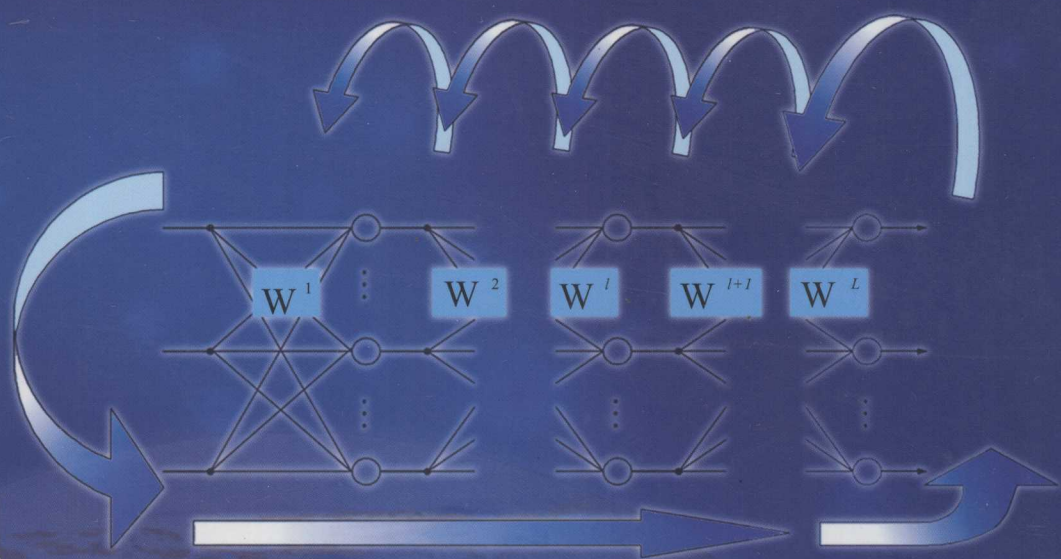


国家自然科学基金项目(40075021)资助

神经网络气象预报 建模理论与应用

金龙 著



气象出版社

国家自然科学基金项目（40075021）资助

神经网络气象预报 建模理论方法与应用

金 龙 著

气象出版社

内 容 简 介

本书详细介绍了作者近 10 年来从事人工神经网络气象预报建模理论与预报应用的主要研究成果。主要内容包括:神经网络主要模型的基本理论和方法,神经网络的各种短期气候预测模型,时间序列的神经网络预报建模,神经网络的混合预报模型,神经网络的数值预报产品释用、能见度预报等应用研究,神经网络预报模型的学习矩阵构造、泛化性能和过拟合等关键问题的解决方法,书中还给出了大量气象预报应用研究的计算实例,具有较强的实用性,便于应用和推广。

本书可供从事气象预报研究,业务预报以及水文、海洋、环境、地震、经济和市场预测研究领域的研究人员及相关院校的本科生、研究生使用。

图书在版编目(CIP)数据

神经网络气象预报建模理论与应用/金龙著.

北京:气象出版社,2005.2

ISBN 7-5029-3928-8

I. 神… II. 金… III. 神经网络-应用-气象预报 IV. P457

中国版本图书馆 CIP 数据核字(2005)第 015476 号

出版者:气象出版社

地 址:北京市海淀区中关村南大街 46 号

网 址: <http://cmp.cma.gov.cn>

邮 编: 100081

E-mail: qxcbs@263.net

电 话: 总编室: 010-68407112 发行部: 010-62175925

责任编辑: 张锐锐 李太宇

终 审: 周诗健

封面设计: 袁丽军 张建永

印刷者: 北京市金瀑印刷有限责任公司

发行者: 气象出版社

开 本: 787×1092 1/16

印 张: 14.0

字 数: 360 千字

版 次: 2004 年 12 月第一版 2004 年 12 月第一次印刷

书 号: 7-5029-3928-8/P·1400

印 数: 1~1500

定 价: 40.00 元

前 言

随着人类对自身大脑功能认识的不断深入,以模拟大脑结构和思维方式为基础的人工神经网络理论方法已成为国际上多学科交叉发展的前沿领域。随着神经网络理论方法研究的不断深入,该理论方法已在众多学科领域取得了大量应用研究成果。大气学科的天气预报、短期气候预测问题大多具有复杂性、非线性、高维性以及观测序列样本大多含有噪声和不精确性等特点,而神经网络方法是一种具有大规模并行处理能力、自适应学习、非线性映射及强容错能力等多种优良性能的新兴交叉学科的新方法。特别是近年来,神经网络方法与小波分析、模糊理论和遗传算法的融合发展,为众多学科的应用研究提供了重要基础,这也为天气预报,特别是短期气候预测等气象预报建模研究和应用提供了新的理论方法。目前,国内外大气科学领域中,也开展了很多有关神经网络方法的气象预报、分析应用研究工作,但是至今未见有关神经网络的气象预报建模理论方法专著出版。

本书是作者近 10 年来从事神经网络气象预报建模理论方法研究与业务预报应用的工作总结。全书共分八章,第一章简单介绍了人工神经网络技术的主要发展背景,重点叙述了国内外大气科学的神经网络应用研究状况;第二、三章分别介绍神经网络的两种主要模型——前馈神经网络和反馈神经网络(Hopfield 网络);第四章介绍神经网络定性、定量嵌套,神经网络集成,最优子集及具有门限变量的多种神经网络短期气候预测模型;第五章以时间序列和多元分析为基础,提出了神经网络的混合预报模型;第六章讨论时间序列的神经网络预报建模方法和应用;第七章介绍神经网络方法在最大可能蒸发计算、农田土壤水分预报、大雾低能见度预报、数字识别和数值预报产品释用预报等方面的应用研究;第八章针对神经网络预报建模的最关键问题,介绍怎样防止神经网络预报模型过拟合现象,提高泛化性能的实用方法。

本书的出版得到了国家自然科学基金项目(40075021)资助。本书的内容包含了作者近 10 年来主持的多项课题的研究成果。在此期间,罗莹高级工程师、陈宁工程师、袁成松工程师,博士生导师缪启龙教授、副教授秦伟良博士,还有作者的研究生金健、李永华、陆虹、农吉夫、吴建生、王业宏、林健玲和博士生林开平等先后参加了相关课题的研究工作。另外,高级工程师罗莹参与了全书的校稿,研究生金健、吴建生、李永华、农吉夫、林健玲、王业宏和林开平分别参与了本书部分相关内容的初稿撰写,图表制作、修改和文字输入、校对工作,作者在此一并致以衷心的感谢。

由于人工神经网络的预报建模理论方法和长、中、短期天气预报技术涉及多学科交叉的众多学科领域,加之作者水平有限,书中一定有不少错误和遗漏,敬请各位专家、读者批评指正。

金 龙

2004 年 6 月 22 日

目 录

前言	
第一章 绪论	(1)
1.1 人工神经网络的研究意义	(1)
1.2 人工神经网络技术的主要发展背景	(2)
1.3 我国神经网络研究发展概况	(3)
1.4 神经网络方法在大气科学中的应用研究	(4)
第二章 前馈神经网络	(9)
2.1 神经元模型和网络结构	(9)
2.2 感知器	(16)
2.3 自适应线性神经网络	(35)
2.4 多层前馈网络与 BP (Back Propagation) 算法	(39)
第三章 反馈神经网络	(48)
3.1 前馈网络与反馈网络	(48)
3.2 Hopfield 网络概况	(49)
3.3 离散 Hopfield 网络 (DHNN)	(53)
3.4 连续 Hopfield 网络 (CHNN)	(66)
第四章 神经网络短期气候预测模型	(75)
4.1 定性定量嵌套的短期气候预测模型	(75)
4.2 两种等级趋势预报模型的对比分析	(85)
4.3 神经网络的集成预报	(89)
4.4 最优子集的神经网络预报建模	(96)
4.5 具有门限变量的神经网络预报应用	(101)
4.6 基于奇异谱分析重建序列的神经网络多步预测模型	(105)
第五章 神经网络混合预报模型	(111)
5.1 神经网络混合预报模型的建模理论分析	(111)
5.2 汛期降水的神经网络混合预报模型应用分析	(112)
5.3 神经网络混合预报模型的性能分析	(116)
5.4 台风频数的混合预报模型应用分析	(120)
5.5 农田旱涝神经网络混合预报模型的预报应用	(124)
5.6 月降水量神经网络混合预报模型的预报研究	(129)
第六章 时间序列神经网络预报建模	(143)
6.1 多步预测的神经网络气候预报模型	(143)

6.2	年平均气温的神经网络多输出预报模型	(149)
6.3	农田土壤湿度的神经网络预报	(154)
6.4	各种最优定阶准则自回归模型的神经网络预报方法研究	(158)
第七章	神经网络方法的其它气象应用研究	(162)
7.1	水面蒸发神经网络计算方法	(162)
7.2	农田土壤湿度的人工神经网络诊断预测系统	(164)
7.3	人工神经网络的数字识别系统模型	(168)
7.4	高速公路能见度的神经网络预报方法应用	(176)
7.5	数值预报产品的神经网络释用预报	(178)
第八章	神经网络预报建模的关键技术	(186)
8.1	神经网络的泛化性能和过拟合问题	(186)
8.2	预报模型泛化性能的改进方法和原理	(188)
8.3	提高网络泛化性能的系统降维实用方法	(191)
附录	神经网络计算程序	(202)
参考文献		(208)

第一章 绪 论

1.1 人工神经网络的研究意义

普通计算机的发展,在计算速度、存贮数据等方面的能力已远远超过了人脑,但是人脑的形象思维能力则是一般电脑所不能及的。从 20 世纪 80 年代开始,以模拟人类大脑结构和思维方式为基础的人工神经网络技术在全世界范围得到了迅速发展。该学科是以模仿人脑神经元在进行形象思维时,脑神经元之间是如何传递、接受和综合信号为出发点的。实际上,人脑是由大量神经细胞(也称为神经元),按不同的结合方式组成的非常复杂的神经元网络。其中每个神经元可看作一个小的处理单元,通过神经网络中的各神经元及其连接强弱,按外部激励信号作出兴奋或抑制的自适应变化,使大脑具有学习、记忆和思维等各种智能。

由于人工神经网络理论方法的研究,是以模仿人脑的结构和思维方式为基础的,所以随着人类对自身大脑奥秘的不断探索和模拟,也不断地推动了人工神经网络学科的发展。并且到目前为止,虽然人工神经网络方法的研究只是对人类大脑结构的低级近似模仿,但是它已经在对外来信息的自适应学习、数据的并行处理以及信息的分布存贮等方面与人脑有相似之处。并且人工神经网络技术已在信号处理、图像、语音识别和记忆、预测及优化等众多学科领域的应用研究中显示了很好的应用前景和良好性能。

人工神经网络方法在处理实际问题时主要包括两个过程,即学习训练过程和记忆联想过程。在做学习训练时,是让网络根据外部输入信息和目标输出,按某种规则(即学习算法)对网络处理单元的连接权值进行调节,当实际输出与目标输出达到一定精度标准时,网络就完成了其预定的学习训练任务,相应的各神经元连接权值确定。而所谓的记忆联想,是对训练好的网络输入一个信号,网络根据确定的连接权值可以正确地给出相应的记忆输出。

根据人工神经网络原理,人们还进一步研制开发了所谓的第六代计算机即神经计算机,由于这种计算机是基于模仿人脑结构和功能而构成的一种新型计算机,它与传统的数字计算机明显不同。神经网络计算机由大量的单个处理单元组成一个庞大的网络系统,通过模仿人脑神经系统的组成结构和活动机理,能够通过学习训练对外部事物进行识别,并且神经计算机与传统的数字计算机处理问题的方法也有较明显的差别。传统的数字计算机在应用于数值计算方面显示了很强的能力,并且它是基于计算方法的顺序处理技术,而神经计算机是采用并行处理技术,通过对研究对象的输入作自适应的学习训练而得到求解问题的结果。同时,神经计算机在处理信息时具有很强的容错性,当研究问题的输入信息部分缺损或信息模糊时仍可以得到比较完整的结果。由于所谓的第六代计算机其原理就是通常所述的人工神经网络,其最重要的特点是它以模仿人类大脑结构和功能来进行信息处理

的, 因此, 这为人们研究和处理大量的各类科学问题提供了新的有效手段和方法, 它的应用研究也将会大力推动科学技术的快速前进。

1.2 人工神经网络技术的主要发展背景

目前, 较为公认的人工神经网络的开拓性研究工作是 20 世纪 40 年代初, 由美国神经心理学家 Warren McCulloch 和数学家 Walter Pitts 合作提出的阈值加权和二值神经元模型即著名的 M-P 模型。该模型是从信息处理的角度, 总结了生物神经元的一些基本生理特征, 采用数学方法建立的一种人工神经元模型。这一模型的提出, 为人们开展人工神经网络的研究提供了重要的基础。人工神经网络早期研究的另一项重要工作是 1949 年心理学家 D.O.Hebb 提出的 Hebb 学习规则, 该学习规则主要是表征了两个神经元处于同一状态时, 神经元轴突的连接强度将得到加强。相反, 当两个神经元处于不同状态时, 神经元轴突的连接强度会减弱。这种神经元连接强度变化规律的提出是人工神经网络学习算法的开创性工作, 对人工神经网络研究发展具有重要意义。50 年代后期, 计算机专家 Frank Rosenblatt 首次提出了一种智能人工神经网络系统即感知器 (Perceptron) 模型。这种感知器虽然比较简单, 但是它已初步具备了学习能力、并行处理和连续计算等一些神经网络的基本功能。并且这种单级感知器也已被用来作为信息处理等实际问题的求解。后来 Berrard Widrew 和 Marcian Haff 又进一步开发研制了自适应单元 (Adaline) 模型。但是随着人工神经网络模型研究和应用的不断深入, M.L.Minsky 和 S.Papert 通过对感知器这样一类简单网络系统的系统研究发现, 感知器的研究虽然对神经元网络的研究有重要的推动作用, 但是它在解决一些问题时还存在一定的局限性。这种比较简单的网络系统对于一些比较简单的线性可分问题有较好的处理能力, 但是无法解决比较复杂的非线性分类问题。并且在 1969 年, 由 MIT 出版社出版发表了他们对感知器研究全面认识的总结性专著——《Perceptron》。当时虽然已经认识到如果要使网络系统能解决非线性分类问题和比较复杂的高阶问题, 神经网络系统的结构必需增加隐层, 构造出多层网络结构系统, 但是由于当时认为要想设计提出一种有效的, 适用于多层神经元系统的学习算法有相当的难度, 由于这些相关研究问题一时难以获得有效进展, 随后一段时期神经网络的应用和理论研究迅速进入了发展缓慢的低潮期。

在 1969 年 M.L.Minsky 和 S.Papert 发表“感知器”专著后一直到 20 世纪 70 年代末, 相对而言世界范围从事神经网络研究的学者人数急剧减少, 尽管如此, 仍有一些科学家仍然坚持不懈地进行神经网络模型和学习算法的相关研究, 提出了许多有意义的理论和方法。其中在人工神经网络系统研究中有重要影响的美国波士顿大学自适应系统的创始人 S.Grossberg 等经过多年系统地研究用数学描述人的心理和认知活动后, 提出了自适应共振理论。芬兰学者 T.Kohonen 提出了自组织特征映射, 而 J.Anderson 从具有神经元突触激活的联想记忆模型出发, 与 Silvetstein 等人在线性联想记忆的基础上进一步提出了盒中脑 (Brain-state-in-a-Box) 模型, 即 BSB 模型等; 而日本学者 K.Fukushima 提出了认知机网络模型理论, 并先后构造了两种认知机模型; 1979 年, 日本东京大学的学者又提出了联想记忆模型等, 这些研究都为后来神经网络研究的发展提供了重要的理论基础。

在 20 世纪 80 年代初,神经网络研究开始走出低谷,迎来了一个新的研究高潮,其中最重要的研究成果是美国加州工学院物理学家 John Hopfield 提出了一种新的神经网络模型,即著名的 Hopfield 网络模型。他在网络的稳定性研究方面引入了李雅普诺夫函数(能量函数)作为判据,使得网络模型具有联想记忆能力,可进行优化问题的求解。其中更为令人瞩目的是 Hopfield 利用该网络模型较好地解决了 TSP (Travelling Saleman Problem,巡回推销商)问题,这为采用神经网络方法解决优化计算提供了新方法,同时这也为神经计算机(Neural Computer)的开发研究提供了重要的基础。

在 1986~1988 年,心理学家 L.L.McClelland 和 D.E.Rumelhart 经过共同努力,研究了并行分布式信息处理的方法,发表了《并行分布处理》(PDP, Parallel Distribution Processing)专著。并进一步通过对神经网络的结构、神经网络节点函数和学习算法的系统研究,提出了著名的多层网络的误差逆传播学习算法即 BP 算法。该算法迄今为止仍然是一种应用最为广泛的神经网络学习算法。该算法的主要设计思想是,将输入信号通过隐层和输出层节点的处理计算得到的网络实际输出进一步与期望输出相比较,并计算实际输出与期望输出的误差,将误差作为修改权值的依据反向传播至输入层,再修正各层的权系数,并且反复这一过程,直到实际输出与期望输出的误差达到预先设定的误差收敛标准,从而获得最终的网路权值。由于 BP 算法解决了多层网络的学习算法问题,因此在许多学科的实际应用问题中获得了成功,如模式识别、信号处理等,并从实践上证明了神经网络确实具备强大的功能,从而大大推动了神经网络的研究热潮。

近十几年来,神经网络的研究仍然是国际上许多科学家所关注的热点研究课题。许多学科的研究人员都在试图尝试应用各种神经网络模型解决本学科的一些难点问题以及一些交叉学科的复杂问题。目前,有关神经网络研究和应用研究的专著和各类专业期刊上发表的相关研究成果还相当多,各种神经网络理论计算软件及神经计算机、神经网络芯片的研究也有相当的研究成果报导。这些均给神经网络的未来发展展示了很好的前景。但是,尽管如此,我们还必须清楚地认识到,迄今为止,人工神经网络研究所取得的成果,还只是比较简单地在模仿人脑生物神经元处理信息的一些低级功能,人类对自身大脑的研究还有很多未知的东西,因此对于人工神经网络研究还有待于更为长期和艰苦的努力。

1.3 我国神经网络研究发展概况

我国是在 1990 年由中国电子学会等八个国家一级学会联合在北京召开了第一届中国神经网络学术会议。并且于 1991 年在南京正式成立了“全国人工神经网络委员会”。

在此之前,北京大学于 1988 年在国内召开了第一次关于神经网络的讨论会,并邀请美国海军实验室的华裔学者斯华龄教授等一些国际知名学者作了专题报告。我国的神经网络研究工作虽然较国际上相比起步较晚,但是随着国内众多相关学科的学者进入这一领域的研究,我国在神经网络的理论研究和应用研究,都有了较为快速的发展。

全国神经网络委员会自 1990 年起每年召开一次全国性神经网络学术大会。并且“全国人工神经网络委员会”也由当初的 8 个国家一级学会发展为 15 个国家一级学会,其中包括中国电子学会、计算机学会、自动化学会、物理学会、生物物理学会、心理学会、通

信学会、人工智能学会、运筹学会、电工学会、数学学会、生物工程学会、电机学会、光学学会和中华医学会。这对我国神经网络的研究起到了重要的推动作用。每届年会几乎都有数百篇研究论文进行学术讨论交流,其中在北京、南京和香港等地都承办了神经网络国际学术交流会议,大大提高和加快了我国有关神经网络各领域的研究水平,也标志着我国神经网络研究工作大规模地走向世界。我国的国家级研究项目:“八六三”高技术计划、国家自然科学基金项目及国家攀登计划项目等都有专项开展有关神经网络理论和应用的研究项目。我国人工神经网络委员会主席,吴佑寿院士在 1999 年汕头大学召开的第 10 届全国神经网络学术交流大会上,报告的我国神经网络研究“十年回顾”,全面系统地总结了我国在这一领域研究取得的重要成果,同时也指出了与国际相比我国在这一领域存在的差距,以及未来研究方向和可以借鉴的国际上的先进经验和规划。

1.4 神经网络方法在大气科学中的应用研究

自 20 世纪 80 年代后期以来,人工神经网络的理论和应用研究都有了较大的发展,其重要原因之一是神经网络方法在图像、信号处理、模式识别、智能控制和计算机技术等众多学科领域取得了大量的应用研究成果。在大气科学中,有关神经网络方法的应用研究相对而言,国内比国外开展的稍晚些。在国外,较早开展气象预报神经网络应用研究的是美国 Neural Ware 公司,该公司于 1987 年开发设计了一种人工神经网络晴雨预报系统。

该预报系统使用的是一个有输入层、中间层和输出层的三层 Madaline 神经网络模型。预报系统的输入是采用了地面气压和风向资料,而模型输出即为今天、明天和后天的逐日晴雨预报。该预报系统利用 4 月、5 月的逐日资料作为训练样本集,对模型作学习训练。以 6 月份资料作为检测集资料,对预报系统作实际预报检验。结果发现预报系统对 6 月份逐日晴雨预报的准确率还略高于当地预报员的业务预报水平,并认为如果进一步将中、高空的气象要素加入网络则有可能进一步提高预报系统的准确性。后来 Neural Ware 公司又采用 BP 网络进一步开发了另一种神经网络系统 (Neural Network Profession II/plus),并利用该系统对卫星云图的冷锋,雹暴云团及 MCC 四类云系进行云的分类、特征识别。结果发现,该系统对四类云的训练集样本的识别能力可达到 100%,而对于检测集样本的识别能力也在 90%至 100%之间。在强对流天气的预报研究工作方面, Dohald w.McCANN 采用最常用的前馈网络模型进行了强雷暴事件的预报研究。研究发现,由于雷暴产生的机制非常复杂,并具有很强的非线性变化特征,因此,利用一些显著相关因子建立的 BP 网络预报模型,经检验具有很好的预报效果。在热带气旋强度预报方面,Jorg Jin Baik 和 Hong Sub Hwang 采用多元回归方法和神经网络方法进行了热带气旋强度预报的比较,结果认为神经网络方法有很大的预报潜力,预报效果优于回归方法。David Siverman 和 John A.Draccp 合作利用 BP 网络模型对美国加利福尼亚地区降水长期预报进行了研究,他们将加州分为 7 个区域分别建立预报模型,结果发现,多数预报模式对降水场的预报结果与实况场分布配合相当好。Andrew R.Dean 和 Brian H.Fiedler 则采用线性回归方法和非线性神经网络方法进行了机场云(雾)消的预报试验。其预报结果表明,非线性神经网络方法的预报技巧评分比气候预报高出 0.25,而线性回归预报的技巧评分比气候预报高出 0.20。并且

为了进一步分析,他们还根据机场早晨的实测温度,采用线性回归方法和非线性神经网络方法分别作机场下午的温度预报,结果显示,非线性神经网络方法的预报技巧评分为 0.446 而线性回归方法为 0.290。Lei Shi 根据 NOAA-15 卫星探测资料,采用前馈网络模型(Backpropagation neural networks)进行了大气温度廓线和对流层大气变量(包括对流层顶温度、高度和气压)的反演。结果表明,从地面到 10 hPa 高度上的温度廓线反演结果非常理想,并且他还将这些反演的结果与线性回归方法所得结果进行了比较,发现神经网络反演的结果在大气对流层的各个高度上均好于线性回归方法。J.Shao 利用一个三层的神经网络模型,进行了冬季公路表面冰冻温度的临近预报(提前 3~6 h 预报),并且将这样的神经网络预报模型与数值预报模式的预报作对比分析,对比分析又分为两种环境条件,一种是正常的普通环境条件,另一种是环境条件比较复杂和恶劣的条件(主要是挪威、英格兰、瑞士和奥地利等冬季温度低、变化大的地区)。对比试验结果表明,两种环境条件下,神经网络预报模型对公路的冰冻预报均方误差均得到改进,其中在一般的环境条件下,神经网络预报模型预报误差减小的程度比恶劣环境条件的预报误差减小程度要小。Tony Hall 利用神经网络方法对德克萨斯州的达拉斯的一个区域进行了两年的降水概率预报和降水定量预报,并与该区域 36 个站的雨量观测网资料进行了对比检验分析。所得结果是,预报与实测降水量的相关系数达到 0.95,而 1~25 mm 的降水预报平均 Ts 评分达到 0.63 以上,其中最高值超过 0.86,显示了有很好的发展前景。Liu Hongping 等提出了一种采用自适应神经网络开展降水预报的方法,这一方案以雷达观测资料作为自适应网络预报模型的输入,对网络模型作学习训练,然后用新的观测资料作降水预测。实验资料的计算结果表明,这种自适应网络预报模型对降水有相当高的预报精度,这为有效利用雷达探测资料进行降水预报提供了新的思路。

Garen Marzban 和 Arthur Witt 分别建立了两种不同的冰雹神经网络预报模型,一种用于预报强雹的大小,另一种模型则用于进行冰雹的分类预报。实际预报结果表明,用于预报强雹大小的神经网络预报模型优于相应的传统预报方法,而用于分类预报的神经网络预报模型,在进行分为三种不同大小的冰雹分类预报时,对最大和最小的两类有较好的分辨能力,而对中等大小的强雹则无区分能力。E.Moreau 和 C.Mallet 等利用神经网络方法进行了反演大气中液态水含量的研究。Tim Bellerby 利用雷达降水资料和气象卫星处理数据,进行神经网络方法的降水预报研究后指出,综合应用雷达降水资料和气象卫星处理数据可以得到时空分辨率较高的降水预报。Roret J.Abrahart 和 Linda See,针对两个蓄水区的河水流量预报问题,用神经网络方法和滑动平均自回归方法进行了预报对比分析,结果表明神经网络预报方法的预报结果更令人鼓舞。G.Mihalakakou 和 M.Santamouris 等利用神经网络方法对太阳辐射时间序列进行短期预报试验,并采用了多步滞后的预报技术,将神经网络方法与自回归预报模型预报结果进行比较,结果发现,神经网络预报模型的预报效果要远好于自回归模型。Mohana Rangan 和 S.M.Sabbcb 采用人工神经网络方法进行了蒸发模拟计算,结果表明,利用气候观测数据和神经网络方法可以成功地模拟蒸发作用。另外,Graham A.Balls 和 Gina E.Sanders 等还采用神经网络方法进行了植物和环境相互作用的模拟研究。

William W.Hsieh 和 Tang Benyang 系统地总结分析了经验统计方法在大气学科和海洋

学科中应用发展的 4 个不同阶段后认为,神经网络在气象学和海洋学的应用研究虽然是一种新的方法,但是该方法在这两个领域的应用初期存在 3 个方面的主要问题:第一,在气象学和海洋学中一些短期观测数据存在非线性不稳定;第二,如何处理空间上大范围的场资料;第三,由神经网络方法所得结果难以得到直观的解释。然而,最新的一些研究结果表明对于上述问题,已有一些相应的解决办法:第一,用总体平均值可以有效地控制非线性不稳定;第二,用 PCA 方法可以对大范围的空间场资料进行压缩;第三,有关隐层数的问题可以用相空间解释,而谱分析也有助于理解非线性神经网络的关系。并认为随着人工神经网络方法的不断发展,该方法可以增强传统资料分析和预报方法的能力,成为更加通用和强大的工具。例如,神经网络方法已被成功地应用于厄尔尼诺预报。同时神经网络方法还可用于资料同化计算,这样使得神经网络方法与动力模式相联系起来,可以进行一些新的混合神经网络动力模式研究。

在国内的大气科学业务科研工作中也广泛开展了有关人工神经网络的应用研究。1994 年张承福曾利用神经网络方法开展了不同区域范围的暴雨预报研究,认为神经网络方法确实可以通过对网络的学习训练,从原始数据中提取足够的分类信息,达到较好的预报准确率。蔡煜东根据月平均最低气温日数等 4 个单站气象要素作为模型的输入,进行了 BP 神经网络的本站汛期降水期预报试验。

1995 年,姜天戟和袁曾任利用美国 Neural Ware 公司在开发研制 Madaline 神经网络天气预报系统时采用的美国匹兹堡机场相同的气象观测数据,以 BP 算法作为基础,提出了三种不同激活函数组成的神经网络预报模型,同样进行了今天、明天和后天的逐日晴雨预报试验。结果表明,标准的 BP 网络模型,带动量项的 BP 模型及最优组合 BP 模型的预报准确率和网络学习训练的收敛速率均比 Neural Ware 公司采用 Madaline 神经网络模型建立的天气预报系统要优越。盛永宽进行了神经网络与气候模式相结合的试验研究,他采用当时世界上主要的十五个模式之一的改进的全球大气环流与海洋混合模式进行了月、季、年尺度的降水短期气候预测。并在气候模式系统误差调整时,尝试结合应用神经网络方法。他将气候模式输出的环流场、模式输出的海温场、模式输入的降水场以及长序列的实测环流场与实测降水量场建了多个 BP 网络模型。最后由组合网络预报模型作出逐月降水预测。试验研究是以我国 160 站,1990~1994 年 5~8 月,共 20 个月的降水量进行预测试验的,由于网络的训练样本不多,且神经网络的预测能力还与网络结构、算法等多种参数有关,因此从实际试验结果看到,预测准确率不高,但是雨带的主要特征已能够预测出了,并认为,随着样本序列的增加和新资料的加入,映射关系的扩展,模式性能有望得到提高。

1996 年,金龙等进行了农田旱涝的神经网络预报方法研究以及采用人工神经网络方法进行水面蒸发量的计算试验。周曾奎和韩桂荣等通过建立的人工神经网络台风预报系统,可作出未来台风移动方向的预报。郭光和严绍瑾利用 BP 网络模型对我国东部 6 个城市的汛期降水进行了模拟预报试验。金龙采用人工神经网络方法,对 5 种最优定价准则确定的月平均气温预报模型进行了集成预报研究。李祚泳和邓新民应用人工神经网络理论建立了城市环境质量的 B-P 网络评价模型。

1997 年,金龙等进行了神经网络汛期降水预报性能与逐步回归预报准确率的对比分析。秦伟良和金龙应用小波神经网络方法进行了非线性时间序列的预报建模研究。曹杰和

谢应齐根据联想记忆神经网络的基本原理,提出了一种基于混沌理论的联想记忆神经网络模型,并用该预报模型对单站的月平均气温时间序列进行了预报试验。

1998 年,白慧卿和方宗义等采用人工神经网络方法进行了 GMS 云图中的四类云系识别,并将神经网络方法对云图类型的识别能力与传统的统计识别方法进行了比较,认为神经网络方法更适合于云系的特征识别。金龙和栾建林等提出了利用前馈网络的多输出模型,对多个逐步回归定量预报方程进行了定性的集成预报方法研究,试验结果表明,神经网络的定性集成预报效果要好于各个子预报方程。李祚泳和万彩萍为了选择与预报量有密切关系的因子和有代表性的样本,尝试了用 BP 神经网络方法对因子和样本进行优选的试验,并将该方法与主成分集合选择变量方法进行了比较,结果表明,两种方法的优选结果基本一致。

1999 年,金龙和陈宁等针对统计天气预报中如何确定集成预报中各子方法的权系数问题,采用神经网络方法进行了预报建模研究。梅钰利用一些常规气象观测数据作为神经网络预报模型的输入,开展了辐射雾的预报试验。王成刚和吴宝骏等进行了鲁西南地区西南涡降水量级的神经网络预报应用试验,发现 BP 网络的学习性能较好,但网络参数稍有变化会引起推广能力振荡。陈宁和金龙等提出了采用最优子集方法构造神经网络统计预报模型输入学习矩阵的方法,并用该方法进行了汛期降水预报的预报效果分析。

2000 年,金龙和秦伟良等采用小波分析与神经网络相结合的方法,并以高斯导函数作为母小波,进行了一种新的非线性长期预报模型研究。张初和蒋国荣,根据赤道地区的海温观测资料与西太平洋副热带高压面积指数之间的关系,建立一个 BP 神经网络预报模型,预报未来 3 个月的副高面积指数变化趋势和走向。熊秋芬和王丽等用数值预报方法、天气学方法和以数值预报产品为基础的神经网络方法进行了 1999 年汛期雨量的预报比较,结果发现区域数值预报模式和神经网络方法有较好的预报效果。金龙和苗春生等利用神经网络灵活可变的网络拓扑结构,分别构造了定性和定量相嵌套的神经网络预报模型。

2001 年,胡江林和张礼平等以前期 500 hPa 月平均高度场、海平面气压和太平洋海温场的扩展自然正交展开的前几个主要模态的时间系数,作为神经网络模型的输入,建立了汛期降水的神经网络短期气候预报模型。通过对 41 年历史样本的交叉检验,预报技巧评分平均为 0.246。陈云浩和史培军等在对城市降雨系统时序规律分析的基础上,提出了基于人工神经网络的降水时序分析与集成预报模型。并以上海市旬平均雨量为例,分析比较了神经网络集成预报与移动平均法、指数滑动预报法三种预报方法的差别,认为神经网络集成预报方法比其它两种预报方法的预报精度高。金龙和秦伟良等提出了利用均生函数方法构造神经网络的学习矩阵输入,建立了一种新的短期气候预报模型。张初在西太平洋副高指数的预报中,采用前传神经网络模型和回归网络模型进行了建模研究,并进一步与传统的统计回归预报方法进行了比较,结果认为,回归网络预报模型较其它两种方法对预报量在总体上的描述和预测要略为准确。梁益同和胡江林在应用神经网络方法进行 NOAA 卫星遥感图像的洪涝灾害水体识别时发现,神经网络方法比阈值法具有更高的精度和效率。

2002 年,曹晓钟和王强利用一种改进的 BP 网络模型提出了基于相空间、时间相关和模式相关的三种资料优化处理方案,认为神经网络方法比时间序列方法应用更广泛,是

时间序列方法的扩充。熊秋芬和胡江林等应用神经网络方法探讨了利用静止卫星多通道资料估算地面降水量的方法,通过个例计算的对比分析表明,神经网络方法提供的客观定量降水估算平均相关系数为 0.57,该方法比传统的阈值法或非线性方法有更大的优越性,能更好地分析和揭示卫星资料与降水之间复杂的非线性关系,对降水的估算结果可直接用于洪水业务预报。金龙和罗莹等利用人工神经网络方法结合均生函数和多元分析方法,提出了一种同时考虑预报量自身显著变化周期和环流因子对预报量未来变化影响的混合预报模型。尚松浩和毛晓敏等通过对农田土壤水分与气象、土壤特性和作物长势分析,建立了冬小麦田间墒情的 BP 网络预报模型,并用模型作一步预报,效果较好。王伶和胡筱敏等根据气象要素与大气污染物之间的关系,建立了空气污染的神经网络预报模型。

2003 年,金龙和林熙等采用模糊数学与神经网络相结合的模块化模糊神经网络方法,开展了数值预报产品的模糊神经网络预报建模研究。在逐日的气温预报对比分析中发现,模块化模糊神经网络预报模型的预报能力要明显高于常规的完全预报(PP)方法。王业宏和金龙在短期气候预测研究中,首先通过对初选的预报因子作自然正交展开,并进一步选取与预报量相关程度高的主成分作为神经网络预报模型的输入,建立了一种新的神经网络短期气候预测模型。罗莹和金龙等根据预报量与预报因子之间的不同相关关系,提出了利用门限值和门限变量建立神经网络的旱涝灾害预报方法。金龙和况雪源等针对神经网络预报建模中,预报模型容易出现“过拟合”和学习能力好而泛化性能差的关键问题,采用降维处理技术,提出了如何确定网络结构,防止神经网络预报模型在对学习样本作学习训练时出现“过拟合”现象和提高神经网络的泛化性能的方法。吴建生和金龙用遗传算法优化神经网络的连接权和网络结构,并在遗传进化过程中,采取保留最佳个体的方法进行短期气候预测建模研究。该方法克服了由于神经网络初始权值的随机性和确定网络结构过程中可能造成网络振荡,以及网络容易陷入局部解的问题。

第二章 前馈神经网络

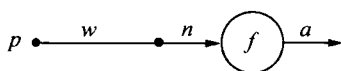
2.1 神经元模型和网络结构

2.1.1 神经元模型

一个神经元由输入、运算体和输出构成。其中，输入部分为神经元的外部输入，一般是一个向量；运算体部分对输入部分进行处理，包括求加权和、将加权和输入传递函数进行处理等；神经元的输出是标量，它取决于神经元的输入和传递函数。

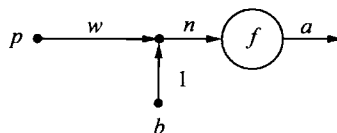
图 2.1 (a) 是一个含一个标量输入、无阈值的最简单的神经元。标量输入 p 通过一个节点和标量权 w 相乘得标量乘积 $n = wp$ ； n 又作为传递函数 f 的唯一自变量，传递函数 f 产生了神经元的唯一标量输出 $a = f(wp)$ 。

图 2.1 (b) 比 (a) 多了一个阈值 b 。一般情况下，人们把这个阈值 b 看作是与固定输入 1 相对应的权。在节点处，阈值和加权输入求和得 $n = wp + b$ 。因此，神经元输出为 $a = f(wp + b)$ 。



$$a = f(wp)$$

(a) 无阈值单输入神经元



$$a = f(wp + b)$$

(b) 含阈值单输入神经元

图 2.1 单输入神经元

可见，在有阈值的神经元中，当 $b = 0$ 时表现为无阈值神经元。

传递函数 f 通常是阶跃 (Hard-Limit) 函数或 S 形 (Sigmoid) 函数等。

值得注意的是神经元的权 w 和阈值 b 都是可以调整的量，而神经网络的基本思想就是调整这些参数，使网络能够产生各种不同的效果。这样，我们就可以训练网络，让它通过调整权和阈值来完成人们需要解决的问题。事实上，对这些参数的调整，我们可以通过算法让网络自动完成。

在神经元中起重要作用的传递函数 (Transfer Function) 在不同的文献中也常被称为激活函数、激励函数、作用函数、转移函数、传输函数或限幅函数等，其作用是对输入进行函数映射，从而产生输出。传递函数是神经元功能的核心，由大量神经元组成的神经网络的功效除了和网络结构相关外，在很大程度上还取决于所选取的传递函数。

最常用的传递函数有阶跃 (Hard-Limit) 函数、纯恒等线性 (Pure-Linear) 函数和对数 S 形 (Logarithmic-Sigmoid) 函数三种 (图 2.2)。

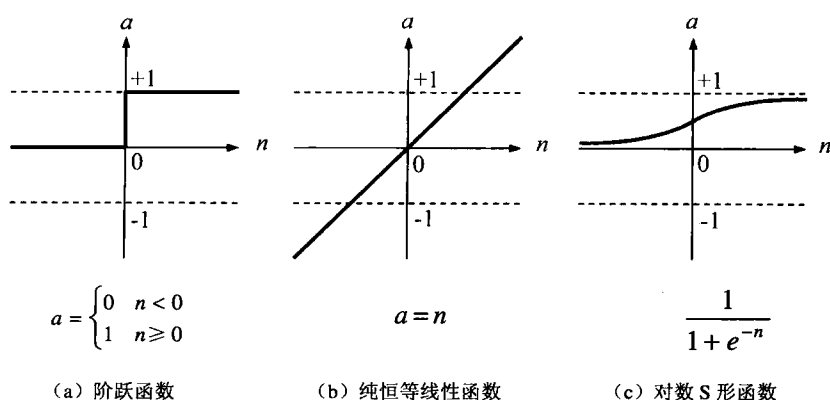


图 2.2 三种常用的传递函数

阶跃函数将输出限制为 0 或 1，应用于感知器网络的决策分类。其函数关系为：

$$a = \begin{cases} 1 & n \geq 0 \\ 0 & n < 0 \end{cases} \quad (2.1)$$

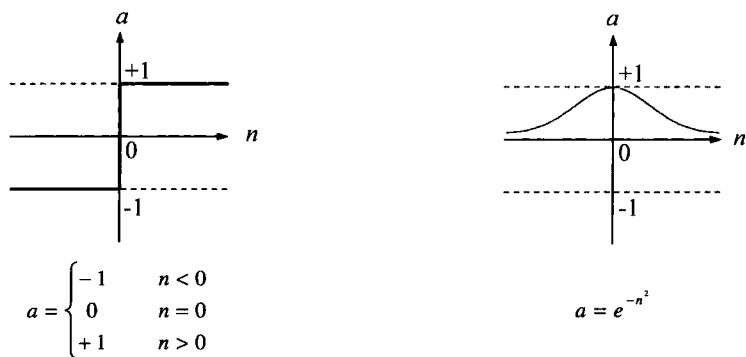
纯恒等线性函数被应用于线性网络，可以用它来构造自适应滤波器。其函数关系为：

$$a = n \quad (2.2)$$

对数 S 形函数能将任意输入压缩限制在 0 和 1 之间，并具有非线性放大增益，因其可微性而普遍应用于 BP 网络。其函数关系为：

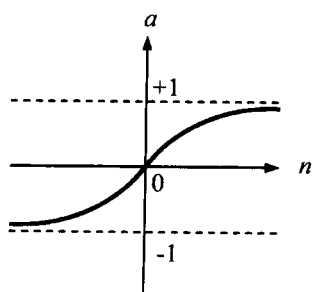
$$a = \frac{1}{1 + e^{-n}} \quad (2.3)$$

另外，还有一些常用的传递函数（如图 2.3）。



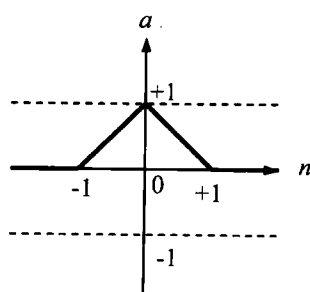
(a) 对称阶跃 (Symmetric Hard-Limit) 函数

(b) 径向基 (Radial Basis) 函数



$$a = \frac{e^n - e^{-n}}{e^n + e^{-n}}$$

(c) 双曲正切 S 形 (Hyperbolic Tangent Sigmoid) 函数



$$a = \begin{cases} 1 - |n| & -1 \leq n \leq 1 \\ 0 & n < -1 \text{ 或 } n > 1 \end{cases}$$

(d) 三角基 (Triangular Basis) 函数

图 2.3 几种常用的传递函数

神经网络模型是线性还是非线性，是由其神经元的传递函数的线性或非线性决定的。

由前面第一章的介绍知道, 人工神经网络的开拓性研究工作是美国心理学家 McCulloch 和数学家 Pitts 于 1943 年首次提出的神经元模型, 简称 M-P 模型。下面我们首先介绍这种 M-P 模型 (见图 2.4)。

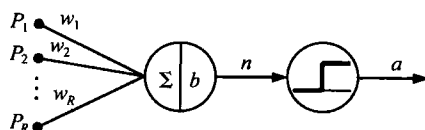


图 2.4 M-P 模型

在 M-P 模型中, 模型的输入是一个 $R \times 1$ 维实数向量 $\mathbf{p} = (p_1, p_2, \dots, p_R)^T$; 其连接权也是一个 $R \times 1$ 维向量 $\mathbf{w} = (w_1, w_2, \dots, w_R)^T$; 具有阈值 b ; 传递函数采用阶跃函数, 因此神经元输出是一个二值量 0 或 1, 分别代表神经元的抑制和兴奋。模型的输入、输出关系可以概括为:

$$n = \mathbf{w}^T \mathbf{p} + b \quad (2.4)$$

$$a = \begin{cases} 0 & n < 0 \\ 1 & n \geq 0 \end{cases} \quad (2.5)$$

早期的 M-P 模型中的权 $w_i = 1$, 它可以用来完成一些逻辑操作。图 2.5 (a)、(b) 分别表示了用这个早期模型表示的逻辑与、或关系 (其中, p 的取值为 0 或 1)。

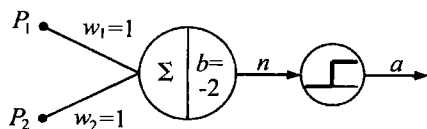
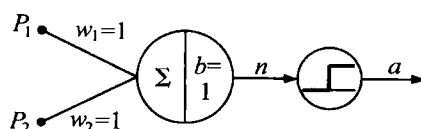
(a) 逻辑与 ($a = p_1 \wedge p_2$)(b) 逻辑或 ($a = p_1 \vee p_2$)

图 2.5 M-P 模型的逻辑关系表达