Journal of Computational Mathematics
Vol.39, No.2, 2021, 159–177.
http://www.global-sci.org/jcm
doi:10.4208/jcm.1907-m2018-0173

A NEW ADAPTIVE SUBSPACE MINIMIZATION THREE-TERM CONJUGATE GRADIENT ALGORITHM FOR UNCONSTRAINED OPTIMIZATION*

Keke Zhang and Hongwei Liu
School of Mathematics and Statistics, Xidian University, Xi'an 710126, China
Email: Coco_xidian@163.com, hwliu@mail.xidian.edu.cn
Zexian Liu[1)]
State Key Laboratory of Scientific and Engineering Computing, Institute of Computational Mathematics and Scientific/Engineering computing, AMSS, Chinese Academy of Sciences, Beijing 100190, China
Email: liuzexian2008@163.com

Abstract

A new adaptive subspace minimization three-term conjugate gradient algorithm with nonmonotone line search is introduced and analyzed in this paper. The search directions are computed by minimizing a quadratic approximation of the objective function on special subspaces, and we also proposed an adaptive rule for choosing different searching directions at each iteration. We obtain a significant conclusion that the each choice of the search directions satisfies the sufficient descent condition. With the used nonmonotone line search, we prove that the new algorithm is globally convergent for general nonlinear functions under some mild assumptions. Numerical experiments show that the proposed algorithm is promising for the given test problem set.

Mathematics subject classification: 90C30.
Key words: Conjugate gradient method, Nonmonotone line search, Subspace minimization, Sufficient descent condition, Global convergence.

1. Introduction

Conjugate gradient (CG) methods have attracted special attention for solving nonlinear equations and nonlinear optimization problems, especially when the dimension of the problem is large. Due to the simplicity and limited memory requirement [1], CG methods are widely used in many optimization fields with the following form:

$$\min_{x\in R^n} f\left(x\right), \tag{1.1}$$

where the objective function $f: R^n \to R$ is continuously differentiable.

By starting from an initial point $x_0 \in R^n$, the CG methods generate a sequence $\{x_k\}$ as

$$x_{k+1} = x_k + \alpha_k d_k, \qquad k \geq 0, \tag{1.2}$$

where $\alpha_k > 0$ is the steplength and d_k is the search direction defined by

$$d_{k+1} = -g_{k+1} + \beta_k d_k, \;\; d_0 = -g_0, \tag{1.3}$$

* Received August 10, 2018 / Revised version received February 18, 2019 / Accepted July 24, 2019 / Published online December 18, 2019 /
1) Corresponding author

here $g_k = g(x_k) = \nabla f(x_k)$, and β_k is a scalar called CG parameter.

For general objective functions, different choices for the parameter β_k result in different CG methods [2], and their theoretical properties and numerical effects can be significantly different [3]. Some well-known choices for β_k are called HS, FR, PRP, DY formula [4–8], and are given by

$$\beta_k^{HS} = \frac{g_{k+1}^T y_k}{d_k^T y_k},\ \ \beta_k^{FR} = \frac{\|g_{k+1}\|^2}{\|g_k\|^2},\ \ \beta_k^{PRP} = \frac{g_{k+1}^T y_k}{\|g_k\|^2},\ \ \beta_k^{DY} = \frac{\|g_{k+1}\|^2}{d_k^T y_k},$$

where $y_k = g_{k+1} - g_k$, $s_k = x_{k+1} - x_k$ and $\|.\|$ denotes the Euclidean norm. More recent reviews on nonlinear CG methods can be found in [2, 9–12].

For nonlinear CG methods, the steplength α_k is usually determined by an inexact line search technique that is essential to the global convergence of CG methods. In this paper, we use the nonmonotone Wolfe line search proposed by Zhang and Hager (ZH) [13]:

$$f(x_k + \alpha_k d_k) \le C_k + \delta \alpha_k \nabla f(x_k)^T d_k, \tag{1.4}$$

$$\nabla f(x_k + \alpha_k d_k)^T d_k \ge \sigma \nabla f(x_k)^T d_k, \tag{1.5}$$

where $0 < \delta < \sigma < 1$ and C_k is a convex combination of the function values $f(x_0), f(x_1), \cdots, f(x_k)$.

In this paper, we are particularly interested in three-term CG methods by using the subspace minimization technique. To improve the efficiency of the classical CG methods, a type of three-term CG methods have been widely studied recently. The first general three-term CG method was proposed by Beale [14] as follows:

$$d_{k+1} = -g_{k+1} + \beta_k d_k + \gamma_k d_t, \tag{1.6}$$

where $\beta_k = \beta_k^{HS}$ (or β_k^{FR}, β_k^{DY} etc.), γ_k is a parameter, $1 \le t < k$.

In [15], Nazareth presented another three-term CG method where the direction is computed as

$$d_{k+1} = -y_k + \frac{y_k^T y_k}{y_k^T d_k} d_k + \frac{y_{k-1}^T y_k}{y_{k-1}^T d_{k-1}} d_{k-1}, \tag{1.7}$$

with $d_{-1} = 0,\ d_0 = 0$. If the objective function $f(x)$ is a convex quadratic function, then, for any stepsize α_k, the search directions generated by (1.7) are conjugate subject to the Hessian of $f(x)$ even without exact line search.

More researches on three-term CG method can be referred to [16–21]. As the scale of optimization problems is getting larger and larger, the subspace technique has become increasingly favored by many researches.

Yuan and Stoer [22] put forward a search direction by using a subspace minimization technique. The direction can be expressed as follows:

$$d_{k+1} = \mu_k g_{k+1} + \upsilon_k s_k, \tag{1.8}$$

where μ_k and υ_k are parameters. The corresponding method can be viewed as the subspace minimization CG method (SMCG). This method reduces to the classical CG method if the objective function is quadratic and the line search is exact.

In addition, based on the research of the above method, Andrei [23] proposed an accelerated subspace minimization three-term CG algorithm (TTS), where the search direction d_k is computed by

$$d_{k+1} = -g_{k+1} + a_k s_k + b_k y_k, \tag{1.9}$$