

神经网络结构设计的 理论与方法

魏海坤 编著

国防工业出版社

<http://www.ndip.cn>

神经网络结构设计的 理论与方法

魏海坤 编著

国防工业出版社

·北京·

内 容 简 介

本书是作者在东南大学多年从事神经网络理论及应用研究教学的基础上编写的神经网络设计研究生课程教材。

本书在介绍常用的神经元模型和权值学习算法的基础上,系统全面地介绍了神经网络,尤其是前馈神经网络结构设计的原理和方法。具体内容包括:影响神经网络泛化能力的各种因素,如结构复杂性、样本、初始权值、学习时间、先验知识等;这些因素如何影响神经网络的性能;常用的神经网络结构优化设计方法,包括剪枝算法(权衰减法、灵敏度计算方法、相关性剪枝方法等)、构造算法(级连相关算法、资源分配网络等)、进化方法等;常用的神经网络参数优化设计方法,包括最优停止方法、主动学习方法、样本输入中加噪声、神经网络集成等。这些方法绝大部分是目前国内外广泛使用的最典型的前馈神经网络设计方法。本书介绍了它们的实现原理和仿真例子,并提供大部分算法的 MATLAB 实现代码。

本书适用于自动化、信号处理等学科的工程技术人员、高年级学生、研究生及教师使用。

图书在版编目(CIP)数据

神经网络结构设计的理论与方法 / 魏海坤编著. —北京:国防工业出版社, 2005.2
ISBN 7-118-03679-X

I. 神... II. 魏... III. 人工神经网络-结构设计 IV. TP183

中国版本图书馆 CIP 数据核字(2004)第 122275 号

国防工业出版社出版发行

(北京市海淀区紫竹院南路 23 号)

(邮政编码 100044)

天利华印刷装订有限公司印刷

新华书店经售

开本 787×1092 1/16 印张 15¼ 347 千字

2005 年 2 月第 1 版 2005 年 2 月北京第 1 次印刷

印数:1—3000 册 定价:21.00 元

(本书如有印装错误,我社负责调换)

国防书店:(010)68428422

发行邮购:(010)68414474

发行传真:(010)68411535

发行业务:(010)68472764

前 言

神经网络应用时必须首先解决结构设计问题,目前国内尚无系统介绍神经网络结构设计理论与方法的教材或专著。

本书在介绍常用的神经元模型和权值学习算法的基础上,重点介绍了神经网络结构设计的原理和方法,包括:影响神经网络泛化能力的因素;这些因素如何影响神经网络的性能;常用的神经网络结构优化设计方法,包括剪枝算法(权衰减法、灵敏度计算方法、相关性剪枝方法等)、构造算法(级连相关算法,资源分配网络等)、进化方法等;常用的神经网络参数优化设计方法,包括最优停止方法、主动学习方法、神经网络集成等。这些方法绝大部分是目前国内外广泛使用的最典型的神经网络设计方法。本书介绍它们的实现原理和仿真例子,并提供大部分算法的 MATLAB 实现代码。

本书适用于自动化、信号处理等学科的工程技术人员、高年级学生、研究生及教师使用。

由于时间仓促,作者水平有限,本书难免存在疏漏或不妥之处,欢迎读者批评指正。

目 录

第 1 章 概论	1
1.1 神经网络简介	1
1.2 神经网络研究内容	2
1.3 本书内容	2
1.4 参考资料简介	3
第 2 章 基本的神经元及其学习规则	5
2.1 神经元模型	5
2.1.1 大脑神经细胞	5
2.1.2 MP 模型	7
2.1.3 一般神经元模型	8
2.2 神经元学习算法	11
2.2.1 Hebb 学习规则	12
2.2.2 离散感知器学习规则	12
2.2.3 δ 学习规则	13
2.2.4 Widrow-Hoff 学习规则	14
2.3 单个神经元解决问题的能力	15
2.3.1 单个神经元的分类能力	15
2.3.2 多个神经元的分类能力	15
2.4 神经网络的拓扑结构	16
2.4.1 前向神经网络	16
2.4.2 反馈神经网络	18
第 3 章 多层感知器网络	19
3.1 两层感知器网络	19
3.1.1 单神经元用于两分类	19
3.1.2 多输出两层感知器用于多分类	23
3.2 线性阈值单元组成的多层感知器	24
3.3 BP 网及 BP 算法	25
3.3.1 BP 网络结构	25
3.3.2 BP 学习算法	26
3.3.3 BP 算法应用例子	30

3.3.4	BP 网和 BP 算法应用要点	33
3.3.5	BP 算法的改进	36
第 4 章	径向基函数神经网络	40
4.1	RBF 网结构和工作原理	40
4.2	RBF 网的生理学基础	42
4.3	RBF 网的数学基础	43
4.3.1	内插问题	43
4.3.2	正则化网络	44
4.4	RBF 网常用的学习算法	45
4.4.1	聚类方法	46
4.4.2	梯度训练方法	47
4.4.3	正交最小二乘(OLS)学习算法	48
4.5	RBF 网的学习动态	49
4.5.1	定义	50
4.5.2	主要结论	52
4.5.3	算例	58
4.5.4	RBF 网的学习动态与 RBF 网设计	62
4.6	仿真例子	64
4.7	RBF 网的特点及其他问题	65
第 5 章	Hopfield 网络	67
5.1	连续 Hopfield 网络	67
5.1.1	连续 Hopfield 网络原理	67
5.1.2	Hopfield 网络的稳定性	68
5.2	离散 Hopfield 网络	71
5.2.1	离散 Hopfield 网络原理	71
5.2.2	离散 Hopfield 网络的稳定性	72
5.3	Hopfield 网络应用	73
5.3.1	连续 Hopfield 网络与优化计算	73
5.3.2	离散 Hopfield 网络与联想存储	73
5.4	应用和仿真实例	75
5.4.1	连续 Hopfield 网络应用实例:TSP 问题	75
5.4.2	离散 Hopfield 网络仿真:字符存储	77
第 6 章	自组织特征映射	80
6.1	生物系统中的竞争	80
6.2	SOFM 结构	81
6.3	SOFM 的学习算法	82

6.4	仿真例子	84
第7章	神经网络的泛化理论	88
7.1	神经网络的泛化理论简介	88
7.2	泛化误差的偏差-标准差分解	89
7.3	结构复杂性和样本复杂性对神经网络泛化能力的影响	90
7.3.1	线性阈值神经网络	90
7.3.2	函数逼近神经网络	91
7.4	正则化方法对泛化能力的影响	92
7.5	神经网络集成对泛化能力的影响	94
7.6	样本输入中加噪声对泛化能力的影响	96
7.7	其他因素对泛化能力的影响	97
第8章	神经网络的参数优化设计	100
8.1	主动学习	100
8.1.1	原理	100
8.1.2	仿真例子:三角形概念学习	101
8.2	在样本输入中添加随机噪声	103
8.2.1	噪声添加方法	103
8.2.2	仿真例子	103
8.3	神经网络集成	105
8.3.1	原理	105
8.3.2	仿真例子:广义异或问题	106
8.4	基于先验知识的泛化方法	107
8.5	最优停止法	107
8.5.1	原理	107
8.5.2	仿真例子:Hermit 多项式逼近	108
第9章	神经网络构造方法	110
9.1	神经网络构造方法简介	110
9.2	级连相关算法	111
9.2.1	BP 算法收敛速度慢的原因	111
9.2.2	网络结构	112
9.2.3	权值学习算法	114
9.2.4	算法实现	115
9.2.5	算法讨论	115
9.2.6	仿真结果:双螺旋分类问题	115
9.3	资源分配网络	119
9.3.1	RAN 网络结构	119

9.3.2	学习算法	120
9.3.3	算法实现	121
9.3.4	RAN 的优点与缺点	122
9.3.5	仿真: Hermit 多项式在线学习	122
第 10 章	神经网络剪枝方法	125
10.1	权衰减法	125
10.1.1	权消去法剪枝原理	125
10.1.2	正则化系数 λ 的动态修改策略	129
10.1.3	试验例子: 非线性系统辨识	129
10.2	灵敏度计算方法	131
10.2.1	Skeletonization 方法原理	131
10.2.2	仿真例子: 非线性系统结构辨识	134
10.3	相关性剪枝方法	136
10.3.1	隐节点之间相关情况分析	136
10.3.2	隐节点合成方法	138
10.3.3	仿真例子	140
第 11 章	进化神经网络方法	142
11.1	进化神经网络简介	142
11.2	进化优选算法	142
11.2.1	OLS 算法设计 RBF 网的缺点	142
11.2.2	进化优选算法的数学基础	143
11.2.3	进化优选算法(ESA 算法)	146
11.2.4	算法的效率分析	148
11.2.5	ESA 算法实现	149
11.2.6	仿真例子	150
第 12 章	混杂神经网络设计	153
12.1	资源优化网络	153
12.1.1	资源优化网络简介	153
12.1.2	RON 在线 RBF 网设计方法	154
12.1.3	仿真研究	159
12.1.4	RON 的优缺点	163
12.2	神经网络的结构分解方法	164
12.2.1	神经网络的结构分解方法简介	164
12.2.2	神经网络的结构分解原理	164
12.2.3	基于结构分解的神经网络设计方法	167
12.2.4	应用例子	169

附录 A: 用于三分类的 BP 算法程序	175
附录 B: 用于函数逼近的 BP 算法程序	180
附录 C: 基于聚类的 RBF 网设计算法	183
附录 D: 基于梯度法的 RBF 网设计算法	186
附录 E: 基于 OLS 的 RBF 网设计算法	188
附录 F: 连续 Hopfield 网络解决 TSP 的程序	191
附录 G: 样本输入加噪声训练程序	194
附录 H: 神经网络集成训练程序	198
附录 I: 最优停止法训练程序	200
附录 J: 级连相关算法	204
附录 K: 资源分配网络算法	209
附录 L: 权消去法	213
附录 M: Skeletonization 方法	217
附录 N: 隐节点合成算法	222
参考文献	229

第 1 章 概 论

1.1 神经网络简介

“神经网络”或“人工神经网络”是指用大量的简单计算单元(即神经元)构成的非线性系统,它在一定程度和层次上模仿了人脑神经系统的信息处理、存储及检索功能,因而具有学习、记忆和计算等智能处理功能。神经网络具有一些显著的特点:具有非线性映射能力;不需要精确的数学模型;擅长从输入输出数据中学习有用知识;容易实现并行计算;由于神经网络由大量简单计算单元组成,因而易于用软硬件实现;等等。正因为神经网络是一种模仿生物神经系统构成的新的信息处理模型,并具有独特的结构,所以人们期望它能解决一些用传统方法难以解决的问题。

目前,神经网络的理论和应用研究得到了极大的发展,而且已经渗透到几乎所有的工程应用领域。但是,人工神经网络的发展过程并不是一帆风顺的。神经网络研究大致经历了以下几个阶段。

1. 初始发展期

1943年,W. McCulloch 和 W. Pitts 提出了 MP 模型¹,从而给出了神经元的最基本模型及相应的工作方式。

1949年,神经生物学家 D. Hebb 发现,脑细胞之间的通路在参与某种活动时将被加强²。这重要规则给出了生理学与心理学间的联系,被称为 Hebb 学习规则。该规则至今还被许多神经网络学习算法所使用。

1957年,F. Rosenblat 提出了感知器模型³,这是一个由线性阈值神经元组成的前馈神经网络,可用于分类。

1960年,B. Widrow 和 M. Hoff 提出了自适应线性单元⁴,这是一种连续取值的神经网络,可用于自适应系统。

2. 低潮时期

1969年,人工智能的创始人 M. Minsky 和 S. Papert 发表了《感知器》(《Perceptrons》)一书⁵,在该书中,他们指出:单层 Perceptron 只能作线性划分,多层 Perceptron 不能给出一种学习算法,因此无实用价值。由于 Minsky 和 Papert 在人工智能领域的地位,该书在人工神经网络研究人员间产生了极大的反响,神经网络研究自此陷入低潮。

但是,即便在神经网络研究的低潮时期,也有一些人仍在兢兢业业地研究神经网络,并得到了一些重要成果。其中最著名的是 1982年由加州理工大学教授 H. Hopfield 提出的 Hopfield 神经网络⁶。在这个用运算放大器搭成的反馈神经网络中,Hopfield 借用 Lyapunov 能量函数的原理,给出了网络的稳定性判据,并为著名的组合优化问题——旅行商问题(TSP)提供了一个新的解决方案。Hopfield 网络可用于联想存储、优化计算等领域。

3. 高潮时期

1985年,Rumelhart等人给出了多层感知器的权值训练的误差反向传播学习算法(BP算法)⁷,从而解决了Minsky认为不能解决的多层感知器的学习问题。自此引导了神经网络的复兴,神经网络研究也进入了一个崭新的发展阶段。

1.2 神经网络研究内容

当前神经网络研究内容主要包括神经网络理论研究、神经网络实现技术和神经网络应用研究3个方面。

1. 神经网络理论研究

神经网络理论研究侧重于寻找合适的神经网络模型和学习算法。其中,模型研究是指构造合适的单个神经元模型,及确定神经元之间的连接方式,并探讨它所适用的场合;学习算法研究是指在神经网络模型的基础上,找出一种调整神经网络结构和权值的算法,并满足学习样本的要求,同时具有较快的学习速度。神经网络理论研究的另一个重要内容是从理论上分析常用的神经网络设计方法对泛化能力的影响。

2. 神经网络实现技术研究

神经网络实现技术研究主要是探讨利用电子、光学、光电、生物等技术实现神经计算机的途径,包括利用传统计算机技术实现模拟神经计算机,及新型神经计算机体系结构的研究等。

3. 神经网络应用研究

神经网络应用研究是探讨如何利用神经网络解决实际工程问题。人们可以在几乎所有的领域中发现神经网络应用的影子。当前神经网络的主要应用领域有:模式识别、故障检测、智能机器人、非线性系统辨识和控制、市场分析、决策优化、物资调用、智能接口、知识处理和认知科学等。

1.3 本书内容

与人脑的作用机理类似,一个神经网络完成任务的过程包括学习(训练)过程和使用(回忆或联想)过程。一般来说,只有经过充分学习的神经网络人们才会放心使用。因此寻找好的神经网络的学习算法,即确定最优网络结构和参数的方法是人们最关心的。对一个神经网络学习算法来说,衡量其性能优劣的指标有如下几方面:

(1) 泛化能力:即一个训练好的神经网络到实际中使用是否有好的效果。这无疑是神经网络最重要的性能指标。

(2) 时间复杂性:即训练一个固定结构的神经网络所需要的时间。

(3) 空间复杂性:即算法计算机实现时占用的内存空间,一般与神经网络的结构复杂程度密切相关。

(4) 在线学习能力:如果神经网络的学习过程和使用过程是分别进行的,即先学习后使用,则称为离线学习;如果这两个过程是同时进行的,即边学习边使用,则称为在线学习。至于采用哪种学习机制由具体应用决定。

(5) 其他指标,如能否用硬件实现、算法的稳定性、神经网络模型的鲁棒性等。

在这些指标中,人们最关心的是神经网络的泛化能力。神经网络具有学习能力,即能够自动从一组样本(有导或无导学习)中提取有用信息。尽管给定训练数据,并选择一种学习算法,任何人都能使用这种现成的算法训练神经网络,但事实上,由于训练后神经网络的泛化性能与神经网络设计过程密切相关,因此本书的重点是讲述基于泛化的神经网络的设计方法。

为了设计一个性能满意的神经网络,通常使用两类设计方法:结构优化设计方法和参数优化设计方法。参数优化设计方法通常选择一个结构“足够大”的初始网络,然后通过优化网络参数来达到满意的性能;结构优化设计方法则通过设计满足精度要求的结构尽可能小的网络,以保证网络的泛化性能。这两类方法本书都将详细介绍,并尽可能给出这些方法的 MATLAB 实现代码,供相关人员参考。

在以后各章中将顺序介绍以下内容。

1. 神经网络基本理论

第 2 章介绍基本的神经元及其学习规则。

2. 基本前向神经网络学习算法

第 3 章和第 4 章介绍多层感知器网络和径向基函数神经网络。

3. 最常用的反馈网络及自组织网络

第 5 章和第 6 章介绍 Hopfield 网络及自组织特征映射。

4. 神经网络的泛化理论

第 7 章介绍影响神经网络泛化能力的因素,如结构、样本、初始权、学习时间等,以及定量分析一些重要的神经网络设计方法(正则化、神经网络集成、样本输入中加噪声等)对泛化能力的影响。本书主要介绍这些理论的最新研究成果,并且给出已有的定性或定量的结果。

5. 神经网络参数优化设计方法

第 8 章介绍神经网络的参数优化设计方法,包括主动学习、神经网络集成、基于先验知识的泛化方法、最优停止等。

6. 神经网络结构优化设计方法

第 9 章~第 12 章也是本书的最重要的内容之一。介绍主要的前馈神经网络的结构设计方法,如剪枝算法、构造算法、进化算法等。本书在介绍各类算法的原理和实现方法的同时,也给出了算法的实现代码。

1.4 参考资料简介

教材:尽管很少见到全面介绍神经网络结构设计的理论与方法的教材或专著,但有关神经网络基本理论的好教材却非常多,如教材^{8,9,10,11}等均可供参考。

期刊:可供查阅的期刊很多,国内期刊如《电子学报》、《自动化学报》、《软件学报》、《计算机学报》、《控制理论与应用》、《控制与决策》、《信息与控制》等;神经网络研究的专业国际期刊包括《Neural Computation》、《IEEE Transaction on Neural Networks》、《Neural Networks》、《Advances in Neural Information Processing Systems》(简称 NIPS)、《Neurocomputing》、《Neural

Processing Letters》等。这些期刊中,《Neural Computation》是目前最有影响的神经网络专业杂志。

网络数据库:绝大部分上述国内外期刊的电子版资料均可通过网络数据库检索,如中国期刊网(CNKI)、IEEE/IEE、Elsevier、EBSCO、ACM等。

机器学习数据库:最著名的是美国加州大学欧文分校的机器学习数据库,网址为<http://www.ics.uci.edu/mlearn/MLRepository.html>。这些数据库均提供了许多实际工程应用和理论研究中的训练样本,可以免费下载,研究人员可以用这些样本测试自己的算法,并与其他同类算法对比。

第 2 章 基本的神经元及其学习规则

本章先介绍了大脑神经细胞的组成及神经元之间的信息传递过程,在此基础上给出了简化的神经元数学模型,并讲述了神经元模型中基函数和激活函数的类型。然后讲述单个神经元的学习规则,包括 Hebb 学习规则、 δ 学习规则、Widrow-Hoff 学习规则等。由于单个神经元能力有限,设置不能解决异或问题,而多个神经元可以轻而易举地解决这个问题,因此引入了神经网络的概念,并介绍了常用的神经网络结构。

2.1 神经元模型

2.1.1 大脑神经细胞

1. 神经细胞的组成

神经细胞又称为神经元(neuron),是大脑神经组织的主要成分。大脑神经元的数量庞大,形态多样,结构复杂。大脑神经元在生理上具有感受刺激、传导冲动和产生反应等功能。

神经元包括胞体(soma)和突起两部分,其中突起又分树突(dendrite)和轴突(axon)两种。

1) 胞体

神经元的胞体在于脑和脊髓的灰质及神经节内,是神经元的代谢和营养中心。胞体的结构与一般细胞相似,有核仁、细胞膜、细胞质和细胞核。其中细胞膜是一个敏感而易兴奋的膜,在膜上有各种受体(receptor)和离子通道(ionic channel)。形成突触部分的细胞膜增厚。膜上受体可与相应的化学物质神经递质结合,使膜的离子通透性及膜内外电位差发生改变,从而使胞膜产生相应的生理活动:兴奋或抑制。

2) 突起

神经元的突起是神经元胞体的延伸部分,由于形态结构和功能的不同,可分为树突和轴突。

(1) 树突

树突是从胞体发出的一个至多个突起,呈放射状。靠近胞体部分较粗,经反复分支而变细,形如树枝状。树突具有接受刺激并将冲动传入细胞体的功能。

(2) 轴突

轴突是一根长神经纤维,其主要功能是将神经冲动由胞体传至其他神经元。轴突传导神经冲动的起始部位是在轴突的起始段,沿轴膜进行传导。每个神经元只有一根轴突。

3) 突触

神经元与神经元之间的连接点,称为突触(synapse)。它是神经元之间的传递信息关键性结构。突触可分两类,即化学性突触(chemical synapse)和电突触(electrical synapse)。

(1) 化学性突触

化学性突触位于轴突末端,呈球状或环状膨大,附在另一个神经元的胞体或树突表面。通常一个神经元有许多突触,可接受多个神经元传来的信息。根据神经元种类不同,突触数目为几千个到几十万个不等。

化学性突触由3部分组成:突触前部、突触间隙和突触后部。突触前部和突触后部相对应的细胞膜较其余部位略增厚,分别称为突触前膜和突触后膜,两膜之间的狭窄间隙称为突触间隙。在突触前膜部位的胞浆内,含有许多突触小泡(synaptic vesicle)。突触小泡内含有化学物质,称为神经递质(neurotransmitter)。各种神经递质在胞体内合成,形成小泡,通过轴突运输到轴突末端。突触后膜具有受体和化学门控的离子通道。根据突触前膜和后膜的胞质面致密物质厚度不同,可将突触分为I型和II型:一般认为I型突触是兴奋性突触,II型突触是一种抑制性突触。

化学性突触的特征是:一侧神经元通过出胞作用释放小泡内的神经递质到突触间隙,相对应一侧的神经元(或效应细胞)的突触后膜上有相应的受体。具有这种受体的细胞称为神经递质的效应细胞或靶细胞,这就决定了化学突触传导为单向性。突触的前后膜是两个神经膜特化部分,维持两个神经元的结构和功能,实现机体的统一和平衡。故突触对内、外环境变化很敏感,如缺氧、酸中毒、疲劳和麻醉等,可使兴奋性降低;茶碱、碱中毒等则可使兴奋性增高。

(2) 电突触

电突触是神经元间传递信息的最简单形式。此时,轴突末端无突触小泡,传导不需要神经递质,是以电流传递信息,传递神经冲动一般均为双向性。神经细胞间电阻小,通透性好,局部电流极易通过。电突触功能有双向快速传递的特点,传递空间减少,传递更有效。

电突触对内、外环境变化很敏感。在疲劳、缺氧、麻醉或酸中毒情况下,可使兴奋性降低;而在碱中毒时,可使兴奋性增高。

2. 神经元之间的信息传递

图 2.1 所示为神经元之间的连接示意图。

如前所述,神经元之间通过突触传递信息。当神经冲动传至突触前膜时,突触小泡移向突触前膜,以胞吐方式释放小泡内的神经递质,其中部分神经递质与突触后膜上的相应受体结合,后膜内外两侧的离子分布状况发生改变,呈现兴奋性或抑制性变化,从而影响突触后神经元(或效应细胞)的活动。使突触后膜发生兴奋的突触,称兴奋性突触(excitatory synapse),而使后膜发生抑制的称抑制性突触(inhibitory synapse)。突触的兴奋或抑制决定于神经递质及其受体的种类。

由于一个神经元通常有许多突触,其中有些是兴奋性的,有些是抑制性的,如果兴奋性突触活动强度总和超过抑制性突触活动强度总和,并达到一定阈值,就能使该神经元的轴突起始段发生动作电位,产生神经冲动。出现神经冲动时,则该神经元呈现兴奋,反之,则表现为抑制。

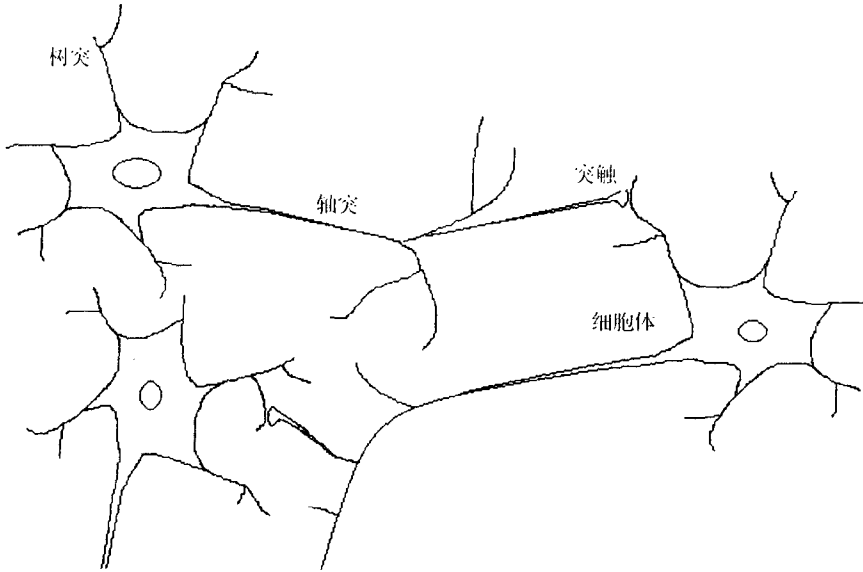


图 2.1 神经元之间的连接

2.1.2 MP 模型

人工神经网络的第一个数学模型是由 McCulloch 和 Pitts 建立的¹。该模型是基于这样一种思想：神经细胞的工作方式是或者兴奋，或者抑制。基于这个思想，McCulloch 和 Pitts 在神经元模型中引入了硬极限函数，该函数形式后来被其他神经网络（多层感知器、离散 Hopfield 网络）采用。

由于神经元之间的信号连接强度取决于突触状态，因此在 MP 模型中，神经元的每个突触的活动强度用一个固定的实数即权值模拟。于是每个神经元模型都可以从数十个甚至数百个其他神经元接收信息，产生神经兴奋和冲动；同时，在其他条件不变的情况下，不论何种刺激，只要达到阈值以上，就能产生一个动作电位。但如果输入总和低于阈值，则不能引起任何可见的反应。

图 2.2 所示为 MP 模型示意图。图中 x_1, x_2, \dots, x_n 为神经元的输入 w_1, w_2, \dots, w_n 为相应的连接权值， T 为神经元的兴奋阈值， y 为神经元的输出。神经元的输出取二值函

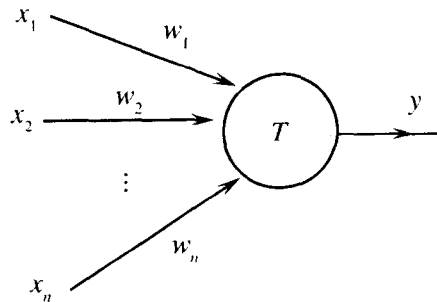


图 2.2 MP 模型

数,即

$$y = \begin{cases} 1, & \sum_{i=1}^n w_i x_i \geq T \\ 0, & \sum_{i=1}^n w_i x_i < T \end{cases} \quad (2.1)$$

式中, x_i 表示神经元的第 i 个输入; w_i 表示神经元的第 i 个输入权值; y 表示神经元的输出; T 表示神经元的阈值; n 为输入个数。

单个 MP 神经元模型可以实现与、或、与非、或非等二值逻辑运算(但不能实现异或运算)。另外,该模型曾因说明了人工神经网络可通过简单的计算产生相当复杂的行为,从而引起极大的轰动,但它是一种静态神经元,即结构固定,权值无法调节,因此缺乏一个关键性的要素,即学习能力。

2.1.3 一般神经元模型

由于 MP 模型过于简单,而且权值不能学习,因此需要更复杂的灵活性更高的神经元模型。图 2.3 所示为一个具有 n 个输入的通用的神经元模型。与 MP 模型一样, $\mathbf{x} = (x_1, x_2, \dots, x_n)^T$ 为神经元输入, $\mathbf{w} = (w_1, w_2, \dots, w_n)^T$ 为可调的输入权值, θ 为偏移信号,用于建模神经元的兴奋阈值。 $u(\cdot)$ 和 $f(\cdot)$ 分别表示神经元的基函数和激活函数(也称为神经元函数、挤压函数或活化函数)。基函数 $u(\cdot)$ 是一个多输入单输出函数 $u = u(\mathbf{x}, \mathbf{w}, \theta)$; 激活函数的一般作用是对基函数输出 u 进行“挤压”: $y = f(u)$, 即通过非线性函数 $f(\cdot)$ 将 u 变换到指定范围内。

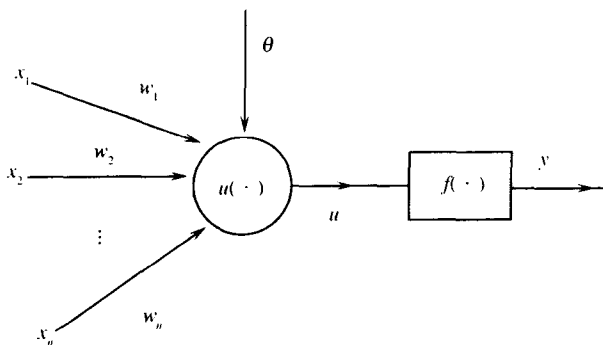


图 2.3 通用神经元模型

下面介绍常用的基函数及激活函数的类型。

1. 基函数类型

1) 线性函数

绝大多数神经网络都采用这种基函数形式,如多层感知器(MLP)、Hopfield 网络等。采用线性函数时,基函数输出 u 为输入和阈值的加权和,即

$$u = \sum_{j=1}^n w_j x_j - \theta = \mathbf{x}^T \mathbf{w} - \theta \quad (2.2)$$

在多维空间中,该基函数形状是一个超平面。