DOI：10.13718/j.cnki.xdzk.2017.07.015

基于 PRP 公式修正的有效共轭梯度算法[①]

林穗华

广西民族师范学院 数学与计算机科学学院，广西 崇左 532200

摘要： 给出一种非负且带有调比因子的修正 PRP 共轭梯度法参数公式. 基于该共轭参数公式，采用 SWP 线搜索的对应共轭梯度算法满足充分下降性，采用 WWP 线搜索的对应谱共轭梯度算法保持下降性. 在常规假设条件下，证明了算法的全局收敛性，数值实验结果表明算法是有效的.

关 键 词： 无约束优化；共轭梯度法；谱共轭梯度法；全局收敛性

中图分类号：O221.2 **文献标志码：A** **文章编号：** 1673－9868(2017)07－0097－07

考虑无约束优化问题：

$$\min\{f(x) \mid x \in \mathbb{R}^n\} \tag{1}$$

其中目标函数 $f: \mathbb{R}^n \longrightarrow \mathbb{R}$ 连续可微，梯度函数 $\boldsymbol{g}(x) \triangleq \nabla f(x)$，$\boldsymbol{g}_k \triangleq \boldsymbol{g}(\boldsymbol{x}_k)$，$f_k \triangleq f(\boldsymbol{x}_k)$. 共轭梯度法是求解大型优化问题(1) 的一种有效方法，迭代格式如下

$$x_{k+1} = \boldsymbol{x}_k + \alpha_k \boldsymbol{d}_k \tag{2}$$

$$\boldsymbol{d}_k = \begin{cases} -\boldsymbol{g}_k & k = 1 \\ -\boldsymbol{g}_k + \beta_k \boldsymbol{d}_{k-1} & k \geqslant 2 \end{cases} \tag{3}$$

其中：$\boldsymbol{d}_k$ 为搜索方向，α_k 为步长因子，β_k 为共轭方向调控参数. 共轭参数的选取影响着算法的收敛性和计算效果. 著名的共轭梯度法 PRP,FR 的 β_k 表达式如下

$$\beta_k^{\mathrm{PRP}} = \frac{\boldsymbol{g}_k^{\mathrm{T}}(\boldsymbol{g}_k - \boldsymbol{g}_{k-1})}{\|\boldsymbol{g}_{k-1}\|^2} \qquad \beta_k^{\mathrm{FR}} = \frac{\|\boldsymbol{g}_k\|^2}{\|\boldsymbol{g}_{k-1}\|^2}$$

其中 $\|\cdot\|$ 为欧氏范数. PRP 方法具备优秀的数值特性，但即使采用精确线搜索，对非凸目标函数仍不能确保其收敛性[1]，不少研究者对 β_k^{PRP} 公式作了非负性的改进[2-6]，解决了收敛性问题. 文献[3] 对 β_k^{PRP} 公式修改得到

$$\beta_k^{\mathrm{WYL}} = \frac{\boldsymbol{g}_k^{\mathrm{T}}\left(\boldsymbol{g}_k - \dfrac{\|\boldsymbol{g}_k\|}{\|\boldsymbol{g}_{k-1}\|}\boldsymbol{g}_{k-1}\right)}{\|\boldsymbol{g}_{k-1}\|^2} \tag{4}$$

$0 \leqslant \beta_k^{\mathrm{WYL}} \leqslant 2\beta_k^{\mathrm{FR}}$，WYL 方法继承了 PRP 方法优秀的数值特性. 文献[6] 对公式(4) 再修改得到

$$\beta_k^{\mathrm{VPRP}} = \frac{\boldsymbol{g}_k^{\mathrm{T}}\left(\boldsymbol{g}_k - \dfrac{\boldsymbol{g}_k^{\mathrm{T}}\boldsymbol{g}_{k-1}}{\|\boldsymbol{g}_{k-1}\|^2}\boldsymbol{g}_{k-1}\right)}{\|\boldsymbol{g}_{k-1}\|^2} \tag{5}$$

$0 \leqslant \beta_k^{\mathrm{VPRP}} \leqslant \beta_k^{\mathrm{FR}}$，$\beta_k^{\mathrm{VPRP}}$ 与 β_k^{WYL} 一样满足文献[2] 为 PRP 方法提出的性质(＊)，文献[7] 还将公式(5) 的分子部分推广到其他共轭梯度法公式上，也取得了较好的效果.

① 收稿日期：2016－08－15

基金项目：国家自然科学基金项目(11261006)；广西高校科研项目(ZD2014143)；广西重点培育学科(应用数学)建设项目(桂教科研[2013]16)；广西民族师范学院科研项目(2013RCGG002).

作者简介：林穗华(1973－)，女，广西龙州人，副教授，主要从事最优化理论与方法研究.

基于文献[3，6]的思路，本文考虑对 β_k^{PRP} 公式作如下改进

$$\beta_k = \frac{\boldsymbol{g}_k^{\mathrm{T}}\left(\boldsymbol{g}_k - \dfrac{(\boldsymbol{g}_k^{\mathrm{T}}\boldsymbol{g}_{k-1})^2}{\|\boldsymbol{g}_k\|\,\|\boldsymbol{g}_{k-1}\|^3}\boldsymbol{g}_{k-1}\right)}{\|\boldsymbol{g}_{k-1}\|^2} \tag{6}$$

易得 $0 \leqslant \beta_k \leqslant 2\beta_k^{FR}$，但 β_k 不再满足性质(*). 受文献[8]的启发，对公式(6)再修正得到

$$\beta_k^N = \frac{\boldsymbol{g}_k^{\mathrm{T}}\left(\boldsymbol{g}_k - \dfrac{(\boldsymbol{g}_k^{\mathrm{T}}\boldsymbol{g}_{k-1})^2}{\|\boldsymbol{g}_k\|\,\|\boldsymbol{g}_{k-1}\|^3}\boldsymbol{g}_{k-1}\right)}{\eta_k\|\boldsymbol{g}_{k-1}\|^2} \tag{7}$$

其中

$$\eta_k = 1 - \mu\min\left\{0,\ \frac{(\boldsymbol{g}_k^{\mathrm{T}}\boldsymbol{g}_{k-1})^3}{\|\boldsymbol{g}_k\|^3\,\|\boldsymbol{g}_{k-1}\|^3}\right\} \qquad \mu \geqslant 1$$

建立在参数 β_k^N 公式上的相应共轭梯度算法和谱共轭梯度算法具有良好的收敛性和较为理想的数值表现.

1 修正 PRP 共轭梯度算法及谱共轭梯度算法

基于参数公式(7)，采用强 Wolfe-Powell 线搜索(SWP)的修正 PRP 共轭梯度算法如下：

算法 1

步骤 1 给定初值 $\boldsymbol{x}_1 \in \mathbb{R}^n$，$\delta \in (0, 0.5)$，$\sigma \in (\delta, 0.5)$，$\mu \geqslant 1$，$\varepsilon \geqslant 0$，$\boldsymbol{d}_1 := -\boldsymbol{g}_1$，$k := 1$. 若 $\|\boldsymbol{g}_k\| \leqslant \varepsilon$，停止.

步骤 2 计算步长 α_k 满足以下 SWP 线搜索准则：

$$f(\boldsymbol{x}_k + \alpha_k\boldsymbol{d}_k) - f_k \leqslant \delta\alpha_k\boldsymbol{g}_k^{\mathrm{T}}\boldsymbol{d}_k \tag{8}$$

$$|\boldsymbol{g}(\boldsymbol{x}_k + \alpha_k\boldsymbol{d}_k)^{\mathrm{T}}\boldsymbol{d}_k| \leqslant -\sigma\boldsymbol{g}_k^{\mathrm{T}}\boldsymbol{d}_k \tag{9}$$

步骤 3 计算 $x_{k+1} = \boldsymbol{x}_k + \alpha_k\boldsymbol{d}_k$. 若 $\|\boldsymbol{g}_{k+1}\| \leqslant \varepsilon$，停止.

步骤 4 由(7)式计算参数 β_{k+1}，由(3)式计算 $\boldsymbol{d}_{k+1}$.

步骤 5 $k := k+1$，转步骤 2.

谱共轭梯度法是一种特殊共轭梯度法[9]，通过调整谱参数 θ_k 和共轭参数 β_k，易使设计的算法满足下降性和收敛性. 为降低线搜索条件，受文献[10-16]的启发，设计基于参数公式(7)并采用弱 Wolfe-Powell 线搜索(WWP)的修正 PRP 谱共轭梯度算法如下：

算法 2

步骤 1 给定初值 $x_1 \in \mathbb{R}^n$，$\delta \in (0, 0.5)$，$\sigma \in (\delta, 0.5)$，$\mu \geqslant 1$，$\varepsilon \geqslant 0$，$\boldsymbol{d}_1 := -\boldsymbol{g}_1$，$k := 1$. 若 $\|\boldsymbol{g}_k\| \leqslant \varepsilon$，停止.

步骤 2 计算步长 α_k 满足 WWP 线搜索准则：(8)式和下式

$$\boldsymbol{g}(\boldsymbol{x}_k + \alpha_k\boldsymbol{d}_k)^{\mathrm{T}}\boldsymbol{d}_k \geqslant \sigma\boldsymbol{g}_k^{\mathrm{T}}\boldsymbol{d}_k \tag{10}$$

步骤 3 计算 $x_{k+1} = \boldsymbol{x}_k + \alpha_k\boldsymbol{d}_k$. 若 $\|\boldsymbol{g}_{k+1}\| \leqslant \varepsilon$，停止.

步骤 4 由(7)式计算参数 β_{k+1}，由下式计算搜索方向 $\boldsymbol{d}_{k+1}$：

$$\boldsymbol{d}_{k+1} = -\theta_{k+1}\boldsymbol{g}_{k+1} + \beta_{k+1}\boldsymbol{d}_k \tag{11}$$

其中

$$\theta_{k+1} = \frac{\boldsymbol{d}_k^{\mathrm{T}}\boldsymbol{y}_k}{\|\boldsymbol{g}_k\|^2} \qquad \boldsymbol{y}_k = \boldsymbol{g}_{k+1} - \boldsymbol{g}_k \tag{12}$$

步骤 5 $k := k+1$，转步骤 2.

算法的收敛性证明需要共轭梯度法的常规假设条件：

(i) 设水平集 $\Omega = \{\boldsymbol{x} \in \mathbb{R}^n \mid f(\boldsymbol{x}) \leqslant f(\boldsymbol{x}_1)\}$ 有界，其中 $\boldsymbol{x}_1$ 为初始点.

(ii) $f(x)$ 在 Ω 的某邻域 N 上连续可微且导数满足 Lipschitz 条件，即存在常数 $L > 0$，使得

$$\|\boldsymbol{g}(\boldsymbol{x}) - \boldsymbol{g}(\boldsymbol{y})\| \leqslant L\|\boldsymbol{x} - \boldsymbol{y}\| \qquad \forall\, \boldsymbol{x}, \boldsymbol{y} \in N$$

本文以下算法的收敛性分析中均假设 $\|\boldsymbol{g}_k\| \neq 0$，否则目标函数的稳定点已获得，算法自动终止.

2　修正 PRP 共轭梯度算法的全局收敛性

引理 1　考虑参数 $\sigma\in\left(0,\dfrac{1}{2}\right)$，$\mu\geqslant 1$，则存在常数 $c_1=\dfrac{1-2\sigma}{1-\sigma}\in(0,1)$，$c_2=\dfrac{1}{1-\sigma}\in(1,2)$，使算法 1 产生的方向序列$\{\boldsymbol{d}_k\}$满足充分下降条件

$$-c_2\|\boldsymbol{g}_k\|^2\leqslant\boldsymbol{g}_k^{\mathrm{T}}\boldsymbol{d}_k\leqslant-c_1\|\boldsymbol{g}_k\|^2\qquad\forall k\geqslant 1\tag{13}$$

证　因 β_k^N 的调比因子

$$\eta_k=1-\mu\min\left\{0,\frac{(\boldsymbol{g}_k^{\mathrm{T}}\boldsymbol{g}_{k-1})^3}{\|\boldsymbol{g}_k\|^3\|\boldsymbol{g}_{k-1}\|^3}\right\}\geqslant 1-\frac{(\boldsymbol{g}_k^{\mathrm{T}}\boldsymbol{g}_{k-1})^3}{\|\boldsymbol{g}_k\|^3\|\boldsymbol{g}_{k-1}\|^3}\geqslant 0$$

所以由(7)式，可知

$$0\leqslant\beta_k^N\leqslant\beta_k^{\mathrm{FR}}\tag{14}$$

当 $k=1$ 时，$\boldsymbol{d}_1=-\boldsymbol{g}_1$，$\boldsymbol{g}_1^{\mathrm{T}}\boldsymbol{d}_1=-\|\boldsymbol{g}_1\|^2$，显然(13)式成立. 假设 $k>1$ 时(13)式成立，则 $\boldsymbol{g}_k^{\mathrm{T}}\boldsymbol{d}_k<0$.

(3)式两端与 $\boldsymbol{g}_{k+1}^{\mathrm{T}}$ 作内积，并利用(9)和(14)式，可得

$$-\|\boldsymbol{g}_{k+1}\|^2+\sigma\beta_{k+1}^N\boldsymbol{g}_k^{\mathrm{T}}\boldsymbol{d}_k\leqslant\boldsymbol{g}_{k+1}^{\mathrm{T}}\boldsymbol{d}_{k+1}=-\|\boldsymbol{g}_{k+1}\|^2+\beta_{k+1}^N\boldsymbol{g}_{k+1}^{\mathrm{T}}\boldsymbol{d}_k\leqslant-\|\boldsymbol{g}_{k+1}\|^2-\sigma\beta_{k+1}^N\boldsymbol{g}_k^{\mathrm{T}}\boldsymbol{d}_k\tag{15}$$

利用(15)式及(14)式，可得

$$-\|\boldsymbol{g}_{k+1}\|^2+\sigma\beta_{k+1}^{\mathrm{FR}}\boldsymbol{g}_k^{\mathrm{T}}\boldsymbol{d}_k\leqslant\boldsymbol{g}_{k+1}^{\mathrm{T}}\boldsymbol{d}_{k+1}\leqslant-\|\boldsymbol{g}_{k+1}\|^2-\sigma\beta_{k+1}^{\mathrm{FR}}\boldsymbol{g}_k^{\mathrm{T}}\boldsymbol{d}_k$$

上式两边除以 $-\|\boldsymbol{g}_{k+1}\|^2$，可得

$$1-\sigma\frac{-\boldsymbol{g}_k^{\mathrm{T}}\boldsymbol{d}_k}{\|\boldsymbol{g}_k\|^2}\leqslant\frac{-\boldsymbol{g}_{k+1}^{\mathrm{T}}\boldsymbol{d}_{k+1}}{\|\boldsymbol{g}_{k+1}\|^2}\leqslant 1+\sigma\frac{-\boldsymbol{g}_k^{\mathrm{T}}\boldsymbol{d}_k}{\|\boldsymbol{g}_k\|^2}\tag{16}$$

利用(16)式递推，可得

$$c_1=\frac{1-2\sigma}{1-\sigma}\leqslant\frac{1-2\sigma+\sigma^{k+1}}{1-\sigma}\leqslant\frac{-\boldsymbol{g}_{k+1}^{\mathrm{T}}\boldsymbol{d}_{k+1}}{\|\boldsymbol{g}_{k+1}\|^2}\leqslant\frac{1-\sigma^{k+1}}{1-\sigma}\leqslant\frac{1}{1-\sigma}=c_2$$

由归纳法知，引理成立.

由引理 1 及文献[1]引理 1.4.1，可得以下引理.

引理 2　若假设(i),(ii)成立，则算法 1 产生的序列$\{\boldsymbol{g}_k,\boldsymbol{d}_k\}$满足 Zoutendijk 条件：$\displaystyle\sum_{k\geqslant 1}\frac{\|\boldsymbol{g}_k\|^4}{\|\boldsymbol{d}_k\|^2}<+\infty$.

定理 1　若假设(i),(ii)成立，$\{\boldsymbol{g}_k\}$为算法 1 产生的序列，则有

$$\lim_{k\to\infty}\inf\|\boldsymbol{g}_k\|=0$$

证　若定理结论不成立，即

$$\lim_{k\to\infty}\inf\|\boldsymbol{g}_k\|>0$$

则存在常量 $\gamma>0$，使得

$$\|\boldsymbol{g}_k\|\geqslant\gamma\qquad\forall k\geqslant 1\tag{17}$$

(3)式两端取模平方，可得

$$\|\boldsymbol{d}_k\|^2=(\beta_k^N)^2\|\boldsymbol{d}_{k-1}\|^2-2\beta_k^N\boldsymbol{g}_k^{\mathrm{T}}\boldsymbol{d}_{k-1}+\|\boldsymbol{g}_k\|^2\tag{18}$$

(18)式两边除以 $\|\boldsymbol{g}_k\|^4$，再由(9)式和(13),(14),(17)式，可得

$$\begin{aligned}\frac{\|\boldsymbol{d}_k\|^2}{\|\boldsymbol{g}_k\|^4}&=(\beta_k^N)^2\frac{\|\boldsymbol{d}_{k-1}\|^2}{\|\boldsymbol{g}_k\|^4}-2\beta_k^N\frac{\boldsymbol{g}_k^{\mathrm{T}}\boldsymbol{d}_{k-1}}{\|\boldsymbol{g}_k\|^4}+\frac{1}{\|\boldsymbol{g}_k\|^2}\leqslant\\&(\beta_k^{\mathrm{FR}})^2\frac{\|\boldsymbol{d}_{k-1}\|^2}{\|\boldsymbol{g}_k\|^4}+2\beta_k^{\mathrm{FR}}\frac{|\boldsymbol{g}_k^{\mathrm{T}}\boldsymbol{d}_{k-1}|}{\|\boldsymbol{g}_k\|^4}+\frac{1}{\|\boldsymbol{g}_k\|^2}\leqslant\\&\frac{\|\boldsymbol{d}_{k-1}\|^2}{\|\boldsymbol{g}_{k-1}\|^4}+2\sigma\frac{-\boldsymbol{g}_{k-1}^{\mathrm{T}}\boldsymbol{d}_{k-1}}{\|\boldsymbol{g}_k\|^2\|\boldsymbol{g}_{k-1}\|^2}+\frac{1}{\|\boldsymbol{g}_k\|^2}\leqslant\\&\frac{\|\boldsymbol{d}_{k-1}\|^2}{\|\boldsymbol{g}_{k-1}\|^4}+\frac{1}{\|\boldsymbol{g}_k\|^2}\left(1+\frac{2\sigma}{1-\sigma}\right)\leqslant\\&\frac{1+\sigma}{1-\sigma}\sum_{i=1}^{k}\frac{1}{\|\boldsymbol{g}_i\|^2}\leqslant\frac{(1+\sigma)k}{(1-\sigma)\gamma^2}\end{aligned}\tag{19}$$

由(19)式，可得

$$\sum_{k\geqslant 1}\frac{\|\boldsymbol{g}_k\|^4}{\|\boldsymbol{d}_k\|^2}\geqslant\frac{1-\sigma}{1+\sigma}\gamma^2\sum_{k\geqslant 1}\frac{1}{k}=+\infty$$

这与引理 2 矛盾，由反证法知定理成立.

3 修正 PRP 谱共轭梯度算法的全局收敛性

引理 3 考虑算法 2 产生的序列$\{\boldsymbol{g}_k,\boldsymbol{d}_k,\beta_k^N\}$，则 $\forall k\geqslant 1$ 有

$$\boldsymbol{g}_k^{\mathrm{T}}\boldsymbol{d}_k<0,\ 0\leqslant\beta_{k+1}^N\leqslant\frac{\boldsymbol{g}_{k+1}^{\mathrm{T}}\boldsymbol{d}_{k+1}}{\boldsymbol{g}_k^{\mathrm{T}}\boldsymbol{d}_k}\tag{20}$$

证 当 $k=1$ 时，$\boldsymbol{d}_1=-\boldsymbol{g}_1$，则 $\boldsymbol{g}_1^{\mathrm{T}}\boldsymbol{d}_1=-\|\boldsymbol{g}_1\|^2<0$. 假设 $k\geqslant 1$ 时 $\boldsymbol{g}_k^{\mathrm{T}}\boldsymbol{d}_k<0$，从而由(10)与(12)式得 $\boldsymbol{d}_k^{\mathrm{T}}\boldsymbol{y}_k\geqslant(\sigma-1)\boldsymbol{g}_k^{\mathrm{T}}\boldsymbol{d}_k>0$. 下证 $\boldsymbol{g}_{k+1}^{\mathrm{T}}\boldsymbol{d}_{k+1}<0$.

(11)式两端与 $\boldsymbol{g}_{k+1}$ 作内积，可得

$$\boldsymbol{g}_{k+1}^{\mathrm{T}}\boldsymbol{d}_{k+1}=-\theta_{k+1}\|\boldsymbol{g}_{k+1}\|^2+\beta_{k+1}\boldsymbol{g}_{k+1}^{\mathrm{T}}\boldsymbol{d}_k$$

由(14)式知 $0\leqslant\beta_{k+1}^N\leqslant\beta_{k+1}^{\mathrm{FR}}$. 若 $\beta_{k+1}^N>0$，则上式可得

$$\boldsymbol{g}_{k+1}^{\mathrm{T}}\boldsymbol{d}_{k+1}=\frac{\|\boldsymbol{g}_{k+1}\|^2}{\|\boldsymbol{g}_k\|^2}(-\boldsymbol{d}_k^{\mathrm{T}}\boldsymbol{y}_k)+\beta_{k+1}\boldsymbol{g}_{k+1}^{\mathrm{T}}\boldsymbol{d}_k\leqslant\beta_{k+1}(-\boldsymbol{d}_k^{\mathrm{T}}\boldsymbol{y}_k+\boldsymbol{g}_{k+1}^{\mathrm{T}}\boldsymbol{d}_k)=\beta_{k+1}\boldsymbol{g}_k^{\mathrm{T}}\boldsymbol{d}_k<0\tag{21}$$

若 $\beta_{k+1}=0$，则结合 $\boldsymbol{d}_k^{\mathrm{T}}\boldsymbol{y}_k>0$ 可得

$$\boldsymbol{g}_{k+1}^{\mathrm{T}}\boldsymbol{d}_{k+1}=-\frac{\boldsymbol{d}_k^{\mathrm{T}}\boldsymbol{y}_k}{\|\boldsymbol{g}_k\|^2}\|\boldsymbol{g}_{k+1}\|^2<0$$

由归纳法知，$\forall k\geqslant 1$，$\boldsymbol{g}_k^{\mathrm{T}}\boldsymbol{d}_k<0$ 成立. 再由(21)式可得

$$0\leqslant\beta_{k+1}\leqslant\frac{\boldsymbol{g}_{k+1}^{\mathrm{T}}\boldsymbol{d}_{k+1}}{\boldsymbol{g}_k^{\mathrm{T}}\boldsymbol{d}_k}$$

引理得证.

引理 3 说明算法 2 产生的搜索方向 $\boldsymbol{d}_k$ 满足下降性. 根据文献[1] 引理 1.4.1，易得以下结论.

引理 4 若假设(i),(ii)成立，$\{\boldsymbol{g}_k,\boldsymbol{d}_k\}$ 为算法 2 产生的序列，则

$$\sum_{k\geqslant 1}\frac{(\boldsymbol{g}_k^{\mathrm{T}}\boldsymbol{d}_k)^2}{\|\boldsymbol{d}_k\|^2}<+\infty$$

定理 2 若假设(i),(ii)成立，$\{\boldsymbol{g}_k\}$ 为算法 2 生成的序列，则有

$$\lim_{k\to\infty}\inf\|\boldsymbol{g}_k\|=0$$

证 若定理不成立，则存在常数 $r>0$，使任意 $k\geqslant 1$ 有 $\|\boldsymbol{g}_k\|\geqslant r$.

由(11)式移项得 $\boldsymbol{d}_k+\theta_k\boldsymbol{g}_k=\beta_k\boldsymbol{d}_{k-1}$，两端取模平方，再除以$(\boldsymbol{g}_k^{\mathrm{T}}\boldsymbol{d}_k)^2$，并利用(20)式，可得

$$\begin{aligned}\frac{\|\boldsymbol{d}_k\|^2}{(\boldsymbol{g}_k^{\mathrm{T}}\boldsymbol{d}_k)^2}&=(\beta_k^N)^2\frac{\|\boldsymbol{d}_{k-1}\|^2}{(\boldsymbol{g}_k^{\mathrm{T}}\boldsymbol{d}_k)^2}-\frac{2\theta_k}{\boldsymbol{g}_k^{\mathrm{T}}\boldsymbol{d}_k}-\frac{\theta_k^2\|\boldsymbol{g}_k\|^2}{(\boldsymbol{g}_k^{\mathrm{T}}\boldsymbol{d}_k)^2}\leqslant\\&\frac{\|\boldsymbol{d}_{k-1}\|^2}{(\boldsymbol{g}_{k-1}^{\mathrm{T}}\boldsymbol{d}_{k-1})^2}+\frac{1}{\|\boldsymbol{g}_k\|^2}-\left(\frac{1}{\|\boldsymbol{g}_k\|}+\frac{\theta_k\|\boldsymbol{g}_k\|}{\boldsymbol{g}_k^{\mathrm{T}}\boldsymbol{d}_k}\right)^2\leqslant\\&\frac{\|\boldsymbol{d}_{k-1}\|^2}{(\boldsymbol{g}_{k-1}^{\mathrm{T}}\boldsymbol{d}_{k-1})^2}+\frac{1}{\|\boldsymbol{g}_k\|^2}\leqslant\\&\sum_{i=1}^{k}\frac{1}{\|\boldsymbol{g}_i\|^2}\leqslant\frac{k}{r^2}\end{aligned}$$

从而可得

$$\sum_{k\geqslant 1}\frac{(\boldsymbol{g}_k^{\mathrm{T}}\boldsymbol{d}_k)^2}{\|\boldsymbol{d}_k\|^2}\geqslant r^2\sum_{k\geqslant 1}\frac{1}{k}=+\infty$$

上式与引理 4 矛盾，由反证法知定理成立.

4　数值实验

为检验本文提出的修正 PRP 共轭梯度法的计算效果，给出算法 1 和算法 2 的数值实验结果如表 1. 测试函数源于文献[17]，算法运行环境为 Matlab 2011b 和 Windows7 操作系统. 终止条件为 $\| \boldsymbol{g}_k \| \leqslant 10^{-6}$，或迭代次数超过 9 999. 表 1 中 $NI/NF/NG$ 分别表示算法迭代次数、目标函数计算迭代次数、梯度函数计算迭代次数，其它符号说明如下：

算法 1 中参数 $\mu=1.15$，$\delta=0.001$，$\sigma=0.25$；算法 2 中参数 $\mu=4.2$，$\delta=0.1$，$\sigma=0.9$；PRP 与 VPRP 采用 SWP 线搜索，$\delta=0.001$，$\sigma=0.25$.

表 1　数值实验结果

测试函数	维数/维	算法 1/次 $NI/NF/NG$	算法 2/次 $NI/NF/NG$	VPRP/次 $NI/NF/NG$	PRP/次 $NI/NF/NG$
ROSE	2	50/143/111	51/101/74	45/116/90	29/206/75
FROTH	2	17/33/28	35/55/45	24/51/44	15/127/24
BEALE	2	24/50/41	42/66/53	25/52/41	13/27/23
JENSAM	2	18/37/29	22/37/29	14/31/23	16/34/28
HELIX	3	113/196/178	78/115/98	67/142/116	52/108/87
BARD	3	39/74/60	44/73/59	38/77/62	65/126/105
GAUSS	3	4/8/6	5/9/7	5/10/8	4/57/6
GULF	3	2/52/3	2/52/3	2/52/3	2/52/3
SING	4	489/890/754	202/322/265	388/760/608	130/256/201
WOOD	4	234/507/426	304/540/416	228/478/388	139/275/234
KOWOSB	4	92/178/146	171/292/232	53/112/89	217/433/348
BIGGS	6	2577/5078/4021	F	F	F
OSB2	11	398/713/629	1140/1630/1377	F	457/838/730
WATSON	20	9519/18204/14559	F	F	F
ROSEX	8	53/155/128	73/130/99	32/129/97	26/188/66
ROSEX	50	39/103/81	74/151/110	61/150/117	24/178/87
ROSEX	100	40/115/87	91/209/130	45/135/102	F
SINGX	8	453/879/698	156/249/205	314/605/483	174/394/275
PEN1	2	78/180/147	73/202/115	44/145/102	34/240/108
PEN2	4	403/1008/766	1112/1988/1502	717/1702/1280	124/377/274
PEN2	50	173/378/312	305/575/423	189/443/347	1620/3337/2658
VARDIM	2	11/13/13	8/11/11	11/13/13	4/9/8
VARDIM	50	19/37/37	39/56/56	19/37/37	15/34/34
TRIG	3	22/44/37	35/55/45	21/46/38	22/48/39
TRIG	50	53/95/87	61/94/76	47/80/75	48/86/77
TRIG	100	62/130/108	84/121/103	61/105/99	57/108/99
BV	3	17/34/27	34/46/41	16/32/26	F
BV	10	136/232/213	330/457/394	399/688/616	118/202/183
BV	500	719/903/902	234/318/272	537/690/689	314/416/415
BV	1 000	99/139/138	20/27/23	107/151/150	42/61/60
BV	2 000	2/4/3	2/4/3	2/4/3	2/4/3
IE	3	8/14/13	12/17/15	7/14/13	6/13/11
IE	50	6/13/12	12/17/15	6/13/12	6/12/11
IE	100	7/15/15	13/18/16	7/15/15	7/14/14

续表1

测试函数	维数/维	算法1/次 NI/NF/NG	算法2/次 NI/NF/NG	VPRP/次 NI/NF/NG	PRP/次 NI/NF/NG
IE	200	7/8/8	9/12/11	7/11/11	7/13/12
IE	500	9/12/12	13/17/16	8/12/12	8/15/14
IE	1 000	9/13/13	12/16/15	8/13/13	10/17/16
IE	2 000	9/13/13	14/18/17	9/14/14	8/13/13
TRID	3	17/31/27	24/35/29	17/28/25	14/25/21
TRID	50	29/36/36	36/49/43	30/38/38	29/43/43
TRID	100	35/52/52	35/50/42	31/37/37	31/51/51
TRID	200	38/51/51	37/50/44	33/52/52	31/49/48
TRID	500	34/44/44	35/47/42	34/43/43	33/50/48
TRID	1 000	40/48/47	37/51/45	37/58/58	33/48/48
TRID	2 000	42/48/48	38/52/46	34/43/43	34/57/57
BAND	3	10/14/14	12/17/15	10/14/14	15/19/19
BAND	50	36/44/44	30/39/36	31/41/41	22/83/34
BAND	100	32/44/43	29/38/35	26/36/36	23/42/38
BAND	200	31/41/41	27/37/34	33/40/40	F
BAND	500	48/57/57	28/38/35	27/36/36	21/35/33
BAND	1 000	81/90/90	30/40/37	25/35/35	22/38/37
BAND	2 000	35/44/44	29/39/36	46/54/54	24/42/39
LIN	2	1/3/3	3/5/5	1/3/3	1/3/3
LIN	50	1/3/3	1/3/3	1/3/3	1/3/3
LIN	500	1/3/3	1/3/3	1/3/3	1/3/3
LIN	1 000	1/3/3	1/3/3	1/3/3	1/3/3
LIN	2 000	1/3/3	1/3/3	1/3/3	1/3/3
LIN1	2	10/11/11	6/8/8	10/11/11	2/4/4
LIN1	10	1/3/3	3/5/5	1/3/3	1/3/3
LIN0	10	1/3/3	3/5/5	1/3/3	1/3/3

注：F表示计算失效或终止时未达计算精度.

本文对PRP公式进行非负修正并添加适当的调比因子.基于新参数公式的两个算法分别适用SWP和WWP步长线搜索，具有良好的收敛性.虽然新参数 β_k^N 不再满足性质(*)，但数值实验结果显示算法1优于表中其他方法，能解决更多的问题.同时，谱共轭梯度形式的算法2比标准共轭梯度形式的算法1收敛条件弱，但计算效果却没算法1好，如何调整两个方向调控参数，使谱共轭梯度算法达到更佳的计算效果，需要进一步的研究.

参考文献：

[1] 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，2001：30-42.

[2] GILBERT J C，NOCEDAL J. Global Convergence Properties of Conjugate Gradient Methods for Optimization [J]. SIAM J Optim，1992(2)：21-42.

[3] WEI Z X，YAO S W，LIU L Y. The Convergence Properties of Some New Conjugate Gradient Methods [J]. Appl Math Comput，2006，183(2)：1341-1350.

[4] YU G H，GUAN L T，LIU L Y. Global Convergence of Modified Polak-Ribiere-Polyak Conjugate Gradient Methods with Sufficient Descent Property [J]. J Ind Mang Optim，2008，4：565-579.

[5] 黎 勇.一类修正PRP共轭梯度法的全局收敛性及其数值试验结果[J].西南大学学报(自然科学版)，2011，33(11)：23-28.

[6] HUANG H D, LI Y J, WEI Z X. Global Convergence of a Modified PRP Conjugate Gradient Method [J]. J Math Res Exposition, 2010, 30(1): 141－148.

[7] DU X W, ZHANG P, MA W Y. Some Modified Conjugate Gradient Methods for Unconstrained Optimization [J]. J Comput Appl Math, 2016, 305(1): 92－114.

[8] JIANG X Z, JIAN J B. Two Modified Nonlinear Conjugate Gradient Methods with Disturbance Factors for Unconstrained Optimization [J]. Nonlinear Dyn, 2014, 77(1－2): 387－394.

[9] 简金宝，江羡珍，尹江华. 非线性共轭梯度法研究进展 [J]. 玉林师范学院学报，2016，37(2)：3－10.

[10] 林穗华，黄　海. 一个新的谱共轭梯度法 [J]. 工程数学学报，2014，31(6)：837－846.

[11] 李智群，林浦任，韦增欣. 一种新的求解无约束优化问题的谱共轭梯度法 [J]. 西南大学学报(自然科学版)，2016，38(7)：115－120.

[12] 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.

[13] 黎　勇，韦增欣. 一种自动充分下降的共轭梯度法 [J]. 西南师范大学学报(自然科学版)，2016，41(5)：36－40.

[14] 赛·闹尔再，张慧玲. 修正 LS 共轭梯度方法及其收敛性 [J]. 西南师范大学学报(自然科学版)，2016，41(7)：20－26.

[15] 张元园，张俊容，谢秉磊. 一种修正的 HS 共轭梯度法及其全局收敛性 [J]. 西南师范大学学报(自然科学版)，2012，37(8)：40－45.

[16] 林穗华. Wolfe 线搜索下的修正 FR 谱共轭梯度法 [J]. 山东大学学报(理学版)，2017，52(4)：6－12.

[17] MOREÈ J J, GARBOW B S, HILLSTROME K E. Testing Unconstrained Optimization Software [J]. ACM Trans Math Software, 1981, 7: 17－41.

Efficient Conjugate Gradient Algorithms Based on a Modified PRP Formula

LIN Sui-hua

College of Mathematics and Computer Science, Guangxi Normal University for Nationalities, Chongzuo Guangxi 532200, China

Abstract: A non-negative modified PRP conjugate gradient method parameter formula with disturbance is presented in this paper. Based on the new conjugate parameter formula, the corresponding conjugate gradient algorithm with the SWP line search satisfies the sufficient descent, and the corresponding spectral conjugate gradient algorithm with the WWP line search keeps decreasing. Under conventional assumptions, the global convergence of the two algorithms is proved. The results of a numerical experiment show that the new algorithms are effective.

Key words: unconstrained optimization; conjugate gradient method; spectral conjugate gradient method; global convergence

责任编辑　张　枸