3.4 Multivariate normal distributions

3.4.1 Density functions and characteristic functions

3.4 Multivariate normal distributions

3.4.1 Density functions and characteristic functions

The pdf of n-dimensional normal distribution $N(\boldsymbol{a}, B)$ (B is positive definite symmetric matrix) :

$$p(\boldsymbol{x}) = \frac{1}{(2\pi)^{n/2}|B|^{1/2}} \exp\{-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{a})' B^{-1} (\boldsymbol{x}-\boldsymbol{a})\}.$$

3.4 Multivariate normal distributions

3.4.1 Density functions and characteristic functions

The pdf of n-dimensional normal distribution $N(\boldsymbol{a}, B)$ (B is positive definite symmetric matrix) :

$$p(\boldsymbol{x}) = \frac{1}{(2\pi)^{n/2}|B|^{1/2}} \exp\{-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{a})'B^{-1}(\boldsymbol{x}-\boldsymbol{a})\}.$$

Its c.f. is

$$f(\boldsymbol{t}) = \exp(i\boldsymbol{t}'\boldsymbol{a} - \frac{1}{2}\boldsymbol{t}'B\boldsymbol{t}),$$

i.e.,

$$f(t_1, \cdots, t_n) = \exp(i\sum_{k=1}^{n} a_k t_k - \frac{1}{2}\sum_{l=1}^{n}\sum_{s=1}^{n} b_{ls} t_l t_s).$$

Proof. Write $B = LL'$ $(L = B^{1/2})$. Let $\boldsymbol{\eta} = L^{-1}(\boldsymbol{\xi} - \boldsymbol{a})$. Then by Theorem 2 in §2.5, the pdf of $\boldsymbol{\eta}$ is

$$p_{\boldsymbol{\eta}}(\boldsymbol{y}) = p(\boldsymbol{x})|L| \ \ \Big(\text{where} \quad \boldsymbol{x} = L(\boldsymbol{y} + \boldsymbol{a})\Big)$$

Proof. Write $B = LL'$ $(L = B^{1/2})$. Let $\boldsymbol{\eta} = L^{-1}(\boldsymbol{\xi} - \boldsymbol{a})$. Then by Theorem 2 in §2.5, the pdf of $\boldsymbol{\eta}$ is

$$
\begin{aligned}
& p_{\boldsymbol{\eta}}(\boldsymbol{y}) = p(\boldsymbol{x})|L| \ \ \big(\text{where} \quad \boldsymbol{x} = L(\boldsymbol{y} + \boldsymbol{a})\big) \\
= \ \ & \frac{1}{(\sqrt{2\pi})^{n/2}} \exp\{-\frac{1}{2}(\boldsymbol{x} - \boldsymbol{a})'(L')^{-1}L^{-1}(\boldsymbol{x} - \boldsymbol{a})\}
\end{aligned}
$$

Proof. Write $B = LL'$ $(L = B^{1/2})$. Let $\boldsymbol{\eta} = L^{-1}(\boldsymbol{\xi} - \boldsymbol{a})$. Then by Theorem 2 in §2.5, the pdf of $\boldsymbol{\eta}$ is

$$
\begin{aligned}
& p_{\boldsymbol{\eta}}(\boldsymbol{y}) = p(\boldsymbol{x})|L| \ \Big(\text{where} \quad \boldsymbol{x} = L(\boldsymbol{y} + \boldsymbol{a})\Big) \\
= \ & \frac{1}{(\sqrt{2\pi})^{n/2}} \exp\{-\frac{1}{2}(\boldsymbol{x} - \boldsymbol{a})'(L')^{-1}L^{-1}(\boldsymbol{x} - \boldsymbol{a})\} \\
= \ & \frac{1}{(\sqrt{2\pi})^{n/2}} \exp\{-\frac{1}{2}\boldsymbol{y}'\boldsymbol{y}\} = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}} \exp\{-\frac{y_i^2}{2}\},
\end{aligned}
$$

i.e., $\eta_1, \cdots, \eta_n$ i.i.d. $\sim N(0,1)$. From Property 3′ in §3.3 it follows that

$$
f_{\boldsymbol{\eta}}(\boldsymbol{t}) = \prod_{i=1}^{n} e^{-\frac{t_i^2}{2}} = \exp\{-\frac{1}{2}\boldsymbol{t}'\boldsymbol{t}\}.
$$

Also $\boldsymbol{\xi} = L\boldsymbol{\eta} + \boldsymbol{a}$. It follows that

$$\begin{aligned} f(\boldsymbol{t}) &= Ee^{i\boldsymbol{t}'\boldsymbol{\xi}} = e^{i\boldsymbol{t}'\boldsymbol{a}} Ee^{i\boldsymbol{t}'L\boldsymbol{\eta}} = e^{i\boldsymbol{t}'\boldsymbol{a}} Ee^{i(L'\boldsymbol{t})'\boldsymbol{\eta}} \\ &= e^{i\boldsymbol{t}'\boldsymbol{a}} \exp\{-\frac{1}{2}(L'\boldsymbol{t})'(L'\boldsymbol{t})\} \\ &= e^{i\boldsymbol{t}'\boldsymbol{a}} \exp\{-\frac{1}{2}\boldsymbol{t}'LL'\boldsymbol{t})\} \\ &= e^{i\boldsymbol{t}'\boldsymbol{a}} \exp\{-\frac{1}{2}\boldsymbol{t}'B\boldsymbol{t}\} \\ &= \exp\{i\boldsymbol{t}'\boldsymbol{a} - \frac{1}{2}\boldsymbol{t}'B\boldsymbol{t}\}. \end{aligned}$$

When B is non-negative definite,

$$f(\boldsymbol{t}) = \exp(i\boldsymbol{t}'\boldsymbol{a} - \frac{1}{2}\boldsymbol{t}'B\boldsymbol{t})$$

is also a c.f.. In fact, Write $B = LL'$, if $\boldsymbol{\eta} = N(\boldsymbol{0}, \boldsymbol{I}_{n\times n})$, then the c.f. of $\boldsymbol{\xi} = L\boldsymbol{\eta} + \boldsymbol{a}$ is $f(\boldsymbol{t})$.

When B is non-negative definite,

$$f(\boldsymbol{t}) = \exp(i\boldsymbol{t}'\boldsymbol{a} - \frac{1}{2}\boldsymbol{t}'B\boldsymbol{t})$$

is also a c.f.. In fact, Write $B = LL'$, if $\boldsymbol{\eta} = N(\mathbf{0}, \boldsymbol{I}_{n\times n})$, then the c.f. of $\boldsymbol{\xi} = L\boldsymbol{\eta} + \boldsymbol{a}$ is $f(\boldsymbol{t})$.

We call the corresponding distribution a singular normal distribution or a degenerate normal distribution. When the rank of B is r $(r < n)$, it is actually only a distribution in r dimensional subspace.

3.4.2 Properties

1. Any sub-vector $(\xi_{l_1}, \cdots, \xi_{l_k})'$ of $\boldsymbol{\xi}$ also follows normal distribution as $N(\tilde{\boldsymbol{a}}, \tilde{B})$, where $\tilde{\boldsymbol{a}} = (a_{l_1}, \cdots, a_{l_k})'$, $\tilde{B}$ is a $k \times k$ matrix consisting of elements in both $l_1, \cdots, l_k$ rows and $l_1, \cdots, l_k$ columns in B. $N(\boldsymbol{a}, B)$ has expected value $\boldsymbol{a}$, covariance matrix B.

3.4.2 Properties

1. Any sub-vector $(\xi_{l_1}, \cdots, \xi_{l_k})'$ of $\boldsymbol{\xi}$ also follows normal distribution as $N(\tilde{\boldsymbol{a}}, \tilde{B})$, where $\tilde{\boldsymbol{a}} = (a_{l_1}, \cdots, a_{l_k})'$, $\tilde{B}$ is a $k \times k$ matrix consisting of elements in both $l_1, \cdots, l_k$ rows and $l_1, \cdots, l_k$ columns in B. $N(\boldsymbol{a}, B)$ has expected value $\boldsymbol{a}$, covariance matrix B.

3.4.2 Properties

1. Any sub-vector $(\xi_{l_1}, \cdots, \xi_{l_k})'$ of $\boldsymbol{\xi}$ also follows normal distribution as $N(\tilde{\boldsymbol{a}}, \tilde{B})$, where $\tilde{\boldsymbol{a}} = (a_{l_1}, \cdots, a_{l_k})'$, $\tilde{B}$ is a $k \times k$ matrix consisting of elements in both $l_1, \cdots, l_k$ rows and $l_1, \cdots, l_k$ columns in B. $N(\boldsymbol{a}, B)$ has expected value $\boldsymbol{a}$, covariance matrix B.

Proof. In the cf of $\boldsymbol{\xi}$: $f_{\boldsymbol{\xi}}(\boldsymbol{t}) = \exp\left\{i\boldsymbol{t}'\boldsymbol{a} - \frac{1}{2}\boldsymbol{t}'B\boldsymbol{t}\right\}$, setting all t_j except $t_{l_1}, \cdots, t_{l_k}$ to be 0 yields the cf of $(\xi_{l_1}, \cdots, \xi_{l_k})'$: $\exp\left\{i\tilde{\boldsymbol{t}}'\tilde{\boldsymbol{a}} - \frac{1}{2}\tilde{\boldsymbol{t}}'\tilde{B}\tilde{\boldsymbol{t}}\right\}$.

2. $N(\boldsymbol{a}, \boldsymbol{B})$ has expected value $\boldsymbol{a}$, covariance matrix B.

2. $N(\boldsymbol{a}, \boldsymbol{B})$ has expected value $\boldsymbol{a}$, covariance matrix B.

2. $N(\boldsymbol{a}, \boldsymbol{B})$ has expected value $\boldsymbol{a}$, covariance matrix B.

Proof. If $\boldsymbol{B}$ is non-singular, the proof is already given in Section 3.2. When $\boldsymbol{B}$ is singular, suppose $\boldsymbol{\xi} \sim N(\boldsymbol{a}, \boldsymbol{B})$, $\boldsymbol{\eta} \sim N(\mathbf{0}, \boldsymbol{I})$, $\boldsymbol{\xi}$ and $\boldsymbol{\eta}$ are independent.

2. $N(\boldsymbol{a}, \boldsymbol{B})$ has expected value $\boldsymbol{a}$, covariance matrix B.

Proof. If $\boldsymbol{B}$ is non-singular, the proof is already given in Section 3.2. When $\boldsymbol{B}$ is singular, suppose $\boldsymbol{\xi} \sim N(\boldsymbol{a}, \boldsymbol{B})$, $\boldsymbol{\eta} \sim N(\boldsymbol{0}, \boldsymbol{I})$, $\boldsymbol{\xi}$ and $\boldsymbol{\eta}$ are independent. Then the cf of $\zeta =: \boldsymbol{\xi} + \boldsymbol{\eta}$ is

$$f_{\zeta}(\boldsymbol{t}) = f_{\boldsymbol{\xi}}(\boldsymbol{t}) f_{\boldsymbol{\eta}}(\boldsymbol{t}) = \exp\left\{ i\boldsymbol{t}'\boldsymbol{a} - \frac{1}{2}\boldsymbol{t}'\boldsymbol{B}\boldsymbol{t} - \frac{1}{2}\boldsymbol{t}'\boldsymbol{I}\boldsymbol{t} \right\}$$
$$= \exp\left\{ i\boldsymbol{t}'\boldsymbol{a} - \frac{1}{2}\boldsymbol{t}'(\boldsymbol{B} + \boldsymbol{I})\boldsymbol{t} \right\}.$$

It follows that $\zeta \sim N(\boldsymbol{a}, \boldsymbol{B}+\boldsymbol{I})$ and $\boldsymbol{B}+\boldsymbol{I}$ is non-singular.

It follows that $\zeta \sim N(\boldsymbol{a}, \boldsymbol{B}+\boldsymbol{I})$ and $\boldsymbol{B}+\boldsymbol{I}$ is non-singular.So

$$E\zeta = \boldsymbol{a} \quad \text{and} \quad Var\zeta = \boldsymbol{B}+\boldsymbol{I}.$$

It follows that $\boldsymbol{\zeta} \sim N(\boldsymbol{a}, \boldsymbol{B}+\boldsymbol{I})$ and $\boldsymbol{B}+\boldsymbol{I}$ is non-singular.So

$$E\boldsymbol{\zeta} = \boldsymbol{a} \quad \text{and} \quad Var\boldsymbol{\zeta} = \boldsymbol{B}+\boldsymbol{I}.$$

On the other hand,

$$E\boldsymbol{\zeta} = E\boldsymbol{\xi} + E\boldsymbol{\eta} = E\boldsymbol{\xi} + \mathbf{0}$$

and

$$Var\boldsymbol{\zeta} = Var\boldsymbol{\xi} + Var\boldsymbol{\eta} = Var\boldsymbol{\xi} + \boldsymbol{I}.$$

It follows that $\zeta \sim N(\boldsymbol{a}, \boldsymbol{B}+\boldsymbol{I})$ and $\boldsymbol{B}+\boldsymbol{I}$ is non-singular.So

$$E\boldsymbol{\zeta} = \boldsymbol{a} \quad \text{and} \quad Var\boldsymbol{\zeta} = \boldsymbol{B}+\boldsymbol{I}.$$

On the other hand,

$$E\boldsymbol{\zeta} = E\boldsymbol{\xi} + E\boldsymbol{\eta} = E\boldsymbol{\xi} + \mathbf{0}$$

and

$$Var\boldsymbol{\zeta} = Var\boldsymbol{\xi} + Var\boldsymbol{\eta} = Var\boldsymbol{\xi} + \boldsymbol{I}.$$

Hence

$$E\boldsymbol{\xi} = \boldsymbol{a} \quad \text{and} \quad Var\boldsymbol{\xi} = \boldsymbol{B}.$$

3. $\xi_1, \cdots, \xi_n$ with joint normal distribution are mutually independent iff they are pairwise uncorrelated. (Proof. Omitted.)

4. Suppose $\boldsymbol{\xi} = (\xi_1, \cdots, \xi_n) \sim N(\boldsymbol{a}, B)$, $C = (c_{ij})_{m\times n}$ is an $m \times n$ matrix, then

$$\boldsymbol{\eta} = C\boldsymbol{\xi} + \boldsymbol{\mu} \sim N(C\boldsymbol{a} + \boldsymbol{\mu}, CBC'),$$

an m-dimensional normal distribution.

4. Suppose $\boldsymbol{\xi} = (\xi_1, \cdots, \xi_n) \sim N(\boldsymbol{a}, B)$, $C = (c_{ij})_{m\times n}$ is an $m \times n$ matrix, then

$$\boldsymbol{\eta} = C\boldsymbol{\xi} + \boldsymbol{\mu} \sim N(C\boldsymbol{a} + \boldsymbol{\mu}, CBC'),$$

an m-dimensional normal distribution.

4. Suppose $\boldsymbol{\xi} = (\xi_1, \cdots, \xi_n) \sim N(\boldsymbol{a}, B)$, $C = (c_{ij})_{m \times n}$ is an $m \times n$ matrix, then

$$\boldsymbol{\eta} = C\boldsymbol{\xi} + \boldsymbol{\mu} \sim N(C\boldsymbol{a} + \boldsymbol{\mu}, CBC'),$$

an m-dimensional normal distribution.

Proof.

$$f_{\boldsymbol{\eta}}(\boldsymbol{t}) \;=\; Ee^{i\boldsymbol{t}'(C\boldsymbol{\xi}+\boldsymbol{\mu})} = e^{i\boldsymbol{t}'\boldsymbol{u}}Ee^{i(C'\boldsymbol{t})'\boldsymbol{\xi}}$$

4. Suppose $\boldsymbol{\xi} = (\xi_1, \cdots, \xi_n) \sim N(\boldsymbol{a}, B)$, $C = (c_{ij})_{m\times n}$ is an $m \times n$ matrix, then

$$\boldsymbol{\eta} = C\boldsymbol{\xi} + \boldsymbol{\mu} \sim N(C\boldsymbol{a} + \boldsymbol{\mu}, CBC'),$$

an m-dimensional normal distribution.

Proof.

$$\begin{aligned} f_{\boldsymbol{\eta}}(\boldsymbol{t}) &= Ee^{i\boldsymbol{t}'(C\boldsymbol{\xi}+\boldsymbol{\mu})} = e^{i\boldsymbol{t}'\boldsymbol{u}}Ee^{i(C'\boldsymbol{t})'\boldsymbol{\xi}} \\ &= e^{i\boldsymbol{t}'\boldsymbol{u}}f_{\boldsymbol{\xi}}(C'\boldsymbol{t}) \end{aligned}$$

4. Suppose $\boldsymbol{\xi} = (\xi_1, \cdots, \xi_n) \sim N(\boldsymbol{a}, B)$, $C = (c_{ij})_{m\times n}$ is an $m \times n$ matrix, then

$$\boldsymbol{\eta} = C\boldsymbol{\xi} + \boldsymbol{\mu} \sim N(C\boldsymbol{a} + \boldsymbol{\mu}, CBC'),$$

an m-dimensional normal distribution.

Proof.

$$\begin{aligned} f_{\boldsymbol{\eta}}(\boldsymbol{t}) &= Ee^{i\boldsymbol{t}'(C\boldsymbol{\xi}+\boldsymbol{\mu})} = e^{i\boldsymbol{t}'\boldsymbol{u}}Ee^{i(C'\boldsymbol{t})'\boldsymbol{\xi}} \\ &= e^{i\boldsymbol{t}'\boldsymbol{u}}f_{\boldsymbol{\xi}}(C'\boldsymbol{t}) \\ &= \exp\{i\boldsymbol{t}'(\boldsymbol{C}\boldsymbol{a}+\boldsymbol{\mu}) - \frac{1}{2}\boldsymbol{t}'CBC'\boldsymbol{t}\}. \end{aligned}$$

5. $\boldsymbol{\xi}$ is normally distributed iff any linear combination of its components follows normal distributions. Specifically, let $\boldsymbol{l} = (l_1, \cdots, l_n)'$ be any n dimensional real vector, then

$$\boldsymbol{\xi} \sim N(\boldsymbol{a}, B) \Leftrightarrow \zeta = \boldsymbol{l}'\boldsymbol{\xi} \sim N(\boldsymbol{l}'\boldsymbol{a}, \boldsymbol{l}'B\boldsymbol{l})$$

$$\Leftrightarrow \zeta = \sum_{j=1}^{n} l_j \xi_j \sim N(\sum_{j=1}^{n} l_j a_j, \sum_{j=1}^{n} \sum_{k=1}^{n} l_j l_k b_{jk})$$

Proof."$\Longrightarrow$" (A special case of Property 4).
Actually,

$$f_\zeta(t) = Ee^{it\boldsymbol{l}'\boldsymbol{\xi}}$$

Proof."$\Longrightarrow$" (A special case of Property 4).
Actually,

$$
\begin{aligned}
f_\zeta(t) &= Ee^{it\boldsymbol{l}'\boldsymbol{\xi}} = \exp\left\{i(t\boldsymbol{l})'\boldsymbol{a} - \frac{1}{2}(t\boldsymbol{l})')B(t\boldsymbol{l})\right\} \\
&= \exp\left\{it(\boldsymbol{l}'\boldsymbol{a}) - \frac{1}{2}t^2\boldsymbol{l}'B\boldsymbol{l}\right\}.
\end{aligned}
$$

Proof."$\Longrightarrow$" (A special case of Property 4).

Actually,

$$\begin{aligned} f_\zeta(t) &= Ee^{it\boldsymbol{l}'\boldsymbol{\xi}} = \exp\left\{i(t\boldsymbol{l})'\boldsymbol{a} - \frac{1}{2}(t\boldsymbol{l})')B(t\boldsymbol{l})\right\} \\ &= \exp\left\{it(\boldsymbol{l}'\boldsymbol{a}) - \frac{1}{2}t^2\boldsymbol{l}'B\boldsymbol{l}\right\}. \end{aligned}$$

So $\zeta = \boldsymbol{l}'\boldsymbol{\xi} \sim N(\boldsymbol{l}'\boldsymbol{a}, \boldsymbol{l}'B\boldsymbol{l})$.

Proof."$\Longrightarrow$" (A special case of Property 4).
Actually,

$$\begin{aligned} f_\zeta(t) &= Ee^{it\boldsymbol{l}'\boldsymbol{\xi}} = \exp\left\{i(t\boldsymbol{l})'\boldsymbol{a} - \frac{1}{2}(t\boldsymbol{l})')B(t\boldsymbol{l})\right\} \\ &= \exp\left\{it(\boldsymbol{l}'\boldsymbol{a}) - \frac{1}{2}t^2\boldsymbol{l}'B\boldsymbol{l}\right\}. \end{aligned}$$

So $\zeta = \boldsymbol{l}'\boldsymbol{\xi} \sim N(\boldsymbol{l}'\boldsymbol{a}, \boldsymbol{l}'B\boldsymbol{l})$.
"$\Longleftarrow$" First, by assumption, each ξ_k is normal. So its mean and variance exists, and then $Cov\{\xi_k, \xi_j\}$ exists.

Proof."$\Longrightarrow$" (A special case of Property 4).
Actually,

$$
\begin{aligned}
f_\zeta(t) &= Ee^{it\boldsymbol{l}'\boldsymbol{\xi}} = \exp\left\{i(t\boldsymbol{l})'\boldsymbol{a} - \frac{1}{2}(t\boldsymbol{l})')B(t\boldsymbol{l})\right\} \\
&= \exp\left\{it(\boldsymbol{l}'\boldsymbol{a}) - \frac{1}{2}t^2\boldsymbol{l}'B\boldsymbol{l}\right\}.
\end{aligned}
$$

So $\zeta = \boldsymbol{l}'\boldsymbol{\xi} \sim N(\boldsymbol{l}'\boldsymbol{a}, \boldsymbol{l}'B\boldsymbol{l})$.
"$\Longleftarrow$" First, by assumption, each ξ_k is normal. So its mean and variance exists, and then $Cov\{\xi_k, \xi_j\}$ exists. Denote $\boldsymbol{a} = E\boldsymbol{\xi}$ and $B = Var\boldsymbol{\xi}$. We want to show that $\boldsymbol{\xi} \sim N(\boldsymbol{a}, B)$.

For any $\boldsymbol{t}$, let $\zeta = \boldsymbol{t}'\boldsymbol{\xi}$. By assumption, ζ is normal.

For any $\boldsymbol{t}$, let $\zeta = \boldsymbol{t}'\boldsymbol{\xi}$. By assumption, ζ is normal. On the other hand, $E\zeta = \boldsymbol{t}'E\boldsymbol{\xi} = \boldsymbol{t}'\boldsymbol{a}$ and $Var\zeta = \boldsymbol{t}'(Var\boldsymbol{\xi})\boldsymbol{t} = \boldsymbol{t}'B\boldsymbol{t}$.

For any $\boldsymbol{t}$, let $\zeta = \boldsymbol{t}'\boldsymbol{\xi}$. By assumption, ζ is normal. On the other hand, $E\zeta = \boldsymbol{t}'E\boldsymbol{\xi} = \boldsymbol{t}'\boldsymbol{a}$ and $Var\zeta = \boldsymbol{t}'(Var\boldsymbol{\xi})\boldsymbol{t} = \boldsymbol{t}'B\boldsymbol{t}$. It follows that

$$\zeta \sim N(\boldsymbol{t}'\boldsymbol{a}, \boldsymbol{t}'B\boldsymbol{t}).$$

Hence

$$f_{\boldsymbol{\xi}}(\boldsymbol{t}) \;=\; Ee^{i\boldsymbol{t}'\boldsymbol{\xi}} = f_\zeta(1)$$

For any $\boldsymbol{t}$, let $\zeta = \boldsymbol{t}'\boldsymbol{\xi}$. By assumption, ζ is normal. On the other hand, $E\zeta = \boldsymbol{t}'E\boldsymbol{\xi} = \boldsymbol{t}'\boldsymbol{a}$ and $Var\zeta = \boldsymbol{t}'(Var\boldsymbol{\xi})\boldsymbol{t} = \boldsymbol{t}'B\boldsymbol{t}$. It follows that

$$\zeta \sim N(\boldsymbol{t}'\boldsymbol{a}, \boldsymbol{t}'B\boldsymbol{t}).$$

Hence

$$\begin{aligned} f_{\boldsymbol{\xi}}(\boldsymbol{t}) &= Ee^{i\boldsymbol{t}'\boldsymbol{\xi}} = f_{\zeta}(1) \\ &= \exp\left\{i\boldsymbol{t}'\boldsymbol{a} - \frac{1}{2}\boldsymbol{t}'B\boldsymbol{t}\right\}. \end{aligned}$$

So, $\boldsymbol{\xi} \sim N(\boldsymbol{a}, B)$.

4. Assume that $\boldsymbol{\xi} \sim N(\boldsymbol{a}, B)$, $\boldsymbol{\xi} = (\boldsymbol{\xi}_1', \boldsymbol{\xi}_2')'$, where $\boldsymbol{\xi}_1, \boldsymbol{\xi}_2$ are k and $n-k$-dimensional sub-vectors of $\boldsymbol{\xi}$ respectively, and

$$B = \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix}.$$

Then $\boldsymbol{\xi}_1 \sim N(\boldsymbol{a}_1, B_{11})$, $\boldsymbol{\xi}_2 \sim N(\boldsymbol{a}_2, B_{22})$; and, $\boldsymbol{\xi}_1$ and $\boldsymbol{\xi}_2$ are independent if and only if $B_{12} = \mathbf{0}$ (resp. $B_{21} = \mathbf{0}$), i.e.,
$Cov\{\boldsymbol{\xi}_1, \boldsymbol{\xi}_2\} = E\left[(\boldsymbol{\xi}_1 - E\boldsymbol{\xi}_1)(\boldsymbol{\xi}_2 - E\boldsymbol{\xi}_2)'\right] = \mathbf{0}$.

Proof. The first of conclusion is obvious. And also, it is obvious that, if $\boldsymbol{\xi}_1$ and $\boldsymbol{\xi}_2$ are independent, then

$$B_{12} = E(\boldsymbol{\xi}_1 - E\boldsymbol{\xi}_1)E(\boldsymbol{\xi}_2 - E\boldsymbol{\xi}_2)' = \mathbf{0}.$$

Proof. The first of conclusion is obvious. And also, it is obvious that, if $\boldsymbol{\xi}_1$ and $\boldsymbol{\xi}_2$ are independent, then

$$B_{12} = E(\boldsymbol{\xi}_1 - E\boldsymbol{\xi}_1)E(\boldsymbol{\xi}_2 - E\boldsymbol{\xi}_2)' = \mathbf{0}.$$

Conversely, if $B_{12} = \mathbf{0}$ and $B_{21} = \mathbf{0}$, then

$$f_{\boldsymbol{\xi}}(\boldsymbol{t}) \quad = \quad \exp\left\{ i\boldsymbol{a}'\boldsymbol{t} - \frac{1}{2}\boldsymbol{t}' B\boldsymbol{t} \right\}$$

Proof. The first of conclusion is obvious. And also, it is obvious that, if $\boldsymbol{\xi}_1$ and $\boldsymbol{\xi}_2$ are independent, then

$$B_{12} = E(\boldsymbol{\xi}_1 - E\boldsymbol{\xi}_1)E(\boldsymbol{\xi}_2 - E\boldsymbol{\xi}_2)' = \mathbf{0}.$$

Conversely, if $B_{12} = \mathbf{0}$ and $B_{21} = \mathbf{0}$, then

$$\begin{aligned} f_{\boldsymbol{\xi}}(\boldsymbol{t}) &= \exp\left\{i\boldsymbol{a}'\boldsymbol{t} - \frac{1}{2}\boldsymbol{t}'B\boldsymbol{t}\right\} \\ &= \exp\left\{i\boldsymbol{a}_1'\boldsymbol{t}_1 + i\boldsymbol{a}_2'\boldsymbol{t}_2 - \frac{1}{2}\boldsymbol{t}_1'B_{11}\boldsymbol{t}_1 - \frac{1}{2}\boldsymbol{t}_2'B_{22}\boldsymbol{t}_2\right\} \end{aligned}$$

Proof. The first of conclusion is obvious. And also, it is obvious that, if $\boldsymbol{\xi}_1$ and $\boldsymbol{\xi}_2$ are independent, then

$$B_{12} = E(\boldsymbol{\xi}_1 - E\boldsymbol{\xi}_1)E(\boldsymbol{\xi}_2 - E\boldsymbol{\xi}_2)' = \mathbf{0}.$$

Conversely, if $B_{12} = \mathbf{0}$ and $B_{21} = \mathbf{0}$, then

$$\begin{aligned} f_{\boldsymbol{\xi}}(\boldsymbol{t}) &= \exp\left\{ i\boldsymbol{a}'\boldsymbol{t} - \frac{1}{2}\boldsymbol{t}'B\boldsymbol{t} \right\} \\ &= \exp\left\{ i\boldsymbol{a}_1'\boldsymbol{t}_1 + i\boldsymbol{a}_2'\boldsymbol{t}_2 - \frac{1}{2}\boldsymbol{t}_1'B_{11}\boldsymbol{t}_1 - \frac{1}{2}\boldsymbol{t}_2'B_{22}\boldsymbol{t}_2 \right\} \\ &= f_{\boldsymbol{\xi}_1}(\boldsymbol{t}_1)f_{\boldsymbol{\xi}_2}(\boldsymbol{t}_2). \end{aligned}$$

5. Assume that $\boldsymbol{\xi} \sim N(\boldsymbol{a}, B)$, $\boldsymbol{\xi} = (\boldsymbol{\xi}_1', \boldsymbol{\xi}_2')'$, where $\boldsymbol{\xi}_1, \boldsymbol{\xi}_2$ are k and $n-k$-dimensional sub-vectors of $\boldsymbol{\xi}$ respectively,
$$B = \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix}$$
is positive definite and $\boldsymbol{\xi}_1 \sim N(\boldsymbol{a}_1, B_{11})$, $\boldsymbol{\xi}_2 \sim N(\boldsymbol{a}_2, B_{22})$. Then conditioning on $\boldsymbol{\xi}_1 = \boldsymbol{x}_1$, the conditional distribution of $\boldsymbol{\xi}_2$ is a normal distribution
$$N(\boldsymbol{a}_2 + B_{21}B_{11}^{-1}(\boldsymbol{x}_1 - \boldsymbol{a}_1), B_{22} - B_{21}B_{11}^{-1}B_{12}).$$

Proof. Let

$$\boldsymbol{\eta} = \boldsymbol{\xi}_2 - \boldsymbol{a}_2 - B_{21}B_{11}^{-1}(\boldsymbol{\xi}_1 - \boldsymbol{a}_1).$$

Then $(\boldsymbol{\xi}_1, \boldsymbol{\eta})$ is still normal random vector, and $\boldsymbol{\xi}_2 = \boldsymbol{a}_2 + B_{21}B_{11}^{-1}(\boldsymbol{\xi}_1 - \boldsymbol{a}_1) + \boldsymbol{\eta}$.

Proof. Let

$$\boldsymbol{\eta} = \boldsymbol{\xi}_2 - \boldsymbol{a}_2 - B_{21}B_{11}^{-1}(\boldsymbol{\xi}_1 - \boldsymbol{a}_1).$$

Then $(\boldsymbol{\xi}_1, \boldsymbol{\eta})$ is still normal random vector, and $\boldsymbol{\xi}_2 = \boldsymbol{a}_2 + B_{21}B_{11}^{-1}(\boldsymbol{\xi}_1 - \boldsymbol{a}_1) + \boldsymbol{\eta}$. It is easily seen that $E\boldsymbol{\eta} = \mathbf{0}$ and

$$\begin{aligned} Var\boldsymbol{\eta} =& B_{22} - 2B_{21}B_{11}^{-1}B_{12} + B_{21}B_{11}^{-1}B_{11}(B_{21}B_{11}^{-1})' \\ =& B_{22} - B_{21}B_{11}^{-1}B_{12} \widehat{=} \boldsymbol{\Sigma}. \end{aligned}$$

It follows that $\boldsymbol{\eta} \sim N(\mathbf{0}, \boldsymbol{\Sigma})$.

Also,

$$E\boldsymbol{\eta}(\boldsymbol{\xi}_1 - \boldsymbol{a}_1)' = B_{21} - B_{21}B_{11}^{-1}B_{11} = \mathbf{0}.$$

Also,

$$E\boldsymbol{\eta}(\boldsymbol{\xi}_1 - \boldsymbol{a}_1)' = B_{21} - B_{21}B_{11}^{-1}B_{11} = \mathbf{0}.$$

It follows that $\boldsymbol{\xi}_1$ and $\boldsymbol{\eta}$ are independent.

Also,

$$E\boldsymbol{\eta}(\boldsymbol{\xi}_1 - \boldsymbol{a}_1)' = B_{21} - B_{21}B_{11}^{-1}B_{11} = \mathbf{0}.$$

It follows that $\boldsymbol{\xi}_1$ and $\boldsymbol{\eta}$ are independent.

$$\boldsymbol{\eta}\Big|_{\boldsymbol{\xi}_1 = \boldsymbol{x}_1} \sim N(\mathbf{0}, \boldsymbol{\Sigma}).$$

Also,

$$E\boldsymbol{\eta}(\boldsymbol{\xi}_1 - \boldsymbol{a}_1)' = B_{21} - B_{21}B_{11}^{-1}B_{11} = \mathbf{0}.$$

It follows that $\boldsymbol{\xi}_1$ and $\boldsymbol{\eta}$ are independent.

$$\boldsymbol{\eta}\Big|_{\boldsymbol{\xi}_1=\boldsymbol{x}_1} \sim N(\mathbf{0}, \boldsymbol{\Sigma}).$$

It follows that

$$\begin{aligned}\boldsymbol{\xi}_2\Big|_{\boldsymbol{\xi}_1=\boldsymbol{x}_1} &= \boldsymbol{a}_2 + B_{21}B_{11}^{-1}(\boldsymbol{x}_1 - \boldsymbol{a}_1) + \boldsymbol{\eta}\Big|_{\boldsymbol{\xi}_1=\boldsymbol{x}_1} \\ &\sim N(\boldsymbol{a}_2 + B_{21}B_{11}^{-1}(\boldsymbol{x}_1 - \boldsymbol{a}_1), \boldsymbol{\Sigma}).\end{aligned}$$

Example

Suppose $\xi_1, \ldots, \xi_n$ be i.i.d. normal $N(\mu, \sigma^2)$ random variables. Let

$$\overline{\xi} = \frac{\sum_{k=1}^n \xi_k}{n}, \quad \widehat{\sigma}^2 = \frac{1}{n}\sum_{k=1}^n (\xi_k - \overline{\xi})^2.$$

Show that $\overline{\xi}$ and $\widehat{\sigma}^2$ are independent.

Proof. Since $(\overline{\xi}, \xi_1 - \overline{\xi}, \ldots, \xi_n - \overline{\xi})$ is a linear transform of the normal vector $(\xi_1, \ldots, \xi_n)$, so it is also a normal vector.

Proof. Since $(\overline{\xi}, \xi_1 - \overline{\xi}, \ldots, \xi_n - \overline{\xi})$ is a linear transform of the normal vector $(\xi_1, \ldots, \xi_n)$, so it is also a normal vector. On the other hand,

$$\begin{aligned} Cov\{\overline{\xi}, \xi_k - \overline{\xi}\} =& Cov\{\overline{\xi}, \xi_k\} - Var\{\overline{\xi}\} \\ =& \frac{1}{n}\sigma^2 - \frac{1}{n}\sigma^2 = 0. \end{aligned}$$

Proof. Since $(\overline{\xi}, \xi_1 - \overline{\xi}, \ldots, \xi_n - \overline{\xi})$ is a linear transform of the normal vector $(\xi_1, \ldots, \xi_n)$, so it is also a normal vector. On the other hand,

$$\begin{aligned} Cov\{\overline{\xi}, \xi_k - \overline{\xi}\} =& Cov\{\overline{\xi}, \xi_k\} - Var\{\overline{\xi}\} \\ =& \frac{1}{n}\sigma^2 - \frac{1}{n}\sigma^2 = 0. \end{aligned}$$

Hence $\overline{\xi}$ and $(\xi_1 - \overline{\xi}, \ldots, \xi_n - \overline{\xi})$ are independent.

Proof. Since $(\overline{\xi}, \xi_1 - \overline{\xi}, \ldots, \xi_n - \overline{\xi})$ is a linear transform of the normal vector $(\xi_1, \ldots, \xi_n)$, so it is also a normal vector. On the other hand,

$$\begin{aligned} Cov\{\overline{\xi}, \xi_k - \overline{\xi}\} =& Cov\{\overline{\xi}, \xi_k\} - Var\{\overline{\xi}\} \\ =& \frac{1}{n}\sigma^2 - \frac{1}{n}\sigma^2 = 0. \end{aligned}$$

Hence $\overline{\xi}$ and $(\xi_1 - \overline{\xi}, \ldots, \xi_n - \overline{\xi})$ are independent. So $\overline{\xi}$ and $\widehat{\sigma}^2$ are independent.

Example 1. Assume $\boldsymbol{\xi} = (\xi_1, \xi_2)' \sim N(a_1, a_2, \sigma^2, \sigma^2, r)$, prove $\eta_1 = \xi_1 + \xi_2$ and $\eta_2 = \xi_1 - \xi_2$ are independent, and find respective distributions of η_1, η_2.

Solution. Since (η_1, η_2) is a linear transform of (ξ_1, ξ_2), so (η_1, η_2) follows a normal distribution.

Solution. Since (η_1, η_2) is a linear transform of (ξ_1, ξ_2), so (η_1, η_2) follows a normal distribution. Also, $E\eta_1 = a_1 + a_2$, $E\eta_2 = a_1 - a_2$,

Solution. Since (η_1, η_2) is a linear transform of (ξ_1, ξ_2), so (η_1, η_2) follows a normal distribution. Also, $E\eta_1 = a_1 + a_2$, $E\eta_2 = a_1 - a_2$,

$$\begin{aligned} Var\eta_1 &= Var\xi_1 + Var\xi_2 + 2Cov\{\xi_1, \xi_2\} \\ &= 2\sigma^2 + 2r\sigma\sigma = 2\sigma^2(1+r), \end{aligned}$$

Solution. Since (η_1, η_2) is a linear transform of (ξ_1, ξ_2), so (η_1, η_2) follows a normal distribution. Also, $E\eta_1 = a_1 + a_2$, $E\eta_2 = a_1 - a_2$,

$$\begin{aligned} Var\eta_1 &= Var\xi_1 + Var\xi_2 + 2Cov\{\xi_1, \xi_2\} \\ &= 2\sigma^2 + 2r\sigma\sigma = 2\sigma^2(1 + r), \\ Var\eta_2 &= Var\xi_1 + Var\xi_2 - 2Cov\{\xi_1, \xi_2\} \\ &= 2\sigma^2 - 2r\sigma\sigma = 2\sigma^2(1 - r), \end{aligned}$$

Solution. Since (η_1, η_2) is a linear transform of (ξ_1, ξ_2), so (η_1, η_2) follows a normal distribution. Also, $E\eta_1 = a_1 + a_2$, $E\eta_2 = a_1 - a_2$,

$$\begin{aligned}
Var\eta_1 &= Var\xi_1 + Var\xi_2 + 2Cov\{\xi_1, \xi_2\} \\
&= 2\sigma^2 + 2r\sigma\sigma = 2\sigma^2(1+r), \\
Var\eta_2 &= Var\xi_1 + Var\xi_2 - 2Cov\{\xi_1, \xi_2\} \\
&= 2\sigma^2 - 2r\sigma\sigma = 2\sigma^2(1-r), \\
Cov\{\eta_1, \eta_2\} &= Var\xi_1 - Var\xi_2 = 0.
\end{aligned}$$

Solution. Since (η_1, η_2) is a linear transform of (ξ_1, ξ_2), so (η_1, η_2) follows a normal distribution. Also, $E\eta_1 = a_1 + a_2$, $E\eta_2 = a_1 - a_2$,

$$\begin{aligned}
Var\eta_1 &= Var\xi_1 + Var\xi_2 + 2Cov\{\xi_1, \xi_2\} \\
&= 2\sigma^2 + 2r\sigma\sigma = 2\sigma^2(1+r), \\
Var\eta_2 &= Var\xi_1 + Var\xi_2 - 2Cov\{\xi_1, \xi_2\} \\
&= 2\sigma^2 - 2r\sigma\sigma = 2\sigma^2(1-r), \\
Cov\{\eta_1, \eta_2\} &= Var\xi_1 - Var\xi_2 = 0.
\end{aligned}$$

So η_1 and η_2 are independent, and

$\eta_1 \sim N(a_1 + a_2, 2\sigma^2(1+r))$, $\eta_2 \sim N(a_1 - a_2, 2\sigma^2(1-r))$.