

INDEX OF HADAMARD MULTIPLICATION BY POSITIVE MATRICES II

Gustavo Corach* and Demetrio Stojanoff†

Gustavo Corach

Instituto Argentino de Matemática
Saavedra 15 (1083), Buenos Aires, Argentina
e-mail: gcorach@ciudad.com.ar

and

Depto. de Matemática, Facultad de Ingeniería,
Universidad de Buenos Aires, Argentina.

Demetrio Stojanoff

Depto. de Matemática
UNLP, 1 y 50 (1900), La Plata, Argentina
e-mail: demetrio@mate.dm.uba.ar

Abstract

For each $n \times n$ positive semidefinite matrix A we define the minimal index $I(A) = \max\{\lambda \geq 0 : A \circ B \succeq \lambda B \text{ for all } B \succeq 0\}$ and, for each norm N , the N -index $I_N(A) = \min\{N(A \circ B) : B \succeq 0 \text{ and } N(B) = 1\}$, where $A \circ B = [a_{ij}b_{ij}]$ is the Hadamard or Schur product of $A = [a_{ij}]$ and $B = [b_{ij}]$ and $B \succeq 0$ means that B is a positive semidefinite matrix. A comparison between these indexes is done, for different choices of the norm N . As an application we find, for each bounded invertible selfadjoint operator S on a Hilbert space, the best constant $M(S)$ such that $\|STS + S^{-1}TS^{-1}\| \geq M(S)\|T\|$ for all $T \succeq 0$.

Key words: Hadamard product, positive semidefinite matrices, norm inequalities.

AMS subject classifications: Primary 47A30, 47B15.

*Supported by UBACYT TX79 and PIP 4463 (CONICET)

†Supported by UBACYT TW49 and PIP 4463 (CONICET)

1 Introduction

Given $A = [a_{ij}], B = [b_{ij}] \in M_n = M_n(\mathbb{C})$, the algebra of $n \times n$ matrices over \mathbb{C} , denote by $A \circ B$ the Hadamard product $[a_{ij}b_{ij}]$. In this paper $A \succeq 0$ means that A is positive semidefinite ; $P_n = \{A \in M_n : A \succeq 0\}$ denotes the set of positive semidefinite matrices.

Every $A \in M_n$ defines a linear map $\Phi_A : M_n \rightarrow M_n$ given by $\Phi_A(B) = A \circ B$, for $B \in M_n$. By Schur's product theorem [22] (see also [14, 7.5.3]) $A \circ B \in P_n$ if $A, B \in P_n$, so that Φ_A is a positive linear map. Actually it is completely positive, i.e., the inflation map $\Phi_A^{(m)}$, which acts entrywise as Φ_A on $M_m(M_n)$, is positive for all $m \in \mathbb{N}$; see [20, Prop. 1.2]. In [23], the second author studied conditions under which

$$\max\{\lambda \geq 0 : \Phi_A(B) \succeq \lambda B, \forall B \in P_n\} = \inf\{\|\Phi_A(B)\| : B \in P_n, \|B\| = 1\}.$$

The problem comes from the theory of conditional expectations. A *conditional expectation* on a C^* -algebra \mathcal{A} is a norm one projection $\mathcal{E} : \mathcal{A} \rightarrow \mathcal{A}$ such that $\mathcal{E}(\mathcal{A})$ is a sub- C^* -algebra of \mathcal{A} . Every conditional expectation \mathcal{E} satisfies the condition

$$\sup\{\lambda \geq 0 : \|\mathcal{E}(a)\| \geq \lambda\|a\|, \forall a \in \mathcal{A}^+\} = \sup\{\lambda \geq 0 : \mathcal{E}(a) \geq \lambda a, \forall a \in \mathcal{A}^+\}, \quad (1)$$

where $\mathcal{A}^+ = \{c \in \mathcal{A} : c \succeq 0\}$. The inverse of this number is called the *index* of \mathcal{E} and it is useful in the classification of inclusions of subalgebras of C^* -algebras. Note that a conditional expectation is completely positive. If $\mathcal{E} : \mathcal{A} \rightarrow \mathcal{A}$ is a completely positive map that is not a conditional expectation, (1) fails in general and the problem arises of characterizing those \mathcal{E} such that (1) holds.

For $A \in P_n$ define the *minimal index* $I(A) = \max\{\lambda \geq 0 : A \circ B \succeq \lambda B, \forall B \in P_n\}$ and the *N-index* $I_N(A) = \max\{\lambda \geq 0 : N(A \circ B) \geq \lambda N(B), \forall B \in P_n\}$ for any given norm N on M_n . We are mainly concerned with Schatten norms $\|\cdot\|_p$ for $p = 1, 2$, and ∞ ; we use the shorter notations I_1, I_2 , and I_{sp} for $I_{\|\cdot\|_1}, I_{\|\cdot\|_2}$, and $I_{\|\cdot\|_\infty}$, respectively. I_{sp} is called the *spectral index*.

If $\mathcal{A} = M_n$, every conditional expectation \mathcal{E} has the form $\mathcal{E}(C) = U\Phi_A(U^*CU)U^*$, where $U \in M_n$ is unitary and $A \in P_n$ is a direct sum of matrices whose diagonal entries are all equal to one. In this case, $\text{Ind}(\mathcal{E})^{-1} = 1/k = I_{sp}(A) = I(A)$, where k is the number of diagonal blocks of A . We remove the inverse in our definition of minimal and N -index in order to avoid complications when the index is zero.

For references on the norm of Φ_A , see [2], [3], [9], [10], [11], [17], [19], [20] and references included therein. There is an extensive bibliography about the index of conditional expectations; see [21] and its references. For a deep study of the index theory of completely positive maps on operator algebras, see [5] and [12].

This paper compares these notions of index and investigates how to compute them. The results obtained are useful in the study of certain operator inequalities. Recall that if $L(\mathcal{H})$ is the algebra of bounded linear operators on a Hilbert space \mathcal{H} and $S \in L(\mathcal{H})$ is a selfadjoint invertible operator, then

$$\|STS + S^{-1}TS^{-1}\| \geq 2\|T\|$$

for all $T \in L(\mathcal{H})$ [4]. It is natural to ask whether 2 is the best constant for each fixed S . Using a reduction to the finite dimensional case and a criterion for computing $I_{sp}(B)$ for matrices $B \in P_n$ such that $B \geq 0$, in terms of the principal submatrices of B (see Corollary 4.6), we are able to find for each S , the best constant $M(S)$ such that $\|STS + S^{-1}TS^{-1}\| \geq M(S)\|T\|$ for all $T \succeq 0$.

In this paper we write $A \geq 0$ for matrices (or vectors) with nonnegative entries. We write $A \succeq B$ or $A \geq B$ if $A - B \succeq 0$ or $A - B \geq 0$, respectively. $R(A)$ is the range of A and $\ker A$ is the kernel of A , where A is thought of as acting on \mathbb{C}^n . A^T is the transpose matrix of A , $\bar{A} = [\bar{a}_{ij}]$ is the conjugate matrix of A , and $A^* = \bar{A}^T$. $\rho(A)$ is the spectral radius of A and A^\dagger is the Moore Penrose pseudoinverse of A . Throughout, p denotes the vector $(1, \dots, 1)^T$ and E denotes the matrix pp^T , which has all its entries equal to 1.

Section 2 contains some elementary characterizations of the minimal index. We prove that, for a given $A \in P_n$, $I(A) > 0$ if and only if $p \in R(A)$; and, in this case, $I(A)^{-1}$ is the spectral radius of $A^\dagger E$.

Section 3 is devoted to a comparison of the minimal index with the spectral index. The main result in this section is the following: if $A \in P_n$, $A \geq 0$, and there exists a vector $u \in A^{-1}(\{p\})$ such that $u \geq 0$, then $I(A) = I_{sp}(A)$. The converse holds if $I(A) \neq 0$, without the hypothesis that $A \geq 0$.

In Section 4 we compare the indexes associated with the spectral and the Frobenius norms. The main result here is that $I_2(A) = I_{sp}(\bar{A} \circ A)^{1/2}$ for every $A \in P_n$. As a consequence of the proof of this result we compute $I_{sp}(B)$ for matrices $B \in P_n$ such that $B \geq 0$, in terms of the principal submatrices of B (see Corollary 4.6). This criterion is the main tool used in Section 5, where we compute the minimal and spectral indexes of $\Lambda = [\lambda_i \lambda_j + \frac{1}{\lambda_i \lambda_j}]$ for any n -tuple of positive numbers $\lambda_1, \dots, \lambda_n$ and use them to find, for each bounded Hermitian invertible operator S on a Hilbert space \mathcal{H} , the number

$$M(S) = \inf\{\|STS + S^{-1}TS^{-1}\| : T \succeq 0, \|T\| = 1\}. \quad (2)$$

For example, if $\|S\| \leq 1$, then $M(S) = \|S\|^2 + \|S\|^{-2}$.

We thank the referees, who helped us to improve several statements and proofs throughout the paper. In particular, the present proofs of Lemma 3.1, item 1 of Theorem 3.3, and the inclusion of the spectral radius in Proposition 2.5 are due to them.

2 Elementary properties of the index

Let us give more detailed definitions:

Definition 2.1 The Hadamard **minimal** index of $A \in P_n$ is

$$\begin{aligned} I(A) &= \max \{ \lambda \geq 0 : A \circ B \succeq \lambda B \quad \forall B \in P_n \} \\ &= \max \{ \lambda \geq 0 : (\Phi_A - \lambda Id)B \succeq 0 \quad \text{for all } B \in P_n \} \\ &= \max \{ \lambda \geq 0 : A - \lambda E \succeq 0 \}. \end{aligned}$$

The last equality follows from the fact that for $C \in M_n$, the map Φ_C is positive if and only if $C \succeq 0$.

Definition 2.2 Given a norm N in M_n , the Hadamard N -index for $A \in P_n$ is

$$\begin{aligned} I_N(A) &= \max \{ \lambda \geq 0 : N(A \circ B) \geq \lambda N(B) \quad \forall B \in P_n \} \\ &= \min \{ N(A \circ B) : B \in P_n \text{ and } N(B) = 1 \}. \end{aligned}$$

The index associated with the spectral norm $\|\cdot\|$ is denoted by $I_{sp}(\cdot)$; we call it the *spectral index*. The index associated with the Frobenius norm $\|\cdot\|_2$ is denoted by $I_2(\cdot)$.

Example 2.3 Let $A = [a_{ij}]$ and $B = [b_{ij}] \in P_n$. Then, if $\|\cdot\|_1$ denotes the trace norm,

$$\|B\|_1 = \text{tr}(B) = \sum_{i=1}^n b_{ii} \quad \text{and} \quad \|A \circ B\|_1 = \text{tr}(A \circ B) = \sum_{i=1}^n a_{ii} b_{ii}.$$

From these identities it is easy to see that, if $I_1(\cdot)$ denotes the associated index, then $I_1(A) = \min_{1 \leq i \leq n} a_{ii}$ for every $A \in P_n$.

Remark 2.4 Estimation of the N -index of a matrix A can be seen as an inequality, namely, $N(A \circ B) \geq I_N(A)N(B)$ for every $B \in P_n$. It would also be interesting to get such inequalities without the restriction $B \succeq 0$ (of course, for matrices A without zero entries). But in this case, the constant involved is the inverse of the norm induced by N of the map Φ_C , where $c_{ij} = a_{ij}^{-1}$. The computation of such norms is well known (see [11], [20] [19], [10], [17], [9]). For the index associated with the Frobenius norm, the computation of an infimum without the restriction $B \succeq 0$ becomes trivial, but with this restriction it is certainly not trivial, as shown in Theorem 4.3.

The minimal index $I(A)$

The index $I(\cdot)$ is called minimal because $I(A) \leq I_N(A)$ for every unitary invariant norm N . Indeed, given $B \in P_n$, then $A \circ B \succeq I(A)B$ and, by Weyl's monotonicity theorem, $s_i(A \circ B) \geq I(A)s_i(B)$, $1 \leq i \leq n$ (where s_i denote the i -th singular value). Therefore $N(A \circ B) \geq I(A)N(B)$ by Ky Fan's dominance theorem; see [15, 3.5.9].

Given $B, C \succeq 0$ the following relation holds:

$$\max\{\alpha \geq 0 : \alpha C \preceq B\} = \|C^{1/2}B^\dagger C^{1/2}\|^{-1} = \rho(B^\dagger C)^{-1} \quad (3)$$

In fact, if B is non singular, (3) follows from [14, 7.7.3]. If B has rank $r < n$, there exist a unitary matrix U and $\Lambda = \begin{bmatrix} \Lambda_1 & 0 \\ 0 & 0 \end{bmatrix}$ such that Λ_1 is an r -by- r invertible matrix

and $B = U\Lambda U^*$. If $\alpha \geq 0$ and $B \succeq \alpha C$ then, setting $D = \begin{bmatrix} D_{11} & D_{12} \\ D_{12}^* & D_{22} \end{bmatrix} = U^*CU$, we

get $\begin{bmatrix} \Lambda_1 & 0 \\ 0 & 0 \end{bmatrix} = \Lambda \succeq \alpha D$ so that $D_{22} = 0$ and, then, $D_{12} = 0$. Therefore $\Lambda_1 \succeq \alpha D_{11}$ and, by the non singular case, $\rho(\Lambda_1^{-1} D_{11}) \leq \alpha$. The result follows by observing that $\rho(\Lambda_1^{-1} D_{11}) = \rho(B^\dagger C)$. Observe also that the block structure of D and the invertibility of Λ_1 imply the inclusion $R(C) \subset R(B)$.

Taking $B = A$ and $C = E$ in (3) we get $I(A) = \max\{\alpha \geq 0 : A \succeq \alpha E\} = \rho(A^\dagger E)^{-1}$ for every $A \in P_n$ such that $p \in R(A)$. This proves part of the following result.

Proposition 2.5 *Let $A \in P_n$. Then $I(A) \neq 0$ if and only if the vector p belongs to $R(A)$. In this case, for any vector y such that $Ay = p$, we have*

$$I(A) = \rho(A^\dagger E)^{-1} = \langle y, p \rangle^{-1} = \left(\sum_{i=1}^n y_i \right)^{-1}. \quad (4)$$

Proof. By definition, $I(A) \neq 0$ if and only if there exists $\alpha > 0$ such that $A \succeq \alpha E$. By the comments following (3), this means that $R(E) \subset R(A)$ or, since p spans $R(E)$, that $p \in R(A)$. Finally, $I(A)^{-1} = \rho(A^\dagger E) = \rho(A^\dagger p p^T) = \rho(p^T A^\dagger p) = p^T A^\dagger p = \langle A^\dagger p, p \rangle$, and $A(A^\dagger p) = p$. If y is any vector such that $Ay = p$ then $y - A^\dagger p \in \ker A = R(A)^\perp$, so $\langle y, p \rangle = \langle A^\dagger p, p \rangle$.

Proposition 2.6 *Let $A \in P_n$. Then $I(A) = \min \{ \langle z, Az \rangle : \sum_{i=1}^n z_i = 1 \}$.*

Proof. If $\langle z, p \rangle = 1$, then $\langle z, Az \rangle \geq I(A) \langle z, Ez \rangle = I(A) z^* p p^* z = I(A) \langle z, p \rangle^2 = I(A)$. If $p \in R(A)$, let $x \in \mathbb{C}^n$ be such that $Ax = p$. Then $z = I(A)x$ satisfies $\langle z, p \rangle = I(A) \langle x, p \rangle = 1$ and $\langle z, Az \rangle = I(A) \langle z, p \rangle = I(A)$ by Proposition 2.5. If $p \notin R(A) = (\ker A)^\perp$, then there exists $z \in \ker A$ such that $\langle z, p \rangle = 1$ and $\langle z, Az \rangle = 0 = I(A)$. ■

Remark 2.7 Using proposition (3.9) of [23] and the results of this section, it is easy to see that, for all $A \in P_n$ and $m \in \mathbb{N}$, the inflation matrix $A^{(m)} = E_m \otimes A$ (where $E_m \in P_m$ has all its entries equal to 1) satisfies $I_{sp}(A^{(m)}) = I_{sp}(A)$ and $I(A^{(m)}) = I(A)$. Note that the inflation map $\Phi_A^{(m)} = \Phi_{A^{(m)}}$. Therefore the indexes of Φ_A are invariant under inflations and are invariants of Φ_A as a completely positive map.

$I_N(A)$ for general norms

Let $A \in P_n$ and let N be a norm in M_n . If $I_N(A) = 0$, there is some positive semidefinite matrix C such that $N(C) = 1$ (so $C \neq 0$) and $N(A \circ C) = 0$. But then $A \circ C = 0$, so $c_{ij} = 0$ whenever $a_{ij} \neq 0$. If all $a_{ii} \neq 0$, then all $c_{ii} = 0$, which forces $C = 0$. This contradiction shows that if all $a_{ii} \neq 0$, then $I_N(A) > 0$. Conversely, if some $a_{ii} = 0$ just take $C = E_{ii}/N(E_{ii})$, so $I(A) \leq N(A \circ C) = 0$. Thus,

$$I_N(A) > 0 \quad \text{if and only if all} \quad a_{ii} > 0. \quad (5)$$

Let $J \subseteq \{1, 2, \dots, n\}$ and let A_J denote the principal submatrix of A associated with J . Then, minimality ensures that

$$I_N(A) \leq I_N(A_J). \quad (6)$$

Remark 2.8 Let $A \in P_n$. Then, it can be shown that for every unitary invariant norm N , the following properties hold:

1. If A has rank one, then $I_N(A) = \min_{1 \leq i \leq n} A_{ii}$.
2. If A is positive and diagonal, then $I_N(D) = N'(D^{-1})^{-1} = \Phi'(a_{11}^{-1}, \dots, a_{nn}^{-1})^{-1}$, where N' is the dual norm of N and Φ' is the symmetric gauge function on \mathbb{R}^n associated with N' ; see [7, Chapter IV].

[23, Proposition 3.2] tells us that

$$I_{sp}(A) = \inf \{ I_{sp}(D) : A \preceq D \text{ and } D \text{ is diagonal} \}, \quad (7)$$

and one could hope that a similar formula holds for any norm, but it does not. In fact, our Corollary 4.4 says that, for every $A \in P_n$ and the Frobenius norm,

$$\begin{aligned} I_2(A) &= \inf \{ (\sum_1^n D_{ii}^{-2})^{-1/2} : D \text{ is diagonal and } A \circ \bar{A} \preceq D^2 \} \\ &= \inf \{ I_2(D) : D \text{ is diagonal and } A \circ \bar{A} \preceq D^2 \}. \end{aligned} \quad (8)$$

Note that the condition $A \circ \bar{A} \preceq D^2$ is strictly less restrictive than $A \preceq D$ (the reverse implication follows from Schur's theorem). Nevertheless, equation (8) allows one to compute the 2-index for every positive semidefinite matrix using only diagonal matrices. We intend to study this type of characterizations of $I_N(A)$ for general norms in a forthcoming paper.

3 $I(A) = I_{sp}(A)$.

In this section we characterize those matrices $A \in P_n$ such that $I(A) = I_{sp}(A)$. In [23] it is shown that for $A = \begin{pmatrix} a & b \\ \bar{b} & c \end{pmatrix} \in P_2$,

$$0 \neq I_{sp}(A) = I(A) \Leftrightarrow b \in \mathbb{R} \text{ and } 0 \leq b \leq \min\{a, c\} \neq 0. \quad (9)$$

This is easily seen to be equivalent to the conditions

1. $A \geq 0$.
2. There exists a vector $z \geq 0$ such that $Az = (1, 1)^T$ (if A is invertible, this means that $A^{-1}(1, 1)^T \geq 0$).

We prove that, for positive semidefinite matrices of any size with nonnegative entries, condition 2 is equivalent to the identity $I_{sp}(A) = I(A)$. But first we need two lemmas:

Lemma 3.1 *Let $A \in P_n$ and $L = \{z \in \mathbb{R}^n : \sum_i z_i = 1\}$. Consider the sets*

$$V_1 = \{z \in L : \langle Az, z \rangle = I(A)\} \quad \text{and} \quad V_2 = \{z \in L : Az = I(A)p\}.$$

Then $V_1 = V_2 \neq \emptyset$. Moreover, any local extreme point of the map $G : L \rightarrow \mathbb{R}$ given by $G(z) = \langle Az, z \rangle$, belongs to V_2 .

Proof. It is clear that $V_2 \subseteq V_1$. By Proposition 2.6, $I(A) \leq \min\{\langle Av, v \rangle : v \in L\}$. Then the map $G : L \rightarrow \mathbb{R}$ given by $G(z) = \langle Az, z \rangle = \sum_{i,j} a_{ij} z_j z_i$, is differentiable and bounded from below. Thus G must have a minimum, which is also a critical point. Let the columns of $X \in M_{n,n-1}$ be a basis for the orthogonal complement of p . Then we seek the unconstrained minimum of

$$\phi(\xi) = G(X\xi + p/n) = \langle A(X\xi + p/n), (X\xi + p/n) \rangle = (X + p/n)^T A(X\xi + p/n)$$

over all $\xi \in \mathbb{R}^{n-1}$. But $\nabla \phi(\xi) = 2X^T A(X\xi + p/n) = 0$ says that at a critical point ξ_0 , $Az_0 = \lambda p$ for some λ , where $z_0 \equiv X\xi_0 + p/n \in L$. But, in that case,

$$I(A) \leq \langle Az_0, z_0 \rangle = \lambda \langle p, z_0 \rangle = \lambda.$$

If $I(A) = 0$ then $\lambda = 0$, because $p \notin R(A)$, by Proposition 2.5. If $I(A) > 0$ then also $\lambda = I(A)$, because $y = \lambda^{-1} z_0$ satisfies $Ay = p$ and

$$\lambda = \langle Az_0, z_0 \rangle = \lambda^2 \langle Ay, y \rangle = \lambda^2 I(A)^{-1}.$$

So $\xi_0 \in \mathbb{R}^{n-1}$ is a critical point of ϕ if and only if $z_0 = X\xi_0 + p/n \in V_2$. Since each local extreme must be a critical point, this shows that $\emptyset \neq V_1 \subseteq V_2$ and the final assertion is proved ■

Lemma 3.2 *Let $A \in M_n$, and suppose $x \in \mathbb{C}^n$ with $\|x\| = 1$. Let $y = x \circ \bar{x} = (|x_1|^2, \dots, |x_n|^2)^T$.*

1. *If $Ay = \lambda p$ for some $\lambda \in \mathbb{C}$, then $(A \circ xx^*)x = \lambda x$.*
2. *Conversely, if all $x_i \neq 0$ and $(A \circ xx^*)x = \lambda x$ for some $\lambda \in \mathbb{C}$, then $Ay = \lambda p$.*

If $A \in P_n$, the eigenvalue λ of the matrix $A \circ xx^$ associated with the vector x must be $I(A)$ and $Ay = I(A) p$.*

Proof. Suppose that $Ay = \lambda p$. Then

$$(A \circ xx^*)x = (a_{ij} x_i \bar{x}_j) \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} (\sum_j a_{1j} |x_j|^2) x_1 \\ \vdots \\ (\sum_j a_{nj} |x_j|^2) x_n \end{pmatrix} = \begin{pmatrix} (Ay)_1 x_1 \\ \vdots \\ (Ay)_n x_n \end{pmatrix} = \lambda x. \quad (10)$$

(10) shows that if all $x_i \neq 0$ and $(A \circ xx^*)x = \lambda x$, then $Ay = \lambda p$. If $A \in P_n$ and $I(A) = 0$, then $\lambda = 0$ because $p \notin R(A)$. If $I(A) \neq 0$, then $p \in R(A) = (\ker A)^\perp$. So $Ay \neq 0$ because $1 = \|x\|^2 = \langle p, y \rangle \neq 0$. Then $\lambda \neq 0$. If $z = \lambda^{-1}y$, then $Az = p$ and $1 = \langle p, y \rangle = \lambda \langle Az, z \rangle = \lambda I(A)^{-1}$, by Proposition 2.5, $\lambda = I(A)$. ■

Theorem 3.3 *Let $A \in P_n$.*

1. *If $I_{sp}(A) = I(A) \neq 0$, then there exists a vector $u \geq 0$ such that $Au = p$.*
2. *If $A \geq 0$ and there exists a vector $u \geq 0$ such that $Au = p$, then $I_{sp}(A) = I(A)$.*

Proof.

1. Observe that $I(A)B \preceq A \circ B \preceq \|A \circ B\| I$. By Lemma 2.1 of [23], there exists $x \in \mathbb{R}^n$ such that $\|x\| = 1$ and $I_{sp}(A) = \|A \circ xx^*\|$. So, if $y = x \circ x$, then $\langle y, p \rangle = 1$ and

$$I(A)xx^T \preceq D_x A D_x \preceq I(A)I,$$

which implies that

$$I(A) = I(A)(x^T x)^2 \leq x^T D_x A D_x x = y^T A y \leq I(A)x^T x = I(A).$$

We have $\langle Ay, y \rangle = I(A)$, $y \geq 0$ and $\langle y, p \rangle = 1$. Then, by Lemma 3.1, $Ay = I(A)p$. Take $u = I(A)^{-1}y$.

2. Let u be a nonnegative vector such that $Au = p$. Let $y = I(A)u$ and $x = (y_1^{1/2}, \dots, y_n^{1/2})^T$. Note that $\|x\|^2 = \langle y, p \rangle = 1$. By Lemma 3.2 we know that x is an eigenvector of $A \circ xx^*$ with eigenvalue $I(A)$. Recall that always $I(A) \leq I_{sp}(A)$.

Case 1. Suppose that x has strictly positive entries. Since $A \circ xx^* \geq 0$, it is well known (see Corollary 8.1.30 of [14]) that the eigenvalue $I(A)$ of x must be the spectral radius of $A \circ xx^*$. Since $A \circ xx^* \in P_n$ we deduce that $I(A) = \|A \circ xx^*\| \geq I_{sp}(A)$.

Case 2. Let $J = \{i : x_i \neq 0\}$, A_J the principal submatrix of A determined by the indexes of J and similarly define x_J . Then x_J is an eigenvector of $A_J \circ x_J x_J^*$ with eigenvalue $I(A)$. Note also that $A_J \circ x_J x_J^* \succeq I(A_J)x_J x_J^*$ and $x_J x_J^*(x_J) = \|x_J\|^2 x_J = x_J$. Then

$$0 \leq \langle (A_J \circ x_J x_J^* - I(A_J)x_J x_J^*)x_J, x_J \rangle = I(A) - I(A_J)$$

and, by the definition of I , $I(A_J) = I(A)$. Now, as in Case 1, we can deduce that

$$I(A) = I(A_J) = \|A_J \circ x_J x_J^*\| \geq I_{sp}(A_J) \geq I_{sp}(A),$$

where the last inequality holds by (6). ■

Corollary 3.4 *Let $A \in P_n$ such that $A \geq 0$ and $I_{sp}(A) = I(A)$. Let u be a nonnegative vector such that $Au = p$ and $y = I(A) u$.*

1. *Let $x = (y_1^{1/2}, \dots, y_n^{1/2})^T$. Then $\|x\| = 1$ and $\|A \circ xx^*\| = I_{sp}(A)$.*
2. *Let $J = \{i : u_i \neq 0\}$ and denote by A_J the principal submatrix of A determined by J . Then $I(A) = I(A_J) = I_{sp}(A_J) = I_{sp}(A)$.*

Proof. This follows from the proof of Theorem 3.3. ■

Remark 3.5 In Theorem 3.3(2), the hypothesis that $A \geq 0$ is essential. Indeed, consider $A = \begin{pmatrix} 2 & -1 \\ -1 & 1 \end{pmatrix}$ and $u = (2, 3)^T$. Then $Au = (1, 1)^T$ but $1/5 = I(A) \neq I_{sp}(A) = 1$. For $A \in P_n$, we conjecture that $I(A) = I(sp, A) \neq 0$ implies that $A \geq 0$, as in the 2×2 case.

4 $I_{sp}(A)$ and $I_2(A)$

In this section we study the relation between the indexes associated with the spectral and Frobenius norms. In Lemma 2.1 of [23] it is shown that the index $I_{sp}(\cdot)$ is always attained at rank one projections. The index $I_1(\cdot)$ has the same property (see Example 2.3). It is natural to conjecture that the same result holds for any unitary invariant norm N . We show that the conjecture is true for the Frobenius norm:

Proposition 4.1 *Let $A \in P_n$. Then there exists an $x \in \mathbb{C}^n$, such that $\|x\| = 1$ and $I_2(A) = \|A \circ xx^*\|_2$. That is, $I_2(A)$ is attained at a rank one projection.*

Proof. Let $\lambda = \max\{\mu \geq 0 : \|A \circ B\|_2 \geq \mu\|B\|_2 \text{ for all } B \in P_n \text{ with rank } 1\}$. By its definition $\lambda \geq I_2(A)$. Let us prove that $\|A \circ B\|_2 \geq \lambda\|B\|_2$ for all $B \in P_n$. Indeed, for $B \succeq 0$, write $B = \sum_{i=1}^k B_i$, where each B_i has rank one, $B_i \in P_n$, and $B_i B_j = 0$ if $i \neq j$. Then

$$\lambda^2 \|B\|_2^2 = \lambda^2 \sum_{i=1}^k \|B_i\|_2^2 \leq \sum_{i=1}^k \|A \circ B_i\|_2^2.$$

On the other hand, using $\text{tr}(XY) \geq 0$ for positive semidefinite matrices X and Y ,

$$\begin{aligned} \|A \circ B\|_2^2 &= \text{tr}((A \circ B)^*(A \circ B)) = \sum_{ij} \text{tr}(A \circ B_i)^*(A \circ B_j) \\ &\geq \sum_i \text{tr}(A \circ B_i)^*(A \circ B_i) = \sum_i \|A \circ B_i\|_2^2. \end{aligned}$$

■

Proposition 4.2 *Let $A \in P_n$.*

1. *There exists a nonnegative vector x such that $\|x\| = 1$ and $\|A \circ xx^*\|_2 = I_2(A)$.*

2. Any such vector x satisfies $(A \circ \bar{A} \circ xx^*)x = I(B_J)x$, where $B = A \circ \bar{A}$ and $J = \{i : x_i \neq 0\}$.

Proof. Let y be a unit vector such that $\|A \circ yy^*\|_2 = I_2(A)$. Let $x_i = |y_i|$. It is easily checked that $\|x\| = 1$ and $\|A \circ xx^*\|_2 = \|A \circ yy^*\|_2 = I_2(A)$, which proves 4.2(1). Let $B = A \circ \bar{A} \in P_n$. Let y be a nonnegative unit vector and let $z = (y_1^2, \dots, y_n^2)^T$. Then

$$\|A \circ yy^*\|_2^2 = \sum_{i,j} |a_{ij}|^2 y_i^2 y_j^2 = \sum_{i,j} b_{ij} z_j z_i = \langle Bz, z \rangle$$

and $\sum_1^n z_i = 1$. Then $\|A \circ yy^*\|_2 = I_2(A)$ if and only if $\langle Bz, z \rangle$ is the minimum of the map $G(v) = \langle Bv, v \rangle$ restricted to the simplex $\Delta = \{v \in \mathbb{R}^n : v \geq 0 \text{ and } \sum_1^n v_i = 1\}$. Using Lemma 3.1, we know that if z belongs to the interior Δ° of Δ , then z is a local extremum of G in the plane $L = \{z \in \mathbb{R}^n : \sum_i z_i = 1\}$, so $Bz = I(B)p$.

If the vector x of item 1 satisfies $x_i > 0$ for all i , then $z = x \circ x \in \Delta^\circ$ and $Bz = I(B)p$. By Lemma 3.2, $(A \circ \bar{A} \circ xx^*)x = I(B)x$, showing 4.2(2) in this case. If some $x_i = 0$, let $J = \{i : x_i \neq 0\}$, let B_J be the principal submatrix of B determined by the indexes of J , and similarly define x_J . Then $I_2(A) = \|A \circ xx^*\|_2 = \|A_J \circ x_J x_J^*\|_2 \geq I_2(A_J)$ and

$$I_2(A) = I_2(A_J) = \|A_J \circ x_J x_J^*\|_2,$$

because the converse inequality always holds by (6). Note that, by its definition, x_J has no zero entries. By the previous case, x_J is an eigenvector of $B_J \circ x_J x_J^*$ with eigenvalue $I(B_J)$. But clearly $B \circ xx^*$ has zeroes outside $J \times J$, so x is an eigenvector of $B \circ xx^*$ if and only if x_J is an eigenvector of $B_J \circ x_J x_J^*$. Note that the eigenvalue of x is always $I(B_J)$. ■

Theorem 4.3 *Let $A \in P_n$. Then $I_2(A) = I_{sp}(\bar{A} \circ A)^{1/2}$.*

Proof. If $B = \bar{A} \circ A$ and $y \in \mathbb{C}^n$ with $\|y\| = 1$, then

$$\|A \circ yy^*\|_2^2 = \sum_{i,j} |a_{ij}|^2 |y_i|^2 |y_j|^2 = \langle (B \circ yy^*)y, y \rangle \leq \|B \circ yy^*\|.$$

Therefore $I_2(A)^2 \leq I_{sp}(B)$. On the other hand, let x be a nonnegative unit vector such that $I_2(A)^2 = \|A \circ xx^*\|_2^2$ and $J = \{i : x_i \neq 0\}$. Then, by Proposition 4.2, $(B \circ xx^*)x = I(B_J)x$ and

$$I_2(A)^2 = \|A \circ xx^*\|_2^2 = \langle (B \circ xx^*)x, x \rangle = I(B_J).$$

But x_J is a unit eigenvector of $B_J \circ x_J x_J^*$ with strictly positive entries. So, by Lemma 3.2, $B_J(x_J \circ x_J) = I(B_J)(1, \dots, 1)^T$. Suppose that $I_2(A) \neq 0$. Then $I(B_J) \neq 0$, $B_J \geq 0$, the vector $u = I(B_J)^{-1}(x_J \circ x_J)$ has strictly positive entries, and $B_J u = (1, \dots, 1)^T$. Hence we can apply Theorem 3.3 to B_J and, by (6),

$$I(B_J) = I_{sp}(B_J) \geq I_{sp}(B) \geq I_2(A)^2 = I(B_J).$$

If $I_2(A) = 0$, then (5) ensures that some $a_{ii} = 0$, so $I_{sp}(B) = 0$ by (5). ■

Corollary 4.4 *Let $A \in P_n$. Then*

$$\begin{aligned} I_2(A) &= \inf \{ (\sum_1^n d_{ii}^{-2})^{-1/2} : D \text{ is positive diagonal and } A \circ \bar{A} \preceq D^2 \} \\ &= \inf \{ I_2(D) : D \text{ is positive diagonal and } A \circ \bar{A} \preceq D^2 \}. \end{aligned}$$

Proof. This is a direct consequence of Theorem 4.3 and Proposition 3.2 of [23]. ■

Remark 4.5 In Theorem 4.3 we get information about $A \in P_n$ using $B = \bar{A} \circ A$. But it can also be used to get information about any $B \in P_n$ with $B \geq 0$, using $A = (b_{ij}^{1/2})$. Unfortunately it may certainly happen that $A \notin P_n$. Nevertheless this obstruction can be removed in the following way: given a selfadjoint (but not necessarily positive semidefinite) matrix $A \in M_n$, consider the index

$$I_2(A) = \min \{ \|A \circ xx^*\|_2 : \|x\| = 1 \},$$

which, by Proposition 4.1, is consistent with Definition 2.2 when $A \succeq 0$. A careful inspection of the proofs of Proposition 4.2 and Theorem 4.3 shows that they remain true using this new index if the condition “ $A \in P_n$ ” is replaced by “ $A = A^*$ and $B = \bar{A} \circ A \in P_n$ ”. Note that Lemma 3.1, Lemma 3.2, and Theorem 3.3 are applied only to the positive semidefinite matrix B or its principal submatrices. The inequality $I_2(A) \leq I_2(A_J)$ in (6) (which is also used in the proofs) remains valid for this new index. This observation is useful in order to avoid the unpleasant condition “ $A = (b_{ij}^{1/2}) \in P_n$ ” in the following result.

Corollary 4.6 *Suppose $B \in P_n$ and $B \geq 0$. Then there exists a subset J_0 of $\{1, 2, \dots, n\}$ such that $I_{sp}(B) = I_{sp}(B_{J_0}) = I(B_{J_0})$. Therefore*

$$I_{sp}(B) = \min \{ I_{sp}(B_J) : I_{sp}(B_J) = I(B_J) \}.$$

If $A = (b_{ij}^{1/2})$ (which may be not positive semidefinite), then J_0 can be characterized as $J_0 = \{i : x_i \neq 0\}$ for any unit vector x such that $I_2(A) = \|A \circ xx^\|_2$. Also $I_{sp}(B) = \|B \circ xx^*\| = \langle By, y \rangle$, where $y = (|x_1|^2, \dots, |x_n|^2)^T$.*

Proof. Use Remark 4.5 and the proof of Theorem 4.3. ■

5 An Operator Inequality

In this section we compute the indexes of a particular class of matrices and, as an application, we get a new operator inequality, closely related to the inequality proved in [8]; see also [4], [18].

Let $x = (\lambda_1, \dots, \lambda_n)^T \in \mathbb{R}_+^n$, $S = \{\lambda_1, \dots, \lambda_n\}$, and

$$\Lambda = \Lambda_x = (\lambda_i \lambda_j + \frac{1}{\lambda_i \lambda_j})_{ij} \in P_n.$$

Observe that Λ has rank 1 or 2.

5.1 Computation of $I(\Lambda)$

1. If all λ_i are equal, then $\Lambda = (\lambda_1^2 + \lambda_1^{-2}) E$ and $I(\Lambda) = \lambda_1^2 + \lambda_1^{-2}$.
2. If $\#S > 1$, then the range of Λ is spanned by the independent vectors $x = (\lambda_1, \dots, \lambda_n)^T$ and $y = (\lambda_1^{-1}, \dots, \lambda_n^{-1})^T$, because $\Lambda = xx^* + yy^* = [xy][xy]^*$ has rank 2.
3. If $\#S = 2$, say $S = \{\lambda, \mu\}$, then $p = ax + by$, with $a = (\lambda + \mu)^{-1}$ and $b = \lambda\mu(\lambda + \mu)^{-1}$. If a vector z satisfies $\Lambda z = p$, then

$$p = \Lambda z = (xx^* + yy^*)z = \langle z, x \rangle x + \langle z, y \rangle y.$$

Therefore

$$I(\Lambda) = \langle z, p \rangle^{-1} = (\langle z, x \rangle^2 + \langle z, y \rangle^2)^{-1} = \frac{(\lambda + \mu)^2}{1 + \lambda^2 \mu^2} = I(\Lambda_0),$$

where the last equality is shown in Remark 4.3 of [23].

4. If $\#S > 2$, it is easy to see that p is not in the subspace spanned by x and y . Then $I(\Lambda)$ must be zero by Proposition 2.5 .

Note that $I(\Lambda) \neq 0$ if and only $\#S \leq 2$.

5.2 Computation of $I_{sp}(\Lambda)$

We shall compute $I_{sp}(\Lambda)$ using Corollary 4.6 and therefore use the principal submatrices of Λ , which are matrices of the same type. Let $J \subset \{1, 2, \dots, n\}$, let $S_J = \{\lambda_j : j \in J\}$, and let x_J be the induced vector. Then $\Lambda_J = \Lambda_{x_J}$ and so $I_{sp}(\Lambda_J) \neq 0$. Suppose that $I_{sp}(\Lambda_J) = I(\Lambda_J)$. Then $\#S_J \leq 2$ by 5.1. If $\#S_J = 2$, let $i_1, i_2 \in J$ be such that $\lambda_{i_1} \neq \lambda_{i_2}$. By Theorem 3.3 there exists a vector $y \in \mathbb{R}^J$ such that $y \geq 0$ and $\Lambda_J y = p_J$. Let $z_1 = \Sigma\{y_k : k \in J \text{ and } \lambda_k = \lambda_{i_1}\} \geq 0$ and $z_2 = \Sigma\{y_j : j \in J \text{ and } \lambda_j = \lambda_{i_2}\} \geq 0$. Easy computations show that $\Lambda_{\{i_1, i_2\}}(z_1, z_2)^T = (1, 1)^T$. Then, by Theorem 3.3 and 5.1,

$$I_{sp}(\Lambda_J) = I(\Lambda_J) = \frac{(\lambda_{i_1} + \lambda_{i_2})^2}{1 + \lambda_{i_1}^2 \lambda_{i_2}^2} = I(\Lambda_{\{i_1, i_2\}}) = I_{sp}(\Lambda_{\{i_1, i_2\}}).$$

Therefore, in order to compute $I_{sp}(\Lambda)$ using Corollary 4.6, we need to consider only the diagonal entries of Λ and some of the principal submatrices of size 2×2 . If $\lambda_i \neq \lambda_j$, (9) ensures that

$$I_{sp}(\Lambda_{\{i, j\}}) = I(\Lambda_{\{i, j\}}) \Leftrightarrow \lambda_i \lambda_j + \frac{1}{\lambda_i \lambda_j} \leq \min\{\lambda_i^2 + \frac{1}{\lambda_i^2}, \lambda_j^2 + \frac{1}{\lambda_j^2}\}.$$

If $\lambda_i < \lambda_j$, this condition is equivalent to

$$\lambda_i^2 \leq \frac{1}{\lambda_i \lambda_j} \leq \lambda_j^2. \tag{11}$$

In particular, this implies that $\lambda_i < 1 < \lambda_j$. Then, by Corollary 4.6,

$$I_{sp}(\Lambda) = \min\{M_1, M_2\} \quad (12)$$

where $M_1 = \min_i \lambda_i^2 + \lambda_i^{-2} = \min_i \Lambda_{ii}$ and

$$M_2 = \inf \left\{ \frac{(\lambda_i + \lambda_j)^2}{1 + \lambda_i^2 \lambda_j^2} : \lambda_i < 1 < \lambda_j \text{ and } \lambda_i^2 \leq \frac{1}{\lambda_i \lambda_j} \leq \lambda_j^2 \right\}.$$

For example, if all $\lambda_i \geq 1$ (or all $\lambda_i \leq 1$), then by (11), $I_{sp}(\Lambda) = M_1 = \min_i \lambda_i^2 + \lambda_i^{-2}$. On the other hand, if $\lambda \neq 1$ and $x = (\lambda, \lambda^{-1})^T$, then

$$I_{sp}(\Lambda_x) = M_2 = \frac{\lambda^2 + \lambda^{-2}}{2} + 1 < M_1 = \lambda^2 + \lambda^{-2}.$$

Proposition 5.3 *Let \mathcal{H} be a Hilbert space and let S be a bounded selfadjoint invertible operator on \mathcal{H} . Let $M(S)$ be the best constant such that*

$$\|STS + S^{-1}TS^{-1}\| \geq M(S)\|T\| \quad \text{for all } 0 \preceq T \in L(\mathcal{H}).$$

Then $M(S) = \min\{M_1(S), M_2(S)\}$, where

$$M_1(S) = \min_{\lambda \in \sigma(S)} \lambda^2 + \lambda^{-2} \quad \text{and}$$

$$M_2(S) = \inf \left\{ \frac{(|\lambda| + |\mu|)^2}{1 + \lambda^2 \mu^2} : \lambda, \mu \in \sigma(S), |\lambda| < |\mu| \text{ and } \lambda^2 \leq \frac{1}{|\lambda \mu|} \leq \mu^2 \right\}.$$

In particular, if $\|S\| \leq 1$ (or $\|S^{-1}\| \leq 1$), then

$$M(S) = \|S\|^2 + \|S\|^{-2} \quad (\text{resp. } \|S^{-1}\|^2 + \|S^{-1}\|^{-2}).$$

Proof. We follow the same steps as in [8]. By taking the polar decomposition of S , we can assume that $S > 0$, because the unitary part of S is also the unitary part of S^{-1} ; it commutes with S and S^{-1} and it preserves norms. Note that we must change $\sigma(S)$ by $\sigma(|S|) = \{|\lambda| : \lambda \in \sigma(S)\}$.

By the spectral theorem, we can assume that $\sigma(S)$ is finite, because S can be approximated in norm by operators S_n such that each $\sigma(S_n)$ is a finite subset of $\sigma(S)$, $\sigma(S_n) \subset \sigma(S_{n+1})$ for all $n \in \mathbb{N}$ and $\cup_n \sigma(S_n)$ is dense in $\sigma(S)$. Then $M(S_n)$ (and $M_i(S_n)$, $i = 1, 2$) converges to $M(S)$ (resp. $M_i(S)$, $i = 1, 2$).

We can suppose also that $\dim \mathcal{H} < \infty$, by choosing an appropriate net of finite rank projections $\{P_F\}_{F \in \mathcal{F}}$ that converges strongly to the identity and replacing S , T by $P_F S P_F$, $P_F T P_F$. Indeed, the net may be chosen in such a way that $S P_F = P_F S$ and $\sigma(P_F S P_F) = \sigma(S)$ for all $F \in \mathcal{F}$. Note that for every $A \in L(H)$, $\|P_F A P_F\|$ converges to $\|A\|$.

Finally, we can suppose that S is diagonal by a unitary change of basis in \mathbb{C}^n . In this case, if $\lambda_1, \dots, \lambda_n$ are the eigenvalues of S (with multiplicity) and $x = (\lambda_1, \dots, \lambda_n)^T$, then

$$STS + S^{-1}TS^{-1} = \Lambda_x \circ T.$$

None of our reductions (unitary equivalences and compressions) change the fact that $0 \preceq T$. Now the statement follows from formula (12). If $\|S\| \leq 1$ then $M(S) = M_1(S)$, because $M_2(S)$ is the infimum of the empty set. Clearly $M_1(S)$ is attained at the element $\lambda \in \sigma(S)$ such that $|\lambda| = \|S\|$. ■

References

- [1] A. Albert, Conditions for positive and nonnegative definiteness in terms of pseudoinverses, *SIAM J. Appl. Math.* 17: 434-440, 1969.
- [2] T. Ando, Majorizations and inequalities in matrix theory, *Linear Algebra Appl.* 199: 17-67, 1994.
- [3] T. Ando and K. Okubo, Induced norms of the Schur multiplier operator, *Linear Algebra Appl.* 147: 181-199, 1991.
- [4] E. Andruchow, G. Corach and D. Stojanoff, Geometric operator inequalities, *Linear Algebra Appl.* 258: 295-310, 1997.
- [5] M. Baillet, Y. Denizeau and J. F. Havet, Indice d'une esperance conditionnelle, *Compositio Math.* 66: 199-236, 1988.
- [6] J. Baksalary, B. Schipp, and G. Trenkler, Some further results on Hermitian-matrix inequalities, *Linear Algebra Appl.* 160: 119-129, 1992.
- [7] R. Bhatia, *Matrix Analysis*, Springer-Verlag, New York, 1997.
- [8] G. Corach, H. Porta and L. Recht, An operator inequality, *Linear Algebra Appl.* 142: 153-158, 1990.
- [9] C. C. Cowen, K. E. Debro and P. D. Sepansky, Geometry and the norms of Hadamard multipliers, *Linear Algebra Appl.* 218: 239-249, 1995.
- [10] C. C. Cowen, M. A. Dritschel and R. C. Penney, Norms of Hadamard multipliers, *SIAM J. Matrix. Anal. Appl.* 15: 313-320, 1994.
- [11] C. Davis, The norm of the Schur product operation, *Numer. Math.* 4: 343-344, 1962.
- [12] Y. Denizeau and J. F. Havet, Correpondances d'indice fini I: indice d'un vecteur, *J. Operator Theory*, 32: 111-156, 1994.
- [13] C. FitzGerald & R. Horn, On the structure of Hermitian-symmetric inequalities, *J. London Math. Soc. (2)* 15: 419-430, 1977 .
- [14] R. A. Horn, C. R. Johnson, *Matrix Analysis*, Cambridge University Press, Cambridge 1990.

- [15] R. A. Horn, C. R. Johnson, *Topics in matrix analysis*, Cambridge University Press, Cambridge 1991.
- [16] R. A. Horn, On inequalities between Hermitian and symmetric forms, *Linear Alg. Appl.* 11: 189-218, 1975.
- [17] R. A. Horn, Norm bounds for Hadamard products and an arithmetic - geometric mean inequality for unitarily invariant norms, *Linear Algebra Appl.* 223/224: 119-129, 1995.
- [18] L. Livshits and S. C. Ong, On the invertibility properties of the map $T \mapsto STS^{-1} + S^{-1}TS$ and operator-norm inequalities, *Linear Algebra Appl.* 183: 117-129, 1993.
- [19] R. Mathias, Matrix completions, norms and Hadamard products, *Proc. Amer. Math. Soc.* 117: 905-918, 1993.
- [20] V. I. Paulsen, S. C. Power and R. R. Smith, Schur products and matrix completions, *J. Functional Analysis*, 85: 151-178, 1989.
- [21] S. Popa, Classification of subfactors and their endomorphisms, AMS-CBMS, No. 86, Providence, 1995.
- [22] I. Schur, Bemerkungen zur theorie de beschränkten bilineareformen mit unendlich vielen veränderlichen, *J. Reine Angew. Math.* 140: 1-28, 1911.
- [23] D. Stojanoff, Index of Hadamard multiplication by positive matrices. *Linear Algebra Appl.*, 290: 95-108, 1999.