NUMERICAL MATHEMATICS
A Journal of Chinese Universities (English Series)

Vol. **16** No. 1, pp. 92-96
February 2007

A Hybrid Method for Nonlinear Least Squares Problems

Zhongyi Liu[1,*] and Linping Sun[2]

[1] *School of Sciences, Hehai University, Nanjing 210098, China.*
[2] *Department of Mathematics, Nanjing University, Nanjing 210093, China.*

Received November 29, 2005; Accepted (in revised version) May 18, 2006

Abstract. A negative curvature method is applied to nonlinear least squares problems with indefinite Hessian approximation matrices. With the special structure of the method, a new switch is proposed to form a hybrid method. Numerical experiments show that this method is feasible and effective for zero-residual, small-residual and large-residual problems.

Key words: Nonlinear least squares; switch; hybrid method; negative curvature; BP factorization.

AMS subject classifications: 65K05, 90C30

1 Introduction

Consider nonlinear least squares problems

$$\min_{x\in R^n} F(x)=\frac{1}{2}f(x)^Tf(x)=\frac{1}{2}\sum_{i=1}^{m}f_i(x)^2 \tag{1}$$

where $m\geq n$, $f:\mathbb{R}^n\to\mathbb{R}^m\in C^2(\Omega)$, $\Omega\in\mathbb{R}^n$ is an open convex set and $f_i(x)$ is the component function of $f(x)$. The gradient of $F(x)$ is

$$g(x)=J(x)^Tf(x), \tag{2}$$

where $J(x)$ is the Jacobian matrix of $f(x)$, and the Hessian matrix is

$$G(x)=J(x)^TJ(x)+\sum_{i=1}^{m}f_i(x)\nabla^2 f_i(x).$$

Set

$$M(x)=J(x)^TJ(x),\quad W(x)=\sum_{i=1}^{m}f_i(x)\nabla^2 f_i(x). \tag{3}$$

*Correspondence to: Zhongyi Liu, School of Sciences, Hehai University, Nanjing 210098, China. Email: zhyi@hhu.edu.cn

Then

$$G(x) = M(x) + W(x). \tag{4}$$

Using the special structures of the object function $F(x)$ and the Hessian matrix $G(x)$, many effective methods have been developed. Among them a fundamental method is Gauss-Newton method which neglects the nonlinear term $W(x)$ in $G(x)$. In other words, a search direction is given by

$$J(x_k)^T J(x_k) p_k = -J(x_k)^T f(x_k). \tag{5}$$

The following theorem shows the convergence of the Gauss-Newton method.

Theorem 1.1. *Suppose that $F(x) \in C^2(\Omega)$, x^* is a local minimum of (1), $J(x)$ and $G(x)$ are Lipschitz continuous in Ω, and for all $x \in \Omega$, $J(x)$ is of full rank. If $\|J(x)\| \leq \delta$, $\|(J(x)^T J(x))^{-1}\| \leq \tau$, where δ and τ are constants, then Gauss-Newton iteration is well-defined for all $x \in \Omega$, and*

$$\|x^{(k+1)} - x^*\| \leq \|(J(x^*)^T J(x^*))^{-1} W(x^*)\| \|x^{(k)} - x^*\| + \mathcal{O}(\|x^{(k)} - x^*\|^2). \tag{6}$$

From the theorem above, whether Gauss-Newton method can succeed depends on whether the neglected term $W(x)$ is important, that is to say, whether $W(x)$ is a small part in $G(x)$. The Gauss-Newton method has quadratic rate of convergence for zero residual problems where $f(x^*) = 0$ or $W(x^*) = 0$.

The search direction can also be obtained by

$$(J(x^{(k)})^T J(x^{(k)}) + \lambda_k I) p^{(k)} = -J(x^{(k)})^T f(x^{(k)}) \tag{7}$$

where the nonnegative scalar λ_k is used to make $J(x^{(k)})^T J(x^{(k)}) + \lambda_k I$ positive definite. This formula is first proposed by Levenberg [4] and Marquardt [5], and is therefore called Levenberg-Marquardt method.

Another method takes advantage of $W(x)$ in $G(x)$, which is necessary for large residuals. One of this type of methods is due to Dennis-Gay-Welsh [6]. Since

$$\nabla^2 f_i(x^{(k+1)}) s^{(k)} = \nabla f_i(x^{(k+1)}) - \nabla f_i(x^{(k)}), \tag{8}$$

we have

$$f_i(x^{(k+1)}) \nabla^2 f_i(x^{(k+1)}) s^{(k)} = f_i(x^{(k+1)}) (J_{k+1} - J_k)^T e_i, \tag{9}$$

which leads to

$$\sum_{i=1}^{m} f_i(x^{(k+1)}) \nabla^2 f_i(x^{(k+1)}) s^{(k)} = (J_{k+1} - J_k)^T f^{(k+1)}. \tag{10}$$

Set $y^\sharp = (J_{k+1} - J_k)^T f^{(k+1)}$. Then W_{k+1} satisfies

$$W_{k+1} s = y^\sharp. \tag{11}$$

The Dennis-Gay-Welsh method gave the updating formula for W_k and scale strategy as follows:

$$W_{k+1} = \tau W_k + \frac{(y^\sharp - \tau W_k s) y^T + y (y^\sharp - \tau W_k s)^T}{y^T s} - \frac{(y^\sharp - \tau W_k s)^T s}{(y^T s)^2} y y^T, \tag{12}$$