

修正 HS 共轭梯度法在大规模信号重构问题中的应用

陈风华, 李双安

(郑州工商学院公共基础教学部, 河南 郑州 451400)

摘要: 本文研究了压缩感知在大规模信号恢复问题中应用的问题. 利用修正 HS 共轭梯度法及光滑化方法, 获得了具有较好重构效果的算法. 数值实验表明用修正 HS 共轭梯度法解决大规模信号恢复问题是可行的.

关键词: 压缩感知; 修正 HS 共轭梯度法; 稀疏信号; Nesterov 光滑技术

MR(2010) 主题分类号: 90C05; 65K05 中图分类号: O221.1

文献标识码: A 文章编号: 0255-7797(2016)06-1291-08

1 引言

已知 $\bar{u} \in \mathbf{R}^n$ 是稀疏信号, $A \in \mathbf{R}^{m \times n} (m \ll n)$ 是线性算子, $b \in \mathbf{R}^m$ 是观测向量, 它们满足下面的关系式 $b = A\bar{u}$. 原信号 \bar{u} 从线性系统中重构是压缩感知问题中稀疏信号重构的核心问题, 然而该线性方程是欠定的, 它有无穷多组解.

在压缩感知问题中, 若先验知道原信号在某个变换下是可压缩的 (或是稀疏的), 则可以通过如下一个最小 l_0 范数问题恢复信号 \bar{u}

$$\min_{u \in \mathbf{R}^n} \|u\|_0 \quad \text{s.t. } Au = b, \quad (1.1)$$

这里 l_0 范数表示向量非零元素的个数. 但是求解最小 l_0 范数问题是一个 NP 难问题, 即最小 l_0 范数问题的求解需要列举原信号中非零值的所有排列组合情况, 计算量巨大, 相关研究见文献 [1].

此后, 人们转而求解最小 l_1 范数问题, 而且证明了在一定的条件下, 最小 l_1 范数问题和最小 l_0 范数问题的解是等价的, 见文献 [2]. 因此问题 (1.1) 转化为求解下面的最小 l_1 范数问题

$$\min_{u \in \mathbf{R}^n} \|u\|_1 \quad \text{s.t. } Au = b. \quad (1.2)$$

最小 l_1 范数问题的求解是一个凸优化问题, 这个问题能够转化为线性规划问题, 见文献 [3, 4]. 然而在实际中求解线性规划问题 (1.2) 是很难的. 在压缩感知问题中, 当矩阵 A 是大且稠密的, 线性规划问题 (1.2) 常常是病态的. 设矩阵 A 是行满秩的, 问题 (1.2) 可转化为 l_1 正则化最小二乘问题 (见文献 [5])

$$\min_{u \in \mathbf{R}^n} \lambda \|u\|_1 + \frac{1}{2} \|Au - b\|_2^2, \quad (1.3)$$

*收稿日期: 2015-12-16 接收日期: 2016-04-15

基金项目: 国家自然科学基金 (11061011; 11361018); 河南省高等学校重点科研项目 (17A110032); 河南省教育厅科学技术研究重点项目 (12B110011).

作者简介: 陈风华 (1982-), 女, 湖北仙桃, 讲师, 主要研究方向: 优化理论与算法研究.

这里 λ 是正则化参数.

2 算法

因为 $\|u\|_1$ 是关于 u 的非光滑凸函数, 所以对于问题 (1.3), 次梯度基本方法的效果可能很差. 本文通过求解问题 (1.3) 的一个适当的光滑形式, 来求解 ℓ_1 最小化问题 (1.2).

由文献 [6], 利用 Nesterov 光滑技术, 光滑化 $\|u\|_1$ 为如下光滑函数

$$f_\tau(u) = \sum_{i=1}^n H(u(i)),$$

其中

$$H(y) = \begin{cases} \frac{y^2}{2\tau}, & |y| \leq \tau; \\ |y| - \frac{\tau}{2}, & |y| > \tau, \end{cases}$$

表示 Hüber 惩罚函数, τ 为光滑参数, 相关研究见文献 [3].

光滑函数 $f_\tau(u)$ 是凸的, 梯度 $\nabla f_\tau(u)$ 是 Lipchitz 连续的, 且其分量为

$$\nabla f_\tau(u(i)) = \begin{cases} \frac{u(i)}{\tau}, & |u(i)| \leq \tau, \\ \text{sign}(u(i)), & |u(i)| > \tau \end{cases} \quad (i = 1, 2, \dots, n),$$

其中

$$\text{sign}(x) = \begin{cases} 1, & x > 0; \\ 0, & x = 0; \\ -1, & x < 0. \end{cases}$$

令 $F(u) = \lambda f_\tau(u) + \frac{1}{2} \|Au - b\|_2^2$, 则问题 (1.3) 转化为如下无约束光滑凸规划问题

$$\min_{u \in \mathbf{R}^n} F(u), \quad (2.1)$$

$F(u)$ 是可微的, 其梯度 $\nabla F(u) = \lambda \nabla f_\tau(u) + A^T(Au - b)$ 是 Lipchitz 连续的. 为方便, 记 $g_k = \nabla F(u_k)$, $y_k = \nabla F(u_{k+1}) - \nabla F(u_k) = g_{k+1} - g_k$.

求解问题 (2.1), 共轭梯度法的迭代公式如下

$$u_{k+1} = u_k + \alpha_k d_k, \quad (2.2)$$

$$d_k = \begin{cases} -g_k, & k = 0; \\ -g_k + \beta_k d_{k-1}, & k \geq 1, \end{cases} \quad (2.3)$$

其中 α_k 为搜索步长, d_k 为搜索方向, β_k 为参数.

共轭梯度法的关键是参数 β_k 的选取, 常用的 β_k 公式有如下几种

$$\beta_k^{HS} = \frac{g_k^T y_{k-1}}{d_{k-1}^T y_{k-1}}, \quad \beta_k^{FR} = \frac{\|g_k\|^2}{\|g_{k-1}\|^2}, \quad \beta_k^{DY} = \frac{\|g_k\|^2}{d_{k-1}^T y_{k-1}}, \quad \beta_k^{PRP} = \frac{g_k^T y_{k-1}}{\|g_{k-1}\|^2}.$$

一些文献对 β_k 的构造及算法的收敛性做了大量的研究, 见文献 [7, 8]. 文献 [8] 指出, PRP, HS 方法在数值计算中有很好的表现, 但是即使使用精确线搜索 PRP 方法也不会有全

局收敛性. FR, DY 方法则具有较好的收敛性质. 为寻求兼具良好收敛性和优秀数值结果的算法, 许多研究者在经典共轭梯度法的基础上推出新的 β_k 公式.

本文我们提出用修正 HS 共轭梯度法解决大规模信号恢复问题 (2.1), 且在适当的假设条件下证明了本文提出的算法具有全局收敛性, 最后数值实验结果表明用修正 HS 共轭梯度法解决大规模信号恢复问题是可行的.

本文构造如下搜索方向

$$d_k = \begin{cases} -g_k, & k = 0; \\ -g_k + \beta_k d_{k-1} - \theta_k g_k, & k \geq 1, \end{cases} \quad (2.4)$$

其中

$$\beta_k = \frac{g_k^T y_{k-1}}{d_{k-1}^T y_{k-1}}, \quad \theta_k = \frac{g_k^T y_{k-1}}{d_{k-1}^T y_{k-1}} \frac{g_k^T d_{k-1}}{\|g_k\|^2}, \quad y_{k-1} = g_k - g_{k-1} = \nabla F(u_k) - \nabla F(u_{k-1}).$$

引理 2.1 对任一线搜索, 搜索方向由 (2.4) 式定义, 则有

$$g_k^T d_k = -\|g_k\|^2. \quad (2.5)$$

证 由 (2.4) 式, 有

$$g_k^T d_k = -g_k^T g_k + \frac{g_k^T y_{k-1}}{d_{k-1}^T y_{k-1}} g_k^T d_{k-1} - \frac{g_k^T y_{k-1}}{d_{k-1}^T y_{k-1}} \frac{g_k^T d_{k-1}}{\|g_k\|^2} g_k^T g_k = -g_k^T g_k.$$

即 $g_k^T d_k = -\|g_k\|^2$. 这表明搜索方向 d_k 是目标函数 $F(u)$ 在 u_k 处的下降方向, 若使用精确线搜索, 有 $g_k^T d_{k-1} = 0$, 即 $\theta_k = 0$, 则本文提出的修正 HS 共轭梯度法退化成标准的 HS 共轭梯度法.

算法 2.1 问题 (2.1) 的修正 HS 共轭梯度算法:

步骤 0 给定初始点 $u_0 \in \mathbf{R}^n$, 终止误差 $0 \leq \epsilon \ll 1$, $\tau_0 = 0.6$, $0 < \rho_1 < \rho_2 < 1$, 令 $k := 0$;

步骤 1 计算 $g_k = \nabla F(u_k)$, 若 $\|g_k\| < \epsilon$, 且光滑参数 $\tau_k \leq \epsilon$, 算法停止; 否则, 转下一步;

步骤 2 计算搜索方向 d_k , 搜索方向由 (2.4) 式定义;

步骤 3 计算步长 α_k :

$$F(u_k + \alpha_k d_k) \leq F(u_k) + \rho_1 \alpha_k \nabla F(u_k)^T d_k; \quad (2.6)$$

$$\nabla F(u_k + \alpha_k d_k)^T d_k \geq \rho_2 \nabla F(u_k)^T d_k; \quad (2.7)$$

步骤 4 更新: 令 $u_{k+1} = u_k + \alpha_k d_k$, 更新光滑参数 $\tau_{k+1} = \frac{1}{4} \tau_k$;

步骤 5 循环: 置 $k := k + 1$, 转步骤 1.

3 全局收敛性

本文提出用修正 HS 共轭梯度法解决大规模信号恢复问题 (2.1), 且在适当的假设条件下证明了本文提出的算法具有全局收敛性.

为证明算法 2.1 具有全局收敛性, 先作如下基本假设.

H1 水平集 $\Omega = \{x \in \mathbf{R}^n | F(x) \leq F(x_0)\}$ 有界.

H2 在 Ω 的某邻域 N 内, $F(x)$ 连续可微, 且其梯度 $\nabla F(x) = g(x)$ 是 Lipschitz 连续的, 即存在一常数 $L > 0$ 使得

$$\|g(x) - g(y)\| \leq L\|x - y\|, \forall x, y \in N. \quad (3.1)$$

由假设 H1、H2 知, 存在一常数 $M > 0$ 满足

$$\|g(x)\| \leq M, \forall x \in \Omega. \quad (3.2)$$

参考文献 [8, 9], 有如下两个引理.

引理 3.1 [9] 若假设条件 H1、H2 成立, 则算法是良定的.

注 由 (2.6) 式知目标函数值序列 $\{F(x_k)\}$ 是一下降序列, 有 $\sum_{k \geq 0} \alpha_k |g_k^T d_k| < +\infty$, 结合 (3.2) 式有

$$\lim_{k \rightarrow \infty} \alpha_k \|d_k\| = 0. \quad (3.3)$$

引理 3.2 [8] 若假设条件 H1、H2 成立, 对任一共轭梯度法的迭代形式 (2.2), 其中 d_k 满足 $g_k^T d_k < 0$ 且 α_k 满足条件 (2.6) 和 (2.7), 则

$$\sum_{k \geq 0} \frac{(g_k^T d_k)^2}{\|d_k\|^2} < +\infty. \quad (3.4)$$

式 (3.4) 结合式 (2.5) 有

$$\sum_{k \geq 0} \frac{\|g_k\|^4}{\|d_k\|^2} < +\infty. \quad (3.5)$$

定理 3.3 如果假设条件 H1、H2 成立, 算法 2.1 产生的迭代序列为 $\{u_k\}$, 若存在常数 $r > 0$ 满足

$$d_k^T y_k \geq r, \quad (3.6)$$

则 $\lim_{k \rightarrow \infty} \inf \|g_k\| = 0$.

证 由 (2.4) 式定义的 d_k 及式 (3.1)、(3.2)、(3.6) 有

$$\|d_k\| \leq \|g_k\| + \frac{2\|g_k\|\|y_{k-1}\|d_{k-1}\|}{d_{k-1}^T y_{k-1}} \leq M + \frac{2ML\alpha_{k-1}\|d_{k-1}\|}{r} \|d_{k-1}\|.$$

由 (3.3) 式知存在一常数 $t \in (0, 1)$, 存在一正整数 k_0 , 使得当 $k \geq k_0$ 有

$$\frac{2ML}{r} \alpha_{k-1} \|d_{k-1}\| \leq t.$$

因此 $\forall k > k_0$ 有

$$\begin{aligned} \|d_k\| &\leq M + t\|d_{k-1}\| \leq M + t(M + t\|d_{k-2}\|) \\ &\leq M(1 + t + t^2 + \cdots + t^{k-k_0-1}) + t^{k-k_0}\|d_{k-k_0}\| \\ &\leq \frac{M}{t-1} + \|d_{k_0}\|. \end{aligned}$$

令 $m = \max\{\|d_1\|, \|d_2\|, \cdots, \|d_{k_0}\|, \frac{M}{t-1} + \|d_{k_0}\|\}$, 则

$$\|d_k\| \leq m, \forall k. \quad (3.7)$$

式 (3.5) 结合式 (3.7) 有

$$\sum_{k \geq 0} \|g_k\|^4 < +\infty,$$

故 $\liminf_{k \rightarrow \infty} \|g_k\| = 0$.

在压缩感知理论中, 当 $\{\min \|u\|_1 : Au = b\}$ 有唯一解时, 精确恢复才会实现. 参考文献 [5], 有下面的定理.

定理 3.4 设 ℓ_1 范数最小化问题 $\min\{\|u\|_1 : Au = b\}$ 有唯一最优解, 算法 2.1 产生的迭代序列为 $\{u_k\}$, 当 $\|u\|_1$ 的光滑参数 $\tau_k \rightarrow 0$ 时, 则 $\lim_{k \rightarrow \infty} u_k = u^*$, 这里 $u^* = \arg \min\{\|u\|_1 : Au = b\}$.

证 记第 k 步的目标函数 $F_k(u)$ 的最小值点为 u_k^* , 即 $u_k^* = \arg \min_{u \in \mathbf{R}^n} F_k(u)$.

由 Nesterov 光滑技术得到的函数 $f_{\tau_k}(u_k)$ 的光滑性有 $f_{\tau_k}(u_k) \geq \|u_k\|_1 - \frac{n\tau_k}{2}$, 故

$$\lambda \|u_k\|_1 \leq \lambda f_{\tau_k}(u_k) + \frac{\lambda n \tau_k}{2} \leq \lambda f_{\tau_k}(u_k) + \frac{1}{2} \|Au_k - b\|_2^2 + \frac{\lambda n \tau_k}{2} = f_k(u_k) + \frac{\lambda n \tau_k}{2}.$$

因为 $F_k(u)$ 是凸函数, 有 Lipschitz 连续梯度, 由文献 [5] 引理 2.3 有, Lipschitz 常数满足

$$F_k(u_k) \leq F_k(u_k^*) + \frac{\tau_k^2}{2L_k},$$

所以

$$\lambda \|u_k\|_1 \leq F_k(u_k^*) + \frac{\tau_k^2}{2L_k} + \frac{\lambda n \tau_k}{2}. \tag{3.8}$$

故当光滑参数 $\tau_k \rightarrow 0$ 时, (3.8) 式表明序列 $\{u_k\}$ 有极限点. 令 u^* 表示这个序列的任意极限点, 由文献 [5] 结合定理 3.3, 可证 u^* 是可行点, 且是问题 (1.2) 的 KKT 点. 又问题 (1.2) 是凸规划问题, 故 u^* 是问题 (1.2) 的最优解.

4 数值实验

实验中, 参数选取 $\rho_1 = 0.2$, $\rho_2 = 0.7$, 误差精度 $\epsilon = 10^{-8}$. A 为 $m \times n$ 维随机矩阵, u_{exact} 表示 n 维 k 稀疏原信号, 满足 $Au_{\text{exact}} = b$, u 表示本文算法恢复出来的信号, 相对残差 $\frac{\|Au - b\|_2}{\|b\|_2}$, 相对误差 $\frac{\|u - u_{\text{exact}}\|_2}{\|u_{\text{exact}}\|_2}$. 图 1 和图 2 中的圈和点分别表示原信号和经由我们的算法恢复出来的信号.

表 1、图 1 选取 A 的维数都是 512×1024 , 表 2 对应的正则化参数 $\lambda = 0.1$, 图 2 选取 A 的维数分别是 256×512 , 512×1024 , 1024×2048 , 2048×4096 .

表 4.1: λ 按递减规则产生的数值结果

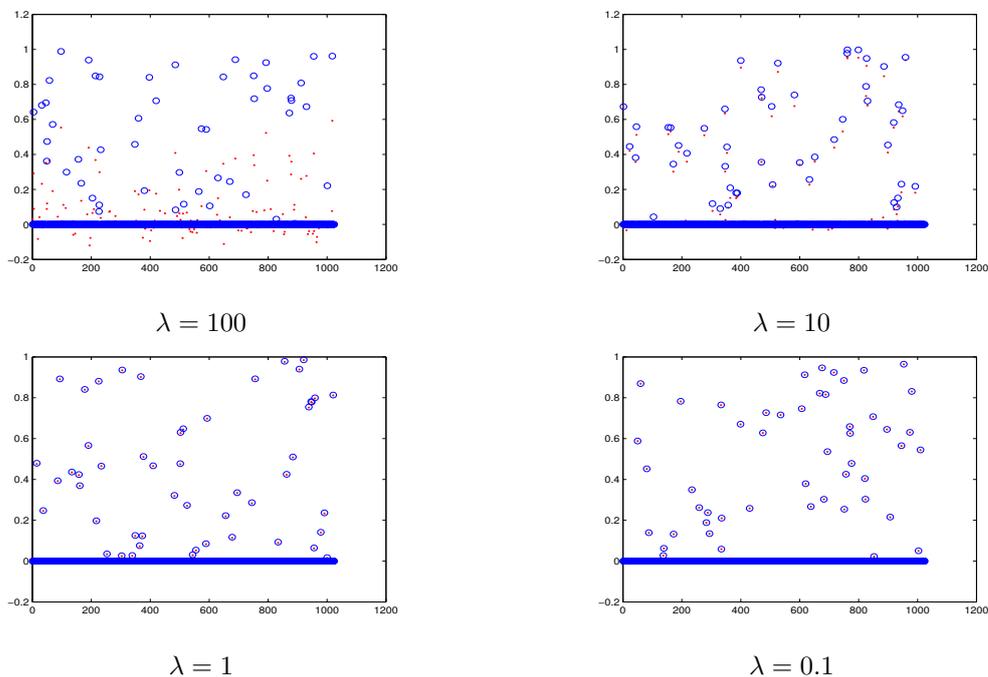
λ	100	10	1	0.1
相对残差	0.0698	0.0698	0.0051	5.0066e-004
相对误差	0.1876	0.1603	0.0059	0.0020
运行时间	119s	96s	77s	67s

表 4.2: 信号维数按递增规则产生的数值结果

$m \times n$	256×512	512×1024	1024×2048	2048×4096
相对残差	8.9212e-004	5.0066e-004	1.7119e-004	1.5245e-004
相对误差	0.0037	0.0020	0.0011	6.5435e-004
运行时间	21s	67s	209s	903s

表 4.3: 信号维数按递增规则产生的数值结果

$m \times n$	算法	相对残差	相对误差	运行时间
128×256	MHSCG	0.0018	0.0040	9s
	HSCG	0.0020	0.0073	10s
	PRPCG	0.0025	0.0053	13s
256×512	MHSCG	8.5860e-004	0.0031	20s
	HSCG	9.3935e-004	0.0066	21s
	PRPCG	12.1786e-004	0.0058	25s
512×1024	MHSCG	4.4897e-004	0.0011	66s
	HSCG	5.6281e-004	0.0027	72s
	PRPCG	7.9732e-004	0.0013	57s
1024×2048	MHSCG	2.7606e-004	2.7606e-004	207s
	HSCG	3.7640e-004	8.9788e-004	225s
	PRPCG	3.8776e-004	7.7406e-004	213s

图 4.1: 四幅图分别表示 λ 不同取值下的信号恢复效果图

5 结论

本文提出用基于 Nesterov 光滑技术和修正 HS 共轭梯度法的压缩感知重构算法解决大规模信号恢复问题, 最后做了两类数值实验, 第一类反映了随着正则化参数 λ 的变化, 相对残差、相对误差、运行时间的变化, 以及随着原信号的维数的变化, 相对残差、相对误差、运行时间的变化; 第二类实验用本文提出的修正 HS 共轭梯度法与 HS 共轭梯度法、PRP 共轭梯

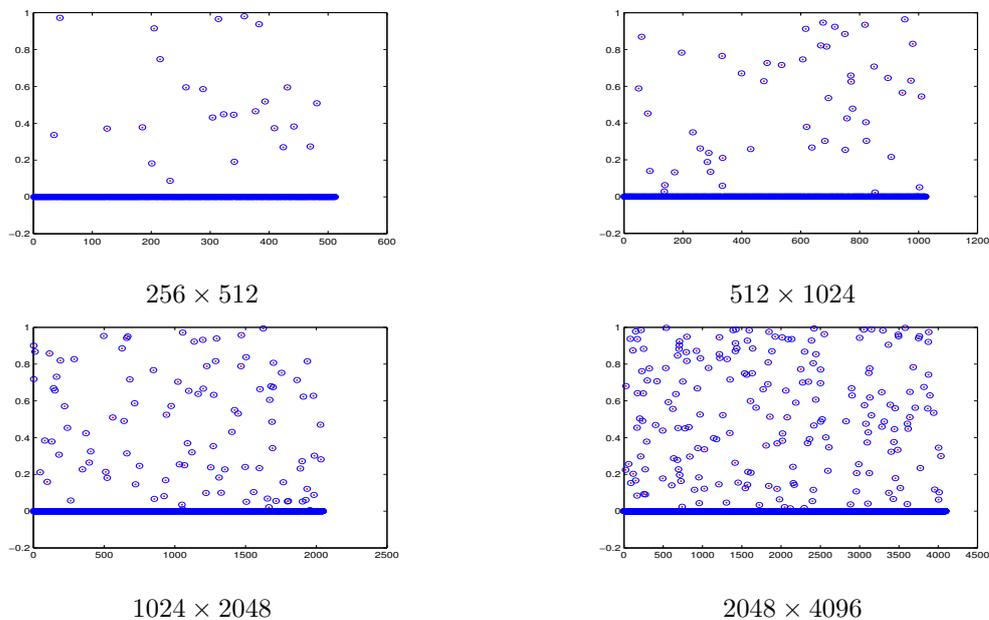


图 4.2: 四幅图分别表示信号维数不同取值下的信号恢复效果图

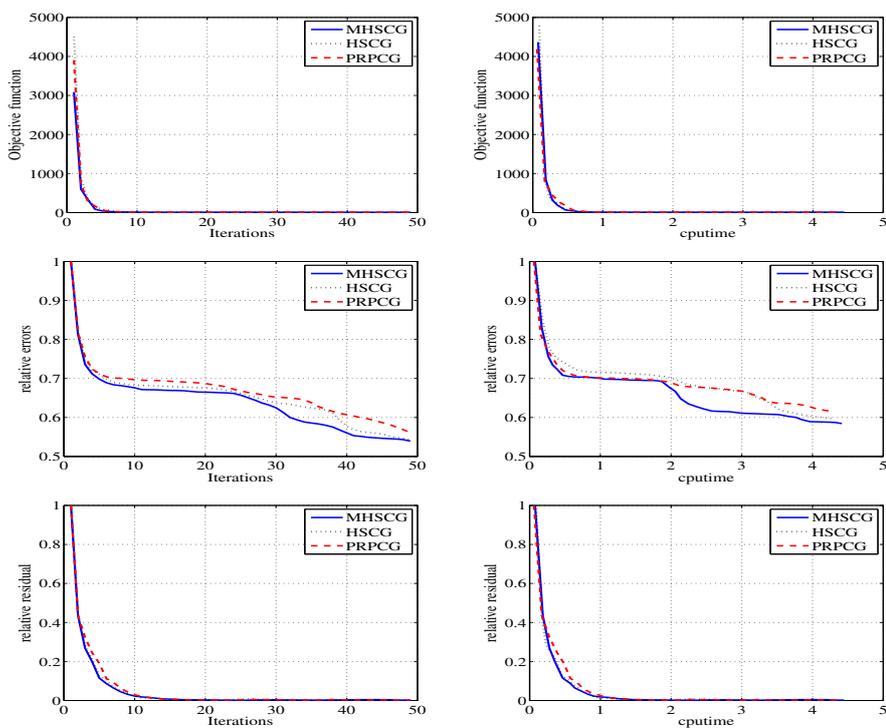


图 4.3: 修正 HS 共轭梯度法 (图中标为 MHSCG) 与 HS 共轭梯度法 (图中标为 HSCG)、PRP 共轭梯度法 (图中标为 PRPCG) 作比较, 上图分别反映了目标函数值、相对误差、相对残差随迭代步、运行时间的变化.

度法解决大规模信号恢复问题作对比, 实验结果表明用修正 HS 共轭梯度法解决大规模信号恢复问题是可行的. 该算法在图像处理中有直接的应用价值. 相关研究见文献 [10, 11].

参 考 文 献

- [1] Zhu H, Xiao Y, Wu S Y. Large sparse signal recovery by conjugate gradient algorithm based on smoothing technique[J]. *Comp. Math. Appl.*, 2013, 66(1): 24–32.
- [2] Donoho D L. For most large underdetermined systems of linear equations the minimal 1-norm solution is also the sparsest solution[J]. *Commun. Pure Appl. Math.*, 2006, 59(6): 797–829.
- [3] Candes E J, Tao T. Near Optimal Signal Recovery From Random Projections And Universal Encoding Strategies[J]. *IEEE Trans. Inform. The.*, 2007, 52(12): 5406–5425.
- [4] Donoho L D. Compressed sensing[J]. *IEEE Trans. Inform. The.*, 2006, 52(4): 1289–1306.
- [5] Aybat S N, Iyengar G. A first-order smoothed penalty method for compressed sensing[J]. *SIAM J. Optim.*, 2011, 21(1): 287–313.
- [6] Nesterov Y. Smooth minimization of non-smooth functions[J]. *Math. Prog.*, 2005, 103(1): 127–152.
- [7] Al Baali M. Descent property and global convergence of the Fletcher-Reeves method with inexact line search[J]. *IMA J. Numer. Anal.*, 1985, 5(1): 121–124.
- [8] Ioannis E, Livieris, Panagiotis Pintelas. Globally convergent modified Perry’s conjugate gradient method[J]. *Appl. Math. Comp.*, 2012, 218(18): 9197–9207.
- [9] Zhang L, Zhou W, Li D H. A descent modified Polak-Ribiere-Polyak conjugate gradient method and its global convergence[J]. *Ima. J. Numer. Anal.*, 2006, 26(4): 629–640.
- [10] 陈风华, 李双安. BFGS 校正拟牛顿法解决大规模信号恢复问题 [J]. *数学杂志*, 2015, 35(3): 727–734.
- [11] 房明磊, 朱志斌. 互补约束规划问题的一个广义梯度投影法 [J]. *数学杂志*, 2011, 31(4): 685–694.

THE APPLICATION OF A MODIFIED HS CONJUGATE GRADIENT METHOD FOR LARGE-SCALE SIGNAL RECONSTRUCTION PROBLEM

CHEN Feng-hua, LI Shuang-an

(Teaching Department of the Public Infrastructure, Zhengzhou Technology and Business University,
Zhengzhou 451400, China)

Abstract: In this paper we study application about the compressed sensing in large-scale signal recovery problem. By the modified HS conjugate gradient method and smoothing technique, the algorithm which possesses better reconstruction effect is obtained. Preliminary numerical results show that our algorithm is suitable for solving large-scale sparse signal recovery problems.

Keywords: compressed sensing; modified HS conjugate gradient method; sparse signal; Nesterov’s smoothing technique

2010 MR Subject Classification: 90C05; 65K05