Applied Mathematics and Computation 259 (2015) 460–469

Contents lists available at ScienceDirect

Applied Mathematics and Computation

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

A new regularized quasi-Newton algorithm for unconstrained optimization ☆

Hao Zhang, Qin Ni *

College of Science, Nanjing University of Aeronautics and Astronautics, Nanjing 210031, China

ARTICLE INFO

Keywords:
Unconstrained optimization
Quasi-Newton method
Regularization

ABSTRACT

In this paper, we present a new regularized quasi-Newton algorithm for unconstrained optimization. In this algorithm, an adaptive quadratic term is employed to regularize the quasi-Newton model. At each iteration, the trial step is obtained by solving an unconstrained quadratic subproblem. The global convergence and superlinear convergence of the new algorithm are established under reasonable assumptions. The numerical results show that new algorithm is effective.

1. Introduction

In this paper, we consider the following unconstrained optimization problem:

$$\min_{x\in R^n} f(x), \tag{1}$$

For simplicity, we denote the gradient $\nabla f(x_k)$ by g_k, the Hessian $\nabla^2 f(x_k)$ by H_k. Throughout this paper, $\|.\|$ denotes the l_2-norm.

Trust-region [3–7] and line-search methods are two commonly-used convergence schemes for unstrained optimization. In 2009, Cartis et al. [1,2] presented an adaptive cubic regularization algorithm (denoted by ACR). In this algorithm, the trial step s_k is obtained by solving the following subproblem

$$\min_{s\in R^n} f(x_k) + g_k^T s + \frac{1}{2} s^T B_k s + \frac{1}{3}\mu_k \|s\|^3, \tag{2}$$

where B_k is the Hessian H_k or its approximation and μ_k is a nonnegative adaptive parameter. If the reduction of $f(x)$ is desirable, the value of μ_k is decreased; otherwise it is increased. Provided that the objective function $f(x)$ is continuously differentiable and bounded below, and B_k is bounded above for all k, the ACR iterates converge at least one limit that is first-order critical. Furthermore, when ∇f is uniformly continuous, the ACR algorithm is globally convergent to a first order critical point.

Although (2) is an unconstrained problem, it is expensive to solve it. According to Theorem 3.1 in [1], any s_k^* is a global minimizer of (2) over R^n if and only if it satisfies the system of equations

☆ This work was supported by the National Natural Science Foundation of China (11071117) and the Natural Science Foundation of Jiangsu Province (BK20141409).

* Corresponding author.
E-mail address: niqfs@nuaa.edu.cn (Q. Ni).

http://dx.doi.org/10.1016/j.amc.2015.02.032
0096-3003/

$$\begin{cases} (B_k + \mu_k^* I)s_k^* = -g_k, \\ \mu_k^* = \mu_k \|s_k^*\|, \end{cases} \tag{3}$$

where $B_k + \mu_k^* I$ is positive semidefinite. Because that μ_k^* is dependent on s_k^*, an iterative method is used to find s_k^* which satisfies (3). Additionally, $B_k + \mu_k^* I$ is ensured to be positive semidefinite at each iteration.

In [9], Ueda et al. presented a regularized Newton method (denoted by RNM) for solving (1). The cubic term in (2) is replaced by a quadratic one, and the subproblem is

$$\min_{s \in R^n} f(x_k) + g_k^T s + \frac{1}{2} s^T H_k s + \frac{1}{2}\sigma_k \|s\|^2, \tag{4}$$

where σ_k is chosen such that $H_k + \sigma_k I$ is positive definite. At each iteration, the rial step s_k is computed by $s_k = -(H_k + \sigma_k I)^{-1} g_k$ which is the global minimizer of (4). This method has global convergence and superlinear rate of convergence under appropriate conditions.

In this paper, we extend the above modification to quasi-Newton method. In order to obtain more general results, we consider the following quadratic approximated model of $f(x)$:

$$m_k(s) = f(x_k) + g_k^T s + \frac{1}{2} s^T B_k s + \frac{1}{2}\sigma_k \|Q_k s\|^2, \tag{5}$$

where B_k is a symmetric matrix which is the Hessian H_k or its approximation, Q_k is a nonsingular matrix,

$$\sigma_k = \frac{1}{\lambda_{min}(Q_k^T Q_k)} max\{0, -\lambda_{min}(B_k)\} + \mu_k, \tag{6}$$

μ_k is a dynamic positive parameter which is similar to the parameter in the ACR algorithm. At each iteration, the trial step s_k is obtained by solving

$$\min_{s \in R^n} m_k(s), \tag{7}$$

s_k^* is the global minimizer of (7) over R^n if and only if it satisfies the following linear system:

$$(B_k + \sigma_k Q_k^T Q_k)s_k^* = -g_k. \tag{8}$$

If $B_k = H_k$ and $Q_k = I$, then (5) is reduced to (4).

The rest of this paper is organized as follows. In Section 2, a description of the new algorithm is given and global convergence is investigated. In Section 3, we establish the superlinear convergence under reasonable assumptions. The numerical results of new algorithm are reported in Section 4.

2. Algorithm and convergence analysis

At first, we give a description of new quasi-Newton algorithm.

Algorithm 2.1. Regularized quasi-Newton algorithm

Step 0: Given $1 > \eta_2 > \eta_1 > 0$, $\gamma_3 > \gamma_2 \geqslant 1 > \gamma_1 > 0$, $\mu_0 > 0$, $x_0 \in R^n$, $B_0 \in R^{n \times n}$ is positive definite, $Q_0 \in R^{n \times n}$ is nonsingular; set $k = 0$ and

$$\sigma_0 = \frac{1}{\lambda_{min}(Q_0^T Q_0)} max\{0, -\lambda_{min}(B_0)\} + \mu_0.$$

Compute $f(x_0)$, $g_0 = \nabla f(x_0)$.

Step 1: If $\|g_k\| = 0$, stop with $x^* = x_k$.
Step 2: Determine s_k by solving (7), compute $f(x_k + s_k)$ and

$$\rho_k = \frac{f(x_k) - f(x_k + s_k)}{f(x_k) - m_k(s_k)}. \tag{9}$$

Step 3: Set

$$x_{k+1} = \begin{cases} x_k + s_k, & \text{if } \rho_k \geqslant \eta_1, \\ x_k, & \text{if } \rho_k < \eta_1. \end{cases} \tag{10}$$

If $\rho_k \geqslant \eta_1$, then set $f(x_{k+1}) = f(x_k + s_k)$, otherwise set $f(x_{k+1}) = f(x_k)$.