基于神经网络的波动方程行波法求解方法

文◆重庆理工大学机械工程学院　熊星宇

引言

行波法是一种波动方程的常见解法，基本思想是将波动方程转化为沿 X 轴两个方向传播的两个函数。鉴于人工神经网络可以建立输入、输出关系进行函数预测，本文提出一种基于神经网络的求行波函数的方法。利用波动方程的初始位移和速度条件生成数据集，根据行波法展开的格式构造人工神经网络预测模型。训练完成后的神经网络具有和函数一样的功能，在波动数值模拟中可以作为波动方程的数值解。经验证，该方法计算规模小，处理速度快，精度较高。

1 行波法简介

一维波动方程是一个偏微分方程，通常用于描述弦、声波和电磁波等波动现象[1]。首先，一维波动方程是一个线性方程，意味着波动现象可以叠加和分解成多个简单的波动模式。其次，波动方程是一个二阶偏微分方程，包含了时间和空间两个变量，描述了波动的演化过程。最后，它具有解析功能，可以通过数学方法求解出波的行为[2–4]。一维波动方程在物理学、工程学和应用数学等领域中有着广泛应用。在物理学中，它被用于描述弦、声波和电磁波等波动现象的传播和干涉。在工程学中，它被应用于建筑结构的振动分析、声学工程和光学器件设计等方面。在应用数学中，一维波动方程的研究为数学物理学和偏微分方程的发展奠定了重要的基础[5]。

一维波动方程的研究起源可以追溯到17世纪。当时，物理学家如达朗贝尔和欧拉开始研究弦的振动现象，并提出了一维波动方程的初步形式。随着时间的推移，通过实验和理论推导，一维波动方程的表达形式和解析方法已经非常完善。

一维波动方程的形式给出如下。

$$u_{tt}=c^2u_{xx} \tag{1}$$

式（1）中，$u(x,t)$ 表示波在 x 位置和 t 时刻的强度，c 表示波速。达朗贝尔给出了波动方程的一般解，公式如下。

$$u=f(x-ct)+g(x+ct) \tag{2}$$

式（2）中，$f(x,t)$ 和 $g(x,t)$ 为任意两个可微分的单变量函数，分别对应于右传播波和左传播波。通常，结合定解条件可以从式（2）导出特解。假设有如下的无界波动问题。

$$\begin{cases} u_{tt}=c^2u_{xx} \\ u|_{t=0}=\varphi(x), u_t|_{t=0}=\psi(x) \end{cases} \tag{3}$$

从式（2）和式（3）可以给出以下方程。

$$\begin{cases} f(x)+g(x)=\varphi(x) \\ cf'(x)-cg'(x)=\psi(x) \end{cases} \tag{4}$$

根据达朗贝尔公式给出的对应初始位移速度初始条件的波动方程解如下。

$$u=\frac{1}{2}\left[\varphi(x+ct)+\varphi(x-ct)\right]+\frac{1}{2c}\int_{x-ct}^{x+ct}\psi(z)dz \tag{5}$$

式（5）中，$\varphi(x+ct)$ 和 $\varphi(x-ct)$ 分别表示了波速为 c、向 x 轴负方向的左行波、向 x 轴正方向的右行波。积分项表示了初始速度对波动的影响。本文将通过机器学习方法给出两个训练完成的神经网络，用于替代式（2）的两个函数，实现波动方程的求解。

2 神经网络模型

人工神经网络（Artificial Neural Networks，ANNs）是一种模拟生物神经系统的计算模型。自从

【作者简介】熊星宇（1997—），男，重庆人，硕士研究生，研究方向：计算力学。

20 世纪 40 年代提出以来，人工神经网络已经经历了多个演进阶段，如今已在各个领域得到广泛应用[6-7]。人工神经网络具备并行分布式处理能力，可以同时处理多个输入数据并产生相应的输出，在处理大规模数据和复杂问题时具有显著的优势[8]。此外，人工神经网络具备自适应性和学习能力。通过反复调整神经元之间的连接权重，神经网络可以从输入数据中提取出关键特征，并逐渐优化模型的性能，使神经网络能够处理各种复杂的非线性问题，并具备一定的泛化能力[9-10]。

鉴于人工神经网络的以上优点，将函数自变量和函数值组成数据集训练人工神经网络，使用人工神经网络表达各种复杂函数。首先，构造表示式（2）中左行波函数 f 与右行波函数 g 的人工神经网络。其次，以两个神经网络为基础，搭建按照满足式（4）关系的要求组装成新的神经网络模型，搭建以初始条件的范围横坐标为输入，以初始条件为输出的神经网络模型。使用具有多层隐藏层的神经网络，通过堆叠多个隐藏层提取更高级别的特征表示。理论上，针对具有不同初始条件的波动方程算例构造数据集，通过调整神经网络的规模和参数，使神经网络满足各种行波函数，解决各种一维波动问题。

构造两个结构相同但参数不同的神经网络链，分别用于表示函数 f 与 g。该神经网络链采用三层神经网络层结构，输入数组维度与输出数组维度均为 1。每个神经网络层具有 16 个节点，采用双曲正切函数作为激活函数。

基于代表函数 f 与 g 的已构造的神经网络链，应构造表示函数 f' 与 g' 的神经网络。通过微分计算公式定义 f' 与 g'，即定义一个极小值 $\Delta=1E-5$，计算公式如下。

$$\begin{aligned} f' &= \frac{f(x+\Delta)-f(x)}{\Delta} \\ g' &= \frac{g(x+\Delta)-g(x)}{\Delta} \end{aligned} \tag{6}$$

表示函数 $f(x)$ 的神经网络示意图如图 1 所示，在图 1 所示神经网络链前增加函数层对输入添加极小值 Δ，构成新的神经网络链。将代表 $f(x+\Delta)$ 及 $f(x)$ 的神经网络链接到同一输入口，在末尾添加神经网络层做差和除以极小值 Δ，构造的神经网络模型用于表示 f'。函数 g' 同理，

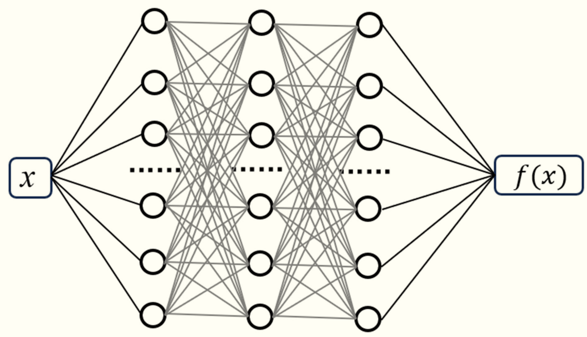


图 1 表示函数 $f(x)$ 的神经网络示意图

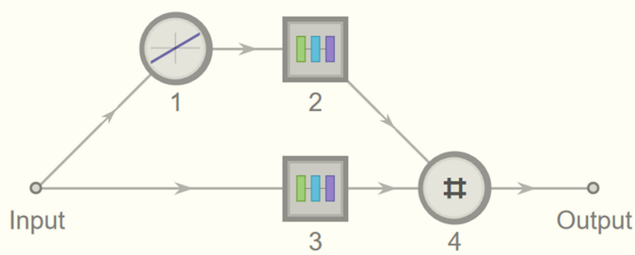


图 2 表示函数 $f'(x)$ 的神经网络

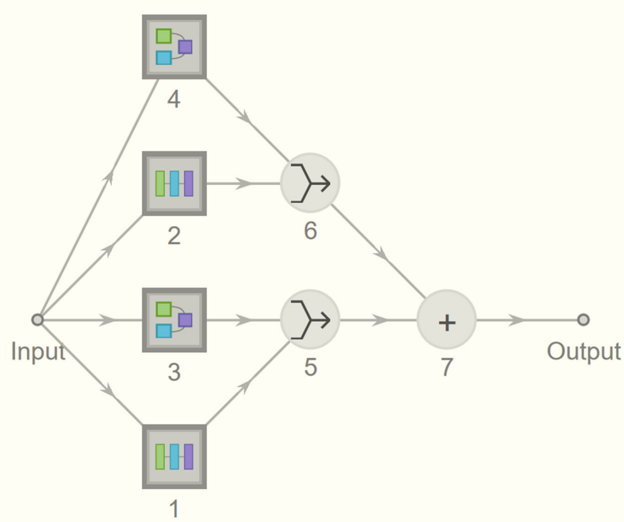


图 3 拟合初始条件的人工神经网络

表示函数 $f'(x)$ 的神经网络如图 2 所示，图中神经网络层 1 表示为输入增加极小值 Δ，层 2、3 表示代表函数 f 的神经网络链，层 4 表示将两个输入做差并除以极小值 Δ 的神经网络层。

在神经网络中，再添加一层无偏置的节点数为 1 的线性层，将该层学习率设置为 0，预设其中的权重表示参数 c。将链和模型组合，以使整个神经网络满足式（4）给出的输入输出关系。将两个神经网络链的输出相加对应输出端的位移初始条件，将图 2 所示的函数 f' 和 g' 的神经网络做差对应速度初始条件。这样一来，通过初始条件的坐标 X 及对应的输出 $\varphi(x)$ 和 $\psi(x)$，就可以训练出与函数 f 与 g 具有近似功能的神经网络。在输出层使用的损失函数采用均方根误差，拟合初始条件的人工神经网络如图 3 所示。图 3 中层 1、2 分别表示代表函数 f 和 g 的神经网络链，层 3、4 分别表示代表函数 f' 和 $-g'$ 的神经网络，层 5、6 表示将输入前后连接到一个向量，层 7 表示将输入相加后输出。

3 数值算例验证

针对式（7）所示的波动方程，使用具有式（8）所示的初始位移和速度的初始条件作为算例

$$u_{xx}=u_{tt} \tag{7}$$

$$\varphi=\begin{cases} e^{-x^2}, |x|\le 10 \\ 0, \quad |x|>10 \end{cases} \qquad (8)$$

$$\psi=0$$

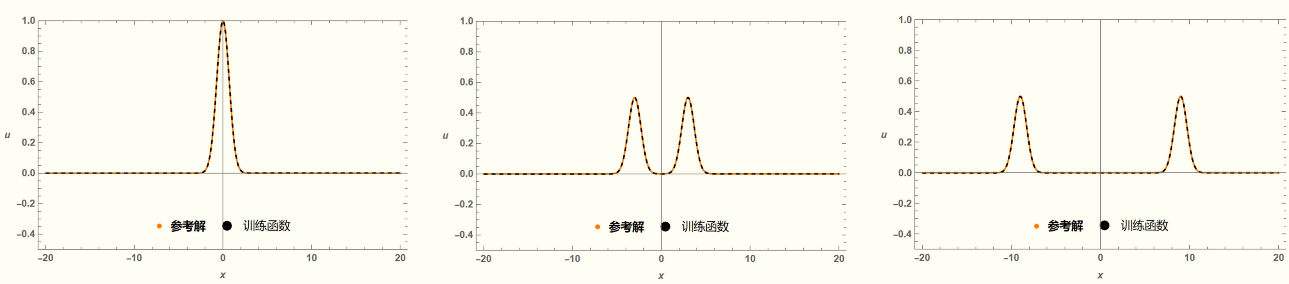
使用式（8）所示的初始条件构造数据集。初始条件函数具有紧支撑的特点，构造数据集应至少包括全部的坐标范围。需要注意的是，若初始条件位移和速度在边界临近处变化较快，应将紧支撑区域外的数值为 0 的部分计算域纳入数据集，否则会导致神经网络链在紧支撑的坐标范围外错误的预测函数值。样本输入是坐标 X，输出分别对应 $\varphi(x)$ 和 $\psi(x)=0$，即输入数据样本长度为 1，输出数据规模样本长度为 2。此过程中取 $\Delta x=0.001$，数据集规模为 $2E+04$，训练最大回合设置为 2000，学习率设置为 0.001，在数据集中抽取 30% 作为训练中的验证集，若 Loss 在 20 回合内没有明显下降则停止训练。最终，得到了表示初始条件的整个神经网络模型以及两个训练完成的神经网络链，它们分别代表函数 f 与 g。整体神经网络模型训练完成后，得到初始条件的对比结果（见图 4）。

无需知道具体的神经网络链内部参数，将其组装为式（2）所示的格式，用来展示算例波动随时间的变化情况。同时，使用有限差分法给出了式（5）所代表的参考解。算例 1 各时刻波形对比如图 4.b 和图 4.c 所示，橙色虚线表示差分方法计算的参考解，黑色虚线表示人工神经网络得出的函数解。在 $t=0$ 到 $t=9$ 的计算时间内，波形吻合情况良好。

此外，在上述算例基础上添加了初始速度，如式（9）所示。

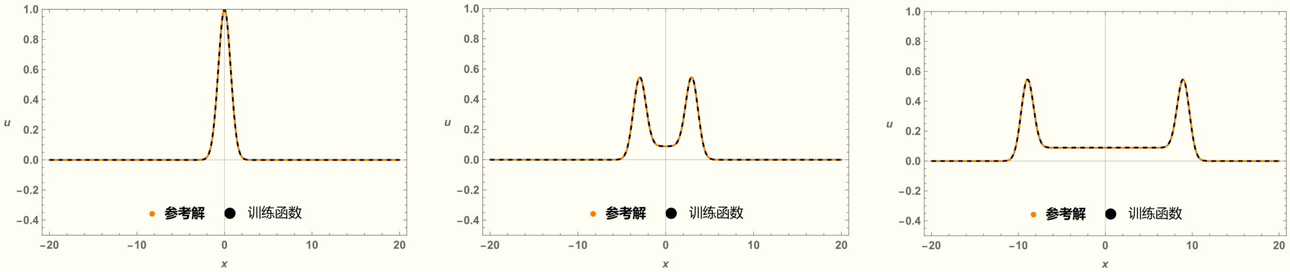
$$\varphi=\begin{cases} e^{-x^2}\ \square|x|\le \\ 0, \quad |x|>10 \end{cases}$$
$$\psi=\begin{cases} 0.1e^{-x^2}, |x|\le 10 \\ 0, \qquad |x|>10 \end{cases} \qquad (9)$$

为取得更好的训练效果，适当增加了图 1 所示神经网络链中层数量增加至 5 个，每一层的节点数量增加至 64 个，数据集的坐标范围取值提高为 [–20,20]。其余设置相同的情况下，完成神经网络训练。使用差分方法计算出了波动方程在 $t\in[0,10]$ 和 $x\in[-20,20]$ 范围内的解。训练的初始速度对比如图 5 所示，给出了神经网络方法计算出的初始速度与实际初始条件的对比，算例 2 各时刻波形对比如图 6 所示，给出了各时刻通过两种方法计算出的波形，其中 $t=0$ 展示了神经网络使用初始条件数据集的训练效果。观察到训练的神经网络模型无法很好的拟合初始速度函数曲线，即所求的左行波与右



a.t=0　　b.t=3　　c.t=9

图 4 算例 1 各时刻波形对比



a.t=0　b.t=3　c.t=9

图 6 算例 2 各时刻波形对比

行波函数的零时刻时间一阶偏导拟合程度不佳。但是在图 6 所示的波形对比图中，可以看出波形在计算时刻内拟合较好。

结语

参照波动方程的行波法求解办法，为求出分解为左行波与右行波形式的波动方程，本文提出了一种基于神经网络的求解行波函数方法。以深度学习的神经网络链表示未知的左行波和右行波函数，按照满足初始条件格式搭建完整的神经网络。使用初始条件构造数据集，经数据集训练后的神经网络可以等效行波函数预测出结果，给出对应时间和位置的波动数值，即给出波动方程的解。经数值算例验证，使用此方

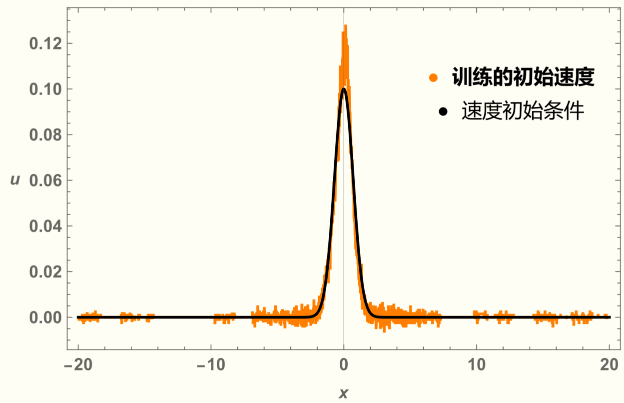


图 5 训练的初始速度对比

法解出的波形与差分方法的解在各时刻均吻合较好。该方法在完成神经网络模型建立后，可以快速解析具有各种初始条件的波动方程。

引用

[1] Joseph B.Keller,李家春.线性波传播理论的发展和前景[J].力学进展,1980(4):101-118.
[2] 廖振鹏.近场波动问题的有限元解法[J].地震工程与工程振动,1984,4(2):1-14.
[3] C.埃蒙,吕仲林.波动方程与模型[J].石油地球物理勘探,1979(5):48-71.
[4] 阎超,于剑,徐晶磊,等.CFD模拟方法的发展成就与展望[J].力学进展,2011,41(5):562-589.
[5] 李红敬,文祯中.自然科学概论[M].南京:南京大学出版社,2019.
[6] 焦李成,杨淑媛,刘芳,等.神经网络七十年:回顾与展望[J].计算机学报,2016,39(8):1697-1716.
[7] 刘建伟,刘媛,罗雄麟.深度学习研究进展[J].计算机应用研究,2014,31(7):1921-1930+1942.
[8] 马秀麟,邬彤,鲍建樟,等.信息处理与数据科学[M].北京:北京师范大学出版社,2021.
[9] 自动化技术、计算机技术[J].中国无线电电子学文摘,2011,27(6):163-242.
[10] 丁丁,刘文哲,盛常冲,等.神经网络架构搜索研究进展与展望[J].国防科技大学学报,2023,45(6):100-131.