Ch. 6 Multivariate Normal Distribution

► Goldberger, Ch. 7-2,3,4

Wackerly et al. chapter 5.10

Yale Note 7, 9

► Bivariate Normal Distribution

► Random Vector

► Multivariate Normal Distribution

1. Bivariate Normal Distribution

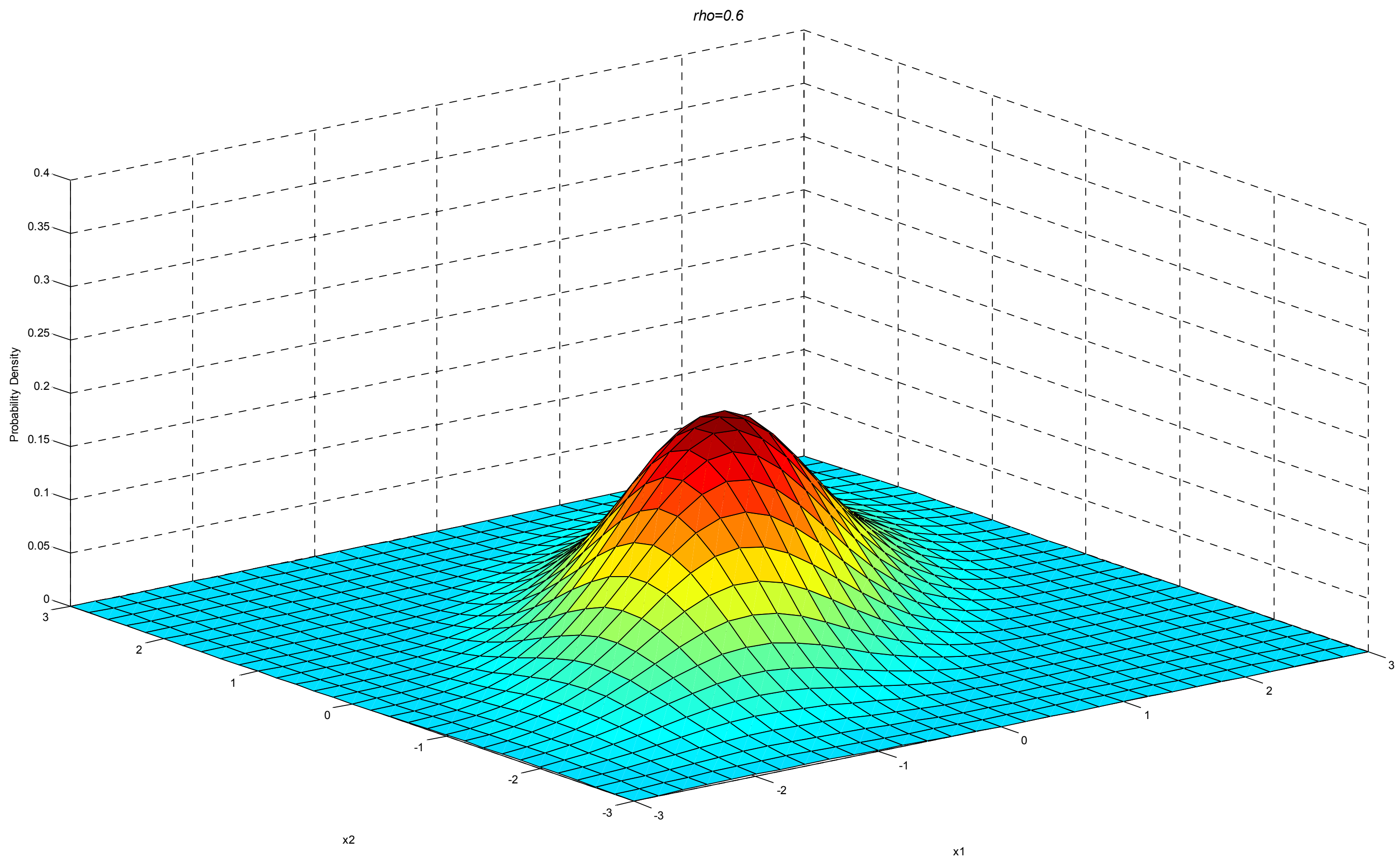
(Definition) The *bivariate normal distribution* is defined by

$$f(x,y)=\frac{1}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho^2}}\exp\left[-\frac{1}{2\left(1-\rho^2\right)}\left\{\left(\frac{x-\mu_X}{\sigma_X}\right)^2+\left(\frac{y-\mu_Y}{\sigma_Y}\right)^2-2\rho\left(\frac{x-\mu_X}{\sigma_X}\right)\left(\frac{y-\mu_Y}{\sigma_Y}\right)\right\}\right]$$

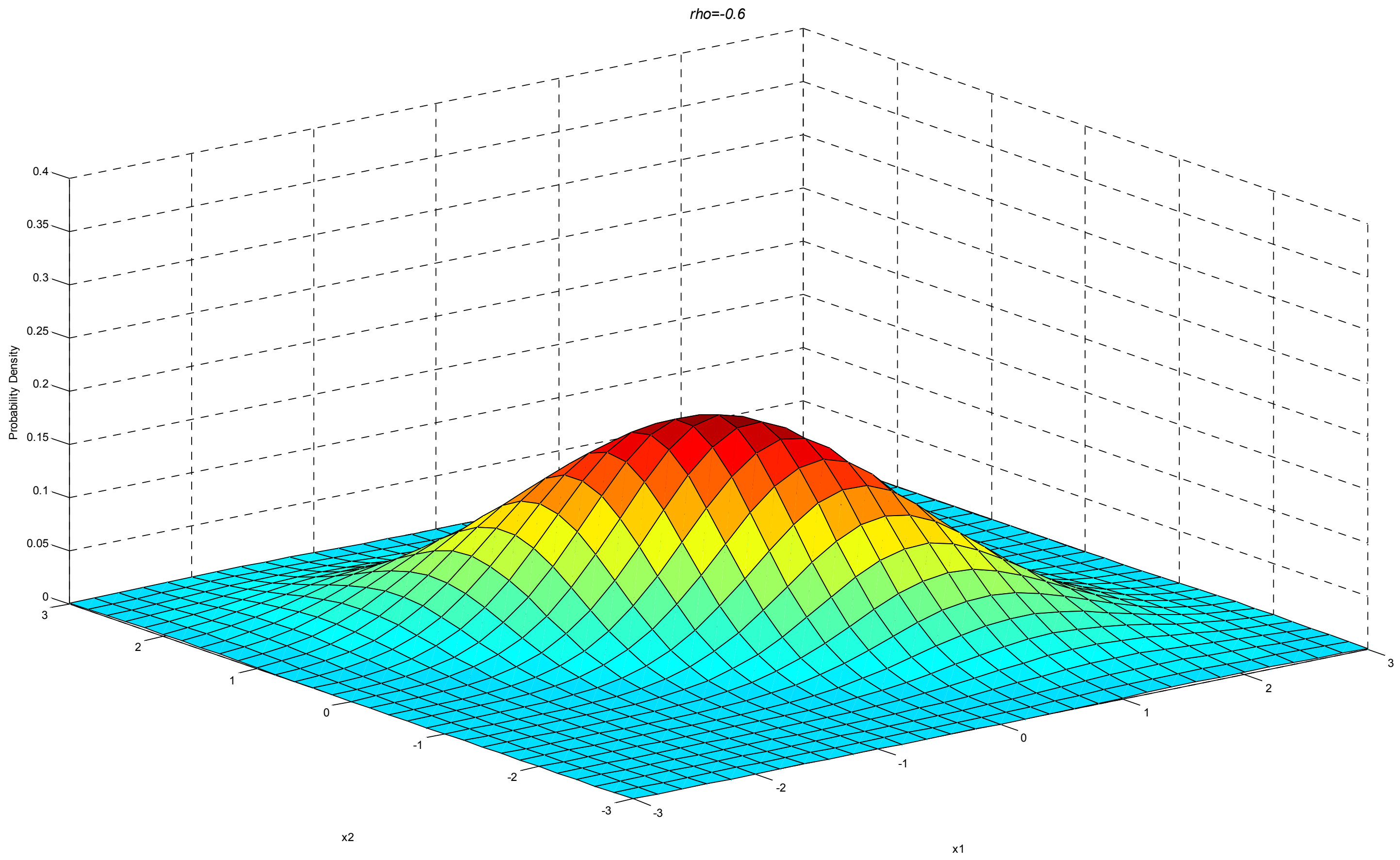
for $-\infty<x<\infty, -\infty<y<\infty$

$$\Rightarrow \begin{pmatrix} X \\ Y \end{pmatrix} \sim BVN\left(\mu_X,\mu_Y,\sigma_X^2,\sigma_Y^2,\rho\right)$$

① BVN(0, 0, 1, 1, 0.6)


rho=0.6
Probability Density
0.4
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
3
2
1
0
-1
-2
-3
-3
-2
-1
0
1
2
3
x2
x1

② BVN(0, 0, 1, 1, -0.6)


rho=-0.6
Probability Density
0
0.05
0.1
0.15
0.2
0.25
0.3
0.35
0.4
x2
x1
-3
-2
-1
0
1
2
3

◎ Notation (in vector form):

Let $\mathbf{X} = \begin{bmatrix} X_1 \\ X_2 \end{bmatrix}$, $\quad \boldsymbol{\mu} = \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}$, $\quad \Sigma = \begin{bmatrix} \sigma_1{}^2 & \sigma_{12} \\ \sigma_{21} & \sigma_2{}^2 \end{bmatrix} = \begin{bmatrix} \sigma_1{}^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2{}^2 \end{bmatrix}$,

Then $f(\mathrm{x}) = (2\pi)^{-1} |\Sigma|^{-1/2} \exp\left(-\frac{1}{2} (\mathbf{x} - \boldsymbol{\mu})' \boldsymbol{\Sigma}^{-1} (\mathbf{x} - \boldsymbol{\mu}) \right)$.

(Theorem) If $\begin{pmatrix} X \\ Y \end{pmatrix} \sim BVN\left(\mu_X, \mu_Y, \sigma_X^2, \sigma_Y^2, \rho\right)$, then

① $X \sim N(\mu_X, \sigma_X^2)$ and $Y \sim N(\mu_Y, \sigma_Y^2)$.

② $Y|X \sim N\left(E\left(Y|X\right), V\left(Y|X\right)\right)$

where $E\left(Y|X\right) = \alpha + \beta X$ with $\beta = \dfrac{\sigma_{XY}}{\sigma_X^2}$, $\alpha = \mu_Y - \beta\mu_X$ and $V\left(Y|X\right) = \sigma_Y^2\left(1-\rho^2\right)$.

③ $Corr(X,Y) = \rho$.

(Proof)

- If $\begin{pmatrix} X \\ Y \end{pmatrix} \sim BVN\left(\mu_X, \mu_Y, \sigma_X^2, \sigma_Y^2, \rho\right)$ and $\rho = 0$,

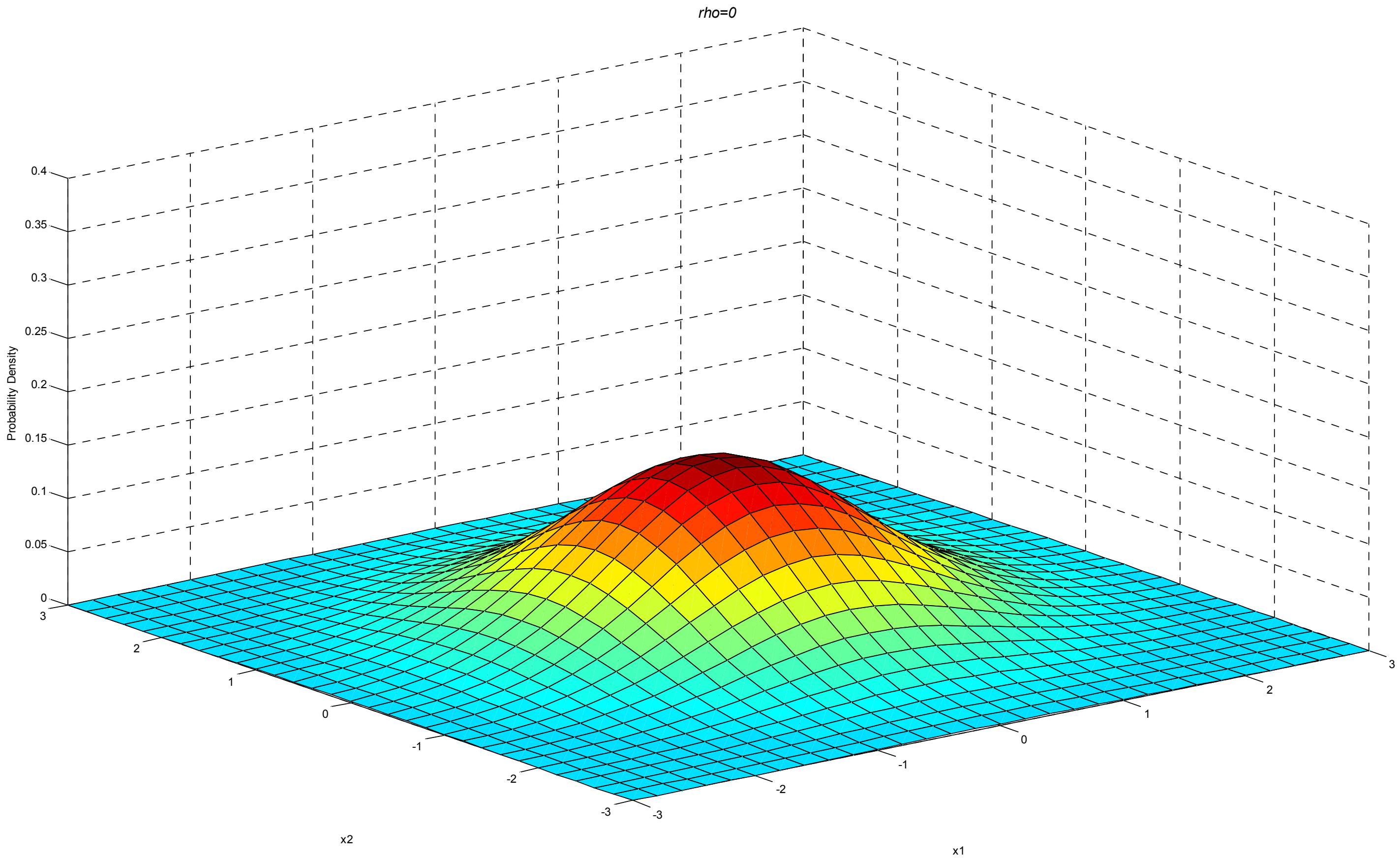
☞ $f(x, y) = f(x) f(y)$; so, independent.

► We know that independence $\Rightarrow$ uncorrelatedness($\sigma_{XY} = 0,\ \rho = 0$),

but uncorrelatedness $\Rightarrow$ independence, NOT in general.

► However, with $\begin{pmatrix} X \\ Y \end{pmatrix} \sim BVN\left(\mu_X, \mu_Y, \sigma_X^2, \sigma_Y^2, \rho\right)$, independence $\Leftrightarrow$ uncorrelatedness.

③ BVN(0, 0, 1, 1, 0)


rho=0
Probability Density
0.4
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
x2
3
2
1
0
-1
-2
-3
x1
-3
-2
-1
0
1
2
3

- If $\begin{pmatrix} X_1 \\ X_2 \end{pmatrix} \sim BVN\left(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho\right)$, $Y_1 = a_1 + b_1 X_1 + c_1 X_2,\ Y_2 = a_2 + b_2 X_1 + c_2 X_2$ and $b_1 c_2 - b_2 c_1 \neq 0$, then $\begin{pmatrix} Y_1 \\ Y_2 \end{pmatrix} \sim BVN(\bullet)$.

(Special case) $\mu_X = 0, \sigma_X^2 = 1, \mu_Y = 0, \sigma_Y^2 = 1$.

Then, $g(z_1, z_2) = \dfrac{1}{2\pi\sqrt{1-\rho^2}} \exp\left(-\dfrac{1}{2(1-\rho^2)} \left(z_1^2 + z_2^2 - 2\rho z_1 z_2 \right) \right)$

; <u>*standard bivariate normal distribution*</u>

► $\rho = 0$:

2. Random Vector

(1) Random Vector

Let $\mathbf{X}_{(n\times 1)} = \begin{pmatrix} X_1 \\ X_2 \\ \vdots \\ X_n \end{pmatrix}$, $\boldsymbol{\mu}_{(n\times 1)} = \begin{pmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_n \end{pmatrix}$, $\Sigma_{(n\times n)} = \begin{pmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1n} \\ \sigma_{21} & \sigma_{22} & \cdots & \sigma_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{n1} & \sigma_{n2} & \cdots & \sigma_{nn} \end{pmatrix}$

where σ_{ij} = element of ith row & jth column.

① $E(X_i) = \mu_i$ for all $i = 1, 2, \cdots, n \quad \Leftrightarrow \quad E(\mathbf{X}) = \boldsymbol{\mu}$; *mean vector*.

② $V(\mathbf{X}) = E\left[(\mathbf{X}-\boldsymbol{\mu})(\mathbf{X}-\boldsymbol{\mu})'\right] = \Sigma \;\Leftrightarrow\; V(X_i) = \sigma_i^2,\ Cov(X_i, X_j) = \sigma_{ij}$, $i = 1, 2, \cdots, n$.

► Diagonal terms are variance and off-diagonal terms are covariance.

So, $E\left[(\mathbf{X}-\boldsymbol{\mu})(\mathbf{X}-\boldsymbol{\mu})'\right] = \Sigma$; *variance-covariance matrix*.

2. Expectation and Variance for Functions of Random Vectors

For random vector $\boldsymbol{X}_{(n\times 1)}$ with $E(\boldsymbol{X}) = \boldsymbol{\mu}_{(n\times 1)}$ and $V(\boldsymbol{X}) = \sum_{(n\times n)}$.

① (Scalar Linear Function)

Let $z_{(1\times 1)} = g + \boldsymbol{h}'\boldsymbol{X}$ where $g_{(1\times 1)}$, $\boldsymbol{h}_{(n\times 1)}$: constant,

then $E(z) = g + \boldsymbol{h}'\boldsymbol{\mu}$, and $V(z) = \boldsymbol{h}'\sum\boldsymbol{h}$.

(proof)

② (Vector Linear Function)

Let $\boldsymbol{z}_{(k\times 1)} = \boldsymbol{g} + \boldsymbol{H}\boldsymbol{X}$ where $\boldsymbol{g}_{(k\times 1)}$, $\boldsymbol{H}_{(k\times n)}$: constant,

then $E(\boldsymbol{z}) = \boldsymbol{g} + \boldsymbol{H}\boldsymbol{\mu}$, and $V(\boldsymbol{z}) = \boldsymbol{H}\sum\boldsymbol{H}'$.

(proof)

③ (Pair of Vector Linear Functions)

Let $\boldsymbol{z}_{1(m\times 1)} = \boldsymbol{g}_1 + \boldsymbol{H}_1\boldsymbol{X}$ and $\boldsymbol{z}_{2(k\times 1)} = \boldsymbol{g}_2 + \boldsymbol{H}_2\boldsymbol{X}$,

where $\boldsymbol{g}_{1(m\times 1)}$, $\boldsymbol{g}_{2(k\times 1)}$, $\boldsymbol{H}_{1(m\times n)}$, $\boldsymbol{H}_{2(k\times n)}$: constants,

then $C(\boldsymbol{z}_1, \boldsymbol{z}_2) = \boldsymbol{H}_1 \sum \boldsymbol{H}_2{}'$.

(proof)

(Example)

$$\mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} \sim N(\boldsymbol{\mu}, \Sigma),\ \boldsymbol{\mu} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix},\ \Sigma = \begin{pmatrix} 2 & -1 & 1 \\ -1 & 5 & 1 \\ 1 & 1 & 3 \end{pmatrix}.$$

① $g = 3,\ h = \begin{pmatrix} 2 \\ 1 \\ -1 \end{pmatrix},\ z = g + h'\mathbf{y},\quad E(z),\ V(z) = ?$

(Ans) $E(z) = 4,\ V(z) = 10$

② $\mathbf{g} = \begin{pmatrix} 1 \\ -1 \\ 1 \\ 3 \end{pmatrix},\ \mathbf{H} = \begin{pmatrix} 1 & 1 & 1 \\ -1 & 1 & 2 \\ 2 & 0 & 1 \\ 1 & -1 & 2 \end{pmatrix},\ \mathbf{Z} = \mathbf{g} + \mathbf{H}\mathbf{y},\quad E(\mathbf{Z}), V(\mathbf{Z}) = ?$

(Ans) $E(\mathbf{Z}) = \begin{pmatrix} 7 \\ 6 \\ 6 \\ 8 \end{pmatrix}, \quad V(\mathbf{Z}) = \begin{pmatrix} 14 & 11 & 13 & 9 \\ 15 & 19 & 8 & 5 \\ 9 & 4 & 15 & 16 \\ 5 & 5 & 12 & 19 \end{pmatrix}$

③ $\mathbf{g}_2 = \begin{pmatrix} 2 \\ 1 \\ -1 \end{pmatrix}, \mathbf{H}_2 = \begin{pmatrix} 1 & 0 & 1 \\ -1 & 1 & 1 \\ 0 & 1 & 2 \end{pmatrix}, \ \mathbf{Z}_2 = \mathbf{g}_2 + \mathbf{H}_2\mathbf{y}, \quad C(\mathbf{Z}, \mathbf{Z}_2) = ?$

(Ans) $C(\mathbf{Z}, \mathbf{Z}_2) = \begin{pmatrix} 9 & 6 & 15 \\ 7 & 13 & 20 \\ 10 & -1 & 9 \\ 9 & -1 & 8 \end{pmatrix}$

3. Multivariate Normal Distribution

(Definition) $\mathbf{X}$ is *multivariate normal distribution* with mean $\boldsymbol{\mu}$, and variance-covariance matrix Σ, if $f(\mathbf{X})=(2\pi)^{-\frac{n}{2}}|\Sigma|^{-\frac{1}{2}}\exp\left[-\frac{1}{2}(\mathbf{X}-\boldsymbol{\mu})'\Sigma^{-1}(\mathbf{X}-\boldsymbol{\mu})\right]$.

► $\mathbf{X}\sim MVN(\boldsymbol{\mu},\Sigma)$.

① $n=1$

☞ $f(x)$

② $n=2$

☞ $f(\mathbf{X})$

◎ Let's partition into 2 parts as follows: $\mathbf{X} \sim MVN(\boldsymbol{\mu}, \Sigma)$.

$$\mathbf{X} = \begin{pmatrix} \mathbf{X}_{1(n_1\times 1)} \\ \mathbf{X}_{2(n_2\times 1)} \end{pmatrix}, \quad \boldsymbol{\mu} = \begin{pmatrix} \boldsymbol{\mu}_{1(n_1\times 1)} \\ \boldsymbol{\mu}_{2(n_2\times 1)} \end{pmatrix}, \quad \Sigma = \begin{pmatrix} \Sigma_{11(n_1\times n_1)} & \Sigma_{12(n_1\times n_2)} \\ \Sigma_{21(n_2\times n_1)} & \Sigma_{22(n_2\times n_2)} \end{pmatrix}$$ where $n_1 + n_2 = n$.

① $\mathbf{X}_1 \sim MVN(\boldsymbol{\mu}_1, \Sigma_{11})$, $\mathbf{X}_2 \sim MVN(\boldsymbol{\mu}_2, \Sigma_{22})$.

② $\mathbf{X}_2 | \mathbf{X}_1 \sim MVN\left(E(\mathbf{X}_2|\mathbf{X}_1), V(\mathbf{X}_2|\mathbf{X}_1)\right)$,

where $E(\mathbf{X}_2|\mathbf{X}_1) = \alpha + B'\mathbf{X}_1$ with $B = \Sigma_{11}^{-1}\Sigma_{12}, \alpha = \boldsymbol{\mu}_2 - B'\boldsymbol{\mu}_1$

$V(\mathbf{X}_2|\mathbf{X}_1) = \Sigma_{22} - B'\Sigma_{11}B$.

③ If $\Sigma_{12} = \Sigma'_{21} = 0$, $\mathbf{X}_1$ and $\mathbf{X}_2$ are independent.

(Example)

$$\mathbf{y}=\begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} \sim N(\boldsymbol{\mu},\Sigma),\ \boldsymbol{\mu}=\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix},\ \Sigma=\begin{pmatrix} 2 & -1 & 1 \\ -1 & 5 & 1 \\ 1 & 1 & 3 \end{pmatrix}$$

① $E\left(y_3 \middle| y_1, y_2\right)=\frac{5}{3}+\frac{2}{3}y_1+\frac{1}{3}y_2.$

② $E\left(y_3 \middle| y_1=1, y_2=1\right)=\frac{8}{3}.$

④ Linear functions of multivariate normal vectors are also multivariate normal.

• If $\mathbf{X} \sim N(\boldsymbol{\mu}, \Sigma)$ and $\mathbf{Z} = \mathbf{g} + H\mathbf{X}$ where $\mathbf{g}$ and H are nonrandom and H has full row rank, then $\mathbf{Z} \sim N(\mathbf{g} + H\boldsymbol{\mu}, H\Sigma H')$.

(Note) Full row rank =

(# of rows$\leq$# of column, and rows are linearly independent)

(Example) Short rank

$$\mathbf{g} = 0, \mathbf{H} = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}, \text{ then } \mathbf{X} = \begin{pmatrix} \mathbf{X}_1 \\ \mathbf{X}_2 \end{pmatrix} = \begin{pmatrix} X_1 + X_2 \\ X_1 + X_2 \end{pmatrix}.$$

Joint density can not be defined.

☞ Bivariate linear combination

⑤ Functions of a standard normal vector:

If $\mathbf{Z} = \begin{pmatrix} Z_1 \\ Z_2 \\ \vdots \\ Z_n \end{pmatrix} \sim N\left(\mathbf{0}, I_n\right) \;\Rightarrow\; \left(\begin{array}{l} E\left(Z_i\right)=0 \quad \textbf{for all } i \\ V\left(Z_i\right)=1 \quad \textbf{for all } i, \quad C\left(Z_i, Z_j\right)=0 \quad \textbf{for all } i \neq j \end{array} \right.'$

then $Z_1, Z_2, \cdots, Z_n$ are standard normal and independent.

$$f\left(\mathbf{Z}\right) = f(Z_1) f(Z_2) \cdots f(Z_n) = \prod_{i=1}^{n} f\left(Z_i\right) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{Z_i^2}{2}\right)$$

$$= (2\pi)^{-\frac{n}{2}} \exp\left(-\frac{\mathbf{Z}'\mathbf{Z}}{2}\right)$$

ⓐ $W = \mathbf{Z}'\mathbf{Z} = \sum_{i=1}^{n} Z_i^2 \sim \chi^2(n)$

ⓑ $v = \dfrac{W_1 / n_1}{W_2 / n_2} \sim F(n_1, n_2)$ where $W_1 = \mathbf{Z}_1'\mathbf{Z}_1 \sim \chi^2(n_1),\ \ W_2 = \mathbf{Z}_2'\mathbf{Z}_2 \sim \chi^2(n_2)$, $\mathbf{Z} = \begin{pmatrix} \mathbf{Z}_1 \\ \mathbf{Z}_2 \end{pmatrix}$.

ⓒ For any $Z_i \in \mathbf{Z}_1$, $\quad g = \dfrac{Z_i}{\sqrt{W_2 / n_2}} = \dfrac{N(0,1)}{\sqrt{\chi^2(n_2) / n_2}} \sim t(n_2)$

(Note) $g^2 = \{t(n_2)\}^2 = \dfrac{Z_i^2}{W_2 / n_2} \sim F(1, n_2)$.

(Theorem) If $\mathbf{X}_{(n\times 1)} \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, then the quadratic form

$$Y = (\mathbf{X} - \boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\mathbf{X} - \boldsymbol{\mu}) \sim \chi^2(n).$$