Wishart分布とその性質

多変量正規分布の推定とBox-Cox変換

山北倫太郎

June 24, 2025

目次

7.1 はじめに

- $X = \begin{pmatrix} x_1' \\ \vdots \\ x_n' \end{pmatrix}$ は標本行列を表す 。
- $\bar{x}$ と S は、それぞれ μ と Σ の一貫性のある不偏推定量を提供する 。
  - $n\bar{x} = X'1$ (標本平均ベクトル($\bar{x}$))
  - ここで、1は n 次元のベクトルで、すべての要素が1である。
  - $(n-1)S = X'X - n\bar{x}\bar{x}'$ (標本共分散行列(S))
  - Sは母集団の共分散行列の不偏推定量であり、$X'X$ は標本行列の転置と自身の積である。

7.1 はじめに

- 7.2節では、$x_1, \ldots, x_n$ が独立同分布で $x \sim N_p(\mu, \Sigma)$ かつ $\Sigma > 0$ の場合の μ と Σ の最尤推定量が導出される 。
- $\bar{x}$ と S の同時分布に関する基本的な結果が命題7.1で証明される 。
- 7.3節では、Wishart分布の基本的な特性が研究される 。
- 7.4節では、データの多変量正規性を高めるためのBox-Cox変換が提示される 。

7.2 $\bar{x}$ と S の同時分布

- 正規性がある場合、$\bar{x}$ と S はいくつかの点で「最適」である 。
- $V = (n-1)S$ とする 。
- X の確率密度関数は様々な方法で記述できる 。

$$f(X) = (2\pi)^{-np/2}|\Sigma|^{-n/2}\exp\left[-\frac{1}{2}\sum_{i=1}^{n}(x_i-\mu)'\Sigma^{-1}(x_i-\mu)\right] \tag{7.1}$$

$$= (2\pi)^{-np/2}|\Sigma|^{-n/2}e^{-\frac{n}{2}\mu'\Sigma^{-1}\mu}\mathrm{etr}\left[-\frac{1}{2}\mathrm{tr}\ \Sigma^{-1}X'X + n\mu'\Sigma^{-1}\bar{x}\right]$$

$$= (2\pi)^{-np/2}|\Sigma|^{-n/2}\mathrm{etr}\left[-\frac{1}{2}[V + n(\bar{x}-\mu)(\bar{x}-\mu)']\Sigma^{-1}\right]$$

1行目

確率変数ベクトル $X = (x_1, \ldots, x_p)'$ が平均ベクトル μ と共分散行列 Σ を持つ多変量正規分布に従う場合、その確率密度関数 $f(x)$ は次のように表される。

p次元多変量正規分布の確率密度関数 (p.d.f.)

$$f(x) = (2\pi)^{-p/2}|\Sigma|^{-1/2} \exp\left(-\frac{1}{2}(x-\mu)'\Sigma^{-1}(x-\mu)\right)$$

- ここで、x は確率変数ベクトル X がとりうる値を示す p 次元の列ベクトルである。
- μ は、各確率変数の平均値を要素とする p 次元の平均ベクトルである。
- Σ は、$p \times p$ の共分散行列であり、各確率変数間の分散と共分散を表す対称な正定値行列である。

2行目

exp[tr()]=etr()の記法を用いると、確率密度関数は次のように表される 。

$$
\begin{aligned}
&(x_i - \mu)'\Sigma^{-1}(x_i - \mu) \\
&= (x_i' - \mu')\Sigma^{-1}(x_i - \mu) \\
&= x_i'\Sigma^{-1}x_i - x_i'\Sigma^{-1}\mu - \mu'\Sigma^{-1}x_i + \mu'\Sigma^{-1}\mu \\
&= x_i'\Sigma^{-1}x_i - 2x_i'\Sigma^{-1}\mu + \mu'\Sigma^{-1}\mu
\end{aligned}
$$

- ここで、$\mu'\Sigma^{-1}x_i$ はスカラー値であり、スカラーの転置は自分自身なので、$\mu'\Sigma^{-1}x_i = (x_i'\Sigma^{-1}\mu)'$ です。
  したがって、$x_i'\Sigma^{-1}\mu$ と $\mu'\Sigma^{-1}x_i$ は同じスカラー値を表します。

2行目

exp[tr()]=etr()の記法を用いると、確率密度関数は次のように表される 。

$$
\begin{aligned}
&(x_i - \mu)'\Sigma^{-1}(x_i - \mu) \\
&= (x_i' - \mu')\Sigma^{-1}(x_i - \mu) \\
&= x_i'\Sigma^{-1}x_i - x_i'\Sigma^{-1}\mu - \mu'\Sigma^{-1}x_i + \mu'\Sigma^{-1}\mu \\
&= x_i'\Sigma^{-1}x_i - 2x_i'\Sigma^{-1}\mu + \mu'\Sigma^{-1}\mu
\end{aligned}
$$

ここで、$\mu'\Sigma^{-1}x_i$ はスカラー値であり、スカラーの転置は自分自身なので、$\mu'\Sigma^{-1}x_i = (x_i'\Sigma^{-1}\mu)'$ です。
したがって、$x_i'\Sigma^{-1}\mu$ と $\mu'\Sigma^{-1}x_i$ は同じスカラー値を表します。

2行目

- μ'は平均ベクトルμの転置を表し、μはp次元の列ベクトルなので、μ'は1行p列の行ベクトルになります。
- Σ^{-1}は共分散行列Σの逆行列を表し、Σはp*pの正方行列なので、Σ^{-1}もp*pの正方行列になります。
- x_iはp次元の列ベクトルであり、x_i'はその転置で1行p列の行ベクトルになります。

よって、$(x_i - \mu)'\Sigma^{-1}(x_i - \mu)$はスカラー値であります。
また、$\mu'\Sigma^{-1}\mu$は、$e^{-\frac{1}{2}n\mu'\Sigma^{-1}\mu}$の形で指数関数に含まれます。

1項目

- 行列のトレースは、行列の対角成分の総和であり、$\mathrm{tr}(A) = \sum_i a_{ii}$ で定義される。

トレースの性質

- $\mathrm{tr}(AB) = \mathrm{tr}(BA)$
- $\mathrm{tr}(A) = \sum_{i=1}^n a_{ii}$, ここで $A = (a_{ij})$ は $n \times n$ 行列である。

$$
\begin{aligned}
\sum_{i=1}^n x_i' \Sigma^{-1} x_i &= \sum_{i=1}^n \mathrm{tr}\left(\Sigma^{-1} x_i x_i'\right) \\
&= \mathrm{tr}\left(\Sigma^{-1} \sum_{i=1}^n x_i x_i'\right) \\
&= \mathrm{tr}(\Sigma^{-1} X' X)
\end{aligned}
$$

1項目

- $X = \begin{pmatrix} x_1' \\ \vdots \\ x_n' \end{pmatrix}$ であるから、$X'X = \sum_{i=1}^{n} x_i x_i'$ となる。
- ここで、x_i' は x_i の転置を表し、$x_i x_i'$ は x_i の外積を表す。
- Σ^{-1} は共分散行列の逆行列であり、$x_i x_i'$ は x_i の外積を表す。

第2項目

$$\sum_{i=1}^{n} -2x_i'\Sigma^{-1}\mu = -2(\sum_{i=1}^{n} x_i')\Sigma^{-1}\mu$$
$$= -2\left(\sum_{i=1}^{n} x_i'\right)\Sigma^{-1}\mu$$

$\displaystyle\sum_{i=1}^{n} x_i' = (x_1' + \cdots + x_n') = (n\bar{x})'$であるので

$$= -2(n\bar{x})'\Sigma^{-1}\mu$$
$$= -2n\bar{x}'\Sigma^{-1}\mu$$

第3項目

$\sum_{i=1}^{n} \mu' \Sigma^{-1} \mu$はスカラー値であり、これが n 回足されます。

$$\sum_{i=1}^{n} \mu' \Sigma^{-1} \mu = n\mu' \Sigma^{-1} \mu$$

したがって、これら3つの項を合わせると、指数部分は次のようになります。

$$\begin{aligned}\sum_{i=1}^{n} (x_i - \mu)' \Sigma^{-1} (x_i - \mu) &= \sum_{i=1}^{n} x_i' \Sigma^{-1} x_i - 2n\bar{x}' \Sigma^{-1} \mu + n\mu' \Sigma^{-1} \mu \\ &= \mathrm{tr}(\Sigma^{-1} X' X) - 2n\bar{x}' \Sigma^{-1} \mu + n\mu' \Sigma^{-1} \mu\end{aligned}$$

まとめると

$$
\begin{aligned}
&-\frac{1}{2}\left[\mathrm{tr}(\Sigma^{-1}X'X) - 2n\bar{x}'\Sigma^{-1}\mu + n\mu'\Sigma^{-1}\mu\right] \\
&= -\frac{1}{2}\mathrm{tr}(\Sigma^{-1}X'X) + n\bar{x}'\Sigma^{-1}\mu - \frac{n}{2}\mu'\Sigma^{-1}\mu
\end{aligned}
$$

となり、2行目の指数部分と完全に一致します。

3行目

$$
\begin{aligned}
\sum_{i=1}^{n}(x_i-\mu)'\Sigma^{-1}(x_i-\mu) &= \sum_{i=1}^{n}\left((x_i-\bar{x})+(\bar{x}-\mu)\right)'\Sigma^{-1}\left((x_i-\bar{x})+(\bar{x}-\mu)\right) \\
&= \sum_{i=1}^{n}(x_i-\bar{x})'\Sigma^{-1}(x_i-\bar{x})+\sum_{i=1}^{n}(\bar{x}-\mu)'\Sigma^{-1}(\bar{x}-\mu) \\
&\quad + 2\sum_{i=1}^{n}(x_i-\bar{x})'\Sigma^{-1}(\bar{x}-\mu)
\end{aligned}
$$

3行目 (続き)

ここで、最後の項は

$$2\left(\sum_{i=1}^{n}(x_i-\bar{x})\right)'\Sigma^{-1}(\bar{x}-\mu)=2(n\bar{x}-n\bar{x})'\Sigma^{-1}(\bar{x}-\mu)=0$$

となるので、

$$\sum_{i=1}^{n}(x_i-\mu)'\Sigma^{-1}(x_i-\mu)=\sum_{i=1}^{n}(x_i-\bar{x})'\Sigma^{-1}(x_i-\bar{x})+n(\bar{x}-\mu)'\Sigma^{-1}(\bar{x}-\mu)$$

$$V=(n-1)S=\sum_{i=1}^{n}(x_i-\bar{x})(x_i-\bar{x})'$$

証明

証明

$$\begin{aligned}\sum_{i=1}^{n}(x_i-\bar{x})(x_i-\bar{x})' &= \sum_{i=1}^{n}\left[x_i x_i' - x_i\bar{x}' - \bar{x}x_i' + \bar{x}\bar{x}'\right] \\ &= \sum_{i=1}^{n} x_i x_i' - \sum_{i=1}^{n} x_i\bar{x}' - \sum_{i=1}^{n}\bar{x}x_i' + \sum_{i=1}^{n}\bar{x}\bar{x}' \\ &= X'X - n\bar{x}\bar{x}' - n\bar{x}\bar{x}' + n\bar{x}\bar{x}' \\ &= X'X - n\bar{x}\bar{x}' \\ &= (n-1)S\end{aligned}$$

したがって、$\sum_{i=1}^{n}(x_i-\bar{x})(x_i-\bar{x})' = (n-1)S = V$ である。

証明（各項の詳細な計算）

- $\sum_{i=1}^{n} \boldsymbol{x}_i \boldsymbol{x}_i'$: 標本行列 X は各行が $\boldsymbol{x}_i'$（$\boldsymbol{x}_i$ は列ベクトル）として定義されるので、$X = \begin{pmatrix} \boldsymbol{x}_1' \\ \boldsymbol{x}_2' \\ \vdots \\ \boldsymbol{x}_n' \end{pmatrix}$。その転置は $X' = (\boldsymbol{x}_1\ \boldsymbol{x}_2\ \dots\ \boldsymbol{x}_n)$ となる。したがって、$X'X = \sum_{i=1}^{n} \boldsymbol{x}_i \boldsymbol{x}_i'$。
- $\sum_{i=1}^{n} \boldsymbol{x}_i \overline{\boldsymbol{x}}'$: $\overline{\boldsymbol{x}}'$ は和のインデックス i に依存しないので、$\sum_{i=1}^{n} \boldsymbol{x}_i \overline{\boldsymbol{x}}' = (\sum_{i=1}^{n} \boldsymbol{x}_i)\overline{\boldsymbol{x}}' = n\overline{\boldsymbol{x}}\overline{\boldsymbol{x}}'$。
- $\sum_{i=1}^{n} \overline{\boldsymbol{x}} \boldsymbol{x}_i'$: $\overline{\boldsymbol{x}}$ も和のインデックス i に依存しないので、$\sum_{i=1}^{n} \overline{\boldsymbol{x}} \boldsymbol{x}_i' = \overline{\boldsymbol{x}}(\sum_{i=1}^{n} \boldsymbol{x}_i') = n\overline{\boldsymbol{x}}\overline{\boldsymbol{x}}'$。
- $\sum_{i=1}^{n} \overline{\boldsymbol{x}}\overline{\boldsymbol{x}}'$: これは n 回足し合わせるので、$n\overline{\boldsymbol{x}}\overline{\boldsymbol{x}}'$。

7.2 $\bar{x}$ と S の同時分布

- $(X'X, \bar{x})$ (または $(S, \bar{x})$ のような一対一関数) は、(Σ, μ) の最小十分完全統計量である 。
- Rao-Blackwell/Lehmann-Schefféの定理により、不偏推定量の中で $(\bar{x}, S)$ は最小分散を持つ 。
- $n-1 \geq p$ の場合の最尤推定量 (MLE) $\hat{\mu}$ と $\hat{\Sigma}$ を得るには、対数尤度関数を最小化する 。

$$\ln|\Sigma| + \mathrm{tr}\,\frac{1}{n}V\Sigma^{-1} + (\bar{x}-\mu)'\Sigma^{-1}(\bar{x}-\mu) \tag{7.2}$$

- 最尤推定量は $\hat{\mu} = \bar{x}$ と $\hat{\Sigma} = \frac{1}{n}V$ である 。

7.2 $\bar{x}$ と S の同時分布 (続き)

- 一般的な結果として、$\bar{x}$ は $N_p(\mu, \Sigma/n)$ に従う 。
- 標本行列 X を Z を用いて表す。$X \overset{d}{=} ZA' + 1\mu'$, ここで $Z \sim N_n^p(0, I_n \otimes I_p)$ かつ $\Sigma = AA'$ 。
- $\bar{x}$ と S_x の分布は $\bar{x}$ と S_z の分布に等しい 。
- $P = n^{-1}11'$ と $Q = I - n^{-1}11'$ は直交射影である 。
- $PZ \perp QZ$ であるため、$\bar{z} \perp S_z$ である 。
- $Q = HH'$ は直交基底を与えるため、$(n-1)S_z = Z'HH'Z = U'U$ となる 。

7.2 $\bar{x}$ と S の同時分布 (続き)

定義 7.1 Wishart分布

$W \sim W_p(m)$ ならば $W \stackrel{d}{=} \sum_{i=1}^m z_i z_i'$, ここで z_i は独立同分布で $N_p(0, I)$ に従う。
$V \sim W_p(m, \Sigma)$ ならば $V \stackrel{d}{=} AWA'$, ここで $\Sigma = AA'$ かつ $W \sim W_p(m)$。

命題 7.1

x_i が独立同分布で $N_p(\mu, \Sigma)$ に従う場合 $(i = 1, \ldots, n)$、

- $\bar{x} \sim N_p(\mu, \Sigma/n)$
- $(n-1)S \sim W_p(n-1, \Sigma)$
- $\bar{x} \perp S$

7.3 Wishart分布の性質 - 補題

補題 7.1

$Z = (z_{ij}) \in \mathbb{R}^{n\times n}$ が独立同分布の $N(0,1)$ に従う場合、$P(|Z|=0)=0$。

$n=1$ の場合は z_{11} が絶対連続分布を持つため、結果は成立する。
Z を以下のように分割する。

$$Z = \begin{pmatrix} z_{11} & z_{12}' \\ z_{21} & Z_{22} \end{pmatrix}$$

$Z_{22} \in \mathbb{R}^{(n-1)\times(n-1)}$ に対して結果が成立すると仮定すると、

$$\begin{aligned} P(|Z|=0) &= P(|Z|=0, |Z_{22}| \neq 0) + P(|Z|=0, |Z_{22}|=0) \\ &= P(z_{11} = z_{12}' Z_{22}^{-1} z_{21}, |Z_{22}| \neq 0) \\ &= E[P(z_{11} = z_{12}' Z_{22}^{-1} z_{21}, |Z_{22}| \neq 0 | z_{12}, z_{21}, Z_{22})] = 0 \end{aligned}$$

7.3 Wishart分布の性質 - 系

系 7.1

$Z = (z_{ij}) \in \mathbb{R}^{n \times n}$ が独立同分布の $N(0,1)$ に従う場合、$P(|Z| = t) = 0, \forall t$。

$P(|Z| = t) = E[P(z_{11} = z_{12}' Z_{22}^{-1} z_{21} + t/|Z_{22}|, |Z_{22}| \neq 0 | z_{12}, z_{21}, Z_{22})] = 0$。
補題7.1と系7.1は、Z が任意の絶対連続分布を持つ場合にも有効である。

命題 7.3

$W \sim W_p(m)$ かつ $m \geq p$ ならば、W は確率1で非特異である。

$W \stackrel{d}{=} Z'Z$ であり、$Z' = (z_1, \ldots, z_m)$ かつ z_i は独立同分布の $N_p(0, I)$ に従う。

$\text{rank } W \stackrel{d}{=} \text{rank } Z'Z = \text{rank } Z \geq \text{rank } (z_1, \ldots, z_p)$ は確率1で p となる。

したがって、rank W は確率1で p となる。