3. 多重正規分布

概要: 正規分布は統計の中でもっとも頻繁に利用される重要な確率密度分布である。これは、1 変数に対してのものであるが、これを多変数まで拡張したものがどうなるか、をこの章では示す。まず説明変数が独立だと仮定するところからスタートする。これを変数変換して、変数間で相互作用がある場合の確率密度分布を得る。

キーワード: 正規分布; 多変数

3.1. 序

正規分布は統計で利用される最も基本的な確率密度分布である。これは、基本的に 1 変数に対してのものであった。しかし、一般には我々は多くの変数を扱う状況が多い。この場合、それらの変数が取る値を予想したい。多くの変数が独立であれば、それは各変数の取る確率をかけ合わせればいい。しかし、一般には変数間には相互作用がある。この場合の確率密度分布をこの章では議論する。

3.2. 1 変数の場合の復習

まず、正規分布の復習から始める。

母集団のデータは N 個からなると仮定する。

一変数の確率密度分布 $f(x)$ は

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^{(2)}}} \exp\left[-\frac{(x-\mu)^2}{2\sigma^{(2)}} \right] \tag{1}$$

ここで、μ は平均、 σ^2 は分散であり、

$$\mu = \frac{1}{N}\sum_{i=1}^{N} x_i \tag{2}$$

$$\sigma^{(2)} = \frac{1}{N}\sum_{i=1}^{N} (x_i - \mu)^2 \tag{3}$$

である。これらは、確率密度分からも以下のように導かれる。

$$\int_{-\infty}^{\infty} xf(x)\,dx = \int_{-\infty}^{\infty} x\frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{(x-\mu)^2}{2\sigma^2} \right] dx = \mu \tag{4}$$

$$\int_{-\infty}^{\infty} (x-\mu)^2 f(x)\,dx = \int_{-\infty}^{\infty} (x-\mu)^2 \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{(x-\mu)^2}{2\sigma^2} \right] dx = \sigma^2 \tag{5}$$

標準偏差σは分散のルートとして

$$\sigma = \sqrt{\sigma^{(2)}} \tag{6}$$

として与えられる。

ここで、変数変換

$$z = \frac{x-\mu}{\sigma} \tag{7}$$

を施し、zに対する確率密度分布を$g(z)$と置く。対応する変数領域に対するは両確率とも等しいとして、

$$f(x)\Delta x = g(z)\Delta z \tag{8}$$

を得る。Δx はxの増分であり、 Δz はそれに対応するzの増分である。

x が x_1 から x_2に変動し、 z はそれに呼応してz_1から z_2に変動するものとする。すると、対応する変換式は

$$z_1 = \frac{x_1}{\sigma} - \frac{\mu}{\sigma} \tag{9}$$

$$z_2 = \frac{x_2}{\sigma} - \frac{\mu}{\sigma} \tag{10}$$

したがって、

$$\Delta z = \frac{\Delta x}{\sigma} \tag{11}$$

となる。ここで、

$$\Delta x \equiv x_2 - x_1, \Delta z \equiv z_2 - z_1 \tag{12}$$

と置いてある。したがって、

$$\begin{aligned} f(x)\Delta x &= \frac{1}{\sqrt{2\pi\sigma^{(2)}}} \exp\left[-\frac{(x-\mu)^2}{2\sigma^{(2)}}\right]\Delta x \\ &= \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{z^2}{2}\right)\sigma\Delta z \\ &= \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{z^2}{2}\right)\Delta z \\ &= g(z)\Delta z \end{aligned} \tag{13}$$

となる。よって、規格化値に対する確率密度分布は

$$g(z) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{z^2}{2}\right) \tag{14}$$

となる。これを標準正規分布と呼ぶ。これは、z は標準正規分布 $N\left[0,1^2\right]$ に従うと表現する。

また、この変数の 2 乗を扱う。

$$u^2=\left(\frac{x-\mu}{\sigma}\right)^2 \tag{15}$$

は自由度 1 の χ^2 分布に従う。したがって、

$$u^2\le\chi^2\left(1,P\right) \tag{16}$$

として、この集団に属するかどうかを判断する。以後はこちらのほうを判断基準を採用する。

3.3.　2 変数の場合

次に二変数 X_1、X_2 の場合を扱う。この場合は、データはペアで $\left(x_1,x_2\right)$ の形式で存在する。それぞれの平均を μ_1、μ_2、分散を σ_1^2、σ_2^2 と置く。

この二変数 X_1、X_2 が独立の場合は、新たなデータセットの平方距離の和を求め

$$u_1^2+u_2^2=\frac{\left(x_1-\mu_1\right)^2}{\sigma_1^2}+\frac{\left(x_2-\mu_2\right)^2}{\sigma_2^2} \tag{17}$$

を評価すればいい。この場合は自由度 2 の χ^2 分布に従うから

$$u_1^2+u_2^2\le\chi^2\left(2,P\right) \tag{18}$$

の条件を満足するか判断すればいい。

しかし、一般には二変数 X_1、X_2 は独立ではなく、相関がある。この相関をどのように取り入れればいいであろうか？

まず、最初に独立な変数 z_1 と z_2 を考える。これらは、標準正規分布に従うとする。したがって、

$$\begin{aligned}
f\left(z_1,z_2\right)dz_1dz_2&=f\left(z_1\right)f\left(z_2\right)dz_1dz_2\\
&=\frac{1}{\sqrt{2\pi}}e^{-\frac{z_1{}^2}{2}}\frac{1}{\sqrt{2\pi}}e^{-\frac{z_2{}^2}{2}}dz_1dz_2\\
&=\frac{1}{2\pi}\exp\left[-\frac{1}{2}\left(z_1{}^2+z_2{}^2\right)\right]dz_1dz_2
\end{aligned} \tag{19}$$

となる。

次に、これらの独立な変数の線形結合を考える。それは以下で与えられるものとする。

$$\begin{cases}x_1=a_{11}z_1+a_{12}z_2+\mu_1\\ x_2=a_{21}z_1+a_{22}z_2+\mu_2\end{cases} \tag{20}$$

ただし、

$$a_{11}a_{22}-a_{12}a_{21}\ne0 \tag{21}$$

とする。これは行列表現で

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} + \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix} \tag{22}$$

で与えられる。これを最初の独立変数について解くと

$$\begin{aligned} \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} &= \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}^{-1} \begin{pmatrix} x_1 - \mu_1 \\ x_2 - \mu_2 \end{pmatrix} \\ &= D \begin{pmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{pmatrix} \begin{pmatrix} x_1 - \mu_1 \\ x_2 - \mu_2 \end{pmatrix} \end{aligned} \tag{23}$$

となる。こで、 D は変換行列

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \tag{24}$$

の行列式で

$$D \equiv \frac{1}{a_{11}a_{22} - a_{12}a_{21}} \tag{25}$$

で与えられる。したがって

$$\begin{cases} z_1 = a_{22}D(x_1 - \mu_1) - a_{12}D(x_2 - \mu_2) \\ z_2 = -a_{21}D(x_1 - \mu_1) + a_{11}D(x_2 - \mu_2) \end{cases} \tag{26}$$

となる。対応するヤコビアンは

$$\begin{aligned} |J| &= \begin{vmatrix} \dfrac{\partial z_1}{\partial x_1} & \dfrac{\partial z_1}{\partial x_2} \\ \dfrac{\partial z_2}{\partial x_1} & \dfrac{\partial z_2}{\partial x_2} \end{vmatrix} \\ &= \begin{vmatrix} a_{22}D & -a_{12}D \\ -a_{21}D & a_{11}D \end{vmatrix} \\ &= \left| a_{22}a_{11}D^2 - a_{12}a_{21}D^2 \right| \\ &= |D| \end{aligned} \tag{27}$$

で与えられるから、確率密度は

$$\begin{aligned} & f(z_1, z_2) dz_1 dz_2 \\ &= f(x_1, x_2) dx_1 dx_2 \\ &= \frac{|D|}{\left(\sqrt{2\pi}\right)^2} \exp\left\{ -\frac{\left[a_{22}(x_1 - \mu_1) - a_{12}(x_2 - \mu_2)\right]^2 + \left[-a_{21}(x_1 - \mu_1) + a_{11}(x_2 - \mu_2)\right]^2}{2} D^2 \right\} dx_1 dx_2 \end{aligned} \tag{28}$$

と変換される。ここで、 x_1 と x_2 の平均と分散、共分散は以下のように評価される。

$$E[x_1] = E[a_{11}z_1 + a_{12}z_2 + \mu_1] = a_{11}E[z_1] + a_{12}E[z_2] + \mu_1 = \mu_1 \tag{29}$$

$$E[x_2] = E[a_{21}z_1 + a_{22}z_2 + \mu_2] = a_{21}E[z_1] + a_{22}E[z_2] + \mu_2 = \mu_2 \tag{30}$$

$$
\begin{aligned}
E\left[\left(x_1-\mu_1\right)^2\right] &= E\left[\left(a_{11}z_1+a_{12}z_2+\mu_1-\mu_1\right)^2\right] \\
&= a_{11}{}^2+a_{12}{}^2 \equiv \sigma_{11}^{(2)}
\end{aligned}
\tag{31}
$$

$$
\begin{aligned}
E\left[\left(x_2-\mu_2\right)^2\right] &= E\left[\left(a_{21}z_1+a_{22}z_2+\mu_2-\mu_2\right)^2\right] \\
&= a_{21}{}^2+a_{22}{}^2 \equiv \sigma_{22}^{(2)}
\end{aligned}
\tag{32}
$$

$$
\begin{aligned}
Cov\left[x_1,x_2\right] &= E\left[\left(x_1-\mu_1\right)\left(x_2-\mu_2\right)\right] \\
&= E\left[\left(a_{11}z_1+a_{12}z_2+\mu_1-\mu_1\right)\left(a_{21}z_1+a_{22}z_2+\mu_2-\mu_2\right)\right] \\
&= a_{11}a_{21}E\left[z_1{}^2\right]+a_{12}a_{22}E\left[z_2{}^2\right]+\left(a_{11}a_{22}+a_{12}a_{21}\right)E\left[z_1z_2\right] \\
&= a_{11}a_{21}+a_{12}a_{22} \equiv \sigma_{12}^{(2)}
\end{aligned}
\tag{33}
$$

$$
\begin{aligned}
Cov\left[x_2,x_1\right] &= E\left[\left(x_2-\mu_2\right)\left(x_1-\mu_1\right)\right] \\
&= \sigma_{12}^{(2)} \\
&\equiv \sigma_{21}^{(2)}
\end{aligned}
\tag{34}
$$

相関係数ρは

$$
\begin{aligned}
\rho &= \frac{\sigma_{12}^{(2)}}{\sqrt{\sigma_{11}^{(2)}}\sqrt{\sigma_{22}^{(2)}}} \\
&= \frac{a_{11}a_{21}+a_{12}a_{22}}{\sqrt{a_{11}{}^2+a_{12}{}^2}\sqrt{a_{21}{}^2+a_{22}{}^2}}
\end{aligned}
\tag{35}
$$

で与えられる。したがって、

$$
\begin{aligned}
1-\rho^2 &= \frac{\sigma_{11}^{(2)}\sigma_{22}^{(2)}-\sigma_{12}^{(2)}\sigma_{21}^{(2)}}{\sigma_{11}^{(2)}\sigma_{22}^{(2)}} \\
&= \frac{\sigma_{11}^{(2)}\sigma_{22}^{(2)}-\left[\sigma_{12}^{(2)}\right]^2}{\sigma_{11}^{(2)}\sigma_{22}^{(2)}} \\
&= \frac{\left(a_{11}{}^2+a_{12}{}^2\right)\left(a_1{}^2+a_{22}{}^2\right)-\left(a_{11}a_{21}+a_{12}a_{22}\right)^2}{\sigma_{11}{}^{(2)}\sigma_{22}{}^{(2)}} \\
&= \frac{\left(a_{11}a_{22}-a_{12}a_{21}\right)^2}{\sigma_{11}^{(2)}\sigma_{22}^{(2)}}
\end{aligned}
\tag{36}
$$

を得る。これから行列式は

$$
|D| = \frac{1}{\left|a_{11}a_{22}-a_{12}a_{21}\right|} = \frac{1}{\sqrt{\left(1-\rho^2\right)\sigma_{11}^{(2)}\sigma_{22}^{(2)}}}
\tag{37}
$$

となる。

$f\left(x_1,x_2\right)$ の指数部は以下のように評価される。

$$
\begin{aligned}
&-\frac{\left[a_{22}\left(x_1-\mu_1\right)-a_{12}\left(x_2-\mu_2\right)\right]^2+\left[-a_{21}\left(x_1-\mu_1\right)+a_{11}\left(x_2-\mu_2\right)\right]^2}{2}D^2 \\
&=-\frac{1}{2}\left[\left(a_{22}{}^2+a_{21}{}^2\right)\left(x_1-\mu_1\right)^2-2\left(a_{12}a_{22}+a_{21}a_{11}\right)\left(x_1-\mu_1\right)\left(x_2-\mu_2\right)+\left(a_{12}{}^2+a_{11}{}^2\right)\left(x_2-\mu_2\right)^2\right]D^2 \\
&=-\frac{1}{2}\left[\sigma_{22}^{(2)}\left(x_1-\mu_1\right)^2-2\rho\sqrt{\sigma_{11}^{(2)}\sigma_{22}^{(2)}}\left(x_1-\mu_1\right)\left(x_2-\mu_2\right)+\sigma_{11}^{(2)}\left(x_2-\mu_2\right)^2\right]\frac{1}{\sigma_{11}^{(2)}\sigma_{22}^{(2)}\left(1-\rho^2\right)} \\
&=-\frac{1}{2\left(1-\rho^2\right)}\left[\sigma_{22}^{(2)}\frac{\left(x_1-\mu_1\right)^2}{\sigma_{11}^{(2)}}-2\rho\frac{\left(x_1-\mu_1\right)\left(x_2-\mu_2\right)}{\sqrt{\sigma_{11}^{(2)}\sigma_{22}^{(2)}}}+\frac{\left(x_2-\mu_2\right)^2}{\sigma_{22}^{(2)}}\right]
\end{aligned}
$$

(38)

したがって $f\left(x_1,x_2\right)$ は

$$
f\left(x_1,x_2\right)=\frac{1}{2\pi\sqrt{\left(1-\rho^2\right)\sigma_{11}^{(2)}\sigma_{22}^{(2)}}}\exp\left\{-\frac{1}{2\left(1-\rho^2\right)}\left[\frac{\left(x_1-\mu_1\right)^2}{\sigma_{11}^{(2)}}-2\rho\frac{\left(x_1-\mu_1\right)\left(x_2-\mu_2\right)}{\sqrt{\sigma_{11}^{(2)}\sigma_{22}^{(2)}}}+\frac{\left(x_2-\mu_2\right)^2}{\sigma_{22}^{(2)}}\right]\right\}
$$

(39)

となる。これは、$N\left[\mu_1,\sigma_{11}^{(2)}\right]$ および$N\left[\mu_2,\sigma_{22}^{(2)}\right]$の正規分布に従い、お互いの相関係数が$\rho$である二つの変数の確率密度である。

ここで、さらに変数変換

$$
\begin{cases}
u_1=\dfrac{x_1-\mu_1}{\sqrt{\sigma_{11}^{(2)}}} \\
u_2=\dfrac{x_2-\mu_2}{\sqrt{\sigma_{22}^{(2)}}}
\end{cases}
\tag{40}
$$

と置くと

$$
\begin{aligned}
&f\left(x_1,x_2\right)dx_1dx_2 \\
&=\frac{1}{2\pi\sqrt{\left(1-\rho^2\right)\sigma_{11}^{(2)}\sigma_{22}^{(2)}}}\exp\left[-\frac{1}{2\left(1-\rho^2\right)}\left(\frac{\left(x_1-\mu_1\right)^2}{\sigma_{11}^{(2)}}-2\rho\frac{\left(x_1-\mu_1\right)\left(x_2-\mu_2\right)}{\sqrt{\sigma_{11}^{(2)}\sigma_{22}^{(2)}}}+\frac{\left(x_2-\mu_2\right)^2}{\sigma_{22}^{(2)}}\right)\right]dx_1dx_2 \\
&=\frac{1}{2\pi\sqrt{\left(1-\rho^2\right)\sigma_{11}^{(2)}\sigma_{22}^{(2)}}}\exp\left[-\frac{1}{2\left(1-\rho^2\right)}\left(u_1^2-2\rho u_1u_2+u_2^2\right)\right]\sigma_1\sigma_2du_1du_2 \qquad (41) \\
&=\frac{1}{2\pi\sqrt{\left(1-\rho^2\right)}}\exp\left[-\frac{u_1^2-2\rho u_1u_2+u_2^2}{2\left(1-\rho^2\right)}\right]du_1du_2 \\
&=f\left(u_1,u_2\right)du_1du_2
\end{aligned}
$$

となる。

これから

$$
f\left(u_1,u_2\right)=\frac{1}{2\pi\sqrt{1-\rho^2}}\exp\left[-\frac{u_1^2-2\rho u_1u_2+u_2^2}{2\left(1-\rho^2\right)}\right] \qquad (42)
$$

を得る。

相関行列は

$$
R=\begin{pmatrix}1 & \rho\\ \rho & 1\end{pmatrix} \qquad (43)
$$

と置けるから、

$$
f\left(u_1,u_2\right)=\frac{1}{2\pi\sqrt{|R|}}\exp\left(-\frac{D^2}{2}\right) \qquad (44)
$$

となる。ただし、

$$
\begin{aligned}
D^2&=\left(u_1,u_2\right)\begin{pmatrix}1 & \rho\\ \rho & 1\end{pmatrix}^{-1}\begin{pmatrix}u_1\\ u_2\end{pmatrix} \\
&=\left(u_1,u_2\right)R^{-1}\begin{pmatrix}u_1\\ u_2\end{pmatrix}
\end{aligned} \qquad (45)
$$

である。

ここで、

$$
\begin{aligned}
D^2 &= \begin{pmatrix} u_1 & u_2 \end{pmatrix} R^{-1} \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} \\
&= \frac{u_1^2 - 2\rho u_1 u_2 + u_2^2}{1-\rho^2} \\
&= \frac{1}{1-\rho^2}\left[\frac{\left(x_1-\mu_1\right)^2}{\sigma_{11}^{(2)}} + \frac{\left(x_2-\mu_2\right)^2}{\sigma_{22}^{(2)}} - 2\rho \frac{\left(x_1-\mu_1\right)}{\sqrt{\sigma_{11}^{(2)}}} \frac{\left(x_2-\mu_2\right)}{\sqrt{\sigma_{22}^{(2)}}} \right] \\
&= \frac{1}{\left(1-\rho^2\right)\sigma_{11}^{(2)}\sigma_{22}^{(2)}} \begin{pmatrix} x_1-\mu_1 & x_2-\mu_2 \end{pmatrix} \begin{pmatrix} \sigma_{22}^{(2)} & -\rho\sqrt{\sigma_{11}^{(2)}\sigma_{22}^{(2)}} \\ -\rho\sqrt{\sigma_{11}^{(2)}\sigma_{22}^{(2)}} & \sigma_{11}^{(2)} \end{pmatrix} \begin{pmatrix} x_1-\mu_1 \\ x_2-\mu_2 \end{pmatrix} \\
&= \begin{pmatrix} x_1-\mu_1 & x_2-\mu_2 \end{pmatrix} \begin{pmatrix} \sigma_{11}^{(2)} & \sigma_{12}^{(2)} \\ \sigma_{12}^{(2)} & \sigma_{22}^{(2)} \end{pmatrix}^{-1} \begin{pmatrix} x_1-\mu_1 \\ x_2-\mu_2 \end{pmatrix} \\
&= \left(\boldsymbol{x}-\boldsymbol{\mu}\right)^T \Sigma^{-1} \left(\boldsymbol{x}-\boldsymbol{\mu}\right)
\end{aligned}
\tag{46}
$$

ただし

$$
\boldsymbol{x}-\boldsymbol{\mu} = \begin{pmatrix} x_1-\mu_1 \\ x_2-\mu_2 \end{pmatrix} \tag{47}
$$

$$
\Sigma = \begin{pmatrix} \sigma_{11}^{(2)} & \sigma_{12}^{(2)} \\ \sigma_{12}^{(2)} & \sigma_{22}^{(2)} \end{pmatrix} \tag{48}
$$

である。

さらに

$$
\begin{aligned}
\left|\Sigma\right| &= \begin{vmatrix} \sigma_{11}^{(2)} & \sigma_{12}^{(2)} \\ \sigma_{12}^{(2)} & \sigma_{22}^{(2)} \end{vmatrix} \\
&= \sigma_{11}^{(2)}\sigma_{22}^{(2)} - \sigma_{12}^{(2)} \\
&= \sigma_{11}^{(2)}\sigma_{22}^{(2)}\left(1 - \frac{\sigma_{12}^{(2)}}{\sigma_{11}^{(2)}\sigma_{22}^{(2)}}\right) \\
&= \sigma_{11}^{(2)}\sigma_{22}^{(2)}\left(1-\rho^2\right) \\
&= \sigma_{11}^{(2)}\sigma_{22}^{(2)}\left|R\right|^2
\end{aligned}
\tag{49}
$$

となる。

以上より、

$$
\begin{aligned}
f\left(u_1,u_2\right)du_1du_2 &= \frac{1}{2\pi\sqrt{\left|R\right|}}\exp\left(-\frac{D^2}{2}\right)du_1du_2 \\
&= \frac{1}{2\pi\sqrt{\left|\Sigma\right|}}\exp\left(-\frac{D^2}{2}\right)dx_1dx_2 \\
&= f\left(x_1,x_2\right)dx_1dx_2
\end{aligned}
\tag{50}
$$

となる。ただし、

$$D^2 = \begin{pmatrix} u_1 & u_2 \end{pmatrix} R^{-1} \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} = \begin{pmatrix} x_1 - \mu_1 & x_2 - \mu_2 \end{pmatrix} \Sigma^{-1} \begin{pmatrix} x_1 - \mu_1 \\ x_2 - \mu_2 \end{pmatrix} \tag{51}$$

である。つまり、D^2 は変数に変換に依存せず、同じ表式になる。

この場合は、そのデータがこの集合に属するかは

$$D^2 \le \chi^2(2, P) \tag{52}$$

を満足しているかで判断できる。

3.4. 多変量の場合

次に m 個の独立で正規化された多変数 $Z_1, Z_2, \cdots, Z_m$ の場合を扱う。この場合は、データはペアで $(z_1, z_2, \cdots, z_m)$ で存在する。それぞれの平均は 0 で、分散は 1 である。これに対する確率密度分布は、それぞれが独立であるから、

$$\begin{aligned} & f(z_1, z_2, \cdots, z_m) dz_1 dz_2 \cdots dz_m \\ &= f(z_1) f(z_2) \cdots f(z_m) dz_1 dz_2 \cdots dz_m \\ &= \frac{1}{\sqrt{2\pi}} e^{-\frac{z_1^2}{2}} \frac{1}{\sqrt{2\pi}} e^{-\frac{z_2^2}{2}} \cdots \frac{1}{\sqrt{2\pi}} e^{-\frac{z_m^2}{2}} dz_1 dz_2 \cdots dz_m \\ &= \frac{1}{\left(\sqrt{2\pi}\right)^m} \exp\left[-\frac{1}{2}\left(z_1^2 + z_2^2 + \cdots + z_m^2 \right) \right] dz_1 dz_2 \cdots dz_m \end{aligned} \tag{53}$$

となる。

ここで、変数変換

$$\mathbf{X} = A\mathbf{Z} + \boldsymbol{\mu} \tag{54}$$

を行う。ただし、

$$\mathbf{X} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{pmatrix} \tag{55}$$

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mm} \end{pmatrix} \tag{56}$$

$$\mathbf{Z}=\begin{pmatrix} z_1 \\ z_2 \\ \vdots \\ z_m \end{pmatrix} \tag{57}$$

$$\boldsymbol{\mu}=\begin{pmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_m \end{pmatrix} \tag{58}$$

である。

これから、

$$\begin{aligned} E[\mathbf{X}] &= E[A\mathbf{Z}+\boldsymbol{\mu}] \\ &= AE[\mathbf{Z}]+E[\boldsymbol{\mu}] \\ &= \boldsymbol{\mu} \end{aligned} \tag{59}$$

である。

さらに、

$$\begin{aligned} E\left[(x_1-\mu_1)^2\right] &= E\left[(a_{11}z_1+a_{12}z_2+\cdots+a_{1m}z_m+\mu_1-\mu_1)^2\right] \\ &= a_{11}{}^2E\left[z_1{}^2\right]+a_{12}{}^2E\left[z_2{}^2\right]+\cdots+a_{1m}{}^2E\left[z_m{}^2\right] \\ &\quad +2a_{11}a_{12}E[z_1z_2]+2a_{11}a_{13}E[z_1z_3]+\cdots+2a_{11}a_{1m}E[z_1z_m] \\ &\quad +2a_{12}a_{13}E[z_2z_3]+2a_{12}a_{14}E[z_2z_4]+\cdots+2a_{12}a_{1m}E[z_2z_m] \\ &\quad +\cdots \\ &\quad +2a_{1(m-1)}a_{1m}E\left[z_{(m-1)}z_m\right] \\ &= a_{11}{}^2+a_{12}{}^2+\cdots+a_{1m}{}^2 \end{aligned} \tag{60}$$

これから、

$$\sigma_{ii}^{(2)}=E\left[(x_i-\mu_i)^2\right]=a_{i1}{}^2+a_{i2}{}^2+\cdots+a_{im}{}^2 \tag{61}$$

$$\begin{aligned} &E[(x_1-\mu_1)(x_2-\mu_2)] \\ &= E[(a_{11}z_1+a_{12}z_2+\cdots+a_{1m}z_m+\mu_1-\mu_1)(a_{21}z_1+a_{22}z_2+\cdots+a_{2m}z_m+\mu_2-\mu_2)] \\ &= a_{11}a_{21}+a_{12}a_{22}+\cdots+a_{1m}a_{2m} \end{aligned} \tag{62}$$

となる。これから、

$$\begin{aligned} \sigma_{ij}^{(2)} &= E\left[(x_i-\mu_i)(x_j-\mu_j)\right] \\ &= a_{i1}a_{j1}+a_{i2}a_{j2}+\cdots+a_{im}a_{jm} \end{aligned} \tag{63}$$

となる。

一方、

$$AA^T = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mm} \end{pmatrix} \begin{pmatrix} a_{11} & a_{21} & \cdots & a_{m1} \\ a_{12} & a_{22} & \cdots & a_{m2} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1m} & a_{2m} & \cdots & a_{mm} \end{pmatrix} \tag{64}$$

より、この(i, j)成分は

$$AA^T(i, j) = \sum_{k=1}^{m} a_{ik} a_{jk} \tag{65}$$

である。したがって、AA^Tは変数$(x_1, x_2, \cdots, x_m)$の共分散行列となる。すなわち

$$AA^T(i, j) = \sigma_{ij}^{(2)} \tag{66}$$

である。そこで、

$$\Sigma = AA^T \tag{67}$$

と置く。ここで、

$$\Sigma = \begin{pmatrix} \sigma_{11}^{(2)} & \sigma_{12}^{(2)} & \cdots & \sigma_{1m}^{(2)} \\ \sigma_{21}^{(2)} & \sigma_{22}^{(2)} & \cdots & \sigma_{2m}^{(2)} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{m1}^{(2)} & \sigma_{m2}^{(2)} & \cdots & \sigma_{mm}^{(2)} \end{pmatrix} \tag{68}$$

である。

さらに

$$\mathbf{Z} = A^{-1}(\mathbf{X} - \boldsymbol{\mu}) \tag{69}$$

となる。これから

$$dz_1 dz_2 \cdots dz_m = \left| A^{-1} \right| dx_1 dx_2 \cdots dx_m \tag{70}$$

ここで、

$$\Sigma^{-1} = \left(AA^T \right)^{-1} \tag{71}$$

$$
\begin{aligned}
|\Sigma| &= \left|AA^T\right| \\
&= |A|\left|A^T\right| \\
&= |A|^2
\end{aligned}
\tag{72}
$$

$$
\left|A^{-1}\right| = \frac{1}{|A|} \tag{73}
$$

より、確率変数の微小領域には

$$
dz_1 dz_2 \cdots dz_m = \frac{1}{\sqrt{|\Sigma|}} dx_1 dx_2 \cdots dx_m \tag{74}
$$

の関係がある。したがって、確率密度の変換は

$$
\begin{aligned}
& f\left(z_1, z_2, \cdots, z_m\right) dz_1 dz_2 \cdots dz_m \\
&= \frac{1}{\left(\sqrt{2\pi}\right)^m} \exp\left[-\frac{1}{2}\left(z_1{}^2 + z_2{}^2 + \cdots + z_m{}^2\right)\right] dz_1 dz_2 \cdots dz_m \\
&= \frac{1}{\left(\sqrt{2\pi}\right)^m} \exp\left[-\frac{1}{2}\left(\mathbf{Z}^T \mathbf{Z}\right)\right] dz_1 dz_2 \cdots dz_m \\
&= \frac{1}{\left(\sqrt{2\pi}\right)^m} \exp\left[-\frac{1}{2}\left(\mathbf{Z}^T \mathbf{Z}\right)\right] \frac{1}{\sqrt{|\Sigma|}} dx_1 dx_2 \cdots dx_m \\
&= \frac{1}{\left(\sqrt{2\pi}\right)^m} \exp\left[-\frac{1}{2}\left(\left[A^{-1}\left(\mathbf{X}-\boldsymbol{\mu}\right)\right]^T A^{-1}\left(\mathbf{X}-\boldsymbol{\mu}\right)\right)\right] \frac{1}{\sqrt{|\Sigma|}} dx_1 dx_2 \cdots dx_m \\
&= \frac{1}{\left(\sqrt{2\pi}\right)^m \sqrt{|\Sigma|}} \exp\left[-\frac{1}{2}\left(\left(\mathbf{X}-\boldsymbol{\mu}\right)^T \left[A^{-1}\right]^T A^{-1}\left(\mathbf{X}-\boldsymbol{\mu}\right)\right)\right] dx_1 dx_2 \cdots dx_m \\
&= \frac{1}{\left(\sqrt{2\pi}\right)^m \sqrt{|\Sigma|}} \exp\left[-\frac{1}{2}\left(\left(\mathbf{X}-\boldsymbol{\mu}\right)^T \left[A^T\right]^{-1} A^{-1}\left(\mathbf{X}-\boldsymbol{\mu}\right)\right)\right] dx_1 dx_2 \cdots dx_m \\
&= \frac{1}{\left(\sqrt{2\pi}\right)^m \sqrt{|\Sigma|}} \exp\left[-\frac{1}{2}\left(\left(\mathbf{X}-\boldsymbol{\mu}\right)^T \left[AA^T\right]^{-1} \left(\mathbf{X}-\boldsymbol{\mu}\right)\right)\right] dx_1 dx_2 \cdots dx_m \\
&= \frac{1}{\left(\sqrt{2\pi}\right)^m \sqrt{|\Sigma|}} \exp\left[-\frac{1}{2}\left(\left(\mathbf{X}-\boldsymbol{\mu}\right)^T \Sigma^{-1} \left(\mathbf{X}-\boldsymbol{\mu}\right)\right)\right] dx_1 dx_2 \cdots dx_m
\end{aligned}
\tag{75}
$$

となる。

以上から、相関のある多変数の確率密度分布は

$$
f\left(x_1, x_2, \cdots, x_m\right) = \frac{1}{\left(\sqrt{2\pi}\right)^m \sqrt{|\Sigma|}} \exp\left[-\frac{1}{2}\left(\left(\mathbf{X}-\boldsymbol{\mu}\right)^T \Sigma^{-1} \left(\mathbf{X}-\boldsymbol{\mu}\right)\right)\right] \tag{76}
$$

ただし、

$$\mathbf{X} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{pmatrix} \tag{77}$$

$$\boldsymbol{\mu} = \begin{pmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_m \end{pmatrix} \tag{78}$$

$$\Sigma = \begin{pmatrix} \sigma_{11}^{(2)} & \sigma_{12}^{(2)} & \cdots & \sigma_{1m}^{(2)} \\ \sigma_{21}^{(2)} & \sigma_{22}^{(2)} & \cdots & \sigma_{2m}^{(2)} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{m1}^{(2)} & \sigma_{m2}^{(2)} & \cdots & \sigma_{mm}^{(2)} \end{pmatrix} \tag{79}$$

である。

3.5. 多変量のデータセットが集合のものか、そうでないか

多変量の場合の確率密度分布が得られた。したがって、新たなデータセット$\left(x_1, x_2, \cdots, x_m\right)$がこの集合内のものであるかどうか判断したい。我々の得ている確率密度分布は確率ではなく確率密度であることに留意しなければならない。したがって、何か対応する領域を設定なければならない。この場合、その領域を設定することは困難である。

しかし、指数の中身はもともと

$$z_1^{(2)} + z_2^{(2)} + \cdots + z_m^{(2)} \tag{80}$$

を変数変換したものである。この量自体は自由度mのχ^2分布に従う。

したがって、

$$\left(\mathbf{X} - \boldsymbol{\mu}\right)^T \Sigma^{-1} \left(\mathbf{X} - \boldsymbol{\mu}\right) \le \chi^2\left(m, P\right) \tag{81}$$

が満足されるかどうかをみればいい。もし、この関係を満足していれば、そのデータはこの集合に属しているとみなせ、満足していなければ、この集合に属していないと見做すことができる。

3.6. まとめ

この章のまとめをする。

相関のある多変数の確率密度分布は

$$f\left(x_1, x_2, \cdots, x_m\right) = \frac{1}{\left(\sqrt{2\pi}\right)^m \sqrt{|\Sigma|}} \exp\left[-\frac{1}{2}\left(\left(\mathbf{X}-\boldsymbol{\mu}\right)^T \Sigma^{-1} \left(\mathbf{X}-\boldsymbol{\mu}\right)\right)\right]$$

で与えられる。ただし、

$$\mathbf{X} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{pmatrix}$$

$$\boldsymbol{\mu} = \begin{pmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_m \end{pmatrix}$$

$$\Sigma = \begin{pmatrix} \sigma_{11}^{(2)} & \sigma_{12}^{(2)} & \cdots & \sigma_{1m}^{(2)} \\ \sigma_{21}^{(2)} & \sigma_{22}^{(2)} & \cdots & \sigma_{2m}^{(2)} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{m1}^{(2)} & \sigma_{m2}^{(2)} & \cdots & \sigma_{mm}^{(2)} \end{pmatrix}$$

である。