Numer. Math. Theor. Meth. Appl.
doi: 10.4208/nmtma.OA-2023-0136

Vol. **17**, No. 2, pp. 534-554
May 2024

Error Analysis of the Mixed Residual Method for Elliptic Equations

Kai Gu, Peng Fang, Zhiwei Sun and Rui Du*

School of Mathematical Science, Soochow University, Suzhou 215000, China

Received 27 October 2023; Accepted (in revised version) 20 February 2024

Abstract. We present a rigorous analysis of the convergence rate of the deep mixed residual method (MIM) when applied to a linear elliptic equation with different types of boundary conditions. The MIM has been proposed to solve high-order partial differential equations in high dimensions. Our analysis shows that MIM outperforms deep Ritz method and deep Galerkin method for weak solution in the Dirichlet case due to its ability to enforce the boundary condition. However, for the Neumann and Robin cases, MIM demonstrates similar performance to the other methods. Our results provide valuable insights into the strengths of MIM and its comparative performance in solving linear elliptic equations with different boundary conditions.

AMS subject classifications: 65N12, 65N15
Key words: Deep mixed residual method, deep neural network, error analysis, elliptic equations, Rademacher complexity.

1. Introduction

Partial differential equations (PDEs) are widely used in various fields such as science, engineering, and finance [1, 11] to model complex physical phenomena. Traditional numerical methods such as the finite difference method and finite element method (FEM) have been successful in solving low-dimensional PDEs. However, when it comes to high-dimensional problems, the curse of dimensionality becomes a bottleneck, making classical methods inapplicable. In recent years, deep learning-based methods [7] have emerged as a promising approach for tackling high-dimensional PDEs. Notable examples of these methods include the deep Ritz method (DRM) [4, 8, 10], the deep Galerkin method (DGM) [20], the physics informed neural network method [19], the weak adversarial network method [22], the deep least-squares methods [3], least-squares ReLU neural network method [2] and the local deep learning method [21]. These techniques have shown significant potential for solving high-dimensional PDEs, offering an attractive alternative to traditional methods.

*Corresponding author. *Email addresses:* durui@suda.edu.com (R. Du), 20234207018@stu.suda.edu.com (K. Gu), pengfang010809@163.com (P. Fang), 20194007008@stu.suda.edu.com (Z. Sun)

Despite the success of deep learning methods such as DGM and DRM, they still suffer from convergence and efficiency issues, especially when dealing with boundary conditions. Recently, the (deep) mixed residual method (MIM) [18] has been proposed to address these limitations. MIM is a novel approach that reformulates high-order PDEs into a first-order system. This idea is similar to the local discontinuous Galerkin method and the mixed FEM. MIM also uses the residual of the first-order system in the least-squares sense as the loss function, which is closely related to the least-squares FEM. Importantly, MIM can directly enforce exact boundary conditions, including mixed boundary conditions, which are difficult to deal with in previous deep learning methods [17]. In comparison to DGM and DRM, MIM provides better numerical approximation in most tested cases. However, rigorous analysis of the convergence rate of MIM is currently limited, which hampers our understanding of its theoretical properties and practical performance.

In the realm of error analysis, extensive investigations have been carried out with respect to DRM [5, 6, 16], PINN [12, 16], and the Deep Galerkin Method for weak solutions (DGMW) [14]. Regarding the MIM, the initial analysis was conducted in [15].

This study focused on the two-layer neural network scenario and revealed that convergence could be achieved by increasing the number of training samples, independent of the network size. Additionally, it demonstrated that MIM provides more accurate approximations of high-order derivatives compared to DRM, as confirmed through numerical experiments. In this article, we extend the analysis to the case of multi-layer neural networks and rigorously investigate the convergence rate of MIM when applied to linear elliptic equations with different kinds of boundary conditions. To evaluate its performance, we compare the results to existing findings for the Deep Ritz Method (DRM) [13] and the Deep Galerkin Method for weak solutions (DGMW) [14]. We employ the same methodology that has been successfully applied to DRM with sigmoid and tanh functions [13] and RELU [6], as well as to DGMW [14].

Specifically, we decompose the total error in MIM into three distinct components: the approximation error, the statistical error, and the optimization error. The results of the approximation error can be obtained by a recent work [9] which provides the dependency of the error upon the network parameters. The derivation of the statistical error is the main contribution of this work, which uses Rademacher complexity to estimate the number of needed sampling points. The optimization error contains information on the landscape structure of the loss function and is beyond the scope of this work. Although the main idea is similar to previous work [13, 14], there are some differences worth mentioning:

(i) MIM uses the residual of the first-order system in the least-squares sense as the loss function, which needs an additional insensible variation condition on the coefficients of the elliptic problem to make the loss function non-degenerate.

(ii) The neural network that MIM uses has a multi-dimensional output function, which makes the approximation analysis slightly different.