ISSN 1746-7659, England, UK
Journal of Information and Computing Science
Vol. 8, No. 1, 2013, *pp*. 075-080

A New Spectral Conjugate Gradient Method and Its Global Convergence

SUN Min [1] and LIU Jing [2]

[1] Department of Mathematics and Information Science，Zaozhuang University，Shandong 277160，China

[2] School of Mathematics and Statistics, Zhejiang University of Finance and Economics, Hangzhou, 310018, China

(*Received October05, 2012, accepted January04, 2013*)

Abstract. Combining the advantage of spectral-gradient method, a spectral conjugate gradient method for the global optimization is presented. This method has the property that the generated search direction is sufficiently descent without utilizing the line search. The proof of the convergence of the proposed method is given. Numerical experiments show that the method is efficient and feasible.

Keywords: *spectral conjugate gradient method; sufficient descent; global convergence*

1. Introduction

The unconstrained nonlinear optimization problems have many important applications in various fields. By the penalty function or merit function, many constrained optimization problems can be transformed into the unconstrained optimization problems. In this paper, we consider the following unconstrained optimization problem

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

where $f(x): R^n \to R^1$ is a continuously differentiable function. The conjugate gradient methods are quite efficient for problem (1), especially when its scale is quite large. Compared with the Newton method, the most important property of the conjugate gradient method is that it doesn't need to compute and store matrices. The iterative process of the conjugate gradient method is given by

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

where x_k is the current iterate, and α_k is the steplength, d_k is the descent direction of the objective function $f(x)$ at x_k, which is defined by

$$d_k = \begin{cases} -g_k, k = 0; \\ -g_k + \beta_k g_{k-1}, k \geq 1. \end{cases}$$

where g_{k-1} denotes $\nabla f(x_{k-1})$, and β_k is a parameter, which results in distinct conjugate gradient methods.

The following are some well-known β_k :

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

where $\|\cdot\|$ denotes the Euclidean norm.. Many researchers have studied the global convergence of the conjugate gradient methods, and obtained many meaningful results[1-5]. Recently, Birgin and Martinez[6] proposed a spectral type conjugate graident method, and the numerical results in [6] indicates that he performance of the spectral method is more efficiency than that of non-spectral method.. Then，Lu Aiguo further proposed a variant spectral-type FR conjugate gradient, which descent direction is defined by：

$$d_k = \begin{cases} -g_k, k = 0; \\ -\rho_k g_k + \beta_k^{VFR} d_{k-1}, k > 0. \end{cases}$$

where ρ_k, β_k^{VFR} are defined by $\rho_k = \frac{| g_k^T d_{k-1} | - g_{k-1}^T d_{k-1}}{\| g_{k-1} \|^2}, \beta_k^{VFR} = \frac{\| g_k \| | g_k^T g_{k-1} |}{\| g_{k-1} \|^3}$. Under the Wolfe line search or the Armijo line search, the method is global convergent. Motivated by the idea of [6,7], in this paper, we will combine the spectral gradient method with the conjugate gradient method , and propose a new spectral conjugate gradient method for problem (1). The parameter β_k in the new method is defined by

$$\beta_k^{NFR} = \frac{(g_k^T g_{k-1})^2}{\| g_{k-1} \|^4}. \tag{3}$$

Under some mild conditions, the global convergence of the proposed method is obtained, and the numerical tests are also given to show the efficiency of the proposed method.

The paper is organized as follows: In the next section, we will propose our algorithm. In Section 3, we will prove its global convergence. In Section 4, we will give the numerical tests.

2. The new spectral conjugate gradient method

Firstly, from the definition of β_k^{NFR}, we can obtain

$$0 \le \beta_k^{NFR} \le \frac{\|g_{k-1}\|^2}{\| g_{k-1} \|^2} = \beta_k^{FR}. \tag{4}$$

Thus, utilizing Zhang et. al.[8] ‘s idea, in order to let the iterative direction d_k satisfy the sufficient descent condition $g_k^T d_k = -\| g_k \|^2$, we only need to guarantee the following equality hold:

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

From (3) and $g_{k-1}^T d_{k-1} = -\| g_{k-1} \|^2$， we have

$$\theta_k = \frac{((g_{k-1}^T g_k)^2 g_k - \|g_k\|^2 \|g_{k-1}\|^2 g_{k-1})^T d_{k-1}}{\|g_{k-1}\|^4 \|g_k\|^2} \tag{5}$$

Thus, we obtain the following iterative direction

$$d_k = \begin{cases} -g_k, k = 0; \\ -\theta_k g_k + \beta_k^{NFR} d_{k-1}, k > 0. \end{cases} \tag{6}$$

Then, we can propose the following spectral conjugate gradient method (SCGM)):

Step 0: Choose $x_0 \in R^n$, $\gamma \in (0,1), \rho \in (0,1), \mu \ge 0,$ and set $k := 0;$

Step 1: If $\|g_k\| = 0$, then stop; Otherwise, go to Step 2;

Step 2: From（6）to compute the iterative direction d_k, and set α_k is the largest α in $\{1, \rho, \rho^2, \cdots\}$ such that

$$f(x_k + \alpha d_k) \le f(x_k) + \gamma \alpha g_k{}^T d_k - \mu \alpha^2 \|d_k\|^2. \tag{7}$$

Set $x_{k+1} = x_k + \alpha_k d_k$, $k = k+1$, and go to Step 1。

3. The global convergence of SCGM

Lemma 3.1 For any $k \ge 0$, we have

$$g_k{}^T d_k = -\| g_k \|^2 . \tag{8}$$

Proof. From the definition of θ_k, it is easy to prove (8)。

To prove the global convergence, we need the following assumptions:

(H1) The function $f(x)$ is bounded from below in the level set $L_0 = \{x \mid f(x) \le f(x_0)\}$.

(H2) $g(x)$ is Lipschitz continuous in some convex set which contains the level set L_0, and the Lipschitz constant is denoted by L.

Lemma 3.2. For any $k \ge 0$, there is a constant $c > 0$, such that

$$\alpha_k \ge c \frac{\|g_k\|^2}{\|d_k\|^2}. \tag{9}$$

Proof. Its proof is similar to that of Lemma 3.1 in [8].

From (8) and (9)，we can deduce the following Zoutendijk condition

$$\sum_{k=0}^{\infty} \frac{\|g_k\|^4}{\|d_k\|^2} < +\infty. \tag{10}$$

In the following, we will prove the global convergence of the proposed method.

Theorem 3.1. Under the assumption (H1) and (H2), the sequence $\{x_k\}$ generated by SCGM is global convergent. That is:

$$\liminf_{k \to \infty} \| g_k \| = 0. \tag{11}$$

Proof. If the conclusion doesn't hold, then there is a constant $\varepsilon > 0$, such that

$$\|g_k\| \ge \varepsilon, \forall k \ge 0.$$

From (6), we have

$$\|d_k\|^2 = (\beta_k^{NFR})^2 \|d_{k-1}\|^2 - 2\theta_k g_k^T d_k - \theta_k^2 \|g_k\|^2.$$

Hence, by (4) and (8), we obtain

$$\|d_k\|^2 \le \frac{\|g_k\|^4}{\|g_{k-1}\|^4} \|d_{k-1}\|^2 - 2\theta_k g_k^T d_k - \theta_k^2 \|g_k\|^2.$$

Then, dividing $\|g_k\|^4$ on the both sides of the above inequality, we can get

$$\frac{\|d_k\|^2}{\|g_k\|^4} \le \frac{\|d_{k-1}\|^2}{\|g_{k-1}\|^4} + \frac{2\theta_k}{\|g_k\|^2} - \theta_k^2 \frac{\|g_k\|^2}{\|g_k\|^4}$$

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

$$\le \frac{\|d_{k-1}\|^2}{\|g_{k-1}\|^4} + \frac{1}{\|g_k\|^2}$$

$$\le \sum_{i=0}^{k-1} \frac{1}{\|g_i\|^2} \le \frac{k}{\varepsilon^2}.$$

which implies that

$$\sum_{k\ge1} \frac{\|g_k\|^4}{\|d_k\|^2} \ge \varepsilon^2 \sum_{k\ge1} \frac{1}{k} = +\infty.$$

and this contradicts with the Zoutendijk condition (10). Therefore, we can get the conclusion (11).

4. Numerical results

In the following numerical experiment, we set the parameters $\gamma = 10^{-3}, \mu = 10^{-8}, \rho = 0.5$. The code is written by Matlab7.1, and is executed on a portable computer with PIV2.8GHz. The new method proposed in this paper is denoted by SCGM., and the method in paper [7] is denoted by VSCG。

Problem 1 $f(x) = x_1^2 + x_2^2 + 2x_3^2 + x_4^2 - 5(x_1 + x_2) - 21x_3 + 7x_4.$

Problem 2 $f(x) = (1-x_1)^2 + \sum_{i=2}^{9}(x_i^2 - x_{i+1})^2 + (1-x_{10})^2.$

Problem 3 $f(x) = e^{x_1} + x_1^2 + 2x_1x_2 + 4x_2^2.$

Our numerical results are listed in the form NF/NG/IT/CPU, where NF denotes the number of function evaluations, and NG denotes the number of gradient evaluation. IT denotes the number of iterations, and CPU denotes the CPU time

Table 1 The numerical results of the two methods(the left is SCGM, and the right is VSCG)

Problem	n	NF/NG/IT/CPU	Problem	n	NF/NG/IT/CPU
1	4	8/4/3/0/01	1	4	82/28/27/0.02
2	10	2354/668/667/0.1102	2	10	2884/195/194/0.0801
3	2	78/25/24/0.01	3	2	110/28/27/0.02

The numerical results in Table 1 indicate that the performance of SCGM is t is more efficiency than that of VSCG.

5. References

[1] Shi Zhenjun, Shen Jie, Convergence of Liu-Story conjugate gradient method. European Journal of Operational Research, 2007,182: 552-560.

[2] Shi Zhenjun, A new memory gradient method under exact line search. Asia-Pacific J. Oper. Res., 2003, 20(2): 275--284.

[3] Shi Zhenjun, A new Super-memory gradient method for unconstrained optimization. Advance in Mathcmatics, 2006, 35(3): 265-274.

[4] Zhang Li, Zhou Weijun, Li Donghui A descent modified Polak-Ribiere-Polyak conjugate gradient method and its global convergence. IMA Journal of Numerical Analysis,2006,26: 629-640.

[5] Yu Zhensheng. Global convergence of a memory gradient method without line search. Journal of Applied Mathematics and Computing, 2008, 26:545-553.

[6] E.Birgin, J.Martinez, A spectral conjugate gradient method for unconstrained minimization problem. Appl. Math. Optim. 2001, 43:117-128.

[7] Lu Aiguo, Liu Hongwei, Zhang Xiuyun, Cong Weijie, A variant spectral-type FR conjugate gradient method and its global convergence. Applied Mathematics and Computation, 2010, doi:10.1016/j.amc.2010.12.028.

[8] Zhang Li, Zhou Weijun, Li Donghui Global convergence of a modified Fletcher-Reeves conjugate gradient method with Armijo line search. Numer.Math. 2006, 104:561-572.