Cauchy–Schwarz Inequality and Riccati Equation for Positive Semidefinite Matrices

Masatoshi Fujii

Abstract By the use of the matrix geometric mean #, the matrix Cauchy–Schwarz inequality is given as $Y^*X \le X^*X \# U^*Y^*YU$ for $k \times n$ matrices X and Y, where $Y^*X = U|Y^*X|$ is a polar decomposition of Y^*X with unitary U. In this note, we generalize Riccati equation as follows: $X^*A^\dagger X = B$ for positive semidefinite matrices, where $A^\dagger$ is the Moore–Penrose generalized inverse of A. We consider when the matrix geometric mean $A \# B$ is a positive semidefinite solution of $XA^\dagger X = B$. For this, we discuss the case where the equality holds in the matrix Cauchy–Schwarz inequality.

2000 Mathematics Subject Classification 47A64, 47A63, 15A09

1 Introduction

One of the most important inequalities in functional analysis is the Cauchy–Schwarz inequality. It is originally an integral inequality, but is usually expressed as follows: Let H be a Hilbert space with inner product $\langle \cdot, \cdot \rangle$. Then

$$|\langle x, y\rangle| \le \langle x, x\rangle^{1/2} \langle y, y\rangle^{1/2} \quad \text{for } x, y \in H. \tag{1.1}$$

Matrix versions of the Cauchy–Schwarz inequality have been discussed by Marshall and Olkin [7], see also Bhatia and Davis [2] for operator versions.

Now we note that its right-hand side of (1.1) is the geometric mean of $\langle x, x\rangle$ and $\langle y, y\rangle$. From this viewpoint, Fujii [3] proposed a matrix Cauchy–Schwarz inequality by the use of the matrix geometric mean #, see [5, Lemma 2.6]. Let X and Y be $k \times n$

M. Fujii (✉)
Department of Mathematics, Osaka Kyoiku University, Kashiwara, Osaka, Japan
e-mail: mfujii@cc.osaka-kyoiku.ac.jp

D. Andrica, T. M. Rassias (eds.), *Differential and Integral Inequalities*, Springer Optimization and Its Applications 151,
https://doi.org/10.1007/978-3-030-27407-8_9

matrices and $Y^*X = U|Y^*X|$ a polar decomposition of an $n \times n$ matrix Y^*X with unitary U. Then

$$|Y^*X| \leq X^*X\#U^*Y^*YU,$$

where the matrix geometric mean # is defined by

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

for positive definite matrices A and B, see [6].

On the other hand, the original definition of it for operators is given by Ando [1] as follows: For $A, B \geq 0$, it is defined by

$$A\#B = \max\left\{X \geq 0; \begin{pmatrix} A & X \\ X & B \end{pmatrix} \geq 0\right\}.$$

Here a bounded linear operator A acting on a Hilbert space H is positive, denoted by $A \geq 0$, if $\langle Ax, x\rangle \geq 0$ for all $x \in H$. It is obvious that a matrix A is positive semidefinite if and only if $A \geq 0$, and A is positive definite if and only if $A > 0$, i.e., A is positive and invertible. It is known that if $A > 0$, then they coincide, that is,

$$\max\left\{X \geq 0; \begin{pmatrix} A & X \\ X & B \end{pmatrix} \geq 0\right\} = A^{1/2}(A^{-1/2}BA^{-1/2})^{1/2}A^{1/2}$$

holds for any $B \geq 0$.

Another approach of geometric mean is the Riccati equation. For $A > 0$ and $B \geq 0$, $A\#B$ is the unique solution of the Riccati equation

$$XA^{-1}X = B.$$

This fact is easily checked by multiplying $A^{-1/2}$ on both sides. For importance of Riccati equation, we refer [8]. Throughout this paper, we restrict our attention to positive semidefinite matrices, by which we can consider the generalized inverse $X^\dagger$ in the sense of Moore–Penrose even if they are not invertible. Among others, we generalize the Riccati equation to

$$XA^\dagger X = B.$$

In this paper, we discuss order relations between $A\#B$ and $A^{1/2}((A^{1/2})^\dagger B(A^{1/2})^\dagger)^{1/2}A^{1/2}$ for positive semidefinite matrices A and B. As an application, we discuss the case where the equality holds in matrix Cauchy–Schwarz inequality. Finally we generalize some results in our previous paper [4] by the use of the generalized inverse $X^\dagger$.

2 A Generalization of Formula for Geometric Mean

Since $A\#B = A^{1/2}(A^{-1/2}BA^{-1/2})^{1/2}A^{1/2}$ for invertible A, the geometric mean $A\#B$ for positive semidefinite matrices A and B might be expected the same formulae as for positive definite matrices, i.e.,

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

As a matter of fact, the following result is mentioned by Fujimoto and Seo [5]. For convenience, we cite it as Theorem FS:

Theorem FS *Let A and B be positive semidefinite matrices. Then*

$$A\#B \le A^{1/2}((A^{1/2})^{\dagger}B(A^{1/2})^{\dagger})^{1/2}A^{1/2},$$

If the kernel inclusion $\ker A \subset \ker B$ *is assumed, then the equality holds in above.*

We remark that the point of its proof is that A and B are expressed as $A = A_1 \oplus 0$ and $B = B_1 \oplus 0$ on $\operatorname{ran} A \oplus \ker A$, respectively, and $A^{\dagger} = (A_1)^{-1} \oplus 0$.

Now Theorem FS has an improvement in the following way. Below, let P_A be the projection onto $\operatorname{ran} A$, the range of A.

Theorem 2.1 *Let A and B be positive semidefinite matrices. Then*

$$A\#B \le A^{1/2}((A^{1/2})^{\dagger}B(A^{1/2})^{\dagger})^{1/2}A^{1/2},$$

In particular, the equality holds in above if and only if $P_A = AA^{\dagger}$ commutes with B.

To prove it, we cite the following lemma:

Lemma 2.2 *If* $\begin{pmatrix} A & X \\ X^* & B \end{pmatrix} \ge 0$, *then* $X = AA^{\dagger}X = P_AX$ *and* $B \ge XA^{\dagger}X$.

Proof The assumption implies that

$$\begin{pmatrix} (A^{1/2})^{\dagger} & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} A & X \\ X^* & B \end{pmatrix} \begin{pmatrix} (A^{1/2})^{\dagger} & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} P_A & (A^{1/2})^{\dagger}X \\ X^*(A^{1/2})^{\dagger} & B \end{pmatrix} \ge 0.$$

Moreover, since

$$\begin{aligned} 0 &\le \begin{pmatrix} 1 & -(A^{1/2})^{\dagger}X \\ 0 & 1 \end{pmatrix}^* \begin{pmatrix} P_A & (A^{1/2})^{\dagger}X \\ X^*(A^{1/2})^{\dagger} & B \end{pmatrix} \begin{pmatrix} 1 & -(A^{1/2})^{\dagger}X \\ 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} P_A & 0 \\ 0 & B - X^*A^{\dagger}X \end{pmatrix}, \end{aligned}$$

we have $B \ge X^*A^{\dagger}X$.

Next we show that $X = P_A X$, which is equivalent to $\ker A \subseteq \ker X^*$. Suppose that $Ax = 0$. Putting $y = -\frac{1}{\|B\|} X^* x$, we have

$$\begin{aligned} 0 &\le \left(\begin{pmatrix} A & X \\ X^* & B \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}, \begin{pmatrix} x \\ y \end{pmatrix} \right) \\ &= (Xy, x) + (X^* x, y) + (By, y) \\ &= -\frac{2}{\|B\|} \|X^* x\|^2 + \frac{1}{\|B\|^2} (BX^* x, X^* x) \\ &\le -\frac{\|X^* x\|^2}{\|B\|} \le 0. \end{aligned}$$

Hence we have $X^* x = 0$, that is, $\ker A \subseteq \ker X^*$ is shown.

Proof of Theorem 2.1 For the first half, it suffices to show that if $\begin{pmatrix} A & X \\ X & B \end{pmatrix} \ge 0$, then

$$X \le A^{1/2}((A^{1/2})^\dagger B (A^{1/2})^\dagger)^{1/2} A^{1/2}$$

because of Ando's definition of the geometric mean. We here use the facts that $(A^{1/2})^\dagger = (A^\dagger)^{1/2}$, and that if $\begin{pmatrix} A & X \\ X & B \end{pmatrix} \ge 0$ for positive semidefinite X, then $X = AA^\dagger X = P_A X$ and $B \ge XA^\dagger X$ by Lemma 2.2.

Now, since $B \ge XA^\dagger X$, we have

$$(A^{1/2})^\dagger B (A^{1/2})^\dagger \ge [(A^{1/2})^\dagger X (A^{1/2})^\dagger]^2,$$

so that Löwner–Heinz inequality implies

$$[(A^{1/2})^\dagger B (A^{1/2})^\dagger]^{1/2} \ge (A^{1/2})^\dagger X (A^{1/2})^\dagger.$$

Hence it follows from $X = P_A X$ that

$$A^{1/2}[(A^{1/2})^\dagger B (A^{1/2})^\dagger]^{1/2} A^{1/2} \ge X.$$

Namely we have $Y = A^{1/2}[(A^{1/2})^\dagger B (A^{1/2})^\dagger]^{1/2} A^{1/2} \ge A\#B$.

Next suppose that $\ker A \subset \ker B$. Then we have $\operatorname{ran} B \subset \operatorname{ran} A$ and so

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

Therefore, putting $C = (A^{1/2})^\dagger B (A^{1/2})^\dagger$, since

$$Y = A^{1/2}((A^{1/2})^\dagger B (A^{1/2})^\dagger)^{1/2} A^{1/2} = A^{1/2} C^{1/2} A^{1/2},$$

we have

$$\begin{pmatrix} A & Y \\ Y & B \end{pmatrix} = \begin{pmatrix} A^{1/2} & 0 \\ 0 & A^{1/2} \end{pmatrix} \begin{pmatrix} I & C^{1/2} \\ C^{1/2} & C \end{pmatrix} \begin{pmatrix} A^{1/2} & 0 \\ 0 & A^{1/2} \end{pmatrix} \geq 0,$$

which implies that $Y \leq A\#B$ and thus $Y = A\#B$ by combining the result $Y \geq A\#B$ in the first paragraph.

Now we show the second half. Notation as in above. If $P_A = AA^\dagger (= A^{1/2}(A^{1/2})^\dagger)$ commutes with B, we have $P_A B P_A \leq B$. Therefore we have

$$\begin{pmatrix} A & Y \\ Y & B \end{pmatrix} \geq \begin{pmatrix} A & Y \\ Y & P_A B P_A \end{pmatrix} = \begin{pmatrix} A^{1/2} & 0 \\ 0 & A^{1/2} \end{pmatrix} \begin{pmatrix} I & C^{1/2} \\ C^{1/2} & C \end{pmatrix} \begin{pmatrix} A^{1/2} & 0 \\ 0 & A^{1/2} \end{pmatrix} \geq 0,$$

which implies that $Y \leq A\#B$ and hence $Y = A\#B$.

Conversely assume that the equality holds. Then $\begin{pmatrix} A & Y \\ Y & B \end{pmatrix} \geq 0$. Hence we have

$$B \geq YA^\dagger Y = A^{1/2} C A^{1/2} = P_A B P_A,$$

which means P_A commutes with B, cf. Lemma 2.2.

3 Solutions of a Generalized Riccati Equation

Noting that $A\#B = A^{1/2}(A^{-1/2}BA^{-1/2})^{1/2}A^{1/2}$ for invertible A, the geometric mean $A\#B$ is the unique solution of the Riccati equation $XA^{-1}X = B$ if $A > 0$, see [8] for an early work. So we consider it for positive semidefinite matrices by the use of the Moore–Penrose generalized inverse, that is,

$$XA^\dagger X = B$$

for positive semidefinite matrices A, B.

Theorem 3.1 *Let A and B be positive semidefinite matrices satisfying the kernel inclusion* $\ker A \subset \ker B$. *Then $A\#B$ is a solution of a generalized Riccati equation*

$$XA^\dagger X = B.$$

Moreover, the uniqueness of its solution is ensured under the additional assumption $\ker A \subset \ker X$.

Proof We first note that $(A^{1/2})^\dagger = (A^\dagger)^{1/2}$ and $P_A = P_{A^\dagger}$. Putting $X_0 = A\#B$, either Theorem FS or 2.1 says that

$$X_0 = A^{1/2}[(A^{1/2})^\dagger B(A^{1/2})^\dagger]^{1/2} A^{1/2}.$$

Therefore we have

$$\begin{aligned} X_0 A^\dagger X_0 &= A^{1/2}[(A^{1/2})^\dagger B(A^{1/2})^\dagger]^{1/2} P_A [(A^{1/2})^\dagger B(A^{1/2})^\dagger]^{1/2} A^{1/2} \\ &= A^{1/2}[(A^{1/2})^\dagger B(A^{1/2})^\dagger] A^{1/2} \\ &= P_A B P_A = B \end{aligned}$$

Since $\operatorname{ran} X_0 \subset \operatorname{ran} A^{1/2}$, X_0 is a solution of the equation.

The second part is proved as follows: If X is a solution of $XA^\dagger X = B$, then

$$(A^{1/2})^\dagger X A^\dagger X (A^{1/2})^\dagger = (A^{1/2})^\dagger B (A^{1/2})^\dagger,$$

so that

$$(A^{1/2})^\dagger X (A^{1/2})^\dagger = [(A^{1/2})^\dagger B (A^{1/2})^\dagger]^{1/2}.$$

Hence we have

$$P_A X P_A = A^{1/2}[(A^{1/2})^\dagger B(A^{1/2})^\dagger]^{1/2} A^{1/2} = X_0.$$

Since $P_A X P_A = X$ by the assumption, $X = X_0$ is obtained.

As an application, we give a simple proof of the case where the equality holds in matrix Cauchy–Schwarz inequality, see [5, Lemma 2.5].

Corollary 3.2 *Let X and Y be $k \times n$ matrices and $Y^*X = U|Y^*X|$ a polar decomposition of an $n \times n$ matrix Y^*X with unitary U. If $\ker X \subset \ker YU$, then*

$$|Y^*X| = X^*X \# U^*Y^*YU$$

if and only if $Y = XW$ for some $n \times n$ matrix W.

Proof Since $\ker X^*X \subset \ker U^*Y^*YU$, the preceding theorem implies that $|Y^*X|$ is a solution of a generalized Riccati equation, i.e.,

$$U^*Y^*YU = |Y^*X|(X^*X)^\dagger |Y^*X| = U^*Y^*X(X^*X)^\dagger X^*YU,$$

or consequently

$$Y^*Y = Y^*X(X^*X)^\dagger X^*Y.$$

Noting that $X(X^*X)^\dagger X^*$ is the projection P_X, we have $Y^*Y = Y^*P_XY$ and hence

$$Y = P_XY = X(X^*X)^\dagger X^*Y$$

by $(Y - P_XY)^*(Y - P_XY) = 0$, so that $Y = XW$ for $W = (X^*X)^\dagger X^*Y$.

4 Geometric Mean in Operator Cauchy–Schwarz Inequality

The origin of Corollary 3.2 is the operator Cauchy–Schwarz inequality due to Fujii [3] as in below. Let $B(H)$ be the C^*-algebra of all bounded linear operators acting on a Hilbert space H.

OCS Inequality *If $X, Y \in B(H)$ and $Y^*X = U|Y^*X|$ is a polar decomposition of Y^*X with a partial isometry U, then*

$$|Y^*X| \le X^*X \# U^*Y^*YU.$$

In his proof of it, the following well-known fact due to Ando [1] is used: For $A, B \ge 0$, the geometric mean $A\#B$ is given by

$$A\#B = \max\left\{X \ge 0; \begin{pmatrix} A & X \\ X & B \end{pmatrix} \ge 0\right\}$$

First of all, we discuss the case $Y^*X \ge 0$ in (OCS). That is,

$$Y^*X \le X^*X\#Y^*Y$$

is shown: Noting that $Y^*X = X^*Y \ge 0$, we have

$$\begin{pmatrix} X^*X & X^*Y \\ Y^*X & Y^*Y \end{pmatrix} = \begin{pmatrix} X & Y \\ 0 & 0 \end{pmatrix}^* \begin{pmatrix} X & Y \\ 0 & 0 \end{pmatrix} \ge 0,$$

which means $Y^*X \le X^*X\#Y^*Y$.

The proof for a general case is presented by applying the above: Noting that $(YU)^*X = |Y^*X| \ge 0$, it follows that

$$|Y^*X| = (YU)^*X \le X^*X\#(YU)^*YU.$$

Incidentally, we can give a direct proof to the general case as follows:

$$\begin{pmatrix} X^*X & |Y^*X| \\ |Y^*X| & U^*Y^*YU \end{pmatrix} = \begin{pmatrix} X & YU \\ 0 & 0 \end{pmatrix}^* \begin{pmatrix} X & YU \\ 0 & 0 \end{pmatrix} \ge 0.$$

Related to matrix Cauchy–Schwarz inequality, the following result is obtained by Fujimoto–Seo [5]:

Let $\mathbb{A} = \begin{pmatrix} A & C \\ C^* & B \end{pmatrix}$ be positive definite matrix. Then $B \geq C^*A^{-1}C$ holds. Furthermore it is known by them:

Theorem 4.1 *Let $\mathbb{A}$ be as in above and $C = U|C|$ a polar decomposition of C with unitary U. Then*

$$|C| \leq U^*AU \,\#\, C^*A^{-1}C.$$

Proof It can be also proved as similar as in above: Since $|C| = U^*C = C^*U$, we have

$$\begin{pmatrix} U^*AU & |C| \\ |C| & C^*A^{-1}C \end{pmatrix} = \begin{pmatrix} A^{1/2}U & A^{-1/2}C \\ 0 & 0 \end{pmatrix}^* \begin{pmatrix} A^{1/2}U & A^{-1/2}C \\ 0 & 0 \end{pmatrix} \geq 0.$$

The preceding result is generalized a bit by the use of the Moore–Penrose generalized inverse, for which we note that $(A^{1/2})^\dagger = (A^\dagger)^{1/2}$ for $A \geq 0$:

Theorem 4.2 *Let $\mathbb{A}$ be of form as in above and positive semidefinite, and $C = U|C|$ a polar decomposition of C with unitary U. If* $\operatorname{ran} C \subseteq \operatorname{ran} A$, *then*

$$|C| \leq U^*AU \,\#\, C^*A^\dagger C.$$

Proof Let P_A be the projection onto the range of A. Since $P_AC = C$ and $C^*P_A = C^*$, we have $|C| = U^*P_AC = C^*P_AU$. Hence it follows that

$$\begin{pmatrix} U^*AU & |C| \\ |C| & C^*A^\dagger C \end{pmatrix} = \begin{pmatrix} A^{1/2}U & (A^\dagger)^{1/2}C \\ 0 & 0 \end{pmatrix}^* \begin{pmatrix} A^{1/2}U & (A^\dagger)^{1/2}C \\ 0 & 0 \end{pmatrix} \geq 0.$$

5 Solutions of Generalized Algebraic Riccati Equation

Following after [4], we discuss solutions of a generalized algebraic Riccati equation. Incidentally P_X means the projection onto the range of a matrix X.

Lemma 5.1 *Let A and B be positive semidefinite matrices and T an arbitrary matrix. Then W is a solution of a generalized Riccati equation*

$$W^*A^\dagger W = B + T^*AT$$

if and only if $X = W + AT$ is a solution of a generalized algebraic Riccati equation

$$X^*A^\dagger X - T^*P_AX - X^*P_AT = B.$$

Proof Put $X = W + AT$. Then it follows that

$$X^* A^\dagger X - T^* P_A X - X^* P_A T = W^* A^\dagger W - T^* AT,$$

so that we have the conclusion.

Theorem 5.2 *Let A and B be positive semidefinite matrices. Then W is a solution of a generalized Riccati equation*

$$W^* A^\dagger W = B \quad \textit{with } \operatorname{ran} W \subseteq \operatorname{ran} A$$

if and only if $W = A^{1/2} U B^{1/2}$ *for some partial isometry* U *such that* $U^*U \geq P_B$ *and* $UU^* \leq P_A$.

Proof Suppose that $W^* A^\dagger W = B$ and $\operatorname{ran} W \subseteq \operatorname{ran} A$. Since $\|(A^{1/2})^\dagger W x\| = \|B^{1/2} x\|$ for all vectors x, there exists a partial isometry U such that $UB^{1/2} = (A^{1/2})^\dagger W$ with $U^*U = P_B$ and $UU^* \leq P_A$. Hence we have

$$A^{1/2} U B^{1/2} = P_A W = W.$$

The converse is easily checked: If $W = A^{1/2} U B^{1/2}$ for some partial isometry U such that $U^*U \geq P_B$ and $UU^* \leq P_A$, then $\operatorname{ran} W \subseteq \operatorname{ran} A$ and

$$W^* A^\dagger W = B^{1/2} U^* P_A U B^{1/2} = B^{1/2} U^* U B^{1/2} = B.$$

Corollary 5.3 *Notation as in above. Then X is a solution of a generalized algebraic Riccati equation*

$$X^* A^\dagger X - T^* X - X^* T = B$$

with $\operatorname{ran} X \subseteq \operatorname{ran} A$ *if and only if* $X = A^{1/2} U (B + T^* AT)^{1/2} + AT$ *for some partial isometry* U *such that* $U^*U \geq P_{B+T^*AT}$ *and* $UU^* \leq P_A$.

Proof By Lemma 5.1, X is a solution of a generalized algebraic Riccati equation $X^* A^\dagger X - T^* P_A X - X^* P_A T = B$ if and only if $W = X - AT$ is a solution of $W^* A^\dagger W = B + T^* AT$. Since $\operatorname{ran} X \subseteq \operatorname{ran} A$ if and only if $\operatorname{ran} W \subseteq \operatorname{ran} A$, we have the conclusion by Theorem 5.2.

References

1. T. Ando, *Topics on Operator Inequalities*, Lecture Note (Hokkaido University, Hokkaido, 1978)
2. R. Bhatia, C. Davis, More operator versions of the Schwarz inequality. Commun. Math. Phys. **215**, 239–244 (2000)
3. J.I. Fujii, Operator-valued inner product and operator inequalities. Banach J. Math. Anal. **2**, 59–67 (2008)

4. J.I. Fujii, M. Fujii, R. Nakamoto, Riccati equation and positivity of operator matrices. Kyungpook Math. J. **49**, 595–603 (2009)
5. M. Fujimoto, Y. Seo, Matrix Wielandt inequality via the matrix geometric mean. Linear Multilinear Algebra **66**, 1564–1577 (2018)
6. F. Kubo, T. Ando, Means of positive linear operators. Math. Ann. **246**, 205–224 (1980)
7. A.W. Marshall, I. Olkin, Matrix versions of the Cauchy and Kantorovich inequalities. Aequationes Math. **40**, 89–93 (1990)
8. G.K. Pedersen, M. Takesaki, The operator equation $THT = K$. Proc. Amer. Math. Soc. **36**, 311–312 (1972)