高斯分布随机变量及其性质

- 中心极限定理
- 高斯分布的随机变量
- N 维高斯随机变量的统计独立特性
- 高斯随机变量的线性变换
- 高斯分布的随机变量的条件分布和边缘分布

1．引言．中心极限定理

给定 n 个独立的随机变量 x_i， $i=1,2,\cdots$n，它们的和为：$x=x_1+x_2+\cdots+x_n$，x 的均值为 $\eta=\eta_1+\eta_2+\cdots+\eta_n$，方差为 $\sigma^2={\sigma_1}^2+{\sigma_2}^2+\cdots+{\sigma_n}^2$，在一定的条件下，当 n 趋于无穷时，x 的概率密度函数 f(x)趋向于具有相同均值和方差的高斯(正态)分布：

$$f(x)\approx\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{(x-\eta)^2}{2\sigma^2}}$$

中心极限定理逼近的性质以及对一个给定误差所需的随机变量的数目 n 依赖于概率密度函数 $f_i(x)$。

2 高斯分布的随机变量

典型高斯分布的随机变量的概率密度与特征函数的描述。

$f_\xi(x)$，

$$\Phi_\xi(u)=E\left\{e^{ju\xi}\right\}=\int_{-\infty}^{\infty}e^{jux}f_\xi(x)dx$$

2.1 一元高斯随机变量

一元高斯随机变量 N(0,1)，均值为零、方差为 1，其概率密度和特征函数：

$$f_\xi(x)=\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}$$

$$\Phi_\xi(u)=e^{-\frac{u^2}{2}}$$

一元高斯随机变量 N(μ, σ²)，均值为μ、方差为σ²，其概率密度和特征函数：

$$f_{\xi}(x)=\frac{1}{\sqrt{2\pi\sigma^{2}}}e^{-\frac{(x-\mu)^{2}}{2\sigma^{2}}}$$

$$\Phi_{\xi}(u)=e^{j\mu u-\frac{\sigma^{2}u^{2}}{2}}$$

2.2 二元高斯随机变量

二元高斯随机变量ξ_1,ξ_2，均值为零、协方差矩阵为：

$$B=\begin{pmatrix}1 & r\\ r & 1\end{pmatrix}，\quad B^{-1}=\frac{1}{1-r^{2}}\begin{pmatrix}1 & -r\\ -r & 1\end{pmatrix}，\quad |B|=1-r^{2}$$

其二元概率密度和特征函数为：

$$f_{\xi_1\xi_2}(x_1,x_2)=\frac{1}{2\pi\sqrt{1-r^{2}}}\exp\left(-\frac{1}{2(1-r^{2})}[x_1^{2}-2rx_1x_2+x_2^{2}]\right)$$

$$\Phi_{\xi_1\xi_2}(u_1,u_2)=\exp\left(-\frac{1}{2}[u_1^{2}+2ru_1u_2+u_2^{2}]\right)$$

二元高斯随机变量ξ_1,ξ_2，其均值、协方差矩阵为，

$$E\begin{pmatrix}\xi_1\\ \xi_2\end{pmatrix}=\begin{pmatrix}\mu_1\\ \mu_2\end{pmatrix}=\mu,$$

$$B=\begin{pmatrix}\sigma_1^{2} & r\sigma_1\sigma_2\\ r\sigma_1\sigma_2 & \sigma_2^{2}\end{pmatrix},\qquad |B|=\sigma_1^{2}\sigma_2^{2}\left(1-r^{2}\right)$$

$$\begin{aligned}B^{-1}&=\frac{1}{\sigma_1^{2}\sigma_2^{2}\left(1-r^{2}\right)}\begin{pmatrix}\sigma_2^{2} & -r\sigma_1\sigma_2\\ -r\sigma_1\sigma_2 & \sigma_1^{2}\end{pmatrix}\\ &=\frac{1}{\left(1-r^{2}\right)}\begin{pmatrix}1/\sigma_1^{2} & -r/\sigma_1\sigma_2\\ -r/\sigma_1\sigma_2 & 1/\sigma_2^{2}\end{pmatrix}\end{aligned}$$

其二元概率密度和特征函数为

$$f_{\xi_1\xi_2}(x,y)=\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-r^{2}}}\times$$
$$\exp\left(-\frac{1}{2(1-r^{2})}\left[\left(\frac{x_1-\mu_1}{\sigma_1}\right)^{2}-2r\left(\frac{x_1-\mu_1}{\sigma_1}\right)\left(\frac{x_2-\mu_2}{\sigma_2}\right)+\left(\frac{x_2-\mu_2}{\sigma_2}\right)^{2}\right]\right)$$

$$\Phi_{\xi\eta}(u_1,u_2)=\exp\left(j\left(\mu_1u_1+\mu_2u_2\right)-\frac{1}{2}[\sigma_1^2u_1^2+2r\sigma_1\sigma_2u_1u_2+\sigma_2^2u_2^2]\right)$$

2.3 n 元高斯随机变量

n 元高斯随机变量ξ，其均值、协方差矩阵（正定的）为，

$$\mu=\begin{pmatrix}\mu_1\\ \mu_2\\ \vdots\\ \mu_n\end{pmatrix},\qquad B=\begin{pmatrix}b_{11} & b_{12} & \cdots & b_{1n}\\ b_{12} & b_{22} & \cdots & b_{2n}\\ \vdots & \vdots & & \vdots\\ b_{n1} & b_{n2} & \cdots & b_{nn}\end{pmatrix}$$

其 n 元概率密度和特征函数为

$$f_\xi(\mathbf{x})=\frac{1}{\left[(2\pi)^n|B|\right]^{1/2}}\exp\left(-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})^TB^{-1}(\mathbf{x}-\boldsymbol{\mu})\right)$$

$$\Phi_\xi(U)=\exp\left(j\boldsymbol{\mu}^{\mathrm{T}}\mathbf{u}-\frac{1}{2}\mathbf{u}^{\mathrm{T}}B\mathbf{u}\right)$$

其中，

$\mathbf{x}=\left(x_1\ x_2\ \cdots x_n\right)^T$， $\mathbf{u}=\left(u_1\ u_2\ \cdots u_n\right)^T$

考虑到矩阵是 B 正定对称的，则存在一个非奇异矩阵 L，使得 $B=LL^T$，作线性变换 L，

$\mathbf{y}=L^{-1}(\mathbf{x}-\boldsymbol{\mu}_X)$， $\mathbf{x}=L\mathbf{y}+\boldsymbol{\mu}_X$，

$$B^{-1}=(LL^t)^{-1}=(L^t)^{-1}\cdot L^{-1}=(L^{-1})^tL^{-1}$$

$$\begin{aligned}(\mathbf{x}-\boldsymbol{\mu}_X)^TB^{-1}(\mathbf{x}-\boldsymbol{\mu}_X)&=(\mathbf{x}-\boldsymbol{\mu}_X)^T(L^t)^{-1}\cdot L^{-1}(\mathbf{x}-\boldsymbol{\mu}_X)\\&=[L^{-1}(\mathbf{x}-\boldsymbol{\mu}_X)]^T[L^{-1}(\mathbf{x}-\boldsymbol{\mu}_X)]\\&=\mathbf{y}^T\mathbf{y}\end{aligned}$$

对应这个变换的雅可比行列式是$\dfrac{\partial\mathbf{x}}{\partial\mathbf{y}}=|L|=|B|^{1/2}$

$$\begin{aligned}f_\eta(Y)&=\frac{1}{\left[(2\pi)^n|B|\right]^{1/2}}\exp\left(-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu}_X)^TB^{-1}(\mathbf{x}-\boldsymbol{\mu}_X)\right)\cdot|B|^{1/2}\\&=\frac{1}{\left[(2\pi)^n\right]^{1/2}}\exp\left(-\frac{1}{2}\mathbf{y}^T\mathbf{y}\right)\end{aligned}$$

$$= \frac{1}{\left[(2\pi)^n\right]^{1/2}} \exp\left(-\frac{1}{2} \sum_{n=1}^{N} y_n^2 \right)$$

显然有

$$\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f_\xi(X) dX = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f_\eta(Y) dY$$

$$= \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f_\eta(Y) dy_1 dy_2 \cdots dy_N = 1$$

2.4 n 元高斯随机变量的特征函数的计算

考虑以下的矩阵运算

$$j\mathbf{u}^T\mathbf{x} = j\mathbf{u}^T(L\mathbf{y} + \boldsymbol{\mu}_X) = j\mathbf{u}^T L\mathbf{y} + j\mathbf{u}^T\boldsymbol{\mu}_X$$
$$= jS^T\mathbf{y} + j\mathbf{u}^T\boldsymbol{\mu}_X$$

其中：$S^T = \mathbf{u}^T L, \qquad S = L^T\mathbf{u}$

$$j\mathbf{u}^T\mathbf{x} - \frac{1}{2}(\mathbf{x} - \boldsymbol{\mu}_X)^T B^{-1}(\mathbf{x} - \boldsymbol{\mu}_X)$$
$$= j\mathbf{u}^T\mathbf{x} - \mathbf{y}^{\mathbf{T}}\mathbf{y}/2$$
$$= j\mathbf{u}^T\boldsymbol{\mu}_X + jS^T\mathbf{y} - \mathbf{y}^T\mathbf{y}/2$$
$$= j\mathbf{u}^T\boldsymbol{\mu}_X - (\mathbf{y} - jS)^T(\mathbf{y} - jS)/2 - S^T S/2$$

N 元高斯随机变量的特征函数是:

$$\Phi_\xi(\mathbf{u}) = E\left\{\exp(j\mathbf{u}^T\mathbf{x})\right\} = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \exp(j\mathbf{u}^T\mathbf{x}) f_\xi(\mathbf{x}) d\mathbf{x}$$

$$= \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \exp(j\mathbf{u}^T\mathbf{x}) \frac{1}{\left[(2\pi)^n |B|\right]^{1/2}} \exp\left(-\frac{1}{2}(\mathbf{x} - \boldsymbol{\mu})^T B^{-1}(\mathbf{x} - \boldsymbol{\mu}) \right) d\mathbf{x}$$

$$= \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \frac{1}{\left[(2\pi)^n |B|\right]^{1/2}} \exp\left(j\mathbf{u}^T\mathbf{x} - \frac{1}{2}(\mathbf{x} - \boldsymbol{\mu})^T B^{-1}(\mathbf{x} - \boldsymbol{\mu}) \right) d\mathbf{x}$$

$$= \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \frac{1}{\left[(2\pi)^n\right]^{1/2}} \exp[j\mathbf{u}^T\boldsymbol{\mu}_X - S^T S/2]$$

$$\exp[-(\mathbf{y} - jS)^T(\mathbf{y} - jS)/2] d\mathbf{y}$$

$$= \exp[j\mathbf{u}^T\boldsymbol{\mu}_X - S^T S/2]$$
$$= \exp[j\mathbf{u}^T\boldsymbol{\mu}_X - \mathbf{u}^T L L^T\mathbf{u}/2]$$
$$= \exp[j\mathbf{u}^T\boldsymbol{\mu}_X - \mathbf{u}^T B\mathbf{u}/2]$$

$$\Phi_{\xi}(\mathbf{u}) = \exp\left(j\mathbf{u}^T \boldsymbol{\mu}_X - \frac{1}{2}\mathbf{u}^T B\mathbf{u} \right)$$

当协方差矩阵是非负定的，可以证明若它的秩为 r<n，它的概率分布集中在 r 维子空间上，这种分布是退化正态分布，或奇异正态分布。

3 N 维高斯随机变量的统计独立特性

3.1 定理 1

N 维随机变量 ξ_1，ξ_2， ，ξ_N，相互统计独立的充要条件是它们两两互不相关。

证明，

首先证明必要性。

- 若 N 维随机变量 ξ_1，ξ_2， ，ξ_N，相互统计独立，则它们 N 个高斯分布随机变量的概率密度函数，等于它们各自概率密度函数的乘积。
- N 维随机变量 ξ_1，ξ_2， ，ξ_N，它们的特征函数等于各自特征函数的乘积。对比高斯分布特征函数的表达式，它们的协方差矩阵是对角矩阵。
- N 维随机变量 ξ_1，ξ_2， ，ξ_N，它们的协方差矩阵是对角矩阵，它们的互相关为零，它们是统计独立的。

其次证明充分性。

- 若 N 维随机变量 ξ_1，ξ_2， ，ξ_N，是两两不相关，它们的协方差矩阵是对角矩阵。
- 若 N 维随机变量 ξ_1，ξ_2， ，ξ_N，协方差矩阵是对角矩阵，它们的特征函数等于各自特征函数的乘积，它们是相互统计独立。

3.2 定理 2

若 ξ 是高斯分布的随机矢量，ξ_1，ξ_2 是两个子矢量，$\xi =（\xi_1 \ \ \xi_2）^T$ 它们的协方差矩阵是 $\mathbf{B} = \begin{pmatrix} \mathbf{B}_{11} & \mathbf{B}_{12} \\ \mathbf{B}_{21} & \mathbf{B}_{22} \end{pmatrix}$，其中 $\mathbf{B}_{11}$ 和 $\mathbf{B}_{22}$ 分别是 ξ_1，ξ_2 的协方差矩阵，$\mathbf{B}_{12}$ 和 $\mathbf{B}_{21}$ 分别是 ξ_1，ξ_2 的互协方差矩阵。$\mathbf{B}_{12}=(\mathbf{B}_{21})^H$，$\xi_1$，$\xi_2$ 相互统计独立的充要条件是 $\mathbf{B}_{12}=0$

证明：

首先证明必要性。

若ξ$_1$，ξ$_2$相互统计独立，它们之间的任意两个分量都统计独立，它们之间的任意两个分量的协方差都是零，相应的协方差矩阵 $B_{12}=0$，$B_{21}=0$。

其次证明充分性。

若 $\mathbf{B}_{12}=0$，$\mathbf{B}_{21}=0$，相应ξ的相关矩阵 $\mathbf{B}=\begin{pmatrix}\mathbf{B}_{11} & \mathbf{0}\\ \mathbf{0} & \mathbf{B}_{22}\end{pmatrix}$。

令 $\mathbf{u}=(\mathbf{u}_1\ \mathbf{u}_2)^T$，与相应分量的维数与ξ＝（ξ$_1$　ξ$_2$）^{T}一致，它们的特征函数，

$$\begin{aligned}\Phi_{\xi}(\mathbf{u})&=\exp\left(j\mathbf{u}^{\mathrm{T}}\boldsymbol{\mu}-\mathbf{u}^{\mathrm{T}}\mathbf{B}\mathbf{u}/2\right)\\&=\exp\left(j\mathbf{u}_1^{\mathrm{T}}\boldsymbol{\mu}_1+j\mathbf{u}_2^{\mathrm{T}}\boldsymbol{\mu}_2-[\mathbf{u}_1^{\mathrm{T}}\mathbf{B}_{11}\mathbf{u}_1+\mathbf{u}_2^{\mathrm{T}}\mathbf{B}_{22}\mathbf{u}_2]/2\right)\\&=\exp\left(j\mathbf{u}_1^{\mathrm{T}}\boldsymbol{\mu}_1-[\mathbf{u}_1^{\mathrm{T}}\mathbf{B}_{11}\mathbf{u}_1]/2\right)\cdot\exp\left(j\mathbf{u}_2^{\mathrm{T}}\boldsymbol{\mu}_2-[\mathbf{u}_2^{\mathrm{T}}\mathbf{B}_{22}\mathbf{u}_2]/2\right)\\&=\Phi_{\xi_1}(\mathbf{u}_1)\cdot\Phi_{\xi_2}(\mathbf{u}_2)\end{aligned}$$

等于两个子矢量的特征函数的乘积，因此这两个子矢量是相互独立的。

4 高斯随机变量的线性变换

4.1 高斯随机变量的线性组合

设ξ＝（ξ$_1$，ξ$_2$，　，ξ$_N$）是 N 维随机矢量，其数学期望是μ＝（μ$_1$，μ$_2$，　，μ$_N$)，协方差矩阵是 B。

高斯随机变量ξ各个分量线性组合

$$\eta=\sum_{n=1}^{N}a_n\xi_n=\mathbf{a}^T\boldsymbol{\xi}，\quad \mathbf{a}^{\mathrm{T}}=（a_1，a_2，\quad，a_N）$$

高斯随机变量线性组合的均值，

$$\begin{aligned}E\{\eta\}&=E\left\{\sum_{n=1}^{N}a_n\xi_n\right\}=\sum_{n=1}^{N}a_nE[\xi_n]\\&=\sum_{n=1}^{N}a_n\mu_n=\mathbf{a}^T\boldsymbol{\mu}\end{aligned}$$

高斯随机矢量线性组合的协方差，

$$
\begin{aligned}
B_\eta &= E\left\{\sum_{n=1}^{N} a_n(\xi_n-\mu_n)\cdot\sum_{m=1}^{N} a_m(\xi_m-\mu_m)\right\}\\
&=\sum_{n=1}^{N}\sum_{m=1}^{N} a_n a_m E\{(\xi_m-\mu_m)\cdot(\xi_n-\mu_n)\}\\
&=\sum_{n=1}^{N}\sum_{m=1}^{N} a_n a_m b_{nm}\\
&=\mathbf{a}^T\mathbf{B}\mathbf{a}
\end{aligned}
$$

4.2 高斯随机变量的线性变换

设ξ＝（ξ$_1$，ξ$_2$，　，ξ$_N$）是 N 维随机矢量，其数学期望是μ＝（μ$_1$，μ$_2$，　，μ$_N$)，协方差矩阵是 B。

线性变换 **C**，是 M*N 的矩阵，ξ经过线性变换 **C** 得到η＝**C**ξ，

均值：

$$
E\boldsymbol{\eta}=E\{\mathbf{C}\boldsymbol{\xi}\}=\mathbf{C}E\{\boldsymbol{\xi}\}=\mathbf{C}\boldsymbol{\mu}_\xi,
$$

协方差矩阵：

$$
\begin{aligned}
D\{(\boldsymbol{\eta}-E\boldsymbol{\eta}\}&=E\left\{(\boldsymbol{\eta}-E\boldsymbol{\eta})\overline{(\boldsymbol{\eta}-E\boldsymbol{\eta})}^T\right\}\\
&=E\left\{(\mathbf{C}\boldsymbol{\xi}-E\mathbf{C}\boldsymbol{\xi})\overline{(\mathbf{C}\boldsymbol{\xi}-E\mathbf{C}\boldsymbol{\xi})}^T\right\}\\
&=E\left\{\mathbf{C}(\boldsymbol{\xi}-\boldsymbol{\mu}_\xi)\overline{(\boldsymbol{\xi}-\boldsymbol{\mu}_\xi)}^T\mathbf{C}^T\right\}\\
&=\mathbf{C}E\left\{(\boldsymbol{\xi}-\boldsymbol{\mu}_\xi)\overline{(\boldsymbol{\xi}-\boldsymbol{\mu}_\xi)}^T\right\}\mathbf{C}^T\\
&=\mathbf{C}\mathbf{B}\mathbf{C}^T
\end{aligned}
$$

4.3 定理 1

设ξ＝（ξ$_1$，ξ$_2$，　，ξ$_N$）是 N 维随机矢量，其数学期望是，μ＝（μ$_1$，μ$_2$，　，μ$_N$)，协方差矩阵是 **B**。ξ服从 N 元高斯分布的充要条件是它的任意一个线性组合

$\varsigma=\sum_{n=1}^{N} a_n\xi_n=\mathbf{a}^T\boldsymbol{\xi}$ 服从一元高斯分布。

证明：

首先证明必要性。

如果ξ的任意一个线性组合 $\varsigma=\sum_{n=1}^{N} a_n\xi_n=\mathbf{a}^T\boldsymbol{\xi}$ 服从一元高斯分布。

考虑到 ζ 的均值是 $\boldsymbol{\mu}_{\varsigma} = \mathbf{a}^T \boldsymbol{\mu}_{\xi}$， ζ 的方差是 $\mathbf{B}_{\varsigma} = \mathbf{a}^T \mathbf{B}_{\xi} \mathbf{a}$， 则 ξ 的特征函数是，

$$
\begin{aligned}
& E\left\{\exp\left(j\mathbf{u}^T \cdot \boldsymbol{\xi}\right)\right\} \\
& = E\left\{\exp\left(j\sum_{n=1}^{N} u_n \xi_n\right)\right\} \\
& = E\left\{\exp\left(ju_0\sum_{n=1}^{N} u'_n \xi_n\right)\right\} = E\left\{\exp\left(ju_0 \mathbf{u}'^T \boldsymbol{\xi}\right)\right\} \\
& = \exp\left(ju_0 \mathbf{u}'^T \cdot \boldsymbol{\mu}_{\xi} - \mathbf{u}'^T \mathbf{B}_{\xi} \mathbf{u}' u_0^2 / 2\right) \\
& = \exp\left(j\mathbf{u}^T \cdot \boldsymbol{\mu}_{\xi} - \mathbf{u}^T \mathbf{B}_{\xi} \mathbf{u} / 2\right)
\end{aligned}
$$

上述推导的第一步，是按照矢量点积的表达式写出的，

上述推导的第二步，是将 $\mathbf{u} = u_0 \mathbf{u}'$ 代入表达式的，

上述推导的第三步，是鉴于 ζ 的任意一个线性组合服从一元高斯分布，因而写出相应的特征函数表达式，

上述推导的第四步，是再次将 $\mathbf{u} = u_0 \mathbf{u}'$ 代入表达式的，

推导的结果说明 ξ 的特征函数具有高斯矢量的形式，必要性得到证明。

其次证明充分性。

如果 ξ 是 N 维高斯随机矢量，它的任意一个线性组合 $\varsigma = \sum_{n=1}^{N} a_n \xi_n = \mathbf{a}^T \boldsymbol{\xi}$ 的特征函数是，

$$
\begin{aligned}
E\left\{\exp\left(ju\varsigma\right)\right\} & = E\left\{\exp\left(ju\sum_{n=1}^{N} a_n \xi_n\right)\right\} \\
& = E\left\{\exp\left(j\sum_{n=1}^{N} (ua_n) \xi_n\right)\right\} \\
& = E\left\{\exp\left(ju\mathbf{a}^T \cdot \boldsymbol{\xi}\right)\right\} \\
& = \exp\left(ju\mathbf{a}^T \mu_{\xi} - u\mathbf{a}^T \mathbf{B}_{\xi} \mathbf{a} u / 2\right) \\
& = \exp\left(ju(\mathbf{a}^T \boldsymbol{\mu}_{\xi}) - u^2 (\mathbf{a}^T \mathbf{B}_{\xi} \mathbf{a}) / 2\right) \\
& = \exp\left(juE(\varsigma) - u^2 D(\varsigma) / 2\right)
\end{aligned}
$$

上述推导的第四步，是按照高斯随机矢量的特征函数表达式写出的，

推导的结果说明 ζ 的特征函数具有高斯变量的形式，充分性性得到证明。

4.4 定理 2

设ξ＝（ξ$_1$，ξ$_2$， ，ξ$_N$）是 N 维随机矢量，其数学期望是，μ＝（μ$_1$，μ$_2$， ，μ$_N$），协方差矩阵是 **B**，服从高斯分布 N(μ,**B**)。**C** 是 M×N 矩阵，ξ经过线性变换 **C** 得到η＝**C**ξ，η为 M×1 列矢量，它服从 M 元高斯分布 N(**C**μ,**CBC^T**)。

证明：

因为ξ＝（ξ$_1$，ξ$_2$， ，ξ$_N$）是 N×1 维随机列矢量，**C** 是 M×N 矩阵，所以η＝**C**ξ是 M×1 列矢量。

其次，考虑 M×1 列矢量 **t**，相应η的特征函数是：

$$
\begin{aligned}
E\left\{\exp\left(j\mathbf{t}^T\boldsymbol{\eta}\right)\right\} &= E\left\{\exp\left(j\mathbf{t}^T\mathbf{C}\boldsymbol{\xi}\right)\right\} \\
&= E\left\{\exp\left(j(\mathbf{C}^T\mathbf{t})^T\boldsymbol{\xi}\right)\right\} \\
&= \exp\left\{j(\mathbf{C}^T\mathbf{t})^T\boldsymbol{\mu}-(\mathbf{C}^T\mathbf{t})^T\mathbf{B}(\mathbf{C}^T\mathbf{t})/2\right\} \\
&= \exp\left\{j\mathbf{t}^T\mathbf{C}\boldsymbol{\mu}-\mathbf{t}^T\mathbf{C}\mathbf{B}\mathbf{C}^T\mathbf{t}/2\right\} \\
&= \exp\left\{j\mathbf{t}^T(\mathbf{C}\boldsymbol{\mu})-\mathbf{t}^T(\mathbf{C}\mathbf{B}\mathbf{C}^T)\mathbf{t}/2\right\}
\end{aligned}
$$

由高斯分布随机矢量的线性变换的性质知，η的均值是 **C**μ，协方差是 **CBC**^{T}，而η的特征函数是 $\exp\left\{j\mathbf{t}^T(\mathbf{C}\boldsymbol{\mu})-\mathbf{t}^T(\mathbf{C}\mathbf{B}\mathbf{C}^T)\mathbf{t}/2\right\}$，故ξ经过线性变换 **C** 得到η＝**C**ξ是高斯随机变量。

4.5 定理 2 推论

设ξ＝（ξ$_1$，ξ$_2$， ，ξ$_N$）是 n 维高斯随机矢量，服从高斯分布 N(μ,**B**)。设存在一个正交变换 **U**，使得η＝**U**ξ是一个具有独立高斯分布分量的高斯分布的随机矢量，它的数学期望是 **U**μ，方差分量是协方差矩阵 B 的特征值。

证明：

对于一个实对称的协方差矩阵 **B**，存在特征值 d_i，和特征矢量 u_i，i=1,2, ,N，有

$$\mathbf{B}\mathbf{u}_i = d_i\mathbf{u}_i,$$

$$\mathbf{U}=\left(\mathbf{u}_1{}^T\ \mathbf{u}_2{}^T\quad\quad \mathbf{u}_N{}^T\right)^T \quad \mathbf{U}^T=\left(\mathbf{u}_1\ \mathbf{u}_2\quad\quad \mathbf{u}_N\right),$$

$$\mathbf{U}\mathbf{U}^T=\mathbf{I},$$

$$\mathbf{B}\mathbf{U}^T = \mathbf{B}\left(\mathbf{u}_1\ \mathbf{u}_2 \quad \cdots \quad \mathbf{u}_N\right) = \left(d_1\mathbf{u}_1\ d_2\mathbf{u}_2 \quad \cdots \quad d_N\mathbf{u}_N\right)$$

$$= \left(\mathbf{u}_1\ \mathbf{u}_2 \quad \cdots \quad \mathbf{u}_N\right)\begin{pmatrix} d_1 & & & \\ & d_2 & & \\ & & \ddots & \\ & & & d_N \end{pmatrix}$$

$$= \mathbf{U}^T\mathbf{D}$$

$$\mathbf{U}\mathbf{B}\mathbf{U}^T = \mathbf{D}$$

$$\mathbf{B} = \mathbf{U}^T\mathbf{D}\mathbf{U}$$

$\mathbf{U}\mathbf{B}\mathbf{U}^T = \mathbf{D}$，U 作为高斯随机矢量的线性变换矩阵，可以得到 η＝Uξ 的数学期望是 Uμ，协方差矩阵是 D。即 $E\left[U\xi\right] = UE\left[\xi\right] = U\mu_\xi$

5 高斯分布的随机变量的条件分布和边缘分布

5.1 均方误差最小的条件估值

设 ξ 和 η 是两个随机矢量，两者存在联合分布，设 η 是观察矢量，通过 η 对 ξ 进行估值，求均方误差最小的估值 $\hat{\boldsymbol{\xi}}$。

$$E\left\{\left\|\boldsymbol{\xi} - \hat{\boldsymbol{\xi}}(\boldsymbol{\eta})\right\|^2 / \boldsymbol{\eta} = \mathbf{y}\right\} = \min E\left\{\left\|\boldsymbol{\xi} - \mathbf{k}\right\|^2 / \boldsymbol{\eta} = \mathbf{y}\right\}$$

解：

计算估值的均方误差

$$\begin{aligned}
&E\left\{\left\|\boldsymbol{\xi} - \mathbf{k}\right\|^2 / \boldsymbol{\eta} = \mathbf{y}\right\} \\
&= E\left\{\left\|\boldsymbol{\xi} - E\left\{\boldsymbol{\xi}/\boldsymbol{\eta}\right\} + E\left\{\boldsymbol{\xi}/\boldsymbol{\eta}\right\} - \mathbf{k}\right\|^2 / \boldsymbol{\eta} = \mathbf{y}\right\} \\
&= E\left\{\left\|E\left\{\boldsymbol{\xi}/\boldsymbol{\eta}\right\} - \mathbf{k}\right\|^2 / \boldsymbol{\eta} = \mathbf{y}\right\} \\
&\quad + E\left\{\left\|[\boldsymbol{\xi} - E\left\{\boldsymbol{\xi}/\boldsymbol{\eta}\right\}] \cdot [E\left\{\boldsymbol{\xi}/\boldsymbol{\eta}\right\} - \mathbf{k}]\right\| / \boldsymbol{\eta} = \mathbf{y}\right\} \\
&\quad + E\left\{\left\|[E\left\{\boldsymbol{\xi}/\boldsymbol{\eta}\right\} - \mathbf{k}] \cdot [\boldsymbol{\xi} - E\left\{\boldsymbol{\xi}/\boldsymbol{\eta}\right\}]\right\| / \boldsymbol{\eta} = \mathbf{y}\right\} \\
&\quad + E\left\{\left\|\boldsymbol{\xi} - E\left\{\boldsymbol{\xi}/\boldsymbol{\eta}\right\}\right\|^2 / \boldsymbol{\eta} = \mathbf{y}\right\} \\
&= E\left\{\left\|E\left\{\boldsymbol{\xi}/\boldsymbol{\eta}\right\} - \mathbf{k}\right\|^2 / \boldsymbol{\eta} = \mathbf{y}\right\} \\
&\quad + E\left\{\left\|\boldsymbol{\xi} - E\left\{\boldsymbol{\xi}/\boldsymbol{\eta}\right\}\right\|^2 / \boldsymbol{\eta} = \mathbf{y}\right\}
\end{aligned}$$

为了使均方误差最小，应使估值 $\mathbf{k} = E\left\{\boldsymbol{\xi}/\boldsymbol{\eta}\right\}$

即均方误差最小的条件估值是条件均值：$\hat{\xi}(\eta = y) = E\{\xi/\eta = y\}$

5.2 二元高斯分布随机变量的条件分布和边缘分布

二元高斯随机变量协方差矩阵三角化分解和逆矩阵

二元高斯随机变量ξ_1,ξ_2，均值为零、协方差矩阵可表示为，

$$B=\begin{pmatrix}\sigma_1^2 & r\sigma_1\sigma_2\\ r\sigma_1\sigma_2 & \sigma_2^2\end{pmatrix},$$

协方差矩阵行列式可表示为，

$$|B|=\begin{vmatrix}\sigma_1^2 & r\sigma_1\sigma_2\\ r\sigma_1\sigma_2 & \sigma_2^2\end{vmatrix}=\sigma_1^2\sigma_2^2(1-r^2)$$

协方差矩阵可进行三角化分解，

$$\begin{aligned}B&=\begin{pmatrix}\sigma_1^2 & r\sigma_1\sigma_2\\ r\sigma_1\sigma_2 & \sigma_2^2\end{pmatrix}\\
&=\begin{pmatrix}1 & 0\\ r\sigma_2/\sigma_1 & 1\end{pmatrix}\begin{pmatrix}\sigma_1^2 & r\sigma_1\sigma_2\\ 0 & \sigma_2^2-r^2\sigma_2^2\end{pmatrix}\\
&=\begin{pmatrix}1 & 0\\ r\sigma_2/\sigma_1 & 1\end{pmatrix}\begin{pmatrix}\sigma_1^2 & 0\\ 0 & \sigma_2^2-r^2\sigma_2^2\end{pmatrix}\begin{pmatrix}1 & r\sigma_2/\sigma_1\\ 0 & 1\end{pmatrix}\end{aligned}$$

则，协方差矩阵的逆是，

$$B^{-1}=\begin{pmatrix}1 & -r\sigma_2/\sigma_1\\ 0 & 1\end{pmatrix}\begin{pmatrix}1/\sigma_1^2 & 0\\ 0 & 1/(1-r^2)\sigma_2^2\end{pmatrix}\begin{pmatrix}1 & 0\\ -r\sigma_2/\sigma_1 & 1\end{pmatrix}$$

$$B^{-1}=\frac{1}{1-r^2}\begin{pmatrix}1/\sigma_1^2 & -r/\sigma_1\sigma_2\\ -r/\sigma_1\sigma_2 & 1/\sigma_2^2\end{pmatrix}$$

二元高斯随机变量条件分布和边缘分布

考虑到

$$\begin{aligned}(x_1\ x_2)B^{-1}\begin{pmatrix}x_1\\ x_2\end{pmatrix}&=(x_1\ x_2)\begin{pmatrix}1 & -r\sigma_2/\sigma_1\\ 0 & 1\end{pmatrix}\\
&\quad\begin{pmatrix}1/\sigma_1^2 & 0\\ 0 & 1/(1-r^2)\sigma_2^2\end{pmatrix}\begin{pmatrix}1 & 0\\ -r\sigma_2/\sigma_1 & 1\end{pmatrix}\begin{pmatrix}x_1\\ x_2\end{pmatrix}\end{aligned}$$

$$\left(x_1\ x_2\right)B^{-1}\begin{pmatrix}x_1\\x_2\end{pmatrix}=\left(x_1\quad x_2-x_1 r\sigma_2/\sigma_1\right)$$

$$\begin{pmatrix}1/\sigma_1^2 & 0\\ 0 & 1/(1-r^2)\sigma_2^2\end{pmatrix}\begin{pmatrix}x_1\\ x_2-x_1 r\sigma_2/\sigma_1\end{pmatrix}$$

$$\left(x_1\ x_2\right)B^{-1}\begin{pmatrix}x_1\\x_2\end{pmatrix}=x_1{}^2/\sigma_1^2+\left(x_2-x_1 r\sigma_2/\sigma_1\right)^2/(1-r^2)\sigma_2^2$$

即，

$$\left(x_1\ x_2\right)B^{-1}\begin{pmatrix}x_1\\x_2\end{pmatrix}=\frac{1}{1-r^2}\left(x_1\ x_2\right)\begin{pmatrix}1/\sigma_1^2 & -r/\sigma_1\sigma_2\\ -r/\sigma_1\sigma_2 & 1/\sigma_2^2\end{pmatrix}\begin{pmatrix}x_1\\x_2\end{pmatrix}$$

$$=\frac{1}{1-r^2}\left(x_1\ x_2\right)\begin{pmatrix}x_1/\sigma_1^2 & -x_2 r/\sigma_1\sigma_2\\ -x_1 r/\sigma_1\sigma_2 & x_2/\sigma_2^2\end{pmatrix}$$

$$=\frac{1}{1-r^2}[x_1{}^2/\sigma_1^2-x_2 r/\sigma_1\sigma_2-x_2 r/\sigma_1\sigma_2+x_2{}^2/\sigma_2^2]$$

二元高斯随机变量ξ_1,ξ_2，条件分布和边缘分布是，

$$f_{\xi_1\xi_2}(x_1,x_2)=\frac{1}{2\pi\sqrt{\left(1-r^2\right)\sigma_1^2\sigma_2^2}}$$

$$\exp\left(-\frac{1}{2}[x_1{}^2/\sigma_1^2+\left(x_2-x_1 r\sigma_2/\sigma_1\right)^2/(1-r^2)\sigma_2^2]\right)$$

$$=\frac{1}{\sqrt{2\pi\sigma_1^2}}\exp\left(-\frac{1}{2}x_1{}^2/\sigma_1^2\right)\cdot$$

$$\frac{1}{\sqrt{2\pi\left(1-r^2\right)\sigma_2^2}}\exp\left(-\frac{1}{2}\left(x_2-x_1 r\sigma_2/\sigma_1\right)^2/(1-r^2)\sigma_2^2\right)$$

$$=f_{\xi_1}(x_1)\cdot f_{\xi_2/\xi_1}(x_2/x_1)$$

$$f_{\xi_2/\xi_1}(x_2/x_1)$$

$$=\frac{1}{\sqrt{2\pi\left(1-r^2\right)\sigma_2^2}}\exp\left(-\frac{1}{2}\left(x_2-x_1 r\sigma_2/\sigma_1\right)^2/(1-r^2)\sigma_2^2\right)$$

$$f_{\xi_1}(x_1)=\frac{1}{\sqrt{2\pi\sigma_1^2}}\exp\left(-\frac{1}{2}x_1{}^2/\sigma_1^2\right)\cdot$$

二元高斯随机变量ξ_1,ξ_2，条件均值和条件期望值是，

$$E\left\{\xi_2/\xi_1=x_1\right\}=E\left\{\xi_2\right\}+x_1 r\sigma_2/\sigma_1$$

5.3 二个多元高斯分布随机变量的条件分布和边缘分布

二个多元高斯随机变量协方差矩阵三角化分解和逆矩阵

二个多元高斯随机变量Ξ_1,Ξ_2，均值为零，联合协方差矩阵可进行三角化分解，

$$B=\begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix}$$
$$=\begin{pmatrix} I_n & 0 \\ B_{21}B_{11}^{-1} & I_m \end{pmatrix}\begin{pmatrix} B_{11} & B_{12} \\ 0 & B_{22}-B_{21}B_{11}^{-1}B_{12} \end{pmatrix}$$
$$=\begin{pmatrix} I_n & 0 \\ B_{21}B_{11}^{-1} & I_m \end{pmatrix}\begin{pmatrix} B_{11} & 0 \\ 0 & B_{22}-B_{21}B_{11}^{-1}B_{12} \end{pmatrix}\begin{pmatrix} I_n & B_{11}^{-1}B_{12} \\ 0 & I_m \end{pmatrix}$$

则，协方差矩阵的逆是，

$$B^{-1}=\begin{pmatrix} I_n & -B_{11}^{-1}B_{12} \\ 0 & I_m \end{pmatrix}\begin{pmatrix} B_{11}^{-1} & 0 \\ 0 & (B_{22}-B_{21}B_{11}^{-1}B_{12})^{-1} \end{pmatrix}\begin{pmatrix} I_n & 0 \\ -B_{21}B_{11}^{-1} & I_m \end{pmatrix}$$

二个多元高斯随机变量Ξ_1,Ξ_2，协方差矩阵的行列式可表示为，

$$|B|=\begin{vmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{vmatrix}=\left|\begin{pmatrix} I_n & 0 \\ B_{21}B_{11}^{-1} & I_m \end{pmatrix}\begin{pmatrix} B_{11} & 0 \\ 0 & B_{22}-B_{21}B_{11}^{-1}B_{12} \end{pmatrix}\begin{pmatrix} I_n & B_{11}^{-1}B_{12} \\ 0 & I_m \end{pmatrix}\right|$$
$$=|B_{11}||B_{22}-B_{21}B_{11}^{-1}B_{12}|$$

二个多元高斯随机变量条件分布和边缘分布，考虑到

$$\begin{pmatrix}\mathbf{x}_1^T & \mathbf{x}_2^T\end{pmatrix}B^{-1}\begin{pmatrix}\mathbf{x}_1 \\ \mathbf{x}_2\end{pmatrix}=\begin{pmatrix}\mathbf{x}_1^T & \mathbf{x}_2^T\end{pmatrix}\begin{pmatrix} I_n & -B_{11}^{-1}B_{12} \\ 0 & I_m \end{pmatrix}$$
$$\begin{pmatrix} B_{11}^{-1} & 0 \\ 0 & (B_{22}-B_{21}B_{11}^{-1}B_{12})^{-1} \end{pmatrix}\begin{pmatrix} I_n & 0 \\ -B_{21}B_{11}^{-1} & I_m \end{pmatrix}\begin{pmatrix}\mathbf{x}_1 \\ \mathbf{x}_2\end{pmatrix}$$

$$\begin{pmatrix}\mathbf{x}_1^T & \mathbf{x}_2^T\end{pmatrix}B^{-1}\begin{pmatrix}\mathbf{x}_1 \\ \mathbf{x}_2\end{pmatrix}=\begin{pmatrix}\mathbf{x}_1^T & \mathbf{x}_2^T-\mathbf{x}_1^TB_{11}^{-1}B_{12}\end{pmatrix}$$
$$\begin{pmatrix} B_{11}^{-1} & 0 \\ 0 & (B_{22}-B_{21}B_{11}^{-1}B_{12})^{-1} \end{pmatrix}\begin{pmatrix}\mathbf{x}_1 \\ \mathbf{x}_2-B_{21}B_{11}^{-1}\mathbf{x}_1\end{pmatrix}$$

$$\begin{pmatrix}\mathbf{x}_1^T & \mathbf{x}_2^T\end{pmatrix}B^{-1}\begin{pmatrix}\mathbf{x}_1 \\ \mathbf{x}_2\end{pmatrix}=\mathbf{x}_1^TB_{11}^{-1}\mathbf{x}_1$$
$$+(\mathbf{x}_2^T-\mathbf{x}_1^TB_{11}^{-1}B_{12})(B_{22}-B_{21}B_{11}^{-1}B_{12})^{-1}(\mathbf{x}_2-B_{21}B_{11}^{-1}\mathbf{x}_1)$$

二个多元高斯随机变量Ξ_1,Ξ_2的条件分布和边缘分布是，

$$f_{\Xi_1\Xi_2}(\mathbf{x}_1,\mathbf{x}_2)=\frac{1}{(2\pi)^{N/2}\left|B\right|^{N/2}}\exp\left\{-\frac{1}{2}[\mathbf{x}_1{}^T B_{11}{}^{-1}\mathbf{x}_1]-\frac{1}{2}\right.$$

$$\left.[(\mathbf{x}_2^T-\mathbf{x}_1^T B_{11}^{-1}B_{12}][B_{22}-B_{21}B_{11}^{-1}B_{12}]^{-1}[\mathbf{x}_2-B_{21}B_{11}^{-1}\mathbf{x}_1]\right\}$$

$$=\frac{1}{(2\pi)^{n/2}\left|B_{11}\right|^{n/2}}\exp\left\{-\frac{1}{2}[\mathbf{x}_1{}^T B_{11}{}^{-1}\mathbf{x}_1]\right\}\cdot\frac{1}{(2\pi)^{m/2}\left|B_{22}-B_{21}B_{11}^{-1}B_{12}\right|^{m/2}}$$

$$\exp\left\{-\frac{1}{2}[(\mathbf{x}_2^T-\mathbf{x}_1^T B_{11}^{-1}B_{12}][B_{22}-B_{21}B_{11}^{-1}B_{12}]^{-1}[\mathbf{x}_2-B_{21}B_{11}^{-1}\mathbf{x}_1]\right\}$$

$$=f_{\Xi_1}(\mathbf{x}_1)\cdot f_{\Xi_2/\Xi_1}(\mathbf{x}_2/\mathbf{x}_1)$$

$$f_{\Xi_2/\Xi_1}(\mathbf{x}_2/\mathbf{x}_1)=\frac{1}{(2\pi)^{m/2}\left|B_{22}-B_{21}B_{11}^{-1}B_{12}\right|^{m/2}}$$

$$\exp\left\{-\frac{1}{2}[(\mathbf{x}_2^T-\mathbf{x}_1^T B_{11}^{-1}B_{12}][B_{22}-B_{21}B_{11}^{-1}B_{12}]^{-1}[\mathbf{x}_2-B_{21}B_{11}^{-1}\mathbf{x}_1]\right\}$$

$$f_{\Xi_1}(\mathbf{x}_1)=\frac{1}{(2\pi)^{n/2}\left|B_{11}\right|^{n/2}}\exp\left\{-\frac{1}{2}[\mathbf{x}_1{}^T B_{11}{}^{-1}\mathbf{x}_1]\right\}\cdot$$

二个多元高斯随机变量Ξ_1,Ξ_2，条件均值和条件期望值、条件方差是，

$$E\{\Xi_2/\Xi_1=\mathbf{x}_1\}=E\{\Xi_2\}+B_{21}B_{11}^{-1}\mathbf{x}_1$$

$$Var\{\Xi_2/\Xi_1=\mathbf{x}_1\}=B_{22}-B_{21}B_{11}^{-1}B_{12}$$