r-adaptivity, deep learning and the deep Ritz method

Simone Appella

Supervisors: Prof. Chris Budd, Dr. Tristan Pryer
Department of Mathematical Sciences
University of Bath
sa2407@bath.ac.uk

Introduction

Physics-informed neural networks (PINNs) have recently become popular as a means of solving ODEs and PDEs by using the tools of deep learning (DL). PINNs are usually referred to as *'meshfree methods'*, as they do rely on the use of collocation points. The adoption of a moving mesh (*r-adaptive*) strategy enables to increase the accuracy of the numerical solution. The aim of this method is to equidistribute the error over the mesh elements.

We will show that different DL settings allow us to solve different tasks, such as one-dimensional mesh equidistribution, convection-dominated ODEs, and the Poisson's equation over an L-shaped domain.

Equidistribution of $u(x)$ by using a deep neural network (DNN)

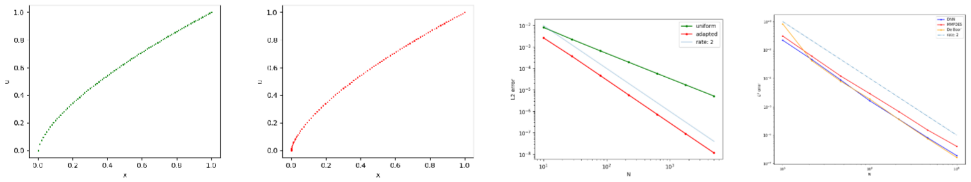
The input of the DNN is one point ξ of a uniform mesh and the output is the new point x given by

$$x(\xi,\Theta) = f_L \circ f_{L-1} \circ \cdots \circ f_0, \quad f_i = \sigma(W_i f_{i-1} + b_i),\ f_0 = \xi \text{ and } \Theta = \{W_i, b_i\}_{i=1}^{L} \quad i = 1, \ldots, L.$$

We train the DNN and enforce **equidistribution** by minimising the loss function $L(\Theta)$ as the upper bound of the L^2 error between exact solution and piecewise linear interpolant:

$$||u - \Pi_1 u||^2_{L^2([0,1])} \leq C \sum_{i}^{N} \left(h_i m_{i+1/2}\right)^5 \equiv L(\Theta),$$

$$\cup_{i=0}^{N}[x_i, x_{i+1}] = [0,1],\ h_i = x_{i+1} - x_i,\ m_{i+1/2} = \frac{m(x_i) + m(x_{i+1})}{2} \quad \text{with } m(x) = (1 + u_{xx}^2)^{1/5}.$$



Numerical Result for $u(x) = x^2/3$: DNN trained with $N = 300$ using Adam optimizer and $\sigma = tanh(\cdot)$. The equidistributed mesh clusters towards x = 0, where the solution exhibits a singular behaviour. The L^2 convergence rate is optimal even when N is greater than the training sample size.

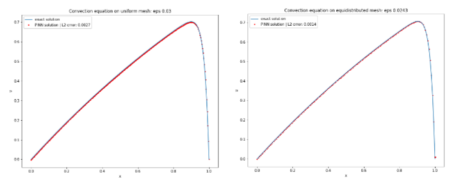
The DNN can be trained to learn the equidistribution process, and outperforms other standard numerical methods.

Convection-dominated ODE

The mesh-less PINN takes randomly sampled points x_k as input and returns the approximate solution $U(x,\Theta)$. We aim to minimise the loss function $L(\Theta) \equiv \sum_{k=1}^{N} |\mathcal{L}[U(x_k,\Theta)] - e^{-x/4}|^2$, where $\mathcal{L}[\cdot]$ is the operator defined as

$$\mathcal{L}[U] = -\varepsilon U_{xx} + \left(1 - \frac{\varepsilon}{2}\right) U_x + \frac{1}{4}\left(1 - \frac{1}{4}\varepsilon\right) U,$$

Homotopy method: Given an initial uniform mesh of $2N$ points, we obtain a mesh clustered towards the singularity by reducing logarithmically ε on the domain $[0,1] = [0, 1-2\varepsilon] \cup [1-2\varepsilon, 1]$.



PINN trained with Adam optimizer ($lr = 1e-3$).
uniform $N = 300$ | equidistributed $N = 150$.

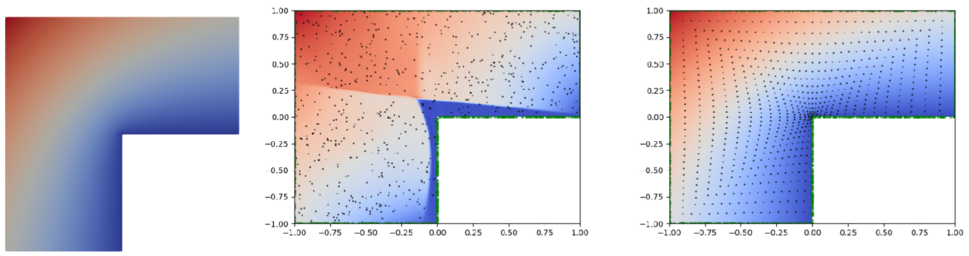
For convective problems, PINNs need homotopy methods to work at all.

Deep Ritz method for the Poisson's equation: $\Delta u(\vec{x}) = 0$ on Ω_L $\quad u_D(r,\theta) = r^{2/3} sin(2\theta/3)$ on $\partial\Omega_L$

The deep Ritz method (DRM) mimics the action of a PINN by solving the minimisation problem

$$U = \arg\min_{v \in H} \mathcal{I}(v), \quad \mathcal{I}(U) = \int_{\Omega} \left(\Delta U(\vec{x}) + f(\vec{x})\right)^2 d\vec{x} + \beta \int_{\partial\Omega} (U(\vec{x}) - u_D)^2 d\vec{x},\ \beta > 0 \text{ penalty parameter.}$$

The exact solution $u(\vec{x})$ has a gradient singularity at the interior corner $\vec{0}$. Given the interior angle ω and the distance from the corner r, the solution for $\theta \approx 0$ is $u(r,\theta) \sim r^{\alpha} f(\theta)$, with $\alpha = \frac{\pi}{\omega}$ and $f(\theta)$ a regular function. We can solve the Monge-Ampère equation locally at the interior corner (semi-analytically) to find the OT-based collocation points. In general, r-adaptivity in $\mathbb{R}^d$ using OT can be obtained by minimising the Wasserstein distance.



Exact solution DRM with randomly sampled points | OT collocation points.

The relative L^2 error is computed by evaluating the approximate solution on a Delaunay mesh - random points: 0.468 | OT-based points: 0.0639.

OT-based r-adaptivity is very effective for 2D problems using the deep Ritz method.

References

[1] C. Budd, W. Huang, and R. Russell: *Adaptivity with moving grids.* In: Acta Numerica 18 (2009).
[2] A. Krishnapriyan et al. : *Characterizing possible failure modes in physics-informed neural networks.* NeurIPS (2021).
[3] E. Weinan and Yu Bing: *The Deep Ritz method: A deep learning numerical algorithm.* In: Communications in Mathematics and Statistics (2017).
[4] S. Appella, C. Budd and T. Pryer: *Adaptive meshes in non-convex domain using h-adaptive and Optimal Transport methos.* In preparation.