Vol. 35 (2015) No. 3 数 学 杂 志

J. of Math. (PRC)

BFGS 校正拟牛顿法解决大规模信号恢复问题

陈凤华,李双安

(河南理工大学万方科技学院,河南郑州 451400)

摘要: 本文采用 BFGS 校正拟牛顿法研究了大规模信号恢复问题 $\min\{||u||_1 : Au = b\}$,这个问题通常被转化为 ℓ_1 正则化最小二乘问题. 利用 Nesterov 光滑化技术对 $||u||_1$ 进行光滑化处理, 原问题被转化为无约束光滑凸规划问题, 最后获得了较好的数值实验结果, 实验结果表明用 BFGS 校正拟牛顿法解决大规模信号恢复问题是可行的.

关键词: 压缩感知; BFGS 校正; 拟牛顿法; 稀疏信号; 光滑优化
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(2015)03-0727-08

1 引言

已知稀疏信号 $\bar{u} \in \mathbf{R}^n$, $A \in \mathbf{R}^{m \times n}$ ($m \ll n$) 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:

$$b = A\overline{u}.$$

原信号 *ū* 从线性系统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 ℓ₀ 范数问题恢复信号:

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

这里 ℓ0 范数表示向量非零元素的个数.

Candés 等人证明了可以通过求解线性优化最小 ℓ_0 范数问题精确重构信号,因为最小 ℓ_0 范数问题的解是最稀疏的解,能够得到原信号的最优稀疏逼近.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 [1–3].

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 [4].因此问题 (1.1) 转化为求解下面的最小 ℓ_1 范数问题:

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

*收稿日期: 2014-08-15 接收日期: 2014-11-15

基金项目:国家自然科学基金 (11361018); 广西杰出青年基金 (2012GXSFFA060003); 河南省教育厅科 学技术研究重点项目 (12B110011).

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

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. Candés, Romberg, Donoho 已经证明了当目标信号 $\overline{u} \in \mathcal{K}$ - 稀疏的,即 n 个分量只有 k 个分量是非 零的,且测量矩阵满足一些规则性条件,理论上稀疏信号能够通过解问题 (1.2) 得到有效恢 复,见文献 [5–7]. 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,问题 的维数可高达 $n = 10^6$,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 ℓ1 正则化最小二乘问题 (见文献 [8]):

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

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 [9], 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

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

表示 Hüber 惩罚函数, τ 为光滑参数, 相关研究见文献 [6]. 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

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

其中

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

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:

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

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

$$s_k = u_{k+1} - u_k, g_k = \nabla F(u_k), y_k = \nabla F(u_{k+1}) - \nabla F(u_k) = g_{k+1} - g_k.$$

求解问题 (2.1), 牛顿法的迭代公式如下:

$$u_{k+1} = u_k + d_k,$$

其中搜索方向 d_k 通过求解如下线性方程而得:

$$\nabla^2 F(u_k) d_k = -g_k. \tag{2.2}$$

牛顿法的优点是具有二阶收敛速度, 但当 Hesse 矩阵 $\nabla^2 F(u_k)$ 不正定时, 不能保证所产 生的方向是目标函数 F(u) 在 u 处的下降方向. 特别地, 当 Hesse 矩阵 $\nabla^2 F(u_k)$ 奇异时, 算法 就无法进行下去.

本文我们提出用 BFGS 校正拟牛顿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 BFGS 校正拟牛顿 法解决大规模信号恢复问题是可行的.

拟牛顿法的基本思想是用某个矩阵 B_k 近似取代 (2.2) 式中的 Hesse 阵 $\nabla^2 F(u_k)$, 通常 B_k 应具有的特点:

1. 对所有的 k, B_k 是对称正定的, 从而使得算法所产生的方向是目标函数 F(u) 在 u_k 处的下降方向;

2. 矩阵 B_k 更新规则相对较简单.

本文搜索方向 dk 通过求解如下线性方程而得

$$B_k d_k = -g_k.$$

其中矩阵 B_k 更新规则如下:

$$B_{k+1} = B_k - \frac{B_k s_k s_k^T B_k}{s_k^T B_k s_k} + \frac{y_k y_k^T}{y_k^T s_k},$$
(2.3)

该公式称为 BFGS 校正公式.

算法 2.1 问题 (2.1) 的 BFGS 校正拟牛顿算法:

步骤 0: 初始点 $u_0 \in \mathbb{R}^n$, 终止误差 $0 \le \epsilon \ll 1$, 初始对称正定矩阵 $B_0 = I_n(n \text{ 阶单位阵})$, $\tau_0 = 0.8$, 且 $0 < \rho_1 < \rho_2 < 1$, 令 k := 0.

步骤 1: 计算 $g_k = \nabla F(u_k)$, 若 $||g_k|| < \epsilon$, 且光滑参数 $\tau_k \le \epsilon$, 算法停止; 否则, 转下一步. **步骤 2**: 通过求解线性方程得搜索方向 d_k .

$$B_k d_k = -g_k.$$

步骤 3: 由 Wolfe 线搜索计算步长 α_k .

$$F(u_k + \alpha_k d_k) \le F(u_k) + \rho_1 \alpha_k \nabla F(u_k)^T d_k;$$
(2.4)

$$\nabla F(u_k + \alpha_k d_k)^T d_k \ge \rho_2 \nabla F(u_k)^T d_k.$$
(2.5)

步骤 5:循环:置 k := k + 1,返回步骤 1.

3 全局收敛性

本文提出用 BFGS 校正拟牛顿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.

引理 3.1 设 B_k 对称正定, B_{k+1} 由 BFGS 校正公式 (2.3) 确定, 则 B_{k+1} 对称正定的充 要条件是 $y_k^T s_k > 0$.

证 详细证明参考文献 [10].

引理 3.1 表明, 若初始矩阵 B_0 对称正定且在迭代过程中保持 $y_k^T s_k > 0$ ($\forall k \ge 0$), 则由 BFGS 校正公式产生的矩阵序列 { B_k } 是对称正定的, 从而方程 $B_k d_k = -g_k$ 有唯一解 d_k , 故 有

$$d_k^T B_k d_k = -g_k^T d_k,$$

因为 $d_k^T B_k d_k > 0$, 所以 $g_k^T d_k < 0$, 即 d_k 是目标函数 F(u) 在 u_k 处的下降方向.

引理 3.2 若在 BFGS 校正拟牛顿算法中采用 Wolfe 搜索准则,则有 $y_k^T s_k > 0$. 证 因为

$$y_k^T s_k = (g_{k+1} - g_k)^T \alpha_k d_k = \alpha_k (g_{k+1}^T d_k - g_k^T d_k),$$

由算法步骤 3 中 (2.5) 式有

$$g_{k+1}^T d_k \ge \rho_2 g_k^T d_k,$$

故

$$y_k^T s_k \ge \alpha_k (\rho_2 - 1) g_k^T d_k,$$

又因为 $\rho_2 < 1, g_k^T d_k < 0, 故 y_k^T s_k > 0.$

引理 3.3 若 l₁ 范数 ||u||₁ 按光滑化技术得到的光滑函数为

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

而

$$H(u(i)) = \begin{cases} \frac{u^2(i)}{2\tau}, & |u(i)| \le \tau, \\ |u(i)| - \frac{\tau}{2}, & |u(i)| > \tau, \end{cases}$$

则有 $f_{\tau}(u) \leq ||u||_1$.

$$\mathbf{\ddot{u}} \, \, \overset{}{=} \, |u(i)| \leq \tau \, \, \mbox{tr}, \, f_{\tau}(u) = \sum_{i=1}^{n} \frac{u^{2}(i)}{2\tau} = \frac{1}{2\tau} \sum_{i=1}^{n} u^{2}(i) \leq \frac{1}{2\tau} \cdot \tau \sum_{i=1}^{n} |u(i)| = \frac{1}{2} ||u||_{1} \leq ||u||_{1}; \\ \overset{}{=} \, |u(i)| > \tau \, \mbox{tr}, \, f_{\tau}(u) = \sum_{i=1}^{n} (|u(i)| - \frac{\tau}{2}) = ||u||_{1} - \frac{n\tau}{2} \leq ||u||_{1}.$$

综上有 $f_{\tau}(u) < ||u||_{1}.$

定理 3.4 设 { u_k } 表示由 BFGS 校正拟牛顿法迭代产生的序列, 当 $||u||_1$ 的光滑参数 $\tau_k \to 0$ 时, 则 { u_k } 是有界序列. 令 u^* 表示 { u_k } 的任意极限点, 则 u^* 是 ℓ_1 范数最小化问题 (1.2) 的最优解.

证 对于第 *k* 步的光滑参数 τ_k , 令 $u_k^* = \arg \min_{u \in \mathbf{R}^n} F_k(u)$ 表示第 *k* 步的目标函数 $F_k(u)$ 的 最小值点, *v* 表示满足 Av = b 的任意向量. 因为 $u_k^* = \arg\min_{u \in \mathbf{R}^n} F_k(u)$,所以 $F_k(u_k^*) \leq F_k(v)$.又Av = b,故

$$F_k(v) = \lambda f_{\tau_k}(v) + \frac{1}{2} ||Av - b||_2^2 = \lambda f_{\tau_k}(v).$$

因为 $F_k(u) = \lambda f_{\tau_k}(u) + \frac{1}{2} ||Au - b||_2^2, ||Au - b||_2^2 \ge 0,$ 故

$$F_k(u) \ge \lambda f_{\tau_k}(u). \tag{3.1}$$

而由引理 3.3 有 $f_{\tau_k}(v) \leq ||v||_1$, 则

$$F_k(u_k^*) \le F_k(v) = \lambda f_{\tau_k}(v) \le \lambda ||v||_1.$$

$$(3.2)$$

结合式 (3.1)、(3.2) 及 $f_{\tau_k}(u)$ 的光滑性有

$$\|u_k\|_1 \le f_{\tau_k}(u_k) + \frac{n\tau_k}{2} \le \frac{1}{\lambda} F_k(u_k) + \frac{n\tau_k}{2} \le \frac{1}{\lambda} \lambda \|v\|_1 + \frac{n\tau_k}{2} = \|v\|_1 + \frac{n\tau_k}{2},$$

故当光滑参数 $\tau_k \to 0$ 时, $\{u_k\}$ 是有界序列.

另一方面,参考文献 [8],可证序列 $\{u_k\}$ 的极限点 u^* 是可行点,且极限点 u^* 是 ℓ_1 范数 最小化问题 (1.2)的 KKT 点.因为 $||u||_1$ 是凸的,对于问题 (1.2) 是凸规划问题, KKT 条件 对于最优性是充分的,故 $u^* \in \ell_1$ 范数最小化问题 (1.2)的最优解.

在压缩感知理论中, 当 ℓ_1 范数最小化问题 min{ $||u||_1 : Au = b$ } 有唯一解时, 精确恢复才 会实现. 参考文献 [8], 有下面的推论.

推论 3.5 设 ℓ_1 范数最小化问题 min{ $||u||_1 : Au = b$ } 有唯一最优解, 令 { u_k } 表示由 BFGS 校正拟牛顿法迭代产生的序列, 当 $||u||_1$ 的光滑参数 $\tau_k \to 0$ 时, 则 $\lim_{k\to\infty} u_k = u^*$, 这里 $u^* = \arg\min\{||u||_1 : Au = b\}.$

4 数值实验

实验中,参数选取 $\rho_1 = 0.5$, $\rho_2 = 0.8$, 误差精度 $\epsilon = 10^{-8}$. A 为 $m \times n$ 维随机矩阵, x_{exact} 表示 n 维 k 稀疏原信号, 满足 $Ax_{\text{exact}} = b$, x 表示本文算法恢复出来的信号, 相对残差 $\frac{\|Ax-b\|_2}{\|b\|_2}$, 相对误差 $\frac{\|x-x_{\text{exact}}\|_2}{\|x_{\text{exact}}\|_2}$. 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 256 × 512, 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 1: λ 按递减规则产生的数值结果

λ	50	10	1	0.1
相对残差	0.3872	0.0669	0.0086	8.7323e-004
相对误差	0.5578	0.1454	0.0174	0.0081
运行时间	113s	118s	119s	99s

表 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	8.7323e-004	5.1328e-004	2.2377e-004	1.4549e-004
相对误差	0.0081	0.0039	9.8647 e-004	4.1573e-004
运行时间	99s	215s	627s	942s





图 1: 四幅图分别表示 λ 不同取值下的信号恢复效果图.

本文用 BFGS 校正拟牛顿法与对称秩 1 校正拟牛顿法解决大规模信号恢复问题作比较, 其中 对称秩 1 校正公式如下:

$$B_{k+1} = B_k + \frac{(y_k - B_k s_k)(y_k - B_k s_k)^T}{(y_k - B_k s_k)^T s_k}$$

5 结论

本文提出用 BFGS 校正拟牛顿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





图 3: BFGS 校正拟牛顿法 (图中标为 BFGS) 与对称秩 1 校正拟牛顿法 (图中标为 sr1) 作比较, 上图分别反映了目标函数值、相对误差随迭代步的变化.

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 BFGS 校正拟牛顿法与对称秩1校正拟牛顿法解决大规模信号恢复问题作对比,实验结果表明用 BFGS 校正拟牛顿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 [11, 12].

参考文献

- Mishali M, Elder Y C. From theory to practice: sub-nyquist sampling of sparse wideband analog signals[J]. IEEE J. Selected Topics Sign. Proc., 2009, 4(2): 375–391.
- [2] Duarte M, Daeenport M, Baraniuk R. Single-pixel imaging via compressive sampling [J]. IEEE Sign. Proc. Magzine, 2008, 25(2): 83–91.
- [3] Zhu Hong, Xiao Yunhai, Wu Soon-Yi. Large sparse signal recovery by conjugate gradient algorithm based on smoothing technique[J]. Comput. Math. Appl., 2013, 66: 24–32.
- [4] Donoho L D. For most large underdetemind systems of linear equations, the minimal-norm solution is also the sparsest solution[J]. Comm. Pure Appl. Math, 2006, 59: 907–934.
- [5] Candes E, Romberg J. Quantitative robust uncertainty principles and optimally sparse decompositions [J]. Found. Comput. Math., 2006, 6: 227–254.
- [6] Candes E, Tao T. Near optimal signal recovery from random projections: universal encoding strategies[J]. IEEE Trans. Info. Theory, 2006, 52: 5406–5425.
- [7] Donoho L D. Compressed sensing[J]. IEEE Trans. Inform. Theory, 2006, 52: 1289–1306.
- [8] Aybat S N, Iyengar G. A first-order smoothed penalty method for compressed sensing[J]. SIAM J. Optim., 2011, 21(1): 287–313.
- [9] Nesterov Y. Smooth minimization of non-smooth functions[J]. Math. Prog., 2005, 103: 127–152.
- [10] 马昌凤. 最优化方法及其 Matlab 程序设计 [M]. 北京: 科学出版社, 2010.
- [11] 景书杰, 苗荣, 李少娟. 一个新的 MBFGS 信赖域算法 [J]. 数学杂志, 2014, 34(3): 569-576.
- [12] 郑莲,金茂明. 解变分不等式的一种二次投影迭代算法 [J]. 数学杂志, 2013, 33(5): 902-908.

LARGE-SCALE SPARSE SIGNAL RECOVERY BY QUASI-NEWTON METHOD OF BFGS CORRECTION

CHEN Feng-hua, LI Shuang-an

(Wanfang Institute of Sci. and Tech., Henan Polytechnic University, Zhengzhou 451400, China)

Abstract: In this paper we study the lage-scale sparse signal recovery problem such as $\min\{||u||_1 : Au = b\}$, adopting the quasi-Newton method of BFGS correction. This problem is usually transformed into ℓ_1 -regularized least-squares programs. By using the Nesterov's smoothing method for $||u||_1$, the original problem is transformed into an unconstrained smoothing convex programming. Further the numerical solution of the algorithm is obtained. Preliminary numerical results show that our algorithm is feasible for solving large-scale sparse signal recovery problems.

Keywords: compressed sensing; BFGS correction; quasi-Newton method; sparse signal; smooth optimization

2010 MR Subject Classification: 90C05; 65K05