* 多次元Gauss分布

$x \in \mathbb{R}^D$ を生成する.

pdf. $\mathcal{N}(x|\mu,\Sigma) = \dfrac{1}{\sqrt{(2\pi)^D \det\Sigma}} \exp\left(-\dfrac{1}{2}(x-\mu)^T\Sigma^{-1}(x-\mu)\right)$.

$\mu \in \mathbb{R}^D$: 平均パラメータ.

$\Sigma \in M_D(\mathbb{R})$: 正定値. 共分散行列.

Remark $A \in M_D(\mathbb{R})$: 対称行列 が 正定値 (positive definite)

$\overset{\text{def.}}{\Longleftrightarrow}$ $^\forall x \in \mathbb{R}^D$, $x \neq 0$ に対し, $x^T A x > 0$.

1次元では,
$x^T A x = Ax^2 > 0$, $^\forall x \neq 0 \to A > 0$.
正定値とは「正の行列」といった
気持ちのもの.

・対数表示.

$\log \mathcal{N}(x|\mu,\Sigma)$

$= -\dfrac{1}{2}\left(D\log 2\pi + \log\det\Sigma\right) - \dfrac{1}{2}(x-\mu)^T\Sigma^{-1}(x-\mu)$

$= -\dfrac{1}{2}\left((x-\mu)^T\Sigma^{-1}(x-\mu) + \log\det\Sigma + D\log 2\pi\right)$.

・多次元正規分布では, 以下が成立:

Prop. $X = (X_1, \cdots, X_D)^T \sim \mathcal{N}(\mu,\Sigma)$ としたとき,

$X_1, \cdots, X_D$: independent $\Longleftrightarrow$ $X_1, \cdots, X_D$ は無相関

つまり, Σ : 対角行列.

pf. $(\Rightarrow)$ 多次元正規分布でなくても成立する.

$(\Leftarrow)$ $\Sigma = \mathrm{diag}(\sigma_1^2, \cdots, \sigma_D^2)$ とする. $\Sigma^{-1} = \mathrm{diag}(\sigma_1^{-2}, \cdots, \sigma_D^{-2})$.

$$\log \mathcal{N}(x|\mu,\Sigma) = -\frac{1}{2}\left((x-\mu)^T\Sigma^{-1}(x-\mu) + \log\det\Sigma + D\log 2\pi\right)$$

$$= -\frac{1}{2}\left(\sum_{d=1}^{D}\frac{(x_d-\mu_d)^2}{\sigma_d^2} + \sum_{d=1}^{D}\log\sigma_d^2 + \sum_{d=1}^{D}\log 2\pi\right)$$

$$= \sum_{d=1}^{D} -\frac{1}{2}\left(\frac{(x_d-\mu_d)^2}{\sigma_d^2} + \log\sigma_d^2 + \log 2\pi\right)$$

$$= \sum_{d=1}^{D}\log\mathcal{N}(x_d|\mu_d,\sigma_d^2) = \log\prod_{d=1}^{D}\mathcal{N}(x_d|\mu_d,\sigma_d^2). \quad \blacksquare$$

・平均.

$$\mathbb{E}[X] = \int_{\mathbb{R}^D} x\,\mathcal{N}(x|\mu,\Sigma)\,dx$$

$$= \int_{\mathbb{R}^D}(x-\mu)\mathcal{N}(x|\mu,\Sigma)\,dx + \mu\underbrace{\int_{\mathbb{R}^D}\mathcal{N}(x|\mu,\Sigma)\,dx}_{=1}$$

$$= \mu + \int_{\mathbb{R}^D}(x-\mu)\frac{1}{\sqrt{(2\pi)^D\det\Sigma}}\exp\left(-\frac{1}{2}(x-\mu)^T\Sigma^{-1}(x-\mu)\right)dx$$

ここで, Σ : 正定値対称 なので, Σ^{-1} : 正定値対称.

$\because$ Σ : sym. なので $\Sigma = \Sigma^T$. $\Sigma^{-1}\Sigma = \Sigma^{-1}\Sigma^T = I_D$

$(\Sigma^T)^{-1} = \Sigma^{-1} = (\Sigma^{-1})^T$. $\therefore \Sigma^{-1}$: sym.

また, A : 正定値 $\Leftrightarrow$ A の全ての固有値が正 に注意して,

$\lambda_1, \cdots, \lambda_D > 0$: Σ の eigenvalue

$\Sigma v = \lambda v$

$v = \lambda\Sigma^{-1}v$

$\Rightarrow \lambda_1^{-1}, \cdots, \lambda_D^{-1} \underbrace{> 0}_{\text{pos. def.}}$: Σ^{-1} : eigenvalue.

$\Sigma^{-1}v = \frac{1}{\lambda}v$.

$U^TU = UU^T = I_D$

このとき，$\exists U$：直交行列 s.t.

Σ^{-1}：対称行列だから．

$$\Sigma^{-1} = U^T \mathrm{diag}(\lambda_1^{-1}, \cdots, \lambda_D^{-1}) U.$$ これより

$$\Sigma^{-1} = U^T \mathrm{diag}(\lambda_1^{-\frac{1}{2}}, \cdots, \lambda_D^{-\frac{1}{2}}) \underbrace{U \cdot U^T}_{=I_D} \mathrm{diag}(\lambda_1^{-\frac{1}{2}}, \cdots, \lambda_D^{-\frac{1}{2}}) U$$

$$=: \Sigma^{-\frac{1}{2}} \cdot \Sigma^{-\frac{1}{2}}$$ とかける．$\left(\Sigma^{-\frac{1}{2}} : \text{pos.def.}\right)$

そこで，$\boldsymbol{z} = \Sigma^{-\frac{1}{2}}(\boldsymbol{x} - \boldsymbol{\mu})$ とおく．すると，

1次元で $z = \frac{x-\mu}{\sqrt{\sigma}}$ とおいたのと対応している

$$\boldsymbol{x} = \Sigma^{\frac{1}{2}} \boldsymbol{z} + \boldsymbol{\mu}.$$

$$(\boldsymbol{x} - \boldsymbol{\mu})^T \Sigma^{-1} (\boldsymbol{x} - \boldsymbol{\mu}) = \left(\Sigma^{-\frac{1}{2}}(\boldsymbol{x} - \boldsymbol{\mu})\right)^T \left(\Sigma^{-\frac{1}{2}}(\boldsymbol{x} - \boldsymbol{\mu})\right)$$

$$= \boldsymbol{z}^T \boldsymbol{z}.$$

変数変換のJacobianは，

$$J(\boldsymbol{z}) = \det\left(\frac{\partial \boldsymbol{x}}{\partial \boldsymbol{z}}\right)$$

ベクトルの微分公式．$\frac{\partial}{\partial \boldsymbol{x}} A\boldsymbol{x} = A^T$

A：対称なら $\frac{\partial}{\partial \boldsymbol{x}} A\boldsymbol{x} = A$.

$$= \det \Sigma^{\frac{1}{2}}$$

$$= \det\left(U^T \mathrm{diag}(\lambda_1^{\frac{1}{2}}, \cdots, \lambda_D^{\frac{1}{2}}) U\right)$$

$\det(AB) = (\det A)(\det B)$

$$= \det U^T \cdot \det\left(\mathrm{diag}(\lambda_1^{\frac{1}{2}}, \cdots, \lambda_D^{\frac{1}{2}})\right) \cdot \det U$$

$$= \det \underbrace{U^T U}_{=I_D} \cdot \left(\lambda_1^{\frac{1}{2}} \cdot \cdots \cdot \lambda_D^{\frac{1}{2}}\right)$$

$$= 1 \cdot (\lambda_1 \cdot \cdots \cdot \lambda_D)^{\frac{1}{2}}$$

$$= (\det \Sigma)^{\frac{1}{2}} \quad (>0)$$

$\therefore\ \mathbb{E}[\mathbb{X}]$

$$= \mu + \frac{1}{\sqrt{(2\pi)^D}} \int_{\mathbb{R}^D} \frac{\Sigma^{\frac{1}{2}} \mathbb{z}}{\cancel{\sqrt{\det \Sigma}}} \exp\left(-\frac{\mathbb{z}^T \mathbb{z}}{2}\right) \cancel{|J(\mathbb{z})|}\, d\mathbb{z}$$

$$= \mu + \frac{1}{\sqrt{(2\pi)^D}} \int_{\mathbb{R}^D} \Sigma^{\frac{1}{2}} \mathbb{z} \exp\left(-\frac{\mathbb{z}^T \mathbb{z}}{2}\right) d\mathbb{z}$$

$$\frac{\partial}{\partial \mathbb{z}} \exp\left(-\frac{\mathbb{z}^T \mathbb{z}}{2}\right) = -\mathbb{z} \exp\left(-\frac{\mathbb{z}^T \mathbb{z}}{2}\right)$$

$$= \mu - \frac{\Sigma^{\frac{1}{2}}}{\sqrt{(2\pi)^D}} \int_{\mathbb{R}^D} \frac{\partial}{\partial \mathbb{z}} \exp\left(-\frac{\mathbb{z}^T \mathbb{z}}{2}\right) d\mathbb{z}$$

ここで、積分項の第i成分について、

$$\int_{\mathbb{R}^D} \frac{\partial}{\partial z_i} \exp\left(-\frac{\mathbb{z}^T \mathbb{z}}{2}\right) d\mathbb{z}$$

$$= \int_{\mathbb{R}^D} \frac{\partial}{\partial z_i} \exp\left(-\frac{1}{2} \sum_{d=1}^{D} z_d^2\right) d\mathbb{z}$$

$$= \int_{\mathbb{R}^D} \frac{\partial}{\partial z_i} \prod_{d=1}^{D} \exp\left(-\frac{z_d^2}{2}\right) d\mathbb{z}$$

d≠iの項とd=iの項に分けてみる。

$$= \underbrace{\prod_{d \neq i} \left(\int_{-\infty}^{\infty} \exp\left(-\frac{z_d^2}{2}\right) dz_d \right)}_{= \mathrm{const.}} \times \int_{-\infty}^{\infty} \frac{\partial}{\partial z_i} \exp\left(-\frac{z_i^2}{2}\right) dz_i$$

$$\propto \left[\exp\left(-\frac{z_i^2}{2}\right) \right]_{-\infty}^{\infty} = 0$$

$$\therefore \int_{\mathbb{R}^D} \frac{\partial}{\partial \mathbb{z}} \exp\left(-\frac{\mathbb{z}^T \mathbb{z}}{2}\right) d\mathbb{z} = 0$$

以上より $\mathbb{E}[\mathbb{X}] = \mu$.

（別の計算方法） こちらの方が楽

$\mathbb{E}[X]$

$$= \mu + \int_{\mathbb{R}^D} (x-\mu) \frac{1}{\sqrt{(2\pi)^D \det \Sigma}} \exp\left(-\frac{1}{2}(x-\mu)^T \Sigma^{-1}(x-\mu)\right) dx$$

ここで, A：対称のとき $\frac{\partial}{\partial x} x^T A x = 2Ax$ を用いると.

↳ 対称でないときは $\frac{\partial}{\partial x} x^T A x = (A+A^T)x$.

$$\frac{\partial}{\partial \mu} \exp\left(-\frac{1}{2}(x-\mu)^T \Sigma^{-1}(x-\mu)\right)$$

合成関数の微分

$$= \exp\left(-\frac{1}{2}(x-\mu)^T \Sigma^{-1}(x-\mu)\right) \cdot \frac{\partial}{\partial \mu}\left(-\frac{1}{2}(x-\mu)^T \Sigma^{-1}(x-\mu)\right)$$

$$= \exp\left(-\frac{1}{2}(x-\mu)^T \Sigma^{-1}(x-\mu)\right)\left(-\frac{1}{\cancel{2}} \cdot \cancel{2}\Sigma^{-1}(x-\mu) \cdot \underbrace{\frac{\partial}{\partial \mu}(x-\mu)}_{=-I_D}\right)$$

$$= \Sigma^{-1}(x-\mu) \exp\left(-\frac{1}{2}(x-\mu)^T \Sigma^{-1}(x-\mu)\right).$$

$$\therefore \int_{\mathbb{R}^D} (x-\mu) \frac{1}{\sqrt{(2\pi)^D \det \Sigma}} \exp\left(-\frac{1}{2}(x-\mu)^T \Sigma^{-1}(x-\mu)\right) dx$$

$$= \Sigma \int_{\mathbb{R}^D} \frac{1}{\sqrt{(2\pi)^D \det \Sigma}} \frac{\partial}{\partial \mu} \exp\left(-\frac{1}{2}(x-\mu)^T \Sigma^{-1}(x-\mu)\right) dx$$

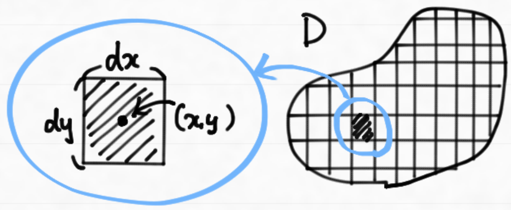
$$= \Sigma \frac{\partial}{\partial \mu} \int_{\mathbb{R}^D} \frac{1}{\sqrt{(2\pi)^D \det \Sigma}} \exp\left(-\frac{1}{2}(x-\mu)^T \Sigma^{-1}(x-\mu)\right) dx$$

$$= \Sigma \frac{\partial}{\partial \mu} \underbrace{\int_{\mathbb{R}^D} \mathcal{N}(x \mid \mu, \Sigma)\, dx}_{=1} = 0$$

$$\therefore \mathbb{E}[X] = \mu.$$

Remark　Jacobian の意味合いについて.（2次元で考える）

・重積分 $\int_D f(x,y)\,dxdy$:

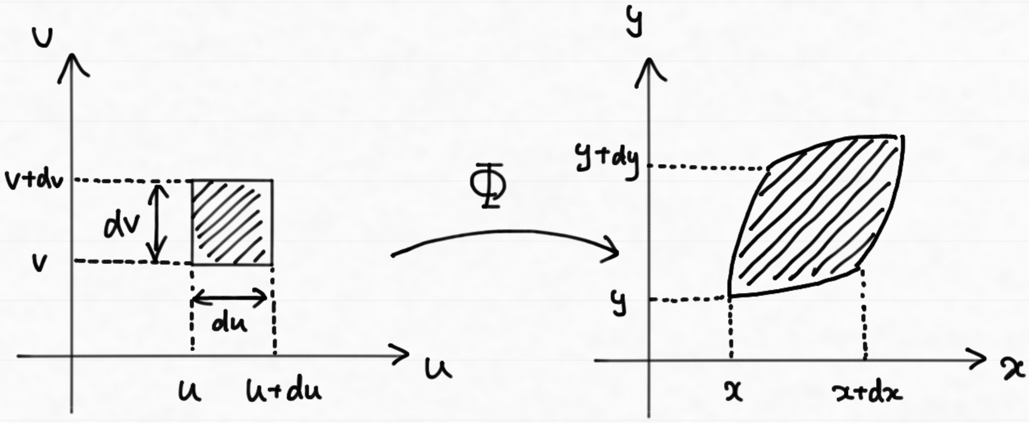


Dを細かく四角形に分割したときにできる微小四角形の面積 $dxdy$ とその中の点 (x,y) での関数 f の値 $f(x,y)$ の積を足しあわせたもの（の極限値）

・変数変換 $\Phi:(u,v)\mapsto(x,y)$ を $x=x(u,v)$, $y=y(u,v)$ とするとき,

u,v-平面の微小長方形 $[u,u+du]\times[v,v+dv]$ は, x,y-平面に写ると

どう歪むか?



変形後の図形は, du, dv が小さいときは近似的に平行四辺形

とみることができる. この面積を $dxdy$ とする. x,y の全微分を考えると,

$$\left\{\begin{aligned} dx &= \frac{\partial x}{\partial u}du+\frac{\partial x}{\partial v}dv \\ dy &= \frac{\partial y}{\partial u}du+\frac{\partial y}{\partial v}dv \end{aligned}\right. \iff \begin{pmatrix} dx \\ dy \end{pmatrix} = \underbrace{\begin{pmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{pmatrix}}_{=:J\ (\text{Jacobi行列})} \begin{pmatrix} du \\ dv \end{pmatrix}.$$

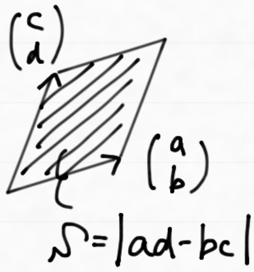
とおく.

これより、

$$\begin{pmatrix} du \\ 0 \end{pmatrix} \mapsto \begin{pmatrix} \frac{\partial x}{\partial u} \\ \frac{\partial y}{\partial u} \end{pmatrix} du , \quad \begin{pmatrix} 0 \\ dv \end{pmatrix} \mapsto \begin{pmatrix} \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial v} \end{pmatrix} dv$$

であるから。

変形後の平行四辺形の面積は、



$$\begin{aligned} dxdy &= \left| \frac{\partial x}{\partial u} du \frac{\partial y}{\partial v} dv - \frac{\partial y}{\partial u} du \frac{\partial x}{\partial v} dv \right| \\ &= \left| \frac{\partial x}{\partial u} \frac{\partial y}{\partial v} - \frac{\partial y}{\partial u} \frac{\partial x}{\partial v} \right| du\,dv \\ &= |\underbrace{\det J}_{\text{Jacobian}}| \, du\,dv . \end{aligned}$$

ということで、$x = x(u,v)$, $y = y(u,v)$ と変数変換するときには

微小四角形の面積が $dxdy = |\det J|\, dudv$ という関係をみたす。

- 多次元の場合も同様のことを考えるとわかる。 $\square$

- $$\mathbb{E}[XX^T] = \int_{\mathbb{R}^D} xx^T \mathcal{N}(x|\mu, \Sigma)\, dx$$

$$= \int_{\mathbb{R}^D} (x-\mu)(x-\mu)^T \mathcal{N}(x|\mu, \Sigma)\, dx + \underbrace{\left(\int_{\mathbb{R}^D} x \mathcal{N}(x|\mu, \Sigma)\, dx \right)}_{=\mu} \mu^T$$

$$+ \mu \underbrace{\left(\int_{\mathbb{R}^D} x^T \mathcal{N}(x|\mu, \Sigma)\, dx \right)}_{=\mu^T} - \mu\mu^T \underbrace{\int_{\mathbb{R}^D} \mathcal{N}(x|\mu, \Sigma)\, dx}_{=1}$$

$$= \mu\mu^T + \int_{\mathbb{R}^D} (x-\mu)(x-\mu)^T \frac{1}{\sqrt{(2\pi)^D \det \Sigma}} \exp\left(-\frac{1}{2}(x-\mu)^T \Sigma^{-1}(x-\mu)\right) dx$$

$$= \mu\mu^T + \frac{\Sigma^{\frac{1}{2}}}{\sqrt{(2\pi)^D}} \left(\int_{\mathbb{R}^D} zz^T \exp\left(-\frac{z^Tz}{2}\right) dz \right) \Sigma^{\frac{1}{2}}.$$

[Claim] $\int_{\mathbb{R}^D} zz^T \exp\left(-\frac{z^Tz}{2}\right) dz = \sqrt{(2\pi)^D} I_D$ を用いると,

$$\mathbb{E}[XX^T] = \mu\mu^T + \frac{\Sigma^{\frac{1}{2}}}{\sqrt{(2\pi)^D}} \sqrt{(2\pi)^D} I_D \Sigma^{\frac{1}{2}} = \mu\mu^T + \Sigma.$$

- 以下, [Claim] を示す.

pf. $\{e_d\}_{d=1}^D$: $\mathbb{R}^D$ の標準基底. $z = \sum_{d=1}^D z_d e_d$ と表せる.

$$\int_{\mathbb{R}^D} zz^T \exp\left(-\frac{z^Tz}{2}\right) dz$$

$$= \int_{\mathbb{R}^D} \left(\sum_{d=1}^D z_d e_d\right)\left(\sum_{d=1}^D z_d e_d\right)^T \exp\left(-\frac{1}{2}\sum_{i=1}^D z_i^2\right) dz_1 \cdots dz_D$$

$$= \int_{\mathbb{R}^D} \left(\sum_{d=1}^D \sum_{d'=1}^D z_d z_{d'} e_d e_{d'}^T\right) \exp\left(-\frac{1}{2}\sum_{i=1}^D z_i^2\right) dz_1 \cdots dz_D$$

$$= \sum_{d=1}^D \sum_{d'=1}^D e_d e_{d'}^T \int_{\mathbb{R}^D} z_d z_{d'} \exp\left(-\frac{1}{2}\sum_{i=1}^D z_i^2\right) dz_1 \cdots dz_D.$$

あとはこの積分を実行する.

$d \neq d'$ のとき、

$$\int_{\mathbb{R}^D} z_d z_{d'} \exp\left(-\frac{1}{2}\sum_{i=1}^{D} z_i^2\right) dz_1 \cdots dz_D$$

$$= -\int_{\mathbb{R}^D} z_d \frac{\partial}{\partial z_{d'}} \exp\left(-\frac{1}{2}\sum_{i=1}^{D} z_i^2\right) dz_1 \cdots dz_D$$

d'以外

$$= -\int_{\mathbb{R}^{D-1}} z_d \left(\int_{-\infty}^{\infty} \frac{\partial}{\partial z_{d'}} \exp\left(-\frac{1}{2}\sum_{i=1}^{D} z_i^2\right) dz_{d'}\right) dz_{\backslash d'}$$

$$= -\int_{\mathbb{R}^{D-1}} z_d \underbrace{\left[\exp\left(-\frac{1}{2}\sum_{i=1}^{D} z_i^2\right)\right]_{-\infty}^{\infty}}_{=0} dz_{\backslash d'}$$

$$= 0.$$

$$\boxed{\begin{aligned}\int_{-\infty}^{\infty} x^2 e^{-\frac{x^2}{2}} dx &= \frac{1}{2}\int_{-\infty}^{\infty} e^{-\frac{x^2}{2}} dx \\ &= \frac{\sqrt{\pi}}{2}.\end{aligned}}$$

$$\therefore \int_{\mathbb{R}^D} \mathbf{z}\mathbf{z}^T \exp\left(-\frac{\mathbf{z}^T\mathbf{z}}{2}\right) d\mathbf{z}$$

$$= \sum_{d=1}^{D} \mathbf{e}_d \mathbf{e}_d^T \int_{\mathbb{R}^D} z_d^2 \exp\left(-\frac{1}{2}\sum_{i=1}^{D} z_i^2\right) dz_1 \cdots dz_D.$$

$$= \sum_{d=1}^{D} \mathbf{e}_d \mathbf{e}_d^T \int_{\mathbb{R}^D} z_d^2 \prod_{i=1}^{D} \exp\left(-\frac{z_i^2}{2}\right) dz_1 \cdots dz_D.$$

$$= \sum_{d=1}^{D} \mathbf{e}_d \mathbf{e}_d^T \left(\prod_{i \neq d} \underbrace{\int_{-\infty}^{\infty} \exp\left(-\frac{z_i^2}{2}\right) dz_i}_{=\sqrt{2\pi}}\right) \underbrace{\int_{-\infty}^{\infty} z_d^2 \exp\left(-\frac{z_d^2}{2}\right) dz_d}_{=\sqrt{2\pi}}$$

$$= \sum_{d=1}^{D} \mathbf{e}_d \mathbf{e}_d^T \sqrt{(2\pi)^D} = \sqrt{(2\pi)^D} I_D.$$

$$\mathbb{V}[X] = \mathbb{E}[XX^T] - \mu\mu^T = \mu\mu^T + \Sigma - \mu\mu^T = \Sigma.$$

(別の計算方法) ← 結構テクニカル。以下ではΣで微分しているが、大人しくΣをCholesky分解した方がよさそう。

$$\mathbb{E}[XX^T]$$

$$= \mu\mu^T + \int_{\mathbb{R}^D} (x-\mu)(x-\mu)^T \frac{1}{\sqrt{(2\pi)^D \det\Sigma}} \exp\left(-\frac{1}{2}(x-\mu)^T\Sigma^{-1}(x-\mu)\right) dx.$$

$$\frac{\partial}{\partial\Sigma}\exp\left(-\frac{1}{2}(x-\mu)^T\Sigma^{-1}(x-\mu)\right)$$ を考えたい。

Def. (Hadamard積)

$A=(a_{ij}), B=(b_{ij}) \in M_{m,n}(\mathbb{R})$ に対し、AとBの

Hadamard積 $A \odot B$ を $A \odot B := (a_{ij}b_{ij})$ で定める。 ← 要素ごとの積 $\square$

Prop. $X \in M_n(\mathbb{R})$. $A \in M_{p,n}(\mathbb{R})$, $B \in M_{n,p}(\mathbb{R})$

(1) Xの各要素が独立のとき、$\frac{\partial}{\partial X}\mathrm{tr}(AXB) = A^TB^T$

(2) X:対称のとき、$\frac{\partial}{\partial X}\mathrm{tr}(AXB) = A^TB^T + BA - (A^TB^T)\odot I_n$.

$\square$

pf. $X=(x_{ij})$, $A=(a_{ij})$, $B=(b_{ij})$ とする。

$$\frac{\partial}{\partial x_{ij}}\mathrm{tr}(AXB) = \frac{\partial}{\partial x_{ij}}\sum_{k=1}^{p}\sum_{s=1}^{n}\sum_{t=1}^{n} a_{ks}x_{st}b_{tk}$$

$$= \sum_{k=1}^{p}\sum_{s=1}^{n}\sum_{t=1}^{n} a_{ks}\frac{\partial x_{st}}{\partial x_{ij}} b_{tk}.$$

(1) $\frac{\partial x_{st}}{\partial x_{ij}} = \delta_{si}\delta_{tj}$ であることから、

$$\frac{\partial}{\partial x_{ij}}\operatorname{tr}(AXB) = \sum_{k=1}^{p}\sum_{s=1}^{n}\sum_{t=1}^{n} a_{ks} b_{tk} \delta_{si} \delta_{tj}$$
$$= \sum_{k=1}^{p} b_{jk} a_{ki}$$
$$= (BA)_{ji} = (A^T B^T)_{ij}$$
$$\therefore \frac{\partial}{\partial X}\operatorname{tr}(AXB) = A^T B^T.$$

(2) $i \geqq j$ とする.

$s \geqq t$ のとき　　$t \geqq s$ のとき　　($s=t$ のとき) 対角項は2回足しているので1回分引く.

$$\frac{\partial x_{st}}{\partial x_{ij}} = \delta_{si}\delta_{tj} + \delta_{sj}\delta_{ti} - \delta_{si}\delta_{tj}\delta_{ij}$$
$$\therefore \frac{\partial}{\partial x_{ij}}\operatorname{tr}(AXB)$$
$$= \sum_{k=1}^{p}\sum_{s=1}^{n}\sum_{t=1}^{n} a_{ks} b_{tk} \delta_{si}\delta_{tj} + \sum_{k=1}^{p}\sum_{s=1}^{n}\sum_{t=1}^{n} a_{ks} b_{tk} \delta_{sj}\delta_{ti}$$
$$- \sum_{k=1}^{p}\sum_{s=1}^{n}\sum_{t=1}^{n} a_{ks} b_{tk} \delta_{si}\delta_{tj}\delta_{ij}$$
$$= \sum_{k=1}^{p} b_{jk} a_{ki} + \sum_{k=1}^{p} b_{ik} a_{kj} - \left(\sum_{k=1}^{p} b_{jk} a_{ki}\right)\delta_{ij}$$
$$= (A^T B^T)_{ij} + (BA)_{ij} - (A^T B^T)_{ij} (I_n)_{ij}$$
$$\therefore \frac{\partial}{\partial X}\operatorname{tr}(AXB) = A^T B^T + BA - (A^T B^T) \odot I_n. \quad \blacksquare$$

<u>Cor.</u> $X \in M_n(\mathbb{R})$. $\mathbf{a}, \mathbf{b} \in \mathbb{R}^n$.

$\mathbf{a}^T X \mathbf{b} = \operatorname{tr}(\mathbf{a}^T X \mathbf{b})$ より, 上の Prop. を用いるとわかる.

(1) X の各要素が独立のとき, $\frac{\partial}{\partial X} \mathbf{a}^T X \mathbf{b} = \mathbf{a}\mathbf{b}^T$

(2) X : 対称のとき, $\frac{\partial}{\partial X} \mathbf{a}^T X \mathbf{b} = \mathbf{a}\mathbf{b}^T + \mathbf{b}\mathbf{a}^T - (\mathbf{a}\mathbf{b}^T) \odot I_n$. $\square$

Prop. $X=(x_{ij})\in M_n(\mathbb{R})$：正則．$\dfrac{\partial X^{-1}}{\partial x_{ij}} = -X^{-1}\dfrac{\partial X}{\partial x_{ij}}X^{-1}$． $\square$

pf. $XX^{-1}=I_n$ の両辺を x_{ij} で微分．積の微分により

$$\frac{\partial X}{\partial x_{ij}}X^{-1} + X\frac{\partial X^{-1}}{\partial x_{ij}} = O.$$

$$\therefore \frac{\partial X^{-1}}{\partial x_{ij}} = -X^{-1}\frac{\partial X}{\partial x_{ij}}X^{-1}.$$ ▨

Prop. $X=(x_{ij})\in M_n(\mathbb{R})$：正則，$X^{-1}=(\xi_{ij})$．

$f: M_n(\mathbb{R})\to\mathbb{R}$：滑らか．

(1) Xの各要素が独立のとき，$\dfrac{\partial f(X)}{\partial X} = -(X^{-1})^T\dfrac{\partial f(X)}{\partial X^{-1}}(X^{-1})^T$.

(2) X：対称のとき，$\tilde{X} := X^{-1}\left(\dfrac{\partial f(X)}{\partial X^{-1}} + \dfrac{\partial f(X)}{\partial X^{-1}}\odot I_n\right)X^{-1}$ として

$$\frac{\partial f(X)}{\partial X} = -\tilde{X} + \frac{1}{2}\tilde{X}\odot I_n.$$ $\square$

pf. (1) $\dfrac{\partial}{\partial x_{ij}}$ それぞれが独立なので，$\dfrac{\partial}{\partial \xi_{kl}}$ それぞれは独立．

chain rule より．$\dfrac{\partial f(X)}{\partial x_{ij}} = \displaystyle\sum_{k=1}^n\sum_{l=1}^n \frac{\partial \xi_{kl}}{\partial x_{ij}}\frac{\partial f(X)}{\partial \xi_{kl}}$． $f(X)$を各ξ_{kl}の関数とみる．

$\dfrac{\partial X^{-1}}{\partial x_{ij}} = -X^{-1}\dfrac{\partial X}{\partial x_{ij}}X^{-1}$ なので，

$$\frac{\partial \xi_{kl}}{\partial x_{ij}} = -\sum_{s=1}^n\sum_{t=1}^n \xi_{ks}\left(\frac{\partial X}{\partial x_{ij}}\right)_{st}\xi_{tl} = -\sum_{s=1}^n\sum_{t=1}^n \xi_{ks}\frac{\partial x_{st}}{\partial x_{ij}}\xi_{tl}.$$

Xの各要素が独立なので，$\dfrac{\partial x_{st}}{\partial x_{ij}} = \delta_{si}\delta_{tj}$.

$$\therefore \frac{\partial \xi_{kl}}{\partial x_{ij}} = -\sum_{s=1}^n\sum_{t=1}^n \xi_{ks}\delta_{si}\delta_{tj}\xi_{tl} = -\xi_{ki}\xi_{jl}.$$

$$\therefore \frac{\partial f(X)}{\partial x_{ij}} = -\sum_{k=1}^n\sum_{l=1}^n \xi_{ki}\frac{\partial f(X)}{\partial \xi_{kl}}\xi_{jl} = -\left((X^{-1})^T\frac{\partial f(X)}{\partial X^{-1}}(X^{-1})^T\right)_{ij}.$$

(2) X: 対称のとき, X^{-1}: 対称.

$$\left(\because \; XX^{-1} = I_n, \; (X^{-1})^T X^T = I_n^T, \; (X^{-1})^T X = I_n. \quad \therefore X^{-1} = (X^{-1})^T \right)$$

このとき独立な ξ_{kl} は $k \geq l$ となる $\frac{n(n+1)}{2}$ コ. $i \geq j$ とすると,

chain rule より, $\dfrac{\partial f(X)}{\partial x_{ij}} = \displaystyle\sum_{k=1}^{n} \sum_{l=1}^{k} \frac{\partial \xi_{kl}}{\partial x_{ij}} \frac{\partial f(X)}{\partial \xi_{kl}} =: \sum_{k \geq l} \frac{\partial \xi_{kl}}{\partial x_{ij}} \frac{\partial f(X)}{\partial \xi_{kl}}$.

やはり, $\dfrac{\partial \xi_{kl}}{\partial x_{ij}} = -\displaystyle\sum_{s=1}^{n} \sum_{t=1}^{n} \xi_{ks} \frac{\partial x_{st}}{\partial x_{ij}} \xi_{tl}$ が成立.

X: 対称より,

$$\frac{\partial x_{st}}{\partial x_{ij}} = \delta_{si}\delta_{tj} + \delta_{sj}\delta_{ti} - \delta_{si}\delta_{tj}\delta_{ij}.$$

$$\therefore \frac{\partial \xi_{kl}}{\partial x_{ij}} = -\sum_{s=1}^{n} \sum_{t=1}^{n} \xi_{ks} (\delta_{si}\delta_{tj} + \delta_{sj}\delta_{ti} - \delta_{si}\delta_{tj}\delta_{ij}) \xi_{tl}$$

$$= -\xi_{ki}\xi_{jl} - \xi_{kj}\xi_{il} + \xi_{ki}\xi_{jl}\delta_{ij}.$$

$$\therefore \frac{\partial f(X)}{\partial x_{ij}} = \sum_{k \geq l} \left(-\xi_{ki}\xi_{jl} - \xi_{kj}\xi_{il} + \xi_{ki}\xi_{jl}\delta_{ij} \right) \frac{\partial f(X)}{\partial \xi_{kl}}$$

和の範囲を $k \geq l$ から広げるため, $l \geq k$ と $l = k$ のときを加え, 半分にする.

$$= \frac{1}{2} \left(\sum_{k=1}^{n} \sum_{l=1}^{n} \left(-\xi_{ki}\xi_{jl} - \xi_{kj}\xi_{il} + \xi_{ki}\xi_{jl}\delta_{ij} \right) \frac{\partial f(X)}{\partial \xi_{kl}} + \sum_{k=1}^{n} \left(-\xi_{ki}\xi_{jk} - \xi_{kj}\xi_{ik} + \xi_{ki}\xi_{jk}\delta_{ij} \right) \frac{\partial f(X)}{\partial \xi_{kk}} \right)$$

X^{-1}: 対称

$$= \frac{1}{2} \left(\sum_{k=1}^{n} \sum_{l=1}^{n} (-2 + \delta_{ij}) \xi_{ik} \frac{\partial f(X)}{\partial \xi_{kl}} \xi_{lj} + \sum_{k=1}^{n} (-2 + \delta_{ij}) \xi_{ik} \frac{\partial f(X)}{\partial \xi_{kk}} \xi_{kj} \right)$$

$$= -\left(\left(X^{-1}\frac{\partial f(X)}{\partial X^{-1}}X^{-1}\right)_{ij} + \left(X^{-1}\left(\frac{\partial f(X)}{\partial X^{-1}}\odot I_n\right)X^{-1}\right)_{ij}\right)$$

$$+ \frac{1}{2}\left(\left(X^{-1}\frac{\partial f(X)}{\partial X^{-1}}X^{-1}\right)_{ij} + \left(X^{-1}\left(\frac{\partial f(X)}{\partial X^{-1}}\odot I_n\right)X^{-1}\right)_{ij}\right)(I_n)_{ij}.$$

$$\therefore \frac{\partial f(X)}{\partial X} = -\left(X^{-1}\frac{\partial f(X)}{\partial X^{-1}}X^{-1} + X^{-1}\left(\frac{\partial f(X)}{\partial X^{-1}}\odot I_n\right)X^{-1}\right)$$

$$+ \frac{1}{2}\left(X^{-1}\frac{\partial f(X)}{\partial X^{-1}}X^{-1} + X^{-1}\left(\frac{\partial f(X)}{\partial X^{-1}}\odot I_n\right)X^{-1}\right)\odot I_n$$

$$= -X^{-1}\left(\frac{\partial f(X)}{\partial X^{-1}} + \frac{\partial f(X)}{\partial X^{-1}}\odot I_n\right)X^{-1}$$

$$+ \frac{1}{2}\left(X^{-1}\left(\frac{\partial f(X)}{\partial X^{-1}} + \frac{\partial f(X)}{\partial X^{-1}}\odot I_n\right)X^{-1}\right)\odot I_n. \qquad \blacksquare$$

<u>Prop.</u> $X=(x_{ij}) \in M_n(\mathbb{R})$: 正則.

(1) Xの各要素が独立のとき, $\dfrac{\partial}{\partial X}\det X = (\det X)(X^{-1})^T$

(2) X : 対称のとき, $\dfrac{\partial}{\partial X}\det X = (\det X)\left(2X^{-1} - X^{-1}\odot I_n\right).$ $\square$

pf. (1) X_{ij} : Xの第i行第j列を除いてできる小行列.

$\Delta_{ij} := (-1)^{i+j}\det X_{ij}$: (i,j)-余因子 とする.

第i行についてのLaplace展開 : $\det X = \sum_{t=1}^{n} x_{it}\Delta_{it}$ を用いて,

$$\frac{\partial}{\partial x_{ij}}\det X = \frac{\partial}{\partial x_{ij}}\sum_{t=1}^{n} x_{it}\Delta_{it} = \sum_{t=1}^{n}\frac{\partial x_{it}}{\partial x_{ij}}\Delta_{it}.$$

Xの各要素が独立のとき, $\dfrac{\partial x_{it}}{\partial x_{ij}} = \delta_{ii}\delta_{tj} = \delta_{tj}$.

$$\therefore \frac{\partial}{\partial x_{ij}}\det X = \sum_{t=1}^{n}\delta_{tj}\Delta_{it} = \Delta_{ij} = (\det X)((X^{-1})^T)_{ij}.$$

(2) $\frac{\partial}{\partial x_{ij}} \det X = \frac{\partial}{\partial x_{ij}} \sum_{t=1}^{n} x_{it}\Delta_{it}$

$x_{ji} = x_{ij}$ が残っている.

$$= \sum_{t=1}^{n} \frac{\partial x_{it}}{\partial x_{ij}}\Delta_{it} + \sum_{t=1}^{n} x_{it}\frac{\partial \Delta_{it}}{\partial x_{ij}} .$$

X: 対称より,

$$\frac{\partial x_{it}}{\partial x_{ij}} = \delta_{ii}\delta_{tj} + \delta_{ij}\delta_{ti} - \delta_{ii}\delta_{tj}\delta_{ij}$$

$$= \delta_{tj} + (\delta_{ti} - \delta_{tj})\delta_{ij}$$

第2項について, $i \neq j$ なら $\delta_{ij} = 0$.
$i = j$ なら $\delta_{ti} - \delta_{tj} = \delta_{ti} - \delta_{ti} = 0$.

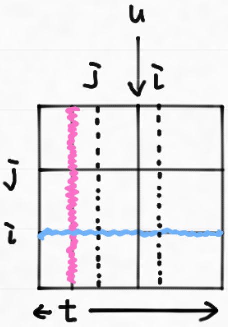
$$= \delta_{tj}$$

$$\therefore \sum_{t=1}^{n} \frac{\partial x_{it}}{\partial x_{ij}}\Delta_{it} = \sum_{t=1}^{n} \delta_{tj}\Delta_{it} = \Delta_{ij}$$

$X^{-1} = (X^{-1})^T$.

$$= (\det X)\left(\left(X^{-1}\right)^T\right)_{ij} = (\det X)\left(X^{-1}\right)_{ij} .$$

$\frac{\partial \Delta_{ii}}{\partial x_{ii}} = 0$ (∵ X_{ii} には x_{ii} がない) である.



$i > j$ として.

$x_{ij} = x_{ji}$

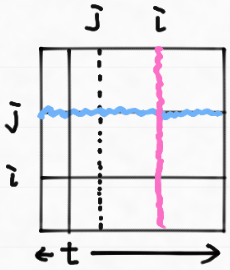
$$\frac{\partial \Delta_{it}}{\partial x_{ij}} = (-1)^{i+t}\frac{\partial}{\partial x_{ij}} \det X_{it} = (-1)^{i+t}\frac{\partial}{\partial x_{ji}} \underline{\det X_{it}}$$

第j行について
Laplace展開

$$= (-1)^{i+t}\frac{\partial}{\partial x_{ji}}\left(\sum_{u=1}^{t-1} x_{ju}(-1)^{j+u}\det(X_{it})_{ju} + \sum_{u=t+1}^{n} x_{ju}(-1)^{j+u-1}\det(X_{it})_{ju}\right)$$

$$= (-1)^{i+t}\left(\sum_{u=1}^{t-1} \delta_{ui}(-1)^{j+u}\det(X_{it})_{ju} + \sum_{u=t+1}^{n} \delta_{ui}(-1)^{j+u-1}\det(X_{it})_{ju}\right)$$

$$= \begin{cases} (-1)^{j+t}\det(X_{it})_{ji} & (i < t) \\ 0 & (i = t) \\ (-1)^{j+t-1}\det(X_{it})_{ji} & (i > t) \end{cases} .$$



∴ $i > j$ のとき

$$\sum_{t=1}^{n} x_{it}\frac{\partial \Delta_{it}}{\partial x_{ij}}$$

$$= \sum_{t=1}^{i-1} x_{it}(-1)^{j+t-1}\det(X_{it})_{ji} + \sum_{t=i+1}^{n} x_{it}(-1)^{j+t}\det(X_{it})_{ji}$$

$$= (-1)^{i+j}\left(\sum_{t=1}^{i-1} x_{it}(-1)^{i+t-1}\det(X_{ji})_{it} + \sum_{t=i+1}^{n} x_{it}(-1)^{i+t}\det(X_{ji})_{it}\right)$$

$$= (-1)^{j+i}\det X_{ji} = \Delta_{ji} = (\det X)(X^{-1})_{ij}$$

$$\therefore \frac{\partial}{\partial x_{ij}}\det X = (\det X)(X^{-1})_{ij} + (\det X)(X^{-1})_{ij}(1-\delta_{ij})$$

$$= (\det X)\left(2(X^{-1})_{ij} - (X^{-1})_{ij}(I_n)_{ij}\right).$$

$$\therefore \frac{\partial}{\partial X}\det X = (\det X)\left(2X^{-1} - X^{-1}\odot I_n\right).$$ ▨

- $C(\Sigma) := \dfrac{1}{\sqrt{(2\pi)^D\det\Sigma}}$ とおく. $y := x-\mu$ と変数変換するとき

Jacobian は $J = \det I_D = 1$.

$$C(\Sigma)\int_{\mathbb{R}^D}(x-\mu)(x-\mu)^T\exp\left(-\frac{1}{2}(x-\mu)^T\Sigma^{-1}(x-\mu)\right)dx.$$

$$= C(\Sigma)\int_{\mathbb{R}^D} yy^T\exp\left(-\frac{1}{2}y^T\Sigma^{-1}y\right)dy.$$

ここで, $\dfrac{\partial}{\partial\Sigma}\exp\left(-\dfrac{1}{2}y^T\Sigma^{-1}y\right)$ ↓ chain rule

$$= \exp\left(-\frac{1}{2}y^T\Sigma^{-1}y\right)\times\left(-\frac{1}{2}\frac{\partial}{\partial\Sigma}\left(y^T\Sigma^{-1}y\right)\right).$$

$f(\Sigma) := y^T \Sigma^{-1} y$ とおいて.

$$\frac{\partial}{\partial \Sigma}\left(y^T \Sigma^{-1} y\right)$$

$$= -\Sigma^{-1}\left(\frac{\partial f}{\partial \Sigma^{-1}} + \frac{\partial f}{\partial \Sigma^{-1}} \odot I_D\right)\Sigma^{-1}$$

$$+ \frac{1}{2}\left(\Sigma^{-1}\left(\frac{\partial f}{\partial \Sigma^{-1}} + \frac{\partial f}{\partial \Sigma^{-1}} \odot I_D\right)\Sigma^{-1}\right) \odot I_n .$$

$\dfrac{\partial f}{\partial \Sigma^{-1}} = yy^T + yy^T - yy^T \odot I_D = 2yy^T - yy^T \odot I_D$ なので,

$$\frac{\partial f}{\partial \Sigma^{-1}} + \frac{\partial f}{\partial \Sigma^{-1}} \odot I_D = 2yy^T - yy^T \odot I_D + \left(2yy^T - yy^T \odot I_D\right) \odot I_D$$

$$= 2yy^T \cancel{- yy^T \odot I_D} + \cancel{2yy^T \odot I_D} \cancel{- yy^T \odot I_D}$$

$$= 2yy^T .$$

$$\therefore \frac{\partial}{\partial \Sigma}\left(y^T \Sigma^{-1} y\right) = -2\Sigma^{-1} yy^T \Sigma^{-1} + \left(\Sigma^{-1} yy^T \Sigma^{-1}\right) \odot I_D$$

$$\therefore \frac{\partial}{\partial \Sigma} \exp\left(-\frac{1}{2} y^T \Sigma^{-1} y\right)$$

$$= \Sigma^{-1} yy^T \Sigma^{-1} \exp\left(-\frac{1}{2} y^T \Sigma^{-1} y\right)$$

$$- \frac{1}{2}\left(\Sigma^{-1} yy^T \Sigma^{-1}\right) \odot I_D \exp\left(-\frac{1}{2} y^T \Sigma^{-1} y\right)$$

$$\therefore yy^T \exp\left(-\frac{1}{2} y^T \Sigma^{-1} y\right)$$

$$= \Sigma\left(\frac{\partial}{\partial \Sigma} \exp\left(-\frac{1}{2} y^T \Sigma^{-1} y\right)\right)\Sigma$$

$$+ \frac{1}{2} \Sigma\left(\left(\Sigma^{-1} yy^T \Sigma^{-1}\right) \odot I_D\right)\Sigma \exp\left(-\frac{1}{2} y^T \Sigma^{-1} y\right).$$

$$\therefore\ C(\Sigma)\int_{\mathbb{R}^D} yy^T \exp\left(-\frac{1}{2}y^T\Sigma^{-1}y\right)dy.$$

$$= C(\Sigma)\Sigma\left(\int_{\mathbb{R}^D}\frac{\partial}{\partial\Sigma}\exp\left(-\frac{1}{2}y^T\Sigma^{-1}y\right)dy\right)\Sigma$$

$$+\frac{C(\Sigma)}{2}\Sigma\left(\int_{\mathbb{R}^D}(\Sigma^{-1}yy^T\Sigma^{-1})\odot I_D\exp\left(-\frac{1}{2}y^T\Sigma^{-1}y\right)dy\right)\Sigma.$$

$$\int_{\mathbb{R}^D}\frac{\partial}{\partial\Sigma}\exp\left(-\frac{1}{2}y^T\Sigma^{-1}y\right)dy$$

$$=\frac{\partial}{\partial\Sigma}\left(C(\Sigma)^{-1}\underbrace{\int_{\mathbb{R}^D}C(\Sigma)\exp\left(-\frac{1}{2}y^T\Sigma^{-1}y\right)dy}_{=1}\right)$$

$$=\frac{\partial}{\partial\Sigma}\sqrt{(2\pi)^D\det\Sigma}$$

$$=\sqrt{(2\pi)^D}\cdot\frac{1}{2}\cdot\frac{1}{\sqrt{\det\Sigma}}\cdot(\det\Sigma)\left(2\Sigma^{-1}-\Sigma^{-1}\odot I_D\right)$$

$$= C(\Sigma)^{-1}\left(\Sigma^{-1}-\frac{1}{2}\Sigma^{-1}\odot I_D\right).$$

$$\int_{\mathbb{R}^D}(\Sigma^{-1}yy^T\Sigma^{-1})\odot I_D\exp\left(-\frac{1}{2}y^T\Sigma^{-1}y\right)dy$$

$$=\left(\int_{\mathbb{R}^D}(\Sigma^{-1}yy^T\Sigma^{-1})\exp\left(-\frac{1}{2}y^T\Sigma^{-1}y\right)dy\right)\odot I_D.$$

$z=\Sigma^{-\frac{1}{2}}y$ とおくと，この変換のJacobianは $(\det\Sigma)^{\frac{1}{2}}$.

$$\therefore\int_{\mathbb{R}^D}(\Sigma^{-1}yy^T\Sigma^{-1})\exp\left(-\frac{1}{2}y^T\Sigma^{-1}y\right)dy$$

$$=\Sigma^{-\frac{1}{2}}\left(\int_{\mathbb{R}^D}zz^T\exp\left(-\frac{z^Tz}{2}\right)(\det\Sigma)^{\frac{1}{2}}dz\right)\Sigma^{-\frac{1}{2}}$$

$$=\Sigma^{-\frac{1}{2}}\sqrt{(2\pi)^D}\sqrt{\det\Sigma}\,I_D\,\Sigma^{-\frac{1}{2}}=C(\Sigma)^{-1}\Sigma^{-1}.$$

$$\therefore\ C(\Sigma)\int_{\mathbb{R}^D} yy^T \exp\left(-\frac{1}{2}y^T\Sigma^{-1}y\right)dy.$$

$$= \Sigma\left(\Sigma^{-1}-\frac{1}{2}\Sigma^{-1}\odot I_D\right)\Sigma+\frac{1}{2}\Sigma\left(\Sigma^{-1}\odot I_D\right)\Sigma$$

$$= \Sigma - \cancel{\frac{1}{2}\Sigma\left(\Sigma^{-1}\odot I_D\right)\Sigma} + \cancel{\frac{1}{2}\Sigma\left(\Sigma^{-1}\odot I_D\right)\Sigma}$$

$$= \Sigma.$$

$$\therefore\ \mathbb{E}[XX^T] = \mu\mu^T + \Sigma.$$

- 実は，Σで微分するかわりにΣ^{-1}で微分する方が計算は楽．

$$\frac{\partial}{\partial\Sigma^{-1}}\exp\left(-\frac{1}{2}y^T\Sigma^{-1}y\right)$$

$$= \exp\left(-\frac{1}{2}y^T\Sigma^{-1}y\right)\left(-\frac{1}{2}\left(2yy^T - yy^T\odot I_D\right)\right)$$

$$= -yy^T\exp\left(-\frac{1}{2}y^T\Sigma^{-1}y\right)+\frac{1}{2}\left(yy^T\odot I_D\right)\exp\left(-\frac{1}{2}y^T\Sigma^{-1}y\right)$$

$$\therefore\ C(\Sigma)\int_{\mathbb{R}^D} yy^T\exp\left(-\frac{1}{2}y^T\Sigma^{-1}y\right)dy$$

$$= -C(\Sigma)\int_{\mathbb{R}^D}\frac{\partial}{\partial\Sigma^{-1}}\exp\left(-\frac{1}{2}y^T\Sigma^{-1}y\right)dy$$

$$+\frac{1}{2}C(\Sigma)\int_{\mathbb{R}^D}\left(yy^T\odot I_D\right)\exp\left(-\frac{1}{2}y^T\Sigma^{-1}y\right)dy$$

$$= -C(\Sigma)\frac{\partial}{\partial\Sigma^{-1}}C(\Sigma)^{-1}$$

$$+\frac{1}{2}C(\Sigma)\left(\Sigma^{\frac{1}{2}}\left(\int_{\mathbb{R}^D}\Sigma^{-\frac{1}{2}}yy^T\Sigma^{-\frac{1}{2}}\exp\left(-\frac{1}{2}y^T\Sigma^{-1}y\right)dy\right)\Sigma^{\frac{1}{2}}\right)\odot I_D$$

$$= -\cancel{C(\Sigma)}\cancel{\sqrt{(2\pi)^D}}\left(-\frac{1}{2}\frac{1}{\cancel{\left(\sqrt{\det\Sigma^{-1}}\right)^3}}\cancel{\det\Sigma^{-1}}\left(2\Sigma - \Sigma\odot I_D\right)\right)$$

$$+\frac{1}{2}\cancel{C(\Sigma)}\left(\Sigma^{\frac{1}{2}}\cancel{\sqrt{(2\pi)^D}\sqrt{\det\Sigma}}\, I_D\,\Sigma^{\frac{1}{2}}\right)\odot I_D$$

$$= \Sigma - \cancel{\frac{1}{2}\Sigma\odot I_D} + \cancel{\frac{1}{2}\Sigma\odot I_D}$$

$$= \Sigma.$$

- ΣをCholesky分解するのがおそらく一番楽か？

- エントロピー.

$$H[\mathcal{N}(x|\mu,\Sigma)] = -\mathbb{E}[\log\mathcal{N}(x|\mu,\Sigma)]$$

$$= \frac{1}{2}\left(\mathbb{E}\left[(x-\mu)^T\Sigma^{-1}(x-\mu)\right] + \log\det\Sigma + D\log 2\pi\right).$$

$$\mathbb{E}\left[(x-\mu)^T\Sigma^{-1}(x-\mu)\right]$$

$$= \mathbb{E}\left[\mathrm{tr}\left((x-\mu)^T\Sigma^{-1}(x-\mu)\right)\right]$$

$$x^TAx = \mathrm{tr}(x^TAx) = \mathrm{tr}(Axx^T)$$

$$= \mathbb{E}\left[\mathrm{tr}\left(\Sigma^{-1}(x-\mu)(x-\mu)^T\right)\right]$$

$$= \mathrm{tr}\left(\mathbb{E}\left[\Sigma^{-1}(x-\mu)(x-\mu)^T\right]\right)$$

$$= \mathrm{tr}\left(\Sigma^{-1}\mathbb{E}\left[(x-\mu)(x-\mu)^T\right]\right) = \mathrm{tr}\left(\Sigma^{-1}\Sigma\right) = \mathrm{tr}\left(I_D\right) = D.$$

$$\therefore H[\mathcal{N}(x|\mu,\Sigma)] = \frac{1}{2}\left(\log\det\Sigma + D\left(\log 2\pi + 1\right)\right).$$

・ $p(x)=\mathcal{N}(x|\mu,\Sigma)$ と $q(x)=\mathcal{N}(x|\hat{\mu},\hat{\Sigma})$ の KL-divergence.

$$\mathrm{KL}[q(x)\|p(x)] = -\mathrm{H}[\mathcal{N}(x|\hat{\mu},\hat{\Sigma})] - \mathbb{E}_q[\log\mathcal{N}(x|\mu,\Sigma)].$$

$$\mathbb{E}_q[\log\mathcal{N}(x|\mu,\Sigma)]$$

$$= -\frac{1}{2}\left(\mathbb{E}_q[(x-\mu)^T\Sigma^{-1}(x-\mu)] + \log\det\Sigma + D\log 2\pi\right)$$

$$\mathbb{E}_q[(x-\mu)^T\Sigma^{-1}(x-\mu)]$$

$$= \mathbb{E}_q\left[\mathrm{tr}\left((x-\mu)^T\Sigma^{-1}(x-\mu)\right)\right] = \mathbb{E}_q\left[\mathrm{tr}\left((x-\mu)(x-\mu)^T\Sigma^{-1}\right)\right]$$

$$= \mathbb{E}_q\left[\mathrm{tr}\left((x-\hat{\mu}+\hat{\mu}-\mu)(x-\hat{\mu}+\hat{\mu}-\mu)^T\Sigma^{-1}\right)\right]$$

$$= \mathbb{E}_q\left[\mathrm{tr}\left(\left((x-\hat{\mu})(x-\hat{\mu})^T + (x-\hat{\mu})(\hat{\mu}-\mu)^T + (\hat{\mu}-\mu)(x-\hat{\mu})^T + (\hat{\mu}-\mu)(\hat{\mu}-\mu)^T\right)\Sigma^{-1}\right)\right]$$

$$= \mathrm{tr}\left(\mathbb{E}_q\left[\left((x-\hat{\mu})(x-\hat{\mu})^T + (x-\hat{\mu})(\hat{\mu}-\mu)^T + (\hat{\mu}-\mu)(x-\hat{\mu})^T + (\hat{\mu}-\mu)(\hat{\mu}-\mu)^T\right)\Sigma^{-1}\right]\right)$$

$$= \mathrm{tr}\left(\left(\mathbb{E}_q[(x-\hat{\mu})(x-\hat{\mu})^T] + (\hat{\mu}-\mu)(\hat{\mu}-\mu)^T\right)\Sigma^{-1}\right)$$

$$= \mathrm{tr}\left(\left((\mu-\hat{\mu})(\mu-\hat{\mu})^T + \hat{\Sigma}\right)\Sigma^{-1}\right).$$

$$\therefore\ \mathrm{KL}[q(x)\|p(x)]$$

$$= -\mathrm{H}[\mathcal{N}(x|\hat{\mu},\hat{\Sigma})] - \mathbb{E}_q[\log\mathcal{N}(x|\mu,\Sigma)].$$

$$= -\frac{1}{2}\left(\log\det\hat{\Sigma} + D\left(\cancel{\log 2\pi} + 1\right)\right)$$

$$+\frac{1}{2}\left(\mathrm{tr}\left(\left((\mu-\hat{\mu})(\mu-\hat{\mu})^T + \hat{\Sigma}\right)\Sigma^{-1}\right) + \log\det\Sigma + \cancel{D\log 2\pi}\right)$$

$$= -\frac{1}{2}\left(\mathrm{tr}\left(\left((\mu-\hat{\mu})(\mu-\hat{\mu})^T + \hat{\Sigma}\right)\Sigma^{-1}\right) + \log\frac{\det\Sigma}{\det\hat{\Sigma}} - D\right).$$