

模糊回声状态网络

彭 宇¹, 王建民^{1,2}, 彭喜元¹

(1. 哈尔滨工业大学自动化测试与控制研究所, 黑龙江哈尔滨 150080;
2. 哈尔滨理工大学电子科学与技术系, 黑龙江哈尔滨 150001)

摘 要: 针对基于梯度下降的模糊递归神经网络训练效率低、容易陷入局部极小的缺点, 本文基于回声状态网络(Echo State Networks, ESNs)和 TS 模型提出一种新的模糊模型结构——模糊回声状态网络(Fuzzy Echo State Networks, FESNs). FESNs 由多条 TS 类型的模糊规则组成, 规则后件采用 ESNs 网络. 研究表明, TS 模型和 ESN 都可以看做是 FESN 模型的某种特例, 而且 FESNs 具有较强的非线性映射能力、局部反馈以及学习算法稳定等特点. 同时, 其模型参数确定方法与经典 TS 模型以及 ESN 一样可以归结为一个线性回归问题, 大大减少了网络训练的计算量. 仿真实验表明, 与经典 TS 模型相比, FESNs 在不显著增加建模时间情况下可有效提高建模精度.

关键词: 回声状态网络; 储备池; TS 模型; 动态系统建模

中图分类号: TP183 **文献标识码:** A **文章编号:** 0372-2112 (2011) 07-1538-07

Fuzzy Echo State Networks

PENG Yu¹, WANG Jian-min^{1,2}, PENG Xi-yuan¹

(1. Automatic Test and Control Institute, Harbin Institute of Technology, Harbin, Heilongjiang 150080, China;
2. Department of Electronics Science & Technology, Harbin University of Science and Technology, Harbin, Heilongjiang 150001, China)

Abstract: Recurrent fuzzy neural networks, which is usually trained by gradient descent, have some inherent shortcomings such as inefficiency of training process, local minimum. In this paper, we proposed a novel fuzzy neural network based on Echo state network and TS fuzzy model, called fuzzy echo state network (FESN), which is a generality of both TS modal and ESN. FESNs consist of several fuzzy IF-THEN rules, each of which has a ESN as consequent part. We illustrate that FESNs have some interesting characteristics, such as better nonlinear mapping capacity, local feedback and stable learning, which results in that FESNs can deal with dynamics of nonlinear system. Furthermore, similar to the TS model and ESN, parameters of the FESNs can be determined by solving a linear regression problem, which dramatically reduce the computing burden of the training process. Experiments shows that FESN can effectively enhance the accuracy of modeling dynamical system at the expense of not using excessive additional time compared with TS model and ESN.

Key words: echo state networks; reservoir; TS model; modeling dynamical system

1 概述

无论是自动控制、模式识别还是时间序列分析和预测问题, 其实现的基础都是必须首先建立系统精确的模型. TS 模糊模型^[1]因模型结构和参数确定方法简单而吸引了大量研究人员的关注, 同时也在实际环境中得到广泛的应用. 尽管理论上已经证明, TS 模型能够以任意精度逼近任意的非线性系统, 但是由于采用线性函数作为模糊规则后件, 模型的非线性映射能力明显受到了制约. 对于非线性较强的系统, 需要使用非常多的规则才能精确地建立系统模型. 同时为了使 TS 模型能够处理

与时序相关的动态系统问题, 还需要对 TS 模型做出修正, 使模型结构上形成某种反馈结构. 近年来, 有大量的研究涉及在保留 TS 模型的原有优点的基础上, 提高 TS 模型的表达能力, 同时使其适合处理与时序相关的动态系统问题^[2-7].

文献[2]提出了一种动态反馈模糊神经网络 (Locally Feedback Dynamic Fuzzy Neural Network, LF-DFNN). LF-DFNN 在 TS 模型框架之下, 采用具有局部反馈的无限冲击响应多层感知机 (IIR-MLP) 作为模糊规则后件, 显著提高了输入输出的非线性映射能力. 与其他的风速预测模型相比, LF-DFNN 模型显示出更好的准确性. 类似的

文献还有文献[3~5],这些工作的共同特点是通过引入递归神经网络(Recurrent Neural Networks, RNNs)作为模糊规则的后件,从而显著地提高了模型的非线性映射能力.但是,由于模型参数(连接权)需要通过误差梯度最小的算法进行训练,与经典的 TS 模型相比,模型的训练效率大大降低(递归神经网络的训练本身就极为困难),同时无法避免局部极小的问题.

文献[6]提出了一种类似 TS 模型结构的模糊递归神经网络,该模型采用的规则形式及后件的形式都与经典 TS 模型一致.为了处理动态系统问题,该模型需要额外引入若干节点以形成反馈结构.引入反馈结构的目的是提高模型处理动态系统建模的能力,但由于仍然采用的是线性后件,模型的非线性映射能力依然受到限制,而且由于额外节点的引入,使得模型采用的模糊规则会显著增加,导致计算效率的下降.

鉴于此,本文基于模糊 TS 模型和回声状态网络(Echo State Networks, ESNs)^[8~11],提出一种新的模型结构,称之为模糊回声状态网络(Fuzzy Echo State Networks, FESNs).FESNs 采用 TS 类型的模糊规则,并采用 ESNs 做为规则后件.ESNs 中的“储备池”不但引入了局部反馈同时将输入从低维输入空间映射到高维状态空间,显著提高了模型的非线性映射能力.而且,FESNs 模型参数的确定与经典 TS 模型参数确定方式一致,最终归结为一个线性回归问题,保持了经典 TS 模型参数确定简单的优点.

回声状态网络是一种全新的递归神经网络,该网络采用大规模随机布置的神经元作为信息处理单元,与传统的 RNNs 相比,ESNs 采用储备池(Reservoir)作为信息处理单元,只有部分的网络连接权值需要训练,而且训练算法基于均方误差最小(线性回归问题),与传统的基于误差梯度最小的训练算法相比,训练效率大大提高.

本文的贡献在于提出了 FESNs 模型的框架,解决经典 TS 模型表达能力不强的问题,同时解决了 TS 模型加 RNNs 结构的递归模糊神经网络训练效率低的问题.在此基础上提出 FESNs 模型的参数确定方法,并通过实验初步验证了该模型在动态系统建模问题中的有效性.

2 回声状态网络的数学模型

本节介绍经典回声状态网络的数学模型.这里所说的“经典”回声状态网络由 Jaeger 在文献[8]中提出,与其他改进和修正的模型相区分.

2.1 回声状态网络的数学模型

回声状态网络的一般结构如图 1 所示^[8].其状态更新方程表示为

$$\mathbf{x}(t) = \boldsymbol{\phi}(\mathbf{W}^{in}\mathbf{u}(t) + \mathbf{W}\mathbf{x}(t-1) + \mathbf{W}^{back}\mathbf{d}(t-1)) \quad (1)$$

其中 $\boldsymbol{\phi}(\cdot) = (\phi_1, \dots, \phi_N)'$ 表示神经元激活函数, ' 表示矩阵转置. $\mathbf{x}(t)$ 表示储备池状态向量在 t 时刻的取值, $\mathbf{d}(t-1)$ 表示系统在 $t-1$ 时刻的输出. $\mathbf{u}(t)$ 表示 t 时刻的输入向量. \mathbf{W}^{in} , \mathbf{W} 和 \mathbf{W}^{back} 分别是输入、储备池和输出的连接权矩阵.通常情况下,回声状态网络的输出由下式给出

$$\mathbf{d}(t) = \boldsymbol{\phi}^{out}(\mathbf{W}^{out}\mathbf{x}(t)) \quad (2)$$

其中 $\boldsymbol{\phi}^{out} = (\phi_1^{out}, \dots, \phi_M^{out})$ 表示输出函数,根据问题的不同 $\boldsymbol{\phi}^{out}$ 可以取线性函数或者 S 型函数. \mathbf{W}^{out} 表示输出连接权矩阵,需要通过在线或者离线的算法训练得到.如果系统具有 L 个输入单元, N 个内部神经元, M 个输出单元,则有 $\mathbf{W}^{in} \in \mathbf{R}^{N \times M}$, $\mathbf{W} \in \mathbf{R}^{N \times N}$, $\mathbf{W}^{back} \in \mathbf{R}^{N \times 1}$, $\mathbf{W}^{out} \in \mathbf{R}^{M \times (N)}$, \mathbf{R} 表示实数集.

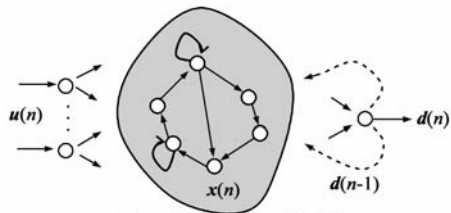


图1 经典ESNs网络结构^[8]

2.2 ESNs 的训练

ESNs 的训练过程就是根据给定的输入输出训练样本对 $(\mathbf{u}(t), \mathbf{d}(t), t = 1, 2, \dots, T)$, 确定系统输出连接权矩阵 \mathbf{W}^{out} 的过程.不失一般性,这里设 $\mathbf{W}^{back} = 0$.回声状态网络的训练过程可以分为两个阶段:采样(Sampling)阶段和权值计算(Weight Computation)阶段^[10].

2.2.1 采样

采样阶段首先选定网络的初始状态 $\mathbf{x}(0)$, 通常选择为 0.训练输入样本 $(\mathbf{u}(t), n = 1, 2, \dots, T)$ 经过输入连接权矩阵 \mathbf{W}^{in} 加入到网络,根据式(1)完成对状态 $\mathbf{x}(t)$ 和输出 $\mathbf{d}(t)$ 的计算与收集.为了消除初始状态对于训练过程的影响,舍弃网络最初运行时前 l 个状态,从某一时刻 l 开始收集网络的状态,并以向量 $\mathbf{s}(i) (i = l, l+1, \dots, T)$ 为行构造系统状态矩阵 $\mathbf{B} = (b_{ij})_{(T-l+1) \times (N)}$, 同时相应的输出样本 $\mathbf{d}(t)$ 也被收集,并构成一个列向量 $\mathbf{D} = (d_{ij})_{(T-l+1) \times 1}$.

2.2.2 权值计算

根据在采样阶段得到的网络的状态矩阵 \mathbf{B} 和输出样本向量 \mathbf{D} , 就可以计算输出连接权矩阵 \mathbf{W}^{out} .考虑用网络输出 $\hat{\mathbf{d}}(t)$ 逼近实际样本输出 $\mathbf{d}(t)$, 即

$$\mathbf{d}(t) \approx \hat{\mathbf{d}}(t) = \mathbf{W}^{out}\mathbf{x}(t)$$

从而输出权值 \mathbf{W}^{out} 的训练问题,可以归结为如下的优化问题:

$$\min_{\mathbf{W}^{out}} \frac{1}{T-l+1} \sum_{t=l}^T (\mathbf{d}(t) - \mathbf{W}^{out}\mathbf{x}(t))^2 \quad (3)$$

从数学的观点看,这是一个线性回归问题,进而可

以归结为求矩阵 B 的逆矩阵问题,即

$$W^{out} = B^{-1}T$$

至此,ESNs 网络训练已经完成.

3 FESNs 的设计

经典 TS 模型^[1]由多条模糊规则组成,具体需要多少条规则与问题的复杂性有关,其中每一条规则的形式如下

R^i : IF $u_1(t)$ is $A_i^1(u_1(t)), \dots, u_M(t)$ is $A_i^L(u_L(t))$

THEN $y_i(t) = a_i^0 + a_i^1 u_1(t) + \dots + a_i^L u_L(t)$

其中 $R^i (i=1,2,\dots,r)$ 表示模型的第 i 条规则,整个模型共有 r 条规则组成. $u_i(t)$ 是第 i 维输入在 t 时刻的取值, L 表示输入向量的维数,为了叙述方便,定义 $\mathbf{u}(t) = [u_1(t), u_2(t), \dots, u_L(t)]'$. $A_i^j(u_j(t))$ 表示规则 R^i 中定义在第 j 维输入上的隶属函数, $j=1, \dots, L$. $y_i(t)$ 表示规则 R^i 对应的输出. 参数 a_i^0, \dots, a_i^M 是规则 R^i 后件线性方程对应的参数.

TS 模型在局部采用一个线性关系表达输入和输出之间的映射关系,然后再采用“加权平均”的方式将每个线性关系“粘”起来. 本文提出的 FESNs 采用与经典 TS 模型类似的模型结构(如图 2 所示),不同之处在于 FESNs 采用 ESNs 作为规则的后件,即在局部采用 ESNs 提高模型的非线性映射能力. 本节将提出 FESNs 的框架结构以及模型参数的确定方法,同时结合模型结构以及参数确定方法介绍 FESNs 的主要特点.

3.1 FESN 的结构

与经典 TS 模型不同, FESNs 的规则后件是一个回声状态网络,其具体形式如下

$$\begin{aligned} R^i: & \text{ IF } \mathbf{u}(t) \text{ is } A_i(\mathbf{u}(t)) \\ & \text{ THEN } (ESN)_i \end{aligned} \quad (4)$$

其中, $(ESN)_i$ 表示规则 R^i 对应的 ESN,其定义如下

$$\begin{cases} \mathbf{x}_i(t) = \phi_i(\mathbf{W}_i^{in} \mathbf{u}(t) + \mathbf{W}_i \mathbf{x}_i(t-1) + \mathbf{W}_i^{back} \mathbf{d}_i(t-1)) \\ \hat{\mathbf{y}}_i(t) = \mathbf{W}_i^{out} [1; \mathbf{x}_i(t)] \end{cases} \quad (5)$$

$A_i(\mathbf{u}(t))$ 表示定义在 M 维输入空间上的模糊隶属函数,这里关于隶属函数的定义方式与经典 TS 模型稍有不同,二者的关系可以表示为

$$A_i(\mathbf{u}(t)) = A_i^1(u_1(t)) \wedge \dots \wedge A_i^L(u_L(t))$$

其中,符号 \wedge 表示取乘积或者取最小运算, L 表示输入向量的维数. $\mathbf{x}_i(t) = [x_1(t), x_2(t), \dots, x_{N_i}(t)]^T$ 表示规则 R^i 所对应的 ESNs 的 N_i 维状态向量, $\mathbf{W}_i^{in} (N_i \times M)$ 、 $\mathbf{W}_i (N_i \times N_i)$ 、 $\mathbf{W}_i^{back} (N_i \times M)$ 和 $\mathbf{W}_i^{out} (M \times (N_i + L))$ 分别表示规则 R^i 所对应 ESNs 的输入连接权矩阵、内部连接权矩阵、反馈连接权以及输出连接权矩阵. $\mathbf{W}_i^{back} (N_i \times$

$M)$ 表示输出到储备池的反馈连接权,在具体应用中该连接权是否需要,可根据实际情况选择. $\phi_i = (\phi_i^1, \phi_i^2, \dots, \phi_i^{N_i})$ 表示规则 R^i 所对应的 ESNs 的神经元激活函数. FESNs 的结构如图 2 所示(为了简单起见,只画出具有 2 条模糊规则以及 1 个输出的情形), β_i 表示规则 R^i 对应的隶属权重,其定义见 3.2.2 节. 图中的虚线框(深色阴影部分)表示 ESNs, ESNs 的储备池(浅色阴影部分)中存在大量随机布置的神经元,整个 ESNs 作为规则的后件.

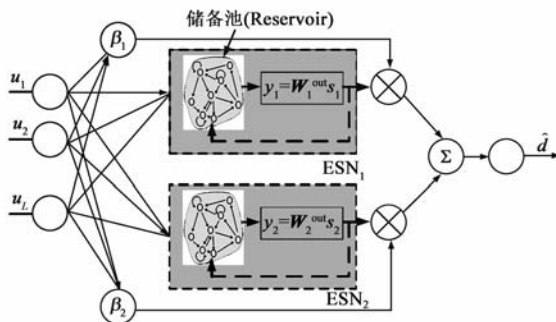


图2 FESNs 的结构框图

分析图 2、式(4)和(5),不难发现, FESNs 的核心思想在于:在模糊规则的后件中引入储备池,从而将输入从低维的输入空间映射到了高维的状态空间. 储备池内部神经元之间存在反馈连接,从而在模型中引入了局部反馈,提高了模型处理动态系统建模问题的能力. 这里需要强调的是储备池中神经元的连接权矩阵 \mathbf{W}_i , 输入到储备池连接权矩阵 \mathbf{W}_i^{in} , 以及输出到储备池的反馈连接权矩阵 \mathbf{W}_i^{back} 是在整个网络初始化过程中随机产生的.

不难发现,上述结构的 FESNs 具备如下特点:

- (1) 通过引入储备池,将低维输入映射到高维状态空间(与核方法中核函数的思想类似),提高了模型的非线性映射能力;
- (2) 根据每条规则所处理的局部特性不同, FESNs 可以为每一条规则中的储备池设置不同的参数 (\mathbf{W}_i^{in} 、 \mathbf{W}_i 和 \mathbf{W}_i^{back} 等),提高模型的非线性映射能力以及处理具有不同特性叠加在一起的动态系统的建模能力;
- (3) 输入通过随机产生的连接权 \mathbf{W}_i^{in} 与储备池中的神经元直接相连,使得储备池可以直接捕捉并记忆输入信号携带的信息;
- (4) 通过储备池引入局部反馈,提高 FESNs 处理与时序相关的动态系统建模和预测能力.

由于储备池的引入,每条规则处理非线性问题的能力将显著提高,可以显著降低建模过程中使用的模糊规则数. 避免模糊建模过程出现的“维数灾难”问题.

根据 FESN 的定义(4),容易发现,如果 FESN 的输出为 1 维变量,即 $M=1$,那么,如果令 $\mathbf{W}_i^{in} = I$, $\mathbf{W}_i = 0$,

$W_i^{back} = 0 (i = 1, 2, \dots, r)$, 那么 FESN 将退化为经典的 TS 模型. 换句话说, TS 模型只是 FESN 在某些特征取值下的一个特例. FESN 模型是 TS 模型的推广.

同理, 如果在 FESN 的定义(4)中, 令 $r = 1$, 即只有一条规则, 同时不考虑模糊集合, 只考虑清晰的集合, FESN 将退化为 ESN 模型, 换句话说, ESN 模型也是 FESN 模型的一个特例, FESN 模型同时也是 ESN 模型的一个推广.

3.2 FESNs 参数确定

以上介绍了 FESNs 的结构以及 FESNs 主要结构特点, 下面介绍如何通过求解一个线性回归问题确定 FESNs 参数. 这种方法与经典的模糊递归神经网络相比具有更高的效率, 同时可以避免陷入局部最优的问题.

3.2.1 前件参数的确定

FESNs 模型的前件参数与 TS 模型前件参数的确定方法一样, 即用于 TS 模型前件参数的确定方法都可以用于 FESNs 前件参数的确定. 采用 TS 模型对系统进行建模一般有两种方式, 一种是根据系统的组成结构、工作原理以及物理特性, 由领域专家根据实际经验直接给出模型的前件参数, 这种方法随着系统复杂程度的不断提高, 难度变得越来越大. 另一种方法被称为系统辨识, 系统辨识方法根据网络的输入输出数据, 采用人工智能方法建立与系统真实物理模型等价的数学模型. 本文关注第二种方法. 采用系统辨识的方法建立复杂动态系统的 TS 模型的经典方法可以参考文献[1, 12, 13]. 本文介绍一种简单的通过模糊 C 均值聚类^[14] (Fuzzy C-Means, FCMs) 确定系统前件参数的方法.

模糊 C 均值聚类是一种常用的确定模糊规则前件的方法. 通过对输入变量进行 FCMs 即可提取输入信号的有用信息, 从而确定模糊规则前件的所有参数. 聚类过程中采用的聚类中心数就是模型模糊规则数. FCMs 算法的目标函数为^[14]

$$Q = \sum_{k=1}^p \sum_{i=1}^r a_{i,k}^m \| \mathbf{u}(k) - \mathbf{v}_i \|^2$$

其中 $\mathbf{u}(k)$ 表示 k 时刻的输入向量, \mathbf{v}_i 表示第 i 个聚类中心, m 被称为指数权重, m 的取值会影响隶属度函数的模糊化程度. 令 $A_i = (a_{i,k})_{(p \times 1)} (k = 1, 2, \dots, p)$ 就是规则前件的模糊隶属函数.

3.2.2 后件参数的确定

FESNs 模型需要确定的后件参数主要是每条规则中 ESNs 网络的输出连接矩阵 W_i^{out} , 其他的后件参数 W_i^{in} 、 W_i 和 W_i^{back} 是根据具体问题在模型初始化过程中随机产生的. 以下介绍参数 W_i^{out} 的确定方法.

假设网络由具有式(4)形式的 r 条规则组成, 那么模型输出 $\hat{\mathbf{y}}(t)$ 由式(6)给出

$$\begin{aligned} \hat{\mathbf{y}}(t) &= \frac{\sum_{i=1}^r A_i(\mathbf{u}(t)) (W_i^{out} [1; \mathbf{x}_i(t)])}{\sum_{i=1}^r A_i(\mathbf{u}(t))} \\ &= \frac{\sum_{i=1}^r A_i(\mathbf{u}(t)) (w_{i,0}^{out} + w_{i,1}^{out} x_1^i(t) + \dots + w_{i,N_i}^{out} x_{N_i}^i(t))}{\sum_{i=1}^r A_i(\mathbf{u}(t))} \end{aligned} \quad (6)$$

其中 $W_i^{out} = (w_{i,0}^{out}, w_{i,1}^{out}, \dots, w_{i,N_i}^{out})$; $i = 1, 2, \dots, r$; $\mathbf{x}_i(t)$ 由式(5)给出.

$$\text{令} \quad \beta_i(t) = \frac{A_i(\mathbf{u}(t))}{\sum_{i=1}^r A_i(\mathbf{u}(t))}$$

则式(6)可化简为

$$\hat{\mathbf{y}}(t) = \sum_{i=1}^r \beta_i(t) (w_{i,0}^{out} + w_{i,1}^{out} x_1^i(t) + \dots + w_{i,N_i}^{out} x_{N_i}^i(t)) \quad (7)$$

如果给定系统的一组输入和输出样本, $\{\mathbf{u}(t); \mathbf{y}(t), t = 1, 2, \dots, p\}$, 采用最小二乘法就可以确定 FESNs 模型的后件参数 W_i^{out} . 令

$$\begin{aligned} \mathbf{x}(t) &= [\beta_1(t), \beta_1(t) x_1^1(t), \dots, \\ &\quad \beta_1(t) x_{N_1}^1(t), \dots, \beta_r(t) x_{N_r}^r(t)] \end{aligned}$$

同时为了消除暂态效应(网络的初始权值等对结果的影响)的影响, 舍弃前 $l-1$ 个样本, 并构造如下 3 个矩阵(向量),

$$\mathbf{X} = [x(l); x(l+1); \dots; x(T)]$$

$$\mathbf{Y} = [y(l); y(l+1); \dots; y(T)]$$

$$\begin{aligned} \mathbf{W} &= [w_{1,0}^{out}, w_{1,1}^{out}, \dots, w_{1,N_1}^{out}, w_{2,0}^{out}, w_{2,1}^{out}, \dots, w_{2,N_2}^{out}, \\ &\quad \dots, w_{R,0}^{out}, w_{R,1}^{out}, \dots, w_{R,N_R}^{out}] \end{aligned}$$

其中 $[\cdot; \cdot]$ 表示矩阵元素的垂直拼接.

则 FESNs 模型的参数 \mathbf{W} 可以由(8)求得

$$\mathbf{W} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{Y} \quad (8)$$

算法的本质是一个均方误差最小的凸优化问题, 不存在局部极小的问题. 克服了经典模糊递归神经网络训练效率低, 容易陷入局部极小的问题.

4 实验验证

为了验证 FESNs 方法的有效性, 首先结合一个动态系统建模问题对其基本特性进行仿真实验分析, 再通过一个实际的移动通信话务量时间序列预测问题证明其应用价值.

4.1 动态系统建模问题

4.1.1 问题描述

考虑如下的动态系统的建模问题^[15]

$$v_{i+1} = (1 + v_i^2)^{-1} (-0.9v_i + x_i) \quad (9)$$

选定系统输入信号 x_i 为正弦激励信号 $x_i = 0.75\sin(\pi(i-1)/180)$, 其中 $i = 1, 2, \dots, 500$, 输出 v_i 的初始状态为 $v_0 = 0.1$. 同时, 为了检测 FESNs 噪声抑制的能力, 在输出 v_i 基础上加入均值为 0 均匀分布的噪声信号 z_i , 噪声信号的上下限取为 ± 005 , 即实际使用的数据为 $\{(x_i, y_i = v_i + z_i), i = 1, 2, \dots, 500\}$. 前 500 个数据用于建模, 接下来 500 个样本数据用于测试. 采用均方根误差作为误差评价指标, 其定义为:

$$prmse = \sqrt{\frac{1}{p-l} \sum_{i=m}^p (y_i - \hat{y}_i)^2}$$

为了衡量 FESNs 方法的建模精度和效率, 基于以上的实验数据, 分别采用经典 ESNs 方法、经典 TS 模型以及 BP 神经网络进行对比实验.

4.1.2 实验参数设置

实验过程首先需要确定模糊规则前件的隶属函数, 实验过程中采用两条模糊规则, 对样本数据进行 FCMs 聚类, 得出的隶属函数如图 3 所示 (多次实验中的 1 次).

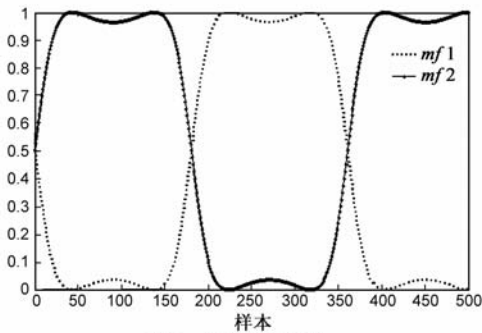


图3 模糊隶属函数

实验过程中 FESN 每条规则后件的 ESN 选取相同的储备池参数, 具体数值设置如下: 储备池规模 $N = 30$, 内部连接权矩阵谱半径 $SR = 0.75$, 输入尺度 $IS = 0.2$, 分别选取了 leaky^[14] 和 tanh (双曲正切) 两种类型的神经元进行实验.

经典 ESN 模型的参数设置与 FESNs 中每条规则后件 ESNs 网络参数设置一致; 经典 TS 模型采用与 FESNs 相同的模糊规则数(2).

BP 神经网络模型参数选取如下: 输入单元数为 5 个, 输出单元数为 1 个, 隐层神经元数为 7, 神经元激活函数为双曲正切函数, 输出层神经元选择线性激活函数.

表 1 给出了 FESNs 与经典 ESNs、TS 模型以及 BP 神经网络三种方法对系统 (9) 进行建模和预测时性能对比情况. 表中数据为进行 50 次实验所得到的平均值, 括号内给出的是方差.

表 1 FESN 与其他三种方法的性能比较

#	训练误差	测试误差	训练时间(s)
FESN(leaky)	0.0298(0.0022)	0.0178(0.0313)	0.5823
FESN(tanh)	0.0285(0.0053)	5.6×10^{-4} (0.0012)	0.4680
ESN 网络	0.0301(2.7×10^{-4})	0.5194(1.3161)	0.1463
TS 模型	0.0308(7.3×10^{-4})	0.0245(0.0068)	0.02875
BP 神经网络	0.0343(0.002)	0.0363(0.0270)	1.8946

4.1.3 实验结果分析

从建模准确率角度看(训练误差), FESNs 显示出更高的建模准确率, 但与其他几种方法差别不大, 训练误差值以及标准差的差别都很小, 这与理论分析的结果是一致的, 即以上几种方法都能够以任意精度逼近复杂的动态系统.

从预测准确率角度看, FESNs 具有最好的性能. 采用 tanh 类型神经元的 FESNs 在预测准确率上甚至较其他方法有 2 个数量级的提高. 与经典 TS 模型相比, 预测精度有 40 倍左右的提高. 与经典 ESNs 方法相比, 预测精度有更多的提高, 大约有 900 倍. 与 BP 神经网络相比, 也有大约 60 倍左右的提高, 与提出 FESNs 方法的初衷一致, 显著提高建模的精度.

从模型效率角度看, TS 模型效率是最高的, FESNs 次之, BP 是最差的. 原因在于 TS 模型采用的是线性后件、模型结构最为简单有关. 而 BP 神经网络由于需要采用基于梯度下降法进行训练才能确定其网络参数, 因而具有最差的效率. FESN 虽然引入了储备池, 但是不会显著降低建模效率, 这是因为由于储备池参数 (W , W^{in} , W^{back}) 是在模型初始化时随机产生的, 不需要训练, 需要确定的参数依然是通过最小二乘法直接获得的, 与经典 TS 模型相比其效率并没有显著下降, 与 BP 神经网络相比有明显提高.

总之, FESNs 在建模的效率上与经典 TS 模型基本一致, 但其非线性映射能力显著提高, 甚至超过 BP 神经网络.

此外, 与经典 ESNs 相比, FESNs 具有明显的噪声抑制能力. 图 4 和图 5 分别给出了采用 ESNs 和 FESNs 对式 (9) 进行建模和预测情况. 通过对比可以发现, FESNs

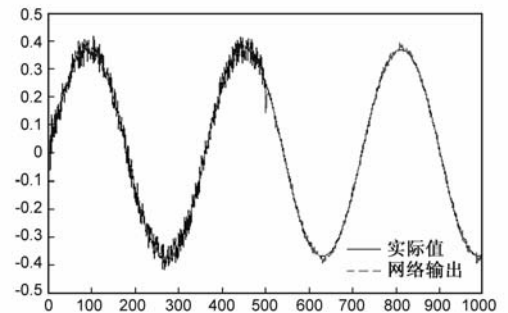


图4 ESN训练和预测误差

除了具有更小的训练和测试误差外,明显去除了样本数据中加入噪声的影响.这种噪声的抑制能力对于实际应用是至关重要的,因为实际通过测试系统获得数据或多或少都会带有一定的噪声,如果模型不具有噪声抑制能力,会导致建立的模型过多地受到噪声的影响,而偏离真实数据,影响后续对其进行的各种处理.

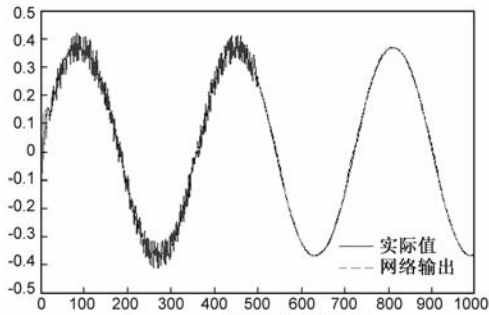


图5 FESN训练和测试误差

4.2 移动通信话务量时间序列预测

移动通信话务量的变化在一定程度上体现了话音信道被占用的强度,是网络管理和优化的重要技术指标之一.但是,话务量时间序列数据具有强非线性、高噪声、非平稳等特点^[14],传统的时间序列建模和预测方法一般难以取得让人满意的预测效果.本文考虑采用 FESNs 对移动通信话务量时间序列进行建模和预测,以验证 FESNs 在实际应用中是有效的.

4.2.1 问题描述

实验所采用的数据来自中国移动黑龙江有限公司,其提供的话务数据以 1 小时为采集间隔,包含全省所有小区共约 500,000 条记录.数据属性包括记录时间、小区名称、话务量、信道拥塞率、切换成功率等.其中话务量变化规律和趋势预测是我们现阶段的分析目标.按照一定时间尺度(一般是小时)进行采集和计算的话务量数据序列是典型的时间序列.

4.2.2 实验参数设置

实验以话务小区为单位进行,选取 2008 年 4 月 1 日到 4 月 29 日的 696 个话务数据作为训练样本,预测其后 1 天(4 月 30 日)24 小时的话务量变化情况.首先将时间序列的预测问题转换为动态系统的建模问题.为此,需要构造 FESNs 的输入和输出向量,具体方法如下: $\mathbf{u}(k) = [d(k), d(k+24), d(k+48), d(k+72), d(k+96)]$ 输出 $\mathbf{y}(k) = d(k+97)$, 其中 $\{d(k), k = 1, \dots, 696\}$ 为已知的话务序列.利用 FESNs 建立 $\mathbf{u}(k)$ 和 $\mathbf{y}(k)$ 之间的映射关系.预测过程实际上就是利用上述映射关系,根据已知的序列值计算新的序列值的过程.实验采用 4 条模糊规则,其他储备池参数设置如下:储备池规模 $N = 30$, 内部连接权矩阵谱半径 $SR = 0.65$, 输入尺度 $IS = 0.01$, 实验中选取了 leaky^[16] 和双曲正切两

种类型的激活函数进行了实验.选择数据特点不同的 131 个话务小区进行实验.采用的误差评价指标为

$$\epsilon = \sqrt{\frac{1}{(p-l)\sigma^2} \times \sum_{n=l+1}^p (\hat{y}(n) - y(n))^2}$$

其中 σ 表示样本方差.

4.2.3 实验结果分析

图 6 给出了某话务小区话务量预测结果,显示出 FESN 对话务量预测有效性.实验针对具有不同特点的 131 个小区展开,训练和测试误差的平均值为 0.253 和 0.3201,可以满足移动通信对话务量建模和预测的精度要求,方差分别为 0.0066 和 0.0202,说明 FESN 方法具有一定的适应能力.

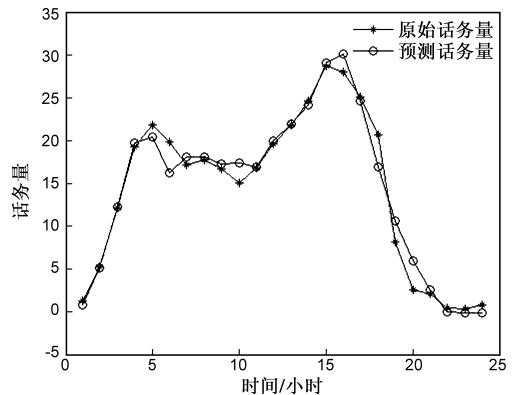


图6 某小区话务量预测结果

5 结论

针对经典 TS 模型采用线性函数作为模糊规则后件导致其非线性映射能力不强,以及基于梯度下降模糊神经网络训练效率低、容易陷入局部最优的问题,本文采用回声状态网络作为模糊规则的后件,提出了模糊回声状态网络. FESNs 通过引入储备池,将输入从低维输入空间映射到高维的状态空间,有效地提高了模型的非线性映射能力,同时,利用储备池内在的递归连接关系,在模型中引入了局部反馈,从而进一步提高了处理与时序相关的动态系统的建模能力.

另一方面,针对采用递归神经网络作为模糊规则后件,模型训练过程中计算量过大、容易陷入局部最优等问题,提出了通过求解线性回归问题确定 FESNs 参数的方法.

最后,通过仿真实验和移动通信话务时间序列预测问题中的应用,验证了 FESNs 方法的基本特性和实际应用价值,后续研究工作还将对模型参数的选取原则等问题做深入探讨.

FESNs 的应用领域还可进一步扩展至复杂非线性动态系统的自动控制、模式识别、故障诊断等更广泛的应用领域.

参考文献

- [1] T Takagi, M Sugeno. Fuzzy identification of systems and its applications to modeling and control[J]. IEEE Transactions on System, Man, Cybernetics, 1985, 15(1): 116 – 132.
- [2] T G Barbounis, J B Theocharis. A locally recurrent fuzzy neural network with application to the wind speed prediction using spatial correlation[J]. Neurocomputing, 2007, 70(7 – 9): 1525 – 1542.
- [3] Daniel Graves, Witold Pedrycz. Fuzzy prediction architecture using recurrent neural networks[J]. Neurocomputing, 2009, 72(7 – 9): 1668 – 1678.
- [4] Y Gao, M J Er. NARMAX time series model prediction: Feed-forward and recurrent fuzzy neural network approaches[J]. Fuzzy Sets and System, 2005, 150(2): 331 – 350.
- [5] Ken-ichi Funahashi, Yuichi Nakamura. Approximation of dynamical systems by continuous time recurrent neural networks[J]. Neural Networks, 1993, 6(6): 801 – 806.
- [6] Chia-Feng Juang. A TSK-type recurrent fuzzy network for dynamic systems processing by neural network and genetic algorithms[J]. IEEE Transactions on Fuzzy Systems, 2002, 10(2): 155 – 170.
- [7] Shang Ming Zhou, Li Da Xu. A new type of recurrent fuzzy neural network for modeling dynamics systems[J]. Knowledge-Based Systems, 2001, 14(5 – 6): 243 – 251.
- [8] H Jaeger. The “Echo State” Approach to Analyzing and Training Recurrent Neural Networks[R]. GMD Report 148, GMD – German National Research Institute for Computer Science, 2001.
- [9] M Lukosevicius, H Jaeger. Reservoir computing approaches to recurrent neural network training[J]. Computer Science Reviews, 2009, 3(3): 127 – 149.
- [10] H Jaeger. Tutorial on Training Recurrent Neural Networks, Covering BPTT, RTRL, EKF and the “Echo State Network” Approach[R]. GMD Report 159, German National Research Center for Information Technology, 2002.
- [11] Yanbo Xue, Le Yang, Simon Haykin. Decoupled echo state networks with lateral inhibition[J]. Neural Networks, 2007, 20(3): 365 – 376.
- [12] M Sugeno, G T Kang. Structure identification of fuzzy model[J]. Fuzzy Sets and Systems, 1986, 28(1): 15 – 33.
- [13] M Sugeno. Fuzzy Control[M]. Tokyo: Nikan Kougyou Shinbunsha Publisher, 1988.
- [14] C Bezdek. Pattern Recognition with Fuzzy Objective Function Algorithms[M]. PJlenum, New Youk, 1981.
- [15] Y Zhang, W Li. Gegenbauer orthogonal basis neural network and its weights-direct-determination methods[J]. Electronics Letters, 2009, 45(23): 1184 – 1185.
- [16] H Jaeger, M Lukosevicius, D Popovici, Udo Siewert. Optimization and applications of echo state networks with leaky-integrator neurons[J]. Neural Networks, 2007, 20(3): 335 – 352.
- [17] 孙圣和, 赵保军. 一种模糊人工神经网络控制器的设计与研究[J]. 电子学报, 1996, 24(10): 42 – 45.
Sun Shenghe, Zhao Baojun. Designing and researching for a FANNC[J]. Acta Electronica Sinica, 1996, 24(10): 42 – 45. (in Chinese)

作者简介



彭宇 男, 教授, 博士生导师, 1973年6月生于陕西西安, 哈尔滨工业大学自动化测试与控制研究所副所长, 主要研究方向为测试诊断技术、无线传感器网络技术和数据挖掘技术等。
E-mail: pengyu@hit.edu.cn



王建民 男, 1976年生于吉林长春, 现为哈尔滨工业大学自动化测试与控制系博士研究生, 主要研究方向为时间序列分析、数据挖掘、故障诊断。
E-mail: wjmfuzzy@126.com