文章编号:1000－2367(2018)04－0107－06 DOI:10.16366/j.cnki.1000－2367.2018.04.017

一类新的 WYL 型共轭梯度法及其全局收敛性

董晓亮[a],李卫军[b]

(北方民族大学 a.数学与信息科学学院;b.网络信息中心,银川 750021)

摘 要:共轭梯度法是求解大规模无约束优化问题的一类重要的优化方法,该方法具有全局收敛性和存储量小的优点. 提出了一类修正的 Wei-Yao-Liu 型三项共轭梯度法,该方法扩大了其中参数的选择范围,在强 Wolfe 搜索下满足充分下降条件和全局收敛性. 初步的数值试验说明了算法的有效性.

关键词:共轭梯度法;全局收敛性;充分下降条件

中图分类号:O224 **文献标志码**:A

共轭梯度法是求解无约束优化问题 $\min\limits_{x\in R^n} f(x)$ 的一类重要方法,其迭代公式为:

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

$$d_1=-g_1,d_k=-g_k+\beta_k d_{k-1},(k\geq 2), \tag{2}$$

其中,$g_k=\nabla f(x_k)$,d_k 是搜索方向,β_k 是参数,步长 α_k 满足的 Wolfe 搜索:

$$f(x_k+\alpha_k d_k)-f(x_k)\leq \rho\alpha_k g_k^{\mathrm{T}} d_k, \tag{3}$$

$$g(x_k+\alpha_k d_k)^{\mathrm{T}} d_k\geq \sigma g_k^{\mathrm{T}} d_k. \tag{4}$$

其中 $0<\rho<\sigma<1$. 经典方法有 FR 方法、HS 方法、PRP 方法、DY 方法、DL 方法和 HZ 方法,其收敛性分析详见文献[1－6].

充分下降条件在下降算法的收敛性分析中的作用很重要[7－8],也即(5)式成立

$$-d_k^{\mathrm{T}} g_k\geq c_1\|g_k\|^2(\forall k\in\mathbf{N},c_1>0\text{ 是某个常数}). \tag{5}$$

在上述方法中,HS 方法和 PRP 方法数值性能最为优越,但对于非凸函数,即使用精确搜索也无法全局收敛.文献[2]证明了在精确搜索下对于一致凸函数极小化时,当 s_k 趋于零 PRP 方可以建立全局收敛性,但同时 Powell 在该文也举出反例说明对于非一致凸函数时 PRP 方法可能在几个稳定点附近循环,无法保证 PRP 方法的收敛性.

近来,文献[9]提出了 MHS,WYL 和 MLS 等方法,对应的参数 β_k 为

$$\beta_k^{\mathrm{MHS}}=\frac{g_k^{\mathrm{T}}\overline{y_{k-1}}}{d_{k-1}^{\mathrm{T}}y_{k-1}},\beta_k^{\mathrm{WYL}}=\frac{g_k^{\mathrm{T}}\overline{y_{k-1}}}{\|g_{k-1}\|^2},\beta_k^{\mathrm{MLS}}=-\frac{g_k^{\mathrm{T}}\overline{y_{k-1}}}{d^{\mathrm{T}_{k-1}}g_{k-1}}, \tag{6}$$

其中 $\overline{y_{k-1}}=g_k-\frac{\|g_k\|}{\|g_{k-1}\|}g_{k-1}$. 在曲率条件(10)式中,他们做出参数 σ 的下述限制:$\sigma\in\left[0,\frac{1}{3}\right]$,$\sigma\in\left[0,\frac{1}{4}\right]$ 和 $\sigma\in\left[0,\frac{1}{4}\right]$ 的限制,并证明了全局收敛性.其相关研究进展详见文献[10－14].

本文结合三项共轭梯度法的特点,考虑对 WYL 方法做出适当修正,扩大参数 σ 的适用范围,同时满足充分下降条件和全局收敛性。具体地,给出的搜索方向为

收稿日期:2017-10-06;**修回日期**:2018-04-28.

基金项目:国家自然科学基金青年基金(11601012);宁夏自然科学基金(NZ17103);宁夏高校科研基金(NGY2016134;NGY2016143);北方民族大学科研项目(2016SXKY05,2016SXKY06);北方民族大学重大专项项目(ZDZX201804).

作者简介(通信作者):董晓亮(1981－),男,甘肃静宁人,北方民族大学副教授,博士,研究方向为数学规划,E-mail:dongxl@nun.edu.cn.

$$d_k^{\mathrm{MWYL}}=-g_k+\beta_k^{\mathrm{MWYL}}d_{k-1}+\theta_k^{\mathrm{MWYL}}g_{k-1}. \tag{7}$$

其中,参数 $c>0$ 为常数,以及

$$\beta_k^{\mathrm{MWYL}}=\frac{g_k^{\mathrm{T}}\left(g_k-\frac{\|g_k\|}{\|g_{k-1}\|}g_{k-1}\right)}{\|g_{k-1}\|^2+\max\{\alpha_{k-1},c\}\|d_{k-1}\|^2}, \tag{8}$$

$$\theta_k^{\mathrm{MWYL}}=\frac{g_k^{\mathrm{T}}d_{k-1}}{\|g_{k-1}\|^2+\max\{\alpha_{k-1},c\}\|d_{k-1}\|^2}\cdot\frac{\|g_k\|}{\|g_{k-1}\|}. \tag{9}$$

1 算法及其充分下降条件

基于上文提出的搜索方向(7)式,本节给出新的共轭梯度法,约记为 MHZ 方法:

算法 1

步骤 1 选定初始步 x_1,参数 $\rho\in(0,1)$,$\delta\in(\rho,1)$,$c\geq\frac{\sigma^2}{4}$ 及 $\varepsilon>0$,置 $k=1$;

步骤 2 计算 $g_k=\nabla f(x_k)$.若 $\|g_k\|<\varepsilon$,则停止计算;否则由(7)式计算 d_k;

步骤 3 按照强 Wolfe 搜索计算步长 α_k:即满足不等式(3)式和

$$|g(x_k+\alpha_k d_k)^{\mathrm{T}}d_k|\leq\sigma|g_k^{\mathrm{T}}d_k|; \tag{10}$$

步骤 4 计算 $x_{k+1}=x_k+\alpha_k d_k$,置 $k=k+1$,转步骤 2.

下面,给出一个引理,说明算法 1 中的搜索方向 d_k 可满足充分下降条件.

引理 1 若 d_k 由(7)式、算法 1 定义,则 d_k 满足充分下降条件 $-d_k^{\mathrm{T}}g_k\geq\left(\frac{\sigma}{2\sqrt{c}}-1\right)\|g_k\|^2$.

证明 $k=1$ 时,易得 $d_1{}^{\mathrm{T}}g_1=-\|g_1\|^2$.当 $k>1$ 时,在(2)式两边同时与 g_k 做内积得

$$g_k^{\mathrm{T}}d_k=-\|g_k\|^2+\frac{g_k^{\mathrm{T}}d_{k-1}}{\|g_{k-1}\|^2+\max\{\alpha_{k-1},c\}\|d_{k-1}\|^2}\|g_k\|^2\leq$$

$$\left(\frac{\sigma|g_{k-1}{}^{\mathrm{T}}d_{k-1}|}{\|g_{k-1}\|^2+c\|d_{k-1}\|^2}-1\right)\|g_k\|^2\leq\left(\frac{\sigma}{2\sqrt{c}}-1\right)\|g_k\|^2, \tag{11}$$

其中,上面的第 2 个不等式是利用了 $u^{\mathrm{T}}v\leq\frac{1}{2}(\|u\|^2+\|v\|^2)$ 的结论.

2 算法的收敛性分析

本节证明算法 1 的全局收敛性.这里总是假设存在 $\varepsilon>0$,使得 $\forall k\in\mathbf{N}$,有 $\|g_k\|>\varepsilon$ 成立;否则,$f(x)$ 的一个稳定点已经找到.同时,本文作如下必要的假设 H:

(H1) $f(x)$ 在水平集 $L_0=\{x\in\mathbf{R}^n\mid f(x)\leq f(x_1)\}$ 有界;(H2) $g(x)$ 在水平集 L_0 内连续可微且满足 Lipschitz 条件,即存在 $L>0$,使 $\|g(x)-g(y)\|\leq L\|x-y\|$,$\forall x,y\in L_0$.

由假设(H1)和(H2)知,$\exists\gamma>0$,$B>0$,满足 $\forall x\in L_0$,$\|g(x)\|\leq\gamma$,$\|x\|\leq B$.

定理 1 若假设 H 成立,d_k 由算法 1 定义,则存在常数 $\bar{c}=1+2L+\frac{1}{2\sqrt{c}}>0$ 满足

$$\|d_k\|\leq\bar{c}\|g_k\|\quad\forall k\in\mathbf{N}. \tag{12}$$

证明 首先对 β_k 进行估计.根据(8)式中参数 β_k^{MWYL} 的定义、Cauchy 不等式和(10)式得到

$$|\beta_k^{\mathrm{MWYL}}|=\left|\frac{g_k^{\mathrm{T}}\left(g_k-\frac{\|g_k\|}{\|g_{k-1}\|}g_{k-1}\right)}{\|g_{k-1}\|^2+\max\{\alpha_{k-1},c\}\|d_{k-1}\|^2}\right|\leq\frac{\|g_k-g_{k-1}\|+\left\|g_{k-1}-\frac{\|g_k\|}{\|g_{k-1}\|}g_{k-1}\right\|}{\|g_{k-1}\|^2+\max\{\alpha_{k-1},c\}\|d_{k-1}\|^2}\|g_k\|\leq$$

$$2\frac{\|g_k-g_{k-1}\|}{\|g_{k-1}\|^2+\max\{\alpha_{k-1},c\}\|d_{k-1}\|^2}\|g_k\|\leq2\frac{L\alpha_{k-1}\|d_{k-1}\|}{\alpha_{k-1}\|d_{k-1}\|^2}\|g_k\|=2\frac{L\|g_k\|}{\|d_{k-1}\|}. \tag{13}$$

其次,可以估计 $|\theta_k^{\mathrm{MWYL}}|$ 如下:

$$|\theta_k^{\mathrm{MWYL}}|=\frac{|g_k^{\mathrm{T}}d_{k-1}|}{\|g_{k-1}\|^2+\max\{\alpha_{k-1},c\}\|d_{k-1}\|^2}\cdot\frac{\|g_k\|}{\|g_{k-1}\|}\le\frac{\sigma|g_{k-1}^{\mathrm{T}}d_{k-1}|}{\|g_{k-1}\|^2+c\ \|d_{k-1}\|^2}\ \frac{\|g_k\|}{\|g_{k-1}\|}\le$$

$$\sigma\ \frac{\|g_{k-1}\|\cdot\|d_{k-1}|}{\|g_{k-1}\|^2+c\|d_{k-1}\|^2}\ \frac{\|g_k\|}{\|g_{k-1}\|}\le\frac{\sigma}{2\sqrt{c}}\ \frac{\|g_k\|}{\|g_{k-1}\|}\tag{14}$$

综合上述结果,根据(7)式中 d_k 的定义得

$$\|d_k^{\mathrm{MWYL}}\|\le\|g_k\|+|\beta_k^{\mathrm{MWYL}}|\cdot\|d_{k-1}\|+|\theta_k^{\mathrm{MWYL}}|\cdot\|g_{k-1}\|\le\left(1+2L+\frac{\sigma}{2\sqrt{c}}\right)\|g_k\|.\tag{15}$$

下面的定理被称为 Zoutendijk 条件,通常被用于共轭梯度法的收敛性证明中.

定理 2　文献[15]考虑一般的共轭梯度法,如果假设(H)成立,其中的步长 α_k 满足 Wolfe 搜索,搜索方向 d_k 是下降的,则有

$$\sum_{k\ge1}\frac{(g_k^{\mathrm{T}}d_k)^2}{\|d_k\|^2}<\infty.\tag{16}$$

定理 3　考虑假设(H)成立,迭代点列 $\{x_k\}$ 由算法 1 生成,如果存在 $\|g_k\|\ge\epsilon$,那么

$$\liminf_{k\to\infty}\|g_k\|=0.\tag{17}$$

证明　采用反证法进行证明. 如果(17)式不成立,则存在 $\epsilon>0$ 使得

$$\|g_k\|\ge\epsilon,\forall\ k\in\mathbf{N};\tag{18}$$

由 Zoutendijk 和充分下降条件(5)式,则有

$$\sum_{k\ge1}(\bar{c})^{-2}\|g_k\|^2\le\sum_{k\ge1}\frac{\|g_k\|^4}{\|d_k\|^2}<\left(\frac{\sigma}{2\sqrt{c}}-1\right)\sum_{k\ge1}\frac{(g_k^{\mathrm{T}}d_k)^2}{\|d_k\|^2}<+\infty\ ,\tag{19}$$

其中 第 1 个不等式来自(12)式,第 2 个不等式源自 Zoutendijk 条件.

3　数值实验

本节比较了 PRP 方法、MYL 方法、HZ 方法和本文提出的 MWYL 方法的数值性能.所有的算法用 Fortran 语言编写程序,所有试验的测试环境相同,采取的搜索方式相同,结果如表 1 所示.所有测试问题的维数都是 1 000,一些参数的取值如下:强 Wolfe 搜索的参数为 $\rho=10^{-4}$,$\sigma=0.1$.收敛准则 $\|g_k\|\le10^{-6}$. MWYL 方法中,参数选取为 $c=0.08$,“(调用)函数次数=目标函数计算次数+5×梯度的计算次数”.初步的实验结果显示,本文算法与以上 3 类数值性能优越的算法相媲美,可以有效求解测试问题,因而是有效的.

表 1　几类方法的数值结果与比较

函数名称	方法	迭代次数	函数次数	最优值 $f(x^*)$	最优解的梯度的模 $\|g(x^*)\|$
Extended Penalty	HZ	21	388	0.883 19E+03	0.250 06E-06
Extended Penalty	MWYL	14	90	0.883 19E+03	0.570 76E-06
Extended Penalty	WYL	18	100	0.883 19E+03	0.638 86E-06
Extended Penalty	PRP	17	85	0.883 19E+03	0.597 76E-07
Raydan 2	HZ	2	7	0.100 00E+04	0.544 85E-06
Raydan 2	MWYL	4	10	0.100 00E+04	0.620 40E-06
Raydan 2	WYL	5	12	0.100 00E+04	0.219 51E-06
Raydan 2	PRP	5	106	0.100 00E+04	0.111 15E-12
Generalized Tridiagonal 1	HZ	23	73	0.997 21E+03	0.985 84E-06
Generalized Tridiagonal 1	MWYL	23	73	0.997 21E+03	0.937 28E-06
Generalized Tridiagonal 1	WYL	24	76	0.997 21E+03	0.749 80E-06
Generalized Tridiagonal 1	PRP	24	78	0.997 21E+03	0.919 23E-06
Generalized Tridiagonal 2	HZ	55	239	0.111 49E+01	0.884 03E-06

续表

函数名称	方法	迭代次数	函数次数	最优值 $f(x^*)$	最优解的梯度的模 $\|g(x^*)\|$
Generalized Tridiagonal 2	MWYL	48	195	0.111 49E+01	0.854 10E-06
Generalized Tridiagonal 2	WYL	68	263	0.111 49E+01	0.880 84E-06
Generalized Tridiagonal 2	PRP	51	221	0.111 49E+01	0.665 14E-06
Diagonal 5	HZ	2	7	0.693 15E+03	0.454 61E-08
Diagonal 5	MWYL	4	10	0.693 15E+03	0.446 67E-07
Diagonal 5	WYL	5	12	0.693 15E+03	0.282 46E-07
Diagonal 5	PRP	3	9	0.693 15E+03	0.790 53E-06
Extended Himmelblau	HZ	8	39	0.223 86E-15	0.188 64E-06
Extended Himmelblau	MWYL	12	52	0.362 36E-14	0.483 92E-06
Extended Himmelblau	WYL	14	58	0.671 18E-15	0.226 56E-06
Extended Himmelblau	PRP	9	44	0.177 86E-14	0.436 93E-06
Extended Block-Diagonal	HZ	26	117	0.785 35E-13	0.391 10E-06
Extended Block-Diagonal	MWYL	22	87	0.185 68E-12	0.597 74E-06
Extended Block-Diagonal	WYL	24	81	0.103 94E-12	0.443 40E-06
Extended Block-Diagonal	PRP	28	212	0.121 12E-13	0.515 29E-06
Extended Quadratic Penal	HZ	10	81	0.399 00E+04	0.435 77E-06
Extended Quadratic Penal	MWYL	15	186	0.399 00E+04	0.610 82E-06
Extended Quadratic Penal	WYL	17	277	0.399 00E+04	0.482 49E-06
Extended Quadratic Penal	PRP	19	109	0.399 00E+04	0.601 91E-06
Extended Tridiagonal 2	HZ	34	104	0.389 34E+03	0.503 18E-06
Extended Tridiagonal 2	MWYL	36	112	0.389 34E+03	0.999 02E-06
Extended Tridiagonal 2	WYL	44	136	0.389 34E+03	0.531 18E-06
Extended Tridiagonal 2	PRP	33	100	0.389 34E+03	0.747 89E-06
DIXMAANA (CUTE)	HZ	7	25	0.100 00E+01	0.104 14E-07
DIXMAANA (CUTE)	MWYL	8	27	0.100 00E+01	0.100 71E-06
DIXMAANA (CUTE)	WYL	10	33	0.100 00E+01	0.300 89E-06
DIXMAANA (CUTE)	PRP	9	30	0.100 00E+01	0.109 46E-06
DIXMAANB (CUTE)	HZ	6	23	0.100 00E+01	0.635 57E-06
DIXMAANB (CUTE)	MWYL	6	22	0.100 00E+01	0.664 97E-07
DIXMAANB (CUTE)	WYL	6	22	0.100 00E+01	0.509 26E-07
DIXMAANB (CUTE)	PRP	7	25	0.100 00E+01	0.603 38E-07
DIXMAANC (CUTE)	HZ	7	26	0.100 00E+01	0.169 05E-06
DIXMAANC (CUTE)	MWYL	6	23	0.100 00E+01	0.707 43E-06
DIXMAANC (CUTE)	WYL	6	23	0.100 00E+01	0.242 26E-06
DIXMAANC (CUTE)	PRP	7	26	0.100 00E+01	0.269 60E-06
Partial Perturbed Quadra	HZ	20	61	0.253 34E-14	0.814 25E-06
Partial Perturbed Quadra	MWYL	32	97	0.317 00E-14	0.995 59E-06
Partial Perturbed Quadra	WYL	58	175	0.397 92E-14	0.978 94E-06
Partial Perturbed Quadra	PRP	20	61	0.253 34E-14	0.814 25E-06
Broyden Tridiagonal	HZ	66	287	0.712 53E+00	0.824 39E-06
Broyden Tridiagonal	MWYL	74	272	0.712 53E+00	0.980 24E-06
Broyden Tridiagonal	WYL	121	403	0.712 53E+00	0.977 51E-06
Broyden Tridiagonal	PRP	60	226	0.712 53E+00	0.984 55E-06
EDENSCH (CUTE)	HZ	25	174	0.600 33E+04	0.462 14E-06
EDENSCH (CUTE)	MWYL	39	1 341	0.600 33E+04	0.798 51E-06
EDENSCH (CUTE)	WYL	38	969	0.600 33E+04	0.524 61E-06
EDENSCH (CUTE)	PRP	38	1 148	0.600 33E+04	0.643 96E-06
VARDIM (CUTE)	HZ	8	22	−0.107 97E+03	0.159 84E-07
VARDIM (CUTE)	MWYL	9	24	−0.107 97E+03	0.980 86E-06

续表

函数名称	方法	迭代次数	函数次数	最优值 $f(x^*)$	最优解的梯度的模 $\|g(x^*)\|$
VARDIM (CUTE)	WYL	11	28	−0.107 97E+03	0.408 57E-06
VARDIM (CUTE)	PRP	13	35	−0.107 97E+03	0.308 77E-06
Diagonal 6	HZ	2	7	0.222 04E-12	0.544 85E-06
Diagonal 6	MWYL	4	10	0.222 04E-12	0.620 40E-06
Diagonal 6	WYL	5	12	0.000 00E+00	0.219 51E-06
Diagonal 6	PRP	3	9	0.222 04E-12	0.690 56E-06
DIXMAANF (CUTE)	HZ	22	80	0.110 82E+04	0.570 13E-06
DIXMAANF (CUTE)	MWYL	40	1 360	0.110 82E+04	0.727 30E-06
DIXMAANF (CUTE)	WYL	31	487	0.110 82E+04	0.815 48E-06
DIXMAANF (CUTE)	PRP	31	390	0.110 82E+04	0.656 42E-06
DIXMAANG (CUTE)	HZ	12	48	0.978 32E-19	0.571 43E-09
DIXMAANG (CUTE)	MWYL	24	80	0.109 55E-12	0.651 99E-06
DIXMAANG (CUTE)	WYL	26	87	0.952 39E-13	0.619 16E-06
DIXMAANG (CUTE)	PRP	11	42	0.370 38E-14	0.128 35E-06
DIXMAANI (CUTE)	HZ	8	28	0.125 51E-18	0.898 47E-09
DIXMAANI (CUTE)	MWYL	12	38	0.823 09E-14	0.222 03E-06
DIXMAANI (CUTE)	WYL	8	26	0.143 18E-13	0.255 56E-06
DIXMAANI (CUTE)	PRP	9	29	0.816 85E-18	0.183 56E-08
DIXMAANJ (CUTE)	HZ	10	67	0.545 42E-23	0.609 67E-10
DIXMAANJ (CUTE)	MWYL	16	90	0.347 80E-15	0.316 41E-06
DIXMAANJ (CUTE)	WYL	16	82	0.206 81E-15	0.286 57E-06
DIXMAANJ (CUTE)	PRP	9	56	0.774 29E-19	0.427 02E-08
FLETCHCR (CUTE)	HZ	14	55	0.158 31E-13	0.284 83E-06
FLETCHCR (CUTE)	MWYL	16	62	0.152 57E-12	0.909 78E-06
FLETCHCR (CUTE)	WYL	17	66	0.646 85E-14	0.171 03E-06
FLETCHCR (CUTE)	PRP	17	69	0.127 26E-13	0.242 23E-06
Extended DENSCHNB (CUTE)	HZ	3	10	−0.480 45E+03	0.105 38E-06
Extended DENSCHNB (CUTE)	MWYL	3	10	−0.480 45E+03	0.108 05E-07
Extended DENSCHNB (CUTE)	WYL	3	10	−0.480 45E+03	0.101 08E-07
Extended DENSCHNB (CUTE)	PRP	5	16	−0.480 45E+03	0.415 22E-08
Extended DENSCHNF (CUTE)	HZ	4	24	−0.250 00E+00	0.315 98E-12
Extended DENSCHNF (CUTE)	MWYL	4	13	−0.250 00E+00	0.365 13E-12
Extended DENSCHNF (CUTE)	WYL	4	13	−0.250 00E+00	0.143 94E-12
Extended DENSCHNF (CUTE)	PRP	7	22	−0.250 00E+00	0.947 93E-13
SINQUAD (CUTE)	HZ	8	38	0.386 60E+03	0.233 72E-06
SINQUAD (CUTE)	MWYL	9	42	0.386 60E+03	0.371 05E-06
SINQUAD (CUTE)	WYL	12	48	0.386 60E+03	0.976 22E-06
SINQUAD (CUTE)	PRP	10	41	0.386 60E+03	0.835 68E-06
BIGGSB1 (CUTE)	HZ	31	98	0.434 34E-13	0.956 84E-06
BIGGSB1 (CUTE)	MWYL	32	101	0.531 16E-13	0.703 65E-06
BIGGSB1 (CUTE)	WYL	43	134	0.665 39E-13	0.840 50E-06
BIGGSB1 (CUTE)	PRP	30	95	0.195 63E-13	0.360 46E-06
Scaled Quadratic SQ1	HZ	3	10	0.200 00E+00	0.856 60E-07
Scaled Quadratic SQ1	MWYL	3	10	0.200 00E+00	0.610 17E-08
Scaled Quadratic SQ1	WYL	2	7	0.200 00E+00	0.298 15E-07
Scaled Quadratic SQ1	PRP	5	16	0.200 00E+00	0.155 34E-07
Scaled Quadratic SQ2	HZ	24	135	0.510 27E-16	0.403 44E-06
Scaled Quadratic SQ2	MWYL	22	129	0.410 74E-16	0.357 17E-06
Scaled Quadratic SQ2	WYL	24	154	0.291 93E-15	0.948 40E-06
Scaled Quadratic SQ2	PRP	26	168	0.165 57E-15	0.754 45E-06

4 结束语

本文研究了一类修正的 Wei-Yao-Liu 型三项共轭梯度法,该方法扩大了其中参数的选择范围,在强 Wolfe 搜索下满足充分下降条件和全局收敛性. 经过测试函数的试验对比,发现修正的算法在一定程度上可以和 HZ 方法接近,说明该算法的有效性. 如何从条件数的角度去分析迭代矩阵的特征值,去减小误差和选择最优参数,是今后要进一步研究的问题.

参 考 文 献

[1] Fletcher R,Reeves C.Function minimization buconjugate gradients[J].Compute J,1964,7:149-154.
[2] Gilbert J C,Nocedal J.Global convergence of conjugate gradient methods for optimization[J].SIAM J Optim,1992,2:21-42.
[3] Dai Y,Yuan Y.A Nonlinear conjugate gradient with a strong global convergence property[J].SIAM J Optim,2000,10:177-182.
[4] Dai Y,Liao L.New conjugacy conditions and related nonlinear conjugate gradient methods[J].Appl Math Opt,2001,43:87-101.
[5] Hager W W,Zhang H.A survey of nonlinear conjugate gradient methods[J].Pacific J Optim,2006,2:35-58.
[6] Hager W W,Zhang H.A new conjugate gradient method with guaranteed descent and an efficient line search[J].SIAM J Optim,2005,16(1):170-192.
[7] Zhang L,Zhou W J,Li D H.A descent modified Polak-Ribière-Polyak conjugate gradient method and its global convergence[J].IMA J Numer Anal,2006,26(4):629-640.
[8] Dai Y H,Kou C X.A nonlinear conjugate gradient algorithm with an optimal property and an improved Wolfe line search[J].SIAM J Optim,2013,23(1):296-320.
[9] 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.
[10] Dai Z F,Wen F H.Global convergence of a modified Hestenes-Stiefel nonlinear conjugate gradient method with Armijo line search[J].Numer Algor,2012,59(1):79-93.
[11] 董晓亮,高岳林,何郁波.一类混合的 FR-PC 共轭梯度法及其全局收敛性[J].河南师范大学学报(自然科学版),2010,38(2):42-44.
[12] 孙中波,段复建.一类无约束优化的非单调共轭梯度法[J].河南师范大学学报(自然科学版),2010,38(1):12-15.
[13] 董晓亮,杨喜美,黄元元.一类 Armijo 搜索下新的共轭梯度法及其全局收敛性[J].河南师范大学学报(自然科学版),2015,43(6):25-29.
[14] 何郁波,董晓亮.混合互补问题的光滑类 Broyden 拟牛顿算法[J].河南师范大学学报(自然科学版),2010,38(6):27-30.
[15] Wolfe P.Convergence conditions for ascent methods[J].SIAM Rev,1969,11(2):226-235.

Global convergence of a new Wei-Yao-Liu type conjugate gradient method

Dong Xiaoliang[a], Li Weijun[b]

(a.School of Mathematics and Information;b.Network Information Technology Center,Beifang Minzu University,Yinchuan 750021,China)

Abstract:Due to the features of strong global convergence properties and low memory requirement, conjugate gradient methods constitute an active choice for efficiently solving the large-scale unconstrained optimization problems. In this paper, an improved Wei-Yao-Liu type three-term conjugate gradient method is proposed, in which the scope of the involved parameter is enlarged. With the proper conditions, the sufficient descent condition and global convergence of the presented method are satisfied with the strong Wolfe conditions. preliminary computational results show that the improved method is effcient and can be used to deal with some test problems.

Keywords:conjugate gradient method; global convergence; sufficient descent condition

[责任编校 陈留院]