《现代应用数学方法》丛书 (6)

科学出版社

21

数学天元基金

神经网络系统理论及其应用

沈世镒著

《现代应用数学方法》丛书（6）

神经网络系统理论及其应用

沈世镒 著

科学出版社

内 容 简 介

本书以前馈网络、Hopfield神经网络系统与玻尔兹曼机模型为基础，系统介绍神经网络系统的理论及其应用．本书对以上模型进行了较系统与严格的讨论，对它们的模型、学习算法、性能指标与命题给以定量的数学描述、论证与计算．关于神经网络系统的应用，除了传统的优化与识别问题外，本书重点介绍了统计分析与股市分析中的神经网络计算问题．

本书可供从事数学、统计、电子、智能计算机与计算机应用的本科生、研究生、教师及有关的工程技术人员学习参考．

图书在版编目（CIP）数据

神经网络系统理论及其应用/沈世镒著．-北京：科学出版社，1998.
（《现代应用数学方法》丛书；6/方开泰主编）
ISBN 7-03-006695-2
Ⅰ．神…　Ⅱ．沈…　Ⅲ．神经网络　Ⅳ．TP18
中国版本图书馆CIP数据核字（98）第09095号

科学出版社出版
北京东黄城根北街16号
邮政编码：100717

中国科学院印刷厂印刷

新华书店北京发行所发行　各地新华书店经售

*

1998年11月第一版　开本：850×1168　1/32
1998年11月第一次印刷　印张：3 3/4
印数：1—1 800　字数：89 000

定价：7.50元

（如有印装质量问题，我社负责调换〈科印〉）

《现代应用数学方法》丛书

总　　序

应用数学的发展与自然科学和社会科学广泛的交叉和渗透密切相关．一方面，它为形形色色的物理、化学、生物、社会等现象提供描述和分析的数学工具．另一方面，这些实际问题的解决又为数学学科的发展提供了动力和永不枯竭的源泉．许多成功的应用数学方法，如解非线性方程的牛顿-高斯法、曲线拟合的最小二乘法、线性规划的单纯形法等，成了当今应用数学工作者手中不可缺少的工具．它们之所以有如此强大的生命力，原因在于方法本身有坚实的理论基础，同时又有鲜明的应用背景，能用于不同的领域．因此，成功的应用数学方法是理论联系实际的桥梁和纽带．

我国的数学要达到世界先进水平，要对人类有所贡献，一个重要的方面是要有一批独创的应用数学方法．《现代应用数学方法》丛书的出版，希望能为鼓励和促进我国的数学工作者创造或介绍更多的现代应用数学方法增加一个舞台．

这套丛书的宗旨是介绍现代应用数学方法．这些方法应该是目前世界上最先进的，或者是我国独创的，或者是国外已经普遍使用但国内知之甚少的方法．丛书着重阐明所介绍方法的应用背景和思想，避开深奥的数学论证，力求深入浅出、图文并茂．有数值及应用性的例子，使读者易于理解和使用，丛书要求短小精练，突出新的方法，不求齐全，一般 10 万字左右．书末所附的文献将指出方法的理论背景以及最近的进展，以便读者进一步深入研究．

本丛书的出版得到国家“天元”项目的资助，得到科学出版社的大力支持，是全体编委努力的成果．我们要特别感谢许多作者在百忙中为丛书撰写文稿，付出了辛勤的劳动．我们希望这套

丛书的出版对我国应用数学的发展起到促进作用，衷心地希望丛书成为广大读者的良师益友.

胡国定（南开大学）

方开泰（中国科学院应用数学所）

引　　言

神经网络系统（以下简记为NNS）理论研究的重要意义已为许多科学家所承认，不少人把它看作未来智能计算机发展的一个主流方向．但是从神经网络系统发展的40年历史来看，它的发展是不平衡的，曾多次出现起伏．自80年代以来，由于Hopfield模型的提出，神经网络系统理论又引起许多科学工作者的兴趣，形成近代科学发展的一个热点．至80年代中期与90年代初期，神经网络系统理论的研究工作发展迅速，出现了许多模型与应用问题，也提出了一系列迫切需要解决的问题，形成多学科的交叉研究局面．但另一方面，目前人工神经网络系统的功能与人脑的智能相比较，差距还是很大，其主要原因还是硬件实现的限制，按目前集成电路的集成水平，可实现1000个神经元的互联网络（有10^6条网络连线），这与人体具有10^{10}—10^{11}个神经元的规模相比有很大的差距．另外，在神经网络的理论研究方面，还有许多问题没有解决，如人体神经系统的网络结构模型与信息处理功能的数据描述等．同时对现有的人工神经网络系统模型的学习规则与性能指标中的许多问题也未很好解决，如Hopfield神经网络系统，多层感知器模型中的若干问题等．这些问题影响了神经网络系统理论的进一步发展，因此需各方面的工作者解决．

本书的目的是以前馈网络、Hopfield神经网络系统（以下简记HNNS）与玻尔兹曼（Boltzmann）机模型为基础，系统介绍神经网络系统的理论及其应用．本书的主要特点是对以上模型进行较系统与严格的讨论，对它们的模型、学习算法、性能指标与命题给以定量的数学描述、论证与计算，为进一步深入研究神经网络系统理论打好基础．关于神经网络系统的应用问题，除了传统的优化与识别问题外，我们重点介绍了统计分析与股市分析中的神

经网络计算问题.

书中许多结果为我们近期所得. 统计分析中的神经网络计算问题中的基本思路与结果是作者与方开泰教授合作的成果. 书中大部分算法我们均已进行模拟计算，并有很好的效果. 具有工科大专以上的读者都可读懂本书中的大部分内容. 部分内容需有一定的线性代数、概率论、随机过程与统计方面的知识. 我们将在各章中作具体说明.

本书的出版获“数学天元出版基金”资助，并获国家自然科学基金项目：批准号，19671048；国家自然科学基金重大项目：“金融数学，金融工程与金融管理”；教育部博士点基金支持，特此感谢.

本书在写作过程中，曾获章祥荪、方开泰、程侃教授的帮助，其中神经网络在非线性回归计算中的应用部分是本人与方开泰教授合作的结果，特此予以说明并表示感谢. 另外，对罗跃虎教授帮我校对书稿表示感谢.

沈世镒

1997 年 10 月

有关记号

X,Y,Z,U,V:集合记号,状态空间.

A,B,C,D:集合记号,或子集合记号;矩阵记号.

$A\cup B, A\cap B$(或 $A\cdot B$),$A\backslash B$:分别为集合 A 与 B 的并、交、差集.

$A\times B$:集合 A 与 B 的积,这时 $A\times B=\{(a,b):a\in A, b\in B\}$.

A^c:集合 A 的余集,这时 $A^c=X\backslash A$.

ϕ:空集记号.

$F_2=\{0,1\}$:二元域.

$\{0,1\},\{\pm 1\}=\{-1,+1\}$:二元集合(或二元空间).

$R=(-\infty,\infty)$:实数空间.

$Z=\{\cdots,-1,0,1,2,\cdots\}$:全体整数集合.

$Z_0=\{0,1,2,\cdots\}$:全体非负整数集合.

$Z_+=\{1,2,3,\cdots\}$:全体自然数集合.

$N=\{1,2,\cdots,n\}, K=\{1,2,\cdots,k\}$:非负整数集合.

英文大写字母有时也表示算子或矩阵,如 $T(\cdot)$ 为算子;$M=[m_{i,j}]_{n\times n}$表示 $n\times n$ 矩阵.

$\alpha=\alpha(k)=\{i_1,i_2,\cdots,i_k\}$:集合 N 中的子集,其中 $1\leqslant i_1<i_2<\cdots<i_k\leqslant n$.

$Z_q=\{0,1,\cdots,q-1\}$:q-元环或域.

$GF(q)$:q-元域.

x,y,z,u,v:集合 X,Y,Z,U,V 中的元素.

$X^n=\prod_{i=1}^{n}X_i\ (X_i=X)$:$X$ 的 n 维乘积空间.

Y^n,Z^n,U^n,V^n:与 X^n 类似定义.

$x^n=(x_1,x_2,\cdots,x_n)$:X^n 空间中的向量,其中 $x_i\in X$.

y^n,z^n,u^n,w^n:与 x^n 类似定义.

$\phi^n = (\overbrace{0,0,\cdots,0}^{n\text{个}})$：$n$ 维零向量.

$1^n = (\overbrace{1,1,\cdots,1}^{n\text{个}})$：$n$ 维幺向量.

$x(\alpha) = (x_i : i \in \alpha) = (x_{i_1}, x_{i_1}, \cdots, x_{i_\alpha})$：$x^n$ 的子向量.

σ^n：数 σ 的 n 次幂，当 σ 为固定常数时.

$T^n = \overbrace{T(T(\cdots T(}^{n\text{个}}\ \overbrace{)\cdots\ )\ }^{n\text{个}})$：算子 T 的 n 次幂.

$W = [\omega_{ij}]_{n\times m}$：$n \times m$ 矩阵，或 n 行，m 列矩阵.

$W^x = [\omega_{ji}]_{m\times n}$：$W$ 的转置矩阵.

$I^n = [\delta_{i,j}]_{n\times n}$：$n$ 阶幺矩阵，其中 $\delta_{i,j} = \begin{cases} 1, \text{如果 } i = j, \\ 0, \text{否则}. \end{cases}$

$w_{i\cdot} = (w_{i,1}, w_{i,2}, \cdots, w_{i,n})$：$W$ 矩阵的行向量.

$w_{\cdot j} = (w_{1,j}, w_{2,j}, \cdots, w_{n,j})$：$W$ 矩阵的列向量.

$(x^n, y^n) = \sum_{i=1}^{n} x_i y_i$：向量 x^n 与 y^n 的内积.

$\| Q \|$：如果 Q 是矩阵 W 时，$\| W \|$ 为 W 矩阵的行列式.
如果 Q 是集合 A 时，$\| A \|$ 为集合 A 的元素个数.
如果 Q 是向量 x^n 时，$\| x^n \|$ 为向量 x^n 的模.

$|a|$：数 a 的绝对值.

$d_H(x^n, y^n) = \sum_{i=1}^{n} d_H(x_i, y_i)$：离散向量 x^n 与 y^n 的 Hamming 距离.
这时 $d_H(x_i, y_i) = 0$ 或 1，分别在 x_i 与 y_i 相等或不等时.

$d_H(x^n) = \sum_{i=1}^{n} d_H(x_i)$：离散向量 x^n 的 Hamming 势.
这时 $d_H(x_i)$ 等于 0 或 1，分别在 x_i 为零或不为零时.

$d(x^n, y^n) = \sqrt{\sum_{i=1}^{n} (x_i - y_i)^2}$：向量 x^n 与 y^n 的欧几里得距离.

$d(x^n) = \sqrt{\sum_{i=1}^{n} (x_i)^2}$：向量 x^n 的欧几里得势(或长度).

$\mathrm{Sgn}(z)$：z 的符号函数，这时 $\mathrm{Sgn}(z) = 1$ 或 -1，分别在 $z \geq 0$ 或 < 0

时.

$\exp(z)$, $\exp_2(z)$：指数函数，分别以 e 与 2 为底.

$\ln(z)$, $\log_2(z)$：对数函数，分别以 e 与 2 为底.

ξ,η,ζ（或 x^*,y^*,z^*）：分别取值于 X,Y,Z 空间的随机变量.

(ξ,η)：取值于 $X\times Y$ 空间的二维随机向量.

$\xi^n=(\xi_1,\xi_2,\cdots,\xi_n)$：取值于 X^n 的随机向量.

$p(A)$：事件 A 的概率.

$p(z)$：$\zeta=z$ 的概率或概率分布密度.

$P_r\{\xi\in A\}$：$\xi\in A$ 的概率.

$P_r\{\xi\in A|\eta\in B\}$：$\xi\in A$ 关于 $\eta\in B$ 的条件概率.

$p(A|B)$：事件 A 关于事件 B 的条件概率.

$p(x|y)$：$\xi=x$ 关于 $\eta=y$ 的条件概率或条件概率分布密度.

$\bar{p}=(p_1,p_2,\cdots,p_n)$：离散随机变量的概率分布，其中 $p_i\geqslant 0$，

$$\sum_{i=1}^{n} p_i=1.$$

$p(\cdot)=\{p(x),x\in X\}$：连续随机变量的概率分布密度，其中

$$p(x)\geqslant 0,\ \int_X p(x)dx=1.$$

$E\{\zeta\}$：随机变量 ζ 的数学期望（或均值）.

$\mathrm{Var}\{\zeta\}$：随机变量 ζ 的方差.

$\mathrm{Cov}(\xi,\eta)=E\{(\xi-E\xi)(\eta-E\eta)\}$：随机变量 (ξ,η) 的协方差.

$\mathrm{Conv}(D)$：集合 D 的凸闭包.

$S(D)$：集合 D 的凸闭包 $\mathrm{Conv}(D)$ 的边界面.

$L(D)$：过集合 D 中所有点的超平面.

$LL(D)$：由集合 D 决定的线性空间.

$G=\{A,L\}$：图，其中 A 为图 G 的点集，L 为图 G 的弧集.

$H(\varepsilon)=-\varepsilon\log_2\varepsilon-(1-\varepsilon)\log_2(1-\varepsilon)$：二元 Shannon 熵.

$H(p)=\sum_{i=1}^{n} p_i\log_2 p_i$：离散型随机变量（或概率分布）的 Shannon 熵.

$H(p)=\int_R p(x)\log_2 p(x)dx$：连续型随机变量（或概率分布）的 Shannon 熵.

$D(p;q)=\sum_{i=1}^{n} q_i \log_2 \frac{q_i}{p_i}$：概率分布 p 对 q 的互熵(Kullback 熵).

$I(\xi;\eta)=\sum_{i=1}^{n}\sum_{j=1}^{n} p_{ij}\log_2 \frac{p_{ij}}{p_i q_i}$：随机变量 ξ 与 η 的交互信息，其中 $p_{ij}=P_r\{\xi=i,\eta=j\}$，$p_i=P_r\{\xi=i\}$，$q_j=P_r\{\eta=j\}$.

目　录

第一章 概 论

§1.1 神经网络系统理论的研究目标与发展简史

1. 神经网络系统理论的研究目标

智能计算机的发展是当今人类社会所面临的一项迫切而又重大的科技问题. 智能计算机发展的一个突出点是关于人脑功能的模拟问题. 计算机的发展为信息技术的革命带来决定性的影响，它的存储能力、计算速度与其他若干逻辑功能等都已大大超过了人脑的能力，但是还有许多人脑功能为数值计算机所远远不及，尤其是关于人的若干智能问题，如人的识别与判决能力，数值计算机还远远不能达到. 因此，智能计算机的发展还需有一个相当漫长的过程. 神经网络系统理论就是以人脑的智能功能为研究对象且以人体神经细胞的信息处理方法为背景的智能计算机与智能计算理论. 概括人体神经细胞的构造与信息处理功能有以下几个特点：

(1) 每一个神经细胞是一个简单的信息处理单元，它可由自身与外部条件决定它的状态，形成一定的输入、输出规则，人们称之为激励规则. 由于生物学与医学的发展，对单细胞的构造与功能特征已有较清楚的了解，这为我们研究神经网络系统理论提供了最基本的模型背景，对此我们在下一节中还要详细描述.

(2) 神经细胞之间按一定的方式相互连接，构成神经网络系统，且按一定的规则进行信息传递与存储. 但从医学生理角度讲，人体神经细胞之间相互连接的方式与它们的数据处理功能还不清楚. 在人工神经网络系统理论中，有许多假设性的构造模型.

(3) 在人体的发育生长过程中，神经网络系统可按已发生的

事件积累经验，从而不断修改该系统的网络连接规则与存储数据．这种网络连接规则与数据存储方式有一定的稳定性与可塑性，它可以保证人类知识经验的积累与修正．这种修改方式人们称之为学习规则．人类（或其他生物）通过学习过程最终能达到正确计算（或识别、判决等）的目标，这种功能被称为学习与训练功能．这种学习与训练功能是人类具有高度智能化的一个主要特征．

神经网络系统理论研究的意义就在于它以模拟人体神经系统为自己的研究目标，并具有上述人体神经系统的基本特征．因此，它的研究工作为许多科学家所承认，不少人把它看作未来智能计算机发展的一个主流方向．本书的主要目的是讨论神经网络系统运算算法的理论基础与它们的各种应用．这种运算算法可在现有普通计算机上通过硬件与软件编程来实现，也可在未来神经网络计算机上快速实现，最终实现智能计算机与智能计算的目标．

2. 神经网络系统理论研究的发展简史

神经网络系统理论的发展历史是不平衡的．自 1943 年心理学家 McCulloch 与数学家 Pitts 提出神经元生物学模型（简称 MP-模型）以来，至今已有 50 多年的发展历史了．在这 50 多年的发展历史中，它的发展大体上可分为以下几个阶段．

自 1943 年 MP-模型产生起，至 60 年代止，这一段时间可称为神经网络系统理论的初期阶段，这时的主要特点是多种网络模型的产生与学习算法的确定，如 1944 年由 Hebb 提出的 Hebb 学习规则，这个规则至今仍是神经网络系统理论学习算法的一个基本规则．1957 年由 Rosenblatt 提出的感知器（perceptron）模型，1962 年由 Widrow 提出的自适应线性元件（adaline）等，这些模型与算法无疑丰富了神经网络系统理论．

60 年代末至 70 年代，神经网络系统理论的发展处于一个低潮时期，造成这个情况的原因是神经网络系统理论的发展出现了本质上的困难，即电子线路交叉极限的困难（对 n 个神经元就有 n^2 条连线）．这在当时条件下，对神经元的数量 n 的大小受到极大的

限制，因此它不可能去完成高度智能化的计算任务．同时，神经网络系统理论本身也有许多不完备之处．因此，神经网络理论与应用研究工作进展缓慢．另一方面，这一时期正是数值计算机发展的全盛时期，无论在硬件、软件、技术应用与商品市场方面都取得了突飞猛进的进展，因此使大批有才能的科学家的注意力都在发展数值机上了．

80年代，关于智能计算机的发展道路问题日趋迫切地提到讨论日程上来．因为计算机的集成度已日趋极限状态，但数值计算机的智能水平与人脑相比较仍有极大的差距，因此，人们就需要从新的角度来思考智能计算机的发展道路问题，这样神经网络系统理论就应运而生．1982年美国加州物理学家J. J. Hopfield提出了Hopfield神经网络系统（简称HNNS）模型，提出了能量函数，稳定性等概念，并在计算TSP（售货员送货路线）问题上取得进展．因此，80年代后期至90年代初，神经网络系统理论形成一个发展热点，多种模型、算法与应用问题被提出，许多国家纷纷增加经费，完成了许多有意义的工作．目前，神经网络系统理论与技术发展大体分以下三个方面进行．

首先在硬件技术方面，一些先进工业国家，如美国与日本均已实现了规模有1000个神经元的网络系统，因此该系统有极高的运算速度，并在股票数据分析中得到应用．另外，为克服电子线路交叉极限的困难，世界各国都在研究除了电子元件以外的神经网络系统，如光电子元件与生物元件等．

在神经网络系统理论研究方面，主要进展有如：Boltzmann机理论的研究，细胞网络的提出，性能指标的分析等．

神经网络系统理论的应用研究主要是在模式识别（包括语音与图像识别）、经济管理、优化控制等方面；与数学、统计中的多个学科分支发生联系，有如线性与非线性规划问题、数值逼近、统计计算等．另外，在其他信息处理问题中也有多种应用，如数据压缩、编码、密码与股市分析等领域．其内容十分丰富．