Journal of Computational Mathematics, Vol.14, No.2, 1996, 143-158.

A CLASS OF FACTORIZED QUASI-NEWTON METHODS FOR NONLINEAR LEAST SQUARES PROBLEMS*

C.X. Xu X.F. Ma M.Y. Kong
(*Department of Mathematics, Xi'an Jiaotong University, Xi'an, China*)

Abstract

This paper gives a class of descent methods for nonlinear least squares solution. A class of updating formulae is obtained by using generalized inverse matrices. These formulae generate an approximation to the second part of the Hessian matrix of the objective function, and are updated in such a way that the resulting approximation to the whole Hessian matrix is the convex class of Broyden-like updating formulae. It is proved that the proposed updating formulae are invariant under linear transformation and that the class of factorized quasi-Newton methods are locally and superlinearly convergent. Numerical results are presented and show that the proposed methods are promising.

1. Introduction

This paper deals with the problem of minimizing a sum of squares of nonlinear functions

$$f(x) = \frac{1}{2}\sum_{i=1}^{m}(r_i(x))^2 = \frac{1}{2}r(x)^T r(x) \tag{1}$$

where $r_i(x), i = 1, 2, \cdots, m$ are twice continuously differentiable, $m \geq n$, $r(x) = (r_1(x), r_2(x), \cdots, r_m(x))^T$ and "T" denotes transpose. Nonlinear least squares problem is a kind of important optimization problems and is appeared in many fields such as scientific experiments, maximum likelihood estimation, solution of nonlinear equations, pattern recognition and etc. The derivatives of the function $f(x)$ are given by

$$g(x) = \bigtriangledown f(x) = A(x)^T r(x) \tag{2}$$

$$G(x) = \bigtriangledown^2 f(x) = A(x)^T A(x) + \sum_{i=1}^{m} r_i(x) \bigtriangledown^2 r_i(x) \tag{3}$$

where $A \in R^{m\times n}$ is the Jacobian matrix of $r(x)$ and its elements are $a_{ij} = \partial r_i(x)/\partial x_j$, $i = 1, 2, \cdots, m, j = 1, 2, \cdots, n$.

Various iterative methods for problem (1) are available and can be divided into two kinds, trust region methods and descent methods. Trust region methods are globally

* Received August 23, 1994.

convergent, but complicated in implementation. In this paper we consider Newton-like descent methods. Suppose that $x^{(k)}$ is a current estimation of the minimum point x^*. A descent direction $d^{(k)}$ is assigned to $x^{(k)}$ by solving a system

$$B_k d^{(k)} = -g^{(k)} \tag{4}$$

and a new estimate point is generated by

$$x^{(k+1)} = x^{(k)} + \alpha_k d^{(k)} \tag{5}$$

where $g^{(k)} = g(x^{(k)}), B_k$ is a symmetric positive definite approximation to the Hessian matrix $G_k = G(x^{(k)}), \alpha_k$ is a step length determined by line search. An ideal choice for the step length is

$$\alpha_k = \arg\min_{\alpha>0} f(x^{(k)} + \alpha d^{(k)}). \tag{6}$$

Since the ideal choice of the step length is impracticable and unnecessary, inexact line searches are usually carried out to give a step length satisfying

$$f(x^{(k)} + \alpha_k d^{(k)}) \leq f(x^{(k)}) + \rho\alpha_k {g^{(k)}}^T d^{(k)} \tag{7}$$

$$| \, g(x^{(k)} + \alpha_k d^{(k)})^T d^{(k)} \, | \leq -\sigma {g^{(k)}}^T d^{(k)}. \tag{8}$$

With $\rho \in (0, \frac{1}{2})$ and $\sigma \in (\rho, 1)$, an interval of acceptable α values always exists and an efficient line search strategy to find such a step length can be found in [2].

Different choices for B_k in (4) generates different descent methods. For example, the Gauss-Newton method with $B_k = A_k^T A_K$ and the quasi-Newton methods with B_k being obtained from qausi-Newton updating formulae are well known. Since B_k in the Gauss-Newton method is obtained by neglecting the second part of G_k, the method is expected to perform well when residuals at x^* are small enough or the function $r_i(x), i = 1, 2, \cdots, m$ are close to linear. The quasi-Newton methods such as the BFGS method and the DFP method approximate the whole Hessian matrix by using quasi-Newton update formulae and information obtained from first derivative values. However the quasi-Newton methods do not take account of the special structure of the problem.

Another kind of descent methods for nonlinear least squares is the hybrid method between the Gauss-Newton and the quasi-Newton method[1,10]. Depending upon the outcome of a certain test, the method chooses B_k to be either the Gauss-Newton matrix or the result of applying an updating formula to B_{k-1}. Numerical experiments[10] show that hybrid methods match or improve on the better of the Gauss-Newton and the quasi-Newton methods for every test problem and therefore give reliable, superlinearly convergent methods that contain the best features of both the Gauss-Newton and the quasi-Newton methods.

Since the Jacobian matrix $A(x)$ is usually calculated analytically or numerically in nonlinear lesat squares algorithms, the first portion $A(x)^T A(x)$ of $G(x)$ is always readily