

### 2.3.6 ウィシャート分布

分布関数

$$\mathcal{W}(\mathbf{\Lambda}|\nu, \mathbf{W}) = C_{\mathcal{W}}(\nu, \mathbf{W}) |\mathbf{\Lambda}|^{\frac{\nu-D-1}{2}} \exp \left\{ -\frac{1}{2} \text{Tr}(\mathbf{W}^{-1} \mathbf{\Lambda}) \right\} \quad (2.85)$$

ここで、 $\mathbf{\Lambda}, \mathbf{W}$  はそれぞれ  $D \times D$  の行列。

以後の計算では、見やすさのために、しばしば  $D$  の代わりに  $n$  と書くことにする。

規格化定数の求め方

Multivariate Gamma function 形式への変換

規格化定数  $C(= C_{\mathcal{W}}^{-1})$  は

$$C = \int_{\mathbf{\Lambda} > 0} d\mathbf{\Lambda} |\mathbf{\Lambda}|^{\frac{\nu-n-1}{2}} \exp \left\{ -\frac{1}{2} \text{Tr}(\mathbf{W}^{-1} \mathbf{\Lambda}) \right\} \quad (\text{A.6.1})$$

ここで、積分範囲は  $\mathbf{\Lambda}$  が正定値対称行列の範囲内で行うことに注意。

まず、 $\mathbf{W}^{-1}$  が正値対称行列であることから、下三角行列  $\mathbf{K}$  を使ってコレスキー分解が出来る

$$\mathbf{W}^{-1} = \mathbf{K} \mathbf{K}^T \quad (\text{A.6.2})$$

これを  $C$  の式に代入すると、

$$C = \int_{\mathbf{\Lambda} > 0} d\mathbf{\Lambda} |\mathbf{\Lambda}|^{\frac{\nu-n-1}{2}} \exp \left\{ -\frac{1}{2} \text{Tr}(\mathbf{K}^T \mathbf{\Lambda} \mathbf{K}) \right\} \quad (\text{A.6.3})$$

次に、積分変数を  $\mathbf{\Lambda}$  から  $\mathbf{X} = \mathbf{K}^T \mathbf{\Lambda} \mathbf{K}$  に変換する。

いま、 $\mathbf{K}$  の  $ij$  成分を  $k_{ij}$ ,  $\mathbf{\Lambda}$  の  $ij$  成分を  $\lambda_{ij}$  とし、

$$\mathbf{K} = \begin{pmatrix} k_{11} & 0 \\ \mathbf{k} & \mathbf{K}_1 \end{pmatrix} \quad (\text{A.6.4})$$

$$\mathbf{\Lambda} = \begin{pmatrix} \lambda_{11} & \boldsymbol{\lambda}^T \\ \boldsymbol{\lambda} & \mathbf{\Lambda}_1 \end{pmatrix} \quad (\text{A.6.5})$$

とすると、

$$\begin{aligned}
d\mathbf{X} &= d(\mathbf{K}^T \Lambda \mathbf{K}) \\
&= d \left[ \begin{pmatrix} k_{11} & \mathbf{k}^T \\ 0 & \mathbf{K}_1^T \end{pmatrix} \begin{pmatrix} \lambda_{11} & \boldsymbol{\lambda}^T \\ \boldsymbol{\lambda} & \Lambda_1 \end{pmatrix} \begin{pmatrix} k_{11} & 0 \\ \mathbf{k} & \mathbf{K}_1 \end{pmatrix} \right] \\
&= d \left[ \begin{pmatrix} k_{11}\lambda_{11} + \mathbf{k}^T \boldsymbol{\lambda} & k_{11}\boldsymbol{\lambda}^T + \mathbf{k}^T \Lambda_1 \\ \mathbf{K}_1^T \boldsymbol{\lambda} & \mathbf{K}_1^T \Lambda_1 \end{pmatrix} \begin{pmatrix} k_{11} & 0 \\ \mathbf{k} & \mathbf{K}_1 \end{pmatrix} \right] \\
&= d \begin{pmatrix} (k_{11}\lambda_{11} + \mathbf{k}^T \boldsymbol{\lambda})k_{11} & (k_{11}\boldsymbol{\lambda}^T + \mathbf{k}^T \Lambda_1)\mathbf{K}_1 \\ \mathbf{K}_1^T (\boldsymbol{\lambda}k_{11} + \Lambda_1 \mathbf{k}) & \mathbf{K}_1^T \Lambda_1 \mathbf{K}_1 \end{pmatrix} \\
&= d(k_{11}\lambda_{11}k_{11} + \mathbf{k}^T \boldsymbol{\lambda}k_{11}) \wedge d(k_{11}\boldsymbol{\lambda}^T \mathbf{K}_1 + \mathbf{k}^T \Lambda_1 \mathbf{K}_1) \wedge d(\mathbf{K}_1^T \Lambda_1 \mathbf{K}_1) \\
&= d(k_{11}\lambda_{11}k_{11}) \wedge d(k_{11}\boldsymbol{\lambda}^T \mathbf{K}_1) \wedge d(\mathbf{K}_1^T \Lambda_1 \mathbf{K}_1) \\
&= k_{11}^{n+1} k_{22} \cdots k_{nn} (d\lambda_{11} \wedge d\lambda_{21} \wedge \cdots \wedge d\lambda_{n1}) d(\mathbf{K}_1^T \Lambda_1 \mathbf{K}_1) \\
&= k_{11}^{n+1} k_{22}^{n+1} \cdots k_{nn}^{n+1} (d\lambda_{11} \wedge \cdots \wedge d\lambda_{n1}) \wedge (d\lambda_{22} \wedge \cdots \wedge d\lambda_{n2}) \wedge \cdots \wedge d\lambda_{nn} \\
&= |\mathbf{K}|^{n+1} d\Lambda \\
&= |\mathbf{W}|^{-\frac{n+1}{2}} d\Lambda \tag{A.6.6}
\end{aligned}$$

ここで  $n(n+1)/2$  次元体積積分の体積素に微分形式（というかウェッジ積）を使った。<sup>1</sup>

これを、元の  $C$  の式 (A.6.1) に代入すると、 $|\Lambda| = |\mathbf{X}|/|\mathbf{K}^T \mathbf{K}| = |\mathbf{X}||\mathbf{W}|$  に注意すれば、

$$\begin{aligned}
C &= \int_{\Lambda > 0} d\Lambda |\Lambda|^{\frac{\nu-n-1}{2}} \exp \left\{ -\frac{1}{2} \text{Tr} [\mathbf{W}^{-1} \Lambda] \right\} \\
&= |\Sigma|^{\frac{\nu}{2}} \int_{\mathbf{X} > 0} d\mathbf{X} |\mathbf{X}|^{\frac{\nu-n-1}{2}} \exp \left\{ -\frac{1}{2} \text{Tr} [\mathbf{X}] \right\} \tag{A.6.7}
\end{aligned}$$

となる。この式が、求めたかった Multivariate gamma function 形式。<sup>2</sup>

## 積分

(A.6.7) の積分を行うにあたって、 $\mathbf{X}$  は正定値対称行列であることから、

<sup>1</sup>微分形式での積分については <http://hooktail.sub.jp/differentialforms/DiffFormsLineVol/> を参照

<sup>2</sup>[https://en.wikipedia.org/wiki/Multivariate\\_gamma\\_function](https://en.wikipedia.org/wiki/Multivariate_gamma_function)

下三角行列  $\mathbf{L}$

$$\mathbf{L} = \begin{pmatrix} l_{11} & 0 & \cdots & 0 \\ l_{21} & l_{22} & & 0 \\ \vdots & \vdots & \ddots & \vdots \\ l_{n1} & l_{n2} & \cdots & l_{nn} \end{pmatrix}$$

を用いてコレスキー分解

$$\mathbf{X} = \mathbf{L}\mathbf{L}^T$$

できる。

このとき、参考文献<sup>3</sup>により、 $\mathbf{X}$ での積分を $\mathbf{L}$ での積分に変数変換出来る。(証明は長くなるのでここでは割愛します。ヤコビアンを丁寧に計算すれば、 $\mathbf{L}$ が下三角であることを使って以下の結果を導けます)

このとき積分の $n(n+1)/2$ 次元体積素は

$$d\mathbf{\Lambda} = \left| \frac{\partial \mathbf{\Lambda}}{\partial \mathbf{L}} \right|_J d\mathbf{L} = 2^n \prod_{i=1}^n l_{ii}^{n-i+1} d\mathbf{L} \quad (\text{A.6.8})$$

となり、積分範囲の対角成分は

$$0 \leq l_{ii} < \infty \quad (1 \leq i \leq n) \quad (\text{A.6.9})$$

非対角成分は

$$-\infty < l_{ij} < \infty \quad (1 \leq j < i \leq n) \quad (\text{A.6.10})$$

となる。(この積分範囲を以後、 $\mathbf{L}^+$ とする)

結局、結局規格化定数は

$$\begin{aligned} C &= |\mathbf{W}|^{\frac{\nu}{2}} \int_{\mathbf{\Lambda} > 0} d\mathbf{\Lambda} |\mathbf{\Lambda}|^{\frac{\nu-n-1}{2}} \exp \left\{ -\frac{1}{2} \text{Tr}(\mathbf{\Lambda}) \right\} \\ &= |\mathbf{W}|^{\frac{\nu}{2}} \int_{\mathbf{L}^+} d\mathbf{L} \left| \frac{\partial \mathbf{\Lambda}}{\partial \mathbf{L}} \right| |\mathbf{\Lambda}|^{\frac{\nu-n-1}{2}} \exp \left\{ -\frac{1}{2} \text{Tr}(\mathbf{\Lambda}) \right\} \\ &= |\mathbf{W}|^{\frac{\nu}{2}} \int_0^\infty dl_{11} \cdots dl_{nn} \int_{-\infty}^\infty dl_{21} \cdots dl_{nn-1} \\ &\quad 2^n \prod_{i=1}^n l_{ii}^{n-i+1} |\mathbf{L}\mathbf{L}^T|^{\frac{\nu-n-1}{2}} \exp \left\{ -\frac{1}{2} \text{Tr}(\mathbf{L}\mathbf{L}^T) \right\} \end{aligned} \quad (\text{A.6.11})$$

---

<sup>3</sup><https://home.hiroshima-u.ac.jp/~wakaki/lecture/probstatbasicA19/JofMatrix.pdf> の Theorem 0.2

となる。

積分部分を考えたいので、 $|\mathbf{W}|^{\frac{\nu}{2}}$  を外した  $D$  を

$$C = |\mathbf{W}|^{\frac{\nu}{2}} D \quad (\text{A.6.12})$$

とすると、

$$D = \int_0^\infty dl_{11} \cdots dl_{nn} \int_{-\infty}^\infty dl_{21} \cdots dl_{nn-1} \\ 2^n \prod_{i=1}^n l_{ii}^{n-i+1} |\mathbf{L}\mathbf{L}^T|^{\frac{\nu-n-1}{2}} \exp \left\{ -\frac{1}{2} \text{Tr}(\mathbf{L}\mathbf{L}^T) \right\} \quad (\text{A.6.13})$$

となる。

$D$  の積分を帰納的に実行するため、 $\mathbf{L}$  を次の様に分解する

$$\mathbf{L} = \begin{pmatrix} l_{11} & 0 \\ \mathbf{l} & \mathbf{L}_1 \end{pmatrix} \quad (\text{A.6.14})$$

ここで、 $\mathbf{L}_1$  は  $(n-1) \times (n-1)$  の下三角行列。

まず、 $\text{Tr}$  部分について考えると

$$\begin{aligned} \text{Tr}(\mathbf{L}\mathbf{L}^T) &= \text{Tr} \left[ \begin{pmatrix} l_{11} & 0 \\ \mathbf{l} & \mathbf{L}_1 \end{pmatrix} \begin{pmatrix} l_{11} & \mathbf{l}^T \\ 0 & \mathbf{L}_1 \end{pmatrix} \right] \\ &= \text{Tr} \begin{pmatrix} l_{11}^2 & l_{11}\mathbf{l}^T \\ l\mathbf{l}_1 & \mathbf{L}_1^T + \mathbf{L}_1\mathbf{L}_1^T \end{pmatrix} \\ &= l_{11}^2 + l_{21}^2 + \cdots + l_{n1}^2 + \text{Tr}(\mathbf{L}_1\mathbf{L}_1^T) \end{aligned} \quad (\text{A.6.15})$$

次に行列式部分について考えると

$$|\mathbf{L}\mathbf{L}^T| = |\mathbf{L}||\mathbf{L}^T| = l_{11}^2 |\mathbf{L}_1| |\mathbf{L}_1^T| \quad (\text{A.6.16})$$

式 (A.6.15) と式 (A.6.16) を式 (A.6.13) に代入すると

$$D = D_1 \int_0^\infty dl_{22} \cdots dl_{nn} \int_{-\infty}^\infty dl_{23} \cdots dl_{nn-1} \\ 2^{n-1} \prod_{i=2}^n l_{22}^{n-i+1} |\mathbf{L}_1\mathbf{L}_1^T|^{\frac{(\nu-1)-n}{2}} \exp \left\{ -\frac{1}{2} \text{Tr}(\mathbf{L}_1\mathbf{L}_1^T) \right\} \quad (\text{A.6.17})$$

ここで  $D_1$  は、 $l_{11}, \dots, l_{1n}$  での積分を実行したもので、次の様になる。

$$\begin{aligned}
D_1 &= \int_0^\infty dl_{11} \int_{-\infty}^\infty dl_{21} \cdots dl_{n1} \\
&\quad 2l_{11}^\nu l_{11}^{\nu-n-1} \exp \left\{ -\frac{1}{2}(l_{11}^2 + l_{21}^2 + \cdots + l_{n1}^2) \right\} \\
&= 2(2\pi)^{\frac{n-1}{2}} \int_0^\infty dl_{11} l_{11}^{\nu-1} \exp \left\{ -\frac{1}{2}l_{11}^2 \right\} \\
&= 2(2\pi)^{\frac{n-1}{2}} \int_0^\infty dt (2t)^{\frac{\nu-2}{2}} e^{-t} \\
&= 2^{\frac{\nu}{2}} (2\pi)^{\frac{n-1}{2}} \int_0^\infty t^{\frac{\nu}{2}-1} e^{-t} \\
&= 2^{\frac{\nu}{2}} (2\pi)^{\frac{n-1}{2}} \Gamma\left(\frac{\nu}{2}\right)
\end{aligned} \tag{A.6.18}$$

同様に、次は  $l_{21}, \dots, l_{2n}$  での積分を実行すると、積分の次元が  $n-1$  になっていることと、非積分関数の  $\nu$  を  $\nu-1$  に読みかえれば同じ形の積分になっているので、帰納的に積分していくことで

$$\begin{aligned}
D &= \left[ 2^{\frac{\nu}{2}} (2\pi)^{\frac{n-1}{2}} \Gamma\left(\frac{\nu}{2}\right) \right] \left[ 2^{\frac{\nu-1}{2}} (2\pi)^{\frac{n-2}{2}} \Gamma\left(\frac{\nu-1}{2}\right) \right] \\
&\quad \cdots \left[ 2^{\frac{\nu-(n-1)}{2}} (2\pi)^0 \Gamma\left(\frac{\nu-(n-1)}{2}\right) \right] \\
&= 2^{\frac{\nu n}{2}} \pi^{\frac{n(n-1)}{4}} \prod_{i=1}^n \Gamma\left(\frac{\nu-(i-1)}{2}\right)
\end{aligned} \tag{A.6.19}$$

を得る。

この式を (A.6.12) に代入することで、結局

$$C = |\mathbf{W}|^{\frac{\nu}{2}} 2^{\frac{\nu n}{2}} \pi^{\frac{n(n-1)}{4}} \prod_{i=1}^n \Gamma\left(\frac{\nu-(i-1)}{2}\right) \tag{A.6.20}$$

よって、式 (2.85) の  $C_{\mathcal{W}}$  は、

$$\begin{aligned}
C_{\mathcal{W}} &= C^{-1} \\
&= |\mathbf{W}|^{-\frac{\nu}{2}} 2^{-\frac{\nu n}{2}} \pi^{-\frac{n(n-1)}{4}} \prod_{i=1}^n \frac{1}{\Gamma\left(\frac{\nu-(i-1)}{2}\right)}
\end{aligned} \tag{A.6.21}$$

これを対数表示して、テキストと合わせて  $n$  を  $D$  に、 $i$  を  $d$  に書き直せば、以下を再現する。

$$\begin{aligned} \ln C_{\mathcal{W}}(\nu, \mathbf{W}) &= -\frac{\nu}{2} \ln |\mathbf{W}| - \frac{\nu D}{2} \ln 2 - \frac{D(D-1)}{4} \ln \pi \\ &\quad - \sum_{d=1}^D \ln \Gamma \left( \frac{\nu - (d-1)}{2} \right) \end{aligned} \quad (2.87)$$

分布関数の対数表示

(2.85) 式の対数を取れば、

$$\ln \mathcal{W}(\mathbf{\Lambda} | \nu, \mathbf{W}) = \frac{\nu - D - 1}{2} \ln |\mathbf{\Lambda}| - \frac{1}{2} \text{Tr}(\mathbf{W}^{-1} \mathbf{\Lambda}) + \ln C_{\mathcal{W}}(\nu, \mathbf{W}) \quad (2.86)$$

基本的な期待値

$\mathbf{\Lambda}$  の期待値

$i, j$  成分の期待値として考えると  $\mathbf{M}^{-1} = \mathbf{W}$  を導入して

$$\begin{aligned} \langle \Lambda_{ij} \rangle &= C_{\mathcal{W}}(\nu, \mathbf{W}) \int_{\mathbf{\Lambda} > 0} d\mathbf{\Lambda} \Lambda_{ij} |\mathbf{\Lambda}|^{\frac{\nu-D-1}{2}} \exp \left\{ -\frac{1}{2} \text{Tr}[\mathbf{W}^{-1} \mathbf{\Lambda}] \right\} \\ &= -2C_{\mathcal{W}}(\nu, \mathbf{M}^{-1}) \frac{\partial}{\partial M_{ji}} \int_{\mathbf{\Lambda} > 0} d\mathbf{\Lambda} |\mathbf{\Lambda}|^{\frac{\nu-D-1}{2}} \exp \left\{ -\frac{1}{2} \text{Tr}[\mathbf{M} \mathbf{\Lambda}] \right\} \\ &= -2C_{\mathcal{W}}(\nu, \mathbf{M}^{-1}) \frac{\partial}{\partial M_{ji}} \frac{1}{C_{\mathcal{W}}(\nu, \mathbf{M}^{-1})} \\ &= 2 \frac{\partial}{\partial M_{ji}} \ln C_{\mathcal{W}}(\nu, \mathbf{M}^{-1}) \\ &= \nu \frac{\partial}{\partial M_{ji}} \ln |\mathbf{M}| \\ &= \nu (\mathbf{M}^{-1})_{ij} \\ &= \nu W_{ij} \end{aligned} \quad (\text{A.6.22})$$

よって、テキストの式を再現する。

$$\langle \mathbf{\Lambda} \rangle = \nu \mathbf{W} \quad (2.89)$$

$\ln |\mathbf{\Lambda}|$  の期待値

$$\begin{aligned}
\langle \ln \mathbf{\Lambda} \rangle &= C_{\mathcal{W}}(\nu, \mathbf{W}) \int_{\mathbf{\Lambda}} d\mathbf{\Lambda} \ln |\mathbf{\Lambda}| |\mathbf{\Lambda}|^{\frac{\nu-D-1}{2}} \exp \left\{ -\frac{1}{2} \text{Tr} \mathbf{W}^{-1} \mathbf{\Lambda} \right\} \\
&= C_{\mathcal{W}}(\nu, \mathbf{W}) \int_{\mathbf{\Lambda}} d\mathbf{\Lambda} \ln |\mathbf{\Lambda}| e^{\frac{\nu-D-1}{2} \ln |\mathbf{\Lambda}|} \exp \left\{ -\frac{1}{2} \text{Tr} \mathbf{W}^{-1} \mathbf{\Lambda} \right\} \\
&= 2C_{\mathcal{W}}(\nu, \mathbf{W}) \frac{\partial}{\partial \nu} \int_{\mathbf{\Lambda}} d\mathbf{\Lambda} e^{\frac{\nu-D-1}{2} \ln |\mathbf{\Lambda}|} \exp \left\{ -\frac{1}{2} \text{Tr} \mathbf{W}^{-1} \mathbf{\Lambda} \right\} \\
&= 2C_{\mathcal{W}}(\nu, \mathbf{W}) \frac{\partial}{\partial \nu} \frac{1}{C_{\mathcal{W}}(\nu, \mathbf{W})} \\
&= -2 \frac{\partial}{\partial \nu} \ln C_{\mathcal{W}}(\nu, \mathbf{W}) \\
&= -2 \left( -\frac{1}{2} \ln |\mathbf{W}| - \frac{D}{2} \ln 2 - \frac{1}{2} \sum_{d=1}^D \psi \left( \frac{\nu - (d-1)}{2} \right) \right) \\
&= \ln |\mathbf{W}| + D \ln 2 + \sum_{d=1}^D \psi \left( \frac{\nu - (d-1)}{2} \right) \tag{A.6.23}
\end{aligned}$$

ここで  $\psi$  はディガンマ関数。これを書き直せば、テキストの式 (2.90) を再現する

$$\langle \ln \mathbf{\Lambda} \rangle = \sum_{d=1}^D \psi \left( \frac{\nu + 1 - d}{2} \right) + D \ln 2 + \ln |\mathbf{W}| \tag{2.90}$$

エントロピー

エントロピーの定義式 (2.10) に、上記の期待値  $\langle \mathbf{\Lambda} \rangle$ ,  $\langle \ln |\mathbf{\Lambda}| \rangle$  を入れれば、

$$\begin{aligned}
H[\mathcal{W}(\mathbf{\Lambda}|\nu, \mathbf{W})] &= -\langle \ln \mathcal{W}(\mathbf{\Lambda}|\nu, \mathbf{W}) \rangle \\
&= -\frac{\nu - D - 1}{2} \langle \ln |\mathbf{\Lambda}| \rangle + \frac{1}{2} \text{Tr}(\mathbf{W}^{-1} \langle \mathbf{\Lambda} \rangle) - \ln C_{\mathcal{W}}(\nu, \mathbf{W}) \\
&= -\frac{\nu - D - 1}{2} \langle \ln |\mathbf{\Lambda}| \rangle + \frac{\nu D}{2} - \ln C_{\mathcal{W}}(\nu, \mathbf{W}) \tag{2.91} \\
&= -\frac{\nu - D - 1}{2} \left[ \sum_{d=1}^D \psi \left( \frac{\nu + 1 - d}{2} \right) + D \ln 2 + \ln |\mathbf{W}| \right] + \frac{\nu D}{2} \\
&\quad + \frac{\nu}{2} \ln |\mathbf{W}| + \frac{\nu D}{2} \ln 2 + \frac{D(D-1)}{4} \ln \pi + \sum_{d=1}^D \ln \Gamma \left( \frac{\nu + 1 - d}{2} \right) \tag{A.6.24}
\end{aligned}$$

となる。