

神经网络

信号处理

罗发龙 李衍达 著

电子工业出版社

73.821
951

神经网络信号处理

罗发龙 李衍达 著

電子工業出版社

9310346

(京)新登字055号

内 容 提 要

神经网络信号处理是在数字信号处理、神经网络等基础上发展起来的一门新兴交叉学科,它在自动控制、机器人、地震勘探、通信、雷达、声纳等领域中有十分广泛的应用前景。本书系统地阐述了神经网络信号处理的基本理论、方法、模型与应用及有关的最新研究成果。主要内容有:神经网络基本模型与理论,基于神经网络的检测、滤波与参数估计,神经网络信号重构与系统辨识,神经网络自适应信号处理与阵列信号处理,以及神经网络在各种信息系统中的应用等。

本书可供从事信号与信息处理、电子工程、自动控制、计算机、电路与系统、模式识别、雷达与通信工程专业的科技人员和教师阅读、参考,也非常适合作为上述专业的研究生和二年级大学生的教材。

神经网络信号处理

罗发龙 李衍达 著

责任编辑:周 琰

*

电子工业出版社出版(北京市万寿路)
电子工业出版社发行 各地新华书店经销
北京市顺义县李史山印刷厂印刷

*

开本: 787×1092 毫米 1/16 印张: 12 字数: 295 千字
1993年10月第1版 1993年10月第1次印刷
印数: 1—4,000 册 定价: 9.80 元
ISBN7-5053-2010-6/TP·495

03501141

前言

神经网络的研究已在国内外广泛兴起,成为政府、科研机构、生产和开发厂家十分重视的课题,各国都投入了大量的人力、物力来从事有关神经网络应用、理论、模型和实现等方面的研究,其中,神经网络的应用,特别是在信号处理(包括模式识别和智能控制)领域中的应用,极为引人注目。美国国防部的DARPA (Defense Advanced Research Projects Agency) 组织众多科学家和军事系统专家,经过多方严格而又周密的论证,决定把基于神经网络的自动目标识别、声纳信号处理、语音识别、地震勘探信号处理作为研究的最重点,并认为它们是最有前景、且能取得重大突破的应用领域。我国实施的有关自然科学基金重大项目也是将基于神经网络的智能信号处理作为重点研究内容的。日本、法国、加拿大等国的重大科学计划,也都是把神经网络在信号处理中的应用研究放在最重要的位置。了解一下信号处理界所陷入的困境和神经网络特有的信息处理机理,就容易理解人们为什么如此热衷于基于神经网络的信号处理研究。在信号处理中,无论是信号的获取、传输,还是识别、变换、存储、分类、建模与显示,所采用的技术、手段、理论、方法和系统都是与数字电子计算机的原理、结构和发展紧密相关的,由于这种计算机固有的程序式数字化串行处理方式以及局域式地址存储特征所受到的限制,使得在信号处理许多领域内的研究陷入了困境。例如,在信号处理的重要领域信号的检测、估计与滤波中,所谓的最优处理与其需要的运算量之间的矛盾在快速变化的环境中几乎达到了不可调和的程度,也就是说,要达到最优处理性能,则完成运算量所需要的计算机数目常常大得令人不能接受。如果要针对线性处理的许多不足而采用非线性处理,则所需的系统更是复杂庞大,这不得不使人们以牺牲处理性能为代价来换取算法运算量的减少和成本的降低。在模式识别方面,以特征提取为基础的方法也遇到了极大的困难,如何表示和提取特征、需要多少特征,都有很大的盲目性和低效率。识别过程必须经历从特征提取、数据压缩到判决几个阶段,所需的运算使得系统难以满足实时性要求。在自动控制领域,对被控系统进行辨识并修正参数的方法只能较好地应用于线性系统或一类可线性化的简单非线性系统,很难推广到复杂的非线性系统,人们只好在此基础上做大量的近似或简化,这就难免失去应有的准确性。

为此,人们期盼着一种新的理论和技术的诞生。

神经网络研究的兴起正是在这样的大气候下出现的。神经网络是由大量处理单元广泛互连而成的网络,它是在现代神经生物学和认知科学对人类信息处理研究成果的基础上提出来的。在信号、信息处理机制上,它与传统的数字计算机有着根本的不同,它具有大规模并行模拟处理、连续时间动力学和网络全局作用等特点,信息的存储体现在神经元之间连接的分布上,存储区和操作区合二为一。神经网络具有很强的自适应和学习能力、鲁棒性和容错能力,从而可以代替复杂耗时的传统算法,使信号处理过程更接近于人类思维活动。利用神经网络的高度并行运算能力,可以实时实现难以用数字计算和技术实现的最优信号处理算法。利用神经网络分布式信息存储和并行处理的特点,可以避开模式识别方法中建模与特征提取的过程,从而消除由于模型不符和特征选择不当带来的影响,并实现实时识别,以提高识别

系统的性能。充分地将这些神经网络特性应用于控制领域,可使控制系统的智能化向前迈进一大步。神经网络不仅是信号处理的工具,而且还是一种新的方法论。利用它可以研究出许多与传统信号处理所根本不同的新理论、新方法。

面对这样的形势,对于信号处理界的学者、研究人员,特别是对那些一直从事传统信号处理的人来说,非常希望能够通过某种系统的培训和学习,搞清楚:(1)什么是神经网络,(2)神经网络有什么特点,(3)神经网络能应用在信号处理哪些方面,(4)基于神经网络的信号处理的优势在哪里,(5)神经网络信号处理已有一些什么成就,(6)如何有效地进行更进一步的研究等一系列问题。目前,国内外已出版了较多有关神经网络的著作,这些著作都是以神经网络的基本模型和基本理论为主要内容的,很少针对信号处理的具体问题进行叙述。有关神经网络的论文也数以万计,但它们比较零散,有些是针对某个非常具体的问题进行讨论的。因此,很需要有一本系统阐述神经网络信号处理的理论和方法、适合于从事信号处理的研究人员的参考书。本书正是旨在填补这一空白。我们希望通过以下九章内容的讨论,能够对前面所述的六个问题有一个比较全面的回答,也能使读者对基于神经网络的信号处理这种新理论、新方法有较为系统的理解和认识。愿读者和作者一起努力,将我国的神经网络信号处理研究推向前进。

本书的具体安排如下:

第一章介绍几种适于信号处理的神经网络模型及有关网络稳定性、学习算法、动态性能等几方面的基本理论。

第二章结合最小二乘、全局最小二乘、递归最小二乘等自适应算法,着重讨论基于神经网络的滤波理论与方法。

第三章叙述神经网络用于一维谱估计、多维谱估计、高阶谱估计、多通道谱估计的模型和方法。

第四章是有关空域参数估计的内容,主要包括基于神经网络的最大似然法、传播算子法、MUSIC法方向估计。

第五章研究各种噪声背景下信号的神经网络检测。

第六章是神经网络自适应信号处理,主要涉及各种自适应主元学习算法。

第七章讨论基于神经网络的信号重构原理和算法。

第八章介绍和比较几种基于神经网络的系统辨识算法,将以非线性系统的辨识为主要内容。

第九章介绍神经网络与模糊系统的结合——模糊神经网络及其在信号处理中的应用。

本书得到了攀登计划、国家自然科学基金重大项目和国家博士后科学基金的资助。西安电子科技大学的保铮教授为本书的出版予以了极大的支持,在此表示衷心的感谢。

作 者

于清华园

一九九三年三月三十日

目 录

前 言

第一章 神经网络基本模型和基本理论	(1)
§1.1 Hopfield 神经网络	(1)
1.1.1 离散型 Hopfield 神经网络	(1)
1.1.2 连续状态 Hopfield 神经网络	(2)
§1.2 多层感知器及BP学习算法	(6)
§1.3 自组织神经网络	(8)
§1.4 高阶神经网络	(9)
§1.5 概率神经网络(PNN)	(10)
本章小结	(11)
第二章 神经网络滤波系统	(13)
§2.1 LS 准则下的神经网络自适应滤波	(13)
§2.2 实现RLS算法的神经网络	(18)
§2.3 带线性约束的LS算法的神经网络求解	(21)
§2.4 TLS 算法的神经网络实时实现	(22)
2.4.1 TLS算法简介	(22)
2.4.2 TLS算法的神经网络实时实现	(23)
§2.5 神经网络用于一类非线性自适应滤波	(28)
§2.6 神经网络非线性滤波基础	(29)
§2.7 神经网络时间序列预测	(30)
§2.8 基于神经网络的噪声相消	(31)
§2.9 非理想因素对神经网络性能的影响	(32)
2.9.1 连接强度失配对TH网络性能的影响	(32)
2.9.2 神经元的延迟对TH网络性能的影响	(33)
本章小结	(35)
第三章 基于神经网络的谱估计	(36)
§3.1 最大熵一维谱估计的神经网络实现	(36)
§3.2 正弦信号参量的神经网络实时估计	(39)
3.2.1 Pisarenko法简介	(39)
3.2.2 一种神经网络模型	(40)
3.2.3 神经网络用于Pisarenko谱估计计算	(43)
3.2.4 单个正弦波参量的神经网络估计	(47)
§3.3 最大熵多维谱估计的神经网络方法	(50)
§3.4 基于神经网络的高阶谱估计	(53)
§3.5 基于神经网络的多通道谱估计	(56)
3.5.1 多通道谱估计	(56)
3.5.2 一种二维神经网络模型	(58)
3.5.3 二维神经网络实时实现多通道谱估计算法	(60)

3.5.4 二维神经网络用于矩阵求逆	(63)
本章小结	(67)
第四章 神经网络方向估计	(68)
§4.1 神经网络用于ML法和传播算子算法	(68)
4.1.1 问题描述	(68)
4.1.2 ML法和交替投影ML法	(69)
4.1.3 Marcos“传播算子”算法	(70)
4.1.4 几种方向估计算法的神经网络实现	(72)
4.1.5 模拟举例	(75)
§4.2 基于神经网络的ML空间参数估计	(77)
4.2.1 实数ML方向估计	(77)
4.2.2 神经网络结构	(78)
4.2.3 神经网络实现ML方向估计	(80)
4.2.4 模拟举例	(82)
§4.3 MUSIC法的神经网络实现	(82)
4.3.1 MUSIC法方向估计	(82)
4.3.2 特征分解神经网络结构	(83)
4.3.3 神经网络求解信号子空间	(84)
§4.4 基于空间分割的神经网络方向估计方法	(88)
4.4.1 窄带信号源	(88)
4.4.2 宽带信号源	(91)
4.4.3 模拟举例	(93)
本章小结	(93)
第五章 基于神经网络的信号检测	(94)
§5.1 神经网络二元检测	(94)
5.1.1 似然比最佳检测原理	(94)
5.1.2 神经网络检测器	(95)
5.1.3 非高斯背景下信号的神经网络检测	(96)
§5.2 强噪声背景下信号的神经网络检测	(97)
§5.3 瞬时信号的神经网络检测	(100)
§5.4 运动目标信号的神经网络检测	(101)
本章小结	(102)
第六章 神经网络自适应信号处理	(103)
§6.1 最大主元的自适应提取	(103)
§6.2 多主元的自适应提取	(108)
§6.3 Sanger主元学习算法	(112)
§6.4 一类局部多主元分析算法	(113)
§6.5 APEX神经网络学习算法	(116)
§6.6 MCA算法	(119)
§6.7 应用举例	(122)
本章小结	(124)
第七章 基于神经网络的信号重构	(125)

§7.1 神经网络实现最大熵信号重构算法	(125)
7.1.1 最大熵信号重构算法	(125)
7.1.2 神经网络求解最大熵信号重构问题	(126)
§7.2 病态反卷积问题的神经网络求解	(131)
§7.3 二进制信号的神经网络重构	(133)
本章小结	(138)
第八章 神经网络与系统辨识	(139)
§8.1 系统辨识的基础	(139)
§8.2 多层前馈神经网络用于系统辨识	(140)
§8.3 循环神经网络用于系统辨识	(146)
§8.4 几种神经网络辨识算法的比较	(148)
本章小结	(153)
第九章 模糊神经网络信号处理	(154)
§9.1 模糊集理论简介	(154)
9.1.1 隶属概念	(154)
9.1.2 模糊子集的简单运算	(154)
9.1.3 模糊向量及其笛卡尔乘积	(155)
9.1.4 模糊矩阵与模糊关系	(155)
9.1.5 常见的模糊条件语句及其对应的模糊关系 R	(157)
9.1.6 精确量与模糊量的相互转换	(157)
9.1.7 模糊度量	(158)
9.1.8 模糊熵	(159)
§9.2 结构性知识的神经网络与模糊集表示	(159)
§9.3 自适应模糊神经网络	(160)
9.3.1 网络的构成	(160)
9.3.2 自适应算法	(161)
§9.4 模糊神经网络分类器	(162)
9.4.1 模糊神经网络的结构与权值的修改	(163)
9.4.2 输入的模糊化	(164)
§9.5 神经网络与模糊推理协作系统	(166)
9.5.1 构造系统的基本原理	(167)
9.5.2 具体实现方法	(169)
9.5.3 应用于目标系统	(171)
9.5.4 学习后的神经网络权值变化的解释	(171)
9.5.5 结论	(172)
本章小结	(173)
参考文献	(174)

第一章 神经网络基本模型和基本理论

现有的神经网络模型已达近百种。随着应用研究的不断深入,新的模型也在不断推出。在信号处理领域,应用最多的是以下四种基本模型和它们的改进型,即 Hopfield 神经网络、多层感知器、自组织神经网络和概率神经网络。神经网络的基本理论主要涉及稳定性、学习算法、动态性能等几个方面,理论与模型是紧密相关的。本章我们着重介绍以上四种基本模型和有关的理论。

§1.1 Hopfield神经网络

Hopfield 神经网络通常分为离散型和连续型两种。

1.1.1. 离散型 Hopfield 神经网络^{[7][9]}

离散型 Hopfield 神经网络如图 1.1 所示。网络由 N 个神经元互连而成。神经元的输出为离散值 1 和 0 (或 -1), 它们分别代表神经元的激活和抑制状态。假若第 i 个神经元在 t 时刻的状态用 $s_i(t)$ 表示, 那么由 N 个神经元构成的神经网络在 t 时刻的状态为一 N 维向量 $\mathbf{S}(t) = [s_1(t), s_2(t), \dots, s_N(t)]^T$ 。这种神经网络的各神经元相互连接, 其连接强度用权值表示。 T_{ij} 代表神经元 i 与神经元 j 的连接权值, $N \times N$ 维矩阵 $\mathbf{T} = \{T_{ij}\}$ 称之为权矩阵。每个神经元都有一阈值 (或称门限值) θ_i 。权矩阵 $\mathbf{T}$ 和阈矢量 $\boldsymbol{\theta} = [\theta_1, \theta_2, \dots, \theta_N]^T$ 定义唯一一个 N 维的离散型 Hopfield 神经网络。

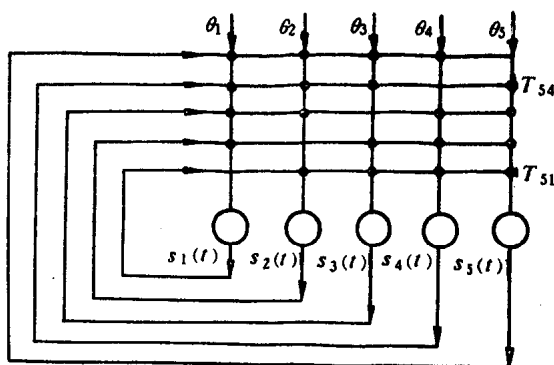


图1.1 离散型 Hopfield 神经网络

稳定性、存储容量、吸引域半径和收敛时间是度量这种神经网络性能的四大指标。如果存在某一时刻 t , 对于 $\forall \Delta t > 0$ 有

$$\mathbf{S}(t + \Delta t) = \mathbf{S}(t) \quad (1.1)$$

则称网络是稳定的, 对应的 $\mathbf{S}(t)$ 为稳定点。存储容量是指网络的稳定点个数。凡是经过一定时间能进入某稳定状态的初始状态集合称之为该稳定点的吸引域, 吸引域半径是指包含在吸引域的最大 Hamming 球半径。稳定点越多, 网络的联想和识别能力就越强。稳定点在空间分布越均匀, 吸引域越大, 则网络的容错能力和自适应能力就越强。收敛时间是网络从某一起始状态进入稳点状态 (或某一极限环) 所需的时间。这四大指标不仅取决于权矩阵 $\mathbf{T}$ 和阈矢量 $\boldsymbol{\theta}$, 而且与网络的工作方式紧密相关。工作方式通常有三种:

(1) 串行(异步)方式: 在某一时刻, 只有某一神经元的状态按(1.2)式发生变化, 而其余 $N-1$ 个神经元的状态保持不变;

$$s_i(t+1) = \text{sgn} \left[\sum_{j=1}^N T_{ij} s_j(t) - \theta_i \right] \quad (1.2)$$

$$s_j(t+1) = s_j(t) \quad j \neq i$$

(2) 部分并行方式: 在某一时刻, 部分神经元的状态按(1.3)式发生变化, 而其余神经元的状态保持不变;

$$s_i(t+1) = \text{sgn} \left[\sum_{j=1}^N T_{ij} s_j(t) - \theta_i \right] \quad (1.3)$$

$$i = i_1, i_2, \dots, i_h$$

$$s_j(t+1) = s_j(t)$$

$$j \neq i_1, i_2, \dots, i_h$$

(3) 全并行(同步)方式: 在某一时刻, 所有的神经元状态均按(1.4)式发生变化。

$$s_i(t+1) = \text{sgn} \left[\sum_{j=1}^N T_{ij} s_j(t) - \theta_i \right] \quad (1.4)$$

$$i = 1, 2, \dots, N$$

关于离散型 Hopfield 神经网络在不同工作方式下的性能已有如下结论(不失一般性, 下面的讨论均设 $\theta = 0$):

结论1.1: 若权矩阵为对称阵, 那么, 网络在同步工作方式下必收敛到一稳定状态或周期为 2 的极限环。

结论1.2: 若权矩阵为对称阵, 且对角线元素非负, 那么, 网络在异步方式下必收敛于一稳定状态。

结论1.3: 若权矩阵为正交投影矩阵, 则在同步方式下必收敛到一稳定状态。

结论1.4: 在稳定性分析中, 同步方式工作的神经网络可以等价于另一个异步方式工作的神经网络。

这几个结论的证明见参考文献[8]、[9]、[10]。在不同权矩阵和不同工作方式下, 离散型 Hopfield 神经网络的稳定性、存储容量、吸收域半径和收敛时间等性能的研究是当前该领域的一个热点, 尤其是在权矩阵元素值为离散值 1 和 -1 的情况, 研究更为引人注目。因为由此构成的网络易于硬件实现。Piret^[12]和杨义先^[13]博士等人已给出了若干有意义的新结果。

1.1.2. 连续状态 Hopfield 神经网络^[14]

连续状态 Hopfield 神经网络如图 1.2 所示。它同样由 N 个神经元互连而成, 但神经元的输出不再是离散值 0 和 1, 而可以在某一区间连续变化。 $g_i(u)$ 为第 i 个神经元的传递函数, 通常取连续、有界和单调增长的 Sigmoid 函数

$$g_i(u) = \frac{1}{2} \left[1 + \tanh \left(\frac{u}{u_0} \right) \right] \quad (1.5)$$

R_i 和 C_i 为第 i 个神经元的输入电阻和输入电容, I_i 为输入偏置电流。根据基尔霍夫定律, 网络具有下列动态方程;

$$\begin{cases} C_i \frac{dU_i(t)}{dt} = \sum_{j=1}^N T_{ij} V_j(t) - \frac{U_i(t)}{R_i} + I_i \\ V_i(t) = g_i(U_i(t)) \end{cases} \quad (1.6)$$

$i=1, 2, \dots, N$

$U_i(t)$ 和 $V_i(t)$ 分别为神经元 i 的输入和输出电压。Hopfield 通过定义能量函数

$$E(t) = -\frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N T_{ij} V_j(t) V_i(t) - \sum_{i=1}^N V_i(t) I_i + \sum_{i=1}^N \frac{1}{R_i} \int_0^{V_i(t)} g^{-1}(V) dV \quad (1.7)$$

并利用动态方程 (1.6) 式, 给出了如下关于网络稳定性的定理:

定理 1.1 若网络中神经元传递函数为单调增长的连续有界函数, 且 $T_{ij} = T_{ji}$, 则有

$$\frac{dE(t)}{dt} \leq 0, \text{ 当且仅当 } \frac{dV_i(t)}{dt} = 0 \text{ 时, } \frac{dE(t)}{dt} = 0$$

证明:

$$\begin{aligned} \frac{dE(t)}{dt} &= -\frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N \left[T_{ij} \frac{dV_j(t)}{dt} V_i(t) + T_{ji} V_j(t) \frac{dV_i(t)}{dt} \right] \\ &\quad - \sum_{i=1}^N I_i \frac{dV_i(t)}{dt} + \sum_{i=1}^N \frac{U_i(t)}{R_i} \cdot \frac{dV_i(t)}{dt} \end{aligned}$$

利用关系 $T_{ij} = T_{ji}$, 上式变为

$$\begin{aligned} \frac{dE(t)}{dt} &= -\sum_{i=1}^N \sum_{j=1}^N T_{ij} V_j(t) \frac{dV_i(t)}{dt} + \sum_{i=1}^N \frac{U_i(t)}{R_i} \frac{dV_i(t)}{dt} - \sum_{i=1}^N I_i \frac{dV_i(t)}{dt} \\ &= -\sum_{i=1}^N \frac{dV_i(t)}{dt} \left[\sum_{j=1}^N T_{ij} V_j(t) + I_i - \frac{U_i(t)}{R_i} \right] \end{aligned} \quad (1.8)$$

将 (1.6) 式代入 (1.8) 式, 则有

$$\frac{dE(t)}{dt} = -\sum_{i=1}^N \frac{dV_i(t)}{dt} C_i \frac{dU_i(t)}{dt} = -\sum_{i=1}^N C_i \left(\frac{dV_i(t)}{dt} \right)^2 [g_i^{-1}(V_i(t))]'$$

从 $g_i(u)$ 的单调增长特性可知 $[g_i^{-1}(V_i(t))]' \geq 0$, 又因 $C_i > 0$, 所以上式右边各项非负, 即

$$\frac{dE}{dt} \leq 0, \text{ 当且仅当 } \frac{dV_i(t)}{dt} = 0 \text{ 时, } \frac{dE(t)}{dt} = 0$$

证毕。

此定理说明网络的能量函数 $E(t)$ 是单调下降的, 若 $E(t)$ 有下界, 那么网络必是稳定的, 且稳定点对应着能量函数 $E(t)$ 的极小点。

显然, 任何一个目标函数可以化为 (1.7) 式的优化问题, 均可以由这种连续状态的 Hopfield 神经网络来求解。由于这种网络具有连续时间动力学性质, 因此, 求解时间在电路时常数数量级 (微秒至纳秒数量级)。Hopfield 这种通过引入能量函数使得具体的网络和

优化问题直接对应的工作被看作是开拓性的, 它把神经网络的计算理论提到了一个新的高度。

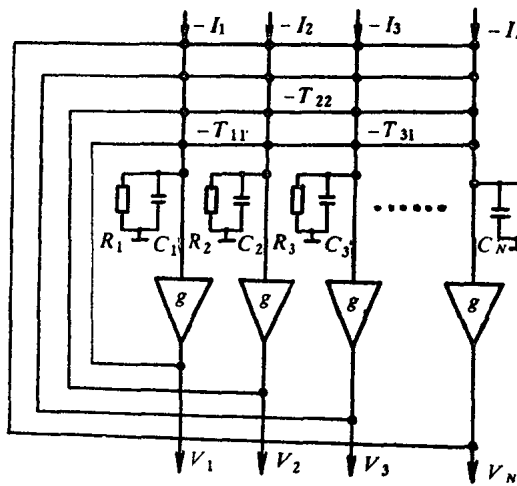


图1.2 连续状态Hopfield神经网络

诸多的优化问题都可以用建立在这个理论基础之上的网络来求解。不同类型的优化问题只是所用的网络模型不同, 如现有的线性规划神经网络^[15]、非线性规划神经网络^[16]、Cellular^[13]网络等等。

对于连续状态 Hopfield 神经网络, 有两个问题是非常重要的, 其一, (1.7) 式中 $E(t)$ 是否有下界; 其二, 神经元的输入电压 $U_i(t)$ 是否有界。它们是网络稳定(或说能求解优化问题)和网络能用硬件实现的必要条件, 为此我们有第二个定理。

定理1.2 若系统的传递函数连续且有界, 则 $U_i(t)$ 和 $E(t)$ 均有界, 且界为

$$U_{\max} = \max_i \left\{ |U_i(0)| + R_i |I_i| + R_i \sum_{j=1}^N |T_{ij}| \right\} \quad (1.9)$$

$$E_{\max} = \sum_{i=1}^N \sum_{j=1}^N |T_{ij}| + \sum_{i=1}^N |I_i| + \sum_{i=1}^N \frac{1}{R_i} g_i(U_{i\max}) U_{i\max} \quad (1.10)$$

其中

$$U_{i\max} = |U_i(0)| + R_i \left[\sum_{j=1}^N |T_{ij}| + |I_i| \right] \quad (1.11)$$

$U_i(0)$ 为第 i 个神经元的初始电压。

在证明这个定理之前, 我们先证明动态方程 (1.6) 式的解可以写成如下积分形式:

$$U_i(t) = U_i(0) e^{-\frac{t}{R_i C_i}} + \int_0^t e^{-\frac{t-\tau}{R_i C_i}} \left[\sum_{j=1}^N T_{ij} V_j(\tau) + I_i \right] / C_i d\tau \quad (1.12)$$

(1.12) 式两边对 t 求导可得

$$\begin{aligned} \frac{dU_i(t)}{dt} &= -\frac{U_i(0)}{R_i C_i} e^{-\frac{t}{R_i C_i}} + \left[\sum_{j=1}^N T_{ij} V_j(t) + I_i \right] / C_i \\ &\quad - \frac{1}{R_i C_i} \int_0^t e^{-\frac{t-\tau}{R_i C_i}} \left[\sum_{j=1}^N T_{ij} V_j(\tau) + I_i \right] / C_i d\tau \\ &= -\frac{U_i(t)}{R_i C_i} + \left[\sum_{j=1}^N T_{ij} V_j(t) + I_i \right] / C_i \end{aligned}$$

显然上式满足 (1.6) 式, 即 (1.12) 式为 (1.6) 式的解。

定理 1.2 的证明: 从 (1.12) 式有

$$|U_i(t)| \leq |U_i(0) e^{-\frac{t}{R_i C_i}}| + \int_0^t e^{-\frac{t-\tau}{R_i C_i}} \left| \sum_{j=1}^N T_{ij} V_j(\tau) + I_i \right| / C_i d\tau$$

$$\begin{aligned}
&\leq \left| U_i(0) e^{-\frac{t}{R_i C_i}} \right| + \int_0^t e^{-\frac{t-\tau}{R_i C_i}} \left[\left| \sum_{j=1}^N T_{ij} V_j(\tau) \right| + |I_i| \right] / C_i d\tau \\
&\leq \left| U_i(0) e^{-\frac{t}{R_i C_i}} \right| + (H_i + \hat{I}_i) \int_0^t e^{-\frac{t-\tau}{R_i C_i}} d\tau \\
&\leq |U_i(0)| + R_i C_i (H_i + \hat{I}_i)
\end{aligned} \tag{1.13}$$

其中

$$\hat{I}_i = |I_i| / C_i \tag{1.14}$$

$$H_i = \max_{\tau} \left| \sum_{j=1}^N T_{ij} V_j(\tau) \right| / C_i \leq \sum_{j=1}^N |T_{ij}| \max_{\tau} |V_j(\tau)| / C_i$$

由于
则

$$|V_j(\tau)| = |g_j(U_j(\tau))| \leq 1,$$

$$H_i \leq 1 / C_i \sum_{j=1}^N |T_{ij}| \tag{1.15}$$

将(1.14)和(1.15)两式代入(1.13)式有

$$|U_i(t)| \leq |U_i(0)| + R_i \left[\sum_{j=1}^N |T_{ij}| + |I_i| \right] = U_{i,\max} \tag{1.16}$$

令

$$U_{\max} = \max_i U_{i,\max}$$

则

$$\max_i |U_i(t)| \leq U_{\max} \quad (i=1, 2, \dots, N)$$

(1.9)式得证。

又, 将能量函数(1.7)式写成

$$E(t) = E_1(t) + E_2(t) \tag{1.17}$$

其中

$$E_2(t) = \sum_{i=1}^N \frac{1}{R_i} \int_0^{V_i(t)} g_i^{-1}(V) dV$$

$$E_1(t) = -\frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N T_{ij} V_i(t) V_j(t) - \sum_{i=1}^N I_i V_i(t)$$

而

$$|E_2(t)| \leq \sum_{i=1}^N \frac{1}{R_i} \left| \int_0^{V_i(t)} g_i^{-1}(V) dV \right| \leq \sum_{i=1}^N \frac{1}{R_i} \int_0^{V_i(t)} |g_i^{-1}(V)| dV \tag{1.18}$$

根据 $U_i(t)$ 的有界性有

$$|g_i^{-1}(V_i(t))| = |U_i(t)| \leq |U_i(0)| + R_i \left[\sum_{j=1}^N |T_{ij}| + |I_i| \right] = U_{i,\max}$$

及

$$|V_i(t)| \leq g_i(U_{i,\max})$$

将这两式代入(1.18)式可得

$$|E_2(t)| \leq \sum_{i=1}^N \frac{1}{R_i} \int_0^{g_i(U_{i,\max})} |g_i^{-1}(V)| dV \leq \sum_{i=1}^N \frac{1}{R_i} g_i(U_{i,\max}) U_{i,\max}$$

令

$$E_{2,\max} = \sum_{i=1}^N \frac{1}{R_i} g_i(U_{i,\max}) U_{i,\max}$$

则有

$$\max |E_2(t)| \leq E_{2,\max} \quad (1.19)$$

因为

$$|E(t)| \leq |E_1(t)| + |E_2(t)|, \text{ 且 } |V_i(\tau)| \leq 1$$

所以

$$\max_t |E(t)| \leq \sum_{i=1}^N \sum_{j=1}^N |T_{ij}| + \sum_{i=1}^N |I_i| + E_{2,\max} = E_{\max}$$

(1.10)式得证。证毕。

需要指出的是, 信号处理的许多问题, 其目标函数只能映射成下面的形式

$$E = -\frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N T_{ij} V_i V_j - \sum_{i=1}^N I_i V_i$$

而非精确的(1.7)式, 因此在用连续状态 Hopfield 神经网络求解这些问题时, 希望积分项

$$\sum_{i=1}^N \frac{1}{R_i} \int_0^{V_i(t)} g_i^{-1}(V) dV \text{ 小到可以忽略。这样, 电阻 } R_i \text{ 的大小和函数 } g_i(u) \text{ 的形式选择很重要。}$$

它们的确定主要根据(1.19)式。显然, R_i 越大, $E_{2,\max}$ 就越小, 但 R_i 太大会导致电路的时常数增大, 从而使网络的收敛时间(求解时间)变长, 因此在实际应用中, 几方面的因素必须综合考虑。

连续状态 Hopfield 神经网络及其改进型已广泛应用于信号分解、自适应滤波、谱估计、A/D变换等信号处理的许多领域。有关这些应用我们将在后面章节里陆续介绍。

§1.2 多层感知器及BP学习算法

多层感知器是一种典型的前馈神经网络, 它通常由输入层、输出层和若干隐层组成, 如图 1.3 所示。每一层包含有若干个神经元。第 k 层中第 j 个神经元具有下列输入输出关系:

$$y_j^{(k)} = f_j^{(k)} \left(\sum_{i=1}^{N_{k-1}} W_{ij}^{(k-1)} y_i^{(k-1)} - \theta_j^{(k)} \right) \quad (1.20)$$

$$j=1, 2, \dots, N_k, \quad k=1, 2, \dots, M$$

式中 $W_{ij}^{(k-1)}$ 为第 $k-1$ 层中第 i 个神经元到第 k 层第 j 个神经元的连接强度(权值); $\theta_j^{(k)}$ 为对应神经元的阈值; $f_j^{(k)}$ 为神经元的传递函数(通常是一非线性函数, 并且每层中各神经元都取相同的形式); N_k 为第 k 层神经元的数目, M 表示总层数。 $M=1$ 为单层感知器。目前研究最多的是 $M=3$, 即只含有一个隐层。

利用(1.20)式, 可以求出网络总的输入 $y_1^{(0)}, y_2^{(0)}, \dots, y_{N_0}^{(0)}$ 与输出 $y_1^{(M)}, y_2^{(M)}, \dots, y_{N_M}^{(M)}$ 之

间的关系式。实际上,代表输入输出之间变换关系的有关信息主要分布在神经元之间的连接强度上,不同的连接强度反映着不同的输入输出关系。因此,这种神经网络具有分布式存储信息的特点。

多层感知器属于映射型神经网络。它的信息处理功能实际上是把 N_0 维欧几里德空间中的子空间 A 映射成 N_M 维欧几里德空间中的子空间 $f(A)$,即 $f: A \subset R^{N_0} \rightarrow R^{N_M}$ 。下面的定理从数学上证实了这种映射的成立。

定理1.3 对于 $\forall \varepsilon > 0$ 和任意的 L_2 函数 $f: [0,1]^{N_0} \subset R^{N_0} \rightarrow R^{N_M}$,总存在着一个三层的感知器,使得感知器完成的映射与 f 的均方差小于 ε 。

此定理说明一个三层感知器可以以任意精度逼近 L_2 函数。多层感知器所具有的这种非线性变换特性为它的广泛应用奠定了理论基础。

定理1.4 隐层神经元数目为 $2N_0+1$ 的三层感知器可以精确实现任意的连续映射 $f: [0,1]^{N_0} \subset R^{N_0} \rightarrow R^{N_M}$ 。

这就是著名的Kolmogorov定理。定理1.3和1.4的证明详见文献[18]和[19]。

在实际应用中,往往已知的是 $A \subset [0,1]^{N_0}$ 中若干个样本集 $Y_1^0, Y_2^0, \dots, Y_I^0$ 和对应变换 $\hat{Y}_1^M, \hat{Y}_2^M, \dots, \hat{Y}_I^M$,而非准确知道其映射关系 f (或者说找不出 f 的显式表达式)。此时,网络的连接权值可通过求解下列优化问题学习得到:

$$\min_W \sum_{i=1}^I \|Y_i^M - \hat{Y}_i^M\|^2 \quad (1.21)$$

其中

$$Y_i^0 = [y_{i1}^{(0)}, y_{i2}^{(0)}, \dots, y_{iN_0}^{(0)}]$$

$$Y_i^M = [y_{i1}^{(M)}, y_{i2}^{(M)}, \dots, y_{iN_M}^{(M)}] \quad i=1,2,\dots,I$$

$$\hat{Y}_i^M = [\hat{g}_{i1}^{(M)}, \hat{g}_{i2}^{(M)}, \dots, \hat{g}_{iN_M}^{(M)}]$$

$y_{ij}^{(0)}$ 为第 i 个样本集上输入层第 j 个神经元的输入值; $\hat{g}_{ij}^{(M)}$ 和 $y_{ij}^{(M)}$ 分别为对应的变换值和网络输出层第 j 个神经元的输出。

所谓的权值学习就是不断地调整 W ,使得 Y_i^M 和 $\hat{Y}_i^M$ 的空间距离最小,以期最终的 W 所确定的网络能在最小二乘意义上逼近所对应的映射关系 f 。反向传播学习算法(简称BP算法)是最简单、最实用的一种。如果多层感知器中各神经元的传递函数均取为

$$f_1(x) = \frac{1}{1 + e^{-(x-\theta)}}$$

其中 θ 为一固定参量,那么,BP算法的具体步骤为:

- (1) 初始化,即随机地设置各层权系数 $W_{ij}^{(k)}$ 和阈值 $\theta_j^{(k)}$;
- (2) 利用输入 $Y_1^0, Y_2^0, \dots, Y_I^0$ 和(1.20)式计算各层实际的输出 $Y_1^1, Y_2^1, \dots, Y_I^1 (k=$

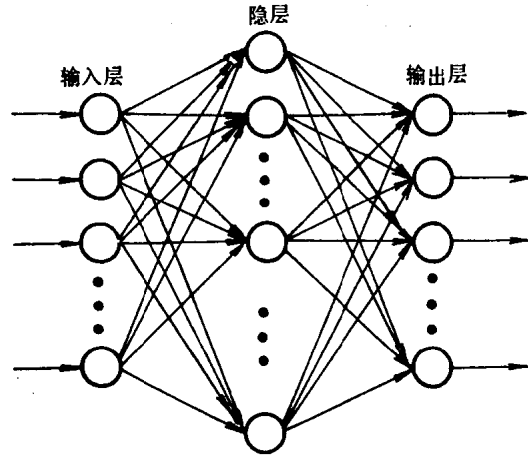


图1.3 多层感知器

1, 2, \dots, M);

(3) 利用下式调整权系数

$$W_{ij}^{(k-1)}(t+1) = W_{ij}^{(k-1)}(t) + \eta \sum_{h=1}^I \delta_h^{(k)} y_h^{(k)} \quad (1.22)$$

其中 $W_{ij}^{(k-1)}(t)$ 为当前权; $W_{ij}^{(k-1)}(t+1)$ 为修正后的权; η 为步长, 它的选择并没有固定的准则; $y_h^{(k)}$ 为第 k 层第 j 个神经元在第 h 个样本输入时由 (1.20) 式算出的值; $\delta_h^{(k)}$ 称之为误差传播项, 当 $k=M$ 时, 对应着输出层有

$$\delta_h^{(M)} = (\hat{y}_h^{(M)} - y_h^{(M)}) f'_1(y_h^{(M)}) \quad (1.23)$$

当 $k < M$ 时, 对应着隐层有

$$\delta_h^{(k)} = f'_1(y_h^{(k)}) \sum_{l=1}^{N_{k+1}} \delta_{hl}^{(k+1)} W_{jl}^{(k)}(t) \quad (1.24)$$

(4) 计算误差 ε_1

$$\varepsilon_1 = \sum_{h=1}^I \|\hat{\mathbf{Y}}_h^M - \mathbf{Y}_h^M\|^2 = \sum_{h=1}^I \sum_{j=1}^{N_M} (\hat{y}_{hj}^{(M)} - y_{hj}^{(M)})^2$$

(5) 如果 $\varepsilon_1 > \varepsilon$ (ε 为预定的误差)。则返回至步骤 (2) 继续学习; 如果 $\varepsilon_1 \leq \varepsilon$, 则认为已收敛而停止学习, 由此时的连接权值构成的网络, 就可在 ε 误差范围内反映对应的输入输出关系。

关于 BP 学习算法需要说明几点:

(1) 第 (2) 和 (3) 两步中, 样本的输入和权值的调整可以结合起来进行, 即每次输入一个样本后, 先可用 (1.22) 式进行权调整, 然后再转至第 (2) 步, 输入另外一个样本, 并执行第 (3) 步, 直到 I 个样本全部输入完。最后的误差由下式给出

$$\varepsilon_1 = \|\hat{\mathbf{Y}}_I^M - \mathbf{Y}_I^M\|^2 = \sum_{l=1}^{N_M} (y_{Il}^{(M)} - \hat{y}_{Il}^{(M)})^2$$

模拟结果表明, 这种循环套循环的方法可以加快收敛速度。

(2) 由于 BP 算法是一种梯度算法, 所以不能保证连接权值收敛于全局最优解。

(3) BP 算法的收敛速度太慢。现已有许多的改进型, 如变步长的 BP 算法、最小二乘学习算法等等^[22]。

§1.3 自组织神经网络^[23]

多层感知器的学习和分类是以已知一定的先验知识为条件的, 即网络权值的调整是在监督情况下进行的。而在实际应用中, 有时并不能提供所需的先验知识, 这就需要网络具有能够自学习的功能。Kohonen 提出的自组织特征映射图就是这种具有自学习功能的神经网络。这种网络是基于生理学和脑科学研究成果提出的。脑神经科学研究表明: 传递感觉的神经元排列是按某种规律有序进行的^[24], 这种排列往往反映所感受的外部刺激的某些物理特征。例如, 在听觉系统中, 神经细胞和纤维是按照其最敏感的频率分布而排列的。为此, Kohonen

认为,神经网络在接受外界输入时,将会分成不同的区域,不同的区域对不同的模式具有不同的响应特征,即不同的神经元以最佳方式响应不同性质的信号激励,从而形成一种拓扑意义上的有序图。这种有序图也称之为特征图,它实际上是一种非线性映射关系,它将信号空间中各模式的拓扑关系几乎不变地反映在这张图上,即各神经元的输出响应上。由于这种映射是通过无监督的自适应过程完成的,所以也称它为自组织特征图。

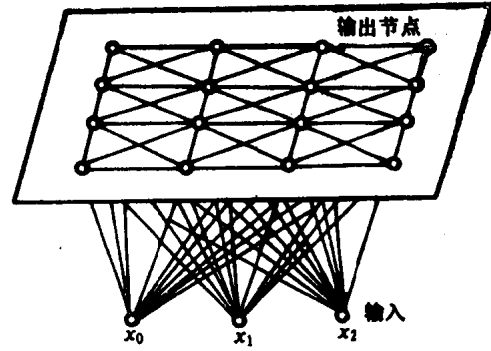


图1.4 自组织神经网络

如图 1.4 所示,在这种网络中,输出节点与其邻域其它节点广泛相连,并互相激励。

输入节点和输出节点之间通过强度 $w_{ij}(t)$

相连接。通过某种规则,不断地调整 $w_{ij}(t)$,使得在稳定时,每一邻域的所有节点对某种输入具有类似的输出,并且这种聚类的概率分布与输入模式的概率分布相接近。

完成自组织特征映射的算法较多。下面给出一种常用的自组织算法:

- (1) 权值初始化并选定邻域的大小;
- (2) 输入模式;

- (3) 计算空间距离 $d_j = \sum_{i=0}^{N-1} [x_i(t) - w_{ij}(t)]^2$, 式中 $x_i(t)$ 是 t 时刻 i 节点的输入, $w_{ij}(t)$

是输入节点 i 与输出节点 j 的连接强度, N 为输入节点的数目;

- (4) 选择节点 j^* , 它满足 $\min_j d_j$

- (5) 按下式改变 j^* 和其邻域节点的连接强度

$$w_{ij}(t+1) = w_{ij}(t) + \eta(t) [x_i(t) - w_{ij}(t)] \quad (1.25)$$

$j \in j^*$ 的邻域, $0 \leq i \leq N-1$

式中 $\eta(t)$ 称之为衰减因子;

- (6) 返回到第 (2) 步, 直至满足 $[x_i(t) - w_{ij}(t)]^2 < \epsilon$ (ϵ 为给定的误差)。

通过这种无导师的学习, 稳定后的网络输出就对输入模式生成自然的特征映射, 从而达到自动聚类的目的。

§1.4 高阶神经网络

高阶神经网络也经常用于信号处理中, 尤其是用于非线性信号处理及不变性模式识别中。图1.5是一高阶神经网络的结构示意图。 k 阶神经网络第 i 个神经元的输出为

$$y(i) = f_i[T_0(i) + T_1(i) + \dots + T_k(i)] \quad (1.26)$$

式中 $f_i(\cdot)$ 为神经元的传递函数, k 是神经元输入加权项的最高幂次, 即为阶数, $T_k(i)$ ($k=1, 2, \dots, k$) 是第 i 个神经元的第 k 阶项, 是输入矢量 k 项分量乘积的加权和。例如:

$$\text{一阶项 } T_1(i) = \sum_j W_1(i, j) x(j)$$