

密级：公开

编号：21110588

中山大學

博士学位论文

典型群作用下的矩阵分布

Matrix Distribution under Classical Group Actions

学位申请人： 王浩铭

导师姓名及职称： 黄辉教授 蒋智超教授

专业名称： 统计学

2025年 11月 24 日

中山大学博士学位论文

典型群作用下的矩阵分布

Matrix Distribution under Classical Group

Actions

专业：统计学

申请人：王浩铭

指导教师：黄辉教授，蒋智超教授

答辩委员会（签名）

主席：

成员：
潘文亮
姜云华
易欣

2025年5月31日

论文原创性声明

本人郑重声明：所呈交的毕业论文，是本人在导师的指导下，独立进行研究工作所取得的成果。除文中已经注明引用的内容外，本论文不包含任何其他个人或集体已经发表或撰写过的作品成果。对本论文的研究做出重要贡献的个人和集体，均已在文中以明确方式标明。本人完全意识到本声明的法律结果由本人承担。

学位论文作者签名：王浩铭

日期：2025年11月25日

学位论文使用授权声明

本人完全了解中山大学有关保留、使用学位论文的规定，即：学校有权保留学位论文并向国家主管部门或其指定机构送交论文的电子版和纸质版；有权将学位论文用于非赢利目的的少量复制并允许论文进入学校图书馆、院系资料室被查阅；有权将学位论文的内容编入有关数据库进行检索；可以采用复印、缩印或其他方法保存学位论文；可以为建立了馆际合作关系的兄弟高校用户提供文献传递服务和交换服务。

保密论文保密期满后，适用本声明。

学位论文作者签名：王浩铭

导师签名：蒋智超

日期：2025年11月25日

日期：2025年11月25日

典型群作用下的矩阵分布

专业: 统计学

博士生: 王浩铭

指导教师: 黄辉教授, 蒋智超教授

摘要

本文提出了四类矩阵正态分布, 作为变量间效应 (Ψ) 与样本间效应 (Φ) 可分离模型 ($\Phi \otimes \Psi$) 的推广. 根据精度矩阵的特定张量形式, 得到了样本方差和协方差的联合分布, 进而得到了相应的积矩分布, 矩阵 t 分布, 与矩阵 F 分布. 一些已知的结果, 包括非中心威沙特分布和正态二次型分布, 现在作为主要定理的推论出现. 此外, 通过证明一个适应的、右连续的、非递减的、具有单位跳跃且初值为零的整数值随机过程具有极小可料强度当且仅当它是绝对连续测度变换下的标准泊松过程, 我们设计了点过程的重要性采样与拟合优度检验算法. 最后, 我们将方法应用于元素丰度与银河系演化巡天数据集 (SAGES), 给出了 RA=25.21, DEC=34.17 天区的 12 个老年红超巨星的候选 ID, 可供天文学实证研究.

关键词: 矩阵正态分布, 不变多项式, 经典特殊函数, 点过程, 天文数据

Matrix Distribution under Classical Group Actions

Major: Statistics

Name: Wang Haoming

Supervisors: Prof. Huang Hui, Prof. Jiang Zhichao

Abstract

This thesis introduces four matrix normal distributions extending the separable covariance $\Phi \otimes \Psi$ with potentially variable-level (Ψ) and/or sample-level (Φ) correlations. The joint distribution of sample variances and covariances, leading to the product moment distribution, matrix t distribution, and matrix F distribution, is considered when precision matrices admit a specific tensor form. Several well-known results, including the non-central Wishart distribution and normal quadratic forms, now appear as corollaries. Besides, by proving that an adapted, right-continuous, non-decreasing, integer-valued process with unit jumps and starting at zero has a minimal predictable intensity if and only if it is a standard Poisson process under an absolutely continuous transformation of measures, we design algorithms for importance sampling and goodness-of-fit test of point processes. Finally, applying these methods to the Stellar Abundance and Galactic Evolution Survey (SAGES) dataset, we provide twelve candidate IDs for old red supergiants in the region RA=25.21, DEC=34.17, which can be used for further astronomical studies.

Keywords: Matrix normal distribution, Invariant polynomial, Classical special function, Point process, Astronomical data

目录

第 1 章 引言	3
1.1 问题背景	3
1.2 相关研究	5
1.3 本文工作	8
第 2 章 矩形坐标与球面坐标	9
2.1 矩形坐标	10
2.2 球面坐标	11
第 3 章 典型群作用下的不变多项式	13
3.1 对称群	13
3.2 典型群	18
3.3 计算若干积分	20
第 4 章 正态矩阵与抽样分布	23
4.1 矩阵正态分布	23
4.2 积矩分布	24
4.3 矩阵 t 分布	32
4.4 矩阵 F 分布	34
4.5 数值模拟	37
第 5 章 点过程的极小可料强度	41
5.1 极小可料强度	41
5.2 典则分解	48
第 6 章 天文案例分析	53
6.1 图像去噪	53
6.2 局部峰值检测	54
6.3 拟合优度检验	55
6.4 判别分析	56
参考文献	65

本文约定

记号约定

x, y, z, \dots	标量或向量.
A, B, C, \dots	矩阵.
$GL(n)$	n 阶可逆方阵的一般线性群.
$O(n)$	n 阶实方阵的正交群.
$U(n)$	n 阶复方阵的酉群.
$SL(n)$	n 阶可逆方阵中满足行列式为 +1 的子群. $SO(n), SU(n)$ 类似.
$Sp(n)$	$2n$ 阶复方阵的辛群.
$G_{n,p}$	n 维线性空间中所有 p 维子空间的格拉斯曼流形.
$V_{n,p}$	$n \times p$ 矩阵的斯蒂菲尔流形 $X'X = I_p$.
$\mathfrak{R}_{\text{I,II,III,IV}}$	四类典型域.
$\mathfrak{C}_{\text{I,II,III,IV}}$	四类典型域的特征流形.
dX	矩阵 X 的微分形式.
(dX)	矩阵微分形式的外积, $\bigwedge_{i,j}^n$ 矩阵, $\bigwedge_{i \leq j}^n$ 对称方阵, $\bigwedge_{i < j}^n$ 斜对称方阵.
$[dX]$	紧群上的归一化哈尔测度.
λ, μ, ν, \dots	整数分拆.
$(a)^\kappa$	上阶乘幂. 类似地, $(a)_\kappa$ 表示下阶乘幂.
$\Gamma_p(a)$	多元伽马函数.
$B_p(a, b)$	多元贝塔函数.
$\mathcal{F}_s, s \leq t$	样本空间 Ω 上的 σ -代数流.
\mathcal{F}_∞	包含所有 $\mathcal{F}_t, t \geq 0$ 的最小 σ -代数.
$\mathcal{F}_s^N, s \leq t$	由过程 N 生成的 σ -代数流.

经典特殊函数

1. ${}_0F_0$ 指数函数.

- (a) χ^2 分布.
- (b) 样本协方差分布 (威沙特分布).
- (c) 样本协方差的特征根分布.

2. ${}_1F_0$ 二项式级数.

- (a) 方差比 F 分布.
- (b) 两样本协方差的特征根分布 (总体协方差不相等 $\Sigma_1 \neq \Sigma_2$).

3. ${}_0F_1$ 贝塞尔函数.

- (a) 非中心 χ^2 分布.
- (b) 非中心样本协方差分布 (非中心威沙特分布).
- (c) 非中心样本协方差的特征根分布 (总体协方差已知).

4. ${}_1F_1$ 合流超几何函数.

- (a) 非中心 F 分布.
- (b) 非中心多元贝塔分布 (狄利克雷分布).
- (c) 非中心两样本协方差的特征根分布.

5. ${}_2F_1$ 高斯超几何函数.

- (a) 多重相关系数.
- (b) 典则相关系数.

第1章 引言

1.1 问题背景

伴随着科技的进步和生产资料的积累, 以矩阵为对象的分布走进了人们的视野. 传统的多元统计分析, 比如多元时间序列的向量自回归模型, 作为向量的序列在时间上具有相关性, 更重要的是, 在不同分量之间也具有相关性. 经典的 t 统计量, F 统计量都是针对单变量或者是向量定义的, 不能满足实际生产生活的需求. 这一困难体现在多元时序协方差结构的张量形式的复杂性. 具体来说, 它们往往是多维度、方差异质和非平稳的. 这为我们数学建模带来了挑战.

一种可取的办法是从随机过程理论出发, 将时间序列视作二阶随机场, 时间维度视作第一指标, 向量维度视作第二指标, 来研究它们的性质. 经典的卡夫宁-洛伊夫定理和傅里叶分析技术为这一方法的数学基础提供了有力保障^[1-2]. 另一办法则是考虑典型群作用下的矩阵分布, 这里的典型群一般指正交群、酉群、辛群等. 这一办法更为直接, 不用引入冗余超参数, 直接对面板数据建模, 但其数学理论却更为艰深, 用到的技术也更加精细. 此外, 近来多元统计分析的教材比如^[3-4], 包括更经典的^[5-7], 大都只讨论了独立同分布多元正态样本的抽样分布, 例如威沙特分布, 但对一般非独立同分布的情形, 现有文献则相对较少^[3-4].

本文将阐述两种方法之间的联系, 通过引入多参数随机场的矩形坐标来推导具有相依条目的矩阵正态分布及其抽样分布. 其分布理论的构建依赖典型矩阵变换下的不变多项式, 这是贯穿本文的核心概念. 通常来说, 直接在典型群上积分往往比较困难, 我们需要在恰当的多项式空间中选择一组规范正交基来表示拉普拉斯-贝特拉米算子的解, 这其中的典型解就是带 (zonal) 多项式, 数学物理方程中称作球调和. 这组基扮演了类似极坐标下径向-角向分解中的角向部分, 使得我们能够方便地进行积分运算. 由于这一方面的参考文献还相对较少, 我们将尽可能地从叙述历史的角度出发, 介绍我们所需要用到的工具办法. 可参考的材料包括: 典型群表示理论^[8-10], 经典多元统计分析^[11-17], 带多项式与超几何函数^[18-21]. 阅读这些材料对理解本文的主要工具和思想是有裨益的.

矩阵分布常常出现在多维、非马尔可夫、连续参数的随机过程当中, 其中一个典型的代表就是点过程. 点过程的定义与我们常说的信息流密切相关, 信息流应当包含所有当下可能发生的事件全体. 火山爆发之前一般伴随着地壳震动, 将火山爆发的时刻视作停时, 这一时刻的停止信息流就应当包含火山爆发之前由地

壳震动引起的所有可能发生的事件全体. 停时则是近代随机过程发展中出现的核心概念. 比如一个积分制比赛, 负者失分而胜者得分, 参与者在初始积分失尽时停止比赛, 则比赛停止的时刻, 或参与者积分为零的时刻就是一个停时. 在柯尔莫哥洛夫的公理体系中, 信息流或称滤子就是概率空间上一族单调递增的子西格玛代数, 条件期望则是可测函数在子西格玛代数上的投影. 可料过程是指那些仅依赖于先于当下时刻的信息的随机过程. 比如风雨交加的夜晚我们总是先看见闪电再听到雷声, 自然界中的雷声就是一个可料过程. 这一定义总是依赖于信息流的. 所谓极小滤子问题, 就是在概率空间上对所有滤子寻找一个关于包含偏序的极小元, 使得点过程的强度保持可料性质. 因为由集合论中的公理极小滤子总是存在的. 极小滤子可以很大, 甚至超出由过程全体生成的西格玛代数, 但如果极小滤子是由过程自然生成的西格玛代数流, 它显然就是最小的.

相较于已被广泛研究的滤子扩张问题, 例如^[22-29], 有关滤子限制问题的相关研究依旧很少^[30-33]. 与前人的工作相比, 我们并不假设零集完备条件, 这使得经典的哈恩-巴拿赫定理得以应用于极小滤子问题. 通过研究一个适应的、右连续的、非减的、具有单位跳跃且初值为零的整值随机过程的极小可料强度, 我们肯定了布里莫^[34]于1972年提出的一个猜想. 该定理有两个推论^[35], 一个是任意霍克斯过程都是绝对连续测度变换下的标准泊松过程, 另一推论则是考克斯过程可以分解为非齐次泊松过程与奇异点过程之和, 在假设分解独立的前提下. 这为拟合优度检验和重要性采样提供了有力保障, 填补了相关领域的理论空白.

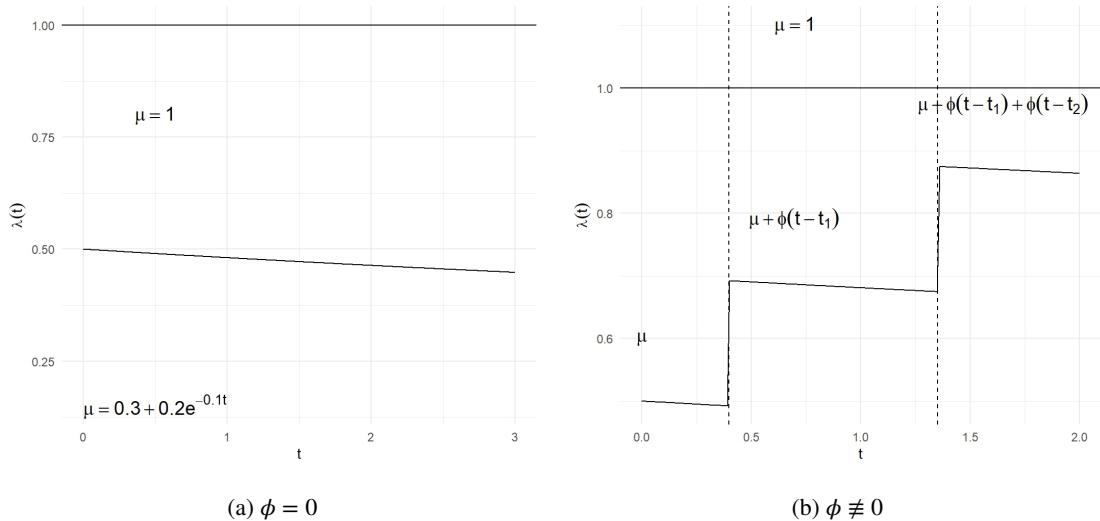


图 1.1 霍克斯过程在两种情形下强度函数 λ_t 随时间 t 变化的单个样本轨道: (a) $\mu(t) = 0.3 + 0.2e^{-0.1t}$; (b) $\mu(t) = 0.1 + 0.2e^{-0.1t}$ 且 $\phi(t) = 0.2$.

1.2 相关研究

在单变量统计分析中, χ^2 分布, t 分布和 F 分布等抽样分布都是基于独立同分布正态样本的最常见的统计量. 戈塞(笔名“Student”)[36]最早研究了来自正态总体样本标准差的分布, 并据此得到了样本均值与样本标准差之比的分布, 也即 χ^2 分布和 t 分布. 这为多变量统计分析奠定了基础.

在两变元的情形, 除了总体的均值向量和方差向量, 我们还需要估计总体分量间的两两相关系数或回归系数. 不同于单变元情形, 费歇尔[37]最早得到了来自独立同分布二元正态总体的样本方差与样本协方差的联合分布. 卡尔·皮尔逊[38]和罗马诺夫斯基[39]在费歇尔工作的基础上得到了相应的积矩分布. 这是威沙特推广 χ^2 分布至矩阵情形的发轫.

在多变元的情形, 威沙特[40]最早在中心情形(考虑样本均值 $M, M = 0$)给出了来自多元正态总体独立同分布样本的积矩分布. 后人如安德森[41](rank(M) ≤ 2) 和詹姆斯[42](rank(M) 任意) 给出了积矩分布在非中心情形的推广,

$$\begin{aligned} W_p(S; n) &= \frac{1}{2^{\frac{np}{2}} \Gamma_p(\frac{n}{2}) |\Psi|^{\frac{n}{2}}} \text{etr}\left(-\frac{1}{2}\Psi^{-1}S\right) |S|^{\frac{n-p-1}{2}}, \text{ 当 } M = 0; \\ &\quad \times \text{etr}\left(-\frac{1}{2}\Omega\right) {}_0F_1\left(\frac{n}{2}; \frac{1}{4}\Omega\Psi^{-1}S\right), \text{ 当 } M \neq 0; \end{aligned} \quad (1.1)$$

其中 $\text{etr}(Z) = \exp \text{tr}(Z)$, $|Z| = \det(Z)$, $\Gamma_p(a)$ 是多元伽马函数,

$$\Gamma_p(a) = \int_{A>0} \text{etr}(-A) |A|^{a-\frac{p+1}{2}} (dA) = \pi^{\frac{p(p-1)}{4}} \prod_{i=1}^p \Gamma\left(a - \frac{i-1}{2}\right), \quad (1.2)$$

其中 a 的实部 $\Re(a) > \frac{1}{2}(p-1)$, $A > 0$ 意味着积分取遍所有 $p \times p$ 实对称正定矩阵, 且第二个等号右边的 Γ 代表通常意义下的伽马函数, ${}_0F_1(a; Z)$ 是矩阵参量超几何函数. 在(1.1)中, n 是样本的大小, p 是变量的数目, Ψ 是 p 元正态总体的协方差矩阵, Ω 是样本均值的非中心参数, $S = (s_{ij})$ 是由 $\frac{1}{2}p(p+1)$ 个样本矩构成的 $p \times p$ 对称矩阵,

$$\Omega = \Psi^{-1}M'M, \quad S = X'X,$$

其中 n 个 p 维正态向量按行排列成 $n \times p$ 样本矩阵 X . 当 $p = 1$ 时, (1.1)式就退化成样本方差乘上自由度为 n 的皮尔逊非中心卡方分布

$$\begin{aligned} \bar{x} &= \frac{1}{n} \sum_{i=1}^n x_i, \\ s^2 &= \sum_{i=1}^n (x_i - \bar{x})^2 + n\bar{x}^2, \end{aligned}$$

其中 x_1, \dots, x_n 代表来自 $N(\mu, \sigma^2)$ 的样本, 这里的 μ 和 σ 代表正态总体的均值和标准差. 此时, 样本方差

$$D^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2$$

与样本均值 \bar{x} 独立. 戈塞定义的 t 分布即是它们的比值,

$$t = \frac{\bar{x}}{D}.$$

近些年来, 许多作者^[3-4,43-52]尝试将威沙特的工作从独立同分布正态向量推广到椭球等高分布族, 其中不同样本之间可能存在相关性. 在多元统计分析中, 非独立样本引出了一个根本性的问题, 那就是样本间相关系数或回归系数往往非零. 如果坚持采用最精确的假设, 在假设正态分布的前提下, 确定 n 个样本在 p 个变量中的样本方差和协方差通常需要估计 $\frac{1}{2}np(np - 1)$ 个两两相关系数或回归系数, 这些参数的数目近似于 $O((np)^2)$. 为了减轻轻过拟合的负担, 一种常见的方法是假设总体协方差或者精度矩阵由特定的张量结构决定, 例如文献^[53]中的椭球等高分布假设. 从这个意义上讲, 我们只需要考虑 $\frac{1}{2}p(p + 1)$ 个二次型的分布,

$$\begin{aligned} & y'_1 A_{11} y_1, y'_1 A_{12} y_2, \dots, y'_1 A_{1p} y_p, \\ & y'_2 A_{22} y_2, \dots, y'_2 A_{2p} y_p, \\ & \vdots \\ & y'_p A_{pp} y_p, \end{aligned} \tag{1.3}$$

这对应于四种总体精度矩阵的张量形式 $T_1, T_{1\frac{1}{2}}, T_2$ 和 T_3 . 然而, 正如文献^[53]所指出的, 在多元椭球等高分布的谱分解中独立性仍然是一个未解决的问题. 本文的 4.1 节将利用精度矩阵的四种嵌套类型的谱分解来解决这一问题. 这一推广建立了正态假设下四种张量记号与椭球等高分布之间的一一对应关系. 基于此, 我们在 4.2 节推导了这些正态矩阵二次型的分布, 并由此得到了相应的矩母函数与特征根的联合分布. 这将包括非中心情形在内的二次型分布从多元正态总体推广到更为一般的椭球等高总体.

卡特里^[43]首先得到了齐次二次型的联合分布, 即 $A_{ij} = A$, 其中 A 是一个固定的精度矩阵, 其超几何函数项为双矩阵参量超几何函数 ${}_0F_0$. 在卡特里的工作^[43]中, 二次型的精确分布并没有像文献^[44]中那样被表示成带多项式的级数, 但却引入了讨厌参数 $q_{ij} > 0$ 和超几何函数 ${}_0F_0$. 由于(1.1)式中的中心威沙特密度函数类似于伽马密度, 超几何函数 ${}_0F_0$ 并未显式出现, 这种表达式在实践中难以应用. 此外, 如詹姆斯^[12,15]所述, 在中心情形, 双矩阵参量的超几何函数经常出

现在矩母函数和特征根分布当中, 而非密度函数, 并且这些函数通常难以计算. 因此, 将近三十年后, 马泰^[48]才通过累积量 (cumulants) 重新考察了 $p = 2$ 时的异质二次型分布. 本文确定了一般的 $p > 2$ 情形下 $\frac{1}{2}p(p+1)$ 个异质二次型的中心分布, 且不引入超几何函数和不必要的参数 q_{ij} . 只有在非中心情形下, 这些异质二次型的分布才表示为单矩阵参量超几何函数 ${}_0F_1$, 这与(1.1)式保持一致. 为了理论自洽, 这些矩阵必须满足 $A'_{ij} = A_{ji}$, 且 $A_{ij}(i \neq j)$ 的对角线元素均为零. 作为处理维数灾难的主要技术, 我们需要考虑同时对角化 $T_1, T_{\frac{1}{2}}, T_2$ 和 T_3 来减少未知参数数目, 而这些概念代表着多元统计分析从独立样本假设到相依样本假设的转变.

这一推广的另一动机来自已被诸多学者研究的矩阵 t 分布^[3,54-55]. 尽管有些复杂, 单变量情形的非中心 t 分布可以通过合流超几何函数 ${}_1F_1$ 来表示, 但是该函数的数值计算往往较为困难. 鉴于此问题的复杂性, 顿内特^[56]将单变量 t 分布推广到双变量中心情形, 随后约翰^[57], 克希尔斯加尔^[58]和狄克^[59]分别将其推广到多变量 t 分布, 三者均针对中心情形. 至于非中心情形, 现有关于 t 分布在多变量情形和矩阵变量情形的推广都仅考虑了中心比值再加上固定均值的定义, 以避免引入合流项 ${}_1F_1$, 这一点可见文献^[3,54-55,60-61]. 然而, 如果总体均值是一个未知参数, 则中心统计量通常是无法精确得到的. 因此, 即使我们对总体知之甚少, 也需要构建恰当的均值统计量, 尤其是在非中心情形下. 为了与非中心单变量 t 分布的定义相吻合, 本文在 4.3 节引入了具有非中心矩阵正态分子的矩阵 t 分布. 由此推导出矩阵 t 分布的密度函数也具有类似的合流超几何函数项 ${}_1F_1$. 从均值参数的敏感性分析中可以看出, 该密度函数与单变量 t 分布相对接近.

除此之外, 在方差分析或判别分析中, 求最小特征根的分布被康斯坦丁^[16]称作是非常困难的. 而贝伦斯-费歇尔问题则是这一问题的核心, 特别是涉及组间独立但方差异质的情形. 其精确检验的构造要用到形如(1.3)的两个独立二次型之比. 在非中心威沙特情形(1.1), 卡特里^[62], 戴维斯^[63], 平川文子^[64], 筑瀬靖子^[65]和古普塔^[66]分别计算了比统计量的概率密度函数, 其具有超几何函数项 ${}_1F_0$. 事实上, 该结果对更一般的左球分布类也成立, 其中心情形由劳特^[67]得到. 要得到非独立且非中心情形下比统计量的分布, 我们首先应得到形如(1.3)的 $\frac{1}{2}p(p+1)$ 个二次型的非中心分布. 本文在 4.4 节给出了矩阵 F 分布的定义, 并证明了此情形下比统计量也具有类似的超几何函数项 ${}_1F_0$. 此外, F 统计量能够同时考虑均值与方差的波动信息, 基于特征根的统计量在统计应用中扮演着重要角色. 本文同时给出了矩阵 F 分布最大最小特征根的分布, 较之单个积矩分布的最大最小特征根的分布^[62,68-70], 这一形式更为简洁.

矩阵分布广泛出现在天文学与点过程的实际应用中。事实上，点过程是一类特殊的半鞅。根据杜布-梅耶分解定理，点过程可以分解为局部鞅和可料有界变差过程之和。布朗运动、泊松过程、伊藤扩散等都是半鞅的典型例子。我们知道半鞅在测度变换下仍是半鞅，本文则是通过测度变换研究了时变半鞅的一个问题。对于连续半鞅，蒙洛衣^[71]证明了任何连续局部鞅都是时变布朗运动，尽管可能需要扩大滤过概率空间。根据佐恩引理，概率空间上存在一个最小滤子。对于布朗运动，许多作者已经解决了这个问题，例如角古静夫^[22]，吉萨诺夫^[23]，克拉克^[24]，诺维科夫^[25]，风卷纪彦^[26]等。与布朗运动类似，泊松过程是许多随机应用中的又一原型。事实上，点过程作为一个适应的、右连续的、非减的、具有单位跳跃且初值为零的整值随机过程，可以表示为时变泊松过程。然而，适当的滤子扩张仍然是必要的。布里莫^[34]曾提问点过程的极小滤子何时是由过程自然生成的，并猜想当且仅当它是绝对连续测度变换下的标准泊松过程。本文 5.1 节对此猜想作出了肯定的回答。

虽然在很多情况下，滤子扩张有助于我们表示点过程，但它也给这类过程的分析带来了一些麻烦，例如包含更多的零集。滤子限制问题是滤子扩张问题的对偶，后者的有关文献见^[27-29]。极小滤子问题最早由雅各德^[30]研究，随后的推广见布里莫^[31]，达利和维尔琼斯^[32]，卡尔^[33]。他们的方法均基于风险函数的条件期望。例如^[32]中的定理 14.1 和^[33]中的定理 2.31。在该方法中，不可避免地需要满足狄拉歇尔^[72]意义下的零集完备假设，即所有终端零集都包含在初始滤子中。然而，基于零集完备的显式构造限制了极小滤子的实用性。本文采用泛函分析中的哈恩-巴拿赫定理证明了极小可料强度的存在性。该定理仅涉及半鞅的闭子空间的性质，使得我们能够摆脱不必要的零集完备假设，从而获得自然的滤子限制。这一定理在点过程的拟合优度检验与重要性采样中有重要应用。

1.3 本文工作

本文将 np 个正态变量根据精度矩阵的奇异值分解分为四类 $T_1, T_{1\frac{1}{2}}, T_2$ 和 T_3 ，并据此推广了相关抽样分布，如积矩分布、矩阵 t 分布和矩阵 F 分布。其优势在于类似^[53] 定理 3.6.7 中椭球等高分布的简洁表示，有助于理解更一般的对称分布。此外，本文还证明了一个点过程具有极小可料强度当且仅当它是绝对连续测度变换下的标准泊松过程。这肯定了布里莫^[34] 第 48 页提出的一个猜想。通过天文学案例分析，我们给出了元素丰度与银河系演化巡天数据集 (SAGES) 当中 RA=25.21, DEC=34.17 天区的 12 个老年红超巨星的候选 ID，可供天文学实证研究。

第2章 矩形坐标与球面坐标

设 $T \subseteq [0, \infty)$, $S \subseteq \mathbb{R}^d$ 是欧式空间中的紧集. 复值随机场 $Z(t, x), t \in T, x \in S$ 称作是二阶的, 如果

$$\mathbf{E} \iint_{T \times S} |Z(t, x)|^2 dt dx < \infty.$$

随机变量列 X_n 称作依二阶矩收敛到 X , 如果 $\lim_{n \rightarrow \infty} \mathbf{E}|X_n - X|^2 \rightarrow 0$.

定义 2.1 ($T_1, T_{\frac{1}{2}}, T_2$ 与 T_3). 设 $Z(t, x)$ 的均值为零, $\mathbf{E}[Z(t, x)] = 0, t \in T, x \in S$.

T_1 如果存在希尔伯特空间 $L^2(S)$ 上的一组标准正交基 $\{\psi_n(x)\}_{n=1}^\infty$, 使得对任意 $t \in T$, $Z(t, x)$ 可以在依二阶矩收敛的意义下, 关于 $x \in S$ 一致地展开为级数

$$\sum_{n=1}^{\infty} \xi_n(t) \psi_n(x)$$

其中 $\{\xi_n(x)\}_{n=1}^\infty$ 是 S 上的一组零均值随机场, 满足 $\mathbf{E}[\xi_n(t) \overline{\xi_{n'}(t)}] = c_n(t) \delta_{n,n'}$.

这里的 $c_n(t)$ 是 T 上的非负实值函数, 满足对任意 $t \in T$, $c_n(t)$ 关于 $n \downarrow 0$.

$T_{\frac{1}{2}}$ 同 T_1 . 如果额外地, 对任意 $t, t' \in T$, $\mathbf{E}[\xi_n(t) \overline{\xi_{n'}(t')}] = c_n(t) \delta_{n,n'} \delta_{t,t'}$.

T_2 同 T_1 . 如果额外地, 存在希尔伯特空间 $L^2(T)$ 上的一组标准正交基 $\{\phi_m(t)\}_{m=1}^\infty$, 使得对任意 n , $\xi_n(t)$ 可以在依二阶矩收敛的意义下, 关于 $t \in T$ 一致地展开为级数

$$\sum_{m=1}^{\infty} \eta_{nm} \phi_m(t)$$

其中 $\{\eta_{nm}\}_{n,m=1}^\infty$ 是两两不相关的零均值随机变量, 满足 $\mathbf{E}(\eta_{nm} \overline{\eta_{n'm'}}) = c_{nm} \delta_{n,n'} \delta_{m,m'}$.

这里的 $\{c_{nm}\}_{n,m=1}^\infty$ 是非负实数列, 满足对任意 n , ω_{nm} 关于 $m \downarrow 0$.

T_3 同 T_2 . 如果额外地, 存在非负实数列 $\{\sigma_m^2\}_{m=1}^\infty$ 和 $\{\tau_n^2\}_{n=1}^\infty$ 分别满足 $\sigma_m^2 \downarrow 0$ 和 $\tau_n^2 \downarrow 0$ 以及 $c_{nm} = \tau_n^2 \sigma_m^2$.

(*) 同 T_3 . 如果额外地, 存在随机变量列 $\{\lambda_m\}_{m=1}^\infty$ 和 $\{\mu_n\}_{n=1}^\infty$, 使得

$$\eta_{nm} = \mu_n \lambda_m$$

并且它们满足 $\mathbf{E}(\lambda_m \overline{\lambda_{m'}}) = \sigma_m^2 \delta_{m,m'}$ 和 $\mathbf{E}(\mu_n \overline{\mu_{n'}}) = \tau_n^2 \delta_{n,n'}$.

记 $c_i(t) = \mathbf{E}|\xi_i(t)|^2$, $c_i(t_1, t_2) = \mathbf{E}[\xi_i(t_1) \xi_i(t_2)]$ 与 $c_{ij} = \mathbf{E}|\eta_{ij}|^2$. 根据本征正交分解定理^[2] 第二卷, 第 37.3 节, 我们有定义 2.1 的充分必要刻画, 证明见^[73].

表 2.1 $T_1, T_{\frac{1}{2}}, T_2$ 与 T_3 的比较.

T_1	$\Leftrightarrow E[Z(t, x_1)\overline{Z(t, x_2)}] = \sum_i c_i(t) \psi_i(x_1) \overline{\psi_i(x_2)}$.
	$\Leftrightarrow \psi_i(x)$ 是 $C(t, t, x_1, x_2)$ 的特征函数, 与 t 无关.
$T_{\frac{1}{2}}$	$\Leftrightarrow E[Z(t_1, x_1)\overline{Z(t_2, x_2)}] = \sum_i c_i(t_1, t_2) \psi_i(x_1) \overline{\psi_i(x_2)}$.
	$\Leftrightarrow \psi_i(t)$ 是 $C(t_1, t_2, x_1, x_2)$ 的特征函数, 与 t_1, t_2 无关.
T_2	$\Leftrightarrow E[Z(t_1, x_1)\overline{Z(t_2, x_2)}] = \sum_i \sum_j c_{ij} \phi_j(t_1) \overline{\phi_j(t_2)} \psi_i(x_1) \overline{\psi_i(x_2)}$.
	$\Leftrightarrow c_{ij}, \phi_j(t) \psi_i(x)$ 是 $C(t_1, t_2, x_1, x_2)$ 的特征值与特征向量.
T_3	$\Leftrightarrow E[Z(t_1, x_1)\overline{Z(t_2, x_2)}] = C_2(t_1, t_2) C_1(x_1, x_2)$.
	$\Leftrightarrow \tau_j^2, \phi_j(t), \sigma_i^2, \psi_i(x)$ 分别是 $C_2(t_1, t_2), C_1(x_1, x_2)$ 的特征值与特征向量.
(*)	\Leftrightarrow 存在 $X(x), Y(t)$ 使得 $Z(t, x) = Y(t)X(x)$, $E[Y(t_1)\overline{Y(t_2)}] = C_2(t_1, t_2)$, 和 $E[X(x_1)\overline{X(x_2)}] = C_1(x_1, x_2)$.

2.1 矩形坐标

用符号 $A \otimes B$ 代表矩阵 $A = (a_{ij})$ 和矩阵 B 的克罗内克积, 即分块矩阵 $(a_{ij}B)$. 用符号 $\text{vec}(A)$, $A = (a_1, a_2, \dots, a_n)$ 代表矩阵 A 的向量化 $[a_1; a_2; \dots; a_n]$. 设 $\mathbb{R}, \mathbb{C}, \mathbb{O}$ 分别代表实数域、复数域和四元数体.

定义 2.2 (矩形坐标). 考虑 $\mathbb{F} = \mathbb{R}, \mathbb{C}, \mathbb{O}$ 上的 $np \times np$ 分块矩阵 Θ 如下

$$\Theta = \begin{bmatrix} \Theta_{11} & \Theta_{12} & \dots & \Theta_{1n} \\ \Theta_{21} & \Theta_{22} & \dots & \Theta_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \Theta_{n1} & \Theta_{n2} & \dots & \Theta_{nn} \end{bmatrix}$$

其中 Θ_{ii} ($i = 1, 2, \dots, n$) 是维数为 p 的正定矩阵. 所谓矩形坐标, 是指 $p \times p$ 矩阵 $B = (b_1, b_2, \dots, b_p)$, 满足 $\bar{b}_i' b_j = \delta_{ij}$, 使得 Θ 具有如下结构

T_1 如下展开式成立

$$\Theta = \sum_{j=1}^p \sum_{j'=1}^p A_{jj'} \otimes B_{jj'}, \quad (2.1)$$

其中 $B_{jj'} = b_i \bar{b}_{j'}'$, 且 $A_{jj'}(k, l) = \bar{b}_j' \Theta_{kl} b_{j'}$ 满足

$$A_{jj'}(k, k) = 0 (j \neq j').$$

换句话说, $A_{jj'} (j \neq j')$ 的对角线元素为零.

$T_{1\frac{1}{2}}$ 展开式 T_1 成立, 且

$$A_{jj'} = A_{jj} \delta_{jj'}, \quad (2.2)$$

也就是说, $A_{jj'} (j \neq j')$ 为零.

T_2 展开式 T_1 成立, 且存在另一组 $n \times n$ 矩形坐标 $A = (a_1, a_2, \dots, a_n)$, 满足

$$\bar{a}'_i a_j = \delta_{ij}, \text{使得}$$

$$\Theta = \sum_{i=1}^n \sum_{j=1}^p \gamma_{ij} A_i \otimes B_j, \quad (2.3)$$

其中 $A_i = a_i \bar{a}'_i$, $B_j = b_j \bar{b}'_j$.

T_3 展开式 T_2 成立, 且存在正常数 α_i, β_j 使得 $\gamma_{ij} = \alpha_i \beta_j$, 或等价地, 下式成立

$$\Theta = \Phi^{-1} \otimes \Psi^{-1}, \quad (2.4)$$

其中 $\Phi^{-1} = \sum_{i=1}^n \alpha_i a_i \bar{a}'_i$, $\Psi^{-1} = \sum_{j=1}^p \beta_j b_j \bar{b}'_j$.

从表2.1不难看出

定理 2.1. 矩形坐标是唯一的.

另一个推论是勋伯格^[74]的离散化情形.

定理 2.2. Σ 是正定的当且仅当 $A_{jj} (j = 1, 2, \dots, p)$ 是正定的.

更具体地, 我们可以利用正定函数与高斯过程的一一对应来验证球面上含时正定函数的系数定正这一充要性质, 这一点容易从球谐函数的加法定理看出. 因此, 表2.1中的充要条件还可作为研究二阶随机过程理论的一个有力工具.^[75]

2.2 球面坐标

设 n, p 是正整数且 $n \geq p$. 不失一般性地, 假设 Z 是 $n \times p$ 列满秩实矩阵.

定义 2.3 (球面坐标). 任意 $n \times p$ 矩阵 Z 可唯一地分解为 $Z = H R^{\frac{1}{2}}$, 其中 $H (n \times p)$ 满足 $H' H = I_p$, $R (p \times p)$ 是对称正定矩阵. 我们称 H 为 Z 的球面坐标. 特别地, $H = Z(Z' Z)^{-\frac{1}{2}}$, $R = Z' Z$.

我们把满足 $H' H = I_p$ 这样条件的 $n \times p$ 矩阵集合称作斯蒂菲尔流形, 记作 $V_{n,p}$. 对一组球面坐标 $H (n \times p)$, 我们总可以把它扩充成一个正交矩阵 $K = (H, H_{\perp}) \in O(n)$, 其中 $H_{\perp} \in V_{n,n-p}$ 的列向量与 $H \in V_{n,p}$ 的列向量两两正交. 詹姆斯^[11]公式(8.19)最早证明了

命题 2.3 (球面分解). $(dZ) = 2^{-p}|R|^{\frac{1}{2}(n-p-1)} \cdot (dR) \cdot (dK)$, 其中 $(dK) = (H'dH) \cdot (H'_\perp dH)$.

让我们来逐项解释定理2.3的几何意义. 首先, $(H'dH)$ 描述了对象 H 的列向量在其张成的 p 维子空间内部的旋转. $H'dH$ 是一个 $p \times p$ 斜对称矩阵, 其独立参数个数为 $\frac{p(p-1)}{2}$. 这些参数对应了在固定 p 维子空间下, 正交基 H 在该子空间内的所有可能旋转, 也就是 $O(p)$ 的体积元. 而 $(H'_\perp dH)$ 描述了 p 维子空间在 n 维空间中的方向变化. 具体而言, $H'_\perp dH$ 是一个 $(n-p) \times p$ 矩阵, 其独立参数个数为 $p(n-p)$. 这些参数描述了 p 维子空间如何相对于其正交补 H_\perp 变化, 即子空间在格拉斯曼流形 $G_{n,p}$ 中的运动. 换句话说, 两者共同描述了斯蒂菲尔流形 $V_{n,p}$ 的几何结构, 其中 $V_{n,p}$ 是以格拉斯曼流形 $G_{n,p}$ 为底空间, 以正交群 $O(p)$ 为纤维的丛. 一般来说, 如果不要求 R 是对称矩阵, 满足 $R'R = Z'Z$ 的 R 不一定是唯一的. 于是我们有下面的定义.

定义 2.4 (上三角坐标). 任意 $n \times p$ 矩阵 Z 可唯一地分解为 $Z = HT$, 其中 $H (n \times p)$ 满足 $H'H = I_p$, $T (p \times p)$ 是对角线元素为正的上三角矩阵, 即 $t_{ii} > 0 (i = 1, 2, \dots, p)$. 我们称 H 为 Z 的上三角坐标.

命题 2.4 (上三角分解). $(dZ) = 2^{-p}|T'T|^{\frac{n-p-1}{2}} d(T'T) \cdot (dK)$.

命题2.4的证明. 取 $T'T$ 为 $Z'Z$ 的科列斯基分解, 由定理2.3可得. \square

另外两种常用的坐标是对角坐标和奇异值坐标, 分别针对方阵和一般情形.

定义 2.5 (对角坐标). 设 Z 是 $n \times n$ 正定矩阵. 则 Z 可唯一地分解为 $Z = H\Lambda H'$, 其中 $H (n \times n)$ 满足 $H'H = I_n$, 且 $\Lambda (n \times n)$ 是正对角矩阵, 即 $\lambda_{ii} > 0 (i = 1, 2, \dots, n)$. 我们称 H 为 Z 的对角坐标.

命题 2.5 (对角分解). $(dZ) = \prod_{i < j}^n (\lambda_i - \lambda_j) \wedge_1^n d\lambda_i \cdot (dH)$.

命题2.5的证明. 见津村善郎^[76] 定理 2.1. \square

定义 2.6 (奇异值坐标). 任意 $n \times p$ 矩阵 Z 可唯一地分解为 $Z = H_1\Lambda H_2'$, 其中 $H_1 (n \times n)$ 满足 $H_1'H_1 = I_n$, $H_2 (p \times p)$ 满足 $H_2'H_2 = I_p$, 且 $\Lambda (n \times p)$ 满足 $\lambda_{ii} > 0 (i = 1, 2, \dots, p)$. 我们称 H_1, H_2 为 Z 的奇异值坐标.

命题 2.6 (奇异值分解). $(dZ) = \prod_{i=1}^p \lambda_i^{n-p} \prod_{i < j}^n (\lambda_i^2 - \lambda_j^2) \wedge_1^p d\lambda_i \cdot (dH_1) \cdot (dH_2)$.

命题2.6的证明. 见津村善郎^[76] 定理 2.2. \square

第3章 典型群作用下的不变多项式

3.1 对称群

3.1.1 置换基本定理

用 $S(n)$ 代表 n 个不同符号的集合 $\{a_1, a_2, \dots, a_n\}$ 上的所有置换, 即同构映射的全体. 形式上, $\sigma \in S(n)$ 可以写作

$$\sigma = \begin{pmatrix} a_1 & a_2 & \cdots & a_n \\ \sigma(a_1) & \sigma(a_2) & \cdots & \sigma(a_n) \end{pmatrix}.$$

如果存在 $1 \leq k \leq n$, 和集合中的 k 个不同元素 b_1, b_2, \dots, b_k , 使得 $\sigma(b_1) = b_2, \sigma(b_2) = b_3, \dots, \sigma(b_k) = b_1$, 且 $\sigma(a) = a, a \neq b_1, b_2, \dots, b_k$, 那么 σ 就称作长度为 k 的轮换, 记作

$$\sigma = (a_1 a_2 \dots a_k).$$

如果 $(a_1 a_2 \dots a_k)$ 和 $(b_1 b_2 \dots b_l)$ 是两个没有公共元素的轮换, 它们称作不交轮换. 显然, 不交轮换对乘法, 即映射的复合, 满足交换律. 长度为 2 的轮换称作对换.

命题 3.1 (置换基本定理). 任意置换都可以分解为两两不交轮换的乘积, 它们在仅相差顺序和平凡轮换的意义下唯一.

命题 3.2 (轮换基本定理). 任意长度为 k 的轮换都可以分解为 $k - 1$ 个对换的乘积, 且同一置换的不同对换分解中, 对换的个数具有相同的奇偶性.

3.1.2 杨图

用 $\lambda = (\lambda_1, \lambda_2, \dots, \lambda_k)$ 代表正整数 n 的一个分拆, 记为 $\lambda \vdash n$, 满足 $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_k \geq 0, \lambda_1 + \lambda_2 + \dots + \lambda_k = n$. 此时称 n 为 λ 的权重, 记作 $|\lambda| = n$. 整数分拆称作是有效的, 如果 $\lambda_k > 0$, 此时 k 称作 λ 的长度, 记作 $l(\lambda)$. 比方说, $n = 4$ 的所有有效分拆为

$$(4), (3, 1), (2, 2), (2, 1, 1), (1, 1, 1, 1).$$

对任意有效分拆 λ , 其对应的杨图是一个由方格构成的左对齐图形, 其行数等于分拆的长度, 第 i 行方格数等于分拆的第 i 个分量 λ_i .

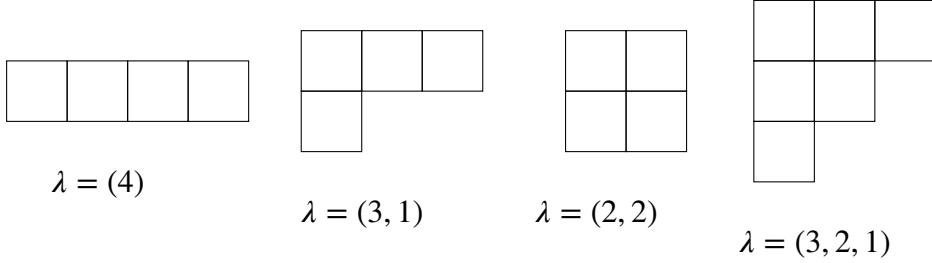


图 3.1 有效分拆的杨图示例.

任意杨图, 交换其行和列得到的新杨图称为原杨图的共轭杨图, 对应的分拆称为原分拆的共轭分拆, 记为 λ' , 其中 λ' 的第 i 个分量 λ'_i , 等于原分拆 λ 中不小于 i 的分量个数, 即原杨图第 i 列的方格数.

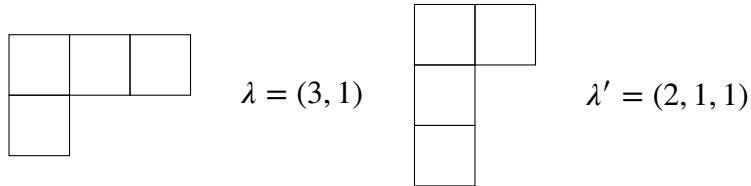


图 3.2 共轭杨图的示例.

特别地, 若 $\lambda = \lambda'$, 则称 λ 为自共轭分拆, 对应的杨图为自共轭杨图. 例如, $\lambda = (3, 2, 1)$ 的杨图转置后仍为自身, 故 $\lambda = (3, 2, 1)$ 是自共轭分拆.

命题 3.3 (杨图与分拆的一一对应). 对任意正整数 n , 映射 $\lambda \mapsto$ 杨图(λ) 是 n 的所有有效分拆集合到 n 对应的所有杨图集合的双射.

用 $\lambda = (\lambda_1, \lambda_2, \dots, \lambda_k)$ 和 $\mu = (\mu_1, \mu_2, \dots, \mu_l)$ 代表两个有效分拆, 如果对所有正整数 i , $\mu_i \leq \lambda_i$, 则称 μ 是 λ 的子分拆, 记为 $\mu \subseteq \lambda$. 当 $\mu \subseteq \lambda$ 时, 斜分拆定义为 $\nu = \lambda/\mu$, 其对应的斜杨图是 λ 的杨图与 μ 的杨图的差集. 若斜杨图中任意两个方格可通过相邻方格(上下左右)连接, 则称为连通斜杨图. $\nu = (3, 2, 1)/(1, 1)$ 和 $\nu = (4, 3)/(2, 1)$ 都是连通斜杨图.

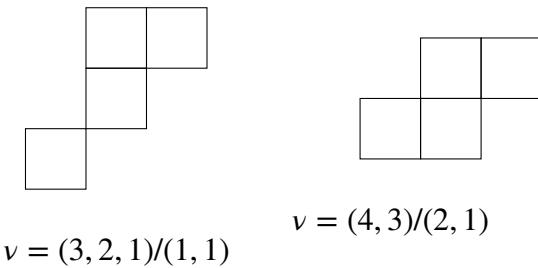


图 3.3 斜杨图的示例.

斜杨图的填充是从斜杨图的方格到正整数集的映射 $T : \text{斜杨图}(\nu) \rightarrow \mathbb{N}_+$. 半标准填充 T 需满足两项条件

1. 每行的数字从左到右非严格递增, 即对同一行的方格 $T(i, j) \leq T(i, j + 1)$.
2. 每列的数字从上到下严格递增, 即对同一列的方格 $T(i, j) < T(i + 1, j)$.

半标准填充的权重定义为向量 $\alpha = (\alpha_1, \alpha_2, \dots)$, 其中 α_k 表示数字 k 在填充中出现的次数, 记为 $T \in \text{SSYT}(\nu, \alpha)$. 标准填充是特殊的半标准填充, 除满足上述行、列条件外, 还要求填充的数字为 $1, 2, \dots, |\nu|$ (无重复). 此时数字在每行、每列均严格递增, 标准填充的集合记为 $\text{SYT}(\nu)$. 斜杨图 ν 的填充映射个数称为 ν 的填充数.

李特尔伍德-理查森填充是一类特殊的半标准填充, 除满足半标准填充的行、列条件外, 还需满足反向格子置换条件, 将填充的数字按列优先顺序(从右到左读取每一行, 每行从下到上)排列得到序列 $w = w_1 w_2 \cdots w_{|\nu|}$, 对任意 $1 \leq t \leq |\nu|$ 和任意 $k \geq 1$, 序列前 t 项中数字 $k + 1$ 的出现次数不超过数字 k 的出现次数, 这类填充记为 $\text{LR}(\nu/\lambda, \mu)$, 其中 ν, λ, μ 为整数分拆, 且 $|\nu| = |\lambda| + |\mu|$.

3.1.3 对称多项式

设 $\text{GL}(n)$ 是由 n 阶可逆方阵构成的一般线性群. 我们定义 $U \in \text{GL}(n)$ 的对称克罗内克幂如下. 首先构造 k 个 U 的通常意义上的克罗内克积 $U^{\otimes k} = U \otimes U \otimes \cdots \otimes U$. 它作用在 k 次张量空间 $\mathbb{C}^n \otimes \mathbb{C}^n \otimes \cdots \otimes \mathbb{C}^n$, 维度为 n^k . 然后引入对称群 $S(k)$ 的对称化算子 P_k ,

$$P_k(v_1 \otimes v_2 \otimes \cdots \otimes v_k) = \frac{1}{k!} \sum_{\sigma \in S(k)} v_{\sigma(1)} \otimes v_{\sigma(2)} \otimes \cdots \otimes v_{\sigma(k)}$$

即对所有置换取平均, 保留张量的对称分量. 由此定义 $U^{[k]} = P_k \circ U^{\otimes k} \circ P_k$, 它是 $U^{\otimes k}$ 在对称化算子作用下的像, 维度为

$$\binom{n+k-1}{k} = \frac{n(n+1)\dots(n+k-1)}{k!},$$

这一维度表示“从 n 个元素中进行 k 次有重复抽样的组合数”.

设 $\kappa = (k_1, k_2, \dots, k_n) \vdash k$. 所有单项式 $z_1^{k_1} z_2^{k_2} \cdots z_n^{k_n}$ 的集合构成 $n \times n$ 对称矩阵空间上解析函数空间的一个完备系, 且不同次数(即 $|\kappa| \neq \kappa'$)的单项式彼此正交. 当变量 $z = (z_1, z_2, \dots, z_n)$ 经 $\text{GL}(n)$ 中元素 U 变换为 $w = zU$ 时, 诱导出 $z^{[k]}$ 到 $w^{[k]}$ 的线性变换, 而 $z^{[k]}$ 是由标准化后的 k 次单项式构成的向量, 其分量为

$$\sqrt{\frac{k!}{k_1! k_2! \cdots k_n!}} z_1^{k_1} z_2^{k_2} \cdots z_n^{k_n},$$

并按照经典的字典序排列, 即 $\kappa < \kappa'$ 仅当存在某个索引 i ($1 \leq i \leq n$) 使得 $k_1 = k'_1, \dots, k_{i-1} = k'_{i-1}$ 且 $k_i < k'_i$. 在这样的表示下, 有

$$w^{[k]} = z^{[k]} U^{[k]}$$

其中 $U^{[k]}$ 正是上述定义的 U 的 k 次对称克罗内克幂, 它满足 $GL(n)$ 在 k 次对称多项式空间上群表示的定义, 同时保持单位元和乘法封闭.

关于表示矩阵的迹, 存在以下重要恒等式^[9]定理 1.4.2

命题 3.4 (华罗庚). $\text{tr}((U^{[2]})^{[k]}) = \sum \chi_{2k_1, 2k_2, \dots, 2k_n}(U).$

其中 $\chi_{2k_1, 2k_2, \dots, 2k_n}(U)$ 是 $GL(n)$ 在所有 $n \times n$ 对称矩阵集合上的不可约表示的特征标, 即表示矩阵的迹. 这些特征标 $\chi_{2k_1, 2k_2, \dots, 2k_n}$ 可以通过下面将要引入的带多项式的线性方程组来确定, 可见外尔^[8]第 181 页.

另一方面, 置换群 $S(2k)$ 的元素自然作用于对称张量空间 $V = S^{\otimes k}(S^{\otimes 2}(\mathbb{C}^n))$, 在该空间上定义出不可约表示 ω_{2k} , 同见外尔^[8]定理 4.4.F. 用 $H = S(2)^k \rtimes S(k)$ 代表对称群 $S(2)$ 的 k 次幂和对称群 $S(k)$ 的半直积. 用 $(\chi_{2k})^{GL}, (\omega_{2k})^H$ 分别代表在群 $GL(n)$ 和 $S(2k)$ 的子群 H 的作用下, V (或 $(\mathbb{C}^n)^{\otimes 2k}$) 上的不变子空间.

命题 3.5 (对偶原理). $V = \bigoplus_{\kappa \vdash k, \ell(\kappa) \leq n} (\chi_{2k})^{GL} \otimes (\omega_{2k})^H.$

3.1.4 特征标

在多元统计分析中, $\binom{n+k-1}{k}$ 维 k 次对称多项式空间中有一组常用的基, 称作带多项式. 它有多种不同的定义方式, 这里我们采用文献^[19]中的定义. 为此我们需要计算一些特别简单的情形, 比如单位矩阵的特征标.

命题 3.6. 设 $n \geq 2$. 定义 $D(x_1, \dots, x_n) = \prod_{i < j}^n (x_i - x_j)$. 如果 $h_1(x), \dots, h_n(x)$ 为实解析函数, 则有

$$\lim_{\substack{x_1 \rightarrow x \\ x_n \rightarrow x}} \frac{\begin{vmatrix} h_1(x_1) & \dots & h_n(x_1) \\ \vdots & \ddots & \vdots \\ h_1(x_n) & \dots & h_n(x_n) \end{vmatrix}}{D(x_1, \dots, x_n)} = \frac{(-1)^{\frac{1}{2}n(n-1)}}{1! 2! \dots (n-1)!} \begin{vmatrix} h_1(x) & \dots & h_n(x) \\ h'_1(x) & \dots & h'_n(x) \\ \vdots & \ddots & \vdots \\ h_1^{(n-1)}(x) & \dots & h_n^{(n-1)}(x) \end{vmatrix}.$$

命题 3.6 的证明. 不失一般性, 假设 x_1, \dots, x_n 均趋于 0. 该引理可由带余项的麦克劳林级数推出. \square

对于 $\kappa \vdash k$, 引入记号

$$M_{k_1, \dots, k_n}(x_1, \dots, x_n) = \begin{vmatrix} x_1^{k_1+n-1} & \dots & x_n^{k_1+n-1} \\ x_1^{k_2+n-2} & \dots & x_n^{k_2+n-2} \\ \vdots & \ddots & \vdots \\ x_1^{k_n} & \dots & x_n^{k_n} \end{vmatrix},$$

显然有 $M_{0, \dots, 0}(x_1, \dots, x_n) = D(x_1, \dots, x_n)$. 定义

$$N(k_1, \dots, k_n) = \lim_{\substack{x_1 \rightarrow x \\ x_n \rightarrow x}} \frac{M_{k_1, \dots, k_n}(x_1, \dots, x_n)}{D(x_1, \dots, x_n)}.$$

由引理3.6可得,

$$N(k_1, \dots, k_n) = \frac{D(k_1 + n - 1, k_2 + n - 2, \dots, k_n)}{D(n - 1, n - 2, \dots, 1, 0)}.$$

设 U 是 $n \times n$ 对称矩阵. 用 $(a)^k = a(a+1)\dots(a+k-1)$ 代表升阶乘幂, 以及 $(a)^\kappa = \prod_{i=1}^m (a - \frac{1}{2}(i-1))^{k_i}$. 定义

$$\begin{aligned} C_\kappa(I_n) &= \frac{2^k (\frac{1}{2}n)^\kappa}{(2k-1)!!} \chi_{2\kappa}(1), \\ C_\kappa(U) &= C_\kappa(I_n) |U|^{k_n} \int_{O(n)} \prod_{i=1}^{n-1} |HXH'(1, \dots, i)|^{k_i - k_{i+1}} (dH). \end{aligned} \tag{3.1}$$

这里的 $(1, \dots, i)$ 代表取主子矩阵, $\chi_{2k_1, \dots, 2k_n}(1)$ 可由引理3.6计算得到.

命题 3.7 (线性化公式). $C_\lambda(U) \cdot C_\mu(U) = \sum_{|\nu|=|\lambda|+|\mu|} g_{\lambda, \mu}^\nu C_\nu(U)$.

我们需要说明, n 元对称多项式空间存在多组基, 所以上面的系数可以通过已知系数的基变换来得到, 比如舒尔多项式的线性化系数. 一些典型的对称多项式基包括杰克、舒尔、带多项式等, 而舒尔的线性化公式是已知的, 称作李特尔伍德-理查德森系数. 具体来说, 对有效分拆 $\lambda \vdash n$, 定义其舒尔多项式为

$$s_\lambda(x_1, \dots, x_n) = \frac{M_{k_1, \dots, k_n}(x_1, \dots, x_n)}{D(x_1, \dots, x_n)}.$$

而带多项式 $C_\lambda(x_1, \dots, x_n)$ 作为对称多项式总可以展开为舒尔多项式的线性组合

$$C_\lambda(x_1, \dots, x_n) = \sum_\kappa P_\lambda^\kappa s_\kappa(x_1, \dots, x_n),$$

其中 $P = (P_\lambda^\kappa)$ 是上三角可逆的变换矩阵. 于是乘积系数 $g_{\lambda, \mu}^\nu$ 可以通过舒尔多项式的线性化系数 $c_{\lambda, \mu}^\nu$ 与矩阵 P 的线性关系得到

$$g_{\lambda, \mu}^\nu = \sum_{\alpha, \beta, \gamma} (P^{-1})_\lambda^\alpha (P^{-1})_\mu^\beta c_{\alpha, \beta}^\gamma P_\gamma^\nu.$$

其中 $c_{\lambda,\mu}^v$ 为非负整数, 其值等于斜杨图 v/λ 的李特尔伍德-理查德森填充数. 弗雷塔雷^[77]和库什纳^[78]分别对两种特殊情形, 即 $\mu = (1^k)$ 和 $\mu = (k)$ 给出了系数 $g_{\lambda,\mu}^v$ 的递推公式. 这两种情形的舒尔形式也称作皮埃里公式.

3.2 典型群

典型域 \mathfrak{R} 的特征流形 \mathfrak{C} 是 \mathfrak{R} 边界上的一部分, 凡 \mathfrak{R} 内的解析函数都在 \mathfrak{C} 上取极大绝对值, 且对 \mathfrak{C} 上的任一点, \mathfrak{R} 中一定有一解析函数在该点取极大值.

华罗庚^[9]最早计算出了典型域 $\mathfrak{R}_I, \mathfrak{R}_{II}, \mathfrak{R}_{III}, \mathfrak{R}_{IV}$ 上的泊松核的形式, 与典型域上的泊松积分公式与调和函数的定义有着密切的联系.

命题 3.8 (典型域上的泊松核).

1. 对适合于 $I_m - Z\bar{Z}' > 0$ 的 $m \times n$ 的复矩阵 Z 全体 \mathfrak{R}_I , 我们把它称作第 I 类典型域, 其 $\frac{1}{2}$ 实维数为 mn . 此外, \mathfrak{R}_I 上的泊松核具有以下形式

$$P(Z, U) = \frac{1}{\text{vol}(\mathfrak{C}_I)} \cdot \frac{|I - Z\bar{Z}'|^n}{|I - Z\bar{U}'|^{2n}}, \quad (3.2)$$

中 U 过适合于 $U\bar{U}' = I$ 的矩阵全体 \mathfrak{C}_I , 其实维数为 $m(2n - m)$. 特别地, 当 $m = n$ 时, 上式可简化为

$$P(Z, U) = \frac{1}{\text{vol}(\mathfrak{C}_I)} \cdot \frac{|I - Z\bar{Z}'|^n}{|Z - U|^{2n}}.$$

2. 对适合于 $I - Z\bar{Z} > 0$ 的 n 阶复对称方阵 Z 全体 \mathfrak{R}_{II} , 我们把它称作第 II 类典型域, 其 $\frac{1}{2}$ 实维数为 $\frac{1}{2}n(n + 1)$. 此外, \mathfrak{R}_{II} 上的泊松核具有以下形式

$$P(Z, U) = \frac{1}{\text{vol}(\mathfrak{C}_{II})} \cdot \frac{[I - Z\bar{Z}]^{\frac{n+1}{2}}}{|I - Z\bar{U}|^{n+1}}, \quad (3.3)$$

其中 U 过适合于 $U\bar{U} = I$ 的对称方阵全体 \mathfrak{C}_{II} , 其实维数为 $\frac{1}{2}n(n + 1)$.

3. 对适合于 $I + Z\bar{Z} > 0$ 的 n 阶斜对称复方阵 Z 全体 \mathfrak{R}_{III} , 我们把它称作第 III 类典型域, 其 $\frac{1}{2}$ 实维数为 $\frac{1}{2}n(n - 1)$. 当 $n = 2k$ 为偶数时, \mathfrak{R}_{III} 上的泊松核具有以下形式

$$P(Z, U) = \frac{1}{\text{vol}(\mathfrak{C}_{III})} \cdot \frac{[I + Z\bar{Z}]^{\frac{n-1}{2}}}{|I + Z\bar{U}|^{n-1}}, \quad (3.4)$$

其中 U 过所有斜对称酉方阵构成的集合 $\mathfrak{C}_{\text{III}}$. 此时, $\mathfrak{C}_{\text{III}}$ 的实维数为 $\frac{1}{2}n(n - 1)$. 当 $n = 2k + 1$ 为奇数时,

$$P(Z, U) = \frac{1}{\text{vol}(\mathfrak{C}_{\text{III}})} \cdot \frac{|I + Z\bar{Z}|^{\frac{n}{2}}}{|I + Z\bar{U}|^n}, \quad (3.5)$$

其中 U 过所有形如 VDV' 斜对称方阵全体 $\mathfrak{C}_{\text{III}}$, 其中 V 是酉方阵,

$$D = \left(\begin{array}{c|c} J_k & \\ \hline & 0 \end{array} \right)$$

是标准辛矩阵 J_k 与 0 的分块对角矩阵. 此时, $\mathfrak{C}_{\text{III}}$ 的实维数为 $\frac{1}{2}n(n + 1) - 1$.

4. 对于第 IV 类典型域 \mathfrak{R}_{IV} , 其由适合于

$$|zz'|^2 + 1 - 2\bar{z}z' > 0, \quad |zz'| < 1$$

的 $n(> 2)$ 维复向量 z 全体构成, $\frac{1}{2}$ 实维数为 n , 且具有泊松核

$$P(z, \xi) = \frac{1}{\text{vol}(\mathfrak{C}_{\text{IV}})} \cdot \frac{||zz'|^2 + 1 - 2\bar{z}z'|^{\frac{n}{2}}}{|(z - \xi)(z - \xi)'|^n}, \quad (3.6)$$

其中 ξ 过所有形如 $e^{i\theta}(x_1, \dots, x_n)$ 的矢量全体 \mathfrak{C}_{IV} , 且 x_1, \dots, x_n 是适合于 $x_1^2 + \dots + x_n^2 = 1$ 的实数. 于是 \mathfrak{C}_{IV} 的实维数为 $n - 1$.

对于任意在 \mathfrak{C} 上连续的函数 $f(U)$, 由泊松核定义出一个 \mathfrak{R} 上的解析函数

$$f(Z) = \int_{\mathfrak{C}} P(Z, U)f(U)dU,$$

称作 \mathfrak{R} 上的调和函数. 华罗庚还证明了, \mathfrak{R} 上的调和函数满足拉普拉斯方程

$$\sum_{\alpha, \beta=1}^n \sum_{j, k=1}^n \left(\delta_{\alpha\beta} - \sum_{l=1}^n x_{l\alpha} \bar{x}_{l\beta} \right) \left(\delta_{jk} - \sum_{\gamma=1}^n x_{j\gamma} \bar{x}_{k\gamma} \right) \frac{\partial^2 u}{\partial x_{j\alpha} \partial \bar{x}_{k\beta}} = 0.$$

命题 3.9 (詹姆斯^[79]). 带多项式 $C_\kappa(U)$ 是 \mathfrak{R}_{II} 上的调和函数.

事实上, 在满足一定的归一化条件后, 比如

$$[\text{tr}(U)]^k = \sum_{\kappa} C_{\kappa}(U),$$

带多项式可以由拉普拉斯方程和归一化条件唯一确定, 见穆尔海德^[7] 定义 7.2.1. 詹姆斯^[79] 据此得到了带多项式的一个递推公式. 设 u_1, \dots, u_n 是 U 的特征值.

$$C_{\kappa}(U) = \sum c_{\kappa, \lambda} u_1^{\lambda_1} \dots u_n^{\lambda_n},$$

$$c_{\kappa,\lambda} = \sum_{\lambda < \mu \leq \kappa} c_{\kappa,\mu} [(l_i + t) - (l_j - t)] / (\rho_\kappa - \rho_\lambda)$$

其中和式过所有 $\mu = (\lambda_1, \dots, \lambda_i + t, \dots, \lambda_j - t, \dots, \lambda_n)$, $t = 1, \dots, (\lambda_{i-1} - \lambda_i) \wedge (\lambda_j - \lambda_{j+1})$, 且 $\rho_\kappa = \sum_{i=1}^n k_i(k_i - i)$,

在下文中, 我们还需要用到典型域及其特征流形的体积, 我们整理了文献^[9]和一些常用的已知结果, 这些体积公式整理在了下面的表格和推论中.

表 3.1 四类典型域及其特征流形的体积.

典型域	体积	特征流形	体积
\mathfrak{R}_I	$\frac{(m-1)! \dots 2!1!(n-1)! \dots 2!1!}{(m+n-1)! \dots 2!1!} \pi^{mn}$	\mathfrak{C}_I	$\frac{2^n}{(m-1)!(m-2)! \dots (m-n)!} \pi^{\frac{mn-n(n-1)}{2}}$
\mathfrak{R}_{II}	$\frac{2!4! \dots (2n-2)!}{n!(n+1)!(n+2)! \dots (2n-1)!} \pi^{\frac{1}{2}n(n+1)}$	\mathfrak{C}_{II}	$\frac{2^n}{1!2! \dots (n-1)!} \pi^{\frac{n(n+1)}{2}}$
\mathfrak{R}_{III}	$\frac{2!4! \dots (2n-4)!}{(n-1)!n! \dots (2n-3)!} \pi^{\frac{1}{2}n(n-1)}$	\mathfrak{C}_{III}	(偶) $\frac{1!3!(n-3)!}{1!2! \dots (n-2)!} 2^{\frac{n}{2}} \pi^{\frac{n^2}{8}}$ (奇) $\frac{1!3!(n-2)!}{1!2! \dots (n-1)!} 2^{\frac{n-1}{2}} \pi^{\frac{(n+1)^2}{4}} - 1$
\mathfrak{R}_{IV}	$\frac{1}{2^{n-1}n!} \pi^n$	\mathfrak{C}_{IV}	$\frac{2}{\Gamma(\frac{n}{2})} \pi^{\frac{n}{2}}$

推论 3.10. $\text{vol}(O(n)) = \frac{2^n}{\prod_{j=1}^n \Gamma(\frac{j}{2})} \pi^{\frac{n(n+1)}{4}}$, $\text{vol}(U(n)) = \frac{2^n}{\prod_{j=1}^n \Gamma(j)} \pi^{\frac{n(n+1)}{2}}$,
 $\text{vol}(Sp(n)) = \frac{2^n}{\prod_{j=1}^n \Gamma(2j)} \pi^{n(n+1)}$.

推论 3.11. $\text{vol}(V_{n,p}) = \frac{2^p}{\prod_{j=1}^p \Gamma(\frac{n-j+1}{2})} \pi^{\frac{np-p(p-1)}{4}}$ (实) = $\frac{2^p}{\prod_{j=1}^p \Gamma(n-j+1)} \pi^{np-\frac{p(p-1)}{2}}$ (复),

$$\text{vol}(G_{n,p}) = \frac{\prod_{j=1}^n \Gamma(\frac{j}{2})}{\prod_{j=1}^p \Gamma(\frac{n-j+1}{2})} \pi^{\frac{(n-p)p}{2}}$$
 (实) = $\frac{\Gamma(j)}{\prod_{j=1}^p \Gamma(n-j+1)} \pi^{(n-p)p}$ (复).

3.3 计算若干积分

多元伽马函数, 记作 $\Gamma_p(a)$, 通过下式定义

$$\Gamma_n(a) = \int_{A>0} \text{etr}(-A) |A|^{a-\frac{n+1}{2}} (dA), \quad (3.7)$$

其中 a 的实部, 记作 $\Re(a) > \frac{1}{2}(n-1)$, 且积分过所有 $n \times n$ 实对称正定矩阵.

另一个与其紧密相关的函数称作多元贝塔函数, 记作 $B_n(a, b)$, 定义为

$$B_n(a, b) = \int_{0 < X < I} |X|^{a-\frac{n+1}{2}} |I-X|^{b-\frac{n+1}{2}} (dX), \quad (3.8)$$

其中 $\Re(a), \Re(b) > \frac{1}{2}(n-1)$, 其积分过所有使得 $X, I-X$ 正定的 $n \times n$ 实对称矩阵 X . 它与多元伽马函数通过如下等式联系

$$B_n(a, b) = \frac{\Gamma_n(a)\Gamma_n(b)}{\Gamma_n(a+b)}. \quad (3.9)$$

首先, 我们需要说明单矩阵参量超几何函数可由赫兹^[80]中的递归关系给出. 1963年, 康斯坦丁^[16]通过引入带多项式求解了超几何函数所遵循的方程组, 这使得利用超几何函数的实际计算变得可行.

定义 3.1. 设 X 是 $n \times n$ 实对称正定矩阵. 那么

$${}_pF_q(a_1, \dots, a_p; b_1, \dots, b_q; X) = \sum_{k=0}^{\infty} \sum_{\kappa \vdash k} \frac{1}{k!} \frac{(a_1)^{\kappa} \dots (a_p)^{\kappa}}{(b_1)^{\kappa} \dots (b_q)^{\kappa}} C_{\kappa}(X). \quad (3.10)$$

一些特殊的情形包括

$${}_0F_0(X) = \text{etr}(X) = \sum_{k=0}^{\infty} \sum_{\kappa \vdash k} \frac{C_{\kappa}(X)}{k!}, \quad (3.11)$$

$${}_1F_0(a; X) = |I - X|^{-a} = \sum_{k=0}^{\infty} \sum_{\kappa \vdash k} (a)^{\kappa} \frac{C_{\kappa}(X)}{k!}. \quad (3.12)$$

此外, 双矩阵参量的超几何函数也可以由带多项式来定义.

定义 3.2. 设 X, Y 分别是 $n \times n, m \times m$ 实对称正定矩阵 ($n \geq m$). 那么

$${}_pF_q(a_1, \dots, a_p; b_1, \dots, b_q; X, Y) = \sum_{k=0}^{\infty} \sum_{\kappa \vdash k} \frac{1}{k!} \frac{(a_1)^{\kappa} \dots (a_p)^{\kappa}}{(b_1)^{\kappa} \dots (b_q)^{\kappa}} \frac{C_{\kappa}(X)C_{\kappa}(Y)}{C_{\kappa}(I_n)}. \quad (3.13)$$

值得一提的是, 当两个矩阵参量具有相同的大小时, 如下等式成立

$$\begin{aligned} & \int_{O(n)} {}_pF_q(a_1, \dots, a_p; b_1, \dots, b_q; XHYH')[dH] \\ &= {}_pF_q(a_1, \dots, a_p; b_1, \dots, b_q; X, Y), \end{aligned} \quad (3.14)$$

其中 X, Y 是 $n \times n$ 对称矩阵, 积分过所有 $n \times n$ 正交矩阵 $O(n)$, 归一化哈尔测度定义为 $[dH] = \frac{1}{\text{vol}(O(n))}(dH) = \frac{\Gamma_p(p/2)}{2^p \pi^{p/2}}(dH)$. (推论3.10)

如果两个矩阵大小不同, 类似的公式由如下引理给出.

引理 3.12 (清水康希^[20]). 令 A 是 $n \times n$, B 是 $p \times p$ 实对称矩阵. 如果 $n \geq p$,

$$\int_{O(n)} {}_0F_0(AH_1BH_1')[dH] = {}_0F_0(A, B),$$

$$H = (H_1, H_2), \text{ 其中 } H_1 \text{ 为 } n \times p,$$

其中积分过所有 $n \times n$ 矩阵构成的正交群 $O(n) = \{H : H'H = I_n\}$.

下面这条引理将用于导出非中心分布.

引理 3.13 (詹姆斯^[14]). 令 X 是 $n \times p$ 实矩阵.

$$\int_{O(n)} \text{etr}(XH'_1)(dH) = {}_0F_1\left(\frac{1}{2}n; \frac{1}{4}X'X\right),$$

$H = (H_1, H_2)$, 其中 H_1 为 $n \times p$,

其中积分过类似引理3.12中的正交群 $O(n)$.

下面这条引理用于证明主要定理4.4.

引理 3.14 (卡特里^[43]). 令 A 是 $n \times n$ 实对称矩阵, B 是 $p \times p$ 实对称正定矩阵, 且 $n > p - 1$. 对任意固定的 $p \times p$ 实对称正定矩阵 S , 我们有

$$\int_{X'X=S} \text{etr}(AXBX')(dX) = \frac{\pi^{\frac{np}{2}}}{\Gamma_p(\frac{n}{2})} |S|^{\frac{n-p-1}{2}} {}_0F_0(A, BS).$$

下面这条引理用于计算拉普拉斯变换.

引理 3.15 (康斯坦丁^[16]). 令 Y, Z 是 $p \times p$ 对称矩阵, $\kappa \vdash k$ 且 $\Re(Z) > 0$. 那么

$$\int_{X>0} \text{etr}(-XZ)|X|^{a-\frac{n+1}{2}} C_\kappa(XY)(dX) = (a)^\kappa \Gamma_n(a) |Z|^{-a} C_\kappa(YZ^{-1}).$$

其中 $\Re(a) > \frac{1}{2}(p-1)$.

下面这条引理用于计算最大最小特征值的分布.

引理 3.16 (戴维斯^[81-82]). 设 $R > 0$ 是 $m \times m$ 矩阵. 有关 ${}_pF_q$ 的不完备伽马积分

$$\begin{aligned} \int_0^R \text{etr}(-AS)|S|^{c-\frac{m+1}{2}} {}_pF_q(a_1, \dots, a_p; b_1, \dots, b_q; BS)(dS) &= \frac{|R|^c}{B_m(c, \frac{m+1}{2})} \\ &\times {}_{p+1}F_{q+1}\left(c, a_1, \dots, a_p; c + \frac{1}{2}(m+1), b_1, \dots, b_q; -AR, BR\right), \end{aligned}$$

如果 R 满足 $0 < R < I$, 则有关 ${}_pF_q$ 的不完备贝塔积分

$$\begin{aligned} \int_0^R |S|^{c-\frac{m+1}{2}} |I-S|^{d-\frac{m+1}{2}} {}_pF_q(a_1, \dots, a_p; b_1, \dots, b_q; AS)(dS) &= \frac{|R|^d}{B_m(d, \frac{m+1}{2})} \\ &\times {}_{p+2}F_{q+1}\left(c, -d + \frac{m+1}{2}, a_1, \dots, a_p; c + \frac{1}{2}(m+1), b_1, \dots, b_q; A, AR\right), \end{aligned}$$

其中 $\Re(c), \Re(d) > \frac{1}{2}(m-1)$.

第4章 正态矩阵与抽样分布

近些年来,许多作者将独立同分布的多元正态总体推广到一般矩阵总体,比如达维德^[83],科比^[84],古德尔^[85]讨论的矩阵正态分布,具有密度函数

$$F_{n,p}(X) = \frac{1}{(2\pi)^{\frac{np}{2}} |\Phi|^{\frac{p}{2}} |\Psi|^{\frac{n}{2}}} \text{etr} \left[-\frac{1}{2} \Phi^{-1}(X - M)\Psi^{-1}(X - M)' \right], \quad (4.1)$$

其中 Φ 和 Ψ 分别是 $n \times n$ 和 $p \times p$ 实对称正定矩阵, M 是 $n \times p$ 固定矩阵, 这与文献^[53]中的向量椭球等高分布相一致. 然而, (4.1)式是用同质方差来定义的, 即总体不同列之间具有相同的协方差, 通过适当的矩阵变换 $Y = \Phi^{-\frac{1}{2}} X$, 其形式等价于独立样本. 这是^[53]中向量椭球等高分布的一个特例. 而椭球等高分布作为近代多元统计分析发展的核心概念, 其谱分解在^[53]中被称作是一个未解决的问题. 本文建立了^[53]中的椭球分类与总体精度矩阵的四类张量形式 $T_1, T_{1\frac{1}{2}}, T_2$ 和 T_3 之间的一一对应关系. 后者直接假设精度矩阵的谱结构, 对应于四类充分且必要的随机表示, 从而解决了这一困难, 其中(4.1)式作为特例 T_3 出现.

在本书中, 我们用 $A = (a_1, a_2, \dots, a_p)$ 代表矩阵 A 按列排列, 用 $A = [a_1; a_2; \dots; a_n]$ 代表矩阵 A 按行排列. 无论向量还是矩阵, 一律不加粗. 以下定义均为实矩阵.

4.1 矩阵正态分布

定义 4.1 (矩阵正态分布). $n \times p$ 矩阵总体 X 称作是正态矩阵, 如果 $\text{vec}(X')$ 是 np 维正态向量.

定理 4.1. 考虑具有一般正态条目 $n \times p$ 正态矩阵总体 X , 其关于 $\mathbb{R}^{n \times p}$ 上的勒贝格测度具有概率密度函数 $p(X)$. 记 X 的精度矩阵为 $\Theta_i \in T_i$, $i = 1, 1\frac{1}{2}, 2, 3$.

$$p(X) = \frac{|\Theta_1|^{\frac{1}{2}}}{(2\pi)^{\frac{np}{2}}} \text{etr} \left(-\frac{1}{2} \sum_{j=1}^p A_{jj} X B_{jj} X' - \sum_{j=1}^p \sum_{j'=j+1}^p A_{jj'} X B_{jj'} X' \right), \quad (T_1)$$

$$p(X) = \frac{|\Theta_{1\frac{1}{2}}|^{\frac{1}{2}}}{(2\pi)^{\frac{np}{2}}} \text{etr} \left(-\frac{1}{2} \sum_{j=1}^p A_{jj} X B_{jj} X' \right), \quad (T_{1\frac{1}{2}})$$

$$p(X) = \frac{|\Theta_2|^{\frac{1}{2}}}{(2\pi)^{\frac{np}{2}}} \text{etr} \left(-\frac{1}{2} \sum_{i=1}^n \sum_{j=1}^p \gamma_{ij} A_i X B_j X' \right), \quad (T_2)$$

$$p(X) = \frac{1}{(2\pi)^{\frac{np}{2}} |\Phi|^{\frac{p}{2}} |\Psi|^{\frac{n}{2}}} \text{etr} \left(-\frac{1}{2} \Phi^{-1} X \Psi^{-1} X' \right). \quad (T_3)$$

定理4.1的证明. 由 $\text{tr}(AXBX') = \text{vec}(X')'(A \otimes B) \text{vec}(X')$ 和定义直接得到. \square

定理 4.2. $T_1 \supset T_{1\frac{1}{2}} \supset T_2 \supset T_3$. 特别地, T_3 即(4.1).

定理4.2的证明. 表2.1的直接推论. \square

定理 4.3. 考虑 $n \times p$ 矩阵正态总体 $X \sim T_i$, $i = 1, 1\frac{1}{2}, 2, 3$. $X'X$ 以概率 1 是正定的当且仅当 $n > p - 1$.

定理4.3的证明. 以下思路来自迪克斯特拉^[86]. 只需要证明 $X'X$ 是以概率 1 非退化的当且仅当 $n > p - 1$. 令 x_1, \dots, x_p 是 X 的列向量, 而 x_1, \dots, x_n 是 X 的行向量.

$$\mathbf{P}(x_1, \dots, x_p \text{ 线性相关})$$

$$\begin{aligned} &\leq \sum_{i=1}^p \mathbf{P}(x_i \text{ 是 } x_1, \dots, x_{i-1}, x_{i+1}, x_p \text{ 的线性组合}) \\ &= \sum_{i=1}^p \mathbf{P}\left(\text{存在 } b_1, \dots, b_{i-1}, b_{i+1}, b_p \text{ 使得 } x_i = \sum_{i' \neq i} b_{i'} x_{i'}\right) \\ &\leq \sum_{j=1}^p \min_{i=1,2,\dots,n} \mathbf{P}\left(\text{存在 } b_1, \dots, b_{j-1}, b_{j+1}, b_p \text{ 使得 } x_{ij} = \sum_{j' \neq j} b_{j'} x_{ij'}\right) \\ &\leq p \cdot \min_{i=1,2,\dots,n} \mathbf{P}(\text{存在列向量 } b \text{ 使得 } x_i b' = 0) = p \cdot 0 = 0 \end{aligned}$$

其中在最后一行中, 我们用到了事实 x_i 具有独立分量, 从而 p 维正态分布落在 $p - 1$ 维子空间中的概率为零. \square

4.2 积矩分布

卡特里^[43], 坎茨, 约翰逊和鲍伊德^[44], 马泰^[48], 还有卡罗洛佩拉, 冈萨雷斯-法里亚斯和巴拉克里希南^[51]分别研究了(4.1)式对应的矩阵正态总体的样本方差与样本协方差的分布, 进而得到了相应的积矩阵分布. 通过引入矩阵参量广义超几何函数 ${}_pF_q$, 卡特里^[43]于 1966 年严格计算了来自总体 $N_{n,p}(0, \Phi, \Psi)$ 的同质二次型 $S = X'X$ 的联合概率密度,

$$V_{n,p}(S) = \frac{\text{etr}(-q^{-1}\Psi^{-1}S) |S|^{\frac{n-p-1}{2}}}{2^{\frac{np}{2}} \Gamma_p(\frac{n}{2}) |\Phi|^{\frac{p}{2}} |\Psi|^{\frac{n}{2}}} {}_0F_0\left(I - \frac{1}{2}q\Phi^{-1}, q^{-1}\Psi^{-1}S\right), \quad (4.2)$$

这里的 q 是任意正常数. 当 $\Phi = I$ 时, (4.2)式就退化成(1.1)式 ($M = 0$). 尽管卡塔里得到了中心情形 $N_{n,p}(0, \Phi, \Psi)$ 的联合密度, 但关于非中心总体 $N_{n,p}(M, \Phi, \Psi)$, $M \neq 0$ 的密度, 相关结果仍非常匮乏.

另一方面, (4.1)式是用同质方差来定义的, 即总体不同列之间具有相同的协方差, 通过适当的矩阵变换 $Y = \Phi^{-\frac{1}{2}}X$ 后, 其形式等价于独立样本. 马泰^[48]曾利用矩母函数的累积量来考察形如 $y_1 A_{11} y_1, y_2 A_{22} y_2$ 和 $y_1 A_{12} y_2$ 的方差异质二次型的独立性, 其中 $y_i, i = 1, 2$ 是 n 维独立正态向量, 且 $A'_{11} = A_{11}, A'_{22} = A_{22}$ 是实对称矩阵. 然而, 马泰的结果仅考虑了两样本的情况, 当变量数目 > 2 时, 有关结果几乎空白. 为了研究马泰定义的异质二次型的联合密度, 我们同时考虑 $\frac{1}{2}p(p+1)$ 个二次型的联合分布.

定义 4.2 (积矩分布). $(X + M)'(X + M)$ 称作是积矩, 其中 $X \sim T_i, i = 1, 1\frac{1}{2}, 2, 3$ 是正态矩阵, M 是固定矩阵. 特别地, 当 $X \sim N_{n,p}(M, I_n, \Psi)$, $X'X$ 服从非中心威沙特分布(1.1).

本节的主要目标是证明积矩分布具有类似威沙特分布(1.1)的密度.

定理 4.4. 设 $X \in T_1$ 的矩形坐标为 $B = (b_1, b_2, \dots, b_p)$. 假设 $p \times p$ 实对称矩阵 $U = (u_{ij})$ 满足 $u_{ij} = \text{tr}(A_{ij})$ 且其对角线上的元素皆为正, 即 $u_{ii} > 0 (i = 1, 2, \dots, p)$. 那么当 $n > p - 1$ 时, $S = X'X$ 的概率密度函数仅依赖于 $T = (t_{ij}), t_{ij} = \text{tr}(B_{ij}S)$,

$$\begin{aligned} & \frac{|\Theta_1|^{\frac{1}{2}}}{2^{\frac{np}{2}} \Gamma_p(\frac{n}{2})} \text{etr}\left(-\frac{1}{2}UT\right) |T|^{\frac{n-p-1}{2}}, \text{ 当 } M = 0; \\ & \times \text{etr}\left(-\frac{1}{2}\Omega\right) {}_0F_1\left(\frac{n}{2}; \frac{1}{4}\Delta T\right), \text{ 当 } M \neq 0; \end{aligned} \quad (4.3)$$

其中 $\Omega = \sum_{i,j=1}^p B'_{ij} M' A_{ij} M$ 和 $\Delta = \sum_{i,j,k,l=1}^p B'_{ij} M' A'_{ij} A_{kl} M B_{kl}$.

推论 4.5. (4.2)式对任意 $n > p - 1$ 成立.

注 4.1. 定理4.4的中心部分 ($M = 0$) 推广了卡特里的结果 (推论4.5), 即(4.2)式至形如(1.3)式的 $\frac{1}{2}p(p+1)$ 个正态二次型.

注 4.2. 定理4.4的中心部分 ($M = 0$) 与(4.2)式实际上是等价的. 通过积分, $\frac{1}{2}p(p+1)$ 个二次型 $y_i A_{ij} y_j (i \leq j)$, 其中 $A_{ij} = A'_{ji}$ 可以化简为(4.2)式. 反过来, 如果定理4.4成立, 那么(4.2)式对任意 $n > p - 1$ 成立.

定理4.4的证明组织如下. 首先, 我们将根据中心部分和推论4.5的等价性来证明定理4.4的非中心部分, 然后再证明等价性. 由于推论4.5借助于命题4.3是已知的, 因此定理4.4的证明就完成了.

定理4.4的证明. 首先假设定理4.4的中心部分对任意 $n > p - 1$ 成立. 对矩阵分解 $X = HZ$, 其中 Z 是上三角矩阵且 $H'H = I_p$. 由定理2.4,

$$dX = 2^{-p} |Z'Z|^{\frac{n-p-1}{2}} d(Z'Z) \cdot (dK), \text{ 其中 } K = (H, H_{\perp}) \in O(n).$$

因此, (4.1)式能够重写为

$$\begin{aligned} & \text{etr}\left(-\frac{1}{2}\Omega\right) \int_{O(n)} \text{etr}\left(-\frac{1}{2} \sum_{i,j=1}^p (A_{ij} M B_{ij}) Z' H'\right) (dK) \\ &= \text{etr}\left(-\frac{1}{2}\Omega\right) {}_0F_1\left(\frac{n}{2}; \frac{1}{4} \sum_{i,j,k,l=1}^p (B'_{ij} M' A'_{ij} A_{kl} M B_{kl}) Z' Z\right). \end{aligned}$$

因此, 如果中心部分成立, 则根据引理3.13定理4.4成立.

但是注意到, 根据注记4.2, 中心部分也可以由推论4.5和命题4.3得到, 只需要对 $Y = X B^{-1} = (y_1, y_2 \dots, y_p)$ 的 $\frac{1}{2}p(p+1)$ 个二次型,

$$\begin{aligned} & y'_1 A_{11} y_1, y'_1 A_{12} y_2, \dots, y'_1 A_{1p} y_p, \\ & y'_2 A_{22} y_2, \dots, y'_2 A_{2p} y_p, \\ & \vdots \\ & y'_p A_{pp} y_p, \end{aligned}$$

逐项积分, 其中 $y'_i A_{ij} y_j$ 对概率密度函数的贡献为

$${}_0F_0\left(I - \frac{q_{ij}}{2} A_{ij}, q_{ij}^{-1} y_i y'_j\right) = {}_0F_0\left(I - \frac{q_{ij}}{2} A_{ij}, q_{ij}^{-1} y'_j y_i\right).$$

这里我们发现 $y'_j y_i = t_{ij} = \text{tr}(B_{ij} S)$, 于是根据以下超几何函数的性质

$$\begin{aligned} {}_0F_0(X, cY) &= {}_0F_0(cX, Y) \\ {}_0F_0(X, I) &= {}_0F_0(X) = \text{etr}(X), \end{aligned}$$

所有包含 q_{ij} 的项都互相抵消. 这便完成了定理4.4的证明. \square

推论4.5的证明. 令 $A_{ij} = \beta_j \delta_{ij} \Phi^{-1}$ 使得 $\Psi^{-1} = \sum_{j=1}^p \beta_j b_j b'_j$, 其中 β_j, b_j 是 Ψ^{-1} 的特征根与特征向量. 那么 $u_{ij} = \beta_j \delta_{ij} \text{tr}(\Phi^{-1})$ 以及

$$\begin{aligned} \text{etr}\left(-\frac{1}{2}UT\right) &= \text{etr}\left(-\frac{1}{2} \text{tr}(\Phi^{-1}) \sum_{j=1}^p \beta_j t_{jj}\right) = \text{etr}\left(-\frac{1}{2} \Phi^{-1} \text{tr}(\Psi^{-1} S)\right) \\ &= \text{etr}\left(-q^{-1} \Psi^{-1} S\right) {}_0F_0\left(I - \frac{q}{2} \Phi^{-1}, q^{-1} \Psi^{-1} S\right), \end{aligned}$$

其中在证明第二个等式时我们用到了定义 $T = B S B'$. \square

4.2.1 积矩系数的矩母函数

定理4.6. 假设同定理4.4. $S = (s_{ij})$ 的矩母函数为

$$\begin{aligned} E \exp \left(\sum_{i \leq j} \gamma_{ij} s_{ij} \right) &= |\Theta_1|^{\frac{1}{2}} |U|^{-\frac{n}{2}} {}_1F_0 \left(\frac{n}{2}; W \right), \text{ 当 } M = 0; \\ &\times \text{etr} \left(-\frac{1}{2} \Omega \right) \text{etr} \left(\frac{1}{2} \Delta U^{-1} (I - W)^{-1} \right), \text{ 当 } M \neq 0; \end{aligned} \quad (4.4)$$

其中 $W = U^{-\frac{1}{2}} B R B' U^{-\frac{1}{2}}$, $2R = \Gamma + I$, 且 $\Gamma = (\gamma_{ij})$ 对称.

定理4.6的证明. 假设定理4.4对任意 $n > p - 1$ 成立. 为了证明定理4.6对任意 $n > p - 1$ 成立, 考察变换 $S \mapsto T = B' S B$, 此时矩母函数应为

$$\begin{aligned} &\frac{|\Theta_1|^{\frac{1}{2}}}{2^{\frac{np}{2}} \Gamma_p(\frac{n}{2})} \int_{T>0} \text{etr} \left(-\frac{1}{2} (U - B R B') T \right) |T|^{\frac{n-p-1}{2}} \\ &\quad \times \text{etr} \left(-\frac{1}{2} \Omega \right) {}_0F_1 \left(\frac{n}{2}; \frac{1}{4} \Delta T \right) (dT). \end{aligned}$$

根据 ${}_0F_1$ 的定义和引理3.15, 上述积分化简为

$$|\Theta_1|^{\frac{1}{2}} |U - B R B'|^{-\frac{n}{2}} \text{etr} \left(-\frac{1}{2} \Omega \right) \text{etr} \left(\frac{1}{2} \Delta (U - B R B')^{-1} \right).$$

令 $W = U^{-\frac{1}{2}} B R B' U^{-\frac{1}{2}}$ 并注意到 ${}_1F_0(a; W) = |I - W|^{-a}$, 即得(4.4)式. \square

推论4.7. (4.2)式的矩母函数为

$$|\Phi|^{-\frac{p}{2}} |W|^{-\frac{n}{2}} {}_1F_0 \left(\frac{n}{2}; qI - \frac{1}{2} \Phi^{-1}, W^{-1} \right),$$

其中 $W = I - q\Psi^{\frac{1}{2}} R \Psi^{\frac{1}{2}}$, $2R = \Gamma + I$, 且 $\Gamma = (\gamma_{ij})$ 对称.

证明见卡特里^[43]. 因此, 定理4.6以更紧凑的形式概括了正态二次形式的经典结果, 其中 ${}_1F_0$ 项类似于威沙特矩母函数. 关于确定矩母函数中指数 n 的容许范围留在本节末尾讨论.

4.2.2 积矩系数的特征根分布

定理4.8. 假设同定理4.4. S 的特征根 l_1, l_2, \dots, l_p 的联合分布为

$$\begin{aligned} &\frac{\pi^{\frac{p^2}{2}} |\Theta_1|^{\frac{1}{2}}}{2^{\frac{np}{2}} \Gamma_p(\frac{n}{2}) \Gamma_p(\frac{p}{2})} \prod_{i < j}^p (l_i - l_j) \prod_{i=1}^p (l_i)^{\frac{n-p-1}{2}} {}_0F_0 \left(-\frac{1}{2} U, L \right), \text{ 当 } M = 0; \\ &\quad \times \text{etr} \left(-\frac{1}{2} \Omega \right) {}_0F_1 \left(\frac{n}{2}; \frac{1}{4} \Delta, L \right), \text{ 当 } M \neq 0; \end{aligned}$$

其中 $L = \text{diag}(l_i)$, $l_1 > l_2 > \dots > l_p > 0$; 别处为零.

关于特征根的联合分布, 存在以下重要结论^[7]定理3.2.17.

命题4.9. 设 S 是 $p \times p$ 实对称正定矩阵, 密度为 $f(S)$. 则 S 的特征根 l_1, l_2, \dots, l_p 的联合分布为

$$\frac{\pi^{p^2/2}}{\Gamma_p(\frac{p}{2})} \prod_{i < j}^p (l_i - l_j) \int_{O(p)} f(H L H')(dH),$$

其中 $L = \text{diag}(l_1, \dots, l_p)$, $l_1 > l_2 > \dots > l_p > 0$; 别处为零.

定理4.8的证明. 根据命题4.3和定理4.4, 分解矩阵 $X = HZ$, 其中 Z 是上三角矩阵且 $H'H = I_p$, 将 H 扩充为 $K = (H, H_\perp) \in O(n)$, 我们有

$$\begin{aligned} & \text{etr}\left(-\frac{1}{2}\Omega\right) \int_{O(p)} \text{etr}\left(-\frac{1}{2} \sum_{i,j=1}^p (A_{ij} M B_{ij}) Z' H'\right) (dK) \\ &= \text{etr}\left(-\frac{1}{2}\Omega\right) {}_0F_1\left(\frac{n}{2}; \frac{1}{4} \sum_{i,j,k,l=1}^p (B'_{ij} M' A'_{ij} A_{kl} M B_{kl}) Z' Z\right). \end{aligned}$$

非中心部分的证明于是由引理3.13得到. 而中心部分则是引理3.12的直接推论.

□

推论4.10. (4.2)式的特征根 l_1, l_2, \dots, l_p 的联合分布为

$$\frac{\pi^{\frac{p^2}{2}}}{2^{\frac{np}{2}} \Gamma_p(\frac{n}{2}) \Gamma_p(\frac{p}{2}) |\Phi|^{\frac{p}{2}} |\Psi|^{\frac{n}{2}}} \prod_{i=1}^p (l_i)^{\frac{n-p-1}{2}} \prod_{i < j}^p (l_i - l_j) {}_0F_0\left(-\frac{1}{2}\Phi^{-1}, \Psi^{-1}L\right)$$

其中 $L = \text{diag}(l_i)$, $l_1 > l_2 > \dots > l_p > 0$; 别处为零.

推论4.10的证明. 根据前述引理, 引入 $K = [H_1, H_2] \in O(n)$, 其中 H_1 是满足 $XH = H_1Q$ 的 $n \times p$ 矩阵, 其中 Q 上三角. 根据引理3.13, 我们有

$$\begin{aligned} p(L) &\propto \int_{O(n)} (dK) \int_{Q'Q=L} \text{etr}\left(-\frac{1}{2}\Phi^{-1} H_1 Q \Psi^{-1} Q' H_1'\right) (dQ) \\ &\propto \int_{O(n)} {}_0F_0\left(-\frac{1}{2} H_1' \Phi^{-1} H_1, \Psi^{-1} L\right) (dK). \end{aligned} \tag{4.5}$$

那么乘上首项系数和范德蒙德行列式 $\prod_{i < j}^p (l_i - l_j)$ 之后即得推论4.10. □

为了说明这与(1.1)式的积分仅当 $\Phi = I$ 时等价, 考虑(4.1)式关于 $X'X = HLH'$ 和 H 的积分. 事实上, 通过引入 $Z = XH$, 根据引理3.14,

$$\begin{aligned} p(L) &\propto \int_{O(p)} (dH) \int_{Z'Z=L} \text{etr}\left(-\frac{1}{2}\Phi^{-1} Z H \Psi^{-1} H' Z'\right) (dZ) \\ &\propto |L|^{\frac{n-p-1}{2}} \int_{O(p)} {}_0F_0\left(-\frac{1}{2}\Phi^{-1}, H \Psi^{-1} H' L\right) (dH). \end{aligned}$$

比较首项系数, 这简化为詹姆斯^[15]关于威沙特分布特征根分布的经典结果,

$$\frac{\pi^{\frac{p^2}{2}}}{2^{\frac{np}{2}} \Gamma_p(\frac{n}{2}) \Gamma_p(\frac{p}{2}) |\Psi|^{\frac{n}{2}}} \prod_{i=1}^p (l_i)^{\frac{n-p-1}{2}} \prod_{i < j}^p (l_i - l_j) {}_0F_0\left(-\frac{1}{2}\Psi^{-1}, L\right), \quad (4.6)$$

其中 $L = \text{diag}(l_1, l_2, \dots, l_p)$, $l_1 > l_2 > \dots > l_p$; 别处为零. \square

4.2.3 $p = 3$ 情形时的例子

以下三个例子均是定理4.4在 $p = 3$ 情形的特例.

例 4.1 (X 具有独立行). 设采样总体的概率密度为

$$p(X) = \frac{\prod_{r=1}^n |B_r|^{\frac{1}{2}}}{(2\pi)^{\frac{3n}{2}}} \exp\left(-\frac{1}{2} \sum_{r=1}^n \mathbf{x}'_r B_r \mathbf{x}_r\right). \quad (4.7)$$

其中 B_r 是 $\mathbf{x}_r = (x_{r1}, x_{r2}, x_{r3})$ 的 3×3 协方差矩阵的逆, $X = [\mathbf{x}_1; \mathbf{x}_2; \dots; \mathbf{x}_n]$ 是具有独立正态分布行的矩阵.

考虑从样本中得到的如下统计量

$$\begin{aligned} n\bar{x}_1 &= \sum_{r=1}^n x_{r1}, \quad n\bar{x}_2 = \sum_{r=1}^n x_{r2}, \quad n\bar{x}_3 = \sum_{r=1}^n x_{r3}, \\ ns_1^2 &= \sum_{r=1}^n (x_{r1} - \bar{x}_1)^2, \quad ns_2^2 = \sum_{r=1}^n (x_{r2} - \bar{x}_2)^2, \quad ns_3^2 = \sum_{r=1}^n (x_{r3} - \bar{x}_3)^2, \\ nr_{ij}s_i s_j &= \sum_{r=1}^n (x_{ri} - \bar{x}_i)(x_{rj} - \bar{x}_j), \quad (i, j = 1, 2, 3). \end{aligned}$$

为了变换体积元为 $\bar{x}_1, \bar{x}_2, \bar{x}_3, s_1, s_2, s_3$, 根据^[76]中的已知结果, 我们通过引入三个辅助角 θ_{11}, θ_{12} 和 θ_{22} 来计算变换对体积元的贡献, 它们代表 3×3 正交群的生成元, 我们具有如下表达式

$$\begin{aligned} dp &= \frac{n^{\frac{3(n-2)}{2}} \prod_{r=1}^n |B_r|^{\frac{1}{2}}}{2^{\frac{3(n-1)}{2}} \pi^{\frac{3}{2}} \Gamma(\frac{n-1}{2}) \Gamma(\frac{n-2}{2}) \Gamma(\frac{n-3}{2})} \exp\left(-\frac{1}{2} \sum_{r=1}^n \mathbf{x}'_r B_r \mathbf{x}_r\right) s_1^{n-2} s_2^{n-2} s_3^{n-2} \\ &\quad \times \sin^{n-3} \theta_{11} \sin^{n-3} \theta_{12} \sin^{n-4} \theta_{22} d\bar{x}_1 d\bar{x}_2 d\bar{x}_3 ds_1 ds_2 ds_3 d\theta_{11} d\theta_{12} d\theta_{22}. \end{aligned} \quad (4.8)$$

通过对(4.8)中的 $\theta_{11}, \theta_{12}, \theta_{22}$ 积分, 我们得到了当 $B_1 = B_2 = \dots = B$ 时三个

样本方差和三个积矩系数联合分布的对称表达式

$$\begin{aligned}
 A &= |B_{11}|/|B|, \quad B = |B_{22}|/|B|, \quad C = |B_{33}|/|B|, \\
 F &= |B_{12}|/|B|, \quad G = |B_{13}|/|B|, \quad H = |B_{23}|/|B|, \\
 a &= s_1^2, \quad b = s_2^2, \quad c = s_3^2, \quad f = s_1 s_2 \cos \theta_{11}, \quad g = s_1 s_3 \cos \theta_{12}, \\
 h &= s_2 s_3 (\sin \theta_{11} \sin \theta_{12} \cos \theta_{22} + \cos \theta_{11} \cos \theta_{12}),
 \end{aligned} \tag{4.9}$$

$$dp = \frac{1}{\pi^{\frac{3}{2}} \Gamma(\frac{n-1}{2}) \Gamma(\frac{n-2}{2}) \Gamma(\frac{n-3}{2})} \cdot \left| \begin{array}{ccc} A & F & G \\ F & B & H \\ G & H & C \end{array} \right|^{\frac{n-1}{2}} \cdot \left| \begin{array}{ccc} a & f & g \\ f & b & h \\ g & h & c \end{array} \right|^{\frac{n-5}{2}} \times e^{-Aa-Bb-Cc-2Hh-2Gg-8Ff} dadbcdfdgdh,$$

其中 B_{ij} 是 B 的 (i, j) 余子式, 乘以公共的常数因子 $(n/2)$.

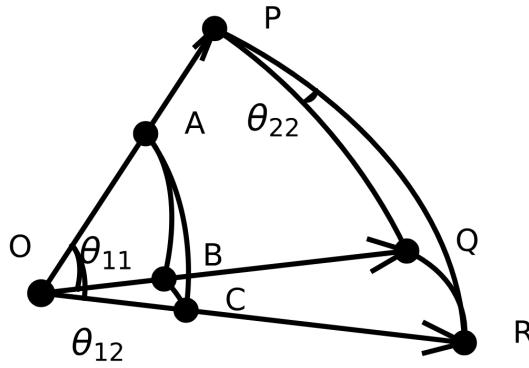


图 4.1 威沙特的几何推导中使用的辅助角 $\theta_{11}, \theta_{12}, \theta_{22}$.

事实上, 威沙特^[40]首先根据 1925 年之前费希尔, 皮尔逊和罗马诺夫斯基的著名结果推导得到了公式(4.9). 在这里, 我们利用津村善郎^[76]的矩阵分解技术来得出了完全相同的结果.

例 4.2 (X 具有独立列). 与(4.7)相比, 当总体具有独立列时积矩分布的计算有所不同. 例如, 新的偏相关系数和多元相关系数以及回归系数需要估计. 因此, 如果 x_1, x_2, x_3 是独立的列, 并且总体中的条目是联合正态分布的, 则总体的一个简化的形式可以写作

$$p(X) = \frac{|\Theta_{1\frac{1}{2}}|^{\frac{1}{2}}}{(2\pi)^{\frac{n}{2}}} \exp \left(-\frac{1}{2} x'_1 A_{11} x_1 - \frac{1}{2} x'_2 A_{22} x_2 - \frac{1}{2} x'_3 A_{33} x_3 \right), \tag{4.10}$$

其中 $\Theta_{1\frac{1}{2}} = A_{11} \otimes E_{11} + A_{22} \otimes E_{22} + A_{33} \otimes E_{33}$, 3×3 矩阵 E_{ij} 除开 (i, j) 元素为一外其余全为零; $n \times n$ 矩阵 A_{11} (类似的, A_{22}, A_{33}) 是变量 x_1 的协方差矩阵的逆 (类

似的, x_2, x_3). 因此, 通过引入任意常数 $q > 0$, (4.10)能够重新写作,

$$p(X) = \frac{|\Theta_{1\frac{1}{2}}|^{\frac{1}{2}}}{(2\pi)^{\frac{3n}{2}}} \exp \left[- \sum_{i=1}^3 q^{-1} x'_i x_i + \sum_{i=1}^3 q^{-1} x'_i \left(I - \frac{q}{2} A_{ii} \right) x_i \right]. \quad (4.11)$$

根据引理3.14, 我们能够积去指数中的讨厌参数 $q^{-1} x'_1 (I - \frac{q}{2} A_{11}) x_1$ (类似的, x_2, x_3),

$$\begin{aligned} \text{etr}(X) &= {}_0F_0(X), \\ \int_{O(n)} {}_0F_0(AH_1BH'_1)(dH) &= {}_0F_0(A, B), \\ H &= [H_1, H_2], \text{ 且 } H_1 \text{ 为 } n \times 3, \\ \int_{X'X=S} \text{etr}(AXBX')(dX) &= \frac{\pi^n}{\Gamma_3(\frac{n}{2})} |S|^{\frac{n-4}{2}} {}_0F_0(A, BS), \end{aligned} \quad (4.12)$$

这一结果, 通过积分 $q^{-1} X'(I - \frac{q}{2} A)X$ (假设 $A_{11} = A_{22} = A_{33} = A$), 具有简洁的矩阵表达式,

$$p(X'X) = \frac{|A|^{\frac{3}{2}}}{2^{\frac{3n}{2}} \Gamma_3(\frac{n}{2})} \text{etr}(-q^{-1} X'X) |X'X|^{\frac{n-3}{2}} {}_0F_0\left(I - \frac{q}{2} A, q^{-1} X'X\right).$$

如果 $A_{11} \neq A_{22}$ (或 $\neq A_{33}$), 我们仍令 $s_{11} = x'_1 x_1, s_{22} = x'_2 x_2, s_{33} = x'_3 x_3$, 其余 s_{12}, s_{13}, s_{23} 类似. 因此, 对任意 $q_{11}, q_{22}, q_{33} > 0$ 和 $S = (s_{ij})$, 结果变为

$$\begin{aligned} p(S) &= \frac{|\Theta_{1\frac{1}{2}}|^{\frac{1}{2}}}{2^{\frac{3n}{2}} \Gamma_3(\frac{n}{2})} \exp \left(- \sum_{i=1}^3 q_{ii}^{-1} s_{ii} \right) |S|^{\frac{n-4}{2}} \prod_{i=1}^3 {}_0F_0\left(I - \frac{q_{ii}}{2} A_{ii}, q_{ii}^{-1} s_{ii}\right) \\ &= \frac{|\Theta_{1\frac{1}{2}}|^{\frac{1}{2}}}{2^{\frac{3n}{2}} \Gamma_3(\frac{n}{2})} \text{etr}\left(-\frac{1}{2} A_{11}s_{11} - \frac{1}{2} A_{22}s_{22} - \frac{1}{2} A_{33}s_{33}\right) |S|^{\frac{n-4}{2}}, \end{aligned} \quad (4.13)$$

(4.13)式不该与(4.9)式混淆. 因此, 假设每列都是联合正态的, 并且假设总体密度为 (4.10), 可以显式计算出矩阵正态总体的积矩系数 s_{11}, s_{22}, s_{33} 的联合概率密度. 另外指出, 这一方法也适用于更一般的 T_1 分布.

例 4.3 (X 具有一般正态条目). 假设我们没有正态总体的先验知识, 只知道矩阵总体中的每个条目都是正态分布的. 在这种情况下, 我们需要做一些假设来减少未知参数数量. 例如, 假设我们的 $N = np$ 个样本值来自 N 个不同的正态总体, 则存在至少 N 个方差和 $\frac{1}{2}N(N-1)$ 个两两相关系数或回归系数需要估计, 因此未知参数数量为 $\frac{1}{2}N(N+1)$, 即近似 $O(N^2)$. 倘若不减少参数数量, 就不可能得出任何有意义的结论. 从表4.1中, 我们可以看到在四个矩阵正态总体 $T_1, T_{1\frac{1}{2}}, T_2$ 和 T_3 中自由度显著降低, 其等价于命题4.2所述的精度矩阵的四种嵌套类型的谱分解.

表 4.1 四类常见的矩阵正态分布 $T_1, T_{1\frac{1}{2}}, T_2, T_3$

类型	$-2 \log(\text{M.G.F.})$	D.O.F.	类别	M.G.F.
T_1	$\sum_{j=1}^p \sum_{j'=1}^p t'_j A_{jj'} t_{j'}$	$\frac{1}{2}n(n-1)p^2 + np$	LE	$\phi(t'_i A_{ij} t_j)$
$T_{1\frac{1}{2}}$	$\sum_{j=1}^p t'_j A_{jj} t_j$	$\frac{1}{2}n(n+1)p$	ME	$\phi(t'_i A_{ii} t_i)$
T_2	$\sum_{i=1}^n \sum_{j=1}^p \gamma_{ij} t'_j A_{ij} t_j$	np	ME	$\phi(t'_i A_{ii} t_i)$
T_3	$\sum_{i=1}^n \sum_{j=1}^p \alpha_i \beta_j t'_j A_{ii} t_j$	$n+p$	VE	$\phi(\sum t'_i A_{ii} t_i)$

其中 $T = (t_1, \dots, t_n)$ 和 $X = (x_1, \dots, x_n)$ 都是 $n \times p$ 矩阵, $A'_{ij} = A_{ji}$, 且矩母函数 M.G.F. 定义为 $E \exp(\sum t_{ri} x_{ri})$. 此处的记号 LE, ME, VE 与^[53]中略有不同.

虽然以上定理4.4, 4.6, 4.8只考虑了实值分布的情形, 但复值分布的结果也应当是类似的. 而且, $T_{1\frac{1}{2}}, T_2, T_3$ 情形的结果亦可作为推论. 特别地, 定理4.4的 T_3 情形推广了卡特里^[43]的已知结果, 并且未在中心情形引入双矩阵参数超几何函数 ${}_0F_0$ 和讨厌参数 $q_{ij} > 0$. 关于积矩分布, 我们提出以下三个问题供读者思考.

问题 4.11. 能否构造一个 3×3 左椭球分布并非 $T_{1\frac{1}{2}}$? [提示: 任意两个 2×2 对称矩阵都可交换, 所以考虑 3×3 矩阵.]

问题 4.12. 能否使用数学归纳法证明定理4.4? [提示: 结合引理3.14和文献^[87].]

问题 4.13. 能否证明(4.4)式 ($M = 0$) 定义了一个矩母函数当且仅当 $n \in \{0, 1, 2, \dots, p-1\} \cup (p-1, \infty)$ 而(4.4)式 ($M \neq 0$) 定义了一个矩母函数当且仅当额外地, $n \geq \max\{\text{rank}(\Omega), \text{rank}(\Delta)\}$ 当 $n < p-1$? [提示: 非平凡. 见^[88]问题 6.10.]

问题4.13的回答 (若真) 则退化成文献^[89]的猜想, 仅当 $A_{jj'} = \beta_j I \delta_{jj'}$ 或 $\Sigma = I \otimes \Psi$. 退化情形的问题4.13已经被许多作者解决, 比如^[90]中的递归方法和^[91]中的延拓方法. 我们想要得到的是问题4.13关于 T_1 及其类似情形的肯定的答案, 所以它仍是有价值和值得探讨的.

4.3 矩阵 t 分布

我们所熟知的单变元 t 分布最早由戈塞^[36]引入, 其非中心分布涉及合流超几何函数 ${}_1F_1$. 通常, 与 ${}_1F_1$ 有关的数值积分都较难求得显式解. 鉴于此问题的复杂性, 现有关于多元或矩阵 t 分布的定义都采用的是中心分布的比与固定均值的和, 以避免在分布中引入合流项 ${}_1F_1$, 这一点可见文献^[3, 54-55, 60-61]. 然而, 如果总体均值是一个未知参数, 则中心统计量往往难以显示构造. 因此, 采取恰当的定义是

有必要的. 在本文中, 为了与单变量非中心 t 分布的定义相吻合, 我们引入具有非中心矩阵正态分子的矩阵 t 分布如下.

定义 4.3. $m \times p$ 矩阵总体 Z 称作 t 矩阵, 如果它能够写作 $Z = (X + M)S^{-\frac{1}{2}}$ 的形式, 其中 M 是 $m \times p$ 固定矩阵, $X \in N_{m,p}(0, I_m, I_p)$ 是正态矩阵, $S \in T_1$ 是积矩.

引理 4.14. 考虑正定矩阵 $C (p \times p)$, 以及固定矩阵 $D (p \times p)$,

$$\begin{aligned} & \int_{X>0} \text{etr}(-CX'X + DX')|X'X|^{a-\frac{p+1}{2}}(dX) \\ &= \Gamma_p\left(a - \frac{1}{2}\right)|C|^{-(a-\frac{1}{2})}{}_1F_1\left(a - \frac{1}{2}; \frac{1}{2}p; \frac{1}{4}D'DC^{-1}\right) \end{aligned} \quad (4.14)$$

其中 $\Re(a) > \frac{1}{2}(p-1)$, 积分过所有 $p \times p$ 实对称正定矩阵.

引理 4.14 的证明. 令 $Y = XC^{\frac{1}{2}}$, $E = DC^{-\frac{1}{2}}$. 由极坐标分解, 任意 $n \times p$ 矩阵 Y 可以唯一地分解为 $Y = HR^{\frac{1}{2}}$, $H = Y(Y'Y)^{-\frac{1}{2}}$, $R = Y'Y$, 其中 $n \times p$ 角向分量 $H \in V_{n,p}$. 根据定理 2.3, $(dY) = 2^{-p}|R|^{-\frac{1}{2}}(dR)(dK)$. 因此, (4.14) 式变为

$$2^{-p}|C|^{-a+\frac{1}{2}} \int_{R>0} (dR) \int_{O(n)} \text{etr}(-R + H'EER^{\frac{1}{2}})|R|^{a-\frac{p}{2}-1}(dK).$$

由引理 3.13, 这一积分又化简为

$$2^{-p}|C|^{-a+\frac{1}{2}} \int_{R>0} |R|^{a-\frac{p}{2}-1} \text{etr}(-R) {}_0F_1\left(\frac{1}{2}n; \frac{1}{4}E'EER\right)(dR).$$

由引理 3.15, 上式即合流超几何函数 ${}_0F_1$. □

定理 4.15. 非中心 t 矩阵 Z 具有如下密度

$$\begin{aligned} & \frac{\Gamma_p\left(n + \frac{1}{2}m - \frac{1}{2}\right)}{(2)^{\frac{1}{2}(n+m)p} \pi^{\frac{1}{2}mp} \Gamma_p(\frac{n}{2})} |U + Z'Z|^{-n-\frac{m}{2}-\frac{1}{2}} \text{etr}\left(-\frac{1}{2}M'M\right) \\ & \times {}_1F_1\left(n + \frac{1}{2}m - \frac{1}{2}, \frac{1}{2}p; \frac{1}{4}M'ZZ'MB'(U + Z'Z)^{-1}B\right). \end{aligned} \quad (4.15)$$

定理 4.15 的证明. 根据独立性假设, S 和 X 的联合密度由下式给出

$$\frac{\pi^{-\frac{1}{2}mp}}{(2)^{\frac{1}{2}(n+m)p} \Gamma_p(\frac{n}{2})} \text{etr}\left(-\frac{1}{2}X'X - \frac{1}{2}UT\right) |T|^{\frac{n-p-1}{2}}.$$

令 $Z = (X + M)S^{-\frac{1}{2}}$, 则变换 $(X, S) \mapsto (Z, S)$ 的雅可比为 $|S|^{\frac{m}{2}}$. 用 X 替换 Z , 再乘以雅可比 $|S|^{\frac{m}{2}}$, 我们得到了 Z 与 S 的联合密度

$$\frac{\pi^{-\frac{1}{2}mp}}{(2)^{\frac{1}{2}(n+m)p} \Gamma_p(\frac{n}{2})} \text{etr}\left(-\frac{1}{2}M'M + S^{\frac{1}{2}}Z'M - \frac{1}{2}(U + Z'Z)BSB'\right) |S|^{\frac{n+m-p-1}{2}},$$

这里我们用到了定义 $S = B'TB$. 根据 4.14, 上式过所有 $p \times p$ 正定矩阵 $S^{\frac{1}{2}}$ 积分的结果正是(4.15). □

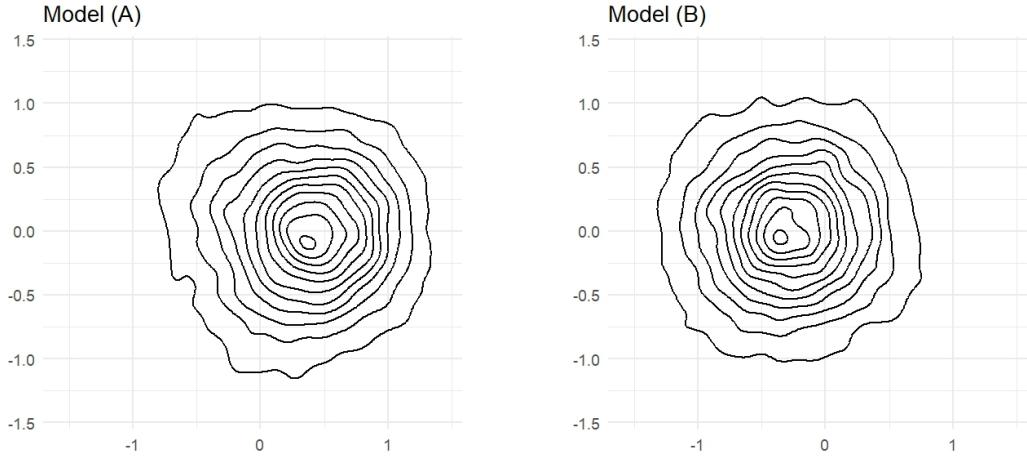


图 4.2 矩阵 t 分布的 2 维等高线图. (A) $Z = (X + M)S^{-\frac{1}{2}}$, (B) $Z = XS^{-\frac{1}{2}} + M$. 其中 $X \sim N_{2,2}(0, I_2, I_2)$, $S \sim W_{2,2}(0, I_2)$, 且 $M = (1.0, 0.2; -0.3, 0.5)$.

4.4 矩阵 F 分布

通过引入矩形坐标, 我们已经得到了积矩分布的密度, 并构造了对应的 t 统计量. 但 t 分布不是方阵, 涉及矩的积分也往往是发散的, 因而基于特征值的矩阵统计方法也就适用于 t 统计量. 于是我们将注意力转向两个独立积矩分布之比, 即 F 分布, 其最大最小特征根分布被广泛应用于方差分析. 这其中的工作包括杉山健一^[92], 卡特里^[62], 戴维斯^[63], 平川文子^[64]等. 较之单个积矩分布的最大最小特征根^[62,68-70], F 统计量能够同时考虑均值与方差的波动信息, 其分布也更易求得. 除此之外, 迹统计量^[93]与行列式统计量^[94]也都是一些典型的基于特征根的统计量. 我们将在本章的最后一节使用数值模拟分析这些统计量的功效.

定义 4.4. $n \times p$ 矩阵总体 Z 称作中心 F 矩阵, 如果它可以写成 $Z = S_1 S_2^{-1}$, 其中 $S_1 \in T_i$ 为非中心积矩, $S_2 \in T_{i'}$ 为另一中心积矩, 具有相同的矩形坐标 B . 如果 S_1 是非中心的但 S_2 是中心的, 且 S_1 与 S_2 具有相同的精度矩阵 $\Theta \in T_i$, 则 S 称作非中心 F 矩阵.

定理 4.16. 中心 F 矩阵 S 具有密度函数

$$p(S) = \frac{|\Theta_1|^{\frac{1}{2}} |\Theta_2|^{\frac{1}{2}} \Gamma_p(\frac{n_1+n_2+p+1}{2})}{2^{(n_1+n_2)p} B_p(\frac{n_1}{2}, \frac{n_2}{2}) \Gamma_p(\frac{p+1}{2}) |U_2|^{\frac{n_1+n_2}{2}}} |S|^{\frac{n_1-p-1}{2}} \\ \times {}_1F_1 \left(\frac{n_1+n_2}{2}; \frac{p+1}{2}; -\frac{1}{2}(U_1 U_2^{-1} + S) \right), \quad (4.16)$$

定理4.16的证明. 依定义, $S_1 = B'T_1B$ 和 $S_2 = B'T_2B$ 具有联合分布

$$\begin{aligned} p(S_1, S_2) &= \frac{|\Theta_1|^{\frac{1}{2}} |\Theta_2|^{\frac{1}{2}}}{2^{(n_1+n_2)p} \Gamma_p(\frac{n_1}{2}) \Gamma_p(\frac{n_2}{2})} \text{etr}\left(-\frac{1}{2} U_2 T_2\right) |T_2|^{\frac{n_2-p-1}{2}} \\ &\quad \times \text{etr}\left(-\frac{1}{2} U_1 T_1\right) |T_1|^{\frac{n_1-p-1}{2}}. \end{aligned} \quad (4.17)$$

矩阵变换 $(S_1, S_2) \mapsto (S_1 S_2^{-1}, S_2)$ 具有雅可比行列式 $|S_2|^{-(p+1)/2}$. 因此, S 与 S_2 根据(4.17)式的联合分布为

$$\begin{aligned} p(S, S_2) &= \frac{|\Theta_1|^{\frac{1}{2}} |\Theta_2|^{\frac{1}{2}}}{2^{(n_1+n_2)p} \Gamma_p(\frac{n_1}{2}) \Gamma_p(\frac{n_2}{2})} \text{etr}\left(-\frac{1}{2}(U_2 + U_1 S) T_2\right) \\ &\quad \times |S|^{\frac{n_1-p-1}{2}} |T_2|^{\frac{n_1+n_2-p-1}{2}}. \end{aligned} \quad (4.18)$$

由贝塔积分与超几何函数的性质

$$\begin{aligned} {}_1F_1(a; a; X) &= {}_0F_0(X) = \text{etr}(X), \\ \int_{S>0} |S|^{a-\frac{p+1}{2}} |I - S|^{b-\frac{p+1}{2}} {}_0F_0(RS) dS &= B(a, b)^{-1} {}_1F_1(a; b; R), \end{aligned}$$

在(4.18)式中积去 S_2 , 我们就得到了 S 的密度分布(4.16). \square

定理 4.17. 令 $S = S_1 S_2^{-1}$ 是非中心 F 矩阵.

1. $S(I + S)^{-1} = S_1(S_1 + S_2)^{-1}$ 的特征根 f_1, f_2, \dots, f_p 的联合分布为

$$\begin{aligned} &\frac{\pi^{p^2/2} |\Theta|}{2^{(n_1+n_2)p} B_p(\frac{n_1}{2}, \frac{n_2}{2}) \Gamma_p(\frac{p}{2}) |U|^{\frac{n_1+n_2}{2}}} \prod_{i<j}^p (f_i - f_j) |F|^{\frac{n_1}{2}} |I - F|^{\frac{n_1-p-1}{2}} \\ &\quad \times \text{etr}\left(-\frac{1}{2} \Omega\right) {}_1F_1\left(\frac{n_1+n_2}{2}; \frac{n_1}{2}; \Delta U^{-1}, F\right), \end{aligned} \quad (4.19)$$

其中 Ω, Δ 的定义类似定理4.4, $F = \text{diag}(f_1, f_2, \dots, f_p)$, $f_1 > f_2 > \dots > f_p$; 别处为零.

2. 最大特征根 f_1 的分布为

$$\begin{aligned} P(f_1 < x) &= \frac{|\Theta| x^{\frac{n_1 p}{2}}}{2^{(n_1+n_2)p} B_p(\frac{n_1}{2}, \frac{n_2}{2}) B_p(\frac{n_1}{2}, \frac{p+1}{2}) |U|^{\frac{n_1+n_2}{2}}} \\ &\quad \times \text{etr}\left(-\frac{1}{2} \Omega\right) {}_2F_1\left(a, b; c; \Delta U^{-1}, \Delta U^{-1} R\right). \end{aligned} \quad (4.20)$$

其中 $a = -\frac{1}{2}(n_2 - p - 1)$, $b = \frac{1}{2}(n_1 + n_2)$, $c = \frac{1}{2}(n_2 + p + 1)$.

3. 最小特征根 f_p 的分布为

$$1 - P(f_p \leq y) = P(f_p > y) = \frac{|\Theta|(1-y)^{\frac{n_2 p}{2}}}{2^{(n_1+n_2)p} B_p(\frac{n_1}{2}, \frac{n_2}{2}) B_p(\frac{n_2}{2}, \frac{p+1}{2}) |U|^{\frac{n_1+n_2}{2}}} \times \text{etr}\left(-\frac{1}{2}\Omega\right) {}_2F_1(a, b; c; \Delta U^{-1}, \Delta U^{-1} R). \quad (4.21)$$

其中 $a = -\frac{1}{2}(n_1 - p - 1)$, $b = \frac{1}{2}(n_1 + n_2)$, $c = \frac{1}{2}(n_1 + p + 1)$.

定理4.17的证明. 直接计算 $S_1 = B'T_1B$ 和 $S_2 = B'T_2B$ (模去非中心项 $\text{etr}\left(-\frac{1}{2}\Omega\right)$) 的联合分布

$$\begin{aligned} p(S_1, S_2) &= \frac{|\Theta|}{2^{(n_1+n_2)p} \Gamma_p(\frac{n_1}{2}) \Gamma_p(\frac{n_2}{2})} \text{etr}\left(-\frac{1}{2}UT_2\right) |T_2|^{\frac{n_2-p-1}{2}} \\ &\quad \times \text{etr}\left(-\frac{1}{2}UT_1\right) |T_1|^{\frac{n_1-p-1}{2}} {}_0F_1\left(\frac{n_1}{2}; \Delta T_1\right). \end{aligned} \quad (4.22)$$

雅可比变换 $(S_1, S_2) \mapsto (S_1 S_2^{-1}, S_2)$ 的行列式也是类似的,

$$\begin{aligned} p(S, S_2) &= \frac{|\Theta|}{2^{(n_1+n_2)p} \Gamma_p(\frac{n_1}{2}) \Gamma_p(\frac{n_2}{2})} \text{etr}\left(-\frac{1}{2}U(S+I)T_2\right) \\ &\quad \times |S|^{\frac{n_1-p-1}{2}} |T_2|^{\frac{n_1+n_2-p-1}{2}} {}_0F_1\left(\frac{n_1}{2}; \Delta ST_2\right). \end{aligned} \quad (4.23)$$

对(4.23)式积分我们有 S 的概率密度函数

$$\begin{aligned} p(S) &= \frac{|\Theta|}{2^{(n_1+n_2)p} B_p(\frac{n_1}{2}, \frac{n_2}{2}) |U|^{\frac{n_1+n_2}{2}}} |I+S|^{-\frac{n_1+n_2}{2}} |S|^{\frac{n_1-p-1}{2}} \\ &\quad \times {}_1F_1\left(\frac{n_1+n_2}{2}; \frac{n_1}{2}; \Delta U^{-1}(I+S^{-1})^{-1}\right), \end{aligned} \quad (4.24)$$

考虑变形 $S(I+S)^{-1} = S_1(S_1+S_2)^{-1}$, 简单计算知 $S \mapsto S(I+S)^{-1}$ 的雅可比行列式为 $|I+S|^{-(p+1)}$,

$$\begin{aligned} p(S) &= \frac{|\Theta|}{2^{(n_1+n_2)p} B_p(\frac{n_1}{2}, \frac{n_2}{2}) |U|^{\frac{n_1+n_2}{2}}} |S|^{\frac{n_1-p-1}{2}} |I-S|^{-\frac{n_2-p-1}{2}} \\ &\quad \times {}_1F_1\left(\frac{n_1+n_2}{2}; \frac{n_1}{2}; \Delta U^{-1}S\right), \end{aligned} \quad (4.25)$$

通过在正交群上积分, 我们得到了统计量 $S_1(S_1 S_2)^{-1}$ 的特征根分布

$$\begin{aligned} &\frac{\pi^{p^2/2} |\Theta|}{2^{(n_1+n_2)p} B_p(\frac{n_1}{2}, \frac{n_2}{2}) \Gamma_p(\frac{p}{2}) |U|^{\frac{n_1+n_2}{2}}} \prod_{i < j}^p (f_i - f_j) |F|^{\frac{n_1-p-1}{2}} |I-F|^{\frac{n_1-p-1}{2}} \\ &\quad \times {}_1F_1\left(\frac{n_1+n_2}{2}; \frac{n_1}{2}; \Delta U^{-1}, F\right), \end{aligned} \quad (4.26)$$

其中 $F = \text{diag}(f_1, f_2, \dots, f_p)$, $f_1 > f_2 > \dots > f_p$; 别处为零.

根据非中心分布的结果, 我们的目标是计算积分

$$\begin{aligned} P(S < R) &= \frac{|\Theta|}{2^{(n_1+n_2)p} B_p(\frac{n_1}{2}, \frac{n_2}{2}) |U|^{\frac{n_1+n_2}{2}}} \int_0^R |S|^{\frac{n_1-p-1}{2}} |I - S|^{\frac{n_2-p-1}{2}} \\ &\quad \times {}_1F_1\left(\frac{n_1+n_2}{2}; \frac{n_1}{2}; \Delta U^{-1} S\right) dS \\ &= \frac{|\Theta| |R|^{\frac{n_1}{2}}}{2^{(n_1+n_2)p} B_p(\frac{n_1}{2}, \frac{n_2}{2}) B_p(\frac{n_1}{2}, \frac{p+1}{2}) |U|^{\frac{n_1+n_2}{2}}} \\ &\quad \times {}_2F_1(a, b; c; \Delta U^{-1}, \Delta U^{-1} R). \end{aligned}$$

其中 $a = -\frac{1}{2}(n_2 - p - 1)$, $b = \frac{1}{2}(n_1 + n_2)$, $c = \frac{1}{2}(n_2 + p + 1)$. 因此, 令 $R = xI_m$ 使得 $f_1 < x$ 等价于 $S < xI_m$, 最大特征根 f_1 的概率分布函数为

$$\begin{aligned} P(f_1 < x) &= \frac{|\Sigma_1| x^{\frac{n_1 p}{2}}}{2^{(n_1+n_2)p} B_p(\frac{n_1}{2}, \frac{n_2}{2}) B_p(\frac{n_1}{2}, \frac{p+1}{2}) |U|^{\frac{n_1+n_2}{2}}} \\ &\quad \times {}_2F_1(a, b; c; \Delta U^{-1}, \Delta U^{-1} R). \end{aligned}$$

where $a = -\frac{1}{2}(n_2 - p - 1)$, $b = \frac{1}{2}(n_1 + n_2)$, $c = \frac{1}{2}(n_2 + p + 1)$. 同样, 最小特征根 f_p 是 $I - S$ 的最大特征根. 通过对称性, 我们便知道了 f_1 的概率分布函数. \square

4.5 数值模拟

考虑两个相互独立的矩阵正态总体 $X_1 \sim N_{n,p}(M_1; A, B)$, $X_2 \sim N_{n,p}(M_2; I, B)$, 其中 $M_1 = \mu_1 \mathbf{1}_{n \times p}$, $M_2 = \mu_2 \mathbf{1}_{n \times p}$. 根据矩阵 F 分布密度, 我们比较了霍特林 $T^2 = \text{tr}(S)$ 、威尔克斯 $\Lambda = |S|$ 、洛伊 $\lambda_{\max}(S)$ 和似然比统计量 (安德森-布哈达方法) 在不同样本量 n 和变量数目 p 下关于以下三条假设的功效,

- $H_0: \mu_1 = \mu_2 = 0, A = \gamma I (\gamma > 0)$;
- $H_1: \mu_1 = \mu_2 = \mu$, 方差未知但相等;
- $H_2: \mu_1 \neq \mu_2$, 方差不相等.

表 4.2 (a) $\mu_1 = \mu_2 = 0$ ($A = \gamma I$, $B = I$), $n = 30$, $p = 2$ 的精确功效分析.

检验统计量	$\gamma = 1$	$\gamma = 2$	$\gamma = 3$	$\gamma = 4$	$\gamma = 5$	$\gamma = 6$
霍特林 $T^2 = \text{tr}(S)$	0.051	0.847	0.982	1.000	1.000	1.000
威尔克斯 $\Lambda = S $	0.049	0.821	0.963	0.993	0.998	1.000
洛伊 $\lambda_{\max}(S)$	0.048	0.805	0.951	0.990	0.997	0.999

表 4.3 (b) $\mu_1 = \mu_2 = \mu$ ($\gamma = 1$), $n = 30$, $p = 2$ 的精确功效分析.

检验统计量	$\mu = 0$	$\mu = 0.5$	$\mu = 1.0$	$\mu = 1.5$	$\mu = 2.0$	$\mu = 2.5$
霍特林 T^2	0.050	0.328	0.873	0.972	0.998	1.000
威尔克斯 Λ	0.048	0.302	0.842	0.961	0.994	0.999
洛伊最大根	0.047	0.284	0.821	0.953	0.988	0.998

表 4.4 使用线性逼近的渐近功效分析 (安德森-布哈达方法).

(c) $\mu_1 = 0.1, \mu_2 = 0$ ($A = \gamma I$, $B = I$)						
检验统计量	$\gamma = 1$	$\gamma = 2$	$\gamma = 3$	$\gamma = 4$	$\gamma = 5$	$\gamma = 6$
似然比 $-2 \ln \frac{L_1}{L_0}$	0.056	0.043	0.039	0.034	0.031	0.029
(c) $\mu_1 = 0.3, \mu_2 = 0$ ($A = \gamma I$, $B = I$).						
检验统计量	$\gamma = 1$	$\gamma = 2$	$\gamma = 3$	$\gamma = 4$	$\gamma = 5$	$\gamma = 6$
似然比 $-2 \ln \frac{L_1}{L_0}$	0.999	0.993	0.983	0.972	0.956	0.941
(c) $\mu_1 = 0.8, \mu_2 = 0$ ($A = \gamma I$, $B = I$).						
检验统计量	$\gamma = 1$	$\gamma = 2$	$\gamma = 3$	$\gamma = 4$	$\gamma = 5$	$\gamma = 6$
似然比 $-2 \ln \frac{L_1}{L_0}$	1.000	1.000	1.000	1.000	1.000	1.000

注: $N = 30$, $n = 2, 1000$ 次蒙特卡洛模拟.

随着样本量 n 和自变量维数 p 的增大, 霍特林 T^2 统计量的大小均维持在显著性水平 $\alpha = 0.05$ 附近, 而功效随非中心参数偏离零的程度增大而递增, 总体表现较佳. 威尔克斯 Λ 和最大根统计量次之. 而似然比统计则是在大样本情形表现出较满意的效果.

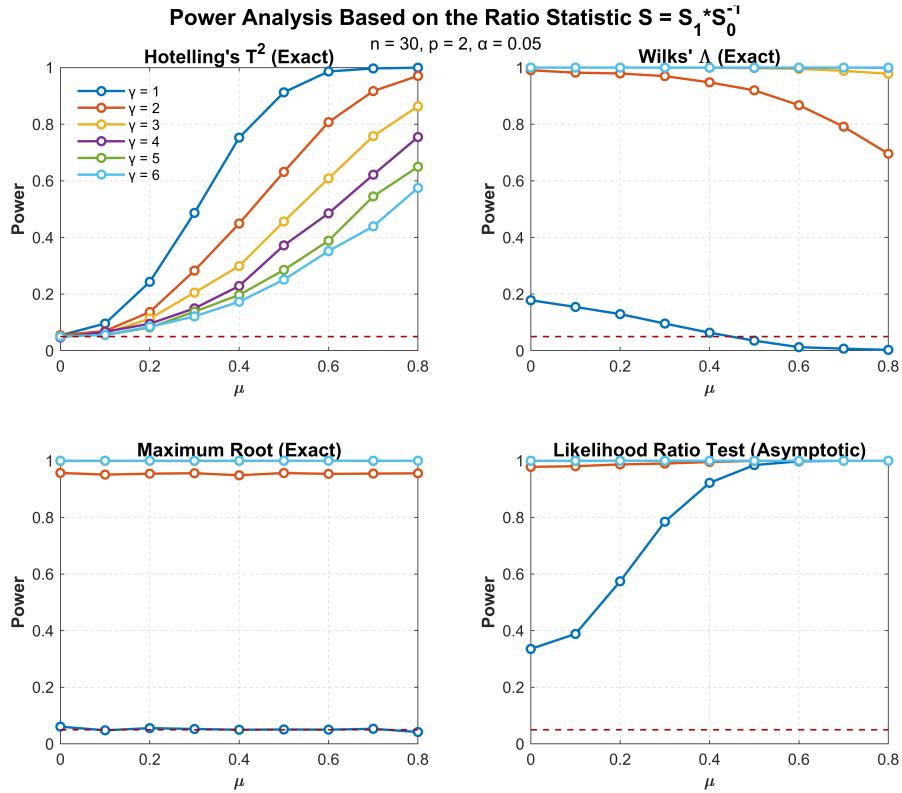


图 4.3 基于霍特林 T^2 (左上)、威尔克斯 Λ (右上)、洛伊最大根(左下)和似然比(右下)关于 $n = 30, p = 2, \alpha = 0.05$ 的功效分析. 在 3000 次蒙特卡罗模拟中, 前三个统计量的分布为精确分布, 最后一个统计量的分布为线性近似的 χ^2 分布.

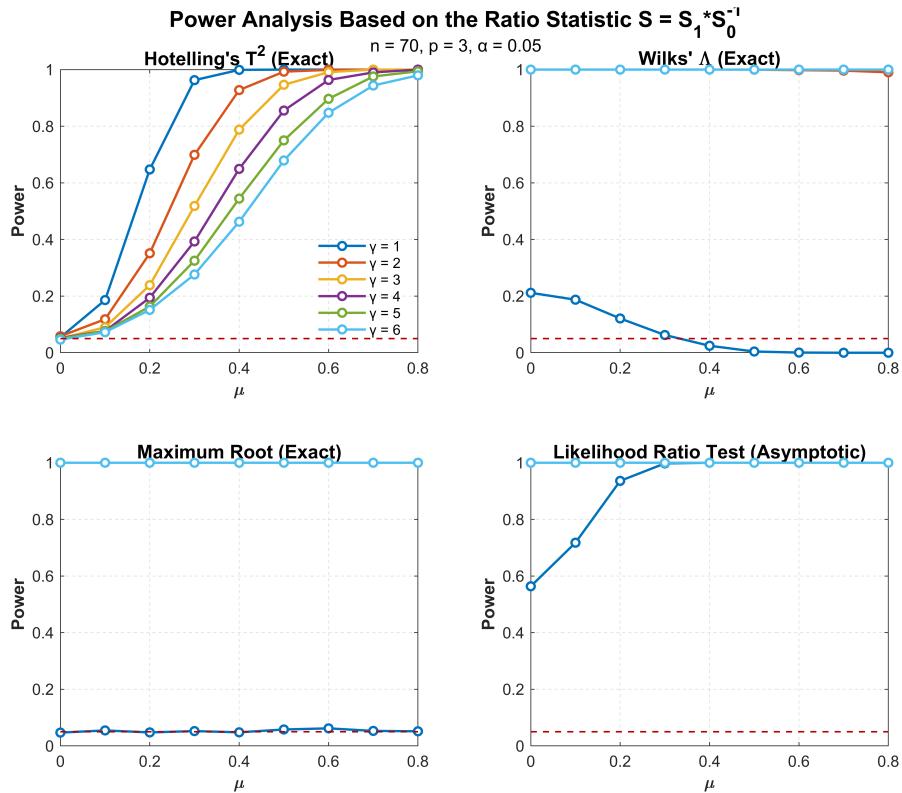
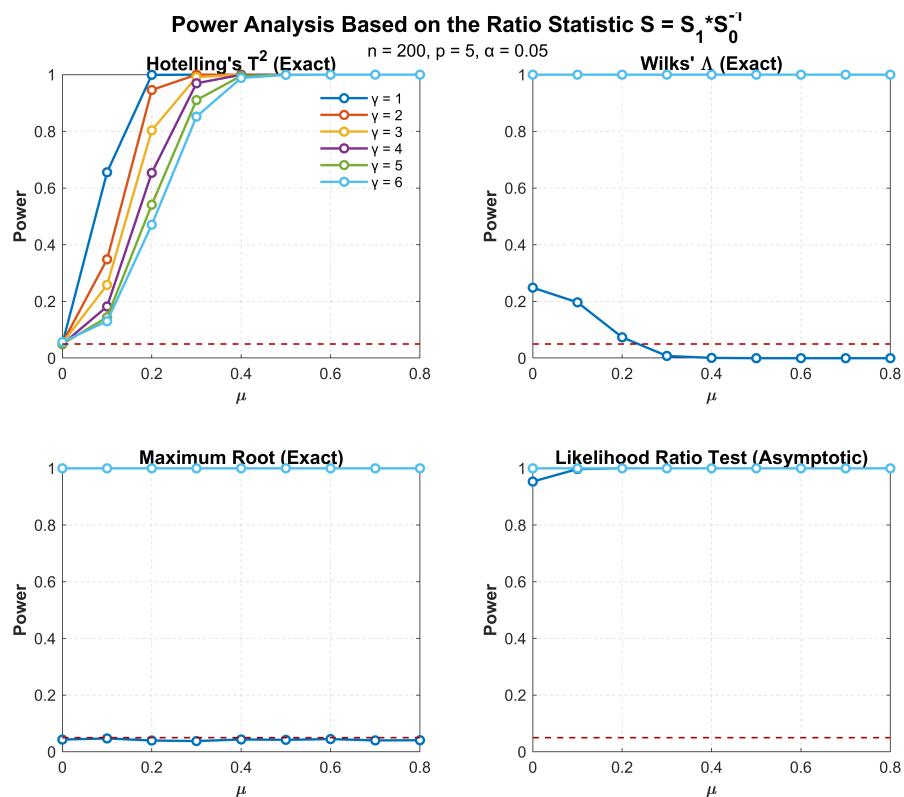


图 4.4 $n = 70, p = 3, \alpha = 0.05$ 的功效分析. 与上图类似.

图 4.5 $n = 200, p = 5, \alpha = 0.05$ 的功效分析. 与上图类似.

第5章 点过程的极小可料强度

在本章中, 我们将研究点过程的重要性采样与拟合优度检验. 一方面, 时变泊松过程已被广泛应用于点过程的拟合优度检验, 但严格的理论基础涉及到证明极小滤子的存在性, 这使得布里莫^[34]1972年的提问有意义, 也自然地引出了极小可料强度的定义. 另一方面, 绝对连续的测度变换在重要性采样方面也有重要应用. 为此, 我们先研究点过程在测度变换下的性质.

设 $\mathcal{F}_t, t \geq 0$ 是抽象样本空间 Ω 上包含所有可能结果的单调非降 σ -代数子集族, 其中 $\mathcal{F}_s \subseteq \mathcal{F}_t \subseteq \mathcal{F} (t > s)$, 包含于所有可能发生事件的 σ -代数 \mathcal{F} . 记 $\{\mathcal{F}_t : t \geq 0\}$ 为 \mathbf{F} . 我们关心过程 $\underline{N} = \{N(\omega, t), \mathbf{F}, \mathbf{P}\}$, 对于所有 ω , $N(\omega, t)$ 是 $[0, \infty)$ 上的分片常值函数, 其中 \mathbf{P} 是定义在 σ -代数 \mathcal{F} 上的概率测度. 这里假设对于任意 $t \geq 0$ 和任意实数轴的子集 Γ , 集合 $\{\omega : N(\omega, t) \in \Gamma\} \in \mathcal{F}_t$.

5.1 极小可料强度

\underline{N} 称作强度为 λ 的泊松过程, 如果它满足以下条件: 对于任意 $t \geq 0$ 和实数轴上的任意博雷尔集 Γ ,

(P0) (适应) $\{\omega : N(\omega, t) \in \Gamma\} \in \mathcal{F}_t$;

(P1) (右连续) $\lim_{t \rightarrow s^+} N(\omega, s) = N(\omega, s)$ ($f(x+)$ 表示右极限 $\lim_{y \rightarrow x^+} f(y)$, $f(x-)$ 表示左极限 $\lim_{y \rightarrow x^-} f(y)$);

(P2) (非减) $N(\omega, t) \geq N(\omega, s), t > s$;

(P3) (整值) $N(\omega, t) \in \{0, \pm 1, \pm 2, \dots\}$;

(P4) (零初值) $N(\omega, 0) = 0$;

(P5) (单位跳跃) $\Delta N(\omega, s) = N(\omega, s) - N(\omega, s-) \in \{0, 1\}$, 对于固定的 ω ; 并且如果

(P6) (泊松律) $\int_{\Omega} \exp[iu(N(\omega, t) - N(\omega, s))] \mathbf{P}(d\omega) = \exp[\lambda(t-s)(e^{iu} - 1)]$;

(P7) (独立增量) $N_t - N_s \perp \mathcal{F}_s$, 其中 $0 \leq s \leq t$.

事实上, 根据伦伊^[95]的结果, 泊松过程的两两独立性(符号 \perp 表示这一点)意味着 n 分量独立性. 定义 $T_n(\omega) = \inf\{t \geq 0 : N(\omega, t) \geq n\}$. 那么泊松过程可以

表示为

$$N(\omega, t) = \sum_{n=1}^{\infty} \mathbf{1}\{T_n(\omega) \leq t\}, \quad (5.1)$$

其中 $U_{n+1}(\omega) = T_{n+1}(\omega) - T_n(\omega)$, 其中 $n = 1, 2, \dots$, 服从均值为 $1/\lambda$ 的独立同分布指数分布.

(5.1)式中的定义可以推广至任意停时列 $0 (= T_0(\omega)) \leq T_1(\omega) \leq T_2(\omega) \leq \dots$, 其中 $T_n(\omega)$ 趋于无穷, 且对于所有 t 和 n , 都有 $\{\omega : T_n(\omega) \leq t\} \in \mathcal{F}_t$. 由于对于固定 t , (5.1)式是示性函数 $\{\omega : T_n(\omega) \leq t\} \in \mathcal{F}_t$ 之和, 因此它是适应的; 由于示性函数非负, 且对于固定 ω , 它具有右连续性, 因此(5.1)式也是右连续的. 它显然是非减的, 且值为整数. 然而, 它不一定从零开始, 也不一定有单位跳跃, 因为两个这样的 $T_n(\omega)$ 和 $T_{n'}(\omega), n \neq n'$ 可能在 Ω 的某个子集 $\Omega' \in \mathcal{F}$ 上相等, 使得 $P(\Omega') > 0$. 除此以外, 满足 (P0)-(P5) 的过程 \underline{N} 称为简单点过程. 如果一个适应过程在两个变量上关于所有左连续适应过程生成的 $\Omega \times [0, \infty)$ 上的最小 σ -代数是联合可测的, 则称该适应过程是可料的. 简单点过程 \underline{N} 的补偿子是唯一可料的非减过程 $A(\omega, t)$, 其中 $A(\omega, 0) = 0$, 使得 $N(\omega, t) - A(\omega, t)$ 是 F -局部鞅. 简单点过程局部具有可积变差 (参见^[27]定理 3.22 之前第 66 页的定义), 因此其补偿子始终存在. 例如, 泊松过程的补偿子是 $A(\omega, t) = \lambda t$.

对于简单点过程 $\underline{N} = \{N(\omega, t), F, P\}$, 如果 F 与所有 $N(\cdot, s), s \leq t$ 生成的自然滤子相一致, 则我们称简单点过程 $\underline{N} = \{N(\omega, t), F, P\}$ 是非预期的, 并且其补偿器 $A(\omega, t)$ 是自然的. 根据佐恩引理, 存在一个极小滤子 F , 它是所有滤子 F 在偏序集 (S, \leq) 中的极小元, 满足

(S1) N 是 F -适应的,

(S2) A 是 F -可料的,

(S3) $N - A$ 是一个 F -局部鞅,

其中 $G \leq F \Leftrightarrow \mathcal{G}_t \in G \subseteq \mathcal{F}_t \in F$, 对所有 $t \geq 0$ 都成立. 这是因为 S 中满足条件的滤子降链的非空交再次满足 (S1)-(S3), 从而给出了佐恩引理的下界.^[34]提出了一个问题: 如果存在一个概率测度 P_0 , 相对于该测度, 点过程 $\underline{N} = \{N(\omega, t), F, P\}$ 绝对连续, 而 $\underline{N}^0 = \{N(\omega, t), F, P_0\}$ 是标准泊松过程, 那么极小滤子 $F \neq F^N$ 是否成立? 在本章节, 我们尝试对他的问题做出简要的回答: 这是不可能的.

回顾一下简单点过程 \underline{N} 不一定是有序的, 这意味着对于任何停时 $\tau(\omega) \geq 0$,

$$\lim_{h \rightarrow 0} \frac{1}{h} P(N(\cdot, \tau(\cdot) + h) - N(\cdot, \tau(\cdot))) > 1 \mid \mathcal{F}_{\tau}(\omega) = 0, \quad (5.2)$$

而有序性意味着简单性(例如,习题 3.3.2 中的混合泊松过程^[32]). 这里, \mathcal{F}_τ 表示 \mathcal{F} 的停止 σ -子代数,

$$\mathcal{F}_\tau = \{A : A \cap \{\omega : \tau(\omega) \leq t\} \in \mathcal{F}_t, \text{ for all } t \geq 0\},$$

有序点过程 \underline{N} 的强度 $\lambda(\omega, t)$ 是一个非负过程, 对于任何可料停时 $\tau(\omega) \geq 0$, 满足

$$\lim_{h \rightarrow 0} \frac{1}{h} \mathbf{P}(N(\cdot, \tau(\cdot) + h) - N(\cdot, \tau(\cdot)) = 1 \mid \mathcal{F}_\tau)(\omega) = \lambda(\omega, \tau(\omega)). \quad (5.3)$$

这里, 我们使用了与^[34]不同的定义, 其考虑将(5.2)和(5.3)中的停时替换为固定的 $t \geq 0$. 我们的定义的优点在于, 对两个这样的过程 $\lambda(\omega, t)$ 和 $\lambda'(\omega, t)$, 它要求了更强的不可区分唯一性, 即

$$\mathbf{P}(\lambda(\cdot, t) \neq \lambda'(\cdot, t) \text{ 存在 } t \geq 0) = 0.$$

这由可料截口定理(^[96]定理 4.85)保证, 而在布里莫的定义中, 只假设版本唯一性, 即

$$\mathbf{P}(\lambda(\cdot, t) \neq \lambda'(\cdot, t)) = 0, \text{ 对任意 } t \geq 0.$$

设 \mathbf{F} 为适应滤子, \mathbf{F}^N 为简单点过程 \underline{N} 生成的自然滤子. 另设 \mathcal{F}_∞ 为包含所有 $\mathcal{F}_t, t \geq 0$ 的 \mathcal{F} 的最小 σ -子代数. 我们假设 $\lambda(\omega, t)$ 满足

(A0) $\lambda(\omega, t)$ 是 \mathbf{F} -适应的;

(A1) $\lambda(\omega, t)$ 对于 ω 和 t 都是可测的;

(A2) $\lambda(\omega, t)$ 对于固定的 ω 和 t 都是严格正的;

(A3) 对于每个 t 几乎必然存在 $\int_0^t \lambda(\omega, s) ds < \infty$;

(A4) 对于每个 t 几乎必然存在 $\int_0^t \log \lambda(\omega, s) N(\omega, ds)$.

从 $\lambda(\omega, t)$ 中我们定义 $M(\lambda)$ 、 $A(\lambda)$ 和 $\zeta_s^t(\lambda)$ 如下:

$$M(\lambda)(\omega, t) = N(\omega, t) - A(\lambda)(\omega), \text{ 其中 } A(\lambda)(\omega, t) = \int_0^t \lambda(\omega, s) ds,$$

且 $\zeta_s^t(\lambda)(\omega) = \int_s^t [1 - \lambda(\omega, u)] du + \int_s^t \log \lambda(\omega, u) N(\omega, du).$

对于 $t \in [0, \infty]$, 设

$$Q(d\omega) = \exp[\zeta_0^t(\lambda)(\omega)] P(d\omega), \text{ 在 } \mathcal{F}_t \text{ 上.} \quad (5.4)$$

定理 5.1. 假设 $\underline{N} = \{N(\omega, t), \mathbf{F}, \mathbf{P}\}$ 是强度为 1 的泊松过程, 且 $\lambda(\omega, t)$ 满足 (A0)-(A4). 在(5.4)中取 $t = \infty$. 我们假设 $\mathbf{Q}(\Omega) = 1$. 对于任意 \mathbf{G} 使得 $\mathbf{F}^N \leq \mathbf{G} \leq \mathbf{F}$, 存在不可区分意义下唯一的 \mathbf{G} 可料过程 $\lambda^{\mathbf{G}}(\omega, t)$, 使得 $\underline{M}(\lambda^{\mathbf{G}}) = \{\underline{M}(\lambda^{\mathbf{G}})(\omega, t), \mathbf{G}, \mathbf{Q}\}$ 为鞅. 特别地, 如果 $\lambda(\omega, t)$ 是 \mathbf{G} 可料的, 那么它与 $\lambda^{\mathbf{G}}(\omega, t)$ 是不可区分的.

注 5.1. 根据截口定理和停止定理 ([96] 定理 6.43), 任何非负 \mathbf{F} -可测过程 $\lambda(\omega, t)$ 都有一个 \mathbf{G} -可料投影 ${}^p\lambda(\omega, t)$, 使得对于任何有限非负可料停时 $\tau(\omega) \geq 0$,

$$\mathbf{E}_{\mathbf{P}}[\lambda(\omega, \tau(\omega)) | \mathcal{G}_{\tau-}](\omega) = {}^p\lambda(\omega, \tau(\omega)) \quad (5.5)$$

其中 $\mathcal{G}_{0-} = \mathcal{G}_0$ 且

$$\mathcal{G}_{\tau-} = \sigma(\{A \cap \{\omega : t < \tau(\omega)\} : t \geq 0, A \in \mathcal{G}_t\} \cup \mathcal{G}_0).$$

我们可以很容易地证明, 对于任何趋向于 $\tau(\omega)$ 的非负非减停时列 $\tau_n(\omega)$, $\mathcal{G}_{\tau-}$ 是包含所有 \mathcal{G}_{τ_n} 的最小 σ -代数 ([27] 定理 3.6). 因此, 定理 5.1 可以重述如下: ${}^p\lambda(\omega, t)$ 是不可区分意义下唯一的 \mathbf{G} 可料过程, 使得 $\underline{M}({}^p\lambda) = \{\underline{M}({}^p\lambda)(\omega, t), \mathbf{G}, \mathbf{Q}\}$ 为鞅.

下文中, 当不致引起混淆时, 我们将简写这些记号: $N_t(\omega) = N(\omega, t)$, 有时省略 ω ; $\mathbf{E}_{\mathbf{P}}[Z] = \int_{\Omega} Z(\omega) \mathbf{P}(d\omega)$, 有时省略下标 \mathbf{P} .

引理 5.2. 假设 $\underline{N} = \{N(\omega, t), \mathbf{F}, \mathbf{P}\}$ 满足 (P0)-(P5) 且

$$\int_{\Omega} N(\omega, t) \mathbf{P}(d\omega) < \infty, \text{ 对于每个 } t \geq 0, \quad (5.6)$$

假设 $\underline{\lambda} = \{\lambda(\omega, t), \mathbf{F}, \mathbf{P}\}$ 满足 (A0)-(A4). 以下断言是等价的.

1. 对于任意非负的 \mathbf{F} -可料过程 $C(\omega, t)$,

$$\int_{\Omega} \mathbf{P}(d\omega) \left[\int_0^{\infty} C(\omega, t) N(\omega, dt) \right] = \int_{\Omega} \mathbf{P}(d\omega) \left[\int_0^{\infty} C(\omega, t) A(\omega, dt) \right].$$

2. $N(\omega, t) - A(\omega, t)$ 是 \mathbf{P} 下的 \mathbf{F} -鞅.

3. $N(\omega, t) - A(\omega, t)$ 是 \mathbf{P} 下的 \mathbf{F} -局部鞅.

引理 5.2 的证明. (3) \Rightarrow (1). 如果 $N(\omega, t) - A(\omega, t)$ 是局部鞅, 则 $A(\omega, t)$ 是可积的, 因为在(5.6)中要求了 $N(\omega, t)$ 对每个 $t \geq 0$ 都是可积的. 让我们选取一个趋于无穷大的非负停时列 $\tau_n(\omega)$ 和一个可料的生成元

$$C(\omega, u) = \mathbf{1}\{\omega \in \Gamma\} \mathbf{1}\{s < u \leq t\}$$

, 其中 $\Gamma \in \mathcal{F}_s$,

$$\begin{aligned} E[(N_{t \wedge \tau_n} - N_{s \wedge \tau_n}); \Gamma] &= E\left[\int_0^\infty C_u dN_{u \wedge \tau_n}\right] \\ &= E\left[\int_0^\infty C_u dA_{u \wedge \tau_n}\right] = E[(A_{t \wedge \tau_n} - A_{s \wedge \tau_n}); \Gamma], \end{aligned} \quad (5.7)$$

其中(5.7)中的第一个和最后一个期望取自 Ω 的子集 Γ . 设 n 趋向于 ∞ , 则由单调类定理可得 (1). 如果 (5.7) 对任意可料过程 $C(\omega, u)$ 成立, 则通过 $N(\omega, t)$ 的可积条件(5.6)证明 $N(\omega, t)$ 再次为鞅. 逆命题 (1) \Rightarrow (2) 也成立. 由于任意鞅都是局部鞅, 因此 (2) \Rightarrow (3) 成立. \square

引理 5.3. 假设同定理5.1. 如果我们令 $Z(\omega, t) = \exp[\zeta_0^t(\omega)]$, 则 $E[Z_\infty] = 1$ 意味着 $\underline{Z} = \{Z(\omega, t), \mathcal{F}, \mathbf{P}\}$ 是一致可积鞅, 即 \underline{Z} 是鞅, 并且

$$1. \exists C > 0, \sup_t E[|Z_t|] \leq C;$$

$$2. \forall \varepsilon > 0, \exists \delta > 0 \text{ 使得 } \forall \Gamma \in \mathcal{F}_\infty, P(\Gamma) \leq \delta, \sup_t E[|Z_t|; \Gamma] < \varepsilon.$$

引理5.3的证明. 将 $Z(\omega, t) = \exp[\zeta_0^t(\omega)]$ 重写为

$$Z_t = \exp\left[\int_0^t \log \lambda(\omega, s) N(\omega, ds)\right] / \exp\left[\int_0^t (\lambda(\omega, s) - 1) ds\right] (= X_t/Y_t).$$

我们将证明 \underline{Z} 是鞅. 通过微分, 我们发现

$$\begin{aligned} dZ &= Y dX + X dY = Y(dX - X(\lambda - 1)ds)(= Y dM), \\ M(\omega, t) &= X(\omega, 0) + \int_0^t [X(\omega, s) - X(\omega, s-)][N(\omega, ds) - ds]. \end{aligned}$$

这里, \underline{M} 是鞅. 由于 \underline{Y} 料, 因此 \underline{Z} 是局部鞅. 设其对应的约化停时列为 $\tau_n(\omega) \geq 0$. 根据法图引理,

$$E[Z_t | \mathcal{F}_s] \leq \liminf_{n \rightarrow \infty} E[Z_{t \wedge \tau_n} | \mathcal{F}_s] = \liminf_{n \rightarrow \infty} Z_{s \wedge \tau_n} = Z_s, \quad t \geq s$$

因此, \underline{Z} 是非负上鞅. 非负上鞅几乎必然收敛于一个非负随机变量, 即 Z_∞ . 由于 $E[Z_t | \mathcal{F}_s]$ 的取值范围为 0 到 Z_s , 因此 $E[Z_\infty] = E[Z_0] = 1$ 意味着等号成立, 于是 \underline{Z} 是鞅. \square

一致可积性由以下基本引理得出, 其证明见^[27]第一章定理 11.

引理 5.4. 设 $\underline{Z} = \{Z(\omega, t), \mathcal{F}, \mathbf{P}\}$ 为鞅. 以下三个条件等价.

1. 存在一个随机变量 Z_∞ 使得 $\lim_{t \rightarrow \infty} E|Z_t - Z_\infty| = 0$.
2. 存在一个随机变量 Z_∞ 使得 $E[Z_\infty] < \infty$ 且 $Z(\omega, t) = E[Z_\infty | \mathcal{F}_t](\omega)$.
3. $\underline{\underline{Z}} = \{Z(\omega, t), \mathbf{F}, \mathbf{P}\}$ 一致可积.

引理 5.5. 在 (A0)-(A4) 下, $\underline{\underline{M}}(\lambda) = \{M(\lambda)(\omega, t), \mathbf{F}, \mathbf{Q}\}$ 为鞅.

引理 5.5 的证明. 根据引理 5.2, 我们需要证明, 对于任意非负的 \mathbf{F} -可料过程 $C(\omega, t)$,

$$E_{\mathbf{Q}}[\int_0^\infty C_t dN_t] = E_{\mathbf{Q}}[\int_0^\infty C_t \lambda_t dt]. \quad (5.8)$$

根据 \mathbf{Q} 的定义, (5.8) 等价于

$$E_{\mathbf{P}}[Z_\infty \int_0^\infty C_t dN_t] = E_{\mathbf{P}}[Z_\infty \int_0^\infty C_t \lambda_t dt], \quad (5.9)$$

结合福比尼定理和引理 5.3, 可得

$$E_{\mathbf{P}}[\int_0^\infty Z_t C_t dN_t] = E_{\mathbf{P}}[\int_0^\infty Z_t C_t \lambda_t dt]. \quad (5.10)$$

通过构造 $Z(\omega, t)$, 在跳跃时 $t = T_n(\omega)$, $Z(\omega, t) = Z(\omega, t-) \lambda(\omega, t)$, 否则, $Z(\omega, t) = Z(\omega, t-)$. (5.10) 式的左边简化为

$$E_{\mathbf{P}}[\int_0^\infty Z_{t-} C_t \lambda_t dN_t] = E_{\mathbf{P}}[\int_0^\infty Z_{t-} C_t \lambda_t dt] \quad (5.11)$$

这里, 我们用到了 $N(\omega, t) - t$ 在 \mathbf{P} 下为 \mathbf{F} -鞅这一事实. 根据(5.11)即得(5.10), 因为半实直线上的任何可数集的勒贝格测度均为零. \square

引理 5.6. 如果定理 5.1 对 \mathbf{G} 成立, 则它对任何满足 $\mathbf{G} \leq \mathbf{G}' \leq \mathbf{F}$ 的 \mathbf{G}' 也成立.

引理 5.6 的证明. 直接来自引理 5.4. \square

因此, 我们可以集中注意力证明定理 5.1 对 $\mathbf{G} = \mathbf{F}^N$. 以下引理是^[97]第三章中定理 1.21 的一个版本. 我们不作证明.

引理 5.7. 假设同引理 5.2. 设 \mathbf{P} 和 \mathbf{Q} 分别是 $(\Omega, \mathcal{F}_\infty^N)$ 上的两个概率测度, 其中 $N(\omega, t)$ 具有共同的 \mathbf{F}^N 补偿子. 则 \mathbf{P} 和 \mathbf{Q} 在 \mathcal{F}_∞^N 上相等.

我们将利用哈恩-巴拿赫定理证明点过程的可料表示性质. 证明思路来自^[98].

引理 5.8. 假设同定理 5.1, 并且引理 5.2 中的某个陈述成立. 如果 $\underline{\underline{L}} = \{L(\omega, t), \mathbf{F}^N, \mathbf{P}\}$ 是一致可积鞅, 则存在不可区分意义下唯一的 \mathbf{F}^N -可料过程 $H(\omega, t)$, 使得除开一个 \mathbf{P} -零集,

$$L(\omega, t) = L(\omega, 0) + \int_0^t H(\omega, s)[N(\omega, ds) - A(\omega, ds)], \quad \text{对于每个 } t. \quad (5.12)$$

引理5.8的证明. 设 $L^1(\mathbf{P})$ 由(5.12)中 $H(\omega, t)$ 遍历所有 \mathbf{F}^N -可料过程的过程 $L(\omega, t)$ 组成. 显然, $L^1(\mathbf{P}) \subseteq M_{\text{loc}}(\mathbf{P})$, \mathbf{P} 下所有局部鞅的集合. 我们的目标是证明包含关系是等号. 如若不然, 由于 $L^1(\mathbf{P})$ 是一个真闭子空间, 根据哈恩-巴拿赫定理, 存在一个连续线性泛函 $f \in M_{\text{loc}}(\mathbf{P})^*$, 使得对于所有 $L \in L^1(\mathbf{P})$, $f(L) = 0$ 不恒等式为零. 但是, 根据^[99]定理10.21中的经典的哈代-BMO对偶性, 存在一个非零过程 $K \in BMO$ 使得

$$E([L, K]_\infty) = 0, \text{ 对任意 } L \in L^1(\mathbf{P}).$$

注意, $BMO \subseteq M_{\text{loc}}^\infty(\mathbf{P})$, 即 \mathbf{P} 下所有有界局部鞅的集合. 因此存在一个停止时间 τ 使得 $|K_{t \wedge \tau}| \leq M$. 为简便起见, 取 $M = 1$. 简单计算可知

$$E(L_\infty K_{0 \wedge \tau}) = E(L_\infty K_{\infty \wedge \tau}) = E([L, K_{t \wedge \tau}]_\infty) = E([L_{t \wedge \tau}, K]_\infty) = 0$$

因为 $L_{t \wedge \tau} \in L^1(\mathbf{P})$. 因此, $K_0 = K_{0 \wedge \tau} = 0$, 因为 $\mathcal{F}_0^N = \{\emptyset, \Omega\}$, 且我们可以取 $L \equiv 1$. 定义

$$Z_t = 1 + \frac{1}{2}K_{t \wedge \tau}, \quad \frac{1}{2} \leq Z_t \leq \frac{3}{2}.$$

特别地, 我们有 $E[Z_t] = E[Z_0] = 1$, 且当 $L \in L^1(\mathbf{P})$ 且 $L(\omega, 0) = 0$ 时, $E(L_\infty Z_\infty) = 0$. 这表明, 由 $d\mathbf{P}' = Z_\infty d\mathbf{P}$ 定义的概率测度 \mathbf{P}' 满足 $\mathbf{P}' \sim \mathbf{P}$. 根据引理5.2, $\underline{\underline{M}}' = \{M(\omega, t), \mathbf{F}^N, \mathbf{P}'\}$ 为鞅, 其中 $M(\omega, t) = N(\omega, t) - A(\omega, t)$. 根据引理5.7, $\underline{\underline{P}}' = \mathbf{P}$, 因此 $Z_\infty = 1$, 即 $K_\tau = 0$. 由于 τ 可以任意大, 因此 $K \equiv 0$, 矛盾. \square

引理5.9. $L(\omega, t) = E_{\mathbf{P}}[dQ/d\mathbf{P}|\mathcal{F}_t^N]$ 是 \mathbf{F}^N -一致可积鞅.

引理5.9的证明. 由引理5.4. \square

引理5.10. 在(A0)-(A4)条件下, $\underline{\underline{M}}(\lambda^N) = \{M(\lambda^N)(\omega, t), \mathbf{F}^N, Q\}$ 为鞅.

定理5.10的证明. 根据鞅表示引理5.8, 存在一个 \mathbf{F}^N -可料过程 $H(\omega, t)$ 使得

$$L(\omega, 0) = 1, \quad L(\omega, t) = 1 + \int_0^t H(\omega, s)(N(\omega, ds) - ds)\mathbb{1} \quad (5.13)$$

根据(5.13), 在跳跃时 $t = T_n(\omega)$, $L(\omega, t) = L(\omega, t-) + H(\omega, t)$, 否则, $L(\omega, t) = L(\omega, t-)$. 我们在扩充半实直线 $[0, \infty]$ 上引入 $\tau(\omega) = \inf\{t : L(\omega, t) = 0\}$, 对于给定的 ω , 它在 ∞ 处严格为正, 并且可能取值无穷. 此外, 根据可测截口定理(^[96]), $\tau(\omega)$ 也是一个 \mathbf{F}^N -停时. 由此可知

$$\lambda^N(\omega, t) = [1 + H(\omega, t)L^{-1}(\omega, t-)]\mathbb{1}\{\omega : \tau(\omega) > t\} \quad (5.14)$$

是一个非负的 \mathbf{F}^N -可料过程, 满足(5.8). 根据引理5.2, $\underline{\underline{M}}(\lambda^N) = \{M(\lambda^N)(\omega, t), \mathbf{F}^N, Q\}$ 是鞅. \square

唯一性同样来自引理5.2. 结合引理5.5和引理5.10, 我们得到以下结论.

引理 5.11. 假设同定理 5.1. $\lambda^N(\omega, t)$ 是唯一的 F^N -可料过程, 使得定理5.1成立.

事实上, 我们也有定理5.1的逆定理.

定理 5.12. 假设 $\underline{N} = \{N(\omega, t), F^N, P\}$ 是强度为 1 的泊松过程, 而 $\lambda(\omega, t)$ 是一个满足 (A0)-(A4) 的 F^N -可料过程. 设 Q 是 $(\Omega, \mathcal{F}_\infty^N)$ 上的另一个概率测度, 使得 $\underline{M}(\lambda) = \{M(\lambda), F^N, Q\}$ 为鞅. 则(5.4)对于 $\lambda(\omega, t)$ 和 $t \in [0, \infty]$ 成立.

定理5.12的证明. 根据引理5.7, 只需证明, 对于每个 $t \geq 0$, Q 在 \mathcal{F}_t 上的限制 Q_t 与概率测度 \tilde{Q}_t 相同, 该概率测度由 $\tilde{Q}_t(d\omega) = \exp[\zeta_0^t(\lambda)(\omega)]P(d\omega)$ 在 \mathcal{F}_t^N 给出. 根据引理5.2, 只需要证明, 对于任意非负的 F^N -可料过程 $C(\omega, s)$, 在 $[0, t]$ 之外为零,

$$E_{\tilde{Q}_t} \left[\int_0^t C_s dN_s \right] = E_{\tilde{Q}_t} \left[\int_0^t C_s \lambda_s ds \right]. \quad (5.15)$$

根据 \tilde{Q}_t 的定义, (5.15)变为

$$E_P \left[Z_t \int_0^t C_s dN_s \right] = E_P \left[Z_t \int_0^t C_s \lambda_s ds \right], \quad (5.16)$$

结合福比尼定理和引理5.3, 可得

$$E_P \left[\int_0^t Z_s C_s dN_s \right] = E_P \left[\int_0^t Z_s C_s \lambda_s ds \right]. \quad (5.17)$$

根据引理5.5中的相同推理, (5.17)的左边可变形为

$$E_P \left[\int_0^t Z_{s-} C_s \lambda_s dN_s \right] = E_P \left[\int_0^t Z_{s-} C_s \lambda_s ds \right] \quad (5.18)$$

其中我们利用以下事实, $N(\omega, t) - t$ 是 P 下的 F^N -鞅. 根据(5.18)可得(5.17), 因为任意 $[0, t]$ 上的可数集的勒贝格测度为零. 因此, Q 在 \mathcal{F}_t 上的限制 Q_t 与 \tilde{Q}_t 具有相同补偿子. (5.4), 对固定 $t \geq 0$ 成立. 根据引理 5.4, (5.4)对 $t = \infty$ 也成立. \square

推论 5.13. 一个适应的、右连续的、非递减的、具有单位跳跃且初值为零的整值随机过程具有极小可料强度当且仅当它是绝对连续测度变换下的标准泊松过程.

5.2 典则分解

霍克斯过程是一类特殊的点过程, 也称作自激励过程, 其增量可能不相互独立. 霍克斯过程在时刻 t 之后的时刻到达数目和到达时间取决于时刻 t 之前的事件到达数目 N_t 以及之前的到达时间 $0 = T_0 < T_1 < T_2 < \dots < T_{N_t}$,

$$\lambda(\omega, t) = \mu(t) + \sum_{n: T_n(\omega) < t} \phi(t - T_n(\omega)),$$

其中 μ 是定义在 $[0, \infty)$ 上的非负局部可积函数, 称为背景强度函数; ϕ 是定义在实直线上的非负连续函数, 满足

$$m = \int_0^\infty \phi(t) dt < \infty,$$

且当 $t < 0$ 时 $\phi(t) = 0$. 霍克斯过程在各种统计应用中发挥着重要作用. 例如, 它可以模拟事件到达率随时间波动的情况, 比如客户访问或网络请求, 这些波动可能是由内部因素引起的. 由于霍克斯过程的强度依定义是极小可料的, 因而

定理 5.14. 霍克斯过程是绝对连续测度变换下的标准泊松过程.

这一定理为霍克斯过程的重要性采样提供了理论基础. 主要定理的另一推论则是考克斯过程, 也称为双随机泊松过程, 通常在给定某个随机测度时作为随机测度引入. 事实上, ^[100] 定义了此类测度空间上的完备可分度量, 从而允许条件期望. 这里我们采用不同的定义. 假设 $\underline{N} = \{N(\omega, t), \mathcal{F}, \mathbf{P}\}$ 满足 (P0)-(P5). 考克斯过程 \underline{N} 由其补偿子 $A(\omega, t)$ 刻画

- (C0) (适应) $N(\omega, t), A(\omega, t)$ 适应于 \mathcal{F} ;
- (C1) (初始适应) \mathcal{F}_0 包含由所有 $A(\cdot, t), t \geq 0$ 生成的 σ -代数;
- (C2) (条件泊松律) 随机变量 $N(\cdot, t) - N(\cdot, s), 0 \leq s < t$ 在 $n = 0, 1, 2, \dots$ 时满足条件泊松律

$$\mathbf{P}(N(\cdot, t) - N(\cdot, s) = n | \mathcal{F}_s)(\omega) = \frac{(A(\omega, t) - A(\omega, s))^n e^{A(\omega, s) - A(\omega, t)}}{n!};$$

- (C3) (条件独立增量) 随机变量 $N(\cdot, t) - N(\cdot, s)$ 和 $N(\cdot, u) - N(\cdot, v)$ 在给定 \mathcal{F}_0 时是条件独立的, 其中 $0 \leq s < t \leq u < v$.

由定理5.1, 定理5.12和勒贝格分解定理, 我们得到

定理 5.15. 任意满足 (C0)-(C3) 的考克斯过程都可以分解为三部分之和

$$N + A + M, \tag{5.19}$$

其中 N 是具有强度的考克斯过程, A 几乎必然关于勒贝格测度奇异, M 是鞅. 特别地, 如果 N 独立于 A , 则 N 为非齐次泊松过程.

引理 5.16. 具有自然强度的考克斯过程是泊松过程.

引理5.16的证明. 对任意 $t \geq 0$, $A_t \in \mathcal{F}_0^N = \{\emptyset, \Omega\}$. 因而 A_t 为确定函数. \square

引理 5.17. 一个适应的、右连续的、非减的过程, 其勒贝格分解中的绝对连续分量和奇异分量都是适应的.

引理 5.17 的证明. 设 $D[0, \infty)$ 为所有在 $[0, \infty)$ 上的右连左极函数构成的空间, 且该函数具有^[101] 意义下的可度量拓扑 (参见^[102] 第 12 节). 定义 $F(\omega) \subseteq D[0, \infty) \times D[0, \infty)$ 为适应的、右连续的、非减过程 A 的勒贝格分解 $A = B + C$ 中的所有对 (B, C) , 其中 B 绝对连续, C 关于勒贝格测度奇异. 对于给定的 $\omega \in \Omega$, $F(\omega)$ 非空. 根据可测截面定理 (^[96] 第三章定理 81), 我们可以选取一个可测过程 $(B, C) : \Omega \rightarrow D[0, \infty) \times D[0, \infty)$, 使得勒贝格分解成立. B 是适应的这一事实由

$$\begin{array}{ccc} \Omega & \xrightarrow{(B,C)} & D[0, \infty) \times D[0, \infty) \\ & \searrow B_t & \downarrow \pi_t^1 \\ & & [0, \infty) \end{array}$$

交换图得出, 其中 $\pi_t^1 : (B, C) \mapsto B_t$ 是从 $D[0, \infty) \times D[0, \infty)$ 到 $[0, \infty)$ 的可测投影. 同样, 我们可以使用另一个可测投影 $\pi_t^2 : (B, C) \mapsto C_t$ 从 $D[0, \infty) \times D[0, \infty)$ 到 $[0, \infty)$, 并利用类似的交换关系, 得出 C 也是适应的. \square

定理 5.19 的证明. 设 $(N, A) = \{N(\omega, t), A(\omega, t), F, P\}$ 为考克斯过程. 如果 F^N 是由 $N(\cdot, s), 0 \leq s \leq t$ 生成的自然滤子, 且 $A^N(\omega, t)$ 是 $A(\omega, t)$ 在 F^N 上的可料投影 (参见注记 5.1 中的定义), 则 $N(\omega, t) - A^N(\omega, t)$ 是 F^N -鞅. 根据引理 5.17, 存在一个 F^N -可料集合 $E \subseteq \Omega \times [0, \infty)$, 使得 $N(\omega, t)$ 在该集合上分解为

$$N^s(\omega, t) = \int_0^t \mathbf{1}_E(\omega, s) N(\omega, ds), \quad N^c(\omega, t) = N(\omega, t) - N^s(\omega, t).$$

此外, 点过程的勒贝格奇异分量 $N^s(\omega, t)$ 具有 F^N -可料补偿子 $A^{N,s}(\omega, t)$. 运用杜布-梅耶分解定理于 $N^s(\omega, t)$, 可得分解公式 (5.19). 如果 $N^c(\omega, t)$ 与 $A^s(\omega, t)$ 独立, 则 $N^c(\omega, t)$ 也与 E 独立, 因为 E 是 $A^s(\omega, t)$ 的支撑集. 这意味着 F^N -可料集 E 分解为 $E = \Omega \times E'$, 其中 E' 是 $[0, \infty)$ 的子集. 故 E 是 F^{N^c} -可料的, 且引理 5.16 适用于 $N^c(\omega, t)$, 也就是说, $N^c(\omega, t)$ 是非齐次泊松过程. \square

引理 5.18. 点过程 $N(\omega, t)$ 在区间

$$(s, t), \quad [s, t), \quad (s, t], \quad [s, t], \quad s < t$$

上几乎必然是常数, 当且仅当其补偿子 $A(\omega, t)$ 也是如此.

引理5.18的证明. 设 $\bar{N}(\omega, t) = N(\omega, t) - N(\omega, s)$. 如果 $N(\omega, t)$ 几乎必然是常数, 则 $\bar{N}(\omega, t)$ 是一个恒等于零的点过程. 根据杜布-梅耶分解定理, 其补偿子恒等于零. 反之, 如果补偿子 $A(\omega, t)$ 几乎必然是常数, 则 $\bar{N}(\omega, t)$ 是一个补偿子为 $\bar{A}(\omega, t) \equiv 0$ 的点过程. 根据引理5.2中的命题(1), $\bar{N}(\omega, t) \equiv 0$. \square

定理5.19. 霍克斯-考克斯过程是非齐次泊松过程.

定理5.19的证明. 假设 $\underline{N} = \{N(\omega, t), F, P\}$ 是一个霍克斯过程. 我们将证明 $\phi \equiv 0$. 如果存在 $t_0 \geq 0$ 使得 $\phi(t_0) \neq 0$, 则存在 $\delta > 0$ 使得 $\phi > 0$ 在 $[t_0, t_0 + \delta]$ 上成立, 因为 ϕ 在 $[0, \infty)$ 上连续. 我们断言点过程 $N(\omega, t)$ 在区间 $(0, t_0 + \delta]$ 上以正概率发生跳跃, 因为否则 $A(\omega, t + \delta) > 0$, 这将与引理5.18相矛盾. 然而这是不可能的, 因为 $N(\omega, t)$ 会在区间 $(0, t_0 + \delta]$ 内的某个时刻 t' 发生跳跃, 而 $\phi(t - t') \neq 0$ 在 $[t' + t_0, t' + t_0 + \delta] \subseteq (t_0, 2(t_0 + \delta)]$ 上成立. 因此 $\lambda(\omega, t)$ 在 $(t_0, 2(t_0 + \delta)]$ 上不是确定性的, 与引理5.16矛盾. \square

最后, 我们给出点过程的重要性采样与拟合优度检验算法的伪代码.

Algorithm 1 一般点过程的重要性采样

Require: 目标点过程 N_t , 目标强度 λ_t^P , 提议强度 λ_t^Q , 样本数为 M

Ensure: 目标过程下统计量的无偏估计

- 1: **for** $m = 1$ to M **do**
- 2: 根据提议强度 λ_t^Q 模拟样本路径 $\{z_1^{(m)}, \dots, z_{n_m}^{(m)}\}$ ($z = (t, x)$)
- 3: 使用吉萨诺夫公式计算似然权重
- 4: **if** 存在簇或树状结构 **then**
- 5: 应用簇分解或边缘校正保护区域来调整路径
- 6: **end if**
- 7: 存储加权样本 $(\{z_1^{(m)}, \dots, z_{n_m}^{(m)}\}, W^{(m)})$
- 8: **end for**
- 9: 计算过程统计量 f 的蒙特卡罗估计量

$$\hat{E}[f(N)] = \frac{1}{M} \sum_{m=1}^M W^{(m)} f(\{z_1^{(m)}, \dots, z_{n_m}^{(m)}\})$$

Algorithm 2 纯时点过程的拟合优度检验

Require: 观测点过程事件时间 t_1, t_2, \dots, t_n , 拟合强度函数 $\lambda(t)$, 显著性水平 α

Ensure: 检验结果和 p 值

- 1: 计算累积强度函数 $\Lambda(t) = \int_0^t \lambda(s)ds$
 - 2: 计算变换后的时间 $\tau_i = \Lambda(t_i)$, 其中 $i = 1, \dots, n$
 - 3: 计算变换后的间隔时间 $s_i = \tau_i - \tau_{i-1}$ for $i = 1, \dots, n$ ($\tau_0 = 0$)
 - 4: 计算经验分布函数 $F_n(s)$ 和理论分布函数 $F(s) = 1 - e^{-s}$ (单位指数分布)
 - 5: 计算 K-S 统计量 $D = \sup_s |F_n(s) - F(s)|$
 - 6: 通过查表计算 p 值: 比较 D 与 K-S 分布在样本大小 n 下的临界值
 - 7: **if** p 值 $< \alpha$ **then**
 - 8: 拒绝 H_0 , 接受 H_1 : 数据不是来自某一点过程
 - 9: **else**
 - 10: 接受 H_0 : 数据来自某一点过程
 - 11: **end if**
 - 12: **返回** p 值和检验结果
-

第6章 天文案例分析

在光学波段的天文观测中, 图像数据通常以灰度矩阵的形式呈现。特别是在可见光和近红外等波段, 观测到的信号往往会受到天空背景、探测器噪声等因素的影响。尤其是低光强或弱光源条件下, 来自天体的少量光子会对观测数据产生较大影响。在这种情况下, 泊松统计方法变得尤为重要, 且由于天体辐射的特点, 光子的数量通常是稀疏的。因此, 在进行源检测时, 泊松分布提供了一个较为合适的模型, 能够有效地分离噪声与信号。^[103]

元素丰度与银河系演化巡天 (Stellar Abundance and Galactic Evolution Survey, SAGES) 是一个利用 SAGES 测光系统开展的北天多波段测光巡天。本数据包括 DR1(斯图尔德天文台基特峰站 2.3 米博克望远镜观测的 uv 波段) 和 Dr1s(在南山一米望远镜观测的 gri 波段)。国家天文台 2024 年公开的 SAGES 巡天第一次数据发布及其增补超过 10.5GB。该项调查旨在研究恒星丰度和星系演化, 可在线供公众访问[<https://nadu.china-vo.org/res/r100876>]^[104-105]。我们运用相关工具对获取的 u, v, g, r, i 波段进行了点过程拟合优度检验和判别分析, 根据从 Gaia 数据库中查询到的数据, 判别分析的结果显示了 RA=25.21, DEC=34.17 天区中候选老年红超巨星在 SAGES 数据集中对应的 ID。

6.1 图像去噪

我们这里以 i 波段为例进行描述统计分析。原始图像中亮度信息是对天体光学强度的直接反映, 其记录了该探测器单元上到达的光子数。天文图像数据中的噪声来源有多种, 我们所考虑的数据集当中的噪声主要是来自宇宙射线的背景噪声与来自地球大气层的外部噪声。我们通过二维高斯核对原始图像进行卷积, 可以实现对原始图像的平滑, 作为对真实数据的预处理。设 I 为原始图像, 平滑后的图像 I' 可以通过下面的公式计算得到

$$I'(x, y) = \sum_{u=-k}^k \sum_{v=-k}^k I(x-u, y-v) \cdot G(u, v)$$

其中 k 为调节参数, G 是二维高斯核函数。

6.2 局部峰值检测

接下来我们对平滑后的图像进行局部峰值检测

$$I(i, j) > I(i - 1, j), \quad I(i, j) > I(i + 1, j),$$

$$I(i, j) > I(i, j - 1), \quad I(i, j) > I(i, j + 1),$$

$$I(i, j) > \delta * I_{\max}$$

其中 δ 是我们选定的阈值, I_{\max} 是平滑后图像像素的最大值. 通过引入与图像数据相同尺寸的二进制掩码矩阵, 其中值为 1 的区域代表感兴趣的目标区域, 值为 0 的区域代表需要忽略的背景. 将掩码与图像数据逐像素相乘, 再对掩码区域中的非零像素值求和以计算该区域内的总光子计数.

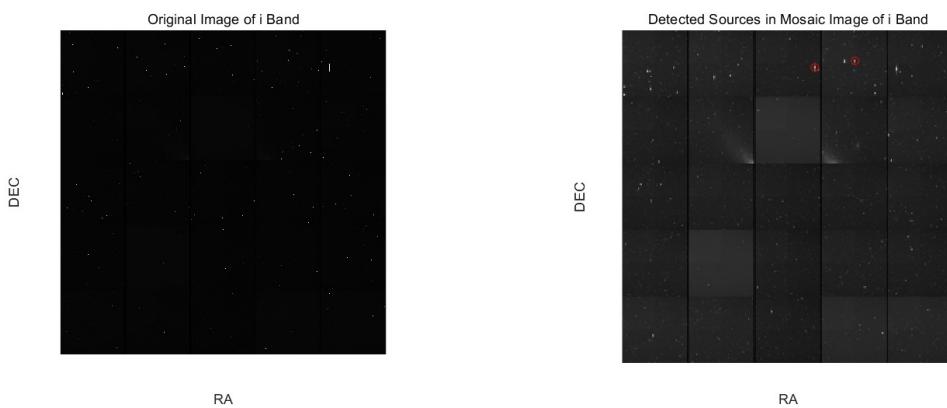


图 6.1 i 波段源检测的对比图像

这里用到的方法事实上是基于天文图像处理中计数值通常是整数, 这是由光子检测的离散性和探测器的工作原理决定的, 因为天文图像通常记录的是个别光子的到达事件, 统计特性服从泊松分布. 根据这一方法, 我们可以通过求泊松分布累计分布函数的逆来给出置信水平 α 下检测的置信区间.

表 6.1 源检测泊松逆分布的置信区间

Threshold = 0.95 for i Band

Source 1 in Mosaic Image: Counts = 69526, 95% CI = [69010, 70043]

Source 2 in Mosaic Image: Counts = 70775, 95% CI = [70254, 71297]

6.3 拟合优度检验

利佩雷 K 统计量是用于分析二维点模式是否具有完全空间随机性 (CSR) 的工具. 其公式为

$$\hat{K}(t) = \frac{1}{\text{区域面积}} \sum_{i \neq j} I(d_{ij} < t)$$

其中, d_{ij} 表示点 i 和点 j 之间的距离, $I(d_{ij} < t)$ 是指在距离 t 内的点对. 对于 CSR 点模式, $\hat{K}(t)$ 应满足 $\hat{K}(t) \approx \pi t^2$. 偏离该期望值的结果表明点模式不满足 CSR, 可能存在簇聚或排斥现象. 我们对平滑后的 i 波段图像进行利佩雷 K 统计分析.

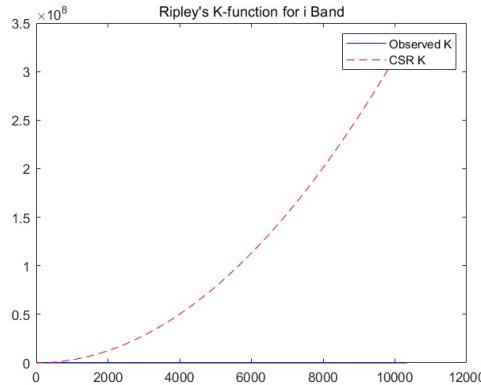


图 6.2 利佩雷 $\hat{K}(t)$ 统计量.

从图像可以看出, CSR 呈现二次函数增长的趋势, 而 i 波段的亮度则始终保持一个低水平. 描述性分析的结果显示点模式不满足 CSR, 所以我们还需要做进一步的分析. 通过绘制不同波段星等和有效温度的热力图, 我们发现有效温度与 u, v, g, r, i 各个波段的星等具有很强的相关性, 这示意我们通过前面介绍的多元统计分析方法来研究它们之间的关系.

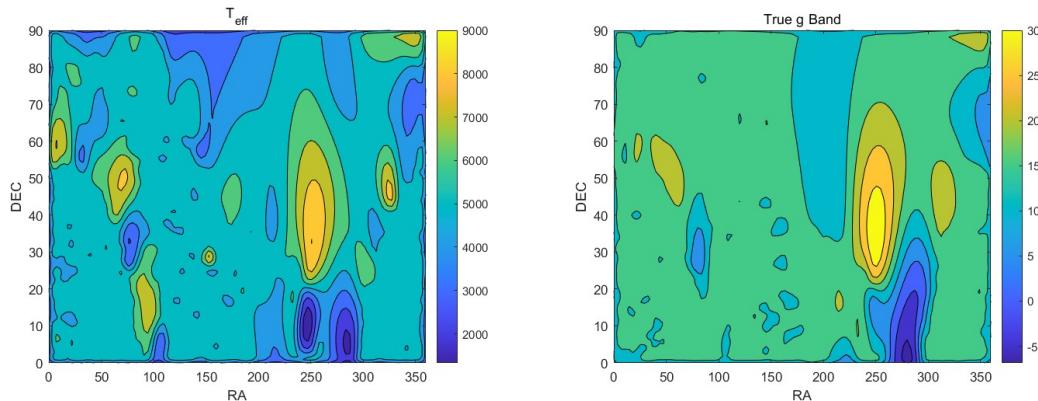


图 6.3 有效温度与 g 波段的对比图像

6.4 判别分析

根据描述性统计分析, 我们将在赤道坐标为 RA = 25.21 和 Dec = 34.17 的天区内推测候选天体. 考虑恒星有效温度的对数 $\log T_{\text{eff}}$ 对 u, v, g, r, i 波段星等

$$\log T_{\text{eff}} \sim u + v + g + r + i$$

在 RA = 25.21 和 Dec = 34.17 天区中关于不同观测时间的主成分回归分析.

第 1 主成分: 贡献率 = 82.40%, 累计贡献率 = 82.40%

第 2 主成分: 贡献率 = 14.79%, 累计贡献率 = 97.19%

第 3 主成分: 贡献率 = 1.42%, 累计贡献率 = 98.61%

第 4 主成分: 贡献率 = 0.77%, 累计贡献率 = 99.39%

第 5 主成分: 贡献率 = 0.61%, 累计贡献率 = 100.00%

其中第一主成分为 g, 第二主成分为 g-r, 前两个主成分累计荷载达到了 95%.

根据回归分析的结果, 我们做了系数是否显著区别于零的 F 检验. 方差分析的结果显示不同波段的星等具有显著性差异.

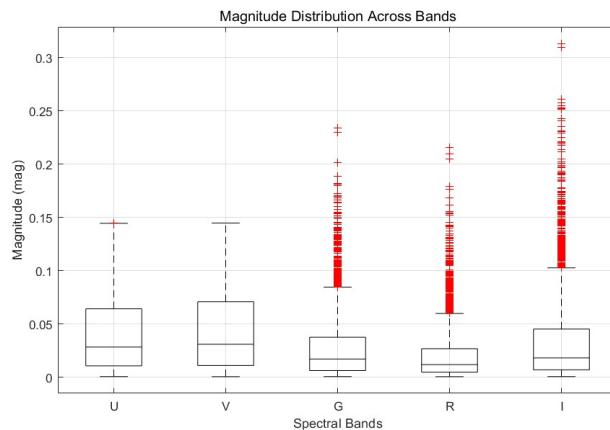


图 6.4 单因素方差分析.

通过 Gaia 获得的老年红超巨星数据我们训练了线性判别模型, 并在由实际数据构成的测试集上进行了预测. 结果显示了 12 个可能的老年红超巨星候选, 我们将其对应的天体 ID、赤经 ($^{\circ}$)、赤纬 ($^{\circ}$)、Gaia G 星等、Gaia BP-RP 颜色指数和基于后验判别概率计算的置信度 (%) 列在下面的表格中, 供天文学研究者参考.

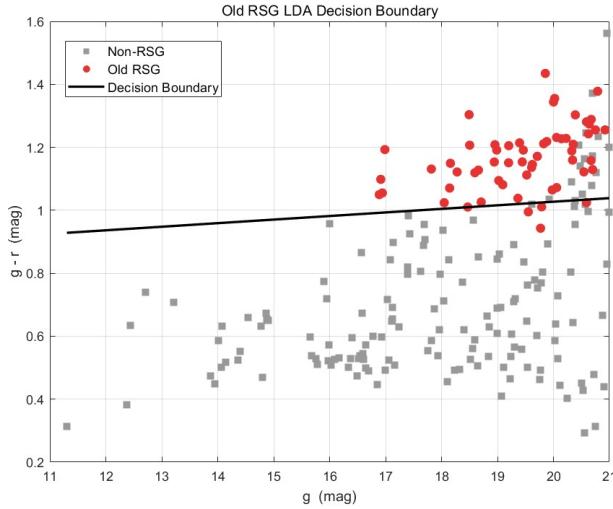


图 6.5 Gaia 训练集上的线性判别决策边界.

表 6.2 RA=25.21,DEC=34.17 天区老年红超巨星判别结果表

天体 ID	赤经 (°)	赤纬 (°)	Gaia G 星等	Gaia BP-RP 颜色指数	置信度 (%)
SAGE014230.5+342136	384.41	34.36	19.59	0.93	99.986
SAGE014251.1+333515	385.69	33.59	18.47	2.04	59.802
SAGE014231.2+342258	384.45	34.38	19.59	0.93	99.986
SAGE014231.9+342614	384.49	34.44	19.59	0.93	99.986
SAGE014231.4+342351	384.46	34.40	19.59	0.93	99.986
SAGE014226.9+340443	384.18	34.08	19.44	0.99	99.987
SAGE014226.3+340147	384.15	34.03	19.44	0.99	99.987
SAGE014224.4+341139	384.02	34.19	19.44	0.99	99.987
SAGE014224.9+340315	384.06	34.05	19.44	0.99	99.987
SAGE014224.0+340734	384.00	34.13	19.44	0.99	99.987
SAGE014226.5+340434	384.16	34.08	19.44	0.99	99.987
SAGE014230.0+341919	384.37	34.32	19.59	0.93	99.986

注: 1. 赤经单位为度 (°), 由时角 (1 小时 = 15°) 转换而来. 2. BP-RP 颜色指数越大, 恒星表面温度越低 (越“红”). 3. 置信度基于线性判别模型后验概率计算, $\geq 90\%$ 为高可靠判别, $< 60\%$ 为低可靠判别.

参考文献

- [1] KARHUNEN K. Über lineare methoden in der wahrscheinlichkeitsrechnung[J]. Ann. Acad. Sci. Fennicae, 1947, 37(1).
- [2] LOÈVE M. Probability theory I - II[M]. 4th ed. Springer-Verlag New York, 1977, 1978.
- [3] GUPTA A K, NAGAR D K. Matrix variate distributions[M]. Chapman and Hall/CRC, 2018.
- [4] MATHAI A M, PROVOST S B, HAUBOLD H J. Multivariate statistical analysis in the real and complex domains[M]. Springer Nature, 2022.
- [5] ANDERSON T W. An introduction to multivariate statistical analysis: Vol. 2[M]. Wiley New York, 1958.
- [6] SRIVASTAVA M S, KHATRI C G. An introduction to multivariate statistics[M]. Elsevier Science Ltd, 1979.
- [7] MUIRHEAD R J. Aspects of multivariate statistical theory[M]. John Wiley & Sons, 1982.
- [8] WEYL H. The classical groups: their invariants and representations: Vol. 45[M]. Princeton university press, 1946.
- [9] 华罗庚. 多复变数函数论中的典型域的调和分析[M]. 科学出版社, 1958.
- [10] MACDONALD I G. Symmetric functions and Hall polynomials[M]. Oxford university press, 1998.
- [11] JAMES A T. Normal multivariate analysis and the orthogonal group[J]. Ann. Math. Stat., 1954, 25(1): 40-75.
- [12] JAMES A T. The distribution of the latent roots of the covariance matrix[J]. Ann. Math. Stat., 1960, 31: 151-158.
- [13] JAMES A T. The distribution of noncentral means with known covariance[J]. Ann. Math. Stat., 1961, 32: 874-882.
- [14] JAMES A T. Zonal polynomials of the real positive definite symmetric matrices[J]. Ann. Math., 1961, 74(2): 456-469.
- [15] JAMES A T. Distributions of matrix variates and latent roots derived from normal samples [J]. Ann. Math. Stat., 1964, 35: 475-501.

- [16] CONSTANTINE A G. Some non-central distribution problems in multivariate analysis[J]. Ann. Math. Stat., 1963, 34(4): 1270-1285.
- [17] CONSTANTINE A G. The distribution of Hotelling's generalized T_0^2 [J]. Ann. Math. Statist., 1966, 37: 215-225.
- [18] TAKEMURA A. Zonal polynomials: Vol. 4[M]. Inst. Math. Stat., 1984.
- [19] GROSS K I, RICHARDS S P, D. Special functions of matrix argument. I: Algebraic Induction, Zonal Polynomials, and Hypergeometric Functions[J]. Trans. Amer. Math. Soc., 1987, 301(2): 781-811.
- [20] SHIMIZU K. Distribution theory of eigenvalues for a singular beta-Wishart matrix[D]. Tokyo Rika University, 2022.
- [21] RICHARDS D S P. Functions of matrix argument[M]//NIST Digital Library of Mathematical Functions. 2024.
- [22] KAKUTANI S. On equivalence of infinite product measures[J]. Ann. Math., 1948, 49(1): 214-224.
- [23] GIRSANOV I V. On transforming a certain class of stochastic processes by absolutely continuous substitution of measures[J]. Theor. Prob. Appl., 1960, 5(3): 285-301.
- [24] CLARK J M. The representation of functionals of Brownian motion by stochastic integrals [J]. Ann. Math. Stat., 1970: 1282-1295.
- [25] NOVIKOV A A. On an identity for stochastic integrals[J]. Theor. Prob. Appl., 1972, 17(4): 761-765.
- [26] KAZAMAKI N. Continuous exponential martingales and BMO: Vol. 1579[M]. Springer-Verlag, 1994.
- [27] PROTTER P E. Stochastic integration and differential equations[M]. 2nd ed. Springer, 2005.
- [28] JEULIN T. Semi-martingales et grossissement d'une filtration: Vol. 833[M]. Springer, 2006.
- [29] TELLA P D, JEANBLANC M. Martingale representation in the enlargement of the filtration generated by a point process[J]. Stoch. Proc. Appl., 2021, 131: 103-121.
- [30] JACOD J. Multivariate point processes: predictable projection, Radon-Nikodym derivatives, representation of martingales[J]. Z. für W., 1975, 31: 235-253.
- [31] BRÉMAUD P M. Point processes and queues[M]. Springer, 1981.

- [32] DALEY D J, VERE-JONES D. An introduction to the theory of point processes Vol.I-II[M]. Springer, 2003, 2007.
- [33] KARR A. Point processes and their statistical inference[M]. Routledge, 2017.
- [34] BRÉMAUD P M. A martingale approach to point processes: UCB/ERL M345[D]. EECS Department, University of California, Berkeley, 1972.
- [35] 王浩铭. 关于点过程的极小可料强度[J]. 休斯顿数学杂志, 2025.
- [36] STUDENT. The probable error of a mean[J]. Biometrika, 1908, 6(1): 1-25.
- [37] FISHER R A. Frequency distribution of the values of the correlation coefficient in samples from an indefinitely large population[J]. Biometrika, 1915, 10(4): 507-521.
- [38] PEARSON K. Further contributions to the theory of small samples[J]. Biometrika, 1925, 17 (1-2): 176-200.
- [39] ROMANOVSKY V. On the moments of the standard deviation and of the correlation coefficient in samples from normal[J]. Metron, 1925, 5(4): 3-46.
- [40] WISHART J. The generalised product moment distribution in samples from a normal multivariate population[J]. Biometrika, 1928, 20A(1/2): 32-52.
- [41] ANDERSON T W. The non-central Wishart distribution and certain problems of multivariate statistics[J]. Ann. Math. Stat., 1946, 17: 409-431.
- [42] JAMES A T. The non-central Wishart distribution[J]. Proc. Roy. Soc. Lond. A, 1955, 229: 364 - 366.
- [43] KHATRI C G. On certain distribution problems based on positive definite quadratic functions in normal vectors[J]. Ann. Math. Stat., 1966, 37(2): 468 - 479.
- [44] KOTZ S, JOHNSON N L, BOYD D W. Series representations of distributions of quadratic forms in normal variables. I: Central case. II: Non-central case[J]. Ann. Math. Stat., 1967, 38 (3): 823-837, 838 - 848.
- [45] GRAYBILL F A, MILLIKEN G. Quadratic forms and idempotent matrices with random elements[J]. Ann. Math. Stat., 1969, 40(4): 1430-1438.
- [46] CACOULLOS T, KOUTRAS M. Quadratic forms in spherical random variables: Generalized noncentral χ^2 distribution[J]. Naval Res. Logist. Quart., 1984, 31(3): 447-461.
- [47] 方开泰, 范剑青, 许建伦. 随机投影阵的二次型分布及其应用[J]. 应用概率统计, 1987, 3 (4): 289-297.

- [48] MATHAI A. On bilinear forms in normal variables[J]. Ann. Inst. Stat. Math., 1992, 44: 769-779.
- [49] 夏应存. 椭球等高分布族中 T_0^2 的精确分布[J]. 应用概率统计, 1994, 10(2): 119-124.
- [50] 秦怀振, 崔恒建, 李勇. 二次型之比服从 F 分布的充分必要条件[J]. 应用概率统计, 1998, 14(1): 79-84.
- [51] CARO-LOPERA F, GONZÁLEZ-FARÍAS G, BALAKRISHNAN N. On generalized Wishart distributions I: Likelihood ratio test for homogeneity of covariance matrices, II: Sphericity test[J]. Sankhyā A, 2014, 76(2): 179-194, 195-218.
- [52] SHIMIZU K, HASHIGUCHI H. Heterogeneous hypergeometric functions with two matrix arguments and the exact distribution of the largest eigenvalue of a singular beta-Wishart matrix[J]. J. Multivar. Anal., 2021, 183: 104714.
- [53] 方开泰, 张尧庭. 广义多元分析[M]. 科学出版社, 1993.
- [54] PHILLIPS P C B. The distribution of matrix quotients[J]. J. Multivar. Anal., 1985, 16(1): 157-161.
- [55] DÍAZ-GARCÍA J A, GUTIÉRREZ-JÁIMEZ R. Matricvariate and matrix multivariate t distributions and associated distributions[J]. Metrika, 2012, 75(7): 963-976.
- [56] DUNNETT C W, SOBEL M. A bivariate generalization of Student's t -distribution, with tables for certain special cases[J]. Biometrika, 1954, 41(1-2): 153-169.
- [57] JOHN S. On the evaluation of the probability integral of the multivariate t-distribution[J]. Biometrika, 1961, 48: 409-417.
- [58] KSHIRSAGAR A M. Some extensions of the multivariate t -distribution and the multivariate generalization of the distribution of the regression coefficient[J]. Math. Proc. Cambridge, 1961, 57: 80 - 85.
- [59] DICKEY J M. Matricvariate generalizations of the multivariate t -distribution and the inverted multivariate t -distribution[J]. Ann. Math. Stat., 1967, 38(2): 511 - 518.
- [60] LIN P E. Some characterizations of the multivariate t distribution[J]. J. Multivar. Anal., 1972, 2(3): 339-344.
- [61] KOTZ S, NADARAJAH S. Multivariate t distributions and their applications[M]. Cambridge University Press, 2004.

- [62] KHATRI C G. On the exact finite series distribution of the smallest or the largest root of matrices in three situations[J]. *J. Multivar. Anal.*, 1972, 2(2): 201-207.
- [63] DAVIS A W. On the marginal distributions of the latent roots of the multivariate beta matrix [J]. *Ann. Math. Stat.*, 1972, 43(5): 1664-1670.
- [64] HIRAKAWA P. The distribution of the latent roots in manova model[J]. *Tokyo Rika Univ. Math.*, 1976, 12(2): 65-73.
- [65] CHIKUSE Y. Distributions of some matrix variates and latent roots in multivariate Behrens-Fisher discriminant analysis[J]. *Ann. Stat.*, 1981, 9(2): 401 - 407.
- [66] GUPTA A K, KABE D G. A multiple integral involving zonal polynomials[J]. *Appl. Math. Lett.*, 2004, 17(6): 671-675.
- [67] LÄUTER J, GLIMM E, KROPPF S. Multivariate tests based on left-spherically distributed linear scores[J]. *Ann. Stat.*, 1998, 26(5): 1972-1988.
- [68] ROY S N. On a heuristic method of test construction and its use in multivariate analysis[J]. *Ann. Math. Stat.*, 1953, 24(2): 220-238.
- [69] KRISHNAIAH P R, CHANG T C. On the exact distribution of the smallest root of the Wishart matrix using zonal polynomials[J]. *Ann. Inst. Stat. Math.*, 1970: 293-295.
- [70] JOHNSTONE I M, NADLER B. Roy' s largest root test under rank-one alternatives[J]. *Biometrika*, 2017, 104(1): 181-193.
- [71] MONROE I. Processes that can be embedded in Brownian motion[J]. *Ann. Prob.*, 1978, 6 (1): 42 - 56.
- [72] DELLACHERIE C. Capacités et processus stochastiques: Vol. 67[M]. Springer-Verlag, Berlin., 1972.
- [73] 王浩铭. 一类时空点过程分离性的假设检验[D]. 中山大学, 2021.
- [74] SCHOENBERG I J. Positive definite functions on spheres[J]. *Duke Math. J.*, 1942, 9(1): 96 - 108.
- [75] BERG C, PORCU E. From Schoenberg coefficients to Schoenberg functions[J]. *Constr. Approx.*, 2017, 45.
- [76] TUMURA Y. The distributions of latent roots and vectors[J]. *Tokyo Rika Univ. Math.*, 1965.
- [77] VRETARE L. Recurrence formulas for zonal polynomials[J]. *Math. Z.*, 1985, 188: 1432-1823.

- [78] KUSHNER H B. The linearization of the product of two zonal polynomials[J]. SIAM J. Math. Anal., 1988, 19(3): 687-717.
- [79] JAMES A T. Calculation of zonal polynomial coefficients by use of the Laplace-Beltrami operator[J]. Ann. Math. Stat., 1968, 39: 1711-1718.
- [80] HERZ C S. Bessel functions of matrix argument[J]. Ann. Math., 1955, 61(3): 474-523.
- [81] DAVIS A W. Invariant polynomials with two matrix arguments extending the zonal polynomials: Applications to multivariate distribution theory[J]. Ann. Inst. Stat. Math., 1979, 31: 465-485.
- [82] DAVIS A W. On the construction of a class of invariant polynomials in several matrices, extending the zonal polynomials[J]. Ann. Inst. Stat. Math., 1981, 33: 297-313.
- [83] DAWID A P. Spherical matrix distributions and a multivariate model[J]. J. R. Stat. Soc. B, 1977, 39(2): 254-261.
- [84] KABE D G. Classical statistical analysis based on a certain hypercomplex multivariate normal distribution[J]. Metrika, 1984, 31(1): 63-76.
- [85] GOODALL C R, MARDIA K V. Multivariate aspects of shape theory[J]. Ann. Stat., 1993: 848-866.
- [86] DYKSTRA R L. Establishing the positive definiteness of the sample covariance matrix[J]. Ann. Math. Stat., 1970, 41(6): 2153-2154.
- [87] HSU P L. A new proof of the joint product moment distribution[J]. Math. Proc. Phil. Soc., 1939, 35(2): 336–338.
- [88] EATON M. Multivariate statistics: A vector space approach[M]. Institute of Mathematical Statistics, 1983.
- [89] PEDDADA S D, RICHARDS D S P. Proof of a conjecture of M. L. Eaton on the characteristic function of the Wishart distribution[J]. Ann. Prob., 1991, 19(2): 868-874.
- [90] LETAC G, MASSAM H. The Laplace transform $(\det s)^{-p} \exp \text{tr}(s^{-1}w)$ and the existence of non-central Wishart distributions[J]. J. Multivar. Anal., 2018, 163(1): 96-110.
- [91] MAYERHOFER E. On Wishart and noncentral Wishart distributions on symmetric cones [J]. Trans. Amer. Math. Soc., 2019, 371(10): 7093-7109.

- [92] SUGIYAMA T. Distribution of the largest latent root and the smallest latent root of the generalized B statistic and F statistic in multivariate analysis[J]. Ann. Math. Stat., 1967, 38(4): 1152-1159.
- [93] WILKS S S. Certain generalizations in the analysis of variance[J]. Biometrika, 1932, 24(3/4): 471-494.
- [94] HOTELLING H. The generalization of Student's ratio[J]. Ann. Math. Stat., 1931, 2: 360-387.
- [95] RÉNYI A. Remarks on the Poisson process[J]. Stud. Math. Hungarica, 1967, 2: 119-123.
- [96] DELLACHERIE C, MEYER P A. Probabilities and potential. Vol. I-III[M]. North-Holland, 1978-1988.
- [97] JACOD J, SHIRYAEV A. Limit theorems for stochastic processes: Vol. 288[M]. Springer Science & Business Media, 2013.
- [98] 屈田兴, 金治明. 关于半鞅的可料表示性[J]. 国防科技大学报, 2010, 32(1): 159-162.
- [99] 何声武, 汪嘉冈, 严加安. 半鞅与随机分析[M]. 科学出版社, 1995.
- [100] PROHOROV Y V. Convergence of random processes and limit theorems in probability theory [J]. Theor. Prob. Appl., 1956, 1(2): 157-214.
- [101] SKOROKHOD A V. Limit theorems for stochastic processes[J]. Theor. Prob. Appl., 1956, 1(3): 261-290.
- [102] BILLINGSLEY P. Convergence of probability measures[M]. John Wiley & Sons, 2013.
- [103] BABU G J, FEIGELSON E D. Spatial point processes in astronomy[J]. J. Stat. Plan. Infer., 1996, 50(3): 311-326.
- [104] 范舟, 赵刚, 王炜, 赵景昆, 郑捷, 谈克峰, 刘玉娟, 何维, 宋轶晗, 姜晓军, 马路. SAGE 巡天介绍 I: 测光系统和数据处理[J]. 天文学进展, 2018, 36(2): 101-112.
- [105] 赵景昆, 赵刚, 范舟, 谈克峰, 宋轶晗, 王奇勋, 王炜. SAGE 巡天介绍 II: 恒星大气参数估计[J]. 天文学进展, 2020, 38(1): 60-68.

