

文章编号 1000-5013(2001) 04-0397-04

神经网络近似结构分析的最佳样本集

张云波

(华侨大学土木工程系, 泉州 362011)

摘要 基于多层神经网络映射存在定理, 建立近似结构分析的人工神经网络模型. 构造不同样本集进行计算机数值试验, 分析不同样本集学习后模型的模拟精度. 算例分析表明, 当采用适当的样本集训练神经网络时, 得出的可靠性指标可满足精度要求同时训练时间又最少.

关键词 神经网络, 结构分析, 最佳样本集

中图分类号 TP 183: TU 3

文献标识码 A

近年来结构的近似分析取得一定的进展^[1~3], 人工神经网络技术的应用是其中之一. 由于神经网络是一种近似的模拟方法, 其精度如何直接影响到结构近似分析结果的精度. 本文选取不同样本集经过大量模拟试验, 提出最佳样本集的选取方法. 算例表明, 由最佳样本集训练的模型直接进行可靠性指标计算, 其精度满足要求同时训练时间又最少.

1 人工神经网络模型

人工神经网络是由大量神经元相互连接而成的网络. 根据连接方式, 神经网络通常分成两大类, 即没有反馈的前向神经网络和相互结合型网络. 前向神经网络是由输入层、一层或多层的隐含层和输出层组成, 每一层的神经元只接受前一层神经元的输出. 而相互结合型神经网络中的任意两个神经元之间, 都可能存在连接. 因此, 输入信号要在神经元之间反复传递. 它从某一初始状态开始, 经过若干次的变化, 渐渐趋于某一稳定状态或者进入周期振荡等其它状态.

BP 神经网络是一单向传播的多层前向神经网络. 网络除输入输出结点外, 有一层或多层的隐含层结点, 同层结点间无任何联结. 输入数据 X_1, X_2, \dots, X_n 从输入层结点, 依次经过各隐含层结点, 然后到达输出层结点, 从而得到输出数据 Y_1, Y_2, \dots, Y_m . 由于同层结点间无任何藕合, 故每层结点的输出只影响下一层结点的输出. 每个结点表示单个神经元, 其对应的传递函数常选 Sigmoid 型函数. 有时, 输出层中的结点之传递函数为线性函数.

Kolmogorov 研究指出, 对于任意一个连续函数 $\Phi: I^n \rightarrow R^m, Y = \Phi(X)$, (I 是单位闭区间 $[0, 1]$), 都可用一个三层神经网络精确地实现. 并且, 此网络的输入层有 n 个神经元, 隐层有

$2n+1$ 个神经元, 输出层有 m 个神经元的三层神经网络结构. Funahashi 进一步对每层神经元函数的选取进行研究后指出, 在三层网络中, 只要对隐层采用非线性递增映射函数, 输入输出层采用线性映射函数, 就可以用该网络对任意连续函数 $Y=\Phi(X)$ 进行逼近.

基于 Kolmogorov 多层神经网络映射存在定理, 可以用多层神经网络, 建立输入变量 X 和输出变量 Z 之间的映射关系. 即通过一定数量的样本点(图1)对网络进行训练, 当误差达到要求后, 所训练的神经网络模型可实现对函数的仿真.

2 样本对模型精度的影响

由样本集训练后的神经网络模型与真实函数的拟合精度取决于学习样本. 作者通过大量的数值试验表明, 当有足够的样本点(图1), 且均匀地分布在解空间时, 所拟合的函数与真实函数具有足够的精度. 为了研究样本对模型精度的影响, 本文选取一个球面函数为

$$Z = \frac{22\,500 - (X - 100)^2 - (Y - 100)^2}{2} \quad (0 \leq X \leq 100, 0 \leq Y \leq 100). \quad (1)$$

构造一个输入层为2个结点, 隐层有5个结点, 输出层为1个结点 BP 神经网络结构. 神经元映射函数为 Sigmoid 函数 $f(x) = \frac{1}{1 + e^{-x}}$ 的三层神经网络, 用以拟合该球面函数. 采用自适应学习速率, 结合附加动量法的反向传播网络进行训练. 对变量 X, Y 在其取值范围内, 分别按 $n(n=2 \sim 8)$ 等分分割. 正交化后的数据点, 代入式(1)得到 Z . 以此为试验样本点, 样本点的数量 $L = m^{n+1}$ (其中 m 为变量个数, n 为等分数). 如 $n=2$ 时的试验样本点为 $X = [0, 0, 0, 50, 50, 50, 100, 100, 100]$ 和 $Y = [0, 50, 100, 0, 50, 100, 0, 50, 100]$, 对应 $Z = [50, 100, 111.8, 100, 132.29, 141.42, 111.80, 141.42, 150]$.

为了研究训练后模型的精度, 选取 $X = [15, 42.5, 57.5, 85], Y = [15, 42.5, 57.5, 85]$. 正交化后, 产生16个样本点作为测试评价所训练模型精度的输入样本. 表1给出测试评价样本

表1 不同样本集预测值

X	Y	Z	Z_2	Z_2/Z	Z_3	Z_3/Z	Z_4	Z_4/Z	Z_5	Z_5/Z
15.0	15.0	89.722	82.749	0.922	87.768	0.978	89.210	0.994	88.588	0.987
15.0	42.5	109.402	104.552	0.956	110.748	1.012	110.578	1.011	110.039	1.006
15.0	57.5	116.055	113.741	0.980	118.232	1.0119	117.447	1.012	116.930	1.008
15.0	85.0	122.678	126.892	1.034	125.550	1.023	124.738	1.017	124.429	1.014
42.5	15.0	109.402	116.679	1.067	117.244	1.072	112.979	1.033	111.370	1.018
42.5	42.5	126.046	132.880	1.054	132.253	1.049	129.982	1.031	129.641	1.029
42.5	57.5	131.862	136.950	1.039	134.998	1.024	133.706	1.014	133.655	1.014
42.5	85.0	137.727	141.130	1.025	136.532	0.991	136.686	0.992	136.662	0.992
57.5	15.0	116.055	126.720	1.092	125.708	1.083	121.487	1.0467	119.553	1.030
57.5	42.5	131.862	138.826	1.053	137.812	1.045	136.322	1.034	136.195	1.033
57.5	57.5	137.432	141.051	1.026	139.493	1.015	139.026	1.012	139.379	1.014
57.5	85.0	143.069	142.999	1.000	139.890	0.978	140.679	0.983	141.135	0.987

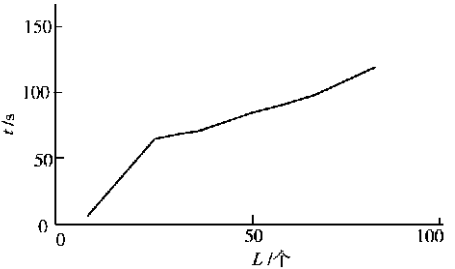


图1 样本点数量与训练时间的关系

续表

X	Y	Z	Z_2	Z_2/Z	Z_3	Z_3/Z	Z_4	Z_4/Z	Z_5	Z_5/Z
85.0	15.0	122.678	130.764	1.066	130.106	1.061	126.855	1.034	124.257	1.013
85.0	42.5	137.727	141.715	1.029	141.227	1.025	140.777	1.022	140.581	1.021
85.0	57.5	143.069	143.045	1.000	142.562	0.997	142.944	0.099	143.554	1.003
85.0	85.0	148.492	143.784	0.968	142.097	0.963	144.064	0.970	145.087	0.977
最大值				1.092		1.083		1.047		1.033
最小值				0.922		0.963		0.970		0.977
标准差				0.046		0.035		0.021		0.016

的精确值 Z 和预测值 Z_n . 由于篇幅的关系, 表1仅列出 $n=2$ 到5的模型预测值以及预测值 Z_n 与精确值 Z 的比值 Z_n/Z . 图1为样本点数量与训练时间的关系. 从表1可以看出, 当 $n=2$ 时, 模型误差在10%以内; $n=4$ 时, 模型误差在5%以内; 而 $n=5$ 时, 模型误差在4%以内. 由此可见, 随着训练样本集的样本点越多, 模型就越精确, 但训练时间就越长.

3 算例

图2所示的为门式平面框架计算简图. 各单元弹性模量均为 $E=2\times10^6\text{ kPa}$, 单元截面惯性矩与截面面积的关系为 $I_i=\alpha A_i^2 (i=1, 2)$. 随机变量取单元截面积 A_1, A_2 以及外荷载 P . 其统计特性见表2. 若最大变形控制取结点3

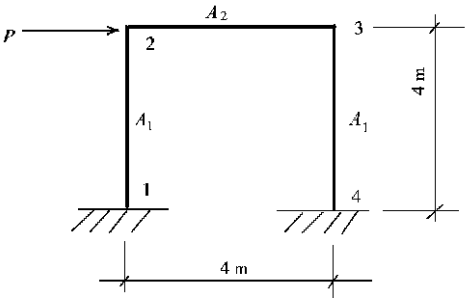


图2 门式框架计算简图

的水平控制位移 u , 试建立模拟极限状态方程的神经网络模型并求其可靠性指标. 其极限状态方程为 $Z=G(A, P, u)=10-u^3(A_1, A_2, P)=0$, 其中 $u=10\text{ mm}$ 为允许最大位移.

构造一个输入层为3个结点, 隐层为7个结点, 输出层为1个结点 BP 神经网络. 按上述方法产生样本集训练. 当误差达到要求后, 所训练的神经网络即可模拟该极限状态方程. 然后, 用传统的结构可靠性分

表2 随机变量统计特性

随机变量	μ_i	σ_i	分布类型	α
A_1/m^2	0.36	0.036	对数正态	0.083 33
A_2/m^2	0.18	0.018	对数正态	0.166 67
P/kN	20	5.0	极值 型	-

析方法计算其可靠性指标 β (表3). 为了便于比较, 表3同时给出文 [6, 4] 的可靠性指标 β 及相应验算点的计算结果. 从表3可以看出, 当采用 $n=2$ 产生的样本点时, 其模型即可满足精度要求. 同时, 网络训练时间最短, 此样本集称之为最佳样本集.

表3 各种方法验算点和可靠性指标 β

验算点	文 [4]	文 [6]	本文解			
			$n=2$	$n=3$	$n=4$	$n=5$
A_1	0.315 7	0.311 464	0.318	0.316 15	0.315 49	0.313 22
A_2	0.169 7	0.164 658	0.170	0.167 96	0.167 85	0.169 15
P	37.387	37.257 54	37.832	36.825	37.048	36.993
β	2.840 5	2.807 134	2.841 9	2.811 8	2.840 4	2.852 8

4 结束语

本文基于神经网络的结构近似分析,讨论样本点的选取问题.通过大量的数值试验表明,基于神经网络技术结构近似分析功能函数的拟合精度,取决于样本点的分布和数量.当有足够的样本点时,所拟合的结构功能函数相当精确.但是,计算量却大大增加,学习效率随之降低.此外,对样本点的取值范围应根据所求解问题来决定,对结构优化设计来说,其取值范围应根据设计变量的可行域来取值.而对结构可靠性分析来说,为了保证设计验算点落在此范围内,根据工程中的 3σ 原则,当给定 X 的取值范围为 $[\mu_x - 3\sigma_x, \mu_x + 3\sigma_x]$ 时, X 取值在其范围的概率为99.7%.因此,对于每一个设计变量均在 $[\mu_x - 3\sigma_x, \mu_x + 3\sigma_x]$ 范围内取值.

参 考 文 献

- 1 Faravelli L. Response-surface approach for reliability analysis[J]. Journal of Engineering Mechanics ASCE, 1989, 115(12): 2 763 ~ 2 781
- 2 陆金桂,余俊,王浩等.基于人工神经网络的结构近似分析方法的研究[J].中国科学(A辑),1994, (6): 653 ~ 658
- 3 李向京,胡云昌.基于遗传算法-神经网络混合训练技术的结构近似可靠分析方法[J].土木工程学报, 2000, (5): 40 ~ 45
- 4 赵国藩.工程结构可靠性理论与应用[M].大连:大连理工大学出版社,1996.1 ~ 350

A Study on Optimal Sample Set Based on Approximate Structural Analysis of Neural Network

Zhang Yunbo

(Dept. of Civil Eng., Huaqiao Univ., 362011, Quanzhou)

Abstract Based on the theorem of the existence of multilayered neural network mapping, a model of artificial neural network is set up for approximate structural analysis. Different sample sets are constructed for computer-ized numerical experimentation, and the accuracy of simulating post-learning model of different sample sets is analysed. As indicated by analysing examples of computation, when proper sample set is adopted to train neural network, a reliability index satisfying the need of accuracy can be obtained and the train time is minimal.

Keywords neural network, structural analysis, optimal sample set