Simpósio de Integração Acadêmica

Inteligência Artificial: A Nova Fronteira da Ciência Brasileira

SIA UFV Virtual 2020

UFV Universidade Federal de Viçosa

APLICAÇÕES DE REDES NEURAIS PROFUNDAS NA SOLUÇÃO DE EQUAÇÕES DIFERENCIAIS PARCIAIS DA FÍSICA TEÓRICA

Universidade Federal de Itajubá (UNIFEI) - Instituto de Física e Química (IFQ)

Daniel Mendes - adsmendesdaniel@unifei.edu.br; **Alan Pavan (orientador)-** alan@unifei.edu.br;

Física Computacional, Ciências Exatas e da Terra; Trabalho de Pesquisa.

Palavras-chave: ***machine learning, deep learning, redes neurais, física computacional, física teórica*.**

Introdução

A aplicação de redes neurais na solução de problemas em Física tem se popularizado bastante nos últimos anos. Propostas por Raissi, *et al* em 2017, as *Physics Informed Neural Networks* (PINN) são redes neurais treinadas para realizar tarefas de aprendizado supervisionado enquanto respeitam "leis da física" impostas por equações diferenciais parciais não lineares (EDPs), que modelam problemas físicos. Com elas é possível resolver dois tipos de problemas baseando-se em dados: Obter uma solução numérica da EDP e descobrir, numericamente, a equação diferencial parcial em si. Portanto, ambas as possibilidades são de interesse aos físicos.

Objetivos

As PINN podem se mostrar boas candidatas para se inferirmos soluções numéricas para EDPs que possuem soluções, conhecidas ou não, de áreas da Física como Gravitação e Relatividade Geral. Em especial, buscar soluções para as equações diferenciais resultantes da perturbação de buracos negros.

Estes problemas necessitam de estratégias mais elaboradas e eficientes para o treinamento das PINN já que suas EDPs são complexas. Testes primários mostraram que o treinamento padrão leva a um tempo de computação impraticável.

Métodologia

Para se tratar um problema usando as PINN deve-se escrever a EDP na forma genérica da equação

$$h_t + N[h(t,x)] = 0, x \in \Omega, t \in [0,T], \quad (1)$$

onde $h(t,x)$ representa a solução da EDP, que será aproximada por uma rede neural profunda, $N[\cdot]$ é um operador diferencial (não) linear e Ω é um subconjunto de $\mathbb{R}^D$ que representa as condições de contorno. Uma função de custo é construída com a equação (1) e as condições de contorno e iniciais.

Para treinarmos a rede devemos minimizar a equação

$$MSE = MSE_h + MSE_f \quad (2)$$

onde MSE_h é o erro médio quadrático da solução numérica da rede neural com relação aos dados das condições iniciais e de contorno, e MSE_f é o erro da solução numérica com relação à EDP dada, com derivadas calculadas através de diferenciação automática. Raissi, *et al* propõe, para o treino da rede, a utilização de um otimizador conhecido como L-BFGS. Porém, para equações mais complexas, essa estratégia torna o tempo de computação excessivamente longo.

Desta forma, propomos a criação de um laço que fará uma etapa de treino com o otimizador ADAM, para tentar escapar de mínimos locais, e uma etapa com L-BFGS para refinar o resultado.

Além disso, propomos que uma simplificação da EDP seja feita, excluindo termos ou alterando parâmetros, para que a PINN gere uma curva mais simples e, consequentemente, mais fácil de ser obtida. Em seguida retorna-se a EDP original e treina-se novamente a PINN pré-treinada, para que se consiga obter uma curva mais elaborada mais facilmente.

O laço poderá ser interrompido quando uma determinada precisão ou tempo de computação forem alcançados.

Resultados e Conclusões

As estratégias por nós apresentadas acima foram testadas com a equação de um oscilador harmônico amortecido forçado por possuírem um resultado semelhante ao obtido em perturbações de buracos negro. Elas demonstraram uma redução de até 54% no tempo de treinamento da PINN. Abaixo apresentamos a evolução dos resultados obtidos para a oscilação (azul) e sua derivada (laranja) ao longo das épocas de treinamento.

Figura 1 - O treinamento proposto por Raissi leva 4.68 horas para obter MSE menor que 5.0×10^{-6} em uma GPU Tesla K80.

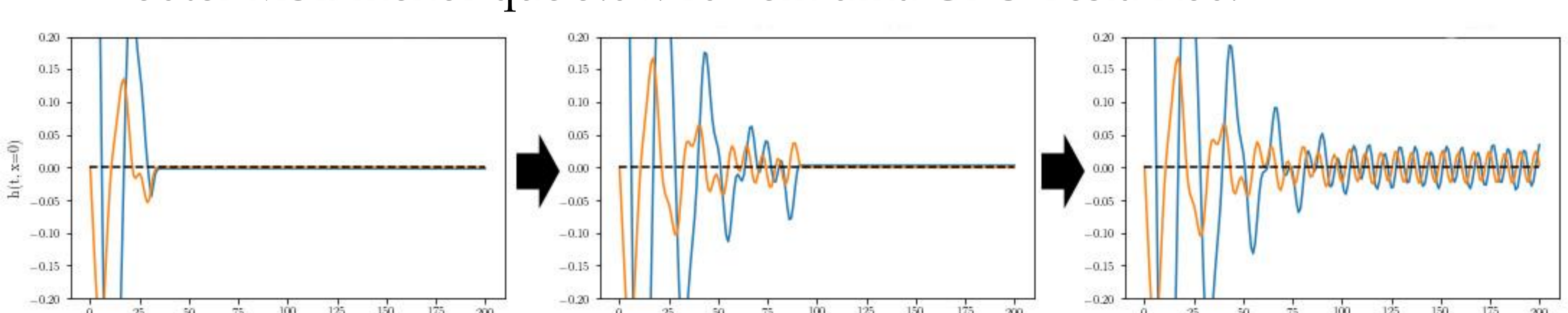
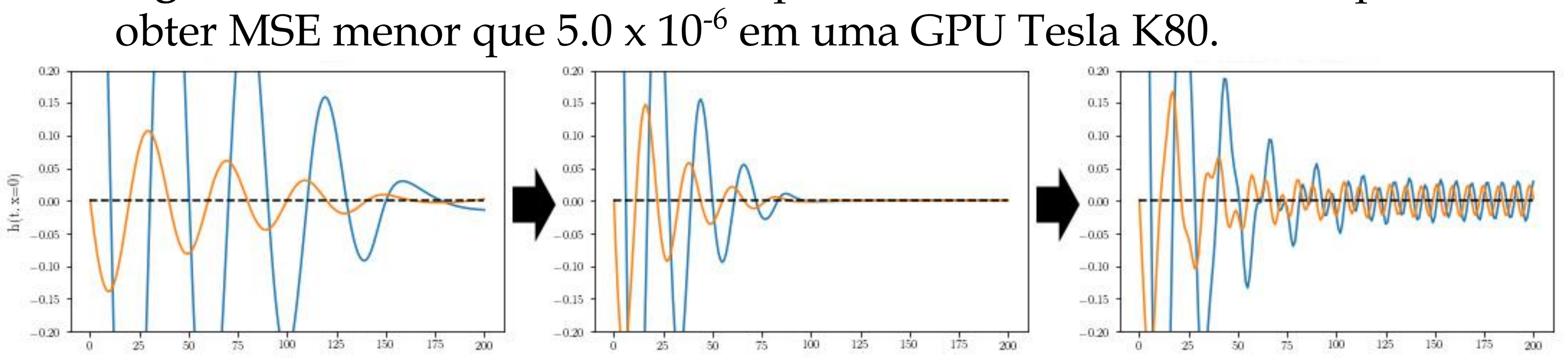


Figura 2 - Nosso método leva aproximadamente 2.16 horas para obter MSE menor que 5.0×10^{-6} em uma GPU Tesla K80.



Fonte: Daniel Mendes (2020)

Testes com a EDP resultante da perturbação escalar de um buraco negro de Schwarzschild estão em andamento.

Bibliografia

RAISSI, Maziar, et al. "Physics Informed Deep Learning (Part I): Data-driven Solutions of Nonlinear Partial Differential Equations". *arXiv:1711.10561, 2017.*

Apoio Financeiro

CNPq Conselho Nacional de Desenvolvimento Científico e Tecnológico