DOI: 10.54254/2753-8818/9/20240764

Research on function optimization problem based on Newton's method

Junlong He

Nanjing Foreign Language School Xianlin Campus, Nanjing, Jiangsu, China, 210000

769748774@qq.com

Abstract. Optimization is important because it helps improve efficiency, reduce waste, and achieve better outcomes in various fields such as business, engineering, and science. It enables organizations and individuals to make the most of their resources and capabilities, ultimately leading to increased productivity, cost savings, and better overall performance. The main focus of this paper is on studying how to optimize results when solving mathematical problems. We use optimization in mathematical problems to achieve the best possible solution or outcome by efficiently utilizing available resources and constraints. This helps us make informed decisions, save time, and find solutions that are both effective and efficient. The main focus of this article is to explore what Newton's method is, its fundamental principles, and its practical applications in solving mathematical problems. It delves into how Newton's method can be employed to achieve optimization within mathematical problem-solving contexts. The article discovers that utilizing Newton's method to solve mathematical problems for computational optimization is a favorable choice.

Keywords: Optimization, Newton's Method, Convergence.

1. Introduction

Convex optimization holds a significant position in the field of mathematical programming. Once a practical problem is formulated as a convex optimization problem, it generally implies that the corresponding problem has been thoroughly solved, a property not possessed by non-convex optimization problems. Its applications are very extensive, and many optimization problems in machine learning are solved through convex optimization. In non-convex optimization, convex optimization also plays a crucial role, as many non-convex optimization problems can be transformed into convex optimization problems for resolution [1]. In recent years, the issues surrounding function optimization using the Newton method have garnered significant attention in the fields of numerical computation and optimization. Within this realm, there exist several specialized terms that warrant elucidation. Firstly, the Newton method is an iterative numerical approach utilized for solving nonlinear equations and optimization problems. Its fundamental concept involves gradually approximating the optimal solution by constructing the Taylor series expansion of the function, thereby expediting the convergence process. Secondly, function optimization refers to the pursuit of variable values that yield extremal values of the objective function while adhering to specified constraints. The focus of this paper is to explore the effectiveness of Newton-based function optimization in the context of large-scale high-dimensional datasets and intricate constraints. The research methodology encompasses theoretical analysis,

DOI: 10.54254/2753-8818/9/20240764

numerical simulations, and practical case validations. This paper delves deeply into the advantages and drawbacks of this method from perspectives such as algorithm convergence, numerical stability, and computational complexity. Furthermore, this paper endeavors to refine and optimize the method with respect to prevailing issues. This research holds substantial theoretical and practical significance. Theoretically, a profound comprehension of Newton-based function optimization algorithms contributes to the enrichment of the theoretical framework in the field of optimization. From a practical standpoint, the enhancement of optimization methodologies will yield heightened efficiency and performance in domains such as machine learning and engineering optimization. To sum up, this study can offer novel insights and methodologies for addressing intricate problems, thereby propelling the further development and application of optimization algorithms in practical scenarios.

2. The fundamental principle of the Newton's method

Firstly we have to choose an initial point as the starting point for iterations, and then approximate the objective function near the current point using a Taylor series expansion, representing the function as a polynomial, and next treat the expanded polynomial as the objective function and find its minimum (or maximum) value at the current point, obtaining a linear approximation solution. Finally, utilize the obtained minimum (or maximum) value as the next iteration point and repeat the above steps. The approach to finding the extrema of a function is to search for points where the derivative is equal to zero, which involves solving an equation [2]. Specifically, given the current iteration point x, perform a Taylor series expansion of the objective function to obtain:

$$f(x+\triangle x) \approx f(x) + f'(x)\triangle x + \frac{1}{2}f''(x)\triangle x^2 \quad (1)$$

$$0 = f'(x) + f''(x)\triangle x \quad (2)$$

$$x_{new} = x + \triangle x \quad (3)$$

Repeat the iteration until a predefined convergence criterion is met, such as the function's change becoming sufficiently small or reaching the maximum iteration count.

2.1. The basic idea and mathematical principles of the Newton's method

Newton's method, also known as the Newton-Raphson method, is an iterative numerical technique used to find the roots of a real-valued function [3]. The method starts with an initial guess for the root and then iteratively refines the guess by using the tangent line to the function at that point. The intersection of the tangent line with the x-axis provides a better approximation to the root. This process is repeated until a desired level of accuracy is achieved.

2.2. Application of Newton's method in optimization problems

Newton's method excels in solving nonlinear optimization problems. For instance, in machine learning, it finds widespread use in training neural networks. Its rapid convergence and utilization of second-order information allow Newton's method to quickly identify local minima in high-dimensional parameter spaces. Furthermore, Newton's method is also applied in image processing, such as image registration and image reconstruction.

Despite its strengths in many cases, Newton's method has certain limitations. Firstly, computing and storing the Hessian matrix can be memory-intensive, especially in high-dimensional problems. Secondly, the Hessian matrix might not be positive definite, leading to algorithm instability. Additionally, Newton's method is sensitive to initial values, and the choice of initial values can impact the algorithm's convergence. To address the limitations of Newton's method, researchers have proposed various improvement strategies. For instance, quasi-Newton methods approximate the inverse of the Hessian matrix to reduce memory consumption. Regularized Newton's methods can handle the issue of non-positive definiteness. Furthermore, stochastic Newton's methods approximate gradients and Hessian matrices through random sampling, thus accelerating computation.

DOI: 10.54254/2753-8818/9/20240764

3. Application of Newton's method in function minimization problems

3.1. Transforming function minimization problem into an optimization problem

Newton's method, like gradient descent, is an optimization algorithm used to maximize the likelihood function in logistic regression for classification problems. It is a numerical method employed to find the root of a function [4]. Consider a real-valued, continuously differentiable function $f(x)$. Our objective is to determine the x value at which $f(x)$ achieves its minimum value. Initially, we perform a second-order Taylor expansion of $f(x)$ around the current point x_0:

$$f(x) \approx f(x_0) + f'(x_0)(x - x_0) + \frac{1}{2} f''(x_0)(x - x_0)^2 \tag{4}$$

Here, f' and f'' denote the first and second derivatives of $f(x)$, respectively. To locate the extremal points of $f(x)$, we aim to nullify the derivative term in the expansion, which yields:

$$f'(x_0) + f''(x_0)(x - x_0) = 0 \tag{5}$$

thereby solving for the next iteration point x_1:

$$x_1 = x_0 - \frac{f'(x_0)}{f''(x_0)} \tag{6}$$

By iteratively applying the above procedure, we can progressively approach the extremal points of the function. The strength of Newton's method lies in the substantial reduction of the function value with each iteration, particularly as the optimal solution is approached. Nonetheless, it's important to acknowledge that Newton's method is not without limitations. Firstly, it may become trapped in local minima, impeding the identification of global minima. Secondly, the method necessitates the computation of second derivatives, which can become computationally complex and time-consuming in high-dimensional problems. To address these issues, researchers have proposed various improved variants of Newton's method, such as quasi-Newton methods, aimed at mitigating the constraints of the classical approach.

3.2. Steps and algorithm for finding minimum using Newton's method

Not all equations have closed-form solutions, or their solutions can be extremely complex, making them difficult to solve. Newton's method can be used for iterative solutions [5]. Consider the function $f(x) = x^2 + 2x + 1$. We will use Newton's method to find the local minimum of this function.

Step 1: Choose Initial Point Let's start with an initial point $x_0 = 2$.

Step 2: Compute First and Second Derivatives. First derivative of the function: $f'(x) = 2x + 2$.

Second derivative of the function: $f''(x) = 2$.

Step 3: Update Iteration Point Using Newton's method update formula, the new iteration point $x_1 = x_0 - f'(x_0)/f''(x_0)$:

$$x_1 = 2 - \frac{(2 * 2 + 2)}{2} = 1 - 2 = -1 \tag{7}$$

Step 4: Repeat Iterations We can continue iterating and compute x_2, x_3, and so on, until the convergence criteria are met.

Continuing the iterations: $x_2 = -0.5 x_3 = -0.25 \ldots$

As the iterations progress, the value of x will gradually approach the local minimum. Note that the effectiveness of iterations may depend on the choice of the initial point and the convergence criteria.

3.3. Analysis of convergence and convergence speed

The convergence analysis of Newton's iterative method includes local convergence and global convergence [6]. Newton's method theoretically exhibits quadratic convergence, which means that the error at each iteration diminishes quadratically as the solution approaches the optimum. However, the practical convergence can be influenced by several factors including the choice of initial point, the nature

DOI: 10.54254/2753-8818/9/20240764

of the function, and algorithmic parameters. The choice of the initial point can impact convergence. If the initial point is far from the optimal solution, more iterations might be needed to achieve convergence. A well-chosen initial point can expedite the convergence process.

The geometric properties and curvature of the function also affect convergence. Newton's method typically performs well for convex functions, as they have a single local minimum. However, in the presence of multiple local minima, Newton's method might converge to one of these local minima and might not escape it.

The convergence criteria of Newton's method need to be carefully set to avoid premature termination or excessive iteration. Common convergence criteria include a small change in the iteration point or reaching a certain number of iterations. Newton's method generally has a faster convergence speed compared to first-order optimization methods like gradient descent because it leverages second-order derivative information. However, in practice, the convergence speed can be influenced by the following factors:

Computational Cost: Computing second-order derivatives (Hessian matrix) can be computationally expensive, especially for high-dimensional problems. In such cases, each iteration of Newton's method can be costly.

Step Length Updates: The step length in Newton's method is determined by the second-order derivative. If the second-order derivative approaches zero at certain points, it can lead to excessively large or unstable step lengths, potentially causing divergence.

Global Convergence: Newton's method doesn't guarantee global convergence, so in some cases, it might converge to a local optimum instead of the global optimum.

4. Conclusion

This paper applied Newton's method to solve the problem of finding the minimum value of a function. Initially, we explored the process of transforming the problem of finding the minimum value of a function into an optimization problem, which provided us with a clear direction for our subsequent research. Subsequently, we elaborated on the specific steps and algorithm of utilizing Newton's method to solve for the minimum value of a function. This encompassed everything from problem modeling to the iterative solving process, enabling readers to gain a deep understanding of the practical application of this method. In this section, we also employed mathematical derivations and practical examples to illustrate the feasibility and effectiveness of the algorithm. An essential aspect of the paper focused on the analysis of convergence and convergence speed. We conducted an in-depth study of the convergence of Newton's method-based optimization problem, discussing whether the algorithm can converge to the global minimum and delving into the characteristics of the convergence process. Additionally, we emphasized the matter of convergence speed, analyzing the algorithm's iterative efficiency in various scenarios, and exploring potential optimization strategies. To sum up the findings of this research paper, we have gained a profound understanding of the utilization of Newton's method in solving function optimization problems. Newton's method, as a classical optimization algorithm, has demonstrated its remarkable prowess in tackling function minimum value problems. However, we also acknowledged the possibility of encountering convergence difficulties or slow convergence rates in certain scenarios. Therefore, future research could consider refining the algorithm further and integrating it with other optimization techniques to more effectively address practical problems. In conclusion, the insights provided by this research paper serve as a valuable reference and inspiration for delving deeper into the realm of function optimization.

References

[1] Dabing Doctor X, Deep Learning/Machine Learning Fundamentals in Mathematics, Convex Optimization, Hessian, Newton's Method, 2018/01/20, https://blog.csdn.net/xbinworld/article/details/79113218

[2] GlassySky0816, The principles and implementation steps of Newton's method, 2021/11/05, https://blog.csdn.net/qq_38784098/article/details/121170050

DOI: 10.54254/2753-8818/9/20240764

[3] Shengxin small rabbit, Newton's method for finding the minimum of a function, 2022/05/10, https://blog.csdn.net/weixin_60737527/article/details/124681773
[4] Norstc, The Newton's method for solving optimization problems in machine learning, 2018/08/04, https://blog.csdn.net/a493823882/article/details/81416213
[5] Mister Zhu, Mathematics and Algorithms: Two Applications of Newton's Method - Root Finding and Optimization, 2021/12/24, https://blog.csdn.net/u011754972/article/details/122124536
[6] Machine of heart, A Discussion on the Convergence of Newton's Iterative Method, 2021/10/17, https://blog.csdn.net/a735148617/article/details/120809458