



HAL
open science

On the Convergence Analysis of a Two-Step Modification of the Gauss-Newton Method

Roman Iakymchuk, Stepan Shakhno

► **To cite this version:**

Roman Iakymchuk, Stepan Shakhno. On the Convergence Analysis of a Two-Step Modification of the Gauss-Newton Method. 2014. hal-00960300v1

HAL Id: hal-00960300

<https://hal.science/hal-00960300v1>

Preprint submitted on 17 Mar 2014 (v1), last revised 30 May 2014 (v2)

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

On the Convergence Analysis of a Two-Step Modification of the Gauss-Newton Method

Roman Iakymchuk^{1,2*} and Stepan Shakhno^{3**}

¹ Sorbonne Universits, UPMC Univ Paris 06, ICS, F-75005 Paris, France

² Sorbonne Universits, UPMC Univ Paris 06, UMR 7606, LIP6, F-75005 Paris, France

³ Ivan Franko National University of Lviv, Universitetska Str. 1, 79000 Lviv, Ukraine

We investigate convergence of the two-step modification of Gauss-Newton method applying the generalized Lipschitz condition for the first and second order derivatives. The convergence order as well as the convergence radius of the method are studied and the uniqueness ball of solution of the nonlinear least squares problem is proved. Finally, we carry out numerical experiments on a set of well-known functions.

Copyright line will be provided by the publisher

1 Introduction

Let us consider the nonlinear least squares problem:

$$\min f(x) := \frac{1}{2} F(x)^T F(x), \tag{1}$$

where F is a Fréchet differentiable operator defined on R^n with its values on R^m , $m \geq n$.

For solving the problem (1), we consider a two-step modification of the Gauss-Newton method

$$\begin{cases} x_{n+1} = x_n - [F'(z_n)^T F'(z_n)]^{-1} F'(z_n)^T F(x_n), \\ y_{n+1} = x_{n+1} - [F'(z_n)^T F'(z_n)]^{-1} F'(z_n)^T F(x_{n+1}), \end{cases} \quad n = 0, 1, 2, \dots \tag{2}$$

where $z_n = (x_n + y_n)/2$; $x_0 = y_0$ are given. This method was initially proposed, but in a different form, by Bartish et al. [2]. The advantage of the method is that it uses only one inverse over two function evaluations.

The main focus of our study is to analyze the convergence of the method (2). The local convergence of the same was examined in [1] using the classical Lipschitz condition for the second order derivatives. Instead, we study the convergence of the above-mentioned method using the Lipschitz condition with L average [4] for derivatives; such conditions are called the generalized Lipschitz conditions, where L is not a constant, but an integrable function. The convergence analysis of the method (2) using the Lipschitz condition with L average provides the following advantages over the corresponding results in [1]: weaker convergence conditions, finer estimates on the distances involved, and larger radius of convergence.

2 Convergence Analysis and the Uniqueness of Solution

In this section, we investigate the convergence and the radius of convergence of the method (2) as well as prove the uniqueness of solution.

Theorem 2.1 *Let $F : R^n \rightarrow R^m$, $m \geq n$, be a Fréchet differentiable operator in $D \in R^n$. Assume that (1) has a solution x_* in $B(x_*, r) = \{x \in D : \|x - x_*\| < r\}$ and there exists a Fréchet derivative $F'(x_*)$ that has full column rank. Suppose that $F(x)$ has Fréchet derivatives $F'(x)$ and $F''(x)$ on $B(x_*, r)$ that satisfy the Lipschitz condition with L and N average:*

$$\|F'(x) - F'(x_*)\| \leq \int_0^{\rho(x)} L(u) du, \quad \|F''(x) - F''(y)\| \leq \int_0^{\|x-y\|} N(u) du,$$

where $x, y \in B(x_*, r)$, $\rho(x) = \|x - x_*\|$; $L(u)$ and $N(u)$ are nondecreasing functions. Also, the radius $r > 0$ satisfies

$$\frac{\beta \int_0^r L(u) du}{1 - \beta \int_0^{2r} L(u) du} + \frac{\sqrt{2}\alpha\beta^2 \int_0^{2r} L(u) du}{r(1 - \beta \int_0^{2r} L(u) du)} \leq 1.$$

* Corresponding author: e-mail: roman.iakymchuk@lip6.fr, Phone: +33 (0)1 44 27 87 73

** e-mail: s.shakhno@franko.lviv.ua, Phone: +38(0)3 22 39 43 91

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

$$\begin{aligned}\rho(x_{k+1}) &= \|x_{k+1} - x_*\| \leq \gamma \rho(x_k)^3 + \eta \rho(x_k) \rho(y_k) + \theta [\rho(x_k) + \rho(y_k)]/2; \\ \rho(y_{k+1}) &= \|y_{k+1} - x_*\| \leq \gamma \rho(x_{k+1})^3 + \zeta \rho(x_{k+1})(\rho(x_k) + \rho(y_k) + \rho(x_{k+1})) + \theta [\rho(x_k) + \rho(y_k)]/2; \\ r_{k+1} &= \max\{\|x_{k+1} - x_*\|, \|y_{k+1} - x_*\|\} \leq q r_k \leq \dots \leq q^{k+1} r_0, \quad q = \max\{\gamma r^2 + \eta r + \theta, \gamma r^2 + 3\zeta r + \theta\} < 1,\end{aligned}$$

where

$$\begin{aligned}\gamma &= \frac{\beta \int_0^{\rho(x_0)} N(u)(\rho(x_0) - u)^2 du}{8 \left(1 - \beta \int_0^{\rho(x_0)} L(u) du\right) \rho(x_0)^3}; \eta = \frac{\beta \int_0^{\rho(y_0)/2} L(u) du}{\left(1 - \beta \int_0^{\rho(x_0)} L(u) du\right) \rho(y_0)/2}; \zeta = \frac{\beta \int_0^{z_0} L(u) du}{\left(1 - \beta \int_0^{\rho(x_0)} L(u) du\right) 2z_0}; \\ \theta &= \frac{\sqrt{2}\alpha\beta^2 \int_0^{\rho(x_0)} L(u) du}{\left(1 - \beta \int_0^{\rho(x_0)} L(u) du\right) \rho(x_0)}; z_0 = (2\rho(x_0) + \rho(x_1))/2, \alpha = \|F'(x_*)\|, \beta = \|(F'(x_*)^T F'(x_*))^{-1} F'(x_*)^T\|.\end{aligned}$$

Corollary 2.2 Order of convergence of the iterative process (2) in case of zero residual is equal $1 + \sqrt{2}$.

Theorem 2.3 Suppose x_* satisfies (1), $F(x)$ has a continuous derivative $F'(x)$ in $B(x_*, r)$. Besides that, $F'(x_*)$ has full rank and $F'(x)$ satisfies Lipschitz condition with L average:

$$\|F'(x) - F'(x_*)\| \leq \int_0^{\rho(x)} L(u) du,$$

where $x \in B(x_*, r)$, $\rho(x) = \|x - x_*\|$, and $L(u)$ is a nondecreasing function. Let $r > 0$ satisfies

$$\beta \int_0^r L(u)(r - u) du + \alpha \beta_0 \int_0^r L(u) du \leq r,$$

where α and β are defined in (3) and $\beta_0 = \|[F'(x_*)^T F'(x_*)]^{-1}\|$. Then x_* is a unique solution of the problem (1) in $B(x_*, r)$.

3 Numerical Results

We carried out a set of well-known tests, comparing the number of iterations under which Gauss-Newton method and its one-step modification converge to the exact solution, see Tab. 1. We used the same initial points for all methods but different stopping conditions such as $\|F'^T(x_k)F(x_k)\| \leq 10^{-8}$ or $\|A_k^T F(x_k)\| \leq 10^{-8}$, when the divided differences were applied. To conclude, the considered method (2) is efficient as it does not use analytically given derivatives, but rather an approximation of the Jacobian with the divided differences of the first order; the latter does not require additional computations to be performed.

Table 1: The number of iterations to solution.

Example	Gauss-Newton m-d	Secant m-d [3]	Two-step Gauss-Newton m-d
Rosenbrock func.	3	3	2
Kowalik-Osborne func.	10	17	10
Box-3D func.	6	8	5
Gnedenko-Veibull dist.	6	8	4
Freidenstein-Ross func.	44	19	8
Wood func.	51	74	49

Acknowledgements This work undertaken (partially) in the framework of CALSIMLAB is supported by the public grant ANR-11-LABX-0037-01 overseen by the French National Research Agency (ANR) as part of the ‘‘Investissements d’Avenir’’ program (reference: ANR-11-IDEX-0004-02).

References

- [1] M. Ia. Bartish and A. I. Chyurko. *Visnyk of Lviv Univ. Ser. Appl. Math. and Infor.*, 1:3–7, 1999.
- [2] M. Ia. Bartish, A. I. Chyurko, and S. M. Shakhno. *Visnyk of Lviv Univ. Ser. Mech. Math.*, 42:35–38, 1995.
- [3] S. M. Shakhno and O. P. Gnatyshyn. *Applied Mathematics and Computation*, 161:253–264, 2005.
- [4] X. Wang. *IMA Journal of Numerical Analysis*, 20:123–134, 2000.