Hesse Geometry

Chapter 6

6.1 確率分布族の双対構造

def 6.1.1

$\mathcal{X}$: discrete (確率空間 contable) set or $\mathbb{R}^m$, $\mathcal{X}$ 上の func $p(x)$ が

(i) $\forall x \in \mathcal{X}$, $p(x) \geq 0$

(ii) $\mathcal{X}$: discrete $\Rightarrow \sum_{x \in \mathcal{X}} p(x) = 1$

$\mathcal{X}$: $\mathbb{R}^m$ $\Rightarrow \int_{\mathcal{X}} p(x) = 1$.

memo categorical な prob.

をみたすとき prob dist

記号の簡約化の為に discrete でも $\int$ を用いる

$\mathcal{X}$ 上の func $f(x)$ に対して exp value $E[f]$

$$E[f] := \int_{\mathcal{X}} f(x) p(x) dx$$

p に対する期待値作用素 $E[\sim] := \int p [\sim] dx$

$\mathcal{X}$ 上に $\lambda := [\lambda^1, \cdots, \lambda^n] \in \Delta$ を parameter とする prob dist の族

(パラメタ空間) (次元 : finite)

$\mathcal{P} := \{ p(x;\lambda) \mid \lambda \in \Delta \}$ が与えられて次の条件をみたすとする

(P_1) : $\Delta \underset{op}{\subseteq} \mathbb{R}^n$ 正規分布なら μ, σ の2つ ($\mathbb{R} \times \mathbb{R}_{\geq 0}$)

(P_2) : $p(x;\lambda)$ は λ に関して C^∞ - class Δ : global coordinate

(P_3) : x に関する int と λ^i に関する diff は交換可能

def 6.1.2

$\mathcal{P} := \{ p(x;\lambda) \mid \lambda \in \Delta \}$ を $(P_1) \sim (P_3)$ をみたす prob dist の族

$p_\lambda := p(x;\lambda)$ に関して期待値をとる作用素 E_λ $E_\lambda[-] = \int p(x) [\sim] dx$

$l_\lambda := l(x;\lambda) := \ln p(x;\lambda)$ とするとき. (対数尤度関数)

$$g_{ij}(\lambda) := E_\lambda \left[\frac{\partial l_\lambda}{\partial \lambda^i} \frac{\partial l_\lambda}{\partial \lambda^j} \right] = \int_{\mathcal{X}} \frac{\partial l}{\partial \lambda^i}(x;\lambda) \frac{\partial l}{\partial \lambda^j}(x;\lambda) p(x;\lambda) dx$$

で def される行列 $g := [g_{ij}]_{ij}$ を Fisher 情報行列

$\int_{\mathcal{X}} p(x;\lambda)\,dx = 1$ の両辺を λ^i で微分して

$$0 = \frac{\partial}{\partial \lambda^i}\int_{\mathcal{X}} p(x;\lambda)\,dx \underset{(P_3)}{=} \int_{\mathcal{X}} \frac{\partial p}{\partial \lambda^i}(x;\lambda)\,dx$$

$$= \int_{\mathcal{X}} \frac{\partial l}{\partial \lambda^i}(x;\lambda)\, p(x;\lambda)\,dx \quad \left(= \int \frac{\partial l}{\partial p}\frac{\partial p}{\partial \lambda^i} p = \int \underbrace{\frac{\partial}{\partial p} \ln p}_{=\frac{1}{p}} \frac{\partial p}{\partial \lambda^i} p = \int \frac{\partial p}{\partial \lambda^i} \right)$$

より $E_\lambda\left[\frac{\partial l_\lambda}{\partial \lambda^i}\right] = 0$

λ^j で更に微分して

$$0 = \frac{\partial}{\partial \lambda^j}\int_{\mathcal{X}} \frac{\partial l}{\partial \lambda^i}(x;\lambda)\, p(x;\lambda)\,dx$$

$$= \int_{\mathcal{X}} \frac{\partial}{\partial \lambda^j}\left(\frac{\partial l}{\partial \lambda^i}(x;\lambda)\, p(x;\lambda) \right) dx$$

$$= \int_{\mathcal{X}} \left(\frac{\partial^2 l}{\partial \lambda^j \partial \lambda^i}(x;\lambda) \right) p(x;\lambda)\,dx + \int_{\mathcal{X}} \frac{\partial l}{\partial \lambda^i}(x;\lambda)\frac{\partial p}{\partial \lambda^j}(x;\lambda)\,dx$$

$$= \int_{\mathcal{X}} \frac{\partial^2 l}{\partial \lambda^j \partial \lambda^i}(x;\lambda)\, p(x;\lambda)\,dx + \underbrace{\int_{\mathcal{X}} \frac{\partial l}{\partial \lambda^i}(x;\lambda)\frac{\partial l}{\partial \lambda^j}(x;\lambda)\, p(x;\lambda)\,dx}_{g_{ij}}$$

より $g_{ij} = -E_\lambda\left[\frac{\partial^2 l_\lambda}{\partial \lambda^i \partial \lambda^j}\right] = -\int_{\mathcal{X}} \frac{\partial^2 l}{\partial \lambda^j \partial \lambda^i}(x;\lambda)\, p(x;\lambda)\,dx$

n 次元 vect $[c^i]$ に対して ($c^i \neq 0$)

$$\sum_{i,j} g_{ij} c^i c^j = \sum_{i,j} \int_{\mathcal{X}} \frac{\partial l}{\partial \lambda^i}\frac{\partial l}{\partial \lambda^j} p(x;\lambda)\, c^i c^j\,dx$$

0 になりうる.

$$= \int_{\mathcal{X}} \left(\sum_i c^i \frac{\partial l}{\partial \lambda^i}(x;\lambda) \right)^2 p(x;\lambda)\,dx \geq 0$$

よって g は Λ 上で半正定値であるが，さらに prob dist の族 $\mathcal{S}$ が

(P_4) $g = [g_{ij}(\lambda)]$ が domain Λ 上で正定値 $X \neq 0$ に対して $g(X, X) > 0$

も更に仮定する.

$$\sum_{i,j}\left(-\int_{\mathcal{X}} \frac{\partial^2 l}{\partial \lambda^i \partial \lambda^j}(x;\lambda)\, p(x;\lambda) \right) c^i c^j$$

Fisher 情報行列 $g = [g_{ij}]$ を Δ 上の Riemann metric とみなして

その Levi-Civita connection の Christoffel Γ^i_{jk} を求める.

$$\Gamma_{kij} = g_{kp}\Gamma^p_{ij} = \frac{1}{2}\left(\frac{\partial g_{ik}}{\partial\lambda^j} + \frac{\partial g_{jk}}{\partial\lambda^i} - \frac{\partial g_{ji}}{\partial\lambda^k}\right)$$ を求める

$$\frac{\partial g_{ij}}{\partial\lambda^k} = \frac{\partial}{\partial\lambda^k}\left(\int_{\mathcal{X}} \frac{\partial l_\lambda}{\partial\lambda^i}\frac{\partial l_\lambda}{\partial\lambda^j}\, p\, dx\right)$$

$$= \int\left(\frac{\partial^2 l_\lambda}{\partial\lambda^k\partial\lambda^i}\frac{\partial l_\lambda}{\partial\lambda^j}p + \frac{\partial l_\lambda}{\partial\lambda^i}\frac{\partial^2 l_\lambda}{\partial\lambda^k\partial\lambda^j}p + \frac{\partial l_\lambda}{\partial\lambda^i}\frac{\partial l_\lambda}{\partial\lambda^j}\frac{\partial p}{\partial\lambda^k}\right)dx$$

$$= \int\left(\frac{\partial^2 l_\lambda}{\partial\lambda^k\partial\lambda^i}\frac{\partial l_\lambda}{\partial\lambda^j}p + \frac{\partial l_\lambda}{\partial\lambda^i}\frac{\partial^2 l_\lambda}{\partial\lambda^k\partial\lambda^j}p + \frac{\partial l}{\partial\lambda^i}\frac{\partial l_\lambda}{\partial\lambda^j}\frac{\partial l}{\partial\lambda^k}p\right)dx$$

$$= E_\lambda\left[\frac{\partial^2 l_\lambda}{\partial\lambda^k\partial\lambda^i}\frac{\partial l_\lambda}{\partial\lambda^j}\right] + E_\lambda\left[\frac{\partial l_\lambda}{\partial\lambda^i}\frac{\partial^2 l_\lambda}{\partial\lambda^k\partial\lambda^j}\right] + E_\lambda\left[\frac{\partial l}{\partial\lambda^i}\frac{\partial l_\lambda}{\partial\lambda^j}\frac{\partial l}{\partial\lambda^k}\right]$$

$$2\Gamma_{kij} = \left(E_\lambda[ji,k] + E_\lambda[\cancel{i,jk}] + E_\lambda[i,k,j]\right)$$

$$+\left(E_\lambda[ij,k] + E_\lambda[\cancel{j,ik}] + E[\cancel{j,i,k}]\right)$$

$$-\left(E_\lambda[\cancel{ki,j}] + E_\lambda[\cancel{i,kj}] + E[\cancel{i,j,k}]\right)$$

$$\therefore\ \underbrace{\Gamma_{kij}}_{} \overset{=\Gamma_{kji}}{} = E_\lambda\left[\frac{\partial^2 l_\lambda}{\partial\lambda^i\partial\lambda^j}\frac{\partial l_\lambda}{\partial\lambda^k}\right] + \frac{1}{2}E_\lambda\left[\frac{\partial l_\lambda}{\partial\lambda^i}\frac{\partial l_\lambda}{\partial\lambda^j}\frac{\partial l_\lambda}{\partial\lambda^k}\right]$$

→ torsion free 性

def より明らか

$$T_{ijk} := \frac{1}{2}E_\lambda\left[\frac{\partial l_\lambda}{\partial\lambda^i}\frac{\partial l_\lambda}{\partial\lambda^j}\frac{\partial l_\lambda}{\partial\lambda^k}\right]$$ は3階共変対称 tensor なので

$$\Gamma(t)_{kij} := \Gamma_{kij} - tT_{kij},\quad \Gamma^i_{jk}(t) := g^{ip}\Gamma(t)_{pjk}$$

とおけば $\Gamma^i_{jk}(t)$ は torsion free $\nabla(t)$ を定める.

$$\left(\begin{aligned}\Gamma^i_{jk}(t) - \Gamma^i_{kj}(t) &= g^{ip}\Gamma(t)_{pjk} - g^{ip}\Gamma(t)_{pkj}\\ &= g^{ip}(\Gamma_{pjk} - tT_{pjk}) - g^{ip}(\Gamma_{pkj} - tT_{pkj})\\ &= 0\end{aligned}\right)$$

$$\Gamma(t)_{jki} + \Gamma(-t)_{ikj} = \Gamma_{jki} - t\cancel{T_{jki}} + \Gamma_{ikj} + t\cancel{T_{ikj}}$$

$$= E_\lambda[ik,j] + \frac{1}{2}E_\lambda[j,k,i] + E_\lambda[kj,i] + \frac{1}{2}E_\lambda[k,j,i]$$

$$= E[ki,j] + E[i,kj] + E[i,j,k]$$

$$= \partial_k g_{ij}$$ $X\to\partial_k,\ Y\to\partial_i,\ Z\to\partial_j$

より $Xg(Y,Z) = g(\nabla^{(\alpha)}_X Y, Z) + g(Y, \nabla^{(-\alpha)}_X Z)$　　Codazzi mfd (stru)

となるので $\nabla^{(\alpha)}, \nabla^{(-\alpha)}$ は Fisher metric g に関して dual connect

def 6.1.3

prob dist の族 $\mathcal{P} = \{p(x;\theta) \mid \theta \in \Theta\}$ が $\mathcal{X}$ 上の func $C(x)$,

$F_1(x), \cdots, F_n(x)$, Θ 上の func $\varphi(\theta)$ によって

$$p(x;\theta) = \exp\Big(\underbrace{C(x) + F_i(x)\theta^i}_{\{\theta^i\}\text{の線形部分}} - \varphi(\theta)\Big)$$

で表されているとき、$\mathcal{P}$ を指数型分布族

def 2.5.1
mfd M 上の torsion free D と Riemann metric g の pair (D,g) が
Codazzi eq をみたすとき (D,g) を Codazzi stru, (M,D,g) を Codazzi mfd
lemma 2.5.1 で def された D' を D の g に関する dual connect
(D',g) を (D,g) の dual Codazzi stru

$\mathcal{P}$ が exp型分布族に対して　　$p=0$ とならない限り p は exp型

$$\frac{\partial l_\theta}{\partial \theta^i}(x;\theta) = \frac{\partial}{\partial \theta^i}\ln p(x;\theta) = F_i(x) - \frac{\partial \varphi}{\partial \theta^i}$$

$$\frac{\partial^2 l_\theta}{\partial \theta^j \partial \theta^i}(x;\theta) = -\frac{\partial \varphi}{\partial \theta^j \partial \theta^i}$$

$$\int \exp\big(C(x) + F_i(x)\theta^i - \varphi(\theta)\big) = 1$$

$$\Gamma^{(+1)}_{ki,j} = \Gamma_{ki,j} - T_{ki,j} = E_\theta[ij,k]$$

$C(x) = F_i = 0$ でなくとも $C + F_i\theta^i$ は linear なので

$$= \int_{\mathcal{X}} \frac{\partial^2 l_\theta}{\partial \theta^i \partial \theta^j}\frac{\partial l_\theta}{\partial \theta^k} p(x;\theta)\,dx$$

$\varphi = -\ln p$ の凸性のみが重要

$$= -\frac{\partial^2 \varphi}{\partial \theta^i \partial \theta^j}\int_{\mathcal{X}}\frac{\partial l_\theta}{\partial \theta^k} p(x;\theta)\,dx$$

→ $\ln p$ の凹性が問題になる。

$$= -\frac{\partial^2 \varphi}{\partial \theta^i \partial \theta^j}\underbrace{E_\theta\left[\frac{\partial l_\theta}{\partial \theta^k}\right]}_{=0} = 0$$

どのくらい強い制限なのか、

よって $\Gamma^{(+1)}_{ki,j}$ は flat で $\{\theta^i\}$ は $\nabla^{(+1)}$ に関する affine chart.

$$g_{ij} = -\int_{\mathcal{X}} \frac{\partial^2 l_\theta}{\partial \theta^i \partial \theta^j} p(x;\theta)\,dx = \int \frac{\partial^2 \varphi(\theta)}{\partial \theta^i \partial \theta^j} p(x;\theta)\,dx$$

$$= \frac{\partial^2 \varphi}{\partial \theta^i \partial \theta^j}\int p(x;\theta)\,dx = \frac{\partial^2 \varphi}{\partial \theta^i \partial \theta^j}$$

→ φ は凸関数という制限が必要

より $\left(\nabla^{(+1)}, g = \left[\frac{\partial^2 \varphi}{\partial \theta^i \partial \theta^j}\right]\right)$ は Θ 上の Hesse stru.

e.g 6.1.1　1次元正規分布

$\mathcal{X}=\mathbb{R}$ 上の prob dist 族　　non cptな hyp space

$$p(x;\lambda) := \frac{1}{\sqrt{2\pi}\sigma}\exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right),\quad \lambda\in\Lambda := \{(\mu,\sigma) \mid \mu\in\mathbb{R},\ \sigma\in\mathbb{R}_{>0}\}$$

を平均がμ, 標準偏差がσの1次元正規分布という.（パラメタ2つ）

$$p(x;\lambda) = \frac{1}{\sqrt{2\pi}\sigma}\exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$$

$$= \ln\exp\left(\frac{1}{\sqrt{2\pi}\sigma}\right)\exp\left(-\frac{x^2-2x\mu+\mu^2}{2\sigma^2}\right)$$

$$= \exp\left(-x^2\frac{1}{2\sigma^2} + x\frac{\mu}{\sigma^2} - \left(\frac{\mu^2}{2\sigma^2} + \ln\sqrt{2\pi}\sigma\right)\right)$$

$F_1(x) := -x^2$, $F_2(x) := x$, $\theta^1 := 1/2\sigma^2$, $\theta^2 := \mu/\sigma^2$,

$$\varphi(\theta) = \frac{\mu^2}{2\sigma^2} + \ln\sqrt{2\pi}\sigma = \frac{(\theta^2)^2}{4\theta^1} + \frac{1}{2}\ln\left(\frac{\pi}{\theta^1}\right)$$ とおくと

$$p(x;\theta) = \exp\{F_1(x)\theta^1 + F_2(x)\theta^2 - \varphi(\theta)\}$$

$\theta\in\Theta = \{(\theta^1,\theta^2)\in\mathbb{R}^+\times\mathbb{R}\}$

1次元正規分布は exp型分布族　　φの凸性を調べる必要がある.

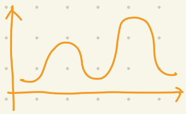
同様に多項分布族, Poisson分布族, 指数分布族も exp型であることが分かる.

que 6.1.1

$\mathcal{X}$上の prob dist の族 $\mathcal{P} := \{p(x;\lambda);\ \lambda\in\Lambda\}$ が $\mathcal{X}$上の $(n+1)$コの prob dist

$p_1(x),\cdots,p_{n+1}(x)$によって

$$p(x;\lambda) = \sum_i^n \lambda^i p_i(x) + \left(1-\sum_i^n\lambda^i\right)p_{n+1}(x)$$



と表されているとき, mix型分布族という.

幾つかの prob dist を合わせて（パラメタ調整して）としたもの

(i) $$\frac{\partial^2 l_\lambda}{\partial\lambda^i\partial\lambda^j} = \frac{\partial}{\partial\lambda^i}\left(\frac{\partial}{\partial\lambda^j}\ln\left(\sum_i^n\lambda^i p_i(x) + \left(1-\sum_i^n\lambda^i\right)p_{n+1}(x)\right)\right)$$

$$= \frac{\partial}{\partial\lambda^i}\left(\frac{p_j(x)-p_{n+1}(x)}{\underbrace{\sum_i^n\lambda^i p_i(x) + \left(1-\sum_i^n\lambda^i\right)p_{n+1}(x)}_{p(x;\lambda)}}\right)$$

$$= \frac{-(p_j(x)-p_{n+1}(x))(p_i(x)-p_{n+1}(x))}{\left(\sum_k^n \lambda^k p_k(x) + \left(1-\sum_k^n \lambda^k\right) p_{n+1}(x)\right)^2}$$

$$= -\frac{\partial \ell_\lambda}{\partial \lambda^i}\frac{\partial \ell_\lambda}{\partial \lambda^j}$$

(ii) $\Gamma^{(-1)}_{kij} = \Gamma_{kij} + T_{kij}$

$\Gamma_{kij} = E_\lambda\left[\frac{\partial^2 \ell_\lambda}{\partial\lambda^i\partial\lambda^j}\frac{\partial \ell_\lambda}{\partial\lambda^k}\right] + \frac{1}{2}E_\lambda\left[\frac{\partial \ell_\lambda}{\partial\lambda^i}\frac{\partial \ell_\lambda}{\partial\lambda^j}\frac{\partial \ell_\lambda}{\partial\lambda^k}\right]$

$T_{ijk} := \frac{1}{2}E_\lambda\left[\frac{\partial \ell_\lambda}{\partial\lambda^i}\frac{\partial \ell_\lambda}{\partial\lambda^j}\frac{\partial \ell_\lambda}{\partial\lambda^k}\right]$

$$= E_\lambda[ij, k] + E_\lambda[i, j, k]$$

$$= -E_\lambda[i, j, k] + E[i, j, k]$$

$$= 0$$

$g_{ij}(\lambda) := E_\lambda\left[\frac{\partial \ell_\lambda}{\partial\lambda^i}\frac{\partial \ell_\lambda}{\partial\lambda^j}\right] = \int_{\mathcal{X}} \frac{\partial \ell}{\partial\lambda^i}(x;\lambda)\frac{\partial \ell}{\partial\lambda^j}(x;\lambda)\, p(x;\lambda)\, dx$

よって $\nabla^{(-1)}$ は flat で $\{\lambda^i\}$ は $\nabla^{(-1)}$ に関する affine chart.

(iii) $$g_{ij}(\lambda) = \int_{\mathcal{X}} \frac{p_i(x)-p_{n+1}(x)}{p(x;\lambda)}\frac{p_j(x)-p_{n+1}(x)}{\cancel{p(x;\lambda)}}\, \cancel{p(x;\lambda)}\, dx$$

$$= ?$$

$$= \sum_{k=1}^{n+1}\frac{(\delta_{ik}-\delta_{n+1,k})(\delta_{jk}-\delta_{n+1,k})}{\lambda^k}$$

$$= \frac{1}{\lambda^i}\delta_{ij} + \frac{1}{\lambda^{n+1}}$$

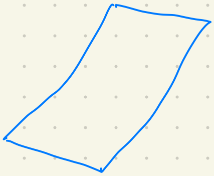
パラメタが cpt なら affine mfd として torus も考えることができるが、よく分からない

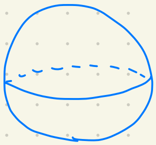
→ 応用上 $\mathbb{R}^n$ しか (affine) としては出てこない.

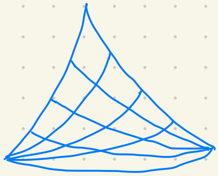
2-dimにおいてRiemann mfdでconst curvature cは規格化すると

$c = +1, 0, -1$のいずれかである

$c = 0$: flat, $c = +1$: sphere, $c = -1$: hypobolic plane (双曲平面).



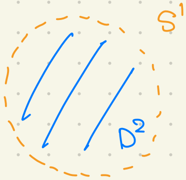




$c > 0$: 縮まる (cpt) というイメージで、Euclid spにembedding可能.

$c < 0$: 拡がる (non cpt) 〃 不可能 (自己交叉する)

non cptな $\mathbb{R}$ を $(-1, 1)$ に有界化するように双曲平面を有界化すると円盤のようになる



無限に大きな平面を含む理想境界

一般に (nonconstな) curvatureでもRiemann mfdであれば ideal boundary $\simeq S^{n-1}$

双曲平面をリサン変換群で割ると (適切な同一視をすると) genus n (≥ 2) の torusが得られる

6.2 正規分布から導かれる Hesse 構造

n次元実対称行列の集まり $\mathcal{S}_n$, 正定値行列の集まり $\mathcal{S}_n^+ \subset \mathcal{S}_n$

n項列vect $x, \mu \in \mathbb{R}^n$, $\sigma \in \mathcal{S}_n$ に対して

$$p(x;\mu,\sigma) := (2\pi)^{-\frac{n}{2}} (\det \sigma)^{-\frac{1}{2}} \exp\left(-\frac{{}^t(x-\mu)\sigma^{-1}(x-\mu)}{2}\right)$$

とすると $\{p(x;\mu,\sigma) \mid (\mu,\sigma) \in \mathbb{R}^n \times \mathcal{S}_n^+\}$ は (μ,σ) をパラメタとする $\mathbb{R}^n$上の prob dist の族. これを正規分布族という

V: finite vect sp. $\Omega \subset V$ linear map $\rho: \Omega \to \mathcal{S}_n$ が inj で

(A1) $\rho(\omega) \in \mathcal{S}_n^+$ for all $\omega \in \Omega$

をみたしているとする. このとき

$$p(x;\mu,\omega) := (2\pi)^{-\frac{n}{2}} (\det \rho(\omega))^{-\frac{1}{2}} \exp\left(-\frac{{}^t(x-\mu)\rho(\omega)(x-\mu)}{2}\right)$$

とすると $\{p(x;\mu,\omega) \mid (\mu,\omega) \in \mathbb{R}^n \times \Omega\}$ は (μ,ω) をパラメタとする $\mathbb{R}^n$上の prob dist の族. これを導かれる prob dist の族.

prop 6.2.1

ρから導かれる prob dist の族 $\{p(x;\mu,\sigma) \mid (\mu,\omega) \in \mathbb{R}^n \times \Omega\}$ は

$\theta := \rho(\omega)\mu \in \mathbb{R}^n$, $\omega \in \Omega$ をパラメタとする exp型分布族で Fisher情報行列は

$$\varphi(\theta,\omega) := \frac{1}{2}\{{}^t\theta\, \rho(\omega)^{-1}\theta - \ln(\det \rho(\omega))\}$$

に関する Hesse metric $g = Dd\varphi$ と一致. (D: $\mathbb{R}^n \times V$上の canonical flat)

proof

V の basis $\{v^1, \cdots, v^m\}$ $x = (x^i)$, $\mu = (\mu^i) \in \mathbb{R}^n$, $\omega = \omega_\alpha v^\alpha \in \Omega$ に対して

$F^\alpha(x) := -\frac{1}{2}{}^t x\, \rho(v^\alpha) x$, $\theta := \rho(\omega)\mu$ とおけば

$$\begin{aligned} p(x;\mu,\omega) &= p(x;\theta,\sigma) \\ &= \exp(\theta_j x^j + \omega_\alpha F^\alpha(x) - \varphi(\theta,\omega) - \frac{n}{2}\ln 2\pi\} \end{aligned}$$

より exp型. Hesse metric $g = Dd\varphi$ と一致

φ の Legendre trans φ' は $\varphi' = \frac{1}{2}\ln\det\rho(\omega) - \frac{n}{2}$

prop 6.2.2

ρ から導かれる prob dist の族 $\{p(x;\mu,\sigma) \mid (\mu,\omega) \in \mathbb{R}^n \times \Omega\}$ の divergence は

$$2\mathcal{D}(p,q) = {}^t(\mu(p)-\mu(q))\,\rho(\omega(p))\,(\mu(p)-\mu(q)) + \mathrm{Tr}(\rho(\omega(p))\rho(\omega(q))^{-1})$$
$$- \ln\det(\rho(\omega(p))\rho(\omega(q))^{-1}) - n$$