NUMERICAL MATHEMATICS
A Journal of Chinese Universities (English Series)

Vol. **15** No. 4, pp. 357-366
November 2006

A Globally Convergent Polak-Ribière-Polyak Conjugate Gradient Method with Armijo-Type Line Search†

Gaohang Yu[1,*], Lutai Guan[1] and Zengxin Wei[2]

[1] *Department of Scientific Computation and Computer Applications, Sun Yat-Sen University, Guangzhou 510275, China.*
[2] *College of Mathematics and Information Science, Guangxi University, Nanning 530004, China.*

Received November 3, 2005; Accepted (in revised version) May 11, 2006

Abstract. In this paper, we propose a globally convergent Polak-Ribière-Polyak (PRP) conjugate gradient method for nonconvex minimization of differentiable functions by employing an Armijo-type line search which is simpler and less demanding than those defined in [4,10]. A favorite property of this method is that we can choose the initial stepsize as the one-dimensional minimizer of a quadratic model $\Phi(t) := f(x_k) + tg_k^T d_k + \frac{1}{2}t^2 d_k^T Q_k d_k$, where Q_k is a positive definite matrix that carries some second order information of the objective function f. So, this line search may make the stepsize t_k more easily accepted. Preliminary numerical results show that this method is efficient.

Key words: Unconstrained optimization; conjugate gradient method; nonconvex minimization; global convergence.

AMS subject classifications: 65K05, 90C30, 49M37

1 Introduction

The nonlinear conjugate gradient method is designed to solve the following unconstrained optimization problem

$$\min\{f(x) \mid x \in \mathbb{R}^n\}, \tag{1}$$

where $f : \mathbb{R}^n \to \mathbb{R}$ is a smooth nonlinear function whose gradient will be denoted by $g(x)$. We consider only the case where the method is implemented without regular restarts. The iterative formula of the conjugate gradient method is given by

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

*Correspondence to: Gaohang Yu, Department of Scientific Computation and Computer Applications, Sun Yat-Sen University, Guangzhou 510275, China. Email: maghyu@163.com
†This work is supported by the Chinese NSF grants 60475042, Guangxi NSF grants 0542043 and the Foundation of Advanced Research Center of Zhongshan University and Hong Kong.

where t_k is a stepsize which is computed by carrying out a line search, and d_k is the search direction defined by

$$d_k = \begin{cases} -g_k & \text{if} \quad k = 1, \\ -g_k + \beta_k d_{k-1} & \text{if} \quad k \geq 2, \end{cases} \tag{3}$$

where β_k is a scalar, and g_k denotes $g(x_k)$. Three well-known formulas for β_k are given as follows:

$$\beta_k^{HS} = \frac{g_k^T(g_k - g_{k-1})}{(g_k - g_{k-1})^T d_{k-1}} \quad \text{(Hestenses-Stiefel [13], 1952);} \tag{4}$$

$$\beta_k^{FR} = \frac{g_k^T g_k}{g_{k-1}^T g_{k-1}} \quad \text{(Fletcher-Reeves [8], 1964);} \tag{5}$$

$$\beta_k^{PRP} = \frac{g_k^T(g_k - g_{k-1})}{g_{k-1}^T g_{k-1}} \quad \text{(Polak-Ribière-Polyak [16,17], 1969).} \tag{6}$$

The convergence behavior of the above formulas with some line search conditions has been studied by many authors (e.g. [1-6,9-12,18-22]). In practical computations, it is generally believed that the conjugate gradient method is preferred to the relatively exact line searches. As a result, in the existing convergence analysis and the implementations of the conjugate gradient method, the strong Wolfe conditions are often imposed on the line search (see e.g. [1,6,14]). The FR method (5) with an exact line search was proved to be globally convergent on general functions by Zoutendijk [22]. Al-Baali [1] extended this result to the strong Wolfe line search. However, the PRP method (6) and the HS method (4) with the exact line search are not globally convergent, see Powell's counterexample [18]. Recent studies show that one can analyze the conjugate gradient method under several practical line searches other than the strong Wolfe line search (see e.g. [7,9,10,21]).

For example, the nonlinear conjugate gradient method in [7] converges globally provided the stepsize satisfies the weak Wolfe conditions, namely,

$$f(x_k + t_k d_k) - f(x_k) \leq \delta t_k g_k^T d_k, \tag{7}$$

and

$$g(x_k + t_k d_k)^T d_k \geq \sigma g_k^T d_k, \tag{8}$$

where $1 > \sigma > \delta > 0$. Under the weak Wolfe line search conditions and the sufficient descent condition $g_k^T d_k \leq -c\|g_k\|^2$, where $c \in (0,1)$, Gilbert and Nocedal [9] established the global convergence result of the PRP and HS methods by restricting the scalar β_k to be nonnegative.

Grippo and Lucidi [10] proposed new line search conditions (GLS) which were designed to match the requirements of the convergence theory to ensure that the PRP method is globally convergent for nonconvex problems. The line search in [10] is as follows: for any given $\mu > 0$ and $\rho \in (0,1)$, find

$$t_k = \max\left\{\rho^j \frac{\mu|g_k^T d_k|}{\|d_k\|^2}; \quad j = 0, 1, ...\right\} \tag{9}$$

such that t_k satisfies

$$f(x_{k+1}) \leq f(x_k) - \delta t_k^2 \|d_k\|^2 \tag{10}$$

and

$$-c_1\|g(x_k + t_k d_k)\|^2 \leq g(x_k + t_k d_k)^T d_{k+1} \leq -c_2\|g(x_k + t_k d_k)\|^2, \tag{11}$$

where $0 < c_2 < 1 < c_1$ are constants. Recently, Grippo and Lucidi [11] reconsidered the convergence properties of the PRP method and pointed out that $\lim_{k\to\infty} \|x_{k+1} - x_k\| = 0$ is a essentially sufficient condition for convergence that implies $\lim_{k\to\infty} \inf \|g_k\| = 0$. Further, they