

# 基于神经网络的测量模型的建立及检验

朱坚民<sup>1</sup>, 王中宇<sup>2</sup>, 吕延庆<sup>2</sup>, 周福章<sup>2</sup>, 宾鸿赞<sup>1</sup>

(1. 华中理工大学机械科学与工程学院, 湖北 武汉 430074;

2. 洛阳工学院机械与电子工程系, 河南 洛阳 471039)

**摘要:** 针对间接测量数学模型难于建立的问题, 提出了基于径向基函数神经网络的非参数测量模型的建立方法。利用对训练样本的聚类结果来确定基函数的中心, 使得神经网络在较少的训练样本条件下仍可获得较高的建模精度。应用实例验证了本方法的有效性。

**关键词:** 测量模型; 径向基函数; 神经网络; 聚类

**中图分类号:** TP391.41 **文献标识码:** A

## 1 引言

许多复杂工业对象重要输出变量的测量有时因无相应传感器或测量滞后较大等因素而无法实时检测, 但较好反映该量的其他变量却能直接检测到, 它们之间的关系可以用一个多元函数(线性的或非线性的)来描述, 这个映射函数称之为测量模型, 故测量模型的建立是对这些输出变量在线监测与控制的先决条件。

建立测量模型可以从两个方面入手, 一是分析工业过程的机理得到机理模型; 另一种是通过反映工业过程运行的数据直接建立模型。由于机理方程推导和运算的复杂性, 通常采用第二种方法。目前, 研究实验数据相互关系, 建立模型的方法主要是回归法<sup>[1]</sup>。在具有测量模型为近似线性等先验知识的条件下, 用线性回归方法可以得到较高的建模精度。但绝大多数被研究对象普遍具有非线性, 在这种情况下, 往往通过变量代换的方法把非线性回归转化成线性回归, 或者将非线性关系表为(或展成)幂级数(多项式), 再通过变量代换转化为多元线性回归问题。这种方法的主要局限性有: 1. 必须预先确定回归模型的结构; 2. 自变量的选择无法在线增减; 3. 在对象非线性较强时建模精度不高。

神经网络由于具有很强的非线性映射能力, 使其在多维非线性函数逼近、多维非线性系统建

模等领域得到了成功的应用, 但也存在着一定的局限性, 主要是神经网络训练时要求有较多数量的训练样本, 且对样本的分布具有一定的要求。若训练样本较少或样本分布不合理时, 建模精度不高。

针对以上问题, 本文提出了一种基于径向基函数(Radial Basis Function)神经网络(简称RBF网络)的非参数测量模型的建立方法。同时, 为了保证在测量数据样本较少时同样可以获得较高的建模精度, 训练RBF网络时采用k-均值聚类算法来确定网络隐层节点基函数的中心, 通过实例建模与检验, 取得了理想的结果。

## 2 神经网络模型的选择

在非线性建模领域应用得较为成功的神经网络模型主要有: 多层前向BP(Back Propagation)网络(简称BP网络), 径向基函数(RBF)网络等。用神经网络进行建模时, 它能够有效地处理测量过程的非线性和滞后, 同时它不需要很多工业过程的先验知识, 而是根据输入输出数据直接建模。BP网络理论上可以以任意精度逼近任意非线性映射, 但在实际应用中存在以下问题: 1. 存在局部最优问题; 2. 网络初始权值及学习速率的选取对网络的收敛速度及稳定性有极大的影响; 3. 网络隐层节点个数的确定需要人的介入, 难以在算法中确定下来。以上问题的存在, 使BP网络的建模

精度、建模速度及所建模型的通用性大打折扣。RBF 网络模型的出现,使 BP 网络存在的问题得到了较好的解决,它不但可以获得全局最优点,有较高的逼近精度,而且迭代次数少,算法具有很好的通用性<sup>[2]</sup>。

### 3 径向基函数(RBF)网络

#### 3.1 RBF 网络的基本原理及其算法

RBF 神经网络是一种两层前传网络,其网络结构如图 1 所示,隐层执行的是一种固定不变的非线性变换,将输入空间映射到一个新的空间,输出层在该新的空间中实现线性组合,从而实现  $f: X^n \rightarrow Y^m$  的非线性映射,可调节的参数就是该线性组合器的权。

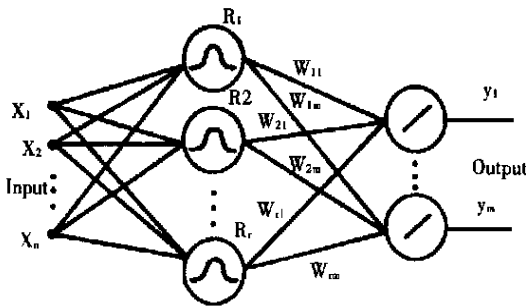


Fig. 1 Topological structure of RBF

图 1 中,隐层各单元的输出为

$$R_i = \mathcal{Q}(X - c_i) \quad i = 1, \dots, r \quad (1)$$

其中,  $\mathcal{Q}(\cdot)$  为网络隐层节点高斯基函数,取  $\mathcal{Q}(v) = \exp(-v^2 \cdot \ln 2)$ ,  $R_i$  是隐层第  $i$  个单元的输出,  $X$  是输入模式,  $c_i$  是隐层第  $i$  个单元基函数的中心,  $r$  为隐层节点数。设  $w_{ij}$  为第  $i$  个隐层单元到第  $j$  个输出层单元的连接权值,则网络的输出可用式(2)来表达:

$$y_j(X) = \sum_{i=1}^r w_{ij} R_i \quad j = 1, \dots, m \quad (2)$$

#### 3.2 RBF 网络的学习训练算法

训练神经网络的目的在于寻找合适的权值,使式(3)表示的输出误差平方和最小。

$$\min \|\hat{Y} - \hat{W}R\|^2 \quad (3)$$

式中,  $\|\cdot\|$  为欧氏范数,  $\hat{Y}$  为期望输出矩阵,  $\hat{W}$  为权值矩阵,  $R = [R_1, R_2, \dots, R_r]$  为隐层输出矩阵。则网络权值的调整可转化为式(4)所示的最小二乘问题(式中  $\epsilon$  为噪声干扰阵)。

$$\hat{Y} = \hat{W}R + \epsilon \quad (4)$$

当隐层单元基函数的中心确定之后,隐层单元的输出可由式(1)求出,这样网络的连接权就可通过

求解线性方程组来确定。

本文的应用实例中讨论的是双输入单输出的情形,此时 RBF 网络要逼近的映射关系如式(5)所示。

$$f(X) = w_0 + \sum_{i=1}^r w_i \mathcal{Q}(X - c_i) \quad (5)$$

学习就是根据训练样本确定隐层节点基函数的中心  $c_i (i = 1, \dots, r)$  及权值  $w_i (i = 0, \dots, r)$  的过程。

#### 3.3 确定 RBF 网络高斯基中心的 $k$ -均值聚类算法

事实上,对 RBF 网络性能有重要影响的关键因素就是基函数中心的选取。在实践中,可以将中心选为数据的某个子集,当然这种子集的选取应当是对输入适当抽样的。如果从数据点中任意选取中心,这样构造出来的网络性能一般是不能令人满意的,特别是在数据样本较少且样本分布不合理时,中心可能靠得太近产生近似线性相关,从而带来数值上的病态条件,使网络泛化能力变差。针对这种情况,为改善网络的泛化能力,对已知的训练样本点先进行聚类,使  $c_i$  尽可能均匀地对输入数据抽样,再将各聚类的中心作为隐层节点径向基函数的中心。聚类算法采用  $k$ -均值聚类算法,其中,类的数目  $k$  采用 RPCL(Rival Penalized Competitive Learning)方法确定<sup>[3]</sup>,这是一种无监督学习方法,不仅简单,而且性能良好。

$k$ -均值聚类算法的基本步骤是<sup>[4]</sup>:

(1) 先选中  $k$  个样本点作为初始聚类点。为合理选择这  $k$  个聚类点,对所有  $n$  个样本点作以欧氏距离  $d$  为半径的“超球”,计算各点之间的距离。把与某点的距离小于  $d$  的点的个数作为该点的“密度”,并取  $n$  个样本中“密度”最大的前  $k$  个点作为初始聚类点。

(2) 对于剩下的  $n - k$  个样本点,分别计算每点到各聚类点的距离,并将该点归入最近的凝聚点所代表的类,然后计算各类的重心(均值),作为新的凝聚点。

(3) 不断重复步骤(2),直到下列两个条件之一满足时停止迭代。

(a) 前后两次聚类结果相同。

(b) 前后两次凝聚点位置基本上重合。

### 4 非参数测量模型建立的应用实例

为比较多元线性回归和 RBF 神经网络建模的精度,对测量模型是线性和非线性的情况,根据

已有的测量数据用上述三种方法建立了各自的模型,并用实际测量数据对所建的模型进行了检验。

#### 4.1 已知测量模型为线性的情况

采用文献[1]中第127页例5-12的29组测

量数据,由该文献知某变量 $Y$ 受变量 $X_1$ 、 $X_2$ 的影响,通过试验获得表1中的一批数据,且根据经验 $Y$ 与 $X_1$ 、 $X_2$ 成线性关系。

Table 1 Measurement data

order	$X_1$	$X_2$	$Y$	order	$X_1$	$X_2$	$Y$
1	15.58	1.95	1.34	16	17.88	2.52	2.41
2	10.68	1.37	1.27	17	13.38	1.43	1.69
3	15.62	2.39	1.56	18	14.21	2.27	1.59
4	15.78	1.14	1.48	19	16.80	1.41	1.19
5	13.22	1.85	1.40	20	16.38	1.78	2.44
6	16.44	1.32	1.82	21	10.81	1.32	1.35
7	11.40	2.05	0.85	22	17.26	1.31	1.57
8	16.17	1.11	1.40	23	14.92	1.42	1.64
9	14.03	1.47	1.15	24	18.14	2.13	1.64
10	15.67	1.38	1.89	25	18.15	1.20	2.34
11	12.74	1.35	0.87	26	10.31	0.98	0.65
12	11.73	1.33	1.53	27	11.40	1.27	1.19
13	14.84	1.09	1.25	28	12.57	0.87	2.06
14	13.73	1.27	2.47	29	17.61	1.21	1.57
15	15.12	1.78	1.83				

取表1中的前20组数据分别用线性回归<sup>[1]</sup>、RBF网络基函数中心任选、RBF网络基函数中心由聚类算法确定等方法分别建立了测量模型。其中,采用最后一种方法建模时,聚类中心分别为: $c_1 = [16.0550 \ 1.5600]$ 、 $c_2 = [13.1891 \ 1.5691]$ 、 $c_3 = [17.8800 \ 2.5200]$ ,聚为3类,权值分别为: $w_0$

$= 1.1612$ 、 $w_1 = 0.5568$ 、 $w_2 = 0.4490$ 、 $w_3 = 1.1934$ 。当隐层节点基函数中心从训练样本中选取时,训练后隐层节点的个数为19个,由于篇幅所限,具体中心值及网络权值不再给出。为了检验各种方法的建模精度,利用表1中第21~29号九组数据来对各自模型进行检验,其结果如表2所示。

Table 2 Comparing results of linear regression model and RBF model

Order $r$	$X_1$	$X_2$	$Y$	Linear regression		RBF model		RBF model	
				model		(without clustering)		(after clustering)	
				Prediction	error	Prediction	error	Prediction	error
1	10.81	1.32	1.35	1.12	0.23	1.47	-0.12	1.17	0.18
2	17.26	1.31	1.57	1.83	-0.26	-0.87	2.44	1.69	-0.12
3	14.92	1.42	1.64	1.59	0.05	1.62	0.02	1.44	0.20
4	18.14	2.13	1.64	2.03	-0.39	1.24	0.40	2.21	-0.57
5	18.15	1.20	2.34	1.91	0.43	-1.67	4.01	1.52	0.82
6	10.31	0.98	0.65	1.02	-0.37	0.63	0.02	1.16	-0.51
7	11.40	1.27	1.19	1.18	0.01	1.81	-0.62	1.21	-0.02
8	12.57	0.87	2.06	1.26	0.80	0.69	1.37	1.41	0.65
9	17.61	1.21	1.57	1.85	-0.28	-1.91	3.48	1.60	-0.03
				error sum of squares:		error sum of squares:		Error sum of squares:	
				1.315		36.581		1.768	

从表2中可看出,由于 $Y$ 与 $X_1$ 、 $X_2$ 存在近似线性关系,故线性回归模型获得了较理想的预测效果。对于RBF神经网络模型来说,由于训练样本太少,且样本分布不甚合理,神经网络模型没有得到很好的训练,使得在有些预测样本的预测误差较小而另外一些预测样本的预测误差很大,故

此时神经网络模型的预测精度与线性回归模型相比较低。但将RBF网络基函数的中心由训练样本聚类确定后建立的神经网络模型的预测精度与未聚类相比有了较大程度的改善,尽管训练样本数量较小,其预测精度已与线性回归模型的预测精度相差无几。

## 4.2 无测量模型先验知识的情形

为了便于与上面的应用实例进行比较, 我们仍采用对被测量有两个主要影响变量的模型, 设

计实验获得 49 组数据如表 3 所示。与表 1 数据不同的是, 无有关被测量  $Y$  与变量  $X_1, X_2$  关系的先验知识。

Table 3 Measurement data

order	$X_1$	$X_2$	$Y$	order	$X_1$	$X_2$	$Y$
1	16.76	1.69	1.64	26	12.17	1.22	1.85
2	11.62	1.94	1.91	27	17.00	2.05	1.84
3	16.70	0.83	0.90	28	11.09	0.82	0.75
4	15.45	1.44	2.60	29	17.15	1.13	1.09
5	16.65	1.65	1.70	30	12.38	1.92	2.49
6	15.67	1.52	2.49	31	12.27	1.59	2.22
7	12.43	1.12	1.93	32	10.51	2.48	2.02
8	11.54	1.95	1.81	33	14.66	1.51	3.27
9	12.42	1.72	2.40	34	14.12	1.36	3.38
10	11.20	1.98	1.71	35	13.46	1.18	2.91
11	13.02	2.26	3.27	36	14.63	2.09	3.65
12	16.82	1.80	1.68	37	14.23	1.88	3.72
13	13.97	2.32	3.79	38	11.67	1.44	1.60
14	16.57	1.89	1.85	39	16.26	1.95	2.09
15	16.54	1.92	1.91	40	13.68	1.76	3.59
16	12.73	1.29	2.49	41	16.35	0.90	1.32
17	12.72	1.70	2.71	42	14.82	0.88	2.63
18	15.81	1.32	2.17	43	13.32	1.31	2.94
19	16.70	1.76	1.74	44	16.99	0.82	0.81
20	12.96	1.99	3.10	45	16.14	2.45	2.42
21	14.37	1.55	3.50	46	17.92	2.14	1.99
22	15.55	1.85	2.78	47	13.50	1.64	3.30
23	16.35	2.42	2.35	48	11.71	1.89	1.74
24	14.18	2.29	3.84	49	12.56	2.43	2.87
25	11.38	2.46	2.01				

与上例相似, 取表 3 中前 40 组数据分别用线性回归、RBF 网络基函数中心任选和 RBF 网络基函数中心由聚类算法确定分别建立测量模型,

$$\begin{aligned}
 c_1 &= [13.6800 \ 1.7600], & c_2 &= [11.5075 \ 1.8275], \\
 c_5 &= [10.9450 \ 2.4700], & c_6 &= [12.5067 \ 1.7800], \\
 c_9 &= [14.0750 \ 2.3050], & c_{10} &= [16.2600 \ 1.9500], \\
 c_{13} &= [13.4600 \ 1.1800], & c_{14} &= [12.2700 \ 1.5900], \\
 c_{17} &= [17.0000 \ 2.0500], & c_{18} &= [12.7300 \ 1.2900], \\
 c_{21} &= [14.2300 \ 1.8800], & c_{22} &= [14.6600 \ 1.5100], \\
 c_{25} &= [16.9250 \ 0.9800].
 \end{aligned}$$

权值分别为:

$$\begin{aligned}
 w_0 &= 1.8863, & w_1 &= -2.7360, & w_2 &= -1.2683, \\
 w_6 &= -14.5859, & w_7 &= -0.9462, & w_8 &= -6.6717, \\
 w_{12} &= 5.5168, & w_{13} &= -2.2601, & w_{14} &= 14.2266, \\
 w_{18} &= 5.5459, & w_{19} &= 101.3145, & w_{20} &= -2.0467, \\
 w_{24} &= -232.4115, & w_{25} &= -8.9003
 \end{aligned}$$

当隐层节点基函数中心从训练样本中选取时, 训练后隐层节点的个数为 26 个, 其具体中心值及网络权值不再给出。为了检验各种方法的建

其中, 采用最后一种方法建模时, 聚类个数为 25, 25 个聚类中心如下:

$$\begin{aligned}
 c_3 &= [14.6300 \ 2.0900], & c_4 &= [14.2450 \ 1.4550], \\
 c_7 &= [11.0900 \ 0.8200], & c_8 &= [12.3000 \ 1.1700], \\
 c_{11} &= [15.5000 \ 1.6450], & c_{12} &= [12.9900 \ 2.1250], \\
 c_{15} &= [15.7400 \ 1.4200], & c_{16} &= [16.7600 \ 1.6900], \\
 c_{19} &= [16.5550 \ 1.9050], & c_{20} &= [16.3500 \ 2.4200], \\
 c_{23} &= [16.6750 \ 1.7050], & c_{24} &= [16.8200 \ 1.8000],
 \end{aligned}$$

$$\begin{aligned}
 w_3 &= 2.7503, & w_4 &= 9.9766, & w_5 &= 0.5736, \\
 w_9 &= 0.7132, & w_{10} &= -23.1489, & w_{11} &= 7.0071, \\
 w_{15} &= -1.5724, & w_{16} &= 360.3788, & w_{17} &= 30.9134, \\
 w_{21} &= -2.4698, & w_{22} &= -8.8644, & w_{23} &= -229.7519,
 \end{aligned}$$

模精度, 利用表 3 中第 41~49 号九组数据来对各自模型进行检验, 其结果如表 4 所示。

Table 4 Comparing results of linear regression model and RBF model

Order	$X_1$	$X_2$	$Y$	Linear regression model		RBF model (without clustering)		RBF model (after clustering)	
				Prediction	error	prediction	error	Prediction	error
1	16.35	0.90	1.32	1.73	-0.41	1.19	0.13	1.26	0.06
2	14.82	0.88	2.63	1.75	0.88	2.08	0.55	2.49	0.14
3	13.32	1.31	2.94	2.09	0.85	2.96	-0.02	2.95	-0.01
4	16.99	0.82	0.81	1.66	-0.85	0.79	0.02	0.74	0.07
5	16.14	2.45	2.42	2.84	-0.42	2.40	0.02	2.41	0.01
6	17.92	2.14	1.99	2.59	-0.60	1.07	0.92	2.08	-0.09
7	13.50	1.64	3.30	2.32	0.98	3.38	-0.08	3.38	-0.08
8	11.71	1.89	1.74	2.54	-0.80	1.91	-0.17	1.92	-0.18
9	12.56	2.43	2.87	2.90	-0.03	2.59	0.28	2.70	0.17
				error sum of squares: 4.525		error sum of squares: 1.281		Error sum of squares: 0.104	

从表 4 中可看出, 由于此时被测量  $Y$  与影响量  $X_1, X_2$  不成线性关系, 采用线性回归模型得不到较好的预测效果。但随着神经网络训练样本的增加, 神经网络模型的建模精度得到了较大的提高, 特别是基函数中心由样本聚类确定的 RBF 网络模型经过 40 个样本点的训练就获得了理想的预测结果。

## 5 结 论

本文提出了基于径向基函数神经网络的非

### 参考文献:

- [1] 费业泰. 误差理论与数据处理[M]. 北京: 机械工业出版社, 1987.
- [2] 徐秉铮, 等. 神经网络理论与应用[M]. 广州: 华南理工大学出版社, 1994.
- [3] Krzyzak Xu A, Oja E. Rival penalized competitive learning for clustering analysis, RBF Net, and Curve Detection [J]. IEEE Transactions on Neural Networks, 1993, 4(4): 636-649.
- [4] 王碧泉, 陈祖萌. 模式识别理论、方法和应用[M]. 北京: 地震出版社, 1989.

## Establishment and verification of measurement model based on neural networks

ZHU Jian-min<sup>1</sup>, WANG Zhong-yu<sup>2</sup>, LU Yan-qing<sup>2</sup>, ZHOU Fu-zhang<sup>2</sup>, BIN Hong-zan<sup>1</sup>

- (1. College of Mechanical Science & Engineering, Huazhong University of Science & Technology, Wuhan, 430074, China;
2. Department of Mechanical & Electrical Engineering, Luoyang Institute of Technology, Luoyang, 471039, China)

**Abstract:** Aiming at the difficulty of the model establishment of indirect measurement, a new method of measurement model establishment based on radial basis function neural networks is presented. In order to obtain a high modeling precision when the size of training sample is small, the center of radial basis function is decided by clustering results of the training sample, then, the effectiveness of this method is proved by application examples.

**Key words:** measurement model; radial basis function; neural networks; clustering

作者简介: 朱坚民(1968-), 男, 江苏如东人, 讲师, 现为华中理工大学机械科学与工程学院现代制造技术研究所博士研究生。主要研究领域为工业生产过程测量与控制, 发表论文 10 余篇。

参数测量模型的建立方法, 该建模方法不需要对象的先验知识, 不管对象模型是线性或非线性均根据测量数据直接建模, 回避了传统回归建模方法需事先确定回归模型结构的难点。网络基函数的中心由输入样本用  $K$ -均值聚类算法确定后, 使得测量模型在小样本训练的情况下仍具有较强的泛化能力及较高的建模精度, 尤其适用于那些机理复杂、影响因素众多且具有严重非线性的测量模型的建立。