

第 1 章 绪论

1.1 引言

人工神经网络理论的研究是当前许多工程研究领域的热点话题，它的应用领域非常广泛，仅在与工业自动化相关的领域中，如模式识别 (Narendra et al. 1990; Cooper et al. 1992)、信号处理 (Tank et al. 1986)、知识工程、专家系统、优化组合、智能控制 (Narendra et al. 1992) 等方面，已经有许多神经网络应用成功的论文报道。另外，许多非自动化工程领域，如应用声学、气象预报、情报分类、工程力学、土木工程、建筑结构等也开始注意到了人工神经网络这一有力的工具，并已逐步应用它来解决各类工程问题 (Poggio et al. 1990; Raol et al. 1996; Delgado et al. 1995; Lippman et al. 1987; Antsaklis 1995)。

没有一个在控制系统领域工作的人会意识不到关于人工神经网络专题的论文、杂志、会议和国际会议数量的增长。这清楚地表明了大量学者正在对与神经网络有关的概念发生兴趣，并希望将相关的算法用于各种应用领域。控制系统已成为一种令人关注的应用领域。

就像其他发展领域（例如自适应控制和专家系统等）一样，在研究过程中可能会出现两种极端的观点，其中的任何一种观点都会对学科发展造成阻碍。第一种这样的观点是不加考虑地接受新事物，或者新概念和新名词，并且不科学地加以应用。Hunt 等 (1992) 在其所发表的综述性论文中对其有十分形象的描述：

这就产生了这种形式的论文：“X 是近年来的热门话题，因此

可以替代其他所有的方法，我们将其应用于某种控制器之中，得到了 X 控制器，这种控制器被证明是最好的控制器，我们对其作了仿真，得到了引人注目的结果。”这里，X 是自调整、专家系统、人工神经网络或任何其他热门话题。这种论文倾向于用拟人化的评论来说明成功的显著性。

当然，第二种极端的观点是将这些发展领域当作纯粹的夸张加以丢弃，不准备应用它们。

如果我们想在任何研究领域取得进展的话，这些极端的观点都必须加以避免。相反，需要用已有的成熟技术对这些新观点加以理性的评估。必须避免两种极端的态度，并需要切实地引导一条中间途径，或者说是辩证的途径，使新观点和已有的理论加以融合，这样才能得到真正的、而不是虚幻的进展。

人工神经网络理论为什么能在如此众多的工程研究领域获得关注，并且能取得如此多的成功应用呢？尤其是，为什么自动控制研究领域对其如此关注？这就是本章的论述着力想说明的问题。作为一个合格的工程研究人员，我们尤其需要避免的是在头脑中产生下列公式：

因为“该项理论是当前的时髦话题，且论文相对较多”。

所以“我的研究中也应用该项理论”。

如果这样，我们所注意到的就仅仅是该项热门理论的名词本身，而不是该理论能获得成功应用的实质所在了。

作为智能控制领域的研究人员，作者很自然地关注到了人工神经网络的动态收敛特性及反馈特性。但是，目前流行的各种多层前馈式神经网络模型及其运算代价较高的学习算法在此领域并不具有很好的特征。随着作者研究工作的深入开展，我们的注意力就逐步集中到了具有动态反馈自收敛特性的反馈式神经网络（如 Hopfield 神经网络）上了。经过多年的系统研究，我们充分证明了这种神经网络在控制领域的适用特征，同时也取得了一些相

关的理论及仿真实验结果，最终导致了此书的诞生。

人工神经网络的研究最初起源于信号处理研究领域，似乎很难直接为控制领域的研究人员所用而那些读过 Wiener 的著作《控制论》的人 would 知道，控制、信息和神经学科在控制论的范围内一度被看成是同一个主题。但从那时起，计算科学（包括人工智能）及神经生物学趋向于沿各自的不同方向发展，这就阻碍了对各方向所取得的研究成果进行交流。尤其是各方向中不同的术语和注释现在已经成了进行有效思想交流的障碍。

我们必须致力于为正在控制系统领域工作的研究人员去除这一障碍，用控制工程师所熟悉的语言和符号来给出人工神经网络的基本思想和技术。这有两方面的好处，既能使这种技术更易被控制工程师接受，又能使控制系统的知识应用于人工神经网络领域。

本章力图从人工神经网络的研究历史出发，详细论述其适用于控制工程领域的主要特征，并针对当前自动控制系统所遇到的难点问题说明为什么人工神经网络理论能在其中获得如此广泛的关注。

1.2 人工神经网络理论的研究简史 及标志性研究成果

人工神经网络 (Artificial Neural Network, ANN) 这一名词是相对于生物学中所说的生物神经网络系统而言的。它的提出者的目的就在于用一些简单的数学模型来对生物神经网络进行描述，并在一定的算法指导下，使其能在某种程度上模拟生物神经网络所具有的智能行为，解决用传统算法所不能胜任的智能信息处理和智能计算问题 (焦李成 1990, 1993; Hunt et al. 1992)。

人工神经网络的研究始于 1943 年 至今已经历了 50 多年的漫长历程。它的研究并不是从一开始就受到广泛关注的，而是经

历了一条从兴起到萧条，又从萧条到兴盛的曲折发展道路。具体说来，大致可分为以下几个阶段。

1.2.1 初创期

由心理学家 W. S. McCulloch 和数学家 W. Pitts 于 1943 年所提出的 M-P 模型，一般被公认为是神经网络研究的开创性成果。该模型研究的是用逻辑的数学工具来把客观事件用形式神经元进行表达。该模型类似于现在的 阈值单元模型，但是其连接权是不作调整的。该模型的提出不仅具有开创意义，还为以后的研究工作提供了一定的形式化依据。

该段时期的另一重要成果是由心理学家 D. O. Hebb 所提出的神经元之间突触强度的调整规则假说。他认为神经网络学习过程的变化体现为网络神经元之间突触连接权重的变化，而连接权的强度调整应正比于两个互连神经元之间激活值输出的乘积，这就是有名的 Hebb 规则。它是 Hebb 借鉴对大脑神经细胞、人类学习行为和条件反射的观察研究的成果而提出的。该规则至今仍在各种神经网络模型中起重要作用。

该段时期的神经网络理论研究是探索性的，但同时也是开创性的。许多成果至今仍对神经网络的理论研究有着重要影响。

1.2.2 初始兴盛期

该段时期的神经网络理论研究基本上确立了从系统的角度研究人工神经网络的基础。其代表性的研究成果是由 F. Rosenblatt 于 20 世纪 50 年代末所提出的感知器 (Perceptron) 模型。这是历史上第一个具有完整意义的神经网络模型，该模型已初步具备了并行处理、分布式存储和学习等神经网络的一些基本特征，并给出了二层感知器的收敛定理证明。

由 B. Widrow 和 M. E. Hoff 于 20 世纪 60 年代初所提出的自适应线性单元 (Adaline) 网络在自适应系统 如自适应滤波、预测和模式识别等的研究中得到了很好的结果。这是一种连续取值的线

性网络模型，与当时占主导地位的以顺序离散符号推理为基本特征的人工智能理论完全不同。它的提出引起了许多研究者的兴趣。在该段时期，人工神经网络的研究开始受到人们的重视，研究工作进入了初始的兴盛时期。

1.2.3 低潮期

由于以逻辑推理为基础的人工智能理论和 Von Neumann 型数字计算机正处于全盛的发展时期，掩盖了发展新型智能计算理论和新型智能技术的必要性。同时也由于处于发展初始阶段的神经网络理论还有一定的缺陷，使人工神经网络理论研究走入了一个缓慢发展的低潮阶段。其标志是著名的人工智能学者——美国麻省理工学院的 M. L. Minsky 和 S. A. Papert 所著的 *Perceptron* 一书的出版。书中研究了以 *Perceptron* 模型为代表的神经网络，在数学上作了深入研究，指出了其局限性，证明了单层感知器只能用于线性问题的求解，而对于即使是简单的“异或”这样的非线性问题便无能为力，并指出还不能从理论上证明将单层感知器网络扩展为多层的意义所在。实际上，该论点的缺陷在于未用辩证思维的模式来思考问题，现在我们已经知道，解决这一问题的方法很简单，只要将神经网络的层数增加即可。

但即使处于低潮期，神经网络的研究并未终止。在此期间所提出的自组织映射理论、自适应共振理论和神经认知机模型等都以后的神经网络研究产生了重大的影响。

1.2.4 复兴期

由于神经网络研究者取得的突出成果，同时也由于传统的人工智能理论和 Von Neumann 型计算机在许多智能信息处理问题上遇到了挫折，另外由于科学技术的发展为人工神经网络的物质实现提供了基础，促使人们将人工神经网络作为未来新一代智能计算机的实现模式之一加以研究。

该阶段的标志之一是美国加州理工学院的生物物理学家 J. J.

Hopfield 于 1982 年和 1984 年所提出的一种新型神经网络模型——Hopfield 神经网络模型以及该模型的电子线路实现。这种神经网络借鉴了物理学中自旋玻璃子研究所采用的哈密尔顿算子,首次引入了网络能量函数的概念,并给出了稳定性证明。在 1986 年他于 *IEEE Circuit and System* 刊物上所发表的论文中,该神经网络被用来成功地解决了一个著名的组合优化问题——旅行商问题,同时,用该模型能够很好地解决线性规划问题。该项研究引起了全世界相关领域研究人员的广泛关注。这也是本书作者所关注的模型。

该阶段的另一重要成果是 D. E. Rumelhart 和 J. L. McClelland 等人所在的并行分布处理研究(PDP)小组所提出的 BP(误差反向传播)学习算法。该算法解决了在多层神经网络学习训练过程中,中间隐含层各连接权重的调节方法问题,从而突破了 M. L. Minsky 等人所持悲观论点的前提条件,达到了预期的学习目的,成为迄今为止在神经网络研究领域影响最大的一种学习算法,至今仍得到广泛的应用。

1.2.5 高潮期

自 Hopfield 神经网络模型和 BP 算法被提出之后,很快掀起了人工神经网络研究的全球性热潮。在此期间,各种神经网络模型相继被提出,其中著名的有:1988 年 L. O. Chua 所提出的 CNN(细胞神经网络)模型,1984 年 Hinton 所提出的 Boltzmann 机, S. Grossberg 等提出的 ART 自适应共振理论, J. S. Albus 的 CMAC(小脑模型)网络及 R. M. Sanner 等人提出的 RBF 高斯径向基函数网络等。同时各种学习算法也得到了发展。其应用已很快渗透到计算机图像处理、语音处理、优化计算、智能控制等领域,并取得了很大的进展。

现在可以认为,人工神经网络的研究正在转入高潮期的快速稳定发展阶段。在理论研究上正在进一步深入,并开发新的网络

数理理论；在应用研究方面，正在进一步进行其软件模拟和硬件实现的研究，并迅速扩展其应用领域，取得了更广泛的成果。

通过对人工神经网络研究历史的回顾，可以得到一个很深刻的体会：一项新理论能否很快得到工程研究领域的认可并获得广泛应用，不仅与该理论本身的严密性、实用性和可行性有关，还与工程应用领域所处的技术发展阶段和应用需求有关。作为一个工程研究人员，应该面向实际的工程应用找到问题，并力图用最合适的科技手段加以解决

1.3 符号主义和连接主义

近年来人们对大脑机制和结构的研究兴趣逐渐增强，从而导致了基于神经网络这种生物背景的新型计算模式的迅速发展，该模型最直接的应用领域就在于解决模式识别、快速信号处理和自适应等复杂问题。连接主义的概念就取自其生物背景，认为这种模拟生物系统连接模式的智能模式能够优于其他一切人工智能模式。

作为连接主义研究的标志性事件，19世纪40年代早期，该领域的先期研究者 McCulloch 和 Pitts(1943)研究了基于神经元模型的几个基本部件互连的潜能和容量。其他人，如 Hebb(1949)则致力于神经系统的自适应规律研究。Rosenblatt(1958)创造了感知器这一名称，并发现了在以后引起许多研究人员注意的一种网络结构。Minsky 和 Papert(1969)对感知器进行了严密分析，他们证实了许多神经网络的独特性能，并指出了一些模型的局限性。19世纪70年代，Grossberg的工作引起了人们的注意，该工作在基于心理学的证据基础上，提出了具有新颖特征的几种非线性动态系统结构。Hopfield又应用一种特殊的非线性动态网络结构解决了一类最优化技术问题——TSP问题(Hopfield 1982, 1986)。在1986年，并行分布处理小组(PDP小组)公布了一系列的研究结论

和算法, Rumelhart 等(1986)对相关领域起到了强劲的激励作用,并促进了该领域的许多后续研究。由于 PDP 小组的工作许多研究结论使得我们现在能用一些定义很好的神经网络结构来解决许多种类的问题(Hecht - Nielsen 1988)。

而符号主义的研究则直接源于人工智能 (Artificial Intelligence, AI) 研究领域。在 19 世纪 50 年代人工智能理论诞生之后,起先它的发展是和神经网络研究基本同步进行的。自 1956 年 Dartmouth 夏季 AI 研究计划之后,人工智能领域的黑箱法和神经网络的一些基本算法同时得到了迅速的发展。但自 19 世纪 70 至 80 年代,人工智能的典型模式——专家系统模型得到迅速发展之后,就有一些人工智能研究领域的人员提出,基于符号表示和符号推理的专家系统模式能够替代一切现存的智能系统,通过符号化的处理模式能够达到一切智能模拟的目的,甚至能够做到对人力的替代。这在当时,甚至现在看来,都是有些不切实际的。事实上,具有冷静的逻辑思维能力和辩证唯物主义观点的研究人员都知道,一切客观存在的事物都有其内在的规律和特点,也有符合其生存和运行的典型环境,用其中的一种来完全替代另一种显然是不可能的。

这里,我们姑且不论这两种研究思路各自的优劣,我们暂时只能凭自己的兴趣和努力对其中的一种进行较为系统的研究。如果我们的研究能够对相关领域的人员有一定的参考价值的话,这将是我们的最大快乐。

第 2 章 人工神经网络的统一描述 及典型结构

2.1 人工神经元的统一描述

要确切地对各种人工神经网络的外在特性进行描述，就必须寻求其基本组成单元——人工神经元的统一描述形式，并对各神经元的连接方式加以定义。连接主义结构的基本处理单元在神经生理学的类比分析中常被叫做神经元，但其他名称如感知器 (Perceptron) (Rosenblatt 1958) 或自适应线性神经元 (Adaline) (Widrow, Hoff 1960) 等也常被使用。下面我们描述一个标准的通用神经元模型，本书后面的论述可以证明，许多有名的神经网络结构都处于该标准模型的描述范围内 (邓志东 1994; Seinowski et al. 1988; Hunt et al. 1992; Guez et al. 1988; Kraft et al. 1990; Amari 1990; Michel et al. 1989)。

许多基本的神经网络处理单元都可被认为有以下三个部分 (见图 2-1)：

- (1) 加权求和器；
- (2) 线性动态单输入 / 单输出系统；
- (3) 非动态的非线性函数映射关系。

这些部分在下面将依次讨论。

2.1.1 加权求和器

加权求和器可用下式表示：

$$v_i(t) = \sum_{j=1}^N a_{ij}y_j(t) + \sum_{k=1}^M b_{ik}u_k(t) + w_i \quad (2-1)$$

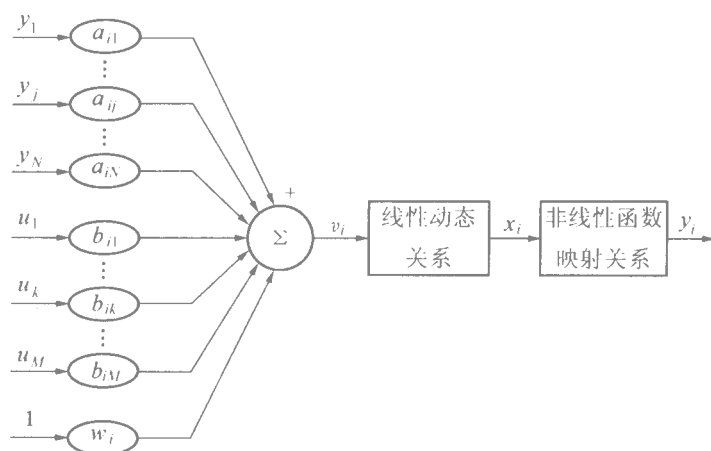


图 2-1 神经元的基本模型

式 (2-1) 给出了针对所有单元输出 y_i 、外部输入 u_k 和常数 w_i 的加权求和 v_i ，这些加权求和元素能够方便地用矢量矩阵的方式加以描述。将 N 个加权求和值 v_i 存放在列矢量 V 中， N 个输出 y_i 存放在矢量 Y 中， M 个输入 u_k 存放在矢量 U 中， N 个常数 w_i 存放在矢量 W 中，方程 (2-1) 就能用矢量矩阵的形式重写如下：

$$v(t) = Ay(t) + Bu(t) + w \quad (2-2)$$

这里， $N \times N$ 矩阵 A 中的第 ij 个元素为 a_{ij} ， $N \times M$ 矩阵 B 中的第 ik 个元素为 b_{ik} ，常数 w 可被包括进输入 u_k 中，但将它们分开来表示是有用的， w_i 表示与神经网络阈值相关的加权量。

2.1.2 线性动态系统

线性动态单输入 / 单输出系统包含输入 v_i 和输出 x_i ，用传递函数的形式可以描述为：

$$\tilde{x}_i(s) = H(s)\tilde{v}_i(s) \quad (2-3)$$

式中 横线表示 Laplace 变换。在时域中，方程 (2-3) 就变成：

$$x_i = \int_{-\infty}^t h(t-t') v_i(t') dt' \quad (2-4)$$

这里 $H(s)$ 和 $h(t)$ 形成了一对 Laplace 变换对。 $H(s)$ 最常用的五种选择为：

$$\begin{aligned} H(s) &= 1 \\ H(s) &= \frac{1}{s} \\ H(s) &= \frac{1}{1+sT} \\ H(s) &= \frac{1}{\alpha_0 s + \alpha_1} \\ H(s) &= e^{-sT} \end{aligned} \quad (2-5)$$

分别对应于：

$$\begin{aligned} h(t) &= \delta(t) \\ h(t) &= \begin{cases} 0 & t < 0 \\ 1 & t \geq 0 \end{cases} \\ h(t) &= \frac{1}{T} e^{-t/T} \\ h(t) &= \frac{1}{\alpha_0} e^{-(\alpha_1/\alpha_0)t} \\ h(t) &= \delta(t-T) \end{aligned} \quad (2-6)$$

式中, δ 是 Dirac-delta 函数。在时域中, 相应的输入 / 输出关系为：

$$\begin{aligned} x_i(t) &= v_i(t) \\ \dot{x}_i(t) &= v_i(t) \\ T\dot{x}_i(t) + x_i(t) &= v_i(t) \\ \alpha_0\dot{x}_i(t) + \alpha_1 x_i(t) &= v_i(t) \\ x_i(t) &= v_i(t-T) \end{aligned} \quad (2-7)$$

显然, 前面的三个式子只是第四式的特殊形式。

如果用到离散时间动态系统，则可表示为：

$$\alpha_0 x_i(t+1) + \alpha_1 x_i(t) = v_i(t) \quad (2-8)$$

式中 i 为积分时间序列。

2.1.3 非动态的非线性函数

非动态的非线性函数 $g(\cdot)$ 给出了传递函数输出 x_i 和单元输出 y_i 之间的相互关系：

$$y_i = g(x_i) \quad (2-9)$$

这里还有一些关于这些函数的习惯性分类：

- (1) 可导 / 不可导函数；
- (2) 脉冲 / 阶跃函数；
- (3) 正均值 / 零均值函数。

其中，第一类区分了光滑函数和不光滑函数。光滑函数在一些自适应算法，如反向传播算法（BP 算法）（Rumelhart et al. 1986）中是必须的，而非连续函数（即稳态函数）在给出实际的二进制输出时是必须的。

第二类区分了仅在输入接近于零时有显著输出值的函数与只在零值附近变化显著的函数。

第三类指阶梯型函数。正均值函数的值域从 0 变化到 1 相应的定义域从 $-\infty$ 到 $+\infty$ ，零均值函数的值域从 -1 变化到 +1，相应的定义域亦为 $-\infty$ 到 $+\infty$ 。

表 2-1 中给出了一些标准的非线性函数。注意，在表中所给出的函数之间有着很强的联系。S 型函数和 Th 函数是相似的；S 型函数从 0 变到 1；而 Th 函数从 -1 变到 +1。另外，阈值函数相当于 S 型函数和 Th 函数的高增益极限状态。

当然，可以将可导的阶梯状函数通过求导取反来得到脉冲状函数。同样，脉冲状函数能通过对平移的阶梯状函数进行求导得到。

表 2-1 典型非线性函数 $g(x)$

名称	形式	特征
阈值函数	$x > 0$ 时 y 为 +1 否则为 0	不可导, 阶梯状, 正均值
阈值函数	$x > 0$ 时 y 为 +1 否则为 -1	不可导, 阶梯状, 零均值
S 型函数	$\frac{1}{1 + e^{-x}}$	可导, 阶梯状, 正均值
双曲正切函数	$\text{Tanh}(x)$	可导, 阶梯状, 零均值
高斯型函数	$e^{(-x^2/\sigma^2)}$	可导, 脉冲状

众所周知, 处于神经网络系统不同部位的神经元具有不同的特征。例如, 视觉运动肌的神经元具有 S 型的响应特征 而那些在视觉区的神经元则具有高斯型响应特征。前一种神经元叫做分类编码器, 并且呈现单调的映射特征。第二种神经元则展现出最大响应特性, 覆盖整个测量空间, 所起的是数值编码的作用 (Ballard 1988) 还有其他一些有用的响应函数, 如对数型和指数型函数, 但它们的生物基础还未被报道。每一种基本单元的表达方式都有其优缺点。例如, 具有高斯型响应特征的神经元似乎具有表达多维变量的能力, 这种功能实现起来比具 S 型响应特征的神经元要容易得多。

2.2 人工神经网络连接权的作用

神经元自身从 计算或表达能力方面来说并不具有很强的功能, 但它们的互相连接使我们能表达不同的变量关系, 同时得到强大的处理功能。前面所讨论的神经元的三个部分能以各种不同的方式加以连接。例如, 如果神经元都是非动态的 $[H(s) = 1]$ 那么神经元的排列可以用代数方程的集合加以表示。通过连接方程 (2-2)、(2-3) 和 (2-9) 可以得到:

$$\begin{aligned}x(t) &= Ay(t) + Bu(t) + w \\y(t) &= g(x(t))\end{aligned}\quad (2-10)$$

式中 x 是 N 个 x_i 元素组成的矢量, 而 $g(x)$ 是一个分量为 $g(x_i)$ 的矢量。另外, 如果每个神经元都具有一阶低通动态特性 $[H(s) = \frac{1}{Ts + 1}]$, 那么, 神经元的排列能以微分方程的集合来加以

描述:

$$\begin{aligned}T\dot{x}(t) + x(t) &= Ay(t) + Bu(t) + w \\y(t) &= g(x(t))\end{aligned}\quad (2-11)$$

显然 方程(2-10) 的解就是方程 (2-11) 的稳态解。

方程 (2-11) 的离散时间动态表达为:

$$\begin{aligned}Tx(t+1) + (1-T)x(t) &= Ay(t) + Bu(t) + w \\y(t) &= g(x(t))\end{aligned}\quad (2-12)$$

式中 t 为积分时间序列。

这种神经网络的特性显然取决于内部连接矩阵 A 和 $H(s)$ 的阶数。下面, 我们来讨论一些重要的神经网络结构。

2.3 典型人工神经网络结构

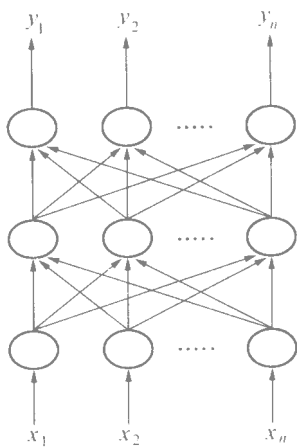
图 2-2 中给出了五种典型的神经网络结构。

(a) 前馈式网络

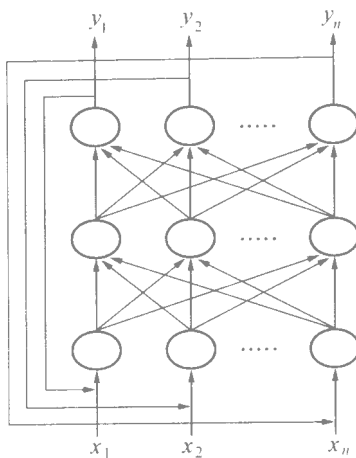
这种网络结构是分层排列的, 每一层的神经元输出只和下一层神经元相连。这种网络结构特别适用于 BP 算法 如今已得到了非常广泛的应用。

(b) 输出反馈的前馈式网络

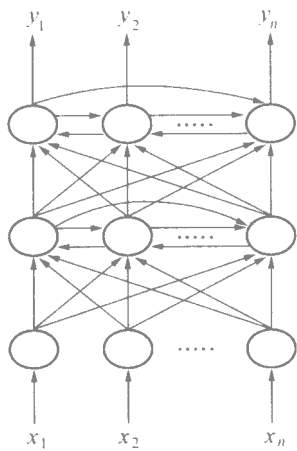
这种网络结构与前馈式网络的不同之处在于这种网络结构存在着一个从输出层到输入层的反馈回路。这种网络结构适用于顺序型的模式识别问题 如 Fukushima 所提出的网络模型结构。



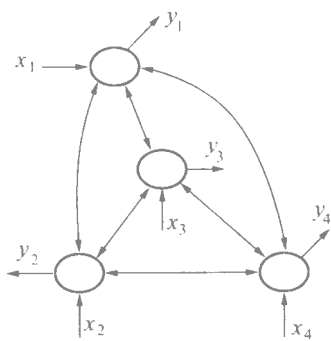
(a) 前馈式网络



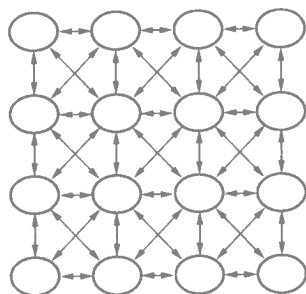
(b) 输出反馈的前馈式网络



(c) 前馈式内层互连网络



(d) 反馈型全互连网络



(e) 反馈型局部互连网络

图 2-2 典型神经网络结构示例

(c) 前馈式内层互连网络

这种网络结构中，同一层之间存在着相互关联，神经元之间有相互制约的关系，但从层与层之间的关系来看还是前馈式的网络结构。许多自组织神经网络大多具有这种结构，如 ART 网络等。

(d) 反馈型全互连网络

这种网络中，每个神经元的输出都和其他神经元相连，从而形成了动态的反馈关系，如 Hopfield 网络。这种网络结构具有关于能量函数的自寻优能力，这也正是本文研究所采用的网络类型。

(e) 反馈型局部互连网络

这种网络中，每个神经元只和其周围若干层的神经元发生互连关系，形成局部反馈，从整体上看是一种网格状结构，如 L. O. Chua 提出的细胞神经网络。这种网络结构特别适合于图像信息的加工和处理，在控制中的应用尚未见报道。

通过以上叙述，我们对神经网络的基本单元及其互连方式有了了解。各种神经网络的结构特性应基本上能按照以上模型框架进行概括，或在以上模型的基础上加以适当拓展即可描述。

2.4 静态多层前馈网络

这种网络一般是非动态的，即方程 (2-3) 中的 $H(s) = 1$ 。连

接矩阵 A 是如此构成的：各输出被按层分解，所以一层中的某神经元输入只接受前一层的神元输出（或者，在第一层中，只接受网络输入）。在这种网络中没有反馈。例如，在一个三层网络中，每一层包括 N 个神经元。我们将方程 (2-10) 中的网络矢量 X, Y, U 和 W 分成以下形式：

$$\begin{bmatrix} x^1(t) \\ x^2(t) \\ x^3(t) \end{bmatrix} = A \begin{bmatrix} y^1(t) \\ y^2(t) \\ y^3(t) \end{bmatrix} + B \begin{bmatrix} u^1(t) \\ u^2(t) \\ u^3(t) \end{bmatrix} + \begin{bmatrix} w^1 \\ w^2 \\ w^3 \end{bmatrix} \quad (2-13)$$

这里的上标表示网络中的相应层数。网络的 A 矩阵和 B 矩阵的结构如下所示：

$$A = \begin{bmatrix} 0_{NN} & 0_{NN} & 0_{NN} \\ A^2 & 0_{NN} & 0_{NN} \\ 0_{NN} & A^3 & 0_{NN} \end{bmatrix}, B = \begin{bmatrix} B^1 & 0_{NM} & 0_{NM} \\ 0_{NM} & 0_{NM} & 0_{NM} \\ 0_{NM} & 0_{NM} & 0_{NM} \end{bmatrix} \quad (2-14)$$

这里 0_{NN} 是 $N \times N$ 的零矩阵，而 0_{NM} 是 $N \times M$ 的零矩阵， A^2 和 A^3 为 $N \times N$ 的权重矩阵，而 B^1 为 $N \times M$ 的权重矩阵。对于第一层，我们得到：

$$\begin{aligned} x^1(t) &= B^1 u^1(t) + w^1 \\ y^1(t) &= g(x^1(t)) \end{aligned} \quad (2-15)$$

对于第二、第三层，我们得到：

$$\begin{aligned} x^l(t) &= A^l y^{l-1}(t) + w^l \\ y^l(t) &= g(x^l(t)) \end{aligned} \quad (2-16)$$

式中 $l=2, 3$ 。

当运用表 2-1 中的不同非线性函数 $g(x)$ 时，可以得到不同的特性。应用 S 型函数是较传统的做法 (Rumelhart et al. 1986)。

多层网络的这种连接使输入和输出之间更复杂的非线性映射成为可能。这种性能可以被用作分类器，或者用以表示变量之间的复杂的非线性关系。这样，将其用于各类非线性控制和学习控