Applied Mathematics and Computation 245 (2014) 36–43

Contents lists available at ScienceDirect

Applied Mathematics and Computation

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

New hybrid conjugate gradient method for unconstrained optimization ☆

J.K. Liu [a,b], S.J. Li [a,*]

[a] College of Mathematics and Statistics, Chongqing University, Chongqing 401331, China
[b] School of Mathematics and Statistics, Chongqing Three Gorges University, Chongqing 404100, China

ARTICLE INFO

Keywords:
Unconstrained optimization
Conjugate gradient method
Hybrid conjugate gradient method
Global convergence

ABSTRACT

In this paper, we propose a new hybrid conjugate gradient method for solving unconstrained optimization problems. The proposed method can be viewed as a convex combination of Liu–Storey method and Dai–Yuan method. An remarkable property is that the search direction of this method not only satisfies the famous D–L conjugacy condition, but also accords with the Newton direction with suitable condition. Furthermore, this property is not dependent on any line searches. Under the strong Wolfe line searches, the global convergence of the proposed method is established. Preliminary numerical results also show that our method is robust and effective.

1. Introduction

In unconstrained optimization, we minimize an objective function that depends on real variable, with no restrictions at all on the values of these variables. The mathematical formulation is

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

where $x \in R^n$ is a real vector with $n \geqslant 1$ component, and $f : R^n \to R$ is a smooth function and its gradient g is available.

When applied to solve the problem (1), starting from an initial guess $x_0 \in R^n$, conjugate gradient (CG) method usually generates a sequence $\{x_k\}$ as

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

where x_k is the current iterate, $\alpha_k > 0$ is called a stepsize determined by some line searches. d_k is the search direction defined by

$$d_{k+1} = -g_{k+1} + \beta_k d_k, \quad d_0 = -g_0, \tag{3}$$

where $g_k = \nabla f(x_k)$, and β_k is an important parameter. The different choices for the parameter β_k correspond to different CG methods. Over the years, many variants of this scheme are proposed, and some are widely used in practice. For instance,

☆ This research was partially supported by the National Natural Science Foundation of China (Grant No.: 11171362), Doctoral Foundation (Doctoral tutor category) of Ministry of Education of China (Grant No.: 20120191110031) and the fund of Scientific research in Southeast University (the support project of fundamental research).

* Corresponding author.
E-mail address: lisj@cqu.edu.cn (S.J. Li).

http://dx.doi.org/10.1016/j.amc.2014.07.096
0096-3003/

Fletcher–Reeves method (FR) [1], Polak–Ribière–Polyak method (PRP) [2,3], Hestenes–Stiefel method (HS) [4], Liu–Storey method (LS) [5], Dai–Yuan method (DY)[6], Conjugate-Descent method (CD) [7], etc. These methods are identical when f is a strongly convex quadratic function and the line search is exact, since the gradients are mutually orthogonal, and the parameters β_k in these methods are equal. When applied to general nonlinear functions with inexact line searches, however, the behavior of theses methods is marked different.

It is well known that FR, DY and CD methods have strong convergent properties, but they may not perform well in practice due to jamming. Moreover, although PRP, HS and LS methods may not converge in general, they often perform better. Naturally, people try to devise some new methods, which have the advantages of these two kinds of methods. So far various hybrid methods have been proposed. For example, in 1990, Touati-Ahmed and Storey [8] firstly proposed a hybrid conjugate gradient method using PRP and FR methods. Soon afterwards, Hu and Storey [9], Gilbert and Nocedal [10] further studied other hybrid schemes about PRP and FR methods. In order to improve the application of DY method in practice, Dai and Yuan [11] combined DY method with HS method to propose two hybrid CG methods. Recently, Andrei [12] introduced a new hybrid CG method (denoted as HYBRID method) based on HS and DY methods for large-scaled unconstrained optimization problems. The main feature of this hybrid method is that the search direction is the Newton direction. Surprisingly, this hybrid method outperforms some sophisticated conjugate gradient methods for many problems.

LS method usually performs better in practice than DY method, and DY method has stronger convergent properties than LS method as previously discussed. In order to take in the advantages of LS and DY methods and establish a more efficient and robust algorithm, and inspired by the work of Andrei [12], we propose a new hybrid CG method based on LS and DY methods for solving unconstrained optimization problems with suitable conditions. The parameter β_k in the proposed method is computed as a convex combination of β_k^{LS} and β_k^{DY}, i.e.,

$$\beta_k = (1-\theta_k)\beta_k^{LS} + \theta_k\beta_k^{DY}, \tag{4}$$

where $\beta_k^{LS} = -\frac{g_{k+1}^T y_k}{d_k^T g_k}$, $\beta_k^{DY} = \frac{\|g_{k+1}\|^2}{d_k^T y_k}$, $y_k = g_{k+1} - g_k$ and $\theta_k \in [0,1]$. By choosing the appropriate value of the parameter θ_k in the convex combination, the search direction d_k of our algorithm not only is the Newton direction, but also satisfies the famous D–L conjugacy condition proposed by Dai and Liao [13]. This feature is not dependent on any line search. Simultaneously, under the strong Wolfe line searches, we prove the global convergence of our algorithm. The numerical results also show the feasibility and effectiveness of our algorithm.

The paper is organized as follows. In the next section, we obtain the parameter θ_k using some approaches and give our specific algorithm. The sufficient descent property of the proposed method is also discussed under the suitable conditions. In Section 3, some assumptions are given, and the global convergence of the proposed method is established. Preliminary numerical results are presented in Section 4. Finally, we make a summary for our paper.

2. A new hybrid conjugate gradient method

In CG method, the traditional conjugacy condition $d_{k+1}^T y_k = 0$, which depends on the exact line search, plays an important role in the convergence analysis and numerical calculation. However, in practical computation, one normally uses some inexact line searches instead of the exact line search to obtain the stepsize α_k. In this case, $d_{k+1}^T y_k$ may be not equal to zero. Moreover, this condition may have some disadvantages (for instance, see [14]). For this reason, Dai and Liao [13] proposed a weaker conjugacy condition, i.e.,

$$d_{k+1}^T y_k = -t s_k^T g_{k+1}, \quad t \geqslant 0, \tag{5}$$

where $s_k = x_{k+1} - x_k$. This is the famous D–L conjugacy condition which implies that traditional conjugacy condition holds if the line search is exact. Simultaneously, they also demonstrated that it seems more reasonable to replace traditional conjugacy condition with the D–L conjugacy condition in practice. Hence, using the similar way in [15], we firstly select the parameter θ_k so that the search direction d_{k+1} satisfies the famous D–L conjugacy condition.

By taking the inner product of (3) with the vector y_k^T and computing β_k by (4), we obtain

$$\begin{aligned} d_{k+1}^T y_k &= -g_{k+1}^T y_k + \beta_k d_k^T y_k = -g_{k+1}^T y_k + (1-\theta_k)\beta_k^{LS} \cdot d_k^T y_k + \theta_k \beta_k^{DY} \cdot d_k^T y_k \\ &= -g_{k+1}^T y_k - (1-\theta_k) \cdot \frac{g_{k+1}^T y_k}{d_k^T g_k} \cdot d_k^T y_k + \theta_k \cdot \frac{\|g_{k+1}\|^2}{d_k^T y_k} \cdot d_k^T y_k. \end{aligned}$$

From the above equality and (5), after some algebra we have

$$\theta_k^{DL} = \frac{g_{k+1}^T y_k \cdot d_k^T g_{k+1} - t \cdot s_k^T g_{k+1} \cdot d_k^T g_k}{\|\, g_{k+1} \,\|^2 \cdot d_k^T g_k + g_{k+1}^T y_k \cdot d_k^T y_k}. \tag{6}$$

It is well-known that the Newton method has quadratical convergent property for solving unconstrained optimization problems. To some extent, this depends on its search direction. Hence, we assume that $\nabla^2 f(x)^{-1}$ is existent at each iterative point for the objective function f, and choose the parameter θ_k so that the search direction d_{k+1} defined by (3) satisfies the condition of the Newton direction, i.e.,