DOI: 10.13656/j.cnki.gxkx.2005.04.011

A New Conjugate Gradient Method for Unconstrained Optimization*
无约束优化中新的共轭梯度方法

Wei Zengxin, Wu Xiaoping, Zhao Yan
韦增欣,武小平,赵 岩

(College of Mathematics and Information Science, Guangxi University, Nanning, Guangxi, 530004, China)
(广西大学数学与信息科学学院,广西南宁 530004)

Abstract A new nonlinear conjugate gradient formula for solving unconstrained optimization problem is proposed. The formula satisfies the sufficient descent condition, and the method in which the proposed formula and the weak Wolfe conditions are used is global convergence. Preliminary numerical results show that the method is promising.
Key words unconstrained optimization, conjugate gradient, Wolfe line search, sufficient descent property, global convergence
摘要:提出一个新的解决无约束最优化问题的非线性共轭梯度公式,该公式满足充分下降条件,采用这个公式和弱 Wolfe条件的方法是全局收敛的.初始的数值结果表明,该方法是有前景的.
关键词:无约束优化 共轭梯度 Wolfe线搜索 充分下降性 全局收敛性
中图法分类号: O224 文献标识码: A 文章编号: 1005-9164(2005)04-0276-06

1 Introduction

The nonlinear conjugate gradient (CG) method is quite suitable for solving large-scale unconstrained optimization problem

$$\min\{f(x) \mid x \in R^n\}, \tag{1.1}$$

where $f: R^n \to R$ is a twice continuously differentiable function whose gradient is denoted by $g: R^n \to R^n$, due to its simplicity and low memory requirement. Its iterative formula is given by

$$x_{k+1} = x_k + \alpha_k d_k, \tag{1.2}$$

where α_k is a step size which is computed by carrying out a line search, and d_k is the search direction defined by

$$d_k = \begin{cases} -g_k, & \text{if } k = 1, \\ -g_k + \beta_k d_{k-1}, & \text{if } k \geqslant 2, \end{cases} \tag{1.3}$$

where β_k is a scalar and g_k denotes $g(x_k)$. There are at least six well-knowm formulae for β_k, which are given below:

$$\beta_k^{HS} = \frac{g_k^T(g_k - g_{k-1})}{(g_k - g_{k-1})^T d_{k-1}}, \tag{1.4}$$

$$\beta_k^{FR} = \frac{g_k^T g_k}{g_{k-1}^T g_{k-1}}, \tag{1.5}$$

$$\beta_k^{PRP} = \frac{g_k^T(g_k - g_{k-1})}{g_{k-1}^T g_{k-1}}, \tag{1.6}$$

$$\beta_k^{CD} = -\frac{g_k^T g_k}{d_{k-1}^T g_{k-1}}, \tag{1.7}$$

$$\beta_k^{LS} = -\frac{g_k^T(g_k - g_{k-1})}{d_{k-1}^T g_{k-1}}, \tag{1.8}$$

$$\beta_k^{DY} = \frac{g_k^T g_k}{(g_k - g_{k-1})^T d_{k-1}}. \tag{1.9}$$

To establish the global convergence results of the above CG methods, the step size α_k is usually required to satisfy some line search conditions, such as the weak Wolfe line search

$$f(x_k + \alpha_k d_k) - f(x_k) \leqslant \delta \alpha_k g_k^T d_k, \tag{1.10}$$

$$g(x_k + \alpha_k d_k)^T d_k \geqslant \sigma g_k^T d_k, \tag{1.11}$$

where $\delta \in (0, \frac{1}{2})$ and $\sigma \in (\delta, 1)$, and the strong

收稿日期: 2004-12-30
作者简介:韦增欣(1963-),男,广西武鸣人,教授,主要从事最优化理论与方法研究
* 国家自然科学基金(10161002)和广西自然科学基金(0542043)资助。

Wolfe line search (1.10) and

$$|g(x_k + t_k d_k)^T d_k| \leqslant -\sigma g_k^T d_k, \qquad (1.12)$$

where $\delta \in (0, \frac{1}{2})$ and $\sigma \in (\delta, 1)$.

Considerable attentions have been made on the global convergence behaviors for the above methods. Zoutendijk[1] proved that the FR method with exact line search is globally convergent. Al-Baali[2] extented this result to the strong Wolfe line search conditions. In Reference [3], Dai and Yuan proposed the DY method which produces a descent search direction at every iteration and converges globally provided that the line search satisfies the weak Wolfe conditions. However, the global convergence has not been established for the PRP method with the strong Wolfe line search conditions.

Under the sufficient descent condition

$$g_k^T d_k \leqslant -c\|g_k\|^2, \qquad (1.13)$$

for some constant $c > 0$, Gilbert and Nocedal[4] was in another way to discuss the global convergence of the PRP method with the weak Wolfe line search. In Reference [4], the parameter β_k in (1.6) is not allowed to be negative, that is

$$\beta_k^+ = \max\{0, \beta_k^{PRP}\}. \qquad (1.14)$$

By using a complicated line search, Gilbert and Nocedal[4] was able to establish the global convergence result of the PRP and the HS methods by restricting the scalar β_k to be nonnegative.

In this paper, we present our new CG formula and its properties in section 2, the global convergence result of this new method in section 3, preliminary numerical results in section 4, and conclusion in the last section.

2 New conjugate gradient formula and its property

The sufficient descent condition (1.13) is an important property in the literature of analyzing the global convergence of the CG methods, so we intend to find a β_k such that d_k satisfies the condition (1.13). In the following, we propose a β_k and prove that it has such property.

A definition of a descent sequence (or a sufficient descent sequence) is given below. A sequence $\{\beta_k\}$ is called a descent sequence (or a sufficient descent sequence) for the CG methods if there exists a constant $\xi \in (0,1)$ (or $\xi \in [0,1)$) such that for all $k \geqslant 2$,

$$\beta_k g_k^T d_{k-1} \leqslant \xi \|g_k\|^2. \qquad (2.1)$$

In Reference [5], the authors proposed a variation of the FR formula (VFR):

$$\beta_k^{VFR} = \frac{\mu_1 \|g_k\|^2}{\mu_2 |g_k^T d_{k-1}| + \mu_3 \|g_{k-1}\|^2}, \qquad (2.2)$$

where $\mu_1 \in (0, \infty)$, $\mu_2 \in [\mu_1 + \varepsilon, +\infty)$, $\mu_3 \in (0, +\infty)$ and ε is an any given positive constant. In Reference [6], the authors designed the following variation of the PRP formula:

$$\beta_k^{VPRP} = \frac{\mu_1 (\|g_k\|^2 - |g_k^T g_{k-1}|)}{\mu_2 |g_k^T d_{k-1}| + \mu_3 \|g_{k-1}\|^2}, \qquad (2.3)$$

where the definitions of μ_1, μ_2, μ_3 are as same as those given in the formula (2.2). In order to ensure the nonnegative of the parameter β_k, it is defined that

$$\beta_k^{VPRP+} = \max\{0, \beta_k^{VPRP}\}. \qquad (2.4)$$

It is easy to prove that $\{\beta_k^{VFR}\}$ and $\{\beta_k^{VPRP}\}$ all are descent sequence (with $\xi = \frac{1}{\mu_2}$). Further the formulae (2.2) and (2.4) make d_k possess the sufficient descent property.

Enlightened by the above ideas, we propose a new β_k as follows

$$\beta_k^{NPRP} = \frac{\lambda \mu_1 \|g_k\|^2 + (1-\lambda)\mu_1 (\|g_k\|^2 - |g_k^T g_{k-1}|)}{\mu_2 |g_k^T d_{k-1}| + \mu_1 \|g_{k-1}\|^2}, \qquad (2.5)$$

where $\lambda \in (0, 1)$ and the definitions of μ_1, μ_2 are as same as those in the formula (2.2). In order to ensure the nonnegative of the parameter β_k, we define

$$\beta_k^{NPRP+} = \max\{0, \beta_k^{NPRP}\}, \qquad (2.6)$$

thus if a negative of β_k^{NPRP} occurs, this strategy will restart the iteration along the steepest direction.

The following two propositions show that the $\{\beta_k^{NPRP+}\}$ is a descent sequence, and so that d_k satisfies the sufficient descent condition (1.13).

Proposition 2.1 Suppose that β_k is defined by the formulae (2.5) and (2.6), then we have that

$$\beta_k^{NPRP+} \leqslant \hat{\sigma} \frac{\|g_k\|^2}{|g_k^T d_{k-1}|}, \qquad (2.7)$$

where $0 < \hat{\sigma} = \frac{1}{\mu_2} < 1$.

Proof It is clear that the inequality (2.7) holds when $\beta_k^{NPRP+} = 0$.

Now we consider the case where $\beta_k^{NPRP+} = \beta_k^{NPRP}$. So we have

$$\beta_k^{NPRP+} =$$

$$\frac{\lambda\_{1}\|g_k\|^2 + (1-\lambda)\_{1}(\|g_k\|^2 - |g_k^T g_{k-1}|)}{\_{2}|g_k^T d_{k-1}| + \_{1}\|g_{k-1}\|^2} \leqslant$$

$$\frac{\lambda\_{1}\|g_k^2\| + (1-\lambda)\_{1}\|g_k\|^2}{\_{2}|g_k^T d_{k-1}|} \leqslant$$

$$\frac{\_{1}}{\_{2}} \frac{\|g_k\|^2}{|g_k^T d_{k-1}|} = \hat{e}\frac{\|g_k\|^2}{|g_k^T d_{k-1}|}.$$

Hence U_k^{NPRP+} can make the inequality (2.7) hold. Furthermore $\{U_k^{NPRP+}\}$ is a descent sequence without any line search.

Proposition 2.2 Suppose that U_k is defined by the formulae (2.5) and (2.6), then d_k satisfies the sufficient descent condition (1.13) for all $k \geqslant 1$, where $c = (1 - \frac{\_1}{\_2})$.

Proof For any $k > 1$, suppose that $g_{k-1}^T d_{k-1} < 0$.

If $U_k^{NPRP+} = 0$, then $d_k = -g_k$. So we have

$$g_k^T d_k = -\|g_k\|^2 \leqslant -c\|g_k\|^2,$$

where $c = (1 - \frac{\_1}{\_2})$.

Otherwise, from the definition of U_k^{NPRP+} and Proposition 2.1, we can obtain

$$g_k^T d_k = g_k^T [-g_k + (\frac{\lambda\_{1}\|g_k\|^2 + (1-\lambda)\_{1}(\|g_k\|^2 - |g_k^T g_{k-1}|)}{\_{2}|g_k^T d_{k-1}| + \_{1}\|g_{k-1}\|^2}) \cdot d_{k-1}] \leqslant -\|g_k\|^2 + (\frac{\lambda\_{1}\|g_k\|^2 + (1-\lambda)\_{1}(\|g_k\|^2 - |g_k^T g_{k-1}|)}{\_{2}|g_k^T d_{k-1}| + \_{1}\|g_{k-1}\|^2}) \cdot |g_k^T d_{k-1}| \leqslant -\|g_k\|^2 + \frac{\_1}{\_2}\frac{\|g_k\|^2}{|g_k^T d_{k-1}|}|g_k^T d_{k-1}| \leqslant$$

$$-(1 - \frac{\_1}{\_2})\|g_k\|^2 = -c\|g_k\|^2.$$

For $g_k^T d_1 = -\|g_1\|^2$, we can deduce that d_k has the sufficient descent condition (1.13) for all $k \geqslant 1$.

3 Algorithm and global convergence

Assumption A The level set

$$\Omega = \{x \in R^n | f(x) \leqslant f(x_1)\}$$

is bounded.

Assumption B The gradient $g(x)$ is Lipschitz continuous, i.e, there exists a constant $L > 0$ such that for any $x, y \in \Omega$,

$$\|g(x) - g(y)\| \leqslant L\|x - y\|.$$

Now we give the algorithm:

Algorithm 3.1

Step 0 Given $x_1 \in R^n$, set $d_1 = -g_1, k = 1$. If $g_1 = 0$, then stop.

Step 1: Find a $t_k > 0$ satisfying the weak Wolfe conditions (1.10) and (1.11).

Step 2 Let $x_{k+1} = x_k + t_k d_k$ and $g_{k+1} = g(x_{k+1})$. If $\|g_{k+1}\| = 0$, then stop.

Step 3 Compute U_{k+1}^{NPRP+} by the formulae (2.5) and (2.6). Then generate d_{k+1} by the condition (1.3).

Step 4 Set $k =: k + 1$, go to step 1.

Since $\{f(x_k)\}$ is a decreasing sequence, it is clear that the sequence $\{x_k\}$ is contained in Ω, and there exists a constant f^*, such that

$$\lim_{k\to\infty} f(x_k) = f^*. \tag{3.1}$$

By using the Assumptions A and B, we can deduce that there exists $M > 0$ such that

$$\|g_k\| \leqslant M, \forall x \in \Omega. \tag{3.2}$$

The following important result was obtained by Zoutendijk[1] and Wolfe[7,8].

Lemma 3.1 Suppose that $f(x)$ is bounded below, and $g(x)$ satisfies the Lipschitz condition. Consider any iteration method of the form (1.2), where d_k satisfies $d_k^T g_k < 0$ and t_k is obtained by the weak Wolf line search. Then

$$\sum_{k=1}^{\infty} \frac{(g_k^T d_k)^2}{\|d_k\|^2} < +\infty. \tag{3.3}$$

The following lemma was obtained by Dai and Yuan[9].

Lemma 3.2 Assumption that a positive series $\{a_i\}$ satisfies the following inequality for all k

$$\sum_{i=1}^{k} a_i \geqslant lk + c, \tag{3.4}$$

where $l > 0$ and c are constants. Then we have

$$\sum_{i\geqslant 1} \frac{a_i^2}{i} = +\infty, \tag{3.5}$$

and

$$\sum_{k\geqslant 1} \frac{a_k^2}{\sum_{i=1}^{k} a_i} = +\infty. \tag{3.6}$$

Theorem 3.1 Suppose that Assumptions A and B hold, $\{x_k\}$ is a sequence generated by Algorithm 3.1, then we have

$$\lim_{k\to\infty}\inf \|g_k\| = 0 \tag{3.7}$$

holds.

Proof The condition (1.3) indicates that for all $k \geqslant 2$,

$$d_k + g_k = U_k d_{k-1}. \tag{3.8}$$

Squaring both sides of Formula (3.8), we obtain

$$\|d_k\|^2 = -\|g_k\|^2 - 2g_k^T d_k + U_k^2\|d_{k-1}\|^2. \tag{3.9}$$

Suppose that $U_k^{NPRP+} = U_k^{NPRP}$ in Formula (2.6). Then we have

$$\|d_k\|^2 = -\|g_k\|^2 - 2g_k^T d_k + (\beta_k^{NPRP+})^2\|d_{k-1}\|^2 = -\|g_k\|^2 - 2g_k^T d_k + \left(\frac{\lambda\mu_1\|g_k\|^2 + (1-\lambda)\mu_1(\|g_k\|^2 - |g_k^T g_{k-1}|)}{\mu_2|g_k^T d_{k-1}| + \mu_1\|g_{k-1}\|^2}\right)^2.$$

$$\|d_{k-1}\|^2 \leqslant -\|g_k\|^2 - 2g_k^T d_k + \|g_k\|^4\frac{\|d_{k-1}\|^2}{\|g_{k-1}\|^4},$$

that is

$$h_k \leqslant h_{k-1} - \frac{1}{\|g_k\|^2} + \frac{2V_k}{\|g_k\|^2}, \qquad (3.10)$$

where $h_k = \frac{\|d_k\|^2}{\|g_k\|^4}$ and $V_k = -\frac{g_k^T d_k}{\|g_k\|^2}$.

Note that $h_1 = \frac{1}{\|g_1\|^2}$ and $V_1 = 1$, it follows Form (3.10) that

$$h_k \leqslant -\sum_{i=1}^{k}\frac{1}{\|g_i\|^2} + 2\sum_{i=1}^{k}\frac{|V_i|}{\|g_i\|^2}. \qquad (3.11)$$

Suppose that the conclusion (3.7) is not held. Then, there exists a positive scalar γ such that for all $k \geqslant 1$,

$$\|g_k\| \geqslant \gamma. \qquad (3.12)$$

Thus, it follows Forms (3.2) and (3.12) that

$$h_k \leqslant -\frac{k}{M^2} + \frac{2}{\gamma^2}\sum_{i=1}^{k}|V_i|. \qquad (3.13)$$

Further, we have

$$h_k \leqslant \frac{2}{\gamma^2}\sum_{i=1}^{k}|V_i|. \qquad (3.14)$$

On the other hand, using $h_k \geqslant 0$, the relation (3.13) implies that

$$\sum_{i=1}^{k}|V_i| \geqslant \frac{\gamma^2 k}{2M^2}. \qquad (3.15)$$

Using Lemmas 3.2 and (3.14), it follows that

$$\sum_{k\geqslant 1}\frac{(g_k^T d_k)^2}{\|d_k\|^2} = \sum_{k\geqslant 1}\frac{V_k^2}{h_k} = +\infty, \qquad (3.16)$$

which contradicts to the Zoutendijk condition (3.3). This shows that (3.7) holds, and the proof of this theorem is complete.

From the proof of the above theorem, we can conclude that any conjugategradient method with the formula β_k^{NPRP+} and some step size techniques which ensure the Zoutendijk condition (3.3) holds is globally convergent. In particular, the formula β_k^{NPRP+} with the strong Wolfe conditions can generate a globally convergent method.

4 Numerical results

PRPSWP the PRP formula with the strong Wolfe conditions, where $\delta = 0.01$, $\sigma = 0.1$.

PRP+ SWP the PRP+ formula with the strong Wolfe conditions, where $\delta = 0.01$, $\sigma = 0.1$.

NPRP+ SWP the new CG formula β_k^{NPRP+} given in this paper with the strong Wolfe conditions, where $\delta = 0.01$, $\sigma = 0.1$, $\lambda = 0.3$, $\mu_1 = 1$, $\mu_2 = 3$.

NPRP+ WWP the new CG formula β_k^{NPRP+} given in this paper with the weak Wolfe conditions, where $\delta = 0.01$, $\sigma = 0.1$, $\lambda = 0.3$, $\mu_1 = 1$, $\mu_2 = 3$.

Further, the results of the performed comparison with the original PRP method are given. The problems that we tested are from Reference [10]. For each tested problem, the termination condition is

$$\|g(x_k)\| \leqslant 10^{-5}.$$

Table 1 shows the computation results, where the columns have the following meanings:

Problem: the name of the test problem in MATLAB;

Dim: the dimension of the problem;

NI: the number of iterations;

NF: the number of function evaluations;

NG: the number of gradient evaluations.

In order to rank the iterative numerical methods, we compute the total numbers of function and gradient evaluations by the following formula

$$N_{total} = NF + m * NG, \qquad (4.1)$$

where m is an integer. According to the results on automatic differentiation[10,11], the value of m can be set to $m = 5$.

Since the PRPSWP method is one of the commonly efficient CG methods, we compare the PRP+ SWP, the NVPRP+ SWP and the NVPRP+ WWP methods with the PRPSWP method as follows: for each tested example i, compute the total numbers of function evaluations and gradient evaluations required by the evaluated method j $(EM(j))$ and the PRPSWP method by the formula (4.1), and denote them by $N_{total,i}(EM(j))$ and $N_{total,i}$(PRPSWP); then calculate the ratio

$$r_i(EM(j)) = \frac{N_{total,i}(EM(j))}{N_{total,i}(\text{PRPSWP})}. \qquad (4.2)$$

If $EM(j_0)$ does not work for example i_0, we replace the $r_{i_0}(EM(j_0))$ by a positive constant τ which is defined as follows

$$\tau = \max\{r_i(EM(j)) : (i,j) \notin S_1\},$$

Table 1 Test results for the PRPSWP, PRP^+ SWP, $NPRP^+$ SWP, $NPRP^+$ WWP

Problem	Dim	PRPSWP NI/NF/NG	PRP^+ SWP NI/NF/NG	$NPRP^+$ SWP NI/NF/NG	$NPRP^+$ WWP NI/NF/NG
ROSE	2	29/502/65	22/394/60	28/318/67	28/355/59
FROTH	2	12/30/20	10/28/20	13/83/25	28/57/41
BADSCP	2	–	41/509/100	34/324/100	55/268/98
BADSCB	2	13/80/22	11/123/22	16/82/25	19/86/29
BEALE	2	9/126/21	9/173/20	20/96/36	19/44/31
JENSAM	2	–	–	11/172/18	13/30/18
HELIX	3	49/255/83	32/265/55	42/283/66	34/81/53
BARD	3	23/98/37	27/152/43	17/134/28	53/98/74
GAUSS	3	4/57/6	4/57/6	3/7/4	3/7/4
MEYER	3	–	–	–	3/7/4
GULF	3	1/2/2	1/2/2	1/2/2	1/2/2
BOX	3	–	–	–	–
SING	4	199/611/338	49/155/79	27/351/44	27/351/44
WOOD	4	169/1103/302	101/549/195	56/128/95	85/223/132
KOWOSB	4	55/300/94	51/249/79	16/282/25	16/282/25
BD	4	–	–	–	–
OSB_1	5	–	–	–	–
BIGGS	6	264/875/423	–	43/486/244	43/486/244
OSB_2	11	254/1061/418	250/1011/412	–	476/1279/644
WATSON	20	2795/7733/4425	2143/5780/3396	1523/3836/2394	1513/3351/2233
ROSEX	8	23/402/59	25/371/62	32/329/72	34/88/59
	50	31/533/77	24/492/60	32/324/62	38/186/68
	100	28/337/74	35/514/101	25/401/63	25/186/63
SINGX	4	199/611/338	49/155/79	56/128/95	56/128/95
PEN_1	2	5/18/12	6/20/14	5/18/12	5/18/12
PEN_2	4	12/134/28	12/136/27	10/81/25	11/75/18
	50	613/2795/1063	136/898/282	129/824/252	129/364/227
VARDIM	2	3/9/7	3/9/7	3/9/7	3/9/7
	50	10/52/36	10/52/36	10/52/36	10/52/36
TRIG	3	12/81/24	14/131/25	13/228/26	16/182/22
	50	41/279/72	41/230/72	39/372/71	41/167/52
	100	46/342/87	46/341/85	63/352/110	63/204/86
BV	3	12/25/16	12/25/16	8/17/11	8/17/11
	10	75/241/117	75/241/117	158/533/254	177/335/233
IE	3	5/12/7	5/12/7	5/12/7	5/12/7
	50	6/13/7	5/11/6	5/11/6	5/11/6
	100	6/13/8	6/13/8	6/13/8	6/13/8
	200	6/13/8	6/13/8	6/13/8	6/12/7
	500	6/13/8	6/13/8	6/13/8	6/13/8
TRID	3	10/75/16	13/33/19	13/77/17	18/35/19
	50	26/55/31	26/55/31	28/58/35	26/51/29
	100	30/67/36	30/67/36	31/68/38	29/63/34
	200	30/66/36	30/66/36	30/66/39	29/58/33
BAND	3	9/68/13	10/23/17	7/64/12	7/17/10
	50	18/183/24	16/331/25	16/184/23	17/182/22
	100	18/183/24	16/373/26	17/375/27	20/142/27
	200	19/283/27	17/340/27	17/147/25	17/327/22
LIN	2	1/3/3	1/3/3	1/3/3	1/3/3
	50	1/3/3	1/3/3	1/3/3	1/3/3
	500	1/3/3	1/3/3	1/3/3	1/3/3
	1000	1/3/3	1/3/3	1/3/3	1/3/3
LIN_1	2	1/51/2	1/51/2	1/51/2	1/51/2
	10	1/3/3	1/3/3	1/3/3	1/3/3
LIN_2	4	1/3/3	1/3/3	1/3/3	1/3/3

where

$S_1 = \{(i,j):$ method j does not work for example $i\}$.

The geometric mean of these ratios for method $EM(j)$ over all the test problems is defined by

$$r(EM(j)) = \left(\prod_{i \in S} r_i(EM(j))\right)^{1/|S|}, \qquad (4.3)$$

where S denotes the set of the test problems and $|S|$ is the number of elements in S. One advantage of the above rule is that, the comparison is relative and hence does not to be dominated by a few problems for which the method requires a great deal of function evaluations and gradient functions.

According to the above rules, it is clear that r (PRPSWP) = 1. The values of r (PRP+ SWP), r (NPRP+ SWP) and r (NPRP+ WWP) are listed in Table 2. As we can see from Table 2, the new method is much better than the PRP method.

Table 2 Relative efficiency of PRPSWP, PRP+ SWP, NPRP+ SWP, NPRP WWP

PRPSWP	PRP+ SWP	NPRP+ SWP	NPRP+ WWP
1	0.9049	0.8526	0.7725

5 Conclusion

By the combination of the variation of the formulae β_k^{FR} and β_k^{PRP}, we have found the new formula possesses the following features: (1) $\{\beta_k^{NPRP+}\}$ is a descent sequence without any line search; (2) the new method possesses the sufficient descent property and converge globally; (3) the strategy will restart the iteration automatically along the steepest descent direction if a negative value of β_k^{NPRP} occurs; (4) the initial numerical results are promising.

Acknowledgements

The authors are very grateful to the referees for their valuable comments and suggestions on the early version of the paper.

References

[1] Zoutendijk G. Nonlinear Programming Computational methods [M]. In: Integer and Nonlinear Programming. Amsterdam: North-Holland, 1970. 37-86.

[2] Al-Baali M. Descent property and global convergence of the Flecher-Reeves method with inexact line search [J]. IMA J Numer Anal, 1985, (5): 121-124.

[3] Dai Y, Yuan Y. An efficient hybrid conjugate gradient method for unconstrained optimization [J]. Annals of Optimizations Research, 2001, (103): 33-47.

[4] Gibert J C, Nocedal J. Global convergence properties of conjugate gradient methods for optimization [J]. SIAM Journal on Optimization, 1992, (2): 21-42.

[5] Wei Z, Li G, Qi L. New nonlinear conjugate gradient forumula for large-scale unconstrained optimization problems [D]. Department of Applied Mathematics, Hung Hom: The Hong kong Polytechnic University, 2002.

[6] Yu G, Zhao Y, Wei Z. A modification of the PRP conjugate gradient method, Department of Mathematics and Information Science [D]. Nanning: Guangxi University, 2003.

[7] Wolfe P. Convergence conditions for ascent methods [J]. SIAM Rew, 1969, (11): 226-235.

[8] Wolfe P. Convergence conditions for ascent methods, II: Some corrections [J]. SIAM Review, 1971, (13): 185-188.

[9] Dai Y, Yuan Y. Nonlinear Conjugate Methods [M]. Shanghai: Science Press of Shanghai, 2000. 24-26.

[10] Powell M J D. Nonconvex Minimization Calculations and the Conjugate Gradient Methds [M]. In: Lecture notes in Mayhematics. Berlin: Springer-Verlag, 1984. 122-141.

[11] Fletcher R, Reeves C. Function minimization by conjugate gradient [J]. Compute J, 1964, (7): 149-154.

(责任编辑: 蒋汉明 黎贞崇)

抗精神病药物思瑞康有助戒毒

美国匹兹堡大学医学院的专家给予213例门诊戒毒病人服用25mg装的思瑞康,每4h服用1~2粒,以帮助抵抗毒瘾发作。如果病人对这一剂量的思瑞康产生了耐药性,医生就会为其加量。41%的病人最终完成了疗程,他们能至少持续5天不吸毒。此后,该医学院对107例戒毒病人进行问卷调查,其中79例病人报告说思瑞康帮助他们缓解了毒瘾;52例病人称该药缓解了毒瘾发作时的焦虑症状;24例病人认为该药有助于减轻疼痛;22例病人感到失眠症状减少;14例病人觉得胃口大增。只有4例病人感觉没有效果,另外7例病人因无法忍受药物的副作用放弃了治疗。

(据科学网)