International Journal of Education and Humanities
ISSN: 2770-6702 | Vol. 25, No. 2, 2026

Newton Iteration: Reconstructing Its Construction and Convergence Mechanism

Huiqin Chen*

School of Artificial Intelligence, Jianghan University, Wuhan, 430056, China
* **Corresponding author:** Huiqin Chen (Email: 1280542695@qq.com)

Abstract: For the single-variable nonlinear equation $f(x) = 0$, standard textbooks usually introduce Newton's method in the following way: the Taylor expansion is truncated to the first order, the remainder term is dropped to construct an iterative scheme, and an independent error analysis is subsequently performed to prove quadratic convergence. This "approximate first, verify later" narrative divides construction and proof into two separate lines of reasoning, loosely joined merely by the statement "neglect higher-order terms." Consequently, students struggle with two foundational questions: what justifies discarding the remainder to yield a valid iteration, and why does this specific approximation achieve quadratic convergence unlike alternative approximations? This paper proposes a reversed instructional sequence. First, we present the full Taylor expansion equipped with the second-order Lagrange remainder and clarify the order hierarchy between the linear leading part and remainder. Second, provided that the derivative does not vanish, the next iterate is rigorously defined as the unique root of the resulting linear equation. Third, we derive that the updated error is of second order with respect to the previous error. We conclude that quadratic convergence of Newton's method is not a fortuitous result confirmed retrospectively in the proof phase; instead, it is predetermined at the construction phase by the truncation order. By integrating the justification "why the approximation is admissible" during construction with the accuracy check in the proof stage, learners can grasp the core mechanism instead of retaining only superficial understanding.

Keywords: Newton's Method, Single-Variable Nonlinear Equation, Quadratic Convergence, Truncation Precision, Teaching Logic.

1. Introduction

The Problem: Everyone Can Do This Work, So Why Re-examine It?

For the single-variable nonlinear equation $f(x) = 0$, nearly all textbooks present Newton's iteration $x_{k+1} = x_k - \frac{f(x_k)}{f'(x_k)}$ with the explanation that "f is expanded into the first-order Taylor series at x_k and higher-order terms are neglected", followed by an independent convergence proof to establish quadratic convergence. This routine presentation is widely adopted and appears correct on the surface.

Nevertheless, a long-standing flaw exists within this teaching logic: the conventional paradigm follows the pattern of approximate first, verify later. A first-order linear approximation is constructed by dropping the remainder to obtain the iterative scheme, and a standalone error analysis is subsequently performed to prove quadratic convergence under suitable assumptions. Scheme construction and convergence proof constitute two separate lines of reasoning, hastily linked merely by the statement "neglect higher-order terms".

Naturally, students raise two key questions: what justifies dropping the remainder while still producing a valid iteration? Why does this particular approximation yield quadratic convergence, unlike other approximation strategies? If the term discarded during construction bears no intrinsic connection to the convergence order obtained in the proof, quadratic convergence becomes a property confirmed only in hindsight instead of one predetermined at the very moment the iteration is built.

This paper does not propose a brand-new numerical method. Instead, it revisits this well-known topic and transforms the teaching paradigm from "construct without full insight and then verify" into "grasp the essential principle prior to construction". Our core argument can be summarized in one sentence: The quadratic convergence of Newton's method is not obtained incidentally in the proof phase. Instead, it is predetermined during the construction phase by the truncation order. The order of the term discarded in the construction procedure governs the intrinsic convergence order of the iterative method.

In what follows, we complete the full derivation following the logical sequence: "clarify the order of the remainder → rigorously define the subsequent iterate → prove that the updated error is a second-order infinitesimal"[1].

2. Preliminary Tools: Put "What Order the Remainder Is" on the Table

A critical defect within conventional instruction is that the remainder is simply discarded without prior clarification of its mathematical nature. The starting point of our work is to write the remainder explicitly and identify its order.

Second-order Taylor expansion for single-variable functions (Lagrange remainder)[4]

Suppose $f(x)$ is twice continuously differentiable. It holds that

$f(x) = f(x_k) + f'(x_k)(x - x_k) + \frac{1}{2}f''(\xi)(x - x_k)^2$, where ξ lies between x_k and x. The remainder term $\frac{1}{2}f''(\xi)(x - x_k)^2$ is a quantity of second order with respect to the increment $(x - x_k)$. Two assumptions are required for this tool: f is twice continuously differentiable and $f'(x_k) \neq$

0 . The first condition ensures the existence and order identification of the remainder; the second guarantees the unique solvability of the next iterate.

3. The Construction Stage: Rather Than Arbitrarily Discarding the Remainder, We Explicitly Identify the Term to Be Dropped

3.1. Complete Decomposition

We perform the Taylor expansion at x_k with the complete second-order remainder:

$$f(x) = f(x_k) + f'(x_k)(x - x_k) + \frac{1}{2}f''(\xi)(x - x_k)^2$$

Define the linear function

$$l(x) = f(x_k) + f'(x_k)(x - x_k)$$

Rearrangement yields

$$f(x) - l(x) = R(x) = O((x - x_k)^2)$$

3.2. Core Idea Behind the Construction

The linear component $l(x)$ completely carries the first-order information, whereas the remainder $R(x)$is a pure second-order infinitesimal. The original nonlinear equation $f(x) = 0$ can be rewritten equivalently as $l(x) + R(x) = 0$.

This constitutes the fundamental justification for Newton's approximation: within a neighbourhood of the exact root, under the assumptions $f'(x) \neq 0$ and twice continuous differentiability of f, f'' is bounded on this neighbourhood. Provided that the increment is sufficiently small, the remainder represents a second-order quantity with respect to the increment, while the linear principal part is of first order. Consequently, substituting the original nonlinear equation with its linear counterpart discards only second-order infinitesimals while fully retaining all first-order information.

It must be emphasized that this procedure is not arbitrarily omitting higher-order terms. Instead, we first explicitly identify that the remainder is a second-order quantity, and only thereafter carry out the construction. Both the truncated term and its order are clearly stated before the iteration scheme is built.

3.3. Rigorous Definition of the Iterate

Solving $l(x) = 0$alone is insufficient; we must first verify that this equation uniquely determines a point. Provided $f'(x_k) \neq 0$ at the current iterate x_k, the linear equation

$$l(x) = 0, \quad i.e. \quad f(x_k) + f'(x_k)(x - x_k) = 0$$

admits exactly one solution. We define this unique solution to be the subsequent iterate x_{k+1}:

$$x_{k+1} = x_k - \frac{f(x_k)}{f'(x_k)}$$

Conclusion for the construction stage (validity argument only, convergence not yet proven): No first-order information is altered; merely second-order infinitesimals are neglected. Since the next iterate is uniquely solved from a linear equation under the non-vanishing derivative condition, this construction has the potential to achieve high-order convergence. This distinguishes Newton's method from naive rough approximations.

4. The Proof Stage: Verifying That the Truncated Term Is Indeed Second-Order and That the Updated Error Depends Quadratically on the Old Error

His section exclusively employs the iterate defined legitimately in Section 3: x_{k+1} satisfies $l(x_{k+1}) = 0$. No extra assumptions are introduced.

(1) Taylor expansion evaluated at the true root

Let x^* denote the true root satisfying$f(x^*) = 0$. Expand f about x_k:

$$0 = f(x^*) = f(x_k) + f'(x_k)(x^* - x_k) + \frac{1}{2}f''(\xi_k)(x^* - x_k)^2.$$

where ξ_k lies strictly between x_k and x^*.

(2) Subtract the iteration equation

From our construction rule,

$$0 = f(x_k) + f'(x_k)(x_{k+1} - x_k)$$

Subtract the two equations. Denote iteration $e_k = x_k - x^*$. Notice $x^* - x_k = -e_k$; the negative sign vanishes after squaring inside the quadratic remainder:

$$f'(x_k)(x^* - x_{k+1}) = -\frac{1}{2}f''(\xi_k)e_k^2$$
$$|x^* - x_{k+1}| = O(e_k^2) \quad i.e. \quad e_{k+1} = O(e_k^2).$$

The left-hand side represents the updated error increment, and the right-hand side is quadratic in the prior error, a second-order infinitesimal[2][3].

5. Conclusion: Why We Re-examine This Derivation

Returning to the opening question: this derivation is widely known, so why reconstruct it step by step?

Traditional teaching often adopts a shortsighted treatment: students are taught to perform a first-order approximation, drop the remainder informally to build an iteration formula, and then run a separate error analysis to confirm quadratic convergence. Since the nature of the discarded term is never clarified during scheme construction, the quadratic convergence order appears as an afterthought appended in the proof. Students memorize the statement "Newton's method converges quadratically," yet fail to grasp the origin of this second-order convergence.

This paper reverses the conventional teaching sequence:

(1) Write out the complete Taylor expansion and separate the linear principal part from the second-order remainder;

(2) Clarify the truncation level: only second-order small quantities are discarded while all first-order information is preserved - this grounds the validity of our construction;

(3) Define iterates via linear equations, with uniqueness guaranteed under non-zero derivative to secure rigorous construction;

(4) Compare successive errors and rigorously derive that the new error is a second-order quantity of the previous error.

From this perspective, quadratic convergence is not discovered accidentally at the proof stage. Instead, it is predetermined at the construction stage by the precision of our truncation strategy. If only second-order terms are discarded, the resulting error will be quadratic in the old error. By contrast, any construction that damages the first-order principal part can at most yield first-order convergence. Only by linking the justification for approximation at the construction phase to accuracy validation in the proof phase can learners capture the core mechanism rather than superficial fragments.

The value of this reconstruction lies not in disputing the well-known final result, but in connecting construction logic with convergence mechanism. Students thereby recognize that quadratic convergence of Newton's method is already determined at the very beginning of Taylor expansion instead of being a coincidental outcome obtained in the final verification step.

Supplementary Note: Extension to Multivariate Functions

Multivariate functions also admit Taylor expansions, which enables us to generalize the whole reasoning by analogy. This reveals the universality of our core principle: the truncation order selected during construction governs the achievable convergence order proven afterwards, and the rule applies beyond single-variable scenarios. Consider a scalar-valued multivariate function $f(\boldsymbol{x})$. Its second-order Taylor expansion centered at the current iterate $\boldsymbol{x}_k$ reads

$$f(\boldsymbol{x}) = f(\boldsymbol{x}_k) + \nabla f(\boldsymbol{x}_k)^T(\boldsymbol{x} - \boldsymbol{x}_k) + \frac{1}{2}(\boldsymbol{x} - \boldsymbol{x}_k)^T \boldsymbol{H}(\boldsymbol{\xi})(\boldsymbol{x} - \boldsymbol{x}_k),$$

where $\nabla f(\boldsymbol{x}_k)$ stands for the gradient vector and $\boldsymbol{H}(\boldsymbol{\xi})$ denotes the Hessian matrix.

We follow exactly the same reasoning as in the one-dimensional setting: retain the linear principal part and drop the second-order remainder. Setting the next iterate as the unique solution to the linearized equation

$$\nabla f(\boldsymbol{x}_k)^T(\boldsymbol{x} - \boldsymbol{x}_k) + f(\boldsymbol{x}_k) = 0,$$

we arrive at the iteration formula (the Newton system for vector-valued equations $\boldsymbol{F}(x) = 0$):

$$\boldsymbol{x}_{k+1} = \boldsymbol{x}_k - \boldsymbol{H}^{-1}(\boldsymbol{x}_k)\nabla f(\boldsymbol{x}_k).$$

Similarly, the updated error $\boldsymbol{e}_{k+1}$ is a second-order quantity with respect to $\boldsymbol{e}_k$. This confirms that multivariate Newton's method naturally inherits quadratic convergence from the identical truncation principle.

References

[1] Deuflhard, P. (2011). *Newton methods for nonlinear problems: Affine invariance and adaptive algorithms*. Springer-Verlag.

[2] Higham, N. J. (2002). *Accuracy and stability of numerical algorithms* (2nd ed.). Society for Industrial and Applied Mathematics (SIAM).

[3] Ortega, J. M., & Rheinboldt, W. C. (1970). *Iterative solution of nonlinear equations in several variables*. Academic Press.

[4] Taylor, A. E., & Mann, W. R. (1983). *Advanced calculus* (3rd ed.). John Wiley & Sons.