# 同ランク半正定値行列の平均について

**Means for fixed rank PSD matrices**

藤井　淳一（大阪教育大学）
Jun Ichi Fujii (Osaka Kyoiku Univ.)

## 1 Introduction.

よく知られているように、可逆性作用素、特に正定値行列の場合、リーマン（フィンスラー）幾何学の測地線として久保-安藤平均 [18] の意味での**幾何平均**の path

$$A\#_t, B = A^{\frac{1}{2}}\left(A^{-\frac{1}{2}}BA^{-\frac{1}{2}}\right)^t A^{\frac{1}{2}}$$

および、**相対作用素エントロピー** [10]

$$\mathsf{S}(A|B) = A^{\frac{1}{2}}\log\left(A^{-\frac{1}{2}}BA^{-\frac{1}{2}}\right)A^{\frac{1}{2}}$$

がその初期速度ベクトルとして位置づけられた [13, 14, 11, 7, 4, 17]。半正定値の一般平均 $\mathsf{m}$ の場合には、久保-安藤平均では、その単調性と上反連続性から、

$$A\,\mathsf{m}\,B = \operatorname*{s-lim}_{\varepsilon\searrow 0}(A+\varepsilon)\,\mathsf{m}(B+\varepsilon)$$

とするのが常であった。しかしこれはもはや純粋な多様体の幾何学の範疇からは逸脱し、特に多変数化において様々な支障をきたすことは、[17] で示した。特に、相対作用素エントロピーは、発散することのないように条件付けが必要になってくる。「可逆性」は通常の幾何学的な考察において重要なファクターなのである。

しかし、ある偶然から幾何学的に半正定値な行列の幾何平均の話を Bonnabel-Sepulchure [6] が導入していることを知った。これは、同ランクに限って、Grassman 多様体の測地線で「正定値に変換して」平均を導入するものであった。関連で調べているうちに、同ランクの射影について Batzies-Hüper-Machado-Leite [5] も幾何平均の path を、明示的に表現していることを知った。当人同士は知らないようだが、実は前者の射影版が後者になっていることがわかった。さらに、この方法で、一般の久保-安藤平均 $\mathsf{m}$ を、$\tilde{\mathsf{m}}$ に広げられることもわかった。[6] で指摘されているように（彼らは幾何平均限定)、正定値の場合には久保-安藤平均と一致するが、半正定値の時は一般には二者は異なり、しかも、相対作用素エントロピーは常に存在することになる。ここでは、この modify された平均 $\tilde{\mathsf{m}}$ について論じたいと思う。ただし話を単純にするため実行列に話を限るが、複素化はそれほど難しくない（無限次元化は多少の注意を有する）と思われるので、容易に拡張できる形で述べたい。

## 2 幾何学的背景

まず $p \leqq n$ についての、 **Stiefel 多様体** $\mathscr{V}_{n,p}$ に注目する：

$$\mathscr{V}_{n,p} = \{V:\ n \times p \mid V^*V = I_p\} = \mathcal{O}(n)/\mathcal{O}(n-p).$$

[8] のように、$\mathcal{O}(n-p)$ を $\mathcal{O}(n)$ の中に $I_p \oplus \mathcal{O}(n-p)$ として埋め込むと、商空間の定義とつながる：

$$\left\{ Q \begin{pmatrix} I_p & O \\ O & Q_{n-p} \end{pmatrix} \middle| \begin{matrix} Q \in \mathcal{O}(n) \\ Q_{n-p} \in \mathcal{O}(n-p) \end{matrix} \right\} = \begin{pmatrix} Q_1 & * \\ Q_2 & * \end{pmatrix} \cong \left\{ V = \begin{pmatrix} Q_1 \\ Q_2 \end{pmatrix} :\ n \times p \middle| V^*V = I_p \right\}.$$

ただし行列のブロック分解は $Q = \begin{pmatrix} Q_1 & Q_3 \\ Q_2 & Q_4 \end{pmatrix}$ とする。つまり、$I_n$ のファイバーは、 $I_p \oplus \mathcal{O}(n-p)$ と同型である。直交群 $\mathcal{O}(n)$ の Lie 環は skew-hermitian 全体 $\mathfrak{so}(n)$ で、これも $I$ での接空間である。ファイバー部分の接空間を垂直空間 $O \oplus \mathfrak{so}(n-p)$ として幾何学的な「接続」が考えられるので、接ベクトル $\Delta \in \mathfrak{so}(n)$ は水平空間に属し、$\Delta$ は

$$\Delta = \begin{pmatrix} A & -B^* \\ B & C \end{pmatrix} - \begin{pmatrix} O & O \\ O & C \end{pmatrix} = \begin{pmatrix} A & -B^* \\ B & O \end{pmatrix}$$

($A \in \mathfrak{so}(p), C \in \mathfrak{so}(n-p)$) とみなせる。したがって、$\mathscr{V}_{n,p}$ の初期接ベクトル $\Delta$、初期値 $Y$ の測地線は、微分方程式 $\ddot{Y} = -Y\dot{Y}^*\dot{Y}$ and $Y(t)^*Y(t) = I_p$ を満たすものであるが、

$$Y(t) = (Y, I_{n,n-p})e^{t\Delta} \begin{pmatrix} I_p \\ O \end{pmatrix}$$

とあらわせる。ただし、$I_{n,n-p} = \begin{pmatrix} I_{n-p} \\ O \end{pmatrix}$ である, [8].

Stiefel 多様体は等質空間で一般には対称空間ではない。そこで、さらに Riemann 対称空間になる **Grassmann 多様体**

$$\mathscr{G}_{n,p} = \mathscr{V}_{n,p}/\mathcal{O}(p) \cong \mathcal{O}(n)/\mathcal{O}(p) \times \mathcal{O}(n-p)$$

が様々に研究されている。定義としても、様々な手法があり、上記以外にも、部分空間として、したがって射影の空間として、もしくは symmetries の空間としての位置づけもある。 [19, 20, 21, 3, 1, 2]。無限次元化するには後者の方が都合がいいが、ここでは触れない。

Stiefel 多様体と比較すると、 $\mathcal{O}(p) \times \mathcal{O}(n-p)$ を $\mathcal{O}(n)$ に、$\mathcal{O}(p) \oplus \mathcal{O}(n-p)$ として埋め込んでやると、ファイバーに接する垂直空間は $\mathfrak{so}(p) \oplus \mathfrak{so}(n-p)$ となって、水平な接ベクトルは

$$\Delta = \begin{pmatrix} A & -B^* \\ B & C \end{pmatrix} - \begin{pmatrix} A & O \\ O & C \end{pmatrix} = \begin{pmatrix} O & -B^* \\ B & O \end{pmatrix}$$

と考えられる。

すると、$U_1, U_2 \in \mathscr{G}_{n,p}$ の測地線の一つの表現は、

$$Y(t) = (Y, I_{n,n-p})e^{t\begin{pmatrix} O & -B^* \\ B & O \end{pmatrix}} \begin{pmatrix} I_p \\ O \end{pmatrix}$$

となるが、このように初期値と初期接ベクトルによる表現でなく、両端を決めて path にしたい。そこで、$\Sigma = \Sigma^*$ に対する **CS-decomposition**

$$\exp \begin{pmatrix} O & -\Sigma \\ \Sigma & O \end{pmatrix} = \begin{pmatrix} \cos\Sigma & -\sin\Sigma \\ \sin\Sigma & \cos\Sigma \end{pmatrix}$$

によって、角度の行列を導入し、測地線の速度を調整して（$t = 1$ で終点になるよう）正規化すれば、$U_1^* U_2 = \mathrm{diag}\,(\sigma_j) = \cos\Theta$ （$0 \leq \Theta \leq I$）となるとき、

$$\begin{aligned} U(t) &= U_1 \cos t\Theta + (1 - U_1 U_1^*) U_2 \sin^\dagger \Theta \sin t\Theta \\ &= U_1 \cos t\Theta + (U_2 - U_1 \cos\Theta) \sin^\dagger \Theta \sin t\Theta \end{aligned}$$

で与えられる [6] 。ただし、† は、**Moore-Penrose 一般化逆行列**である。

## 3　同ランク PSD 行列の平均 — BS 平均—

ここで、久保-安藤平均 $\mathsf{m}$ 若しくは、その path $\mathsf{m}_t$ を考える（$t \in [0,1]$）。通常対称的な path を考えるので、$A\,\mathsf{m}_{1-t}B = B\,\mathsf{m}_t A$ を仮定し、$\mathsf{m} = \mathsf{m}_{\frac{1}{2}}$ は、対称平均とすることが多い。

同ランクの半正定値行列 $A, B$ の時は、$A$ の 0 以外の $p$ 個の固有値の単位固有ベクトルで、$V_A$ を作って $V_A^* A V_A$ で対角化し、$B$ でも同様に考えて、$V_A^* V_B$ の特値値は $[0,1]$ に収まるので、特異値 $\sigma_j$ を降順に並べた $p$ 次対角行列 $\Theta$ を考え、その対角化を導く $O_A, O_B \in \mathcal{O}(p)$; $\Theta = O_A^* V_A^* V_B O_B$ について、$U_1 = U_A = V_A O_A$, $U_2 = U_B = V_B O_B$ とすることになる。するとこの測地線は補間的であり、逆向き測地線が $U(1-t)$ で与えられるなど、性質のいいものであることがわかる。

そこで、$p$ 次正定値行列 $R_A^2 = U_A^* A U_A$, $R_B^2 = U_B^* B U_B$ によって、$\mathsf{m}$ についての **BS-mean** $\tilde{\mathsf{m}}$ を

$$A\,\tilde{\mathsf{m}}_t B = U(t)\Big(R_A^2\,\mathsf{m}_t R_B^2\Big)U(t)^*.$$

と定義することができる（これは何らかの測地線とは限らない）。path の対称性は、測地線の対称性によって遺伝し、$A\,\tilde{\mathsf{m}}_t B = B\,\tilde{\mathsf{m}}_{1-t} A$ がすぐわかる。ほかにも様々な特徴を持つ：

**Theorem 1.** *BS* 平均 $A\,\tilde{\mathsf{m}}_t B$ は以下の性質を持つ：

1. 遺伝性： $\mathrm{ran}\,A = \mathrm{ran}\,B$ ならば、久保-安藤平均と等しい：$A\,\tilde{\mathsf{m}}_t B = A\,\mathsf{m}_t B.$
   特に $A\,\tilde{\mathsf{m}}_t A = A$ であり、可逆な場合も久保-安藤平均になる。

*2.* 同次性*:* $(cA)\,\tilde{\mathsf{m}}_t(cB) = c(A\,\tilde{\mathsf{m}}_t B)$ （$c > 0$）.

*3.* ユニタリ不変*:* $(U^*AU)\,\tilde{\mathsf{m}}_t(U^*BU) = U^*(A\,\tilde{\mathsf{m}}_t B)U$ （$U$:ユニタリ）.

*4.* 単調性*:* $A \le A',\, B \le B'$ ならば、 $A\,\tilde{\mathsf{m}}_t B \le A'\,\tilde{\mathsf{m}}_t B'$.

*5.* *path* 補間性の遺伝*:* $(A\,\tilde{\mathsf{m}}_t B)\,\tilde{\mathsf{m}}_r(A\,\tilde{\mathsf{m}}_s B) = A\,\tilde{\mathsf{m}}(1-r)t + rsB$ （$s\,, t,\, r \in [0,1]$）.

特殊な性質として、久保-安藤平均とは違って一般化逆行列との親和性がある:

**Theorem 2.** 準随伴性： $A\,\tilde{\mathsf{m}}_t^* B = \left(A^\dagger\,\tilde{\mathsf{m}}_t B^\dagger\right)^\dagger$.

しかし、値域をそろえるために変換されるので、「トランス不等式」は成立しない。

## 4 射影の BS 平均と BHML 平均

Batzies-Hüper-Machado-Leitesection[5] らの射影 $P, Q$ の幾何平均の path は、BS 平均であることが確認できるが、彼らは公式として、

$$P(t) = P\tilde{\#}_t Q = e^{t\Omega} P e^{-t\Omega}, \qquad \Omega = \frac{\log(2Q-I)(2P-I)}{2}.$$

のように path を明示した。「同ランク」という仮定があると、値域を測地線による変換でそろえた場合、値域内では双方単位行列になってしまい、元の平均部分はすべて単位行列なので、測地線部分で決まるので、いかなる平均の path も上記と一致することもわかる。

$$P\,\tilde{\mathsf{m}}_t Q = e^{t\Omega} P e^{-t\Omega}.$$

ただし、当然ながら、少しでも射影と異なると平均も異なってくる。

ところで、$1\,\mathsf{m}\,0 = 0\,\mathsf{m}\,1 = 0$ となる平均については、$P\,\mathsf{m}\,Q = P \wedge Q$ であることが知られているので、値域がずれて共通部分が自明になれば、久保安藤平均では消えてしまうが、次の例のように、BS 平均では「角度情報」がしっかり残るという利点がある：

**Example.** 2 つの同ランク射影 $P = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$, $Q = \begin{pmatrix} \cos^2\theta & \cos\theta\sin\theta \\ \cos\theta\sin\theta & \sin^2\theta \end{pmatrix}$ （$Q$ は、角度 $\theta$ の部分空間に対応する射影）について、

$$2P - I = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \qquad 2Q - I = \begin{pmatrix} \cos 2\theta & \sin 2\theta \\ \sin 2\theta & -\cos 2\theta \end{pmatrix}$$

なので、

$$(2Q-I)(2P-I) = \begin{pmatrix} \cos 2\theta & -\sin 2\theta \\ \sin 2\theta & \cos 2\theta \end{pmatrix}, \quad \sqrt{(2Q-I)(2P-I)} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$$

となって、$\left((2Q-I)(2P-I)\right)^{\frac{t}{2}} = \begin{pmatrix} \cos t\theta & -\sin t\theta \\ \sin t\theta & \cos t\theta \end{pmatrix}$ となるので、path は、

$$P(t) = \begin{pmatrix} \cos t\theta & -\sin t\theta \\ \sin t\theta & \cos t\theta \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \cos t\theta & \sin t\theta \\ -\sin t\theta & \cos t\theta \end{pmatrix} = \begin{pmatrix} \cos^2 t\theta & \cos t\theta \sin t\theta \\ \cos t\theta \sin t\theta & \sin^2 t\theta \end{pmatrix}.$$

となって、角度 $t\theta$ に対応する射影になっている。

## 5 concluding remarks

さらに、初期微分係数として $\mathsf{m}$ についての solidarity

$$\tilde{\mathsf{S}}_{\mathsf{m}}(A|B) = \frac{d(A\,\tilde{\mathsf{m}}_t B)}{dt}\bigg|_{t=0} = \lim_{\varepsilon \searrow 0} \frac{A\,\tilde{\mathsf{m}}_\varepsilon B - A}{\varepsilon}$$

特に、$\mathsf{m} = \#$ の場合の相対作用素エントロピーも常に定義できる。

実際にはこの方向で無理やり rank が違うときの平均も考えられるので、可逆でない場合すべてカバーできるが、それはもはや幾何学的には意味を持たないかもしれない。

**謝辞：** 本研究は科研費 (課題番号:19K03542) の助成を受けたものである。

## 参考文献

[1] P.-A.Absil, R.Mahorny and R.Sepulchre, “Optimization Algorithms on Matrix Manifolds”, Princeton Univ. Press, 2008.

[2] P.-A.Absil, R.Mahony and R.Sepulchre, Riemannian geometry of Grassmann manifolds with a view on algorithmic computation, Acta Applicandae Mathematicae, **80**, 199–220 (2004).

[3] E.Andruchow, The Grassmann manifold of a Hilbert space, Actas del XII Congreso Dr. A.A.R.Monteiro , 2014, 41–55 (2013).

[4] R.Bhatia and J.A.R.Holbrook: Riemannian geometry and matrix geometric means, Linear Algebra Appl., **423**, 594–618 (2006).

[5] E.Batzies, H.Hüper, L.Machado and F.Silva Leite, Geometric mean and geodesic regression on Grassmanians, Linear Algebra Appl., **466**, 83–101 (2015).

[6] S.Bonnabel and R.Sepulchure, Riemannian metric and geometric mean for positive semidefinite matrices of fixed rank, SIAM, J. Matrix Anal. Appl., **31(3)**, 1055–1070 (2009).

[7] G.Corach and A.L.Maestripieri: Differential and metrical structure of positive operators, Positivity, **3**, 297–315 (1999).

[8] A.Edelman, T.A.Arias and S.T.Smith, The geometry of algorithms with orthogonality constraints, SIAM, J. Matrix Anal. Appl., **20(2)**, 303–353 (1998).

[9] J.I.Fujii, Izumino's view of operator means, Math. Japon., **33**, 671–675 (1988).

[10] J.I.Fujii and E.Kamei, Relative operator entropy in noncommutative information theory, Math. Japon., **34**, 341–348 (1989).

[11] J.I.Fujii, M.Fujii and Y.Seo, An extension of the Kubo -Ando theory: Solidarities, Math. Japon., **35**, 509–512 (1990).

[12] J.I.Fujii, Operator means and the relative operator entropy. Operator theory and complex analysis (Sapporo, 1991), 161–172, Oper. Theory Adv. Appl., **59**, Birkhäuser, Basel, 1992.

[13] J.I.Fujii, Structure of Hiai-Petz parametrized geometry for positive definite matrices, Linear Algebra Appl., **432**, 318–326 (2010).

[14] J.I.Fujii, Path of quasi-means as a geodesic, Linear Algebra Appl., **434**, 542–558 (2011).

[15] J.I.Fujii, Interpolationality for symmetric operator means, Sci. Math. Japon., **75**, 267–274 (2012).

[16] J.I.Fujii, Moore-Penrose inverse and operator mean, to appear in Sci. Math. Japon., (on line 2017-15).

[17] J.I.Fujii and Y.Seo, The relative operator entropy and the Karcher mean, Linear Algebra and Appl., **542**, 4–34 (2018).

[18] F.Kubo and T.Ando, Means of positive linear operators, Math. Ann., **246**, 205–224 (1980).

[19] H.Porta and L.Recht, Minimality of geodesics in Grassmann manifolds, Proc. Amer. Math. Soc., **100**, 464–464 (1987).

[20] A.T.Smith, Covariance, subspace, and intrinsic Cramér-Rao bounds, IEEE Transactions on Signal Processing, **53(5)**, 1610–1630 (2005).

[21] B.Vandereycken, P.-A.Absil and S.Vandewalle, Embedded geometry of the set of symmetric positive semidefinite matrices of fixed rank, 2009 IEEE/SP 15th Workshop on Statistical Signal Processing.