On a two-step method for the nonlinear least squares problem with decomposition of operator

R.P. Iakymchuk[1], S.M. Shakhno[2], H.P. Yarmola[2]
[1] KTH Royal Institute of Technology, Sweden
[2] Ivan Franko National University of Lviv, Ukraine

1 Introduction

Let us consider the nonlinear least squares problem [2]:

$$\min_{x\in\mathbb{R}^n} \frac{1}{2}(F(x)+G(x))^T(F(x)+G(x)), \tag{1}$$

where the residual function $F+G$ is defined on $\mathbb{R}^n$ with its values on $\mathbb{R}^m$ and it is nonlinear by x; F is a continuously differentiable function; G is a continuous function, differentiability of which in general is not required.

We propose a two-step iterative method, for solving the problem (1), which considers the decomposition of the nonlinear operator, as follows

$$\begin{cases} x_{k+1} = x_k - [A_k^T A_k]^{-1} A_k^T (F(x_k) + G(x_k)), \\ y_{k+1} = x_{k+1} - [A_k^T A_k]^{-1} A_k^T (F(x_{k+1}) + G(x_{k+1})), \quad k = 0, 1, \dots, \end{cases} \tag{2}$$

where $A_k = F'((x_k + y_k)/2) + G(x_k, y_k)$; $F'(x_k)$ is a Fréchet derivative of $F(x)$; $G(x_k, y_k)$ is the divided difference of the first-order of the function $G(x)$ at points x_k, y_k; x_0, y_0 are given starting points. In case of $m = n$, the problem (1) converges to solving a system of n nonlinear equations with n unknown and the method (2) to the two-step method [4]. In case when $G(x) = 0$, we obtain the two-step modification of the Gauss-Newton method [1]; when $F(x) = 0$, we obtain the two-step method with divided differences [3].

We investigate the convergence of this method under the classical Lipschitz condition for the first- and second-order derivatives of the differentiable part and for the first-order divided differences of the nondifferentiable part of the decomposition. The convergence order as well as the convergence radius of the method are studied and the uniqueness ball of the solution of the nonlinear least squares problem is examined.

2 Convergence analysis of the method (2)

Theorem 1. *Let $F+G : \mathbb{R}^n \to \mathbb{R}^m$, $m \geq n$, be continuous, where F is a twice Fréchet differentiable operator and G is a continuous operator on a subset $D \subseteq \mathbb{R}^n$. Assume that the problem (1) has a solution $x_* \in D$ and an operator $F'(x_*)+G(x_*, x_*)$ has full rank. Suppose that Fréchet derivatives $F'(x)$ and $F''(x)$ satisfy the Lipschitz conditions on $B(x_*, r) = \{x \in D : \|x - x_*\| \leq R\}$*

$$\begin{aligned} \|F'(x) - F'(y)\| &\leq L\|x - y\|, \\ \|F''(x) - F''(y)\| &\leq N\|x - y\|, \end{aligned} \tag{3}$$

and the function G has the first-order divided difference $G(x, y)$ and

$$\|G(x, y) - G(u, v)\| \leq M(\|x - u\| + \|y - v\|) \tag{4}$$

for all $x, y, u, v \in D$; L, N, and M are non-negative numbers.

Also, the radius $r > 0$ is a root of the equation

$$\beta N p^2 + 120\beta T p + 48\sqrt{2}\alpha\beta^2 T - 24 = 0,$$

where $2\sqrt{2}\alpha\beta^2 T < 1$.

Then, for all $x_0, y_0 \in B(x_*, r)$ the sequences $\{x_k\}$ and $\{y_k\}$, which are generated by the method (2), are well defined, remain in $B(x_*, r)$ for all $k \geq 0$, and converge to x_* such that

$$\begin{aligned} \rho(x_{k+1}) &\leq \frac{\beta}{1 - \beta T \tau_k}\big((N/24)\rho(x_k)^3 + T\rho(x_k)\rho(y_k) + \sqrt{2}\alpha\beta T \tau_k\big), \\ \rho(y_{k+1}) &\leq \frac{\beta}{1 - \beta T \tau_k}\big((N/24)\rho(x_k)^3 + T(\rho(x_{k+1}) + \rho(x_k) + \rho(y_k)) \times \\ &\times \rho(x_{k+1}) + \sqrt{2}\alpha\beta T \tau_k\big), \\ r_{k+1} &= \max\{\rho(x_{k+1}), \rho(y_{k+1})\} \leq q r_k \leq \cdots \leq q^{k+1} r_0, \end{aligned}$$

where

$$0 < q = \frac{\beta\big((N/24)\rho(x_0)^2 + T(2\rho(x_0) + \rho(y_0)) + \sqrt{2}\alpha\beta T \tau_0 / r_0\big)}{1 - \beta T \tau_0} < 1,$$

$\rho(x) = \|x - x_*\|$, $\tau_k = \|x_k - x_*\| + \|y_k - x_*\|$, $r_0 = \max\{\rho(x_0), \rho(y_0)\}$, $z_k = (x_k + y_k)/2$, $\alpha = \|F(x_*) + G(x_*)\|$, $\beta = \|(A_*^T A_*)^{-1} A_*^T\|$, $A_* = F'(x_*) + G(x_*, x_*)$, $T = \dfrac{L + 2M}{2}$, $\beta T \tau_0 < 1$.

Convergence order of the iterative method (2) in case of zero residual is equal to $1 + \sqrt{2}$.

Theorem 2. *Suppose x_* satisfies (1) and $F(x)$ has a continuous derivative $F'(x)$ and $G(x)$ has a divided difference $G(x, y)$ in $B(x_*, r)$. Moreover, $F'(x_*) + G(x_*, x_*)$ has full rank; $F'(x)$ satisfies the Lipschitz condition as in (3); the divided difference $G(x, y)$ satisfies the Lipschitz condition as in (4). Let $r > 0$ satisfies $\beta(Lr/2 + M) + \alpha\beta_0(L + 2M) \leq 1$, where $\beta_0 = \|(F'(x_*) + G(x_*, x_*))^T (F'(x_*) + G(x_*, x_*))\|$. Then, x_* is a unique solution of the problem (1) in $B(x_*, r)$.*

3 Numerical experiments

We carried out a set experiments on test problems and compared the number of iterations under which the two-step Secant method

$$\begin{cases} x_{k+1} = x_k - [H(x_k, y_k)^T H(x_k, y_k)]^{-1} H(x_k, y_k)^T H(x_k), \\ y_{k+1} = x_{k+1} - [H(x_k, y_k)^T H(x_k, y_k)]^{-1} H(x_k, y_k)^T H(x_{k+1}), \ k = 0, 1, \ldots \end{cases} \tag{5}$$

and the method (2) converge to the solution; $H(x) \equiv F(x) + G(x)$. We used the same initial points for all methods and the following stopping criteria: $\|x_{k+1} - x_k\| \leq \varepsilon$.

Let us denote $h(x) = (H(x))^T H(x)$. Below we list several examples.

Example 1. $n = 3, \quad m = 15;$

$H_i(x_1, x_2, x_3) = x_1 e^{-\frac{x_2(t_i - x_3)^2}{2}} - y_i + (y_i - 1)|x_1^2 - x_3 + t_i x_3 x_2^2 + 1|$,
$t_i = 4 - i/2$, $i = \overline{1, m}$,
$y = (0.0009, 0.0044, 0.0175, 0.0540, 0.1295, 0.2420, 0.3521, 0.3989,$
$0.3521, 0.2420, 0.1295, 0.0540, 0.0175, 0.0044, 0.0009)$,

$$x_* = (1,\ 0,\ 1), \quad h(x_*) = 0.$$

Example 2. $n = 2, \quad m = 8;$

$H_i(x_1, x_2) = 1 - e^{-\left(\frac{t_i}{x_1}\right)^{x_2}} - y_i + 0.01 t_i \left|\dfrac{x_1}{x_2} - x_1\right|$, $i = \overline{1, m}$,
$t = (0.1, 0.5, 0.7, 1.0, 1.2, 1.7, 2.2, 4.5)$,
$y = (0.005, 0.1175, 0.2173, 0.3939, 0.5132, 0.7643, 0.9111, 0.99961)$,

$$x_* \approx (1.439857,\ 1.962064), \quad h(x_*) = 0.001082.$$

In Table 1 we present the amount of iterations spent by each method to compute an approximation to the solution of both examples with the accuracy of $\varepsilon = 10^{-7}$. The additional initial point y_0 we calculated by setting y_0 to $x_0 + 0.0001$. Since methods (2) and (5) for problems with zero residual have convergence order $1+\sqrt{2}$, we obtain the same number of iterations for them. However, the method (2) has advantages for problems with nonzero residual.

Table 1: Number of iterations for solving Examples 1-2.

Example	**The initial approximation** x_0	**Combined method (2)**	**Secant method** (5)
1	(0.7, 0.01, 0.7)	4	4
	(0.6, -0.1, 1.4)	7	8
	(1.4, -0.1, 0.6)	6	6
2	(1.4, 2)	7	10
	(2, 1.3)	38	43
	(1.1, 2.2)	11	34

References

[1] Bartish M. Ia. About one modification of the Gauss-Newton method / M. Ia. Bartish, A. I. Chypurko, S. M. Shakhno // Visnyk of Lviv Univ. Ser. Mech. Math. – 1995. – Vol. 42. – P. 35–38 (in Ukrainian).

[2] Shakhno S. M. An iterative method for solving nonlinear least squares problems with nondifferentiable operator / S. M. Shakhno, R. P. Iakymchuk, H. P Yarmola // Mat. Stud. – 2017. – Vol. 48. – P. 97–107.

[3] Shakhno S. M. On a difference method with superquadratic convergence for solving nonlinear least squares problems / S. M. Shakhno, O. P. Gnatyshyn, R. P. Iakymchuk // Visnyk of Lviv Univ. Ser. Appl. Math. Inform. – 2007. – Vol. 13. – P. 51–58 (in Ukrainian).

[4] Shakhno S. M. On the two-step method for solving nonlinear equations with nondifferentiable operator / S. M. Shakhno, H. P Yarmola // Proc. Appl. Math. Mech. – 2012. – **1**. – P. 617–618.