多元统计推断

纳拉扬·吉里 著
张尧庭 等 译

一九七九年四月

目　录

序 临

本书是多元分析（特别是使用不变性方法的多元正态分布）理论与应用两个方面的最新介绍。它是为具有大学数学和统计知识的读者编写的。书中各处均用取自应用领域的生动数据加以说明。鉴于多元分析入门教程的一般特点，我们试图多收入一些与特定论题有关的例子与提示。这里介绍的内容是由作者的一个多元分析研究生一学年课程的讲稿发展起来的，该讲稿曾在蒙特利尔大学等处多次使用。每章都附有大量有关的问题和完备的文献，每章末尾的习题还收入了有关复多元正态总体的类似结果。

不变性是关于某些变换的对称性的数学术语。不变性的概念在统计推断中由来已久。Hunt和Stein的第二次世界大战结束前夕未出版的著作，即竭力推崇这一概念在统计检验一般类的结构中的应用前景与深远意义。现在不变性概念已成为证明许多统计检验方法的优良性的十分有力的工具。如果一个具有唯一解的问题在某一变换下是不变的，那末这个解在变换下也应当是不变的，这是一个公认的原理。借助不变性讨论多元分析的另一令人信服的理由是，常用的检验法则大部分是以似然比检验原理为基础的；在关于所论参数空间和概率密度函数的一个较弱的限制下，似然比检验对于保持检验问题不变的变换群是几乎不变的。

多元分析这一广阔领域的取材与论述颇不容易。在这里，作者以自己在统计方面讲授各种水平的研究生课程及大学课程的经验为指导，亦以自己十五年来在这个领域从事并指导研究

工作的经验为依据。本书介绍了多元分析的基本工具，讨论了它们的理论基础，这样就使读者有可能为在这个领域从事进一步探讨研究作好准备。

第I章包括一些有关特征根、特征向量、实和复矩阵的分块子矩阵的特殊结果，还包括在多元分析中常用的关于实和复矩阵的一些特殊定理。

第II章讨论群论及对不变统计检验法则的研究颇为有用的有关结果，还给出了一些对推导多元样本分布有用的特殊变换的雅可比行列式。

第III章专门讨论多元分布的基本概念和统计假设检验中的不变性原理。对于不变性与充分性，不变性与无偏性，不变性与最优检验，不变性与最紧检验之间的相互关系，本章也作了研究。

第IV章通过概率密度函数，通过一个简单的特征化研究实的多元正态分布。第二种方法简化了多元理论，且由此能够从一元理论不经进一步分析即得出适当的推广。本章还包括实多元正态分布、凝聚椭球与轴、回归、多重相关与偏相关等方面的一些特征化。复多元正态分布的类似结果也包括在内。

多元正态分布参数的极大似然估计及其优良性构成了第V章的主题。

第VI章包括实数情形下基本多元样本分布的系统推导。这些结果的复数情形作为问题列入。

具有已知和未知协方差矩阵的多元正态总体的均值向量的检验法与置信域以及它们的优良性在第VII章讨论。

第VIII章专门研究与多元正态总体的协方差矩阵和均值向量相联系的检验法的系统推导以及它们的优良性。

第IX章包括判别分析的近代论述。判别分析的简要历史也

包括在内。

第五章包括关于复多元正态分布的不同协方差模型及其分析。主分量模型，因子模型，典型相关与时间序列也在这里讲述。

我们认为本书宜作为多元分析的基础教材，每周3学时，两学期授完。

第1章 向量与矩阵代数

1.0 引 言

多元分析的研究需要向量和矩阵代数的知识，有关向量和矩阵的一些基本结果将在本章讨论。这些结果中有的仅予陈述而不给证明，证明可从 Giri（1974），Mac Lane 与 Birkoff（1967），Markus 与 Mine（1967），Perlis（1952）或关于矩阵代数的任何课本中获得。

1.1 向 量

向量是有序的 p- 元组 $x_1, \cdots, x_p$，记成

$$x = \begin{pmatrix} x_1 \\ \vdots \\ x_p \end{pmatrix}$$

具体地称为 p- 维列向量，为简单起见，称为 p- 向量或向量。x 的转置记为 $x' = (x_1, \cdots, x_n)$。若一向量的一切分量都是零，称为零向量 0。几何上，一个 p- 向量表示 p- 维欧氏空间 E^p 的一个点 $A = (x_1, \cdots, x_p)$ 或以 A 为端点的有向线段

$\overline{OA}$。全体 p-向量的集合记为 V^p。显然若向量的全部分量是实数，则 $V^p = E^p$。对于任何两个向量 $x=(x_1,\cdots,x_p)'$ 与 $y=(y_1,\cdots,y_p)'$，我们定义向量和 $x+y=(x_1+y_1,\cdots,x_p+y_p)'$ 且由一常数 a 的数乘为

$$ax=(ax_1,\cdots,ax_p)'.$$

显然，向量加法是可结合与可交换的运算，即 $x+y=y+x$，$(x+y)+z=x+(y+z)$，其中 $z=(z_1,\cdots,z_p)'$，且数乘是一可分配的运算，即对常数 a，b，$(a+b)x=ax+bx$。对于 x，$y\in V^p$，$x+y$ 与 ax 也属于 V^p。因此，对常数 a，b，$a(x+y)=ax+ay$，且 $a(bx)=b(ax)=abx$。

量 $x'y=y'x=\sum_i^p x_iy_i$ 称作 V^p 中两个向量 x，y 的内积。向量 x 与它自身的内积记作 $\|x\|^2=x'x$，其中 $\|x\|$ 称作 x 的范数。范数的一些几何意义是：

(i) $\|x\|^2$ 是在 V^p 中点 x 到原点的距离的平方，

(ii) 两点 $(x_1,\cdots,x_p)$，$(y_1,\cdots,y_p)$ 之间的距离的平方由 $\|x-y\|^2$ 给出，

(iii) 两个向量 x，y 之间的夹角 θ 由 $\cos\theta=(x/\|x\|)'(y/\|y\|)$ 给定。

定义 1.1.1 <u>正交向量</u> V^p 中两个向量 x，y 是正交的，当且仅当 $x'y=y'x=0$。V^p 中的向量组是正交的，如果其中的向量两两正交。

几何上看，两个向量 x，y 正交，当且仅当 x，y 的夹角是 $90°$。一个正交向量 x，如果 $\|x\|^2=1$，称为标准正交向量。

定义 1.1.2 <u>向量的投影</u> x，$y\in V^p$，x 在 y $(\neq 0)$ 上的投影为 $\|y\|^{-2}(x'\cdot y)y$。

若 $\overline{OA}=x$，$\overline{OB}=y$，且 P 是点 A 到 OB 的垂足，则 $\overline{OP}=\|y\|^{-2}(x'y)y$，其中 O 是 E^p 的原点。对于两个正交的向量 x，

y，x到y上的投影为零。

定义1.1.3　V^p的一组向量$\alpha_1,\cdots,\alpha_R$，如果无一向量能表示成其它向量的线性组合，则称为是线性无关的。

于是，若$\alpha_1,\cdots,\alpha_R$线性无关，则不存在不全为零的常数$c_1,\cdots,c_R$，满足$c_1\alpha_1+\cdots+c_R\alpha_R=0$．可以验证$V^p$中的一组正交向量是线性无关的。

定义1.1.4　由一个向量组张成的向量空间　设$\alpha_1,\cdots,\alpha_R$是V^p中向量组，则由一切能表示成$\alpha_1,\cdots,\alpha_R$的线性组合的向量与零向量构成的集合叫做由$\alpha_1,\cdots,\alpha_R$张成的向量空间V.

于是，若$\alpha,\beta\in V$，则对常数a,b，$a\alpha+b\beta$与$a\alpha$也属于V．此外，由$\alpha_1,\cdots,\alpha_p\in V^p$，$\alpha_1,\cdots,\alpha_p$的任何线性组合也属于$V^p$，故得$V\subset V^p$．所以$V$是$V^p$的线性子空间。

定义1.1.5　向量空间的基　向量空间V的基就是张成V的线性无关向量组。

V^p中单位向量$\varepsilon_1=(1,0,\cdots,0)'$，$\varepsilon_2=(0,1,0,\cdots,0)'$，$\cdots$，$\varepsilon_p=(0,\cdots,0,1)'$形成$V^p$的基。

如果$\alpha_1,\cdots,\alpha_R$张成V，则$\alpha_1,\cdots,\alpha_R$的子集形成V的基。

定理1.1.1　每个向量空间V都有基，且V的两个基有同样多个元素。

定理1.1.2　设向量空间V由向量$\alpha_1,\cdots,\alpha_R$张成。任一元素$\alpha\in V$能唯一地表示成$\alpha=\sum_1^R c_i\alpha_i$（$c_1,\cdots,c_R$是不全为零的常数），当且仅当$\alpha_1,\cdots,\alpha_R$是$V$的基。

定义1.1.6　向量的坐标　若$\alpha_1,\cdots,\alpha_R$是向量空间V的基，且若$\alpha\in V$被唯一地表示为$\alpha=\sum_1^R c_i\alpha_i$，$c_1,\cdots,c_R$是常数，则向量$\alpha_i$的系数$c_i$称作$\alpha$关于基$\alpha_1,\cdots,\alpha_R$的第$i$个坐标。

定义1.1.7　向量空间的秩　向量空间的基中向量的个数

称作V的秩或维数。

1.2 矩 阵

定义1.2.1 矩阵 实矩阵A是元素a_{ij}（实数）的一个有次序的矩形排列

$$A=\begin{pmatrix} a_{11} & \cdots & a_{1q} \\ \vdots & & \vdots \\ a_{p1} & \cdots & a_{pq} \end{pmatrix} \cdots\cdots\cdots\cdots\cdots\cdots (1.1)$$

且记为$A_{p\times q}=(a_{ij})$.

p行q列的矩阵称作$p\times q$维（p乘q）矩阵，行数总写在前边。若$p=q$，则称之为p维方阵。p-维列向量是$p\times 1$维矩阵。同样维数的两个矩阵$A_{p\times q}$，$B_{p\times q}$，如果$a_{ij}=b_{ij}$，$i=1,\cdots,p$，$j=1,\cdots,q$，则称它们是相等的（记作$A=B$）。如果a_{ij}全下是零，则称A为零矩阵，记为粗体的零 —— 0。$p\times q$维矩阵A的转置是$q\times p$维矩阵A'：

$$A'=\begin{pmatrix} a_{11} & \cdots & a_{p1} \\ \vdots & & \vdots \\ a_{1q} & \cdots & a_{pq} \end{pmatrix} \cdots\cdots\cdots\cdots\cdots\cdots (1.2)$$

它是由A的行列交换而得到的。显然$(A')'=A$. 方阵A若有$A=A'$则称为是对称的，若$A=-A'$，就是反对称的。反对称矩阵的对角线元素是零。有时我们用记号"$p\times q$维矩阵"A而不用$A_{p\times q}$。

对任意两个同样维数$p\times q$的矩阵$A=(a_{ij})$与$B=(b_{ij})$，我们定义矩阵和$A+B$为$p\times q$维矩阵$(a_{ij}+b_{ij})$. 矩阵$A-B$看作与$A+B$意义相同，只是在用(+)号的地方改用(−)号。显然$(A+B)'=A'+B'$，$A+B=B+A$，且对任意

三个矩阵 A, B, C，有 $(A+B)+C=A+(B+C)$. 于是矩阵和运算是可交换和可结合的。

对任一矩阵 A 和任一常数 C，数乘 CA 定义为 $CA=AC=(Ca_{ij})$. 显然 $(CA)'=CA'$，数乘是可分配的运算。

两个矩阵 $A_{p\times q}=(a_{ij})$ 与 $B_{q\times r}=(b_{ij})$ 的矩阵积是矩阵 $C_{p\times r}=AB=(C_{ij})$，其中

$$c_{ij}=\sum_{k=1}^{q} a_{ik}b_{kj}, \quad i=1,\cdots,p, \ j=1,\cdots,r. \qquad \cdots\cdots(1.3)$$

积 AB 当 A 的行数等于 B 的列数时才有定义，一般地 $AB\neq BA$. 此外，$(AB)'=B'A'$. 当积都有定义时矩阵的积是满足分配律和结合律的，即对任意三个矩阵 A, B, C，

(i) $A(B+C)=AB+AC$（分配律），

(ii) $(AB)C=A(BC)$（结合律）.

定义 1.2.2 对角阵 如果一个方阵的对角线以外的元素都是零，则称方阵为对角阵。

定义 1.2.3 单位阵 对角线元素全部是 1 的对角阵，称作单位阵，记为 I。

对任一方阵 A，$AI=IA=A$.

定义 1.2.4 三角阵 A 是方阵，如果 $j<i$ 时，$a_{ij}=0$ 则称 A 为上三角阵，如果 $j>i$ 时，$a_{ij}=0$，则称 A 为下三角阵。

定义 1.2.5 正交阵 A 是方阵，如果 $AA'=A'A=I$，则称 A 为正交阵。

对应于任意 $p\times p$ 维方阵有一个数量 $|A|$，或 $\det A$，称作 A 的行列式，定义为

$$|A|=\sum_{\pi}\sigma(\pi)a_{1\pi(1)}a_{2\pi(2)}\cdots a_{p\pi(p)}, \qquad \cdots\cdots(1,4)$$

其中π取遍列指标$(1, 2, \cdots, p)$的全部$p!$个置换，$\delta(\pi)=1$，当从标准次序到$\pi(1), \cdots, \pi(p)$逆序的个数是偶数时；$\delta(\pi)=-1$，当这种逆序的个数是奇数时。在一个实际的置换中，逆序的个数就是将一个元素放到按标准次序$1, 2, \cdots, p$应在它前面的数的后面去的总次数。从第Ⅲ章起，我们统一用符号$\det A$表示行列式，把$\|$当作绝对值的符号。

定义1.2.6　子式与代数余子式　对任一$p\times p$维矩阵$A=(a_{ij})$．元素a_{ij}的子式是划去A的i行与j列后所形成的矩阵的行列式，以$(-1)^{i+j}$乘a_{ij}的子式得到的量，称作a_{ij}的代数余子式，记为A_{ij}。

其对角线元素也是A的对角线元素的（A的）$i\times i$维子矩阵的行列式称作A的i阶主子式。领先的主子式集就是阶分别为$1, 2, \cdots, p$的p个主子式的集合，其中i阶主子式的矩阵是$i+1$阶主子式的矩阵的子矩阵，$i=1, \cdots, p-1$。

容易验证，对任一$p\times p$维方阵$A=(a_{ij})$，

$$|A|=\sum_{j=1}^{p}a_{ij}A_{ij}=\sum_{i=1}^{p}a_{ij}A_{ij}, \quad\cdots\cdots\cdots\cdots(1.5)$$

而对$j\neq j', i\neq i'$，

$$\sum_{i=1}^{p}a_{ij}A_{ij'}=\sum_{j=1}^{p}a_{ij}A_{i'j}=0\ 。\quad\cdots\cdots\cdots\cdots(1.6)$$

此外，若A对称，则$A_{ij}=A_{ji}$对一切i, j成立。对一具有对角线元素a_{ii}的$p\times p$维三角阵或对角阵A，$|A|=\prod_{i=1}^{p}a_{ii}$。如果$A$的任二列（或行）交换，则$|A|$只改变符号；若$A$的二列（或行）相等或成比例，则$|A|=0$。

定义1.2.7　非奇异阵　A是方阵，若$|A|\neq 0$，则称A是非奇异阵。若$|A|=0$，则称A为奇异阵。

非奇异阵的列与行都是线性无关的。因为对任二方阵 A, B，$|AB|=|A||B|$，我们得到两个非奇异阵的积是非奇异阵。然而两个非奇异阵的和不一定是非奇异阵。这样一个平庸的例子就是 $A=-B$，其中 A、B 都是非奇异阵。

定义 1.2.8 逆矩阵 $p\times p$ 维非奇异阵 A 的逆是唯一的矩阵 A^{-1}，使得 $A^{-1}A=AA^{-1}=I$。

设 A_{ij} 是 A 的元素 a_{ij} 的代数余子式，且

$$C=\begin{pmatrix} \frac{A_{11}}{|A|} & \cdots & \frac{A_{1p}}{|A|} \\ \vdots & & \vdots \\ \frac{A_{p1}}{|A|} & \cdots & \frac{A_{pp}}{|A|} \end{pmatrix}, \qquad (1.7)$$

从 (1.6) 与 (1.7) 我们得 $AC'=I$．因此 $A^{-1}=C'$．矩阵的逆仅对非奇异阵有定义．且若 A 是对称的则 A^{-1} 也是对称的。此外 $|A^{-1}|=(|A|)^{-1}$，$(A')^{-1}=(A^{-1})'$ 且 $(AB)^{-1}=B^{-1}A^{-1}$。

1.3 矩阵的秩与迹

设 A 是 $p\times q$ 维矩阵。设 $R(A)$ 是由 A 的行所张成的向量空间，$C(A)$ 是由 A 的列所张成的向量空间。空间 $R(A)$ 称作 A 的行空间，且它的秩 $r(A)$ 叫作 A 的行秩。空间 $C(A)$ 称作 A 的列空间，而它的秩 $c(A)$ 叫作 A 的列秩。对任何矩阵 A，$r(A)=c(A)$。

定义 1.3.1 矩阵的秩 矩阵 A 的行秩与列秩的共同值称作矩阵的秩，且记为 $\rho(A)$。

对任意 $p\times q$ 维矩阵，若 $q<p$，$\rho(A)$ 可从 0 到 q 变化。如果 $\rho(A)=q$，则 A 称作满秩矩阵。零矩阵 0 的秩是 0。

对任意两个使 AB 有定义的矩阵 A，B，AB 的列是 A 的列的线性组合。于是 AB 的线性无关的列的个数不能超过 A 的线性无关的列的个数，因此 $\rho(AB)\leqslant\rho(A)$。类似地，考虑 AB 的行可论证 $\rho(AB)\leqslant\rho(B)$。因此 $\rho(AB)\leqslant\min(\rho(A),\rho(B))$.

定理 1.3.1　若 A, B, C 分别是 $p\times q$ 维，$p\times p$ 维，$q\times q$ 维矩阵，且 B, C 是非奇异阵[①]则 $\rho(A)=\rho(AC)=\rho(BA)=\rho(BAC)$.

定义 1.3.2　矩阵的迹　$p\times p$ 维方阵 $A=(a_{ij})$ 的迹定义为它的对角线元素的和，记为 $trA=\sum_1^p a_{ii}$.

显然 $trA=trA'$，$tr(A+B)=tr(A)+tr(B)$，此外 $trAB=trBA$，只要 AB，BA 均有定义。因此对任何正交矩阵 θ，$tr\theta'A\theta=tr\,A\theta\theta'=trA$.

1.4　二次型和正定矩阵

实变易 $x_1,\cdots,x_p$ 的二次型是一个形如 $Q=\sum_{i=1}^p\sum_{j=1}^p\cdot a_{ij}x_ix_j$ 的表示式，其中 a_{ij} 是实常数。

记 $x=(x_1,\cdots,x_p)'$，$A=(a_{ij})$，我们可记 $Q=x'Ax$，不失一般性，我们可取二次型 Q 中的矩阵 A 是对称阵。因 Q 是一数易，有

$$Q=Q'=x'A'x=\frac{1}{2}(Q+Q')=x'((A+A')/2)x,$$

且 $\frac{1}{2}(A+A')$ 是对称阵。

定义 1.4.1　正定阵　一个方阵 A 或相应的二次型 $x'Ax$，若对一切 $x\neq0$ 都有 $x'Ax>0$，则称为正定的。如果对一切 x，都有 $x'Ax\geqslant0$，则称为半正定的。

---

① B，C 是非奇异阵这一条件是我们加的。无此条件，定理显然不能成立。——译者注。

矩阵A或相应的二次型$x'Ax$，当$-x'Ax$是正定或半正定时，分别称为负定的或半负定的。

例 1.4.1

$$(x_1, x_2)\begin{pmatrix} 2 & 1 \\ 1 & 3 \end{pmatrix}(x_1, x_2)' = 2x_1^2 + 2x_1x_2 + 3x_2^2$$
$$= 2(x_1 + \frac{1}{2}x_2)^2 + \frac{5}{2}x_2^2 > 0$$

对一切$x_1 \neq 0$或$x_2 \neq 0$成立，因此矩阵$\begin{pmatrix} 2 & 1 \\ 1 & 3 \end{pmatrix}$是正定的。

1.5 特征根与特征向量

$p \times p$维的方阵$A = (a_{ij})$的特征根，就是特征方程

$$|A - \lambda I| = 0 \quad \cdots\cdots\cdots\cdots (1.8)$$

的根，其中λ是实的[①]。显然，这是λ的p次方程，故恰好有p个根。如果A是对角阵，则其对角线元素就是它的特征根。一般地，可将(1.8)写成

$$(-\lambda)^p + (-\lambda)^{p-1}S_1 + (-\lambda)^{p-2}S_2 + \cdots + (-\lambda)S_{p-1} + |A| = 0 \quad \cdots\cdots\cdots (1.9)$$

其中S_i是A的一切i阶主子式的和。特别有$S_1 = trA$. 于是A的特征根的积等于$|A|$，A的特征根的和等于t_rA. 满足

$$(A - \lambda I)x = 0 \quad (1.10)$$

的非零向量$x = (x_1, \cdots, x_p)'$称作矩阵A关于特征根λ的特征向量。很明显，若x是矩阵A关于特征值λ的特征向量，则任一数乘$cx\ (c \neq 0)$，也是A关于特征值λ的特征向量。因为对任何$p \times p$维正交矩阵Θ，

---

① 这里的λ可以取复数。否则下面有p个根的结论就不一定成立。——译者注。

$$|\Theta A\Theta' - \lambda I| = |\Theta A\Theta' - \lambda\Theta\Theta'| = |A - \lambda I|,$$

矩阵A的特征根关于变换$A \longrightarrow \Theta A\Theta'$是保持不变的。

定理 1.5.1 如果A是（$p \times p$ 阶的）实对称矩阵，则它的特征根都是实数。

证明 设λ是A的复特征根，且设$x + iy$，$x = (x_1, \cdots, x_p)'$，$y = (y_1, \cdots, y_p)'$是关于λ的特征向量（复的）。则由（1.10）

$$A(x + iy) = \lambda(x + iy), \quad (x - iy)'A(x + iy) = \lambda(x'x + y'y),$$

但是

$$(x - iy)'A(x + iy) = x'Ax + y'Ay,$$

因此我们得到λ必定是实数。证完。

注意 关于复特征根的特征向量Z一定是复的，否则从$AZ = \lambda Z$会推出实向量等于复向量。

定理 1.5.2 一个对称矩阵的关于不同特征根的特征向量是正交的。

证明 设λ_1，λ_2是对称（实）矩阵A的两个不同的特征根，且设$x = (x_1, \cdots, x_p)'$，$y = (y_1, \cdots, y_p)'$分别是关于λ_1，λ_2的特征向量，则

$$Ax = \lambda_1 x, \quad Ay = \lambda_2 y.$$

所以

$$y'Ax = \lambda_1 y'x, \quad x'Ay = \lambda_2 x'y.$$

于是

$$\lambda_1 x'y = \lambda_2 x'y$$

因为$\lambda_1 \neq \lambda_2$，故得 $x'y = 0$. 证完。

设λ是对称正定矩阵A的特征根，设x是关于λ的特征向量，则