Journal of Computational Mathematics, Vol.22, No.1, 2004, 11–20.

A CLASS OF REVISED BROYDEN ALGORITHMS WITHOUT EXACT LINE SEARCH *1)

Ding-guo Pu Sheng-hua Gui Wei-wen Tian
(*Department of Mathematics, Shanghai Tongji University, Shanghai 200331, China*)

Abstract

In this paper, we discuss the convergence of the Broyden algorithms with revised search direction. Under some inexact line searches, we prove that the algorithms are globally convergent for continuously differentiable functions and the rate of local convergence of the algorithms is one-step superlinear and n-step second-order for uniformly convex objective functions.

Key words: Variable metric algorithms, Line search, Convergence, Convergence rate.

1. Introduction

The Broyden family of algorithms remains a standard workhorse for minimization. These methods share the properties of finite termination on strictly convex quadratic functions, a superlinear rate of convergence on general strictly convex functions, and no need to store or evaluate the second derivative matrix. (see [2, 4, 1, 5, 6, 7]). However, there are several unsolved problems for the Broyden algorithms. In this paper we propose a new class of variable metric algorithms with revised search directions. We prove that the algorithms are convergent for the continuously differentiable objective functions. Also the new algorithms are superlinear and n-step second order convergent for uniformly convex functions when the line searches are inexact, but satisfy some search conditions.

These algorithms are iterative. Given a starting point x_1 and an initial positive definite matrix B_1, they generate a sequence of points $\{x_k\}$ and a sequence of matrices of $\{B_k\}$ which are given by following (1) and (2)

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

where $\alpha_k > 0$ is the step factor, d_k is the search direction satisfying

$$-d_k = H_k g_k + \|Q_k H_k g_k\| R_k g_k,$$

where g_k is the gradient of $f(x)$ at x_k, H_k is the inverse of B_k, $\{Q_k\}$ and $\{R_k\}$ are two sequences of positive definite matrices which are uniformly bounded. All eigenvalues of these matrices are included in $[q, r]$, $0 < q \le r$, *i.e.*, for all k and $x \in R^n, x \neq 0$

$$q\|x\|^2 \le x^T Q_k x \le r\|x\|^2; \quad q\|x\|^2 \le x^T R_k x \le r\|x\|^2.$$

If $g_k = 0$, the algorithms terminate, otherwise let

$$B_{k+1} = B_k - \frac{B_k s_k s_k^T B_k}{s_k^T B_k s_k} + \frac{y_k y_k^T}{s_k^T y_k} + \phi(s_k^T B_k s_k) v_k v_k^T \tag{2}$$

* Received July 1, 2001; final revised March 16, 2003.
1) This research is supported by National Natural Science Foundation of China, Ministry of Education P. R. C and the Research and development foundation of Shanghai Education Commission.

where $y_k = g_{k+1} - g_k, v_k = y_k(s_k^T y_k)^{-1} - B_k s_k(s_k^T B_k s_k)^{-1}$ and $\phi \in [0,1]$. In the above algorithms, if $\phi = 0$ we call it revised BFGS algorithm, or RBFGS algorithm and if $\phi = 1$ we call it revised DFP algorithm, or RDFP algorithm.

The matrix H_{k+1} denotes the inverse of B_{k+1}, the recurrence formula of H_{k+1} is

$$H_{k+1} = H_k - \frac{H_k y_k y_k^T H_k}{y_k^T H_k y_k} + \frac{s_k s_k^T}{s_k^T y_k} + \frac{\rho \mu_k \mu_k^T}{y_k^T H_k y_k}, \tag{3}$$

where

$$\mu_k = H_k y_k - \frac{y_k^T H_k y_k}{s_k^T y_k} s_k \tag{4}$$

and $\rho \in [0,1]$, the relationship of ρ and ϕ is

$$\phi = \frac{(1-\rho)(s_k^T y_k)^2}{(1-\rho)(s_k^T y_k)^2 + \rho y_k^T H_k y_k s_k^T B_k s_k}.$$

In this paper, the line searches are not required to be exact. In order to guarantee descentness of the objective function values and the convergence of the algorithms, we must give some conditions for determining α_k. We use Wolfe conditions on line searches,

$$f(x_k) - f(x_{k+1}) \geq \zeta_0(-g_k^T s_k) \tag{5}$$

and

$$|g_{k+1}^T s_k| \leq \theta_0(-g_k^T s_k), \tag{6}$$

where ζ_0 and θ_0 be two constants satisfying $0 < \zeta_0 \leq \theta_0 < 1/2$. We always try $\alpha_k = 1$ first in choosing the step length.

Using the mathematical induction it is easy to imply that B_k and H_k are positive definite matrices if H_1 and B_1 are positive definite matrices.

If no ambiguities are arisen we may drop the subscript of the characters, for example, g, x, R denote g_k, x_k, R_k, and use subscript $*$ to denote the amounts obtained by the next iteration, *i.e.*, g_*, x_*, R_* denote g_{k+1}, x_{k+1}, R_{k+1}, respectively.

For simplicity we let

$$\begin{aligned} U_k &= \frac{-g_k^T H_k y_k}{y_k^T H_k y_k}; \ V_k = \frac{y_k^T H_k y_k}{s_k^T y_k}; \qquad W_k = \frac{-g_k^T d_k}{y_k^T d_k} = \frac{-g_k^T s_k}{s_k^T y_k}; \\ Z_k &= H_k g_k + \frac{-g_k^T H_k y_k}{s_k^T y_k} s_k \\ &= \frac{\|Q_k H_k g_k\| y_k^T R_k g_k}{s_k^T y_k} s_k - \|Q_k H_k g_k\| R_k g_k. \end{aligned} \tag{7}$$

The paper is outlined as follows: Section 2 gives several convergence results without the convexity assumption. Section 3 gives some results for convex objective functions. In Sections 4, we prove that the algorithms are linearly convergent for $\phi \in [0,1)$ in detail. In Section 5, we prove that our algorithms are one-step superlinearly convergent, then give the quadratical convergence of the algorithms without detail proof.

Throughout this paper the vector norms are Euclidian.

2. Results Without Convexity Assumption

In this section, we assume: