Arithmetic-geometric mean inequalities for matrices

T. ANDO, Hokusei Gakuen University, Sapporo, Japan

The arithmetic-geometric mean inequality for positive numbers a, $b > 0$ or the Schwarz inequality for complex numbers x, y

$$\sqrt{ab} \le \frac{a+b}{2} \quad \text{or} \quad |xy| \le \frac{|x|^2+|y|^2}{2}$$

is the most fundamental inequality. One of its generalization is the Young inequality

$$a^t b^{1-t} \le ta + (1-t)b \quad (0 < t < 1)$$

or

$$|xy| \le \frac{1}{p}|x|^p + \frac{1}{q}|y|^q \quad \left(p, q > 1; \frac{1}{p} + \frac{1}{q} = 1\right).$$

If scalars a, b are replaced with matrices A, B, there is no problem in understanding $A, B \ge 0$ as positive semi-definiteness of A and B. Correspondingly, the order relation, *Löwner order*, $X \ge Y$ between two Hermitian matrices is induced by the cone of positive semi-definite matrices.

There are, however, several directions for matrix generalizations of the Schwarz inequality as well as the Young one. We take up two of them.

The first one is an approach using matrix inequalities, that is, to discuss the inequalities with respect to the Löwner order while the second one is to discuss the eigen (or singular) value inequalities and their variants.

In the direction of matrix inequalities a serious obstacle is in that for $A, B \geq 0$ the product *AB* is Hermitian (and positive semi-definite) only if *A* and *B* are commutative. If *A* and *B* are commutative, through simultaneous diagonalization, the situations are almost the same as the scalar case.

There are several natural ways of defining a positive definite *geometric mean* for (non-commuting) positive definite *A*, *B* > 0. One candidate is

$$\mathbf{G}(A, B) \equiv \exp\left(\frac{\log A + \log B}{2}\right).$$

It is rather surprising to see that with this definition the arithmetic-geometric mean inequality is not valid in general. The only valid inequality is

$$\log \mathbf{G}(A, B) \leq \log\left(\frac{A + B}{2}\right).$$

The other candidate, denoted by $A\#B$, is

$$A\#B \equiv A^{1/2} \cdot (A^{-1/2} B A^{-1/2})^{1/2} \cdot A^{1/2},$$

for which the arithmetic-geometric mean inequality is valid. This geometric mean $A\#B$ admits several nice characterizations. First $A\#B$ is proved to be the maximum of $X \geq 0$ for which

$$\begin{bmatrix} A & X \\ X & B \end{bmatrix} \geq 0.$$

This implies that $A\#B$ is a unique positive definite solution of the matrix quadratic equation $XA^{-1}X = B$. It is seen from this that despite its asymmetric appearance $A\#B$ coincides with $B\#A$.

There are many evidences which support rightness of this definition of geometric mean. $A\#B$ becomes the midpoint of a geodesic curve connecting A and B with respect to a natural Riemannian metric on the cone of positive definite matrices. Also there is an iteration scheme leading to this geometric mean $A\#B$ using the arithmetic means and the harmonic means. Starting from $X_0 \equiv A$ and $Y_0 \equiv B$, define

$$X_n \equiv \frac{X_{n-1} + Y_{n-1}}{2}, \qquad Y_n \equiv \left\{ \frac{(X_{n-1})^{-1} + (Y_{n-1})^{-1}}{2} \right\}^{-1} \qquad n = 1, 2, \ldots$$

Then it is proved that both X_n and Y_n converge to $A\#B$. In the same spirit a possible matrix generalization of $a^t b^{1-t}$ may be

$$A\#_t B \equiv A^{1/2}\cdot(A^{-1/2}BA^{-1/2})^{1-t}\cdot A^{1/2}$$

for which the Young inequality is valid.

Let us turn to the eigen (or singular) value inequalities. For matrices *X*, *Y* the singular value inequalities $\sigma_k(X) \leq \sigma_k(Y)\ k = 1, 2, \ldots,$ where the singular values are arranged in non-increasing order, is equivalent to the existence of a unitary matrix *U* such that $|X| \leq U^*|Y|U$. In this connection the Schwarz inequality or more generally the Young inequality is valid in the following form ; for general matrices *X*, *Y* and for $1 < p, q < \infty$ with $1/p + 1/q = 1$ there is a unitary matrix *U* such that

$$|XY| \leq U^* \left\{\frac{1}{p}|X|^p + \frac{1}{q}|Y|^q\right\} U.$$

A weaker comparison is based on (weak) *majorization* relation among singular values, defined as

$$\sum_{j=1}^{k}\sigma_j(X) \leq \sum_{j=1}^{k}\sigma_j(Y) \qquad k = 1, 2, \ldots$$

This is equivalent to the existence of a finite number of unitary matrices $U_j\ (j = 1, 2, \ldots, N)$ and a finite number of non-negative numbers $\alpha_j\ (j = 1, 2, \ldots, N)$ with $\sum_{j=1}^{N}\alpha_j = 1$ such that

$$|X| \leq \sum_{j=1}^{N} \alpha_j U_j^* |Y| U_j.$$

It is, however, more convenient to formulate this condition as $|||X||| \leq |||Y|||$ for all unitarily invariant norms $|||\cdot|||$. There is the Young inequality of the form; for $A, B \geq 0$ and for general matrix *X*

$$|||A^t X B^{1-t} + A^{1-t} X B^t||| \leq |||AX + XB||| \qquad (0 < t < 1).$$

Up to this point, for both matrix and singular value inequalities, the most basic fact is that we are treating a *pair* of matrices. We shall discuss, when $n \geq 3$, what is a reasonable definition of *geometric mean* of $A_j \geq 0 \ (j = 1, 2, \ldots, n)$ and what kind of arithmetic-geometric mean inequalities are expected.