Available online at www.sciencedirect.com
ScienceDirect
Applied Mathematics and Computation 189 (2007) 302–313
APPLIED MATHEMATICS AND COMPUTATION
www.elsevier.com/locate/amc

A new version of the Liu–Storey conjugate gradient method ☆

Chun-ming Tang [a,b,*], Zeng-xin Wei [b], Guo-yin Li [b]

[a] College of Science, Shanghai University, 200444, Shanghai, China
[b] College of Mathematics and Information Science, Guangxi University, 530004, Nanning, China

Abstract

In this paper, the global convergence of a new version of the Liu–Storey conjugate gradient method is discussed. This method combines the Liu–Storey conjugate gradient formula and a new inexact line search. We prove that the new method is globally convergent. Some preliminary numerical results show that the corresponding algorithm is efficient.

Keywords: Conjugate gradient; Liu–Storey; Inexact line search; Nonconvex optimization; Global convergence

1. Introduction

We consider the following unconstrained optimization problem:

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

where $f : \mathbb{R}^n \to \mathbb{R}$ is a smooth and nonlinear function.

During the last several decades, the research of the conjugate gradient methods was a very active topic due to the simplicity of their iteration and their excellent numerical performance in solving large scale optimization problems, see e.g. [15,7,11,3]. For the kth iterate x_k, the conjugate gradient methods generate a new point x_{k+1} by

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

where λ_k is a steplength determined by a certain line search, and d_k is a search direction defined by

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

☆ Project supported by Chinese NSF Grants 10161002, Guangxi NSF Grants 0542043 and Guangxi University research foundation X061057.
* Corresponding author. Address: College of Mathematics and Information Science, Guangxi University, 530004, Nanning, China.
E-mail addresses: cmtang@gxu.edu.cn (C.-m. Tang), zxwei@gxu.edu.cn (Z.-x. Wei), gyli@math.cuhk.edu.hk (G.-y. Li).

0096-3003/$ - see front matter
doi:10.1016/j.amc.2006.11.098

where β_k is a scalar, and $g_k = g(x_k)$ denotes the gradient of $f(x)$ at x_k. There are many formulas to generate β_k, such as the Fletcher–Reeves (FR) formula [4]

$$\beta_k^{\mathrm{FR}} = \frac{g_k^{\mathrm{T}} g_k}{g_{k-1}^{\mathrm{T}} g_{k-1}}, \tag{1.4}$$

the Polak–Ribiere–Polyak (PRP) formula [12,13]

$$\beta_k^{\mathrm{PRP}} = \frac{g_k^{\mathrm{T}}(g_k - g_{k-1})}{g_{k-1}^{\mathrm{T}} g_{k-1}}, \tag{1.5}$$

the Liu–Storey (LS) formula [7]

$$\beta_k^{\mathrm{LS}} = -\frac{g_k^{\mathrm{T}}(g_k - g_{k-1})}{d_{k-1}^{\mathrm{T}} g_{k-1}}, \tag{1.6}$$

and the Dai–Yuan (DY) formula [3]

$$\beta_k^{\mathrm{DY}} = \frac{g_k^{\mathrm{T}} g_k}{(g_k - g_{k-1})^{\mathrm{T}} d_{k-1}}. \tag{1.7}$$

The convergence behavior of various conjugate gradient formulas with different line search conditions [1,6,8,16–18] has been widely studied by many authors (see [7,10–16], and the references therein). Recently, Wei et al. [18] have proposed a modification of the Armijo steplength procedure

$$f(x_k + \rho^j d_k) - f(x_k) \leqslant \alpha\rho^j g_k^{\mathrm{T}} d_k - \frac{1}{2}(\rho^j)^2 d_k^{\mathrm{T}} B_k d_k \tag{1.8}$$

by using the function

$$f_k(x) = f(x) + \frac{1}{2}(x - x_k)^{\mathrm{T}} B_k (x - x_k)$$

to replace f in the following standard Armijo line search:

$$f(x_k + \rho^j d_k) - f(x_k) \leqslant \alpha\rho^j g_k^{\mathrm{T}} d_k,$$

where $\alpha \in (0,1)$, $\rho \in (0,1)$ and B_k is a simple and positive definite matrix.

Based on the line search (1.8), Wei et al. [16] proposed further a new Armijo-type line search technique (ATLS) for PRP method, and for simplicity, they set $B_k = mI$, where m is a fixed positive scalar for all k. More precisely, the Armijo-type line search for the PRP method given in [16] is as follows.

Find $\lambda_k = \rho^{j_k}$ such that j_k is the smallest nonnegative integer j satisfying

$$f(x_k + \rho^j d_k) - f(x_k) \leqslant \alpha\rho^j g_k^{\mathrm{T}} d_k - \frac{m}{2}(\rho^j)^2 d_k^{\mathrm{T}} d_k \tag{1.9}$$

and

$$g(x_k + \rho^j d_k)^{\mathrm{T}} Q_k^{\mathrm{PRP}}(j) \leqslant -c\|g(x_k + \rho^j d_k)\|^2, \tag{1.10}$$

where Q_k^{PRP} is defined by

$$Q_k^{\mathrm{PRP}}(j) = -g(x_k + \rho^j d_k) + \frac{g(x_k + \rho^j d_k)^{\mathrm{T}}(g(x_k + \rho^j d_k) - g_k)}{\|g_k\|^2} d_k. \tag{1.11}$$

Under weak conditions, they proved that the PRP method with ATLS (1.9)–(1.11) has the following sufficient descent property:

$$g_k^{\mathrm{T}} d_k \leqslant -c\|g_k\|^2 \tag{1.12}$$

for some constant $c \in (0,1)$, and thus the global convergence of the method can be proved.