Mathematical
Inequalities
& Applications
Volume 12, Number 2 (2009), 239–254

MATRIX VERSIONS OF YOUNG'S INEQUALITY

TOMAŽ KOSEM

(*Communicated by R. Mathias*)

Abstract. A matrix majorization version of the general Young's inequality $xy \leqslant \Phi(x)+\Psi(y)$ is presented. Multivariate Young's inequality is extended to the matrix setting by means of the geometric mean of positive semidefinite matrices. Also some refined Hilbert-Schmidt norm generalizations of Young's inequality are given and a symmetrized Young's inequality for unitarily invariant norms is proved.

1. Preliminaries

Let $\mathscr{M}_n$ be the space of $n \times n$ complex matrices. A norm $|||\cdot|||$ on $\mathscr{M}_n$ is called *unitarily invariant* if $|||UAV||| = |||A|||$ for all $A,U,V \in \mathscr{M}_n$ with U,V unitary. Notation $A \leqslant B$ denotes the Löwner partial order, i.e., $A \leqslant B$ if and only if $B-A$ is a positive-semidefinite matrix, where A,B are Hermitian. If not otherwise stated, a capital letter will always denote a complex $n \times n$ matrix. Notations $\lambda(A)$ and $s(A)$ will denote vectors of eigenvalues and singular values of a matrix A, respectively. We assume the reader's familiarity with the theory of majorization and Ky Fan principles (dominance, maximum, majorization) (see e.g. [5], [17]). Notation $\langle \cdot,\cdot \rangle$ will mean a standard scalar product on $\mathbb{C}^n$. Vectors from $\mathbb{C}^n$ will be denoted by small letters.

Whenever the term Young's inequality turns up, we usually think of

$$|ab| \leqslant \frac{|a|^p}{p} + \frac{|b|^q}{q}, \quad a,b \in \mathbb{C}, \quad \frac{1}{p}+\frac{1}{q}=1,$$

but this is only a very special case of a general Young's inequality, though the most important one. In order to present the general version, we recall some definitions and facts from [15].

DEFINITION 1.1. A *Young's function* is a convex function $\Phi : \mathbb{R} \to [0,\infty]$, which satisfies conditions $\Phi(x) = \Phi(-x)$, $\Phi(0) = 0$ and $\lim_{x\to\infty}\Phi(x) = +\infty$. The function $\Psi : \mathbb{R} \to [0,\infty]$, defined by $\Psi(y) = \sup\{x|y| - \Phi(x) : x \geqslant 0\}$, is called the *complementary function*.

Mathematics subject classification (2000): 15A45, 15A60.
Keywords and phrases: unitarily invariant norm, Young's inequality, Young's function, positive-semidefinite matrix, eigenvalue, majorization, convex function, geometric mean, Hilbert-Schmidt norm, Hadamard product.

It follows from the definition that $\Psi(0) = 0$, $\Psi(-y) = \Psi(y)$ and Ψ is a convex function satisfying $\lim_{y\to\infty}\Psi(y) = +\infty$. It is also evident that the pair (Φ,Ψ) satisfies *Young's inequality*

$$xy \leqslant \Phi(x)+\Psi(y), \quad x,y \in \mathbb{R}.$$

The complementary function Ψ is the smallest convex Young's function that satisfies the Young's inequality.

It is a well known fact that the complementary function of Young's function $\Phi(x) = \frac{|x|^p}{p}, p > 1$ is $\Psi(y) = \frac{|y|^q}{q}$, where $\frac{1}{p}+\frac{1}{q} = 1$. The example $\Phi(x) = |x|$ shows that the complementary function of a continuous Young's function on $\mathbb{R}$ can be a jump function. Namely, in this case

$$\Psi(y) = \begin{cases} 0, & |y| \leqslant 1 \\ \infty, & |y| > 1 \end{cases}.$$

The definiton of the complementary function presented above is simple but not informatory. Convexity enables us to obtain an alternative definition. Young's function $\Phi : [0,\infty) \to [0,\infty]$ admits an integral representation of the form

$$\Phi(x) = \int_0^x \phi(t)\,dt, \quad x \geqslant 0$$

where $\phi : [0,\infty) \to [0,\infty]$ is a nondecreasing left continuous function with $\phi(0) = 0$. If $\Phi(x) = +\infty$ for $x > a$ then $\phi(x) = +\infty$, $x > a > 0$. We let

$$\psi(u) = \inf\{t : \phi(t) > u\}, \quad u \geqslant 0.$$

Then $\psi(0) = 0$ and ψ is a nondecreasing Borel function. Define

$$\Psi(y) = \int_0^y \psi(u)\,du, \quad y \geqslant 0.$$

Then Ψ is a Young's function and it follows from the next theorem that it is complementary to Φ (see [15]).

THEOREM 1.2. *Let $\Phi : [0,\infty) \to [0,\infty]$ be a Young's function and let Ψ be associated to Φ as above. Then they satisfy Young's inequality*

$$xy \leqslant \Phi(x)+\Psi(y), \quad x,y \geqslant 0,$$

with equality when $y = \phi(x)$ or $x = \psi(y)$.

Let us consider again the famous special case of Young's inequality for a while and survey some of its matrix generalizations. Bhatia and Kittaneh [7] obtained several matrix versions of arithmetic-geometric mean inequality, for instance

$$s_j(AB) \leqslant s_j(\tfrac{1}{2}(A^2+B^2)),$$

where A,B are positive semidefinite, and consequently

$$|||AB||| \leqslant \left|\left|\left|\tfrac{1}{2}(A^2+B^2)\right|\right|\right|$$

for any unitarily invariant norm. They also observed

$$|||AB||| \leqslant \tfrac{1}{4} \left|\left|\left|(|A|+|B^*|)^2\right|\right|\right|$$

for *all* matrices A,B and

$$\lambda_j(AB) \leqslant \lambda_j(\tfrac{1}{2}(A^2+B^2)),$$

where A,B are positive definite matrices, and consequently

$$\left|\left|\left|B^{1/2}AB^{1/2}\right|\right|\right| \leqslant \left|\left|\left|\tfrac{1}{2}(A^2+B^2)\right|\right|\right|.$$

Furuta and Yanagida [9] proved an operator version of Young's inequality.

THEOREM 1.3. *Let A and B be positive invertible operators on a Hilbert space $\mathscr{H}$. Then the following inequality holds for $0 \leqslant \lambda \leqslant 1$:*

$$(1-\lambda)A+\lambda B \geqslant A^{1/2}(A^{-1/2}BA^{-1/2})^{\lambda}A^{1/2} \geqslant ((1-\lambda)A^{-1}+\lambda B^{-1})^{-1}.$$

This theorem could be viewed as an operator version of the inequality between generalized arithmetic, geometric and harmonic mean. The proof is very short and is reduced to verifying the scalar inequality $(1-\lambda)+\lambda x \geqslant x^{\lambda}$ for $x \geqslant 0$ and $0 \leqslant \lambda \leqslant 1$. Using their result consecutively it is easy to see the following corollary.

COROLLARY 1.4. *Let $A_1,A_2,\ldots,A_n$ be positive invertible operators on a Hilbert space $\mathscr{H}$. Then the following inequality holds for $0 \leqslant \lambda_i, i=1,2,\ldots,n$ with $\lambda_n > 0$ and $\sum_{i=1}^{n}\lambda_i = 1$:*

$$\begin{aligned}\lambda_1A_1+\lambda_2A_2+\cdots+\lambda_nA_n &\geqslant G_n(A_1,A_2,\ldots,A_n;\lambda_1,\lambda_2,\ldots,\lambda_n)\\ &\geqslant (\lambda_1A_1^{-1}+\lambda_2A_2^{-1}+\cdots+\lambda_nA_n^{-1})^{-1},\end{aligned}$$

where $G_n(-;-)$ is defined recursively by

$$G_n(A_1,A_2,\ldots,A_n;\lambda_1,\lambda_2,\ldots,\lambda_n)$$

$$:= G_2\left(A_1, G_{n-1}\left(A_2,\ldots,A_n;\tfrac{\lambda_2}{\lambda_2+\cdots+\lambda_n},\ldots,\tfrac{\lambda_n}{\lambda_2+\cdots+\lambda_n}\right);\lambda_1,\lambda_2+\cdots+\lambda_n\right),$$

$$G_2(A,B;1-\lambda,\lambda) := A^{1/2}(A^{-1/2}BA^{-1/2})^{\lambda}A^{1/2}.$$

Ando [1] proved a singular value version of Young's inequality.

THEOREM 1.5. *Let $p,q > 1$ with $\frac{1}{p}+\frac{1}{q}=1$. Then for any pair A,B of matrices there is a unitary matrix U depending on A,B such that*

$$U^*|AB^*|U \leqslant \tfrac{1}{p}|A|^p+\tfrac{1}{q}|B|^q.$$

The inequality in the theorem can be reformulated into

$$s_j(AB^*) \leqslant s_j(\tfrac{1}{p}|A|^p+\tfrac{1}{q}|B|^q),\ j=1,2,\ldots,n.$$

Similarly as in Furuta-Yanagida case we can extend Ando's result.

COROLLARY 1.6. *Let $A_1, A_2, \ldots, A_n$ be positive semidefinite matrices and $p_1, p_2, \ldots, p_n > 1$ real numbers satisfying $\frac{1}{p_1} + \frac{1}{p_2} + \cdots + \frac{1}{p_n} = 1$. Then there exist unitary matrices $U_1, U_2, \ldots, U_n$ depending on $A_1, A_2, \ldots, A_n$ such that*

$$|A_1U_1A_2U_2 \cdots A_nU_n| \leqslant \frac{1}{p_1}A_1^{p_1} + \frac{1}{p_2}A_2^{p_2} + \cdots + \frac{1}{p_n}A_n^{p_n}.$$

In fact, U_1 may be chosen to be the identity matrix.

In the section that follows we will obtain a majorization version of general Young's inequality for eigenvalues, i.e.

$$\lambda(AB) \prec_w \lambda(\Phi(A) + \Psi(B)),$$

where A, B are positive semidefinite matrices and also consider the case of equality. In the third section the multivariate version of Young's inequality will be treated and the noncommutative analogue of a product of nonnegative numbers will be the n^{th} power of the geometric mean of n positive semidefinite matrices. The fourth section is devoted to Young's inequalities for Hilbert-Schmidt norm, when the derivative of Φ is a convex continuous function, and the last section deals with a symmetrized Young's inequality.

2. Young's inequality for eigenvalues

We begin this section with a short lemma (see [4, p. 69]), which will be useful in the sequel.

LEMMA 2.1. *Let A be a positive definite matrix and x, y arbitrary vectors. Then*

$$\langle Ax, x\rangle\langle A^{-1}y, y\rangle \geqslant |\langle x, y\rangle|^2.$$

Proof. Write $A = B^2$, where B is a positive definite matrix. Then

$$\langle Ax, x\rangle\langle A^{-1}y, y\rangle = \|Bx\|^2 \left\|B^{-1}y\right\|^2.$$

But

$$\|Bx\| \left\|B^{-1}y\right\| \geqslant |\langle Bx, B^{-1}y\rangle| = |\langle x, y\rangle|$$

by Cauchy-Schwarz inequality, which yields the desired result.

An alternative proof uses Cauchy-Schwarz inequality for inner product $[\cdot,\cdot] := \langle A\cdot,\cdot\rangle$. □

The following two lemmas (see [3] and references therein) will also be needed.

LEMMA 2.2. *Let A be a Hermitian matrix with $\sigma(A)$ contained in I and let f be a convex function on I. Then for every unit vector x*

$$f(\langle Ax, x\rangle) \leqslant \langle f(A)x, x\rangle.$$

LEMMA 2.3. *Let A be a Hermitian matrix. Then*

$$\sum_{j=1}^{k} \lambda_j(A) = \max \sum_{j=1}^{k} \langle Ax_j, x_j \rangle \quad k = 1, 2, \ldots, n,$$

where the maximum is taken over all choices of orthonormal vectors $x_1, x_2, \ldots, x_n$.

Having these lemmas at our disposal, it is not difficult to prove the next theorem. Similar steps and ideas of the proof can also be detected in [3].

THEOREM 2.4. *Let A and B be positive semidefinite matrices. Then for a continuous Young's function Φ with a continuous complementary function Ψ the following inequality holds:*

$$\lambda(AB) \prec_w \lambda(\Phi(A) + \Psi(B)).$$

REMARK. The continuity assumption on functions Φ, Ψ is adopted here for the sake of simplicity and clarity. In the general case the intervals of finiteness of Φ and Ψ dictate the choice of matrices A and B, respectively.

Proof. Let us prove first the special case when B is invertible. The general case will easily follow from the special one by continuity argument. Let x be a *unit eigenvector* for $B^{1/2}AB^{1/2}$. We will establish the following:

$$\langle B^{1/2}AB^{1/2}x, x \rangle \leqslant \langle Ax, x \rangle \langle Bx, x \rangle. \tag{$\star$}$$

We notice that $\langle B^{1/2}AB^{1/2}x, x \rangle$ is an *eigenvalue* for $B^{1/2}AB^{1/2}$ corresponding to eigenvector x.

Without loss of generality we may hence assume that $B^{1/2}AB^{1/2} = D$, where D is a diagonal matrix $\operatorname{diag}(d_1, d_2, \ldots, d_n)$ and x is the first standard unit vector $e_1 = (1, 0, \ldots, 0)$. Then $A = B^{-1/2}DB^{-1/2}$ and

$$\langle Ae_1, e_1 \rangle = \langle DB^{-1/2}e_1, B^{-1/2}e_1 \rangle$$

Write $B^{-1/2}e_1 = (y_1, \ldots, y_n)$. Then

$$\langle DB^{-1/2}e_1, B^{-1/2}e_1 \rangle = \langle (d_1y_1, \ldots, d_ny_n), (y_1, \ldots, y_n) \rangle \geqslant d_1|y_1|^2 = d_1 \langle B^{-1/2}e_1, e_1 \rangle^2.$$

We also have $\langle Be_1, e_1 \rangle \geqslant \langle B^{1/2}e_1, e_1 \rangle^2$ by Lemma 2.2. Consequently

$$\langle Ae_1, e_1 \rangle \langle Be_1, e_1 \rangle \geqslant d_1 \langle B^{-1/2}e_1, e_1 \rangle^2 \langle B^{1/2}e_1, e_1 \rangle^2 \geqslant d_1 \langle e_1, e_1 \rangle^4 = d_1 = \langle De_1, e_1 \rangle$$

by Lemma 2.1.

If the matrix $B^{1/2}AB^{1/2}$ is not diagonal or $x \neq e_1$, we choose a unitary matrix U such that $B^{1/2}AB^{1/2} = U^*DU$, where D is a diagonal matrix and $Ux = e_1$. Then $D = \tilde{B}^{1/2}\tilde{A}\tilde{B}^{1/2}$ for $\tilde{B} = UBU^*, \tilde{A} = UAU^*$ and we can use the special case.

Let $x_1, x_2, \dots x_k$ be the orthonormal eigenvectors, corresponding to eigenvalues $\lambda_1 \geqslant \lambda_2 \geqslant \cdots \geqslant \lambda_k$ of $B^{1/2}AB^{1/2}$. Then we have

$$\begin{aligned}
\sum_{i=1}^{k} \lambda_i(AB) &= \sum_{i=1}^{k} \lambda_i(B^{1/2}AB^{1/2}) \\
&= \sum_{i=1}^{k} \langle B^{1/2}AB^{1/2}x_i, x_i\rangle \\
&\leqslant \sum_{i=1}^{k} \langle Ax_i, x_i\rangle \langle Bx_i, x_i\rangle \quad \text{(by } (\star)\text{)} \\
&\leqslant \sum_{i=1}^{k} (\Phi(\langle Ax_i, x_i\rangle) + \Psi(\langle Bx_i, x_i\rangle) \quad \text{(scalar Young's inequality)} \\
&\leqslant \sum_{i=1}^{k} (\langle \Phi(A)x_i, x_i\rangle) + \langle \Psi(B)x_i, x_i\rangle) \quad \text{(by Lemma 2.2)} \\
&= \sum_{i=1}^{k} \langle (\Phi(A) + \Psi(B))x_i, x_i\rangle \\
&\leqslant \sum_{i=1}^{k} \lambda_i(\Phi(A) + \Psi(B)) \quad \text{(by Lemma 2.3)}
\end{aligned}$$

In the general case, when B is not necessarily invertible, we take $\varepsilon > 0$ and define $B_\varepsilon := B + \varepsilon I$, which is clearly invertible and we get by the special case

$$\lambda(AB_\varepsilon) \prec_w \lambda(\Phi(A) + \Psi(B_\varepsilon)).$$

Since $\lambda(AB_\varepsilon)$ and $\lambda(\Phi(A) + \Psi(B_\varepsilon))$ converge to $\lambda(AB)$ and $\lambda(\Phi(A) + \Psi(B))$, respectively, as $\varepsilon \downarrow 0$, the proof is completed. □

We immediately obtain the following corollaries.

COROLLARY 2.5. *Let A, B be arbitrary matrices, Φ and Ψ as above. Then for every unitarily invariant norm $|||\cdot|||$*

$$\left|\left|\left| |AB^*|^2 \right|\right|\right| \leqslant \left|\left|\left| \Phi(|A|^2) + \Psi(|B|^2) \right|\right|\right|.$$

COROLLARY 2.6. *Let A, B be positive semidefinite matrices and $p, q > 1$ such that $\frac{1}{p} + \frac{1}{q} = 1$. Then for every unitarily invariant norm $|||\cdot|||$*

$$\left|\left|\left| B^{1/2}AB^{1/2} \right|\right|\right| \leqslant \left|\left|\left| \tfrac{1}{p}A^p + \tfrac{1}{q}B^q \right|\right|\right|.$$

The case of equality is considered in the following theorem.

THEOREM 2.7. *Let A, B be positive semidefinite matrices and suppose that Φ, Ψ are such that $\Phi' = f, \Psi' = f^{-1}$, where $f : [0, \infty) \to [0, \infty)$ is a bijective increasing continuous function with $f(0) = 0$. Then*

$$\lambda_k(AB) = \lambda_k(\Phi(A) + \Psi(B)),\ k = 1, 2, \dots, n$$

if and only if $B = f(A)$.

Proof. Considering the proof of Theorem 2.4 we obtain the chain of equalities

$$\begin{aligned}\lambda_k(AB) &= \lambda_k(B^{1/2}AB^{1/2}) = \langle B^{1/2}AB^{1/2}x_k, x_k\rangle = \langle Ax_k, x_k\rangle\langle Bx_k, x_k\rangle \\ &= \Phi(\langle Ax_k, x_k\rangle) + \Psi(\langle Bx_k, x_k\rangle) = \langle \Phi(A)x_k, x_k\rangle + \langle \Psi(B)x_k, x_k\rangle,\end{aligned}$$

from where we conclude

$$\begin{aligned}\Phi(\langle Ax_k, x_k\rangle) &= \langle \Phi(A)x_k, x_k\rangle \\ \Psi(\langle Bx_k, x_k\rangle) &= \langle \Psi(B)x_k, x_k\rangle\end{aligned}$$

which yields the fact that x_k is also an eigenvector for A and B. Namely, Φ and Ψ are strictly convex functions due to the properties of f, hence $\langle Ax_k, x_k\rangle$ is a trivial convex combination of eigenvalues of A, i.e., $\langle Ax_k, x_k\rangle$ is an eigenvalue of A and analogously for $\langle Bx_k, x_k\rangle$. So we can diagonalize A and B simultaneously: without loss of generality we take $x_j = e_j,\ j = 1,2,\ldots,n$ and $A = \mathrm{diag}(a_1, a_2, \ldots, a_n), B = \mathrm{diag}(b_1, b_2, \ldots, b_n)$. Then $a_k b_k = \Phi(a_k) + \Psi(b_k),\ k = 1,2,\ldots,n$, which is possible only if $b_k = f(a_k),\ k = 1,2,\ldots,n$. But this last condition means exactly $B = f(A)$. $\square$

3. Multivariate Young's inequality

There are at least two possible multidimensional generalizations of Young's inequality ([8, 14]). For our purpose the generalization of Cooper [8] will be appropriate.

THEOREM 3.1. *Let $\phi_1, \phi_2, \ldots, \phi_n$ be continuous, increasing functions defined for $x \geqslant 0$. Let $F_i(x) := x\phi_i(x)$ and suppose that*

$$\prod_{i=1}^{n} F_i^{-1}(x) = x$$

for all $x \geqslant 0$. Then

$$\prod_{i=1}^{n} a_i \leqslant \sum_{i=1}^{n} \int_0^{a_i} \phi_i(x)\,dx$$

for all $a_1, a_2, \ldots, a_n \geqslant 0$.

Cooper's theorem can be proved using geometrically more intuitive Oppenheim's inequality ([14]).

THEOREM 3.2. *If, for $i = 1,2,\ldots,n$, the function f_i is a continuous non-negative increasing function of x for $x \geqslant 0$, then, provided that at least one of the numbers $f_1(0), \ldots, f_n(0)$ is zero,*

$$\prod_{i=1}^{n} f_i(t_i) \leqslant \sum_{i=1}^{n} \int_0^{t_i} \prod_{j \neq i} f_j\, df_i,$$

where $t_i \geqslant 0$ and the integrals are taken in the sense of Riemann-Stieltjes.

Cooper's inequality now follows from Oppenheim's result by considering a curve in $\mathbb{R}^n$ with parametrization $x_i = F_i^{-1}(t)$. Oppenheim's inequality has a nice geometrical interpretation. Consider the curve in n-dimensional Euclidean space with rectangular coordinates $x_i = f_i(t)$, where $f_i(t)$ is a continuous, non-negative, and increasing for $t \geqslant 0$. The integral

$$\int_0^{t_i} \prod_{j \neq i} f_j \, d f_i$$

represents the volume V_i bounded by the coordinate planes, other than x_i and the cylinders which project the curve onto these coordinate planes.

The assumption $\prod_{i=1}^n F_i^{-1}(x) = x$ in Theorem 3.1 implies $\phi_i(0) = 0$ and $\phi_i(x) \to \infty$ as $x \to \infty$ for all i. The following lemma (see [8]) convinces us that Theorem 3.1 is truly a generalization of the inequality

$$ab \leqslant \int_0^a \phi(t)\,dt + \int_0^b \phi^{-1}(s)\,ds,$$

where $a,b \geqslant 0$ and ϕ is an appropriate function.

LEMMA 3.3. *If $\phi : [0,\infty) \to [0,\infty)$ is a bijective continuous increasing functions, satisfying $\phi(0) = 0$ and $\psi = \phi^{-1}$, then functions F_1, F_2, defined by $F_1(x) := x\phi(x)$ and $F_2(x) := x\psi(x)$, satisfy*

$$F_1^{-1}(x) F_2^{-1}(x) = x$$

for all $x \geqslant 0$ and conversely.

In [2] authors listed what properties should be required for a reasonable geometric mean $G(A,B,C)$ of three positive definite matrices A,B,C. It is clear what the corresponding conditions would be for k matrices when $k > 3$.

P1 *Consistency with scalars.* If A,B,C commute then $G(A,B,C) = (ABC)^{1/3}$.

P1' This implies $G(A,A,A) = A$.

P2 *Joint homogeneity.* $G(aA,bB,cC) = (abc)^{1/3} G(A,B,C) \quad (a,b,c > 0)$.

P2' This implies $G(aA,aB,aC) = aG(A,B,C) \quad (a > 0)$.

P3 *Permutation invariance.* For any permutation $\pi(A,B,C)$ of (A,B,C) we have $G(A,B,C) = G(\pi(A,B,C))$.

P4 *Monotonicity.* The map $(A,B,C) \mapsto G(A,B,C)$ is monotone, i.e., if $A \geqslant A_0, B \geqslant B_0, C \geqslant C_0$, then $G(A,B,C) \geqslant G(A_0,B_0,C_0)$ in the positive semidefinite ordering.

P5 *Continuity from above.* If $\{A_n\},\{B_n\},\{C_n\}$ are monotonic decreasing sequences (in the positive semidefinite ordering) converging to A,B,C, respectively, then $\{G(A_n,B_n,C_n)\}$ converges to $G(A,B,C)$.

P6 *Congruence invariance.* $G(S^*AS, S^*BS, S^*CS) = S^*G(A,B,C)S$ for any invertible S.

Once a geometric mean for three positive definite matrices is defined so as to satisfy P1-P6, by monotonicity we can uniquely extend the definition of $G(A,B,C)$ for every triple of positive semidefinite matrices (A,B,C) by setting

$$G(A,B,C) = \lim_{\varepsilon \downarrow 0} G(A+\varepsilon I, B+\varepsilon I, C+\varepsilon I).$$

We can derive a stronger form of P6 with help of P4 and P5:

P6' $G(S^*AS, S^*BS, S^*CS) \geqslant S^*G(A,B,C)S$ for all S.

The geometric mean of two positive definite matrices is uniquely defined via

$$G(A,B) = A\#B := A^{1/2}(A^{-1/2}BA^{-1/2})^{1/2}A^{1/2}$$

but in the case of k-tuples ($k>2$) we have at least two alternative definitions. In [2] an inductive definition of a geometric mean was introduced. Suppose we have defined the geometric mean $G(X_1,\ldots,X_k)$ of k positive definite matrices $X_1,\ldots,X_k$. Consider the transformation on $(k+1)$-tuples of positive definite matrices $A=(A_1,\ldots,A_{k+1})$ by

$$T(A) := (G((A_i)_{i\neq 1}), G((A_i)_{i\neq 2}), \ldots, G((A_i)_{i\neq k+1})).$$

We define the sequence $\{T^r(A)\}_{r=1}^{\infty}$. The limit of this sequence exists and has the form $(\tilde{A},\ldots,\tilde{A})$. We define $G(A_1,\ldots,A_{k+1})$ to be $\tilde{A}$. So defined geometric mean satisfies properties P1-P6.

Kosaki proposed the following definition of a weighted geometric mean ([2, p. 324]). Let $\alpha_j \geqslant 0$ $(j=1,2,\ldots,k)$ satisfy $\sum_{j=1}^k \alpha_j = 1$. For positive definite matrices A_j $(j=1,2,\ldots,k)$ we define

$$(A_1,\ldots,A_k;\alpha_1,\ldots,\alpha_k) := \frac{1}{\prod_{j=1}^k \Gamma(\alpha_j)} \int_{\Delta_k} \left\{ \sum_{j=1}^k \lambda_j A_j^{-1} \right\}^{-1} \left\{ \prod_{j=1}^k \lambda_j^{\alpha_j - 1} \right\} d\lambda_1 \cdots d\lambda_k,$$

where Δ_k is a standard $(k-1)$-simplex in $\mathbb{R}^k$, i.e.,

$$\Delta_k := \left\{ (\lambda_1,\ldots,\lambda_k) : \lambda_j \geqslant 0, \sum_{j=1}^k \lambda_j = 1 \right\}.$$

Kosaki's geometric mean is then

$$G_K(A_1,\ldots,A_k) = (A_1,\ldots,A_k; \tfrac{1}{k},\ldots,\tfrac{1}{k}).$$

The weighted geometric mean $(A_1,\ldots,A_k;\alpha_1,\ldots,\alpha_k)$ has an expected property

$$(A_1,\ldots,A_k;\alpha_1,\ldots,\alpha_k) = A_1^{\alpha_1} \cdots A_k^{\alpha_k},$$

when $\{A_j : j=1,\ldots,k\}$ is a commuting k-tuple. It is not hard to prove that G_K satisfies properties P1–P6.

From now on the term *geometric mean* will stand for an arbitrary mapping G from a set of n-tuples of positive semidefinite $n\times n$ matrices into a set of positive semidefinite $n\times n$ matrices satisfying P1, P2 and P6'. We have already seen that such mappings do exist. Before stating the matrix generalization of Cooper's theorem we prove the following lemma, which is analogous to the key fact $(\star)$, used in the proof of Theorem 2.4

LEMMA 3.4. *Let $A_1,A_2,\dots,A_n$ be positive semidefinite matrices and $G(A_1,A_2,\dots,A_n)$ their geometric mean. Let x be an arbitrary vector. Then*

$$\langle G(A_1,A_2,\dots,A_n)x,x\rangle^n \leqslant \langle A_1x,x\rangle\langle A_2x,x\rangle\cdots\langle A_nx,x\rangle.$$

Proof. Take S a matrix whose first column is x and has zeroes elsewhere. Then $S^*XS=\langle Xx,x\rangle E_{11}$ (E_{11} is a matrix whose (1,1)-entry is 1 and has zeroes elsewhere) and using properties P1, P2 and P6' we obtain

$$\begin{aligned}G(S^*A_1S,S^*A_2S,\dots,S^*A_nS) &= \sqrt[n]{\langle A_1x,x\rangle\langle A_2x,x\rangle\cdots\langle A_nx,x\rangle}E_{11}\\ &\geqslant \langle G(A_1,A_2,\dots,A_n)x,x\rangle E_{11},\end{aligned}$$

therefore the proof is completed. □

Imitating the last part of the proof of Theorem 2.4 and applying Lemma 3.4 we can prove

THEOREM 3.5. *Let $A_1,A_2,\dots,A_n$ be positive semidefinite matrices and $G(A_1,A_2,\dots,A_n)$ their geometric mean. Let ϕ_i,F_i be as above. Define $\Phi_i(a):=\int_0^a\phi_i(x)\,dx$. Then*

$$|||G(A_1,A_2,\dots,A_n)^n||| \leqslant |||\Phi_1(A_1)+\Phi_2(A_2)+\cdots+\Phi_n(A_n)|||$$

for every unitarily invariant norm.

When we consider pairs instead of n-tuples, the inequality from Lemma 3.4

$$\langle G(A,B)x,x\rangle^2 \leqslant \langle Ax,x\rangle\langle Bx,x\rangle$$

can be seen directly. To show this, we first recall that in this case $G(A,B)$ is uniquely defined ($G(A,B)=A\#B:=A^{1/2}(A^{-1/2}BA^{-1/2})^{1/2}A^{1/2}$ when A,B invertible and $\lim_{\varepsilon\downarrow 0}(A+\varepsilon I)\#(B+\varepsilon I)$ otherwise) and that the block matrix

$$\begin{bmatrix} A & A\#B\\ A\#B & B\end{bmatrix}$$

is positive semidefinite. But then 2-by-2 matrix

$$\begin{bmatrix} \langle Ax,x\rangle & \langle (A\#B)y,x\rangle\\ \langle (A\#B)x,y\rangle & \langle By,y\rangle\end{bmatrix}$$

is positive semidefinite for arbitrary vectors x,y. We take $y=x$ and compute the determinant, which is nonnegative and this completes the proof.

At this moment we are able to provide another short proof of Lemma 2.1. Since $A\#A^{-1} = I$, the 2-by-2 matrix

$$\begin{bmatrix} \langle Ax,x\rangle & \langle y,x\rangle \\ \langle x,y\rangle & \langle A^{-1}y,y\rangle \end{bmatrix}$$

is positive semidefinite for arbitrary vectors x,y, therefore $\langle Ax,x\rangle\langle A^{-1}y,y\rangle \geqslant |\langle x,y\rangle|^2$.

Now we have two different types of Young's inequality for matrices:

$$\left|\left|\left|B^{1/2}AB^{1/2}\right|\right|\right| \leqslant |||\Phi(A)+\Psi(B)|||$$

and

$$\left|\left|\left|(A\#B)^2\right|\right|\right| \leqslant |||\Phi(A)+\Psi(B)|||.$$

Are they comparable? As we shall see, the second one is weaker than the first one, that is, we have $\left|\left|\left|(A\#B)^2\right|\right|\right| \leqslant \left|\left|\left|B^{1/2}AB^{1/2}\right|\right|\right|$. To establish this inequality we need the following proposition [5, Prop. IX.1.1].

PROPOSITION 3.6. *Let A,B be any two matrices such that the product AB is normal. Then, for every unitarily invariant norm, we have*

$$|||AB||| \leqslant |||BA|||.$$

Now we are ready to prove the aforementioned inequality.

PROPOSITION 3.7. *Let A,B be positive semidefinite matrices. Then*

$$\left|\left|\left|(A\#B)^2\right|\right|\right| \leqslant \left|\left|\left|B^{1/2}AB^{1/2}\right|\right|\right|$$

for every unitarily invariant norm.

Proof. Let us assume that A,B are invertible. The general case will follow by continuity argument. It suffices to prove

$$|||A\#B||| \leqslant \left|\left|\left|A^{1/2}B^{1/2}\right|\right|\right| = \left|\left|\left|\,\left|A^{1/2}B^{1/2}\right|\,\right|\right|\right| = \left|\left|\left|\left(B^{1/2}AB^{1/2}\right)^{1/2}\right|\right|\right|,$$

due to the principle of majorization. One of the equivalent definitions of the geometric mean of two positive definite matrices says that $A\#B = A^{1/2}UB^{1/2}$, where U is any unitary matrix that makes the right-hand side positive definite. We can let $U = (A^{-1/2}BA^{-1/2})^{1/2}A^{1/2}B^{-1/2}$. So by Proposition 3.6

$$|||A\#B||| = \left|\left|\left|A^{1/2}UB^{1/2}\right|\right|\right| \leqslant \left|\left|\left|UB^{1/2}A^{1/2}\right|\right|\right| = \left|\left|\left|B^{1/2}A^{1/2}\right|\right|\right| = \left|\left|\left|A^{1/2}B^{1/2}\right|\right|\right|.$$

When A,B are merely positive semidefinite, we let $A_\varepsilon := A+\varepsilon I$ and similarly for B, where $\varepsilon > 0$. Since $\left|\left|\left|(A_\varepsilon\#B_\varepsilon)^2\right|\right|\right|$ tends to $\left|\left|\left|(A\#B)^2\right|\right|\right|$ as $\varepsilon \downarrow 0$ and the analogous statement holds for $\left|\left|\left|B^{1/2}AB^{1/2}\right|\right|\right|$, we are finished. $\square$

4. Young's inequality for Hilbert-Schmidt norm

Hirzallah and Kittaneh [10] obtained refined matrix Young's inequality for Hilbert-Schmidt norm.

THEOREM 4.1. *Let A,B be positive semidefinite and X arbitrary. If $p,q>1$ with $\frac{1}{p}+\frac{1}{q}=1$, then*

$$\left\|\frac{1}{p}A^pX+\frac{1}{q}XB^q\right\|_2^2 \geqslant \frac{1}{r^2}\left\|A^pX-XB^q\right\|_2^2+\|AXB\|_2^2,$$

where $r=\max(p,q)$.

This theorem together with its corollaries has motivated the following results. Their proofs can be worked up using similar techniques as those applied in the proofs in [10] and are therefore omitted. The role of [10, Lemma 1], i.e.

$$\left(\frac{a^p}{p}+\frac{b^q}{q}\right)^2 \geqslant a^2b^2+\frac{1}{r^2}(a^p-b^q)^2,$$

is played by the following

LEMMA 4.2. *Let $f:[0,\infty)\to[0,\infty)$ be strictly increasing convex function satisfying $f(0)=0$. Let $\Phi(x):=\int_0^x f(t)\,dt$ and $\Psi(y):=\int_0^y f^{-1}(s)\,ds$. Then for $a,b>0$*

$$\Phi(a)+\Psi(b)\geqslant ab+\frac{1}{2}(a-f^{-1}(b))^2\frac{b}{f^{-1}(b)},$$

$$(\Phi(a)+\Psi(b))^2\geqslant a^2b^2+a(a-f^{-1}(b))^2\frac{b^2}{f^{-1}(b)}.$$

Proof. Assume that $a>f^{-1}(b)$ (it is trivial in the case $a=f^{-1}(b)$ and very similar if $a<f^{-1}(b)$). Then the expression $\phi(a)+\psi(b)-ab$ can be estimated with an area of a rectangular triangle with catheti of length $a-f^{-1}(b)$ and $\alpha(a-f^{-1}(b))$, respectively, where $\alpha=\frac{b}{f^{-1}(b)}$.

The second inequality can easily be derived from the first one:

$$\begin{aligned}(\Phi(a)+\Psi(b))^2-a^2b^2&=(\Phi(a)+\Psi(b)-ab)(\Phi(a)+\Psi(b)+ab)\\&\geqslant\left(\frac{1}{2}(a-f^{-1}(b))^2\frac{b}{f^{-1}(b)}\right)(2ab)\\&=a(a-f^{-1}(b))^2\frac{b^2}{f^{-1}(b)}.\qquad\square\end{aligned}$$

THEOREM 4.3. *Let A,B be positive definite, X arbitrary and f,Φ,Ψ as above. Then*

$$\|\Phi(A)X+X\Psi(B)\|_2^2\geqslant\|AXB\|_2^2+\left\|A^{1/2}(AX-Xf^{-1}(B))g(B)\right\|_2^2,$$

where $g(t):=\frac{t}{\sqrt{f^{-1}(t)}}$.

COROLLARY 4.4. *Let A,B be positive definite, X arbitrary and f,Φ,Ψ as above. Then*

$$\|\Phi(A)X + X\Psi(B)\|_2 = \|AXB\|_2$$

if and only if $AX = Xf^{-1}(B)$.

COROLLARY 4.5. *Let A,B be positive definite and f,Φ,Ψ as above. Then*

$$s_j(\Phi(A) + \Psi(B)) = s_j(AB) \text{ for } j = 1,2,\ldots,n$$

if and only if $A = f^{-1}(B)$.

We end this section with an inequality for Hilbert-Schmidt norm, estimating the right hand side of Young's inequality. The following scalar inequality is needed in the proof (see [16]).

LEMMA 4.6. *Let $f : [0,\infty) \to [0,\infty)$ be strictly increasing function satisfying $f(0) = 0$. Let Φ and Ψ be defined as above. Then for $a,b > 0$*

$$\Phi(a) + \Psi(b) \leqslant af(a) - f(a)f^{-1}(b) + f^{-1}(b)b.$$

THEOREM 4.7. *Let A,B be positive semidefinite and f,Φ,Ψ as above. Then*

$$\|\Phi(A)X + X\Psi(B)\|_2 \leqslant \left\|Af(A)X - f(A)Xf^{-1}(B) + Xf^{-1}(B)B\right\|_2.$$

5. A symmetrized Young's inequality

If we define a new function $\Sigma := \Phi + \Psi$, where Φ,Ψ are complementary Young's functions, we immediately obtain the inequality

$$2|ab| \leqslant \Sigma(|a|) + \Sigma(|b|),$$

which may be seen as a symmetrized form of Young's inequality. This inequality is trivially extended to

$$2|abx| \leqslant (\Sigma(|a|) + \Sigma(|b|))|x|$$

and we will prove a matrix version of this inequality.

This section is an adaptation of a part of [12], from where we recall the following definitions. Given $A \in \mathscr{M}_n$ we define the linear map $S_A : \mathscr{M}_n \to \mathscr{M}_n$ by $S_A(B) = A \circ B$ (Hadamard product). Let $\|\cdot\|$ denote the spectral norm on $\mathscr{M}_n$ and $\|S_A\|$ the induced norm of S_A, i.e.

$$\|S_A\| := \max\{\|A \circ B\| : \|B\| \leqslant 1\}.$$

We also define a partial order $\leqslant_\circ$ on $\mathscr{M}_n$ by

$$A \leqslant_\circ B \iff \|A \circ X\| \leqslant \|B \circ X\| \text{ for all } X \in \mathscr{M}_n.$$

THEOREM 5.1. *Let $x \in \mathbb{R}_+^n$ and Φ, Ψ as above. Then for matrices*

$$K(x) := \left[\Phi(x_i) + \Psi(x_i) + \Phi(x_j) + \Psi(x_j)\right]_{i,j=1}^n, \quad L(x) := \left[2x_ix_j\right]_{i,j=1}^n$$

we have

$$K(x) \geqslant_\circ L(x).$$

Proof. To determine whether $K(x) \geqslant_\circ L(x)$ is equivalent to determining whether

$$\left\| S_{K(x)^{(-1)} \circ L(x)} \right\| \leqslant 1.$$

$K(x)^{(-1)}$ is a Hilbert matrix $\left[\frac{1}{\alpha_i + \alpha_j}\right]_{i,j=1}^n$, where $\alpha_i := \Phi(x_i) + \Psi(x_i)$, therefore it is positive semidefinite. Since $L(x)$ is positive semidefinite too and since the main diagonal entries of $K(x)^{(-1)} \circ L(x)$ are all $\leqslant 1$, we conclude (as in [12]) that $\left\| S_{K(x)^{(-1)} \circ L(x)} \right\| \leqslant 1$. □

COROLLARY 5.2. *Let $A, B, X \in \mathscr{M}_n$ with A, B positive semidefinite. Let Φ, Ψ be as above. Then*

$$\begin{aligned} 2\,|||AXB||| &\leqslant |||(\Phi(A) + \Psi(A))X + X(\Phi(B) + \Psi(B))||| \\ &= |||(\Phi(A)X + X\Psi(B)) + (\Psi(A)X + X\Phi(B))|||. \end{aligned}$$

The proof is similar to the one in [12] and is therefore omitted.

We recall from [11] that the inequality

$$|||AXB||| \leqslant \left|\left|\left| \frac{1}{p}A^pX + \frac{1}{q}XB^q \right|\right|\right|$$

does not hold and the same is true for

$$|||AXB||| \leqslant |||\Phi(A)X + X\Psi(B)|||.$$

Nevertheless, we still have a "weak matrix Young's inequality"

$$|||AXB||| \leqslant \kappa_p \left|\left|\left| \frac{1}{p}A^pX + \frac{1}{q}XB^q \right|\right|\right|$$

with a certain constant $\kappa_p \geqslant 1$ depending only upon p. Analogously,

$$|||AXB||| \leqslant \kappa_q \left|\left|\left| \frac{1}{q}A^q X + \frac{1}{p} X B^p \right|\right|\right|$$

Taking adjoints and considering an optimality of κ_p we notice that $\kappa_p = \kappa_q$ for $\frac{1}{p} + \frac{1}{q} = 1$.

Summing last two inequalities gives

$$2\,|||AXB||| \leqslant \kappa_p \left(\left|\left|\left| \frac{1}{p}A^p X + \frac{1}{q} X B^q \right|\right|\right| + \left|\left|\left| \frac{1}{q}A^q X + \frac{1}{p} X B^p \right|\right|\right| \right). \qquad (\dagger)$$

By Corollary 5.2, we have

$$2\,|||AXB||| \leqslant \left|\left|\left| \left(\frac{1}{p}A^p X + \frac{1}{q} X B^q \right) + \left(\frac{1}{q}A^q X + \frac{1}{p} X B^p \right) \right|\right|\right|,$$

which is a considerable improvement of $(\dagger)$, since an upper estimate for κ_p from [11] is $\dfrac{1}{\cos(\frac{\pi}{2}(\frac{1}{p} - \frac{1}{q}))}$.

Acknowledgement. The author was supported by the ARRS, Slovenian Research Agency.

REFERENCES

[1] T. ANDO, *Matrix Young inequalities*, Operator theory in function spaces and Banach lattices, 33-38, Oper. Theory Adv. Appl., **75**, Birkhäuser, Basel, 1995.

[2] T. ANDO, C. K. LI, R. MATHIAS, *Geometric means*, Linear Algebra Appl., **385** (2004), 305–334.

[3] J. S. AUJLA, F. C. SILVA, *Weak majorization inequalities and convex functions*, Linear Algebra Appl., **369** (2003), 217–233.

[4] E. F. BECKENBACH, R. BELLMAN, *Inequalities*, Second revised printing, Ergebnisse der Mathematik und ihrer Grenzgebiete. Neue Folge, Band 30, Springer-Verlag, New York, Inc., 1965.

[5] R. BHATIA, *Matrix Analysis*, GTM 169, Springer-Verlag, New York, 1997.

[6] R. BHATIA, J. HOLBROOK, *Riemannian geometry and matrix geometric means*, Linear Algebra Appl., **413** (2006), 594–618.

[7] R. BHATIA, F. KITTANEH, *Notes on matrix arithmetic-geometric mean inequalities*, Linear Algebra Appl., **308** (2000), 203–211.

[8] R. COOPER, *Notes on certain inequalities (1)*, J. London Math. Soc., **2** (1927), 17–21.

[9] T. FURUTA, M. YANAGIDA, *Generalized means and convexity of inversion for positive operators*, Amer. Math. Monthly, **105** (1998), 258-259.

[10] O. HIRZALLAH, F. KITTANEH, *Matrix Young inequalities for the Hilbert-Schmidt norm*, Linear Algebra Appl., **308** (2000), 77–84.

[11] H. KOSAKI, *Arithmetic-geometric mean and related inequalities for operators*, J. Funct. Anal., **156**, 2 (1998), 429–451.

[12] R. MATHIAS, *An arithmetic-geometric-harmonic mean inequality involving Hadamard products*, Linear Algebra Appl., **184** (1993), 71–78.

[13] D. S. MITRINOVIĆ, *Analytic inequalities*, Die Grundlehren der mathematischen Wissenschaften, Band 1965, Springer-Verlag, New York-Berlin, 1970.

[14] A. OPPENHEIM, *Note on Mr. Cooper's Generalization of Young's Inequality*, J. London Math. Soc., **2** (1927), 21–23.
[15] M. M. RAO, Z. D. REN, *Theory of Orlicz spaces*, Monographs and Textbooks in Pure and Applied Mathematics, 146, Marcel Dekker, Inc., New York, 1991.
[16] D. RÜTHING, *On Young's inequality*, Internat. J. Math. Ed. Sci. Tech., **25**, 2 (1994), 161-164.
[17] X. ZHAN, *Matrix inequalities*, LNM 1790, Springer-Verlag, Berlin, 2002.

(Received May 1, 2007)

Tomaž Kosem
Institute of Mathematics, Physics and Mechanics
University of Ljubljana
Jadranska 19
SI-1000 Ljubljana
Slovenia
e-mail: tomaz.kosem@fmf.uni-lj.si