

PINNuclear-Neutrons 项目笔记

刘洋¹

1. 四川师范大学数学科学学院, 四川成都 610066

E-mail: mathliuyang@163.com

通信作者简介: 刘洋 (1997-), 男, 研究生, 主要研究方向: 偏微分方程与数学物理

摘要 PINNuclear-Neutrons 项目旨在学习和复现核反应堆中子相关问题的解决方法, 使用物理反应神经网络 (PINN) 和深度机器学习技术。我们的主要目标是重新实现刘东老师团队发表的“基于 PINN 深度机器学习技术求解多维中子学扩散方程”, 以更准确地模拟中子传输。鼓励对核能技术和深度学习感兴趣的研究人员和学生参与到这个项目中。在各个子目录中, 您将找到有关每个主题的详细说明和示例。

关键词 基于物理信息指引的神经网络模型; 深度学习; 核反应堆; 中子学扩散方程

引言

随着人工智能技术的迅猛发展, 深度学习神经网络 (DNNs) 在解决复杂偏微分方程 (PDE) 的领域引起了广泛的兴趣。在核反应堆设计和工程领域, 中子学扩散模型是核反应堆工程设计的重要组成部分。传统的中子扩散模型采用有限差分法 (FDM) 和有限元法 (FEM) 等数值方法进行求解, 这些方法在工程实践中取得了显著的成功。然而, 随着深度学习技术的不断发展, 一种新的方法已经崭露头角, 即基于物理信息指导的神经网络模型 (PINN), 它在解决微分方程方面展现出巨大潜力。DNNs 首次用于解微分方程是在 1998 年由 Lagaris 提出的 [1], 后来由 Raissi 和 Karniadakis 进一步发展成为基于物理信息的神经网络 (PINNs) [2], 将方程和边界条件等物理信息作为损失函数引入到神经网络训练当中, 通过这种方法将问题的物理信息考虑进去, 不再仅仅依靠数据驱动的方法来使得网络拟合偏微分方程的解函数; 同时, 此研究也指出如果将方程中的系数当作未知参数也可以完成发现方程的任务。此后, 便有许多研究围绕基于物理信息的神经网络展开, 而这些研究通常可以总结为调整神经网络结构, 更改损失函数的形式, 挑选不同的激活函数, 采用不同的优化策略。PINN 模型充分利用了深度学习神经网络的通用逼近能力, 同时能够处理多维高阶复杂微分方程。与传统数值方法相比, PINNs 具有多方面的优势, 包括无需事先收集大量训练数据、适用于正向和反向问题、无空间几何限制等。因此, PINNs 已经引起了广泛的研究兴趣, 并在热传递、结构动力学、流体力学、固体力学以及核反应堆动力学等多个领域取得了重要突破。

本 PINNuclear-Neutrons 项目笔记旨在深入理解 PINN 模型的原理和应用，通过推导和复现刘东老师的研究工作 [3]，进一步探讨 PINN 模型在核反应堆设计和工程中的潜在应用。我们将重点关注 PINN 模型的工作原理、关键参数的影响以及可能的优化方向，以便更好地理解和应用这一创新方法。通过本学习笔记，我们希望能够为核反应堆设计和工程领域的研究人员提供有关 PINNs 的深入见解，以应对复杂问题和挑战。

1 相关理论和技术

本章的主要内容是介绍相关理论和技术，首先介绍基于物理信息的神经网络，通过这种方法可以将物理信息引入到神经网络当中，然后介绍了本文方法中所用到的二阶优化函数 L-BFGS，为了网络更快更好的优化。最后介绍了核反应堆物理的相关知识，为后文的方法介绍做好理论上的准备。

1.1 基于物理信息的神经网络

使用神经网络逼近函数从万能逼近定理被提出之后就一直是科学计算的重要方向，然而随着数据时代的到来，大数据在人工智能领域取得了巨大的成功，于是人们想通过数据驱动的方式来拟合函数，但对于某些难以求解的函数，函数的解析解和数值解很难获得，无法大量产生训练神经网络所需要的数据。并且数据驱动的方法还会降低模型的泛用性，只有在有监督的点的位置能达到对应的精度，而在做推断位置的点所求出的值精度就会与训练时相差很多，直至采点够密集才能使求解精度达到对应要求。

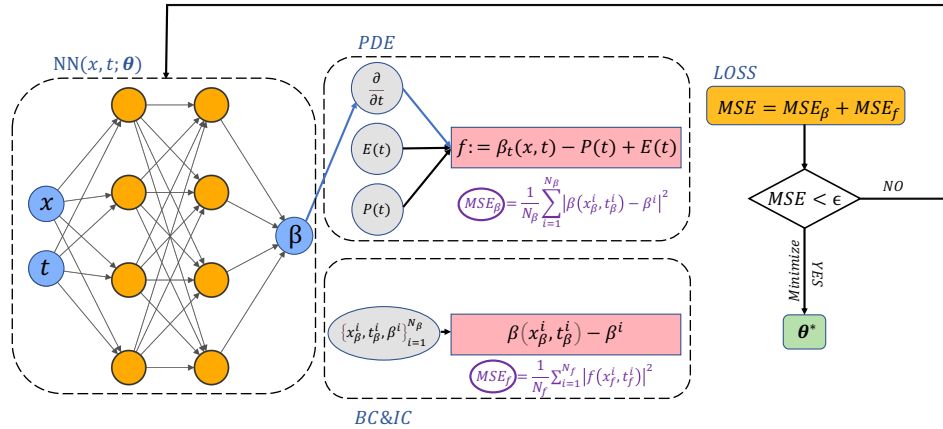


图 1 PINN 网络结构图
Figure 1 PINN Network Structure

为此 Raissi 等人在 2019 年提出了基于物理信息的神经网络 (PINN) [2]，通过将物理方程条件作为损失函数，将物理方程信息参与到神经网络的训练过程之中。当网络训练时，神经网络迭代优化所反向传播的不仅仅有数值上的损失，同时还有物理方程的损失，这样就可以对神经网络所表示的空间进行限制，逐渐去逼近真正的解空间。而且有了物理信息的引入，网络对数据的要求也逐渐降低，不再完全依赖数据驱动，在有很少的数据，甚至没有数据的情况下依然能的到很好的回归效

果,大大增加了网络的泛用性和稳定性。具体神经网络结构如下图 1 所示。如图中所示定义损失函数为:

$$MSE = MSE_u + MSE_{BC,IC} + MSE_R \quad (1)$$

其中 MSE_R 是解的数值约束,只对能取到解的值的点进行约束,通过数值约束来使得神经网络拟合对应的函数; $MSE_{BC,IC}$ 是初边值条件的约束,当一个点来自于初边值中任意一个点时,需要满足其对应的初边值条件,通过边值条件来限制解空间; MSE_u 是方程本身的约束,对于内部区域的点,都需要满足微分方程,通过微分方程来限制解空间。当损失函数整体趋于零时,可以认为神经网络很好的拟合了满足初边值条件和方程约束的解,所以此时解方程问题变成了满足对应损失函数的优化问题,可以使用神经网络的优化方法来求得对应方程的解。

1.2 L-BFGS 优化算法

常用的人工智能优化算法如 Adam,SGD 等都是在一阶梯度下降法 (Gradient Descent) 上改进的,而 L-BFGS 是基于二阶优化算法牛顿法改进的。牛顿法作为二阶的优化算法收敛速度要远快于一阶算法的,但是牛顿法需要计算海森矩阵的逆矩阵,运算成本很高,所以针对这个问题,人们提出了 BFGS 算法和 L-BFGS 算法降低内存的需求,在介绍此算法前,先介绍牛顿法的过程。

设 $f(x)$ 的零点为 \tilde{x} , 随机选取初始点 x_0 , 做 $f(x)$ 的切线得到 $y = f(x_0) + f'(x_0)(x - x_0)$, 令 $y = 0$, 得到 $x_1 = x_0 - \frac{f(x_0)}{f'(x_0)}$, 由此我们可以得到递推公式 $x_{n+1} =$ 用目标函数 $f(x)$ 的二阶泰勒展开求驻点, 即 $f(x) \approx f(x_n) + f'(x_n)(x - x_n) + \frac{1}{2}f''(x_n)(x - x_n)^2$, 对其进行求导可得 $f'(x_n) + f''(x_n)(x - x_n) = 0$, 由此更新 x 来得到是目标函数最小的点。

当 $f(x)$ 是多元函数时, x 与 $f'(x)$ 为向量, 记 $f'(x)$ 为 g , $f''(x)$ 为海森矩阵记为 H , 则牛顿法在多元情况下的迭代公式变为 $x_{k+1} = x_k - g_k H_k^{-1}$, $k = 0, 1, \dots$, 其中 H_k^{-1} 是海森矩阵的逆, 而直接求海森矩阵的逆十分困难, 所以 BFGS 算法提出通过迭代来近似 H_k^{-1} , 即:

$$B_{k+1}^{-1} = \left(I - \frac{\delta_k y_k^T}{y_k^T \delta_k} \right) B_k^{-1} \left(I - \frac{y_k \delta_k^T}{y_k^T \delta_k} \right) + \frac{\delta_k \delta_k^T}{y_k^T \delta_k} \quad (2)$$

其中 B_{k+1}^{-1} 是对 H_{k+1}^{-1} 的逼近, $B_0^{-1} = I$, $\delta_k = x_{k+1} - x_k$, $y_k = g_{k+1} - g_k$ 。利用 B_k^{-1} 的迭代公式, 每次都需要储存 B , 当数据维度很大时, 对计算机的内存需求也变得很高, 于是 L-BFGS 方法提出只储存 δ_k 与 y_k , 每个 B_k^{-1} 都可由 $\delta_1, \dots, \delta_k$ 与 y_1, \dots, y_k 迭代计算得出, 但当 k 过于大时, 计算机内存仍然会面临内存不足的风险, 于是在算法中可以设定一个超参数 N , 当 k 超过 N 时, 我们释放前 $k - N$ 项, 仅用最近的 N 项来逼近 B_k^{-1} , 虽然会损失部分精度, 但可以在有限的内存条件下完成函数的优化。

L-BFGS 仍然面临一个问题, 就是当初始点与函数最优解的点相距过远时, L-BFGS 算法可能会收敛到鞍点, 所以在本文实际应用时会先用 Adam 优化算法寻找一个最优解可能存在的区域, 再使用 L-BFGS 优化方法加快收敛的速度, 得到更好的优化结果。

1.3 核反应堆物理

核反应堆是一种能以可控方式实现自续链式核反应的装置。根据原子核产生能量的方式, 可以分为裂变反应堆和聚变反应堆两种 [4]。当今世界上已建成和广泛使用的反应堆都是裂变反应堆, 聚

变反应堆目前尚处于研究设计阶段。裂变反应堆通过把一个重核裂变为两个中等质量核而释放能量。它是由核燃料、冷却剂、慢化剂、结构材料和吸收剂等材料组成的一个复杂系统。按用途不同, 裂变反应堆可分为生产堆、实验堆和动力堆。按冷却剂或慢化剂的种类不同可分为轻水堆、重水堆、气冷堆和液态金属冷却快中子增殖堆。按引起裂变反应的中子能量不同, 又可分为热中子反应堆和快中子反应堆。

1.4 本章小结

本章主要对本文所需要用到的理论基础与技术进行了简单地介绍。了解了本文方法将物理信息引入神经网络的理论基础, 并对本文使用的网络结构的理论基础进行介绍, 最后介绍了一种不常用的二阶优化算法 L-BFGS, 为后文的方法介绍做好了理论方面的准备。

2 结束语

本文主要介绍了 PINNuclear-Neutrons 项目的学习笔记, 主要是对 PINNuclear-Neutrons 项目的学习过程进行了总结, 复习了论文中的算例代码, 并对项目中的一些关键代码进行了解读, 对项目中的一些关键参数进行了调整, 以便更好地理解论文中的算法。

参考文献

- 1 LAGARIS, I. E., LIKAS, A., AND FOTIADIS, D. I. Artificial Neural Networks for Solving Ordinary and Partial Differential Equations. *IEEE Trans. Neural Netw.* 9, 5 (Sept./1998), 987–1000.
- 2 RAISSI, M., PERDIKARIS, P., AND KARNIADAKIS, G. E. Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *Journal of Computational Physics* 378 (Feb. 2019), 686–707.
- 3 刘东, 罗琦, 唐雷, 安萍, AND 杨帆. 基于 PINN 深度机器学习技术求解多维中子学扩散方程. *核动力工程* 43, 2 (2022), 1–8.
- 4 谢仲生. 核反应堆物理分析, 第五版 ed. 西安交通大学出版社, 2020.

PINNuclear-Neutrons Project Notes

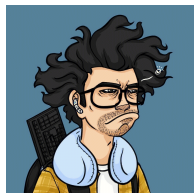
Liu Yang¹

1. School of Mathematical Sciences, Sichuan Normal University, Chengdu 610068, Sichuan

E-mail: mathliuyang@163.com

Abstract PINNuclear-Neutrons project aims to learn and reproduce the solution of neutron related problems in nuclear reactor, using physical reaction neural network (PINN) and deep machine learning technology. Our main goal is to re-implement the "Solving Multidimensional Neutron Diffusion Equation Based on PINN Deep Machine Learning Technology" published by Liu Dong's team, so as to simulate neutron transport more accurately.

Keywords PINN; Deep Learning; Nuclear Reactor; Neutron Diffusion Equation



Liu Yang was born in Sichuan in 1997. He is currently a master's student in the School of Mathematical Sciences, Sichuan Normal University. His research interests include partial differential equations and mathematical physics.

