

修正FR谱共轭梯度法在信号恢复问题中的应用*

晁丽佳, 张永富

(内蒙古民族大学 数理学院, 内蒙古 通辽 028000)

摘要:基于压缩感知理论,为解决稀疏信号恢复问题,在经典的FR共轭梯度法分析的基础上,给出一种基于谱梯度方法的修正FR谱共轭梯度法,该方法结合了FR、DY、CD等6种经典共轭梯度法的共轭参数部分特点。另外,证明了所给方法在Wolfe线搜索下具有下降方向且全局收敛,通过与其他2种算法的数值结果比较显示,本文算法具有一定优势。同时,信号重构结果表明,提出的方法重构信噪比高,耗时少,能有效的应用于稀疏信号的恢复问题。

关键词:信号恢复;谱共轭梯度法;全局收敛性;Wolfe线搜索

中图分类号:O224

文献标志码:A

文章编号:1672-6693(2023)05-0011-08

1 预备知识

利用压缩感知恢复信号,考虑方程组:

$$\mathbf{Ax} = \mathbf{y}, \quad (1)$$

式中: $\mathbf{A} \in \mathbf{R}^{m \times n}$, $\mathbf{x} \in \mathbf{R}^n$, $\mathbf{y} \in \mathbf{R}^m$ 。式(1)是一个欠定方程组,在压缩感知理论中, \mathbf{x} 是 n 维采集的信号,通过 \mathbf{A} 将原信号压缩为 m 维的 \mathbf{y} ,由于 $m \ll n$,这就提高了存储和采样的效率。

稀疏信号重构问题在压缩感知理论中应用广泛,即利用式(1)中压缩后的向量 \mathbf{y} 恢复原始信号 \mathbf{x} 。信号重构是 l_0 -范数优化问题,但是求解 l_0 -范数最小化是NP难问题,计算量大且有时无法直接求解^[1]。相关研究表明,在一定条件下, l_0 -范数最小化问题可对应于如下的 l_1 -范数问题:

$$\min \|\mathbf{x}\|_1, \text{ s. t. } \mathbf{Ax} = \mathbf{y}, \quad (2)$$

式中: $\mathbf{A} \in \mathbf{R}^{m \times n}$ 为观测矩阵, $\mathbf{x} \in \mathbf{R}^n$ 为原始稀疏信号, $\mathbf{y} \in \mathbf{R}^m$ 为观测向量。这样的转换是压缩感知研究中的一个跃进^[2-4]。

求解问题(2)可转化为求解 l_1 -范数正则化问题^[5]:

$$\min \lambda \|\mathbf{x}\|_1 + \frac{1}{2} \|\mathbf{Ax} - \mathbf{y}\|_2^2, \quad (3)$$

式中: $\lambda > 0$ 为正则化参数。由于 $\|\mathbf{x}\|_1$ 非光滑,借助文献[1]中的Nesterov光滑技术对 $\|\mathbf{x}\|_1$ 进行光滑化处理:

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

式中: $H(\mathbf{y}) = \begin{cases} \frac{\mathbf{y}^2}{2\tau}, & |\mathbf{y}| \leq \tau \\ |\mathbf{y}| - \frac{\tau}{2}, & |\mathbf{y}| > \tau \end{cases}$, $\tau > 0$ 为光滑参数。令 $F(\mathbf{x}) = \lambda f_\tau(\mathbf{x}) + \frac{1}{2} \|\mathbf{Ax} - \mathbf{y}\|_2^2$,则求解问题(3)转化为

求解无约束凸优化问题:

$$\min F(\mathbf{x}). \quad (4)$$

函数 $F(\mathbf{x})$ 是可微的,梯度为 $\nabla F(\mathbf{x}) = \lambda \nabla f_\tau(\mathbf{x}) + \mathbf{A}^T(\mathbf{Ax} - \mathbf{y})$,记 $\mathbf{g}_k = \nabla F(\mathbf{x}_k)$ 。

* 收稿日期:2022-05-07 修回日期:2022-08-16 网络出版时间:2023-06-25T11:14

资助项目:国家自然科学基金面上项目(No. 11961053);内蒙古民族大学博士科研启动基金项目(No. BS424);内蒙古民族大学科学技术研究项目(No. NMDYB17154)

第一作者简介:晁丽佳,女,研究方向为运筹学与控制论,E-mail:2687533268@qq.com;通信作者:张永富,男,副教授,博士,E-mail:zhyf88888@163.com

网络出版地址:https://link.cnki.net/urlid/50.1165.N.20230623.1114.006

解决问题(4)的方法有牛顿法、拟牛顿法、共轭梯度法等,其中共轭梯度法因具有数据储存占用空间小、迭代简单的特点被普遍应用于求解大型无约束极小化问题。共轭梯度法的一般迭代格式如下:

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \alpha_k \mathbf{d}_k, \quad (5)$$

$$\mathbf{d}_k = \begin{cases} -\mathbf{g}_k, & k=1, \\ -\mathbf{g}_k + \beta_k \mathbf{d}_{k-1}, & k \geq 2. \end{cases} \quad (6)$$

式中: \mathbf{x}_k 为第 k 次迭代点, α_k 为步长, β_k 为共轭参数, \mathbf{d}_k 为搜索方向。

经典的共轭梯度法可划分为 2 类:一类是 Fletcher 等人^[6]提出的 FR 方法、Dai 等人^[7]提出的 DY 方法以及 Fletcher^[8]提出的共轭下降法(CD),这些方法对应的共轭参数 β_k 的公式分别为: $\beta_k^{\text{FR}} = \frac{\|\mathbf{g}_k\|^2}{\|\mathbf{g}_{k-1}\|^2}$, $\beta_k^{\text{DY}} = \frac{\|\mathbf{g}_k\|^2}{\mathbf{d}_{k-1}^T \mathbf{y}_{k-1}}$,

$\beta_k^{\text{CD}} = \frac{\|\mathbf{g}_k\|^2}{-\mathbf{d}_{k-1}^T \mathbf{g}_{k-1}}$ 。文中出现的 \mathbf{y}_{k-1} 均为 $\mathbf{y}_{k-1} = \mathbf{g}_k - \mathbf{g}_{k-1}$ 。以上 3 个方法具有很好的全局收敛性质,但数值表现却欠佳^[9]。另一类方法有 Polak-Ribiere 和 Polyak^[10-11]提出的 PRP 方法、Hestenes 等人^[12]提出的 HS 方法以及 Liu 等人^[13]提出的 LS 方法,所对应的共轭参数 β_k 的公式分别为: $\beta_k^{\text{PRP}} = \frac{\mathbf{g}_k^T \mathbf{y}_{k-1}}{\|\mathbf{g}_{k-1}\|^2}$, $\beta_k^{\text{HS}} = \frac{\mathbf{g}_k^T \mathbf{y}_{k-1}}{\mathbf{d}_{k-1}^T \mathbf{y}_{k-1}}$, $\beta_k^{\text{LS}} =$

$\frac{\mathbf{g}_k^T \mathbf{y}_{k-1}}{-\mathbf{d}_{k-1}^T \mathbf{g}_{k-1}}$ 。这 3 种方法虽然数值表现优秀,但理论收敛性需要添加特定的条件才能证明^[9]。

一直以来,为寻求一种计算性能良好、理论收敛性又较为满意的方法,许多学者不断探索,对共轭参数 β_k 进行了大量研究,得到了诸多令人比较满意的结果。

21 世纪初, Bergin 等人^[14]将共轭梯度法与谱梯度法进行结合,提出谱共轭梯度法,搜索方向为:

$$\mathbf{d}_k = \begin{cases} -\mathbf{g}_k, & k=1, \\ -\theta_k \mathbf{g}_k + \beta_k \mathbf{d}_{k-1}, & k \geq 2. \end{cases} \quad (7)$$

式中: θ_k 为谱参数。特别地,当 $\theta_k \equiv 1$ 时,式(7)即为式(6)。近 10 年的研究表明,谱共轭梯度法比传统的共轭梯度法计算效果更好^[15-18],因此谱共轭梯度法成为研究者的热点研究课题之一。为了在谱方法和共轭梯度法之间找到恰当的平衡以保持 2 种算法的计算优点,从而选择不同的 θ_k 和 β_k ,以此提高谱共轭梯度法的运算效率与收敛结果^[19]。

文献[20]讨论了如下 2 种参数的谱共轭梯度法,并结合了 2 种线搜索方法证明此方法是收敛的:

$$\beta_k = \frac{\mathbf{g}_k^T \left(\mathbf{g}_k - \frac{\mathbf{g}_k^T \mathbf{d}_{k-1}}{\|\mathbf{d}_{k-1}\|^2} \mathbf{d}_{k-1} \right)}{\|\mathbf{g}_{k-1}\|^2},$$

$$\theta_k^{(1)} = \frac{\mathbf{d}_{k-1}^T \mathbf{y}_{k-1}}{\|\mathbf{g}_{k-1}\|^2}, \theta_k^{(2)} = \theta_k^{(1)} - \frac{\mathbf{g}_k^T \mathbf{d}_{k-1}}{\|\mathbf{g}_{k-1}\|^2} \cos^2 r_k. \quad (8)$$

式中: r_k 为向量 \mathbf{g}_k 与 \mathbf{d}_{k-1} 的夹角。另外,文献[20]也给出了谱共轭梯度法与 FR 共轭梯度法的数值结果的比较,得出由式(5)、(7)、(8)所构成的谱共轭梯度法更有效。

基于文献[20]给出的共轭参数和谱参数修正方法,文献[21]讨论了如下 2 种参数的谱共轭梯度法,并证明了这 2 种方法在 Wolfe 搜索下的收敛性:

$$\beta_k = \frac{\mathbf{g}_k^T \left(\mathbf{g}_k - \frac{\mathbf{g}_k^T \mathbf{g}_{k-1}}{\|\mathbf{g}_{k-1}\|^2} \mathbf{g}_{k-1} \right)}{\max\{\|\mathbf{g}_{k-1}\|^2, \mathbf{d}_{k-1}^T \mathbf{y}_{k-1}\}}, \theta_k^1 = \frac{\mathbf{d}_{k-1}^T \mathbf{y}_{k-1}}{\max\{\|\mathbf{g}_{k-1}\|^2, \mathbf{d}_{k-1}^T \mathbf{y}_{k-1}\}}, \theta_k^2 = \theta_k^1 - \frac{\mathbf{g}_k^T \mathbf{d}_{k-1}}{\max\{\|\mathbf{g}_{k-1}\|^2, \mathbf{d}_{k-1}^T \mathbf{y}_{k-1}\}} \cos^2 \varphi_k.$$

式中: φ_k 为向量 \mathbf{g}_k 与 \mathbf{g}_{k-1} 的夹角。文献[21]将提出的谱共轭梯度法与文献[20]中 2 个算法及 FR 共轭梯度法的数值结果进行比较,显示文献[21]给出的方法迭代步数最少,用时最短。

基于上述方法,并结合 6 种经典的共轭梯度算法(即 FR、DY、CD、PRP、HS 和 LS)的共轭参数 β_k 部分特点,本文提出了一种修正 FR 谱共轭梯度法,简称 XZFR 算法。

2 XZFR 算法

2.1 下降性及 XZFR 算法

本文给出的共轭参数和谱参数具体格式如下:

$$\beta_k^{\text{XZFR}} = \frac{\mathbf{g}_k^T \left(\mathbf{g}_k - \frac{\mathbf{g}_k^T \mathbf{y}_{k-1}}{\|\mathbf{y}_{k-1}\|^2} \mathbf{y}_{k-1} \right)}{\max \{ \|\mathbf{g}_{k-1}\|^2, \mathbf{d}_{k-1}^T \mathbf{y}_{k-1}, -\mathbf{g}_{k-1}^T \mathbf{d}_{k-1} \}}, \quad (9)$$

$$\theta_k = \frac{\mathbf{d}_{k-1}^T \mathbf{y}_{k-1}}{\max \{ \|\mathbf{g}_{k-1}\|^2, \mathbf{d}_{k-1}^T \mathbf{y}_{k-1}, -\mathbf{g}_{k-1}^T \mathbf{d}_{k-1} \}}. \quad (10)$$

文中采用了标准的 Wolfe 线搜索:

$$f(\mathbf{x}_k + \alpha_k \mathbf{d}_k) \leq f(\mathbf{x}_k) + \rho \alpha_k \mathbf{g}_k^T \mathbf{d}_k, \quad (11)$$

$$\mathbf{g}(\mathbf{x}_k + \alpha_k \mathbf{d}_k)^T \mathbf{d}_k \geq \sigma \mathbf{g}_k^T \mathbf{d}_k. \quad (12)$$

式中: $0 < \rho < \sigma < 1$.

算法 1 (XZFR 算法) 步骤 1, 给定初值 $\mathbf{x}_1 \in \mathbf{R}^n$, $\delta \in (0, \frac{1}{2})$, $\epsilon \geq 0$, $\mathbf{d}_1 = -\mathbf{g}_1$, $k=1$;

步骤 2, 若 $\|\mathbf{g}_k\| \leq \epsilon$, 停止; 否则计算满足式(11)、(12)的 Wolfe 线搜索准则的步长 α_k ;

步骤 3, 计算 $\mathbf{x}_{k+1} = \mathbf{x}_k + \alpha_k \mathbf{d}_k$. 若 $\|\mathbf{g}_{k+1}\| \leq \epsilon$, 停止;

步骤 4, 由式(9)、(10)分别式计算参数 β_{k+1} 及 θ_{k+1} , 由式(7)计算 \mathbf{d}_{k+1} ;

步骤 5, $k=k+1$, 转入步骤 2.

引理 1 考虑 XZFR 算法, 其中 $\{\beta_k, \theta_k, \alpha_k\}$ 分别由式(9)~(12)决定, 则有 $\mathbf{d}_k^T \mathbf{g}_k < 0$ 且 $\beta_k \geq 0$.

证明 若 $k=1, \mathbf{d}_1 = -\mathbf{g}_1$, 则有

$$\mathbf{d}_1^T \mathbf{g}_1 = -\|\mathbf{g}_1\|^2 < 0 \quad (13)$$

成立. 假设对于 $k > 1$, 有 $\mathbf{d}_{k-1}^T \mathbf{g}_{k-1} < 0$ 成立, 由式(12)得 $\mathbf{d}_{k-1}^T \mathbf{y}_{k-1} = \mathbf{d}_{k-1}^T (\mathbf{g}_k - \mathbf{g}_{k-1}) \geq (\sigma - 1) \mathbf{g}_{k-1}^T \mathbf{d}_{k-1} > 0$.

设 ψ_k 为向量 \mathbf{g}_k 与 \mathbf{y}_{k-1} 的夹角, 则 $\cos \psi_k = \frac{\mathbf{g}_k^T \mathbf{y}_{k-1}}{\|\mathbf{g}_k\| \|\mathbf{y}_{k-1}\|}$.

1) 当 $\mathbf{d}_{k-1}^T \mathbf{y}_{k-1} = \max \{ \|\mathbf{g}_{k-1}\|^2, -\mathbf{g}_{k-1}^T \mathbf{d}_{k-1}, \mathbf{d}_{k-1}^T \mathbf{y}_{k-1} \}$ 时, 对式(7)两端分别与 \mathbf{g}_k 作内积, 得:

$$\begin{aligned} \mathbf{d}_k^T \mathbf{g}_k &= -\theta_k \|\mathbf{g}_k\|^2 + \beta_k \mathbf{d}_{k-1}^T \mathbf{g}_k = -\frac{\mathbf{d}_{k-1}^T \mathbf{y}_{k-1}}{\mathbf{d}_{k-1}^T \mathbf{y}_{k-1}} \|\mathbf{g}_k\|^2 + \frac{\mathbf{g}_k^T \left(\mathbf{g}_k - \frac{\mathbf{g}_k^T \mathbf{y}_{k-1}}{\|\mathbf{y}_{k-1}\|^2} \mathbf{y}_{k-1} \right)}{\mathbf{d}_{k-1}^T \mathbf{y}_{k-1}} \mathbf{d}_{k-1}^T \mathbf{g}_k = \\ &= -\|\mathbf{g}_k\|^2 + \frac{\|\mathbf{g}_k\|^2 - \frac{(\mathbf{g}_k^T \mathbf{y}_{k-1})^2}{\|\mathbf{y}_{k-1}\|^2}}{\mathbf{d}_{k-1}^T \mathbf{y}_{k-1}} \mathbf{d}_{k-1}^T \mathbf{g}_k = -\|\mathbf{g}_k\|^2 + \frac{\|\mathbf{g}_k\|^2 (1 - \cos^2 \psi_k)}{\mathbf{d}_{k-1}^T \mathbf{y}_{k-1}} \mathbf{d}_{k-1}^T \mathbf{g}_k \leq \\ &= \frac{\|\mathbf{g}_k\|^2 (1 - \sigma \cos^2 \psi_k)}{\mathbf{d}_{k-1}^T \mathbf{y}_{k-1}} \mathbf{d}_{k-1}^T \mathbf{g}_{k-1} < 0, \end{aligned} \quad (14)$$

$$\beta_k = \frac{\mathbf{g}_k^T \left(\mathbf{g}_k - \frac{\mathbf{g}_k^T \mathbf{y}_{k-1}}{\|\mathbf{y}_{k-1}\|^2} \mathbf{y}_{k-1} \right)}{\mathbf{d}_{k-1}^T \mathbf{y}_{k-1}} = \frac{\|\mathbf{g}_k\|^2 (1 - \cos^2 \psi_k)}{\mathbf{d}_{k-1}^T \mathbf{y}_{k-1}} \geq 0.$$

2) 当 $\|\mathbf{g}_{k-1}\|^2 = \max \{ \|\mathbf{g}_{k-1}\|^2, -\mathbf{g}_{k-1}^T \mathbf{d}_{k-1}, \mathbf{d}_{k-1}^T \mathbf{y}_{k-1} \}$ 时, 有 $\frac{\|\mathbf{g}_k\|^2}{\|\mathbf{g}_{k-1}\|^2} \geq \beta_k = \frac{\|\mathbf{g}_k\|^2 (1 - \cos^2 \psi_k)}{\|\mathbf{g}_{k-1}\|^2} \geq 0$.

若 $\beta_k > 0$, 将式(7)两端分别与 \mathbf{g}_k 作内积得:

$$\begin{aligned} \mathbf{d}_k^T \mathbf{g}_k &= -\theta_k \|\mathbf{g}_k\|^2 + \beta_k \mathbf{d}_{k-1}^T \mathbf{g}_k = -\frac{\mathbf{d}_{k-1}^T \mathbf{y}_{k-1}}{\|\mathbf{g}_{k-1}\|^2} \|\mathbf{g}_k\|^2 + \beta_k \mathbf{d}_{k-1}^T \mathbf{g}_k \leq \\ &= \beta_k (-\mathbf{d}_{k-1}^T \mathbf{y}_{k-1}) + \beta_k \mathbf{g}_k^T \mathbf{d}_{k-1} = \beta_k \mathbf{g}_{k-1}^T \mathbf{d}_{k-1} < 0. \end{aligned} \quad (15)$$

若 $\beta_k = 0$, 结合 $\mathbf{d}_{k-1}^T \mathbf{y}_{k-1} > 0$ 可知, $\mathbf{d}_k^T \mathbf{g}_k = -\frac{\mathbf{d}_{k-1}^T \mathbf{y}_{k-1}}{\|\mathbf{g}_{k-1}\|^2} \|\mathbf{g}_k\|^2 < 0$.

3) 当 $-\mathbf{g}_{k-1}^T \mathbf{d}_{k-1} = \max \{ \|\mathbf{g}_{k-1}\|^2, -\mathbf{g}_{k-1}^T \mathbf{d}_{k-1}, \mathbf{d}_{k-1}^T \mathbf{y}_{k-1} \}$ 时,

$$\beta_k = \frac{\mathbf{g}_k^T \left(\mathbf{g}_k - \frac{\mathbf{g}_k^T \mathbf{y}_{k-1}}{\|\mathbf{y}_{k-1}\|^2} \mathbf{y}_{k-1} \right)}{-\mathbf{g}_{k-1}^T \mathbf{d}_{k-1}} = \frac{\|\mathbf{g}_k\|^2 (1 - \cos^2 \psi_k)}{-\mathbf{g}_{k-1}^T \mathbf{d}_{k-1}} \geq 0. \quad (16)$$

同理,有 $\frac{\|\mathbf{g}_k\|^2}{-\mathbf{g}_{k-1}^T \mathbf{d}_{k-1}} \geq \beta_k \geq 0$ 。

若 $\beta_k > 0$,则有:

$$\mathbf{d}_k^T \mathbf{g}_k = -\theta_k \|\mathbf{g}_k\|^2 + \beta_k \mathbf{d}_{k-1}^T \mathbf{g}_k = -\frac{\mathbf{d}_{k-1} \mathbf{y}_{k-1}}{-\mathbf{g}_{k-1}^T \mathbf{d}_{k-1}} \|\mathbf{g}_k\|^2 + \beta_k \mathbf{d}_{k-1}^T \mathbf{g}_k \leq \beta_k (-\mathbf{d}_{k-1}^T \mathbf{y}_{k-1}) + \beta_k \mathbf{g}_k^T \mathbf{d}_{k-1} < 0;$$

若 $\beta_k = 0$,则有 $\mathbf{d}_k^T \mathbf{g}_{k-1} = -\theta_k \|\mathbf{g}_k\|^2 = -\frac{\mathbf{d}_{k-1}^T \mathbf{y}_{k-1}}{-\mathbf{d}_{k-1}^T \mathbf{g}_{k-1}} \|\mathbf{g}_k\|^2$ 。由于 $\mathbf{d}_{k-1}^T \mathbf{g}_{k-1} < 0$ 且 $\mathbf{d}_{k-1}^T \mathbf{y}_{k-1} > 0$,因此有 $\mathbf{d}_k^T \mathbf{g}_k < 0$

成立。

由数学归纳法, $\forall k \geq 1, \mathbf{d}_k^T \mathbf{g}_k < 0$ 都成立,同时 $\beta_k \geq 0$ 成立。 证毕

定理 1 考虑 XZFR 算法,步长 α_k 满足式(11)、(12)即 Wolfe 线搜索条件,可得:

$$0 \leq \beta_k \leq \frac{\mathbf{g}_k^T \mathbf{d}_k}{\mathbf{g}_{k-1}^T \mathbf{d}_{k-1}}. \quad (17)$$

证明 前面已经证明 $\beta_k \geq 0$ 成立,下面证明式(17)的右端等式成立。

1) 当 $\mathbf{d}_{k-1}^T \mathbf{y}_{k-1} = \max\{\|\mathbf{g}_{k-1}\|^2, -\mathbf{g}_{k-1}^T \mathbf{d}_{k-1}, \mathbf{d}_{k-1}^T \mathbf{y}_{k-1}\}$ 时,有 $\beta_k = \frac{\|\mathbf{g}_k\|^2 (1 - \cos^2 \psi_k)}{\mathbf{d}_{k-1}^T \mathbf{y}_{k-1}} \leq \frac{\|\mathbf{g}_k\|^2 (1 - \sigma \cos^2 \psi_k)}{\mathbf{d}_{k-1}^T \mathbf{y}_{k-1}}$ 。

由式(14)的证明过程可得 $\beta_k \leq \frac{\mathbf{g}_k^T \mathbf{d}_k}{\mathbf{d}_{k-1}^T \mathbf{g}_{k-1}}$ 。

2) 当 $\|\mathbf{g}_{k-1}\|^2 = \max\{\|\mathbf{g}_{k-1}\|^2, -\mathbf{g}_{k-1}^T \mathbf{d}_{k-1}, \mathbf{d}_{k-1}^T \mathbf{y}_{k-1}\}$ 时,由式(15)得,引理结论成立。

3) 当 $-\mathbf{g}_{k-1}^T \mathbf{d}_{k-1} = \max\{\|\mathbf{g}_{k-1}\|^2, -\mathbf{g}_{k-1}^T \mathbf{d}_{k-1}, \mathbf{d}_{k-1}^T \mathbf{y}_{k-1}\}$ 时,由式(16)得,引理结论成立。 证毕

2.2 全局收敛性

本文需假设 H:

(H1) 水平集 $L = \{\mathbf{x} \in \mathbf{R}^n : F(\mathbf{x}) \leq F(\mathbf{x}_1)\}$ 有界;

(H2) 在水平集 $L_1 = \{\mathbf{x} \in \mathbf{R}^n : F(\mathbf{x}) \leq F(\mathbf{x}_1)\}$ 的一个邻域 U 内, $F(\mathbf{x})$ 连续可微,梯度函数 $\mathbf{g}(\mathbf{x})$ 满足 Lipschitz 条件,即存在常数 $L > 0$ 使 $\|\mathbf{g}(\mathbf{x}) - \mathbf{g}(\mathbf{y})\| \leq L \|\mathbf{x} - \mathbf{y}\|, \forall \mathbf{x}, \mathbf{y} \in U$ 。

引理 2 (Zoutendijk 条件) 设目标函数 $F(\mathbf{x})$ 满足假设 H,考虑 XZFR 算法,其中 \mathbf{d}_k 满足 $\mathbf{d}_k^T \mathbf{g}_k < 0$,步长因子 α_k 满足式(12)、(13),则有:

$$\sum_{k \geq 1} \frac{(\mathbf{g}_k^T \mathbf{d}_k)^2}{\|\mathbf{d}_k\|^2} < +\infty. \quad (18)$$

定理 2 若假设 H 成立,考虑 XZFR 算法,由 Wolfe 条件(11)、(12)生成 α_k 且引理 1 成立, θ_k 是式(10)生成的序列,则:

$$\liminf_{k \rightarrow \infty} \|\mathbf{g}_k\| = 0. \quad (19)$$

证明 若假设式(18)不成立。设存在常数 $c > 0$,使得对任意的 $k \geq 1$,有 $\|\mathbf{g}_k\| \geq c, k = 1, 2, \dots$

对式(7)移项,两端取模平方并结合式(17)可得:

$$\|\mathbf{d}_k\|^2 = \beta_k^2 \|\mathbf{d}_{k-1}\|^2 - 2\theta_k \mathbf{d}_k^T \mathbf{g}_k - \theta_k^2 \|\mathbf{g}_k\|^2 \leq \frac{(\mathbf{g}_k^T \mathbf{d}_k)^2}{(\mathbf{g}_{k-1}^T \mathbf{d}_{k-1})^2} \|\mathbf{d}_{k-1}\|^2 - 2\theta_k \mathbf{d}_k^T \mathbf{g}_k - \theta_k^2 \|\mathbf{g}_k\|^2.$$

对上式两端同时除以 $(\mathbf{g}_k^T \mathbf{d}_k)^2$,得:

$$\frac{\|\mathbf{d}_k\|^2}{(\mathbf{g}_k^T \mathbf{d}_k)^2} \leq \frac{\|\mathbf{d}_{k-1}\|^2}{(\mathbf{g}_{k-1}^T \mathbf{d}_{k-1})^2} - \frac{2\theta_k}{\mathbf{g}_k^T \mathbf{d}_k} - \frac{\theta_k^2 \|\mathbf{g}_k\|^2}{(\mathbf{g}_k^T \mathbf{d}_k)^2} = \frac{\|\mathbf{d}_{k-1}\|^2}{(\mathbf{g}_{k-1}^T \mathbf{d}_{k-1})^2} - \left(\frac{1}{\|\mathbf{g}_k\|} + \frac{\theta_k \|\mathbf{g}_k\|}{\mathbf{g}_k^T \mathbf{d}_k} \right)^2 + \frac{1}{\|\mathbf{g}_k\|^2} \leq \frac{\|\mathbf{d}_{k-1}\|^2}{(\mathbf{g}_{k-1}^T \mathbf{d}_{k-1})^2} + \frac{1}{\|\mathbf{g}_k\|^2}.$$

又因为 $\frac{\|\mathbf{d}_1\|^2}{(\mathbf{g}_1^T \mathbf{d}_1)^2} = \frac{1}{\|\mathbf{g}_1\|^2}$,由上式递推得: $\frac{\|\mathbf{d}_k\|^2}{(\mathbf{g}_k^T \mathbf{d}_k)^2} \leq \sum_{i=1}^k \frac{1}{\|\mathbf{g}_i\|^2} \leq \frac{k}{c^2}$,即 $\frac{(\mathbf{g}_k^T \mathbf{d}_k)^2}{\|\mathbf{d}_k\|^2} \geq \frac{c^2}{k}$,从而 $\sum_{k \geq 1} \frac{(\mathbf{g}_k^T \mathbf{d}_k)^2}{\|\mathbf{d}_k\|^2} = +\infty$ 。

这与引理 2 矛盾,故式(19)成立。 证毕

3 数值实验

3.1 非线性方程组

为检测新方法的有效性,将对经典 FR 算法和文献[21]中 ZFR1 算法,在文献[22]中选取算例,测试版本

为 Windows 10 专业版,Inter(R) Core(TM) i5-8250U CPU 1.60 GHz 8.00 GB 内存。文中搜索方法全部采用标准 Wolfe 线搜索,搜索参数为 $\rho=0.1, \sigma=0.9, \epsilon=10^{-6}$,终止条件为 $\|\mathbf{g}_k\| \leq \epsilon$ 。

从表 1 可以看出,与经典 FR 共轭梯度法和 ZFR1 算法相比,XZFR 算法迭代步骤更少,运算时间更短,效果更明显。

表 1 算法的数值结果

Tab. 1 The numerical results of the algorithm

测试问题	维数	FR 方法	ZFR1	XZFR
Quadratic QF2	10	90/455/0.014 645	40/153/0.004 454	39/160/0.003 297
	20	100/796/0.012 245	54/256/0.006 666	50/243/0.003 476
Generalized Tridiagonal 2	10	190/1 140/0.025 134	16/50/0.005 681	11/37/0.005 3
	150	—	18/58/0.006 298	13/42/0.005 47
Generalized Tridiagonal 1	40	61/308/0.029 753	27/109/0.009 872	26/109/0.008 271
	400	—	31/123/0.013 761	24/100/0.008 613
Extended Himmelblau	4 000	—	33/135/0.054 222	25/105/0.040 862
	10	24/126/0.006 442	20/101/0.004 699	18/97/0.003 283
	500	25/132/0.015 154	24/121/0.016 299	19/103/0.004 94
	1 000	25/132/0.019 065	24/121/0.018 243	21/114/0.005 37
	10 000	29/153/0.074 884	26/131/0.145 101	21/114/0.046 308

注:表中每个算法下的 3 组数据依次为对应算法迭代次数、目标函数迭代次数和 CPU 时间(单位:s)。

3.2 稀疏信号重构

为了测试 XZFR 算法在稀疏信号重构问题上的有效性,选取经典 FR 方法进行对比,主要测试随着原信号维数的变化,均方误差、相对误差、信噪比以及时间的变化。设定 $m = \frac{n}{2}, \mathbf{x} \in \mathbf{R}^n$ 为稀疏度 $K = \lceil 0.05m \rceil$ ($\lceil \cdot \rceil$ 表示向下取整)的原始稀疏信号, $\mathbf{A} \in \mathbf{R}^{m \times n}$ 为 Gauss 矩阵,观测向量为 $\mathbf{y} = \mathbf{Ax} + \omega$,其中 ω 为均值为 0、方差为 10^{-4} 的随机噪声。测试中各参数取值为 $\lambda = 0.01, \rho = 0.01, \sigma = 0.9, \tau_0 = 0.6$,且 $\tau_{k+1} = \tau_k$,终止条件为 $\frac{\|f(\mathbf{x}_{k+1}) - f(\mathbf{x}_k)\|}{\|f(\mathbf{x}_{k+1})\|} < 10^{-5}$ 。采用均方误差 \mathbf{x}_{MSE} 、相对误差 e 、信噪比 r_{SNR} 作为信号重构的评判标准:

$$\mathbf{x}_{\text{MSE}} = \frac{1}{n} \|\mathbf{x} - \mathbf{x}_*\|_2^2, e = \frac{\|\mathbf{x}_* - \mathbf{x}\|_2}{\|\mathbf{x}\|_2}, r_{\text{SNR}} = 10 \log_{10} \left(\frac{\|\mathbf{x}\|_2^2}{\|\mathbf{x} - \mathbf{x}_*\|_2^2} \right),$$

式中: \mathbf{x}_* 表示恢复信号。恢复结果如表 2 所示。

表 2 信号恢复结果对比

Tab. 2 Comparison of signal recovery results

$m \times n$	FR					XZFR				
	k	\mathbf{x}_{MSE}	e	r_{SNR}	CPU 时间/s	k	\mathbf{x}_{MSE}	e	r_{SNR}	CPU 时间/s
312×624	259	4.76×10^{-5}	0.042 5	27.432	4.953	150	1.77×10^{-5}	0.025 9	31.718	2.656
624×1 248	265	4.33×10^{-5}	0.033 8	29.427	19.531	171	1.54×10^{-5}	0.020 2	33.893	7.265
1 248×2 496	203	1.32×10^{-4}	0.059 0	24.587	51.375	168	1.72×10^{-5}	0.021 3	33.428	30.000
2 048×4 096	285	7.71×10^{-5}	0.045 3	26.875	200.92	165	1.37×10^{-5}	0.019 1	34.377	73.609

图 1 为 $n=2 496$ 时,FR 算法原始信号与恢复信号的对比,图 2 为相同维数下 XZFR 算法的原始信号与恢复信号的对比。图中星号与实线顶端重合表示信号恢复的效果较好。

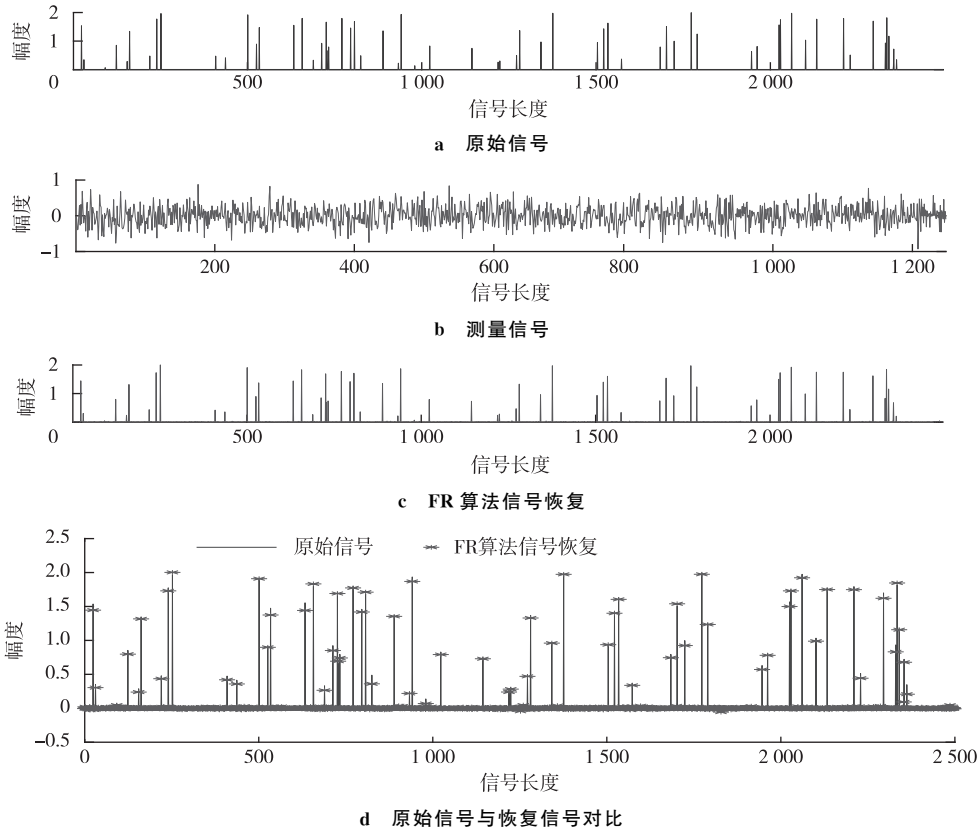


图 1 FR 算法信号恢复结果

Fig. 1 FR algorithm signal recovery results

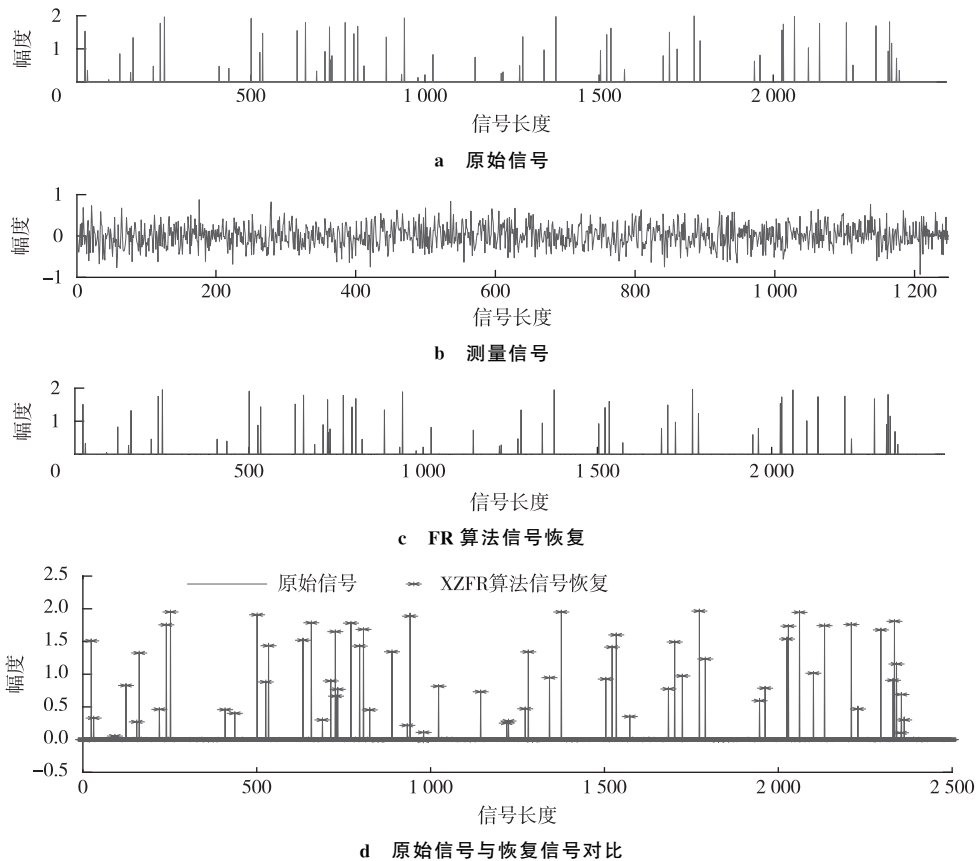


图 2 XZFR 算法信号恢复结果

Fig. 2 XZFR algorithm signal recovery results

通过对表 2 和图 2 的观察,XZFR 算法恢复信号所用时间较短,步骤较少,相对误差和均方误差较小,信噪比较大,信号恢复精度比较好,证明了将 XZFR 算法应用到稀疏信号恢复中是可行的。

参考文献:

- [1] 李双安. 共轭梯度法在大规模信号重构问题中的应用[D]. 桂林: 桂林电子科技大学, 2015.
LI S A. Application of conjugate gradient method in large scale signalreconstruction[D]. Guilin: Guilin University of Electronic Technology, 2015.
- [2] CHEN S S, DONOHO D L, SAUNDERS M A. Atomic decomposition by basis pursuit[J]. SIAM Review, 2011, 43(1): 129-159.
- [3] 郑丹青. 基于压缩感知的信号观测和重构算法研究[D]. 南京: 南京邮电大学, 2016.
ZHENG D Q. Research on signal observation and reconstruction algorithm based on compressed sensing[D]. Nanjing: Nanjing University of Posts and Telecommunications, 2016.
- [4] 周健, 张永平, 张功萱. 基于海量物联网数据的压缩感知及其并行处理[J]. 微电子学与计算机, 2012, 29(11): 116-119.
ZHOU J, ZHANG Y P, ZHANG G X. Research of parallel compressed sensing for mass data in internet of things[J]. Microelectronics & Computer, 2012, 29(11): 116-119.
- [5] 王慧敏. 共轭梯度法在信号恢复问题中的应用[D]. 呼和浩特: 内蒙古大学, 2019.
WANG H M. Application of conjugate gradient method in signal recovery problem[D]. Hohhot: Inner Mongolia University, 2019.
- [6] FLETCHER R, REEVES C M. Function minimization by conjugate gradients[J]. Computer Journal, 1964, 7: 149-154.
- [7] DAI Y H, YUAN Y. A nonlinear conjugate gradient method with a strong global convergence property[J]. SIAM Journal on Optimization, 1999, 10(1): 177-182.
- [8] FLETCHER R. Practical Methods of Optimization vol. 1: Unconstrained Optimization[J]. New York: John Wiley & Sons, 1987.
- [9] 杨雪, 李锋. 非线性优化问题中的共轭梯度法研究综述[J]. 昆明学院学报, 2021, 43(3): 59-66.
YANG X, LI F. A review of Conjugate gradient method in nonlinear optimization problems[J]. Journal of Kunming University, 2021, 43(3): 59-66.
- [10] POLAK E, RIBIERE G. Note sur la convergence de méthodes de directions conjuguées[J]. Esaim Mathematical Modelling & Numerical Analysis, 1969, 3(16): 35-43.
- [11] POLYAK B T. The conjugate gradient method in extremal problems[J]. Ussr Computational Mathematics & Mathematical Physics, 1969, 9(4): 94-112.
- [12] HESTENES M. Methods of conjugate gradients for solving linear systems[J]. Journal of Research of the National of Standards, 1952, 49.
- [13] LIU Y, STOREY C. Efficient generalized conjugate gradient algorithms, part 1[J]. Journal of Optimization Theory and Applications, 1991, 69(1): 129-137.
- [14] BIRGIN E G, MARTÍNEZ J M. A spectral conjugate gradient method for unconstrained optimization[J]. Applied Mathematics and Optimization, 2001, 43(2): 117-128.
- [15] 江羨珍, 廖伟, 简金宝, 毋晓迪. 一个带重启步的改进 PRP 型谱共轭梯度法[J]. 数学物理学报, 2022, 42(1): 216-227.
JIANG X Z, LIAO W, JIANG J B, WU X D. A modified PRP spectral Conjugate gradient method with restart step[J]. Acta Mathematica Scientia, 2022, 42(1): 216-227.
- [16] 王松华, 黎勇, 吴加其. 一种新型线搜索下的修正 3 项 LS 谱共轭梯度法[J]. 安徽大学学报(自然科学版), 2019, 43(4): 40-44.
WANG S H, LI Y, WU J Q. Modified 3-term LS spectral conjugate gradient method under a new line search[J]. Journal of Anhui University (Natural Science Edition), 2019, 43(4): 40-44.
- [17] 王森森, 张俊容, 韩信, 等. 一类具有充分下降性的混合型谱共轭梯度法[J]. 西南大学学报(自然科学版), 2017, 39(5): 139-144.
WANG S S, ZHANG J R, HAN X, et al. A class of mixed spectral conjugate gradient method with sufficient descent property [J]. Journal of Southwest University (Natural Science Edition), 2017, 39(5): 139-144.
- [18] Nezhadhossein S. A Modified Descent Spectral Conjugate Gradient Method for Unconstrained Optimization[J]. Iranian Journal of Science and Technology, Transactions A: Science, 2020: 1-12.
- [19] 胡朝明. 几类谱共轭梯度方法理论及数值行为研究[D]. 长沙: 中南大学, 2012.
HU Z M. Study on the theory and numerical behavior of several spectral conjugate gradient methods[D]. Changsha: Central South University, 2012.

- [20] 林穗华. 求解无约束优化问题的两个谱共轭梯度法的全局收敛性[J]. 重庆师范大学学报(自然科学版), 2015, 32(2): 1-6.
LIN S H. Global convergence of two spectral conjugate gradient method for unconstrained optimization problems[J]. Journal of Chongqing Normal University (Natural Science), 2015, 32(2): 1-6.
- [21] 夏丽娜, 朱志斌. Wolfe 线搜索下的两类修正 FR 谱共轭梯度法[J]. 应用数学, 2021, 34(3): 647-656.
XIA L N, ZHU Z B. Two kinds of modified FR conjugate gradient method under wolfe line search[J]. Mathematica Applicata, 2021, 34(3): 647-656.
- [22] ANDREI N. An unconstrained optimization test functions collection[J]. Advanced Modeling and Optimization, 2008, 10(1): 147-161.

Operations Research and Cybernetics

Application of Modified FR Conjugate Gradient Method in Signal Recovery

CHAO Lijia, ZHANG Yongfu

(College of Mathematics and Physics, Inner Mongolia University for Nationalities, Tongliao Inner Mongolia 028000, China)

Abstract: Based on compressive sensing theory, it aims to solve the problem of sparse signal recovery. On the basis of the analysis of the classical FR conjugate gradient method, a modified FR spectral conjugate gradient method based on the spectral gradient method is proposed. This method combines some characteristics of conjugate parameters of six classical conjugate gradient methods, such as FR, DY, and CD. Additionally, it is proved that the given method has a descending direction and global convergence under the Wolfe line search. Comparing with the numerical results of the other two algorithms, it is shown that the given optimization algorithm has some advantages. The signal reconstruction results show that the proposed method has high signal-to-noise ratio and less time-consuming, and can be effectively applied to the recovery of sparse signals.

Keywords: signal restored; spectral conjugate gradient method; global convergence; wolfe line search

(责任编辑 黄 颖)