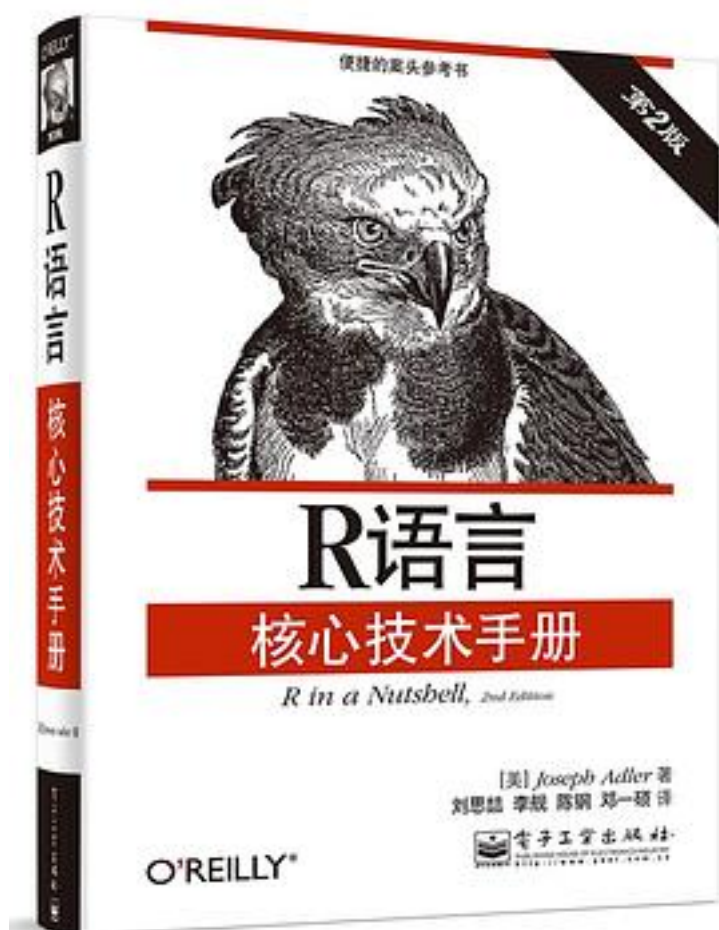


R语言核心技术手册（第2版）



[R语言核心技术手册（第2版）_下载链接1](#)

著者:Joseph Adler（约瑟夫 阿德勒）

出版者:电子工业出版社

出版时间:2014-7-1

装帧:

isbn:9787121237867

R
是一款优秀的开源统计应用语言，它直观、易用、低成本，而且还有庞大的社区支持，

随着数据挖掘技术的兴起，R语言得到了广泛的应用。《R语言核心技术手册（第2版）》介绍从安装R软件到基本语法以及应用的全过程，可以帮助你全面地学习和使用R。《R语言核心技术手册（第2版）》共6部分，26章，基本涵盖了R语言的所有功能，而且提供了大量的实例说明运用R语言绘图、分析数据以及拟合统计模型的过程。书中虽然涉及很多统计学理论和知识，但并不是《R语言核心技术手册（第2版）》的重点。

作为第2版，本书增加了一些处理数据的新章节，将绘图章节集中放在“可视化篇”，同时针对R的版本变化做了一些升级。

《R语言核心技术手册（第2版）》适合从事数据挖掘、机器学习、统计及数据分析的人士阅读。

作者介绍:

作者：约瑟夫·阿德勒（Joseph Adler）拥有多年数据挖掘和数据分析经验，曾就职于DoubleClick、美国运通和VeriSign公司。约瑟夫毕业于麻省理工学院，期间取得了计算机科学和电子工程的学士及硕士学位。他拥有多项计算机安全和密码学领域的专利，并且是《棒球技巧》（Baseball Hacks）的作者。目前，他在LinkedIn公司任高级数据科学家。

译者：刘思喆 中国人民大学统计学院科班出身，国内资深R领域专家，拥有10年R语言使用经验，中国R语言会议联合发起人。《153分钟学会R》作者，《R reference card》译者。“统计之都”理事会成员，R语言版版主；电信、互联网、彩票行业资深数据挖掘专家。CSDN大数据技术论坛、DATA World Forum、中国人民大学数据挖掘中心特邀讲师。

目录: I 基础篇	
第1章 获取和安装R	3
R版本	3
R的安装	3
Windows	4
Mac OS X	5
Linux和UNIX系统	5
第2章 R的用户界面	7
R的图形用户界面	7
Windows	8
Mac OS X	8
Linux和UNIX程序	9
R控制台	11
命令行编辑	13
批处理模式	14
在Excel中使用R	15
RStudio	17
其他运行R的方式	17
第3章 简短的示例	19
基本操作	19
函数	21
变量	22
数据结构简介	25

- 对象和类28
- 模型和公式30
- 图表32
- 获得帮助37
- 第4章R包39
 - R包概览39
 - 列示本地库中的R包40
 - 加载R包42
 - 在Windows和Linux系统下加载R包42
 - 在Mac OS X系统下加载R包42
 - 搜索R包资源库43
 - 搜索网络上的R包资源库44
 - 基于R界面搜寻和安装包44
 - 从其他资源库安装R包47
 - 定制R包47
 - 创建包目录48
 - 创建R包49
- II 语言篇
 - 第5章R语言概览53
 - 表达式53
 - 对象54
 - 符号54
 - 函数55
 - 在赋值语句中，对象会被复制56
 - R中一切皆为对象57
 - 特殊值58
 - NA 58
 - Inf和-Inf 58
 - NaN 59
 - NULL 59
 - 强制转换59
 - R解释器60
 - 观察R是如何工作的62
 - 第6章R语法65
 - 常量65
 - 数值向量65
 - 字符向量67
 - 符号67
 - 运算符69
 - 运算顺序70
 - 赋值操作71
 - 表达式72
 - 分离型表达式72
 - 括号72
 - 花括号73
 - 控制结构74
 - 条件语句74
 - 循环75
 - 访问数据结构79
 - 数据结构操作符79
 - 通过整数向量引用79
 - 通过逻辑向量引用82
 - 通过名字进行引用83
 - R编程标准84

- 第7章R对象 86
 - 基本对象类型86
 - 向量89
 - 列表91
 - 其他对象91
 - 矩阵92
 - 数组92
 - 因子93
 - 数据框95
 - 公式96
 - 时间序列98
 - Shingle 对象99
 - 日期和时间对象100
 - 连接对象100
 - 属性101
 - 类104
 - 第8章符号和环境 106
 - 符号106
 - 环境107
 - 全局环境108
 - 环境和函数109
 - 调用堆栈109
 - 在不同的环境中对函数求值110
 - 向环境中添加对象113
 - 异常113
 - 提示错误114
 - 捕获错误115
 - 第9章函数 117
 - 函数的关键字117
 - 参数117
 - 返回值119
 - 函数参数119
 - 匿名函数120
 - 函数的属性122
 - 参数顺序和具名实参124
 - 副作用125
 - 改变其他环境125
 - 输入/输出126
 - 图形126
 - 第10章面向对象编程 127
 - R的面向对象编程概览128
 - 核心概念128
 - 实现的例子129
 - R的面向对象编程：S4 135
 - 类的定义135
 - 对象的新建137
 - 槽的存取138
 - 对象的操作138
 - 创建强制转换方法138
 - 方法139
 - 方法的管理140
 - 基本类型141
 - 更多的帮助142
 - 守旧派的OOP：S3 142

- S3 的类143
- S3 方法144
- 在S4 的类中使用S3 的类145
- 查找隐藏的S3 方法145
- III 数据篇
- 第11 章数据的存取和编辑 149
- 在R 中输入数据149
- 用R 命令输入数据149
- 用图形界面输入数据150
- 保存和读入R 对象153
- 用save 保存对象153
- 从外部文件导入数据154
- 文本文件155
- 其他软件163
- 导出数据164
- 从数据库获取数据164
- 导出然后导入165
- 数据库连接包165
- RODBC 166
- DBI 177
- TSDBI 181
- 从Hadoop 中获取数据182
- 第12 章准备数据 183
- 合并数据集183
- 粘贴数据结构184
- 通过共同字段合并数据188
- 数据转换189
- 变量重新赋值190
- 转换函数190
- 对对象的每个元素进行函数运算191
- 数据分段197
- shingle 197
- Cut 197
- 利用分组变量合并对象198
- 子集199
- 中括号索引的方式199
- subset 函数200
- 随机抽样200
- 汇总函数202
- tapply 与aggregate 202
- 用rowsum 聚合表格205
- 计数206
- 数据修整209
- 数据清洗218
- 查找和删除重复数据219
- 排序219
- IV 可视化篇
- 第13 章图形 225
- R Graphics 概述225
- 散点图226
- 时间序列231
- 柱状图233
- 饼图238
- 分类数据绘图239

- 三维数据245
- 绘制分布图252
- 箱线图256
- 画图设备259
- 自定义图形260
- 绘图函数常见参数260
- 图形参数260
- 基本图形函数271
- 第14 章Lattice 绘图280
 - 历史280
 - lattice 包概述281
 - lattice 的工作原理281
 - 例子281
 - 使用lattice 函数284
 - 定制面板函数285
 - 高级lattice 函数286
 - 单一的网格作图287
 - 二元网格作图312
 - 三元图321
 - 其他图形326
 - 定制lattice 图328
 - lattice 函数的常用参数328
 - trellis.skeleton 329
 - 指定如何绘制坐标轴330
 - 参数331
 - plot.trellis 336
 - strip.default 337
 - simpleKey 338
 - 低级函数339
 - 低级绘图函数339
 - 面板函数340
- 第15 章ggplot2 341
 - 一个简短的介绍341
 - 图形语法345
 - 一个更复杂的例子：医保数据350
 - 快速绘图360
 - 用ggplot2 绘图361
 - 更多信息365
- V 统计篇
 - 第16 章数据分析369
 - 描述性统计369
 - 相关系数和协方差372
 - 主成分分析375
 - 因子分析379
 - bootstrap 重抽样380
 - 第17 章概率分布382
 - 正态分布382
 - 常见分布的参数384
 - 分布函数族385
 - 第18 章统计检验389
 - 连续型数据390
 - 基于正态分布的检验390
 - 不依赖分布的检验405
 - 离散数据408

比例检验408
二项式检验409
列联表检验410
列联表非参数检验416
第19章功效检验417
实验设计示例417
t 检验实验设计418
比例实验设计419
方差分析设计420
第20章回归模型422
简单的线性模型示例422
拟合模型424
指定模型的工具函数425
获取模型信息425
更新模型431
lm 函数的详述432
最小二乘回归的假设434
稳健回归和阻力回归436
子集选取和Shrinkage 回归438
变量的逐步选取438
岭回归439
Lasso 和最小角回归440
弹性网络442
主成分回归和偏最小二乘回归442
非线性模型443
广义线性模型443
glmnet 包446
非线性最小二乘449
生存模型450
平滑456
样条线456
拟合多项式曲面458
核平滑459
回归的机器学习算法460
回归树模型462
MARS 算法473
神经网络479
投影寻踪回归483
广义可加模型486
支持向量机488
第21章分类模型 490
线性分类模型490
logistic 回归490
线性判别分析495
对数线性模型499
机器学习分类模型500
k 近邻500
分类树模型502
神经网络506
支持向量机507
随机森林507
第22章机器学习509
购物篮分析509
聚类514

距离度量514
聚类算法515
第23章时间序列分析519
自相关函数519
时间序列模型520
VI 其他主题
第24章优化R程序性能527
R程序性能的测量527
时间测定527
性能分析528
监控内存的使用530
内存性能分析531
优化你的R代码531
使用向量操作531
R中查找的性能534
使用数据库查询大数据集541
内存预分配541
清理内存541
大数据集的函数543
加速R的其他方法543
R字节码编译器543
高性能的R版本546
第25章Bioconductor 551
例子552
加载原始的表达数据552
从GEO读取数据557
匹配表型数据559
分析表达数据560
关键的Bioconductor包564
数据结构568
eSet 568
AssayData 570
AnnotatedDataFrame 571
MIAME 571
Bioconductor包使用的其他类572
如何进一步学习573
Bioconductor之外的资源574
教程574
课程574
相关图书575
第26章R和Hadoop 576
R和Hadoop 576
Hadoop简介577
RHadoop 582
Hadoop streaming 597
了解更多600
一些其他的用R做并行计算的包601
Segue 601
doMC 602
从哪里我们可以了解更多602
参考文献603
索引606
• • • • • ([收起](#))

标签

R

R语言

数据分析

计算机

统计

数据挖掘

数学/统计/数据

编程

评论

R语言扛鼎之作！！！！

买了本实体书，感觉非常好！觉得可以让很多R语言的书失去市场了。

速读

工具书，什么都提到一点，但都讲得不深…

工具书已入

工具书啊工具书。

内容多，但都比较浅，大致浏览一下，也只能当工具书用了

关键内容有些简练，实现有扎实的概率 统计
和机器学习的底子会好一点，但是基础部分讲的很好，适合初学者

不错的手册，之前有在Coursera上刷课，刷了十几门感觉还是比较懵懂，不得不说需要更系统的学习。常读常新。工具书。

入门看看

当手册翻翻还行

译者很好，都是优秀的人，然而书并未感到很有用，还是R自己的help好……

数据可视化…从入门到放弃有一说一是一本好的工具书

做入门书挺好的！

Cookbook类，放在案头备查的工具书

需要与被需要之间存在一种矛盾关系，一旦成为事果，必定有无数的事因。

整体还行，会有一些R语法的讲解，不是只有统计部分

适合和《R语言实战》一起看,适合查阅

书内容全面 讲得也比较细致 定位案头书 恰到好处

厚

[R语言核心技术手册（第2版）_下载链接1](#)

书评

1、in a nutshell 是简单说的意思，开始时以为nutshell是一种计算机语言，还以为该书比较难，所以一直搁在那里没看。其实很简单，作为一本入门级的书非常适合，书中对一些概念性的东西做了详细的介绍，比较赋值符号“<-”,packages,如何写packages,函数的语法等，比《The R bo...

[R语言核心技术手册（第2版）_下载链接1](#)