A general Hua-type matrix equality and its applications

Minghua Lin*, Henry Wolkowicz†

January 7, 2013

Abstract

We present a very general Hua-type matrix equality. Among several applications of the proposed equality, we give a matrix version of the Aczél inequality.

Keywords: matrix equality, Schur complement, Aczél inequality, Löwner partial order.
AMS subjects classification 2010: 15A60; 15A24.

1 Hua-type matrix equality

Let $\mathbb{M}_{m\times n}$ be the set of all complex matrices of size $m \times n$ with $\mathbb{M}_n = \mathbb{M}_{n\times n}$. For $A \in \mathbb{M}_{m\times n}$, we denote the conjugate transpose of A by A^* and call A strictly contractive if $I - A^*A$ is positive definite, where I is the identity matrix of appropriate size. If $A \in \mathbb{M}_n$ is Hermitian positive (semi)definite, then we write $A(\geq) > 0$. Also, we identify $A > (\geq)B$ with $A - B > (\geq)0$, called the Löwner partial order.

The starting point of this paper is the following Hua matrix equality, which arises in studying the theory of functions of several variables.

Theorem 1.1. *[2] Let $A, B \in \mathbb{M}_{m\times n}$ be strictly contractive. Then*

$$(I - B^*B) - (I - B^*A)(I - A^*A)^{-1}(I - A^*B) = -(A - B)^*(I - AA^*)^{-1}(A - B). \quad (1.1)$$

Paige et al. [9] gave a new proof of (1.1) using the technique of Schur complements and extended it to the following form. (For simplicity, we do not consider generalized inverses in this paper. We refer the reader to [4, 9] for details with generalized inverses.)

Theorem 1.2. *[9] Let X, Y, W and $Z \in \mathbb{M}_{m\times n}$. Then*

$$(I + W^*Z) - (I + W^*Y)(I + X^*Y)^{-1}(I + X^*Z) = (W - X)^*(I + Y X^*)^{-1}(Z - Y), \quad (1.2)$$

where we assume all the relevant inverses exist.

---

*Research partially supported by The Natural Sciences and Engineering Research Council of Canada.
Email: mlin87@ymail.com
†Research partially supported by The Natural Sciences and Engineering Research Council of Canada and a grant from AFOSR. Email: hwolkowicz@uwaterloo.ca

In this paper, we present a matrix equality that is more general than (1.2).

Theorem 1.3. *Let* $X, Y, W, Z \in \mathbb{M}_{m\times n}$ *and* $R, T, U, V \in \mathbb{M}_n$*. Then*

$$\begin{aligned}(R^*T + W^*Z) &- (R^*V + W^*Y)(U^*V + X^*Y)^{-1}(U^*T + X^*Z)\\ &= (W - XU^{-1}R)^*\Big(I + (YV^{-1})(XU^{-1})^*\Big)^{-1}(Z - YV^{-1}T),\end{aligned} \tag{1.3}$$

where we assume all the relevant inverses exist.

Proof. Compute

$$\begin{aligned}
& (R^*T + W^*Z) - (R^*V + W^*Y)(U^*V + X^*Y)^{-1}(U^*T + X^*Z)\\
= \;& (R^*T + W^*Z)\\
& -\Big[R^*V + (U^{-1}R)^*X^*Y - (U^{-1}R)^*X^*Y + W^*Y\Big](U^*V + X^*Y)^{-1}(U^*T + X^*Z)\\
= \;& (R^*T + W^*Z) - (U^{-1}R)^*(U^*T + X^*Z)\\
& -(-(U^{-1}R)^*X^* + W^*)Y(U^*V + X^*Y)^{-1}(U^*T + X^*Z)\\
= \;& (W - XU^{-1}R)^*Z - (W - XU^{-1}R)^*YV^{-1}\Big(I + (XU^{-1})^*(YV^{-1})\Big)^{-1}(T + (XU^{-1})^*Z)\\
= \;& (W - XU^{-1}R)^*\Big[Z - \Big(I + (YV^{-1})(XU^{-1})^*\Big)^{-1}(YV^{-1}T + YV^{-1}(XU^{-1})^*Z)\Big]\\
= \;& (W - XU^{-1}R)^*\Big(I + (YV^{-1})(XU^{-1})^*\Big)^{-1}(Z - YV^{-1}T),
\end{aligned}$$

where we have used an easily verified formula $B(I + A^*B)^{-1} = (I + BA^*)^{-1}B$ in the fourth equality. $\square$

We may also give a proof of (1.3) using the techniques of Schur complements. The argument goes as follows:

Let

$$P := \begin{bmatrix} U & R\\ X & W\end{bmatrix}^* \begin{bmatrix} V & T\\ Y & Z\end{bmatrix} = \begin{bmatrix} U^*V + X^*Y & U^*T + X^*Z\\ R^*V + W^*Y & R^*T + W^*Z\end{bmatrix}.$$

The Schur complement of $U^*V + X^*Y$ in P is

$$(R^*T + W^*Z) - (R^*V + W^*Y)(U^*V + X^*Y)^{-1}(U^*T + X^*Z).$$

On the other hand, the following equivalent transformation on P preserves the Schur complement of $U^*V + X^*Y$; see [9, Lemma 1]:

$$\begin{bmatrix} I & 0\\ -(U^{-1}R)^* & I\end{bmatrix} P \begin{bmatrix} I & -V^{-1}T\\ 0 & I\end{bmatrix} = \begin{bmatrix} U^*V + X^*Y & X^*(Z - YV^{-1}T)\\ (W - XU^{-1}R)^*Y & (W - XU^{-1}R)^*(Z - YV^{-1}T)\end{bmatrix} := Q.$$

It remains to note that the Schur complement of $U^*V + X^*Y$ in Q is

$$\begin{aligned}
& (W - XU^{-1}R)^*(Z - YV^{-1}T) - (W - XU^{-1}R)^*Y(U^*V + X^*Y)^{-1}X^*(Z - YV^{-1}T)\\
= \;& (W - XU^{-1}R)^*\Big[I - Y(U^*V + X^*Y)^{-1}X^*\Big](Z - YV^{-1}T)\\
= \;& (W - XU^{-1}R)^*\Big[I - (YV^{-1})\Big(I + (XU^{-1})^*(YV^{-1})\Big)^{-1}(XU^{-1})^*\Big](Z - YV^{-1}T)\\
= \;& (W - XU^{-1}R)^*\Big(I + (YV^{-1})(XU^{-1})^*\Big)^{-1}(Z - YV^{-1}T).
\end{aligned}$$

A quick observation is that letting $R = T = U = V = I$ in (1.3), implies that (1.2) follows. The next two corollaries are also readily seen.

Corollary 1.4. *[12, Theorem 1.16] Let $A, B, X \in \mathbb{M}_n$. Then*

$$\begin{aligned}(AA^* + BB^*) &= (B + AX)(I + X^*X)^{-1}(B + AX)^* \\ &\quad +(A - BX^*)(I + XX^*)^{-1}(A - BX^*)^*. \end{aligned} \tag{1.4}$$

Proof. Putting $U = V = I$, $R = T = B^*$, $W = Z = A^*$ and $Y = X$ in (1.3), we obtain (1.4). □

Corollary 1.5. *[10, Theorem 3.2] Let $A, B, X \in \mathbb{M}_n$. Then*

$$\begin{aligned}(XX^* - BB^*) &= (X - BA^*)(I - AA^*)^{-1}(X - BA^*)^* \\ &\quad -(B - XA)(I - A^*A)^{-1}(B - XA)^*. \end{aligned} \tag{1.5}$$

Proof. Putting $U = V = I$, $R = T = X^*$, $W = -Z = B^*$ and $X = -Y = A^*$ in (1.3), we obtain (1.5). □

The next result, which we believe to be of interest in its own right, slightly generalizes [11, Theorem 1].

Theorem 1.6. *Let $A, B, C, D, E, F \in \mathbb{M}_n$ such that*

$$\begin{bmatrix} A & B \\ B^* & C \end{bmatrix} \geq 0, \quad \begin{bmatrix} D & E \\ E^* & F \end{bmatrix} \geq 0.$$

If moreover, $A > D$, $C > F$, and $\operatorname{rank}\left(\begin{bmatrix} A & B \\ B^* & C \end{bmatrix}\right) \leq n$, *then the following holds*

$$\begin{bmatrix} (A - D)^{-1} & (B^* - E^*)^{-1} \\ (B - E)^{-1} & (C - F)^{-1} \end{bmatrix} \geq 0, \tag{1.6}$$

where the inverses are well-defined.

Proof. We may write

$$\begin{bmatrix} A & B \\ B^* & C \end{bmatrix} = \begin{bmatrix} R^*R & R^*U \\ U^*R & U^*U \end{bmatrix}, \quad \begin{bmatrix} D & E \\ E^* & F \end{bmatrix} = \begin{bmatrix} W^*W & W^*X \\ X^*W & X^*X \end{bmatrix},$$

for some $R, U \in \mathbb{M}_n$, and $W, X \in \mathbb{M}_{m\times n}$. As $A, C > 0$, thus R, U are invertible. Putting $Z = -W$, $Y = -X$, $V = U$ and $T = R$ in (1.3), we have

$$\begin{aligned}&(R^*R - W^*W) - (R^*U - W^*X)(U^*U - X^*X)^{-1}(U^*R - X^*W) \\ &= -(W - XU^{-1}R)^*\Big(I - (XU^{-1})(XU^{-1})^*\Big)^{-1}(W - XU^{-1}R). \end{aligned} \tag{1.7}$$

Note that $U^*U > X^*X$ implies $I > (XU^{-1})^*(XU^{-1})$ and so $I > (XU^{-1})(XU^{-1})^*$, i.e., the right hand side of (1.7) is nonpositive definite. Therefore,

$$(R^*R - W^*W) \leq (R^*U - W^*X)(U^*U - X^*X)^{-1}(U^*R - X^*W),$$

i.e.,

$$(A-D) \le (B-E)(C-F)^{-1}(B^*-E^*). \tag{1.8}$$

As $A-D>0$, the above inequality guarantees the existence of $(B-E)^{-1}$. Taking the inverse on both sides of (1.8), we get

$$(A-D)^{-1} \ge (B^*-E^*)^{-1}(C-F)(B-E)^{-1},$$

and so (1.6) follows. □

Remark 1.7. The condition that rank $\left(\begin{bmatrix} A & B \\ B^* & C \end{bmatrix}\right) \le n$ in Theorem 1.6 is necessary. Otherwise, $B-E$ may not be invertible. As a quick example, consider

$$\begin{bmatrix} A & B \\ B^* & C \end{bmatrix} = \begin{bmatrix} 2I & 0 \\ 0 & I \end{bmatrix}, \quad \begin{bmatrix} D & E \\ E^* & F \end{bmatrix} = \begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix}.$$

2 Matrix Aczél inequality

If $A, B > 0$, then the geometric mean of A and B, denoted by $A\sharp B$, is the positive definite solution of the Ricatti equation $XB^{-1}X = A$ and has the explicit expression

$$A\sharp B = B^{1/2}(B^{-1/2}AB^{-1/2})^{1/2}B^{1/2}. \tag{2.1}$$

From here, we find that $A\sharp B = B\sharp A$, $(A\sharp B)^{-1} = A^{-1}\sharp B^{-1}$, and the monotonicity property: $A\sharp B \ge A\sharp C$ whenever $B \ge C > 0$ and $A > 0$. One of the motivations for such a geometric mean is of course the following matrix arithmetic-geometric mean inequality:

$$\frac{A+B}{2} \ge A\sharp B.$$

A remarkable property of the geometric mean is a maximal characterization by Pusz-Woronovicz [8]:

Theorem 2.1. *Let $A, B > 0$. Then*

$$A\sharp B = \max\left\{ X \;\middle|\; \begin{bmatrix} A & X \\ X^* & B \end{bmatrix} \ge 0, X = X^* \right\}. \tag{2.2}$$

The *maximum* here is in the sense of the Löwner partial order. With Theorem 2.1, the geometric mean (2.1) is also valid for $A, B \ge 0$. An equivalent possibility is

$$A\sharp B = \lim_{\epsilon\to 0} A\sharp(B+\epsilon I).$$

Applying this maximal characterization to the summation of positive semidefinite matrices $\begin{bmatrix} A_i & A_i\sharp B_i \\ A_i\sharp B_i & B_i \end{bmatrix}$, $i = 1, \ldots, n$, we get

$$\left(\sum_{i=1}^n A_i\right) \sharp \left(\sum_{i=1}^n B_i\right) \ge \sum_{i=1}^n A_i\sharp B_i. \tag{2.3}$$

The inequality (2.3) is a matrix Cauchy-Schwarz inequality [3], as it resembles the scalar Cauchy-Schwarz inequality: if $a_i, b_i \geq 0$, $i = 1, \ldots, n$, then

$$\left(\sum_{i=1}^{n} a_i\right)^{1/2} \left(\sum_{i=1}^{n} b_i\right)^{1/2} \geq \sum_{i=1}^{n} \sqrt{a_i b_i}. \tag{2.4}$$

A complement to (2.4) is the Aczél inequality [1]: if $a_i, b_i \geq 0$, $i = 0, 1, \ldots, n$, such that $a_0 \geq \sum_{i=1}^{n} a_i$ and $b_0 \geq \sum_{i=1}^{n} b_i$, then

$$\left(a_0 - \sum_{i=1}^{n} a_i\right)^{1/2} \left(b_0 - \sum_{i=1}^{n} b_i\right)^{1/2} \leq \sqrt{a_0 b_0} - \sum_{i=1}^{n} \sqrt{a_i b_i}. \tag{2.5}$$

There are also operator or matrix versions of the Aczél inequality; see [7] and the references therein. In this section, we present a new matrix Aczél inequality that is analogous to (2.3).

Theorem 2.2. *Let $A_i, B_i \geq 0$, $i = 0, 1, \ldots, n$, such that $A_0 \geq \sum_{i=1}^{n} A_i$ and $B_0 \geq \sum_{i=1}^{n} B_i$. Then*

$$\left(A_0 - \sum_{i=1}^{n} A_i\right) \sharp \left(B_0 - \sum_{i=1}^{n} B_i\right) \leq A_0 \sharp B_0 - \sum_{i=1}^{n} A_i \sharp B_i. \tag{2.6}$$

We need a few lemmas to prove (2.6).

Lemma 2.3. *[5, Lemma 2.2] Let $A > 0$ and B be any Hermitian matrix in $\mathbb{M}_n$. Then*

$$A \sharp (B A^{-1} B) \geq B. \tag{2.7}$$

Proof. We provide a short proof here for completeness. It is easy to see that

$$\begin{bmatrix} A & B \\ B & BA^{-1}B \end{bmatrix} \geq 0.$$

Now by (2.2), the desired inequality follows. □

Lemma 2.4. *Let $A, C \geq 0$, $B > 0$ be such that $A \leq CB^{-1}C$. Then*

$$A \sharp B \leq C. \tag{2.8}$$

Proof. We may assume $A, C > 0$. The general case follows from a continuity argument. Since $A \leq CB^{-1}C$ implies $A^{-1} \geq C^{-1}BC^{-1}$, the monotonicity of the geometric mean then gives

$$A^{-1} \sharp B^{-1} \geq (C^{-1}BC^{-1}) \sharp B^{-1} \geq C^{-1}, \tag{2.9}$$

where the second inequality is by Lemma 2.3. Now (2.8) follows by taking the inverse on both sides of (2.9). □

Proof of Theorem 2.2. We assume $A_0 > \sum_{i=1}^n A_i$ and $B_0 > \sum_{i=1}^n B_i$ in Theorem 2.2. The general case follows from a continuity argument. First, note that

$$\begin{bmatrix} A_0 & A_0 \sharp B_0 \\ A_0 \sharp B_0 & B_0 \end{bmatrix} \geq 0, \quad \begin{bmatrix} \sum_{i=1}^n A_i & \sum_{i=1}^n A_i \sharp B_i \\ \sum_{i=1}^n A_i \sharp B_i & \sum_{i=1}^n B_i \end{bmatrix} \geq 0.$$

Also, as $(A_0 \sharp B_0) A_0^{-1} (A_0 \sharp B_0) = B_0$, we get rank $\begin{bmatrix} A_0 & A_0 \sharp B_0 \\ A_0 \sharp B_0 & B_0 \end{bmatrix} = n$. Thus, the condition of Theorem 1.6 is satisfied. By (1.8), we have

$$\left(A_0 - \sum_{i=1}^n A_i\right) \leq \left(A_0 \sharp B_0 - \sum_{i=1}^n A_i \sharp B_i\right) \left(B_0 - \sum_{i=1}^n B_i\right)^{-1} \left(A_0 \sharp B_0 - \sum_{i=1}^n A_i \sharp B_i\right).$$

The monotonicity property of the geometric mean and (2.3) imply

$$A_0 \sharp B_0 \geq \left(\sum_{i=1}^n A_i\right) \sharp \left(\sum_{i=1}^n B_i\right) \geq \sum_{i=1}^n A_i \sharp B_i,$$

i.e., $A_0 \sharp B_0 - \sum_{i=1}^n A_i \sharp B_i \geq 0$.

By Lemma 2.4, it follows that

$$\left(A_0 - \sum_{i=1}^n A_i\right) \sharp \left(B_0 - \sum_{i=1}^n B_i\right) \leq A_0 \sharp B_0 - \sum_{i=1}^n A_i \sharp B_i.$$

□

We end the paper with several remarks.

Remark 2.5. Although Theorem 2.2 is stated and proved in the language of matrices, the proofs go through without any change in the context of linear operators on a Hilbert space.

Remark 2.6. The matrix Cauchy-Schwarz inequality (2.3) has been stated for accretive-dissipative matrices, with a generalized Löwner partial order involved; see [6, Corollary 2.2]. It is thus natural to ask whether Theorem 2.2 also has such an extension.

Remark 2.7. A matrix reverse Cauchy-Schwarz inequality has been considered in [3]. It is also of interest to consider a reverse direction to that of (2.4). We leave it for interested readers.

References

[1] J. Aczél, Some general methods in the theory of functional equations in one variable. New applications of functional equations, (in Russian), Uspekhi Mat. Nauk, 11 (1956) 3-68.

[2] L.-K. Hua, Inequalities involving determinants, (in Chinese), Acta Math. Sinica 5 (1955) 463-470. See also Trans. Amer. Math. Soc. Ser. II 32 (1963) 265-272.

[3] E.-Y. Lee, A matrix reverse Cauchy-Schwarz inequality, Linear Algebra Appl. 430 (2009) 805-810.

[4] M. Lin, Q. Wang, Remarks on Hua's matrix equality involving generalized inverses, Linear Multilinear Algebra 59 (2011) 1059-1067.

[5] M. Lin, Notes on Anderson-Taylor type inequality, submitted to Electron. J. Linear Algebra.

[6] M. Lin, Arithmetic, geometric, and harmonic means for accretive-dissipative matrices, preprint. `http://arxiv.org/abs/1206.0370`

[7] M.S. Moslehian, Operator Aczél inequality, Linear Algebra Appl. 434 (2011) 1981-1987.

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

[9] C.C. Paige, G.P.H. Styan, B.-Y.Wang, F. Zhang, Hua's matrix equality and Schur complements, Int. J. Inform. Syst. Sci. 4(1) (2008) 124-135.

[10] Z.-Z. Yan, Schur complements and determinant inequalities, J. Math. Inequal. 2 (2009) 161-167.

[11] G. Xu, C. Xu, F. Zhang, Contractive matrices of Hua type, Linear Multilinear Algebra 59 (2011) 159-172.

[12] F. Zhang. The Schur Complement and its Applications. Numerical Methods and Algorithms, Springer-Verlag, New York, 2005.

Minghua Lin
Department of Applied Mathematics,
University of Waterloo,
Waterloo, ON, N2L 3G1, Canada.
mlin87@ymail.com

Henry Wolkowicz
Department of Combinatorics and Optimization,
University of Waterloo,
Waterloo, ON, N2L 3G1, Canada.
hwolkowicz@uwaterloo.ca