Applied Mathematics and Computation 240 (2014) 339–347

Contents lists available at ScienceDirect

Applied Mathematics and Computation

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

Spectral method and its application to the conjugate gradient method ☆

Dongyi Liu [a,b,*], Liping Zhang [a], Genqi Xu [a,b]

[a] Department of Mathematics, Tianjin University, Tianjin 300072, PR China
[b] Center for Applied Mathematics of Tianjin University, Tianjin 300072, PR China

ARTICLE INFO

Keywords:
Conjugate gradient method
Descent property
Spectral analysis
Global convergence

ABSTRACT

A new method used to prove the global convergence of the nonlinear conjugate gradient methods, the spectral method, is presented in this paper, and it is applied to a new conjugate gradient algorithm with sufficiently descent property. By analyzing the descent property, several concrete forms of this algorithm are suggested. Under standard Wolfe line searches, the global convergence of the new algorithm is proven for nonconvex functions. Preliminary numerical results for a set of 720 unconstrained optimization test problems verify the performance of the algorithm and show that the new algorithm is competitive with CG_DESCENT algorithm.

1. Introduction and main idea

The classical nonlinear conjugate gradient (NCG) method with line searches is as follows:

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

and

$$d_0 = -g_0, \quad d_{k+1} = -g_{k+1} + \beta_k d_k, \quad \forall k \geqslant 0 \tag{2}$$

where $g_k = g(x_k) = \nabla f(x_k)$, which is a well-known method for the large-scale unconstrained optimization problem

$$\min\{f(x) : x \in R^n\}.$$

To guarantee the global convergence properties of NCG methods, the fundamental assumptions about the objective function $f(x)$ are

H1. f is bounded below in R^n and continuously differentiable in a neighborhood $\mathcal{N}$ of the level set $\mathcal{L} = \{x : f(x) \leqslant f(x_0)\}$, where x_0 is the starting point of the iteration.
H2. The gradient of f is Lipschitz continuous in $\mathcal{N}$, that is, there exists a constant $L > 0$ such that

$$\|\nabla f(\bar{x}) - \nabla f(x)\| \leqslant L\|\bar{x} - x\|, \quad \forall \bar{x}, x \in \mathcal{N}.$$

☆ The work of the authors research is supported by the Natural Science Foundation of China Grant NSFC-61174080 and by the Seed Foundation of Tianjin University.

* Corresponding author at: Department of Mathematics, Tianjin University, Tianjin 300072, PR China.
E-mail addresses: dyliu@tju.edu.cn (D. Liu), tjdxzlp@163.com (L. Zhang), gqxu@tju.edu.cn (G. Xu).

http://dx.doi.org/10.1016/j.amc.2013.12.094
0096-3003/

And the standard Wolfe line search strategy [20,21] usually is needed as follows:

$$f(x_k + \alpha_k d_k) \leqslant f(x_k) + c_1 \alpha_k d_k^{\mathrm{T}} g_k \tag{3}$$

and

$$d_k^{\mathrm{T}} g(x_k + \alpha_k d_k) \geqslant c_2 d_k^{\mathrm{T}} g_k \tag{4}$$

where $0 < c_1 < c_2 < 1$. Of course other type line searches are also often used, such as the strong Wolfe line search, Goldstein type line-search and Armijo-type line search and so on [8,13,24].

In addition, the descent property [1]

$$d_k^{\mathrm{T}} g_k < 0, \tag{5}$$

or the sufficiently descent property [11]

$$d_k^{\mathrm{T}} g_k \leqslant -c_0 \|g_k\|^2 \text{ with } c_0 > 0 \tag{6}$$

is also a necessary condition for the global convergence. The Zoutendijk condition [25] is another important condition often used to prove the global convergence of NCG methods, such as [5,8]. Under the condition that the level set $\mathcal{L}$ is bounded, the Zoutendijk condition is utilized in [3,18,19,23,24].

Gilbert and Nocedal [11] introduced the so-called property $(*)$ and proved the convergence of the modified PRP method (PRP+). Later on, their results were generalized by Dai et al. [6]. The similar idea was also used in [7,9,12,22] to proved the global convergence under the assumption that the level set $\mathcal{L}$ is bounded. These algorithms mentioned above focused on the update for β_k in (2) to guarantee that the corresponding algorithms converge globally.

In this paper, we reformulate the line search directions of NCG methods as follows:

$$d_0 = -g_0, \quad d_k = -M_k g_k, \quad \forall k \geqslant 1 \tag{7}$$

where M_k is called the conjugate gradient iteration matrix. We develop new algorithms by selecting the suitable iteration matrix M_k and prove the global convergence by estimating the eigenvalues of $M_k^{\mathrm{T}} M_k$. So, this proof method is called the spectral method.

In what follows, we introduce in the spectral condition theorem for an objective function satisfying H1 and H2, which generalizes Theorem 4.1 in [15].

Theorem 1.1. *Assume the objective function $f(x)$ satisfy* H1 *and* H2. *For a NCG method determined by (1) and (7) which satisfies the sufficiently descent condition (6) and implements the standard Wolfe line searchs (3) and (4), if*

$$\sum_{k=1}^{\infty} (\Lambda_k)^{-1} = +\infty \tag{8}$$

where Λ_k is the maximum eigenvalue of $M_k^{\mathrm{T}} M_k$, then either $g_k = 0$ for some $k \geqslant 1$, or

$$\liminf_{k\to\infty} \|g_k\| = 0. \tag{9}$$

Moreover, if $\Lambda_k \leqslant \widetilde{\Lambda}$ where $\widetilde{\Lambda}$ is a positive constant, then

$$\lim_{k\to\infty} \|g_k\| = 0. \tag{10}$$

Proof. Assume that $g_k \neq 0,\ \forall k \geqslant 1$ and $\liminf_{k\to\infty} \|g_k\| \neq 0$, then there exists $\gamma > 0$ such that $\|g_k\| > \gamma,\ \forall k \geqslant 1$, and the sufficiently descent condition (6) implies that $d_k \neq 0,\ \forall k \geqslant 1$. It follows from (7) that

$$\|d_k\|^2 = g_k^{\mathrm{T}} M_k^{\mathrm{T}} M_k g_k \leqslant \Lambda_k \|g_k\|^2. \tag{11}$$

Thus, according to (6) and the above inequality, it can be deduced that

$$\cos^2 \theta_k = \frac{(-d_k^{\mathrm{T}} g_k)^2}{\|d_k\|^2 \|g_k\|^2} \geqslant c_0^2 \frac{\|g_k\|^2}{\|d_k\|^2} \geqslant \frac{c_0^2}{\Lambda_k}$$

where θ_k is the angle between d_k and $-g_k$. Thus,

$$\sum_{k\geqslant 1} \|g_k\|^2 \cos^2 \theta_k \geqslant \gamma \sum_{k=1}^{\infty} \frac{c_0^2}{\Lambda_k} = \infty,$$

which contradicts to the Zoutendijk's condition (see also Theorem 2.1 in [11]),

$$\sum_{k\geqslant 1} \|g_k\|^2 \cos^2 \theta_k = \sum_{k=1}^{\infty} \frac{(g_k^{\mathrm{T}} d_k)^2}{\|d_k\|^2} < \infty. \tag{12}$$