Complex Anal. Oper. Theory (2018) 12:969–986
https://doi.org/10.1007/s11785-018-0762-0

Complex Analysis and Operator Theory

Generalized Product of Two Square Matrices and Application for Some Algebraic Equations

Mustapha Raïssouli[1,2] · Ibrahim Al-Subaihi[1]

Received: 1 December 2016 / Accepted: 9 January 2018 / Published online: 19 January 2018

Abstract In this paper, a generalized matrix product is introduced and related properties are studied as well. Afterwards, we show how our approach can be applied to the so-called Sylvester and Lyaponov matrix equations for obtaining their related solutions in terms of the generalized matrix product. Numerical examples illustrating the theoretical study are also discussed.

Keywords Square matrix · Generalized matrix product · Matrix-convexity · Matrix equations · Sylvester equation · Lyapunov equation

Mathematics Subject Classification 15A60 · 15A24 · 26A51 · 47A64

1 Introduction and Basic Notions

Let $n \geq 2$ be an integer. We denote by $\mathcal{M}_n$ the space of $n \times n$ matrices with real entries, equipped with the classical norm

$$\forall A \in \mathcal{M}_n \qquad \|A\| = \sup_{\|u\|=1} \|Au\|,$$

Communicated by Izchak Lewkowicz.

✉ Mustapha Raïssouli
raissouli.mustapha@gmail.com

Ibrahim Al-Subaihi
alsubaihi@hotmail.com

1 Department of Mathematics, Faculty of Science, Taibah University, P.O.Box 30097, Al Madinah Al Munawwarah Zip Code 41477, Kingdom of Saudi Arabia

2 Science Faculty, Department of Mathematics, Moulay Ismail University, Meknes, Morocco

where $\|u\|$ denotes the euclidian norm of $u \in \mathbb{R}^n$.

By $\mathcal{S}_n$ we denote the subspace of $n \times n$ symmetric matrices. The set of $n \times n$ symmetric positive semi-definite matrices, denoted by $\mathcal{S}_n^+$, is a closed convex cone of $\mathcal{S}_n$ and thus positive semi-definiteness induces a partial ordering on $\mathcal{S}_n$, namely the Löwner order, defined by: $A \leq B$ if and only if $A, B \in \mathcal{S}_n$ and $B - A \in \mathcal{S}_n^+$. It is well-known that

$$\forall A \in \mathcal{S}_n^+ \qquad \|A\| = \sup_{\|u\|=1} \langle Au, u\rangle,$$

where $\langle ., .\rangle$ denotes the classical inner product of $\mathbb{R}^n$. We then infer that, if $A, B \in \mathcal{S}_n^+$ are such that $A \leq B$ then $\|A\| \leq \|B\|$.

By $\mathcal{S}_n^{+*}$ we denote the open convex cone of $n \times n$ symmetric positive definite matrices. Since the space $\mathcal{M}_n$ is with finite dimension then $A \in \mathcal{S}_n^{+*}$ if and only if $A \in \mathcal{S}_n^+$ and A is invertible. In another way, $\mathcal{S}_n^{+*}$ is the topological interior of $\mathcal{S}_n^+$.

Let $\Phi : \mathcal{C} \subset \mathcal{S}_n \longrightarrow \mathcal{S}_n$ be a matrix-map. We say that Φ is monotone increasing if $A \leq B$ implies $\Phi(A) \leq \Phi(B)$, where $\Phi(A)$ and $\Phi(B)$ are defined via functional calculus as usual. If moreover $\mathcal{C}$ is convex, we say that Φ is convex if for all $A, B \in \mathcal{C}$ and all real number $t \in [0, 1]$ we have

$$\Phi\Big((1-t)A + tB\Big) \leq (1-t)\Phi(A) + t\Phi(B).$$

The matrix-map Φ is said to be monotone decreasing (resp. concave) if $-\Phi$ is monotone increasing (resp. convex). As standard examples of such matrix-maps, we mention the following, see [3] and the related references cited therein.

Example 1.1 (i) Let $\Phi(X) = X^p$ for all $X \in \mathcal{S}_n^{+*}$, where p is a real number. It is well-known that the matrix-map Φ is monotone increasing and concave for all $p \in (0, 1)$, monotone decreasing and convex for each $p \in (-1, 0)$, convex not monotone for every $p \in (1, 2]$.

(ii) The matrix-map $X \longmapsto Log\ X$ is monotone increasing and concave on $\mathcal{S}_n^{+*}$ while $X \longmapsto exp\ X$ is neither monotone nor convex.

Otherwise, let $A, B \in \mathcal{S}_n^{+*}$. The geometric mean $G(A, B) \in \mathcal{S}_n^{+*}$ of A and B is defined as the unique positive definite matrix solution of the algebraic Riccati equation: find $Z \in \mathcal{S}_n^{+*}$ such that $ZA^{-1}Z = B$. It is well-known that $G(A, B)$ is explicitly given by

$$\begin{aligned} G(A, B) &= A^{1/2}\left(A^{-1/2}BA^{-1/2}\right)^{1/2}A^{1/2} \\ &= B^{1/2}\left(B^{-1/2}AB^{-1/2}\right)^{1/2}B^{1/2} = G(B, A). \end{aligned} \tag{1.1}$$

Obviously, $G(I, A) = G(A, I) = A^{1/2}$, where I denotes the $n \times n$ matrix identity. We recall that for fixed $A \in \mathcal{S}_n^{+*}$, the matrix map $X \longmapsto G(A, X)$ is monotone increasing concave and continuously differentiable on $\mathcal{S}_n^{+*}$.

As usual, $G(A, B)$ is extended for $A, B \in \mathcal{S}_n^+$ by setting

$$G(A, B) := \lim_{\epsilon \downarrow 0} G\big(A + \epsilon I, B + \epsilon I\big).$$

The following lemma, which will be needed in the sequel, asserts that $G(A, B)$ can be explicitly computed whenever A and B are 2×2 symmetric positive definite matrices.

Lemma 1.1 [1] *Let $A, B \in \mathcal{S}_2^{+*}$ and set $a = \sqrt{det\ A},\ b = \sqrt{det\ B}$. Then we have*

$$G(A, B) = \frac{\sqrt{ab}}{\sqrt{det\big(bA + aB\big)}}\big(bA + aB\big).$$

In particular,

$$A^{1/2} = \frac{\sqrt{a}}{\sqrt{det(A + aI)}}\big(A + aI\big). \tag{1.2}$$

For more details about properties, applications, some extensions and numerical computations of $G(A, B)$, we refer the interested reader to [1,2,10,11] and the related references cited therein.

2 Generalized Matrix Product

We preserve the same notations as previous. Let $\mathcal{C}$ be a nonempty subset of $\mathcal{S}_n$. We say that $\mathcal{C}$ satisfies:

- The property (P) if for all $A, B \in \mathcal{C}$ and $t > 0$, we have $A + tB \in \mathcal{C}$.
- The property (P_0) if for all $A \in \mathcal{C}$, $B \in \mathcal{S}_n$ and $|t|$ enough small, we have $A + tB \in \mathcal{C}$.

The following examples explain the previous terminologies.

Example 2.1 (i) If $\mathcal{C}$ is a convex cone (in particular a subspace) of $\mathcal{S}_n$ then $\mathcal{C}$ satisfies the property (P). In particular, if $\mathcal{C} = \mathcal{S}_n,\ \mathcal{C} = \mathcal{S}_n^+$ or $\mathcal{C} = \mathcal{S}_n^{+*}$ then $\mathcal{C}$ satisfies (P).

(ii) If $\mathcal{C}$ is open then it satisfies (P_0). For instance, $\mathcal{C} = \mathcal{S}_n^{+*}$ satisfies (P_0) but $\mathcal{C} = \mathcal{S}_n^+$ does not.

Example 2.2 Let $\mathcal{C}$ be defined as follows

$$\mathcal{C} = \{X \in \mathcal{S}_n, X \geq I\}.$$

It is easy to see that $\mathcal{C}$ satisfies (P), although $\mathcal{C}$ is not a cone.

Now, the following definition may be stated.

Definition 2.1 Let $\mathcal{C}$ be a nonempty subset of $\mathcal{S}_n$ satisfying (P_0) and $\Phi : \mathcal{C} \longrightarrow \mathcal{S}_n$ be a given map. For $A \in \mathcal{C}$ and $B \in \mathcal{S}_n$, we set

$$\big[A, B\big]_\Phi = \lim_{t \downarrow 0} \frac{\Phi(A + tB) - \Phi(A)}{t},$$

provided this limit exists. In this case, $[A, B]_\Phi$ is called the (generalized) Φ-matrix product of A and B.

The previous terminology "generalized matrix product" can be justified by the fact that $[A, B]_\Phi$ extends the commutative product of A and B, since $[A, B]_\Phi = AB + BA$ for $\Phi(X) = X^2$.

Clearly, $[A, B]_\Phi$ is the directional derivative of Φ in the direction B at A. If Φ is differentiable at A, with gradient $\nabla\Phi(A)$, then $[A, B]_\Phi$ exists for every $B \in \mathcal{S}_n$, and we have

$$\big[A, B\big]_\Phi = \nabla\Phi(A)(B).$$

On another hand, if the operator map $t \longmapsto \Phi(A + tB)$ is differentiable on a neighborhood of 0 then, by Hopital rule, we have

$$\big[A, B\big]_\Phi = \frac{d}{dt}\Phi(A + tB)\Big|_{t=0}.$$

It is also worth mentioning that if $B \in \mathcal{S}_n^+$ and Φ is monotone increasing then $[A, B]_\Phi \in \mathcal{S}_n^+$, since for $t > 0$ (enough small) we have

$$B \geq 0 \Longrightarrow A + tB \geq A \Longrightarrow \Phi(A + tB) \geq \Phi(A) \Longrightarrow \big[A, B\big]_\Phi \geq 0.$$

An interesting other situation is that where $\Phi : \mathcal{C} \longrightarrow \mathcal{S}_n$ is convex (resp. concave). The next result explains this latter situation.

Theorem 2.1 *Let $\mathcal{C}$ be a nonempty convex subset of $\mathcal{S}_n$ satisfying (P) and $\Phi : \mathcal{C} \longrightarrow \mathcal{S}_n$ be a convex map. Then, for all $A, B \in \mathcal{C}$, the matrix-function*

$$(0, \infty) \ni t \longmapsto \frac{\Phi(A + tB) - \Phi(A)}{t} \tag{2.1}$$

is monotone increasing. That is,

$$t_1 \geq t_2 > 0 \Longrightarrow \frac{\Phi(A + t_2 B) - \Phi(A)}{t_2} \leq \frac{\Phi(A + t_1 B) - \Phi(A)}{t_1}.$$

Proof If $t_1 \geq t_2 > 0$ then we can write

$$\Phi(A + t_2 B) - \Phi(A) = \Phi\Big(\frac{t_2}{t_1}\big(A + t_1 B\big) + \left(1 - \frac{t_2}{t_1}\right) A\Big) - \Phi(A).$$

This, with the fact that Φ is convex and $0 < t_2/t_1 \leq 1$, yields

$$\Phi(A + t_2 B) - \Phi(A) \leq \frac{t_2}{t_1}\Phi\big(A + t_1 B\big) + \Big(1 - \frac{t_2}{t_1}\Big)\Phi(A) - \Phi(A).$$

We then deduce, after simple manipulation,

$$\frac{\Phi(A+t_2B)-\Phi(A)}{t_2} \le \frac{\Phi(A+t_1B)-\Phi(A)}{t_1},$$

which is the desired result. □

The above theorem, with the definition of $[A, B]_\Phi$, immediately implies the following corollary.

Corollary 2.2 *Let $\mathcal{C}$ and Φ be as in the above theorem and let $A, B \in \mathcal{C}$. If $[A, B]_\Phi$ exists then*

$$\big[A, B\big]_\Phi = \inf_{t>0} \frac{\Phi(A+tB)-\Phi(A)}{t}, \tag{2.2}$$

where the "inf" is taken for the Löwner order.

Corollary 2.3 *Let $\mathcal{C}$ and Φ be as in Theorem 2.1. Let $A, B \in \mathcal{C}$ be such that $A-B \in \mathcal{C}$. Then $[A, B]_\Phi$ exists and satisfies*

$$\Phi(A) - \Phi(A-B) \le [A, B]_\Phi \le \Phi(A+B) - \Phi(A). \tag{2.3}$$

Proof We first show the left side of (2.3). The following identity

$$A = \frac{1}{1+t}(A+tB) + \frac{t}{1+t}(A-B)$$

is obviously satisfied for all $t > 0$ and all A, B. If moreover $A - B \in \mathcal{C}$ and Φ is convex then we have

$$\Phi(A) \le \frac{1}{1+t}\Phi(A+tB) + \frac{t}{1+t}\Phi(A-B).$$

It follows that

$$\Phi(A) + t\Phi(A) \le \Phi(A+tB) + t\Phi(A-B),$$

or equivalently

$$t\Phi(A) - t\Phi(A-B) \le \Phi(A+tB) - \Phi(A),$$

or again

$$\Phi(A) - \Phi(A-B) \le \frac{\Phi(A+tB)-\Phi(A)}{t}.$$

We then deduce that $[A, B]_\Phi$ exists and the desired matrix-inequality follows by letting $t \downarrow 0$, or by using (2.2). This, with (2.2), yields the right matrix-inequality of (2.3). The proof is so complete. □

The next corollary is immediate from the above one.

Corollary 2.4 *Let $\mathcal{C}$ be a subspace of $\mathcal{S}_n$ and $\Phi : \mathcal{C} \longrightarrow \mathcal{S}_n$ be a convex map. Then, for all $A, B \in \mathcal{C}$, $[A, B]_\Phi$ exists and (2.3) holds true.*

Remark 2.1 (i) It is easy to see that if $\mathcal{C}$ is a convex cone then $X \longmapsto \big[A, X\big]_\Phi$, for fixed $A \in \mathcal{C}$, is positively homogeneous and so it is sub-additive.

(ii) If the map $\Phi : \mathcal{C} \longrightarrow \mathcal{S}_n$ is concave then the matrix-function (2.1) is monotone decreasing. So, analog of (2.2) holds, with "sup" instead of "inf", and (2.3) is reversed. For fixed $A \in \mathcal{C}$, the map $X \longmapsto \big[A, X\big]_\Phi$ is concave.

Now, we present the following example illustrating the above.

Example 2.3 (i) Let $\Phi : \mathcal{S}_n \longrightarrow \mathcal{S}_n$ be defined by $\Phi(X) = XPX$, where $P \in \mathcal{S}_n^+$ is a fixed matrix. The map Φ is convex and it is simple to see that, for all $A, B \in \mathcal{S}_n$, we have

$$[A, B]_\Phi = APB + BPA = (AP)B + B(AP)^T.$$

In particular, if $\Phi(X) = X^2$ then one has $[A, B]_\Phi = AB + BA$.

(ii) Let $\Phi : \mathcal{S}_n^{+*} \longrightarrow \mathcal{S}_n^{+*}$ with $\Phi(X) = X^{-1}$. The map Φ is convex and it is not hard to see that

$$[A, B]_\Phi = -A^{-1}BA^{-1}.$$

(iii) Let $\Phi : \mathcal{S}_n^{+} \longrightarrow \mathcal{S}_n^{+}$ be such that $\Phi(X) = X^{1/2}$. Here, Φ is concave. We can easily see that $Z := [A^2, B]_\Phi$ satisfies the algebraic Lyapunov equation $AZ + ZA = B$.

3 Computation of $[A, B]_\Phi$ for $\Phi(X) = X^p$ and $\Phi(X) = \log X$

This section will be devoted to the computation of $[A, B]_\Phi$ for some special maps Φ. In particular, the case $\Phi(X) = X^p$, with p real number, will be discussed. For this case, we write $[A, B]_p$ instead of $[A, B]_\Phi$. With this, Example 2.3 gives

$$\big[A, B\big]_2 = AB + BA, \ \big[A, B\big]_{-1} = -A^{-1}BA^{-1}, \ A\big[A, B\big]_{1/2} + \big[A, B\big]_{1/2}A = B.$$

Otherwise, we set $A_s = A + sI$ for all $s \geq 0$. In particular $A_0 = A$ and $A_1 = A + I$. We first state the next auxiliary lemma.

Lemma 3.1 *Let $A, B \in \mathcal{M}_n$ and $p \geq 1$ be an integer. Then we have*

$$(A + tB)^p = A^p + t\sum_{i=1}^{p} A^{p-i}BA^{i-1} + t\,\epsilon_t(A, B), \tag{3.1}$$

where $\epsilon_t(A, B) \to 0$ *as* $t \to 0$. *If moreover* A *is invertible then*

$$(A + tB)^{-p} = A^{-p} - t \sum_{i=1}^{p} A^{i-p-1} B A^{-i} + t\ \eta_t(A, B), \tag{3.2}$$

with $\eta_t(A, B) \to 0$ *as* $t \to 0$.

Proof Expansion (3.1) follows from a simple mathematical induction on $p \geq 1$. Detail is simple and therefore omitted here. See also ([7], Lemma 2.3, page 951).

To prove (3.2) we first write

$$(A + tB)^{-p} = \Big((A + tB)^{-1}\Big)^p = \Big(A^{-1} - tA^{-1}BA^{-1} + t\epsilon_t(A, B)\Big)^p.$$

The desired expansion follows by applying (3.1) to this latter form, so completes the proof. □

Before stating another auxiliary lemma, we need some notation. For $A, B \in \mathcal{S}_n^{+*}$, $s \in [0, 1]$ and $t \geq 0$ we set

$$F_{A,B}(s, t) = \Big(f_A(s) + g(s, t)B\Big)^{-1},$$

where $s \longmapsto f_A(s) \in \mathcal{S}_n^{+*}$ is a continuous operator function with respect to $s \in [0, 1]$ and $g : [0, 1] \times [0, \infty) \longrightarrow [0, \infty)$ is a (positive) continuous function such that $g(s, 0) = 0$.

Lemma 3.2 *With the above, the following equality holds*

$$\lim_{t \downarrow 0} \sup_{s \in [0,1]} \|F_{A,B}(s, t) - F_{A,B}(s, 0)\| = 0.$$

Proof Using the obvious identity $X^{-1} - Y^{-1} = X^{-1}(Y - X)Y^{-1}$, valid for all invertible matrices X and Y, we easily check that

$$F_{A,B}(s, t) - F_{A,B}(s, 0) = -g(s, t)F_{A,B}(s, t)BF_{A,B}(s, 0). \tag{3.3}$$

Otherwise, it is clear that $f_A(s) + g(s, t)B \geq f_A(s)$ and therefore

$$F_{A,B}(s, t) \leq F_{A,B}(s, 0).$$

This, with the fact that $F_{A,B}(s, t) \in \mathcal{S}_n^+$ and $F_{A,B}(s, t) \in \mathcal{S}_n^+$, implies that

$$\|F_{A,B}(s, t)\| \leq \|F_{A,B}(s, 0)\|.$$

This, with (3.3), yields

$$\begin{aligned}\|F_{A,B}(s,t)-F_{A,B}(s,0)\| &= g(s,t)\|F_{A,B}(s,t)BF_{A,B}(s,0)\|\\ &\le g(s,t)\|F_{A,B}(s,t)\|\|B\|\|F_{A,B}(s,0)\|\\ &\le g(s,t)\|B\|\|F_{A,B}(s,0)\|^2.\end{aligned}$$

The real-function $s \longmapsto \|F_{A,B}(s,0)\|^2$ is continuous on the compact [0, 1] and so

$$0 \le \sup_{s\in[0,1]} \|F_{A,B}(s,0)\|^2 < \infty.$$

We can therefore write

$$\begin{aligned}0 &\le \lim_{t\downarrow 0} \sup_{s\in[0,1]} \|F_{A,B}(s,t)-F_{A,B}(s,0)\|\\ &\le \lim_{t\downarrow 0} g(s,t)\|B\|\Big(sup_{s\in[0,1]}\|F_{A,B}(s,0)\|^2\Big) = 0,\end{aligned}$$

since g is a continuous function with $g(s,0) = 0$. The desired result follows, so completes the proof of the lemma. □

Now, we are in position to state our desired claim recited as follows.

Theorem 3.3 *Let $A, B \in \mathcal{S}_n^{+*}$ and p be a real number. Then the following assertions hold:*

(1) If $p \ge 1$ is integer then we have

$$[A,B]_p = \sum_{i=1}^{p-1} A^{p-i}BA^{i-1} \quad and \quad [A,B]_{-p} = -\sum_{i=1}^{p} A^{i-p-1}BA^{-i}. \tag{3.4}$$

(2) If $0 < p < 1$ then one has

$$[A,B]_p = BA^{p-1} + \frac{sin(p\pi)}{\pi} A \int_0^\infty s^{p-1}[A_s,B]_{-1} ds. \tag{3.5}$$

(3) If p is such that $p = i_p + r_p$ where i_p is the integer part of p and $0 < r_p < 1$ then

$$[A,B]_p = [A,B]_{i_p} A^{r_p} + A^{i_p}[A,B]_{r_p}, \tag{3.6}$$

where $[A,B]_{i_p}$ and $[A,B]_{r_p}$ can be computed by (3.4) and (3.5), respectively.

Proof (1) Matrix-equalities (3.4) follow from (3.1) and (3.2), with the definition of $[.,.]_p$, respectively.

(2) Let $0 < p < 1$ and $A \in \mathcal{S}_n^{+*}$. We first recall that (see [6,8] for instance)

$$A^p = \frac{sin(p\pi)}{\pi} A \int_0^\infty s^{p-1}\big(sI + A\big)^{-1} ds. \tag{3.7}$$

We can then write

$$(A+tB)^p - A^p = \frac{sin(p\pi)}{\pi} A \int_0^\infty s^{p-1}\Big(\big(A_s+tB\big)^{-1} - A_s^{-1}\Big)ds$$
$$+t\frac{sin(p\pi)}{\pi} B \int_0^\infty s^{p-1}\big(A_s+tB\big)^{-1}ds. \qquad (3.8)$$

Writing

$$\big(A_s+tB\big)^{-1} - A_s^{-1} = -t\big(A_s+tB\big)^{-1}BA_s^{-1},$$

equality (3.8) yields

$$\frac{(A+tB)^p - A^p}{t} = -\frac{sin(p\pi)}{\pi} A \int_0^\infty s^{p-1}\big(A_s+tB\big)^{-1}BA_s^{-1}ds$$
$$+\frac{sin(p\pi)}{\pi} B \int_0^\infty s^{p-1}\big(A_s+tB\big)^{-1}ds$$

Now, setting

$$F_{A,B}(s,t) = \big(A_s+tB\big)^{-1},\ F_{A,B}(s,0) = A_s^{-1},$$

we are in position of Lemma 3.2 which tells us that $F_{A,B}(.,t)$ converges uniformly to $F_{A,B}(.,0)$ with respect to $t \downarrow 0$. We can therefore write

$$\lim_{t\downarrow 0}\frac{(A+tB)^p - A^p}{t} = -\frac{sin(p\pi)}{\pi} A \int_0^\infty s^{p-1}\lim_{t\downarrow 0}\big(A_s+tB\big)^{-1}BA_s^{-1}ds$$
$$+\frac{sin(p\pi)}{\pi} B \int_0^\infty s^{p-1}\lim_{t\downarrow 0}\big(A_s+tB\big)^{-1}ds,$$

which, with (3.7) again and the fact that $A_s^{-1}BA_s^{-1} = -[A_s, B]_{-1}$, yields the desired result.

(3) Let p be such that $p = i_p + r_p$ then we can write

$$\big(A+tB\big)^{i_p+r_p} - A^{i_p+r_p} = \Big(\big(A+tB\big)^{i_p} - A^{i_p}\Big)\big(A+tB\big)^{r_p}$$
$$+A^{i_p}\Big(\big(A+tB\big)^{r_p} - A^{r_p}\Big).$$

Dividing by $t > 0$ and then letting $t \downarrow 0$ we obtain (3.6) after a simple manipulation. The proof of the theorem is so completed. □

Now, we will be interested by $[A, B]_\Phi$ when $\Phi(X) = \log\ X$, case for which we write $[A, B]_{\log}$.

Theorem 3.4 *For all* $A, B \in \mathcal{S}_n^{+*}$ *we have*

$$\big[A, B\big]_{\log} = \int_0^1 \Big((1-s)I + sA\Big)^{-1} B\Big((1-s)I + sA\Big)^{-1} ds$$
$$= -\int_0^1 \big[(1-s)I + sA, B\big]_{-1} ds.$$

Proof Let $A, B \in \mathcal{S}_n^{+*}$. It is well-known that, see [11] for instance

$$\log\ A = \int_0^1 \frac{I - \Big((1-s)I + sA\Big)^{-1}}{s} ds.$$

We then deduce

$$\big[A, B\big]_{\log} = \lim_{t\downarrow 0} \frac{\log(A+tB) - \log\ A}{t} = \lim_{t\downarrow 0} \int_0^1 \frac{F_{A,B}(s,0) - F_{A,B}(s,t)}{ts} ds,$$

where

$$F_{A,B}(s,t) = \Big((1-s)I + sA + stB\Big)^{-1},$$
$$F_{A,B}(s,t) - F_{A,B}(s,0) = -st F_{A,B}(s,t) B F_{A,B}(s,0).$$

We are in the situation of Lemma 3.2 which, by similar way as above, yields

$$\big[A, B\big]_{\log} = -\int_0^1 \lim_{t\downarrow 0} \Big(F_{A,B}(s,0) B F_{A,B}(s,t)\Big) ds$$
$$= -\int_0^1 F_{A,B}(s,0) B F_{A,B}(s,0) ds,$$

and the desired result follows, so completing the proof. □

4 Computation of $[A, B]_{1/2}$ for 2×2-Matrices

In this section, we will be interested by computing explicitly $[A, B]_{1/2}$ when A and B are two 2×2-symmetric positive semi-definite matrices. The main result of this section is recited as follows.

Theorem 4.1 *Let* $A \in \mathcal{S}_2^{+}$ *and* $B \in \mathcal{S}_2$, *with* $det\ A = 1$. *Then there holds*

$$\Big[A, B\Big]_{1/2} = \big(det\ A_1\big)^{-1/2} A_1\big(\tilde{x} I + \tilde{X}\big), \tag{4.1}$$

where $A_1 = A + I$, $\tilde{x}$ *and* $\tilde{X}$ *are given by*

$$\tilde{x} = \tilde{x}(A, B) := \frac{1}{4} Tr(A^{-1}B) - \frac{1}{2} Tr\ \tilde{X},$$

and

$$\tilde{X} = \tilde{X}(A, B) := A_1^{-1}\Big(B + \frac{1}{2} Tr(A^{-1}B)I\Big).$$

To establish the previous theorem, we first need to prove a list of lemmas which we will state in the following.

Lemma 4.2 *Let* $A, B \in \mathcal{S}_n$ *with* A *invertible. Then we have*

$$det(A + tB) = (det\ A)\Big(1 + t\ Tr(A^{-1}B)\Big) + t\ \epsilon_t(A, B),$$

where $\epsilon_t(A, B))$ *tends to* 0 *when* $t \to 0$.

Proof The real-map $d : X \longmapsto det\ X$ is differentiable on $\mathcal{M}_n$ with gradient

$$\nabla d(X)(Y) = \langle com\ X, Y \rangle,$$

where $com\ X$ denotes the co-matrix of X and $\langle ., . \rangle$ is the classical inner product of $\mathcal{M}_n$ defined by $\langle X, Y \rangle = Tr(X^T Y)$. We then deduce, for all $A, B \in \mathcal{S}_n$,

$$det(A + tB) = det\ A + t\langle com\ A, B \rangle + t\ \epsilon_t(A, B).$$

If further A is invertible then $com\ A = A^{-1}(det\ A)$ and the desired result follows after a simple manipulation. □

Remark 4.1 For simplifying the writing, the notation ϵ_t throughout the following represents a quantity (depending generally on A and B) which tends to 0 when $t \to 0$. As usual, ϵ_t has not the same (expression) in each statement.

Lemma 4.3 *Let* $A \in \mathcal{S}_n^+$ *and* $B \in \mathcal{S}_n$, *with* $det\ A = 1$. *If we set*

$$c_t = \sqrt{det\big(A + tB\big)}$$

then we have

$$\sqrt{c_t} = 1 + \frac{t}{4} Tr(A^{-1}B) + t\ \epsilon_t. \tag{4.2}$$

$$\sqrt{det\big(A + tB + c_t I\big)} = \sqrt{det\ A_1}\Big(1 + \frac{t}{2} Tr\ \tilde{X}\Big) + t\ \epsilon_t, \tag{4.3}$$

where A_1 *and* $\tilde{X}$ *are as in the statement of the previous theorem.*

Proof By Lemma 4.2 we have

$$det(A + tB) = 1 + t\ Tr(A^{-1}B) + t\ \epsilon_t$$

and so

$$c_t = 1 + \frac{t}{2}Tr(A^{-1}B) + t\ \epsilon_t \tag{4.4}$$

and

$$\sqrt{c_t} = 1 + \frac{t}{4}Tr(A^{-1}B) + t\ \epsilon_t,$$

and so (4.2) is proved. Otherwise, with (4.4) we have

$$det(A + tB + c_t I) = det\Big(A + I + t\left(B + \frac{1}{2}Tr(A^{-1}B)I\right) + t\ \epsilon_t\Big),$$

which with Lemma 4.2 again becomes

$$det(A + tB + c_t I) = (det\ A_1)\big(1 + t\ Tr\tilde{X}\big) + t\ \epsilon_t,$$

from which we deduce (4.3), so completes the proof. □

Lemma 4.4 *Let $A \in \mathcal{S}_2^+$ and $B \in \mathcal{S}_2$, with $det\ A = 1$. Then we have*

$$\big(A + tB\big)^{1/2} = \big(det\ A_1\big)^{-1/2} A_1\Big(I + t\big(\tilde{x}I + \tilde{X}\big)\Big) + t\ \epsilon_t. \tag{4.5}$$

Proof By Lemma 1.1 one has

$$\big(A + tB\big)^{1/2} = \frac{\sqrt{c_t}}{\sqrt{det(A + tB + c_t I)}}\big(A + tB + c_t I\big).$$

According to Lemma 4.3 we then deduce

$$\big(A + tB\big)^{1/2} = \frac{1 + \frac{t}{4}Tr(A^{-1}B) + t\ \epsilon_t}{\sqrt{det\ A_1}\Big(1 + \frac{t}{2}Tr\ \tilde{X}\Big) + t\ \epsilon_t}\big(A + tB + c_t I\big).$$

By an elementary manipulation we then have

$$\begin{aligned}\big(A + tB\big)^{1/2} &= \big(det\ A_1\big)^{-1/2}\big(1 + \frac{t}{4}Tr(A^{-1}B) + t\ \epsilon_t\big)\\ &\quad \big(1 - \frac{t}{2}Tr\tilde{X} + t\ \epsilon_t\big)\big(A + tB + c_t I\big)\\ &= \big(det\ A_1\big)^{-1/2}\big(1 + t\ \tilde{x} + t\ \epsilon_t\big)\big(A + tB + c_t I\big).\end{aligned}$$

This, with (4.4), yields

$$\big(A + tB\big)^{1/2} = \big(det\ A_1\big)^{-1/2}\big(1 + t\ \tilde{x} + t\ \epsilon_t\big)\Big(A + I + tB + \frac{t}{2}Tr(A^{-1}B)I + t\ \epsilon_t\Big),$$

or again

$$\left(A+tB\right)^{1/2}=\left(det\ A_1\right)^{-1/2}A_1\Big(I+t\big(\tilde{x}I+\tilde{X}\big)\Big)+t\ \epsilon_t,$$

and the proof is completed. □

Now, using the above lemmas we are in position to establish our previous theorem. Indeed, by definition we have

$$\left[A,B\right]_{1/2}=\lim_{t\downarrow 0}\frac{(A+tB)^{1/2}-A^{1/2}}{t},$$

which, with (4.5) and (1.2), yields (after a simple reduction)

$$\left[A,B\right]_{1/2}=\lim_{t\downarrow 0}\Big(\left(det\ A_1\right)^{-1/2}A_1\big(\tilde{x}I+\tilde{X}\big)+\epsilon_t\Big)=\left(det\ A_1\right)^{-1/2}A_1\big(\tilde{x}I+\tilde{X}\big),$$

so completes the proof of Theorem 4.1.

Corollary 4.5 *Let $A\in\mathcal{S}_2^{+*}$ and $B\in\mathcal{S}_2$. Then we have*

$$\Big[A,B\Big]_{1/2}=\sqrt{a}\Big(det\ A_a\Big)^{-1/2}A_a\Big(\tilde{y}I+\tilde{Y}\Big),$$

where $a=\sqrt{det\ A}$, $A_a=A+aI$ and $\tilde{y}$, $\tilde{Y}$ are defined as follows

$$\tilde{y}=\tilde{y}(A,B):=\frac{1}{4}Tr(A^{-1}B)-\frac{1}{2}Tr\ \tilde{Y},$$

and

$$\tilde{Y}=\tilde{Y}(A,B):=A_a^{-1}\Big(B+\frac{a}{2}Tr(A^{-1}B)I\Big).$$

Proof Let $a=\sqrt{det\ A}$ and set $C=A/a$, $D=B/a$. Then it is easy to see that

$$[A,B]_{1/2}=[aC,aD]_{1/2}=\sqrt{a}[C,D]_{1/2},$$

with $det\ C=1$. Applying Theorem 4.1 for $[C,D]_{1/2}$, we obtain the desired result after simple computation and manipulation. Detail is simple and therefore omitted here. □

The following example, illustrating the above, will be needed in the sequel.

Example 4.1 Let us consider

$$A=\begin{pmatrix}89 & -60\\ -60 & 41\end{pmatrix},\quad B=\begin{pmatrix}398 & -257\\ -257 & 166\end{pmatrix}.$$

Executing a MATLAB program, we obtain the following results

$$
\begin{aligned}
a &= 7.000000000000027,\\
A_a &= \begin{pmatrix} 96.00000000000003 & -60.00000000000000 \\ -60.00000000000000 & 48.00000000000003 \end{pmatrix},\\
\tilde{Y} &= \begin{pmatrix} 4.511904761904752 & -1.285714285714292 \\ 0.2857142857142776 & 2.226190476190464 \end{pmatrix},\\
\tilde{y} &= -2.083333333333330,\\
[A, B]_{1/2} &= \begin{pmatrix} 17.99999999999999 & -11.00000000000000 \\ -11.00000000000000 & 6.999999999999993 \end{pmatrix} \approx \begin{pmatrix} 18 & -11 \\ -11 & 7 \end{pmatrix}.
\end{aligned}
$$

Remark 4.2 As usual, $[A, B]_{1/2}$ previously defined for $A \in \mathcal{S}_n^{+*}$ can be extended for $A \in \mathcal{S}_n^{+}$ by setting

$$[A, B]_{1/2} = \lim_{\epsilon \downarrow 0} [A + \epsilon I, B]_{1/2}.$$

5 Application 1: Sylvester Matrix Equation

Let $A \in \mathcal{S}_n^{+*}$ and $B \in \mathcal{S}_n$ be given and consider the following matrix equation

$$\text{Find } \ Z \in \mathcal{S}_n \ \text{ such that } \ AZ + ZA = B. \tag{5.1}$$

Such matrix equation, known in the literature as Sylvester equation, arises in various contexts and contributes as good tool for solving many scientific problems. For an approach solving general Sylvester equation by using derivative, we can consult [4].

Here, we will see how our present approach can be applied for the previous matrix equation in the aim to give an explicit form of its solution in terms of the generalized matrix product. Precisely the following result may be stated.

Theorem 5.1 *Let $A \in \mathcal{S}_n^{+*}$ and $B \in \mathcal{S}_n$. Then the Eq. (5.1) has one and only one solution in $Z \in \mathcal{S}_n$ given by $Z = [A^2, B]_{1/2}$.*

Proof If for a matrix $M = (m_{ij}) \in \mathcal{M}_n$ we denote by

$$vect\ M = \big(m_{11}, m_{21}, \ldots, m_{12}, m_{22}, \ldots, m_{nn}\big) \in \mathbb{R}^{n^2}$$

then the matrix equation $AZ + ZA = B$ can be written in the form of a linear system as well

$$\Big(I \otimes A + A \otimes I\Big) z = b, \quad \text{with } \ z = vect\ Z, \ b = vect\ B,$$

where the notation $\otimes$ refers to the so-called Kronecker product (or tentorial product) between two matrices. Such linear system can be solved if and only if $I \otimes A + A \otimes I$

is invertible. Since $A \in \mathcal{S}_n^{+*}$ then $I \otimes A + A \otimes I \in \mathcal{S}_{n^2}^{+*}$ and so the matrix equation $[A, Z]_2 := AZ + ZA = B$ has one and only one solution $Z \in \mathcal{S}_n$. Now, writing

$$\big(A + tZ\big)^2 = A^2 + tB + t\epsilon_t, \quad \epsilon_t = \epsilon_t(A, B) \to 0 \text{ as } t \to 0,$$

and taking the root side by side, by remarking that $A + tZ \in \mathcal{S}_n^{+*}$ for t enough small, we obtain

$$A + tZ = \Big(A^2 + tB + t\epsilon_t\Big)^{1/2}.$$

Since the map $r : X \longmapsto X^{1/2}$ is continuously differentiable on $\mathcal{S}_n^{+*}$ then we obtain

$$A + tZ = \Big(A^2 + tB + t\epsilon_t\Big)^{1/2} = A + t\nabla r(A^2)\big(B + \epsilon_t\big) + t\eta_t, \;\; \eta_t \to 0 \text{ as } t \to 0,$$

where $\nabla r(A^2)$ denotes the gradient of the map r at A^2. It follows that, after simplification and then by letting $t \downarrow 0$,

$$Z = \nabla r(A^2)(B) = \big[A^2, B\big]_{1/2}.$$

The proof of the lemma is completed. □

In [5], the authors showed that, if $A \in \mathcal{S}_n^{+*}$ and $B \in \mathcal{S}_n^{+}$ then the solution Z of the next matrix equation

$$A^2 Z + ZA^2 = AB + BA \tag{5.2}$$

is always positive semi-definite. Applying the above theorem to this matrix equation we immediately obtain the following.

Corollary 5.2 *The solution of the matrix Eq. (5.2) is given by*

$$Z = \Big[A^4, AB + BA\Big]_{1/2} \in \mathcal{S}_n^{+}.$$

6 Application 2: Lyapunov Matrix Equation

Let $M \in \mathcal{M}_n, \;\; B \in \mathcal{S}_n^{+}$. It is often of interest to solve the equation

$$\text{Find} \;\; Z \in \mathcal{S}_n^{+} \;\; \text{such that} \quad MZ + ZM^T = B. \tag{6.1}$$

Such matrix equation, known in the literature as the Lyapunov equation, occurs in the theory of stability and also arises in the theory of structures, see [9] for instance.

Before giving an explicit solution of (6.1) in terms of the generalized matrix product, we state the next needed lemma.

Lemma 6.1 *Let $\Phi : \mathcal{S}_n^{+*} \longrightarrow \mathcal{S}_n^{+*}$ be defined by $\Phi(X) = XPX$, where $P \in \mathcal{S}_n^{+*}$ is a fixed matrix. Let $A \in \mathcal{S}_n^{+*}$ and $B \in \mathcal{S}_n^{+}$. Then the equation:*

$$\textit{Find} \quad Z \in \mathcal{S}_n^{+} \ \textit{ such that } \ \big[A, Z\big]_{\Phi} = B$$

has one and only one solution given by

$$Z = \big[APA, B\big]_{\Phi^{-1}}, \quad \textit{with } \ \Phi^{-1}(X) = G\big(P^{-1}, X\big) \ \textit{ for all } \ X \in \mathcal{S}_n^{+*},$$

where G is the geometric matrix mean defined through (1.1).

Proof By Example 2.3, (i) the Eq. (6.1) can be written as

$$(AP)Z + Z(AP)^T = B.$$

Since $A, P \in \mathcal{S}_n^{+*}$ and $AP = A^{1/2}\big(A^{1/2}PA^{1/2}\big)A^{-1/2}$ then $Sp(AP) \subset (0, \infty)$ and so the $n^2 \times n^2$-matrix $I \otimes (AP) + (AP) \otimes I$ is invertible. It follows that the equation $[A, Z]_\Phi = B$ has one and only one solution $Z \in \mathcal{S}_n$. Further, $[A, Z]_\Phi = B$ is equivalent to

$$\Phi\big(A + tZ\big) = \Phi(A) + tB + t\,\epsilon_t, \quad \epsilon_t \to 0 \ \text{ as } t \to 0.$$

The map $\Phi^{-1} : X \longmapsto G(P^{-1}, X)$, inverse of Φ, is continuously differentiable on $\mathcal{S}_n^{+*}$ and so we can write

$$A + tZ = A + t\nabla\Phi^{-1}\big(\Phi(A)\big)(B + \epsilon_t) + t\,\eta_t, \quad \eta_t \to 0 \ \text{ as } \ t \to 0.$$

Similarly to the proof of Theorem 5.1 we deduce

$$Z = \nabla\Phi^{-1}\big(\Phi(A)\big)(B) = \Big[\Phi(A), B\Big]_{\Phi^{-1}} = \Big[APA, B\Big]_{\Phi^{-1}} \in \mathcal{S}_n.$$

Since $B \in \mathcal{S}_n^{+}$ and the matrix map Φ^{-1} is monotone increasing then $Z \in \mathcal{S}_n^{+}$. The proof of the lemma is completed. □

It is well known that every diagonalizable $n \times n$ matrix M with $Sp(M) \subset (0, \infty)$ can be written as product of two (symmetric) positive definite $n \times n$ matrices. Indeed, if $M = Q^{-1}DQ$ with D diagonal positive semi-definite, then we can write $M = AP$ with $A = Q^{-1}DQ^{-T}$ and $P = Q^TQ$.

We now are in position to state the following result concerning explicit solution of the Lyapunov Eq. (6.1).

Theorem 6.2 *Let $M \in \mathcal{M}_n$ be diagonalizable with $Sp(M) \subset (0, \infty)$ and $B \in \mathcal{S}_n^{+}$. Let $M = AP$ be a decomposition of M, with $A, P \in \mathcal{S}_n^{+*}$. Then the Lyapunov Eq.* (6.1) *has as solution*

$$Z = P^{-1/2}\Big[P^{1/2}M^2P^{-1/2}, P^{1/2}BP^{1/2}\Big]_{1/2}P^{-1/2}. \tag{6.2}$$

Proof Let $M = AP$ be as assumed. Then, Eq. (6.1) is equivalent to

$$APZ + ZPA = B.$$

According to Example 2.3, (i) this latter matrix equation can be written in the next equivalent form

$$[A, Z]_{\Phi} = B, \quad \text{with} \quad \Phi(X) = XPX \quad \text{for all} \quad X \in \mathcal{S}_n^{+*}.$$

This, with Lemma 6.1, yields

$$Z = [APA, B]_{\Phi^{-1}} \in \mathcal{S}_n^{+}. \tag{6.3}$$

Now, if we remark that the inverse map of $X \longmapsto \Phi(X) = XPX$ is

$$X \longmapsto \Phi^{-1}(X) = G(P^{-1}, X) = P^{-1/2}\Big(P^{1/2}XP^{1/2}\Big)^{1/2}P^{-1/2},$$

then (6.3) becomes (after simple manipulation)

$$\begin{aligned} Z &= \lim_{t\downarrow 0} P^{-1/2}\frac{\Big(P^{1/2}\big(APA + tB\big)P^{1/2}\Big)^{1/2} - \Big(P^{1/2}APAP^{1/2}\Big)^{1/2}}{t}P^{-1/2} \\ &= P^{-1/2}\Big[P^{1/2}APAP^{1/2}, P^{1/2}BP^{1/2}\Big]_{1/2}P^{-1/2}. \end{aligned}$$

To complete the proof we write

$$P^{1/2}APAP^{1/2} = P^{1/2}APAPP^{-1/2} = P^{1/2}M^2P^{-1/2},$$

and the desired result follows. □

Remark 6.1 (i) From (6.3) we deduce that if moreover $B \in \mathcal{S}_n^{+*}$ then so is Z, since Φ^{-1} is monotone increasing.

(ii) It is obvious that the decomposition $M = AP$, with A and P as above, is not unique. If the $n^2 \times n^2$-matrix $I \otimes M + M \otimes I$ is invertible then (6.1) has one and only one solution and in this case its solution Z given by (6.2) does not depend on the choice of P.

Finally, we present the following example illustrating the above.

Example 6.1 We search $Z \in \mathcal{S}_2^{+*}$ such that $MZ + ZM^T = B$ with,

$$M = \begin{pmatrix} 7 & -4 \\ -7 & 5 \end{pmatrix} \in \mathcal{M}_2, \quad B = \begin{pmatrix} 50 & -41 \\ -41 & 34 \end{pmatrix} \in \mathcal{S}_2^{+*}.$$

It is easy to see that $M = AP$ with

$$A = \begin{pmatrix} 2 & 1 \\ 1 & 4 \end{pmatrix} \in \mathcal{S}_2^{+*}, \quad P = \begin{pmatrix} 5 & -3 \\ -3 & 2 \end{pmatrix} \in \mathcal{S}_2^{+*}, \quad P^{1/2} = \begin{pmatrix} 2 & -1 \\ -1 & 1 \end{pmatrix}.$$

Simple computation leads to

$$P^{1/2} M^2 P^{-1/2} = \begin{pmatrix} 89 & -60 \\ -60 & 41 \end{pmatrix} \in \mathcal{S}_2^{+*}, \quad P^{1/2} B P^{1/2} = \begin{pmatrix} 398 & -257 \\ -257 & 166 \end{pmatrix} \in \mathcal{S}_2^{+*}.$$

Thanks to Example 4.1 with (6.2), we then obtain

$$Z = P^{-1/2} \begin{pmatrix} 18 & -11 \\ -11 & 7 \end{pmatrix} P^{-1/2} = \begin{pmatrix} 3 & -1 \\ -1 & 2 \end{pmatrix} \in \mathcal{S}_2^{+*},$$

which is the researched solution.

For this example, it is very easy to see that $I \otimes M + M \otimes I$ is invertible. According to Remark 6.1, another decomposition of $M = AP$ gives then the same solution.

References

1. Ando, T., Li, C.K., Mathias, R.: Geometric means. Linear Algebra Appl. **38**, 305–334 (2004)
2. Atteia, M., Raïssouli, M.: Self dual operators on convex functionals, geometric mean and square root of convex functionals. J Convex Anal **8**, 223–240 (2001)
3. Bhatia, R.: Matrix analysis. Verlag, New York (1997)
4. Bhatia, R., Uchiyama, M.: The operator equation $\sum_{i=1}^{n} A^{n-1} X B^{i} = Y$. Expo. Math. **27**, 251–255 (2009)
5. Chan, N.N., Kwong, M.K.: Hermitian matrix inequalities and a conjecture. Am. Math. Monthly **92**, 533–541 (1985)
6. Cardoso, J.R.: Computation of the Pth root and its Fréchet derivative by integrals. Electr. Trans. Num. An. **39**, 414–436 (2012)
7. Furuta, T.: Positive sem-definite solution of the operator equation $\sum_{j=1}^{n} A^{n-j} X A^{j-1} = B$. Linear Algebra Appl. **432**, 949–955 (2009)
8. Kato, T.: Perturbation theory for linear operators. Springer, Berlin (1984)
9. Lefschetz, S., Lasalle, J.P.: Stability by Lyapunov direct method. Academic Press, New York (1987)
10. Raïssouli, M., Leazizi, F.: Continued fraction expansion of the geometric matrix mean and applications. Linear Algebra Appl. **359**, 37–57 (2003)
11. Raïssouli, M., Bouziane, H.: Arithmetico-geometrico-harmonic functional mean in the sense of convex analysis. Ann. Sci. Math. Québec **30**(1), 79–107 (2006)