

国外实用统计丛书

应用线性统计模型

上册 (英文影印版·原书第5版)

Applied Linear Statistical Models

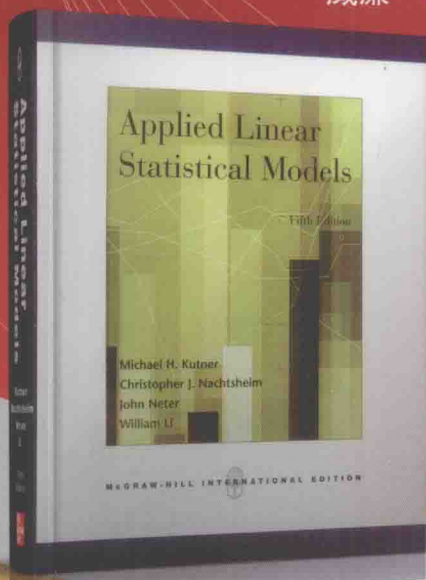
迈克尔 H.库特纳 (Michael H.Kutner)

[美] 克里斯托弗 J.纳彻舍姆(Christopher J.Nachtsheim)

约翰·内特 (John Neter)

威廉·李 (William Li)

著



Mc
Graw
Hill
Education

Mc
Graw
Hill
Education



机械工业出版社
CHINA MACHINE PRESS

国外实用统计丛书

应用线性统计模型 上册

(英文影印版·原书第5版)

迈克尔 H. 库特纳 (Michael H. Kutner)

[美]

克里斯托弗 J. 纳彻舍姆 (Christopher J. Nachtsheim)

约翰·内特 (John Neter)

威廉·李 (William Li)

机械工业出版社

本书分为三部分:第1部分简单线性回归,内容涉及单个预测变量的线性回归、利用回归和相关分析做推断、诊断和修正测度、回归分析的联合推断和其他论题以及简单线性回归分析的矩阵法等内容;第2部分多重线性回归、内容涉及多重回归I、多重回归II、定量和定性预测变量的回归模型、构建回归模型I、构建回归模型II、构建回归模型III、时序数据中的自相关等内容;第3部分非线性回归,内容涉及非线性回归的引入和神经网络、Logistic回归、泊松回归和广义线性模型等内容。本书篇幅适中,例子涉及各个应用领域,在介绍统计思想方面比较突出,数据丰富。

本书适用于高等院校统计学专业和理工科各专业本科生和研究生作为教材使用。

Michael H. Kutner, John Neter, Christopher J. Nachtsheim, William Li

Applied Linear Statistical Models

ISBN: 007310874X

Copyright © 2005 by McGraw-Hill Education.

All Rights reserved. No part of this publication may be reproduced or transmitted in any form or by any means, electronic or mechanical, including without limitation photocopying, recording, taping, or any database, information or retrieval system, without the prior written permission of the publisher.

This authorized English reprint edition is jointly published by McGraw-Hill Education and China Machine Press. This edition is authorized for sale in the People's Republic of China only, excluding Hong Kong, Macao SAR and Taiwan.

Copyright ©2015 by McGraw-Hill Education and China Machine Press.

版权所有。未经出版人事先书面许可,对本出版物的任何部分不得以任何方式或途径复制或传播,包括但不限于复印、录制、录音,或通过任何数据库、信息或可检索的系统。

本授权英文影印版由麦格劳-希尔(亚洲)教育出版公司和机械工业出版社合作出版。此版本经授权仅限在中华人民共和国境内(不包括香港特别行政区、澳门特别行政区和台湾地区)销售。

版权 ©2015 由麦格劳-希尔(亚洲)教育出版公司和机械工业出版社所有。

本书封面贴有 McGraw-Hill Education 公司防伪标签,无标签者不得销售。

北京市版权局著作权合同登记号:图字 01-2013-2948。

图书在版编目(CIP)数据

应用线性统计模型:原书第5版.上册:英文/(美)库特纳(Kutner, M.H.)等著.一影印本.
—北京:机械工业出版社,2015.3
(国外实用统计丛书)

书名原文:Applied linear statistical models

ISBN 978-7-111-49068-5

I. ①应… II. ①库… III. ①线性回归-统计模型-英文 IV. ①O212.1

中国版本图书馆CIP数据核字(2014)第310373号

机械工业出版社(北京市百万庄大街22号 邮政编码100037)

策划编辑:汤嘉 责任编辑:汤嘉

责任校对:孟令磊 封面设计:张静

责任印制:常天培

北京机工印刷厂印刷(三河市南杨庄国丰装订厂装订)

2016年4月第1版第1次印刷

184mm×260mm·47.75印张·1插页·1187千字

标准书号:ISBN 978-7-111-49068-5

定价:98.00元

凡购本书,如有缺页、倒页、脱页,由本社发行部调换

电话服务

网络服务

服务咨询热线:010-88379833 机工官网:www.cmpbook.com

读者购书热线:010-88379649 机工官博:weibo.com/cmp1952

教育服务网:www.cmpedu.com

封面无防伪标均为盗版

金书网:www.golden-book.com

英文影印版序

本书是在美国大学中广泛使用的教材，已经再版至第5版，不仅深受广大师生的欢迎，而且有很大的影响，已逐步成为经典。

由于篇幅较大，故其英文影印版分为上、下两册。本书深入地介绍了“应用线性统计模型”这门课程中几乎所有的关键知识，但是读起来并不艰深晦涩。书中用深入浅出的方式来讲解相关概念，同时配有大量的例题、习题以及实际案例帮助学生理解知识点。同时在帮助学生独立地解决实际问题方面，本书给人留下很深刻的印象。

本书图文并茂，许多例子和习题都是经过精心挑选的，来源于生活和工程实践，丰富的数据也都取材于实际案例。因此，本书不仅适用于统计专业，也可作为商业、计量经济学等专业的参考书。

本书叙述比较详尽，内容比国内教材丰富，篇幅较大，因此作为教材时可适当选取主要内容讲授，其余可作为学生自学使用。

Preface

Linear regression models are widely used today in business administration, economics, engineering, and the social, health, and biological sciences. Successful applications of these models require a sound understanding of both the underlying theory and the practical problems that are encountered in using the models in real-life situations. While *Applied Linear Regression Models*, Fourth Edition, is basically an applied book, it seeks to blend theory and applications effectively, avoiding the extremes of presenting theory in isolation and of giving elements of applications without the needed understanding of the theoretical foundations.

The fourth edition differs from the third in a number of important respects.

1. We have reorganized the chapters for better clarity and flow of topics. Material from the old Chapter 15 on normal correlation models has been integrated throughout the text where appropriate. Much of the material is now found in an expanded Chapter 2, which focuses on inference in regression analysis. Material from the old Chapter 7 pertaining to polynomial and interaction regression models and from old Chapter 11 on qualitative predictors has been integrated into a new Chapter 8 called, “Regression Models for Quantitative and Qualitative Predictors.” Material on model validation from old Chapter 10 is now fully integrated with updated material on model selection in a new Chapter 9 entitled, “Building the Regression Model I: Model Selection and Validation.”
2. We have added material on important techniques for data mining, including regression trees and neural network models, in Chapters 11 and 13, respectively.
3. We have made extensive revisions to the problem material. Problem data sets are generally larger and more challenging, and we have included a large number of new case data sets in Appendix C. In addition, we have added a new category of chapter exercises, called Case Studies. These are open-ended problems that require students, given an overall objective, to carry out complete analyses of the various case data sets in Appendix C. They are distinct from the material in the Problems and Projects sections, which frequently ask students to simply carry out specific analytical procedures.
4. The chapter on logistic regression (Chapter 14) has been extensively revised and expanded to include a more thorough treatment of logistic, probit, and complementary log-log models, logistic regression residuals, model selection, model assessment, logistic regression diagnostics, and goodness of fit tests. We have also developed new material on polytomous (multicategory) nominal logistic regression models and polytomous ordinal logistic regression models.
5. We have expanded the discussion of model selection methods and criteria. The Akaike information criterion and Schwarz Bayesian criterion have been added, and a greater emphasis is placed on the use of cross-validation for model selection and validation.
6. We have substantially expanded the amount of graphic presentation, including much greater use of scatter plot matrices, three-dimensional rotating plots, three-dimensional response surface and contour plots, and conditional effects plots.
7. Throughout the text, we have made extensive revisions in the exposition on the basis of classroom experience to improve the clarity of the presentation.

We have included in this book not only the more conventional topics in regression, but also topics that are frequently slighted, though important in practice. We devote three chapters (Chapter 9–11) to the model-building process for regression, including computer-assisted selection procedures for identifying good subsets of predictor variables and validation of the chosen regression model. Two chapters (Chapters 8 and 14) are devoted to indicator variables, covering both response and predictor indicator variables. The use of residual analysis and other diagnostics for examining the appropriateness of a regression model is a recurring theme throughout this book. So is the use of remedial measures that may be helpful when the model is not appropriate. In the analysis of the results of a study, we give greater emphasis to the use of estimation procedures than to significance tests, because estimation is often more meaningful in practice. Also, since practical problems seldom are concerned with a single inference, we stress the use of simultaneous inference procedures.

Theoretical ideas are presented to the degree needed for good understanding in making sound applications. Proofs are given in those instances where we feel they serve to demonstrate an important method of approach. Emphasis is placed on a thorough understanding of the regression models, particularly the meaning of the model parameters, since such understanding is basic to proper applications. A wide variety of examples and cases is presented to illustrate the use of the theoretical principles, to show the great diversity of applications of regression models, and to demonstrate how analyses are carried out for different problems.

We use “Comments” sections in each chapter to present additional discussion and matters related to the mainstream of development. In this way, the basic ideas in a chapter are presented concisely and without distraction.

Applications of regression models frequently require extensive computations. We take the position that a computer is available in most applied work and that almost every computer user has access to program packages for regression analysis. Hence, we explain the basic mathematical steps in fitting a regression model but do not dwell on computational details.

This approach permits us to avoid many complex formulas and enables us to focus on basic principles. We make extensive use in this text of computer capabilities for performing computations and preparing graphic plots, and we illustrate a variety of computer printouts and plots and explain how they are used for analysis.

A selection of problems is provided at the end of each chapter. Here readers can reinforce their understanding of the methodology and use the concepts learned to analyze data. We have been careful to supply data-analysis problems that typify genuine applications. In most problems the calculations are best handled on a computer. To facilitate data entry, a compact disk is provided with the text that includes the data sets for all examples, problems, exercises, projects, and case studies, as well as for the data sets in Appendix C. The README.TXT file on the compact disk provides information about the identification of the data sets.

We assume that the reader of *Applied Linear Regression Models*, Fourth Edition, has had an introductory course in statistical inference, covering the material outlined in Appendix A. Should some gaps in the reader’s background exist, the relevant portions of an introductory text can be studied, or the instructor of the class may use supplemental materials for covering the missing segments. Appendix A is primarily intended as a reference of basic statistical results for continuing use as the reader progresses through the book.

Calculus is not required for reading *Applied Linear Regression Models*, Fourth Edition. In a number of instances, we use calculus to demonstrate how some important results are obtained, but these demonstrations are confined to supplementary comments and can

be omitted without any loss of continuity. Readers who do know calculus will find these comments in natural sequence so that the benefits of the mathematical developments are obtained in their immediate context. Some basic elements of matrix algebra are needed for multiple regression. Chapter 5 introduces these elements of matrix algebra in the context of simple regression for easy learning.

Applied Linear Regression Models, Fourth Edition, is intended for use in undergraduate and graduate courses in regression analysis and in second courses in applied statistics. The extent to which material presented in this text is used in a particular course depends upon the amount of time available and the objectives of the course. The basic elements of regression are covered in Chapters 1, 2, 3, 4 (Sections 4.1–4.3), 5, 6, 7, 8, 9, and 10. Chapters 11, 12, 13, and 14 can be covered as time permits and interests dictate.

This book can also be used for self-study by persons engaged in the fields of business administration, economics, engineering, and the social, health, and biological sciences who desire to obtain competence in the application of regression models.

An *Instructor Solutions Manual*, containing detailed solutions to all numerical problems and analytical exercises, is available from the publisher, McGraw-Hill/Irwin, for use by instructors.

A *Student Solutions Manual* is included on the compact disk for use by students. The Student Solutions Manual provides intermediate and final numerical results for easy self-checking of solutions for selected problems. We use an asterisk (*) in front of the problem number to designate the problems for which the solutions appear in the Student Solutions Manual.

Available new for this Fourth Edition is a *SAS and SPSS Program Solution Manual* for examples contained in the text. The manual was prepared by Dr. William Replogle and Dr. William Johnson, both of the University of Mississippi Medical School. The manual uses data files to present solutions, output, and discussion of necessary steps for students to solve examples selected from the book using SAS and SPSS.

The *Student Solutions Manual* and all of the data files on the compact disk can also be downloaded from the book's website at: www.mhhe.com/KutnerALRM4e. A list of errata for the book as well as some useful, related links will also be maintained at this address.

A book such as this cannot be written without substantial assistance from numerous persons. We are indebted to the many contributors who have developed the theory and practice discussed in this book. We also would like to acknowledge appreciation to our students, who helped us in a variety of ways to fashion the method of presentation contained herein. We are grateful to the many users of *Applied Linear Statistical Models* and *Applied Linear Regression Models*, who have provided us with comments and suggestions based on their teaching with these texts. We are also indebted to Professors James E. Holstein, University of Missouri, and David L. Sherry, University of West Florida, for their review of *Applied Linear Statistical Models*, First Edition; to Professors Samuel Kotz, University of Maryland at College Park, Ralph P. Russo, University of Iowa, and Peter F. Thall, The George Washington University, for their review of *Applied Linear Regression Models*, First Edition; to Professors John S. Y Chiu, University of Washington, James A. Calvin, University of Iowa, and Michael F. Driscoll, Arizona State University, for their review of *Applied Linear Statistical Models*, Second Edition; to Professor Richard Anderson-Sprecher, University of Wyoming, for his review of *Applied Linear Regression Models*, Second Edition; and to Professors Alexander von Eye, The Pennsylvania State University, Samuel Kotz, University of

Maryland at College Park, and John B. Willett, Harvard University, for their review of *Applied Linear Statistical Models*, Third Edition; to Professors Jason Abrevaya, University of Chicago, Frank Alt, University of Maryland, Vitoria Chen, Georgia Tech, Rebecca Doerge, Purdue University, Mark Henry, Clemson University, Jim Hobert, University of Florida, Ken Koehler, Iowa State University, Chii-Dean Lin, University of Massachusetts Amherst, Mark Reiser, Arizona State University, Lawrence Ries, University of Missouri Columbia, and Ehsan Soofi, University of Wisconsin Milwaukee, for their reviews of *Applied Linear Regression Models*, Third Edition, or *Applied Linear Statistical Models*, Fourth Edition. These reviews provided many important suggestions, for which we are most grateful.

In addition, valuable assistance was provided by Professors Richard K. Burdick, Arizona State University, R. Dennis Cook, University of Minnesota, W. J. Conover, Texas Tech University, Mark E. Johnson, University of Central Florida, Dick DeVeaux, Williams College, and by Drs. Richard I. Beckman, Los Alamos National Laboratory, Ronald L. Iman, Sandia National Laboratories, and Brad Jones, SAS Institute. We are most appreciative of their willing help. We are also indebted to the 88 participants in a survey concerning *Applied Linear Regression Models*, Second Edition, the 76 participants in a survey concerning *Applied Linear Statistical Models*, Third Edition, and the 73 participants in a survey concerning *Applied Linear Regression Models*, Third Edition, or *Applied Linear Statistical Models*, Fourth Edition. Helpful suggestions were received in these surveys, for which we are thankful.

Weiyong Zhang and Vincent Agboto assisted us in the development of new problem material, and Lexin Li and Yingwen Dong prepared the revised *Instructor Solutions Manual* and *Student Solutions Manual* under considerable time pressure. Amy Hendrickson provided much-needed LaTeX expertise. We are most grateful to these individuals for their invaluable help and assistance. We also wish to thank the various members of the Carlson Executive MBA Program classes of 2003 and 2004, notably Mike Ohmes, Trevor Bynum, Baxter Stephenson, Zakir Salyani, Sanders Marvin, Trent Spurgeon, Nate Ogzawalla, David Mott, Preston McKenzie, Bruce DeJong, and Tim Kensok, for their contributions of interesting and relevant case study data and materials.

Finally, our families bore patiently the pressures caused by our commitment to complete this revision. We are appreciative of their understanding.

Michael H. Kutner

Christopher J. Nachtsheim

John Neter

William Li

目 录

英文影印版序

前言

第 1 部分

简单线性回归 1

第 1 章

单个预测变量的线性回归 2

1.1 变量之间的关系 2

两个变量之间的函数关系 2

两个变量之间的统计关系 3

1.2 回归模型及其用法 5

历史起源 5

基本概念 5

构建回归模型 7

回归分析的应用 8

回归和因果律 8

计算机的使用 9

1.3 具有不确定误差项分布的

简单线性回归模型 9

模型的形式语句 9

模型的主要特征 9

回归参数的意义 11

回归模型的替代形式 11

1.4 回归分析的数据 12

观测数据 12

试验数据 13

完全随机设计 13

1.5 回归分析步骤概述 13

1.6 回归函数估计 15

最小二乘法 15

均值响应的点估计 21

残差 22

拟合回归曲线的性质 23

1.7 误差项方差 σ^2 的估计 24

σ^2 的点估计 24

1.8 正态误差回归模型 26

模型 26

极大似然参数估计 27

引用的参考文献 33

问题 33

练习题 37

课题 38

第 2 章

利用回归和相关分析做推断 40

2.1 关于 β_1 的推断 40

b_1 的抽样分布 41

$(b_1 - \beta_1) / s\{b_1\}$ 的抽样分布 44

β_1 的置信区间 45

β_1 的检验 47

2.2 关于 β_0 的推断 48

b_0 的抽样分布 48

$(b_0 - \beta_0) / s\{b_0\}$ 的抽样分布 49

β_0 的置信区间 49

2.3 进行 β_0 和 β_1 推断的注意事项 50

正态偏差效应 50

对置信系数和误差风险的解释 50

X 水平间距 50

检验的功效 50

2.4 $E\{Y_h\}$ 的区间估计 52

\hat{Y}_h 的抽样分布 52

$(\hat{Y}_h - E\{Y_h\}) / s\{\hat{Y}_h\}$ 的

抽样分布 54

$E\{Y_h\}$ 的置信区间 54

2.5 新观测值的预测 55

参数已知时 $Y_{h(\text{new})}$ 的

预测区间 56

参数未知时 $Y_{h(\text{new})}$ 的

预测区间 57

给定 X_h 的 m 个新观测值的

均值预测 60

2.6 回归线的置信带 61

2.7 回归分析的分差分析法 63

总平方和的划分 63

自由度的分解 66

Contents

Preface

PART ONE

SIMPLE LINEAR REGRESSION 1

Chapter 1

Linear Regression with One Predictor Variable 2

- 1.1 Relations between Variables 2
 - Functional Relation between Two Variables* 2
 - Statistical Relation between Two Variables* 3
- 1.2 Regression Models and Their Uses 5
 - Historical Origins* 5
 - Basic Concepts* 5
 - Construction of Regression Models* 7
 - Uses of Regression Analysis* 8
 - Regression and Causality* 8
 - Use of Computers* 9
- 1.3 Simple Linear Regression Model with Distribution of Error Terms Unspecified 9
 - Formal Statement of Model* 9
 - Important Features of Model* 9
 - Meaning of Regression Parameters* 11
 - Alternative Versions of Regression Model* 12
- 1.4 Data for Regression Analysis 12
 - Observational Data* 12
 - Experimental Data* 13
 - Completely Randomized Design* 13
- 1.5 Overview of Steps in Regression Analysis 13
- 1.6 Estimation of Regression Function 15
 - Method of Least Squares* 15
 - Point Estimation of Mean Response* 21
 - Residuals* 22
 - Properties of Fitted Regression Line* 23
- 1.7 Estimation of Error Terms Variance σ^2 24
 - Point Estimator of σ^2* 24
- 1.8 Normal Error Regression Model 26
 - Model* 26
 - Estimation of Parameters by Method of Maximum Likelihood* 27

Cited References 33

Problems 33

Exercises 37

Projects 38

Chapter 2

Inferences in Regression and Correlation Analysis 40

- 2.1 Inferences Concerning β_1 40
 - Sampling Distribution of b_1* 41
 - Sampling Distribution of $(b_1 - \beta_1)/s\{b_1\}$* 44
 - Confidence Interval for β_1* 45
 - Tests Concerning β_1* 47
- 2.2 Inferences Concerning β_0 48
 - Sampling Distribution of b_0* 48
 - Sampling Distribution of $(b_0 - \beta_0)/s\{b_0\}$* 49
 - Confidence Interval for β_0* 49
- 2.3 Some Considerations on Making Inferences Concerning β_0 and β_1 50
 - Effects of Departures from Normality* 50
 - Interpretation of Confidence Coefficient and Risks of Errors* 50
 - Spacing of the X Levels* 50
 - Power of Tests* 50
- 2.4 Interval Estimation of $E\{Y_h\}$ 52
 - Sampling Distribution of \hat{Y}_h* 52
 - Sampling Distribution of $(\hat{Y}_h - E\{Y_h\})/s\{\hat{Y}_h\}$* 54
 - Confidence Interval for $E\{Y_h\}$* 54
- 2.5 Prediction of New Observation 55
 - Prediction Interval for $Y_{h(\text{new})}$ when Parameters Known* 56
 - Prediction Interval for $Y_{h(\text{new})}$ when Parameters Unknown* 57
 - Prediction of Mean of m New Observations for Given X_h* 60
- 2.6 Confidence Band for Regression Line 61
- 2.7 Analysis of Variance Approach to Regression Analysis 63
 - Partitioning of Total Sum of Squares* 63
 - Breakdown of Degrees of Freedom* 66

均方	66
方差分析表	67
期望均方	68
$\beta_1 = 0$ 与 $\beta_1 \neq 0$ 的 F 检验	69
2.8 一般线性检验方法	72
全模型	72
降阶模型	72
检验统计量	73
总结	73
2.9 X 与 Y 的线性相关关系描述测度	74
决定系数	74
R^2 极限	75
相关系数	76
2.10 应用回归分析研究	77
2.11 正态相关模型	78
回归模型和相关模型的区别	78
二元正态分布	78
条件推断	80
相关系数推断	83
Spearman 等级相关系数	87
引用的参考文献	89
问题	89
练习题	97
课题	98
第 3 章	
诊断和修正测度	100
3.1 预测变量的诊断	100
3.2 残差	102
残差的性质	102
半学生化的残差	103
模型偏差的残差法	103
3.3 残差诊断	103
回归函数的非线性	104
误差方差的不定性	107
离群值的存在性	108
误差项的非独立性	108
误差项的非正态性	110
重要预测变量的省略	112
最后的注释	114
3.4 残差检验概述	114
随机性检验	114
方差不变性检验	115
离群值检验	115
正态性检验	115
3.5 正态性的相关检验	115
3.6 误差方差齐性检验	116
Brown-Forsythe 检验	116
Breusch-Pagan 检验	118
3.7 失拟的 F 检验	119
假设	119
符号	121
全模型	121
降阶模型	123
检验统计量	123
ANOVA (方差分析) 表	124
3.8 修正测度概述	127
回归函数的非线性	128
误差方差的不定性	128
误差项的非独立性	128
误差项的非正态性	128
重要预测变量的省略	129
离群观测值	129
3.9 变换	129
针对非线性关系的变换	129
非正态和异方差误差下的变换	132
Box-Cox 变换	134
3.10 回归函数的图形探索	137
Lowess 方法	138
拟合回归函数的光滑曲线的应用	139
3.11 案例: 钶测量	141
引用的参考文献	146
问题	146
练习题	151
课题	152
实例研究	153

- Mean Squares* 66
 - Analysis of Variance Table* 67
 - Expected Mean Squares* 68
 - F Test of $\beta_1 = 0$ versus $\beta_1 \neq 0$* 69
 - 2.8** General Linear Test Approach 72
 - Full Model* 72
 - Reduced Model* 72
 - Test Statistic* 73
 - Summary* 73
 - 2.9** Descriptive Measures of Linear Association between X and Y 74
 - Coefficient of Determination* 74
 - Limitations of R^2* 75
 - Coefficient of Correlation* 76
 - 2.10** Considerations in Applying Regression Analysis 77
 - 2.11** Normal Correlation Models 78
 - Distinction between Regression and Correlation Model* 78
 - Bivariate Normal Distribution* 78
 - Conditional Inferences* 80
 - Inferences on Correlation Coefficients* 83
 - Spearman Rank Correlation Coefficient* 87
 - Cited References 89
 - Problems 89
 - Exercises 97
 - Projects 98
- Chapter 3**
- Diagnostics and Remedial Measures 100**
- 3.1** Diagnostics for Predictor Variable 100
 - 3.2** Residuals 102
 - Properties of Residuals* 102
 - Semistudentized Residuals* 103
 - Departures from Model to Be Studied by Residuals* 103
 - 3.3** Diagnostics for Residuals 103
 - Nonlinearity of Regression Function* 104
 - Nonconstancy of Error Variance* 107
 - Presence of Outliers* 108
 - Nonindependence of Error Terms* 108
 - Nonnormality of Error Terms* 110
 - Omission of Important Predictor Variables* 112
 - Some Final Comments* 114
 - 3.4** Overview of Tests Involving Residuals 114
 - Tests for Randomness* 114
 - Tests for Constancy of Variance* 115
 - Tests for Outliers* 115
 - Tests for Normality* 115
 - 3.5** Correlation Test for Normality 115
 - 3.6** Tests for Constancy of Error Variance 116
 - Brown-Forsythe Test* 116
 - Breusch-Pagan Test* 118
 - 3.7** F Test for Lack of Fit 119
 - Assumptions* 119
 - Notation* 121
 - Full Model* 121
 - Reduced Model* 123
 - Test Statistic* 123
 - ANOVA Table* 124
 - 3.8** Overview of Remedial Measures 127
 - Nonlinearity of Regression Function* 128
 - Nonconstancy of Error Variance* 128
 - Nonindependence of Error Terms* 128
 - Nonnormality of Error Terms* 128
 - Omission of Important Predictor Variables* 129
 - Outlying Observations* 129
 - 3.9** Transformations 129
 - Transformations for Nonlinear Relation Only* 129
 - Transformations for Nonnormality and Unequal Error Variances* 132
 - Box-Cox Transformations* 134
 - 3.10** Exploration of Shape of Regression Function 137
 - Lowess Method* 138
 - Use of Smoothed Curves to Confirm Fitted Regression Function* 139
 - 3.11** Case Example—Plutonium Measurement 141
 - Cited References 146
 - Problems 146
 - Exercises 151
 - Projects 152
 - Case Studies 153

第 4 章

回归分析的联合推断和其他论题 154

- 4.1 β_0 和 β_1 的联合估计 154
 - 联合估计的预备知识 154
 - Bonferroni 联合置信区间 155
- 4.2 平均响应的联合估计 157
 - Working-Hotelling 方法 158
 - Bonferroni 方法 159
- 4.3 新观测值的联合预测区间 160
- 4.4 通过原点的回归 161
 - 模型 161
 - 推断 161
 - 过原点回归应用中的重要注意事项 164
- 4.5 测量误差效应 165
 - Y 的测量误差 165
 - X 的测量误差 165
 - Berkson 模型 167
- 4.6 反预测 168
- 4.7 X 水平的选择 170
 - 引用的参考文献 172
 - 问题 172
 - 练习题 175
 - 课题 175

第 5 章

简单线性回归分析的矩阵法 176

- 5.1 矩阵 176
 - 矩阵的定义 176
 - 方阵 178
 - 向量 178
 - 转置 178
 - 矩阵相等 179
- 5.2 矩阵的加减法 180
- 5.3 矩阵的乘法 182
 - 矩阵的数乘 182
 - 矩阵的乘法 182
- 5.4 特殊矩阵 185
 - 对称矩阵 185
 - 对角矩阵 185

所有元素为 1 的向量和矩阵 187

零向量 187

- 5.5 矩阵的线性相关和秩 188
 - 线性相关 188
 - 矩阵的秩 188
- 5.6 矩阵的逆 189
 - 求逆矩阵 190
 - 逆矩阵的应用 192
- 5.7 矩阵的一些基本结果 193
- 5.8 随机向量与矩阵 193
 - 随机向量或矩阵的期望 193
 - 随机向量的方差 - 协方差矩阵 194
 - 一些基本结果 196
 - 多元正态分布 196
- 5.9 具有矩阵项的简单线性回归模型 197
- 5.10 回归参数的最小二乘估计 199
 - 正规方程 199
 - 估计回归系数 200
- 5.11 拟合值与残差 202
 - 拟合值 202
 - 残差 203
- 5.12 方差结果分析 204
 - 平方和 204
 - 二次型的平方和 205
- 5.13 回归分析推断 206
 - 回归系数 207
 - 平均响应 208
 - 新观测值的预测 209
 - 引用的参考文献 209
 - 问题 209
 - 练习题 212

第 2 部分

多重线性回归 213

第 6 章

多重回归 I 214

- 6.1 多重回归模型 214

Chapter 4

Simultaneous Inferences and Other Topics in Regression Analysis 154

- 4.1 Joint Estimation of β_0 and β_1 154
 - Need for Joint Estimation* 154
 - Bonferroni Joint Confidence Intervals* 155
- 4.2 Simultaneous Estimation of Mean Responses 157
 - Working-Hotelling Procedure* 158
 - Bonferroni Procedure* 159
- 4.3 Simultaneous Prediction Intervals for New Observations 160
- 4.4 Regression through Origin 161
 - Model* 161
 - Inferences* 161
 - Important Cautions for Using Regression through Origin* 164
- 4.5 Effects of Measurement Errors 165
 - Measurement Errors in Y* 165
 - Measurement Errors in X* 165
 - Berkson Model* 167
- 4.6 Inverse Predictions 168
- 4.7 Choice of X Levels 170
 - Cited References* 172
 - Problems* 172
 - Exercises* 175
 - Projects* 175

Chapter 5

Matrix Approach to Simple Linear Regression Analysis 176

- 5.1 Matrices 176
 - Definition of Matrix* 176
 - Square Matrix* 178
 - Vector* 178
 - Transpose* 178
 - Equality of Matrices* 179
- 5.2 Matrix Addition and Subtraction 180
- 5.3 Matrix Multiplication 182
 - Multiplication of a Matrix by a Scalar* 182
 - Multiplication of a Matrix by a Matrix* 182
- 5.4 Special Types of Matrices 185
 - Symmetric Matrix* 185
 - Diagonal Matrix* 185

Vector and Matrix with All Elements

Unity 187

Zero Vector 187

- 5.5 Linear Dependence and Rank of Matrix 188
 - Linear Dependence* 188
 - Rank of Matrix* 188
- 5.6 Inverse of a Matrix 189
 - Finding the Inverse* 190
 - Uses of Inverse Matrix* 192
- 5.7 Some Basic Results for Matrices 193
- 5.8 Random Vectors and Matrices 193
 - Expectation of Random Vector or Matrix* 193
 - Variance-Covariance Matrix of Random Vector* 194
 - Some Basic Results* 196
 - Multivariate Normal Distribution* 196
- 5.9 Simple Linear Regression Model in Matrix Terms 197
- 5.10 Least Squares Estimation of Regression Parameters 199
 - Normal Equations* 199
 - Estimated Regression Coefficients* 200
- 5.11 Fitted Values and Residuals 202
 - Fitted Values* 202
 - Residuals* 203
- 5.12 Analysis of Variance Results 204
 - Sums of Squares* 204
 - Sums of Squares as Quadratic Forms* 205
- 5.13 Inferences in Regression Analysis 206
 - Regression Coefficients* 207
 - Mean Response* 208
 - Prediction of New Observation* 209
 - Cited Reference* 209
 - Problems* 209
 - Exercises* 212

PART TWO

MULTIPLE LINEAR REGRESSION 213

Chapter 6

Multiple Regression I 214

- 6.1 Multiple Regression Models 214

- 需要多个预测变量 214
- 两个预测变量的一阶模型 215
- 多个预测变量的一阶模型 217
- 一般线性回归模型 217
- 6.2 含矩阵项的一般线性回归模型 222**
- 6.3 回归系数的估计 223**
- 6.4 拟合值和残差 224**
- 6.5 方差结果分析 225**
 - 平方和与均方 225
 - 回归关系的 F 检验 226
 - 多重测定系数 226
 - 多重相关系数 227
- 6.6 回归参数的推断 227**
 - β_k 的区间估计 228
 - β_k 的检验 228
 - 联合推断 228
- 6.7 平均响应估计和新观测值的预测 229**
 - $E\{Y_h\}$ 的区间估计 229
 - 回归面的置信区间 229
 - 平均响应的联合置信区间 230
 - 新观测值 $Y_{h(\text{new})}$ 的预测 230
 - 对 X_h 的 m 个新观测值的均值预测 230
 - g 个新观测值的预测 231
 - 隐外推法的注意事项 231
- 6.8 诊断和修正测度 232**
 - 散点图矩阵 232
 - 三维散点图 233
 - 残差图 233
 - 正态性的相关检验 234
 - 误差方差不变性的
 - Brown-Forsythe 检验 234
 - 误差方差不变性的
 - Breusch-Pagan 检验 234
 - 失拟 F 检验 235
 - 残差测度 236
- 6.9 例子：两个预测变量的多重回归 236**
 - 背景 236
 - 基本计算 237
 - 估计回归函数 240
 - 拟合值和残差 241
 - 模型的合理性分析 241
 - 方差分析 243
 - 回归参数的估计 245
 - 平均响应的估计 245
 - 新观测值的预测极限 247
 - 引用的参考文献 248
 - 问题 248
 - 练习题 253
 - 课题 254
- 第 7 章**
- 多重回归 II 256**
- 7.1 附加平方和 256**
 - 基本思想 256
 - 定义 259
 - SSR 分解成附加平方和 260
 - 包含 SSR 分解的 ANOVA 表 261
- 7.2 利用附加平方和进行回归系数检验 263**
 - 单个 $\beta_k=0$ 的检验 263
 - 多个 $\beta_k=0$ 的检验 264
- 7.3 回归系数检验的总结 266**
 - 全部 $\beta_k=0$ 的检验 266
 - 单个 $\beta_k=0$ 的检验 267
 - 多个 $\beta_k=0$ 的检验 267
 - 其他检验 268
- 7.4 部分测定系数 268**
 - 两个预测变量 269
 - 一般情形 269
 - 部分相关系数 270
- 7.5 标准多重回归模型 271**
 - 正规方程计算的舍入误差 271
 - 回归系数的比较缺失 272
 - 相关变换 272
 - 标准回归模型 273
 - 变换变量的 $X'X$ 矩阵 274

- Need for Several Predictor Variables* 214
 - First-Order Model with Two Predictor Variables* 215
 - First-Order Model with More than Two Predictor Variables* 217
 - General Linear Regression Model* 217
 - 6.2** General Linear Regression Model in Matrix Terms 222
 - 6.3** Estimation of Regression Coefficients 223
 - 6.4** Fitted Values and Residuals 224
 - 6.5** Analysis of Variance Results 225
 - Sums of Squares and Mean Squares* 225
 - F Test for Regression Relation* 226
 - Coefficient of Multiple Determination* 226
 - Coefficient of Multiple Correlation* 227
 - 6.6** Inferences about Regression Parameters 227
 - Interval Estimation of β_k* 228
 - Tests for β_k* 228
 - Joint Inferences* 228
 - 6.7** Estimation of Mean Response and Prediction of New Observation 229
 - Interval Estimation of $E\{Y_h\}$* 229
 - Confidence Region for Regression Surface* 229
 - Simultaneous Confidence Intervals for Several Mean Responses* 230
 - Prediction of New Observation $Y_{h(\text{new})}$* 230
 - Prediction of Mean of m New Observations at X_h* 230
 - Predictions of g New Observations* 231
 - Caution about Hidden Extrapolations* 231
 - 6.8** Diagnostics and Remedial Measures 232
 - Scatter Plot Matrix* 232
 - Three-Dimensional Scatter Plots* 233
 - Residual Plots* 233
 - Correlation Test for Normality* 234
 - Brown-Forsythe Test for Constancy of Error Variance* 234
 - Breusch-Pagan Test for Constancy of Error Variance* 234
 - F Test for Lack of Fit* 235
 - Remedial Measures* 236
 - 6.9** An Example—Multiple Regression with Two Predictor Variables 236
 - Setting* 236
 - Basic Calculations* 237
 - Estimated Regression Function* 240
 - Fitted Values and Residuals* 241
 - Analysis of Appropriateness of Model* 241
 - Analysis of Variance* 243
 - Estimation of Regression Parameters* 245
 - Estimation of Mean Response* 245
 - Prediction Limits for New Observations* 247
 - Cited Reference 248
 - Problems 248
 - Exercises 253
 - Projects 254
- ## Chapter 7
- ### Multiple Regression II 256
- 7.1** Extra Sums of Squares 256
 - Basic Ideas* 256
 - Definitions* 259
 - Decomposition of SSR into Extra Sums of Squares* 260
 - ANOVA Table Containing Decomposition of SSR* 261
 - 7.2** Uses of Extra Sums of Squares in Tests for Regression Coefficients 263
 - Test whether a Single $\beta_k = 0$* 263
 - Test whether Several $\beta_k = 0$* 264
 - 7.3** Summary of Tests Concerning Regression Coefficients 266
 - Test whether All $\beta_k = 0$* 266
 - Test whether a Single $\beta_k = 0$* 267
 - Test whether Some $\beta_k = 0$* 267
 - Other Tests* 268
 - 7.4** Coefficients of Partial Determination 268
 - Two Predictor Variables* 269
 - General Case* 269
 - Coefficients of Partial Correlation* 270
 - 7.5** Standardized Multiple Regression Model 271
 - Roundoff Errors in Normal Equations Calculations* 271
 - Lack of Comparability in Regression Coefficients* 272
 - Correlation Transformation* 272
 - Standardized Regression Model* 273
 - $X'X$ Matrix for Transformed Variables* 274

估计标准回归系数 275

7.6 多重共线性及其影响 278

不相关预测变量 279

预测变量完全相关问题的性质 281

多重共线性的影响 283

更有效诊断多重共线性的需求 289

引用的参考文献 289

问题 289

练习题 292

课题 293

第 8 章

定量和定性预测变量的回归模型 294

8.1 多项式回归模型 294

多项式模型的运用 294

一个预测变量：二阶 295

一个预测变量：三阶 296

一个预测变量：高阶 296

两个预测变量：二阶 297

三个预测变量：二阶 298

多项式回归模型的实现 298

实例 300

多项式回归的进一步注释 305

8.2 关联回归模型 306

关联效应 306

具有线性效应的关联回归模型的解释 306

具有曲线效应的关联回归模型的解释 309

关联回归模型的实现 311

8.3 定性预测 313

两类定性预测 314

回归系数的解释 315

多类定性预测 318

时间序列应用 319

8.4 使用指标变量的一些注意事项 321

指示变量与分配代码 321

指示变量与定量变量 322

指示变量的其他编码 323

8.5 定量和定性预测交互关系的建模 324

回归系数法 324

8.6 更多复杂模型 327

多个定性预测变量 328

只含定性预测变量 329

8.7 两个或多个回归函数的比较 329

肥皂生产线的例子 330

仪器校准研究的例子 334

引用的参考文献 335

问题 335

练习题 340

课题 341

实例研究 342

第 9 章

构建回归模型 I：模型的选择与评价 343

9.1 模型构建过程概述 343

数据收集 343

数据整理 346

初始模型的研究 346

解释变量的简化 347

模型的细化和选择 349

模型评价 350

9.2 外科的例子 350

9.3 模型选择的准则 353

R_p^2 或 SSE_p 准则 354

R_{adj}^2 或 MSE_p 准则 355

Mallows 的 C_p 准则 357

AIC_p 和 SBC_p 准则 359

PRESS_p 准则 360

9.4 模型选择的自动搜索过程 361

最佳子集算法 361

逐步回归法 364