Journal of Computational Mathematics, Vol.26, No.3, 2008, 390–403.

ON THE SEPARABLE NONLINEAR LEAST SQUARES PROBLEMS*

Xin Liu

LSEC, ICMSEC, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, and Graduate University of Chinese Academy of Sciences, Beijing 100190, China

Email: liuxin@lsec.cc.ac.cn

Yaxiang Yuan

LSEC, ICMSEC, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing 100190, China

Email: yyx@lsec.cc.ac.cn

Dedicated to Professor Junzhi Cui on the occasion of his 70th birthday

Abstract

Separable nonlinear least squares problems are a special class of nonlinear least squares problems, where the objective functions are linear and nonlinear on different parts of variables. Such problems have broad applications in practice. Most existing algorithms for this kind of problems are derived from the variable projection method proposed by Golub and Pereyra, which utilizes the separability under a separate framework. However, the methods based on variable projection strategy would be invalid if there exist some constraints to the variables, as the real problems always do, even if the constraint is simply the ball constraint. We present a new algorithm which is based on a special approximation to the Hessian by noticing the fact that certain terms of the Hessian can be derived from the gradient. Our method maintains all the advantages of variable projection based methods, and moreover it can be combined with trust region methods easily and can be applied to general constrained separable nonlinear problems. Convergence analysis of our method is presented and numerical results are also reported.

Mathematics subject classification: 65K05, 65H10.

Key words: Separable nonlinear least squares problem, Variable projection method, Gauss-Newton method, Levenberg-Marquardt method, Trust region method, Asymptotical convergence rate, Data fitting.

1. Introduction

In this paper, we consider the separable nonlinear least squares problem, which is derived from the following special nonlinear data fitting problem

$$y_i = \sum_{j=1}^{p} a_j \phi_j(b, t_i), \quad (i = 1, 2, \cdots, m), \tag{1.1}$$

where $\phi_j(b,t)(j = 1, \cdots, p)$ are real functions defined on $\Re^{q+1}$, t_i and y_i $(i = 1, \cdots, m)$ are given data, $a \in \Re^p$ and $b \in \Re^q$ are parameters to be decided, and p and q are two positive

* Received February 16, 2008 / Revised version received February 18, 2008 / Accepted March 3, 2008 /

integers satisfying $p + q = n$. In practice, it is usual to have $m > n$, which is also assumed in this paper. Data fitting problem (1.1) can be written as general nonlinear equations as follows

$$f_i(a, b) = 0, \quad i = 1, 2, \cdots, m, \tag{1.2}$$

where

$$f_i(a, b) = y_i - \sum_{j=1}^{p} a_j \phi_j(b, t_i), \quad i = 1, 2, \cdots, m. \tag{1.3}$$

It is easy to see that the nonlinear equations (1.2) are linear to variable set a. Models of this type are very common and have a variety of applications in different fields, such as inverse problems, signal analysis, medical and biological imaging, neural networks, robotics and vision, telecommunications, electrical and electronics engineering, environmental sciences and time series analysis, differential equations and dynamical systems, etc. (see, e.g., [4, 5]). It is natural to consider the nonlinear least squares formulation

$$\min_{a \in R^p, b \in R^q} \frac{1}{2} \sum_{i=1}^{m} (f_i(a, b))^2, \tag{1.4}$$

because (1.4) is equivalent to (1.2) when the latter has solutions and numerical methods for (1.4) can be used for solving (1.2) (see, e.g., [2, 12, 16, 19]).

If y_i $(i = 1, \cdots, m)$ depend on b, the data fitting problem (1.1) is generalized to the standard form of separable nonlinear least squares problems

$$\min_{a \in R^p, b \in R^q} \psi(a, b) = \frac{1}{2} || \ y(b) - \Phi(b) a ||_2^2, \tag{1.5}$$

where $y : R^q \mapsto R^m$; $\Phi : R^q \mapsto R^{m \times p}$, with $(\Phi(b))_{ij} = \phi_j(b, t_i)$, $1 \le i \le m$, $1 \le j \le p$.

Golub and Pereyra [4] proposed the variable projection method for (1.4), which is (1.5) with $y(b) = y_0$. The main idea of their variable projection method is as follows. For any fixed $b \in R^q$, (1.5) reduces to a linear least squares problem and we can obtain the least-norm solution

$$\hat{a}(b) = \Phi^+(b) y_0, \tag{1.6}$$

where $\Phi^+(b)$ is the Moore-Penrose inverse of $\Phi(b)$. Substituting (1.6) into (1.4), we have

$$\min_{a \in R^p, b \in R^q} \psi(a, b) = \min_{b \in R^q} \frac{1}{2} || \ y_0 - \Phi(b) \Phi^+(b) y_0 ||_2^2 = \min_{b \in R^q} \frac{1}{2} || P_{\Phi(b)}^{\perp} y_0 ||_2^2. \tag{1.7}$$

Here, $P_{\Phi(b)}^{\perp}$ is the orthogonal projector from R^m to the null space of $\Phi(b)^T$. Thus, we derive a reduced problem (1.7). A solution $\hat{b}$ of (1.7) can be obtained by applying any nonlinear least squares algorithm, and consequently $\hat{a}(\hat{b})$ can be defined by (1.6). In Golub and Pereyra [5], the computing and storage methods for the Fréchet derivative of the orthogonal projector are also developed, and it is proved that the separating variables approach led to the same solution set as that of the original problem when $\Phi(b)$ has a constant rank. The main feature of the variable projection methods is the elimination of the linear variables, which leads to three main advantages over the standard Gauss-Newton method: less iteration steps to convergence; less initial guess; decreasingly ill-conditioned if the whole problem is.

Kaufman[6] simplified the Jacobian formula of the orthogonal projector in Golub and Pereyra's method. It has extensively demonstrated that savings of up to 25% are achieved