

神经网络稳定性理论

钟守铭 刘碧森 王晓梅 范小明 编著

神经网络稳定性理论

钟守铭 刘碧森 王晓梅 范小明 编著

科学出版社

北 京

内 容 简 介

本书介绍了神经网络、微分方程稳定性、泛函分析的基本理论和概念、Hopfield 型神经网络的稳定性理论、细胞神经网络的稳定性理论、二阶神经网络的稳定性理论、随机神经网络的稳定性理论以及神经网络的应用. 本书在选材时注重新颖性,反映了近年来神经网络稳定性理论的最新研究成果,写作时体现了通俗性与简洁性,论述深入浅出.

本书可作为应用数学专业的高年级大学生或研究生的教科书或参考书,也适合从事神经网络理论研究的科研人员阅读.

图书在版编目(CIP)数据

神经网络稳定性理论/钟守铭,刘碧森,王晓梅,范小明编著. —北京:科学出版社,2008

ISBN 978-7-03-021126-2

I. 神… II. 钟… III. 人工神经网络—运动稳定性理论—高等学校—教材 IV. TP183 0175.13

中国版本图书馆 CIP 数据核字(2008)第 071768 号

责任编辑:吕 虹 赵彦超/责任校对:赵桂芬

责任印制:赵德静/封面设计:陈 敬

科 学 出 版 社 出 版

北京东黄城根北街 16 号

邮政编码:100717

<http://www.sciencep.com>

新 蕾 印 刷 厂 印 刷

科学出版社发行 各地新华书店经销

*

2008 年 6 月第 一 版 开本:B5(720×1000)

2008 年 6 月第一次印刷 印张:18 3/4

印数:1—3 000 字数:354 000

定价:56.00 元

(如有印装质量问题,我社负责调换<路通>)

前 言

人乃万物之灵长,起关键作用的物理实体就是人脑.人脑是造物主在这个宇宙中最完美的创造,它具有最精巧的结构、最奥妙的运行、最完善的功能、最快的判断决策速度和大量的记忆实体.长期以来,科学家们一直想揭开人脑工作机制的奥秘,以便能制造出接近人脑功能的机器来解决问题.到目前为止,人们虽然对脑神经结构、机制和功能等方面已做了大量研究工作,取得了许多有建设性的成果,但远远还没有认识清楚.这是一项艰巨的任务,是需要各个领域的科学家长期共同努力方能实现的目标.神经网络模型,更精确地说,人工神经网络模型只是生物神经系统高度简化后的一种近似;它是用大量的简单神经元广泛连成的一种计算结构,属于非线性动力学系统;它在不同程度、不同层次上模拟人脑神经系统的结构及其信息处理、存储和检索等功能.神经网络是一门新兴交叉学科,自1982年美国加州理工学院物理学家Hopfield提出Hopfield型网络模型和1988年美国电子学家Chua和Yang提出细胞神经网络(cellular neural network, CNN)以来,神经网络研究取得了前所未有的良好局面,进入了快速发展时期,并成为人工智能中最为活跃的分支之一,在很多领域得到了广泛应用.例如,模式识别与图像处理、时间序列分析、信号处理和控制等.

稳定性概念的历史可以追溯到17世纪,Torricelli原理中用到过这个概念.但是稳定性的严格的数学定义和一般理论迟迟没有出现,直到1892年俄国数学家Lyapunov在他的博士学位论文《运动稳定性的一般问题》中首次给出了运动稳定性精确的数学定义和一般方法,从而奠定了稳定性理论的基础.

本书旨在发展新的学科方向,寻求新的研究领域.因为神经网络模型是一个非线性动力学系统,所以本书力图用稳定性的理论和方法来系统地研究神经网络模型.在本书撰写之前,国内还没有一本较为系统地专门研究神经网络稳定性理论的研究生教材.本书试图将国际上近二十年有关神经网络稳定性理论的主要研究成果系统地呈现给读者.

本书由钟守铭教授主编,各章节执笔者分别为:刘碧森(第1章),王晓梅(第2,3章),范小明(第4~6章).

本书的出版获得电子科技大学学科建设经费的资助,并得到了电子科技大学研究生院、应用数学学院和科学出版社的大力支持,对此表示衷心的感谢.在本书的编写过程中,博士生杨军阅读了部分手稿,提出了不少宝贵意见,在此表示诚挚的感谢,同时也对关心帮助本书出版的老师们表示感谢.

由于作者水平有限,加之时间仓促,疏漏之处在所难免,恳请同行专家及读者批评指正!

作 者

2007年9月

目 录

第 1 章 绪 论	1
1.1 神经网络简介	1
1.1.1 神经网络的概念	2
1.1.2 神经网络的优点	3
1.1.3 人脑简介	5
1.1.4 神经元模型	7
1.1.5 神经网络的网络结构	17
1.1.6 知识表示	20
1.1.7 人工智能和神经网络	28
1.1.8 两个重要的神经网络模型	32
1.2 微分方程稳定性理论基础	39
1.2.1 微分方程稳定性的基本概念	39
1.2.2 稳定性的判据与 LaSalle 不变原理	46
1.3 泛函分析简介	52
1.3.1 度量空间的概念和度量空间中的点集	52
1.3.2 线性空间上的范数、 L^p 空间及其几个重要不等式	54
1.3.3 连续映射、稠密性、完备性	56
1.3.4 不动点定理	58
1.3.5 内积空间、Hilbert 空间上的有界线性算子	60
第 2 章 Hopfield 型神经网络的稳定性	63
2.1 无时滞的 Hopfield 型神经网络的稳定性分析	63
2.1.1 Hopfield 型神经网络平衡点的存在性和唯一性	64
2.1.2 Hopfield 型神经网络平衡点的渐近稳定性	66
2.2 有时滞的 Hopfield 型神经网络的稳定性分析	67
2.3 Hopfield 型神经网络的 k 稳定性分析	81
第 3 章 细胞神经网络的稳定性	87
3.1 无时滞的细胞神经网络的稳定性分析	87
3.2 有时滞的细胞神经网络的稳定性分析	98
3.3 细胞神经网络的周期解	118
3.4 广义细胞神经网络	136

第 4 章 二阶神经网络的稳定性	140
4.1 二阶神经网络简介	140
4.2 二阶 Hopfield 型神经网络的局部稳定性分析	143
4.3 二阶 Hopfield 型神经网络的全局稳定性分析	150
4.3.1 全局渐近稳定性分析	150
4.3.2 全局指数稳定性分析及收敛速度的估计	152
4.3.3 例子	156
4.4 具有反应扩散的二阶 Hopfield 型神经网络的稳定性分析	157
4.4.1 全局渐近稳定性分析	158
4.4.2 全局指数稳定性分析及收敛速度的估计	161
4.5 具有时滞的二阶 Hopfield 型神经网络的稳定性	165
4.5.1 平衡点的存在性	166
4.5.2 全局渐近稳定的若干充分条件	168
4.5.3 全局指数稳定性	174
4.5.4 推广到变时滞情形	175
4.6 具有时滞的二阶 Hopfield 型神经网络的正不变集与吸引子	177
4.7 带反应扩散项和时滞的二阶 Hopfield 型神经网络的稳定性	180
4.7.1 全局渐近稳定性	182
4.7.2 全局指数稳定性	187
4.8 具有时滞的二阶 Hopfield 型神经网络的周期解及其稳定性	189
4.8.1 常时滞情形	189
4.8.2 变时滞情形	193
4.9 具有时滞的脉冲二阶 Hopfield 型神经网络	195
第 5 章 随机神经网络的稳定性	202
5.1 随机神经网络简介	202
5.1.1 随机神经网络的发展	202
5.1.2 GRNN 模型描述	203
5.1.3 RNN 的应用	204
5.1.4 其他随机网络	205
5.1.5 研究前景	207
5.2 随机 Hopfield 型神经网络的稳定性	208
5.3 随机二阶神经网络的稳定性	213
5.4 随机时滞神经网络的稳定性	216
5.4.1 几乎必然指数稳定性	217
5.4.2 均方指数稳定性	219

5.5 具有可变时滞的随机 Hopfield 型神经网络	223
5.6 随机细胞神经网络	227
5.6.1 与时滞无关的稳定性判据	227
5.6.2 与时滞有关的稳定性判据	231
第 6 章 神经网络的应用	233
6.1 神经网络在工业控制领域中的应用	233
6.1.1 人工神经网络简介	233
6.1.2 人工神经网络技术与焊接过程的结合及其相关范例	233
6.1.3 人工神经网络的巨大优势及在焊接过程中的初步应用	237
6.2 神经网络在手写体字符识别中的应用	238
6.2.1 手写体字符识别分类中的神经网络模型	239
6.2.2 神经网络分类器集成方法	240
6.2.3 神经网络与其他优化算法的结合	242
6.2.4 未来的研究方向	243
6.3 基于神经网络的全自动模式识别跟踪系统	243
6.3.1 神经网络模式识别的原理	244
6.3.2 神经网络模式识别的实现	244
6.3.3 识别跟踪系统实验测试与数据分析	246
6.4 基于神经网络不变性模式识别的人脸识别系统	249
6.4.1 最优化神经网络	249
6.4.2 人脸识别系统的基本结构	250
6.4.3 仿真实验	252
6.5 人工神经网络在临床医学中的应用	253
6.6 神经网络在城市综合业务故障诊断中的应用	256
6.6.1 支撑环境	256
6.6.2 软件设计实例	257
6.6.3 实验结果实例分析	258
6.7 基于人工神经网络的商业银行贷款风险预警研究	260
6.7.1 前向三层 BP 网络商业银行信贷风险预警系统的构造	260
6.7.2 贷款风险预警信号知识处理	262
6.7.3 实验设计	263
6.7.4 BP 网络模型的计算机实现	264
6.8 Hopfield 型神经网络应用实例分析	264
6.8.1 Hopfield 型神经网络简介	264
6.8.2 离散 Hopfield 型神经网络在模式识别中的应用	265

6.8.3 连续 Hopfield 型神经网络在优化计算中的应用	267
6.9 基于 Hopfield 型神经网络的遥感图像超分辨率识别算法	269
6.9.1 Hopfield 型神经网络实现的图像超分辨率目标识别处理算法分析 ...	270
6.9.2 Hopfield 型神经网络目标识别处理的主要实验过程分析	273
6.9.3 Hopfield 型神经网络的图像超分辨率处理的实验结果分析	274
6.10 应用 Hopfield 型神经网络抑制高频地波雷达中短波通信干扰	275
6.10.1 Hopfield 型神经网络	275
6.10.2 基于 LMS 的自适应抑制干扰法	275
6.10.3 抑制干扰 Hopfield 型神经网络	277
6.10.4 仿真结果的比较	279
参考文献	281

第 1 章 绪 论

神经网络的稳定性理论在过去二十多年中的发展是十分活跃的. 神经网络理论建立以后, 由于要求算法的收敛性, 从而开始了对神经网络模型的稳定性的研究, 神经动力学就是以研究神经网络的数学模型为主要内容的. 本书主要介绍神经动力学的稳定性理论. 为了研究神经网络稳定性, 需要一些神经网络理论和数学知识, 必须先介绍神经网络的基本理论、微分方程稳定性理论和泛函分析的一些知识.

1.1 神经网络简介

神经网络技术是 20 世纪末迅速发展起来的一门高新技术. 由于神经网络具有良好的非线性映射能力、自学习适应能力和并行信息处理能力, 为解决不确定非线性系统的建模和控制问题提供了一条新的思路, 因而吸引了国内外的众多学者和工程技术人员从事神经网络控制的研究, 并取得了丰硕成果, 提出了许多成功的理论和方法, 使神经网络控制逐步发展成为智能控制的一个重要分支.

神经网络控制的基本思想就是从仿生学角度来模拟人的神经系统的运作方式, 从而使得机器具有人脑那样的感知、学习和推理能力. 它将控制系统看成是由输入到输出的一个映射, 利用神经网络的学习能力和适应能力实现系统的映射特性, 从而完成对系统的建模和控制. 它使模型和控制的概念更加一般化. 理论上讲, 基于神经网络的控制系统具有一定的学习能力, 能够更好地适应环境和系统特性的变化, 非常适合于复杂系统的建模和控制. 特别是当系统存在不确定性因素时, 更体现了神经网络方法的优越性.

神经网络在控制领域受到重视主要归功于它的非线性映射能力、自学习适应能力、联想记忆能力、并行信息处理方式及其良好的容错性能. 应用神经网络时, 人们总期望它有非常快的全局收敛特性、大范围的映射泛化能力和较少的实现代价.

非线性控制系统早期的研究是针对一些特殊的、基本的系统而言的, 其代表性的理论有: 相平面法、描述函数法、绝对稳定性理论、Lyapunov 稳定性理论、输入输出稳定性理论等. 自 20 世纪 80 年代以来, 非线性科学越来越受到人们重视, 数学中的非线性分析、非线性泛函、物理学中的非线性动力学, 发展都很迅速. 与此同时, 非线性系统理论也得到了蓬勃发展, 有更多的控制理论专家转入到非线

性系统的研究,更多的工程师力图用非线性系统理论构造控制器,并取得了一定成就.神经网络方法是主要方法中的一种.

1.1.1 神经网络的概念

当人们认识到人脑的工作方式与传统的数字计算机的工作方式完全不同的时候,关于神经网络的研究就开始了.人脑是一个高度复杂的、非线性的和并行的信息处理系统.人脑能够组织它的组成成分,即神经元,以比今天已有的最快的计算机还要快许多倍的速度进行特定的计算(如模拟识别、感知和运动神经控制).例如,考虑人类视觉,这是一个信息处理的任务(Marr,1982;Leving,1985;Churchland et al.,1992).视觉系统功能是我们提供了一个关于周围环境的表示,并且更重要的是提供给我们和环境交互所需的信息.具体地讲,完成一个感知识别任务(如识别一张被嵌入陌生场景的熟悉的脸)人脑需要100~200ms,而一台传统的计算机却要花费几天时间才能完成一个相对简单得多的任务.

再举一个例子,考虑一只蝙蝠的声呐.声呐就是一个活动回声定位系统.除了提供目标(如飞行的昆虫)有多远的信息外,蝙蝠的声呐可以搜集目标的相对速度、目标大小、目标不同特征的大小及它的方位角和仰角的信息(Suga,1990a,1990b).所有信息都是从目标的回声中提取,而所有需要的复杂神经计算只在李子般大小的脑中完成.事实上,一只回声定位的蝙蝠可以灵巧地以很高的成功率追捕和捕捉目标,这一点可以让雷达或声呐工程师们自叹不如.

那么,人脑或蝙蝠的脑是如何做到这一点的呢?脑具有精巧的构造和通过“经验”而建立自己规则的能力.确实,经验是经时间积累的,人脑在人出生后头两年内发生了最戏剧性的发展(即硬连接),但是发展将超越这个阶段并继续进行.

一个“发展中”的神经元是与可塑的人脑同义的.可塑性允许一个发展中的神经系统适应它的周边环境.可塑性似乎是人脑中作为信息处理单元的神经元功能的关键.同样,它在人工神经元组成的神经网络中亦是如此.形式最普通的神经网络就是对人脑完成特定任务或感兴趣的功能的方法进行建模的机器;网络一般用电子器件实现或者用软件在数字计算机上模拟.为了获得好的结果,神经网络使用一个很庞大的简单计算单元间的相互连接,这些简单计算单元称为“神经元”或“处理单元”.据此给出将神经网络看作一种自适应机器的定义.

一个神经网络是一个由简单处理元构成的规模宏大的并行分布式处理器.天然具有存储经验知识和使之可用的特性.神经网络在两个方面与人脑相似:

- (1) 神经网络获取的知识是从外界环境中学习得来的.
- (2) 互连神经元的连接强度,即突触权值,用于存储获取的知识.

用于完成学习过程的程序称为算法,其功能是以有序的方式来改变网络的

突触权值以获得想要的设计目标。

突触权值修改提供神经网络设计的传统方法。这种方法和线性自适应滤波器理论很近似。滤波器理论已经很好地建立起来并成功应用在很多领域。但是,神经网络修改它自身的拓扑结构亦是可能的,这也和人脑的神经元会死亡、新的突触连接会生长的情况相适应。

1.1.2 神经网络的优点

神经网络的计算能力明显有以下两点:① 大规模并行分布结构;② 神经网络学习能力以及由此而来的泛化能力。泛化是指神经网络对不在训练(学习)集中的数据可以产生合理的输出。这两种信息处理能力让神经网络可以解决一些当前还不能处理的复杂(大型)问题。但是在实践中,神经网络不能单独作出解答,它们需要被整合在一个协调一致的系统工程方法中。具体地讲,一个复杂的问题往往被分解成若干个相对简单的任务,而神经网络处理与其能力相符的子任务。但是,在建立一个可以模拟人脑的计算机结构(如果可能)之前还有很长的路要走,认识这一点很重要。

神经网络具有下列性质和能力:

(1) 非线性。一个人工神经元可以是线性或者非线性的。一个由非线性神经元互联而成的神经网络自身是非线性的,并且非线性是一种分布于整个网络中的特殊性质。非线性是一个很重要的性质,特别是如果产生输入信号(如语音信号)内部的物理机制是天生非线性的。

(2) 输入输出映射。有教师的学习,也称为有监督学习,这是一种流行的学习范例。教师和神经网络同时要对从周围环境中抽取出来的训练样本作出判断,教师可以根据自身掌握的一些知识为神经网络提供对训练样本的期望响应,神经网络的参数可以在训练向量和误差信号的综合影响下进行调整。误差信号可以定义神经网络实际响应和期望响应之差。这种调整可以反复进行,其目的就是用神经网络模拟教师。在某种统计意义下,可以认为这种模拟是最优的。使用训练集中的很多例子重复神经网络的训练,直到网络到达没有显著的突触权值修正的稳定状态为止。先前用过的例子可能还要在训练期间以不同的顺序重复使用。因此对当前问题,网络通过建立输入输出映射从例子中进行学习。这样一个方法使人想起了无参数统计推断的研究,它是非模型估计的统计处理的分支,或者从生物学角度看,称为 *tabula rasa* 学习(Geman et al., 1992)。这里使用“非参数”表示的一个事实是,没有对输入数据的统计模型做任何先验假设。例如,考虑一个模式分类任务,这里要求把代表具体物体或事件的输入信号分到几个预先分好的类中去。在这个问题的非参数方法中,要求利用例子集“估计”输入信号空间中模式分类任务的任意判决边界,并且不使用概率分布模式。有监督学习

范例隐含了一个类似的观点,这提示神经网络的输入输出映射和非参数统计推断之间的一个相近的类比。

(3) 适应性. 神经网络嵌入了一个调整自身突触权值以适应外界变化的能力. 特别是,一个在特定运行环境下接受训练的神经网络,对环境条件不大的变化可以容易进行重新训练. 而且,当它在一个时变环境(即它的统计特性随时间变化)中运行时,网络突触权值就可以设计成随时间变化. 用于模式识别、信号处理和控制的神经网络与它的自适应能力耦合,就可以变成能力进行自适应模式识别、自适应信号处理和自适应控制的有效工具. 作为一个一般规则,在保证系统保持稳定时一个系统的自适应性越好,当要求在一个时变环境下运行时它的性能就越有鲁棒性. 但是,需要强调的是,自适应性不一定导致鲁棒性,实际可能相反. 例如,一个暂态自适应系统可能变化过快,以致对寄生干扰有反应,这将引起系统性能的急剧恶化. 为最大限度地实现自适应性,系统的主要时间常数应该长到可以忽略寄生干扰,短到可以反映环境的重要变化. 这是一个稳定的可塑性环境.

(4) 证据响应. 在模式识别问题中,神经网络可以设计成既能提供不限于选择哪一个特定模式的信息,也提供决策的置信度的信息. 后者可以用来拒判那些出现得过于模糊的模式.

(5) 背景信息. 神经网络的特定结构和激发状态代表知识. 网络中每一个神经元都潜在地受网络中其他所有神经元全局活动的影响. 因此,背景信息自然由一个神经网络处理.

(6) 容错性. 一个以硬件形式实现后的神经网络有天生容错的潜质,或者鲁棒计算的能力,意即它的性能在不利条件下逐渐下降. 例如,一个神经元或它的连接损坏了,存储模式的回忆在质量上被削弱. 但是,由于网络信息存储的分布特性,在网络的总体响应严重恶化之前这种损坏是分散的. 因此,原则上,一个神经网络的性能显示了一个缓慢恶化而不是灾难性的失败. 有一些关于鲁棒性计算的经验证据,但通常它是不可控的. 为了确保网络事实上的容错性,有必要在设计训练网络的算法时采用正确的度量.

(7) VLSI 实现. 神经网络的大规模并行性使它具有快速处理某些任务的潜在能力. 这一特性使得神经网络很适合用超大规模集成(very large scale integrated, VLSI)技术实现. VLSI 的一个特殊优点是,提供一个以高度分层的方式捕捉真实复杂性行为的方法.

(8) 分析和设计的一致性. 基本上,神经网络作为信息处理器具有通用性. 这样说是在这样的意义下,即涉及神经网络应用的所有领域都使用同样的记号. 这种特征以不同的方式表现出来.

(i) 神经元. 不管形式如何,在所有的神经网络中都代表一个相同部分;

(ii) 这种共性使得在不同应用中的神经网络共享相同的理论和学习算法成为可能;

(iii) 模块化网络可以用模块的无缝集成来实现.

有了神经生物学的启示,对人脑及其组织的结构层次作简要的考察看来是合适的.

1.1.3 人脑简介

人的神经系统可看作三个阶段系统,如图 1.1.1 所描绘的框图. 系统的中央是人脑,由神经网络表示,它连续地接收信息,感知它并作出适当的决定. 图中有两组箭头,从左到右的箭头表示携带信息的信号通过系统向前传输,从右到左的箭头表示系统中的反馈. 感受器把人体或外界环境的刺激转换成电冲击,对神经网络(大脑) 传送信息. 神经网络的效应器转换神经网络产生的电冲击为可识别的响应作为系统输出.

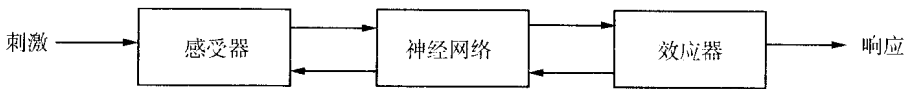


图 1.1.1 神经系统的框图表示

1911 年,Romony Cajal 首先引入把神经元作为人脑结构成分的思想. 由于他的这一开创性工作,人们理解人脑就简单多了. 通常,神经元比硅逻辑门要慢 5 ~ 6 个数量级;硅逻辑门中的事件发生在纳秒(10^{-9} s) 级,而在神经中的时间发生在毫秒(10^{-3} s) 级. 但是人脑由运行速度相对较慢的神经元构成,神经元(神经细胞) 数目确实惊人,而且它们之间具有大量互联. 估计人的皮质大约有 100 亿个神经元和大约 6 亿兆个突触或连接(Shepherd et al., 1990). 脑中的网络是高效结构. 特别地,脑的能量效率每秒每个操作大约为 10^{-16} J,而今天所用的最好的计算机的相应值是每秒每个操作大约 10^{-6} J(Faggin, 1991).

突触是调节神经元之间相互作用的基本结构和功能单元. 最普遍的一类突触是化学突触,它运行如下:前突触过程释放发送器物质,扩散到神经元之间的突触连接,然后作用于后突触过程. 这样突触就完成了突触前端的电信号和化学信号的转换,然后返回突触后端电信号(Shepherd et al., 1990). 用电学术语,这样的元素称为非互逆的两端口设备. 在传统的神经组织描述中,仅假设突触是一个简单的连续,能加载兴奋或抑制,但不同时作用在接受神经元.

前面曾提到过,可塑性允许发展神经系统以适应周边环境(Eggermont, 1990;Churchland et al., 1992). 在成年人的脑中,可塑性可以解释两个机能:创建神经元之间的连接和修改连接;轴突触有光滑的表面,较少的分支,比较长,而树突正相反(之所以这样称呼是因为它和树相似),它有不规则的表面和更多的

分支(Freeman, 1975). 脑中的不同部分有很多种形状和大小的神经元, 图 1. 1. 2 是一种锥形细胞, 它在脑皮层中是常见的. 和其他许多神经元一样, 它从树突刺激接收大部分信号; 可以从图 1. 1. 2 中看到树突片段细节. 锥形细胞可以有一万个或更多的突触与其他细胞连接, 它可以投射到数以千计的目标细胞.

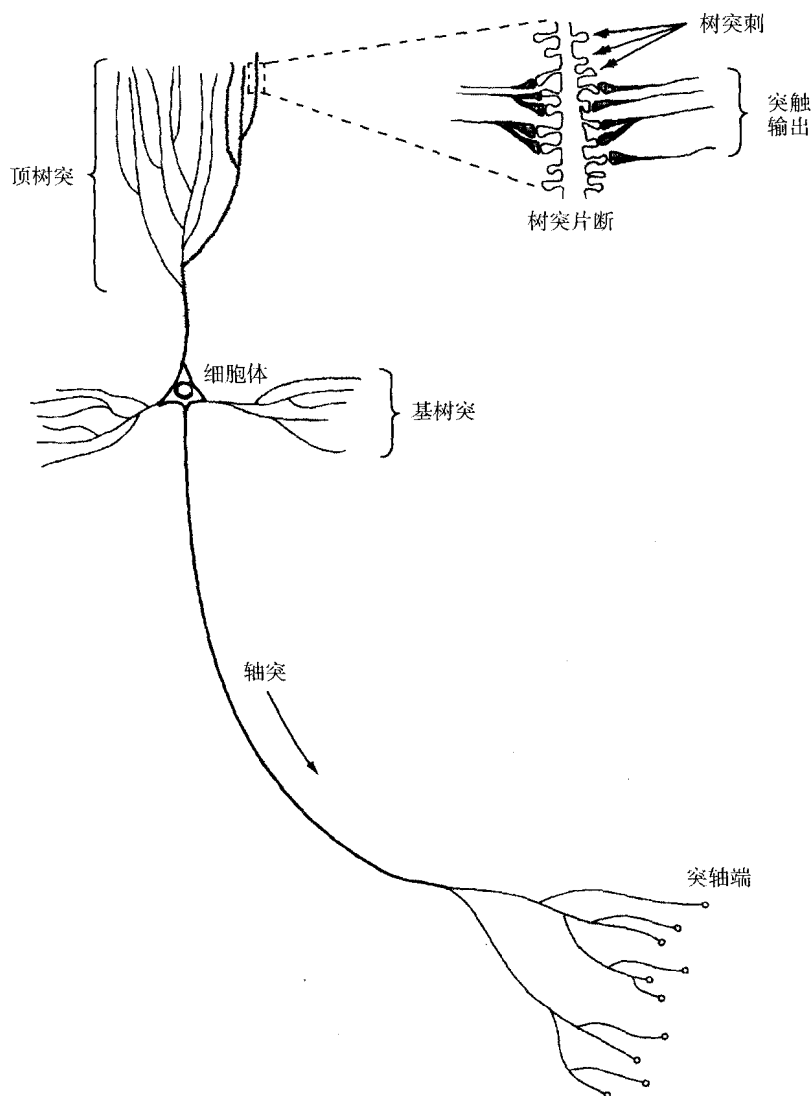


图 1. 1. 2 锥形细胞

大多数神经元把它们的输出转化成一系列简短的电压脉冲编码. 这些脉冲一般称为动作电位或冲击, 产生于神经元细胞体或其附近并以恒定的电压和振幅穿越个体神经元. 神经元很细, 有很高的电阻和非常大的电容, 二者分布于轴突中. 因此可以用 RC 传输线路来建模, 用“线路方程”这个术语来描述轴突中的

信号传播. 对传播机制的分析揭示电压在传输中随距离指数衰减, 在到达一端时会变得很小. 动作电位提供了克服这个问题的方法(Anderson, 1995).

在人脑中, 有大小解剖组织之分, 机能也有高下之别. 图 1.1.3 显示脑组织交织水平的层次结构, 这已经在广泛的关于脑局部区域的分析工作中显现出来了(Shepherd et al. , 1990; Churchland et al. , 1992). 突触表示最基本的层次, 其活动依赖于分子和离子. 其后的层次有神经微电路、树突树和最后的神经元. 神经微电路指突触集成, 组织成可以产生感兴趣的功能操作的连接模式. 它就像一个由晶体管集成的硅片, 最小的尺寸用微米(μm) 度量, 最快的操作速度用毫秒(ms) 度量, 神经微电路被组织成属于神经元个体的树突树的树突子单元. 整个神经元约为 $100\mu\text{m}$ 大小, 包含几个树突子单元. 局部电路(约 1mm 大小) 处在其次的复杂性水平, 由具有相似或不同性质的神经元组成, 这些神经元集成完成脑部解剖图组成, 牵涉脑中不同部分的多个区域.

局部解剖图由响应输入感知器信息所组成. 它们经常被组织成片束状, 如同在上丘中一样. 上丘中视觉、听觉和人体触觉区以层邻接的方式放置, 使得空间中相应点的刺激处于各层的下面或上面. 由 Brodal(1981) 作出的大脑皮质的细胞结构图清楚地表明, 不同的感觉信息(运动、触觉、视觉、听觉等) 被有序地映射到大脑皮层的相应位置. 在复杂性的最后一级, 局部解剖图和其他的区域间电路成为中央神经系统传递特定行为的媒介.

认识到 Brodmal 所描绘的结构分层组织是人脑的独有特征非常重要. 在数字计算机中找不到这种结构, 在人工神经网络中也无法近似地重构它们. 但是, 仍在向图 1.1.3 中描述的类似的分级计算的层状结构缓慢推进. 用以构造神经网络的人工神经元和人脑中的神经元相比确实比较初级. 目前能设计的网络和人脑中初级的局部电路和区域间电路相当. 但是, 真正令人满意的是过去 20 年间在许多前沿有了显著进步. 以神经生物类比为灵感的源泉, 加上具有的理论和技术工具的这些财富, 未来 10 年对神经网络的理解一定会更加深入.

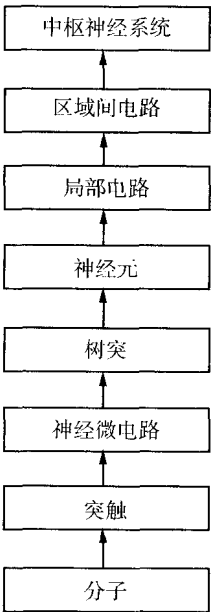


图 1.1.3 脑组织的
分层结构

1.1.4 神经元模型

神经元是神经网络操作的基本信息处理单位. 它是(人工)神经网络的设计基础. 在这里给出神经元模型的三种基本元素:

(1) 突触或连接链, 每一个都由其权值或者强度作为特征. 特别地, 在连到神经元 k 的突触 j 上的输入信号 x_j 被乘以 k 的突触权重 w_{kj} . 注意突触权值 w_{kj} 的下标的写法很重要. 第一个下标指查询神经元, 第二个下标指权值所在的突触的输入端. 和人脑中的突触不一样, 人工神经元的突触权值有一个范围, 可以取正值也可以取负值.

(2) 加法器, 用于求输入信号被神经元的相应突触加权的和. 这个操作构成一个线性组合器.

(3) 激活函数, 用来限制神经元输出振幅. 激活函数也称为压制函数, 由于它将输出信号压制(限制)到允许范围之内的一定值. 通常, 一个神经元输出的正常幅度范围可写成单位闭区间 $[0, 1]$ 或者另一种区间 $[-1, 1]$.

图 1.1.4 的神经元模型也包括一个外部偏置, 记为 b_k . 偏置的作用是根据其为正或为负, 相应地增加或降低激活函数的网络输入.

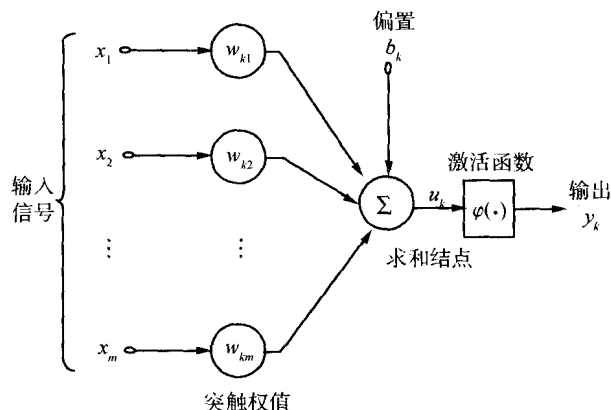


图 1.1.4 神经元的非线性模型

用数学术语, 可以用如下两个方程描述一个神经元 k :

$$u_k = \sum_{j=1}^m w_{kj} x_j, \quad (1.1.1)$$

$$y_k = \varphi(u_k + b_k), \quad (1.1.2)$$

其中, $x_1, \dots, x_m$ 是输入信号, $w_{k1}, w_{k2}, \dots, w_{km}$ 是神经元 k 的突触权值, u_k 是输入信号的线性组合器的输出, 偏置为 b_k , 激活函数 $\varphi(\cdot)$, y_k 是神经元输出信号. 偏置 b_k 的作用是对图 1.1.4 模型中的线性组合器的输出 u_k 作仿射变换, 如下所示:

$$v_k = u_k + b_k. \quad (1.1.3)$$

特别地, 根据偏置 b_k 取正或取负, 神经元 k 的诱导局部域或激活电位 v_k 和线性组合器输出 u_k 的关系如图 1.1.5 所示. 注意到由于这个仿射变换的作用, v_k 与 u_k 的图形不再经过原点.