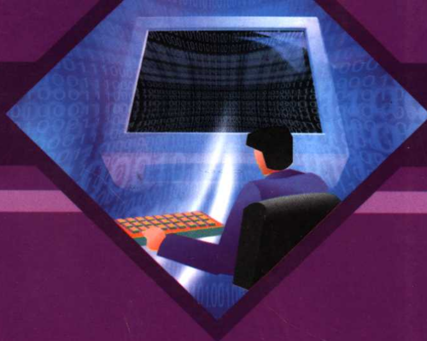

中国地质大学（武汉）研究生系列教材

实用多元统计分析

SHIYONG DUOYUAN TONGJI FENXI

主　编　向东进　副主编　李宏伟　刘小雅

中国地质大学（武汉）研究生教材建设基金资助

中国地质大学出版社

实用多元统计分析

主　编：向东进

副主编：李宏伟　刘小雅

中国地质大学出版社

内容简介

本书以介绍实用数理统计方法为目的，内容包括：矩阵代数与概率统计的一些基础知识、回归分析、判别分析、聚类分析、主成分分析与典型相关分析、因子分析与对应分析，书中例题大都是各个领域的实际应用问题，基本上是利用功能强大的统计软件 SPSS 计算的。

本书主要为理工科非数学专业研究生编写，也可以作为信息、管理类等有关专业的高年级本科生教材，还可以作为统计方法应用工作者的参考书籍。

图书在版编目(CIP)数据

实用多元统计分析/向东进主编；李宏伟，刘小雅副主编．—武汉：中国地质大学出版社，2005.9

ISBN 7-5625-2047-X

Ⅰ．实…

Ⅱ．向…

Ⅲ．多元统计-分析方法

Ⅳ．O212

实用多元统计分析

向东进　主　编

李宏伟　刘小雅　副主编

责任编辑：方　莉　　　　责任校对：张咏梅

出版发行：中国地质大学出版社（武汉市洪山区鲁磨路 388 号）　　邮编：430074

电话：(027)87482760　传真：87481537　　E-mail：cbb @ cug.edu.cn

经　销：全国新华书店　　http://www.cugp.cn

开本：787 毫米×1092 毫米　1/16　　字数：300 千字　印张：11.75

版次：2005 年 9 月第 1 版　　印次：2005 年 9 月第 1 次印刷

印刷：中国地质大学出版社印刷厂　　印数：1—1 000 册

ISBN 7-5625-2047-X/O·73　　定价：25.00 元

“研究生系列教材”总序

在中国地质大学研究生院建院二十周年来临之际，第一批反映我校研究生教学与科学研究成果的“研究生系列教材”出版了，这是我校研究生教育发展过程中的一件大事，可喜可贺！

随着我校研究生招生规模的不断扩大，如何保证研究生的培养质量是我们必须积极思考并努力着手解决的问题。这套研究生系列教材的及时出版，正是一个很有力的举措。研究生教材建设是保证和提高研究生培养质量的重要手段，是反映一个学校教师队伍的学术水平和教学水平的宏观尺度，更是具有战略性意义的基本建设。各门课程必须有高质量的教材，才能使研究生通过学习，掌握各门学科的坚实的基础理论和系统的专门知识，为从事科学研究工作打下良好的基础。因此，我校研究生院筹集资金设立了“研究生教材建设基金”，资助出版“研究生系列教材”以满足本校各学科研究生教学的需要，促进我校研究生教材建设工作，提高研究生培养质量。

由于研究生具有人才的高层次性、培养的超前性和学习的研究性等特点，这就要求研究生教材并不是本科生教材的简单深化和延续，而应该结合学校的学科专业结构和特色来编写系统性、新颖性、适用性融为一体的研究生教材。这套“研究生系列教材”以具有我校特色的研究生课程教材为主，既有基础理论教材，又有研究生专业课教材，准备在今后数年内分批次出版。“研究生系列教材”总的特色是从我校研究生的教学实际需要出发，根据各门课程在各专业研究生培养中的地位和作用，在内容上求新、求深、求精。专业课程教材还要力求高起点，反映科学规律，追踪该学科专业的发展前沿，反映国内外的最新研究成果。

虽然，我们的主观愿望是尽可能组织编写出一套特色鲜明、适用性强的高质量“研究生系列教材”，但由于我校研究生教材建设工作起步不久，经验不足，已出版的教材质量尚待在使用中检验，敬请校内外专家学者及读者不吝指教，我们将非常感谢。

姚书振

中国地质大学(武汉)研究生院　院长

2005 年 5 月 20 日

目　　录

第一章　绪　论

§1.1　什么是多元统计分析

科学研究是一个反复学习的过程.无论是研究自然现象还是社会现象,人们往往以解释某种现象或者预测某种变化趋势作为目标,然后通过收集数据并分析数据对这些目标进行检验,通常还会对所研究的现象提出一个改进的解释.在这个过程中,常常需要同时观测多个指标.例如,要衡量一个地区的经济发展状况,需要观测的指标有:总产值、利润、效益、劳动生产率、万元产值能耗、固定资产、流动资金周转率、物价、信贷、税收等等;在医学诊断中,要判断某人是否患病,也需要做多项指标的体检,如:血压、脉搏跳动次数、白血球、体温等等.像这样需要处理多个变量的观测数据的情况,如何进行有效的分析和研究呢?如果用一元统计方法,则必须把多个变量分开分析,一次处理一个变量,这样做有时候也许比较方便有效,但由于这种方法忽视了诸多变量间可能存在的相关性,因此,一般会丢失很多信息.另一种方法就是本书要讨论的多元统计方法,它同时对多个变量的观测数据进行分析.这样的分析对诸变量之间的关系、相依性和相对重要性都能提供有用信息.

许多多元分析方法的基本依据是多元正态分布的概率模型.因此,多元统计分析包括的主要内容有:多元正态总体的参数估计和假设检验以及常用的统计方法.这些方法是:多元数据图表示法、聚类分析、判别分析、回归分析、主成分分析、因子分析、对应分析、典型相关分析、路径分析和多维标度法等.本书主要介绍多元分析的各种常用方法.随着计算机技术的发展,已经产生了一些功能十分强大的标准统计分析软件包,从而使那些原本相当复杂方法的实现步骤变得容易起来.SPSS 因其功能强大、容易学习、操作方便而备受国内科技工作者的青睐.本书中大多数应用实例是经由 SPSS 计算而得.

1.1.1　研究的内容和方法

英国著名统计学家肯德尔(Kendall)把多元统计分析所研究的内容和方法概括为以下几个方面:

(1)数据或结构化简．通过变量变换等方法将相互依赖的关系变成不相关,或者把高维空间的数据投影到低维空间,在不损失有价值信息的前提下尽可能简单地将被研究的现象描述出来,并希望这种表示能够很容易地进行解释.主成分分析、因子分析和对应分析都属于这类方法.

(2)分类和分组．即根据所测量的特征将一些"类似的"对象或变量分组,同时给出好的分组法则.如聚类分析和判别分析是解决这些问题的统计方法.

(3)研究变量间的依赖关系．人们对变量间关系的本质感兴趣.变量间是否相互独立?还

是有一个或多个变量依赖于其他变量？如果是后者，那么这种关系又是怎样的？能否建立变量之间的定量关系并用于预测或控制？可用回归分析方法解决这样的问题．如果要分析两组变量之间的相互关系，可用典型相关分析．

(4)多元数据的统计推断．主要是考虑参数估计和假设检验问题．尤其是与多元正态分布有关的估计和假设检验问题．通过假设检验，证实某些假设条件的合理性或支持原有的某种理论．

(5)多元统计分析的理论基础．多元统计分析的理论基础包括多维随机向量，以及各种多元统计量，推导它们的分布并研究其性质，研究它们的抽样分布理论．

1.1.2　多元统计分析的应用

多元统计方法是科学研究的一种重要工具，其应用颇为广泛，下面我们列举一些这样的实际问题．

1. 医学

(1)在一项关于癌症患者对放射疗法反应的研究中，对一批患者测量六项反应指标．因为难于同时解释所有变量的观测值，于是人们需要寻找刻画患者反应的一个简单综合性指标．可以采用主成分分析方法达到这一数据化简的目的，而且不损失数据中蕴涵的主要信息．

(2)对视觉刺激(如闪光)的反应情况能够从头皮上用电脑记录下来．在一项关于视觉系统多重硬化效应研究中，研究人员采用多元分析中的判别分析方法，建立数值判别法则，把由视觉疾病引起多重硬化的病人与没有这种病的人区分开来．

(3)随机抽取 200 名患有抑郁症的病人，按照测量到的指标，可以将他们分成不同的类型，这里可以采用聚类分析方法．如果对阑尾炎患者进行诊断，则可根据病人的多种症状(体温、白血球、恶心、呕吐、腹部压痛感等)，采用判别分析方法判别此人患何种类型阑尾炎．

2. 商业和经济

(1)保险公司希望能够通过收集投保单位的一些金融变量的数据，以便应用多元分析中的判别分析建立一个判别准则，将投保者分为有偿付能力的和“预后不良”的两类，然后采取相应的补救措施，以防止后者破产．

(2)对我国 30 个省市自治区的社会情况进行分析，选取一些能反映社会情况的代表性指标，如：人口密度、城市和农村的平均每人每月收入和支出情况、居住面积、城市绿化率等等．首先根据这些指标将 30 个省市自治区进行分类(聚类分析)，然后根据分类结果对社会情况进行综合评价．

(3)根据人均国民收入、人均工农业产值、人均消费水平等多种指标来判别一个国家的经济发展程度所属的类型．

(4)如何研究国民收入变量与投资性变量之间的相关关系？如何研究企业的经济效益指标与其资金、利税等主要财务指标之间的关系？可用典型相关分析．

(5)如何考察某产品的多个质量指标与影响产品质量的因素之间的关系？如何考察商品的销售量与商品价格、消费者的收入等等之间的关系？可否对今后的销售情况进行预测？可以采用多元回归分析方法．

(6)上市公司的经营业绩是投资者很关心的问题，对股票价格有直接的影响．如何对上市公司的经济效益进行评价，供投资者择股参考呢？有人从上市公司的财务报表中选取了 7 项指

标(每股税后利润、净资产利润率、资金利润率、主营业务利润率、应收账款周转率、税后利润增长率、主营业务收入增长率)做分析. 研究者利用主成分分析确定了反映上市公司赢利水平、发展能力、经营能力的三个综合指标,对上市公司做了排序,供投资者参考.

3. 教育学

(1)如何对高考的成绩做因素分析? 学生入学后的成绩与高考的各门成绩有何关系?

(2)如何对教师的教学活动做出合理的评价?研究人员选取了反映教学态度、教学水平、教学方法和教学效果几个方面的共计 13 项指标,运用主成分分析方法,确定评价教师教学质量的综合指标,并运用聚类分析方法对教师做了分类.

4. 体育科学

林登于 1977 年对奥林匹克运动会十项全能成绩用因子分析方法确定了四个基本体力因子:短跑速度、臂力,长跑耐力、腿力. 在这里,目的是确定观测变量(田径运动成绩)对潜在变量(体力因子)的依赖关系.

5. 生物学和生态学

(1)在植物育种研究中,当植物收获时,我们要选择种子,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. 育种者的目标就是在最短的时间内,使遗传基因达到最大. 可以用因子分析方法和判别分析方法.

(2)一项研究探讨了中国七星瓢虫在黄海、渤海的群聚与近期气象条件的关系,另一项研究对 1 000 个类似的鱼类样本进行分类,根据测量的特征,如体重、身长、鳍数、鳍长、头宽等,将这些鱼分成若干个不同品种.

6. 环境保护

(1)有学者研究了洛杉矶地区大气污染物的浓度. 在一项研究中,在较长的一段时间内,每天测量与污染物有关的七个变量. 空气污染的程度在一周内是否固定不变,或周末与平时是否有显著差异,要知道这些庞杂的测量数据是否能用一种容易解释的方式加以归纳化简. 可以用假设检验和数据化简方法.

(2)为了了解某大型化工厂对环境的污染程度,可在厂区建立多个监测点,每天定时测定各种污染气体的浓度. 可用统计方法将厂区按污染情况分为严重污染、一般污染和轻微污染三类;并为今后建立监测点提供经济合理的方案.

7. 地质学

(1)在一项关于沉积物体积等级分布的研究中,多元分析被用于构造 10 个变量的两个线性函数,以便地质学家区分五种沉积环境,所获得的结果使人们大大减少了为区别不同沉积类型所必须做的一些实验室工作.

(2)在一项关于金矿地球化学特征的研究中,研究人员应用相关分析、因子分析和回归分析方法确定了金的指示性元素,划分了成矿阶段,对未知地区进行了含矿性预测.

8. 心理学

在一项冒险行为的研究中,随机地分配每个学生去接受三种不同训练或“处理”的一种,然后,用同一测验的两种形式去检查他们,对错误反应给予高、低两种处罚分数. 这样的测验得分与带有不同风险的训练方案有关. 我们需要考察的是:在用测验得分度量可察觉的风险方面,不同的训练是否确有差别? 需要利用假设检验方法.

9. 气象学

研究人员为了确定树木年龄与各种气候参数之间的定量关系，首先确定了年轮所包含的气候信息的类型，然后找出公元 1 700 年以来的气候异常现象. 他们应用多元分析方法把海量的数据化简到人们能够处理的范围之内. 又重新定义了几个新变量，并对它们做较为简单的统计分析和解释.

10. 社会学

调查青年对婚姻家庭的态度，如：对文化和职业的要求，对经济收入的态度，对老人的责任，对相貌的重视等等作为主要因素分析，以便进行正确引导和思想教育.

11. 考古学

考古学家根据挖掘出来的人头盖骨的高、宽等特征来判断是男或女，根据挖掘出的动物牙齿的有关测试指标，判别它是属于哪一类动物，是哪一时代的.

12. 军事科学

现代战争中的空情信息复杂多变，如何合理地区分空中目标，以提取其主要信息，是空情信息处理的一个重要方面. 可以基于多元统计的思想，采用聚类分析和判别分析的方法，对空中目标进行区分、辨别，为指挥人员提供作战决策的依据.

13. 文学

关于《红楼梦》一书的作者一直存在争议，一般认为前 80 回为曹雪芹所作，后 40 回为高鹗所续. 对此，能否从数学上做出论证？复旦大学李贤平教授带领他的学生们将 120 回看成是 120 个样本，选定与情节无关的虚词作为变量，并以每一回里这些变量出现的次数作为数据，然后用聚类方法进行分类，果然将 120 回分为两类，即前 80 回为一类，后 40 回为另一类. 之后还论证前 80 回确为曹雪芹所写，后 40 回却不是高鹗一人所写. 这个论证在红学界引起了很大轰动.

§1.2 利用多元统计方法解决实际问题的大致步骤

利用多元统计分析方法解决实际问题时，我们往往要做好以下几个步骤的工作.

1. 提炼具体问题，确定欲达目标

各个领域的实际问题，涉及的变量往往很多，变量之间的关系也复杂多样，由于种种局限使得我们不可能在研究中面面俱到，我们只能集中精力抓主要矛盾. 因此，提炼问题，确定切实可行的目标就成为首要步骤.

2. 根据定性理论，设置指标变量

对于实际的研究问题，确定了目标之后，立即面临了一个具体问题：在该项研究中应涉及哪些指标变量？这要求依据与研究对象有关的定性理论进行分析研究. 对于研究自己不熟悉领域的问题，在选择变量时要与有关专家合作. 另外还要注意，漏掉重要变量肯定会影响研究效果，但也不是引入的变量越多越好. 关于这一点，我们在介绍具体的多元统计分析方法时，将会有深入的解释. 总之，指标变量的确定是一项非常重要的工作，一般不能一次就完全确定，通常要经过反复试算才能选定一些合适的变量.

3. 收集、整理统计数据

统计模型的建立基于指标变量的样本统计数据，样本数据的质量对统计模型的应用效果

有至关重要的影响.

常用的样本数据分为时间序列数据和横截面数据.时间序列数据是按时间顺序排列的统计数据.对于收集到的时间序列数据,我们要注意数据的可比性,如历年的国民收入数据,是不是按可比价格计算的?对于没有可比性的统计数据要进行整理.在应用一些多元分析方法时,要消除时间序列数据的序列相关性.横截面数据是在同一时间截面上的统计数据,采用这种数据时,要注意消除异方差性.

在整理统计数据时,不仅要对一些变量的数据进行折算、差分,有时还要用到把数据对数化、标准化等变换方式,有时还须注意剔除个别异常值,必要时,还要利用插值方法把空缺的数据补齐.

4. 根据目标和数据,选择统计方法,构造理论模型

在收集到所设置变量的数据之后,就要确定适当的数学模型来描述研究对象.在这个步骤中,要将研究对象所遵循的基本原理与数学方法相结合.如做回归分析时,一方面可以画出因变量与自变量的散点图,另一方面也要依据有关理论知识,确定回归模型.有时候,我们无法根据所获得的信息确定模型的形式,这时可以采用不同的形式进行计算机模拟,从不同的模拟结果中,选择较好的一个作为理论模型.

5. 进行统计计算,估计模型参数,求得一些结果

此工作现在基本上都由统计软件来做.

6. 模型的检验与修改

当模型的未知参数估计出来,各种计算也已完成之后,已经初步建立了一个统计模型,但这个统计模型是否真正揭示了所研究问题的实质呢?还有待对模型进行检验.

检验过程包括统计检验和模型实际意义的检验.在没有通过检验的情况下,需要对模型做修改,有时要经过反复修改,方能得到一个理想模型.

7. 统计模型的应用

当一个实际问题的统计模型通过了各种检验之后,我们就可以运用它做进一步的分析研究.需要说明的是,在统计模型的应用中,要注意将定性分析与定量分析有机结合起来.这是因为统计方法只是从事物外在性的数量上去研究问题,这种纯数量的关系是否反映了事物的本质,还有待专门学科的研究才能下定论.

第二章 矩阵代数基础知识

§2.1 矩阵代数的基础知识

2.1.1 矩阵的概念及其运算

定义 2.1 将 $m \times n$ 个实数 $a_{ij}(i=1,2,\cdots,m;j=1,2,\cdots,n)$ 排成一个数阵

$$A=\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}$$

称 A 为 $m \times n$ 矩阵. 常记为 $A=(a_{ij})_{m\times n}$，a_{ij} 表示是 A 的元素，m、n 分别表示 A 的行数和列数.

元素全为零的矩阵称为零矩阵，记为 O.

$m=1$ 的矩阵称为行向量；$n=1$ 的矩阵称为列向量.

$m=n$ 的矩阵称为 n 阶方阵. 元素 $a_{ii}(i=1,2,\cdots,n)$ 称为方阵主对角线上的元素，其他元素称为非对角元素.

主对角线上的元素全为 1，非对角元素全为零的方阵称为单位矩阵，记为 I.

定义 2.2 把矩阵 A 的行元素换成同序数的列元素得到的一个新矩阵，称为 A 的转置矩阵，

记为
$$A^{\mathrm{T}}=\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}^{\mathrm{T}}=\begin{pmatrix} a_{11} & a_{21} & \cdots & a_{m1} \\ a_{12} & a_{22} & \cdots & a_{m2} \\ \vdots & \vdots & & \vdots \\ a_{1n} & a_{2n} & \cdots & a_{mn} \end{pmatrix}$$

$A^{\mathrm{T}}=A$ 的方阵称为对称矩阵. $A^{\mathrm{T}}=-A$ 的方阵称为反对称矩阵.

定义 2.3 两个行数、列数都相等的矩阵称为同型矩阵.

两个同型矩阵 $A=(a_{ij})_{m\times n}$、$B=(b_{ij})_{m\times n}$ 的各个对应元素分别相等，即 $a_{ij}=b_{ij}(i=1,2,\cdots,m;j=1,2,\cdots,n)$，称 A 与 B 相等，记为 $A=B$.

两个同型矩阵 $A=(a_{ij})_{m\times n}$、$B=(b_{ij})_{m\times n}$ 的和、差定义为 $A\pm B=(a_{ij}\pm b_{ij})_{m\times n}$.

数 k 与矩阵 A 的乘积定义为 $kA=(ka_{ij})_{m\times n}$.

矩阵 $A=(a_{ij})_{m\times p}$ 与矩阵 $B=(b_{ij})_{p\times n}$ 的乘积定义为 $AB=(c_{ij})_{m\times n}$，其中 $c_{ij}=\sum\limits_{k=1}^{p}a_{ik}b_{kj}$，$(i=1,2,\cdots,m;j=1,2,\cdots,n)$.

上述转置运算以及另外四种运算符合以下的规律：

$$A+(-1)A=O,(A\pm B)^{\mathrm{T}}=A^{\mathrm{T}}\pm B^{\mathrm{T}},(AB)^{\mathrm{T}}=B^{\mathrm{T}}A^{\mathrm{T}}$$
$$(A^{\mathrm{T}})^{\mathrm{T}}=A,A(BC)=(AB)C,A(B+C)=AB+AC$$

2.1.2　矩阵的微分运算

定义 2.4　设 $X=(x_{ij}(t))_{m\times n}$ 是矩阵，它的各个元素是 t 的函数，下列矩阵是它的微分矩阵：

$$\frac{\mathrm{d}X}{\mathrm{d}t}=\begin{bmatrix} x'_{11}(t) & x'_{12}(t) & \cdots & x'_{1n}(t) \\ x'_{21}(t) & x'_{22}(t) & \cdots & x'_{2n}(t) \\ \vdots & \vdots & & \vdots \\ x'_{m1}(t) & x'_{m2}(t) & \cdots & x'_{mn}(t) \end{bmatrix}=(x'_{ij}(t))_{m\times n}$$

由定义 2.4 可得以下运算公式：

$$\frac{\mathrm{d}(X+Y)}{\mathrm{d}t}=\frac{\mathrm{d}X}{\mathrm{d}t}+\frac{\mathrm{d}Y}{\mathrm{d}t}$$
$$\frac{\mathrm{d}(XY)}{\mathrm{d}t}=\frac{\mathrm{d}X}{\mathrm{d}t}Y+X\frac{\mathrm{d}Y}{\mathrm{d}t}$$

定义 2.5　设多元函数 $f(x_1,x_2,\cdots,x_n)$ 是可微的，记 n 维向量 $\boldsymbol{x}=(x_1,x_2,\cdots,x_n)^{\mathrm{T}}$，称向量

$$\frac{\partial f(\boldsymbol{x})}{\partial \boldsymbol{x}}=(\frac{\partial f}{\partial x_1},\frac{\partial f}{\partial x_2},\cdots,\frac{\partial f}{\partial x_n})^{\mathrm{T}}\text{ 是函数 } f(\boldsymbol{x}) \text{ 的微分向量}.$$

由定义 2.5 有下述微分运算结论.

定理 2.1　设 $\boldsymbol{x}=(x_1,x_2,\cdots,x_n)^{\mathrm{T}}$，已知向量 $\boldsymbol{a}=(a_1,a_2,\cdots,a_n)^{\mathrm{T}}$，则 $\frac{\partial(\boldsymbol{x}^{\mathrm{T}}\boldsymbol{a})}{\partial \boldsymbol{x}}=\boldsymbol{a}$.

证　因 $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{a}=a_1x_1+a_2x_2+\cdots+a_nx_n$，根据定义 5 有

$$\frac{\partial(\boldsymbol{x}^{\mathrm{T}}\boldsymbol{a})}{\partial \boldsymbol{x}}=\left(\frac{\partial(\boldsymbol{x}^{\mathrm{T}}\boldsymbol{a})}{\partial x_1},\frac{\partial(\boldsymbol{x}^{\mathrm{T}}\boldsymbol{a})}{\partial x_2},\cdots,\frac{\partial(\boldsymbol{x}^{\mathrm{T}}\boldsymbol{a})}{\partial x_n}\right)^{\mathrm{T}}$$
$$=(a_1,a_2,\cdots,a_n)^{\mathrm{T}}=\boldsymbol{a}$$

故定理结论成立.□

特别地，当 $\boldsymbol{a}=\boldsymbol{x}$ 时，有 $\frac{\partial(\boldsymbol{x}^{\mathrm{T}}\boldsymbol{x})}{\partial \boldsymbol{x}}=2\boldsymbol{x}$.

因为 $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{x}=x_1^2+x_2^2+\cdots+x_n^2$，所以

$$\frac{\partial(\boldsymbol{x}^{\mathrm{T}}\boldsymbol{x})}{\partial \boldsymbol{x}}=\left(\frac{\partial(\boldsymbol{x}^{\mathrm{T}}\boldsymbol{x})}{\partial x_1},\frac{\partial(\boldsymbol{x}^{\mathrm{T}}\boldsymbol{x})}{\partial x_2},\cdots,\frac{\partial(\boldsymbol{x}^{\mathrm{T}}\boldsymbol{x})}{\partial x_n}\right)^{\mathrm{T}}$$
$$=(2x_1,2x_2,\cdots,2x_n)^{\mathrm{T}}=2\boldsymbol{x}\quad\Box$$

定理 2.2　设 $\boldsymbol{x}=(x_1,x_2,\cdots,x_n)^{\mathrm{T}}$，$\boldsymbol{B}=(b_{ij})_{n\times n}$ 是一对称矩阵，则 $\frac{\partial(\boldsymbol{x}^{\mathrm{T}}\boldsymbol{B}\boldsymbol{x})}{\partial \boldsymbol{x}}=2\boldsymbol{B}\boldsymbol{x}$.

证　注意到 $\boldsymbol{B}$ 是对称矩阵，故

$$\boldsymbol{x}^{\mathrm{T}}\boldsymbol{B}\boldsymbol{x}=\sum_{i=1}^{n}\sum_{j=1}^{n}b_{ij}x_ix_j=x_1(\sum_{j=1}^{n}b_{1j}x_j)+x_2(\sum_{j=1}^{n}b_{2j}x_j)+\cdots+x_n(\sum_{j=1}^{n}b_{nj}x_j)$$
$$=b_{11}x_1^2+b_{22}x_2^2+\cdots+b_{nn}x_n^2+2b_{12}x_1x_2+\cdots+2b_{n-1,n}x_{n-1}x_n$$

所以对每一个变量 $x_i(i=1,2,\cdots,n)$，有

$$\frac{\partial(\boldsymbol{x}^{\mathrm{T}}\boldsymbol{B}\boldsymbol{x})}{\partial x_i}=\sum_{j=1}^{n}b_{ij}x_j+x_1b_{1i}+x_2b_{2i}+\cdots+x_nb_{ni}$$

$$= \sum_{j=1}^{n} x_j b_{ji} + \sum_{j=1}^{n} b_{ij} x_j = 2\sum_{j=1}^{n} b_{ij} x_j \qquad (i = 1,2,\cdots,n)$$

故向量$\dfrac{\partial(\boldsymbol{x}^{\mathrm T}\boldsymbol{B}\boldsymbol{x})}{\partial \boldsymbol{x}} = \left(2\sum_{j=1}^{n} b_{1j}x_j, 2\sum_{j=1}^{n} b_{2j}x_j, \cdots, 2\sum_{j=1}^{n} b_{nj}x_j\right)^{\mathrm T} = 2\boldsymbol{B}\boldsymbol{x}$ □

定理 2.3 设 $\boldsymbol{x} = (x_1, x_2, \cdots, x_n)^{\mathrm T}$ 是 n 维向量，$\boldsymbol{y} = (y_1, y_2, \cdots, y_m)^{\mathrm T}$ 是 m 维向量，$\boldsymbol{B} = (b_{ij})_{n\times m}$ 是 $n \times m$ 矩阵，则$\dfrac{\partial(\boldsymbol{x}^{\mathrm T}\boldsymbol{B}\boldsymbol{x})}{\partial \boldsymbol{x}} = \boldsymbol{B}\boldsymbol{y}$，$\dfrac{\partial(\boldsymbol{x}^{\mathrm T}\boldsymbol{B}\boldsymbol{y})}{\partial \boldsymbol{y}} = \boldsymbol{B}^{\mathrm T}\boldsymbol{x}$.

证 因为 $\boldsymbol{x}^{\mathrm T}\boldsymbol{B}\boldsymbol{y} = x_1(\sum_{j=1}^{m} b_{1j}y_j) + x_2(\sum_{j=1}^{m} b_{2j}y_j) + \cdots + x_n(\sum_{j=1}^{m} b_{nj}y_j)$

$$= y_1(\sum_{i=1}^{n} b_{i1}x_i) + y_2(\sum_{i=1}^{n} b_{i2}x_i) + \cdots + y_m(\sum_{i=1}^{n} b_{im}x_i)$$

所以向量$\dfrac{\partial(\boldsymbol{x}^{\mathrm T}\boldsymbol{B}\boldsymbol{y})}{\partial \boldsymbol{x}} = \left(\dfrac{\partial(\boldsymbol{x}^{\mathrm T}\boldsymbol{B}\boldsymbol{y})}{\partial x_1}, \dfrac{\partial(\boldsymbol{x}^{\mathrm T}\boldsymbol{B}\boldsymbol{y})}{\partial x_2}, \cdots, \dfrac{\partial(\boldsymbol{x}^{\mathrm T}\boldsymbol{B}\boldsymbol{y})}{\partial x_n}\right)^{\mathrm T}$

$$= \left(\sum_{j=1}^{m} b_{1j}y_j, \sum_{j=1}^{m} b_{2j}y_j, \cdots, \sum_{j=1}^{m} b_{nj}y_j\right)^{\mathrm T} = \boldsymbol{B}\boldsymbol{y}$$

$$\frac{\partial(\boldsymbol{x}^{\mathrm T}\boldsymbol{B}\boldsymbol{y})}{\partial \boldsymbol{y}} = \left(\frac{\partial(\boldsymbol{x}^{\mathrm T}\boldsymbol{B}\boldsymbol{y})}{\partial y_1}, \frac{\partial(\boldsymbol{x}^{\mathrm T}\boldsymbol{B}\boldsymbol{y})}{\partial y_2}, \cdots, \frac{\partial(\boldsymbol{x}^{\mathrm T}\boldsymbol{B}\boldsymbol{y})}{\partial y_m}\right)^{\mathrm T}$$

$$= \left(\sum_{i=1}^{n} b_{i1}x_i, \sum_{i=1}^{n} b_{i2}x_i, \cdots, \sum_{i=1}^{n} b_{im}x_i\right)^{\mathrm T} = \boldsymbol{B}^{\mathrm T}\boldsymbol{x} \quad \square$$

2.1.3 方阵的行列式与逆矩阵

定义 2.6 由 n 阶方阵的 $\boldsymbol{A} = (a_{ij})_{n\times n}$ 元素组成的行列式叫做方阵的行列式，记作 $|\boldsymbol{A}|$. 如果 $|\boldsymbol{A}| \neq 0$，称 $\boldsymbol{A}$ 为非奇异阵，否则称为奇异阵.

定义 2.7 对于 n 阶方阵 $\boldsymbol{A}$，如果存在 n 阶方阵 $\boldsymbol{B}$，使得 $\boldsymbol{AB} = \boldsymbol{BA} = \boldsymbol{I}$，则称矩阵 $\boldsymbol{A}$ 是可逆的，并称 $\boldsymbol{B}$ 是 $\boldsymbol{A}$ 的逆矩阵，记作 $\boldsymbol{A}^{-1} = \boldsymbol{B}$.

方阵行列式和逆矩阵的重要性质有：

1° $|k\boldsymbol{A}| = k^n|\boldsymbol{A}|$，其中 $\boldsymbol{A}$ 是 n 阶方阵，k 是常数.

2° $|\boldsymbol{AB}| = |\boldsymbol{A}| \cdot |\boldsymbol{B}|$，其中 $\boldsymbol{A}$、$\boldsymbol{B}$ 均为方阵.

3° $|\boldsymbol{A}^{\mathrm T}| = |\boldsymbol{A}|$，$|\boldsymbol{A}^{-1}| = (|\boldsymbol{A}|)^{-1}$.

4° 如果方阵是可逆的，则逆矩阵是惟一的.

5° 方阵可逆的充分必要条件是其为非奇异阵.

6° $(\boldsymbol{AB})^{-1} = \boldsymbol{B}^{-1}\boldsymbol{A}^{-1}$，$(\boldsymbol{A}^{\mathrm T})^{-1} = (\boldsymbol{A}^{-1})^{\mathrm T}$.

2.1.4 方阵的特征值与特征向量

定义 2.8 对于 n 阶方阵 $\boldsymbol{A}$，如果存在数 λ 和 n 维非零向量 $\boldsymbol{x}$，使得 $\boldsymbol{A}\boldsymbol{x} = \lambda\boldsymbol{x}$，则称数 λ 为方阵 $\boldsymbol{A}$ 的特征值，$\boldsymbol{x}$ 为对应于 λ 的特征向量.

方程组 $\boldsymbol{A}\boldsymbol{x} = \lambda\boldsymbol{x}$，即$(\boldsymbol{A} - \lambda\boldsymbol{I})\boldsymbol{x} = \boldsymbol{0}$ 有非零解 $\boldsymbol{x}$ 的充要条件是系数行列式等于零. 即 $|\boldsymbol{A} - \lambda\boldsymbol{I}| = 0$. 此式称为 $\boldsymbol{A}$ 的特征方程. 等式左端

$$|\boldsymbol{A} - \lambda\boldsymbol{I}| = \begin{vmatrix} a_{11} - \lambda & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} - \lambda & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} - \lambda \end{vmatrix}$$

是一个关于未知量 λ 的 n 次多项式，称为 A 的特征多项式.

特征值和特征向量是两个极为重要的概念，它们在本书所介绍的判别分析、主成分分析、典型相关分析中起重要作用. 下面是一些有关的结论：

1° n 阶方阵 $\boldsymbol{A}$ 有且恰有 n 个特征值. 这些特征值可以是实数，也可以是复数.

2° 设 n 阶方阵 $\boldsymbol{A}=(a_{ij})_{n\times n}$ 的特征值分别为 $\lambda_1,\lambda_2,\cdots,\lambda_n$，则 $\lambda_1+\lambda_2+\cdots+\lambda_n=a_{11}+a_{22}+\cdots+a_{nn}$.

3° n 阶方阵 $\boldsymbol{A}$ 与 $\boldsymbol{A}^{\mathrm{T}}$ 有相同的特征值.

4° 如果 $\boldsymbol{A}$ 是 $m\times n$ 阶矩阵，$\boldsymbol{B}$ 是 $n\times m$ 阶矩阵，则 $\boldsymbol{AB}$ 和 $\boldsymbol{BA}$ 有相同的非零特征值.

5° 实对称矩阵的特征值都是实数. 因此，实对称矩阵的特征值可以按大小次序排成 $\lambda_1\geqslant\lambda_2\geqslant\cdots\geqslant\lambda_n$.

6° 实对称矩阵 $\boldsymbol{A}$ 的不同特征值所对应的特征向量是正交的. 因此，存在正交矩阵 $\boldsymbol{P}$，使得 $\boldsymbol{P}^{-1}\boldsymbol{AP}=\Lambda$，其中 Λ 是以 $\boldsymbol{A}$ 的 n 个特征值为对角元素的对角矩阵.

7° 实对称矩阵 $\boldsymbol{A}$ 为正定矩阵的充分必要条件是 $\boldsymbol{A}$ 的特征值都是正的. 因此，矩阵 $\boldsymbol{A}$ 是非奇异的.

8° 如果 $\boldsymbol{A}$ 是 n 阶正定矩阵，$\boldsymbol{B}$ 是 $n\times m(m\leqslant n)$ 阶矩阵，且 $R(\boldsymbol{B})=m$，则 $\boldsymbol{B}^{\mathrm{T}}\boldsymbol{AB}$ 是正定阵.

2.1.5　矩阵的秩、迹与正交矩阵

定义 2.9　矩阵 $\boldsymbol{A}$ 的行(或列)向量组的秩称为这个矩阵的秩，记作 $R(\boldsymbol{A})$.

矩阵 $\boldsymbol{A}$ 的秩可反映该矩阵的很多特性，它有以下主要性质：

1° $R(\boldsymbol{A})=R(\boldsymbol{A}^{\mathrm{T}})$.

2° $R(\boldsymbol{AB})\leqslant\min\{R(\boldsymbol{A}),R(\boldsymbol{B})\}$.

3° 如果方阵 $\boldsymbol{B}$、$\boldsymbol{C}$ 都是非奇异的，则 $R(\boldsymbol{BAC})=R(\boldsymbol{A})$.

4° n 阶方阵 $\boldsymbol{A}$ 为非奇异阵的充分必要条件是 $R(\boldsymbol{A})=n$，并称 $\boldsymbol{A}$ 是满秩的.

定义 2.10　n 阶方阵 $\boldsymbol{A}$ 主对角线上的元素之和称为这个矩阵的迹，记为 $\mathrm{tr}\boldsymbol{A}=\sum_{i=1}^{n}a_{ii}$.

任意方阵的迹具有下列性质：

1° $\mathrm{tr}\boldsymbol{A}=\mathrm{tr}\boldsymbol{A}^{\mathrm{T}}$.

2° $\mathrm{tr}(\boldsymbol{A}+\boldsymbol{B})=\mathrm{tr}\boldsymbol{A}+\mathrm{tr}\boldsymbol{B}$.

3° $\mathrm{tr}(\boldsymbol{AB})=\mathrm{tr}(\boldsymbol{BA})$.

4° $\mathrm{tr}\boldsymbol{A}=\lambda_1+\lambda_2+\cdots+\lambda_n$，其中 $\lambda_i(i=1,2,\cdots,n)$ 是矩阵 $\boldsymbol{A}$ 的特征值.

定义 2.11　对于 n 阶方阵 $\boldsymbol{A}$，如果 $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}=\boldsymbol{AA}^{\mathrm{T}}=\boldsymbol{I}$，则称 $\boldsymbol{A}$ 为正交矩阵.

正交矩阵有下列性质：

1° 若 $\boldsymbol{A}$ 是正交矩阵，则 $\boldsymbol{A}^{-1}=\boldsymbol{A}^{\mathrm{T}}$.

2° 正交矩阵的逆矩阵仍然是正交矩阵.

3° 两个正交矩阵的积仍然是正交矩阵.

4° 设 $\boldsymbol{A}$ 是正交矩阵，$\boldsymbol{B}$ 与 $\boldsymbol{A}$ 是同阶方阵，则 $\mathrm{tr}(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{BA})=\mathrm{tr}(\boldsymbol{B})$.

2.1.6　二次型与正定阵

定义 2.12　含有 n 个变量 $x_1,x_2,\cdots,x_n$ 的二次齐次函数

$$f(x_1,x_2,\cdots,x_n)=\sum_{i=1}^{n}\sum_{j=1}^{n}a_{ij}x_ix_j=\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}$$

称为二次型，式中 $\boldsymbol{x}=(x_1,x_2,\cdots,x_n)^{\mathrm{T}}$，$\boldsymbol{A}=(a_{ij})_{n\times n}$ 是一个对称矩阵. 如果 a_{ij} 是复数，称 $f(\boldsymbol{x})$ 为复二次型；如果 a_{ij} 是实数，称 $f(\boldsymbol{x})$ 为实二次型.

如果存在可逆的线性变换 $\boldsymbol{x}=\boldsymbol{C}\boldsymbol{y}$，使得二次型只含平方项，即形如 $f=k_1y_1^2+k_2y_2^2+\cdots+k_ny_n^2$，称为二次型的标准形.

定义 2.13　设有实二次型 $f(\boldsymbol{x})=\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}$，如果对一切非零 n 维列向量 $\boldsymbol{x}$，有 $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}>0$，则称实二次型 $f(\boldsymbol{x})=\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}$ 是正定的，相应的对称矩阵 $\boldsymbol{A}$ 是正定的；如果对一切非零 n 维列向量 $\boldsymbol{x}$，有 $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}<0$，则称实二次型 $f(\boldsymbol{x})=\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}$ 是负定的，相应的对称矩阵 $\boldsymbol{A}$ 是负定的.

判断二次型 $f(\boldsymbol{x})=\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}$ 是否正定的主要方法有：

1° 实二次型 $f(\boldsymbol{x})=\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}$ 为正定的充分必要条件是：它的标准形中 n 个系数皆正.

2° 对称矩阵 $\boldsymbol{A}$ 为正定的充分必要条件是：$\boldsymbol{A}$ 的特征值皆正.

3° 对称矩阵 $\boldsymbol{A}$ 为正定的充分必要条件是：$\boldsymbol{A}$ 的各阶顺序主子式皆正，即

$$a_{11}>0,\begin{vmatrix}a_{11}&a_{12}\\a_{21}&a_{22}\end{vmatrix}>0,\cdots,\begin{vmatrix}a_{11}&\cdots&a_{1n}\\\vdots&&\vdots\\a_{n1}&\cdots&a_{nn}\end{vmatrix}>0$$

4° 对称矩阵 $\boldsymbol{A}$ 为负定的充分必要条件是：奇数阶主子式为负，而偶数阶主子式为正，即

$$(-1)^k\begin{vmatrix}a_{11}&\cdots&a_{1k}\\\vdots&&\vdots\\a_{k1}&\cdots&a_{kk}\end{vmatrix}>0,\qquad k=1,2,\cdots,n$$

§ 2.2　线性方程组

2.2.1　线性方程组解的结构

设非齐次方程组 $\boldsymbol{A}\boldsymbol{x}=\boldsymbol{b}$　　(2.1)

其中　$\boldsymbol{A}=\begin{pmatrix}a_{11}&a_{12}&\cdots&a_{1n}\\a_{21}&a_{22}&\cdots&a_{2n}\\\vdots&\vdots&&\vdots\\a_{m1}&a_{m2}&\cdots&a_{mn}\end{pmatrix}$，$\boldsymbol{x}=(x_1,x_2,\cdots,x_n)^{\mathrm{T}}$，$\boldsymbol{b}=(b_1,b_2,\cdots,b_m)^{\mathrm{T}}$.

记增广系数矩阵 $\boldsymbol{B}=\begin{pmatrix}a_{11}&a_{12}&\cdots&a_{1n}&b_1\\a_{21}&a_{22}&\cdots&a_{2n}&b_2\\\vdots&\vdots&&\vdots&\vdots\\a_{m1}&a_{m2}&\cdots&a_{mn}&b_m\end{pmatrix}$，与方程组(2.1)相应的齐次方程组为

$$\boldsymbol{A}\boldsymbol{x}=\boldsymbol{0}\qquad(2.2)$$

关于方程组解的几个重要结论是：

1° 方程组(2.1)有解的充要条件是 $R(\boldsymbol{A})=R(\boldsymbol{B})$.

2° 当 $m=n$ 时，方程组(2.1)有惟一解的充要条件是 $\boldsymbol{A}$ 为非奇异阵. 这时，解为 $\boldsymbol{x}=\boldsymbol{A}^{-1}\boldsymbol{b}$.

3° 当 $m\neq n$ 时，如果 $R(\boldsymbol{A})=r<n$，则方程组(2.2)有无穷多组解，且解向量的全体构成 $n-r$ 维向量空间.

4° 方程组(2.2)的通解加上方程组(2.1)的一个特解，构成方程组(2.1)的通解.

在多元分析中，经常要进行的运算有：求矩阵的行列式或逆；解线性方程组；做某种递推运算等.下面介绍主要的计算方法.

2.2.2 高斯消去法

设 $m=n$ 时，非齐次线性方程组(2.1)中 $\boldsymbol{A}$ 为非奇异阵.为了书写方便，我们记 $\boldsymbol{A}=(a_{ij}^{(0)})_{n\times n}$，列向量 $\boldsymbol{b}$ 的元素记为 $\boldsymbol{b}=(a_{1,n+1}^{(0)},a_{2,n+1}^{(0)},\cdots,a_{n,n+1}^{(0)})^{\mathrm{T}}$.

假定 $a_{11}^{(0)}\neq 0$(否则可适当调换方程的次序，使得第一个方程的第一个系数非零)，把第一个方程的 $-\dfrac{a_{21}^{(0)}}{a_{11}^{(0)}}$ 倍加到第二个方程上，消去第二个方程中的 x_1；类似地，把第一个方程的 $-\dfrac{a_{i1}^{(0)}}{a_{11}^{(0)}}$ 倍加到第 i 个方程上，消去第 i 个方程中的 $x_1(i=3,4,\cdots,n)$，得到

$$\begin{cases} a_{11}^{(0)}x_1+a_{12}^{(0)}x_2+a_{13}^{(0)}x_3+\cdots+a_{1n}^{(0)}x_n=a_{1,n+1}^{(0)} \\ \qquad a_{22}^{(1)}x_2+a_{23}^{(1)}x_3+\cdots+a_{2n}^{(1)}x_n=a_{2,n+1}^{(1)} \\ \qquad \cdots\cdots\cdots\cdots\cdots\cdots \\ \qquad a_{n2}^{(1)}x_2+a_{n3}^{(1)}x_3+\cdots+a_{nn}^{(1)}x_n=a_{n,n+1}^{(1)} \end{cases} \tag{2.3}$$

再假定 $a_{22}^{(1)}\neq 0$(否则可适当调换后 $n-1$ 个方程的次序，使得第二个方程的第一个系数非零)，用上述同样的步骤在后 $n-2$ 个方程中消去 x_2，得到

$$\begin{cases} a_{11}^{(0)}x_1+a_{12}^{(0)}x_2+a_{13}^{(0)}x_3+\cdots+a_{1n}^{(0)}x_n=a_{1,n+1}^{(0)} \\ \qquad a_{22}^{(1)}x_2+a_{23}^{(1)}x_3+\cdots+a_{2n}^{(1)}x_n=a_{2,n+1}^{(1)} \\ \qquad\qquad a_{33}^{(2)}x_3+\cdots+a_{3n}^{(2)}x_n=a_{3,n+1}^{(2)} \\ \qquad\qquad \cdots\cdots\cdots\cdots\cdots\cdots \\ \qquad\qquad a_{n3}^{(2)}x_3+\cdots+a_{nn}^{(2)}x_n=a_{n,n+1}^{(2)} \end{cases} \tag{2.4}$$

其中

$$a_{ij}^{(2)}=\begin{cases} a_{ij}^{(1)}-a_{i2}^{(1)}a_{2j}^{(1)}/a_{22}^{(1)}, & i=3,4,\cdots,n;\quad j=3,4,\cdots,n,n+1 \\ 0, & i=3,4,\cdots,n;\quad j=1,2 \end{cases} \tag{2.5}$$

按照上述步骤进行 $n-1$ 次，得到方程组

$$\begin{cases} a_{11}^{(0)}x_1+a_{12}^{(0)}x_2+a_{13}^{(0)}x_3+\cdots+a_{1n}^{(0)}x_n=a_{1,n+1}^{(0)} \\ \qquad a_{22}^{(1)}x_2+a_{23}^{(1)}x_3+\cdots+a_{2n}^{(1)}x_n=a_{2,n+1}^{(1)} \\ \qquad\qquad a_{33}^{(2)}x_3+\cdots+a_{3n}^{(2)}x_n=a_{3,n+1}^{(2)} \\ \qquad\qquad \cdots\cdots\cdots\cdots\cdots\cdots \\ \qquad\qquad\qquad a_{nn}^{(n-1)}x_n=a_{n,n+1}^{(n-1)} \end{cases} \tag{2.6}$$

从式(2.6)的第 n 个方程解出 $x_n=a_{n,n+1}^{(n-1)}/a_{nn}^{(n-1)}$，代回到第 $n-1$ 个方程，可解出 x_{n-1}，类似地，依次回代直到从第一个方程中解出 x_1，从而得到方程组(2.6)，也就是方程组(2.1)的一个解向量.这就是高斯消去法.

此方法可用对增广矩阵 $\boldsymbol{B}$ 作线性变换表示成为