Journal of Computational Mathematics, Vol.12, No.4, 1994, 380–386.

A QUASI-NEWTON ALGORITHM WITHOUT CALCULATING DERIVATIVES FOR UNCONSTRAINED OPTIMIZATION*[1)]

Sun Lin-ping
(*Department of Mathematics, Nanjing University, Jiangsu, China*)

Abstract

A new algorithm for unconstrained optimization is developed, by using the product form of the OCSSR1 update. The implementation is especially useful when gradient information is estimated by difference formulae. Preliminary tests show that new algorithm can perform well.

1. Introduction

We consider the unconstrained optimization problem

$$Min\ f(x) \tag{1.1}$$

where $f(x): \mathbb{R}^n \to \mathbb{R}$ is a real continuously differentiable function.

Many algorithms have been proposed for solving (1.1). Typically, given both an approximation H to $[\nabla^2 f(x)]^{-1}$ and g the gradient $\nabla f(x)$ at the current point x, a quasi-Newton algorithm starts each iteration by taking a step

$$x_+ = x - \alpha Hg, \tag{1.2}$$

where the steplenth $\alpha > 0$ is chosen so that

$$f(x) - f(x_+) \geq \sigma\alpha g^T Hg \tag{1.3a}$$

and

$$|\ g_+^T Hg\ | \leq \tau g^T Hg \tag{1.3b}$$

are satisfied, where $\sigma \in (0, 1/2)$ and $\tau \in (\sigma, 1)$; and then to form H_+, an estimate of $[\nabla^2 f(x)]^{-1}$ by using an updating formula satisfying the quasi-Newton condition

$$H_+ y = s, \tag{1.4}$$

where $s = x_+ - x$ and $y = g_+ - g$. For the SSR1 update

$$H_+ = \theta H + \frac{(s - \theta Hy)(s - \theta Hy)^T}{(s - \theta Hy)^T y}, \tag{1.5}$$

* Received August 18, 1993.
[1)] The Project Supported by National Natural Science Foundation of China.

Osborne and Sun[9] propose a new algorithm (OCSSR1) with Davidon's optimal condition, in which the scaling factor θ is taken either

$$\theta_1 = c/b - \sqrt{c^2/b^2 - c/a} \tag{1.6a}$$

or

$$\theta_1 = c/b + \sqrt{c^2/b^2 - c/a}, \tag{1.6b}$$

where $a = y^T Hy, b = s^T y$ and $c = s^T H^{-1} s$. Preliminary numerical tests show that the OCSSR1 method compares favourably with good implementations of the BFGS method[5,12,14].

Here, a new implementation of the OCSSR1 algorithm is developed by writing the expression for the SSR1 update in product form. If the derivatives are available, the implementation is equivalent to the OCSSR1 algorithm. Moreover, it is especially useful when gradient information is estimated by finite difference formulae, in this case, the algorithm can perform well. In Section 2, an expression of the SSR1 update in product form is derived. In Section 3, an algorithm using the OCSSR1 update without calculating derivatives is outlined. Numerical results are contained in Section 4.

In this paper, the following notations are used: I is an unit matrix; $\kappa(.)$ denotes the condition number of a matrix; Tr(.) is the trace of a matrix; det(.) denotes the determinant of a square matrix.

2. The SSR1 Formula in Product Form

The SSR1 update (1.5) can be written in product form. That is

$$H_+ = (I + uv^T)\theta H(I + uv^T)^T, \tag{2.1}$$

where

$$u = \mu(s - \theta Hy), \tag{2.2}$$

$$v = (H^{-1}s - \theta y)/\theta \tag{2.3}$$

and

$$\mu = \frac{-\theta \pm \sqrt{(c\theta - b\theta^2)/(b - a\theta)}}{c - 2b\theta + a\theta^2}. \tag{2.4}$$

Let $H_+ = C_+ C_+^T$ and $H = CC^T$, from (2.1) we have

$$C_+ = \sqrt{\theta}[C + u(C^T v)^T]. \tag{2.5}$$

Remark 2.1. Osborne and Sun[9] point out that if $\theta > 0$ and $s^T y > 0$, then H_+ is positive definite if and only if

$$\theta \notin [s^T y/y^T Hy, s^T H^{-1} s/s^T y]. \tag{2.6}$$

Thus, in (2.4), it always holds that $(c\theta - b\theta^2)/(b - a\theta) > 0$. In addition, $c - 2b\theta + a\theta^2 > 0$ because $ac \geq b^2$.