一九八三年　无锡轻工业学院学报　(第四期)
JOURNAL OF THE WXUI INSTITUTE OF LIGHT INDUTSRY　(总第五期)

奇异正态分布的概率元素和广义密度函数

南京工学院　鲁国斌　　无锡轻工业学院　吴有炜

摘　　要

本文讨论奇异正态分布，给出了它在整个空间上的概率元素和分布集中的低维子空间。定义了奇异正态分布的广义密度函数。

设$\vec{a}$是n维实向量，Σ是n维非负定实对称矩阵，则由特征函数

$$\phi(\vec{t})=\exp\{i\vec{a}^{\tau}\vec{t}-\frac{1}{2}\vec{t}^{\tau}\Sigma\vec{t}\}$$

所决定的分布称为n元正态分布，记为$N_n(\vec{a},\Sigma)$。

当$|\Sigma|\neq 0$时，$\vec{Y}\sim N_n(\vec{a},\Sigma)$是连续型随机向量，其密度函数为

$$f(\vec{y})=\frac{1}{(2\pi)^{n/2}|\Sigma|^{1/2}}\exp\{-\frac{1}{2}(\vec{y}-\vec{a})^{\tau}\Sigma^{-1}(\vec{y}-\vec{a})\} \quad (1)$$

当$|\Sigma|=0$时，$N_n(\vec{a},\Sigma)$称为奇异正态分布，此时(1)式无意义。下面讨论其在整个空间上概率元素的分布情况。

概率元素$dF(\vec{y})$为R^n中微元的Lebesgue测度$d\vec{y}$在概率空间$(\Omega,\mathfrak{F},P)$上的概率测度，对应随机向量$\vec{Y}$的分布函数显然为

$$F(\vec{y})=\int_{\substack{u_i<y_i\\ i=1,2,\cdots,n}}dF(\vec{u})$$

若$\vec{Y}$是连续型随机向量，$f(\vec{y})$是其密度函数，则$\vec{Y}$的概率元素

$$dF(\vec{y})=f(\vec{y})d\vec{y},\quad d\vec{y}=dy_1dy_2\cdots dy_n$$

它可以粗略地看成在空间区域

$$\vec{y}-\frac{1}{2}d\vec{y}<\vec{u}<\vec{y}+\frac{1}{2}d\vec{y}$$

上的概率微量。

若$\vec{Y}$是离散型随机向量，$p\{\vec{Y}=\vec{y}_j\}=p_j$，$j=1,2,\cdots$，则

本文收到日期 1983 年 5 月 23 日

$$dF(\vec{y})=\begin{cases}p_j, & \text{当 } \vec{y}=\vec{y}_j \text{ 时}, \ j=1,2,\cdots;\\ 0, & \text{其它。}\end{cases}$$

引入 Penrose(1955 年)定义的广义逆矩阵。

对于任意的 $n\times m$ 阵 A，A 的广义逆矩阵 A^+ 是一个 $m\times n$ 阵，满足条件：

a) $AA^+A=A$； b) $A^+AA^+=A^+$；

c) $(AA^+)^\tau=AA^+$； d) $(A^+A)^\tau=A^+A$。

可以证明这样的 A^+ 对于 A 是唯一存在的。

对于任意实对称阵 Σ，存在正交阵 B，使 $B^\tau\Sigma B=\begin{pmatrix}\lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n\end{pmatrix}$，$\lambda_i(i=1,2,\cdots,n)$ 是 Σ 的特征值，令

$$\lambda_i^*=\begin{cases}\lambda_i^{-1}, & \lambda_i\neq 0;\\ 0, & \lambda_i=0,\end{cases}\qquad i=1,2,\cdots,n。$$

容易验证

$$\Sigma^+=B\begin{pmatrix}\lambda_1^* & & & \\ & \lambda_2^* & & \\ & & \ddots & \\ & & & \lambda_n^*\end{pmatrix}B^\tau$$

下面给出本文主要定理：

定理 1 设 $\vec{Y}\sim N_n(\vec{a},\Sigma)$，$\operatorname{rank}(\Sigma)=r<n$，则

(i) $\vec{Y}$ 的概率元素为

$$dF(\vec{y})=\begin{cases}\dfrac{1}{(2\pi)^{r/2}\sqrt{\lambda_1\cdots\lambda_r}}\exp\left\{-\dfrac{1}{2}(\vec{y}-\vec{a})^\tau\Sigma^+(\vec{y}-\vec{a})\right\}d\vec{y}^*, & \text{当 } \vec{y}=\vec{a}+\Sigma\vec{c} \text{ 时};\\ 0, & \text{其它。}\end{cases}$$

其中，$\lambda_1,\cdots,\lambda_r$ 是 Σ 的非零特征值，$\vec{c}$ 是任意 n 维常向量，$d\vec{y}^*$ 是 r 维超平面 $\{\vec{y}\mid\vec{y}=\vec{a}+\Sigma\vec{c},\ \forall\vec{c}\}$ 上微元在 R^r 中的 L 测度。

(ii) 对 $S\in B_n$，有 $p\{\vec{Y}\in S\}=p\{\vec{Y}^*\in S^*\}$。

这里，$S^*=\{\vec{y}^*\mid\vec{y}^*=B_1^\tau(\vec{y}-\vec{a}),\ \vec{y}\in S\}$；$B=[B_1\,\vdots\,B_2]$ 是正交阵，使得 $B^\tau\Sigma B=$ $\left(\begin{array}{ccc|c}\lambda_1 & & & \\ & \ddots & & 0\\ & & \lambda_r & \\ \hline & 0 & & \begin{matrix}0 & & \\ & \ddots & \\ & & 0\end{matrix}\end{array}\right)$，$\vec{Y}^*\sim N_r\left(\vec{0},\begin{pmatrix}\lambda_1 & & \\ & \ddots & \\ & & \lambda_r\end{pmatrix}\right)$

[证明] 因 $\operatorname{rank}(\Sigma)=r<n$，$N_n(\vec{a},\Sigma)$ 是奇异正态分布。Σ 是非负定实对称阵，必有正交阵 B，使得 $B^\tau\Sigma B=[B_1\,\vdots\,B_2]^\tau\Sigma[B_1\,\vdots\,B_2]=\left(\begin{array}{ccc|c}\lambda_1 & & & \\ & \ddots & & \\ & & \lambda_r & \\ \hline & & & \begin{matrix}0 & & \\ & \ddots & \\ & & 0\end{matrix}\end{array}\right)$，易见 $\Sigma B_2=0$，$B_2^\tau B_1=0$，

$B_1^\tau\Sigma B_1=\begin{pmatrix}\lambda_1 & & \\ & \ddots & \\ & & \lambda_r\end{pmatrix}$，$B_2^\tau\Sigma B_2=0$，作变换 $\vec{U}=B^\tau(\vec{Y}-\vec{a})=\begin{pmatrix}B_1^\tau(\vec{y}-\vec{a})\\ B_2^\tau(\vec{y}-\vec{a})\end{pmatrix}\triangleq\begin{pmatrix}\vec{Y}^*\\ \vec{u}_0\end{pmatrix}$

则$\vec{U}\sim N_n(\vec{0},\begin{pmatrix}\lambda_1 & & & & \\ & \ddots & & & \\ & & \lambda_r & & \\ & & & 0 & \\ & & & & \ddots \\ & & & & & 0\end{pmatrix})$，$\vec{Y}^*\sim N_r(\vec{0},\begin{pmatrix}\lambda_1 & & \\ & \ddots & \\ & & \lambda_r\end{pmatrix})$，$\vec{Y}^*$的密度函数为

$$g_1(\vec{y}^*)=\frac{1}{(2\pi)^{r/2}\sqrt{\lambda_1\cdots\lambda_r}}\exp\left\{-\frac{1}{2}\vec{y}^{*\tau}(B_1^\tau\Sigma B_1)^{-1}\vec{y}^*\right\},$$

其概率元素为$dG_1(\vec{y}^*)=g_1(\vec{y}^*)d\vec{y}^*$，$d\vec{y}^*=dy_1^*\cdots dy_r^*$为$R^r$中L测度下的微分。

$\vec{u}_o\sim N_{n-r}(\vec{0},B_2^\tau\Sigma B_2)=N_{n-r}(\vec{0},O)$，可见$\vec{u}_o$服从正态分布当协方差阵为零阵时的极限情况，即$p\{\vec{u}_o=0\}=1$，所以$\vec{u}_o$的概率元素为

$$dG_2(\vec{u}_o)=\begin{cases}1, & \text{当}\vec{u}_o=\vec{0}\text{时；}\\ 0, & \text{其它}\end{cases}$$

而$\vec{u}_o$与$\vec{Y}^*$是相互独立的，因为$cov(\vec{Y}^*,\vec{u}_o)=cov[B_1^\tau(\vec{Y}-\vec{a}),B_2^\tau(\vec{Y}-\vec{a})]=cov(B_1^\tau\vec{Y},B_2^\tau\vec{Y})=B_1^\tau cov(\vec{Y},\vec{Y})B_2=B_1^\tau\Sigma B_2=0$，因此$\vec{u}$的概率元素$dG(\vec{u})=dG_1(\vec{y}^*)\cdot dG_2(\vec{u}_o)$，即

$$dG(\vec{u})=\begin{cases}\frac{1}{(2\pi)^{r/2}\sqrt{\lambda_1\cdots\lambda_r}}\exp\left\{-\frac{1}{2}\vec{y}^{*\tau}(B_1^\tau\Sigma B_1)^{-1}\vec{y}^*\right\}d\vec{y}^*, & \text{当}\vec{u}_o=\vec{0}\text{时；}\\ 0, & \text{其它}\end{cases}$$

又$\vec{u}_o=0\Leftrightarrow B_2^\tau(\vec{y}-\vec{a})=0\Leftrightarrow B_2^\tau\vec{y}=B_2^\tau\vec{a}$。

非齐次线性方程组$B_2^\tau\vec{y}=B_2^\tau\vec{a}$的通解可表成$\vec{y}=\vec{y}_o+\vec{a}$，其中$\vec{y}_o$是齐次线性方程组$B_2^\tau\vec{y}=0$的通解。

$\because$ $B_2^\tau B_1=0$，$rank(B_2)=n-r$，$rank(B_1)=r$，

$\therefore$ $B_2^\tau\vec{y}=0$的解空间为$M(B_1)$，这儿$M(B_1)$表示B_1的r个列向量$\vec{b}_1,\cdots,\vec{b}_r$为基底的r维线性子空间。

又以$M(\Sigma)$表示Σ诸列向量张成的线性子空间。

$\because$ $\Sigma\vec{b}_i=\lambda_i\vec{b}_i(i=1,2,\cdots,r)$，$\lambda_i>0$，　$\therefore\vec{b}_i=\frac{1}{\lambda_i}\Sigma\vec{b}_i$，即$\vec{b}_i\in M(\Sigma)(i=1,2,\cdots,r)$。又$rank(\Sigma)=r$，　$\therefore dim[M(\Sigma)]=r$，故$M(\Sigma)=M(B_1)$，由此知$B_2^\tau\vec{y}=0$的通解为$\vec{y}_o=\Sigma\vec{c}$（$\vec{c}$为任意n维常向量，因此$\vec{u}_o=0\Leftrightarrow\vec{y}=\Sigma\vec{c}+\vec{a}$　（$\vec{c}$任意）。

$$\because\ (B_1^\tau\Sigma B_1)^{-1}=\begin{pmatrix}\frac{1}{\lambda_1} & & & \\ & \frac{1}{\lambda_2} & & \\ & & \ddots & \\ & & & \frac{1}{\lambda_r}\end{pmatrix}\triangleq\begin{pmatrix}\lambda_1^* & & \\ & \ddots & \\ & & \lambda_r^*\end{pmatrix},\quad\therefore\quad\Sigma^+=B\begin{pmatrix}\lambda_1^* & & & & \\ & \ddots & & & \\ & & \lambda_r^* & & \\ & & & 0 & \\ & & & & \ddots \\ & & & & & 0\end{pmatrix}B^\tau=[B_1\vdots B_2]$$

$$\begin{pmatrix}\lambda_1^* & & & & \\ & \ddots & & & \\ & & \lambda_r^* & & \\ & & & 0 & \\ & & & & \ddots \\ & & & & & 0\end{pmatrix}[B_1\vdots B_2]^\tau=B_1(B_1^\tau\Sigma B_1)^{-1}B_1^\tau。$$

设$\vec{Y}$的分布函数为$F(\vec{y})$，$\vec{U}$的分布函数为$G(\vec{u})$，由$\vec{U}=B^\tau(\vec{Y}-\vec{a})$是可逆变换，故

对任意的 $\vec{x}=(x_1,\cdots,x_n)^\tau$，有

$$G(\vec{x})=G(x_1,\cdots,x_n)=p\{u_1<x_1,\cdots,u_n<x_n\}=p\{\vec{b}_1^\tau(\vec{Y}-\vec{a})<x_1,\cdots,\vec{b}_n(\vec{Y}-\vec{a})<x_n\}$$

$$=\int\limits_{\substack{\vec{b}_i^\tau(\vec{Y}-\vec{a})<x_i\\ i=1,2,\cdots,n}} dF(\vec{y}),\quad 而\ G(\vec{x})=\int\limits_{\substack{u_i<x_i\\ i=1,\cdots,n}} dG(\vec{u}),$$

比较两式知 $dF(\vec{y})=dG(\vec{u})$，于是，$\vec{Y}$ 的概率元素为

$$dF(\vec{y})=\begin{cases}\dfrac{1}{(2\pi)^{r/2}\sqrt{\lambda_1\cdots\lambda_r}}\exp\left\{-\dfrac{1}{2}(\vec{y}-\vec{a})^\tau\Sigma^+(\vec{y}-\vec{a})\right\}d\vec{y}^*, & 当\ \vec{y}=\vec{a}+\Sigma\vec{c}\ 时；\\ 0, & 其它\end{cases}$$

这里 $d\vec{y}^*$ 是 R^n 中 r 维超平面 $\{\vec{y}\mid\vec{y}=\vec{a}+\Sigma\vec{c}\}$ 上微元在 R^r 中的 L 测度，$\vec{c}$ 是任意 n 维常向量，结论(i)获证。

对于确定的 $\vec{a}$、Σ，$\mathrm{rank}(\Sigma)=r$，由结论(i)的证明中知，$\vec{Y}\sim N_n(\vec{a},\Sigma)$ 的分布集中于由 $\vec{a}$，Σ 唯一确定的超平面 $\{\vec{y}\mid\vec{y}=\vec{a}+\Sigma\vec{c}\}=\{\vec{y}\mid B_2^\tau(\vec{y}-\vec{a})=0\}$ 上，$\vec{Y}^*=B_1^\tau(\vec{Y}-\vec{a})$ 与 $\vec{u}_o=B_2^\tau(\vec{Y}-\vec{a})$ 独立，$p\{\vec{u}_o=0\}=1$，∴对任意集合 $S\in B_n$ 有

$$p\{\vec{Y}\mid\vec{Y}\in S\}=p\{\vec{Y}\mid\vec{Y}\in S,B_2^\tau(\vec{Y}-\vec{a})=0\}=p\left\{\begin{bmatrix}\vec{Y}^*\\ \vec{u}_o\end{bmatrix}\middle|\ \vec{Y}\in S,\vec{Y}^*=B_1^\tau(\vec{Y}-\vec{a}),\right.$$

$$\left.\vec{u}_o=B_2^\tau(\vec{y}-\vec{a})=0\right\}=p\{\vec{Y}^*\mid\vec{Y}\in S,\vec{Y}^*=B_1^\tau(\vec{Y}-\vec{a})\}\cdot p\{\vec{u}_o\mid\vec{u}_o=0\}=p\{\vec{Y}^*\mid\vec{Y}\in S^*\}$$

这儿 $S^*=\{\vec{y}^*=B_1^\tau(\vec{y}-\vec{a}),\vec{y}\in S\}$

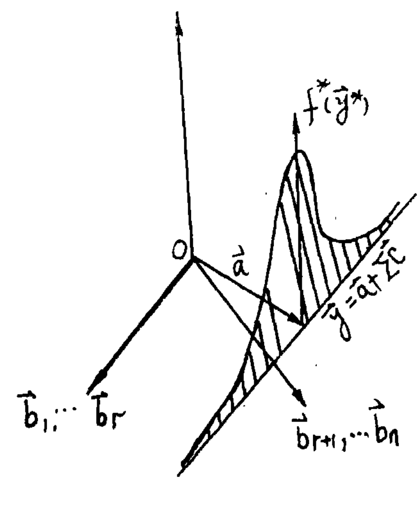

图 1

$\vec{Y}^*\sim N_r\left(0,\begin{pmatrix}\lambda_1 & & \\ & \ddots & \\ & & \lambda_r\end{pmatrix}\right)$ 是 r 维超平面 $\{\vec{y}^*\mid\vec{y}^*=B_1^\tau(\vec{y}-\vec{a})\}$ 上的非奇异正态分布。结论(ii)成立。定理证毕。

由结论(i)，$\vec{Y}$ 的概率元素还可以由左图直观地表示出来，即奇异正态随机向量 $\vec{Y}$ 的概率分布集中在由 $\vec{a}$，Σ 唯一确定的 r 维超平面 $\{\vec{y}\mid\vec{y}=\vec{a}+\Sigma\vec{c},\ \forall\vec{c}\}$ 上，其上相应地有一个非奇异的正态分布 $\vec{Y}^*\sim N_r\left(0,\begin{pmatrix}\lambda_1 & & \\ & \ddots & \\ & & \lambda_r\end{pmatrix}\right)$，当 $\vec{a}\in M(\Sigma)$ 时，超平面就是 $M(\Sigma)$，当 $\vec{a}\overline{\in}M(\Sigma)$ 时，超平面由 $M(\Sigma)$ 平移 $\vec{a}$ 而得。

下面给出利用结论(ii)解决奇异正态的概率计算问题的实例。

〔例〕 设 $\vec{Y}=\begin{bmatrix}x\\ y\end{bmatrix}\sim N_2\left(\begin{bmatrix}2\\ 1\end{bmatrix},\begin{bmatrix}1&1\\ 1&1\end{bmatrix}\right)$，求 $p\{(x,y)\mid x^2+y^2\leqslant1\}$

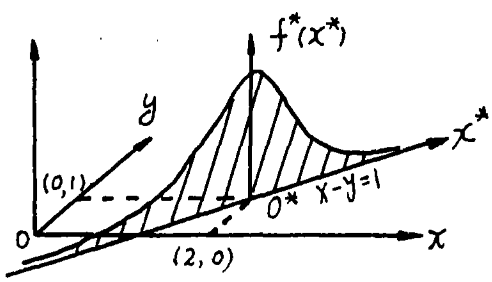

图 2

解：$\lambda_1=2$，$\lambda_2=0$，　$B_1=\begin{bmatrix}\frac{1}{\sqrt{2}}\\ \frac{1}{\sqrt{2}}\end{bmatrix}$，　$B_2=\begin{bmatrix}\frac{1}{\sqrt{2}}\\ -\frac{1}{\sqrt{2}}\end{bmatrix}$，

$\begin{bmatrix}x\\ y\end{bmatrix}=\begin{bmatrix}2\\ 1\end{bmatrix}+\begin{bmatrix}1 & 1\\ 1 & 1\end{bmatrix}\begin{bmatrix}C_1\\ C_2\end{bmatrix}(\forall C_1,C_2)$，即直线 $x-y=1$

$\vec{Y}$的概率元素

$$dF(x,y)=\begin{cases}\dfrac{1}{2\sqrt{\pi}}e^{-\frac{1}{8}(x+y-3)^2}dx^*, & \text{当 } x-y=1\text{时；}\\ 0, & \text{其它}\end{cases}$$

令 $x^*=B_1^{\tau}(\vec{Y}-\vec{a})=\frac{1}{\sqrt{2}}[1,\ 1]\begin{bmatrix}x-2\\ y-1\end{bmatrix}$，有 $x^*\sim N_1(0,2)$，又圆周 $x^2+y^2=1$ 与直线 $x-y=1$ 的交点为 $(1,0)$，$(0,-1)$，相应地 $x_1^*=-\sqrt{2}$，$x_2^*=-2\sqrt{2}$，由定理 1 结论(ii)立即得到 $p\{(x,y)\mid x^2+y^2\leqslant 1\}=p\{x^*\mid -2\sqrt{2}\leqslant x^*\leqslant -\sqrt{2}\}=0.135905$。

最后，讨论奇异正态分布在非古典意义下的密度函数。

在古典的“每点对应一个函数值”的函数概念下，无法表示集中质量相应的密度分布，引入广义函数后，我们就可以逾越这个障碍，现在引入 (Dirac) δ 函数给出奇异正态分布的广义密度函数。

对于在 R^n 内连续且绝对可积的函数 $\varphi(\vec{x})$，定义泛函 $\delta_n(\vec{x}-\vec{c})$，使 $\int_{R^n}\varphi(\vec{x})\delta_n(\vec{x}-\vec{c})d\vec{x}=\varphi(\vec{c})$。

δ 函数可以视为普通函数弱收敛意义下的极限。可以证明

$$\lim_{\substack{\lambda_i\to 0\\ i=r+1,\cdots,n\\ \lambda_i>0}}\frac{1}{(2\pi)^{(n-r)/2}\sqrt{\lambda_{r+1}\cdots\lambda_n}}\exp\left\{-\frac{1}{2}(\vec{\xi}-\vec{c})^{\tau}\begin{bmatrix}\lambda_{r+1} & & \\ & \ddots & \\ & & \lambda_n\end{bmatrix}^{-1}(\vec{\xi}-\vec{c})\right\}\overset{(弱)}{=\!=\!=}\delta_{n-r}(\vec{\xi}-\vec{c}) \tag{2}$$

即对于任意在 R^{n-r} 内连续且绝对可积的函数

$$\varphi(\vec{\xi}),\ \varphi(\vec{c})=\lim_{\substack{\lambda_i\to 0\\ i=r+1,\cdots n\\ \lambda_i>0}}\int_{R^{n-r}}\frac{1}{(2\pi)^{n-r/2}\sqrt{\lambda_{r+1}\cdots\lambda_n}}\exp\left\{-\frac{1}{2}(\vec{\xi}-\vec{c})^{\tau}\begin{pmatrix}\lambda_{r+1}&&\\&\ddots&\\&&\lambda_n\end{pmatrix}^{-1}(\vec{\xi}-\vec{c})\right\}\varphi(\vec{\xi})\,d\vec{\xi}。$$

对n维离散型随机向量$\vec{Y}$，若$p\{\vec{Y}=\vec{Y}_j\}=p_j$，$j=1,2,\cdots$，定义$\vec{Y}$的广义密度函数为$f(\vec{y})=\sum_j p_j\delta_n(\vec{y}-\vec{y}_j)$。如此定义的密度函数，同样具有性质：

$$1^\circ\ f(\vec{y})\geqslant 0;\qquad 2^\circ\int_{R^n}f(\vec{y})\,d\vec{y}=1$$

粗略地，可以认为$\delta_n(\vec{y}-\vec{a})=\begin{cases}\infty^n, & 当\ \vec{y}=\vec{a}\ 时;\\ 0, & 其它,\end{cases}$

把定理1的证明稍加相应更动，即得

定理2　随时向量$\vec{Y}\sim N_n(\vec{a},\Sigma)$，$\text{rank}(\Sigma)=r<n$，

则$\vec{Y}$的广义密度函数为

$$f(\vec{y})=\frac{1}{(2\pi)^{r/2}\sqrt{\lambda_1\cdots\lambda_r}}\exp\left\{-\frac{1}{2}(\vec{y}-a)^{\tau}\Sigma^{+}(\vec{y}-\vec{a})\right\}\delta_{n-r}(B_2^{\tau}\vec{y}-B_2^{\tau}\vec{a})$$

这儿$\lambda_1,\cdots,\lambda_r$是$\Sigma$的非零特征值，短阵$B_2$的诸列向量是$\Sigma$的零特征值所对应的（n－r）个彼此正交的特征么矢。

定理2给出了奇异正态分布的广义密度函数，它的几何意义由（2）式清晰可见，表示非奇异正态分布的密度函数(在常义下)在超平面$B_2^{\tau}\vec{y}=B_2^{\tau}\vec{a}$的垂直方向上方差趋于零时极限状况。

The Element of Probabilty of the Singular Multinormal Distribution and its Generalized Density Function

Lu Guo-bing　Wu Yu-wei

Abstract

This paper deals with the singular multinormal distribution. Its element of probability in whole space is given and the lower-dimension sub-space is uniquely determined in which the total probability is located. The generalized density function of the singular normal distribution is also deducted in this paper.