Applied Mathematics and Computation 231 (2014) 179–186

Contents lists available at ScienceDirect

Applied Mathematics and Computation

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

A modified Wei–Yao–Liu conjugate gradient method for unconstrained optimization ☆

Hai Huang *, Suihua Lin

Department of Mathematics and Computer Science, Guangxi Normal University for Nationalities, Chongzuo, Guangxi 532200, PR China

ARTICLE INFO

Keywords:
Unconstrained optimization
Conjugate gradient method
Sufficient descent condition
Global convergence

ABSTRACT

In this paper, we give a modified Wei–Yao–Liu conjugate gradient method (Wei et al., 2006 [18]), which will reduce to the original Wei–Yao–Liu method, and possess the sufficient descent property without any line search. Furthermore, we prove that the presented method is globally convergent for nonconvex functions with the weak Wolfe–Powell line search. In a similar way, we also extend these results to the modified Liu–Storey method. Preliminary numerical results show that the proposed methods are effective for the given test problems.

1. Introduction

Conjugate gradient methods are a class of important methods for unconstrained optimization, especially when the dimension is large. We consider the following problem:

$$\min\{f(x) \mid x \in \Re^n\}, \tag{1.1}$$

where $f : \Re^n \to \Re$ is smooth and its gradient g is available.

The iterate of conjugate gradient method for solving (1.1) is given by

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

where α_k is a step-size which is computed by carrying out some line search, and d_k is the search direction defined by

$$d_k = \begin{cases} -g_k & \text{if } k = 1, \\ -g_k + \beta_k d_{k-1} & \text{if } k \geqslant 2, \end{cases} \tag{1.3}$$

where $g_k = \nabla f(x_k)$, $\beta_k \in \Re$ is a scalar and can be defined by

$$\beta_k^{HS} = \frac{g_k^T y_{k-1}}{d_{k-1}^T y_{k-1}}, \quad \beta_k^{FR} = \frac{\|g_k\|^2}{\|g_{k-1}\|^2}, \quad \beta_k^{PRP} = \frac{g_k^T y_{k-1}}{\|g_{k-1}\|^2},$$

$$\beta_k^{CD} = -\frac{\|g_k\|^2}{d_{k-1}^T g_{k-1}}, \quad \beta_k^{LS} = -\frac{g_k^T y_{k-1}}{d_{k-1}^T g_{k-1}}, \quad \beta_k^{DY} = \frac{\|g_k\|^2}{d_{k-1}^T y_{k-1}}$$

☆ This work was supported by the Scientific Research Fund of Guangxi Normal University for Nationalities (No. ybxm200909) and the Research Program of Guangxi Department of Education (No. 201012MS215).

* Corresponding author.
E-mail addresses: hhl6973@yahoo.cn (H. Huang), linsuihuah@163.com (S. Lin).

0096-3003/$ - see front matter © 2014 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.amc.2014.01.012

or by other formulae (e.g. [7,16,22–24]), where $y_{k-1} = g_k - g_{k-1}$ and $\|\cdot\|$ stands for the Euclidean norm of vectors. The corresponding methods are respectively called HS [9], FR [4], PRP [13,14], CD (conjugate descent [5]), LS [11] and DY [2] conjugate gradient method.

Generally, in the analysis and implementation of these conjugate gradient methods, the step-size α_k is required to satisfy some line search conditions. The above mentioned conjugated gradient methods are identical when f is a strongly convex quadratic function and the line search is exact, the behavior of these methods differs markedly when f is a general function and the line search is inexact (see [1–9,11,13–17,19,23,24]). Usually, two major inexact line searches are the weak Wolfe–Powell (WWP) line search, i.e.,

$$\begin{cases} f(x_k + \alpha_k d_k) - f(x_k) \leqslant \delta \alpha_k g_k^T d_k, \\ g(x_k + \alpha_k d_k)^T d_k \geqslant \sigma g_k^T d_k, \end{cases} \tag{1.4}$$

and the strong Wolfe–Powell (SWP) line search, namely,

$$\begin{cases} f(x_k + \alpha_k d_k) - f(x_k) \leqslant \delta \alpha_k g_k^T d_k, \\ |g(x_k + \alpha_k d_k)^T d_k| \leqslant -\sigma g_k^T d_k, \end{cases} \tag{1.5}$$

where $\delta \in (0, \frac{1}{2})$ and $\sigma \in (\delta, 1)$.

In addition, the sufficient descent condition

$$g_k^T d_k \leqslant -c\|g_k\|^2, \quad \text{for all } k \geqslant 1 \text{ and some constant } c \in (0,1) \tag{1.6}$$

is often used in the literature to analyze the global convergence of conjugate gradient methods with inexact line search, and this condition may be crucial for conjugate gradient methods. It has been relaxed to the descent condition ($g_k^T d_k < 0$) in the convergence analysis of any conjugate gradient method. For example, the PRP is generally regarded as the most efficient conjugate gradient method in practical computation, but its global convergence has not been established with exact line search, weak (or strong) Wolfe–Powell line search (see [1,15]). Dai [1] gave an example which shows that even when the objective function is strongly convex and is sufficiently small, the PRP method may still fail by generating an ascent search direction. Gilbert and Nocedal [6] established the global convergence result of the PRP method, by restricting the scalar β_k to be nonnegative, i.e., $\beta_k^{PRP^+} = \max\{0, \beta_k^{PRP}\}$, and the sufficient descent condition (1.6) is used.

Motivated by β_k nonnegative and better descent property, we present a new modified nonlinear conjugate gradient method based on Wei–Yao–Liu [18] formula in the next section, and establish global convergence results for relatively algorithm in Section 3. The preliminary numerical results are contained in Section 4.

2. Modified formula and algorithm

Recently, Wei et al. [18] gave a variant of the PRP method which we call the WYL method, that is,

$$\beta_k^{WYL} = \frac{g_k^T\left(g_k - \frac{\|g_k\|}{\|g_{k-1}\|} g_{k-1}\right)}{\|g_{k-1}\|^2} \tag{2.1}$$

it is obvious to see that $0 \leqslant \beta_k^{WYL} \leq 2\beta_k^{FR}$, the WYL method and PRP method have similar feature, which performs essentially a restart if a bad direction occurs. The WYL method was extended to a variant of the LS method by Yao et al. [20], that is,

$$\beta_k^{MLS} = \frac{g_k^T\left(g_k - \frac{\|g_k\|}{\|g_{k-1}\|} g_{k-1}\right)}{-g_{k-1}^T d_{k-1}}. \tag{2.2}$$

If $\sigma \in (0, \frac{1}{4})$ in the SWP line search (1.5), Huang et al. [10] have proved that the WYL method satisfies the sufficient descent condition (1.6). If $\sigma \in (0, \frac{1}{2})$ in the SWP line search (1.5), Yao et al. [20] proved that the MLS method also can produce sufficient descent directions. Although the WYL and MLS methods possess nice property as the PRP$^+$ method [6] has, these methods may not be descent methods if the WWP line search (1.4) is used.

Yu et al. [21] propose a nonlinear conjugate gradient formula as follows

$$\beta_k^N(\mu) = \begin{cases} \frac{\|g_k\|^2 - |g_k^T g_{k-1}|}{\|g_{k-1}\|^2 + \mu |g_k^T d_{k-1}|} & \text{if } \|g_k\|^2 \geqslant |g_k^T g_{k-1}|, \\ 0 & \text{otherwise}, \end{cases} \tag{2.3}$$

where $\mu > 1$. They show that $0 \leqslant \beta_k^N(\mu) \leqslant \beta_k^{FR}$, and the method possesses attractive property that the sufficient descent condition (1.6) holds without any line search. In fact, the term $\mu|g_k^T d_{k-1}|$ in the denominator of (2.3) plays an important role in enhancing descent. Zhang [25] discussed improved scheme based on Wei–Yao–Liu conjugate gradient method, propose formula