Anonymized Deep Ritz Method for Elliptic Differential-Difference Equations*

Submission ID 1002

Abstract

In this paper, the deep Ritz method is used for the numerical solution of the first boundary-value problem for differential-difference equations. The variational formulation of the problem allows one to use the Deep Ritz method. The numerical results were compared with exact solutions where it is possible, that justifies the use of this method. Authors are anonymized for double-blind review.

1 Introduction

Elliptic differential-difference equations were studied by Alexander Skubachevskiĭ and his students (see survey [7]). Under some assumptions, these equations are equivalent to elliptic problems with nonlocal boundary conditions, which arise in the physics of plasma (see paper by A. Bitsadze and A. Samarskiĭ [1]). Despite the voluminous literature on these problems, only a few papers discuss the numerical solutions of such equations and only in the 1D case (see paper by A. Kamenskiĭ [4] and literature therein). The reason for this lack of numerical methods for nonlocal problems lies in the properties of their solutions. Typically, the smoothness of the solutions is violated inside the domain (the first derivative does not exist), there are examples when it doesn't exist almost everywhere in the domain (see Example 8.2 in [7]). In this case, the grid or finite element methods are useless, and the numerical solution of such equations was an open problem for many years.

In this paper, we investigate the possibility of applying the deep Ritz method for the numerical solution of elliptic differential-difference equations. The deep Ritz method was introduced by W. E and B. Yu in paper [3]. It uses a neural network to approximate the unknown function and automated differentiation for minimization of the energy functional.

We apply it to problems in 1D and 2D with known analytical solutions and also to a differential-difference equation in a disk, where no analytic solution is known. The numerical experiments demonstrate the efficiency of this approach.[1]

*The full version of the paper can be accessed at `redacted-url`

[1]The code for all examples in this paper can be found on `redacted-url`.

2 Statement of the Problem

We consider the following boundary-value problem for a differential-difference equation in a bounded domain $Q \subset \mathbb{R}^n$. The smoothness of the boundary ∂Q will be motivated by the ellipticity conditions. For now, we can assume it is Lipschitz.

$$-\sum_{i,j=1}^{n} \left(R_{ijQ} u_{x_j}\right)_{x_i} + \sum_{i=1}^{n} R_{iQ} u_{x_i} + R_{0Q} u = f(x) \tag{2.1}$$

for $x \in Q$ with the boundary condition

$$u(x) = 0, \qquad x \in \partial Q. \tag{2.2}$$

Where $f \in L_2(Q)$, and operators $R_{ij,Q}$, R_{iQ}, and R_{0Q} are defined below. Introduce bounded difference operators R_{ij}, $R_i \colon L_2(\mathbb{R}^n) \to L_2(\mathbb{R}^n)$ by the formulas

$$\begin{aligned} (R_{ij}u)(x) &= \sum_{h \in M} a_{ijh} u(x+h), \quad i,\, j = 1..n, \\ (R_i u)(x) &= \sum_{h \in M} a_{ih} u(x+h), \quad i = 0..n. \end{aligned} \tag{2.3}$$

Here, a_{ijh}, a_{ih} are real numbers, the set M consists of a finite number of vectors $h \in \mathbb{R}^n$ with integer coordinates.

We introduce an operator of extension of a function by zero outside Q: $I_Q \colon L_2(Q) \to L_2(\mathbb{R}^n)$; an operator of the projection of a function onto Q: $P_Q \colon L_2(\mathbb{R}^n) \to L_2(Q)$; then the operators R_{ijQ}, $R_{iQ} \colon L_2(Q) \to L_2(Q)$ defined by the formulas $R_{ijQ} = P_Q R_{ij} I_Q$, $R_{iQ} = P_Q R_i I_Q$.

Let $\mathring{H}^1(Q)$ be the Sobolev space of real-valued functions from $L_2(Q)$ having all generalized derivatives of the first order from $L_2(Q)$ and vanishing on ∂Q.

Introduce a bilinear form $a[v, w]$ in $L_2(Q)$ with domain $\mathring{H}^1(Q)$ by the formula

$$\begin{aligned} a[v,\, w] &= \sum_{i,j=1}^{n} \left(R_{ijQ} v_{x_j},\, w_{x_i}\right)_{L_2(Q)} \\ &+ \sum_{i=1}^{n} \left(R_{iQ} v_{x_i},\, w\right)_{L_2(Q)} + \left(R_{0Q} v,\, w\right)_{L_2(Q)}. \end{aligned} \tag{2.4}$$

The difference operators R_{ijQ}, R_{iQ}, $R_{0Q} \colon L_2(Q) \to L_2(Q)$ are bounded. Therefore, it follows that there exists a constant $c_0 > 0$ such that for $v, w \in \mathring{H}^1(Q)$

$$|a[v, w]| \le c_0 \|v\|_{\mathring{H}^1(Q)} \|w\|_{\mathring{H}^1(Q)}. \tag{2.5}$$

Definition 2.1. *The form $a[v, w]$ is said to be* coercive *and equation (2.1)* elliptic (strongly), *if there exist numbers $c_1 > 0$ and $c_2 \ge 0$ such that for $v \in \mathring{H}^1(Q)$*

$$a[v, v] \ge c_1 \|v\|^2_{\mathring{H}^1(Q)} - c_2 \|v\|^2_{L_2(Q)}. \tag{2.6}$$

For the necessary and sufficient conditions of coercivity in algebraic form see Sec. 3 in [7] (it is also true in case ∂Q is Lipschitz). Further, we shall assume that the form $a[v, w]$ is coercive.

Definition 2.2. *A function $u \in \mathring{H}^1(Q)$ is said to be a weak solution to problem (2.1)-(2.2), if*

$$a[u, v] = (f, v)_{L_2(Q)}, \quad \forall v \in \mathring{H}^1(Q). \tag{2.7}$$

From (2.5) and (2.6), we get the following result.

Lemma 2.1. *If $c_2 = 0$, then $a[v, v]$ defines an equivalent norm on $\mathring{H}^1(Q)$.*

Our goal is to find a numerical solution for problem (2.1)-(2.2). In the next section, we will use the Deep Ritz algorithm under the assumption of strict ellipticity with $c_2 = 0$.

3 Solution via Variational Formulation

The following theorem gives a variational formulation of the problem (2.1)-(2.2).

Theorem 3.1. *Assume that the form $a[v, w]$ is coercive with $c_2 = 0$. Then there exists a unique function $u \in \mathring{H}^1(Q)$ that is the minimizer of the following functional:*

$$E(u) = a[u, u] - 2(u, f)_{L_2(Q)}. \tag{3.8}$$

This function is the weak solution of the problem (2.1)-(2.2).

Proof. The proof follows the same steps as Theorem 14 in [5, Ch 4 §1] using Lemma 2.1. □

From Theorem 3.1, it follows that a sequence $u_n \in \mathring{H}^1(Q)$ that minimizes functional (3.8) should converge in $\mathring{H}^1(Q)$ to the weak solution of problem (2.1)-(2.2). We will use the deep Ritz algorithm to find an approximation of the minimizer of the energy functional $E(u)$.

The first idea of the deep Ritz method is to approximate the solution u by a neural network with parameters θ. To satisfy the boundary conditions, we search for the solution in the following form:

$$u_\theta(x) = \alpha(x) v_\theta(x). \tag{3.9}$$

Here, $\alpha(x)$ is a smooth function, such that, $\alpha(x) \neq 0$ for $x \in Q$ and $\alpha(x) = 0$ for $x \in \partial Q$; v_θ is a neural network with input shape n and output shape 1.

Thus, the functional $E(u)$ becomes a function of θ :

$$J(\theta) = E(u_\theta), \tag{3.10}$$

and any optimization method can be used to minimize it. In particular, one can use the built-in optimizors in neural network frameworks. The only problem remains to calculate integrals in (2.4).

The second idea of the Deep Ritz method is to use Monte Carlo estimations for the integrals (see, e.g., [6, Ch. 3]). Then a one-sample loss function has the following form:

$$\begin{aligned} L_m\big(x^{(m)}, \theta\big) &= \sum_{i,j=1}^{n} \big(R_{ijQ} u_{\theta_k}\big)_{x_j} \big(u_{\theta_k}\big)_{x_i} \\ &+ \sum_{i=1}^{n} \big(R_{iQ} u_{\theta_k}\big)_{x_i} u_{\theta_k}\big(x^{(m)}\big) \\ &+ \big(R_{0Q} u_{\theta_k}\big) u_{\theta_k}\big(x^{(m)}\big) - 2 f\big(x^{(m)}\big) u_{\theta_k}\big(x^{(m)}\big). \end{aligned} \tag{3.11}$$

We generate a set of points (grid) $x^{(m)}$, $m = 1..M$ and approximate the energy functional as follows:

$$J(\theta) \approx \frac{1}{M} \sum_{m=1}^{M} L_m\big(x^{(m)}, \theta\big). \tag{3.12}$$

This approximation corresponds to a mini-batch in terms of stochastic gradient descent.

The final algorithm for the numerical solution of problem is as follows.

Algorithm 3.1. Initialize parameters of NN θ_0
for $k = 1 \ldots$ n_epoch **do**
  $x^{(m)} \sim Uniform(\overline{Q})$, $m = 0, \ldots, M$
  $J \leftarrow$ (3.12)
  Update $\theta_{k+1} \leftarrow \theta_k - \eta \nabla J$
end for
return $\theta_{\text{n\_epoch}}$

The numerical experiments demonstrate the efficiency of this approach for solving the differential-difference equations (2.1) with Dirichlet boundary conditions (2.2).

4 Experimental Results

4.1 Example 1: 1D case

Consider the following problem:

$$\begin{cases} -(R_Q u)'' = 1, & 0 \le x \le 2, \\ u(0) = u(2) = 0, \end{cases} \tag{4.13}$$

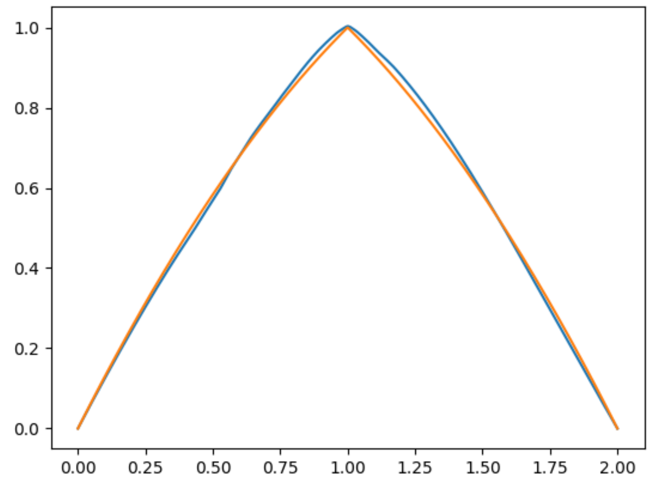

Figure 1: Example 1: higher curve is the numerical solution, the other curve is the exact solution.

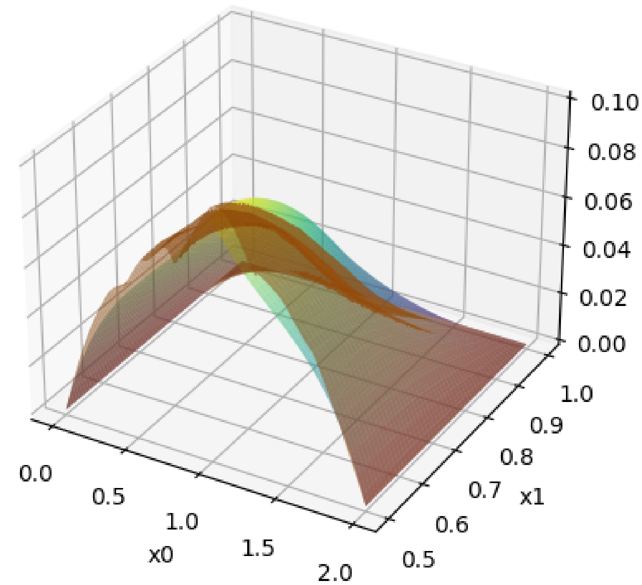

Figure 2: Example 2: uneven plot corresponds to the numerical solution.

where operator R_Q is defined by $R_Q = P_Q R I_Q$ with $Q = (0,\ 2)$ and

$$(Rv)(x) = v(x) + bv(x+1) + bv(x-1), \quad |b| < 1. \tag{4.14}$$

This nonlocal problem is equivalent to the following equation with nonlocal boundary conditions:

$$\begin{cases} -w'' = 1, & 0 \le x \le 2, \\ w(0) = w(2) = bw(1). \end{cases} \tag{4.15}$$

It can be solved and the solution is given by

$$w(x) = -\frac{x^2}{2} + x + \frac{b}{2(1-b)}. \tag{4.16}$$

By Theorem 7.2 in [7], $w = R_Q u$, and $u = R_Q^{-1} w$, that is plotted in fig. See, e.g., [7, Sec. 2] for how to invert the difference operator R_Q.

4.2 Example 2: 2D problem

Consider the following problem in $Q = (0,2) \times (0,1)$:

$$\begin{cases} -\Delta(R_Q u) = \sin(\pi x_2), & x \in Q, \\ u|_{x_1=0} = u|_{x_1=2} = u|_{x_2=0} = u|_{x_2=1} = 0. \end{cases} \tag{4.17}$$

Where $R_Q = P_Q R I_Q$ with

$$\begin{aligned}(Rv)(x_1, x_2) = v(x_1, x_2) + b\,(v(x_1 - 1, x_2) \\ + v(x_1 + 1, x_2)).\end{aligned} \tag{4.18}$$

After substitution $u(x) = u_1(x_1) u_2(x_2)$ with $u_2(x_2) = \sin(\pi x_2)$ this problem is equivalent to the following equation with nonlocal conditions:

$$\begin{cases} -\Delta w = \sin(\pi x_2), & x \in Q, \\ w|_{x_1=0} = w|_{x_1=2} = bw|_{x_1=1}, \\ w|_{x_2=0} = w|_{x_2=1} = 0, \end{cases} \tag{4.19}$$

that can be solved, and after applying R_Q^{-1} the solution of problem (4.17) can be found.

$$w(x) = \frac{1}{\pi^2} + \frac{(b-1)}{\pi^2} \frac{\sinh \pi(2 - x_1) + \sinh x_1}{\sinh 2\pi - 2b \sinh \pi} \sin \pi x_2. \tag{4.20}$$

The exact and approximate solutions are plotted in fig.

4.3 Example 3: 2D problem in a disc

Finally, consider a problem

$$\begin{cases} -\Delta(R_Q u) = 1, & x \in Q = \{|x| < 1\}, \\ u|_{|x|=1} = 0. \end{cases} \tag{4.21}$$

Where $R_Q = P_Q R I_Q$ and operator R is defined by equality (4.18).

Table 1 shows additional details about the above experiments.

Table 1: Example table.

Example	# Layers	# Samples	$\|u - u_\theta\|_{C(Q)}$
1D	14	1600	0.009
2D rectangular	30	2^{13}	0.01
2D circle	30	2^{13}	—

We noticed, that in 2D case the number of samples for Monte Carlo estimation is less important than the depth of the neural network (we use ResNet with ReLu

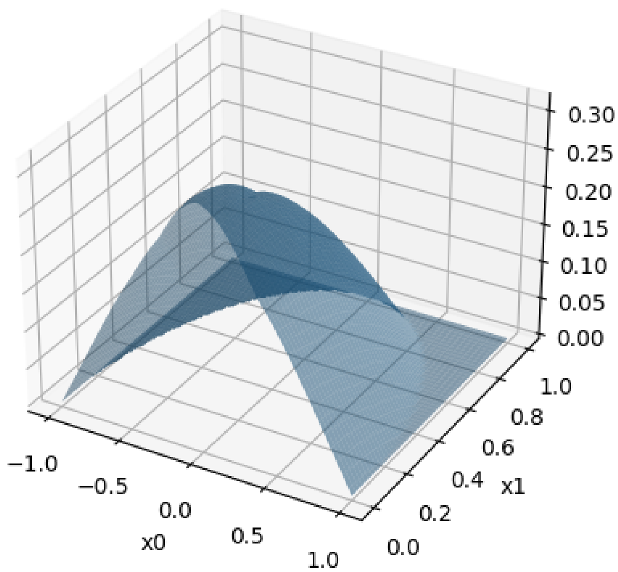

Figure 3: Example 3: numerical solution.

activation). In 1D case, the increase of the number of layers resulted in a worse convergence, but it is possible that after more epochs it would converge better. The last observation we have done is that decrease of the learning rate with the number of iterations results in lower values of the loss function. We used Adam optimizer and residual architecture of the neural network.

5 Conclusion

We have demonstrated the feasibility of using the deep Ritz method for solving elliptic differential-difference equations. Our numerical experiments show that the method is capable of providing accurate solutions in both one-dimensional and two-dimensional cases with known analytical solutions. Additionally, the method performs well in a more complex case where no analytical solution is available. These results suggest that the deep Ritz method is a promising approach for solving elliptic differential-difference equations and potentially other types of nonlocal problems.

In paper [2], a regularization term was added to the loss function and the convergence of the deep Ritz method was discussed, but the error estimation and model architecture are open problems.

References

[1] A. V. Bitsadze and A. A. Samarskiĭ, *Some elementary generalizations of linear elliptic boundary value problems*, Dokl. Akad. Nauk SSSR, 185 (1969), pp. 739–740.

[2] P. Dondl, J. Müller, and M. Zeinhofer, *Uniform convergence guarantees for the deep ritz method for nonlinear problems*, Adv. Cont. Discr. Mod., 49 (2022).

[3] W. E and B. Yu, *The deep ritz method: A deep learning-based numerical algorithm for solving variational problems*, Commun. Math. Stat., 6 (2018), pp. 1–12.

[4] A. G. Kamenskiĭ, *Approximate solution of second-order functional-differential equations*, Differentsial'nye Uravneniya, 19 (1983), pp. 2148–2153. In Russian.

[5] V. P. Mikhailov, *Partial Differential Equations*, MIR Publishers, Moscow, 1978.

[6] C. P. Robert and G. Casella, *Monte Carlo Statistical Methods*, Springer, New-York, 2 ed., 2004.

[7] A. L. Skubachevskiĭ, *Boundary value problems for elliptic functional-differential equations and their applications*, Russian Math. Surveys, 71 (2016), pp. 801–906.