昆明学院学报　2021，43（3）：59～66　ISSN 1674－5639　CN 53－1211/G4
Journal of Kunming University

非线性优化问题中的共轭梯度法研究综述

杨　雪，李　锋*
（云南师范大学 数学学院，云南 昆明 650500）

摘要：共轭梯度法是解决大规模线性等式和非线性优化问题的重要方法之一，其突出优点是收敛性强、迭代步骤简单、对存储要求低．因此，通过查阅大量国内外文献，介绍几种经典共轭梯度法及其改进算法的发展趋势和收敛性，并阐述混合共轭梯度法的研究热点以及共轭梯度法在图像复原和压缩感知中的应用，同时探讨了现有算法的局限性及其未来的研究方向．

关键词：混合共轭梯度法；收敛性；压缩感知；图像复原
中图分类号：O221　**文献标识码**：A　**文章编号**：1674－5639（2021）03－0059－08
DOI：10.14091/j.cnki.kmxyxb.2021.03.013

Review of Conjugate Gradient Method in Nonlinear Optimization Problems

YANG Xue，LI Feng*
（College of Mathematics，Yunnan Normal University，Kunming，Yunnan，China　650500）

Abstract：Conjugate gradient method is one of the important methods to solve large linear equation and nonlinear optimization problems. Its outstanding advantages are strong convergence，simple iterative steps and low storage requirements. Therefore，through the study of extensive domestic and foreign literature，we introduced several classical Conjugate gradient methods and the developing trend of its improving algorithm and convergence；stated the research focus of mixed conjugate gradient method and the application of conjugate gradient method in image restoration and compressed sensing. At the same time，the shortcoming of existing algorithms and possible research directions in the future were given.

Key words：mixed conjugate gradient method；convergence；compressed sensing；image restoration

优化问题与我们的工作和生活息息相关，比如每个公司都需要考虑“在一定成本下，如何使利润最大化或者损失最小化”的问题，此时，可以利用最优化方法去解决类似的实际优化问题．而优化问题存在有约束条件和无约束条件两种情况，也可将其称之为有约束优化和无约束优化．在处理实际问题时，通常会将有约束优化问题转变为无约束优化问题求解．因此，无约束优化一直是优化问题的研究重点．对于无约束优化问题，若目标函数连续可微，则可以利用最速下降法、牛顿法、拟牛顿法、共轭梯度法等方法对函数进行求解．这些方法都是利用目标函数的导数信息，然后采取一定的迭代格式求出目标函数的最优值[1]．

在上述方法中，最简单的方法是最速下降法，但其面对一些复杂的问题时会受到锯齿现象的影响导致收敛速度很慢．牛顿法和拟牛顿法的优势是收敛速度快，但由于其需要计算和储存矩阵且搜索方向的求解复杂，为大规模优化问题的求解增加了难度[2]．而共轭梯度法在计算时仅需求目标函数的一阶导数，且与最速下降法相比提高了收敛速度，同时又避免了牛顿法对存储空间的要求和求解 Hesse 矩阵及其逆矩阵的缺点．基于上述原因，在处理实际问题时共轭梯度法受到人们的广泛喜爱．

收稿日期：2020－07－08
作者简介：杨雪（1996—），女，云南元江人，硕士研究生，主要从事最优化理论与算法研究．
*通讯作者：李锋（1964—），男，云南红河人，教授，硕士，主要从事最优化理论与算法研究，E-mail：lf_2364@126.com.

随着大数据时代的来临，亟须对海量数据进行分析应用，于是对大规模优化问题的求解提出新的要求．而具有收敛性强，且算法简单、易操作的共轭梯度法恰恰是大规模问题的重要求解方法之一[2]．因此，为了处理某些优化问题，研究共轭梯度法具有重要意义．

1 共轭梯度法的发展及其收敛性

1.1 共轭梯度法的发展

对于无约束优化问题： $\min\{f(x)\mid x\in \boldsymbol{R}^n\}$，

其中 f：$R^n\to R$ 连续可微，$g_k=\nabla f(x_k)$．给定初始值后，共轭梯度法的迭代公式为：

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

1952 年 Hestenes 等[3]提出了 HS 共轭梯度法，该方法不需要求逆矩阵或矩阵分解，而是以迭代的方式求线性方程的解，且结果具有有限终止性．对于一个线性方程组

$$\boldsymbol{A}x=b,\ x\in \boldsymbol{R}^n,\tag{2}$$

其中，$\boldsymbol{A}$ 是对称正定矩阵，且 $\boldsymbol{A}\in \boldsymbol{R}^{m\times n}$，$b\in \boldsymbol{R}^{\mathrm{m}}$．

可将（2）式看作无约束严格凸二次函数的极小值问题 $\min f(x)=1/2x^TAx-b^Tx$ 的驻点条件（梯度等于 0），即 $\nabla f(x)=\boldsymbol{A}x-b=0$．

在解决了线性方程组的问题之后，如何将该方法推广到非线性领域值得关注．于是 Fletcher 等[4]将 HS 法的思想运用到非线性优化问题，并提出了 FR 法，FR 法可先在精确线搜索下求解凸二次函数最小值问题，之后他们又将其推广到一般函数最小值的求解中[1]；1969 年 Polak 等[5]和 Polyak[6]提出了 PRP 方法，该方法通过对 FR 方法中 β_k 的改动，改善了 FR 方法的收敛速度慢的情况；1987 年 Fletcher[7]在其著作中引入了 CD 法，即共轭下降法，从而得到了关于下降方向更好的性质；1991 年 Liu 等[8]在 PRP 方法的基础上提出了 LS 法；1999 年戴彧虹等[9]提出了在 Wolfe 线搜索条件下的 DY 方法；2001 年戴彧虹等[10]提了 DL 法，该方法利用拟牛顿方程及 HS 方法的共轭性使得新算法在收敛性和下降方向上均有所改善；2004 年 Yabe 等[11]根据 DL 方法的思想得到了 YT 方法．

不同的共轭梯度法区别在于 β_k 不同，现将以上方法中的 β_k 计算公式归纳如下：

$$\beta_k^{HS}=\frac{g_k^Ty_k}{d_{k-1}^Ty_k},\ \beta_k^{FR}=\frac{g_{k+1}^Tg_{k+1}}{g_k^Tg_k};\tag{3}$$

$$\beta_k^{PRP}=\frac{g_k^Ty_k}{g_{k-1}^Tg_{k-1}},\ \beta_k^{CD}=\frac{g_k^Tg_k}{-d_{k-1}^Tg_{k-1}};\tag{4}$$

$$\beta_k^{LS}=\frac{g_k^Ty_k}{-d_{k-1}^Tg_k},\ \beta_k^{DY}=\frac{g_k^Tg_k}{d_{k-1}^Ty_{k-1}};\tag{5}$$

$$\beta_k^{DL}=\frac{g_k^T(y_k-ts_k)}{d_{k-1}^Ty_k},\ \text{其中 } t\geqslant 0;\tag{6}$$

$$\beta_k^{YT}=\frac{g_k^T(z_k-ts_k)}{d_{k-1}^Tz_k}.\tag{7}$$

其中 $s_k=x_{k+1}-x_k$，$y_k=g_k-g_{k-1}$，$z_k=y_k+\rho(\frac{\theta_k}{s_k^T\mu_k}u_k)$．

1.2 共轭梯度法的收敛性

1.2.1 FR 方法

起初对 FR 方法的收敛性分析工作是在精确线搜索基础上展开，Zoutendijk[12]给出了精确线搜索下 FR 方法对一般非凸函数全局收敛的证明，之后，Powell[13]给出 FR 方法全局有效性，同时证明了精确线搜索下的 FR 方法若在某一步产生一个小步长则后面生成的步长也可能很小．因此，FR 方法收敛可能很慢．

此外，Powell[13] 的分析也对 FR 方法数值表现不佳提供了一些依据.

由于精确线搜索计算量较大，实际计算通常采用非精确搜索确定迭代步长. 1985 年，Al-Baali[14] 首先分析了 FR 方法在非精确搜索下的收敛性，且证明当参数满足 $0<\rho<\sigma<1/2$；1/2 时，FR 方法在强 Wolfe 线搜索下满足充分下降性及全局收敛性；Liu 等[15] 继续研究了 FR 方法在强 Wolfe 线搜索下的收敛性，证明了当 $\sigma=1/2$ 它也是充分下降且全局收敛的；1996 年戴彧虹等[16] 给出了 FR 方法在 $\sigma=1/2$ 的强 Wolfe 线搜索时全局收敛的新证明，并还举出一个反例证明：当 $\sigma=1/2$ 时 FR 方法可能因为产生上升方向而导致算法失败；此外戴彧虹等[17] 还提出新的广义线搜索，并给出 FR 方法在广义线搜索下的全局收敛性；之后，戴彧虹等[18] 在没有充分下降条件的情况下建立了 FR 方法和 PRP 方法的收敛性.

1.2.2　DY 方法

由于前人对 FR 方法和 CD 方法的收敛性研究大多停留在精确线搜索或强 Wolfe 线搜索上，如何在较弱的线搜索条件下提出一种具有全局收敛性的共轭梯度法显得尤为重要. 于是戴彧虹等[9] 对原来的方法进行深入研究后，提出一种在 Wolfe 线搜索下能自动保持下降条件且全局收敛的共轭梯度法，并将其称之为 DY 方法. 之后，戴彧虹[19] 分析了 DY 方法的内在性质，给出了该方法在远离最优点时大部分迭代点列都满足充分下降条件的证明.

1.2.3　PRP 方法

与其他共轭梯度法相比，PRP 方法具有良好的数值表现. 对于普通非线性函数，当出现连续产生小步长的情况时，共轭参数 $\beta_k^{PR}\approx 0$，这相当于对算法进行了一次重新启动，克服了 FR 方法数值表现不好的缺点. 但对于一般非凸函数，PRP 方法的收敛性表现不佳[20]. 为了改善这种情况，Powell[21] 建议令 PRP 方法中的共轭参数为：

$$\beta_k=\max\{\beta_k^{PRP},\ 0\}. \tag{8}$$

之后，Gilbert 等[22] 根据 Powell[21] 的建议，在一定条件下给出了一般非凸函数在改进 PRP 方法下的收敛性. 2006 年 Wei 等[23] 给出了 PRP 方法的一种变体 β_k^*，并称为 WYL 方法，同时给出了新算法在几种线搜索下的收敛性. 其中：

$$\beta_k^*=\frac{g_k^T\bar{y}_k}{\|g_{k-1}\|^2},\quad \bar{y}_k=g_k-\frac{\|g_k\|}{\|g_{k-1}\|}g_{k-1}. \tag{9}$$

延续 WYL 方法的思想，2007 年 Yao 等[24] 将这种方法推广到 HS 和 LS 方法上，并提出新的算法，其形式如下：

$$\beta_k^{MHS}=\frac{g_k^T\bar{y}_k}{d_{k-1}^Ty_k},\quad \beta_k^{MLS}=\frac{g_k^T\bar{y}_k}{-d_{k-1}^Tg_{k-1}}. \tag{10}$$

2009 年 Zhang[25] 对 WYL 方法进行改进，提出一种收敛性更好的 NPRP 算法，新算法同时具有 PRP 方法的数值表现和 FR 方法的收敛性. 观察两种方法的共轭参数易知，作者将 WYL 方法中的 $g_k^Tg_{k-1}$ 替换成 $|g_k^Tg_{k-1}|$，且严格凸的目标函数在精确线搜索下 NPRP 法可简化为 FR 方法. 此外，证明了当参数 σ 在强 Wolfe 条件中限制在（0，1/2）时，NPRP 方法的充分下降性和在强 Wolfe 条件下全局收敛于非凸极小化. 出于相同的动机，她将结果推广到 HS 方法，并将改进的共轭梯度法称为 NHS 法. 其中：

$$\beta_k^{NPRP}=\frac{\|g_k\|^2-\dfrac{\|g_k\|}{\|g_{k-1}\|}|g_k^Tg_{k-1}|}{\|g_{k-1}\|^2}; \tag{11}$$

$$\beta_k^{NHS}=\frac{\|g_k\|^2-\dfrac{\|g_k\|}{\|g_{k-1}\|}|g_k^Tg_{k-1}|}{d_{k-1}^Ty_{k-1}}. \tag{12}$$

2012 年 Dai 等[26] 受到文献［27］提出的新算法能够产生充分下降条件的启发下，对 Zhang[25] 的 NPRP 方法进行了改进，使改进后的 NPRP 方法在下降方向上具有更好的性能，并将该方法称为 DPRP 法. 因此，DPRP 方法和 NPRP 方法相比较，不仅对任意线搜索具有充分的下降性，而且在 Wolfe 线搜索或

Armijo 线搜索下非凸极小化都保持全局收敛性．此外，将这一结果推广到 NHS 方法，并称为 DHS 方法．其中：

$$\beta_k^{DPRP}=\frac{\|g_k\|^2-\frac{\|g_k\|}{\|g_{k-1}\|}|g_k^Tg_{k-1}|}{\mu|g_k^Td_{k-1}|+\|g_{k-1}\|^2},\ \mu>1; \tag{13}$$

$$\beta_k^{DHS}=\frac{\|g_k\|^2-\frac{\|g_k\|}{\|g_{k-1}\|}|g_k^Tg_{k-1}|}{\mu|g_k^Td_{k-1}|+d_{k-1}^Ty_{k-1}},\ \mu>1. \tag{14}$$

2 共轭梯度法的研究现状及热点

目前，对于共轭梯度法的研究主要有两种思路：一种是直接改进共轭参数 β_k；另一种是利用混合方法能够取长补短的特点将不同的方法进行混合．众所周知，证明 FR 法、DY 法和 CD 法的收敛性时，全局收敛定理只需要满足 Lipschitz 假设，而不需要有界性假设．这决定了它们虽然计算效果一般，却具有很好的全局收敛性．另外，由于 PRP 法、HS 法和 LS 法能够自动调节 β_k 避开干扰步长，具有内置重新启动功能，因此在一般情形下可能不收敛，但是它们却具有良好的数值表现．正是因为混合共轭梯度法具有结合两类方法优缺点的特点，所以从混合共轭梯度法的思路出发构造出一种既具有好的收敛性质，又具有优秀的数值表现的新算法具有深远的意义．

混合共轭梯度法一般是将其他共轭梯度法与经典共轭梯度法混合．1990 年 Touati-Ahmed 等[28]首先通过将 FR 方法和 PRP 方法投影，引入了混合共轭梯度法．这是对混合共轭梯度法的一次有效尝试，新方法兼具 FR 方法和 PRP 方法的特点，并克服了 FR 方法数值表现一般的缺点，其中：

$$\beta_k=\max\{0,\ \min\{\beta_k^{PRP},\ \beta_k^{FR}\}\}, \tag{15}$$

为扩展 PRP 更新参数的允许选择范围，同时保持全局收敛性，Gilbert 等[22]提出：

$$\beta_k=\max\{-\beta_k^{FR},\ \min\{\beta_k^{PRP},\ \beta_k^{FR}\}\}. \tag{16}$$

该算法的数值表现虽然比 FR 方法好，但却不如 PRP 方法．由于 DY 方法具有比 FR 方法更好的全局收敛性，因此戴彧虹等[9]沿着这种思路给出一种混合 DY 方法和 HS 方法的新算法：

$$\beta_k=\max\{0,\ \min\{\beta_k^{HS},\ \beta_k^{DY}\}\}, \tag{17}$$

而且其以大量的数值实验证明该算法明显优于 HS 方法和 DY 方法，特别在面对较为复杂的问题时数值表现也比 PRP 方法好．由于受文献［22］的启发，何晓旭等[29]在 2016 年提出一种将 FR 和 PRP 算法混合的新算法．该算法充分利用二者的优势，下降方向 d_k 由新算法产生且不依赖于任何一种线搜索条件．其中：

$$\beta_k^*=\max\{\min[-c\beta_k^{PRP},\ \beta_k^{FR}],\ \min[\beta_k^{FR},\ \beta_k^{PRP}]\}; \tag{18}$$

$$c=\frac{1-\gamma}{1+\gamma},\ \gamma\in\left[\frac{1}{2},\ 1\right]; \tag{19}$$

$$d^{(k)}=\begin{cases}-g_k,\ k=0,\\ \theta_kg_k+\beta_k^*d^{(k-1)},\ k\geqslant 1,\end{cases}\quad \theta_k=1+\beta_k^*\frac{d_{k-1}^Tg_k}{\|g_k\|^2}. \tag{20}$$

为了延续 DPRP 法[26]、DY 法[9]等多种共轭梯度法的优点，韩信等[30]在 2017 年提出了一种新 HZW 法，其在发表的文章中不仅给出了新算法良好的收敛性证明，而且用文献［31］中的 11 个测试函数对它的数值表现进行了验证，其中：

$$\beta_k^{HZW}=\frac{\|g_k\|^2-\frac{\|g_k\|}{\|g_{k-1}\|}\max\{0,\ g_k^Tg_{k-1}\}}{\max\{\mu|g_k^Td_{k-1}|+\|g_{k-1}\|^2d_{k-1}^Ty_{k-1}\}},\ \mu>1. \tag{21}$$

除此之外，刘玲等[32]在 2016 年提出了一种新的带参数的共轭梯度法，并给出了新算法全局收敛性证明．新算法最大的特点是搜索方向满足充分下降性却不依赖任何线搜索，但引入参数的选取会对最后的数值效果造成一定的影响．其中 β_k 为：

$$\beta_k^{MNSE}=\frac{\mu\left|g_k^T y_{k-1}\right|}{\left|d_{k-1}^T g_k\right|-\lambda d_{k-1}^T g_{k-1}},\ 0<\mu\leqslant 1,\ 0<\lambda\leqslant 1; \tag{22}$$

$$d_k=\begin{cases}-g_k,\ k=0,\\ -\left(1+\beta_k^{MNEW}\dfrac{g_k^T d_{k-1}}{\|g_k\|^2}\right)g_k+\beta_k^{MNEW}d_{k-1},\ k\geqslant 1.\end{cases} \tag{23}$$

与此同时，利用两种方法的凸组合构造新算法的尝试也在进行．Andrei[33]给出了一种新的混合共轭梯度算法，其中：

$$\beta_k^C=(1-\theta_k)\beta_k^{HS}+\theta_k\beta_k^{DY},\ \theta_k\in[0,1], \tag{24}$$

对于这种方法，一个必须解决的问题就是凸组合参数 θ_k 确定．在文献［33］中，作者先假设搜索方向 d_k 为牛顿方向，据此得出 θ_k 的表达式为：

$$\theta_k=\frac{s_k^T\nabla^2 f(x_{k+1})g_{k+1}-s_k^T g_{k+1}-\dfrac{y_k^T g_{k+1}}{y_k^T s_k}s_k^T\nabla^2 f(x_{k+1})s_k}{\left[\dfrac{g_{k+1}^T g_{k+1}}{y_k^T s_k}-\dfrac{y_k^T g_{k+1}}{y_k^T s_k}\right]s_k^T\nabla^2 f(x_{k+1})s_k}, \tag{25}$$

由于表达式中涉及 Hessian 矩阵的计算，因此假设序列（s_k，y_k）满足割线方程 $\nabla^2 f(x_{k+1})s_k=y_k$，其中，$s_k=x_{k+1}-x_k$，$y_k=g_{k+1}-g_k$，经过代数运算后可得到：

$$\theta_k=-\frac{s_k^T g_{k+1}}{g_k^T g_{k+1}}. \tag{26}$$

类似也可以得到 d_k 的表达式，最终得到的 d_k 在一定条件下为近似牛顿方向．但这种方法的局限性在于不能保证生成的 d_k 一定是下降方法，因此可能导致数值结果不稳定．除此之外，作者还给出了新算法在一定条件下的收敛性证明以及数值实验表现．

Liu 等[34]在 Andrei[33]的启发下，综合考虑 LS 方法数值表现及 DY 方法的收敛性，提出了基于 LS 和 DY 方法凸组合的一种新的混合共轭梯度法，该方法下的搜索方向满足 Dai 等[10]提出的 D-L 共轭条件：

$$d_{k+1}^T y_k=-ts_k^T g_{k+1},\ t\geqslant 0. \tag{27}$$

特别值得注意的一点是，Liu 等[34]借鉴文献［33］证明思路发现了一个事实．即当 D-L 共轭条件的参数 $t=1$ 时，满足 D-L 共轭条件的搜索方向中的凸组合参数 θ_k^{DL} 和搜索方向符合牛顿方向时的凸组合参数 θ_k^{NT} 相等．因此取 $t=1$ 时得到的搜索方向 d_k 就能具有很好的性质，这为后面研究新算法又提供了新思路．

2019 年 Mtagulwa 等[35]借鉴 Alhawarat[36]和文献［34］的思想，构造了 NPRP 与 FR 方法的凸组合的新混合共轭梯度法．该方法继承了 PRP、FR 和 NPRP 共轭梯度的特点，并将 β_k 更新为：

$$\beta_k^{EPF}=\begin{cases}\beta_k^{PRP},\ \text{if}\ \|g_k\|^2>\left|g_k^T g_{k-1}\right|,\\ \beta_k^{CCM},\ \text{otherwise}.\end{cases} \tag{28}$$

其中 $\beta_k^{CCM}=(1-\theta_k)\beta_k^{NPRP}+\theta_k\beta_k^{FR}$，$\theta_k\in[0,1]$，$\beta_k^{NPRP}$ 由文献［25］给出，d_k 的定义为：

$$d^{(k)}=\begin{cases}-g_k,\ k=0,\\ -g_k-\beta_k^{EPF}\dfrac{d_{k-1}^T g_k}{\|g_k\|^2}g_{k-1}+\beta_k^{EPF}d_{k-1},\ k\geqslant 1.\end{cases} \tag{29}$$

此外，关于凸组合的其他混合共轭梯度法还可以参见 Djordjevic[37-38]提出的 LS 和 CD 方法凸组合的混合共轭梯度法及 LS 和 FR 方法凸组合的混合共轭梯度法．

目前，对混合共轭梯度法研究的方法已有很多，但对不同的混合方法而言，其优缺点、原理、算法特点以及收敛性特征等仍存在差异．因此对混合共轭梯度法进行深入研究仍具有重要意义，在未来研究中争取探索出更多兼具好的收敛性质及良好数值表现的混合共轭梯度法．

3　共轭梯度法的应用

由于共轭梯度法能够高效、简洁地处理一些大规模线性等式和非线性优化问题，在处理实际问题时，

其应用范围广泛，如：图像复原、压缩感知问题和求解大规模非线性互补问题等.

3.1 共轭梯度法在图像复原中的应用

图像是获取信息的重要方式之一，为高效、快捷、精准地从图像中获取重要信息，对图像复原进行研究具有重要应用价值. 图像复原的基本任务是最大限度地保留原始图像的边缘信息和去除污染图像中噪声及模糊部分. 对污染图像的处理时，要先根据不同的污染情况建立相应的原始图像和污染图像的数学模型，然后利用污染图像信息，对建立好的数学模型进行反解[39]. 因此图像复原本质上可以看成是数学中反问题的研究范畴. 在得到一个数学模型后，可以先对最小化模型进行离散处理，然后再利用最优化方法对离散后的模型进行求解. 而离散模型其中的一个子问题需要对大规模线性方程组进行求解，此时就可以利用共轭梯度法来实现.

2020 年 Yuan 等[40]将最速下降算法和 LS 方法进行凸组合，提出了一种修正的共轭算法. 新算法的搜索方向在没有任何线搜索技术的情况下，具有充分下降性和信赖域性质，在一定的条件及回溯线搜索技术下建立了全局收敛性[41]. 此外，为了改善目标函数值会随迭代次数增加而缩减的情况，作者对步长采用了加速系统[42]，并将该算法成功的应用到图像复原中，而且数值实验表明，该算法与其他方法相比具有一定的优势. 同年，Cao 等[43]提出了一种修正 PRP 共轭梯度法，由新算法定义的搜索方向具有充分下降性和信赖域特征. 由于这两个特别好的特性，因此能够更简单地证明非凸函数在 Armijo 线搜索下的全局收敛性. 同时作者还用数值实验证明了在图像复原背景下新算法同 PRP 方法相比具有较大优越性.

3.2 共轭梯度法在压缩感知中的应用

2006 年 Donoho[44]和 Candes 等[45]的两篇突破性的论文使压缩感知进入了大众视野. 随着现代科学技术的迅猛发展，压缩感知方法已经广泛应用在电子工程、医学、计算机科学等领域. 它利用一个高度不完备的线性测量对稀疏信号进行采样，这时获取的是被“压缩”后的数据. 然后，通过使用高效的方法对“压缩”的数据进行重构就能恢复原稀疏信号[46]. 在进行第 2 步重构时，就可利用共轭梯度法恢复原始信号.

在处理信号恢复模型时，由于 l_1-范数是非光滑凸函数，给 l_1-范数最小化问题的求解带来了一定难度. 李双安[47]利用 Nesterov 光滑技术把 l_1-范数近似为一个光滑函数，原问题就转变成求解光滑无约束凸规划问题，然后利用改进的 HS 算法求解该问题，从而达到大规模稀疏信号恢复的目的.

与前述类似，王慧敏等[48]为解决 l_0-范数非凸非连续无法用凸优化算法求解的问题，基于 Gauss 函数提出一个新的光滑函数逼近 l_0-范数. 因此原模型就转化成可用凸优化算法求解的凸优化问题，作者在研究中选择 PRP 法完成了信号恢复过程.

4 小结

共轭梯度法是解决大规模线性等式和非线性优化问题的重要方法之一，其突出优点是收敛性强、迭代步骤简单、对存储要求低. 本文分别介绍了几种经典共轭梯度法及其改进算法的发展变化和收敛性，然后描述了混合共轭梯度法的研究现状以及共轭梯度法在图像复原和压缩感知中的应用.

现有的共轭梯度法在实践中取得了优异的表现，但部分算法尚存在易受参数影响、适用于特定的函数、可能需要在特定条件下证明收敛性等局限性，因此凸组合法中凸组合参数的选取、对新共轭梯度法的优化以及在更弱的搜索条件下证明算法的收敛性等问题有待于后期进一步的研究和完善. 除此之外，结合谱方法的谱共轭梯度法中谱参数的选择、三项混合共轭梯度法等也值得今后继续挖掘和拓展.

[参考文献]

[1] 陈宝林. 最优化理论与算法 [M]. 北京：清华大学出版社，2005.

[2] 戴彧虹，袁亚湘. 非线性共轭梯度法 [M]. 上海：上海科学技术出版社，2000.

[3] HESTENES M R，STIEFEL E L. Methods of conjugate gradients for solving linear systems [J]. Journal of Research of the National Bureau of Standards，1952，49 (6)：409 - 436.

[4] FLETCHER R，REEVES C M. Function minimization by conjugate gradients [J]. The Computer Journal，1964，7 (2)：

149 – 154.

[5] POLAK E, RIBIERE G. Note sur la convergence de méthodes de directions conjuguées [J]. Rev Francaise Informat Recherche Opertionelle, 1969, 16 (3): 35 – 43.

[6] POLYAK B T. The conjugate gradient methods in extreme problems [J]. USSR Computational Mathematics and Mathematical Physics, 1969, 9 (4): 94 – 112.

[7] FLETCHER R. Practical methods of optimization: Vol 1 [M]. New York: John Wiley and Sons, 1987.

[8] LIU Y, STOREY C. Efficient generalized conjugate gradient algorithms: Part 1 [J]. Optimization Theory and Applications, 1991, 69 (1): 129 – 137.

[9] DAI Y H, YUAN Y X. A nonlinear conjugate gradient method with a strong global convergence property [J]. SIAM Journal on Optimization, 1999, 10 (1): 177 – 182.

[10] DAI Y H, LIAO L Z. New conjugacy conditions and related nonlinear conjugate gradient methods [J]. Applied Mathematics and Optimization, 2001, 43 (1): 87 – 101.

[11] YABE H, TAKANO M. Global convergence properties of nonlinear conjugate gradient methods with modified secant condition [J]. Computational Optimization & Applications, 2004, 28 (2): 203 – 225.

[12] ZOUTENDIJK G. Nonlinear programming, computational methods [J]. Integer and Nonlinear Programming, 1970, 143: 37 – 86.

[13] POWELL M J. Restart procedures for the conjugate gradient method [J]. Mathematical Programming, 1977, 12 (1): 241 – 254.

[14] AL-BAALI M. Descent property and global convergence of the fletcher: reeves method with inexact line search [J]. IMA Journal of Numerical Analysis, 1985, 5 (1): 121 – 124.

[15] LIU G H, HAN J Y, YIN H X. Global convergence of the fletcher-reeves algorithm with inexact line search [J]. Applied Mathematics (A Journal of Chinese Universities), 1995, 10 (1): 75 – 82.

[16] DAI Y H, YUAN Y. Convergence properties of the fletcher-reeves method [J]. IMA Journal of Numerical Analysis, 1996, 16 (2): 155 – 164.

[17] DAI Y H, YUAN Y. Convergence of the fletcher-reeves method under a generalized wolfe search [J]. Numerical Mathematics a Journal of Chinese Universities, 1996 (2): 142 – 148.

[18] DAI Y H, HAN J, LIU G, et al. Convergence properties of nonlinear conjugate gradient methods [J]. SIAM Journal on Optimization, 2000, 10 (2): 345 – 358.

[19] DAI Y H. New properties of a nonlinear conjugate gradient method [J]. Numerische Mathematik, 2001, 89 (1): 83 – 98.

[20] POWELL M J D. Nonconvex minimization calculations and the conjugate gradient method [J]. Lecture Notes in Mathematics, 1984, 106 (1): 122 – 141.

[21] POWELL M J D. Convergence properties of algorithms for nonlinear optimization [J]. SIAM Review, 1986, 28 (4): 487 – 500.

[22] GILBERT J C, NOCEDAL J. Globally convergence properties of conjugate gradient methods for optimization [J]. SIAM Journal Optimization, 1992, 2 (1): 21 – 42.

[23] WEI Z, YAO S, LIU L. The convergence properties of some new conjugate gradient methods [J]. Applied Mathematics and Computation, 2006, 183 (2): 1341 – 1350.

[24] YAO S W, WEI Z X, HUANG H. A note about WYL's conjugate gradient method and its applications [J]. Applied Mathematics and Computation, 2007, 191 (2): 381 – 388.

[25] ZHANG L. An improved Wei-Yao-Liu nonlinear conjugate gradient method for optimization computation [J]. Applied Mathematics and Computation, 2009, 215 (6): 2269 – 2274.

[26] DAI Z, WEN F. Another improved Wei-Yao-Liu nonlinear conjugate gradient method with sufficient descent property [J]. Applied Mathematics and Computation, 2012, 218 (14): 7421 – 7430.

[27] WEI Z, LI G, QI L. New nonlinear conjugate gradient formulas for large-scale unconstrained optimization problems [J]. Applied Mathematics and Computation, 2006, 179 (2): 407 – 430.

[28] TOUATI-AHMED D, STOREY C. Efficient hybrid conjugate gradient techniques [J]. Journal of Optimization Theory and Applications, 1990, 64 (2): 379 – 397.

[29] 何晓旭，殷守林，赵志刚. 一种新的无优化约束问题的混合 FR 和 PRP 共轭梯度算法 [J]. 沈阳师范大学学报（自然科学版），2016，34 (1): 92 – 95.

[30] 韩信. 求解无约束优化问题的混合共轭梯度算法 [D]. 重庆：西南大学，2017.

[31] ANDREI N. An unconstrained optimization test functions collection [J]. Advanced Modeling and Optimization，2008，10 (1)：147 – 161.

[32] 刘玲，田志远，商春雷. 带参数的混合共轭梯度算法及其收敛性研究 [J]. 青岛大学学报 (自然科学版)，2016，29 (3)：16 – 20，24.

[33] ANDREI N. Another hybrid conjugate gradient algorithm for unconstrained optimization [J]. Numerical Algorithms，2008，47 (2)：143 – 156.

[34] LIU J K，LI S J. New hybrid conjugate gradient method for unconstrained optimization [J]. Applied Mathematics & Computation，2014，245 (5)：36 – 43.

[35] MTAGULWA P，KAELO P. An efficient modified PRP-FR hybrid conjugate gradient method for solving unconstrained optimization problems [J]. Applied Numerical Mathematics，2019，45 (1)：111 – 120.

[36] ALHAWARAT A，MAMAT M，RIVAIE M，et al. An efficient hybrid conjugate gradient method with the strong wolfe-powell line search [J]. Mathematical Problems in Engineering，2015，2015 (13)：1 – 7.

[37] DJORDJEVI S S. New hybrid conjugate gradient method as a convex combination of LS and CD methods [J]. Filomat，2017，31 (6)：1813 – 1825.

[38] DJORDJEVI S S. New hybrid conjugate gradient method as a convex combination of LS and FR methods [J]. Acta MathematicaScientia，2019，39 (1)：216 – 230.

[39] 黄丽丽. 图像复原中若干问题的正则化模型与算法 [D]. 南京：南京理工大学，2013.

[40] YUAN G L，LI T T，HU W J. A conjugate gradient algorithm for large-scale nonlinear equations and image restoration problems [J]. Applied Numerical Mathematics，2020，147：129 – 141.

[41] LI X，RUAN Q. A modified PRP conjugate gradient algorithm with trust region for optimization problems [J]. Numerical Functional Analysis & Optimization，2011，32 (5)：496 – 506.

[42] YUAN G，LU X. A new backtracking inexact BFGS method for symmetric nonlinear equations [J]. Computers & Mathematics with Applications，2008，55 (1)：116 – 129.

[43] CAO J Y，WU J Z. A conjugate gradient algorithm and its applications in image restoration [J]. Journal of Inequalities and Applications，2020，247：12 – 25.

[44] DONOHO D L. Compressed sensing [J]. IEEE Transactions on Information Theory，2006，52 (4)：1289 – 1306.

[45] CANDES E J，ROMBERG J，TAO T. Robust uncertainty principles：Exact signal reconstruction from highly incomplete Fourier information [J]. IEEE Transactions on Information Theory，2006，52 (2)：489 – 509.

[46] 龙忱. 稀疏信号恢复中的 SMV 与 MMV 问题研究 [D]. 长沙：国防科学技术大学，2015.

[47] 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.

[48] 王慧敏，乌彩英. PRP 共轭梯度法在信号恢复问题中的应用 [J]. 内蒙古大学学报 (自然科学版)，2019，50 (2)：126 – 132.