多変量正規分布

１．多変量正規分布の誘導

$u_i \sim N(\nu_i, 1^2)$,　独立　　$i = 1, 2, \cdots, n$　期待値 ν_i、分散 1 の正規分布の同時分布は、

$$f(u_1, u_2, \cdots, u_n)du_1 du_2 \cdots du_n = \Pi_{i=1}^{n} \frac{1}{\sqrt{2\pi}} exp\left\{-\frac{1}{2}(u_i - \nu_i)^2\right\} du_i$$

$$f(\boldsymbol{u})d\boldsymbol{u} = \left(\frac{1}{\sqrt{2\pi}}\right)^n exp\left\{-\frac{1}{2}(\boldsymbol{u} - \boldsymbol{\nu})'(\boldsymbol{u} - \boldsymbol{\nu})\right\} d\boldsymbol{u}$$

ここで正則行列 A を用いて $\boldsymbol{u}$ を独立でない要素に関する分布に拡張することを考える。

$$\begin{aligned}
f(\boldsymbol{u})d\boldsymbol{u} &= \left(\frac{1}{\sqrt{2\pi}}\right)^n exp\left\{-\frac{1}{2}(\boldsymbol{u} - \boldsymbol{\nu})'(\boldsymbol{u} - \boldsymbol{\nu})\right\} d\boldsymbol{u} \\
&= \left(\frac{1}{\sqrt{2\pi}}\right)^n exp\left\{-\frac{1}{2}(\boldsymbol{u} - \boldsymbol{\nu})'A'A'^{-1}A^{-1}A(\boldsymbol{u} - \boldsymbol{\nu})\right\} d\boldsymbol{u} \\
&= \left(\frac{1}{\sqrt{2\pi}}\right)^n exp\left\{-\frac{1}{2}(A\boldsymbol{u} - A\boldsymbol{\nu})'A'^{-1}A^{-1}(A\boldsymbol{u} - A\boldsymbol{\nu})\right\} d\boldsymbol{u} \\
&= \left(\frac{1}{\sqrt{2\pi}}\right)^n exp\left\{-\frac{1}{2}(A\boldsymbol{u} - A\boldsymbol{\nu})'(AA')^{-1}(A\boldsymbol{u} - A\boldsymbol{\nu})\right\} d\boldsymbol{u}
\end{aligned}$$

$\boldsymbol{x} = A\boldsymbol{u} \Rightarrow \boldsymbol{u} = A^{-1}\boldsymbol{x}$なる変換を行うと、　$d\boldsymbol{u} = |A|^{-1}d\boldsymbol{x}$,　付録参照

$$= \left(\frac{1}{\sqrt{2\pi}}\right)^n |A|^{-1} exp\left\{-\frac{1}{2}(\boldsymbol{x} - A\boldsymbol{\nu})'(AA')^{-1}(\boldsymbol{x} - A\boldsymbol{\nu})\right\} d\boldsymbol{x}$$

$\boldsymbol{x}$ の期待値ベクトルと分散共分散行列は、

$$\begin{aligned}
&E(\boldsymbol{x}) = E(A\boldsymbol{u}) = A\boldsymbol{\nu} \equiv \boldsymbol{\mu} \\
&E\{(\boldsymbol{x} - A\boldsymbol{\nu})(\boldsymbol{x} - A\boldsymbol{\nu})'\} = E\{A(A^{-1}\boldsymbol{x} - \boldsymbol{\nu})(A^{-1}\boldsymbol{x} - \boldsymbol{\nu})'A'\} \\
&\quad = E\{A(\boldsymbol{u} - \boldsymbol{\nu})(\boldsymbol{u} - \boldsymbol{\nu})'A'\} = AA' \equiv \boldsymbol{\Sigma} \\
&|A|^{-1} = \sqrt{|A'^{-1}||A^{-1}|} = \sqrt{|\Sigma^{-1}|} = |\Sigma|^{-\frac{1}{2}}
\end{aligned}$$

それゆえ多変量正規分布に従う $\boldsymbol{x}$ の密度関数は一般に次式で表わされる。

$$f(\boldsymbol{x}) = f(\boldsymbol{x}; \boldsymbol{\mu}, \boldsymbol{\Sigma}) = \left(\frac{1}{\sqrt{2\pi}}\right)^n |\boldsymbol{\Sigma}|^{-\frac{1}{2}} exp\left\{-\frac{1}{2}(\boldsymbol{x} - \boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\boldsymbol{x} - \boldsymbol{\mu})\right\} \quad \textbf{(1.1)}$$

このとき、$\boldsymbol{x} \sim N_n(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ と書くこともある。

２．多変量正規分布の性質

(1) $\boldsymbol{x}$ の 1 次結合は 1 変数の正規分布に従う。
$\boldsymbol{x}$ の 1 次結合の特性関数は、

$$\begin{aligned}
&E[exp\{it\boldsymbol{l}'\boldsymbol{x}\}] = \int\int\cdots\int e^{it\boldsymbol{l}'\boldsymbol{x}} f(\boldsymbol{x})d\boldsymbol{x} \\
&= \int\int\cdots\int exp\{(it\boldsymbol{l}'\boldsymbol{x}\} \left(\frac{1}{\sqrt{2\pi}}\right)^n |\boldsymbol{\Sigma}|^{-\frac{1}{2}} exp\left\{-\frac{1}{2}(\boldsymbol{x} - \boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\boldsymbol{x} - \boldsymbol{\mu})\right\} d\boldsymbol{x} \\
&= \int\int\cdots\int \left(\frac{1}{\sqrt{2\pi}}\right)^n |\boldsymbol{\Sigma}|^{-\frac{1}{2}} exp\left[-\frac{1}{2}\left\{(\boldsymbol{x} - \boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\boldsymbol{x} - \boldsymbol{\mu}) - 2it\boldsymbol{l}'\boldsymbol{x}\right\}\right] d\boldsymbol{x}
\end{aligned}$$

ここで、$\boldsymbol{u} = \boldsymbol{\Sigma}^{-\frac{1}{2}}(\boldsymbol{x} - \boldsymbol{\mu})$,　$\boldsymbol{x} = \boldsymbol{\Sigma}^{\frac{1}{2}}(\boldsymbol{u} + \boldsymbol{\mu})$,　$|\boldsymbol{\Sigma}^{-\frac{1}{2}}|d\boldsymbol{x} = d\boldsymbol{u}$　なる変換を行うと、

$$\begin{aligned}
&= \int\int\cdots\int \left(\frac{1}{\sqrt{2\pi}}\right)^n exp\left[-\frac{1}{2}\left\{\boldsymbol{u}'\boldsymbol{u} - 2it\boldsymbol{l}'(\boldsymbol{\Sigma}^{\frac{1}{2}}\boldsymbol{u} + \boldsymbol{\mu})\right\}\right] d\boldsymbol{u} \\
&= \int\int\cdots\int \left(\frac{1}{\sqrt{2\pi}}\right)^n exp\left[-\frac{1}{2}\left\{(\boldsymbol{u}' - it\boldsymbol{l}'\boldsymbol{\Sigma}^{\frac{1}{2}})'(\boldsymbol{u} - it\boldsymbol{\Sigma}^{\frac{1}{2}}\boldsymbol{l})\right\} + \frac{1}{2}(it)^2\boldsymbol{l}'\boldsymbol{\Sigma}\boldsymbol{l} + it\boldsymbol{l}'\boldsymbol{\mu}\right] d\boldsymbol{u} \\
&= exp\left\{\frac{1}{2}(it)^2\boldsymbol{l}'\boldsymbol{\Sigma}\boldsymbol{l} + it\boldsymbol{l}'\boldsymbol{\mu}\right\} = exp\left\{-\frac{1}{2}t^2\boldsymbol{l}'\boldsymbol{\Sigma}\boldsymbol{l} + it\boldsymbol{l}'\boldsymbol{\mu}\right\}
\end{aligned}$$

これは、平均 $\boldsymbol{l}'\boldsymbol{\mu}$ 分散 $\boldsymbol{l}'\boldsymbol{\Sigma}\boldsymbol{l}$ の正規分布の特性関数であるから、$\boldsymbol{l}'\boldsymbol{x} \ \sim N(\boldsymbol{l}'\boldsymbol{\mu}, \boldsymbol{l}'\boldsymbol{\Sigma}\boldsymbol{l})$ に従う。

【注意】

ここでは (1.1) に従う分布を多変量正規分布としたが、"任意の一次結合の分布が1変量の正規分布に従う"という定義の方が退化した (分散共分散行列が正則でない) 多変量正規分布にも拡張できるので、今後こちらの定義を用いる場合もある。

(2) 多変量正規分布 $\boldsymbol{x}$ の特性関数

先の式で、$t=1$、$\boldsymbol{l}=\boldsymbol{t}$ とすれば多変量正規分布の特性関数を得る。

$$E[exp\{i\boldsymbol{t}'\boldsymbol{x}\}] = exp\left\{\frac{1}{2}i\boldsymbol{t}'\boldsymbol{\Sigma}i\boldsymbol{t} + i\boldsymbol{t}'\boldsymbol{\mu}\right\} = exp\left\{-\frac{1}{2}\boldsymbol{t}'\boldsymbol{\Sigma}\boldsymbol{t} + i\boldsymbol{t}'\boldsymbol{\mu}\right\}$$

(3) 周辺分布は多変量正規分布である。

n 次元の確率ベクトルを p 次元ベクトル $\boldsymbol{x}^{(1)}$ と q 次元のベクトル $\boldsymbol{x}^{(2)}$ に分け、確率密度関数を行列式が1の正則行列 $\mathbf{E}$ で変換する。

ここで、 $\mathbf{E} = \begin{pmatrix} \mathbf{I}_p & \mathbf{O} \\ -\boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1} & \mathbf{I}_q \end{pmatrix}$, $\quad \boldsymbol{\Sigma} = \begin{pmatrix} \boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12} \\ \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22} \end{pmatrix}$, $\quad \mathbf{E}^{-1} = \begin{pmatrix} \mathbf{I}_p & \mathbf{O} \\ \boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1} & \mathbf{I}_q \end{pmatrix}$

密度関数は、

$$f(\boldsymbol{x};\boldsymbol{\mu},\boldsymbol{\Sigma}) = \left(\frac{1}{\sqrt{2\pi}}\right)^n |\boldsymbol{\Sigma}|^{-\frac{1}{2}} exp\left\{-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\boldsymbol{x}-\boldsymbol{\mu})\right\} d\boldsymbol{x}$$
$$= \left(\frac{1}{\sqrt{2\pi}}\right)^n |\boldsymbol{\Sigma}|^{-\frac{1}{2}} exp\left\{-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu})'\mathbf{E}'\mathbf{E}'^{-1}\boldsymbol{\Sigma}^{-1}\mathbf{E}^{-1}\mathbf{E}(\boldsymbol{x}-\boldsymbol{\mu})\right\} d\boldsymbol{x}$$
$$= \left(\frac{1}{\sqrt{2\pi}}\right)^n |\boldsymbol{\Sigma}|^{-\frac{1}{2}} exp\left\{-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu})'\mathbf{E}'(\mathbf{E}'^{-1}\boldsymbol{\Sigma}^{-1}\mathbf{E}^{-1})\mathbf{E}(\boldsymbol{x}-\boldsymbol{\mu})\right\} d\boldsymbol{x}$$
$$= \left(\frac{1}{\sqrt{2\pi}}\right)^n |\boldsymbol{\Sigma}|^{-\frac{1}{2}} exp\left\{-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu})'\mathbf{E}'(\mathbf{E}\boldsymbol{\Sigma}\mathbf{E}')^{-1}\mathbf{E}(\boldsymbol{x}-\boldsymbol{\mu})\right\} d\boldsymbol{x}$$
$$= \left(\frac{1}{\sqrt{2\pi}}\right)^n |\boldsymbol{\Sigma}|^{-\frac{1}{2}} exp\left[-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu})'\mathbf{E}'\left\{\begin{pmatrix} \mathbf{I}_p & \mathbf{O} \\ -\boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1} & \mathbf{I}_q \end{pmatrix}\boldsymbol{\Sigma}\begin{pmatrix} \mathbf{I}_p & -\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{\Sigma}_{12} \\ \mathbf{O} & \mathbf{I}_q \end{pmatrix}\right\}^{-1}\mathbf{E}(\boldsymbol{x}-\boldsymbol{\mu})\right] d\boldsymbol{x}$$
$$= \left(\frac{1}{\sqrt{2\pi}}\right)^n |\boldsymbol{\Sigma}|^{-\frac{1}{2}} exp\left[-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu})'\mathbf{E}'\begin{pmatrix} \boldsymbol{\Sigma}_{11} & \mathbf{O} \\ \mathbf{O} & \boldsymbol{\Sigma}_{22}-\boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{\Sigma}_{12} \end{pmatrix}^{-1}\mathbf{E}(\boldsymbol{x}-\boldsymbol{\mu})\right] d\boldsymbol{x}$$
$$= \left(\frac{1}{\sqrt{2\pi}}\right)^p |\boldsymbol{\Sigma}_{11}|^{-\frac{1}{2}} exp\left\{-\frac{1}{2}(\boldsymbol{x}^{(1)}-\boldsymbol{\mu}^{(1)})'\boldsymbol{\Sigma}_{11}^{-1}(\boldsymbol{x}^{(1)}-\boldsymbol{\mu}^{(1)})\right\} d\boldsymbol{x}^{(1)}$$
$$\times\left(\frac{1}{\sqrt{2\pi}}\right)^q |\boldsymbol{\Sigma}_{22}-\boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{\Sigma}_{12}|^{-\frac{1}{2}} exp\left\{-\frac{1}{2}(\boldsymbol{x}^{(2)}-\boldsymbol{\nu}^{(2)})'(\boldsymbol{\Sigma}_{22}-\boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{\Sigma}_{12})^{-1}(\boldsymbol{x}^{(2)}-\boldsymbol{\nu}^{(2)})\right\} d\boldsymbol{x}^{(2)}$$

このとき、$\boldsymbol{\nu}^{(2)} = \boldsymbol{\mu}^{(2)} + \boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{x}^{(1)}$である。

すなわち、

$$f(\boldsymbol{x};\boldsymbol{\mu},\boldsymbol{\Sigma})d\boldsymbol{x} = f(\boldsymbol{x}^{(1)};\boldsymbol{\mu}^{(1)},\boldsymbol{\Sigma}_{11})d\boldsymbol{x}^{(1)}\times f(\boldsymbol{x}^{(2)};\boldsymbol{\mu}^{(2)}+\boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{x}^{(1)},\boldsymbol{\Sigma}_{22}-\boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{\Sigma}_{12})d\boldsymbol{x}^{(2)}$$

と変形できる。

$$\int\int\cdots\int f(\boldsymbol{x}^{(2)};\boldsymbol{\mu}^{(2)}+\boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{x}^{(1)},\boldsymbol{\Sigma}_{22}-\boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{\Sigma}_{12})d\boldsymbol{x}^{(2)} = 1$$

であるから、

$\boldsymbol{x}^{(1)}$ の周辺分布は多変量正規分布 $\quad N_p(\boldsymbol{\mu}^{(1)},\boldsymbol{\Sigma}_{11}) \quad$ であることが分かる。

(4) $\boldsymbol{x}^{(1)} = \boldsymbol{x}_0^{(1)}$ を与えたときの $\boldsymbol{x}^{(2)}$ の条件付き分布は多変量正規分布である。

$$f(\boldsymbol{x}^{(2)}|\boldsymbol{x}_0^{(1)}) = \frac{f(\boldsymbol{x}_0^{(1)},\boldsymbol{x}^{(2)})}{f(\boldsymbol{x}_0^{(1)})}$$

$$= \frac{f(\boldsymbol{x}_0^{(1)}; \boldsymbol{\mu}^{(1)}, \boldsymbol{\Sigma}_{11}) \times f(\boldsymbol{x}^{(2)}; \boldsymbol{\mu}^{(2)} + \boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{x}_0^{(1)}, \boldsymbol{\Sigma}_{22} - \boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{\Sigma}_{12})}{f(\boldsymbol{x}_0^{(1)}; \boldsymbol{\mu}^{(1)}, \boldsymbol{\Sigma}_{11})}$$

$$= f(\boldsymbol{x}^{(2)}; \boldsymbol{\mu}^{(2)} + \boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{x}_0^{(1)}, \boldsymbol{\Sigma}_{22} - \boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{\Sigma}_{12})$$

すなわち、$\boldsymbol{x}^{(1)} = \boldsymbol{x}_0^{(1)}$ を与えたときの $\boldsymbol{x}^{(2)}$ の条件付き分布は

$$N_q(\boldsymbol{\mu}^{(2)} + \boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{x}_0^{(1)}, \boldsymbol{\Sigma}_{22} - \boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{\Sigma}_{12})$$

である。

3．付録 ヤコビアン (Jacobi 行列式) について

いま、m 個の変数 $\boldsymbol{x}' = (x_1, x_2, \cdots, x_m)$ から n 個の変数 $\boldsymbol{y}' = (y_1, y_2, \cdots, y_n)$ への変換 $x_1 = f_1(\boldsymbol{y}), x_2 = f_2(\boldsymbol{y}), \cdots, x_m = f_m(\boldsymbol{y})$ を考える。この変換である点 $(\boldsymbol{x}_0, \boldsymbol{y}_0)$ の近傍の振る舞いを線形近似したとき得られる$\boldsymbol{x}$を$\boldsymbol{y}$で偏微分した行列、

$$\begin{pmatrix} \dfrac{\partial f_1(\boldsymbol{y})}{\partial y_1} & \dfrac{\partial f_1(\boldsymbol{y})}{\partial y_2} & \cdots & \dfrac{\partial f_1(\boldsymbol{y})}{\partial y_n} \\ \dfrac{\partial f_2(\boldsymbol{y})}{\partial y_1} & \dfrac{\partial f_2(\boldsymbol{y})}{\partial y_2} & \cdots & \dfrac{\partial f_2(\boldsymbol{y})}{\partial y_n} \\ \vdots & \vdots & \ddots & \vdots \\ \dfrac{\partial f_m(\boldsymbol{y})}{\partial y_1} & \dfrac{\partial f_m(\boldsymbol{y})}{\partial y_2} & \cdots & \dfrac{\partial f_m(\boldsymbol{y})}{\partial y_n} \end{pmatrix}$$

を得る。これをヤコビ行列 (*Jacobi matrix*) という。

$m = n$ のときヤコビ行列は正方行列となり、その行列の行列式が正則の場合 (逆変換が存在する場合)、この行列式 (または行列式の絶対値) をヤコビアン (ヤコビ行列式、まれにヤコビ関数という) といい、多重積分の変数変換時に必要とされる。この値は変換時の面積または体積の変換比に相当する。

特に $f_i(\boldsymbol{y})$ が線形関数 (1 次式) の場合、

$$\begin{cases} x_1 &= a_{11}y_1 + a_{12}y_2 + \cdots + a_{1n}y_n \\ x_2 &= a_{21}y_1 + a_{22}y_2 + \cdots + a_{2n}y_n \\ & \vdots \\ x_n &= a_{n1}y_1 + a_{n2}y_2 + \cdots + a_{nn}y_n \end{cases}$$

このとき、ヤコビアン $J = \dfrac{d\boldsymbol{x}}{d\boldsymbol{y}} = \left|\dfrac{\partial \boldsymbol{x}'}{\partial \boldsymbol{y}'}\right|$ は線形変換に用いられた行列の行列式になり、

$$J = \left|\frac{\partial \boldsymbol{x}'}{\partial \boldsymbol{y}'}\right| = \left|\frac{\partial(x_1, x_2, \cdots, x_n)}{\partial(y_1, y_2, \cdots, y_n)}\right| = \begin{vmatrix} \dfrac{\partial x_1}{\partial y_1} & \dfrac{\partial x_1}{\partial y_2} & \cdots & \dfrac{\partial x_1}{\partial y_n} \\ \dfrac{\partial x_2}{\partial y_1} & \dfrac{\partial x_2}{\partial y_2} & \cdots & \dfrac{\partial x_2}{\partial y_n} \\ \vdots & \vdots & \ddots & \vdots \\ \dfrac{\partial x_n}{\partial y_1} & \dfrac{\partial x_n}{\partial y_2} & \cdots & \dfrac{\partial x_n}{\partial y_n} \end{vmatrix} = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}$$

従って、$\boldsymbol{x} = \mathbf{A}\boldsymbol{y}$ なる変換のヤコビアンは形式的に両辺を微分して、

$d\boldsymbol{x} = |\mathbf{A}|d\boldsymbol{y}$ であり、$\dfrac{d\boldsymbol{x}}{d\boldsymbol{y}} = |\mathbf{A}|$、$\dfrac{d\boldsymbol{y}}{d\boldsymbol{x}} = |\mathbf{A}|^{-1}$ となる。