普通高等教育"十一五"规划教材

21世纪大学数学创新教材

◎丛书主编　陈　化

数学建模与数学实验

汪晓银　周保平　主编

科学出版社
www.sciencep.com

普通高等教育“十一五”规划教材

21世纪大学数学创新教材

丛书主编 **陈 化**

数学建模与数学实验

汪晓银 周保平 主 编

科 学 出 版 社

北 京

内 容 简 介

本书通过实例与算法程序设计介绍了常用的数学建模方法，包括多元统计、时间序列分析、线性与非线性规划、多目标规划与目标规划、图论、动态规划、排队论、智能优化算法、微分与差分、模糊数学、神经网络、计算机仿真、灰色系统和层次分析法. 全书将建模技术与数学实验融为一体，注重数学建模思想介绍，重视数学软件（SAS、MATLAB、LINGO）在实际问题中的应用. 全书案例丰富，通俗易懂，便于自学.

本书既可以作为高等学校数学建模与数学实验课程的教材，也可作为本科生、研究生数学建模竞赛的培训教材或参考书籍，也是科学研究人员一本有价值的参考书籍.

图书在版编目（CIP）数据

数学建模与数学实验/汪晓银，周保平主编. 一北京：科学出版社，2010.2
普通高等教育"十一五"规划教材. 21世纪大学数学创新教材
ISBN 978-7-03-026502-9

Ⅰ. 数… Ⅱ. ①汪…②周… Ⅲ. ①数学模型－高等学校－教材 ②高等数学－实验－高等学校－教材 Ⅳ. ①O141.4 ②O13－33

中国版本图书馆 CIP 数据核字（2010）第 015571 号

责任编辑：王雨舸 曾 莉/责任校对：董艳辉
责任印制：彭 超/封面设计：苏 波

科学出版社 出版
北京东黄城根北街 16 号
邮政编码：100717
http://www.sciencep.com

武汉市新华印刷有限责任公司印刷
科学出版社发行 各地新华书店经销

*

2010 年 1 月第 一 版 开本：B5（720×1000）
2010 年 1 月第一次印刷 印张：17 1/2
印数：1—5 000 字数：338 000

定价：28.00 元

（如有印装质量问题，我社负责调换）

《21 世纪大学数学创新教材》丛书编委会

《21世纪大学数学创新教材》丛书序

《21世纪大学数学创新教材》为大学本科数学系列教材，大致划分为公共数学类、专业数学类两大块，创新是其主要特色和要求. 经组编委员会审定，列选科学出版社普通高等教育"十一五"规划教材.

一、组编机构

《21世纪大学数学创新教材》丛书由多所985和211大学联合组编：

丛书主编 陈 化

常务副主编 樊启斌

副 主 编 吴传生 何 穗 刘安平

丛书编委 （按姓氏笔画为序）

王卫华 王展青 刘安平 严国政 李 星

杨瑞琰 肖海军 吴传生 何 穗 汪晓银

陈 化 罗文强 赵东方 黄樟灿 梅全雄

彭 放 彭斯俊 曾祥金 谢民育 樊启斌

二、教材特色

创新是本套教材的主要特色和要求，创造双重品牌：

先进. 把握教改、课改动态和学科发展前沿，学科、课程的先进理念、知识和方法原则上都要写进教材或体现在教材结构及内容中.

知识与方法创新. 重点教材、高层次教材，应体现知识、方法、结构、内容等方面的创新，有所建树，有所创造，有所贡献.

教学实践创新. 教材适用，教师好教，学生好学，是教材的基本标准. 应紧跟和引领教学实践，在教学方法、教材结构、知识组织、详略把握、内容安排上有独到之处.

继承与创新. 创新须与继承相结合，是继承基础上的创新；创新需转变为参编者、授课者的思想和行为，避免文化冲突.

三、指导思想

遵循国家教育部高等学校数学与统计学教学指导委员会关于课程教学的基本要求，力求教材体系完整，结构严谨，层次分明，深入浅出，循序渐进，阐述精炼，富有启发性，让学生打下坚实的理论基础. 除上述一般性要求外，还应具备下列特点：

(1) 恰当融入现代数学的新思想、新观点、新结果,使学生有较新的学术视野.

(2) 体现现代数学创新思维,着力培养学生运用现代数学软件的能力,使教材真正成为基于现代数学软件的、将数学软件融合到具体教学内容中的现代精品教材.

(3) 在内容取舍、材料组织、叙述方式等方面具有较高水准和自身特色.

(4) 数学专业教材要求同步给出重要概念的英文词汇,章末列出中文小结,布置若干道(少量)英文习题,并要求学生用英文解答. 章末列出习题和思考题,并列出可进一步深入阅读的文献. 书末给出中英文对照名词索引.

(5) 公共数学教材具有概括性和简易性,注重强化学生的实验训练和实际动手能力,加强内容的实用性,注重案例分析,提高学生应用数学知识和数学方法解决实际问题的能力.

四、主编职责

丛书组编委员会和出版社确定全套丛书的编写原则、指导思想和编写规范,在这一框架下,每本教材的主编对本书具有明确的责权利:

1. 拟定指导思想

按照丛书的指导思想和特色要求,拟出编写本书的指导思想和编写说明.

2. 明确创新点

教改、课改动态,学科发展前沿,先进理念、知识和方法,如何引入教材;知识和内容创新闪光点及其编写方法;教学实践创新的具体操作;创新与继承的关系把握及其主客体融合.

3. 把握教材质量

质量是图书的生命,保持和发扬科学出版社"三高"、"三严"的传统特色,创出品牌;适用性是教材的生命力所在,应明确读者对象,篇幅要结合大部分学校对课程学时数的要求.

4. 掌握教材编写环节

(1) 把握教材编写人员水平,原则上要求博士、副教授以上,有多年课程教学经历,熟悉课程和学科领域的发展状况,有教材编写经验,有扎实的文字功底.

(2) 充分注意著作权问题,不侵犯他人著作权.

(3) 讨论、拟定教材提纲,并负责编写组的编写分工、协调与组织.

(4) 拟就内容简介、前言、目录、样章,统稿、定稿,确定交稿时间.

(5) 负责出版事宜,敦促编写组成员使用本教材,并优先选用本系列教材.

《21世纪大学数学创新教材》组编委员会

2009年6月

前　言

数学的应用正向几乎所有的科学领域渗透. 除了自然科学、工程技术、农业科学等领域外，还出乎意料地渗透到语言学、社会学、历史学等许多人文科学和其他领域，运用数学解决实际问题已经显得越来越重要.

“学数学，用数学”一直是我们的教学理念. 1997～2008 年，我们开展了以教学内容和课程体系改革为主体的，以注重增强大学生“用数学”的意识，培养学生“用数学”的能力为目标的教学改革. 改革解决了教学内容与课程体系的设置，建立了完善的数学应用推广机制，每年学习数学建模的人数从 1999 年的 15 人上升到 2008 年的 1400 人. 而且，最近几年我们的国家建模成绩显著提升，初步显现了我们的教学改革成效.

但是我们发现，数学应用能力的培养仍然是大学数学课堂教学中最薄弱的环节之一，缺乏实用有效、便于自学的教材的矛盾日益体现. 为了满足广大师生对数学建模技术的迫切需求，我们编写了这本教材.

本书得到了湖北省教改项目“农林高校数理化基础课实践教学体系的创新与实践”的大力支持，建立提高大学生的实践动手能力与创新能力的教学体系是本项目改革的重点. 数学学习的目标不仅仅是为了锻炼学生的计算能力，更重要的是提高学生运用数学解决实际问题的能力. 要提高这种能力必须要大力推广和普及数学建模方法与数学软件. 本书就是进行这种普及和推广所依赖的重要工具之一.

在本书中，我们依照科学研究、大学生科技创新以及国家建模竞赛所需要的数学建模方法，通过实例与算法程序设计介绍了多元统计、时间序列分析、线性与非线性规划、多目标规划与目标规划、图论、动态规划、排队论、智能优化算法、微分与差分、模糊数学、神经网络、计算机仿真、灰色系统和层次分析法等多种技术. 本书内容详实，通俗易懂，便于自学，其主要特点有：

第一，本书所有程序都在计算机上进行过调试和优化，运行可靠；

第二，本书案例是我们多年来的教学总结，具有代表性；

第三，本书案例丰富、图文并茂、条理清晰；

第四，全书将建模技术与数学实验融为一体，注重数学建模思想介绍，重视数学软件(SAS、MATLAB、LINGO)在实际中的应用.

本书由汪晓银(华中农业大学)、周保平(塔里木大学)担任主编. 全书共 9 章，第 1 章、第 2 章介绍了统计建模与 SAS 编程，由汪晓银、成峰、何丽娟编写；第 3 章介绍了线性规划、非线性规划、多目标规划和目标规划以及 LINGO 编程，由任兴龙、宋双

营编写；第 4 章介绍了图论与 MATLAB 编程，由周保平、任兴龙编写；第 5 章介绍了动态规划与排队论，由李治、齐立美（塔里木大学）编写；第 6 章介绍了智能优化算法，由谭劲英、王邦菊、杨前雨编写；第 7 章介绍了微分方程与差分方程及 MATLAB 编程，由侯志敏、徐思、周林编写；第 8 章介绍了模糊数学方法及 MATLAB 编程，由方红、李雪菲编写；第 9 章介绍了神经网络、计算机仿真、灰色系统和层次分析法，分别由石峰、徐艳玲、汪晓银、胡汉涛（塔里木大学）编写. 汪晓银、周保平审阅全稿，并对全稿进行了排版校对.

为了方便读者更好的学习本书，我们将所有建模技术的软件实现代码放到华中数学建模网 http://www.shumo.cn/大学数学实验栏目，供读者下载.

本书是集体智慧的结晶，是多年来教学的总结. 但限于水平所限，书中难免有错误之处，恳请读者批评指正.

编　者

2009 年 12 月 22 日

目　录

第1章　多元统计

在工业、农业、医学、气象、环境以及经济、管理等诸多领域中，常常需要同时观测多个指标. 在对多个随机变量的观测数据进行有效地分析和研究时，一种做法是将多个随机变量分开分析，每次分析一个变量；另一种做法是同时进行分析研究. 显然前者做法虽然有时是有效的，但分开处理不仅会丢掉很多信息，往往也不容易取得很好的研究结果. 因此，本章通过定性与定量分析、理论与实例介绍几种常用的多元统计方法.

1.1　多元回归

多元回归分析是研究多个变量之间关系的回归分析方法，按因变量和自变量的数量对应关系可划分为一个因变量对多个自变量的回归分析（简称为“一对多”回归分析）及多个因变量对多个自变量的回归分析（简称为“多对多”回归分析），按回归模型类型可划分为线性回归分析和非线性回归分析. 数学建模中建立变量之间互相影响的模型需要用到这类方法.

1.1.1　多元线性回归

1. 多元线性回归的概念

设自变量 $x_1,x_2,\cdots,x_p$ 的观测值 $x_{i1},x_{i2},\cdots,x_{ip}$ 及因变量 y 对应的观测值 y_i 满足关系式：

$$y_i=\beta_0+\sum_j\beta_j x_{ij}+\varepsilon_i,\quad i=1,2,\cdots,n;j=1,2,\cdots,p$$

其中，$\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 是相互独立且都服从正态分布 $N(0,\sigma^2)$ 的随机变量，且

$$\boldsymbol{Y}=\begin{bmatrix}y_1\\y_2\\\vdots\\y_n\end{bmatrix},\quad \boldsymbol{X}=\begin{bmatrix}1&x_{11}&x_{21}&\cdots&x_{p1}\\1&x_{12}&x_{22}&\cdots&x_{p2}\\\vdots&\vdots&\vdots&&\vdots\\1&x_{1n}&x_{2n}&\cdots&x_{pn}\end{bmatrix},\quad \boldsymbol{\beta}=\begin{bmatrix}\beta_1\\\beta_2\\\vdots\\\beta_p\end{bmatrix},\quad \boldsymbol{e}=\begin{bmatrix}\varepsilon_1\\\varepsilon_2\\\vdots\\\varepsilon_n\end{bmatrix}$$

根据最小二乘法，得

$$\hat{\boldsymbol{\beta}}=(\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X})^{-1}(\boldsymbol{X}^{\mathrm{T}}\boldsymbol{Y})$$

2. 回归方程显著性检验

与一元线性回归方程相类似，多元线性回归方程的总离差平方和 SST 也可以分

解为剩余平方和 SSE 与回归平方和 SSR，即

$$SST=SSE+SSR$$

理论上可以证明：当原假设 H_0 为 $\beta_1=\beta_2=\cdots=\beta_p=0$ 且 H_0 成立时，

$$F=\frac{SSR/p}{SSE/(n-p-1)}\sim F(p,n-p-1), \qquad \hat{\sigma}^2=MSE=\frac{SSE}{n-p-1}$$

为 σ^2 的无偏估计量.

因此，给出显著性水平 α，即可进行回归方程的显著性检验.

3. 回归系数的显著性检验

对自变量系数的检验通常使用 t 检验法.

设随机变量 $x_1,x_2,\cdots,x_n$ 对应的系数为 $b_1,b_2,\cdots,b_n$. 各 x_i 都服从正态分布，所以 b_j 也服从正态分布，且

$$E(b_j)=\beta_j, \qquad D(b_j)=\sigma^2 c_{jj}, \qquad \frac{b_j-\beta_j}{\sqrt{\sigma^2 c_{jj}}}\sim N(0,1)$$

其中，c_{jj} 是正规方程组的系数矩阵的逆矩阵中第 j 行第 j 列的元素.

可以提出原假设 H_0 为 $\beta_j=0$ 且 H_0 成立时，由 $\frac{SSE}{\sigma^2}$ 服从 $\chi^2(n-p-1)$ 分布推出：

$$F_j=\frac{b_j^2/c_{jj}}{SSE/(n-p-1)}\sim F(1,n-p-1), \qquad t_j=\frac{b_j/\sqrt{c_{jj}}}{\sqrt{SSE/(n-p-1)}}\sim t(1,n-p-1)$$

因此给出显著性水平 α 即可进行回归常数与回归系数的检验，得到各个系数显著的结论. 其中 F 的显著性以 $\Pr>F$ 表示，t 的显著性以 $\Pr>|t|$ 表示，而统计量 $F=t^2$，因此，$\Pr>F$ 与 $\Pr>|t|$ 应是等价的. 通常使用 $\Pr>|t|$ 小于 α 时拒绝原假设 H_0，认为系数不为 0；否者接受原假设 H_0，认为系数为 0，系数没有通过检验.

例 1.1 某品种水稻糙米含镉量 y(mg/kg)与地上部生物量 x_1(10 g/盆)及土壤含镉量 x_2(100 mg/kg)的 8 组观测值见表 1.1. 试建立多元线性回归模型.

表 1.1 某水稻糙米含镉量的观测值

x_1	1.37	11.34	9.67	0.76	17.67	15.91	15.74	5.41
x_2	9.08	1.89	3.06	10.2	0.05	0.73	1.03	6.25
y	4.93	1.86	2.33	5.78	0.06	0.43	0.87	3.86

编写程序如下：

```
/*数据段*/
data ex;  /*表示数据集为 ex*/
input x1-x2 y@@;  /*@@表示连续输入数据*/
cards;
/*数据省略,见华中数学建模网 http://www.shumo.cn/大学数学实验栏目*/
;
```

```
/*程序段*/
proc reg;  /*调用回归模块*/
model y=x1/cli;  /*对 y 关于 x1 做回归,/cli 表示求预测值与预测区间*/
run;
```

运行结果如下：

（1）回归方程显著性检验.

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	1	31.46291	31.46291	494.06	<.0001
Error	6	0.38209	0.06368		
Corrected Total	7	31.84500			

Root MSE	0.25235	R-Square	0.9880
Dependent Mean	2.51500	Adj R-Sq	0.9860
Coeff Var	10.03390		

由 Analysis of Variance 表可知，其 F Value＝494.06，Pr＞F 的值小于 0.0001，远小于 0.05，故拒绝原假设，接受备择假设，认为 y 与 x1，x2 之间具有显著的线性相关关系；由 R-Square 的值为 0.988 可知该方程的拟合度很高，样本观察值有 98.8％的信息可以用回归方程进行解释，故拟合效果较好，认为 y 与 x1，x2 之间具有显著的线性相关关系.

（2）参数显著性检验.

Parameter Estimates

Variable	DF	Parameter Estimate	Standard Error	t Value	Pr > |t|
Intercept	1	3.61051	0.95915	3.76	0.0131
x1	1	-0.19828	0.05822	-3.41	0.0191
x2	1	0.20675	0.09769	2.12	0.0879

由 Parameter Estimates 表可知，对自变量 x2，t 检验值为 t＝2.12，Pr＞|t|的值等于 0.0879，大于 0.05，因此，接受原假设 $H_0:\beta_2=0$，认为 x2 的系数应为 0，说明 x2 的系数没有通过检验. 为此，需要在程序 model y＝x1 x2 中去掉 x2.

再次运行得到如下结果：

Parameter Estimates

Variable	DF	Parameter Estimate	Standard Error	t Value	Pr > |t|
Intercept	1	5.62117	0.16580	33.90	<.0001
x1	1	-0.31911	0.01436	-22.23	<.0001

由 Parameter Estimates 表可知，对常数检验 t 值为 t＝33.9，Pr＞|t|的值小于 0.0001，远小于 0.05，说明截距项通过检验，估计值为 5.62117.

对自变量 x1 检验 t 值为 t=−22.23，Pr>|t|的值小于 0.0001，远小于 0.05，说明 x1 的系数通过检验，估计值为−0.31911.

以上结果表明所有变量的系数均通过检验，于是该线性模型即可得到．然而，许多实际问题中可能还会出现某几个变量的系数并没有通过检验，此时，可以在原程序中的 model y=x1−x2 中去掉没用通过的变量，直到所有的系数均通过检验；或者使用逐步回归方法，让软件自动保留通过检验的变量.

(3) 拟合区间.

Output Statistics

Obs	Dep Var y	Predicted Value	Std Error Mean Predict	95% CL Predict		Residual
1	4.9300	5.1840	0.1496	4.4662	5.9018	-0.2540
2	1.8600	2.0024	0.0922	1.3451	2.6598	-0.1424
3	2.3300	2.5353	0.0892	1.8804	3.1903	-0.2053
4	5.7800	5.3786	0.1567	4.6518	6.1055	0.4014
5	0.0600	-0.0176	0.1447	-0.7294	0.6942	0.0776
6	0.4300	0.5441	0.1258	-0.1459	1.2340	-0.1141
7	0.8700	0.5983	0.1241	-0.0898	1.2864	0.2717
8	3.8600	3.8948	0.1087	3.2224	4.5671	-0.0348

以上为样本的拟合结果，其中 Dep Var y 为因变量的原始值，Predicted Value 为 y 的拟合值，95% CL Predict 为拟合值 95%的拟合区间，Residual 为残差．例如，第一组原函数值为 4.93，拟合区间为[4.4662，5.9018]，残差为−0.254.

综合以上分析可以得到回归方程：

$$y=-0.31911x_1+5.62117$$

1.1.2 多元非线性回归

建立多元非线性回归方程在科学研究中应用广泛，其重要方法是将非线性回归方程转化为线性回归方程．转化时应首先选择适合的非线性回归形式，并将其线性化，再确定线性化回归方程的系数，最后确定非线性回归方程中未知的系数或参数.

对于实际问题，多元非线性回归模型与一元非线性回归类似，首先应对原始数据作图或者通过经验，观察数据发展情况，选择适当的函数进行拟合.

例 1.2 已知 1978～2006 年全国生产总值 GDP(y)、第一产业生产总值(x_1)、第二产业生产总值(x_2)、工业生产总值(x_3)、第三产业生产总值(x_4)，请建立 y 对 x_1～x_4 的回归模型.

通过经济模型经验，选用对数线性模型．输入命令：

```
data ex;
input y x1-x4@@;
y1=log(y);z1=log(x1);z2=log(x2);z3=log(x3);z4=log(x4);
/*对数据做变化，取自然对数后再做回归分析*/
cards;
```

```
/*数据省略,见华中数学建模网 http://www.shumo.cn/ 大学数学实验栏目*/
;
proc reg;  /*调用回归模块*/
model y1=z1 z2 z3 z4/cli;  /*也可写成 y1=z1-z4/cli;*/
proc corr;var z1-z4;  /*求相关系数矩阵*/
run;
```

运行结果如下：

(1) 回归方程显著性检验.

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	3	59.29570	19.76523	73270.5	<.0001
Error	25	0.00674	0.00026976		
Corrected Total	28	59.30244			

Root MSE	0.01642	R-Square	0.9999
Dependent Mean	5.09388	Adj R-Sq	0.9999
Coeff Var	0.32243		

由 Analysis of Variance 表可知，其 F Value＝73270.5，Pr＞F 的值小于 0.0001，远小于 0.05，故拒绝原假设，接受备择假设，认为 y1 与 z1，z2，z3，z4 之间具有显著的线性相关关系；由 R-Square 的值为 0.9999 可知该方程的拟合度很高，样本观察值有 99.99%的信息可以用回归方程进行解释，故拟合效果较好.

(2) 参数显著性检验.

Parameter Estimates

Variable	DF	Parameter Estimate	Standard Error	t Value	Pr > |t|
Intercept	1	0.96645	0.06662	14.51	<.0001
z1	1	0.09854	0.04497	2.19	0.0384
z2	1	0.67951	0.07739	8.78	<.0001
z3	1	-0.02654	0.07474	-0.36	0.7256
z4	1	0.24800	0.04929	5.03	<.0001

由 Parameter Estimates 表可知，对自变量 z3 检验 t 值为 $t=-0.36$，Pr>|t|的值等于 0.7256，大于 0.05，因此接受原假设 $H_0:\beta_3=0$ 认为 z3 的系数应为 0，说明 z3 的系数没有通过检验. 为此，需要在程序 model y1＝z1 z2 z3 z4 中去掉 z3.

再次运行得到如下结果：

Parameter Estimates

Variable	DF	Parameter Estimate	Standard Error	t Value	Pr > |t|
Intercept	1	0.96744	0.06539	14.79	<.0001
z1	1	0.09768	0.04412	2.21	0.0362
z2	1	0.65586	0.03873	16.93	<.0001
z4	1	0.24696	0.04833	5.11	<.0001