Journal of Computational Mathematics, Vol.23, No.4, 2005, 357–372.

CONVERGENCE PROPERTIES OF MULTI-DIRECTIONAL PARALLEL ALGORITHMS FOR UNCONSTRAINED MINIMIZATION *1)

Cheng-xian Xu   Yue-ting Yang
(*Faculty of Sciences, Xi'an Jiaotong University, Xi'an 710049, China*)

Abstract

Convergence properties of a class of multi-directional parallel quasi-Newton algorithms for the solution of unconstrained minimization problems are studied in this paper. At each iteration these algorithms generate several different quasi-Newton directions, and then apply line searches to determine step lengths along each direction, simultaneously. The next iterate is obtained among these trail points by choosing the lowest point in the sense of function reductions. Different quasi-Newton updating formulas from the Broyden family are used to generate a main sequence of Hessian matrix approximations. Based on the BFGS and the modified BFGS updating formulas, the global and superlinear convergence results are proved. It is observed that all the quasi-Newton directions asymptotically approach the Newton direction in both direction and length when the iterate sequence converges to a local minimum of the objective function, and hence the result of superlinear convergence follows.

Mathematics subject classification: 65K05.
Key words: Unconstrained minimization, Multi-directional parallel quasi-Newton method, Global convergece, Superlinear convergence.

1. Introduction

This paper concerns with quasi-Newton methods for unconstrained nonlinear minimization

$$\min \quad f(x), \tag{1.1}$$

where $f:\ R^n \rightarrow R$ is assumed to be twice continuously differentiable. Starting from an initial point x_1 and an initial symmetric positive definite matrix B_1, a quasi-Newton method generates sequences $\{x_k\}$ and $\{B_k\}$ by the iteration

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

and an updating formula for B_k, where α_k is a step length and d_k is a descent search direction that is generated by solving the following system of equations

$$B_k d_k = -g_k,$$

$g_k = \nabla f(x_k)$ is the gradient of $f(x)$ at x_k. B_k is an $n \times n$ symmetric matrix that approximates the Hessian $G(x) = \nabla^2 f(x)$ of $f(x)$ at x_k, and satisfies the so-called quasi-Newton equation

$$B_k s_{k-1} = y_{k-1} \tag{1.3}$$

with $s_{k-1} = x_k - x_{k-1}$ and $y_{k-1} = g_k - g_{k-1}$.

* Received November 12, 2003; final revised January 11, 2005.
1) This work is supported by National Science Foundation of China: 10231060.

Various updating formulae that satisfy equation (1.3) exist, and one of the most widely used class of updates was the Broyden family (see [3])

$$B_{k+1}(\phi) = B_k - \frac{B_k s_k s_k^T B_k}{s_k^T B_k s_k} + \frac{y_k y_k^T}{s_k^T y_k} + \phi(s_k^T B_k s_k) u_k u_k^T, \tag{1.4}$$

where ϕ is a scale parameter and

$$u_k = \frac{y_k}{s_k^T y_k} - \frac{B_k s_k}{s_k^T B_k s_k}.$$

The computational characters and convergence properties of quasi-Newton methods in Broyden family have been widely studied (see [6], [7], [9], [10], [11], [12], [13], [14], [16], [24], [25]).

One of the most widely used quasi-Newton update is the BFGS update

$$B_{k+1}^{BFGS} = B_k - \frac{B_k s_k s_k^T B_k}{s_k^T B_k s_k} + \frac{y_k y_k^T}{s_k^T y_k}, \tag{1.5}$$

that is obtained by setting $\phi = 0$ in (1.4), denoted by $B_{k+1}(0)$. Broyden, Dennis and More in [5] proved that the BFGS method with unit step length for all k is superlinearly convergent provided that the initial point x_1 and the initial Hessian approximation B_1 are sufficiently accurate. Powell in [24] proved the global convergence for the BFGS method when it is applied to convex functions and the step length α_k satisfies the Wolfe conditions for all k. Furthermore, if the function $f(x)$ is strictly convex and the step length $\alpha_k = 1$ is taken whenever it satisfies the Wolfe conditions, the result of Broyden, Dennis and More in [5] applies, i.e., the convergence rate is superlinear. These convergence properties of the BFGS method have been extended to the convex Broyden class, except for the DFP method, $(0 \le \phi < 1)$ by Ritter [26], Byrd, Nocedal and Yuan [7], and to the preconvex Broyden class $(\phi_{k0} < \phi < 1)$ by Byrd, Liu and Nocedal [8] where

$$\phi_{k0} = (s_k^T y_k)^2 / [(s_k^T y_k)^2 - s_k^T B_k s_k y_k^T B_k^{-1} y_k] < 0. \tag{1.6}$$

The easiest update in (1.4) is the symmetric rank one (**SR1**) update

$$B_{k+1}^{SR1} = B_k + \frac{(y_k - B_k s_k)(y_k - B_k s_k)^T}{(y_k - B_k s_k)^T s_k},$$

that is obtained by setting

$$\phi = s_k^T y_k / (y_k - B_k s_k)^T s_k \stackrel{\triangle}{=} \hat{\phi}_k \tag{1.7}$$

in (1.4), denoted by $B_{k+1}(\hat{\phi}_k)$. The drawback of the SR1 update is that the matrix B_{k+1}^{SR1} may not be positive definite or it may even not well defined when the denominator approaches zero. However, some recent works on the SR1 method have sparked renewed interesting in this updating formula (see [9], [21], [18] and [1]). It is proved in [9] that the sequence $\{B_k\}$ generated by the SR1 update converges to the actual Hessian $G(x^*)$ at the solution x^*, provided that the search directions $\{d_k\}$ are uniformly linearly independent, and that the denominators in the SR1 update are always sufficiently different from zero, and that the iterates $\{x_k\}$ converges to x^*. Moreover, numerical tests (see [9] and [29]) show that in comparison with the BFGS update, the SR1 update generates more accurate Hessian approximations. Khalfan, Byrd and Schnabel in [18] provided a proof of $(n + 1)$-step super-linear convergence result for the SR1 method under an assumption that the updating matrices $\{B_k\}$ are positive definite for all k and bounded asymptotically.

Based on the idea of obtaining more accurate Hessian approximation in the direction s_{k-1} through using more available function value information in updating formulae, Zhang and Xu in [32] proposed a modification to quasi-Newton equation (1.3)

$$B_k s_{k-1} = (1 + \frac{\theta_{k-1}}{s_{k-1}^T y_{k-1}}) y_{k-1} \stackrel{def}{=} \hat{y}_{k-1},$$