Numer. Math. Theor. Meth. Appl. Vol. **17**, No. 2, pp. 310-330
doi: 10.4208/nmtma.OA-2023-0133 May 2024

Efficiently Training Physics-Informed Neural Networks via Anomaly-Aware Optimization

Jiacheng Li, Min Yang* and Chuanjun Chen

School of Mathematics and Information Sciences, Yantai University, Yantai, China

Received 24 October 2023; Accepted (in revised version) 24 January 2024

Abstract. Physics-Informed Neural Networks (PINNs) encounter challenges in dealing with imbalanced training losses, especially when there are sample points with extremely high losses. This can make the optimization process unstable, making it challenging to find the correct descent direction during training. In this paper, we propose a progressive learning approach based on anomaly points awareness to improve the optimization process of PINNs. Our approach comprises two primary steps: the awareness of anomaly data points and the update of training set. Anomaly points are identified by utilizing an upper bound calculated from the mean and standard deviation of the feedforward losses of all training data. In the absence of anomalies, the parameters of the PINN are optimized using the default training data; however, once anomalies are detected, a progressive exclusion method aligned with the network learning pattern is introduced to exclude potentially unfavorable data points from the training set. In addition, intermittent detection is employed, rather than performing anomaly detection in each iteration, to balance performance and efficiency. Extensive experimental results demonstrate that the proposed method leads to substantial improvement in approximation accuracy when solving typical benchmark partial differential equations. The code is accessible at `https://github.com/JcLimath/Anomaly-Aware-PINN`.

AMS subject classifications: 35Q68, 68T07, 68W25

Key words: Imbalanced losses, anomaly detection, progressive learning, physics-informed, neural networks.

1. Introduction

Scientific computing is a critical tool for enhancing human understanding and driving positive change worldwide. One significant area within this field is solving partial differential equations (PDEs), which find extensive applications across multiple disciplines. While traditional numerical methods like finite element methods [4], finite

*Corresponding author. *Email addresses:* `jiachenglimath@hotmail.com` (J. Li), `yang@ytu.edu.cn` (M. Yang), `cjchen@ytu.edu.cn` (C. Chen)

volume methods [30], and finite difference methods [11] have continuously evolved over decades to address PDEs, they still encounter challenges, especially when dealing with high-dimensional and multi-scale PDEs [53]. In recent years, deep learning has made groundbreaking advancements in many fields, including computer vision [42] and natural language processing [51]. Combining deep learning methods for solving PDEs has also gained increasing attention due to their powerful nonlinear approximation capabilities [7, 13–15, 19, 20, 34, 45].

One particularly noteworthy development is the emergence of the Physical-Informed Neural Network (PINN) paradigm [3, 6, 8, 27, 31, 39, 46, 47, 52]. The core concept underpinning PINNs revolves around the integration of partial differential equations and their associated conditions as soft constraints within the loss function of a neural network [21, 22, 25, 38]. This innovative approach facilitates an iterative optimization process, allowing for the continuous refinement of the approximation as time progresses. PINNs offer several advantages over traditional numerical methods, including its grid-free nature, capacity to handle high-dimensional equations, and seamless integration with data. PINNs have demonstrated promising results and has found applications in diverse fields, including fluid dynamics [32, 36], bioengineering [24, 33], and the electrical power industry [29]. However, the standard version of PINN faces challenges during training. A major challenge is its difficulty in effectively dealing with unbalanced training losses. In some cases, the losses for specific points within the computational domain can be orders of magnitude higher than others, leading to their dominance during the training process. In such scenarios, the vanilla PINN struggles to find the optimal direction for optimization [1, 16, 43, 49].

Previous research has tackled the challenge of addressing the loss imbalance by adjusting the weights of various loss terms. Some works [5, 44] relied on manually designed hyperparameters to fine-tune the weights of loss terms. However, due to their heavy dependence on prior knowledge, these methods may not always yield optimal solutions. Consequently, there has been a growing interest in the adaptive weighting approach. For instance, gradient statistics were employed to harmonize the interactions among different terms in the loss function [40], while a novel adaptive weighting approach based on the Neural Tangent Kernel (NTK) was introduced by Wang *et al.* [41]. In [2], the loss function was modified into a dimensionless form, with the parameters determined using least squares weighted residuals method. However, these efforts primarily centered on adjusting the weights of different loss components, possibly overlooking the imbalances originating from individual data points. Recent endeavors have sought improvements in designing the loss weights for individual points. Drawing inspiration from soft multiplicative attention masks, a method was introduced by McClenny *et al.* [28], where trainable weights are assigned to each individual training point. In [16], quantiles of the residuals were used to adjust the weights, leading to a redistribution of weights of points with extreme losses towards the median. Notably, a distinct approach was proposed by Wang *et al.* [43], where instead of assigning greater weights to sample points with larger losses, priority is given to the learning of easy low-loss sample points. This approach significantly enhances the performance of