



国际电气工程先进技术译丛

CRC Press  
Taylor & Francis Group

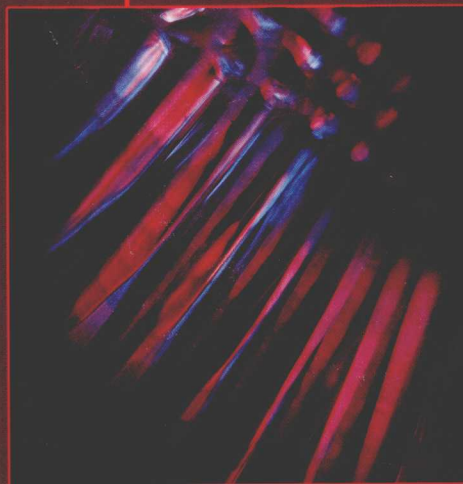
# 神经网络在应用科学和工程中的应用——从基本原理到复杂的模式识别

**Neural Networks for Applied  
Sciences and Engineering—  
From Fundamentals to Complex  
Pattern Recognition**

(美) Sandhya Samarasinghe 著  
史晓霞 陈一民 李军治 等译



机械工业出版社  
CHINA MACHINE PRESS



国际电气工程先进技术译丛

# 神经网络在应用科学和工程中的应用

——从基本原理到复杂的模式识别

(美) Sandhya Samarasinghe 著

史晓霞 陈一民 李军略 等译



机械工业出版社

社址：北京市西城区百万庄大街24号  
邮编：100037  
电话：(010) 68330660  
电传：(010) 68330660  
网址：http://www.cmpbook.com  
电子邮箱：mailto:custserv@cmpbook.com

本书是研究神经网络对科学数据进行模式识别的一种探索,主要介绍了神经网络的基本概念,并以图示的形式用大量的实例和个案研究结合计算机仿真对神经网络用于模式识别的各种方法进行阐述与对比。本书所涉及的学科领域包括生物学、经济学、应用科学、工程、计算和商业等,研究了神经网络在线性和非线性预报、分类、聚类和预测方面的应用,并对模型开发的所有阶段和结果进行了阐述,包括数据预处理、数据维数约简、输入选择、模型开发和验证、模型不确定性评估以及对输入、误差和模型参数的灵敏度分析。本书内容清晰明了,并结合了大量实例使得全书更加容易理解。

本书适合作为高等院校相关专业的大学生、研究生的教材及有关研究人员的参考用书。

Neural Networks for Applied Sciences and Engineering—From Fundamentals to Complex Pattern Recognition, ISBN: 978-0-8493-3375-0, edited by Sandhya Samarasinghe, published by CRC Press, part of Taylor & Francis Group LCC.

Copyright© 2007 by CRC Press.

Authorized translation from the English language edition published by CRC Press, part of Taylor & Francis Group LCC.

All Rights Reserved.

本书中文简体字版由机械工业出版社出版,未经出版者书面允许,本书的任何部分不得以任何方式复制或抄袭。

版权所有,翻印必究。

本书版权登记号:图字01-2007-5495号

## 图书在版编目(CIP)数据

神经网络在应用科学和工程中的应用:从基本原理到复杂的模式识别/(美)萨马拉辛荷(Samarasinghe. s.)著;史晓霞等译. —北京:机械工业出版社,2009.7

(国际电气工程先进技术译丛)

书名原文:Neural Networks for Applied Sciences and Engineering—From Fundamentals to Complex Pattern Recognition

ISBN 978-7-111-27585-5

I. 神… II. ①萨…②史… III. 神经网络—应用—研究 IV. TP183

中国版本图书馆CIP数据核字(2009)第115777号

机械工业出版社(北京市百万庄大街22号 邮政编码100037)

策划编辑:靳平 责任编辑:任鑫

版式设计:霍永明 责任校对:陈延翔

封面设计:马精明 责任印制:洪汉军

三河市宏达印刷有限公司印刷

2010年1月第1版第1次印刷

169mm×239mm·28.75印张561千字

标准书号:ISBN 978-7-111-27585-5

定价:88.00元

凡购本书,如有缺页、倒页、脱页,由本社发行部调换

电话服务

网络服务

社服务中心:(010) 88361066

销售一部:(010) 68326294

销售二部:(010) 88379649

读者服务部:(010) 68993821

门户网:<http://www.cmpbook.com>

教材网:<http://www.cmpedu.com>

封面无防伪标均为盗版

## 译者序

神经网络模式识别是一门崭新的学科，近年来受到普遍的关注和重视，它的应用已经成为了神经网络理论应用最成功的一个。与传统的模式识别相比，神经网络模式识别方法表现出来较强的自组织学习能力、容错性、鲁棒性以及非线性处理等优点，使得它在各个领域得到了广泛的应用。

本书图文并茂，实例众多，所举的实例针对性强，分析透彻，特别突出理论内容的实际应用。本书首先给出了神经网络的大量基础概念和基本原理，并通过不同学科领域（包括生物学、经济学、应用科学、工程、计算和商业等）的实例和个案研究，用神经网络软件与实践相结合的方式给出了大量的图例说明。这样既体现了神经网络的理论性和实践性，也开阔了研究者的视野。

本书涵盖了神经网络在建模方面的许多重要领域的应用，可作为自学研究者或者相关科学研究者的参考用书。本书还可为应用科学家和工程师使用，也可作为这些领域学生的教科书。

本书的翻译与出版，得到许多同仁、朋友的真诚支持和帮助。本书的第1~5、7章由史晓霞翻译，第6章由史晓霞、李军治翻译，第8、9章由陈一民、樊清翻译，全书由史晓霞统稿。

在本书的翻译过程中，得到了许多同事的帮助与支持，特别是庄俊华、张俊红、魏东、张立权、胡玉玲、谭志、刘静纨、李锐、张亚峰、范红霞、沈倩倩、张世杰、赵亮、王健、马振宇等，在此表示衷心的感谢！

本书翻译稿虽经多次审阅修正，但不当与疏漏之处在所难免，诚望广大读者批评指正。

译者

## 前 言

本书是通过神经网络对科学数据进行模式识别的一种探索。本书最大的一个亮点是通篇贯穿了对神经网络概念的大量可视化描述。初衷是既要更广阔的背景去满足研究人员和学生的需要，又要将神经网络融入多学科的科学背景。过去的七年里，我一直在新西兰的林肯大学从事包括生物学、经济学、应用科学、工程、计算和商业等不同背景的神经网络研究生教学工作。我通过把神经网络和其内部细节变得明晰的思路展开介绍，从而使学生们在学习上树立信心。这种可视化的描述对于理解困难的数学概念而言是一种很有价值的工具，所以本书也是这些研究成果以及本人对于神经网络的探究兴趣的反映。

在应用背景上，我想提供一种合理的理论背景。我的经验证明，使用神经网络软件与实践应用相结合的学习方法会得到很好的效果。另外，在本书中通过实际辅导环节作为理论的补充也是很成功的。

本书的读者对象是具有一些数学和基础计算背景的从事应用领域研究方面的高年级学生和研究生。书中不仅以简单表达与可视化辅助相结合，将一个网络拆分成为可能，用以理解数学概念及相关推导，而且介绍被许多人认为是“黑箱”的神经网络内部运作的重要性。

第1章开始以神经网络在科学数据分析中扮演的角色进行了介绍性的讨论，并提出了本书的具体安排。许多科学家都对用传统的统计学方法去判定神经网络的优越性感兴趣，所以书中对与神经网络相关的统计学方法给出了详细的论述。第2章阐述了两种方法在线性数据分析方面的等效性，然后开始建立基本神经网络概念的坚实基础，逐渐地进行更深入的理解。

第3章到第5章论述了使用多层神经网络分析的非线性数据，以及多层神经网络对非线性模式的识别。多层神经网络是一种带有非线性处理能力的神经元层的网络。本书提出了这些网络的广义收敛性。如果与鲁棒性和反复实验需要的相关局限性进行处理，则它们在系统建模方面的潜力和有效性就会被提高。这几章对基于统计学方法的神经网络在非线性的建模方面的优势进行了阐述，并对网络的非线性处理、网络训练和网络结构的优化进行了详细的论述。为了使其容易理解，这几章提供了具体的实例和个案研究。这些资料适合独立学习和

课堂讲解。

第6章对数据测定和预处理进行了讨论,包括统计学方法的有效数目,一些可以用普通的统计学程序获得。本章的目的是为有效模型开发选取相关的和独立的输入,同时这些理论可以在解决实际问题时与统计学软件一起使用。

第7章讨论神经网络的不确定性评估,多采用统计学方法,即使用严格的统计学理论来检验神经网络。尽管神经网络是很强的非线性过程,但评价它们的鲁棒性手段是很局限的。本章为了严格评估神经网络不确定性,涉及了网络参数、误差和灵敏性,将神经网络引入贝叶斯统计学的内容。本章提供的这些资料需要对简单的、共同的和条件概率有一定的了解,正如在第3章到第5章所讲述的神经网络的概念。这章提出的不确定性评估对在神经网络模型中获得置信度是很有价值的,并可以将其用于决策的制定。

以我的经验,学生尤其感兴趣的是自组织映射——自主网络用于发现未知聚类和多维数据的关系。第8章<sup>①</sup>以循序渐进的方式提出这些内容,以强调这些重要概念,可用作教学资料和自学。其中重要概念逐步提出,很多自主数据聚类的特点和一些统计学聚类方法的关系都用实例进行了阐述。本章涵盖的主题包括竞争学习和拓扑维持、一维和二维映射、映射训练和验证、映射质量评估、映射中的聚类构成,并以图示发展了自组织映射。

最后一章是用神经网络处理线性和非线性时间序列预测。它广泛地涵盖了用于短期和长期时间序列预测的循环网络概念和反馈网络概念,这些资料的大部分内容可同时用作讲课资料和自学资料。大部分实例个案研究都强调了所介绍的新概念,并阐述了线性神经网络和相关典型统计学方法的相似性,同时对非线性神经网络的优势进行了演示。

本书所提到的实例主要是基于神经网络的 *Mathematica* 发展,并且 *Mathematica* 机器学习框架——两个 *Mathematica* 附加程序和一种 *NeuroShell2*<sup>TM</sup> 商业软件,在万维网上可获得很多商业的和免费的神经网络程序作为本书素材的补充。

我希望本书中的资料和基本原理可以帮助读者应用和理解本书没有涉及的各种网络,并且能在快速成长的领域有其他新的发展。我的经验已经表明,本书可以帮助很多不同方向的研究者对神经网络进行学习和将其应用到他们各自的领域。希望读者也能发现他们自己的方向。

### 怎样使用这本书

本书覆盖了用神经网络在建模方面的许多重要领域,可作为集中研究的教

① 原文为第6章。——译者注

科书或者自学研究者的参考书。本书可为应用科学家和工程师所用,也可作为这些领域学生的教科书。这些内容提要可以跨越两个学期:第2章到第5章覆盖第一学期的内容,第6章到第9章用在第二学期。尽管从第2章到第7章的素材是完全连续的而且具有逻辑性,第8章依赖于第2章提出的内容,最后一章与第3章到第7章的内容有很大的关系。在过去我已经使用了 NeuroShell2 (美国公司, Ward 系统)、Neural Connection<sup>TM</sup> (美国公司, SPSS) 和 NeuroSolutions<sup>TM</sup> (Neuro Dimension 公司) 软件,每种都适合于班级指导老师和独立研究者使用。

在多学科学习者中,他们所期望成果是很广泛的,一些人希望在他们的研究领域使用神经网络作为工具,所以对于他们而言,用户界面友好的软件是很必要的。对于那些将神经网络内容用于实验的研究者,则 Mathematica (Wolfram Research 公司)、MATLAB<sup>®</sup> (美国 MathWorks 公司) 和 C++ 等编程环境是有用的。

## 致 谢

首先，我必须向我的丈夫、新西兰林肯大学计算机建模和仿真系统的教授 Don Kulasiri 表达衷心的感谢，是他鼓励我去从事神经网络领域的研究，并鼓舞我写这本书。不仅对我写这本书，而且在我作为学者去发展我的事业方面，他在精神上、专业上和个人的支持上都是无价的。

我在美国的新泽西州的普林斯顿大学的普林斯顿学院时写了本书的方案和前面几章。我要感谢 Simon Levin 教授，生物学的 Moffette 教授和普林斯顿大学的生物复杂性研究中心的主任，是他们邀请我在 2004 年去做一名访问学者。

我的研究生 Erica Wang，准备了所有本书中用到的图，我对她所做的贡献表示诚挚的感谢。我也要对在多方面对我提供帮助的神经网络的研究者表示感谢。特别感谢法国 Pierre et Marie Curie 大学的 Filipe Aires 博士和瑞典 Chalmers 科技大学的 Jonas Sjoberg 教授（Mathematica 的神经网络的开拓者）。

最后，感谢我父母的鼓励和支持，尤其是我的父亲，他在我的教育过程中长期的奉献是无价的。

## 作者简介

**Sandhya Samarasinghe** 在俄罗斯的 Lumumba 大学获得机械工程的理学硕士 (Hons)，在美国的 Virginia Tech 获得了工程学硕士和博士学位。目前她是新西兰林肯大学自然工程系的高级讲师，并成为先进计算解决方案中心的奠基人之一。她的研究包括神经网络、统计学、软计算和人工智能、统计学方法、计算机视觉的利用，以及用于解决工程学、生物学、经济学、环境和自然系统及应用科学的实际问题的复杂系统建模。**Samarasinghe** 博士参与了很多领域的科学与工业项目，并且在那些领域的国际期刊和国际会议上发表了很多文章。她在神经网络方面的研究重点包括理论理解与进展和实际应用。她参与了许多国家和国际基金组织的资助申请评审、杂志文章的评审、组织国际会议，并且是编委会成员。她还曾经在美国斯坦福大学、普林斯顿大学及澳大利亚国家科学和工业研究组织 (CSIRO) 做过巡讲。

# 目 录

译者序

前言

致谢

作者简介

## 第 1 章 从数据到模型：理解生物学、生态学和自然系统的复杂性和挑战 ..... 1

|                |   |
|----------------|---|
| 1.1 概述 .....   | 1 |
| 1.2 本书安排 ..... | 3 |
| 参考文献 .....     | 5 |

## 第 2 章 神经网络基础和线性数据分析模型 ..... 9

|                                                     |    |
|-----------------------------------------------------|----|
| 2.1 概述 .....                                        | 9  |
| 2.2 神经网络及其能力 .....                                  | 9  |
| 2.3 生物学的启示 .....                                    | 13 |
| 2.4 神经元信息处理的建模 .....                                | 15 |
| 2.5 神经元模型和学习策略 .....                                | 15 |
| 2.5.1 作为一个简单分类器的阈值神经元 .....                         | 16 |
| 2.5.2 神经元和神经集合的学习模型 .....                           | 19 |
| 2.5.2.1 Hebbian 学习 .....                            | 19 |
| 2.5.2.2 无监督学习或竞争学习 .....                            | 20 |
| 2.5.2.3 有监督学习 .....                                 | 21 |
| 2.5.3 作为分类器的有监督学习的感知器 .....                         | 21 |
| 2.5.3.1 感知器学习算法 .....                               | 22 |
| 2.5.3.2 基于大量现实数据集的感知器实例：根据测定的成长年轮<br>直径辨识鱼的起源 ..... | 28 |
| 2.5.3.3 统计学中带有线性判别函数分析的感知器比较 .....                  | 30 |
| 2.5.3.4 多种分类中的多输出感知器 .....                          | 31 |
| 2.5.3.5 使用感知器的高维分类 .....                            | 35 |
| 2.5.3.6 感知器小结 .....                                 | 36 |
| 2.5.4 用于线性分类和预报的线性神经元 .....                         | 36 |
| 2.5.4.1 利用 delta 规则的学习 .....                        | 37 |
| 2.5.4.2 作为分类器的线性神经元 .....                           | 40 |
| 2.5.4.3 作为预报能力子集的线性神经元的分类属性 .....                   | 42 |
| 2.5.4.4 实例：作为预报器的线性神经元 .....                        | 43 |
| 2.5.4.5 线性预报的实例：预报一个家庭的热流 .....                     | 48 |

|                                             |           |
|---------------------------------------------|-----------|
| 2.5.4.6 线性神经元模型与线性回归的比较 .....               | 48        |
| 2.5.4.7 实例:多输入线性神经元模型——提高一个家庭的热流预报精确度 ..... | 49        |
| 2.5.4.8 一个多输入线性神经元与多重线性回归的比较 .....          | 49        |
| 2.5.4.9 多线性神经元模型 .....                      | 50        |
| 2.5.4.10 多重线性神经网络与正则相关性分析的比较 .....          | 50        |
| 2.5.4.11 线性神经元和线性网络小结 .....                 | 50        |
| 2.6 小结 .....                                | 51        |
| 习题 .....                                    | 51        |
| 参考文献 .....                                  | 52        |
| <b>第3章 用于非线性模式识别的神经网络 .....</b>             | <b>54</b> |
| 3.1 概述 .....                                | 54        |
| 3.2 非线性神经元 .....                            | 56        |
| 3.2.1 神经元激励函数 .....                         | 57        |
| 3.2.1.1 S形函数 .....                          | 57        |
| 3.2.1.2 高斯函数 .....                          | 58        |
| 3.2.2 实例:利用非线性神经元对人口增长建模 .....              | 59        |
| 3.2.3 非线性神经元与非线性回归分析的比较 .....               | 61        |
| 3.3 单输入多层非线性网络 .....                        | 61        |
| 3.3.1 用单一非线性隐含层神经元处理 .....                  | 61        |
| 3.3.2 实例:用多非线性神经元建立循环现象模型 .....             | 65        |
| 3.3.2.1 实例1:逼近一个方波 .....                    | 66        |
| 3.3.2.2 实例2:为物种的季节性迁移建立模型 .....             | 71        |
| 3.4 两输入的多层感知器网络 .....                       | 74        |
| 3.4.1 用非线性神经元处理二维输入 .....                   | 75        |
| 3.4.2 网络输出 .....                            | 78        |
| 3.4.3 实例:二维预报和分类 .....                      | 79        |
| 3.4.3.1 实例1:二维非线性函数逼近 .....                 | 79        |
| 3.4.3.2 实例2:二维非线性分类模型 .....                 | 80        |
| 3.5 用非线性多层感知器网络为多维数据建模 .....                | 83        |
| 3.6 小结 .....                                | 83        |
| 习题 .....                                    | 84        |
| 参考文献 .....                                  | 85        |
| <b>第4章 神经网络对非线性模式的学习 .....</b>              | <b>87</b> |
| 4.1 概述 .....                                | 87        |
| 4.2 非线性模式识别中网络的监督训练 .....                   | 87        |
| 4.3 梯度下降法和误差最小化 .....                       | 88        |
| 4.4 BP学习 .....                              | 89        |
| 4.4.1 实例:BP训练——手工计算 .....                   | 90        |

|         |                                             |     |
|---------|---------------------------------------------|-----|
| 4.4.1.1 | 关于输出神经元权值的误差梯度                              | 92  |
| 4.4.1.2 | 关于隐含神经元权值的误差梯度                              | 95  |
| 4.4.1.3 | 梯度下降法在 BP 学习中的应用                            | 98  |
| 4.4.1.4 | 批量学习                                        | 98  |
| 4.4.1.5 | 学习率和权值更新                                    | 100 |
| 4.4.1.6 | 遍历(在线)学习                                    | 103 |
| 4.4.1.7 | 动量法                                         | 103 |
| 4.4.2   | 实例: BP 学习的计算实验                              | 106 |
| 4.4.3   | 具有多隐含神经元的单输入单输出网络                           | 109 |
| 4.4.4   | 多输入、多隐含神经元和单输出网络                            | 109 |
| 4.4.5   | 多输入、多隐含神经元多输出网络                             | 110 |
| 4.4.6   | 实例: BP 学习个案学习——解决一个复杂的分类问题                  | 111 |
| 4.5     | delta-bar-delta 学习方法                        | 117 |
| 4.5.1   | 实例: 用 delta-bar-delta 的网络训练——手工计算           | 118 |
| 4.5.2   | 实例: 带有动量的 delta-bar-delta 方法——手工计算          | 120 |
| 4.5.3   | delta-bar-delta 的网络训练——计算机实验                | 121 |
| 4.5.4   | delta-bar-delta 方法和 BP 方法的比较                | 122 |
| 4.5.5   | 实例: 用 delta-bar-delta 方法的网络训练——一个案例研究       | 122 |
| 4.6     | 最速下降法                                       | 125 |
| 4.6.1   | 实例: 用最速下降法的网络训练——手工计算                       | 125 |
| 4.6.2   | 实例: 用最速下降法的网络训练——计算机实验                      | 126 |
| 4.7     | 误差最小和权值最优的二阶方法                              | 127 |
| 4.7.1   | QuickProp                                   | 127 |
| 4.7.1.1 | 实例: QuickProp 的网络训练——手工计算                   | 128 |
| 4.7.1.2 | 实例: QuickProp 的网络训练——计算机实验                  | 129 |
| 4.7.1.3 | QuickProp 与最速下降法、delta-bar-delta 和 BP 方法的比较 | 130 |
| 4.7.2   | 误差最小的二阶方法的基本概念                              | 131 |
| 4.7.3   | 高斯-牛顿方法                                     | 132 |
| 4.7.3.1 | 采用高斯-牛顿方法的网络训练——手工计算                        | 134 |
| 4.7.3.2 | 实例: 采用高斯-牛顿方法的网络训练——计算机实验                   | 136 |
| 4.7.4   | LM 方法                                       | 137 |
| 4.7.4.1 | 实例: 用 LM 方法的网络训练——手工计算                      | 138 |
| 4.7.4.2 | 用 LM 方法的网络训练——计算机实验                         | 139 |
| 4.7.5   | 一阶和二阶方法在使误差最小效果上的比较                         | 140 |
| 4.7.6   | 一阶和二阶学习方法收敛性的比较                             | 141 |
| 4.7.6.1 | BP 方法                                       | 142 |
| 4.7.6.2 | 最速下降法                                       | 143 |
| 4.7.6.3 | 高斯-牛顿方法                                     | 144 |
| 4.7.6.4 | LM 方法                                       | 145 |
| 4.8     | 小结                                          | 146 |

|                                 |            |
|---------------------------------|------------|
| 习题                              | 146        |
| 参考文献                            | 147        |
| <b>第5章 从数据中抽取可靠模式的神经网络模型的实现</b> | <b>148</b> |
| 5.1 概述                          | 148        |
| 5.2 偏差-方差折中                     | 149        |
| 5.3 改善神经网络的泛化                   | 149        |
| 5.3.1 及早停止法的说明                  | 150        |
| 5.3.1.1 任意初始权值的影响               | 154        |
| 5.3.1.2 训练好网络的权值结构              | 156        |
| 5.3.1.3 随机采样的影响                 | 157        |
| 5.3.1.4 模型复杂性的影响: 隐含神经元的数目      | 161        |
| 5.3.1.5 及早停止法小结                 | 163        |
| 5.3.2 正规化法                      | 164        |
| 5.4 通过修剪减小网络结构的复杂性              | 169        |
| 5.4.1 最优脑部损伤 (OBD)              | 170        |
| 5.4.2 基于网络灵敏度方差的网络修剪            | 175        |
| 5.4.2.1 修剪权值的方差无效性应用的阐述         | 177        |
| 5.4.2.2 基于灵敏度的方差无效性修剪隐含神经元      | 179        |
| 5.5 网络对权值扰动的鲁棒性                 | 181        |
| 5.6 小结                          | 184        |
| 习题                              | 185        |
| 参考文献                            | 185        |
| <b>第6章 数据探测、维数约简和特征提取</b>       | <b>187</b> |
| 6.1 概述                          | 187        |
| 6.1.1 实例: 关于相关输入数据的木头的热导率       | 188        |
| 6.2 数据可视化                       | 189        |
| 6.2.1 散点图和柱状图的相关性               | 189        |
| 6.2.2 并行可视化                     | 190        |
| 6.2.3 多维数据到二维平面的映射              | 190        |
| 6.3 变量间的相关性与协方差                 | 191        |
| 6.4 数据的规范化                      | 193        |
| 6.4.1 标准化                       | 193        |
| 6.4.2 简单值域调整                    | 194        |
| 6.4.3 白化——多维相关数据的规范化            | 194        |
| 6.5 选择相关输入                      | 197        |
| 6.5.1 用于变量选择的统计手段               | 198        |
| 6.5.1.1 偏相关                     | 198        |
| 6.5.1.2 多元回归与最优子集回归             | 199        |

|                                           |     |
|-------------------------------------------|-----|
| 6.6 维数约简与特征析取 .....                       | 200 |
| 6.6.1 多重共线性 .....                         | 200 |
| 6.6.2 主成分分析 .....                         | 200 |
| 6.6.3 偏最小二乘回归 .....                       | 204 |
| 6.7 异常点检测 .....                           | 205 |
| 6.8 噪声 .....                              | 206 |
| 6.9 个案研究: 阐述一个实际问题的输入选择和维数约简 .....        | 206 |
| 6.9.1 数据预处理和初步建模 .....                    | 207 |
| 6.9.2 基于 PCA 的神经网络建模 .....                | 210 |
| 6.9.3 隐含神经元对非基于 PCA 方法和基于 PCA 方法的影响 ..... | 213 |
| 6.9.4 案例研究小结 .....                        | 214 |
| 6.10 小结 .....                             | 214 |
| 习题 .....                                  | 215 |
| 参考文献 .....                                | 215 |
| <b>第 7 章 使用贝叶斯统计的神经网络模型的不确定性评估</b> .....  | 217 |
| 7.1 概述 .....                              | 217 |
| 7.2 利用贝叶斯统计评估权值的不确定性 .....                | 218 |
| 7.2.1 质量判据 .....                          | 219 |
| 7.2.2 结合贝叶斯统计估计权值不确定性 .....               | 220 |
| 7.2.3 多元输出的目标固有不确定性 .....                 | 223 |
| 7.2.4 权值的概率密度函数 .....                     | 224 |
| 7.2.5 用于阐述权值分布产生的实例 .....                 | 226 |
| 7.3 使用贝叶斯统计评估神经网络输出的不确定性 .....            | 230 |
| 7.3.1 阐述输出误差不确定性的实例 .....                 | 230 |
| 7.3.1.1 总的网络输出误差 .....                    | 231 |
| 7.3.1.2 误差相关性和协方差矩阵 .....                 | 231 |
| 7.3.1.3 误差协方差的统计分析 .....                  | 231 |
| 7.3.1.4 总输出误差分解为模型误差和固有噪声 .....           | 233 |
| 7.4 评估网络输出对输入的灵敏度 .....                   | 237 |
| 7.4.1 确定反馈网络里输入对输出影响的方法 .....             | 237 |
| 7.4.1.1 基于权值大小的方法 .....                   | 237 |
| 7.4.1.2 灵敏度分析 .....                       | 238 |
| 7.4.2 实例: 评估输入对输出影响的方法的比较 .....           | 238 |
| 7.4.3 灵敏度的不确定性 .....                      | 239 |
| 7.4.4 阐述网络对输入灵敏度的不确定性评估实例 .....           | 240 |
| 7.4.4.1 输入和输出的 PCA 分解 .....               | 240 |
| 7.4.4.2 基于 PCA 的神经网络回归 .....              | 244 |
| 7.4.4.3 神经网络灵敏度 .....                     | 246 |
| 7.4.4.4 输入灵敏度的不确定性 .....                  | 247 |

|            |                                          |            |
|------------|------------------------------------------|------------|
| 7.4.4.5    | PCA 调整的雅可比                               | 251        |
| 7.4.4.6    | 案例研究小结                                   | 254        |
| 7.5        | 小结                                       | 255        |
|            | 习题                                       | 255        |
|            | 参考文献                                     | 256        |
| <b>第8章</b> | <b>应用自组织映射的方法发现数据中的未知聚类</b>              | <b>258</b> |
| 8.1        | 概述                                       | 258        |
| 8.2        | 无监督网络的结构                                 | 259        |
| 8.3        | 无监督网络的学习过程                               | 260        |
| 8.4        | 竞争学习过程的实施                                | 260        |
| 8.4.1      | 基于神经元激励的获胜者选择                            | 261        |
| 8.4.2      | 基于对输入向量距离的获胜者选择                          | 261        |
| 8.4.3      | 竞争学习例子                                   | 263        |
| 8.4.3.1    | 回归学习和批量学习的比较                             | 264        |
| 8.4.3.2    | 对获胜者选择中所涉及的计算的说明                         | 264        |
| 8.4.3.3    | 网络训练                                     | 266        |
| 8.5        | 自组织特性映射                                  | 268        |
| 8.5.1      | 自组织映射网络中的学习过程                            | 268        |
| 8.5.1.1    | 邻域几何形状的选择                                | 269        |
| 8.5.1.2    | 自组织映射的训练                                 | 269        |
| 8.5.1.3    | 邻点强度                                     | 269        |
| 8.5.1.4    | 例子: 训练具有邻域特性的自组织网络                       | 270        |
| 8.5.1.5    | 邻接矩阵和从获胜者到邻点之间的距离                        | 272        |
| 8.5.1.6    | 随着迭代缩减邻域的尺寸                              | 274        |
| 8.5.1.7    | 学习率衰减                                    | 276        |
| 8.5.1.8    | 包含学习率和邻域衰减后的权值更新                         | 276        |
| 8.5.1.9    | 递归和批量训练与 K 平均值聚类的关系                      | 276        |
| 8.5.1.10   | 自组织映射训练中的两个阶段                            | 277        |
| 8.5.1.11   | 例子: 通过手工计算说明自组织映射学习                      | 277        |
| 8.5.1.12   | SOM 实例研究: 结合牛奶品质对牛奶场的奶牛群乳腺炎<br>状况进行判断    | 283        |
| 8.5.2      | 二维自组织映射的实例: 基于鱼鳞年轮的直径对加拿大和阿拉斯加<br>鲑鱼进行聚类 | 286        |
| 8.5.2.1    | 映射结构和初始化                                 | 286        |
| 8.5.2.2    | 映射训练                                     | 287        |
| 8.5.2.3    | U 矩阵                                     | 293        |
| 8.5.3      | 映射初始化                                    | 294        |
| 8.5.4      | 实例: 在多维数据上训练二维映射                         | 295        |
| 8.5.4.1    | 数据可视化                                    | 295        |

|              |                           |            |
|--------------|---------------------------|------------|
| 8.5.4.2      | 映射结构和训练                   | 296        |
| 8.5.4.3      | U 矩阵                      | 301        |
| 8.5.4.4      | 由映射捕获的输入概率密度的点估计          | 301        |
| 8.5.4.5      | 量化误差                      | 303        |
| 8.5.4.6      | 从映射中对输入数据检索的精确性           | 304        |
| 8.5.5        | 在映射上形成聚类                  | 305        |
| 8.5.5.1      | 聚类的方法                     | 306        |
| 8.5.5.2      | 在训练过的映射上说明聚类的实例           | 307        |
| 8.5.5.3      | 通过 Ward 方法在映射上寻找最优化的聚类    | 310        |
| 8.5.5.4      | 通过 K 平均值聚类去寻找最优聚类         | 312        |
| 8.5.6        | 对训练过的映射的确认                | 314        |
| 8.5.6.1      | n-折交叉确认                   | 316        |
| 8.6          | 进化的自组织映射                  | 319        |
| 8.6.1        | 映射的成长的细胞结构                | 320        |
| 8.6.1.1      | 将输入数据投射到映射上的神经元之间的位置的质心方法 | 323        |
| 8.6.2        | 通过控制的成长进行动态自组织映射 (GSOM)   | 325        |
| 8.6.2.1      | 实例: 动态自组织映射的应用            | 328        |
| 8.6.3        | 进化树                       | 333        |
| 8.7          | 小结                        | 337        |
|              | 习题                        | 337        |
|              | 参考文献                      | 339        |
| <b>第 9 章</b> | <b>神经网络在时间序列预测中的应用</b>    | <b>341</b> |
| 9.1          | 概述                        | 341        |
| 9.2          | 应用统计模型和神经网络模型进行时间序列线性预测   | 343        |
| 9.2.1        | 实例研究: 炉温控制                | 345        |
| 9.2.1.1      | 多步超前线性预测                  | 347        |
| 9.3          | 用于非线性时间序列预测的神经网络          | 348        |
| 9.3.1        | 聚焦时间滞后和动态驱动的递归网络          | 349        |
| 9.3.1.1      | 聚焦时间滞后前馈网络                | 350        |
| 9.3.1.2      | 时空时间滞后网络                  | 351        |
| 9.3.2        | 例子: 时空时间滞后网络——炉内温度调节      | 354        |
| 9.3.2.1      | 通过神经元 NARx 模型进行单步预测       | 355        |
| 9.3.2.2      | 通过神经元 NARx 模型进行多步预测       | 356        |
| 9.3.3        | 实例研究: 河水流量预测              | 358        |
| 9.3.3.1      | 河流流量预测的线性模型               | 361        |
| 9.3.3.2      | 对于河水流量预测的非线性神经元 NARx 模型   | 363        |
| 9.3.3.3      | 输入敏感度                     | 367        |
| 9.4          | 混合线性 (ARIMA) 和非线性神经网络模型   | 369        |
| 9.4.1        | 实例研究: 对每年太阳黑子数量的预测        | 370        |

|                                                           |     |
|-----------------------------------------------------------|-----|
| 9.5 运用最简化结构概念进行网络结构的自泛化 .....                             | 371 |
| 9.5.1 实例研究:应用自泛化神经网络模型预测空气污染 .....                        | 372 |
| 9.6 广义神经网络 .....                                          | 375 |
| 9.6.1 实例研究:应用一般神经网络进行短期负荷预测 .....                         | 379 |
| 9.7 动态驱动递归网络 .....                                        | 381 |
| 9.7.1 具有隐含神经元反馈的递归网络 .....                                | 382 |
| 9.7.1.1 封装长期记忆 .....                                      | 382 |
| 9.7.1.2 Elman 网络的结构和实施 .....                              | 383 |
| 9.7.1.3 训练递归网络 .....                                      | 385 |
| 9.7.1.4 网络训练实例:手工计算 .....                                 | 389 |
| 9.7.1.5 递归学习网络应用实例研究:降雨量流量模型 .....                        | 393 |
| 9.7.1.6 递归网络的两步超前预测 .....                                 | 396 |
| 9.7.1.7 实时递归学习结构实例研究:两步超前河流流量预测 .....                     | 397 |
| 9.7.2 有输出反馈的递归网络 .....                                    | 399 |
| 9.7.2.1 在有输出反馈的递归网络中封装长期记忆 .....                          | 399 |
| 9.7.2.2 有输出、误差反馈和外生输入的递归网络 (NARIMAx) 的应用:<br>短期温度预测 ..... | 401 |
| 9.7.2.3 具有输出反馈的递归网络的训练 .....                              | 404 |
| 9.7.3 完全递归网络 .....                                        | 405 |
| 9.7.3.1 完全递归网络应用实例研究:短期电力负荷预测 .....                       | 407 |
| 9.8 时间序列预测中的偏差和方差 .....                                   | 409 |
| 9.8.1 将全部误差分解为偏差分量和方差分量 .....                             | 410 |
| 9.8.2 实例:对偏差-方差分离的说明 .....                                | 411 |
| 9.9 长期预测 .....                                            | 417 |
| 9.9.1 实例研究:用多神经网络 (MNN) 进行长期预测 .....                      | 419 |
| 9.10 时间序列预测的输入选择 .....                                    | 421 |
| 9.10.1 从非线性非独立变量中进行输入选择 .....                             | 423 |
| 9.10.1.1 偏交互信息方法 .....                                    | 423 |
| 9.10.1.2 广义回归神经网络 .....                                   | 424 |
| 9.10.1.3 对于输入选择的自组织映射 .....                               | 426 |
| 9.10.1.4 输入选择的遗传算法 .....                                  | 427 |
| 9.10.2 时间序列预测输入选择方法的实际应用 .....                            | 430 |
| 9.10.3 输入选择实例研究:对于预测河流含盐度的输入选择 .....                      | 432 |
| 9.11 小结 .....                                             | 435 |
| 习题 .....                                                  | 437 |
| 参考文献 .....                                                | 438 |
| 附录 .....                                                  | 440 |
| A.1 线性代数、向量、矩阵 .....                                      | 440 |
| A.2 矩阵 .....                                              | 442 |
| 参考文献 .....                                                | 444 |