

多元统计分析是统计学中一个非常重要的分支,它起源于20世纪初,1928年J.维希特(J. Wishart)发表论文《多元正态总体样本协差阵的精确分布》,被认为是多元统计分析历史的开始。20世纪30年代,R.A.费希尔(R.A. Fisher)、H.霍特林(H. Hotelling)等人做了一系列奠基性的工作,使多元统计分析在理论上得到迅速发展。

Multivariate statistical analysis

多元统计分析

花 靳刘蕊 王 兢 ◎ 著

20世纪40年代,多元统计分析在心理学、教育学、生物学等领域获得了一些应用。但由于当时需要大量的计算,加上第二次世界大战的影响,使其发展受到阻碍,甚至停滞了相当长的时间。20世纪50年代中期开始,随着电子计算机的发展和普及,多元统计分析在地质、气象、生物、医学、经济分析等诸多领域得到了广泛的应用,同时也促进了理论的发展。

各种统计软件(如SAS、SPSS等)的开发和使用,使实际工作者利用多元统计分析方法解决实际问题更加简单方便。20世纪80年代起在我国许多领域拉开了多元统计分析应用的帷幕,并在其理论研究和应用上取得了显著成就。对实际领域中的研究者和高等院校的学生来说,要学习掌握多元统计分析的各种模型和方法,手头有一本既有多元统计分析理论又有实际案例分析的参考书是非常必要的,为此我们编著了这本《多元统计分析》。



多元统计分析

符想花 靳刘蕊 王兢 ◎ 著

 郑州大学出版社

图书在版编目(CIP)数据

多元统计分析/符想花,靳刘蕊,王兢著. —郑州:郑州大学出版社,2009.12

ISBN 978 - 7 - 5645 - 0170 - 9

I. 多… II. ①符…②靳…③王… III. 多元分析:统计分析
IV. 0212.4

中国版本图书馆 CIP 数据核字 (2009) 第 199137 号

郑州大学出版社出版发行

郑州市大学路 40 号

出版人:王 锋

全国新华书店经销

郑州美联印刷有限公司印制

开本:787 mm × 1 092 mm

印张:20.25

字数:365 千字

版次:2009 年 12 月第 1 版

邮政编码:450052

发行部电话:0371 - 66966070

1/16

印次:2009 年 12 月第 1 次印刷

书号:ISBN 978 - 7 - 5645 - 0170 - 9 定价:36.00 元

本书如有印装质量问题,由本社负责调换

前 言

多元统计分析是统计学中一个非常重要的分支,它起源于20世纪初,1928年J. 维希特(J. Wishart)发表论文《多元正态总体样本协差阵的精确分布》,被认为是多元统计分析历史的开始。20世纪30年代,R. A. 费希尔(R. A. Fisher)、H. 霍特林(H. Hotelling)等人做了一系列奠基性的工作,使多元统计分析在理论上得到迅速发展。20世纪40年代,多元统计分析在心理学、教育学、生物学等领域获得了一些应用。但由于在应用时需要大量的计算,加上第二次世界大战的影响,使其发展受到影响,甚至停滞了相当长的时间。20世纪50年代中期开始,随着电子计算机的发展和普及,多元统计分析在地质、气象、生物、医学、经济分析等许多领域得到了广泛的应用,同时也促进了理论的发展。各种统计软件(如SAS, SPSS等)的开发和使用,使实际工作者利用多元统计分析方法解决实际问题更加简单方便。20世纪80年代起在我国许多领域拉开了多元统计分析应用的帷幕,并在其理论研究和应用上取得了显著成就。对实际领域中的研究者和高等院校的学生来说,要学习掌握多元统计分析的各种模型和方法,手头有一本既有多元统计分析理论又有实际案例分析的参考书是非常必要的,为此我们编著了这本《多元统计分析》。

本书的写作突出了以下几点:①对数学基础知识的要求较低,读者只需掌握线性代数、概率论与数理统计、微积分的基本知识,即便如此,本书还是专门对矩阵代数知识作了简单的回顾。②在一元统计分析的基础上深入浅出地介绍多元统计分析的内容,这有助于读者对内容的理解和掌握。③本书在系统介绍多元统计分析基本思想和分析问题方法的同时,结合社会科学和经济学等领域的案例进行分析,将理论和实际密切地结合在了一起。④多元统计分析方法的

应用离不开计算机及数据处理与分析软件,本书的案例分折基本上是借助于 SPSS 软件实现的,而且在各章的最后一节结合案例介绍了其在 SPSS 软件上的实现过程,并对部分输出结果做了解释和说明。⑤在每章的后面都附有一定数量的思考与练习题,目的是让读者进一步掌握多元统计分折方法及其在 SPSS 软件上的实现过程。

本书共分 12 章。第一章绪论,介绍多元统计分折的作用及应用范围。第二章对多元统计分折中常用的矩阵代数知识做一些简单的回顾和介绍。第三章和第四章介绍多元统计分折的基本概念和基本理论,这些内容是在复习一元统计分折内容的基础上自然引申出来的。第五章和第六章主要介绍分类问题,包括聚类分折和判别分折。第七章到第九章介绍数据简化问题,包括主成分分折、因子分折和对应分折。第十章和第十一章阐述两组变量之间的相互关系,包括典型相关分折和多重多元回归分折。第十二章是定性资料的统计分折。

读者按每章所述方法对实际问题进行计算、分折,相信一定会有收获的。

由于作者水平有限,书中难免有不足之处,恳请读者批评指正。

作者

2009 年 4 月

目 录

第一章 绪 论	(1)
第一节 多元统计分析的作用	(1)
第二节 主要内容安排	(4)
第二章 矩阵代数	(6)
第一节 矩阵及基本运算	(6)
第二节 行列式、逆矩阵和矩阵的秩	(11)
第三节 特征根、特征向量和矩阵的迹	(14)
第四节 正定阵和非负定阵	(16)
第五节 矩阵代数运算的上机实现	(17)
第三章 多元正态分布	(27)
第一节 一元统计分析中的有关概念	(27)
第二节 多元统计分析中的基本概念	(30)
第三节 多元正态分布的定义及基本性质	(36)
第四节 多元正态分布的参数估计	(40)
第五节 $\bar{X}$ 和 S 的抽样分布	(42)
第六节 随机向量数字特征的上机实现	(43)
第四章 多元正态总体均值向量和协方差阵的假设检验	(53)
第一节 均值向量的检验	(53)
第二节 协方差阵的检验	(66)
第三节 有关检验的上机实现	(68)
第五章 聚类分析	(82)
第一节 聚类分析的基本思想	(82)
第二节 相似性度量	(83)
第三节 系统聚类方法	(87)
第四节 其他聚类法	(109)
第五节 聚类分析的上机实现	(115)

第六章 判别分析	(128)
第一节 距离判别	(128)
第二节 贝叶斯(Bayes)判别	(141)
第三节 费希尔判别	(146)
第四节 逐步判别	(153)
第五节 判别分析的上机实现	(161)
第七章 主成分分析	(177)
第一节 主成分分析的基本思想及数学模型	(177)
第二节 主成分分析的几何意义	(178)
第三节 主成分的推导及其性质	(180)
第四节 主成分分析中有关问题的讨论	(185)
第五节 主成分分析的应用	(190)
第六节 主成分分析的上机实现	(196)
第八章 因子分析	(204)
第一节 因子分析的基本理论	(204)
第二节 因子分析的数学模型	(205)
第三节 因子载荷矩阵的估计方法	(208)
第四节 因子旋转	(211)
第五节 因子得分	(214)
第六节 因子分析的应用	(217)
第七节 因子分析的上机实现	(221)
第九章 对应分析	(229)
第一节 对应分析的基本思想	(229)
第二节 列联表及列联表分析	(230)
第三节 对应分析的基本原理	(232)
第四节 对应分析的应用	(236)
第五节 对应分析的上机实现	(241)
第十章 典型相关分析	(252)
第一节 典型相关分析及基本思想	(252)
第二节 总体典型相关	(253)
第三节 样本典型相关	(256)
第四节 典型相关系数的显著性检验	(258)

第五节 典型相关系数的计算步骤	(259)
第十一章 多重多元回归分析	(264)
第一节 多重多元线性回归分析	(264)
第二节 双重筛选逐步回归分析	(272)
第十二章 定性资料的统计分析	(278)
第一节 定性变量数量化	(278)
第二节 对数线性模型	(279)
第三节 Logistic 回归	(283)
第四节 对数线性模型和 Logistic 回归模型的上机实现	(290)
附录一	(302)
附录二	(307)
参考文献	(316)

第一章 绪 论

多元统计分析是统计学中一个非常重要的分支,它起源于20世纪初,1928年J. 维希特(J. Wishart)发表论文《多元正态总体样本协差阵的精确分布》,被认为是多元统计分析历史的开始。20世纪30年代,R. A. 费希尔(R. A. Fisher)、H. 霍特林(H. Hotelling)、许宝禄以及S. N. 罗伊(S. N. Roy)等人作了一系列奠基性的工作,使多元统计分析在理论上得到迅速发展。20世纪40年代,多元统计分析在心理学、教育学、生物学等领域获得了一些应用。但由于在应用时需要大量的计算,加上第二次世界大战的影响,使其发展受到影响,甚至停滞了相当长的时间。20世纪50年代中期开始,随着电子计算机的发展和普及,多元统计分析在地质、气象、生物、医学、经济分析等许多领域得到了广泛的应用,同时也促进了理论的发展。各种统计软件(如SAS,SPSS等)的开发和使用,使实际工作者利用多元统计分析方法解决实际问题更加简单方便。20世纪80年代起在我国许多领域拉开了多元统计分析应用的帷幕,并在其理论研究和应用上取得了显著成就。

第一节 多元统计分析的作用

一、什么是多元统计分析

在工业、农业、医学、气象、环境以及经济、管理等众多领域,常常需要同时观测多个指标来分析和研究问题。例如,在经济管理中,要对国有企业资本金绩效进行评价,需观测净资产收益率、总资产报酬率、总资产周转率、流动资产周转率、资产负债率、已获利息倍数、销售增长率和资本积累率等多个指标。在统计学上,通常将指标称为变量,变量有确定性变量和随机性变量。由于受某些确定性因素的影响,现象的量沿着某一方向持续变化,这样的量就是确定性变量。如由于科学技术的不断提高和医疗卫生条件的不断改善,人类的平均寿命在不断延长,因此,从长期看,人的平均寿命是一确定性变量。而有些变量的变动受许多因素变动的影响,变量值的大小没有明确的方向,这样的变量就是随机性变量。如净资产收益率、总资产报酬率、总资产周转率等在不同的企业表现为不同的数值,这些变量都是随机性变量。

如何同时对多个随机变量的观测数据进行有效的分析和研究呢?传统的方法是把多个随机变量分开进行分析和研究,一次只处理一个指标(变量),这种方法忽视了指标之间可能存在的相关性,不仅会丢失很多信息,也不能取得较好的研究结果。而现代的分析方法是对多个变量同时进行分析研究,通过对多个随机变量观测数据的分析,来研究变量之间的相互关系以及揭示变量内在的变化规律。一元统计分析是研究一个随机变量统计规律的科学;而多元统计分析是研究多个随机变量之间相互依赖关系及内在统计规律的一门统计学科。

二、多元统计分析方法的作用

(一)能够简化数据和数据结构

要认识客观现象,往往需要从多角度、多方面进行系统的相互联系的考察,这必然要用到统计指标体系,构成统计指标体系的多个指标,各有侧重地解释、说明着同一事物,多重共线性的出现不可避免。为此,在不损失数据蕴含信息量的情况下,通过变换数据和构造模型剔除指标间相互制约的成分,尽可能简单地将被研究现象描述出来,就可达到简化数据和数据结构的目的。主成分分析、因子分析等均可达到这样的目的。

(二)能够进行分类或分组

通过统计分组将性质相同的单位(或者变量)归为一组,将性质不同的单位(或者变量)归为不同的组,有利于对问题做深入细致的分析和研究。根据所测量到的某些特征数据,同时对研究现象进行分类或分组,是多元统计分析的另一目的所在。聚类分析和判别分析等可达到这样的目的。

(三)能够研究变量间的相互依赖关系

人们对变量间关系的本质感兴趣:是否所有的变量都相互独立?还是一个变量或一些变量依赖于另一个或一些变量?如果是后者,这种依赖关系是怎样的?相关分析和回归分析就可解决此类问题。

(四)能够进行预测

在生产与生活中,探索多变量系统内在的客观规律性及其与外部环境的关系,进行预测预报,以实现对系统的最优控制,是应用多元统计分析技术的主要目的,而多个变量之间关系的建立,为从一些变量观测值预测另一个或一些变量值提供了可能。

(五)能够进行假设检验

检验多元总体参数表示的某种统计假设,可以证实某种假设的合理性或支持事先树立的某种信念。

三、多元统计分析在现实生活中经常处理的问题

多元统计分析作为科学研究的重要工具,在自然科学、社会科学等诸多方面有着广泛的应用。

(一)在经济分析中经常处理的问题

1. 要对我国大型工业企业的经济效益进行综合评价,其做法是选取能够反映经济效益的代表性指标,如:百元固定资产实现利税、资金产值率、资金利税率、全员劳动生产率等,根据这些指标对大型工业企业分类,根据分类结果可给出对企业的评价。对大型工业企业分类应用的是聚类分析法。

2. 给出美国、日本、英国等经济发达国家和印度、罗马尼亚、南非等发展中国家反映其经济发展程度的若干指标,如:人均国内生产总值、人均收入、人均消费支出等,要根据这些指标判断中国所属类型,应用的是判别分析法。

3. 要对我国大型工业企业经济效益进行综合评价,评选出前 500 强,其做法同样是选取能够反映经济效益的代表性指标,如:百元固定资产实现利税、资金产值率、资金利税率、全员劳动生产率等,根据这些指标采用主成分分析或因子分析法即可达到此目的。

4. 某一产品质量的好坏可用多个指标进行衡量,而影响产品质量好坏的因素亦有多个,要揭示影响产品质量的多个因素与衡量产品质量的多个指标之间的依赖关系,需要用典型相关分析法,如果它们之间存在相关关系,这时要用多重多元回归分析法建立回归模型,以便进行预测预报。

5. 不同原料生产的两种产品,其产品寿命有无显著性差异,可进行假设检验。

(二)在其他领域研究中的应用

1. 在一项癌症患者对放射性疗法反应的研究中,对一批患者测量多项反应指标,因为难以同时解释所有变量的观测值,于是人们需要寻找刻画患者的一个简单综合性指标,主成分分析或因子分析能够帮助研究者达到这一目标,且不损失数据中蕴含的许多信息。

2. 在地质勘探中,根据岩石标本中的多种特征来判别地质的地质年代,有矿还是无矿,若有矿,是何矿种等。

3. 在考古学中,根据挖掘出的动物牙齿的有关测试指标,判别它属于哪一类动物牙齿,是哪一个时代的动物。

4. 在植物育种学中,植物收获时,要选择种子,以保证在一些特征上,下一代优于上一代,这时往往要测量、评价很多特征,将这些特征通过变量变换变为

综合指标,根据综合指标得分的高低来选择种子。

5. 在社会学中,对当代大学生择偶标准(对文化程度、职业、居住条件、收入、相貌等的要求)进行调查,在此基础上做主要因素分析,以便对他们的婚姻家庭观作正确引导。

6. 在文学研究中,同样可以应用多元分析方法。众所周知,《红楼梦》一书共 120 回,一般认为,前 80 回为曹雪芹所写,后 40 回为高鹗所续,但长期以来对这一结论存在争议。1985、1986 年复旦大学的李贤平教授,带领他的学生对《红楼梦》一书采用多元统计分析的方法进行了研究,他们创造性地将 120 回看成 120 个样品,并确定与情节无关的虚词作为变量,让学生数出每一回里各变量出现的次数,用聚类分析的方法进行了分类,分类的结果是前 80 回为一类,后 40 回为一类。为了判断前 80 回是否为曹雪芹所写,他们找到曹雪芹的其他著作,做了类似的分析计算,证实用词手法完全相同,结果断定前 80 回为曹雪芹一人所写。而后 40 回是否为高鹗所写? 论证结果是并非如此。他们用多元统计分析的方法支持了红学界的观点,使红学界大为赞叹。

尽管上述问题不同、领域不同,但从对数据的处理上看,却有很多相似之处。正是多元统计分析方法能够应用于各个领域,才使得该学科能够伴随着计算技术的普及而得到广泛的应用。

第二节 主要内容安排

本书共分十二章。第一章是绪论,介绍多元统计分析的作用及其应用范围。第二章是矩阵代数,在多元统计分析中,需要用到有关矩阵代数知识,对矩阵代数知识作一些简单的回顾和介绍,将为以后各章的学习带来很大的方便。第三章和第四章介绍多元统计分析的基本概念和基本理论,包括:四个主要的统计量分布——多元正态分布、Wishart 分布、Hotelling T^2 分布、Wilks 分布,多元正态总体的参数估计和假设检验。而这些内容都是一元统计分析中相应内容的推广,因此这两章的内容是在复习一元统计分析内容的基础上自然引申出来的,这样读者在学习时就不会感觉抽象和困难。

第五章和第六章阐述分类问题。聚类分析是研究“物以类聚”的一种现代多元统计分析方法,判别分析是在研究对象用某种方法分好若干类的情况下,确定新样品属于已知类别中哪一类的多元统计分析方法。实际应用时这两种方法往往结合起来使用,因为判别分析要求对新样品进行判别归类之前,应先知对总体分了几类,然后建立判别式,根据一定的判别规则对新样品进行

判别归类。如果在判别分析之前,并不知道事先划分了几类,这时可先做聚类分析然后再做判别分析。

第七章到第九章介绍主成分分析、因子分析和对应分析。主要阐述数据简化问题,将具有错综复杂关系的变量(或样品)综合成数量较少的综合指标,用较少的综合指标以尽可能简单地表示所研究的对象,又不至于损失很多有价值的信息。

第十章和第十一章介绍两组变量之间的相互关系,包括典型相关分析和多重多元回归分析。典型相关分析能够反映两组变量之间相互线性依赖关系,但根据典型相关关系不能通过自变量的取值来求解因变量的可能取值,要解决这一问题,需要多重多元回归分析。这两章内容是在简单相关和一元回归分析的基础上引申而来的。

第十二章是定性资料的统计分析。在实际应用中,不可避免地要涉及到定性变量,例如人的性别、人的职业、天气状况、产品等级等。对诸如这些定性变量给以相应的数量描述,从而进行有关的统计分析,是定性资料统计分析要解决的问题。但本章并不详细介绍这方面的理论、方法,而是初步反映一下这方面的内容,目的是展示进一步可学的知识,以便更好地解决实际问题。

本书除介绍统计分析方法的理论外,还有统计分析方法在 Excel、SPSS 软件中的实现,而且各章还配有相应的练习,读者按每章所述方法去计算、分析,相信一定会有收获的。

【思考与练习】

1. 简述多元统计分析的作用。
2. 举例说明多元统计分析在实际工作中的应用。

第二章 矩阵代数

矩阵代数是研究多元统计分析的重要工具,为学习方便,本章将对书中需要的有关矩阵代数知识进行回顾和介绍。

第一节 矩阵及基本运算

一、矩阵的定义

将 $n \times p$ 个实数 $a_{11}, a_{12}, \dots, a_{1p}; a_{21}, a_{22}, \dots, a_{2p}; \dots; a_{n1}, a_{n2}, \dots, a_{np}$ 排列成如下形式的矩形数表,记为 A

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1p} \\ a_{21} & a_{22} & \cdots & a_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{np} \end{pmatrix}$$

则称 A 为 $n \times p$ 阶矩阵。一般记为 $A = (a_{ij})_{n \times p}$, 其中 a_{ij} 是 A 中第 i 行第 j 列的元素。

如果 A 的所有元素都为 0, 则称 A 为 $n \times p$ 阶零矩阵, 记为 $A = 0_{n \times p}$ 或 $A = 0$ 。

当 $n = p$ 时, 称 A 为 p 阶方阵, 此时

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1p} \\ a_{21} & a_{22} & \cdots & a_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ a_{p1} & a_{p2} & \cdots & a_{pp} \end{pmatrix} = (a_{ij})_{p \times p}$$

称 $a_{11}, a_{22}, \dots, a_{pp}$ 为 A 的对角线元素, 其他元素称为非对角线元素。

如果一个方阵只有对角线元素不为零, 非对角线元素全为零, 则称 A 为对角矩阵, 记为

$$A = \begin{pmatrix} a_{11} & 0 & \cdots & 0 \\ 0 & a_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_{pp} \end{pmatrix} = \text{diag}(a_{11}, a_{22}, \dots, a_{pp})$$

如果一个 p 阶对角矩阵中, 对角线元素均为 1, 即

$$A = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}$$

称 A 为 p 阶单位阵, 可记为 $A = I_p$ 或 $A = I$ 。

如果一个 p 阶方阵中, 当 $i > j$ 时, 所有元素均为零, 则

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1p} \\ 0 & a_{22} & \cdots & a_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_{pp} \end{pmatrix}$$

称 A 为上三角矩阵; 当 $i < j$ 时, 所有元素均为零, 则称 A 为下三角矩阵。

若将矩阵 A 的行与列互换, 则称其所形成的矩阵为 A 的转置矩阵, 记为 A' , 即

$$A' = \begin{pmatrix} a_{11} & a_{21} & \cdots & a_{n1} \\ a_{12} & a_{22} & \cdots & a_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1p} & a_{2p} & \cdots & a_{np} \end{pmatrix}$$

若 A 是方阵, 且 $A = A'$, 则 A 为对称矩阵。

只有一行的矩阵称为行矩阵, 也叫行向量, 如:

$$a = (a_{11}, a_{12}, \cdots, a_{1p})$$

同样, 只有一列的矩阵称为列矩阵, 也叫列向量, 如:

$$a = \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{n1} \end{pmatrix}$$

二、矩阵的运算

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

$$A + B = (a_{ij} + b_{ij})_{n \times p}$$

若 c 为一常数, 则它与 $A = (a_{ij})_{n \times p}$ 的积定义为

$$cA = (ca_{ij})_{n \times p}$$

若 $A = (a_{ik})_{n \times p}$, $B = (b_{kj})_{p \times r}$, 则 A 与 B 的积定义为

$$A \times B = \left(\sum_{k=1}^p a_{ik} b_{kj} \right)_{n \times r}$$

一般情况下, $AB \neq BA$ 。

从上述定义中, 容易得出如下运算规律:

1. $A + (-1)A = 0$

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

3. $(A')' = A$

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

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

6. $A(B+C) = AB + AC$

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

8. $AI = IA = A$

9. 对任意常数 d , 恒有 $d(A+B) = dA + dB$

10. 对任意常数 d , 有 $d(AB) = (dA)B = A(dB)$

其中 A, B, C 是适合运算的矩阵。

例 2.1 设 $A = \begin{pmatrix} 0 & 3 & 1 \\ 1 & -1 & 1 \end{pmatrix}$, $B = \begin{pmatrix} 1 & -2 & -3 \\ 2 & 5 & 1 \end{pmatrix}$, $C = \begin{pmatrix} 1 & 1 \\ 2 & 3 \end{pmatrix}$

则

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

(2) $A+B = \begin{pmatrix} 0+1 & 3-2 & 1-3 \\ 1+2 & -1+5 & 1+1 \end{pmatrix} = \begin{pmatrix} 1 & 1 & -2 \\ 3 & 4 & 2 \end{pmatrix}$

(3) $(A+B)' = A' + B' = \begin{pmatrix} 1 & 3 \\ 1 & 4 \\ -2 & 2 \end{pmatrix}$

(4) $4A = \begin{pmatrix} 0 & 12 & 4 \\ 4 & -4 & 4 \end{pmatrix}$

(5) $CA = \begin{pmatrix} 1 & 1 \\ 2 & 3 \end{pmatrix} \begin{pmatrix} 0 & 3 & 1 \\ 1 & -1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 2 \\ 3 & 3 & 5 \end{pmatrix}$

三、正交矩阵

若方阵 A 满足 $AA' = A'A = I$, 则称 A 为正交矩阵。如:

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & -1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ -\frac{2}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} \end{pmatrix}, \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$$

正交矩阵有如下性质:

1. 正交矩阵为非奇异矩阵 ($|A| \neq 0$), 存在逆矩阵 A^{-1} , 且 A^{-1} 仍为正交矩阵。

2. 若 A 为正交矩阵, 则它的转置矩阵等于它的逆矩阵, 即 $A' = A^{-1}$ 。

3. 如果 $A = (a_{ij})_{p \times p}$ 和 $B = (b_{ij})_{p \times p}$ 都是 p 阶正交矩阵, 则

(1) AB 仍是 p 阶正交矩阵

$$(2) \sum_{k=1}^p a_{ik} a_{jk} = \begin{cases} 1 & (i=j) \\ 0 & (i \neq j) \end{cases}, \sum_{k=1}^p a_{ki} a_{kj} = \begin{cases} 1 & (i=j) \\ 0 & (i \neq j) \end{cases}$$

即正交矩阵的 p 个行向量相互正交, p 个列向量也彼此正交。

若 T 为正交矩阵, 则称 $T'AT$ 为矩阵 A 的正交变换, 对称矩阵经过正交变化后仍为对称矩阵。

四、投影矩阵

若方阵 A 满足 $A^2 = A$, 则称 A 为幂等矩阵。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。

五、矩阵的分块

矩阵的分块是在处理阶数较高的矩阵时常用的方法。分块矩阵是: 对于任意一个矩阵 A , 按某种需要用纵线和横线将它们划分成若干块低阶的矩阵, 也可以看作是以所分成的子块为元素的矩阵。设 $A = (a_{ij})_{n \times p}$, 将它分成四块, 即

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

其中 $A_{11} = (a_{ij})_{k \times l}$, $A_{12} = (a_{ij})_{k \times (p-l)}$, $A_{21} = (a_{ij})_{(n-k) \times l}$, $A_{22} = (a_{ij})_{(n-k) \times (p-l)}$ 。