Workshop

eman ta zabal zazu

Universidad del País Vasco — Euskal Herriko Unibertsitatea

Taller

Coding Deep Neural Networks for PDEs
Session II: PINNs and VPINNs

Professors: Carlos Uriarte*, Ángel Javier Omella*, David Pardo*

Organizers: Ignacio Muga†, Paulina Sepúlveda†

*Research Group on Applied Mathematical Modelling, Statistics, and Optimization (MATHMODE) at UPV/EHU

†IMA Numerics Research Group at PUCV

Follow us on: @mathmode.science, @ima.numerics.pucv†*

Biosfera Lodge, Olmué. January 19-24, 2025

PINNs: Not-So-Named Original Approach

For a problem of the form

$$\begin{cases} Au^* = f, & \text{in } \Omega, \\ Bu^* = g, & \text{in } \partial\Omega, \end{cases}$$

the loss function was considered as

$$\mathcal{L}(u) = \int_\Omega \|Au - f\|^2 + \int_{\partial\Omega} \|Bu - g\|^2.$$

Original authors considered: a feed-forward neural network to approximate u, collocation points for quadrature, finite differences for differentiation, and a quasi-Newton method for optimization.

Dissanayake, M. G., & Phan-Thien, N. (1994). Neural-network-based approximations for solving partial differential equations. *Communications in Numerical Methods in Engineering*, 10(3), 195-201.

PINNs: Actual Approach

For a problem of the form

$$\begin{cases} Au^* = f, & \text{in } \Omega, \\ Bu^* = g, & \text{in } \partial\Omega, \end{cases}$$

the loss function was considered as

$$\mathcal{L}(u) = \int_\Omega (Au - f)^2 + \lambda \int_{\partial\Omega} (Bu - g)^2, \qquad \lambda \geq 0 \text{ (user-chosen for appropriate balance)}$$

$$\mathcal{L}(u(\theta)) \approx \frac{\mathrm{Vol}(\Omega)}{K_{int}} \sum_{k=1}^{K_{int}} (Au(\theta)(x_k) - f(x_k))^2 + \lambda \frac{\mathrm{Vol}(\partial\Omega)}{K_b} \sum_{k=1}^{K_b} (Bu(\theta)(x_k) - g(x_k))^2$$

where $u(\theta)$ is a neural network, x_k is a randomly and uniformly sampled integr. point in either Ω or $\partial\Omega$.

Raissi, M., Perdikaris, P., & Karniadakis, G. E. (2019). Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. Journal of Computational Physics, 378, 686-707.

Why Do PINNs Actually Work?

- They rely on the **strong (least-squares) formulation of PDEs** with regularization for BCs.

- ...as long as $A : \mathbb{U} \longmapsto L^2(\Omega)$, $f \in L^2(\Omega)$, $B : \mathbb{U} \longrightarrow L^2(\partial\Omega)$, and $g \in L^2(\partial\Omega)$ are well defined and determine in a compatible fashion a well-posed problem.

- We might restrict to the case where $\mathbb{U}$ **absorbs homogeneous BCs**, i.e.,

$$\mathbb{U} = \{u \in L^2(\Omega) : Au \in L^2(\Omega), Bu = 0\}.$$

For non-homogeneous problems, the seeking space becomes $g + \mathbb{U} = \{g + u : u \in \mathbb{U}\}$.

✓ This avoids the balancing challenge in the loss function. $\boxed{\mathcal{L}(u) = \|Au - f\|^2_{L^2(\Omega)}}$

✗ This is straightforward when using Dirichlet-only BCs, but it is not in general.

Existence and Uniqueness

Let

- $\mathbb{U}$ be a Banach space,
- $A : \mathbb{U} \longrightarrow L^2(\Omega)$ be a PDE operator,
- $f \in L^2(\Omega)$.

When does $Au^* = f$ admit a unique solution? Assume

1. A is **linear**.
2. A is **bounded** (thus, A is continuous), i.e., there exists $0 < \mu < \infty$ such that

$$\|Au\|_{L^2(\Omega)} \leq \mu \|u\|_{\mathbb{U}}, \qquad \forall u \in \mathbb{U}.$$

3. A is **bounded from below** (thus, A is into), i.e., there exists $0 < \alpha < \mu$ such that

$$\alpha \|u\|_{\mathbb{U}} \leq \|Au\|_{L^2(\Omega)}, \qquad \forall u \in \mathbb{U}.$$

4. A', the adjoint of A, is bounded from below (thus, A is onto) OR $f \in A(\mathbb{U})$.

Robustness

- Under the previous conditions,

Robustness

$$\frac{1}{\mu}\|\underbrace{Au-f}_{\text{residual}}\|_{L^2(\Omega)} \leq \|\underbrace{u-u^*}_{\text{error}}\|_{\mathbb{U}} \leq \frac{1}{\alpha}\|\underbrace{Au-f}_{\text{residual}}\|_{L^2(\Omega)}, \qquad u \in \mathbb{U}.$$

- For the **favorable graph-norm choice** for $\mathbb{U}$ — possibly not faithful for a physical problem of interest,

$$\|\cdot\|_{\mathbb{U}} := \|A\ \cdot\|_{L^2(\Omega)},$$

we obtain the superb relation "*residual minimization equals error minimization*": $\mu = \alpha = 1$.

Summary

- PINNs rely on the **strong formulation** of a PDE, $Au = f$.
- The main term of the loss function is $\|Au - f\|^2_{L^2(\Omega)}$.
- Loss minimization is **robust** in scenarios with absorbed BCs. Indeed, it can be interpreted as equal to error minimization in the graph norm. ✓
- In non-absorbed BC scenarios, regularizing addends might lead to **unbalancedness** and unclear robustness. Non-Dirichlet BCs are challenging to absorb. ✗

Residual Operator

A PDE in variational form can be read as vanishing a residual operator $\mathcal{R} : \mathbb{U} \longrightarrow \mathbb{V}$,

$$\mathcal{R}(u^*) = 0$$

Poisson's equation ($A = -\Delta$): $-\Delta u = f$

- **Strong form:** $\mathcal{R} : H^2(\Omega) \cap H_0^1(\Omega) \to L^2(\Omega)$,

$$\boxed{\mathcal{R}(u) = Au - f = -\Delta u - f = 0}$$

- **Weak form:** $\mathcal{R} : H_0^1(\Omega) \to H_0^1(\Omega)$,

$$\boxed{(\mathcal{R}(u), v)_{H_0^1(\Omega)} = \int_\Omega \nabla u \cdot \nabla v - f\ v = 0 \quad \forall v \in H_0^1(\Omega)}$$

Residual Minimization

Then, solving a PDE in variational form can be read as

$$u^* = \arg\min_{u \in \mathbb{U}} \|\mathcal{R}(u)\|_{\mathbb{V}}^2$$

Poisson's equation ($A = -\Delta$): $-\Delta u = f$

- **Strong form:** $\mathcal{R} : H^2(\Omega) \cap H_0^1(\Omega) \to L^2(\Omega)$,

$$\boxed{u^* = \arg \min_{u \in H^2(\Omega) \cap H_0^1(\Omega)} \int_\Omega (Au - f)^2}$$

- **Weak form:** $\mathcal{R} : H_0^1(\Omega) \to H_0^1(\Omega)$,

$$\boxed{u^* = \arg \min_{u \in H_0^1(\Omega)} \int_\Omega \nabla u \cdot \nabla \mathcal{R}(u) - f\, \mathcal{R}(u)}$$

Under the hypotheses of the Babuska-Lax-Milgram Theorem,

Robustness

$$\frac{1}{\mu}\|\mathcal{R}(u)\|_{\mathbb{V}} \leq \|u - u^*\|_{\mathbb{U}} \leq \frac{1}{\alpha}\|\mathcal{R}(u)\|_{\mathbb{V}}$$

- μ continuity (sup-sup) constant,
- α weak coercivity (inf-sup) constant.

Babuška, I. (1971). Error-bounds for finite element method. Numerische Mathematik, 16(4), 322-333.

Challenge

$\mathcal{R}(u)$ is typically **uncomputable/unavailable**.

- In strong formulations: $\mathcal{R}(u) = Au - f$.
- In weak formulations: $\mathcal{R}(u)$ might be unknown or impractical.
  (e.g., in SPD problems, $\mathcal{R}(u) = u - u^*$ when the inner product is the bilinear form).

Discretization of the test space

- We consider a conforming discretization for $\mathbb{V}$:

$$\mathbb{V}_M := \operatorname{span}\{v_m\}_{m=1}^{M} \subset \mathbb{V}$$

- We orthogonally project $\mathcal{R}(u) \in \mathbb{V}$ onto $\mathbb{V}_M$, producing $\mathcal{R}_M(u) \in \mathbb{V}_M$.

Loss Function

$$\begin{aligned}
\mathcal{R}(u) &= \mathcal{R}_M(u) + \mathcal{R}(u) - \mathcal{R}_M(u), \qquad \mathcal{R}_M(u) \perp_{\mathbb{V}} \mathcal{R}(u) - \mathcal{R}_M(u), \\
\underbrace{\|\mathcal{R}(u)\|_{\mathbb{V}}^2}_{\substack{\text{Ideal } \mathcal{L} \\ \text{(e.g., in PINNs)}}} &= \underbrace{\|\mathcal{R}_M(u)\|_{\mathbb{V}}^2}_{\substack{\text{Discrete } \mathcal{L} \\ \text{(in VPINNs)}}} + \underbrace{\|\mathcal{R}(u) - \mathcal{R}_M(u)\|_{\mathbb{V}}^2}_{\substack{\text{Discretization error} \\ \text{(oscillation term)}}}
\end{aligned}$$

Rojas, S., Maczuga, P., Muñoz-Matute, J., Pardo, D., & Paszyński, M. (2024). Robust Variational Physics-Informed Neural Networks. Computer Methods in Applied Mechanics and Engineering, 425, 116904.

Computation

In vector form:

$$\boxed{\underbrace{\|\mathcal{R}(u)\|_{\mathbb{V}}^2}_{\text{Ideal } \mathcal{L}} \approx \underbrace{\|\mathcal{R}_M(u)\|_{\mathbb{V}}^2}_{\text{Practical } \mathcal{L}} = \mathbf{r}(u)^T \boldsymbol{G}^{-1}\mathbf{r}(u) =: \|\mathbf{r}(u)\|_{\boldsymbol{G}^{-1}}^2}$$

- $\mathbf{G} := [(v_m, v_n)_{\mathbb{V}}]_{n,m=1}^M$: Gram matrix
- $\mathbf{r}(u) := [(\mathcal{R}(u), v_m)_{\mathbb{V}}]_{m=1}^M$: Discrete residual

Poisson's equation: $-\Delta u = f$

- **Strong form:** $\mathbf{r}(u) = \left[\int_\Omega -\Delta u \cdot v_m - f \ v_m\right]_{m=1}^M$, solving a strong PDE with VPINNs is dumb?!

- **Weak form:** $\mathbf{r}(u) = \left[\int_\Omega \nabla u \cdot \nabla v_m - f \ v_m\right]_{m=1}^M$,

Disclaimer

- The original VPINN proposals considered $\mathbf{G} = \mathbf{I} = \mathbf{G}^{-1}$ obviating the $\mathbb{V}$-metric,

  Kharazmi, E., Zhang, Z., & Karniadakis, G. E. (2019). Variational physics-informed neural networks for solving partial differential equations. arXiv preprint arXiv:1912.00873.

  Kharazmi, E., Zhang, Z., & Karniadakis, G. E. (2021). hp-VPINNs: Variational physics-informed neural networks with domain decomposition. Computer Methods in Applied Mechanics and Engineering, 374, 113547.

  which leads to **unrobustness**...

- ...unless the test basis functions are **purposely chosen orthonormal**. The Deep Fourier Residual Method is a VPINN method with an orthonormal test space discretization.

  Taylor, J. M., Pardo, D., & Muga, I. (2023). A deep Fourier residual method for solving PDEs using neural networks. Computer Methods in Applied Mechanics and Engineering, 405, 115850.

Summary

- VPINNs rely on the **weak formulation** of a PDE. It allows dealing with problem specifications that are "illegal" in strong formulations. ✓

- Absorbing BCs is more natural. ✓

- The loss function entails a **discretization** of the test space — unlike PINNs. The **robustness** holds partially. The **curse of dimensionality** arises.[1] ✗

[1] For proposals addressing distinct trial and test spaces with independent neural networks, see WeakPINNs, Weak Adversarial Networks (WANs), or the Deep Double Ritz Method (D^2RM).