多元统计分析

吴密霞　刘春玲　编著

科学出版社

多元统计分析

吴密霞　刘春玲　编著

科　学　出　版　社

北　京

内 容 简 介

本书系统介绍了多元统计分析的基本理论、方法和相关最新发展. 全书共分11 章. 第1章主要介绍了多元分析的发展及其主要研究内容. 第2章讨论矩阵论方面的补充知识和多元正态分布的相关重要定理. 第3章介绍了多元正态随机矩阵的几种重要分布. 第4章和第5章介绍了多元正态总体的参数估计和检验问题以及缺失数据情形的相关研究. 第6章介绍了几类多元线性模型的估计和检验问题. 第 7 章至第 11 章依次介绍了主成分分析、因子分析、典型相关分析、判别分析，以及聚类分析的基本思想和应用.

本书可作为高等学校数学、数理统计或统计、生物统计、计量经济等相关专业的高年级本科生、研究生的选修课教材，也可作为数学、生物、医学、金融、经济等领域的教师或科技工作者的参考书.

图书在版编目(CIP)数据

多元统计分析/吴密霞，刘春玲编著. —北京：科学出版社，2014
ISBN 978-7-03-040496-1

Ⅰ. ①多… Ⅱ. ①吴… ②刘… 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(2014)第 080929 号

责任编辑：李 欣／责任校对：朱光兰
责任印制：赵德静／封面设计：陈 敬

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

中国科学院印刷厂 印刷

科学出版社发行 各地新华书店经销
*
2014年6月第 一 版 开本：720×1000 1/16
2014年6月第一次印刷 印张：13 1/4
字数：267 000

定价：58.00 元

(如有印装质量问题，我社负责调换)

前　　言

多元统计分析是数理统计学中一个重要分支, 在生物、医学、经济、金融、环境科学、抽样调查及工程技术领域得到愈来愈广泛应用. 因此, 多元统计分析已被列入国内外多数高等院校的数理统计、生物统计、计量经济等专业的高年级本科生、研究生所学习或研究的内容之一.

目前, 关于多元统计分析的优秀专著和教材有许多, 如 Anderson(1957,1984), Mardia, Kshirsagar(1972), Gnanadesikan(1977,1997), Kent, Bibby(1979), Muirhead (1982), 王学仁, 王松桂 (1990), 张尧庭, 方开泰 (2003), Morrison(2005), Johnson (2005), Johnson,Wichern(2007), 王静龙 (2008), 何晓群 (2008), Hardle, Simar(2012) 等. 每本书各有特点. 经过多年学习和仔细研读之后, 本书作者受益匪浅, 结合 5 年连续讲授"多元统计分析"研究生课程的心得, 编写了本书.

本书系统介绍了多元统计分析的基本理论、方法及其相关最新发展. 全书共分 11 章. 第 1 章主要介绍了多元分析的发展和主要研究内容. 第 2 章讨论矩阵论方面的补充知识和多元正态分布的相关重要定理. 第 3 章介绍了多元正态随机矩阵的几种重要分布. 第 4 章介绍了多元正态总体的均值向量和协方差阵的极大似然估计, 以及缺失情形的相关研究成果. 第 5 章分别介绍了单总体、多总体下多元正态分布的均值向量、协方方差阵检验问题. 第 6 章介绍了几类多元线性模型的估计和检验问题. 第 7 章至第 11 章依次介绍了主成分分析、因子分析、典型相关分析、判别分析, 以及聚类分析的基本思想, 并给出实例分析.

借本书出版之际, 特别感谢第一作者的导师王松桂教授, 感谢他多年来在科研上给予的指导、鼓励, 此书的编写也得益于研究生期间王松桂教授讲授的"多元统计分析"课, 不但使作者对多元分析产生了浓厚的兴趣, 同时打下了扎实的基础, 尤其是本书第 3 章的许多定理的证明就采用了当时王松桂教授讲授课上所用的方法. 同时也特别感谢第二作者的两位导师: 中国科学院数学与系统科研究院的陈希孺院士和王启华研究员, 感谢两位作者博士后期间的两位指导老师: 美国国家健康研究院 Kai-Fun Yu 研究员和 Aiyi Liu 研究员, 感谢北京工业大学的杨振海教授、张忠占教授、薛留根教授、李寿梅教授、程维虎教授、陈立萍副教授等各位老师多年来给予的大力支持和帮助. 感谢北京工业大学统计专业 2012 级研究生赵延同学和李晓红同学、2013 级研究生刘青同学, 他们提供了书中的部分数据分析. 在此也特别感谢第一作者的女儿孙铭岳, 感谢她在本书的编著阶段给予的支持和配合.

另外, 本书的写作和出版得到了国家自然科学基金 (11171011)、北京市自然科

学基金 (1132007)、北京市教委科技项目 (km201410005011)、北京市人才强教计划"青年骨干教师培养计划"项目 (PHR20110820)、北京工业大学京华人才培养计划项目和校基础基金的资助. 作者愿借此机会表示诚挚的谢意.

尽管该手稿已经作为北京工业大学统计专业研究生的"多元统计分析"课程讲义连续讲授四次, 期间不断修改和调整多次, 但由于编者水平所限, 书中错误或不当之处在所难免, 恳请国内同行及广大读者不吝赐教.

北京工业大学　吴密霞
香港理工大学　刘春玲
2014 年 2 月 28 日

符　号　表

$\triangleq$	“定义为”或“记为”
$\overset{d}{=}$	“同分布于”
$\boldsymbol{A} \geqslant 0$	$\boldsymbol{A}$为对称半正定方阵
$\boldsymbol{A} > 0$	$\boldsymbol{A}$为对称正定方阵
$\mathrm{rk}(\boldsymbol{A})$	矩阵$\boldsymbol{A}$的秩
$\mid \boldsymbol{A} \mid$	矩阵$\boldsymbol{A}$的行列式
$\mathrm{tr}(\boldsymbol{A})$	方阵 $\boldsymbol{A}$ 的迹
$\lambda_i(\boldsymbol{A})$	$\boldsymbol{A}$的第 i 个顺序特征根
$\mathcal{M}(\boldsymbol{A})$	矩阵$\boldsymbol{A}$的列向量张成的子空间
$\mathbf{1}_n' = (1, \cdots, 1)$	元素皆为 1 的 $n \times 1$ 列向量
$\boldsymbol{J}_n = \mathbf{1}_n \mathbf{1}_n'$	元素皆为 1 的 $n \times n$ 矩阵
$\mathrm{Vec}(\boldsymbol{A})$	将 $\boldsymbol{A}$ 的列向量依次排成的列向量
$\mathrm{diag}(a_1, \cdots, a_n)$	对角元素分别为 $a_1, \cdots, a_n$ 的对角矩阵
$\boldsymbol{A} \otimes \boldsymbol{B}$	矩阵 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 的 Kronecher 乘积
$E(\boldsymbol{X})$	随机变量或向量 $\boldsymbol{X}$ 的均值
$\mathrm{Var}(X)$	随机变量 X 的方差
$\mathrm{Cov}(X, Y)$	随机变量或向量 X, Y 的协方差
$\boldsymbol{u} \sim N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$	均值为 $\boldsymbol{\mu}$, 协方差阵为 $\boldsymbol{\Sigma}$ 的 p 维正态向量
$W_p(n, \boldsymbol{\Sigma})$	自由度为 n 的 Wishart 分布
$T_{p,n}^2$	Hotelling-T^2 统计量
MSE	均方误差
MVUE	最小方差无偏估计
LRT	似然比检验
UIT	并交检验
$\det(\boldsymbol{A})$	矩阵 $\boldsymbol{A}$ 的行列式

目　　录

第 1 章　概　　述

1.1　一元统计分析的局限性

对待同时观察的多个变量 $X=(X_1,X_2,\cdots,X_p)$, 通常有如下两种分析方法来处理多变量的观测数据. 一种方法就是把多个变量单独进行研究 —— 一元统计分析. 这主要针对这些变量之间相互独立的情形, 但若这些变量之间相关, 这种方法就只考虑了它们各自的边缘分布特征, 丢失变量之间的相关信息, 故使得分析结果往往低效或无效. 因此, 在变量相关情形, 另一种方法就是把这多个变量一起研究 —— 多元统计分析. 此时, $X=(X_1,X_2,\cdots,X_p)$ 被看作 p 维空间的随机点. 下面用一个简单的例子来揭示一元统计分析方法的局限性.

例 1.1.1 (王静龙, 2008)　英国著名统计学家 K.Pearson 曾进行了一项研究, 研究家庭成员间的相似性, 其中一项是研究父子身高, 分别记为 X,Y. 他测量了 1078 个父亲及其成年儿子的身高, 经计算, 父亲的平均身高为 172.7cm, 标准差 6.86cm, 儿子的平均身高为 175.3cm, 标准差 6.86cm, 它们之间的相关系数为 0.5. 本例的目标是得到一个区域 $\mathcal{D}$, 使得

$$P((X,Y)\in\mathcal{D})=95\% \ . \tag{1.1.1}$$

一般情形, 人的身高和体重的生理测量值都服从正态分布, 故假设父亲和儿子的身高的联合分布为正态分布. 由于样本量很大, 不妨令总体均值、标准差和相关系数分别等于样本均值、标准差和相关系数, 即 $\mu_1=172.7\text{cm}$, $\mu_2=175.3\text{cm}$, $\sigma=6.86\text{cm}$, $\rho=0.5$. 于是有 $X\sim N(172.7,6.86^2)$, $Y\sim N(175.3,6.86^2)$.

依据一元统计分析方法, 忽视相关系数, 即认为 X 和 Y 独立, 故

$$P((X,Y)\in\mathcal{D})=P((X)\in\mathcal{D}_1)\times P((Y)\in\mathcal{D}_2)=95\%,$$

这里, $\mathcal{D}=\mathcal{D}_1\times\mathcal{D}_2$. 由于 X 和 Y 变量对等, 故令

$$P(X\in\mathcal{D}_1)=P(Y\in\mathcal{D}_2)=\sqrt{95\%}\ =97.47\%.$$

于是得

$$\begin{cases}\mathcal{D}_1=\left(172.7-6.86U_{1-0.9747/2},172.7+6.86U_{1-0.9747/2}\right)=(159.29,186.11),\\ \mathcal{D}_2=\left(175.3-6.86U_{1-0.9747/2},175.3+6.86U_{1-0.9747/2}\right)=(161.89,188.71).\end{cases} \tag{1.1.2}$$

因此, $\mathcal{D}$ 为正方形. 但我们知道家庭成员中父亲的身高与其儿子的身高正相关, 所以(1.1.2)成立, 并不能保证(1.1.1)成立. 这是一个很明显的不足, 因为父亲高的, 儿子往往也比较高; 同样父亲矮的, 儿子也常常会矮. 因此, 正方形左上角和右下角发生的可能性很小, 没必要保留. 故区域 $\mathcal{D}$ 取椭圆更为恰当 (图 1.1).

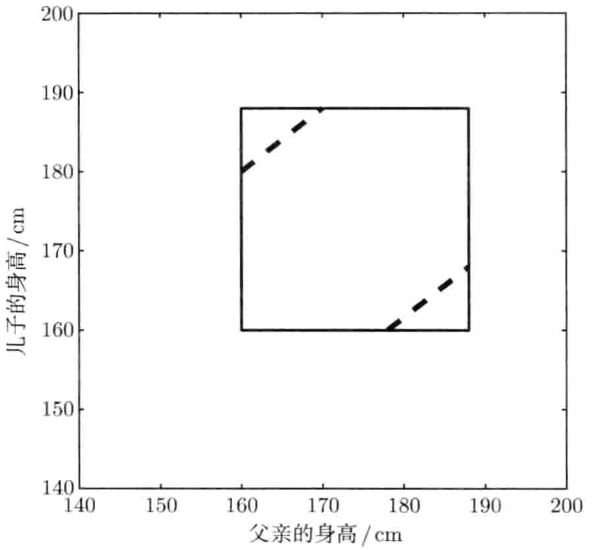

图 1.1

下面采用多元方法, 将父亲的身高和儿子的身高一起研究, 从而构建椭圆区域 $\mathcal{D}$. 设 $Z=(X,Y)\sim N_2(\mu_1,\mu_2,\sigma_1^2,\sigma_2^2,\rho)$, 其密度函数为

$$f(x,y)=\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}\exp\left\{-\frac{1}{2(1-\rho^2)}\left[\frac{(x-\mu_1)^2}{\sigma_1^2}-\frac{2\rho(x-\mu_1)(y-\mu_2)}{\sigma_1\sigma_2}+\frac{(y-\mu_2)^2}{\sigma_2^2}\right]\right\}.$$

如同一元正态分布, 二元正态分布的密度函数曲面也是单峰对称的. 注意到二元正态分布的密度函数中, 当 $\rho>0$ 时, (X,Y) 的值集中在方向自左向右向上 (斜率大于 0) 的直线附近; 当 $\rho<0$ 时, (X,Y) 的值集中在方向自左向右向上 (斜率小于 0) 的直线附近; 当 $\rho=0$ 时, (X,Y) 的值集中在方向自平行于其中一条坐标轴 (斜角等于 0° 或者 90°) 的直线附近. 因此取使得(1.1.1)成立的区域 $\mathcal{D}$ 的边界为密度等高椭圆曲线比较好, 即

$$\mathcal{D}=\left\{(x,y):\frac{(x-\mu_1)^2}{\sigma_1^2}-\frac{2\rho(x-\mu_1)(y-\mu_2)}{\sigma_1\sigma_2}+\frac{(y-\mu_2)^2}{\sigma_2^2}\leqslant c\right\},$$

见图 1.2, 其中 c 满足条件(1.1.1). 由于 $\sigma_1=\sigma_2=\sigma=6.86$, $\rho=0.5$, 故上区域可被简写成

$$D^*=\left\{(x,y):(x-172.7)^2-(x-172.7)(y-175.3)+(y-175.3)^2\leqslant c^*\right\},$$

其中 $c^*=c\sigma^2=\left(2(1-\rho^2)\ln 20\right)6.86^2=211.47$.

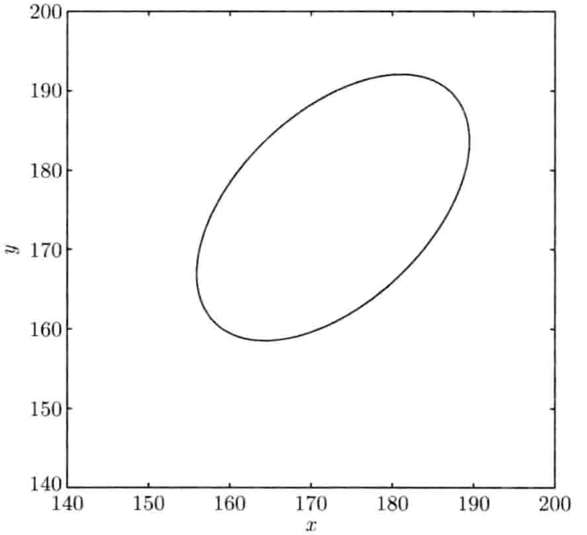

图 1.2

事实上, 记 $\boldsymbol{z}=(x,y)'$, $\boldsymbol{\mu}=(\mu_1,\mu_2)'$,

$$\boldsymbol{\Sigma}=\sigma^2\begin{pmatrix}1 & \rho\\ \rho & 1\end{pmatrix}.$$

$\boldsymbol{Z}=(X,Y)$ 的密度函数可表达为矩阵形式

$$f(x,y)=\frac{1}{2\pi\sqrt{|\boldsymbol{\Sigma}|}}\exp\left\{-\frac{1}{2}(\boldsymbol{z}-\boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\boldsymbol{z}-\boldsymbol{\mu})\right\}. \tag{1.1.3}$$

令 $\boldsymbol{w}=\boldsymbol{\Sigma}^{-1/2}(\boldsymbol{z}-\boldsymbol{\mu})$, $R^2=c/(1-\rho^2)$. 于是

$$\mathcal{D}=\left\{(\boldsymbol{z}-\boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\boldsymbol{z}-\boldsymbol{\mu})=\boldsymbol{w}'\boldsymbol{w}\leqslant R^2\right\},$$

利用极坐标变换, 令 $w_1=r\cos\theta$, $w_2=r\sin\theta$, 则

$$\begin{aligned}\iint_D f(x,y)dxdy&=\frac{1}{2\pi}\int\int_{\boldsymbol{w}'\boldsymbol{w}\leqslant R^2}\exp\left(-\frac{\boldsymbol{w}'\boldsymbol{w}}{2}\right)dw_1dw_2\\&=\frac{1}{2\pi}\int_0^{2\pi}d\theta\int_0^R\exp\left(-\frac{r^2}{2}\right)rdr\end{aligned}$$

$$= (1 - e^{-R^2/2}) = 0.95.$$

因此, $R^2 = 2\ln 20$. 于是

$$c = (1 - \rho^2)R^2 = 2(1 - \rho^2)\ln 20.$$

结合事实 $c^* = c\sigma^2$, 便可求得 $c^* = 211.47$. 另外, 可以证明所求得 $\mathcal{D}^*$ 是所有满足条件(1.1.1)的 $\mathcal{D}$ 中面积最小的.

1.2 多元分析的目标

多元分析研究的对象是多个变量同时测量的数据, 研究的目标是应用多元分析方法和技术实现以下几个目标 (Johnson, Wichern, 2007):

- 数据降维和结构简化. 为了更容易解释所研究的对象, 就需要在不损失主要信息的情况下尽可能地将结构简单化. 如将太空扫描仪收集到的多光谱图像数据转换成二维可视图片或影像.
- 分组和归类. 根据某些测量特征, 将相似的个体或变量合理分组. 另外, 建立某些判别准则, 可将新个体归到相应的类中. 如保险公司推出一个新产品, 需要通过市场调研数据寻找该产品潜在的消费人群. 制定一个判别准则, 来判断个体是否酗酒. 这需要借助的多元分析的方法分别为聚类分析和判别分析.
- 调查变量间的依赖关系. 多个变量间的关系自然是研究者感兴趣的问题, 如果变量间相互独立, 这时多元分析问题就可降为一元统计分析. 若不独立, 则需要知道某个或某些变量是如何依赖其他变量的. 如某种病与哪些因素有关, 最理想的是能建立它们间的关系模型.
- 预测. 多元变量间关系研究的一个主要目的是预测, 即基于其他变量的观测值来预测感兴趣的某个或某些变量的取值. 如从会计和财务的几个数据来识别出潜在的资不抵债的保险公司.
- 假设检验. 应用中往往需要对多元总体的参数的具体假设进行检验, 从统计角度来验证该假设合理与否. 如通过测量大城市中环境污染的几个变量, 来判断该地区污染一周内大致保持不变, 还是工作日与周末显著不同.

解决以上问题, 主要涉及的多元统计分析基本方法有总体参数的统计推断、主成分分析、因子分析、典型相关分析、判别分析和聚类分析. 在应用中这些方法并非单独使用, 往往需要相互配合, 最终得到有效的分析结果. 由于多元统计分析处理的是多维数据, 故其计算上远比一元分析要复杂得多, 难度大得多. 然而, 目前的许多统计软件已经有相当成熟的相应软件包, 这使得多元分析各方法的实施更加容易.

目前, 多元分析中的大部分理论和方法都是基于正态分布假设下得到的, 故矩阵运算和正态分布的性质对真正掌握多元分析的基本方法起着重要的作用. 因此,

本书在介绍以上六种方法前, 分别在第 2 章和第 3 章介绍了多元分析中涉及的矩阵知识、多元正态分布性质以及几种重要的多元分布.

第 2 章　矩阵与正态向量

2.1　矩阵运算知识

2.1.1　Kronecker 乘积与向量化运算

本节我们要介绍矩阵的两种特殊运算: Kronecker 乘积与向量化运算, 它们在线性模型、多元统计分析等分支的参数估计理论中有特别重要的应用.

定义 2.1.1　设 $\boldsymbol{A}=(a_{ij})$ 和 $\boldsymbol{B}=(b_{ij})$ 分别为 $m\times n$, $p\times q$ 的矩阵, 定义矩阵 $\boldsymbol{C}=(a_{ij}\boldsymbol{B})$, 这是一个 $mp\times nq$ 的矩阵, 称为 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 的Kronecker乘积, 记为 $\boldsymbol{C}=\boldsymbol{A}\otimes\boldsymbol{B}$, 即

$$\boldsymbol{A}\otimes\boldsymbol{B}=\begin{pmatrix} a_{11}\boldsymbol{B} & a_{12}\boldsymbol{B} & \cdots & a_{1n}\boldsymbol{B} \\ a_{21}\boldsymbol{B} & a_{22}\boldsymbol{B} & \cdots & a_{2n}\boldsymbol{B} \\ \vdots & \vdots & & \vdots \\ a_{m1}\boldsymbol{B} & a_{m2}\boldsymbol{B} & \cdots & a_{mn}\boldsymbol{B} \end{pmatrix}. \tag{2.1.1}$$

Kronecker 乘积具有下列性质:

(a) $\boldsymbol{0}\otimes\boldsymbol{A}=\boldsymbol{A}\otimes\boldsymbol{0}=\boldsymbol{0}$,

(b) $(\boldsymbol{A}_1+\boldsymbol{A}_2)\otimes\boldsymbol{B}=(\boldsymbol{A}_1\otimes\boldsymbol{B})+(\boldsymbol{A}_2\otimes\boldsymbol{B})$, $A\otimes(\boldsymbol{B}_1+\boldsymbol{B}_2)=(\boldsymbol{A}\otimes\boldsymbol{B}_1)+(\boldsymbol{A}\otimes\boldsymbol{B}_2)$,

(c) $(\alpha\boldsymbol{A})\otimes(\beta\boldsymbol{B})=\alpha\beta(\boldsymbol{A}\otimes\boldsymbol{B})$,

(d) $(\boldsymbol{A}_1\otimes\boldsymbol{B}_1)(\boldsymbol{A}_2\otimes\boldsymbol{B}_2)=(\boldsymbol{A}_1\boldsymbol{A}_2)\otimes(\boldsymbol{B}_1\boldsymbol{B}_2)$,

(e) $(\boldsymbol{A}\otimes\boldsymbol{B})'=\boldsymbol{A}'\otimes\boldsymbol{B}'$,

(f) $(\boldsymbol{A}\otimes\boldsymbol{B})^{-1}=\boldsymbol{A}^{-1}\otimes\boldsymbol{B}^{-1}$.

定义 2.1.2　设随机阵 $\boldsymbol{X}=(\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_n)$ 是一个 $p\times n$ 矩阵, 把矩阵 $\boldsymbol{X}$ 按列向量 $\boldsymbol{x}_1,\cdots,\boldsymbol{x}_n$ 依次排成一个 $pn\times 1$ 的向量

$$\mathrm{Vec}(\boldsymbol{X})=\begin{pmatrix} \boldsymbol{x}_1 \\ \vdots \\ \boldsymbol{x}_n \end{pmatrix},$$

则称 $\mathrm{Vec}(\boldsymbol{X})$ 为矩阵 $\boldsymbol{X}$ 的向量化运算, 称 Vec( ) 为矩阵的向量化符号.

向量化运算具有下列性质:

(a) $\mathrm{Vec}(\boldsymbol{A}+\boldsymbol{B})=\mathrm{Vec}(\boldsymbol{A})+\mathrm{Vec}(\boldsymbol{B})$,

(b) $\mathrm{Vec}(\alpha\boldsymbol{A})=\alpha\mathrm{Vec}(\boldsymbol{A})$, 这里 α 为数,

(c) $\mathrm{tr}(\boldsymbol{AB})=(\mathrm{Vec}(\boldsymbol{A}'))'\mathrm{Vec}(\boldsymbol{B})$,

(d) 设 $\boldsymbol{a}$ 和 $\boldsymbol{b}$ 分别为 $n\times 1$ 和 $m\times 1$ 向量, 则 $\mathrm{Vec}(\boldsymbol{ab}')=\boldsymbol{b}\otimes\boldsymbol{a}$,

(e) $\mathrm{Vec}(\boldsymbol{ABC})=(\boldsymbol{C}'\otimes\boldsymbol{A})\mathrm{Vec}(\boldsymbol{B})$.

下面只证明性质 (e), 其余留作练习.

设 $\boldsymbol{C}_{m\times n}=(c_{ij})=(\boldsymbol{c}_1,\cdots,\boldsymbol{c}_n)$, $\boldsymbol{B}=(\boldsymbol{b}_1,\cdots,\boldsymbol{b}_m)$, 依定义, 有

$$\begin{aligned}(\boldsymbol{C}'\otimes\boldsymbol{A})\mathrm{Vec}(\boldsymbol{B})&=\begin{pmatrix}c_{11}\boldsymbol{A} & c_{21}\boldsymbol{A} & \cdots & c_{m1}\boldsymbol{A}\\ c_{12}\boldsymbol{A} & c_{22}\boldsymbol{A} & \cdots & c_{m2}\boldsymbol{A}\\ \vdots & \vdots & & \vdots\\ c_{1n}\boldsymbol{A} & c_{2n}\boldsymbol{A} & \cdots & c_{mn}\boldsymbol{A}\end{pmatrix}\begin{pmatrix}\boldsymbol{b}_1\\ \boldsymbol{b}_2\\ \vdots\\ \boldsymbol{b}_m\end{pmatrix}\\ &=\begin{pmatrix}\boldsymbol{A}\boldsymbol{\Sigma}c_{j1}\boldsymbol{b}_j\\ \boldsymbol{A}\boldsymbol{\Sigma}c_{j2}\boldsymbol{b}_j\\ \vdots\\ \boldsymbol{A}\boldsymbol{\Sigma}c_{jn}\boldsymbol{b}_j\end{pmatrix}\\ &=\begin{pmatrix}\boldsymbol{ABc}_1\\ \boldsymbol{ABc}_2\\ \vdots\\ \boldsymbol{ABc}_n\end{pmatrix}=\mathrm{Vec}(\boldsymbol{ABC}).\end{aligned}$$

性质 (e) 得证.

在多元统计分析中, 经常需要考虑对称随机阵的向量化运算. 设 $\boldsymbol{A}=(a_{ij})_{p\times p}$ 为对称随机阵, 因其只含有 $p(p+1)/2$ 个不同的随机变量, 故只需对矩阵 $\boldsymbol{A}$ 的下三角部分向量化成 $p(p+1)/2\times 1$ 的向量. 记

$$\mathrm{Svec}(\boldsymbol{A})=(a_{11},a_{21},\cdots,a_{p1},a_{22},\cdots,a_{p2},\cdots,a_{pp})',$$

则称 $\mathrm{Svec}(\boldsymbol{A})$ 为对称矩阵 $\boldsymbol{A}$ 的向量化运算.

2.1.2 矩阵分解

定理 2.1.1 (Schmidt 三角化分解) $\boldsymbol{A}$ 为 $n\times m$ 阵 $(n\geqslant m)$, 则 $\boldsymbol{A}$ 可分解为

$$\boldsymbol{A}=\boldsymbol{QR},$$

其中 $\boldsymbol{Q}$ 为 $n\times m$ 列正交阵, $\boldsymbol{R}$ 为上三角阵, 其对角元素非负, 当 $\mathrm{rk}(\boldsymbol{A})=m$ 时, $\boldsymbol{R}$ 的对角元素全为正.