

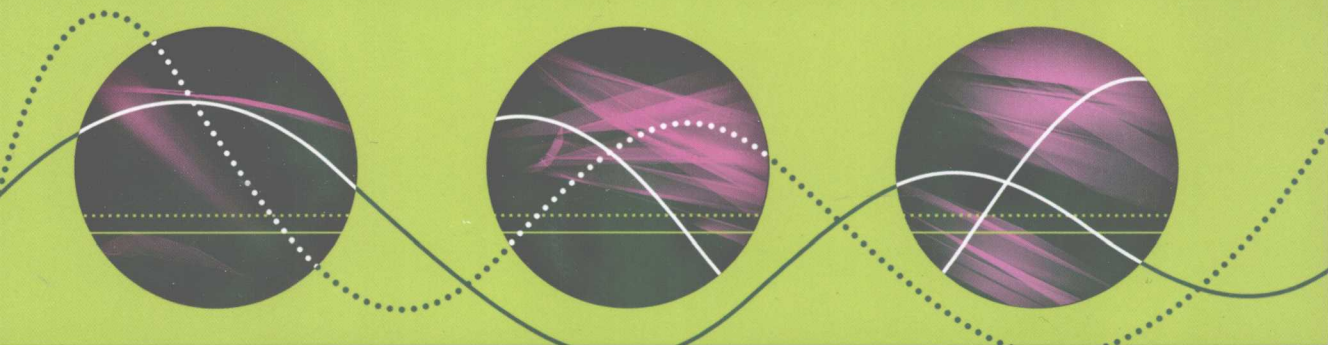
<http://www.phei.com.cn>



自动控制技术应用丛书

神经模糊 控制理论及应用

李国勇 编著



电子工业出版社
PUBLISHING HOUSE OF ELECTRONICS INDUSTRY

自动控制技术应用丛书

神经模糊控制理论及应用

李国勇 编著

電子工業出版社

Publishing House of Electronics Industry

北京·BEIJING

内 容 简 介

本书系统地论述了神经网络和模糊逻辑的基本概念、工作原理、控制算法,以及利用 MATLAB 语言、MATLAB 工具箱函数和 Simulink 对其实现的方法。该书取材先进实用,讲解深入浅出,各章均有相应的例题和练习题,并提供了大量用 MATLAB/Simulink 实现的仿真实例,便于读者掌握和巩固所学知识,让学生在有限的时间内,掌握神经网络和模糊逻辑的基本原理与应用技术。

本书可作为高等院校自动化、计算机、电子与信息、机电工程等专业研究生和高年级本科生的教材,也可作为从事智能控制与智能系统研究、设计和应用的科学技术人员的参考用书。

未经许可,不得以任何方式复制或抄袭本书之部分或全部内容。

版权所有,侵权必究。

图书在版编目(CIP)数据

神经模糊控制理论及应用/李国勇编著. —北京:电子工业出版社,2009.1

(自动控制技术应用丛书)

ISBN 978-7-121-07537-7

I. 神… II. 李… III. 神经网络-应用-模糊控制 IV. TB535

中国版本图书馆 CIP 数据核字(2008)第 157464 号

策划编辑:张 榕

责任编辑:韩玲玲

印 刷:北京市海淀区四季青印刷厂

装 订:涿州市桃园装订有限公司

出版发行:电子工业出版社

北京市海淀区万寿路 173 信箱 邮编 100036

开 本:787×1092 1/16 印张:21.5 字数:550 千字

印 次:2009 年 1 月第 1 次印刷

印 数:4000 册 定价:39.80 元

凡所购买电子工业出版社图书有缺损问题,请向购买书店调换。若书店售缺,请与本社发行部联系,联系及邮购电话:(010)88254888。

质量投诉请发邮件至 zlt@phei.com.cn,盗版侵权举报请发邮件至 dbqq@phei.com.cn。

服务热线:(010)88258888。

前 言

针对神经网络和模糊逻辑的迅速推广应用, MathWorks 公司在其 MATLAB 版中添加了神经网络和模糊逻辑工具箱。该工具箱由长期从事神经网络和模糊逻辑研究与开发工作的有关专家和技术人员编制。工具箱提供了许多进行神经网络和模糊逻辑设计和分析的工具函数, 这给用户带来了极大的方便, 即使不了解算法的本质, 也可以直接应用功能丰富的函数来实现自己的目的。由于其编程简单, 所以给使用者节省了大量的编程时间, 使其能够把更多的精力投入到网络设计而不是具体程序的实现上。

本书本着把当前国际控制界最为流行的面向工程与科学计算的高级语言——MATLAB 与神经网络和模糊逻辑理论结合起来的宗旨而编写。全书共分 2 篇 6 章, 书中详细地介绍了神经网络和模糊逻辑两大智能控制方法, 且分别介绍了三种利用 MATLAB 实现其仿真的方法, 即利用 MATLAB 编程计算、利用 MATLAB 中工具箱函数编程计算和利用 Simulink 仿真计算。其中, 第一篇详细地介绍了十几种常用神经网络的拓扑结构、工作方式和学习算法及其利用 MATLAB 实现的方法、MATLAB 神经网络工具箱函数及图形用户界面的使用方法、基于 Simulink 的神经网络的系统模块和神经网络控制系统。第二篇系统地论述了模糊逻辑理论的基本原理、MATLAB 模糊逻辑工具箱函数及图形用户界面的使用方法、基于 Simulink 的模糊逻辑的系统模块、模糊神经网络和模糊聚类及其利用 MATLAB 实现的方法。

本教材适用学时数为 32 ~ 48 (2 ~ 3 学分), 章节编排具有相对的独立性, 利于教师与学生取舍, 便于不同层次院校的不同专业选用, 以适应不同教学学时的需要。教材内容完善、新颖、有利于学生能力的培养。

本书可作为高等院校自动化、计算机、电子与信息、机电工程、电气信息类其他各专业本科生和研究生教材。鉴于本书的通用性和实用性较强, 故也可作为从事智能控制及相关领域的教学、研究、设计人员和工程技术人员的参考用书。

由于作者水平有限, 书中仍难免有疏漏与不当之处, 故恳请有关专家、同行和广大读者批评指正。

作 者

目 录

第一篇 神经网络理论及其 MATLAB 实现

第 1 章 神经网络理论	(3)
1.1 神经网络的基本概念	(4)
1.1.1 生物神经元的结构与功能特点	(4)
1.1.2 人工神经元模型	(5)
1.1.3 神经网络的结构	(7)
1.1.4 神经网络的工作过程	(8)
1.1.5 神经网络的学习	(8)
1.1.6 神经网络的分类	(11)
1.2 典型神经网络的模型	(11)
1.2.1 MP 模型	(11)
1.2.2 感知机神经网络	(13)
1.2.3 自适应线性神经网络	(18)
1.2.4 BP 神经网络	(20)
1.2.5 径向基神经网络	(29)
1.2.6 竞争学习神经网络	(33)
1.2.7 学习向量量化神经网络	(43)
1.2.8 Elman 神经网络	(44)
1.2.9 Hopfield 神经网络	(45)
1.2.10 Boltzmann 神经网络	(60)
1.3 神经网络的训练	(63)
小结	(67)
练习题 1	(67)
第 2 章 MATLAB 神经网络工具箱	(68)
2.1 MATLAB 神经网络工具箱函数	(68)
2.1.1 神经网络工具箱中的通用函数	(68)
2.1.2 感知机神经网络 MATLAB 函数	(71)
2.1.3 线性神经网络 MATLAB 函数	(84)
2.1.4 BP 神经网络 MATLAB 函数	(91)
2.1.5 径向基神经网络 MATLAB 函数	(104)
2.1.6 自组织神经网络 MATLAB 函数	(110)
2.1.7 学习向量量化神经网络 MATLAB 函数	(126)
2.1.8 Elman 神经网络 MATLAB 函数	(131)

2.1.9 Hopfield 神经网络 MATLAB 函数	(134)
2.1.10 Demos 演示功能	(140)
2.2 MATLAB 神经网络工具箱的图形用户界面	(141)
2.2.1 神经网络编辑器	(141)
2.2.2 神经网络拟合工具	(151)
2.3 基于 Simulink 的神经网络模块	(155)
2.3.1 模块的设置	(155)
2.3.2 模块的生成	(156)
2.4 神经网络在系统预测和故障诊断中的应用	(159)
2.4.1 系统输入/输出数据的处理	(159)
2.4.2 基于神经网络的系统预测	(160)
2.4.3 基于神经网络的故障诊断	(173)
小结	(181)
练习题 2	(181)
第 3 章 神经网络控制系统	(182)
3.1 神经网络控制理论	(182)
3.1.1 神经控制的基本原理	(182)
3.1.2 神经网络在控制中的主要作用	(183)
3.1.3 神经网络控制系统的分类	(184)
3.2 基于 Simulink 的三种典型的神经网络控制系统	(194)
3.2.1 神经网络模型预测控制	(194)
3.2.2 反馈线性化控制	(200)
3.2.3 模型参考控制	(203)
小结	(207)
练习题 3	(207)

第二篇 模糊逻辑理论及其 MATLAB 实现

第 4 章 模糊逻辑理论	(211)
4.1 模糊逻辑理论的基本概念	(211)
4.1.1 模糊集合及其运算	(211)
4.1.2 模糊关系及其合成	(218)
4.1.3 模糊向量及其运算	(220)
4.1.4 模糊逻辑规则	(221)
4.1.5 模糊逻辑推理	(223)
4.2 模糊逻辑控制系统的基本结构	(230)
4.2.1 模糊控制系统的组成	(230)
4.2.2 模糊控制器的基本结构	(231)
4.2.3 模糊控制器的维数	(231)
4.2.4 模糊控制中的几个基本运算操作	(232)

4.3 模糊逻辑控制系统的基本原理	(232)
4.3.1 模糊化运算	(232)
4.3.2 数据库	(233)
4.3.3 规则库	(236)
4.3.4 模糊推理	(238)
4.3.5 清晰化计算	(240)
4.4 离散论域的模糊控制系统的设计	(241)
4.5 具有 PID 功能的模糊控制器	(246)
小结	(247)
练习题 4	(247)
第 5 章 MATLAB 模糊逻辑工具箱	(249)
5.1 MATLAB 模糊逻辑工具箱简介	(249)
5.1.1 模糊逻辑工具箱的功能特点	(249)
5.1.2 模糊推理系统的基本类型	(250)
5.1.3 模糊逻辑系统的构成	(251)
5.2 利用 MATLAB 模糊逻辑工具箱建立模糊推理系统	(251)
5.2.1 模糊推理系统的管理函数	(251)
5.2.2 模糊语言变量及其语言值	(255)
5.2.3 模糊语言变量的隶属度函数	(256)
5.2.4 模糊规则的建立与修改	(264)
5.2.5 模糊推理计算与去模糊化	(268)
5.2.6 模糊推理在控制系统中的应用	(270)
5.3 MATLAB 模糊逻辑工具箱的图形用户界面	(276)
5.3.1 模糊推理系统编辑器(Fuzzy)	(276)
5.3.2 隶属度函数编辑器(Mfedit)	(278)
5.3.3 模糊规则编辑器(Ruleedit)	(278)
5.3.4 模糊规则浏览器(Ruleview)	(279)
5.3.5 模糊推理输入/输出曲面浏览器(Surfview)	(279)
5.4 基于 Simulink 的模糊逻辑的系统模块	(281)
小结	(285)
练习题 5	(285)
第 6 章 模糊神经和模糊聚类及其 MATLAB 实现	(286)
6.1 基于 Mamdani 模型的模糊神经网络	(286)
6.1.1 模糊系统的 Mamdani 模型	(286)
6.1.2 系统结构	(288)
6.1.3 学习算法	(289)
6.2 基于 Takagi-Sugeno 模型的模糊神经网络	(291)
6.2.1 模糊系统的 Takagi-Sugeno 模型	(291)
6.2.2 系统结构	(292)

6.2.3 学习算法	(294)
6.3 自适应神经模糊系统及其 MATLAB 实现	(296)
6.3.1 采用网格分割方式生成模糊推理系统	(296)
6.3.2 自适应神经模糊系统的建模	(297)
6.3.3 自适应神经模糊推理系统在建模中的应用	(299)
6.3.4 自适应神经模糊推理系统的图形界面编辑器	(303)
6.4 模糊聚类及其 MATLAB 实现	(309)
6.4.1 模糊 C-均值聚类函数	(309)
6.4.2 模糊减法聚类函数	(310)
6.4.3 基于减法聚类的模糊推理系统建模函数	(312)
6.4.4 模糊 C-均值和减法聚类的图形用户界面	(314)
小结	(316)
练习题 6	(317)
附录 A MATLAB 程序清单	(318)
附录 B MATLAB 函数一览表	(325)
附录 C MATLAB 函数分类索引	(331)
参考文献	(333)

The background of the page is a light gray with a subtle, repeating pattern of stylized circles and swirls. A dark gray horizontal band runs across the middle of the page, serving as a backdrop for the title text.

第一篇 神经网络理论及其MATLAB实现



第 1 章 神经网络理论

人脑是一部不寻常的智能机,它能以惊人的高速度解释感觉器官传来的含糊不清的信息。它能觉察到喧闹房间内的窃窃私语,能够识别出光线暗淡的胡同中的一张面孔,更能通过不断的学习而产生伟大的创造力。古今中外,许许多多科学家为了揭开大脑机能的奥秘,从不同的角度进行着长期不懈的努力和探索,由此逐渐形成了一个多学科交叉的前沿技术领域——神经网络(Neural Network)。

人工神经网络的研究可以追溯到 1800 年 Frued 的精神分析学时期,那时他已经做了一些初步工作。1913 年出现了人工神经系统的第一个实践,即由 Russell 描述的水力装置。1943 年美国心理学家 Warren S McCulloch 与数学家 Walter H Pitts 合作,用逻辑的数学工具研究客观事件在形式神经网络中的描述,从此开创了对神经网络的理论研究。他们在分析、总结神经元基本特性的基础上,首先提出神经元的数学模型,简称 MP 模型。从脑科学研究来看,MP 模型不愧为第一个用数理语言描述脑的信息处理过程的模型。后来 MP 模型经过数学家的精心整理和抽象,最终发展成一种有限自动机理论,再一次展现了 MP 模型的价值。此模型沿用至今,直接影响着这一领域的研究进展。1949 年心理学家 D. O. Hebb 提出关于神经网络学习机理的“突触修正假设”,即突触联系效率可变的假设,现在多数学习机仍遵循这一学习规则。1957 年, Frank Rosenblatt 首次提出并设计制作了著名的感知机(Perceptron),第一次从理论研究转入过程实现阶段,掀起了研究人工神经网络的高潮。今天,随着科学技术的迅猛发展,神经网络正以极大的魅力吸引着世界上众多专家、学者为之奋斗,在世界范围内再次掀起了神经网络的研究热潮。难怪有关国际权威人士评论指出,目前对神经网络的研究,其重要意义不亚于第二次世界大战时对原子弹的研究。

人工神经网络特有的非线性适应性信息处理能力,克服了传统人工智能方法对于直觉,如模式、语音识别、非结构化信息处理方面的缺陷,使之在神经专家系统、模式识别、智能控制、组合优化、预测等领域得到成功应用。人工神经网络与其他传统方法相结合,将推动人工智能和信息处理技术的不断发展。近年来,人工神经网络正向着模拟人类认知的道路上更加深入地发展,且与模糊系统、遗传算法、进化机制等结合,形成计算智能,成为人工智能的一个重要方向,将在实际应用中得到发展。

使用神经网络的主要优点是能够自适应样本数据,当数据中有噪声、形变和非线性时,它也能够正常地工作,很容易继承现有的领域知识,使用灵活,能够处理来自多个资源和决策系统的数据;提供简单工具进行自动特征选取,产生有用的数据表示,可作为专家系统的前端(预处理器)。此外,神经网络还能提供十分快的优化过程(尤其以硬件直接实现网络时),而且可以加速联机应用程序的运行速度。当然,过分夸大神经网络的应用能力也是不恰当的,毕竟它不是无所不能的。这就需要在实际工作中具体分析问题,合理选择。

基于神经网络的控制称为神经网络控制(NNC),简称神经控制(Neuro Control, NC)。这一新词是在国际自控联杂志《自动化》(Automatica)1994 年 No. 11 中首次使用的,最早源于

1992 年 H. Tolle 和 E. Ersu 的专著《Neuro Control》。基于神经网络智能模拟用于控制,是实现智能控制的一种重要形式,近年来获得了迅速发展。本章介绍神经网络的基本概念、基本结构,神经控制系统的组成及其实现神经控制的基本方法。

1.1 神经网络的基本概念

1.1.1 生物神经元的结构与功能特点

神经生理学和神经解剖学证明了人的思维是由人脑完成的。神经元是人脑组成的最基本单元,它能够接受并处理信息,人脑大约由 $10^{11} \sim 10^{12}$ 个神经元组成,其中每个神经元约与 $10^4 \sim 10^5$ 个神经元通过突触连接,因此,人脑是一个复杂的信息并行加工处理巨系统。探索脑组织的结构、工作原理及信息处理的机制,是整个人类面临的一项挑战,也是整个自然科学的前沿领域。

1. 生物神经元的结构

生物神经元,也称神经细胞,是构成神经系统的基本单元。生物神经元主要由细胞体、树突和轴突构成,其基本结构如图 1-1 所示。

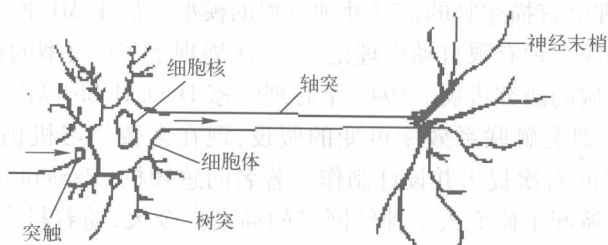


图 1-1 生物神经元的基本结构

(1) 细胞体

细胞体由细胞核、细胞质与细胞膜等组成。一般其直径为 $5 \sim 100 \mu\text{m}$, 大小不等。细胞体是生物神经元的主体,它是生物神经元的新陈代谢中心,同时还负责接收并处理从其他神经元传递过来的信息。细胞体的内部是细胞核,外部是细胞膜,细胞膜外是许多外延的纤维。细胞膜内外有电位差,称为膜电位,膜外为正,膜内为负。

(2) 轴突

轴突是由细胞体向外伸出的所有纤维中最长的一条分枝。每个生物神经元只有一个轴突,其长度最大可达 1m 以上。轴突相当于生物神经元的输出电缆,它通过尾部分的许多神经末梢及梢端的突触向其他生物神经元输出神经冲动。

(3) 树突

树突是由细胞体向外伸出的除轴突外的其他纤维分枝,长度一般均较短,但分枝很多。它相当于神经元的输入端,用于接收从四面八方传来的神经冲动。

突触是轴突的终端,是生物神经元之间的连接接口,每一个生物神经元约有 $10^4 \sim 10^5$ 个突触。一个生物神经元通过其轴突的神经末梢,经突触与另一个生物神经元的树突连接,以实现信息的传递。

2. 生物神经元的功能特点

从生物控制论的观点来看,作为控制和信息处理基本单元的生物神经元,具有以下功能特点。

(1) 时空整合功能

生物神经元对于不同时间通过同一突触传入的信息,具有时间整合功能;对于同一时间通过不同突触传入的信息,具有空间整合功能。两种功能相互结合,是使生物神经元具有时空整合的输入信息处理功能。

(2) 动态极化性

在每一种生物神经元中,信息都是以预知的确定方向流动的,即从生物神经元的接收信息部分(细胞体、树突)传到轴突的起始部分,再传到轴突终端的突触,最后再传给另一个生物神经元。尽管不同的生物神经元在形状及功能上都有明显的不同,但大多数生物神经元都是按这一方向进行信息流动的。

(3) 兴奋状态与抑制状态

生物神经元具有两种常规工作状态,即兴奋状态与抑制状态。所谓兴奋状态,是指生物神经元对输入信息经整合后使细胞膜电位值升高,且超过了动作电位的阈值,此时产生神经冲动并由轴突输出。抑制状态是指对输入信息整合后,细胞膜电位值下降到低于动作电位的阈值,从而导致无神经冲动输出。

(4) 结构的可塑性

突触传递信息的特性是可变的,也就是说它随着神经冲动传递方式的变化,传递作用强弱不同,从而形成了生物神经元之间连接的柔性。这种特性又称为生物神经元结构的可塑性。

(5) 脉冲与电位信号的转换

突触界面具有脉冲与电位信号的转换功能。沿轴突传递的电脉冲是等幅的、离散的脉冲信号,而细胞膜的电位变化为连续的电位信号,这两种信号是在突触接口进行变换的。

(6) 突触延期和不应期

突触对信息的传递具有时延和不应期的特点,即在相邻的两次输入之间需要一定的时间间隔,在此期间,无激励,不传递信息,称为不应期。

(7) 学习、遗忘和疲劳

由于生物神经元结构的可塑性,突触的传递作用有增强、减弱和饱和的情况。所以,神经细胞也具有相应的学习、遗忘和疲劳效应(饱和效应)。

1.1.2 人工神经元模型

生物神经元经抽象化后,可得到如图1-2所示的一种人工神经元模型,它有三个基本要素。

1. 连接权

连接权对应于生物神经元的突触,各个神经元之间的连接强度由连接权的权值表示,权值为正表示激活,为负表示抑制。

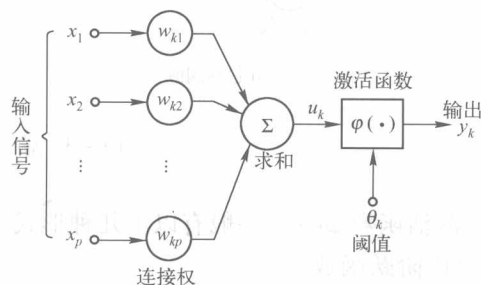


图1-2 人工神经元模型

2. 求和单元

用于求取各输入信号的加权和(线性组合)。

3. 激活函数

激活函数起非线性映射的作用,并将神经元输出幅度限制在一定范围内,一般限制在(0,1)或(-1,1)之间。激活函数也称传输函数。

此外还有一个阈值 θ_k (或偏值 $b_k = -\theta_k$)。

以上作用可分别以数学式表达出来:

$$u_k = \sum_{j=1}^p w_{kj}x_j, v_k = net_k = u_k - \theta_k, y_k = \varphi(v_k)$$

式中, $x_1, x_2, \dots, x_p$ 为输入信号,它相当于生物神经元的树突,为人工神经元的输入信息; $w_{k1}, w_{k2}, \dots, w_{kp}$ 为神经元 k 的权值; u_k 为线性组合结果; θ_k 为阈值; $\varphi(\cdot)$ 为激活函数; y_k 为神经元 k 的输出,它相当于生物神经元的轴突,为人工神经元的输出信息。

若把输入的维数增加一维,则可把阈值 θ_k 包括进去,即

$$v_k = \sum_{j=0}^p w_{kj}x_j, y_k = \varphi(v_k)$$

此处增加了新的连接,其输入 $x_0 = \mp 1$, 权值 $w_{k0} = \theta_k$ (或 b_k), 如图 1-3 所示。

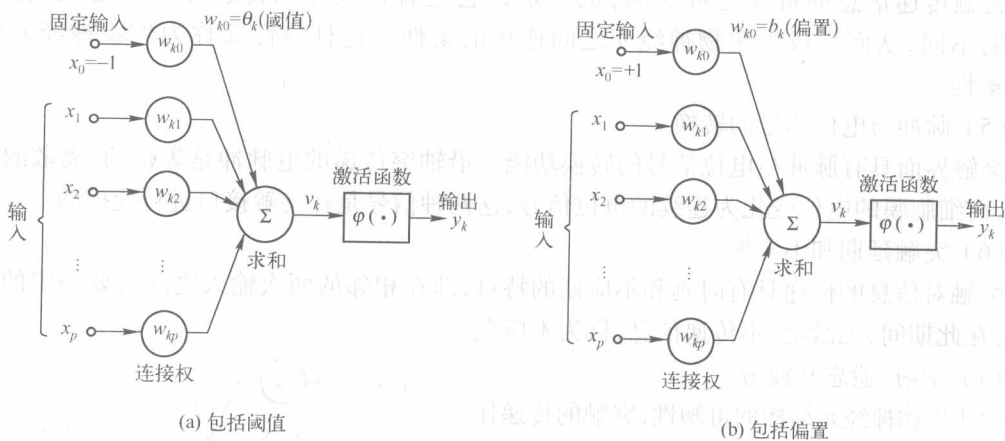


图 1-3 输入扩维后的神经元模型

激活函数 $\varphi(\cdot)$, 一般有以下几种形式。

① 阶跃函数

函数表达式:

$$y = \varphi(x) = \begin{cases} 1, & x \geq 0, \\ -1, & x < 0. \end{cases}$$

② 分段线性函数

函数表达式:

$$y = \varphi(x) = \begin{cases} 1, & (x \geq 1), \\ \frac{1}{2}(1+x), & (-1 < x < 1), \\ -1, & (x \leq -1). \end{cases}$$

③ Sigmoid 型函数

最常用的 Sigmoid 型函数为

$$\varphi(x) = \frac{1}{1 + \exp(-ax)}$$

式中,参数 a 可控制其斜率。

Sigmoid 型函数也简称为 S 型函数,上式表示的是一种非对称型 S 函数。

另一种常用的 Sigmoid 型函数为双曲正切对称型 S 函数,即

$$\varphi(x) = \tanh\left(\frac{1}{2}x\right) = \frac{1 - \exp(-x)}{1 + \exp(-x)}$$

这类函数具有平滑和渐近线,并保持单调性。

1.1.3 神经网络的结构

人工神经网络(Artificial Neural Networks, ANN)是由大量人工神经元广泛互连而组成的,它可用来模拟脑神经系统的结构和功能。人工神经网络可以看成是以人工神经元为节点,用有向加权弧连接起来的有向图。在此有向图中,人工神经元(以下在不易引起混淆的情况下,人工神经元简称神经元)就是对生物神经元的模拟,而有向加权弧则是轴突—突触—树突对的模拟。有向加权弧的权值表示相互连接的两个神经元间相互作用的强弱。

人工神经网络是生物神经网络的一种模拟和近似。它主要从两个方面进行模拟。一个是从生理结构和实现机理方面进行模拟,它涉及生物学、生理学、心理学、物理及化学等诸多基础学科。由于生物神经网络的结构和机理相当复杂,所以现在距离完全认识它们还相差甚远。另外一个是从功能上加以模拟,即尽量使得人工神经网络具有生物神经网络的某些功能特性,如学习、识别、控制等功能。本书仅讨论后者。从功能上来看,人工神经网络(以下简称神经网络, NN)根据连接方式主要分为前馈型网络和反馈型网络两类。

1. 前馈型网络

前馈型神经网络是整个神经网络体系中最常见的一种网络,其网络中各个神经元接受前一级的输入,并输出到下一级,网络中没有反馈,如图 1-4 所示。其节点分为两类,即输入单元和计算单元,每一计算单元可有任意个输入,但只有一个输出(它可耦合到任意多个其他节点作为输入)。通常,前馈型网络可分为不同的层,第 i 层的输入只与第 $i-1$ 层的输出相连,输入和输出节点与外界相连,而其他中间层称为隐含层,它们是一种强有力的学习系统,其结构简单而易于编程。从系统的观点看,前馈型神经网络是一静态非线性映射,通过简单非线性

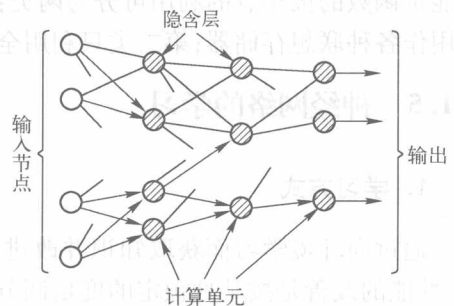


图 1-4 前馈型网络

处理的复合映射可获得复杂的非线性处理能力。但从计算的观点看,前馈型神经网络并非是一种强有力的计算系统,不具备丰富的动力学行为。大部分前馈型神经网络是学习网络,并不注意系统的动力学行为,但它们的分类能力和模式识别能力一般强于其他类型的神经网络。

2. 反馈型网络

反馈型神经网络又称递归网络(或回归网络)。在反馈型网络中(Feedback NNs),输入信号决定反馈系统的初始状态,然后系统经过一系列状态转移后,逐渐收敛于平衡状态,这样的平衡状态就是反馈型网络经计算后输出的结果。由此可见,稳定性是反馈网络中最重要的问题之一。如果能找到网络的 Lyapunov 函数,则能保证网络从任意的初始状态都能收敛到局部最小点。反馈型神经网络中的所有节点都是计算单元,同时也可接受输入,并向外界输出,可画成一个无向图,如图 1-5(a)所示,其中每个连接弧都是双向的,也可画成图 1-5(b)的形式。若总单元数为 n ,则每一个节点有 $n-1$ 个输入和一个输出。

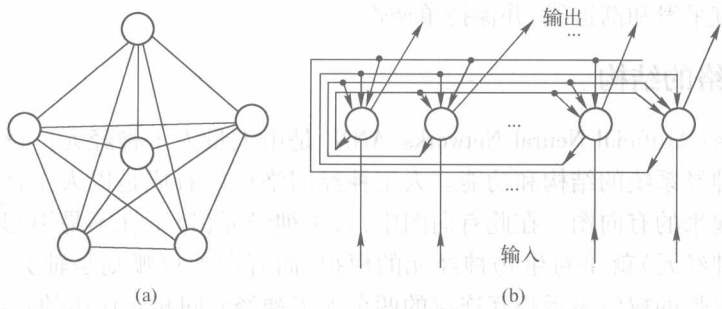


图 1-5 单层全连接反馈型网络

1.1.4 神经网络的工作过程

神经网络的工作过程主要分为两个阶段:第一个阶段是学习期,此时各计算单元状态不变,各连接权上的权值可通过学习来修改;第二个阶段是工作期,此时各连接权固定,计算单元变化,以达到某种稳定状态。

从作用效果看,前馈型网络主要是函数映射,可用于模式识别和函数逼近。反馈型网络按对能量函数的极小点的利用可分为两类:第一类是能量函数的所有极小点都起作用,这一类主要用作各种联想存储器;第二类只利用全局极小点,主要用于求解最优化问题。

1.1.5 神经网络的学习

1. 学习方式

通过向环境学习而获取知识并改进自身性能是神经网络的一个重要特点。在一般情况下,性能的改善是按某种预定的度量调节自身参数(如权值)而随时间逐步达到的神经网络的学习方式(按环境所供信息的多少分)有以下三种。

(1) 有监督学习(有教师学习)

这种学习方式需要外界存在一个“教师”,他可对一组给定输入提供应有的输出结果(正确

答案), 这组已知的输入/输出数据称为训练样本集。学习系统可根据已知输出与实际输出之间的差值(误差信号)来调节系统参数, 如图 1-6 所示。在有监督学习中, 学习规则由一组描述网络行为的训练集给出:

$$\{\mathbf{x}^{(1)}, \mathbf{t}^{(1)}\}, \{\mathbf{x}^{(2)}, \mathbf{t}^{(2)}\}, \dots, \{\mathbf{x}^{(N)}, \mathbf{t}^{(N)}\}$$

其中, $\mathbf{x}^p$ 为网络的第 p 个输入数据向量; $\mathbf{t}^p$ 为对应 $\mathbf{x}^p$ 的目标输出向量; N 为训练集中的样本数。

当输入作用到网络时, 网络的实际输出与目标输出相比较, 然后学习规则调整网络的权值和阈值, 从而使网络的实际输出越来越接近于目标输出。

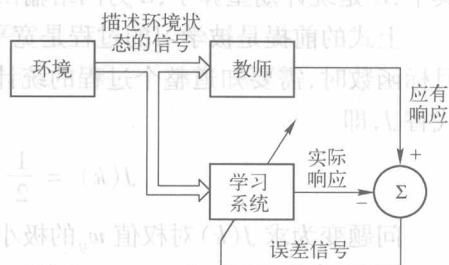


图 1-6 有监督学习框图

(2) 无监督学习(无教师学习)

无监督学习时不存在外部教师, 学习系统完全按照环境所提供数据的某些统计规律来调节自身参数或结构(这是一种自组织过程), 以表示外部输入的某种固有特性(如聚类, 或某种统计上的分布特征), 如图 1-7 所示。在无监督学习中, 仅仅根据网络的输入调整网络的权值和阈值, 它没有目标输出。乍一看这种学习似乎并不可行, 不知道网络的目的是什么, 还能够训练网络吗? 实际上, 大多数这种类型的算法都是要完成某种聚类操作, 学会将输入模式分为有限的几种类型。这种功能特别适合于诸如向量量化等应用问题。

(3) 强化学习(再励学习)

强化学习介于上述两种情况之间, 外部环境对系统的输出结果给出评价(奖或罚)而不是给出正确答案, 学习系统通过强化那些受奖励的动作来改善自身性能, 如图 1-8 所示。强化学习与有监督学习类似, 只是它不像有监督学习一样为每一个输入提供相应的目标输出, 而仅仅是给出一个级别。这个级别(或评分)是对网络在某些输入序列上的性能测度。当前, 强化学习要比有监督学习少见, 它最适合控制系统应用领域。



图 1-7 无监督学习框图

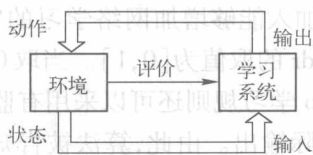


图 1-8 强化学习框图

2. 学习算法

(1) δ 学习规则(误差纠正学习规则)

若 $y_i(k)$ 为输入 $\mathbf{x}(k)$ 时神经元 i 在 k 时刻的实际输出, $t_i(k)$ 表示相应的期望输出, 则误差信号可写为

$$e_i(k) = t_i(k) - y_i(k)$$

误差纠正学习的最终目的是使某一基于 $e_i(k)$ 的目标函数达最小, 以使网络中每一输出单元的实际输出在某种统计意义上最逼近于期望输出。一旦选定了目标函数形式, 误差纠正学习就成为一个典型的最优化问题。最常用的目标函数是均方误差判据, 其定义为

$$J = E \left\{ \frac{1}{2} \sum_{i=1}^L (t_i - y_i)^2 \right\}$$