

PACKT
PUBLISHING

异步图书
www.epubit.com

使用R优化交易策略，并建立自己的管理策略

量化金融R语言 高级教程

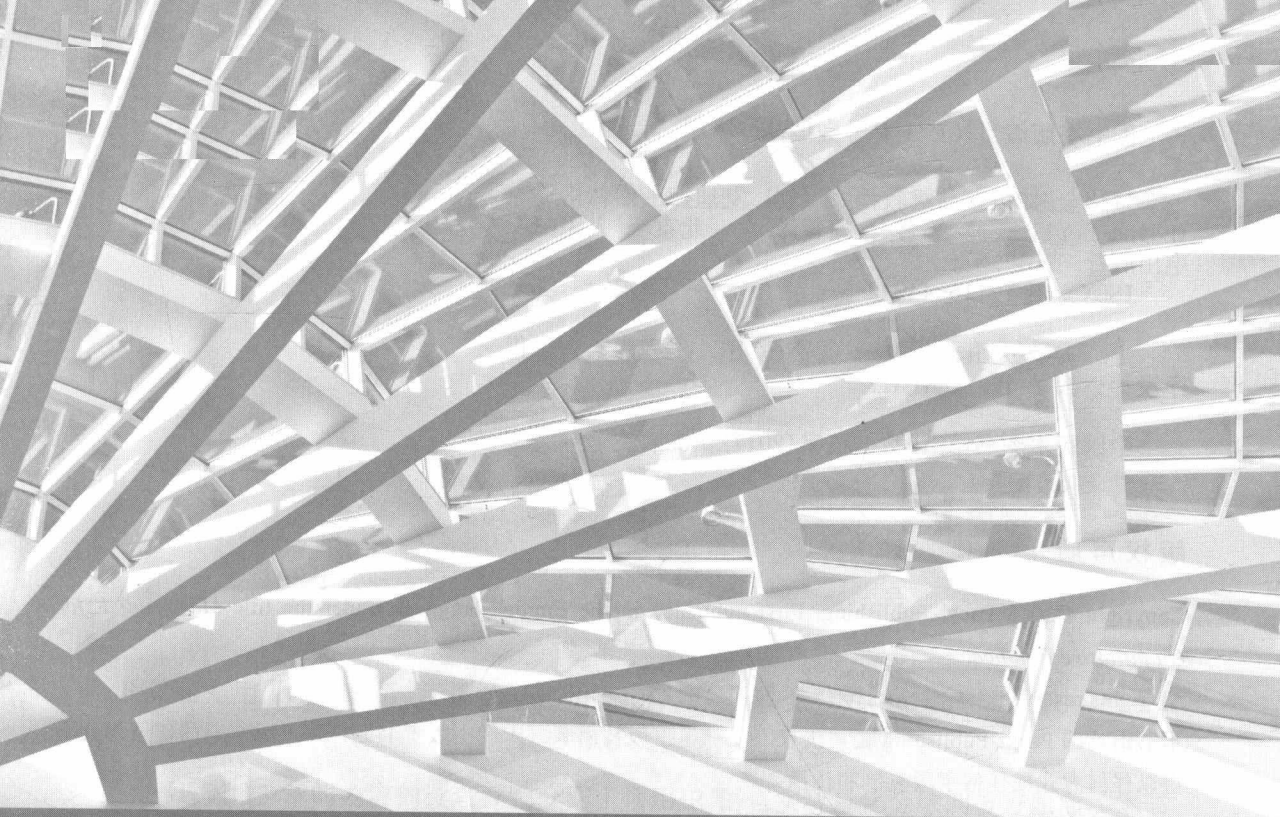
Mastering R for
Quantitative Finance

[匈牙利] Edina Berlinger 等 著
高蓉 译

中国工信出版集团



人民邮电出版社
POSTS & TELECOM PRESS



量化金融R语言 高级教程

[匈牙利] Edina Berlinger 等 著
高蓉 译

人民邮电出版社
北 京

图书在版编目 (C I P) 数据

量化金融R语言高级教程 / (匈) 艾迪娜·伯林格
(Edina Berlinger) 等著 ; 高蓉译. — 北京 : 人民邮
电出版社, 2017.5
ISBN 978-7-115-44982-5

I. ①量… II. ①艾… ②高… III. ①程序设计—教材 IV. ①TP312

中国版本图书馆CIP数据核字(2017)第054188号

版权声明

Copyright ©2015 Packt Publishing. First published in the English language under the title *Mastering R for Quantitative Finance*.

All rights reserved.

本书由英国 Packt Publishing 公司授权人民邮电出版社出版。未经出版者书面许可, 对本书的任何部分不得以任何方式或任何手段复制和传播。

版权所有, 侵权必究。

-
- ◆ 著 [匈牙利] Edina Berlinger 等
 - 译 高 蓉
 - 责任编辑 胡俊英
 - 责任印制 焦志炜
 - ◆ 人民邮电出版社出版发行 北京市丰台区成寿寺路 11 号
 - 邮编 100164 电子邮件 315@ptpress.com.cn
 - 网址 <http://www.ptpress.com.cn>
 - 北京市艺辉印刷有限公司印刷
 - ◆ 开本: 800×1000 1/16
 - 印张: 18.25
 - 字数: 350 千字 2017 年 5 月第 1 版
 - 印数: 1—2 500 册 2017 年 5 月北京第 1 次印刷
 - 著作权合同登记号 图字: 01-2016-3960 号
-

定价: 79.00 元

读者服务热线: (010)81055410 印装质量热线: (010)81055316

反盗版热线: (010)81055315

广告经营许可证: 京东工商广字第 8052 号

内容提要

R 语言是用于统计分析、绘图的语言和操作环境，是属于 GNU 系统的一个自由、免费、源代码开放的软件。它是一个用于统计计算和统计制图的优秀工具。

本书通过 13 章的内容向读者详细介绍了使用 R 语言实现量化金融的方方面面。本书包括实证金融（第 1~4 章）、金融工程（第 5~7 章）、交易策略优化（第 8~10 章）和银行管理（第 10~13 章）等主题。

本书的目标读者是那些既熟悉基本金融概念又具有一定编程能力的人。通过阅读本书，读者可以了解 R 语言与量化金融相关的各类知识和编程技巧。

译者序

自 20 世纪 50 年代马可维茨开创投资组合理论以来，金融学开始与数学模型和数据分析紧密融合，逐渐形成了高度数理化的现代金融学。时至今日，无论是对传统的金融学理论，还是对如日中天的量化投资，又或是对方兴未艾的金融科技（Fintech），数学模型与数据科学源源不断地为这些金融理论和实践提供着发展工具，一起构建了复杂深奥、丰富多彩的量化金融世界。由于数学模型与数据分析本身的复杂性，量化金融技术一直被公众视为数学家或华尔街的“火箭科学家”的秘密武器。

近年来，开源软件得到了迅速推广，促进了各行各业对数据科学的学习与应用。R 就是这样一种开源数据分析工具。它灵活易用，数据分析能力强大，从技术上降低了量化金融技术的学习和应用难度，目前在世界范围内被广泛地采用，在中国也拥有大量用户。但是，可以结合量化金融与数据科学的图书，在市场上依然很少，而本套书正是基于这种需求而作的关于量化金融的学习教程。本套书选择 R 语言作为工具，通过 R 语言学习量化金融，帮助读者快速掌握知识并加以实践应用，为各种金融问题提供实践解决方案，同时还可使用第三方贡献的免费 R 包。

本套书内容广泛、取舍精当，涵盖了实证金融、金融工程、交易策略与银行管理等内容，可以帮助读者全面学习金融理论与技术。本套书既包含金融时间序列分析、资产定价和期权定价等理论知识，也包括投资组合管理、信用风险管理等实践知识，还涉及金融学的前沿领域以及 2008 年金融危机后发展起来的金融网络分析理论。本套书包括两本，分别是《量化金融 R 语言初级教程》和《量化金融 R 语言高级教程》。学习《量化金融 R 语言初级教程》不需要读者精通 R 和金融理论，但需要对金融领域具有一定了解。学习《量化金融 R 语言高级教程》要求读者具备一定的 R 编程能力和金融学基本概念，建议读者在学

习“初级教程”的基础上再学习“高级教程”。

作为一名金融学教师，我认为本套书是通向量化金融世界的一把“金钥匙”，可以有效地促进金融知识和计算技术的传播。我很荣幸能够承担本书的翻译工作。在此特别感谢胡俊英编辑，胡编辑认真专业的审核工作，有力保证了本书翻译的如期完成，我亦受益匪浅。同时特别感谢天津理工大学的李茂老师，书中多处的专业词汇敲定，都来自与李茂老师的讨论结果。当然，文责自负。同时，感谢杭州电子科技大学 2016 年高等教育研究资助项目 YB201631 “投资学教学与 R 软件应用”的支持。

——高蓉，2017 年 3 月

译者简介

高蓉，任教于杭州电子科技大学经济学院。博士毕业于南开大学经济学院，本科毕业于南开大学数学科学学院。研究领域包括资产定价、实证金融、数据科学应用。已出版教材《实验投资学》，译著《数据科学入门》，发表学术论文数篇。

作者简介

Edina Berlinger 拥有布达佩斯考文纽斯大学的博士学位。她是一位助理教授，讲授公司财务、金融学和金融风险管理。她还是大学金融系的领导，也是匈牙利科学院金融分会的主席。她的专业领域涉及信贷系统、风险管理以及最近的网络分析。她已经领导过几个研究项目：学生贷款设计、流动性管理、异质代理模型和系统风险。

“本工作由匈牙利科学院的动量项目（LP-004/2010）支持。”

Ferenc Illés 拥有罗兰大学的数学硕士学位。毕业之后的一些年中，他开始研究精算和金融数学，而且他即将开始在布达佩斯考文纽斯大学的博士学习。最近几年，他在银行业工作。目前他正在开发使用 R 的统计模型。他的兴趣与大型网络以及计算复杂性有关。

Milón Badics 拥有布达佩斯考文纽斯大学的硕士学位。现在，他是一名博士生，并且是 PADS 博士奖学金项目的成员。他讲授金融计量学，而且他的研究主题是使用数据挖掘方法的时间序列预测、金融信号处理以及利率模型的数值敏感分析。2014 年 5 月，他在由匈牙利证券交易所组织的 X. Kochmeister 奖项的竞赛中获胜。

Ádám Banai 从布达佩斯考文纽斯大学得到投资分析和风险管理的硕士学位。他加入了匈牙利国家银行（Magyar Nemzeti Bank, MNB, 匈牙利的中央银行）的金融稳定性部门。从 2013 年起，他成为金融系统分析理事会（MNB）应用研究和压力测试部门的领导。自 2011 年起，他也是布达佩斯考文纽斯大学的博士生。他的主要研究领域是偿付能力压力测试、资金流动性风险和系统风险。

Gergely Daróczi 是一位狂热的 R 包开发者，并且是一家位于 Rapporter 的 R 网络应用公司的创始人和 CTO。他同时也在攻读社会学博士学位，并且目前作为 R 开发者领导在洛杉矶的 CARD.com 工作。如果算上讲授统计学和从事数据分析项目的几年时间，他大约已

经有 10 年的 R 编程环境的工作经验。Gergely 是《量化金融 R 语言初级教程》(*Introduction to R for Quantitative Finance*) 的合著者, 目前除了一些关于社会科学的杂志文章和报告, 他同时还忙于另一本 Packt 出版社的图书《精通 R 语言数据分析》(*Mastering Data Analysis with R*)。他对那本书的贡献是审阅并负责 R 源代码的格式。

Barbara Dömötör 是布达佩斯考文纽斯大学金融系的一名助理教授。在 2008 年开始博士学习之前, 她曾为多家跨国银行工作。她的博士论文与公司的套期保值有关。她撰写了关于公司财务、金融风险管理和投资分析的讲义。她的主要研究领域是公司财务、金融风险管理和公司的套期保值。

Gergely Gabler 自 2014 年起是匈牙利国家银行 (MNB) 金融监管单位的商业模型分析部门领导。在这之前, 自 2008 年起, 他曾经是匈牙利 Erste 银行宏观经济研究部门的领导人。他在 2009 年毕业于布达佩斯考文纽斯大学, 并获得金融数学的硕士学位。自 2010 年起, 他在布达佩斯考文纽斯大学任客座讲师, 同时也在 MCC 学院做高等研究的讲座。他预计会在 2015 年结束 CFA 考试, 并成为一名持证入。

Dániel Havran 是一名匈牙利科学院经济和区域研究中心经济研究所的博士后研究人员。他同时作为布达佩斯考文纽斯大学兼职助理教授, 在那里他讲授公司财务 (本科、博士) 以及信用风险管理 (硕士)。在 2011 年, 他获得了布达佩斯考文纽斯大学的经济学博士学位。

“我非常感谢匈牙利科学院博士后奖学金计划的支持。”

Péter Juhász 拥有布达佩斯考文纽斯大学的工商管理博士学位, 同时也持有 CFA 证书。作为一名助理教授, 他讲授公司财务、商业估值、Excel 的 VBA 编程以及沟通技巧。他的研究领域涉及无形资产的估值、商业表现分析和建模以及政府采购和体育管理。他曾写过一些文章和书的某些章节, 主要关于匈牙利公司的财务表现。同时, 他也定期为中小企业服务, 而且在安永商业学院的 EMEA (欧洲、中东和非洲) 区域任高级培训师。

István Margitai 是 CEE (中东欧) 区域一家主要银行集团的资产负债管理团队的分析师。他主要处理方法论问题、产品建模以及内部转移定价等主题。在 2009 年, 他开启了在匈牙利的资产负债管理的职业生涯, 并收获了战略流动性管理和流动性计划的经验。他在布达佩斯考文纽斯大学主修投资和风险管理。他的研究兴趣是银行业的微观经济学、市场微观结构以及订单驱动市场的流动性。

Balázs Márkus 从事金融衍生品工作已经超过 10 年。他曾经交易过多种类型的衍生品, 从碳互换到国债期货的期权。他是布达佩斯 Raiffeisen 银行外汇衍生品部门的领导。他是 Pallas Athéné Domus 环境科学基金会顾问委员会的一员、匈牙利国家银行的兼职分析师以

及一家小型的证券自营和顾问公司 Nitokris 有限公司的常务董事。目前，他正在布达佩斯考尔纽斯大学攻读动态对冲作用的博士学位，同时他还是那里的一名教学助理。

Péter Medvegyev 拥有布达佩斯 Marx Károly 大学的经济学硕士学位。在 1977 年毕业之后，他开始了匈牙利管理发展中心的顾问工作。他在 1985 年获得了经济学博士学位。自 1993 年开始，他为布达佩斯考文纽斯大学数学系工作。他在考文纽斯大学的教学经历涵盖随机过程、数理金融以及其他多门数学专业课。

Julia Molnár 是布达佩斯考文纽斯大学的一名博士学位候选人。她的主要研究兴趣包括金融网络、系统风险以及零售银行业的金融技术创新。自 2011 年起，她为 McKinsey & Company 工作，在那里她参与了银行业领域的多项数字和创新研究。

Balázs Árpád Szűcs 是布达佩斯考文纽斯大学的金融学博士生，并同时在该大学的金融系任研究助理。他拥有投资分析和风险管理的硕士学位。他的研究兴趣包括最优执行、市场微观结构和日内交易量预测。

Ágnes Tuza 拥有布达佩斯考文纽斯大学的应用经济学学位，而且是巴黎高等商学院（HEC Paris）的转学生。她的工作经验包括为摩根斯坦利从事结构化产品估值，同时承担波士顿咨询集团的管理顾问一职。她是一名活跃的外汇交易者，并且为 Gazdaság 电视台拍摄了一个月的投资思想的直播，在节目里她经常用到技术分析，这一主题自她 15 岁起就开始感兴趣。她曾经是维尔纽斯大学多门金融相关科目的助教。

Tamás Vadász 拥有布达佩斯考文纽斯大学的经济学硕士学位。毕业之后，他从事金融服务业的顾问工作。目前，他正在进行金融学博士学位的学习，他的主要研究兴趣包括金融经济学和银行业的风险管理。他在考文纽斯大学教的课程包括金融计量学、投资学和公司财务。

Kata Váradi 自 2013 年起是布达佩斯考文纽斯大学金融系的助理教授。作为金融学学生，Kata 毕业于 2009 年，并在 2012 年其毕业论文关于匈牙利股票市场的市场流动性风险分析通过答辩，获得了布达佩斯考文纽斯大学的博士学位。她的研究领域是市场流动性、固定收益证券以及医疗保健系统的网络。除了做研究，她也积极从事教学。她主要讲授公司财务、投资学、估值以及跨国金融管理。

Ágnes Vidovics-Dancs 是一位博士学位候选人，并且是布达佩斯考文纽斯大学的助理教授。此前她的工作是匈牙利政府债务管理局的初级风险管理师。她的主要研究领域是政府债务管理（一般）以及主权危机和违约（特别的）。她持有 CEFA 和 CIIA 证书。

审稿人简介

Matthew Gilbert 是一名量化分析师，在加拿大多伦多的加拿大养老基金投资公司（CPPIB）的全球宏观组工作。他拥有多伦多大学的量化金融硕士学位，以及皇后大学的应用数学和机械工程本科学位。

Hari Shanker Gupta 博士是一位算法交易系统开发领域的高级量化研究分析师。在此之前，他是位于印度班加罗尔的印度科学研究院（Indian Institute of Science, IISc）的博士后。他在 IISc 获得了应用数学和科学计算的博士学位。他在印度瓦拉纳西的巴纳拉斯印度大学（Banaras Hindu University, BHU）完成了硕士学位。在瓦拉纳西，他因为杰出的表现，获得过 4 次金奖。

Hari 已经在数学和科学计算的著名期刊上发表过 5 篇研究论文。他拥有在数学、统计和计算领域的工作经验。经验主题包括数值方法、偏微分方程、数理金融、随机分析、数据分析、时间序列分析、有限差分以及有限元方法。他擅长数学软件 Matlab、统计编程语言 R、Python 和编程语言 C。

他曾经为 Packt 出版的《量化金融 R 语言初级教程》（*Introduction to R for Quantitative Finance*）和《Python 数据分析学习指南》（*Learning Python Data Analysis*）两本书审稿。

Ratan Mahanta 拥有计算金融的硕士学位。目前他就职于 GPSK 投资集团，是一名高级量化分析师。对于卖方和风险顾问公司的量化交易和开发等方面，他拥有 3 年半的相关工作经验。他在 Github 开源平台为量化交易领域编写算法代码。他是自我激励、求知欲旺盛并且努力工作的人，热爱解决市场、技术、研究和设计交叉领域的困难问题。目前，他正在开发高频交易策略和量化交易策略。

量化交易：外汇、股权、期货与期权以及关于衍生品的工程。

算法：偏微分方程、随机微分方程、有限差分方法、蒙特卡洛以及机器学习。

代码：R 编程、RStudio 的 Shiny、C++、Matlab、HPC 以及科学计算。

数据分析：大数据分析（EOD 到 TBT）、Bloomberg、Quandl 以及 Quantopian。

策略：波动率套利、香草和奇异期权建模、趋势跟踪、均值回复、协整、蒙特卡洛模拟、在线价值、压力测试、高夏普比的买方策略、信用风险建模以及信用评分。

他也审阅过 Packt 出版社的《精通 R 语言科学计算》(*Mastering Scientific Computing with R*)，目前，他正在审阅的图书是 Packt 出版社的《R 语言机器学习手册》(*Maching Learning with R Cookbook*)。

前言

本书是我们的前一本书《量化金融 R 语言初级教程》(*Introduction to R for Quantitative Finance*)的续作。本书是为那些希望学习 R 语言来建立更高级量化金融模型的读者而写的。本书包括实证金融(第 1~4 章)、金融工程(第 5~7 章)、交易策略优化(第 8~10 章)和银行管理(第 10~13 章)等主题。

本书内容

第 1 章,“时间序列分析”讨论了一些重要的概念,比如,协整(结构化)、向量自回归模型、脉冲-响应函数、基于非对称 GARCH 模型的波动率建模以及信息影响曲线。

第 2 章,“因素模型”讲述了如何建立多因子模型并实施。借助主成分分析方法,可以识别出 5 个独立因子解释资产回报率。为了给出例证,我们基于真实市场数据重新建立了 Fama-French 模型。

第 3 章,“成交量预测”讲述了日内成交量预测模型,及其基于 DJIA 指数数据的 R 语言实现。模型没有使用成交率,而是使用了换手率,并从动态成分中分解出季节成分,进而分别预测了这两个成分。

第 4 章,“大数据—高级分析”使用 R 访问开源数据,并在大数据上执行多种分析。为了给出例证,我们将 K-均值聚类 and 线性回归运用到了大数据上。

第 5 章,“FX 衍生品”推广了用于衍生品定价的 Black-Scholes 模型。Margrabe 公式是对 Black-Scholes 模型的一种扩展,可以用来编程实现对股票期权、货币期权、外汇期权和

quanto 期权的定价。

第 6 章，“利率衍生品和模型”给出了利率模型和利率衍生品的概览。使用 Black 模型对上限和上限单元定价，并给出了利率模型，如 Vasicek 模型和 CIR 模型。

第 7 章，“奇异期权”介绍了奇异期权，解释了它们与普通香草期权的关系，并对任意衍生品定价函数估计了它们对应的希腊字母表示，并且更详细地考察了一种特殊的奇异期权——无触发二元期权。

第 8 章，“最优对冲”分析了衍生品对冲过程中的一些应用问题，这些问题由组合在离散时间上重新安排和交易成本引起。为了找到最优对冲策略，使用了不同的数值算法。

第 9 章，“基本面分析”研究了如何基于基本面建立交易策略。为了选出收益最佳的股票，一方面，根据公司过去的表现创建它们的聚类；另一方面，借助决策树区分出表现卓越的股票。基于这些结果，定义股票选择的规则并实施了回测。

第 10 章，“技术分析、神经网络和对数优化组合”给出了技术分析和一些相应策略的概述，如神经网络和对数优化组合；并以动态的设定研究了预测单资产（比特币）价格、优化交易择时以及（NYSE 股票）组合配置的问题。

第 11 章，“资产和负债管理”讲述了如何使用 R 语言支持银行的资产和负债管理。关注点在于数据生成、利率风险的度量和报告、流动性的风险管理以及无到期存款行为的建模。

第 12 章，“资本充足率”总结了巴塞尔协议的原则，而且借助历史的、delta-正态的、蒙特卡洛模拟的方法计算了在险价值（VaR），用以确定银行的资本充足性。同时也对信用和操作风险的特定问题有所涉猎。

第 13 章，“系统风险”讲述了两方法，这些方法基于网络理论，并有助于识别系统性重要的金融机构，分别是核心-边缘模型和传染模型。

Gergely Daróczi 审核了本书大多数章节的程序代码，感谢其为本书做出的贡献。

阅读本书之前的准备工作

首先，你需要在自己的计算机上安装 R 控制台。本书提供的所有代码示例需要在 R 控制台上运行。你可以免费下载软件，并找到适用于所有主要操作系统的安装指导。尽管我们没有包括高级主题，比如，如何在整合发展环境中使用 R 语言，但是对于 Emacs、Eclipse、vi 或者 Notepad++ 以及其他的编辑器有许多极好的插件，并且我们高度推荐你试一试 RStudio，这是一款致力于 R 语言的免费开源 IDE。

除了安装 R 的操作平台，我们还会使用一些用户贡献的 R 包，可以轻易从 CRAN (Comprehensive R Archive Network) 进行安装。为了安装 R 包，可以在 R 控制台使用 `install.packages` 命令，显示如下：

```
> install.packages('Quantmod')
```

这个包安装完毕之后，需要先装载在当前的 R 会话中，再进行使用：

```
> library (Quantmod)
```

在 R 的主页，你可以找到免费的入门文章和手册。

目标读者

本书的目标读者是那些既熟悉基本金融概念又具有一定编程能力的人。但是，即使你具备量化金融的基础，或者你已经拥有一些 R 的编程经验，本书同样可以给你新的启示。即使你已经是相关领域的专家，本书也能帮助你迅速切入其他领域。不过，如果你希望完美地掌握各章的内容，需要具备中级水平的量化金融知识，并且还需要掌握关于 R 的相关知识。以上这些技能可以从本书的前作——《量化金融 R 语言初级教程》中获得。

排版约定

在本书中，你会发现一些不同的文本样式，用以区别不同种类的信息。这里举例说明其中一些样式，以及它们的含义。

命令行的输入和输出的格式如下：

```
#generate the two time series of length 1000
set.seed(20140623)           #fix the random seed
N <- 1000                     #define length of simulation
x <- cumsum(rnorm(N))         #simulate a normal random walk
gamma <- 0.7                  #set an initial parameter value
y <- gamma * x + rnorm(N)     #simulate the cointegrating series
plot(x, type='l')             #plot the two series
lines(y,col="red")
```

新术语和重要的词以黑体表示。你在屏幕上看到的文字，例如，在菜单或对话框中，

就像这样出现在文本中：“另一种有用的可视化练习是看对数尺度上的密度。”



警告或者重要的注解出现在这样的图标中。



提示或者技巧像这样出现。

读者反馈

我们始终欢迎读者的反馈。如果你对本书有任何想法，喜欢或者不喜欢什么，请让我们知道。读者的反馈对我们来说非常重要，这样我们才能出版读者最需要的图书。

一般性的反馈，请通过电子邮件发送到 feedback@packtpub.com，在邮件的主题中请注明书名。

如果你精通某个领域而且有兴趣写作或者促成一本书，请参考我们的作者指南 www.packtpub.com/authors。

客户支持

现在，你是一位自豪的 Packt 图书的拥有者，我们会竭尽全力帮助你充分利用手中的书。

下载示例代码

可以使用你的账户从 <http://www.packtpub.com> 下载所有已购买的 Packt 图书的示例代码文件。如果你从其他地方购买了本书，可以访问 <http://www.packtpub.com/support> 并且注册，我们会通过电子邮件把文件发送给你。

勘误表

虽然我们已经尽力确保本书的内容正确，但疏漏之处依旧在所难免。如果你在我国的图书中发现错误，无论是文本还是代码，希望能告知我们，我们不胜感激。这样做可以减

少其他读者的困扰，帮助我们改进本书的后续版本。如果你发现任何错误，请访问 <http://www.packtpub.com/submit-errata> 并提交，选择你的书，单击勘误表提交表单的链接，并输入详细的说明。勘误一经核实，你的提交将被接受，此勘误将上传到本公司网站或者添加到现有勘误表中。

要查看之前提交的勘误，请到 <https://www.packtpub.com/books/content/support> 并在搜索框中输入书名，请求的信息会在勘误部分出现。

侵权行为

互联网上的侵权材料是所有媒体都要面对的问题。在 Packt，我们非常重视保护版权和许可证。如果你发现我们的作品在互联网上以任何形式被非法复制，请立即为我们提供网址或者网站名称，以便我们能够寻求补救。

请把可疑盗版材料发送到 copyright@packtpub.com。

非常感谢你帮助我们保护作者，以及我们给你带来有价值内容的能力。

问题

如果你对本书内容有疑问，不管是哪个方面的，都可以通过 questions@packtpub.com 来联系我们，我们将尽最大努力来解决问题。