Journée Scientifique en Mathématiques AEO'2015
Souk-Ahras Le Dimanche 10 mai 2015

A new method for large-scale unconstrained optimization

Belloufi Mohammed [(1)*], Benzine Rachid [(2)]

[1] *Laboratory (LiM), Mohamed Cherif Messaadia University, Souk Ahras, Algeria.*
[2] *LANOS, Badji Mokhtar University, Annaba, Algeria.*

[*] *Corresponding author : mbelloufi_doc@yahoo.fr*

Abstract: The conjugate gradient (CG) method has played a special role in solving large-scale nonlinear optimization due to the simplicity of their iterations and their very low memory requirements. In this paper, a new nonlinear conjugate gradient method was proposed for large-scale unconstrained optimization. It is important that the proposed method produce sufficient descent search direction at every iteration with the strong Wolfe line searches, and the global convergence for general non-convex functions can be guaranteed. The numerical results show that one of our new CG methods is very encouraging.

Key word:Unconstrained optimization, Conjugate gradient method, strong Wolfe line search, Global convergence.

1 Introduction

Consider the unconstrained optimization problem

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

where $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ is continuously differentiable. The line search method usually takes the following iterative formula

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

The iterative formula of the conjugate gradient method is given by (1.2), where α_k is a steplength which is computed by carrying out a line search, and d_k is the search direction defined by

$$d_{k+1} = \begin{cases} -g_k & \text{si } k = 1 \\ -g_{k+1} + \beta_k d_k & \text{si } k \geq 2 \end{cases} \tag{1.3}$$

where β_k is a scalar and $g(x)$ denotes $\nabla f(x)$.
Conjugate gradient methods differ in their way of defining the scalar parameter β_k. In the literature, there have been proposed several choices for β_k which give rise to distinct conjugate gradient methods. The most well known conjugate gradient methods are the Hestenes–Stiefel (HS) method [11], the Fletcher–Reeves (FR) method [9], the Polak-Ribière-Polyak (PR) method [20,22], the Conjugate Descent method(CD) [8], the Liu-Storey (LS) method [14], the Dai-Yuan (DY) method [6], and Hager and Zhang (HZ) method [12].
The main aim of this note is to show that the descent property holds for all k and the global convergence is achieved for an inexact line search.
This paper is organized as follows. In the next section, the New algorithms are stated and descent property is presented. The global convergence of the new methods are established in Section 3. Numerical results and a conclusion are presented in Section 4 and in Section 5, respectively.

2 New Algorithm and descent property

In this section, we give the specific form of the proposed conjugate gradient method as follows. Then we can define the following descent direction as the search direction in (1.2),

$$d_k^{BB} = \begin{cases} -\frac{g_k}{\|g_k\|^2} & \text{if } k = 1 \\ -\frac{1}{\|g_k\|^2} g_k + d_{k-1} & \text{if } k \geq 2 \end{cases} \tag{2.1}$$

The following theorem indicates that, in the inexact case, the search direction d_k satisfies descent property.

Théorèm 2.1 *If the steplength α_k is computed by the Wolfe line search with $\delta < \sigma < \frac{1}{2}$, , then for the proposed conjugate gradient method, the inequality*

$$-\Sigma_{j=0}^{k-1} \sigma^j \leq g_k^T d_k \leq -2 + \Sigma_{j=0}^{k-1} \sigma^j \tag{2.2}$$

holds for all k, and hence the descent property

$$g_k^T d_k < 0, \forall k \tag{2.3}$$

holds, as long as $g_k \neq 0$.

3 Global convergence

In order to establish the global convergence of the proposed method, we assume that the following assumption always holds, i.e. Assumption 3.1 :
Assumption 3.1 :
Let f be twice continuously differentiable, and the level set $L = \{x \in \mathbb{R}^n \mid f(x) \leq f(x_1)\}$ be bounded

Théorèm 3.1 *Suppose that x_1 is a starting point for which Assumption 3.1 holds. Consider the New method (1.2) and (2.1). If the steplength α_k is computed by the strong Wolfe line search with $\delta < \sigma < \frac{1}{2}$, then the method is globally convergent, i.e.,*

$$\liminf_{k \longrightarrow \infty} \|g_k\| = 0 \tag{3.1}$$

4 Numerical results and discussions

In this section we report some numerical results obtained with an implementation of the $CGBB$ algorithm. For our numerical tests, we used test functions and Fortran programs from ([01],[03]). Considering the same criterias as in ([02]), the code is written in Fortran and compiled with f90 on a Workstation Intel Pentium 4 with 2 GHz. We selected a number of 105 unconstrained optimization test functions in generalized or extended form [17] (some from CUTE library [03]). For each test function we have taken twenty (20) numerical experiments with the number of variables increasing as $n = 2, 10, 30, 50, 70, 100, 300, 500, 700, 900, 1000, 2000, 3000, 4000, 5000,\ 6000, 7000, 8000, 9000, 10000$. The algorithm implements the Wolfe line search conditions (1.3) and (1.4), and the same stopping criterion $\|\nabla f(x_k)\| < 10^{-6}$. In all the algorithms we considered in this numerical study the maximum number of iterations is limited to 100000.
The comparisons of algorithms are given in the following context. Let f_i^{ALG1} and f_i^{ALG2} be the optimal value found by ALG1 and ALG2, for problem $i = 1, ..., 962$, respectively. We say that, in the particular problem i, the performance of ALG1 was better than the performance of ALG2 if:

$$\left| f_i^{ALG1} - f_i^{ALG2} \right| < 10^{-3}$$

and the number of iterations, or the number of function-gradient evaluations, or the CPU time of ALG1 was less than the number of iterations, or the number of function-gradient evaluations, or the CPU time corresponding to ALG2, respectively.
In a performance profile plot, the top curve corresponds to the method that solved the most problems in a time that was within a factor τ of the best time. The percentage of the test problems for which a method is the fastest is given on the left axis of the plot. The right side of the plot gives the percentage of the test problems that were successfully solved by these algorithms, respectively. Mainly, the right side is a measure of the robustness of an algorithm.
In the set of numerical experiments we compare $CGBB$ algorithm to *Steepest descent algorithm*, $CG\_DESCNET$, PRP and FR conjugate gradient methods.

Conclusion: In this paper, we have proposed a new and simple d_k that is easy to implement. We have also provided proof that this method converges globally with strong Wolfe line search.The presented numerical results illustrated the efficiency and robustness of our proposed method.
Our future work is concentrated on studying the convergence properties and numerical performance of our proposed method using different inexact line searches

Références

[1] N. Andrei, An unconstrained optimization test functions collection, Adv. Modell. Optim. 10 (2008) 147–161.

[2] N. Andrei, Another conjugate gradient algorithm for unconstrained optimization. Annals of Academy of Romanian Scientists, Series on Science and Technology of Information, vol. 1, nr.1, 2008, pp.7-20

[3] I. Bongartz, A. Conn, N. Gould, P. Toint, Cute: constrained and unconstrained testing environments, ACM Transaction on Mathematical Software 21 (1995) 123–160.

[4] R. Fletcher, Practical Method of Optimization, second ed., Unconstrained Optimization, vol. I, Wiley, New York, 1997.

[5] M.R. Hestenes, E. Stiefel, Method of conjugate gradient for solving linear equations, J. Res. Nat. Bur. Stand. 49 (1952) 409–436.

[6] W.W. Hager, H. Zhang, A new conjugate gradient method with guaranteed descent and an efficient line search, SIAM Journal on Optimization 16 (2005) 170–192.

[7] B.T. Polyak, The conjugate gradient method in extreme problems, USSR Comp. Math. Math. Phys. 9 (1969) 94–112.

[8] M.J.D. Powell, Nonconvex minimization calculations and the conjugate gradient method, Lecture Notes in Mathematics, 1066, Springer-Verlag, Berlin, 1984, pp. 122–141.

[9] M. Raydan, The Barzilai and Borwein gradient method for the large scale unconstrained minimization problem. SIAM Journal of Optimization, 7 (1997) 26–33.

[10] P.Wolfe, Convergence conditions for ascent méthods. *Siam Review* ,11,pp.226-235 ,(1969) .

[11] G. Zoutendijk, Nonlinear programming computational methods, in: J. Abadie (Ed.), Integer and Nonlinear Programming, North-Holland, Amsterdam, 1970, pp. 3