Wishart分布引论

刘金山 编著

科学出版社
www.sciencep.com

Wishart 分布引论

刘金山　编著

科学出版社
北　京

内 容 简 介

本书系统地给出 Wishart 统计分布理论的一些基本结果，并在此基础上介绍一些现代发展结果．主要内容有：作为预备介绍常用的矩阵代数知识．引进微分外积形式工具，并介绍 Haar 不变测度和矩阵积分；讨论多元正态和矩阵正态分布，并由此引进中心 Wishart 分布，讨论它的性质、矩量、Bartlett 分解和特征值的联合分布，并介绍逆 Wishart 分布和矩阵 β 分布；通过带状多项式矩阵变量超几何函数引进非中心 Wishart 分布，讨论它的性质和特征值的分布；将 Wishart 分布理论推广到球对称矩阵分布，讨论与其相关的矩阵 t 和 F 分布；一般地讨论正态矩阵二次型的分布，并给出其密度的级数表达形式．

本书可作为概率统计、生物统计和计量经济等相关学科专业的高年级本科生、硕士或博士研究生教材．也可作为高校教师、研究人员和科技人员的科研参考书．

图书在版编目(CIP)数据

Wishart 分布引论/刘金山编著．—北京：科学出版社，2005

ISBN 7-03-014631-X

Ⅰ.W…　Ⅱ.刘…　Ⅲ.分布-理论　Ⅳ.O211.3

中国版本图书馆 CIP 数据核字(2004)第 121525 号

责任编辑：陈玉琢　吕　虹　祖翠娥／责任校对：宋玲玲
责任印制：钱玉芬／封面设计：王　浩

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
源海印刷有限责任公司印刷

科学出版社发行　各地新华书店经销
*

2005 年 3 月第 一 版　开本：B5(720×1000)
2005 年 3 月第一次印刷　印张：18 3/4
印数：1—2 000　字数：352 000

定价：46.00 元

(如有印装质量问题，我社负责调换〈环伟〉)

序

随着电子信息技术的飞速发展，在生物医学、经济金融、工程技术和管理等领域工作的人们越来越需要并且有能力处理大型数据分析问题. 于是，作为处理大型数据问题的现代统计分支之一 —— 多元统计分析在这些领域的应用日益广泛和深入. 这些来自实践的需求大大地刺激了多元统计的理论与方法的研究. Wishart(威沙特) 分布作为多元统计分布理论的核心和众多统计方法的理论基础，自然吸引了统计学界的极大兴趣和注意力，统计学家们发表了许多研究成果. 限于篇幅，一般多元统计分析的教科书只讨论很有限的基础性内容，而大量的结果则散布在国内外各种期刊上，这给研究者和应用者带来很大的不便，十分需要有一部系统总结这些成果的专著. 现在呈现在读者面前的这本书《 Wishart 分布引论》正是适应了这一客观需要.

作者刘金山教授自 20 世纪 80 年代以来一直从事统计学研究，在国内外学术期刊上发表了一系列论文. 本书的写作以他近 20 年来的研究经历作基础. 本书内容丰富，涵盖了各种形式的 Wishart 分布，其中一部分是作者近期的研究成果. 本书的另一个特点是自封性和系统性都很强，前面几章给出了必要的预备性数学工具，使后续章节叙述能够简洁. 读者既可以把此书当作教科书，也可用作科学研究参考书.

鉴于 Wishart 分布理论在多元统计乃至整个现代统计学中的重要地位，我认为此书对广大统计理论研究人员来讲，几乎是他们书架上不可缺少的一本书. 这可免除他们对许多文献的翻检之劳，而方便地从此书中找到他们所需的东西. 我相信，这正是作者撰写这本书的初衷.

王松桂

2004 年 6 月于北京

前　言

多元统计分析以多元正态分布为其主要理论基础，而基于多元正态分布的样本协方差矩阵服从 Wishart 分布. 因而 Wishart 分布在几乎所有多元统计推断中起核心作用，它是多元统计分析中的最典型内容. 除了理论上的核心作用外， Wishart 分布在计量经济、生物、医学、工农业和社会科学等许多领域中都有重要应用.

多元统计分析中关于统计量分布的推导一直是一个十分重要的问题. 中心和非中心 Wishart 分布以及与它们有关的一些统计量的分布一直是这个问题的焦点. 在历史上，为了寻求统计量的分布发展了许多求分布密度的方法. 例如，矩法，直接变换法和随机变量变换法. 但是这些方法都不能有效地解决求非中心 Wishart 分布问题.

从 20 世纪 50 年代中期到 60 年代末，通过 James 等人的一系列工作，发现非中心 Wishart 分布与局部紧集上的 Haar 不变测度有着密切关系，并由此给出了一种由带状多项式矩阵变量超几何函数表达的密度形式. 与 Haar 不变测度积分使用差不多同时， James 等人引进了微分外积形式工具. 实际上，矩阵积分中的变量代换在几何上看来只是一个坐标系的变换，雅可比行列式相当于体积元改变的比值，而微分外积形式正好是处理这一类问题的有效工具.

Wishart 分布的内容极为丰富，除经典内容外，有大量的研究结果散布在统计和计量经济等多种学术杂志上. 但迄今为止尚未见关于这种矩阵分布理论的专著出现.

本书系统地给出 Wishart 统计分布理论的一些基本内容和近代发展结果. 第 1 章作为预备介绍多元分析中常用的一些矩阵代数知识，引进微分外积工具，并导出一些常见变换的雅可比行列式，还讨论了 Haar 不变测度和矩阵积分. 第 2 章介绍多元正态分布，包括向量正态及其二次型的分布和矩阵正态分布. 第 3 章系统地介绍中心 Wishart 分布，讨论它的性质、矩量、 Bartlett 分解、相关和回归矩阵的分布、特征值和特征向量的分布，并介绍逆 Wishart 分布及与 Wishart 分布密切相关的矩阵 β 分布. 第 4 章引进带状多项式矩阵变量超几何函数，并由此推导非中心 Wishart 分布密度，讨论它的一些基本性质，并推导它的特征值和特征向量的分布. 第 5 章将 Wishart 分布推广到比矩阵正态分布广泛的球对称分布类中，讨论与球对称分布有关的矩阵 t 分布和矩阵 F 分布. 第 6 章则一般地讨论矩阵正态分布的二次型分布，包括中心和非中心分布，推导它们及它们的函数的密度，给出密度的级数表达形式，并讨论二次型的 Wishart 性和独立性问题.

本书讨论中主要采用矩阵和微分外积方法，使推导过程显得较为简洁，易于读

者理解和掌握. 各章均配备了练习题, 因而既可以作为从事该领域研究人员的参考书, 也可作为概率统计、生物统计和计量经济等相关专业的教材之用.

本书的完成得到了作者的老师和许多同行的支持和帮助. 作者特别感谢陈希孺院士、王松桂教授、张尧庭教授、方开泰教授和王静龙教授, 他们的渊博学识和杰出工作对作者完成本书产生了很大影响, 特别是王松桂教授对作者从事统计学研究和顺利完成本书给予了巨大支持, 并且还在百忙中为本书作序, 对此作者深表感谢. 本书的顺利出版还要感谢国家自然科学基金 (编号: 10271033) 及华南农业大学校长科学基金的资助和科学出版社的良好合作.

由于水平所限, 书中一定有不少错误和缺点, 恳请批评指正.

作　者

2004 年 6 月

目　录

第 1 章　预　　备

本章给出一些常用矩阵概念和结论，对于那些基本和熟知的结论，我们只叙述而不给其证明．本章将介绍矩阵导数和微分，引进微分外积方法，并给出一些常见变换的雅可比 (行列式)．最后讨论不变测度，并介绍一些特殊的矩阵函数积分．

§1.1　基本矩阵概念

1.1.1　定义和基本性质

定义 1.1.1　一个 $n \times m$ 矩阵 A 是由 nm 个数值 $a_{11}, a_{12}, \cdots, a_{nm}$ 按照一定顺序排成的矩形阵列，即

$$A = \begin{bmatrix} a_{11} & \cdots & a_{1m} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nm} \end{bmatrix}, \tag{1.1.1}$$

简记为 $A = (a_{ij})$ 或 $A_{n \times m}$．

若不特别指出，我们总假定所讨论的矩阵为实数矩阵，虽然有许多结果对于复数矩阵也成立．如果 $m = n$, 则称 A 为 m 阶矩阵或方阵．m 阶单位矩阵记为 I_m 或 I. 矩阵 A 的转置记为 A'. 如果 $A = A'$, 称 A 为对称阵．若分别用 a_i 和 $a'_{(j)}$ 表示 A 的第 i 列和第 j 行，则 A 可表示为

$$A = (a_1, \cdots, a_m) = \begin{bmatrix} a'_{(1)} \\ \vdots \\ a'_{(n)} \end{bmatrix}. \tag{1.1.2}$$

如果一个 m 阶矩阵 A 的主对角线以下元素全为 0, 则称 A 为上三角阵，记为 $A \in UT(m)$. 若 A 的主对角线上而元素全为 0, 则称 A 为下三角阵，记为 $A \in LT(m)$. 如果 A 既是上三角阵又是下三角阵，则它是一个对角阵；若它的对角元素为 $a_{11}, \cdots, a_{mm}$, 则可记它为 $A = \mathrm{diag}(a_{11}, \cdots, a_{mm})$.

一个 m 阶矩阵 A 称为正交的，如果它满足 $AA' = A'A = I_m$, 记为 $A \in O(m)$, 这里 $O(m)$ 表示所有 m 阶正交矩阵集合．

定义 1.1.2　m 阶矩阵 A 的行列式定义为

$$\det(A) = \sum_{\pi} (-1)^{\varepsilon_\pi} a_{1j_1} a_{2j_2} \cdots a_{mj_m}, \tag{1.1.3}$$

其中 $\sum_{\pi}$ 表示对 $(1,\cdots,m)$ 的所有 $m!$ 个排列 $\pi=(j_1,\cdots,j_m)$ 求和， ε_π 是排列 π 的逆序数.

下面是行列式的一些基本性质：

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

(2) $\det(\alpha A)=(\alpha)^m\det(A)$.

(3) $\det(AB)=\det(A)\det(B)$.

(4) $\det(AA')\geqslant 0$, $\det(A'A)\geqslant 0$.

(5) $\det\begin{bmatrix} A & C \\ 0 & B\end{bmatrix}=\det\begin{bmatrix} A & 0 \\ D & B\end{bmatrix}=\det(A)\det(B)$, 其中 A,B 是方阵.

(6) $\det(I_n+AB)=\det(I_m+BA)$, 其中 A 是 $n\times m$ 的， B 是 $m\times n$ 的.

(7) 若 $T=(t_{ij})$ 是 m 阶上 (或下) 三角阵，则 $\det(T)=\prod_{i=1}^{m}t_{ii}$.

(8) 若 H 是 m 阶正交矩阵，即 $H\in O(m)$, 则 $\det(H)=\pm 1$.

定义 1.1.3　如果 $\det(A)\neq 0$, 则称 A 是可逆的，或非奇异的，此时存在唯一矩阵 B 使得 $AB=BA=I$, 称 B 为 A 的逆矩阵，记作 $B=A^{-1}$.

下面给出一些常用结果：

(1) 设 A 是 m 阶非奇异方阵，分块为

$$A=\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22}\end{bmatrix}, \tag{1.1.4}$$

其中 A_{11} 是 $p\times p$ 的，则

(a) 当 $\det(A_{11})\neq 0$ 时， $A_{22.1}=A_{22}-A_{21}A_{11}^{-1}A_{12}$ 非奇异，且有

$$\det(A)=\det(A_{11})\det(A_{22.1}), \tag{1.1.5}$$

$$A^{-1}=\begin{bmatrix} A_{11}^{-1}+A_{11}^{-1}A_{12}A_{22.1}^{-1}A_{21}A_{11}^{-1} & -A_{11}^{-1}A_{12}A_{22.1}^{-1} \\ -A_{22.1}^{-1}A_{21}A_{11}^{-1} & A_{22.1}^{-1}\end{bmatrix}. \tag{1.1.6}$$

(b) 当 $\det(A_{22})\neq 0$ 时， $A_{11.2}=A_{11}-A_{12}A_{22}^{-1}A_{21}$ 非奇异，且有

$$\det(A)=\det(A_{22})\det(A_{11.2}), \tag{1.1.7}$$

$$A^{-1}=\begin{bmatrix} A_{11.2}^{-1} & -A_{11.2}^{-1}A_{12}A_{22}^{-1} \\ -A_{22}^{-1}A_{21}A_{11.2}^{-1} & A_{22}^{-1}+A_{22}^{-1}A_{21}A_{11.2}^{-1}A_{12}A_{22}^{-1}\end{bmatrix}. \tag{1.1.8}$$

(2) 设矩阵 A,B 非奇异，而 C 和 D 是适当阶数的矩阵，则

$$(A+CBD)^{-1}=A^{-1}-A^{-1}CB[B+BDA^{-1}CB]^{-1}BDA^{-1}, \tag{1.1.9}$$

特别是，当取 $B=1$, 而 $C=u$ 和 $D'=-v$ 均为向量时，有

$$(A-uv')^{-1}=A^{-1}+\frac{1}{1-v'A^{-1}u}A^{-1}uv'A^{-1}. \tag{1.1.10}$$

(3) 设 A 是由 (1.1.2) 表示的 $n\times m$ 矩阵，则

$$A'A=(a_i'a_j)_{m\times m}=\sum_{j=1}^{n}a_{(j)}a_{(j)}'. \tag{1.1.11}$$

定义 1.1.4 称矩阵 A 的秩为 r, 记为 $rk(A)=r$. 如果 A 至少有一个 r 阶非零子式，而它的所有 $r+1$ 阶子式全为零，这里所谓 A 的 r 阶子式是指它的 r 阶子矩阵的行列式.

矩阵的秩具有以下基本性质：

(1) $rk(A)\leqslant\min\{n,m\}$, 这里 A 是 $n\times m$ 矩阵.

(2) $rk(A)=0$, 当且仅当 $A=0$.

(3) $rk(A)=rk(A')=rk(A'A)=rk(AA')$.

(4) $rk(AB)\leqslant\min\{rk(A),rk(B)\}$.

(5) $rk(A+B)\leqslant rk(A)+rk(B)$.

(6) 设 A 为 m 阶方阵，则 $rk(A)=m$, 当且仅当 A 非奇异.

(7) $rk(ABC)=rk(B)$, 这里 A 和 C 均为非奇异矩阵.

(8) 设 A 和 B 分别是 $n\times m$ 和 $m\times p$ 矩阵，满足 $AB=0$, 则 $rk(B)\leqslant m-rk(A)$.

定义 1.1.5 m 阶矩阵 A 的迹定义为它的主对角元素之和，记为 $\mathrm{tr}(A)$.

关于矩阵的迹有以下基本事实：

(1) $\mathrm{tr}(A)=\mathrm{tr}(A')$.

(2) $\mathrm{tr}(AB)=\mathrm{tr}(BA)$.

(3) $\mathrm{tr}(cA)=c\mathrm{tr}(A)$.

为了记号简便，以后用 $\mathrm{etr}(A)$ 表示 $\exp\{\mathrm{tr}(A)\}$.

1.1.2 特征值和特征向量

定义 1.1.6 设 A 是 m 阶矩阵， A 的特征值定义为特征方程

$$\det(A-\lambda I)=0 \tag{1.1.12}$$

的根 λ, 满足 $Ax=\lambda x$ 的非零向量 x 称为 A 的相应于 λ 的特征向量.

(1.1.12) 式的左边是 λ 的 m 次多项式，由多项式理论知道，它必有 m 个根 (可能有重根或复数根), 一般记为 λ_i 或 $\lambda_i(A)$, $i=1,2,\cdots,m$.

下面列出一些常用结果：

(1) 如果 A 是 m 阶实对称阵, 则它的 m 个特征值均为实数, 可按大小顺序排为 $\lambda_1 \geqslant \lambda_2 \geqslant \cdots \geqslant \lambda_m$.

(2) 相应于 A 的不同特征值的特征向量线性无关, 相应于对称矩阵的不同特征值的特征向量相互正交.

(3) A 和 A' 有相同的特征值.

(4) AB 和 BA 有相同的非零特征值.

(5) 设 λ 是 m 阶矩阵 A 的特征值, x 是相应于 λ 的特征向量, 并设 $\phi(A) = a_0I + a_1A + \cdots + a_nA^n$, 则 $\phi(\lambda) = a_0 + a_1\lambda + \cdots + a_n\lambda^n$ 和 x 分别是 $\phi(A)$ 的特征值和相应特征向量.

(6) 设 A 可逆, λ 是 A 的特征值, 则 λ^{-1} 是 A^{-1} 的特征值.

(7) 若 A 为正交阵, 则 $|\lambda_i| = 1, i = 1, 2, \cdots, m$.

(8) 若 A 是 m 阶矩阵, 则有 $\det(A) = \prod\limits_{i=1}^{m} \lambda_i$, $\mathrm{tr}(A) = \sum\limits_{i=1}^{m} \lambda_i$.

(9) 若 $f(\lambda)$ 是 A 的特征多项式, 即 $f(\lambda) = \det(A - \lambda I)$, 则 $f(A) = 0$.

定义 1.1.7　设 A 和 B 都是 m 阶矩阵, 若存在可逆阵 P, 使得 $B = P^{-1}AP$, 则称 B 与 A 相似, 记为 $B \sim A$.

相似矩阵有以下性质:

(1) 相似矩阵有相同的特征值, 因而也有相同的迹和行列式.

(2) m 阶矩阵 A 相似于对角阵 $\Lambda = \mathrm{diag}(\lambda_1, \cdots, \lambda_m)$, 当且仅当 A 有 m 个线性无关的特征向量.

(3) 若 $A = PBP^{-1}$, 即 $A \sim B$, 则 $A^k = PB^kP^{-1}$, 这里 k 是非负整数. 若 $\phi(x)$ 是一个多项式, 则

$$\phi(A) = P\phi(B)P^{-1},$$

特别是, 若 $A \sim \Lambda = \mathrm{diag}(\lambda_1, \cdots, \lambda_m)$, 则存在可逆矩阵 P 使得 $A = P\Lambda P^{-1}$, $A^k = P\Lambda^k P^{-1}$, $\phi(A) = P\phi(\Lambda)P^{-1}$, 其中

$$\Lambda^k = \mathrm{diag}(\lambda_1^k, \cdots, \lambda_m^k), \qquad \phi(\Lambda) = \mathrm{diag}(\phi(\lambda_1), \cdots, \phi(\lambda_m)).$$

(4) 若 A 为 m 阶对称阵, 则必存在一个正交阵 $\Gamma \in O(m)$, 使

$$\Gamma' A \Gamma = \mathrm{diag}(\lambda_1, \cdots, \lambda_m) \triangleq \Lambda,$$

其中 $\lambda_1, \cdots, \lambda_m$ 是 A 的特征值, Γ 的第 i 列为相应于 λ_i 的特征向量, $i = 1, \cdots, m$.

(5) 设 A, B 是两个 m 阶对称阵, 则存在正交阵 Q, 使得 $Q'AQ$ 和 $Q'BQ$ 同时为对角阵, 当且仅当 $AB = BA$.

1.1.3 非负定矩阵

定义 1.1.8 设 A 是 $m\times m$ 对称阵. 若对任何 $x\neq 0, x\in R^m$, 总有 $x'Ax>0(<0)$, 则称 A 为正 (负) 定矩阵, 记为 $A>0(<0)$; 若对任何 $x\in R^m$, 总有 $x'Ax\geqslant 0(\leqslant 0)$, 则称 A 是非负 (正) 定矩阵, 记为 $A\geqslant 0(\leqslant 0)$. 非负 (正) 定矩阵也可称为半正 (负) 定的.

关于非负定矩阵有以下一些性质:

(1) $A>0(\geqslant 0)$, 当且仅当它的特征值全为正 (非负).

(2) $BB'\geqslant 0$ 且 $B'B\geqslant 0$, 对一切矩阵 B 成立.

(3) 设 $A\geqslant 0$ 是 $m\times m$ 矩阵, B 是 $p\times m$ 矩阵, $p\leqslant m$, $rk(B)=r$, 则 $BAB'\geqslant 0$. 若 $A>0$, 则 $BAB'>0$, 当且仅当 $r=p$.

(4) 若 $A>0, B>0$, 且 $A-B>0$, 则 $B^{-1}-A^{-1}>0$, 且 $\det(A)>\det(B)$, 此时我们记 $A>B$. 若 $A\geqslant 0, B\geqslant 0$, 且 $A-B\geqslant 0$, 则 $\det(A)\geqslant\det(B)$, 此时我们记 $A\geqslant B$.

(5) 若 $A>0$, 将其如 (1.1.4) 分块, 则 $A_{11}>0, A_{22}>0, A_{11.2}=A_{11}-A_{12}A_{22}^{-1}A_{21}>0, A_{22.1}=A_{22}-A_{21}A_{11}^{-1}A_{12}>0$.

(6) 设 A,B 是两个 m 阶对称阵, 且 $B>0$, 则存在可逆阵 Q, 使得 $A=Q'\Lambda Q$, $B=Q'Q$, 其中 $\Lambda=\mathrm{diag}(\mu_1,\cdots,\mu_m)$, μ_i 为 AB^{-1} 的特征值.

(7) 设 A,B 是两个 m 阶非负定阵, 则存在可逆阵 Q, 使得 $Q'AQ$ 和 $Q'BQ$ 皆为对角阵.

在非负定阵中有一类重要矩阵, 如下定义.

定义 1.1.9 若矩阵 A 对称且幂等, 即 $A=A'$, 且 $A=A^2$, 则称 A 为投影阵.

投影阵有以下一些主要性质:

(1) 若 A 是投影阵, 则 $\mathrm{tr}(A)=rk(A)$.

(2) 若 A 是投影阵, 则 $I-A$ 也是投影阵.

(3) 若 A 是秩为 r 的投影阵, 则 A 有 r 个特征值为 1, 其余特征值全为 0. 故满秩的投影阵必为单位阵 I.

(4) 若 A,B 都是投影阵, 且 $A+B=I$, 则 $AB=BA=0$.

§1.2 矩阵分解

矩阵分解在多元统计分析中特别有用, 现归纳如下:

(1) 如果 A 是一个有实特征值的 $m\times m$ 实矩阵, 则 A 可分解为

$$A=H'TH, \tag{1.2.1}$$

其中 $H\in O(m)$, 而 $T\in UT(m)$, 其对角元素是 A 的特征值.

(2) 如果 A 是一个 $m\times m$ 对称矩阵，其特征值为 $\lambda_1,\cdots,\lambda_m$, 则

$$A=H'\Lambda H, \tag{1.2.2}$$

其中 $H\in O(m),\Lambda=\mathrm{diag}(\lambda_1,\cdots,\lambda_m)$. 如果记 $H=(h_1,\cdots,h_m)'$, 则 h_i 是相应于 A 的特征值 λ_i 的特征向量. 若特征值 $\lambda_1,\cdots,\lambda_m$ 互不相等，则表达式 (1.2.2) 除 H 的第一行元素改变符号外是唯一的. 我们可以把 (1.2.2) 改写为

$$A=\sum_{i=1}^{m}\lambda_i h_i h_i', \tag{1.2.3}$$

此式称为 A 的谱分解. 当 $\lambda_1=\cdots=\lambda_r=1,\lambda_{r+1}=\cdots=\lambda_m=0$ 时，上述分解变为

$$A=\sum_{i=1}^{r}h_i h_i'=(h_1,\cdots,h_r)(h_1,\cdots,h_r)'\triangleq H_1'H_1.$$

(3) 如果 $A>0(\geqslant 0)$, 则存在 $A^{\frac{1}{2}}>0(\geqslant 0)$ 使得 $A=A^{\frac{1}{2}}A^{\frac{1}{2}}$, 这里 $A^{\frac{1}{2}}$ 可以分解为 $A^{\frac{1}{2}}=H'\Lambda^{\frac{1}{2}}H$, 其中 H 和 Λ 如在 (1.2.2) 中. 当 $A>0$ 时，一般地有

$$f(A)=H'f(\Lambda)H, \tag{1.2.4}$$

其中 $f(\Lambda)=\mathrm{diag}(f(\lambda_1),\cdots,f(\lambda_m))$, 而 $f(x)$ 是一个 Borel 函数.

(4) 设 $m\times m$ 矩阵 $A\geqslant 0$, $rk(A)=r$, 则

(a) 存在一个秩为 r 的 $m\times r$ 或 $m\times m$ 阵 B, 使得

$$A=BB'. \tag{1.2.5}$$

(b) 存在一个 $m\times m$ 的非奇异阵 C, 使得

$$A=C\begin{bmatrix} I_r & 0\\ 0 & 0\end{bmatrix}C'. \tag{1.2.6}$$

(c) 存在 $T\in TU(m)$, 其对角元素非负，使得

$$A=T'T, \tag{1.2.7}$$

上式称为楚列斯基 (Cholesky) 分解. 若 $A>0$, 则 T 的对角元素全为正，且此时分解是唯一的.

(5) 设 A 是 $n\times m$ 矩阵 $(n\geqslant m)$, 则存在 $n\times m$ 阵 Q, 满足 $Q'Q=I_m$ 和 $m\times m$ 阵 $B\geqslant 0$, 使得

$$A=QB. \tag{1.2.8}$$

若 $rk(A)=m$, 则 $B>0$, 上式还可以进一步分解为

$$A=H\begin{bmatrix} I_m \\ 0 \end{bmatrix}B,$$

其中 $H\in O(n)$.

(6) (Schmidt 三角化分解) 设 A 为 $n\times m$ 阵 $(n\geqslant m)$, 则 A 可分解为

$$A=QR, \tag{1.2.9}$$

其中 Q 为 $n\times m$ 列正交阵, 即 $Q'Q=I_m$, $R\in UT(m)$, 其对角元素非负. 当 $rk(A)=m$ 时, R 的对角元素全为正.

(7) 设 A 和 B 分别是 $k\times m$ 和 $k\times n$ 矩阵 $(n\geqslant m)$, 则 $AA'=BB'$, 当且仅当存在 $m\times n$ 列正交矩阵 Q, 使得 $AQ=B$.

(8) 设 A 是 $n\times m$ 矩阵 $(n\geqslant m)$, 则 A 可以分解为

$$A=Q\Lambda V, \tag{1.2.10}$$

其中 Q 如在 (1.2.9) 中, $\Lambda=\mathrm{diag}(\lambda_1,\cdots,\lambda_m)$, $V\in O(m)$. 由于

$$A'A=V'\Lambda Q'Q\Lambda V=V'\Lambda^2V,$$

$\lambda_1^2,\cdots,\lambda_p^2$ 是 $A'A$ 的特征值, V 的行向量是 $A'A$ 的特征向量. 又因为 $AA'=Q\Lambda^2Q'$, 可知 Q 的列向量是 AA' 的特征向量. (1.2.10) 还可以进一步分解为

$$A=H\begin{bmatrix} \Lambda \\ 0 \end{bmatrix}V, \tag{1.2.11}$$

其中 $H\in O(n)$,V 和 Λ 如在 (1.2.10) 中.

(9) 设 $A_1,\cdots,A_k$ 为 $m\times m$ 对称矩阵, 且 $A_iA_j=0, i\neq j, i,j=1,\cdots,k$, 则存在 $H\in O(m)$, 使得 $H'A_iH=\Lambda_i$, 这里 Λ_i 为对角阵, $i=1,\cdots,k$.

(10) 设 A 和 B 均为 $m\times m$ 矩阵, $A>0, B'=B$, 则存在 $m\times m$ 非奇异阵 H, 使得

$$A=HH', \qquad B=H\Lambda H', \tag{1.2.12}$$

其中 $\Lambda=\mathrm{diag}(\lambda_1,\cdots,\lambda_m)$ 是 $A^{-1}B$ 的特征值. 若 $B>0$ 且 $\lambda_1,\cdots,\lambda_m$ 互不相等, 则 H 除首行元素符号改变外是唯一的.

(11) (满秩分解) 设 A 为 $n\times m$ 阵, $rk(A)=r$, 则存在列满秩阵 $B_{n\times r}$ 和行满秩阵 $C_{r\times m}$, 使得 $A=BC$.

(12) (秩分解) 设 A 为 $n\times m$ 阵, 则存在两个可逆阵 $P_{n\times n}$ 和 $Q_{m\times m}$, 使得

$$A=P\begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}Q, \tag{1.2.13}$$

其中 $r=rk(A)$.

(13) 若 A 是 $m\times m$ 对称幂等阵, 且 $rk(A)=k$, 则存在 $m\times k$ 矩阵 H 满足 $H'H=I_k$, 使得

$$H'AH=I_k, \qquad HH'=A. \tag{1.2.14}$$

(14) 设 A 是 $m\times m$ 非奇异阵, 若 A 的 i 阶主子式非零 $(i=1,\cdots,m)$, 则

$$A=TU, \tag{1.2.15}$$

其中 $T\in LT(m)$, $U\in UT(m)$, 且 $u_{ii}=1,\ i=1,\cdots,m$.

§1.3　矩阵向量化和 Kronecker 积

记 $e_i(n)$ 为第 i 个元素为 1 其余元素全为 0 的 n 维向量, 即 $e_i(n)=(0,\cdots,0,1,0,\cdots,0)'$, 并记 $E_{ij}(n,m)=e_i(n)e_j(m)'$. 为简单起见, 我有时分别记它们为 e_i 和 E_{ij}. 我们有以下基本结果:

(1) $e_i'e_j=\delta_{ij}$, 这里 δ_{ij} 为 Kronecker 算子, 即 $\delta_{ii}=1, \delta_{ij}=0\ (i\neq j)$.

(2) $E_{ij}e_k=\delta_{jk}e_i,\ e_k'E_{ij}=\delta_{ki}e_j'$.

(3) $E_{ij}E_{rs}=\delta_{jr}E_{is}$.

(4) $I=\sum\limits_i E_{ii}=\sum\limits_i e_ie_i'$.

(5) $E_{ij}'(n,m)=E_{ji}(m,n)$.

1.3.1　矩阵向量化

所谓矩阵向量化就是将矩阵按列拉成一个向量.

定义 1.3.1　设 $A=(a_1,\cdots,a_m)$ 是 $n\times m$ 矩阵, 定义一个 nm 维向量

$$\mathrm{vec}(A)=\begin{bmatrix} a_1 \\ \vdots \\ a_m \end{bmatrix}, \tag{1.3.1}$$

其中 "vec" 称为拉直算子.

显然,

$$\mathrm{vec}(A')=\begin{bmatrix} a_{(1)} \\ \vdots \\ a_{(m)} \end{bmatrix},$$