Rong Master student at the Research Center for Machine Life and Intelligence, Guangzhou University. Master student at the Mathematics and Information Science College, Guangzhou University. She received her bachelor degree in mathematics from the School of Lingnan Normal University in 2014. Her main research interest is sparse information processing.)



李海洋 广州大学机器生命与智能研究中心教授. 广州大学数学与信息科学学院教授. 2008 年获得陕西师范大学基础数学博士学位. 主要研究方向为量子逻辑理论, 机器学习理论和稀疏信息处理.

E-mail: fplihaiyang@126.com; haiyangli@gzhu.edu.cn

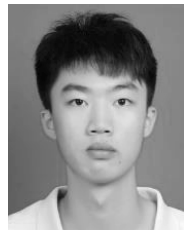
(LI Hai-Yang Professor at the Research Center for Machine Life and Intelligence, Guangzhou University. Professor at the Mathematics and Information Science College, Guangzhou University. He received his Ph.D. degree in fundamental mathematics from Shaanxi Normal University in 2008. His research interest covers quantum logic theory, machine learning theory, and sparse information processing.)



凌军 广州大学机器生命与智能研究中心博士研究生. 广州大学数学与信息科学学院博士研究生. 2019 年获得南昌大学概率论与数理统计硕士学位. 主要研究方向为小目标运动检测, 算子半群理论和人工智能.

E-mail: gdalingjun@163.com

(LING Jun Ph.D. candidate at the Research Center for Machine Life and Intelligence, Guangzhou University. Ph.D. candidate at the Mathematics and Information Science College, Guangzhou University. He received his master degree in probability theory and mathematical statistics from Nanchang University in 2019. His research interest covers small target detection, operator semigroup theory, and artificial intelligence.)



陈浩 广州大学机器生命与智能研究中心硕士研究生. 广州大学数学与信息科学学院硕士研究生. 2015 年获得广州大学数学学士学位. 主要研究方向为小目标运动检测, 人工智能和计算神经学.

E-mail: gdchenhao@126.com

(CHEN Hao Master student at the Research Center for Machine Life and Intelligence, Guangzhou University. Master student at the Mathematics and Information Science College, Guangzhou University. He received his bachelor degree in mathematics from Guangzhou University in 2015. His research interest covers small target motion detection, artificial intelligence, and computational neuroscience.)



彭济根 广州大学机器生命与智能研究中心教授. 广州大学数学与信息科学学院教授. 1998 年获得西安交通大学计算数学博士学位. 主要研究方向为非线性泛函分析, 机器学习理论, 稀疏优化和碰撞检测的数学理论与方法. 本文通信作者.

E-mail: jgpeng@gzhu.edu.cn

(PENG Ji-Gen Professor at the Research Center for Machine Life and Intelligence, Guangzhou University. Professor at the Mathematics and Information Science College, Guangzhou University. He received his Ph.D. degree in applied mathematics from Xi'an Jiaotong University in 1998. His research interest covers nonlinear functional analysis, machine learning theory, sparse optimization, and mathematical theory and methods for collision detection. Corresponding author of this paper.)