

基于小波神经网络的自适应控制器设计

蔡吉刚（控制理论与控制工程）

指导教师：李树荣（教授）

摘要

小波神经网络是基于小波分析理论和神经网络理论建立起来的一种网络体系结构。小波变换具有良好的时频局部特性，而神经网络则具有自学习功能和良好的容错能力，小波神经网络由于较好的结合了两者的优点而具有强大的优势。本论文主要研究了非线性系统基于小波神经网络的控制器的设计问题，研究的主要内容可概括如下：

首先，针对一类未知非线性系统，设计了一种基于小波神经网络的模型参考自适应控制器，并提出了一种新的在线学习参数的混合训练算法。通过离线学习系统的特性，得到小波神经网络控制器的初始参数，然后使用混合训练算法在线修正控制律，实现自适应控制。仿真结果验证了控制方案的有效性，并通过与常规 BP 神经网络的对比，说明了小波神经网络作为控制器的优越性。

接着，对一类单输入单输出（SISO）仿射非线性系统，利用反馈线性化理论，设计了一种基于小波神经网络的自校正控制器。该方法用小波神经网络逼近系统的不确定性，网络的权值变化律通过构造 Lyapunov 函数并经过严格逐步推导得出，保证了所设计的控制器闭环系统的稳定性，并考虑了在系统状态未知时，通过设计高增益观测器以实现系统的输出自适应控制。通过对一类化工过程（CSTR）的仿真实例，说明了所设计控制器的有效性。

最后，考虑了一类多输入多输出（MIMO）非线性系统的自适应

控制问题，设计了基于小波神经网络的自适应控制器。由 Lyapunov 稳定性理论得出权值在线调整规律，并证明了闭环系统是一致最终有界（UUB）的。通过仿真，验证了所设计控制器的有效性。

关键词：小波神经网络，自适应控制，在线学习算法，仿射非线性系统，反馈线性化

Design of Adaptive Controller Based on Wavelet Neural Network

CAI Ji-gang (Control Theory and Control Engineering)

Directed by Professor LI Shu-rong

Abstract

Wavelet neural network (WNN) is built based on wavelet transform theory and neural network theory. Wavelet transform has excellent time frequency characteristic, while neural network possesses self-learning and fault tolerant ability. This paper is to study the design problem of wavelet neural network controller.

First, for a class of unknown nonlinear system, a model reference adaptive controller is designed based on wavelet neural network and a new online parameter training algorithm is presented. Offline system learning is used to attain the initial parameters of WNN and mixed training algorithm is to modify online control law for adaptive control. The simulation results verify the effectiveness of the controller and super to the traditional back propagation (BP) neural network.

Then, for a class of single input and single output (SISO) affine nonlinear system, a WNN adaptive controller is designed based on feedback linearization method. This method uses WNN to approximate system uncertainty and strictly deduces the network weight tuning law by constructing Lyapunov function in order to make sure the stability of closed system. When system state is unknown, high gain observer is

applied for output adaptive control. The simulation results on a continuous stirred-tank reactor (CSTR) system show that the design is effective.

At last, WNN based adaptive control is design for a class of multiple inputs and multiple outputs (MIMO) nonlinear system. From Lyapunov stability theory the weight online tuning law is obtained and it is proved that the closed system is uniformly ultimately bounded (UUB). The simulation results verify the effectiveness of the controller.

Key words: Wavelet neural network, Adaptive control, Online training algorithm, Affine nonlinear system, Feedback linearization

独创性声明

本人声明所呈交的论文是我个人在导师指导下进行的研究工作及取得的研究成果。尽我所知，除了文中特别加以标注和致谢的地方外，论文中不包含其他人已经发表或撰写过的研究成果，也不包含为获得中国石油大学或其它教育机构的学位或证书而使用过的材料。与我一同工作的同志对本研究所做的任何贡献均已在论文中作了明确的说明并表示了谢意。

签名： 蔡书刚 06年 5月 20日

关于论文使用授权的说明

本人完全了解中国石油大学有关保留、使用学位论文的规定，即：学校有权保留送交论文的复印件及电子版，允许论文被查阅和借阅；学校可以公布论文的全部或部分内容，可以采用影印、缩印或其他复制手段保存论文。

(保密论文在解密后应遵守此规定)

学生签名： 蔡书刚 06年 5月 20日
导师签名： 李树华 06年 5月 20日

第1章 前言

控制科学是 20 世纪中形成和发展起来的新学科，其应用已遍及众多的技术和社会领域，其发展经历了从古典控制理论到现代控制理论，再从现代控制理论到现在的智能控制和大系统控制诸阶段。

经典控制理论所涉及的对象是线性单回路系统，系统的数学模型比较单纯。现代控制理论的研究主要以状态空间为基础，所研究的系统推广到多输入多输出系统，同时既可以是线性的、定常的，也可以是简单非线性的、时变的。

然而，现代科学技术的发展和重大进步，对控制科学提出了更新、更高的要求，传统控制理论在实际应用中遇到了很多的困难，这包括：(1) 传统控制系统的设计和分析是建立在已知系统精确数学模型基础上的，而实际系统由于存在复杂性、非线性、时变性、不确定性和不完全性等因素，一般难以获得精确的数学模型；(2) 研究这些系统时，必须提出并遵循一些比较苛刻的假设，而这些假设在应用中往往与实际不相吻合；(3) 对某些复杂和包含不确定性的对象，根本无法以传统的数学模型来表示，即无法解决模型问题；(4) 为了提高性能，传统控制系统可能变得很复杂，降低了系统的可靠性。面对上述困难，控制工作者先后开辟了三条途径：系统辨识、自适应控制、智能控制。

自适应控制^[1]理论研究的目的是使控制器能适应被控对象参数缓慢变化或对象特性难以确知的情况，自适应控制能够根据输入输出数据，在线的调整控制器的参数。传统的自适应控制理论从一开始就有如下两个缺点：在理论上，只能分析参数的慢时变过程；在应用对象上，仅在线性系统自适应控制上取得了较大进展。对于非线性系统，传统方法几乎无能为力，人们正期望着非线性及时变的控制理论能有重大突破。

面对以上问题，自动控制发展的一个新方向就是智能控制。目前，

智能控制的理论体系尚未建立,但是神经网络控制、模糊控制、专家系统、遗传算法已成为公认的智能控制方法,并已取得了很多成功的应用^[2~4]。将人工智能引入控制系统中,把控制理论的分析与人工智能的灵活结合起来并采用先进的计算机技术,这就诞生了智能控制。智能控制的研究和发展表明把人工智能和反馈控制理论的结合对解决某些复杂系统的控制问题是有效的,但还有待突破性的进展。

现在,人们对于神经网络控制已经开展了广泛的研究^[5~12]。神经网络的本质非线性、逼近性、学习性和适应性等特点,使得神经网络用于非线性系统控制有着独特的优势,这也为长期困扰人们的非线性自适应控制的诸多问题的解决带来了希望。从目前看,神经网络用于控制系统设计主要是针对系统的非线性和不确定性进行的,它不需要被控对象的数学模型,只需对神经网络进行在线或离线训练,然后进行控制系统设计。神经网络由于在结构选择上的灵活性,克服了经典自适应控制基于数学模型的方法。因此,神经网络自适应控制器的研究为非线性系统控制提供了广阔的前景。

近年来,大批学者对神经网络控制进行了深入的研究,基于神经网络的自适应控制迅速发展并出现了丰富的成果,利用 Lyapunov 稳定理论探讨神经网络自适应控制稳定性的成果也不断涌现^[9~13]。

但是,传统的 BP 等神经网络由于收敛速度慢,并且容易陷入局部极小点,所以人们一直致力于研究和寻求好的网络结构和学习算法。小波神经网络^{[14][15]}是基于小波分析理论和神经网络理论建立起来的网络结构,由于它引入了小波函数作为隐节点函数,使得其具有良好的时频局部化特性,因此无论在函数的逼近能力上,还是收敛速度和避免局部最优问题上都优于其它网络。

如今,对于小波神经网络控制的研究已经成为广大学者的研究热点^[16~26]。控制界权威 Astrom 和 Benveniste 甚至认为小波逼近可能成为解决非线性系统建模、辨识与控制器设计的最新非参数方法^[19]。目前,小波神经网络用于控制系统的研究已经在系统辨识与建模^[20]、自适应控制^{[21][22]}、预测控制^{[23][24]}、故障检测与诊断^[25]等方面取得很多的成

果。

小波分析为神经网络的研究应用提供了坚实的理论基础,但小波神经网络的理论研究毕竟刚刚起步,迄今还存在许多有待解决的问题:小波基的选取问题;小波神经网络的收敛性,鲁棒性,计算复杂度等问题;在多维输入情况下,如何有效的利用小波分析理论及其相关理论来优化网络结构等。

同时,对于自适应控制系统的稳定性、鲁棒性的研究,始终是一个具有挑战性的课题。如何把小波神经网络用于控制中,同时保证系统的稳定性及其它性能指标等问题,对神经网络控制也是一个挑战,如何设计和证明稳定的自适应神经网络控制器也是理论上的难题。所以,设计稳定的小波神经网络控制器具有一定的理论价值和实际意义。

本论文主要致力于研究基于小波神经网络的自适应控制器的设计,所做的主要工作如下:

首先,给出论文中用到的一些数学基础知识、非线性控制理论的基本概念及定理和小波神经网络的结构、算法及相关理论。

接着,提出了一种改进的小波神经网络参数训练算法,即采用带有动量项的梯度下降法和递推最小二乘法分别对参数进行在线学习,以提高学习速度,并针对一类非线性自回归滑动平均(NARMAX)模型,设计了基于小波神经网络的模型参考自适应控制器。仿真结果验证了此控制方案的有效性,并通过对比说明了该算法优于常规BP神经网络控制。

然后,对于一类SISO仿射非线性系统,基于小波神经网络,利用反馈线性化的设计方法,先设计了一种新的神经网络自适应控制器,然后基于Lyapunov稳定性理论证明了系统的一致最终有界性。并考虑了在系统状态未知时,通过设计高增益观测器,实现系统的输出控制问题。给出了一类化工过程CSTR的跟踪实例,验证了所设计控制器的适应性和鲁棒性。

最后,对一类MIMO非线性系统进行了研究,设计了基于小波

神经网络的自适应控制器，通过 Lyapunov 理论证明了闭环系统是一致最终有界的。通过仿真，表明了所设计的控制器是有效的。

总之，小波神经网络自适应控制的设计和实现是一个十分有意义的研究课题。对于小波神经网络的结构选择，自适应控制方案的选取，和对整个控制系统的闭环稳定性、鲁棒性的研究也一直是人们所关注的热点所在。

第2章 预备知识

本章介绍了与本论文有关的基础知识, 共分三部分内容:

第一部分介绍了本论文中所用到的一些基本的数学知识; 第二部分给出了非线性系统控制中的有关理论和常用的一些结论; 第三部分介绍小波神经网络的相关理论。由于很多定理、结论都是比较容易得到的, 就省略了相关的证明, 并且不特别指出出处。

2.1 数学基础

定义 2.1.1: 开集、列紧集和紧集

称子集 $S \subset R^n$ 是开的, 如果对于每一个 $x \in S$, 都存在 x 的 ε 邻域:

$N(x, \varepsilon) = \{Z \in R^n \mid \|Z - x\| < \varepsilon\}$, 使得: $N(x, \varepsilon) \subset S$; 称一个集合是闭

的, 当且仅当其在 R^n 中的补集是开集; 是凸的, 如果对于任意 $x, y \in S$

和 θ , $0 < \theta < 1$, 使得点 $\theta x + (1 - \theta)x \in S$ 。

设有度量空间 X , E 是 X 的子集。如果 E 中的每个序列都含有收敛的子列, 且收敛于 E 中, 则称 E 是 X 的列紧集。若有开集族 $\{G_\lambda, \lambda \in \Lambda\} \supset E$, 则称该开集族是 E 的开覆盖, 若 E 的任何开覆盖, 均存在 E 的有穷子覆盖则称 E 是 X 的紧集。简单地说, 一个集合如果它是有界的闭集则它是紧的。

定义 2.1.2: 对于 n 维向量 x 的 p 范数记作 $\|x\|_p$, 且有:

$$\|x\|_p = \left\{ \sum_{i=1}^n |x_i|^p \right\}^{1/p}, \quad x_i \text{ 指 } x \text{ 的第 } i \text{ 个分量。常用的向量范数有:}$$

$$(1) \quad p=1, \quad \|x\|_1 = |x_1| + |x_2| + \cdots + |x_n|$$

(2) $p = 2$, $\|x\|_2$ 为欧氏范数

(3) $\|x\|_\infty = \lim_{p \rightarrow \infty} \|x\|_p = \max_i |x_i|$

矩阵 $A = (a_{ij})_{m \times n} \in C^{m \times n}$ 的诱导 p -范数为: $\|A\|_p = \sup_{\|x\|_p=1} \|Ax\|_p$, 常用

的诱导范数分别为: $p = 1, 2, \infty$

$$\|A\|_1 = \max_j \sum_{i=1}^m |a_{ij}| \quad (\text{列的绝对值之和最大值})$$

$$\|A\|_2 = \max_i \sqrt{\lambda_i(A^H A)} \quad (A^H A \text{ 最大特征根的平方根})$$

$$\|A\|_\infty = \max_i \sum_{j=1}^n |a_{ij}| \quad (\text{行的绝对值之和最大值})$$

矩阵 A 的 Frobenius 范数, 简称 F -范数, 又称欧氏范数, 如下定义:

$$\|A\|_F^2 = \sum_i \sum_j a_{ij}^2 = \text{tr}(A^T A)$$

这里 $\text{tr}(\cdot)$ 表示矩阵的迹(对角元之和), F -范数虽不是诱导范数, 但是它和向量的 2-范数是相容的, 即:

$$\|Ax\|_2 \leq \|A\|_F \|x\|_2$$

定义 2.1.3: 对于时间函数 $u(t)$ 的范数, 其 L_p 范数定义为:

$$\|u\|_p = \left(\int |u(t)|^p d\tau \right)^{1/p}, \quad p \in [1, \infty)$$

(1) $p = 1$, $\|u(t)\|_1 = \int |u(t)| d\tau$, 若 $\int |u| d\tau < \infty$, 则称 $u(t)$ 的 L_1 范数存在, 记为: $u(t) \in L_1$, 或称 $u(t)$ 是绝对可积的。

(2) $p = 2$, $\|u(t)\|_2 = \left(\int |u(t)|^2 d\tau \right)^{1/2}$, 若 $\|u(t)\|_2 < \infty$, 则称 $u(t)$ 的

L_2 范数存在，记为： $u(t) \in L_2$ ，或称 $u(t)$ 是平方可积的。

(3) $p = \infty$ ， $\|u(t)\|_\infty = \sup_{t>0} |u(t)|$ ，若 $\|u(t)\|_\infty < \infty$ ，则称 $u(t)$ 的 L_∞ 范

数存在，并记为： $u(t) \in L_\infty$ ，或称 $u(t)$ 是有界的。

对于 $\|u\|_p$ ，具有性质：若 $\|u(t)\|_1 < \infty$ 且 $\|u(t)\|_\infty < \infty$ ，则有 $\|u(t)\|_2 < \infty$ ，

简单的说：如果信号 $u(t)$ 是绝对可积的，而且是有界的，则必定是平方可积的。

定义 2.1.4：对于矩阵 $A \in R^{n \times n}$ ，称 A 为半正定矩阵，并记为 $A \geq 0$ ，

若对于任意的 n 维向量 $x \in R^n$ 都有 $x^T A x \geq 0$ ；称 A 为正定矩阵，若对

某一标量 α 和任意的 n 维向量 $x \in R^n$ 都有 $x^T A x \geq \alpha x^T x \geq \alpha \|x\|^2$ ，并记为

$A > 0$ ；称 A 为对称半正定矩阵，若 $A \geq 0$ 且 $A = A^T$ ；称 A 为对称正定

矩阵，若 $A > 0$ 且 $A = A^T$ 。对称半正定矩阵 A 的二次型可用下式估计：

$$\lambda_{\min}(A) \|x\|^2 \leq x^T A x \leq \lambda_{\max}(A) \|x\|^2 \quad (2.1.1)$$

其中 $\lambda_{\min}(A)$ ， $\lambda_{\max}(A)$ 分别表示 A 的最小、最大特征值。

定理 2.1.1：Barbalat's 定理

令 $f(t)$ 是一个可微的函数，如果 $\lim_{t \rightarrow \infty} f(t) = k < \infty$ 且 $\dot{f}(t)$ 是一致连续的，则：

$$\lim_{t \rightarrow \infty} \dot{f}(t) = 0$$

推论 1：如果 $f(t)$ 是一致连续的，且使得： $\lim_{t \rightarrow \infty} \int_0^t f(\tau) d\tau$ 存在且有限，则： $f(t) \rightarrow 0$ 当 $t \rightarrow \infty$ 。

推论 2：如果 $f(t)$ ， $\dot{f}(t) \in L_\infty$ ，且对于某个 $p \in [1, \infty]$ 有 $f(t) \in L_p$ ，

则：当 $t \rightarrow \infty$ 时 $f(t) \rightarrow 0$ 。

推论 3: 对于可微的函数 $f(t)$ ，如果 $\lim_{t \rightarrow \infty} f(t) = k < \infty$ 且 $\dot{f}(t)$ 存在，

则： $f(t) \rightarrow 0$ 当 $t \rightarrow \infty$ 。

定理 2.1.2: Bellman-Gronwall 定理

令： $x(\cdot), a(\cdot), u(\cdot): R^+ \rightarrow R^+$ ， $T \geq 0$ ，若：

$x(t) \leq \int_0^t a(\tau)x(\tau)d\tau + u(t)$ ， $\forall t \in [0, T]$ ，则有：

$$x(t) \leq \int_0^t a(\tau)u(\tau)e^{\int_{\tau}^t a(\sigma)d\sigma} d\tau + u(t), \forall t \in [0, T] \quad (2.1.2)$$

若 $u(\cdot)$ 是可微的，则：

$$x(t) \leq u(0)e^{\int_0^t a(\sigma)d\sigma} + \int_0^t \dot{u}(\tau)e^{\int_{\tau}^t a(\sigma)d\sigma} d\tau, \forall t \in [0, T] \quad (2.1.3)$$

在本论文中，除非特别说明，对向量都是使用 2-范数，对矩阵均使用 F -范数。

2.2 非线性控制系统的基础知识

定义 2.2.1: 一致最终有界 (Uniformly Ultimately Bounded, UUB)

对于系统 $\dot{x} = f(x)$ ， $x \in R^n, x(t_0) = x_0$ 称其解是 UUB 的，如果对于任意的 R^n 中的一个初始状态 x_0 ，存在一个常数 $\mu > 0$ 和数 $T(\mu, x_0)$ 使得 $\|x(t)\| < \mu$ 对于所有的 $t > t_0 + T(\cdot)$ 都成立。

定义 2.2.2: 仿射非线性系统

对于如下的非线性系统：

$$\text{状态方程: } \dot{x}(t) = f(x(t)) + \sum_{i=1}^m g_i(x(t))u_i(t) \quad (2.2.1)$$

$$\text{输出方程: } y(t) = h(x) \quad (2.2.2)$$

这里 $x \in R^n$ 为状态向量, $u_i (i=1, \dots, m)$ 为控制输入, $f(x(t))$ 和 $g_i(x(t))$ 为 n 维向量函数, $h(x)$ 为 m 维输出函数向量。从状态方程和输出方程可以看出, 该非线性系统对 $h(x)$ 是非线性的, 而对控制量 $u_i (i=1, \dots, m)$ 却是线性的, 称这种系统为仿射非线性系统。

定义 2.2.3: 李导数, 李括号

一个标量函数 $\lambda(x)$ 沿着一个向量场 $f(x) = [f_1, \dots, f_n] \in R^n$ 的方向导数定义为:

$$L_f \lambda(x) := \frac{\partial \lambda}{\partial x} f(x) = \sum_{i=1}^n \frac{\partial \lambda}{\partial x_i} f_i(x), \quad x \in R^n \quad (2.2.3)$$

称为函数 $\lambda(x)$ 沿向量场 $f(x)$ 的李导数, 或记为形式: $\langle d\lambda(x), f(x) \rangle$ 。

李导数可重复进行, 可采用如下记法:

$$L_g L_f \lambda(x) = \frac{\partial L_f \lambda(x)}{\partial x} g(x) \quad (2.2.4)$$

$$L_f^k \lambda(x) = L_f L_f^{k-1} \lambda(x) = \frac{\partial L_f^{k-1} \lambda(x)}{\partial x} f(x) \quad (2.2.5)$$

一个向量场 $g(x) \in R^n$ 沿着另外一个向量场 $f(x)$ 的方向导数为:

$$ad_f g = \frac{\partial g(x)}{\partial x} f - \frac{\partial f(x)}{\partial x} g \quad \text{或记为: } [f, g] \in R^n \quad (2.2.6)$$

称所得到的新的向量场定义为 $g(x)$ 对于 $f(x)$ 的李括号。李括号也可以重复计算, 记:

$$ad_f^k g(x) = ad_f ad_f^{k-1} g(x) = [f(x), ad_f^{k-1} g(x)] \quad (2.2.7)$$

定义 2.2.4: 相对阶

对于如下的一个单输入单输出 (SISO) 仿射非线性系统:

$$\begin{cases} \dot{x} = f(x) + g(x)u \\ y = h(x) \end{cases} \quad (2.2.8)$$

其中, $x \in R^n$, $f(x), g(x) \in R^n$ 均为光滑的向量函数, $h(x)$ 是一个光滑的标量函数。称上面的系统在 x_0 具有相对阶 r , 如果:

$$(i) \quad L_g L_f^k h(x) = 0, x \in U_{x_0}, k < r - 1 \quad (2.2.9)$$

$$(ii) \quad L_g L_f^{r-1} h(x_0) \neq 0 \quad (2.2.10)$$

设系统在某个时间 t_0 的状态为 $x(t_0) = x_0$, 将输出 $y(t) = h(x(t))$ 对时间求微分得:

$$\dot{y} = \frac{\partial h}{\partial x} \frac{\partial x}{\partial t} = \frac{\partial h}{\partial x} (f(x) + g(x)u) = L_f h(x) + L_g h(x)u$$

若系统的相对阶 r 大于 1, 则有 $L_g h(x) = 0$, 因此有 $\dot{y} = L_f h(x)$, 对

$$\text{有 } k < r: \frac{d^k y}{dt^k} = L_f^k h(x), \text{ 而 } \left. \frac{d^r y(t)}{dt^r} \right|_{t=t_0} = L_f^r h(x_0) + L_g L_f^{r-1} h(x_0)u(t_0).$$

故系统的相对阶 r 恰是保证 u 出现而需要在 $t = t_0$ 微分 $y(t)$ 的次数。

同样对于如下的多输入多输出 (MIMO) 非线性系统:

$$\begin{cases} \dot{x} = f(x) + \sum_{i=1}^m g_i(x)u_i \\ [y_1, \dots, y_m] = [h_1(x), \dots, h_m(x)] \end{cases} \quad (2.2.11)$$

称该系统在 x_0 处具有相对阶 $\{r_1, \dots, r_m\}$, 如果:

$$(i) \quad L_{g_j} L_f^k h_j(x) = 0, \quad 1 \leq j \leq m, k < r_j - 1 \quad (2.2.12)$$

$$(ii) \text{ 矩阵 } A(x_0) = \begin{bmatrix} L_{g_1} L_f^{n-1} h_1(x_0) & \cdots & L_{g_m} L_f^{n-1} h_1(x_0) \\ \vdots & \ddots & \vdots \\ L_{g_1} L_f^{r-1} h_m(x_0) & \cdots & L_{g_m} L_f^{r-1} h_m(x_0) \end{bmatrix} \quad (2.2.13)$$

是非奇异的。

定义 2.2.5: SISO 非线性系统精确反馈线性化

考虑系统(2.2.8), 对于输出 $y = h(x)$, 假定系统的相对阶 $r = n$,

则有:

$$\begin{aligned} \dot{y} &= L_f h(x) \\ &\vdots \\ y^{(n-1)} &= L_f^{n-1} h(x) \\ y^{(n)} &= L_f^n h(x) + L_g L_f^{n-1} h(x) u \end{aligned} \quad (2.2.14)$$

记 $z = (y, \dots, y^{(n-1)})^T$, 则 $z = \Phi(x)$, 若 $\frac{\partial \Phi}{\partial x}$ 非奇异, 则称 $z = \Phi(x)$ 构成

一坐标变换; 若 $\frac{\partial \Phi}{\partial x} \Big|_{x=x_0}$ 非奇异, 则称 $z = \Phi(x)$ 构成一局部坐标变换。

则在新坐标下, 系统 (2.2.8) 可变为如下标准型:

$$\begin{cases} \dot{z}_i = z_{i+1}, 1 \leq i \leq n-1 \\ \dot{z}_n = L_f^n h(x) + L_g L_f^{n-1} h(x) u = b(z) + a(z) u \\ y = z_1 \end{cases} \quad (2.2.15)$$

若采用以下控制:

$$u = -\frac{b(z)}{a(z)} + \frac{v}{a(z)} = \frac{(-L_f^n h(x) - \alpha_{n-1} L_f^{n-1} h(x) - \cdots - \alpha_0 h(x) + y_d)}{L_g L_f^{n-1} h(x)} \quad (2.2.16)$$

y_d 是设计的输入信号, 则在坐标变换 z 下系统成为一能控能观系统:

$$\begin{cases} \dot{z}_i = z_{i+1}, 1 \leq i \leq n-1 \\ \dot{z}_n = -\alpha_0 z_1 - \cdots - \alpha_{n-1} z_n + y_d \\ y = z_1 \end{cases} \quad (2.2.17)$$

若在新系统中选择 α_i 使得多项式 $\Gamma(s) = s^n + \alpha_{n-1}s^{n-1} + \dots + \alpha_0$ 是 Hurwitz 的, 就可以使得系统对 y_d 是渐近跟踪的, 且通过选择合适的 α_i 可以任意配置系统的极点。如何选择 $h(x)$ 满足相对阶为 n 呢, 由相对阶的定义可知 $h(x)$ 应满足如下方程:

$$\frac{\partial h(x)}{\partial x} [g(x), ad_f g(x), \dots, ad_f^{n-2} g(x)] = 0 \quad (2.2.18)$$

定理 2.2.1: Frobenius 定理

方程 (2.2.18) 有解的充分必要条件是:

- (i) n 个矢量 $g(x_0), ad_f g(x_0), \dots, ad_f^{n-2} g(x_0)$ 是线性无关的;
- (ii) 分布 $\{g(x), ad_f g(x), \dots, ad_f^{n-2} g(x)\}$ 在 x_0 是对合的。

上面假设相对阶为 $r = n$, 当 $y = h(x)$ 不满足 (2.2.18) 时, 即 $r < n$, 则在 x_0 附近可找到一个局部的坐标变换:

$$Z = \begin{pmatrix} z_1 \\ z_2 \\ \vdots \\ z_n \end{pmatrix} = \Phi(x) = \begin{pmatrix} h(x) \\ \vdots \\ L_f^{r-1} h(x) \\ \Phi_{r+1}(x) \\ \vdots \\ \Phi_n(x) \end{pmatrix}, \quad L_g \Phi_i = 0, r+1 \leq i \leq n \quad (2.2.19)$$

此时在新坐标下, 系统 (2.2.8) 变为:

$$\begin{cases} \dot{z}_i = z_{i+1}, 1 \leq i \leq r-1 \\ \dot{z}_r = L_f^r h(x) + L_g L_f^{r-1} h(x)u = b(z) + a(z)u \\ y = z_1 \end{cases} \quad (2.2.20)$$

$$(\dot{z}_{r+1}, \dots, \dot{z}_n)^T = q(z) \quad (2.2.21)$$

这里光滑的标量函数 $\Phi_{r+1}, \dots, \Phi_n$ 应使 $\Phi(x)$ 为一局部的坐标变换。

令:

$$u = -\frac{b(z)}{a(z)} + \frac{v}{a(z)} \quad (2.2.22)$$

可以设计 v 使得系统是输入输出稳定的。如何保证系统的内部稳定性呢? 令 $\xi = (z_1, \dots, z_r)^T$, $\eta = (z_{r+1}, \dots, z_n)^T$ 则:

$$\dot{\xi} = A\xi + Bv, \quad y = (1, 0, \dots, 0)\xi \quad (2.2.23)$$

$$\dot{\eta} = q(\xi, \eta) \quad (2.2.24)$$

易知子系统 ξ 是能控的, 则闭环系统的稳定性取决于子系统 η 。 $\dot{\eta} = q(0, \eta)$ 是系统约束在 $y(t) = 0$ 上时的一个内部动力系统, 因此称之为系统(2.2.8)的一个零动力系统。

定理 2.2.2: 对于系统(2.2.8), 若在系统的某个平衡点处其相对阶 $r < n$, 其零动态系统 $\dot{\eta} = q(0, \eta)$ 是(局部)渐近稳定的, 那么令:

$$u = -\frac{b(z)}{a(z)} + \frac{v}{a(z)} = \frac{(-L_f^r h(x) - \alpha_{r-1} L_f^{r-1} h(x) - \dots - \alpha_0 h(x))}{L_g L_f^{r-1} h(x)} \quad (2.2.25)$$

其中 α_i , ($i = 0, \dots, r-1$) 为对应 r 阶 Hurwitz 多项式系数, 那么整个闭环系统是(局部)渐近稳定的。

定义 2.2.6: MIMO 非线性系统的线性化

对于系统(2.2.11) 如下选择坐标系:

$$\begin{aligned} z_1^1 &= h_1(x), z_2^1 = L_f h(x), \dots, z_n^1 = L_f^{n-1} h(x) \\ z_1^2 &= h_1(x), z_2^2 = L_f h(x), \dots, z_n^2 = L_f^{n-1} h(x) \\ &\dots \end{aligned} \quad (2.2.26)$$

令 $r_1 + r_2 + \dots + r_m = r \leq n$, 若 $r \neq n$ 我们可以选择 $n-r$ 个函数 $\eta(x), \dots, \eta_{n-r}$ 使得它们和(2.2.19)一起构成一个坐标变换(微分同胚)。则在新坐标下:

$$\begin{cases} \dot{z}_j^i = z_{j+1}^i, & 1 \leq j \leq r_i - 1 \\ \dot{z}_{r_i}^i = b^i(z, \eta) + \sum_{k=1}^m a_k^i(z, \eta) u_k \\ y_i = z_1^i \end{cases} \quad (2.2.27)$$

$\eta \in R^{n-r}$ ，且由下列方程定义：

$$\dot{\eta} = q(z, \eta) + \sum_{j=1}^m p_j^i(z, \eta) u_j \quad (2.2.28)$$

解如下的代数方程：

$$\begin{bmatrix} \dot{z}_{r_1}^1 \\ \vdots \\ \dot{z}_{r_m}^m \end{bmatrix} = \begin{bmatrix} b_1(z) \\ \vdots \\ b_m(z) \end{bmatrix} + \begin{bmatrix} a_1^1(z) & a_2^1(z) & \cdots & a_m^1(z) \\ \vdots & \vdots & \ddots & \vdots \\ a_1^m(z) & a_2^m(z) & \cdots & a_m^m(z) \end{bmatrix} \begin{bmatrix} u_1 \\ \vdots \\ u_m \end{bmatrix} = \begin{bmatrix} v_1 \\ \vdots \\ v_m \end{bmatrix} \quad (2.2.29)$$

由于 $A(z)$ 非奇异，从而可解出

$$u = A^{-1}(z)(-b(z) + v) \quad (2.2.30)$$

同理，若系统零动态 $\dot{\eta} = q(0, \eta) + p(0, \eta)[-A^{-1}(0, \eta)b(0, \eta)]$ 是渐近稳定的，那么只要令： $v_i = (s^{r_i} + c_{r_i-1}s^{r_i-1} + \cdots + c_1s + c_0)z_1^i, i = 1, \dots, m$ ，其中： $s^{r_i} + c_{r_i-1}s^{r_i-1} + \cdots + c_1s + c_0$ 是 Hurwitz 的，则所得闭环系统是渐近稳定的。

2.3 小波神经网络的模型结构及其算法

2.3.1 小波分析理论

小波分析是 80 年代中期发展起来的一门新技术，它是数学理论中调和和分析技术发展的最新成果，被公认为是工具和方法上的重大突

破。它在时域和频域同时具有良好的局部化性能，有一个灵活可变的时间-频率窗，这在理论和实用中都有重要的意义，已被成功应用于逼近理论、微分方程、分形式别、计算机视觉、非线性科学等方面。

本部分仅介绍与本文相关的小波理论^{[26][27]}知识。

定义 2.3.1: 引入函数 $\varphi(x) \in L^2(R)$ ，如果此函数满足容许条件：

$$C_\varphi = \int_{\mathbb{R}} \frac{|\hat{\varphi}(\omega)|^2}{|\omega|} d\omega < \infty \quad (2.3.1)$$

则称此函数为母小波或基小波。式中， $\hat{\varphi}(\omega)$ 为 $\varphi(x)$ 的 Fourier 变换。

由容许条件知，用做母小波的函数 $\varphi(x)$ 至少要满足 $\hat{\varphi}(\omega=0)=0$ ，即：

$$\int_{\mathbb{R}} \varphi(x) dx = 0 \quad (2.3.2)$$

也就是说 $\hat{\varphi}(\omega)$ 必须带有带通性质，且 $\varphi(x)$ 必须是正负交替的震荡波形，使得其平均值为0。小波函数波形如图 2-1。

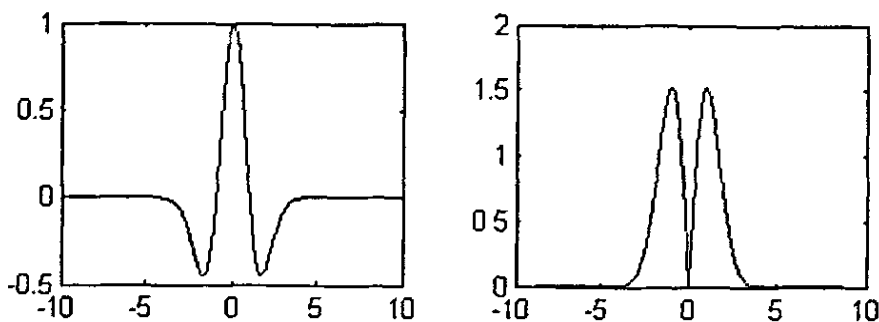


图 2-1 一小波函数波形及其傅立叶变换

母小波经平移和伸缩后产生的小波基定义如下：

$$\varphi_{a,b}(x) = \frac{1}{\sqrt{a}} \varphi\left(\frac{x-b}{a}\right) \quad (a, b \in R, a \neq 0) \quad (2.3.3)$$

a 称为伸缩因子或尺度因子， b 称为平移因子。在实际应用中，

往往要把上述的连续小波离散化，即令 $a = 2^n, b = m2^n$ ，则：

$$\varphi_{n,m}(x) = \frac{1}{\sqrt{2^n}} \varphi(2^{-n}x - m) \quad (n, m \in \mathbb{Z}) \quad (2.3.4)$$

定义 2.3.2: 对任意 $f(x) \in L^2(\mathbb{R})$, 其连续小波变换定义为:

$$(W_\varphi f)(a,b) = |a|^{-\frac{1}{2}} \int_{-\infty}^{\infty} f(x) \overline{\varphi\left(\frac{x-b}{a}\right)} dx \quad (2.3.5)$$

或写成内积形式:

$$(W_\varphi f) = \langle f, \varphi_{a,b} \rangle \quad (2.3.6)$$

函数 $f(x)$ 可以由它的小波变换进行重构, 重构公式为:

$$f(x) = \frac{1}{C_\varphi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (W_\varphi f)(a,b) \varphi_{a,b} \frac{da}{a^2} db \quad (2.3.7)$$

定义 2.3.3: 设函数 $\varphi: \mathbb{R}^n \mapsto \mathbb{R}$ 是一个母小波函数, 由 $\varphi(\cdot)$ 产生的一个可数函数族:

$$\Phi = \{ \det(D_k^{\frac{1}{2}}) \varphi[D_k x - t_k] : t_k \in \mathbb{R}^n, D_k = \text{diag}(d_k), d_k \in \mathbb{R}_+, k \in \mathbb{Z} \}$$

记 $\varphi_k = \det(D_k^{\frac{1}{2}}) \varphi[D_k x - t_k]$. 如果存在两个数 $c_{\min} > 0$ 和 $c_{\max} < \infty$,

使得对于任意的函数 $f \in L^2(\mathbb{R}^n)$ 都有

$$c_{\min} \|f\|^2 \leq \sum_{\varphi_k \in \Phi} |\langle \varphi_k, f \rangle|^2 \leq c_{\max} \|f\|^2 \quad (2.3.8)$$

则称该小波满足框架条件. 对于任意的函数 $f \in L^2(\mathbb{R}^n)$ 和 $\varepsilon > 0$ 必存在一个函数序列 $\{\varphi_k : \varphi_k \in \Phi, k=1,2,\dots,n\}$ 以及一组实数 $w_k, k=1,2,\dots,n$, 使得

$$\left\| f(x) - \sum_{k=1}^n w_k \varphi_k(x) \right\|_2^2 < \varepsilon \quad (2.3.9)$$

上式说明了在满足框架条件下, 小波族在 $L^2(\mathbb{R}^n)$ 上的逼近能力。

定义 2.3.4: 非平凡函数 $w(t)$ 被称为是一个窗函数, 如果 $w(t) \in L^2(\mathbb{R})$, 且 $tw(t) \in L^2(\mathbb{R})$ 。表征窗函数的两个参数是窗函数的中心和半径。

由于基小波函数 $\varphi(t)$ 及其 Fourier 变换 $\hat{\varphi}(\omega)$ 都有很快的衰减, 故可作为窗函数。它们在时间-频率域上定义了一个时频窗:

$$(b + at^* - a\Delta\varphi, b + at^* + a\Delta\varphi) \times \left(\frac{w^*}{a} - \frac{1}{a}\Delta\hat{\varphi}, \frac{w^*}{a} + \frac{1}{a}\Delta\hat{\varphi} \right) \quad (2.3.10)$$

其中心为 $(b + at^*, \frac{w^*}{a})$, 窗的高度(频窗)和宽度(时窗)分别为 $2\frac{1}{a}\Delta\hat{\varphi}, 2a\Delta\varphi$ 。窗函数决定的窗口是对信号 $f(t)$ 局部性的一次刻画, 小波函数提供了信号 $f(t)$ 在时段 $(b + at^* - a\Delta\varphi, b + at^* + a\Delta\varphi)$ 和频段 $(\frac{w^*}{a} - \frac{1}{a}\Delta\hat{\varphi}, \frac{w^*}{a} + \frac{1}{a}\Delta\hat{\varphi})$ 时的“含量”。正是小波的这种变窗特性, 使它能够表示各种不同频率分量的信号, 特别是具有突变性质的信号。

2.3.2 小波神经网络概述

小波神经网络 (Wavelet Neural Network, WNN), 在有些文献中也称小波网络 (Wavelet Network, WN), 是小波分析理论与神经网络理论相结合的产物, 它起源于小波分解, 是近年来在小波分析研究获得突破的基础上提出的一种前馈型神经网络。

小波分析自 80 年代提出, 被公认为是傅立叶分析的突破性发展。小波变换具有时频局部特性和变焦特性, 而神经网络具有自学习、自适应、鲁棒性和推广能力, 如何把两者的优势结合起来, 一直是人们所关注的问题^{[28][29]}。一种方法是用小波分析作为神经网络的前置处理

阶段,为神经网络提供输入特征向量,以完成分类、函数逼近等功能;另一种即所谓的小波神经网络,把小波变换与神经网络有机结合起来,充分继承了二者的优点。

目前,对小波神经网络的研究主要集中在以下两个方面:

(1) 对小波神经网络自身的研究,包括小波神经网络的结构设计方法、逼近能力、学习算法及其收敛性、鲁棒性等方面的研究。

(2) 对小波神经网络的应用研究,如信号分析、图像处理、边缘检测、数据压缩、自动控制、模式识别、故障诊断等自然科学,以及金融、证券、股票、经济指标预测的社会、经济学科的诸多领域。

小波神经网络与其它前馈神经网络相比有以下特点:

(1) 小波神经网络可以看作是利用神经网络的拓扑结构来表达小波级数的表达式,也可看作是利用小波函数代替神经网络中神经元的激励函数而形成的网络结构;

(2) 小波神经网络属于局部逼近网络,而BP网络则是一种全局逼近网络。局部逼近网络与全局逼近网络相比,具有收敛速度快、易适应新数据、可以避免较大的外推误差等优点;

(3) 小波基元及整个网络结构的确定有可靠的理论根据,可避免BP网络结构设计的盲目性;

(4) 网络的权系数线性分布和学习目标函数的凸性,使网络训练过程从根本上避免了局部最优等非线性优化问题;

(5) 有较强的函数学习能力和推广能力。

小波神经网络的优良品质提高了系统的智能水平,使控制系统具有更快的速度和更强的鲁棒性,显示出它独特的优势。目前神经网络控制器设计方法可以分为以下几种:基于自适应方式的神经网络控制系统、由神经网络单独构成的控制系统、基于常规控制原理的神经网络控制、神经网络智能控制、神经网络优化控制。

由于小波神经网络的逼近特性只有在一定紧集上才可以保证,所以在使用小波神经网络控制器时,它有其自己特定的稳定性及鲁棒性问题。我们可以从以下的问题考虑:

(1) 参数漂移: 参数漂移是由于函数逼近误差的存在, 使网络的权值随着时间的增大而趋于无穷大, 这个问题已经通过采用修正自适应规则及引入目标算法得到解决, 或根据权值与输出的线性关系用线性方法求解。

(2) 控制的可行性: 现在, 大多数的神经网络控制是基于反馈线性化的, 对于一非线性系统 $\dot{x} = f(x) + g(x)u$, 其中 $f(x), g(x)$ 未知, 所设计的控制器形式一般为 $u = \hat{g}^{-1}(-\hat{f} + v)$, 要使控制器存在, 必须保证 \hat{g}^{-1} 是非奇异的。

(3) 初始条件和参考信号: 小波神经网络只有在一定紧集上才可以保证函数近似的精确性, 这就对系统状态的初始条件有了限制。如果初始状态不在集合中, 系统就可能不稳定。

(4) 设计参数的选取: 从应用的角度考虑, 一种好的控制器设计应该给出控制器参数的范围及这些参数的选择与系统鲁棒性的关系。

对应于用作控制器的小波神经网络, 我们在进行结构设计和网络训练时, 必须考虑以下几个问题^{[15][30~34]}:

(1) 隐层节点数的选择: 隐层节点个数的多少直接影响逼近精度和响应的速度, 但目前并没有严格的理论方法, 多采用折中的办法。常采用以下几种原则来选取:

① 先设较少的节点, 进行网络学习, 检查误差, 逐渐增加节点数, 直到误差不再明显减小;

② 先设较多的节点, 采用将网络权值考虑在内的误差代价函数进行训练, 训练结束后去掉连接权值太小的节点;

③ 根据逼近函数的时频空间区域来选择小波基个数。

(2) 网络参数的初始化: 对于神经网络的训练来说, 初始值的选取不当可能直接导致网络收敛很慢, 甚至会发散。对于小波神经网络的参数初始化尤为重要, 因为当选择很小的伸缩因子时, 会引起小波太局部化, 使得价值函数的梯度在感兴趣的区域非常的小。对小波神

经网络的初始化来说, 要充分考虑到网络的输入空间的区域特性。

(3) 基小波的选取: 根据不同的网络结构和不同的逼近函数, 小波函数的选取也不相同。常用的小波函数表达形式有如下几种:

① Haar 小波: 它是最简单的母小波,

$$\varphi(x) = \begin{cases} 1 & 0 \leq x < 0.5 \\ -1 & 0.5 \leq x < 1 \\ 0 & \text{其它} \end{cases}$$

Haar 小波是一种正交函数系, 有 $\int_{-\infty}^{\infty} \varphi(x)\varphi(x-n)dx = 0, n = 0, \pm 1, \pm 2, \dots$

② 表达式为 $\varphi(x) = -xe^{-x^2/2}$ 的小波

③ Morlet 小波: $\varphi(x) = \cos(1.75x)e^{-x^2/2}$

④ 合成小波: 由常规神经网络的 Sigmoid 函数经过平移和叠加来构成小波神经网络的激励函数, $\varphi(x) = s(x+2) - 2s(x) + s(x-2)$,

$$s(x) = (1 + e^{-x})^{-1}.$$

神经网络比较常用的有如下定理:

定理 2.3.1: 具有任意数目的隐节点的二层网络可以一致逼近欧氏空间 R^n 紧子集上的连续函数 $f \in C(R^n, R^m)$ 。

定理 2.3.2: 具有任意数目隐单元的三层网络可以一致逼近紧集上的连续函数或按 L_2 逼近紧集上的平方可积函数。

定理 2.3.3: 函数逼近原理

若 $f(x)$ 是定义在一紧集 Ω 上的连续函数, $f(W, x): R^l \times R^n \rightarrow R$ 是对 W 和 x 连续的一个逼近函数, 则函数近似问题就是对于某种机制(如距离 d), 确定一个最优的参数 W^* , 对于一个可接受的 ε 使得:

$$d(f(W^*, x), f(x)) \leq \varepsilon \tag{2.3.10}$$

2.3.3 小波神经网络的主要结构和算法

目前研究小波神经网络的文献中最广泛采用的一种结构形式如图 2-2 所示，它由输入层、输出层和一个隐层组成。

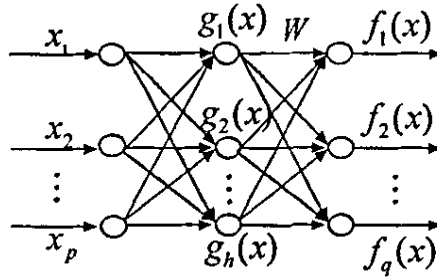


图 2-2 小波神经网络结构

根据基函数 $g_k(x)$ 和学习参数的选取不同，如上图所示的小波神经网络又可分为如下三种形式的网络：

(1) 连续小波神经网络^{[14][15]}：

$$f_i(x) = \sum_{j=1}^h W_{j,i} g_j(x) \quad (2.3.11)$$

其中， $g_j(x) = \prod_{i=1}^p \varphi\left(\frac{x_i - m_{i,j}}{d_{i,j}}\right) = \varphi(D_j x - \vec{m}_j)$ ， $1 \leq i \leq q, 1 \leq j \leq h$ ，

$D_j = \text{diag}(d_{1j}^{-1}, \dots, d_{pj}^{-1})$, $\vec{m}_j = [d_{1j}^{-1} m_{1j}, \dots, d_{pj}^{-1} m_{pj}]^T$ ，参数 m_{ij}, d_{ij} 在网络学习中与权值 W_{ij} 一起通过某种算法进行修正。

(2) 离散仿射小波神经网络^[35]：

$$f(x) = \sum W_{j,k} \varphi_{j,k} \quad (2.3.12)$$

其中， $g(x) = \prod_{i=1}^p g(x_i) = \prod_{i=1}^p \varphi(2^j x_i - k)$ 。根据函数 $f(x)$ 的时频特

性确定 j, k 的范围后，网络的可调参数只有权值，其与输出呈线性，

可通过最小二乘法或其他优化法修正权值, 使网络充分逼近 $f(x)$ 。

(3) 离散正交小波神经网络^[36]。

$$f(x) = \sum d_{j,k} \phi_{j,k}(x) + \sum C_{j,k} \varphi_{j,k}(x) \quad (2.3.13)$$

网络隐节点由小波节点 φ 和尺度节点 ϕ 构成。这种形式的小波网络的可行性基于 Daubechies 对紧支撑正交小波存在性及其构造的研究及 Mallat 的多分辨分析理论。

针对上面不同的网络模型, 人们提出了不同的网络训练算法。Zhang Qinghua^[14]给出了连续小波神经网络的概念和算法并提出用随机梯度下降法训练网络, 但这种算法与一般的反向传播算法类似。此学习算法利用了母小波函数导数, 对非显式小波网络无法学习。何正友^[14]等提出了把信号在小波基的各平移和伸缩上的展开同时输入到神经网络分类器的改进小波神经网络模型, 解决了具有非显式小波函数的网络学习问题。

钱峻^[32]把模型结构确定和参数估计方法与小波网络相结合, 给出了小波网络的结构确定和权系数估计算法, 并保证模型保持 BIC 准则下的最优结构, 后来又出现了正交搜索法等。沈雪勤等^[37]针对神经元个数太多、网络学习收敛速度较慢的问题, 在时频分析的基础上引入了能量密度的概念, 提出了一种新的小波神经网络模型。焦李成^[38]和李衍达^[39]等人研究了小波神经网络与模糊逻辑的联系, 用隶属函数表示权重值, 构造了模糊小波神经网络模型。

张邦礼等^[40]研究了用于非线性函数逼近的小波神经网络结构设计方法, 并分析了算法的收敛性, 在其后的论文中, 进一步讨论了仿射小波神经网络的逼近原理, 并针对小波神经网络中普遍存在的基元过多问题提出了网络结构设计的时(空)域“分解-综合”方法。

万建^[35]等提出了优化组合小波基元激励函数的算法, 对于小波基参数的选取和网络结构的确定给出了确定方法。其算法如下:

(1) 根据样本估计对象的输出信号 f 的时频集中区域。设学习样

本集 $l = \{(x_p, y_p), p = 1, 2, \dots\}$, $x_p \in A \subset [x_{\min}, x_{\max}]$, f 的频谱在 $\omega > 0$ 上的 ε 支撑为^[41]:

$$[\omega_{\min}, \omega_{\max}] = \min\{[\omega_1, \omega_2] / \int_{\omega \in (0, +\infty) \cap [\omega_1, \omega_2]} |\hat{f}(\omega)|^2 d\omega < \varepsilon \|f\|^2\}$$

则可得 f 的时频集中区域 $Q_\varepsilon(f)$ 。根据频谱的对称性, 仅考虑 $\omega > 0$ 的部分可得: $Q_\varepsilon^+(f) = [x_{\min}, x_{\max}] \times [\omega_{\min}, \omega_{\max}]$, 式中 ω_{\min} 和 ω_{\max} 可以利用样本集估计得到。

(2) 根据对象输出信号的频率集中区域确定小波基元个数。设选取的小波函数 $\varphi(x)$ 的正频率分布范围为 $[\omega_0, \omega_1]$, 则相应的小波基函数 $\varphi_{mn}(x) = 2^{n/2} \varphi(2^n x - m)$ 的正频率分布范围为 $[2^n \omega_0, 2^n \omega_1]$, 由此可以得到覆盖整个被逼近函数频率空间的 n_{\min} 和 n_{\max} 为:

$$n_{\min} = \text{floor}[\log_2 \omega_{\min} - \log_2 \omega_0], n_{\max} = \text{ceil}[\log_2 \omega_{\max} - \log_2 \omega_1]$$

其中, ceil 和 floor 分别为向 $-\infty$ 和 $+\infty$ 方向取整, 则小波基元的个数应选为 $N = n_{\max} - n_{\min} + 1$ 。

(3) 根据对象的输入信号集中区域确定小波基元。当选定了伸缩系数 n_i 后, 则通过合理选择平移系数使在每个分辨率 2^{n_i} 下所得到的的小波基函数序列能覆盖被逼近函数的整个时间集中区域。设选取的小波基函数 $\varphi(x)$ 的时间分布范围为 $[x_0, x_1]$ 。则相应小波基函数 $\varphi_{mn_i}(x) = 2^{n_i/2} \varphi(2^{n_i} x - m)$ 的时域范围为 $[2^{-n_i}(x_0 + m), 2^{-n_i}(x_1 + m)]$, 由此得到对每个 n_i 值为覆盖被逼近函数整个时间区域的平移系数 m 的取值范围 $m_{\min} = \text{floor}[2^{n_i} x_{\min} - x_0]$, $m_{\max} = \text{ceil}[2^{n_i} x_{\max} - x_1]$ 。

文献[42]给出了选择隐层节点数的一种方法，即使 Akaike's 的最终预测误差准则（FPE）达到最小。FPE 定义如下：

$$FPE(y) = \frac{1 + \frac{Q}{L}}{2 \left(1 - \frac{Q}{M}\right)} \sum_{t=1}^L (y(t) - \hat{y}(t))^2 \quad (2.3.14)$$

其中 M 是训练样本的长度， L 是隐层结点的个数， Q 是网络中需要调整的参数的个数， $Q = 2N \cdot L + L$ ，其中 N 为输入向量的维数。

文献[31]给出了小波神经网络的参数初始化方法，为了使小波伸缩系覆盖输入向量的整个范围，伸缩因子和平移因子必须满足下式：

$$b_j + a_j t^* - a_j \Delta\varphi = \sum_{i=1}^{N_j} x_{i,\min}, \quad b_j + a_j t^* + a_j \Delta\varphi = \sum_{i=1}^{N_j} x_{i,\max}$$

因此，可以得到：

$$\begin{cases} a_j = (\sum_{i=1}^{N_j} x_{i,\max} - \sum_{i=1}^{N_j} x_{i,\min}) / (2\Delta\varphi) \\ b_j = (\sum_{i=1}^{N_j} x_{i,\max} (\Delta\varphi - t^*) + \sum_{i=1}^{N_j} x_{i,\min} (\Delta\varphi + t^*)) / (2\Delta\varphi) \end{cases} \quad (2.3.15)$$

在本论文中，选用连续参数小波神经网络，选取的母小波函数为

$$\varphi(x) = -x \exp\left(-\frac{1}{2}x^2\right), \quad \text{其时域和频域窗口的中心和半径均为 } 0 \text{ 和}$$

1.224 744 87。

第3章 基于小波神经网络的模型参考自适应控制器设计

3.1 引言

对于一般的控制问题, 都可以归结为求解系统的输入, 使系统的输出在此输入下, 使得某一性能指标达到最小。对于跟踪控制问题可以归结为在 t 时刻, 求解控制律 $u(t)$, 使得在 $t+1$ 时刻系统的输出 $y_p(t+1) = y_m(t+1)$, 其中 $y_m(t+1)$ 为参考模型的输出。在实际应用中由于被控对象的表达式未知, 因此不能通过对象直接求取 u 。

面对上述问题, 一种解决的方法就是利用神经网络来逼近系统的非线性, 它不需要对被控对象进行建模, 而是通过离线、在线训练获得对象的控制律。由于小波神经网络具有良好非线性、学习性和自适应性, 我们可以设计基于小波神经网络的控制器, 先利用模型和对象的先验知识进行离线学习, 获得控制器的控制特性, 然后将控制器加到实际系统上, 通过在线学习, 实现系统的跟踪控制。

本章提出了一种改进的小波神经网络参数训练算法, 采用带有动量项的梯度下降法和递推最小二乘法分别对参数进行在线学习和计算, 并对于一类非线性自回归滑动平均 (Nonlinear Auto Regressive Moving Average model with eXogeneous inputs, NARMAX) 模型, 设计了基于小波神经网络的在线自适应控制器。仿真结果验证了此控制方案的有效性, 并通过与常规BP神经网络的对比, 说明了小波神经网络作为控制器的优越性。

3.2 问题的提出

非线性动态系统的 NARMAX 模型是 Billings 等于 1982 年提出的^[43], 并于 1985-1989 年基于 Nerode 的有限实现理论及泰勒定理对该模型描述一般非线性的普适性给予了证明^{[44][45]}。该模型概括了以前研究的几乎所有非线性模型, 如双线性模型、Hammerstein 模型、Wiener 模型、非线性时间序列模型、ARMAX 模型、输出仿射模型等, 并在实际控制过程中如化工、海洋、电力工程的获得了成功应用, 形成了国际上颇具影响的 Billings 学派。

许多非线性系统都可以通过一定的方式转化为此类模型的一些特例。在本章针对这一类问题, 我们将考虑设计一个稳定的控制器, 通过学习直接估计出系统的控制信号, 而不需要对系统进行辨识。

考虑如下 NARMAX 模型:

$$y_p(t+1) = f(y_p(t), y_p(t-1), \dots, y_p(t-n+1), u(t), u(t-1), \dots, u(t-m+1)) \quad (3.2.1)$$

其中, $y_p(t)$ 为输出, $u(t)$ 为输入, $f(\cdot)$ 为未知非线性函数, n 和 m 为系统的已知阶次。

假设参考模型为如下形式的线性方程:

$$y_m(t+1) = A_m y_m(t) + B_m u_m(t) \quad (3.2.2)$$

其中, $A_m(z) = a_0 + a_1 z^{-1} + \dots + a_{n-1} z^{n-1}$

$$B_m(z) = b_0 + b_1 z^{-1} + \dots + b_{m-1} z^{m-1}$$

u_m 和 y_m 分别为参考输入和参考输出。

控制的目的是: 在被控对象未知的情况下, 设计一个控制器, 通过学习直接估计动态系统所需的控制输入信号, 使得系统的输出能够跟踪参考模型的输出。

3.3 小波神经网络控制结构设计

由于被控对象的表达式未知，因此不能通过对象直接求取控制 u 。为了实现被控对象的输出对参考模型的跟踪控制，采用小波神经网络和参考模型相结合的控制方式，设计小波神经网络自适应控制器 (Wavelet Neural Network Controller, WNNC)，结构如图3-1所示。

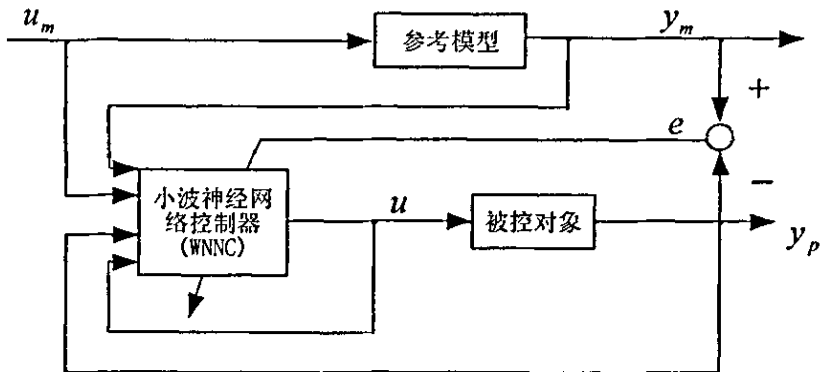


图3-1 基于小波神经网络的自适应控制系统结构

在图3-1中，小波神经网络控制器的输入 x 由四部分组成， $x = [u_m, y_m, u, y_p]^T$ ，如图3-2所示。

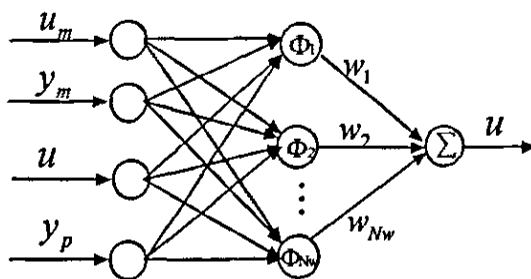


图 3-2 小波神经网络控制器结构

控制律 u 由小波神经网络控制器的输出来获得：

$$u = \sum_{j=1}^{N_w} w_j \Phi_j(x) \tag{3.3.1}$$

其中：

$$\Phi_j(x) = \prod_{k=1}^{N_j} \varphi(z_{jk}), \quad z_{jk} = \frac{x_k - m_{jk}}{d_{jk}} \quad (3.3.2)$$

x 为网络输入向量, N_i, N_w 分别为输入层和隐含层神经元的个数, w_j 为隐含层和输出层之间的连接权值, m_{jk}, d_{jk} 为平移因子和伸缩因子, φ 为母小波函数。

3.4 控制器参数在线训练算法

所设计的小波神经网络控制器的训练参数包括 $\theta = \{m_{jk}, d_{jk}, w_j\}$ 。对于确定的网络结构, 通过对其进行训练, 使给定的目标函数最优, 从而找到满足要求的权值、伸缩因子和平移因子参数。一般情况下, 找到最优参数比较困难, 只能找到真实函数的一个近似。为此, 给定一个逼近精度 $\varepsilon > 0$, 只要找到 $\theta^* = \{m^*, d^*, w^*\}$, 使得控制对象的输出与期望输出的误差小于 ε , 便认为小波神经网络控制器 $u = f(x, \theta^*)$ 对被控对象达到了有效的控制。

假设被控对象的输出和参考模型的输出误差是:

$$e(t) = y_m(t) - y_p(t) \quad (3.4.1)$$

定义如下的目标函数:

$$J(\theta) = \frac{1}{2} (e(t))^2 = \frac{1}{2} (y_m(t) - y_p(t))^2, \quad (3.4.2)$$

为了提高控制的精度和响应速度, 便于在线训练参数, 本章提出了基于梯度下降和递推最小二乘相结合的在线混合训练算法。算法描述如下。

3.4.1 平移因子和伸缩因子训练算法

常规的 BP 算法存在如下缺点: 首先, 为了极小化误差, 学习速

率要选得足够小，但小的学习速率又使得学习过程很慢；其次，大的学习速率又可能导致学习过程的震荡，从而收敛不到期望解；另外学习过可能收敛到局部极小点或在误差函数的平稳段停止不前。

一种简单的改进方法是在权值的调整律中加入动量项，以平滑权值的变化，它的引入可以提高收敛速度和改善动态性能。借鉴这一思想，我们对小波神经网络中的平移因子 m_{jk} 和伸缩因子 d_{jk} 的训练采用带有动量项的梯度下降法，以加快其学习速度，并避免局部小问题出现的概率。

m_{jk} , d_{jk} 的更新方法如下：

$$\Delta m_{jk}(t) = -\eta \frac{\partial J(\theta)}{\partial m_{jk}} + \alpha \Delta m_{jk}(t-1) \quad (3.4.3)$$

$$\Delta d_{jk}(t) = -\eta \frac{\partial J(\theta)}{\partial d_{jk}} + \alpha \Delta d_{jk}(t-1) \quad (3.4.4)$$

其中 η 为学习率，也就是按梯度法搜索的步长，一般取为小于 1 的正值常数。 α 为动量因子，它的大小决定过去参数变化对目前值的影响程度，其作用为记忆上次参数的变化方向，抑制系统可能产生的振荡，起平滑的作用。选择适当的动量因子，可以降低反传算法对误差表面的灵敏度，使网络避免陷入局部最小点。

对 (3.4.3) 和 (3.4.4) 中的 m_{jk} , d_{jk} 求偏导数得：

$$\frac{\partial J}{\partial m_{jk}} = -\frac{\partial J}{\partial y_p} \frac{\partial y_p}{\partial u} \frac{\partial u}{\partial m_{jk}} \quad (3.4.5)$$

$$\frac{\partial J}{\partial d_{jk}} = -\frac{\partial J}{\partial y_p} \frac{\partial y_p}{\partial u} \frac{\partial u}{\partial d_{jk}} \quad (3.4.6)$$

其中，

$$\frac{\partial u}{\partial m_{jk}} = -\frac{w_j}{d_{jk}} \frac{\partial \Phi_j}{\partial z_{jk}} \Big|_{x=x^t} \quad k=1, \dots, N_t, \quad j=1, \dots, N_w \quad (3.4.7)$$

$$\frac{\partial u}{\partial d_{jk}} = -\frac{w_j}{d_{jk}} z'_{jk} \frac{\partial \Phi_j}{\partial z_{jk}} \Big|_{x=x^t} \quad k=1, \dots, N_t, \quad j=1, \dots, N_w \quad (3.4.8)$$

$$\frac{\partial \Phi_j}{\partial z_{jk}} \Big|_{x=x^t} = \varphi(z'_{j1}) \varphi(z'_{j2}) \cdots \varphi'(z'_{jk}) \cdots \varphi(z'_{jN_t}) \quad (3.4.9)$$

$$\varphi'(z'_{jk}) = \frac{d\phi(z)}{dz} \Big|_{z=z'_{jk}} = z'_{jk} \quad (3.4.10)$$

x^t 表示 t 时刻网络的输入, $\frac{\partial y_p}{\partial u}$ 为未知非线性对象的 Jacobi 矩阵,

由于被控对象表达式未知, 故不可求取, 可以用下面的近似方法求解

$$\frac{\partial y_p}{\partial u} = \frac{y_p(t) - y_p(t-1)}{u(t) - u(t-1)}.$$

3.4.2 权值训练算法

常规BP神经网络的权值训练一般用基于梯度下降的BP算法, 由于收敛速度比较慢, 所以不利于在线训练。从式 (3.3.1) 可以看出, 网络的输出 u 和权值 w_j 之间呈线性关系, 故可用线性化的方法进行求解, 因此, 本章对于权值的训练采用递推最小二乘法求取, 这样就保证了解的唯一性, 既避免了陷入局部最小, 又避免了复杂的矩阵求导运算, 从而更利于在线计算以达到实时控制的目的。

先介绍最小二乘算法的思想。假设网络的训练样本为 (X_k, Y_k) ,

$k=1, 2, \dots, m$ 。 X_k 为输入样本, 可以表示为:

$$X_k = [x_{1,k}, x_{2,k}, \dots, x_{N_t,k}]^T \quad (3.4.11)$$

Y_k 为期望输出，可以表示为：

$$Y_k = [y_{1,k}, y_{2,k}, \dots, y_{N_0,k}]^T \quad (3.4.12)$$

设 $\bar{X}_k = [\Phi_{1,k}, \Phi_{2,k}, \dots, \Phi_{N_w,k}]^T = \Phi_k(X_k)$ 为输入样本数据 X_k 经小波变换后的输出。 \bar{X}_k 和 Y_k 之间的变换关系可用如下方程描述：

$$Y_m' = W_m \bar{X}_m' \quad (3.4.13)$$

即：

$$\begin{bmatrix} y_{1,1} & y_{1,2} & \dots & y_{1,m} \\ y_{2,1} & y_{2,2} & \dots & y_{2,m} \\ \vdots & \vdots & \ddots & \vdots \\ y_{N_0,1} & y_{N_0,2} & \dots & y_{N_0,m} \end{bmatrix} = \begin{bmatrix} w_{1,1} & w_{2,1} & \dots & w_{N_0,1} \\ w_{1,2} & w_{2,2} & \dots & w_{N_0,2} \\ \vdots & \vdots & \ddots & \vdots \\ w_{1,N_w} & w_{2,N_w} & \dots & w_{N_0,N_w} \end{bmatrix}^T \times \begin{bmatrix} \Phi_{1,1} & \Phi_{1,2} & \dots & \Phi_{1,m} \\ \Phi_{2,1} & \Phi_{2,2} & \dots & \Phi_{2,m} \\ \vdots & \vdots & \ddots & \vdots \\ \Phi_{N_w,1} & \Phi_{N_w,2} & \dots & \Phi_{N_w,m} \end{bmatrix}$$

$Y_{i,m}$ 表示对应 m 个样本第 i 个输出端的输出矢量，它可以表示为：

$$Y_{i,m} = \bar{X}_m'^T W_i \quad (3.4.14)$$

假定矩阵 $\bar{X}_m' \bar{X}_m'^T$ 为非奇异，则式 (3.4.14) 的最小二乘解为：

$$W_{i,m} = (\bar{X}_m' \bar{X}_m'^T)^{-1} \bar{X}_m' Y_{i,m} \quad (3.4.15)$$

在本章 i 取 1。为了便于在线计算，采用最小二乘解的递推形式，由递推最小二乘思想，它可以表示为^[46]：

$$W(t) = W(t-1) + (1 + \bar{X}^T(t)P(t-1)\bar{X}(t))^{-1} P(t-1)\bar{X}(t)e(t) \quad (3.4.16)$$

式中：

$$P(t-1) = (\bar{X}(t-1)\bar{X}^T(t-1))^{-1} \quad (3.4.17)$$

$$e(t) = y_m(t) - y_p(t) \quad (3.4.18)$$

$$P(t) = P(t-1) - \left(1 + \bar{X}^T(t)P(t-1)\bar{X}(t)\right)^{-1} P(t-1)\bar{X}(t)\bar{X}^T(t)P(t-1) \quad (3.4.19)$$

其中， $\bar{X}(t)$ 为第 t 个采样时刻的输入矢量 $x(t)$ 对应的隐含层的输出 $\bar{X}(t) = [\Phi_1(t), \Phi_2(t), \dots, \Phi_{N_c}(t)]^T$ 。初始化矩阵 $P_0 = \gamma I$ ，其中 γ 为比较大的正实数。

上述算法要求矩阵 $\bar{X}_m \bar{X}_m^T$ 为非奇异，可以证明它是正半定的，但未必是正定的。因此 $\bar{X}_m \bar{X}_m^T$ 的非奇异就无法保证，其结果是 $P(t)$ 的递推计算可能存在不稳定性。解决方法是将(3.4.15)中的 $(\bar{X}_m \bar{X}_m^T)^{-1}$ 以 $(\bar{X}_m \bar{X}_m^T)^{-1} + \mu I$ 代替， μ 为很小的正实数， I 为单位矩阵。从而(3.4.15)重写为：

$$W_{l,m} = (\bar{X}_m \bar{X}_m^T + \mu I)^{-1} \bar{X}_m Y_{l,m} \quad (3.4.20)$$

式(3.4.17)重写为：

$$P(t-1) = (\bar{X}(t-1)\bar{X}^T(t-1) + \mu I)^{-1} \quad (3.4.21)$$

3.5 仿真分析

考虑非线性对象：

$$y_p(k+1) = 0.75 \left\{ \frac{y_p(k)y_p(k-1)[y_p(k)+2.5]}{1+y_p^2(k)+y_p^2(k-1)} \right\} + u(k) \quad (3.5.1)$$

参考模型为线性降阶形式：

$$y_m(k+1) = 0.5y_m(k) + 0.3u_m(k) \quad (3.5.2)$$

控制输入为：

$$u_m(k) = \begin{cases} \sin(2\pi kT) + \sin(4\pi kT) & 0 < k < 200 \\ \sin(4\pi kT) + \sin(6\pi kT) & 200 \leq k \leq 400 \end{cases} \quad (3.5.3)$$

其中 k 为采样时刻， T 为采样时间。

不加控制时，该系统对参考输入的开环响应为图3-3。其中虚线为参考模型的输出，实线为对象输出，可看出对象有很强的非线性。

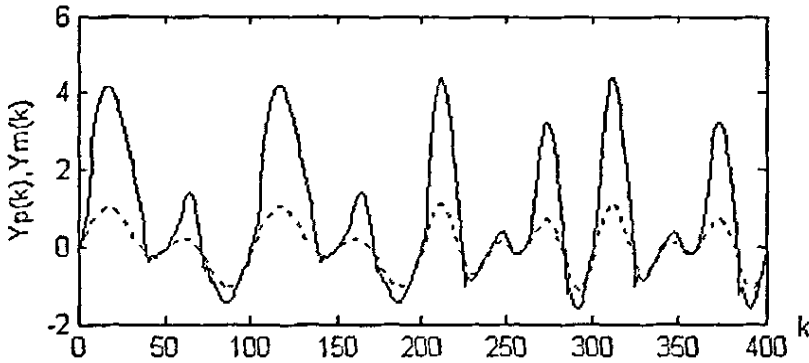


图3-3 系统的开环响应

(Y_p “—”: 对象的输出曲线, Y_m “- -”: 参考模型的输出)

为了说明小波神经网络的优越性，同时采用BP神经网络进行控制进行比较。BP神经网络控制采用形如图3-1的控制结构，仅把控制器进行替换，BP控制器权值的训练算法采用梯度下降法。

(1) 网络参数初始化

网络的初始化对于神经网络的训练来说非常重要，初始值的选取不当可能直接导致网络收敛很慢，甚至会发散。对于小波神经网络的参数初始化尤为重要，不能同一般的神经网络一样初始化为很小的随机数。当选取很小的伸缩参数时，会引起小波太局部化，使得目标函数的梯度参数在感兴趣的区域非常的小。对小波神经网络的初始化一般来说，要充分考虑到小波神经网络输入空间的区域特性。

我们选取的基小波函数为 $\varphi(x) = -x \exp\left(-\frac{1}{2}x^2\right)$ ，根据采样所确定

的输入区间，我们用 $[\alpha_k, \beta_k]$ 表示输入向量中第 k 个元素的区间，各参数初始化如下^[16]：

$m_{jk} = \frac{1}{2}(\alpha_k + \beta_k)$ ， $d_{jk} = 0.2(\beta_k - \alpha_k)$ ， $w_j (k=1, \dots, N_w)$ 初始化为小的随机数。

对于BP神经网络的权值 w_1 和 w_2 初始化为小的随机数。

(2) 控制器参数初始化

为保证控制效果，在实行在线控制前，首先利用开环响应时产生的样本数据对控制器进行离线训练，以获得系统的基本特性，从而得到控制器在线训练参数的初始值。离线学习的参数 m_{jk} ， d_{jk} 训练方法采用3.4.1中提出的算法，此时权值 w_j 的训练同样采用梯度下降法如下：

$$\Delta w_j(t) = -\eta \frac{\partial J(\theta)}{\partial w_j} + \alpha \Delta w_j(t-1) \quad (3.5.4)$$

$$\frac{\partial J}{\partial w_j} = -\frac{\partial J}{\partial y_p} \frac{\partial y_p}{\partial u} \frac{\partial u}{\partial w_j} \quad (3.5.5)$$

$$\frac{\partial u}{\partial w_j} = \Phi_j(x'), \quad j=1, \dots, N_w \quad (3.5.6)$$

小波神经网络选用的 $N_w = 3$ ，学习率取 $\eta = 0.0001$ ，设定误差为0.001，训练误差的变化曲线如图3-4，可看出在365次后达到了设定的精度值。

BP神经网络采用单隐层的网络结构，隐单元选用的神经元个数为3，选用Sigmoid函数，设定误差为0.001，训练1000次的误差曲线如图3-5所示，可以看出，此时误差几乎不再收敛，而是产生震荡，达不到设定的误差值。

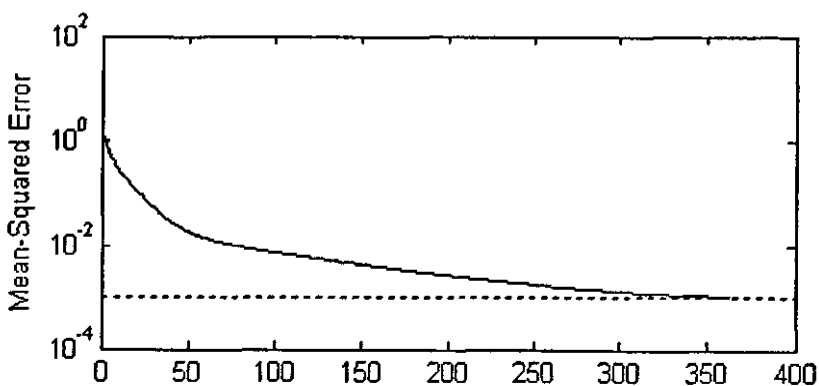


图3-4 WNN训练误差曲线（横坐标为训练次数）

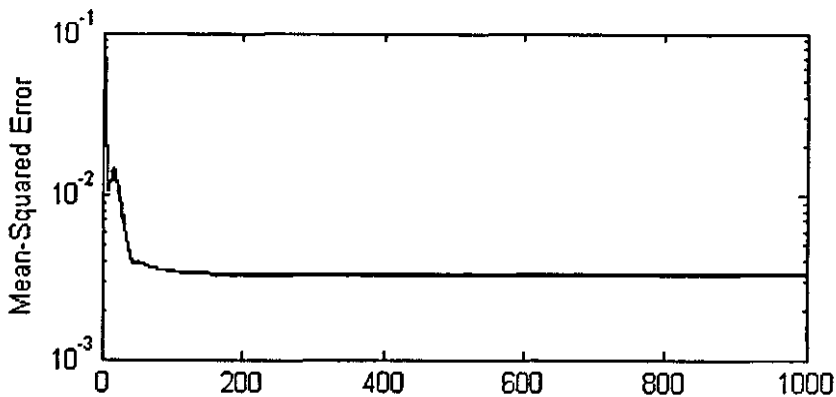


图3-5 BP训练误差曲线（横坐标为训练次数）

训练后的网络输出和实际输出的曲线如图3-6和图3-7所示，小波神经网络输出的误差均方差（Mean-Squared Error） $TMSE = 9.9724e-004$ ，而BP神经网络的 $TMSE = 0.0482$ 。

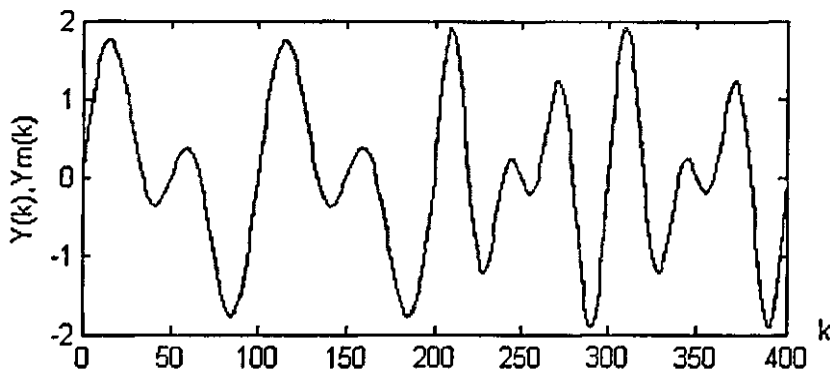


图3-6 WNN的逼近效果

(Y “—”: 小波神经网络的输出, Ym “—”: 实际输出曲线)

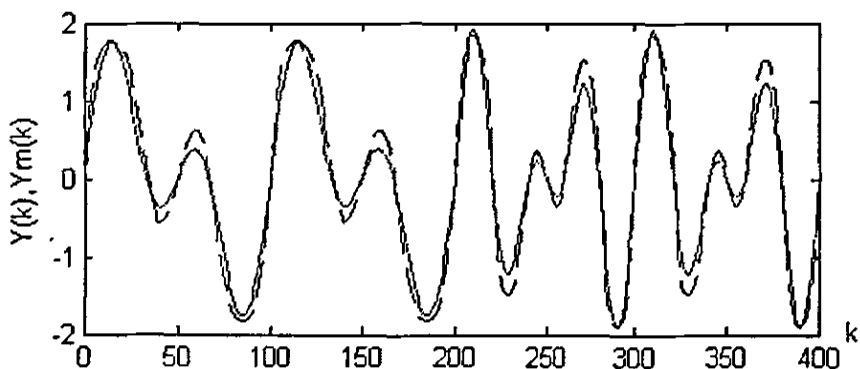


图3-7 BP神经网络的逼近效果

(Y “—”: BP神经网络的输出, Ym “—”: 实际输出曲线)

可见小波神经网络可以很好的学习对象的特性，并且在很少的训练次数内可以使误差达到一个很高的精度。

(3) 在线控制

离线训练后，将小波神经网络控制器加到系统中去，此时将离线训练好的参数 $\theta = \{m_{jk}, d_{jk}, w_j\}$ 保存为网络的初始参数对系统进行在线控制。控制算法采用本文3.4提出的在线训练算法，取 $N_w = 3$ ， $T = 0.01$ ， $\eta = 0.2$ ， $\alpha = 0.1$ ， $P = [0.25, 0, 0; 0, 0.25, 0; 0, 0, 0.25]$ 。跟踪误差的均方差 $MSE = 0.0067$ 。

BP神经网络控制，同样将离线训练的参数保存为在线控制的初始值，采样时间为 $T = 0.01$ ，学习律 $\eta = 0.2$ ，跟踪误差的均方差 $MSE = 0.0215$ 。

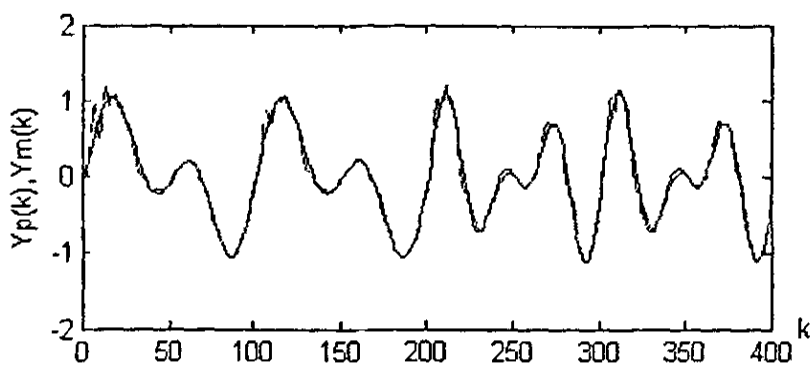


图3-8 小波神经网络的控制效果

(Y_p “—”: 被控对象的输出, Y_m “—”: 参考模型的输出)

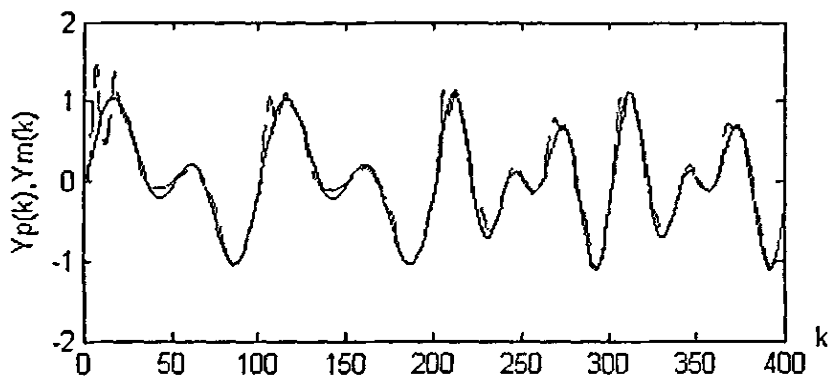


图3-9 BP神经网络的控制效果

(Y_p “—”: 对象的输出曲线, Y_m “—”: 参考模型的输出)

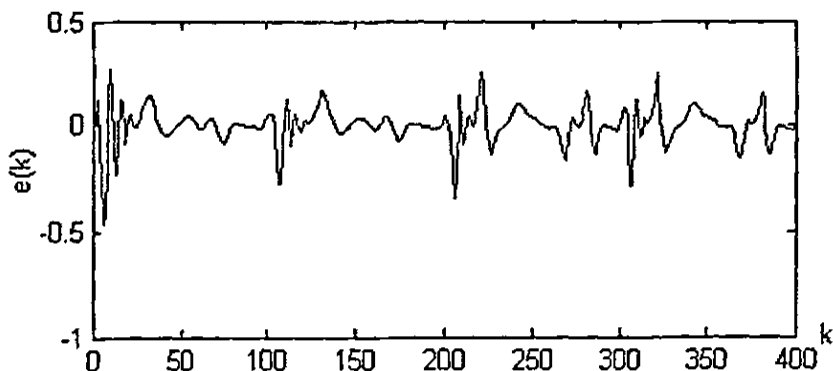


图3-10 WNN被控对象输出与参考模型输出误差

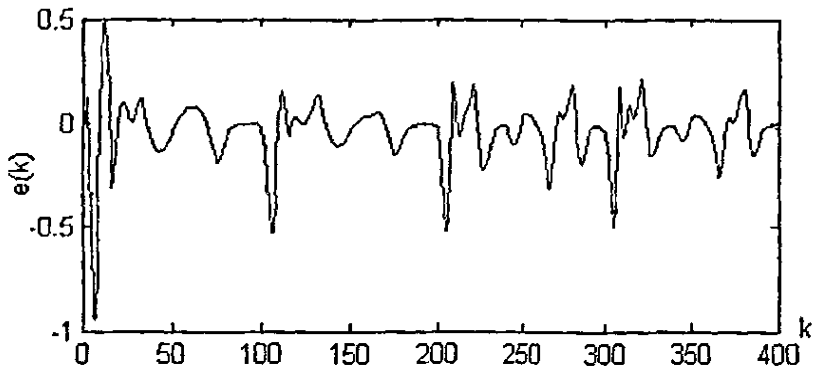


图3-11 BP被控对象输出与参考模型输出误差

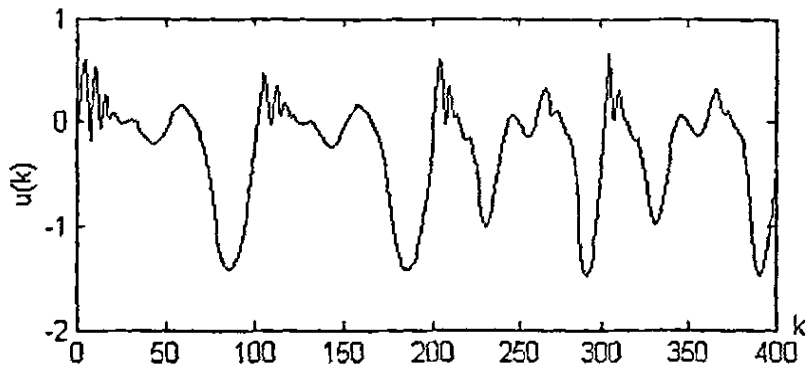


图3-12 WNNC的控制作用

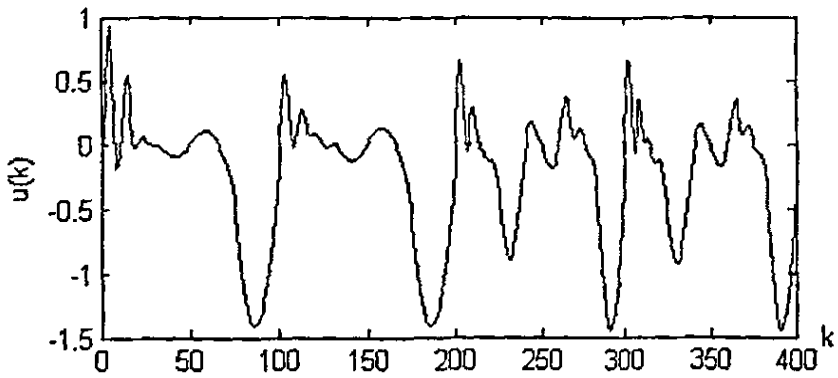


图3-13 BP的控制作用

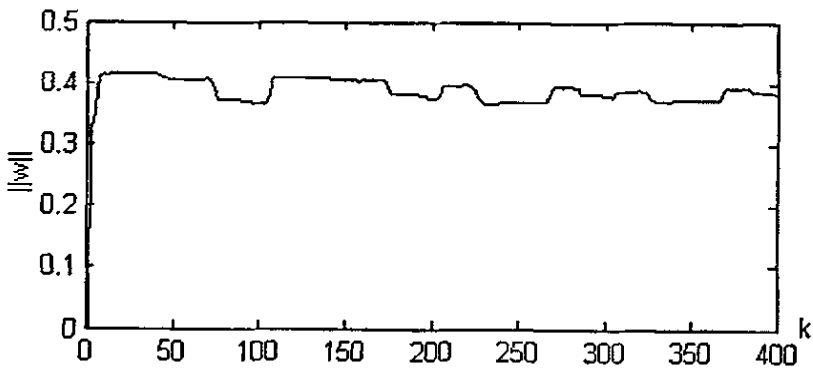


图3-14 WNNC权值范数 $\|w\|$

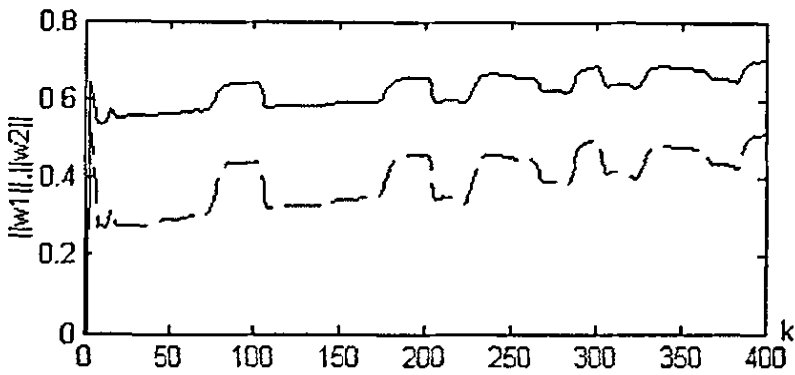


图3-15 BP的权值范数 ($\|w_1\|$ (“—”), $\|w_2\|$ (“--”))

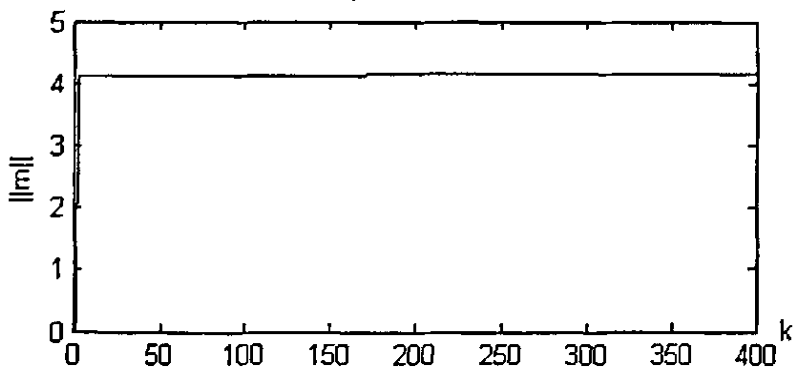
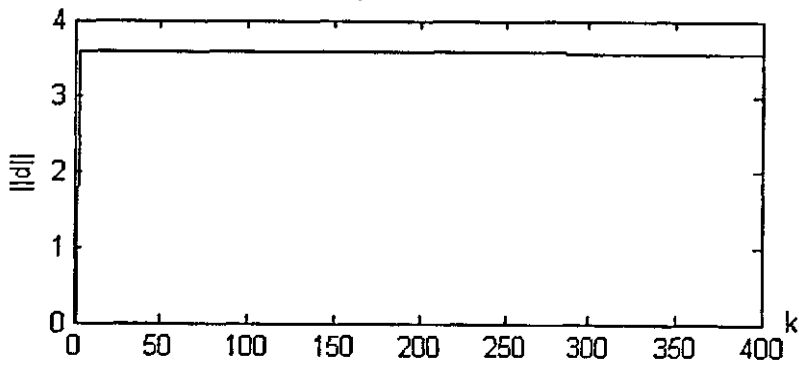


图3-16 WNNC的 $\|m\|$ 变化

图3-17 WNNC的 $\|a\|$ 变化

通过仿真实验，可以看出控制器的权值、平移因子、伸缩因子和控制作用都是有界的，并且通过与BP神经网络对比，可以看出在函数逼近上，小波神经网络要明显优于BP神经网络，同时所设计的小波神经网络控制器采用提出的在线混合训练算法可实现在线自适应控制，且优于BP控制器，同时这种小波神经网络控制器的控制结构实现简单，易于调整，控制精度高。

第4章 SISO 系统小波神经网络自适应控制器设计及应用

4.1 引言

反馈线性化的提出使得非线性控制理论获得飞速的发展，反馈线性化的建立是将非线性系统通过微分同胚变换成一个标准型，这需要系统方程的精确描述。当系统模型存在不确定性的时候，我们可以用神经网络来逼近其非线性函数，设计基于神经网络的自适应控制器。

在基于神经网络的自适应控制系统中，神经网络的权值是利用网络的输出误差进行修正，而自适应控制律则是利用神经网络的输出进行综合，因此利用神经网络实现非线性系统自适应控制时需要考虑的关键问题是网络权值修正规则的收敛性和闭环系统的稳定性。

对于一类仿射系统，人们采用 Lyapunov 稳定性理论，将神经网络与反馈线性化控制相结合，给出了一些稳定的直接的自适应神经网络设计方法^[47-51]。然而，这些非线性系统的神经网络自适应控制器都是基于状态反馈的控制器，并且都是假设系统的状态已知，或直接可以测量的，对于系统状态不完全已知或不可直接测量的情况却少有考虑。

根据系统输出设计反馈控制器的一种设计方案是基于系统的观测器设计输出反馈控制，即：由系统输出设计一个状态观测器估计系统的状态，再对估计状态应用状态反馈控制律实现系统的输出反馈控制目的。文献[52]提出了一种基于高增益观测器的非线性系统的输出反馈控制设计方法。

同时，这些控制器的设计都是基于常规 BP 神经网络的设计方法，对于小波神经网络的应用研究，却很少涉及。

基于上述设计思想，本章利用反馈线性化方法考虑了一类单入单

出(SISO)非线性系统的自适应神经网络输出跟踪控制问题,利用小波神经网络的优良品质,不需辨识系统的模型,而是直接设计在线的控制器,网络权值的自适应修正规则是基于 Lyapunov 稳定性理论实现的,保证了闭环系统的稳定性及闭环信号的有界性。最后,通过在一类化工过程 CSTR 上仿真,验证了控制器的有效性。

4.2 问题的提出

考虑如下一类 SISO 非线性系统

$$\begin{cases} \dot{x}_1 = x_2 \\ \vdots \\ \dot{x}_{n-1} = x_n \\ \dot{x}_n = f(x) + g(x)u \\ y = x_1 \end{cases} \quad (4.2.1)$$

其中,状态 $x \in R^n$, 输出 $y \in R$, 输入 $u \in R$, $f(x), g(x) \in R^n$ 为未知非线性光滑函数。

假设系统(4.2.1)满足下面条件:

假设 1: 在包含 x_0 的某个紧集 $\Omega_{x_0} \in R^n$ 上, $g(x) \neq 0$, 即 $g(x)$ 严格为正或为负。

假设 2: 系统的期望输出 y_d 和它的导数是光滑且有界的已知函数, 记 $y_r = [y_d, y_d^{(1)}, \dots, y_d^{(n-1)}]^T$, $y_R = [y_r^T, y_r^{(n)}]^T$, Y_r 和 Y_R 分别为 R^n 和 R^{n+1} 中的紧集, 且 $y_r \in Y_r$, $y_R \in Y_R$, 对所有的 $t > 0$ 。

控制的目的是如何设计控制 $u(t)$, 使得系统(4.2.1)的输出 $y(t)$ 在容许的精度内跟踪一条给定的轨线 $y_d(t)$ 。

4.3 小波神经网络控制器设计

令变量 $e_1 = x_1 - y_d$, $e_i = x_i - y_d^{(i-1)}$, $i = 2, \dots, n$, 结合式 (4.2.1)

则向量 $e = [e_1, \dots, e_n]^T$ 的动态方程为:

$$\begin{aligned} \dot{e}_1 &= \dot{x}_1 - y_d^{(1)} = x_2 - y_d^{(1)} = e_2 \\ \dot{e}_2 &= \dot{x}_2 - y_d^{(2)} = x_3 - y_d^{(2)} = e_3 \\ &\vdots \\ \dot{e}_n &= \dot{x}_n - y_d^{(n)} = f(x) + g(x)u - y_d^{(n)} \end{aligned}$$

即有:

$$\begin{cases} \dot{e} = Ae + b[f(x) + g(x)u - y_d^{(n)}] \\ y_e = ce = e_1 \end{cases} \quad (4.3.1)$$

其中, 矩阵 A 和向量 b 分别为:

$$A = \begin{bmatrix} 0 & 1 & 0 & 0 & \dots & 0 \\ 0 & 0 & 1 & 0 & \dots & 0 \\ \vdots & \vdots & 0 & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \quad b = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix}, \quad c = [1 \quad 0 \quad \dots \quad 0]$$

显然, 系统 (4.3.1) 是完全可控的, 我们选择 k 使得 $A_c = A - bk$ 是严格 Hurwitz 矩阵, 即矩阵的特征根都在左半平面。则 (4.3.1) 可重写为:

$$\begin{cases} \dot{e} = A_c e + b[ke + f(x) + g(x)u - y_d^{(n)}] \\ y_e = ce = e_1 \end{cases} \quad (4.3.2)$$

当系统 (4.2.1) 中的 $f(x)$, $g(x)$ 未知或存在不确定性时, 我们可以利用两个小波神经网络 WNN_f, WNN_g 分别对其进行逼近, 选用同第三章所述结构形式一样的连续小波神经网络 (表达式见第三章, 此

处不再赘述)，权值的训练在算法上本章是基于 Lyapunov 稳定性理论经过严格推导给出。

假设 3：在紧集 $\Omega_{x_0} \in R^n$ 上，存在如下的网络逼近：

$$\begin{aligned} f(x) &= w_f^{*T} \Phi_f(x) + d_f \\ g(x) &= w_g^{*T} \Phi_g(x) + d_g \end{aligned} \quad (4.3.3)$$

其中， w_f^* 和 w_g^* 均为理想的有界权值向量，仅作为理论分析用，实际应用中可用 w_f 和 w_g 分别对其估计， d_f 和 d_g 均为理想的逼近误差，且满足 $|d_f| \leq \bar{d}_f$ ， $|d_g| \leq \bar{d}_g$ ， $\Phi_f(x)$ 和 $\Phi_g(x)$ 分别为两个基小波函数的输出。

假设 4：存在 $\hat{\Omega}_g$ 使得 $\hat{g}(x, w_g) = w_g^T \Phi_g(x) \neq 0$ ，对于 $\forall x \in \Omega_{x_0}$ ， $\forall w_g \in \hat{\Omega}_g$ 。

于是，由反馈线性化理论可设计系统 (4.2.1) 的小波神经网络控制器为：

$$\hat{u} = \frac{-w_f^T \Phi_f(x) + v(t)}{w_g^T \Phi_g(x)} \quad (4.3.4)$$

其中， $v(t) = y_d^{(n)} - \alpha_{n-1}e_n - \dots - \alpha_0e_1$ 为辅助输入信号， e 为跟踪误差向量，系数 $\alpha_{n-1}, \dots, \alpha_0$ 的选取使得多项式 $\Gamma(s) = s^n + \alpha_{n-1}s^{n-1} + \dots + \alpha_0$ 为 Hurwitz 多项式。在此假设 $|\hat{u}| \leq \bar{u}$ ， \bar{u} 为控制作用的幅值。

令 $k = [\alpha_0, \dots, \alpha_{n-1}]$ ，则式 (4.3.4) 可重写为：

$$\hat{u} = \frac{-w_f^T \Phi_f(x) + y_d^{(n)} - ke}{w_g^T \Phi_g(x)} \quad (4.3.5)$$

定义估计误差为：

$$\begin{aligned}\tilde{w}_f &= w_f - w_f^* \\ \tilde{w}_g &= w_g - w_g^*\end{aligned}\tag{4.3.6}$$

把 (4.3.5) 式代入 (4.3.2) 得:

$$\begin{aligned}\dot{e} &= A_c e + b[ke + f(x) + g(x)\hat{u} - y_d^{(n)}] \\ &= A_c e + b[ke + f(x) + (g(x) - \hat{g}(x, w_g))\hat{u} + \hat{g}(x, w_g)\hat{u} - y_d^{(n)}] \\ &= A_c e + b[ke + w_f^{*T}\Phi_f(x) + d_f + (w_g^{*T}\Phi_g(x) + d_g - w_g^T\Phi_g(x))\hat{u} \\ &\quad + (w_f^T\Phi_f(x) + y_d^{(n)} - ke) - y_d^{(n)}] \\ &= A_c e + b[-\tilde{w}_f^T\Phi_f(x) - \tilde{w}_g^T\Phi_g(x)\hat{u} + d_g\hat{u} + d_f]\end{aligned}\tag{4.3.7}$$

设 $P = P^T > 0$ 是下列 Lyapunov 方程的解:

$$PA_c + A_c^T P = -Q, \quad Q = Q^T > 0\tag{4.3.8}$$

我们定义如下的小波神经网络权值自适应律:

$$\begin{aligned}\dot{w}_f &= \dot{\tilde{w}}_f = \gamma_f [2e^T P b \Phi_f(x) - k_f w_f] \\ \dot{w}_g &= \dot{\tilde{w}}_g = \gamma_g [2e^T P b \Phi_g(x)\hat{u} - k_g w_g]\end{aligned}\tag{4.3.9}$$

其中, $\gamma_f, \gamma_g, k_f, k_g$ 均为大于零的权值增益。

定理: 在满足假设 1—假设 4 的条件下, 如果小波神经网络的权值自适应修正规则满足 (4.3.9) 式, 则由系统 (4.2.1) 和控制律 (4.3.5) 构成的闭环系统是一致最终有界的 (UUB)。

证明: 构造如下的 Lyapunov 函数:

$$V = e^T P e + \frac{1}{2} \tilde{w}_f^T \gamma_f^{-1} \tilde{w}_f + \frac{1}{2} \tilde{w}_g^T \gamma_g^{-1} \tilde{w}_g\tag{4.3.10}$$

则 V 对时间的导数为:

$$\begin{aligned}\dot{V} &= \dot{e}^T P e + e^T P \dot{e} + \frac{1}{2} \dot{\tilde{w}}_f^T \gamma_f^{-1} \tilde{w}_f + \frac{1}{2} \tilde{w}_f^T \gamma_f^{-1} \dot{\tilde{w}}_f + \frac{1}{2} \dot{\tilde{w}}_g^T \gamma_g^{-1} \tilde{w}_g + \frac{1}{2} \tilde{w}_g^T \gamma_g^{-1} \dot{\tilde{w}}_g \\ &= [e^T A_c^T + b^T [-\tilde{w}_f^T \Phi_f(x) - \tilde{w}_g^T \Phi_g(x)\hat{u} + d_g\hat{u} + d_f]] P e + e^T P [A_c e\end{aligned}$$

$$\begin{aligned}
 & + b[-\tilde{w}_f^T \Phi_f(x) - \tilde{w}_g^T \Phi_g(x) \hat{u} + d_g \hat{u} + d_f]] + \dot{\tilde{w}}_f^T \gamma_f^{-1} \tilde{w}_f + \dot{\tilde{w}}_g^T \gamma_g^{-1} \tilde{w}_g \\
 = & e^T (A_c^T P + P A_c) e + 2e^T P b [-\tilde{w}_f^T \Phi_f(x) - \tilde{w}_g^T \Phi_g(x) \hat{u} + d_g \hat{u} + d_f] \\
 & + \dot{\tilde{w}}_f^T \gamma_f^{-1} \tilde{w}_f + \dot{\tilde{w}}_g^T \gamma_g^{-1} \tilde{w}_g \\
 = & -e^T Q e + 2e^T P b [-\tilde{w}_f^T \Phi_f(x) - \tilde{w}_g^T \Phi_g(x) \hat{u} + d_g \hat{u} + d_f] + \dot{\tilde{w}}_f^T \gamma_f^{-1} \tilde{w}_f \\
 & + \dot{\tilde{w}}_g^T \gamma_g^{-1} \tilde{w}_g \tag{4.3.11}
 \end{aligned}$$

把 (4.3.9) 式中的权值调整律代入 (4.3.11) 得:

$$\begin{aligned}
 \dot{V} = & -e^T Q e + 2e^T P b [-\tilde{w}_f^T \Phi_f(x) - \tilde{w}_g^T \Phi_g(x) \hat{u} + d_g \hat{u} + d_f] \\
 & + [2e^T P b \Phi_f(x) - k_f w_f] \tilde{w}_f + [2e^T P b \Phi_g(x) \hat{u} - k_g w_g] \tilde{w}_g \\
 = & -e^T Q e + 2e^T P b [d_f + d_g \hat{u}] - [k_f \tilde{w}_f^T w_f + k_g \tilde{w}_g^T w_g] \tag{4.3.12}
 \end{aligned}$$

对上式放大不等式:

(1) 记 $\lambda_{\min}(Q)$ 为矩阵 Q 的最小特征值, 由第二章的预备知识得:

$$-e^T Q e \leq -\lambda_{\min}(Q) \|e\|^2 \tag{4.3.13}$$

$$(2) \quad 2e^T P b [d_f + d_g \hat{u}] \leq 2 \|P b\| \|e\| (|d_f| + |d_g| |\hat{u}|) \leq 2 \|P b\| \|e\| d_{fg} \tag{4.3.14}$$

其中, $d_{fg} = |d_f| + |d_g| |\hat{u}|$;

$$\begin{aligned}
 (3) \quad -\tilde{w}_f^T w_f & = -\tilde{w}_f^T w_f + (\tilde{w}_f^T w_f^* - \tilde{w}_f^T w_f^*) \\
 & = -\tilde{w}_f^T (w_f - w_f^*) - \tilde{w}_f^T w_f^* \\
 & = -(\|\tilde{w}_f\|^2 + \tilde{w}_f^T w_f^*)
 \end{aligned}$$

$$\leq -(\|\tilde{w}_f\|^2 - \|\tilde{w}_f\| \|w_f^*\|) \quad (4.3.15)$$

同理有：

$$-\tilde{w}_g^T w_g = -(\|\tilde{w}_g\|^2 + \tilde{w}_g^T w_g^*) \leq -(\|\tilde{w}_g\|^2 - \|\tilde{w}_g\| \|w_g^*\|) \quad (4.3.16)$$

将式 (4.3.13) (4.3.14) (4.3.15) (4.3.16) 代入式 (4.3.12) 得：

$$\begin{aligned} \dot{V} \leq & -\lambda_{\min}(Q)\|e\|^2 + 2\|Pb\|\|e\|d_{fg} + k_f\|\tilde{w}_f\|\|w_f^*\| + k_g\|\tilde{w}_g\|\|w_g^*\| \\ & - [k_f\|\tilde{w}_f\|^2 + k_g\|\tilde{w}_g\|^2] \end{aligned} \quad (4.3.17)$$

令：

$$c_0 = 2\|Pb\|\|e\|d_{fg} + k_f\|\tilde{w}_f\|\|w_f^*\| + k_g\|\tilde{w}_g\|\|w_g^*\|$$

于是有：

$$\dot{V} \leq -\lambda_{\min}(Q)\|e\|^2 - [k_f\|\tilde{w}_f\|^2 + k_g\|\tilde{w}_g\|^2] + c_0 \quad (4.3.18)$$

对 (4.3.10) 式放大不等式可得：

$$V \leq \lambda_{\max}(P)\|e\|^2 + \frac{1}{2\gamma_f}\|\tilde{w}_f\|^2 + \frac{1}{2\gamma_g}\|\tilde{w}_g\|^2 \quad (4.3.19)$$

其中， $\lambda_{\max}(P)$ 为矩阵 P 的最大特征值。

由式 (4.3.18) 和 (4.3.19) 可得：

$$\dot{V} \leq -\lambda_0 V + c_0 \quad (4.3.20)$$

λ_0 由下式给出：

$$\lambda_0 = \min\{\lambda_{\min}(Q)/\lambda_{\max}(P), 2k_f\gamma_f, 2k_g\gamma_g\}$$

其中， P 和 Q 均为实对称正定矩阵，故 $\lambda_{\max}(P)$ 和 $\lambda_{\min}(Q)$ 均为

正数。由 (4.3.20) 知在 $V \leq \frac{c_0}{\lambda_0}$ 所确定的紧集外， $\dot{V} < 0$ 。由 (4.3.20)，

并根据 Bellman-Gronwall 定理可以得出:

$$0 \leq V(t) \leq \frac{c_0}{\lambda_0} + \left(V(0) - \frac{c_0}{\lambda_0} \right) e^{-\lambda_0 t}, \forall t \geq 0$$

可以看出, $V(t)$ 最终有界于 $\frac{c_0}{\lambda_0}$, 这说明闭环系统的所有信号都是一致最终有界的, 这里 e 为自然常数。

所以, 对于任意紧集 $\Omega_x \subset R^n$, 存在控制器 u , 只要 $x(0) \in \Omega_x$, 闭环系统的所有信号都是一致最终有界的 (UUB)。

本节需说明的几个问题:

(1) 本章为了叙述方便, 研究了一类形如 (4.2.1) 定义的仿射非线性系统。对于一般的仿射非线性系统, 如果它是可精确反馈线性化的系统, 则由微分几何理论可知, 存在一个微分同胚使得系统变化成式 (4.2.1) 的形式, 因此仍可应用本节给出的小波神经网络输出反馈控制器设计方法。

(2) 对于式 (4.2.1) 系统设计的重点在于 $g(x)$ 作为控制器的分母, 为了使控制器可以实现, 因此逼近 $g(x)$ 的小波神经网络 $w_g^T \Phi_g(x)$ 在整个控制过程中的输出不能为零, 当其输出为 0 时, 对其加一很小的常数使其不为零, 从而使得所设计的控制器存在。

(3) 通过 4.5 节仿真说明, 在存在干扰的情况下, 闭环系统仍能稳定在设定的工作点, 这说明虽然在设计控制器时我们没有考虑到干扰的情况, 但所得到的闭环控制系统仍然具有良好的适应性和鲁棒性。

4.4 小波神经网络自适应输出控制器设计

由上一章节可知, 小波神经网络控制器的输入为系统的所有状态, 所以以上控制器的设计是基于系统状态的反馈控制方案。若系统

的状态不能直接观测时，则需要构造状态观测器，由状态观测器的状态取代原系统的状态。

本节研究的是系统状态不完全可测得情况下，如何实现系统的输出控制，这就需要对控制器所涉及的状态进行估计。对于非线性系统，构造状态观测器是一个比较复杂的问题，这里构造如下的一种高增益观测器^[52]：

$$\begin{cases} \dot{\hat{e}}_i = \hat{e}_{i+1} + \frac{\alpha_i}{\varepsilon}(e_1 - \hat{e}_1), 1 \leq i \leq n-1 \\ \dot{\hat{e}}_n = \frac{\alpha_n}{\varepsilon}(e_1 - \hat{e}_1) \end{cases} \quad (4.4.1)$$

其中， ε 是设计的比较小的正参数。

参数 $\alpha_i > 0, 1 \leq i \leq n$ 使得如下的多项式

$\Gamma(s) = s^n + \alpha_1 s^{n-1} + \dots + \alpha_{n-1} s + \alpha_n$ 为 Hurwitz 多项式，即多项式的所有根均在左半平面。

由于高增益观测器的引入，闭环系统可能出现冲击现象，为了避免这种现象，我们可以在感兴趣的范围内对控制器进行饱和处理。另外，由于引入高增益观测器，也可能产生峰值现象，为避免这种现象，把 (4.4.1) 式变为如下的奇异摄动模型：

$$\begin{cases} \varepsilon \dot{q}_i = q_{i+1} + \alpha_i(e_1 - q_1), 1 \leq i \leq n-1 \\ \varepsilon \dot{q}_n = \alpha_n(e_1 - q_1) \end{cases} \quad (4.4.2)$$

其中： $\dot{\hat{e}}_i = \frac{q_i}{\varepsilon^{i-1}}, 1 \leq i \leq n$ 。

这样，我们就可以用状态观测器的值代替系统的状态作为小波神经网络的输入。

4.5 CSTR 系统小波神经网络控制器设计

4.5.1 CSTR 系统模型描述

CSTR (Continuous Stirred-Tank Reactor) 是聚合反应中广泛使用的重要系统，具有很强的非线性，被广泛的用作各种控制理论和算法的仿真对象，对其仿真能否取得满意的结果可以看作是对算法进行评价的一个基准。

CSTR 系统的结构如图 4-1 所示：

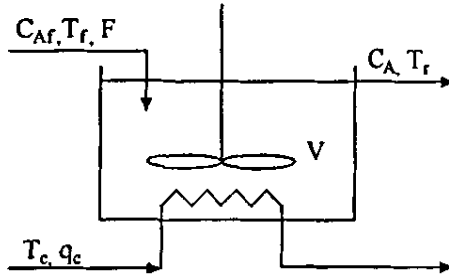


图4-1 CSTR系统结构

CSTR 的工作过程如下：浓度为 C_{Af} ，温度为 T_f 的物料以流量 F 进入反应器，在反应器中发生化学反应 $A \xrightarrow{K} B \pm \Delta H$ ，此过程为放热反应，产生的热量会降低反应速度，因此必须加入冷却剂（流量 q_c ，温度为 T_c ）带走热量，冷却釜内温度，保证产品的浓度得到控制。

由于该对象为高度非线性的对象，在控制器为线性的 PI 调节器时，很难满足要求，表现为在设定点附近阻尼比变化，系统振荡厉害。

文献[53]给出了 CSTR 的动态方程：

$$\begin{cases} V \frac{dC_A}{dt'} = F(C_{Af} - C_A) - V k_0 \exp\left(-\frac{E}{R_g T_r}\right) C_A \\ V \rho C_p \frac{dT_r}{dt'} = \rho C_p F(T_f - T_r) - V(-\Delta H) k_0 \exp\left(-\frac{E}{R_g T_r}\right) C_A - UA_h(T_r - T_c) \end{cases}$$

(4.5.1)

式中, C_A 为反应物的浓度 (mol/L), T_r 为温度 (K), k_0 为反应时间常数, E/R_g 为反应激活能 (K), C_p 质量定压热容 ($\text{cal}/(\text{g}\cdot K)$), ΔH 反应热 (cal/mol), UA_h 热交换系数。

引入如下参数:

$$\phi = \frac{E}{R_g T_{f0}}, \quad B_h = \frac{(-\Delta H)C_{Af0}}{\rho C_p T_{f0}}, \quad D_a = \frac{k_0 e^{-\phi} V}{F_0}, \quad \beta = \frac{UA_h}{F_0 \rho C_p}$$

式中, F_0 , C_{Af0} 和 T_{f0} 分别为进料率, 进料组分和温度的某一确定参考值。相关的一些无因式化变量定义如下:

$$t = \frac{t' F_0}{V}, \quad x_1 = \frac{C_{Af0} - C_A}{C_{Af0}}, \quad x_2 = \frac{T_r - T_{f0}}{T_{f0}}, \quad u = \frac{T_c - T_{c0}}{T_{f0}} \phi$$

式中 T_c 为温度的某一确定参考值。

于是, CSTR 的动态方程可无因式化为:

$$\begin{cases} \dot{x}_1 = -x_1 + D_a(1-x_1)\exp\left(\frac{x_2}{1+x_2/\phi}\right) \\ \dot{x}_2 = -x_2 + B_h D_a(1-x_1)\exp\left(\frac{x_2}{1+x_2/\phi}\right) + \beta(u-x_2) \end{cases} \quad (4.5.2)$$

式中, x_1 和 x_2 分别为无因式化后的反应物浓度和温度; u 为无因式化后的冷剂温度; D_a 为 Damkohler 系数; ϕ 为无因式化后的活化能; B_h 为反应热; β 为传热系数。

4.5.2 CSTR 系统小波神经网络控制器设计

CSTR 过程的控制目标是用 u 来控制 x_1 , 其过程参数选择如下:

$$D_e = 0.072, \phi = 20.0, B_h = 8.0, \beta = 0.3$$

所以系统的动态方程变为:

$$\begin{cases} \dot{x}_1 = -x_1 + 0.072(1-x_1)\exp\left(\frac{x_2}{1+x_2/20}\right) \\ \dot{x}_2 = -1.3x_2 + 0.576(1-x_1)\exp\left(\frac{x_2}{1+x_2/20}\right) + 0.3u \\ y = x_1 \end{cases} \quad (4.5.3)$$

应用第二章的基本知识, 对于系统 (4.5.3) 求其相对阶:

$$f(x) = \begin{bmatrix} 0.072(1-x_1)\exp\left(\frac{x_2}{1+x_2/20}\right) - x_1 \\ 0.576(1-x_1)\exp\left(\frac{x_2}{1+x_2/20}\right) - 1.3x_2 \end{bmatrix}, \quad h(x) = x_1, \quad g(x) = [0 \quad 0.3]^T$$

则由相对阶的定义得:

$$L_g h(x) = \frac{\partial h(x)}{\partial x} g(x) = 0$$

$$L_f h(x) = \frac{\partial h(x)}{\partial x} f(x) = 0.072(1-x_1)\exp\left(\frac{x_2}{1+x_2/20}\right) - x_1$$

$$L_g L_f h(x) = \frac{\partial L_f h(x)}{\partial x} g(x)$$

$$= \left[-0.072 \exp\left(\frac{x_2}{1+x_2/20}\right) - 1 \quad 0.072(1-x_1)\exp\left(\frac{x_2}{1+x_2/20}\right) \right]$$

$$* \frac{400}{(20+x_2)^2} \Big] * [0 \quad 0.3]^T$$

$$= 0.3 * 0.072(1-x_1)\exp\left(\frac{x_2}{1+x_2/20}\right) \cdot \frac{400}{(20+x_2)^2} \neq 0$$

所以, 系统(4.5.3)的相对阶为2, 对其进行坐标变换: 令 $z_1 = x_1$,

$z_2 = \dot{x}_1$ ，则 (4.5.3) 变为如下形式：

$$\begin{cases} \dot{z}_1 = z_2 \\ \dot{z}_2 = f(z) + g(z)u \\ y = z_1 \end{cases} \quad (4.5.4)$$

由于化工过程中的建模不确定性，从而使得 $f(z)$ 和 $g(z)$ 常常是不确定的，假设 $f(z)$ 和 $g(z)$ 未知，且在坐标变换的局部 $g(z) \neq 0$ 。假设系统的跟踪信号为 y_d ，利用 4.3 部分的分析， $e = [z_1 - y_d \quad z_2 - \dot{y}_d]^T$ ，则可为系统设计小波神经网络控制器如下：

$$\hat{u} = \frac{-w_f^T \Phi_f(z) + \ddot{y}_d - ke}{w_g^T \Phi_g(z)} \quad (4.5.5)$$

小波神经网络的权值调整律由式 (4.3.9) 给出。

设计控制结构如图 4-2 所示。

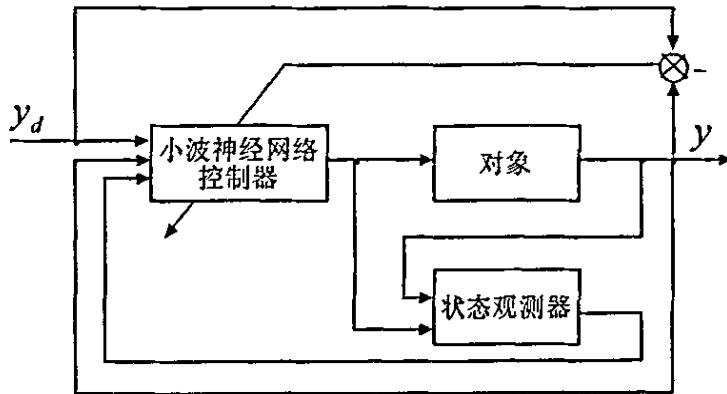


图4-2 小波神经网络控制结构图

4.5.3 仿真分析

文献[54]指出系统 (4.5.3) 是一个不稳定的非线性系统，随着 u (温度) 在 $[-2,2]$ 中的变化，有三个平衡点，但有一个不稳定的平衡点，在该平衡点出，系统增益改变符号，因此用常规的 PID 是无法进

行控制的。控制的目的是使系统从一个稳定的平衡点($y_d = 0.144$)运动到不稳定的平衡点($y_d = 0.445$)。我们利用前面的分析设计了如式 (4.4.5) 的小波神经网络控制器, 并根据 4.4 节设计了一高增益观测器。

首先测试所设计的状态观测器, 是否能很好的估计状态, 如图 4-3 所示, 图 b 是图 a 的放大显示。

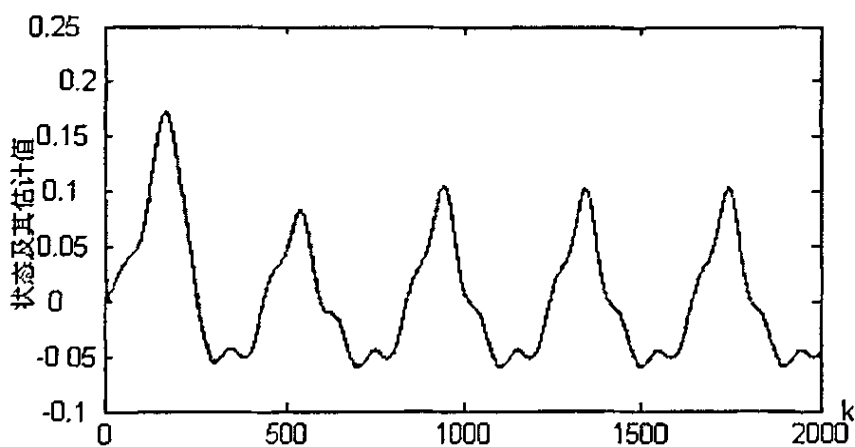


图4-3a 系统状态 z_2 (—) 和其估计值 \hat{z}_2 (---) (整体)

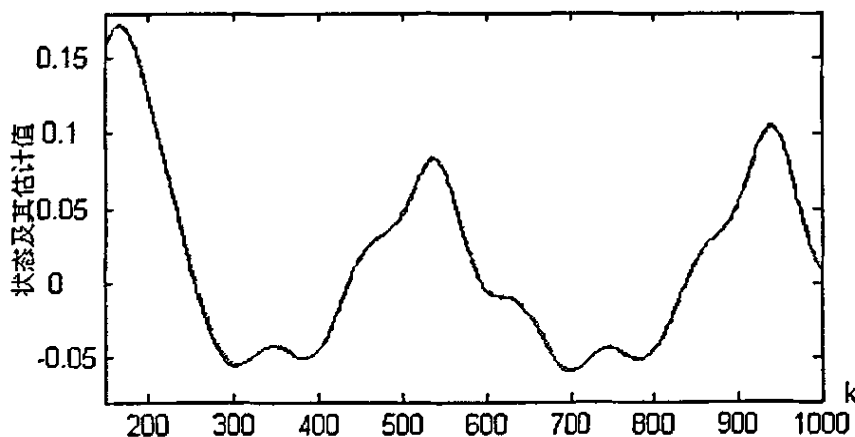


图4-3b 系统状态 z_2 (—) 和其估计值 \hat{z}_2 (---) (细节)

从上图可以看出，设计的状态观测器可以很好的估计系统的状态。

仿真参数设定：

参考信号： $y_d(t) = 0.445$ ；

系统的初始状态： $x(0) = [0.144, 0.8869]^T$ ；

观测器的初始状态： $z(0) = [0.144, 0]^T$ ； $\varepsilon = 0.02$ 且 $\alpha_1 = \alpha_2 = 1$ ；

控制器的参数： $k = [10, 4]$ ； $\gamma_f = 3, \gamma_g = 0.6, k_f = 0.08, k_g = 0.08$ ；

两个小波神经网络 WNN_f, WNN_g 的隐含层分别取 10 个小波神经元，平移和伸缩因子通过输入输出空间的时频特性来确定，平移因子的选择要靠近工作点，伸缩因子根据工作点的范围乘以经验系数。权值初始化为小的随机数，控制器的输入为对象的输出 y 和观测器得到的 \hat{y} 。积分步长 $T_s = 0.01$ 时间单位（因文献[54]未给出具体时间单位）。

(1) 系统跟踪阶跃分析

把设定值从初始的稳定平衡点的输出值（ $y_d = 0.144$ ）改为不稳定的输出（ $y_d = 0.445$ ），系统的输出相应曲线及控制作用、权值范数如图 4-4 到图 4-7。

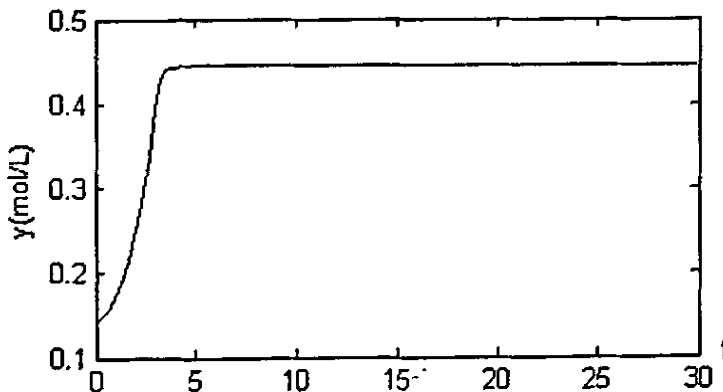


图4-4 控制器跟踪阶跃信号输出响应

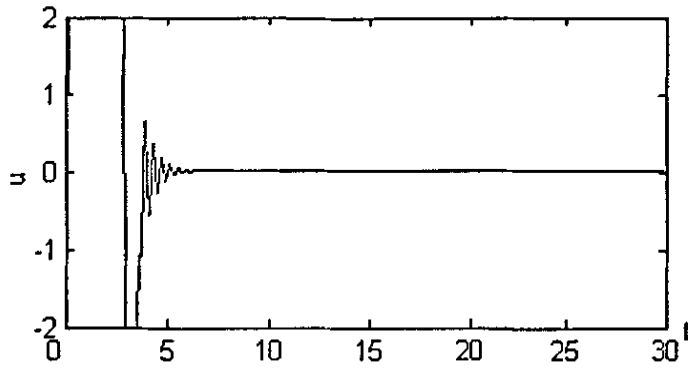


图4-5 跟踪阶跃信号的控制输入

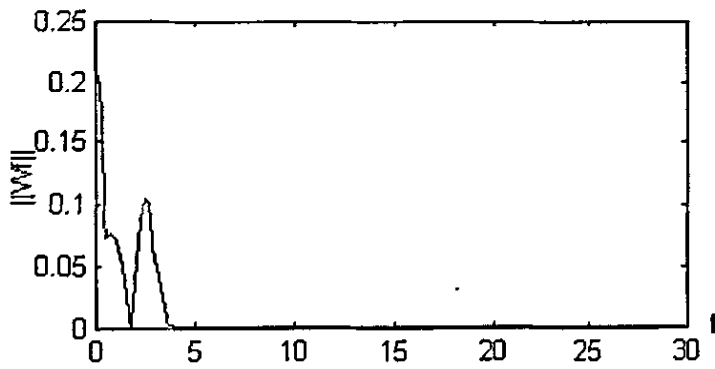


图4-6 跟踪阶跃信号时 $\|w_f\|$ 的变化

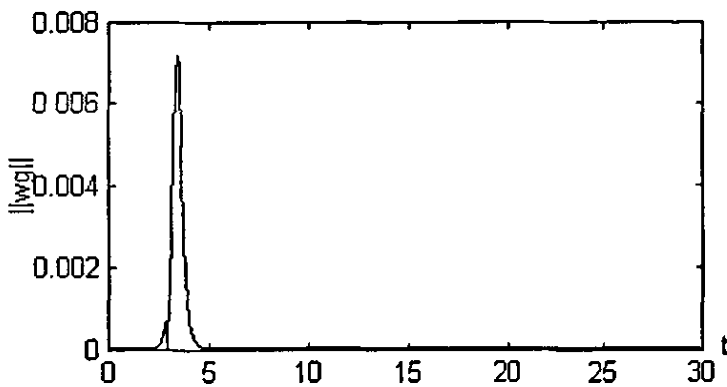


图4-7跟踪阶跃信号时 $\|w_g\|$ 的变化

可以看出设计的小波神经网络控制器有很好的伺服性能，输出能很好的跟踪期望值，权值和控制信号也是有界的。

(2) 系统的抗干扰分析

系统从稳定的工作点 ($y_d = 0.144$) 变化到在不稳定的平衡点 ($y_d = 0.445$) 后, 在时刻 10~15 之间, 把活化能 ϕ 从 20 变为 28, 在时刻 15 之后再恢复到 20。图 4-8 到图 4-12 为对应的仿真曲线 (其中图 4-8b 为图 4-8a 的局部放大), 可以看出系统的输出、控制作用和网络的权值均由于干扰而出现波动, 但系统能很快的恢复到设定值 ($y_d = 0.445$), 并表现出良好的动态性能。

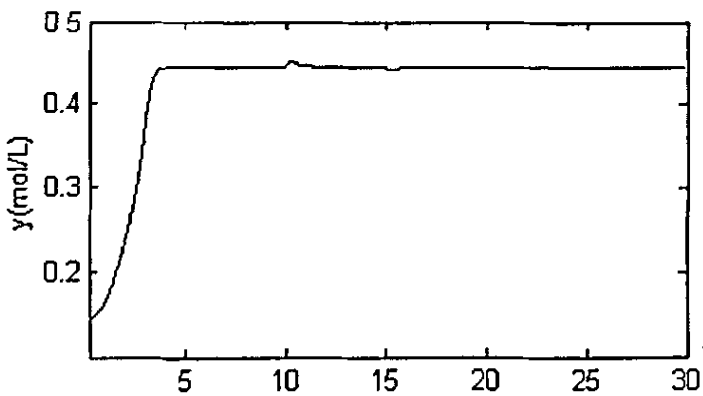


图4-8a 系统参数存在扰动时的输出响应 (整体)

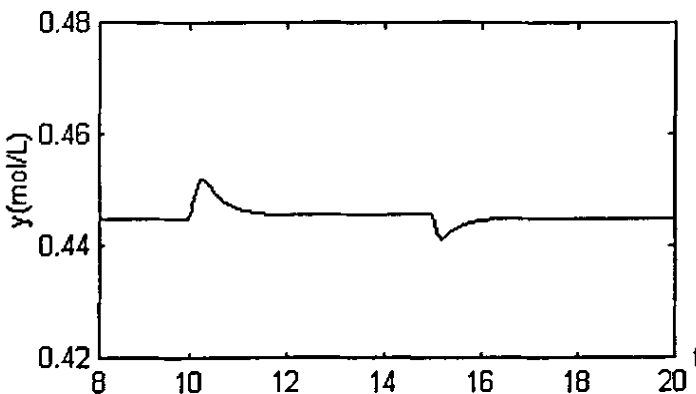


图4-8b 系统参数存在扰动时的输出响应 (细节)

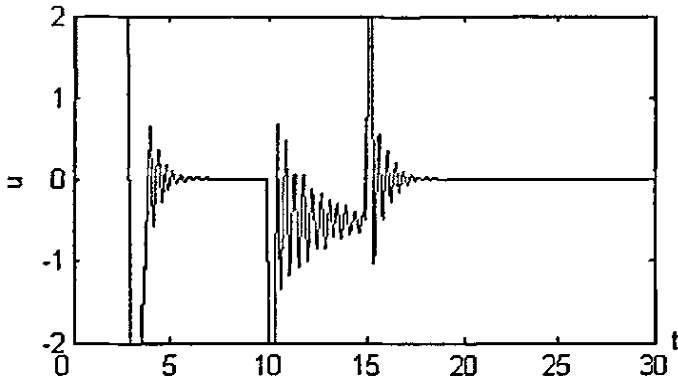


图 4-9 系统参数存在扰动时的控制输入

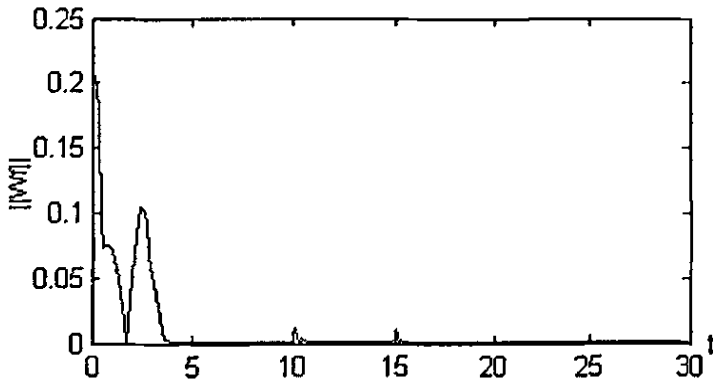


图4-10 系统参数存在扰动时 $\|w_f\|$ 的变化

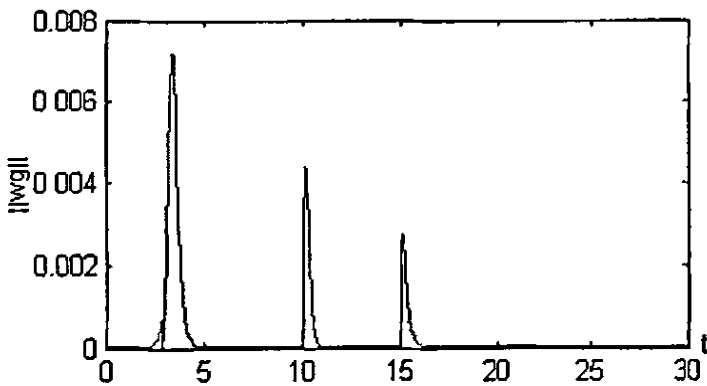


图4-11 系统参数存在扰动时 $\|w_g\|$ 的变化

仿真曲线说明，在存在干扰的情况下，闭环系统仍能稳定的工作

在不稳定的平衡点，这说明虽然在设计控制器时我们没有考虑到干扰的情况，但所得到的闭环控制系统有良好的适应性和鲁棒性，同时权值和控制信号也是有界的。

第 5 章 MIMO 系统小波神经网络自适应控制器设计

非线性多变量系统在系统和控制领域具有一定的挑战性，神经网络控制提供了一种有效的方法来解决非线性系统问题，特别是在对控制对象不完全认知或不确定性进行的情况，但大多数基于神经网络的系统设计都是针对单输入单输出系统，对于 MIMO 系统的讨论较少也存在一定的难度，对于 MIMO 仿射非线性系统，人们已经提出了一些解决方法^{[49][55~57]}。

在上一章，我们研究了一类 SISO 仿射非线性系统基于小波神经网络的自适应控制问题。小波神经网络的良好非线性和并行处理能力使得其能够处理复杂系统，应用小波神经网络作为控制器可以避免繁琐的回归矩阵计算，且不需要整个动态模型结构的先验知识。本章在第四章的基础上研究了 MIMO 非线性系统基于小波神经网络的自适应控制器的设计问题，基于 Lyapunov 函数推导出网络权值的控制率，从而保证了闭环系统的稳定性和权值的有界性，最后通过仿真实例说明了控制器的有效性。

5.1 问题的提出

考虑如下 MIMO 非线性系统：

$$\begin{cases} \dot{x}_1^1 = x_2^1 \\ \dot{x}_{i-1}^1 = x_i^1, i = 3, \dots, r_1 \\ \dot{x}_{r_1}^1 = f_1(x, \eta) + g_{11}(x, \eta)u_1 + g_{12}(x, \eta)u_2 \cdots + g_{1m}(x, \eta)u_m \\ \vdots \\ \dot{x}_1^m = x_2^m \\ \dot{x}_{i-1}^m = x_i^m, i = 3, \dots, r_m \\ \dot{x}_{r_m}^m = f_m(x, \eta) + g_{m1}(x, \eta)u_1 + g_{m2}(x, \eta)u_2 \cdots + g_{mm}(x, \eta)u_m \\ \dot{\eta} = f_\eta(x, \eta) + g_{\eta 1}(x, \eta)u_1 + g_{\eta 2}(x, \eta)u_2 \cdots + g_{\eta m}(x, \eta)u_m \end{cases} \quad (5.1.1)$$

$$y = [x_1^1, \dots, x_1^m]^T$$

其中, $x \in R^p$, $p = r_1 + r_2 + \dots + r_m$, $y \in R^m$, $u \in R^m$, 系统为 n 维, $r = [r_1, r_2, \dots, r_m]$ 为相对阶。

$$\text{令: } A = \text{diag}(A_1, A_2, \dots, A_m), \quad A_i = \begin{bmatrix} 0 & I_{r_i-1} \\ 0 & 0 \end{bmatrix}, \quad i = 1, 2, \dots, m$$

$$B = [B_1, B_2, \dots, B_m]^T, \quad B_i \text{ 为 } p \text{ 阶单位矩阵的第 } \sum_{j=1}^i r_j \text{ 列,}$$

$$i = 1, 2, \dots, m$$

$$F = [f_1, f_2, \dots, f_m]^T$$

$$G = \begin{bmatrix} g_{11} & \dots & g_{1m} \\ \vdots & \ddots & \vdots \\ g_{m1} & \dots & g_{mm} \end{bmatrix}$$

则上述系统 (5.1.1) 可化为如下形式:

$$\begin{cases} \dot{x} = Ax + B[F(x, \eta) + G(x, \eta)u] \\ \dot{\eta} = f_{\eta}(x, \eta) + g_{\eta 1}(x, \eta)u_1 + g_{\eta 2}(x, \eta)u_2 \dots + g_{\eta m}(x, \eta)u_m \\ y = [x_1^1, \dots, x_1^m]^T \end{cases} \quad (5.1.2)$$

假设 1: 在闭集 $\Omega_{x, \eta} \subset R^n$ 上, $G(x, \eta)$ 非奇异, 即 $|G(x, \eta)| \neq 0$ 。

假设 2: 系统 (5.1.2) 要跟踪如下的期望模型:

$$\dot{z} = Az + B \begin{bmatrix} -\alpha_1 z + v_1 \\ \vdots \\ -\alpha_m z + v_m \end{bmatrix} \quad (5.1.3)$$

$$\alpha_1 = [\alpha_{11}, \dots, \alpha_{1r_1}, 0, \dots, 0], \dots, \alpha_m = [0, \dots, 0, \alpha_{m1}, \dots, \alpha_{mr_m}] \quad (5.1.4)$$

其中 α_i 使得:

$$\begin{aligned} \Gamma_1(s) &= s^n + \alpha_{1n} s^{n-1} + \dots + \alpha_{12} s + \alpha_{11} \\ &\vdots \\ \Gamma_m(s) &= s^{r_m} + \alpha_{mr_m} s^{r_m-1} + \dots + \alpha_{m2} s + \alpha_{m1} \end{aligned}$$

均为 Hurwitz 的, v_1, v_2, \dots, v_m 为有界的参考输入。

控制目的是: 设计一自适应控制器, 使得系统 (5.1.2) 的状态 x 在容许的精度内能渐进的跟踪 (5.1.3) 的轨线 z 。

5.2 小波神经网络自适应控制器设计

假设 3: 系统 (5.1.2) 状态 x 的稳态值为 \bar{x} , 且系统的零动态 $\dot{\eta} = f_\eta(\bar{x}, \eta) + g_{\eta 1}(\bar{x}, \eta)u_1 + g_{\eta 2}(\bar{x}, \eta)u_2 \dots + g_{\eta m}(\bar{x}, \eta)u_m$ 是局部渐进稳定的。

当 $f_i(x, \eta)$, $g_j(x, \eta)$ ($i=1,2,\dots,m, j=1,2,\dots,m$) 未知时, 可用小波神经网络对它们进行逼近, 选用同第三章所述结构形式的连续小波神经网络 (表达式见第三章, 此处不再赘述), 且假设在理想情况下有:

$$F = \hat{F}(W_f^*, x, \eta) + \varepsilon_f = [\Phi^T w_{f_1}^*, \dots, \Phi^T w_{f_m}^*]^T + \varepsilon_f \quad (5.2.1)$$

$$G = \hat{G}(W_g^*, x, \eta) + \varepsilon_g = \begin{bmatrix} \Phi^T w_{g_{11}}^* & \dots & \Phi^T w_{g_{1m}}^* \\ \vdots & \ddots & \vdots \\ \Phi^T w_{g_{m1}}^* & \dots & \Phi^T w_{g_{mm}}^* \end{bmatrix} + \varepsilon_g \quad (5.2.2)$$

其中: $w_{f_i}^*$ 和 $w_{g_j}^*$ ($i, j=1,2,\dots,m$) 为小波神经网络的理想权值向量, 仅用作理论分析, 实际上不可得。 $\Phi \in R^{(N_w+1) \times N_w}$ 为网络的小波函数输出向量, N_w 为网络的小波基函数个数, ε_f 和 ε_g 为相应维数的网络逼近误差。

记 w_{f_i} 和 w_{y_i} 为理想权值向量的估计值, \hat{W}_f 和 \hat{W}_y 分别为 w_{f_i} 和 w_{y_i} 组成的集合。

定义:

$$\begin{aligned} \tilde{w}_{f_i} &= w_{f_i} - w_{f_i}^* \\ \tilde{w}_{y_i} &= w_{y_i} - w_{y_i}^* \end{aligned} \quad (5.2.3)$$

根据反馈线性化理论, 在满足假设 1、假设 2 和假设 3 的情况下, 可设计小波神经网络控制器如下:

$$\hat{u} = \hat{G}^{-1}(\hat{W}_g, x, \eta) \left\{ -\hat{F}(\hat{W}_f, x, \eta) + \begin{bmatrix} -\alpha_1 x + v_1 \\ \vdots \\ -\alpha_m x + v_m \end{bmatrix} \right\} \quad (5.2.4)$$

则有:

$$\begin{aligned} \begin{bmatrix} \dot{x}_1^1 \\ \vdots \\ \dot{x}_m^m \end{bmatrix} &= \begin{bmatrix} \Phi^T w_{f_1}^* \\ \vdots \\ \Phi^T w_{f_m}^* \end{bmatrix} + \varepsilon_f + \hat{G}(W_g^*, x, \eta) \hat{u} + \varepsilon_g \hat{u} \\ &= \begin{bmatrix} \Phi^T w_{f_1}^* \\ \vdots \\ \Phi^T w_{f_m}^* \end{bmatrix} + \varepsilon_f + \varepsilon_g \hat{u} + [\hat{G}(W_g^*, x, \eta) - \hat{G}(\hat{W}_g, x, \eta)] \hat{u} + \hat{G}(\hat{W}_g, x, \eta) \hat{u} \\ &= \begin{bmatrix} \Phi^T w_{f_1}^* \\ \vdots \\ \Phi^T w_{f_m}^* \end{bmatrix} + \left(\begin{bmatrix} \Phi^T w_{11}^* & \cdots & \Phi^T w_{1m}^* \\ \vdots & \ddots & \vdots \\ \Phi^T w_{m1}^* & \cdots & \Phi^T w_{mm}^* \end{bmatrix} - \begin{bmatrix} \Phi^T w_{11} & \cdots & \Phi^T w_{1m} \\ \vdots & \ddots & \vdots \\ \Phi^T w_{m1} & \cdots & \Phi^T w_{mm} \end{bmatrix} \right) \hat{u} \\ &\quad + \varepsilon_f + \varepsilon_g \hat{u} - \hat{F}(\hat{W}_f, x, \eta) + \begin{bmatrix} -\alpha_1 x + v_1 \\ \vdots \\ -\alpha_m x + v_m \end{bmatrix} \\ &= - \begin{bmatrix} \Phi^T \tilde{w}_{f_1} \\ \vdots \\ \Phi^T \tilde{w}_{f_m} \end{bmatrix} - \begin{bmatrix} \sum_{i=1}^m \Phi^T \tilde{w}_{1i} \hat{u}_i \\ \vdots \\ \sum_{i=1}^m \Phi^T \tilde{w}_{mi} \hat{u}_i \end{bmatrix} + \varepsilon_f + \varepsilon_g \hat{u} + \begin{bmatrix} -\alpha_1 x + v_1 \\ \vdots \\ -\alpha_m x + v_m \end{bmatrix} \end{aligned}$$

令：

$$\begin{bmatrix} \Delta_1 \\ \vdots \\ \Delta_m \end{bmatrix} = - \begin{bmatrix} \Phi^T \tilde{w}_{f_1} - \sum_{i=1}^m \Phi^T \tilde{w}_{1i} \hat{u}_i \\ \vdots \\ \Phi^T \tilde{w}_{f_m} - \sum_{i=1}^m \Phi^T \tilde{w}_{mi} \hat{u}_i \end{bmatrix} \quad (5.2.5)$$

则：

$$\begin{bmatrix} \dot{x}'_1 \\ \vdots \\ \dot{x}'_m \end{bmatrix} = \begin{bmatrix} \Delta_1 \\ \vdots \\ \Delta_m \end{bmatrix} + \varepsilon_f + \varepsilon_g \hat{u} + \begin{bmatrix} -\alpha_1 x + v_1 \\ \vdots \\ -\alpha_m x + v_m \end{bmatrix} \quad (5.2.6)$$

定义：

$$e = x - z \quad (5.2.7)$$

为系统与模型的跟踪误差。

则：

$$\begin{aligned} \dot{e} = \dot{x} - \dot{z} &= Ax + B(F(x, \eta) + G(x, \eta)u) - Az - B \begin{bmatrix} -\alpha_1 z + v_1 \\ \vdots \\ -\alpha_m z + v_m \end{bmatrix} \\ &= A(x - z) + B \left(\begin{bmatrix} \Delta_1 \\ \vdots \\ \Delta_m \end{bmatrix} + \varepsilon_f + \varepsilon_g \hat{u} + \begin{bmatrix} -\alpha_1 x + v_1 \\ \vdots \\ -\alpha_m x + v_m \end{bmatrix} \right) - B \begin{bmatrix} -\alpha_1 z + v_1 \\ \vdots \\ -\alpha_m z + v_m \end{bmatrix} \\ &= A(x - z) + B \left(\begin{bmatrix} \Delta_1 \\ \vdots \\ \Delta_m \end{bmatrix} + \varepsilon_f + \varepsilon_g \hat{u} \right) - B \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_m \end{bmatrix} (x - z) \end{aligned}$$

令：

$$A_e = A - B \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_m \end{bmatrix}$$

则：

$$\dot{e} = A_e e + B \begin{bmatrix} \Delta_1 \\ \vdots \\ \Delta_m \end{bmatrix} + B(\varepsilon_f + \varepsilon_g \hat{u}) \quad (5.2.8)$$

定理：系统 (5.1.2) 在满足假设 1、假设 2 的条件下，选取控制器为 (5.2.4)，并采用如下的自适应律：

$$\begin{aligned} \dot{w}_{f_i} &= \dot{\tilde{w}}_{f_i} = \gamma_f [2e^T PB_i \Phi(x) - 2k \|PB\| \cdot \|e\| w_{f_i}] \\ \dot{w}_{g_j} &= \dot{\tilde{w}}_{g_j} = \gamma_g [2e^T PB_i \Phi(x) \hat{u}_j - 2k \|PB\| \cdot \|e\| w_{g_j}], i, j = 1, 2, \dots, m \end{aligned} \quad (5.2.9)$$

其中， γ_f, γ_g 均为大于零的权值增益，则所得的小波神经网络自适应闭环控制系统是 UUB 的。

证明：选择正定的对称矩阵 P, Q 使得 $PA_e + A_e^T P = -Q$ ，并构造如下的 Lyapunov 函数：

$$V = e^T P e + \frac{1}{2} \sum_{i=1}^m \tilde{w}_{f_i}^T \gamma_f^{-1} \tilde{w}_{f_i} + \frac{1}{2} \sum_{i=1}^m \sum_{j=1}^m \tilde{w}_{g_j}^T \gamma_g^{-1} \tilde{w}_{g_j} \quad (5.2.10)$$

则 V 对时间的导数为：

$$\dot{V} = -e^T Q e + 2e^T PB \left(\begin{bmatrix} \Delta_1 \\ \vdots \\ \Delta_m \end{bmatrix} + (\varepsilon_f + \varepsilon_g \hat{u}) \right) + \sum_{i=1}^m \dot{\tilde{w}}_{f_i}^T \gamma_f^{-1} \tilde{w}_{f_i} + \sum_{i=1}^m \sum_{j=1}^m \dot{\tilde{w}}_{g_j}^T \gamma_g^{-1} \tilde{w}_{g_j} \quad (5.2.11)$$

代入 (5.2.9) 式中的小波神经网络权值调整律得：

$$\begin{aligned} \dot{V} &= -e^T Q e + 2e^T PB \left(\begin{bmatrix} \Delta_1 \\ \vdots \\ \Delta_m \end{bmatrix} + (\varepsilon_f + \varepsilon_g \hat{u}) \right) + 2 \sum_{i=1}^m e^T PB_i \Phi + 2 \sum_{i=1}^m \sum_{j=1}^m e^T PB_i \Phi \hat{u}_j \\ &\quad - 2 \sum_{i=1}^m \sum_{j=1}^m k \|PB\| \cdot \|e\| w_{g_j}^T \tilde{w}_{g_j} - 2 \sum_{i=1}^m k \|PB\| \cdot \|e\| w_{f_i}^T \tilde{w}_{f_i} \\ &= -e^T Q e + 2e^T PB(\varepsilon_f + \varepsilon_g \hat{u}) - 2 \sum_{i=1}^m k \|PB\| \cdot \|e\| w_{f_i}^T \tilde{w}_{f_i} - 2 \sum_{i=1}^m \sum_{j=1}^m k \|PB\| \cdot \|e\| w_{g_j}^T \tilde{w}_{g_j} \end{aligned}$$

$$(5.2.12)$$

下面对上式进行放大不等式：

(1) 记 $\lambda_{\min}(Q)$ 为矩阵 Q 的最小特征值，由第二章的预备知识得：

$$-e^T Q e \leq -\lambda_{\min}(Q) \|e\|^2 \quad (5.2.13)$$

(2) 记 $\max_{x \in \Omega_x, \eta} \|\varepsilon_f + \varepsilon_g \hat{u}\| = \varepsilon_{wmm}$ ，所以

$$2e^T P B (\varepsilon_f + \varepsilon_g \hat{u}) \leq 2 \|e\| \cdot \|P B\| \varepsilon_{wmm} \quad (5.2.14)$$

(3) $-w_{f_i}^T \tilde{w}_{f_i} = -w_{f_i}^T \tilde{w}_{f_i} + (w_{f_i}^{*T} \tilde{w}_{f_i} - w_{f_i}^{*T} \tilde{w}_{f_i})$

$$= -(w_{f_i} - w_{f_i}^*)^T \tilde{w}_{f_i} - w_{f_i}^{*T} \tilde{w}_{f_i}$$

$$= -\tilde{w}_{f_i}^T \tilde{w}_{f_i} - w_{f_i}^{*T} \tilde{w}_{f_i}$$

$$\leq -\left(\|\tilde{w}_{f_i}\|^2 - \|w_{f_i}^*\| \cdot \|\tilde{w}_{f_i}\|\right)$$

$$= -\left(\|\tilde{w}_{f_i}\| - \frac{1}{2}\|w_{f_i}^*\|\right)^2 + \frac{1}{4}\|w_{f_i}^*\|^2 \quad (5.2.15)$$

同样有：

$$-w_{g_j}^T \tilde{w}_{g_j} = -(\tilde{w}_{g_j} + w_{g_j}^*)^T \tilde{w}_{g_j} \leq -\|\tilde{w}_{g_j}\|^2 + 2\|w_{g_j}^*\| \cdot \|\tilde{w}_{g_j}\|$$

$$= -\left(\|\tilde{w}_{g_j}\| - \frac{1}{2}\|w_{g_j}^*\|\right)^2 + \frac{1}{4}\|w_{g_j}^*\|^2 \quad (5.2.16)$$

所以：

$$\begin{aligned} \dot{V} \leq & -\lambda_{\min}(Q) \|e\|^2 + 2\|e\| \cdot \|P B\| \varepsilon_{wmm} - 2k \|P B\| \cdot \|e\| \cdot \sum_{i=1}^m \left[\left(\|\tilde{w}_{f_i}\| - \frac{1}{2}\|w_{f_i}^*\| \right)^2 \right. \\ & \left. - \frac{1}{4}\|w_{f_i}^*\|^2 \right] - 2k \|P B\| \cdot \|e\| \cdot \sum_{i=1}^m \sum_{j=1}^m \left[\left(\|\tilde{w}_{g_j}\| - \frac{1}{2}\|w_{g_j}^*\| \right)^2 - \frac{1}{4}\|w_{g_j}^*\|^2 \right] \end{aligned}$$

$$\begin{aligned}
 &= -\lambda_{\min}(Q)\|e\|^2 - 2k\|PB\|\cdot\|e\|\cdot\left[\sum_{i=1}^m\left(\|\tilde{w}_{f_i}\| - \frac{1}{2}\|w_{f_i}^*\|\right)^2 + \sum_{i=1}^m\sum_{j=1}^m\left(\|\tilde{w}_{y_j}\| - \frac{1}{2}\|w_{y_j}^*\|\right)^2\right] \\
 &\quad + 2\|e\|\cdot\|PB\|\left(\varepsilon_{wnn} + \sum_{i=1}^m\frac{1}{4}\|w_{f_i}^*\|^2 + \sum_{i=1}^m\sum_{j=1}^m\frac{1}{4}\|w_{y_j}^*\|^2\right) \tag{5.2.17}
 \end{aligned}$$

记： $\varepsilon = \varepsilon_{wnn} + \sum_{i=1}^m\frac{1}{4}\|w_{f_i}^*\|^2 + \sum_{i=1}^m\sum_{j=1}^m\frac{1}{4}\|w_{y_j}^*\|^2$ 则有 $\varepsilon \geq 0$ ，所以 (5.2.17)

变为：

$$\begin{aligned}
 \dot{V} &\leq -\lambda_{\min}(Q)\|e\|^2 - 2k\|PB\|\cdot\|e\|\cdot\left[\sum_{i=1}^m\left(\|\tilde{w}_{f_i}\| - \frac{1}{2}\|w_{f_i}^*\|\right)^2 + \sum_{i=1}^m\sum_{j=1}^m\left(\|\tilde{w}_{y_j}\| - \frac{1}{2}\|w_{y_j}^*\|\right)^2\right] \\
 &\quad + 2\|e\|\cdot\|PB\|\varepsilon
 \end{aligned}$$

由于 Q 为正定的对称实矩阵，故有 $\lambda_{\min}(Q) \geq 0$ ，且 k 为一正常数。

要想使得 $\dot{V} < 0$ 成立，我们可以分下面几种情况讨论。

(1) 当 $-\lambda_{\min}(Q)\|e\|^2 + 2\|e\|\cdot\|PB\|\varepsilon < 0$ 时，有 $\|e\| > \frac{2\varepsilon\|PB\|}{\lambda_{\min}(Q)}$ ；

(2) 当 $-2k\|PB\|\cdot\|e\|\cdot\sum_{i=1}^m\left(\|\tilde{w}_{f_i}\| - \frac{1}{2}\|w_{f_i}^*\|\right)^2 + 2\|e\|\cdot\|PB\|\varepsilon < 0$ 时，有

$\sum_{i=1}^m\left(\|\tilde{w}_{f_i}\| - \frac{1}{2}\|w_{f_i}^*\|\right)^2 > \frac{\varepsilon}{k}$ ，所以只需 $\|w_{f_i}\| > \sqrt{\frac{\varepsilon}{k \cdot m}} - \frac{1}{2}\|w_{f_i}^*\|$ 即可保证上式成立。

(3) 当 $-2k\|PB\|\cdot\|e\|\cdot\sum_{i=1}^m\sum_{j=1}^m\left(\|\tilde{w}_{y_j}\| - \frac{1}{2}\|w_{y_j}^*\|\right)^2 + 2\|e\|\cdot\|PB\|\varepsilon < 0$ 时，同

上只需 $\|w_{y_j}\| > \sqrt{\frac{\varepsilon}{k \cdot m^2}} - \frac{1}{2}\|w_{y_j}^*\|$ 即可保证上式成立。

由上面分析可知，在上述 3 种情况下均可设计可行的小波神经网络

络自适应控制器，即保证 $\dot{V} < 0$ ，并且当系统的零动态（局部）渐近稳定，则所设计的自适应控制系统是一致最终有界（UUB）的。

当 $\hat{G}(\hat{W}_g, x, \eta)$ 不可逆时，可令 (5.2.4) 式中的 $\hat{G}^{-1}(\hat{W}_g, x, \eta)$ 为 $(\hat{G}^T \hat{G} + RI)^{-1} \hat{G}^T$ ， I 为相应维数的单位阵， R 为使得上面的求逆存在的小常数。

5.3 仿真分析

下面以一个两输入两输出对象为例进行自适应控制器设计研究，假设研究对象的数学模型如下：

$$\begin{cases} \dot{x}_1 = x_2 + (1 + x_1^2)u_1 + x_2 u_2 \\ \dot{x}_2 = -x_1(x_2 + 1) - x_2 u_1 + u_2 \end{cases} \quad (5.3.1)$$

即：

$$\begin{aligned} \dot{x} &= F(x) + G(x)u \\ y &= x \end{aligned} \quad (5.3.2)$$

其中，

$$F(x) = \begin{bmatrix} x_2 \\ -x_1(x_2 + 1) \end{bmatrix}, G(x) = \begin{bmatrix} 1 + x_1^2 & x_2 \\ -x_2 & 1 \end{bmatrix} \quad (5.3.3)$$

显然，在闭集 $\Omega_x \subset R^n$ 上， $G(x)$ 非奇异，即满足假设 1。

易知上述系统的相对阶为： $[r_1 \ r_2] = [1 \ 1]$ 。由第二部分分析可设计如下的小波神经网络控制器：

$$\hat{u} = G_{wnn}^{-1} \left\{ -F_{wnn} + \begin{bmatrix} -\alpha_1 x + v_1 \\ -\alpha_2 x + v_2 \end{bmatrix} \right\} \quad (5.3.4)$$

由 $A = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$, $B = [B_1 \ B_2] = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$, $K = \begin{bmatrix} -\alpha_1 \\ -\alpha_2 \end{bmatrix}$, $Q = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ 可以

求解 $A_e = A - BK, PA_e + A_e^T P = -Q$ 。

则由第二部分的结论知若令小波神经网络的权值调整律如下，

$$\begin{aligned}
 \dot{\tilde{w}}_{F1} &= \gamma_F [e^T PB_1 \Phi(x) - k \|PB\| \cdot \|e\| w_{F1}] \\
 \dot{\tilde{w}}_{F2} &= \gamma_F [e^T PB_2 \Phi(x) - k \|PB\| \cdot \|e\| w_{F2}] \\
 \dot{\tilde{w}}_{G11} &= \gamma_G [e^T PB_1 \Phi(x) u_1 - k \|PB\| \cdot \|e\| w_{G11}] \\
 \dot{\tilde{w}}_{G12} &= \gamma_G [e^T PB_1 \Phi(x) u_2 - k \|PB\| \cdot \|e\| w_{G12}] \\
 \dot{\tilde{w}}_{G21} &= \gamma_G [e^T PB_2 \Phi(x) u_1 - k \|PB\| \cdot \|e\| w_{G21}] \\
 \dot{\tilde{w}}_{G22} &= \gamma_G [e^T PB_2 \Phi(x) u_2 - k \|PB\| \cdot \|e\| w_{G22}]
 \end{aligned} \tag{5.3.5}$$

则系统是一致最终有界的，其中 $e = [(x_1 - y_{1d}) \quad (x_2 - y_{2d})]^T$ 。

系统 (5.3.1) 在零输入下，相平面如图所示：

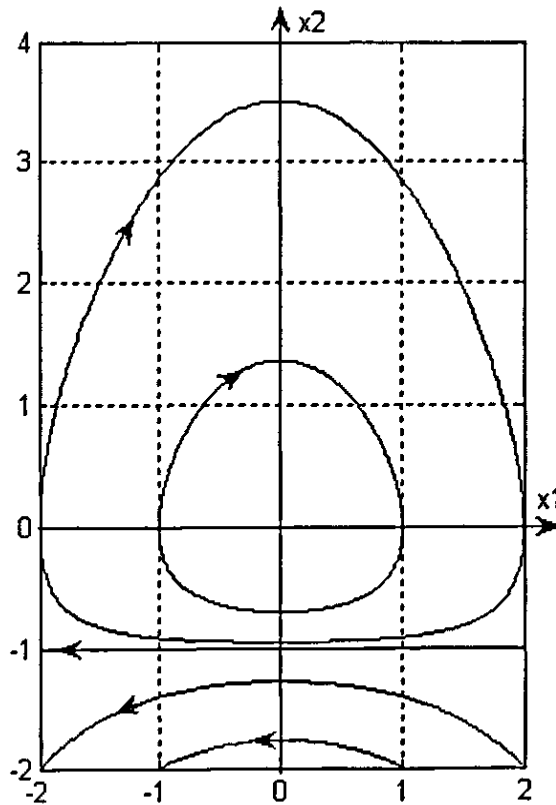


图5-1 被控对象相平面图

系统在 $x_2 > -1$ 的区域均为临界稳定，在 $x_2 \leq -1$ 的区域不稳定。

控制的目的是使系统从某一临界稳定点运动到平衡点 $(0, 0)$ 。

仿真参数设定：

参考信号： $y_d(t) = [0, 0]^T$ ；

系统的初始状态： $x(0) = [0.6, 0.4]^T$ ；

控制器的参数： $\alpha_1 = [1, 0]$ ； $\alpha_2 = [0, 10]$ ； $\gamma_F = 2, \gamma_G = 2, k = 0.6$ ；

其中 α_1, α_2 的取值满足假设 2。

两个小波神经网络 WNN_f, WNN_g 的隐层分别取 9 个小波神经元，平移因子和伸缩因子预先给定，权值初始化为小的随机数，网络的输入为输出 y ，采样时间 $T_s = 0.01$ 。

仿真结果分析：系统的输出响应、控制作用、权值范数的相应曲线如图 5-2 到图 5-5 (k 表示采样时刻)。

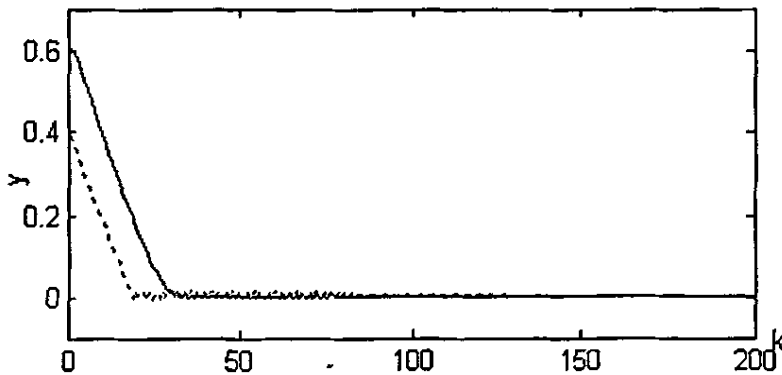


图5-2 控制器跟踪阶跃信号输出响应

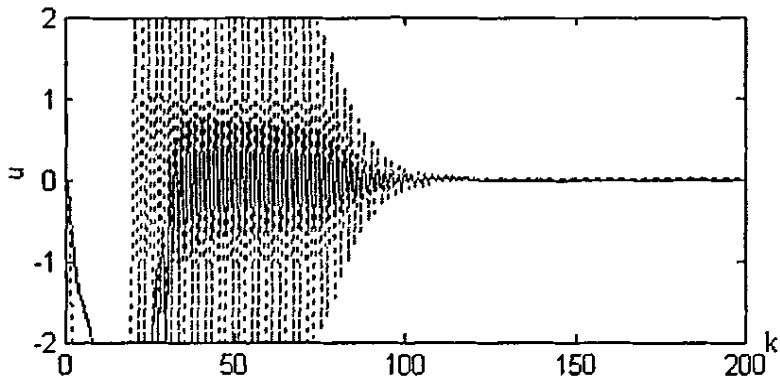


图5-3 跟踪阶跃信号的控制输入

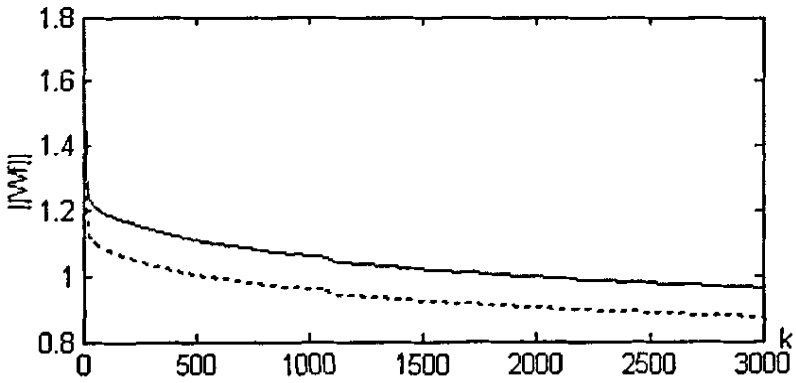


图5-4 跟踪阶跃信号 WNN_F 的权值范数

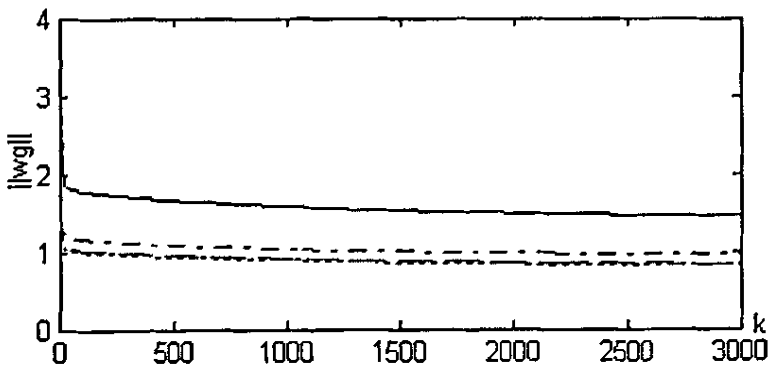


图5-5 跟踪阶跃信号 WNN_G 的权值范数

可以看出设计的小波神经网络控制器有很好的伺服性能，输出能很好的跟踪期望值，权值和控制信号也是有界的。与SISO自适应控制

不同的是，权值范数没有减小到0。在式 (5.3.5) 中可以看到，当残差 e 减小时， $\dot{\tilde{w}}_F, \dot{\tilde{w}}_G$ 减小，当 $e=0$ 时， $\dot{\tilde{w}}_F=0, \dot{\tilde{w}}_G=0$ 。而在式 (4.3.9) 中，当 $e=0$ 时， $\dot{\tilde{w}}_F, \dot{\tilde{w}}_G$ 仍然与 w_F, w_G 负相关，使得 w_F, w_G 逐渐收敛到0。

第 6 章 总结

本论文基于小波神经网络,研究了非线性系统的自适应控制问题,并且提出了新的在线训练算法,然后把所设计的控制器进行仿真实验和应用于一化工过程 CSTR 中。经过以上的研究工作,得到如下结论:

(1)连续小波神经网络的训练对初始参数和步长的选取十分敏感,选取合适的参数和步长会使得网络快速收敛,虽然依靠时频分析理论可以对初始参数进行估计,但目前并没有理论上的确定方法,仍需借助一定的经验和尝试来选择。

(2)所提出的在线混合训练算法将改进 BP 算法的思想用于对平移因子和伸缩因子的训练,同时利用递推最小二乘算法对权值进行训练,二者的有机结合,使得网络收敛速度加快,实时性强,便于在线控制的应用。

(3)小波神经网络用于自适应控制中不仅对系统参数可以进行估计,而且还可以逼近系统中不确定的非线性映射,控制器的参数是直接设计的,不用预先对系统模型进行辨识,因此控制是直接自适应的,且算法直接简单。

(4)对于一类仿射非线性系统,两种控制器的设计均构造了 Lyapunov 函数,经逐步理论推导,得出控制器的权值在线调整规律。由于基于 Lyapunov 稳定性理论,所以能保证闭环系统的稳定性,并可证明闭环系统的所有信号都是一致最终有界的,并且得到的闭环控制系统具有良好的适应性和鲁棒性。

(5)对于仿射系统仅仅考虑了小波神经网络中权值的自适应律的设计问题,若再考虑基于 Lyapunov 稳定性的平移和伸缩因子的自适应律的控制,控制效果会更好,但对于这两个因子的设计相对比较困难,这将在以后的工作中继续研究。

(6) 由于对一般非线性系统研究的难度很大, 所以本论文只以一类 NARMAX 模型和一类仿射非线性系统为例进行了研究。如何把对所设计的控制器进行拓展以应用于更广泛的非线性系统, 是以后深入研究的重点。

总之, 大量的理论研究和仿真结果以及许多成功的应用表明, 小波神经网络在非线性控制中具有独特的优势和广阔的前景, 但对于小波神经网络的基础理论分析还不完善, 要对其进行深入研究, 还有赖于小波理论和非线性科学的最新进展。

参考文献

- [1] 韩曾晋. 自适应控制. 北京:清华大学出版社, 1995
- [2] 白玫. 智能控制理论综述. 华北水利水电学报, 2002; 23(1): 58-62
- [3] 焦李成. 神经网络的应用与实现. 西安:西安电子科技大学出版社, 1993
- [4] 王永骥, 涂键. 神经元网络控制. 北京:机械工业出版社, 1998
- [5] 汪镭. 人工神经网络理论在控制领域中的应用综述. 同济大学学报, 2001; 29(3): 357-361
- [6] Venugopal G. Krishnapura, Arthur Jutan. A neural adaptive controller. *Chemical Engineering Science*, 2000; 55: 3803-3812
- [7] Xiang-sun Zhang, Xin-jian Zhuo, Zhu-jun Jing. An adaptive neural network model for nonlinear programming problems. *Acta Mathematicae Applicatae Sinica*, 2002; 18(3): 377-388
- [8] Shuzhi S. Ge and cong Wang. Direct adaptive NN control of a class of nonlinear systems. *IEEE transaction on neural networks*, January 2002; 13(1)
- [9] S. S. Ge and C. C. Hang. Nonlinear adaptive control using neural networks and its application to CSTR systems. *Journal of Process Control*, 1998; 9: 313-323
- [10] Dan Wang, Jie Huang. Adaptive neural network control for a class of uncertain nonlinear systems in pure-feedback form. *Automatic*, 2002; 38: 1365-1372
- [11] G. A. Rorithakis. Stable adaptive neuro-control design via lyapunov function derivative estimation. *Automatic*, 2001; 37: 1213-1221
- [12] J. Stoev and J. Y. Choi. Adaptive control for output feedback nonlinear systems in the presence of modeling error. *Automatic*, 2002; 38:

1761-1767

[13] J. Q. Gong and Bin Yao. Neural network adaptive robust control of nonlinear systems in semi-strict feedback form. *Automatic*, 2001; 37: 1149-1160

[14] Zhang Q., Benveniste A.. Wavelet networks. *IEEE Transactions on Neural Networks*, 1992; 3(6): 899-898

[15] Y. Oussar, I. Rivals, L. Personnaz. Training wavelet neural networks for nonlinear dynamic input-output modeling. *Neuro computing*, 1998; 20: 173-188

[16] Huang Dexian, Wang Jingchun, Jin Yihui. Application research of wavelet neural networks in process control, *Journal of Tsinghua university (science and technology)*, 1999; 1

[17] Yusuf Oysal, A. Serdar Yilmaz, Etem Koklukaya. A dynamic wavelet network based adaptive load frequency control in power system. *Electrical Power and Energy Systems*, 2005; 27: 21-29

[18] Qing-xiang Guo, Lei Liu, Wen-sheng Cai. Driving force prediction for inclusion complexation of α -cyclodextrin with benzene derivatives by a wavelet neural network. *Chemical Physics Letters*, 1998; 190: 514-518

[19] A. Benveniste and K. J. Astrom. Meeting the challenge of computer science in the industrial application of control. *Automatica*, 1993; (29)5: 1169-1175

[20] Daniel Ho W C, Jinhua Xu. Wavelet neural network for continuous time nonlinear system identification, 14th World Congress of IFAC, Beijing, 1999

[21] 王忠勇, 张端金, 方浩等. 一种基于小波神经网络的自校正方法. *控制理论与应用*, 2001; 18(2): 270-272

[22] 方浩, 薛培鼎. 一种基于小波神经网络的自适应控制方法. *西安交通大学学报*, 2000; 34(2)

[23] 黄显林, 王永福等. 基于小波神经网络的非线性系统预测控制研

- 究. 黑龙江自动化技术与应用, 1999; 18(5)
- [24] Huang Dexian, Jin Yihui. The application of wavelet network neural network to nonlinear predictive control, Proceeding of the 1997 IEEE international Conference on Neural Network, 1997: 724-727
- [25] 吕柏权. 一种基于小波网络的故障检测方法. 控制理论与应用, 1998, 15(5): 802-805
- [26] 程正兴. 小波分析算法与应用. 西安交通大学出版社, 1998
- [27] [美] 崔锦泰著. 小波分析导论. 程正兴译. 西安交通大学出版社, 1995
- [28] 朱春明, 黄道平. 小波分析在控制中的应用状况. 计算机测量与控制, 2003; 11(10)
- [29] 张红英, 吴斌. 小波神经网络的研究及其展望. 西南工学院学报, 2003; 17(1)
- [30] Jinhua Xu, Daniel W. C. Ho. A basis selection algorithm for wavelet neural networks. Neurocomputing, 2002; 48: 681-689
- [31] 赵学智, 邹春华, 陈统坚等. 小波神经网络的参数初始化研究. 华南理工大学学报, 2003; 31(2)
- [32] 钱峻, 邵惠鹤. 小波神经网络结构设计新算法. 上海交通大学学报, 1999; 33(40): 422-424
- [33] 张治国, 孙德波. 小波神经网络建模研究. 计算机测量与控制, 2003; 11(2)
- [34] 何正友, 钱清泉. 小波神经网络改进结构及其算法[J]. 西南交通大学学报, 1994; 34(4): 436-440
- [35] 万建, 徐德民. 一种小波神经网络结构及其学习算法研究. 系统工程与电子应用, 2002; 24(3)
- [36] 徐晓霞, 陈涛. 基于正交最小二乘算法的小波神经网络. 电子学报, 1998; 26(10)
- [37] 沈雪勤. 能量密度在正交小波神经网络中的应用. 1997年中国神经计算科学大会论文集, 南京, 1997: 613-616

- [38] 王岭, 焦李成. 区间估计的FWNN及其区间学习算法. 电子学报, 1998; 26(4): 41-45
- [39] 张良杰, 李衍达. 基于模糊逻辑与神经网络技术的智能控制器自动设计系统的研究. 信息与控制, 25(3): 129-136
- [40] 张邦礼, 李银国, 曹长修. 小波神经网络的构造及其算法的鲁棒性分析. 重庆大学学报(自然科学版), 1995; 18(6): 88-95
- [41] Pati Y C, Krishnaprasad P S. Analysis and Synthesis of Feedforward Neural Networks Using Discrete Affine Wavelet Transformation[J]. IEEE Trans. On Neural Networks, 1993; 4(1): 73-85
- [42] Tan Y. H. Nonlinear Dynamic system identification based on wavelet neural network. Journal of Guilin Institute of Electronic Technology, 1999; 19(1): 1-6
- [43] Billings S A, Leontaritis I J. Parameter Estimation Techniques for Nonlinear Systems. Proc. 6th IFAC symposium on Identification and System Parameter Estimation, Washington, 1992: 505-510
- [44] Leontaritis I J, Billings S A. Input-output Parametric Models for Nonlinear Systems. Part1: Deterministic Nonlinear Systems; Part2: Stochastic non-linear systems. Int. J. Control, 1985; 41: 303-344; Model selection and validation methods for nonlinear system. Ibid., 1987; 45: 311-341
- [45] S Chen, S A Billings. Representation of Non-linear System: the NARMAX Model. Int. J. Control, 1989; 49: 1013-1032
- [46] Karayiannis N B, Venetsanopoulos A N. Artificial Neural Networks: Learning Algorithms, Performance Evaluation and Applications[M]. New York: Kluwer Academic Publishers, 1993
- [47] T. Zhang, S. S. Ge, C. C. Hang. Neural-based direct adaptive control for a class of general nonlinear systems, Int. J. system science, 1997; 28(10): 1011-1020
- [48] S. S. Ge, C. C. Hang, T. Zhang. Neural-based adaptive control

- design for general nonlinear system and its application to process control. Proc. American Control, Philadelphia, Pennsylvania, 1998; 73-77
- [49] Yu-Min CHENG, Bor-Sen CHEN, Fu-Yuan SHIAU. Adaptive wavelet network control design for nonlinear system. Proc. Natl. Sci. Coune. ROC(A), 1998; 22(6): 783-799
- [50] Man Z H, Wu H R, Palaniswami M. An Adaptive Tracking Controller Using Neural Networks for a class of Nonlinear Ststems. IEEE TNN, 1998; 9(5): 947-954
- [51] 牛玉刚, 赵建丛. 一类非线性系统的自适应神经跟踪控制. 系统工程与电子技术. 2000;22(12)
- [52] 佟绍成, 李庆国, 柴天佑. 基于神经网络的一类非线性系统自适应输出跟踪. 自动化学报. 2005;5
- [53] 何剑春, 王慧燕. CMAC 网络建模在非线性预测控制中的应用. 控制与决策. 2002;17(1)
- [54] Jixian Zhan, Masaru. The Muliti-Step Predictive Control of Nonlinear SISO Processes With a Neural Model Predictive Control Method. Computer chem. 1997; 21(2): 201-210
- [55] 李树荣, 时海涛, 王维波等. 基于径向基神经网络的精馏塔产品质量自适应控制. 石油大学学报（自然科学版）, 2002;26(6)
- [56] S. S. Ge. Robust adaptive NN feedback linearization control of nonlinear systems. International journal of system science, 1996; 27(12): 1327-1338
- [57] Ramon R. Costa, Liu Hsu. Lyapunov-based adaptive control of MIMO systems. Automatic, 2003; 39:1251-1257

致谢

首先感谢我的导师李树荣教授，感谢他对我谆谆教诲和辛勤培养，感谢他在各方面给我的诸多关怀。李老师严谨的治学态度和博大精深的科学知识深深影响着我，他精益求精的科学态度和敏锐的学术思想不断培养我开拓创新的能力，他孜孜不倦的敬业精神及科研上锐意进取的作风将使我终生受益。

感谢实验室各位同学的热心帮助与鼓励，在和他们的共同学习过程中，他们给了我很多有益的启发，使我受益匪浅。

感谢我的父母，感谢他们对我完成学业提供的莫大支持和理解！
感谢在我的求学经历中所有给予我无私帮助和支持的人们！

个人简历、在学期间的研究成果

个人简历:

蔡吉刚, 男, 1978年6月6日出生, 2002年7月毕业于石油大学(华东)信息与控制工程学院自动化系, 获工学学士学位。

在学期间的研究成果:

- [1] 李树荣, 蔡吉刚. 基于小波神经网络的自适应控制器设计. 中国石油大学学报(自然科学版), 已录用.
- [2] 纪友芳, 王平, 蔡吉刚. 智能化参数整定方法研究及应用. 2006 中国控制与决策学术年会. 已录用.