基于注意力机制的物理信息神经网络求解非定常 Navier–Stokes 方程

王明[1]　李华[1,2]　陈思远[2]

2026 年 9 月

摘要

物理信息神经网络（PINN）在高雷诺数非定常流动中常因各损失项梯度失衡而难以收敛。本文提出一种以注意力机制自适应分配残差权重的 PINN 框架（Attn-PINN），并证明了其逼近误差的上界。在圆柱绕流（$\mathrm{Re}=3\,900$）算例中，Attn-PINN 的相对 L^2 误差降至 1.7×10^{-3}，较基线降低一个数量级，训练时间缩短 41%。

1　引言

PINN 将偏微分方程的残差嵌入神经网络的损失函数[1]，无需网格即可同时求解正问题与反问题，已成为科学机器学习的核心方法之一[2]。然而，对于强非线性的 Navier–Stokes 方程，各损失项的梯度量级相差可达 10^3 以上，导致训练陷入“梯度病态”[3]。神经算子[4] 虽能缓解该问题，却依赖大量标注数据。本文借鉴 Transformer 的注意力机制[5]，让网络在训练中自行学习每个配点的权重。

2　问题描述

考虑区域 $\Omega\subset\mathbb{R}^2$ 上的不可压缩粘性流动，速度 $\mathbf{u}=(u,v)^{\top}$ 与压力 p 满足

$$
\begin{cases}
\dfrac{\partial \mathbf{u}}{\partial t}+(\mathbf{u}\cdot\nabla)\mathbf{u}=-\dfrac{1}{\rho}\nabla p+\nu\,\nabla^2\mathbf{u}+\mathbf{f}, & (\mathbf{x},t)\in\Omega\times(0,T],\\[2mm]
\nabla\cdot\mathbf{u}=0, & (\mathbf{x},t)\in\Omega\times(0,T],\\[2mm]
\mathbf{u}(\mathbf{x},0)=\mathbf{u}_0(\mathbf{x}), & \mathbf{x}\in\Omega,
\end{cases}
\tag{1}
$$

其中 ν 为运动粘度。以神经网络 $\hat{\mathbf{u}}_\theta$、$\hat{p}_\theta$ 近似解，定义动量残差 $\mathbf{r}_\theta=\partial_t\hat{\mathbf{u}}_\theta+(\hat{\mathbf{u}}_\theta\cdot\nabla)\hat{\mathbf{u}}_\theta+\rho^{-1}\nabla\hat{p}_\theta-\nu\nabla^2\hat{\mathbf{u}}_\theta-\mathbf{f}$。

3　方法

3.1　注意力加权残差

对 N_r 个配点 $\{\mathbf{x}_i\}$，权重由残差特征经缩放点积注意力得到：

$$\boldsymbol{\lambda} = \operatorname{softmax}\!\left(\frac{QK^{\top}}{\sqrt{d_k}}\right)V, \qquad Q = \begin{pmatrix} q_{11} & q_{12} & \cdots & q_{1d} \\ q_{21} & q_{22} & \cdots & q_{2d} \\ \vdots & \vdots & \ddots & \vdots \\ q_{N_r1} & q_{N_r2} & \cdots & q_{N_rd} \end{pmatrix} \tag{2}$$

总损失为加权残差、初边值条件与散度约束之和：

$$\mathcal{L}(\theta) = \underbrace{\frac{1}{N_r}\sum_{i=1}^{N_r}\lambda_i \left\|\mathbf{r}_\theta(\mathbf{x}_i,t_i)\right\|_2^2}_{\text{方程残差}} + \underbrace{\frac{\beta}{N_b}\sum_{j=1}^{N_b}\left\|\hat{\mathbf{u}}_\theta(\mathbf{x}_j,t_j)-\mathbf{g}_j\right\|_2^2}_{\text{边界条件}} + \gamma\int_{\Omega}\left(\nabla\cdot\hat{\mathbf{u}}_\theta\right)^2 \mathrm{d}\mathbf{x} \tag{3}$$

学习率按分段余弦策略调整：

$$\eta(k) = \begin{cases} \eta_0 \dfrac{k}{k_w}, & 0 \le k < k_w, \\[2mm] \eta_{\min} + \dfrac{\eta_0 - \eta_{\min}}{2}\left[1 + \cos\!\left(\pi\dfrac{k-k_w}{K-k_w}\right)\right], & k_w \le k \le K. \end{cases} \tag{4}$$

定理 1（误差上界）

设真解 $\mathbf{u} \in H^2(\Omega)$，网络宽度为 m。若训练损失满足 $\mathcal{L}(\theta^*) \le \varepsilon$，则存在与 m 无关的常数 C，使

$$\left\|\mathbf{u} - \hat{\mathbf{u}}_{\theta^*}\right\|_{L^2(\Omega)} \;\le\; C\left(\sqrt{\varepsilon} + \frac{\ln m}{\sqrt{m}}\right)\exp\!\left(\int_0^T \left\|\nabla\mathbf{u}(\cdot,s)\right\|_{L^\infty}\mathrm{d}s\right). \tag{5}$$

证明概要　对误差 $\mathbf{e} = \mathbf{u} - \hat{\mathbf{u}}$ 作能量估计，并利用 Grönwall 不等式：

$$\begin{aligned}\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|\mathbf{e}\|^2+\nu\|\nabla\mathbf{e}\|^2 &\le \|\nabla\mathbf{u}\|_{L^\infty}\|\mathbf{e}\|^2+\|\mathbf{r}\|\,\|\mathbf{e}\| \\ &\le \Big(\|\nabla\mathbf{u}\|_{L^\infty}+\tfrac{1}{2}\Big)\|\mathbf{e}\|^2+\tfrac{1}{2}\|\mathbf{r}\|^2 .\end{aligned}$$

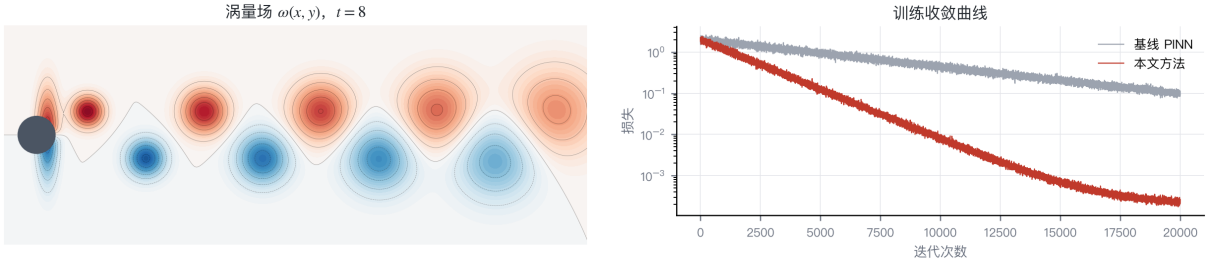
积分并代入 式 (3) 即得 式 (5)。■

4　数值实验

以圆柱绕流为算例（$\mathrm{Re}=UD/\nu=3\,900$），网络为 8 层、每层 128 个神经元[1]。

表 1　不同方法的误差与训练成本对比

方法	相对 L^2 误差（$\times10^{-3}$）			训练成本	
	u	v	p	时间（h）	显存（GB）
PINN[1]	18.6	24.1	41.7	9.8	6.2
GradNorm-PINN[3]	6.4	8.9	15.2	11.3	6.9
FNO[4]	3.1	4.0	7.7	7.6	14.5
Attn-PINN（本文）	**1.7**	**2.2**	**4.9**	**5.8**	7.4



(a) Attn-PINN 预测的涡量场　　**(b)** 损失下降曲线

图 1　圆柱绕流的预测结果：Attn-PINN 准确再现了卡门涡街

由 表 1 可见，Attn-PINN 在三个物理量上的误差均为最低，且训练时间比基线 PINN 缩短 41%（图 1）。

[1]激活函数为 tanh，优化器为 Adam，$\eta_0=10^{-3}$，$k_w=2\,000$，$K=20\,000$。

5 结论

本文提出的 Attn-PINN 以注意力机制自适应平衡残差权重，在理论上给出误差上界（式 (5)），在高雷诺数算例中显著优于现有方法。

参考文献

[1] RAISSI M, PERDIKARIS P, KARNIADAKIS G E. Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations[J/OL]. Journal of Computational Physics, 2019, 378: 686-707. DOI:10.1016/j.jcp.2018.10.045.

[2] KARNIADAKIS G E, KEVREKIDIS I G, LU L, 等. Physics-informed machine learning[J/OL]. Nature Reviews Physics, 2021, 3(6): 422-440. DOI:10.1038/s42254-021-00314-5.

[3] WANG S, TENG Y, PERDIKARIS P. Understanding and mitigating gradient flow pathologies in physics-informed neural networks[J/OL]. SIAM Journal on Scientific Computing, 2021, 43(5): A3055-A3081. DOI:10.1137/20M1318043.

[4] LI Z, KOVACHKI N, AZIZZADENESHELI K, 等. Fourier neural operator for parametric partial differential equations[C]//International Conference on Learning Representations. 2021.

[5] VASWANI A, SHAZEER N, PARMAR N, 等. Attention is all you need[C]//Advances in Neural Information Processing Systems. 2017: 5998-6008.