高等学校试用

多元统计分析

孙文爽　陈兰祥　编

教育出版社

高等学校试用教材

多元统计分析

孙文爽　陈兰祥　编

高等教育出版社

（京）112号

内容提要

本书为工科院校应用数学专业的教材.

全书共有十一章：矩阵理论；多元正态分布；Wishart 分布和 T^2 统计量，Λ-统计量的分布；多元正态分布参数的估计；统计假设检验；多元线性统计模型；判别分析；聚类分析；主成分分析；因子分析；典型相关分析。

高等学校试用教材

多元统计分析

孙文爽 陈兰祥 编

*

高等教育出版社出版

新华书店总店科技发行所发行

中国科学院印刷厂印装

*

开本 850×1168 1/32 印张 16.875 字数 430 000

1994年6月第1版 1994年6月第1次印刷

印数 0001—1 445

ISBN7-04-004693-8/O·1323

定价 7.50 元

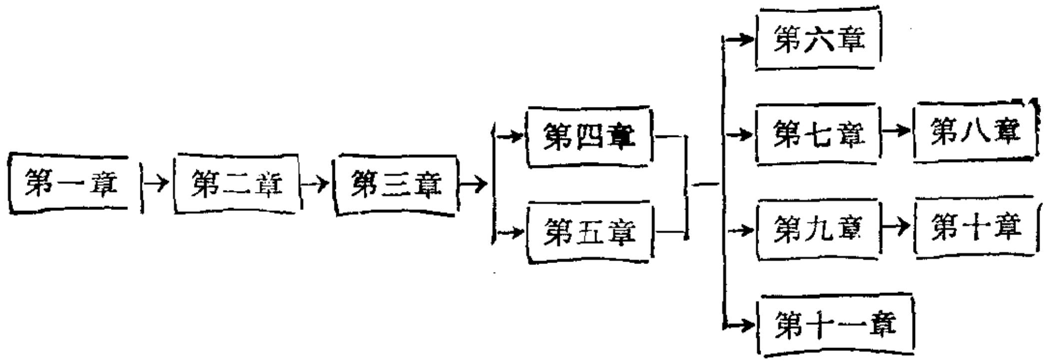
前　言

“多元统计分析”是数理统计学近30年来迅速地发展起来的一个分支，特别是由于现代电子计算机的广泛运用，多元统计方法已广泛应用到许多社会科学领域和自然科学学科之中去，取得了很大的进展．目前许多院校的有关专业都已开设或者准备开设这门课程，迫切需要一本合适的既阐明多元统计理论又介绍各种多元统计方法的教材．为了适应社会的需要，我们在国家教委高等工业学校应用数学专业教材编审委员会的关心支持下编写了本教材供教学使用．

本教材是作者根据自编讲义，经过多年使用，广泛听取了学生意见，并吸收了国内外该课程教材和教学上的许多长处不断修改而写成的，并经过国家教委高等工业学校应用数学专业教材编审委员会评审后定为应用数学专业多元统计分析课程的试用教材．它也可供其他有关专业学生或研究生作为教材和参考书．同时也为气象水文，生物，医药，经济和工农业的科技人员和工程技术人员提供了一本实用多元统计方法的参考书．

全书共分十一章．第一章介绍了“多元统计分析”教程中所需要的有关矩阵理论的知识．第二章介绍了数据分析所依据的基本概率模型——多元正态分布及它的有关性质．第三章介绍了三个重要的分布，即 Wishart 分布、Hotelling T^2 分布和 Wilks Λ 分布及有关性质．第四章和第五章分别对多元正态分布有关的参数估计和假设检验的概念以及在实际上的运用进行了讨论．而从第六章开始直至第十一章依次介绍了近年来发展起来的卓有成效的各种新的多元统计方法，包括多元回归分析，多元方差分析，判别分析，聚类分析，主成分分析，因子分析，对应分析和典型相关分析等内容．

上述章节的关系可由下列框图表示：



本书有以下几个特点：

1）“理论与实践并重”. 为了使读者更好地掌握多元统计分析的各种方法，一方面用相当篇幅讲清最基本的多元统计理论，使读者能理解统计模型的数学处理和统计推断的理论基础. 另一方面对各种统计方法都强调其背景和实际意义，说明该方法的统计思想，并在此基础上建立起统计模型，推导出模型的解，然后再归纳出统计推断的方法和步骤. 书中还列举了地质、生物、医药、经济、农业、教育等领域内的大量实例.

2）本书在理论证明中全部用矩阵处理，有些定理采用了较新的证明方法，如变换 Jacobian 的计算和 Wishart 分布密度的简捷证法等，不仅便于教师讲解，节约教学时间，而且学生易于接受. 经过实践，效果较好.

3）为便于读者自学和保持教材的连贯性，本书第一章给出了所需的矩阵理论知识. 除了大样本的渐近分布以及个别定理我们直接给出结论外，绝大部分的定理证明在本书范围内是完整和严谨的，读者不必查阅许多参考资料.

4）为了便于不同类型的学校和不同学时的教学安排提供选择，本书在内容安排上有各种方案. 对于偏重于应用的专业可略去*部分不讲（ *的内容有整节，某些节中的若干段落或某定理的证明）而不影响教学的连贯性. 对于一些较重要的但难度较大的定理证明都在相应各章的附录中给出，供教师和学生参考，如果

课时较多或者理论上要求较高的专业也可安排为教学内容。

5）为了使学生能熟练地掌握本教程的内容，每一章都配备了一定数量理论和应用方面的习题，部分较难的习题还附有提示。

本书所需教学时间大致如下：

第一章 8—10 小时

第二章 6—8 小时

第三章 6—8 小时

第四章 4 小时

第五章 6—10 小时

第六章 8—10 小时

第七章 6 小时

第八章 8—10 小时

第九章 4—6 小时

第十章 4—8 小时

第十一章 4 小时

略去*后基本内容大致需用 64 小时。

需要说明的是：在本书中定义，定理，引理，推论，例子，公式均按节编号，例如第五章第二节的第 3 个定理，记为定理 2.3。如果在第五章内引用它，就写根据“定理 2.3”，而在第七章中引用它时，则写明根据“第五章定理 2.3”．余类推．

有些内容未详加说明需要参考查阅有关文献．例如我们标明参考[8]的第四章，表示在文献目录[8]中的第四章可以找到有关的内容．

本书在编写过程中始终得到王学仁，马逢时两位教授的关心和支持．另外余明书、范大茵、陈祖荫、史道济和郭福星诸位专家对本书进行了极其认真负责的评审，提出了许多宝贵的意见．本书能够被国家教委高等工业学校应用数学专业教材编审委员会评审通过后推荐出版，上述诸位同志的指导帮助起了十分重要的作用，谨在此向他们表示最诚挚的感谢．

由于作者水平有限，书中难免会有缺点和错误，恳请专家和读者批评指正。

云 南 大 学 统 计 系　孙文爽
同济大学应用数学系　陈兰祥
一九九二年九月十日

主要符号说明

x 小写字母表示变量或随机变量
$\boldsymbol{x}$ 黑体小写字母表示向量或随机向量 （视上，下文而定）
X 大写字母表示矩阵或随机矩阵

$A>0$　A 为对称正定方阵

$A\geqslant 0$　A 为对称半正定(非负定)方阵

$A\geqslant B$　指 $A\geqslant 0, B\geqslant 0$ 且 $A-B\geqslant 0$

$|A|$　方阵 A 的行列式

$\mathrm{rk}(A)$　矩阵 A 的秩

$\mathrm{tr}(A)$　方阵 A 的迹

$\mathscr{L}(A)$　矩阵 A 的列向量张成的子空间

$\mathscr{L}^{\perp}(A)$　$\mathscr{L}(A)$ 的正交子空间

$\|\boldsymbol{a}\|$　向量 $\boldsymbol{a}$ 的长度

$\mathbf{1}_n$　所有分量为 1 的 n 维列向量

P_z　在 $\mathscr{L}(Z)$ 上的投影矩阵

$\mathrm{ch}_i(A)$　方阵 A 的第 i 大的特征根

$J(\boldsymbol{x}\rightarrow\boldsymbol{y})$　向量 $\boldsymbol{x}$ 到向量 $\boldsymbol{y}$ 的变换 Jacobian 行列式

EX　随机矩阵 X 的数学期望矩阵

$\mathrm{var}(x)$　随机变量 x 的方差

$\mathrm{cov}(\boldsymbol{x},\boldsymbol{y})$　随机向量 $\boldsymbol{x}$ 和 $\boldsymbol{y}$ 的协方差矩阵

$\mathrm{cov}(\boldsymbol{x})$　随机向量 $\boldsymbol{x}$ 的自协方差矩阵

$\mathrm{vec}(A)$　将 A 的诸行向量依次排列后进行转置得到的拉长向量

$N(\mu,\sigma^2)$　数学期望为 μ，方差为 σ^2 的一维正态分布

$N_p(\boldsymbol{\mu},\Sigma)$　数学期望向量为 $\boldsymbol{\mu}$，自协方差矩阵为 Σ 的 p 维正态分布

$N_{p,n}(M,\Sigma,V)$　参数为 M,Σ,V 的 $p\times n$ 阶矩阵正态分布

$W_p(n,\Sigma,\Omega)$　非中心参数为 Ω 的 Wishart 分布

$W_p(n,\Sigma)$　中心 Wishart 分布

$\chi^2(p,\lambda)$　非中心参数为 λ 的自由度为 p 的 χ^2 分布

$\chi^2(p)$ 自由度为 p 的中心 χ^2 分布

$F(n,m,\lambda)$ 非中心参数为 λ、自由度为 n,m 的 F 分布

$F(n,m)$ 自由度为 n,m 的中心 F 分布

$T^2(p,n,\lambda)$ 非中心参数为 λ、自由度为 p,n 的 T^2 分布

$T^2(p,n)$ 自由度为 p,n 的中心 T^2 分布

$t(p)$ 自由度为 p 的 t 分布

$\Lambda(p,n,l)$ 参数为 p,n,l 的 (Wilks) Λ 分布

$B(n,p)$ 二项分布

$PN(n,p_1,\cdots,p_r)$ 多项分布

$\chi^2_\alpha(p)$ $\chi^2(p)$ 分布的上侧 α 分位数

$t_\alpha(p)$ $t(p)$ 分布的上侧 α 分位数

$F_\alpha(n,m)$ $F(n,m)$ 分布的上侧 α 分位数

LS 估计 最小二乘估计

MLE 极大似然估计

MSE$(\hat{\theta})$ 估计向量 $\hat{\theta}$ 的均方误差

$\xrightarrow{P}$ 按概率收敛

$\xrightarrow{L}$ 按分布收敛

$d(x,y)$ 两点 x,y 之间的距离

$D(A,B)$ A 类和 B 类之间的距离

$\mathrm{etr}(A)$ $\exp(\mathrm{tr}A)$ 或者 $e^{\mathrm{tr}A}$ 的缩写

目　　录

第一章　矩 阵 理 论

在本章中我们将介绍“多元统计分析”课程中所需要矩阵的有关理论，凡是在通常线性代数教材中已有的材料我们将不加证明或者略加说明地进行引用. 虽然本章的内容并非本教材的中心内容，但其有关的结论对于以后各章的学习都是十分重要的. 在本章中我们引进了在通常线性代数教材中很少提及的一些矩阵运算，如拉长向量 vec$A$ 将帮助我们把随机矩阵的讨论化为熟知的随机向量的讨论. 而 Kronecker 乘积 $A\otimes B$ 在处理随机矩阵的协方差矩阵时有着十分简明的意义. 不同的 Jacobian 行列式的导出对于求出各种随机矩阵相应的分布密度是至关重要的，希望读者能熟练掌握.

§1.1　矩阵的有关定义及其运算

（一）　矩阵概念及其运算

令 R 是由实数组成的域，在 R 中任一组数字的四则运算总能进行. 设 $a_{11},\cdots,a_{pq}$ 是 R 中 pq 个元素，称由这些元素组成的 p 行 q 列矩形方块

$$\begin{pmatrix} a_{11} & a_{12}\cdots a_{1q} \\ a_{21} & a_{22}\cdots a_{2q} \\ & \cdots\cdots \\ a_{p1} & a_{p2}\cdots a_{pq} \end{pmatrix}$$

为一个 $p\times q$ 阶**矩阵**，通常记为大写字母 $A=(a_{ij})_{p\times q}$，其中 a_{ij} 是 A 的第 i 行第 j 列上的元素. 如果 A 的诸元素都为 0，则称 A 是一个 $p\times q$ 阶**零矩阵**，记为 $O_{p\times q}$，在不致误解时可简记为 O. 如果 $p=q$，称 A 为 p 阶**方阵**. 当 $p=1$ 时也称 A 为 q 维**行向量**，当 $q=1$ 时也称 A 为 p 维**列向量**.

以后总是用

$$\alpha=\begin{pmatrix}a_1\\a_2\\\vdots\\a_p\end{pmatrix}$$

表示由元素 $a_1,\cdots,a_p$ 组成的列向量,而用 $\alpha'=(a_1,a_2,\cdots,a_p)$ 表示由元素 $a_1,a_2,\cdots,a_p$ 组成的行向量.

如果 p 阶方阵中所有非对角线上元素全为0(即 $a_{ij}=0,i\neq j$),则称 A 为 p 阶**对角阵**. 对角线元素为 $\lambda_1,\cdots,\lambda_p$ 的对角阵记为 Λ 或者 $\mathrm{diag}(\lambda_1,\cdots,\lambda_p)$. 特别当 λ_i 全为1时,相应的对角阵称为 p 阶**单位方阵**,可简记为 I_p.

称在主对角线下方诸元素全为0的方阵为**上三角矩阵**,称在主对角线上方诸元素全为0的方阵为**下三角矩阵**.

称将矩阵 A 的诸行(列)换成相应的诸列(行)所得到的新矩阵为 A 的**转置矩阵**,记为 A'. 从而若 $A=(a_{ij})_{p\times q}$, 则 $A'=(a_{ji})_{q\times p}$.

如果方阵 A 满足 $A=A'$, 称 A 为**对称矩阵**(或对称阵).

有时我们把一个 $p\times q$ 阶矩阵 A 表示成分块形式

$$A=\begin{pmatrix}A_{11}&\cdots&A_{1n}\\\vdots&&\vdots\\A_{m1}&\cdots&A_{mn}\end{pmatrix}$$

其中 A_{ij} 是 A 的一个 $p_i\times q_j$ 阶子矩阵 $(i=1,\cdots,m,j=1,\cdots,n)$ 且 $\sum\limits_{i=1}^{m}p_i=p$, $\sum\limits_{j=1}^{n}q_j=q$.

下面介绍矩阵的运算.

两矩阵 A,B (具有相同的列数和行数)的和规定为

$$A+B=(a_{ij})+(b_{ij})=(a_{ij}+b_{ij})$$

一实数 λ 与矩阵 A 的乘积规定为

$$\lambda\cdot A=A\cdot\lambda=(\lambda a_{ij})$$

如果矩阵 A 的列数等于B的行数，即 $A=(a_{ij})_{p\times q}$，$B=(b_{jk})_{q\times r}$，则规定A与B的积为

$$A\cdot B=\left(\sum_{j=1}^{q}a_{ij}b_{jk}\right)_{p\times r}$$

即$A\cdot B$是一个$p\times r$阶矩阵，它的第(i,k)元素是$\sum\limits_{j=1}^{q}a_{ij}b_{jk}$。

可以验证上述运算具有以下性质：

1) $A+B=B+A$.

2) $(A+B)+C=A+(B+C)$.

3) $A+(-1)A=O$.

4) $(\lambda+\mu)A=\lambda A+\mu A$.

5) $\lambda(A+B)=\lambda A+\lambda B$.

6) $\lambda(\mu A)=(\lambda\mu)A$.

7) $(A+B)C=AC+BC$.

8) $(AB)C=A(BC)$.

9) $I\cdot A=A\cdot I=A$.

若A和B为同阶上(下)三角矩阵，那么它们的和 $A+B$ 以及它们的积 $A\cdot B$ 也是同类型的上(下)三角矩阵。

转置矩阵具有下列性质：

1) $(A')'=A$.

2) $(A+B)'=A'+B'$.

3) $(A\cdot B)'=B'\cdot A'$.

若矩阵A和B具有相同的行数与列数且按同样方式分块，即$A=(A_{ij})$，$B=(B_{ij})$，那么

$$A+B=(A_{ij}+B_{ij}).$$

若$m\times n$阶矩阵A，$n\times p$阶矩阵B分别划分为$A=(A_{ij})$，$B=(B_{jk})$，其中A_{ij}为$m_i\times n_j$阶矩阵，B_{jk}为$n_j\times p_k$阶矩阵，且$\sum\limits_i m_i=m$，$\sum\limits_j n_j=n$，$\sum\limits_k p_k=p$。那么

$$AB=\left(\sum_{j} A_{ij}B_{jk}\right)$$

（二） 矩阵的直积

如果A是 $m\times n$ 阶矩阵，B是 $p\times q$ 阶矩阵，它们的**直积**(或者称 **Kronecker 积**)是如下定义的 $mp\times nq$ 阶矩阵

$$A\otimes B=(a_{ij}B)$$

其中 $A=(a_{ij})$. 它表示以 $a_{ij}B$ 为它的第 (i,j)块的分块矩阵.

直积运算具有以下性质:

1) $0\otimes A=A\otimes 0=O$.

2) $(A_1+A_2)\otimes B=A_1\otimes B+A_2\otimes B$.

3) $A\otimes(B_1+B_2)=A\otimes B_1+A\otimes B_2$.

4) $\alpha A\otimes\beta B=\alpha\beta\cdot A\otimes B$.

5) $A_1A_2\otimes B_1B_2=(A_1\otimes B_1)(A_2\otimes B_2)$.

6) $(A\otimes B)'=A'\otimes B'$.

证明 我们仅证明5)式,其余请读者自行完成.

设 $A_1=(a_{ij})_{p\times m}$, $A_2=(c_{jk})_{m\times n}$, B_1 是 $s\times t$ 阶矩阵，B_2是 $t\times u$ 阶矩阵. 于是

$$A_1A_2\otimes B_1B_2=\left(\sum_{j=1}^{m} a_{ij}c_{jk}\right)_{p\times n}\otimes B_1B_2=\left(\sum_{j=1}^{m}a_{ij}c_{jk}B_1B_2\right)$$

是一个 $ps\times nu$ 阶矩阵，它的第 (i,k) 块是 $\sum\limits_{j=1}^{m} a_{ij}c_{jk}B_1B_2$. 而

$$A_1\otimes B_1=(a_{ij}B_1),\quad A_2\otimes B_2=(c_{jk}B_2)$$

按分块矩阵乘法有

$$(A_1\otimes B_1)(A_2\otimes B_2)=\left(\sum_{j=1}^{m} a_{ij}c_{jk}B_1B_2\right)$$

从而5)式成立.

令 vec$A$ 表示把 $p\times q$ 阶矩阵的诸行向量依次排列成一个拉长的行向量再加以转置而得到的 pq 维列向量,即

$$\mathrm{vec}A=(a_{11},a_{12},\cdots,a_{1q},a_{21},a_{22},\cdots,a_{2q},\cdots,a_{p1},\cdots,a_{pq})'$$

通过这种记号可以把有关矩阵的运算化为向量的运算．这时，对于矩阵方程

$$AXB=C \tag{1.1}$$

其中 A,X,B,C 分别为 $m\times n$ 阶，$n\times r$ 阶，$r\times q$ 阶，$m\times q$ 阶矩阵，利用直积记号可以改写成另一等价形式：

$$(A\otimes B')\mathrm{vec}X=\mathrm{vec}C \tag{1.2}$$

对此只须注意到把(1.1)和(1.2)具体写成线性方程组后都是

$$\sum_{j=1}^{n}\sum_{k=1}^{r}a_{ij}x_{jk}b_{kl}=c_{il}\qquad(i=1,\cdots,m,l=1,\cdots,q).$$

由此可得

定理 1.1 $\mathrm{vec}(AXB)=(A\otimes B')\mathrm{vec}X$。

例如，对于向量 $\boldsymbol{x}$ 和 $\boldsymbol{y}$ 有

$$\mathrm{vec}(\boldsymbol{x}\cdot\boldsymbol{y}')=\boldsymbol{x}\otimes\boldsymbol{y}.$$

（三） 向量的相关和 Schmidt 正交化方法

为了对矩阵有更深刻的理解，下面我们复习一些向量空间（线性空间）的知识．

设 R^p 表示实数域 R 上的全部 p 维向量，那么任一 p 维向量可以表示为 R^p 空间上的一个点，它作为特殊矩阵，可以在 R^p 上定义加法和数乘运算．显然 R^p 关于这些运算是封闭的，我们称 R^p 是一个**向量空间**．

对于 $p\times q$ 阶矩阵 $A\doteq(a_{ij})$，我们记它的列向量为 $\boldsymbol{a}_1,\cdots,\boldsymbol{a}_q$，行向量为 $\boldsymbol{a}'_{(1)},\cdots,\boldsymbol{a}'_{(p)}$，那么 A 可以写成

$$A=(\boldsymbol{a}_1,\cdots,\boldsymbol{a}_q)\ \text{或者}\ A=(\boldsymbol{a}_{(1)},\cdots,\boldsymbol{a}_{(p)})' \tag{1.3}$$

从而矩阵 A 既可看成 R^p 空间中的 q 个点，又可看成 R^q 空间中的 p 个点．另外由于矩阵 A 与 $\mathrm{vec}A$ 是一一对应的，它又可看成是 R^{pq} 空间中的一个点．

对于一组向量 $\boldsymbol{a}_1,\cdots,\boldsymbol{a}_q$，如果存在不全为 0 的一组实数