Journal of Mathematical Systems, Estimation, and Control
Vol. 5, No. 1, 1995, pp. 1–16

A Modification of the Trust-Region Gauss-Newton Method to Solve Separable Nonlinear Least Squares Problems*

Christine Böckmann

Abstract

A modification of the trust-region Gauss-Newton method with a two-parameter approximation for the solution of separable nonlinear least squares problems is described and analysed. Global convergence results are presented. The new method of a regularized variable projection leads to a numerically stable implementation.

Key words: curve fitting, nonlinear least squares estimation, trust-region Gauss-Newton method, variable projection method, global convergence

AMS Subject Classifications: 65H10, 65F20, 65D10

1 Introduction

In mathematical modelling nonlinear formulations are continuously used. Numerous problems in physics, mechanical engineering, chemistry and medicine lead to a special nonlinear regression problem; specifically, we have to fit the model function

$$\eta(a, x, t) := \sum_{j=1}^{l} a_j \varphi_j(x, t) + \varphi_o(x, t) \tag{1}$$

to the given data $(t^i, y_i) \in R^k \times R^1$ $(i = 1, ..., m)$, x and a are to be estimated. In this paper we consider the analysis of models which are solely "linear combinations" of certain nonlinear functions. We are often faced with this problem in the field of physics and biology in evaluating spectroscopic data, in particular, if one attempts to fit given data in a least

*Received September 16, 1992; received in final form September 9, 1993. Summary appeared in Volume 5, Number 1, 1995.

squares sense to a nonlinear model, as for example in an exponential fitting. These problems also arise in some specific mathematical formulations when studying, e.g., approximations to solutions of parabolic linear partial differential equation, cf. Gilliam, Lund, Martin [10] or specific cases for the identification of linear ordinary differential equations where an exponential interpolation is used, cf. Ammar, Dayawansa, Martin [1]. The nonlinear functions φ_j $(j = 1, ..., l)$ are given, whereas

$$\{\Phi(x)\}_{ij} := \varphi_j(x, t^i) \quad (i = 1, ..., m; j = 1, ..., l) \tag{2}$$

is a matrix function $\Phi : x \in D_x \subset R^n \to L(R^l, R^m)$. We face the problem, however, to estimate $a = (a_1, ..., a_l)^T \in R^l$ and $x = (x_1, ..., x_n)^T \in R^n$ in the least squares sense. It appears to be necessary to determine

$$\min\{r_1(a, x) : a \in R^l, x \in R^n\} \tag{3}$$

$$\text{whereas} \qquad r_1(a, x) := \frac{1}{2}\sum_{i=1}^{m}(y_i - \eta(a, x, t^i))^2 = \frac{1}{2}\|H(a, x)\|^2$$

$$\text{with} \qquad H(a, x) := z(x) - \Phi(x)a$$

$$\text{and} \qquad z(x) := (y_1 - \varphi_o(x, t^1), ..., y_m - \varphi_o(x, t^m))^T.$$

The norm is always the Euclidean vector norm, or the spectral norm for matrices. The problem (3) is a nonlinear least squares problem of the dimension $l + n$. Golub and Pereyra [11] proved that this problem can be reduced to a nonlinear least squares problem of the dimension n comprising $x \in R^n$ only, and a linear least squares problem comprising $a \in R^l$ only. A variable projection algorithm using the Gauss-Newton method to solve the problem (3), was developed by Golub and Pereyra. Kaufman [12] proposed a simplification of this algorithm, which proved to be more efficient when using a computer. Ruhe and Wedin [17] offered a more general analysis of different algorithms for separable problems. It was proved that the Gauss-Newton algorithm when applied to both (6) as well as the original problem have the same asymptotic convergence rate. In particular, both converge quadratically for the zero residual problem. This is in contrast to the naive algorithm when minimizing over a and x alternatively; which always converges linearly. They also proved that the simplified algorithm of Kaufman has roughly the same asymptotic convergence rate as the one proposed by Golub and Pereyra, cf. Björck [2].

2 The Method of Regularized Variable Projection

Let us first fix $x \in R^n$. We then obtain a linear least squares problem in $a \in R^l$

$$\min\{r_1(a, x) : a \in R^l\}. \tag{4}$$

The solution with the smallest Euclidean vector norm is

$$a^*(x) := \Phi(x)^+ z(x). \tag{5}$$

This is a function of $x \in R^n$, where $\Phi(x)^+$ is the Moore-Penrose pseudoinverse of $\Phi(x)$. Substituting (5) in (3) leads to a reduced nonlinear least squares problem in $x \in R^n$ of the dimension n,

$$\min\{r_2(x) : x \in R^n\} \tag{6}$$

$$\begin{aligned} \text{whereas} \qquad & r_2(x) := r_1(a^*(x), x) = \frac{1}{2}\ \|G(x)\|^2 \\ \text{with} \qquad & G : x \in D_x \subset R^n \to G(x) \in R^m, \\ \text{and} \qquad & G(x) := (I - \Phi(x)\Phi(x)^+)z(x). \end{aligned}$$

This is advantageous, because the number of iteration steps in general is smaller and the iteration is more stable, i.e. the variable projection algorithm is able to solve problems which other methods not using separability may not be able to solve, cf. Krogh [13]. By solving (6) we get an optimal $x^* \in R^n$. So the first subproblem is a nonlinear least squares problem in $x \in R^n$ and another advantage is that no initial values for the linear parameters have to be provided, whereas the second subproblem is simply a linear least squares problem of finding $a^* \in R^l$. Now the question is raised, which method is favourable for the solution of (6). The method of Kaufman is useful, although there are two other suggestions. The process of deriving $G(x)$ is extensive. Therefore our algorithm is at first without derivative. A finite-difference approximation with function evaluations for $DG(x)$ is used. Consequently it is necessary to compute the function evaluations effectively. For every evaluation the correction $\Phi(x)^+ z(x)$ is needed, although not explicitly. We have to solve the problem

$$\min\{\|z(x) - \Phi(x)a\|^2 : a \in R^l\} \tag{7}$$

effectively and stably e.g. with the QR-factorization by either Householder or Givens. If $\Phi(x)$ is singular or the pseudorank is smaller than l, i.e. $\Phi(x)$ is nearly rank deficient, then we will need e.g. the singular value decomposition of $\Phi(x)$ for every function evaluation. Again this requires great algebraical expense. Let us now assume

$$rk(\Phi(x)) = r = (constant) \leq l \tag{8}$$

for all $x \in D_x$. The idea is to determine $a := a_R(x, \varepsilon)$ in a Tikhonov-like sense. The linear least squares problem is replaced by

$$\min\{\|z(x) - \Phi(x)a\|^2 + \varepsilon\|a\|^2 : a \in R^l\} \tag{9}$$

with $\varepsilon > 0$ as regularization parameter. The solution is effectively possible by QR-factorization. What has been described so far is well-known. In the following let us describe what we would like to call the method of regularized variable projection from another angle. This method may also be described as a method with an approximation $F(x, \varepsilon)$ for the function evaluation $G(x)$. Then

$$F : (x, \varepsilon) \in D_x \times D_\varepsilon \subset R^n \times R^1_+ \to F(x, \varepsilon) \in R^m \tag{10}$$

$$\text{where} \qquad F(x, \varepsilon) := z(x) - \Phi(x) a_R(x, \varepsilon) = (I - \Phi(x)\tilde{\Psi}(x, \varepsilon))z(x) \ ,$$

$$\text{and} \qquad \tilde{\Psi}(x, \varepsilon) := \begin{cases} \Phi(x)^+ & : \quad \varepsilon = 0 \\ \Psi(x, \varepsilon) & : \quad \varepsilon \in (0, \varepsilon_o], \varepsilon_o > 0 \end{cases}$$

with $\Psi(x, \varepsilon) := (\varepsilon I + \Phi(x)^T \Phi(x))^{-1} \Phi(x)^T$ and $[0, \varepsilon_o] \subset D_\varepsilon$. The parameter ε is decisive for the quality of the approximation. In this paper we propose a modification of the Trust-Region Gauss-Newton method with a two-parameter approximation.

3 Preliminaries

The trust-region Gauss-Newton method is commonly used for solving nonlinear least squares problems

$$\min\{\frac{1}{2}\|G(x)\|^2 : x \in R^n\}. \tag{11}$$

The solution proceeds iteratively

$$x^{k+1} := x^k + s^k, k = 0, 1, 2, ... \tag{12}$$

where the correction s^k is an approximate solution of the constrained linear least squares problem

$$\min\{\|DG(x^k)s + G(x^k)\| : \|s\| \le \Delta_k\}. \tag{13}$$

Here $DG(x^k)$ denotes the Jacobian matrix of $G(x^k)$ at the current iterate x^k and $\Delta_k > 0$ is the trust-region radius. Let

$$s_{GN} := -DG(x^k)^+ G(x^k) \tag{14}$$

be the generalized Gauss-Newton step so that the solution of the unconstrained linear least squares problem

$$\min \|DG(x^k)s + G(x^k)\| \tag{15}$$

has minimal norm $\|s\|$. It is well-known that the Gauss-Newton step $s = s_{GN}$ solves (13) if $\|s_{GN}\| \le \Delta_k$. Otherwise, problem (13) has the unique solution $s = s(\lambda)$ where $s(\lambda)$ is defined as the solution of the unconstrained regularized linear least squares problem

$$\min\{\|DG(x^k)s + G(x^k)\|^2 + \lambda\|s\|^2 : s \in R^n\} \tag{16}$$

and $\lambda > 0$ is determined from the scalar equation

$$\beta(\lambda) := \|s(\lambda)\| - \Delta_k = 0. \tag{17}$$

An appropriate algorithm for solving (16) and (17) was described by Moré [14], which has become part of most trust-region codes. Since $\beta(\lambda)$ is nonlinear in λ, (17) has to be solved iteratively, too, so that the solution $s(\lambda)$ of (16) has to be repeatedly evaluated for different values of λ, as reported by Moré [14] and confirmed by Gay [9]. Let us now consider a common function $F(x, \varepsilon)$

$$F : D_x \times D_\varepsilon \subset R^n \times R^1_+ \rightarrow R^m, m \geq n \tag{18}$$

and define $$f(x,\varepsilon) := \frac{1}{2}\|F(x,\varepsilon)\|^2$$

where f is bounded below and then

$$\min\{f(x, 0) : x \in R^n\} \tag{19}$$

has to be determined.

For a given $(x^o, \varepsilon_o) \in R^n \times R^1_+$

$$W_o := \{x \in R^n : \exists\varepsilon \in [0, \varepsilon_o] \quad \text{with} \quad f(x,\varepsilon) \leq f(x^o, \varepsilon_o)\} \tag{20}$$

is the level set of F. If F is differentiable with respect to x, so is f and the gradient of f

$$df(x,\varepsilon) := DF(x,\varepsilon)^T F(x,\varepsilon) \tag{21}$$

is of the above form.

Assumption 3.1 *Let us agree that the function F in (18) fulfills the assumption (3.1), if the following conditions are satisfied*

i) *The level set W_o is compact for a given (x^o, ε_o). There exists $D_o := conv(W_o) \subset D_x$ with D_x open.*

ii) *The derivative $DF(x, \varepsilon)$ exists with respect to x on D_x and is bounded, i.e.*

$$\|DF(x,\varepsilon)\| \leq M_1 \tag{22}$$

for all $(x, \varepsilon) \in D_o \times D_\varepsilon$ with $M_1 > 0$.

iii) *The functions F and DF are Lipschitz-continuous at the point $\varepsilon = 0$. There are two constants $L_o, L_1 > 0$ with*

$$\|F(x,\varepsilon) - F(x,0)\| \leq L_o\varepsilon \tag{23}$$

$$\text{and} \qquad \|DF(x,\varepsilon) - DF(x,0)\| \leq L_1\varepsilon \tag{24}$$

for all $(x,\varepsilon),(x,0) \subset D_o \times D_\varepsilon$.

iv) *The function DF is also Lipschitz-continuous with respect to x. There exists a constant $L_2 > 0$ with*

$$\|DF(x,\varepsilon) - DF(y,\varepsilon)\| \leq L_2\|x-y\| \tag{25}$$

for all $(x,\varepsilon),(y,\varepsilon) \in D_o \times D_\varepsilon$.

Definition 3.1 *The approximation A of DF*

$$A : D_x \times D_\varepsilon \times D_h \subset R^n \times R^1_+ \times R^n \to L(R^n, R^m)$$

is said to be a uniformly strongly consistent approximation, or shorter a (C,r)-approximation if

$$\|A(x,\varepsilon,h) - DF(x,\varepsilon)\| \leq C\|h\| \tag{26}$$

for all $(x,\varepsilon,h) \in D_x \times D_\varepsilon \times D_h^r$ with two constants $C, r > 0$.

$D_h \subset R^n$ is called the discretisation domain, where $0 \in R^n$ is an accumulation point of D_h and

$$D_h^r := \{h \in D_h : \|h\| \leq r\} \subset D_h. \tag{27}$$

The constant C is independent of ε. From our assumptions the (C,r)-approximation exists, cf. Schwetlick [18]. The discretisation of the gradient is given by

$$b := b(x,\varepsilon,h) = A(x,\varepsilon,h)^T F(x,\varepsilon). \tag{28}$$

Lemma 3.1 *Let the function F satisfy the assumption (3.1) and A be a (C,r)-approximation on $W_o \times D_\varepsilon$. Then there exist constants K_1, K_2, K_3 and $M_2 > 0$ in such a way that we have for all $(x,\varepsilon,h) \in W_o \times D_\varepsilon \times D_h^r$ and for all $(x,\varepsilon) \in W_o \times D_\varepsilon$*

i)

$$\|b(x,\varepsilon,h) - df(x,\varepsilon)\| \leq K_1\|h\|, \tag{29}$$

ii)

$$\|df(x,\varepsilon) - df(x,0)\| \leq K_2\varepsilon, \tag{30}$$

iii)

$$||df(x,\varepsilon) - df(y,\varepsilon)|| \leq K_3||x-y||, \tag{31}$$

iv)

$$||A(x,\varepsilon,h)|| \leq M_2. \tag{32}$$

Proof: The proof is simple. Indeed, when the conditions (20), (27) and the inequalities (22), (23), (24), (25) and (26) are satisfied, we deduce the assertions (29), (30), (31) and (32). □

Definition 3.2 *The function $f : D_x \times D_\varepsilon \subset R^n \times R^1_+ \rightarrow R^1$ is called parameter monotone on $D_o \times D_\varepsilon$, if*

$$f(x,\varepsilon_1) \leq f(x,\varepsilon_2) \tag{33}$$

for all $(x,\varepsilon_1),(x,\varepsilon_2) \in D_o \times D_\varepsilon$ with $\varepsilon_1 \leq \varepsilon_2$.

4 Algorithm: Statement and Convergence

Following (13) we now have to determine

$$\min\{m_k : s \in R^n, ||s|| \leq \Delta_k\} \tag{34}$$

$$\text{with} \qquad m_k(s) := ||A(x^k,\varepsilon_k,h^k)s + F(x^k,\varepsilon_k)||.$$

Let us use the notations

$$\text{ared} := f(x^k,\varepsilon_k) - f(x^k+s^k,\varepsilon_k)$$

for the actual reduction of the nonlinear function,

$$\text{pred} := f(x^k,\varepsilon_k) - \frac{1}{2}m_k(s^k)^2$$

for the predicted reduction of the linear model and

$$\rho_k := \frac{\text{ared}}{\text{pred}}$$

for the ratio measuring the agreement between the linear model and the nonlinear function. The iteration step is successful if $\rho_k \geq \eta_1$ otherwise unsuccessful. In the latter case the trust-region radius Δ_k has to be reduced. After a successful iteration step, Δ_k is increased, if $\rho_k > \eta_2$, $0 < \eta_1 < \eta_2 < 1$, compare step 7.

Lemma 4.1 *Let $\tilde{s}$ be a solution to (34). Then*

$$\text{pred} \geq \frac{1}{2}||b|| \min(\Delta, \frac{||b||}{||A||^2}). \tag{35}$$

In fact, the inequality in Lemma (4.1) is obtained by Powell's "dog-leg" step [16]. This inequality is the main ingredient used to show that the sequence of gradients tends to zero for the following algorithm.

Lemma 4.2 *Let the function F satisfy the assumption (3.1) and let A be a (C, r)-approximation on $W_o \times D_\varepsilon$. Then there exists a constant $K_4 > 0$ so that for all $(x, \varepsilon, h) \in W_o \times D_\varepsilon \times D_h^r$*

$$|\text{ared} - \text{pred}| \leq K_4(\Delta + ||h||)(1 + \Delta)\Delta. \tag{36}$$

Proof: The assertion (36) can be deduced by using Taylor's theorem therefore employing condition (20) and the inequalities (22), (25), (26) and (32).

□

Algorithm 4.1

step 0 Choose $x^0 \in R^n, \Delta_o > 0, \varepsilon_{-1} > 0, j_{max} \in \{1, 2, ...\}, \beta_1, \beta_2 \in (0, 1)$, $0 < \eta_1 < \eta_2 < 1, 0 < \gamma_1 < 1 < \gamma_2$, tol and set $k := 0$.

step 1 Choose ε_k with $0 \leq \varepsilon_k \leq \varepsilon_{k-1}$ and $h^k \in D_h$ with $h^k \neq 0$.

step 2 Compute $F^k := F(x^k, \varepsilon_k)$.
<u>If</u> $F^k = 0$
<u>then</u> $N := k$, <u>stop</u>
<u>else</u> compute $A_k := A(x^k, \varepsilon_k, h^k)$ and $b^k := A_k^T F^k$, set $j := 0$.

step 3 <u>If</u> $||b^k|| >$tol <u>then</u> <u>goto</u> step 5.

step 4 Choose $\tilde{\varepsilon}_k$ with $0 \leq \tilde{\varepsilon}_k \leq \beta_2\varepsilon_k$ and $\tilde{h}^k \in D_h$ with $||\tilde{h}^k|| \leq \beta_1||h^k||$, $\tilde{h}^k \neq 0$, set $\varepsilon_k := \tilde{\varepsilon}_k, h^k := \tilde{h}^k$, <u>goto</u> step 2.

step 5 Compute s^k as minimum-norm-solution of

$$\min\{||F^k + A_k s|| : ||s|| \leq \Delta_k\},$$

set $j := j + 1$.

step 6 Compute $F(x^k + s^k, \varepsilon_k)$ and ρ_k.

step 7 <u>If</u> $\rho_k < \eta_1$
<u>then</u> <u>if</u> $j < j_{\max}$
<u>then</u> $\Delta_k := \gamma_1\Delta_k$, <u>goto</u> step 5
<u>else</u> <u>goto</u> step 4

$$\begin{array}{l} \underline{\text{else}}\ x^{k+1} := x^k + s^k,\ \underline{\text{if}}\ \rho_k > \eta_2 \\ \qquad\qquad\qquad \underline{\text{then}}\ \Delta_{k+1} := \gamma_2 \Delta_k \\ \qquad\qquad\qquad \underline{\text{else}}\ \Delta_{k+1} := \Delta_k \\ k := k+1,\ \underline{\text{goto}}\ \text{step 1.} \end{array}$$

It should be added, it might happen that $b^k = 0$ in step 3 even though $df(x^k, \varepsilon_k) \neq 0$. Then b^k is not a good approximation and we have to choose a smaller ε and $\|h\|$ in step 4. The constant j_{max} controls the approximations A_k and $F(x^k, \varepsilon_k)$. If we have decreased Δ_k j_{max}-times successively without success, we will have to choose a smaller ε and $\|h\|$ in step 4, too.

Theorem 4.1 *Assume the function $F : D_x \times D_\varepsilon \subset R^n \times R^1_+ \to R^m$ satisfy the assumption (3.1), the approximation A of DF on $W_o \times D_\varepsilon$ be a (C, r)-approximation, and the function f be parameter monotone. Then the algorithm (4.1) is practicable and produces either a finite or an infinite sequence $\{x^k\}$ whereas the sequence $\{f(x^k, \varepsilon_k)\}$ is strictly monotone decreasing and exactly one of the following three cases holds:*

i) *The algorithm finishes after N iterations because $F(x^N, \varepsilon_N) = 0$.*

ii) *After M iterations a cycle occurs where $\|h^M\|, \varepsilon_M$ and Δ_M are decreased unsuccessfully in its progress. Then x^M is a stationary point of f, i.e. $df(x^M, 0) = 0$.*

iii) *The algorithm is infinite. In case the parameters h^k and ε_k are chosen in such a way that*

$$\lim_{k\to\infty} \|h^k\| = \lim_{k\to\infty} \varepsilon_k = 0 \tag{37}$$

then

$$\lim_{k\to\infty} b^k = \lim_{k\to\infty} df(x^k, \varepsilon_k) = \lim_{k\to\infty} df(x^k, 0) = 0. \tag{38}$$

Proof: At first we will consider the case $df(x^k, 0) \neq 0$ and show that the inner cycles between the steps 3, 4 and 2 and between the steps 7, 4 and 2 are finite. Because of inequality (30) there exists a $j_o > 0$ so that $df(x^k, \varepsilon_{k_j}) \neq 0$ for all $j \geq j_o$ and furthermore by inequality (29) then exists $j_1 \geq j_o$ so that $\|b^k\| > \sigma > 0$ for all $j \geq j_1$. So the first inner cycle is finite. By (36), (35) and (32) we obtain

$$\begin{aligned} |\rho_k - 1| &\leq \frac{K_4(\Delta_k + \|h^k\|)(1+\Delta_k)\Delta_k}{|\text{pred}|} \\ &\leq \frac{K_4(\Delta_k + \|h^k\|)(1+\Delta_k)\Delta_k}{\frac{1}{2}\sigma \min(\Delta_k, \sigma/M_2^2)} \end{aligned} \tag{39}$$

for all $j \geq j_1$. Because of the instructions of decreasing $\varepsilon_k, \|h^k\|$ and Δ_k the sequence $|\rho_k - 1|$ tends to zero and so ρ_k tends to 1 so that the second cycle is finite, too. Assume the algorithm has been iterated k times. Then this may be assumed to be true

$$f(x^k, \varepsilon_k) < f(x^{k-1}, \varepsilon_{k-1}) < ... < f(x^0, \varepsilon_o).$$

Case 1: In step 2 $F(x^k, \varepsilon_k) = 0$. Then $N := k$ and the algorithm stops. Condition (33) implies $F(x^k, 0) = 0$.

Case 2: Although $F(x^k, \varepsilon_k) \neq 0$, an inner cycle begins in the iteration step $M := k$. Then x^M is a stationary point, i.e. $df(x^M, 0) = 0$. This need not necessarily lead to $F(x^k, 0) = 0$.

Case 3: Let $F(x^k, \varepsilon_k) \neq 0$ and $df(x^k, 0) \neq 0$ so that with h^k, ε_k and Δ_k the new iteration can be formed. If step 4 is run through i-times then with condition (33) we have

$$f(x^k, \varepsilon_{k_i}) \leq f(x^k, \beta_2^i \varepsilon_k) < f(x^k, \varepsilon_k).$$

Moreover $x^{k+1} = x^k + s^k$ including $b^k \neq 0$ and $\rho_k \geq \eta_1$. This implies that

$$\begin{aligned} f(x^k, \varepsilon_k) - f(x^{k+1}, \varepsilon_{k+1}) &\geq f(x^k, \varepsilon_k) - f(x^{k+1}, \varepsilon_k) \geq \eta_1 \text{pred} \\ &\geq \frac{1}{2}\eta_1 \|b^k\| \min(\Delta_k, \frac{\|b^k\|}{\|A_k\|^2}) > 0. \end{aligned} \tag{40}$$

Indeed $x^{k+1} \in W_o$. The sequence $\{f(x^k, \varepsilon_k)\}$ is strictly monotone decreasing and bounded below. The second part of the proof is indirect and follows a well-known idea, cf. Thomas [20] or Sorensen [19]. Supposing that a subsequence $\{x^{k_j}\} \subset \{x^k\}$ exists so that $\|b^{k_j}\| \geq \xi > 0$ for all j. We select an integer l_j corresponding to each j so that

$$l_j = \max\{l \in [k_j, k_{j+1}) : \|b^i\| \geq \frac{\xi}{2}, k_j \leq i \leq l\}. \tag{41}$$

The existence is guaranteed because $l_j = k_j$ is always possible. By the condition (33) and the inequalities (35) and (32) we obtain

$$\begin{aligned} f(x^l, \varepsilon_l) - f(x^{l+1}, \varepsilon_{l+1}) &\geq f(x^l, \varepsilon_l) - f(x^{l+1}, \varepsilon_l) \\ &\geq \eta_1 \text{pred} \\ &\geq \frac{1}{4}\eta_1 \xi \min(\Delta_l, \frac{\xi}{2M_2^2}). \end{aligned} \tag{42}$$

Furthermore we get that

$$f_{k_j} - f_{l_j+1} = \sum_{l=k_j}^{l_j} f_l - f_{l+1} \geq \frac{1}{4}\eta_1 \xi \min(\sum_{l=k_j}^{l_j} \Delta_l, \frac{\xi}{2M_2^2}). \tag{43}$$

So it can be deduced

$$\lim_{j\to\infty} \sum_{l=k_j}^{l_j} \Delta_l = 0$$

because f is bounded below and strictly monotone decreasing. By

$$\|x^{k_j} - x^{l_j+1}\| \leq \sum_{l=k_j}^{l_j} \|s^l\| \leq \sum_{l=k_j}^{l_j} \Delta_l$$

we obtain

$$\lim_{j\to\infty} \|x^{k_j} - x^{l_j+1}\| = 0. \tag{44}$$

Since (29), (30) and (31) are satisfied it follows

$$\begin{aligned} \|b^{k_j} - b^{l_j+1}\| &= \|b^{k_j} - df(x^{k_j}, \varepsilon_{k_j}) + df(x^{k_j}, \varepsilon_{k_j}) - df(x^{l_j+1}, \varepsilon_{k_j}) \\ &\quad + df(x^{l_j+1}, \varepsilon_{k_j}) - df(x^{l_j+1}, 0) + df(x^{l_j+1}, 0) \\ &\quad - df(x^{l_j+1}, \varepsilon_{l_j+1}) + df(x^{l_j+1}, \varepsilon_{l_j+1}) - b^{l_j+1}\| \\ &\leq K_1\|h^{k_j}\| + K_3\|x^{k_j} - x^{l_j+1}\| + K_2\varepsilon_{k_j} \\ &\quad + K_2\varepsilon_{l_j+1} + K_1\|h^{l_j+1}\|. \end{aligned} \tag{45}$$

Moreover, because of (44) and (45) there exists an index $j_o > 0$ so that $\|b^{k_j} - b^{l_j+1}\| < \xi/4$ for all $j \geq j_o$. Therefore

$$\|b^{k_j}\| = \|b^{k_j} - b^{l_j+1} + b^{l_j+1}\| \leq \frac{\xi}{4} + \frac{\xi}{2} < \xi \tag{46}$$

for all $j \geq j_o$. Obviously, this is a contradiction and so with (29) and (30) the proof is complete. □

We shall only briefly describe the local asymptotic properties. Let the assumptions of Theorem(4.1) be satisfied. Roughly speaking, if the sequence $\{x^k\}$ converges to x^* with $df(x^*, 0) = 0$ and if the model has a sufficiently small residuum at x^* and $rk(DF(x^*, 0)) = n$ then $\|s_{GN}\| \leq \Delta_k$, i.e. we have that $\lambda_k = 0$ for all $k \geq k_o$, cf. (15) and (16). This means that the algorithm(4.1) for $k \geq k_o$ becomes the undamped Gauss-Newton method with a two-parameter approximation and then it even converges superlinearly, i.e. faster than linearly. For a closer look at the proof of local convergence results, cf. Böckmann [4].

5 Implementation

The algorithm developed in Section 4 has been applied mainly for parameter estimation in models whose parameters are separated, i.e. the algorithm is applied to function (10). The approximation parameter ε is used as the regularization parameter of the variable projection. It is easy to show that function (10) satisfies the assumption (3.1) and the properties (26) and (33) are fulfilled under simple assumptions on $\Phi(x)$ and $z(x)$. Roughly speaking, $\Phi(x)$ and $z(x)$ have to be differentiable, $D\Phi(x)$ and $Dz(x)$ have to be Lipschitz-continuous and condition (8) has to be fulfilled. For the detailed proof of these properties, cf. Böckmann [4].

We compared three methods for solving problem (1). The first and the second ones are methods without separation, cf. (3). They are common trust-region Gauss-Newton methods with a one-parameter approximation for the derivative of $\tilde{H}(a,x) = y - \Phi(x)a$ with $\varphi_o = 0$. The first method uses a complete approximation for $D\tilde{H}$. The second method on the other hand uses a subapproximation. Because of the special structure of $\tilde{H}(a,x)$, we determine

$$D\tilde{H}(a,x) = [D\tilde{H}_a | D\tilde{H}_x] = [-\Phi(x) | D\Phi(x)a]$$

and only the second part has to be discretized. The third method is the presented algorithm (4.1) with a two-parameter approximation. For the implementation we use the LINPACK-routines, cf. Dongarra, Bunch, Moler, Stewart [7] and the NL2SOL-algorithm, cf. Dennis, Gay, Welsch [6]. We used our algorithms for solving different problems known from literature for a comparison. The results for some of these problems are presented below.

Example 1: Willers [21]

$$\eta(a,x,t) = a_1 + a_2 e^{x_1 t}, \qquad m = 10$$

t	2	4	6	8	10
y	92.4	86.2	80.5	75.2	70.3

t	12	14	16	18	20
y	65.8	61.6	57.7	54.1	50.8

1. initial value vector: $[10, 100, (-0,01)]^T$
2. initial value vector: $[-145, 250, (-0.01)]^T$

Example 2: Golub/Pereyra [11]

$$\eta(a,x,t) = a_1 + a_2 e^{x_1 t} + a_3 e^{x_2 t}, \qquad m = 33$$

For the measuring series cf. [11].
Initial value vector: $[0.4, 1.5, -1.5, (0.01, 0.02)]^T$

Example 3: Ruhe/Wedin [17]

$$\eta(a,x,t) = a_1 + a_2 \frac{1}{t+x_1}, \qquad m = 9$$

t	0	0.15625	0.3125	0.625	
y	20182	19585	19190	17746	
$\overline{y}$	20100	19237	18228	16630	
t	1.25	2.5	5	10	20
y	15244	12177	9175	6406	4970
$\overline{y}$	13826	10748	8200	6287	4946

1. initial value vector: $[2000, 60000, (3)]^T$
2. initial value vector: $[3000, 30000, (3)]^T$

Example 4: Ekenberg [8] / Clair,Rigler [5]

$$\eta(a,x,t) = a_1 exp\{-\frac{4ln2(x_1-t)^2}{x_2^2}\} + a_2 exp\{-\frac{4ln2(x_3-t)^2}{x_4^2}\}$$

Two measuring series with different parameters were generated.

1) m=57, t=0(0.1)5.6

$$\begin{aligned} a &= [65.97176, 76.66948]^T \\ x &= [3.97588, 0.61526, 2.52642, 0.87850]^T \end{aligned}$$

initial value vector: $[70, 80, (3.2111, 1.7813, 3.0817, 1.7795)]^T$

2) m=71, t=0(0.1)7

$$\begin{aligned} a &= [57.5361, 68.62627]^T \\ x &= [2.50158, 1.46932, 2.25775, 0.74416]^T \end{aligned}$$

initial value vector: $[60, 70, (3.2111, 1.7813, 3.0817, 1.7795)]^T$

Example 5: Ottoy, Vansteenkiste [15]

$$\eta(a,x,t) = a_1 + a_2 tanh(x_1(lnt - x_2)), \qquad m = 50$$

The measuring series were generated by $t = 0.2(0.2)10$, $a = [200, 150]^T$ and $x = [3, 1]^T$.
initial value vector: $[180, 180, (7, 2)]^T$

Example 6:

$$\eta(a,x,t) = a_1 e^{x_1 t} cos(x_2 t) + a_2 e^{x_1 t} sin(x_2 t), \qquad m = 5$$

t	0.5	1	1.5	2	2.33
y	5.3	-2.3	-9	2.2	13.2

1. initial value vector: $[1.5, 4.5, (0.3, 2)]^T$
2. initial value vector: $[1.5, 4.5, (0.3, 2)]^T$

Example 7: chemical problem

$$\eta(a,x,t) = a_1 \frac{t}{1 + x_1 t + x_2 t^2} + a_2 \frac{t^2}{1 + x_1 t + x_2 t^2}, \qquad m = 24$$

See Böckmann [3] for the measuring series.
1. initial value vector: $[800, 1500, (0.0595, 1.846)]^T$
2. initial value vector: $[304.976, 174.65, (0.09776, 1.9159)]^T$
3. initial value vector: $[150, 200, (0.09398, 2.274)]^T$

Table 5.1 Runtime (R) and Number of Iteration Steps (I)

Example	*method* 1		*method* 2		*method* 3	
	R	I	R	I	R	I
1.1	2	5	1	5	1	3
1.2	8	24	8	24		
2	6	4	5	4	5	4
3.1	3	11	3	11	<1	3
3.2	4	15	4	15	1	4
4.1	d		d		26	9
4.2	d		d		36	11
5	54	41	47	43	46	41
6.1	4	8	3	8	2	5
6.2	d		d		2	5
7.1	4	6	4	6	3	4
7.2	5	7	5	7	4	5
7.3	4	5	3	5	3	4

The type of the approximation of $D\tilde{H}$ does not affect the number of iteration steps, but as expected the runtime is different. Method 2 is somewhat better. This fact is to be seen in Table 5.1, columns 1 and 2. Table 5.1, column 3 shows the advantage of separation. The number of iteration steps is smaller and so is the runtime. The examples 4.1, 4.2 and 6.2 cannot be solved by employing methods 1 or 2. This is remarkable, since solutions could be obtained using method 3. Moreover, method 3 is able to solve example 4.1, but it has to be taken into consideration that ε_o is truly greater than 0. So the reduction of the dimension and the regularization of the variable projection stabilize the computation at a low expense. This is very important. The algorithm is user-friendly, because it does need neither an initial value for $a \in R^l$ nor the derivative DF.

Our numerical experience shows that the method of regularized variable projection provides an efficient and stable algorithm for solving separable nonlinear least squares problems.

Acknowledgments

I would like to take the opportunity to thank Hubert Schwetlick and Oswald Knoth, Martin-Luther-Universitt Halle-Wittenberg, for their suggestions and constant encouragement.

References

[1] G. Ammar, W. Dayawansa and C. Martin. Exponential interpolation: theory and numerical algorithms, *Applied Mathematics and Computation* **44** (1991), 189-232.

[2] Å. Björck. Solution of Equations in R^n Vol.I, in *Handbook of Numerical Analysis* (P.G. Ciarlet and J.l. Lions, eds.) Amsterdam: Elsevier Science (North Holland), 1990.

[3] Ch. Böckmann. Ein ableitungsfreies Verfahren vom Gauss–Newton–Typ zur Lösung von nichtlinearen Quadratmittelproblemen mit separierten Variablen, Ph. D. thesis, Technische Universität Dresden, 1984.

[4] Ch. Böckmann. Curve fitting and identification of physical spectra, preprint, September 1992.

[5] D.C.ST. Clair and A.K. Rigler. An algorithm for least squares analysis of spectroscopic data, *BIT* **19** (1979), 448-456.

[6] J.E. Dennis, Jr., D.M. Gay and R.E. Welsch. An adaptive nonlinear least squares algorithm, *TOMS* **7** (1981), 348-368.

[7] J.J. Dongarra, J.R. Bunch, C.B. Moler and G.W. Stewart. LINPACK Users' Guide, *Society for Industrial and Applied Mathematics*, 1979.

[8] B. Ekenberg. Curve fitting by use of graphic display, *BIT* **15** (1975), 385-393.

[9] D.M. Gay. Computing optimal locally constrained steps, *SIAM J. Sci. Statist. Comput.* **2** (1981), 186-197.

[10] D. Gilliam, J. Lund and C. Martin. Inverse parabolic problems and discrete orthogonality, *Numerische Mathematik* **59** (1991), 361-383.

[11] G.H. Golub and V. Pereyra. The differentiation of pseudo-inverses and nonlinear least squares problems whose variables separate, *SIAM J. Numer. Anal.* **10** (1973), 413-432.

[12] L. Kaufman. Variable projection methods for solving separable nonlinear least squares problems, *BIT* **15** (1975), 49-57.

[13] F.T. Krogh. Efficient implementation of a variable projection algorithm for nonlinear least squares, *Comm. ACM* **17** (1974), 167-169.

[14] J.J. Moré. The Levenberg–Marquardt algorithm: Implementation and theory, in *Lecture Notes in Mathematics 630* (G.A. Watson, ed.) New York: Springer-Verlag, 1978.

[15] J.P. Ottoy and G.C. Vansteenkiste. A computer program for non-linear curve fitting, *Adv. Eng. Software* **3** (1981), 55-61.

[16] M.J.D. Powell. A hybrid method for nonlinear equations, in *Numerical Methods for Nonlinear Algebraic Equations* (P. Rabinowitz, ed.) London: Gordon and Breach, 1970.

[17] A. Ruhe and P-Å. Wedin. Algorithms for separable nonlinear least squares problems, *SIAM Review* **22** (1980), 318-337.

[18] H. Schwetlick. *Numerische Lösung nichtlinearer Gleichungen.* Berlin: VEB Deutscher Verlag der Wissenschaften, 1979.

[19] D.C. Sorensen. Newton's method with a model trust region modification, *SIAM J. Numer. Anal.* **19** (1982), 409-426.

[20] S.W. Thomas. Sequential estimation techniques for quasi-Newton algorithms, Ph. D. thesis, Department of Computer Science, Cornell University, Report TR75-227, 1975.

[21] F.A. Willers. *Methoden der praktischen Analysis.* Berlin: Walter de Gruyter und Co., 1957.

Fachbereich Mathematik, Universitt Potsdam, Am Neuen Palais 10, D-14415 Potsdam, Germany

Communicated by Clyde F. Martin