

MATLAB  
开发案例系列图书



# MATLAB

## 统计分析与应用： 40个案例分析

谢中华 编著

 北京航空航天大学出版社

MATLAB 开发实例系列图书

# MATLAB 统计分析与应用： 40 个案例分析

谢中华 编著

北京航空航天大学出版社

## 内 容 简 介

本书从实际应用的角度出发,以大量的案例详细介绍了 MATLAB 环境下的统计分析与应用。本书主要内容包括:利用 MATLAB 制作统计报告或报表;从文件中读取数据到 MATLAB;从 MATLAB 中导出数据到文件;数据的平滑处理、标准化变换和极差归一化变换;生成一元和多元分布随机数;蒙特卡洛方法;参数估计与假设检验;Copula 理论及应用实例;方差分析;基于回归分析的数据拟合;聚类分析;判别分析;主成分分析;因子分析;图像处理中的统计应用等。

本书可以作为高等院校本科生、研究生的统计学相关课程的教材或教学参考书,也可作为从事数据分析与数据管理的研究人员的参考用书。

## 图书在版编目(CIP)数据

MATLAB 统计分析与应用:40 个案例分析/谢中华编

著. —北京:北京航空航天大学出版社,2010.6

ISBN 978-7-5124-0084-9

I. ①M… II. ①谢… III. ①统计分析—计算机辅助  
计算—软件包,MATLAB IV. ①C812

中国版本图书馆 CIP 数据核字(2010)第 079424 号

版权所有,侵权必究。

## MATLAB 统计分析与应用:40 个案例分析

谢中华 编著

责任编辑 陈守平

\*

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:bhpress@263.net 邮购电话:(010)82316936

北京市媛明印刷厂印装 各地书店经销

\*

开本:787×1092 1/16 印张:27.75 字数:710 千字

2010 年 6 月第 1 版 2010 年 6 月第 1 次印刷 印数:4 000 册

ISBN 978-7-5124-0084-9 定价:55.00 元

# 前 言

MATLAB、SAS、Spss、Splus、R 语言等软件都可用作统计计算与分析,在这些软件中, MATLAB 的功能无疑是最强大的。MATLAB 有“草稿纸式”的编程语言,还有包罗万象的工具箱,易学易用,用户不仅可以调用其内部函数作“傻瓜式”的计算,还可以根据自己的算法进行扩展编程。可以说,它就是计算软件中的“航空母舰”。试问读者朋友们,你们是想拥有一艘普通的“战舰”,还是想拥有一艘无所不能的“航空母舰”呢?

在我们的生活中,统计无处不在,大到国计民生,小到个人起居,无不与统计息息相关。与统计有关的论著也如春日繁花。但就目前情况来看,市面上有关统计与 MATLAB 结合的论著并不多见,并且大多只是 MATLAB 统计工具箱的中文翻译,或者在概率论与数理统计的教材里加了一些 MATLAB 代码,它们普遍存在的问题就是缺乏具体的案例分析,并且在统计的应用方面缺乏创新。本书仅以较少篇幅介绍 MATLAB 统计工具箱函数的调用方法,将通过大量的案例分析介绍 MATLAB 在统计方面的应用。本书内容分 12 章,另有 2 个附录,共涉及 40 个大的案例,其中有些大案例下还包含了一些小的案例。本书章节是这样安排的:第 1 章,利用 MATLAB 生成 Word 和 Excel 文档;第 2 章,数据的导入与导出;第 3 章,数据的预处理;第 4 章,生成随机数;第 5 章,参数估计与假设检验;第 6 章, Copula 理论及应用实例;第 7 章,方差分析;第 8 章,数据拟合;第 9 章,聚类分析;第 10 章,判别分析;第 11 章,主成分分析;第 12 章,因子分析;附录 A,图像处理中的统计应用案例;附录 B, MATLAB 统计工具箱函数索引。其中利用 MATLAB 与 Word、Excel 接口技术生成 Word 和 Excel 文档属作者原创,利用这一技术可以很方便的生成各种统计报告或报表。另外,本书还涉及 5 个基于统计方法的图像处理案例,包括从图像资料中提取绘图数据并进行曲线拟合,灰度图像和真彩图像的分割,从固定背景视频中识别运动目标,手写体数字识别,图像压缩等。这些都是其他统计软件很难解决的问题,也是传统教材没有涉及的问题。

笔者长期从事本科生“概率论与数理统计”、“多元统计分析”,硕士研究生“数理统计”,博士研究生“应用数学基础”等课程的教学。在教学中,作者把 MATLAB 引入课堂,深受学生欢迎。本书是作者长期教学经验的总结。

笔者长期活跃于研学论坛、仿真论坛和振动论坛的 MATLAB 版面,以及 MATLAB 中文论坛的各版面,笔者编写的“利用 MATLAB 生成 Word 和 Excel 文档”、“猫追耗子的动画演示”、“概率统计实验演示系统”等 MATLAB 程序在各论坛间广泛流传。笔者认为这些论坛是学习 MATLAB 的好地方,论坛上的很多问题都是经典的、共性的、案例式的。笔者把自己学习 MATLAB 的经历总结成三个词语:“纸上谈兵”,“闭门造车”和“改革开放”。刚接触 MATLAB 时,由于没有电脑可用,只能天天泡在图书馆里看 MATLAB 教程;虽然笔记记了一大本,但是收获甚微,这段经历纯属“纸上谈兵”。后来有了自己的电脑,就以极大的热情投入到 MATLAB 的学习中,编写了大量的 MATLAB 程序,在实践中积累了一些 MATLAB 的使用经验;但是由于缺乏与 MATLAB 高手们的交流,这段经历也只能是“闭门造车”。再后来,由于查资料的缘故,“误入”论坛这片“桃花源”,从此进入了一个新天地;笔者也从“闭门造车”走

向“改革开放”。笔者在论坛里神交了很多高手,体会到了与高手过招的乐趣;这种过招是一种付出和收获共存的过程,通过回答别人的问题提高了自己的能力,通过学习别人的帖子收获了别人的经验。到如今,蓦然回首,发现自己竟也成了别人眼中的“高手”!

针对本书,MATLAB 中文论坛(<http://www.ilovematlab.cn/>)特别提供了读者与作者在线交流的平台,笔者希望在这个平台上与广大读者做面对面的交流,解决大家在阅读此书过程中遇到的问题,分享彼此的学习经验。本书涉及的所有源程序将放到 MATLAB 中文论坛的在线交流平台上,供读者自由下载。这些源程序在 MATLAB R2009a(即 MATLAB 7.8)下经过了验证,均能够正确执行,读者可将自己的 MATLAB 版本更新至 MATLAB R2009a 及其以后的版本,以避免出现不必要的问题。另外,为了使读者能够顺利阅读此书,笔者还编写了“MATLAB 编程简介”和“常用统计理论介绍”等相关的基础内容,也将其放在 MATLAB 中文论坛的在线交流平台上,供读者自由下载。

在本书的写作过程中,笔者得到了北京航空航天大学出版社陈守平编辑、MATLAB 中文论坛创始人 math(张延亮)和研学论坛 MATLAB 版版主 rocwoods(吴鹏)的支持与鼓励,陈守平和 math 编辑为本书提出了宝贵的修改意见。在此,向他们表示最真诚的谢意!

本书的写作得到了天津科技大学理学院和数学系领导及同事们的支持与帮助,还得到了天津科技大学损伤生物力学和车辆安全工程中心的阮世捷、李海岩老师以及赵玮、包永涛、丁成、唐小兵、顾玉龙、姜颖飞等研究生的帮助与鼓励,在此一并表示最诚挚的感谢!

最后,还要感谢我的家人,她们默默地为我付出,支持我顺利完成本书的写作,在此,向我的家人表示最衷心的感谢!

由于作者水平有限,书中的疏漏和不当之处,恳请广大读者和同行批评指正! 作者邮箱: xiezhzh@tust.edu.cn。本书勘误网址: <http://www.ilovematlab.cn/thread-79642-1-1.html>。

谢中华

2010 年 1 月于天津滨海新区

# 目 录

|                                                    |    |
|----------------------------------------------------|----|
| 第 1 章 利用 MATLAB 生成 Word 和 Excel 文档 .....           | 1  |
| 1.1 组件对象模型(COM) .....                              | 1  |
| 1.1.1 什么是 COM .....                                | 1  |
| 1.1.2 COM 接口 .....                                 | 2  |
| 1.2 MATLAB 中的 ActiveX 控件接口技术 .....                 | 3  |
| 1.2.1 actxcontrol 函数 .....                         | 3  |
| 1.2.2 actxcontrollist 函数 .....                     | 4  |
| 1.2.3 actxcontrolselect 函数 .....                   | 5  |
| 1.2.4 actxserver 函数 .....                          | 7  |
| 1.2.5 利用 MATLAB 调用 COM 对象 .....                    | 8  |
| 1.2.6 调用 actxserver 函数创建组件服务器 .....                | 17 |
| 1.3 案例 1:利用 MATLAB 生成 Word 文档 .....                | 20 |
| 1.3.1 调用 actxserver 函数创建 Microsoft Word 服务器 .....  | 20 |
| 1.3.2 建立 Word 文本文档 .....                           | 21 |
| 1.3.3 插入表格 .....                                   | 26 |
| 1.3.4 插入图片 .....                                   | 31 |
| 1.3.5 保存文档 .....                                   | 37 |
| 1.3.6 完整代码 .....                                   | 37 |
| 1.4 案例 2:利用 MATLAB 生成 Excel 文档 .....               | 43 |
| 1.4.1 调用 actxserver 函数创建 Microsoft Excel 服务器 ..... | 43 |
| 1.4.2 新建 Excel 工作簿 .....                           | 43 |
| 1.4.3 获取工作表对象句柄 .....                              | 45 |
| 1.4.4 插入、复制、删除、移动和重命名工作表 .....                     | 45 |
| 1.4.5 页面设置 .....                                   | 47 |
| 1.4.6 选取工作表区域 .....                                | 48 |
| 1.4.7 设置行高和列宽 .....                                | 49 |
| 1.4.8 合并单元格 .....                                  | 49 |
| 1.4.9 边框设置 .....                                   | 50 |
| 1.4.10 设置单元格对齐方式 .....                             | 53 |
| 1.4.11 写入单元格内容 .....                               | 53 |
| 1.4.12 插入图片 .....                                  | 54 |
| 1.4.13 保存工作簿 .....                                 | 59 |
| 1.4.14 完整代码 .....                                  | 59 |

|                           |     |
|---------------------------|-----|
| <b>第 2 章 数据的导入与导出</b>     | 64  |
| 2.1 案例 3:从 TXT 文件中读取数据    | 64  |
| 2.1.1 利用数据导入向导导入 TXT 文件   | 65  |
| 2.1.2 调用高级函数读取数据          | 70  |
| 2.1.3 调用低级函数读取数据          | 82  |
| 2.2 案例 4:把数据写入 TXT 文件     | 92  |
| 2.2.1 调用 dlmread 函数写入数据   | 92  |
| 2.2.2 调用 fprintf 函数写入数据   | 93  |
| 2.3 案例 5:从 Excel 文件中读取数据  | 97  |
| 2.3.1 利用数据导入向导导入 Excel 文件 | 97  |
| 2.3.2 调用 xlsread 函数读取数据   | 98  |
| 2.4 案例 6:把数据写入 Excel 文件   | 101 |
| <b>第 3 章 数据的预处理</b>       | 104 |
| 3.1 案例 7:数据的平滑处理          | 104 |
| 3.1.1 smooth 函数           | 104 |
| 3.1.2 smoohts 函数          | 107 |
| 3.1.3 medfilt1 函数         | 109 |
| 3.2 案例 8:数据的标准化变换         | 111 |
| 3.2.1 标准化变换公式             | 112 |
| 3.2.2 标准化变换的 MATLAB 实现    | 112 |
| 3.3 案例 9:数据的极差归一化变换       | 114 |
| 3.3.1 极差归一化变换公式           | 114 |
| 3.3.2 极差归一化变换的 MATLAB 实现  | 114 |
| <b>第 4 章 生成随机数</b>        | 117 |
| 4.1 案例 10:生成一元分布随机数       | 117 |
| 4.1.1 均匀分布随机数和标准正态分布随机数   | 117 |
| 4.1.2 RandStream 类        | 120 |
| 4.1.3 常见一元分布随机数           | 124 |
| 4.1.4 任意一元分布随机数           | 127 |
| 4.2 案例 11:生成多元分布随机数       | 134 |
| 4.3 案例 12:蒙特卡洛方法          | 136 |
| 4.3.1 有趣的蒙提霍尔问题           | 136 |
| 4.3.2 抽球问题的蒙特卡洛模拟         | 137 |
| 4.3.3 用蒙特卡洛方法求圆周率 $\pi$   | 139 |
| 4.3.4 用蒙特卡洛方法求积分          | 143 |
| 4.3.5 街头骗局揭秘              | 146 |
| <b>第 5 章 参数估计与假设检验</b>    | 150 |
| 5.1 案例 13:常见分布的参数估计       | 150 |
| 5.2 案例 14:正态总体参数的检验       | 151 |



|       |                                     |     |
|-------|-------------------------------------|-----|
| 5.2.1 | 总体标准差已知时的单个正态总体均值的 $U$ 检验 .....     | 151 |
| 5.2.2 | 总体标准差未知时的单个正态总体均值的 $t$ 检验 .....     | 153 |
| 5.2.3 | 总体标准差未知时的两个正态总体均值的比较 $t$ 检验 .....   | 154 |
| 5.2.4 | 总体均值未知时的单个正态总体方差的 $\chi^2$ 检验 ..... | 155 |
| 5.2.5 | 总体均值未知时的两个正态总体方差的比较 $F$ 检验 .....    | 156 |
| 5.3   | 案例 15: 分布的拟合与检验 .....               | 157 |
| 5.3.1 | 案例描述 .....                          | 157 |
| 5.3.2 | 描述性统计量 .....                        | 159 |
| 5.3.3 | 统计图 .....                           | 162 |
| 5.3.4 | 分布的检验 .....                         | 165 |
| 5.3.5 | 最终结论 .....                          | 178 |
| 5.4   | 案例 16: 核密度估计 .....                  | 179 |
| 5.4.1 | 经验密度函数 .....                        | 179 |
| 5.4.2 | 核密度估计 .....                         | 179 |
| 5.4.3 | 核密度估计的 MATLAB 实现 .....              | 182 |
| 5.4.4 | 核密度估计的案例分折 .....                    | 183 |
| 第 6 章 | Copula 理论及应用实例 .....                | 187 |
| 6.1   | Copula 函数的定义与基本性质 .....             | 187 |
| 6.1.1 | 二元 Copula 函数的定义及性质 .....            | 187 |
| 6.1.2 | 多元 Copula 函数的定义及性质 .....            | 188 |
| 6.2   | 常用的 Copula 函数 .....                 | 189 |
| 6.2.1 | 正态 Copula 函数 .....                  | 189 |
| 6.2.2 | $t$ -Copula 函数 .....                | 189 |
| 6.2.3 | 阿基米德 Copula 函数 .....                | 189 |
| 6.3   | Copula 函数与相关性度量 .....               | 191 |
| 6.3.1 | Pearson 线性相关系数 $r$ .....            | 191 |
| 6.3.2 | Kendall 秩相关系数 $\tau$ .....          | 191 |
| 6.3.3 | Spearman 秩相关系数 $\rho_s$ .....       | 192 |
| 6.3.4 | 尾部相关系数 $\lambda$ .....              | 192 |
| 6.3.5 | 基于 Copula 函数的相关性度量 .....            | 193 |
| 6.3.6 | 基于常用二元 Copula 函数的相关性度量 .....        | 193 |
| 6.4   | 案例 17: 沪深股市日收益率的二元 Copula 模型 .....  | 196 |
| 6.4.1 | 案例描述 .....                          | 196 |
| 6.4.2 | 确定边缘分布 .....                        | 197 |
| 6.4.3 | 选取适当的 Copula 函数 .....               | 202 |
| 6.4.4 | 参数估计 .....                          | 203 |
| 6.4.5 | 与 Copula 有关的 MATLAB 函数 .....        | 204 |
| 6.4.6 | 案例的计算与分析 .....                      | 207 |



|                                |     |
|--------------------------------|-----|
| <b>第 7 章 方差分析</b>              | 214 |
| 7.1 案例 18:单因素一元方差分析            | 214 |
| 7.1.1 单因素一元方差分析的 MATLAB 实现     | 214 |
| 7.1.2 案例分析                     | 217 |
| 7.2 案例 19:双因素一元方差分析            | 222 |
| 7.2.1 双因素一元方差分析的 MATLAB 实现     | 222 |
| 7.2.2 案例分析                     | 224 |
| 7.3 案例 21:多因素一元方差分析            | 227 |
| 7.3.1 多因素一元方差分析的 MATLAB 实现     | 227 |
| 7.3.2 案例分析一                    | 228 |
| 7.3.3 案例分析二                    | 233 |
| 7.4 案例 20:单因素多元方差分析            | 237 |
| 7.4.1 单因素多元方差分析的 MATLAB 实现     | 237 |
| 7.4.2 案例分析                     | 239 |
| 7.5 案例 22:非参数方差分析              | 242 |
| 7.5.1 非参数方差分析的 MATLAB 实现       | 242 |
| 7.5.2 Kruskal-Wallis 检验的案例分析   | 245 |
| 7.5.3 Friedman 检验的案例分析         | 249 |
| <b>第 8 章 数据拟合</b>              | 251 |
| 8.1 案例 23:一元线性回归分析             | 251 |
| 8.1.1 数据的散点图                   | 252 |
| 8.1.2 调用 regress 函数作一元线性回归分析   | 253 |
| 8.1.3 调用 regstats 函数作一元线性回归分析  | 258 |
| 8.1.4 调用 robustfit 函数作稳健回归     | 263 |
| 8.2 案例 24:一元非线性回归分析            | 267 |
| 8.2.1 数据的散点图                   | 267 |
| 8.2.2 调用 nlinfit 函数作一元非线性回归分析  | 268 |
| 8.2.3 利用曲线拟合工具 cftool 作一元非线性拟合 | 275 |
| 8.3 案例 25:多重回归分析               | 283 |
| 8.3.1 调用自编 reglm 函数作多重回归分析     | 284 |
| 8.3.2 调用 stepwise 函数作逐步回归      | 287 |
| <b>第 9 章 聚类分析</b>              | 290 |
| 9.1 聚类分析简介                     | 290 |
| 9.1.1 距离和相似系数                  | 290 |
| 9.1.2 系统聚类法                    | 292 |
| 9.1.3 K 均值聚类法                  | 296 |
| 9.1.4 模糊 C 均值聚类法               | 297 |
| 9.2 案例 26:系统聚类法的案例分析           | 298 |

|                                       |     |
|---------------------------------------|-----|
| 9.2.1 系统聚类法的 MATLAB 函数 .....          | 298 |
| 9.2.2 样品聚类案例 .....                    | 307 |
| 9.2.3 变量聚类案例 .....                    | 312 |
| 9.3 案例 27:K 均值聚类法的案例分析 .....          | 316 |
| 9.3.1 K 均值聚类法的 MATLAB 函数 .....        | 316 |
| 9.3.2 K 均值聚类法案例 .....                 | 320 |
| 9.4 案例 28:模糊 C 均值聚类法的案例分析 .....       | 324 |
| 9.4.1 模糊 C 均值聚类法的 MATLAB 函数 .....     | 324 |
| 9.4.2 模糊 C 均值聚类法案例 .....              | 324 |
| <b>第 10 章 判别分析</b> .....              | 329 |
| 10.1 判别分析简介 .....                     | 329 |
| 10.1.1 距离判别 .....                     | 329 |
| 10.1.2 贝叶斯判别 .....                    | 332 |
| 10.1.3 Fisher 判别 .....                | 332 |
| 10.2 案例 29:距离判别法的案例分析 .....           | 334 |
| 10.2.1 classify 函数 .....              | 334 |
| 10.2.2 案例分析 .....                     | 336 |
| 10.3 案例 30:贝叶斯判别法的案例分析 .....          | 340 |
| 10.3.1 NaiveBayes 类 .....             | 340 |
| 10.3.2 案例分析 .....                     | 342 |
| 10.4 案例 31:Fisher 判别法的案例分析 .....      | 345 |
| 10.4.1 Fisher 判别分析的 MATLAB 实现 .....   | 345 |
| 10.4.2 案例分析 .....                     | 350 |
| <b>第 11 章 主成分分析</b> .....             | 354 |
| 11.1 主成分分析简介 .....                    | 354 |
| 11.1.1 主成分分析的几何意义 .....               | 354 |
| 11.1.2 总体的主成分 .....                   | 355 |
| 11.1.3 样本的主成分 .....                   | 357 |
| 11.1.4 关于主成分表达式的两点说明 .....            | 359 |
| 11.2 主成分分析的 MATLAB 函数 .....           | 359 |
| 11.2.1 pcacov 函数 .....                | 359 |
| 11.2.2 princomp 函数 .....              | 359 |
| 11.2.3 pcares 函数 .....                | 360 |
| 11.3 案例 32:从协方差矩阵或相关系数矩阵出发求解主成分 ..... | 361 |
| 11.3.1 调用 pcacov 函数作主成分分析 .....       | 361 |
| 11.3.2 结果分析 .....                     | 363 |
| 11.4 案例 33:从样本观测值矩阵出发求解主成分 .....      | 363 |
| 11.4.1 调用 princomp 函数作主成分分析 .....     | 365 |

|                              |                           |     |
|------------------------------|---------------------------|-----|
| 11.4.2                       | 结果分析                      | 370 |
| 11.4.3                       | 调用 pcacres 函数重建观测数据       | 373 |
| <b>第 12 章 因子分析</b>           |                           | 374 |
| 12.1                         | 因子分析简介                    | 374 |
| 12.1.1                       | 基本因子分析模型                  | 374 |
| 12.1.2                       | 因子模型的基本性质                 | 375 |
| 12.1.3                       | 因子载荷阵和特殊方差阵的估计            | 376 |
| 12.1.4                       | 因子旋转                      | 378 |
| 12.1.5                       | 因子得分                      | 379 |
| 12.1.6                       | 因子分析中的 Heywood 现象         | 380 |
| 12.2                         | 因子分析的 MATLAB 函数           | 380 |
| 12.3                         | 案例 34:基于协方差矩阵或相关系数矩阵的因子分析 | 383 |
| 12.4                         | 案例 35:基于样本观测值矩阵的因子分析      | 389 |
| 12.4.1                       | 读取数据                      | 390 |
| 12.4.2                       | 调用 factoran 函数作因子分析       | 391 |
| <b>附录 A 图像处理中的统计应用案例</b>     |                           | 397 |
| A.1                          | 案例 36:基于图像资料的数据重建与拟合      | 398 |
| A.1.1                        | 案例描述                      | 398 |
| A.1.2                        | 重建图像数据                    | 398 |
| A.1.3                        | 曲线拟合                      | 401 |
| A.2                          | 案例 37:基于 K 均值聚类的图像分割      | 402 |
| A.2.1                        | 灰度图像分割案例                  | 402 |
| A.2.2                        | 真彩图像分割案例                  | 403 |
| A.3                          | 案例 38:基于中位数算法的运动目标检测      | 405 |
| A.3.1                        | 案例描述                      | 405 |
| A.3.2                        | 中位数算法原理                   | 406 |
| A.3.3                        | 本案例的 MATLAB 实现一           | 407 |
| A.3.4                        | 本案例的 MATLAB 实现二           | 408 |
| A.4                          | 案例 39:基于贝叶斯判别的手写体数字识别     | 409 |
| A.4.1                        | 样本图片的预处理                  | 409 |
| A.4.2                        | 创建朴素贝叶斯分类器对象              | 411 |
| A.4.3                        | 判别效果                      | 412 |
| A.5                          | 案例 40:基于主成分分析的图像压缩与重建     | 414 |
| A.5.1                        | 基于主成分分析的图像压缩与重建原理         | 414 |
| A.5.2                        | 图像压缩与重建的 MATLAB 实现        | 415 |
| <b>附录 B MATLAB 统计工具箱函数大全</b> |                           | 419 |
| <b>参考文献</b>                  |                           | 432 |

# 案例目录

|                                   |     |
|-----------------------------------|-----|
| 案例 1:利用 MATLAB 生成 Word 文档 .....   | 20  |
| 案例 2:利用 MATLAB 生成 Excel 文档 .....  | 43  |
| 案例 3:从 TXT 文件中读取数据 .....          | 64  |
| 案例 4:把数据写入 TXT 文件 .....           | 92  |
| 案例 5:从 Excel 文件中读取数据 .....        | 97  |
| 案例 6:把数据写入 Excel 文件 .....         | 101 |
| 案例 7:数据的平滑处理 .....                | 104 |
| 案例 8:数据的标准化变换 .....               | 111 |
| 案例 9:数据的极差归一化变换 .....             | 114 |
| 案例 10:生成一元分布随机数 .....             | 117 |
| 案例 11:生成多元分布随机数 .....             | 134 |
| 案例 12:蒙特卡洛方法 .....                | 136 |
| 案例 13:常见分布的参数估计 .....             | 150 |
| 案例 14:正态总体参数的检验 .....             | 151 |
| 案例 15:分布的拟合与检验 .....              | 157 |
| 案例 16:核密度估计 .....                 | 179 |
| 案例 17:沪深股市日收益率的二元 Copula 模型 ..... | 196 |
| 案例 18:单因素一元方差分析 .....             | 214 |
| 案例 19:双因素一元方差分析 .....             | 222 |
| 案例 21:多因素一元方差分析 .....             | 227 |
| 案例 20:单因素多元方差分析 .....             | 237 |
| 案例 22:非参数方差分析 .....               | 242 |
| 案例 23:一元线性回归分析 .....              | 251 |
| 案例 24:一元非线性回归分析 .....             | 267 |
| 案例 25:多重回归分析 .....                | 283 |
| 案例 26:系统聚类法的案例分析 .....            | 298 |
| 案例 27:K 均值聚类法的案例分析 .....          | 316 |
| 案例 28:模糊 C 均值聚类法的案例分析 .....       | 324 |
| 案例 29:距离判别法的案例分析 .....            | 334 |
| 案例 30:贝叶斯判别法的案例分析 .....           | 340 |

|                                   |     |
|-----------------------------------|-----|
| 案例 31: Fisher 判别法的案例分析 .....      | 345 |
| 案例 32: 从协方差矩阵或相关系数矩阵出发求解主成分 ..... | 361 |
| 案例 33: 从样本观测值矩阵出发求解主成分 .....      | 363 |
| 案例 34: 基于协方差矩阵或相关系数矩阵的因子分析 .....  | 383 |
| 案例 35: 基于样本观测值矩阵的因子分析 .....       | 389 |
| 案例 36: 基于图像资料的数据重建与拟合 .....       | 398 |
| 案例 37: 基于 K 均值聚类的图像分割 .....       | 402 |
| 案例 38: 基于中位数算法的运动目标检测 .....       | 405 |
| 案例 39: 基于贝叶斯判别的手写体数字识别 .....      | 409 |
| 案例 40: 基于主成分分析的图像压缩与重建 .....      | 414 |

# 第 1 章

## 利用 MATLAB 生成 Word 和 Excel 文档

利用 MATLAB 生成统计报告或报表是一件非常有意义的事情。做这个事情的原因是笔者每个学期末都要做好几张试卷分析,工作单调重复,并且周围很多人都有类似的困扰。特别是当统计工笔者面对大量相同格式的统计报表时,所做的工作往往只是重复性劳动,此时就需要一个模板,可以每次自动导入数据,自动生成人们想要的报告,把人们从繁重、重复的工作中解脱出来。基于这个考虑,本章先介绍一点编程理论,然后以案例的形式详细介绍如何利用 MATLAB 生成 Word 和 Excel 文档。读者可以参照本章内容,尝试生成自己想要的统计报告或报表。

本章主要内容包括:微软组件对象模型(COM),基于 COM 技术的 ActiveX 控件接口技术,利用 MATLAB 生成 Word 文档,利用 MATLAB 生成 Excel 文档。

### 1.1 组件对象模型(COM)

#### 1.1.1 什么是 COM

事物的发展总是处在不断的新老交替中,软件的开发也不例外。辛辛苦苦开发出一个比较大的应用程序交给客户使用,过一段时间后应用程序需要升级,需要添加一些新的特性,怎么办?再组织人力物力重新编写所有代码,重新编译?这显然不太现实。为此,微软公司提出了组件对象模型(Component Object Model,COM),它是微软公司为了使软件开发更加符合人的行为方式而提出的一种规范。在这种规范下,单个应用程序被分隔成多个独立的部分,即组件(component)。这种做法的好处是可以随着技术的发展而用新的组件取代已有的组件,此时的应用程序不再是一个一出生就命中注定会过时的静态程序,而是随时可以用新组件取代旧组件而“返老还童”的动态程序。

COM 组件是由以 Win32 动态链接库(DLL)或可执行文件(.exe)的形式发布的可执行代码组成的,其必须满足以下几个条件。

- ① 为了使客户在应用程序的运行中能够将组件替换掉,组件必须动态链接。
- ② 它们必须隐藏(或封装)其内部实现细节。而各组件是通过接口连接在一起的,接口不能变,所以组件要实现封装。
- ③ 封装之后的组件以二进制的形式发布。
- ④ 开发组件的编程语言必须被隐藏起来,也就是说 COM 组件是与语言完全无关。
- ⑤ 组件必须可以在不妨碍已有用户的情况下升级,也就是说一个组件的新版本必须既能够同老版本的用户一起使用,也要能够和新版本用户一起使用。
- ⑥ 组件在网络上的位置必须可以被透明地重新分配。对远程机器上的组件同本地机器上的组件的处理方式没有差别,组件以及使用它的应用程序应能够在同一进程中、不同进程中

或不同的机器上运行。

需要注意的是,COM 是一种软件开发的规范,并不是一种计算机语言。把 COM 与 DLL (动态链接库)相比或相提并论也是不合适的,实际上 COM 是使用了 DLL 来给组件提供动态链接的能力。COM 也不是像 Win32 API 那样的一个函数集。对于 DLL 形式的组件对象,在实际调用时,组件代码被载入应用程序的进程中,所以这种 COM 组件也被称作进程内(in-proc)组件。可执行文件(.exe)形式的组件在实际应用中可以运行于自己独立的进程中,所以也称为进程外(out-of-proc)组件。MATLAB 编译产生的 COM 对象都是 DLL 文件,也就是进程内组件。

## 1.1.2 COM 接口

通过接口可以访问 COM 对象(组件对象类或组件的实例)的属性和方法,这也是访问 COM 对象的唯一途径。对于 COM 对象来说,接口就是一个包含函数指针数组的内存结构,每一个数组元素对应一个由组件所实现函数的地址,也称为虚函数。一个 COM 组件可以有多个接口,客户通过这些接口与组件打交道,但通常客户并不知道组件的全部接口。客户可以通过 COM 对象的基本接口来查询其他的接口。COM 对象有两种类型的基本接口。

### 1. IUnknown 接口

IUnknown 接口是所有接口的基础,其他的接口是 IUnknown 接口的继承。它负责两项工作:

#### (1) 接口查询

IUnknown 中定义了一个名为 QueryInterface 的函数,负责得到该组件的其他接口的指针。

#### (2) 生存期控制

IUnknown 的另外两个成员函数 AddRef 和 Release 负责管理该组件的生存期。AddRef 和 Release 实现的是一种名为引用计数的内存管理技术,COM 组件将维护一个被称为引用计数的数值,当客户从组件取得一个接口时,该引用计数值加 1,当客户用完这个接口时,引用计数值减 1,当引用计数值为零时,组件即可将自己从内存中删除。

### 2. IDispatch 接口

IDispatch 接口将接收一个函数名并执行它。许多应用程序如 Microsoft Word、Microsoft Excel 以及一些解释性语言如 Visual Basic 和 Java 等在控制组件时都采用了自动化的方法。自动化是建立在 COM 基础上的,一个自动化服务器实际上就是一个实现了 IDispatch 接口的 COM 组件,而一个自动化控制器则是一个通过 IDispatch 接口同自动化服务器进行通信的 COM 客户。自动化控制器不会直接调用自动化服务器实现那些函数,而是通过 IDispatch 接口的成员函数实现对服务器中函数的间接调用。

IDispatch 中有两个主要的函数 GetIDsOfNames 和 Invoke,其中 GetIDsOfNames 将读取一个函数的名称并返回其调度标识(一个长整数),每一函数对应唯一的调度标识。为执行某个函数,自动化控制程序将该函数的调度标识传给 Invoke 函数,Invoke 可以将调度标识作为函数指针数组的索引,这一点同常规 COM 接口是类似的。



## 1.2 MATLAB 中的 ActiveX 控件接口技术

ActiveX 控件技术是建立在 COM 之上,由微软公司推出的共享程序数据和功能的技术。据微软权威的软件开发指南 MSDN(microsoft developer network)的定义,ActiveX 控件以前也叫做 OLE 控件或 OCX 控件,它是一些组件,可以将其插入到 web 网页或其他应用程序中。该组件的客户接口使它可以对客户的行为作出响应,它有属性、方法、事件三个特征,其中属性决定了显示效果,方法是用来完成特定的任务所提供的外部调用函数,事件一般是用户启动的某种行为,用来给应用控件的程序发送消息,以做出某种响应动作。

MATLAB 作为一种面向对象的编程语言,它支持 COM 技术,利用 MATLAB 的 COM 编译器可以把 MATLAB 开发的应用程序转换成方便使用的 COM 组件。可以利用微软 ActiveX 控件技术,在 MATLAB 中调用其他应用程序,比如 Microsoft Word 或 Microsoft Excel 等,此时 Microsoft Word 或 Microsoft Excel 被作为组件,是服务器程序,MATLAB 是控制器程序,是客户。当然也可以把 MATLAB 作为组件,用其他应用程序(例如 Microsoft Excel)进行调用,这时候 MATLAB 成了服务器,其他应用程序成了客户。

本节只介绍 MATLAB 作为 COM 客户,利用 MATLAB 调用其他组件。MATLAB 中提供了 actxcontrol 函数用来在当前图形窗口中创建 ActiveX 控件,提供了 actxcontrollist 函数用来查看系统上当前安装的所有 COM 控件,提供了 actxcontrolselect 函数用来通过图形界面的方式在图形窗口中创建控件,还提供了 actxserver 函数用来创建 COM 自动化服务器。下面分别进行详细介绍。

### 1.2.1 actxcontrol 函数

actxcontrol 函数用来在当前图形窗口中创建 ActiveX 控件,其调用格式有以下几种:

1) `h = actxcontrol('progid')`

在图形窗口中创建一个 ActiveX 控件,返回控件的默认接口 h。控件的程序标识符 progid 决定了所创建控件的类型,它是一个字符串,由控件供应商提供。

2) `h = actxcontrol('progid','param1',value1,...)`

利用可选的成对出现的参数名称(param)和参数值(value)创建一个 ActiveX 控件。可选的参数名包括:position(位置矢量,见第 3 种调用格式中的说明),parent(指向父窗口、模型或命令窗口的句柄),callback(响应事件句柄名称),filename(设定控件的初始状态为以前保存的控件状态),licensekey(ActiveX 控件的授权许可密钥)。

3) `h = actxcontrol('progid', position)`

创建一个 ActiveX 控件,其位置和大小由矢量 position 指定,position 的格式为[left, bottom, width, height],其中 left 和 bottom 为控件左下角相对于图形窗口左下角的偏移距离,width 和 height 分别表示控件的宽度和高度,它们的单位均为像素。

4) `h = actxcontrol('progid', position, fig_handle)`

在句柄值为 fig\_handle 的图形窗口中指定位置 position 处创建一个 ActiveX 控件,默认图形窗口句柄为 gcf(当前图形窗口)。

5) `h = actxcontrol('progid',position,fig_handle,event_handler)`

创建一个响应事件的 ActiveX 控件。当事件被触发时,控件通过调用一个 M 函数来响应事件。输入参数 event\_handler 用来指定事件所对应的 M 函数,它可以是函数名字符串或函数句柄。

6) `h = actxcontrol('progid',position,fig_handle,event_handler,'filename')`

用前 4 个参数创建一个 ActiveX 控件,将它的初始状态设置为与前面保存的控件一样。MATLAB 从 filename 指定的文件中载入初始状态。如果不想指定事件控制器参数 event\_handler,可以用空字符串作为第 4 个输入,progid 参数必须与所保存的控件的 progid 相匹配。

event\_handler 参数有多种合法格式:

①用一个元胞数组{'event','eventhandler';'event2','eventhandler2';...}作为 event\_handler 参数的值,数组的每一行指定一个事件/控制函数对。'event'可以是包含事件名称的字符串,或者作为事件标识符的数值,'eventhandler'是事件所对应的 M 函数的函数名字符串或函数句柄。这样可以为控件支持的每个事件指定一个不同的事件控制过程。

②仍使用元胞数组,但是每一个'event'对应的'eventhandler'都相同。这样可以为控件支持的每个事件指定一个相同的事件控制过程。

③不用元胞数组,而是用一个函数名字符串或函数句柄作为 event\_handler 的值,这样就指定了一个公共过程来控制所有事件。

**【例 1.1】** 调用 actxcontrol 函数创建一个日历控件,如图 1-1 所示。

```
% 新建一个图形窗口,指定图形窗口大小,返回图形窗口句柄 f
>> f = figure('position',[360 278 535 410]);
% 在新建的图形窗口中创建一个日历控件,并设置控件的大小
>> cal = actxcontrol('mscal.calendar',[0 0 535 410],f)

cal =

    COM.mscal_calendar
```

本例创建了一个日历控件,并返回了控件的句柄值 cal,可以通过 get(cal)访问该控件的所有属性,也可以通过 set 函数修改控件属性值。关于 get 和 set 函数的调用格式,后面将作讨论。

## 1.2.2 actxcontrollist 函数

actxcontrollist 函数用来查看系统上当前安装的所有 COM 控件。

```
>> C = actxcontrollist % 查看系统上当前安装的所有 COM 控件
```

```
C =

    'Adobe PDF Reader'    'AcroPDF.PDF.1'    [1x54 char]
           [1x29 char]           [1x21 char]    [1x57 char]
    'COMNSView Class'    'COMSNAP.COMNSView.1'    [1x31 char]
    'CTreeView 控件'    [1x25 char]    [1x29 char]
           [1x45 char]           [1x21 char]    [1x66 char]
    ...
```