

神经网络导论

王晓梅 编著



科学出版社

神经网络导论

王晓梅 编著

科学出版社

北京

内 容 简 介

本书共 5 章, 第 1 章主要介绍神经网络、微分系统稳定性理论和泛函分析的基本理论和概念; 第 2 章介绍神经网络的基本模型及算法; 第 3 章介绍后期比较热门的三种神经网络, 即 Hopfield 神经网络、细胞神经网络与双向联想(BAM)神经网络的模型及动力学问题; 第 4 章介绍复杂神经网络模型及动力学问题; 第 5 章介绍神经网络的应用. 因为神经网络模型是一个非线性动力学系统, 有些内容让读者难以读懂, 所以本书写作时力求内容简洁、通俗, 论述深入浅出, 系统地介绍了神经网络模型、算法及动力学问题.

本书既可作为数学专业或工科专业的本科生教材, 也可作为研究生教材或供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

神经网络导论/王晓梅编著. —北京: 科学出版社, 2016.12

ISBN 978-7-03-051155-3

I. ①神… II. ①王… III. ①人工神经网络—研究 IV. ①TP183

中国版本图书馆 CIP 数据核字(2016) 第 318203 号

责任编辑: 李静科 / 责任校对: 彭 涛
责任印制: 张 伟 / 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>

北京中石油彩色印刷有限责任公司印刷

科学出版社发行 各地新华书店经销

*

2017 年 1 月第 一 版 开本: 720 × 1000 1/16

2017 年 1 月第一次印刷 印张: 19 1/4

字数: 378 000

定价: 98.00 元

(如有印装质量问题, 我社负责调换)

前 言

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,因而吸引了国内外众多的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,模拟人神经系统的运作方式,使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。它使模型和控制的观念更加一般化。从理论上讲,基于神经网络的控制系统的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,人们总期望它具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、非线性泛函及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,并取得了一定的成就。神经网络方法是主要方法中的一种。

对读者来讲,本书作为人工神经网络的入门课程,重点介绍人工神经网络及其网络模型,使读者了解智能系统描述的基本模型,掌握人工神经网络的基本概念与各种基本网络模型的结构、特点、典型训练算法、运行方式,掌握软件实现方法;然后将学生引入人工神经网络应用的研究领域,通过实验进一步体会有关模型的用法和性能,获取一些初步的经验,同时使读者了解人工神经网络的有关研究思想,从中学习问题的求解方法。对高级研究者可以查阅适当的参考文献,将所学的知识与自己未来研究课题(包括研究生论文阶段的研究课题)结合起来,达到既丰富学习

内容,又具有一定的研究和应用的目的。

本书针对本科高年级学生或研究生用书的实际情况,精选了《神经网络稳定性理论》(钟守铭等编著)部分内容,并对内容进行了优化和增删,同时结合相关的参考文献和自身的研究领域,增添了基本网络模型的结构、特点、典型训练算法、运行方式及典型问题等,并加进了最新的一些研究成果。本书注重内容和体系的整体优化,浅显易懂,语言通俗简洁,重视培养学生应用神经网络知识解决实际问题的能力。

本书由电子科技大学王晓梅副教授主编和执笔。

本书的出版获得了电子科技大学学科建设和新编特色教材项目经费的资助,并得到了电子科技大学数学科学学院和科学出版社的大力支持。在本书的编写过程中,电子科技大学钟守铭教授对本书进行了评审,并提出了不少宝贵的意见,在此一并表示衷心的感谢。电子科技大学于雪梅硕士阅读了本书的部分手稿,对编辑格式和内容进行了仔细查阅,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。

由于编者水平有限,不足之处在所难免,恳请读者批评指正!

作 者

2016年6月

目 录

前言

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 神经网络的发展	4
1.1.3 神经网络的优点	5
1.1.4 人工神经元模型	7
1.1.5 神经网络的网络结构	14
1.2 微分方程稳定性理论基础	17
1.2.1 微分方程的基本知识	18
1.2.2 微分、积分不等式	19
1.2.3 Lyapunov 函数相关定义和定理	21
1.2.4 稳定性的基本定义和定理	27
1.2.5 Lyapunov 直接法的基本定理	39
1.2.6 构造 Lyapunov 函数的基本方法	48
第 2 章 神经网络基本模型	53
2.1 M-P 模型	53
2.1.1 MP 模型的概念	53
2.1.2 标准 M-P 模型	53
2.1.3 时延 M-P 模型	54
2.1.4 改进的 M-P 模型	55
2.2 感知器模型	56
2.2.1 问题背景	56
2.2.2 感知器的概念	57
2.2.3 单层感知器神经元模型	58
2.2.4 单层感知器工作原理	58
2.2.5 单层感知器用于模式识别	60
2.2.6 多层感知器神经元	62
2.2.7 感知器的学习规则	64
2.2.8 感知器的局限性	74

2.2.9	本节小结	74
2.3	自适应线性神经元模型	74
2.3.1	线性神经网络模型	75
2.3.2	线性神经网络的学习	76
2.3.3	线性神经网络的 MATLAB 仿真程序设计	78
2.4	BP 神经网络模型	81
2.4.1	BP 神经元及 BP 网络模型	81
2.4.2	BP 网络的学习	82
2.4.3	BP 网络学习算法	82
2.4.4	理论与实例	85
2.4.5	BP 网络的局限性	95
2.4.6	BP 网络的 MATLAB 仿真程序设计	95
2.4.7	BP 网络应用实例	97
2.5	径向基函数神经网络模型简介	102
2.5.1	径向基网络模型	102
2.5.2	径向基网络的创建与学习过程	104
2.5.3	其他径向基神经网络	105
2.5.4	径向基网络的 MATLAB 仿真程序设计	107
第 3 章	常用神经网络模型及动力学问题	109
3.1	Hopfield 神经网络模型及动力学问题	109
3.1.1	无时滞的 Hopfield 神经网络模型及动力学问题	109
3.1.2	有时滞的 Hopfield 神经网络模型及动力学问题	115
3.1.3	Hopfield 神经网络的 k -稳定性分析	121
3.2	细胞神经网络模型及动力学问题	129
3.2.1	无时滞的细胞神经网络的平衡点及稳定性	129
3.2.2	有时滞的细胞神经网络的平衡点及稳定性	141
3.2.3	无时滞细胞神经网络的周期解及稳定性	149
3.2.4	有时滞细胞神经网络的周期解及稳定性	156
3.2.5	广义细胞神经网络简介	162
3.3	BAM 神经网络模型及动力学问题	166
3.3.1	无时滞 BAM 神经网络模型及稳定性	166
3.3.2	具有连续时滞的 BAM 神经网络模型及稳定性	171
3.3.3	具有连续和离散时滞的混杂 BAM 神经网络模型及动力学问题	177
第 4 章	复杂神经网络模型及动力学问题	188
4.1	二阶 Hopfield 神经网络模型及动力学问题	188

4.1.1	二阶神经网络模型	188
4.1.2	无时滞的二阶 Hopfield 神经网络的局部稳定性分析	191
4.1.3	无时滞的二阶 Hopfield 神经网络的全局稳定性分析	199
4.1.4	具有时滞的二阶 Hopfield 神经网络的稳定性	204
4.2	具有扩散的神经网络模型和动力学问题	208
4.2.1	具有反应扩散的二阶 Hopfield 神经网络全局渐近稳定性分析	210
4.2.2	具有反应扩散的二阶 Hopfield 神经网络全局指数稳定性分析及收敛速度的估计	212
4.3	脉冲神经网络系统的动力学问题	218
4.3.1	BAM 系统的平衡点的存在性及指数稳定性	218
4.3.2	具有时滞的脉冲二阶 Hopfield 神经网络模型及动力学问题	230
4.4	随机神经网络模型及动力学问题	235
4.4.1	随机 Hopfield 神经网络模型及动力学问题	235
4.4.2	随机细胞神经网络模型及动力学问题	246
第 5 章	神经网络的应用	254
5.1	神经网络应用于模式识别	255
5.1.1	神经网络模式识别的基本知识	255
5.1.2	神经网络在手写体字符识别中的应用	257
5.1.3	基于 Hopfield 神经网络的遥感图像超分辨率识别算法	259
5.1.4	神经网络的全自动模式识别跟踪系统	265
5.1.5	RBF 神经网络应用于股票预测	269
5.2	神经网络在优化计算中的应用	273
5.2.1	连续 Hopfield 在优化计算中的应用	273
5.2.2	神经网络与其他优化算法的结合	278
5.3	神经网络应用与知识处理	280
5.4	神经网络在医学中的应用	284
参考文献		289
附录	神经网络工具箱函数	293
A.1	工具箱函数索引	293
A.2	工具箱函数详解	297

第 1 章 绪 论

人们通常所说的神经网络指的是人工神经网络 (artificial neural network, ANN), ANN 是由多个非常简单的处理单元 (神经元) 彼此按某种方式相互连接而成的计算系统, 该系统靠其状态对外部输入信息的动态响应来处理信息, 是以大脑的生理研究成果为基础模拟生物神经网络进行信息处理的一种数学模型. 其目的在于模拟大脑的某些机理与机制, 实现一些特定的功能. 人工神经网络模型主要考虑网络连接的拓扑结构、神经元的特征、学习规则等. 神经动力学就是以研究神经网络的数学模型为主要内容, 自从神经网络理论建立以后, 由于要求算法的收敛性, 人们开始了对神经网络模型动力学行为的研究, 为了研究神经动力学行为, 必须先介绍神经网络的基本理论.

1.1 神经网络简介

神经网络控制的基本思想就是从仿生学角度模拟人脑神经系统的运作方式, 使机器具有和人脑一样的感知、学习和推理功能, 将控制系统看成是一个由输入到输出的映射, 利用神经网络的学习能力和适应能力来实现系统的映射特性, 从而完成对系统的建模和控制. 目前, 神经网络的研究内容十分广泛, 这也反映出多学科交叉技术领域的特点. 其中主要研究工作概括为以下四个方面:

(1) 生物原型研究:

从生理学、心理学、脑科学、病理学、解剖学等生物科学方面研究神经细胞、神经系统、神经网络的生物原型结构及其功能机理.

(2) 理论模型的研究与建立:

在生物原型研究的基础上建立神经元、神经网络的理论模型, 其中包括知识模型、概念模型、数学模型、物理化学模型等.

(3) 研究网络模型及其算法:

在理论模型研究的基础上建立具体的神经网络模型, 以实现计算机仿真或者准备制作硬件, 其中包括网络学习算法的研究. 这方面的工作也称为技术模型研究.

(4) 神经网络应用系统研究:

在网络模型和算法研究的基础上, 通过神经网络组成实际的应用系统, 例如, 完成某类信号处理或模式识别的功能、制造机器人、构建专家系统等.

神经网络由于其大规模并行、分布式存储和处理、容错性、自组织和自适应能力以及联想功能等特点,已成为解决问题的强有力的工具,特别是非常适合处理需要同时考虑诸多因素和条件的、不精确或者模糊的信息处理问题.例如,面对缺乏物理解释和统计解释、数据由非线性机制产生、观察的数据中存在着统计变化等棘手问题时,神经网络往往能够提供非常有效的解决办法.另一方面,神经网络对突破现有科学技术的瓶颈,更深入地探索研究非线性等复杂现象具有非常重大的意义.此外,神经网络理论在实际中的应用也取得了令人瞩目的发展,特别是在信息处理、智能控制、模式识别、非线性优化、生物医学工程等方面都有重要的应用实例.根据 Mritin T.Hagen 等的归纳总结,神经网络在实际生活中也有着诸多应用.例如,汽车自动驾驶系统、汽车调度和路线系统、动画和特效设计、产品最优化、货币价格预测、脑电图和心电图分析、石油和天然气的勘探等.我们相信随着神经网络研究的进一步深入和拓展,特别是神经网络作为一种智能方法与其他学科技术领域更为紧密的结合,神经网络将具有更为广阔的应用前景.

1.1.1 神经网络的概念

神经网络的全称是人工神经网络,是由大量的、非常简单的处理单元(神经元)彼此按某种方式互连而成的复杂网络系统,它通过对人脑的抽象、简化和模拟反映人脑功能的基本特性,是一个高度复杂的非线性动力学系统.神经网络的研究是以人脑的生理结构为基础来研究人的智能行为,模拟人脑信息处理能力.神经网络的发展与数理科学、神经科学、计算机科学、人工智能、信息科学、分子生物学、控制论、心理学等相关,因此,神经网络是一门特别活跃的边缘交叉学科.神经网络控制的基本思想就是从仿生学角度模拟人脑神经系统的运作方式,使机器具有和人脑一样的感知、学习和推理功能,将控制系统看成是一个由输入到输出的映射,利用神经网络的学习能力和适应能力来实现系统的映射特性,从而完成对系统的建模和控制.为了更好地理解什么是神经网络,下面简单介绍人脑结构.

人的神经系统可看作三个阶段系统,如同图 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,感知它并做出适当的决定.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.神经网络的效应器将电冲击通过神经网络转换为可识别的响应作为系统输出.

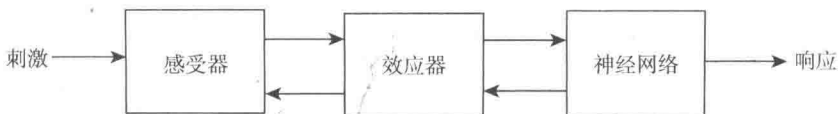


图 1.1.1 神经系统的框图表示

人脑神经系统的基本单元是神经细胞,即生物神经元,人脑神经系统约由 10^{11} 个神经元构成,每个神经元与约 10^4 个其他神经元相连接.神经细胞与人体中其他细胞的关键区别在于神经细胞具有产生、处理和传递信号的能力.一个神经元的构造主要包括细胞体、树突、轴突和突触,如图 1.1.2 所示.

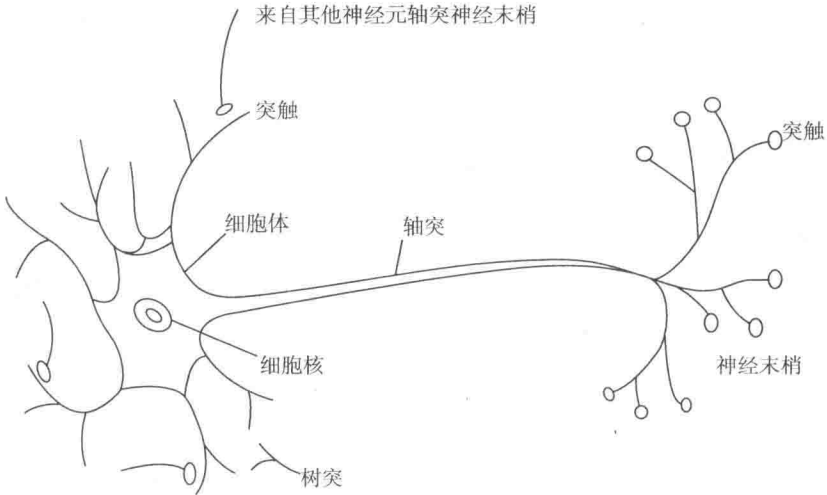


图 1.1.2 生物神经元示意图

细胞体: 由细胞核、细胞质和细胞膜等组成.

树突: 从细胞体延伸出来像树枝一样向四处分散开来的许多突起,称之为树突,其作用是感受其他神经元的传递信号.

轴突: 一般每个神经元从细胞体伸出一根粗细均匀、表面光滑的突起,长度从几微米到 1m 左右,称为轴突,它的功能是传来自细胞体的神经信息.在高等动物的神经细胞中,大多数神经元都有轴突.

突触: 轴突末端有许多细的分枝,称之为神经末梢,每一根神经末梢可以与其他神经元连接,其连接的末端称为突触.

神经元之间的连接是靠突触实现的,主要有:轴突与树突、轴突与细胞体、轴突与轴突、树突与树突等连接形式.

神经细胞单元的信息是宽度和幅度都相同的脉冲串,若某个神经细胞兴奋,其轴突输出的脉冲串的频率就高;若某个神经细胞抑制,其轴突输出的脉冲串的频率就低,甚至无脉冲发出.根据突触对下一个神经细胞的功能活动的影响,突触又可分为兴奋性的和抑制性的两种.兴奋性的突触可能引起下一个神经细胞兴奋,抑制性的突触使下一个神经细胞抑制.

神经细胞的细胞膜将细胞体内外分开,从而使细胞体内外有不同的电位,一般

内部电位比外部低,其内外电位差称为膜电位.突触使神经细胞的膜电位发生变化,且电位的变化是可以累加的,该神经细胞膜电位是它所有突触产生的电位总和.当该神经细胞的膜电位升高到超过一个阈值时,会产生一个脉冲,从而总和的膜电位直接影响该神经细胞兴奋发放的脉冲数.

突触传递信息需要一定的延迟,对温血动物,延迟时间为 0.3~1.0ms.一般每个神经细胞的轴突连接 100~1000 个其他神经细胞,神经细胞的信息就这样从一个神经细胞传到另一个神经细胞,且这种传播是正向的,不允许逆向传播.

1.1.2 神经网络的发展

神经网络主要经过早期阶段、过渡期、新高潮期、热潮期四个阶段:

(1) 早期阶段指的是 1913 年到 20 世纪 60 年代末. 60 年代中期神经网络的研究处于低潮,在这期间研究人员提出了许多神经元模型和学习规则. 1913 年人工神经系统的第一个实践是由 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,提出了 M-P 模型; 1949 年心理学家 D.O. Hebb 提出了突触联系效率可变的假设,这种假设就是调整权值; 1957 年 F. Rosenblatt 设计制作了著名的感知器; 1962 年 Bernard Widrow 和 Marcian Hoff 提出了自适应线性元件网络.

(2) 20 世纪 60 年代末到 20 世纪 70 年代为过渡期,这期间神经网络研究跌入了一个低潮.人们开始发现感知器存在一些缺陷,例如它不能解决异或问题,因而使研究工作陷入了谷底.难能可贵的是,在这一时期,仍有众多学者在极端艰难的环境下持之以恒地对神经网络进行研究.例如, Grossberg 提出了自适应共振理论; Kohonen 提出了自组织映射; Fukushima 提出了神经认知机网络理论; Anderson 提出了 BSB 模型; Webos 提出了 BP 理论; 还有这一时期日本福岛彦帮的认知机模型和日本东京大学的中野馨的联想记忆模型等.

(3) 新高潮期指的是 20 世纪 80 年代,这一时期主要是 Hopfield 神经网络模型,引入了“计算能量函数”的概念,给出了网络稳定性判断依据,有力地推动了神经网络的研究与发展.

(4) 热潮期指的是 20 世纪 80 年代后期,1986 年 Rumelhart 和 McClelland 等提出并行分布处理 (PDP) 的理论,与此同时还提出了多层次网络的误差反向传播学习算法,简称 BP 算法.这种算法从实践上证明神经网络具有很强的运算能力,可以完成许多学习任务,解决许多具体问题,自 1986 年以来,在控制领域,将神经网络与传统控制技术相结合取得了许多令人鼓舞的结果.神经网络理论的应用研究已经渗透各个领域,并在智能控制、模式识别、自适应滤波和信号处理、非线性优化、传感技术和机器人、生物医学工程等方面取得了令人鼓舞的进展.这些成就加强了人们对神经网络系统的进一步认识,引起了世界许多国家的科学家、研究机构及企

业界人士的关注,也促使不同学科的科学工作者联合起来,从事神经网络理论、技术开发及应用于现实的研究。

1.1.3 神经网络的优点

在学习神经网络优点之前,先了解神经网络的计算能力和性质,计算能力有两点:

(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的(大型)问题。但是在实践中,神经网络不能单独做出解答,它们需要被整合在一个协调一致的系统工程方法中。具体来讲,一个复杂问题往往被分解成若干相对简单的任务,而神经网络则处理与其能力相符的子任务。

神经网络具有下列性质和能力:

(1) 非线性。一个神经元可以是线性或者是非线性的。一个由非线性神经元互连而成的神经网络,其自身是非线性的,并且非线性是一种分布于整个网络中的特殊性质。

(2) 输入输出映射。有监督学习或有教师学习是一个学习的流行范例,涉及使用带标号的训练样本或任务例子对神经网络的突触权进行修改。每个样本由一个唯一的输入信号和相应的期望响应组成。从一个训练集中随机选取一个样本给网络,网络就调整它的突触权值(自由参数),以最小化期望响应和由输入信号以适当的统计准则产生的实际响应之间的差别。使用训练集中的例子很多,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。先前使用过的例子可能还要在训练期间以不同顺序重复使用,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。

(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地进行重新训练。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。

(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,也能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) 背景的信息。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

经元潜在地都受到网络中所有其他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。比如, 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实验证据, 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) 超大规模集 (very-large-scale-integrated, 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 VLSI 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, 即涉及神经网络的应用的所有领域都使用同样记号。这种特征以不同的方式表现出来:

① 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。

② 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。

③ 模块化网络可以用模块的无缝集成来实现。

基于神经网络以上能力和性质, 所以神经网络具有以下优点:

(1) 分布式存储信息. 其信息的存储分布在不同的位置, 神经网络是用大量神经元之间的连接及对各连接权值的分布来表示特定的信息, 从而使网络在局部网络受损或输入信号因各种原因发生部分畸变时, 仍然能够保证网络的正确输出, 提高网络的容错性和鲁棒性。

(2) 并行协同处理信息. 神经网络中的每个神经元都可以根据接收到的信息进行独立的运算和处理, 并输出结果, 同一层中的各个神经元的输出结果可被同时计算出来, 然后传输给下一层做进一步处理, 这体现了神经网络并行运算的特点。这一特点使网络具有非常强的实时性。虽然单个神经元的结构极其简单, 功能有限, 但大量神经元构成的网络系统所能实现的行为是极其丰富多彩的。

(3) 信息处理与存储合二为一, 神经网络的每个神经元都兼有信息处理和存储功能。神经元之间连接强度的变化, 既反映了对信息的记忆, 同时又与神经元对激励的响应一起反映了对信息的处理。

(4) 对信息的处理具有自组织、自学习的特点, 便于联想、综合和推广。神经网络的神经元之间的连接强度用权值大小来表示, 这种权值可以通过对训练样本的学习而不断变化, 而且随着训练样本量的增加和反复学习, 这些神经元之间的连接强

度会不断增加,从而提高神经元对这些样本特征的反应灵敏度。

1.1.4 人工神经元模型

神经元是神经网络操作的基本信息处理单位,是(人工)神经网络的设计基础。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 的突触权值 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,正值表示兴奋性突触,负值表示抑制性突触。

(2) 加法器,用于求输入信号中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(励)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以写成单位闭区间 $[0, 1]$ 或者另一种区间 $[-1, 1]$ 。

图 1.1.3 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。

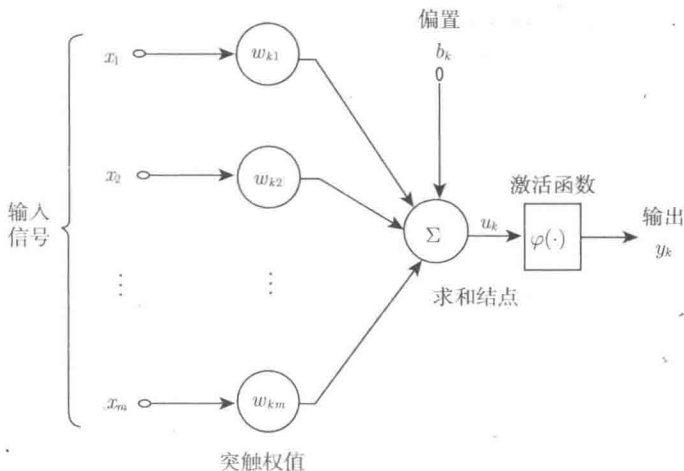


图 1.1.3 神经元的非线性模型

用数学术语,可以用如下对方程描述一个神经元 k :

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

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

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 , 激活函数是 $\varphi(\cdot)$; y_k 是神经元输出信号. 偏置 b_k 的作用是对图 1.1.3 模型中的线性组合器的输出 u_k 作仿射变换, 如下所示:

$$v_k = Wp + b = \sum_{j=1}^m w_{kj}x_j + b_k. \quad (1.1.3)$$

特别地, 根据偏置 b_k 取正或取负, 神经元 k 的诱导局部或激活电位 v_k 和线性组合器输出 u_k 的关系如图 1.1.4 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.

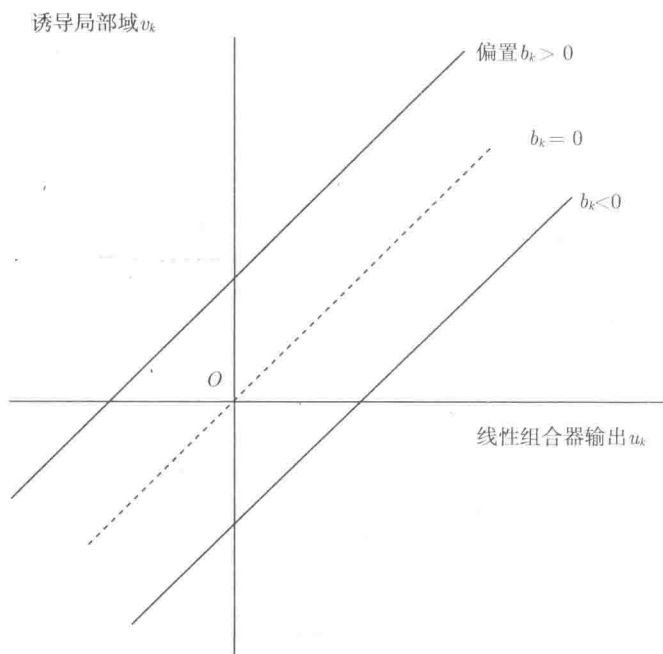


图 1.1.4 偏置产生的仿射变换 (注意 $u_k = 0$ 时 $v_k = b_k$)

偏置 b_k 是人工神经元 k 的外部参数. 结合方程 (1.1.1) 和 (1.1.3) 得到如下公式:

$$v_k = \sum_{j=0}^M w_{kj}x_j, \quad (1.1.4)$$

$$y_k = \varphi(v_k). \quad (1.1.5)$$

在 (1.1.4) 中, 我们加上一个新的突触, 其输入是

$$x_0 = 1, \quad (1.1.6)$$

权值是

$$w_{k0} = b_k, \quad (1.1.7)$$

因此得到了神经元 k 的新型模型图 1.1.5. 在这个图形中, 偏置的作用是做两件事: ①添加新的固定的输入 $+1$; ②添加新的等于偏置 b_k 的突触权值. 虽然形式上图 1.1.3 和图 1.1.5 的模型不相同, 但在数学上它们是等价的.

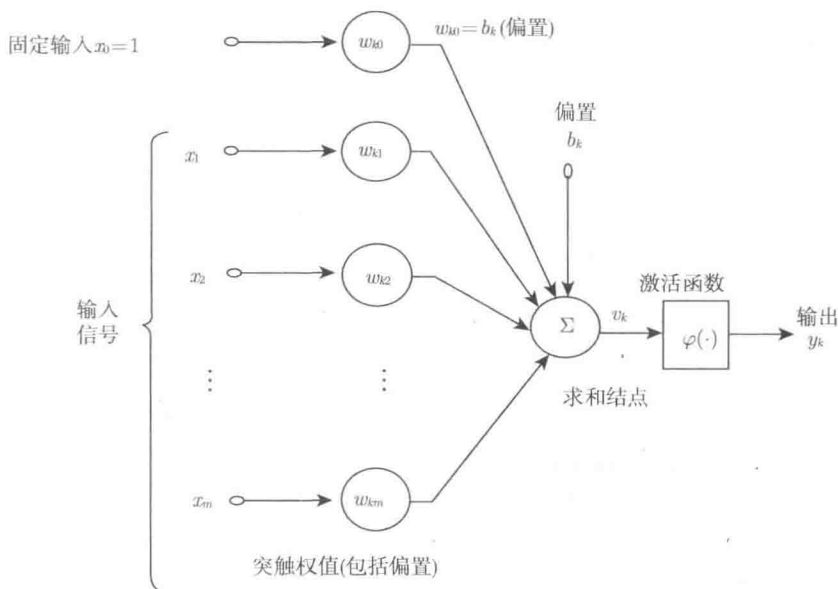


图 1.1.5 神经元的另一个非线性模型

1. 传输 (激活) 函数类型

传输函数, 记为 $f(v)$, 通过诱导局部域 v 定义神经元输出. 这里主要介绍三种基本的传输函数:

(1) 阈值函数. 这种传输函数如图 1.1.6(a) 所示, 可写为

$$\varphi(v) = \begin{cases} 1, & v \geq 0, \\ 0, & v < 0. \end{cases} \quad (1.1.8)$$

相应地, 在神经元 k 使用这种阈值函数, 其输出可表示为

$$y_k = \begin{cases} 1, & v_k \geq 0, \\ 0, & v_k < 0, \end{cases} \quad (1.1.9)$$