Journal of Computational Mathematics, Vol.17, No.6, 1999, 645–652.

CONVERGENCE PROPERTIES OF A MODIFIED BFGS ALGORITHM FOR MINIMIZATION WITH ARMIJO-GOLDSTEIN STEPLENGTHS*1)

Nai-yang Deng   Zheng-feng Li
(*Department of Mathematics, Beijing Agricultural Engineering University, Beijing 100083, China*)

Abstract

The line search strategy is crucial for an efficient unconstrained optimization algorithm. One of the reason why the Wolfe line searches is recommended lies in that it ensures positive definiteness of BFGS updates. When gradient information has to be obtained costly, the Armijo-Goldstein line searches may be preferred. To maintain positive difiniteness of BFGS updates based on the Armijo-Goldstein line searches, a slightly modified form of BFGS update is proposed by I.D. Coope and C.J. Price (Journal of Computational Mathematics, 13 (1995), 156–160), while its convergence properties is open up to now. This paper shows that the modified BFGS algorithm is globally and superlinearly convergent based on the Armijo-Goldstein line searches.

Key words: BFGS methods, Convergence, Superlinear convergence.

1. Introduction

Assume that we are finding the minimizer of the following unconstrained optimization problem

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

and assume the current point is x_k. To calculate x_{k+1} from x_k by a line search method, the following iteration

$$x_{k+1} = x_k + \lambda_k p_k, \quad k = 1, 2, \cdots \tag{1.2}$$

is applied. In the BFGS algorithm the search direction p_k is chosen so that $B_k p_k = g_k$, where $g_k = \nabla f(x_k)$, the matrices B_k are defined by the update formula recurrently

$$B_{k+1} = B_k - \frac{B_k s_k s_k^T B_k}{s_k^T B_k s_k} + \frac{y_k y_k^T}{s_k^T y_k} \tag{1.3}$$

* Received December 5, 1995.
1) Work supported by the National Natural Science Foundation of China and the Natural Science Foundation of Beijing.

$$s_k = x_{k+1} - x_k \tag{1.4}$$

$$y_k = g_{k+1} - g_k \tag{1.5}$$

It is well known that if B_1 is positive definite and

$$s_k^T y_k > 0 \tag{1.6}$$

then all matrices B_k, $k = 1, 2, \cdots$, generated by (1.3) are positive definite. One of the line search strategies is the Wolfe line searches which require the steplength $\lambda_k > 0$ to satisfy the inequalities

$$f(x_k + \lambda_k p_k) \leq f(x_k) + \alpha \lambda_k g_k^T p_k \tag{1.7}$$

$$g(x_k + \lambda_k p_k)^T p_k \geq \beta g_k^T p_k \tag{1.8}$$

where α and β are constants that satisfy $0 < \alpha < \beta < 1$ and $\alpha < 1/2$. It is easy to show that condition (1.8) implies that

$$s_k^T y_k \geq (\beta - 1) s_k^T g_k > 0 \tag{1.9}$$

so that the BFGS updating formula can be applied with positive definiteness being maintained automatically. A disadvantage is that to test condition (1.8) requires an extragradient evaluation at each trial value for λ_k. When gradient information has to be obtained costly, the Armijo-Goldstein line searches

$$\alpha_2 \lambda_k p_k^T g_k \leq f(x_{k+1}) - f(x_k) \leq \alpha_1 \lambda_k p_k^T g_k \tag{1.10}$$

may be preferred, where $0 < \alpha_1 < 1/2 < \alpha_2 < 1$. However, condition (1.10) does not ensure that $s_k^T y_k > 0$. To maintain positive definiteness of BFGS updates based on the Armijo-Goldstein line searches, a slightly modified form of BFGS update is proposed by I.D. Coope and C.J. Price in [2]. They require the quadratic model, $q_k(x)$, interpolating the data $q_k(x_k) = f(x_k)$, $q_k(x_{k+1}) = f(x_{k+1})$, and $\nabla q_k(x_k) = g_k$. Let

$$z_k = y_k + \frac{2(f(x_{k+1}) - f(x_k) - s_k^T g_k) - s_k^T y_k}{s_k^T s_k} s_k \tag{1.11}$$

Applying the standard BFGS update (1.3), they derive their modified BFGS update with z_k replacing y_k

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

Notice the condition (1.10), we have

$$s_k^T z_k = s_k^T y_k + 2(f(x_{k+1}) - f(x_k)) - s_k^T g_k) - s_k^T y_k = 2(f(x_{k+1}) - f(x_k)) - s_k^T g_k)$$