Geometric means inequalities and their extensions to Lie groups

by

Sima Ahsani

A dissertation submitted to the Graduate Faculty of
Auburn University
in partial fulfillment of the
requirements for the Degree of
Doctor of Philosophy

Auburn, Alabama
December 15, 2018

Keywords: Positive definite matrices, geometric means, matrix norms, semisimple Lie group, semisimple Lie algebra, Kostant pre-order

Approved by

Tin-Yau Tam, Chair, Professor of Mathematics
Ming Liao, Co-chair, Professor of Mathematics
Randall R. Holmes, Professor of Mathematics
Huajun Huang, Professor of Mathematics

Abstract

This dissertation has two main parts.

1. After reviewing the Riemannian structure of the space of $n \times n$ positive definite matrices, $\mathbb{P}_n$, and the geometric mean in terms of geodesic, t-geometric mean, we present some inequalities of Dinh, Ahsani, and Tam [15] involving t-geometric mean in the context of $\mathbb{P}_n$. Some very recent geometric inequalities of Lemos and Soares [24] are also presented.

2. After reviewing some preliminary materials of Lie groups and Lie algebras, we obtain extensions of the inequalities of Lemos and Soares in the context of semisimple Lie groups.

Acknowledgments

Many people have helped me in the way of obtaining my Ph.D. Firstly, I want to thank my parents who are a constant encouragement to me in my life and have always supported my dreams. Although it was not towards their beliefs, they never stopped me when I decided to continue my education out of my country. Now, as a parent, I understand and value their sacrifices. That is why at any stage of my life I am truly grateful to my parents.

I would like to thank my advisor Dr. Tin-Yau Tam for his patience, supports, and advice. He has been kind to me and has always listened to me patiently. Things have become easier and clearer to me because of his instruction and valuable advice. Dr. Tam kindly gave me opportunity to assist him in reviewing and editing papers. He also supported me to attend many conferences.

I wish to thank my committee members, Dr. Randall Holmes, Dr. Huajun Huang, and Dr. Ming Liao (Co-chair). They have always been open and helpful when I had questions or needed their help. I am grateful for having opportunity to learn from each of them. Without their help, it would have been impossible for me to finish this dissertation.

I want to say thank you to my husband, Reza Hemmati. He has been extremely supportive during my graduate education at Auburn University. Words can not express my appreciation for what he has done for me during these years. He made it possible for me to continue my research by taking care of our daughter, Arianna, during weekends and whenever I was attending conferences. He has been patient with me when I got frustrated. I know life was challenging for both of them during my PhD study. So, at the end, I would like to dedicate this work to my love, Reza, and joy of life, my little lovely daughter, Arianna.

Table of Contents

List of Figures

Chapter 1

Introduction

The main topics of this dissertation are matrix geometric means inequalities and their extensions to semisimple Lie groups. The challenge is to define geometric mean for two positive definite matrices and then in the context of Lie group. The Riemannian geometric point of view gives us a good understanding of the geometric mean of two positive definite matrices. Connection between geometric mean and differential geometry, for the first time, was discussed by Moakher in 2005 in order to find an appropriate definition for the geometric mean for more than two matrices. Indeed, the geometric mean of two matrices is the midpoint of the geodesic joining those matrices. In some areas, such as machine learning and optimization, positive definite matrices are essential tools. The view of P_n as a subset of a Euclidean space is not helpful, and it is more useful to consider it as a Riemannian manifold with an appropriate metric [34]. Furthermore, this view provides natural extension to Lie groups [25].

The following is the organization of this dissertation.

In Chapter 2, we review the Riemannian manifold structure of $\mathbb{P}_n$. The point on $\gamma(s)$ at $s = t$ is denoted by $A\#_t B$, where $\gamma : [0,1] \to \mathbb{P}_n$ is the geodesic joining $A, B \in \mathbb{P}_n$ and $A\#_t B$ is called the t-geometric mean of A and B.

In Chapter 3, we discuss the concept of log-majorization and related inequalities. Then we provide some results for t-geometric mean in [15].

In Chapter 4, we approach the geometric mean from the operator theory and present some very recent log-majorization inequalities of Lemos and Soares [24].

In Chapter 5, we provide some norm inequalities for the matrix geometric mean in Dinh, Ahsani, and Tam [15]. Audenaert's result [5] is the motivation of some of the results in [15].

It also motivated us to ask some questions. In fact, some of these questions were stated as a theorem in [15] and some of them were posted as a conjecture. At the end of this chapter we present some affirmative solutions under some conditions. But the conjecture is still unsolved.

In Chapter 6, we give a brief review of semisimple Lie groups and Lie algebras. We discuss Cartan decomposition, Iwasawa decomposition, and Complete Multiplicative Jordan Decomposition (CMJD). Then we introduce Kostant pre-order and an important characterization by Kostant.

In Chapter 7, we extend some log-majorization inequalities of Lemos and Soares [24] in terms of Kostant pre-order for semisimple Lie groups.

Chapter 2

Geometric mean from differential geometry point of view

Let $\mathbb{R}_+$ denote the set of positive real numbers. For $a, b \in \mathbb{R}_+$, the geometric mean is

$$\sqrt{ab}.$$

Operator theorists and physicists were the first scientists considering geometric mean extension for matrices. A good platform for the extension is $\mathbb{P}_n$, the set of all $n \times n$ positive definite matrices. Pusz and Woronowicz [32] defined the geometric mean of $A, B \in \mathbb{P}_n$ as

$$A\#B := A\#_{\frac{1}{2}}B = A^{1/2}(A^{-1/2}BA^{-1/2})^{1/2}A^{1/2}. \tag{2.0.1}$$

It turns out that this is the mid point of the geodesic joining A and B when we view $\mathbb{P}_n$ as a Riemannian manifold. We would like to provide some background information of $\mathbb{P}_n$ and most of the material in this chapter can be found in Bhatia [7].

2.1 $\mathbb{P}_n$ as a Riemannian manifold

In this section we will show that $\mathbb{P}_n$ is a Riemannian manifold which is differentiable manifold equipped with a Riemannian metric. Let $\mathcal{M}$ be a differentiable manifold. A *Rimannian metric* on $\mathcal{M}$ is given by an inner product on each tangent space $T_A\mathcal{M}$ which depends smoothly on the original point A.

Let $\mathbb{C}_{n\times n}$ be the set of all $n \times n$ complex matrices equipped the inner product $\langle A, B\rangle = \operatorname{tr} A^*B$ which induces the Frobenius norm

$$||A||_2 = \sqrt{\operatorname{tr} A^*A}.$$

Note that $\mathbb{C}_{n\times n}$ is an $2n^2$-dimensional vector space over $\mathbb{R}$. The set of all $n \times n$ Hermitian matrices, $\mathbb{H}_n$, is an n^2-dimensional real subspace of $\mathbb{C}_{n\times n}$.

The exponential of $A \in \mathrm{GL}_n(\mathbb{C})$ is given by

$$\exp A := e^A = \sum_{k=0}^{\infty} \frac{1}{k!} A^k. \tag{2.1.1}$$

Let $A \in \mathrm{GL}_n(\mathbb{C})$ be given such that $||A - I_n|| < 1$ for some matrix norm $|| \cdot ||$. The logarithm of A is defined by

$$\log A = \sum_{k=0}^{\infty} (-1)^{k+1} \frac{(A - I_n)^k}{k}. \tag{2.1.2}$$

The exponential map $\exp_A : T_A\mathbb{P}_n \to \mathbb{P}_n$ at $A \in \mathbb{P}_n$ maps $h \in T_A\mathbb{P}_n$ to $B := \exp_A h$ that lies on the geodesic $\gamma(t) : [0, 1] \to \mathbb{P}_n$ with $\gamma(0) = A, \gamma(1) = B$ and $\gamma^{'}(0) = h$.

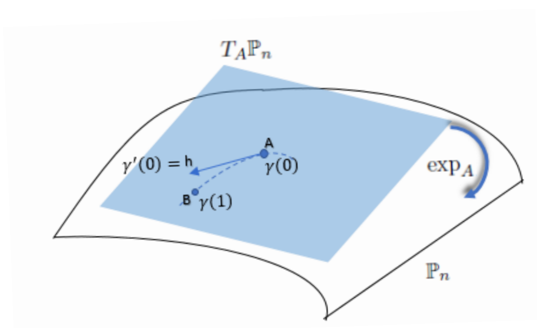

Figure 2.1: Geodesic $\gamma(t)$ with the starting point $\gamma(0) = A$ and the end point $\gamma(1) = B$

Identifying $T_A\mathbb{P}_n$ with $\mathbb{H}_n$, the inner product in $T_A\mathbb{P}_n$ is given by

$$\langle X, Y\rangle_A = \mathrm{tr}\,(A^{-1/2}XA^{-1}YA^{-1/2}), \tag{2.1.3}$$

where $X, Y \in T_A\mathbb{P}_n$. Thus, the induced norm on the tangent space is

$$||X||_A = \sqrt{\mathrm{tr}\,(A^{-1}X)^2} = ||A^{-1/2}XA^{-1/2}||_2.$$

The inner product (2.1.3) leads to a Riemannian metric on $\mathbb{P}_n$. Let $\gamma : [0,1] \to \mathbb{P}_n$ be a differentiable curve. At any point $\gamma(t_0)$ where, $t_0 \in [0,1]$, there is a tangent vector $\dot{\gamma}(t_0)$ in the tangent space $T_{\gamma(t_0)}\mathbb{P}_n$ and

$$\|\dot{\gamma}(t_0)\| = \|\gamma(t_0)^{-1/2}\gamma(\dot{t_0})\gamma(t_0)^{-1/2}\|_2,$$

in which $\|\cdot\|$ is the norm induced by the inner product on $T_{\gamma(t_0)}\mathbb{P}_n$. Therefore, the length of the curve γ is

$$L(\gamma) = \int_0^1 \|\dot{\gamma}(t)\|dt.$$

For a piecewise differentiable path $\gamma : [a,b] \to \mathbb{P}_n$ in $\mathbb{P}_n$, the length of γ is defined as

$$\mathrm{L}(\gamma) = \int_a^b ds\ dt,$$

where $ds = \|\dot{\gamma}(t)\|$. We define

$$\delta_R(A,B) := \inf\{\mathrm{L}(\gamma) : \gamma \text{ is a path from } A \text{ to } B\}, \tag{2.1.4}$$

which is a metric on $\mathbb{P}_n$ and we would like to work out the explicit form of $\delta_R(A,B)$.

2.2 Parametrization of geodesic joining two points in $\mathbb{P}_n$

Definition 2.2.1. Let $X \in \mathrm{GL}_n(\mathbb{C})$. A transformation $\Gamma_X : \mathbb{P}_n \to \mathbb{P}_n$ defined as

$$\Gamma_X(A) = X^*AX, \quad X \in \mathbb{P}_n,$$

is called a *congruence transformation.*

Some of the nice properties of Γ_X are:

- $\Gamma_X(A) \geq 0$ whenever $A \geq 0$.

- $\Gamma_X(A+B) = \Gamma_X(A) + \Gamma_X(B)$.
- $\Gamma_U(AB) = \Gamma_U(A)\Gamma_U(B)$, where U is unitary.
- If $A \leq B$, then $\Gamma_X(A) \leq \Gamma_X(B)$.

In fact, Γ_X is a bijection of $\mathbb{P}_n$ to itself and is an isometry for the length $L(\gamma)$ ([7, Lemma 6.1.1]). It is an isometry for the metric $\delta_R(A, B)$, i.e.,

$$\mathrm{L}(\Gamma_X \circ \gamma) = \gamma$$

and

$$\delta_R(\Gamma_X(A), \Gamma_X(B)) = \delta_R(A, B).$$

For each $A \in \mathbb{P}_n$, there exists a unique $B \in \mathbb{H}_n$ such that $A = \exp B$, or equivalently, $\log A = B$. Therefore, if $\gamma(t)$ is a path joining A and B in $\mathbb{P}_n$, then $\beta(t)$ is a path joining $\log A$ and $\log B$ in $\mathbb{H}_n$.

Note that if $A, B \in \mathbb{P}_n$ commute, then the geodesic joining $A, B \in \mathbb{P}_n$ and the straight line segment $[\log A, \log B] \in \mathbb{H}$ coincide and we have

Proposition 2.2.2. *([7, Proposition 6.5.1])*

$$\delta_R(A, B) = ||\log A - \log B||_2. \tag{2.2.1}$$

We would like to have explicit expressions of the Riemannian distance between $A, B \in \mathbb{P}_n$ and the parametrization of the geodesic joining A and B when $AB \neq BA$. Consider $\beta(t)$, the straight line segment joining $\log A$ and $\log B$ in $\mathbb{H}_n$. The natural parametrization of $\beta(t)$ is

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

A natural way to parameterize geodesic $\gamma(t)$ when A and B commute is

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

Consider the matrices I and $A^{-1/2}BA^{-1/2}$. Since they commute, the parametrization of geodesic $[I, A^{-1/2}BA^{-1/2}]$ is $\gamma_0(t) = (A^{-1/2}BA^{-1/2})^t$. Applying $\Gamma_{A^{1/2}}$,

$$\gamma(t) = \Gamma_{A^{1/2}}(\gamma_0(t)) = A^{1/2}(A^{-1/2}BA^{-1/2})^t A^{1/2} \tag{2.2.2}$$

is the parametrization of the geodesic between $\Gamma_{A^{1/2}}I = A$ and $\Gamma_{A^{1/2}}(A^{-1/2}BA^{-1/2}) = B$. We call $\gamma(t)$ the *t-geometric mean* of A and B and from now on we denote it by

$$A\#_tB := A^{1/2}(A^{-1/2}BA^{-1/2})^t A^{1/2}.$$

It can be seen that the *matrix geometric mean*

$$A\#B := A\#_{\frac{1}{2}}B$$

is the unique midpoint between A and B and

$$A\#B = B\#A.$$

Moreover, the Riemaniann distance between A and B in $\mathbb{P}_n$ is

$$\begin{aligned}\delta_R(A,B) &= \delta_R(\Gamma_{A^{1/2}}I, \Gamma_{A^{1/2}}(A^{-1/2}BA^{-1/2})) \\ &= \delta_R(I, A^{-1/2}BA^{-1/2}) \\ &= ||\log I - \log A^{-1/2}BA^{-1/2}||_2 \\ &= ||\log A^{-1/2}BA^{-1/2}||_2 \\ &= \Big(\sum_{i=1}^{n} \log^2 \lambda_i(A^{-1}B)\Big)^{1/2}.\end{aligned}$$

Chapter 3

Geodesic convexity of t-geometric means

In this chapter we will present several results by Dinh, Ahsani, and Tam [15] on the geometric mean of two positive definite matrices.

Let $A \in \mathbb{C}_{n\times n}$. Denote by $\lambda(A) = (\lambda_1(A), \dots, \lambda_n(A))$ the vector of eigenvalues of A and usually we arrange its entries in non-increasing order with respect to their moduli, i.e.,

$$|\lambda_1(A)| \geq \cdots \geq |\lambda_n(A)|.$$

Both A^*A and AA^* are positive semidefinite and have the same spectrum counting multiplicities. The eigenvalues $s_1(A), \dots, s_n(A)$ of $\sqrt{A^*A}$ (or $\sqrt{AA^*}$) are called the *singular values* of A. Denoted by $s(A) = (s_1(A), \dots, s_n(A))$ the vector of singular values and arrange the singular values in non-increasing order, i.e.,

$$s_1(A) \geq \cdots \geq s_n(A).$$

3.1 Majorization and log-majoration

Let $x, y \in \mathbb{R}^n$. We denote by $x^\downarrow$ the vector that has the same component as x has, but is sorted in a non-increasing order, i.e., $x_1^\downarrow \geq \cdots \geq x_n^\downarrow$. Let $x, y \in \mathbb{R}^n$. We say that x is *majorized* by y, denoted by $x \prec y$, if

$$\sum_{i=1}^k x_i^\downarrow \leq \sum_{i=1}^k y_i^\downarrow, \quad k = 1, 2, \dots, n-1, \tag{3.1.1}$$

$$\sum_{i=1}^n x_i = \sum_{i=1}^n y_i. \tag{3.1.2}$$

We say that x is *weakly majorizied* by y, denoted by $x \prec_w y$ if (3.1.1) holds. We say that x is *log-majorized* by y, denoted by $x \prec_{\log} y$, if

$$\begin{aligned} \prod_{i=1}^{k} x_i^{\downarrow} &\le \prod_{i=1}^{k} y_i^{\downarrow}, \quad k = 1, 2, \ldots, n-1, \\ \prod_{i=1}^{n} x_i &= \prod_{i=1}^{n} y_i \end{aligned}$$

We remark that if $x, y \in \mathbb{R}_+^n$, then $x \prec_{\log} y$ if and only if $\log x \prec \log y$.

3.2 Compound matrices

We now provide a quick review of compound matrices which will be used later. Suppose $A \in \mathbb{C}_{n\times n}$. For index sets $\alpha \subset \{1, \ldots, n\}$ and $\beta \subset \{1, \ldots, n\}$, we denote by $A[\alpha|\beta]$ the submatrix of A whose entries lie in the rows of A indexed by α and the columns indexed by β. For all $1 \le k \le n$, the *k-th compound* of A is defined as the $\binom{n}{k} \times \binom{n}{k}$ complex matrix $C_k(A)$ whose elements are given by

$$C_k(A)_{\alpha,\beta} = \det A[\alpha|\beta], \quad \alpha, \beta \in Q_{k,n} \tag{3.2.1}$$

where

$$Q_{k,n} = \{\omega = (\omega(1), \ldots, \omega(k)) : 1 \le \omega(1) < \cdots < \omega(k) \le n\}$$

is the set of increasing sequences of length k chosen from $\{1, \ldots, n\}$. In particular, $C_1(A) = A$ and $C_n(A) = \det A$.

Compound matrices have many nice properties which are listed below. See [27, 28, 30] for proofs.

Theorem 3.2.1. *Let $A, B \in \mathbb{C}_{n\times n}$. Let $s(A) = (s_1, \ldots, s_n)$ and $\lambda(A) = (\lambda_1, \ldots, \lambda_n)$ denote the vector of singular values of A in non-increasing order and the vector of eigenvalues* *of A whose absolute values are in non-increasing order, respectively. Then the following statements are true.*

(1) $C_k(A^*) = [C_k(A)]^*$.

(2) If $A = (a_{ij})$ *is upper triangular, then so is* $C_k(A)$ *and its diagonal entries are*

$$\prod_{j=1}^{k} a_{\omega(j),\omega(j)}$$

for all $\omega \in Q_{k,n}$.

(3) The eigenvalues of $C_k(A)$ *are*

$$\prod_{j=1}^{k} \lambda_{\omega(j)}$$

for all $\omega \in Q_{k,n}$.

(4) The singular values of $C_k(A)$ *are*

$$\prod_{j=1}^{k} s_{\omega(j)}$$

for all $\omega \in Q_{k,n}$.

(5) If A is unitary, then so is $C_k(A)$.

(6) If A is positive semidefinite, then so is $C_k(A)$.

(7) $C_k(AB) = C_k(A)C_k(B)$, *which is called the Binet-Cauchy Theorem. Thus the map* $C_k : \mathrm{GL}_n(\mathbb{C}) \to \mathrm{GL}_{\binom{n}{k}}(\mathbb{C})$ *is a group homomorphism.*

Compound matrix can be used to prove the following theorem of Weyl, which states that log-majorization is the relation between the eigenvalues and singular values of $A \in \mathbb{C}_{n\times n}$.

Theorem 3.2.2. *(Weyl [6])*

Let $A \in \mathbb{C}_{n\times n}$. *Then* $|\lambda(A)| \prec_{\log} s(A)$.

Proof. Note that

$$s_1(A) = \max\{|(Ax, y)| : x, y \in \mathbb{C}^n, \|x\|_2 = \|y\|_2 = 1\}.$$

Suppose $Ax = \lambda_1 x$ for some unit vector $x \in \mathbb{C}^n$. Then

$$|\lambda_1| = |(Ax, x)| \leq s_1(A).$$

Apply this result to the k-th compound of A to conclude that for all $k = 1, \ldots, n-1$,

$$\lambda_1(A) \cdots \lambda_k(A) = \lambda_1(C_k(A)) \leq s_1(C_k(A)) \leq s_1(A) \cdots s_k(A).$$

The equality $\lambda_1(A) \cdots \lambda_n(A) = s_1(A) \cdots s_n(A)$ follows by determinantal consideration. □

3.3 Matrix geometric mean and log-majorization

The following interesting results can be found in Bhatia and Grover [8, p.730].

Theorem 3.3.1. *(Bhatia and Grover [8, p.730])*
Let $A, B \in \mathbb{P}_n$. For any $t \in [0, 1]$ and $s > 0$,

$$\begin{aligned} \lambda(A \#_t B) &\prec_{\log} \lambda\left(e^{(1-t)\log A + t\log B}\right) \\ &\prec_{\log} \lambda\left(B^{ts/2} A^{(1-t)s} B^{ts/2}\right)^{1/s} \\ &= \lambda\left(A^{(1-t)s} B^{ts}\right)^{1/s}. \end{aligned} \tag{3.3.1}$$

The first inequality is a result of Ando and Hiai [2, Corollary 2.3] as the complementary counterpart of the famous Golden-Thompson inequality for $A, B \in \mathbb{H}_n$:

$$\operatorname{tr} e^{A+B} \leq \operatorname{tr} e^A e^B.$$

The second inequality follows from a result of Araki [4].

Recall that a norm $|||\cdot|||$ on $\mathbb{C}_{n\times n}$ is *unitarily invariant* if for all $U, V \in \mathrm{U}(n)$ and $A \in \mathbb{C}_{n\times n}$

$$|||\,UAV|||\; = |||\,A|||\,,$$

where $\mathrm{U}(n)$ is the unitary group. For example the spectral norm $\|\cdot\|$, i.e., $\|A\| = s_1(A)$, $A \in \mathbb{C}_{n\times n}$, is unitarily invariant. Ky Fan k-norm, $|||\,A|||_{(k)} = s_1(A) + \cdots + s_k(A)$, $A \in \mathbb{C}_{n\times n}$, is another example.

The following is the well-known Ky Fan Dominance Theorem.

Theorem 3.3.2. *(Ky Fan Dominance Theorem [6, p. 93])*
Given $A, B \in \mathbb{C}_{n\times n}$, $s(A) \prec_w s(B)$ if and only if $|||\,A|||\; \le |||\,B|||$ for all unitarily invariant norms $|||\cdot|||$, where $s(A)$ denotes the vector of singular values of A.

Proposition 3.3.3. *(Dinh, Ahsani, Tam [15, Proposition 2.2])*
Let $A, B \in \mathbb{P}_n$ and $t \in [0,1]$ and $s > 0$. For all unitarily invariant norms $|||\cdot|||$ on $\mathbb{C}_{n\times n}$,

$$|||\,A\#_t B|||\; \le |||\,\left(B^{ts/2}A^{(1-t)s}B^{ts/2}\right)^{1/s}|||\; \le |||\,\left(A^{(1-t)s}B^{ts}\right)^{1/s}|||\,. \qquad (3.3.2)$$

In particular, with $s = 1$, $t = 1/2$,

$$|||\,A^2\#B^2|||\; \le \min\{|||\,A^{1/2}BA^{1/2}|||\,, |||\,B^{1/2}AB^{1/2}|||\,\} \le \min\{|||\,AB|||\,, |||\,BA|||\,\}$$

and

$$|||\,(A\#B)^2|||\; \le \min\{|||\,AB|||\,, |||\,BA|||\,\}.$$

Proof. Since $A\#_t B \in \mathbb{P}_n$ for all $t \in [0,1]$,

$$s(A\#_t B) = \lambda(A\#_t B) \prec_{\log} \lambda\left(B^{ts/2}A^{(1-t)s}B^{ts/2}\right)^{1/s}$$

for all $t \in [0,1]$ and $s > 0$, by Theorem 3.3.1. By Theorem 3.2.2 we have

$$\lambda\left(B^{ts/2}A^{(1-t)s}B^{ts/2}\right)^{1/s} \prec_{\log} s\left(B^{ts/2}A^{(1-t)s}B^{ts/2}\right)^{1/s}.$$

Then apply Theorem 3.3.2 to have

$$|||\, A\#_t B||| \leq |||\, \left(B^{ts/2}A^{(1-t)s}B^{ts/2}\right)^{1/s} |||\,.$$

The second inequality follows from

$$s\left(B^{ts/2}A^{(1-t)s}B^{ts/2}\right)^{1/s} = \lambda\left(B^{ts/2}A^{(1-t)s}B^{ts/2}\right)^{1/s} = \lambda\left(A^{(1-t)s}B^{ts}\right)^{1/s} \prec_{\log} s\left(A^{(1-t)s}B^{ts}\right)^{1/s}$$

and Theorem 3.3.2. So, we have (3.3.2). The last inequality follows from Theorem 3.3.1. □

Let $S \subseteq \mathbb{R}^n$ be a finite set. The *convex hull* of S, denoted by $\operatorname{conv} S$, is the smallest convex set containing S. Let S_n be the symmetric group of order n, let $S_n x := \{\sigma x : \sigma \in S_n\}$ be the orbit of x under the action of the symmetric group S_n. For any $x, y \in \mathbb{R}^n$, it is known [6, 20] that $x \prec y$ is equivalent to $\operatorname{conv} S_n x \subset \operatorname{conv} S_n y$.

A subset S of $\mathbb{P}_n$ is called a *geodesically convex set* if for every $A, B \in S$, the geodesic joining A and B is contained in S. Let X be a subset of $\mathbb{P}_n$. The *geodesic convex hull* of X is $\bigcap\{S \in \mathbb{P}_n : S \supseteq X \text{ and geodesically convex set}\}$, i.e., the smallest geodesically convex set containing X.

Let us recall a result of Thompson, which motivates the study of Theorem 3.3.8.

Theorem 3.3.4. *(Thompson [37, Theorem 12])*
Let $\omega = (\omega_1, \dots, \omega_n) \in \mathbb{R}^n$. Let S be the set of Hermitian matrices with prescribed eigenvalues $\omega_1, \dots, \omega_n$, i.e.,

$$S = \{A \in \mathbb{H}_n : \lambda(A) = \omega\}. \tag{3.3.3}$$

The convex hull of S is the set of all Hermitian matrices with spectrum consisting $\rho = (\rho_1, \dots, \rho_n)$ satisfying $\rho \prec \omega$.

For any $A \in \mathbb{P}_n$ define,

$$M(A) = \{B \in \mathbb{P}_n : \lambda(B) \prec_{\log} \lambda(A)\} \subset \mathbb{P}_n. \tag{3.3.4}$$

The map $\exp : \mathbb{H}_n \to \mathbb{P}_n$ is a diffeomorphism and its inverse $\log : \mathbb{P}_n \to \mathbb{H}_n$ is defined. Therefore, the image of $M(A)$ under the map log is

$$\log(M(A)) = \{H \in \mathbb{H}_n : \lambda(H) \prec \lambda(\log A)\}. \tag{3.3.5}$$

This set is convex in $\mathbb{H}_n$. By Theorem 3.3.4, it is the convex hull of the set

$$S = \{X \in \mathbb{H}_n : \lambda(X) = \lambda(\log(A))\}. \tag{3.3.6}$$

In other words, it consists of all Hermitian matrices with spectrum coincided with $\lambda(\log A)$. Noted that in general $M(A)$ is not closed under the usual matrix addition and as a result $M(A)$ is not convex in $\mathbb{P}_n$ when it is viewed as a subset of the Euclidean space $\mathbb{C}_{n\times n}$.

Theorem 3.3.5. *(Dinh, Ahsani, Tam [15, Theorem 2.3])*
Given $A \in \mathbb{P}_n$. $M(A)$ is geodesically convex with respect to the Riemannian structure of $\mathbb{P}_n$. In other words, if $B, C \in M(A)$, then the geodesic joining B and C lies in $M(A)$. So

$$M(A) = \{A\#_t B : t \in [0,1], B \in \mathbb{P}_n, \lambda(B) \prec_{\log} \lambda(A)\}. \tag{3.3.7}$$

Proof. Let $\|\cdot\|$ denote the spectral norm on $\mathbb{C}_{n\times n}$, which is unitarily invariant. By Proposition 3.3.3, if $t \in (0,1)$, then

$$\|B\#_t C\| \le \|B^{1-t}C^t\| = \|(B^{\frac{1-t}{t}})^t C^t\| \le \|B^{\frac{1-t}{t}}C\|^t$$

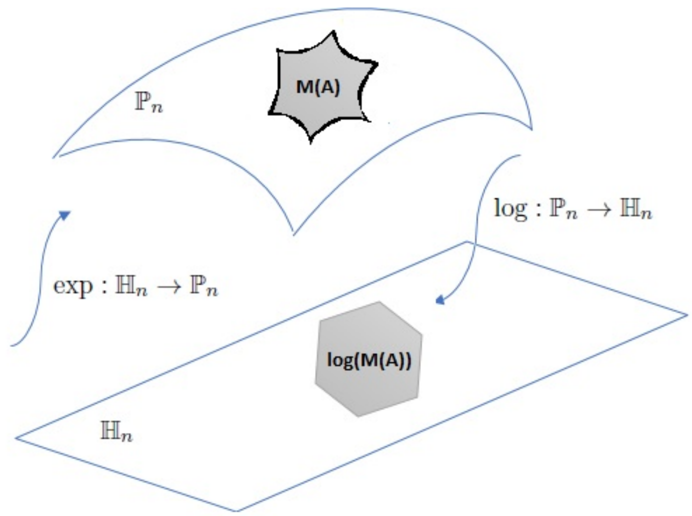

Figure 3.1: $M(A)$ is geodesically convex in $\mathbb{P}_n$.

by [6, Theorem IX.2.1]. Now

$$\|B\#_t C\| \le \|B^{\frac{1-t}{t}} C\|^t \le \|B^{\frac{1-t}{t}}\|^t \|C\|^t = \|B\|^{1-t}\|C\|^t, \tag{3.3.8}$$

which is no greater than $\|A\|$ since $\lambda(B), \lambda(C) \prec_{\log} \lambda(A)$. So

$$\lambda_1(B\#_t C) = s_1(B\#_t C) = \|B\#_t C\| \le \|A\| = s_1(A) = \lambda_1(A). \tag{3.3.9}$$

Denote by $C_k(X)$ the k-th compound of $X \in \mathbb{C}_{n\times n}$, $k = 1, \ldots, n$. Note that for any $X, Y \in \mathbb{P}_n$,

$$\begin{aligned} C_k(X\#_t Y) &= C_k(X^{1/2}(X^{-1/2}YX^{-1/2})^t X^{1/2}) \\ &= C_k(X^{1/2})C_k((X^{-1/2}YX^{-1/2})^t)C_k(X^{1/2}) \\ &= C_k^{1/2}(X)(C_k^{-1/2}(X)C_k(Y)C_k^{-1/2}(X))^t C_k^{1/2}(X) \\ &= C_k(X)\#_t C_k(Y). \end{aligned} \tag{3.3.10}$$

In other words, C_k respects $\#_t$ in $\mathbb{P}_n$. By Theorem 3.2.1, the $\binom{n}{k}$ eigenvalues of $C_k(X)$, where $X \in \mathbb{C}_{n\times n}$, are the $\binom{n}{k}$ possible products of any k eigenvalues of X. So

$$\lambda_1(C_k(B\#_t C)) = \prod_{i=1}^{k} \lambda_i(B\#_t C), \quad k = 1, \ldots, n-1, \tag{3.3.11}$$

and

$$\det(B\#_t C) = (\det B)^{1-t}(\det C)^t = \det A. \tag{3.3.12}$$

Applying (3.3.8) on $C_k(B)$ and $C_k(C)$ that are both positive definite, we have

$$\begin{aligned}
\prod_{i=1}^{k} \lambda_i(B\#_t C) &= \lambda_1(C_k(B)\#_t C_k(C)) \quad \text{by (3.3.11)} \\
&= \|C_k(B)\#_t C_k(C)\| \\
&\le \|C_k(B)\|^t \|C_k(C)\|^{1-t} \quad \text{by (3.3.8)} \\
&= (\prod_{i=1}^{k} \lambda_i(B))^t (\prod_{i=1}^{k} \lambda_i(C))^{1-t} \\
&\le \prod_{i=1}^{k} \lambda_i(A), \quad i = 1, \ldots, n-1.
\end{aligned}$$

Together with (3.3.12), we conclude that $\lambda(B\#_t C) \prec_{\log} \lambda(A)$. □

Corollary 3.3.6. *(Dinh, Ahsani, Tam [15, Corollary 2.4])*
If $A, B \in \mathbb{P}_n$ such that $B \prec_{\log} A$, then for all $t \in [0,1]$, we have $A\#_t B \prec_{\log} A$, or equivalently, $|||\, A\#_t B ||| \le |||\, A |||$ for all unitarily invariant norms $|||\cdot|||$ on $\mathbb{C}_{n\times n}$.

We are going to show that $M(A)$ has a nice geometric description and our proof requires a lemma, which is of independent interest. Let $\alpha, \beta \in \mathbb{R}^n$. We say that β is a pinch of α [29, p.17] if

$$\beta = (\lambda I + (1-\lambda)Q)\alpha,$$

where Q is the permutation matrix that interchanges two coordinates. It is well known that if $\beta \prec \alpha$, then β can be obtained by applying at most n pinches consecutively, starting from

α. The converse is clearly true. Now let $\alpha, \beta \in \mathbb{R}^n_+$, where $\mathbb{R}^n_+$ denotes the set of all positive n-tuples. If $\beta \prec_{\log} \alpha$, then β can be obtained by applying at most n pinches multiplicatively in the following sense. We say that β is a *geometric pinch* of α if

$$\text{diag}\,(\beta_1, \dots, \beta_n) = (Q^\top \text{diag}\,(\alpha_1, \dots, \alpha_n) Q) \#_t \text{diag}\,(\alpha_1, \dots, \alpha_n)$$

for some $t \in [0, 1]$, and some transposition matrix Q.

Lemma 3.3.7. *(Dinh, Ahsani, Tam [15, Lemma 2.5])*
Let $\alpha, \beta \in \mathbb{R}^n_+$, where $\mathbb{R}^n_+$ denotes the set of all positive n-tuples. If $\beta \prec_{\log} \alpha$, then β can be obtained by applying at most n geometric pinches consecutively, starting from α.

Proof. Since $\log \beta \prec \log \alpha$, $\log \beta$ can be obtained by at most n pinches from $\log \alpha$. Let $\log \hat{\alpha}$ be a pinch of $\log \alpha$. Without loss of generality, we may assume that the pinch occurs on the first two coordinates. So $(\hat{\alpha}_1, \hat{\alpha}_2) \prec_{\log} (\alpha_1, \alpha_2)$ and thus $\hat{\alpha}_1 = \alpha_1^t \alpha_2^{1-t}$ and $\hat{\alpha}_2 = \alpha_2^t \alpha_1^{1-t}$ for some $t \in [0, 1]$. Let P denote the matrix corresponding to the transposition switching the first two coordinates. Then

$$\begin{aligned} & (P^\top \text{diag}\,(\alpha_1, \alpha_2, \alpha_3, \dots, \alpha_n) P) \#_t \text{diag}\,(\alpha_1, \alpha_2, \alpha_3, \dots, \alpha_n) \\ = \;& \text{diag}\,(\alpha_2, \alpha_1, \alpha_3, \dots, \alpha_n) \#_t \text{diag}\,(\alpha_1, \alpha_2, \alpha_3, \dots, \alpha_n) \\ = \;& \text{diag}\,(\alpha_1^t \alpha_2^{1-t}, \alpha_2^t \alpha_1^{1-t}, \alpha_3, \dots \alpha_n) \\ = \;& \hat{\alpha}. \end{aligned}$$

Then repeat the process to conclude that there exist $t_1, \dots, t_k \in [0, 1]$ and transposition matrices $P_1, \dots, P_k$ such that

$$\text{diag}\,\alpha^{(i+1)} := (P_i^\top (\text{diag}\,\alpha^{(i)}) P_i) \#_{t_i} \text{diag}\,\alpha^{(i)}, \quad i = 1, \dots, k,$$

where $\alpha^{(1)} := \alpha$ and $\alpha^{(k+1)} := \beta$. $\square$

The unitary similarity orbit of $A \in \mathbb{P}_n$, denoted by $O(A)$, is the set of all matrices in $\mathbb{P}_n$ that are unitarily similar to A, i.e., $O(A) = \{U^*AU : U \in \mathrm{U}(n)\}$, where $\mathrm{U}(n)$ is the unitary group. By the Spectral Theorem of Hermitian matrices, one can see that $O(A)$ is equal to

$$S = \{B \in \mathbb{P}_n : \lambda(B) = \lambda(A)\}.$$

The set S is the collection of $B \in \mathbb{P}_n$ whose spectrum coincides with that of A. The following result gives a nice geometric relation between $M(A)$ and $O(A)$.

Theorem 3.3.8. *(Dinh, Ahsani, Tam [15, Theorem 2.6])*
The set $M(A)$ is the geodesic convex hull, denoted by $G(A)$, of the orbit $O(A)$.

Proof. By the Spectral Theorem of Hermitian matrices, it is easy to see that $O(A) := \{UAU^* : U \in \mathrm{U}(n)\}$ is equal to

$$\{B \in \mathbb{P}_n : \lambda(B) = \lambda(A)\} \subset M(A).$$

So $G(A) \subset M(A)$ as $M(A)$ is geodesically convex by Theorem 3.3.5. Thus it suffices to show that $M(A) \subset G(A)$. Let $B \in M(A)$, that is, $\lambda(B) \prec_{\log} \lambda(A)$. Since $M(A)$, $O(A)$ and thus $G(A)$ are invariant under unitarily similarity, we may assume that $A = \mathrm{diag}\,(\alpha_1, \ldots, \alpha_n)$. Let $\lambda(B) = (\beta_1, \ldots, \beta_n)$ and let $U \in \mathrm{U}(n)$ such that $B = U^*\mathrm{diag}\,(\beta_1, \ldots, \beta_n)U$. By Lemma 3.3.7, there exist $t_1, \ldots, t_k \in [0, 1]$ and transposition matrices $P_1, \ldots, P_k$ such that

$$\mathrm{diag}\,\alpha^{(i+1)} := (P_i^\top(\mathrm{diag}\,\alpha^{(i)})P_i)\#_{t_i}\mathrm{diag}\,\alpha^{(i)}, \quad i = 1, \ldots, k,$$

where $\alpha^{(1)} := \alpha$ and $\alpha^{(k+1)} := \beta$. It is easy to see that

$$V^*(C\#_t D)V = (V^*CV)\#_t(V^*DV)$$

for all $V \in \mathrm{U}(n)$, $C, D \in \mathbb{P}_n$. So

$$\begin{aligned} B &= U^* \mathrm{diag}\,(\beta_1, \ldots, \beta_n) U \\ &= U^* (P_k^\top (\mathrm{diag}\,\alpha^{(k)}) P_k) \#_{t_k} \mathrm{diag}\,\alpha^{(k)}) U \\ &= (U^* P_k^\top (\mathrm{diag}\,\alpha^{(k)}) P_k U) \#_{t_k} (U^* \mathrm{diag}\,\alpha^{(k)} U). \end{aligned}$$

Then use induction on k to show that $B \in G(A)$ as $U^* P_k^\top (\mathrm{diag}\,\alpha^{(k)}) P_k U$ and $U^* \mathrm{diag}\,\alpha^{(k)} U \in G(A)$ since $G(A)$ is invariant under unitarily similarity. □

Chapter 4

Geometric mean from operator theory point of view

The differential geometry view of the geometric mean of two positive definite matrices gives a geometric insight of geometric mean in terms of geodesic when $\mathbb{P}_n$ is considered as a Riemannian manifold. However, Loewner order $\leq$ on $\mathbb{P}_n$ or $\mathbb{H}_n$ plays an important role in operator theory which is not present in the differential geometry approach. Recall the Loewer order: given $A, B \in \mathbb{H}_n$, $B \leq A$ means $A - B \geq 0$, i.e, positive semidefinite. One important property that we expect a geometric mean of two matrices in $\mathbb{P}_n$ to have is *monotonicity* with respect to the Loewer order. There are also important inequalities and nice properties of geometric mean in operator theory. Most of the material in this chapter can be found in [7]. We start with the following definition of matrix mean.

4.1 Matrix mean

A *matrix mean* is a map $M : \mathbb{P}_n \times \mathbb{P}_n \mapsto \mathbb{P}_n$ that satisfies the following conditions:

1. $A \leq B$ implies $A \leq M(A, B) \leq B$.

2. (transformation property) $M(X^*AX, X^*BX) = X^*M(A, B)X$, for $A, B \in \mathbb{P}_n$ and a nonsingular matrix X.

3. (symmetry property) $M(A, B) = M(B, A)$.

4. (monotonicity property) if $A_1 \leq A_2$, then $M(A_1, B) \leq M(A_2, B)$.

5. $M(\cdot, \cdot)$ is continuous in each argument.

Recall that the geometric mean of $A, B \in \mathbb{P}_n$ is

$$A\#B := A\#_{\frac{1}{2}}B = A^{1/2}(A^{-1/2}BA^{-1/2})^{1/2}A^{1/2},$$

which Pusz and Woronowicz formulated it in [32] and characterized some important properties of (2.0.1). One can verify that the matrix geometric mean satisfies the above matrix mean conditions. In addition, for basic properties of a geometric mean one can see Ando [1]. The following theorem and the important properties of Γ_X also enable us to extend the definition of geometric mean from $\mathbb{R}_+$ to $\mathbb{P}_n$.

Theorem 4.1.1. *([20, Corollary 7.6.5])*
Let $A, B \in \mathbb{P}_n$. There exists a nonsingular matrix X such that $\Gamma_X(A) = I$, and $\Gamma_X(B) = D_B$, where D_B is a diagonal matrix.

Therefore, by Theorem 4.1.1, we have

$$\begin{aligned}
A\#B &= (X^*)^{-1}IX^{-1}\#(X^*)^{-1}D_BX^{-1} \\
&= (X^*)^{-1}I_AX^{-1}\#(X^*)^{-1}D_BX^{-1} \\
&= (X^*)^{-1}(I\#D_B)X^{-1} \\
&= (X^*)^{-1}(I\#D_B)X^{-1} \\
&= (X^*)^{-1}(D_B)^{1/2}X^{-1}. \qquad (4.1.1)
\end{aligned}$$

Now, let $X = A^{-1/2}U$. We have $\Gamma_X(B) = U^*(A^{-1/2}BA^{-1/2})U = D_A$, $\Gamma_X(A) = I$. By (4.1.1) and the fact that $D_A^{1/2} = U^*(B^{-1/2}AB^{-1/2})^{1/2}U$, we have the geometric mean of two matrices as

$$A\#B = A^{1/2}\left(A^{-1/2}BA^{-1/2}\right)^{1/2}A^{1/2} = \Gamma_{A^{1/2}}\Big((\Gamma_{A^{-1/2}}(B))^{1/2}\Big). \qquad (4.1.2)$$

There is a nonsingular matrix S that simultaneously diagonalizes A and B by congruence,

$$A = \Gamma_S(D_A), \quad B = \Gamma_S(D_B).$$

Therefore,

$$A\#B = \Gamma_S(D_A D_B)^{1/2}$$

and it is independent of the choice of S. Since computing $(D_A D_B)^{1/2}$ is easy, it enables us to compute $A\#B$. Some important properties of $A\#B$ are listed:

1. $A\#B = B\#A$.

2. $A\#B$ is the unique positive solution of the Riccati equation $XA^{-1}X = B$.

3. $A\#B$ has an extremal property

$$A\#B = \max \left\{ X : X = X^*, \begin{bmatrix} A & X \\ X^* & B \end{bmatrix} \geq 0 \right\}. \tag{4.1.3}$$

4. There exists a unitary matrix U such that $A\#B = A^{1/2}UB^{1/2}$. Since U is unitary, $A^{1/2}UB^{1/2}$ is positive definite and as a result $A\#B$ is positive definite.

5. $A\#B$ has the transformation property, i.e., $\Gamma_X(A)\#\Gamma_X(B) = \Gamma_X(A\#B)$.

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

7. When $A, B \in \mathbb{C}_{2\times 2}$, $A\#B = \frac{A+B}{\sqrt{\det(A+B)}}$.

8. For any $\alpha \geq 0$, $(\alpha A)\#B = \sqrt{\alpha}(A\#B)$.

9. For $A_1, A_2, B_1, B_2 \in \mathbb{P}_n$, if $A_1 \leq B_1$ and $A_2 \leq B_2$, then $A_1\#A_2 \leq B_1\#B_2$.

10. $\operatorname{tr}(A\#B) \leq \operatorname{tr}(A^{1/2}B^{1/2})$.

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

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

A map $\Phi : \mathbb{P}_n \times \cdots \times \mathbb{P}_n \to \mathbb{P}_n$ is called *convex* if

$$\Phi(\lambda A_1 + (1-\lambda)B_1, \ldots, \lambda A_k + (1-\lambda)B_k) \leq \lambda\Phi(A_1, \ldots, A_k) + (1-\lambda)\Phi(B_1, \ldots, B_k).$$

Theorem 4.1.2. *([6, Theorem 1.3.3])*
The block matrix $M = \begin{bmatrix} A & C \\ C^* & B \end{bmatrix}$, *where* $A, B \in \mathbb{P}_n$, *and* $C \in \mathbb{C}_{n\times n}$, *is positive semidefinite if and only if* $B \geq C^*A^{-1}C$.

Theorem 4.1.3. *(Pusz and Woronowicz [32])*
Let $A, B \in \mathbb{P}_n$. *The map* $(A, B) \to A\#B$ *is concave.*

Proof. Let $C = \begin{bmatrix} A & A\#B \\ A\#B & B \end{bmatrix}$. Since

$$(A\#B)A^{-1}(A\#B) = A^{1/2}(A^{-1/2}B^{1/2}A^{-1/2})A^{1/2} = B, \tag{4.1.4}$$

by Theorem 4.1.2, C is positive semidefinite. Let $C_1 = \begin{bmatrix} A_1 & A_1\#B_1 \\ A_1\#B_1 & B_1 \end{bmatrix}$ and $C_2 = \begin{bmatrix} A_2 & A_2\#B_2 \\ A_2\#B_2 & B_2 \end{bmatrix}$, where A_1, A_2, B_1 and B_2 are positive semidefinite. We have

$$\lambda C_1 + (1-\lambda)C_2 \geq 0.$$

i.e.,

$$\begin{bmatrix} \lambda A_1 + (1-\lambda)A_2 & \lambda(A_1\#B_1) + (1-\lambda)A_2\#B_2) \\ \lambda(A_1\#B_1) + (1-\lambda)(A_2\#B_2) & \lambda B_1 + (1-\lambda)B_2 \end{bmatrix}.$$

Therefore, by (4.1.3), we have

$$(\lambda A_1 + (1-\lambda)A_2)\#(\lambda B_1 + (1-\lambda)B_2) \geq \lambda(A_1\#B_1) + (1-\lambda)A_2\#B_2).$$

$\square$

As a consequence, for $A_i, B_i,\ i = 1, \ldots, n$, we have

$$\sum_{i=1}^{n} A_i \# B_i \leq \sum_{i=1}^{n} A_i \# \sum_{i=1}^{n} B_i. \tag{4.1.5}$$

4.2 Matrix geometric mean inequalities of Lemos and Soares

Very recently Lemos and Soares [24] obtained log-majorization relations that involve matrix geometric mean.

Theorem 4.2.1. *(Lemos and Soares [24, p. 26])*
Let $A, B \geq 0$. We have

$$\lambda(A(A\#B)B(A\#B)) \prec_{\log} \lambda(A^2B^2). \tag{4.2.1}$$

Note that the matrices $A(A\#B)B(A\#B)$ and A^2B^2 on both sides are not necessarily positive semidefinite. Since the spectra of AB and BA coincide, we have

$$\lambda(A(A\#B)B(A\#B)) = \lambda(A^{1/2}(A\#B)B^{1/2}B^{1/2}(A\#B)A^{1/2}) \tag{4.2.2}$$

$$= \lambda^2(|A^{1/2}(A\#B)B^{1/2}|), \tag{4.2.3}$$

and

$$\lambda(A^2B^2) = \lambda((AB)(BA)) = \lambda^2(|AB|).$$

As a result, we have the inequality

$$\lambda(|A^{1/2}(A\#B)B^{1/2}|) \prec_{\log} \lambda(|AB|). \tag{4.2.4}$$

They also asked if the following log-majorization holds:

Question:

$$\lambda(|A^t(A\#_tB)B^{1-t}|) \prec_{\log} \lambda(|AB|). \tag{4.2.5}$$

Lemos and Soares also proved the following theorem.

Theorem 4.2.2. *(Lemos and Soares [24, p. 25])*
Let $A, B, X \geq 0$ and $t \in [0,1]$. We have

$$\lambda((A\#_t B)X(A\#_{1-t}B)X) \prec_{\log} \lambda(AXBX). \tag{4.2.6}$$

For $X = I$ and $t = 1/2$, we have

$$\lambda((A\#B)^2) \prec_{\log} \lambda(AB). \tag{4.2.7}$$

Indeed, by using the transformation property of t-geometric mean we can obtain the following log-majorization inequality from Theorem 4.2.2, but we would like to prove it by using their techniques.

Theorem 4.2.3. *Let $A, B \geq 0$. For any $X \in \mathrm{GL}_n(\mathbb{C})$ and $t \in [0,1]$ we have*

$$\lambda_1(\Gamma_X(A\#_t B)\Gamma_X(A\#_{1-t}B)) \prec_{\log} \lambda_1(\Gamma_X(A)\Gamma_X(B)). \tag{4.2.8}$$

Proof. Since the k-th compound matrix is multiplicative and respects complex conjugate transpose, we have

$$\begin{aligned} & C_k(X^*(A\#_t B)XX^*(A\#_{1-t}B)X) \\ = \;& C_k(X^*)C_k(A\#_t B)C_k(X)C_k(X^*)C_k(A\#_{1-t}B)C_k(X) \\ = \;& \Gamma_{C_k(X)}(C_k(A)\#_t C_k(B))\Gamma_{C_k(X)}(C_k(A)\#_{1-t}C_k(B)), \end{aligned}$$

and

$$\begin{aligned} C_k(X^*AXX^*BX) &= (C_k(X))^*C_k(A)(C_k(X))(C_k(X))^*C_k(B)C_k(X) \\ &= \Gamma_{C_k(X)}C_k(A)\Gamma_{C_k(X)}C_k(B). \end{aligned}$$

Moreover, $\det \Gamma_X(A\#_t B)\Gamma_X(A\#_{1-t}B) = \det \Gamma_X(A)\Gamma_X(B)$. Therefore, it suffices to show that

$$\lambda_1(\Gamma_X(A\#_t B)(A\#_{1-t}B)X) \le \lambda_1(\Gamma_X(A)\Gamma_X(B)). \tag{4.2.9}$$

Let $X = I$. From $\lambda_1(AB) \le 1$ we have $AB \le I$. As a result, $A \le B^{-1}$ and $B \le A^{-1}$. Thus, by joint monotonicity of matrix geometric mean we have

$$A\#_t B \le B^{-1}\#_t A^{-1} = (A\#_{1-t}B)^{-1}. \tag{4.2.10}$$

Therefore, we have

$$\lambda_1((A\#_t B)(A\#_{1-t}B)) \le \lambda_1(AB). \tag{4.2.11}$$

For any $X \in \mathbb{C}_{n\times n}$ we have

$$X^*(A\#_t B)X \le X^*AX\#_t X^*BX$$

and the equality holds when X is nonsingular. Now, Let $X \in \mathrm{GL}_n(\mathbb{C})$. We have

$$\begin{aligned} & \lambda_1(X^*(A\#_t B)XX^*(A\#_{1-t}B)X) \\ = \ & \lambda_1((X^*AX\#_t X^*BX)(X^*AX\#_{1-t}X^*BX)) \\ \le \ & \lambda_1((X^*AX)(X^*BX)). \end{aligned}$$

□

It can be seen that when $X \ge 0$ by choosing $X = C$, we have $C^* = C$ and

$$\lambda(C^2(A\#_t B)C^2(A\#_{1-t}B)) \prec_{\log} \lambda(C^2AC^2B), \tag{4.2.12}$$

Furthermore, by choosing $X = C^{1/2}$ we can see

$$\lambda(C(A\#_t B)C(A\#_{1-t}B)) \prec_{\log} \lambda(CACB), \tag{4.2.13}$$

in which is the same as Theorem 4.2.2.

4.3 More inequalities for t-geometric mean

In the previous section, we saw the log-majorization

$$\lambda((A\#_t B)(A\#_{1-t}B)) \prec_{\log} \lambda(AB),$$

for $A, B \in \mathbb{P}_n$. Geometrically, if we travel on the geodesic $\gamma(t)$, joining A and B, from point A towards B and in the meanwhile from B to A, then we will have two points at time t, namely, $A\#_t B$ and $B\#_t A$. The above log-majorization gives a relationship between the product of these two points and product of the original two points A and B. Now, let us

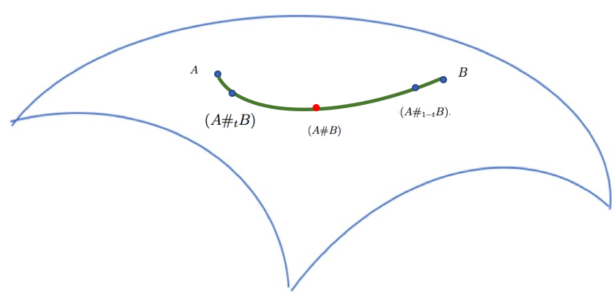

Figure 4.1: We travel from A and B towards the midpoint $A\#B$.

consider the endpoints $A\#_t B$ and $A\#_{1-t}B$ instead of endpoints A and B. Thus the midpoint $A\#B$ remains the midpoint of $A\#_t B$ and $A\#_{1-t}B$. Hence, we use the inequality for new endpoints $A\#_t B$ and $A\#_{1-t}B$ and midpoint $A\#B$. Therefore, for any two positive definite

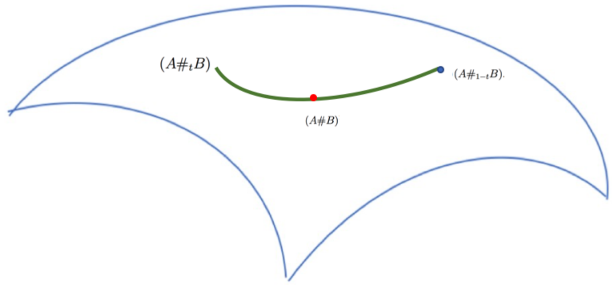

Figure 4.2: The midpoint of the geodesic joining $A\#_t B$ and $A\#_{1-t} B$ is still $A\#B$.

matrices and any $t \in [0, 1]$, we have the following log-majorization inequality

$$\lambda((A\#B)^2) \prec_{\log} \lambda((A\#_t B)(A\#_{1-t} B)). \tag{4.3.1}$$

This shows the beauty of geometric perspective because proving the inequality (4.3.1) is not easy. On the other hand, we have the following theorem of Ando and Hiai.

Theorem 4.3.1. *(Ando and Hiai [2, Theorem 2.1])*
For $A, B \in \mathbb{P}_n$ and $t \in [0, 1]$, we have

$$\lambda(A^r \#_t B^r) \quad \prec_{\log} \quad \lambda((A\#_t B)^r), \quad r \geq 1, \tag{4.3.2}$$

$$\lambda((A\#_t B)^r) \quad \prec_{\log} \quad \lambda(A^r \#_t B^r), \quad 0 \leq r \leq 1. \tag{4.3.3}$$

As a result, by (4.3.1) and (4.3.2) for $r = 2$ we have the following log-majorization inequality:

$$\lambda(A^2 \# B^2) \prec_{\log} \lambda((A\#_t B)(A\#_{1-t} B)). \tag{4.3.4}$$

Furthermore, by the above discussions and considering the endpoints $A^{1/2}$ and $B^{1/2}$ we have

$$\lambda(A\#B) \prec_{\log} \lambda((A^{1/2} \# B^{1/2})^2) \prec_{\log} \lambda(A^{1/4} B^{1/2} A^{1/4}),$$

which is another way to approach the inequality in ([8, Theorem 3]) for $t = 1/2$.

Theorem 4.3.2. *(Order preserving inequality, Furuta's inequality)*
If $X \geq Y \geq 0$, then for each $r \geq 0$ we have

$$(Y^r X^p Y^r)^{(1/q)} \geq Y^{(p+2r)/q}, \tag{4.3.5}$$

and

$$X^{(p+2r)/q} \geq (X^r Y^p X^r)^{1/q}, \tag{4.3.6}$$

for all $p \geq 0$ and $q \geq 1$ and $(1+2r)q \geq p+2r$.

Theorem 4.3.3. *[14, p. 324]*
Let $A, B \geq 0$ and $t \in [0,1]$. We have

$$\lambda(A^{1/2}(A\#_t B)A^{1/2}) \prec_{\log} \lambda(A^{1-t/2}B^t A^{1-t/2}). \tag{4.3.7}$$

Proof. It can be seen that

$$C_k(A^{1/2}(A\#_t B)A^{1/2}) = (C_k(A))^{1/2}(C_k(A)\#_t C_k(B))(C_k(A))^{1/2},$$

and

$$C_k(A^{1-t/2}B^t A^{1-t/2}) = (C_k(A))^{1-t/2})(C_k(B))^t(C_k(A))^{1-t/2}.$$

Moreover, $\det(A^{1/2}(A\#_t B)A^{1/2}) = \det(A^{1-t/2}B^t A^{1-t/2})$. It sufices to show that

$$\lambda_1(A^{1/2}(A\#_t B)A^{1/2}) \leq \lambda_1(A^{1-t/2}B^t A^{1-t/2}). \tag{4.3.8}$$

Thus we need to prove the following inequality

$$A^{1-t/2}B^t A^{1-t/2} \leq I \quad \text{implies} \quad A^{1/2}(A\#_t B)A^{1/2} \leq I. \tag{4.3.9}$$

Let $A^{1-t/2}B^tA^{1-t/2} \leq I$, or equivalently $B^t \leq A^{t-2}$. By using the Furuta Inequality (4.3.6), for $X = A^{t-2}, Y = B^t, q = p = t^{-1}$ and $r = 1/2(2-t)$, we have $p + 2r = \frac{2}{t(2-t)}$ and

$$\left((A^{t-2})^{1/(2(2-t))}(B^t)^{t^{-1}}(A^{t-2})^{1/(2(2-t))}\right)^{\frac{1}{t^{-1}}} \leq ((A^{t-2})^{2/t(2-t)})^t, \qquad (4.3.10)$$

which implies $A^{1/2}(A\#_tB)A^{1/2} \leq I$. □

Since the log-majorization implies weak majorization and for any $A \in \mathbb{C}_{n\times n}$, $\sum_{i=1}^n \lambda_i(A) = \operatorname{tr} A$, the above log-majorization inequalities and the property $\operatorname{tr}(AB) = \operatorname{tr}(BA)$ imply the following trace inequalities

$$\operatorname{tr}(A(A\#_tB)) \leq \operatorname{tr}(A^{2-t}B^t), \quad [10] \qquad (4.3.11)$$

$$\operatorname{tr}(A^2\#B^2) \leq \operatorname{tr}((A\#_tB)(A\#_{1-t}B)), \qquad (4.3.12)$$

$$\operatorname{tr}((A\#_tB)(A\#_{1-t}B)) \leq \operatorname{tr}(AB). \quad [10] \qquad (4.3.13)$$

Chapter 5

On the norm inequality of t-geometric mean of matrices

In this chapter we will discuss motivations of the questions which lead us to the conjecture in Dinh, Ahsani, Tam [15]. Furthermore, we will bring some extensions of the related results in [13] and a solution for the conjecture under a special condition.

Theorem 5.0.1. *(Araki [4], [6, Theorem IV.2.10])*
Let $A, B \in \mathbb{P}_n$. For all unitary invariant norms $|||\,.|||$ we have

$$|||\, B^t A^t B^t ||| \leq |||\, (BAB)^t |||\,, \quad 0 \leq t \leq 1 \tag{5.0.1}$$

$$|||\, (BAB)^t ||| \leq |||\, B^t A^t B^t |||\,, \quad t \geq 1. \tag{5.0.2}$$

5.1 Motivation

In 1998, Bhatia and Kittaneh in [9] proved the following matrix subadditivity inequality

$$|||\, A^m + B^m ||| \leq |||\, (A+B)^m |||\,,$$

where $A, B \geq 0$, and m is an integer. They also showed the analogus inequalities when m is replaced by a positive real number r

$$|||\, \mathrm{A^r} + \mathrm{B^r} ||| \leq |||\, (A+B)^r |||\,, \quad 1 \leq r \leq \infty, \tag{5.1.1}$$

$$|||\, A^r + B^r ||| \geq |||\, (A+B)^r |||\,, \quad 0 \leq r \leq 1. \tag{5.1.2}$$

Later in 1999, Ando and Zhan proved a subadditivity inequality for operator concave functions [3]. In 2007, Bourin and Uchiyama extended their work to all concave functions [12].

Theorem 5.1.1. *([12, Theorem 1.1])*

Let A and B be positive semidefinite matrices and let $|||\,.\,|||$ be any unitary invariant norm. We have the following assertions

(I) *For a concave function $f : [0, \infty) \to [0, \infty)$ we have*

$$|||\, f(A + B)|||\ \leq |||\, f(A) + f(B)|||\,, \tag{5.1.3}$$

which can be stated for a family of positive definite matrices $\{A_i\}_{i=1}^{m}$:

$$|||\, f(A_1 + \cdots + A_m)|||\ \leq |||\, f(A_1) + \cdots + f(A_m)|||\,. \tag{5.1.4}$$

(II) For a convex function $f : [0, \infty) \to [0, \infty)$ with $g(0) = 0$, we have

$$|||\, f(A) + f(B)|||\ \leq |||\, f(A + B)|||\,, \tag{5.1.5}$$

$$|||\, f(A_1) + \cdots + f(A_m)|||\ \leq |||\, f(A_1 + \cdots + A_m)|||\,. \tag{5.1.6}$$

These inequalities were starting points for matrices subadditivity inequalities and raised some related questions by Bourin in 2009 [11]:

Question (I) (Bourin-2009):

Given $A, B \geq 0$ and $p, q \geq 0$, does the following inequality hold or not?

$$|||\, A^{p+q} + B^{p+q}|||\ \leq |||\, (A^p + B^p)(A^q + B^q)|||\,? \tag{5.1.7}$$

In addition, does the following inequality hold in general?

$$|||\, A^p + B^q|||\ \leq |||\, (A^{p_1} + B^{q_1}) \ldots (A^{P_n} + B^{q_n})|||\,?$$

where $p = \sum_{i=1}^n p_i$ and $q = \sum_{i=1}^n q_i$ with $p_i, q_i \geq 0, \quad i = 1, 2, \dots n.$

Furthermore, he asked the following question for the case $n = 2$ and $p_1 = q_1 = p$ and $p_2 = q_2 = q$:

$$|||\, A^{p+q} + B^{p+q} |||\, \leq |||\, (A^p + B^p)^{\frac{1}{2}} (A^q + B^q)(A^p + B^p)^{\frac{1}{2}} |||\, ?$$

Question (II) (Bourin-2009):

Given $A, B \geq 0$ and $p, q \geq 0$. Does the following inequality hold or not?

$$|||\, A^p B^q + B^p A^q |||\, \leq |||\, A^{p+q} + B^{p+q} |||\, .$$

In 2013 Hayanjeh and Kittaneh [18] gave an affirmative answer to this question for the Frobenius norm and the trace norms and conjectured that the above inequalities are true for all unitary invariant norms and for all commuting positive operators.

Conjecture 5.1.2. *([18], Hayanjeh and Kittaneh-2013)*

Given $A_i, B_i \geq 0,\ i = 1, 2$ and $A_i B_i = B_i A_i$. The following inequalities are true for all unitary invariant norms.

$$|||\, A_1 B_1 + A_2 B_2 |||\, \leq |||\, (A_1 + A_2)(B_1 + B_2) |||\, ,$$

$$|||\, A_1 B_1 + A_2 B_2 |||\, \leq |||\, (A_1 + A_2)^{\frac{1}{2}} (B_1 + B_2)(A_1 + A_2)^{\frac{1}{2}} |||\, .$$

Finally in 2015, Koenraad Audenaert in ([5, Theorem 1.3]) proved the general case for positive semidefinite matrices A_i and B_i, $i = 1, \cdots, m$, such that, for each i, A_i commuting with B_i.

$$|||\, \sum_{i=1}^m A_i B_i |||\, \leq |||\, (\sum_{i=1}^m A_i^{1/2} B_i^{1/2})^2 |||\, \leq |||\, \big(\sum_{i=1}^m A_i\big)\big(\sum_{i=1}^m B_i\big) |||\, . \tag{5.1.8}$$

As a consequence, it can be seen that for any $f, g : [0,\infty) \to [0,\infty)$ the inequality

$$||| \sum_{i=1}^{m} f(A_i)g(A_i)||| \leq ||| \big(\sum_{i=1}^{m} f(A_i)\big)\big(\sum_{i=1}^{m} g(A_i)\big)||| \,,$$

holds for all unitary invariant norms. By choosing $f(x) = x^{p_1}$, $g(x) = x^{p_2}$ and $m = 2$, Audenaert gave an affirmative answer to the question of Bourin (5.1.7). Recently, in 2017, Hayajneh et al., [17] proved a sharper inequality than (5.1.8)

$$||| (\sum_{i=1}^{m} A_i^{1/2} B_i^{1/2})^2 ||| \leq \big(\sum_{i=1}^{m} A_i\big)^{1/2}\big(\sum_{i=1}^{m} B_i\big)\big(\sum_{i=1}^{m} A_i\big)^{1/2} ||| \,, \tag{5.1.9}$$

and as a result,

$$||| \sum_{i=1}^{m} A_i B_i ||| \leq ||| \big(\sum_{i=1}^{m} A_i\big)^{1/2}\big(\sum_{i=1}^{m} B_i\big)\big(\sum_{i=1}^{m} A_i\big)^{1/2} ||| \,. \tag{5.1.10}$$

Now, regarding the inequality (5.1.8), we investigate that does a similar inequality hold in the noncommuting case when matrix multiplication is replaced with the geometric mean? Symmetric properties of geometric mean, $A\#B = B\#A$ and $A^2\#B^2 = (A\#B)^2 = AB$, for commuting matrices A and B beside the properties (5.2.4) and (5.2.5) motivate our work. In this regard, we have some questions in [15].

Question (I):

$$||| \sum_{i=1}^{m} A_i^2 \# B_i^2 ||| \leq ||| \sum_{i=1}^{m} A_i B_i ||| \,, \tag{5.1.11}$$

and

$$||| \sum_{i=1}^{m} (A_i \# B_i)^2 ||| \leq ||| \sum_{i=1}^{m} A_i B_i ||| \,, \tag{5.1.12}$$

where $A_i, B_i \in \mathbb{P}_n, i = 1, \ldots, m$. By creating counter examples we showed in [15] that these inequalities are not true, while the computational results show the following inequalities are true.

Question (II):

$$|||\sum_{i=1}^{m}(A_i\#B_i)^2||| \le |||\Big(\sum_{i=1}^{m}A_i\Big)\Big(\sum_{i=1}^{m}B_i\Big)|||, \tag{5.1.13}$$

$$|||\sum_{i=1}^{m}A_i^2\#B_i^2||| \le |||\Big(\sum_{i=1}^{m}A_i\Big)\Big(\sum_{i=1}^{m}B_i\Big)|||, \tag{5.1.14}$$

where $A_i, B_i \in \mathbb{P}_n, i = 1, \dots, m$.

The inequality (5.1.13) was proved in [15] and was extended to the general case in [13] as the following, but the inequality (5.1.14) was posted in [15] as a conjecture.

Theorem 5.1.3. *[14, Theorem 2.2]*

Let $A_i, B_i \in \mathbb{P}_n, i = 1, \cdots, m, p > 0$ and $r \ge 1$. We have

$$|||(\sum_{i=1}^{m}(A_i\#_tB_i)^r||| \le |||\Big(\Big(\sum_{i=1}^{m}A_i\Big)^{rtp/2}\Big(\sum_{i=1}^{m}B_i\Big)^{(1-t)pr}\Big(\sum_{i=1}^{m}A_i\Big)^{rtp/2}\Big)^{\frac{1}{p}}|||. \tag{5.1.15}$$

For $t = 1/2, r = 2$ and $p = 1$ the above inequality becomes

$$|||(\sum_{i=1}^{m}(A_i\#B_i)^2||| \le |||\Big(\sum_{i=1}^{m}A_i\Big)^{1/2}\Big(\sum_{i=1}^{m}B_i\Big)\Big(\sum_{i=1}^{m}A_i\Big)^{1/2}|||, \tag{5.1.16}$$

which is sharper than the Inequality (5.1.13) in Question(II).

Proof. For $A_i, B_i \in \mathbb{P}_n, i = 1, \dots, n$, we have,

$$\begin{aligned}
|||\sum_{i=1}^{m}A_i\#_tB_i||| &\le |||\sum_{i=1}^{m}A_i\#_t\sum_{i=1}^{m}B_i|||, \qquad \text{concavity of } t\text{-geometric mean}\\
&\le |||\Big(\Big(\sum_{i=1}^{m}B_i\Big)^{\frac{tp}{2}}\Big(\sum_{i=1}^{m}A_i\Big)^{(1-t)p}\Big(\sum_{i=1}^{m}B_i\Big)^{\frac{tp}{2}}\Big)^{\frac{1}{p}}|||, \quad \text{by (3.3.2)}\\
&\le |||\Big(\Big(\sum_{i=1}^{m}A_i\Big)^{(1-t)p}\Big(\sum_{i=1}^{m}B_i\Big)^{tp}\Big)^{\frac{1}{p}}|||.
\end{aligned}$$

As a consequence, it can be seen that for $r \geq 1$, we have

$$\begin{aligned}
|||\sum_{i=1}^{m}(A_i\#_t B_i)^r||| &\leq |||\Big(\sum_{i=1}^{m}A_i\#_t B_i\Big)^r||| \\
&\leq |||\Big(\sum_{i=1}^{m}A_i\#_t \sum_{i=1}^{m}B_i\Big)^r||| \\
&\leq |||\Big(\Big(\sum_{i=1}^{m}B_i\Big)^{\frac{tp}{2}}\Big(\sum_{i=1}^{m}A_i\Big)^{(1-t)p}\Big(\sum_{i=1}^{m}B_i\Big)^{\frac{tp}{2}}\Big)^{\frac{r}{p}}||| \\
&\leq |||\Big(\Big(\sum_{i=1}^{m}B_i\Big)^{\frac{rtp}{2}}\Big(\sum_{i=1}^{m}A_i\Big)^{r(1-t)p}\Big(\sum_{i=1}^{m}B_i\Big)^{\frac{rtp}{2}}\Big)^{\frac{1}{p}}|||. \quad \text{by (5.0.2)}
\end{aligned}$$

□

Moreover, the left hand side of the inequalities (5.1.14) and (5.1.12) is not comparable in general. See [15] for presented examples. In addition, numerical calculation supports the following inequality.

Conjecture 5.1.4. Let $A_i, B_i \in \mathbb{P}_n, i = 1, 2, \cdots, m$. We have

$$||\sum_{i=1}^{m} A_i^2\#B_i^2||_2 \leq ||\sum_{i=1}^{m}(A_i\#B_i)^2||_2. \tag{5.1.17}$$

Remark: Let $A_i, B_i \in \mathbb{P}_n$, $i = 1, \cdots, m$. The inequality

$$|||\sum_{i=1}^{m} A_i^2\#\sum_{i=1}^{m} B_i^2||| \leq |||(\sum_{i=1}^{m} A_i)^2\#(\sum_{i=1}^{m} B_i)^2||| \tag{5.1.18}$$

is not true in general. While, numerical computation shows that the inequality holds for the trace norm and the Frobenius norm.

Remark 5.1.5. $\mathbb{P}_n$ is a Riemannian manifold, but as a subset of $\mathrm{GL}_n(\mathbb{C})$, it does not form a subgroup. As we mentioned, the reason is that the product of two Hermitian matrices A and B is Hermitian if and only if $AB = BA$. To be more precise, for $A, B \in \mathbb{P}_n$, we have

$AB = B^{-1/2}(B^{1/2}AB^{1/2})B^{1/2}$, then

$$AB \sim B^{1/2}AB^{1/2}.$$

Moreover, we have $\lambda(B^{1/2}AB^{1/2}) = \langle B^{1/2}AB^{1/2}X, X\rangle = \langle AB^{1/2}X, B^{1/2}X\rangle \geq 0, X \neq 0$. Therefore, $B^{1/2}AB^{1/2}$ has positive eigenvalues. Furthermore, $(B^{1/2}AB^{1/2})^* = B^{1/2}AB^{1/2}$, and by [20, Theorem 2.5.6, p.135] is diagonalizable. As a result, AB is similar to $B^{1/2}AB^{1/2}$ which has positive eigenvalues and diagonalizable. Also, see [20, Corollary 7.6.2, p.486] for more details. Hence, the eigenvalues of AB coincide with the singular values of a positive definite matrix. In other words, the set

$$\mathbb{P}_n^2 = \{AB \mid A, B \in \mathbb{P}_n\}$$

consists of those matrices that are diagonalizable and have positive eigenvalues. As a consequence, under the condition of commuting A_i and commuting B_i, we have

Proposition 5.1.6. Let $A_i, B_i \in \mathbb{P}_n$, $i = 1, \cdots, m$, such that $A_iA_j = A_jA_i$ for $i, j = 1, \cdots, m$ and $B_iB_j = B_jB_i$ for $i, j = 1, \cdots, m$. For any unitary invariant norm $|||\ .\ |||$ on $\mathbb{C}_{n\times n}$,

$$|||\sum_{i=1}^{m} A_i^2 \# \sum_{i=1}^{m} B_i^2||| \leq |||(\sum_{i=1}^{m} A_i)^2 \# (\sum_{i=1}^{m} B_i)^2|||. \tag{5.1.19}$$

Proof. Let $A_iA_j = A_jA_i$ and $B_iB_j = B_jB_i$ for $i, j = 1, \cdots, m$.

We have,

$$\sum_{\mathrm{i=1}}^{\mathrm{m}} \mathrm{A_i^2} \leq (\sum_{i=1}^{m} A_i)^2, \tag{5.1.20}$$

$$\sum_{i=1}^{m} B_i^2 \leq (\sum_{i=1}^{m} B_i)^2. \tag{5.1.21}$$

The geometric mean of A and B, $A\#B$, is a non-decreasing function of its arguments, i.e., if $A_1 \leq A_2$, then $A_1\#B \leq A_2\#B$, [16, Theorem 2.2]. Therefore, if $A_1 \leq A_2$ and $B_1 \leq B_2$, then, $A_1\#B_1 \leq A_2\#B_2$. Thus, (5.1.19) follows from this fact. □

Theorem 5.1.7. Let $A_i, B_i \in \mathbb{P}_n$, $i = 1, \cdots, m$, which $A_iA_j = A_jA_i$ and $B_iB_j = B_jB_i$ for $i, j = 1, \cdots, m$. For any unitary invariant norm $|||\ .\ |||$ on $\mathbb{C}_{n\times n}$ the following inequalities are true:

$$||| \sum_{i=1}^{m} A_i^2 \# B_i^2 ||| \leq ||| (\sum_{i=1}^{m} A_i)^{\frac{1}{2}} (\sum_{i=1}^{m} B_i)(\sum_{i=1}^{m} B_i)^{\frac{1}{2}} ||| . \tag{5.1.22}$$

Proof. The geometric mean of A and B, $A\#B$ is a concave function. Therefore,

$$\begin{aligned} ||| \sum_{i=1}^{m} A_i^2 \# B_i^2 ||| &\leq ||| \sum_{i=1}^{m} A_i^2 \# \sum_{i=1}^{m} B_i^2 ||| \\ &\leq ||| (\sum_{i=1}^{m} A_i)^2 \# (\sum_{i=1}^{m} B_i)^2 ||| \\ &\leq ||| (\sum_{i=1}^{m} A_i)^{\frac{1}{2}} (\sum_{i=1}^{m} B_i)(\sum_{i=1}^{m} B_i)^{\frac{1}{2}} ||| , \end{aligned}$$

where the second inequality is true due to the Proposition 5.1.6 and the third inequality is due to [8, Inequality number (18)]. □

5.2 More norm inequalities of t-geometric mean of matrices

Recall that Bhatia and Grover ([8, Theorem 3]) showed that the inequality

$$||| A\#_t B ||| \leq ||| B^{\frac{t}{2}} A^{1-t} B^{\frac{t}{2}} ||| \leq ||| A^{1-t} B^t ||| , \tag{5.2.1}$$

holds for $A, B > 0$ and all unitary invariant norms $|||\,.\,|||$. Also, recall that from (3.3.3) we have

$$||| A\#_t B ||| \leq ||| (B^{\frac{tp}{2}} A^{(1-t)p} B^{\frac{tp}{2}})^{\frac{1}{p}} ||| \leq ||| (B^{tp} A^{(1-t)p})^{\frac{1}{p}} ||| , \tag{5.2.2}$$

for the special cases $t = 1/2$ and $p = 1$, we have

$$|||\, A\#B|||\ \leq |||\, (B^{\frac{1}{4}}A^{\frac{1}{2}}B^{\frac{1}{4}})|||\ \leq |||\, A^{\frac{1}{2}}B^{\frac{1}{2}}|||\,, \tag{5.2.3}$$

As a result,

$$|||\, A^2\#B^2|||\quad \leq \quad |||\, B^{\frac{1}{2}}AB^{\frac{1}{2}}|||\ \leq |||\, AB|||\,, \tag{5.2.4}$$

and

$$\begin{aligned} |||\, (A\#B)^2||| \quad &\leq \quad |||\, (B^{\frac{1}{4}}A^{\frac{1}{2}}B^{\frac{1}{4}})^2||| \\ &\leq \quad |||\, B^{\frac{1}{2}}AB^{\frac{1}{2}}|||\ \leq |||\, AB||| \quad \text{by Theorem 5.0.1} \end{aligned} \tag{5.2.5}$$

Another way to approach the last norm inequality is using the recent result of Lemos and Soares (4.2.7). Since $(A\#B) \in \mathbb{P}_n$, we have

$$\begin{aligned} s((A\#B)^2) \quad &= \quad \lambda((A\#B)^2) \\ &\prec_{\log} \quad \lambda(AB) \\ &\prec_{\log} \quad s(AB). \end{aligned}$$

Thus, we have the inequality $|||\, (A\#B)^2|||\ \leq |||\, AB|||\,.$

Similarly, by (4.3.3) and (4.3.1) and using the fact that $A^{1/2}(A\#_t B)A^{1/2}$, $A^2\#B^2 \in \mathbb{P}_n$, we have

$$|||\, A^{1/2}(A\#_t B)A^{1/2}|||\ \leq |||\, A^{2-t}B^t|||\,. \tag{5.2.6}$$

and

$$|||\, A^2\#B^2|||\ \leq |||\, (A\#_t B)(A\#_{1-t}B)|||\,, \tag{5.2.7}$$

and the following theorem

Theorem 5.2.1. *Let $A_i, B_i \in \mathbb{P}_n, i = 1, 2, \ldots, n$. We have*

$$|||\sum_{i=1}^{n}(A_i\#B_i)^2||| \le |||(\sum_{i=1}^{n}A_i\#_t\sum_{i=1}^{n}B_i)(\sum_{i=1}^{n}B_i\#_t\sum_{i=1}^{n}A_i)|||. \tag{5.2.8}$$

Proof. Since $(A_i\#B_i)^2 \in \mathbb{P}_n, i = 1, 2, \ldots, n$, by using (4.3.1) and the above techniques for each $i = 1, 2, \ldots, n$, we have

$$|||(A_i\#B_i)^2||| \le |||(A_i\#_tB_i)(A_i\#_{1-t}B_i)|||. \tag{5.2.9}$$

Now,

$$\begin{aligned}
|||\sum_{i=1}^{n}(A_i\#B_i)^2||| &\le |||\Big(\sum_{i=1}^{n}(A_i\#B_i)\Big)^2|||, \quad \text{apply (5.1.6) for the convex function } t^2\\
&\le |||\Big(\sum_{i=1}^{n}A_i\#\sum_{i=1}^{n}B_i\Big)^2|||, \quad \text{concavity of } t\text{-geometric mean}\\
&\le |||(\sum_{i=1}^{n}A_i\#_t\sum_{i=1}^{n}B_i)(\sum_{i=1}^{n}B_i\#_t\sum_{i=1}^{n}A_i)|||, \quad \text{by (4.3.1).}
\end{aligned}$$

□

Chapter 6

Semisimple Lie groups and Lie algebras

6.1 Lie groups and Lie algebras

A *Lie group* G is both a differentiable manifold and a group such that the map

$$\begin{aligned} G \times G &\to G \\ (g,h) &\mapsto gh^{-1}, \end{aligned}$$

is a smooth map. In other words, the group structure and manifold structure are compatible. The Lie algebra $\mathfrak{g}$ of a Lie group G can be viewed as the tangent space of G at the identity with a bracket operation

$$\begin{aligned} \mathfrak{g} \times \mathfrak{g} &\to \mathfrak{g} \\ (X,Y) &\mapsto [X,Y]. \end{aligned}$$

satisfying the following conditions

1. $[X,Y]$ is bilinear.

2. $[X,X] = 0$ for all $X \in \mathfrak{g}$.

3. The Jacobi identity $[X,[Y,Z]] + [Y,[Z,X]] + [Z,[X,Y]] = 0$ holds for all $X,Y,Z \in \mathfrak{g}$.

The books of Helgason and Knapp [19, 22] are standard references for Lie groups and algebras. The following are well-known examples of Lie groups.

Example 6.1.1. 1. The general linear group $\mathrm{GL}_n(\mathbb{C}) = \{A \in \mathbb{C}_{n\times n} : \det A \neq 0\}$.

2. The special linear group $\mathrm{SL}_n(\mathbb{C}) = \{A \in \mathbb{C}_{n\times n} : \det A = 1\}$.

3. The orthogonal group $\mathrm{O}(n) = \{A \in \mathbb{R}_{n\times n} : A^\top A = I_n\}$ and the special orthogonal group $\mathrm{SO}(n) = \{A \in \mathrm{O}(n) : \det A = 1\}$.

4. The complex orthogonal group $\mathrm{O}_n(\mathbb{C}) = \{A \in \mathrm{GL}_n(\mathbb{C}) : A^\top A = I_n\}$.

5. The unitary group $\mathrm{U}(n) = \{A \in \mathbb{C}_{n\times n} : A^* A = I_n\}$.

6. The complex symplectic matrices $\mathrm{Sp}_{2n}(\mathbb{C}) = \{A \in \mathbb{C}_{2n\times 2n} : A^\top J A = J\}$, where

$$J = \begin{bmatrix} O_n & I_n \\ -I_n & O_n \end{bmatrix}.$$

Example 6.1.2. 1. The Lie algebra of $\mathrm{GL}_n(\mathbb{C})$ is $\mathfrak{gl}_n(\mathbb{C}) = \mathbb{C}_{n\times n}$ with $[A, B] = AB - BA$, $A, B \in \mathbb{C}_{n\times n}$.

2. The Lie algebra of $\mathrm{SL}_n(\mathbb{C})$ is $\mathfrak{sl}_n(\mathbb{C}) = \{A \in \mathbb{C}_{n\times n} : \operatorname{tr} A = 0\}$.

3. The Lie algebra of $\mathrm{O}(n)$ and $\mathrm{SO}(n)$ is $\mathfrak{so}(n) = \{A \in \mathbb{R}_{n\times n} : A^\top = -A\}$.

4. The Lie algebra of $\mathrm{O}_n(\mathbb{C})$ is $\mathfrak{so}_n(\mathbb{C}) = \{A \in \mathbb{C}_{n\times n} : A^\top = -A\}$.

5. The Lie algebra of $\mathrm{U}(n)$ is $\mathfrak{u}(n) = \{A \in \mathbb{C}_{n\times n} : A^* = -A\}$.

6. The Lie algebra of $\mathrm{Sp}_{2n}(\mathbb{C})$ is $\mathfrak{sp}_{2n}(\mathbb{C}) = \{A \in \mathbb{C}_{2n\times 2n} : AJ + A^\top J = 0\}$.

Let $\mathfrak{g}$ and $\mathfrak{h}$ be Lie algebras over the same field. A *Lie algebra homomorphism* (or simply homomorphsim) as a linear transformation $\phi : \mathfrak{g} \to \mathfrak{h}$ that respects the bracket, i.e., for any X and $Y \in \mathfrak{g}$

$$\phi([X, Y]) = [\phi(X), \phi(Y)].$$

Note that a *representation* of the Lie algebra $\mathfrak{g}$ is a Lie algebra homomorphism

$$\rho : \mathfrak{g} \to \mathfrak{gl}(V),$$

where V is a vector space over field $\mathbb{F}$.

The *adjoint representation* of $\mathfrak{g}$ is $\operatorname{ad} : \mathfrak{g} \to \mathfrak{gl}(\mathfrak{g})$ defined by

$$\operatorname{ad} X(Y) = [X, Y], \quad X, Y \in \mathfrak{g}.$$

It is a homomorphism because of the bilinearity of Lie bracket and Jacobi identity.

Let G be a Lie group with Lie algebra $\mathfrak{g}$. A *one-parameter* subgroup of G is a smooth homomorphism from the additive group of real numbers to G. i.e.,

$$\phi : \mathbb{R} \to G$$

There is a one-to-one correspondence between one-parameter subgroup of G and T_eG, [19, p. 103]. Now, by using this fact, we can define an analogue of the matrix exponential for Lie group G as

$$\exp : \mathfrak{g} \quad \to \quad G \tag{6.1.1}$$

$$X \quad \mapsto \quad \phi_X(1), \tag{6.1.2}$$

where $X \in \mathfrak{g}$ and ϕ_X is the one-parameter subgroup corresponding to X. Therefore, for all $t \in \mathbb{R}$,

$$\phi_X(t) = \exp(tX).$$

Let G and H be Lie groups. If $\phi : G \to H$ be a smooth homomorphism, and $d\phi : \mathfrak{g} \to \mathfrak{h}$ is its differential at the identity element, then

$$\phi \circ \exp_{\mathfrak{g}} = \exp_{\mathfrak{h}} \circ d\phi. \tag{6.1.3}$$

Given $g \in G$, let $I_g : G \to G$ be defined by

$$I_g(f) = g^{-1} f g, \quad f \in G$$

It can be seen that I_g is a smooth automorphism. Its differential $(dI_g)_e$ at the identity $e \in G$ is denoted by $\mathrm{Ad}\,(g) : \mathfrak{g} \to \mathfrak{g}$. The *adjoint representation* of G is $\mathrm{Ad} : G \to \mathrm{Aut}\,(\mathfrak{g})$. By (6.1.3), we have

$$\exp(\mathrm{Ad}\,(g)(X)) = g(\exp(X))g^{-1}, \text{where } g \in G,\ X \in \mathfrak{g}.$$

In the special case $G = \mathrm{GL}_n(\mathbb{C})$, we have $\mathrm{Ad}\,(g)(X) = g^{-1}Xg$, where $g \in \mathrm{GL}_n(\mathbb{C})$.

Let $\mathfrak{g}$ be the Lie algebra of the Lie group G. The bilinear form

$$B(X, Y) = \mathrm{tr}\,(\mathrm{ad}\, X \mathrm{ad}\, Y)$$

on $\mathfrak{g} \times \mathfrak{g}$ is called the Killing form of $\mathfrak{g}$. The Lie algebra $\mathfrak{g}$ (and the Lie group G) is called *semisimple* if the Killing form $B(\cdot, \cdot)$ is nondegenerate.

6.2 Cartan decomposition

Let G be a connected semisimple Lie group. Let Θ be a nontrivial involution in $\mathrm{Aut}\,(G)$ and let K be a compact subgroup of G such that K is the fixed point set of Θ. The derived automorphism of Θ is an automorphism of the Lie algebra $\mathfrak{g}$, denoted by $\theta := (d\Theta)_e$. Note that $\theta^2 = 1$. As a result, eigenvalues of θ are ± 1. The $+1$ eigenspace is the Lie algebra $\mathfrak{k}$ of K. Let $\mathfrak{p}$ be the eigenspace corresponding to eigenvalue -1, i.e.,

$$\begin{aligned} \mathfrak{k} &= \{X \in \mathfrak{g} : \theta X = X\}, \\ \mathfrak{p} &= \{X \in \mathfrak{g} : \theta X = -X\}. \end{aligned}$$

Thus, we have a direct sum decomposition of Lie algebra $\mathfrak{g}$:

$$\mathfrak{g} = \mathfrak{k} \oplus \mathfrak{p}. \tag{6.2.1}$$

which is called a *Cartan decomposition* corresponding to the *Cartan involution* $\theta : \mathfrak{g} \to \mathfrak{g}$ with $(X + Y) \mapsto X - Y$, where $X \in \mathfrak{k}$ and $Y \in \mathfrak{p}$. It is known that the Killing form $B(\cdot, \cdot)$ is positive definite on $\mathfrak{p}$ and negative definite on $\mathfrak{k}$. The Cartan decomposition can be lifted to the group level

$$G = PK$$

known as the *global Cartan decomposition*, where

$$P := \exp \mathfrak{p} = \{\exp(X) : X \in \mathfrak{p}\}. \tag{6.2.2}$$

The map

$$\begin{array}{rcl} \mathfrak{p} \times K & \to & G \\ (X, k) & \mapsto & \exp(X)\, k. \end{array}$$

is a diffeomorphism. Therefore, each $g \in G$ can be written as $g = pk$, with $k \in K$ and $p \in P$. For example, we have

$$\mathrm{SL}_n(\mathbb{R}) = \mathcal{SP}_n(\mathbb{R})\mathrm{SO}(n)$$

where $\mathcal{SP}_n(\mathbb{R})$ is the $n \times n$ real positive matrices of determinant one and $\mathrm{SO}(n)$ is the special orthogonal group. This is the classical polar decomposition of $\mathrm{SL}_n(\mathbb{R})$.

6.3 Iwasawa decomposition

In matrix theory the *Gram-Schmidt* orthogonalization process or QR decomposition enables us to write any $A \in \mathrm{GL}_n(\mathbb{C})$ as a product of a unitary matrix and an upper triangular matrix with positive diagonal entries. In 1949 Iwasawa extended the QR decomposition to semisimple Lie groups. Let G be a noncompact connected semisimple Lie group with associated Lie algebra $\mathfrak{g}$ with fixed Cartan decomposition $\mathfrak{g} = \mathfrak{k} \oplus \mathfrak{p}$ and corresponding Cartan involution θ. Let $\mathfrak{a}$ be the maximal abelian subspace of $\mathfrak{p}$. For any $H \in \mathfrak{a}$ consider the self adjoint transformation of $\mathfrak{g}$, $\operatorname{ad} H$. For any real linear functional λ on $\mathfrak{a}$, let

$$\mathfrak{g}_\lambda = \{X \in \mathfrak{g} : \ \operatorname{ad} H(X) = \lambda(H)X \ \text{ for all } \ H \in \mathfrak{a}\}. \tag{6.3.1}$$

If $\lambda \neq 0$ and $\mathfrak{g}_\lambda \neq 0$, we will call λ the *restricted root* of $\mathfrak{g}$ and $\mathfrak{g}_\lambda$ *restricted root space* correspond to λ. We will denote Σ the set of all restricted roots of $\mathfrak{g}$. For $\lambda = 0$ we have

$$\mathfrak{g}_0 = \{X \in \mathfrak{g} : \ (\operatorname{ad} H)X = 0 \ \text{ for all } \ H \in \mathfrak{a}\}. \tag{6.3.2}$$

Let Σ^+ and Σ^- be the set of positive roots and negative roots, respectively. We define

$$\mathfrak{n} = \bigoplus_{\lambda \in \Sigma^+} \mathfrak{g}_\lambda, \quad \mathfrak{n}^- = \bigoplus_{\lambda \in \Sigma^-} \mathfrak{g}_\lambda,$$

which are subalgebras of $\mathfrak{g}$. Note that if $X \in \mathfrak{n}^-$, by writing it as $X = (X + \theta(X)) - \theta(X)$, we see that it is an element of $\mathfrak{k} + \mathfrak{n}$. Since $[H, \theta X] = \theta[\theta H, X] = -\theta[H, X] = -\lambda(H)\theta X$, we have $\theta \mathfrak{g}_\lambda = \mathfrak{g}_{-\lambda}$. i.e., if $\lambda \in \Sigma^+$, then $-\lambda \in \Sigma^-$. Moreover, $\theta \mathfrak{g}_0 = \mathfrak{g}_0$. Therefore,

$$\mathfrak{g}_0 = (\mathfrak{g}_0 \cap \mathfrak{k}) \oplus (\mathfrak{g}_0 \cap \mathfrak{p}).$$

Since $\mathfrak{a}$ is the maximal abelian subspace of $\mathfrak{p}$, if $X \in (\mathfrak{g}_0 \cap \mathfrak{p})$, it implies that $X \in \mathfrak{p}$ and $X \in \mathfrak{g}_0$. Since $X \in \mathfrak{g}_0$, we have $[X, H] = 0$, for all $H \in \mathfrak{a}$. Thus, $X \in \mathfrak{a}$ and $\mathfrak{g}_0 \cap \mathfrak{p} = \mathfrak{a}$. Also, $\mathfrak{g}$ is the orthogonal direct sum $\mathfrak{g} = \mathfrak{g}_0 \bigoplus_{\lambda \in \Sigma} \mathfrak{g}_\lambda$, [22, Proposition 6.40 (a)].

Theorem 6.3.1. *(Iwasawa decomposition of Lie Algebra, [22, Proposition 6.43])*
Each semisimple Lie algebra $\mathfrak{g}$ *admits the following direct sum*

$$\mathfrak{g} = \mathfrak{k} \oplus \mathfrak{a} \oplus \mathfrak{n}.$$

It has a global counterpart for the Lie group G.

Theorem 6.3.2. *(Iwasawa decomposition of Lie Group)*
Let G *be a noncompact semisimple Lie group and let* K, A *and* N *be analytic subgroups of* G *with Lie algebras* $\mathfrak{k}, \mathfrak{a}$ *and* $\mathfrak{n}$*, respectively. Then the following map is a diffeomorphism onto* G*.*

$$\begin{array}{rcl} K \times A \times N & \to & G \\ (k, a, n) & \mapsto & kan. \end{array}$$

Consequently, any element in G can be uniquely written as $g = kan$, with $k \in K, a \in A$ and $n \in N$.

Example 6.3.3. As we have mentioned, the Lie algebra of $G = \mathrm{SL}_n(\mathbb{C})$ is $\mathfrak{sl}_n(\mathbb{C})$. Let $\mathfrak{sl}_n(\mathbb{C}) = \mathfrak{su}(n) + i\mathfrak{su}(n)$ be the corresponding Cartan decomposition with Cartan involution θ, where $\mathfrak{k} = \mathfrak{su}(n)$ is the Lie algebra of skew-Hermitian matrices of zero trace and $\mathfrak{p} = i\mathfrak{su}_n(\mathbb{C})$ is the set of Hermitian matrices of zero trace. We choose

$$\mathfrak{a} = \{X \in \mathfrak{p} : \ X = \mathrm{diag}\,(x_1, \ldots, x_n),\ x_i \in \mathbb{R} \text{ and } \mathrm{tr}\, X = 0 \text{ for } i = 1, \ldots, n\}.$$

Let E_{ij} be an $n \times n$ matrix with ij-th entry equals to 1 and elsewhere 0. For any $H \in \mathfrak{a}$, it can be seen that

$$\mathrm{ad}\, H(E_{ij}) = HE_{ij} - E_{ij}H = (e_i(H) - e_jH)E_{ij},$$

where $e_i(H)$ returns the i-th diagonal entry of H. That is, $e_i - e_j, i \neq j$, are restricted roots of $\mathfrak{sl}_n(\mathbb{C})$ and

$$\mathfrak{g}_{e_i - e_j} = \mathbb{C}E_{ij}.$$

As a result,

$$\mathfrak{n} = \bigoplus_{j>i,\ i\neq j} g_{e_i - e_j},$$

the set of all $n \times n$ strictly upper triangular matrices. So $N := \exp \mathfrak{n}$ is the set of upper triangular matrices with diagonal entries equal to 1. Now $K = \exp \mathfrak{k} = \mathrm{SU}(n)$ and $A = \exp \mathfrak{a}$, the group of diagonal matrices with real and positive entries with determinant 1. Therefore, $\mathrm{SL}_n(\mathbb{C}) = KAN$ in which $K = \mathrm{SU}(n)$ and $AN \subset \mathrm{SL}_n(\mathbb{C})$ is the subgroup of upper triangular with positive diagonal entries, i.e., the Iwasawa decomposition of $\mathrm{SL}_n(\mathbb{C})$ is the QR decomposition.

6.4 Complete Multiplicative Jordan Decomposition

Let G be a semisimple Lie group with Lie algebra $\mathfrak{g}$. An element $X \in \mathfrak{g}$ is called *semisimple* if the linear transformation $\operatorname{ad} X \in \operatorname{End}(\mathfrak{g})$ is diagonalizable over $\mathbb{R}$. Similarly, $X \in \mathfrak{g}$ is called *nilpotent* if $\operatorname{ad} X \in \operatorname{End}(\mathfrak{g})$ is nilpotent, i.e., the eigenvalues of the linear transformation $\operatorname{ad} X$ are all zeros.

Definition 6.4.1. (Elliptic, hyperbolic and unipotent elements)

1. An element $g \in G$ is called *hyperbolic* if it can be written as $g = \exp X$, where $X \in \mathfrak{g}$ is real semisimple.

2. An element $g \in G$ is called *unipotent* if it can be written as $g = \exp X$, where $X \in \mathfrak{g}$ is nilpotent.

3. An element $g \in G$ is called *elliptic* if $\operatorname{Ad} g$ is diagonalizable in $\mathbb{C}$ with eigenvalues of modulus 1.

We state the *Complete Multiplicative Jordan Decomposition* (CMJD) of a semisimple Lie group G.

Theorem 6.4.2. *(CMJD [23, p. 419])*
Let G be a semisimple Lie group. Each element $g \in G$ can be uniquely decomposed as

$$g = e(g)h(g)u(g), \tag{6.4.1}$$

where $e(g), h(g)$ and $u(g)$ are elliptic, hyperbolic and unipotent element of G and three elements commute.

See [19, p.430-431] for the CMJDs of $\mathrm{GL}_n(\mathbb{R})$ and $\mathrm{GL}_n(\mathbb{C})$, which can be obtained by the additive Jordan decomposition.

6.5 Kostant pre-order $\prec_G$

Let G be a noncompact connected semisimple Lie group with Lie algebra $\mathfrak{g}$. Let

$$\mathfrak{g} = \mathfrak{k} \oplus \mathfrak{a} \oplus \mathfrak{n}$$

be the Iwasawa decomposition of $\mathfrak{g}$ and

$$G = KAN$$

be the global Iwasawa decomposition of G [19, 22]. Since

$$K \times A \times N \ni (k, a, n) \mapsto kan \in G$$

is a diffeomorphism between $K \times A \times N$ and G, each $g \in G$ can be uniquely written as $g = kan$, where $k \in K$, $a \in A$, and $n \in N$.

For any $X \in \mathfrak{g}$, the Weyl group orbit of X, denoted by $\mathfrak{w}(X)$, is defined as

$$\mathfrak{w}(X) = \operatorname{Ad} G(X) \cap \mathfrak{a},$$

i.e., it is the set of elements in $\mathfrak{a}$ that is conjugate to X via the adjoint representation of G. It is known that ([23, Proposition 2.4]) $\mathfrak{w}(X)$ is a single W-orbit in $\mathfrak{a}$.

Let $\operatorname{conv} \mathfrak{w}(X)$ be the convex hull of the Weyl group orbit $\mathfrak{w}(X)$. For any $g \in G$, define

$$A(g) := \exp \operatorname{conv} \mathfrak{w}(\log h(g)),$$

where $h(g)$ is the hyperbolic component of g in its CMJD.

Kostant's pre-order on G ([23, p.426]) is defined by setting $f \prec_G g$ if

$$A(f) \subseteq A(g).$$

Note that if f and g are two distinct element of G, we may not have $f \prec_G g$ or $g \prec_G f$. This pre-order induces a partial order on the conjugacy classes of G.

We have the following characterization of Kostant pre-order. In the following example we will show that the Kostant pre-order coincide log-majorization in $\mathrm{SL}_n(\mathbb{C})$.

Example 6.5.1. Let $G = \mathrm{SL}_n(\mathbb{C})$. We will use notations of the example (6.3.3). Let $f, g \in \mathrm{SL}_n(\mathbb{C})$ with sets of eigenvalues $|\alpha| = \{|\alpha_1|, \ldots, |\alpha_n|\}$, with $|\alpha_1| \geq |\alpha_2| \geq, \ldots, \geq |\alpha_n|$, and $|\beta| = \{|\beta_1|, \ldots, |\beta_n|\}$, with $|\beta_1| \geq |\beta_2| \geq, \ldots, \geq |\beta_n|$, respectively. The hyperbolic part of f in CMJD decomposition is a diagonal matrix with diagonal entries $|\alpha_1|, \ldots, |\alpha_n|$, [35,

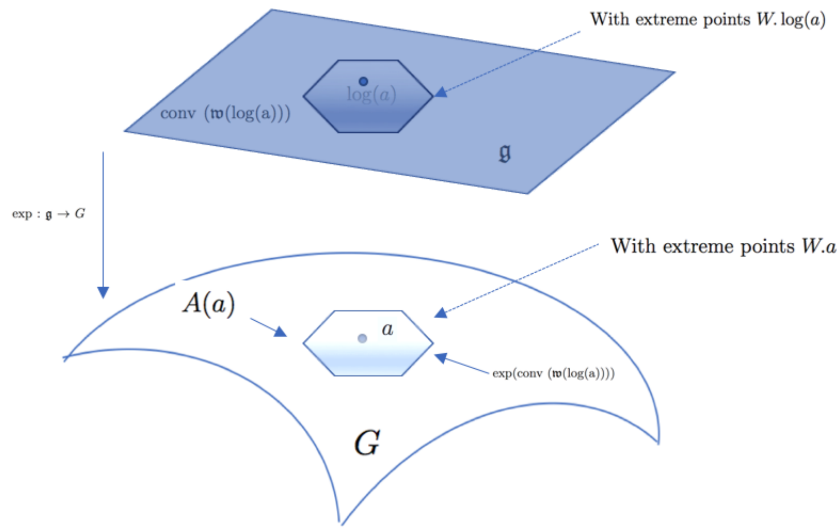

Figure 6.1: $A(a) = \exp(\text{conv }(\mathfrak{w}(\log(a))))$ with extreme points $W.a$.

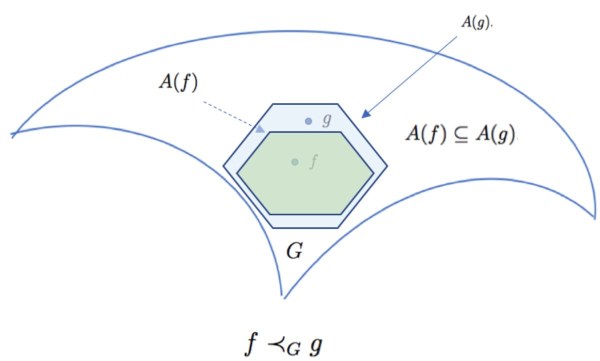

Figure 6.2: $f \prec_G g$ if and only if $A(f) \subseteq A(g)$

p. 543]. Thus we have

$$\begin{aligned} A(f) &= \exp \operatorname{conv} \mathfrak{w}(\log(\operatorname{diag}(|\alpha_1|, \ldots, |\alpha_n|))) \\ &= \exp \operatorname{conv} \{(\operatorname{diag}(\log|\alpha_{\sigma(1)}|, \ldots, \log|\alpha_{\sigma(n)}|), \sigma \in S_n\}, \end{aligned}$$

since the Weyl group W is the symmetric group S_n [22]. Therefore, by the fact that $x \prec y$ if and only if x is in the convex hull of all vectors obtained by permuting the coordinates of y. Thus, we have $A(f) \subseteq A(g)$, if and only if $\log|\alpha| \prec \log|\beta|$, or equivalently, $|\alpha| \prec_{\log} |\beta|$.

Though Kostant's pre-order appears to depend on the choice of $\mathfrak{a}$, it is actually not, due to the following nice characterization by Kostant.

Theorem 6.5.2. *(Kostant, [23, Theorem 3.1])*
Let $f, g \in G$. Then $f \prec g$ if and only if $\rho(\pi(f)) \leq \rho(\pi(g))$ for any irreducible finite dimensional representation $\pi : G \to \mathrm{GL}(V)$ of G, where $\rho(\pi(g))$ denotes the spectral radius of the operator $\pi(g)$.

Chapter 7

Extensions of some results of Lemos and Soares to semisimple Lie groups

In this chapter, G is a noncompact connected semisimple Lie group. So G admits a Cartan decomposition $G = PK$. It could be shown that P can be identified with G/K and regarded as a symmetric space of noncompact type [36, p.112]. Also, G/K has a unique analytic manifold structure with the property that G is a Lie transformation group on G/K under the natural G-action on G/K, see [19, Chapter II, Theorem 4.2]. Let $* : G \to G$ be the diffeomorphism defined by

$$*(g) = g^* = \Theta(g^{-1}),$$

where Θ is the Cartan involution. Then $k^* = k^{-1}$ for $k \in K$ and $p^* = p$ for $p \in P$. The map $G \to P$, $g \mapsto gg^*$, is onto. Because for any $g \in G$, it maps gK to a single point gg^*, it follows that the map

$$\psi : \; G/K \;\to P, \quad gK \mapsto gg^*, \tag{7.0.1}$$

is a bijection. It is in fact a diffeomorphism by the Cartan decomposition $G = PK$. Via ψ, P may be identified with G/K, and so may be regarded as a symmetric space of noncompact type. Note that for $p \in P$, $\psi^{-1}(p) = p^{1/2}K$, and G acts on P by

$$(g, p) \mapsto gpg^*.$$

Theorem 7.0.1. *(Liao, Liu, Tam [25, Proposition 2.3])*
Let $p, q \in P$. The unique geodesic $\alpha(t)$ joining p and q in P has the following parametrization

$$\alpha(t) = p^{1/2}(p^{-1/2}qp^{-1/2})^t p^{1/2}, \quad 0 \le t \le 1. \tag{7.0.2}$$

Proof. Using the identification of P and G/K via the map ψ defined in (7.0.1), the unique geodesic in P from p (at $t = 0$) to q (at $t = 1$) is given by $\gamma(t) = p^{1/2}e^{tY}p^{1/2}$ for some $Y \in \mathfrak{p}$. Because $q = \gamma(1) = p^{1/2}e^{Y}p^{1/2}$, $\gamma(t) = p^{1/2}(p^{-1/2}qp^{-1/2})^{t}p^{1/2}$. □

By comparing with the definition of t-geometric mean (4.1.2) on $\mathbb{P}_n$, we can see that the parametrization (7.0.2) has the same form. Therefore, we define the t-geometric mean of $p, q \in P$ as

$$p\#_t q = p^{1/2}(p^{-1/2}qp^{-1/2})^{1/2}p^{1/2}. \tag{7.0.3}$$

Because the G-action on G/K is isometric,

$$g(p\#_t q)g^* = (gpg^*)\#_t(gqg^*), \quad g \in G. \tag{7.0.4}$$

The symmetric space $P \subseteq G$ inherits the pre-order $\prec_G$ in G. We are going to extend some geometric mean inequalities of Lemos and Soares to P. Recall Theorem 4.2.2 of Lemos and Soares:

$$\lambda((A\#_t B)X(A\#_{1-t}B)X) \prec_{\log} \lambda(AXBX),$$

where $A, B, X \geq 0$ and $t \in [0, 1]$. In particular,

$$\rho(A\#_t B)X(A\#_{1-t}B)X) \leq \rho(AXBX). \tag{7.0.5}$$

where $\rho(\cdot)$ denotes the spectral radius. Now, we extend Theorem 4.2.2 in terms of Kostant's pre-order.

Theorem 7.0.2. *Let $p, q, r \in P$ and $t \in [0, 1]$. We have*

$$(p\#_t q)r(p\#_{1-t}q)r \prec_G prqr. \tag{7.0.6}$$

Proof. Let $\pi : G \to \mathrm{GL}(V)$ be any finite dimensional representation of G. By [23, p. 435], there exists an inner product on V such that for any $x \in P$, $\pi(x)$ is positive definite. Let

$d\pi : \mathfrak{g} \to \mathfrak{gl}(V)$ be the differential of π at the identity element of G. By (6.1.3), for any $X \in \mathfrak{g}$ we have $\pi(\exp_{\mathfrak{g}}(X)) = \exp_V(d\pi(X))$. For any $p \in P := \exp_{\mathfrak{g}} \mathfrak{p}$ there exists $X \in \mathfrak{p}$ such that $p = \exp_{\mathfrak{g}}(X)$. Therefore, for any $r \in \mathbb{R}$, we have

$$\begin{aligned}
\pi(p^r) &= \pi \exp_{\mathfrak{g}}(rX) \\
&= \exp_V(d\pi(rX)) \\
&= \exp_V(r d\pi(X)) \\
&= (\exp_V(d\pi(X)))^r \\
&= (\pi(\exp_{\mathfrak{g}}(X)))^r \\
&= (\pi(p))^r.
\end{aligned}$$

Consequently, we have

$$\begin{aligned}
\pi(p\#_t q) &= \pi(p^{1/2}(p^{-1/2}qp^{-1/2})^t p^{1/2} \\
&= \pi(p)^{1/2}(\pi(p)^{-1/2}\pi(p)\pi(p)^{-1/2})^t \pi(p)^{1/2} \\
&= \pi(p)\#_t \pi(q),
\end{aligned}$$

i.e., π respects the t-geometric mean. Hence

$$\begin{aligned}
\rho\Big(\pi((p\#_t q)r(p\#_{1-t}q)r)\Big) &= \rho\Big(\pi((p\#_t q))\pi(r)\pi((p\#_{1-t}q))\pi(r)\Big) \\
&= \rho\Big((\pi(p)\#_t\pi(q))\pi(r)(\pi(p)\#_{1-t}\pi(q))\pi(r)\Big) \\
&\leq \rho\Big(\pi(p)\pi(r)\pi(q)\pi(r)\Big) \quad \text{by (7.0.5)} \\
&= \rho\Big(\pi(prqr)\Big).
\end{aligned}$$

As a result, by Theorem (6.5.2), we complete the proof. □

Let $G = PK$ be a Cartan decomposition of Lie group G. Therefore, we can write each element $g \in G$ in the form $g = pk$, where $p \in P$ and $k \in K$. For any $g \in G$, in the Cartan decomposition, we will denote p-component by

$$|g| := p(g) = (gg^*)^{1/2}$$

and the k-component by $k(g)$. As indicated in the proof of Theorem (7.0.2), we have the polar decomposition

$$\pi(g) = \pi(pk) = \pi(p)\pi(k), \tag{7.0.7}$$

where $\pi(p)$ and $\pi(k)$ are the positive definite part and unitary part of $\pi(g)$, respectively. It can be seen that

$$\pi(g^*) = \pi(\Theta(pk)^{-1}) = \pi(\Theta(k^{-1})\Theta(p^{-1})) = \pi(k^{-1}p) = \pi(k)^{-1}\pi(p). \tag{7.0.8}$$

On the other hand, we can take the conjugate transpose of $\pi(g)$ as a matrix, i.e., $(\pi(g))^*$, and

$$\begin{aligned}
(\pi(g))^* &= (\pi(p)\pi(k))^* = \theta((\pi(p)\pi(k))^{-1}) \\
&= \theta(\pi(k)^{-1}\pi(p)^{-1})) \\
&= \theta(\pi(k)^{-1})\theta(\pi(p)^{-1}) \\
&= \theta(\pi(k)^{-1})\theta(\pi(p)^{-1}) \\
&= \pi(k)^{-1}\pi(p).
\end{aligned}$$

Therefore, $(\pi(g))^* = \pi(g^*)$. As a result, we have

$$\begin{aligned} |\pi(g)| &= (\pi(g)(\pi(g))^*)^{1/2} \\ &= (\pi(p)\pi(k)\pi(k)^{-1}\pi(p))^{1/2} \\ &= (\pi(p)^2)^{1/2} \\ &= \pi(p) \\ &= \pi(p(g)). \end{aligned}$$

That means, $|\pi(g)|$ is equal to positive definite component of the matrix $\pi(g)$. Using similar technique, we can extend Theorem 4.2.1, another result of Lemos and Soares in the following result. Recall that Theorem 4.2.1 asserts that if $A, B \geq 0$, then

$$\lambda(A(A\#B)B(A\#B)) \prec_{\log} \lambda(A^2B^2).$$

Let L denote the set of hyperbolic elements in G. It is known that ([23, Proposition 6.2])

$$L = P^2 := \{pq : p, q \in P\}.$$

Theorem 7.0.3. *Let $p, q \in P$. We have*

$$p^{1/2}(p\#q)q(p\#q)p^{1/2} \prec_G p^2q^2. \tag{7.0.9}$$

Note that $p^2q^2 \in L$ and $p^{1/2}(p\#q)q(p\#q)p^{1/2} \in P$.

Corollary 7.0.4. Let $p, q \in P$. We have

$$|p^{1/2}(p\#q)q^{1/2}| \prec_G |pq|. \tag{7.0.10}$$

Proof. We have

$$\pi(|g|) = (\pi(gg^*))^{1/2} = (\pi(g)(\pi(g))^*)^{1/2} = |\pi(g)|.$$

Thus to prove (7.0.10), we need to show $\rho(\pi(|g^{1/2}(g\#h)h^{1/2}|)) \leq \rho(\pi(|gh|))$ for all representation π of G. Now

$$\begin{aligned}
\rho(\pi(|p^{1/2}(p\#q)q^{1/2}|)) &= \rho(|\pi(p^{1/2}(p\#q)q^{1/2})|) \\
&= \rho(|\pi(p)^{1/2}(\pi(p)\#\pi(q))\pi(q)^{1/2}|) \\
&\leq \rho(|\pi(p)\pi(q)|) \qquad \text{by (4.2.4)} \\
&= \rho(|\pi(pq)|) \\
&= \rho(\pi(|pq|)).
\end{aligned}$$

Thus, (7.0.10) is established. □

We remark that the above proof technique was used in [26, 25, 33] to establish other inequalities.

Bibliography

[1] T. Ando, Concavity of certain maps on positive definite matrices and applications to Hadamard products, Linear Algebra Appl., **26** (1979), 203–241.

[2] T. Ando, F. Hiai, Log-majorization and complementary Golden-Thompson type inequalities, Second Conference of the International Linear Algebra Society (ILAS) (Lisbon, 1992), Linear Algebra Appl., **197/198** (1994), 113–131.

[3] T. Ando, X. Zhan, Norm inequalities related to operator monotone functions, Math. Ann., **315** (1999), 771–780.

[4] H. Araki, On an inequality of Lieb and Thirring, Lett. Math. Phys., **19** (1990), 167–170.

[5] K.M.R. Audenaert, A norm inequality for pairs of commuting positive semidefinite matrices, Electron. J. Linear Algebra **30** (2015), 80–84.

[6] R. Bhatia, *Matrix Analysis*, Springer, New York, 1997.

[7] R. Bhatia, *Positive Definite Matrices*, Princeton University Press, New Jersey, 2007.

[8] R. Bhatia, P. Grover, Norm inequalities related to the matrix geometric mean, Linear Algebra Appl., **437** (2012), 726–733.

[9] R. Bhatia, F. Kittaneh, Norm inequalities for positive operators, Lett. Math. Phys., **43** (1998), 225–231.

[10] R. Bhatia, Y. Lim, T. Yamazaki, Some norm inequalities for matrix means, Linear Algebra Appl., **501** (2016), 112–122.

[11] J.C. Bourin, Matrix subadditivity inequalities and block-matrices, Internat. J. Math., **20** (2009), 679–691.

[12] J.C. Bourin, M. Uchiyama, A matrix subadditivity inequality for $f(A+B)$ and $f(A)+f(B)$, Linear Algebra Appl., **423** (2007), 512–518.

[13] H. Dinh, An inequality for t-geometric means. Math. Inequal. Appl. **19** (2016), 765–768.

[14] H. Dinh, Some inequalities for the matrix Heron mean. Linear Algebra Appl. **528** (2017), 321–330.

[15] H. Dinh, S. Ahsani, T.Y. Tam, Geometry and inequalities of geometric mean, Czechoslovak Math. J., **66** (2016), 777–792.

[16] M. Fiedler, V. Pták, A new positive definite geometric mean of two positive definite matrices, Linear Algebra Appl., **251** (1997), 1–20.

[17] M. Hayajneh, S. Hayajneh, F. Kittaneh, Remarks on some norm inequalities for positive semidefinite matrices and questions of Bourin, Math. Inequal. Appl., **20** (2017), 225–232.

[18] S. Hayajneh, F. Kittaneh, Trace inequalities and a question of Bourin, Bull. Aust. Math. Soc., **88** (2013), 384–389.

[19] S. Helgason, *Differential Geometry, Lie Groups, and Symmetric Spaces*, Academic Press, New York, 1978.

[20] R. Horn, C.R. Johnson, *Matrix Analysis*, Second Edition, Cambridge University Press, 2013.

[21] F. Kittaneh, M. Lin, Trace inequalities for positive semidefinite block matrices, Linear Algebra Appl., **524** (2017), 153–158.

[22] A.W. Knapp, *Lie Groups Beyond an Introduction*, Second edition, Progress in Mathematics, 140, Birkhäuser, Boston, MA, 2002.

[23] B. Kostant, On convexity, the Weyl group and the Iwasawa decomposition, Ann. Sci. École Norm. Sup., (4) **6** (1973), 413–455.

[24] R. Lemos, G. Soares, Some log-majorizations and an extension of a determinantal inequality, Linear Algebra Appl., **547** (2018), 19–31.

[25] X. Liu, M. Liao, T.Y. Tam, Geometric mean for symmetric spaces of noncompact type, J. Lie Theory, **24** (2014), 725–736.

[26] X. Liu, T.Y. Tam, Extensions of three matrix inequalities to semisimple Lie groups, Special Matrices **2** (2014), 148–154.

[27] M. Marcus, *Finite Dimensional Multilinear Algebra I,* Marcel Dekker, New York, 1973.

[28] M. Marcus, *Finite Dimensional Multilinear Algebra II,* Marcel Dekker, New York, 1975.

[29] A.W. Marshall, I. Olkin, B. C. Arnold, *Inequalities: Theory of Majorization and Its Applications*, 2nd edition, Springer, New York, 2011.

[30] R. Merris, *Multilinear Algebra,* Gordon and Breach Science Publishers, Amsterdam, 1997.

[31] M. Moakher, A differential geometric approach to the geometric mean of symmetric positive-definite matrices, SIAM J. Matrix Anal. Appl., **26** (2005), 735–747.

[32] W. Pusz, S.L. Woronowicz, Functional calculus for sesquilinear forms and the purification map, Rep. Math. Phys., **8** (1975), 159–170.

[33] Z. Sarver, T.Y. Tam, Extension of Wang-Gong monotonicity result in semisimple Lie groups, Special Matrices, **3** (2015), 244–249.

[34] S. Sra, A new metric on the manifold of kernel matrices with application to matrix geometric means, Advances in neural information processing systems, **25** (2012), 144–152.

[35] T.Y. Tam, Some exponential inequalities for semisimple Lie groups, book chapter in *Operators, Matrices and Analytic Functions*, 539–552, Oper. Theory Adv. Appl., 202, Birkhäuser Verlag, Basel, 2010.

[36] T.Y. Tam, X. Liu, *Matrix Inequalities and Their Extensions to Lie Groups*, Monographs and Research in Mathematics, CRC Press, Boca Raton, FL, 2018.

[37] R.C. Thompson, Singular values, diagonal elements, and convexity, SIAM J. App. Math, **32** (1977), 39–63.