Applied Mathematics and Computation 269 (2015) 606–617

Contents lists available at ScienceDirect

Applied Mathematics and Computation

journal homepage: www.elsevier.com/locate/amc

New version of the three-term conjugate gradient method based on spectral scaling conjugacy condition that generates descent search direction

Xiao Liang Dong[a,*], Hong Wei Liu[b], Yu Bo He[c]

[a] School of Mathematics and Information, Beifang University of Nationalities, Yinchuan, 710021, China
[b] School of Mathematics and Statistics, Xidian University, Xi'an 710126, People's Republic of China
[c] Department of Mathematics and Applied Mathematics, Huaihua University, Huaihua 418008, People's Republic of China

ARTICLE INFO

Keywords:
Three-term conjugate gradient method
Global convergence
Sufficient descent condition
Spectral scaling conjugacy condition

ABSTRACT

In this paper, a general form of three-term conjugate gradient method is presented, in which the search directions simultaneously satisfy the Dai–Liao conjugacy condition and sufficient descent property. In addition, the choice for an optimal parameter is suggested in the sense that the condition number of the iteration matrix could arrives at its minimum, which can be regarded as the inheritance and development of the spectral scaling quasi-Newton equation. Different from the existent methods, a new update strategy in constructing the search direction is proposed to establish the global convergence for the general function. Numerical results show our algorithm is practical and effective for the test problems.

1. Introduction

Consider the unconstrained optimization problem

$$\min_{x\in\mathbb{R}^n} f(x). \tag{1}$$

where $f:\mathbb{R}^n\to\mathbb{R}$ is assumed to be continuously differentiable, whose gradient $g(x)=\nabla f(x)$ is available. More precisely, we assume that f satisfies the some basic assumptions:

Assumption 1.1. (i) The level set, defined by $\Omega=\{x\in\mathbb{R}^n|f(x)\le f(x_1)\}$ with x_1 to be the initial point, is bounded; (ii) In some neighborhood Ω_0 of Ω, g is Lipschitz continuous, namely, there exists a constant $L>0$ such that $||g(x)-g(y)||\le L||x-y||$, $\forall x,y\in\Omega_0$.

Clearly, there exist constants $B>0$ and $\gamma>0$ such that

$$||x-y||\le B,\ ||g(x)||\le\gamma,\ \forall x,y\in\Omega. \tag{2}$$

Conjugate gradient methods are a class of very prominent iterative methods for solving large-scale unconstrained optimization problems. The major advantages of this class of algorithms lies in the fact that they possess modest memory requirement and simple computational scheme, that is,

$$x_{k+1}=x_k+s_k,\ s_k=\alpha_k d_k. \tag{3}$$

* Corresponding author. Tel.: +8602981891371.
E-mail addresses: dongxl@stu.xidian.edu.cn (X.L. Dong), hwliu@mail.xidian.edu.cn (H.W. Liu), heyinprc@csu.edu.cn (Y.B. He).

http://dx.doi.org/10.1016/j.amc.2015.07.067
0096-3003/

For convergence analysis of the CG method, the Wolfe line search is commonly chosen to determine the steplength α_k, requiring that

$$f(x_k + \alpha_k d_k) - f(x_k) \le \rho \alpha_k g_k^T d_k, \tag{4}$$

$$g(x_k + \alpha_k d_k)^T d_k \ge \sigma g_k^T d_k, \tag{5}$$

where $0 < \rho < \sigma < 1$ are constants. The search direction d_k is of the form:

$$d_k = \begin{cases} -g_k, & k = 1, \\ -g_k + \beta_k d_{k-1}, & k \ge 2. \end{cases} \tag{6}$$

With a distinct choice of the parameter β_k, the obtained method has different theoretical property and numerical performance. The leading parameters formulate for β_k are called the Polak–Ribière–Polyak [1], Liu–Store [2], Hestenes–Stiefel [3], Fletcher–Reeves [4] Conjugate–Descent [5] and Dai–Yuan [6] formulate. More recent reviews on nonlinear conjugate gradient methods can be found in Hager and Zhang [7].

Here and throughout the paper, we always use $||\cdot||$ to stand for the Euclidian norm of vectors, $y_{k-1} = g_k - g_{k-1}$. We also use $det(A)$, $tr(A)$ and $\kappa_2(A)$ to stand for the determinant, trace and condition number of a square matrix A.

More recently, many researchers highlighted two properties in designing new CG methods, called the conjugacy condition and descent property, which play a crucial role in obtaining global convergence and nice actual performance.

In order to accelerate the CG method, the conjugacy condition is often utilized to obtain the order accuracy in the approximation of the curvature of the function as high as possible. By modifying the HS method, Dai and Liao [8] proposed the Dai–Liao (DL) conjugacy condition

$$d_k^T y_{k-1} = -\xi g_k^T s_{k-1}. \tag{7}$$

Very recently, Babaie–Kafaki and Reza discussed the optimal choice for the parameter ξ, see [9,10].

The latter property is an indispensable factor in the convergence analysis of CG methods. Exactly, a direction d_k satisfies the so-called sufficient descent condition, if there exists a constant $c > 0$ such that

$$d_k^T g_k \le -c\|g_k\|^2, \forall k \in N. \tag{8}$$

The papers by Hager and Zhang [11], Dai and Kou [12], Yu et al. [13] and Cheng [14] proposed different versions conjugate gradient methods together with their global convergence, in which the corresponding search directions satisfy (8) independent of the line search used. Extensive numerical results showed the exploitation of new techniques greatly enhance the numerical efficiency.

The study on the three-term CG method has also made good progress, which take the advantages of the above nice properties and therefore can be viewed as another important computational innovation to solve the problem (1). Beale [15], McGuire and Wolfe [16], Powell [17], Dai and Yuan [18] researched into this topic and established some efficient restart strategies and obtained good computational performance. Recently, Zhang, Zhou and Li proposed a MPRP [19] method, in which the property $-d_k^T g_k = \|g_k\|^2$ always be satisfied independent of any line searches used and the convexity of the objective function. Due to its simplicity and efficiency, the MPRP approach and its variants received great attention, see [20–28] and references therein.

In this paper, we proceed to study the three-term CG method. An example is first presented to explain the reason for the inappropriate adoption of assumption condition of the TTCG method [24]. Then, a new three-term CG method is motivated, derived and analyzed, which preserves the conjugacy and sufficient descent conditions previously discussed. Also, satisfactory of an optimal parameter choice makes it easier for our proposed algorithm converges rapidly to the solution or reduces the condition number.

The paper is organized as follows. Section. 2 describes the motivation and some properties of our method. In Section. 3, an optimal parameter is discussed, and our method is proposed formally. Sections. 4 and 5 are devoted to providing the convergence analysis and numerical results, respectively. Finally, conclusions are made.

2. Motivation and properties

We start this section by describing our motivation, and then present some properties of our proposed method.

More recently, combining the DL conjugacy condition and sufficient descent condition, N. Andrei proposed the TTCG [24] method, which is closely related to the THREECG [23] presented by the same author. In contrast, the TTCG can be globalized not only for the uniformly convex functions but also for the general functions. A mass of numerical experiments showed that they are practicable and efficient for the large-scale test problems.

Acknowledging that these methods have their merits, a boundedness condition that are assumed in the TTCG method seems not hold for all the general objective functions. The reason is that in the convergence analysis, he mistook the sequence $\{s_k^T y_k\}$ for a subsequence $\{s_k^T y_k\}_K$ may converged to zero, particularly during the late iterations when closed to the solution. In fact, if we slightly modify some appropriate conditions, the convergence theorem may be obtained.