Global Analysis and Discrete Mathematics
Volume 2, Issue 2, pp. 145–149
ISSN: 2476-5341

A Note on the Descent Property Theorem for the Hybrid Conjugate Gradient Algorithm CCOMB Proposed by Andrei

Hadi Nosratipour* · **Farhad Sarani**

Received: 4 April 2017 / Accepted: 10 October 2017

Abstract In [1] (Hybrid Conjugate Gradient Algorithm for Unconstrained Optimization J. Optim. Theory Appl. 141 (2009) 249 - 264), an efficient hybrid conjugate gradient algorithm, the CCOMB algorithm is proposed for solving unconstrained optimization problems. However, the proof of Theorem 2.1 in [1] is incorrect due to an erroneous inequality which used to indicate the descent property for the search direction of the CCOMB algorithm. It is also remarkable that the proof of the Theorem 2.2 should be revised. Following the notations in [1], the main goal of this note is to provide some necessary corrections to rectify the mentioned issues.

Keywords Unconstrained optimization · Hybrid conjugate gradient method · Sufficient descent condition · Conjugacy condition

Mathematics Subject Classification (2010) 65K05 · 90C30

1 Introduction

The conjugate gradient method represents a major contribution to the panoply of methods for solving large-scale unconstrained optimization problems. They

*Corresponding author
Hadi Nosratipour
Department of Applied Mathematics, School of Mathematics and Computer Science, Damghan University, Damghan, Iran.
E-mail: hadi.nosratipour@gmail.com, h.nosratipour@std.du.ac.ir

Farhad Sarani
Department of Applied Mathematics, School of Mathematics and Computer Science, Damghan University, Damghan, Iran.
E-mail: f.sarani@std.du.ac.ir

are characterized by low memory requirements and have strong local and global convergence properties [4,3]. The popularity of these methods is remarkable partially due to their simplicity both in their algebraic expression and in their implementation in computer codes, and partially due to their efficiency in solving large-scale unconstrained optimization problems

$$\text{minimize } f(x), \quad \text{Subject to} \quad x \in \mathbb{R}^n. \tag{1}$$

Recently, by a convex combination of PRP conjugate gradient method suggested by Polak and Ribiére [5] and DY conjugate gradient method suggested by Dai and Yuan [2], Andrei [1] proposed the hybrid CG algorithm CCOMB which is numerically efficient. CCOMB is an iterative method in the following form,

$$\begin{aligned} &x_0 \in \mathbb{R}^n \\ &x_{k+1} = x_k + s_k, \quad s_k = \alpha_k d_k \quad k = 0, 1, 2, ..., \end{aligned} \tag{2}$$

where $\alpha_k > 0$ is obtained by line search and the directions d_k are generated as

$$d_{k+1} = -g_{k+1} + \beta_k^N s_k, \quad d_0 = -g_0. \tag{3}$$

where $g_k = \nabla f(x_k)$ and the scalar parameter β_k^N is defined by

$$\beta_k^N = (1-\theta_k)\beta_k^{PRP} + \theta_k \beta_k^{DY} = (1-\theta_k)\frac{y_k^T g_{k+1}}{g_k^T g_k} + \theta_k \frac{g_{k+1}^T g_{k+1}}{y_k^T s_k}, \tag{4}$$

where $y_k = g_{k+1} - g_k$ and θ_k is a scalar parameter satisfying $0 \leq \theta_k \leq 1$, which will be determined in a specific way to be described later. Observe that if, $\theta_k = 0$, then, $\beta_k^N = \beta_k^{PRP}$, and if, $\theta_k = 1$, then, $\beta_k^N = \beta_k^{DY}$. On the other hand, if $0 < \theta_k < 1$, then β_k^N is a convex combination of β_k^{PRP} and β_k^{DY}.

In the CCOMB algorithm, the parameter θ_k is selected in such a way that at every iteration the conjugacy condition $y_k^T d_{k+1} = 0$ is satisfied independently of the line search. Obviously,

$$d_{k+1} = -g_{k+1} + (1-\theta_k)\frac{y_k^T g_{k+1}}{g_k^T g_k} s_k + \theta_k \frac{g_{k+1}^T g_{k+1}}{y_k^T s_k} s_k. \tag{5}$$

Hence, after some algebra effort, one get

$$\theta_k = \frac{(y_k^T g_{k+1})(y_k^T s_k) - (y_k^T g_{k+1})(g_k^T g_k)}{(y_k^T g_{k+1})(y_k^T s_k) - (g_{k+1}^T g_{k+1})(g_k^T g_k)} \tag{6}$$

In the CCOMB algorithm, the steplength α_k is selected to satisfy the so-called Wolfe line search conditions [6,7], requiring that

$$\begin{aligned} f(x_k + \alpha_k d_k) &\leq f(x_k) + \delta \alpha_k g_k^T d_k, \\ g_{k+1}^T d_k &\geq \sigma g_k^T d_k, \end{aligned} \tag{7}$$

where $0 < \delta < \sigma < 1$.

Although the CCOMB algorithm is more robust than PRP and DY conjugate gradient algorithms, the proof of Theorem 2.1, the descent property theorem of the CCOMB, and a part of the proof of Theorem 2.2, the sufficient descent property theorem of the CCOMB, are incorrect. Here, we first state the wrongs occurred in the proofs of Theorems 2.1 and 2.2 in [1] and then, to complete the proof of Theorem 2.1 in [1], we will show that the search directions generated by the CCOMB are descent.

2 Two wrongs in analysis of the CCOMB algorithm

In the proof of Theorem 2.1 in [1], according to $0 < \theta_k < 1$ and Eq. (5) (Eq. (8) in [1]), authors of [1] have claimed that the search direction of the CCOMB is decreasing, that is,

$$g_k^T d_k \leq 0, \quad \forall k \geq 0.$$

Hence, they deduce that

$$g_{k+1}^T d_{k+1} = -\|g_{k+1}\|^2 + (1-\theta_k)\frac{y_k^T g_{k+1}}{g_k^T g_k} g_{k+1}^T s_k + \theta_k \frac{g_{k+1}^T g_{k+1}}{y_k^T s_k} g_{k+1}^T s_k \tag{8a}$$

$$\leq -\|g_{k+1}\|^2 + \frac{y_k^T g_{k+1}}{g_k^T g_k} g_{k+1}^T s_k + \frac{g_{k+1}^T g_{k+1}}{y_k^T s_k} g_{k+1}^T s_k, \tag{8b}$$

but, since the terms of $y_k^T g_{k+1}$ and $g_{k+1}^T s_k$ can be positive or negative, so, the signs of the second and the third parts of the right hand side of (8a) are not clear at all. Hence, the inequality of (8b) is an incorrect conclusion. Therefore, another strategy should be considered to prove the descent property for the search directions generated by the CCOMB, which is the subject of the next section.

On the other hand, in the proof of Theorem 2.2 in [1], the sufficient descent property of the search directions of the CCOMB, that is,

$$g_k^T d_k \leq -c\|g_k\|^2, \quad \forall k \geq 0,$$

with some positive constant c, has been proved. In the proof presented [1], using $y_k^T s_k > 0$, which is due to the curvature condition (inequality (5) in [1]), and

$$g_{k+1}^T s_k = y_k^T s_k + g_k^T s_k < y_k^T s_k \tag{9}$$

the following result has been taken

$$\frac{y_k^T s_k}{g_{k+1}^T s_k} > 1. \tag{10}$$

Note that the term of $g_{k+1}^T s_k$ can be positive or negative. Hence, the inequality (10) is a wrong result and it must be replaced by

$$\frac{g_{k+1}^T s_k}{y_k^T s_k} < 1. \tag{11}$$

3 Descent property for the CCOMB algorithm

The mentioned issues occurred in the analyzing of the CCOMB motivated us to study the descent property for this algorithm. The following theorem guarantees the descent property of the CCOMB and gives a modified proof to Theorem 2.1 in [1].

Theorem 1 *In the algorithm (2), (5) and (6), assume that* α_k *is determined by the Wolfe line search (7). If* $0 < \theta_k < 1$*, then the direction* d_{k+1} *given by (5) is a descent direction.*

Proof. We consider the following situation.

Case 1: $g_{k+1}^T s_k > 0$ and $y_k^T g_{k+1} > 0$;

Case 2: $g_{k+1}^T s_k > 0$ and $y_k^T g_{k+1} < 0$;

Case 3: $g_{k+1}^T s_k < 0$ and $y_k^T g_{k+1} > 0$;

Case 4: $g_{k+1}^T s_k < 0$ and $y_k^T g_{k+1} < 0$.

*Case*1. In this situation, the proof of [1] is correct.

*Case*2. Since $0 < \theta_k < 1$, from (5) we get

$$\begin{aligned} g_{k+1}^T d_{k+1} &= -\|g_{k+1}\|^2 + (1-\theta_k)\frac{y_k^T g_{k+1}}{g_k^T g_k} g_{k+1}^T s_k + \theta_k \frac{g_{k+1}^T g_{k+1}}{y_k^T s_k} g_{k+1}^T s_k \\ &\leq -\|g_{k+1}\|^2 + \frac{g_{k+1}^T g_{k+1}}{y_k^T s_k} g_{k+1}^T s_k \\ &= (-1 + \frac{g_{k+1}^T s_k}{y_k^T s_k})\|g_{k+1}\|^2 \\ &= \frac{g_k^T s_k}{y_k^T s_k}\|g_{k+1}\|^2 \leq 0. \end{aligned}$$

*Case*3. In this situation, since the second and the third parts of the right hand side of (8a) are negative, therefore, it follows from (8) that $g_{k+1}^T d_{k+1} \leq 0$, i.e., d_{k+1} is a descent direction.

*Case*4. Since $0 < \theta_k < 1$, from (5) we get

$$\begin{aligned} g_{k+1}^T d_{k+1} &= -\|g_{k+1}\|^2 + (1-\theta_k)\frac{y_k^T g_{k+1}}{g_k^T g_k} g_{k+1}^T s_k + \theta_k \frac{g_{k+1}^T g_{k+1}}{y_k^T s_k} g_{k+1}^T s_k \\ &\leq -\|g_{k+1}\|^2 + \theta_k \frac{g_{k+1}^T g_{k+1}}{y_k^T s_k} g_{k+1}^T s_k + \frac{y_k^T g_{k+1}}{g_k^T g_k} g_{k+1}^T s_k. \end{aligned}$$

Observe that y_k becomes tiny while $\|g_k\|$ is bounded away from zero. Consequently, the last term in the above inequality becomes negligible. From $0 < \theta_k < 1$ and (11), we have $(1 - \theta_k \frac{g_{k+1}^T s_k}{y_k^T s_k}) > 0$, that is

$$g_{k+1}^T d_{k+1} \leq -(1 - \theta_k \frac{g_{k+1}^T s_k}{y_k^T s_k}) \|g_{k+1}\|^2 \leq 0.$$

□

References

1. N. Andrei, Hybrid conjugate gradient algorithm for unconstrained optimization, J. Optim. Theory Appl 141(2), 249–264 (2009).
2. Y.H. Dai, Y. Yuan, A nonlinear conjugate gradient method with a strong global convergenc eproperty, SIAM Journal Optimization, 10(21), 348–358 (1991).
3. Y.H. Dai, J.Y. Han, G.H. Liu, and D.F. Sun, Convergence properties of nonlinear conjugate gradient methods, SIAM Journal Optimization, 10(21), 348–358 (1991).
4. J.C. Gilbert, J. Nocedal, Global convergence properties of conjugate gradient methods for optimization, SIAM Journal Optimization, 2(1), 21–42 (1992).
5. R. Polak, G. Ribiere, Note sur la convergence de mthodes de directions conjugues, Rev. Fr. Inf. Rech. Oper. 3e Annee, 3(1), 35–43 (1969).
6. P. Wolfe, Convergence conditions for ascent methods, SIAM Rev., 11(2), 226–235 (1969).
7. P. Wolfe, Convergence conditions for ascent methods, II: Some corrections, SIAM Rev., 13(2), 185–188 (1971).