* Wishart分布.

多次元Gauss分布の共分散行列 Σ の逆行列 $\Lambda \in M_D(\mathbb{R})$ (正定値行列.)

(精度行列) を生成する分布.

pdf. $$\mathcal{W}(\Lambda|\nu, W) = C_{\mathcal{W}}(\nu, W)(\det \Lambda)^{\frac{\nu-D-1}{2}} \exp\left(-\frac{1}{2}\mathrm{tr}(W^{-1}\Lambda)\right).$$

正定値性のための条件

$\underline{\nu > D-1}$: 自由度パラメータ, $W \in M_D(\mathbb{R})$: 正定値行列.

$$\log \mathcal{W}(\Lambda|\nu, W) = \frac{\nu-D-1}{2}\log(\det\Lambda) - \frac{1}{2}\mathrm{tr}(W^{-1}\Lambda) + \log C_{\mathcal{W}}(\nu, W)$$

○ $C_{\mathcal{W}}(\nu, W)$ を求めてみる.

以下では, A が正定値行列であることを $A \succ O$ とかく.

正規化項 $C_{\mathcal{W}}(\nu, W)$ の計算には,

$$C_{\mathcal{W}}(\nu, W)\int_{\Lambda \succ O}(\det \Lambda)^{\frac{\nu-D-1}{2}} \exp\left(-\frac{1}{2}\mathrm{tr}(W^{-1}\Lambda)\right)d\Lambda = 1$$

であることを用いる

対称行列ではこの $\frac{D(D+1)}{2}$ コを決めれば行列が一つ定まる.

(なお, $d\Lambda = \underbrace{d\lambda_{11}d\lambda_{21}d\lambda_{22}\cdots\cdots d\lambda_{D1}\cdots d\lambda_{D,D}}_{\text{下三角要素に対応する } \frac{D(D+1)}{2} \text{ コ}}$ である.)

$$\rightarrow \int_{\Lambda \succ O}(\det \Lambda)^{\frac{\nu-D-1}{2}} \exp\left(-\frac{1}{2}\mathrm{tr}(W^{-1}\Lambda)\right)d\Lambda$$

の形の積分を計算する必要がある! この形の積分について考えてみる.

有名な形の積分で, 名前がついている.

Def. (Ingham-Siegel積分)

$X, Q \in M_n(\mathbb{R})$, $Q > O$ とする. 積分

$$I_{n,k}(Q) := \int_{X>O} (\det X)^k \exp(-\mathrm{tr}(XQ))\, dX$$

を Ingham-Siegel積分 とよぶ. □

- $C_W(\nu, W)$ を求めるのに $I_{D, \frac{\nu-D-1}{2}}\left(\frac{1}{2}W^{-1}\right)$ を計算しなければならない.
- Ingham-Siegel積分を行うために, うまく変数変換をしていく必要がある.

そのために多用する定理が, 以下の Cholesky分解.

Th. (Cholesky分解)

$A \in M_n(\mathbb{R})$ が 正定値対称 $(A > O)$

$\Leftrightarrow$ A は, L : 対角項が正の下三角行列で $A = LL^T$ と一意的に

分解(Cholesky分解)できる. □

pf. ($\Leftarrow$) ${}^\forall L$: 対角項が正の下三角行列, ${}^\forall x \in \mathbb{R}^n$ を任意にとる.

このとき, $A = LL^T$ とすると,

$x^T A x = x^T L L^T x = \|L^T x\| \geq 0$. Lの定め方から,

$\|L^T x\| = 0$ となるのは $x = 0$ のときに限る. $\therefore$ A : 正定値対称.

($\Rightarrow$)は煩雑なので略. ■

((⇒)は、線形代数、行列解析、数値解析の本によく書いてある。たとえば、

杉原、室田『線形計算の数理』、山本『行列解析の基礎』など)

Remark この定理が述べているのは、

「正定値行列 と 対角項が正の下三角行列は 1対1に対応する」

ということ。後々の積分の変数変換のときに、積分範囲を考える際に

用いられる。

・ 対角項が正のn次下三角行列全体の集合を $\mathcal{L}_n^+$ と書くことにする。

・ 次の命題も積分範囲を考慮する上で重要。

Prop. $A > O$, R : 正則. $\Rightarrow$ $R^T A R > O$. □

pf. $\forall x \in \mathbb{R}^n$, $x \neq 0$: fix. $A > O$ より $x^T A x > 0$. ―(*)

$x = Ry$ と変換すると、R が正則なので、x と y は 1対1に対応。

(*)より $y^T R^T A R y > 0$. $x \neq 0$ は任意だったから、

y も $y \neq 0$ の範囲をすべてとりうる。 $\therefore R^T A R > O$. ▨

・ 以下では、必要になる変数変換の Jacobian について、事前に計算しておく

Prop. $X, Q \in M_n(\mathbb{R}),\ X, Q > O.$

Qの Cholesky分解を $Q = LL^T$ とする.

$X = (L^T)^{-1} S L^{-1}$ と XをSに変数変換する際の Jacobianは,

$$\frac{\partial X}{\partial S} = (\det Q)^{-\frac{n+1}{2}}$$ $\square$

pf. $S = (s_{ij}),\ X = (x_{ij})$ とする.

逆写像の Jacobian.

$S = L^T X L\ (> O)$ であり, $\dfrac{\partial X}{\partial S} = \left(\dfrac{\partial S}{\partial X}\right)^{-1}$ であるから,

SをXに変数変換するときの Jacobianを考えればよい.

(この変数変換では, $\dfrac{n(n+1)}{2}$コの変数について

$(s_{11}, s_{21}, s_{22}, \cdots, s_{n1}, \cdots, s_{nn}) \to (x_{11}, x_{21}, x_{22}, \cdots, x_{n1}, \cdots, x_{nn})$

と変換が行われており, 従って Jacobi行列 $J(X)$ は

$\dfrac{n(n+1)}{2} \times \dfrac{n(n+1)}{2}$ の行列になっていることに注意する.)

以下, $i \geq j,\ p \geq q$ とする.

$s_{ij} = (L^T X L)_{ij}$

$$\frac{\partial s_{ij}}{\partial x_{pq}} = \frac{\partial}{\partial x_{pq}}\left(\sum_{a=1}^{n}\sum_{b=1}^{n} l_{ai} x_{ab} l_{bj}\right)$$

L: 下三角なので,
$a<i$ では $l_{ai} = 0$.
$b<j$ では $l_{bj} = 0$.

$$= \frac{\partial}{\partial x_{pq}}\left(\sum_{a=i}^{n}\sum_{b=j}^{n} l_{ai} x_{ab} l_{bj}\right)$$

$$= \frac{\partial}{\partial x_{pq}}\left(\sum_{a \geq b} l_{ai} x_{ab} l_{bj} + \sum_{b > a} l_{ai} x_{ba} l_{bj}\right)$$

X: sym.

$\delta_{ij} = \begin{cases} 1 & (i=j) \\ 0 & (i\neq j) \end{cases}$ Kroneckerのデルタ

$$= \sum_{a\geq b} l_{ai} l_{bj} \delta_{ap} \delta_{bq} + \sum_{b>a} l_{ai} l_{bj} \delta_{bp} \delta_{aq}$$

とくに、$(q\leq) p<i$ では $l_{pi} = l_{qi} = 0$, $q<j (\leq i)$ では

$l_{qj} = l_{qi} = 0$ となるので、$\dfrac{\partial s_{ij}}{\partial x_{pq}} = 0 \quad (p<i \text{ or } q<j)$. $-(*)$

$$J_{ip} := \begin{pmatrix} \frac{\partial s_{i1}}{\partial x_{p1}} & \cdots & \frac{\partial s_{i1}}{\partial x_{pp}} \\ \vdots & \ddots & \vdots \\ \frac{\partial s_{ii}}{\partial x_{p1}} & \cdots & \frac{\partial s_{ii}}{\partial x_{pp}} \end{pmatrix} \in M_{i,p}(\mathbb{R}) \quad (i,p = 1,\cdots,n)$$ とすると

Jacobi行列はブロック行列で表せて，$J(X) = \begin{pmatrix} J_{11} & \cdots & J_{1n} \\ \vdots & \ddots & \vdots \\ J_{n1} & \cdots & J_{nn} \end{pmatrix}$.

$(*)$ より，$J_{ip} = O \quad (p<i)$ なので $J(X) = \begin{pmatrix} J_{11} & \cdots & J_{1n} \\ & \ddots & \vdots \\ O & & J_{nn} \end{pmatrix}$.

ブロック三角行列

ゆえに，$\dfrac{\partial X}{\partial S} = \det J(X) = \prod_{i=1}^{n} \det J_{ii}$.

対角項の行列式の積.

$(*)$ より，$\dfrac{\partial s_{ij}}{\partial x_{iq}} = 0 \quad (q<j)$ なので，

$$J_{ii} = \begin{pmatrix} \frac{\partial s_{i1}}{\partial x_{i1}} & \cdots & \frac{\partial s_{i1}}{\partial x_{ii}} \\ & \ddots & \vdots \\ O & & \frac{\partial s_{ii}}{\partial x_{ii}} \end{pmatrix}$$ となり.

三角行列.

$$\det J_{ii} = \prod_{j=1}^{i} \frac{\partial s_{ij}}{\partial x_{ij}} = \prod_{j=1}^{i} l_{ii} l_{jj} = l_{ii}^{i} \prod_{j=1}^{i} l_{jj}.$$

対角項の積.

$$\therefore \frac{\partial X}{\partial S} = \prod_{i=1}^{n}\left(l_{ii}^{\,i}\prod_{j=1}^{i} l_{jj}\right) = \left(\prod_{i=1}^{n} l_{ii}^{\,i}\right)\left(\underbrace{\prod_{i=1}^{n}\prod_{j=1}^{i}}\, l_{jj}\right)$$

$$= \left(\prod_{i=1}^{n} l_{ii}^{\,i}\right)\left(\underbrace{\prod_{j=1}^{n}\prod_{i=j}^{n}}\, l_{jj}\right)$$ 積の順序交換

$$= \left(\prod_{i=1}^{n} l_{ii}^{\,i}\right)\left(\prod_{j=1}^{n} l_{jj}^{\,n-j+1}\right)$$

$$= \prod_{i=1}^{n} l_{ii}^{\,n+1} = \left(\prod_{i=1}^{n} l_{ii}\right)^{n+1}$$

L：下三角より

$\det L = \prod_{i=1}^{n} l_{ii}$ が成立．

$$= (\det L)^{n+1}$$

$$= (\det L \cdot \underbrace{\det L})^{\frac{n+1}{2}}$$

$\det L = \det L^T$．

$$= (\det L \cdot \underbrace{\det L^T})^{\frac{n+1}{2}}$$

$\det AB = \det A \cdot \det B$．

$$= (\det LL^T)^{\frac{n+1}{2}}$$

$$= (\det Q)^{\frac{n+1}{2}}$$

$$\therefore \frac{\partial X}{\partial S} = \left(\frac{\partial S}{\partial X}\right)^{-1} = (\det Q)^{-\frac{n+1}{2}}$$ ▨

<u>Prop.</u> $S \in M_n(\mathbb{R})$, $S > O$．

S の Cholesky 分解を $S = UU^T$ とし，$S = (s_{ij})$, $U = (u_{ij})$ とする．

S を U に変数変換する際の Jacobian は，

$$\frac{\partial S}{\partial U} = 2^n \cdot \prod_{i=1}^{n} u_{ii}^{\,n-i+1}.$$ □

pf. 以下, $i \geq j$, $p \geq q$ とする.

$s_{ij} = (UU^T)_{ij}$

$$\frac{\partial s_{ij}}{\partial u_{pq}} = \frac{\partial}{\partial u_{pq}}\left(\sum_{k=1}^{n} u_{ik}u_{jk}\right)$$

U: 下三角なので, $j < k$ では $u_{jk} = 0$.

$$= \frac{\partial}{\partial u_{pq}}\left(\sum_{k=1}^{j} u_{ik}u_{jk}\right)$$

$$= \sum_{k=1}^{j}\left(\frac{\partial}{\partial u_{pq}}u_{ik}\right)u_{jk} + \sum_{k=1}^{j} u_{ik}\left(\frac{\partial}{\partial u_{pq}}u_{jk}\right)$$

$$= \sum_{k=1}^{j} \delta_{ip}\delta_{kq}u_{jk} + \sum_{k=1}^{j} u_{ik}\delta_{jp}\delta_{kq}$$

とくに, $p > i \ (\geq j)$ では, $\delta_{ip} = \delta_{jp} = 0$,

$q > j$ では, $\delta_{kq} = 0 \ (k = 1, \cdots, j)$

なので, $\dfrac{\partial s_{ij}}{\partial u_{pq}} = 0 \quad (p > i \text{ or } q > j)$ — ①

また, $j = q$ では, $\dfrac{\partial s_{ij}}{\partial u_{pj}} = \displaystyle\sum_{k=1}^{j} \delta_{ip}\delta_{kj}u_{jk} + \sum_{k=1}^{j} u_{ik}\delta_{jp}\delta_{kj}$ — ②

$$J_{ip} := \begin{pmatrix} \frac{\partial s_{i1}}{\partial u_{p1}} & \cdots & \frac{\partial s_{i1}}{\partial u_{pp}} \\ \vdots & \ddots & \vdots \\ \frac{\partial s_{ii}}{\partial u_{p1}} & \cdots & \frac{\partial s_{ii}}{\partial u_{pp}} \end{pmatrix} \in M_{i,p}(\mathbb{R}) \quad (i, p = 1, \cdots, n) \text{ とすると}$$

Jacobi行列はブロック行列で表せて, $J(U) = \begin{pmatrix} J_{11} & \cdots & J_{1n} \\ \vdots & \ddots & \vdots \\ J_{n1} & \cdots & J_{nn} \end{pmatrix}$.

① より $J_{ip} = O \quad (p > i)$ なので $J(U) = \begin{pmatrix} J_{11} & & O \\ \vdots & \ddots & \\ J_{n1} & \cdots & J_{nn} \end{pmatrix}$.

ゆえに, $\dfrac{\partial S}{\partial U} = \det J(U) = \displaystyle\prod_{i=1}^{n} \det J_{ii}$.

再び①より，$\dfrac{\partial s_{ij}}{\partial u_{iq}} = 0 \quad (q > j)$ なので，

$$J_{ii} = \begin{pmatrix} \dfrac{\partial s_{i1}}{\partial u_{i1}} & & O \\ \vdots & \ddots & \\ \dfrac{\partial s_{i1}}{\partial u_{i1}} & \cdots & \dfrac{\partial s_{ii}}{\partial u_{ii}} \end{pmatrix}$$ となり，

$$\det J_{ii} = \prod_{j=1}^{i} \frac{\partial s_{ij}}{\partial u_{ij}}$$

②より

$$= \prod_{j=1}^{i} \left(\sum_{k=1}^{j} \delta_{ii}\delta_{kj} u_{jk} + \sum_{k=1}^{j} u_{ik}\delta_{ji}\delta_{kj} \right)$$

$$= \prod_{j=1}^{i} \left(u_{jj} + u_{ij}\delta_{ji} \right)$$

$$= \left(\prod_{j=1}^{i-1} (u_{jj} + u_{ij}\underbrace{\delta_{ji}}_{=0}) \right)(u_{ii} + u_{ii}\underbrace{\delta_{ii}}_{=1})$$

$$= 2\prod_{j=1}^{i} u_{jj}.$$

$$\therefore \frac{\partial S}{\partial U} = \prod_{i=1}^{n}\left(2\prod_{j=1}^{i} u_{jj}\right) = 2^n \prod_{i=1}^{n}\prod_{j=1}^{i} u_{jj}$$

$$= 2^n \prod_{j=1}^{n}\prod_{i=j}^{n} u_{jj}$$ 積の順序交換

$$= 2^n \prod_{j=1}^{n} u_{jj}^{\,n-j+1}. \quad \blacksquare$$

- 以上でJacobianの計算は終了．あとは積分を実行する．

Th. $Q \in M_n(\mathbb{R})$, $Q > 0$ とする.

$$I_{n,k}(Q) = (\det Q)^{-k-\frac{n+1}{2}} \pi^{\frac{n(n-1)}{4}} \prod_{i=1}^{n} \Gamma\left(k+1+\frac{n-i}{2}\right). \quad \square$$

pf. $I_{n,k}(Q) := \int_{X>0} (\det X)^k \exp(-\mathrm{tr}(XQ))\, dX$ であった.

$Q = LL^T$ と Cholesky 分解する.

$\mathrm{tr}(ABC) = \mathrm{tr}(BCA) = \mathrm{tr}(CAB)$

$$\mathrm{tr}(XQ) = \mathrm{tr}(XLL^T) = \mathrm{tr}(L^T XL).$$

$X = (L^T)^{-1} S L^{-1}$ と変数変換すると, $L^T XL = S$,

$\det(A^{-1}) = (\det A)^{-1}$

$$\det X = \det(L^T)^{-1} \cdot \det S \cdot \det L^{-1} = (\det L \cdot \det L^T)^{-1} \cdot \det S$$

$$= (\det LL^T)^{-1} \det S = (\det Q)^{-1} \det S.$$

変換の Jacobian は, $\dfrac{\partial X}{\partial S} = (\det Q)^{-\frac{n+1}{2}} > 0.$

$\left\{ \det Q = (\det L)^2 > 0 \right.$ ← $\because Q > 0$ より

$\therefore \det L \neq 0$

$X > 0$, L:正則より, $S > 0$.

$\therefore I_{n,k}(Q)$

$$= \int_{S>0} (\det Q)^{-k} (\det S)^k \exp(-\mathrm{tr}(S)) \left|\frac{\partial X}{\partial S}\right| dS$$

$$= (\det Q)^{-k-\frac{n+1}{2}} \int_{S>0} (\det S)^k \exp(-\mathrm{tr}(S))\, dS$$

単位行列.

$$= (\det Q)^{-k-\frac{n+1}{2}} I_{n,k}(I_n)$$

あとは $I_{n,k}(I_n) = \pi^{\frac{n(n-1)}{4}} \prod_{i=1}^{n} \Gamma\left(k+1+\frac{n-i}{2}\right)$ を示せばよい.

積分を実行するため, 行列の要素を使って表してみる.

$I_{n,k}(I_n) = \int_{S>0} (\det S)^k \exp(-\mathrm{tr}(S))\, dS$ で

S を Cholesky 分解して，$S = UU^T$ とする．この変数変換の

Jacobian は，$U = (u_{ij})$ とすると $\dfrac{\partial S}{\partial U} = 2^n \cdot \prod_{i=1}^{n} u_{ii}^{\,n-i+1} > 0$.

← 対角項は正

また，積分範囲は $\{S>0\}$ から $\mathcal{L}_n^+$ にかわる．よって，

$$I_{n,k}(I_n)$$

$$= \int_{\mathcal{L}_n^+} (\det UU^T)^k \exp(-\mathrm{tr}(UU^T)) \left|\frac{\partial S}{\partial U}\right| dU$$

$$= 2^n \cdot \int_{\mathcal{L}_n^+} (\det U)^{2k} \exp(-\mathrm{tr}(UU^T)) \prod_{i=1}^{n} u_{ii}^{\,n-i+1}\, dU.$$

U：下三角行列なので，

$$(\det U)^{2k} = \left(\prod_{i=1}^{n} u_{ii}\right)^{2k} = \prod_{i=1}^{n} u_{ii}^{2k}.$$

$(AB)_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj}$.

$$\mathrm{tr}(UU^T) = \sum_{i=1}^{n} (UU^T)_{ii} = \sum_{i=1}^{n}\sum_{k=1}^{n} u_{ik}u_{ik}$$

$$= \sum_{i=1}^{n}\sum_{k=1}^{n} u_{ik}^2$$

$$= \sum_{i=1}^{n}\sum_{k=1}^{i} u_{ik}^2$$

U：下三角なので $u_{ik} = 0$ $(i<k)$

$\therefore$ $I_{n,k}(I_n)$

$$= 2^n \cdot \int_{\mathcal{L}_n^+} \prod_{i=1}^{n} u_{ii}^{2k} \exp\left(-\sum_{i=1}^{n}\sum_{k=1}^{i} u_{ik}^2\right) \prod_{i=1}^{n} u_{ii}^{\,n-i+1}\, dU$$

$$= 2^n \cdot \int_{\mathcal{L}_n^+} \left(\prod_{i=1}^{n} u_{ii}^{\,2k+n-i+1}\right)\left(\prod_{i=1}^{n}\prod_{k=1}^{i} \exp(-u_{ik}^2)\right) dU.$$

$\mathcal{L}_n^+$ は対角項が正，その他の項は任意の実数をとりうることに注意して，

$$\int_{\mathcal{L}_n^+}\left(\prod_{i=1}^{n} u_{ii}^{2k+n-i+1}\right)\left(\prod_{i=1}^{n}\prod_{k=1}^{i}\exp\left(-u_{ik}^2\right)\right)dU$$

非対角成分.

対角成分.

$$= \prod_{i=1}^{n}\left(\int_0^{\infty} u_{ii}^{2k+n-i+1} e^{-u_{ii}^2} du_{ii}\right)\prod_{i>k}\left(\int_{-\infty}^{\infty} e^{-u_{ik}^2} du_{ik}\right).$$

$$\int_0^{\infty} u_{ii}^{2k+n-i+1} e^{-u_{ii}^2} du_{ii}$$

$t = u_{ii}^2$ とおくと.

$du_{ii} = \dfrac{1}{2\sqrt{t}}dt.$

$$= \int_0^{\infty} t^{k+\frac{n-i+1}{2}} e^{-t}\frac{1}{2}t^{-\frac{1}{2}}dt$$

$$= \frac{1}{2}\int_0^{\infty} t^{k+1+\frac{n-i}{2}-1} e^{-t} dt$$

$$= \frac{1}{2}\Gamma\left(k+1+\frac{n-i}{2}\right).$$

また, $\displaystyle\int_{-\infty}^{\infty} e^{-u_{ik}^2} du_{ik} = \pi^{\frac{1}{2}}.$ ← Gauss積分

以上より,

$$I_{n,k}(I_n)$$

$$= 2^n \cdot \prod_{i=1}^{n}\frac{1}{2}\Gamma\left(k+1+\frac{n-i}{2}\right)\cdot\prod_{i>k}\pi^{\frac{1}{2}}$$

$$= \cancel{2^n}\cdot\frac{1}{\cancel{2^n}}\left(\pi^{\frac{1}{2}}\right)^{\frac{n(n-1)}{2}}\prod_{i=1}^{n}\Gamma\left(k+1+\frac{n-i}{2}\right)$$

$$= \pi^{\frac{n(n-1)}{4}}\prod_{i=1}^{n}\Gamma\left(k+1+\frac{n-i}{2}\right)$$

となるから,

$$I_{n,k}(Q) = (\det Q)^{-k-\frac{n+1}{2}}\pi^{\frac{n(n-1)}{4}}\prod_{i=1}^{n}\Gamma\left(k+1+\frac{n-i}{2}\right). \quad ▨$$

$$\int_{\Lambda>0} \mathcal{W}(\Lambda|\nu, W)\, d\Lambda = 1 \text{ より}$$

$$C_{\mathcal{W}}(\nu, W)\, I_{D, \frac{\nu-D-1}{2}}\left(\frac{1}{2}W^{-1}\right) = 1 \text{ であった.}$$

$$I_{D, \frac{\nu-D-1}{2}}\left(\frac{1}{2}W^{-1}\right)$$

$$= \left(\det \frac{1}{2}W^{-1}\right)^{-\frac{\nu-D-1}{2}-\frac{D+1}{2}} \pi^{\frac{D(D-1)}{4}} \prod_{d=1}^{D} \Gamma\left(\frac{\nu-D-1}{2}+1+\frac{D-d}{2}\right)$$

$$= \left(\frac{1}{2^D}\det W^{-1}\right)^{-\frac{\nu}{2}} \pi^{\frac{D(D-1)}{4}} \prod_{d=1}^{D} \Gamma\left(\frac{\nu+1-d}{2}\right)$$

$$= 2^{\frac{\nu D}{2}} (\det W)^{\frac{\nu}{2}} \pi^{\frac{D(D-1)}{4}} \prod_{d=1}^{D} \Gamma\left(\frac{\nu+1-d}{2}\right).$$

$$\therefore C_{\mathcal{W}}(\nu, W) = \left(2^{\frac{\nu D}{2}} (\det W)^{\frac{\nu}{2}} \pi^{\frac{D(D-1)}{4}} \prod_{d=1}^{D} \Gamma\left(\frac{\nu+1-d}{2}\right)\right)^{-1}.$$

- $\log C_{\mathcal{W}}(\nu, W)$

$$= -\frac{\nu}{2}\log\det W - \frac{\nu D}{2}\log 2 - \frac{D(D-1)}{4}\log\pi - \sum_{d=1}^{D} \log\Gamma\left(\frac{\nu+1-d}{2}\right)$$

- 平均

$$\mathbb{E}[\Lambda] = \int_{\Lambda>0} \Lambda\, \mathcal{W}(\Lambda|\nu, W)\, d\Lambda$$

$$= C_{\mathcal{W}}(\nu, W) \int_{\Lambda>0} \Lambda\, (\det\Lambda)^{\frac{\nu-D-1}{2}} \exp\left(-\frac{1}{2}\mathrm{tr}(W^{-1}\Lambda)\right) d\Lambda.$$

$W>0$ より, $W^{-1}>0$. よって W^{-1} の Cholesky 分解を

$W^{-1} = LL^T$ とおく.

$\mathrm{tr}(W^{-1}\Lambda) = \mathrm{tr}(LL^T\Lambda) = \mathrm{tr}(L^T\Lambda L)$ なので.

$$\mathbb{E}[\Lambda] = C_{\mathcal{W}}(\nu, W)\int_{\Lambda \succ 0} \Lambda (\det\Lambda)^{\frac{\nu-D-1}{2}} \exp\left(-\frac{1}{2}\mathrm{tr}(L^T\Lambda L)\right) d\Lambda .$$

$\to \int_{\Lambda \succ 0} \mathcal{W}(\Lambda | \nu, W)\, d\Lambda = 1$ を使いたいので

積分の中を $\mathcal{W}(\Lambda|\nu, W)$ に変形したい.

[Claim] $\frac{\partial}{\partial L}\mathrm{tr}(L^T\Lambda L) = 2\Lambda L$. □

pf. $L = (l_{ij})$, $\Lambda = (\lambda_{ij})$ とする.

$$\frac{\partial}{\partial l_{pq}}\mathrm{tr}(L^T\Lambda L) = \frac{\partial}{\partial l_{pq}}\left(\sum_{d=1}^{D}(L^T\Lambda L)_{dd}\right)$$

$$= \frac{\partial}{\partial l_{pq}}\left(\sum_{d=1}^{D}\sum_{i=1}^{D}\sum_{j=1}^{D} l_{id}\lambda_{ij}l_{jd}\right)$$

$$= \sum_{d=1}^{D}\sum_{i=1}^{D}\sum_{j=1}^{D}\frac{\partial l_{id}}{\partial l_{pq}}\lambda_{ij}l_{jd} + \sum_{d=1}^{D}\sum_{i=1}^{D}\sum_{j=1}^{D} l_{id}\lambda_{ij}\frac{\partial l_{jd}}{\partial l_{pq}}$$

$$= \sum_{d=1}^{D}\sum_{i=1}^{D}\sum_{j=1}^{D}\lambda_{ij}l_{jd}\delta_{ip}\delta_{dq} + \sum_{d=1}^{D}\sum_{i=1}^{D}\sum_{j=1}^{D} l_{id}\lambda_{ij}\delta_{jp}\delta_{dq}$$

$$= \sum_{j=1}^{D}\lambda_{pj}l_{jq} + \sum_{i=1}^{D} l_{iq}\lambda_{ip}$$

Λ : 対称より

$\lambda_{ip} = \lambda_{pi}$

$$= 2\sum_{i=1}^{D}\lambda_{pi}l_{iq}$$

$$= 2(\Lambda L)_{pq}$$

$\therefore \frac{\partial}{\partial L}\mathrm{tr}(L^T\Lambda L) = 2\Lambda L$. ▨

これより, $\frac{\partial}{\partial L}\exp\left(-\frac{1}{2}\mathrm{tr}(L^T\Lambda L)\right) = -\exp\left(-\frac{1}{2}\mathrm{tr}(L^T\Lambda L)\right)\Lambda L$.

$\therefore \Lambda\exp\left(-\frac{1}{2}\mathrm{tr}(L^T\Lambda L)\right) = -\left(\frac{\partial}{\partial L}\exp\left(-\frac{1}{2}\mathrm{tr}(L^T\Lambda L)\right)\right)L^{-1}$ より,

$\mathbb{E}[\Lambda]$

$$= -C_W(\nu, W)\left(\int_{\Lambda>0} (\det\Lambda)^{\frac{\nu-D-1}{2}} \frac{\partial}{\partial L}\exp\left(-\frac{1}{2}\mathrm{tr}(L^T\Lambda L)\right)d\Lambda\right)L^{-1}$$

$$= -C_W(\nu, W)\left(\frac{\partial}{\partial L}\int_{\Lambda>0} (\det\Lambda)^{\frac{\nu-D-1}{2}} \exp\left(-\frac{1}{2}\mathrm{tr}(L^T\Lambda L)\right)d\Lambda\right)L^{-1}$$

$$= -C_W(\nu, W)\left(\frac{\partial}{\partial L}\frac{1}{C_W(\nu, W)}\right)L^{-1}$$

$$= -\left(\frac{\partial}{\partial L}\log\frac{1}{C_W(\nu, W)}\right)L^{-1} = \left(\frac{\partial}{\partial L}\log C_W(\nu, W)\right)L^{-1}$$

L によらない

$$= \left(\frac{\partial}{\partial L}\left(-\frac{\nu}{2}\log\det W - \cancel{\frac{\nu D}{2}\log 2} - \cancel{\frac{D(D-1)}{4}\log\pi} - \cancel{\sum_{d=1}^{D}\log\Gamma\left(\frac{\nu+1-d}{2}\right)}\right)\right)L^{-1}$$

$$= \left(\frac{\nu}{2}\frac{\partial}{\partial L}\log(\det W)^{-1}\right)L^{-1} = \left(\frac{\nu}{2}\frac{\partial}{\partial L}\log(\det W^{-1})\right)L^{-1}$$

$$= \left(\frac{\nu}{2}\frac{\partial}{\partial L}\log(\det LL^T)\right)L^{-1} = \left(\frac{\nu}{2}\frac{\partial}{\partial L}\log(\det L)^2\right)L^{-1}$$

$$= \left(\nu\frac{\partial}{\partial L}\log\det L\right)L^{-1}$$

$$= \nu(L^T)^{-1}L^{-1}$$

$$= \nu(LL^T)^{-1}$$

$$= \nu(W^{-1})^{-1}$$

$$= \nu W.$$

微分公式：$A>0$, A の各要素が独立.

$$\frac{d}{dA}\log\det A = (A^T)^{-1}.$$

∵ A の固有値：$\lambda_1, \cdots, \lambda_n$.

$$\log\det A = \log\prod_{i=1}^{n}\lambda_i = \sum_{i=1}^{n}\log\lambda_i = \mathrm{tr}\log A.$$

$\frac{d}{dA}\mathrm{tr}(f(A)) = f'(A)^T$ を用いると.

$$\frac{d}{dA}\log\det A = \frac{d}{dA}\mathrm{tr}\log A = (A^{-1})^T = (A^T)^{-1}.$$

$$
\begin{aligned}
&\mathbb{E}[\log\det\Lambda] \\
&= C_W(\nu,W)\int_{\Lambda>0} \underline{\log(\det\Lambda)\,(\det\Lambda)^{\frac{\nu-D-1}{2}}}\exp\left(-\frac{1}{2}\mathrm{tr}(W^{-1}\Lambda)\right)d\Lambda. \\
&= C_W(\nu,W)\int_{\Lambda>0} \underline{2\frac{d}{d\nu}(\det\Lambda)^{\frac{\nu-D-1}{2}}}\exp\left(-\frac{1}{2}\mathrm{tr}(W^{-1}\Lambda)\right)d\Lambda. \\
&= 2C_W(\nu,W)\frac{d}{d\nu}\int_{\Lambda>0}(\det\Lambda)^{\frac{\nu-D-1}{2}}\exp\left(-\frac{1}{2}\mathrm{tr}(W^{-1}\Lambda)\right)d\Lambda. \\
&= 2C_W(\nu,W)\frac{d}{d\nu}\left(\frac{1}{C_W(\nu,W)}\underbrace{\int_{\Lambda>0}\mathcal{W}(\Lambda|\nu,W)d\Lambda}_{=1}\right) \\
&= 2C_W(\nu,W)\frac{d}{d\nu}\left(\frac{1}{C_W(\nu,W)}\right) \\
&= 2\frac{d}{d\nu}\left(\log\frac{1}{C_W(\nu,W)}\right) = -2\frac{d}{d\nu}\log C_W(\nu,W) \\
&= -2\frac{d}{d\nu}\left(-\frac{\nu}{2}\log\det W - \frac{\nu D}{2}\log 2 - \cancel{\frac{D(D-1)}{4}\log\pi} - \sum_{d=1}^{D}\log\Gamma\left(\frac{\nu+1-d}{2}\right)\right) \\
&= \log\det W + D\log 2 + 2\sum_{d=1}^{D}\frac{d}{d\nu}\log\Gamma\left(\frac{\nu+1-d}{2}\right) \\
&= 2\sum_{d=1}^{D}\frac{1}{2}\psi\left(\frac{\nu+1-d}{2}\right) + D\log 2 + \log\det W \\
&= \sum_{d=1}^{D}\psi\left(\frac{\nu+1-d}{2}\right) + D\log 2 + \log\det W.
\end{aligned}
$$

・エントロピー.

$$H[\mathcal{W}(\Lambda|\nu,W)] = -\mathbb{E}[\log \mathcal{W}(\Lambda|\nu,W)]$$

$$= -\frac{\nu-D-1}{2}\mathbb{E}[\log(\det\Lambda)] + \frac{1}{2}\mathbb{E}[\mathrm{tr}(W^{-1}\Lambda)] - \log C_{\mathcal{W}}(\nu,W)$$

$$\downarrow \ \mathbb{E}[\mathrm{tr}(X)] = \mathrm{tr}(\mathbb{E}[X]).$$

$$= -\frac{\nu-D-1}{2}\mathbb{E}[\log(\det\Lambda)] + \frac{1}{2}\mathrm{tr}(W^{-1}\mathbb{E}[\Lambda]) - \log C_{\mathcal{W}}(\nu,W)$$

$$= -\frac{\nu-D-1}{2}\mathbb{E}[\log(\det\Lambda)] + \frac{1}{2}\mathrm{tr}(\nu W^{-1}W) - \log C_{\mathcal{W}}(\nu,W)$$

$$= -\frac{\nu-D-1}{2}\mathbb{E}[\log(\det\Lambda)] + \frac{\nu}{2}\mathrm{tr}(I_D) - \log C_{\mathcal{W}}(\nu,W)$$

$$= -\frac{\nu-D-1}{2}\mathbb{E}[\log(\det\Lambda)] + \frac{\nu D}{2} - \log C_{\mathcal{W}}(\nu,W).$$

・Wishart分布は, χ^2分布の多次元版:

$$\mathbb{X}_i \overset{\text{i.i.d.}}{\sim} \mathcal{N}(\mu,\Sigma) \quad (i=1,\cdots,n).$$

$$\Rightarrow \sum_{i=1}^{n}(\mathbb{X}_i-\mu)(\mathbb{X}_i-\mu)^T \sim \mathcal{W}(n,\Sigma).$$

$$\left(\begin{array}{l} \text{cf.} \quad X_i \overset{\text{i.i.d.}}{\sim} \mathcal{N}(\mu,\sigma^2) \\ \Rightarrow \dfrac{1}{\sigma^2}\sum_{i=1}^{n}(X_i-\mu)^2 \sim \chi_n^2 \\ \chi_n^2 \text{の平均} : n \end{array}\right)$$