运筹学与控制论

DOI:10.11721/cqnuj20150201

求解无约束优化问题的两个谱共轭梯度法的全局收敛性*

林 穗 华
(广西民族师范学院 数学与计算机科学系,广西 崇左 532200)

摘要:谱共轭梯度法含有两个方向调控参数,是一种结合共轭梯度法和谱梯度法的无约束优化方法。本文建立新的共轭参数和谱参数,提出无约束优化问题的两个谱共轭梯度法,这两个新方法在精确线搜索下等价于 FR 共轭梯度法。然后,证明了算法 1 在 Wolfe 线搜索下和算法 2 在 Armijo 线搜索下的全局收敛性,并给出了算法的数值实验结果,验证了算法的有效性。

关键词:无约束优化;谱共轭梯度法;全局收敛性

中图分类号:O224　　**文献标志码:**A　　**文章编号:**1672-6693(2015)02-0001-06

考虑无约束优化问题

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

其中 $f:\mathbf{R}^n\to\mathbf{R}$ 为连续可微目标函数,其梯度函数记为 $g:\mathbf{R}^n\to\mathbf{R}^n$。共轭梯度法具有算法结构简单、计算存储空间需求少等优点,是求解大规模无约束优化问题(1)的有效方法之一,其迭代格式为

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

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

其中 α_k 为由线搜索确定的步长,d_k 为搜索方向,β_k 为共轭参数。

1964 年 Fletcher 和 Reeves 提出首个非线性共轭梯度法—FR 方法[1]以来,著名的 PRP 方法、HS 方法、DY 方法、LS 方法、CD 方法等共轭梯度法相继被推出并受到广泛的研究[2],它们相应的参数 β_k 公式为

$$\beta_k^{\mathrm{PRP}}=\frac{g_k^{\mathrm{T}}y_{k-1}}{\|g_{k-1}\|^2},\beta_k^{\mathrm{HS}}=\frac{g_k^{\mathrm{T}}y_{k-1}}{d_{k-1}^{\mathrm{T}}y_{k-1}},\beta_k^{\mathrm{LS}}=\frac{g_k^{\mathrm{T}}y_{k-1}}{-g_{k-1}^{\mathrm{T}}d_{k-1}},$$

$$\beta_k^{\mathrm{FR}}=\frac{\|g_k\|^2}{\|g_{k-1}\|^2},\beta_k^{\mathrm{DY}}=\frac{\|g_k\|^2}{d_{k-1}^{\mathrm{T}}y_{k-1}},\beta_k^{\mathrm{CD}}=\frac{\|g_k\|^2}{-g_{k-1}^{\mathrm{T}}d_{k-1}},$$

这里 $y_{k-1}=g_k-g_{k-1}$,$\|\cdot\|$ 为欧氏范数。其中,FR 方法、DY 方法和 CD 方法具有良好的收敛性质,而 PRP 方法、HS 方法和 LS 方法则数值性能优良。为寻求兼具良好收敛性与数值表现的算法,不少研究者还在上述 6 个经典参数公式的基础上,对 β_k 进行改进而推出各种不同效果的算法。如文献[3]结合 β_k^{FR} 和 β_k^{PRP} 推出 β_k^{WYL} 公式,文献[4]结合 β_k^{DY} 和 β_k^{HS} 推出 β_k^{WHT} 公式,文献[5]则讨论了新的参数类型 β_k^{new} 公式,分别如下

$$\beta_k^{\mathrm{WYL}}=\frac{g_k^{\mathrm{T}}\left(g_k-\frac{\|g_k\|}{\|g_{k-1}\|}g_{k-1}\right)}{\|g_{k-1}\|^2},\beta_k^{\mathrm{WHT}}=\frac{g_k^{\mathrm{T}}\left(g_k-\frac{g_k^{\mathrm{T}}g_{k-1}}{\|g_{k-1}\|^2}g_{k-1}\right)}{d_{k-1}^{\mathrm{T}}y_{k-1}},\beta_k^{\mathrm{new}}=\frac{g_k^{\mathrm{T}}(g_k-d_{k-1})}{d_{k-1}^{\mathrm{T}}y_{k-1}}。$$

2001 年 Birgin 和 Martinez 结合谱梯度法[6-8]与共轭梯度法提出了一种谱共轭梯度法[9],其方向迭代格式由(3)式推广为

$$d_k=\begin{cases}-g_k,k=1,\\-\theta_k g_k+\beta_k d_{k-1},k>1。\end{cases}\tag{4}$$

* **收稿日期:**2014-02-04　　**修回日期:**2014-12-20　　**网络出版时间:**2015-01-22 11:30
资助项目:广西高校科研项目(No. ZD2014143);广西重点培育学科(应用数学)建设项目(No. 桂教科研[2013]16);广西民族师范学院科研项目(No. 2013RCGG002)
作者简介:林穗华,女,副教授,研究方向为最优化方法及应用,E-mail: linsuihuah@163.com
网络出版地址:http://www.cnki.net/kcms/detail/50.1165.N.20150122.1130.003.html

其中 θ_k 为谱参数。虽然 Birgin-Martinez 谱共轭梯度法的搜索方向 d_k 不能满足下降条件 $g_k^{\mathrm{T}}d_k<0$,算法的收敛性不佳,但谱共轭梯度法迭代格式(4)式含有 2 个方向调控参数,为研究具备下降性的谱共轭梯度法提供了新的思路。文献[10-15]等通过构造适当的共轭参数和谱参数,给出了收敛性和数值结果良好的谱共轭梯度法,其中文献[12]讨论了取如下参数的谱共轭梯度法

$$\beta_k^{\mathrm{MLS}}=\frac{g_k^{\mathrm{T}}\left(g_k-\frac{g_k^{\mathrm{T}}g_{k-1}}{\|g_{k-1}\|^2}g_{k-1}\right)}{-g_{k-1}^{\mathrm{T}}d_{k-1}},\theta_k=1-\frac{g_k^{\mathrm{T}}d_{k-1}(1-\cos^2 r_k)}{g_{k-1}^{\mathrm{T}}d_{k-1}},$$

其中 r_k 为向量 g_k 与 g_{k-1} 的夹角。

受上述文献的启发,考虑如下参数的 2 个谱共轭梯度法

$$\beta_k=\frac{g_k^{\mathrm{T}}\left(g_k-\frac{g_k^{\mathrm{T}}d_{k-1}}{\|d_{k-1}\|^2}d_{k-1}\right)}{\|g_{k-1}\|^2},\theta_k^{(1)}=\frac{d_{k-1}^{\mathrm{T}}y_{k-1}}{\|g_{k-1}\|^2},\theta_k^{(2)}=\theta_k^{(1)}-\frac{g_k^{\mathrm{T}}d_{k-1}}{\|g_{k-1}\|^2}\cos^2 r_k,\tag{5}$$

其中 r_k 为向量 g_k 与 d_{k-1} 的夹角。显然在精确线搜索下有 $g_k^{\mathrm{T}}d_{k-1}=0$,从而 $\beta_k=\beta_k^{\mathrm{FR}},\theta_k^{(1)}=\theta_k^{(2)}=1$,即若采用精确搜索,则(2)、(4)、(5)式对应的谱共轭梯度法等价于 FR 共轭梯度法。以下分别结合 Wolfe 线搜索和标准 Armijo 线搜索建立新的谱共轭梯度法算法,并分析算法的收敛性质。

1 两个谱共轭梯度算法及其下降性

算法 1 步骤 1,给定初值 $x_1\in\mathbf{R}^n,\delta\in\left(0,\frac{1}{2}\right),\sigma\in(\delta,1),\varepsilon\geqslant 0,d_1:=-g_1,k:=1$。若 $\|g_k\|\leqslant\varepsilon$,停止。

步骤 2,计算满足以下 Wolfe 线搜索准则的步长 α_k

$$f(x_k+\alpha_k d_k)-f_k\leqslant\delta\alpha_k g_k^{\mathrm{T}}d_k,\tag{6}$$

$$g(x_k+\alpha_k d_k)^{\mathrm{T}}d_k\geqslant\sigma g_k^{\mathrm{T}}d_k。\tag{7}$$

步骤 3,计算 $x_{k+1}=x_k+\alpha_k d_k$。若 $\|g_{k+1}\|\leqslant\varepsilon$,停止。

步骤 4,由(5)式计算参数 β_{k+1} 及 $\theta_{k+1}=\theta_{k+1}^{(1)}$,由(4)式计算 d_{k+1}。

步骤 5,$k:=k+1$,转入步骤 2。

算法 2 步骤 1,给定初值 $x_1\in\mathbf{R}^n,\rho\in(0,1),\delta\in(0,1),\varepsilon\geqslant 0,d_1:=-g_1,k:=1$。若 $\|g_k\|\leqslant\varepsilon$,停止。

步骤 2,计算满足标准 Armijo 线搜索准则(6)式和(8)式的步长 α_k

$$\alpha_k=\max\{\rho^j,j=0,1,2,\cdots\}。\tag{8}$$

步骤 3,计算 $x_{k+1}=x_k+\alpha_k d_k$。若 $\|g_{k+1}\|\leqslant\varepsilon$,停止。

步骤 4,由(5)式计算参数 β_{k+1} 及 $\theta_{k+1}=\theta_{k+1}^{(2)}$,由(4)式计算 d_{k+1}。

步骤 5,$k:=k+1$,转入步骤 2。

以下均假设 $\|g_k\|\neq 0$,否则算法已找到稳定点而停止。下面引理 1 说明算法 1 产生的搜索方向 d_k 满足下降性。

引理 1 设 $\{g_k,d_k,\beta_k\}$ 为算法 1 生成的序列,则

$$g_k^{\mathrm{T}}d_k<0,0\leqslant\beta_{k+1}\leqslant\frac{g_{k+1}^{\mathrm{T}}d_{k+1}}{g_k^{\mathrm{T}}d_k},\forall k\geqslant 1。\tag{9}$$

证明 当 $k=1$ 时,$d_1=-g_1$,则有 $g_1^{\mathrm{T}}d_1=-\|g_1\|^2<0$。

对 $k\geqslant 1$,假设 $g_k^{\mathrm{T}}d_k<0$,由(7)式易得 $d_k^{\mathrm{T}}y_k\geqslant(\sigma-1)g_k^{\mathrm{T}}d_k>0$。下证 $g_{k+1}^{\mathrm{T}}d_{k+1}<0$。

$$\frac{\|g_{k+1}\|^2}{\|g_k\|^2}\geqslant\beta_{k+1}=\frac{\|g_{k+1}\|^2}{\|g_k\|^2}\left[1-\frac{(g_{k+1}^{\mathrm{T}}d_k)^2}{\|g_{k+1}\|^2\|d_k\|^2}\right]=\frac{\|g_{k+1}\|^2}{\|g_k\|^2}(1-\cos^2 r_{k+1})\geqslant 0。\tag{10}$$

由(4)式知 $d_{k+1}=-\theta_{k+1}g_{k+1}+\beta_{k+1}d_k$,两端与 g_{k+1} 作内积。若 $\beta_{k+1}>0$,则可得

$$g_{k+1}^{\mathrm{T}}d_{k+1}=\frac{\|g_{k+1}\|^2}{\|g_k\|^2}(-d_k^{\mathrm{T}}y_k)+\beta_{k+1}g_{k+1}^{\mathrm{T}}d_k\leqslant\beta_{k+1}(-d_k^{\mathrm{T}}y_k+g_{k+1}^{\mathrm{T}}d_k)=\beta_{k+1}g_k^{\mathrm{T}}d_k<0。\tag{11}$$

若 $\beta_{k+1}=0$,则结合 $d_k^{\mathrm{T}}y_k>0$ 可得 $g_{k+1}^{\mathrm{T}}d_{k+1}=-\frac{d_k^{\mathrm{T}}y_k}{\|g_k\|^2}\|g_{k+1}\|^2<0$。从而,由数学归纳法知,$\forall k\geqslant 1,g_k^{\mathrm{T}}d_k<0$ 成立。

再由(10)、(11)式可得 $0\leqslant\beta_{k+1}\leqslant\frac{g_{k+1}^{\mathrm{T}}d_{k+1}}{g_k^{\mathrm{T}}d_k}$。　证毕

下面引理 2 说明算法 2 产生的搜索方向 d_k 满足充分下降性。

引理 2　设$\{g_k,d_k,\beta_k\}$为算法 2 生成的序列,则

$$g_k^{\mathrm{T}}d_k=-\|g_k\|^2<0,\forall k\geqslant1。\tag{12}$$

证明　若 $k=1$,则 $d_1=-g_1$,有 $g_1^{\mathrm{T}}d_1=-\|g_1\|^2<0$。假设 $g_k^{\mathrm{T}}d_k=-\|g_k\|^2<0$,则由(4)式可得

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

$$-\frac{(d_k^{\mathrm{T}}y_k-g_{k+1}^{\mathrm{T}}d_k\cos^2r_{k+1})}{\|g_k\|^2}\|g_{k+1}\|^2+\frac{\|g_{k+1}\|^2(1-\cos^2r_{k+1})}{\|g_k\|^2}g_{k+1}^{\mathrm{T}}d_k=\frac{\|g_{k+1}\|^2}{\|g_k\|^2}g_k^{\mathrm{T}}d_k=-\|g_{k+1}\|^2<0。$$

由数学归纳法知引理 2 成立。　证毕

2 算法的全局收敛结果

为了证明算法 1 和算法 2 的全局收敛性,要求目标函数 $f(x)$满足如下假设 H

i) $f(x)$在水平集 $\Omega=\{x\in\mathbf{R}^n\mid f(x)\leqslant f(x_1)\}$上有下界。

ii) $f(x)$的梯度函数 $g(x)$在 Ω 上 Lipschitz 连续,即$\exists L>0$,使$\|g(y)-g(x)\|\leqslant L\|y-x\|,\forall x,y\in\Omega$。

引理 3　若假设 H 成立,$\{g_k,d_k\}$为算法 1 生成的序列,则

$$\sum_{k\geqslant1}\frac{(g_k^{\mathrm{T}}d_k)^2}{\|d_k\|^2}<+\infty。\tag{13}$$

证明　由(7)式及假设条件 H 中的 ii),可得 $L\alpha_k\|d_k\|^2\geqslant d_k^{\mathrm{T}}(g_{k+1}-g_k)\geqslant(\sigma-1)g_k^{\mathrm{T}}d_k$,从而可得 $\alpha_k\geqslant\frac{(\sigma-1)}{L}\frac{g_k^{\mathrm{T}}d_k}{\|d_k\|^2}$。

由(6)、(9)式和假设条件 H 中的 i),可知$\{f_k\}$为单调递减的收敛数列,再结合上式可得

$$f_k-f_{k+1}\geqslant-\delta\alpha_kg_k^{\mathrm{T}}d_k\geqslant\frac{\delta(1-\sigma)}{L}\frac{(g_k^{\mathrm{T}}d_k)^2}{\|d_k\|^2}。$$

对上式两端分别求和,并利用$\{f_k\}$的收敛性,可知(13)式成立。　证毕

引理 4　若假设 H 成立,$\{g_k,d_k\}$为算法 2 生成的序列,则

$$\sum_{k\geqslant1}\frac{\|g_k\|^4}{\|d_k\|^2}<+\infty。\tag{14}$$

证明　由 Armijo 线搜索准则可知,$\rho^{-1}\alpha_k$ 不满足步长条件(6)式,即

$$f(x_k+\rho^{-1}\alpha_kd_k)-f_k>\delta\rho^{-1}\alpha_kg_k^{\mathrm{T}}d_k。\tag{15}$$

由微分中值定理,Cauchy-Schwartz 不等式及假设条件 H 中的 ii)知,$\exists t_k\in(0,1)$,使

$$f(x_k+\rho^{-1}\alpha_kd_k)-f_k=\rho^{-1}\alpha_k[g(x_k+t_k\rho^{-1}\alpha_kd_k)-g_k]^{\mathrm{T}}d_k+\rho^{-1}\alpha_kg_k^{\mathrm{T}}d_k\leqslant\rho^{-2}\alpha_k^2L\|d_k\|^2+\rho^{-1}\alpha_kg_k^{\mathrm{T}}d_k。\tag{16}$$

结合(12)、(15)和(16)式,可得

$$\alpha_k\geqslant\frac{(1-\delta)\rho}{L}\frac{\|g_k\|^2}{\|d_k\|^2}。\tag{17}$$

由(12)式可得$\|g_k\|^2=|g_k^{\mathrm{T}}d_k|\leqslant\|g_k\|\cdot\|d_k\|$,从而$\frac{\|g_k\|}{\|d_k\|}\leqslant1$。又由(8)式知 $0<\alpha_k\leqslant1$,结合(17)式,可知对$\forall k\geqslant1$,有

$$\alpha_k\geqslant\min\left\{1,\frac{(1-\delta)\rho}{L}\right\}\frac{\|g_k\|^2}{\|d_k\|^2}。\tag{18}$$

类似引理 3 的证明过程,利用(12)和(18)式易得(14)式成立。　证毕

定理 1　若假设 H 成立,$\{g_k\}$为算法 1 生成的序列,则$\liminf_{k\to\infty}\|g_k\|=0$。

证明　若定理不成立,则存在常数 $r>0$,使任意 $k\geqslant1$ 有

$$\|g_k\|\geqslant r。\tag{19}$$

由(4)式移项得 $d_k+\theta_kg_k=\beta_kd_{k-1}$,两端取模平方并利用(9)式,可得

$$\|d_k\|^2=\beta_k^2\|d_{k-1}\|^2-2\theta_kg_k^{\mathrm{T}}d_k-\theta_k^2\|g_k\|^2\leqslant\frac{(g_k^{\mathrm{T}}d_k)^2}{(g_{k-1}^{\mathrm{T}}d_{k-1})^2}\|d_{k-1}\|^2-2\theta_kg_k^{\mathrm{T}}d_k-\theta_k^2\|g_k\|^2。\tag{20}$$

(20)式两端同时除以$(g_k^{\mathrm{T}}d_k)^2$,可得

$$\frac{\| d_k \|^2}{(g_k^{\mathrm{T}} d_k)^2} \leqslant \frac{\| d_{k-1} \|^2}{(g_{k-1}^{\mathrm{T}} d_{k-1})^2} - \frac{2\theta_k}{g_k^{\mathrm{T}} d_k} - \frac{\theta_k^2 \| g_k \|^2}{(g_k^{\mathrm{T}} d_k)^2} =$$

$$\frac{\| d_{k-1} \|^2}{(g_{k-1}^{\mathrm{T}} d_{k-1})^2} + \frac{1}{\| g_k \|^2} - \left(\frac{1}{\| g_k \|} + \frac{\theta_k \| g_k \|}{g_k^{\mathrm{T}} d_k} \right)^2 \leqslant \frac{\| d_{k-1} \|^2}{(g_{k-1}^{\mathrm{T}} d_{k-1})^2} + \frac{1}{\| g_k \|^2}。$$

由上式递推,并利用 $d_1 = -g_1$ 和(19)式,可得 $\frac{\| d_k \|^2}{(g_k^{\mathrm{T}} d_k)^2} \leqslant \frac{\| d_1 \|^2}{(g_1^{\mathrm{T}} d_1)^2} + \sum_{i=2}^{k} \frac{1}{\| g_i \|^2} = \sum_{i=1}^{k} \frac{1}{\| g_i \|^2} \leqslant \frac{k}{r^2}$,从而可得 $\frac{(g_k^{\mathrm{T}} d_k)^2}{\| d_k \|^2} \geqslant \frac{r^2}{k}$。

上式两端分别求和,可得 $\sum_{k \geqslant 1} \frac{(g_k^{\mathrm{T}} d_k)^2}{\| d_k \|^2} \geqslant r^2 \sum_{k \geqslant 1} \frac{1}{k} = +\infty$,这与引理3矛盾,所以定理成立。 证毕

定理2 若假设H成立,$\{g_k\}$为算法2生成的序列,则$\lim\limits_{k \to \infty} \inf \| g_k \| = 0$。

证明 由(5)和(12)式,可得 $0 \leqslant \beta_k \leqslant \frac{\| g_k \|^2}{\| g_{k-1} \|^2}$,结合(4)式得到

$$\| d_k \|^2 = \beta_k^2 \| d_{k-1} \|^2 - 2\theta_k g_k^{\mathrm{T}} d_k - \theta_k^2 \| g_k \|^2 \leqslant \frac{\| g_k \|^4}{\| g_{k-1} \|^4} \| d_{k-1} \|^2 + (2\theta_k - \theta_k^2) \| g_k \|^2。$$

上式两端同时除以 $\| g_k \|^4$,可得

$$\frac{\| d_k \|^2}{\| g_k \|^4} \leqslant \frac{\| d_{k-1} \|^2}{\| g_{k-1} \|^4} - \frac{(\theta_k - 1)^2}{\| g_{k-1} \|^2} + \frac{1}{\| g_k \|^2} \leqslant \frac{\| d_{k-1} \|^2}{\| g_{k-1} \|^4} + \frac{1}{\| g_k \|^2}。$$

上式递推,并利用 $d_1 = -g_1$,可得

$$\frac{\| d_k \|^2}{\| g_k \|^4} \leqslant \frac{\| d_1 \|^2}{\| g_1 \|^4} + \sum_{i=2}^{k} \frac{1}{\| g_i \|^2} = \sum_{i=1}^{k} \frac{1}{\| g_i \|^2}。$$

类似定理1,利用反证法易得定理成立。 证毕

3 数值实验

本节对将FR方法分别与算法1、算法2进行对比数值实验,测试函数源于文献[16],运行环境为PC 2.80 GHz CPU,1 G RAM,Matlab 6.5+Windows XP操作系统。Wolfe线搜索下的计算参数为 $\delta = 0.01$,$\sigma = 0.1$,$\varepsilon = 10^{-5}$,终止条件为 $\| g_k \| \leqslant \varepsilon$,或迭代次数超过999;Armijo线搜索下的计算参数为 $\delta = 0.001$,$\rho = 0.8$,$\varepsilon = 10^{-5}$,终止条件为 $\| g_k \| \leqslant \varepsilon$,或迭代次数超过10 000。表1中,FR方法采用Wolfe线搜索,NI/NF/NG分别代表算法迭代次数、目标函数迭代次数、梯度迭代次数。表2中,FR方法采用Armijo线搜索,NI/NF/time分别代表算法迭代次数、目标函数迭代次数、CPU时间(单位:s)。

表1 算法1与FR方法的数值结果比较

Tab.1 The numerical results contrast between algrithm 1 and FR method

测试问题	维数	算法1 NI/NF/NG	FR方法 NI/NF/NG
Rosenbrock *	2	41/222/63	84/148/122
Jennrich and Sampson	2	23/46/36	21/46/37
Helical Valley	3	83/294/125	60/150/87
Gaussian	3	7/11/9	4/8/6
Extended Rosenbrock *	100	62/163/93	243/1 310/328
Penalty II *	50	147/658/255	200/784/335
Variably Dimensioned *	50	11/43/36	12/44/37
Trigonometric *	50	70/193/85	207/731/227
Discrete Boundary Value *	10	139/352/183	640/2 209/740
Discrete Integral Equation *	500	6/11/10	7/11/11
Broyden Tridiagonal *	200	30/48/40	61/117/66
Broyden Banded *	3	8/13/11	24/28/27
CPU时间/s		9.344 0e+000	1.382 9e+001

表 2　算法 2 与 FR 方法的数值结果比较

Tab. 2　The numerical results contrast between algrithm 2 and FR method

测试问题	维数	算法 2 NI/NF/time	FR 方法 NI/NF/time
Rosenbrock *	2	164/1682/3.750 0e−001	217/2 785/7.040 0e−001
Jennrich and Sampson *	2	43/303/9.400 0e−002	60/465/1.250 0e−001
Helical Valley *	3	101/942/2.970 0e−001	141/1 369/4.220 0e−001
Gaussian	3	28/85/4.700 0e−002	9/23/1.500 0e−002
Extended Rosenbrock *	100	184/1 872/1.172 0e+000	282/3 819/1.9370e+000
Penalty II *	50	213/1 751/1.891 0e+000	444/3 773/3.250 0e+000
Variably Dimensioned	50	20/455/1.560 0e−001	16/339/1.410 0e−001
Trigonometric *	50	125/128/2.660 0e−001	613/3 041/5.312 0e+000
Discrete Boundary Value *	10	139/640/1.880 0e−001	311/2 005/5.780 0e−001
Discrete Integral Equation	500	33/67/1.542 200e+001	11/21/4.703 000e+000
Broyden Tridiagonal *	200	74/525/5.780 0e−001	920/8 633/9.468 0e+000
Broyden Banded *	3	67/405/1.090 0e−001	94/730/2.190 0e−001
CPU 时间/s		2.060 0e+001	2.690 0e+001

表 1、表 2 中加“*”的测试问题分别表示算法 1、算法 2 的数值结果优于 FR 算法。从 CPU 时间看,采用 Wolfe 线搜索时算法 1 优于 FR 方法,采用 Arimijo 线搜索时算法 2 优于 FR 方法,算法 1 优于算法 2。本文给出的两个谱共轭梯度法具有良好的收敛性质,表中的数值实验结果显示这些算法是有效的,适合于求解非线性无约束优化问题。

参考文献:

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

[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,2001.
Dai Y H,Yuan Y X. Nonlinear conjugate gradient methods [M]. Shanghai:Shanghai Science and Tech-nology Publisher,2001.

[3] Wei Z X,Yao S W,Liu L Y. The convergence properties of some new conjugate gradient methods[J]. Applied Mathematics and Computation,2006,183:1341-1350.

[4] Wei Z X,Huang H D,Tao Y R. A modified Hestenes-Stiefel conjugate gradient method and its convergence[J]. Journal of Mathematical Research & Exposition,2010,30(2):297-308.

[5] 郑希锋,田志远,宋立温. Wolfe 线搜索下一类混合共轭梯度法的全局收敛性[J]. 运筹学学报,2009,13(2):18-24.
Zheng X F,Tian Z Y,Song L W. The global convergence of a mixed conjugate gradient method with the Wolfe line search[J]. OR Transactions,2009,13(2):18-24.

[6] Barzilai J,Borwein J M. Two-point step size gradient methods[J]. IMA J Num Anal,1988(8):141-148.

[7] Raydan M. The Barzilai and Borwein gradient method for the large scale unconstrained minimization problem[J]. SIAM J Optim,1997,7(1):26-33.

[8] Birgin E G,Martimez J M. A spectral conjugate gradient method for unconstrained optimization[J]. Appl Math Optim,2001,43(2):117-128.

[9] 黄海. 无约束优化的修正谱梯度法[J]. 四川师范大学学报:自然科学版,2012,35(3):349-354.
Huang H. Modified spectral gradient method for unconstrained optimization[J]. Journal of Sichuan Normal University:Natural Science,2012,35(3):349-354.

[10] Zhang L,Zhou W J,Li D H. Global convergence of a modified Fletcher-Reeves conjugate gradient method with Armijo-type line search [J]. Numerische Mathematik,2006,104:561-572.

[11] Liu J K. Global convergence of a new spectral PRP conjugate gradient method[J]. Journal of Applied Mathematics and Informatics,2011,29(6):1303-1039.

[12] 胡鹏,杜学武,郭翠峰. 一类修正 LS 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,2012,29(5):13-15.
Hu P,Du X W,Guo C F. Global convergence of modified LS spectral conjugate gradient method[J]. Journal of Chongqing Normal University: Natural Science, 2012, 29 (5):13-15.

[13] 孟继东,杜学武. Armijo 型线搜索一个修正 LS 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,

2012,29(6):6-8.
Meng J D,Du X W. Global convergence of a modified LS conjugate gradient method with an Armijo-Type Line Search[J]. Journal of Chongqing Normal University: Natural Science,2012,29(6):6-8.
[14] 陈龙卫,夏福全,贾朝勇.一类充分下降的谱共轭梯度法[J].重庆师范大学学报:自然科学版,2013,30(4):10-14.
Chen L W,Xia F Q,Jia C Y. A sufficient descent spectral conjugate gradient method[J]. Journal of Chongqing Normal University: Natural Science,2013,30(4):10-14.
[15] 林穗华.求解无约束优化问题的新谱共轭梯度法及其收敛性[J].经济数学,2013,30(4):33-37.
Lin S H. A new spectral conjugate gradient method for unconstrained optimization and its convergence[J]. Journal of Quantitative Economics,2013,30(4):33-37.
[16] Morè J J, Garbow B S, Hillstrom K E. Testing unconstrained optimization software[J]. ACM Transactions on Mathematical Software,1981,7(1):17-41.

Operations Research and Cybernetics

Global Convergence of Two Spectral Conjugate Gradient Methods for Unconstrained Optimization

LIN Suihua
(Department of Mathematics and Computer Science, Guangxi Normal University for Nationalities, Chongzuo Guangxi 532200, China)

Abstract: Spectral conjugate gradient method contains two directions regulatory parameters is a kind of method for unconstrained optimization that combines conjugate gradient method with spectral gradient method. In this paper, based on the new conjugate parameters and spectral parameters, two spectral conjugate gradient methods are proposed; the corresponding methods are equivalent to the FR conjugate gradient method when the line search is exact. Moreover, the global convergence of algorithm 1 with Wolfe line search is proved, the global convergence of algorithm 2 with standard Armijo line search is proved. The given numerical results show that the new methods are efficient.

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

(责任编辑 黄 颖)