

筑波大学大学院理工学研究科

集中講義「数学特別講義 I」(15 時間分)  
(2005 年 1 月 25 日-2005 年 1 月 27 日)

## 講義録

今野 良彦

email: [konno@fc.jwu.ac.jp](mailto:konno@fc.jwu.ac.jp)

home page: <http://mp-w3math.jwu.ac.jp/~konno/stat-t.html>

日本女子大学理学部数物科学科

2005 年 1 月 24 日



# はじめに

$\text{Sym}^+(p, \mathbb{R})$  上のウィシャート分布族についての分布論および推測理論(この講義では, 分散行列の推定問題を統計的決定理論の立場から議論を扱う)は統計的多変量解析のなかでも最も重要な古典的な事項である. これらの古典的な性質は, かなり見通しよく対称錐体上のウィシャート分布族に拡張(等質錐体へもある程度可能)できる. 本講義の第一の目的は, 一般化されたウィシャート分布族への橋渡しという観点から, 古典的な事項を紹介することである. さらに, 時間がゆるせば, 対称錘(すなわち, 自己双対で等質な開凸錐体で正定値対称行列やエルミート対称複素行列をその例として含むもの)がジョルダン代数(すなわち, 可換律のほかにジョルダン積という特殊な結合律をみたす代数系)によって記述される事実を利用すれば, 正定値対称行列上のウィシャート分布とその基本的な性質が任意の対称錘上のウィシャート分布に見通しよく拡張・適応できることを解説する.

## 記号について

$\mathbb{R}$  : 実数

$\mathbb{R}^+$  : 正の実数

$\mathbb{Z}^+$  : 整数

$\mathbb{Z}^+$  : 正の整数

$\mathbb{C}$  : 複素数

$\mathbb{H}$  : 四元数

$\mathbb{R}^{r \times s}$ ,  $M(r, s, \mathbb{R})$  :  $r \times s$  の行列の集合

$M^\times(r, s, \mathbb{R})$  : 成分が実数の  $r \times s$  のフルランクの行列の集合

$GL(p, \mathbb{R})$  : 成分が実数の  $p \times p$  の正則行列の集合

$Sym(p, \mathbb{R})$  : 成分が実数の  $p \times p$  の対称行列の集合

$Sym^+(p, \mathbb{R})$  : 成分が実数の  $p \times p$  の正定値対称行列の集合

$\overline{Sym}^+(p, \mathbb{R})$  : 成分が実数の  $p \times p$  の半正定値対称行列の集合

$Alt(p, \mathbb{R})$  : 成分が実数の  $p \times p$  の歪対称の集合

$O(p, \mathbb{R})$  : 成分が実数の  $p \times p$  の直交行列の集合

$SO(p, \mathbb{R})$  :  $O(p, \mathbb{R})$  の単位元成分

$GT^+(p, \mathbb{R})$  : 対角成分が正の  $p \times p$  の下三角実行の集合

$Herm(p, \mathbb{C})$  : 成分が複素数の  $p \times p$  のエルミート行列の集合

$\overline{Herm}^+(p, \mathbb{C})$  : 成分が複素数の  $p \times p$  の正定値エルミート行列の集合

$Herm^+(p, \mathbb{C})$  : 成分が複素数の  $p \times p$  の半正定値エルミート行列の集合

$\mathcal{V}$  : 実ベクトル空間

$\mathcal{V}^*$  : 実ベクトル空間  $\mathcal{V}$  の双対空間

$\mathcal{L}(\mathcal{V}, \mathcal{W})$  : 実ベクトル空間  $\mathcal{V}$  から実ベクトル空間  $\mathcal{W}$  の線形写像の集合

$\overset{\circ}{C}$  : 集合  $C$  の内部

$\text{Tr}$  :  $\mathcal{L}(\mathcal{V}, \mathcal{V})$  のトレース

$\text{Det}$  :  $\mathcal{L}(\mathcal{V}, \mathcal{V})$  の行列式

# 目次

はじめに	i
<b>第 1 章</b>	<b>1</b>
1.1 ガンマ分布, ウィシャート行列とその分解	1
1.1.1 ガンマ分布	1
1.1.2 正値対称行列の分解	3
1.1.3 等質錐体と対称錐体	5
1.2 行列変換のヤコビアン	8
1.3 分布論	12
<b>第 2 章</b>	<b>15</b>
2.1 自然指数分布族	15
2.2 ウィシャート NEF 族	19
<b>第 3 章</b>	<b>25</b>
3.1 統計的決定問題	25
3.2 共分散行列の推定問題	28
3.2.1 定数リスクのミニマックス推定量	28
3.2.2 直交不変な推定量と SURE 法	30
3.2.3 部分積分の公式	31
3.2.4 直交不変ミニマックス推定量	33
3.2.5 Stein の推定量	34
<b>第 4 章</b>	<b>37</b>
4.1 ショルダン代数と対称錐体	37
4.2 対称錐体上のウィシャート分布	45
<b>付録 A</b>	<b>61</b>
A.1	61
A.2 基本的な事項	62
A.3 図とプログラム	63
<b>付録 B 不変積分について</b>	<b>65</b>
B.1 不変積分と不変測度	65
B.2 乗因子と相対不変積分	70

B.3 変換群と相対不変測度 . . . . .	73
付 録 C 等質錐体上のウイシャート分布	<b>79</b>
C.1 等質錐体上のウイシャート NEF . . . . .	79

# 第1章

## 1.1 ガンマ分布，ウイシャート行列とその分解

### 1.1.1 ガンマ分布

確率変数  $S$  が母数  $(\alpha, \beta)$  を持つガンマ分布とは， $S$  の確率密度関数

$$f_S(x|\alpha, \beta) = \frac{1}{\beta^\alpha \Gamma(\alpha)} x^{\alpha-1} e^{-x/\beta} \mathbb{I}_{(0, \infty)}(x)$$

をもつときをいう．ただし， $\alpha, \beta > 0$  とし，

$$\Gamma(\alpha) = \int_0^\infty x^{\alpha-1} e^{-x} dx, \quad \mathbb{I}_{(0, \infty)}(x) = \begin{cases} 1 & (x \in (0, \infty)), \\ 0 & (\text{その他}) \end{cases}$$

である．ガンマ分布に関してつぎが成立する：

(i) ガンマ分布の積率母関数は

$$M_S(t) = \mathbb{E}[e^{tS}] = \frac{1}{(1 - \beta t)^\alpha}, \quad (t \in (-\infty, 1/\beta))$$

となる．

(ii)  $\mathbb{E}[S] = \alpha\beta$ ,  $\text{VAR}[S] = \alpha\beta^2$  となる．

(iii) 1次元の確率変数  $X$  が正規分布  $N(0, \sigma^2)$  に従う場合を考える．ただし， $\sigma > 0$  とする．このとき， $X^2$  は母数  $(\alpha = 1/2, \beta = 2\sigma^2)$  のガンマ分布従う．

(iv)  $S_1$  と  $S_2$  は独立でそれぞれが母数  $(\alpha_1, \beta)$  と  $(\alpha_2, \beta)$  のガンマ分布に従うとき， $S_1 + S_2$  は母数  $(\alpha_1 + \alpha_2, \beta)$  のガンマ分布に従う．

(v)  $S$  が母数  $(\alpha, \beta)$  を持つガンマ分布に従うとき， $cS$  は母数  $(\alpha, c\beta)$  のガンマ分布に従う．ただし， $c > 0$  は定数である．

(vi) 指数分布族の一員である．

(vii) Lukacs-Olkin-Rubin の性質． $S_1, S_2$  は独立，非退化，かつ非負値確率変数で， $S_1 + S_2 > 0$  とし， $Z = S_1/(S_1 + S_2)$  とおく．このとき，

$$S_1 + S_2 \text{ と } Z \text{ は独立} \iff \text{ある } \alpha_1, \alpha_2, \beta > 0 \text{ が存在して，} S_1 \text{ と } S_2 \text{ は母数 } (\alpha_1, \beta) \text{ と } (\alpha_2, \beta) \text{ のガンマ分布に従う．}$$

(viii) Matsumoto-Yor の性質  $p, a, b > 0$  とする.  $X$  の分布を

$$\mu_{-p,a,b}(dx) = Cx^{-(p+1)}e^{-\frac{1}{2}(ax+\frac{b}{x})} \mathbb{I}_{(0,\infty)}(x)dx$$

とし,  $Y$  の分布を

$$\gamma_{p,2/a}(dy) = C'y^{p-1}e^{-\frac{1}{2}ay} \mathbb{I}_{(0,\infty)}(y)dy$$

とする. ただし,  $dx, dy$  は  $\mathbb{R}$  上のルベーグ測度とし,  $C, C'$  は基準化定数である. いま,

$$\begin{cases} U = \frac{1}{X+Y} \\ V = \frac{1}{X} - \frac{1}{X+Y} \end{cases} \iff \begin{cases} X = \frac{1}{U+V} \\ Y = \frac{V}{U(U+V)} \end{cases}$$

とおいたとき,  $U$  と  $V$  は独立に  $\mu_{-p,b,a}$  と  $\gamma_{p,b/2}$  に従う.

特に, (iii) と (iv) の性質から  $X_1, X_2, \dots, X_n$  が独立同一に正規分布  $N(0, \sigma^2)$  に従うとき,

$$S = \sum_{i=1}^n X_i^2$$

は母数  $(n/2, 2\sigma^2)$  のガンマ分布(すなわち,  $S/\sigma^2$  は自由度  $n$  の  $\chi^2$  分布)に従うことがわかる.

**正規分布の分散の推定**  $\tilde{X}_1, \dots, \tilde{X}_n, \tilde{X}_{n+1}$  が独立同一に正規分布  $N(\mu, \sigma^2)$  に従っていると  
する. ここで,  $\mu (-\infty < \mu < \infty)$  と  $\sigma^2 (0 < \sigma < \infty)$  は未知とする. このとき, 推定量

$$\tilde{S} = \sum_{i=1}^{n+1} (\tilde{X}_i - \bar{\tilde{X}}_{n+1})^2, \quad \bar{\tilde{X}}_{n+1} = \frac{1}{n+1} \sum_{i=1}^{n+1} \tilde{X}_i$$

とおくと,  $\tilde{S}/n$  は  $\sigma^2$  の不偏推定量になることが知られている.  $\tilde{S}$  のみに基づく  $\sigma^2$  の推定を考えた場合,  $\tilde{S}$  は  $\sigma^2 \chi_n^2$  に従うことがわかる. 十分統計量は  $(\tilde{S}, \bar{\tilde{X}}_{n+1})$  であるので, Stein (1964) のように  $\bar{\tilde{X}}_{n+1}$  の情報を  $\sigma^2$  の推定に活用することは意味がある. しかし, 本講義では議論を絞り,  $\tilde{S}$  のみに基づく  $\sigma^2$  の推定問題に限定して議論を進めることにする. 従って,  $\mu = 0$  の設定とする.

$X_1, X_2, \dots, X_n$  が独立同一に正規分布  $N(0, \sigma^2)$  に従っている設定に戻る. このとき,

$$\frac{S}{n} \tag{1.1}$$

は  $\sigma^2$  の推定量として最適性を持つことが期待できる. しかし, 平均ベクトルが 0 の多変量正規分布の共分散行列の推定問題を考えた場合には, 推定量 (1.1) を単純に多次元にしたものは必ずしも 1 次元のときに持っている最適性を保つとは限らない. 分散の推定を考えたときは, 母数空間は通常の積に関して可換であるが, 母数空間が正定値行列の空間となったときは, この空間は行列の積に関して非可換になる.

**複素数値確率変数と複素正規分布**  $U, V$  は実数値確率変数としたとき,  $X = U + iV$  を複素数値確率変数という.  $X, Y$  は複素数値確率変数とし,  $\mathbb{E}[X\bar{X}] = \mathbb{E}[U^2 + V^2] < \infty, \mathbb{E}[Y\bar{Y}] < \infty$  のとき,

$$\begin{aligned} \mathbb{E}(X) &= \mathbb{E}(U) + i\mathbb{E}(V), \\ \text{COV}(X, Y) &= \mathbb{E}((X - \mathbb{E}(X))\overline{(Y - \mathbb{E}(Y))}) \end{aligned}$$

とおく．複素数値確率変数  $X = U + iV$  は，平均 0，分散  $\sigma^2$  ( $\sigma > 0$ ) の複素数値正規分布  $CN(0, \sigma^2)$  に従うとは， $(U, V)'$  が 2 次元正規分布  $N(0, (\sigma^2/2)I_2)$  に従うときにいう．すなわち， $X$  は確率測度

$$P(dx) = \frac{1}{\sigma\pi} \exp(-\bar{x}x/\sigma^2) dx$$

をもつ．ただし， $dx$  は  $\mathbb{R}^2$  上のルベグ測度である． $S = \bar{X}X$  とおいたとき， $S$  は母数  $(1, \sigma^2)$  のガンマ分布 ( $2S/\sigma^2$  は自由度 2 の  $\chi^2$  分布) に従うことがわかる．

### 1.1.2 正值対称行列の分解

1 次元の正規分布に従う確率変数の 2 乗和 (1.1) と同じように考えて，多変量正規分布の 2 乗和を考える．ただし，スカラーにすると面白くないので，行列になるようにする．

$X_1, X_2, \dots, X_n$  を  $\mathbb{R}^p$ -値確率ベクトルとし，独立同一に多変量正規分布  $N_p(0, \Sigma)$  に従うとする．ただし，

$$\Sigma \in \text{Sym}^+(p, \mathbb{R}) = \{S \in \mathbb{R}^{p \times p} : x'Sx > 0 \text{ for } x \in \mathbb{R}^p, x \neq 0\}$$

とし， $x'$  は縦ベクトル  $x$  の転置を表すとする． $X_1$  は確率測度

$$P(dx) = \frac{1}{(2\pi)^{p/2} \text{Det}(\Sigma)^{1/2}} \exp\left\{-\frac{1}{2}x'\Sigma^{-1}x\right\} (dx)$$

をもつ．ただし， $dx$  は  $\mathbb{R}^p$  上のルベグ測度である．

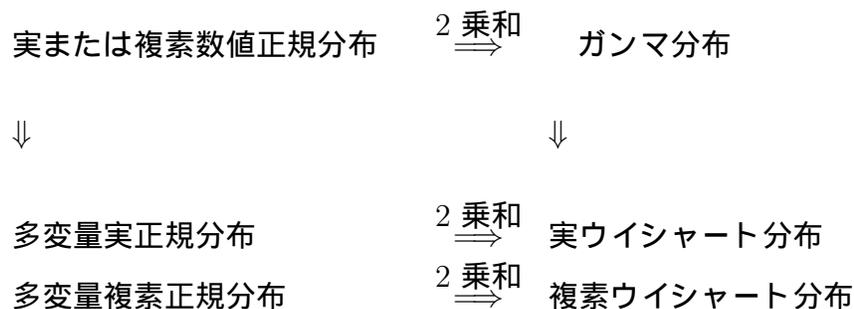
いま，

$$S : p \times p = \sum_{j=1}^n X_j X_j'$$

とおく． $S$  のことをウイシャート行列と呼ぶことにする．定義より， $S \in \overline{\text{Sym}^+(p, \mathbb{R})}$  であるが，

$$\mathbb{P}\{S \in \text{Sym}^+(p, \mathbb{R})\} = 1 \iff n \geq p \quad \text{かつ} \quad \Sigma \in \text{Sym}^+(p, \mathbb{R})$$

となる．Eaton (1983, page 304) を参照．以後は，簡単のために， $S \in \text{Sym}^+(p, \mathbb{R})$  等と書けば，確率 1 で成立することを意味することにする．



ウィシャート行列の分布はガンマ分布の多次元の拡張と考えることができる。したがって、ガンマ分布の性質と同じようなことが成立することが期待される。これについては後ほどみることにする。さらに、行列を考えているので、1 変数の場合にはないことも考えることができる。それはウィシャート行列の分解である。

ウィシャート行列  $S$  のよく知られている分解である。

**極分解** 始めに、分解

$$S = H L H', \quad H \in O(p, \mathbb{R}), \quad L = \text{diag}(\ell_1, \ell_2, \dots, \ell_p) \quad \text{with } \ell_1 \geq \ell_2 \geq \dots \geq \ell_p > 0$$

を考える。ただし、

$$O(p, \mathbb{R}) = \{H \in \mathbb{R}^{p \times p} : H'H = HH' = I_p\}$$

で  $I_p$  は  $p \times p$  の単位行列とする。

**ガウス分解** 次に、分解

$$S = T T', \quad T \in \text{GT}^+(p, \mathbb{R})$$

を考える。 $G := \text{GT}^+(p, \mathbb{R}) \subset \text{GL}(p, \mathbb{R})$  を対角成分が正の下三角行列の集合とする。 $x \in \text{GT}^+(p, \mathbb{R})$  のとき、

$$x = \begin{pmatrix} x_{11} & 0 & \cdots & 0 \\ x_{21} & x_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ x_{p1} & x_{p2} & \cdots & x_{pp} \end{pmatrix}$$

で  $x_{ii} > 0$ ,  $x_{ij} \in \mathbb{R}$  ( $i = 1, 2, \dots, p$ ,  $j = 1, 2, \dots, i$ ) である。

**Partial Iwasawa 分解** 最後に、分解

$$S = \begin{pmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{pmatrix}, \quad S_{11} \in \text{Sym}^+(p_1, \mathbb{R}), \quad S_{22} \in \text{Sym}^+(p_2, \mathbb{R})$$

を考える。ただし、 $p = p_1 + p_2$  で  $p_1, p_2$  は正の整数である。このとき、

$$\begin{aligned} S &= \begin{pmatrix} I_{p_1} & 0 \\ S_{21} S_{11}^{-1} & I_{p_2} \end{pmatrix} \begin{pmatrix} S_{11} & 0 \\ 0 & S_{22 \bullet 1} \end{pmatrix} \begin{pmatrix} I_{p_1} & S_{11}^{-1} S_{12} \\ 0 & I_{p_2} \end{pmatrix} \\ &= \begin{pmatrix} I_{p_1} & S_{12} S_{22}^{-1} \\ 0 & I_{p_2} \end{pmatrix} \begin{pmatrix} S_{11 \bullet 2} & 0 \\ 0 & S_{22} \end{pmatrix} \begin{pmatrix} I_{p_1} & 0 \\ S_{22}^{-1} S_{21} & I_{p_2} \end{pmatrix} \end{aligned} \quad (1.2)$$

となる。ただし、

$$S_{11 \bullet 2} = S_{11} - S_{12} S_{22}^{-1} S_{21} \in \text{Sym}^+(p_1, \mathbb{R}),$$

$$S_{22 \bullet 1} = S_{22} - S_{21} S_{11}^{-1} S_{12} \in \text{Sym}^+(p_2, \mathbb{R})$$

である。

**行列の群**  $p$  次の正方行列  $A = (a_{ij}), B = (b_{ij})$  の演算として, 行列の和と積を

$$A + B = (a_{ij} + b_{ij}), \quad AB = \left( \sum_{k=1}^p a_{ik} b_{kj} \right)$$

を考える. 一般に,  $AB = BA$  は成立しないが, 結合律, 分配律などは成立する:

$$\begin{aligned} (AB)C &= A(BC), \\ (A+B)C &= AC + BC. \end{aligned}$$

ただし,  $A, B, C$  は  $p$  次の正方行列. さらに,

$$AI_p = I_p A = A$$

が成立し,  $A \in \text{GL}(p, \mathbb{R})$  に対して,

$$AB = BA = I_p$$

なる  $B \in \text{GL}(p, \mathbb{R})$  が存在する. これを  $A^{-1}$  と書く. したがって,  $\text{GL}(p, \mathbb{R})$  は行列の積に関して群となる.

$\text{GL}(p, \mathbb{R})$  のつぎの部分集合は部分群となる:

$$\begin{aligned} \text{O}(p, \mathbb{R}) &= \{H \in \mathbb{R}^{p \times p} : H'H = HH' = I_p\}, \\ \text{GT}^+(p, \mathbb{R}) &= \left\{ x = \begin{pmatrix} x_{11} & 0 & \cdots & 0 \\ x_{21} & x_{22} & \cdots & 0 \\ \vdots & \vdots & \vdots & \\ x_{p1} & x_{p2} & \cdots & x_{pp} \end{pmatrix} : x_{ii} > 0, x_{ij} \in \mathbb{R} (i > j) \right\}, \\ \mathcal{D}(p) &= \{\text{diag}(\ell_1, \dots, \ell_p) : \ell_1 \geq \cdots \geq \ell_p > 0\}, \\ \text{GT}^+(p_1, p_2, \mathbb{R}) &= \left\{ \begin{pmatrix} I_{p_1} & 0 \\ A & I_{p_2} \end{pmatrix} : A \in \mathbb{R}^{p_1 \times p_2} \right\} \end{aligned}$$

となる.

### 1.1.3 等質錐体と対称錐体

$\text{GL}(p, \mathbb{R})$  の  $\text{Sym}^+(p, \mathbb{R})$  への作用 (変換) を

$$\text{GL}(p, \mathbb{R}) \times \text{Sym}^+(p, \mathbb{R}) \quad (g, x) \mapsto gxg'$$

で定義する. このとき, 作用は推移的である: すなわち, 任意の  $x, y \in \text{Sym}^+(p, \mathbb{R})$  に対して, ある  $g \in \text{GL}(p, \mathbb{R})$  が存在して,  $x = gy$  とできる. すなわち,  $\text{Sym}^+(p, \mathbb{R})$  は等質な空間となる.

さらに,  $\text{Sym}^+(p, \mathbb{R})$  は凸錐体となる: すなわち,

$$x_1, x_2 \in \text{Sym}^+(p, \mathbb{R}), \lambda_1, \lambda_2 \in \mathbb{R} \implies \lambda_1 x_1 + \lambda_2 x_2 \in \text{Sym}^+(p, \mathbb{R})$$

をみたく、これに加えて、

$$\begin{aligned} \overline{\text{Sym}^+(p, \mathbb{R})} \cap \overline{(-\text{Sym}^+(p, \mathbb{R}))} &= \{0\}, \\ \text{Sym}^+(p, \mathbb{R}) \cup (-\text{Sym}^+(p, \mathbb{R})) &= \text{Sym}(p, \mathbb{R}) \end{aligned}$$

となる。この条件が成り立つとき、固有な (proper) 凸錐体という。

**注意**  $\mathcal{V}$  を  $\mathbb{R}$  上の有限次元ベクトル空間とする。 $\mathcal{V}$  の部分集合  $C$  が、錐体であるとは、

$$x \in C, \lambda > 0 \implies \lambda x \in C$$

をみたくときをいう。 $\mathcal{V}$  の部分集合  $S$  が凸であるとは、

$$x_1, x_2 \in S, 0 < \lambda < 1 \implies \lambda x_1 + (1 - \lambda)x_2 \in S$$

をみたくときをいう。 $C \subset \mathcal{V}$  が凸錐体であるとは、

$$x_1, x_2 \in C, \lambda_1, \lambda_2 \in \mathbb{R} \implies \lambda_1 x_1 + \lambda_2 x_2 \in C$$

をみたくときをいう。□

任意の元  $x, y \in \text{Sym}(p, \mathbb{R})$  に対して、 $x, y$  の内積  $(x|y)$  を

$$(x|y) = \text{Tr}(xy)$$

で定める。 $\text{Sym}(p, \mathbb{R})$  は  $p(p+1)/2$  次元 Euclid 空間の開集合である。開凸錐体  $\text{Sym}^+(p, \mathbb{R})$  の開双対錐体  $\{\text{Sym}^+(p, \mathbb{R})\}^*$  を

$$\{\text{Sym}^+(p, \mathbb{R})\}^* = \{x \in \text{Sym}(p, \mathbb{R}) : (x|y) > 0 \text{ for } y \in \overline{\text{Sym}^+(p, \mathbb{R})} \setminus \{0\}\}$$

で定めたとき、

$$\{\text{Sym}^+(p, \mathbb{R})\}^* = \text{Sym}^+(p, \mathbb{R})$$

となる。この性質を自己双対という。ここで、

$$\text{等質錐体} + \text{自己双対凸錐体} = \text{対称錐体},$$

という。また、

$$\text{等質錐体} \supset \text{対称錐体} \supset \text{Sym}^+(p, \mathbb{R})$$

となる。

**多変量複素正規分布と複素構造を持つ共分散行列をもつ多変量実正規分布** 複素ベクトル  $c = a + ib$  ( $a, b \in \mathbb{R}^p$ ) と複素行列  $C = A + iB$  ( $A, B \in \mathbb{R}^{n \times p}$ ) に対して、

$$[c] = \begin{pmatrix} a \\ b \end{pmatrix}, \quad \{C\} = \begin{pmatrix} A & -B \\ B & A \end{pmatrix}$$

と定める．これについての記述を追加する． $p$ -次元複素ベクトル値確率変数  $X = U + iV$  の共分散行列は

$$\text{COV}(X) = \mathbb{E}((X - \mathbb{E}(X))(X - \mathbb{E}(X))^*) =: \Sigma + iA \in \text{Herm}(p, \mathbb{C})$$

で定める．ただし， $\Sigma \in \text{Sym}(p, \mathbb{R})$ ,  $A \in \text{Alt}(p, \mathbb{R})$  で，立ベクトル  $x$  に対して， $x^* = \bar{x}'$  は随伴ベクトルである．また，

$$\text{COV}([X]) = \begin{pmatrix} \Sigma & -A \\ A & \Sigma \end{pmatrix}$$

となる．このような共分散行列の構造を複素構造という． $p$  次元複素正規分布は，複素構造を持つ共分散行列をもつ  $2p$  次元実正規分布と考えることができる．

## 1.2 行列変換のヤコビアン

命題 1.1  $X, Y$  を functionally independent な  $\mathbb{R}^p$ -値確率ベクトルとし,  $A \in \text{GL}(p, \mathbb{R})$  を定数行列とする. このとき,  $Y = AX$  のヤコビアンは, 符号を除いて,

$$dY = \text{Det}(A) dX$$

となる.

証明:  $X = (x_i), Y = (y_i)$  とする.

$$\frac{\partial Y}{\partial X} = \left( \frac{\partial y_i}{\partial x_j} \right) = \begin{pmatrix} \frac{\partial y_1}{\partial x_1} & \cdots & \frac{\partial y_1}{\partial x_p} \\ \vdots & & \vdots \\ \frac{\partial y_p}{\partial x_1} & \cdots & \frac{\partial y_p}{\partial x_p} \end{pmatrix}$$

からわかる. □

形式的な計算 簡単のために  $p = 2$  とする. 上の定理より

$$dy_1 dy_2 = \text{Det} \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} dx_1 dx_2 \quad (1.3)$$

となる. ただし,  $A = (a_{ij})$  とした.  $y_1$  と  $y_2$  を入れ替えると

$$dy_2 dy_1 = \text{Det} \begin{pmatrix} a_{21} & a_{22} \\ a_{11} & a_{12} \end{pmatrix} dx_1 dx_2 = -\text{Det} \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} dx_1 dx_2$$

となる. これより

$$dy_1 dy_2 = -dy_1 dy_2$$

となる. さらに,  $dy_1(dy_1) = -(dy_1)dy_1$  より  $dy_1 dy_1 = 0$  となる.

外微分形式を用いて,

$$\begin{aligned} dy_1 &= \frac{\partial y_1}{\partial x_1} dx_1 + \frac{\partial y_1}{\partial x_2} dx_2, \\ dy_2 &= \frac{\partial y_2}{\partial x_1} dx_1 + \frac{\partial y_2}{\partial x_2} dx_2, \end{aligned}$$

としたとき,

$$\begin{aligned} dy_1 dy_2 &= \left( \frac{\partial y_1}{\partial x_1} dx_1 + \frac{\partial y_1}{\partial x_2} dx_2 \right) \left( \frac{\partial y_2}{\partial x_1} dx_1 + \frac{\partial y_2}{\partial x_2} dx_2 \right) \\ &= \left( \frac{\partial y_1}{\partial x_1} \right) \left( \frac{\partial y_2}{\partial x_2} \right) dx_1 dx_2 - \left( \frac{\partial y_1}{\partial x_2} \right) \left( \frac{\partial y_2}{\partial x_1} \right) dx_1 dx_2 \\ &= \text{Det} \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} dx_1 dx_2 \end{aligned}$$

となる。(1.3) の積は交代で、くさび積 (wedge product) もしくは外積 (exterior product) と呼ばれる。これを  $\bigwedge_{j=1}^p dx_j$  と書く。したがって、

$$dx_i \wedge dx_j = \begin{cases} -dx_j \wedge dx_i & (i \neq j), \\ 0 & (i = j) \end{cases}$$

をみたすことになる。

**記法**  $X = (x_{ij}) \in \mathbb{R}^{p \times q}$  としたとき、 $dX$  を積分やヤコビアン (符号が残っている) で用いたとき

$$dX = \bigwedge_{i=1}^p \bigwedge_{j=1}^q dx_{ij}$$

と理解し、 $dX$  を確率密度関数に関する積分 (ヤコビアンの絶対値を取る) で用いたとき、

$$dX = \prod_{i=1}^p \prod_{j=1}^q dx_{ij}$$

と理解する。

**命題 1.2**  $X, Y \in GT^+(p, \mathbb{R})$  を functionally independent な確率行列とし、 $A \in GT^+(p, \mathbb{R})$  を定数とする。このとき、 $Y = XA$  のヤコビアンは、符号を除いて、

$$dY = \left\{ \prod_{i=1}^p a_{ii}^{p-i+1} \right\} dX$$

となる。また、 $Y = AX$  のヤコビアンは、符号を除いて、

$$dY = \left\{ \prod_{i=1}^p a_{ii}^i \right\} dX$$

となる。

証明：くさび積を用いて証明すればよい。 □

**命題 1.3**  $X, Y \in \text{Sym}(p, \mathbb{R})$  を functionally independent な確率行列とし、 $A \in GL(p, \mathbb{R})$  を定数とする。このとき、 $Y = AXA'$  のヤコビアンは、符号を除いて、

$$dY = \text{Det}(A)^{p+1} dX$$

となる。

証明：直交行列  $H_1, H_2 \in O(p, \mathbb{R})$  と対角行列  $D = \text{diag}(d_1, \dots, d_p)$  を用いて、 $A = H_1 D H_2$  と書き直す。変換  $S \mapsto ASA'$  は

$$S \mapsto H_2 S H_2' \mapsto D H_2 S H_2' D \mapsto H_1 D H_2 S H_2' D H_1'$$

と書き直せる．一番目と三番目の変換は直交行列に関する変換なので，それらのヤコビアン  
の絶対値は 1 となる． $H_2SH_2' =: B = (b_{ij})$  としたとき，二番目の変換は

$$b_{ij} \mapsto d_i d_j b_{ij}$$

である．したがって， $B$  の次元が  $p(p+1)/2$  となることに注意すれば，ヤコビアンは

$$\prod_{i=1}^p \prod_{j=1}^i d_i d_j = \prod_{i=1}^p d_i^{p+1} = \text{Det}(D)^{p+1}$$

からわかる． □

**命題 1.4**  $X \in \text{Sym}^+(p, \mathbb{R})$  を functionally independent な確率行列とし， $T \in \text{GT}^+(p, \mathbb{R})$  とする．このとき， $X = TT'$  のヤコビアンは，

$$dX = 2^p \left\{ \prod_{i=1}^t a_{ii}^{p-i+1} \right\} dT$$

となる．このとき， $X = T'T$  のヤコビアンは，符号を除いて，

$$dX = 2^p \left\{ \prod_{i=1}^t t_{ii}^{pi} \right\} dT$$

となる．

証明：くさび積を用いて証明すればよい． □

**命題 1.5**  $X \in \text{Sym}(p, \mathbb{R})$  を functionally independent な確率行列とし，ゼロでない異なる固有根  $\ell_1 > \dots > \ell_p$  を持つとする． $H \in O(p, \mathbb{R})$  で  $X = H \text{diag}(\ell_1, \dots, \ell_p) H'$  とする．このとき，この変換のヤコビアンは

$$dX = \left\{ \prod_{i=1}^{p-1} \prod_{j=i+1}^p |\ell_i - \ell_j| \right\} \prod_{i=1}^p d\ell_i H' dH$$

となる．ただし， $dH$  は  $O(p, \mathbb{R})$  上に制約したルベーク測度である．

証明：Mathai (1997, page 117) を参照． □

**命題 1.6**  $p = p_1 + p_2$  とする．関数

$$f : \text{Sym}^+(p, \mathbb{R}) \rightarrow \text{Sym}^+(p_2, \mathbb{R}) \times \mathbb{R}^{p_1 \times p_2} \times \text{Sym}^+(p_2, \mathbb{R})$$

$$\begin{pmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{pmatrix} \mapsto (S_{11 \bullet 2}, S_{12} S_{22}^{-1}, S_{22})$$

の逆写像  $h$  は

$$h : \text{Sym}^+(p_1, \mathbb{R}) \times \mathbb{R}^{p_1 \times p_2} \times \text{Sym}^+(p_2, \mathbb{R}) \rightarrow \text{Sym}^+(p, \mathbb{R})$$

$$(A_{11}, A_{12}, A_{22}) \mapsto \begin{pmatrix} A_{11} + A_{12} A_{22}^{-1} A_{21} & A_{12} A_{22} \\ A_{22} A_{21} & A_{22} \end{pmatrix}$$

である。さらに、写像  $h$  のヤコビアンは

$$\text{Det}(A_{22})^{p_1}$$

となる。

証明：

$$\text{Det} \begin{pmatrix} \frac{\partial S_{11}}{\partial A_{11}} & \frac{\partial S_{11}}{\partial A_{12}} & \frac{\partial S_{11}}{\partial A_{22}} \\ \frac{\partial S_{12}}{\partial A_{11}} & \frac{\partial S_{12}}{\partial A_{12}} & \frac{\partial S_{12}}{\partial A_{22}} \\ \frac{\partial S_{22}}{\partial A_{11}} & \frac{\partial S_{22}}{\partial A_{12}} & \frac{\partial S_{22}}{\partial A_{22}} \end{pmatrix} = \text{Det} \begin{pmatrix} I_r & * & * \\ 0 & \frac{\partial S_{12}}{\partial A_{12}} & * \\ 0 & 0 & I_s \end{pmatrix} = \text{Det} \left( \frac{\partial S_{12}}{\partial A_{12}} \right)$$

と命題 1.1 からわかる。

□

### 1.3 分布論

この節では,  $\Sigma \in \text{Sym}^+(p, \mathbb{R})$  とし,  $n \geq p$  を仮定する. このとき,  $\text{Sym}(p, \mathbb{R})$  上の測度

$$W_p(\Sigma, n)(dx) = \frac{1}{C(p, n)(\text{Det } \Sigma)^{n/2}} (\text{Det } x)^{n/2} \text{Tr} \left( -\frac{1}{2} \Sigma^{-1} x \right) \mathbb{I}_{\text{Sym}^+(p, \mathbb{R})}(x) \frac{dx}{(\text{Det } x)^{(p+1)/2}} \quad (1.4)$$

を考える. ただし,

$$C(p, n) = 2^{np} \pi^{p(p-1)/4} \prod_{i=0}^{p-1} \Gamma((n-i)/2),$$

$$\mathbb{I}_{\text{Sym}^+(p, \mathbb{R})}(x) = \begin{cases} 1 & (x \in \text{Sym}^+(p, \mathbb{R})), \\ 0 & (\text{その他}) \end{cases}$$

である. 確率行列  $S \in \text{Sym}^+(p, \mathbb{R})$  が確率測度 (1.4) を持つ (まだ、確率測度であることはこれまでのこの講義の議論ではわかっていない. また、ウイシャート行列の密度関数になることもこれまでの講義の議論ではわかっていない!) とき,  $S \sim W_p(\Sigma, n)$  と書くことにする.

#### 分割に対する分布

命題 1.7  $S \in \text{Sym}^+(p, \mathbb{R})$  はウイシャート分布  $W_p(\Sigma, n)$  に従う. ただし,  $n \geq p$  を仮定し,  $\Sigma \in \text{Sym}^+(p, \mathbb{R})$  である.

$$S = \begin{bmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{bmatrix}, \quad \Sigma = \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix},$$

$$S_{11 \bullet 2} = S_{11} - S_{12} S_{22}^{-1} S_{21}, \quad \Sigma_{11 \bullet 2} = \Sigma_{11} - \Sigma_{12} \Sigma_{22}^{-1} \Sigma_{21}$$

とする. ここで,  $S_{11}, \Sigma_{11}$  は  $p_1 \times p_1$  とし,  $p_2 = p - p_1$  である. このとき, つぎが成立する.

- (i)  $S_{11 \bullet 2}$  と  $(S_{21}, S_{22})$  は独立,
- (ii)  $S_{11 \bullet 2}$  は  $W_{p_1}(n - p_2, \Sigma_{11 \bullet 2})$  に従う,
- (iii)  $S_{22}$  は  $W_{p_2}(n, \Sigma_{22})$  に従う,
- (iv)  $S_{21} | S_{22}$  は  $N(S_{22} \Sigma_{22}^{-1} \Sigma_{21}, S_{22} \otimes \Sigma_{11 \bullet 2})$  に従う.

証明: 命題 1.6 からわかる. □

系 1.1

$$\int_{\text{Sym}(p, \mathbb{R})} W_p(\Sigma, n)(dx) = 1$$

である.

証明:  $p = 1$  とし,  $\Sigma = \sigma^2$  と書く.

$$W_1(\Sigma, n)(dx) = \frac{1}{2^n \Gamma(n/2) \sigma^n} x^{n/2-1} \exp\left(-\frac{x}{2\sigma^2}\right) \mathbb{I}_{(0, \infty)}(x)$$

となる .  $y = x/\sigma^2$  と変換して ,  $y$  について積分をすれば ,

$$\int_0^\infty W_1(\Sigma, n)(dx) = 1$$

となることがわかる . あとは , 命題 1.7 を繰り返して適用すればよい .  $\square$

### Bartlett 分解

命題 1.8  $S \sim W_p(I_p, n)$  とする .  $S = TT'$  と分解する . ただし ,  $T = (t_{ij}) \in \text{GT}^+(p, \mathbb{R})$  である . このとき ,  $t_{ij}$  は互いに独立であって ,  $t_{ii}^2 \sim \chi_{(n-i+1)}^2$  ,  $t_{ij} \sim N(0, 1)$  ( $i = 1, 2, \dots, p$  ,  $j = 1, 2, \dots, i$ ) に従う .

証明 : 命題 1.4 または 1.7 からわかる .  $\square$

### 固有根の分布

命題 1.9  $S \in \text{Sym}^+(p, \mathbb{R})$  を連続な分布をもつランダム行列で確率測度  $f(S)dS$  を持つとする . ただし ,  $dS$  は  $\text{Sym}^+(p, \mathbb{R})$  上のルベーグ測度とする . このとき ,  $S$  の固有値  $\ell = (\ell_1, \dots, \ell_p)$  ( $\ell_1 \geq \dots \geq \ell_p$ ) の確率測度は

$$\mu(d\ell) = \frac{\pi^{n^2/2}}{\Gamma_p(n/2)} \prod_{i < j} (\ell_i - \ell_j) \int_{\text{O}(p, \mathbb{R})} f(HDH') \mu_{\text{O}(p)}(dH) \prod_{i=1}^p d\ell_i$$

であたえられる . ただし ,

$$\begin{aligned} \mu_{\text{O}(p, \mathbb{R})}(dH) &= \frac{\pi^{n^2/2}}{\Gamma(n/2)} H' dH, \\ \Gamma_p(n/2) &= \pi^{p(p-1)/4} \prod_{i=0}^{p-1} \Gamma((n-i)/2) \end{aligned}$$

である .

証明 : まず ,

$$\int_{\text{O}(p, \mathbb{R})} \mu_{\text{O}(p, \mathbb{R})}(dH) = 1$$

となることが知られていることに注意する . Mathai (1997, page 117) を参照 . すなわち ,  $\text{O}(p, \mathbb{R})$  上の基準化されたハール測度である . あとは , 命題 1.5 よりわかる .  $\square$



# 第2章

## 2.1 自然指数分布族

$\mathcal{V}$  を有限次元実ベクトル空間とし,  $\mathcal{V}^*$  をその双対空間<sup>(2-1)</sup>とする.  $\theta \in \mathcal{V}^*$  に対して,  $x \in \mathcal{V}$  に対応する値を  $\langle \theta, x \rangle$  と書く.  $(\theta, x) \mapsto \langle \theta, x \rangle$  は  $\mathcal{V}^* \times \mathcal{V}$  上の非退化な双一次形式になる.

$\mu$  を  $\mathcal{V}$  上のラドン測度 (必ずしも有限測度ではない) とする.  $\mu$  に対して, 写像  $L_\mu : \mathcal{V}^* \rightarrow [0, \infty]$  を

$$L_\mu(\theta) = \int_{\mathcal{V}} \exp\langle \theta, x \rangle \mu(dx), \quad (\theta \in \mathcal{V}^*)$$

で定める. これを測度  $\mu$  のラプラス変換という. ヘルダーの不等式<sup>(2-2)</sup>より, 集合

$$D(\mu) = \{\theta \in \mathcal{V}^* : L_\mu(\theta) < \infty\}$$

は凸集合となる. さらに,

$$k_\mu(\theta) = \log L_\mu(\theta)$$

は  $D(\mu)$  上の凸関数となる. さらに,  $\mathcal{R}(\mathcal{V})$  を  $\mathcal{V}$  上のラドン測度でつぎの二つの条件をみたす集合とする: (i)  $D(\mu)$  の内部 ( $\Theta(\mu) := \text{Int}(D(\mu))$ ) は空でない, (ii)  $\mu$  は  $\mathcal{V}$  のアファイン平面に集中しない. このとき,  $\mu \in \mathcal{R}(\mathcal{V})$  ならば,  $k_\mu$  は実解析的で,  $\Theta(\mu)$  上で狭義凸関数となる. Brown (1986, page 39) を参照のこと.

$\mu \in \mathcal{R}(\mathcal{V})$  と  $\theta \in \Theta(\mu)$  に対して,  $\mathcal{V}$  上の確率測度  $P$  を

$$P(\theta, \mu)(dx) = \exp\{\langle \theta, x \rangle - k_\mu(\theta)\} \mu(dx)$$

で定める. 確率測度の族

$$\mathcal{F} = \mathcal{F}(\mu) := \{P(\theta, \mu)(dx) : \theta \in \Theta(\mu)\}$$

を  $\mu$  によって生成された自然指数分布族 (NEF) という.

注意 2.1  $\mu_1, \mu_2 \in \mathcal{R}(\mathcal{V})$  に対して

$$\mathcal{F}(\mu_1) = \mathcal{F}(\mu_2) \iff \exists (a, b) \in \mathcal{V}^* \times \mathbb{R} \text{ があって, } \mu_1(dx) = \exp\{\langle a, x \rangle + b\} \mu_2(dx)$$

なぜならば,  $\mu_2 \in \mathcal{F}(\mu_1)$  となるので,  $\exists \theta \in \Theta(\mu_1)$  があって,

$$\mu_2(dx) = P(\theta, \mu_1)(dx)$$

と書けることからわかる. □

**平均と分散**  $\mathcal{L}(\mathcal{V}^*, \mathcal{V})$  は  $\mathcal{V}^*$  から  $\mathcal{V}$  への線形写像全体の空間とする．写像  $\dot{k}_\mu : \mathcal{V}^* \rightarrow \mathcal{V}$  と  $\ddot{k}_\mu : \mathcal{V}^* \rightarrow \mathcal{L}(\mathcal{V}^*, \mathcal{V})$  をつぎのように定める：任意の  $\xi, \eta \in \mathcal{V}^*$  に対して，

$$\langle \eta, \dot{k}_\mu(\theta) \rangle = \dot{k}_\mu(\theta)(\eta) = \left. \frac{d}{dt} k_\mu(\theta + t\eta) \right|_{t=0}$$

で定める．さらに，

$$\langle \xi, \ddot{k}_\mu(\theta)(\eta) \rangle = \left. \frac{d^2}{dsdt} k_\mu(\theta + s\xi + t\eta) \right|_{s=0, t=0}$$

で定める．

補題 2.1  $\theta \in \Theta(\mu)$  に対して，

$$\begin{aligned} \dot{k}_\mu(\theta) &= \int_{\mathcal{V}} x P(\theta, \mu)(dx), \\ \langle \xi, \ddot{k}_\mu(\theta)(\eta) \rangle &= \int_{\mathcal{V}} \langle \xi, x - \dot{k}_\mu(\theta) \rangle \langle \eta, x - \dot{k}_\mu(\theta) \rangle P(\theta, \mu)(dx), \quad (\xi, \eta \in \mathcal{V}^*) \end{aligned}$$

となる．

証明：Brown (1986, page 35) と同様にすれば，任意の  $\eta \in \mathcal{V}^*$  に対して，

$$\left. \frac{d}{dt} \int_C \exp\langle \theta + t\eta, x \rangle \mu(dx) \right|_{t=0} = \int_C \langle \eta, x \rangle \exp\langle \theta, x \rangle \mu(dx)$$

が確認できる．これより，

$$\begin{aligned} \langle \eta, \dot{k}_\mu(\theta) \rangle &= \dot{k}_\mu(\theta)(\eta) = \frac{1}{\int_C \exp\langle \theta + t\eta, x \rangle \mu(dx)} \left. \int_C \langle \eta, x \rangle \exp\langle \theta + t\eta, x \rangle \mu(dx) \right|_{t=0} \\ &= \exp\{-k_\mu(\theta)\} \int_C \langle \eta, x \rangle \exp\langle \theta, x \rangle \mu(dx) \\ &= \int_C \langle \eta, x \rangle \exp\{\langle \theta, x \rangle - k_\mu(\theta)\} \mu(dx) \\ &= \left\langle \eta, \int_C x \exp\{\langle \theta, x \rangle - k_\mu(\theta)\} \mu(dx) \right\rangle \end{aligned}$$

となり，補題の前半部分が示せた．

まず，任意の  $\xi, \eta \in \mathcal{V}^*$  に対して，

$$\left. \frac{d^2}{dsdt} \int_C \exp\langle \theta + s\xi + t\eta, x \rangle \mu(dx) \right|_{(s,t)=(0,0)} = \int_C \langle \xi, x \rangle \cdot \langle \eta, x \rangle \exp\langle \theta, x \rangle \mu(dx)$$

となることが注意する．これより，

$$\begin{aligned} \frac{d^2}{dsdt} k_\mu(\theta + s\xi + t\eta) &= \frac{1}{L_\mu(\theta + s\xi + t\eta)} \frac{d^2}{dsdt} L_\mu(\theta + s\xi + t\eta) \Big|_{(s,t)=(0,0)} \\ &\quad - \frac{1}{\{L_\mu(\theta + s\xi + t\eta)\}^2} \frac{d}{ds} L_\mu(\theta + s\xi + t\eta) \frac{d}{dt} L_\mu(\theta + s\xi + t\eta) \Big|_{(s,t)=(0,0)} \\ &= \int_C \langle \xi, x \rangle \cdot \langle \eta, x \rangle \exp\{\langle \theta, x \rangle - k_\mu(\theta)\} \mu(dx) - \langle \xi, \dot{k}_\mu(\theta) \rangle \cdot \langle \eta, \dot{k}_\mu(\theta) \rangle \end{aligned}$$

よりわかる .

□

**平均母数と分散関数**  $\mathcal{V}$  の部分集合  $\mathcal{M}_{\mathcal{F}}$  を

$$\mathcal{M}_{\mathcal{F}} = \dot{k}_{\mu}(\Theta(\mu))$$

で定める . これを NEF  $\mathcal{F}$  の平均領域という .  $\mu \in \mathcal{R}(\mathcal{V})$  ならば ,  $\dot{k}_{\mu}$  は狭義単調関数なので , 写像  $\theta \mapsto \dot{k}_{\mu}(\theta)$  は  $\Theta(\mu)$  から  $\mathcal{M}_{\mathcal{F}}$  への全単射となる .  $\dot{k}_{\mu}$  の逆写像を定義することができるので , それを  $\psi_{\mu} : \mathcal{M}_{\mathcal{F}} \rightarrow \Theta(\mu)$  ( $m \mapsto \psi_{\mu}(m)$ ) とする .

任意の  $m \in \mathcal{M}_{\mathcal{F}}$  に対して ,

$$\tilde{P}(m, \mathcal{F})(dx) = P(\psi_{\mu}(m), \mu)(dx)$$

とする . すると  $m \mapsto \tilde{P}(m, \mathcal{F})$  ( $\mathcal{M}_{\mathcal{F}} \rightarrow \mathcal{F}$ ) は NEF  $\mathcal{F}$  の新たな母数化となる .

確率測度  $\tilde{P}(m, \mathcal{F})$  の分散作用素を  $V_{\mathcal{F}}(m) : \mathcal{V}^* \rightarrow L(\mathcal{V}^*, \mathcal{V})$  とおけば ,

$$V_{\mathcal{F}}(m) = \ddot{k}_{\mu}(\psi_{\mu}(m)), \quad (m \in \mathcal{M}_{\mathcal{F}})$$

となる . 実際 ,  $\xi, \eta \in \mathcal{V}^*$  に対して ,

$$\begin{aligned} \langle \xi, \ddot{k}_{\mu}(\psi_{\mu}(m))(\eta) \rangle &= \int_{\mathcal{V}} \langle \xi, x - \dot{k}_{\mu}(\psi_{\mu}(m)) \rangle \langle \eta, x - \dot{k}_{\mu}(\psi_{\mu}(m)) \rangle P(\psi_{\mu}(m), \mu)(dx) \\ &= \int_{\mathcal{V}} \langle \xi, x - m \rangle \langle \eta, x - m \rangle \tilde{P}(m, \mathcal{F})(dx) = \langle \xi, V_{\mathcal{F}}(m)(\eta) \rangle \end{aligned}$$

となることよりわかる .

写像  $m \mapsto V_{\mathcal{F}}(m)$  ( $\mathcal{M}_{\mathcal{F}} \rightarrow \mathcal{L}(\mathcal{V}^*, \mathcal{V})$ ) を分散関数とよぶ . また , 分散関数  $V_{\mathcal{F}}$  は写像  $\dot{\psi}_{\mu} : \mathcal{M}_{\mathcal{F}} \rightarrow \mathcal{L}(\mathcal{V}, \mathcal{V}^*)$  を用いて ,

$$V_{\mathcal{F}}(m) = \{\dot{\psi}_{\mu}(m)\}^{-1} \quad (2.1)$$

と表現される . なぜならば , 任意の  $x \in \mathcal{V}$  に対して ,

$$\dot{\psi}_{\mu}(m)(x) = \lim_{t \rightarrow 0} \frac{\psi_{\mu}(m + tx) - \psi_{\mu}(m)}{t} \in \mathcal{V}^*$$

となる .  $\psi_{\mu} : \mathcal{V} \rightarrow \mathcal{V}^*$  から

$$m \mapsto \dot{\psi}_{\mu}(m) \quad (\mathcal{M}_{\mathcal{F}} \rightarrow \mathcal{L}(\mathcal{V}, \mathcal{V}^*))$$

であることがわかる .  $\dot{k}_{\mu}(\psi_{\mu}(m)) = m$  に注意する . 任意の  $x \in \mathcal{V}$  に対して , 合成関数の微分を使えば ,

$$\begin{aligned} \ddot{k}_{\mu}(\psi_{\mu}(m))\dot{\psi}_{\mu}(m)(x) &= \lim_{t \rightarrow 0} \frac{\dot{k}_{\mu}(\psi_{\mu}(m + tx)) - \dot{k}_{\mu}(\psi_{\mu}(m))}{t} \\ &= \lim_{t \rightarrow 0} \frac{m + tx - m}{t} = x \end{aligned}$$

となる . したがって ,

$$\ddot{k}_{\mu}(\psi_{\mu}(m))\dot{\psi}_{\mu}(m) = \text{id}$$

をえる . よって ,

$$\ddot{k}_{\mu}(\psi_{\mu}(m)) = \{\dot{\psi}_{\mu}(m)\}^{-1}$$

がわかる .

例 2.1  $\alpha > 0$  に対して, 測度

$$\mu(dx) = x^{\alpha-1} \mathbb{I}_{(0, \infty)}(x) dx$$

によって生成される自然指数分布族を求める:

$$\begin{aligned} L_\mu(\theta) &= \int e^{\theta x} \mu(dx) = (-\theta)^{-\alpha} \Gamma(\alpha), & (\theta < 0), \\ k_\mu(\theta) &= -\alpha \log(-\theta) + \log \Gamma(\alpha), \\ \mathcal{F}(\mu) &= \{\exp(\theta x - k_\mu(\theta)) \mu(dx) : \theta \in \Theta(\mu) = (-\infty, 0)\} \\ &= \left\{ \frac{(-\theta)^\alpha}{\Gamma(\alpha)} e^{\theta x} x^{\alpha-1} \mathbb{I}_{(0, \infty)}(x) dx : \theta \in \Theta(\mu) \right\}, \\ \dot{k}_\mu(\theta) &= -\frac{\alpha}{\theta}, & \psi_\mu(m) = -\alpha m^{-1}, & (m \in \mathcal{M}_m u = \dot{k}_\mu(\Theta(\mu)) = (0, \infty)) \\ \ddot{k}_\mu(\theta) &= \frac{\alpha}{\theta^2}, & V_{\mathcal{F}}(m) = \ddot{k}_\mu(\theta) = \frac{1}{\alpha} (k_\mu(\theta))^2 = \frac{1}{\alpha} m^2 \end{aligned}$$

と<sup>(2-3)</sup>なる . □

定理 2.1 (Letac and Mora (1990))  $\mathcal{F}_1$  と  $\mathcal{F}_2$  を  $\mathcal{V}$  上の NEF とする .  $\mathcal{M}_{\mathcal{F}_1} \cap \mathcal{M}_{\mathcal{F}_2}$  の空でない開集合上で  $V_{\mathcal{F}_1}$  と  $V_{\mathcal{F}_2}$  が一致すれば ,  $\mathcal{F}_1 = \mathcal{F}_2$  である .

証明 : Jørgensen (1997, page 51) を参照せよ . □

## 2.2 ウィシャート NEF 族

$p \geq 1$  を整数とする .  $\mathcal{V} = \text{Sym}(p, \mathbb{R})$  とし ,  $\mathcal{V}$  上の内積を

$$(\boldsymbol{\theta} | \boldsymbol{x}) = \text{Tr}(\boldsymbol{\theta} \boldsymbol{x}), \quad (\boldsymbol{\theta}, \boldsymbol{x} \in \mathcal{V})$$

とおく .  $\boldsymbol{\theta} \in \mathcal{V}^*$  に対して , ある元  $\boldsymbol{\theta} \in \mathcal{V}$  が一意的に存在して ,

$$\langle \boldsymbol{\theta}, \boldsymbol{x} \rangle = (\boldsymbol{\theta} | \boldsymbol{x})$$

となることに注意する . さらに ,  $\boldsymbol{w} = (w_1, w_2, \dots, w_p)' \in \mathbb{R}^p$  に対して ,  $\boldsymbol{x}_w \in \mathcal{V}$  を

$$\boldsymbol{x}_w = \boldsymbol{w} \boldsymbol{w}'$$

で定める . したがって ,

$$(\boldsymbol{\theta} | \boldsymbol{x}_w) = \sum_{i,j}^p w_i \theta_{ij} w_j, \quad (\boldsymbol{\theta} = (\theta_{ij}))$$

となる .  $\mathbb{R}^p$  上の測度  $(2\pi)^{-p/2} dw_1 dw_2 \cdots dw_p$  に対して , 写像  $\boldsymbol{w} \mapsto \boldsymbol{x}_w$  によって誘導される  $\mathcal{V}$  上の測度を  $\mu_{1/2}$  とする . 測度  $\mu_{1/2}$  のラプラス変換  $L_{\mu_{1/2}}$  はつぎで与えられる :  $\boldsymbol{\theta} \in \mathcal{V}$  に対して ,

$$\begin{aligned} L_{\mu_{1/2}} \left( \frac{\boldsymbol{\theta}}{2} \right) &= \int_{\mathcal{V}} \exp \left( \frac{1}{2} (\boldsymbol{\theta} | \boldsymbol{x}) \right) \mu_{1/2}(d\boldsymbol{x}) \\ &= \int_{\mathbb{R}^p} \exp \left( \frac{1}{2} \sum_{i,j}^p w_i \theta_{ij} w_j \right) \frac{dw_1 dw_2 \cdots dw_p}{(2\pi)^{p/2}} \end{aligned}$$

となる . したがって ,

$$L_{\mu_{1/2}} \left( -\frac{\boldsymbol{\theta}}{2} \right) = \begin{cases} \text{Det}(\boldsymbol{\theta}) & (\boldsymbol{\theta} \in \text{Sym}^+(p, \mathbb{R})), \\ \infty & (\text{その他}) \end{cases}$$

となる .  $\text{Sym}^+(p, \mathbb{R})$  は  $\text{Sym}(p, \mathbb{R})$  の開部分集合なので ,

$$\begin{aligned} \Theta(\mu_{1/2}) &= -\text{Sym}^+(p, \mathbb{R}), \\ k_{\mu_{1/2}}(\boldsymbol{\theta}) &= -\frac{1}{2} \log \text{Det}(-2\boldsymbol{\theta}), \quad (\boldsymbol{\theta} \in -\text{Sym}^+(p, \mathbb{R})) \end{aligned}$$

となる . NEF  $\mathcal{F}(\mu_{1/2})$  , すなわち ,

$$\begin{aligned} \mathcal{F}(\mu_{1/2}) &= \{W(\boldsymbol{\theta}, \mu_{1/2}) : L_{W(\boldsymbol{\theta}, \mu_{1/2})}(\boldsymbol{\eta}) = k_{\mu_{1/2}}(\boldsymbol{\theta} + \boldsymbol{\eta}) - k_{\mu_{1/2}}(\boldsymbol{\theta}), \\ &\quad \text{for } \boldsymbol{\theta} \in -\text{Sym}^+(p, \mathbb{R}), \boldsymbol{\theta} + \boldsymbol{\eta} \in -\text{Sym}^+(p, \mathbb{R})\} \\ &= \{W(\boldsymbol{\theta}, \mu_{1/2}) = \exp\{\langle \boldsymbol{\theta}, \boldsymbol{x} \rangle - k_{\mu_{1/2}}(\boldsymbol{\theta})\} \mu_{1/2}(d\boldsymbol{x})\} \end{aligned}$$

を母数  $1/2$  (自由度  $1$ ) のウィシャート分布という .

この分布族の解釈はつぎのようになる :  $\boldsymbol{Y} = (Y_1, \dots, Y_p)' \in \mathbb{R}^p$  は  $N_p(0, \Gamma)$  に従うとする . このとき ,

$$\boldsymbol{x}_Y = \boldsymbol{Y} \boldsymbol{Y}'$$

は  $W(-\Gamma^{-1}/2, \mu_{1/2})$  に従う . 実際 ,  $\mathbf{s} \in \text{Sym}(p, \mathbb{R})$  に対して ,

$$\mathbb{E} \exp \left( -\frac{1}{2} \sum_{i,j}^p Y_i s_{ij} Y_j \right) = \begin{cases} \frac{\text{Det}(\Gamma^{-1})^{1/2}}{\text{Det}(\mathbf{s} + \Gamma^{-1})^{1/2}} & (\mathbf{s} \in \text{Sym}^+(p, \mathbb{R})), \\ \infty & (\text{その他}) \end{cases}$$

となる . したがって ,  $\boldsymbol{\theta} \in -\text{Sym}^+(p, \mathbb{R})$  に対して ,

$$\begin{aligned} \mathbb{E} \exp(\boldsymbol{\theta} | \mathbf{x}_Y) &= \frac{\text{Det}(\Gamma^{-1})^{1/2}}{\text{Det}(-2\boldsymbol{\theta} + \Gamma^{-1})^{1/2}} \\ &= \exp \left\{ \frac{1}{2} \log \text{Det}(\Gamma^{-1}) - \frac{1}{2} \log \text{Det}(-2\boldsymbol{\theta} + \Gamma^{-1}) \right\} \\ &= \exp \left\{ k_{\mu_{1/2}} \left( \boldsymbol{\theta} - \frac{1}{2} \Gamma^{-1} \right) - k_{\mu_{1/2}} \left( -\frac{1}{2} \Gamma^{-1} \right) \right\} \end{aligned}$$

となる . 一方 ,

$$\begin{aligned} L_{W(-\Gamma^{-1}/2, \mu_{1/2})}(\boldsymbol{\theta}) &= \int_{\mathcal{V}} \exp(\boldsymbol{\theta} | \mathbf{x}) W(-\Gamma^{-1}/2, \mu_{1/2})(d\mathbf{x}) \\ &= \int_{\mathcal{V}} \exp(\boldsymbol{\theta} | \mathbf{x}) \exp \left\{ \left( -\frac{1}{2} \Gamma^{-1} | \mathbf{x} \right) - k_{\mu_{1/2}}(\boldsymbol{\theta}) \right\} \mu_{1/2}(d\mathbf{x}) \\ &= \int_{\mathcal{V}} \exp \left\{ \left( \boldsymbol{\theta} - \frac{1}{2} \Gamma^{-1} | \mathbf{x} \right) - k_{\mu_{1/2}} \left( \boldsymbol{\theta} - \frac{1}{2} \Gamma^{-1} \right) + k_{\mu_{1/2}} \left( \boldsymbol{\theta} - \frac{1}{2} \Gamma^{-1} \right) - k_{\mu_{1/2}}(\boldsymbol{\theta}) \right\} \mu_{1/2}(d\mathbf{x}) \\ &= \exp \left\{ k_{\mu_{1/2}} \left( \boldsymbol{\theta} - \frac{1}{2} \Gamma^{-1} \right) - k_{\mu_{1/2}}(\boldsymbol{\theta}) \right\} \end{aligned}$$

となるので , ラプラス変換が一致することからわかる .

$n > (p-1)/2$  に対して ,  $\mathcal{V}$  上の測度を

$$\mu_n(d\mathbf{x}) = \frac{1}{c(p, n)} \text{Det}(\mathbf{x})^{n-(p-1)/2-1} \mathbb{I}_{\text{Sym}^+(p, \mathbb{R})}(\mathbf{x}) d\mathbf{x}$$

とおく . ただし ,  $d\mathbf{x}$  は  $\mathcal{V}$  上に制限したルベーク速度で  $d\mathbf{x} = \prod_{i \leq j} dx_{ij}$  ,  $(\mathbf{x} = (x_{ij}))$  で

$$c(p, n) = 2^p \pi^{p(p-1)/4} \prod_{i=1}^p \Gamma \left( n - \frac{i-1}{2} \right)$$

である . このとき ,

$$\int_{\mathcal{V}} \exp \left( -\frac{1}{2} (\mathbf{s} | \mathbf{x}) \right) \mu_n(d\mathbf{x}) = \begin{cases} |\text{Det}(\mathbf{s})|^{-n} & (\mathbf{s} \in \text{Sym}(p, \mathbb{R})), \\ \infty & (\text{その他}) \end{cases}$$

となる . このことから

$$\begin{aligned} \Theta(\mu_n) &= -\text{Sym}^+(p, \mathbb{R}), \\ k_{\mu_n}(\boldsymbol{\theta}) &= -n \log \text{Det}(-2\boldsymbol{\theta}), \quad \mathbf{S} \in -\text{Sym}^+(p, \mathbb{R}) \end{aligned}$$

となる .

いま ,

$$\Lambda = \{n > 0 : \exists \mu \in \mathcal{R}(\mathcal{V}) \text{ s.t. } \Theta(\mu) = -\text{Sym}^+(p, \mathbb{R}), k_\mu(\boldsymbol{\theta}) = -n \log \text{Det}(-2\boldsymbol{\theta})\}$$

とおく . 明らかに ,

$$\Lambda \supset \Lambda_0 := \left\{ \frac{1}{2}, 1, \frac{3}{2}, \dots, \frac{p-1}{2} \cup \left( \frac{p-1}{2}, \infty \right) \right\}$$

である . Gindikin (1964) によれば ,

$$\Lambda = \Lambda_0$$

となることが知られている . この証明については , Peddada and Richards (1991) を参照のこと .

以後では ,  $n \in \Lambda$  に対して ,

$$\begin{aligned} \mathcal{F}_n &:= \mathcal{F}(\mu_n) \\ &= \{W(\boldsymbol{\theta}, \mu_n) : L_{W(\boldsymbol{\theta}, \mu_n)}(\boldsymbol{\eta}) = k_{\mu_n}(\boldsymbol{\theta} + \boldsymbol{\eta}) - k_{\mu_n}(\boldsymbol{\theta}), \\ &\quad \text{for } n \in \Lambda, \boldsymbol{\theta} \in -\text{Sym}^+(p, \mathbb{R}), \boldsymbol{\theta} + \boldsymbol{\eta} \in -\text{Sym}^+(p, \mathbb{R})\} \\ &= \{W(\boldsymbol{\theta}, \mu_n)(dx) = \exp\{\langle \boldsymbol{\theta} | \mathbf{x} \rangle - k_{\mu_n}(\boldsymbol{\theta})\} \mu_n(dx)\} \end{aligned}$$

と記し ,  $\mathcal{F}_n$  の元を母数  $n$  の NEF ウィシャート分布族と呼ぶことにする . さらに , 記号の乱用になるが , 簡単に

$$W(\mathbf{m}, \mu_n) = W(\psi_{\mu_n}(\mathbf{m}), \mu_n)$$

とも書くことにする .

定理 2.2  $n \in \Lambda$  とし ,  $\mathcal{F}_n$  を母数  $n$  の NEF ウィシャート分布族とする . このとき ,

$$\mathcal{M}(\mathcal{F}_n) = \text{Sym}^+(p, \mathbb{R})$$

となる . また , 各  $\mathbf{m} \in \text{Sym}^i(p, \mathbb{R})$  に対して , 分散関数  $V_{\mathcal{F}_n}(\mathbf{m}) : \mathcal{V} \rightarrow \mathcal{V}$  は

$$\boldsymbol{\theta} \mapsto \frac{1}{n} \mathbf{m} \boldsymbol{\theta} \mathbf{m}$$

で与えられる . ただし .  $\mathbf{m} \boldsymbol{\theta} \mathbf{m}$  は行列  $\mathbf{m}, \boldsymbol{\theta}$  の積である .

証明 : 分散関数の表現 (2.1) と  $\psi_\mu(\mathbf{m})$  は  $\dot{k}_\mu(\boldsymbol{\theta})$  の逆関数であったことを思い出す . まず ,  $\dot{k}_\mu$  を求める .  $\Theta(\mu_n) = -\text{Sym}^+(p, \mathbb{R})$  上で

$$k_{\mu_n}(\boldsymbol{\theta}) = -n \log \text{Det}(-2\boldsymbol{\theta}), \quad \boldsymbol{\theta} \in -\text{Sym}^*(p, \mathbb{R})$$

より , 補題 A.1(i) より ,  $\mathbf{x} \in \text{Sym}^+(p, \mathbb{R})$  に対して ,

$$\begin{aligned} \dot{k}_{\mu_n}(\boldsymbol{\theta})(\mathbf{x}) &= \frac{d}{dt} k_{\mu_n}(\boldsymbol{\theta} + t\mathbf{x}) = -n \frac{d}{dt} \log \text{Det}(-2\boldsymbol{\theta} - 2t\mathbf{x}) \Big|_{t=0} \\ &= -n \frac{d}{dt} (\{-2\boldsymbol{\theta} - 2t\mathbf{x}\}^{-1} | -2\mathbf{x}) \Big|_{t=0} = -n(\boldsymbol{\theta}^{-1} | \mathbf{x}) \end{aligned}$$

より

$$\dot{k}_{\mu_n}(\boldsymbol{\theta}) = -n\boldsymbol{\theta}^{-1} \quad (2.2)$$

となる．よって，

$$\dot{k}_{\mu_n}(\Theta(\mu_n)) = \text{Sym}^+(p, \mathbb{R})$$

となる．(2.2) から  $\dot{k}_{\mu_n}$  の逆写像  $\psi_{\mu_n}$  は

$$\psi_{\mu_n}(\mathbf{m}) = -n\mathbf{m}^{-1}, \quad (\mathbf{m} \in \text{Sym}^+(p, \mathbb{R})) \quad (2.3)$$

となる．さらに，補題 A.1(ii) より， $\mathbf{x} \in \text{Sym}^+(p, \mathbb{R})$  に対して，

$$\begin{aligned} \dot{\psi}_{\mu_n}(\mathbf{m})(\mathbf{x}) &= \left. \frac{d}{dt} \psi_{\mu_n}(\mathbf{m} + t\mathbf{x}) \right|_{t=0} \\ &= -n \left. \frac{d}{dt} (\mathbf{m} + t\mathbf{x})^{-1} \right|_{t=0} = n(\mathbf{m} + t\mathbf{x})^{-1} \mathbf{x} (\mathbf{m} + t\mathbf{x})^{-1} \Big|_{t=0} \\ &= n\mathbf{m}^{-1} \mathbf{x} \mathbf{m}^{-1} \end{aligned}$$

となる．これより

$$\{\dot{\psi}_{\mu}(\mathbf{m})\}^{-1}(\mathbf{x}) = n^{-1} \mathbf{m} \mathbf{x} \mathbf{m}$$

となる．したがって，

$$V_{\mathcal{F}_n}(\mathbf{m})(\mathbf{x}) = \{\dot{\psi}_{\mu_n}(\mathbf{m})\}^{-1}(\mathbf{x}) = \frac{1}{n} \mathbf{m} \mathbf{x} \mathbf{m}$$

がわかる． □

ここから記号の定義と命題 2.1 が正確さを欠くおそれがあるので注意  
つぎにモデルの変換を考える． $a \in \text{GL}(p, \mathbb{R})$  によるモデルの変換を

$$g_a : W(\mathbf{m}, \mu_n) \mapsto W(g_a^{-1} \mathbf{m}, g_a \mu_n)$$

で定める．ただし， $g_a^{-1} \mu_n(B) = \mu_n(g_a B)$  ( $B \subset \text{Sym}(p, \mathbb{R})$ ) である．この変換によって得たモデルを  $g_a \mathcal{F}_n$  と書くこと（モデルへの作用を同じ記号  $g_a$  と書いているので，多少記号の乱用があることに注意）にする．すなわち，

$$g_a \mathcal{F}_n = \{g_a W(\mathbf{m}, \mu_n) = W(g_a^{-1} \mathbf{m}, g_a^{-1} \mu_n) : W \in \mathcal{F}_n\},$$

である．

命題 2.1  $a \in \text{GL}(p, \mathbb{R})$  に対して， $\text{Sym}(p, \mathbb{R})$  上の変換  $g_a$  を

$$g_a(\mathbf{x}) = \mathbf{a}' \mathbf{x} \mathbf{a}, \quad (\mathbf{x} \in \text{Sym}(p, \mathbb{R}))$$

で定義する．このとき，母数  $n$  ( $n \in \Lambda$ ) のウイシャート分布族 NEF  $\mathcal{F}_n$  は変換  $g_a$  に関して不変である．

証明：  $g_a(\text{Sym}^+(p, \mathbb{R})) = \text{Sym}^+(p, \mathbb{R})$  となるので，  $\mathcal{M}_{\mathcal{F}_n}$  は  $g_a$  に関して不変である．各  $\mathbf{m} \in \text{Sym}^+(p, \mathbb{R})$  に対して，

$$V_{g_a(\mathcal{F}_n)}(\mathbf{m}) = g_a \circ V_{\mathcal{F}_n}(g_a^{-1}(\mathbf{m})) \circ {}^t g_a$$

となる．ただし，  ${}^t g_a$  は，任意の  $\mathbf{x}, \mathbf{y} \in \text{Sym}(p, \mathbb{R})$  に対して，

$$(\mathbf{x} | g_a(\mathbf{y})) = ({}^t g_a(\mathbf{x}) | \mathbf{y})$$

で定める．実際，  $n > (p-1)/2$  の場合のみを考えると，任意の  $\xi, \eta \in \text{Sym}(p, \mathbb{R})$  に対して，

$$\begin{aligned} (\xi | g_a \circ V_{\mathcal{F}_n}(g_a^{-1}(\mathbf{m})) \circ {}^t g_a(\eta)) &= ({}^t g_a(\xi) | V_{\mathcal{F}_n}(g_a^{-1}(\mathbf{m})) \circ {}^t g_a(\eta)) \\ &= \int_{\text{Sym}(p, \mathbb{R})} ({}^t g_a(\xi) | \mathbf{x} - g_a^{-1}(\mathbf{m})) \times ({}^t g_a(\eta) | \mathbf{x} - g_a^{-1}(\mathbf{m})) W(g_a^{-1}(\mathbf{m}), \mu_n)(dx) \\ &= \int_{\text{Sym}(p, \mathbb{R})} (\xi | g_a(\mathbf{x}) - \mathbf{m}) \times (\eta | g_a(\mathbf{x}) - \mathbf{m}) W(g_a^{-1}(\mathbf{m}), \mu_n)(dx) \\ &= \int_{\text{Sym}(p, \mathbb{R})} (\xi | \mathbf{x} - \mathbf{m}) \times (\eta | \mathbf{x} - \mathbf{m}) W(g_a^{-1}(\mathbf{m}), g_a \mu_n)(dx) \\ &= \int_{\text{Sym}(p, \mathbb{R})} (\xi | \mathbf{x} - \mathbf{m}) \times (\eta | \mathbf{x} - \mathbf{m}) g_a W(\mathbf{m}, \mu_n)(dx) \\ &= (\xi | V_{g_a(\mathcal{F}_n)}(\mathbf{m})(\eta)) \end{aligned}$$

となることよりわかる．また，明らかに

$$g_a^{-1}(\mathbf{m}) = (\mathbf{a}')^{-1} \mathbf{m} \mathbf{a}^{-1} = g_{\mathbf{a}^{-1}}(\mathbf{m})$$

である．さらに，  $\theta, \mathbf{x} \in \text{Sym}(p, \mathbb{R})$  に対して，

$$({}^t g_a(\theta) | \mathbf{x}) = (\theta | g_a(\mathbf{x})) = (\theta | \mathbf{a}' \mathbf{x} \mathbf{a}) = \text{Tr}(\theta \mathbf{a}' \mathbf{x} \mathbf{a}) = (\mathbf{a} \theta \mathbf{a}' | \mathbf{x}) = (g_{\mathbf{a}'}(\theta) | \mathbf{x})$$

より

$${}^t g_a = g_{\mathbf{a}'} \tag{2.4}$$

がわかる．定理 2.2 より

$$\begin{aligned} [g_a \circ V_{\mathcal{F}_n}(g_a^{-1}(\mathbf{m})) \circ {}^t g_a](\theta) &= g_a \circ V_{\mathcal{F}_n}(g_a^{-1}(\mathbf{m})) g_{\mathbf{a}'}(\theta) \\ &= g_a \circ V_{\mathcal{F}_n}(g_a^{-1}(\mathbf{m})) (\mathbf{a} \theta \mathbf{a}') \\ &= g_a \left( \frac{1}{n} g_a^{-1}(\mathbf{m}) \mathbf{a} \theta \mathbf{a}' g_a^{-1}(\mathbf{m}) \right) \\ &= \frac{1}{n} g_a \left( (\mathbf{a}')^{-1} \mathbf{m} \mathbf{a}^{-1} \mathbf{a} \theta \mathbf{a}' (\mathbf{a}')^{-1} \mathbf{m} \mathbf{a}^{-1} \right) \\ &= \frac{1}{n} g_a \left( (\mathbf{a}')^{-1} \mathbf{m} \theta \mathbf{m} \mathbf{a}^{-1} \right) \\ &= \frac{1}{n} \mathbf{m} \theta \mathbf{m} = V_{\mathcal{F}_n}(\mathbf{m})(\theta) \end{aligned}$$

となる．よって，

$$V_{g_a(\mathcal{F}_n)} = V_{\mathcal{F}_n}$$

がわかり，  $g_a(\mathcal{F}_n) = \mathcal{F}_n$  がいえた．

□

命題 2.2  $\text{Sym}(p, \mathbb{R})$ -値確率行列  $W$  が  $W(\psi_{\mu_n}(\mathbf{m}), \mu_n)(dx)$  に従うとする . このとき ,  $g_a(W)$  は  $W(\psi_{\mu_n}(g_a(\mathbf{m})), \mu_n)(dx)$  に従う .

証明 :  $\boldsymbol{\theta} \in -\text{Sym}^+(p, \mathbb{R})$  に対して ,

$$\begin{aligned} \mathbb{E} \exp(\boldsymbol{\theta} | \mathbf{a}' \mathbf{W} \mathbf{a}) &= \int_{\text{Sym}(p, \mathbb{R})} \exp(\boldsymbol{\theta} | \mathbf{a}' \mathbf{x} \mathbf{a}) \exp\{(\psi_{\mu_n}(\mathbf{m}) | \mathbf{x}) - k_{\mu_n}(\psi_{\mu_n}(\mathbf{m}))\} \mu_n(dx) \\ &= \int_{\text{Sym}(p, \mathbb{R})} \exp\{({}^t g_a(\boldsymbol{\theta}) + \psi_{\mu_n}(\mathbf{m}) | \mathbf{x}) - k_{\mu_n}(\psi_{\mu_n}(\mathbf{m}))\} \mu_n(dx) \\ &= \exp\{k_{\mu_n}(\boldsymbol{\theta} + ({}^t g_a)^{-1} \psi_{\mu_n}(\mathbf{m})) - k_{\mu_n}(({}^t g_a)^{-1} \psi_{\mu_n}(\mathbf{m}))\} \\ &= \exp\{k_{\mu_n}(\boldsymbol{\theta} + \psi_{\mu_n}(g_a(\mathbf{m}))) - k_{\mu_n}(\psi_{\mu_n}(g_a(\mathbf{m})))\} \end{aligned}$$

がわかる . ただし , 最後から 2 番目の等号は  $k_{\mu_n}(\boldsymbol{\theta}) = -n \log \text{Det}(-2\boldsymbol{\theta})$  から

$$\begin{aligned} &k_{\mu_n}({}^t g_a(\boldsymbol{\theta}) + \psi_{\mu_n}(\mathbf{m})) - k_{\mu_n}(\psi_{\mu_n}(\mathbf{m})) \\ &= -n \log \text{Det}({}^t g_a(\boldsymbol{\theta} + \psi_{\mu_n}(\mathbf{m})) + n \log \text{Det}(\psi_{\mu_n}(\mathbf{m})) \\ &= -n \log \text{Det}(g_a' \boldsymbol{\theta} + \psi_{\mu_n}(\mathbf{m})) + n \log \text{Det}(\psi_{\mu_n}(\mathbf{m})) \\ &= -n \log \text{Det}((\mathbf{a} \boldsymbol{\theta} \mathbf{a}' + \psi_{\mu_n}(\mathbf{m}))(\psi_{\mu_n}(\mathbf{m}))^{-1}) \\ &= -n \log \text{Det}((\boldsymbol{\theta} + \mathbf{a}^{-1} \psi_{\mu_n}(\mathbf{m})(\mathbf{a}')^{-1})(\mathbf{a}^{-1} \psi_{\mu_n}(\mathbf{m})(\mathbf{a}')^{-1})^{-1}) \\ &= -n \log \text{Det}((\boldsymbol{\theta} + g_a^{-1} \psi_{\mu_n}(\mathbf{m}))(g_a^{-1} \psi_{\mu_n}(\mathbf{m}))^{-1}) \\ &k_{\mu_n}(\boldsymbol{\theta} + ({}^t g_a)^{-1} \psi_{\mu_n}(\mathbf{m})) - k_{\mu_n}(({}^t g_a)^{-1} \psi_{\mu_n}(\mathbf{m})) \end{aligned}$$

となることからわかる . また , 最後の等号は , (2.3) と (2.4) から

$$\begin{aligned} ({}^t g_a)^{-1} \psi_{\mu_n}(\mathbf{m}) &= g_a^{-1} \psi_{\mu_n}(\mathbf{m}) = \mathbf{a}^{-1} \psi_{\mu_n}(\mathbf{m})(\mathbf{a}')^{-1} = -n \mathbf{a}^{-1} \mathbf{m}^{-1} \mathbf{a}'^{-1} \\ &= -n(\mathbf{a}' \mathbf{m} \mathbf{a})^{-1} = -n(g_a \mathbf{m})^{-1} = \psi_{\mu_n}(g_a(\mathbf{m})) \end{aligned}$$

となることよりわかる . □

# 第3章

## 3.1 統計的決定問題

統計的決定問題を構成する要素はつぎの5つである：

- (i) 標本空間  $(\mathcal{X}, \mathcal{B}_1)$ ,
- (ii) 母数空間  $(\Theta, \mathcal{B}_2)$ ,
- (iii) 行動空間  $(\mathcal{A}, \mathcal{B}_3)$
- (iv) 統計モデル  $\{P_\theta : \theta \in \Theta\}$  . すなわち, 標本空間  $(\mathcal{X}, \mathcal{B}_1)$  上の確率測度の族,
- (v) 損失関数  $L : \Theta \times \mathcal{A} \rightarrow [0, \infty)$  . ただし,  $L$  は可測関数である .

標本  $x \in \mathcal{X}$  に対して, 行動空間  $(\mathcal{A}, \mathcal{B}_3)$  上の確率測度を対応させる関数  $\delta : x \mapsto \delta(\cdot | x)$  を統計的決定関数という . 各  $x \in \mathcal{X}$  に対して, ある元  $a(x) \in \mathcal{A}$  が存在して,  $\delta(\{a(x)\} | x) = 1$  (すなわち,  $a(x)$  に退化) のとき, 決定関数  $\delta$  は非確率化決定関数という .

与えられた決定関数  $\delta$  に対して,  $\delta$  の危険関数  $R(\theta, \delta)$  を

$$R(\theta, \delta) = \int_{\mathcal{X}} \int_{\mathcal{A}} L(\theta, a) \delta(da | x) P_\theta(dx)$$

で定める . 非確率化決定関数  $\delta$  ( $\delta(\{a(x)\} | x) = 1$ ) に対する危険関数は

$$R(\theta, \delta) = \int_{\mathcal{X}} L(\theta, a(x)) P_\theta(dx)$$

と簡単に書ける .

**決定関数のよさ** 決定関数全体の集合を  $\mathcal{D}$  と書くことにする .  $\delta_1, \delta_2 \in \mathcal{D}$  とする .  $\delta_1$  が  $\delta_2$  と同程度によいとは, すべての  $\theta \in \Theta$  に対して,  $R(\theta, \delta_1) \leq R(\theta, \delta_2)$  が成り立つときにいう .  $\delta_1$  が  $\delta_2$  よりよいとは,  $\delta_1$  は  $\delta_2$  と同程度によく, さらに, すくなくとも一つの  $\theta_0 \in \Theta$  が存在して,  $R(\theta_0, \delta_1) < R(\theta_0, \delta_2)$  が成り立つときにいう .

$\delta_1$  よりよい  $\delta \in \mathcal{D}$  が存在しないとき,  $\delta_1$  は許容的であるという .  $D \subset \mathcal{D}$  を部分集合とする . 任意の  $\delta \in \mathcal{D} \setminus D$  に対して, ある  $\delta' \in D$  が存在して,  $\delta'$  が  $\delta$  よりよい (同程度によい) とき,  $D$  は (本質的) 完備類という .  $D$  が (本質的) 完備類であって,  $D$  のいかなる真部分集合も (本質的) 完備類にならないとき,  $D$  は最小 (本質的) 完備類という .

命題 3.1 行動空間  $\mathcal{A}$  は  $\mathbb{R}^k$  の凸集合とし, 損失関数  $L(\theta, a)$  は  $a$  の凸関数で,  $\|a\| \rightarrow \infty$  のとき,  $L(\theta, a) \rightarrow \infty$  とする . このとき, 非確率化決定関数の全体の集合は本質的完備類である .

証明：信じることにする。 □

命題 3.2  $t$  を十分統計量とする。  $\mathcal{X}$  または  $\mathcal{A}$  は  $\mathbb{R}^k$  の部分集合とする。  $t$  に基づく決定関数全体の集合は本質的完備類である。

証明：信じることにする。 □

**ベイズ解**  $\xi$  を  $(\Theta, \mathcal{B}_2)$  上の確率測度とする。これを事前確率とよぶことにする。  $\delta_1$  が事前確率  $\xi$  に対するベイズ解であるとは、

$$\int_{\Theta} \mathbf{R}(\theta, \delta_1) \xi(d\theta) = \min_{\delta \in \mathcal{D}} \int_{\Theta} \mathbf{R}(\theta, \delta) \xi(d\theta)$$

をみたすときにいう。  $\delta_1$  が広義ベイズ解であるとは、  $(\Theta, \mathcal{B}_2)$  上の確率測度列  $\{\xi_n\}_{n=1}^{\infty}$  が存在して、

$$\int_{\Theta} \mathbf{R}(\theta, \delta_1) \xi_n(d\theta) - \min_{\delta \in \mathcal{D}} \int_{\Theta} \mathbf{R}(\theta, \delta) \xi_n(d\theta) \rightarrow 0, \quad (n \rightarrow \infty)$$

がなりたつときにいう。

**ミニマックス決定関数**  $\delta_1$  がミニマックス決定関数であるは、

$$\max_{\theta \in \Theta} \mathbf{R}(\theta, \delta_1) = \min_{\delta \in \mathcal{D}} \max_{\theta \in \Theta} \mathbf{R}(\theta, \delta)$$

が成り立つときにいう。

**不変性**  $(\mathcal{X}, G, \varphi)$ ,  $(\Theta, \bar{G}, \bar{\varphi})$ ,  $(\mathcal{A}, \tilde{G}, \tilde{\varphi})$  は変換群とする。以後、誤解のおそれがなければ、  $G, \bar{G}, \tilde{G}$  および  $\varphi, \bar{\varphi}, \tilde{\varphi}$  をすべて  $G, \varphi$  とかくことにする。決定問題が  $G$ -不変であるとは、つぎの条件をみたすときをいう。

(i) モデル  $\{P_{\theta} : \theta \in \Theta\}$  は不変。すなわち、

$$gP_{\theta} = P_{g\theta}, \quad (g \in G, \theta \in \Theta)$$

ただし、任意の  $B \in \mathcal{B}_1$  に対して、  $gP_{\theta}(B) = P_{\theta}(g^{-1}B)$  である。

(ii) 損失関数  $L$  は不変。すなわち、

$$L(g\theta, ta) = L(\theta, a), \quad (g \in \Theta, \theta \in \Theta, a \in \mathcal{A}).$$

決定関数  $\delta \in \mathcal{D}$  と  $g \in G$  に対して、決定関数  $g\delta$  を

$$(g\delta)(B|x) = \delta(g^{-1}B|g^{-1}x), \quad (B \in \mathcal{B}_3)$$

で定める。

決定関数  $\delta$  が  $G$ -不変であるとは、すべての  $g \in G$  に対して、  $g\delta = \delta$  が成り立つときにいう。  $\delta$  が非確率化検定 ( $\delta(\{a(x)\}|x) = 1$ ) のとき、

$$\delta \text{ が不変} \iff a \text{ が不変} \iff a(gx) = ga(x), \quad (g \in G)$$

となる。

命題 3.3 決定問題が  $G$ -不変とする．任意の決定関数  $\delta$  に対して，

$$\mathbf{R}(g\theta, g\delta) = \mathbf{R}(\theta, \delta), \quad (g \in G, \theta \in \Theta)$$

さらに，決定関数  $\delta$  が  $G$ -不変ならば，

$$\mathbf{R}(g\theta, \delta) = \mathbf{R}(\theta, \delta), \quad (g \in G)$$

がわかる．

証明 モデルの不変性から，任意の可積分関数  $h$  に対して，

$$\int_{\mathcal{X}} f(gx) P_{\theta}(dx) = \int_{\mathcal{X}} f(x) P_{g\theta}(dx) \quad (3.1)$$

となる．また，任意の決定関数  $\delta$  と可積分関数  $k$  に対して，

$$\int_{\mathcal{A}} k(a) (g\delta)(da|x) = \int_{\mathcal{A}} k(ga) \delta(da|g^{-1}x) \quad (3.2)$$

信じることにする．これらのことより

$$\begin{aligned} \mathbf{R}(g\delta, \theta) &= \int_{\mathcal{X}} \int_{\mathcal{A}} \mathbf{L}(\theta, a) (g\delta)(da|x) P_{\theta}(dx) \\ &= \int_{\mathcal{X}} \int_{\mathcal{A}} \mathbf{L}(ga, \theta) (g\delta)(da|g^{-1}x) P_{\theta}(dx) \\ &= \int_{\mathcal{X}} \int_{\mathcal{A}} \mathbf{L}(a, g^{-1}\theta) (g\delta)(da|g^{-1}x) P_{\theta}(dx) \\ &= \int_{\mathcal{X}} \int_{\mathcal{A}} \mathbf{L}(a, g^{-1}\theta) (g\delta)(da|x) P_{g^{-1}\theta}(dx) = \mathbf{R}(g^{-1}\theta, \delta) \end{aligned}$$

となる．2 番目の等号は (3.2)，3 番目の等号は損失関数の不変性，最後から 2 番目の等号は (3.1) からそれぞれわかる．  $\square$

$G$  が  $\mathcal{X}$  に推移的に作用していれば，固定した元  $\theta_0 \in G$  に対して，

$$\mathbf{R}(\theta, \delta) = \mathbf{R}(\theta, \theta_0)$$

となる．したがって，固定した  $\theta_0$  のもとで最良のものは  $G$ -不変のなかで最良となる．

### 最良 $G$ -不変性とミニマックス性

命題 3.4  $G$  は三角群とし，決定問題は  $G$ -不変とする．このとき，もし，ミニマックス決定関数が存在するならば， $G$ -不変ミニマックス決定関数が存在する．もし， $G$ -不変決定関数  $\delta_1$  が

$$\max_{\theta \in \Theta} \mathbf{R}(\theta, \delta_1) = \min_{\delta \in \mathcal{D}^*} \max_{\theta \in \Theta} \mathbf{R}(\theta, \delta)$$

をみたすならば， $\delta_1$  はミニマックスである．ただし， $\mathcal{D}^*$  は  $G$ -不変決定関数の全体である．

証明ではない注意 この命題は Hunt-Stein の定理と呼ばれるものを簡単にしたものである．Hunt-Stein の定理が成立するための条件として，Amenability 条件( Lehmann and Casella (1998, page 422) を参照 )が知られている．三角群はこの条件をみたしていることが知られている． $\square$

## 3.2 共分散行列の推定問題

$p$ 次元確率変数列  $X_1, X_2, \dots, X_n$  ( $n \geq p$ ) は独立同一に多変量正規分布  $N_p(0, \Sigma)$  に従うと仮定する。ただし,  $\Sigma \in \text{Sym}^+(p, \mathbb{R})$  とする。このとき,  $(X_1, X_2, \dots, X_n)$  に基づき,  $\Sigma$  の推定問題を損失関数

$$\mathbf{L}(\Sigma, a) = \text{Tr}(a\Sigma^{-1}) - \log \text{Det}(a\Sigma^{-1}) - p$$

のもとで考える。ただし,  $a$  は  $\Sigma$  の推定量で  $S = \sum_{i=1}^n X_i X_i'$  を通してのみ  $(X_1, X_2, \dots, X_n)$  に依存する推定量である。これは  $S$  は  $\Sigma$  の十分統計量であることと命題 3.2 から正当化される。この損失関数  $\mathbf{L}$  は James and Stein (1961) で提案された。この損失関数は非常に取り扱いが容易である。これは, つぎの性質をみtas。

- (i)  $\mathbf{L}(\Sigma, a) \geq 0$  で等号成立は  $a = \Sigma$  のときのみ。
- (ii)  $\mathbf{L}(\Sigma, a)$  は第 2 の変数に関して凸関数。
- (iii) 任意の  $g \in \text{GL}(p, \mathbb{R})$  に対して,

$$\mathbf{L}(g\Sigma g', gag') = \mathbf{L}(\Sigma, a)$$

をみtas。

損失関数の性質と命題 3.1 から非確率化推定量のみを考えればよいことがわかる。

損失関数  $\mathbf{L}$  に対応する危険関数  $\mathbf{R}$  を

$$\mathbf{R}(\Sigma, a) = \mathbb{E}[\mathbf{L}(\Sigma, a)]$$

で定義する。

### 3.2.1 定数リスクのミニマックス推定量

自然指数分布族の性質から,  $\Sigma$  の最尤推定量は  $\hat{\Sigma}_{\text{mle}} = n^{-1}S$  となるのが直ちにわかる。

$\mathcal{X} = \text{Sym}^+(p, \mathbb{R})$ ,  $G = \text{GL}(p, \mathbb{R})$ ,  $\varphi(g, x) = gxg'$  とすれば,  $(\mathcal{X}, G, \varphi)$  は変換群となる。群  $G$  は  $\mathcal{X}$  に推移的に作用するので, 最良  $G$ -不変推定量を求めることができる。すなわち,  $\hat{\Sigma}$  を  $G$ -不変とすれば, 命題 3.3 から

$$\mathbf{R}(\Sigma, \hat{\Sigma}) = \mathbf{R}(I_p, \hat{\Sigma})$$

となる。さらに, 任意の  $G$ -不変推定量は  $\hat{\Sigma} = cS$  と表現できることに注意する。ただし,  $c$  は正の定数である。なぜならば,  $g = S^{-1/2}$  ととれば,

$$\hat{\Sigma}(S) = S^{-1/2} \hat{\Sigma}(I_p) S^{-1/2}$$

と表現できる。さらに,  $g \in \mathcal{P}_p \subset \text{GL}(p, \mathbb{R})$  (置換群) に対して,  $\hat{\Sigma}(I_p) = \hat{\Sigma}(gI_p g') = g \hat{\Sigma}(I_p) g'$  となることから,  $\hat{\Sigma}(I_p) = cI_p$  と書けることがわかる。

命題 3.5  $G$ -不変推定量の中で最小のものは

$$\widehat{\Sigma}_{\text{mle}} = \frac{1}{n}S$$

である .

証明：損失関数と推定量の族の  $G$ -不変性と命題 3.3 から  $S \sim W_p(I_p, n)$  としてよい . しかし ,

$$\mathbf{R}(I_p, cS) = \mathbb{E}[c\text{Tr}(S) - p \log c - \log \text{Det } S - p]$$

を  $c$  に関して微分すれば ,

$$c = \frac{p}{\mathbb{E}[\text{Tr}(S)]} = \frac{p}{np}$$

よりわかる . 損失関数の凸性から命題はわかる .  $\square$

$\Sigma_{\text{mle}}$  は 最良  $G$ -不変推定量 である . また ,  $\Sigma_{\text{mle}}$  は  $\text{GT}^+(p, \mathbb{R})$ -不変でもあることに注目する .  $\text{GT}^+(p, \mathbb{R})$  も  $\text{Sym}^+(p, \mathbb{R})$  に推移的に作用するので , 命題 3.3 から , このクラスに属する推定量のリスクも  $\Sigma = I_p$  のもとで計算すればよい . これから ,  $\text{GT}^+(p, \mathbb{R})$ -不変で一番よいものを求めることができる .

定理 3.1 (i)  $S = TT'$  と分解する . ここで ,  $T \in \text{GT}^+(p, \mathbb{R})$  である .  $\text{GT}^+(p, \mathbb{R})$ -不変推定量は

$$TAT'$$

の形で与えられる . ただし ,  $A \in \text{Sym}(p, \mathbb{R})$  は定数である .

(ii)  $\text{GT}^+(p, \mathbb{R})$ -不変推定量のリスクの下限はつぎで与えられる :

$$\mathbf{R}(I_p, TAT') \geq \sum_{i=1}^p \{ \log(n+p-2i+1) - \mathbb{E}[\log \chi_{n-i+1}^2] \}$$

(iii) 推定量  $\Sigma_{\min} = T \text{diag}(1/(n+p-2i+1) | i=1, 2, \dots, p) T'$  は (ii) の下限に達する . したがって , 命題 3.4 から ,  $\Sigma_{\min}$  はミニマックスである .

証明 :  $\text{GT}^+(p, \mathbb{R})$ -不変性から

$$\widehat{\Sigma}(T^{-1}S(T')^{-1}) = T^{-1}\widehat{\Sigma}(I_p)(T')^{-1}$$

から (i) はわかる .

つぎに , 命題 3.4 と (i) から ,  $\text{GT}^+(p, \mathbb{R})$ -不変推定量  $\widehat{\Sigma}$  に対して ,

$$\mathbf{R}(\Sigma, \widehat{\Sigma}) = \mathbf{R}(I_p, \widehat{\Sigma}) = \text{Tr} \{ \mathbb{E}(T'T)A \} - \log \text{Det}(A) - \sum_{i=1}^p \mathbb{E}[\log t_{ii}^2] - p$$

となる . しかし , 命題 1.8 から  $t_{ii}^2 \sim \chi_{n-i+1}^2$  と  $\mathbb{E}[T'T] = \text{diag}(n+p-2i+1 | i=1, 2, \dots, p)$  となる . さらに , 命題 A.1 に注意して ,  $A$  の各成分に関して微分すれば ,

$$\text{diag}(n+p-2i+1 | i=1, 2, \dots, p) - A^{-1} = 0$$

を得る . よって , 損失関数の凸性から (ii) がわかる . (iii) は (ii) の証明をたどればよい .  $\square$

### 3.2.2 直交不変な推定量と SURE 法

$\Sigma$  の直交不変な推定量を考える :

$$\widehat{\Sigma}(HSH') = H\widehat{\Sigma}(S)H', \quad (H \in O(p, \mathbb{R}))$$

このクラスの推定量は

$$\widehat{\Sigma} = H \text{diag}(\phi_i | i = 1, 2, \dots, p) H' \quad (3.3)$$

と書ける . ただし ,  $S = H \text{diag}(\ell_1, \dots, \ell_p) H'$  で  $l = (\ell_1, \dots, \ell_p)$  で  $\phi_i(l)$  は微分可能とする . これは ,  $H = \text{diag}(\pm 1, \dots, \pm 1)$  とすればわかる .  $\phi_i = (1/n)\ell_i$  とすれば ,  $\widehat{\Sigma}_{\text{mle}}$  となることがわかる .

**動機**  $p$  が大きく , さらに ,  $\Sigma = I_p$  としたときには , これはよくないことが Johansson (1982) で知られていることから予想される :  $0 < y < 1$  とする .  $p/n \rightarrow y (n \rightarrow \infty)$  のとき ,

$$\frac{1}{p} \sum_{i=1}^p \mathbb{I}\{\ell_i/p \leq x\} \xrightarrow{P} \int_a^x \frac{1}{u} \frac{\sqrt{(u-a)(b-u)}}{2\pi y u} du \mathbb{I}_{[a,b]}(x) + \mathbb{I}_{(b,\infty)}(x)$$

となる . ただし , 各  $x$  における  $\xrightarrow{P}$  は確率収束を示し ,

$$a = 1 + y - 2\sqrt{y} \quad b = 1 + y + 2\sqrt{y}$$

である .

したがって ,  $i/p$  と  $1 - i/p$  が小さい  $i \in \{1, 2, \dots, p\}$  に対しては ,  $\ell_i/n$  は 1 に高い確率で離れていることがわかる . よって ,  $n$  と  $p$  が近く ,  $\Sigma = I_p$  のとき ,  $1 - i/p$  が小さいもの ( $i$  が  $p$  に近い) に対しては ,  $\ell_i/n$  を縮小し ,  $i/p$  が小さいもの ( $i$  が 1 に近い) に対しては ,  $\ell_i/n$  を拡大するのがよいことが予想される . 問題は , このような修正を行っても  $\Sigma$  に関して一様に  $\widehat{\Sigma}_{\text{mle}}$  よりよいものができるかである . 実際にうまくできることを Stein (1973) で示している .

**SURE 法 : リスクの不偏推定量** 直交群  $O(p, \mathbb{R})$  の  $\text{Sym}^+(p, \mathbb{R})$  への作用は推移的でない . したがって , リスクは一般には定数でない . 直交不変な推定量のリスクを評価するためには , 直接的な評価が困難になる . これを克服する方法が Stein's Unbiased Risk Estimate 法 (SURE 法) である .

まず , 議論を簡単にするために , 推定量のリスクを比較するための本質的な部分を

$$\widetilde{\mathbf{R}}(\Sigma, \widehat{\Sigma}) = \mathbb{E}[\text{Tr}(\widehat{\Sigma}\Sigma^{-1}) - \log \text{Det}(\widehat{\Sigma})] \quad (3.4)$$

とおく . 有限リスクを持つふたつの推定量  $\widehat{\Sigma}_1, \widehat{\Sigma}_2$  に対して ,

$$\mathbf{R}(\Sigma, \widehat{\Sigma}_1) \geq \mathbf{R}(\Sigma, \widehat{\Sigma}_2) \iff \widetilde{\mathbf{R}}(\Sigma, \widehat{\Sigma}_1) \geq \widetilde{\mathbf{R}}(\Sigma, \widehat{\Sigma}_2), \quad (\forall \Sigma \in \text{Sym}^+(p, \mathbb{R}))$$

となることに注意せよ . 以後は , 言葉の乱用となるが ,  $\widetilde{\mathbf{R}}$  のこともリスクと呼ぶことにする .

各推定量のリスクは  $\Sigma$  の関数となっている . リスクの一般的な比較をするために , 直交不変な推定量  $\widehat{\Sigma}$  に対するリスク ( $\Sigma$  の関数) の不偏推定量を求めることを考える . 形は  $\widehat{\Sigma}$  に依存す

るが，ひとたび  $S$  を観測すれば，計算できる量，すなわち，推定量（これを  $\widehat{\mathbf{R}}_{\widehat{\Sigma}}(S)$  と書くことにする）で

$$\widetilde{\mathbf{R}}(\Sigma, \widehat{\Sigma}) = \mathbb{E}[\widehat{\mathbf{R}}_{\widehat{\Sigma}}(S)]$$

をみたくものを構成することを目指す．有限のリスクを持つ任意の直交不変な推定量  $\widehat{\Sigma}$  に対して，幸運にもこのように，リスクの不偏推定量を見つけることができれば，期待値の基本的な性質から

$$\widehat{\mathbf{R}}_{\widehat{\Sigma}_1}(S) \geq \widehat{\mathbf{R}}_{\widehat{\Sigma}_2}(S) \implies \widetilde{\mathbf{R}}(\Sigma, \widehat{\Sigma}_1) \geq \widetilde{\mathbf{R}}(\Sigma, \widehat{\Sigma}_2) \quad (\forall \Sigma \in \text{Sym}^+(p, \mathbb{R}))$$

となる．さらに，ある  $\Sigma$  で

$$\mathbb{P}(\widehat{\mathbf{R}}_{\widehat{\Sigma}_1}(S) > \widehat{\mathbf{R}}_{\widehat{\Sigma}_2}(S)) > 0$$

ならば，

$$\widetilde{\mathbf{R}}(\Sigma, \widehat{\Sigma}_1) > \widetilde{\mathbf{R}}(\Sigma, \widehat{\Sigma}_2) \quad (\exists \Sigma \in \text{Sym}^+(p, \mathbb{R}))$$

がわかるのである．

方針をまとめるとつぎのふたつの段階に分かれる：

- (1) 有限のリスクを持つ任意の直交不変な推定量  $\widehat{\Sigma}$  に対して，そのリスクの不偏推定量  $\widehat{\mathbf{R}}_{\widehat{\Sigma}}(S)$  を構成する．
- (2)  $\widehat{\mathbf{R}}_{\widehat{\Sigma}}(S)$  で推定量の比較を行う．

(1) の段階では，部分積分を用いる．(2) はある種の偏微分不等式をとくことになる．これを解く体系的な手法<sup>(3-1)</sup> は知らない．現段階では，いくつかのパターンを試してみて，解を見つけるにとどまっている．

### 3.2.3 部分積分の公式

部分積分の公式を理解するために，簡単な統計モデルを考える：確率変数  $X$  は連続型分布に従い，確率密度関数

$$P(dx) = \exp\{\theta x - k(\theta)\}w(x)dx$$

を持つとする．ただし， $\theta \in \Theta$  は母数とし， $dx$  は  $\mathbb{R}$  上のルベーグ測度で， $w$  は微分可能，

$$\mathbb{R} = \{x \in \mathbb{R} : \exp\{\theta x - k(\theta)\}w(x) > 0\}$$

とする．さらに，

$$t(x) = -\frac{\dot{w}(x)}{w(x)}$$

とおく．ただし， $\dot{w}$  は  $w$  の導関数とする．

命題 3.6  $g : \mathbb{R} \rightarrow \mathbb{R}$  を可積分かつ可微分で  $\mathbb{E}|g(X)| < \infty$  とする．このとき，

$$\mathbb{E}[g(X)] = \mathbb{E}[(t(X) - \theta)g(X)]$$

が成立する．ただし， $\dot{g}$  は  $g$  の導関数である．

証明:  $\mathbb{E}|\dot{g}(X)| < \infty$  より,  $x \rightarrow \pm\infty$  のとき,

$$g(x) \exp\{\theta x - k(\theta)\}w(x) \rightarrow 0$$

となること<sup>(3-2)</sup>に注意する.

$$\begin{aligned} \mathbb{E}[\dot{g}(X)] &= \int_{\mathbb{R}} \dot{g}(x) \exp\{\theta x - k(\theta)\}w(x) dx \\ &= \int_{\mathbb{R}} g(x) (\exp\{\theta x - k(\theta)\}) (\theta w(x) + \dot{w}(x)) dx \\ &= \mathbb{E}[(t(X) - \theta)g(X)] \end{aligned}$$

となる. □

例 3.1 ガンマ分布を考える:

$$\gamma(dx) = \exp(\theta x - k_{\mu}(\theta))x^{\alpha-1}, dx, \quad (\theta < 0)$$

を考える.

$$t(X) = -\frac{\alpha - 1}{X}$$

となるので,

$$\mathbb{E}[(-\theta)g(X)] = (\alpha - 1)\mathbb{E}[X^{-1}g(X)] + \mathbb{E}[\dot{g}(X)]$$

を得る. □

ここで, もとの設定にもどる.

$$H\Sigma^{-1}H' = (a_{ij}(H))$$

とおき, 推定量 (3.3) を (3.4) に代入すると

$$\tilde{\mathbf{R}}(\Sigma, \hat{\Sigma}) = \mathbb{E} \left[ \sum_{i=1}^p \phi_i(\ell) a_{ii}(H) - \log \phi_i(\ell) \right] \quad (3.5)$$

と書き直せる. したがって, 上の式の右辺の括弧内の第一目のみが母数に依存している. それを部分積分の公式を用いて評価する. ただし, ここで, 損失は  $\ell$  と  $H$  にのみ依存しているので, ウィシャート行列の極分解に対する分布に関して部分積分の公式を使うと非常に見通しよく, 不偏推定量を求めることができる. そのために, 以下の記号を用意する: 命題 1.9 より,  $\ell$  の同時分布は

$$K \exp\left\{-\frac{1}{2} \sum_{j=1}^r \ell_j a_{jj}(H)\right\} (\prod_{j=1}^p \ell_j)^{(n-p-1)/2} \prod_{j>i} (\ell_i - \ell_j) \mu_{O(p, \mathbb{R})} \prod_{j=1}^r d\ell_j.$$

これを踏まえてつぎの記号を用意する:

$$F(\ell) = K (\prod_{j=1}^r \ell_j)^{(n-p-1)/2} \prod_{j>i} (\ell_i - \ell_j)$$

$$G(a) = \int_{\mathcal{X}} \exp\left\{-\frac{1}{2} \sum_{j=1}^p \ell_j a_{ii}(H)\right\} \mu_{O(p)},$$

$$\mathcal{R}_j = \{\ell^{(j-)} = (\ell_1, \dots, \ell_{j-1}, \ell_{j+1}, \dots, \ell_p) \mid \ell_1 > \dots > \ell_{j-1} > \ell_{j+1} > \dots > \ell_p > 0\},$$

とおく. ただし,  $\ell_0 = \infty$  と  $\ell_{p+1} = 0$  した. これらの記号を用いて, つぎの定理を示す.

## 補題 3.1

$$\sum_{j=1}^p \mathbb{E}[\phi_j a_{jj}(H)] = \mathbb{E} \left[ \sum_{j=1}^p \left\{ 2 \frac{\partial \phi_j}{\partial \ell_j} + (n-p-1) \frac{\phi_j}{\ell_j} + 2 \sum_{i \neq j} \frac{\phi_j}{\ell_j - \ell_i} \right\} \right]$$

証明：部分積分から

$$\begin{aligned} \sum_{j=1}^p \mathbb{E}[\phi_j a_{jj}(H)] &= -2 \sum_{j=1}^p \int_{\mathcal{R}_j} \left( \int_{\ell_{j+1}}^{\ell_{j-1}} \varphi_j F(\ell) \frac{\partial}{\partial \ell_j} G(\ell) d\ell_j \right) d\ell^{(j-)} \\ &= 2 \sum_{j=1}^p \int_{\mathcal{R}_j} \int_{\ell_{j+1}}^{\ell_{j-1}} \frac{\partial}{\partial \ell_j} (\phi_j F(\ell)) G(\ell) da_j d\ell^{(j-)} \\ &= 2 \mathbb{E} \left[ \sum_{j=1}^p \frac{\partial}{\partial \ell_j} (\phi_j F(\ell)) \frac{1}{F(\ell)} \right] \\ &= \mathbb{E} \left[ \sum_{j=1}^p \left\{ 2 \frac{\partial \phi_j}{\partial \ell_j} + (n-p-1) \frac{\phi_j}{\ell_j} + 2 \sum_{i \neq j} \frac{\phi_j}{\ell_j - \ell_i} \right\} \right], \end{aligned}$$

となる。

□

補題より，推定量 (3.3) に対して，

$$\tilde{\mathbf{R}}(\Sigma, \hat{\Sigma}) = \mathbb{E} \left[ \sum_{j=1}^p \left\{ 2 \frac{\partial \phi_j}{\partial \ell_j} + (n-p-1) \frac{\phi_j}{\ell_j} + 2 \sum_{i \neq j} \frac{\phi_j}{\ell_j - \ell_i} - \log \phi_j \right\} \right] \quad (3.6)$$

を得る。

## 3.2.4 直交不変ミニマックス推定量

天くだりではあるが，推定量 (3.3) において

$$\phi_j^{(m)} = d_j \ell_j, \quad d_j = \frac{1}{n+p-2j+1} \quad (3.7)$$

としたものを  $\hat{\Sigma}_m$  とする。  $d_1 < \dots < d_p$  で  $(1/p) \sum_{j=1}^p 1/d_j = n$  となっている。

定理 3.2 (3.3) と (3.7) で与えられる推定量  $\hat{\Sigma}_m$  に対して，

$$\mathbf{R}(\Sigma, \hat{\Sigma}_m) \leq - \sum_{j=1}^r \log d_j - \sum_{j=1}^r \mathbb{E}[\log u_j^2],$$

が成立する。ただし，  $u_j$  は  $\chi_{n-j+1}^2$  ( $j = 1, 2, \dots, r$ ) に従う。したがって，  $\hat{\Sigma}_m$  はミニマックス。

証明 まず，

$$\mathbb{E}[\log \text{Det } \hat{\Sigma}_m] = \sum_{j=1}^r \log d_j + \mathbb{E}[\log \text{Det } S] = \sum_{j=1}^r \{ \log d_j + \log u_j^2 \} + \log \text{Det } \Sigma$$

となることに注意する．この式と (3.6) より

$$\begin{aligned}
\mathbf{R}(\Sigma, \widehat{\Sigma}_m) &= \mathbb{E} \left[ 2 \sum_{j=1}^p \sum_{i>j} \frac{d_j \ell_j - d_i \ell_i}{\ell_j - \ell_i} + 2 \sum_{j=1}^p d_j + (n - p - 1) \sum_{j=1}^p d_j \right. \\
&\quad \left. - \sum_{j=1}^p \mathbb{E} \log u_j^2 - \sum_{j=1}^p \log d_j - p \right] \\
&\leq \sum_{j=1}^p (n + p - 2j + 1) d_j - \sum_{j=1}^p \mathbb{E} \log u_j^2 - \sum_{j=1}^p \log d_j - p \\
&= - \sum_{j=1}^p \mathbb{E} \log u_j^2 - \sum_{j=1}^p \log d_j.
\end{aligned}$$

となることからわかる． □

### 3.2.5 Stein の推定量

(3.5) の右辺を  $\phi_i$  に関して形式的に微分すれば，

$$\phi_i = \frac{1}{\mathbb{E}[a_{ii}(H)]}$$

を得る．しかし， $\mathbb{E}[a_{ii}(H)]$  は母数に依存するので，その不偏推定量  $\widehat{a}_{ii}(H)$  を求めて，

$$\phi_i = \frac{1}{\widehat{a}_{ii}}$$

を推定量として考える．

命題 3.7

$$\widehat{a}_{ii} = (n - p - 1) \frac{1}{\ell_i} + 2 \sum_{i \neq j} \frac{1}{\ell_i - \ell_j}$$

のとき，

$$\mathbb{E}[\widehat{a}_{ii}] = \mathbb{E}[a_{ii}(H)]$$

が成立する．

証明：補題 3.1 と同じようにすれば，

$$\mathbb{E}[a_{ii}(H)] = \mathbb{E} \left[ 2 \frac{\partial}{\partial \ell_i} \log F(\ell) \right]$$

となることからわかる．Sheena (1989) を参照． □

推定量 (3.3) において

$$\phi_j^{(s)} = \frac{\ell_j}{(n - p - 1) + \frac{2}{\ell_i} \sum_{i \neq j} \frac{1}{\ell_i - \ell_j}} \quad (3.8)$$

としたものを  $\hat{\Sigma}_s$  とする .

推定量 (3.7) と (3.8) と同

$$\mathbb{P}\left(\phi_1^{(m)} \geq \cdots \geq \phi_p^{(m)}\right) < 1, \quad \mathbb{P}\left(\phi_1^{(s)} \geq \cdots \geq \phi_p^{(s)} > 0\right) < 1$$

である . 推定値としては ,

$$\phi_1^{(m)} \geq \cdots \geq \phi_p^{(m)}$$

でないときには , これらの値をプールして修正し , 望ましい順序をみたすようにする必要がある . このアルゴリズムについては Lin and Perlman (1985) を参照のこと . また , 数値実験の結果 ( Lin and Perlman (1985) を参照 ) から ,  $\hat{\Sigma}_s$  はミニマックスではないと思われる .



# 第4章

## 4.1 ショルダン代数と対称錐体

**Euclidean Jordan 代数** Euclidean Jordan 代数 (EJA と記す) とは, Euclid 空間  $\mathcal{V}$  でジョルダン積 (単位元を  $e$  と書く) と呼ばれる双一次形式

$$(a, b) \mapsto ab, \quad \mathcal{V} \times \mathcal{V} \rightarrow \mathcal{V}$$

でつぎの性質をみたすものを持つときをいう:

- (i) 任意の  $a, b \in \mathcal{V}$  に対して,  $ab = ba$ .
- (ii) 任意の  $a, b \in \mathcal{V}$  に対して,  $a(a^2b) = a^2(ab)$ . ただし,  $aa = a^2$  と書いた.
- (iii) 任意の  $a, b, c \in \mathcal{V}$  に対して,  $(a|bc) = (ab|c)$ . ただし, 内積を  $(a|b)$  は内積である.

EJA  $\mathcal{V}$  が正の次元の二つの EJA  $\mathcal{V}_1$  と  $\mathcal{V}_2$  の集合の直積のとき,  $\mathcal{V}$  は非単純といい, その他のとき,  $\mathcal{V}$  は単純という.

以後では,  $\mathcal{V}$  は有限次元かつ単純な EJA とする.

例 4.1 (i)  $\text{Sym}(p, \mathbb{R})$  に新たな積  $\circ$  を

$$A \circ B = \frac{1}{2}(AB + BA), \quad (\forall A, B \in \text{Sym}(p, \mathbb{R}))$$

で定義する. ただし,  $AB$  は行列の通常積である. このとき, 積  $\circ$  をもつ  $\text{Sym}(p, \mathbb{R})$  は EJA となる. さらに,

$$ABA = 2A \circ (A \circ B) - A^2 \circ B$$

となる.

- (ii)  $\text{Herm}(p, \mathbb{C})$  も同じ.
- (iii)  $\text{Herm}(p, \mathbb{H})$  も同じ.
- (iv) 3 次 Cayley-Hermite 行列全体.
- (v)  $\mathcal{W}$  は内積空間とする. その内積を  $B(w_1|w_2)$  ( $w_1, w_2 \in \mathcal{W}$ ) と書き,  $\mathcal{V} = \mathbb{R} \times \mathcal{W}$  とする.  $\mathcal{V}$  に積を

$$(\lambda_1, w_1)(\lambda_2, w_2) = (\lambda_1\lambda_2 + B(w_1|w_2), \lambda_1w_2 + \lambda_2w_1), \quad (\forall (\lambda_1, w_1), (\lambda_2, w_2) \in \mathcal{V})$$

で定義する.  $\mathcal{V}$  は EJA となる. □

EJA についての重要事項
---------------

(i)  $\mathcal{V}$  は非結合的 .

(ii) べきに関して指数法則が成立する . さらに ,  $\mathcal{V}$  上の  $j$  次の斉次多項式関数  $a_j (1 \leq j \leq r)$  が存在して ,

$$m_x(t) = t^r - a_1(x)t^{r-1} + a_2(x)t^{r-2} + \cdots + (-1)^r a_r(x)$$

が ,  $\mathcal{V}$  のある開集合に属する任意の  $x$  に対して ,  $x$  の最小多項式になることが知られている . 以下では ,

$$\text{tr}(x) = a_1(x), \quad \det(x) = a_r(x)$$

とおき ,  $x$  の EJA に意味の trace と determinant というこにする .

(iii) 左変換と 2 次表現 . 各  $x \in \mathcal{V}$  に対して , 変換  $L$  と  $P$  を

$$\begin{aligned} L(x) &: y \mapsto xy, & \mathcal{V} &\rightarrow \mathcal{V}, \\ P(x) &: y \mapsto 2x(xy) - x^2y, & \mathcal{V} &\rightarrow \mathcal{V} \end{aligned}$$

で定める . したがって ,

$$P(x) = 2L(x)^2 - L(x^2)$$

となる .

(iv) 逆元について . 各  $x \in \mathcal{V}$  に対して ,  $\mathbb{R}[x]$  を  $x$  と  $e$  で生成される部分環とする .  $x \in \mathcal{V}$  が可逆であるとは , 一意的に  $y \in \mathbb{R}[x]$  が存在して ,  $xy = e$  となるときをいう .

(v) 可逆性の条件 .

$$L(x) \text{ が可逆} \implies x \text{ は可逆} . \text{ このとき , } x^{-1} = L(x)^{-1}e$$

$$P(x) \text{ が可逆} \iff x \text{ は可逆} . \text{ このとき , } x^{-1} = P(x)^{-1}x, P(x)^{-1} = P(x^{-1})$$

このことより  $\mathcal{V}$  の可逆元の集合を

$$\mathcal{V}^\times = \{x \in \mathcal{V} : \text{Det } P(x) \neq 0\}$$

と書くことにする .

(vi) べき等元 . 0 でない元  $c \in \mathcal{V}$  がべき等であるとは ,  $c^2 = c$  をみたすときをいう . 0 でないべき等元  $c$  が二つのべき等元の和で表現できないとき , 単純という . 0 でないべき等元  $c_1, c_2$  に対して ,  $c_1c_2 = 0$  のとき ,  $c_1$  と  $c_2$  は直交であるという . べき等元の系  $c_1, \dots, c_r$  が完全直交系であるとは ,  $c_ic_j = 0 (i \neq j)$  かつ  $\sum_{i=1}^r c_i = e$  が成り立つときをいう . 単純なべき等元の完全直交系をジュールダン枠という .

(vii) スペクトル分解 .  $\mathcal{V}$  のランクを  $r$  とする .  $\mathcal{V}$  の各元  $x$  に対して , 0 でないべき等元のジュールダン枠  $c_1, \dots, c_r$  と実数  $\lambda_1, \dots, \lambda_r$  が一意的に存在して ,

$$x = \lambda_1c_1 + \cdots + \lambda_rc_r$$

と表される .

(viii) EJA の意味の trace と determinant .  $x = \lambda_1c_1 + \cdots + \lambda_rc_r$  に対して ,

$$\text{tr}(x) = \sum_{i=1}^r \lambda_i, \quad \det(x) = \prod_{i=1}^r \lambda_i$$

したがって,

$$x \text{ が可逆} \iff \det(x) \neq 0$$

となる.

(ix) 内積. 一般性を失わず, 内積を

$$(x|y) = \text{tr}(xy), \quad (\forall x, y \in \mathcal{V})$$

と入れてよい.

(x) Peirce 分解. 固定したジョルダン枠  $c_1, c_2, \dots, c_r$  に対して,

$$\mathcal{V}_{ii} = \{x \in \mathcal{V} | c_i x = x\} \quad \text{と} \quad \mathcal{V}_{ij} = \{x \in \mathcal{V} | c_i x = (1/2)x \text{ and } c_j x = (1/2)x\} \quad (4.1)$$

とおく. ただし,  $i = 1, 2, \dots, r$  と  $j = i + 1, \dots, r$  である. このとき, 各  $\mathcal{V}_{ij}$ 's ( $i < j$ ) の次元 ( $2g$  と書く) は同じであることが知られている. その次元  $2g$  を Peirce invariant という. さらに,  $\mathcal{V}$  は Peirce 分解

$$\mathcal{V} = \bigoplus_{1 \leq i < j \leq r} \mathcal{V}_{ij}$$

をもつことが知られている.  $v$  を  $\mathcal{V}$  の次元としたとき, これらから

$$v = r + gr(r-1), \quad (4.2)$$

が成立することがわかる.

(xi) Frobenius 変換. べき等元  $c$  と元  $z \in \{x \in \mathcal{V} | cx = (1/2)x\}$  に対して, Frobenius 変換を

$$\tau_c(z) = \exp\{L(z) + 2L(z)L(c) - 2L(c)L(z)\}$$

で定める. ここで, 任意の写像  $A$  に対して,  $\exp(A) = \sum_{j=0}^{\infty} A^j/j!$  とした.

**対称錐体** EJA  $\mathcal{V}$  に対応する部分集合  $\mathcal{V}^+$  を

$$\mathcal{V}^+ = \{x^2 : x \in \mathcal{V}^\times\}$$

で定義する. するとつぎが成立することが知られている:

- (i)  $\mathcal{V}^+$  は空でない.
- (ii)  $\mathcal{V}^+$  は凸錐体である.
- (iii)  $\mathcal{V}^+$  は自己双対である. したがって, 対称錐体.
- (iv)  $\mathcal{V}^+$  は等質である.
- (v)

$$x \in \mathcal{V}^+ \iff P(x) > 0$$

となる.

**$\mathcal{V}^+$  の自己同型群**  $GL(\mathcal{V})$  を一般線形群とする．元  $g \in GL(\mathcal{V})$  に対して， $g^*$  を

$$(g^*x|y) = (x|gy), \quad (\forall x, y \in \mathcal{V})$$

で定める． $\mathcal{V}^+$  の自己同型群  $G(\mathcal{V}^+)$  を

$$G(\mathcal{V}^+) = \{g \in GL(\mathcal{V}) : g\mathcal{V}^+ = \mathcal{V}^+\}$$

で定め，その単位元成分を  $G$  と記す． $G(\mathcal{V}^+)$  は  $GL(\mathcal{V})$  の閉部分群となる． $x \in \mathcal{V}^\times$  ならば， $P(x) \in G(\mathcal{V}^+)$  となる．任意の  $x \in \mathcal{V}^+$  に対して，

$$P(x^{1/2})e = x$$

となる．したがって，

$$\mathcal{V}^+ = G(\mathcal{V}^+)e$$

である．

**直交群**  $O(\mathcal{V})$  を内積  $(x|y)$  に関する直交群とする． $K = O(\mathcal{V}) \cap G(\mathcal{V}^+)$  とおくと，

$$K = \{g \in GL(\mathcal{V}) : ge = e\}$$

となる． $\mathcal{V}$  の自己同型がなす群を  $Aut(\mathcal{V})$  とする：すなわち，

$$Aut(\mathcal{V}) = \{g \in GL(\mathcal{V}) : g(xy) = g(x)g(y), \forall x, y \in \mathcal{V}\}$$

である． $Aut(\mathcal{V}) = GL(\mathcal{V}) \cap O(\mathcal{V})$  となる．さらに， $Aut(\mathcal{V})$  の単位元成分は  $Int(Aut(\mathcal{V})) = K$  となる．

**三角群**  $GL(\mathcal{V})$  の部分群  $GT^+(\mathcal{V})$  が三角群であるとは，各  $t \in GT^+(\mathcal{V})$  つぎをみたすときをいう：任意の  $x \in \mathcal{V}$  に対して，

$$(tx_{kl})_{ij} = \begin{cases} 0, & \text{if } (i, j) < (k, \ell), \\ \lambda_{ij}x_{ij}, & \text{if } (i, j) = (k, \ell), \end{cases}$$

となる．ここで， $x = \sum_{j=1}^r x_j c_j + \sum_{j < k} x_{jk}$  は  $x$  の Peirce 分解で， $x_1, x_2, \dots, x_r$  は正の実数、 $x_{k\ell} \in \mathcal{V}_{k\ell}$  ( $k < \ell$ ) であり， $\lambda_{ij}$  は正の実数である．また， $(i, j) < (k, \ell)$  は辞書的順序を意味する．

三角群  $GT^+(\mathcal{V})T$  はつぎのような表現をもつ：任意の  $u \in \mathcal{V}_+$  に対して， $u = \sum_{i=1}^r u_i c_i + \sum_{i < j} u_{ij}$ ， $u_i > 0$ ， $u_{ij} \in \mathcal{V}_{ij}$  と書き，

$$t(u) = P(b_1)\tau_{c_1}(u^{(1)})P(b_2)\tau_{c_2}(u^{(2)}) \times \cdots \times \tau_{c_{r-1}}(u^{(r-1)})P(b_r), \quad (4.3)$$

とおく．ここで， $u^{(j)} = \sum_{k=j+1}^r u_{jk}$  and  $b_j = c_1 + \cdots + c_{j-1} + u_j c_j + c_{j+1} + \cdots + c_r$  である．このとき，写像  $u \mapsto t(u)$  は  $\mathcal{V}^+$  から  $GT^+(\mathcal{V})$  への全単射となることが知られている．

補題 4.1  $\mathcal{V}$  を有限次元かつ単純な EJA でランク  $r$ , 次元  $v$ , Peirce invariant  $2g$  を持つとする.  $\mathcal{V}^+$  を  $\mathcal{V}$  に対応する対称錐体とする. このとき, つぎが成立する:

(i) 可逆元  $x \in \mathcal{V}^\times$  に対して,

$$P(x)^{-1}x = x \quad \text{と} \quad P(x)^{-1} = P(x^{-1})$$

となる.

(ii) 任意の元  $x \in \mathcal{V}$  に対して,

$$\begin{aligned} \text{Tr } L(x) &= (v/r)\text{tr}(x), \\ \text{Det } P(x) &= (\det x)^{2v/r}, \\ \det(P(y)x) &= (\det y)^2 \det x \end{aligned}$$

となる.

(iii) 任意の元  $u \in \mathcal{V}^+$  と  $x \in \mathcal{V}$  に対して,

$$\begin{aligned} (d/dt)|_{t=0} \det(x + tu) &= \det(x)(x^{-1}|u), \\ (d/dt)|_{t=0} (x + tu)^{-1} &= -P(x)^{-1}u, \\ P(P(y)x) &= P(y)P(x)P(y) \end{aligned}$$

となる.

(iv) 任意の元  $y \in \mathcal{V}^+$  に対して, ある元  $t \in \text{GT}^+(\mathcal{V})$  が存在して,

$$y = t^*e$$

となる. さらに,

$$t^{-1}e = (t^*e)^{-1} = y^{-1}$$

が成立する.

(v) 任意の元  $x \in \mathcal{V}^+$  と  $t \in \text{GT}^+(\mathcal{V})$  に対して,  $z = tx$  とおいたとき,

$$\det(x) = \det(t^{-1}z) = \det(t^{-1}e) \det(z)$$

となる.

(vi)  $\mathcal{V}_{ij}$  ( $i = 1, 2, \dots, r, j = i + 1, \dots, r$ ) を (4.1) で定める.  $i \neq j$  のとき,

$$\mathcal{V}_{ik}\mathcal{V}_{jk} \subset \mathcal{V}_{ij}$$

となる. また,  $\{i, j\} \cap \{k, \ell\} = \emptyset$  のとき,

$$\mathcal{V}_{ij}\mathcal{V}_{kl} = 0$$

となる.

(vii)  $c$  をべき等元とし  $z \in \mathcal{V}_{1/2} = \{y \in \mathcal{V} | cy = (1/2)y\}$  とする.  $x = x_1 \oplus x_{12} \oplus x_0$  はべき等元  $c$  に関する Peirce 分解で  $x_1 \in \{y \in \mathcal{V} | cy = y\}$ ,  $x_0 \in \{y \in \mathcal{V} | cy = 0\}$ , と  $x_{12} \in \mathcal{V}_{1/2}$  としたとき,

$$\begin{aligned} \tau_c(z)(x_1 \oplus x_{12} \oplus x_0) \\ = x_1 \oplus (2L(z)x_1 + x_{12}) \oplus (2L(e - c)L(z)^2x_1 + 2L(e - c)L(z)x_{12} + x_0). \end{aligned}$$

$ga$  が成り立つ . 特に ,  $x \in \mathcal{V}^\times$  のとき ,

$$\tau_c(-2x_1^{-1}x_{12})x = x_1 \oplus (x_0 - P(x_{12})x_1^{-1}) \quad (4.4)$$

が成立する .

(viii)  $c_1$  と  $c_2$  を 0 でない原始的なべき等元とする . このとき ,

$$a, b \in \mathcal{V}_{12} = \{y \in \mathcal{V} \mid c_1y = (1/2)y \text{ and } c_2y = (1/2)y\}$$

ならば ,

$$ab = (1/2)(a|b)(c_1 + c_2)$$

となる .

証明 : (i)-(vii) の証明は Faraut and Korányi (1994) の 2 章を参照のこと . (viii) の証明は Faraut and Korányi (1994) の命題 IV.1.4 をすこし書き直せばよい .  $\square$

例 4.2  $\mathcal{V} = \text{Sym}(p, \mathbb{R})$  ( $p \geq 2$ ) とし ,  $p = p_1 + p_2$  ( $p_1, p_2 \in \mathbb{N}$ ) とする .  $\mathcal{V}$  上のジョルダン積  $\circ$  は

$$x \circ y = \frac{1}{2}(xy + yx), \quad (x, y \in \mathcal{V})$$

で定める . いま ,

$$c = \begin{bmatrix} I_{p_1} & 0 \\ 0 & 0 \end{bmatrix}, \quad e - c = \begin{bmatrix} 0 & 0 \\ 0 & I_{p_2} \end{bmatrix}$$

とし ,

$$\begin{aligned} \mathcal{V}_1 &= \left\{ x = \begin{bmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{bmatrix} \in \text{Sym}(p, \mathbb{R}) : cx = x \right\} \\ &= \left\{ \begin{bmatrix} S_{11} & 0 \\ 0 & 0 \end{bmatrix} : S_{11} \in \text{Sym}(p_1, \mathbb{R}) \right\}, \\ \mathcal{V}_0 &= \left\{ x = \begin{bmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{bmatrix} \in \text{Sym}(p, \mathbb{R}) : cx = 0 \right\} \\ &= \left\{ \begin{bmatrix} 0 & 0 \\ 0 & S_{22} \end{bmatrix} : S_{22} \in \text{Sym}(p_2, \mathbb{R}) \right\}, \\ \mathcal{V}_{12} &= \left\{ x = \begin{bmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{bmatrix} \in \text{Sym}(p, \mathbb{R}) : cx = \frac{1}{2}x \right\} \\ &= \left\{ \begin{bmatrix} 0 & S_{12} \\ S_{21} & 0 \end{bmatrix} : S_{12} = S'_{21} \in M(p_1, p_2, \mathbb{R}) \right\} \end{aligned}$$

となる .

$$x_1 = \begin{bmatrix} S_{11} & 0 \\ 0 & 0 \end{bmatrix}, \quad x_0 = \begin{bmatrix} 0 & 0 \\ 0 & S_{22} \end{bmatrix}, \quad x_{12} = \begin{bmatrix} 0 & S_{12} \\ S_{21} & 0 \end{bmatrix}$$

とおいたとき,

$$P(x_{12})x_1^{-1} = \begin{bmatrix} 0 & 0 \\ 0 & S_{21}S_{11}^{-1}S_{12} \end{bmatrix}$$

となる. したがって,

$$x_0 - P(x_{12})x_1^{-1} = \begin{bmatrix} 0 & 0 \\ 0 & S_{22} - S_{21}S_{11}^{-1}S_{12} \end{bmatrix}$$

を得る. また,

$$2x_1^{-1}x_{12} = \begin{bmatrix} 0 & S_{11}^{-1}S_{12} \\ S_{21}S_{11}^{-1} & 0 \end{bmatrix}$$

である. (4.4) より

$$\tau_c(-2x_1^{-1}x_{12})x = x_1 \oplus (x_0 - P(x_{12})x_1^{-1}) = \begin{bmatrix} S_{11} & 0 \\ 0 & S_{22} - S_{21}S_{11}^{-1}S_{12} \end{bmatrix}$$

となる. これと (1.2) より

$$\begin{aligned} \begin{bmatrix} S_{11} & 0 \\ 0 & S_{22} - S_{21}S_{11}^{-1}S_{12} \end{bmatrix} &= \tau_c \left( \begin{bmatrix} 0 & -S_{11}^{-1}S_{12} \\ -S_{21}S_{11}^{-1} & 0 \end{bmatrix} \right) \begin{bmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{bmatrix} \\ &= \begin{bmatrix} I_{p_1} & 0 \\ -S_{21}S_{11}^{-1} & I_{p_2} \end{bmatrix} \begin{bmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{bmatrix} \begin{bmatrix} I_{p_1} & -S_{11}^{-1}S_{12} \\ 0 & I_{p_2} \end{bmatrix} \end{aligned}$$

となる.

**補題 4.2** ジョルダン枠  $c_1, c_2, \dots, c_r$  に関する, 元  $u \in \mathcal{V}^+$  の Peirce 分解を  $u = \sum_{j=1}^r u_j c_j + \sum_{j < k} u_{jk}$  とし, 元  $t(u) \in \text{GT}^+(\mathcal{V})$  を (4.3) で定める.

(i) 元  $x = \sum_{j=1}^r x_j c_j + \sum_{j < k} x_{jk}$  が  $x = t(u) \sum_{j=1}^r a_j c_j$  の Peirce 分解とする. ただし,  $a_j > 0$  ( $j = 1, 2, \dots, r$ ) である. このとき,

$$\begin{aligned} x_j &= a_j u_j^2 + \frac{1}{2} \sum_{k=1}^{j-1} a_k \|u_{kj}\|^2, \\ x_{jk} &= a_j u_j u_{jk} + 2 \sum_{\ell=1}^{j-1} a_\ell u_{\ell j} u_{\ell k}. \end{aligned}$$

が成立する. さらに,

$$\text{tr}(x) = \sum_{j=1}^r a_j u_j^2 + \frac{1}{2} \sum_{j=1}^r \sum_{k=j+1}^r a_j \|u_{jk}\|^2,$$

となる. ただし,  $\|\cdot\|$  は a Euclidean norm である.

(ii) 特に,  $a_1 = a_2 = \dots = a_r = 1$  のとき,

$$\begin{aligned} \det(x) &= \prod_{j=1}^r u_j^2, \\ \det(x)^{-v/r} dx &= 2^r \prod_{j=1}^r u_j^{-2g(j-1)-1} du, \end{aligned}$$

となる . ただし ,  $dx$  と  $du$  は  $\mathcal{V}$  と  $\mathcal{V}^+$  に制限したルベグ測度とする .

(iii) 元  $a_{ij} \in \mathcal{V}_{ij}$  ( $i < j$ ) としたとき ,

$$\mathrm{tr}(t(u)a_{ij}) = u_j(u_{ij}|a_{ij}) + 2 \sum_{\ell=j+1}^r (u_{i\ell}|u_{j\ell}a_{ij}).$$

が成立する .

証明 : (i) の最初の部分の証明は Faraut and Korányi (1994) の命題 VI.3.8 を書き直せばよい .

(i) の後半は補題 4.1(ii) からわかる . (ii) の証明は Faraut and Korányi (1994) の定理 VI.3.9 と命題 VI.3.10 からわかる . (iii) を最後に示す . まず ,  $\ell \notin \{i, j\}$  に対して ,  $P(b_\ell)a_{ij} = a_{ij}$  となることに注意する . また ,  $\ell \in \{i, j\}$  に対して , 補題 4.1(vii) から  $P(b_\ell)a_{ij} = u_\ell a_{ij}$  と  $\tau_{c_j}(u_{jk})y = y$  となる . ただし ,  $y \in \mathcal{V}_{mn}$  ( $\{m, n\} \cap \{j, k\} = \emptyset$ ) である . (4.3) と上のことから

$$t(u)a_{ij} = P(b_i)\tau_{c_i}(u^{(i)})P(b_j)\tau_{c_j}(u^{(j)})a_{ij}, \quad (4.5)$$

となる . ここで ,  $\tau_{c_{r+1}}(u^{(r+1)}) = L(e)$  ,  $c_{r+1} = 0$  と  $u^{(r+1)} = 0$  とした . さらに , 補題 4.1(vi)-(viii) と  $\mathrm{tr}(y) = 0$  ( $y \in \mathcal{V}_{k\ell}$  ( $k < \ell$ )) から (4.5) の右辺は

$$\begin{aligned} \mathrm{tr}(t(u)a_{ij}) &= \mathrm{tr}[P(b_i)\tau_{c_i}(u^{(i)})P(b_j)\{a_{ij} + 2 \sum_{k=j+1}^r L(e - c_j)L(u_{jk})a_{ij}\}] \\ &= \mathrm{tr}[P(b_i)\tau_{c_i}(u^{(i)})\{u_j a_{ij} + 2 \sum_{k=j+1}^r u_{jk} a_{ij}\}] \\ &= \mathrm{tr}[P(b_i)\{u_j a_{ij} + 2 \sum_{k=j+1}^r u_{jk} a_{ij} \\ &\quad + 2 \sum_{\ell=i+1}^r L(e - c_i)L(u_{i\ell})(u_j a_{ij} + 2 \sum_{k=j+1}^r u_{jk} a_{ij})\}] \\ &= \mathrm{tr}[2 \sum_{\ell=i+1}^r L(e - c_i)\{u_{i\ell}(u_j a_{ij}) + 4 \sum_{k=j+1}^r u_{i\ell}(u_{jk} a_{ij})\}] \\ &= u_j(u_{ij}|a_{ij}) + 2 \sum_{\ell=j+1}^r (u_{i\ell}|u_{j\ell}a_{ij}), \end{aligned}$$

と書き直せ , (iii) の証明を得る . □

## 4.2 対称錐体上のワイシャート分布

$\mathcal{V}$  を単純 EJA とし,  $v$  をその次元,  $r$  をランク,  $2g$  を Peirce invariant とする.

$\mathcal{V}^+$  上のガンマ関数  $n > (r-1)g$  に対して,  $\mathcal{V}^+$  上のガンマ関数を

$$\Gamma_{\mathcal{V}^+}(n) = \int_{\mathcal{V}^+} e^{-\text{tr}(x)} \det(x)^n \det(x)^{-v/r} (dx)$$

で定義する. ここで, Faraut and Korányi (1994) に従い,  $(dx) = \prod_{i=1}^p dx_{ii} \prod_{i < j} d(\sqrt{2}x_{ij})$  としていることに注意する. これは

$$v = r + \frac{2gr(r-1)}{2}$$

より

$$(dx) = 2^{2gr(r-1)/2} dx = 2^{(v-r)/2} dx, \quad dx = \prod_{i=1}^r dx_{ii} \prod_{i < j} dx_{ij}$$

である. したがって, ルベーク測度に関する積分でガンマ関数を定義したものは,

$$2^{-(v-r)/2} \Gamma_{\mathcal{V}^+}(r)$$

となる. また,  $\det(x)^{-v/r} dx$  は  $G$ -不変測度となっている. なぜならば,  $x \in \mathcal{V}^\times$  に対して, 補題 4.1(ii) より  $x \mapsto P(y)x$  のヤコビアンは  $\text{Det } P(y) = (\det y)^{2v/v}$  となる. 一方, 補題 4.1(ii)

$$\det(P(y)x)^{-v/r} d(P(y)x) = \{(\det y)^2 \det x\}^{-v/r} \text{Det } P(y) dx$$

よりわかる. 補題 4.1 (ii) を用いて計算すれば,

$$\Gamma_{\mathcal{V}^+}(n) = (2\pi)^{(v-r)/2} \prod_{i=1}^r \Gamma(n - g(j-1)) \quad (4.6)$$

となることがわかる.

例 4.3  $\mathcal{V}^+ = \text{Sym}^+(p, \mathbb{R})$  とする. このとき,  $r = p$ ,  $v = p(p+1)/2$ ,  $2g = 1$  となる.  $n > (p-1)/2$  に対して,

$$\Gamma_{\text{Sym}^+(p, \mathbb{R})}(n) = 2^{p(p-1)/4} \pi^{p(p-1)/4} \prod_{j=1}^p \Gamma(n - \frac{j-1}{2})$$

となる. したがって,

$$2^{-p(p-1)/4} \Gamma_{\text{Sym}^+(p, \mathbb{R})}(n) = \pi^{p(p-1)/4} \prod_{j=1}^p \Gamma(n - \frac{j-1}{2}) = \Gamma_p(n)$$

が通常の  $\text{Sym}^+(p, \mathbb{R})$  上のガンマ関数に一致することがわかる.

いま,  $\mathcal{V}^+$  上のラドン測度  $\mu$  に対して, そのラプラス変換を

$$L_\mu(\theta) = \int_{\mathcal{V}} \exp(\theta | x) \mu(dx), \quad \theta \in \mathcal{V}$$

で定める．さらに，

$$\Theta(\mu) = \text{Int}(\{\theta \in \mathcal{V} : L_\mu(\theta) < \infty\})$$

とし，キュムラント関数  $k_\mu$  を

$$k_\mu(\theta) = \log L_\mu(\theta), \quad \theta \in \Theta(\mu)$$

で定める．このとき， $\mu$  によって生成される NEF を

$$\mathcal{F}(\mu) = \left\{ P(\theta, \mu)(dx) = \exp\{(\theta|x) - k_\mu(\theta)\} \mu(dx) \right\}$$

で定義する．

命題 4.1  $n > g(r-1)$  とする．測度

$$\mu_n(dx) = \frac{1}{\Gamma_{\mathcal{V}^+}(n)} (\det x)^{n-v/r} \mathbb{I}_{\mathcal{V}^+} dx$$

のラプラス変換を

$$L_{\mu_n}(\theta) = \int_{\mathcal{V}^+} \exp(\theta|x) \mu_n(dx)$$

とする．このとき， $\theta \in -\mathcal{V}$  ならば， $L_{\mu_n}$  は有限で

$$L_{\mu_n}(\theta) = (\det(-\theta))^{-n}$$

で与えられる．

証明： $\theta \in -\mathcal{V}^+$  に対して，ある  $t \in \text{GT}^+(\mathcal{V}^+)$  が存在して，

$$-\theta = t^* e$$

と書ける．また，補題 4.1(iv) より

$$t^{-1} e = -\theta^{-1}$$

となる．積分

$$\int_{\mathcal{V}^+} \exp(\theta|x) (\det x)^n (\det x)^{-v/r} (dx) (dx)$$

において， $z = tx$  と変換する． $(\det x)^{-v/r} (dx)$  は  $G$ -不変測度なので，

$$\begin{aligned} & \int_{\mathcal{V}^+} \exp\{-(t^* e|x)\} (\det x)^n (\det x)^{-v/r} (dx) (dx) \\ &= \int_{\mathcal{V}^+} \exp\{-(e|tx)\} (\det x)^n (\det x)^{-v/r} (dx) (dx) \\ &= \int_{\mathcal{V}^+} \exp\{-\text{tr}(z)\} (\det(t^{-1}z))^n (\det z)^{-v/r} (dz) (dz) \\ &= (\det(t^{-1}e))^n \int_{\mathcal{V}^+} \exp\{-\text{tr}(z)\} (\det z)^n (\det z)^{-v/r} (dz) (dz) \\ &= (\det \theta)^{-n} \Gamma_{\mathcal{V}^+}(n) \end{aligned}$$

となる．最後から 2 番目の等号は

$$\det(t^{-1}z) = \det(t^{-1}e) \det z$$

を用いた．

□

命題 4.1 から， $\Theta(\mu_n) = \mathcal{V}^+$  となる． $\mu_n$  のキユムランド関数  $k_{\mu_n} : \mathcal{V}^+ \rightarrow \mathbb{R}$  を

$$k_{\mu_n}(\theta) = -n \log \det(-\theta), \quad -\theta \in \mathcal{V}^+$$

で定める．

定義 4.1  $n > g(r-1)$  とする． $\mu_n$  によって生成される NEF を

$$\begin{aligned} \mathcal{F}_n &= \mathcal{F}(\mu_n) = \left\{ W_r(\theta, \mu_n) = \exp\{(\theta|x) - k_{\mu_n}(\theta)\} \mu_n(dx) : -\theta \in \mathcal{V}^+ \right\} \\ &= \left\{ W_r(\theta, \mu_n) = \frac{1}{\Gamma_{\mathcal{V}^+}(n)} (\det(-\theta))^n (\det x)^n e^{(\theta|x)} (\det x)^{-v/r} (dx) : -\theta \in \mathcal{V}^+ \right\} \end{aligned}$$

で定める．これを  $\mu_n$  によって生成されたウイシャート族という．

ウイシャート分布の族の分散関数を求める：そのために，はじめに  $\dot{k}_{\mu_n}$  から求める．補題 4.1(iii) から， $x \in \mathcal{V}$  に対して，

$$\dot{k}_{\mu_n}(\theta)(x) = \left. \frac{d}{dt} k_{\mu_n}(\theta + tx) \right|_{t=0} = -n \left. \frac{d}{dt} \log \det(-\theta - tx) \right|_{t=0} = -n(\theta^{-1}|x)$$

となる．よって，

$$(\dot{k}_{\mu_n}(\theta)|x) = \dot{k}_{\mu_n}(x) = (-n\theta^{-1}|x)$$

より

$$\dot{k}_{\mu_n} : -\mathcal{V}^+ \rightarrow \mathcal{V}, \quad \dot{k}_{\mu_n}(\theta) = -n\theta^{-1} \quad (4.7)$$

を得る． $\mathcal{M}_n = \dot{k}_{\mu_n}(\Theta(\mu_n)) = \mathcal{V}^+$  とする． $\dot{k}_{\mu_n}$  の逆写像  $\psi_{\mu_n} : \mathcal{V} \rightarrow -\mathcal{V}$  は

$$\psi_{\mu_n}(m) = -nm^{-1}, \quad (m \in \mathcal{V}^+)$$

で定まる．記号の乱用になるが， $m \in \mathcal{V}^+$  に対して，

$$W_r(m, \mu_n) = W_r(\psi_n(m), \mu_n)$$

とも書く．したがって， $\sigma \in \mathcal{V}^+$  ( $\psi_n(n\sigma) = -\sigma^{-1}$ ) に対して，

$$\begin{aligned} W_r(n\sigma, \mu_n) &= W_r(\psi(n\sigma), \mu_n) = W_r(-\sigma^{-1}, \mu_n) \\ &= \frac{1}{\Gamma_{\mathcal{V}^+}(n)} (\det \sigma)^{-n} (\det x)^n e^{-(\sigma^{-1}|x)} (\det x)^{-v/r} (dx) \end{aligned}$$

となる．したがって， $w$  が確率測度  $W_r(n\sigma, \mu_n)$  に従えば，

$$\mathbb{E}[w] = n\sigma$$

となる．

つぎに分散関数を求める．補題 4.1(iii) より  $x \in \mathcal{V}^+$  に対して，

$$\dot{\psi}_{\mu_n}(m)(x) = \left. \frac{d}{dt} \psi_{\mu_n}(m + tx) \right|_{t=0} = -n \left. \frac{d}{dt} (m + tx)^{-1} \right|_{t=0} = nP(m)^{-1}x$$

となる．よって，

$$\dot{\psi}_{\mu_n} : \mathcal{V}^+ \rightarrow \mathcal{L}(\mathcal{V}, \mathcal{V}), \quad \dot{\psi}_{\mu_n}(m) = nP(m)^{-1}$$

これより

$$\{\dot{\psi}_{\mu_n}(m)\}^{-1} = n^{-1}P(m)$$

となる．したがって，分散関数  $V_{\mathcal{F}_n} : \mathcal{V}^+ \rightarrow \mathcal{L}(\mathcal{V}, \mathcal{V})$  は

$$V_{\mathcal{F}_n}(m)(x) = \{\dot{\psi}_{\mu_n}(m)\}^{-1}(x)n^{-1}P(m)x, \quad (m \in \mathcal{V}^+, x \in \mathcal{V})$$

で与えられる．

例 4.4  $\mathcal{V}^+ = \text{Sym}^+(p, \mathbb{R})$  とする．したがって， $r = p$ ， $v = p(p+1)/2$ ， $2g = 1$  である．この例の中では，行列  $x$  と  $y$  の積を  $xy$  と書き， $x$  と  $y$  のジョルダン積を  $x \circ y$  と書くことにする．すなわち，

$$x \circ y = \frac{1}{2}(xy + yx) \tag{4.8}$$

である．補題 4.1(iii) より

$$(\det x)^{2v/r} = \text{Det } P(x)$$

となる．さらに， $P(x)y = xyx$  となることと命題 1.3 から

$$\text{Det } P(x) = (\text{Det } x)^{p+1}$$

となるので，

$$\det x = \text{Det } (x)$$

がわかる．また，補題 4.1(ii) より

$$\text{Tr } L(x) = \frac{v}{r} \text{tr } x = \frac{p+1}{2} \text{tr } x$$

である．しかし， $c$  を原始的べき等元とすれば，

$$\text{Tr } L(c) = \text{Tr} \begin{pmatrix} 1 & 0 & & 0 \\ 0 & \frac{1}{2} & & 0 \\ & & \cdots & \\ 0 & & \frac{1}{2} & 0 \\ 0 & & & 0 \\ 0 & & & & 0 \end{pmatrix} = 1 + (p-1)\frac{1}{2} = \frac{p+1}{2}$$

となる。したがって、

$$\mathrm{Tr} x = \mathrm{tr} x$$

となる。さらに、

$$(dx) = 2^{p(p-1)/4} \prod_{i=1}^p dx_{ii} \prod_{i<j} dx_{ij} = 2^{p(p-1)/4} dx$$

であった。よって、

$$W_p(n\sigma, \mu_n) = \frac{1}{\Gamma_{\mathcal{V}^+}(n)/2^{p(p-1)/4}} (\mathrm{Det} \sigma)^{-n} (\mathrm{Det} x)^n e^{-(\sigma^{-1}|x)} (\mathrm{Det} x)^{-(p+1)/2} dx$$

となる。さらに、 $w = (w_{ij}) = 2x$  とおけば、

$$dx = \prod_{i=1}^p dx_{ii} \prod_{i<j} dx_{ij} = 2^{p(p+1)/2} \prod_{i=1}^p dw_{ii} \prod_{i<j} dw_{ij} = 2^{p(p+1)/2} dw$$

となる。これより

$$\begin{aligned} W_p(n\sigma, \mu_n) &= \frac{2^{-pn}}{\Gamma_{\mathcal{V}^+}(n)/2^{p(p-1)/4}} (\mathrm{Det} \sigma)^{-n} (\mathrm{Det} w)^n e^{-(1/2)(\sigma^{-1}|w)} (\mathrm{Det} w)^{-(p+1)/2} dw \\ &= \frac{1}{2^{pn} \pi^{p(p-1)/4} \prod_{j=1}^p \Gamma(n - \frac{1}{2}(j-1))} (\mathrm{Det} \sigma)^{-n} (\mathrm{Det} w)^n e^{-\frac{1}{2} \mathrm{Tr}(\sigma^{-1}w)} \\ &\quad \times (\mathrm{Det} w)^{-(p+1)/2} dw \end{aligned}$$

を得る。したがって、

$$\mathbb{E}[w] = 2\mathbb{E}[s] = 2n\sigma$$

となる。 □

例 4.5  $\mathcal{V}^+ = \mathrm{Herm}(p, \mathbb{C})$  とする。  $r = p$ ,  $v = p^2$ ,  $2g = 2$  となる。ジョルダン積 (4.8) を用いる。また、

$$\det x = \mathrm{Det} x, \quad \mathrm{tr} x = \mathrm{Tr} x, \quad \frac{1}{\Gamma_{\mathcal{V}^+}(n)} (dx) = \frac{1}{\pi^{p(p-1)/2} \prod_{i=1}^p \Gamma(g - 2(i-1))} dx$$

である。したがって、

$$W_p(n\sigma, \mu_n) = \frac{1}{\pi^{p(p-1)/2} \prod_{i=1}^p \Gamma(g - 2(i-1))} (\mathrm{Det} \sigma)^{-n} (\mathrm{Det} x)^n \exp(-\mathrm{Tr}(\sigma^{-1}x)) (\mathrm{Det} x)^{-p} dx$$

となる。したがって、

$$\mathbb{E}[w] = n\sigma$$

を得る。 □

例 4.6  $\mathcal{V} = \mathbb{R} \times \mathcal{W}$  とする。ただし、 $\mathcal{W}$  は双一次形式  $B$  をもつ  $(v-1)$ -次元実ベクトル空間とする。したがって、 $\dim \mathcal{V} = v$ ,  $r = 2$ ,  $2g = v - 2$  である。さらに、

$$\mathcal{V}^+ = (\mathbb{R} \times \mathcal{W})^+ = \{(x_0, x_2) \in \mathbb{R} \times \mathcal{W} : x_0^2 - B(x_1|x_2) > 0\}$$

とする．このとき，以下が成立する： $(x_0, x_1) \in (\mathbb{R} \times \mathcal{W})^+$  に対して，

$$\operatorname{tr} x = 2x_0, \quad (4.9)$$

$$\det x = x_0^2 - B(x_1|x_1), \quad (4.10)$$

$$x^{-1} = \frac{1}{\det x}(x_0, -x_1).$$

よって，

$$\begin{aligned} W_2(n\sigma, \mu_n) &= \frac{1}{\Gamma_{\mathcal{V}^+}(n)} (\det \sigma)^{-n} (\det x)^n e^{-(\sigma^{-1}|x)} (\det x)^{-v/2} (dx) \\ &= \frac{1}{\pi^{(v-2)/2} \Gamma(n) \Gamma(n - \frac{v-2}{2})} (\det \sigma)^{-n} (\det x)^n e^{-(\sigma^{-1}|x)} (\det x)^{-v/2} dx \end{aligned}$$

となる．なぜならば， $x = (x_0, x_1)$  に対して，

$$x^2 - 2x_0x + (x_0^2 - B(x_1|x_1))e = 0$$

よりわかる．

$\sigma = (\sigma_0, \sigma_1)$ ,  $e = (1, 0)$  とする． $\det \sigma \neq 0$  のとき，(4.10) より

$$\sigma^{-1} = -\frac{\sigma - \operatorname{tr}(\sigma)e}{\det \sigma}$$

である．これより

$$\sigma^{-1}x = \frac{1}{\det \sigma}(\sigma_0x_0 - B(\sigma_1|x_1), \sigma_0x_1 - x_0\sigma_1)$$

となるので，

$$\operatorname{tr}(\sigma^{-1}x) = \frac{2}{\det \sigma}(\sigma_0x_0 - B(\sigma_1|x_1))$$

となる．よって，

$$\begin{aligned} W_2(n\sigma, \mu_2) &= \frac{1}{\Gamma_{(\mathbb{R} \times \mathcal{W})^+}(n)} \frac{(x_0^2 - B(x_1|x_1))^n}{(\sigma_0^2 - B(\sigma_1|\sigma_1))^2} \exp\left\{-\frac{2}{\sigma_0^2 - B(\sigma_1|\sigma_1)}(\sigma_0x_0 - B(\sigma_1|x_1))\right\} \\ &\quad \times \frac{(dx)}{(x_0^2 - B(x_1|x_1))^{-v/2}} \end{aligned}$$

を得る． □

写像  $\Lambda : \mathbb{R} \times \mathcal{W} \rightarrow \mathcal{L}(\mathbb{R}^p, \mathbb{R}^p)$  を一対一線形写像で

$$\Lambda(xy) = \frac{1}{2}\{\Lambda(x)\Lambda(y) + \Lambda(y)\Lambda(x)\} \quad (4.11)$$

をみたすものとする．

補題 4.3  $\mathcal{V}$  を単純な有限次元 EJA でそのランクを  $r$ ，次元を  $v$ ，Peirce invariant を  $2g$  とする． $\mathcal{V}^+$  を  $\mathcal{V}$  に対応する対称錐体とし，線形写像  $\Lambda : \mathbb{R} \times \mathcal{W} \rightarrow \mathcal{L}(\mathbb{R}^p, \mathbb{R}^p)$  を (4.11) をみたす線形写像とする．このとき，以下が成立する：

- (i) 各元  $x \in \mathcal{V}^+$  に対して， $\Lambda(x^{-1}) = \Lambda(x)^{-1}$
- (ii) 任意の元  $x \in \mathcal{V}$  に対して， $\operatorname{Det} \Lambda(x) = (\det x)^{p/r}$
- (iii)  $x \in \mathcal{V}^+$  とする．任意の  $\xi \in \mathbb{R}^p$  に対して，ある  $y \in \overline{\mathcal{V}^+}$  が存在して，

$$\operatorname{Tr}(\Lambda(x)\xi'\xi) = (x|y).$$

証明 (i) は  $\Lambda(e)$  は単位行列であることと  $\Lambda(x)$  と  $\Lambda(x^{-1})$  は可換であることからわかる。(ii) は Faraut and Korányi (1994, 命題 IV.4.2) と Tolver Jensen (1988, 補題 2) を参照のこと。□

$(\mathbb{R} \times \mathcal{W})^+$  で母数化された多変量正規モデル  $\sigma = (\sigma_0, \sigma_1) \in (\mathbb{R} \times \mathcal{W})^+$  とする。確率行列  $X : m \times p$  は確率測度

$$N(\Lambda(\sigma), m, p) = \frac{1}{(2\pi)^{mp/2} \{\text{Det } \Lambda(\sigma)\}^{m/2}} \exp \left\{ -\frac{1}{2} \text{Tr} (\Lambda(\sigma^{-1})x'x) \right\} dx$$

とする。ただし,  $dx$  は  $\mathbb{R}^{m \times p}$  上のルベグ測度である。

$\sigma$  の最尤推定量を求める。そのために, 任意の  $y \in \mathbb{R} \times \mathcal{W}$  に対して, 元  $w \in \overline{\mathbb{R} \times \mathcal{W}}$  は

$$\text{Tr} (\Lambda(y)x'x) = (y|w)$$

をみたすものとする。  $w$  は  $x$  の依存するので,  $w = Q(x)$  とも書く。  $y = (y_0, y_1)$  として,  $\Lambda$  の線形性,  $\Lambda(e) = \Lambda(1, 0) = I_p$ , および (4.9) を用いると

$$y_0 \text{Tr} (x'x) + \text{Tr} (\Lambda((0, y_1)x'x) = \text{Tr} (\Lambda(y)x'x) = \text{tr} (yw) = y_0 w_0 + B(y_1|w_2)$$

となる。したがって,

$$\begin{cases} w_0 = \text{Tr} (x'x) \\ \text{Tr} (\Lambda((0, y_1)x'x) = B(y_1|w_2) (\forall y_1 \in \mathcal{W}) \end{cases}$$

となる。

$\sigma$  の対数尤度関数  $l(\sigma)$  は

$$\begin{aligned} l(\sigma) &= -\frac{1}{2} \text{Tr} (\Lambda(\sigma^{-1})x'x) - \frac{m}{2} \log \text{Det } \Lambda(\sigma) \\ &= -\frac{1}{2} (\sigma^{-1}|w) - \frac{mp}{4} \log \det \sigma \end{aligned}$$

となる。最後の等号は補題 4.3 からわかる。補題 4.1(iii) から,  $u \in \mathbb{R} \times \mathcal{W}$  に対して,

$$\left. \frac{d}{dt} l(\sigma + tu) \right|_{t=0} = \frac{1}{2} (P(\sigma)^{-1}u|w) - \frac{mp}{4} (\sigma^{-1}|u)$$

となる。よって,

$$(u|P(\sigma)^{-1}w) = \frac{mp}{2} (u|\sigma^{-1}) \iff P(\sigma)^{-1}w = \frac{mp}{2} \sigma^{-1} \iff \sigma = \frac{2}{mp} w$$

となる。したがって,  $\sigma$  の最尤度推定量は

$$\hat{\sigma} = \frac{2}{mp} w, \quad w = (w_0, w_1), \quad \begin{cases} w_0 = \text{Tr} (x'x) \\ \text{Tr} (\Lambda(0, y_1)x'x) = B(y_1|w_2) (\forall y_1 \in \mathcal{W}) \end{cases}$$

である。

最後に,  $w$  の分布を求める: そのために  $w = Q(x)$  のラプラス変換を求める.  $\theta \in \mathbb{R} \times \mathcal{W}$  に対して,

$$\begin{aligned} & \int_{\mathbb{R}^{m \times p}} \exp(\theta | w) N(\sigma, m, p)(dx) \\ &= \frac{1}{(2\pi)^{mp/2} (\text{Det } \Lambda(\sigma))^{m/2}} \int_{\mathbb{R}^{m \times p}} \exp \left\{ -\frac{1}{2} \text{Tr} (\Lambda(\sigma^{-1}) x' x) + \text{Tr} (\Lambda(\theta) x' x) \right\} dx \\ &= \frac{1}{(2\pi)^{mp/2} (\text{Det } \Lambda(\sigma))^{m/2}} \int_{\mathbb{R}^{m \times p}} \exp \left\{ -\frac{1}{2} \text{Tr} (\Lambda(\sigma^{-1} - 2\theta) x' x) \right\} dx \\ &= \{ \text{Det } \Lambda(\sigma) \text{Det } \Lambda(\sigma^{-1} - 2\theta) \}^{m/2} \\ &= \{ \det \sigma \det(\sigma^{-1} - 2\theta) \}^{-mp/4} \end{aligned}$$

となる. 最後の等号は補題 4.3(ii) からわかる.

一方,  $W_2(\frac{mp}{2}\sigma, \mu_{mp/4}) = W_2(-(2\sigma)^{-1}, \mu_{mp/4})$  のラプラス変換を求めてみる:

$$\begin{aligned} & \int_{(\mathbb{R} \times \mathcal{W})^+} \exp(\theta | x) W_2(-(2\sigma)^{-1}, \mu_{mp/4})(dx) dw \\ &= \frac{(\det 2\sigma)^{-mp/4}}{\Gamma_{(\mathbb{R} \times \mathcal{W})^+}(mp/4)} \int_{(\mathbb{R} \times \mathcal{W})^+} (\det x)^{mp/4 - v/2} \exp(-((2\sigma)^{-1} - \theta | x)) (dx) \\ &= \{ (\det 2\sigma) \det((2\sigma)^{-1} - \theta) \}^{-mp/4} \\ &= \{ (\det \sigma) \det(\sigma^{-1} - 2\theta) \}^{-mp/4} \end{aligned}$$

となる. したがって,  $w \sim W_2(\frac{mp}{2}\sigma, \mu_{mp/4}) = W_2(-(2\sigma)^{-1}, \mu_{mp/4})$  となる. これより

$$\mathbb{E}[\hat{\sigma}] = \frac{2}{mp} \mathbb{E}[w] = \sigma$$

がわかる.

### Bartlett 分解

命題 4.2  $w \sim W_r(ne, \mu_n)$  とし,  $w = t(u)e$  と分解する. ただし,  $u \in \mathcal{V}^+$  で  $t(u)$  は (4.3) で与えられる. ジョルダン枠  $c_1, \dots, c_r$  に関する  $u$  を Peirce 分解を

$$u = \sum_{i=1}^r u_i c_i + \sum_{i < j} u_{ij}, \quad u_{ij} \in \mathcal{V}_{ij}$$

と書いたとき,  $(u_i, u_{ij} | i = 1, \dots, r, j = i, \dots, r)$  は互いに独立に

$$u_i^2 \sim \chi_{n-g(i-1)}^2, \quad u_{ij} \sim N_{2g}(0, I_{2g})$$

に従う.

証明: 補題 4.2(ii) を用いる:

$$\begin{aligned} W(ne, \mu_n)(dx) &= \frac{1}{\Gamma_{\mathcal{V}^+}} (\det x)^n e^{-\text{tr } x} (\det x)^{-v/r} (dx) \\ &= \frac{2^r}{(2\pi)^{(n-r)/2} \prod_{i=1}^r \Gamma(n-g(i-1))} \prod_{i=1}^n e^{-u_i^2} u_i^{2n-2g(i-1)-1} du_i \\ &\quad \times \prod_{i < j} \exp\{-(1/2)\|u_{ij}\|\} du_{ij} \end{aligned}$$

となる．最後に， $2u_i^2 = s_i$  と変換すれば， $4u_i du_i = ds_i$  となり，等式  $v = r + gp(p-1)$  と (4.6) に注意して変形をすれば，

$$W(ne, \mu_n)(dx) = \prod_{i=1}^r \frac{1}{2^{n-g(i-1)} \Gamma(n-g(i-1))} e^{-(1/2)s_i} s_i^{n-g(i-1)-1} ds_i \\ \times \prod_{i < j} \frac{1}{(2\pi)^g} \exp\{-(1/2)\|u_{ij}\|\} du_{ij}$$

よりわかる．

□

### 固有根の分布

命題 4.3  $w \sim W_r(n\sigma, \mu_n)$  とする．固定したジョルダン枠  $c_1, \dots, c_r$  に関する  $w$  の固有根  $\ell_1, \dots, \ell_r$  の分布はつぎで与えられる．

$$\text{constant} \frac{(\det \sigma)^n}{\Gamma_{\mathcal{V}^+}} \left( \prod_{i=1}^r \ell_i \right)^{n-v/r} \prod_{i < j} (\ell_i - \ell_j)^{2g} \int_K \exp\left(-\sum_{i=1}^r (k^* \sigma | c_i)\right) dk$$

で与えられる．ただし， $dk$  は  $K$  上の不変測度である．

証明：信じることにする．

□

## 参考文献

- [1] ANDERSEN. H.H., HØBJERRE, M., SØRENSEN, D., and ERIKSEN, P.S. (1995). *Linear and Graphical models*. Springer-Verlag, New York.
- [2] ANDERSSON, S.A.(1975). Invariant normal models. *Ann. Statist.* **3** 132-154.
- [3] ANDERSSON, S.A. (1982). Distributions of maximal invariants using quotient measures. *Ann. Statist* **10** 965–961.
- [4] ANDERSSON. S.A., BRØNS, H.K., and TOLVER JENSEN, S. (1983). Distribution of eigenvalues in multivariate statistical analysis. *Ann. Statist.* **11**. 392–415.
- [5] ANDERSSON, S.A. and PERLMAN, M.D. (1984). Two testing problems relating the real and complex multivariate normal distributions. *J. Multivariate Anal.* **15** 21-51.
- [6] ANDERSSON, S.A. and PERLMAN, M.D. (1993). Lattice models for conditional independence in a multivariate normal distribution. *Ann. Statist* **21** 1318-1358.
- [7] ANDERSSON, S.A. and WOJNAR, G.G.(2001). The Wishart distributions on homogeneous cones. Technical report no. 442, Institut für Mathematik der Universität Augsburg. Available at <http://www.math.uni-augsburg.de/stochastik/reports/welcome.html>.
- [8] ANDERSSON, S.A. and WOJNAR, G.G.(2004). Wishart distributions on homogeneous cones. *J. Theoret. Prob.* **17** 781–818.
- [9] BAR-LEV. S.K. and CASALIS, M. (2003). A classification of reproducible natural exponential families in the broad sense. *J. Theoret. Prob.* **16** 175–196.
- [10] BAR-LEV, S.K., ENIS. P., and Letac, G. (1994). Sampling models which admit a given general exponential families as a conjugate families of priors. *Ann. Statist.* **22** 1555–1585.
- [11] BARNDORFF-NIELSEN. O.E., BLÆSID, P., and ERIKSEN, P.S. (1989). *Decomposition and invariance of measures, and statistical transformation models*. Springer
- [12] BONDER, J. and MILNES, P. (1981). Amenability: A survey for statistical applications of Hunt-Stein and related conditions on groups. *Z. Wahrsch. verw. Gebiete* **57** 103-128.
- [13] BROWN, L.D. (1986). *Foundations of statistical exponential families*. IMS lecture-notes-monograph series.
- [14] CASALIS, M. (1996). The  $2d + 4$  simple quadratic natural exponential families on  $\mathbb{R}^d$ .

*Ann. Statist.* **24** 1828–1854.

[15] CASALIS, M. and LETAC, G. (1994). Characterization of the Jorgensen set in generalized linear models. *it Test* **3** 145–162.

[16] CASALIS, M. and LETAC, G. (1996). The Lukas-Olkin-Rubin characterization of Wishart distributions on symmetric cones. *Ann. Statist.* **24** 763–786.

[17] CONSONNI, G. and VERONESE, P.(2003). Enriched conjugate and reference priors for the Wishart family on symmetric cones. *Ann. Statist.* **31** 1491–1516.

[18] DAS PEDDADA, s. and RICHARDS, D.St.P. (1991). Proof of a conjecture of M.L. Eaton on the characteristic function of the Wishart distribution. *Ann. Prob.* **19** 898–874.

[19] DEY, D.K. and SRINIVASAN, C.(1985). Estimation of a covariance matrix under Stein's loss. *Ann. Statist.* **13** 1581–1591.

[20] EATON, M.L. (1983). *Multivariate statistics*. John Wiley & Sons.

[21] EATON, M.L. (1989). *Group Invariance Application in Statistics*. Regional conference series in Probability and Statistics Vol. 1. Institute of Mathematical Statistics.

[22] FARAUT, J. and KORÁNYI, K.(1994). *Analysis on symmetric cones*. Oxford Science Publications.

[23] FAYBUSOVICH, L.(1997). Euclidean Jordan algebras and interior-point algorithms. *Positivity* **1** 331–357.

[24] FAYBUSOVICH, L. and TSUCHIYA, T. (2003). Primal-dual algorithms and infinite-dimensional Jordan algebras of finite rank. *Math. Program., Ser. B* **97** 471–493.

[25] FARRELL, R.H. (1985). *Multivariate calculation*. Springer.

[26] GOODMAN, N.R.(1963). Statistical analysis based on a certain multivariate complex Gaussian distribution (An introduction). *Ann. Math. Statist.* **34** 152–176.

[27] GRACZYK. P., LETAC, G. and MASSAM, H.(2003). The complex Wishart distribution and symmetric group. *Ann. Statist.* **31** 287–309.

[28] HAFF, L.R. (1980). Empirical Bayes estimation of the multivariate normal covariance matrix. *Ann. Statist.* **8**, 586–597.

- [29] HAFF, L.R. (1982). Identities for the inverse Wishart distribution with computational results in linear and quadratic discrimination. *Sankhyā, ser. B* **44**, 245–258.
- [30] HAFF, L.R. (1992). The variational form of certain Bayes estimators. *Ann. Statist.* **19** 1163–1190.
- [31] HASSAIRI, A. and LAJMI, S. (2001). Riesz exponential families on symmetric cones. *J. Theoret. Prob.* **14** 927–948.
- [32] HASSAIRI, A. and LAJMI, S. (2004). Classification of Riesz exponential families on a symmetric cone by invariance properties. *J. Theoret. Prob.* **17** 521–539.
- [33] HASSAIRI, A., LAJMI, S., and ZINE, R. (2005). Beta-Riesz distributions on symmetric cones. *J. Statist. Plan. Infer.* Available online 16 July 2004.
- [34] HUDON, H.M. (1973). A natural identity for exponential families with application in multiparameter estimation. *Ann. Statist.* **6** 473–484.
- [35] JAMES, W. and STEIN, C. (1961). Estimation with quadratic loss. In: *Proc. Fourth Berkeley Symp. Math. Statist. Prob.* **1** 361–380, Univ. California Press.
- [36] JOHANSSON, D. (1982). Some limit theorems for the eigenvalues of a sample covariance matrix. *J. Multivariate Anal.* **12** 1–38.
- [37] JØRGENSEN, B. (1997). *The theory of dispersion models*. Chapman & Hall.
- [38] KHATRI, C.G. (1965). Classical statistical analysis based on a certain multivariate complex Gaussian distribution. *Ann. Math. Statist.* **36** 98–114.
- [39] KONNO, Y. (2001). Inadmissibility of the maximum likelihood estimator of *J. Multivariate Anal.* **79** 33–51.
- [40] KUBOKAWA, T. and SRIVASTAVA, M.S. (1999). Robust improvement in estimation of a covariance matrix in an elliptically contoured distribution. *Ann. Statist.* **27** 600–609.
- [41] LEHMANN, E.L. and CASELLA, G. (1998), *Theory of point estimation*, 2nd ed. Springer.
- [42] LETAC (1989). A characterization of the Wishart exponential families by an invariance property. *J. Theoret. Prob.* **2** 71–86.
- [43] LETAC, G. and MASSAM, H. (2001). Representations of the Wishart distributions. *Contemporary Mathematics* **261**, 121–142.

- [44] LETAC G. and MASSAM, H. (2001). The normal quasi-Wishart distribution. *Contemporary Mathematics* **287**, 231—239.
- [45] LETAC, G. and MASSAM, H. (1994). A tutorial on noncentral Wishart distributions. Available at <http://www.math.uni-augsburg.de/stochastik/reports/welcome.html>.
- [46] LETAC, G., MASSAM, H., and RICHARDS, D. (2001). An expectation formula for the multivariate Dirichlet distribution. *J. Multivariate Anal.* **77** 113-137.
- [47] LETAC, G. and MORA, M. (1990). Natural real exponential families with cubic variance functions. *Ann. Statist.* **18** 1–37.
- [48] LETAC, G. and WESOŁOWSKI, J. (2009), An independence property for the product of GIG and Gamma laws. *Ann. Prob.* **28** 1371–1383.
- [49] LIN, S.P. and PERLMAN, M.D. (1985). A Monte Carlo comparison of four estimators for a covariance matrix. In *Multivariate Analysis VI* (P.R. Krishnaiah, ed.) 411-426, North-Holland, Amsterdam.
- [50] LOH, W.L. (1988). Estimating covariance matrices. Ph. D. dissertation, Dept. Statist., Stanford Univ.
- [51] LOH, W.L. (1991a). Estimating covariance matrices. *Ann. Statist.* **19** 283–296.
- [52] LOH, W.L. (1991b). Estimating covariance matrices II. *J. Multivariate Anal.* **36** 163–174.
- [53] MASSAM, H.(1994). An Exact decomposition theorem and a unified view of some related distributions for a class of exponential transformation models on symmetric cones. *Ann. Statist.* **32** 389-394.
- [54] MASSAM, H. and NEHER, E.(1997). On transformations and determinants of Wishart variables on symmetric cones. *J. Theoret. Prob.* **10** 867-902.
- [55] MASSAM, H. and NEHER, E.(1998). Estimation and testing for lattice conditional independence models of Euclidean Jordan algebras. *Ann. Statist.* **26** 1051-1082.
- [56] MASSAM, H. and WESOŁOWSKI, J. (2000). An independence property for the product of GIG and Gamma laws. *Ann. Prob.* **28** 1371–1383.
- [57] MASSAM, H. and WESOŁOWSKI, J. (2004). The Matsumoto-Yor property and the structure of th Wishart distributins. To appear it J. Multivariate Anal.

- [58] MATH, A.M. (1997). *Jacobians of matrix transformations and functions of matrix arguments*. World Scientific.
- [59] MCCRIMMON, K. (2004). *A Taste of Jordan Algebra*. Springer.
- [60] MORRIS, C.N. (1982). Natural exponential families with quadratic variance functions. *Ann. Statist.* **10** 65–80.
- [61] MUIRHEAD, R.J. (1982). *Aspects of multivariate statistical theory*. John Wiley & Sons.
- [62] MUIRHEAD, R.J. (1987). Developments in eigenvalue estimation. In *Advances in Multivariate Statistical Analysis* (A.K. Gupta, Ed.), 277–288, Reidel, Boston.
- [63] MUIRHEAD, R.J. and VERATHAWORN, T. (1985). On estimating the latent roots of  $\Sigma_1 \Sigma_2^{-1}$ . In *Multivariate Anal.* (P. R. Krishnaiah, Ed.), Vol. 6, 431–447. North-Holland, Amsterdam.
- [64] PERRON, F. (1992). Minimax estimators of a covariance matrix. *J. Multivariate Anal.* **43** 6–28.
- [65] SHAMAN, P. (1980). The inverted complex Wishart distribution and its application to spectral estimation. *J. Multivariate Anal.* **10** 51–59.
- [66] SHEENA, Y. (1995). Unbiased estimator of risk for an orthogonally invariant estimator of a covariance matrix. *J. Japan Statist. Soc.* **25** 35–48.
- [67] SHEENA, Y. and TAKEMURA, A. (1992). Inadmissibility of non-order-preserving orthogonally invariant estimators of the covariance matrix in the case of Stein's loss. *J. Multivariate Anal.* **41** 117–131.
- [68] SRIVASTAVA, M.S. (2003). Singular Wishart and multivariate Beta distribution. *Ann. Statist.* **31** 1537–1560.
- [69] STEIN, C. (1956). Inadmissibility of the usual estimator for the mean of a multivariate normal distribution. In *Proc. Third Berkeley Symp. Math. Statist. Prob.* **1**, 197–206. Univ. California Press, Berkeley.
- [70] STEIN, C. (1964). Inadmissibility of the usual estimator for the variance of a normal distribution with unknown mean. *Ann. Inst. Statist. Math.* **16** 155–160.
- [71] STEIN, C. (1977). Lectures on the theory of estimation of many parameters. In *Studies in the Statistical Theory of Estimation I* (I. A. Ibragimov and M. S. Nikulin, eds.).

- [72] STEIN, C.M. (1981). Estimation of the mean of a multivariate normal distribution. *Ann. Statist.* **9** 1135–1151.
- [73] TAKEMURA, A. (1984). An orthogonally invariant minimax estimators of the covariance matrix of a multivariate normal populations. *Tsukuba J. Math.* **8** 365-376.
- [74] TOLVER JENSEN, S.(1988). Covariance hypotheses which are linear in both the covariance and the inverse covariance. *Ann. Statist.* **16** 302-322.
- [75] TOLVER JENSEN, S. and MADSEN. J. (2004). Estimation of proportional covariances in the presence of certain linear restrictions. *Ann. Statist.* **32** 219–232.
- [76] UHLING, H. (1994). On singular Wishart and singular multivariate Beta distributions. *Ann. Statist.* **22** 395–405.
- [77] WIJSMAN, R.A. (1990). *Invariant measures on groups and their use in Statistics*. IMS lecture notes-monograph-series.

## 和書

- [78] 赤平昌文 (2003). 統計解析入門，森北 .
- [79] 竹村彰道 (1991). 多変量推測統計の基礎，共立 .



# 付録 A

## A.1

補題 A.1 (Letac (1989)) ある開区間から  $\text{Sym}^+(p, \mathbb{R})$  への写像  $t \mapsto a(t)$  は微分可能とする .  
このとき , 以下が成立する :

$$(i) \frac{d}{dt} \log \text{Det } a(t) = (a^{-1}(t) | \dot{a}(t)),$$

$$(ii) \frac{d}{dt} a^{-1}(t) = -a^{-1} \dot{a}(t) a^{-1},$$

$$(iii) \frac{d}{dt} a^{-2}(t) = -a^{-2}(t) \dot{a}(t) a^{-1}(t) - a^{-1}(t) \dot{(t)} a^{-2}(t).$$

証明 : 信じることにする .

□

## A.2 基本的な事項

**GIG 分布について**  $a, b > 0, \nu \in \mathbb{R}$  に対して,

$$f(x) = \left(\frac{b}{a}\right)^{\nu/2} \frac{x^{\nu-1}}{2K_\nu(\sqrt{ab})} \exp\left(-\frac{1}{2}\left(\frac{a}{x} + bx\right)\right) \mathbb{I}_{(0,\infty)}(x)$$

によって確率密度関数が与えられる確率分布を一般化逆ガウス (GIG) 分布という。ただし,  $K_\nu$  は変形ベッセル関数である。すなわち,

$$K_\nu(z) = \frac{1}{2} \left(\frac{1}{2}\right)^\nu \int_0^\infty t^{-\nu-1} e^{-t-z^2/4t} dt, \quad (z > 0)$$

である。松本 (2004) を参照。

**群の定義** 集合  $G$  が群であるとは, 演算

$$G \times G \rightarrow G \quad ((x, y) \mapsto xy)$$

が定義されていて, つぎの公理をみたつときをいう:

- (i)  $(xy)z = x(yz) \quad (x, y, z \in G),$
- (ii) ある元  $e \in G$  が存在して,  $xe = ex = x \quad (x \in G),$
- (iii) 各元  $x \in G$  に対して, 逆元  $x^{-1} \in G$  が存在して,  $xx^{-1} = x^{-1}x = e.$

**部分群の定義** 群  $G$  の部分集合  $H$  が部分群であるとは,  $H$  が単位元  $e$  を含み,  $G$  の演算に関して群となっているときをいう。したがって,

- (i)  $x, y \in H$  ならば,  $xy \in H,$
- (ii)  $x \in H$  ならば,  $x^{-1} \in H$  をみたさなければならない。

**外微分形式について** ダーリング (1994, page 36, page 178) を参照。

**双対空間について** 田坂 (1991, page 117) を参考に。

## A.3 図とプログラム

大きなウイシャート行列の固有根の経験分布  $n = 130, p = 100$

```
# 倍数
ratio<-1.2
y<-1/ratio
# 次元
no.of.dim<-100
# 標本数
no.of.sample<-trunc(ratio*no.of.dim)
print(no.of.sample)
a<-1+y-2*sqrt(y)
b<-1+y+2*sqrt(y)
winger<-function(x){
return(sqrt((x-a)*(b-x))/(2*pi*y*x))
}
mesh<-0.1
k<-trunc((b-a)/mesh)+1
x<-seq(a,b,by=mesh)
cwinger<-rep(0,k)
for(i in 1:k){
cwinger[i]<-integrate(winger,a,x[i])
}
plot(x,cwinger,xlim=c(0,b+1),ylim=c(0,1),type="l")
```

最尤推定量の固有根の確率密度関数  $p = 4, n = 10, \Sigma = I_4$

```
# MLE の固有根の密度関数の作図
rep.no<-10000
ell.largest<-rep(0,rep.no)
ell.2nd<-rep(0,rep.no)
ell.3rd<-rep(0,rep.no)
ell.smallest<-rep(0,rep.no)
for (i in 1:rep.no){
xw.1<-rnorm(10,0,1)
xw.2<-rnorm(10,0,1)
xw.3<-rnorm(10,0,1)
xw.4<-rnorm(10,0,1)
xw<-matrix(c(xw.1,xw.2,xw.3,xw.4),10,4)
#xw
w<-t(xw)%*%xw
#w
eigen.w<-eigen(w)
#eigen.w
lam<-c(eigen.w$values)
hmat<-matrix(eigen.w$vectors,4,4)
#lam; hmat
ell.largest[i]<-lam[1]/10; ell.2nd[i]<-lam[2]/10; ell.3rd[i]<-lam[3]/10;
+ ell.smallest[i]<-lam[4]/10
}
plot(density(ell.largest),xlim=c(0,4),ylim=c(0,3.0),xlab="",ylab="",
+ main="density of eigenvalues of MLE(dim=4,n=10)")
par(new=T)
plot(density(ell.2nd),xlim=c(0,4),,xlab="",ylab="",ylim=c(0,3.0),main="")
```

```

par(new=T)
plot(density(ell.3rd),xlim=c(0,4),,xlab="",ylab="",ylim=c(0,3.0),main="")
par(new=T)
plot(density(ell.smallest),xlim=c(0,4),,xlab="",ylab="",ylim=c(0,3.0),main="")
mean(ell.1d);mean(ell.2d)
var(ell.1d); mean(ell.2d)
mean(ell.1);mean(ell.2)
var(ell.1); mean(ell.2)

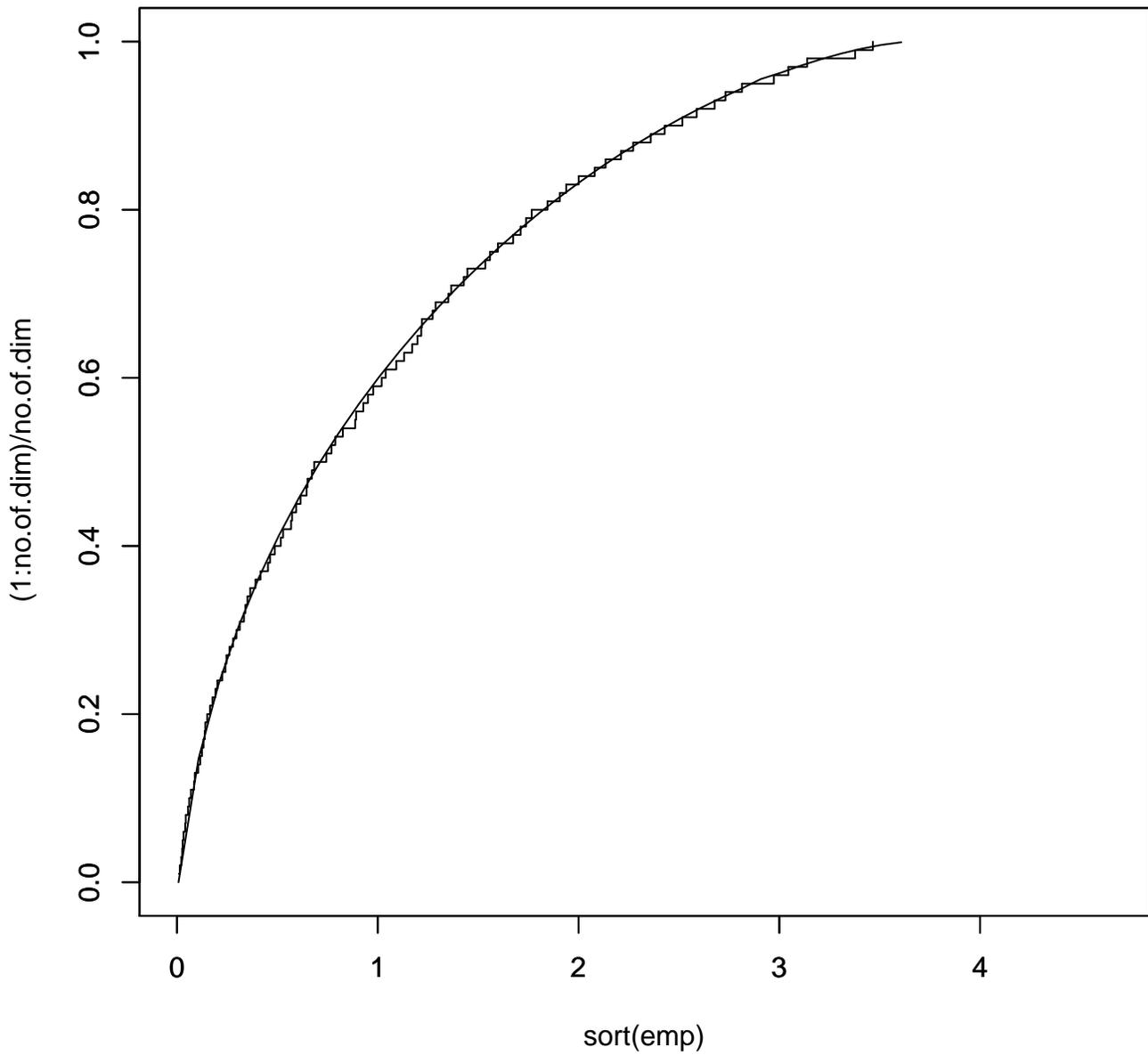
```

DS 推定量の固有根の確率密度関数  $p = 4, n = 10, \Sigma = I_4$

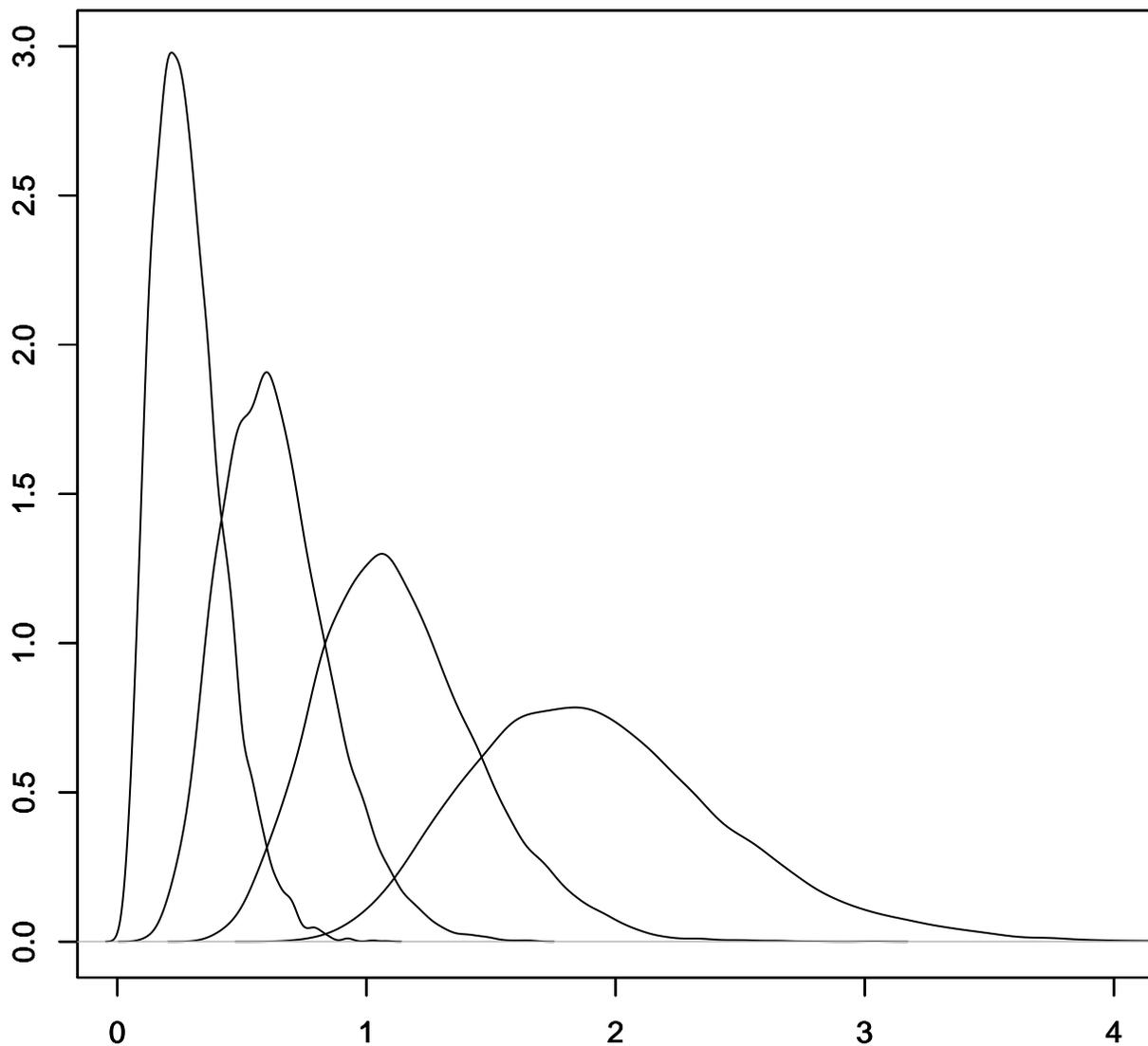
```

rep.no<-100000
ell.largest<-rep(0,rep.no)
ell.2nd<-rep(0,rep.no)
ell.3rd<-rep(0,rep.no)
ell.smallest<-rep(0,rep.no)
ell.1d<-ell.1;ell.2d<-ell.2
#ell.1; ell.2
for (i in 1:rep.no){
  xw.1<-rnorm(10,0,1)
  xw.2<-rnorm(10,0,1)
  xw.3<-rnorm(10,0,1)
  xw.4<-rnorm(10,0,1)
  xw<-matrix(c(xw.1,xw.2,xw.3,xw.4),10,4)
  #xw
  w<-t(xw)%*%xw
  #w
  eigen.w<-eigen(w)
  #eigen.w
  lam<-c(eigen.w$values)
  hmat<-matrix(eigen.w$vectors,4,4)
  #lam; hmat
  #ell.largest[i]<-lam[1]/10; ell.2nd[i]<-lam[2]/10; ell.3rd[i]<-lam[3]/10;
  # ell.smallest[i]<-lam[4]/10
  ell.largest[i]<-lam[1]/13; ell.2nd[i]<-lam[2]/11; ell.3rd[i]<-lam[3]/9;
  + ell.smallest[i]<-lam[4]/7
  #ell.largest[i]<-lam[1]/10; ell.smallest[i]<-lam[4]/10
}
plot(density(ell.largest),xlim=c(0,4),ylim=c(0,3.0),
+ xlab="",ylab="",main="density of eigenvalues of Dey(dim=4,n=10)")
par(new=T)
plot(density(ell.2nd),xlim=c(0,4),,xlab="",ylab="",
+ ylim=c(0,3.0),main="")
par(new=T)
plot(density(ell.3rd),xlim=c(0,4),,xlab="",ylab="",
+ ylim=c(0,3.0),main="")
par(new=T)
plot(density(ell.smallest),xlim=c(0,4),,xlab="",ylab="",ylim=c(0,3.0),main="")
mean(ell.1d);mean(ell.2d)
var(ell.1d); mean(ell.2d)
mean(ell.1);mean(ell.2)
var(ell.1); mean(ell.2)

```



# density of eigenvalues



# density of eigenvalues

