2014年3月6日 第21回 沼津研究会

不定値計量空間の双対幾何

A dual geometry on indefinite inner product spaces

鈴木 達夫 Tatsuo Suzuki*1

Shibaura Institute of Technology

(吉澤 真太郎氏（御殿場基礎科学研究会）との共同研究)

*1 *E-mail address*: suzukita@sic.shibaura-it.ac.jp

1　背景と目的

多変量解析の分野などでは長方形行列を変数とする理論が重要である．
また，行列変数の力学系，双対理論も研究されている([Y] など).

ポテンシャル関数：

$$f(X) = -\log\det(\text{正定値対称行列} + \Sigma_1 X \Sigma_2 X^T) + \cdots, \quad \Sigma_i^T = \Sigma_i > 0\ (i = 1, 2) \tag{1}$$

そこで，今回の研究では不定値対称行列 $J_r = \mathrm{diag}(-1, \cdots, -1, 1, \cdots, 1)$ (-1 が r 個)
および 長方形行列のダブル (V, W) を用いたポテンシャル関数

$$f(V, W) := -\log\det(J_r + WV^T) \tag{2}$$

に対する双対理論の構築を試みた．

Remark 1. V, W がベクトル(n 行 1 列)のとき，

$$\det(J_r + WV^T) = (-1)^r \det(I + J_r WV^T) = (-1)^r (1 + V^T J_r W) = (-1)^r (1 + \langle V, W \rangle_r), \tag{3}$$

$\langle V, W \rangle_r$: 不定値内積.

2　双対平坦構造

一般に，C$^\infty$ 多様体 M のアファイン接続 ∇ に対し，座標系 $[x^i]$ で $\nabla_{\partial_i}\partial_j = 0 \quad (\forall i, j)$ を満たすものを ∇ のアファイン座標系と呼び，アファイン座標系が存在するとき ∇ は平坦であるという．

多様体 M の平坦接続 ∇ と(擬)Riemann 計量の組 (∇, g) が Hesse 構造であるとは，g が ∇ のアファイン座標系に関してある関数 φ の Hesse 形式

$$g_{ij} = \frac{\partial^2 \varphi}{\partial x^i \partial x^j}, \quad \text{すなわち} \quad g = \nabla d\varphi \tag{4}$$

によって表されることである．このとき g を ∇ に関する Hesse 計量，φ を g の ∇ に関する(主)ポテンシャルという．このとき，∇^0 を g の Levi-Civita 接続，$\nabla^* = 2\nabla^0 - \nabla$ とおくと ∇^* も M の平坦接続となり，M 上の任意のベクトル場 X, Y, Z に対し，

$$Xg(Y, Z) = g(\nabla_X Y, Z) + g(Y, \nabla^*_X Z) \tag{5}$$

が成り立つ．この ∇^* を ∇ の双対接続といい，$(g, \nabla,\ \nabla^*)$ を M 上の双対構造，$(M, g, \nabla,\ \nabla^*)$ を双対平坦空間と呼ぶ．

φ をルジャンドル変換した関数を $\varphi^* = \sum_i x^i \dfrac{\partial \varphi}{\partial x^i} - \varphi$ (双対ポテンシャル) とすると，

$g = \nabla^* d\varphi^*$ が成り立ち，g は ∇^* に関しても Hesse 計量となる．

3　行列変数への準備：vec-作用素とテンソル積

$X \in \mathbf{R}^{m \times n}$ **に対して，**vec-**作用素**　vec(X) **を次のように定義する：**

$$X = \begin{pmatrix} x_{11} & \cdots & x_{1n} \\ \vdots & & \vdots \\ x_{m1} & \cdots & x_{mn} \end{pmatrix} \quad \text{に対し，} \quad \mathrm{vec}(X) = \begin{pmatrix} x_{11} \\ \vdots \\ x_{m1} \\ \vdots \\ x_{1n} \\ \vdots \\ x_{mn} \end{pmatrix} \in \mathbf{R}^{mn \times 1} \tag{6}$$

Proposition 1.

$$\mathit{vec}(X^T) = P(m,n)\mathit{vec}(X) \qquad \forall X \in \mathbf{R}^{m \times n} \tag{7}$$

ここで，$P(m,n)$ **は vec-permutation 行列と呼ばれ，次で定義される：**

$$P(m,n) = \sum_{i=1}^{m} \sum_{j=1}^{n} E_{ij} \otimes E_{ij}^T \qquad E_{ij} \text{ は} m \text{行} n \text{列の行列単位} \tag{8}$$

性質：

$$P(m,1) = I_m, \quad P(1,n) = I_n, \quad P(m,n) = P(n,m)^T = P(n,m)^{-1} \tag{9}$$

Hesse**行列を抜き出すために，以下の公式を用いる：**

Proposition 2.

$$B \otimes A = P(k,n)[A \otimes B]P(n,k), \quad A \in \mathbf{R}^{n\times n},\ B \in \mathbf{R}^{k\times k} \tag{10}$$

$$\mathit{tr}\,(AXBY^T) = \mathit{vec}^T(Y)[B^T \otimes A]\mathit{vec}(X) \tag{11}$$

参考文献：

Topics in Matrix Analysis, R.A.Horn and C.R.Johnson, Cambridge University Press(1994).

4　主ポテンシャルとそのHessian

$\mathbf{R}^{(n+1)\times m} \times \mathbf{R}^{(n+1)\times m}$ の開部分多様体

$$M := \{(V,W) \in \mathbf{R}^{(n+1)\times m} \times \mathbf{R}^{(n+1)\times m} |\ \det(J_r + WV^T) > 0\} \tag{12}$$

を考え，M 上の関数

$$f(V,W) := -\log\det(J_r + WV^T) = -\mathrm{tr}\ \log(J_r + WV^T) \tag{13}$$

の２階微分を計算する．

M 内の曲線 $(V,W) = (V(t),W(t))$ に対し，

$$\frac{df}{dt} = -\mathrm{tr}\ \left((J_r + WV^T)^{-1}(\frac{dW}{dt}V^T + W\frac{dV^T}{dt}) \right) \tag{14}$$

ここで $A \equiv J_r + WV^T$ とおく．このとき，接ベクトル $(X,Y) \in T_{(V,W)}M$ に対し１階微分は

$$\frac{df}{dt}(X,Y) = -\mathrm{tr}\ \left[A^{-1}(YV^T + WX^T)\right] = -\mathrm{tr}\ \left[YV^TA^{-1} + A^{-1}WX^T\right] \tag{15}$$

同様にして２階微分を計算すると，$(X_1, Y_1), (X_2, Y_2) \in T_{(V,W)}M$ に対し

$$\frac{d^2 f}{dt^2}(X_1, Y_1; X_2, Y_2)$$

$$= \ \mathrm{tr}\ \left[A^{-1}(Y_1 V^T + W X_1^T) A^{-1}(Y_2 V^T + W X_2^T)\right] + \mathrm{tr}\ \left[A^{-1}(Y_1 X_2^T + Y_2 X_1^T)\right] \quad (16)$$

$$= \ \begin{pmatrix}\mathrm{vec}^T(X_2) & \mathrm{vec}^T(Y_2)\end{pmatrix} \times$$
$$\begin{pmatrix} P(m, n+1)[(A^{-1}W) \otimes (A^{-1}W)^T] & (I + V^T A^{-1} W)^T \otimes A^{-1} \\ (I + V^T A^{-1} W) \otimes (A^{-1})^T & P(m, n+1)[(V^T A^{-1})^T \otimes (V^T A^{-1})] \end{pmatrix} \begin{pmatrix} \mathrm{vec}(X_1) \\ \mathrm{vec}(Y_1) \end{pmatrix} \quad (17)$$

Theorem 3. 上記の $2m(n+1)$ **次の** *Hesse* **行列が正定値であるための必要十分条件は，**

$$P(m, n+1)[(A^{-1}W) \otimes (A^{-1}W)^T] > 0, \quad (18)$$

$$P(m, n+1)\{[(V^T A^{-1})^T \otimes (V^T A^{-1})]$$
$$-[(A^{-1})^T \otimes (I + V^T A^{-1} W)][(A^{-1}W)^T \otimes (A^{-1}W)]^{-1}[(I + V^T A^{-1} W)^T \otimes A^{-1}]\} > 0 \quad (19)$$

5　双対ポテンシャル

(15) より双対変数 (P, Q) を

$$P := -A^{-1}W,\ Q^T := -V^T A^{-1} \quad \text{つまり} \quad P := -A^{-1}W,\ Q := -(A^T)^{-1}V \tag{20}$$

と定義する．

Remark 2. $m = 1$ の場合，すなわち $V = (v_0, v_1, \cdots, v_n)^T,\ W = (w_0, w_1, \cdots, w_n)^T$ のとき，$V^T J_r W =: \langle V, W \rangle_r$ （符号が r の不定値内積）となるので，$A = J_r + WV^T$ の行列式 $|A|$ は，ある種の双対性

$$\det(I + XY^T) = \det(I + Y^T X) \tag{21}$$

と $|J_r| = (-1)^r$ に注意して，

$$|A| = (-1)^r(1 + (V^T J_r W)) = (-1)^r(1 + \langle V, W \rangle_r), \tag{22}$$

さらに A の余因子行列 $\tilde{A} = |A|A^{-1}$ は

$$\tilde{A} = (-1)^r[J_r + \{(V^T J_r W)J_r - (J_r W)(J_r V)^T\}] \tag{23}$$

で与えられる．さらに

$$\{(V^T J_r W)J_r - (J_r W)(J_r V)^T\}W = O \tag{24}$$

が成り立つので，

$$P = -A^{-1}W = -\frac{1}{|A|}\tilde{A}W = (-1)^{r+1}\frac{1}{|A|}J_r W \tag{25}$$

$$= \frac{(-1)^{r+1}}{(-1)^r(1+\langle v,w\rangle_r)}(-w_0,-w_1,\cdots,-w_{r-1},w_r,\cdots,w_n)^T \tag{26}$$

$$= \left(-\frac{\partial}{\partial v_i}\log((-1)^r(1+\langle V,W\rangle_r))\right)_{i=0,1,\cdots,n}, \tag{27}$$

Q の方も同様の式が成り立つ．

さて，$P=-A^{-1}W,\ Q^T=-V^TA^{-1}$ に対し，

$$A = J_r + WV^T = J_r + APQ^TA \tag{28}$$

$$\Rightarrow \qquad A = (J_r + PQ^T)^{-1} \equiv B^{-1} \tag{29}$$

に注意する．

さらに非退化二次形式を

$$\langle (V,W),(P,Q)\rangle := \mathrm{tr}\ \left[WQ^T + PV^T\right] \tag{30}$$

とおき，双対ポテンシャル関数を次のように定義する：

$$\begin{aligned} f^*(P,Q) &:= \langle (V,W),(P,Q)\rangle - f \\ &= -2\mathrm{tr}\ [B^{-1}PQ^T] + \log\det B^{-1} \\ &= -2\mathrm{tr}\ [(J_r + PQ^T)^{-1}PQ^T] - \log\det(J_r + PQ^T) \end{aligned} \tag{31}$$

6　ダイバージェンス

ダイバージェンスは次のように定義される：

$$
\begin{aligned}
D(V_1, W_1 || P_2, Q_2) := {} & f(V_1, W_1) + f^*(P_2, Q_2) - \langle (V_1, W_1), (P_2, Q_2) \rangle \\
= {} & -\log\det(J_r + W_1 V_1^T) \\
& \quad -2\mathrm{tr}\,[(J_r + P_2 Q_2^T)^{-1} P_2 Q_2^T] - \log\det(J_r + P_2 Q_2^T) \\
& \qquad -\mathrm{tr}\,\left[W_1 Q_2^T + P_2 V_1^T\right]
\end{aligned} \tag{32}
$$

$$
\begin{aligned}
= {} & -\log\det(J_r + W_1 V_1^T)(J_r + P_2 Q_2^T) \\
& \quad -\mathrm{tr}\,\left[W_1 Q_2^T + P_2 V_1^T + 2(J_r + P_2 Q_2^T)^{-1} P_2 Q_2^T\right]
\end{aligned} \tag{33}
$$

$$
\begin{aligned}
= {} & -\log\det(J_r + W_1 V_1^T)(J_r + W_2 V_2^T)^{-1} \\
& \quad +\mathrm{tr}\,\left[(J_r + W_2 V_2^T)^{-1}(W_1 V_2^T + W_2 V_1^T - 2 W_2 V_2^T)\right]
\end{aligned} \tag{34}
$$

特に$m=1$**の場合，**$\det(J_r+WV^T)=(-1)^r(1+\langle V,W\rangle_r)$ **と**

$$\begin{aligned}\text{tr}\ \{(J_r+W_2V_2^T)^{-1}W_iV_2^T\} &= \text{tr}\ \{V_2^T(I+J_rW_2V_2^T)^{-1}J_rW_i\}\\ &= (1+V_2^TJ_rW_2)^{-1}V_2^TJ_rW_i \quad (\text{“Bumping lemma”})\\ &= (1+\langle V_2,W_2\rangle_r)^{-1}\langle V_2,W_i\rangle_r \quad (i=1,2),\end{aligned} \tag{35}$$

$$\begin{aligned}\text{tr}\ \{(J_r+W_2V_2^T)^{-1}W_2V_1^T\} &= \text{tr}\ \{V_1^TJ_r(I+W_2V_2^TJ_r)^{-1}W_2\}\\ &= V_1^TJ_rW_2(1+V_2^TJ_rW_2)^{-1} \quad (\text{“Bumping lemma”})\\ &= \langle V_1,W_2\rangle_r(1+\langle V_2,W_2\rangle_r)^{-1} \quad (i=1,2)\end{aligned} \tag{36}$$

を用いて，次のように書きなおせる：

$$D(V_1,W_1||P_2,Q_2)=\frac{\langle V_1,W_2\rangle_r+\langle V_2,W_1\rangle_r-2\langle V_2,W_2\rangle_r}{1+\langle V_2,W_2\rangle_r}-\log\frac{1+\langle V_1,W_1\rangle_r}{1+\langle V_2,W_2\rangle_r} \tag{37}$$

$$=\frac{x_{12}+x_{21}-2x_{22}}{1+x_{22}}-\log\frac{1+x_{11}}{1+x_{22}} \tag{38}$$

ここで $x_{ij}:=\langle V_i,W_j\rangle_r$ **とおいた．**

7 Legendre変換と擬球

$r=1$, $V=W\in\mathbf{R}^{(n+1)\times 1}$, $A^T=(J_r+VV^T)^T=A$, $P=Q=-A^{-1}V$とする．
$(\mathbf{R}^{(n+1)\times 1},\langle\ ,\ \rangle_1)$ をミンコフスキー空間 $\mathbf{R}_1^{n+1}$ と考え，パラメータ$\alpha>0$を導入して次のポテンシャル関数を考える：

$$f_\alpha(V):=-\log\det(J_1+\alpha VV^T)=-\log\{-(1+\alpha\langle V,V\rangle_1)\} \tag{39}$$

変数Vと双対変数Pの関係は，

$$\langle P,P\rangle_1=\frac{4\alpha^2\langle V,V\rangle_1}{(1+\alpha\langle V,V\rangle_1)^2} \tag{40}$$

$\langle V,V\rangle_1$ と$\langle P,P\rangle_1$ の対応は一般の点では一対一ではないが，"極大点" で一対一となり，
$\alpha=\dfrac{1}{c^2}$ とおくと，pseudo-sphere (de Sitter space)

$$S_1^n(c^2)=\{V\in\mathbf{R}_1^{n+1}|\ \langle V,V\rangle_1=c^2\}$$

と

$$S_1^n(\frac{1}{c^2})=\{P\in(\mathbf{R}_1^{n+1})^*|\ \langle P,P\rangle_1=\frac{1}{c^2}\}$$

が対応する．（この対応に付随する幾何構造の理解は今後の課題）

参考文献

[Amari-Nagaoka] **甘利俊一，長岡浩司，情報幾何の方法**, **岩波講座応用数学**12(1993).
[志磨] **志磨裕彦 ヘッセ幾何学**, **裳華房，**2001.
[Y] S. Yoshizawa , *Legendre dualities between matrix subspace flows*, Mathematical System Theory, (2013), 471-478.
[Y-T] S. Yoshizawa and K. Tanabe, *Dual differential geometry associated with the Kullback-Leibler informaion on the Gaussian distributions and its 2-parameter deformations*, SUT Journal of Mathematics, vol.35, No.1 (1999), 113-137.