INTERNATIONAL JOURNAL OF
INFORMATION AND SYSTEMS SCIENCES
Volume 3, Number 1, Pages 1–18

HUA'S MATRIX EQUALITY AND SCHUR COMPLEMENTS

CHRIS PAGE, GEORGE P. H. STYAN, BO-YING WANG, AND FUZHEN ZHANG

Abstract. The purpose of this paper is to revisit Hua's matrix equality (and inequality) through the Schur complement. We present Hua's original proof and two new proofs with some extensions of Hua's matrix equality and inequalities. The new proofs use a result concerning Schur complements and a generalization of Sylvester's law of inertia, each of which is useful in its own right.

Key Words. Contractions, contractive matrices, generalized inverses, Hua matrix, Hua's determinantal inequality, Hua's matrix equality, Hua's matrix inequality, Hua-Marcus inequalities, inertia additivity, matrix inequalities, rank additivity, Schur complement, Sylvester's law of inertia.

1. Introduction

Loo-Keng Hua (1910-1985) was a great mathematician and a Chinese legendary hero. He had little formal education, but made enormous contributions to number theory, algebra, complex analysis, matrix geometry and applied mathematics [10]. He worked with G.H. Hardy for two years and spent several years in USA [8].

The ideas in this paper were motivated by an elegant matrix equality which was established in 1955 by Hua in [14]. Let A and B be $m \times n$ complex matrices, I_n or simply I, be the $n \times n$ identity matrix, and denote by A^* the conjugate transpose of the matrix A. When $I_n - A^*A$ and so $I_m - AA^*$ is nonsingular, Hua showed (in the proof of Theorem 1 of [14]) the matrix equality

$$(1)\qquad \begin{aligned} &(I - B^*B) - (I - B^*A)(I_n - A^*A)^{-1}(I - A^*B) \\ &\qquad = -(A - B)^*(I_m - AA^*)^{-1}(A - B). \end{aligned}$$

Hua's paper [14] does not appear to be widely known or easily available, so it is worthwhile giving his short proof here, followed by a brief explanation:

$$\begin{aligned} &(I - B^*A)(I - A^*A)^{-1}(I - A^*B) - (I - B^*B) \\ &\quad = (I - A^*A)^{-1} - I - B^*A(I - A^*A)^{-1} \\ &\qquad -(I - A^*A)^{-1}A^*B + B^*(I + A(I - A^*A)^{-1}A^*)B \\ &\quad = A^*(I - AA^*)^{-1}A - B^*(I - AA^*)^{-1}A \\ &\qquad -A^*(I - AA^*)^{-1}B + B^*(I - AA^*)^{-1}B \\ &\quad = (B - A)^*(I - AA^*)^{-1}(B - A). \end{aligned}$$

In showing the second identity in this proof, we can use the formulae

$$(2)\qquad (I - A^*A)^{-1} = I + A^*(I - AA^*)^{-1}A$$

Received by the editors December 19, 2006 and, in revised form, March 22, 2007.
2000 *Mathematics Subject Classification.* 15A09, 15A45.
This research was supported in part by the Natural Sciences and Engineering Research Council of Canada, an NSF grant of China, and NSU-FAR Faculty Development Funds.

and

$$(I - A^*A)^{-1}A^* = A^*(I - AA^*)^{-1}$$

which are easily established by multiplying both sides by $I - A^*A$. These well-known matrix identities can be found in, for example, [6, 21, 24].

In establishing (1), Hua [14] took $m = n$ and both A and B to be *strictly contractive*: a matrix A is said to be *contractive* whenever $I - A^*A$ is nonnegative definite, namely the eigenvalues of A^*A (i.e., the singular values of A) all lie in the closed unit interval [0, 1]. Horn & Johnson [13, p. 46] call such matrices A *contractions*; see also Bhatia [2, p. 7]. When $I - A^*A$ is positive definite, all singular values of A lie in the half-open unit interval [0, 1), such an A is said to be *strictly contractive*. But it is clear from the proof that the only restrictions on A and B are that they both be $m \times n$, and $I - A^*A$ be nonsingular.

If $I - A^*A$ is positive definite, it is clear from (1) that (see also Ando [1])

$$I - B^*B \leq (I - B^*A)(I - A^*A)^{-1}(I - A^*B). \tag{3}$$

The matrix partial ordering in (3) is in the Löwner sense: $F \leq G$ means that both F and G are Hermitian matrices of the same size and $G - F$ is nonnegative definite. In particular, $G \geq 0$ means that G is nonnegative definite. Note that, in light of (1), equality holds in (3) if and only if $A = B$.

Hua's main purpose for (1) in [14] was to establish the determinant inequality

$$\det(I - A^*A) \cdot \det(I - B^*B) \leq |\det(I - A^*B)|^2 \tag{4}$$

with both A and B (strictly) contractive. To do this he first proved the now well-known inequality (see, for instance, [16, p. 117]): for Hermitian $M \geq 0$, $S \geq 0$

$$\det(M + S) \geq \det(M) + \det(S). \tag{5}$$

Equality holds in (5) if and only if $M + S$ is singular or $M = 0$ or $S = 0$ (or M and S are scalars). Hua then proved the following inequality (6), by taking

$$M = I - B^*B, \quad S = (A - B)^*(I_m - AA^*)^{-1}(A - B),$$

and using (1)

$$M + S = (I - B^*A)(I_n - A^*A)^{-1}(I - A^*B)$$

to get

$$\begin{aligned} |\det(I - A^*B)|^2 \quad \geq \quad & \det(I - A^*A) \cdot \big(\det(I - B^*B) \\ & + \det((A - B)^*(I_m - AA^*)^{-1}(A - B))\big), \end{aligned} \tag{6}$$

which yields (4), where equality holds if and only if $A = B$ (or $m = n = 1$).

If A and B are square, then

$$|\det(I - A^*B)|^2 \quad \geq \quad \det(I - A^*A) \cdot \det(I - B^*B) + |\det(A - B)|^2.$$

In 1958 Marcus [15] extended (4), and also proved for $I - A^*A$ and $I - B^*B$ nonnegative definite and any complex n-vector x:

$$x^*(I - A^*A)x \cdot x^*(I - B^*B)x \leq |x^*(I - A^*B)x|^2. \tag{7}$$

In 1980 Ando [1] generalized (7), pointing out if A and B are strictly contractive then (7) implies (3).

We shall call the expression (matrix) on the left-hand side of the identity (1) *Hua matrix* and denoted it by H, and the matrix equality (1) *Hua's matrix equality.* We will refer to (3) as *Hua's matrix inequality* and (4) *Hua's determinant inequality.*

Our purpose in this paper is to present a new proof of Hua's matrix equality (1), and a new proof of Hua's matrix inequality (3), both using Schur complements, and we do this in Sections 3 and 4, after presenting necessary background theory in Section 2. Perhaps more interesting than the proofs are the tools we introduce to obtain them — a general result on Schur complements for the first proof, and a generalization of Sylvester's law of inertia for the second. These are more widely useful, and appear to be new to us in the literature. In Section 5 we comment on the most general situation when the matrix A is not necessarily strictly contractive, in Section 6 we use our result on Schur complements to generalize Hua's matrix equality (1) and prove a new inequality.

2. Background Theory

We will use the following material on Schur complements.

When the matrix C in the top-left corner of the partitioned matrix

$$G = \begin{pmatrix} C & D \\ E & F \end{pmatrix} \tag{8}$$

is square and nonsingular, then the *Schur complement of* C *in* G is defined to be

$$(G/C) = F - EC^{-1}D. \tag{9}$$

The term Schur complement and the notation (G/C) in (9) were introduced in 1968 by E. Haynsworth [11] following the seminal 1917 paper [22] by I. Schur.

When C is square and singular, or rectangular, (9) can be generalized. We give a constructive approach that makes the Schur complement, its main properties, and its generalization and the resulting necessary conditions, obvious. If C is $m \times n$ and F is $p \times q$ in (8), and C^- denotes any generalized inverse of C satisfying $CC^-C = C$, then

$$\begin{pmatrix} I_m & 0 \\ -EC^- & I_p \end{pmatrix} \begin{pmatrix} C & D \\ E & F \end{pmatrix} \begin{pmatrix} I_n & -C^-D \\ 0 & I_q \end{pmatrix} = \begin{pmatrix} C & D - CC^-D \\ E - EC^-C & G_1 \end{pmatrix}, \tag{10}$$

where

$$G_1 = F - EC^-D - EC^-(D - CC^-D) = F - EC^-D - (E - EC^-C)C^-D.$$

When C is invertible the off-diagonal blocks in the right-hand side of (10) are zero and G_1 is the Schur complement (9). As we will see, the vanishing of these two blocks gives the Schur complement its useful properties, so we will require in the general case that

$$CC^-D = D \quad \text{and} \quad EC^-C = E, \tag{11}$$

or equivalently that

$$\mathcal{C}(D) \subseteq \mathcal{C}(C) \text{ and } \mathcal{R}(E) \subseteq \mathcal{R}(C) \;\Leftrightarrow\; \operatorname{rank}(C, D) = \operatorname{rank}(C) = \operatorname{rank}\begin{pmatrix} C \\ E \end{pmatrix}, \tag{12}$$

where $\mathcal{C}(\cdot)$ denotes column space (range) and $\mathcal{R}(\cdot)$ row space. Carlson [3] refers to these conditions as the "natural setting for results in generalized Schur complements". We will refer to (11) and (12) as the *natural conditions for* C *in* G, or just the *natural conditions*. We define the *generalized Schur complement* of C in G as G_1 above when these natural conditions (11) hold

$$(G/C) = F - EC^-D. \tag{13}$$

Note that the natural conditions are necessary and sufficient for the generalized Schur complement (13) to be unique [3, 21, 24]. And the natural conditions automatically hold when C is invertible or G is (Hermitian) nonnegative definite.

When either of the natural conditions (11) holds, and C and F are square in (8), we see from (10) that

$$\det(G) = \det(C) \cdot \det(G/C).$$

When both the natural conditions (11) hold, *rank is additive* on the Schur complement (G/C) in the sense that

$$\operatorname{rank}(G) = \operatorname{rank}(C) + \operatorname{rank}(G/C). \tag{14}$$

This rank additivity result is obvious from (10), and was apparently first established in 1946 by Guttman [9] for C square and nonsingular and then extended by Meyer [19] and Marsaglia & Styan [17] to C square and singular or rectangular – but subject to the natural conditions (11) or (12).

We define the *inertia* $\operatorname{In}(M)$ of a Hermitian matrix M to be the ordered triple $\{i_+,\ i_-,\ i_0\}$, where i_+ denotes the number of positive eigenvalues of M, i_- the number of negative eigenvalues, and i_0 the number of zero eigenvalues.

Crucial to studies of inertia is Sylvester's law of inertia:

$$\operatorname{In}(M) = \operatorname{In}(T^*MT), \tag{15}$$

which was first established in 1852 by Sylvester [25]; cf. e.g., Mirsky [20, p. 377]. In (15) the matrix M is Hermitian and T is square and nonsingular; in Section 4 we extend (15) to situations where the matrix T is square and singular or rectangular and show how this generalized Sylvester's law of inertia may be used to provide a new proof of Hua's matrix inequality (3) using Schur complements.

The inertia is additive on the Schur complement (G/C) in the sense that when the matrix G in (8) is Hermitian with C square (and so also Hermitian), then

$$\operatorname{In}(G) = \operatorname{In}(C) + \operatorname{In}(G/C), \tag{16}$$

With C nonsingular this is obvious from (10) and (15). Haynsworth [11] established (16) with C nonsingular while Carlson, Haynsworth & Markham [4] showed that (16) remains true when C is square and singular provided that the (generalized) Schur complement (G/C) is defined as in (13) with the natural conditions (11) or (12) which, since now $C = C^*$ and $E = D^*$ with G Hermitian, can be shown to simplify to

$$CC^-D = D \quad \Leftrightarrow \quad \mathcal{C}(D) \subseteq \mathcal{C}(C) \quad \Leftrightarrow \quad \operatorname{rank}(C, D) = \operatorname{rank}(C). \tag{17}$$

Then with $T = \begin{pmatrix} I & -C^-D \\ 0 & I \end{pmatrix}$, (16) follows from $\operatorname{In}(G) = \operatorname{In}(T^*GT)$ due to (15).

The theory for the Schur complement of F in G can be similarly found by applying permutations to bring F to the top-left corner. For more on Schur complements and their applications see the survey articles by Carlson [3], Cottle [5], Ouellette [21], and Styan [24], and the book edited by Zhang [28].

3. A useful result on Schur complements

For our new proof of Hua's matrix equality (1) we will need the following lemma.

Lemma 1. *Let G be a partitioned matrix defined as in (8), with C $m \times n$ in*

$$\begin{pmatrix} I_m & 0 \\ X_1 & Y_1 \end{pmatrix} \begin{pmatrix} C & D \\ E & F \end{pmatrix} \begin{pmatrix} I_n & X_2 \\ 0 & Y_2 \end{pmatrix} = Z_1 G Z_2.$$

If the natural conditions (11) hold and (G/C) is the Schur complement of C in G, then the natural conditions hold for C in Z_1GZ_2, and the Schur complement (Z_1GZ_2/C) of C in Z_1GZ_2 is invariant for all choices of X_1 and X_2, and

$$(Z_1GZ_2/C) = Y_1(G/C)Y_2. \tag{18}$$

Proof. Multiplying out

$$Z_1GZ_2 = \begin{pmatrix} C & CX_2 + DY_2 \\ X_1C + Y_1E & X_1CX_2 + Y_1EX_2 + X_1DY_2 + Y_1FY_2 \end{pmatrix}$$

and checking the natural conditions for C in Z_1GZ_2, we see that (Z_1GZ_2/C) is well defined and unique. Using (11) and by a direct computation we obtain (18). ∎

Corollary 1. *When Y_1 and Y_2 are invertible, if the natural conditions hold for C in Z_1GZ_2, then they hold for C in G, and $(G/C) = Y_1^{-1}(Z_1GZ_2/C)Y_2^{-1}$.*

Proof. These follow by applying Lemma 1 to $Z_1^{-1}(Z_1GZ_2)Z_2^{-1} = G$. ∎

We note that a similar result to Lemma 1 has appeared in [26, 27]. We now use Lemma 1 to give a new proof of (1). The generalized inverse notation in the proof is for later use.

Proof of (1). Let $I - A^*A$ be nonsingular, so $(I - A^*A)^- = (I - A^*A)^{-1}$. Let

$$H_1 = \begin{pmatrix} I - A^*A & I - A^*B \\ I - B^*A & I - B^*B \end{pmatrix}, \tag{19}$$

$$H_2 = \begin{pmatrix} I & 0 \\ -I & I \end{pmatrix} H_1 \begin{pmatrix} I & -I \\ 0 & I \end{pmatrix} = \begin{pmatrix} I - A^*A & A^*(A-B) \\ (A-B)^*A & -(A-B)^*(A-B) \end{pmatrix}. \tag{20}$$

The Hua matrix H is the Schur complement of $I - A^*A$ in H_1:

$$H = (H_1/(I - A^*A)) = (I - B^*B) - (I - B^*A)(I - A^*A)^-(I - A^*B) \tag{21}$$

and from Lemma 1 this equals the Schur complement of $I - A^*A$ in H_2:

$$(H_2/(I - A^*A)) = -(A-B)^*[I_m + A(I - A^*A)^-A^*](A-B) \tag{22}$$
$$= -(A-B)^*(I_m - AA^*)^-(A-B) \tag{23}$$

from (2) with A and A^* interchanged. So the Hua matrix

$$H = (I - B^*B) - (I - B^*A)(I_n - A^*A)^{-1}(I - A^*B)$$
$$= -(A-B)^*(I_m - AA^*)^{-1}(A-B) \tag{24}$$

and our easy proof of (1) follows at once. ∎

A similar proof is obtained using a technique by Embry [7]. If we replace B by $A - N$ in the right-hand side of (21) then, after some cancellation, it becomes $-N^*[I_m + A(I - A^*A)^{-1}A^*]N$, which is the second expression in (23) since now $N = A - B$. In fact Embry only needed to show (21) could not be nonnegative definite in [7], and did not go on to prove (24).

Our Schur complement approach allows us to generalize Hua's matrix equality (1) to the case of singular $I - A^*A$ with minimal additional effort.

Theorem 1. *Let A and B be be any $m \times n$ complex matrices such that*

$$\mathcal{C}(I - A^*B) \subseteq \mathcal{C}(I - A^*A). \tag{25}$$

Then

$$H = (I - B^*B) - (I - B^*A)(I - A^*A)^-(I - A^*B)$$
$$= -(A-B)^*(I_m - AA^*)^-(A-B). \tag{26}$$

Proof. From (19) we see that (25) is the natural condition for $I - A^*A$ in Hermitian H_1. So (21) is the unique Schur complement of $I - A^*A$ in H_1. Moreover, since

$$\mathrm{rank}(I - A^*B, I - A^*A) = \mathrm{rank}(I - A^*A)$$

equivalently,

$$\mathrm{rank}(A^*A - A^*B, I - A^*A) = \mathrm{rank}(I - A^*A),$$

we see that (25) is equivalent to

$$\mathcal{C}(A^*(A - B)) \subseteq \mathcal{C}(I - A^*A),$$

Thus if (25) holds, the Schur complement of $I - A^*A$ in H_2 is unique. By Lemma 1,

$$H = (H_1/(I - A^*A)) = (H_2/(I - A^*A)).$$

We now only need to show that (22) is equal to (23). Notice that

$$(I - A^*A)[I + A^*(I - AA^*)^- A](I - A^*A) = (I - A^*A)^2 + A^*(I - AA^*)A = I - A^*A.$$

This says that $I + A^*(I - AA^*)^- A$ is a generalized inverse of $I - A^*A$ of the form $(I - A^*A)^-$, completing the proof. This also generalizing equation (2). ∎

We note that the identity (26) is understood as follows: for any generalized inverse $(I - A^*A)^-$ on the left-hand side, there is a generalized inverse $(I_m - AA^*)^-$ on the right-hand side such that the expressions on both sides agree.

We now show Hua's matrix inequality (3) also generalizes to the singular case.

Corollary 2. *Let C^- denote any generalized inverse of C satisfying $CC^-C = C$. If $I - A^*A$ is nonnegative definite and (25) holds, then*

$$I - B^*B \le (I - B^*A)(I - A^*A)^-(I - A^*B). \tag{27}$$

Proof. When (25) holds, there exists $n \times n$ R so that $I - A^*B = (I - A^*A)R$, then

$$A^*(A - B) = (I - A^*B) - (I - A^*A) = (I - A^*A)(R - I),$$

which with (22) shows

$$(H_2/(I - A^*A)) = -(A - B)^*(A - B) - (R - I)^*(I - A^*A)(R - I) \le 0.$$

Since it is equal to the right hand side of (26), the inequality (27) follows. ∎

When the natural conditions (25) hold, we can combine (14) with (19) to obtain

$$\det(H_1) = \det(I - A^*A)\det(H). \tag{28}$$

This is zero if $I - A^*A$ is singular, otherwise the natural conditions must hold and

$$\begin{aligned} (-1)^n\det(H_1) &= \det(I - A^*A)\det[(A - B)^*(I_m - AA^*)^{-1}(A - B)] \\ &= |\det(A - B)|^2, \quad \text{if } A \text{ and } B \text{ are } n \times n, \end{aligned} \tag{29}$$

since A^*A and AA^* have the same nonzero eigenvalues. Note that here A need not be contractive. Whether A and B are square or not, if they are both (strictly) contractive, we see from (6) and the above that (note the "gap" in (4))

$$(-1)^n\det(H_1) \le |\det(I - A^*B)|^2 - \det(I - A^*A) \cdot \det(I - B^*B).$$

As Hua [14] observed, when $I - A^*A$ and $I - B^*B$ are *both* positive definite then $I - A^*B$ is nonsingular and the nonpositive definiteness (no positive eigenvalues) of the Hermitian Hua matrix H defined in (1), see (24), is equivalent to

$$\tilde{H}_1 = \begin{pmatrix} (I - A^*A)^{-1} & (I - B^*A)^{-1} \\ (I - A^*B)^{-1} & (I - B^*B)^{-1} \end{pmatrix} \ge 0. \tag{30}$$

We note that the matrix $I - A^*B$ now appears in the lower left rather than the upper right position as in (19), and that its nonsingularity follows from [18, p. 246]

$$|\lambda(A^*B)| \leq \sigma_{\max}(A^*B) \leq \sigma_{\max}(A) \cdot \sigma_{\max}(B) < 1,$$

Here $\lambda(\cdot)$ denotes any eigenvalue, and $\sigma_{\max}(\cdot)$ the largest singular value.

The inequality (30) follows from the inertia additivity for the symmetrically permuted block matrix,

$$\text{(31)} \qquad \mathrm{In}\begin{pmatrix} (I - B^*B)^{-1} & (I - A^*B)^{-1} \\ (I - B^*A)^{-1} & (I - A^*A)^{-1} \end{pmatrix} = \mathrm{In}[(I - B^*B)^{-1}] + \mathrm{In}(K)$$

where

$$K = (I - A^*A)^{-1} - (I - B^*A)^{-1}(I - B^*B)(I - A^*B)^{-1}.$$

It follows from (21), then (15), that

$$\text{(32)} \qquad (I - B^*A)K(I - A^*B) = -H, \qquad \mathrm{In}(K) = \mathrm{In}(-H).$$

Since the Hua matrix H is nonpositive definite, K is nonnegative definite, (30) follows from (31). Using (14), (32), (28), the determinant of the matrix in (30) is

$$\begin{aligned} \det(\tilde{H}_1) &= \det((I - B^*B)^{-1})\det(K) \\ &= \det(-H)/[\det(I - B^*B) \cdot |\det(I - A^*B)|^2] \\ &= \frac{(-1)^n \det(H_1)}{\det(I - A^*A)\det(I - B^*B)|\det(I - A^*B)|^2}, \end{aligned}$$

which can be combined with (29) and (30) to give more information on $\det(\tilde{H}_1)$.

4. A generalization of Sylvester's law of inertia

We have seen that an elegant approach to proving Hua's matrix and determinant inequalities is to first prove Hua's matrix equality, since the inequalities follow from this. But our early work showed we could prove the matrix inequality directly using a generalized version of Sylvester's law of inertia, and since this is interesting in its own right, we include it here.

Sylvester's law of inertia (15) holds for congruent transformations T^*MT with T square and nonsingular. Here we give a similar result for T square and singular, or rectangular. For the Hermitian matrix M we will write, cf. (15),

$$\begin{aligned} i_+(M) &= \text{the number of positive eigenvalues of } M, \\ i_-(M) &= \text{the number of negative eigenvalues of } M, \\ i_0(M) &= \text{the number of zero eigenvalues of } M, \end{aligned}$$

and $i_\pm(M)$ for either $i_+(M)$ or $i_-(M)$ consistently within an expression.

Lemma 2. *Let M be $p \times p$ Hermitian. Then for any $p \times q$ matrix T of rank r*

$$\text{(33)} \qquad 0 \leq i_-(T^*MT) + i_+(T^*MT) = \mathrm{rank}(T^*MT) \leq \min\{r, \mathrm{rank}(M)\},$$

$$\text{(34)} \qquad i_\pm(M) - (p - r) \leq i_\pm(T^*MT) \leq i_\pm(M).$$

Proof. Since M is Hermitian, (33) follows immediately. Let

$$\lambda_p(M) \leq \cdots \leq \lambda_2(M) \leq \lambda_1(M)$$

denote the ordered, necessarily real, eigenvalues of M. When $r = 0$, (34) is trivially true. So let $r > 0$ and let T have singular value decomposition $T = U\Sigma V^* = U_r\Sigma_r V_r^*$, and let $M_r = U_r^* M U_r$, where U_r is $p \times r$ and V_r is $q \times r$ with $U_r^* U_r = V_r^* V_r = I_r$, while Σ_r is $r \times r$ diagonal positive definite. It follows that

$$V^*T^*MTV = \begin{pmatrix} \Sigma_r M_r \Sigma_r & 0 \\ 0 & 0 \end{pmatrix}. \tag{35}$$

By the Poincaré separation theorem, cf. e.g., [12, Cor. 4.3.16], [23], we have

$$\lambda_{k+p-r}(M) \le \lambda_k(U_r^* M U_r) = \lambda_k(M_r) \le \lambda_k(M), \quad k = 1, 2, \ldots, r. \tag{36}$$

We see from this that $U_r^* M U_r = M_r$ cannot have, respectively, any more negative eigenvalues, or any more positive eigenvalues, than M and so the upper bound in (34) is established. We now consider $i_+(T^*MT) = i_+(M_r)$ (by (35)). If M has exactly $k + p - r$ positive eigenvalues then M_r has at least k positive eigenvalues and so $i_+(M_r) \ge k = i_+(M) - (p-r)$; a similar argument for $i_-(T^*MT) = i_-(M_r)$ holds in terms of the negative eigenvalues and so our proof is complete. ∎

We can use (33) and (34) to obtain bounds on $i_0(T^*MT) = q - i_+(T^*MT) - i_-(T^*MT)$ in special cases. When T has full row rank $r = p$, (34) collapses to

$$i_\pm(T^*MT) = i_\pm(M) \tag{37}$$

and then

$$i_0(T^*MT) = (q-p) + i_0(M). \tag{38}$$

When $q = p = r$ then T is square and nonsingular and (37) and (38) together become Sylvester's law of inertia (15).

We now show how this generalized Sylvester's law of inertia may be used to provide a direct proof of Hua's matrix inequality (3) using Schur complements.

Another Proof of (3). With $m \times n$ strictly contractive A, we define

$$T = \begin{pmatrix} I_n & I_n \\ A & B \end{pmatrix}, \qquad \text{where} \quad \operatorname{rank}(T) = n + \operatorname{rank}(B - A), \tag{39}$$

by rank additivity on the Schur complement $(T/I_n) = B - A$, cf. (14). Then

$$H_1 = T^* \begin{pmatrix} I_n & 0 \\ 0 & -I_m \end{pmatrix} T = \begin{pmatrix} I_n - A^*A & I_n - A^*B \\ I_n - B^*A & I_n - B^*B \end{pmatrix} \tag{40}$$

and the Hua matrix H, cf. (1), is the Schur complement of $I_n - A^*A$ in H_1, cf. (21). But inertia is additive on the Schur complement H, cf. (16), so

$$i_+(H_1) = n + i_+(H) \quad \text{and} \quad i_-(H_1) = i_-(H). \tag{41}$$

This with (33) in Lemma 2 and (39) gives

$$\begin{aligned} 0 \le i_-(H_1) + i_+(H_1) = n + i_-(H) + i_+(H) &\le \operatorname{rank}(T) = n + \operatorname{rank}(B - A), \\ i_-(H) + i_+(H) &\le \operatorname{rank}(B - A). \end{aligned}$$

With (34) in Lemma 2, and (39),

$$i_\pm \begin{pmatrix} I_n & 0 \\ 0 & -I_m \end{pmatrix} - (m - \operatorname{rank}(B - A)) \le i_\pm(H_1) \le i_\pm \begin{pmatrix} I_n & 0 \\ 0 & -I_m \end{pmatrix},$$

where using (41) the right-hand side result for i_+ gives $n + i_+(H) \le n$, and the left-hand side result for i_- gives $\operatorname{rank}(B - A) \le i_-(H)$. So $i_+(H) = 0$ and H has no positive eigenvalues, $i_-(H) = \operatorname{rank}(H) = \operatorname{rank}(B - A)$, and our proof of Hua's matrix inequality (3) using Schur complements is complete. ∎

When $m = n$, (14) with (39) and (40) gives another proof of (29):

$$\det(H_1) = (-1)^n |\det(T)|^2 = (-1)^n |\det(B - A)|^2.$$

5. Hua's matrix without contractiveness

Since inertia is additive on the Schur complement, cf. (16), equations (19) and (21) show for the Hua matrix H that with

$$(42)\quad H_1 = \begin{pmatrix} I - A^*A & I - A^*B \\ I - B^*A & I - B^*B \end{pmatrix}, \qquad \mathrm{In}(H_1) \;=\; \mathrm{In}(I - A^*A) + \mathrm{In}(H)$$

holds for A not necessarily strictly contractive, but it does require that $I - A^*A$ be nonsingular (which we also need for the Hua matrix H to be defined). But from (15) with

$$\begin{aligned} \tilde{H}_2 &= \begin{pmatrix} -I & I \\ I & 0 \end{pmatrix} H_1 \begin{pmatrix} -I & I \\ I & 0 \end{pmatrix} \\ &= \begin{pmatrix} -(A-B)^*(A-B) & (A-B)^*A \\ A^*(A-B) & I - A^*A \end{pmatrix}, \end{aligned}$$

$$(43)\quad \mathrm{In}(H_1) \;=\; \mathrm{In}(\tilde{H}_2) \;=\; \mathrm{In}[-(A-B)^*(A-B)] + \mathrm{In}(I - A^*Q_{A-B}A)$$

holds for *any* complex matrix A not necessarily (strictly) contractive. Here the Hermitian idempotent matrix

$$Q_{A-B} = I - (A-B)[(A-B)^*(A-B)]^-(A-B)^* = I - P_{A-B}$$

is the orthogonal projector onto the null space of $(A-B)^*$ or onto the orthocomplement of the column space of $A-B$, and the natural conditions for the Schur complement $(\tilde{H}_2 / -(A-B)^*(A-B))$ can be seen to hold.

Combining (42) and (43) yields

$$(44)\qquad \mathrm{In}(H) = \{0,\ c,\ n-c\} + \mathrm{In}(I - A^*QA) - \mathrm{In}(I - A^*A)$$

with $I - A^*A$ nonsingular, where

$$c = \mathrm{rank}(A-B) \quad \text{and} \quad Q = Q_{A-B}.$$

Hence, when $I - A^*A$ is nonsingular, we find that

$$(45)\qquad \begin{aligned} i_+(H) &= i_+(I - A^*QA) - i_+(I - A^*A) &&\geq 0 \\ i_-(H) &= c + i_-(I - A^*QA) - i_-(I - A^*A) &&\leq c \\ i_0(H) &= n - c + i_0(I - A^*QA) &&\geq n - c. \end{aligned}$$

The first two equalities and the first inequality in (45) follow at once from (44) while the second inequality (as well as the first) follows from

$$(46)\qquad A^*QA \leq A^*A,$$

which holds since $I - Q$ is Hermitian idempotent and hence nonnegative definite; equality holds in (46) if and only if $QA = A$. The last equality in (45) follows from the nonsingularity of $I - A^*A$.

The inequality (46) shows that QA is strictly contractive whenever A is and then equality holds throughout all of (45). To go the other way we see that when equality holds throughout (45) then $I - A^*QA$ must be nonsingular and, in addition, $i_\pm(I - A^*QA) = i_\pm(I - A^*A)$ or equivalently QA and A have the same numbers of (nonzero) singular values, respectively greater than and less than 1; neither A nor QA need, however, be strictly contractive. As an example, let

$$A = \begin{pmatrix} 2 & 0 \\ 0 & \frac{1}{2} \end{pmatrix}, \quad B = \begin{pmatrix} 2 & 0 \\ 0 & -\frac{1}{2} \end{pmatrix}, \quad A - B = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix},$$

so that

$$Q = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad QA = \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix}, \quad I - A^*QA = \begin{pmatrix} -3 & 0 \\ 0 & 1 \end{pmatrix}, \quad I - A^*A = \begin{pmatrix} -3 & 0 \\ 0 & \frac{3}{4} \end{pmatrix}.$$

The equalities in (45), or equivalently in the inertia equality (44), lead to

$$\text{rank}(H) = \text{rank}(A-B) - [n - \text{rank}(I - A^*QA)] \le \text{rank}(A-B) \tag{47}$$

with equality if and only if $I - A^*QA$ is nonsingular; curiously, this does not (seem to) require that $i_\pm(I - A^*QA) = i_\pm(I - A^*A)$; of course we do need (our earlier assumption) that $I - A^*A$ is nonsingular. When A is strictly contractive, however, the rank of H is maximal—equal to rank$(A - B)$.

6. Generalizing Hua's matrix equality

The factorization of H_1 in (40) and the proof of (1) in Section 3 suggest a nice generalization of Hua's matrix equality (1). The following proof suggests the power of the Schur complement approach used in Section 3.

Theorem 2. *Let W, X, Y, and Z be any $m \times n$ complex matrices such that*

$$\mathcal{C}(I + X^*Z) \subseteq \mathcal{C}(I + X^*Y) \quad \text{and} \quad \mathcal{R}(I + W^*Y) \subseteq \mathcal{R}(I + X^*Y). \tag{48}$$

Then the "generalized Hua matrix"

$$\begin{aligned} H_g &= I + W^*Z - (I + W^*Y)(I + X^*Y)^-(I + X^*Z) \\ &= (W - X)^*(I_m + YX^*)^-(Z - Y). \end{aligned} \tag{49}$$

Proof. Let

$$J_1 = \begin{pmatrix} I & I \\ X & W \end{pmatrix}^* \begin{pmatrix} I & I \\ Y & Z \end{pmatrix} = \begin{pmatrix} I + X^*Y & I + X^*Z \\ I + W^*Y & I + W^*Z \end{pmatrix}.$$

By (12), the inclusions in (48) are the natural conditions for $I + X^*Y$ in J_1. So

$$H_g^{(1)} = I + W^*Z - (I + W^*Y)(I + X^*Y)^-(I + X^*Z)$$

is the unique Schur complement of $I + X^*Y$ in J_1. Furthermore, by Lemma 1,

$$H_g^{(2)} = (W - X)^*[I_m - Y(I + X^*Y)^-X^*](Z - Y)$$

is the unique Schur complement of $I + X^*Y$ in

$$J_2 = \begin{pmatrix} I & 0 \\ -I & I \end{pmatrix} J_1 \begin{pmatrix} I & -I \\ 0 & I \end{pmatrix} = \begin{pmatrix} I + X^*Y & X^*(Z - Y) \\ (W - X)^*Y & (W - X)^*(Z - Y) \end{pmatrix},$$

and $H_g^{(1)} = H_g^{(2)}$. On the other hand,

$$\begin{aligned} &(I_m + YX^*)[I_m - Y(I + X^*Y)^-X^*](I_m + YX^*) = \\ &\qquad (I_m + YX^*)^2 - Y(I + X^*Y)X^* = I_m + YX^* \end{aligned}$$

so $I_m - Y(I + X^*Y)^-X^*$ is a generalized inverse of $I + YX^*$ of the form $(I_m + YX^*)^-$ (which generalizes Duncan's formula (2) further than in Theorem 1). This shows

$$H_g^{(2)} = (W - X)^*(I_m + YX^*)^-(Z - Y)$$

for some generalized inverse $(I_m + YX^*)^-$, so (49) holds, completing the proof. ∎

Whenever $I + X^*Y$ is nonsingular (and so $I_m + YX^*$ is also nonsingular) the natural conditions (48) automatically hold.

When $W = B$, $X = A$, $Y = -A$, and $Z = -B$, then (48) reduces to (25) and (49) reduces to (26), our generalization of (1):

$$\begin{aligned} H &= (I - B^*B) - (I - B^*A)(I_n - A^*A)^-(I - A^*B) \\ &= -(A - B)^*(I_m - AA^*)^-(A - B). \end{aligned}$$

When $W = -B$, $X = A$, $Y = A$, $Z = -B$, however, then (49) yields the *alternate Hua matrix*:

$$\begin{aligned} H_a &= (I + B^*B) - (I - B^*A)(I_n + A^*A)^{-1}(I - A^*B) \\ &= (A + B)^*(I_m + AA^*)^{-1}(A + B), \end{aligned}$$

which holds for *all* $m \times n$ matrices A and B (since $I_n + A^*A$ and $I_m + AA^*$ are *always* positive definite). Hence

$$I + B^*B \geq (I - B^*A)(I + A^*A)^{-1}(I - A^*B)$$

holds for all $m \times n$ matrices A and B, with equality if and only if $A = -B$.

References

[1] T. Ando, Hua-Marcus inequalities, Linear and Multilinear Algebra 8 (1980), 347–352.

[2] R. Bhatia, Matrix Analysis, Springer-Verlag, New York, 1997.

[3] D. Carlson, What are Schur complements anyway?, Linear Algebra Appl. 59 (1984), 188–193 & 4 (1986), 257–275.

[4] D. Carlson, E. Haynsworth & T. Markham, A generalization of the Schur complement by means of the Moore-Penrose inverse, SIAM J. Appl. Math. 26 (1974), 169–175.

[5] R. W. Cottle, Manifestations of the Schur complement, Linear Algebra Appl. 8 (1974), 189–211.

[6] W. J. Duncan, Some devices for the solution of large sets of simultaneous linear equations (with an appendix on the reciprocation of partitioned matrices), The London, Edinburgh and Dublin Philosophical Magazine and Journal of Science, Seventh Series 35 (1944), 660–670.

[7] M. R. Embry, Solution to Problem 5233: Self-adjoint operators on Hilbert space, Amer. Math. Monthly 72 (1965), 1031.

[8] Wannan Gu, Hua Loo-Keng, Hebei People's Publishing House, 1985.

[9] L. Guttman, Enlargement methods for computing the inverse matrix, Ann. Math. Statist. 17 (1946), 336–343.

[10] H. Halberstam, Ed., Loo-Keng Hua selected papers, Springer-Verlag, Berlin, 1983.

[11] E. V. Haynsworth, Determination of the inertia of a partitioned Hermitian matrix, Linear Algebra Appl. 1 (1968), 73–81.

[12] R. A. Horn & C. R. Johnson, Matrix Analysis, Corrected Reprint Edition, Cambridge University Press, 1990.

[13] R. A. Horn & C. R. Johnson, Topics in Matrix Analysis, Corrected Reprint Edition, Cambridge University Press, 1994.

[14] L.-K. Hua, Inequalities involving determinants (in Chinese), Acta Math. Sinica 5 (1955), 463–470. [Translated into English: Transl. Amer. Math. Soc. Ser. II 32 (1963), 265–272.]

[15] M. Marcus, On a determinantal inequality, Amer. Math. Monthly 65 (1958), 266–268.

[16] M. Marcus & M. Minc, A Survey of Matrix Theory and Matrix Inequalities, 2nd Reprint Edition, Dover, New York, 1992.

[17] G. Marsaglia & G. P. H. Styan, Equalities and inequalities for ranks of matrices, Linear and Multilinear Algebra 2 (1974), 269–292.

[18] A. W. Marshall & I. Olkin, Inequalities: Theory of Majorization and its Applications, Academic Press, San Diego, 1979.

[19] C. D. Meyer, Jr., Generalized inverses and ranks of block matrices, SIAM J. Appl. Math. 25 (1973), 597–602.

[20] L. Mirsky, An Introduction to Linear Algebra, 2nd Corrected Reprint Edition, Dover, New York, 1990.

[21] D. V. Ouellette, Schur complements and statistics, Linear Algebra Appl. 36 (1981), 187–295.

[22] I. Schur, Über Potenzreihen, die im Innern des Einheitskreises beschränkt sind: I, J. Reine Angew. Math. 147 (1917), 205–232. [Reprinted in: Issai Schur: Gesammelte Abhandlungen (Alred Brauer & Hans Rohrbach, eds.), Springer-Verlag, Berlin 2 (1973), 137–164.]

[23] A. J. Scott & G. P. H. Styan, On a separation theorem for generalized eigenvalues and a problem in the analysis of sample surveys, Linear Algebra Appl. 70 (1985), 209–224.

[24] G. P. H. Styan, Schur complements and linear statistical models, Proc. First Internat. Tampere Seminar on Linear Statistical Models and their Applications: Tampere, Finland, August–September 1983 (T. Pukkila & S. Puntanen, eds.), Dept. of Mathematical Sciences, University of Tampere, 37–75 (1985).

[25] J. J. Sylvester, A demonstration of the theorem that every homogeneous quadratic polynomial is reducible by real orthogonal substitutions to the form of a sum of positive and negative squares, The London, Edinburgh and Dublin Philosophical Magazine and Journal of Science, Fourth Series 4 (1852), 138–142. [Reprinted in: The Collected Mathematical Papers of James Joseph Sylvester (H. F. Baker, ed.), Cambridge University Press I (1904), 378–381.]

[26] B.-Y. Wang, X. Zhang & F. Zhang, Some inequalities on generalized Schur complements, Linear Algebra Appl. 302-303(1999), 163–172.

[27] F.-Z. Zhang, A matrix identity on the Schur complement, Linear Mutilinear Algebra, Vol. 52(2004), No. 5, pp. 367–373.

[28] F. Zhang. The Schur Complements and Its Applications, Springer, New York, 2005.

School of Computer Science, McGill University, 3480 rue University Street, Montréal, Québec, Canada H3A 2A7

E-mail: chris@cs.mcgill.ca

Department of Mathematics and Statistics, McGill University, 805 ouest, rue Sherbrooke Street West, Montréal, Québec, Canada H3A 2K6

E-mail: styan@math.mcgill.ca

Department of Mathematics, Beijing Normal University, Beijing 100875, China

E-mail: bywang@math.bnu.edu.cn

Division of Math, Science and Technology, Nova Southeastern University, 3301 College Ave, Fort Lauderdale, FL 33314, USA

E-mail: zhang@nova.edu