

文章编号: 1000-8055(2004) 04-0525-05

多宗量导热反问题求解的神经网络法

王秀春, 智会强, 毛 之, 杨增军, 韩 鹏
(河北工业大学 能源与环境工程学院, 天津 300132)

摘要: 将神经网络用于求解多宗量导热反问题, 建立了相应的数学模型, 并给出了具体的网络设计方法。数值模拟结果表明, 此法是可行的、有效的, 并具有较高的精度和较强的通用性, 可推广用于其它反问题的求解。

关 键 词: 航空、航天推进系统; 多宗量; 导热反问题; BP 神经网络; 测试集

中图分类号: TP393. 09; TK124

文献标识码: A

Neural Network Model for Solving Multi-Variable Inverse Heat Conduction Problem

WANG Xiu-chun, ZHI Hui-qiang, MAO Yi-zhi,
YANG Zeng-jun, HAN Peng

(School of Energy and Environment Engineering,
Hebei University of Technology, Tianjin 300132, China)

Abstract: Neural network was applied to solving multi-variable inverse heat conduction problem, and a valid numerical model was presented. The processing method of the inputting data and outputting data and the determining method of the number of inner layers were discussed. By the numerical simulation, the change of the training error and the testing error accompanying the variation of the number of inner layers is developed. The results indicate that the method is practicable and has high precision. In addition, in order to create a neural network which has high level of generalization, test sets for trial and error must be used to monitor the training process. The extension of the method in this paper to other inverse heat conduction problem is straightforward.

Key words: aerospace propulsion system; multi-variables; inverse heat conduction problem; BP neural network; test sets

导热反问题(Inverse Heat Conduction Problem, IHCP)是指通过研究对象内部或边界一个点或多个点的温度值或其随时间的变化来反算边界条件、初始条件、热物性参数、内热源强度或物体的几何条件等未知项。导热反问题广泛存在于航空航天、冶金、机械、核反应堆、地热能勘探、生

物传热等工程领域, 它的研究在工程实践中有着重要作用^[1]。

从数学上讲, 导热反问题是一种不适定问题, 不像适定问题那样在解的存在性、唯一性、稳定性方面有较好的性质, 因此, 求解反问题要比正问题困难的多。目前, 对多维问题研究较少, 解决方法

收稿日期: 2003- 08- 25; 修订日期: 2004- 01- 02

作者简介: 王秀春(1949-), 男, 天津人, 河北工业大学能源与环境工程学院教授, 主要从事工业过程中强化传热, 电机、变压器发热冷却研究、计算传热和导热反问题研究

也不完善,目前的工作多侧重于单宗量的反演,但在工程实际中对热流、温度、换热系数等多宗量的综合反演又比较常见^[2]。已有的反问题求解方法主要有:明确函数法、有限差分法、正则化方法、共轭梯度法^[3,4]等,这些方法主要针对特定的导热反问题,适用范围较窄,收敛性也难于得到保证,对工程应用不利。

本文将人工神经网络(Artificial neural Network, ANN)用于导热反问题的研究,建立了求解多宗量导热反问题的数学模型。数值模拟结果表明,方法可行,并具有相当的通用性。

1 BP 网络概述

近年来,人工神经网络理论得到了飞速发展,已在各个工程领域得到广泛的应用^[5]。人工神经网络是模仿生物神经网络的人工智能系统,用处理单元间的可变权连接模拟生物神经元间的突触行为,各处理单元之间以一定的形式连接成具有特定拓扑结构的网络,主要特点是:分布存储、并行处理,并具有很好的容错性、自组织性和自学习性^[5]。

本文采用BP网络进行求解。BP网络主要用于函数逼近、模式识别、分类、数据压缩等领域。BP网络的隐层传递函数常采用Sigmoid型函数,输出层采用Sigmoid型函数或纯线性函数。

BP网络的学习算法为BP算法,是由Rumelhart、Hinton和Williams于1986年提出的,是一种有监督学习算法^[5],主要由两部分组成:信息的正向传播和误差的反向传播。在正向传播过程中,输入信息从输入经隐层逐层计算传向输出层,每一层神经元的输出作用于下一层神经元的输入,如果在输出层没有的得到期望的输出,则计算输出层的误差变化值,然后转向反向传播,通过网络将误差信号沿原来的连接通路反传回来,以修改各层神经元的权值和域值,如此循环直到达到期望目标。

2 BP 网络应用于导热反问题的理论基础及方法

2.1 理论基础

我们之所以将神经网络用于求解导热反问题,是以下定理为基础的:

定理^[6]: 令 $\Phi(\cdot)$ 为有界非常量的单调增连续函数, I_p 代表 p 维超立方体 $[0, 1]^p$, $C(I_p)$ 表示

定义在 I_p 上的连续函数构成的集合,则给定任意函数 $f(\cdot) \in C(I_p)$ 和 $\epsilon > 0$,存在正整数 M 和一组实常数 α, θ 和 ω ,其中 $i = 1, \dots, M; j = 1, \dots, p$,使得网络输出

$$F(x_1, \dots, x_p) = \sum_{i=1}^M \alpha \Phi \left(\sum_{j=1}^p \omega_j x_j - \theta_i \right)$$

可逼近任意 $f(\cdot)$,即:

$$|F(x_1, \dots, x_p) - f(x_1, \dots, x_p)| < \epsilon$$

上述结果说明,只含有一个隐层的前馈网络是一个通用的函数逼近器^[7],所以,我们可以用单隐层网络来逼近任意函数。由此可以得出结论:可以用神经网络来学习导热反问题所描述的函数关系。

2.2 BP 网络求解导热反问题的方法

由于导热反问题是相对于正问题而言的,而在实际应用中建立正问题的模型是比较容易的。同时正问题的求解方法也比较成熟,因此我们可以利用求解导热正问题的方法来产生BP网络的训练样本,即对研究对象先施加不同的边界条件,通过正问题求解,得到各个边界条件下的温度场。然后以这些温度场的观测点为BP网络的输入,边界条件为目标输出来训练BP网络,训练好后,就可以把研究对象观测点的温度值输入到网络,输出即为边界条件。当然在工程实际中,训练样本亦可以通过实测得到。

3 工程算例

为了验证上述算法的可行性、有效性,本文选择了一个较为复杂的算例和一个相应正问题具有解析解的算例,第一个算例同时求解第一类、第二类、第三类边界条件,第二个算例同时求解第一类、第三类边界条件。本文所选算例是设计型反问题^[8],即求解能达到已知温度场所施加的边界条件,这类反问题在工程中有着广泛的应用,比如在冷冻储藏和冶金铸造中,为了达到预期的凝固或溶解温度,就要求解所需施加的边界条件。

3.1 算例一

3.1.1 数学描述

设有一L型截面的物体,如图1所示,导热系数 λ 为 $35.5 \text{ W}/(\text{m} \cdot \text{K})$,其内有均匀内热源 Φ ,大小为 $1800 \text{ W}/\text{m}^3$,已知其第一、四边界为第二类边界条件,待求量为 q_1, q_2 ,第二、三边界为第一类边界条件,待求量为 T_0 ,第五边界为绝热边界,第六边界为第三类边界条件,待求量为 h_0 ,流

体的温度 T_f 为 20°C 。现由于某种工程需要, 要求 L 型截面物体内的温度场达到某预定值, 求解应该在第一、二、三、四、六边界施加怎样的边界条件值, 才能达到此要求。

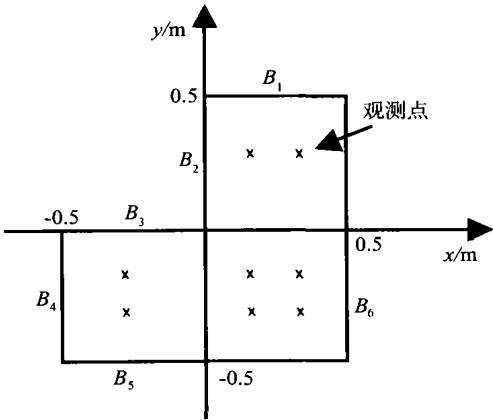


图 1 物理模型
Fig. 1 Physical model

控制方程如下:

$$\lambda \frac{\partial^2 T}{\partial x^2} + \lambda \frac{\partial^2 T}{\partial y^2} + \Phi = 0 \tag{1}$$

$$\lambda \frac{\partial T}{\partial y} \Big|_{(x,y)} = q_1 \quad (x,y) \in B_1 \tag{2}$$

$$T(x,y) = T_0 \quad (x,y) \in B_2 \tag{3}$$

$$T(x,y) = T_0 \quad (x,y) \in B_3 \tag{4}$$

$$\lambda \frac{\partial T}{\partial x} \Big|_{(x,y)} = q_2 \quad (x,y) \in B_4 \tag{5}$$

$$\lambda \frac{\partial T}{\partial y} \Big|_{(x,y)} = 0 \quad (x,y) \in B_5 \tag{6}$$

$$\lambda \frac{\partial T}{\partial x} \Big|_{(x,y)} = h_0(T_w - T_f) \quad (x,y) \in B_6 \tag{7}$$

$$T(x_i, y_i) = T_i \quad (x_i, y_i) \in \Omega \tag{8}$$

其中 q_1, q_2, T_0, h_0 为待求量, 因为在工程中各个待求量的范围一般能够根据实际条件确定, 因此在这里我们假定 $q_1 \in [20, 50], q_2 \in [30, 60], T_0 \in [50, 80], h_0 \in [40, 70]$, 式 (8) 是反问题的补充条件, 即观测点的温度值。

3. 1. 2 BP 网络的设计

3. 1. 2. 1 输入和输出的确定及数据处理

本例的输入即为观测点的温度值, 输入层的神经元个数等于观测点的数目, 经实验确定, 在本例所给的边界条件范围内, 如果有 8 个神经元即

可完成对此反问题的求解; 输出层的神经元个数很容易确定, 即为反演的未知量个数, 本例为 4 个。

为了使各个输入变量一开始便处于同等重要的地位, 同时, 为了消除各输出值对输出误差的影响差异, 需要对样本的输出和输入进行归一化处理^[9]。

3. 1. 2. 2 各层传递函数和隐层单元数的确定

在本例中, 隐层传递函数采用 tansig 函数(正切 sigmoid 函数), 输出层采用纯线性函数。

对 BP 网络而言, 隐层神经元数量的确定是成败的关键, 若隐层神经元数量太少, 则网络所能获取的用以解决问题的信息太少, 若数量太多, 不仅增加训练时间, 而且还会出现“过渡吻合(over-fitting)”问题^[8], 从而导致网络的推广能力急剧下降, 推广能力低下的网络在实际中是没什么应用价值的。鉴于以上因素, 我们采取了以下方法: 首先把样本集分为训练集和测试集, 然后用训练集对具有不同隐节点数的网络进行训练, 在训练的同时不断的用测试集对网络进行测试, 当发现测试误差开始增加时(即此时网络的推广能力开始下降), 便停止训练。同时程序返回具有最小测试误差时的网络, 最后从得到的一系列网络中选取测试误差最小的网络, 此网络即为具有最佳隐层结点数的网络, 同时, 也具有最好的推广能力。

3. 1. 2. 3 对 BP 算法的改进

由于传统的 BP 算法的训练速度较慢, 而且易陷入局部最小, 因此, 我们采用了一种基于数值优化方法的 BP 算法, 即 Levenberg-Marguardt 法^[10], 此法的权值和域值调整公式为:

$$X^{(k+1)} = X^{(k)} + S(X^{(k)})$$

$$S(X^{(k)}) = - (H^{(k)} + \lambda^{(k)} I) \nabla f(X^{(k)})$$

其中: $X^{(k)}$ 为第 k 时层的权值或域值; $H^{(k)}$ 为第 k 时层的海森矩阵; $f(X^{(k)})$ 为目标函数; I 为单位矩阵(和 $H^{(k)}$ 同维); $\lambda^{(k)}$ 为一正数, 程序开始运行时取很大值, 并逐渐减少至零。

此方法的优点是收敛速度极快, 而且稳定性也比较好。

3. 1. 3 仿真步骤

(1) 对不同的边界条件组合, 求解导热正问题, 得到训练样本集, 并把样本集分为训练集和测试集;

(2) 先把网络的隐单元数设为 2, 利用 (1) 中

的数据进行训练,并记录其训练误差和测试误差,然后逐步增加隐节点数,对各个网络进行训练,记录各网络的训练误差和测试误差,最后观察训练误差和测试误差随隐单元数的变化情况;

(3) 从(2)中所得到的数据中选出最好的网络,并用 20 组不在训练集和测试集之内的样本进行测试,分析其结果。

3.1.4 仿真结果

表 1 是网络的训练误差和测试误差随隐节点数的变化情况,从中可以看出,增加隐节点数可以减少训练误差,但测试误差随隐节点数的增多先减小后增加,这说明网络的推广能力随隐节点数的增加先增强后减弱,主要原因是在增加隐节点数目的同时出现了过度吻合现象。因此,在训练网络时,不能刻意的减小训练误差,这样做会损害网络的推广能力,这一点应在网络设计时应引起高度重视。

表 1 误差随隐节点数的变化情况

Table 1 Variety of error with the number of hidden units

隐单元个数	最大训练误差	最大测试误差
2	2.956 6	3.157 6
3	1.765 9	2.054 1
4	0.137 9	0.317 5
5	0.081 3	0.113 8
6	0.022 5	0.025 2
7	0.021 1	0.031 9
8	0.020 3	0.052 7
12	0.019 7	1.548 6

从表 1 中我们选出了测试误差最小的网络,即隐节点数为 6 的网络,并用 20 组样本进行了测试,图 2 是此网络的训练曲线,可见训练误差和测试误差的曲线形状是相似的,当训练到第 335 步时达到了目标误差,此时,测试误差略高于训练误差。图 3 是对 20 组样本的仿真结果,共有 4 条曲线,每条曲线代表一个所求量,最大绝对误差为 0.025,其中温度 T_0 几乎达到了零误差,可见结果是比较精确的。

3.2 算例二

由于篇幅所限,本文对此算例只作简略概述。此问题可简要描述为:为使一矩形区域内达到某

预定温度场,求其右边界和下边界应施加怎样的边界条件(右边界为第三类边界条件,所求量为换热系数,下边界为第一类边界条件,所求量为温度,其它两个边界绝热)。此算例所对应的正问题具有解析解^[11],所以可以得到很准确的训练样本。求解方法和第一个算例相同,我们在这里只给出求解结果的最大绝对误差:0.002 1(换热系数),0.007 9(温度),可见结果是相当准确的。

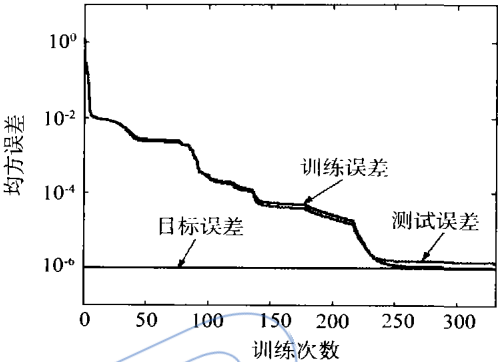


图 2 训练曲线

Fig. 2 Training curve

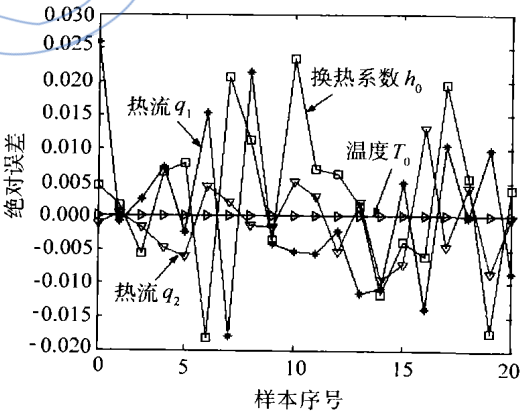


图 3 测试误差图

Fig. 3 Testing error

从理论上来说,具有单隐层的 BP 网络可以逼近任何函数,因此,只要所求的反问题能够得到足够多的训练样本,就可以用 BP 网络来求解。

应当指出,导热反问题有时会在存在不唯一的解,但在工程应用中,常常只需要满足条件的一个解即可。因此,解的不唯一性并不影响本算法的工程应用。另外,由于 BP 网络是通过训练样本来学习反问题所蕴涵的函数关系的,因此要求训练样本要能真实反映相应的函数关系,不应存在过大的误差。

4 结 论

通过数值模拟可以看出, 本文所提出的多宗量导热反问题的反演模型是可行的、有效的, 并具有较高的精度; 本方法具有较强的通用性, 只要所求问题有足够的样本即可。另外, 在网络的训练过程中, 一定要用测试集随时进行监测, 以期得到具有较高推广能力的网络, 只有这样的网络才具有实用意义。

参考文献:

[1] 俞炳丰, 王志刚, 蒋立军, 等. 基于导热反问题分析的推力轴承瞬态油膜温度监控方法[J]. 西安交通大学学报, 2001; 35 (3): 225– 269.

[2] 胡国俊. 传热中的多宗量反演研究[D]. 大连: 大连理工大学, 2002.

[3] Beck J V, Blackwell B, Clair C R St. Inverse Heat Conduction-III Posed Problem[M]. *New York: Wilay*, 1985, 1– 40.

[4] Huang C H, Ozsik M N, Sawaf B. Conjugate Gradient Method for Determining Unknown Contact Conductance During Metal Casting[J]. *Int J. heat Mass Transfer*, 1991, 34(7): 1779 – 1785.

[5] 王文成. 神经网络及其在汽车工程中的应用[M]. 北京: 北京理工大学出版社, 1998: 61– 89.

[6] 阎平凡, 张长水. 人工神经网络与模拟进化计算[M]. 北京: 清华大学出版社, 2000, 13– 37.

[7] Hornik K, Stinchcombe M, White H. Multilayer Feedforward Network are Universal Approximators[J]. *Neural Networks*, 1989, 2: 359– 366.

[8] 黄光远, 李小军. 数学物理反问题[M]. 山东: 山东科学技术出版社, 1986: 27– 50.

[9] 韩力群. 人工神经网络理论、设计及应用[M]. 北京: 化学工业出版社, 2002, 18– 69.

[10] 丛爽. 神经网络、模糊系统及其在运动控制中的应用[M]. 合肥: 中国科学技术大学出版社, 2001, 20– 30.

[11] 奥齐西克 M N. 热传导[M]. 俞昌铭译. 北京: 高等教育出版社, 1983, 58– 72.