Wasit Journal for Pure Science

Journal Homepage: https://wjps.uowasit.edu.iq/index.php/wjps/index
e-ISSN: 2790-5241 p-ISSN: 2790-5233

# A Numerical Technique for Solving Optimization Problems

**Safaa M. Aljassas[1], Ahmed Sabah Al-Jilawi[2]**

[1]First Author Affiliation,
College of Education for girls, University of Kufa, IRAQ

[2]Second Author Affiliation,
College of Basic Education, University of Babylon, IRAQ
Aljelawy2000@yahoo.com

*Corresponding Author: Safaa M. Aljassas

DOI: https://doi.org/10.31185/wjps.92
Received 18 August 2023; Accepted 29 September 2023; Available online 30 December 2023

**ABSTRACT:** The aim of this paper is to calculate a better approximation value (whether it is maximized or minimize) for one- and two-dimensional nonlinear equations using the best numerical optimization algorithms, which is Newton's method. The idea of this technique is based on approximating the function by expanding the Taylor series expansion and iteratively updating the estimate of the optimal solution. we have obtained good results in terms of accuracy and speed of approach, as shown in the examples mentioned. We also mentioned the applications of Newton's method in multiple disciplines, including engineering, physics, economics, finance, computer graphics, machine learning, image processing, and other applications.

**Keywords:** numerical optimization algorithm, Newton's method, nonlinear equations The Hessian matrix, Gradient of a function

## 1. INTRODUCTION

Optimization techniques have been utilized since the era of prominent mathematicians such as Cauchy, Lagrange, and Newton. The contributions of Newton and Leibnitz to the field of calculus facilitated the advancement of differential calculus techniques in the realm of optimization. The foundational principles of the calculus of variations were established by prominent mathematicians such as Bernoulli, Euler, Weistrass, and Lagrange. Lagrange's name has become closely associated with the optimization strategy used to solve limited problems by including unknown multipliers. The technique of steepest descent was initially employed by Cauchy in order to address unconstrained optimization problems. The advent of high-speed digital computers throughout the mid-20th century facilitated the execution of intricate optimization procedures, hence stimulating more investigation into novel methodologies. Subsequent to these remarkable advancements, a substantial corpus of knowledge pertaining to optimization strategies was generated. The emergence of numerous well-defined topics in optimization theory has been a direct result of this achievement. This discussion highlights some significant advancements in the field of numerical methods for unconstrained optimization. [1]

The field of optimization is wide and continually evolving. The characteristics of maximizes and minimizers of functions are contingent upon the availability of mathematical tools. These encompass mathematical techniques derived from calculus, topology, and other geometric concepts. The theoretical framework that underlies contemporary computer optimization techniques is in a constant state of evolution. Some instances of mathematical techniques commonly employed in optimization problems include interior point methods, applications rooted in control theory, and algorithms based on duality. [2].

*Corresponding author: safaam.musa@uokufa.edu.iq

## 2. OPTIMAZATION

Optimization is the mathematical procedure of maximizing or minimizing the objective function that satisfies the specified limitations or chooses the best solution to the problem from the available alternatives [7,8].

### 2.1 Unconstrained Optimization

Consider the scenario wherein the objective function has to be minimized, taking into account real variables that are not subject to any value restrictions. Let us denote the real vector as $x \in R^n$ , comprising components $n \geq 1$, and let represent a differentiable function $f : R^n \rightarrow R$. The unconstrained optimization issue is represented by the following symbolic expression, [4,8]:

$$min_x f(x)$$

### 2.2 Constrained Optimization

Assume a general constrained problem of the following form

$$\min_x f(x)$$

$$\text{subject to } x \in X$$

$$g_i(x) = 0 \qquad i \in \omega$$

$$L_i(x) \leq 0 \qquad i \in \tau$$

where g is the equality constraint, L is the inequality constraint and we called $f$ as an objective function, [4,7].

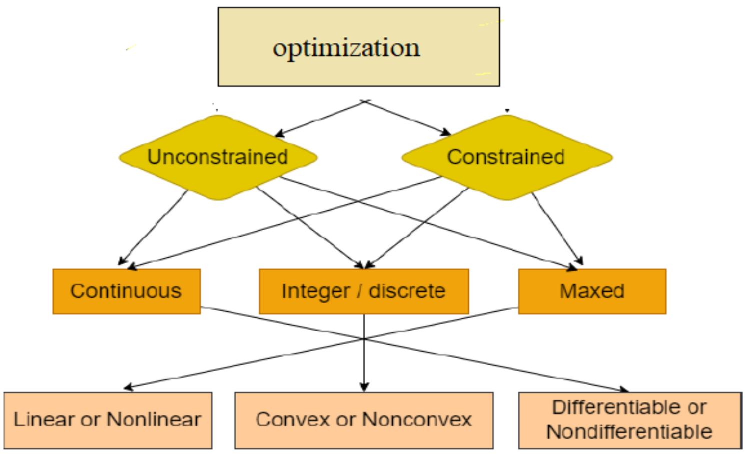


**Fig. 1.** Classification of Optimization

## 3. GRADIENT OF A FUNCTION

Let $f : \mathbb{R}^n \rightarrow \mathbb{R}$ be a continuous function on the open neighborhood D $\subseteq$ $R^n$ then the gradient of $f$ is defined as follows [3].

$$\nabla f = \begin{bmatrix} \frac{\partial f}{\partial x_1} \\ \vdots \\ \frac{\partial f}{\partial x_n} \end{bmatrix} , \quad \text{where} \quad x = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$$

## 4. THE HESSIAN MATRIX

Given $\boldsymbol{f}: \mathbb{R}^{\boldsymbol{n}} \rightarrow \mathbb{R}$ , if $\boldsymbol{\nabla f}$ is differentiable, we say for that $\boldsymbol{f}$ is twice differentiable, and write the second derivative of $\boldsymbol{f}$ as:

$$\nabla^2 f\,(x) = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2} & \cdots & \frac{\partial^2 f}{\partial x_n \partial x_1} \\ \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_1 \partial x_n} & \cdots & \frac{\partial^2 f}{\partial x_n^2} \end{bmatrix} \quad , \quad \text{where } \ x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

The matrix $\nabla^2 f\,(x)$ is called Hessian Matrix of $f$ at $x$, and often too Symbolized by $F'(x)$ ,[3].

## 5. NEWTON METHOD

Suppose the problemis distinguishable twice continuously. Additionally, we assume that the $f$, function $\min_x f\,(x)$ function $f$ has a lower bound. $d^k = -H^{-1} g^k$ indicates the direction of the search based on the information from the first and second derivatives. Here $g^k$ represents the gradient at the point $x^k$ .The iterative approach involves using a Taylor series expansion to approximate a given function $f$ at a given point $x^k$ using a quadratic. Then the lower bound of this quadratic function is determined to obtain $x^{k+1}$ an estimate of the required value. This procedure continues until the specified criterion is met,[7].

$$f(x) \approx f_q(x) = f(x^k) + g^{k^T}(x - x^k) + \tfrac{1}{2}(x - x^k)^T H^T (x - x^k) \quad \ldots(1)$$

where $H^k$ is the Hessian at the current point $x^k$, the new iteration is obtain by minimizing the quadratic function $f_q(x)$ this means that $x^{k+1} = \arg\min_x f_q(x)$ the gradient of $f_q$ denote by $\nabla f_q(x) = 0 \Rightarrow x^{k+1} = x^k - (H^k)^{-1} g^k$ (assuming $H^k$ is in revertible ), if write it in our visual form $x^k + \alpha^k d^k$ , chosen $\alpha^k = 1$ and $d_N{}^{-}(H^k)^{-1} g^k$ is called the Newton direction. The Newton direction is descent direction if H is positive definite $g^{k^T} d_N = -g^{k^T}(H^k)^{-1} g^k < 0$ (it satisfy when H ≻ 0) Consider the problem to minimize $f(x) = \frac{1}{2} x^T H x - c^T x$ where H is a symmetric positive matrix g(x) = 0 $\Rightarrow x^* = H^{-1} c$ is a strict local minimum let $x^0 \in R^0$ be any point$g(x^0) = Hx^0 - c, H(x^0) = H$using classical Newton method $x^1 = x^0 - H^{-1}(Hx^0 - c) = H^{-1} c = x^*$ for a quadratic function, if apply classical Newton method starting from any point can reach the solution in exactly one step, assuming that the initial point is not the optimal point compare this with the behavior of the steepest descent method depends alto on the condition number of the Hessian matrix $H = 1$ converges in one step starting from any point and if the condition number of the a Hessian matrix H > 1 then that the zigzagging takes place for a typical starting point.. [5,9].

In general, the minimum of the perfectly convex quadratic function, (with the invertible Hessian matrix) can be obtained in a single iteration using the classical Newton approach, regardless of the initial starting point.

### 5.1 Newton Method Algorithm

**Step 1**: Initialize $x^0$ and ϵ, set k = 0

- Set an initial guess $x^0$ for the root or stationary point.
- Set a tolerance ∈ to define the convergence criterion.
- Initialize k as 0 to keep track of the number of iterations.

**Step 2:** While $\|g\,(x^k)\| \geq \in$

- This is the main loop that iterates as long as the value of the function $g(x_k)$ is greater than or equal to the specified tolerance ∈ .

a. $d^k = -(H^{k-1}) g^k$

Calculate the update direction $d^k$ as the negation of the product of the inverse of the Hessian matrix $H^{k-1}$ and the gradient $g^k$ of the function f(x) evaluated at the current approximation $x_k$.

b. Find $\alpha^k = 1$. Determine the step size $\alpha^k$ in code, it's directly set to 1.

c. $x^{k+1} = x^k + \alpha^k d^k$

Update the current approximation $x^k$ using the step size $\alpha^k$ and the calculated update direction $d^k$. This produces the next approximation $x^{k+1}$

d. k = k + 1

Increment the iteration counter k to prepare for the next iteration.

**Step 3**. End while.

**Step 4**. output $x^* = x^k$ a stationary of $f(x)$

### 5.2 Definition

A locally convergent iteration optimization algorithm is characterized by the property that, for every solution $x^*$ , there exists a positive value δ such that, for any initial point $x^0 \in B(x^*, \delta)$, the process generates a sequence $\{x^k\}$that converges to $x^* \in R$ .(δ is function of $x^*$). [9].

### 5.3 Theorem:

Consider a function $f: R \to R$ that belongs to the class $C^2$. Let $x^*$ be an element in the real numbers such that $g(x^*) = 0$ and $g'(x^*) > 0$. Assuming that $x^0$ is sufficiently closed to $x^*$, the sequence $\{x^k\}$ generated by the classical Newton method converges to $x^*$ with an order of convergence two (locally convergent) [9].

**Proof:** Since $f: R \to R, f \in C^2$ consider the problem min $f(x)$, and since $x^* \in R$ be such that $g(x^*) = 0$ and $g'(x^*) > 0$ assume that $x^0$ is sufficiently closed to $x^*$ ,suppose apply classical Newton algorithm to minimize $f(x)$ At $k^{th}$ iteration

$$x^{k+1} = x^k - \frac{g(x^k)}{g'(x^k)} \quad \dots (2)$$

$$x^{k+1} - x^* = x^k - x^* - \frac{g(x^k) - g(x^*)}{g'(x^k)}$$

$$= -\frac{\left(g(x^k) - g(x^*) + g'(x^k)(x^* - x^k)\right)}{g'(x^k)} \quad \dots(3)$$

If assume that $f \in C^3 \left(or\ g \in C^2\right)$, then using truncated Taylor series

$$g(x^*) = g(x^k) + g'(x^k)(x^* - x^k) + (0.5) g''(\bar{x}^k)(x^* - x^k)^2 \quad \dots(4)$$

Where $\bar{x}^k \in LS(x^* - x^k)$

Therefore

$$x^{k+1} - x^* = (0.5)\frac{g''(\bar{x}^k)}{g'(x^k)}(x^k - x^*)^2 \quad \dots (5)$$

got relationship between ( $x^{k+1}$ and $x^*$),$\left(x^k\ and\ x^*\right)^2$, if recall the definition of convergence of algorithm, this would turn out to be order 2convergence,

$$|x^{k+1} - x^*| = (0.5)\frac{|g''(\bar{x}^k)|}{|g'(x^k)|}|(x^k - x^*)^2| \quad \dots(6)$$

suppose there exist $\alpha_1$ and $\alpha_2$ such that $\left|g''(\bar{x}^k)\right| < \alpha_1$ for all $\bar{x}^k \in LS\left(x^* - x^k\right)$ $\left|g'(x^k)\right| > \alpha_2$ for all $x^k$ sufficiently close to $x^*$ then

$$|x^{k+1} - x^*| \leq (0.5)\frac{\alpha_1}{\alpha_2}|(x^k - x^*)^2| \qquad \ldots(7)$$

Where $(0.5)\frac{\alpha_1}{\alpha_2} > 0$ and constant (order two convergence if $x^k - x^*$)

Where $(0.5)\frac{\alpha_1}{\alpha_2}(|(x^k - x^*)|)$required to be < 1, then got the distance between $x^{k+1}$ and $x^*$ less than distance between $x^k$ and $x^*$ $|x^{k+1} - x^*| < |x^k - x^*|, \forall k$
Choose $\alpha_1$ and $\alpha_2$ in some way such that the inequality

$$|x^{k+1} - x^*| < |x^k - x^*|, \forall k \qquad \ldots(8)$$

holds. Now, $g(x^*) = 0$ and $g'(x*)$ 0 since $g' \in C^0(g')$ continues $\exists\eta > 0 \ni g'(x) > 0, \forall x \in (x* - \eta, x* + \eta)$

$$\alpha_1 = \max_{x \in \left(x^* - \eta, x^* + \eta\right)} \left|g''(x)\right|$$

$$\alpha_2 = \max_{x \in \left(x^* - \eta, x^* + \eta\right)} \left|g'(x)\right|$$

Therefore

$$\left|0.5\frac{g''\bar{x}^k}{g'\bar{x}^k}\right| \leq (0.5)\frac{\alpha_1}{\alpha_2} \qquad \ldots(9)$$

preferable to choose $x0 \in (x^* - \eta, x^* + \eta)$ also, want $\beta\left|x^k - x^*\right| < 1, \forall k$, that is $|x^k - x^*| < \frac{1}{\beta}, \forall k \Rightarrow x^k \in \left(x^* - \frac{1}{\beta}, x^* + \frac{1}{\beta}\right)$,therefore choose
$x^0 \in (x^* - \eta, x^* + \eta) \cap x^0 \in (x^* - \eta, x^* + \eta)$Now, will show if $x^0$ choose in this regain then $x^k$ converge to $x^*$

$$|x^k - x^*| \leq |x^{k-1} - x^*|^2$$

$$\beta|x^k - x^*| \leq (\beta|x^{k-1} - x^*|)^{2^k}$$

$$|x^k - x^*| \leq \frac{1}{\beta}\underbrace{(|x^{k-1} - x^*|)^{2^k}}_{<1} \Rightarrow \lim_{\beta\to\infty}|x^k - x^*| = 0 \qquad \ldots(10)$$

the problem is that the initialization of $x^0$ require knowledge of $x^*$ and our aim is to minimize *f(x)* to get $x^*$. So, this knowledge of $x^*$ is not there cannot initialize $x^0$ properly, so that, can get global convergence of Newton method. In other words, the Newton method does depend a lot on $x^0$.

## 6. EXAMPLES

### 6.1 One-dimensional non-linear equations

**Example (1)**:- apply the Newton Method to solve the nonlinear equation
max $f(x) = \ln x + \sin(x) + 2$ when Initial value $x_0 = 0.2$

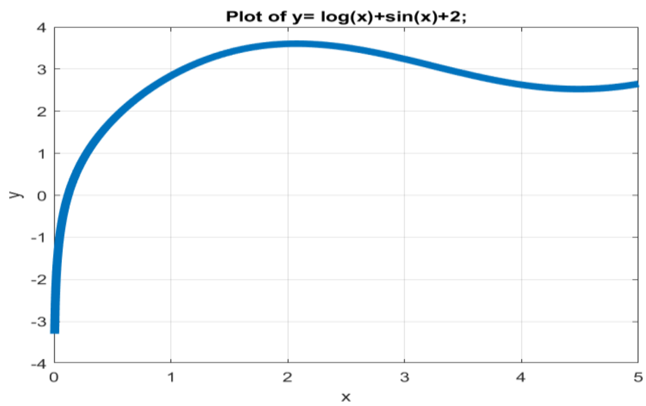


**Fig. 2.** graph of Function $\max f(x) = \ln x + \sin(x) + 2$

**Solution**: - when applying Newton formula $x_{i+1} = x_i - \frac{f(x_i)}{f'(x_i)}$ , $f'(x_i) \neq 0$ on the mentioned equation, and account $f(x_i), f'(x_i), i = 1, 2, \cdots, 6$ we obtained the results recorded in the table (1) which includes i mean number of iterations and values x.

**Table 1.**

| I | value x |
|---|---|
| 0 | 0.02 |
| 1 | 0.05709866 |
| 2 | 0.10063316 |
| 3 | 0.11854476 |
| 4 | 0.12005053 |
| 5 | 0.12005913 |
| 6 | 0.12005913 |

Where the results showed that a good approximate solution appears after six iterations and the maximize function $f(x) = \ln x + \sin(x) + 2$ is 0.12005913.

**Example (2)**: - apply the Newton Method to solve the nonlinear equation $min\, f(x) = xe^x + \frac{1}{x} + 3$ when Initial value $x_0 = 1$

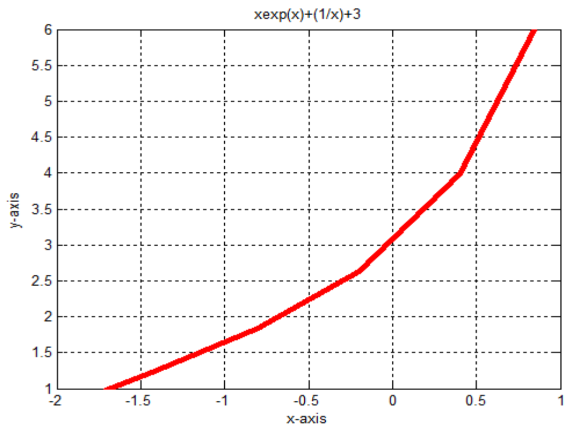


**Fig 3**. graph of Function

$$min\ f\ (x) = xe^x + \frac{1}{x} + 3$$

**Solution** -:when applying Newton formula $x_{i+1} = x_i - \frac{f(x_i)}{f'(x_i)}$ , $f'(x_i) \neq 0$ on the mentioned equation, and account $f(x_i)$ , $f'(x_i)$ ,i=1,2,⋯ ,6 we obtained the results recorded in the table(2) which includes i mean number of iterations and values x.

**Table 2.**

| I | value x |
|---|---|
| 0 | 1 |
| 1 | -0.51429853 |
| 2 | -0.29996104 |
| 3 | -0.35243596 |
| 4 | -0.36364616 |
| 5 | -0.36402677 |
| 6 | -0.36402677 |

Where the results showed that a good approximate solution appears after six iterations and the maximize e function $\min f\ (x) = xe^x + \frac{1}{x} + 3$ is -0.36402677.

### 6.2 Two-dimensional nonlinear equations

**Example (3)** :- apply the Newton Method to solve the nonlinear equation $\min f\ (x) = 2x^2 + 4y^2 + e^{x+y}$ when Initial values

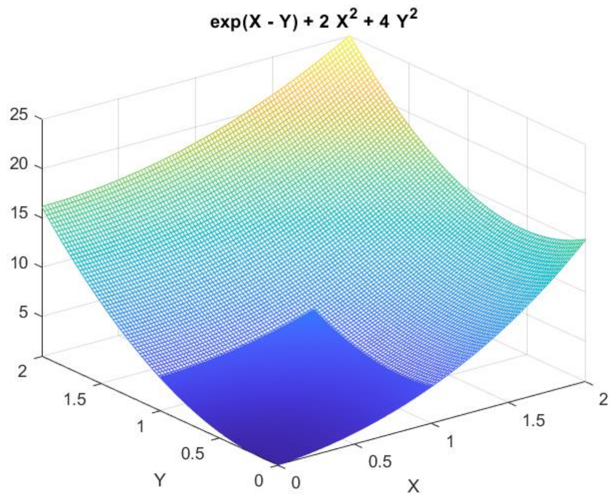


**Fig. 3.** graph of Function $\min f\ (x) = 2x^2 + 4y^2 + e^{x+y}$

Solution: - Gradient of a function = $g_1 = \nabla f = \begin{bmatrix} e^{x+y} + 4x \\ e^{x+y} + 8y \end{bmatrix}$

and The Hessian matrix. $H = \begin{bmatrix} e^{x+y}+4 & e^{x+y} \\ e^{x+y}+8 & e^{x+y} \end{bmatrix}$

We apply Newton formula $x^{k+1} = x^k + \alpha^k d^k$, $d^k = -\left(H^{k-1}\right) g^k$ on the mentioned equation, and account Gradient of a function and The Hessian matrix at the initial values $x_0 = y_0 = 0.5$. we obtained the lowest results: -

The initial value of the objective function: 2.5
The minimum successfully achieved.
The number of iterations required for achieving convergence.: 5
Point of Minima: [-0.18844312399, 0.09422156199]
Objective Function Minimum Value after Optimization: 0.860305

**Table 3.**

| i | Value x | Value y |
|---|---|---|
| 0 | 0.5 | 0.5 |
| 1 | -0.18181818181 | 0.09090909091 |
| 2 | -0.18843583768 | 0.09421791884 |
| 3 | -0.18844312398 | 0.09422156199 |
| 4 | -0.18844312399 | 0.09422156199 |

**Example (4):-** apply the Newton Method to solve the nonlinear equation $\min f(x) = 2x^3 + y^3 + 5xy + 2x - 4y$ when Initial values $x_0 = y_0 = 0$

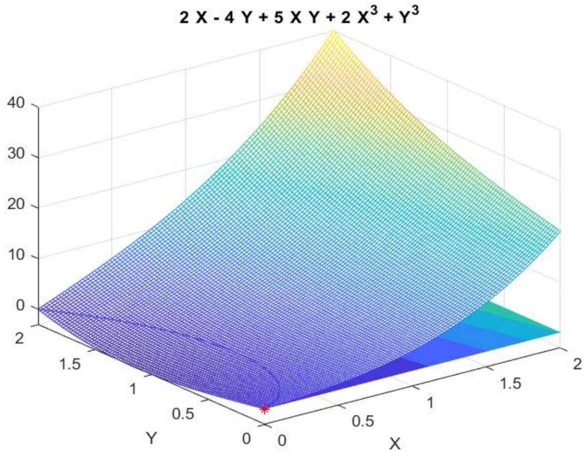


**Fig. 4.** graph of Function $\min f(x) = 2x^3 + y^3 + 5xy + 2x - 4y$

**Solution:** - Gradient of a function = $g_1 = \nabla f = \begin{bmatrix} 6x^2+5y+2 \\ 3y^2+5x-4 \end{bmatrix}$ ,The Hessian matrix. $H = \begin{bmatrix} 12x & 5 \\ 5 & 6y \end{bmatrix}$ we apply Newton formula $x^{k+1} = x^k + \alpha^k d^k$ , $d^k = -\left(H^{k-1}\right) g^k$ on the mentioned equation, and account Gradient of a function and The Hessian matrix at the initial values $x_0 = y_0 = 0$. we obtained the lowest results: -

The initial value of the objective function: 0
The minimum successfully achieved.
The number of iterations required for achieving convergence: 7
Point of Minima: [0.50297974747, - 0.703586351639]
Objective Function Minimum Value after Optimization: 1.957054

**Table 4.**

| i | Value x | Value y |
|---|---|---|
| 0 | 0 | 0 |
| 1 | 0.8 | -0.4 |
| 2 | 0.55820149875 | -0.70374687760 |
| 3 | 0.50442967113 | -0.701869460415 |
| 4 | 0.50298167522 | -0.703586162707 |
| 5 | 0.50297974747 | -0.703586351637 |
| 6 | 0.50297974747 | -0.703586351639 |

**Example (5)**:- apply the Newton Method to solve the nonlinear equation
$\max f(x) = \ln(1+2xy) + \cos(2x+3y)$ when Initial values $x_0 = y_0 = 1$

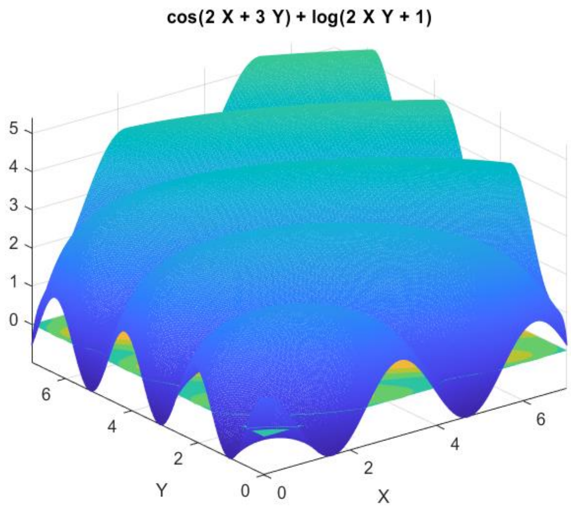


**Fig. 5.** graph of Function $\max f(x) = \ln(1+2xy) + \cos(2x+3y)$

**Solution:** - Gradient of a function $= g_1 = \nabla f = \begin{bmatrix} \frac{2y}{1+2xy} - 2\sin(2x+3y) \\ \frac{2x}{1+2xy} - 3\sin(2x+3y) \end{bmatrix}$,

The Hessian matrix.

$$H=\begin{bmatrix} \frac{-2y^2}{(1+2xy)^2}-4\cos(2x+3y) & \frac{2}{(1+2xy)}-6\cos(2x+3y)-\frac{-4xy}{(1+2xy)^2} \\ \frac{2}{(1+2xy)}-6\cos(2x+3y)-\frac{-4xy}{(1+2xy)^2} & \frac{-4y^2}{(1+2xy)^2}-9\cos(2x+3y) \end{bmatrix}$$

we apply Newton formula $x^{k+1}=x^k+\alpha^k d^k$ , $d^k=-\left(H^{k-1}\right)g^k$ on the mentioned equation, and account Gradient of a function and The Hessian matrix at the initial values $x_0=y_0=0$. we obtained the lowest results: -

The initial value of the objective function: 6
The minimum successfully achieved.
The number of iterations required for achieving convergence: 6
point of minima: [1.538957433913, 2.30843615086]
Objective Function Minimum Value after Optimization: 6.1106935713985

**Table 5.**

| i | Value x | Value y |
|---|---|---|
| 0 | 1 | 1 |
| 1 | 1.98423879028 | 1.69627290318 |
| 2 | 2.29321584657 | 1.551735584573 |
| 3 | 2.30846406768 | 1.538944002054 |
| 4 | 2.30843615080 | 1.538957433967 |
| 5 | 2.30843615086 | 1.538957433913 |

## 7. CONCULUSION

A differentiable function's minimum or maximum can be found using Newton's method, an iterative numerical methodology for optimization. Based on calculus principles, it iteratively improves the approximation of the best solution by using the gradient and Hessian matrix (second derivatives) of the function. Each iteration of the method begins with an initial approximation for the optimum, which is subsequently updated based on the local curvature of the function. If the function is well-behaved and the initial guess is close to the optimum, it converges quickly. In one-dimensional optimization, Newton's approach uses the function's first derivative (gradient) to update the current guess; in multi-dimensional optimization, it uses the gradient plus the Hessian matrix to determine the update direction.

One key advantage of Newton's method is its fast convergence rate. It can efficiently find the optimum in fewer iterations compared to some other optimization algorithms. The method's effectiveness heavily depends on the choice of the initial guess and the nature of the function being optimized. For non-convex functions or poor initial guesses, it may converge to local optima rather than the global optimum.

## REFERNCES


[1] Nagesh Kumar, IIS , Bangalore, , Optimization Methods ",(2013).

[2] Borwein J. M. ,and Lewis A. S.," Convex analysis and Nonlinear ear optimi zation Theory and Example ", Canada V5A IS6,( 1999).

[3] Edwin K.P.C. and Stanislaw H.Ż. "An introduction to optimization", second edition ,John wiley &Sons ,Inc ,(2001).

[4] Jorge Nocedal and Stephen J. Wright, Numerical Optimization, (Second Edition). Springer, 2006.

[5] Josef Kallrath. Mathematical solution techniques the nonlinear world.In Business OptimizationUsing Mathematical Programming, pages 423 - 446.Springer,2021.

[6] Shashi Kant Mishra, Mohammad Esmael Samei, Suvra Kanti Chakraborty, and Bhagwat Ram. On q-variant of dai–yuan conjugate gradient algorithm for unconstrained optimization problems. Nonlinear Dynamics, pages 1–26, 2021.

[7] Nomi,Ammar Imad Nadhim,' New Approaches to Solve Two-sided Optimization Problems in Numerical Analysis', Dissertation of Doctor Submitted to University of Babylon, 2023.

[8] Marjan,Mazin Kareem Kadhim,' Employing Control Mathematical Models of Numerical Differential Optimization 'Dissertation of Doctor Submitted to University of Babylon, 2023.

[9] Xiangli Li, Wenjuan Zhao,and Xiaoliang Dong. A new cg algorithm based on a scaled memoryless bfgs update with adaptive search strategy, and its application to large-scale unconstrained optimization problems. Journal of Computational and Applied Mathematics, page 113670, 2021