ISSN 1746-7659, England, UK
Journal of Information and Computing Science
Vol. 7, No. 3, 2012, *pp*. 212-217

A New Sufficient Descent Conjugate Gradient Method for Unconstrained Optimization

Sun Min [1+] and Liu Jing [2]

[1] School of Mathematics and Statistics, Zaozhuang University, Shandong 277160, China
[2] School of Mathematics and Statistics, Zhejiang University of Finance and Economics, Hangzhou, 310018, China
(*Received March 9, 2012, accepted July 10, 2012*)

Abstract. In this paper, a new conjugate conjugate method with sufficient descent property is proposed for the unconstrained optimization problem. An attractive property of the new method is that the descent direction generated by the method always possess the sufficient descent property, and this property is independent of the line search used and the choice of β_{ki}. Under mild conditions, the global convergence of the new method is proved.

Keywords: conjugate gradient method, sufficient descent property, global convergence

1. Introduction

In this paper, we consider the following unconstrained nonlinear optimization problem

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

where $f(x): R^n \to R^1$ is a continuously differentiable function whose gradient is denoted by $g(x)$.

The conjugate gradient methods are welcome methods for solving optimization problem (1). They are particularly efficient for solving large scale problems due to their simplicity and low storage [1,2]. A nonlinear conjugate gradient method generates a sequence $\{x_k\}$, starting from an initial point $x_0 \in R^n$, using the recurrence

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

where α_k is called the step size which is determined by some line search and d_k is the search direction defined by the rule

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

Here g_k is an abbreviation of $g(x_k)$ and β_k is a scalar which results in distinct conjugate gradient methods. Well known conjugate gradient methods include the Fletcher-Reeves(FR) method[1], the Polak-Ribiere-Polyak(PRP) method[2,3], the Dai-Yuan(DY) method[4] and the Conjugate Descent(CD) method[5] in which β_k are specified by

$$\beta_k^{FR} = \frac{\| g_k\|^2}{\| g_{k-1}\|^2}, \beta_k^{PRP} = \frac{g_k^\top (g_k - g_{k-1})}{d_{k-1}^\top (g_k - g_{k-1})} \quad \beta_k^{DY} = \frac{\| g_k\|^2}{d_{k-1}^\top (g_k - g_{k-1})}, \beta_k^{CD} = -\frac{\| g_k\|^2}{g_{k-1} d_{k-1}}$$

The global convergence of above conjugate gradient methods have bee studied by many researchers, see, for instance [6] and references therein.

Descent property, that is $g_k^\top d_k < 0$, is very important for an iterative method to be globally convergent. If

---

+ Corresponding author. Tel.: +86-0632-3786733.
E-mail address: ziyouxiaodou@163.com.

the step size α_k is determined by exact line search, we have $g_k^\bullet d_{k-1} = 0$. Then from (3), we get

$$g_k^\bullet d_k = -\| g_k \|^2, \tag{4}$$

which indicates that d_k satisfies the sufficient descent condition:

$$g_k^\bullet d_k \le -c\| g_k \|^2,$$

for a positive constant c, independently of line search. However, it is difficult or time consuming to implement an exact line search for seeking step length in practical computation. Therefore, in most existing conjugate gradient methods, the direction d_k determined by (3) may not be descent direction of f at x_k if inexact line search such as the Armijo line search is used.

Recently, Hager and Zhang[11] proposed a new conjugate gradient method which was obtained by modifying the HS method and called CG-DESCENT method. A nice property of the CG-DESCENT method is that it generates sufficient descent directions which is independent of the line search used. Then, Zhang Li et al.[12-13] also proposed some modified conjugate gradient methods which also possess the above property. Specially, at each iteration, the generated directions satisfy $g_k^\bullet d_k = -\| g_k \|^2$.More recently, using the Gram-Schmidt orthogonalization to d_{k-1} and g_k, Cheng and Liu [9] proposed a sufficient descent nonlinear conjugate gradient method defined by

$$d_k = \begin{cases} -g_k, & k = 0, \\ -g_k + \beta_k \left(d_{k-1} - \dfrac{g_k^\bullet d_{k-1}}{\| g_k \|^2} g_k \right), & k \ge 1. \end{cases} \tag{5}$$

It is clear that the direction in (5) satisfies (4). The sufficient descent property is independent of the line search used and the choice of β_k. Moreover, the scheme (5) reduces to the standard conjugate gradient method if exact line search is used.

Similar to conjugate gradient methods, memory gradient methods can also avoid storing and remembering matrices. The main difference between them is that the latter can use the information of the previous multi-step iterations more sufficiently and hence it is helpful to design algorithms with quick convergence rate. See Cragg and Lery[7], Miele and Cantrell[8].

In this paper, we continue Cheng and Liu's research and extend their method to memory gradient type method, which uses the previous m step information sufficiently. Under mild conditions, the global convergence of the new method is proved.

The paper is organized as follows: In Sect.2, we propose the new conjugate gradient method and show the global convergence. In Sect.3, we give the numerical tests and the conclusion is given in Sect.4..

2. Algorithm and Global Convergence

Let m denotes the number of the past iterations remembered. By using the Gram-Schmidt orthogonalization to d_{k-i} and g_k $(i = 1, 2, \cdots, m)$, we define the search direction as follows:

$$d_k = -g_k + \frac{1}{m} \sum_{i=1}^{m} \beta_{ki} \left(d_{k-i} - \frac{g_k^\bullet d_{k-i}}{\| g_k \|^2} g_k \right), \tag{6}$$