



书中所有程序的源代码
可通过扫描此二维码
免费下载。

模式识别与智能计算的 MATLAB实现

(第2版)

许国根 贾 瑛 韩启龙 编著



北京航空航天大学出版社
BEIHANG UNIVERSITY PRESS

MATLAB®
examples

MATLAB®&Simulink® 开发实例系列丛书

模式识别与智能计算的 MATLAB 实现 (第 2 版)

许国根 贾 瑛 韩启龙 编著



配套资料(程序源代码+数据)

北京航空航天大学出版社

内 容 简 介

针对各学科数据信息的特点以及科学工作者对信息处理和数据挖掘技术的要求,本书既介绍了模式识别和智能计算的基础知识,又较为详细地介绍了现代模式识别和智能计算在科学研究中的应用方法和各算法的 MATLAB 源程序。本书可以帮助广大的科学工作者掌握模式识别和智能计算方法,并应用于实际的研究中,提高对海量数据信息的处理及挖掘能力,针对性和实用性强,具有较高的理论和实用价值。

本书可作为高等院校计算机工程、信息工程、生物医学工程、智能机器人、工业自动化、地质、水利、化学和环境等专业研究生、本科生的教材或教学参考书,亦可供有关工程技术人员参考。

图书在版编目(CIP)数据

模式识别与智能计算的 MATLAB 实现 / 许国根,贾瑛,韩启龙编著. -- 2 版. -- 北京:北京航空航天大学出版社,2017.5

ISBN 978-7-5124-2400-5

I. ①模… II. ①许… ②贾… ③韩… III. ①模式识别-计算机辅助计算-Matlab 软件②人工智能-计算机辅助计算-Matlab 软件 IV. ①O235-39②TP183

中国版本图书馆 CIP 数据核字(2017)第 086194 号

版权所有,侵权必究。

模式识别与智能计算的 MATLAB 实现(第 2 版)

许国根 贾 瑛 韩启龙 编著

责任编辑 王慕冰

*

北京航空航天大学出版社出版发行

北京市海淀区学院路 37 号(邮编 100191) <http://www.buaapress.com.cn>

发行部电话:(010)82317024 传真:(010)82328026

读者信箱:goodtextbook@126.com 邮购电话:(010)82316936

北京兴华昌盛印刷有限公司印装 各地书店经销

*

开本:787×1092 1/16 印张:24.25 字数:636 千字

2017 年 8 月第 2 版 2017 年 8 月第 1 次印刷 印数:4 000 册

ISBN 978-7-5124-2400-5 定价:49.00 元

若本书有倒页、脱页、缺页等印装质量问题,请与本社发行部联系调换。联系电话:(010)82317024

作者简介



许国根，毕业于南京大学分析化学专业，长期奋斗在一线的高校资深化学教师。为了实现“数学化学”梦想，一直致力于数学在化学中的应用研究，热衷于

MATLAB、模式识别、数据挖掘、化学计量学等相关知识的学习与应用，撰写过多部介绍MATLAB在化学中应用技巧的书籍。

“在线交流”网络互动参与步骤:

❶ 在MATLAB中文论坛www.iLoveMatlab.cn注册一个会员账号并登录。

❷ 进入本书所在版块

<http://www.ilovematlab.cn/forum-210-1.html>，与作者和同行们在线交流。

第 2 版前言

本书理论联系实际,较为全面地介绍了现代模式识别和智能计算方法及其应用技巧。通过大量的实例,讲解了模式识别和智能计算的理论、算法及编程步骤,并提供了 MATLAB 程序的源代码。通过学习本书,读者能够真正掌握模式识别和智能计算方法并应用于实际研究和工程实践。

本书第 1 版面世以来,受到了读者的好评,并提出了许多宝贵的意见。根据读者反馈的意见和建议,本版对第 1 版作了适当的修订,增加了以下内容:第 14 章介绍了人工鱼群算法、进化计算、人工免疫算法等群体智能算法的理论和方法;第 15 章介绍了仿生模式识别的基本原理及应用。这 2 章的内容由韩启龙、曾宝平、冯锐、王坤、王爽、戴津星等人编写。

本次再版对第 1 版的文字及程序中的一些错误作了修改,为了方便读者学习,还提供了完整的各章节习题的文本文件及相应的程序。读者首先将提供的程序安装到 MATLAB 的搜索目录下(work 文件夹),或复制到任意目录,再通过设置 MATLAB 的路径将安装程序的目录设为 MATLAB 搜索目录,学习时将各习题的文本文件复制到 MATLAB 计算窗口,即可进行习题的运算。

MATLAB 的功能非常多,要想完全掌握它的全部功能非常困难,也没有必要。无论 MATLAB 怎样发展,归根结底它只是一个工具,何况 MATLAB 对于许多算法而言给出的是基本算法,不可能非常完美,总有一个适用的范围和场合。因此学习 MATLAB 更为重要的是通过学习掌握它的基本功能,再结合专业知识和实际情况,对 MATLAB 相应的函数进行修改,或者是自编函数进行算法参数优化,或通过与其他算法的结合等研究,从而解决学习和工作中遇到的复杂问题。因此本书中给出的一些自编函数在 MATLAB 中可能有相应的函数,或者会在今后更高的版本中给出,或者在更高的版本中用法不同,但不影响读者对本书内容的学习,相反能进一步加深对方法及 MATLAB 功能的学习和掌握,进而借助于 MATLAB 的基本功能实现任何算法和方法,摆脱“MATLAB 控”。

本书的再版得到了北京航空航天大学出版社的大力支持,编辑陈守平对本书的内容修订提出了许多宝贵的意见,很多读者也给作者提出了各种有益的建议,在此一并表示衷心的感谢!

由于作者水平有限,虽经修订,书中难免仍存在缺点和错误,敬请读者批评斧正。

作者

2016 年 10 月于西安

第 1 版前言

MATLAB 是功能非常强大的计算机软件,在科学研究和工程实践中得到了广泛的应用。利用它来编制现代模式识别和智能计算等技术的程序,揭开这些在大多数人眼中极为深奥的数学方法的神秘面纱,使每个科学工作者都能非常容易地使用它们来解决实际问题,是作者学习 MATLAB 后,结合实际的科学研究经验产生的一个强烈的愿望。

本书理论联系实际,较为全面地介绍了现代模式识别和智能计算方法及其应用技巧。通过大量实例,讲解了模式识别和智能计算的理论、算法及编程步骤,并提供基于 MATLAB 的源代码。通过本书的学习,读者能够真正掌握模式识别和智能计算并应用于科学研究和工程实践中。

模式识别是当今高科技研究的重要领域之一,它创立于 20 世纪 60 年代,初期属于信息、控制和系统科学领域。模式识别是利用某些特征,对一组对象进行判别或分类,被分类的对象即为模式,分类的过程即为识别。模式识别所涉及的信息往往存在高维、影响因素多、关系复杂等特征,单靠人的思维往往难以有效地确定其规律,需要通过一定的数学方法借助计算机来完成。到了 20 世纪 70 年代后,随着大规模集成电路技术的发展,特别是计算机硬件的飞速发展,无论在理论上,还是在应用上,模式识别技术都有了长足的进步;同时,也推动了以计算机科学为基础的具有智能性质的自动化系统的实际应用,促进了人工智能、专家系统、景物动态分析、图像识别、语音识别等多学科的发展,广泛应用于人工智能、机器人、系统控制、遥感数据分析、生物医学工程、军事目标识别等领域,在国民经济、国防建设、社会发展等各个方面发挥着越来越重要的作用。

在众多学科的科学研究的工程应用中,人们往往通过对研究对象的观察和实验积累了海量的数据信息,并且由于对象的复杂性,使得这些数据具有高维、复杂非线性、强相关性和多噪声等特点。如何从这些数据信息中发现更多、更有价值的关系,找到其内在规律,建立的模型能较好地反映研究对象的实际特征,并具有良好的可理解性,易与先验知识相结合,能适应大规模数据处理的要求,正逐步成为科学工作者关注的焦点。常规数学手段已不能解决这个问题,现代模式识别和智能计算将起到十分重要的核心作用。

现代模式识别和智能计算从已知数据出发,参照相应的数学(或物理、化学)模型或经验规律得到一批特征量,然后进一步进行特征抽取以求得合适的特征量,张成模式空间或特征空间,最后通过模式识别算法进行训练和判别,以揭示已知数据信息中隐含的性质和规律,为研究者提供十分有用的决策信息和过程优化的重要信息。

现代模式识别和智能计算作为一种高效的信息处理技术,解决了众多学科研究和工程应用中的重要问题。例如化学研究中对化合物性质的分类、化合物的构效分析、基于化学方法的疾病诊断、药物的分类、环境质量评价等,都可以通过模式识别和智能计算对一组样本中某些化学组分的分析结果进行分析处理来实现;在机械故障诊断、军事目标识别、遥感数据分析、水利等学科的研究中,模式识别和智能计算也得到了广泛的应用。目前,模式识别和智能计算已经成为科学研究中的一种重要的数据处理手段,应用越来越广泛。

国内外论述模式识别和智能计算的参考书为数不少。由于这一领域涉及深奥的数学理论,而且大部分书籍只是罗列模式识别和智能计算的原理及伪代码,看不到源代码以及算法的

实际效果和各种算法的对比结果,往往使大多数科学工作者感到困惑,不知如何下手。对大多数科学工作者而言,目前确实还缺少一本具有较强的系统性、可比性及实用性的模式识别和智能计算的参考书。基于这点考虑,作者撰写了本书,目的是想通过系统的介绍和实例分析,让众多学科的科学工作者不仅具备模式识别和智能计算的理论,而且能掌握模式识别和智能计算的方法,可以在各自的学科实际研究中加以应用。

本书内容基本涵盖了目前模式识别和智能计算的重要理论和方法,包括了最近十几年来刚刚发展起来的并被实践证明有用的新技术、新理论,如支持向量机、神经网络、决策树、粗糙集理论、模糊集理论和遗传算法等。在介绍各种理论和方法的同时,将不同算法应用于科学研究的实际中。对于每一种模式识别方法,本书都通过实际问题,按照理论基础、实现步骤、编程代码三部分进行阐述,避免空洞的理论说教,着重介绍编程步骤及 MATLAB 源代码,具有较强的指导性和实用性,使读者不至于面对如此丰富的理论和方法无所适从,而是通过了解各种算法的实现思路和方法,体会算法源代码的意义,这样即使所举的实例不属于读者从事的学科,通过举一反三,读者也能掌握模式识别和智能计算并应用于自己从事的科学研究中。

由于至今还没有统一的、有效的能够应用于所有的模式识别的理论,当前的一种普遍看法认为,不存在对所有的模式识别问题都适应的单一模型和解决识别问题的单一技术,即存在着所谓的无免费午餐定理(NFL 定理):若算法 A 在某些函数中的表现超过算法 B,则在这类其他的适应度值函数中,算法 B 的表现就比 A 要好。我们所要做的是把模式识别方法与具体问题结合起来,把模式识别和智能计算与统计学、神经网络、数据挖掘、机器学习、人工智能等学科的先进思想和理论融合在一起,为读者提供一个多种理论和方法的测试平台,并在此基础上,深入掌握各种理论的效能和应用的可能性,互相取长补短,开创模式识别和智能计算在各学科中应用的新局面。

本书的出版得到了张剑、张永勇、崔虎、吴婉娥、郭和军、吕晓猛、李茸、马岚、王焕春、谢拯、慕晓刚、姚汝亮等同仁和研究生的支持,他们在选题确定、章节安排及习题解答等方面给予了很大的帮助;同时,本书的出版也得到了北京航空航天大学出版社的大力支持,编辑陈守平在本书内容编排等许多方面提出了宝贵的意见,在此表示衷心的感谢!

由于作者水平有限,特别是数学知识相对贫乏,书中的错误和不当之处,敬请读者批评斧正。

作 者

2011 年于西安二炮工程大学



配套资料(程序源代码+数据)

本书所有程序的源代码均可通过 QQ 浏览器扫描二维码免费下载。读者也可以通过以下网址下载全部资料:<http://www.buaapress.com.cn/waiyump3/9787512424005/MSSByznjs.rar>。

配套资料下载或与本书相关的其他问题,请咨询北京航空航天大学出版社理工图书分社,电话(010)82317036。

若您对此书内容有任何疑问,可以登录 MATLAB 中文论坛与作者和同行交流。

目 录

第 1 章 绪 论	1
1.1 模式识别的基本概念	1
1.1.1 模式与模式识别的概念	1
1.1.2 模式的特征	1
1.1.3 模式识别系统	2
1.2 模式识别的主要方法	2
1.3 模式识别的主要研究内容	3
1.4 模式识别在科学研究中的应用	3
1.4.1 化合物的构效分析	3
1.4.2 谱图解析	4
1.4.3 材料研究	4
1.4.4 催化剂研究	5
1.4.5 机械故障诊断与监测	5
1.4.6 化学物质源产地判断	6
1.4.7 疾病的诊断与预测	6
1.4.8 矿藏勘探	7
1.4.9 考古及食品工业中的应用	7
第 2 章 统计模式识别技术	8
2.1 基于概率统计的贝叶斯分类方法	8
2.1.1 最小错误率贝叶斯分类	9
2.1.2 最小风险率贝叶斯分类	10
2.2 线性分类器	12
2.2.1 线性判别函数	12
2.2.2 Fisher 线性判别函数	13
2.2.3 感知器算法	14
2.3 非线性分类器	15
2.3.1 分段线性判别函数	15
2.3.2 近邻法	17
2.3.3 势函数法	19
2.3.4 SIMCA 方法	20
2.4 聚类分析	22
2.4.1 模式相似度	22
2.4.2 聚类准则	24
2.4.3 层次聚类法	25
2.4.4 动态聚类法	25

2.4.5	决策树分类器	28
2.5	统计模式识别在科学研究中的应用	29
第 3 章	人工神经网络及模式识别	43
3.1	人工神经网络的基本概念	43
3.1.1	人工神经元	43
3.1.2	传递函数	43
3.1.3	人工神经网络分类和特点	44
3.2	BP 人工神经网络	44
3.2.1	BP 人工神经网络学习算法	44
3.2.2	BP 人工神经网络 MATLAB 实现	46
3.3	径向基函数神经网络 RBF	47
3.3.1	RBF 的结构与学习算法	47
3.3.2	RBF 的 MATLAB 实现	48
3.4	自组织竞争人工神经网络	48
3.4.1	自组织竞争人工神经网络的基本概念	48
3.4.2	自组织竞争神经网络的学习算法	49
3.4.3	自组织竞争网络的 MATLAB 实现	49
3.5	对向传播神经网络 CPN	50
3.5.1	CPN 的基本概念	50
3.5.2	CPN 网络的学习算法	50
3.6	反馈型神经网络 Hopfield	51
3.6.1	Hopfield 网络的基本概念	51
3.6.2	Hopfield 网络的学习算法	52
3.6.3	Hopfield 网络的 MATLAB 实现	53
3.7	人工神经网络技术在科学研究中的应用	53
第 4 章	模糊系统理论及模式识别	72
4.1	模糊系统理论基础	72
4.1.1	模糊集合	72
4.1.2	模糊关系	75
4.1.3	模糊变换与模糊综合评判	77
4.1.4	If...then 规则	78
4.1.5	模糊推理	78
4.2	模糊模式识别的基本方法	79
4.2.1	最大隶属度原则	79
4.2.2	择近原则	79
4.2.3	模糊聚类分析	81
4.3	模糊神经网络	85
4.3.1	模糊神经网络	85
4.3.2	模糊 BP 神经网络	86
4.4	模糊逻辑系统及其在科学研究中的应用	86

若您对此书内容有任何疑问，可以登录 MATLAB 中文论坛与作者和同行交流。

第 5 章 核函数方法及应用	107
5.1 核函数方法	107
5.2 基于核的主成分分析方法	108
5.2.1 主成分分析	108
5.2.2 基于核的主成分分析	110
5.3 基于核的 Fisher 判别方法	112
5.3.1 Fisher 判别方法	112
5.3.2 基于核的 Fisher 判别方法分析	113
5.4 基于核的投影寻踪方法	114
5.4.1 投影寻踪分析	114
5.4.2 基于核的投影寻踪分析	118
5.5 核函数方法在科学研究中的应用	119
第 6 章 支持向量机及其模式识别	130
6.1 统计学习理论基本内容	130
6.2 支持向量机	131
6.2.1 最优分类面	131
6.2.2 支持向量机模型	132
6.3 支持向量机在模式识别中的应用	134
第 7 章 可拓学及其模式识别	142
7.1 可拓学概论	142
7.1.1 可拓工程基本思想	142
7.1.2 可拓工程使用的基本工具	143
7.2 可拓集合	145
7.2.1 可拓集合含义	145
7.2.2 物元可拓集合	146
7.3 可拓聚类预测的物元模型	146
7.4 可拓学在科学研究中的应用	147
第 8 章 粗糙集理论及其模式识别	154
8.1 粗糙集理论基础	154
8.1.1 分类规则的形成	156
8.1.2 知识的约简	157
8.2 粗糙神经网络	158
8.3 系统评估粗糙集方法	158
8.3.1 模型结构	159
8.3.2 综合评估方法	159
8.4 粗糙集聚类方法	160
8.5 粗糙集理论在科学研究中的应用	161
第 9 章 遗传算法及其模式识别	170
9.1 遗传算法的基本原理	170
9.2 遗传算法分析	173

9.2.1	染色体的编码	173
9.2.2	适应度函数	174
9.2.3	遗传算子	175
9.3	控制参数的选择	177
9.4	模拟退火算法	178
9.4.1	模拟退火的基本概念	178
9.4.2	模拟退火算法的基本过程	179
9.4.3	模拟退火算法中的控制参数	180
9.5	基于遗传算法的模式识别在科学研究中的应用	180
9.5.1	遗传算法的 MATLAB 实现	180
9.5.2	遗传算法在科学研究中的应用实例	185
第 10 章 蚁群算法及其模式识别		201
10.1	蚁群算法原理	201
10.1.1	基本概念	201
10.1.2	蚁群算法的基本模型	202
10.1.3	蚁群算法的特点	203
10.2	蚁群算法的改进	203
10.2.1	自适应蚁群算法	203
10.2.2	遗传算法与蚁群算法的融合	204
10.2.3	蚁群神经网络	204
10.3	聚类问题的蚁群算法	205
10.3.1	聚类数目已知的聚类问题的蚁群算法	205
10.3.2	聚类数目未知的聚类问题的蚁群算法	206
10.4	蚁群算法在科学研究中的应用	207
第 11 章 粒子群算法及其模式识别		217
11.1	粒子群算法的基本原理	217
11.2	全局模式与局部模式	218
11.3	粒子群算法的特点	218
11.4	基于粒子群算法的聚类分析	219
11.4.1	算法描述	219
11.4.2	实现步骤	220
11.5	粒子群算法在科学研究中的应用	221
第 12 章 可视化模式识别技术		229
12.1	高维数据的图形表示方法	229
12.1.1	轮廓图	229
12.1.2	雷达图	230
12.1.3	树形图	230
12.1.4	三角多项式图	231
12.1.5	散点图	231
12.1.6	星座图	232

若您对此书内容有任何疑问，可以登录 MATLAB 中文论坛与作者和同行交流。

12.1.7	脸谱图	233
12.2	图形特征参数计算	235
12.3	显示方法	237
12.3.1	线性映射	237
12.3.2	非线性映射	237
第 13 章	灰色系统方法及应用	241
13.1	灰色系统的基本概念	241
13.1.1	灰 数	241
13.1.2	灰数白化与灰度	242
13.2	灰色序列生成算子	242
13.2.1	均值生成算子	242
13.2.2	累加生成算子	243
13.2.3	累减生成算子	243
13.3	灰色分析	244
13.3.1	灰色关联度分析	244
13.3.2	无量纲化的关键算子	244
13.3.3	关联分析的主要步骤	245
13.3.4	其他几种灰色关联度	246
13.4	灰色聚类	247
13.5	灰色系统建模	247
13.5.1	GM(1,1)模型	247
13.5.2	GM(1,1)模型检验	248
13.5.3	残差 GM(1,1)模型	250
13.5.4	GM(1,N)模型	250
13.6	灰色灾变预测	251
13.7	灰色系统的应用	252
第 14 章	人工鱼群等群体智能算法	258
14.1	人工鱼群算法	259
14.1.1	鱼群模式的提出	259
14.1.2	人工鱼的四种基本行为算法描述	259
14.1.3	人工鱼群算法概述	261
14.1.4	各种参数对算法收敛性能的影响	263
14.1.5	人工鱼群算法在科学研究中的应用	264
14.2	人工免疫算法	270
14.2.1	人工免疫算法的生物学基础	270
14.2.2	人工免疫优化算法概述	272
14.2.3	人工免疫算法与遗传算法的比较	276
14.2.4	人工免疫算法在科学研究中的应用	277
14.3	进化计算	281
14.3.1	进化规划算法	283

14.3.2	进化策略算法	284
14.3.3	进化计算在科学研究中的应用	286
14.4	混合蛙跳算法	291
14.4.1	基本原理	291
14.4.2	基本术语	291
14.4.3	算法的基本流程及算子	292
14.4.4	算法控制参数的选择	294
14.4.5	混合蛙跳算法在科学研究中的应用	294
14.5	猫群算法	296
14.5.1	基本术语	296
14.5.2	基本流程	297
14.5.3	控制参数选择	299
14.5.4	猫群算法在科学研究中的应用	299
14.6	细菌觅食算法	300
14.6.1	细菌觅食算法基本原理	301
14.6.2	算法主要步骤与流程	303
14.6.3	算法参数选取	304
14.6.4	细菌觅食算法在科学研究中的应用	306
14.7	人工蜂群算法	307
14.7.1	人工蜂群算法的基本原理	308
14.7.2	人工蜂群算法的流程	309
14.7.3	控制参数选择	311
14.7.4	人工蜂群算法在科学研究中的应用	311
14.8	量子遗传算法	312
14.8.1	量子计算的基础知识	312
14.8.2	量子计算	313
14.8.3	量子遗传算法流程	316
14.8.4	控制参数	318
14.8.5	量子遗传算法在科学研究中的应用	320
14.9	Memetic 算法	321
14.9.1	Memetic 算法的构成要素	321
14.9.2	Memetic 算法的基本流程	322
14.9.3	控制参数选择	322
14.9.4	Memetic 算法在科学研究中的应用	323
第 15 章 仿生模式识别		328
15.1	仿生模式识别基本理论	328
15.1.1	仿生模式识别的连续性规律	328
15.1.2	多自由度神经元	329
15.2	仿生模式识别的数学工具	331
15.2.1	高维空间几何分析基本概念	332

若您对此书内容有任何疑问，可以登录 MATLAB 中文论坛与作者和同行交流。

15.2.2	高维空间中点、线、超平面的关系	333
15.2.3	高维空间几何覆盖理论	334
15.3	仿生模式识别的实现方式	335
15.3.1	高维空间复杂几何形体覆盖	335
15.3.2	多权值神经元的构造	338
15.4	仿生模式识别与传统模式识别的区别	338
15.4.1	认知理论的差别	338
15.4.2	数学模式的差异	339
15.5	仿生模式识别在科学研究中的应用	340
第 16 章	模式识别的特征及确定	348
16.1	基本概念	348
16.1.1	特征的特点	348
16.1.2	特征的类别	348
16.1.3	特征的形成	352
16.1.4	特征选择与提取	353
16.2	样本特征的初步分析	353
16.3	特征筛选处理	357
16.4	特征提取	357
16.4.1	特征提取的依据	357
16.4.2	特征提取的方法	359
16.5	基于 K-L 变换的特征提取	362
16.5.1	离散 K-L 变换	362
16.5.2	离散 K-L 变换的特征提取	363
16.5.3	吸收类均值向量信息的特征提取	363
16.5.4	利用总体熵吸收方差信息的特征提取	364
16.6	因子分析	365
16.6.1	因子分析的一般数学模型	365
16.6.2	Q 型和 R 型因子分析	366
参考文献	372

模式识别(pattern recognition)是当前科学发展中的一门前沿学科,也是一门典型的交叉学科,它的发展与人工智能、计算机科学、传感技术、信息论等学科的研究水平息息相关,相辅相成。

模式识别在数据挖掘、生物特征识别、自动检测、化学、遥感、文本分类、工业自动化、文档与图像分析等领域的应用越来越广泛,同时,模式识别技术的研究也越来越受到各方面的重视,与之相关的市场也越来越庞大,如智能交通管理系统、文字识别系统等都需要模式识别技术的支撑。

1.1 模式识别的基本概念

1.1.1 模式与模式识别的概念

模式识别是人类天生具备的能力。初生婴儿就能辨认自己的父母;大多数 5 岁的孩子便能辨认数字和字母,而且不论其大小、方向、手写体、印刷体,甚至这些数字或字母的一部分被遮住……。随着计算机技术和人工智能技术的迅速发展,人们迫切希望计算机能够完成模式的识别工作,诸如听懂人类所说的话,看懂人类所写的文字,而不论它是手写体还是印刷体……。这些愿望促进了模式识别学科的形成和发展。

模式识别研究的目的是利用计算机对物理对象即“模式”进行分类或描述,在错误率最小的条件下,使识别的结果与客观物体相符合。“模式”是一个内涵十分丰富的概念,可以把凡是人类能用其感官直接或间接接受的外界信息都称为模式,而把具有某些共同特性的模式的集合称为模式类。

1.1.2 模式的特征

在模式识别中,被观察的每个对象称为样本(或样品)。对于一个样本来说,必须确定一些与识别有关的因素,作为研究的根据,每一个因素称为一个特征。例如,医生给病人治病,病人就是样本;疾病特征或病人的一些生理指标即为模式识别的特征。

模式的特征集可用处于同一个特征空间(R^n)的特征向量表示。如果一个样本 \mathbf{X} 有 n 个特征,则 \mathbf{X} 可以看作是一个 n 维的列向量,其中的每个元素称为特征,该向量也因此称为特征向量:

$$\mathbf{X} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = (x_1, x_2, \dots, x_n)^T$$

若有 N 个样本,每个样本具有 n 个特征,则样本集的特征可以用一个 n 行 N 列的矩阵表

