

修士論文

正值行列のノルム不等式と幾何平均

佐々木 賢之介*

2008 年

*東北大学大学院情報科学研究科システム情報科学専攻

目次

1 正値行列とノルム	3
1.1 定義と基本性質	3
1.2 ブロック行列	4
2 Fréchet 微分	7
3 行列平均	9
3.1 正数の平均	9
3.2 幾何平均	9
3.3 古田の不等式	12
3.4 正定値関数	16
3.5 ノルム不等式	17
4 正値行列の幾何平均	23
4.1 リーマン距離	23
4.2 距離空間	29
4.3 幾何平均	34
4.4 ユニタリ不変ノルムに対する不等式	41

序

平均化は数学において昔から行われてきた。例えば2と4の平均は3である、という単純なものを小学校で学習し、高校では相加・相乗平均という平均に関する関係を学習した。学校での学習においてだけでなく、平均は実生活でも我々になじみのあるものである。いくつかある平均の中でも特に幾何平均とも呼ばれる相乗平均に注目し、正値行列における幾何平均を考察するのが本修士論文の第一のテーマである。

正値行列の定義は第1章で述べるが、正値行列全体の集合はコーン状の形を成しており、第4章で定義される距離は直線ではなく、カーブを描く。そのため、3つの正値行列を頂点とする三角形は平面のものとは違う形状だが、中線定理などの基本的な性質と似たものを持つことは興味深い。また、正値行列の性質は多くのノルム不等式によって表される。

本論文では、幾何平均を考察するのに必要な基礎知識を第1章と第2章で紹介する（第1章や第2章についての派生的な内容や証明等は[1]や[2]を参照されたい）。第3章では、行列の平均が持つべき性質について3.1節でまとめ、3.2節では2つの正値行列の幾何平均を導入する。3.3節では正値行列に関する不等式についてまとめた。3.4節で正定値関数を導入し、3.5節ではそれを利用して正値行列に関するノルム不等式を考察する。第4章では、まず4.1節で正値行列全体の集合に距離を導入し、4.2節ではその距離空間としての性質をまとめた。4.3節では3つの正値行列の幾何平均を定義する2つの方式を解説し、その違いや性質について明らかにする。4.4節では物理学と関連のある不等式を紹介する。

謝辞

セミナーや論文作成において、親身に丁寧なアドバイスをくださった指導教官の日合文雄教授をはじめ、同研究室先輩の大野博道さん、宮本拓歩さんに深く感謝致します。本当にありがとうございました。

1 正値行列とノルム

1.1 定義と基本性質

\mathbb{H}_n を $n \times n$ のエルミート行列全体, $\mathbb{H}_n(I)$ を固有値がすべて区間 I に含まれる $n \times n$ のエルミート行列全体とする. \mathcal{H} をヒルベルト空間, $\mathcal{L}(\mathcal{H})$ を \mathcal{H} 上の有界線形作用素全体, \mathbb{M}_n を $n \times n$ の複素行列全体とする. この論文では, ヒルベルト空間はすべて有限次元とする.

定義 1.1 任意の $x \in \mathcal{H}$ に対して, $\langle x, Ax \rangle \geq 0$ のとき, A は半正値であるという. 任意の $x \neq 0$ に対して, $\langle x, Ax \rangle > 0$ のとき, A は正値であるという. それぞれ $A \geq O$, $A > O$ で表す. また, $A - B \geq O$, $A - B > O$ のとき, $A \geq B$, $A > B$ と書く.

命題 1.2

- (1) $A > O$ と, $A \geq O$ かつ A が可逆であることは同値.
- (2) $A \geq O$ ならば, 任意の行列 B に対して $B^*AB \geq O$.

定義 1.3 ヒルベルト空間上の線形作用素のノルムを以下のように定義する.

$$\|A\| = \sup_{\|x\|=1} \|Ax\| = \sup_{\|x\| \leq 1} \|Ax\|.$$

このノルムは以下の性質を持つ.

$$\begin{aligned} \|AB\| &\leq \|A\| \|B\|, \\ \|A\| &= \|A^*\|, \\ \|A\| &= \|UAV\| \quad (U, V \text{ は任意のユニタリ行列}). \end{aligned}$$

最後の性質は, ユニタリ不変性と呼ばれる.

定義 1.4 \mathbb{M}_n 上のノルム $\|\cdot\|$ が, すべてのユニタリ行列 U, V に対して, $\|A\| = \|UAV\|$ を満たすとき, ノルム $\|\cdot\|$ をユニタリ不変ノルムという.

定義 1.5 $n \times n$ 行列 $A = (a_{ij})$, $B = (b_{ij})$ に対して, $A \circ B := (a_{ij}b_{ij})$ と定義する. $A \circ B$ はシューア積あるいはアダマール積と呼ばれる.

命題 1.6 $A \geq O$ のとき, 任意の行列 X に対して, $\|A \circ X\| \leq \max a_{ii} \|X\|$.

命題 1.7 $A > O$ とし, B をエルミート行列とすると, X^*AX と X^*BX が対角行列となるような正則行列 X が存在する. このとき, A と B は X により同時に対角化されるという.

命題 1.8 $A, B \geq O$, $A > B$ のとき, $A^{1/2} > B^{1/2}$.

命題 1.9 $A > O$, $B \geq O$ のとき, $XAX = B$ を満たす解 $X \geq O$ がただひとつ存在する.

命題 1.10 A を \mathcal{H} 上の正値な作用素, X をヒルベルト空間 \mathcal{K} から \mathcal{H} への線形作用素とすると, X^*AX は \mathcal{K} 上の正値行列となる. また, X が可逆のとき, 逆も成り立つ.

定義 1.11 $A, B \in \mathcal{L}(\mathcal{H})$ とする. $B = X^*AX$ となるような \mathcal{H} 上の正則作用素 X が存在するとき, A は B と合同であるといい, $A \sim B$ で表す.

定義 1.12 $A, B \in \mathcal{L}(\mathcal{H})$ とする. $B = U^*AU$ となるような \mathcal{H} 上のユニタリ作用素 U が存在するとき, A は B とユニタリ同値であるといい, $A \simeq B$ で表す.

1.2 ブロック行列

$A, B, C, D \in M_n$ に対して, 2×2 ブロック行列 $\begin{bmatrix} A & B \\ C & D \end{bmatrix} \in M_{2n}$ を考え, その性質を見る.

定義 1.13 行列 A に対し, ユニタリ行列 U と正値行列 P を用いて表される $A = UP$ を A の極分解と呼ぶ. また, $|A| := P = (A^*A)^{1/2}$ と定義する. $|A|$ の固有値を A の特異値と呼び, $s_1(A), \dots, s_n(A)$ で表す.

定義 1.14 行列 A に対し, $S := \text{diag}(s_1(A), \dots, s_n(A))$ (ただし, $s_1(A) \geq \dots \geq s_n(A) \geq 0$) とユニタリ行列 U, V を用いて表される $A = USV$ を A の特異値分解と呼ぶ.

命題 1.15 $\|A\| \leq 1$ と $\begin{bmatrix} I & A \\ A^* & I \end{bmatrix} \geq O$ は同値.

証明 特異値分解を用いて $A = USV$ とすると,

$$\begin{aligned} \begin{bmatrix} I & A \\ A^* & I \end{bmatrix} &= \begin{bmatrix} I & USV \\ V^*SU^* & I \end{bmatrix} \\ &= \begin{bmatrix} U & O \\ O & V^* \end{bmatrix} \begin{bmatrix} I & S \\ S & I \end{bmatrix} \begin{bmatrix} U^* & O \\ O & V \end{bmatrix}. \end{aligned}$$

よって, $\begin{bmatrix} I & A \\ A^* & I \end{bmatrix} \simeq \begin{bmatrix} I & S \\ S & I \end{bmatrix}$ であり, また, $\begin{bmatrix} I & S \\ S & I \end{bmatrix} \simeq \begin{bmatrix} 1 & s_1(A) \\ s_1(A) & 1 \end{bmatrix} \oplus \dots \oplus \begin{bmatrix} 1 & s_n(A) \\ s_n(A) & 1 \end{bmatrix}$ である. 任意の $k = 1, \dots, n$ に対して $\begin{bmatrix} 1 & s_k(A) \\ s_k(A) & 1 \end{bmatrix} \geq 0$ となる必要十分条件は $s_1(A) \leq 1$, つまり $\|A\| \leq 1$. ■

命題 1.16 $A, B > O$ とする. このとき, $\begin{bmatrix} A & X \\ X^* & B \end{bmatrix} \geq O$ と, $\|K\| \leq 1$ である K が存在して $X = A^{1/2}KB^{1/2}$ となることは同値.

証明 $A, B > O$ とすると,

$$\begin{aligned} \begin{bmatrix} A & X \\ X^* & B \end{bmatrix} &\sim \begin{bmatrix} A^{-1/2} & O \\ O & B^{-1/2} \end{bmatrix} \begin{bmatrix} A & X \\ X^* & B \end{bmatrix} \begin{bmatrix} A^{-1/2} & O \\ O & B^{-1/2} \end{bmatrix} \\ &= \begin{bmatrix} I & A^{-1/2}XB^{-1/2} \\ B^{-1/2}X^*A^{-1/2} & I \end{bmatrix}. \end{aligned}$$

$K := A^{-1/2}XB^{-1/2}$ とすると, 命題 1.15 より, $\begin{bmatrix} I & K \\ K^* & I \end{bmatrix} \geq O \Leftrightarrow \|K\| \leq 1$. これでは, $A, B > O$ のときは示せた. $A, B \geq O$ のときは以下のように考える.

$A, B \geq O$ に対して, 任意の $\varepsilon > 0$ を用いると, $A + \varepsilon I, B + \varepsilon I > O$. $\begin{bmatrix} A & X \\ X^* & B \end{bmatrix} \geq O$ のとき,

$$\begin{bmatrix} A + \varepsilon I & X \\ X^* & B + \varepsilon I \end{bmatrix} = \begin{bmatrix} A & X \\ X^* & B \end{bmatrix} + \varepsilon \begin{bmatrix} I & O \\ O & I \end{bmatrix} \geq O$$

なので, $X = (A + \varepsilon I)^{1/2}K_\varepsilon(B + \varepsilon I)$ かつ $\|K_\varepsilon\| \leq 1$ を満たす K_ε が存在する. $\varepsilon \downarrow 0$ として K_ε の集積点 K をとると $X = A^{1/2}KA^{1/2}$ となる. 逆は

$$\begin{bmatrix} A & A^{1/2}KB^{1/2} \\ A^{1/2}KA^{1/2} & B \end{bmatrix} = \begin{bmatrix} A^{1/2} & O \\ O & B^{1/2} \end{bmatrix} \begin{bmatrix} I & K \\ K & I \end{bmatrix} \begin{bmatrix} A^{1/2} & O \\ O & B^{1/2} \end{bmatrix}$$

からわかる. ■

定理 1.17 $A, B > O$ とする. このとき, $\begin{bmatrix} A & X \\ X^* & B \end{bmatrix} \geq O$ と, $A \geq XB^{-1}X^*$ は同値.

証明 合同関係を考えると,

$$\begin{aligned} \begin{bmatrix} A & X \\ X^* & B \end{bmatrix} &\sim \begin{bmatrix} I & -XB^{-1} \\ O & I \end{bmatrix} \begin{bmatrix} A & X \\ X^* & B \end{bmatrix} \begin{bmatrix} I & O \\ -B^{-1}X^* & I \end{bmatrix} \\ &= \begin{bmatrix} A - XB^{-1}X^* & O \\ O & B \end{bmatrix}. \end{aligned}$$

$\begin{bmatrix} A - XB^{-1}X^* & O \\ O & B \end{bmatrix}$ が半正値行列となることは, $A \geq XB^{-1}X^*$ と同値. ■

補題 1.18 $A \geq O$ と $\begin{bmatrix} A & A \\ A & A \end{bmatrix} \geq O$ は同値.

証明 $A = BB^*$ となる行列 B が存在することは, $A \geq O$ であることの必要十分条件. $A \geq O$ のとき

$$\begin{bmatrix} A & A \\ A & A \end{bmatrix} = \begin{bmatrix} A^{1/2} & O \\ A^{1/2} & O \end{bmatrix} \begin{bmatrix} A^{1/2} & A^{1/2} \\ O & O \end{bmatrix} \geq O. \quad \blacksquare$$

系 1.19 任意の A に対して, $\begin{bmatrix} |A| & A^* \\ A & |A^*| \end{bmatrix} \geq O$.

証明 極分解 $A = UP$ を考えると,

$$\begin{aligned} \begin{bmatrix} |A| & A^* \\ A & |A^*| \end{bmatrix} &= \begin{bmatrix} P & PU^* \\ UP & UPU^* \end{bmatrix} \\ &= \begin{bmatrix} I & O \\ O & U \end{bmatrix} \begin{bmatrix} P & P \\ P & P \end{bmatrix} \begin{bmatrix} I & O \\ O & U^* \end{bmatrix}. \end{aligned}$$

ここで, 補題 1.18 を用いればよい. \blacksquare

2 Fréchet 微分

f を区間 I 上の微分可能な関数とする。固有値がすべて I に含まれる $A \in \mathbb{H}_n$ に対して、関数カルキュラス $f(A)$ が定義できる。このとき、 A における導関数 $Df(A)$ は \mathbb{H}_n から \mathbb{H}_n 自身への線形写像である。 $A = \sum \lambda_i P_i$ が A のスペクトル分解のとき、導関数の公式は、任意の $B \in \mathbb{H}_n$ に対して

$$Df(A)(B) = \sum_i \sum_j \frac{f(\lambda_i) - f(\lambda_j)}{\lambda_i - \lambda_j} P_i B P_j \quad (1)$$

となる。 $i = j$ のとき、(1) の商は $f'(\lambda_i)$ を意味する。

この公式は違う方法でも表せる。 $f^{[1]}$ を以下のような $I \times I$ 上の関数とする。

$$\begin{aligned} f^{[1]}(\lambda, \mu) &= \frac{f(\lambda) - f(\mu)}{\lambda - \mu} \quad (\lambda \neq \mu \text{ のとき}), \\ f^{[1]}(\lambda, \lambda) &= f'(\lambda). \end{aligned}$$

これは f の第一商差と呼ばれる。 $A \in \mathbb{H}_n(I)$ に対して、 $f^{[1]}(A)$ を $n \times n$ 行列

$$f^{[1]}(A) := \left[\frac{f(\lambda_i) - f(\lambda_j)}{\lambda_i - \lambda_j} \right] \quad (2)$$

とする ($\lambda_1, \dots, \lambda_n$ は A の固有値)。公式 (1) は、 A を対角化する正規直交基底で表すと、

$$Df(A)(B) = f^{[1]}(A) \circ B \quad (3)$$

と書ける。ただし、 \circ は A の固有ベクトルからなる正規直交基底に関してとる。

定理 2.1 $f \in C^1(I)$, $A \in \mathbb{H}_n$ の固有値がすべて I に含まれるとき、すべての $H \in \mathbb{H}_n$ に対して、

$$Df(A)(H) = f^{[1]}(A) \circ H. \quad (4)$$

証明 $A := U\Lambda U^*$ (Λ は対角行列、 U はユニタリ行列) とする。

$$Df(A)(H) = U \left(f^{[1]}(\Lambda) \circ (U^* H U) \right) U^* \quad (5)$$

を示したい。

(5) の右辺を $Df(A)(H)$ とおく。任意の $f \in C^1(I)$ に対して、 $Df(A)$ は \mathbb{H}_n 上の線形写像である。そして、

$$\|Df(A)(H)\|_2 = \|f^{[1]}(\Lambda) \circ (U^* H U)\|_2. \quad (6)$$

平均値の定理より、 $f^{[1]}(\Lambda)$ のすべての成分の絶対値は $\max_{|t| \leq \|A\|} |f'(t)|$ 以下なので、

$$\|Df(A)(H)\|_2 \leq \max_{|t| \leq \|A\|} |f'(t)| \|H\|_2. \quad (7)$$

H を $A + H$ の固有値がすべて I に含まれるようなエルミート行列とする. $[a, b]$ を $[a, b] \subset I$ かつ A と $A + H$ の固有値がすべて $[a, b]$ に含まれるような閉区間とする. 多項式関数の列 $\{f_n\}$ を $[a, b]$ 上で一様に $f_n \rightarrow f, f'_n \rightarrow f'$ となるようにとる. \mathcal{L} を, \mathbb{H}_n 上の A から $A + H$ への線分とする. このとき,

$$\begin{aligned} & \|f_m(A + H) - f_n(A + H) - (f_m(A) - f_n(A))\|_2 \\ & \leq \|H\|_2 \sup_{X \in \mathcal{L}} \|Df_m(X) - Df_n(X)\|_2 \\ & = \|H\|_2 \sup_{X \in \mathcal{L}} \|Df_m(X) - Df_n(X)\|_2. \end{aligned} \quad (8)$$

ε を任意の正の実数とする. (7) より, 自然数 n_0 が存在して, 任意の $m, n \geq n_0$ に対して,

$$\sup_{X \in \mathcal{L}} \|Df_m(X) - Df_n(X)\|_2 \leq \frac{\varepsilon}{3}, \quad (9)$$

$$\|Df_n(A) - Df(A)\|_2 \leq \frac{\varepsilon}{3}. \quad (10)$$

が成立する. $m \rightarrow \infty$ として, (8) と (9) より,

$$\|f(A + H) - f(A) - (f_n(A + H) - f_n(A))\|_2 \leq \frac{\varepsilon}{3} \|H\|_2. \quad (11)$$

$\|H\|_2$ が十分小さいとき, Fréchet 微分の定義より,

$$\|f_n(A + H) - f_n(A) - Df_n(A)(H)\|_2 \leq \frac{\varepsilon}{3} \|H\|_2. \quad (12)$$

三角不等式より,

$$\begin{aligned} & \|f(A + H) - f(A) - Df(A)(H)\|_2 \\ & \leq \|f(A + H) - f(A) - (f_n(A + H) - f_n(A))\|_2 \\ & \quad + \|f_n(A + H) - f_n(A) - Df_n(A)(H)\|_2 + \|Df(A) - Df_n(A)(H)\|_2. \end{aligned}$$

(10),(11),(12) より, 十分小さい $\|H\|_2$ に対して,

$$\|f(A + H) - f(A) - Df(A)(H)\|_2 \leq \varepsilon \|H\|_2.$$

よって, $Df(A) = Df(A)$. ■

3 行列平均

3.1 正数の平均

正数 a, b に対して, 様々な平均の取り方がある.

$$A(a, b) := \frac{a+b}{2},$$

$$G(a, b) := \sqrt{ab},$$

$$H(a, b) := \left(\frac{a^{-1} + b^{-1}}{2} \right)^{-1}.$$

これらはそれぞれ, 算術平均 (相加平均), 幾何平均 (相乗平均), 調和平均と呼ばれる. また, このような平均を一般に $M(a, b)$ と表す. 平均 $M(a, b)$ は以下の性質を持つ.

- (1) $M(a, b) > 0$,
- (2) $a < b$ のとき, $a \leq M(a, b) \leq b$,
- (3) $M(a, b) = M(b, a)$,
- (4) $M(a, b)$ は a, b に関して単調,
- (5) 任意の $a, b > 0$ と α に対して, $M(\alpha a, \alpha b) = \alpha M(a, b)$,
- (6) $M(a, b)$ は a, b に関して連続.

さらに, $H(a, b) \leq G(a, b) \leq A(a, b)$ である.

不等号 \leq をエルミート行列における $X \leq Y$ と考えて, 正値行列に関して同様の平均を考えてみる. 上の性質の (5) は, $a, b > 0, x (\neq 0) \in \mathbb{C}$ に対して,

$$M(\bar{x}ax, \bar{x}bx) = \bar{x}M(a, b)x$$

とも考えられるので, 行列の平均が満たす性質として,

$$M(X^*AX, X^*BX) = X^*M(A, B)X \quad (X \text{ は正則行列}) \tag{13}$$

を期待する. 実際に行列の算術平均として $A(A, B) = \frac{1}{2}(A+B)$ を考えると, 上の性質をすべて満たす. 同様に調和平均は $H(A, B) = \left(\frac{A^{-1}+B^{-1}}{2} \right)^{-1}$ とできる. しかし幾何平均はそのように簡単にはいかない. それについては以下の節で述べる.

3.2 幾何平均

幾何平均を $G(A, B) = A^{1/2}B^{1/2}$ としても, A と B が交換可能でなければ, $A^{1/2}B^{1/2}$ はエルミート行列にならない. ここで, 命題 1.7 を用いると, 2つの任意の正値行列は同時に対角化することができる. 正値な対角行列 A, B は交換可能なので, その幾何平均は $A^{1/2}B^{1/2}$ とできる.

議論に入る前に、記号をいくつか定義する。 $n \times n$ の正則行列全体を $GL(n)$ とし、行列 X, A に対して $\Gamma_X(A) := X^*AX$ と定める。写像 Γ_X は行列の正值性を保存する。また、行列 A, B に対して、 $\Gamma_X(A, B) := (\Gamma_X(A), \Gamma_X(B))$ と定める。

$A, B > O$ とすると、

$$\Gamma_{A^{-1/2}}(A, B) = (I, A^{-1/2}BA^{-1/2}).$$

$A^{-1/2}BA^{-1/2}$ はエルミート行列なので、 $U^*(A^{-1/2}BA^{-1/2})U$ が対角行列となるようなユニタリ行列 U が存在する。 $D := U^*(A^{-1/2}BA^{-1/2})U$ とすると、

$$\Gamma_{A^{-1/2}U}(A, B) = (I, D).$$

I と D は交換可能なので、 I と D の幾何平均は、

$$G(I, D) = D^{1/2} = U^*(A^{-1/2}BA^{-1/2})^{1/2}U$$

となる。 $A, B > O$ の幾何平均が (13) を満たすことを考慮すると、 A と B の幾何平均は

$$A\#B := G(A, B) = A^{1/2}(A^{-1/2}BA^{-1/2})^{1/2}A^{1/2} \quad (14)$$

となる。 A と B が交換可能ならば、 $A\#B = A^{1/2}B^{1/2}$ である。また、命題 1.9 より、(14) は $XA^{-1}X = B$ のただひとつの正值な解である。両辺の逆をとると、 $XB^{-1}X = A$ 。これより、 $A\#B = B\#A$ であることがわかる。 $A = (A\#B)B^{-1}(B\#A)$ と、定理 1.17 より、

$$\begin{bmatrix} A & A\#B \\ A\#B & B \end{bmatrix} \geq O.$$

一方、 X が $\begin{bmatrix} A & X \\ X & B \end{bmatrix} \geq O$ を満たすようなエルミート行列ならば、定理 1.17 より、 $A \geq XB^{-1}X$ 。よって、

$$B^{-1/2}AB^{-1/2} \geq B^{-1/2}XB^{-1}XB^{-1/2} = (B^{-1/2}XB^{-1/2})^2.$$

両辺の平方根をとり、 $\Gamma_{B^{1/2}}$ を施すと、

$$B^{1/2}(B^{-1/2}AB^{-1/2})^{1/2}B^{1/2} \geq X$$

を得る。つまり、 $\begin{bmatrix} A & X \\ X & B \end{bmatrix} \geq O$ を満たすエルミート行列 X に対して、 $A\#B \geq X$ となる。ここまでをまとめると、以下ようになる。

定理 3.1 $A, B > O$ とし、 $A\#B = A^{1/2}(A^{-1/2}BA^{-1/2})^{1/2}A^{1/2}$ とおく。このとき、以下が成り立つ。

$$(1) \ A\#B = B\#A,$$

(2) $A\#B$ は $XA^{-1}X = B$ のただひとつの正値な解,

$$(3) A\#B = \max \left\{ X \mid X = X^*, \begin{bmatrix} A & X \\ X & B \end{bmatrix} \geq O \right\}.$$

行列の平均が満たすべき6つの条件は、定理3.1から導くことができる。

幾何平均には他の表し方がある。それを以下で紹介する。

命題 3.2 $A, B > O$ に対して,

$$A\#B = A(A^{-1}B)^{1/2}.$$

ただし、 $A^{-1}B$ の固有値はすべて正であるので、解析関数カルキュラスで $(A^{-1}B)^{1/2}$ が定義できる。

証明 $A^{-1/2}BA^{-1/2}$ を変形すると,

$$\begin{aligned} A^{-1/2}BA^{-1/2} &= A^{1/2}A^{-1}BA^{-1/2} \\ &= \left(A^{1/2}(A^{-1}B)^{1/2}A^{-1/2} \right)^2. \end{aligned}$$

両辺の平方根を取ると,

$$\left(A^{-1/2}BA^{-1/2} \right)^{1/2} = A^{1/2}(A^{-1}B)^{1/2}A^{-1/2}.$$

左右から $A^{1/2}$ をかけると,

$$A^{1/2} \left(A^{-1/2}BA^{-1/2} \right)^{1/2} A^{1/2} = A(A^{-1}B)^{1/2}. \quad \blacksquare$$

命題 3.3 $A, B > O$ に対して,

$$A\#B = (AB^{-1})^{1/2}B.$$

証明 命題3.2と同様に、 $A^{-1/2}BA^{-1/2}$ を変形すればよい。 ■

命題 3.4 $A, B > O$ に対して,

$$A\#B = (A+B) \left\{ (A+B)^{-1}A(A+B)^{-1}B \right\}^{1/2}.$$

証明 $X = (X^{-1} + I)^{-1}(I + X)$ より,

$$\begin{aligned} A^{-1}B &= (B^{-1}A + I)^{-1}(I + A^{-1}B) \\ &= \left((A+B)^{-1}(AB^{-1})^{-1/2}(A+B) \right)^2. \end{aligned}$$

両辺の平方根を取ると,

$$(A^{-1}B)^{1/2} = (A+B)^{-1}(AB^{-1})^{-1/2}(A+B).$$

右から順に $(A+B)^{-1}$, $(AB^{-1})^{1/2}$, B , 左から A をかけると,

$$A(A^{-1}B)^{1/2}(A+B)^{-1}(AB^{-1})^{1/2}B = A(A+B)^{-1}B.$$

命題 3.2 と命題 3.3 より,

$$(A\#B)(A+B)^{-1}(A\#B) = A(A+B)^{-1}B.$$

左から $(A+B)^{-1}$ をかけると,

$$\left((A+B)^{-1}(A\#B)\right)^2 = (A+B)^{-1}A(A+B)^{-1}B.$$

両辺の平方根を取ると,

$$(A+B)^{-1}(A\#B) = \left\{(A+B)^{-1}A(A+B)^{-1}B\right\}^{1/2}. \blacksquare$$

命題 3.5 2×2 の $A, B > 0$ が $\det A = \det B = 1$ のとき,

$$A\#B = \frac{A+B}{\sqrt{\det(A+B)}}.$$

証明 命題 3.2 を用いる. $X := (A^{-1}B)^{1/2}$ とすると, $\det X = 1$ なので, X は 2 つの正の固有値 $\lambda, 1/\lambda$ を持つ.

$$\begin{aligned} \det(A+B) &= \det\{A(I+A^{-1}B)\} \\ &= (\det A)\{\det(I+A^{-1}B)\} \\ &= \det(I+A^{-1}B) \\ &= \det(I+X^2) \\ &= (\lambda+1/\lambda)^2. \end{aligned}$$

よって, $\operatorname{tr} X = \sqrt{\det(A+B)}$. ケーリー・ハミルトンの定理より,

$$X^2 - \sqrt{\det(A+B)}X + I = O.$$

左から A をかけて, 両辺を整理すると,

$$A(A^{-1}B)^{1/2} = \frac{A+B}{\sqrt{\det(A+B)}}. \blacksquare$$

3.3 古田の不等式

$A \geq B \geq O$ のとき, $A^2 \geq B^2$ は一般に成り立たない. この節では, $A \geq B \geq O$ に関する不等式について考える.

定理 3.6 $A \geq B \geq O$ のとき, 任意の $0 \leq r \leq 1$ に対して, $A^r \geq B^r$.

証明 $r = 0, r = 1$ のときは明らか. r_1, r_2 を, $A^{r_1} \geq B^{r_1}, A^{r_2} \geq B^{r_2}$ を満たすものとする, 幾何平均の単調性より, $A^{r_1} \# A^{r_2} \geq B^{r_1} \# B^{r_2}$. つまり, $A^{(r_1+r_2)/2} \geq B^{(r_1+r_2)/2}$. $r_1 = 0, r_2 = 1$ とすれば, $A^{1/2} \geq B^{1/2}$ が成立することになる. $r_1 = 0$ と 2 進小数 $1/2$ に対して同様に考えると, $A^{1/4} \geq B^{1/4}$ が成立する. この議論を繰り返し用いることで, 任意の $0 \leq r \leq 1$ に対して $A^r \geq B^r$ が成立する. 連続性より, 任意の $0 \leq r \leq 1$ に対して $A^r \geq B^r$ がいえる. ■

今, $A \geq O, B \geq O$ に対して, 以下の不等式を考える.

- (1) $A^2 \geq B^2$,
- (2) $BA^2B \geq B^4$,
- (3) $(BA^2B)^{1/2} \geq B^2$.

(1) \Rightarrow (2) \Rightarrow (3) は命題 1.2 と命題 1.8 より明らかであるが, $A \geq B$ と仮定しても, (1) と (2) は一般に成立しない. また, (3) には $A \geq B$ という条件が大きく影響していることがわかる.

次に一般的な形を考える.

- (1) $A \geq B \geq O$,
- (2) $A^p \geq B^p$ ($0 \leq p \leq 1$),
- (3) $B^r A^p B^r \geq B^{p+2r}$ ($0 \leq p \leq 1, r \geq 0$),
- (4) $(B^r A^p B^r)^{1/q} \geq (B^{p+2r})^{1/q}$ ($0 \leq p \leq 1, r \geq 0, q \geq 1$).

(1) \Rightarrow (2) \Rightarrow (3) \Rightarrow (4) は命題 1.2 と定理 3.6 より明らか. q にある制限をつけることで, (4) が任意の $p \geq 0$ に対して成立することを述べたのが, 古田の不等式である.

定理 3.7 (古田の不等式) $A \geq B \geq O$ のとき, $p \geq 0, r \geq 0, q \geq 1, q \geq \frac{p+2r}{1+2r}$ に対して,

$$(B^r A^p B^r)^{1/q} \geq B^{(p+2r)/q}. \quad (15)$$

証明 $0 \leq p \leq 1$ のとき, q に制限がなくても (15) が成立することはわかっているので, $p \geq 1$ のときだけを考える. $q \geq \frac{p+2r}{1+2r}$ とすると, $\frac{p+2r}{(1+2r)q} \leq 1$. また, (15) は以下のように書きかえられる.

$$\left((B^r A^p B^r)^{\frac{1+2r}{p+2r}} \right)^{\frac{p+2r}{(1+2r)q}} \geq (B^{1+2r})^{\frac{p+2r}{(1+2r)q}}.$$

$\frac{p+2r}{(1+2r)q} \leq 1$ なので, $p \geq 1, r \geq 0$ に対して,

$$(B^r A^p B^r)^{\frac{1+2r}{p+2r}} \geq B^{1+2r} \quad (16)$$

を示せばよい.

$A, B > 0$ として,

$$B^r A^{p/2} = UP \quad (17)$$

を $B^r A^{p/2}$ の極分解とすると,

$$\begin{aligned} B^r A^p B^r &= B^r A^{p/2} A^{p/2} B^r \\ &= (B^r A^{p/2})(B^r A^{p/2})^* \\ &= UP^2 U^*. \end{aligned}$$

よって, $q > 0$ に対して,

$$(B^r A^p B^r)^{1/q} = UP^{2/q} U^*.$$

(17) を使うと, 次の式を得る.

$$\begin{aligned} B^{-r} (B^r A^p B^r)^{1/q} B^{-r} &= B^{-r} UP^{2/q} U^* B^{-r} \\ &= (A^{p/2} P^{-1} U^*) (UP^{2/q} U^*) (UP^{-1} A^{p/2}) \\ &= A^{p/2} (P^2)^{1/q-1} A^{p/2} \\ &= A^{p/2} (A^{p/2} B^{2r} A^{p/2})^{1/q-1} A^{p/2} \\ &= A^{p/2} (A^{-p/2} B^{-2r} A^{-p/2})^{1-1/q} A^{p/2}. \end{aligned} \quad (18)$$

$0 \leq r < 1/2$ とすると, $A^{2r} \geq B^{2r}$ なので,

$$B^{-2r} \geq A^{-2r}. \quad (19)$$

また, $q = \frac{p+2r}{1+2r}$ とすると, $1 - \frac{1}{q} = \frac{p-1}{p+2r} \leq 1$. (18) と (19) より,

$$B^{-r} (B^r A^p B^r)^{1/q} B^{-r} \geq A^{p/2} (A^{-p/2} A^{-2r} A^{-p/2})^{(p-1)/(p+2r)} A^{p/2} = A.$$

よって,

$$(B^r A^p B^r)^{1/q} \geq B^r A B^r \geq B^{1+2r}. \quad (20)$$

これで $0 \leq r \leq 1/2$ に対して (16) を示したことになるが, r の範囲は以下のように帰納的に拡張できる.

$r \in [0, 1/2]$, $q = (p+2r)/(1+2r)$ に対して,

$$A_1 := (B^r A^p B^r)^{1/q}, \quad B_1 := B^{1+2r} \quad (21)$$

とおく. また, p_1 と r_1 を $p_1 \geq 1$, $r_1 \in [0, 1/2]$ なる任意の数として, $q_1 := (p_1 + 2r_1)/(1 + 2r_1)$ とおく. このとき, (20) に $A = A_1$, $B = B_1$, $q = q_1$, $r = r_1$, $p = p_1$ を代入すると,

$$(B_1^{r_1} A_1^{p_1} B_1^{r_1})^{1/q_1} \geq B_1^{1+2r_1}. \quad (22)$$

特に $p_1 = q$, $r_1 = 1/2$ とすると,

$$(B_1^{1/2} A_1^q B_1^{1/2})^{1/q_1} \geq B_1^2.$$

これに (21) を代入すると,

$$(B^{2r+1/2} A^p B^{2r+1/2})^{1/q_1} \geq B^{2(1+2r)}. \quad (23)$$

$s := 2r + 1/2$ とすると,

$$\begin{aligned} q_1 &= \frac{p_1 + 2r_1}{1 + 2r_1} \\ &= \frac{q + 1}{2} \\ &= \frac{(p + 2r)/(1 + 2r) + 1}{2} \\ &= \frac{p + 2s}{1 + 2s}. \end{aligned}$$

これを用いて (23) を書きかえると,

$$(B^s A^p B^s)^{(1+2s)/(p+2s)} \geq B^{1+2s} \quad (s \in [1/2, 3/2]).$$

これにより, (16) の不等式が $0 \leq r \leq 3/2$ に対して成り立つことを示せた. 以下同様にして, 帰納的に任意の $r \geq 0$ に対して (16) が成り立つ. ■

系 3.8 $A \geq B \geq O$ のとき, $p \geq 0$, $r \geq 0$, $q \geq 1$, $q \geq \frac{p+2r}{1+2r}$ に対して,

$$(A^{p+2r})^{1/q} \geq (A^r B^p A^r)^{1/q}. \quad (24)$$

証明 $A, B > O$ とすると, $B^{-1} \geq A^{-1} > O$ なので, (15) より,

$$(A^{-r} B^{-p} A^{-r})^{1/q} \geq A^{-(p+2r)/q}.$$

両辺の逆をとると,

$$A^{(p+2r)/q} \geq (A^r B^p A^r)^{1/q}. \quad \blacksquare$$

系 3.9 $A \geq B \geq O$ とすると, このとき, $0 < p < \infty$ に対して,

$$A^p \# B^{-p} \geq I.$$

証明 (24) において $r = p/2$, $q = 2$ とすると,

$$A^p \geq (A^{p/2} B^p A^{p/2})^{1/2}.$$

左右から $A^{-p/2}$ を両辺にかけると,

$$\begin{aligned} I &\geq A^{-p/2} (A^{p/2} B^p A^{p/2})^{1/2} A^{-p/2} \\ &= A^{-p} \# B^p \\ &= (A^p \# B^{-p})^{-1}. \quad \blacksquare \end{aligned}$$

3.4 正定値関数

定義 3.10 (i, j) 成分が A_{ij} であるブロック行列を, $[A_{ij}]$ で表す.

定義 3.11 $\{a_n \mid n \in \mathbb{Z}, a_n \in \mathbb{C}\}$ が正値な列とは, 任意の $N \in \mathbb{N}$ に対して,

$$\sum_{r,s=0}^{N-1} a_{r-s} \xi_r \bar{\xi}_s \geq 0 \quad (\xi_0, \xi_1, \dots, \xi_{N-1} \in \mathbb{C}) \quad (25)$$

となることをいう.

(25) は, 任意の $N \in \mathbb{N}$ に対して,

$$\begin{bmatrix} a_0 & a_{-1} & \cdots & a_{-(N-1)} \\ a_1 & a_0 & a_{-1} & \cdots \\ \vdots & & \cdots & a_{-1} \\ a_{N-1} & \cdots & a_1 & a_0 \end{bmatrix} \geq O$$

であることと同値. これより, $a_0 \geq 0, a_n = \bar{a}_{-n}, |a_n| \leq a_0$ がわかる.

定義 3.12 $\varphi: \mathbb{R} \rightarrow \mathbb{C}$ が正定値関数とは, 任意の $N \in \mathbb{N}$ に対して,

$$\sum_{r,s=0}^{N-1} \varphi(x_r - x_s) \xi_r \bar{\xi}_s \geq 0 \quad (x_0, x_1, \dots, x_{N-1} \in \mathbb{R}, \xi_0, \xi_1, \dots, \xi_{N-1} \in \mathbb{C}) \quad (26)$$

となることをいう.

(26) は, 任意の $N \in \mathbb{N}$ に対して,

$$[\varphi(x_r - x_s)] \geq O \quad (x_0, x_1, \dots, x_{N-1} \in \mathbb{R})$$

であることと同値. これより, $\varphi(0) \geq 0, \varphi(-x) = \overline{\varphi(x)}, |\varphi(x)| \leq \varphi(0)$ がわかる.

命題 3.13 $-1 \leq \alpha \leq 1$ に対して, $\cosh \alpha x / \cosh x$ は正定値関数.

証明

$$\cosh x = \prod_{k=1}^{\infty} \left(1 + \frac{4x^2}{(2k+1)^2\pi^2} \right)$$

を用いると,

$$\frac{\cosh \alpha x}{\cosh x} = \alpha \prod_{k=1}^{\infty} \frac{1 + 4\alpha^2 x^2 / (2k+1)^2 \pi^2}{1 + 4x^2 / (2k+1)^2 \pi^2}$$

となるが,

$$\frac{1 + b^2 x^2}{1 + a^2 x^2} = \frac{b^2}{a^2} + \frac{1 - b^2/a^2}{1 + a^2 x^2}, \quad (0 \leq b < a)$$

が正定値関数であることより, $\cosh \alpha x / \cosh x$ は正定値関数. ■

これと同様に,

$$\frac{\sinh x}{x} = \prod_{k=1}^{\infty} \left(1 + \frac{x^2}{k^2\pi^2}\right)$$

を用いることで, 次の系を得る.

系 3.14 $0 < \alpha < 1$ に対して, $\sinh \alpha x / \sinh x$ は正定値関数.

3.5 ノルム不等式

幾何平均と算術平均に関する不等式として, $\sqrt{ab} \leq \frac{1}{2}(a+b)$ がある. この形に似ていて, 正値行列 A, B を使う不等式を考える. 任意の X に対して,

$$\| \|A^{1/2}XB^{1/2}\| \| \leq \frac{1}{2} \| \|AX + XB\| \| . \quad (27)$$

この不等式は, $A = B$ のときだけを証明すれば十分. 実際, 任意の X と $A > O$ に対して

$$\| \|A^{1/2}XA^{1/2}\| \| \leq \frac{1}{2} \| \|AX + XA\| \| \quad (28)$$

を考えると, 任意の X_0 と $A_0, B_0 > O$ に対して, (28) の A と X を

$$A = \begin{bmatrix} A_0 & O \\ O & B_0 \end{bmatrix}, X = \begin{bmatrix} O & X_0 \\ O & O \end{bmatrix}$$

とおくと, (27) が得られる. (28) はユニタリ不変ノルムなので, $A = \text{diag}(\lambda_1, \dots, \lambda_n)$ とすると, $Y = \begin{bmatrix} 2\sqrt{\lambda_i\lambda_j} \\ \lambda_i + \lambda_j \end{bmatrix}$ に対して,

$$A^{1/2}XA^{1/2} = Y \circ \left(\frac{AX + XA}{2} \right) = [\sqrt{\lambda_i}x_{ij}\sqrt{\lambda_j}] \quad (29)$$

となり, また, $Y = \begin{bmatrix} 2\sqrt{\lambda_i\lambda_j} \\ \lambda_i + \lambda_j \end{bmatrix} \sim \begin{bmatrix} 1 \\ \lambda_i + \lambda_j \end{bmatrix}$ である. そして, $\begin{bmatrix} 1 \\ \lambda_i + \lambda_j \end{bmatrix} \geq O$ なので, $Y \geq O$. Y の対角成分が 1 なので, 命題 1.6 より,

$$\| \|A^{1/2}XA^{1/2}\| \| = \| \|Y \circ \left(\frac{AX + XA}{2} \right)\| \| \leq \| \left\| \frac{AX + XA}{2} \right\| \| .$$

これで (28) が得られた.

定理 3.15 任意の X と $A, B > O, 0 \leq v \leq 1$ に対して,

$$2 \| \|A^{1/2}XB^{1/2}\| \| \leq \| \|A^vXB^{1-v} + A^{1-v}XB^v\| \| \leq \| \|AX + XB\| \| . \quad (30)$$

証明 (27) の証明で用いた議論を使う. (30) の右側の不等式を証明するために,

$$Y = [y_{ij}] = \left[\frac{\lambda_i^v \lambda_j^{1-v} + \lambda_i^{1-v} \lambda_j^v}{\lambda_i + \lambda_j} \right] \quad (31)$$

が $0 < v < 1$ で正値行列となることを示す必要がある. y_{ij} を変形すると,

$$y_{ij} = \lambda_i^{1-v} \left(\frac{\lambda_i^{2v-1} + \lambda_j^{2v-1}}{\lambda_i + \lambda_j} \right) \lambda_j^{1-v}$$

となり,

$$Y \sim Z = [z_{ij}] = \left[\frac{\lambda_i^\alpha + \lambda_j^\alpha}{\lambda_i + \lambda_j} \right] \quad (-1 < \alpha < 1)$$

であることがわかる. $\lambda_i > 0$ なので, λ_i を e^{x_i} にかえて z_{ij} を変形すると,

$$\begin{aligned} v_{ij} &= \frac{e^{\alpha x_i} - e^{\alpha x_j}}{e^{x_i} - e^{x_j}} \\ &= \frac{e^{\alpha x_i/2} \cosh \alpha(x_i - x_j)/2}{e^{x_i/2} \cosh(x_i - x_j)/2} \frac{e^{\alpha x_j/2}}{e^{x_j/2}}. \end{aligned}$$

Z の正値性を確かめるには, $-1 < \alpha < 1$ における $\cosh \alpha x / \cosh x$ の正値性を考えればよいが, 命題 3.13 より, $\cosh \alpha x / \cosh x$ は正値関数. よって, (27) の証明での議論と同様に $A = \text{diag}(\lambda_1, \dots, \lambda_n)$ として, $A = B$ のとき,

$$\begin{aligned} A^v X A^{1-v} + A^{1-v} X A^v &= Y \circ (AX + XA) \\ &= \left[\lambda_i^v \lambda_j^{1-v} x_{ij} + \lambda_i^{1-v} \lambda_j^v x_{ij} \right]. \end{aligned}$$

$Y \geq O$, $y_{ii} = 1$ なので,

$$\| \| Y \circ (AX + XA) \| \| \leq \| \| AX + XA \| \|.$$

これで (30) の右の不等式は示された.

左側は,

$$Y = [y_{ij}] = \left[\frac{2\sqrt{\lambda_i \lambda_j}}{\lambda_i^v \lambda_j^{1-v} + \lambda_i^{1-v} \lambda_j^v} \right]$$

とすると, $Y \sim Z = \left[\frac{1}{\lambda_i^\alpha + \lambda_j^\alpha} \right] \geq O$ ($-1 < \alpha < 1$) なので, $Y \geq O$ と $y_{ii} = 1$ より,

$$\begin{aligned} 2 \| \| A^{1/2} X B^{1/2} \| \| &= \| \| Y \circ (A^v X B^{1-v} + A^{1-v} X B^v) \| \| \\ &\leq \| \| A^v X B^{1-v} + A^{1-v} X B^v \| \|. \quad \blacksquare \end{aligned}$$

定理 3.16 任意の X と $A, B > O$, $0 \leq v \leq 1$ に対して,

$$\| \| A^{1/2} X B^{1/2} \| \| \leq \left\| \int_0^1 A^t X B^{1-t} dt \right\| \leq \frac{1}{2} \| \| AX + XB \| \|. \quad (32)$$

証明 $A = B$, $A = \text{diag}(\lambda_1 \dots \lambda_n)$ とし,

$$Y = [y_{ij}] = \left[\frac{\lambda_i^{1/2} \lambda_j^{1/2} (\log \lambda_i - \log \lambda_j)}{\lambda_i - \lambda_j} \right]$$

とすると,

$$A^{1/2} X A^{1/2} = Y \circ \left(\int_0^1 A^t X A^{1-t} dt \right).$$

となり,

$$Y = [y_{ij}] \sim Z = [z_{ij}] = \left[\frac{\log \lambda_i - \log \lambda_j}{\lambda_i - \lambda_j} \right].$$

右辺において $\lambda_i = e^{x_i}$ とおくと,

$$\begin{aligned} \left[\frac{\log \lambda_i - \log \lambda_j}{\lambda_i - \lambda_j} \right] &= \frac{x_i - x_j}{e^{x_i} - e^{x_j}} \\ &= \frac{1}{e^{x_i/2}} \frac{(x_i - x_j)/2}{\sinh(x_i - x_j)/2} \frac{1}{e^{x_j/2}}. \end{aligned}$$

$x/(\sinh x)$ は正値関数なので, $Z \geq O$ である. よって, $Y \geq O$ なので,

$$\| \| Y \circ \left(\int_0^1 A^t X A^{1-t} dt \right) \| \| \leq \| \| \int_0^1 A^t X A^{1-t} dt \| \|.$$

これで左側の証明ができた.

右側の証明も同様に,

$$W = [w_{ij}] = \left[\frac{2(\lambda_i - \lambda_j)}{(\log \lambda_i - \log \lambda_j)(\lambda_i + \lambda_j)} \right]$$

とすると,

$$\int_0^1 A^t X A^{1-t} dt = W \circ \left\{ \frac{1}{2}(AX + XA) \right\}.$$

w_{ij} において $\lambda_i = e^{x_i}$ とすると,

$$\begin{aligned} w_{ij} &= \frac{2(e^{x_i} - e^{x_j})}{(x_i - x_j)(e^{x_i} + e^{x_j})} \\ &= \frac{\tanh(x_i - x_j)/2}{(x_i - x_j)/2} \end{aligned}$$

$(\tanh x)/x$ は正値関数なので, $W \geq O$ である. よって,

$$\| \| W \circ \left\{ \frac{1}{2}(AX + XA) \right\} \| \| \leq \frac{1}{2} \| \| AX + XA \| \|.$$

これで (32) の右側も証明できた. ■

次の2つの不等式は以下の定理等の証明に役に立つ. 任意の A, B, X に対して,

$$2\|A^*XB\| \leq \|AA^*X + XBB^*\|, \quad (33)$$

$$\|A^*XB\|^2 \leq \|AA^*X\| \|XBB^*\|. \quad (34)$$

定理 3.17 任意の X と $A, B \geq O, 0 \leq v \leq 1$ に対して,

$$\|A^vXB^{1-v}\| \leq \|AX\|^v \|XB\|^{1-v}.$$

証明 $v = 0, 1$ のとき明らかに成立する. 帰納法を用いて示す. $\nu := (2k+1)/2^n, \mu := k/2^{n-1}, \rho := 1/2^n$ とする. 今, 分母が 2^{n-1} である任意の有理数 μ と $\lambda := \mu + 2\rho = \nu + \rho$ に対して, 不等式が成り立つと仮定する. このとき,

$$\begin{aligned} \|A^\nu XB^{1-\nu}\| &= \|A^{\mu+\rho}XB^{1-\lambda+\rho}\| \\ &= \|A^\rho(A^\mu XB^{1-\lambda})B^\rho\| \\ &\leq \|A^{2\rho}(A^\mu XB^{1-\lambda})\|^{1/2} \|(A^\mu XB^{1-\lambda})B^{2\rho}\|^{1/2} \\ &= \|A^\lambda XB^{1-\lambda}\|^{1/2} \|A^\mu XB^{1-\mu}\|^{1/2} \\ &\leq \|AX\|^{\lambda/2} \|XB\|^{(1-\lambda)/2} \|AX\|^{\mu/2} \|XB\|^{(1-\mu)/2} \\ &= \|AX\|^{(\lambda+\mu)/2} \|XB\|^{1-(\lambda+\mu)/2} \\ &= \|AX\|^\nu \|XB\|^{1-\nu}. \end{aligned}$$

上の1番目の不等式は (34) による. ■

系 3.18 任意の X と $A, B \geq O, 0 \leq v \leq 1$ に対して,

$$\|A^vXB^v\| \leq \|X\|^{1-v} \|AXB\|^v.$$

証明 任意の $\varepsilon > 0$ に対して $A_\varepsilon := A + \varepsilon I$ とおくと, $A_\varepsilon > O$ なので, A_ε は正則行列. 定理 3.17 より,

$$\begin{aligned} \|A_\varepsilon^vXB^v\| &= \|(A_\varepsilon^{-1})^{1-v}A_\varepsilonXB^{1-(1-v)}\| \\ &\leq \|A_\varepsilon^{-1}A_\varepsilonX\|^{1-v} \|AXB\|^v \\ &= \|X\|^{1-v} \|A_\varepsilonXB\|^v. \end{aligned}$$

$\varepsilon \rightarrow 0$ とすれば, $\|A^vXB^v\| \leq \|X\|^{1-v} \|AXB\|^v$ となる. ■

命題 3.19 任意の X とエルミート行列 H, K に対して,

$$\|(\sin H)X(\cos K) - (\cos H)X(\sin K)\| \leq \|HX - XK\|.$$

証明 (27) の証明で用いた議論を使う. $A = \text{diag}(\lambda_1 \dots \lambda_n)$ とすると,

$$W = [w_{ij}] = \left[\frac{\sin(\lambda_i - \lambda_j)}{\lambda_i - \lambda_j} \right] \geq O$$

に対して,

$$(\sin A)X(\cos A) - (\cos A)X(\sin A) = W \circ (AX - XA).$$

よって,

$$\| \|W \circ (AX - XA)\| \| \leq \| \|AX - XA\| \|. \quad \blacksquare$$

同様の証明で, 次の不等式を得る.

命題 3.20 任意の X とエルミート行列 H, K に対して,

$$\| \|(\sin H)X(\cos K) + (\cos H)X(\sin K)\| \| \leq \| \|HX + XK\| \|.$$

命題 3.21 任意の X と $A, B > O$ に対して,

$$\| \|(\log A)X - X(\log B)\| \| \leq \| \|A^{1/2}XB^{-1/2} - A^{-1/2}XB^{1/2}\| \|.$$

証明 $D := \text{diag}(\sqrt{\lambda_1} \dots \sqrt{\lambda_n})$ とし,

$$v_{ij} := \frac{\log \lambda_i - \log \lambda_j}{\lambda_i - \lambda_j} \quad (i \neq j)$$

$$v_{ii} := \frac{1}{\lambda_i}$$

によって定めた $V = [v_{ij}] \geq O$ を用いて, $W := DVD$ とおくと $W \geq O$. このとき,

$$(\log A)X - X(\log B) = W \circ (A^{1/2}XB^{-1/2} - A^{-1/2}XB^{1/2})$$

となり,

$$\| \|W \circ (A^{1/2}XB^{-1/2} - A^{-1/2}XB^{1/2})\| \| \leq \| \|A^{1/2}XB^{-1/2} - A^{-1/2}XB^{1/2}\| \|. \quad \blacksquare$$

系 3.22 任意の X とエルミート行列 H, K に対して,

$$\| \|HX - XK\| \| \leq \| \|e^{H/2}Xe^{-K/2} - e^{-H/2}Xe^{K/2}\| \|.$$

証明 命題 3.21 において, $A := e^H, B := e^K$ とおけば, $A, B > O$. \blacksquare

命題 3.23 任意の X と $A, B \geq O, 0 \leq v \leq 1$ に対して,

$$\| \|A^vXB^{1-v} - A^{1-v}XB^v\| \| \leq |2v - 1| \| \|AX - XB\| \|.$$

証明 $v = 0, 1$ のとき明らかに成立する. 帰納法を用いて示す. $\nu := (2k+1)/2^n$, $\mu := k/2^{n-1}$, $\rho := 1/2^n$ とする. 今, 分母が 2^{n-1} である任意の有理数 μ と $\lambda := \mu + 2\rho = \nu + \rho$ に対して, 不等式が成り立つと仮定する. このとき,

$$\begin{aligned}
\|A^\nu X B^{1-\nu} - A^{1-\nu} X B^\nu\| &= \|A^{\mu+\rho} X B^{1-\lambda+\rho} - A^{1-\lambda+\rho} X B^{\mu+\rho}\| \\
&= \|A^\rho (A^\mu X B^{1-\lambda} - A^{1-\lambda} X B^\mu) B^\rho\| \\
&\leq \frac{1}{2} \|A^{2\rho} (A^\mu X B^{1-\lambda} - A^{1-\lambda} X B^\mu) + (A^\mu X B^{1-\lambda} - A^{1-\lambda} X B^\mu) B^{2\rho}\| \\
&= \frac{1}{2} \|A^\lambda X B^{1-\lambda} - A^{1-\mu} X B^\mu + A^\mu X B^{1-\mu} - A^{1-\lambda} X B^\lambda\| \\
&\leq \frac{1}{2} \left(\|A^\lambda X B^{1-\lambda} - A^{1-\lambda} X B^\lambda\| + \|A^\mu X B^\mu - A^{1-\mu} X B^\mu\| \right) \\
&\leq \frac{1}{2} (|2\lambda - 1| + |2\mu - 1|) \|AX - XB\|.
\end{aligned}$$

上の1番目の不等式は(33)による. $\nu < 1/2$ のとき,

$$\begin{aligned}
\frac{1}{2} (|2\lambda - 1| + |2\mu - 1|) &= \frac{1}{2} (1 - 2\lambda + 1 - 2\mu) \\
&= 1 - 2 \cdot \frac{\lambda + \mu}{2} \\
&= 1 - 2\nu \\
&= |2\nu - 1|.
\end{aligned}$$

$\nu > 1/2$ のとき,

$$\begin{aligned}
\frac{1}{2} (|2\lambda - 1| + |2\mu - 1|) &= 2\nu - 1 \\
&= |2\nu - 1|.
\end{aligned}$$

$\nu = 1/2$ のとき, 両辺 = 0 で成立. ■

系 3.24 任意の Z に対して,

$$\operatorname{Re} Z := \frac{1}{2}(Z + Z^*), \quad \operatorname{Im} Z := \frac{1}{2i}(Z - Z^*)$$

とおき, 任意の $A \geq 0$, エルミート行列 X , $0 \leq v \leq 1$ に対して,

$$S := A^v X A^{1-v}, \quad T := vAX + (1-v)XA$$

とおく. このとき,

$$\|\operatorname{Re} S\| \leq \|\operatorname{Re} T\|, \quad \|\operatorname{Im} S\| \leq \|\operatorname{Im} T\|.$$

証明 定理 3.15 の 2 番目の不等式と命題 3.23 から容易. ■

4 正値行列の幾何平均

4.1 リーマン距離

\mathbb{M}_n を $\langle A, B \rangle = \text{tr} A^* B$ で定義された内積, $\|A\|_2 = (\text{tr} A^* A)^{1/2}$ で定義されたノルムを持つヒルベルト空間とする. \mathbb{H}_n をエルミート行列全体とすると, \mathbb{H}_n は \mathbb{M}_n の閉部分集合となる. \mathbb{P}_n を正値行列全体とすると, \mathbb{P}_n は \mathbb{H}_n の開部分集合となるので, \mathbb{P}_n は微分可能多様体である. A における距離は微分

$$ds = \|A^{-1/2} dA A^{-1/2}\|_2 = \left[\text{tr} (A^{-1} dA)^2 \right]^{1/2} \quad (35)$$

によって与えられる. これは, \mathbb{P}_n における微分可能な路の長さを計算することに役立つ. $\gamma: [a, b] \rightarrow \mathbb{P}_n$ が微分可能な路のとき, その長さを

$$L(\gamma) = \int_a^b \|\gamma^{-1/2}(t) \gamma'(t) \gamma^{-1/2}(t)\|_2 dt \quad (36)$$

で定義する. 任意の $X \in GL(n)$ に対して, $\Gamma_X(A) = X^* A X$ は \mathbb{P}_n 上の全単射となる. また, $\Gamma_X \circ \gamma$ も \mathbb{P}_n における微分可能な路である.

補題 4.1 任意の $X \in GL(n)$ と微分可能な路 γ に対して,

$$L(\Gamma_X \circ \gamma) = L(\gamma). \quad (37)$$

証明 任意の t に関して

$$\begin{aligned} & \left\| (X^* \gamma(t) X)^{-1/2} (X^* \gamma(t) X)' (X^* \gamma(t) X)^{-1/2} \right\|_2 \\ &= \left[\text{tr} \left((X^* \gamma(t) X)^{-1} (X^* \gamma(t) X)' (X^* \gamma(t) X)^{-1} (X^* \gamma(t) X)' \right)^{1/2} \right]^{1/2} \\ &= \left[\text{tr} X^{-1} \gamma^{-1}(t) \gamma'(t) \gamma^{-1}(t) \gamma'(t) X \right]^{1/2} \\ &= \left[\text{tr} \gamma^{-1}(t) \gamma'(t) \gamma^{-1}(t) \gamma'(t) \right]^{1/2} \\ &= \|\gamma^{-1/2}(t) \gamma'(t) \gamma^{-1/2}(t)\|_2. \end{aligned}$$

t に関して両辺の積分を取れば, $L(\Gamma_X \circ \gamma) = L(\gamma)$ となる. ■

\mathbb{P}_n の任意の 2 点 A, B に対して,

$$\delta_2(A, B) := \inf \{ L(\gamma) \mid \gamma \text{ は } A \text{ から } B \text{ への路} \} \quad (38)$$

とすると, δ_2 は \mathbb{P}_n 上の距離となる.

補題 4.1 より Γ_X は L に対して等長写像であるので, δ_2 に対しても等長写像である. つまり, 任意の $A, B \in \mathbb{P}_n$ と $X \in GL(n)$ に対して

$$\delta_2(\Gamma_X(A), \Gamma_X(B)) = \delta_2(A, B). \quad (39)$$

(38) の下限は, A と B を結ぶただ一つの方法で得られる. その路を「 A から B への測地線」と呼ぶ. 測地線とその長さに関する IEMI と呼ばれる公式があるが, その前に記号を定義する (IEMI は the Infinitesimal Exponential Metric Increasing property の略).

\mathbb{H}_n のある点における指数関数 e^H の導関数を De^H で表す. これは \mathbb{H}_n 上の線形写像で,

$$De^H(K) = \lim_{t \rightarrow 0} \frac{e^{H+tK} - e^H}{t}$$

によって与えられる.

命題 4.2 (IEMI) 任意の $H, K \in \mathbb{H}_n$ に対して

$$\|e^{-H/2} De^H(K) e^{-H/2}\|_2 \geq \|K\|_2. \quad (40)$$

証明 正規直交基底をひとつ選んで, $H := \text{diag}(\lambda_1, \dots, \lambda_n)$ とする. 公式 (3) より,

$$\begin{aligned} De^H(K) &= \begin{bmatrix} e^{\lambda_i} - e^{\lambda_j} \\ \lambda_i - \lambda_j \end{bmatrix} \circ K \\ &= \begin{bmatrix} e^{\lambda_i} - e^{\lambda_j} \\ \lambda_i - \lambda_j \end{bmatrix} k_{ij}. \end{aligned}$$

よって, $e^{-H/2} De^H(K) e^{-H/2}$ の (i, j) 成分は

$$\begin{aligned} &e^{-\lambda_i/2} \cdot \frac{e^{\lambda_i} - e^{\lambda_j}}{\lambda_i - \lambda_j} k_{ij} \cdot e^{-\lambda_j/2} \\ &= \frac{\sinh\{(\lambda_i - \lambda_j)/2\}}{(\lambda_i - \lambda_j)/2} k_{ij}. \end{aligned}$$

$x/(\sinh x)$ は正値関数であるから, (40) を得る. ■

系 4.3 $H(t)$ ($a \leq t \leq b$) を \mathbb{H}_n 内の路 $\gamma(t) := e^{H(t)}$ とすると,

$$L(\gamma) \geq \int_a^b \|H'(t)\|_2 dt. \quad (41)$$

証明 $\gamma'(t) = De^{H(t)}(H'(t))$ なので, (36) と (40) より,

$$\begin{aligned} L(\gamma) &= \int_a^b \|\gamma^{-1/2}(t) \gamma'(t) \gamma^{-1/2}(t)\|_2 dt \\ &= \int_a^b \|e^{-H(t)/2} De^{H(t)}(H'(t)) e^{-H(t)/2}\|_2 dt \\ &\geq \int_a^b \|H'(t)\|_2 dt. \quad \blacksquare \end{aligned}$$

定理 4.4 (EMI) 任意の $A, B \in \mathbb{P}_n$ に対して

$$\delta_2(A, B) \geq \|\log A - \log B\|_2. \quad (42)$$

あるいは, 任意の $H, K \in \mathbb{H}_n$ に対して

$$\delta_2(e^H, e^K) \geq \|H - K\|_2. \quad (43)$$

証明 $\gamma(t)$ を \mathbb{P}_n における A から B への路とすると, $H(t) = \log \gamma(t)$ は \mathbb{H}_n における $\log A$ から $\log B$ への路である. (41) の右辺は \mathbb{H}_n におけるその長さである. これは $\log A$ から $\log B$ を結ぶ線分の長さによって, 下から押さえられる. よって,

$$L(\gamma) \geq \int_a^b \|H'(t)\|_2 dt \geq \|\log A - \log B\|_2.$$

この不等式から (42) を得る. ■

A と B が交換可能であるという条件を加えると, 次に示すように (42) の不等式は等式となる. また, 写像 e^H は \mathbb{H}_n 内の $\log A$ から $\log B$ への線分を, \mathbb{P}_n 内の A から B への測地線へと写す.

ここで, 記号を定義する. \mathbb{H}_n 内の H から K への線分

$$H(t) = (1-t)H + tK, \quad (0 \leq t \leq 1)$$

を $[H, K]$ で表す. また, \mathbb{P}_n 内の A から B への測地線を $[A, B]$ で表す.

命題 4.5 $A, B \in \mathbb{P}_n$ を交換可能な行列とする. このとき, 写像 e^H は $[\log A, \log B] \subset \mathbb{H}_n$ を $[A, B] \subset \mathbb{P}_n$ へ写し,

$$\delta_2(A, B) = \|\log A - \log B\|_2. \quad (44)$$

証明 命題を証明するには, $\gamma(t) = e^{H(t)} = \exp\{(1-t)\log A + t\log B\}$ ($0 \leq t \leq 1$) が (\mathbb{P}_n, δ_2) 内の A から B への唯一の最短路であることを確かめればよい.

A と B は交換可能なので, $\gamma(t) = A^{1-t}B^t$. よって, $\gamma'(t) = (\log B - \log A)\gamma(t)$. (36) より,

$$L(\gamma) = \int_0^1 \|\log A - \log B\|_2 dt = \|\log A - \log B\|_2.$$

(42) より, この路 γ が最短であることがわかる.

次に, A から B への路のうち, γ と同じ長さの路を $\tilde{\gamma}$ とすると, $\tilde{H}(t) = \log \tilde{\gamma}(t)$ は \mathbb{H}_n 内の $\log A$ から $\log B$ への道であり, 系 4.3 より, その長さは $\|\log A - \log B\|_2$ 以下である. ユークリッド空間では, 2 点を結ぶ直線はいちばん短く, ひとつしかない. よって, $\tilde{H}(t)$ は線分 $[\log A, \log B]$ である. ■

$[0, 1]$ の中の区間 $[0, a]$ に対して上の証明を用いると, $[\log A, \log B]$ のパラメーター表示

$$H(t) = (1-t)\log A + t\log B$$

は, すべての区間で等距離に $[A, B]$ へ写像される. つまり, A と B が交換可能なときの $[A, B]$ の自然なパラメーター表示は,

$$\gamma(t) = A^{1-t}B^t$$

によって与えられる (任意の t に対して, $\delta_2(A, \gamma(t)) = t\delta_2(A, B)$). 一般的な場合は, 等距離写像 Γ_X によって得られる.

定理 4.6 任意の $A, B \in \mathbb{P}_n$ に対して, 測地線 $[A, B]$ はただひとつ存在する. そのパラメーター表示は

$$\gamma(t) = A^{1/2} \left(A^{-1/2} B A^{-1/2} \right)^t A^{1/2}, \quad 0 \leq t \leq 1 \quad (45)$$

で表せる. また, 各 t に対して

$$\delta_2(A, \gamma(t)) = t \delta_2(A, B). \quad (46)$$

さらに,

$$\delta_2(A, B) = \|\log A^{-1/2} B A^{-1/2}\|_2. \quad (47)$$

証明 I と $A^{-1/2} B A^{-1/2}$ は交換可能なので, $[I, A^{-1/2} B A^{-1/2}]$ のパラメーター表示は,

$$\gamma_0(t) = \left(A^{-1/2} B A^{-1/2} \right)^t.$$

等距離写像 $\Gamma_{A^{1/2}}$ を両辺に施すと,

$$\gamma(t) = \Gamma_{A^{1/2}}(\gamma_0(t)) = A^{1/2} \left(A^{-1/2} B A^{-1/2} \right)^t A^{1/2}$$

により, $\Gamma_{A^{1/2}}(I) = A$ から $\Gamma_{A^{1/2}}(A^{-1/2} B A^{-1/2}) = B$ への路を得る. $\Gamma_{A^{1/2}}$ は等距離写像なので, この路は測地線 $[A, B]$ である. 命題 4.5 より,

$$\begin{aligned} \delta_2(A, B) &= \delta_2(I, A^{-1/2} B A^{-1/2}) \\ &= \|\log I - \log(A^{-1/2} B A^{-1/2})\|_2 \\ &= \|\log A^{-1/2} B A^{-1/2}\|_2. \quad \blacksquare \end{aligned}$$

(38) で定義した距離は, (47) によって具体的な表示が与えられた. これは多様体 \mathbb{P}_n 上のリーマン距離であり, $\|\cdot\|_2$ の定義より,

$$\delta_2(A, B) = \left(\sum_{i=1}^n \log^2 \lambda_i(A^{-1} B) \right)^{1/2}. \quad (48)$$

ただし, λ_i は $A^{-1} B$ の固有値.

命題 4.7 ある $A, B \in \mathbb{P}_n$ に対して $I \in [A, B]$ とすると, A と B は交換可能で, $[A, B]$ は O を含む直線を写像 e^H で写した等距離像であり,

$$\log B = -\frac{1-\xi}{\xi} \log A \quad \left(\xi := \frac{\delta_2(A, I)}{\delta_2(A, B)} \right). \quad (49)$$

証明 定理 4.6 より,

$$I = A^{1/2} \left(A^{-1/2} B A^{-1/2} \right)^\xi A^{1/2} \quad \left(\xi := \frac{\delta_2(A, I)}{\delta_2(A, B)} \right).$$

よって,

$$B = A^{1/2} A^{-1/\xi} A^{1/2} = A^{-(1-\xi)/\xi}. \quad (50)$$

A と B は交換可能であり, (49) を満たす. 命題 4.5 より, 写像 e^H は線分 $[\log A, \log B]$ を測地線 $[A, B]$ へ等距離に写す. よって, $O = \log I \in [\log A, \log B]$. ■

δ_2 の特徴は半中線定理を満たすことである. その理解の補助として, ヒルベルト空間 \mathcal{H} における中線定理をここで考える. $a, b \in \mathcal{H}$, $m := (a+b)/2$ とする. ある点 $c \in \mathcal{H}$ が与えられたとき, 対角線が $[a, b]$ と $[c, d]$ になるような平行四辺形を考える. 2本の対角線は点 m で交わり, 中線定理は

$$\|a-b\|^2 + \|c-d\|^2 = 2(\|a-c\|^2 + \|b-c\|^2)$$

となる. これを書きかえると,

$$\|c-m\|^2 = \frac{\|a-c\|^2 + \|b-c\|^2}{2} - \frac{\|a-b\|^2}{4}.$$

半中線定理では, 等式ではなく不等式となる. 次の定理は δ_2 の半中線定理である.

定理 4.8 (半中線定理) A, B を \mathbb{P}_n の任意の2点とし, $M := A\#B$ ($[A, B]$ の中点) とすると, 任意の点 $C \in \mathbb{P}_n$ に対して

$$\delta_2^2(M, C) \leq \frac{\delta_2^2(A, C) + \delta_2^2(B, C)}{2} - \frac{\delta_2^2(A, B)}{4}. \quad (51)$$

証明 すべての行列に $\Gamma_{M^{-1/2}}$ を施して, $M = I$ と考える. I は $[A, B]$ の中点なので, 命題 4.7 より,

$$\log B = -\frac{1-\frac{1}{2}}{\frac{1}{2}} \log A = -\log A.$$

また, 命題 4.7 と $I \in [A, B]$ より A と B は交換可能なので, 命題 4.5 より,

$$\delta_2(A, B) = \|\log A - \log B\|_2.$$

$[M, C] = [I, C]$ に対しても同様にして,

$$\delta_2(M, C) = \|\log M - \log C\|_2.$$

ヒルベルト空間 $(\mathbb{H}_n, \|\cdot\|_2)$ における中線定理より,

$$\|\log M - \log C\|_2^2 = \frac{\|\log A - \log C\|_2^2 + \|\log B - \log C\|_2^2}{2} - \frac{\|\log A - \log B\|_2^2}{4}.$$

定理 4.4 より,

$$\delta_2^2(M, C) \leq \frac{\delta_2^2(A, C) + \delta_2^2(B, C)}{2} - \frac{\delta_2^2(A, B)}{4}. \quad \blacksquare$$

ユークリッド空間において、三角形の二辺の中点の距離は、もう一辺の長さの半分に等しい（中点連結定理）。半中線定理を満たす距離が存在する空間においては、中点連結定理は等式ではなく不等式で表される。

命題 4.9 任意の $A, B, C \in \mathbb{P}_n$ に対して

$$\delta_2(A\#B, A\#C) \leq \frac{\delta_2(B, C)}{2}. \quad (52)$$

証明 A, B, C を頂点とする三角形を考える。 $M_1 := A\#B$ とすると、 M_1 は頂点 C の対辺 $[A, B]$ の中点となる。 よって、 (51) より、

$$\delta_2^2(M_1, C) \leq \frac{\delta_2^2(A, C) + \delta_2^2(B, C)}{2} - \frac{\delta_2^2(A, B)}{4}.$$

同様に $M_2 := A\#C$ とすると、 M_2 は三角形 AM_1C における頂点 M_1 の対辺 $[A, C]$ の中点となる。 よって、 (51) より、

$$\delta_2^2(M_1, M_2) \leq \frac{\delta_2^2(M_1, C) + \delta_2^2(M_1, A)}{2} - \frac{\delta_2^2(A, C)}{4}.$$

1 番目の不等式を 2 番目の不等式に代入すると、

$$\delta_2^2(M_1, M_2) \leq \frac{1}{4} \left(\delta_2^2(A, C) + \delta_2^2(B, C) \right) - \frac{1}{8} \delta_2^2(A, B) + \frac{1}{2} \delta_2^2(M_1, A) - \frac{1}{4} \delta_2^2(A, C).$$

$\delta_2(M_1, A) = \frac{1}{2} \delta_2(A, B)$ なので、上の不等式は、

$$\delta_2(M_1, M_2) \leq \frac{\delta_2(B, C)}{2}$$

となる。 ■

(52) の不等式は、もっと一般的な形の不等式の証明に使われる。 $0 \leq t \leq 1$ に対して、

$$A\#_t B := A^{1/2} \left(A^{-1/2} B A^{-1/2} \right)^t A^{1/2} \quad (53)$$

とおく。これは (45) の測地線 $\gamma(t)$ を表すもうひとつの記号である。 $t = 1/2$ のとき、これは幾何平均 $A\#B$ となる。さらに一般的な形は以下で述べる。

系 4.10 $B, C, B', C' \in \mathbb{P}_n$ とし、

$$f(t) := \delta_2(B'\#_t B, C'\#_t C) \quad (54)$$

とすると、この写像 f は $[0, 1]$ 上の凸関数となる。つまり、

$$\delta_2(B'\#_t B, C'\#_t C) \leq (1-t)\delta_2(B', C') + t\delta_2(B, C). \quad (55)$$

証明 f の連続性を最初に示す.

$$\begin{aligned}
|f(t) - f(t_0)| &= |\delta_2(B' \#_t B, C' \#_t C) - \delta_2(B' \#_{t_0} B, C' \#_{t_0} C)| \\
&\leq |\delta_2(B' \#_t B, C' \#_t C) - \delta_2(B' \#_{t_0} B, C' \#_t C)| \\
&\quad + |\delta_2(B' \#_{t_0} B, C' \#_t C) - \delta_2(B' \#_{t_0} B, C' \#_{t_0} C)| \\
&\leq \delta_2(B' \#_t B, B' \#_{t_0} B) + \delta_2(C' \#_t C, C' \#_{t_0} C) \\
&= |t - t_0| \delta_2(B', B) + |t - t_0| \delta_2(C', C) \\
&\rightarrow 0 \quad (t \rightarrow t_0).
\end{aligned}$$

よって f は連続なので, 中点凸関数であることを示せば十分.

$M_1 = B' \# B, M_2 = C' \# C, M = B' \# C$ とする. 命題 4.9 より,

$$\begin{aligned}
\delta_2(M_1, M) &= \delta_2(B' \# B, B' \# C) \leq \frac{\delta_2(B, C)}{2}, \\
\delta_2(M, M_2) &= \delta_2(B' \# C, C' \# C) \leq \frac{\delta_2(B', C')}{2}.
\end{aligned}$$

よって,

$$\begin{aligned}
\delta_2(M_1, M_2) &\leq \delta_2(M_1, M) + \delta_2(M, M_2) \\
&\leq \frac{1}{2} \{ \delta_2(B, C) + \delta_2(B', C') \}.
\end{aligned}$$

よって, f は中点凸関数である. ■

(55) において $B' = C' = A$ とすると, 距離 δ_2 の凸性と呼ばれる以下の定理を得る.

定理 4.11 A, B, C を \mathbb{P}_n の任意の 3 点とすると, 任意の $t \in [0, 1]$ に対して

$$\delta_2(A \#_t B, A \#_t C) \leq t \delta_2(B, C). \quad (56)$$

命題 4.12 固定された任意の点 $A \in \mathbb{P}_n$ に対して, 関数 f を $f(X) = \delta_2^2(A, X)$ によって定める. このとき, $X_1 \neq X_2$ ならば $0 < t < 1$ に対して

$$f(X_1 \#_t X_2) < (1-t)f(X_1) + tf(X_2). \quad (57)$$

証明 $\varphi(t) = f(X_1 \#_t X_2)$ とおくと, 帰納法より,

$$\varphi\left(\frac{k}{2^{n+1}}\right) < \left(1 - \frac{k}{2^n}\right)\varphi(0) + \frac{k}{2^n}\varphi\left(\frac{1}{2}\right) \quad (0 \leq k \leq 2^n).$$

$t \in (0, 1/2]$ のとき,

$$\begin{aligned}
\varphi(t) &\leq (1-2t)\varphi(0) + 2t\varphi\left(\frac{1}{2}\right) \\
&< (1-2t)\varphi(0) + 2t\frac{\varphi(0) + \varphi(1)}{2} \\
&= (1-t)\varphi(0) + t\varphi(1).
\end{aligned}$$

つまり, $f(X_1 \#_t X_2) < (1-t)f(X_1) + tf(X_2)$. $t \in [1/2, 1)$ のときも同様. ■

4.2 距離空間

補題 4.13 写像 e^H は $(\mathbb{H}_n, \|\cdot\|_2)$ から (\mathbb{P}_n, δ_2) への連続写像である。

証明 $\{H_m\}_{m=1}^\infty \subset \mathbb{H}_n, H_m \rightarrow H$ とすると,

$$e^{-H_m} e^H \rightarrow e^{-H} e^H = I.$$

よって, すべての固有値について

$$\lambda_i(e^{-H_m} e^H) \rightarrow 1 \quad (1 \leq i \leq n)$$

となるので, (48) より,

$$\delta_2(e^{H_m}, e^H) = \left(\sum_{i=1}^n \log^2 \lambda_i(e^{-H_m} e^H) \right)^{1/2} \rightarrow 0 \quad (m \rightarrow \infty). \quad \blacksquare$$

命題 4.14 (\mathbb{P}_n, δ_2) は完備な空間である。

証明 $\{A_m\}$ を (\mathbb{P}_n, δ_2) 内のコーシー列とし, $H_m := \log A_m$ とすると, (42) より,

$$\|H_m - H_l\|_2 \leq \delta_2(A_m, A_l) \rightarrow 0 \quad (m, l \rightarrow \infty).$$

よって, $\{H_m\}$ は $(\mathbb{H}_n, \|\cdot\|_2)$ 上のコーシー列. $(\mathbb{H}_n, \|\cdot\|_2)$ は完備距離空間である. よって, $H_m \rightarrow H \in \mathbb{H}_n$ となる H が存在する. 補題 4.13 より,

$$A_m \rightarrow A = e^H \in (\mathbb{P}_n, \delta_2). \quad \blacksquare$$

\mathbb{P}_n は正値行列全体の集合で, 非正則半正値行列全体をなす境界をもつ. \mathbb{P}_n 上の点と非正則行列の関係を次の命題で考えるが, 命題では (53) の $A \#_t B$ を使う. A と B が交換可能ならば, $A \#_t B = A^{1-t} B^t$.

命題 4.15 S を非正則な半正値行列とする. このとき,

$$\|A^{1-t} B^t - S\|_2 \rightarrow 0 \quad (t \rightarrow \infty),$$

$$\delta_2(A^{1-t} B^t, A) \rightarrow \infty \quad (t \rightarrow \infty).$$

となるような交換可能な $A, B \in \mathbb{P}_n$ が存在する.

証明 対角化した行列について考える. $S := \text{diag}(\lambda_1, \dots, \lambda_n)$ とする. ただし, $1 \leq k \leq n$ に対して $\lambda_k \geq 0$ であり, ある k に対しては $\lambda_k = 0$ とする. また, $\lambda_k > 0$ のとき, $\alpha_k = \beta_k := \lambda_k, \lambda_k = 0$ のとき, $\alpha_k := 1, \beta_k := 1/2$ とする. $A := \text{diag}(\alpha_1, \dots, \alpha_n), B := \text{diag}(\beta_1, \dots, \beta_n)$ とすると,

$$\lim_{t \rightarrow \infty} \|A^{1-t} B^t - S\|_2 = 0.$$

また, 距離 δ_2 に対して,

$$\begin{aligned}\delta_2(A^{1-t}B^t, A) &= \|\log A^{1-t}B^t - \log A\|_2 = \|\log A^{-t}B^t\|_2 \\ &= \left(\operatorname{tr}(\log A^{-t}B^t)^*(\log A^{-t}B^t)\right)^{1/2} \\ &= \left(\sum_{k=1}^n |\log \alpha_k^{-t} \beta_k^t|^2\right)^{1/2} \geq |\log 2^{-t}| = t \log 2 \rightarrow \infty \quad (t \rightarrow \infty). \quad \blacksquare\end{aligned}$$

この命題のポイントは, 曲線 $A \#_t B$ は $t=0$ のとき点 A からスタートし, $(\mathbb{H}_n, \|\cdot\|_2)$ では S へ収束するが, (\mathbb{P}_n, δ_2) では ∞ へ発散することである.

\mathbb{P}_n 上の 2 点 A, B が測地線によって結ばれることについて, ここまで見てきた. ここからは, 測地線を発展させた議論をする.

定義 4.16 $\mathcal{K} \subset \mathbb{P}_n$ とする. 任意の $A, B \in \mathcal{K}$ に対して $[A, B] \subset \mathcal{K}$ のとき, 集合 \mathcal{K} は凸であるという.

S が \mathbb{P}_n の部分集合のとき, S の凸包は S を含む最小の凸集合である. S の凸包を $\operatorname{conv}(S)$ で表す. すなわち, $\operatorname{conv}(S)$ は S を含む凸集合全体の共通部分である. $\{A, B\}$ の凸包は明らかに $[A, B]$.

命題 4.17 $S \subset \mathbb{P}_n$ とする. S_m を $S_0 := S$,

$$S_{m+1} := \cup\{[A, B] : A, B \in S_m\}$$

によって定める. このとき,

$$\operatorname{conv}(S) = \bigcup_{m=0}^{\infty} S_m.$$

証明 $S_m \subset \operatorname{conv}(S)$ は明らかなので, $\bigcup_{m=0}^{\infty} S_m \subset \operatorname{conv}(S)$. あとは, $\bigcup_{m=0}^{\infty} S_m$ が凸集合であることを示せばよい. $A \in [A, B]$ なので, $S_m \subset S_{m+1}$. $A, B \in \bigcup_{m=0}^{\infty} S_m$ とすると, $A, B \in S_{m'}$ となるような m' が存在する. よって,

$$[A, B] \subset S_{m'+1} \subset \bigcup_{m=0}^{\infty} S_m. \quad \blacksquare$$

命題 4.18 $A, B, C \in \mathbb{P}_n$ に対して, 以下の条件は同値.

- (1) $\Gamma_X(A), \Gamma_X(B), \Gamma_X(C)$ が交換可能となるような $X \in GL(n)$ が存在する,
- (2) $AB^{-1}C = CB^{-1}A$,
- (3) $A^{-1/2}BA^{-1/2}$ と $A^{-1/2}CA^{-1/2}$ が交換可能.

証明 (1) \Rightarrow (2). XAX^* , XBX^* , XCX^* は交換可能なので,

$$XAX^*(XBX^*)^{-1}XCX^* = XCX^*(XBX^*)^{-1}XAX^*.$$

つまり, $XAB^{-1}CX^* = XCB^{-1}AX^*$. よって $AB^{-1}C = CB^{-1}A$.

(2) \Rightarrow (3). 任意の $Y \in \mathbb{P}_n$ に対して,

$$\Gamma_Y(A)\Gamma_Y(B)^{-1}\Gamma_Y(C) = \Gamma_Y(C)\Gamma_Y(B)^{-1}\Gamma_Y(A).$$

$Y = A^{-1/2}$ とすると,

$$\Gamma_{A^{-1/2}}(B)^{-1}\Gamma_{A^{-1/2}}(C) = \Gamma_{A^{-1/2}}(C)\Gamma_{A^{-1/2}}(B)^{-1}.$$

(3) \Rightarrow (1). $X = A^{-1/2}$ とすればよい. ■

定理 4.19 \mathcal{K} を (\mathbb{P}_n, δ_2) に含まれる閉凸集合とする. このとき, 任意の $A \in \mathbb{P}_n$ に対して, $C \in \mathcal{K}$ が存在して

$$\delta_2(A, C) < \delta_2(A, K) \quad (K \in \mathcal{K}, K \neq C).$$

証明 $\mu := \inf\{\delta_2(A, K) : K \in \mathcal{K}\}$ とすると, $\delta_2(A, C_n) \rightarrow \mu$ となるような $\{C_n\} \subset \mathcal{K}$ が存在する. ある n と m に対して $M := C_n \# C_m$ とすると, \mathcal{K} は凸集合なので, $M \in \mathcal{K}$. 半中線定理 (51) より,

$$\delta_2^2(M, A) \leq \frac{\delta_2^2(C_n, A) + \delta_2^2(C_m, A)}{2} - \frac{\delta_2^2(C_n, C_m)}{4}.$$

これを变形して,

$$\begin{aligned} \delta_2^2(C_n, C_m) &\leq 2\{\delta_2^2(C_n, A) + \delta_2^2(C_m, A)\} - 4\delta_2^2(M, A) \\ &\leq 2\{\delta_2^2(C_n, A) + \delta_2^2(C_m, A)\} - 4\mu^2. \end{aligned} \quad (58)$$

$n, m \rightarrow \infty$ とすると, 右辺 $\rightarrow 0$. よって, $\{C_n\}$ はコーシー列. 命題 4.14 より (\mathbb{P}_n, δ_2) は完備なので, $C_n \rightarrow C \in \mathbb{P}_n$. \mathcal{K} は閉集合なので, $C \in \mathcal{K}$. さらに,

$$\delta_2(A, C) = \lim \delta_2(A, C_n) = \mu \leq \delta_2(A, K) \quad (K \in \mathcal{K}).$$

もし $K \in \mathcal{K}$ が $\delta_2(A, K) = \mu$ をみたすならば, (58) において $C_n = C$, $C_m = K$ とおけば, $\delta_2(C, K) = 0$ を得る. つまり, $C = K$. ■

この命題で与えられる $\pi(A) = C$ という写像は, \mathcal{K} の上への距離射影と呼ばれる.

定理 4.20 \mathcal{K} を (\mathbb{P}_n, δ_2) に含まれる閉凸集合とし, π を \mathcal{K} の上への距離射影とする. A が \mathbb{P}_n の点で, $\pi(A) = C$ ならば, 任意の $D \in \mathcal{K}$ に対して

$$\delta_2^2(A, D) \geq \delta_2^2(A, C) + \delta_2^2(C, D). \quad (59)$$

証明 $\{M_n\} \subset \mathcal{K}$, $M_0 := D$, $M_{n+1} := M_n \# C$ と定めると, $\delta_2(C, M_n) = 2^{-n} \delta_2(C, D)$ となり, $M_n \rightarrow C = M_\infty$ である. 半中線定理 (51) より,

$$2\delta_2^2(A, M_{n+1}) \leq \delta_2^2(A, M_n) + \delta_2^2(A, C) - \frac{1}{2}\delta_2^2(C, M_n).$$

変形すると,

$$\delta_2^2(A, M_n) - \delta_2^2(A, M_{n+1}) \geq \frac{1}{2 \cdot 4^n} \delta_2^2(C, D) + \delta_2^2(A, M_{n+1}) - \delta_2^2(A, C).$$

両辺を $n = 0$ から ∞ まで辺々加えると,

$$\sum_{n=0}^{\infty} \{\delta_2^2(A, M_n) - \delta_2^2(A, M_{n+1})\} \geq \frac{2}{3} \delta_2^2(C, D) + \sum_{n=0}^{\infty} \{\delta_2^2(A, M_{n+1}) - \delta_2^2(A, C)\}.$$

\mathcal{K} は凸集合なので, 任意の n に対して $M_n \in \mathcal{K}$. よって $\delta_2^2(A, M_n) - \delta_2^2(A, C) \geq 0$. 上の左辺の級数は明らかに収束する. よって右辺の級数も収束する.

$d_n := \delta_2^2(A, M_n) - \delta_2^2(A, C)$ とすると, 上の不等式は

$$d_0 = \delta_2^2(A, D) - \delta_2^2(A, C) \geq \frac{2}{3} \delta_2^2(C, D) + \sum_{n=1}^{\infty} d_n$$

となる. D に M_n を代入すると,

$$d_n = \delta_2^2(A, M_n) - \delta_2^2(A, C) \geq \frac{2}{3} \delta_2^2(C, M_n) + \sum_{k=n+1}^{\infty} d_k.$$

よって,

$$\begin{aligned} d_0 &\geq \frac{2}{3} \delta_2^2(C, D) + d_1 + \sum_{k=2}^{\infty} d_k \\ &\geq \frac{2}{3} \delta_2^2(C, D) + \frac{2}{3} \delta_2^2(C, M_1) + 2 \sum_{k=2}^{\infty} d_k \\ &= \frac{2}{3} \delta_2^2(C, D) + \frac{2}{3} \cdot \frac{1}{4} \delta_2^2(C, D) + 2d_2 + 2 \sum_{k=3}^{\infty} d_k \\ &= \frac{2}{3} \left(1 + \frac{1}{4}\right) \delta_2^2(C, D) + 2d_2 + 2 \sum_{k=3}^{\infty} d_k \\ &\geq \frac{2}{3} \left(1 + \frac{1}{4}\right) \delta_2^2(C, D) + 2 \left(\frac{2}{3} \delta_2^2(C, M_2) + \sum_{k=3}^{\infty} d_k \right) + 2 \sum_{k=3}^{\infty} d_k \\ &= \frac{2}{3} \left(1 + \frac{1}{4} + \frac{2}{4^2}\right) \delta_2^2(C, D) + 4 \sum_{k=3}^{\infty} d_k \\ &\geq \dots \end{aligned}$$

$d_n \geq 0$ であるから,

$$d_0 \geq \frac{2}{3} \left(1 + \sum_{n=1}^{\infty} \frac{2^{n-1}}{4^n}\right) \delta_2^2(C, D) = \delta_2^2(C, D). \quad \blacksquare$$

命題 4.21 \mathcal{E} を距離 d のユークリッド空間, $a, b \in \mathcal{E}$ とすると,

$$f(x) = d^2(a, x) + d^2(b, x)$$

によって定義される関数 f は, ただひとつの点 $x_0 = \frac{1}{2}(a + b)$ で \mathcal{E} 上の最小値をとる.

(\mathbb{P}_n, δ_2) においても, 幾何平均によって同様の関係が見られる.

命題 4.22 $A, B \in \mathbb{P}_n$ とし,

$$f(X) := \delta_2^2(A, X) + \delta_2^2(B, X)$$

とすると, 関数 f は \mathbb{P}_n の狭義凸関数で, ただひとつの $X_0 = A \# B$ で最小値をとる.

証明 狭義凸性は命題 4.12 で示した. 半中線定理を任意の X について使うと,

$$\begin{aligned} \delta_2^2(A \# B, X) &\leq \frac{1}{2}f(X) - \frac{1}{4}\delta_2^2(A, B) \\ &= \frac{1}{2}f(X) - \frac{1}{2}f(A \# B). \end{aligned}$$

よって, $f(A \# B) \leq f(X) - 2\delta_2^2(A \# B, X)$. これは, f が $X_0 := A \# B$ で最小値をとることを示している. ■

4.3 幾何平均

前章では, 2つの行列の幾何平均の定義と, それに関連する問題について述べた. この節では, 3つ以上の行列の幾何平均の適切な定義を与え, それに関連した問題について考える. 行列が3つの場合の考え方は, 行列が4つ以上の場合にも使える.

$A_1, A_2, A_3 \geq O$ とする. これらの幾何平均 $G(A_1, A_2, A_3)$ は以下の性質をみたすことが望ましく, このとき $G(A_1, A_2, A_3)$ は半正値行列となる.

- (1) $\{1, 2, 3\}$ の任意の置換 π に対して, $G(A_1, A_2, A_3) = G(A_{\pi(1)}, A_{\pi(2)}, A_{\pi(3)})$.
- (2) $A_1 \leq A'_1$ のとき, $G(A_1, A_2, A_3) \leq G(A'_1, A_2, A_3)$.
- (3) 任意の $X \in GL(n)$ に対して, $G(X^*A_1X, X^*A_2X, X^*A_3X) = X^*G(A_1, A_2, A_3)X$.
- (4) G は連続写像.

(1) は対称性, (2) は単調性, (3) は合同不変性と呼ばれる.

2つの正値行列の幾何平均の定義を考えるときに使うアイデアのうち, 3つの幾何平均へ拡張できるものはない. しかしここで, ユークリッド空間 (\mathcal{E}, d) 上の3点 x_1, x_2, x_3 の算術平均 $\bar{x} := \frac{1}{3}(x_1 + x_2 + x_3)$ を考えてみる. この点 \bar{x} は, 以下の性質によって特徴付けられる.

- (1) \bar{x} は, $\Delta(x_1, x_2, x_3)$ の3本の中線の唯一の交点である. この点を, 三角形の重心と呼ぶ.

- (2) \bar{x} は, $d^2(x, x_1) + d^2(x, x_2) + d^2(x, x_3)$ の最小値をとるような唯一の点である.
- (3) \bar{x} は, $\Delta_1 := \Delta(x_1, x_2, x_3)$ とし, Δ_j の 3 辺の各中点を頂点とする三角形を Δ_{j+1} とするとき, 列 $\{\Delta_n\}$ の唯一の共通部分.

これらの構造を (\mathbb{P}_n, δ_2) で考えるときには, 注意が必要である. まず, 三角形の形状である. 前節では凸包を定義して, 2 点 $A_1, A_2 \in \mathbb{P}_n$ の凸包が測地線 $[A_1, A_2]$ であることを見た. 3 点 A_1, A_2, A_3 の凸包はやや複雑である. 命題 4.17 の記号を用いると, $S = \{A_1, A_2, A_3\}$ のとき, $S_1 = [A_1, A_2] \cup [A_2, A_3] \cup [A_3, A_1]$. しかし S_2 は膨らんだ物体のようになるので, $[A_1, A_2 \# A_3]$, $[A_2, A_1 \# A_3]$, $[A_3, A_1 \# A_2]$ がまったく交わらないこともある. よって, (1) を $\Delta(A_1, A_2, A_3)$ の重心の定義とするのは適当ではない.

次に考えるのは, 任意の $A_1, A_2, A_3 \in \mathbb{P}_n$ に対して, $f(X) = \sum_{j=1}^3 \delta_2^2(A_j, X)$ が最小値をとるような唯一の X_0 が存在するかどうかである. 半中線定理を用いて, そのような点の存在を見る.

$m := \inf f(X)$ とし, \mathbb{P}_n の列 $\{X_r\}$ を $f(X_r) \rightarrow m$ となるようにとる. 半中線定理より, $j = 1, 2, 3$ と任意の r, s に対して,

$$\delta_2^2(X_r \# X_s, A_j) \leq \frac{\delta_2^2(X_r, A_j) + \delta_2^2(X_s, A_j)}{2} - \frac{\delta_2^2(X_r, X_s)}{4}.$$

$j = 1, 2, 3$ に対して辺々加えると,

$$f(X_r \# X_s) \leq \frac{1}{2}(f(X_r) + f(X_s)) - \frac{3}{4}\delta_2^2(X_r, X_s).$$

これより,

$$\begin{aligned} \frac{3}{4}\delta_2^2(X_r, X_s) &\leq \frac{1}{2}(f(X_r) + f(X_s)) - f(X_r \# X_s) \\ &\leq \frac{1}{2}(f(X_r) + f(X_s)) - m. \end{aligned}$$

よって $\{X_r\}$ はコーシー列で, $X_r \rightarrow X_0 \in \mathbb{P}_n$. 明らかに, f は X_0 で最小値をとる. 命題 4.12 より f は狭義凸関数であるから, 唯一の点 X_0 で最小値をとる.

$f(X)$ の最小値をとるような X_0 を $X_0 := \operatorname{argmin} f(X)$ とおき, $\{A_1, A_2, A_3\}$ の重心の定義を以下のように定める.

$$G(A_1, A_2, A_3) := \operatorname{argmin} \sum_{j=1}^3 \delta_2^2(A_j, X). \quad (60)$$

$G(A_1, A_2, A_3)$ が対象性をもち, 3 変数の連続写像であることは定義より明らか. 合同変形 Γ_X は (\mathbb{P}_n, δ_2) の等距離写像なので, G が合同不変であることも明らか.

よって, G はこの節のはじめに述べた, 幾何平均の望ましい 4 つの性質のうち, 3 つをみたすことがわかった. しかし, 単調性についてはまだわかっていない. G の他の性質を以下で述べる.

G の性質を見る前に, 方向微分に関する命題等をいくつか確認する.

定義 4.23 X, Y をバナッハ空間とする. 写像 $f: X \rightarrow Y$ が点 u において方向微分可能とは, 任意の $v \in X$ に対して

$$Df(u)(v) := \lim_{t \rightarrow 0} \frac{f(u+tv) - f(u)}{t}$$

が存在することをいう. このとき, $Df(u)$ は X から Y への有界線形写像となる.

命題 4.24 X, Y_1, Y_2, Z を実バナッハ空間とし, $f: X \rightarrow Y_1, g: X \rightarrow Y_2$ は $u \in X$ で方向微分可能とする. $B: Y_1 \times Y_2 \rightarrow Z$ を有界双線形写像とし, $\varphi(x) = B(f(x), g(x))$ と定める. このとき, $\varphi(x) = B(f(x), g(x))$ で定義される $\varphi: X \rightarrow Z$ について, 任意の $u, v \in X$ に対して,

$$D\varphi(u)(v) = B(Df(u)(v), g(u)) + B(f(u), Dg(u)(v)).$$

証明 仮定より

$$\begin{aligned} & \frac{\varphi(u+tv) - \varphi(u)}{t} \\ &= \frac{B(f(u+tv), g(u+tv)) - B(f(u), g(u))}{t} \\ &= \frac{B(f(u+tv), g(u+tv)) - B(f(u), g(u+tv))}{t} + \frac{B(f(u), g(u+tv)) - B(f(u), g(u))}{t} \\ &= B\left(\frac{f(u+tv) - f(u)}{t}, g(u+tv)\right) - B\left(f(u), \frac{g(u+tv) - g(u)}{t}\right) \\ &\rightarrow B(Df(u)(v), g(u)) + B(f(u), Dg(u)(v)) \quad (t \rightarrow 0). \quad \blacksquare \end{aligned}$$

補題 4.25 φ_1, φ_2 を $(0, \infty)$ 上の連続かつ微分可能な実数値関数とし, 任意の $X \in \mathbb{P}_n$ に対して

$$h(X) := \langle \varphi_1(X), \varphi_2(X) \rangle = \text{tr} \{ \varphi_1(X) \varphi_2(X) \}$$

とする. このとき, h の微分は

$$Dh(X)(Y) = \langle \varphi_1'(X) \varphi_2(X) + \varphi_1(X) \varphi_2'(X), Y \rangle$$

によって与えられる.

証明 命題 4.24 より,

$$Dh(X)(Y) = \langle D\varphi_1(X)(Y), \varphi_2(X) \rangle + \langle \varphi_1(X), D\varphi_2(X)(Y) \rangle.$$

正規直交基底を選んで $X = \text{diag}(\lambda_1, \dots, \lambda_n)$ とすると, (3) より,

$$D\varphi_1(X)(Y) = \left[\frac{\varphi_1(\lambda_i) - \varphi_1(\lambda_j)}{\lambda_i - \lambda_j} \right] \circ Y.$$

よって,

$$\begin{aligned}\langle D\varphi_1(X)(Y), \varphi_2(X) \rangle &= \sum_i \varphi_1'(\lambda_i) y_{ii} \varphi_2(\lambda_i) \\ &= \langle \varphi_1'(X) \varphi_2(X), Y \rangle.\end{aligned}$$

同様に,

$$\langle \varphi_1(X), D\varphi_2(X)(Y) \rangle = \langle \varphi_1(X) \varphi_2'(X), Y \rangle. \blacksquare$$

系 4.26 $h(X) := \|\log X\|_2^2$, $X \in \mathbb{P}_n$ とすると, 任意の $Y \in \mathbb{H}_n$ に対して

$$Dh(X)(Y) = 2\langle X^{-1} \log X, Y \rangle.$$

証明 $h(X) = \|\log X\|_2^2 = \langle \log X, \log X \rangle$ なので, $\varphi_1 = \varphi_2 = \log$ として, 補題 4.25 を使うと,

$$\begin{aligned}Dh(X)(Y) &= \langle X^{-1}(\log X) + (\log X)X^{-1}, Y \rangle \\ &= 2\langle X^{-1} \log X, Y \rangle. \blacksquare\end{aligned}$$

系 4.27 $h(X) := \|\log(A^{-1/2}XA^{-1/2})\|_2^2$, $X \in \mathbb{P}_n$ とすると, 任意の $Y \in \mathbb{H}_n$ に対して

$$Dh(X)(Y) = 2\langle (A^{-1/2}XA^{-1/2})^{-1} \log(A^{-1/2}XA^{-1/2}), A^{-1/2}YA^{-1/2} \rangle. \quad (61)$$

証明 上の補題と同様に考えて, $\varphi = \log t$ とし, $\varphi_1(X) = \varphi_2(X) = \log(A^{-1/2}XA^{-1/2}) = \varphi(A^{-1/2}XA^{-1/2})$ とする. $D\varphi_1(X)(Y) = D\varphi(A^{-1/2}XA^{-1/2})(A^{-1/2}YA^{-1/2})$ より,

$$\begin{aligned}\langle D\varphi_1(X)(Y), \varphi_2(X) \rangle &= \langle \varphi'(A^{-1/2}XA^{-1/2}) \varphi_2(X), A^{-1/2}YA^{-1/2} \rangle \\ &= \langle (A^{-1/2}XA^{-1/2})^{-1} \log(A^{-1/2}XA^{-1/2}), A^{-1/2}YA^{-1/2} \rangle.\end{aligned}$$

同様に,

$$\langle \varphi_1(X), D\varphi_2(X)(Y) \rangle = \langle \log(A^{-1/2}XA^{-1/2}) \cdot (A^{-1/2}XA^{-1/2})^{-1}, A^{-1/2}YA^{-1/2} \rangle. \blacksquare$$

以上の事実を用いて, G に関する性質を見る.

定理 4.28 A_1, A_2, A_3 を \mathbb{P}_n の任意の 3 点とし,

$$f(X) := \sum_{j=1}^3 \delta_2^2(A_j, X) \quad (62)$$

とする. このとき, X における f の微分は, 任意の $Y \in \mathbb{H}_n$ に対して

$$Df(X)(Y) = 2 \sum_{j=1}^3 \langle X^{-1} \log(XA_j^{-1}), Y \rangle \quad (63)$$

で与えられる.

証明 (47) より,

$$f(X) = \sum_{j=1}^3 \|\log(A_j^{-1/2} X A_j^{-1/2})\|_2^2.$$

ここで (61) を用いると,

$$\begin{aligned} Df(X)(Y) &= \sum_{j=1}^3 2 \left\langle (A_j^{-1/2} X A_j^{-1/2})^{-1} \log(A_j^{-1/2} X A_j^{-1/2}), A_j^{-1/2} Y A_j^{-1/2} \right\rangle \\ &= \sum_{j=1}^3 2 \operatorname{tr} \left\{ A_j^{1/2} X^{-1} A_j^{1/2} \log(A_j^{-1/2} X A_j^{-1/2}) A_j^{-1/2} Y A_j^{-1/2} \right\} \\ &= \sum_{j=1}^3 2 \operatorname{tr} \left\{ X^{-1} A_j^{1/2} \log(A_j^{-1/2} X A_j^{-1/2}) A_j^{-1/2} Y \right\} \\ &= \sum_{j=1}^3 2 \operatorname{tr} \{ X^{-1} \log(X A_j^{-1}) Y \}. \end{aligned}$$

上で $A_j^{1/2} \log(A_j^{-1/2} X A_j^{-1/2}) A_j^{-1/2} = \log(X A_j^{-1})$ を用いた. ■

定理 4.29 A_1, A_2, A_3 を任意の正値行列とし,

$$X_0 := G(A_1, A_2, A_3) = \operatorname{argmin} \sum_{j=1}^3 \delta_2^2(A_j, X)$$

とする. このとき, X_0 は

$$\sum_{j=1}^3 X_0^{-1} \log(X_0 A_j^{-1}) = O \tag{64}$$

の唯一の正値な解である.

証明 仮定より, X_0 は (62) に対して $Df(X_0)(Y) = 0$ ($Y \in \mathbb{H}_n$) をみたす唯一の点である. (63) を用いて,

$$Df(X_0)(Y) = 2 \sum_{j=1}^3 \langle X_0^{-1} \log(X_0 A_j^{-1}), Y \rangle = 0.$$

Y は \mathbb{H}_n の任意の点なので, これは

$$\sum_{j=1}^3 X_0^{-1} \log(X_0 A_j^{-1}) = O. \quad \blacksquare$$

と同値.

命題 4.30 A_1, A_2, A_3 が交換可能のとき, $G(A_1, A_2, A_3) = (A_1 A_2 A_3)^{1/3}$.

証明 (42) より.

$$\sum_{j=1}^3 \delta_2^2(A_j, X) \geq \sum_{j=1}^3 \|\log A_j - \log X\|_2.$$

右辺が最小となるのは, $\log X = \frac{1}{3} \sum_i^3 \log A_i$ のときである. A_1, A_2, A_3 は交換可能なので, $X = (A_1 A_2 A_3)^{1/3}$. また, この X は A_1, A_2, A_3 と交換可能なので,

$$\sum_{j=1}^3 \delta_2^2(A_j, X) = \sum_{j=1}^3 \|\log A_j - \log X\|_2.$$

よって $X = (A_1 A_2 A_3)^{1/3}$ は $\sum_{j=1}^3 \delta_2^2(A_j, X)$ を最小にする. ■

(60) までの議論と同様に次がわかる.

命題 4.31 $w := (w_1, w_2, w_3)$, $w_j \geq 0$, $\sum_{j=1}^3 w_j = 1$ とし,

$$f_w(X) := \sum_{j=1}^3 w_j \delta_2^2(A_j, X)$$

とすると, f_w は狭義凸関数で, $X_0 = \operatorname{arccmin} f_w(X)$ となる唯一の X_0 が存在する.

定義 4.32 f_w に最小値を与える点を $G_w(A_1, A_2, A_3)$ とおく. つまり,

$$G_w(A_1, A_2, A_3) := \operatorname{arccmin} f_w(X).$$

$w = (\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$ とすれば, $G_w(A_1, A_2, A_3) = G(A_1, A_2, A_3)$ となる.

命題 4.33 各 $G_w(A_1, A_2, A_3)$ は $\operatorname{conv}(\{A_1, A_2, A_3\})$ の閉包に含まれる.

証明 \mathcal{K} を $\operatorname{conv}(\{A_1, A_2, A_3\})$ の閉包, π を \mathcal{K} の上への距離射影とすると, 定理 4.20 より, 任意の $X \in \mathbb{P}_n$ に対して,

$$\delta_2^2(A_j, X) \geq \delta_2^2(A_j, \pi(X)) + \delta_2^2(\pi(X), X) \geq \delta_2^2(A_j, \pi(X)).$$

よって,

$$f_w(X) \geq f_w(\pi(X)).$$

$\pi(X) \in \mathcal{K}$ より, $f_w(X)$ の最小値は \mathcal{K} の外の点で与えられることはない. ■

次に, この章の最初に見た算術平均の特徴の (3) を参考に, 3 つの行列の幾何平均を考えてみる.

$A_1, A_2, A_3 \in \mathbb{P}_n$ に対して, $A_1^{(0)} := A_1, A_2^{(0)} := A_2, A_3^{(0)} := A_3$,

$$A_1^{(m+1)} := A_1^{(m)} \# A_2^{(m)}, A_2^{(m+1)} := A_2^{(m)} \# A_3^{(m)}, A_3^{(m+1)} := A_3^{(m)} \# A_1^{(m)} \quad (65)$$

として, 三角形 $\{A_1^{(m)}, A_2^{(m)}, A_3^{(m)}\}$ を考える.

定理 4.34 $A_1, A_2, A_3 \in \mathbb{P}_n$ に対して $\{A_1^{(m)}, A_2^{(m)}, A_3^{(m)}\}$ を (65) のように定めると, $\text{conv}(\{A_1^{(m)}, A_2^{(m)}, A_3^{(m)}\})$ からどんな点 X_m を取っても, $\{X_m\}$ はある点 $X \in \text{conv}(\{A_1, A_2, A_3\})$ に収束する. X は X_m の取り方によらない.

証明 \mathbb{P}_n の集合 \mathcal{S} の直径を以下のように定義する.

$$\text{diam } \mathcal{S} := \sup\{\delta_2(X, Y) \mid X, Y \in \mathcal{S}\}.$$

$\text{diam} = M$ のとき, δ_2 の凸性から, $\text{diam}(\text{conv}(\mathcal{S})) = M$.

$\mathcal{K}_m := \text{conv}(\{A_1^{(m)}, A_2^{(m)}, A_3^{(m)}\})$ とする. (52) と上で述べたことから,

$$\text{diam } \mathcal{K}_m \leq 2^{-m} M_0 \quad (M_0 := \text{diam}\{A_1, A_2, A_3\}).$$

$\{\mathcal{K}_m\}$ は減少列. よって, $\{X_m\}$ はコーシー列で, ある点 X に収束する. 任意の X_m は \mathcal{K}_0 の点なので, X は \mathcal{K}_0 の閉包に含まれる.

別に $Y_m \in \mathcal{K}_m$ に対して $Y_m \rightarrow Y$ (Y_m は $\text{conv}(\{A_1^{(m)}, A_2^{(m)}, A_3^{(m)}\})$ の点) とする.

$$\delta_2(X_m, Y_m) \leq \text{diam}(\mathcal{K}_m) \rightarrow 0 \quad (m \rightarrow \infty)$$

であるから,

$$\delta_2(X, Y) \leq \delta_2(X, X_m) + \delta_2(X_m, Y_m) + \delta_2(Y_m, Y) \rightarrow 0 \quad (m \rightarrow \infty).$$

よって, $X = Y$ となり, X の唯一性が示された. ■

定理 4.34 において,

$$G^\#(A_1, A_2, A_3) := X \left(= \lim_{m \rightarrow \infty} X_m \right)$$

とする. これは安藤-Li-Mathias が導入した 3 個の正値行列の幾何平均である. その構成から, $G^\#$ は明らかに対称性と連続性を持つ. また, 幾何平均 $A \# B$ は A と B に関して単調性を持ち,

$$\Gamma_X(A \# B) = \Gamma_X(A) \# \Gamma_X(B)$$

より, 単調性と合同不変性が導かれる.

$G^\#(A_1, A_2, A_3)$ と $G(A_1, A_2, A_3)$ は必ずしも一致しないことが知られている. $G^\#$ は 4 つのすべての性質をみたすが, G が単調性を持つかどうかは未解決である. しかし, A_1, A_2, A_3 が交換可能のときは次が成立する.

命題 4.35 A_1, A_2, A_3 が交換可能のとき, $G^\#(A_1, A_2, A_3) = (A_1 A_2 A_3)^{1/3}$.

証明 まず, A_1, A_2, A_3 が交換可能のとき, $\text{conv}(\{A_1, A_2, A_3\}) = \{A_1^{r_1} A_2^{r_2} A_3^{r_3} \mid r_1, r_2, r_3 \geq 0, r_1 + r_2 + r_3 = 1\}$ であることに注意する. $A_1^{(m)} := A_1^{r_1^{(m)}} A_2^{r_2^{(m)}} A_3^{r_3^{(m)}}$, $A_2^{(m)} := A_1^{s_1^{(m)}} A_2^{s_2^{(m)}} A_3^{s_3^{(m)}}$, $A_3^{(m)} :=$

$A_1^{t_1^{(m)}} A_2^{t_2^{(m)}} A_3^{t_3^{(m)}}$ とすると, $(r_1^{(1)}, r_2^{(1)}, r_3^{(1)}) = (\frac{1}{2}, \frac{1}{2}, 0)$, $(s_1^{(1)}, s_2^{(1)}, s_3^{(1)}) = (0, \frac{1}{2}, \frac{1}{2})$, $(t_1^{(1)}, t_2^{(1)}, t_3^{(1)}) = (\frac{1}{2}, 0, \frac{1}{2})$, となる. そして, 帰納法より, $j = 1, 2, 3$ に対して,

$$|r_j^{(m)} - s_j^{(m)}| \leq \frac{1}{2^m}, |s_j^{(m)} - t_j^{(m)}| \leq \frac{1}{2^m}, |t_j^{(m)} - r_j^{(m)}| \leq \frac{1}{2^m}$$

となることがわかる. さらに

$$r_1^{(m)} + r_2^{(m)} + r_3^{(m)} = s_1^{(m)} + s_2^{(m)} + s_3^{(m)} = t_1^{(m)} + t_2^{(m)} + t_3^{(m)}$$

であるから,

$$r_j^{(m)} \rightarrow \frac{1}{3}, s_j^{(m)} \rightarrow \frac{1}{3}, t_j^{(m)} \rightarrow \frac{1}{3} \quad (m \rightarrow \infty).$$

よって, $G^\#(A_1, A_2, A_3) = (A_1 A_2 A_3)^{1/3}$ となる. ■

4.4 ユニタリ不変ノルムに対する不等式

4.1節で示した $\|\cdot\|$ を用いた不等式のうちのいくつかは, ユニタリ不変ノルム $\|\|\cdot\|\|$ への不等式としても同様に成立する. この節ではそれらについてまとめる.

二つの結果をここで使う. 任意の $\lambda_1, \dots, \lambda_n \geq 0$ に対して,

$$\left[\frac{\lambda_i - \lambda_j}{\sinh(\lambda_i - \lambda_j)} \right] \geq 0.$$

もうひとつは命題 1.6 である. これらを用いて, 命題 (4.2) のユニタリ不変ノルム版を示す.

命題 4.36 (一般化された IEMI) 任意の $H, K \in \mathbb{R}_n$ と任意のユニタリ不変ノルムに対して,

$$\|\|e^{-H/2} D e^H(K) e^{-H/2}\|\| \geq \|\|K\|\|. \quad (66)$$

証明 公式 (3) より,

$$\begin{aligned} e^{-H/2} D e^H(K) e^{-H/2} &= \left[\frac{\sinh \frac{\lambda_i - \lambda_j}{2}}{\frac{\lambda_i - \lambda_j}{2}} k_{ij} \right] \\ &= \left[\frac{\sinh \frac{\lambda_i - \lambda_j}{2}}{\frac{\lambda_i - \lambda_j}{2}} \right] \circ K. \end{aligned}$$

上式においてユニタリ不変ノルムを考えると,

$$\begin{aligned} \|\|K\|\| &= \|\| \left[\frac{\frac{\lambda_i - \lambda_j}{2}}{\sinh \frac{\lambda_i - \lambda_j}{2}} \right] \circ e^{-H/2} D e^H(K) e^{-H/2} \|\| \\ &\leq \max_i \frac{2}{e^{\lambda_i} + e^{-\lambda_i}} \|\|e^{-H/2} D e^H(K) e^{-H/2}\|\| \\ &\leq \|\|e^{-H/2} D e^H(K) e^{-H/2}\|\|. \quad \blacksquare \end{aligned}$$

(36)において, $\|\cdot\|$ をユニタリ不変ノルムに代えると,

$$L_{\|\cdot\|}(A, B) = \int_a^b \|\gamma^{-1/2}(t)\gamma'(t)\gamma^{-1/2}(t)\| dt. \quad (67)$$

$\|X\|$ は X の特異値の対称関数なので, 補題 4.1 は $L_{\|\cdot\|}$ にも用いることができる. また, (38) をユニタリ不変ノルムに代えた

$$\delta_{\|\cdot\|}(A, B) := \inf\{L_{\|\cdot\|}(\gamma) \mid \gamma \text{ は } A \text{ から } B \text{ への路}\} \quad (68)$$

は, \mathbb{P}_n の距離となる. 一般化された IEMI から, 一般化された EMI を導くことができる.

定理 4.37 (一般化された EMI) 任意の $A, B \in \mathbb{P}_n$ に対して

$$\delta_{\|\cdot\|}(A, B) \geq \| \log A - \log B \|. \quad (69)$$

あるいは, 任意の $H, K \in \mathbb{H}_n$ に対して

$$\delta_{\|\cdot\|}(e^H, e^K) \geq \|H - K\|. \quad (70)$$

測地線の唯一性に考えるときには注意が必要となる. \mathbb{H}_n 上でユニタリ不変ノルムが導く距離において, 2点間の線分は唯一の測地線となるという性質を, 多くのユニタリ不変ノルムは持つ. そしてユニタリ不変ノルムがこの性質を持っているとき, \mathbb{P}_n の距離 $\delta_{\|\cdot\|}$ はその性質を引き継ぐことになる. 命題 4.5 や定理 4.6 で得た結果は, ユニタリ不変ノルムを用いて示すことができる. $A, B \in \mathbb{P}_n$ が交換可能のとき, 命題 4.5 より,

$$L_{\|\cdot\|}(A, B) = \| \log A - \log B \|$$

を得て, これを用いて, 定理 4.6 より,

$$\delta_{\|\cdot\|}(A, B) = \| \log A^{-1/2} B A^{-1/2} \|. \quad (71)$$

を得る.

幾何的な不等式と, 物理に関する不等式には関係性がある. H と K をエルミート行列とすると, (70) と (71) より,

$$\|H + K\| \leq \| \log(e^{H/2} e^K e^{H/2}) \|. \quad (72)$$

写像 e^x は \mathbb{R} 上で単調増加凸関数なので, (72) より,

$$\|e^{H+K}\| \leq \|e^{H/2} e^K e^{H/2}\|. \quad (73)$$

以下の2つの不等式は, (73) のノルムを $\|\cdot\|_1$ と $\|\cdot\|$ にしたもので, 物理学でよく知られている. λ_1 を最も大きい固有値とすると,

$$\text{tr } e^{H+K} \leq \text{tr } e^H e^K, \quad (74)$$

$$\lambda_1(e^{H+K}) \leq \lambda_1(e^H e^K). \quad (75)$$

(74) は Golden-Thompson 不等式, (75) は Segal の不等式と呼ばれている.

参考文献

- [1] R. A. Horn and C. R. Johnson, *Matrix Analysis*, Cambridge Univ. Press, 1985.
- [2] R. Bhatia, *Matrix Analysis*, Springer, 1997.
- [3] R. Bhatia, *Positive Definite Matrices*, Princeton Univ. Press, Princeton and Oxford, 2007.
- [4] R. Bhatia and C. Davis, A Cauchy-Schwarz inequality for operators with applications, *Linear Algebra Appl.*, **223/224** (1995), 119-129.
- [5] R. Bhatia and J. A. R. Holbrook, Riemannian geometry and matrix geometric means, *Linear Algebra Appl.*, **413** (2006), 594-618.