UNIVERSITY OF NIS FACULTY OF MECHANICAL ENGINEERING

THE TENTH INTERNATIONAL CONFERENCE
TRANSPORT AND LOGISTICS
https://doi.org/10.46793/TIL2025.248R

til 2025

GAUSS-NEWTON OPTIMIZATION METHOD AND ITS APPLICATIONS

***Predrag RAJKOVIĆ*[1], *Slađana MARINKOVIĆ*[2], *Dragan KOSTIĆ*[3], *Jovan RAJKOVIĆ*[2]**

[1] Fac. Mechanical Engineering, University of Niš, Serbia
[2] Fac. Electronic Engineering, University of Niš, Serbia
[3] Eindhoven University of Technology, Netherlands

ORCID iDs: Predrag RAJKOVIĆ https://orcid.org/0000-0002-2914-0985
Slađana MARINKOVIĆ https://orcid.org/0000-0002-7752-4393
Dragan KOSTIĆ https://
Jovan RAJKOVIĆ https://orcid.org/0009-0003-6282-5542

Abstract

The Gauss-Newton method (G-N method) is a highly efficient iterative method used to solve nonlinear least-squares problems. It can be seen as a modification of the Newton method for finding the minimum value of a function.
For G-N method, we will discuss its convergence, further modifications and applications.
Finally, we present and analyze the efficiancy of this method through numerical experiments: In a biological experiment, which studies the relationship between the concentration of a substrate and the rate of an enzyme-mediated reaction; prognostic score in temperature of a city by the measured average high temperatures per month in a year; approximating and prognosis average annual daily traffic through state roads and so on.

***Keywords**: unconstrained optimization, nonlinear least-squares, Newton method, fitting.*

1 INTRODUCTION

Most optimization problems lead to the minimization of a function with respect to some parameters. This corresponds to a nonlinear programming problem, the solution of which can be obtained only by iterative methods. Nonlinear least-squares approximations are an incredibly useful tool for analyzing sets of data and determining whether a theoretical model is a good fit for data obtained experimentally.
There are plenty of applications for an iterative process for solving a nonlinear least-squares approximation problem.
It can be used as a method of locating a single point or as a way of determining how well a theoretical model fits a set of experimental data points. By solving the system of nonlinear equations, we obtain the best estimates of the unknown variables in a theoretical model. Then we are able to plot this function along with the data points and see how well these data points fit the theoretical equation.
A wide variety of applications with this formulation can be found in mathematical programming literature, e.g., nonlinear inclusions, penalization methods, minimax, and goal programming. Besides its practical applications, this model provides a convenient tool for the study of first and second order optimality conditions in constrained optimization. Optimisation methods are the current workhorse for training neural networks.

2 NEWTON OPTIMIZATION METHOD

Let $x=(x_1,x_2,\dots,x_n)$ be a point in the real space R^n and $\vec{x}$ the corresponding vector.
Notice a strictly convex differentiable function $f: R^n \to \mathrm{R}$ which achieves its minimum in the point x^*:

$$\min_{x\in R^n} f(x) = f(x^*). \tag{1}$$

Let $x^{(k)} \in R^n$ be some point and $f_k = f\left(x^{(k)}\right)$.

From the Taylor series of $f(x)$ around the point $x^{(k)}$, we can extract quadratic approximation $q(x)$ as

$$f(x) \approx q(x) = f_k + (\nabla f_k)^T\left(\vec{x}-\vec{x}^{(k)}\right) + \frac{1}{2}\left(x-x^{(k)}\right)\nabla^2 f_k\left(\vec{x}-\vec{x}^{(k)}\right),$$

where the gradient vector and Hessian matrix are:

$$\nabla f(x) = \begin{bmatrix} \dfrac{\partial f}{\partial x_1} \\ \vdots \\ \dfrac{\partial f}{\partial x_n} \end{bmatrix}, \quad \nabla^2 f(x) = \begin{bmatrix} \dfrac{\partial^2 f}{\partial x_1^2} & \cdots & \dfrac{\partial^2 f}{\partial x_1 \partial x_n} \\ \vdots & & \\ \dfrac{\partial^2 f}{\partial x_n \partial x_1} & & \dfrac{\partial^2 f}{\partial x_n^2} \end{bmatrix}.$$

If $q(x)$ has its local minimum in the point $x^{(k+1)}$, then its gradient is equal to zero, i.e.

$$\nabla q\left(x^{(k+1)}\right) = \nabla f_k + \nabla^2 f_k\left(\vec{x}^{(k+1)} - \vec{x}^{(k)}\right) = 0.$$

Supposing that $\left(\nabla^2 f_k\right)^{-1} = \left(\nabla^2 f\left(x^{(k)}\right)\right)^{-1}$ exists, we can define sequence of iterations

$$\vec{x}^{(k+1)} = \vec{x}^{(k)} - \left(\nabla^2 f_k\right)^{-1}\nabla f_k \qquad (k\in N). \tag{2}$$

This is the *Newton optimization method*.

Theorem 1. *Let $D \subseteq R^n$ and $f(x) \in C^2(D)$ and $x^* \in R^n$ is such that $\nabla f(x^*) = 0$ and $\nabla^2 f(x^*)$ is invertible. For any $x^{(0)} \in D$ such that*

(1) $\det \nabla^2 f\left(x^{(0)}\right) \neq 0$;

(2) $\left\|\left(\nabla^2 f\left(x^{(0)}\right)\right)^{-1}\left(\nabla^2 f(y) - \nabla^2 f(x)\right)\right\| \le L\|x-y\| \qquad (L>0; \forall x,y\in D);$

(3) $\left\|\left(\nabla^2 f\left(x^{(0)}\right)\right)^{-1} \nabla f\left(x^{(0)}\right)\right\| \le b \quad (b>0);$

(4) $2bL \le 1$;

(5) $\varepsilon^* = \frac{1-\sqrt{1-2bL}}{L} \wedge B\left(x^{(0)},\varepsilon^*\right) \subseteq D,$

the sequence $\{x^{(k)}\}$ *generated by (2) is contained in* $B(x^{(0)}, \varepsilon^*)$ *and converges to the unique solution* x^*, *i.e.* $\lim_{k\to\infty} x^{(k)} = x^*$. *The following estimation is valid*

$$\|x^{(k+1)} - x^*\| \leq \frac{1}{2}\|x^{(k)} - x^*\|,$$

and the convergence rate is quadratic:

$$\lim_{k\to\infty} \frac{\|x^{(k+1)} - x^*\|}{\|x^{(k)} - x^*\|^2} = C = \text{const} \notin \{0, \infty\}.$$

Newton's method has the serious drawbacks:
(i) it is not globally convergent toward stationary points;
(ii) The search direction is not defined at points where the Hessian matrix is singular;
(iii) It requires finding the inverse matrix at each step and this may become ill-conditioned for large dimensional problems.

3 GAUSS-NEWTON OPTIMIZATION METHOD

This method is applied to the class of functions that are sums of squares of real functions and it is based on the Hessian approximation in Newton's method.

We want to minimize the function $f: R^n \to \mathrm{R}$ of the form

$$f(x) = \sum_{i=1}^{m} r_i^2(x) \qquad (n \leq m), \tag{3}$$

where all functions $r_i(x)$ are real strictly convex differentiable functions.
Let $\vec{r}(x)$ be the vector of these functions and $\mathrm{J}(x)$ their Jacobian:

$$\vec{r}(x) = \begin{bmatrix} r_1(x) \\ \vdots \\ r_m(x) \end{bmatrix}, \qquad \mathrm{J}(x) = \left[\frac{\partial r_i}{\partial x_j}\right]_{m\times n} = \begin{bmatrix} \frac{\partial r_1}{\partial x_1} & \cdots & \frac{\partial r_1}{\partial x_n} \\ \vdots & & \\ \frac{\partial r_m}{\partial x_1} & & \frac{\partial r_m}{\partial x_n} \end{bmatrix}.$$

Since

$$\frac{\partial f}{\partial x_k} = \sum_{i=1}^{m} 2r_i(x)\frac{\partial r_i}{\partial x_k} \quad (k = 1, \dots, n), \tag{4}$$

it is

$$\nabla f(x) = 2(\mathrm{J}(x))^T \vec{r}(x). \tag{5}$$

By differentiating (4), we get

$$\frac{\partial^2 f}{\partial x_l \partial x_k} = 2\sum_{i=1}^{m} \frac{\partial}{\partial x_l}\left(r_i(x)\frac{\partial r_i}{\partial x_k}\right),$$

i.e.,

$$\frac{\partial^2 f}{\partial x_l \partial x_k} = 2\sum_{i=1}^{m}\left(\frac{\partial r_i}{\partial x_l}\frac{\partial r_i}{\partial x_k} + r_i(x)\frac{\partial^2 r_i}{\partial x_l \partial x_k}\right).$$

Therefore, the Hessian $\nabla^2 f(x)$ can be written in the form

$$\nabla^2 f(x) = 2(\mathrm{J}(x))^T \cdot \mathrm{J}(x) + Q(x), \tag{6}$$

with

$$Q(x) = 2\sum_{i=1}^{m}\left(r_i(x) \cdot \nabla^2 r_i(x)\right).$$

If $r_i(x)$ and $\frac{\partial^2 r_i}{\partial x_l \partial x_k}$ are negligibly small with respect to $\frac{\partial r_i}{\partial x_l}\frac{\partial r_i}{\partial x_k}$ when x is near the minimum point x^* of $f(x)$, we can approximate $\frac{\partial^2 f}{\partial x_l \partial x_k}$ with

$$\frac{\partial^2 f}{\partial x_l \partial x_k} \approx 2\sum_{i=1}^{m}\left(\frac{\partial r_i}{\partial x_l}\frac{\partial r_i}{\partial x_k}\right).$$

Hence, we have approximation for $\nabla^2 f(x)$ in the form

$$\nabla^2 f(x) \approx 2(\mathrm{J}(x))^T \mathrm{J}(x). \tag{7}$$

Remark. This situation occurs when

$$|r_i(x)| \, \|\nabla^2 r_i(x)\| \ll \lambda_{\max}\left((J(x))^T \mathrm{J}(x)\right),$$

where $\|\nabla^2 r_i(x)\|$ denotes the spectral norm $\nabla^2 r_i(x)$, i.e.

$$\|\nabla^2 r_i(x)\| = \sqrt{\lambda_{\max}\left(\nabla^2 r_i(x)^T \nabla^2 r_i(x)\right)},$$

and $\lambda_{\max}$ is the maximal eigenvalue of the corresponding matrices (see Nocedal-Wright [7]).

Let

$$J_k = \mathrm{J}(x^{(k)}).$$

Including (4) and (5) into (1) with assumption

$$\det[(J_k)^T J_k] \neq 0,$$

we yield

$$\vec{x}^{(k+1)} \approx \vec{x}^{(k)} - \left(2J_k^{\mathrm{T}} J_k\right)^{-1} 2J_k^{\mathrm{T}} \vec{r}(x^{(k)}).$$

Now, Gauss-Newton optimization method is

$$\vec{x}^{(k+1)} = \vec{x}^{(k)} - \left(J_k^{\mathrm{T}} J_k\right)^{-1}(J_k)^T \vec{r}(x^{(k)}). \tag{7}$$

Including the Moore-Penrose inverse (see Wedin [10]))

$$J_k^{\dagger} = \left(J_k^{\mathrm{T}} J_k\right)^{-1}(J_k)^T,$$

it can be written in the final simple form

$$\vec{x}^{(k+1)} = \vec{x}^{(k)} - J_k^{\dagger}\vec{r}(x^{(k)}) \tag{8}$$

$$(k = 1, 2, \dots).$$

4 CONVERGENCE, ERROR ESTIMATION AND MODIFICATIONS

Gauss-Newton method (G-N method) was constructed by omitting some members in the Newton method. Since the Newton method has the order of convergence $r_N = 2$, G-N method will have $r_{G-N} \in (1,2)$. But, if $\|Q\|$ is small when $x^{(k)}$ tends to x^*, we can expect acceleration of convergence.
Let us introduce 3 conditions:
(A.1) There is $x^* \in R^n$ such that $J^{\mathrm{T}}(x^*) f(x^*) = 0$;
(A.2) Jacobian matrix $\mathrm{J}(x^*)$ has full rank n;
(A.3) Starting from some $x^{(k)}$, the matrix $\left(J_k^{\mathrm{T}} J_k\right)^{-1} Q(x^{(k)})$ has the spectral radius $\rho < 1$.

Theorem 2. (Wedin [11]) *If the conditions* (A.1), (A.2) *and* (A.3) *are fulfilled, the Gauss-Newton method converges locally to the optimal point* x^* *for every initial point close enough to it.*

The error of $\vec{x}^{(k)}$ can be estimated by the formula

$$\|x^{(k)} - x^*\| \approx \|\nabla f(x^{(k)})\| = \|2J_k^{\mathrm{T}} \vec{r}(x^{(k)})\|. \qquad (9)$$

There are a few problems in applications of previous methods, such as: the computational cost, execution time, efficiency and accuracy. The previous methods can fail if a singular matrix appears. That is why a few modifications were developed: the modified G-N method, truncated G-N method, inexact G-N method (see Gratton [2], Dixon [3], Grippo [4] and Hedarta [5]), q-G-N method (see Marinković&al [6] , Rajković&al [8], Ram&al [9]), and so on.

The *modified G-N method,* in the case that

$$A_k = J_k^{\mathrm{T}} J_k: \quad \det A_k = 0,$$

includes

$$\tilde{A}_k = A_k + E_k,$$

where the matrix $E_k = E(x^{(k)})$ is a positive semidefinite matrix such that $\tilde{A}_k$ is positive definite.

For example, in a variant of this method, if exact solution is not achieved, i.e. $f(x^{(k)}) \neq 0$, the matrix E_k is the diagonal matrix with diagonal elements $\|J_k^{\mathrm{T}} J_k\|$.

The q-Gauss-Newton algorithm is a modification of the Gauss-Newton method based on interchange of the classical by the q-derivative

$$\frac{\partial f}{\partial x} \approx D_q f(x) = \frac{f(x) - f(qx)}{(1-q)x},$$

where q is a given parameter $q \in (0,1)$.

By approximation $J_k \approx J_{q,k}$, we get q-G-N method

$$\vec{x}^{(k+1)} = \vec{x}^{(k)} - \left(J_{q,k}^{\mathrm{T}} J_{q,k}\right)^{-1} \left(J_{q,k}\right)^T \vec{r}(x^{(k)}).$$

The convergence of q-G-N method is assured only when the initial guess is close enough to the point of the minimum (see It is very useful for the functions which do not have continuous derivatives, or they are too complicated for the computation.

5 APPLICATIONS

Example 1. In a biological experiment, which studies the relationship between the concentration of a substrate T and the rate of an enzyme-mediated reaction Y, Table 1 was obtained.

Table 1

i	T	Y
1	0.0380	0.0500
2	0.1940	0.1270
3	0.4250	0.0940
4	0.6260	0.2122
5	1.2530	0.2729
6	2.5000	0.2665
7	3.7400	0.3317

The experiment can be modeled by a curve

$$y(t) = \frac{x_1 t}{x_2 + t},$$

with unknown parameters x_1 and x_2. Let the initial values be $x_1 = 0.9$ and $x_2 = 0.2$ and stopping criteria $\|\nabla f\| \leq 10^{-9}$. Here, we must minimize the sum of squared functions

$$r_i(x_1, x_2) = y_i - \frac{x_1 t_i}{x_2 + t_i} \qquad (i = 1, \dots, 7),$$

i.e., to find

$$\min f(x_1, x_2) = \sum_{i=1}^{7} r_i^2(x_1, x_2).$$

In order to apply Gauss-Newton method, we determine

$$\frac{\partial r_i}{\partial x_1} = \frac{-t_i}{x_2 + t_i}, \qquad \frac{\partial r_i}{\partial x_2} = \frac{x_1 t_i}{(x_2 + t_i)^2} \quad (i = 1, \dots, 7),$$

and

$$\vec{r}(x) = \begin{bmatrix} r_1 \\ \vdots \\ r_7 \end{bmatrix}, \quad \mathrm{J}(x) = \left[\frac{\partial r_i}{\partial x_j}\right]_{7\times 2} = \begin{bmatrix} \frac{\partial r_1}{\partial x_1} & \frac{\partial r_1}{\partial x_2} \\ \vdots & \vdots \\ \frac{\partial r_7}{\partial x_1} & \frac{\partial r_7}{\partial x_2} \end{bmatrix}.$$

In that manner, we evaluate the iterations

$$\vec{x}^{(0)} = \begin{bmatrix} 0.9 \\ 0.2 \end{bmatrix}, \ \vec{x}^{(1)} = \begin{bmatrix} 0.3327 \\ 0.2602 \end{bmatrix},$$

$$\vec{x}^{(2)} = \begin{bmatrix} 0.3428 \\ 0.4261 \end{bmatrix}, \dots, \ \vec{x}^{(9)} = \begin{bmatrix} 0.36184 \\ 0.55627 \end{bmatrix}.$$

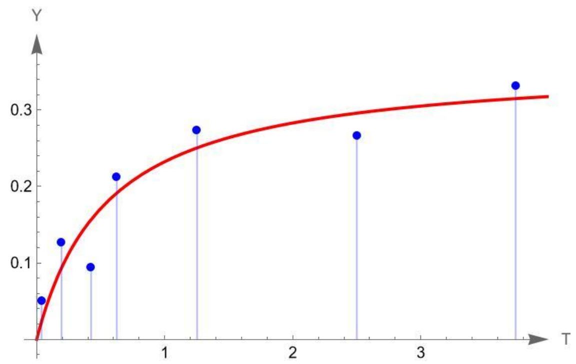

Fig. 1

Let us analyze convergence of the method. Here, the matrix

$$(J_9^{\mathrm{T}} J_9)^{-1} Q(x^{(9)}) = \begin{bmatrix} 0.0915 & -0.0709 \\ 0.52174 & -0.4383 \end{bmatrix},$$

has the eigenvalues $\lambda_1 \approx -0.35548$ and $\lambda_2 \approx 0.008651$.

Hence $\varrho \approx 0.35548 < 1$, what proves the convergence.

Error of the iteration $x^{(9)}$ with respect to the exact solution x^* can be estimated as

$$\|\nabla f(x^{(9)})\| = \|2(J_9)^T \circ \vec{r}(x^{(9)})\| \approx 1.33844 \times 10^{-10}.$$

Example 2. Let $L = \{p_k, q_k\}$ be a given set of $n = 40$ points randomly taken very close to unknown lemniscate. The lemniscate is determined by its center (u, v) and the half-width w. Therefore, consider the functions

$$r_k(u, v, w) = ((p_k - u)^2 + (q_k - v)^2)^2 - w^2\,((p_k - u)^2 - (q_k - v)^2).$$

The optimization problem is

$$\min_{(u,v,w)\in R^3} S(u, v, w) = \sum_{k=1}^{n} r_k^2(u, v, w).$$

For the set of blue points given on Figure 3, the Gauss-Newton method gives

$u \approx 1.05, \ \ v \approx 2.05, \quad w \approx 3.$

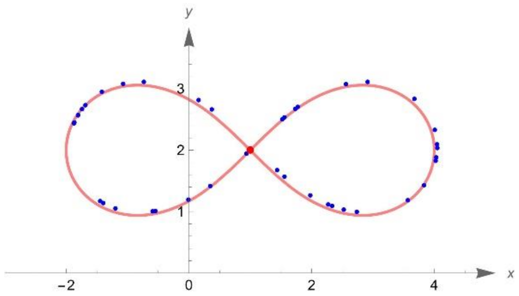

Fig. 2

Example 3. Let $P = \{p_k, q_k\}$ be a given set of $n = 30$ points randomly taken very close to unknown ellipse. We are looking for ellipse with unknown center (u, v) and half-axes a and b. Therefore, consider the functions

$$r_k(u, v, a, b) = \left(\frac{p_k - u}{a}\right)^2 + \left(\frac{q_k - v}{b}\right)^2 - 1.$$

The optimization problem is

$$\min_{(u,v,a,b)\in R^3} S(u, v, a, b) = \sum_{k=1}^{n} r_k^2(u, v, a, b).$$

For the set of blue points given on Figure 3, the Gauss-Newton method gives
$u \approx 1.0794,\ \ v \approx 1.8356,\ a \approx 5.0087,\ b \approx 2.9728$.

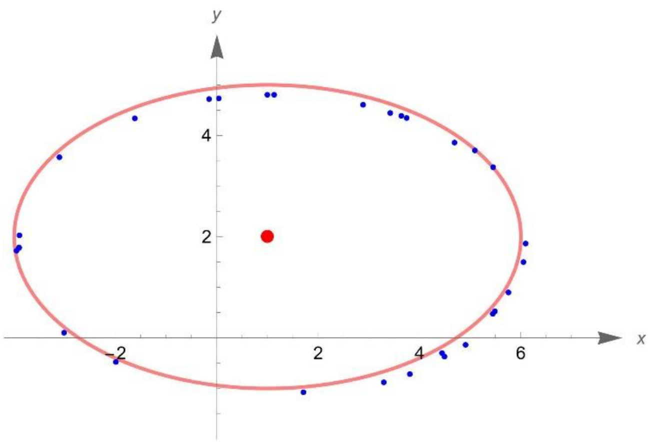

Fig. 3

Example 4. The average high temperatures per month were measured at Fahrenheit in the city of Monroe Louisiana.

Table 2

m	T
1	56
2	60
3	69
4	77
5	84
6	90
7	94
8	94
9	88
10	79
11	68
12	58

It is supposed to have sinusoidal behaviour

$$T = f\ (t; A, w, p, q)\ = A\sin(wt + p) + q.$$

Gauss-Newton method gives

$$A\ \approx 11.05467,\ \ w \approx 0.47580,$$
$$p \approx 10.77126,\ \ \ q \approx 23.61121.$$

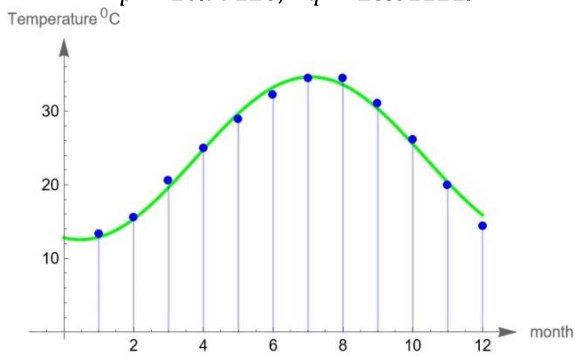

Fig. 4

Remark. If we have obvious increase in temperature, we should use the function

$$T = f\ (t; A, w, p, q)\ = A\sin(wt + p) + qt.$$

Example 5. The problem of approximating average annual daily traffic through IA category state roads in Republic of Serbia, counted at Interchange Niš North - Interchange Niš East and given in the Table 3.

Table 3

Year	Average Annual Daily Traffic
2015	7 025
2016	7 742
2017	8 220
2018	10676
2019	10 225
2020	7 043
2021	10 808
2022	11 665
2023	12 372

We model experiment by a curve

$$y(t) = \frac{a}{t} + b + ct + dt^2,$$

with unknown parameters a, b, c and d.

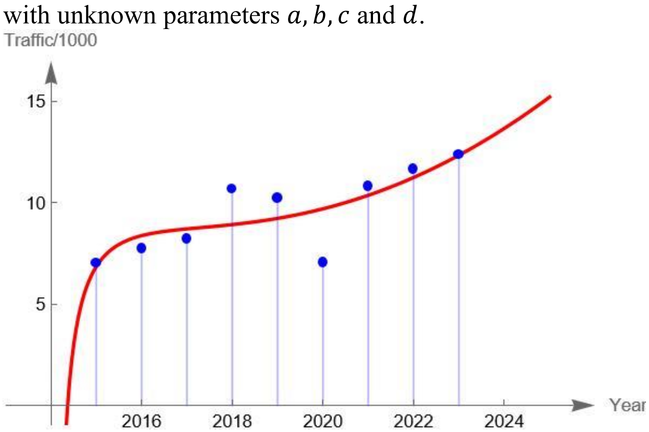

Fig. 5

We have got

$$y_{\mathrm{approx}}(t) = \frac{-4.2919}{t} + 11.9988 - 0.9737t + 0.1184t^2.$$

From this approximation, we can make prognosis $y(2024) \approx 13\,679$. The exact data is $y(2024) = 13\,603$. For the next year, we suppose $y(2025) \approx 15\,224$.

Example 6. The problem of approximating a solution x^* of multidimensional equation (see Argyros [1])

$$\left(\nabla f(x)\right)^T f(x) = 0,$$

where $f(x)$ is a Frechet-differentiable operator defined on R^n with values in $R^m (m \geq n)$ is equivalent to

$$\min g(x) = \left(f(x)\right)^T f(x).$$

It comes from the formula

$$\nabla\left(\left(f(x)\right)^T f(x)\right) = \left(\nabla f(x)\right)^T f(x) + \left(f(x)\right)^T \nabla f(x).$$

We solve it numerically by the Gauss-Newton method.

For example, let $f(x, y) = \begin{bmatrix} x - 2y^2 + 1 \\ x + 3y + 2 \end{bmatrix}$. Starting with the initial value $\vec{x}^{(1)} = [0.1 \;\; 0.2]^T$, we seek for the solution of

$$\left(\nabla f(x)\right)^T f(x) = \begin{bmatrix} 2x + 3y - 2y^2 - 1 \\ 3x + 5y - 4xy + 8y^2 - 6 \end{bmatrix} = 0,$$

as it is shown on the Figure 4. We reduce it to a minimization problem. Applying the Gauss-Newton method, we find

$$x \approx 0.058422, \quad y \approx 0.686141.$$

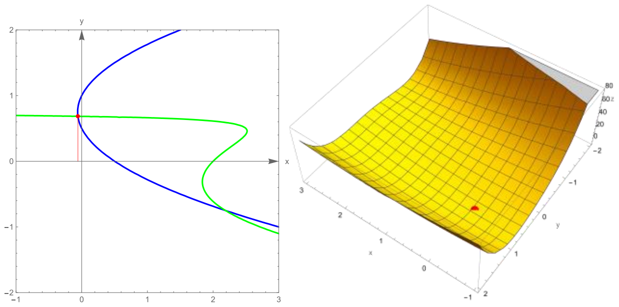

Fig. 6

Example 7. The Newton method does not converge if someone applies it in solving the system (see Hendrata [5]):

$$r_1(x, y) = 10^4 xy - 1 = 0,$$
$$r_2(x, y) = e^{-x} + e^{-y} - 1.0001 = 0,$$

with initial value $(2,2)$. But, the modified Gauss-Newton method gives the approximation of the real solution in 15 iterations with these significant digits:

$$(x^*, y^*) \approx (9.106, 0.00001098).$$

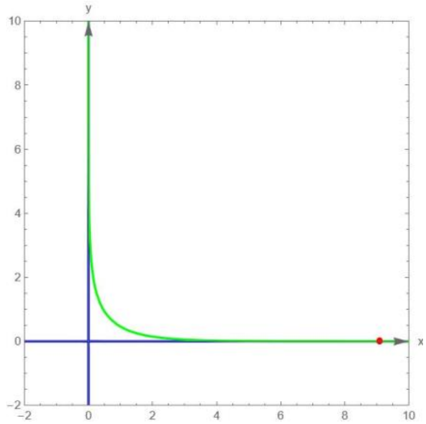

Fig. 7

Example 8. The q-Pochammer symbol is defined by product

$$(a;q)_n = \prod_{k=0}^{n-1}(1 - aq^k) \quad (n \in N \cup \{\infty\}).$$

Here, $(a;q)_0 = 1$. The *Euler identity*

$$\prod_{k=1}^{\infty}(1 - aq^k) = \sum_{k=0}^{\infty}(-1)^k a^k \frac{q^{k(k+1)/2}}{(q;q)_k}, \tag{14}$$

enable us to compute an infinite sum instead of the infinite product

$$\prod_{k=1}^{\infty}(1 - aq^k) \approx \sum_{k=0}^{n}(-1)^k a^k \frac{q^{k(k+1)/2}}{(q;q)_k}.$$

Let us consider the system

$$r_1(x, y) = 2(qx;q)_\infty - (qy^3;q)_\infty = 0,$$
$$r_2(x, y) = (qx^2;q)_\infty + e^{(qy;q)_\infty} - 5 = 0,$$

with $q = 0.75$ and initial value $(0.2, -0.1)$. Their graphs are shown in Figure 5 where $r_1(x, y)$ is blue line and $r_2(x, y)$ is green line.

Here is very hard to use any method which includes the classical derivative. But, the q-Gauss-Newton method gives the approximation of the real solution in 10 iterations with these significant digits:

$$(x^*, y^*) \approx (0.218193, -0.119667).$$

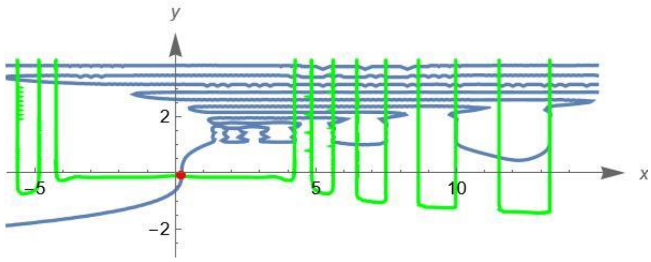

Fig. 8

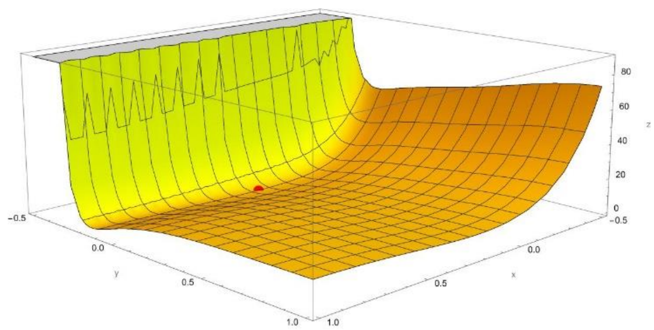

Fig. 9

Example 9. (Nonconvergence case) There are theorems which guarantee convergence of G-N method. In other cases, we are not sure that it leads us to the solution. Here is an counterexample.

Let us apply the Gauss-Newton method with initial value $\vec{x}^{(1)} = [1 \;\; 1]^T$ to the minimization problem

$$\min f(x, y) = (10y + x^2)^2 + (1 + x^2)^2.$$

Here, we have

$$\vec{r}(x) = \begin{bmatrix} 10y + x^2 \\ 1 + x^2 \end{bmatrix}, \quad J(x) = \begin{bmatrix} \frac{\partial r_1}{\partial x} & \frac{\partial r_1}{\partial y} \\ \frac{\partial r_2}{\partial x} & \frac{\partial r_2}{\partial y} \end{bmatrix} = \begin{bmatrix} 2x & 10 \\ 2x & 0 \end{bmatrix}.$$

Hence

$$\vec{r}(x^{(1)}) = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \quad J_1 = \begin{bmatrix} 2 & 10 \\ 2 & 0 \end{bmatrix},$$

and

$$\vec{x}^{(2)} = \vec{x}^{(1)} - (J_1)^{-1}\vec{r}(x^{(1)}) = \begin{bmatrix} 0 \\ 0.1 \end{bmatrix}.$$

But, since $\det[J_2] = 0$, the method does not converge. Even more, it may not be effective when $0 < \det[J_2] \ll 1$.
In these cases, we can apply the modified G-N method. Really, for this problem we will get iterations

$$\vec{x}^{(2)} = \begin{bmatrix} 0.33(-15) \\ 0.1 \end{bmatrix}, \vec{x}^{(3)} = \begin{bmatrix} 0.32(-15) \\ 0.05 \end{bmatrix}, \dots, \vec{x}^{(10)} = \begin{bmatrix} 0 \\ 0.00039 \end{bmatrix}.$$

The exact solution is $\min f(x, y) = f(0,0) = 1.$

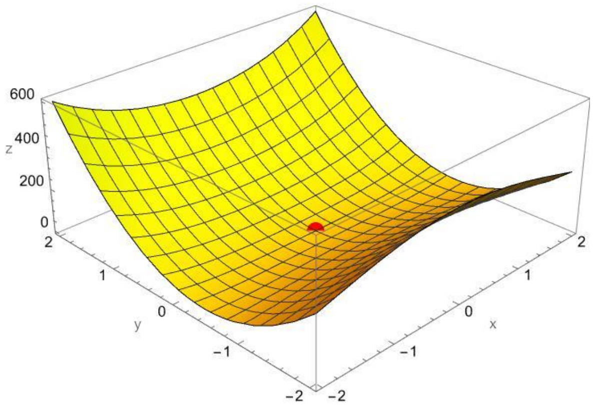

Fig. 10

6 CONCLUSIONS

The Gauss-Newton method is a highly efficient iterative method used to solve nonlinear least-squares problems. This can be seen as a modification of the newton method to find the minimum value of a function. The analysis of Gauss Newton's method involves convergence of simple roots and multiple roots. Gauss Newton's method often converges quickly, especially when the iteration begins to be close enough to the desired root. But if iteration begins far from the searched root, this method can slip off without warning, its convergence may be slow, it probably will not converge at all. Implementation of this method usually detects and overcomes convergence failures.

ACKNOWLEDGEMENTS

This research was financially supported by the Ministry of Education, Science and Technological Development of the Republic of Serbia (Contract No. 451-03-68/2022-14/200109).

REFERENCES

1.Argyros, I.K., Hilout, S., 2012, *On the Gauss-Newton method for solving equations*, Proyecciones Journal of Mathematics, vol. 31, No 1, pp. 11-24.
2.Gratton, S., Lawless, A.S., Nichols, N.K., 2007, *Approximate Gauss–Newton Methods for Nonlinear Least Squares Problems,* SIAM J. Optimization, vol. 18, No. 1, pp. 106–132.
3. Dixon, L.C.W., Price, R.C., 1986, *Numerical Experience with the Truncated Newton Method*, Numerical Optimisation Centre, Hatfield Polytechnic, Technical Report No. 169.
4. Grippo, L., Lampariello, F., Lucidi, S., 1989, *A Truncated Newton Method with Nonmonotone Line Search for Unconstrained Optimization*, Journal of Optimization Theory and Applications: vol. 60, No. 3, pp. 401-419.
5. Hendrata, M., Subramanian, P.K., 2012, *Some remarks on Newton's algorithm* - II, International Journal of Pure and Applied Mathematics: vol. 74, No. 3, pp. 373-391.
6. Marinković, S.D., Rajković, P.M., Stanković, M.S., 2013, *The q-iterative methods in numerical solving of some equations with infinite products*, Facta Universitatis (Niš), Ser. Math. Inform. Vol. 28, No 4, pp. 379–392.
7. Nocedal, J., Wright, S.J., 2006, *Numerical Optimization*, Springer.
8. Rajković, P.M, Marinković, S.D., Stanković, M.S., 2008, *Differential and Integral Calculus of Basic Hypergeometric Functions* (monography in Serbian), Faculty of Mechanical Engineering, Niš.
9. Ram, B., Mishra, S.K., Lai, K.K., Rajković, P.M., 2024, *Unconstrained Optimization and Quantum Calculus*, Book, Series title: Uncertainty and Operations Research, Springer.
10. Wedin, P.A., 1973, *Perturbation theory for pseudo-inverse*, BIT 13, 217–232.
11. Wedin, P.A., 1974, *On the Gauss–Newton method for the nonlinear least-squares problem*, Rep. 24, The Swedish Institute of Applied Mathematics.

Contact address:
Predrag Rajković
University of Niš, Faculty of Mechanical engineering
18000 Niš, A. Medvedeva 14
E-mail: predrag.rajkovic@masfak.ni.ac.rs