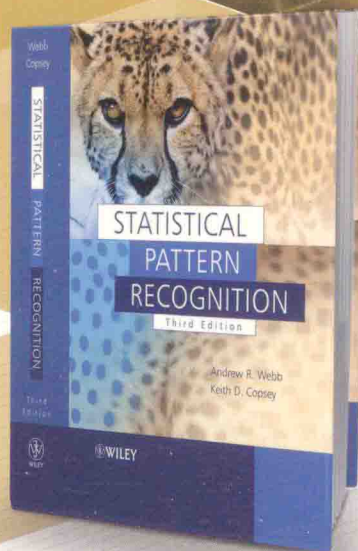


# 统计模式识别 (第三版)

Statistical Pattern Recognition  
Third Edition

【英】 Andrew R. Webb Keith D. Copsey 著  
王萍 译

Statistical Pattern Recognition  
Third Edition



经典译丛·人工智能与智能系统

# 统计模式识别

(第三版)

Statistical Pattern Recognition

Third Edition

[英] Andrew R. Webb 著  
Keith D. Copsey

王萍译

电子工业出版社

Publishing House of Electronics Industry

北京·BEIJING

## 内 容 简 介

本书系统地介绍统计模式识别的理论和新技术,并讨论机器学习领域的诸多问题和相关算法,反映模式识别理论和技术的最新研究进展。其中,大部分识别和分类问题取材于工程学、统计学、计算机科学和社会学等领域的相关应用,并配有应用研究实例。与前版相比,充实或新增了关于估计概率密度的贝叶斯方法、估计概率密度的新的非参数方法、新的分类模型、谱聚类问题、分类规则的归纳法、复杂网络等方面的介绍。

本书注重基本概念、基本方法的讲述,启发性强,且应用实例丰富,适合作为大学高年级和研究生模式识别课程的教材,也适合作为从事模式识别研究和应用工作的相关技术人员的重要参考用书。

Statistical Pattern Recognition, Third Edition, 9780470682272, Andrew R. Webb, Keith D. Copesey.

Copyright © 2011, John Wiley & Sons, Ltd.

All rights reserved. This translation published under license.

Authorized translation from the English language edition published by John Wiley & Sons, Ltd.

本书简体中文字版专有翻译出版权由 John Wiley & Sons, Ltd. 授予电子工业出版社。

未经许可,不得以任何手段和形式复制或抄袭本书内容。

版权贸易合同登记号 图字:01-2012-4995

## 图书在版编目(CIP)数据

统计模式识别:第3版/(英)韦布(Webb, A. R.), (英)科普西(Copesey, K. D.)著;王萍译.

北京:电子工业出版社,2015.1

(经典译丛·人工智能与智能系统)

书名原文:Statistical Pattern Recognition, Third Edition

ISBN 978-7-121-25012-5

I. ①统… II. ①韦… ②科… ③王… III. ①统计模式识别 IV. ①O235

中国版本图书馆CIP数据核字(2014)第280482号

策划编辑:马 岚

责任编辑:马 岚 特约编辑:马爱文

印 刷:三河市鑫金马印装有限公司

装 订:三河市鑫金马印装有限公司

出版发行:电子工业出版社

北京市海淀区万寿路173信箱 邮编 100036

开 本:787×1092 1/16 印张:31.25 字数:800千字

版 次:2004年11月第1版(原著第2版)

2015年1月第2版(原著第3版)

印 次:2015年1月第1次印刷

定 价:89.00元

凡所购买电子工业出版社图书有缺损问题,请向购买书店调换。若书店售缺,请与本社发行部联系,联系及邮购电话:(010)88254888。

质量投诉请发邮件至 zltz@phei.com.cn, 盗版侵权举报请发邮件至 dbqq@phei.com.cn。

服务热线:(010)88258888。

# 译者序

信息时代，无处不有模式识别的需求。概括地讲，模式识别是一门以应用数学为理论基础，利用计算机应用技术，解决实际分类及识别问题的学问。按照研究问题的特点及解决问题的手段特征，通常有统计模式识别和结构模式识别之分，前者以多元统计理论为数学基础，以数据特征的形式对问题进行描述，而后者则以形式语言为数学基础，以结构图元的形式对问题进行描述，它们都致力于将隐含在大量样本中的类间差异的规律归纳出来，并综合成适当的分类、识别乃至预测模型。

从发展的角度看，在传统的、较成熟的分类和识别方法的基础上，模糊数学思想方法的介入，人工神经网络对统计模型类型的丰富、进化算法等一批优秀算法的出现，支持向量机、复杂网络、极度学习和深度学习等一些新方法的提出和介入等，使统计模式识别的研究和应用充满活力。

英国著名学者 Andrew R. Webb 所著《统计模式识别》一书对统计模式识别的理论、概念和方法进行了全面介绍，并在以下方面具有鲜明特点。

**1. 编写体系。**本书以“分类与识别”为主线，在“基本概念-理论分析-方法讲解-应用实例-拓展研究”的框架下，介绍统计模式识别的每一个具体方法；再以应用研究、建议、参考文献等，对由若干方法形成的一类问题进行综述。其中，“拓展研究”能够使读者从知识点伸展到面，进一步了解相关问题的研究动态及人们普遍关注的问题；而“应用研究”则将模式识别技术与广泛的实际问题紧密相联，颇具启发性；“总结”及“建议”凝结了作者的体会和经验，颇具指导性；“参考文献”给出了所列文献与书中内容的联系及其特色。这样的组织格局使读者从局部到全局、从理论到方法、从方法到应用、从研究动态到问题展望，一览无余。

**2. 清晰的分类方法的主线设计。**作者将各种分类器学习方法收纳于统计决策、超特征空间划分这两条主线中，从第2章到第9章，用了共八章的篇幅。统计决策重点解决类概率密度函数的训练，除了非参数法和参数法之外，增加了贝叶斯方法的介绍，特别是按照近邻法-直方图法-核函数法-级数法逐步展开的概率密度估计的讲解，对学习理解、掌握和用好相关技术大有益处；超特征空间划分按照线性和非线性线条展开，自然引出对支持向量机和多层感知器的介绍，规则归纳法反映了模式识别与智能方法的有机联系，搭建起从分类模型的判别分析到可解释规则的桥梁。

**3. 将最新研究方法融入统计模式识别框架。**作者在“分类与识别”主线下带出对统计模式识别概念、新方法（例如人工神经网络、模糊思想用于聚类、支持向量机、新的非参数方法、谱聚类、复杂网络等）的较详尽介绍，使读者能够更深层次地理解它们的构成内涵及其识别行为属性，从而为根据具体问题特点灵活、合理地选用它们提供帮助。

**4. 内容前后呼应。**作者在保持各章节内容相对独立的前提下，特别加强了“谈此及彼”，使读者能够对一种重要方法进行多角度的理解和消化。

**5. 辩证评述和比较性研究。**模式识别问题本身决定了目前实用的模式识别方法和技术没有绝对的好与坏。相信读者会从本书的字里行间领略到作者科学严谨的理论分析及辩证客观

的方法评述，并从中受益。另外，本书特别强调并略加笔墨的“分类器优化组合”、“比较性研究”，近年来受到模式识别学者和专家的重视，值得读者关注。

本书对上一版的大部分章节内容都进行了重新编写和组织，包括内容顺序的梳理和调整，使其内容的模块性更强，分类方法的线条更清晰，与机器学习、数据挖掘及知识发现的关联更紧密；配置了更多的例子和图表，使内容更易读、易理解。

本书的中译本在上版译稿基础上完成。上一版翻译工作由王萍、杨培龙和罗颖昕完成。在这个版本的翻译过程中，范凯波、王娟、王迪、闫春遐和杜雪峰等，在新增内容初译和公式整理等方面提供了帮助。全书由王萍统稿和定稿。

在这里向为本书的翻译工作做出贡献的所有人表示感谢，包括已经毕业的学生杨培龙、罗颖昕和杜雪峰，以及即将毕业的博士生王娟和硕士生闫春遐，在读的博士生范凯波、王迪和石君志。谢谢你们！

由于译者水平所限，译文中难免有疏漏和不妥之处，恳请读者不吝赐教。

王萍

2014年9月

于天津大学

# 前 言

本书介绍统计模式识别的基本理论和技术,其中大部分内容涉及识别和分类问题,并取材于工程学、统计学、计算机科学和社会学等领域的相关文献。在这些文献中,反映了许多当今最有用的模式处理技术,包括许多最新的非参数识别方法和贝叶斯计算方法,本书一并对它们进行介绍,并对使用这些技术方法的起因和支撑这些技术方法的理论展开讨论,以使读者在使用那些流行软件包解决问题时获益最大。本书对各项技术均附以应用研究实例说明之。至于书中涉及的模式识别的应用、对比研究法及理论进展的细节,可以在书后各类文献中找到。

本书内容源自我们对统计模式识别方法进展的研究,以及对传感器数据分析问题的实际应用,针对高年级本科生课程和研究生课程而写,其中有些材料已用于研究生的模式识别课程及模式识别暑期班。本书也是为模式识别领域的实际工作者及其研究者所设计的。作为学习本书内容的先决条件,学习者应具备概率论和线性代数的基本知识,掌握一些基本数学方法(例如,在一些推导中,用于解决具有等式约束和不等式约束问题的拉格朗日数乘法)。本书前版附录提供的一些基本材料可以在本书配套网站([www.wiley.com/go/statistical\\_pattern\\_recognition](http://www.wiley.com/go/statistical_pattern_recognition))找到。

## 范围

本书展现绝大多数常用的统计模式识别方法。然而,模式识别的许多重要研究进展并非局限于统计学文献,而经常呈现于与机器学习交叉的研究领域。因此,打破传统的统计模式识别的框架将是有益的,本书正是这样做的。例如,我们把一些规则归纳方法作为一种补充方法添加进来,以通过决策树归纳掌控探索过程。本书谈到的大多数方法具有一般性,即这些方法并不要求指定数据或应用的特定类型,于是本书内容不涉及大家时常用到的信号(和图像)预处理方法,以及信号(和图像)滤波方法。

## 方法

本书每一章所讨论的方法,均会安排讲述与其相关的基本概念和算法,均会在章末给出引自参考文献的相关方法或分类技术的实际应用,其主要目的是理解方法的基本概念。有时候需要进行一些详细的数学描述,因此有时不得不划一个界限,以掌控把哪个特定主题讨论到多深。本书涉及的大部分主题可以用整本书来论述,于是我们不得不对所拥有的材料进行取舍,因此每一章的最后一节均提供了主要的参考文献。章末所附习题与开卷式问题有所不同,开卷式问题涉及比较冗长的计算机工程项目。

## 第三版的新增内容

本书对前版的许多章节进行了重新编写,并添加了一些新的材料,新增内容特点如下。

- 第3章的内容是新增的,这一章讲述密度估计的贝叶斯法,包括对贝叶斯采样方案的内容拓展、马尔可夫链蒙特卡罗方法、序贯蒙特卡罗采样器和变分贝叶斯法。
- 新增一节专门讲述密度估计的非参数方法。

- 新增规则归纳方法。
- 为分类器的组合方法新增一章。
- 对特征选择内容进行了重新修订，增添了关于特征选择稳定性的章节。
- 新增谱聚类内容。
- 新增一章讲述复杂网络问题，这个问题与社会及计算机网络分析的高增长领域相关。

## 全书梗概

第1章作为统计模式识别的绪论，给出一些名词术语的定义，介绍监督型分类和无监督型分类。就监督型分类而言，有两种研究方法：一种方法基于概率密度函数的运用；另一种方法则基于判别函数的构建。在这一章的最后对模式识别的完整过程进行概括，细节问题则安排在后继章节中讨论。第2章至第4章讨论识别问题的密度函数法。其中，第2章讲解密度函数估计的参数法，它们在贝叶斯法上的进一步拓展安排在第3章，第4章讨论非参数分类器的实现方案，包括被广泛使用的 $k$ 近邻法及与之相关的有效搜索算法。

第5章至第7章研究有监督分类问题的判别函数的构建方法。第5章集中讨论线性判别函数，其中所涉及的大多数判别法(包括优化、正则化和支持向量机)也适用于第6章展开的非线性研究。第6章探讨基于核函数的方法，特别是径向基函数网络和支持向量机，还讨论了基于投影的方法(多层感知器)，这些通常称为神经网络方法。第7章讨论如何使分类函数变为可解释的规则，这种判别方法对一些应用来说非常重要。

第8章讨论分类器的集成方法，即为提高系统的鲁棒性，将多个分类器组合起来。第9章讲述如何测评分类器的性能。

第10章和第11章探讨数据分析和预处理技术(这些工作通常先于第5章至第7章介绍的有监督分类工作，尽管有时可以用来作为有监督方法的后置处理)。第10章讲述特征选择和特征提取方法，它们用以降低描述原始数据特征的维数，这项工作通常是分类器整体设计工作的一部分，只是被人为地将这一模式识别问题划分为相对独立的特征提取过程和模式分类过程。特征提取可以帮助我们深入了解数据结构及分类器需要选用的类型，因此该研究备受瞩目。第11章讲述无监督分类或称聚类问题，即在样本群中找到所存在的结构并借此将其分组的过程。这类技术的工程应用是对图像进行矢量量化及对语音编码。第12章讨论复杂网络问题，所述方法对待分析的数据用图形的数学概念进行表述，所述及问题与社会及计算机网络的关联很显著。

最后，即第13章，讨论一些重要的包括模型选择问题在内的研究课题。

## 本书网站

网站 [www.wiley.com/go/statistical\\_pattern\\_recognition](http://www.wiley.com/go/statistical_pattern_recognition) 对如下问题提供了补充材料：相异测度、估计方法、线性代数、数据分析和基本概率方法。

## 致谢

在编写本书第三版的过程中，我们得到了很多人的帮助。在此特别感谢 East Anglia 大学的 Gavin Cawley 博士所给予的帮助和建议，感谢朋友们和同事们(RSRE, DERA 和 QinetiQ 的自始至终的帮助)，他们对原稿的不同部分提出了许多宝贵意见。还要特别感谢 Anna Skeoch 为第12章提供数据；感谢 Richard Davies 和 John Wiley 的同事们为稿件的最终出版所给予的帮助。Andrew Webb 特别感谢 Rosemary 所给予的爱、支持和耐心。

## 符 号

下面列出一些常用符号, 并使用国际惯例进行标注。例如, 对于变量和变量上的观测值, 我们倾向于使用同一符号。它们的不同含义可以从上下文中明显看出。再者, 我们将  $x$  的密度函数记为  $p(x)$ , 将  $y$  的密度函数记为  $p(y)$ , 即使这两个函数并不相同。我们用黑体的小写字母表示向量, 而用粗体的大写字母表示矩阵。模式识别涉及多学科的研究问题, 因此, 所用符号在书中各章保持一致, 同时与各类文献的常用符号保持一致, 这一点很重要。本书尽可能做到在一章内保持符号的一致性。

$p, d$	变量个数
$C$	类别数
$n$	观测值个数
$n_j$	第 $j$ 类的观测值个数
$\omega_j$	第 $j$ 类的标记
$X_1, \dots, X_p$	$p$ 个随机变量
$x_1, \dots, x_p$	变量 $X_1, \dots, X_p$ 上的观测值
$\mathbf{x} = (x_1, \dots, x_p)^T$	观测值向量
$\mathbf{X} = [\mathbf{x}_1, \dots, \mathbf{x}_n]^T$	$n \times p$ 的数据矩阵
$\mathbf{X} = \begin{bmatrix} x_{11} & \cdots & x_{1p} \\ \vdots & \ddots & \vdots \\ x_{n1} & \cdots & x_{np} \end{bmatrix}$	
$P(\mathbf{x}) = \text{prob}(X_1 \leq x_1, \dots, X_p \leq x_p)$	
$p(\mathbf{x}) = \partial P / \partial \mathbf{x}$	概率密度函数
$p(\mathbf{x}   \omega_j)$	第 $j$ 类的概率密度函数
$p(\omega_j)$	第 $j$ 类的先验概率
$\boldsymbol{\mu} = \int \mathbf{x} p(\mathbf{x}) d\mathbf{x}$	总体均值
$\boldsymbol{\mu}_j = \int \mathbf{x} p(\mathbf{x}   \omega_j) d\mathbf{x}$	第 $j$ 类的均值, $j = 1, \dots, C$
$\mathbf{m} = (1/n) \sum_{i=1}^n \mathbf{x}_i$	样本均值
$\mathbf{m}_j = (1/n_j) \sum_{i=1}^n z_{ji} \mathbf{x}_i$	第 $j$ 类的样本均值, $j = 1, \dots, C$ , 若 $\mathbf{x}_i \in \omega_j$ , 则 $z_{ji} = 1$ , 否则为 0, $n_j$ 为 $\omega_j$ 类的样本数,
	$n_j = \sum_{i=1}^n z_{ji}$



$\hat{\Sigma} = \frac{1}{n} \sum_{i=1}^n (\mathbf{x}_i - \mathbf{m})(\mathbf{x}_i - \mathbf{m})^T$	样本协方差矩阵(极大似然估计)
$n/(n-1)\hat{\Sigma}$	样本协方差矩阵(无偏估计)
$\hat{\Sigma}_j = \frac{1}{n_j} \sum_{i=1}^{n_j} z_{ji}(\mathbf{x}_i - \mathbf{m}_j)(\mathbf{x}_i - \mathbf{m}_j)^T$	第 $j$ 类的样本协方差矩阵(极大似然估计)
$S_j = \frac{n_j}{n_j - 1} \hat{\Sigma}_j$	第 $j$ 类的样本协方差矩阵(无偏估计)
$S_w = \sum_{j=1}^c \frac{n_j}{n} \hat{\Sigma}_j$	总体类内样本协方差矩阵
$S = \frac{n}{n-C} S_w$	总体类内样本协方差矩阵(无偏估计)
$S_B = \sum \frac{n_j}{n} (\mathbf{m}_j - \mathbf{m})(\mathbf{m}_j - \mathbf{m})^T$	类间样本矩阵
$S_B + S_w = \hat{\Sigma}$	
$\ A\ ^2 = \sum_{ij} A_{ij}^2$	
$N(\mathbf{m}, \Sigma)$	均值为 $\mathbf{m}$ 且协方差矩阵为 $\Sigma$ 的正态(或高斯)分布
$N(\mathbf{x}; \mathbf{m}, \Sigma)$	均值为 $\mathbf{m}$ 且协方差矩阵为 $\Sigma$ 的正态分布的概率密度函数
$E[Y X]$	给定 $X$ 下 $Y$ 的期望
$I(\theta)$	指示函数, 若 $\theta$ 为真, 其值为 1, 否则为 0

# 目 录

第 1 章 统计模式识别绪论 .....	1
1.1 统计模式识别 .....	1
1.1.1 引言 .....	1
1.1.2 基本模型 .....	2
1.2 解决模式识别问题的步骤 .....	3
1.3 问题讨论 .....	4
1.4 统计模式识别的方法 .....	5
1.5 基本决策理论 .....	6
1.5.1 最小错误贝叶斯决策规则 .....	6
1.5.2 最小错误贝叶斯决策规则——拒绝分类 .....	8
1.5.3 最小风险贝叶斯决策规则 .....	10
1.5.4 最小风险贝叶斯决策规则——拒绝分类 .....	11
1.5.5 Neyman-Pearson 决策规则 .....	11
1.5.6 最小最大决策 .....	13
1.5.7 讨论 .....	14
1.6 判别函数 .....	15
1.6.1 引言 .....	15
1.6.2 线性判别函数 .....	16
1.6.3 分段线性判别函数 .....	17
1.6.4 广义线性判别函数 .....	17
1.6.5 小结 .....	19
1.7 多重回归 .....	19
1.8 本书梗概 .....	21
1.9 提示及参考文献 .....	21
习题 .....	22
第 2 章 密度估计的参数法 .....	24
2.1 引言 .....	24
2.2 分布参数估计 .....	24
2.2.1 估计法 .....	24
2.2.2 预测法 .....	25
2.3 高斯分类器 .....	25
2.3.1 详述 .....	25
2.3.2 高斯分类器插入估计的推导 .....	27
2.3.3 应用研究举例 .....	28
2.4 处理高斯分类器的奇异问题 .....	29

2.4.1	引言	29
2.4.2	朴素贝叶斯	29
2.4.3	投影到子空间	30
2.4.4	线性判别函数	30
2.4.5	正则化判别分析	31
2.4.6	应用研究举例	32
2.4.7	拓展研究	33
2.4.8	小结	34
2.5	有限混合模型	34
2.5.1	引言	34
2.5.2	混合判别模型	35
2.5.3	正态混合模型的参数估计	36
2.5.4	正态混合模型协方差矩阵约束	38
2.5.5	混合模型分量的数量	39
2.5.6	期望最大化算法下的极大似然估计	40
2.5.7	应用研究举例	44
2.5.8	拓展研究	46
2.5.9	小结	46
2.6	应用研究	47
2.7	总结和讨论	49
2.8	建议	49
2.9	提示及参考文献	49
	习题	50
<b>第3章</b>	<b>密度估计的贝叶斯法</b>	<b>52</b>
3.1	引言	52
3.1.1	基本原理	53
3.1.2	递归计算	53
3.1.3	比例性	54
3.2	解析解	54
3.2.1	共轭先验概率	54
3.2.2	方差已知的正态分布的均值估计	55
3.2.3	多元正态分布的均值及协方差矩阵估计	58
3.2.4	未知类先验概率的情形	62
3.2.5	小结	63
3.3	贝叶斯采样方案	64
3.3.1	引言	64
3.3.2	梗概	64
3.3.3	贝叶斯分类器的采样类型	65
3.3.4	拒绝采样	65
3.3.5	均匀比	66

3.3.6	重要性采样	67
3.4	马尔可夫链蒙特卡罗方法	69
3.4.1	引言	69
3.4.2	吉布斯(Gibbs)采样器	69
3.4.3	Metropolis-Hastings 算法	75
3.4.4	数据扩充	78
3.4.5	可逆跳跃马尔可夫链蒙特卡罗方法	79
3.4.6	切片采样	80
3.4.7	MCMC 举例——正弦噪声估计	82
3.4.8	小结	84
3.4.9	提示及参考文献	85
3.5	贝叶斯判别方法	85
3.5.1	标记训练数据	85
3.5.2	无类别标签的训练数据	86
3.6	连续蒙特卡罗采样	88
3.6.1	引言	88
3.6.2	基本方法	89
3.6.3	小结	92
3.7	变分贝叶斯方法	93
3.7.1	引言	93
3.7.2	描述	93
3.7.3	分解为因子的变分近似	95
3.7.4	简单的例子	96
3.7.5	模型选择中的运用	99
3.7.6	拓展研究与应用	100
3.7.7	小结	101
3.8	近似贝叶斯计算	101
3.8.1	引言	101
3.8.2	ABC 拒绝采样	102
3.8.3	ABC MCMC 采样	103
3.8.4	ABC 总体蒙特卡罗采样	104
3.8.5	模型选择	105
3.8.6	小结	106
3.9	应用研究举例	106
3.10	应用研究	107
3.11	总结和讨论	108
3.12	建议	108
3.13	提示及参考文献	109
	习题	109

第 4 章 密度估计的非参数法 .....	111
4.1 引言 .....	111
4.1.1 密度估计的基本性质 .....	111
4.2 $k$ 近邻法 .....	112
4.2.1 $k$ 近邻分类器 .....	112
4.2.2 推导 .....	114
4.2.3 距离度量的选择 .....	116
4.2.4 最近邻法决策规则的性质 .....	117
4.2.5 线性逼近排除搜索算法 .....	118
4.2.6 分支定界搜索算法: kd 树 .....	120
4.2.7 分支定界搜索算法: ball 树 .....	126
4.2.8 剪辑方法 .....	129
4.2.9 应用研究举例 .....	130
4.2.10 拓展研究 .....	131
4.2.11 小结 .....	132
4.3 直方图法 .....	132
4.3.1 直方图自适应数据 .....	133
4.3.2 独立性假设(朴素贝叶斯) .....	133
4.3.3 Lancaster 模型 .....	134
4.3.4 最大权值相关树 .....	134
4.3.5 贝叶斯网络 .....	137
4.3.6 应用研究举例: 朴素贝叶斯文本分类 .....	139
4.3.7 小结 .....	141
4.4 核函数方法 .....	142
4.4.1 有偏估计 .....	144
4.4.2 延伸到多元 .....	145
4.4.3 平滑参数的选择 .....	146
4.4.4 核函数的选择 .....	148
4.4.5 应用研究举例 .....	148
4.4.6 拓展研究 .....	149
4.4.7 小结 .....	150
4.5 用基函数展开 .....	151
4.6 copula 方法 .....	152
4.6.1 引言 .....	152
4.6.2 数学基础 .....	152
4.6.3 copula 函数 .....	153
4.6.4 copula 概率密度函数的估计 .....	154
4.6.5 简单举例 .....	155
4.6.6 小结 .....	156
4.7 应用研究 .....	157

4.7.1	比较研究	159
4.8	总结和讨论	159
4.9	建议	160
4.10	提示及参考文献	160
	习题	161
<b>第5章</b>	<b>线性判别分析</b>	<b>163</b>
5.1	引言	163
5.2	两类问题算法	163
5.2.1	总体思路	163
5.2.2	感知准则	164
5.2.3	Fisher 准则	167
5.2.4	最小均方误差法	168
5.2.5	拓展研究	172
5.2.6	小结	173
5.3	多类算法	174
5.3.1	总体思路	174
5.3.2	错误修正法	174
5.3.3	Fisher 准则:线性判别分析	175
5.3.4	最小均方误差法	177
5.3.5	正则化	181
5.3.6	应用研究实例	181
5.3.7	拓展研究	182
5.3.8	小结	182
5.4	支持向量机	183
5.4.1	引言	183
5.4.2	两类线性可分数据问题	183
5.4.3	两类线性不可分数据问题	186
5.4.4	多类支持向量机	188
5.4.5	支持向量机回归	189
5.4.6	具体实施	191
5.4.7	应用研究举例	193
5.4.8	小结	194
5.5	logistic 判别	194
5.5.1	两类问题	194
5.5.2	极大似然估计	195
5.5.3	多类 logistic 判别	196
5.5.4	应用研究举例	197
5.5.5	拓展研究	197
5.5.6	小结	198
5.6	应用研究	198

5.7	总结和讨论	198
5.8	建议	199
5.9	提示及参考文献	199
	习题	199
<b>第6章</b>	<b>非线性判别分析——核与投影法</b>	<b>202</b>
6.1	引言	202
6.2	径向基函数	203
6.2.1	引言	203
6.2.2	模型的确定	204
6.2.3	指定函数的形式	205
6.2.4	中心位置	206
6.2.5	平滑参数	207
6.2.6	权值的计算	207
6.2.7	模型阶次的选择	209
6.2.8	简单径向基函数	210
6.2.9	一些调整	210
6.2.10	径向基函数的性质	212
6.2.11	应用研究举例	212
6.2.12	拓展研究	213
6.2.13	小结	214
6.3	非线性支持向量机	214
6.3.1	引言	214
6.3.2	二分类	214
6.3.3	核函数的类型	215
6.3.4	模型选择	216
6.3.5	多类支持向量机	217
6.3.6	概率估计	217
6.3.7	非线性回归	218
6.3.8	应用研究举例	219
6.3.9	拓展研究	219
6.3.10	小结	220
6.4	多层感知器	220
6.4.1	引言	220
6.4.2	多层感知器结构的确定	221
6.4.3	多层感知器权值的确定	222
6.4.4	多层感知器的建模能力	226
6.4.5	logistic 分类	227
6.4.6	应用研究举例	228
6.4.7	贝叶斯多层感知器网络	229
6.4.8	投影寻踪	231

6.4.9 小结 .....	231
6.5 应用研究 .....	232
6.6 总结和讨论 .....	234
6.7 建议 .....	234
6.8 提示及参考文献 .....	235
习题 .....	235
<b>第7章 规则和决策树归纳法 .....</b>	<b>238</b>
7.1 引言 .....	238
7.2 决策树 .....	238
7.2.1 引言 .....	238
7.2.2 决策树的构造 .....	241
7.2.3 拆分规则的选择 .....	241
7.2.4 终止拆分过程 .....	244
7.2.5 为终端节点分配类标签 .....	245
7.2.6 决策树剪枝(含实施示例) .....	245
7.2.7 决策树构造方法 .....	250
7.2.8 其他问题 .....	251
7.2.9 应用研究举例 .....	252
7.2.10 拓展研究 .....	252
7.2.11 小结 .....	253
7.3 规则归纳 .....	253
7.3.1 引言 .....	253
7.3.2 从决策树生成规则 .....	255
7.3.3 用连续覆盖算法进行规则归纳 .....	256
7.3.4 应用研究举例 .....	259
7.3.5 拓展研究 .....	260
7.3.6 小结 .....	260
7.4 多元自适应回归样条 .....	260
7.4.1 引言 .....	260
7.4.2 递归分割模型 .....	260
7.4.3 应用研究举例 .....	263
7.4.4 拓展研究 .....	263
7.4.5 小结 .....	263
7.5 应用研究 .....	264
7.6 总结和讨论 .....	265
7.7 建议 .....	265
7.8 提示及参考文献 .....	266
习题 .....	266



<b>第 8 章 组合方法</b> .....	268
8.1 引言 .....	268
8.2 分类器组合方案特性 .....	269
8.2.1 特征空间 .....	269
8.2.2 层次 .....	272
8.2.3 训练程度 .....	273
8.2.4 成员分类器的形式 .....	273
8.2.5 结构 .....	274
8.2.6 优化 .....	274
8.3 数据融合 .....	274
8.3.1 体系结构 .....	275
8.3.2 贝叶斯方法 .....	276
8.3.3 奈曼-皮尔逊(Neyman-Pearson)公式 .....	277
8.3.4 可训练规则 .....	278
8.3.5 固定规则 .....	278
8.4 分类器组合方法 .....	279
8.4.1 乘积规则 .....	279
8.4.2 和式规则 .....	280
8.4.3 最小、最大及中值组合分类器 .....	281
8.4.4 多数表决 .....	281
8.4.5 Borda 数 .....	282
8.4.6 在类别预测上训练组合分类器 .....	282
8.4.7 叠加归纳 .....	284
8.4.8 专家混合器 .....	284
8.4.9 bagging .....	286
8.4.10 boosting .....	287
8.4.11 随机森林 .....	289
8.4.12 模型平均 .....	290
8.4.13 方法小结 .....	295
8.4.14 应用研究举例 .....	295
8.4.15 拓展研究 .....	296
8.5 应用研究 .....	297
8.6 总结和讨论 .....	297
8.7 建议 .....	298
8.8 提示及参考文献 .....	298
习题 .....	299
<b>第 9 章 性能评价</b> .....	300
9.1 引言 .....	300
9.2 性能评价 .....	300
9.2.1 性能测度 .....	300