Proceedings of the 2023 International Conference on Mathematical Physics and Computational Simulation
DOI: 10.54254/2753-8818/10/20230346

Analysis of applications for Taylor series expansion: Evidence from machine learning, mathematics and engineering

Ding He

Department of mathematics, University of Nottingham, Ningbo, 315100, China

ssydh2@nottingham.edu.cn

Abstract. Contemporarily, the Taylor series expansion is a common mathematical approach that is widely applied in various fields. This study provides a comprehensive overview of the Taylor expansion, a powerful mathematical tool used to represent continuous and differentiable functions as an infinite sum of terms around a specific point. This paper delves into the historical context and the fundamental concepts of the Taylor expansion, followed by its applications in machine learning, mathematics, and engineering. Furthermore, the research highlights the advantages and disadvantages of using the Taylor expansion in different scenarios. According to the analysis, it also discusses recent advancements in utilizing the Taylor expansion to improve computational efficiency and problem-solving across these three fields. In conclusion, despite its limitations, the Taylor expansion has significantly contributed to scientific advancements and will continue to play a vital role in diverse research areas. Overall, these results shed light on guiding further exploration of the Taylor expansion.

Keywords: symplectic Taylor neural networks, differential equations, groundwater models.

1. Introduction

The Taylor expansion is extensively used in machine learning, mathematics, and engineering. Taylor neural networks, within the realm of the symplectic algorithm, offer continuous, long-term predictions for complex Hamiltonian dynamics from sparse, transient input. The method's innovative architecture comprises two sub-networks, both with Taylor series expansion components arranged symmetrically. The expressiveness and symmetry of the expansion aid in fitting Hamiltonian gradients and preserving the symplectic structure. The Taylor-Net architecture and the neural ODEs framework are combined with a fourth order symplectic integrator to learn continuous-time system evolution while maintaining symplectic properties. Despite limited training data and time, the model demonstrates remarkable computational advantages, outperforming previous methods in accuracy, convergence rate, and resilience [1].

Meanwhile, the Taylor expansion plays an important role in differential equations. Numerical methods are frequently employed to handle differential equations that cannot be solved analytically or in closed form [2]. Although numerical methods are commonly used to solve differential equations, they can encounter issues when dealing with non-linearities, higher orders, or stiffness, which may lead to complete failure or significant accumulation of numerical errors. There is a method for solving ordinary differential equations that makes use of evolutionary algorithms. With the help of this technique, one

DOI: 10.54254/2753-8818/10/20230346

may define the concept of a Taylor series matrix, which turns a differential equation into an optimization problem. The coefficients of a series expansion make up the objective function in this problem [3].

A number of numerical algorithms used in engineering rely on the finite difference approximation, which is derived from the Taylor series [4]. For example, groundwater models are utilized to simulate the flow of fluids and the transport of contaminants in the geological subsurface. They are crucial tools for understanding and predicting how water resource systems will react to planned and actual use as well as pollution prevention initiatives. These models contain many variables that have a significant influence on how well the model simulation works. However, not all these parameters can be precisely calibrated against the scarce observational data since many of them are unobservable. Sensitivity analysis (SA) is necessary to identify the sensitive parameters for model calibration. Uncertainty quantification (UQ) is therefore essential to enable informed decision-making based on scientific analysis [5]. Taylor expansion-based adaptive design (TEAD), a method that seeks to lower the computational cost associated with building global surrogate models by adapting the Taylor expansion approach, has been developed by a group of researchers to improve the computational efficiency of surrogate-based global sensitivity analysis (GSA) and uncertainty quantification (UQ) [6].

The Taylor expansion is a mathematical tool that can be used to represent a continuous and differentiable function as an infinite sum of terms around a specific point. It has been widely applied to solve complex problems in various fields such as economics, physics, and engineering. However, the Taylor expansion only applies in the neighbourhood of the unfolded point and may have convergence problems at some points. The Taylor expansion only considers the local behaviour of the function at the unfolded point, not the global behaviour. As a result, care needs to be taken when using Taylor expansions in terms of the accumulation of errors, especially at distant points in the vicinity of the unfolded point. The coefficients of Taylor expansions require the calculation of derivatives of all orders of the function, which may be difficult to calculate for some functions of higher order, limiting the scope of Taylor expansions. This paper will focus on the advantages and disadvantages of Taylor expansion and choose research from three fields, machine learning, mathematics, and engineering, which start from a basic description to application cases, ending with evaluation.

2. Basic descriptions

Zeno of Elea, an ancient Greek philosopher, explored the concept of adding up an infinite series to obtain a finite outcome but ultimately dismissed it as an impossibility [7]. This inquiry led to Zeno's paradox. Aristotle later offered a philosophical solution to the paradox, but the mathematical aspect remained unresolved until Archimedes addressed it. Similarly, Democritus, a Presocratic Atomist, had also tackled the issue before Aristotle. Using the method of exhaustion, Archimedes was able to perform an infinite number of progressive subdivisions and achieve a finite result [8]. A similar technique was independently used a few decades later by Liu Hui [9].

In the 14th century, Madhava of Sangamagrama produced some of the earliest examples of Taylor series (but not the general approach) [10]. While there are no written records of his work, his disciples in the Kerala school of astronomy and mathematics claim that he discovered the Taylor series for the sine, cosine, and arctangent trigonometric functions (known as Madhava series). His followers proceeded to create new series extensions and logical approximations over the following 200 years.

Gregory received a letter from John Collins in the latter half of 1670 that contained various Maclaurin series derived by Isaac Newton, including sin x, cos x, arcsin x, and x cot x. Collins informed Gregory in the letter that Newton had created a generic technique for extending functions in series. Unaware that Newton's method included a challenging procedure of term-by-term integration and lengthy division of series, Gregory set out to find a general approach on his own. After creating a series similar to the general Maclaurin series that comprised series for arctan, tan x, and sec x, Gregory addressed a letter to Collins at the beginning of 1671. Gregory created the series for arctan, tan x, and sec x, but didn't explain how he did it since he believed he had just uncovered Newton's approach. It is only hypothesized that he grasped the overall procedure by looking at some scrawled notes he had made on the back of another letter dated 1671 [11].

DOI: 10.54254/2753-8818/10/20230346

Despite finding the series for arctan, tan x, and sec x at the beginning of 1671 and including them in a letter to Collins, Gregory thought he had just rediscovered Newton's approach and did not explain how he arrived at these series. According to a popular theory, Gregory may have recognized the basic approach by looking at some scribbles he had made on the back of another letter dated 1671.

The Maclaurin series is named for Colin Maclaurin, an Edinburgh professor who, in the middle of the 18th century, produced the particular formulation of the Taylor theorem[12].

The power series is the Taylor series of a function with real or complex values f (x) that can be endlessly differentiated at a real or complex number a,

$$f(a)+\frac{f'(a)}{1!}(x-a)^1+\frac{f''(a)}{2!}(x-a)^2+\frac{f'''}{3!}(x-a)^3+\cdots \tag{1}$$

Or it can be written as:

$$\sum_{n=0}^{\infty}\frac{f^{(n)}(a)}{n!}(x-a)^n \tag{2}$$

where $f^{(n)}(a)$ denotes the *n*th derivate of *f* evaluated at the point *a*. (The derivative of order zero of *f* is defined to be *f* itself and $(x-a)^0$ and 0! are both defined to 1). When $a = 0$, the series is also called a Maclaurin series [13].

3. Applications

3.1. Math

In one study, Researchers in one study thought of using an extension to the Taylor series to reliably obtain finite difference approximations for functions with discontinuities. However, this method can only be applied in situations where the discontinuities are limited to a finite number of isolated positions [14]. They considered a function f(x) that is continuous except at one discrete position $x = x_d$ where it is discontinuous. The behaviour of the function at the point of discontinuity can be characterized by a discontinuity matrix [M], which relates the values of the function and its derivatives on both sides of the discontinuity as depicted in next formula:

$$\varphi_{\mathrm{R}} = \begin{bmatrix} m_{11} & \cdots & m_{1q} \\ \vdots & \ddots & \vdots \\ m_{p1} & \cdots & m_{pq} \end{bmatrix} \varphi_{\mathrm{L}} \tag{3}$$

Using the function [15]:

$$\varphi(\mathrm{x} = x_{i+1}) = e^{qD}\ [\mathrm{M}]\ e^{pD}\varphi(x = x_i) \tag{4}$$

The differential operator d/dx, represented by D, is used to define the extended series in Eq. (2) for non-continuous function illustrations. ETS formulates a function considering its value and derivatives at x = x i+1, resembling the Taylor series but without class limitations. Eq. (3), a bivariate function, uses p for the distance to the discontinuity and q for the distance from it to the evaluation spot. Eq. (4) connects the variables by mapping the real axis x onto a path in the p, q plane, as shown by the dashed line in Figure 2 and determined in Figure 1.

DOI: 10.54254/2753-8818/10/20230346

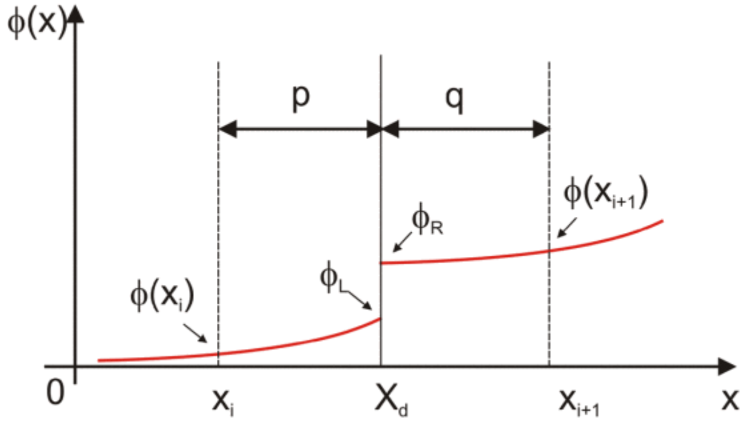

Figure 1. Schematic diagram with definitions of p, q, x_i and Xd.

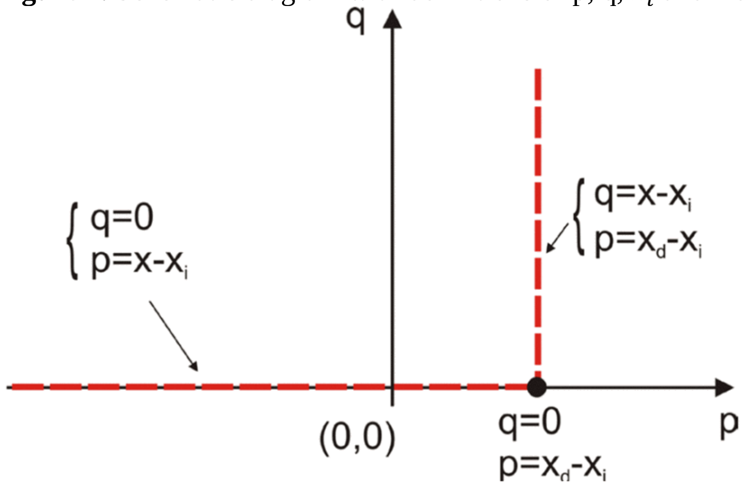

Figure 2. Schematic of mapping onto the p-q plane.

3.2. Machine learning

To enforce the conservative properties of Hamiltonian systems, Greydanus et al. rewrote the loss function using Hamilton's equations [16]. HNNs have been used in various works, including Symplectic Recurrent Neural Networks created by Chen et al. [17] and the Hamiltonian Generative Network created by Toth et al. [18] Zhong developed Symplectic ODE-Net (SymODEN) to better understand the dynamics of systems adhering to Hamiltonian dynamics with control [19]. However, methods that reformulate the loss function face two main issues: difficulty obtaining the temporal derivatives of real-world systems to compute the loss function and a failure to entirely retain the symplectic structure.

To address these issues, a team of researchers developed a distinctive model that incorporates an integrator solver to avoid the need for temporal derivatives and uses a neural network design that predicts a Hamiltonian system's continuous-time development while preserving the system's symplectic structure. The researchers also use neural networks to implement the Taylor series expansion, with each term being symmetric. In addition, they compare the training loss with the current model to evaluate the effectiveness of using the Taylor series as the underlying structure of Taylor-net to ensure nonlinearity.

Experimental results demonstrate that the Taylor series outperforms ReLU, the most popular activation function, in the pendulum, Lotka-Volterra, and Kepler problems. The Taylor series provides a more accurate approximation of the system dynamics and improves forecast accuracy due to its excellent capacity for representation.

DOI: 10.54254/2753-8818/10/20230346

3.3. Engineering

A group of experimenters used an adaptive design based on Taylor expansions to process parameters in a groundwater model to improve accuracy. The adaptive design strategy based on Taylor expansion is a global agent modeling approach. Its core idea is to adaptively adjust the global agent model by gradually adding new sample points to better approximate the true model. Specifically, the method estimates the gradient and curvature information of the real model by performing a first-order Taylor expansion around the current sample point to determine the next optimal sample point. The process is repeated until the error between the global proxy model and the true model reaches a pre-determined threshold. This method can quickly approximate the true model and has high accuracy and reliability.

In contrast to other models, they chose the well-known MmDD strategy [20] in addition to the FLOLAVoronoi strategy and TEAD [21]. The MmDD approach is a space-filling method that chooses new samples by maximizing the distance between them and their nearest neighbors. The FLOLA-Voronoi method is an adaptive design. It balances exploration and exploitation using Voronoi square decomposition and Fuzzy Change of Local Linear Approximation (FLOLA) for non-linear identification. According to several studies, both the MmDD and FLOLAVoronoi methods perform well when creating global surrogate models.

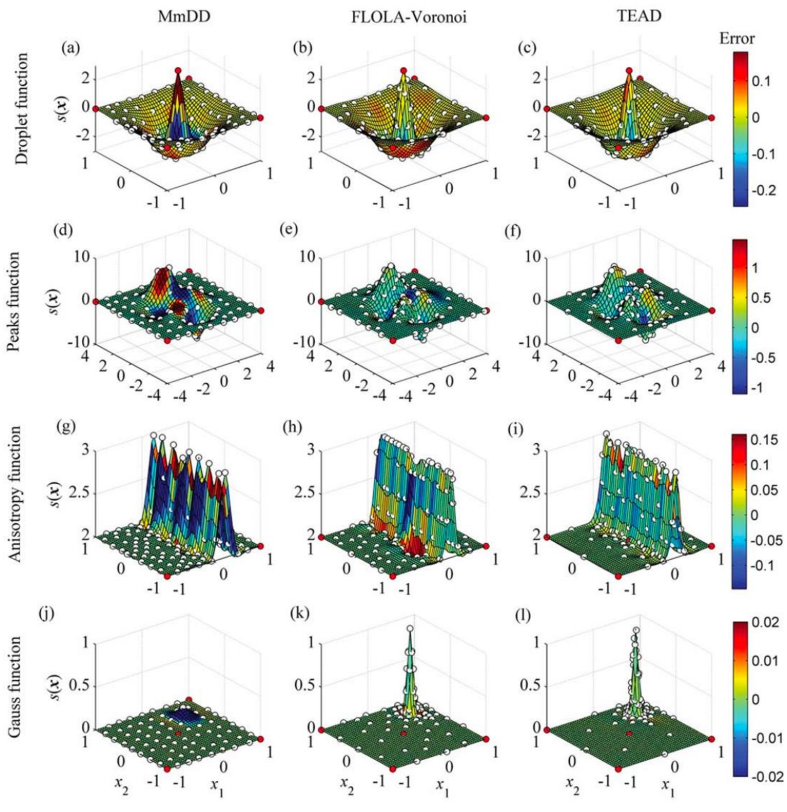

Figure 3. The approximation errors of surrogates for the Droplet (row 1), Peaks (row 2), Anisotropy (row 3), and Gauss functions (row 4) are displayed based on the MmDD (column 1), FLOLA-Voronoi (column 2), and TEAD (column 3) methods. The red dots represent the original samples, while the white dots indicate the new samples utilized in the design.

While developing surrogate functions for the seven test functions, the MmDD, FLOLA-Voronoi, and TEAD algorithms were evaluated for their effectiveness. Based on the three techniques, Figure 3 displays the surrogate functions and the related approximation errors (the discrepancy between the actual function values and the surrogate forecasts). The Droplet, Peaks, Anisotropy, and Gauss functions were all calculated using 81, 85, 125, and 77 samples, respectively. According to the first column in Figure 3, the MmDD technique essentially distributes samples evenly over the parameter domain, which wastes computing resources by oversampling the flat parts of the function and produces approximation errors that are quite substantial. The surrogate performance is somewhat improved by the FLOLA-Voronoi approach, as seen in the second column of Figure 3. In comparison to MmDD, FLOLA-Voronoi gathered more samples in areas with significant function variation and fewer samples in areas with flat topography, which led to a decrease in the approximation error. The drastically fluctuating region along the x1 5 0 line in Figure 3 is an example of a highly non-linear region where errors are still relatively large. Conversely, TEAD significantly enhances the performance of the surrogate through the careful selection and use of sample populations. TEAD employs a significant number of samples in the non-linear zone and a relatively small number in the broad flat part of the function, as illustrated in Figures 3i and 3l, which significantly reduces the total error to a low level.

When creating proxies for the seven functions using MmDD, FLOLA-Voronoi, and TEAD, Figure 3 illustrates the decay of NRMSEs with increasing sample size. The NRMSEs were assessed over a set of validation points. The graphs demonstrate that TEAD is more efficient at reaching the same NRMSE values by utilizing a smaller sample size and is more accurate when compared to MmDD and FLOLA-Voronoi. TEAD also has fewer NRMSE values for the same sample size. According to the statistical findings, TEAD can reduce computational expenses by more than 80% when compared to MmDD and by 40% when compared to FLOLA-Voronoi. Given that running a groundwater model might take many hours or even days, the number of function evaluations saved by TEAD points to a significant increase in the computing effort required for groundwater modeling.

4. Limitations & prospects

The Taylor expansion has been a valuable tool for solving problems in mathematical, engineering, and machine learning applications. However, it is essential to acknowledge its limitations to ensure appropriate use in these fields. The following are some of the limitations of Taylor expansion in various applications, along with references to essays that provide further insight. Firstly, the Taylor expansion can only be applied in the neighborhood of the unfolded point and may have convergence problems at some points, which can result in inaccurate solutions [22]. In some cases, the series may diverge or only converge to the function under specific conditions [23]. Secondly, the Taylor expansion only considers the local behavior of the function at the unfolded point and not the global behavior. It can limit the accuracy of the approximation when applied to distant points from the expansion point [24]. Moreover, the coefficients of the Taylor expansion require the calculation of derivatives of all orders of the function, which may be difficult or even impossible to compute for some higher-order functions, limiting its applicability [25]. Finally, in practical applications, such as engineering and machine learning, the Taylor expansion is usually divide in finite terms. This can result in the accumulation of errors, especially at distant points from the expansion point, leading to imprecise predictions or solutions [26]. To sum up, the Taylor expansion has been a valuable tool in various applications, including mathematics, engineering, and machine learning. However, its limitations, such as convergence issues, local approximation, difficulties with higher-order derivatives, error accumulation, and unsuitability for specific problems, should be considered when applying this method to real-world problems.

5. Conclusion

In conclusion, the Taylor expansion has proven to be a powerful and versatile mathematical tool with a wide range of applications in fields such as machine learning, mathematics, and engineering. Its ability to approximate complex functions has led to advancements in algorithm development, differential equation solutions, and numerical analysis. In particular, its use in symplectic Taylor neural networks, evolutionary algorithms, and Taylor expansion-based adaptive design has demonstrated its potential for addressing challenging problems in these domains. Despite its remarkable capabilities, the Taylor expansion has its limitations. These include convergence issues, the local nature of the approximation, difficulties with calculating higher-order derivatives, and error accumulation in practical applications. When using the Taylor expansion, users must necessarily learn these restrictions beforehand so that they can properly implement the method and maintain precision and efficacy. Overall, the Taylor expansion has made significant contributions to various fields, offering solutions and insights that would have been difficult to achieve otherwise. By acknowledging and addressing its limitations, researchers can continue to harness the power of this mathematical tool to drive further advancements and discoveries.

References

[1] Tong Y et al. 2021 Symplectic Neural Networks in Taylor Series Form for Hamiltonian Systems. Journal of Computational Physics, 437, 110325.
[2] Butcher J C 1987 The numerical analysis of ordinary differential equations: Runge-Kutta and general linear methods. Wiley-Interscience.
[3] Gutierrez N, Lopez A S 2018 Solving ordinary differential equations using genetic algorithms and the Taylor series matrix method. Journal of Physics Communications, 2(11), 115010.
[4] Smith G D 1985 Numerical solution of partial differential equations: finite difference methods. Oxford University Press.
[5] Linde N et al. 2017 On uncertainty quantification in hydrogeology and hydrogeophysics. Advances in Water Resources, 110, 166-181.
[6] Mo S et al. 2017 A Taylor Expansion-Based Adaptive Design Strategy for Global Surrogate Modeling with Applications in Groundwater Modeling. Water Resources Research, 53(12), 10802-10823.
[7] Skovgaard O 1978 Table errata: Handbook of mathematical functions with formulas, graphs, and mathematical tables (Nat. Bur. Standards, Washington, DC, 1964), edited by Milton Abramowitz and Irene A. Stegun. Mathematics of Computation, 32(141), 317.
[8] Kline M 1990 Mathematical Thought from Ancient to Modern Times: Volume 2 (Vol. 2). Oxford University Press.
[9] Sayed B H, George L E 2020 Flat Model for Representing Contiguous UTM Coordinates over Iraq Territory. Iraqi Journal of Science, 908-919.
[10] Dani S G 2012 Ancient Indian mathematics—A conspectus. Resonance, 17, 236-246.
[11] Malet A 1993 James Gregorie on Tangents and the "Taylor" Rule for Series Expansions. Archive for History of Exact Sciences, 97-137.
[12] Taylor B 1969 Methodus Incrementorum Directa et Inversa [Direct and Reverse Methods of Incrementation]. Pearsonianis prostant apud Gul. Innys, 21-23.
[13] Vetter W J 1973 Matrix calculus operations and Taylor expansions. SIAM Review, 15(2), 352-369.
[14] Sujecki S 2013 Application of Extended Taylor Series based Finite Difference Method in photonics. 2013 15th International Conference on Transparent Optical Networks (ICTON). IEEE.
[15] Sujecki S 2013 Extended Taylor Series Interpolation of Physically Meaningful Functions. Optical Quantum Electronics, 45, 53-66
[16] Greydanus S, Dzamba M, Yosinski J 2019 Hamiltonian neural networks. In Advances in Neural Information Processing Systems 32.
[17] Chen Z, Wang S, Yang J, Wu C 2019 Symplectic recurrent neural networks. arXiv preprint arXiv:1909.13334.
[18] Toth P, Kara R, Hajas P, Szabo T 2019 Hamiltonian generative networks. arXiv preprint arXiv:1909.13789.
[19] Zhong Y D, Dey B, Chakraborty A 2019 Symplectic ODE-Net: Learning Hamiltonian Dynamics with Control. arXiv preprint arXiv:1909.12077.
[20] Johnson M E, Moore L M, Ylvisaker D 1990 Minimax and maximin distance designs. Journal of Statistical Planning and Inference, 26(2), 131-148.
[21] Van der Herten J, Couckuyt I, Deschrijver D, Dhaene T 2015 A fuzzy hybrid sequential design strategy for global surrogate modeling of high-dimensional computer experiments. SIAM Journal on Scientific Computing, 37(2), A1020-A1039.
[22] Kung H T 1982 Why systolic architectures? IEEE Computer, 15(1), 37-46.

DOI: 10.54254/2753-8818/10/20230346

[23] Watkins D S 2007. The matrix eigenvalue problem: GR and Krylov subspace methods. SIAM.
[24] Hairer E, Nørsett S P, Wanner G 1993 Solving ordinary differential equations I: Nonstiff problems. Springer.
[25] Bryson A E, Ho Y C 1975 Applied optimal control: Optimization, estimation, and control. Hemisphere Publishing Corporation.
[26] Rivlin T J 1981 An introduction to the approximation of functions. Dover Publications.