Advances in Applied Mathematics 应用数学进展, 2019, 8(3), 512-523
Published Online March 2019 in Hans. http://www.hanspub.org/journal/aam
https://doi.org/10.12677/aam.2019.83057

Hans汉斯

Conjugate Gradient Method for $l_p + l_2$ Norm Problems

Jiaming Zhan, Caiying Wu*

School of Mathematical Sciences, Inner Mongolia University, Hohhot Inner Mongolia
Email: 389806392@qq.com, *wucaiyingyun@163.com

Received: Feb. 28th, 2019; accepted: Mar. 13th, 2019; published: Mar. 20th, 2019

Abstract

A new model is proposed for sparse signal reconstruction. We smooth our model by smoothing absolute value function. Then we use a new tri-term conjugate gradient method to restore sparse signal. The effect of different parameters on sparse signal recovery is tested. The numerical results show that our algorithm is efficient.

Keywords

Sparse Signal Recovery, Conjugate Gradient Method, $l_p + l_2$ Norm, Global Convergence

求解基于$l_p + l_2$范数问题的共轭梯度法

詹佳明，乌彩英*

内蒙古大学数学科学学院，内蒙古 呼和浩特
Email: 389806392@qq.com, *wucaiyingyun@163.com

收稿日期：2019年2月28日；录用日期：2019年3月13日；发布日期：2019年3月20日

摘　要

本文针对稀疏信号的重构问题提出新的函数模型，通过光滑化绝对值函数光滑我们的模型，基于三项共轭梯度法对稀疏信号进行恢复，并试验了不同参数值对稀疏信号恢复效果的影响，数值实验表明本文算法的数值有效性。

*通讯作者。

文章引用：詹佳明，乌彩英．求解基于$l_p + l_2$范数问题的共轭梯度法[J]．应用数学进展, 2019, 8(3): 512-523.
DOI: 10.12677/aam.2019.83057

关键词

稀疏信号恢复，三项共轭梯度法，$l_p + l_2$ 范数，全局收敛性

1. 引言

压缩感知技术在信号重构问题中有着广泛的应用，而稀疏性是压缩感知中的重要问题。2006 年，E. Candes，J. Romberg 和 T. Tao 在文献[1]中提出，若矩阵 A 满足 RIP 条件，则可以通过下述模型(1)精确恢复信号：

$$\min_{x\in R^n} \|x\|_0, \text{s.t.} Ax = b. \tag{1}$$

其中 $A \in R^{m\times n}\ (m < n)$ 为观测矩阵，$b \in R^m$ 为原始信号，$\|x\|_0$ 表示 0 范数，即稀疏信号 x 中非零元素的个数。但该问题是 NP-hard 问题[2]，因此学者们考虑了如下的凸优化问题：

$$\min_{x\in R^n} \|x\|_1, \text{s.t.} Ax = b. \tag{2}$$

(2)称作 1 范数模型，在适当的假设下，由文献[3]的定理 1.3 可知，模型(2)可以较精确地恢复原始信号。

由于问题(2)的凸性，有许多有效算法可求解之，如基追踪法[4]，迭代阈值方法[5]。近年来，p ($0 < p < 1$)拟范数模型受到学者们的青睐，因 p 范数较 1 范数更能得到稀疏解，尤其是徐宗本提出 $p = 0.5$ 时具有较好的计算效果[6]。p 范数模型为：

$$\min_{x\in R^n} \|x\|_p, \text{s.t.} Ax = b. \tag{3}$$

由于 $0 < p < 1$，故问题(3)为非凸问题，且与 0 范数同样是 NP-hard 问题。文献[1]中作者利用光滑逼近 $|\cdot|$ 函数进而光滑逼近 p 范数。受到该文的启发，本文利用文献[7] [8]中的光滑逼近 $|\cdot|$ 函数技术光滑逼近 p 范数，且结合 p 范数与 2 范数各自的优点，利用 p 范数与 2 范数加权的方法弥补由于 p 过小而引起的数值不稳定，从而应用共轭梯度法进行信号恢复，并在适当的假设下，证明了算法的全局收敛性，同时进行了数值测试，测试结果证明我们的方法是有效的。

2. 模型及其性质

我们考虑如下模型：

$$\min_{x\in R^n} \lambda_1 \|x\|_p + \lambda_2 \|x\|_2, \text{s.t.} Ax = b.$$

其中 $\lambda_1, \lambda_2 \in R$。该问题的正则化问题为：

$$\min_{x\in R^n} \lambda_1 \|x\|_p + \lambda_2 \|x\|_2 + \frac{1}{2}\|Ax - b\|_2^2. \tag{4}$$

其模型较 p 范数模型(3)，加入了一项 2 范数项进行调节 p 过小时目标函数的非凸程度，以 $p = 0.3$ 时为例，依次做 $R = \lambda_2/\lambda_1 = 0,1,10,100$ 时的范数图像，如图 1 至图 4 所示：

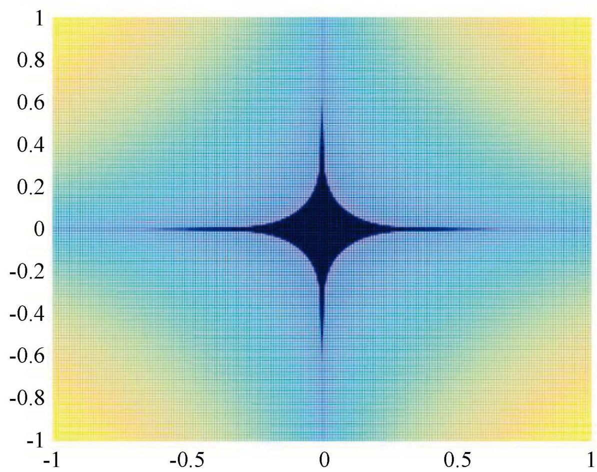

Figure 1. $R = 0$ norm image
图 1. $R = 0$ 范数图像

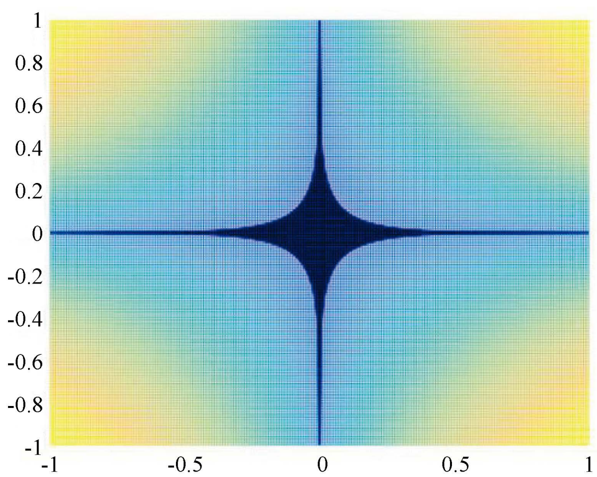

Figure 2. $R = 1$ norm image
图 2. $R = 1$ 范数图像

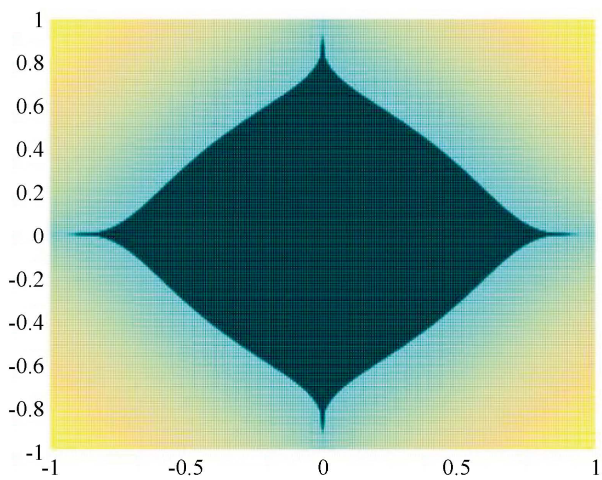

Figure 3. $R = 10$ norm image
图 3. $R = 10$ 范数图像

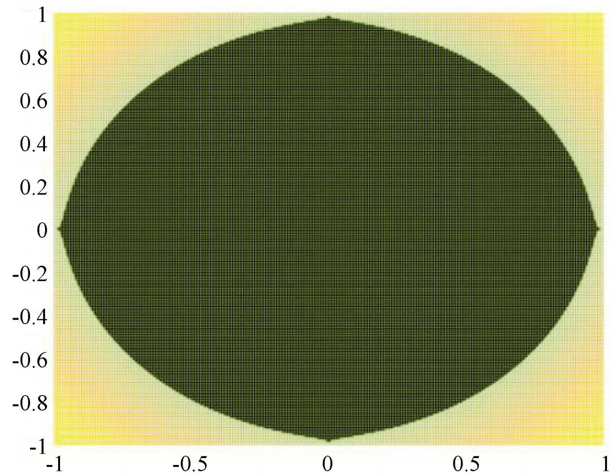

Figure 4. $R = 100$ norm image
图 4. $R = 100$ 范数图像

可见通过适当调整 λ_1, λ_2 取值在保证产生稀疏解的同时可以增强问题的凸性。由 p 范数的定义 $\|x\|_p = \left(\sum_{i=1}^{n} |x_i|^p \right)^{\frac{1}{p}}$ 可知，由于 $|x|$ 的非光滑性，导致 $\|x\|_p$ 是非光滑的，因而我们通过光滑绝对值函数 $|x|$ 对 $\|x\|_p$ 进行光滑逼近。文献[8]中对绝对值函数提出如下两个光滑函数：

$$\varphi_1(\varepsilon, t) = \begin{cases} t, t \ge \dfrac{\varepsilon}{2}, \\ \dfrac{t^2}{\varepsilon} + \dfrac{\varepsilon}{4}, -\dfrac{\varepsilon}{2} < t < \dfrac{\varepsilon}{2} \\ -t, t \le -\dfrac{\varepsilon}{2}. \end{cases} \text{及 } \varphi_2(\varepsilon, t) = \begin{cases} \dfrac{t^2}{2\varepsilon}, |t| \le \varepsilon, \\ |t| - \dfrac{\varepsilon}{2}, |t| > \varepsilon. \end{cases} \tag{5}$$

利用 $\varphi_i(i = 1, 2)$，我们得到的模型为：

$$\min \lambda_1 \phi_i(\varepsilon, x) + \lambda_2 \|x\|_{2,\mu} + \frac{1}{2}\|Ax - b\|_2^2. \tag{6}$$

其中，

$$\phi_i(x) = \left[\sum_{j=1}^{n} \varphi_i(\varepsilon, x_j)^p \right]^{\frac{1}{p}}, i = 1, 2, \tag{7}$$

$$\|x\|_{2,\mu} = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2 + \mu^2}, \mu > 0.$$

为叙述方便，令：

$$G_i(x) = \lambda_1 \phi_i(\varepsilon, x) + \lambda_2 \|x\|_{2,\mu} + \frac{1}{2}\|Ax - b\|_2^2, (i = 1, 2). \tag{8}$$

计算函数 $G_i(x)$ 的梯度：

$$\nabla G_i(x) = \lambda_1 \nabla \phi_i(\varepsilon, x) + \lambda_2 \nabla \|x\|_{2,\mu} + A^{\mathrm{T}}(Ax - b). \tag{9}$$

引理 1 设 D 为 R^n 中的紧集，若 $h: D \to R$ 是 D 上的光滑函数，则 h 在 D 上满足 Lipschitz 条件，即：

$$\|h(x)-h(y)\| \le L\|x-y\|, \forall x,y \in D.$$

其中$L>0$称为Lipschitz常数。

证明：对$\forall x,y \in D$，由$h(x)$的光滑性及Lagrange中值定理可知：

$\left|h(x)-h(y)\right| = \left|\nabla h^{\mathrm{T}}(\xi)(x-y)\right| \le \|\nabla h(\xi)\| \cdot \|x-y\|$，其中$\xi$位于$x,y$之间。

因D为紧集，则$\exists L>0$，$\|\nabla h(\xi)\| \le L$，从而$\|h(x)-h(y)\| \le L\|x-y\|$。

引理2 设$D \subset R^n$为紧集，$f,g: D \to R$在D上Lipschitz连续，则$f \cdot g$在D上Lipschitz连续。

证明：因D紧且f和g均在D上Lipschitz连续，则$\exists M>0, \forall L_1>0, L_2>0$，满足：

$$|f(x)| \le M, |g(x)| \le M, |f(x)-f(y)| \le L_1\|x-y\|, |g(x)-g(y)| \le L_2\|x-y\|, \forall x,y \in D.$$

对$\forall x,y \in D$，有：

$$\begin{aligned}
&|f(x)g(x)-f(y)g(y)| \\
&=|f(x)g(x)-f(y)g(x)+f(y)g(x)-f(y)g(y)| \\
&\le |g(x)||f(x)-f(y)|+|f(y)||g(x)-g(y)| \\
&\le ML_1\|x-y\|+ML_2\|x-y\| \\
&=M(L_1+L_2)\|x-y\|.
\end{aligned}$$

引理3 若$\varphi_i(i=1,2)$由(5)定义，则$\varphi_i'(\varepsilon,t)$在紧集$D$上Lipschitz连续。

证明：对于$i=1$的情况，我们有：

$$\varphi_i'(\varepsilon,t)=\begin{cases}\operatorname{sgn}(t), |t| \ge \dfrac{\varepsilon}{2}, \\ \dfrac{2t}{\varepsilon}, |t| < \dfrac{\varepsilon}{2}.\end{cases}$$

当$\dfrac{\varepsilon}{2}>t_1>t_2>-\dfrac{\varepsilon}{2}$时，有：

$$|\varphi_1'(\varepsilon,t_1)-\varphi_1'(\varepsilon,t_2)| \le \left|\frac{2t_1}{\varepsilon}-\frac{2t_2}{\varepsilon}\right| = \frac{2}{\varepsilon}|t_1-t_2|.$$

当$t_1>\dfrac{\varepsilon}{2}>t_2>-\dfrac{\varepsilon}{2}$时，有：

$$|\varphi_1'(\varepsilon,t_1)-\varphi_1'(\varepsilon,t_2)| = \left|1-\frac{2t_2}{\varepsilon}\right| < \left|\frac{2t_1}{\varepsilon}-\frac{2t_2}{\varepsilon}\right| = \frac{2}{\varepsilon}|t_1-t_2|.$$

当$\dfrac{\varepsilon}{2}>t_1>-\dfrac{\varepsilon}{2}>t_2$时，有：

$$|\varphi_1'(\varepsilon,t_1)-\varphi_1'(\varepsilon,t_2)| = \left|\frac{2t_1}{\varepsilon}+1\right| < \left|\frac{2t_1}{\varepsilon}-\frac{2t_2}{\varepsilon}\right| = \frac{2}{\varepsilon}|t_1-t_2|.$$

当$t_1>\dfrac{\varepsilon}{2}>-\dfrac{\varepsilon}{2}>t_2$时，有：

$$|\varphi_1'(\varepsilon,t_1)-\varphi_1'(\varepsilon,t_2)| = 2 = \frac{2}{\varepsilon}\varepsilon \le \frac{2}{\varepsilon}|t_1-t_2|.$$

当$t_1 > t_2 > \frac{\varepsilon}{2}$或$-\frac{\varepsilon}{2} > t_1 > t_2$时，有：

$$\left|\varphi_1'(\varepsilon,t_1)-\varphi_1'(\varepsilon,t_2)\right| = 0 \le \frac{2}{\varepsilon}\left|t_1-t_2\right|.$$

令$L_1=\frac{2}{\varepsilon}$，同理可证存在常数$L_2$使得$\varphi_2'(t)$在紧集$D$上Lipschitz连续。

最终令$L_m=\max\{L_1,L_2\}$，我们得到：

$$\left|\varphi_i'(\varepsilon,t_1)-\varphi_i'(\varepsilon,t_2)\right| \le L_m\left|t_1-t_2\right|, \forall t_1,t_2\in D, i=1,2.$$

引理4 函数$\phi_i(\varepsilon,x)$由(7)所定义，则$\nabla\phi_i(\varepsilon,x)$在紧集$D$上Lipschitz连续。

证明：计算得：

$$\left[\nabla\phi_i(\varepsilon,x)\right]_j=\left(\sum_{j=1}^{n}\varphi_i^p(\varepsilon,x_j)\right)^{\frac{1-p}{p}}\varphi_i^{p-1}(\varepsilon,x_j)\varphi'(\varepsilon,x_j).$$

因φ_i连续可微，且$\varepsilon\ne 0,\varphi_i\ne 0$，我们有$\left(\sum_{j=1}^{n}\varphi_i^p(\varepsilon,x_j)\right)^{\frac{1-p}{p}}$和$\varphi_i^{p-1}$均连续可微。假设集合$D$有界，则$D$为紧集，由引理1知$\left(\sum_{j=1}^{n}\varphi_i^p(\varepsilon,x_j)\right)^{\frac{1-p}{p}}$及$\varphi_i^{p-1}$在紧集$D$上Lipschitz连续。

同时由引理3可知$\varphi'(\varepsilon,x_j)$在紧集$D$上Lipschitz连续，则由引理2可知：

$\left(\sum_{j=1}^{n}\varphi_i^p(\varepsilon,x_j)\right)^{\frac{1-p}{p}}\cdot\varphi_i^{p-1}(\varepsilon,x_j)\cdot\varphi'(\varepsilon,x_j)$在紧集$D$上Lipschitz连续，记Lipschitz常数为L_D，那么有：

$$\left\|\nabla\phi_i(\varepsilon,x)-\nabla\phi_i(\varepsilon,y)\right\| \le \sum_{j=1}^{n}\left\|\left[\nabla\phi_i(\varepsilon,x)\right]_j-\left[\nabla\phi_i(\varepsilon,y)\right]_j\right\| \le nL_D\left\|x-y\right\|.$$

即$\nabla\phi_i(\varepsilon,x)$在紧集$D$上Lipschitz连续。

引理5 $\nabla\|x\|_2$在紧集D上Lipschitz连续。

证明：计算得$\nabla\|x\|_2=\left[\frac{x_1}{\|x\|_2},\frac{x_2}{\|x\|_2},\cdots,\frac{x_n}{\|x\|_2}\right]$，由引理1可知$\frac{x_j}{\|x\|}$在水平集上光滑，则$\frac{x_j}{\|x\|}$在紧集$D$上Lipschitz连续，记Lipschitz常数为L_c，则：

$$\left\|\nabla\|x\|_2-\nabla\|y\|_2\right\| \le nL_c\left\|x-y\right\|.$$

引理6 目标函数梯度$\nabla G_i(x)$由(9)所定义，则$\nabla G_i(x)$在紧集D上Lipschitz连续。

证明：计算得$\nabla G_i(x)=\lambda_1\nabla\varphi_i(\varepsilon,x)+\lambda_2\nabla\|x\|_{2,\mu}+A^{\mathrm{T}}(Ax-b)$，则：

$$\begin{aligned}
&\left\|\nabla G_i(x)-\nabla G_i(y)\right\| \\
&=\left\|\lambda_1\nabla\phi_i(\varepsilon,x)+\lambda_2\nabla\|x\|_{2,\mu}+A^{\mathrm{T}}(Ax-b)-\lambda_1\nabla\phi_i(\varepsilon,y)-\lambda_2\nabla\|y\|_{2,\mu}-A^{\mathrm{T}}(Ay-b)\right\| \\
&\le \lambda_1 nL_D\left\|x-y\right\|+\lambda_2 nL_C\left\|x-y\right\|+\|A\|_2^2\left\|x-y\right\| \\
&=\left(\lambda_1 nL_D+\lambda_2 nL_C+\|A\|_2^2\right)\left\|x-y\right\|.
\end{aligned}$$

3. 基于非精确线搜索的三项共轭梯度法

求问题 $\min\limits_{x\in R^n} f(x)$ 的共轭梯度法的具体迭代过程如下：

$$x_{k+1} = x_k + \alpha_k d_k,$$
$$d_k = \begin{cases} -g_k, k = 0, \\ -g_k + \beta_{k-1} d_{k-1}, k > 0. \end{cases}$$

其中 x_k 代表当前迭代点，d_k 代表当前搜索方向，α_k 代表搜索步长，$g_k = \nabla G(x_k)$，不同的 β_k 代表不同的共轭梯度算法。Shouqiang Du 和 Miao Chen 在文献[9]中提出了一种新型三项共轭梯度法。具体迭代过程如下：

$$d_k = \begin{cases} -g_k, k = 0, \\ -g_k + \beta_{k-1} d_{k-1} - \theta_{k-1} y_{k-1}, k > 0. \end{cases}$$

其中新加入项 $-\theta_{k-1} y_{k-1}$ 可以每步迭代确保目标函数值严格下降，使计算效率更高.本文将利用该共轭梯度法解决信号恢复问题。

算法 1 (基于非精确线搜索的共轭梯度法)

步 0 给出初始参数 $\mu,\varepsilon,\eta,\delta,\lambda_1,\lambda_2,\rho,\sigma,x_0,\mu_0,\gamma$，置 $k := 0$。

步 1 若满足终止准则 $\left\|\nabla G_i(x)\right\| < \varepsilon$，停止计算，输出结果 $x_k \approx x^*$；否则，转步 2。

步 2 计算搜索方向：

$$d_k = \begin{cases} -g_k, k = 0, \\ -g_k + \beta_{k-1} d_{k-1} - \theta_{k-1} y_{k-1}, k > 0. \end{cases}$$

其中，

$$y_k = g_k - g_{k-1}, \beta_k = \frac{g_k^{\mathrm{T}} y_{k-1}}{-\eta g_{k-1}^{\mathrm{T}} d_{k-1} + \mu\left|g_k^{\mathrm{T}} y_{k-1}\right|}, \theta_k = \frac{g_k^{\mathrm{T}} d_{k-1}}{-\eta g_{k-1}^{\mathrm{T}} d_{k-1} + \mu\left|g_k^{\mathrm{T}} y_{k-1}\right|}.$$

步 3 令 $\alpha_k = \max\left\{\rho^m : m = 0,1,2,\cdots\right\}$，同时满足：

$$G(x_k + \alpha_k d_k) > G(x_k) - 2\sigma(1-\gamma\mu_0)\alpha_k G(x_k).$$

步 4 令 $x_{k+1} = x_k + \alpha_k d_k$，计算 $g_{k+1} = \nabla G(x_{k+1})$。

步 5 令 $k := k+1$，转步 1。

假设 1 令 $c > 0$ 为常数，假设水平集 $L_c = \left\{x_k \mid G(x_k) \le c\right\}$ 是紧集。

定理 1 若 $\{x_k\}$ 是由算法 1 产生的序列，则：

$$\lim_{k\to 0}\left\|\nabla G(x_k)\right\| = 0.$$

证明：由文献[9]中定理 6，本文引理 6 及假设 1 可得。

4. 数值实验

本节实验在 Intel Core i7-6500U 2.50GHz CPU，8G RAM，Windows10 64 位操作系统，MATLAB R2016a 环境下进行，所有数值实验结果为测试十次取平均值。

我们对两个不同的绝对值近似函数 φ_1 和 φ_2 分别在信号维数 n 取 2048，4096，8192 三种情况下，测试随机信号在无噪声和存在噪声干扰时的恢复效果.通过对比信号恢复时间 t，相对误差 e 和迭代步数 k

来比较范数 p 和 R 对算法对信号恢复效果的影响。参数取值如下：A 为 $m\times n$ 的随机高斯矩阵，x 为 n 维随机初始信号，$x0=zeros(n,1)$ 为初始点，$gs=normnd(0,0.01,n/2,1)$ 是一个平均值是 0，标准差是 0.01 的对观测向量 b 的噪声干扰，p 取 0.3，0.4，0.5，0.6，R 取 0，1，10，100 (当 $R=0$ 时代表 2 范数项为 0，即目标函数为 p 范数模型)，其他参数取值如下：

$$\mu=1e-5,\varepsilon=1e-5,\eta=1,\delta=2,\lambda=1e-6*\left\|A^{\mathrm{T}}b\right\|_{\infty},\rho=0.5,\sigma=0.002,\gamma=0.2.$$

1) 我们先对绝对值近似函数 φ_1 的信号恢复问题进行数值实验，结果如表 1 至表 6 所示：

Table 1. Signal dimension n = 2048, no noise interference
表 1. 信号维数 n = 2048，无噪声干扰

R	**0**			**1**			**10**			**100**		
p	***e***	***k***	***t***	***e***	***k***	***t***	***e***	***k***	***t***	***e***	***k***	***t***
0.6	3.0e−7	278	4.84	2.9e−7	256	5.09	2.2e−7	266	5.97	5.5e−7	268	5.32
0.5	5.0e−6	223	3.91	4.2e−6	248	5.10	2.2e−6	243	5.07	8.4e−6	227	4.87
0.4	3.8e−4	376	8.20	3.5e−4	193	4.73	2.5e−4	282	8.62	1.9e−4	321	11.58
0.3	1.0e−2	152	3.89	9.3e−3	142	4.17	2.2e−3	146	4.59	1.3e−2	145	4.94

Table 2. Signal dimension n = 2048, Gauss noise interference
表 2. 信号维数 n = 2048，高斯噪声干扰

R	**0**			**1**			**10**			**100**		
p	***e***	***k***	***t***	***e***	***k***	***t***	***e***	***k***	***t***	***e***	***k***	***t***
0.6	1.8e−3	664	11.6	1.6e−3	564	11.1	1.7e−3	532	10.5	1.8e−3	690	14.9
0.5	7.1e−2	269	12.7	1.4e−3	298	12.8	8.9e−4	274	13.5	1.5e−3	231	2.38
0.4	8.6e−2	167	3.52	6.2e−3	164	4.20	9.2e−4	187	5.96	1.7e−3	125	3.10
0.3	1.2e−2	152	4.43	1.3e−2	144	4.35	2.5e−2	137	3.76	8.1e−3	155	4.54

Table 3. Signal dimension n = 4096, no noise interference
表 3. 信号维数 n = 4096，无噪声干扰

R	**0**			**1**			**10**			**100**		
p	***e***	***k***	***t***	***e***	***k***	***t***	***e***	***k***	***t***	***e***	***k***	***t***
0.6	5.6e−7	270	17.8	8.2e−7	219	17.5	8.0e−7	253	20.0	2.4e−6	205	15.5
0.5	1.6e−5	207	15.3	1.5e−5	226	19.4	8.5e−6	239	19.2	1.2e−5	243	19.8
0.4	7.5e−3	230	21.3	1.1e−3	224	23.2	6.0e−4	213	21.1	6.1e−4	239	24.9
0.3	1.3e−3	149	14.6	2.3e−2	162	18.4	1.5e−2	156	18.1	1.8e−2	150	17.4

Table 4. Signal dimension n = 4096, Gauss noise interference
表 4. 信号维数 n = 4096，高斯噪声干扰

R	**0**			**1**			**10**			**100**		
p	***e***	***k***	***t***	***e***	***k***	***t***	***e***	***k***	***t***	***e***	***k***	***t***
0.6	1.2e−3	613	44.4	1.4e−3	926	75.5	1.3e−3	572	47.7	1.2e−3	714	57.68
0.5	3.2e−4	131	8.98	3.3e−4	128	10.2	3.3e−4	156	12.6	3.1e−4	134	10.94
0.4	1.3e−3	169	14.9	1.2e−3	184	18.3	1.0e−3	162	15.7	1.3e−3	188	18.30
0.3	1.8e−2	156	15.7	1.4e−2	163	18.5	2.2e−2	163	18.7	1.5e−3	163	19.74

Table 5. Signal dimension $n = 8192$, no noise interference
表 5. 信号维数 $n = 8192$，无噪声干扰

R		**0**			**1**			**10**			**100**	
p	***e***	***k***	***t***	***e***	***k***	***t***	***e***	***k***	***t***	***e***	***k***	***t***
0.6	1.7e−6	219	62.42	1.0e−6	212	66.59	6.2e−7	219	66.04	3.5e−6	215	67.41
0.5	3.2e−5	273	89.12	1.3e−5	223	78.24	2.0e−5	222	78.90	2.7e−5	230	82.25
0.4	7.7e−3	107	38.12	6.6e−3	117	43.98	7.1e−3	111	43.61	6.8e−3	116	44.30
0.3	3.1e−2	164	67.77	2.9e−2	172	74.96	2.4e−2	163	71.80	1.9e−2	167	72.89

Table 6. Signal dimension $n = 8192$, Gauss noise interference
表 6. 信号维数 $n = 8192$，高斯噪声干扰

R		**0**			**1**			**10**			**100**	
p	***e***	***k***	***t***	***e***	***k***	***t***	***e***	***k***	***t***	***e***	***k***	***t***
0.6	8.5e−4	622	176.7	8.2e−4	1011	377.4	8.6e−4	998	383.3	8.1e−4	514	175.4
0.5	2.5e−4	134	41.26	1.8e−4	154	54.85	1.8e−4	190	67.90	1.7e−4	157	55.52
0.4	7.2e−3	113	40.62	1.8e−3	219	97.92	1.7e−3	271	126.7	2.9e−4	158	67.28
0.3	8.0e−2	170	73.10	2.4e−2	162	78.82	2.8e−2	168	82.36	2.6e−2	171	82.74

2) 我们再对绝对值近似函数 φ_2 的信号恢复问题进行数值实验，结果如表 7 至表 12 所示：

Table 7. Signal dimension $n = 2048$, no noise interference
表 7. 信号维数 $n = 2048$，无噪声干扰

R		**0**			**1**			**10**			**100**	
p	***e***	***k***	***t***	***e***	***k***	***t***	***e***	***k***	***t***	***e***	***k***	***t***
0.6	4.4e−6	348	5.89	3.9e−6	347	6.14	1.5e−5	328	6.08	5.0e−7	342	5.71
0.5	1.4e−6	274	4.55	1.2e−6	260	4.29	2.1e−6	262	4.22	5.7e−6	228	3.54
0.4	1.4e−3	335	5.82	4.4e−5	248	5.88	6.7e−5	326	7.76	6.6e−5	209	4.83
0.3	8.3e−3	137	3.51	1.7e−3	154	4.04	1.3e−2	168	4.52	1.0e−2	137	3.42

Table 8. Signal dimension $n = 2048$, Gauss noise interference
表 8. 信号维数 $n = 2048$，高斯噪声干扰

R		**0**			**1**			**10**			**100**	
p	***e***	***k***	***t***	***e***	***k***	***t***	***e***	***k***	***t***	***e***	***k***	***t***
0.6	2.8e−3	419	8.53	1.8e−3	489	8.98	1.6e−3	412	6.99	1.7e−3	432	7.45
0.5	7.0e−4	171	2.80	6.9e−4	160	2.64	8.2e−4	160	2.50	8.2e−4	167	2.61
0.4	3.3e−3	168	3.95	5.2e−4	149	3.26	5.2e−4	184	3.97	6.6e−4	136	2.78
0.3	6.2e−2	157	4.41	7.7e−3	148	3.87	7.1e−3	150	3.84	1.2e−2	141	3.59

Table 9. Signal dimension $n = 4096$, no noise interference
表 9. 信号维数 $n = 4096$，外界无噪声干扰

R		**0**			**1**			**10**			**100**	
p	***e***	***k***	***t***	***e***	***k***	***t***	***e***	***k***	***t***	***e***	***k***	***t***
0.6	9.3e−6	195	14.6	8.2e−6	222	16.9	4.0e−5	272	20.8	8.2e−6	287	20.6
0.5	2.7e−5	254	20.0	5.0e−6	236	19.0	2.2e−6	226	18.6	4.9e−6	207	15.6
0.4	1.5e−3	314	31.7	3.9e−4	199	19.1	2.9e−4	240	23.5	1.6e−4	258	25.5
0.3	1.5e−2	140	15.3	1.2e−2	146	15.4	1.1e−2	156	17.1	1.6e−2	134	13.6

Table 10. Signal dimension $n = 4096$, Gauss noise interference
表 10. 信号维数 $n = 4096$，高斯噪声干扰

R	0			1			10			100		
p	e	k	t	e	k	t	e	k	t	e	k	t
0.6	1.1e−2	397	28.6	1.3e−3	376	25.8	1.3e−3	322	22.9	1.1e−3	361	24.6
0.5	1.2e−3	134	10.2	3.9e−4	145	10.2	3.6e−4	145	10.1	3.3e−4	144	9.93
0.4	3.7e−3	251	23.9	6.0e−4	211	19.4	5.6e−4	179	15.1	4.6e−4	208	18.1
0.3	1.4e−2	153	16.4	1.1e−2	153	17.5	1.0e−2	179	20.1	9.6e−2	150	15.7

Table 11. Signal dimension $n = 8192$, no noise interference
表 11. 信号维数 $n = 8192$，无噪声干扰

R	0			1			10			100		
p	e	k	t	e	k	t	e	k	t	e	k	t
0.6	5.2e−7	228	66.9	3.5e−7	247	71.7	5.4e−5	225	66.7	8.0e−7	246	65.7
0.5	1.0e−5	217	69.7	1.6e−6	222	53.3	9.7e−6	229	73.9	8.4e−6	185	55.8
0.4	2.7e−4	213	84.2	1.8e−4	266	77.5	6.1e−4	187	70.7	6.9e−4	241	93.4
0.3	5.3e−2	153	74.7	2.5e−2	167	68.9	2.5e−2	145	60.7	2.3e−2	169	72.6

Table 12. Signal dimension $n = 8192$, Gauss noise interference
表 12. 信号维数 $n = 8192$，高斯噪声干扰

R	0			1			10			100		
p	e	k	t	e	k	t	e	k	t	e	k	t
0.6	7.6e−3	241	76.1	7.7e−4	280	78.8	7.6e−4	291	82.8	8.3e−4	272	75.2
0.5	2.0e−4	135	39.0	1.6e−4	147	41.3	1.9e−4	134	40.3	2.0e−4	133	40.2
0.4	1.1e−3	174	60.7	3.6e−4	154	61.8	6.4e−4	267	87.3	6.1e−4	207	78.3
0.3	4.3e−2	167	66.1	2.8e−2	151	58.2	2.7e−2	157	64.3	2.6e−2	149	59.9

5. 主要结论

通过在不同维度 $n = 2048$，4096，8192 下做信号恢复效果分析，我们主要得出以下结论：

1) 相对于 p 范数模型，$l_p + l_2$ 范数模型具有更好的恢复效果，二者在时间和迭代步数相近的情况下，$l_p + l_2$ 范数模型对初始信号的精度更高。

2) 做 λ_2/λ_1 的值对信号恢复效果影响的横向对比，λ_2/λ_1 的取值对信号恢复的时间和迭代步数影响并不明显；但是 $\lambda_2/\lambda_1 = 1$ 或 $\lambda_2/\lambda_1 = 10$ 时信号的恢复具有更高精度，这是由于当 λ_2/λ_1 取值为 1~10 时 2 范数项对 p 范数的调节使得函数模型的非凸性程度降低，因而提高信号恢复的精度。

3) 做 p 的值对信号恢复效果影响的纵向对比，显然范数 $p = 0.6$ 时信号恢复的精度最高，$p = 0.3$ 时精度最低。

注：1) 文献[10]中提出的模型为

$$\min_{x \in R^n} \lambda \|x\|_1 - \tau \|x\|_2,$$
$$\text{s.t.} Ax = b.$$

即利用 $l_1 - l_2$ 范数加权从而产生稀疏阶，而本文考虑 p 范数比 1 范数更易得到稀疏解，从而利用 $l_p + l_2$ 范数加权进行信号恢复。

2) 文献[10]中的光滑化绝对值函数是文献[8]中 φ_3。

3) 文献[11]中的模型为：

$$\min_{x\in R^n} \lambda_1 \|x\|_p^p + \lambda_2 \|x\|_2^2 + \frac{1}{2}\|Ax-b\|_2^2 .$$

而 $\lambda_1 \|x\|_p^2 + \lambda_2 \|x\|_2^2$ 较 $\lambda_1 \|x\|_p + \lambda_2 \|x\|_2$ 不易产生稀疏解。取 $p = 0.3$，$R = 100$ 对比图形如图 5 和图 6 所示：

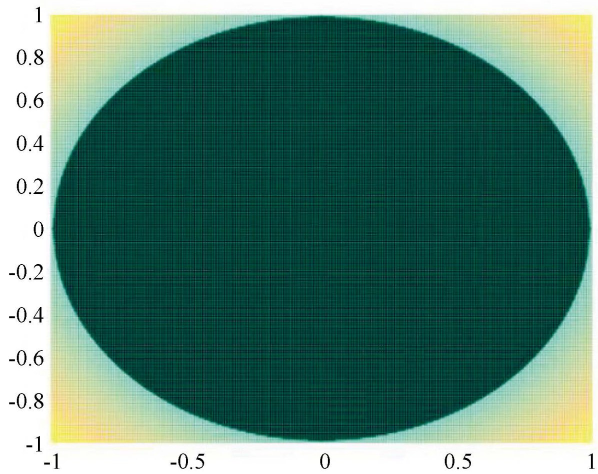

Figure 5. $\lambda_1 \|x\|_p^p + \lambda_2 \|x\|_2^2$ norm image

图 5. $\lambda_1 \|x\|_p^p + \lambda_2 \|x\|_2^2$ 范数图像

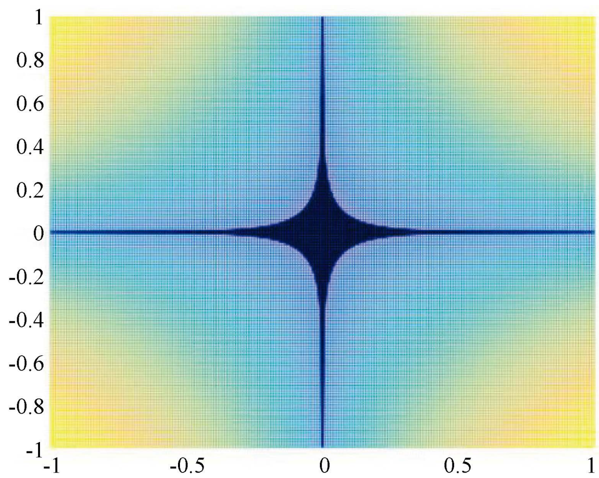

Figure 6. $\lambda_1 \|x\|_p + \lambda_2 \|x\|_2$ norm image

图 6. $\lambda_1 \|x\|_p + \lambda_2 \|x\|_2$ 范数图像

基金项目

内蒙古自然科学基金(2018MS01016)。

参考文献

[1] Candes, E.J., Romberg, J. and Tao, T. (2006) Robust Uncertainty Principles: Exact Signal Reconstruction from Highly

Incomplete Frequency Information. *IEEE Transactions on Information Theory*, **52**, 489-509. https://doi.org/10.1109/TIT.2005.862083

[2] Ge, D., Jiang, X. and Ye, Y. (2011) A Note on the Complexity of L_p Minimization. *Mathematical Programming*, **129**, 285-299. https://doi.org/10.1007/s10107-011-0470-2

[3] Candes, E.J. and Tao, T. (2004) Near Optimal Signal Recovery from Random Projections: Universal Encoding Strategies. *IEEE Transactions on Information Theory*, **52**, 5406-5425. https://doi.org/10.1109/TIT.2006.885507

[4] Chen, S.S., Donoho, D.L. and Saunders, M.A. (1998) Atomic Decomposition by Basis Pursuit. *SIAM Journal on Scientific Computing*, **20**, 33-61. https://doi.org/10.1137/S1064827596304010

[5] Daubechies, I., Defries, M. and De Mol, C. (2004) An Iterative Thresholding Algorithm for Linear Inverse Problems with a Sparsity Constraint. *Communications on Pure and Applied Mathematics*, **57**, 1413-1457. https://doi.org/10.1002/cpa.20042

[6] Xu, Z.B., Chang, X., Xu, F., *et al*. (2012) Regularization: A Thresholding Representation Theory and a Fast Solver. *IEEE Transactions on Neural Networks and Learning Systems*, **23**, 1013-1027. https://doi.org/10.1109/TNNLS.2012.2197412

[7] Jeevan, K., Pant and Lu, W. (2014) New Improved Algorithms for Compressive Sensing Based on l(p) Norm. Circuits and Systems II: Express Briefs. *IEEE Transactions*, **61**, 198-202.

[8] Saheya, B., Yu, C.-H. and Chen, J.-S. (2018) Numerical Comparisons Based on Four Smoothing Functions for Absolute Value Equation. *Journal of Applied Mathematics and Computing*, **56**, 131-149. https://doi.org/10.1007/s12190-016-1065-0

[9] Bakhtawar, B., Zabidin, S., Ahmad, A., *et al*. (2017) A New Modified Three-Term Conjugate Gradient Method with Sufficient Descent Property and Its Global Convergence. *Journal of Mathematics*, **2017**, 1-12.

[10] Liu, C., Chen, Q., Zhou, B., *et al*. (2016) L1- and L2-Norm Joint Regularization Based Sparse Signal Reconstruction Scheme. *Applied Physics A*, **45**, 313-323.

[11] Zhang, Y. and Ye, W.Z. (2017) Sparse Recovery by the Iteratively Reweighted, Algorithm for Elastic, Minimization. *Optimization*, **66**, 1-11. https://doi.org/10.1080/02331934.2017.1359590