DOI: 10.13718/j.cnki.xdzk.2017.05.021

一类具有充分下降性的混合型谱共轭梯度法[①]

王森森， 张俊容， 韩 信， 王逸云

西南大学 数学与统计学院，重庆 400715

摘要：首先基于共轭梯度法的下降性条件，提出了一类结合了 FR 法、WYL 法、PRP 法优点的充分下降的混合型谱共轭梯度法. 在 Wolfe 线搜索下用反证法证明了新的混合型谱共轭梯度法的全局收敛性. 最后通过数值算例，将本文算法与 WYL 法、FR 法进行比较，结果表明新算法在迭代次数与迭代总时间上均优于其他另外两种算法. 算法的全局收敛性和数值效果的优越性表明新算法是有效的.

关 键 词：无约束优化；谱共轭梯度法；全局收敛；Wolfe 线搜索

中图分类号：O224　　**文献标志码**：A　　**文章编号**：1673－9868(2017)05－0139－06

考虑如下的无约束优化问题

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

其中 $f(\boldsymbol{x})$ 是连续可微函数，其梯度向量 $\nabla \boldsymbol{f}(\boldsymbol{x})$ 记为 $\boldsymbol{g}(\boldsymbol{x})$. 共轭梯度法是求解上述无约束优化问题的一种十分有效的算法. 经典共轭梯度法的基本迭代格式为：

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

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

其中：$\boldsymbol{d}_k$ 为搜索方向；α_k 为步长因子，通常可以由 Wolfe 非精确的一维线搜索来确定.

$$\begin{gathered} f(\boldsymbol{x}_k+\alpha_k\boldsymbol{d}_k)\leqslant f(\boldsymbol{x}_k)+\delta\alpha_k\boldsymbol{g}_k^{\mathrm{T}}\boldsymbol{d}_k \\ \boldsymbol{g}(x_k+\alpha_k\boldsymbol{d}_k)^{\mathrm{T}}\boldsymbol{d}_k\geqslant\sigma\boldsymbol{g}_k^{\mathrm{T}}\boldsymbol{d}_k \end{gathered} \tag{4}$$

δ 和 σ 为满足 $0<\delta<\sigma<1$ 的常数；β_k 为标量参数. 关于 β_k 的选择有许多著名的公式，如：

$$\begin{aligned} &\beta_k^{\mathrm{HS}}=\frac{\boldsymbol{g}_k^{\mathrm{T}}\boldsymbol{y}_{k-1}}{\boldsymbol{d}_{k-1}{}^{\mathrm{T}}\boldsymbol{y}_{k-1}} &&\beta_k^{\mathrm{FR}}=\frac{\|\boldsymbol{g}_k\|^2}{\|\boldsymbol{g}_{k-1}\|^2}\\ &\beta_k^{\mathrm{PRP}}=\frac{\boldsymbol{g}_k^{\mathrm{T}}\boldsymbol{y}_{k-1}}{\|\boldsymbol{g}_{k-1}\|^2} &&\beta_k^{\mathrm{DY}}=\frac{\|\boldsymbol{g}_k\|^2}{\boldsymbol{d}_{k-1}{}^{\mathrm{T}}\boldsymbol{y}_{k-1}} \end{aligned} \tag{5}$$

其中 $\boldsymbol{y}_{k-1}=\boldsymbol{g}_k-\boldsymbol{g}_{k-1}$. 上述 4 个公式分别对应 4 种不同的共轭梯度方法. 每种方法又有很多变形，例如 Wei 等[1] 给出了 PRP 的一个变形，参数 β_k 取值为：

---

① 收稿日期：2015－10－29
基金项目：国家自然科学基金项目(11401487).
作者简介：王森森(1990－)，男，河南辉县人，硕士研究生，主要从事最优化理论、算法及应用研究.
通信作者：张俊容，博士，副教授.

$$\beta_k^{\mathrm{WYL}} = \frac{\|\boldsymbol{g}_k\|^2 - \frac{\|\boldsymbol{g}_k\|}{\|\boldsymbol{g}_{k-1}\|}\boldsymbol{g}_k^{\mathrm{T}}\boldsymbol{g}_{k-1}}{\|\boldsymbol{g}_{k-1}\|^2} \tag{6}$$

WYL 法继承了 PRP 法的一些优良性质，例如良好的数值表现，并且 Huang 等[2] 证明了此方法在强 Wolfe 线搜索下具有全局收敛性和下降性.

2001 年 Birgin 和 Martinez[3] 提出了一种谱共轭梯度法，其搜索方向定义如下：

$$\boldsymbol{d}_k = \begin{cases} -\boldsymbol{g}_k & k=1 \\ -\theta_k\boldsymbol{g}_k + \beta_k\boldsymbol{d}_{k-1} & k \geqslant 2 \end{cases} \tag{7}$$

其中 $\theta_k = \frac{\boldsymbol{s}_{k-1}{}^{\mathrm{T}}\boldsymbol{s}_{k-1}}{\boldsymbol{s}_{k-1}{}^{\mathrm{T}}\boldsymbol{y}_{k-1}}$ 为谱系数，$\beta_k = \frac{(\theta_k\boldsymbol{y}_{k-1} - \boldsymbol{s}_{k-1})^{\mathrm{T}}\boldsymbol{y}_{k-1}}{\boldsymbol{d}_{k-1}{}^{\mathrm{T}}\boldsymbol{y}_{k-1}}$，$\boldsymbol{s}_{k-1} = \boldsymbol{x}_k - \boldsymbol{x}_{k-1} = \alpha_{k-1}\boldsymbol{d}_{k-1}$. 但是 Birgin 和 Martinez 提出的谱共轭梯度法的搜索方向 $\boldsymbol{d}_k$ 不满足下降性，当然也不具有全局收敛性. 为了获得全局收敛性，Zhang 等[4] 在此基础上构造出了 FR 型谱共轭梯度法，其搜索方向 $\boldsymbol{d}_k$ 中 $\theta_k = \frac{\boldsymbol{d}_{k-1}{}^{\mathrm{T}}\boldsymbol{y}_{k-1}}{\|\boldsymbol{g}_{k-1}\|^2}$，$\beta_k = \beta_k^{\mathrm{FR}}$. 此方法的一个重要特点是满足充分下降条件

$$\boldsymbol{d}_k^{\mathrm{T}}\boldsymbol{g}_k \leqslant -t\|\boldsymbol{g}_k\|^2 \tag{8}$$

其中 $t>0$ 为常数. 此后 Wang，Cao 等[5-6] 提出了 CD 型谱共轭梯度法；Du 和 Liu 提出了 HS 型谱共轭梯度法[7]；Wan，Huang 等[8-9] 提出了 PRP 型的谱共轭梯度法. 上述方法都满足充分下降条件(8).

本文在上述文献的基础上，提出了一类具有充分下降性的混合型谱共轭梯度法，参数 β_k 为：

$$\beta_k^{\mathrm{WS}} = \frac{\|\boldsymbol{g}_k\|^2 - \max\left\{0, \frac{\|\boldsymbol{g}_k\|}{\|\boldsymbol{g}_{k-1}\|}\boldsymbol{g}_k^{\mathrm{T}}\boldsymbol{g}_{k-1}, \boldsymbol{g}_k^{\mathrm{T}}\boldsymbol{g}_{k-1}\right\}}{\|\boldsymbol{g}_{k-1}\|^2} \tag{9}$$

谱系数 $\theta_k = c + \beta_k^{\mathrm{WS}}\frac{\boldsymbol{g}_k^{\mathrm{T}}\boldsymbol{d}_{k-1}}{\|\boldsymbol{g}_k\|^2}$，$c$ 为大于 0 的参数，然后基于(9)式给出混合型谱共轭梯度法的算法，证明了算法的收敛性，并进行了数值实验.

1　混合型谱共轭梯度法的算法及其性质

本节中，首先给出混合型谱共轭梯度法(算法 1)，然后说明它所具有的一些性质.

算法 1　混合型谱共轭梯度法.

步骤 1　给定初始点 $\boldsymbol{x}_1$ 及精度 ε，计算 $\boldsymbol{g}_1$，若 $\|\boldsymbol{g}_1\| \leqslant \varepsilon$，停止. 否则，转步骤 2.

步骤 2　计算搜索方向 $\boldsymbol{d}_k$

$$\boldsymbol{d}_k = \begin{cases} -\boldsymbol{g}_k & k=1 \\ -\theta_k\boldsymbol{g}_k + \beta_k^{\mathrm{WS}}\boldsymbol{d}_{k-1} & k \geqslant 2 \end{cases} \tag{10}$$

$$\theta_k = c + \beta_k^{\mathrm{WS}}\frac{\boldsymbol{g}_k^{\mathrm{T}}\boldsymbol{d}_{k-1}}{\|\boldsymbol{g}_k\|^2} \tag{11}$$

步骤 3　由(4)式计算步长因子 α_k.

步骤 4　迭代计算 $\boldsymbol{x}_{k+1} = \boldsymbol{x}_k + \alpha_k\boldsymbol{d}_k$，$\boldsymbol{g}_{k+1} = \boldsymbol{g}(\boldsymbol{x}_{k+1})$. 若 $\|\boldsymbol{g}_{k+1}\| \leqslant \varepsilon$，则停止.

步骤 5　$k := k+1$，转步骤 2.

命题 1　由算法 1 产生的序列$\{\boldsymbol{g}_k\}$和$\{\boldsymbol{d}_k\}$满足下降性，即对任意 $k \geqslant 1$，都有 $\boldsymbol{g}_k^{\mathrm{T}}\boldsymbol{d}_k < 0$.

证　由算法 1 得

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

$$-\left(c+\beta_k^{\mathrm{WS}}\frac{\boldsymbol{g}_k^{\mathrm{T}}\boldsymbol{d}_{k-1}}{\|\boldsymbol{g}_k\|^2}\right)\|\boldsymbol{g}_k\|^2+\beta_k^{\mathrm{WS}}\boldsymbol{g}_k^{\mathrm{T}}\boldsymbol{d}_{k-1}=$$
$$-c\|\boldsymbol{g}_k\|^2<0$$

命题 2　对任意 $k\geqslant 1$，参数 β_k^{WS} 满足 $0\leqslant\beta_k^{\mathrm{WS}}\leqslant\dfrac{\|\boldsymbol{g}_k\|^2}{\|\boldsymbol{g}_{k-1}\|^2}$.

证　由参数 β_k^{WS} 的定义可知命题成立.

分析(9)式有：参数 β_k^{WS} 为 FR，WYL，PRP 3 者的混合. 下面分 3 种情况对此说明：

1）当 $\boldsymbol{g}_k^{\mathrm{T}}\boldsymbol{g}_{k-1}\leqslant 0$ 时，

$$\beta_k^{\mathrm{WS}}=\frac{\|\boldsymbol{g}_k\|^2}{\|\boldsymbol{g}_{k-1}\|^2}=\beta_k^{\mathrm{FR}}\tag{12}$$

2）当 $\boldsymbol{g}_k^{\mathrm{T}}\boldsymbol{g}_{k-1}>0$ 且 $\dfrac{\|\boldsymbol{g}_k\|}{\|\boldsymbol{g}_{k-1}\|}\leqslant 1$ 时，

$$\beta_k^{\mathrm{WS}}=\frac{\|\boldsymbol{g}_k\|^2-\boldsymbol{g}_k^{\mathrm{T}}\boldsymbol{g}_{k-1}}{\|\boldsymbol{g}_{k-1}\|^2}=\beta_k^{\mathrm{PRP}}\tag{13}$$

3）当 $\boldsymbol{g}_k^{\mathrm{T}}\boldsymbol{g}_{k-1}>0$ 且 $\dfrac{\|\boldsymbol{g}_k\|}{\|\boldsymbol{g}_{k-1}\|}>1$ 时，

$$\beta_k^{\mathrm{WS}}=\frac{\|\boldsymbol{g}_k\|^2-\dfrac{\|\boldsymbol{g}_k\|}{\|\boldsymbol{g}_{k-1}\|}\boldsymbol{g}_k^{\mathrm{T}}\boldsymbol{g}_{k-1}}{\|\boldsymbol{g}_{k-1}\|^2}=\beta_k^{\mathrm{WYL}}\tag{14}$$

此外由算法 1 的谱系数的选取可知，可通过选取不同的参数 c 来优化算法 1 的数值效果.

2　算法收敛性

为了研究算法的收敛性，下面给出一些基本假设：

(A1) 设 $f(\boldsymbol{x})$ 的水平集 $\Omega=\{\boldsymbol{x}\in\mathbb{R}^n \mid f(\boldsymbol{x})\leqslant f(\boldsymbol{x}_1)\}$ 有界，其中 $\boldsymbol{x}_1$ 为初始迭代点.

(A2) 存在 Ω 的某个邻域 Λ，使得 $f(\boldsymbol{x})$ 在该邻域上连续可微且梯度函数 $g(\boldsymbol{x})$ 满足 Lipschitz 条件，即存在常数 $L>0$，使得

$$\|g(\boldsymbol{x})-g(\boldsymbol{y})\|\leqslant L\|\boldsymbol{x}-\boldsymbol{y}\|\qquad\forall\,\boldsymbol{x},\boldsymbol{y}\in\Lambda\tag{15}$$

引理 1[10]　假设(A1)，(A2)成立，如果搜索方向 $\boldsymbol{d}_k$ 满足下降性，步长 α_k 由 Wolfe 非精确的一维线搜索确定，那么 $\sum\dfrac{(\boldsymbol{g}_k^{\mathrm{T}}\boldsymbol{d}_k)^2}{\|\boldsymbol{d}_k\|^2}<\infty$.

定理 1　如果条件(A1)，(A2)成立，算法 1 产生的序列 $\{\boldsymbol{g}_k\}$ 有

$$\liminf_{k\to\infty}\|\boldsymbol{g}_k\|=0\tag{16}$$

证　假设结论不成立，则存在常数 $\gamma>0$ 使得

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

由(10)式可得

$$\boldsymbol{d}_k+\theta_k\boldsymbol{g}_k=\beta_k^{\mathrm{WS}}\boldsymbol{d}_{k-1}\tag{18}$$

从而有

$$(\boldsymbol{d}_k+\theta_k\boldsymbol{g}_k)^{\mathrm{T}}(\boldsymbol{d}_k+\theta_k\boldsymbol{g}_k)=(\beta_k^{\mathrm{WS}})^2\|\boldsymbol{d}_{k-1}\|^2\tag{19}$$

将(19)式展开得

$$\|\boldsymbol{d}_k\|^2+2\theta_k\boldsymbol{d}_k^{\mathrm{T}}\boldsymbol{g}_k+\theta_k^2\|\boldsymbol{g}_k\|^2=(\beta_k^{\mathrm{WS}})^2\|\boldsymbol{d}_{k-1}\|^2\tag{20}$$

移项可得

$$\|\boldsymbol{d}_k\|^2=-2\theta_k\boldsymbol{d}_k^{\mathrm{T}}\boldsymbol{g}_k-\theta_k^2\|\boldsymbol{g}_k\|^2+(\beta_k^{\mathrm{WS}})^2\|\boldsymbol{d}_{k-1}\|^2\tag{21}$$

由命题 1 的证明，(21) 式可变为

$$\|\boldsymbol{d}_k\|^2 = 2c\theta_k\|\boldsymbol{g}_k\|^2 - \theta_k^2\|\boldsymbol{g}_k\|^2 + (\beta_k^{\mathrm{WS}})^2\|\boldsymbol{d}_{k-1}\|^2 \tag{22}$$

(22) 式两边同时除以$(\boldsymbol{g}_k^{\mathrm{T}}\boldsymbol{d}_k)^2$ 得

$$\frac{\|\boldsymbol{d}_k\|^2}{(\boldsymbol{g}_k^{\mathrm{T}}\boldsymbol{d}_k)^2} = \frac{\|\boldsymbol{d}_k\|^2}{c^2\|\boldsymbol{g}_k\|^4} = \frac{2\theta_k}{c\|\boldsymbol{g}_k\|^2} - \frac{\theta_k^2}{c^2\|\boldsymbol{g}_k\|^2} + \frac{(\beta_k^{\mathrm{WS}})^2}{c^2\|\boldsymbol{g}_k\|^4}\|\boldsymbol{d}_{k-1}\|^2 \tag{23}$$

进而有

$$\begin{aligned}\frac{\|\boldsymbol{d}_k\|^2}{(\boldsymbol{g}_k^{\mathrm{T}}\boldsymbol{d}_k)^2} &\leqslant \left(\frac{\|\boldsymbol{g}_k\|^2}{\|\boldsymbol{g}_{k-1}\|^2}\right)^2 \frac{\|\boldsymbol{d}_{k-1}\|^2}{c^2\|\boldsymbol{g}_k\|^4} - \frac{1}{c^2\|\boldsymbol{g}_k\|^2}(\theta_k^2 - 2c\theta_k + c^2 - c^2) = \\ &\frac{\|\boldsymbol{d}_{k-1}\|^2}{c^2\|\boldsymbol{g}_{k-1}\|^4} + \frac{1}{\|\boldsymbol{g}_k\|^2} - \frac{(\theta_k - c)^2}{c^2\|\boldsymbol{g}_k\|^2} \leqslant \\ &\frac{\|\boldsymbol{d}_{k-1}\|^2}{c^2\|\boldsymbol{g}_{k-1}\|^4} + \frac{1}{\|\boldsymbol{g}_k\|^2} \leqslant \\ &\sum_{i=1}^{k}\frac{1}{\|\boldsymbol{g}_k\|^2} + \frac{1}{c^2\|\boldsymbol{g}_1\|^2} \leqslant \frac{k}{\gamma^2} + l\left(l = \frac{1}{c^2\|\boldsymbol{g}_1\|^2}\right)\end{aligned} \tag{24}$$

综上所述可得

$$\sum \frac{(\boldsymbol{g}_k^{\mathrm{T}}\boldsymbol{d}_k)^2}{\|\boldsymbol{d}_k\|^2} \geqslant \gamma^2 \sum \frac{1}{k + \gamma^2 l} \to +\infty \tag{25}$$

与引理 1 矛盾.

3 数值实验

下面对算法 1 进行数值实验，测试函数中 Example 1 与 Example 2 的表达式为：

$$\begin{gathered} f(\boldsymbol{x}) = \sum_{i=1}^{n-1} 100(x_{1i+1} - x_{1i}^2) + (1 - x_{1n})^2 \\ \boldsymbol{x}_1 = (x_{11}, x_{12}, \cdots, x_{1n}) = (-1, -1.2, \cdots, -1, -1.2) \end{gathered} \tag{26}$$

$$\begin{gathered} f(\boldsymbol{x}) = \sum_{i=1}^{n-2} (x_{1i}^2 + 100x_{1i+1}^2 + 100x_{1i+2}^2) + (1 - x_{11})^2 \\ \boldsymbol{x}_1 = (x_{11}, x_{12}, \cdots, x_{1n}) = (3, 3, \cdots, 3, 3) \end{gathered} \tag{27}$$

其它测试对象均来自文献[11]中的函数. 实验在 PC 机上完成，PC 的配置如下：AMD A4－3300M CPU 1.90 GHz，2.00 GB内存. 程序用 Matlab 编写，运行环境为 Matlab R2010a. 算法中参数的选取为：$\delta = 0.001$，$\sigma = 0.9$. 终止条件为：$\|\boldsymbol{g}_k\| \leqslant 10^{-4}$；或者算法的迭代总时间超过 1 800 s；或者超过规定的迭代次数.

对于算法 1 中参数 c 的选取，以二维的香蕉函数为实验对象进行测试，香蕉函数的表达式如下：

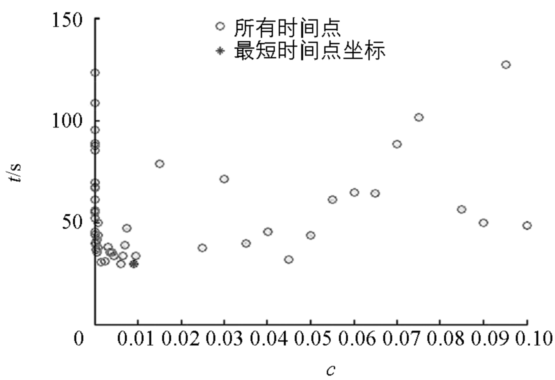

图 1 $c-t$ 图

$$f(\boldsymbol{x}) = 100(x_{12} - x_{11}^2)^2 + (1 - x_{11})^2 \qquad \boldsymbol{x}_1 = (x_{11}, x_{12}) = (-1, 1) \tag{28}$$

选取不同的 c 值($0 < c < 0.1$)，得到不同的迭代时间 t，作出 $c-t$ 图(图 1).

由图 1 可得：当 c 取 0.009 时，迭代时间最短，最短时间点坐标为(0.009，29.498 5). 另外通过数值实验发现，当 c 取值大于 0 且小于 10^6 时，算法 1 仍然有效；但当 c 取值为 0 或大于 10^6 的时候，算法 1 无效.

所以在本文后面的数值实验中选取 $c=0.009$，将算法 1 与 WYL 共轭梯度法及 FR 共轭梯度法进行比较，实验结果见表 1.

在表 1 中 i 的取值 0 和 1 分别表示迭代成功与不成功. 从表 1 可以看出无论是迭代次数、迭代的总时间、还是每次迭代的平均时间，算法 1 数值效果均优于其它另外两种算法. 此外算法 1 可以调控参数 c 来优化数值效果，具有一定的灵活性.

表 1　数值结果

函数名称	算法	维数 / 维	迭代次数 / 次	迭代时间 /s	$\Vert g_k \Vert$	每迭代一次的平均时间 /s	i
Example1	1	60	19	314	3.16×10^{-5}	16.5	0
	WYL	60	34	1190	5.89×10^{-5}	35	0
	FR	60	110	3123	8.57×10^{-5}	28.4	0
Example2	1	100	78	387	9.31×10^{-5}	4.96	0
	WYL	100	179	2866	7.83×10^{-5}	16	0
	FR	100	19	1898	4.81×10^{30}	99.9	1
POWER	1	50	426	1752	8.82×10^{-5}	4.11	0
	WYL	50	279	3601	3.45×10^{-1}	12.9	1
	FR	50	434	3605	9.35×10^{-1}	8.31	1
Diagonal 4	1	200	69	610	5.51×10^{-5}	8.84	0
	WYL	200	45	2127	1.44×10^{1}	47.3	1
	FR	200	69	1488	7.3×10^{-5}	21.6	0
Extended Penalty	1	20	28	474	8.5×10^{-5}	16.9	0
	WYL	20	28	679	5.27×10^{-5}	24.3	0
	FR	20	36	609	6.45×10^{-5}	16.9	0
Extended Rosenbrock	1	10	144	217	9.39×10^{-5}	1.51	0
	WYL	10	644	3606	1.37×10^{-1}	5.6	1
	FR	10	151	861	6.72×10^{1}	5.7	1
Extended White Holst	1	100	42	1535	2.97×10^{-5}	36.6	0
	WYL	100	57	3699	0.36×10^{1}	64.9	1
	FR	100	51	3685	2.27×10^{-2}	52.6	1
Perturbed Quadrtic	1	20	31	283	6.35×10^{-5}	9.11	0
	WYL	20	41	874	2.06×10^{-3}	21.3	1
	FR	20	41	441	1.87×10^{-2}	21.3	1
ENGVAL 1	1	50	91	1182	9.43×10^{-5}	13	0
	WYL	50	35	2137	3.02×10^{-5}	61	0
	FR	50	83	2682	7.61×10^{-5}	32.3	0

参考文献：

[1] WEI Z X, YAO S W, LIU Y X. The Convergence Properties of Some New Conjugate Gradient Methods [J]. Applied Mathematics and Computation, 2006, 183(2): 1341−1350.

[2] HUANG H, WEI Z X, YAO S W. The Proof of the Sufficient Descent Condtion of Wei-Yao-Liu Conjugate Gradient Method under the Strong Wolfe-Powell Line Search [J]. Applied Mathematics and Computation, 2007, 189(2): 1241−1245.

[3] BIRGIN E G, MARTINEZ J M. A Spectral Conjugate Gradient Method for Unconstraiend Optimization [J]. Applied Mathematics and Optimization, 2001, 43(2): 117−128.

[4] ZHANG L，ZHOU W J，Li D H. Global Convergence of a Modified Fletcher-Reeves Conjugate Gradient Method with Armijo-Type Line Search [J]. Numerical Mathematics，2006，104(4)：561－572.

[5] 王开荣，曹 伟，王银河. Armijo 型线搜素下的谱 CD 共轭梯度法 [J]. 山东大学学报(理学版)，2011，45(11)：104－108.

[6] CAO W，WANG K R，WANG Y H. Global Convergence of a Modified Spectral CD Conjugate Gradient Method [J]. Journal of Mathematical Research and Exposition，2011，31(2)：261－268.

[7] DU X L，LIU J K. Global Convergence of a Spectral HS Conjugate Gradient Method [J]. Procedia Engineering，2011，15；1487－1492.

[8] WAN Z，YANG Z L，WANG Y L. New Spectral PRP Conjugdte Gradient Method for Unconstrained Optimization [J]. Applied Mathematics Letters，2011，24(1)：16－22.

[9] 黄 海，林穗华. 一个 PRP 型共轭梯度法的收敛性 [J]. 西南大学学报(自然科学版)，2012，34(3)：22－29.

[10] 戴彧虹，袁亚湘. 非线性共轭梯度法 [M]. 上海：上海科学出版社，2000：10－13.

[11] ANDREI N. An Uconstrained Optimization Test Fuctions Collection [J]. Advanced Modelling and Optimzation，2008，10(1)：147－161.

A Mixed Spectral Conjugate Gradient Method with Sufficient Descent Property

WANG Sen-sen， ZHANG Jun-rong， HAN Xin， WANG Yi-yun

School of Mathematics and Statistics，Southwest University，Chongqing 400715，China

Abstract：First，a mixed sufficiently descent spectral conjugate gradient method is put forward，which satisfies the descent condition. Besides，the method possesses the advantages of FR，WYL and PPR. Then，the global convergence of the new hybrid spectral conjugate gradient method is proved with the reduction to absurdity under the Wolfe line search. Finally，the new algorithm and the existing WYL and FR algorithms are compared in their iterative times and computing time. The comparison results show that the new algorithm is superior to the other two algorithms. The global convergence and the numerical superiority of the new algorithm indicate that it is an effective algorithm which is worth studying.

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

责任编辑 张 枸