Non-commutative perspectives

Edward Effros and Frank Hansen

September 30, 2013

Abstract

We prove that the non-commutative perspective of an operator convex function is the unique extension of the corresponding commutative perspective that preserves homogeneity and convexity.

1 Introduction and preliminaries

Let f be a function defined in the positive (open) half-line. The perspective function $\mathcal{P}_f$ is the function of two variables given by

$$P_f(t,s) = sf(ts^{-1}) \qquad t,s > 0.$$

Depending on the application, we may also consider the function $(t,s) \to \mathcal{P}_f(s,t)$ and denote this as the perspective of f.

If A and B are commuting positive definite matrices then the matrix $\mathcal{P}_f(A,B)$ is well-defined by the functional calculus. Even if A and B do not commute we may by choosing an appropriate ordering define the perspective by setting

$$\mathcal{P}_f(A,B) = B^{1/2} f(B^{-1/2} A B^{-1/2}) B^{1/2}.$$

This expression is well-defined and coincides with $P_f(A,B)$, when A and B commute.

Effros [4] only considered the case where each pair in the argument of the perspective function consists of commuting operators and proved in this way that the perspective of an operator convex function is operator convex as a functions of two variables. Ebadian et. al. [3] noticed that virtually the same proof applies without any commutativity conditions. We include the proof for the convenience of the reader.

Theorem 1.1. *Let f be an operator convex function defined in the positive half-line. The mapping*

$$(A, B) \to \mathcal{P}_f(A, B),$$

defined in pairs of positive definite matrices, is convex.

Proof. Consider positive definite matrices A_1, A_2 and B_1, B_2 and take a real number $\lambda \in [0, 1]$. We set

$$A = \lambda A_1 + (1 - \lambda)A_2 \qquad \text{and} \qquad B = \lambda B_1 + (1 - \lambda)B_2.$$

The matrices

$$X = (\lambda B_1)^{1/2} B^{-1/2} \qquad \text{and} \qquad Y = ((1 - \lambda)B_2)^{1/2} B^{-1/2}$$

satisfy

$$X^* X + Y^* Y = B^{-1/2} \lambda B_1 B^{-1/2} + B^{-1/2}(1 - \lambda) B_2 B^{-1/2} = 1$$

and

$$X^* B_1^{-1/2} A_1 B_1^{-1/2} X + Y^* B_2^{-1/2} A_2 B_2^{-1/2} Y$$
$$= B^{-1/2} \lambda A_1 B^{-1/2} + B^{-1/2}(1 - \lambda) A_2 B^{-1/2} = B^{-1/2} A B^{-1/2}.$$

Hence

$$\mathcal{P}_f(\lambda A_1 + (1 - \lambda)A_2, \lambda B_1 + (1 - \lambda)B_2) = \mathcal{P}_f(A, B)$$

$$= B^{1/2} f(B^{-1/2} A B^{-1/2}) B^{1/2}$$

$$= B^{1/2} f\left(X^* B_1^{-1/2} A_1 B_1^{-1/2} X + Y^* B_2^{-1/2} A_2 B_2^{-1/2} Y\right) B^{1/2}$$

$$\leq B^{1/2} \left(X^* f(B_1^{-1/2} A_1 B_1^{-1/2}) X + Y^* f(B_2^{-1/2} A_2 B_2^{-1/2}) Y\right) B^{1/2}$$

$$= (\lambda B_1)^{1/2} f(B_1^{-1/2} A_1 B_1^{-1/2})(\lambda B_1)^{1/2}$$

$$\qquad\qquad + ((1 - \lambda)B_2)^{1/2} f(B_2^{-1/2} A_2 B_2^{-1/2})((1 - \lambda)B_2)^{1/2}$$

$$= \lambda \mathcal{P}_f(A_1, B_1) + (1 - \lambda)\mathcal{P}_f(A_2, B_2),$$

where we used Jensen's operator inequality [7]. **QED**

There are obvious similarities between the notion of a perspective function and the operator means studied by Kubo and Ando [8]. The crucial difference is that the representing function of an operator mean is operator monotone (and hence operator convex). We are considering operator convex functions, and they are in general not monotone.

Example 1.2. *Consider the operator convex function $f(t) = -\log t$ defined for $t > 0$. The (classical) perspective function is given by*

$$P_f(t, s) = -t\log(t^{-1}s) = t\log t - t\log s,$$

and the non-commutative perspective is then given by

$$\mathcal{P}_f(A, B) = -A^{1/2}\log(A^{-1/2}BA^{-1/2})A^{1/2}.$$

The relative entropy $S(A, B)$ is defined by setting

$$S(A, B) = \operatorname{Tr} A\log A - \operatorname{Tr} A\log B$$

and is known to be a convex function of two variables. For commuting matrices we have

$$S(A, B) = \operatorname{Tr}\mathcal{P}_f(A, B),$$

although the two quantities in general are different.

1.1 Spectral functions

Let $B(\mathcal{H})$ denote the set of bounded linear operators on a Hilbert space $\mathcal{H}$. A function $F\colon \mathcal{D} \to B(\mathcal{H})$ defined in a convex domain $\mathcal{D}$ of normal operators in $B(\mathcal{H})$ is called a spectral function, if it can be written on the form $F(x) = f(x)$ for some real or complex function f defined in a real interval I, where $f(x)$ is obtained by applying the functional calculus for normal operators.

Although this definition appears quite intuitive it contains some hidden assumptions. Firstly, the domain $\mathcal{D}$ should be invariant under unitary transformations and

$$(1) \qquad\qquad F(u^*xu) = u^*F(x)u \qquad x \in \mathcal{D}$$

for every unitary transformation u on $\mathcal{H}$. Secondly, for orthogonal projections p and q on $\mathcal{H}$, the element $pxp + qxq \in \mathcal{D}$ and

$$(2) \qquad\qquad F(pxp + qxq) = pF(pxp)p + qF(qxq)q$$

for arbitrary $x \in B(\mathcal{H})$ such that pxp and qxq are contained in $\mathcal{D}$. An operator function $x \to F(x)$ is a spectral function if and only if (1) and (2) are satisfied, cf. [2, 6].

2 The main result

The notion of spectral function is not immediately extendable to functions of two variables. However, we may consider the two properties of spectral functions noticed by C. Davis as a kind of regularity conditions, and they are readily extendable to functions of more than one variable.

Definition 2.1. *Let* $F\colon \mathcal{D} \to B(\mathcal{H})$ *be a function of two variables defined in a convex domain* $\mathcal{D} \subseteq B(\mathcal{H}) \times B(\mathcal{H})$. *We say that* F *is regular if*

(i) *The domain* $\mathcal{D}$ *is invariant under unitary transformations of* $\mathcal{H}$ *and*

$$F(u^*xu, u^*yu) = u^*F(x,y)u \qquad (x,y) \in \mathcal{D}$$

for every unitary u *on* $\mathcal{H}$.

(ii) *Let* p *and* q *be orthogonal projections on* $\mathcal{H}$. *Then the pair of diagonal block matrices* $(pxp + qxq, pyp + qyq) \in \mathcal{D}$ *and*

$$F(pxp + qxq, pyp + qyq) = pF(pxp, pyp)p + qF(qxq, qyq)q$$

for arbitrary $x, y \in B(\mathcal{H})$ *such that* (pxp, pyp) *and* (qxq, qyq) *are in* $\mathcal{D}$.

The following theorem is related to [5, Theorem 2.2].

Theorem 2.2. *Let* $(A, B) \to F(A, B)$ *be a regular map from pairs of bounded positive semi-definite operators on an infinite dimensional Hilbert space* $\mathcal{H}$ *into* $B(\mathcal{H})$ *satisfying the conditions:*

(i) $F(tA, tB) = tF(A, B) \qquad t > 0$

(ii) $F\left(\dfrac{A_1 + A_2}{2}, \dfrac{B_1 + B_2}{2}\right) \leq \dfrac{F(A_1, B_1) + F(A_2, B_2)}{2}$

(iii) $F(0,0) = 0$, *and* $B \to F(1, B)$ *is continuous on bounded subsets in the strong operator topology, where* 1 *denotes the unit operator on* $\mathcal{H}$.

Then there exists an operator convex function $f\colon \mathbf{R}_+ \to \mathbf{R}$ *such that*

$$F(1, t \cdot 1) = f(t)1 \qquad t > 0.$$

Furthermore,

$$F(A, B) = A^{1/2}f(A^{-1/2}BA^{-1/2})A^{1/2} = \mathcal{P}_f(A, B)$$

for positive definite invertible operators A *and* B.

Proof. The regularity of F entails that

$$u^* F(1, t \cdot 1) u = F(1, t \cdot 1) \qquad t > 0$$

for every unitary u in $B(\mathcal{H})$. Thus $F(1, t \cdot 1)$ commutes with every unitary in $B(\mathcal{H})$ and is therefore of the form

$$F(1, t \cdot 1) = f(t) \cdot 1 \qquad t > 0$$

for some function $f \colon \mathbf{R}_+ \to \mathbf{R}$. If $A = \sum_{i=1}^{n} \lambda_i P_i$ is the spectral decomposition of a finite rank positive definite operator A on $\mathcal{H}$ then

$$
\begin{aligned}
F(1, A) &= \sum_{i=1}^{n} P_i F(P_i, \lambda_i P_i) P_i \\
&= \sum_{i=1}^{n} P_i F(1, \lambda_i \cdot 1) P_i \\
&= \sum_{i=1}^{n} f(\lambda_i) P_i = f(A)
\end{aligned}
$$

(3)

by the regularity of F. Since F is mid-point convex it follows that f is mid-point operator convex and thus operator convex[1].

Let now C be a contraction and consider the unitary block matrices

$$U = \begin{pmatrix} C & (1 - CC^*)^{1/2} \\ (1 - C^*C)^{1/2} & -C^* \end{pmatrix} = \begin{pmatrix} C & D \\ E & -C^* \end{pmatrix}$$

and

$$V = \begin{pmatrix} C & -D \\ E & C^* \end{pmatrix}.$$

It is plain to calculate that

$$\frac{1}{2} U^* \begin{pmatrix} A & 0 \\ 0 & 0 \end{pmatrix} U + \frac{1}{2} V^* \begin{pmatrix} A & 0 \\ 0 & 0 \end{pmatrix} V = \begin{pmatrix} C^* A C & 0 \\ 0 & D A D \end{pmatrix}.$$

---

[1]It is a curiosity that continuity is not required to prove that mid-point operator convexity implies operator convexity.

We then obtain

$$\begin{pmatrix} C^*F(A,B)C & 0 \\ 0 & DF(A,B)D \end{pmatrix}$$

$$= \frac{1}{2}U^* \begin{pmatrix} F(A,B) & 0 \\ 0 & 0 \end{pmatrix} U + \frac{1}{2}V^* \begin{pmatrix} F(A,B) & 0 \\ 0 & 0 \end{pmatrix} V$$

$$= \frac{1}{2}U^*F \left(\begin{pmatrix} A & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} B & 0 \\ 0 & 0 \end{pmatrix} \right) U + \frac{1}{2}VF \left(\begin{pmatrix} A & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} B & 0 \\ 0 & 0 \end{pmatrix} \right) V^*$$

$$= \frac{1}{2}F \left(U^* \begin{pmatrix} A & 0 \\ 0 & 0 \end{pmatrix} U, U^* \begin{pmatrix} B & 0 \\ 0 & 0 \end{pmatrix} U \right) + \frac{1}{2}F \left(V^* \begin{pmatrix} A & 0 \\ 0 & 0 \end{pmatrix} V, V^* \begin{pmatrix} B & 0 \\ 0 & 0 \end{pmatrix} V \right)$$

$$\geq F \left(\frac{1}{2}U^* \begin{pmatrix} A & 0 \\ 0 & 0 \end{pmatrix} U + \frac{1}{2}V^* \begin{pmatrix} A & 0 \\ 0 & 0 \end{pmatrix} V, \frac{1}{2}U^* \begin{pmatrix} B & 0 \\ 0 & 0 \end{pmatrix} U + \frac{1}{2}V^* \begin{pmatrix} B & 0 \\ 0 & 0 \end{pmatrix} V \right)$$

$$= F \left(\begin{pmatrix} C^*AC & 0 \\ 0 & DAD \end{pmatrix}, \begin{pmatrix} C^*BC & 0 \\ 0 & DBD \end{pmatrix} \right)$$

$$= \begin{pmatrix} F(C^*AC, C^*BC) & 0 \\ 0 & F(DAD, DBD) \end{pmatrix},$$

where, in the second equality, we used $F(0,0) = 0$ from condition (iii). In particular, we have proved that

$$(4) \qquad\qquad C^*F(A,B)C \geq F(C^*AC, C^*BC)$$

for contractions C. However, the homogeneity of F then implies (4) for any operator C. In particular, if C is invertible we obtain

$$F(A,B) \geq (C^*)^{-1}F(C^*AC, C^*BC)C^{-1} \geq F(A,B),$$

hence there is equality and thus

$$C^*F(A,B)C = F(C^*AC, C^*BC).$$

For invertible A we therefore obtain

$$A^{-1/2}F(A,B)A^{-1/2} = F(1, A^{-1/2}BA^{-1/2}).$$

If B is positive definite and of finite rank, then so is $A^{-1/2}BA^{-1/2}$ and thus

$$F(1, A^{-1/2}BA^{-1/2}) = f(A^{-1/2}BA^{-1/2})$$

6

by equation (3).

Let g be a continuous function defined in an open interval I. The functional calculus $X \rightarrow g(X)$ is strongly continuous on bounded subsets of self-adjoint operators X with spectra in I, cf. the proof of [1, Lemma 2.2]. Indeed, if (X_i) is a bounded net of operators converging strongly to X, then the inequality

$$\|X^k \xi - X_i^k \xi\| \leq \|X^k \xi - X_i^{k-1} X \xi\| + \|X_i^{k-1} X \xi - X_i^k \xi\|$$
$$\leq \|X^k \xi - X_i^{k-1} X \xi\| + \|X_i^{k-1}\| \cdot \|X \xi - X_i \xi\|,$$

together with an induction argument, shows that (X_i^k) converges strongly to X^k for any natural number k. The assertion then follows by approximating g uniformly by polynomials in a compact subset of I containing the spectrum of X. The continuity condition in (iii) therefore implies

$$F(1, A^{-1/2} B A^{-1/2}) = f(A^{-1/2} B A^{-1/2})$$

and thus

$$F(A, B) = A^{1/2} f(A^{-1/2} B A^{-1/2}) A^{1/2}$$

for arbitrary positive definite operators A and B defined on $\mathcal{H}$. **QED**

Remark 2.3. *It is crucial in the above proof that the regular map $F(A, B)$ is defined for positive semi-definite operators. We are therefore excluding the limiting case,*

$$F(A, B) = A^{1/2} f(A^{-1/2} B A^{-1/2}) A^{1/2} = A B^{-1} A,$$

that appears by setting $f(t) = t^{-1}$ for $t > 0$.

Notice that the above theorem has an obvious counterpart if convexity is replaced by concavity. The theorem states that a non-commutative perspective function, that allows an extension to positive semi-definite operators, is the unique extension of a commutative perspective function to a homogeneous, convex and regular operator mapping. In particular, the geometric operator mean

$$A \# B = A^{1/2} \left(A^{-1/2} B A^{-1/2} \right)^{1/2} A^{1/2}$$

is the only sensible extension of the geometric mean $(t, s) \rightarrow \sqrt{ts}$ of positive numbers to a homogeneous and concave operator mapping.

References

[1] J. Bendat and S. Sherman. Monotone and convex operator functions. *Trans. Amer. Math. Soc.*, 79:58–71, 1955.

[2] C. Davis. A Schwarz inequality for convex operator functions. *Proc. Amer. Math. Soc.*, 8:42–44, 1957.

[3] A. Ebadian, I. Nikoufar, and M.E. Gordji. Perspectives of matrix convex functions. *Proc. Natl. Acad. Sci. USA*, 108(18):7313–7314, 2011.

[4] E.G. Effros. A matrix convexity approach to some celebrated quantum inequalities. *Proc. Natl. Acad. Sci. USA*, 106:1006–1008, 2009.

[5] F. Hansen. Means and concave products of positive semi-definite matrices. *Math. Ann.*, 264:119–128, 1983.

[6] F. Hansen and Pedersen G.K. Jensen's operator inequality. *Bull. London Math. Soc.*, 35:553–564, 2003.

[7] F. Hansen and G.K. Pedersen. Jensen's inequality for operators and Löwner's theorem. *Math. Ann.*, 258:229–241, 1982.

[8] F. Kubo and T. Ando. Means of positive linear operators. *Math. Ann.*, 246:205–224, 1980.

Edward Effros: Mathematics Department, UCLA, Los Angeles, CA 90015.
Email: ege@math.ucla.edu

Frank Hansen: Institute for International Education, Tohoku University, Japan.
Email: frank.hansen@m.tohoku.ac.jp.