高等学校试用教材

多元分析

成都地质学院　王柏钧　程积康　编

地质出版社

高等学校试用教材

多 元 分 析

成都地质学院

王柏钧 程积康 编

地 质 出 版 社

多　元　分　析

成都地质学院

王柏钧　程积康　编

地质矿产部教育司教材室编辑

责　任　编　辑：林　清　湜

地质出版社　出　版

（北京西四）

地质出版社印刷厂印刷

（北京海淀区学院路29号）

新华书店北京发行所发行·全国新华书店经售

开本：850×1168 1/32　印张：7 1/2　字数：194,000

1982年　7月北京第一版·1982年7月北京第一次印刷

印数：1—3,380册　定价：0.92元

统一书号：15038·教143

前　言

一　《多元分析概论》是为数学地质专业学习地质多元统计分析所开设的一门专业基础课，其目的是为了适当加强有关多元地质统计的理论基础。它是在学习了线性代数和概率论等数学课程以后开设的。考虑到地质院校数学地质专业学生和地质系研究生的实际数学水平，在内容选取上力求适当而不过于求全和求深。此书是在数学地质专业用的《多元分析概论》讲义的基础上稍加扩充改写而成的。此书亦可供有关专业参考。

二　本书共分九章和两个附录。它是在多元正态分布模型的基础上进行叙述的，内容有多元正态分布与常用的一元非中心分布；Wishart 分布与 Hotelling 分布；多元回归与相关；独立性检验及正态分布的相等性检验；判别分析；主分量分析；因子分析；典则相关与典则变量；集群分析等。另外考虑到多元分析需要用到较多的矩阵理论和一个因变量的回归理论，而有的学生可能还不完全具备这些知识，因而书本附矩阵理论和一个因变量的多元回归两个附录作为阅读时参考查阅之用。有关多元分析方法在地质方面的应用和程序，我们拟在后继有关多元地质统计分析的讲义中编写，故在本书中没有涉及，有兴趣的读者也可参阅目前已出版的有关数学地质多元统计的书藉。

三　本书的编写大纲曾由吉林大学数学系赵文、南京大学数学系洪再吉、云南大学数学系何湘藩、新疆工学院地质系徐仲平、昆明工学院地质系冉崇英、成都地质学院数学地质专业冯宗礼等同志共同审订。由成都地质学院王柏钧、程积康负责执笔编写。本书的主审单位为吉林大学数学系概率统计教研室，主审人为吉林大学周光亚付教授和赵文老师。对同志们详细审阅初稿和提出的宝贵意见，我们在此表示由衷的感谢。

四　本书编写内容主要参考和取材于以下的一些著作：

1．中国科学院数学研究所：多元分析资料汇编Ⅱ、Ⅲ。

2．N. Giri：多元统计推断，（张尧亭等译）。

3．T. W. Anderson：线性统计推断及其应用。（云南大学数学系译）。

4．四川大学概率论教研室：多元分析中的几个主要分布。

5．C. R. Rao：线性统计推断及其应用。

6．吉林大学概率统计教研室：多元分析资料汇编。

7．K. Fukunaga：统计图形识别导论。

五、有关本教材的习题题解，请参阅吉林大学概率统计教研室编《多元统计分析习题选解》。

六、武汉大学张尧亭老师、应用数学研究所方开泰同志对本书的编曾给予很大的关心和支持，在此深表谢意。

七、由于多元分析这一广阔领域的取材和论述很不容易，而我们在这方面无论业务水平还是经验均感不足，肯定缺点错误不少，敬请同志们评批指正。

编者

1981 年 11 月

目　录

第一章　多元正态分布

§1　记 号 与 准 备

一、记号

在本章介绍多元正态分布及其性质时，用到较多矩阵运算的知识，有关矩阵的一些知识在附录中作了一些介绍。现将常用到的一些矩阵记号简介如下：

随机向量：

$$\boldsymbol{x}=\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_p \end{pmatrix},\quad \boldsymbol{y}=\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_q \end{pmatrix}$$

等，其中 x_i、y_j 为随机变量。

数学期望向量：

$$E\boldsymbol{x}=\begin{pmatrix} Ex_1 \\ Ex_2 \\ \vdots \\ Ex_p \end{pmatrix}=\begin{pmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_p \end{pmatrix}=\boldsymbol{\mu},$$

样本均值向量：

$$\bar{\boldsymbol{x}}=\begin{pmatrix} \bar{x}_1 \\ \bar{x}_2 \\ \vdots \\ \bar{x}_p \end{pmatrix},\ \text{其中}\ \bar{x}_i=\frac{1}{N}\sum_{k=1}^{N}x_{ik}\quad i=1,\cdots,p$$

自协方差阵：

$$\cos(\boldsymbol{x},\boldsymbol{x})=(\sigma_{ij})\quad \text{其中}\quad \sigma_{ij}=Ex_ix_j-(Ex_i)(Ex_j)$$

本书也常用 $\boldsymbol{\Sigma}$ 或 $D(\boldsymbol{x})$ 表示 $\operatorname{cov}(\boldsymbol{x},\boldsymbol{x})$．显然

$$D(\boldsymbol{x})=E(\boldsymbol{x}-E\boldsymbol{x})(\boldsymbol{x}-E\boldsymbol{x})'=E\boldsymbol{x}\boldsymbol{x}'-E\boldsymbol{x}E\boldsymbol{x}'.$$

样本自协方差阵：

$$\boldsymbol{S}=(S_{ij})．\text{其中 } S_{ij}=\frac{\left(\sum_{k=1}^{N}x_{ik}x_{jk}-N\bar{x}_i\bar{x}_j\right)}{(N-1)}.$$

$\boldsymbol{S}$ 也可表示为

$$\boldsymbol{S}=\frac{1}{N-1}\sum_{k=1}^{N}(\boldsymbol{x}_k-\bar{\boldsymbol{x}})(\boldsymbol{x}_k-\bar{\boldsymbol{x}})'=\frac{1}{N-1}\left(\sum_{k=1}^{N}\boldsymbol{x}_k\boldsymbol{x}_k'-N\bar{\boldsymbol{x}}\bar{\boldsymbol{x}}'\right).$$

互协方差阵：

$$\operatorname{cov}(\boldsymbol{x},\boldsymbol{y})=(\sigma_{ij}')，\text{其中 } \sigma_{ij}'=Ex_iy_j-(Ex_i)(Ey_j).$$

$\operatorname{cov}(\boldsymbol{x},\boldsymbol{y})$ 还可以简写为 $c(\boldsymbol{x},\boldsymbol{y})$，且

$$\operatorname{cov}(\boldsymbol{x},\boldsymbol{y})=E(\boldsymbol{x}-E\boldsymbol{x})(\boldsymbol{y}-E\boldsymbol{y})'=E\boldsymbol{x}\boldsymbol{y}'-(E\boldsymbol{x})(E\boldsymbol{y})'.$$

样本互协方差阵：

$$\boldsymbol{S}'=(S_{ij}')，\quad \text{其中 } S_{ij}'=\frac{1}{N-1}\left(\sum_{k=1}^{N}x_{ik}y_{jk}-N\bar{x}_i\bar{y}_j\right).$$

也还可以表为

$$\boldsymbol{S}'=\frac{1}{N-1}\sum_{k=1}^{N}(\boldsymbol{x}_k-\bar{\boldsymbol{x}})(\boldsymbol{y}_k-\bar{\boldsymbol{y}})'=\frac{1}{N-1}\left(\sum_{k=1}^{N}\boldsymbol{x}_k\boldsymbol{y}_k'-N\bar{\boldsymbol{x}}\bar{\boldsymbol{y}}'\right).$$

范围空间：命 $\boldsymbol{A}=(\boldsymbol{a}_1,\boldsymbol{a}_2,\cdots\cdots,\boldsymbol{a}_N)$，而所谓 $\boldsymbol{A}$ 的范围空间 $M(\boldsymbol{A})$，是指由矩阵 $\boldsymbol{A}$ 的列向量所张成的空间。即

$$M(\boldsymbol{A})=\left\{\boldsymbol{x}\,\middle|\;\boldsymbol{x}=\sum_{i=1}^{N}\lambda_i\boldsymbol{a}_i，\ \lambda_i \text{ 为实数}\right\}.$$

还有其它一些常用的矩阵和向量的记号：

对角阵：

$$\operatorname{diag}(\lambda_1,\lambda_2,\cdots,\lambda_n)=\begin{pmatrix}\lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n\end{pmatrix}=\boldsymbol{\Lambda}.$$

向量 $\mathbf{1}$：

$$\mathbf{1}_n=\begin{pmatrix}1\\1\\\vdots\\1\end{pmatrix}_{n\times 1}$$

矩阵 $\boldsymbol{J}$：

$$\boldsymbol{J}=\mathbf{1}_n\mathbf{1}_n'=\begin{pmatrix}1 & 1 & \cdots\cdots & 1\\1 & 1 & \cdots\cdots & 1\\\vdots & \vdots & & \vdots\\1 & 1 & \cdots\cdots & 1\end{pmatrix}_{n\times n}$$

正定与半正定性：

$\boldsymbol{\Sigma}>0$　表示 $\boldsymbol{\Sigma}$ 为正定矩阵，

$\boldsymbol{\Sigma}\geqslant 0$　表示 $\boldsymbol{\Sigma}$ 为半正定矩阵。

二、随机向量与随机矩阵

上面引进了一些常用的记号，如随机向量、数学期望向量、自协方差阵等，在此简述一下有关的定义。

设有 p 个随机变量 $x_1, x_2, \cdots\cdots, x_p$。它们所构成的向量 $\boldsymbol{x}$ 叫做随机向量，而 $m\times n$ 个随机变量所构成的矩阵

$$\boldsymbol{X}=(x_{ij})\quad i=1,\cdots\cdots,m;\ j=1,\cdots\cdots,n.$$

叫做随机矩阵。

〔定义〕　设有随机向量 $\boldsymbol{x}=(x_1,x_2,\cdots\cdots,x_p)'$，如果 $Ex_i(i=1,\cdots\cdots,p)$ 均存在，则称 $E\boldsymbol{x}=(Ex_1,\cdots\cdots,Ex_p)'$ 为随机向量 $\boldsymbol{x}$ 的数学期望，用 $E\boldsymbol{x}$ 表示。如果随机矩阵 $\boldsymbol{X}$ 的各元素的数学期望均存在，则称 $E\boldsymbol{X}=(Ex_{ij})$ 为随机矩阵 $\boldsymbol{X}$ 的数学期望。

〔定义〕　设有随机向量 $\boldsymbol{x}$、$\boldsymbol{y}$，它们的数学期望为 $E\boldsymbol{x}$、$E\boldsymbol{y}$，则称 $E(\boldsymbol{x}-E\boldsymbol{x})(\boldsymbol{y}-E\boldsymbol{y})'$ 为随机向量 $\boldsymbol{x}$ 和 $\boldsymbol{y}$ 的互协方差阵（如果存在的话），记为 $\mathrm{cov}(\boldsymbol{x},\boldsymbol{y})$。当 $\boldsymbol{x}=\boldsymbol{y}$ 时，称之为 $\boldsymbol{x}$ 的自协方差阵，记为 $D(\boldsymbol{x})$，亦即 $\mathrm{cov}(\boldsymbol{x},\boldsymbol{x})$。

由于多元分析是在学习了概率论的基础上进行的，为了简便和节省篇幅起见。除上面例举的多元随机向量有关的几个定义以

外，今后对其他有关多元分布的一些基本概念：如边际（缘）分布、条件分布、各阶矩、独立性等的有关定义，就不再加以一一叙述了。

三、有关的简单运算

有关随机矩阵的一些简单运算，作为性质罗列如下，请读者自行验证。

〔性质 1〕 $\boldsymbol{A}=(a_{ij})$，$\boldsymbol{B}=(b_{ij})$，$\boldsymbol{C}=(c_{ij})$ 为三个常数矩阵，$\boldsymbol{x}$ 为随机变量组成的随机矩阵，

则
$$E(\boldsymbol{AxB}+\boldsymbol{C})=\boldsymbol{A}(E\boldsymbol{x})\boldsymbol{B}+\boldsymbol{C}.$$
在这里当然假设这些矩阵的阶是可以进行运算的，下面性质中记号的解释均类似。

〔性质 2〕 $E(\boldsymbol{Ax}+\boldsymbol{By})=\boldsymbol{A}(E\boldsymbol{x})+\boldsymbol{B}(E\boldsymbol{y}).$

〔性质 3〕 $D(\boldsymbol{x}+\boldsymbol{a})=D(\boldsymbol{x}).$

〔质性 4〕 $\operatorname{cov}(\boldsymbol{Ax},\boldsymbol{By})=\boldsymbol{A}\operatorname{cov}(\boldsymbol{x},\boldsymbol{y})\boldsymbol{B}'.$

〔性质 5〕
$$\begin{aligned}\operatorname{cov}(a\boldsymbol{x}+b\boldsymbol{y},c\boldsymbol{u}+d\boldsymbol{v})=&ac\operatorname{cov}(\boldsymbol{x},\boldsymbol{u})\\&+ad\operatorname{cov}(\boldsymbol{x},\boldsymbol{v})+bc\operatorname{cov}(\boldsymbol{y},\boldsymbol{u})\\&+bd\operatorname{cov}(\boldsymbol{y},\boldsymbol{v}).\end{aligned}$$

〔性质 6〕
$$\begin{aligned}D(a\boldsymbol{x}+b\boldsymbol{y})=&a^2D(\boldsymbol{x})+b^2D(\boldsymbol{y})\\&+ab\operatorname{cov}(\boldsymbol{x},\boldsymbol{y})+ab\operatorname{cov}(\boldsymbol{y},\boldsymbol{x}).\end{aligned}$$

以上 $\boldsymbol{x}$、$\boldsymbol{y}$、$\boldsymbol{u}$、$\boldsymbol{v}$ 是随机向量或矩阵，而 a、b、c、d 为常数。

§2 多元正态分布

一、定义

正态分布在多元分析中如同它在概率论中一样起着非常重要的作用。而关于它定义的方法，一种是用密度定义，其缺点是要求有 $\boldsymbol{\Sigma}^{-1}$；一种是用它的特征函数来定义，虽然它不要求 $\boldsymbol{\Sigma}^{-1}$，但却需要对特征函数加以说明。还有用随机向量分量的线性组合是一元正态变量来加以定义的。这种方式较为简单且具有一般性，但有时用它来证明其性质时，因要说明其分量的线性组合为

正态，推导较繁。下面介绍的是一种刻划正态变量的另一种简单方式。

〔定义〕 若 $x_1, x_2, \cdots\cdots, x_p$ 独立同分布 $N(0,1)$，则称 $\boldsymbol{x}' = (x_1, \cdots\cdots, x_p)$ 为标准正态变量。

〔定义〕 设 $\boldsymbol{x}$ 为标准正态变量，$\boldsymbol{A}$ 为 $n \times p$ 的矩阵，$\boldsymbol{\mu}$ 为 $n \times 1$ 的常向量，若 $\boldsymbol{y} = \boldsymbol{A}\boldsymbol{x} + \boldsymbol{\mu}$，则称 $\boldsymbol{y}$ 遵从多维正态分布，记为 $\boldsymbol{y} \sim N_n(\boldsymbol{\mu}, \boldsymbol{\Sigma})$，此处 $\boldsymbol{\Sigma} = \boldsymbol{A}\boldsymbol{A}'$ 或简记为 N （$\boldsymbol{\mu}$，$\boldsymbol{\Sigma}$）。

显然由定义有

$$E\boldsymbol{x} = 0, \quad \text{而 } D(\boldsymbol{x}) = \boldsymbol{I},$$

及

$$E\boldsymbol{y} = \boldsymbol{\mu}, \quad \text{而 } D(\boldsymbol{y}) = \boldsymbol{\Sigma}.$$

〔定理 1〕 若 $\boldsymbol{z} = \boldsymbol{P}\boldsymbol{y} + \boldsymbol{a}$，而 $\boldsymbol{y} \sim N_n(\boldsymbol{\mu}, \boldsymbol{\Sigma})$，则

$$\boldsymbol{z} \sim N_n(\boldsymbol{a} + \boldsymbol{P}\boldsymbol{\mu}, \boldsymbol{P}\boldsymbol{\Sigma}\boldsymbol{P}').$$

此定理可由定义直接验证。

〔定理 2〕 若 $\boldsymbol{y} \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$，今将 $\boldsymbol{y}$ 剖分为两部分

$$\boldsymbol{y} = \begin{pmatrix} \boldsymbol{y}_{(1)} \\ \boldsymbol{y}_{(2)} \end{pmatrix} \begin{matrix} r \text{ 行} \\ n-r \text{ 行} \end{matrix},$$

相应地将 $\boldsymbol{\mu}$ 与 $\boldsymbol{\Sigma}$ 剖分为两部分

$$\boldsymbol{\mu} = \begin{pmatrix} \boldsymbol{\mu}_{(1)} \\ \boldsymbol{\mu}_{(2)} \end{pmatrix} \begin{matrix} r \\ n-r \end{matrix}, \quad \boldsymbol{\Sigma} = \begin{pmatrix} \boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12} \\ \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22} \end{pmatrix} \begin{matrix} r \\ n-r \end{matrix}$$

则 $\boldsymbol{y}_{(1)} \sim N_r(\boldsymbol{\mu}_{(1)}, \boldsymbol{\Sigma}_{11})$，$\boldsymbol{y}_{(2)} \sim N_{n-r}(\boldsymbol{\mu}_{(2)}, \boldsymbol{\Sigma}_{22})$.

〔证明〕 在前定理中取 $\boldsymbol{P} = (\boldsymbol{I}_r, \boldsymbol{0})_{r\times n}$，$\boldsymbol{a} = 0$，

则

$$\boldsymbol{y}_{(1)} = \boldsymbol{P}\boldsymbol{y}$$

而

$$\boldsymbol{P}\boldsymbol{\Sigma}\boldsymbol{P}' = (\boldsymbol{I}_r, \boldsymbol{0}) \begin{pmatrix} \boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12} \\ \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22} \end{pmatrix} (\boldsymbol{I}_r, \boldsymbol{0})' = \boldsymbol{\Sigma}_{11}$$

故

$$\boldsymbol{y}_{(1)} \sim N_r(\boldsymbol{\mu}_{(1)}, \boldsymbol{\Sigma}_{11})。$$

同理取 $\boldsymbol{P}$ 为 $(\boldsymbol{0}, \boldsymbol{I}_{n-r})$ 则得 $\boldsymbol{y}_{(2)} \sim N_{n-r}(\boldsymbol{\mu}_{(2)}, \boldsymbol{\Sigma}_{22})$.

〔定理 3〕 若 $\boldsymbol{y} \sim N_n(\boldsymbol{\mu}, \boldsymbol{\Sigma})$，则它的特征函数

为

$$E[\exp(i\boldsymbol{t}'\boldsymbol{y})] = \varphi_{\boldsymbol{y}}(\boldsymbol{t}) = \exp\left(i\boldsymbol{t}'\boldsymbol{\mu} - \frac{1}{2}\boldsymbol{t}'\boldsymbol{\Sigma}\boldsymbol{t}\right).$$

此处 $\boldsymbol{t}$ 为一实向量。

〔证明〕 在一元正态变量中，$x \sim N(\mu, \sigma^2)$，则它的特征函数为$E(e^{itx}) = \exp\left(it\mu - \frac{1}{2}\sigma^2 t^2\right)$。而当$x \sim N(0,1)$时，它的特征函数为$\exp\left(\frac{1}{2}t^2\right)$。当$\boldsymbol{x}$为多元标准正态变量时，它的特征函数

$$\varphi_x(\boldsymbol{t}) = E(e^{i\boldsymbol{t}'\boldsymbol{x}}) = E(e^{i\sum_\alpha t_\alpha x_\alpha}) = \prod_\alpha E(e^{it_\alpha x_\alpha})$$

$$= \prod_\alpha e^{-\frac{t_\alpha^2}{2}} = \exp\left(-\frac{1}{2}\boldsymbol{t}'\boldsymbol{t}\right).$$

故 $$\varphi_y(\boldsymbol{t}) = E(e^{i\boldsymbol{t}'\boldsymbol{y}}) = E(e^{i\boldsymbol{t}'(\boldsymbol{Ax}+\boldsymbol{\mu})}) = e^{i\boldsymbol{t}'\boldsymbol{\mu}}E(e^{i\boldsymbol{t}'\boldsymbol{Ax}})$$

$$= e^{i\boldsymbol{t}'\boldsymbol{\mu}}e^{-\frac{1}{2}\boldsymbol{t}'\boldsymbol{AA}'\boldsymbol{t}} = e^{i\boldsymbol{t}'\boldsymbol{\mu}-\frac{1}{2}\boldsymbol{t}'\boldsymbol{\Sigma t}}.$$

〔定理 4〕 若$\boldsymbol{y} \sim N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$，且$\boldsymbol{\Sigma} > 0$，则$\boldsymbol{y}$有分布密度为

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

〔证明〕 因$\boldsymbol{y} = \boldsymbol{Ax} + \boldsymbol{\mu}$故$\boldsymbol{y} - \boldsymbol{\mu} = \boldsymbol{Ax}$，而

$$\boldsymbol{x} \sim N_p(\boldsymbol{0}, \boldsymbol{I}_p).$$

由$\boldsymbol{x}$分量的独立性，有$\boldsymbol{x}$的密度

$$f(\boldsymbol{x}) = \frac{1}{(\sqrt{2\pi})^p}\exp\left(-\frac{1}{2}\boldsymbol{x}'\boldsymbol{x}\right) = \frac{1}{(\sqrt{2\pi})^p}\exp\left(-\frac{1}{2}\sum_{\alpha=1}^p x_\alpha^2\right).$$

作变量置换，注意到$\boldsymbol{\Sigma} = \boldsymbol{AA}'$，$\boldsymbol{\Sigma} > 0$，有

$$rk(\boldsymbol{A}) = p,$$

而变换$\boldsymbol{A}^{-1}(\boldsymbol{y}-\boldsymbol{\mu}) = \boldsymbol{x}$的雅可比行列式为$|\boldsymbol{A}^{-1}| = |\boldsymbol{\Sigma}|^{-\frac{1}{2}}$，代入即有定理的结果。上面是$\boldsymbol{A}$矩阵取$n = p$的情况。

对于$p > n$时的情形，亦有相同的结果。其证法为：令

$$\boldsymbol{U} = \begin{pmatrix}\boldsymbol{Ax}\\ \boldsymbol{Bx}\end{pmatrix} \triangleq \boldsymbol{T}\,\boldsymbol{x} = \begin{pmatrix}\boldsymbol{y}-\boldsymbol{\mu}\\ \boldsymbol{Bx}\end{pmatrix} \triangleq \begin{pmatrix}\boldsymbol{U}_{(1)}\\ \boldsymbol{U}_{(2)}\end{pmatrix},$$

其中$\boldsymbol{B}$的取法为使它满足$\boldsymbol{BA}' = 0$，且$\boldsymbol{BB}' = \boldsymbol{I}_{p-n}$.

容易由 $$\boldsymbol{x} = \boldsymbol{T}^{-1}\boldsymbol{U}$$

及 $$TT'=\begin{pmatrix}A\\B\end{pmatrix}(A',B')=\begin{pmatrix}AA' & \\ & BB'\end{pmatrix}=\begin{pmatrix}\Sigma & \\ & I_{p-n}\end{pmatrix},$$

得 $$x'x=U'(T^{-1})'(T)^{-1}U=U'_{(1)}\Sigma^{-1}U_{(1)}+U'_{(2)}U_{(2)}.$$

由 $x\sim N_p(0,I_p)$，作变量置换，其雅可比行列式

为 $$|T^{-1}|=|\Sigma^{-1}|^{\frac{1}{2}}.$$

故 $U=\begin{pmatrix}U_{(1)}\\U_{(2)}\end{pmatrix}$的密度为

$$\frac{1}{(\sqrt{2\pi})^n|\Sigma|^{\frac{1}{2}}}\exp\left(-\frac{(y-\mu)'\Sigma^{-1}(y-\mu)}{2}\right)\frac{1}{(\sqrt{2\pi})^{p-n}}$$
$$\times\exp\left(-\frac{U'_{(2)}U_{(2)}}{2}\right).$$

再对 $U_{(2)}$ 积分，即得证 $p>n$ 时的结果。

〔例〕$y=(y_1,y_2)'\sim N_2(\mu,\Sigma)$ 令 $\rho=\dfrac{\sigma_{12}}{\sigma_1\sigma_2}$,

则 $$\Sigma=\begin{pmatrix}\sigma_{11} & \sigma_{12}\\ \sigma_{21} & \sigma_{22}\end{pmatrix}=\begin{pmatrix}\sigma_1^2 & \sigma_1\sigma_2\rho\\ \sigma_1\sigma_2\rho & \sigma_2^2\end{pmatrix},$$

$$\Sigma^{-1}=\frac{1}{1-\rho^2}\begin{pmatrix}\sigma_1^{-2} & -\dfrac{\rho}{\sigma_1\sigma_2}\\ -\dfrac{\rho}{\sigma_1\sigma_2} & \sigma_2^{-2}\end{pmatrix}.$$

故密度为

$$f(y_1,y_2)=\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}\exp\left\{-\frac{1}{2(1-\rho^2)}\left[\frac{(y_1-\mu_1)^2}{\sigma_1^2}\right.\right.$$
$$\left.\left.-2\rho\left(\frac{y_1-\mu_1}{\sigma_1}\right)\times\left(\frac{y_2-\mu_2}{\sigma_2}\right)+\left(\frac{y_2-\mu_2}{\sigma_2}\right)^2\right]\right\}$$

讨论：1）当 $\rho=0$ 时$\Longleftrightarrow y_1,y_2$ 相互独立。

2）当 $\rho^2=1$ 时$\Longleftrightarrow y_1,y_2$ 有线性关系，

1）的结论显然成立。而在 2）中，由于 $\rho^2=1$

故 $$|\Sigma|=\sigma_1^2\sigma_2^2(1-\rho^2)=0.$$

而Σ为非正定阵，即存在有 t'，使 $t'\Sigma t=0$ ($t'\neq 0$),

亦即 $$\boldsymbol{t}'\boldsymbol{\Sigma}\boldsymbol{t}=E[\boldsymbol{t}'(\boldsymbol{y}-\boldsymbol{\mu})]^2=0$$

故有 $$t_1(y_1-\mu_1)+t_2(y_2-\mu_2)=0\ (a.s.)$$

反之亦然。

二、多元正态变量的性质

下面讨论一些多元正态变量的进一步的性质。

〔定理 5〕 $\boldsymbol{y}'=(\boldsymbol{y}'_{(1)},\boldsymbol{y}'_{(2)})$，遵从正态分布 $N_n(\boldsymbol{\mu},\boldsymbol{\Sigma})$，其中 $\boldsymbol{y}_{(1)},\boldsymbol{y}_{(2)}$ 的维数分别为 r 和 $n-r$。而 $\boldsymbol{\Sigma}_{11},\boldsymbol{\Sigma}_{12},\boldsymbol{\Sigma}_{22}$ 为 $\boldsymbol{\Sigma}$ 相应的分块，则在 $\boldsymbol{y}'_{(1)}$ 给定后，$\boldsymbol{y}'_{(2)}$ 的条件分布为:

$$\boldsymbol{y}_{(2)}\sim N_{n-r}(\boldsymbol{\mu}_{(2)}+\boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-}(\boldsymbol{y}_{(1)}-\boldsymbol{\mu}_{(1)}),\boldsymbol{\Sigma}_{22}-\boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-}\boldsymbol{\Sigma}_{12})$$

其中 $\boldsymbol{\mu}_{(i)}=E\boldsymbol{y}_{(i)}$ $(i=1,2)$，而 $\boldsymbol{\Sigma}_{11}^{-}$ 为 $\boldsymbol{\Sigma}_{11}$ 的广义逆。

〔证明〕 用 $M(\boldsymbol{A})$ 表示 $\boldsymbol{A}$ 的列向量所张成的范围空间。则对任一向量 $\boldsymbol{l}_1\in R_n$，若 $\boldsymbol{l}_1$ 垂直于 $M(\boldsymbol{\Sigma}_n)$，由 $\boldsymbol{\Sigma}_{12}$ 的定义，可知 $\boldsymbol{l}_1$ 也垂直于 $M(\boldsymbol{\Sigma}_{12})$。故有

$$M^{\perp}(\boldsymbol{\Sigma}_{11})\subset M^{\perp}(\boldsymbol{\Sigma}_{12}),$$

此处 $M^{\perp}(\boldsymbol{\Sigma}_{11})$ 表示 $M(\boldsymbol{\Sigma}_{11})$ 的直交补空间。所以有

$$M(\boldsymbol{\Sigma}_{12})\subset M(\boldsymbol{\Sigma}_{11})$$

因之存在有 $\boldsymbol{B}$，使 $\boldsymbol{\Sigma}_{12}=\boldsymbol{\Sigma}_{11}\boldsymbol{B}'$。

由此可得 $\boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-}\boldsymbol{\Sigma}_{11}=\boldsymbol{B}\boldsymbol{\Sigma}_{11}\boldsymbol{\Sigma}_{11}^{-}\boldsymbol{\Sigma}_{11}=\boldsymbol{B}\boldsymbol{\Sigma}_{11}=\boldsymbol{\Sigma}_{21}$.

现计算 $\boldsymbol{y}_{(1)}-\boldsymbol{\mu}_{(1)}$，与 $\boldsymbol{y}_{(2)}-\boldsymbol{\mu}_{(2)}-\boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-}(\boldsymbol{y}_{(1)}-\boldsymbol{\mu}_{(1)})$ 的协方差。

$$\begin{aligned}&\operatorname{cov}(\boldsymbol{y}_{(2)}-\boldsymbol{\mu}_{(2)}-\boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-}(\boldsymbol{y}_{(1)}-\boldsymbol{\mu}_{(1)}),\boldsymbol{y}_{(1)}-\boldsymbol{\mu}_{(1)})\\&=\boldsymbol{\Sigma}_{21}-\boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-}\boldsymbol{\Sigma}_{11}=0.\end{aligned}$$

由于它们都是正态变量，故此二随机向量为独立。又

$$\begin{aligned}&D(\boldsymbol{y}_{(2)}-\boldsymbol{\mu}_{(2)}-\boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-}(\boldsymbol{y}_{(1)}-\boldsymbol{\mu}_{(1)}))\\&=\boldsymbol{\Sigma}_{22}-\boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-}\boldsymbol{\Sigma}_{12}-\boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-}\boldsymbol{\Sigma}_{12}+\boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-}\boldsymbol{\Sigma}_{11}\boldsymbol{\Sigma}_{11}^{-}\boldsymbol{\Sigma}_{12}\\&=\boldsymbol{\Sigma}_{22}-\boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-}\boldsymbol{\Sigma}_{12}\triangleq\boldsymbol{\Sigma}_{22.1}.\end{aligned}$$

因此 $$\boldsymbol{y}_{(2)}-\boldsymbol{\mu}_{(2)}-\boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-}(\boldsymbol{y}_{(1)}-\boldsymbol{\mu}_{(1)})\sim N_{n-r}(\boldsymbol{0},\ \boldsymbol{\Sigma}_{22.1})$$

由于和 $\boldsymbol{y}_{(1)}$ 独立，故在给定 $\boldsymbol{y}_{(1)}$ 的条件下有

$$\boldsymbol{y}_{(2)}\sim N_{n-r}(\boldsymbol{\mu}_{(2)}+\boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-}(\boldsymbol{y}_{(1)}-\boldsymbol{\mu}_{(1)}),\boldsymbol{\Sigma}_{22.1}).$$

特别令 $\boldsymbol{y}_{(2)}=y_n$ 时，y_n 的条件分布是一维正态分布

$$E(y_n|\boldsymbol{y}(1))=\mu_n+\boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-}(\boldsymbol{y}_{(1)}-\boldsymbol{\mu}_{(1)}).$$

实际上它就是 y_n 关于 $\boldsymbol{y}_{(1)}$ 的回归。而 $\boldsymbol{\Sigma}_{22.1}$ 称为剩余方差，当 $\boldsymbol{\Sigma}_{11}$ 非奇时，

$$\boldsymbol{\Sigma}_{22.1}=\frac{|\boldsymbol{\Sigma}|}{|\boldsymbol{\Sigma}_{11}|}=\frac{1}{\sigma^{pp}}$$

此处 σ^{pp} 为 $\boldsymbol{\Sigma}^{-1}$ 的第 p 个对角线元素。

〔定理 6〕 若 $\boldsymbol{y}\sim N_n(\boldsymbol{\mu},\boldsymbol{\Sigma})$，且 $\boldsymbol{\Sigma}>0$.

则 $$Q=(\boldsymbol{y}-\boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\boldsymbol{y}-\boldsymbol{\mu})\sim\chi_n^2.$$

〔证明〕 $\boldsymbol{y}-\boldsymbol{\mu}=\boldsymbol{A}\boldsymbol{x}$

故 $$Q=\boldsymbol{x}'\boldsymbol{A}'\boldsymbol{\Sigma}^{-1}\boldsymbol{A}\boldsymbol{x}=\boldsymbol{x}'\boldsymbol{A}'(\boldsymbol{A}\boldsymbol{A}')^{-1}\boldsymbol{A}\boldsymbol{x}\triangleq\boldsymbol{x}'\boldsymbol{B}\boldsymbol{x},$$

即 $\boldsymbol{B}=\boldsymbol{A}'(\boldsymbol{A}\boldsymbol{A}')^{-1}\boldsymbol{A}$。显然可以验证 $\boldsymbol{B}'=\boldsymbol{B}$、$\boldsymbol{B}^2=\boldsymbol{B}$，所以 $\boldsymbol{B}$ 为一投影阵，故 $rk(\boldsymbol{B})=tr(\boldsymbol{B})$，其特征值为 1 或 0（参看附录一）。

今计算 $rk(\boldsymbol{B})$，$rk(\boldsymbol{B})=tr(\boldsymbol{B})=tr(\boldsymbol{A}'(\boldsymbol{A}\boldsymbol{A}')^{-1}\boldsymbol{A})$

$$=tr((\boldsymbol{A}\boldsymbol{A}')^{-1}\boldsymbol{A}\boldsymbol{A}')=tr(\boldsymbol{I}_n)=n.$$

故知 $\boldsymbol{B}$ 有 n 个 1 及 $p-n$ 个 0 作为它的特征值。因而存在正交阵 $\boldsymbol{T}$ 使得：

$$\boldsymbol{T}'\boldsymbol{B}\boldsymbol{T}=\begin{pmatrix}\boldsymbol{I}_n & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{0}\end{pmatrix}.$$

令 $\boldsymbol{Z}=\boldsymbol{T}'\boldsymbol{x}$，$\boldsymbol{x}=\boldsymbol{T}\boldsymbol{z}$，代入 $\boldsymbol{Q}$ 有

$$\boldsymbol{Q}=\boldsymbol{z}'\boldsymbol{T}'\boldsymbol{B}\boldsymbol{T}\boldsymbol{z}=\boldsymbol{z}'\begin{pmatrix}\boldsymbol{I}_n & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{0}\end{pmatrix}\boldsymbol{z}\sim\chi_n^2.$$

这是由于 $\boldsymbol{x}\sim N(\boldsymbol{0},\boldsymbol{I})$ 的缘故。

〔定理 7〕 设 $\boldsymbol{y}\sim N_n(\boldsymbol{\mu},\sigma^2\boldsymbol{I}_n)$，$\boldsymbol{P}$ 为对称阵，

则 $$\boldsymbol{Q}=\frac{(\boldsymbol{y}-\boldsymbol{\mu})'\boldsymbol{P}(\boldsymbol{y}-\boldsymbol{\mu})}{\sigma^2}\sim\chi_r^2\Longleftrightarrow\boldsymbol{P}^2=\boldsymbol{P},\ rk(\boldsymbol{P})=r.$$

〔证明〕

充分性，与上定理后半部情况类似，由条件可知它有 r 个特征值为 1，$n-r$ 个特征值为 0。故存在正交阵 $\boldsymbol{T}$，使

$$\boldsymbol{T}'\boldsymbol{P}\boldsymbol{T}=\begin{pmatrix}\boldsymbol{I}_r & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{0}\end{pmatrix}.$$