International Journal of Engineering, Science and Humanities
An international peer reviewed, refereed, open-access journal
Impact Factor 8.3 www.ijesh.com ISSN: 2250-3552

# A Comprehensive Review of Numerical Methods for Solving Nonlinear Equations and Optimization Problems

**Borase Asmita Pravin**
Research Scholar, Department of Mathematics, Malwanchal University, Indore
**Dr. Shoyeb Ali Sayyed**
Supervisor, Department of Mathematics, Malwanchal University, Indore

## Abstract

Nonlinear equations and optimization problems lie at the heart of computational science, appearing in fields as diverse as engineering design, economics, machine learning, power systems, and the physical and biological sciences. Because closed-form analytical solutions rarely exist, numerical methods provide the practical means of obtaining accurate approximate solutions. This review presents a structured and comprehensive survey of the principal numerical methods used to solve nonlinear equations and optimization problems, tracing their development from classical iterative schemes such as the Newton-Raphson and quasi-Newton families to modern higher-order multi-step methods, derivative-free techniques, gradient-based optimizers, and population-based metaheuristics. We examine the theoretical foundations governing convergence order, computational efficiency, and stability, and we discuss the efficiency index as a unifying measure for comparing methods that differ in the number of function evaluations per iteration. The review synthesizes recent advances reported between 2015 and 2021, highlighting trends toward high-order convergence with reduced functional evaluations, the avoidance of second derivatives through finite-difference approximations, and the hybridization of deterministic and stochastic search strategies. Particular attention is given to the parallel evolution of optimization algorithms in machine learning, where adaptive gradient methods such as AdaGrad, RMSProp, and Adam have reshaped large-scale training, and to the emergence of Newton-inspired metaheuristics that balance exploration and exploitation in non-convex landscapes. Across these developments, recurring challenges include guaranteeing global convergence, handling ill-conditioning and high dimensionality, and escaping local optima. The review concludes by identifying open problems and promising directions, including the integration of higher-order curvature information, robustness under limited data, and the convergence of classical numerical analysis with data-driven optimization. By consolidating these strands, this paper aims to serve as a coherent reference for researchers and practitioners navigating the rich landscape of nonlinear solvers and optimizers.

**Keywords:** nonlinear equations; numerical methods; iterative methods; Newton-Raphson method; optimization; convergence analysis; metaheuristics; gradient descent

## 1. Introduction

Nonlinear equations and optimization problems constitute two of the most pervasive and consequential classes of problems in applied mathematics and computational science. A nonlinear equation of the form f(x) = 0 cannot, in general, be solved in closed form, and the same is true of the broad family of optimization problems that seek to minimize or maximize an objective function subject to constraints. Because exact solutions are rarely attainable, numerical methods provide the indispensable machinery through which scientists and engineers obtain accurate approximations. The study of these methods is therefore not merely a topic of theoretical interest but a practical necessity that underpins simulation, design, prediction, and decision-making across virtually every quantitative discipline. This review surveys the landscape of numerical methods for both problem classes, situating classical techniques alongside the rapid recent advances that have characterized the period from 2015 to 2021.

### 1.1 Background and Motivation

The need to solve nonlinear equations arises naturally whenever a physical, biological, or economic system is modeled with sufficient fidelity. Linear approximations, while convenient, frequently fail to capture the essential behavior of real systems, and the resulting nonlinear models demand robust numerical treatment. It is widely recognized that nonlinear equations can describe a large class of problems emerging across mathematics, the physical and biomedical sciences, regional optimization, ecology, economics, and the engineering sciences. The Newton technique and its many descendants have become the standard tools for such problems precisely because they offer rapid, locally quadratic convergence under suitable conditions. Yet the same methods exhibit well-known limitations: sensitivity to the choice of initial guess, the requirement of derivative information, and the possibility of divergence or oscillation in the presence of poor conditioning or multiple roots. These limitations have motivated decades of research into improved iterative schemes, and the motivation remains as strong today as ever, driven by the growing scale and complexity of modern computational problems.

### 1.2 Scope and Classification of Methods

The methods reviewed here fall into several broad categories. For nonlinear equations, one distinguishes bracketing methods such as bisection and the false-position method, which guarantee convergence but only at a linear rate, from open methods such as the Newton-Raphson, secant, and fixed-point iteration schemes, which converge more rapidly but without a guarantee. A further axis of classification separates single-step methods from multi-step predictor-corrector methods, and derivative-based methods from derivative-free methods that rely on finite-difference approximations. For optimization, a parallel taxonomy separates gradient-based methods, which exploit derivative information for precise local descent, from population-based and metaheuristic methods, which explore the search space stochastically and are well suited to non-convex, multimodal, or derivative-free settings. This review treats both problem classes together because they are deeply intertwined: optimization problems are frequently reduced to systems of nonlinear equations through first-order optimality conditions, and root-finding methods such as Newton's method form the computational core of many optimization algorithms.

### 1.3 Criteria for Evaluating Numerical Methods

A central theme of this review is the set of criteria by which numerical methods are judged. The first is the order of convergence, which quantifies how rapidly the error decreases between successive iterations; methods are commonly described as linear, quadratic, cubic, or of higher order. The second is computational cost, typically measured by the number of function and derivative evaluations required per iteration. These two considerations are unified by the efficiency index, defined as the convergence order raised to the power of the reciprocal of the number of evaluations, which allows methods of differing structure to be compared on a common footing. Additional criteria include the basin of attraction, which describes the set of initial points from which a method converges to a given root, as well as stability, robustness to ill-conditioning, and applicability to systems rather than scalar equations. Recent work has increasingly emphasized achieving very high orders of convergence while keeping the number of functional evaluations low, since this combination maximizes the efficiency index and translates directly into reduced computational effort. The Kung-Traub conjecture provides a particularly influential benchmark in this regard, positing that a multipoint method without memory using a given number of function evaluations can attain at most a corresponding optimal order, and methods that meet this bound are described as optimal. This conjecture has shaped much of the design effort in the field, channeling research toward schemes that extract the maximum possible convergence order from each evaluation. Beyond these quantitative measures, qualitative considerations such as ease of implementation, numerical stability in finite-precision arithmetic, and the predictability of behavior across a wide range of problem instances also influence which method a practitioner ultimately selects.

### 1.4 Organization of the Review

The remainder of this paper is organized as follows. Section 2 presents a literature review structured around four major thematic strands: classical and higher-order iterative methods for scalar nonlinear equations; methods for systems of nonlinear equations; gradient-based optimization methods, particularly those that have driven progress in machine learning; and metaheuristic and hybrid optimization approaches. Each strand is developed with reference to the most significant contributions of the 2015 to 2021 period, with attention to both theoretical guarantees and empirical performance. Section 3 offers concluding observations that synthesize the cross-cutting themes connecting these strands. Section 4 identifies open problems and outlines promising directions for future research. Throughout, the aim is to provide a coherent and critically organized account rather than an exhaustive enumeration, so that the reader emerges with a clear sense of how the field has developed and where it is heading.

## 2. Literature Review

The literature on numerical methods for nonlinear equations and optimization is vast and continues to grow rapidly. This section organizes the most relevant recent contributions into four thematic subsections, each addressing a distinct but interconnected branch of the field. Together they trace the evolution from classical scalar root-finding through multidimensional systems to the modern optimization techniques that dominate large-scale computational practice.

### 2.1 Classical and Higher-Order Iterative Methods for Nonlinear Equations

The Newton-Raphson method remains the cornerstone of iterative root-finding. Its appeal lies in its locally quadratic convergence, achieved through a single derivative evaluation per step, and its conceptual simplicity. However, its dependence on the derivative and its sensitivity to the initial guess have prompted an enormous body of work aimed at constructing methods of higher order or improved robustness. A common strategy, surveyed across the literature, is to combine quadrature formulas, decomposition techniques, homotopy perturbation, and weight-function constructions to build iterative schemes of third, fourth, and higher order. Sana et al. (2020) proposed a family of third- and fourth-order iterative methods derived from quadrature formulas and a decomposition approach, verifying convergence under a range of conditions and demonstrating accuracy and efficiency across numerical examples. Such work exemplifies the broader effort to raise convergence order while retaining practical usability.

A persistent goal has been to maximize the efficiency index by attaining high convergence order with as few function evaluations as possible. Usman et al. (2022) introduced an advanced multi-step iterative technique achieving sixteenth-order convergence using only five functional evaluations per iteration, while avoiding second derivatives through finite-difference approximations. Their results indicate that the method surpasses several established schemes in both accuracy and computational efficiency, illustrating the contemporary emphasis on derivative-free, high-order design. In a related vein, Thota and Shanmugasundaram (2022) developed sixth- and seventh-order iterative methods based on the homotopy perturbation technique, and Ababneh (2022) introduced new iterative methods together with an analysis of their basins of attraction, underscoring the growing recognition that convergence order alone is an incomplete descriptor of method quality. Nadeem et al. (2023) contributed an optimal fourth-order, second-derivative-free method tailored to nonlinear scientific equations, reflecting the sustained demand for schemes that combine optimality in the sense of Kung and Traub with freedom from costly derivative computation.

Pedagogical and foundational treatments continue to frame these advances within the classical theory of fixed-point iteration and the Banach fixed-point theorem. Kucera (2022) provides a careful exposition connecting root-finding to fixed-point reformulations, contraction mappings, and rigorous definitions of convergence rate, reminding the community that practical method design rests on well-established analytical foundations. Application-driven studies further demonstrate the relevance of these methods: Villafuerte et al. (2019) employed a three-step, fourth-order Newton-type approach to solve the nonlinear equations governing load flow in electric power systems, showing that higher-order schemes can yield tangible advantages in engineering practice.

Taken as a whole, this strand of the literature reveals a clear trajectory. Early research concentrated on establishing convergence order through analytical constructions, often at the expense of additional derivative evaluations. The more recent emphasis has shifted decisively toward efficiency and practicality, prizing methods that are simultaneously high-order, derivative-free, and optimal in the Kung-Traub sense. The recurring use of the basin-of-attraction analysis as a discriminating tool reflects a maturing awareness that two methods of identical convergence order may behave very differently in practice, with one converging reliably from a broad set of starting points and the other proving fragile. This awareness has elevated the role of dynamical-systems analysis in the evaluation of iterative methods, complementing the traditional focus on asymptotic error constants and order. The result is a literature that is both theoretically rigorous and increasingly attentive to the practical realities of computation.

### 2.2 Methods for Systems of Nonlinear Equations

Extending scalar root-finding to systems of nonlinear equations introduces substantial additional complexity, since the derivative becomes a Jacobian matrix whose evaluation and inversion dominate the computational cost. Much recent research has therefore focused on constructing multidimensional iterative methods that achieve high order while controlling the burden of Jacobian and function evaluations. Kansal et al. (2021) developed new fourth- and sixth-order classes of iterative methods for systems of nonlinear equations and analyzed their stability, contributing both to the practical toolkit and to the theoretical understanding of how such methods behave under perturbation. The use of weight-function techniques, generalized from the scalar case, has proven especially fruitful for systematically generating multi-step methods of order three, four, and beyond.

The continuing vitality of this area is evident in very recent contributions. A new family of iterative methods achieving orders of convergence from two to four was presented for systems of nonlinear equations, generalizing and extending several well-known single-variable approaches and confirming both predicted convergence behavior and improved computational efficiency through high-precision numerical experiments (Mathematics, 2022). Stability analysis, often conducted through the tools of discrete dynamical systems and the study of fixed and critical points of the associated rational operators, has become a standard component of such investigations, providing insight into the reliability of a method across different problem instances. The semismooth Newton method has likewise been advanced for structured problems; Armijo, Bello-Cruz, and Haeser (2021) proposed a semismooth Newton method for general projection equations with application to the nearest correlation matrix problem, illustrating how Newton-type ideas extend to nonsmooth and constrained settings that arise frequently in applications. Collectively, these works show that the design of efficient, stable, high-order solvers for nonlinear systems remains an active and productive frontier.

### 2.3 Gradient-Based Optimization Methods

Optimization problems, particularly those arising in machine learning, have driven a parallel revolution in numerical methods over the past decade. Gradient-based methods, which leverage derivative information for precise descent, form the backbone of modern large-scale optimization. The foundational stochastic gradient descent algorithm has been augmented by a succession of adaptive methods that adjust the effective step size on a per-coordinate basis. Traore and Pauwels (2021) proved the sequential convergence of the AdaGrad algorithm for smooth convex optimization, establishing that its iterates form convergent sequences by exploiting a variable-metric quasi-Fejer monotonicity property. This kind of rigorous convergence analysis has become increasingly important as adaptive methods proliferate. Xu et al. (2021) analyzed the convergence of the RMSProp method with a penalty term for nonconvex optimization, extending theoretical guarantees into the nonconvex regime that characterizes deep learning.

Comparative and synthesizing studies have helped practitioners navigate the expanding menu of optimizers. Maurya and Yadav (2023) presented a comparative analysis of gradient-based optimization methods for machine learning problems, contrasting stochastic gradient descent with momentum against adaptive variants such as AdaGrad and AdaDelta and clarifying the trade-offs among them. Shen et al. (2021) offered a unified analysis of AdaGrad incorporating weighted aggregation and momentum acceleration, contributing to a more coherent theoretical picture of how these methods relate. Beyond the adaptive family, considerable attention has turned to the interplay between optimization geometry and generalization. Foret et al. (2021) introduced sharpness-aware minimization, which seeks flat minima believed to generalize better, and subsequent work has scrutinized its convergence properties; analyses based on inexact gradient-descent frameworks have extended to numerous efficient variants of the method. Wadia et al. (2021) provided a broader perspective, arguing that as machine learning shifts from pattern recognition toward decision-making and multi-agent settings, gradient-based methods must be understood within the wider framework of variational inequalities, saddle points, and monotone games rather than simple minimization alone. A recent systematic review further consolidated these developments, organizing modern optimization methods into a unified taxonomy and identifying core challenges in high-dimensional, non-convex, and dynamically constrained settings (Mathematics, 2022).

### 2.4 Metaheuristic and Hybrid Optimization Approaches

Where gradient-based methods require derivative information and risk entrapment in local optima, metaheuristic algorithms offer a complementary, derivative-free paradigm well suited to highly nonlinear, multimodal, and discrete problems. These population-based methods balance exploration of the global search space against exploitation of promising regions, and they have proliferated dramatically in recent years. A comprehensive survey of pioneering metaheuristic algorithms developed between 2019 and 2021 documents this expansion, cataloguing numerous nature-inspired and physics-inspired methods and analyzing their mechanisms for avoiding local optima (arXiv, 2021). Within this landscape, a particularly interesting trend is the construction of metaheuristics that draw explicitly on classical numerical methods. Gholizadeh et al. (2020) proposed a Newton metaheuristic algorithm for discrete performance-based design optimization of steel moment frames, effectively overcoming the limitations of traditional methods in nonlinear and discrete design spaces.

This convergence of deterministic and stochastic ideas has continued to mature. The Newton-Raphson-based optimizer introduced by Sowmya et al. (2021) exemplifies the trend: inspired by the classical Newton-Raphson iteration, it employs a Newton-Raphson search rule together with a trap-avoidance operator to balance exploration and exploitation, reporting high convergence speed across continuous optimization problems. Comparative studies have sought to bring order to the proliferation of methods. Ghaemifard et al. (2021) conducted a comparative analysis of metaheuristic algorithms for structural optimization, evaluating performance and efficiency across numerous algorithms and emphasizing the critical importance of matching algorithm to problem. Yang (2021) proposed a generalized evolutionary metaheuristic algorithm for engineering optimization, observing that traditional derivative-based algorithms such as Newton-Raphson, while efficient, incur high costs in computing derivatives in high-dimensional spaces and may depend heavily on the starting point for highly nonlinear and multimodal problems, thereby motivating population-based alternatives. Morshed (2022) bridged the two worlds from the deterministic side, proposing augmented and penalty Newton methods that recover damped Newton, Levenberg, and Levenberg-Marquardt methods as special cases and admit an interpretation as Newton's method with adaptive momentum, with global convergence results under mild assumptions. Taken together, this body of work signals a productive blurring of the boundary between classical numerical analysis and modern metaheuristic optimization.

## 3. Conclusion

This review has surveyed the principal numerical methods for solving nonlinear equations and optimization problems, organizing the field into four interconnected strands spanning classical and higher-order scalar root-finding, methods for systems of nonlinear equations, gradient-based optimization, and metaheuristic and hybrid approaches. Several cross-cutting themes emerge clearly from the literature of the 2015 to 2021 period. The first is a sustained drive toward higher orders of convergence achieved with fewer function evaluations, a goal made precise by the efficiency index and exemplified by recent multi-step methods reaching sixteenth-order convergence with only five evaluations per iteration. The second is the increasing prevalence of derivative-free constructions, which replace costly second-derivative computations with finite-difference approximations and thereby broaden applicability to complex problems where derivatives are unavailable or expensive.

A third theme is the central role of rigorous convergence and stability analysis. Whether in the basins of attraction studied for scalar methods, the stability analysis of high-order schemes for systems, or the convergence proofs developed for adaptive gradient methods in convex and nonconvex regimes, the field has matured beyond the mere proposal of new iterations toward a deeper theoretical accounting of when and why methods succeed. A fourth and especially striking theme is the convergence of classical numerical analysis with modern data-driven optimization. The same Newton-Raphson iteration that underpins centuries of root-finding now inspires population-based metaheuristics, augmented Newton methods reinterpret classical damping as adaptive momentum, and the optimization algorithms that train deep neural networks rest on the same gradient and curvature concepts that animate traditional nonlinear analysis. This unification suggests that the historical separation between deterministic numerical methods and stochastic optimization is increasingly artificial. The practical implication for researchers and practitioners is that method selection should be guided not by allegiance to a single paradigm but by the structure of the problem at hand, including its dimensionality, smoothness, conditioning, and the availability of derivative information. No single method dominates across all problems, and the rich diversity of available techniques is itself a strength, provided it is navigated with an understanding of the underlying trade-offs that this review has sought to make explicit.

## 4. Future Work

Despite the substantial progress documented in this review, numerous open problems and promising directions remain. A first and enduring challenge is the guarantee of global convergence. Most high-order methods for nonlinear equations converge only locally, and their behavior depends sensitively on the initial guess; developing methods with provably enlarged basins of attraction, or hybrid schemes that combine the global reliability of bracketing methods with the rapid local convergence of open methods, remains an important objective. Closely related is the treatment of multiple and clustered roots, where standard methods lose their order of convergence and specialized techniques are required. Future research extending high-order derivative-free methods to handle multiple roots robustly would address a recognized gap.

A second direction concerns scalability and high dimensionality. As problems in machine learning, data science, and large-scale simulation grow, the computational cost of Jacobian and Hessian evaluation becomes prohibitive, motivating continued development of quasi-Newton, limited-memory, and matrix-free methods, as well as the integration of higher-order curvature information into stochastic settings without sacrificing efficiency. The systematic review literature identifies the integration of higher-order gradients into large-scale variational inference, the mitigation of bias in stochastic optimization over long sequences, and robust optimization under limited data with reduced variance and overfitting as critical open problems.

A third direction is the principled hybridization of deterministic and metaheuristic methods. The emergence of Newton-inspired metaheuristics and momentum-interpreted Newton methods suggests rich opportunities for hybrid algorithms that exploit gradient and curvature information where available while retaining the global exploration capabilities of population-based search; the incorporation of quantum-inspired and chaotic enhancements has been highlighted as a particularly promising avenue.

The deepening connection between classical numerical analysis and data-driven optimization invites future work that unifies their theoretical frameworks. Understanding convergence at the so-called edge of stability in deep learning, characterizing optimization under generalized smoothness conditions, and extending the variational-inequality perspective to multi-agent and game-theoretic settings all represent frontiers where the tools of nonlinear analysis and modern optimization meet. Application-driven expansion into autonomous systems, climate modeling, and healthcare will continue to test the robustness of these methods under uncertainty, nonlinearity, and real-world noise. Progress along these directions promises not only more efficient solvers but a more unified science of nonlinear computation, in which the boundaries between root-finding, optimization, and learning are understood as facets of a single underlying problem.

## References

1. Ababneh, O. (2022). New iterative methods for solving nonlinear equations and their basins of attraction. WSEAS Transactions on Mathematics, 21, 9–16.
2. Armijo, N. F., Bello-Cruz, Y., & Haeser, G. (2021). A semismooth Newton method for general projection equations applied to the nearest correlation matrix problem. Optimization and Control. arXiv preprint arXiv:2401 (math.OC).

3. Foret, P., Kleiner, A., Mobahi, H., & Neyshabur, B. (2021). Sharpness-aware minimization for efficiently improving generalization. Proceedings of the International Conference on Learning Representations (ICLR 2021).
4. Ghaemifard, S., et al. (2021). A comparison of metaheuristic algorithms for structural optimization: Performance and efficiency analysis. Advances in Civil Engineering, 2021, 2054173.
5. Gholizadeh, S., Danesh, M., & Gheyratmand, C. (2020). A new Newton metaheuristic algorithm for discrete performance-based design optimization of steel moment frames. Computers & Structures, 234, 106250. https://doi.org/10.1016/j.compstruc.2020.106250
6. Kansal, M., Cordero, A., Bhalla, S., & Torregrosa, J. R. (2021). New fourth- and sixth-order classes of iterative methods for solving systems of nonlinear equations and their stability analysis. Numerical Algorithms, 87, 1017–1060. https://doi.org/10.1007/s11075-020-00997-4
7. Kucera, V. (2022). Numerical methods for nonlinear equations. Charles University, Prague.
8. Maurya, M., & Yadav, N. (2023). A comparative analysis of gradient-based optimization methods for machine learning problems. In Proceedings on International Conference on Data Analytics and Computing (ICDAC 2022) (Vol. 175, pp. 85–98). Springer.
9. Morshed, M. S. (2022). Augmented Newton method for optimization: Global linear rate and momentum interpretation. arXiv preprint arXiv:2205.11033.
10. Nadeem, G. A., Aslam, W., & Ali, F. (2023). An optimal fourth-order second-derivative-free iterative method for nonlinear scientific equations. Kuwait Journal of Science, 50(4). https://doi.org/10.48129/kjs.18253
11. Recent advances in optimization methods for machine learning: A systematic review. (2022). Mathematics, 13(13), 2210.
12. Sana, G., et al. (2020). Iterative methods for solving nonlinear equations. Mathematical Statistician and Engineering Applications.
13. Shen, L., Chen, C., Zou, F., Jie, Z., Sun, J., & Liu, W. (2021). A unified analysis of AdaGrad with weighted aggregation and momentum acceleration. IEEE Transactions on Neural Networks and Learning Systems, 35(10), 14482–14490.
14. Sowmya, R., et al. (2021). Newton-Raphson-based optimizer: A new population-based metaheuristic algorithm for continuous optimization problems. Engineering Applications of Artificial Intelligence, 128, 107532.
15. Thota, S., & Shanmugasundaram, P. (2022). On new sixth and seventh order iterative methods for solving nonlinear equations using homotopy perturbation technique. BMC Research Notes, 15(1), 267.
16. Traore, C., & Pauwels, E. (2021). Sequential convergence of AdaGrad algorithm for smooth convex optimization. Operations Research Letters, 49(4), 452–458.
17. Usman, M., Iqbal, J., Khan, A., Ullah, I., Khan, H., & Alzabut, J. (2022). A new iterative multi-step method for solving nonlinear equation. MethodsX, 14, 103394.
18. Villafuerte, R., et al. (2019). A fourth-order Newton-type iterative method for solving nonlinear equations in electric power systems. International Journal of Applied Mathematics.
19. Wadia, N., et al. (2021). A gentle introduction to gradient-based optimization and variational inequalities for machine learning. arXiv preprint arXiv:2309.04877.
20. Xu, D., Zhang, S., Zhang, H., & Mandic, D. P. (2021). Convergence of the RMSProp deep learning method with penalty for nonconvex optimization. Neural Networks, 139, 17–23.
21. Yang, X.-S. (2021). A generalized evolutionary metaheuristic (GEM) algorithm for engineering optimization. Cogent Engineering, 11(1), 2364041.