Nonlinear Least Square Problems with Operator Decomposition

Keywords: least squares problem, decomposition of operator, Gauss-Newton method, Lipschitz conditions, divided difference

1 Introduction

Sometimes the nonlinear function consist of a differentiable and a nondiffirentiable parts. Despite the applicability of iterative-difference methods to such class of problems, it is better to take into account the specificity of given function, which enable to achieve a greater convergence order. One of possible options is using in method the only derivative from differentiable part, and omit the nondiffirentiable part, because full Jacobian does not exist. However, as practice shows, such methods converges much slower. Methods, which are using derivate of differentiable part of operator and divided differences of nondiffirentiable part of operator demonstrate much better results. This approach well recommends itself in solvin nonlinear problems what you can see in I.K. Argyros (2008); E. Catinas (1994) ; ? .

Let us consider the nonlinear least squares problem

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

where the residual function $F+G$ is defined on $\mathbb{R}^p$ 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. If $m=p$, we get system of nonlinear equations (E. Catinas, 1994; ?).

For finding the solution of problem (1), we propose method based on Gauss-Newton method and Potra method (?)

$$\begin{aligned} x_{k+1} &= x_k-\left(A_k^TA_k\right)^{-1}A_k^T\left(F\left(x_k\right)+G\left(x_k\right)\right), k=0,1,2,\ldots \\ A_k &= F'\left(x_k\right)+G\left(x_k,x_{k-1}\right)+G\left(x_{k-2},x_k\right)-G\left(x_{k-2},x_{k-1}\right) \end{aligned} \tag{2}$$

Here $F'\left(x_k\right)$ is Fréchet derivative $F\left(x\right)$ in point $x_k, G\left(x_k,x_{k-1}\right), G\left(x_{k-2},x_k\right), G\left(x_{k-2},x_{k-1}\right)$ are divided differences of first order of function $G\left(x\right)$ at appropriate points (?); x_0, x_{-1}, x_{-2} are given initial approximations. In case when $m=n$, this method reduces to the Newton-Potra methods (?).

In this work, we study the local convergence of the proposed method and show efficiency of this method in comparing with other methods for solving such class of problems.

2 Convergence Analysis

Let us denote $\Omega\left(x^*,r\right)=\{x\in D\subseteq\mathbb{R}^p:\|x-x^*\|<r\}$ as an open ball with the radius r $(r>0)$ at x^*, D is an open convex subset of $\mathbb{R}^p$.

Sufficient conditions of the local convergence of the iterative process (2) are given in the following theorem.

Theorem 1. *Let $F+G:\mathbb{R}^p\longrightarrow\mathbb{R}^m, m\geq p$, be continuous operator, where F is a Fréchet differentiable operator and G is a continuius operator on a subset $D\subseteq\mathbb{R}^n$. Assume that the problem (1) has a solution $x^*\in D$ an exist an inverse operator $\left(A_*^TA_*\right)^{-1}$, where $A_*=F'\left(x^*\right)+G\left(x^*,x^*\right)$, and $\|\left(A_*^TA_*\right)^{-1}\|\leq B$. Suppose that Fréchet derivate $F'\left(x\right)$ satisfies the Lipschitz conditions on D*

$$\|F'\left(x\right)+F'\left(y\right)\|\leq L\|x-y\| \tag{3}$$

and the function G has the first and second order divided differences $G\left(\cdot,\cdot\right)$ and $G\left(\cdot,\cdot,\cdot\right)$ and

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

$$\|G(u,x,y)-G(v,x,y)\| \le N(\|u-v\|) \tag{5}$$

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

Furthermore,

$$\begin{aligned} \|F(x^*)+G(x^*)\| \le \eta, \|F'(x^*)+G(x^*,x^*)\| \le \alpha \\ B(L+2M)\eta \le 1 \end{aligned} \tag{6}$$

and $\Omega = \Omega(x^*, r_*) \subseteq D$, where the radius $r_* > 0$ is a unique root of the equation

$$\begin{aligned} q(r) = B\left[\left(\alpha+(L+2M)r+2Nr^2\right)\left(\left(\frac{1}{2}L+2M\right)r+4Nr^2\right)+(L+2M+2Nr)\eta\right]+ \\ B\left[2\alpha+(L+2M)r+2Nr^2\right]\left((L+2M)r+2Nr^2\right)-1=0 \end{aligned} \tag{7}$$

Then, for all $x_0, x_{-1}, x_{-2} \in \Omega$ the sequence x_k, which are generated by the method (2), is well defined, remain in Ω for all $k \ge 0$, and converges to x^*. Moreover, the following estimates hold for all $k \ge 0$

$$\begin{aligned} &\|x_{k+1}-x^*\| \le g(r_*)\left[\alpha+(L+2M)\|x_k-x^*\|+2N\|x_{k-2}-x^*\|\|x_{k-1}-x^*\|\right]\cdot \\ &\cdot\left[\left(\frac{1}{2}L+M\right)\|x_k-x^*\|+4N\|x_{k-1}-x^*\|\|x_{k-2}-x^*\|\right]\|x_k-x^*\|+ \\ &+\eta\left[(L+2M)\|x_k-x^*\|+2N\|x_{k-2}-x^*\|\|x_{k-1}-x^*\|\right]=C_1\|x_k-x^*\|+C_2\|x_k-x^*\|\|x_{k-1}-x^*\|+ \\ &+C_3\|x_{k-1}-x^*\|\|x_{k-2}-x^*\|+C_4\|x_k-x^*\|^2+C_5\|x_k-x^*\|\|x_{k-1}-x^*\|\|x_{k-2}-x^*\|+ \\ &+C_6\|x_k-x^*\|^3+C_7\|x_k-x^*\|^2\|x_{k-1}-x^*\|\|x_{k-2}-x^*\|+C_8\|x_k-x^*\|^3\|x_{k-1}-x^*\|+ \\ &+C_9\|x_k-x^*\|^2\|x_{k-1}-x^*\|^2\|x_{k-2}-x^*\| \\ &+C_{10}\|x_k-x^*\|^2\|x_{k-1}-x^*\|\|x_{k-2}-x^*\|+C_{11}\|x_k-x^*\|\|x_{k-1}-x^*\|^2\|x_{k-2}-x^*\|^2 \end{aligned} \tag{8}$$

where

$$\begin{aligned} g(r)=B\left[1-B\left[2\alpha+(L+2M)r+2Nr^2\right]\cdot\left[(L+2M)r+2Nr^2\right]\right]^{-1}; C_1=g(r_*)\eta(L+2M); \\ C_2=C_3=g(r_*)\eta N; C_4=g(r_*)\left(\frac{\alpha L}{2}+\alpha M\right); C_5=g(r_*)\alpha 4N; C_6=g(r_*)\frac{1}{2}(L+2M)^2; \\ C_7=g(r_*)4N(L+2M); C_8=C_{10}=g(r_*)\left(\frac{NL}{2}+NM\right); C_9=C_{11}=g(r_*)4N^2 \end{aligned} \tag{9}$$

3 Numerical experiments

Example. $3x^2y+y^2-1+\|x-1\|=0,$

$x^4+xy^3+\|y\|=0,$

$\|x^2-y\|=0,$

$(x^*,y^*) \approx (0.74862800, 0.43039151), f(x^*) \approx 4.0469349 \cdot 10^{-2}$

Numerical solution of the problem we get with the accuracy $\varepsilon = 10^{-8}$:

$\|x_{k+1}-x_k\| \le \varepsilon$

The additional initial points we calculated by

$$(x_{-1},y_{-1}) = \left(x_0-10^{-4}, y_0-10^{-4}\right),$$
$$(x_{-2},y_{-2}) = \left(x_0-2\cdot 10^{-4}, y_0-2\cdot 10^{-4}\right)$$

Example	$(x_0; y_0)$	Method	
		(1)	(2)
1	(1, 0.5)	5	5
	(5, 2.5)	14	11
	(10, 5)	19	14
2	(0.6, 0.4)	14	14
	(3, 2)	21	19
	(6, 4)	25	21

Table 1: Number of iterations for solving test problems.

From obtained results, we can see that the combined Gauss-Newtone-Potra method (2) is more efficient than basic method.

References

[1] I.K. Argyros, 2008. Convergence and Applications of Newton-type Iterations. Springer-Verlag: New York, 506 p. *a document preparation system.*

[2] I.K. Argyros, H. Ren, 2011. A derivative free iterative method method for solving least squares problems. *Numerical Algorithms*, 4(58), 555-571.

[3] E. Catinas, 1994.On some iterative methods for solving nonlinear equations. *Revue d‘Analyse Numerique et de Theorie de l‘Approximation*, 1(23), 47-53.

[4] J.Chen, W.Li, 2005. Convergence of Gauss-Newton method‘s and uniqueness of the solution. *Applied Mathematics and Computation*, 170, 686-705.