Journal of Computational Mathematics, Vol.13, No.2, 1995, 95-107.

NON-QUASI-NEWTON UPDATES FOR UNCONSTRAINED OPTIMIZATION*1)

Ya-xiang Yuan
(*Computing Centre, Academia Sinica, Beijing, China*)

Richard H. Byrd
(*Department of Computer Science, University of Colorado at Boulder, USA*)

Abstract

In this report we present some new numerical methods for unconstrained optimization. These methods apply update formulae that do not satisfy the quasi-Newton equation. We derive these new formulae by considering different techniques of approximating the objective function. Theoretical analyses are given to show the advantages of using non-quasi-Newton updates. Under mild conditions we prove that our new update formulae preserve global convergence properties. Numerical results are also presented.

1. Introduction

Unconstrained optimization is to minimize a nonlinear function $f(x)$ in a finite dimensional space, that is

$$\min_{x\in R^n} \quad f(x) \quad . \tag{1.1}$$

Newton's method for problem (1.1) is iterative and at the $k-$th iteration a current approximation solution x_k is available. The Newton step at the $k-$th iteration is

$$d_k = -(\nabla^2 f(x_k))^{-1}\nabla f(x_k) \quad . \tag{1.2}$$

One advantage of Newton's method is that it convergence quadratically. Assume x^* is a stationary point of (1.1) at which $\nabla^2 f(x^*)$ is non-singular. Then for x_k sufficiently close to x^* we have that

$$||x_k + d_k - x^*|| = O(||x_k - x^*||^2) \quad . \tag{1.3}$$

However Newton's method also has some disadvantages. Firstly the Hessian $\nabla^2 f(x_k)$ may be singular, in that case the Newton step (1.2) is not well defined. Secondly when $\nabla^2 f(x_k)$ is not positive definite the Newton step d_k may not necessarily be a descent

* Received March 28, 1994.

1) This work was supported by the State Key Project "large scale scientific computing" of China.

direction of the objective function. Thirdly the calculation of the Hessian $\nabla^2 f(x_k)$ may be very expensive especially for large scale problems, not to mention that for some problems the Hessian of $f(x)$ is not available.

Quasi-Newton methods are a class of numerical methods that are similar to Newton's method except that the Hessian $(\nabla^2 f(x_k))^{-1}$ is replaced by an $n \times n$ symmetric matrix H_k which satisfies the "quasi-Newton" equation

$$H_k y_{k-1} = s_{k-1} \tag{1.4}$$

where

$$s_{k-1} = x_k - x_{k-1} = \alpha_{k-1} d_{k-1} \tag{1.5}$$

$$y_{k-1} = \nabla f(x_k) - \nabla f(x_{k-1}) \ , \tag{1.6}$$

and $\alpha_{k-1} > 0$ is a step-length which satisfies some line search conditions. Assume H_k is nonsingular, we define $B_k = (H_k)^{-1}$. It is easy to see that the "quasi-Newton step"

$$d_k = -H_k \nabla f(x_k) \tag{1.7}$$

is a stationary point of the following problem:

$$\min_{d \in R^n} \quad \phi_k(d) = f(x_k) + d^T \nabla f(x_k) + \frac{1}{2} d^T B_k d \tag{1.8}$$

which is an approximation to problem (1.1) near the current iterate x_k, since $\phi_k(d) \simeq f(x_k + d)$ for small d. In fact, the definition of $\phi_k(.)$ in (1.8) implies that

$$\phi_k(0) = f(x_k), \tag{1.9}$$

$$\nabla \phi_k(0) = \nabla f(x_k), \tag{1.10}$$

and the quasi-Newton condition (1.4) is equivalent to

$$\nabla \phi_k(x_{k-1} - x_k) = \nabla f(x_{k-1}) \ . \tag{1.11}$$

Thus, $\phi_k(x - x_k)$ is a quadratic interpolation of $f(x)$ at x_k and x_{k-1}, satisfying conditions (1.9)-(1.11). The matrix B_k (or H_k) can be updated so that the quasi-Newton equation is satisfied. One well known update formula is the BFGS formula which updates B_{k+1} from B_k, s_k and y_k in the following way:

$$B_{k+1} = B_k - \frac{B_k s_k s_k^T B_k}{s_k^T B_k s_k} + \frac{y_k y_k^T}{s_k^T y_k} \quad . \tag{1.12}$$

In Yuan (1991), approximate function $\phi_k(d)$ in (1.8) is required to satisfy the interpolation condition

$$\phi_k(x_{k-1} - x_k) = f(x_{k-1}) \ , \tag{1.13}$$