文章编号:1674-2974(2024)11-0033-10 DOI:10.16339/j.cnki.hdxbzkb.2024105

基于物理驱动深度学习的结构形状优化设计

唐和生[1†],李度[1],廖洋洋[1],李荣帅[2]
(1. 同济大学 土木工程学院,上海 200092;
2. 上海建工集团股份有限公司,上海 200080)

摘 要:结构形状的优化设计本质上是一类泛函极值求解问题.在求解高维度泛函极值问题时,传统的变分法往往面临着求解目标函数类型有限、求解过程呈现振荡行为等问题.利用深度学习模型的高维非线性映射能力,建立了一种基于物理驱动深度学习的泛函极值求解模型.首先将描述结构形状优化问题的物理信息(控制方程、初始条件和边界条件等)作为正则化项嵌入深度学习模型中,基于性能目标构建损失函数;采用随机梯度下降法完成深度学习模型的训练,进而实现泛函极值的求解和结构形状的优化设计;通过分析最优曲面和最优拱轴线问题验证模型的有效性,并与遗传算法进行对比,结果表明该模型在小样本的目标任务上具有较高的预测精度和效率.作为一种非参数模型化技术,物理驱动深度学习模型对解决数据采集成本高、难度大的工程问题具有重要意义.

关键词:物理驱动深度学习;形状优化设计;泛函极值;遗传算法
中图分类号:TU318 **文献标志码**:A

Structural Shape Optimization Design Based on Physics-Informed Deep Learning

TANG Hesheng[1†], LI Du[1], LIAO Yangyang[1], LI Rongshuai[2]
(1. College of Civil Engineering, Tongji University, Shanghai 200092, China;
2. Shanghai Construction Group Co., Ltd., Shanghai 200080, China)

Abstract: The optimization design of structural shapes is fundamentally a problem of solving functional extremum. Traditional variational methods often encounter challenges, such as limited functional types and oscillation in the solution process when solving high-dimensional functional extreme value problems. In this paper, a functional extremum numerical solution method based on physics-informed deep learning (PIDL) is proposed by using the high-dimensional nonlinear mapping ability of deep learning model. The method first embeds the physical information (control equations, initial conditions and boundary conditions, etc.) of the shape optimization problem as regularization terms into the deep learning model, and a loss function based on the objective functional extremum is constructed. Then, a random gradient descent algorithm is used to train the deep learning model, further realizing the solution of functional extremum and optimization design of structural shape. The proposed method is verified

* **收稿日期**:2023-12-08
基金项目:国家自然科学基金资助项目(52378184), National Natural Science Foundation of China(52378184)
作者简介:唐和生(1973—),男,安徽安庆人,同济大学研究员,博士
† 通信联系人,E-mail: thstj@tongji.edu.cn

through numerical examples of optimizing the shape of surfaces and arch axes, and a comparative analysis is conducted with the computational results obtained from genetic algorithms. The results demonstrate that the method has high prediction accuracy and efficiency for the target task of small samples. As a non-parametric modeling technology, the method is of great significance for solving engineering problems characterized by high data acquisition costs and data collection challenges.

Key words: physics-informed deep learning; shape optimization design; functional extremum; genetic algorithm

工程中结构的形状优化是在给定约束条件下,在结构形状的优化空间内求解性能目标函数的极值.结构的形状优化问题涉及建筑工程、结构工程、航天航空和岩土工程等领域,可归结为泛函极值问题,即寻找使对应泛函取极值的目标函数.现阶段,基于Euler-Lagrange方程[1],可求解结构一些简单的泛函极值问题,但适用范围有限.随着工程应用场景越来越复杂,传统数值方法在求解结构形状优化问题时,常面临基函数逼近不充分、收敛困难及数值振荡等问题,这使得传统数值方法的通用性受限.

为了弥补传统数值方法的不足,国内外学者提出了一系列求解泛函极值问题的新型数值方法.Mahdy等[2]通过使用Sumudu变换方法实现泛函极值问题求解;Razzaghi[3]采用傅里叶级数和相关的积分矩阵解决泛函极值问题,降低了整体近似误差;Zhou等[4]使用帕德逼近法在函数空间中高效地搜索目标函数以求解泛函极值问题;Gornov等[5]基于微分方程系统的隐凸性,采用多起点搜索法,提出了在非线性非凸最优控制问题中求泛函最小值的方法.上述算法都是解决泛函极值问题的新型数值方法,适用于多种类型的泛函极值问题,且在求解精度和效率方面有显著提升.

此外,传统数值方法可能会受到高维空间的限制,耗费大量计算资源.近年来,很多学者结合启发式算法对泛函极值问题开展了新的研究.Vogt等[6]用提升算法将泛函极值问题转化为合适的高维空间中的凸问题,从而有效求解全局最优值.Pedregal等[7]提出了一种求解泛函极值问题的启发式算法,用于求解多维约束的泛函极值问题.Ye等[8]开发一种退火算法来定位满足欧拉-拉格朗日方程的变分路径,以求解泛函极值问题.Mera等[9]采用遗传算法(Genetic Algorithm, GA)求解边界检测的泛函极值问题,采用实数编码遗传算法最小化目标泛函.Hosseini等[10]提出了多目标遗传算法,以解决物理学中优化问题.Korda等[11]提出了求解泛函极值问题的进化算法,该方法适用于具有预定义的形状,并且能求解满足边界或其他附加条件的泛函极值问题.

相较于启发式算法,深度学习算法展现出更强的鲁棒性、自适应性和广泛的应用范围.随着计算机科学的飞速发展,深度学习在泛函极值问题上的应用日益广泛.根据通用逼近定理[12],多层神经网络能构建非线性高维函数空间,以求解满足约束条件的泛函极值问题的最优解.Weinan等[13]提出的Deep Ritz深度学习方法,利用变分形式对应的优化问题作为损失函数训练神经网络,实现目标函数优化.Araz等[14]基于TensorFlow 1.14.0[15]开发了深度神经网络Elvet框架来求解泛函极值问题.Qiu等[16]提出了SenseNet模型,采用离散应变传感器网络数据重构拓扑形状.Liu等[17]结合深度学习与多导遗传算法,优化了旋翼翼型气动外形.尽管这些深度学习算法能处理传统数值方法难以解决的高维、非线性问题,但仍面临一些共性问题:一是深度学习算法的训练需要大量数据计算资源,且耗时较长;二是深度学习算法模型的黑箱特性导致可解释性和可靠性需进一步研究.

机器学习作为新兴技术,已广泛应用于数值领域.其中,物理驱动深度学习(Physics-Informed Deep Learning, PIDL)模型备受瞩目,该方法将深度学习与物理方程相结合,以更准确地建模和求解物理问题.物理信息神经网络(Physics-Informed Neural Network, PINN)[18]作为PIDL的一种实现,是解决各种高阶偏微分方程相关问题的有效手段.在PINN基础上,国内外学者相继提出了不同的改进模型,如稀疏PINN(SPINN)[19]、神经同质化PINN(NH-PINN)[20]等.相较于传统数值方法,PIDL方法以采样替代网格剖分,以非线性基底(即神经网络)替代多项式

基底,将求解控制方程转化为优化损失函数的问题来找到控制方程解.相较于传统数值方法,PIDL架构的显著优势在于,可以通过采样点一次性求解整个时空域,而此过程可利用GPU大规模并行化来提高计算效率.目前PIDL方法在流体力学、岩土工程、固体力学等领域得到广泛应用[21-23].

基于此,本文建立了一种基于物理驱动深度学习的结构形状优化求解方法.该方法通过神经网络构建函数空间,在传统的深度学习模型中嵌入形状优化问题的物理知识(控制方程、初始条件和边界条件等),实现了物理知识与数据驱动模型的有机结合.通过分析最优曲面和最优拱轴线问题,验证了该方法在结构形状优化中的有效性.

1　基于PIDL的泛函极值求解模型

1.1　形状优化问题的泛函极值描述

一般而言,考虑$y(x)$为待优化的结构几何特征,在外荷载$f(x)$作用下,结构的基本物理信息涵盖控制方程、初始条件和边界条件等,分别对应式(1)~式(3).

控制方程:

$$\Psi\left[y(x),f(x)\right]_{\Omega}=0 \tag{1}$$

边界条件:

$$\mathrm{B}\left[y(x)\right]_{\partial\Omega}=0 \tag{2}$$

初始条件:

$$\mathrm{I}\left[y(x)\right]_{I_0}=0 \tag{3}$$

结构形状优化旨在达成特定最优性能目标,寻求结构最优形状$y^*(x)$,该问题可视作在函数空间H上,满足物理信息条件的泛函极值问题:

$$y^*=\arg\min_{y\in H} L(y) \tag{4}$$

式中:$L(y)$为目标泛函;y^*为最优函数.

1.2　泛函极值问题的PIDL模型

为解决泛函目标函数类型有限的问题,利用PIDL构造的函数空间,提出基于物理信息的深度学习泛函极值求解方法,将优化问题的物理信息嵌入传统神经网络中.由PIDL方法构建神经网络,通过训练神经网络,得到使目标最优的函数.具体步骤如下:首先,采用神经网络构建函数空间,用于逼近最优函数.其次,将形状优化问题的物理信息嵌入神经网络中.最后,构建包含目标函数与物理约束的损失函数.其中,目标函数项旨在反映结构形状的性能优化目标,物理约束项通过物理信息(控制方程、初始条件和边界条件等)对模型施加约束.在模型训练过程中,首先随机初始化曲面函数,再利用随机梯度下降算法训练深度学习模型,通过训练数据逐步迭代调整参数得到最小化损失函数.通过在优化问题的自定义区间、边界条件、初始条件上采样,利用基于深度学习的随机梯度下降算法进行网络训练,求解泛函极值问题,最终获得结构的最优化形状.

1.3　损失函数构建

用泛函极值问题的变分形式所对应的优化问题作为损失函数来训练神经网络.下面构建损失函数:在函数空间H中,考虑所有光滑函数$y(x)$的集合.若测试函数$y(x)$满足式(1)~式(3),则需要找到一个函数$y(x)\in H$,使得定义在函数空间H上的泛函能够达到最小值或最大值.因此目标泛函和其他物理信息约束的损失函数可分别表示为式(5)~式(9),最终得到PIDL泛函极值问题的损失函数,如式(10)所示.

$$L_{\mathrm{F}}=\int_{\Omega}F\left[\hat{y}(x)\right]\mathrm{d}x \tag{5}$$

$$L_{\mathrm{c}}=\int_{\Omega}C\left[\hat{y}(x)\right]\mathrm{d}x \tag{6}$$

$$L_{\mathrm{f}}=\Psi\left[\hat{y}(x),f(x)\right] \tag{7}$$

$$L_{\mathrm{b}}=\mathrm{B}\left[\hat{y}(x)\right] \tag{8}$$

$$L_{\mathrm{I}}=\mathrm{I}\left[\hat{y}(x)\right] \tag{9}$$

$$L_{\mathrm{loss}}=L_{\mathrm{F}}+W_{\mathrm{f}}L_{\mathrm{f}}+W_{\mathrm{c}}L_{\mathrm{c}}+W_{\mathrm{b}}L_{\mathrm{b}}+W_{\mathrm{I}}L_{\mathrm{I}} \tag{10}$$

式中:L_{F}为目标泛函损失,L_{f}为控制方程损失,L_{c}为泛函约束损失,L_{b}为边界条件损失,L_{I}为初始条件损失,W_{f}为控制方程损失权重,W_{c}为泛函约束损失权重,W_{b}为边界条件损失权重,W_{I}为初始条件损失权重.

1.4　自动微分与积分离散化

物理驱动的深度学习方法训练分别采纳了自动微分算法与积分离散化算法,实现微分与积分计算.自动微分利用符号微分处理基本算子,并通过链式法则整合各部分导数,从而精确评估由计算机程序表达的数值函数的导数[24].自动微分机制包含3部分:链式法则、反向传播和有向图结构.在正向传递中计算各节点的值,在反向传递中利用有向图结构,依据链式法则计算梯度向量.

泛函极值问题在样本空间选取采样点,采用辛普森法则[25]进行求解.PIDL方法作为无监督学习方

法,无须人工标注数据点.

1.5 梯度下降训练算法

神经网络训练旨在通过调整参数实现泛函的最小化,训练算法的应用可加快收敛.模型迭代过程采用批量梯度下降法,具体步骤如下:首先,使用TensorFlow提供的自动微分功能计算目标函数的梯度;其次,进行梯度反向传播,结合Adam算法[26]更新参数并进入下一轮迭代.

Adam算法是梯度下降法的一个变种.该算法结合了一阶动量算法与RMSprop算法[27]的优点,具有更快的收敛速度和避免陷入局部最小值的能力.Adam算法通过一阶动量调整训练方向,通过二阶动量调节学习速率.其参数更新见式(11)~式(13),公式对应的符号说明见表1,其中t表示迭代步.

$$m_t = m_{t-1}\cdot\beta_1 + \nabla f\left[x^{(i)}\right]\cdot\left(1-\beta_1\right) \tag{11}$$

$$v_t = v_{t-1}\cdot\beta_2 + \nabla f\left[x^{(i)}\right]^2\cdot\left(1-\beta_2\right) \tag{12}$$

$$\alpha^{(i+1)} = \alpha^{(i)} + m_t\cdot\frac{\eta}{\sqrt{v_t}+\varepsilon} \tag{13}$$

表1 Adam算法符号说明

Tab.1 Symbol description of Adam algorithm

符号	符号说明	符号	符号说明
m_t	损失函数梯度一阶矩	$\nabla f[x^{(i)}]$	目标函数梯度
v_t	损失函数梯度二阶矩	$\alpha^{(i)}$	在第i次迭代时的参数值
β_1	一阶矩的衰减系数	η	学习率
β_2	二阶矩的衰减系数	ε	极小的常量(防止分母为零)

2 算例分析

2.1 最小旋转曲面问题

已知连续曲线通过两固定点,当该曲线绕横轴旋转时,所生成的旋转曲面面积最小,即对于通过点$A(0,\ \cosh(-1))$和$B(2,\ \cosh(1))$的曲线函数$y=y(x)$,求最优函数$y=y^*(x)$,使旋转曲面面积最小.其表达式为:

$$\begin{aligned}&\min\ \ S[y(x)]\equiv\int y(x)\sqrt{1+\left(\frac{\mathrm{d}y(x)}{\mathrm{d}x}\right)^2}\,\mathrm{d}x\\&\text{s.t.}\ \ y(x_1)-y_1=0\\&\qquad y(x_2)-y_2=0\end{aligned} \tag{14}$$

式中:S为面积泛函,y_1和y_2为边界值.

构建损失函数需考虑结构的基本物理信息.该案例中考虑端点约束条件,求一个测试函数$y=\hat{y}(x)$,使得定义在函数空间H上的泛函达到最小值或最大值,可得损失函数如式(15)~式(17)所示.

$$L_{\mathrm{F}} = \int_{\Omega}F\left[\hat{y}(x)\right]\mathrm{d}x = \int_0^2\hat{y}(x)\sqrt{1+\left(\frac{\mathrm{d}\hat{y}(x)}{\mathrm{d}x}\right)^2}\,\mathrm{d}x \tag{15}$$

$$L_{\mathrm{b}} = \mathrm{B}\left[\hat{y}(x)\right] = \left[\hat{y}(x_1)-y_1\right]^2+\left[\hat{y}(x_2)-y_2\right]^2 \tag{16}$$

$$L(y) = L_{\mathrm{F}} + W_{\mathrm{b}}L_{\mathrm{b}} \tag{17}$$

式中:W_{b}=100.

利用PIDL求解最小旋转曲面问题,训练数据为在样本空间[0, 2]上均匀选取的100个样本点,网络训练采用Adam优化器.神经网络结构包含1个神经元的输入层、10个神经元的隐藏层和1个神经元的输出层,且均为全连接层.常用激活函数有Sigmoid函数、Tanh函数和Relu函数等[28],本文采用Sigmoid函数作为激活函数.求解最小旋转曲面问题的网络结构如图1所示,神经网络超参数设置如表2所示.其中x代表自变量输入值,y代表神经网络输出值,参数w、b分别表示权重和偏置,α表示参数集$\{w,b\}$.由该神经网络创建的函数空间表示为H,其维度等于权重与偏置的个数.

$$\begin{cases}\mathbb{R} & \to & \mathbb{R}\\ x & \mapsto & \hat{y}(x;\ \alpha)\end{cases} \tag{18}$$

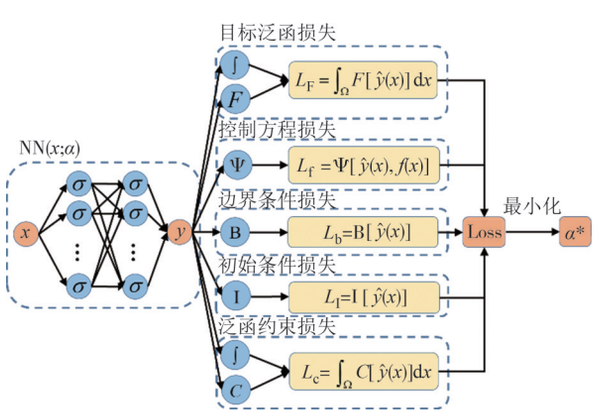


图1 泛函极值问题神经网络图

Fig.1 Neural network diagram for functional extremum problem

显式表达神经网络输出的计算公式如下:

$$\hat{y}(x;\ \alpha) = b_1^{(2)} + \sum_{j=1}^{10}w_{1j}^{(2)}\mathrm{Sigmoid}\left(b_j^{(1)}+w_{j1}^{(1)}x\right) \tag{19}$$

式中:$w_{j1}^{(1)}$、$b_j^{(1)}$与$w_{1j}^{(2)}$、$b_1^{(2)}$分别为第一层、第二层的神经网络的权重与偏置.

表 2　最小旋转曲面问题超参数设置

Tab.2　Hyperparameter settings for the minimal surface of revolution problem

超参数	属性	值
lr	训练学习率	0.001
β_1	一阶矩估计的指数衰减率	0.9
β_2	二阶矩估计的指数衰减率	0.999
ε	用于数值稳定性的常数	1e−7
Epoch	训练总回合数	60 000

训练采用 2.70 GHz Intel/6 核/11 代 Intel Core i5−11 400 H 处理器. 迭代 60 000 次所花费时间为 55.15 s. 训练得到的预测母线表达式如下:

$$\hat{y}(x;\ \alpha) = \boldsymbol{B}^{(2)} + \left(\boldsymbol{W}^{(2)}\right)^{\mathrm{T}}\sigma\left(\boldsymbol{B}^{(1)} + \boldsymbol{W}^{(1)}x\right) \tag{20}$$

式中:σ 表示 Sigmoid 函数,$\boldsymbol{W}^{(1)}$、$\boldsymbol{B}^{(1)}$ 与 $\boldsymbol{W}^{(2)}$、$\boldsymbol{B}^{(2)}$ 分别为第一层、第二层的神经网络的权重与偏置矩阵.

为进行对比研究,本文采用遗传算法搜索最小旋转曲面. 该算法首先随机初始化一定数量的个体,个体基因代表了泰勒展开系数的一组取值. 在每一代中计算旋转曲面面积,并在区间端点引入惩罚项以评估个体适应度. 通过多元竞标赛选择适应度最好的个体,并利用单点交叉生成新的子代. 在子代中进行变异操作,并用新子代替换原有种群中的部分个体,形成新的种群. 通过不断迭代,逐渐优化种群中的个体,寻找具有最优泰勒展开系数的旋转曲面母线,以满足性能优化目标和物理约束条件. 遗传算法的超参数设置如表 3 所示.

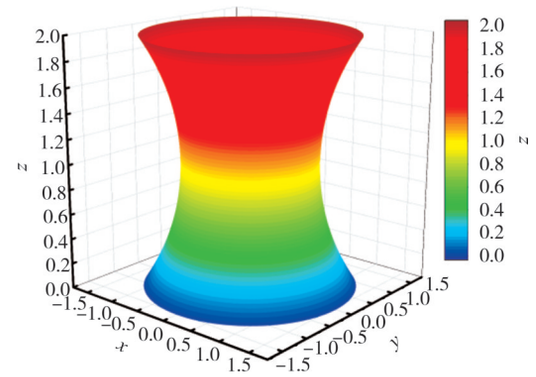
表 3　遗传算法超参数设置

Tab.3　Hyperparameter settings of GA

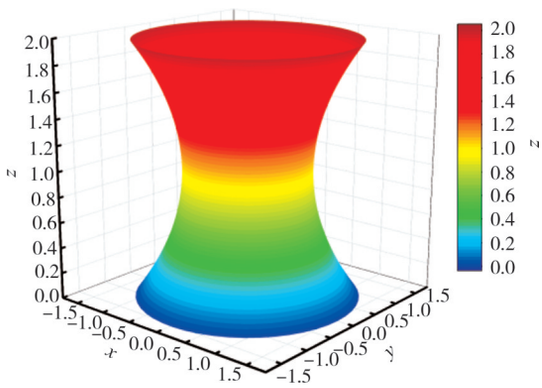
超参数名称	参数取值
种群大小	200
个体基因数量	4
迭代次数	6 000
变异率	0.000 1
变异扰动大小	0.02
多元竞标赛规模	3

遗传算法迭代 6 000 次用时 182.95 s. 最小旋转曲面问题的 PIDL 预测解、GA 预测解和理论计算得到的精确解分别如图 2 所示. 图 3 为最小旋转曲面母线的 PIDL 预测解、GA 预测解和精确解,PIDL 训练迭代收敛曲线如图 4 所示,GA 适应度曲线如图 5 所示.

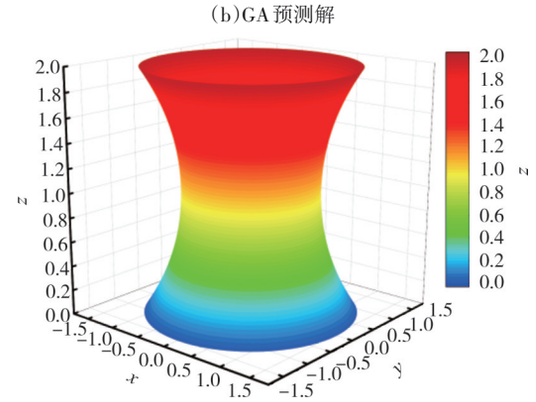
如图 3 所示,与 GA 计算结果相比,PIDL 计算结果更接近精确解. 从图 4 可知,在训练开始阶段损失函数大幅度下降,表明所选择的学习率合适. 训练迭代 60 000 次时损失函数值约为 1×10^{-5}. 从图 5 可知,群体适应度值逐渐上升直至收敛. 采用平方误差评估预测值与精确值之间的误差,其公式为:



(a)PIDL 预测解



(b)GA 预测解



(c)精确解

图 2　最小旋转曲面问题的解

Fig.2　Solutions of minimum rotation surface problem

$$E_{\text{error}} = \left[\hat{y}(x_i;\ \alpha) - y_i\right]^2 \tag{21}$$

图 6、图 7 分别为 PIDL、GA 预测最小旋转曲面母线的平方误差曲线. 由图 6 可知,PIDL 预测结果误差主要集中在[1.2, 1.4]区间内,最大平方误差约为 1.2×10^{-4},同时观察到曲线存在波动. 由图 7 可知,GA

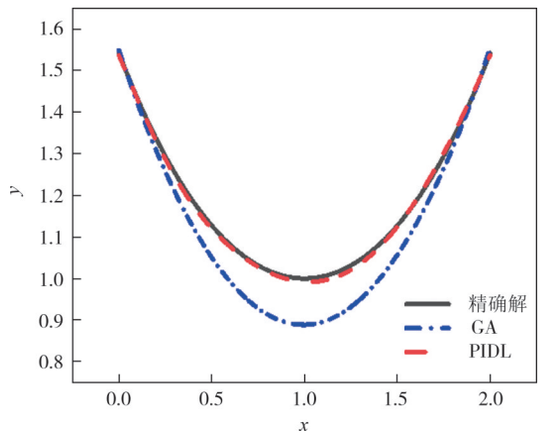


图 3 最小旋转曲面母线的 PIDL 预测解、GA 预测解与精确解

Fig.3 Predictive solution of PIDL, predictive solution of GA and exact solution of the minimum rotation surface generatrix

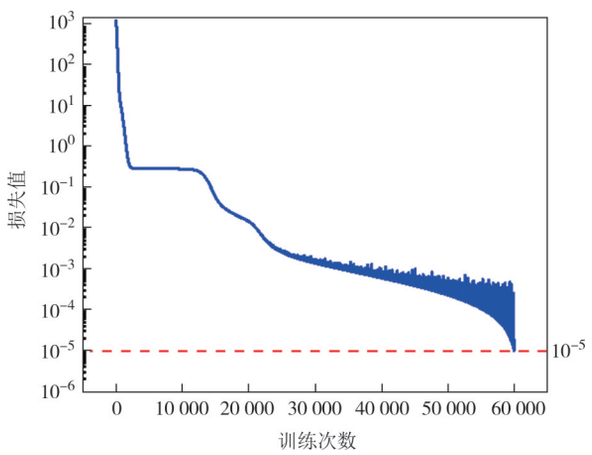


图 4 PIDL 迭代收敛曲线

Fig.4 Iterative convergence curve of PIDL

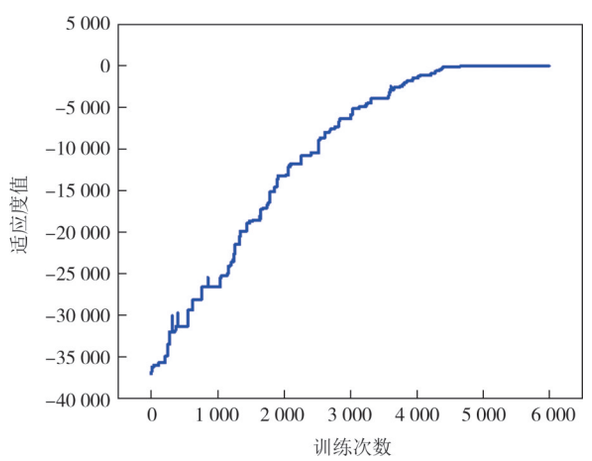


图 5 GA 适应度曲线

Fig.5 Fitness curve of GA

预测结果误差集中在 x=1.0 附近. 最小旋转曲面精确解为 17.677 30, PIDL 与 GA 计算误差对比如表 4 所示. 旋转曲面面积精确结果与 PIDL 预测结果的误差为 0.466%, 与 GA 预测结果的误差为 0.904%, 表明 PIDL 模型具有较高的精度. 同时 PIDL 计算用时 55.15 s, 相较 GA 计算用时更短, 计算效率更高.

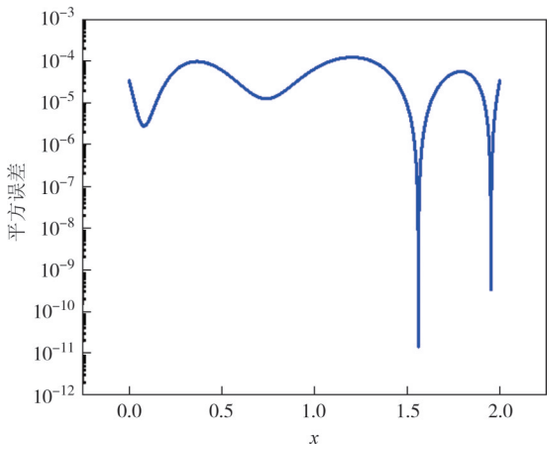


图 6 PIDL 预测母线平方误差曲线

Fig.6 Square error curve of the generatrix predicted by PIDL

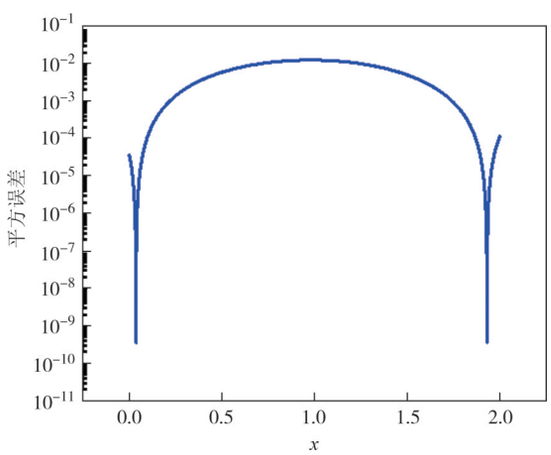


图 7 GA 预测母线平方误差曲线

Fig.7 Square error curve of the generatrix predicted by GA

表 4 最小旋转曲面问题计算误差对比

Tab.4 Comparison of computational errors for minimum rotation surface problem

计算方法	旋转曲面面积	相对误差绝对值/%
PIDL 预测值	17.594 92	0.466
GA 预测值	17.837 02	0.904

2.2 最优拱轴线问题

最优拱轴线问题可描述为: 确定三铰拱的拱轴线, 使其在指定荷载下仅产生轴力. 已知拱上填料的荷载集度计算公式为:

$$q = q_0 + \gamma y \tag{22}$$

建立如图 8 所示的直角坐标系, 三铰拱轴线函数为 y=$y(x)$, 且所有的截面都具有零弯矩. 此外, 三铰拱所受荷载如图 8 所示, 三铰拱的具体参数详见表 5.

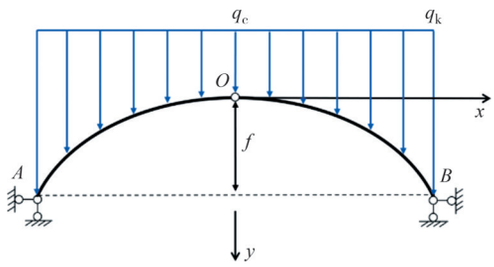


图 8　三铰拱受力图

Fig.8　Force diagram of three-hinge arch

表 5　三铰拱系统的参数

Tab.5　Parameter of three-hinge arch

符号	参数名称	参数取值
γ	填料重度	20 kN/m^2
q_c	拱脚荷载集度	700 kN/m^3
f	拱高度	17.5 m
l	拱跨度	50 m

上述问题可描述为寻找一个最优函数 $y=\hat{y}(x)$，通过点 $O(0,0)$ 满足边界条件，且满足各拱截面弯矩 $M(x)$ 为零. 鉴于对称性，仅考虑 O 点右侧拱轴线. 数学表述如下：

$$\begin{aligned}&\min \int M(x)^2\,\mathrm{d}x\\&\text{s.t. } y(x_0)-y_0=0\\&\quad\frac{\mathrm{d}y(x_0)}{\mathrm{d}x}-\frac{\mathrm{d}y_0}{\mathrm{d}x}=0\end{aligned}\tag{23}$$

式中：$M(x)$ 为对应截面弯矩，y_0 为 O 点对应的边界值.

三铰拱的截面弯矩，可表达为对应简支梁的截面弯矩减去拱脚水平推力产生的弯矩. 计算公式如式(24)所示.

$$M^0(x)-F_{\mathrm{H}}\left[f-\hat{y}(x)\right]=0\tag{24}$$

式中：$M^0(x)$ 为该截面对应简支梁的截面弯矩. 式(22)对 x 求二阶导数可得：

$$F_{\mathrm{H}}\frac{\mathrm{d}^2\hat{y}(x_i)}{\mathrm{d}x^2}-\gamma\hat{y}(x_i)-q_0=0\tag{25}$$

构建损失函数需考虑结构的基本物理信息. 该案例中考虑边界约束条件和控制条件，求一个测试函数 $y=\hat{y}(x)$，使得定义在函数空间 H 上的泛函 $\int M(x)^2\,\mathrm{d}x$ 达到最小值或最大值. 由此可得损失函数如式(26)和式(27)所示.

$$L_{\mathrm{f}}=\Psi\left[\hat{y}(x),f(x)\right]=\sum_{i=1}^{N}\left[F_{\mathrm{H}}\frac{\mathrm{d}^2\hat{y}(x_i)}{\mathrm{d}x^2}-\gamma\hat{y}(x_i)-q_0{}^2\right]^2\tag{26}$$

$$L_{\mathrm{b}}=\mathrm{B}\left[\hat{y}(x)\right]=\left[\hat{y}(x_0)-y_0\right]^2+\left[\frac{\mathrm{d}\hat{y}(x_0)}{\mathrm{d}x}-\frac{\mathrm{d}y_0}{\mathrm{d}x}\right]^2\tag{27}$$

式中：N 为采样点总数. 损失函数如式(28)所示.

$$L(y)=L_{\mathrm{f}}+W_{\mathrm{b}}L_{\mathrm{b}}\tag{28}$$

式中：W_{b}=1.

利用 PIDL 求解最优拱轴线问题，训练数据为在样本空间[0,25]上均匀选取的 100 个采样点. 神经网络结构与最小旋转曲面问题相同，隐藏层中的激活函数采用 Sigmoid 函数，优化器采用 Adam[26]. 神经网络超参数设置如表 6 所示.

表 6　最优拱轴线问题超参数设置

Tab.6　Hyperparameter settings of the optimal arch axis problem

超参数	属性	值
lr	训练学习率	0.001
β_1	一阶矩估计的指数衰减率	0.9
β_2	二阶矩估计的指数衰减率	0.999
ε	用于数值稳定性的常数	1e-7
Epoch	迭代次数	50 000

对于 2.70 GHz Intel/6 核/11 代 Intel Core i5-11400H 处理器，迭代次数为 50 000，用时 16.82 s. 神经网络训练后得到的预测曲线表达式如式(20)所示.

为进行对比研究，采用遗传算法寻找最小旋转曲面. 首先随机初始化一定数量的个体，个体基因为泰勒展开系数的一组取值. 在每一代中计算各采样点处截面弯矩平方和，并引入惩罚项以评估个体适应度. 通过多元竞标赛选择适应度最好的个体. 在子代中进行变异操作，并用新子代替换原有种群中的部分个体. 通过不断迭代，逐渐优化种群中的个体，寻找具有最优泰勒展开系数的最优拱轴线. 遗传算法的超参数设置如表 7 所示.

遗传算法迭代 10 000 次，用时 98.65 s. 最优拱轴

表 7　遗传算法超参数设置

Tab.7　Hyperparameter settings of GA

超参数名称	参数取值
种群大小	300
个体基因数量	3
迭代次数	10 000
变异率	0.000 1
变异扰动大小	0.01
多元竞标赛规模	3

线的PIDL预测解、GA预测解与精确解对比曲线如图9所示.

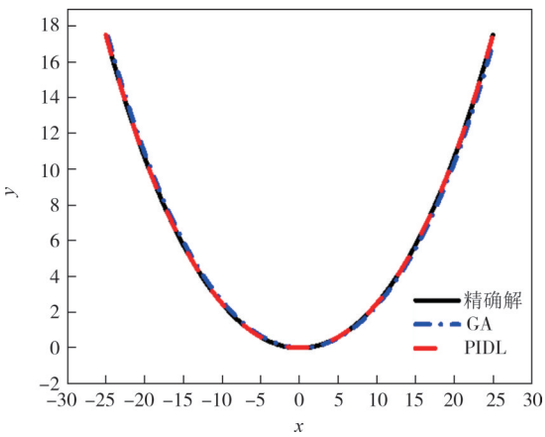


图9 最优拱轴线的PIDL预测解、GA预测解与精确解

Fig.9 PIDL predictive solution, GA prediction solution and exact solution of optimal arch axis

由图9可知,PIDL预测曲线与理论结果高度吻合.图10所示的损失曲线变化表明,迭代25 000次时损失值大幅下降,表明学习率取值合适.当迭代次数为50 000时,损失函数数值约为1×10^{-7},误差在容许范围内.图11为遗传算法适应度变化曲线,群体适应度值随迭代次数上升直至收敛.图12为PIDL预测值的平方误差,平方误差极大值为1.0×10^{-6},同时存在较大波动.图13为遗传算法预测值的平方误差,误差集中在两端点处.最优拱轴线曲面问题精确解为63.764 28,PIDL与GA计算误差对比如表8所示.PIDL预测结果的误差值为0.482%,GA预测结果的误差值为0.965%.这表明PIDL模型具有更高的精度.PIDL计算用时16.82 s,较GA计算效率更高.

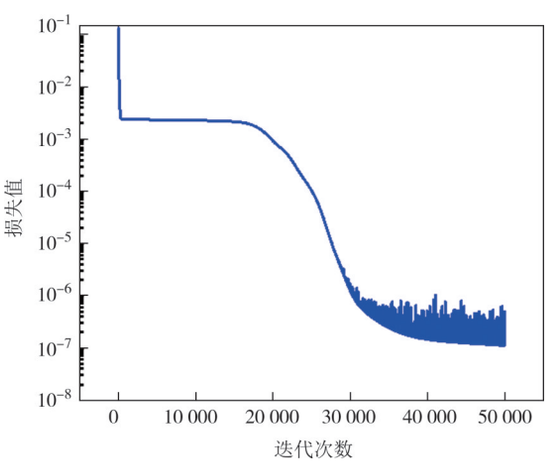


图10 最优拱轴线的PIDL迭代收敛曲线

Fig.10 PIDL iterative convergence curve of optimal arch axis

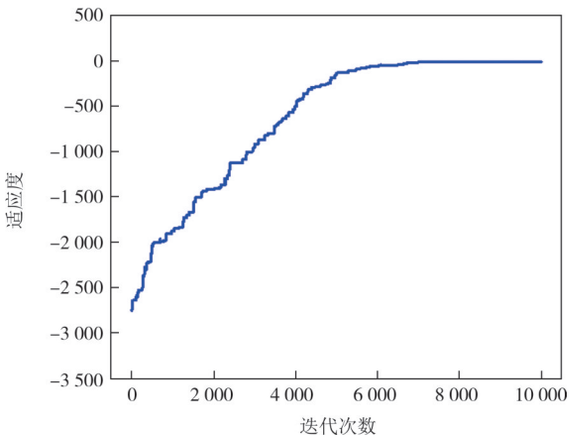


图11 最优拱轴线的GA适应度曲线

Fig.11 GA fitness curve of optimal arch axis

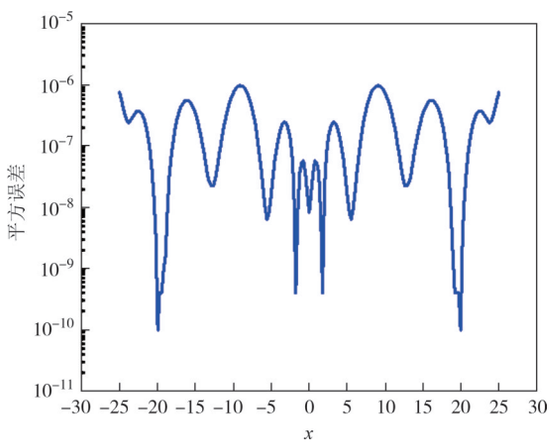


图12 PIDL预测拱轴线的平方误差曲线

Fig.12 Square error curve of the arch axis predicted by PIDL

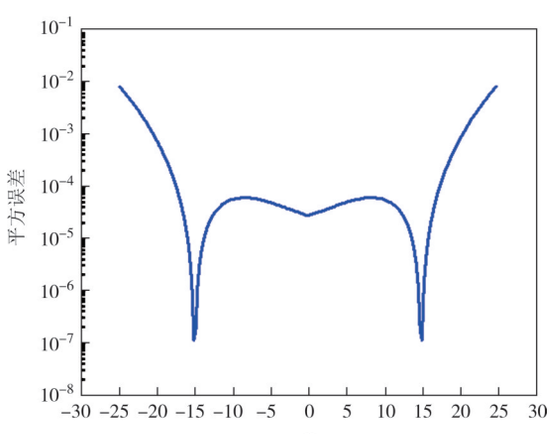


图13 GA预测拱轴线平方误差曲线

Fig.13 Square error curve of the arch axis predicted by GA

表8 最优拱轴线问题计算误差对比

Tab.8 Comparison of computational errors for the optimal arch axis problem

计算方法	拱轴线长度	相对误差绝对值/%
PIDL预测值	63.456 88	0.482
GA预测值	63.149 16	0.965

3 结 论

建立了一种基于物理驱动深度学习的泛函极值求解模型,用于解决形状优化问题.该方法将形状优化问题的物理信息嵌入传统的神经网络模型中.通过性能目标泛函极值构建损失函数,并利用样本空间中的样本点数据进行网络训练,实现泛函极值问题的求解.主要结论如下:

1)建立了泛函极值问题的物理信息深度学习求解方法,并对最优曲面和拱轴线形状问题进行求解,结果与真实值吻合较好,相较于遗传算法,该方法的误差更小且计算效率更高.旋转曲面面积的相对误差绝对值为0.466%,预测拱轴线长度的相对误差绝对值为0.482%,表明该模型具有较高精度.

2)PIDL模型作为一种无网格技术,其优势在于融合数据和物理知识.PIDL无须标签数据即可求解高度非线性非凸的最优化问题,适用于形状优化类的泛函极值问题.该方法对于解决数据收集成本高、数据采集困难的工程问题具有重要意义.未来考虑将其推广至复杂约束的形状优化问题.

3)PIDL在训练过程中不仅能够利用大量的数据样本,还可施加物理信息作为约束.相较于纯数据驱动的机器学习方法,该方法具有较好的泛化性,能够求解不同类型的优化问题.同时,利用边界和初始条件等物理信息来约束模型,避免了过拟合问题,提高了泛化能力.

尽管PIDL相较于遗传算法具有显著优势,但仍存在局限性,如复杂工况下模型训练需要大量计算资源.未来可考虑采用并行计算、自适应学习等技术提高训练效率.为增强PIDL的泛化能力,可结合迁移学习、集成学习等方法,改善模型的鲁棒性和泛化性能.

参考文献

[1] LEVINE W. Optimal control theory: an introduction[J]. IEEE Transactions on Automatic Control, 1972, 17(3): 423.

[2] MAHDY A M S, YOUSSEF E S M. Numerical solution technique for solving isoperimetric variational problems[J]. International Journal of Modern Physics C, 2021, 32(1): 2150002.

[3] RAZZAGHI M. Fourier series direct method for variational problems[J]. International Journal of Control, 1988, 48(3): 887-895.

[4] ZHOU C C, LIU Y. The pade approximant based network for variational problems[J]. arXiv: 2004.00711, 2020.

[5] GORNOV A Y, ZARODNYUK T S, ANIKIN A S, et al. Extension technology and extrema selections in a stochastic multistart algorithm for optimal control problems[J]. Journal of Global Optimization, 2020, 76(3): 533-543.

[6] VOGT T, STREKALOVSKIY E, CREMERS D, et al. Lifting methods for manifold-valued variational problems[M]//Handbook of Variational Methods for Nonlinear Geometric Data. Cham: Springer International Publishing, 2020: 95-119.

[7] PEDREGAL P. A direct algorithm for constrained variational problems in several dimensions[J]. Computers & Mathematics with Applications, 2018, 75(1): 105-121.

[8] YE J X, REY D, KADAKIA N, et al. Systematic variational method for statistical nonlinear state and parameter estimation[J]. Physical Review. E, Statistical, Nonlinear, and Soft Matter Physics, 2015, 92(5): 052901.

[9] MERA N S, ELLIOTT L, INGHAM D B. Numerical solution of a boundary detection problem using genetic algorithms[J]. Engineering Analysis with Boundary Elements, 2004, 28(4): 405-411.

[10] HOSSEINI S E, KARIMI O, ALI ASEMANBAKHSH M. Experimental investigation and multi-objective optimization of savonius wind turbine based on modified non-dominated sorting genetic algorithm-II[J]. Wind Engineering, 2024, 48(3): 446-467.

[11] KORDA V Y, BEREZOVSKY S V, MOLEV A S, et al. Solving variational problems *via* evolutionary algorithm[J]. International Journal of Modern Physics C, 2013, 24(3): 1350009.

[12] HORNIK K, STINCHCOMBE M, WHITE H. Multilayer feedforward networks are universal approximators[J]. Neural Networks, 1989, 2(5): 359-366.

[13] WEINAN E, YU B. The deep ritz method: a deep learning-based numerical algorithm for solving variational problems[J]. Communications in Mathematics and Statistics, 2018, 6(1): 1-12.

[14] ARAZ J Y, CRIADO J C, SPANNOWSKY M. Elvet: a neural network-based differential equation and variational problem solver[J]. arXiv: 2103.14575, 2021.

[15] ABADI M, BARHAM P, CHEN J M, et al. TensorFlow: a system for large-scale machine learning[C]//Proceedings of the 12th USENIX Symposium on Operating Systems Design and Implementation (OSDI' 16). Savannah, GA: USENIX Association, 2016: 265-283.

[16] QIU Y T, ARUNACHALA P K, LINDER C. SenseNet: a physics-

informed deep learning model for shape sensing[J]. Journal of Engineering Mechanics,2023,149(3):04023002.

[17] LIU J Q, CHEN R Q, LOU J H, et al. Deep-learning-based aerodynamic shape optimization of rotor airfoils to suppress dynamic stall[J]. Aerospace Science and Technology,2023,133:108089.

[18] 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]. Journal of Computational Physics, 2019, 378: 686-707.

[19] RAMABATHIRAN A A, RAMACHANDRAN P. SPINN: sparse, physics-based, and partially interpretable neural networks for PDEs[J]. Journal of Computational Physics,2021,445:110600.

[20] LEUNG W T, LIN G, ZHANG Z C. NH-PINN: neural homogenization-based physics-informed neural network for multiscale problems[J]. Journal of Computational Physics, 2022, 470: 111539.

[21] RAISSI M, YAZDANI A, KARNIADAKIS G E. Hidden fluid mechanics: learning velocity and pressure fields from flow visualizations[J]. Science, 2020, 367(6481): 1026-1030.

[22] 张子龙,潘秋景,仉文岗,等. 基于物理信息神经网络的盾构隧道诱发地表沉降预测[J]. 工程力学,2024,41(4):161-173.
ZHANG Z L, PAN Q J, ZHANG W G, et al. Prediction of surface settlements induced by shield tunnelling using physics-informed neural networks[J]. Engineering Mechanics, 2024, 41(4): 161-173.(in Chinese)

[23] 唐和生,何展朋,廖洋洋,等. 基于物理驱动深度迁移学习的薄板力学正反问题[J]. 工程力学, 2023, 40(8): 1-10.
TANG H S, HE Z P, LIAO Y Y, et al. Forward and inverse problems of thin plate mechanics based on physics-informed deep transfer learning[J]. Engineering Mechanics, 2023, 40(8): 1-10.(in Chinese)

[24] BAYDIN A G, PEARLMUTTER B A, RADUL A A, et al. Automatic differentiation in machine learning: a survey[J]. The Journal of Machine Learning Research, 2018, 18(1): 5595-5637.

[25] UDDIN M J, MOHEUDDIN M M, KOWSHER M. A new study of trapezoidal, Simpson's 1/3 and Simpson's 3/8 rules of numerical integral problems[J]. Applied Mathematics and Sciences an International Journal (MathSJ),2019,6(4):1-13.

[26] KINGMA D P, BA J L. Adam: a method for stochastic optimization[C]//ICLR 2015-Conference Track Proceedings, 2015.

[27] HAJI S H, ABDULAZEEZ A M. Comparison of optimization techniques based on gradient descent algorithm: a review[J]. PalArch's Journal of Archaeology of Egypt/Egyptology, 2021, 18(4):2715-2743.

[28] SAMANIEGO E, ANITESCU C, GOSWAMI S, et al. An energy approach to the solution of partial differential equations in computational mechanics *via* machine learning: concepts, implementation and applications[J]. Computer Methods in Applied Mechanics and Engineering,2020,362:112790.