

高等院校精品课系列教材
国家统计局优秀教材
上海普通高校优秀教材
上海市级精品课程教材

应用多元分析

第五版

应用多元统计分析

王学民 编著

Applied Multivariate
Statistical Analysis

上海财经大学出版社

高等院校精品课系列教材
国家统计局优秀教材
上海普通高校优秀教材
上海市级精品课程教材



应用多元统计分析



王学民 编著

多元统计分析是统计学的一个重要分支，它在自然科学、社会科学和经济等各个领域得到了越来越广泛的应用，是一种非常重要且实用的多元数据处理方法。

本书对原第四版作了全面大幅修订，使之在多方面都有了进一步的改进。主要是针对财经类院校统计学专业的本科生和研究生而写的，也可作为其他院校统计学教材或教学参考书。书中的绝大部分内容曾向上海财经大学本科生和研究生分别讲授过二十多届。

本教材有如下一些特点和说明。

(1) 全书对数学基础知识的要求较低，只要求读者掌握初步的微积分、线性代数和概率统计知识。尽管如此，为便于读者能顺利地理解本书，第一章对矩阵代数以及第四章第一节的多元统计理论进行了必要的回顾。

(2) 本书注重对多元数据的处理，力求做到概念清晰、讲解深入浅出，很便于读者深刻理解和自学。

(3) 书中提供的大量例题和习题为读者展示了多元分析的实际应用，每章的例题和习题并非侧重于对基本概念的理解，而是侧重于对多元分析问题的数学技巧和难度。书中有不少例题，其结论较重要或较有用，将在题号后面标注“有用结论”，这些题目的结论本身或许比如何对其解题更为重要，为便于读者的学习（特别是自学的学习）和学习（特别是自学的学习）习题解答。

(4) 本书每一章后面的附录A、B、C、D、E、F、G、H、I、J、K、L、M、N、O、P、Q、R、S、T、U、V、W、X、Y、Z，本教材还配有相应电子版的JMP的应用和SPSS的应用，供读者使用。建议统计专业的读者以SAS为主，JMP(同为SAS公司)和SPSS(同为SAS公司)软件，JMP是较小型软件，本书所有内容的运算都可用SAS实现，运算都可用SAS实现，内容的运算，但JMP的菜单操作非常方便，交互式可视化方面做得也十分简便。两款软件结合使用可以互补，很有益处。

(5) 书中的一些数学证明和理论性较强的内容安排在每章的附录或打“*”的章节(或段落)里，一些定理性较强的内容安排在每章的附录或打“*”的章节(或段落)里，一些定理性较强的内容安排在每章的附录或打“*”的章节(或段落)里。

(6) 此次的主要修订有：(1) 将书名改为“应用多元统计分析”，主要原因是现在是互联网时代，输入关键字“多元统计分析”一般较难搜到“多元分析”，(2) 将正文中

图书在版编目(CIP)数据

应用多元统计分析/王学民编著. —5版. —上海:上海财经大学出版社, 2017. 8

高等院校精品课系列教材

国家统计局优秀教材

上海普通高校优秀教材

上海市级精品课程教材

ISBN 978-7-5642-2748-7/F · 2748

I. ①应… II. ①王… III. ①多元分析-统计分析 IV. ①O212.4

中国版本图书馆 CIP 数据核字(2017)第 125705 号

责任编辑 何苏湘

书籍设计 张克瑶

YINGYONG DUOYUAN TONGJI FENXI

应用多元统计分析

(应用多元分析 第五版)

王学民 编著

上海财经大学出版社出版发行
(上海市中山北一路 369 号 邮编 200083)

网 址: <http://www.sufep.com>

电子邮箱: webmaster@sufep.com

全国新华书店经销

上海叶大印务发展有限公司印刷装订
2017 年 8 月第 5 版 2017 年 8 月第 1 次印刷

787mm×1092mm 1/16 21 印张 537 千字

印数: 42 501—48 500 定价: 40.00 元

前 言

多元统计分析(简称多元分析)是统计学中内容十分丰富、应用性极强的一个重要分支,它在自然科学、社会科学和经济学等各领域得到了越来越广泛的应用,是一种非常重要和实用的多元数据处理方法。

本书对原第四版作了全面大幅修订,使之在多方面都有了进一步的改进和提升。本教材主要是针对财经类院校统计学专业的本科生和研究生而写的,也可作为其他各专业读者的多元统计分析教材或教学参考书。书中的绝大部分内容曾向上海财经大学统计与管理学院的本科生和研究生分别讲授过二十多届。

本教材有如下一些特点和说明:

(1) 全书对数学基础知识的要求较低,只需读者掌握初步的微积分、线性代数和概率统计知识。尽管如此,为便于读者能顺利地阅读本书,书中第一章对矩阵代数以及第四章第一节对一元统计推断知识都作了简要的回顾和介绍。

(2) 本教材以十分细致和深入浅出的方式阐述了多元统计分析的基本概念、统计思想和数据处理方法,在充分考虑到适合财经院校学生使用的前提下进行了严谨的论述。整本书说得清清楚楚,讲得明白明白,很便于读者深刻理解和自学。

(3) 书中提供的大量例题和习题为读者展示了多元分析在社会科学和经济学等领域中的应用,每章的例题和习题安排侧重于对基本概念的理解、知识的实际应用以及该章一些结论的证明需要,并不刻意强调解题的数学技巧和难度。书中有不少例题和习题,其结论较重要或较有用,特在题序号后面标注“有用结论”,这些题目的结论本身或许比如何对其解题更为重要。为便于读者的学习(特别是自学),书后的附录给出了(基本完整的)习题解答。

(4) 本书每一章后面的附录1都附有《SAS的应用》。此外,本教材还配有相应电子版的《JMP的应用》和《SPSS的应用》,这三款统计软件读者可选择使用。建议统计专业的读者以SAS为主,JMP(同为SAS公司的产品)为辅,SAS是大型软件,JMP是较小型软件。本书所有内容的运算都可用SAS实现,而JMP只能实现其大部分内容的运算,但JMP的菜单操作非常方便,交互式可视化方面做得极为出色,数据管理的操作也十分简便。两款软件结合使用可以互补,很有益处。

(5) 书中的一些数学证明和理论性较强的内容被安排在了各章的附录或打“*”的章节(或段落)里,一些理论性较强的习题也打了“*”,读者可视情况决定取舍。另外,为结构显示清晰,在所有的例题和证明的结束处标记有“□”。

(6) 此次的主要修订有:(i) 将原《应用多元分析》更名为《应用多元统计分析》,主要原因是现在是互联网时代,输入关键字“多元统计分析”一般较难搜索到“多元分析”;(ii) 将正文中

一些相对复杂点的推导或证明重新编排了一下,打上“*”或放入各章后面的附录中,此外还在其中增加了一些有助于学习和理解的证明,这些都供读者选读,这样既降低了阅读正文的数学难度,又能满足希望了解更多证明的读者的需要;(iii)删除了原第四版中交互式数据分析菜单系统的使用介绍(只保留了少数很有价值的输出结果介绍)和 § 3.6,新增了 § 7.4 和习题解答;(iv)增加了许多“注”和“脚注”,以利读者更易自学及可更深入地理解或拓展有关知识。此外,一般的改写和局部新增基本上贯穿全书,且有不少作了大幅变化。

全书共分十章。第一章介绍了多元分析中常用的矩阵代数知识,这是全书的基础。第二章至第四章介绍的基本上是一元统计推广到多元统计的内容,主要阐述了多元分布的基本概念和多元正态分布及其统计推断。第五章至第十章是多元统计独有的内容,这部分内容具有很强的实用性,特别是介绍了各种降维技术。涉及的降维方法包括:费希尔判别、主成分分析、因子分析、对应分析和典型相关分析。

读者可进入“上财教学网”:http://bb.shufe.edu.cn/,选择“教师主页”→“统计与管理学院”→“王学民”,从中可下载的配书资料包括:(1)与本教材相配套的PPT,并在此基础上制作了PPT的简化版,简化了本教材的内容,以供2学分的或非统计专业的多元统计分析课程使用;(2)书中(需使用SAS软件运算的)所有例题、习题的数据及SAS程序;(3)与本教材相配套的分章来写的《JMP的应用》和《SPSS的应用》;(4)被删除的原第四版中的一些内容;(5)其他配书资料。这些配书资料可能会不定期地作一些更新。

由于编者水平有限,再加上作者长期以来眼睛一直不能多看书,书中错误、不足之处在所难免,恳请读者批评指正。

王学民
2017年5月

§ 3.5 \bar{x} 和 $(n-1)S$ 的抽样分布	(60)
小 结	(61)
附录 3-1 SAS 的应用	(62)
附录 3-2 § 3.2 中若干性质的证明	(66)
习 题	(68)
第四章 多元正态总体的统计推断	(71)
§ 4.1 一元情形的回顾	(71)
§ 4.2 单个总体均值的推断	(75)
§ 4.3 两个总体均值的比较推断	(82)
§ 4.4 轮廓分析	(86)
§ 4.5 多个总体均值的比较检验(多元方差分析)	(92)
§ 4.6 协方差矩阵相等性的检验	(97)
§ 4.7 总体相关系数的推断	(98)
小 结	(100)
附录 4-1 SAS 的应用	(101)
附录 4-2 若干推导	(106)
附录 4-3 威尔克斯 Λ 分布的定义及基本性质	(109)
习 题	(110)
第五章 判别分析	(114)
§ 5.1 引言	(114)
§ 5.2 距离判别	(115)
§ 5.3 贝叶斯判别	(125)
§ 5.4 费希尔判别	(133)
§ 5.5 逐步判别	(141)
小 结	(145)
附录 5-1 SAS 的应用	(146)
附录 5-2 若干证明	(155)
习 题	(157)
第六章 聚类分析	(160)
§ 6.1 引言	(160)
§ 6.2 距离和相似系数	(160)
§ 6.3 系统聚类法	(163)
§ 6.4 动态聚类法	(184)
小 结	(186)
附录 6-1 SAS 的应用	(187)
附录 6-2 若干公式的推导	(196)

习题	(198)
第七章 主成分分析	(201)
§ 7.1 引言	(201)
§ 7.2 总体的主成分	(202)
§ 7.3 样本的主成分	(209)
§ 7.4 若干补充及主成分应用中需注意的问题	(218)
小结	(224)
附录 7-1 SAS 的应用	(225)
附录 7-2 (7.4.2)式的证明	(228)
习题	(228)
第八章 因子分析	(231)
§ 8.1 引言	(231)
§ 8.2 正交因子模型	(232)
§ 8.3 参数估计	(237)
§ 8.4 因子旋转	(243)
§ 8.5 因子得分	(250)
小结	(255)
附录 8-1 SAS 的应用	(255)
习题	(261)
第九章 对应分析	(263)
§ 9.1 引言	(263)
§ 9.2 行轮廓和列轮廓	(263)
§ 9.3 独立性的检验和总惯量	(267)
§ 9.4 行、列轮廓的坐标	(270)
§ 9.5 对应分析图	(271)
小结	(278)
附录 9-1 SAS 的应用	(278)
附录 9-2 若干推导	(281)
习题	(283)
第十章 典型相关分析	(285)
§ 10.1 引言	(285)
§ 10.2 总体典型相关	(285)
§ 10.3 样本典型相关	(290)
§ 10.4 典型相关系数的显著性检验	(294)
小结	(296)



第一章 矩阵代数

本章我们对书中需要用到的有关矩阵代数知识作一些简单的回顾和介绍,熟悉这些内容将为以后各章的阅读带来很大的方便。如果读者希望对这方面知识有更多、更详细的了解,可参考有关的教科书。

§ 1.1 定义 将 $p \times q$ 个实数 $a_{11}, a_{12}, \dots, a_{pq}$ 排列成的一个有 p 行、 q 列的矩形阵列

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

称为 $p \times q$ 矩阵^①,常记作 $A = (a_{ij}) : p \times q$,其中 a_{ij} 是第 i 行、第 j 列的元素。例如,矩阵

$$A = \begin{pmatrix} 5 & 2 & 9 \\ 3 & 7 & 1 \end{pmatrix}$$

是一个 2×3 矩阵,其中 $a_{11} = 5, a_{12} = 2, a_{13} = 9, a_{21} = 3, a_{22} = 7, a_{23} = 1$ 。

若 $q = 1$,则称 A 为 p 维列向量,记作

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

例如,带有元素 6, 9 和 3 的 3 维列向量可写为

$$a = \begin{pmatrix} 6 \\ 9 \\ 3 \end{pmatrix}$$

若 $p = 1$,则称 A 为 q 维行向量,用转置符号记作

$$a' = (a_1, a_2, \dots, a_q)$$

例如,带有元素 2, 7, 3, 4, 3 的 5 维行向量可写为

^① 本书中矩阵用大写的黑体字母表示,向量用小写的黑体字母表示。

$$a' = (2, 7, 3, 4, 3)$$

$\sqrt{a'a}$ 称为向量 a 的长度, 记作 $\|a\|$ 。若 $\|a\| = 1$, 则称 a 为单位向量。任一非零向量 (即元素不全为零) a 被其长度 $\|a\|$ 相除后即为单位向量, 即 $c = a / \|a\|$ 是一个单位向量。在 p 维欧氏空间 R^p 中, p 维向量 a 既可看作是一个带有方向和长度的矢量, 也可看作是一个带有 p 个分量坐标的点。

若 A 的所有元素全为零, 则称 A 为零矩阵, 记作 0_M 或 0 。例如,

$$0_{23} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

若 $p = q$, 则称 A 为 p 阶方阵, $a_{11}, a_{22}, \dots, a_{pp}$ 称为它的对角线元素, 其他元素 $a_{ij} (i \neq j)$ 称为非对角线元素。

若方阵 A 的对角线下方的元素全为零, 则称 A 为上三角矩阵。显然, $a_{ij} = 0, i > j$ 。例如,

$$\begin{pmatrix} 4 & 1 & 3 \\ 0 & 8 & 9 \\ 0 & 0 & 7 \end{pmatrix}, \begin{pmatrix} 3 & 1 \\ 0 & 4 \end{pmatrix}, \begin{pmatrix} 2 & 5 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

若方阵 A 的对角线上方的元素全为零, 则称 A 为下三角矩阵。显然, $a_{ij} = 0, i < j$ 。

若方阵 A 的所有非对角线元素均为零, 则称 A 为对角矩阵, 可记为 $\text{diag}(a_{11}, a_{22}, \dots, a_{pp})$ 。

例如,

$$\text{diag}(2, 1, 3) = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 3 \end{pmatrix}$$

若 p 阶对角矩阵 A 的所有 p 个对角线元素均为 1, 则称 A 为 p 阶单位矩阵, 记作 I_p 或 I 。例如,

$$I_1 = (1), \quad I_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad I_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

若将矩阵 A 的行与列互换, 则得到的矩阵称为 A 的转置, 记作 A' , 即

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

例如, 若

$$A = \begin{pmatrix} 1 & 3 & 1 \\ 2 & 4 & 6 \end{pmatrix}$$

则

$$A' = \begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 1 & 6 \end{pmatrix}$$

若方阵 A 满足 $A' = A$, 则称 A 为对称矩阵。显然, $a_{ij} = a_{ji}$ 。例如,

$$\begin{pmatrix} 1 & 3 \\ 3 & 4 \end{pmatrix}, \begin{pmatrix} 9 & 0 & 0 \\ 0 & 7 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \begin{pmatrix} 2 & 3 & 5 \\ 3 & 0 & 1 \\ 5 & 1 & 1 \end{pmatrix}$$

§ 1.2 矩阵的运算

若 $A=(a_{ij}):p \times q, B=(b_{ij}):p \times q$, 则 A 与 B 的和定义为

$$A+B=(a_{ij}+b_{ij}):p \times q$$

若 c 为一常数, 则它与 A 的积定义为

$$cA=(ca_{ij}):p \times q$$

若 $A=(a_{ij}):p \times q, B=(b_{ij}):q \times r$, 则 A 与 B 的积定义为

$$AB=\left(\sum_{k=1}^q a_{ik} b_{kj}\right):p \times r$$

从上述定义中容易得出如下规律:

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

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

$$(3) A(B_1+B_2)=AB_1+AB_2.$$

$$(4) A\left(\sum_{i=1}^k B_i\right)=\sum_{i=1}^k AB_i.$$

$$(5) c(A+B)=cA+cB.$$

例 1.2.1 设

$$A=\begin{pmatrix} 1 & 4 & 2 \\ 3 & 1 & 5 \end{pmatrix}, \quad B=\begin{pmatrix} 6 & 0 & 2 \\ 3 & 1 & 4 \end{pmatrix}, \quad C=\begin{pmatrix} 1 & -1 \\ 2 & 3 \end{pmatrix}$$

则

$$(1) A'=\begin{pmatrix} 1 & 3 \\ 4 & 1 \\ 2 & 5 \end{pmatrix}, \quad B'=\begin{pmatrix} 6 & 3 \\ 0 & 1 \\ 2 & 4 \end{pmatrix};$$

$$(2) A+B=\begin{pmatrix} 7 & 4 & 4 \\ 6 & 2 & 9 \end{pmatrix};$$

$$(3) (A+B)'=A'+B'=\begin{pmatrix} 7 & 6 \\ 4 & 2 \\ 4 & 9 \end{pmatrix};$$

$$(4) CA=\begin{pmatrix} 1 & -1 \\ 2 & 3 \end{pmatrix} \begin{pmatrix} 1 & 4 & 2 \\ 3 & 1 & 5 \end{pmatrix}=\begin{pmatrix} 4 & 5 & 7 \\ 11 & 11 & 19 \end{pmatrix}.$$

注意: AC 是没有定义的。

若两个 p 维向量 a 和 b 满足

$$a'b=a_1b_1+a_2b_2+\cdots+a_pb_p=0$$

则称 a 和 b 正交。几何上, 正交向量之间相互垂直。

若方阵 A 满足 $AA'=I$, 则称 A 为正交矩阵。例如,

$$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad \begin{pmatrix} \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \\ -\frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{pmatrix}, \quad \begin{pmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{2 \cdot 1}} & \frac{-1}{\sqrt{2 \cdot 1}} & 0 \\ \frac{1}{\sqrt{3 \cdot 2}} & \frac{1}{\sqrt{3 \cdot 2}} & \frac{-2}{\sqrt{3 \cdot 2}} \end{pmatrix}$$

由 § 1.4 可知, $AA'=I \Leftrightarrow A'=A^{-1} \Leftrightarrow A'A=I$, 因此正交矩阵可有这三种等价定义。

正交矩阵 A 有着较好的几何意义。如将 p 维向量 x 看作是在 R^p 中的一个点, 则 x 的各

分量是该点在相应各坐标轴上的坐标。正交变换 $y=Ax$ 意味着对原 p 维坐标系作一刚性旋转(或称正交旋转,此时 $|A|=1$)或反射(此时 $|A|=-1$)^①, y 的各分量正是该点在新坐标系下的坐标。习题 1.7 是二维情形下的一种直观展示,在三维情形下同样可以有着直观展示。由于

$$y'y = (Ax)'(Ax) = x'A'Ax = x'x$$

因此,在新、旧坐标系下,该点到原点的距离保持不变。

若方阵 A 满足 $A^2=A$, 则称 A 为幂等矩阵。例如,

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}$$

对称的幂等矩阵称为投影矩阵。

矩阵的分块是在处理阶数较高的矩阵时常用的方法。有时,我们把一个大矩阵看成是由一些小矩阵组成的,就像矩阵是由数组成的一样。设 $A=(a_{ij}): p \times q$, 将它分成四块,表示成

$$A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}$$

其中 $A_{11}: k \times l, A_{12}: k \times (q-l), A_{21}: (p-k) \times l, A_{22}: (p-k) \times (q-l)$ 。例如,若

$$A_{11} = \begin{pmatrix} 1 & 3 \\ 2 & 4 \end{pmatrix}, \quad A_{12} = \begin{pmatrix} 3 \\ 1 \end{pmatrix}, \quad A_{21} = (5, 4), \quad A_{22} = (6)$$

则

$$A = \begin{pmatrix} 1 & 3 & 3 \\ 2 & 4 & 1 \\ 5 & 4 & 6 \end{pmatrix}$$

若 A 和 B 有相同的分块,则

$$A+B = \begin{pmatrix} A_{11}+B_{11} & A_{12}+B_{12} \\ A_{21}+B_{21} & A_{22}+B_{22} \end{pmatrix}$$

若 C 为 $q \times r$ 矩阵,分成

$$C = \begin{pmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{pmatrix}$$

其中 $C_{11}: l \times m, C_{12}: l \times (r-m), C_{21}: (q-l) \times m, C_{22}: (q-l) \times (r-m)$, 则有

$$\begin{aligned} AC &= \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} \begin{pmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{pmatrix} \\ &= \begin{pmatrix} A_{11}C_{11}+A_{12}C_{21} & A_{11}C_{12}+A_{12}C_{22} \\ A_{21}C_{11}+A_{22}C_{21} & A_{21}C_{12}+A_{22}C_{22} \end{pmatrix} \end{aligned}$$

例 1.2.2(有用结论) 证明正交矩阵 $A: p \times p$ 的 p 个列向量和 p 个行向量都是一组正交单位向量。

^① 反射相当于旋转后再将其中的一个轴作镜面反射。在三维空间中,经反射后右(左)手坐标系将变为左(右)手坐标系。在统计上,反射和旋转所起的作用完全相同,本书后面将只提旋转,因为它更易于理解。

证明 将矩阵 A 分别按列向量和行向量分块, 并记

$$A = (a_1, a_2, \dots, a_p) = \begin{pmatrix} a'_{(1)} \\ a'_{(2)} \\ \vdots \\ a'_{(p)} \end{pmatrix}$$

由 $A'A = I$, 得

$$\begin{pmatrix} a'_1 \\ a'_2 \\ \vdots \\ a'_p \end{pmatrix} (a_1, a_2, \dots, a_p) = I$$

于是

$$\begin{pmatrix} a'_1 a_1 & a'_1 a_2 & \cdots & a'_1 a_p \\ a'_2 a_1 & a'_2 a_2 & \cdots & a'_2 a_p \\ \vdots & \vdots & \ddots & \vdots \\ a'_p a_1 & a'_p a_2 & \cdots & a'_p a_p \end{pmatrix} = \begin{pmatrix} 1 & & & \\ & 1 & & \\ & & \ddots & \\ 0 & & & 1 \end{pmatrix}$$

故有

$$a'_i a_j = \begin{cases} 1, & \text{若 } i=j \\ 0, & \text{若 } 1 \leq i \neq j \leq p \end{cases}$$

即 a_1, a_2, \dots, a_p 为一组正交单位向量。同理, 由 $AA' = I$ 可证 $a_{(1)}, a_{(2)}, \dots, a_{(p)}$ 也是一组正交单位向量。□

显然, 上例证明过程中的每一步推导都是可逆的。由此可见, 要验证一个方阵是否为正交矩阵只需验证一下其所有列或所有行是否为一组正交单位向量即可, 这比直接从正交矩阵的定义出发进行验证要方便些。

§ 1.3 行列式

p 阶方阵 $A = (a_{ij})$ 的行列式定义为

$$|A| = \sum_{j_1 j_2 \cdots j_p} (-1)^{\tau(j_1 j_2 \cdots j_p)} a_{1j_1} a_{2j_2} \cdots a_{pj_p} \quad (1.3.1)$$

这里 $\sum_{j_1 j_2 \cdots j_p}$ 表示对 $1, 2, \dots, p$ 的所有排列求和, $\tau(j_1 j_2 \cdots j_p)$ 是排列 j_1, j_2, \dots, j_p 中逆序的总数, 称它为这个排列的逆序数, 一个逆序是指在一个排列中一对数的前后位置与大小顺序相反, 即前面的数大于后面的数。例如, $\tau(3142) = 1 + \tau(1342) = 3 + \tau(1234) = 3$ 。

例 1.3.1

$$(1) \begin{vmatrix} 3 & 1 \\ 2 & 6 \end{vmatrix} = 3 \times 6 - 2 \times 1 = 16;$$

分量是 $\begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 2 & 0 \\ 0 & 3 & 0 & 0 \\ 4 & 0 & 0 & 0 \end{pmatrix}$ 是一个四阶行列式, 在展开式中应该有 $4! = 24$ 项, 但不为零的项只有

$a_{14}a_{23}a_{32}a_{41}$ 这一项, 而 $\tau(4321) = 6$, 所以, 该行列式为 $(-1)^6 \times 1 \times 2 \times 3 \times 4 = 24$. \square

由行列式的定义可以得到如下的一些基本性质:

- (1) 若 A 的某行(或列)为零, 则 $|A| = 0$.
- (2) $|A'| = |A|$.
- (3) 若将 A 的某一行(或列)乘以常数 c , 则所得矩阵的行列式为 $c|A|$.
- (4) 若 A 是一个 p 阶方阵, c 为一常数, 则 $|cA| = c^p |A|$.
- (5) 若互换 A 的任意两行(或列), 则行列式符号改变.
- (6) 若 A 的某两行(或列)相同, 则行列式为零.
- (7) 若将 A 的某一行(或列)的倍数加到另一行(或列), 则所得行列式不变.
- (8) 若 A 的某一行(或列)是其他一些行(或列)的线性组合, 则行列式为零.
- (9) 若 A 为上三角矩阵或下三角矩阵或对角矩阵, 则 $|A| = \prod_{i=1}^p a_{ii}$.

(10) 若 A 和 B 均为 p 阶方阵, 则 $|AB| = |A||B|$.

例 1.3.2 设 A 和 B 均为 p 阶方阵, 则 AB 和 BA 有相同的行列式. \square

(11) $|AA'| \geq 0$.

证明 由本章稍后 § 1.7 中的性质(6)和(5)即可证得. \square

(12) 若 A 与 B 都是方阵, 则

$$\begin{vmatrix} A & C \\ 0 & B \end{vmatrix} = \begin{vmatrix} A & 0 \\ C & B \end{vmatrix} = |A||B| \quad (1.3.2)$$

证明 设 $A = (a_{ij}) : p \times p, B = (b_{kl}) : q \times q$, 则由行列式的定义(1.3.1)可得

$$\begin{aligned} \begin{vmatrix} A & C \\ 0 & B \end{vmatrix} &= \sum_{\substack{j_1 \cdots j_p \\ l_1 \cdots l_q}} (-1)^{\tau(j_1 \cdots j_p (p+l_1) \cdots (p+l_q))} a_{1j_1} \cdots a_{pj_p} b_{1l_1} \cdots b_{ql_q} \\ &= \sum_{\substack{j_1 \cdots j_p \\ l_1 \cdots l_q}} (-1)^{\tau(j_1 \cdots j_p)} a_{1j_1} \cdots a_{pj_p} \cdot (-1)^{\tau(l_1 \cdots l_q)} b_{1l_1} \cdots b_{ql_q} \\ &= \sum_{j_1 \cdots j_p} (-1)^{\tau(j_1 \cdots j_p)} a_{1j_1} \cdots a_{pj_p} \cdot \sum_{l_1 \cdots l_q} (-1)^{\tau(l_1 \cdots l_q)} b_{1l_1} \cdots b_{ql_q} \\ &= |A||B| \end{aligned}$$

(13) 若 $A : p \times q, B : q \times p$, 则

$$|I_p + AB| = |I_q + BA| \quad (1.3.3)$$

证明 因为

$$\begin{pmatrix} I_p & A \\ 0 & I_q \end{pmatrix} \begin{pmatrix} I_p & -A \\ B & I_q \end{pmatrix} = \begin{pmatrix} I_p + AB & 0 \\ B & I_q \end{pmatrix}$$

$$\begin{pmatrix} I_p & 0 \\ -B & I_q \end{pmatrix} \begin{pmatrix} I_p & -A \\ B & I_q \end{pmatrix} = \begin{pmatrix} I_p & -A \\ 0 & I_q + BA \end{pmatrix}$$

上述两个等式两边各取行列式, 故得

(6) 若 A 和 C 为非退化方阵, $|I_p + AB| = |I_q + BA|$ □

例 1.3.3(有用结论) 设 x, y 为两个 p 维向量, 则

(7) p 阶方阵 A 是非退化的, 当 $|I_p + xy'| = 1 + y'x$ □

设 A 为 p 阶方阵, 将其元素 a_{ij} 所在的第 i 行与第 j 列划去之后所得 $(p-1)$ 阶矩阵的行列式, 称为元素 a_{ij} 的余子式, 记为 M_{ij} . $A_{ij} = (-1)^{i+j}M_{ij}$ 称为元素 a_{ij} 的代数余子式. 有以下公式成立

$$|A| = \sum_{j=1}^p a_{ij}A_{ij} = \sum_{i=1}^p a_{ij}A_{ij} \quad (1.3.4)$$

例 1.3.4 设

$$A = \begin{pmatrix} 6 & 5 & 10 & 1 \\ 7 & 10 & 7 & 6 \\ 9 & 8 & 12 & 2 \\ 4 & 9 & 11 & 3 \end{pmatrix}$$

则 a_{32} 的余子式为

$$M_{32} = \begin{vmatrix} 6 & 10 & 1 \\ 7 & 7 & 6 \\ 4 & 11 & 3 \end{vmatrix} = -191$$

其代数余子式为

$$A_{32} = (-1)^{3+2}M_{32} = -M_{32} = 191 \quad \square$$

§ 1.4 矩阵的逆

若方阵 A 满足 $|A| \neq 0$, 则称 A 为非退化(或非奇异)方阵; 若 $|A| = 0$, 则称 A 为退化(或奇异)方阵. 设 $A = (a_{ij})$ 是一非退化方阵, 若方阵 C 满足 $AC = I$, 则称 C 为 A 的逆矩阵, 记为 $C = A^{-1}$, A^{-1} 必是一个非退化矩阵. 令

$$B' = (A_{ij}) / |A| \quad (1.4.1)$$

其中 A_{ij} 是 a_{ij} 的代数余子式, 则容易验证 $AB = BA = I$. 由于 $C = BAC = B$, 因此 A^{-1} 是唯一的, 且 $(A^{-1})^{-1} = A$.

逆矩阵具有如下的基本性质:

(1) $AA^{-1} = A^{-1}A = I$.

(2) $(A')^{-1} = (A^{-1})'$.

(3) 若 A 和 C 均为 p 阶非退化方阵, 则

$$(AC)^{-1} = C^{-1}A^{-1}$$

(4) $|A^{-1}| = |A|^{-1}$.

(5) 若 A 是正交矩阵, 则 $A^{-1} = A'$.

(6) 若 $A = \text{diag}(a_{11}, a_{22}, \dots, a_{pp})$ 非退化(即 $a_{ii} \neq 0, i = 1, 2, \dots, p$), 则 $A^{-1} = \text{diag}(a_{11}^{-1}, a_{22}^{-1}, \dots, a_{pp}^{-1})$.

(7) 若 A 和 B 为非退化方阵, 则

$$\begin{pmatrix} \mathbf{A} & \mathbf{0} \\ \mathbf{0} & \mathbf{B} \end{pmatrix}^{-1} = \begin{pmatrix} \mathbf{A}^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{B}^{-1} \end{pmatrix} \quad (1.4.2)$$

例 1.4.1 设

$$\mathbf{A} = \begin{pmatrix} 2 & -1 \\ 3 & 6 \end{pmatrix}$$

于是 $|\mathbf{A}| = 15$, $A_{11} = (-1)^{1+1} \times 6 = 6$, $A_{12} = (-1)^{1+2} \times 3 = -3$, $A_{21} = (-1)^{2+1} \times (-1) = 1$, $A_{22} = (-1)^{2+2} \times 2 = 2$, 故 \mathbf{A} 的逆矩阵为

$$\mathbf{A}^{-1} = \frac{1}{15} \begin{pmatrix} 6 & 1 \\ -3 & 2 \end{pmatrix} = \begin{pmatrix} \frac{2}{5} & \frac{1}{15} \\ -\frac{1}{5} & \frac{2}{15} \end{pmatrix}$$

例 1.4.2 设

$$\mathbf{A} = \begin{pmatrix} 2 & -1 & 0 \\ 3 & 6 & 0 \\ 0 & 0 & -4 \end{pmatrix}$$

则

$$\mathbf{A}^{-1} = \begin{pmatrix} (2 & -1)^{-1} & 0 & 0 \\ 3 & 6 & 0 & 0 \\ 0 & 0 & (-4)^{-1} & 0 \end{pmatrix} = \begin{pmatrix} \frac{2}{5} & \frac{1}{15} & 0 \\ -\frac{1}{5} & \frac{2}{15} & 0 \\ 0 & 0 & -\frac{1}{4} \end{pmatrix}$$

§ 1.5 矩阵的秩

一组同维向量 a_1, a_2, \dots, a_n , 若存在不全为零的常数 c_1, c_2, \dots, c_n , 使得

$$c_1 a_1 + c_2 a_2 + \dots + c_n a_n = \mathbf{0} \quad (1.5.1)$$

则称该组向量线性相关。若向量 a_1, a_2, \dots, a_n 不线性相关, 就称为线性无关。如果等式(1.5.1)成立, 则至少有一个向量 a_i 可以表示成该组其余向量的线性组合, 这意味着该组有“多余”的向量, 而在线性无关的向量组中却没有这种“多余”的向量。

矩阵 \mathbf{A} 的线性无关行向量的最大数目称为行秩, 其线性无关列向量的最大数目称为列秩。矩阵的行秩和列秩必相等, 故统一将其称为 \mathbf{A} 的秩, 记作 $\text{rank}(\mathbf{A})$ 。

矩阵的秩具有下述基本性质:

(1) $\text{rank}(\mathbf{A}) = 0$, 当且仅当 $\mathbf{A} = \mathbf{0}$ 。

(2) 若 \mathbf{A} 为 $p \times q$ 矩阵, 且 $\mathbf{A} \neq \mathbf{0}$, 则 $1 \leq \text{rank}(\mathbf{A}) \leq \min\{p, q\}$ [若 $\text{rank}(\mathbf{A}) = p$, 则称 \mathbf{A} 为行满秩的; 若 $\text{rank}(\mathbf{A}) = q$, 则称 \mathbf{A} 为列满秩的]。

(3) $\text{rank}(\mathbf{A}) = \text{rank}(\mathbf{A}')$ 。

(4) $\text{rank} \begin{pmatrix} \mathbf{A} & \mathbf{0} \\ \mathbf{0} & \mathbf{B} \end{pmatrix} = \text{rank} \begin{pmatrix} \mathbf{0} & \mathbf{A} \\ \mathbf{B} & \mathbf{0} \end{pmatrix} = \text{rank}(\mathbf{A}) + \text{rank}(\mathbf{B})$ 。

(5) $\text{rank}(\mathbf{AB}) \leq \min\{\text{rank}(\mathbf{A}), \text{rank}(\mathbf{B})\}$ 。