2.3.6 ウィシャート分布

分布関数

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

ここで、$\boldsymbol{\Lambda},\boldsymbol{W}$ 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

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

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

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

$$\boldsymbol{W}^{-1} = \boldsymbol{K}\boldsymbol{K}^T \tag{A.6.2}$$

これを C の式に代入すると、

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

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

いま、$\boldsymbol{K}$ の ij 成分を k_{ij}, $\boldsymbol{\Lambda}$ の ij 成分を λ_{ij} とし、

$$\boldsymbol{K} = \begin{pmatrix} k_{11} & 0 \\ \boldsymbol{k} & \boldsymbol{K}_1 \end{pmatrix} \tag{A.6.4}$$

$$\boldsymbol{\Lambda} = \begin{pmatrix} \lambda_{11} & \boldsymbol{\lambda}^T \\ \boldsymbol{\lambda} & \boldsymbol{\Lambda}_1 \end{pmatrix} \tag{A.6.5}$$

とすると、

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

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。[1]

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

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

となる。この式が、求めたかった Multivariate gamma function 形式。[2]

積分

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

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

[2] https://en.wikipedia.org/wiki/Multivariate_gamma_function

下三角行列 $\boldsymbol{L}$

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

を用いてコレスキー分解

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

できる。

このとき、参考文献[3] により、$\boldsymbol{X}$ での積分を $\boldsymbol{L}$ 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、$\boldsymbol{L}$ 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

$$d\boldsymbol{\Lambda} = \left|\frac{\partial \boldsymbol{\Lambda}}{\partial \boldsymbol{L}}\right|_J d\boldsymbol{L} = 2^n \prod_{i=1}^{n} l_{ii}^{n-i+1} d\boldsymbol{L} \tag{A.6.8}$$

となり、積分範囲の対角成分は

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

非対角成分は

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

となる。（この積分範囲を以後、$\boldsymbol{L}^+$ とする）

結局、結局規格化定数は

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

[3] `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、$|\boldsymbol{W}|^{\frac{\nu}{2}}$ を外した D を

$$C = |\boldsymbol{W}|^{\frac{\nu}{2}} D \tag{A.6.12}$$

とすると、

$$\begin{aligned} D = &\int_0^\infty dl_{11} \cdots dl_{nn} \int_{-\infty}^\infty d_{21} \cdots dl_{n\,n-1} \\ &2^n \prod_{i=1}^n l_{ii}^{n-i+1} |\boldsymbol{L}\boldsymbol{L}^T|^{\frac{\nu-n-1}{2}} \exp\left\{-\frac{1}{2}\mathrm{Tr}(\boldsymbol{L}\boldsymbol{L}^T)\right\} \end{aligned} \tag{A.6.13}$$

となる。

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

$$\boldsymbol{L} = \begin{pmatrix} l_{11} & 0 \\ \boldsymbol{l} & \boldsymbol{L}_1 \end{pmatrix} \tag{A.6.14}$$

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

まず、Tr 部分について考えると

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

次に行列式部分について考えると

$$|\boldsymbol{L}\boldsymbol{L}^T| = |\boldsymbol{L}||\boldsymbol{L}^T| = l_{11}^2|\boldsymbol{L}_1||\boldsymbol{L}_1^T| \tag{A.6.16}$$

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

$$\begin{aligned} D = D_1 &\int_0^\infty dl_{22} \cdots dl_{nn} \int_{-\infty}^\infty dl_{23} \cdots dl_{nn-1} \\ &2^{n-1} \prod_{i=2}^n l_{22}^{n-i+1} |\boldsymbol{L}_1\boldsymbol{L}_1^T|^{\frac{(\nu-1)-n}{2}} \exp\left\{-\frac{1}{2}\mathrm{Tr}(\boldsymbol{L}_1\boldsymbol{L}_1^T)\right\} \end{aligned} \tag{A.6.17}$$

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

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

同様に、次は $l_{21},\cdots,l_{2n}$ 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 ν 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

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

を得る。

この式を (A.6.12) に代入することで、結局

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

よって、式 (2.85) の $C_{\mathcal{W}}$ は、

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

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。

$$\ln C_{\mathcal{W}}(\nu, \boldsymbol{W}) = -\frac{\nu}{2}\ln|\boldsymbol{W}| - \frac{\nu D}{2}\ln 2 - \frac{D(D-1)}{4}\ln\pi \\ - \sum_{d=1}^{D}\ln\Gamma\left(\frac{\nu-(d-1)}{2}\right) \tag{2.87}$$

分布関数の対数表示

(2.85) 式の対数を取れば、

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

基本的な期待値

$\boldsymbol{\Lambda}$ の期待値

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

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

よって、テキストの式を再現する。

$$\langle\boldsymbol{\Lambda}\rangle = \nu\boldsymbol{W} \tag{2.89}$$

$\ln|\mathbf{\Lambda}|$ の期待値

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

ここで ψ 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

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

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、

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

となる。