浙江理工大学学报，第 49 卷，第 2 期，2023 年 3 月
Journal of Zhejiang Sci-Tech University
DOI：10.3969/j.issn.1673-3851(n).2023.02.015

一种具有充分下降性的修正 DL 型谱共轭梯度法

李亚敏
（河南开封科技传媒学院经济学院，河南开封 475001）

摘　要：提出了一种大规模无约束优化问题的求解方法，通过修正 Dai-Liao(DL)共轭梯度法的共轭参数和谱共轭梯度法的谱参数，构造了一种修正 DL 型谱共轭梯度法。所选取的谱参数使得每次迭代都自动产生一个不依赖于任何线搜索的下降方向；在常规假设下，利用强 Wolfe 线搜索证明了此方法对一致凸函数是全局收敛的。

关键词：无约束优化；强 Wolfe 线搜索；谱共轭梯度法；谱参数；全局收敛

中图分类号：O221.2　　**文献标志码：**A　　**文章编号：**1673-3851 (2023) 03-0279-06

引文格式：李亚敏．一种具有充分下降性的修正 DL 型谱共轭梯度法[J]．浙江理工大学学报(自然科学)，2023，49(2)：279-284.

Reference Format：LI Yamin. A modified DL-type spectral conjugate gradient method with sufficiently descent property [J]. Journal of Zhejiang Sci-Tech University，2023，49(2)：279-284.

A modified DL-type spectral conjugate gradient method with sufficiently descent property

LI Yamin
(School of Economics，Technology & Media University of Henan Kaifeng，Kaifeng 475001，China)

Abstract：A method for solving large-scale unconstrained optimization problems is proposed. By modifying the conjugate parameter of the Dai-Liao (DL) conjugate gradient method and the spectral parameter of the spectral conjugate gradient method，a modified DL-type spectral conjugate gradient method is constructed. The spectral parameter is selected so that each iteration automatically generates a descent direction that does not depend on any line search. Under conventional assumptions，it is proved that this method is globally convergent for uniformly convex functions by using strong Wolfe line search.

Key words：unconstrained optimization；strong Wolfe line search；spectral conjugate gradient method；spectral parameter；global convergence

0　引　言

共轭梯度法(Conjugate gradient method，CGM)的迭代格式简单，在计算时仅需要目标函数值和梯度函数值，无需计算二阶导数值，储存需求低，是求解无约束优化问题(1)的常用迭代方法之一：

$$\min_{x\in\mathbf{R}^n} f(x) \tag{1}$$

其中：$f(x)$在 $\mathbf{R}^n\rightarrow\mathbf{R}$ 上是连续单调函数。

CGM 的一般迭代格式为：

$$x_{k+1}=x_k+\alpha_k d_k \tag{2}$$

收稿日期：2022－08－01　　网络出版日期：2022－11－01
作者简介：李亚敏(1992－　)，女，河南开封人，助教，硕士，主要从事优化理论方面的研究。

其中：

$$d_k=\begin{cases}-g_k, & k=1;\\ -g_k+\beta_k d_{k-1}, & k>1。\end{cases}$$

其中：x_k 是当前迭代点，α_k 是步长，$g_k=g(x_k)$是目标函数 $f(x)$在 x_k 处的梯度向量。

最著名的 CGM 有 6 种，参见文献[1-7]，本文关注最著名的共轭梯度方法：

$$\beta_k^{\mathrm{PRP}}=\frac{g_k^{\mathrm{T}}y_{k-1}}{\|g_{k-1}\|^2}{}^{[2\text{-}3]},$$

$$\beta_k^{\mathrm{HS}}=\frac{g_k^{\mathrm{T}}y_{k-1}}{d_{k-1}^{\mathrm{T}}y_{k-1}}{}^{[4]},$$

其中：$\|\cdot\|$代表 Euclidean 范数，$y_{k-1}=g_k=g_{k-1}$。

虽然这 6 种 CGM 可以解决诸如信号恢复[8]、图像处理[9]等很多实际问题，但仍不尽如人意。若有理论性质和数值表现都很好的 CGM，就能更好地解决实际问题。已有学者在最著名的 CGM 的基础上进行改进(主要是修正 β_k)，得到了理论和数值表现都好的 CGM，相关研究可参见文献[10-12]。

2001 年，Dai 等[13]提出新的共轭条件：$d_k^{\mathrm{T}}y_{k-1}=-tg_k^{\mathrm{T}}s_{k-1}$，并利用该条件推导出了 DL 共轭梯度法，简称 DL 法。DL 法的共轭参数为：

$$\beta_k^{\mathrm{DL}}=\beta_k^{\mathrm{HS}}-t\,\frac{g_k^{\mathrm{T}}s_{k-1}}{d_{k-1}^{\mathrm{T}}y_{k-1}},$$

其中：$t\geqslant0$，$s_{k-1}=x_k-x_{k-1}$。事实上，当 $t=0$ 时 DL 法即 HS 法，对于 t 的不同选择，数值结果有很大的不同。

2005 年，Hager 等[14]受无记忆 BFGS 方法的启发，提出了一个新的共轭参数：

$$\beta_k^{\mathrm{HZ}}=\frac{g_{k+1}^{\mathrm{T}}y_k}{d_k^{\mathrm{T}}y_k}-2\,\frac{\|y_k\|^2g_{k+1}^{\mathrm{T}}y_k}{(d_k^{\mathrm{T}}y_k)^2},$$

巧合的是，该方法可以看作 DL 法对应于 $t=2\,\dfrac{\|y_k\|^2}{s_k^{\mathrm{T}}y_k}$的一个自适应版本。段复建等[15]提出了参数 t 的一个自适应形式：

$$t=\frac{\|y_k\|^2}{s_k^{\mathrm{T}}y_k}+\theta\,\frac{s_k^{\mathrm{T}}y_k}{\|s_k\|^2}(\theta>0\text{ 是一个变化是参数});$$

并把 β_k^{DL} 修正为：

$$\beta_k^{\mathrm{DL}_+}=\max\left\{\frac{g_{k+1}^{\mathrm{T}}y_k}{d_k^{\mathrm{T}}y_k},0\right\}-\left(\frac{\|y_k\|^2}{s_k^{\mathrm{T}}y_k}+\frac{\|y_k\|}{\|s_k\|}\right)\frac{g_{k+1}^{\mathrm{T}}s_k}{d_k^{\mathrm{T}}y_k},$$

提出了一种自适应 DL 共轭梯度法，并证明了该方法对一致凸函数是全局收敛的。

2012 年，Rivaie 等[16]认为，学者们修正的 β_k 虽能具有良好的理论性质和数值表现，但所给的 β_k 过于复杂，不便于应用，因此提出了一个结构简单的 β_k^{RMIL}：

$$\beta_k^{\mathrm{RMIL}}=\frac{g_k^{\mathrm{T}}(g_k-g_{k-1})}{\|d_{k-1}\|^2}。$$

Rivaie 等[16]利用精确线搜索建立了 RMIL 法的全局收敛性；数值计算结果表明，与其他共轭参数相比，此公式具有更高的计算效率。

2001 年，Birgin 等[17]提出谱共轭梯度法(Spectral conjugate gradient method，SCGM)，其迭代格式为：

$$d_k=\begin{cases}-g_k, & k=1;\\ -\theta_k g_k+\beta_k d_{k-1}, & k>1\end{cases}\tag{3}$$

其中：$\beta_k=\dfrac{(\theta_k y_{k-1}-s_{k-1})^{\mathrm{T}}g_k}{d_{k-1}^{\mathrm{T}}y_{k-1}}$；$\theta_k$ 是谱参数，$\theta_k=\dfrac{\|s_{k-1}\|^2}{y_{k-1}^{\mathrm{T}}s_{k-1}}$。SCGM 有两大缺陷：一是搜索方向 d_k 不满足充分下降条件 $g_k^{\mathrm{T}}d_k<0$，二是算法的全局收敛性也未得到证明。但从数值试验上看，在大多数情况下，SCGM 在目标函数的梯度计算次数、最优函数值、CPU 计算时间等方面优于 CGM[17]。故众多学者着手修

正谱参数 θ_k 和共轭参数 β_k，以得到具有良好的理论性质和数值表现 SCGM，相关研究可参见文献[18-20]。

2016 年，Liu 等[21]在文献[13-14]的基础上，提出了一个 SCGM，称为 SPCG2+法，迭代格式为：

$$d_k=\begin{cases}-g_k, k=1,\\ -\left(1+\beta_k^{\mathrm{SPCG2}}\dfrac{g_k^{\mathrm{T}}d_{k-1}}{\|g_k\|^2}\right)g_k+\beta_k^{\mathrm{SPCG2}}d_{k-1}, & k>1,\end{cases}\quad \beta_k^{\mathrm{SPCG2}}=\beta_k^{\mathrm{HS}}-\frac{g_k^{\mathrm{T}}d_k}{\|d_k\|^2},$$

在强 Wolfe 线搜索下建立了 SPCG2+法的全局收敛性。

谱共轭梯度法作为共轭梯度法的一种拓展形式，是求解大规模无约束优化问题的比较实用的方法之一，目前已有方法各有优缺。通过修正共轭参数和谱参数得到了一种不依赖于任何线搜索都自动下降的修正 DL 型谱共轭梯度法，记为 DL-RSCG 法；并从理论上证明了该方法对一致凸函数是全局收敛的。

1　算法及其下降性

本文给出 DL-RSCG 法的谱参数 θ_k 和共轭参数 $\beta_k^{\mathrm{DL\text{-}R}}$ 为：

$$\theta_k=1+\beta_k^{\mathrm{DL\text{-}R}}\frac{g_k^{\mathrm{T}}d_{k-1}}{\|g_k\|^2} \tag{4}$$

$$\beta_k^{\mathrm{DL\text{-}R}}=\frac{g_k^{\mathrm{T}}(g_k-g_{k-1})}{\|d_{k-1}\|^2}-t\frac{g_k^{\mathrm{T}}s_{k-1}}{d_{k-1}^{\mathrm{T}}(g_k-g_{k-1})} \tag{5}$$

其中：$t\geqslant 0$。显然，当 $t=0$ 时，DL-RSCG 法退化为 RMIL 法。

DL-RSCG 算法步骤：

初始设置：给定初始值 $x_1\in\mathbf{R}^n$，$\varepsilon>0$，令 $0<\rho<\sigma<1$，$t\geqslant 0$。令 $k:=1$，$d_1=-g_1$。

步骤 1：若 $\|g_k\|<\varepsilon$，则停止；否则转步骤 2。

步骤 2：使用强 Wolfe 线搜索计算步长 $\alpha_k>0$，即 α_k 满足：

$$f(x_k+\alpha_k d_k)\leqslant f(x_k)+\rho\alpha_k g_k^{\mathrm{T}}d_k \tag{6}$$

$$|g(x_k+\alpha_k d_k)^{\mathrm{T}}d_k|\leqslant -\sigma g_k^{\mathrm{T}}d_k \tag{7}$$

步骤 3：利用式(3)—(5)，分别计算 d_k、θ_k、$\beta_k^{\mathrm{DL\text{-}R}}$。

步骤 4：令 $x_{k+1}=x_k+\alpha_k d_k$，求 g_{k+1}，令 $k:=k+1$，转步骤 1。

引理 1　设 d_k 由 DL-RSCG 算法生成，则对任意 $k\geqslant 1$，

$$g_k^{\mathrm{T}}d_k=-\|g_k\|^2<0 \tag{8}$$

成立。

证明　a)当 $k=1$ 时，有 $d_1=-g_1$，则有 $g_k^{\mathrm{T}}d_k=g_1^{\mathrm{T}}d_1=-\|g_1\|^2<0$ 成立。

b)当 $k>1$ 时，由式(3)有

$$\begin{aligned}g_k^{\mathrm{T}}d_k&=-\theta_k\|g_k\|^2+\beta_k^{\mathrm{DL\text{-}R}}g_k^{\mathrm{T}}d_{k-1}=-\left(1+\beta_k^{\mathrm{DL\text{-}R}}\frac{g_k^{\mathrm{T}}d_{k-1}}{\|g_k\|^2}\right)\|g_k\|^2+\beta_k^{\mathrm{DL\text{-}R}}g_k^{\mathrm{T}}d_{k-1}\\&=-\|g_k\|^2-\beta_k^{\mathrm{DL\text{-}R}}g_k^{\mathrm{T}}d_{k-1}+\beta_k^{\mathrm{DL\text{-}R}}g_k^{\mathrm{T}}d_{k-1}=-\|g_k\|^2。\end{aligned}$$

综上所述，式(8)得证。

引理 1 表明 DL-RSCG 法的搜索方向 d_k 在所有线搜索下都充分下降。

2　全局收敛性

为验证 DL-RSCG 算法的全局收敛性，需作如下假设：

假设 1　$f(x)$在水平集 $\Omega=\{x\in\mathbf{R}^n \mid f(x)\leqslant f(x_1)\}$上有下界，其中 x_1 为 DL-RSCG 算法的初始值。

假设 2　$f(x)$在水平集 Ω 的一个邻域 U 内连续可微，且梯度函数 $g(x)$满足 Lipschitz 连续，即存在常数 $L>0$，对任意 $x,y\in U$，有

$$\|g(x)-g(y)\|\leqslant L\|x-y\|。$$

定义 1　称连续可微函数 $f(x)$是一致凸函数，若存在一个常数 μ，使得：

$$[g(y)-g(x)]^{\mathrm{T}}(y-x) \geqslant \mu \|y-x\|^2, \forall x, y \in \Omega \tag{9}$$

下面证明 DL-RSCG 算法在强 Wolfe 线搜索下对一致凸函数是全局收敛的。

引理 2 对于一致凸函数 $f(x)$, $k \geqslant 1$ 和 $t \geqslant 0$,有

$$0 < \beta_{k+1}^{\text{DL-R}} \leqslant \left(1+\frac{t\sigma}{\mu}\right) \frac{\|g_{k+1}\|^2}{\|d_k\|^2} \tag{10}$$

成立。

证明 由文献[16]知, $0 \leqslant \beta_{k+1}^{\text{RMIL}} = \dfrac{g_{k+1}(g_{k+1}^{\mathrm{T}}-g_k)}{\|d_k\|^2} \leqslant \dfrac{\|g_{k+1}\|^2}{\|d_k\|^2}$。

由式(2)和式(9),得

$$d_k^{\mathrm{T}}(g_{k+1}-g_k) \geqslant \mu\alpha_k \|d_k\|^2,$$

故

$$|\beta_{k+1}^{\text{DL-R}}| = \left| \frac{g_{k+1}^{\mathrm{T}}(g_{k+1}-g_k)}{\|d_k\|^2} - t\frac{g_{k+1}^{\mathrm{T}} s_k}{d_k^{\mathrm{T}}(g_{k+1}-g_k)} \right| \leqslant \frac{\|g_{k+1}\|^2}{\|d_k\|^2} + t\frac{\alpha_k |g_{k+1}^{\mathrm{T}} d_k|}{|d_k^{\mathrm{T}}(g_{k+1}-g_k)|} \leqslant$$
$$\frac{\|g_{k+1}\|^2}{\|d_k\|^2} + t\frac{\alpha_k |g_{k+1}^{\mathrm{T}} d_k|}{\mu\alpha_k \|d_k\|^2} \leqslant \frac{\|g_{k+1}\|^2}{\|d_k\|^2} + \frac{t\sigma}{\mu}\frac{\|g_{k+1}\|^2}{\|d_k\|^2} \leqslant \left(1+\frac{t\sigma}{\mu}\right)\frac{\|g_{k+1}\|^2}{\|d_k\|^2}。$$

证毕。

引理 3 (Zoutendijk 条件[22])若 $f(x)$ 满足假设 1 和假设 2, d_k 由 DL-RSCG 算法生成, α_k 满足强 Wolfe 线搜索,则有

$$\sum_{k=1}^{\infty} \frac{(g_k^{\mathrm{T}} d_k)^2}{\|d_k\|^2} < +\infty \tag{11}$$

显然,由式(8),Zoutendijk 条件(11)等价于

$$\sum_{k=1}^{\infty} \frac{\|g_k\|^4}{\|d_k\|^2} < +\infty \tag{12}$$

证明 由式(7)和假设 2,可得:

$$L\alpha_k \|d_k\|^2 \geqslant g_{k+1}^{\mathrm{T}} d_k - g_k^{\mathrm{T}} d_k \geqslant (\sigma-1) g_k^{\mathrm{T}} d_k。$$

因此,

$$\alpha_k \geqslant \frac{\sigma-1}{L}\frac{g_k^{\mathrm{T}} d_k}{\|d_k\|^2}。$$

将式(8)代入上式,

$$\alpha_k \geqslant \frac{1-\sigma}{L}\frac{\|g_k\|^2}{\|d_k\|^2}。$$

由式(6)和假设 1,得

$$\sum_{k=1}^{\infty} \frac{\|g_k\|^4}{\|d_k\|^2} \leqslant \frac{L}{\rho(1-\sigma)} \sum_{k=1}^{\infty} \{f(x_k)-f(x_{k+1})\} < +\infty。$$

证毕。

定理 1 若假设 1 和假设 2 成立,且目标函数 $f(x)$ 是一致凸函数,序列 $\{g_k\}$ 由 DL-RSCG 算法产生,则有

$$\liminf_{k\to\infty} \|g_k\| = 0 \tag{13}$$

证明 假设结论不成立,那么必存在常数 $\varepsilon > 0$,使得

$$\|g_k\| \geqslant \varepsilon。$$

改写式(3)得:

$$d_{k+1} = -\theta_{k+1} g_{k+1} + \beta_{k+1}^{\text{DL-R}} d_k,$$

上式两边取平方模得:

$$\|d_{k+1}\|^2=(\beta_{k+1}^{\mathrm{DL\text{-}R}})^2\|d_k\|^2-2\theta_{k+1}g_{k+1}^{\mathrm{T}}d_{k+1}-\theta_{k+1}^2\|g_{k+1}\|^2。$$

上式两边同时除以$(g_{k+1}^{\mathrm{T}}d_{k+1})^2$，得：

$$\begin{aligned}\frac{\|d_{k+1}\|^2}{(g_{k+1}^{\mathrm{T}}d_{k+1})^2}&=\frac{(\beta_{k+1}^{\mathrm{DL\text{-}R}})^2\|d_k\|^2}{(g_{k+1}^{\mathrm{T}}d_{k+1})^2}-\frac{2\theta_{k+1}}{g_{k+1}^{\mathrm{T}}d_{k+1}}-\frac{\theta_{k+1}^2\|g_{k+1}\|^2}{(g_{k+1}^{\mathrm{T}}d_{k+1})^2}\\&=\frac{(\beta_{k+1}^{\mathrm{DL\text{-}R}})^2\|d_k\|^2}{(g_{k+1}^{\mathrm{T}}d_{k+1})^2}-\left(\frac{1}{\|g_{k+1}\|}+\frac{\theta_{k+1}\|g_{k+1}\|}{g_{k+1}^{\mathrm{T}}d_{k+1}}\right)^2+\frac{1}{\|g_{k+1}\|^2}\\&\leqslant\frac{(\beta_{k+1}^{\mathrm{DL\text{-}R}})^2\|d_k\|^2}{(g_{k+1}^{\mathrm{T}}d_{k+1})^2}+\frac{1}{\|g_{k+1}\|^2},\end{aligned}$$

再利用式(8)和式(10)，得：

$$\begin{aligned}\frac{\|d_{k+1}\|^2}{\|g_{k+1}\|^4}&\leqslant\left[\left(1+\frac{t\sigma}{\mu}\right)\frac{\|g_{k+1}\|^2}{\|d_k\|^2}\right]^2\frac{\|d_k\|^2}{\|g_{k+1}\|^4}+\frac{1}{\|g_{k+1}\|^2}\\&\leqslant\left(1+\frac{t\sigma}{\mu}\right)^2\frac{1}{\|d_k\|^2}+\frac{1}{\|g_{k+1}\|^2},\end{aligned}\tag{14}$$

由式(8)和 Cauchy-Schwarz 不等式，有$\|g_k\|\leqslant\|d_k\|$。故$\frac{1}{\|d_i\|}\leqslant\frac{1}{\|g_i\|}$。

注意到当$k=1$时，$d_1=-g_1$，即$\frac{1}{\|d_1\|}=\frac{1}{\|g_1\|}$。所以由式(14)得：

$$\frac{\|d_k\|^2}{\|g_k\|^4}\leqslant M\sum_{i=1}^{k}\frac{1}{\|g_i\|^2}\leqslant\frac{Mk}{\varepsilon^2}。$$

其中$M=\left(1+\frac{t\sigma}{\mu}\right)^2$。所以

$$\sum_{k=1}^{\infty}\frac{\|g_k\|^4}{\|d_k\|^2}\geqslant\sum_{k=1}^{\infty}\frac{\varepsilon^2}{Mk}=+\infty。$$

显然，这与引理 3 中的式(12)矛盾，故

$$\liminf_{k\to\infty}\|g_k\|=0。$$

证毕。

3　结　论

本文在 DL 法和谱共轭梯度法的基础上，把 DL 法的共轭参数 β_k^{DL} 的第一项$\frac{g_k^{\mathrm{T}}y_{k-1}}{d_{k-1}^{\mathrm{T}}y_{k-1}}$修正为$\frac{g_k^{\mathrm{T}}(g_k-g_{k-1})}{\|d_{k-1}\|^2}$，得到了共轭参数 $\beta_k^{\mathrm{DL\text{-}R}}=\frac{g_k^{\mathrm{T}}(g_k-g_{k-1})}{\|d_{k-1}\|^2}-t\ \frac{g_k^{\mathrm{T}}s_{k-1}}{d_{k-1}^{\mathrm{T}}(g_k-g_{k-1})}$，修正谱参数为 $\theta_k=1+\beta_k^{\mathrm{DL\text{-}R}}\frac{g_k^{\mathrm{T}}d_{k-1}}{\|g_k\|^2}$，使得搜索方向 d_k 不依赖于任何线搜索都充分下降，进而得到了一种求解大规模无约束优化问题的修正 DL 型谱共轭梯度法；并在强 Wolfe 线搜索下证明了该方法对一致凸函数是全局收敛的。

参考文献：

[1] Fletcher R, Reeves C M. Function minimization by conjugate gradients[J]. The Computer Journal, 1964, 7(2): 149-154.

[2] Polak E, Ribiere G. Note sur la convergence de méthodes de directions conjuguées[J]. Revue Française d'Informatique et De Recherche Opérationnelle Série Rouge, 1969, 3(16): 35-43.

[3] Polyak B T. The conjugate gradient method in extremal problems[J]. USSR Computational Mathematics and Mathematical Physics, 1969, 9(4): 94-112.

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

[5] 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.
[6] Liu Y, Storey C. Efficient generalized conjugate gradient algorithms, part 1: Theory[J]. Journal of Optimization Theory and Applications, 1991, 69(1):129-137.
[7] Fletcher R. Practical Methods of Optimization, Vol 1: Unconstrained Optimization[M]. New York: John Wiley & Sons, 1987:80-92.
[8] 李丹丹,李远飞,王松华. 一种修正三项 Hestenes-Stiefel 共轭梯度投影算法及其应用[J]. 吉林大学学报(理学版),2022,60(1):64-72.
[9] Jiang X Z, Liao W, Yin J H, et al. A new family of hybrid three-term conjugate gradient methods with applications in image restoration[J]. Numerical Algorithms, 2022, 91(1):161-191.
[10] Waziri M Y, Ahmed K, Halilu A S. A modified PRP-type conjugate gradient projection algorithm for solving large-scale monotone nonlinear equations with convex constraint[J]. Journal of Computational and Applied Mathematics, 2022,407: 114035.
[11] Lotfi M, Hosseini S M. An efficient Dai-Liao type conjugate gradient method by reformulating the CG parameter in the search direction equation[J]. Journal of Computational and Applied Mathematics, 2020, 371:112708.
[12] Mtagulwa P, Kaelo P. An efficient modified PRP-FR hybrid conjugate gradient method for solving unconstrained optimization problems[J]. Applied Numerical Mathematics, 2019, 145:111-120.
[13] 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.
[14] Hager W, Zhang H C. A new conjugate gradient method with guaranteed descent and an efficient line search[J]. SIAM Journal on Optimization, 2005,16:170-192.
[15] 段复建,杨冲,李向利. 一个自适应 Dai-Liao 共轭梯度法[J]. 应用数学, 2022,35(2):336-342.
[16] Rivaie M, Mamat M, June L W, et al. A new class of nonlinear conjugate gradient coefficients with global convergence properties[J]. Applied Mathematics and Computation, 2012, 218(22): 11323-11332.
[17] Birgin E G, Martínez J M. A spectral conjugate gradient method for unconstrained optimization[J]. Applied Mathematics and Optimization, 2001, 43(2): 117-128.
[18] Awwal A M, Kumam P, Abubakar A B. Spectral modified Polak-Ribiére-Polyak projection conjugate gradient method for solving monotone systems of nonlinear equations[J]. Applied Mathematics and Computation, 2019, 362:124514.
[19] Koorapetse M, Kaelo P, Lekoko S, et al. A derivative-free RMIL conjugate gradient projection method for convex constrained nonlinear monotone equations with applications in compressive sensing[J]. Applied Numerical Mathematics, 2021,165:431-441.
[20] Fang X W. A class of new derivative-free gradient type methods for large-scale nonlinear systems of monotone equations [J]. Journal of Inequalities and Applications, 2020, 2020: 93.
[21] Liu H, Yao Y, Qian X Y, et al. Some nonlinear conjugate gradient methods based on spectral scaling secant equations [J]. Computational and Applied Mathematics,2016,35(2):639-651.
[22] 高立. 数值最优化方法[M]. 北京:北京大学出版社,2014:107-108.

(责任编辑:康　锋)