

河道洪水演算的径向基函数神经网络模型

黄国如 ,芮孝芳

(河海大学水资源环境学院 ,江苏 南京 210098)

摘要 :将径向基函数神经网络方法应用于河道洪水演算中 ,并利用最小二乘法来确定模型参数 .结合河道洪水演算的具体方式 ,分别构建基于马斯京根方法和具有预见期的洪水演算方法的径向基函数神经网络模型 .将该模型应用于两条天然河道的洪水演算中 ,计算结果表明 ,该模型运算快速 ,精度较高 ,具有较大的应用价值 .

关键词 :洪水演算 ;马斯京根方法 ;有预见期法 ;径向基函数 ;人工神经网络

中图分类号 :TV122 文献标识码 :A 文章编号 :1000-198X(2003)06-0621-05

河道洪水演算以马斯京根方法为代表 ,其计算过程比较简便 ,精度较高 ,在实际洪水演算中应用得最为广泛 .文献 [1 , 2] 利用运动波数值扩散模拟扩散波物理扩散的概念 ,导出了具有预见期的洪水演算方法 .这两类洪水演算方法均需要首先确定其演算系数 ,但有时难以确定 ,尤其当河道洪水波运动为非线性时 ,洪水演算系数的率定就更加困难 .近年来发展起来的人工神经网络技术具有自学习、自适应和容错性等优点 ,在水文预报中已得到了广泛的应用 ,其中 BP 神经网络应用得较为广泛 [3 ~ 6] .但是 ,BP 网络的学习一般基于梯度法等非线性优化方法 ,学习过程中往往会使其优化准则函数陷入局部极小 .尽管可以采用遗传算法、模拟退火法等来避免局部最小 ,但一般会增加大量的工作量 .而径向基函数人工神经网络方法(Radial Basis Function , RBF - ANN)可以有效地解决以上问题 ,它为多层前向神经网络的学习提供了一种新颖而有效的方法 [7] .本文利用径向基函数人工神经网络方法 ,以马斯京根方法和具有预见期的洪水演算方法为基础 ,建立基于径向基函数神经网络的河道洪水演算模型 .

1 径向基函数神经网络模型及其算法

1.1 RBF-ANN 的结构

RBF-ANN 网络的结构如图 1 所示 ,它通常是一种 3 层前向式网络 ,由输入层、隐层和输出层组成 .输入层有 n 个单元 $X \in R^n$, $X = (x_1 , x_2 , \dots , x_n)^T$;隐层有 d 个单元 $H \in R^d$, $H = (h_1 , h_2 , \dots , h_d)^T$;输出层有 m 个单元 $Y \in R^m$, $Y = (y_1 , y_2 , \dots , y_m)^T$.隐层单元基函数的中心为 $C_j = (c_{1j} , c_{2j} , \dots , c_{nj})$ ($j = 1 , 2 , \dots , d$) ,宽度为 σ_j ($j = 1 , 2 , \dots , d$) ;隐层单元与输出层单元之间的连接权重为 w_{jk} ($j = 1 , 2 , \dots , d ; k = 1 , 2 , \dots , m$) .下标 i , j 和 k 分别表示输入单元、隐层单元和输出单元 , n , d 和 m 分别为输入单元、隐层单元和输出单元的数量 . n 和 m 同实际问题的输入维数和输出维数一致 ,而隐层单元 d 的选择则是一个较为复杂的问题 ,直到现在还没有适宜的方法来确定 .

1.2 径向基函数

Powell 于 1985 年提出了多变量插值的径向基函数方法 ,Broomhead 等人成功地将该方法应用于模式识别 [7] .对于图 1 所示的 RBF-ANN 结构 ,假设样本集中有 P 个学习样本 ,其输入模式为 $X^p = (x_1^p , x_2^p , \dots , x_n^p)^T$,则有如下结果 [7 ~ 10] :

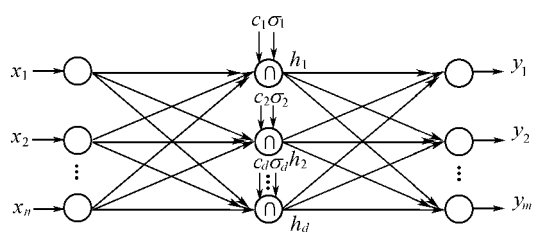


图 1 径向基函数人工神经网络结构
Fig.1 Radial basis function artificial neural network architecture

隐层单元的输出为

$$h_j^p = R_j^p / \sum_{j=1}^d R_j^p \quad (p = 1, 2, \dots, P; j = 1, 2, \dots, d) \quad (1)$$

$$R_j^p = \exp \left[-0.5 \sum_{i=1}^n \left(\frac{x_i^p - c_{ij}}{\sigma_j} \right)^2 \right] \quad (2)$$

式中: R_j^p ——隐层径向基函数,它与数理统计中的高斯正态分布函数形式相同,具有以 $X = C_j$ 为中心的径向对称性质,这也正是径向基函数这一名称的由来; h_j^p ——归一化的隐层径向基函数.

输出单元的输出为

$$y_k^p = \sum_{j=1}^d w_{jk} h_j^p \quad (p = 1, 2, \dots, P; k = 1, 2, \dots, m) \quad (3)$$

从式(1)和式(3)可以看出,在 RBF-ANN 网络中,隐层执行的是一种非线性变换,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,可以调节的是该线性组合的权重.由于输出单元的输出值为隐层输出值,其隐层到输出层之间连接权重的线性组合只有一个参数可供优化,因此,可用最小二乘方法来确定这些参数,这样就不会存在局部极小问题,这正是 RBF 和 BP 神经网络的根本区别所在.

一般认为,RBF 网络所利用的非线性激励函数形式对网络性能的影响并非至关重要,关键因素是基函数中心的选取^[7].

1.3 RBF-ANN 网络训练算法

在式(1)和式(3)中,需要通过网络学习得到的参数为隐层基函数的中心 C_j 、宽度 σ_j 以及隐层到输出层的权重 w_{jk} .一般可以根据以下 3 个步骤分别进行.

a. 隐层基函数中心 C_j

在多变量插值的 RBF 方法中,中心 C_j 的确定通常对应于所有的输入模式 X^p ,这样有多少对输入数据就有多少个隐单元.但是,由于输入数据可能具有类聚性,此时会有许多隐层单元其中心相距很近,导致许多冗余单元,使得不必要地加长了训练时间,而且容易造成“过度拟合”,降低网络的推广能力.为此通常先采用聚类分析技术,对输入数据进行预处理,找出有代表性的数据点(这些数据点不一定位于原始数据点)作为隐层基函数的中心,从而极大地减少隐层的单元数目,降低了网络的复杂性^[10].

RBF 网络学习常用的聚类分析技术是 K-均值算法,其具体过程如下:

第一步,通常将样本集中 P 个学习样本的前 d 个样本分别赋值给隐层基函数中心 C_j ,即 $c_{ij} = x_i^p$ ($i = 1, 2, \dots, n; j = 1, 2, \dots, d; p = j$),使得基函数中心 C_j 初始化.所以隐层单元的数目不能大于样本数目,即 $h \leq P$.

第二步,将所有学习样本 X^p 按照最近的基函数中心分组.对于每个基函数中心 C_j 以及所有学习样本 X^p ,计算每个学习样本 X^p 与每个基函数中心 C_j 之间的差值 $\sum_{i=1}^n (x_i^p - c_{ij})^2$,并且统计出每个学习样本 X^p 在所有基函数中心 C_j 内的最小值 $\min_{j=1}^d \sum_{i=1}^n (x_i^p - c_{ij})^2$,并记满足 $\min_{j=1}^d \sum_{i=1}^n (x_i^p - c_{ij})^2$ 时的样本 X^p 属于 C_j 的子样本集 θ_j .

第三步,重新计算 c_{ij} . $c_{ij} = \frac{1}{M_j} \sum_{p \in \theta_j} x_i^p$, M_j 为子样本集 θ_j 中的样本模式数目.

第四步,若基函数中心不再变化,则停止训练,其基函数中心 C_j 的稳定值即为所求;否则转向第 2 步继续计算直至稳定为止.

b. 隐层基函数宽度 σ_j

在确定基函数中心 C_j 后,可以求得基函数的宽度 σ_j ,它可以采用很多方法确定.这里令它等于基函数中心与子样本集 θ_j 中样本模式之间的平均距离,即

$$\sigma_j = \sqrt{\frac{1}{M_j} \sum_{p \in \theta_j} \sum_{i=1}^n (x_i^p - c_{ij})^2} \quad (4)$$

c. 隐层到输出层权重 w_{jk}

由式(3)可看出,在基函数的参数 C_j 和 σ_j 确定后,输出层的输出仅是权重 w_{jk} 的函数,故此可利用最小二

乘法确定权重系数^[9].假设对第 p 个样本而言,输入层为 X^p ,输出层为 $Y^p=(y_1^p,y_2^p,\dots,y_m^p)$,期望输出为 $T^p=(t_1^p,t_2^p,\dots,t_m^p)$,则取平方误差最小为优化准则函数,即

$$\min E=\frac{1}{2}\sum_{p=1}^P\sum_{k=1}^m(t_k^p-y_k^p)^2$$

(5)

式中 E 为误差,式(5)可写为

$$\min E=\frac{1}{2}\sum_{p=1}^P\sum_{k=1}^m\left(t_k^p-\sum_{j=1}^d w_{jk}h_j^p\right)^2$$

(6)

于是 $\frac{\partial E}{\partial w_{jk}}=-\sum_{p=1}^P\left(t_k^p-\sum_{j=1}^d w_{jk}h_j^p\right)h_j^p=0\quad(j=1,2,\dots,d;k=1,2,\dots,m)$

为了以示区别,将上式中两个 h_j^p 之中的一个改写为 h_r^p ,即为

$$\sum_{j=1}^d\left(\sum_{p=1}^P h_r^p\cdot h_j^p\right)w_{jk}=\sum_{p=1}^P h_r^p\cdot t_k^p\quad(r=1,2,\dots,d;k=1,2,\dots,m)$$

令 $a_{rj}=\sum_{p=1}^P h_r^p\cdot h_j^p\quad c_{rk}=\sum_{p=1}^P h_r^p\cdot t_k^p$

则有 $\sum_{j=1}^d a_{rj}w_{jk}=c_{rk}$

写成矩阵形式为

$$AW=C$$

(7)

$$A=(a_{rj})=\begin{bmatrix} a_{11} & \cdots & a_{1d} \\ \vdots & \ddots & \vdots \\ a_{d1} & \cdots & a_{dd} \end{bmatrix}\quad W=(w_{jk})=\begin{bmatrix} w_{11} & \cdots & w_{1m} \\ \vdots & \ddots & \vdots \\ w_{d1} & \cdots & w_{dm} \end{bmatrix}\quad C=(c_{rk})=\begin{bmatrix} c_{11} & \cdots & c_{1m} \\ \vdots & \ddots & \vdots \\ c_{d1} & \cdots & c_{dm} \end{bmatrix}$$

实际上,式(7)是由 d 个 m 元未知数组成的线性代数方程组,并且系数矩阵 $A=(a_{rj})$ 为对称矩阵,用解对称阵的常用方法就可以得到 $d\times m$ 个权重系数 w_{jk} .

2 实例研究

2.1 径向基函数神经网络模型结构

在实际洪水演算中,马斯京根方法是使用最为广泛的方法,另外,近年来芮孝芳等^[1,2]提出了具有预见期的洪水演算方法,马斯京根洪水演算方法为^[1,2]

$$Q_{\text{断}}(t+1)=C_1Q_{\text{断}}(t)+C_2Q_{\text{断}}(t+1)+C_3Q_{\text{断}}(t)$$

(8)

式中: $Q_{\text{断}}(t),Q_{\text{断}}(t)$ ——河道上、下断面的流量过程; t ——时间; C_1,C_2,C_3 ——系数.

类似地,具有预见期的洪水演算方法为^[1,2]

$$Q_{\text{断}}(t+1)=C_1Q_{\text{断}}(t)+C_2Q_{\text{断}}(t-1)+C_3Q_{\text{断}}(t)$$

(9)

模仿式(8)和(9),分别建立马斯京根和具有预见期的径向基函数神经网络模型,网络结构见表1.人工神经网络的学习过程就是识别表1中的函数关系 f ,相当于求得式(8)和(9)中的系数 C_1,C_2 和 C_3 ,只不过求出来的这些系数只能用一组神经网络结构来表达,而不能显式地表示.

表1 马斯京根方法和具有预见期方法的人工神经网络结构

Table 1 ANN architecture for Muskingum method and the method with a forecast period

方法	函数关系	网络结构
马斯京根	$Q_{\text{断}}(t+1)=f[Q_{\text{断}}(t),Q_{\text{断}}(t+1),Q_{\text{断}}(t)]$	3-8-1
具有预见期	$Q_{\text{断}}(t+1)=f[Q_{\text{断}}(t),Q_{\text{断}}(t-1),Q_{\text{断}}(t)]$	3-8-1

2.2 基本资料及其处理

现以汉江上襄阳—皇庄河段的6场次洪水和西江上梧州—高要河段的8场次洪水来验证径向基函数神经网络方法.其中襄阳—皇庄河段的前4场次洪水以及梧州—高要河段的前6场次洪水用于神经网络学习,其余场次洪水用于校核.根据洪水过程的涨落情况,两个河段的计算时段均取为6h.

由于径向基函数的值位于 $[0, 1]$ 之间, 需要将输入和输出流量数据进行规格化处理, 即

$$y_i = \frac{z_i - z_{\min}}{z_{\max} - z_{\min}} \alpha + \beta$$

(10)

式中: z_i, y_i ——处理前后的变量; $z_{\max}, z_{\min}$ ——变量 z 的最大和最小值; α, β ——参数, 一般取值分别为 0.9 和 0.05.

2.3 计算结果

洪水演算过程按照表 1 的网络结构进行. 由表 1 的网络结构输入输入层和输出层的流量数据, 计算得到的结果见表 2. 从表 2 可以看出, 不论是用于学习的若干场次洪水, 还是用于校核的那几场次洪水, 基于马斯京根方法和具有预见期方法所得到的精度均较高. 同时还可以看出, 这两个演算方法所得到的计算结果非常接近, 这是由于这两种洪水演算方法都是扩散波方程的二阶精度数值解^[1, 2], 精度相仿.

表 2 马斯京根方法和具有预见期方法的洪水演算结果

Table 2 Simulated results of flood volume based on Muskingum method and the method with a forecast period

类型	站名	洪号	马斯京根方法					具有预见期方法				
			洪峰			峰现误差 (时段)	确定性 系数	洪峰			峰现误差 (时段)	确定性 系数
			实测流量 ($\text{m}^3 \cdot \text{s}^{-1}$)	计算流量 ($\text{m}^3 \cdot \text{s}^{-1}$)	误差 /%			实测流量 ($\text{m}^3 \cdot \text{s}^{-1}$)	计算流量 ($\text{m}^3 \cdot \text{s}^{-1}$)	误差 /%		
学 习	皇庄	741 002	18 900	19 360	2.4	0	0.82	18 900	18 950	0.3	1	0.82
		800 720	10 400	10 594	1.9	0	0.91	10 400	10 313	-0.8	0	0.95
		810 820	12 600	13 218	4.9	1	0.96	12 600	12 763	1.3	1	0.97
		810 901	10 900	11 193	2.7	1	0.96	10 900	10 918	0.2	1	0.97
		780 516	36 367	33 308	-8.4	0	0.93	36 367	33 550	-7.7	0	0.94
	高要	800 713	21 700	22 758	4.9	1	0.95	21 700	23 332	7.5	1	0.95
		800 812	25 900	25 500	-1.5	0	0.92	25 900	24 805	-4.2	0	0.96
		820 506	22 600	22 697	0.4	0	0.99	22 600	23 255	2.9	1	0.98
		820 809	21 500	22 315	3.8	1	0.97	21 500	22 727	5.7	1	0.97
		830 616	34 367	33 308	-3.1	0	0.98	34 367	33 550	-2.4	0	0.99
校 核	皇庄	820 727	10 600	10 383	-2.0	-1	0.93	10 600	10 215	-3.6	0	0.95
		830 730	12 100	12 924	6.8	-1	0.95	12 100	12 393	2.4	-1	0.98
	高要	840 515	16 600	14 722	-11.3	-1	0.87	16 600	15 021	-9.5	-1	0.87
		840 525	23 200	22 875	-1.4	0	0.97	23 200	23 457	1.1	1	0.97

3 结 语

河道洪水演算是一个由来已久的研究课题, 而马斯京根方法是应用最为广泛的方法之一, 具有预见期的洪水演算方法是近些年来提出并加以研究的一类方法. 本文基于这两种洪水演算方法, 构建洪水演算的径向基函数神经网络模型, 并且采用最小二乘算法来确定神经网络参数. 两条天然河道的洪水演算实例表明, 该模型计算简便、快速, 精度令人满意.

径向基函数神经网络方法不仅具有良好的推广能力, 而且避免了像 BP 网络那样繁琐、冗长的计算, 使得学习过程可以比通常的 BP 方法快很多. 但也不难看出, 和传统的洪水演算方法相比, 径向基函数神经网络的建模过程同样也是一个非常繁杂的过程. 隐层数目的确定在很大程度上建立在试错的基础之上, 隐层单元参数的确定方法经验性较强, 这些还有待深入研究.

参考文献:

[1] 芮孝芳, 王伶俐. 具有预见期的洪水演算方法研究[J]. 水科学进展, 2000, 11(3): 291—295.
[2] 黄国如, 芮孝芳. 基于运动波数值扩散的洪水演算方法[J]. 河海大学学报(自然科学版), 2001, 29(2): 110—113.
[3] 李致家, 孔祥光. 河道洪水演算的神经网络模型[J]. 河海大学学报, 1997, 25(5): 7—12.
[4] 冯利华. 基于神经网络的洪水预报研究[J]. 自然灾害学报, 2000, 9(2): 45—48.
[5] 苑希民, 刘树坤, 陈浩. 基于人工神经网络的多泥沙洪水预报[J]. 水科学进展, 1999, 10(4): 393—398.

[6] 王建群 ,董增川 ,吴浩云 . 太湖洪水预报的前向人工神经网络全局最优逼近 [J]. 河海大学学报(自然科学版) ,2001 ,29(4) : 84—90 .
[7] 徐秉铮 ,张百灵 ,韦岗 . 神经网络理论与应用 [M]. 广州 :华南理工大学出版社 ,1994 .304—316 .
[8] 杨建基 ,罗先香 . 地下水动态预测的径向基函数法 [J]. 水文 ,2001 ,21(4) :1—3 .
[9] 郭宗楼 ,刘肇 . 环境质量综合评价的径向基函数网络模型 [J]. 环境科学研究 ,1997 ,10(4) :1—5 .
[10] 钟秉林 ,黄仁 . 机械故障诊断学 [M]. 北京 :机械工业出版社 ,1997 .89—91 .

Radial basis function-neural network model for channel flood routing

HUANG Guo-ru , RUI Xiao-fang

(College of Water Resources and Environment , Hohai Univ . , Nanjing 210098 , China)

Abstract :The radial basis function-neural network method is applied to channel flood routing . In combination with the specific modes of channel flood routing , radial basis function-artificial neural network models , based on the Muskingum method and flood routing method with a forecast period , are developed , and the least square method is used to determine the parameters of the models . The models are applied to flood routing for two natural river channels , and the results show that the models have the characteristics of fast calculation , high precision , and high value of application .

Key words :flood routing ; Muskingum method ; method with a forecast period ; radial basis function ; artificial neural network

.....

· 简讯 ·

国际水文科学协会主席竹内邦良先生被聘为河海大学名誉教授

2003 年 9 月 25 日河海大学举行仪式 ,聘任国际水文科学协会主席竹内邦良(Kuniyoshi Takeuchi)先生为河海大学名誉教授 .聘任仪式由水资源环境学院董增川院长主持 ,严以新副校长为竹内邦良先生颁发证书 .竹内邦良先生 1941 年出生 ,1972 年获美国卡罗琳娜大学哲学博士 ,1982 年获东京大学工学博士 ,日本山梨大学水文学教授 ,先后在多个国际学术组织任职 ,如 1998 ~ 2000 年担任国际水资源协会副主席及 UNESCO 国际水文计划(IHP)政府间理事会主席 ,2001 年起任国际水文科学协会主席 ,在该学术领域有较高声望和影响 ,曾 3 次访问河海大学 .

竹内邦良先生对河海大学聘任他为名誉教授感到非常高兴 .他认为 ,中国在近几年取得了很大的进步 ,但中国是个自然灾害频发的国家 ,洪涝灾害尤其严重 .另外现在不仅是中国、日本以及其他很多国家都存在严重的水资源问题 ,因此大家应该加强交流与合作 .他表示希望以他在国际学术组织工作的经验和影响来帮助河海大学进一步加强与国际学术组织的交流 ,并愿意在水资源问题上与中国有更多的合作 .

仪式结束后 ,竹内邦良名誉教授以亚太地区的水问题和全球水资源的开发利用为主题 ,给学校师生作了一个精彩的学术报告 .

河海大学水资源环境学院 陈元芳