2016 年 1 月 重庆师范大学学报(自然科学版) Jan. 2016
第 33 卷 第 1 期 Journal of Chongqing Normal University (Natural Science) Vol. 33 No. 1

Operations Research and Cybernetics DOI:10.11721/cqnuj20160102

A New Spectral Conjugate Gradient Method and Its Global Convergence*

HUANG Liuyu, GAO Shuping, ZHANG Baoyu
(School of Mathematics and Statistics, Xidian University, Xi'an 710126, China)

Abstract: Conjugate gradient method has played a special role for solving large scale optimization problems due to the simplicity of its iteration, convergence property and its low memory requirement. In this paper, a new spectral conjugate gradient method is presented through improving conjugate parameter β_k^{RMIL} of paper. The search direction generated by this method at each iteration is always a descent direction. Under some appropriate conditions, the proposed method is globally convergent with both Armijo line search and Wolfe line search. Numerical experiments show that this method is efficient and feasible.

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

CLC number: O224 **Document code**: A **Artide number**: 1672-6693(2016)01-0017-07

Consider the following unconstrained optimization problem

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

where $f(x):\mathbf{R}^n\to\mathbf{R}^1$ is a continuously differentiable function and its gradient is denoted by $g(x)=\nabla f(x)$.

Conjugate gradient(CG) method is a most famous iterative method for solving large scale problem(1), characterized by the simplicity of its iteration and its low memory requirements. The iterative formula of the CG method is given by

$$x_{k+1}=x_k+\alpha_k d_k, k=0,1,2,\cdots \tag{2}$$

$$d_k=\begin{cases}-g_k, k=0,\\ -g_k+\beta_k d_{k-1}, \ k\geqslant 1,\end{cases} \tag{3}$$

in which α_k is the step-size obtained by some line search, d_k is the search direction, β_k is a scalar called the conjugate parameter, and $g_k=\nabla f(x_k)$. Compared with Newton method[1-2], the CG method does not need to compute and store matrices. However, there exist many theoretical and computational challenges to apply the CG method into solving problem (1). These challenges concern the selection of initial direction, the computation of α_k and β_k, and so forth.

Correspond to different β_k, the most well-known formulas are FR, PRP, HS and DY[3], which β_k are specified by

$$\beta_k^{\text{FR}}=\frac{\|g_k\|^2}{\|g_{k-1}\|^2}, \beta_k^{\text{PRP}}=\frac{g_k^{\text{T}}(g_k-g_{k-1})}{\|g_{k-1}\|^2}, \beta_k^{\text{HS}}=\frac{g_k^{\text{T}}(g_k-g_{k-1})}{(g_k-g_{k-1})^{\text{T}}d_{k-1}}, \beta_k^{\text{DY}}=\frac{\|g_k\|^2}{(g_k-g_{k-1})^{\text{T}}d_{k-1}}$$

respectively.

---

* **Received**: 01-20-2015 **Accepted**: 06-04-2015 **Internet publishing time**:
The first author biography: HUANG Liuyu, female, Research fields: Optimization, E-mail: hly2990@126.com; **Corresponding author**: Gao shuping, professer, E-mail: gaosp@mail.xidian.edu.cn
收稿日期:2015-01-20 **修回日期**:2015-06-04 **网络出版时间**:2015-12-02 13:29
作者简介:黄柳玉,女,研究方向为最优化,E-mail:hly2990@126.com;**通信作者**:高淑萍,教授,E-mail:gaosp@mail.xidian.edu.cn
网络出版地址:http://www.cnki.net/kcms/detail/50.1165.n.20151202.1329.060.html

So far, many researchers have discussed the CG method. Wei et al[4] presented a VPRP method in which β_k is given by

$$\beta_k^{\mathrm{VPRP}}=\frac{\|g_k\|^2-\left(\frac{\|g_k\|}{\|g_{k-1}\|}\right)g_k^{\mathrm{T}}g_{k-1}}{\|g_{k-1}\|^2}. \tag{4}$$

The VPRP method is globally convergent under some line searches such as the exact line search, the Wolfe line search and the Grippo-Lucidi line search. Huang et al[5] modified β_k^{VPRP}[4] as follows

$$\beta_k^{\mathrm{NPRP}}=\frac{g_k^{\mathrm{T}}\left(g_k-\left(\frac{\|g_k\|}{\|g_{k-1}\|}\right)g_{k-1}\right)}{\mu|g_k^{\mathrm{T}}d_{k-1}|+\|g_{k-1}\|^2},\mu\geqslant 0. \tag{5}$$

If $\mu=0$, then $\beta_k^{\mathrm{VPRP}}=\beta_k^{\mathrm{NPRP}}$. Therefore, the NPRP method not only contains all the features of the VPRP, but possesses sufficient descent property without any line search when the parameter $\mu>2$. Rivaie et al[6] proposed a new RMIL method, where the conjugate parameter β_k is defined by

$$\beta_k^{\mathrm{RMIL}}=\frac{g_k^{\mathrm{T}}(g_k-g_{k-1})}{\|d_{k-1}\|^2}. \tag{6}$$

Compared to the other methods under the exact line search, the RMIL method have global convergence, linear convergence rate and the best performance.

Recently, Birgin and Martinez[7] proposed a spectral conjugate gradient(SCG) method by combining the CG method and spectral gradient method[8], in here the search direction d_k is given by

$$d_k=\begin{cases}-g_k,k=0,\\-\theta_k g_k+\beta_k d_{k-1},k\geqslant 1,\end{cases} \tag{7}$$

where $\theta_k=\frac{s_{k-1}^{\mathrm{T}}s_{k-1}}{s_{k-1}^{\mathrm{T}}y_{k-1}}$ is the spectral quotient, $\beta_k=\frac{(\theta_k y_{k-1}-s_{k-1})^{\mathrm{T}}g_k}{d_{k-1}^{\mathrm{T}}y_{k-1}}$, $s_{k-1}=x_k-x_{k-1}$ and $y_{k-1}=g_k-g_{k-1}$. The numerical results[7] indicate that SCG method is more efficiency than CG method. However, the search direction d_k in the SCG method is not a descent direction. Therefore, some modified SCG methods have been developed.

Wan et al[9] proposed a new PRP spectral method, in which the descent property of the search direction can be guaranteed. Gao et al[10] presented a new hybrid SCG method(HSCG), where the spectral quotient θ_k and the conjugate parameter β_k are given by

$$\theta_k^{\mathrm{HSCG}}=1+\beta_k^{\mathrm{HSCG}}\frac{g_k^{\mathrm{T}}d_{k-1}}{\|g_k\|^2}, \tag{8}$$

$$\beta_k^{\mathrm{HSCG}}=\max\{\beta_k^{\mathrm{IPRP}},\min\{\beta_k^{\mathrm{FR}},\beta_k^{\mathrm{PRP}}\}\}, \tag{9}$$

in which $\beta_k^{\mathrm{IPRP}}=\frac{\|g_k\|^2-\left(\frac{\|g_k\|}{\|g_{k-1}\|}\right)|g_k^{\mathrm{T}}g_{k-1}|}{\|g_{k-1}\|^2}$. The HSCG method contains a good convergence of the FR method and a perfect computational efficiency of the PRP method.

In this paper, we improve the conjugate parameter β_k^{RMIL} and give a new search direction d_k, which are respectively defined by

$$\beta_k^{\mathrm{NRMIL}}=\frac{\|g_k\|^2-\left(\frac{\|g_k\|}{\|g_{k-1}\|}\right)|g_k^{\mathrm{T}}g_{k-1}|}{\mu|g_k^{\mathrm{T}}d_{k-1}|+\|d_{k-1}\|^2},\mu>1, \tag{10}$$

$$d_k=\begin{cases}-g_k,k=0,\\-\theta_k^{\mathrm{NRMIL}}g_k+\beta_k^{\mathrm{NRMIL}}d_{k-1},\ k\geqslant 1,\end{cases} \tag{11}$$

where $\theta_k^{\mathrm{NRMIL}}=1+\beta_k^{\mathrm{NRMIL}}\frac{g_k^{\mathrm{T}}d_{k-1}}{g_k{}^2}$. Then, a new SCG method called the NRMIL method is proposed and its global convergence is proved.

1 NRMIL method and its descent property

The steps of NRMIL method are as follows.

Step 0: Given constants $0<\sigma_1<\sigma_2<1$, $\rho\in(0,1)$, $0<\delta_1<\delta_2<1$, $\mu>1$, $\varepsilon>0$. Choose an initial point, $x_0\in\mathbf{R}^n$. Let $k=0$.

Step 1: If $\|g_k\|<\varepsilon$, then the algorithm stops. Otherwise, compute d_k by (11) and β_k^{NRMIL} by (10).

Step 2: Determine the step-size α_k by Armijo line search or Wolfe line search.

1) Armijo line search $\alpha_k=\max\{\rho^j, j=0,1,2,\cdots\}$ satisfying

$$f(x_k+\alpha_k d_k)\leqslant f(x_k)+\delta_1\alpha_k g_k^{\mathrm{T}}d_k-\delta_2\alpha_k^2\|d_k\|^2 \tag{12}$$

2) Wolfe line search

$$f(x_k+\alpha_k d_k)-f(x_k)\leqslant\sigma_1\alpha_k g_k^{\mathrm{T}}d_k, \tag{13}$$

$$g(x_k+\alpha_k d_k)^{\mathrm{T}}d_k\geqslant\sigma_2 g_k^{\mathrm{T}}d_k. \tag{14}$$

Step 3: Let $x_{k+1}=x_k+\alpha_k d_k$ and $k=k+1$. Return to step 1.

The following theorem shows that NRMIL method has descent property.

Theorem 1　Suppose the search direction d_k is obtained by (10) and (11), then

$$g_k^{\mathrm{T}}d_k=-\|g_k\|^2<0 \tag{15}$$

holds for all $k\geqslant0$.

Proof　Firstly, for $k=0$, since $d_0=-g_0$, it is clear that (15) is true.

Secondly, for $k>0$, from (11), it follows that

$$g_k^{\mathrm{T}}d_k=-\theta_k\|g_k\|^2+\beta_k^{\mathrm{MRMIL}}g_k^{\mathrm{T}}d_{k-1}=-\|g_k\|^2-\beta_k^{\mathrm{MRMIL}}g_k^{\mathrm{T}}d_{k-1}+\beta_k^{\mathrm{MRMIL}}g_k^{\mathrm{T}}d_{k-1}=-\|g_k\|^2.$$

Hence, $g_k^{\mathrm{T}}d_k=-\|g_k\|^2<0$ holds for all $k\geqslant0$. The proof is completed.

2 The global convergence of NRMIL method

The following assumptions are very important for the global convergence property.

Assumption 1[11]　$f(x)$ is bounded from below on the level set $\Omega=\{x\in\mathbf{R}^n: f(x)\leqslant f(x_0)\}$, where x_0 is the initial point.

Assumption 2[11]　In some neighborhood N of Ω, $g(x)$ is Lipschitz continuous, that as, there exists a constant $L\geqslant0$ such that

$$\|g(x)-g(y)\|\leqslant L\|x-y\|,\ \forall x,y\in N. \tag{16}$$

The following lemmas are needed to proof the convergence of NRMIL.

Lemma 1　For any $k\geqslant0$, the inequality $\beta_k^{\mathrm{NRMIL}}\leqslant\dfrac{\|g_k\|^2}{\|d_{k-1}\|^2}$ holds.

Proof　From (10), it is easy to prove Lemma 1.

Lemma 2　Suppose Assumption 1 and 2 hold. There exists a constant $T>0$ such that the step-size α_k computed by Armijo line search (12) satisfies

$$\alpha_k\geqslant T\frac{\|g_k\|^2}{\|d_k\|^2}. \tag{17}$$

Proof　Firstly, for $\alpha_k=1$, from (15), the following inequality

$$\|g_k\|\leqslant\|d_k\|. \tag{18}$$

holds. Then, (17) holds for $T=1$.

Secondly, for $\alpha_k\neq1$, there exist a $\rho^{-1}\alpha_k$, such that the line search rule (12) does not hold. Then, it follows that

$$f(x_k+\rho^{-1}\alpha_k d_k)-f(x_k)>\delta_1\rho^{-1}\alpha_k g_k^{\mathrm{T}}d_k-\delta_2\rho^{-2}\alpha_k^2\|d_k\|^2. \tag{19}$$

There exist a $t_k\in(0,1)$, such that $x_k+t_k\rho^{-1}\alpha_k d_k\in N$ and

$$f(x_k+\rho^{-1}\alpha_k d_k)-f(x_k)=\rho^{-1}\alpha_k g(x_k+t_k\rho^{-1}\alpha_k d_k)^{\mathrm{T}}d_k=$$

$$\rho^{-1}\alpha_k g_k^{\mathrm{T}} d_k+\rho^{-1}\alpha_k[g(x_k+t_k\rho^{-1}\alpha_k d_k)-g(x_k)]^{\mathrm{T}} d_k\leqslant\rho^{-1}\alpha_k g_k^{\mathrm{T}} d_k+L\rho^{-2}\alpha_k^2\|d_k\|^2. \tag{20}$$

From equality (15), inequality (19) and (20), we obtain $\alpha_k\geqslant\frac{-(1-\delta_1)g_k^{\mathrm{T}}d_k}{\rho^{-1}(L+\delta_2)\|d_k\|^2}=\frac{(1-\delta_1)\rho\|g_k\|^2}{(L+\delta_2)\|d_k\|^2}=T\frac{\|g_k\|^2}{\|d_k\|^2}$, where $T=\frac{(1-\delta_1)\rho}{L+\delta_2}$. The proof of the desired result has been completed.

Lemma 3 Suppose that Assumption 1 and 2 hold. The search direction d_k is obtained by formula (10) and (11), and the step-size α_k is calculated by the line search rule (12) or (13), (14), then the Zoutendijk condition

$$\sum_{k=0}^{\infty}\frac{\|g_k\|^4}{\|d_k\|^2}<\infty, \tag{21}$$

holds.

Proof If α_k is computed by Armijo line search (12), then $f_k>f_{k+1}$ and

$$\delta_2\alpha_k^2\|d_k\|^2-\delta_1\alpha_k g_k^{\mathrm{T}}d_k\leqslant f_k-f_{k+1}. \tag{22}$$

when $f_k>f_{k+1}$ and Assumption 1, it follows that the sequence $\{f_k\}$ is monotone decreasing and has lower bound. Hence, the sequence $\{f_k\}$ converges.

From equality (15), inequality (17) and (22), it follows that

$$f_k-f_{k+1}\geqslant-\delta_1\alpha_k g_k^{\mathrm{T}}d_k+\delta_2\alpha_k^2\|d_k\|^2\geqslant-\delta_1\alpha_k g_k^{\mathrm{T}}d_k\geqslant T\delta_1\frac{\|g_k\|^4}{\|d_k\|^2}>0.$$

Then, $T\delta_1\sum_{k=0}^{+\infty}\frac{\|g_k\|^4}{\|d_k\|^2}\leqslant\sum_{k=0}^{+\infty}(f_k-f_{k+1})<+\infty$, the conclusion (21) holds.

If α_k is calculated by Wolfe line search (13) and (14), then from inequality (14), it follows that

$$(g_{k+1}-g_k)^{\mathrm{T}}d_k\geqslant(\sigma_2-1)g_k^{\mathrm{T}}d_k.$$

From (2) and Lipschitz condition (16), the following inequality

$$(g_{k+1}-g_k)^{\mathrm{T}}d_k\leqslant\|g_{k+1}-g_k\|\cdot\|d_k\|\leqslant L\|x_{k+1}-x_k\|\cdot\|d_k\|=L\alpha_k\|d_k\|^2$$

holds.

Then,

$$\alpha_k\geqslant\frac{\sigma_2-1}{L}\cdot\frac{g_k^{\mathrm{T}}d_k}{\|d_k\|^2}>0. \tag{23}$$

Hence, by Wolfe condition (13), equality(15) and inequality (23), it follows that

$$f_k-f_{k+1}\geqslant-\sigma_1\alpha_k g_k^{\mathrm{T}}d_k\geqslant\frac{\sigma_1(1-\sigma_2)}{L}\frac{\|g_k\|^4}{\|d_k\|^2}>0. \tag{24}$$

Combining the above inequality (24) with Assumption 1, it is clear that the sequence $\{f_k\}$ is monotone decreasing and has lower bound. Hence, the sequence $\{f_k\}$ converges.

Therefore, $\sum_{k=0}^{m}\frac{\|g_k\|^4}{\|d_k\|^2}\leqslant\frac{L}{\sigma_1(1-\sigma_2)}\cdot\sum_{k=0}^{m}(f_k-f_{k+1})<+\infty$. The proof is completed.

Theorem 2 Suppose Assumption 1 and 2 hold. The search direction d_k is computed by (10) and (11), and the step-size α_k satisfies Armijo line search (12) or Wolfe line search (13) and (14), then $\liminf_{k\to\infty}\|g_k\|=0$ holds.

Proof Suppose that there exist a positive constant $c>0$ such that $\|g_k\|\geqslant c$ for all k. From (3), the following equality

$$\|d_k\|^2=(\beta_k^{\mathrm{NRMIL}})^2\|d_{k-1}\|^2-2\theta_k g_k^{\mathrm{T}}d_k-\theta_k^2\|g_k\|^2, \tag{25}$$

holds.

Dividing the above equality (25) by $\|g_k\|^4$, then together with Lemma 1 and relations (15) and (18), it follows that

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

$$(\beta_k^{\mathrm{NRMIL}})^2 \frac{\|d_{k-1}\|^2}{(g_k^{\mathrm{T}} d_k)^2} + \frac{1}{\|g_k\|^2} \leqslant \frac{\|g_k\|^4}{\|d_{k-1}\|^4} \cdot \frac{\|d_{k-1}\|^2}{\|g_k\|^4} + \frac{1}{\|g_k\|^2} \leqslant \frac{1}{\|d_{k-1}\|^2} + \frac{1}{\|g_k\|^2} \leqslant \frac{1}{\|g_{k-1}\|^2} + \frac{1}{\|g_k\|^2}.$$

Hence, $\frac{\|d_k\|^2}{\|g_k\|^4} \leqslant \sum_{i=0}^{k} \frac{1}{\|g_i\|^2} \leqslant \frac{k+1}{c^2}$. Then, $\sum_{k=0}^{\infty} \frac{\|g_k\|^4}{\|d_k\|^2} \geqslant \sum_{k=0}^{\infty} \frac{c^2}{k+1} = \infty$.

This contradicts the Zoutendijk condition (21). Therefore, the Theorem 2 holds.

3 Numerical experiments

This section provides the results of some numerical experiments. A comparison is made among NRMIL, HSCG[10], RMIL[6] and PRP[3] with Matlab 7.0. Test problems come from [12-13]. The test results for the four algorithms are obtained in Tab. 1 and Tab. 3 respectively with Armijo line search ($\rho=0.49, \delta_1=0.001, \delta_2=0.01, \mu=1.5$) and Wolfe line search ($\sigma_1=0.30, \sigma_2=0.75, \mu=1.5$). Based on PRP, we calculate the average efficiency[14] of each method, which are shown in Tab. 2 and Tab. 4. The following notations is used in Tables.

Problem: the name of the problem; Dim: the dimension of the problem; NI/NF/NG: the number of iterations, function evaluations and gradient evaluations, respectively. The stopping criterion is $\|g_k\| \leqslant 10^{-6}$ and $NI > 2\,000$.

Tab. 1　Numerical results of the four methods with Armijo line search

Problem	Dim	NRMIL	HSCG	RMIL	PRP
		NI/NF/NG	NI/NF/NG	NI/NF/NG	NI/NF/NG
Bohachevsky2	2	26/171/27	46/312/47	35/535/36	37/867/38
Griewank	2	19/44/20	20/40/21	31/62/32	37/615/38
Zakharov	2	17/76/18	21/91/22	20/348/21	23/613/24
Rastrigin	2	23/215/24	32/335/33	14/176/15	22/564/23
Trecanni	2	124/502/125	128/524/129	237/3 514/238	286/4 664/287
Himmelbau	2	31/230/32	39/309/40	38/702/39	36/861/37
Generalized Griewank	3	34/68/35	32/64/33	48/96/49	53/239/54
Schwefel's Problem 1.2	3	35/156/36	41/186/42	31/218/32	51/984/52
Fletcher	4	48/555/49	51/598/52	44/916/45	74/1692/75
Extended Beale	10	85/552/86	58/353/59	102/1 702/103	207/3 238/208
SphereModel	4	4/12/5	4/12/5	10/185/11	7/211/8
	20	5/15/6	5/15/6	13/245/14	9/278/10
Extended Denschnf	20	26/260/27	46/461/47	30/403/31	40/960/41
Raydan 1	50	87/267/88	82/312/83	94/812/95	77/807/78
	100	140/570/141	97/447/98	126/994/127	187/1 806/188
Perturbed Quadratic	20	51/317/52	60/427/61	63/1 062/64	91/2 063/92
	50	89/653/90	88/722/89	81/1 109/82	122/2 113/123
	100	166/1 372/167	205/1 856/206	131/1 427/132	188/3 252/189
Variably dimensioned	20	13/201/14	13/201/14	25/749/26	25/895/26
	200	43/1 125/44	46/1 194/47	91/3 295/92	83/3 237/84
SumSquares	20	50/312/51	78/529/79	74/1 445/75	83/1 757/84
	100	124/1 029/125	123/1 109/124	142/1 751/143	183/3 028/184
	500	461/4 831/462	305/3 309/306	352/5 202/353	433/9 315/434

The average efficiency of the four methods in Tab. 1 is shown in Tab. 2.

Tab. 2　Average efficiency of the four methods with Armijo line search

NRMIL	HSCG	RMIL	PRP
0.314 3	0.347 3	0.624 0	1

From Tab. 2, average efficiency of

NRMIL method is obviously smaller than the others. So, NRMIL method performs the best of the four methods under Armijo line search.

Tab. 3 Numerical results of the four methods with Wolfe line search

Function	Dim	NRMIL	HSCG	RMIL	PRP
		NI/NF/NG	NI/NF/NG	NI/NF/NG	NI/NF/NG
Bohachevsky2	2	21/190/43	26/241/54	23/275/49	26/425/54
Griewank	2	18/67/37	21/81/43	17/129/35	21/361/43
Zakharov	2	18/117/37	21/134/43	20/107/41	28/668/57
Booth	2	24/174/49	26/192/53	28/255/57	37/743/75
Matyas	2	17/68/69	17/68/69	26/427/124	33/872/101
Hosaki	2	8/24/17	8/24/17	18/234/37	15/300/31
McCormick	2	21/118/44	23/130/51	23/159/47	39/726/79
Fletcher	4	35/476/71	35/490/71	36/532/73	39/924/79
Himmelbau	10	19/131/39	20/145/41	20/190/41	29/475/59
Perturbed Quadratic	10	35/247/71	42/320/85	35/254/71	48/698/97
Extended Rosenbrock	20	116/1 355/234	245/2 914/492	142/1 917/286	275/6 214/551
Generalized Quartic	50	23/136/47	24/147/49	23/326/47	33/1 015/67
Raydan 2	100	12/47/25	12/47/25	20/255/41	23/434/47
Variably dimensioned	10	14/214/29	14/214/29	21/603/43	27/993/55
	100	15/454/31	15/454/31	22/906/45	29/1 498/59
SphereModel	20	12/60/25	12/60/25	17/340/35	23/739/47
	200	13/65/27	13/65/27	19/388/39	25/804/51
Extended Denschnf	20	20/247/41	27/341/55	20/324/41	39/1 046/80
	200	21/258/43	29/365/59	22/347/45	43/1 115/88
Extended Beale	10	46/355/93	62/492/125	60/539/121	85/1 242/171
	200	54/416/109	68/533/137	62/555/125	89/1 228/179
Extended White & Holst	500	241/3 083/483	1 039/13 542/2 079	281/3 857/564	757/16 342/1 516

The following table displays average efficiency of the four methods in Tab. 3.

Tab. 4 Average efficiency of the four methods with Wolfe line search

NRMIL	HSCG	RMIL	PRP
0.328 8	0.403 9	0.511 7	1

Through a comparative analysis of the four methods from Tab. 4, we reach a conclusion that NRMIL method performs the best of the four methods with Wolfe line search.

According to the above, we draw a conclusion that convergence property of NRMIL method is the best of the four methods under not only Armijo line search but also Wolfe line search. Consequently, NRMIL method is efficient and feasible.

4 Conclusions

This paper proposes NRMIL method which has a descent search direction and a global convergence under both Armijo line search and Wolfe line search. Although NRMIL method works well from the numerical experiment results, it has some disadvantages, such as, we don't discuss whether the choices of the initial points about the test functions affect the number of iterations, whether the value of the parameter μ in β_k^{NRMIL} influences the convergence rate of NRMIL method, and so on. Therefor, our future work will be concentrated on studying these aspects.

References:

[1] Frontini M, Sormani E. Some variant of Newton's method with third-order convergence[J]. Applied Mathematics and Computation, 2003, 140(2): 419-426.

[2] Zban A. Some new variants of Newton's method[J]. Applied Mathematics Letters, 2004, 17(6): 677-682.

[3] Zhang L, Zhou W J. Two descent hybrid conjugate gradient methods for optimization[J]. Journal of Computational and Applied Mathematics, 2008, 216(1): 251-264.

[4] Wei Z, Yao S, Liu L. The convergence properties of some new conjugate gradient methods[J]. Appl Math Comput 2006, 183(1): 1341-1350.

[5] Huang H, Lin S H. A modified Wei-Yao-Liu conjugate gradient method for unconstrained optimization[J]. Applied Mathematics and Computation, 2014, 231(2): 179-186.

[6] 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.

[7] Birgin E G, Martínez J M. A spectral conjugate gradient method for unconstrained optimization[J]. Applied Mathematics and Optimization, 2001, 43(2): 117-128.

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

[9] Wan Z, Yang Z L, Wang Y L. New spectral PRP conjugate gradient method for unconstrained optimization[J]. Applied Mathematics Letters, 2011, 24(1): 16-22.

[10] Gao H Y, Zhu T X, Xu C L. A new Hybrid conjugate gradient method for unconstrained optimization[J]. Journal of Jilin University: Science Edition, 2012, 50(3): 457-461.

[11] Cao W, Wang K R, Wang Y L. Global convergence of a modified spectral CD conjugate gradient method[J]. Journal of Mathematical Research & Exposition, 2011, 31(2): 261-268.

[12] Andrei N. An unconstrained optimization test functions collection[J]. Adv Model Optim, 2008, 10(1): 147-161.

[13] Moré J J, Garbow B S, Hillstrom K E. Testing unconstrained optimization software[J]. ACM Transactions on Mathematical Software (TOMS), 1981, 7(1): 17-41.

[14] Liu J K, Du X L, Wang K R. A mixed spectral CD-DY conjugate gradient method[EB/OL]. (2012-01-23)[2015-01-15]. http://dx.doi.org/10.1155/2012/569795.

运筹学与控制论

一种新的谱共轭梯度法及其全局收敛性

黄柳玉，高淑萍，张宝玉

(西安电子科技大学 数学与统计学院，西安 710126)

摘要：共轭梯度法因具有迭代简单、收敛性和低内存等优点而在求解大型优化问题中发挥着重要作用。本文对已有文献中的共轭参数 β_k^{RMIL} 进行改进，得到了一种新的谱共轭梯度法。该方法每步迭代产生的搜索方向具有下降性。在适当的条件下，该方法在 Armijo 线搜索和 Wolfe 线搜索下均具有全局收敛性。数值试验表明，该方法可行有效。

关键词：谱共轭梯度法；优化问题；Armijo 线搜索；Wolfe 线搜索；全局收敛性

(责任编辑 黄 颖)