確率統計基礎講義 A

若木宏文

wakaki@hiroshima-u.ac.jp
http://home.hiroshima-u.ac.jp/ wakaki/lecture/index.shtml

2019.4.24

目 次

2.2 多次元正規分布

定義 1

確率密度関数

$$f(\boldsymbol{x};\boldsymbol{\mu},\Sigma)=\frac{1}{(2\pi)^{p/2}}|\Sigma|^{-1/2}\exp\left\{-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu})'\Sigma^{-1}(\boldsymbol{x}-\boldsymbol{\mu})\right\}$$
$$(\boldsymbol{\mu}\in\mathbb{R}^p,\Sigma>0)$$

によって定まる連続型分布を，p 次元正規分布といい $N_p(\boldsymbol{\mu},\Sigma)$ と表す.

注意 2.2.1 $f(\boldsymbol{x};\boldsymbol{\mu},\Sigma)$ は確率密度関数の条件を満たす.

Memo

Memo

標準化

$\boldsymbol{X} \sim N_p(\boldsymbol{\mu}, \Sigma)$ とする. $\Sigma = AA'$ とすると

$$\boldsymbol{Z} = A^{-1}(\boldsymbol{X} - \boldsymbol{\mu}) \sim N_p(\boldsymbol{0}, I_p)$$

が成り立つ. $N_p(\boldsymbol{0}, I_p)$ を p 次元標準という.

逆に, $\boldsymbol{Z} \sim N_p(\boldsymbol{0}, I_p)$ に対して

$$\boldsymbol{X} = \boldsymbol{\mu} + A\boldsymbol{Z} \sim N_p(\boldsymbol{\mu}, \Sigma) \ (\Sigma = AA')$$

が成り立つ.

注. 位置尺度分布族
一般に, $\boldsymbol{Z} \sim g(\boldsymbol{z}) \ (\boldsymbol{z} \in \mathbb{R}^p)$ (確率密度関数)とするとき,

$$\mathcal{G} = \left\{ f(\boldsymbol{z}; \boldsymbol{\mu}, A) = g(\boldsymbol{\mu} + A\boldsymbol{z})|A| \;\middle|\; \boldsymbol{\mu} \in \mathbb{R}^p, A : \text{正則} \right\}$$

を g によって生成される位置尺度分布族という.

Memo

定義 2

多次元正規分布の別定義

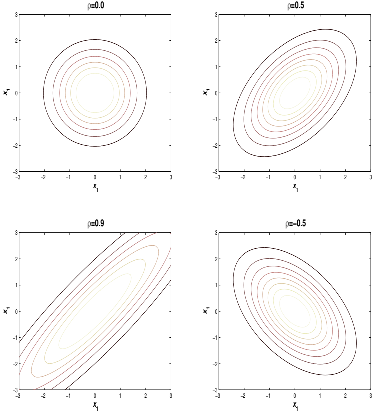
Σ を半正定値行列とする．
p 次元確率ベクトル $\boldsymbol{X} = (X_1, \ldots, X_p)'$ に対し，
$\boldsymbol{X} \sim N_p(\boldsymbol{\mu}, \Sigma)$ とは，$\mathrm{E}[\boldsymbol{X}] = \boldsymbol{\mu}$, $\mathrm{Var}[\boldsymbol{X}] = \Sigma$ で，かつ，
任意の $\boldsymbol{l} \in \mathbb{R}^p$ に対し，$\boldsymbol{l}'\boldsymbol{X}$ が 1 次元正規分布に従うときのことをいう．
定義 2 で Σ が正則でないとき，$N_p(\boldsymbol{\mu}, \Sigma)$ を**退化した正規分布**という．

注 2.2.2. この定義 2 は，定義 1 より広い分布族を定義する．

Memo

分布の特性

1. 多次元正規分布は $\boldsymbol{\mu}$ を中心に対称な分布であり，
   Σ はばらつきと変量ごとの関係の強さを表している．



$$\Sigma = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}$$

分布の特性（続き）

2. $X \sim N_p(\boldsymbol{\mu}, \Sigma)$ のとき, $\mathsf{E}[\boldsymbol{X}] = \boldsymbol{\mu}$, $\mathsf{Var}[\boldsymbol{X}] = \Sigma$.
3. $X \sim N_p(\boldsymbol{\mu}, \Sigma)$ の特性関数は
$$\mathcal{C}(\boldsymbol{t}) = \mathsf{E}[\exp(i\boldsymbol{t}'\boldsymbol{X})] = \exp\left(i\boldsymbol{t}'\boldsymbol{\mu} - \frac{1}{2}\boldsymbol{t}'\Sigma\boldsymbol{t}\right) \ (\boldsymbol{t} \in \mathbb{R}^p)$$
4. $\boldsymbol{b} \in \mathbb{R}^q$ $(q \leq p)$, B : $q \times p$ 行列 ($\mathrm{rank}(B) = q$) とすると
$$\boldsymbol{X} \sim N_p(\boldsymbol{\mu}, \Sigma) \ \Rightarrow \ \boldsymbol{Y} = \boldsymbol{b} + B\boldsymbol{X} \sim N_q(\boldsymbol{b} + B\boldsymbol{\mu}, B\Sigma B')$$
このことから，次が言える.

   $X_1, \ldots, X_p$ が独立で $X_j \sim N(\mu_j, \sigma^2)$ $(j = 1, \ldots, p)$ とし，p 次直交行列 $H = (h_{ij})$ を用いて
   $(Y_1, \ldots, Y_p)' = H(X_1, \ldots, X_p)$ と定めると，$Y_1, \ldots, Y_p$ は独立で $Y_j \sim N(\eta_j, \sigma^2)$ となる.
   ここで，$\eta_j = \sum_{k=1}^{p} h_{jk}\mu_k$.

Memo

Memo

Memo

分布の特性（続き）

5. 再生性：$\boldsymbol{X}_1 \sim N_p(\boldsymbol{\mu}_1, \Sigma_1), \boldsymbol{X}_2 \sim N_p(\boldsymbol{\mu}_2, \Sigma_2)$ で，かつ，$\boldsymbol{X}_1, \boldsymbol{X}_2$ が独立ならば $\boldsymbol{X}_1 + \boldsymbol{X}_2 \sim N_p(\boldsymbol{\mu}_1 + \boldsymbol{\mu}_2, \Sigma_1 + \Sigma_2)$
6. 条件付き分布：$\boldsymbol{X} \sim N_p(\boldsymbol{\mu}, \Sigma)$ のとき，$\boldsymbol{X}, \boldsymbol{\mu}, \Sigma$ を

$$\boldsymbol{X} = \begin{array}{c} p_1 \\ p_2 \end{array} \begin{pmatrix} \boldsymbol{X}_1 \\ \boldsymbol{X}_2 \end{pmatrix}, \quad \boldsymbol{\mu} = \begin{array}{c} p_1 \\ p_2 \end{array} \begin{pmatrix} \boldsymbol{\mu}_1 \\ \boldsymbol{\mu}_2 \end{pmatrix}, \quad \Sigma = \begin{array}{cc} & \begin{array}{cc} p_1 & p_2 \end{array} \\ \begin{array}{c} p_1 \\ p_2 \end{array} & \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{12}' & \Sigma_{22} \end{pmatrix} \end{array}$$

と分割する．このとき，$\boldsymbol{X}_2 = \boldsymbol{x}_2$ を与えたときの $\boldsymbol{X}_1$ の条件付き分布は

$$\boldsymbol{X}_1 | \boldsymbol{X}_2 = \boldsymbol{x}_2 \sim N_{p_1}(\boldsymbol{\mu}_1 + \Sigma_{12}\Sigma_{22}^{-1}(\boldsymbol{x}_2 - \boldsymbol{\mu}_2), \Sigma_{11.2}),$$
$$(\Sigma_{11.2} = \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{12}'),$$

となる．また，$\boldsymbol{X}_2 \sim N_{p_2}(\boldsymbol{\mu}_2, \Sigma_{22})$ となる．

レポート問題３

性質２を，性質３または性質４を用いて示せ.

Memo

Memo

Memo