More Operator Versions of the Schwarz Inequality

Rajendra Bhatia[1], **Chandler Davis**[2]

[1] Indian Statistical Institute, 7, S.J.S. Sansanwal Marg, New Delhi – 110 016, India.
E-mail: rbh@isid.ac.in

[2] Department of Mathematics, University of Toronto, Toronto M5S3G3, Canada.
E-mail: davis@math.toronto.edu

Received: 14 April 2000 / Accepted: 10 May 2000

Abstract: Some new operator versions of the Schwarz inequality are obtained. One of them is a counterpart of the variance-covariance inequality in the context of noncommutative probability.

The Schwarz inequality has appeared in several *avatars*. Some of these are its versions for operators [1, 2, 4, 7–9]. More are presented here.

1. A Variance–Covariance Inequality

Let f and g be random variables – elements of the space $L_2(X,\mu)$, where μ is a probability measure. The *covariance* between f and g is defined as

$$\mathrm{cov}(f,g) = \mathrm{E}(\overline{f}g) - \overline{\mathrm{E}f}\,Eg. \tag{1}$$

where $\mathrm{E}f = \int f d\mu$ denotes the *expectation* of f. The *variance* of f is defined as

$$\mathrm{var}(f) = \mathrm{cov}(f,f) = \mathrm{E}(|f|^2) - |\mathrm{E}f|^2. \tag{2}$$

The inequality

$$|\mathrm{cov}(f,g)|^2 \leq \mathrm{var}(f)\mathrm{var}(g), \tag{3}$$

much used in statistics, is just the Schwarz inequality.

A noncommutative analogue of variance and covariance is defined as follows. Let $\mathcal{B}(\mathcal{H})$ be the space of all (bounded linear) operators on a (complex separable) Hilbert space $\mathcal{H}$. Let Φ be a unital completely positive linear map [10] on $\mathcal{B}(\mathcal{H})$. We define the *covariance* between two operators A and B as

$$\mathrm{cov}(A,B) = \Phi(A^*B) - \Phi(A)^*\Phi(B). \tag{4}$$

The *variance* of A is defined as

$$\mathrm{var}(A) = \mathrm{cov}(A, A) = \Phi(A^*A) - \Phi(A)^*\Phi(A). \tag{5}$$

The famed inequality of Kadison [7], extended in several directions by Choi [4], says that

$$\mathrm{var}(A) \geq 0. \tag{6}$$

[We write $T \geq 0$ to mean that the operator T is positive (semidefinite).] This generalises the simple fact that $\mathrm{var}(f)$ is a nonnegative number.

A good generalisation of the inequality (3) is given by the following theorem.

Theorem 1. *The* 2×2 *operator matrix*

$$\begin{bmatrix} \mathrm{var}(A) & \mathrm{cov}(A, B) \\ \mathrm{cov}(A, B)^* & \mathrm{var}(B) \end{bmatrix} \tag{7}$$

is positive.

Proof. We have to prove that

$$\begin{bmatrix} \Phi(A^*A) & \Phi(A^*B) \\ \Phi(B^*A) & \Phi(B^*B) \end{bmatrix} \geq \begin{bmatrix} \Phi(A)^*\Phi(A) & \Phi(A)^*\Phi(B) \\ \Phi(B)^*\Phi(A) & \Phi(B)^*\Phi(B) \end{bmatrix}. \tag{8}$$

First consider the special case when Φ is the map $\Phi(T) = V^*TV$, where V is an isometry; i.e., $V^*V = I$. Then the inequality (8) can be rewritten as

$$\begin{bmatrix} V^* & 0 \\ 0 & V^* \end{bmatrix} \begin{bmatrix} A^*A & A^*B \\ B^*A & B^*B \end{bmatrix} \begin{bmatrix} V & 0 \\ 0 & V \end{bmatrix}$$
$$\geq \begin{bmatrix} V^* & 0 \\ 0 & V^* \end{bmatrix} \begin{bmatrix} A^*VV^*A & A^*VV^*B \\ B^*VV^*A & B^*VV^*B \end{bmatrix} \begin{bmatrix} V & 0 \\ 0 & V \end{bmatrix}.$$

This will follow from the inequality

$$\begin{bmatrix} A^*A & A^*B \\ B^*A & B^*B \end{bmatrix} \geq \begin{bmatrix} A^*VV^*A & A^*VV^*B \\ B^*VV^*A & B^*VV^*B \end{bmatrix}.$$

This, in turn, can be written as

$$\begin{bmatrix} A^* & 0 \\ B^* & 0 \end{bmatrix} \begin{bmatrix} A & B \\ 0 & 0 \end{bmatrix} \geq \begin{bmatrix} A^* & 0 \\ B^* & 0 \end{bmatrix} \begin{bmatrix} VV^* & 0 \\ 0 & VV^* \end{bmatrix} \begin{bmatrix} A & B \\ 0 & 0 \end{bmatrix}.$$

But as $VV^* \leq I$, this is certainly true. We have proved (8) for the special Φ.

The general case follows from this via the Stinespring dilation theorem: there exists a Hilbert space $\mathcal{K}$, an isometry V of $\mathcal{H}$ into $\mathcal{K}$, and a $*$-homomorphism π of $\mathcal{B}(\mathcal{H})$ into $\mathcal{B}(\mathcal{K})$ such that $\Phi(A) = V^*\pi(A)V$. $\square$

Remark 1. It is well-known [1,4] that $\begin{bmatrix} R & T \\ T^* & S \end{bmatrix} \geq 0$ if and only if R, S are positive and $R \geq TS^{-1}T^*$. Here if S is not invertible S^{-1} is understood to be its generalised inverse. (The same convention is followed in such contexts throughout the paper.) Theorem 1 is thus equivalent to the statement

$$\mathrm{var}(A) \geq \mathrm{cov}(A, B)[\mathrm{var}(B)]^{-1}\mathrm{cov}(A, B)^*. \tag{9}$$

This equivalence will be used repeatedly.

Remark 2. The Schwarz inequality proved by Lieb and Ruskai [8] says that

$$\Phi(A^*A) \geq \Phi(A^*B)\Phi(B^*B)^{-1}\Phi(B^*A), \tag{10}$$

or, equivalently,

$$\begin{bmatrix} \Phi(A^*A) & \Phi(A^*B) \\ \Phi(B^*A) & \Phi(B^*B) \end{bmatrix} \geq 0. \tag{11}$$

The inequality (8) is a considerable strengthening of this result.

Remark 3. Say that a function f is in the Lieb class $\mathcal{L}$ if $f(R) \geq 0$ whenever $R \geq 0$, and $|f(T)|^2 \leq f(R)f(S)$ whenever $\begin{bmatrix} R & T \\ T^* & S \end{bmatrix} \geq 0$. Several examples of such functions may be found in [2, pp. 268–270]. Many Schwarz-type inequalities for such f may thus be obtained from (7). For example, we have

$$\| \operatorname{cov}(A, B)\|^2 \leq \| \operatorname{var}(A)\| \; \| \operatorname{var}(B)\|, \tag{12}$$

for every unitarily invariant norm. This gives a variety of good generalisations of (3).

It is often of interest to weaken the hypothesis that the map Φ be completely positive. We will comment below on this much studied class of maps [4]:

Definition. *Let Φ be a linear map between C^*-algebras. We say that Φ is n-**positive** in case the condition $[A_{ij}]_{ij} \geq 0$ on an $n \times n$ operator matrix A_{ij} implies $[\Phi(A_{ij})]_{ij} \geq 0$.*

Thus ordinary positivity is 1-positivity; complete positivity is n-positivity for all n.

Remark 4. If Φ is assumed only to be a unital positive linear map then the inequality (6) is not always true. It is true under additional hypotheses such as self-adjointness of A. However, the inequalities (6) and (11) are true if Φ is just assumed to be 2-positive. We do not know whether we have (8) under this weaker condition. If Φ, in addition to being 2-positive, has the averaging property [5] $\Phi(A\Phi(B)) = \Phi(A)\Phi(B)$, then the inequality (8) does hold. Another strengthening is given in Remark 7.

Remark 5. When Φ is the identity map, the inequality (10) reduces to

$$A^*A \geq A^*B(B^*B)^{-1}B^*A. \tag{13}$$

An easy proof of this is given in [9]. The operator $B(B^*B)^{-1}B^*$ is idempotent and Hermitian. Hence, $I \geq B(B^*B)^{-1}B^*$; and (13) follows at once.

Remark 6. The argument in the proof of Theorem 1 can be used to show that for any operators $A_1, \ldots, A_n$, the $n \times n$ block operator matrix $[\operatorname{cov}(A_i, A_j)]$ is positive.

Remark 7. The referee has pointed out to us an ingenious proof of (8) (that is, the conclusion of Theorem 1) under the hypothesis that Φ is unital and 4-positive. This will be sketched here. From the easily verified relation

$$\begin{bmatrix} A^*A & A^*B & A^* & A^* \\ B^*A & B^*B & B^* & B^* \\ A & B & I & I \\ A & B & I & I \end{bmatrix} \geq 0$$

and the 4-positivity of Φ follows

$$\begin{bmatrix} \Phi(A^*A) & \Phi(A^*B) & \Phi(A)^* & \Phi(A)^* \\ \Phi(B^*A) & \Phi(B^*B) & \Phi(B)^* & \Phi(B)^* \\ \Phi(A) & \Phi(B) & I & I \\ \Phi(A) & \Phi(B) & I & I \end{bmatrix} \geq 0.$$

Applying again the equivalence in Remark 1, this yields (8). The referee remarks that this improvement can be made equally to the generalisation in Remark 6; here one assumes of Φ only that it is unital and $2n$-positive.

2. An Operator Version of the Wielandt Inequality

Let A be a positive operator on $\mathcal{H}$. For any two vectors x, y in $\mathcal{H}$, we have from the Schwarz inequality

$$|\langle x, Ay\rangle|^2 \leq \langle x, Ax\rangle \langle y, Ay\rangle. \tag{14}$$

A well-known inequality of Wielandt [6, p. 443] gives a much improved inequality in the special case when x and y are orthogonal. If $mI \leq A \leq MI$, and $x \perp y$, then

$$|\langle x, Ay\rangle|^2 \leq \left(\frac{M-m}{M+m}\right)^2 \langle x, Ax\rangle \langle y, Ay\rangle. \tag{15}$$

From this one can derive another well-known result called the Kantorovich inequality: for every unit vector x,

$$\langle x, Ax\rangle \langle x, A^{-1}x\rangle \leq \frac{(M+m)^2}{4Mm}. \tag{16}$$

See [6] for details. We discuss operator versions of these inequalities.

Let A be a positive operator and X, Y any two operators. Replacing the A and B in (13) by $A^{1/2}X$ and $A^{1/2}Y$, respectively, we obtain

$$X^*AY(Y^*AY)^{-1}Y^*AX \leq X^*AX. \tag{17}$$

From this we get for every 2-positive linear map Φ

$$\Phi(X^*AY)\Phi(Y^*AY)^{-1}\Phi(Y^*AX) \leq \Phi(X^*AX). \tag{18}$$

(See Remark 1.) This is an operator version of (14). A similar extension of (15) is given below.

Theorem 2. *Let A be a positive operator on $\mathcal{H}$ with $mI \leq A \leq MI$. Let X, Y be two partial isometries in $\mathcal{H}$ whose final spaces are orthogonal to each other. Let Φ be a 2-positive linear map on $\mathcal{B}(\mathcal{H})$. Then*

$$\Phi(X^*AY)\Phi(Y^*AY)^{-1}\Phi(Y^*AX) \leq \left(\frac{M-m}{M+m}\right)^2 \Phi(X^*AX). \tag{19}$$

Proof. An operator version of (16) is known [3, 9]. It says that for every positive unital linear map Ψ,

$$\Psi(A^{-1}) \leq \frac{(M+m)^2}{4Mm}\Psi(A)^{-1}. \tag{20}$$

Now consider a direct sum decomposition $\mathcal{H} = \mathcal{H}_1 \oplus \mathcal{H}_2$, and a corresponding block decomposition of A as

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}.$$

Then we have

$$A^{-1} = \begin{bmatrix} (A_{11} - A_{12}A_{22}^{-1}A_{21})^{-1} & * \\ * & * \end{bmatrix}.$$

See [6, p. 472]. If we put $\Psi(A) = A_{11}$, we get from (20)

$$(A_{11} - A_{12}A_{22}^{-1}A_{21})^{-1} \leq \frac{(M+m)^2}{4Mm}A_{11}^{-1}.$$

This is equivalent to

$$A_{11} - A_{12}A_{22}^{-1}A_{21} \geq \frac{4Mm}{(M+m)^2}A_{11},$$

and thereby to

$$A_{12}A_{22}^{-1}A_{21} \leq \left(\frac{M-m}{M+m}\right)^2 A_{11}.$$

If X, Y are projections onto $\mathcal{H}_1$ and $\mathcal{H}_2$, respectively, then this inequality can be written as

$$X^*AY(Y^*AY)^{-1}Y^*AX \leq \left(\frac{M-m}{M+m}\right)^2 X^*AX. \tag{21}$$

A minor argument shows that this remains true if X, Y are mutually orthogonal projections whose ranges do not span all of $\mathcal{H}$. This proves the inequality (19) in the special case when Φ is the identity map.

Let $\alpha = (M-m)/(M+m)$. As pointed out in Remark 1, the inequality (21) is equivalent to the statement

$$\begin{bmatrix} \alpha X^*AX & X^*AY \\ Y^*AX & \alpha Y^*AY \end{bmatrix} \geq 0.$$

From this we get the inequality (19) for every 2-positive linear map Φ. □

Remark 8. The referee has proved, under the stronger hypothesis that Φ is 3-positive, a somewhat stronger inequality than our (19).

Remark 9. The inequality (21) was proved recently in [11] by a different argument. In the scalar case generally the Wielandt inequality (15) is used to derive the Kantorovich inequality (16). In our proof for the operator version we have gone in the opposite direction. Similar ideas have been used by F. Zhang [12].

Acknowledgements. The second author thanks the Indian Statistical Institute and NSERC (Canada) for sponsoring his visit to Delhi in January 2000, when this work was done.

References

[1] Ando, T.: Concavity of certain maps on positive definite matrices and applications to Hadamard products. Linear Algebra and its Applications **26**, 203–241 (1979)

[2] Bhatia, R.: Matrix Analysis. Berlin–Heidelberg–New York: Springer, 1997

[3] Bhatia, R., Davis, Ch.: A better bound on the variance. American Mathematical Monthly **107**, 353–356 (2000)

[4] Choi, M.-D.: Some assorted inequalities for positive linear maps on C^*-algebras. Journal of Operator Theory **4**, 271–285 (1980)

[5] Davis, Ch.: Various averaging operations onto subalgebras. Illinois J. Math. **3**, 538–553 (1959)

[6] Horn, R.A., Johnson, C.R.: Matrix Analysis. Cambridge: Cambridge University Press, 1985

[7] Kadison, R.V.: A generalized Schwarz inequality and algebraic invariants for C^*-algebras. Ann. of Math. **56**, 494–503 (1952)

[8] Lieb, E.H., Ruskai, M.B.: Some operator inequalities of the Schwarz type, Advances in Math. **12**, 269–273 (1974)

[9] Marshall, A.W., Olkin, I.: Matrix versions of Cauchy and Kantorovich inequalities. Aequationes Math. **40**, 89–93 (1990)

[10] Stinespring, W.F.: Positive functions on C^*-algebras. Proc. Amer. Math. Soc. **6**, 211–216 (1955)

[11] Wang, S.-G., Ip, W.-C.: A matrix version of the Wielandt inequality and its applications to statistics. Lin. Alg. Appl. **296**, 171–181 (1999)

[12] Zhang, F.: Matrix inequalites in the Löwner ordering by means of block matrices and Schur complements. Preprint

Communicated by H. Araki