ARTICLE IN PRESS

Applied Mathematical Modelling xxx (2014) xxx–xxx

Contents lists available at ScienceDirect

Applied Mathematical Modelling

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

A hybrid conjugate gradient method with descent property for unconstrained optimization ☆

Jinbao Jian [a,b,*], Lin Han [b], Xianzhen Jiang [a]

[a] College of Mathematics and Information Science, Yulin Normal University, Yulin, Guangxi 537000, China
[b] College of Mathematics and Information Science, Guangxi University, Nanning, Guangxi 530004, China

ARTICLE INFO

Article history:
Received 28 February 2013
Received in revised form 17 April 2014
Accepted 18 August 2014
Available online xxxx

Keywords:
Unconstrained optimization
Hybrid conjugate gradient method
Global convergence
Numerical experiments

ABSTRACT

In this paper, based on some famous previous conjugate gradient methods, a new hybrid conjugate gradient method was presented for unconstrained optimization. The proposed method can generate decent directions at every iteration, moreover, this property is independent of the steplength line search. Under the Wolfe line search, the proposed method possesses global convergence. Medium-scale numerical experiments and their performance profiles are reported, which show that the proposed method is promising.

1. Introduction

In this work, we considered the following unconstrained optimization problem:

$$\min\{f(x)|\ x \in R^n\} \tag{1}$$

where $f : R^n \to R$ is continuously differentiable and its gradient $g(x) \triangleq \nabla f(x)$ is available. Conjugate gradient methods (CGMs for short) are very efficient for solving large-scale unconstrained optimization problem (1), especially when the dimension n is large. Generally, the iterates of CGMs for solving problem (1) are obtained by

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

where α_k is a steplength computed by carrying out some suitable exact or inexact line search, and the search direction d_k is generated by

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

with a suitable scalar β_k and $g_k := \nabla f(x_k)$.

☆ Project supported by the Natural Science Foundation of Guangxi Province (Grant Nos. 2014GXNSFFA118001, 2013GXNSFAA019009) and Science Foundation of Guangxi Education Department (Grant Nos. ZD2014107, 2013YB196).

* Corresponding author at: College of Mathematics and Information Science, Yulin Normal University, Yulin, Guangxi 537000, China.
E-mail addresses: jianjb@gxu.edu.cn, jianjb@gmail.com (J. Jian).
URL: http://jians.gxu.edu.cn (J. Jian).

http://dx.doi.org/10.1016/j.apm.2014.08.008
0307-904X/

Please cite this article in press as: J. Jian et al., A hybrid conjugate gradient method with descent property for unconstrained optimization, Appl. Math. Modell. (2014), http://dx.doi.org/10.1016/j.apm.2014.08.008

In the type of CGMs with inexact line search, the steplength α_k is usually computed by the Wolfe line search

$$f(x_k + \alpha_k d_k) \leqslant f(x_k) + \delta\alpha_k g_k^T d_k, \quad g(x_k + \alpha_k d_k)^T d_k \geqslant \sigma g_k^T d_k, \tag{4}$$

or the strong Wolfe line search

$$f(x_k + \alpha_k d_k) \leqslant f(x_k) + \delta\alpha_k g_k^T d_k, \quad \left|g(x_k + \alpha_k d_k)^T d_k\right| \leqslant \sigma|g_k^T d_k|, \tag{5}$$

where parameters δ and σ satisfy $0 < \delta < \sigma < 1$.

In respect to different values of β_k, there are distinct CGMs. the well-known CGMs include the Fletcher–Reeves (FR) method [1], the Polak–Ribière–Polyak (PRP) [2,3], the Hestenes–Stiefel (HS) method [4] and the Dai–Yuan (DY) method [5]. The parameters β_k in these CGMs are specified as follows:

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

$$\beta_k^{HS} = \frac{g_k^T(g_k - g_{k-1})}{d_{k-1}^T(g_k - g_{k-1})}, \quad \beta_k^{DY} = \frac{\|g_k\|^2}{d_{k-1}^T(g_k - g_{k-1})}.$$

where $\|\cdot\|$ stands for the Euclidean norm. Note that the above four formulas for β_k are equivalent to each other if the objective function is a convex quadratic and the steplength α_k is yielded by exact line search. However, in the non-convex case, their theoretical properties and practical performance may be significantly different. The FR method and the DY method have strong convergence properties but usually have poor numerical performance for general objective functions. Under the strong Wolfe line search (5) with the parameter $\sigma < \frac{1}{2}$, Al-Baali [6] proved that the FR method satisfies the sufficient descent condition and converges globally for general objective functions. Dai and Yuan [5] shown that the DY method is descent and globally convergent if the Wolfe line search (4) is used. In contrary, the PRP method and the HS method are generally regarded to be two of the most efficient CGMs in practical computation, but their convergence properties are not so good. In the latest years, based on the above four formulas and their hybridization, many works putting effort into seeking for new CGMs with not only good convergence property but also excellent numerical effect were published, see Refs. [7–26].

In [10], Wei et al. gave a variant of the PRP method, WYL method for short, where the parameter β_k is yielded by

$$\beta_k^{WYL} = \frac{\|g_k\|^2 - \frac{\|g_k\|}{\|g_{k-1}\|} g_k^T g_{k-1}}{\|g_{k-1}\|^2}. \tag{6}$$

One important property of the WYL method is that it inherits the good properties of the PRP method, such as excellent numerical effect. Furthermore, Huang et al. [11] proved that the WYL method satisfies the sufficient descent condition and converges globally under the strong Wolfe line search (5) if the parameter satisfies $\sigma < \frac{1}{4}$.

Yao et al. [12] extended this idea to the HS method and proposed a conjugate gradient method, denote it by YWH method, that is

$$\beta_k^{YWH} = \frac{\|g_k\|^2 - \frac{\|g_k\|}{\|g_{k-1}\|} g_k^T g_{k-1}}{d_{k-1}^T(g_k - g_{k-1})}. \tag{7}$$

Under the strong Wolfe line search (5) with the parameter $\sigma < \frac{1}{3}$, it has been shown that the YWH method can generate sufficient descent directions and converges globally for general objective functions.

In [14], Wei et al. discussed the global convergence of the PRP CGM with inexact line search for nonconvex unconstrained optimization. Yu et al. [15] analyzed the global convergence of modified PRP CGM with sufficient descent property.

Recently, based on works [5,12,17], Jiang et al. [26] proposed a hybrid CGM with

$$\beta_k^{JHJ} = \frac{\|g_k\|^2 - \max\left\{0, \frac{\|g_k\|}{\|d_{k-1}\|} g_k^T d_{k-1}, \frac{\|g_k\|}{\|g_{k-1}\|} g_k^T g_{k-1}\right\}}{d_{k-1}^T(g_k - g_{k-1})}.$$

Under the Wolfe line search, the method possesses global convergence and efficient numerical performance.

Motivated by the ideas on the hybrid methods [22,23,26] and taking into account the good convergence properties of Refs. [1,5] and the good numerical performance [10,12], this paper introduce a new hybrid choice for parameter β_k as follows:

$$\beta_k^{N} = \frac{\|g_k\|^2 - \max\left\{0, \frac{\|g_k\|}{\|g_{k-1}\|} g_k^T g_{k-1}\right\}}{\max\left\{\|g_{k-1}\|^2, \ d_{k-1}^T(g_k - g_{k-1})\right\}}. \tag{8}$$

It is easy to know that $\beta_k^N = \beta_k^{DY}$ or β_k^{FR} or β_k^{WYL} or β_k^{YWH}, so β_k^N is one of hybrids of $\beta_k^{DY}, \beta_k^{FR}, \beta_k^{WYL}$ and β_k^{YWH}. Then, a new hybrid CGM is proposed. The proposed method has a good property that its search directions are always descent under any steplength line search technique, i.e., it is independent of the line search, and which allows us have a wide range of choice of the