Received 19/10/12

ORDER GENERALISED GRADIENT AND OPERATOR INEQUALITIES

SEVER S. DRAGOMIR[1,2] AND EDER KIKIANTY[2,*]

ABSTRACT. We introduce the notion of order generalised gradient for operator-valued functions and state some operator inequalities of Hermite-Hadamard and Jensen types. We discuss the connection between the order generalised gradient and Gâteaux derivative of operator-valued functions. Evidently, we state a characterisation of operator convexity via an inequality concerning the order generalised gradient.

Key words and Phrases: subgradient inequality, operator convex function, operator inequality
MSC 2010: 47A63, 46E40

1. Introduction

Convex functions play a crucial role in many fields of mathematics, most prominently in optimisation theory. There are two main important inequalities which characterise convex functions, namely Jensen's and Hermite-Hadamard's inequalities. In 1905 (1906), Jensen defined convex functions as follows: $f : I \subset \mathbb{R} \to \mathbb{R}$ is a convex function if and only if

$$f\left(\frac{a+b}{2}\right) \le \frac{f(a)+f(b)}{2}, \quad \text{for any } a, b \in I.$$

This inequality is then referred to as Jensen's inequality. Hermite-Hadamard's inequality provides a refinement for Jensen's inequality, namely, for a convex function $f : I \subset \mathbb{R} \to \mathbb{R}$,

$$f\left(\frac{a+b}{2}\right) \le \frac{1}{b-a}\int_a^b f(x)dx \le \frac{f(a)+f(b)}{2}, \quad \text{for any } a, b \in I.$$

We refer the reader to Section 2 for further details regarding these inequalities.

Analogously to the case of real-valued functions, the operator convexity can be characterised by some operator inequalities. Hansen and Pedersen [20] characterise operator convexity via a non-commutative generalisation of Jensen's inequality. If f is a real continuous function on an interval I, and $\mathcal{A}(H)$ is the set of bounded self-adjoint operators on a Hilbert space H with spectra in I, then f is operator convex if and only if

$$f\left(\sum_{i=1}^{n} a_i^* x_i a_i\right) \le \sum_{i=1}^{n} a_i^* f(x_i) a_i$$

for $x_1, \dots x_n \in \mathcal{A}(H)$ and $a_1, \dots a_n \in \mathcal{B}(H)$ with $\sum_{i=1}^{n} a_i^* a_i = \mathbf{1}$. We refer the reader to Section 2 for further details regarding this characterisation.

One of the useful differential properties of convex functions is the fact that their one-sided directional derivatives exist universally [28, p. 213]. Just as the ordinary two-sided directional derivatives of a differentiable function can be described in

[1] School of Engineering and Science, Victoria University, PO Box 14428, Melbourne 8001, Victoria, Australia
[2] School of Computational and Applied Mathematics,University of the Witwatersrand, Private Bag 3 Wits 2050, South Africa
[*] Corresponding Author, E-mail Address: eder.kikianty@wits.ac.za, Phone: +27 11 7176118, Fax:+27 867653471

terms of gradient vectors, the one-sided directional derivatives can be described in terms of "subgradient" vectors [28, p. 213]. A vector x^* is said to be a subgradient of a convex function $f : K \subset \mathbb{R}^n \to \mathbb{R}$ at point x if

$$f(x) - f(y) \geq x^* \cdot (x - y), \quad \text{for all} \;\; y \in K. \tag{1}$$

This condition is referred to as the ***subgradient inequality*** [28, p. 214]. If (1) holds for every $x \in K$ then (1) characterises the convexity of f (see for example, Eisenberg [15, Theorem 1]). Vandergraft [30, Lemma 4.1] mentioned this characterisation of convexity in the settings of partially ordered topological vector space which was used to state a convergence theorem of Newton's method for convex operators. We refer the reader to Section 2 for further details concerning these subgradient inequalities.

In this paper, we introduce the notion of ***order generalised gradient*** (cf. Section 3) for operator-valued functions, which is a generalisation of (1) (without the assumptions of convexity) in the settings of bounded self-adjoint operators on a Hilbert space. Furthermore, we state some inequalities of Hermite-Hadamard and Jensen types for the order generalised gradient in Sections 4 and 5. Finally, in Section 6, we state the connection between the order generalised gradient and Gâteaux derivative of operator-valued functions. Evidently, we state a characterisation of convexity analogues to (1) in the context of operator-valued functions.

2. Inequalities for convex functions

This section serves as a point of reference for known results regarding some inequalities related to convex functions; both real-valued and operator-valued functions.

2.1. Jensen's inequality.

Jensen's inequality for convex functions plays a crucial role in the theory of inequalities due to the fact that other inequalities such as the Arithmetic-Geometric mean, Hölder, Minkowski, and Ky Fan's inequalities, can be obtained as particular cases of it.

Let C be a convex subset of the linear space X and f be a convex function on C. If $\mathbf{p} = (p_1, \ldots, p_n)$ is a probability sequence and $\mathbf{x} = (x_1, \ldots, x_n) \in C^n$, then

$$f\left(\sum_{i=1}^n p_i x_i\right) \leq \sum_{i=1}^n p_i f(x_i). \tag{2}$$

This inequality is referred to as ***Jensen's inequality***. Recently, Dragomir [14] obtained the following refinement of Jensen's inequality:

$$\begin{aligned} f\left(\sum_{j=1}^n p_j x_j\right) &\leq \min_{k\in\{1,\ldots,n\}} \left[(1-p_k) f\left(\frac{\sum_{j=1}^n p_j x_j - p_k x_k}{1-p_k}\right) + p_k f(x_k)\right] \\ &\leq \frac{1}{n}\left[\sum_{k=1}^n (1-p_k) f\left(\frac{\sum_{j=1}^n p_j x_j - p_k x_k}{1-p_k}\right) + \sum_{k=1}^n p_k f(x_k)\right] \\ &\leq \max_{k\in\{1,\ldots,n\}} \left[(1-p_k) f\left(\frac{\sum_{j=1}^n p_j x_j - p_k x_k}{1-p_k}\right) + p_k f(x_k)\right] \\ &\leq \sum_{j=1}^n p_j f(x_j), \end{aligned} \tag{3}$$

where f, x_k and p_k are as defined above. For other refinements of Jensen's inequality, we refer the reader to Pečarić and Dragomir [24] and Dragomir [8].

The above result provides a different approach to the one that Pečarić and Dragomir [24] obtained in 1989

$$
\begin{aligned}
(4)\qquad f\left(\sum_{i=1}^{n} p_i x_i\right) &\le \sum_{i_1,\ldots,i_{k+1}=1}^{n} p_{i_1}\ldots p_{i_{k+1}} f\left(\frac{x_{i_1}+\cdots+x_{i_{k+1}}}{k+1}\right)\\
&\le \sum_{i_1,\ldots,i_k=1}^{n} p_{i_1}\ldots p_{i_k} f\left(\frac{x_{i_1}+\cdots+x_{i_k}}{k}\right)\\
&\le \cdots \le \sum_{i=1}^{n} p_i f(x_i),
\end{aligned}
$$

for $k \ge 1$ and $\mathbf{p}$, $\mathbf{x}$ are as defined above.

If $q_1,\ldots q_k \ge 0$ with $\sum_{j=1}^{k} q_j = 1$, then the following refinement obtained in 1994 by Dragomir [8] also holds:

$$
\begin{aligned}
(5)\qquad f\left(\sum_{i=1}^{n} p_i x_i\right) &\le \sum_{i_1,\ldots,i_k=1}^{n} p_{i_1}\ldots p_{i_k} f\left(\frac{x_{i_1}+\cdots+x_{i_k}}{k}\right)\\
&\le \sum_{i_1,\ldots,i_k=1}^{n} p_{i_1}\ldots p_{i_k} f(q_1 x_{i_1}+\cdots+q_k x_{i_{k+1}})\\
&\le \sum_{i=1}^{n} p_i f(x_i),
\end{aligned}
$$

where $1 \le k \le n$ and $\mathbf{p}$, $\mathbf{x}$ are as defined above.

For more refinements and applications related to Ky Fan's inequality, the Arithmetic-Geometric mean inequality, the generalised triangle inequality, the f-divergence measures etc., we refer the readers to [5], [6], [7], [8], [9], [12], and [14].

2.2. **Hermite-Hadamard's inequality.** The following inequality also holds for any convex function f defined on $\mathbb{R}$

$$
(6)\qquad (b-a)f\left(\frac{a+b}{2}\right) \le \int_a^b f(x)dx \le (b-a)\frac{f(a)+f(b)}{2}, \quad a,b\in\mathbb{R}.
$$

It was firstly discovered by Ch. Hermite in 1881 in the journal *Mathesis* [21]. However, this result was nowhere mentioned in the mathematical literature and was not widely known as Hermite's result [26].

E.F. Beckenbach, a leading expert on the history and the theory of convex functions wrote that this inequality was proven by J. Hadamard in 1893 [3]. In 1974, D.S. Mitrinović found Hermite's note in *Mathesis* [21]. Since (6) was known as Hadamard's inequality, the inequality is now commonly referred to as ***Hermite-Hadamard's inequality*** [26].

Hermite-Hadamard inequality's has been extended in many different directions. One of the extensions of this inequality is in the vector space settings. Firstly, we start with the following definitions and notation: let X be a vector space and x, y be two distinct vectors in X. We define the segment generated by x and y to be the set:

$$[x,y] := \{(1-t)x + ty, \quad t\in[0,1]\}.$$

For any function real-valued function f defined on the segment $[x,y]$ there exists an associated function $g_{x,y} : [0,1] \to \mathbb{R}$, with

$$g_{x,y}(t) = f[(1-t)x + ty].$$

We remark that f is convex on $[x,y]$ if and only if g is convex on $[0,1]$. For any convex function defined on a segment $[x,y] \subset X$, we have the *Hermite-Hadamard* *integral inequality* (cf. Dragomir [10, p. 2] and Dragomir [11, p. 2]):

$$f\left(\frac{x+y}{2}\right) \le \int_0^1 f[(1-t)x+ty]dt \le \frac{f(x)+f(y)}{2}, \; x,y \in X \tag{7}$$

which can be derived by the classical Hermite-Hadamard inequality (6) for the convex function $g_{x,y} : [0,1] \to \mathbb{R}$. Consider the function $f(x) = \|x\|^p$ ($x \in X$ and $1 \le p < \infty$), which is convex on X, then we have the following norm inequality (derived from (7)) [25, p. 106]

$$\left\|\frac{x+y}{2}\right\|^p \le \int_0^1 \|(1-t)x+ty\|^p dt \le \frac{\|x\|^p + \|y\|^p}{2} \tag{8}$$

for any $x, y \in X$.

2.3. **Non-commutative generalisation of Jensen's inequality.** Hansen [16] discussed Jensen's operator inequality for operator monotone functions. Motivated by Aujla's work [1] on the matrix convexity of functions of two variables, Hansen [17] characterised operator convex functions of two variables in terms of a non-commutative generalisation of Jensen's inequality (cf. [17, Theorem 3.1]). A simplified proof of this result formulated for matrices is given in Aujla [2]. The case for several variable is given in Hansen [18]. The case for self-adjoint elements in the algebra M_n of n-square matrices is given in Hansen and Pedersen [19]. Finally, Hansen and Pedersen [20] presented a generalisation of the above results for self-adjoint operators defined on a Hilbert space.

Theorem 1. *If f is a real continuous function on an interval I, and $\mathcal{A}(H)$ is the set of bounded self-adjoint operators on a Hilbert space H with spectra in I, the following conditions are equivalent:*

(i) *f is operator convex;*
(ii) $f(\sum_{i=1}^n a_i^* x_i a_i) \le \sum_{i=1}^n a_i^* f(x_i) a_i$ *for* $x_1, \ldots x_n \in \mathcal{A}(H)$ *and* $a_1, \ldots a_n \in \mathcal{B}(H)$ *with* $\sum_{i=1}^n a_i^* a_i = \mathbf{1}$;
(iii) $f(v^*xv) \le v^* f(x) v$ *for any* $x \in \mathcal{A}(H)$ *and any isometry* $v \in \mathcal{B}(H)$;
(iv) $pf(pxp + s(\mathbf{1}-p))p \le pf(x)p$ *for* $x \in \mathcal{A}(H)$ *, projection* $p \in \mathcal{B}(H)$*, every self-adjoint operator* x *with spectrum in* I *and* $s \in I$*.*

2.4. **Subgradient inequality.** Recall the following definition of subgradient:

Definition 2 (Rockafellar [28])**.** A vector x^* is said to be a subgradient of a convex function $f : K \subset \mathbb{R}^n \to \mathbb{R}$ at point x if

$$f(x) - f(y) \ge x^* \cdot (x-y), \quad \text{for all } \; y \in K.$$

The following theorem is a useful characterisation of convexity is (cf. Eisenberg [15, Theorem 1]):

Theorem 3 (Eisenberg [15])**.** *If U is a nonempty open subset of $\mathbb{R}^n$, $f : U \to \mathbb{R}$ is a differentiable function on U, and K is a convex subset of U, then f is convex on K if and only if*

$$f(x) - f(y) \ge (x-y)^T f'(y), \quad \text{for all } \; x,y \in K \tag{9}$$

where $f'(y)$ denotes the gradient of f at y.

This theorem has been generalised and employed in obtaining optimality conditions of a non-differentiable minimax programming problem in complex spaces (cf. Lai and Liu [22]). Note that $(x-y)^T f'(y)$ can be written as $f'(y) \cdot (x-y)$, which can be interpreted as the directional derivative of f at point y in $x-y$ direction.

Vandergraft [30, Lemma 4.1] mentioned this characterisation of convexity in the settings of partially ordered topological vector space which was used to state a convergence theorem of Newton's method for convex operators. Following is the characterisation.

Theorem 4 (Vandergraft [30]). *Let Z, W be partially ordered topological vector spaces. If $F : D \subset Z \to W$ is Gâteaux differentiable in the convex domain D, then F is convex if and only if*

$$F'(z)h \leq F(z+h) - f(z) \tag{10}$$

for all z, h such that $z \in D^0$, $z+h \in D$. Here $F'(z)h$ denotes the Gâteaux derivative of F at z on h direction; and D^0 denotes the interior of the set D.

Damm and Hinrichsen [4] defined a concave mapping on an ordered Banach space, to prove a non-local convergence result for Newton's method applied to a class of concave operator equations with resolvent positive derivatives. Following is the definition.

Definition 5. Let X be a real Banach space ordered by a pointed convex closed cone C; and consider a nonlinear mapping $f : \text{dom } f \subset X \to X$. Let $D_+ \subset D \subset \text{dom } f$. We say that f is D_+-concave on D if we can attach a bounded linear operator $T_x : X \to X$ at each point $x \in D$, such that for all $x \in D$

$$f(y) - f(x) \leq T_x(y-x), \quad \text{for all } y \in D_+. \tag{11}$$

It is shown that equality holds in (11) when f is Gâteaux differentiable at y, and furthermore, T_x is f'_y, i.e. the Gâteaux derivative of f at y.

3. Order generalised gradient

Throughout the paper, we use the following notation. Let $\mathcal{B}(H)$ be the Banach algebra of all bounded linear operators on the Hilbert space $(H, \langle \cdot, \cdot \rangle)$, $\mathcal{A}(H)$ be the linear subspace of all self-adjoint operators on H.

We denote $\mathcal{P}_+(H) \subset \mathcal{A}(H)$ the convex cone of all positive definite operators defined on H, that is $P \in \mathcal{P}_+(H)$ if and only if $\langle Px, x \rangle \geq 0$, and for all $x \in H$, $\langle Px, x \rangle = 0$ implies $x = 0$. This gives a partial ordering (we refer to it as operator order) on $\mathcal{A}(H)$, where two elements $A, B \in \mathcal{A}(H)$ satisfy $A \leq B$ if and only if $B - A \in \mathcal{P}_+(H)$.

Definition 6. Let $\mathcal{C}$ be a convex set in $\mathcal{A}(H)$. A function $f : \mathcal{C} \to \mathcal{A}(H)$ has the function $\nabla_f : \mathcal{C} \times \mathcal{A}(H) \to \mathcal{A}(H)$ as an ***order generalised gradient*** if

$$f(A) - f(B) \geq \nabla_f(B, A-B), \quad \text{for any } A, B \in \mathcal{C} \tag{12}$$

in the operator order of $\mathcal{A}(H)$.

Remark 7. We remark that in (12), if f is a real-valued differentiable function on an open set $U \subset \mathbb{R}^n$, and ∇_f is the gradient of f, then (12) becomes (9). We also note that there is no assumption of convexity at this point. We discuss the convexity case in Section 6.

Proposition 8. *If $Q \in \mathcal{A}(H)$ and $f : \mathcal{A}(H) \to \mathcal{A}(H)$, $f(A) = QA^2Q$ then*

$$\nabla_f(B, X) := Q(BX + XB)Q \tag{13}$$

is an order generalised gradient for f.

Proof. Observe that $BX + XB \in \mathcal{A}(H)$ and if $P \in \mathcal{A}(H)$ then $P(BX+XB)P \in \mathcal{A}(H)$. We need to prove that

$$f(A) - f(B) \geq \nabla_f(B, A-B)$$

for any $A, B \in \mathcal{A}(H)$, that is,

$$QA^2Q - QB^2Q \geq Q[B(A-B) + (A-B)B]Q. \tag{14}$$

Since

$$Q[B(A-B)+(A-B)B]Q = QBAQ - QB^2Q + QABQ - QB^2Q$$

thus (14) is equivalent with

$$QA^2Q - QB^2Q \geq QBAQ - QB^2Q + QABQ - QB^2Q$$

which is also equivalent to

$$Q(A-B)^2Q \geq 0$$

which always holds. This completes the proof. □

We denote by $\mathcal{P}(H) \subset \mathcal{A}(H)$ the convex cone of all nonnegative operators defined on H.

Proposition 9. *If $P \in \mathcal{P}(H)$, then the function $f : \mathcal{A}(H) \to \mathcal{A}(H)$, $f(A) = APA$ has*

$$\nabla_f(B,X) := XPB + BPX \tag{15}$$

as an order generalised gradient.

Proof. Observe that $XPB + BPX \in \mathcal{A}(H)$. We need to prove that

$$\begin{aligned} APA - BPB &\geq (A-B)PB + BP(A-B) \\ &= APB - BPB + BPA - BPB \end{aligned}$$

that is,

$$APA - APB - BPA + BPB \geq 0.$$

But $APA-APB-BPA+BPB = (A-B)P(A-B)$ and since $(A-B)P(A-B) \geq 0$; and this completes the proof. □

Recall $\mathcal{P}_+(H) \subset \mathcal{A}(H)$ the convex cone of all positive definite operators defined on H, that is $P \in \mathcal{P}_+(H)$ if and only if $\langle Px, x\rangle \geq 0$, and for all $x \in H$, $\langle Px, x\rangle = 0$ implies $x = 0$.

Proposition 10. *Let $f : \mathcal{P}_+(H) \to \mathcal{A}(H)$ defined by*

$$f(A) = QA^{-1}Q$$

where $Q \in \mathcal{A}(H)$. The function $\nabla_f : \mathcal{P}_+(H) \times \mathcal{P}_+(H) \to \mathcal{A}(H)$ with

$$\nabla_f(B,X) = -QB^{-1}XB^{-1}Q$$

is an order generalised gradient for f.

Proof. For $B \in \mathcal{P}_+(H)$, $B^{-1} \in \mathcal{P}_+(H)$ then $B^{-1}XB^{-1} \in \mathcal{P}_+(H)$ for any $X \in \mathcal{P}_+(H)$ and thus $QB^{-1}XB^{-1}Q \in \mathcal{P}_+(H)$ showing that $\nabla_f(B,X) \in \mathcal{A}(H)$. We need to prove that

$$QA^{-1}Q - QB^{-1}Q \geq -QB^{-1}(A-B)B^{-1}Q$$

that is

$$QA^{-1}(B-A)B^{-1}Q + QB^{-1}(A-B)B^{-1}Q \geq 0$$

or equivalently

$$QA^{-1}(B-A)B^{-1}Q - QB^{-1}(B-A)B^{-1}Q \geq 0$$

or

$$Q(A^{-1}-B^{-1})(B-A)B^{-1}Q \geq 0.$$

But

$$\begin{aligned} Q(A^{-1}-B^{-1})(B-A)B^{-1}Q &= Q(A^{-1}-B^{-1})AA^{-1}(B-A)B^{-1}Q \\ &= Q(A^{-1}-B^{-1})A(A^{-1}-B^{-1})Q \geq 0 \end{aligned}$$

which is true since for $A \in \mathcal{P}_+(H)$ we have that

$$(A^{-1} - B^{-1})A(A^{-1} - B^{-1}) \geq 0$$

and $Q \in \mathcal{A}(H)$. □

4. Inequalities for the order generalised gradient

We start this section by the following definition.

Definition 11. The order generalised gradient $\nabla_f : \mathcal{C} \times \mathcal{A}(H) \to \mathcal{A}(H)$ is

- **operator convex** if
$$\nabla_f(B, \alpha X + \beta Y) \leq \alpha \nabla_f(B, X) + \beta \nabla_f(B, Y)$$
for any $B \in \mathcal{C}$, $X, Y \in \mathcal{A}(H)$ and $\alpha, \beta \geq 0$ with $\alpha + \beta = 1$.
- **operator sub-additive** if
$$\nabla_f(B, X + Y) \leq \nabla_f(B, X) + \nabla_f(B, Y)$$
for any $B \in \mathcal{C}$ and $X, Y \in \mathcal{A}(H)$.
- **positive homogeneous** if
$$\nabla_f(B, \alpha X) = \alpha \nabla_f(B, X)$$
for any $B \in \mathcal{C}$, $X \in \mathcal{A}(H)$ and $\alpha \geq 0$.
- **operator linear** if
$$\nabla_f(B, \alpha X + \beta Y) = \alpha \nabla_f(B, X) + \beta \nabla_f(B, Y)$$
for any $B \in \mathcal{C}$, $X, Y \in \mathcal{A}(H)$ and $\alpha, \beta \in \mathbb{R}$.

It can be seen that if $\nabla_f(\cdot, \cdot)$ is operator linear then it is positive homogeneous and sub-additive. If $\nabla_f(\cdot, \cdot)$ is positive homogeneous and sub-additive, then it is operator convex.

Theorem 12. *Let $f : \mathcal{C} \to \mathcal{A}(H)$ be operator convex and $\nabla_f : \mathcal{C} \times \mathcal{A}(H) \to \mathcal{A}(H)$ be an order generalised gradient for f. Then for any $A, B, \in \mathcal{C}$ and $t \in [0, 1]$ we have the inequalities*

$$\begin{aligned} &-(1-t)\nabla_f(B, -t(B-A)) - t\nabla_f(A, (1-t)(B-A)) \\ &\geq tf(A) + (1-t)f(B) - f(tA + (1-t)B) \\ &\geq \nabla_f(tA + (1-t)B, 0). \end{aligned}$$

Proof. If we write the definition of ∇_f for B instead of A we get

$$f(B) - f(A) \geq \nabla_f(A, B - A)$$

which is equivalent with

$$-\nabla_f(A, B - A) \geq f(A) - f(B).$$

Therefore for any $A, B \in \mathcal{C}$ we have the gradient inequalities

$$-\nabla_f(A, B - A) \geq f(A) - f(B) \geq \nabla_f(B, A - B). \tag{16}$$

Since $\mathcal{C}$ is a convex set, then by (16) we have

$$\begin{aligned} -\nabla_f(A, (1-t)(B-A)) &\geq f(A) - f(tA + (1-t)B) \\ &\geq \nabla_f(tA + (1-t)B, -(1-t)(B-A)) \end{aligned} \tag{17}$$

and

$$\begin{aligned} -\nabla_f(B, -t(B-A)) &\geq f(B) - f(tA + (1-t)B) \\ &\geq \nabla_f(tA + (1-t)B, t(B-A)) \end{aligned} \tag{18}$$

for any $t \in (0, 1)$.

If we multiply (17) with t and (18) with $(1-t)$ and add the obtained inequalities, then we get

$$\begin{aligned}&-t\nabla_f(A,(1-t)(B-A))-(1-t)\nabla_f(B,-t(B-A))\\&\geq tf(A)+(1-t)f(B)-f(tA+(1-t)B)\\&\geq t\nabla_f(tA+(1-t)B,-(1-t)(B-A))+(1-t)\nabla_f(tA+(1-t)B,t(B-A)).\end{aligned}$$

Since $\nabla_f(\cdot,\cdot)$ is operator convex, we also know that

$$\begin{aligned}&t\nabla_f(tA+(1-t)B,-(1-t)(B-A))+(1-t)\nabla_f(tA+(1-t)B,t(B-A))\\&\geq\nabla_f(tA+(1-t)B),-t(1-t)(B-A)+(1-t)t(B-A))\\&\geq\nabla_f(tA+(1-t)B,0)\end{aligned}$$

which completes the proof. □

Corollary 13. *Under the assumptions of Theorem 12*

(1) If $\nabla_f(\cdot,\cdot)$ is positive homogeneous, then we have

$$\begin{aligned}&-t(1-t)[\nabla_f(B,A-B)+\nabla_f(A,B-A)]\\&\geq tf(A)+(1-t)f(B)-f(tA+(1-t)B)\geq 0.\end{aligned}\tag{19}$$

(2) If $\nabla_f(\cdot,\cdot)$ is operator linear then

$$\begin{aligned}&t(1-t)[\nabla_f(B,B-A)-\nabla_f(A,B-A)]\\&\geq tf(A)+(1-t)f(B)-f(tA+(1-t)B)\geq 0.\end{aligned}\tag{20}$$

Corollary 14 (Hermite-Hadamard type inequality)**.** *Under the assumptions of Theorem 12, if ∇_f is positive homogeneous then we have the following inequality*

$$\begin{aligned}&-\frac{1}{6}[\nabla_f(B,A-B)+\nabla_f(A,B-A)]\\&\geq\frac{f(A)+f(B)}{2}-\int_0^1 f(tA+(1-t)B)\,dt\geq 0.\end{aligned}\tag{21}$$

We obtain (21) by integrating (19) over $t\in[0,1]$.

Example 15. 1. If we consider the function $f(A)=QA^2Q$ with $Q\in\mathcal{A}(H)$ then

$$\begin{aligned}\nabla_f(B,X)-\nabla_f(A,X)&=Q(BX+XB)Q-Q(AX+XA)Q\\&=Q[(B-A)X+X(B-A)]Q.\end{aligned}$$

For $X=B-A$ we then get

$$\nabla_f(B,B-A)-\nabla_f(A,B-A)=2Q(B-A)^2Q.$$

Applying inequality (20) we have

$$2t(1-t)Q(B-A)^2Q\geq Q[tA^2+(1-t)B^2-(tA+(1-t)B)^2]Q\geq 0\tag{22}$$

for any $A,B\in\mathcal{A}(H)$ and $Q\in\mathcal{A}(H)$.

2. If we consider the function $f(A)=APA$ with $P\in\mathcal{P}(H)$, thus

$$\begin{aligned}\nabla_f(B,X)-\nabla_f(A,X)&=XPB+BPX-XPA-APX\\&=XP(B-A)+(B-A)PX.\end{aligned}$$

If $X=B-A$ we then get

$$\nabla_f(B,B-A)-\nabla_f(A,B-A)=2(B-A)P(B-A).$$

Applying the inequality (20) we have

$$\begin{aligned}&2t(1-t)(B-A)P(B-A)\\&\geq tAPA+(1-t)BPB-(tA+(1-t)B)P(tA+(1-t)B)\geq 0\end{aligned}$$

for any $A,B\in\mathcal{A}(H)$ and $P\in\mathcal{P}(H)$.

3. For $f(A) = QA^{-1}Q$ with $Q \in \mathcal{A}(H)$ and $A \in \mathcal{P}_+(H)$ then

$$\nabla_f(B, X) - \nabla_f(A, X) = -QB^{-1}XB^{-1}Q + QA^{-1}XA^{-1}Q.$$

For $X = B - A$ we get

$$\begin{aligned}
&\nabla_f(B, B - A) - \nabla_f(A, B - A) \\
&= -QB^{-1}(B - A)B^{-1}Q + QA^{-1}(B - A)A^{-1}Q \\
&= -Q(B^{-1} - B^{-1}AB^{-1})Q + Q(A^{-1}BA^{-1} - A^{-1})Q \\
&= QA^{-1}BA^{-1}Q + QB^{-1}AB^{-1}Q - QB^{-1}Q - QA^{-1}Q.
\end{aligned}$$

By (20) we then have the inequality

$$\begin{aligned}
&t(1 - t)[QA^{-1}BA^{-1}Q + QB^{-1}AB^{-1}Q - QB^{-1}Q - QA^{-1}Q] \\
&\geq tQA^{-1}Q + (1 - t)QB^{-1}Q - Q(tA + (1 - t)B)^{-1}Q \geq 0
\end{aligned}$$

for any $A, B \in \mathcal{P}_+(H)$ and $Q \in \mathcal{A}(H)$.

5. Jensen type inequality for the order generalised gradient

In this section, we will state inequalities of Jensen type for the order generalised gradient.

Theorem 16. *Let $f : \mathcal{C} \subset \mathcal{A}(H) \to \mathcal{A}(H)$ be a function that possesses $\nabla_f : \mathcal{C} \times \mathcal{A}(H) \to \mathcal{A}(H)$ as an order generalised gradient. Then for any $A_i \in \mathcal{C}$, $i \in \{1, \ldots, n\}$ and $p_i \geq 0$ with $P_n := \sum_{i=1}^n p_i > 0$ we have the inequalities*

$$\begin{aligned}
&-\frac{1}{P_n}\sum_{j=1}^n p_j \nabla_f\left(A_j, \frac{1}{P_n}\sum_{i=1}^n p_i A_i - A_j\right) \\
&\geq \frac{1}{P_n}\sum_{j=1}^n p_j f(A_j) - f\left(\frac{1}{P_n}\sum_{i=1}^n p_i A_i\right) \\
&\geq \frac{1}{P_n}\sum_{j=1}^n p_j \nabla_f\left(\frac{1}{P_n}\sum_{i=1}^n p_i A_i, A_j - \frac{1}{P_n}\sum_{i=1}^n p_i A_i\right).
\end{aligned} \tag{23}$$

Proof. From the definition of order generalised gradient we have

$$-\nabla_f(A, B - A) \geq f(A) - f(B) \geq \nabla_f(B, A - B). \tag{24}$$

Now, if we choose $A = A_j$, $j \in \{1, \ldots, n\}$ and $B = (1/P_n)\sum_{i=1}^n p_i A_i$ in (24), then we get

$$\begin{aligned}
&-\nabla_f\left(A_j, \frac{1}{P_n}\sum_{i=1}^n p_i A_i - A_j\right) \\
&\geq f(A_j) - f\left(\frac{1}{P_n}\sum_{i=1}^n p_i A_i\right) \\
&\geq \nabla_f\left(\frac{1}{P_n}\sum_{i=1}^n p_i A_i, A_j - \frac{1}{P_n}\sum_{i=1}^n p_i A_i\right)
\end{aligned} \tag{25}$$

for any $j \in \{1, \ldots, n\}$. We obtain the desired inequalities (23) by multiplying the inequalities in (25) by $p_j \geq 0$ and taking the sum over j from 1 to n; and divide the resulted inequalities by P_n. □

Corollary 17. *Under the assumptions of Theorem 16, we have the following results:*

(1) If $\nabla_f : \mathcal{C} \times \mathcal{A}(H)$ is convex, then

$$\frac{1}{P_n}\sum_{j=1}^n p_j f(A_j) - f\left(\frac{1}{P_n}\sum_{i=1}^n p_i A_i\right) \geq \nabla_f\left(\frac{1}{P_n}\sum_{i=1}^n p_i A_i, 0\right) \tag{26}$$

(2) If ∇_f is linear then $\nabla_f(B,0) = 0$ for any B and we get the Jensen's inequality:

$$\frac{1}{P_n}\sum_{j=1}^{n} p_j f(A_j) - f\left(\frac{1}{P_n}\sum_{i=1}^{n} p_i A_i\right) \geq 0. \tag{27}$$

(3) If ∇_f is linear, we have

$$\begin{aligned} &\frac{1}{P_n}\sum_{j=1}^{n} p_j \nabla_f\left(A_j, A_j - \frac{1}{P_n}\sum_{i=1}^{n} p_i A_i\right) \\ \geq\ & \frac{1}{P_n}\sum_{j=1}^{n} p_j f(A_j) - f\left(\frac{1}{P_n}\sum_{i=1}^{n} p_i A_i\right) \geq 0. \end{aligned} \tag{28}$$

Theorem 18. *Under the assumptions of Theorem 16, we have the following results:*

$$\begin{aligned} &\frac{1}{P_n}\sum_{j=1}^{n} p_j f(A_j) - \frac{1}{P_n}\sum_{j=1}^{n} p_j \nabla_f(A, A_j - A) \\ \geq\ & f(A) \\ \geq\ & \frac{1}{P_n}\sum_{j=1}^{n} p_j f(A_j) + \frac{1}{P_n}\sum_{j=1}^{n} p_j \nabla_f(A_j, A - A_j). \end{aligned}$$

Proof. From (24) we also have

$$-\nabla_f(A, A_j - A) \geq f(A) - f(A_j) \geq \nabla_f(A_j, A - A_j). \tag{29}$$

If we multiply (29) by $p_j \geq 0$ and take the sum over j from 1 to n, then

$$\begin{aligned} -\frac{1}{P_n}\sum_{j=1}^{n} p_j \nabla_f(A, A_j - A) &\geq f(A) - \frac{1}{P_n}\sum_{j=1}^{n} p_j f(A_j) \\ &\geq \frac{1}{P_n}\sum_{j=1}^{n} p_j \nabla_f(A_j, A - A_j) \end{aligned}$$

which completes the proof. □

Remark 19. If ∇_f is linear in Theorem 18, then we get simpler inequalities such as:

$$f(A) \geq \frac{1}{P_n}\sum_{j=1}^{n} p_j f(A_j) + \frac{1}{P_n}\sum_{j=1}^{n} p_j \nabla_f(A_j, A) - \frac{1}{P_n}\sum_{j=1}^{n} p_j \nabla_f(A_j, A_j);$$

and

$$f(A) \leq \frac{1}{P_n}\sum_{j=1}^{n} p_j f(A_j) - \frac{1}{P_n}\sum_{j=1}^{n} p_j \nabla_f(A, A_j) + \frac{1}{P_n}\sum_{j=1}^{n} p_j \nabla_f(A, A).$$

Therefore, if $A \in \mathcal{A}(H)$ is such that

$$\frac{1}{P_n}\sum_{j=1}^{n} p_j \nabla_f(A_j, A) \geq \frac{1}{P_n}\sum_{j=1}^{n} p_j \nabla_f(A_j, A_j),$$

then we have Slater type inequality (cf. Slater [29]):

$$f(A) \geq \frac{1}{P_n}\sum_{j=1}^{n} p_j f(A_j).$$

6. Connection with Gâteaux derivative

In this section, we consider the connection between the order generalised gradient and the Gâteaux derivative. We refer the reader to Dragomir [13] for some inequalities of Jensen type, involving Gâteaux derivatives for convex functions in linear spaces.

Let $\mathcal{C} \subset \mathcal{A}(H)$ be a convex set and $f : \mathcal{C} \to \mathcal{A}(H)$ is said to be ***operator convex*** if for all $t \in [0,1]$, and $A, B \in \mathcal{C}$, we have

$$f[(1-t)A + tB] \leq (1-t)f(A) + tf(B).$$

We start with the following lemmas.

Lemma 20. *Let $F : \mathbb{R} \to \mathcal{B}(H)$ be a function such that $\lim_{t\to 0^\pm} F(t)$ exists. Then, both $\lim_{t\to 0^\pm} F(t)$ are bounded linear operators and*

$$\left\langle \left[\lim_{t\to 0^\pm} F(t)\right] x, y \right\rangle = \lim_{t\to 0^\pm} \langle F(t)x, y\rangle$$

for all nonzero $x, y \in H$.

Proof. We provide the proof for the right-sided limit, as the proof for the left-sided limit follows similarly. Let $\varepsilon > 0$ and for $x, y \in H$, where $x, y \neq 0$, set $\varepsilon_0 = \varepsilon/(\|x\|_H \|y\|_H)$. Since $\lim_{t\to 0^+} F(t) = L$, there exists δ_0 such that

$$\|F(t) - L\|_{\mathcal{B}(H)} < \varepsilon_0$$

when $0 < t < \delta_0$. Note that $L \in \mathcal{B}(H)$ since $\mathcal{B}(H)$ is a Banach space, hence $F(t) - L$ is also a bounded linear operator. Now, we have

$$\begin{aligned} |\langle F(t)x, y\rangle - \langle Lx, y\rangle| &\leq \|(F(t) - L)x\|_H \|y\|_H \\ &\leq \|F(t) - L\|_{\mathcal{B}(H)} \|x\|_H \|y\|_H < \varepsilon_0 \|x\|_H \|y\|_H = \varepsilon; \end{aligned}$$

which completes the proof. □

Lemma 21. *Let $f : \mathcal{A}(H) \to \mathcal{A}(H)$ be operator convex and $A \in \mathcal{A}(H)$. Then for all $B \in \mathcal{A}(H)$, both limits*

$$(\nabla_- f(A))(B) = \lim_{t\to 0^-} \frac{f(A+tB) - f(A)}{t}$$

and

$$(\nabla_+ f(A))(B) = \lim_{t\to 0^+} \frac{f(A+tB) - f(A)}{t}$$

exist and are bounded self-adjoint operators.

Proof. Fix an arbitrary $B \in \mathcal{A}(H)$, and let

$$G(t) = \frac{f(A+tB) - f(A)}{t}, \quad t \in \mathbb{R} \setminus \{0\}.$$

We want to show that G is nondecreasing. Let $0 < t_1 < t_2$, then

$$\begin{aligned} f(A + t_1 B) - f(A) &= f\left[\frac{t_1}{t_2}(A + t_2 B) + \left(1 - \frac{t_1}{t_2}\right) A\right] - f(A) \\ &\leq \frac{t_1}{t_2} f(A + t_2 B) + \left(1 - \frac{t_1}{t_2}\right) f(A) - f(A) \\ &= \frac{t_1}{t_2}[f(A + t_2 B) - f(A)]. \end{aligned}$$

Thus,

$$\frac{f(A + t_1 B) - f(A)}{t_1} \leq \frac{f(A + t_2 B) - f(A)}{t_2}.$$

Also,

$$\begin{aligned}\frac{f(A-t_2B)-f(A)}{-t_2} &= -\frac{f[A+t_2(-B)]-f(A)}{t_2} \\ &\leq -\frac{f[A+t_1(-B)]-f(A)}{t_1} = \frac{f(A-t_1B)-f(A)}{-t_1}.\end{aligned}$$

Note also that

$$\begin{aligned}f(A) = f\left(\frac{2A+t_1B-t_1B}{2}\right) &= f\left[\frac{1}{2}(A+t_1B)+\frac{1}{2}(A-t_1B)\right] \\ &\leq \frac{1}{2}f(A+t_1B)+\frac{1}{2}f(A-t_1B).\end{aligned}$$

which implies that

$$2f(A) \leq f(A+t_1B)+f(A-t_1B);$$

and thus

$$f(A+t_1B)-f(A) \geq -[f(A-t_1B)-f(A)],$$

which implies that

$$\frac{f(A+t_1B)-f(A)}{t_1} \leq \frac{f(A-t_1B)-f(A)}{-t_1}.$$

By the above expositions, we conclude that G is nondecreasing on $\mathbb{R}\setminus\{0\}$. This proves that both $(\nabla_- f(A))(B)$ and $(\nabla_+ f(A))(B)$ exist and are bounded linear operators by Lemma 20. Note that for all $t \in \mathbb{R}$, $t \neq 0$ and $A, B \in \mathcal{A}(H)$,

$$\frac{f[B+t(A-B)]-f(B)}{t}$$

is a self-adjoint operator. Let $x, y \in H$, then Lemma 20 gives us

$$\begin{aligned}&\left\langle \left[\lim_{t\to 0^\pm}\frac{f[B+t(A-B)]-f(B)}{t}\right]x, y\right\rangle \\ &= \lim_{t\to 0^\pm}\left\langle \left[\frac{f[B+t(A-B)]-f(B)}{t}\right]x, y\right\rangle \\ &= \lim_{t\to 0^\pm}\left\langle x, \left[\frac{f[B+t(A-B)]-f(B)}{t}\right]y\right\rangle \\ &= \left\langle x, \lim_{t\to 0^\pm}\left[\frac{f[B+t(A-B)]-f(B)}{t}\right]y\right\rangle,\end{aligned}$$

which completes the proof. □

Theorem 22. *Let $\mathcal{C} \subset \mathcal{A}(H)$ be a convex set and $f : \mathcal{C} \to \mathcal{A}(H)$ be operator convex. Then,*

$$(\nabla_\pm f(B))(A-B) = \lim_{t\to 0^\pm}\frac{f[B+t(A-B)]-f(B)}{t} \tag{30}$$

is an order generalised gradient.

Proof. Let $t \in (0,1)$, and $A, B \in \mathcal{C}$. Since f is operator convex, we have

$$\begin{aligned}\frac{f[B+t(A-B)]-f(B)}{t} &= \frac{f[(1-t)B+tA]-f(B)}{t} \\ &\leq \frac{(1-t)f(B)+tf(A)-f(B)}{t} = f(A)-f(B)\end{aligned}$$

This is equivalent to

$$K := f(A)-f(B)-\frac{f[B+t(A-B)]-f(B)}{t} \in \mathcal{P}_+(H).$$

Note that for all $x \in H$

$$\left\langle \left[\lim_{t\to 0^{\pm}} K \right] x, x \right\rangle = \lim_{t\to 0^{\pm}} \langle Kx, x\rangle ,$$

by Lemma 20. Since $K \in \mathcal{P}_+(H)$, $\langle Kx, x\rangle \geq 0$ and thus $\langle [\lim_{t\to 0^{\pm}} K]\, x, x\rangle \geq 0$ which implies that

$$\lim_{t\to 0^{\pm}} \left[f(A) - f(B) - \frac{f[B + t(A - B)] - f(B)}{t} \right] \in \mathcal{P}_+(H).$$

Therefore,

$$\big(\nabla_+ f(B)\big)(A - B) = \lim_{t\to 0^{\pm}} \frac{f[B + t(A - B)] - f(B)}{t} \leq f(A) - f(B)$$

Lemma 21 gives us

$$\big(\nabla_- f(B)\big)(A - B) \leq \big(\nabla_+ f(B)\big)(A - B)$$

which implies that

$$\big(\nabla_- f(B)\big)(A - B) \leq f(A) - f(B).$$

Thus both $\nabla_+ f$ and $\nabla_- f$ are order generalised gradient. □

Proposition 23. *Let $f : \mathcal{A}(H) \to \mathcal{A}(H)$ be operator convex and $A \in \mathcal{A}(H)$. The right Gâteaux derivative is sub-additive, i.e.*

$$(\nabla_+ f(A))(B + C) \leq (\nabla_+ f(A))(B) + (\nabla_+ f(A))(C)$$

for any $B, C \in \mathcal{A}(H)$. The left Gâteaux derivative is super-additive, i.e.

$$(\nabla_- f(A))(B + C) \geq (\nabla_- f(A))(B) + (\nabla_- f(A))(C)$$

for any $B, C \in \mathcal{A}(H)$.

Proof. Since f is operator convex, we have the following for any $B, C \in \mathcal{A}$ and $t > 0$

$$\begin{aligned} \frac{f[A + t(B + C)] - f(A)}{t} &= \frac{f[\frac{1}{2}(A + 2tB) + \frac{1}{2}(A + 2tC)] - f(A)}{t} \\ &\leq \frac{f(A + 2tB)) - f(A)}{2t} + \frac{f(A + 2tC) - f(A)}{2t} \end{aligned}$$

By a similar argument to the proof of Theorem 22, we conclude that

$$\begin{aligned} (\nabla_+ f(A))(B + C) &= \lim_{t\to 0^+} \frac{f[A + t(B + C)] - f(A)}{t} \\ &\leq \lim_{t\to 0^+} \frac{f(A + 2tB)) - f(A)}{2t} + \lim_{t\to 0^+} \frac{f(A + 2tC) - f(A)}{2t} \\ &= (\nabla_+ f(A))(B) + (\nabla_+ f(A))(C) \end{aligned}$$

as desired. The proof for the left Gâteaux derivative follows similarly. □

Remark 24. We remark that the Gâteaux (lateral) derivative(s) is always positive homogeneous with respect to the second variable, i.e. for any function $f : \mathcal{A}(H) \to \mathcal{A}(H)$ and fixed $A \in \mathcal{A}(H)$

$$(\nabla_{\pm} f(A))(\alpha B) = \alpha(\nabla_{\pm} f(A))(B)$$

for all $\alpha \geq 0$ and $B \in \mathcal{A}(H)$. The Gâteaux derivative, on the other hand is always homogeneous with respect to the second variable, i.e. for any function $f : \mathcal{A}(H) \to \mathcal{A}(H)$ and fixed $A \in \mathcal{A}(H)$

$$(\nabla f(A))(\alpha B) = \alpha(\nabla f(A))(B)$$

for all $\alpha \in \mathbb{C}$ and $B \in \mathcal{A}(H)$.

The following result restates Theorem 3 in the settings of operator-valued functions.

Corollary 25. *Let $\mathcal{C} \subset \mathcal{A}(H)$ be a convex set and $f : \mathcal{C} \to \mathcal{A}(H)$ be a Gâteaux differentiable function. Then f is operator convex if and only if*

$$(\nabla f(B))(A-B) = \lim_{t\to 0} \frac{f[B+t(A-B)]-f(B)}{t} \tag{31}$$

is an order generalised gradient.

Proof. For any $A, B \in \mathcal{C}$, if f is operator convex then by Theorem 22

$$(\nabla_{\pm} f(B))(A-B) = \lim_{t\to 0^{\pm}} \frac{f[B+t(A-B)]-f(B)}{t}$$

are order generalised gradients. Since f is assumed to be Gâteaux differentiable, both limits are equal, hence

$$(\nabla f(B))(A-B) = \lim_{t\to 0} \frac{f[B+t(A-B)]-f(B)}{t}$$

is an order generalised gradient for any $A, B \in \mathcal{C}$. Conversely, we have the following inequality

$$(\nabla f(B))(A-B) \le f(A) - f(B)$$

for any $A, B \in \mathcal{C}$. Let $C, D \in \mathcal{C}$, $t \in (0,1)$; and choose $A = C$ and $B = tC+(1-t)D$, we have

$$(1-t)(\nabla f[tC+(1-t)D])(C-D) \le f(C) - f[tC+(1-t)D]. \tag{32}$$

Let $A = D$ and $B = tC + (1-t)D$, we have

$$(-t)(\nabla f[tC+(1-t)D])(C-D) \le f(D) - f[tC+(1-t)D]. \tag{33}$$

Multiply (32) by t and (33) by $(1-t)$, and add the resulting inequalities to obtain

$$f[tC+(1-t)D] \le tf(C) + (1-t)f(D)$$

which completes the proof. □

The following result follows by Corollary 14 and employing the fact that the Gâteaux lateral derivatives are positive homogenous.

Corollary 26 (Hermite-Hadamard type inequality)**.** *Let $f : \mathcal{C} \subset \mathcal{A}(H) \to \mathcal{A}(H)$ be operator convex. The following inequality holds*

$$\begin{aligned} & -\frac{1}{6}[(\nabla_{\pm} f(B))(A-B) + (\nabla_{\pm} f(A))(B-A)] \\ \ge \quad & \frac{f(A)+f(B)}{2} - \int_0^1 f(tA+(1-t)B)dt \ge 0. \end{aligned}$$

The above inequality also holds for ∇_f when f is Gâteaux differentiable.

Example 27. (1) We note that the function $f(x) = -\log(x)$ is operator convex. The log function is (operator) Gâteaux differentiable with the following explicit formula for the derivative (cf. Pedersen [27, p. 155]):

$$(\nabla \log(A))(B) = \int_0^{\infty} (sI+A)^{-1} B (sI+A)^{-1} ds$$

for $A, B \in \mathcal{A}(H)$ and I the identity operator. Thus, we have the following inequality

$$\begin{aligned} & \frac{1}{6}\left[\int_0^\infty (sI+B)^{-1}(A-B)(sI+B)^{-1}ds \right. \\ & \left. + \int_0^\infty (sI+A)^{-1}(B-A)(sI+A)^{-1}ds\right] \\ \geq \; & -\frac{\log(A)+\log(B)}{2} + \int_0^1 \log(tA+(1-t)B)dt \geq 0. \end{aligned}$$

(2) We consider the operator convex function $f(x) = x\log(x)$, and using the following representation (cf. Pedersen [27, p. 155]):

$$\log(t) = \int_0^\infty \frac{t-1}{(s+t)(s+1)}ds,$$

and note the fact that $\frac{d}{dt}t\log(t) = \log(t)$, we have

$$(\nabla f(A))(B) = \int_0^\infty \frac{1}{s+1}(sI+A)^{-1}(A-I)B\ ds.$$

Then, we have the following inequalities

$$\begin{aligned} & -\frac{1}{6}\left[\int_0^\infty \frac{1}{s+1}(sI+B)^{-1}(B-I)(A-B)\ ds \right. \\ & \left. + \int_0^\infty \frac{1}{s+1}(sI+A)^{-1}(A-I)(B-A)\ ds\right] \\ \geq \; & \frac{A\log(A)+B\log(B)}{2} - \int_0^1 [tA+(1-t)B]\log(tA+(1-t)B)dt \geq 0. \end{aligned}$$

The following results follow by Theorems 16 and 18.

Corollary 28 (Jensen type inequality)**.** *Let $f : \mathcal{C} \subset \mathcal{A}(H) \to \mathcal{A}(H)$ be operator convex. Then for any $A_i \in \mathcal{C}$, $i \in \{1,\ldots,n\}$ and $p_i \geq 0$ with $P_n := \sum_{i=1}^n p_i > 0$ we have the inequalities*

$$\begin{aligned} & -\frac{1}{P_n}\sum_{j=1}^n p_j(\nabla_\pm f(A_j))\left(\frac{1}{P_n}\sum_{i=1}^n p_iA_i - A_j\right) \\ & \geq \frac{1}{P_n}\sum_{j=1}^n p_jf(A_j) - f\left(\frac{1}{P_n}\sum_{i=1}^n p_iA_i\right) \\ & \geq \frac{1}{P_n}\sum_{j=1}^n p_j\left(\nabla_\pm f\left(\frac{1}{P_n}\sum_{i=1}^n p_iA_i\right)\right)\left(A_j - \frac{1}{P_n}\sum_{i=1}^n p_iA_i\right). \end{aligned}$$

We also have

$$\begin{aligned} & \frac{1}{P_n}\sum_{j=1}^n p_jf(A_j) - \frac{1}{P_n}\sum_{j=1}^n p_j(\nabla_\pm f(A))(A_j - A) \\ \geq \; & f(A) \\ \geq \; & \frac{1}{P_n}\sum_{j=1}^n p_jf(A_j) + \frac{1}{P_n}\sum_{j=1}^n p_j(\nabla_\pm f(A_j)(A - A_j). \end{aligned}$$

The above inequalities also holds for ∇_f when f is Gâteaux differentiable.

Example 29. (1) We have the following inequalities for the operator convex function $f(x) = -\log(x)$:

$$\frac{1}{P_n}\sum_{j=1}^{n} p_j \int_0^\infty (sI+A_j)^{-1}\left(\frac{1}{P_n}\sum_{i=1}^{n} p_i A_i - A_j\right)(sI+A_j)^{-1} ds$$
$$\geq -\frac{1}{P_n}\sum_{j=1}^{n} p_j \log(A_j) + \log\left(\frac{1}{P_n}\sum_{i=1}^{n} p_i A_i\right)$$
$$\geq -\frac{1}{P_n}\sum_{j=1}^{n} p_j \int_0^\infty \left(sI + \frac{1}{P_n}\sum_{i=1}^{n} p_i A_i\right)^{-1}\left(A_j - \frac{1}{P_n}\sum_{i=1}^{n} p_i A_i\right)\left(sI + \frac{1}{P_n}\sum_{i=1}^{n} p_i A_i\right)^{-1} ds$$

and

$$-\frac{1}{P_n}\sum_{j=1}^{n} p_j \log(A_j) + \frac{1}{P_n}\sum_{j=1}^{n} p_j \int_0^\infty (sI+A)^{-1}(A_j - A)(sI+A)^{-1} ds$$
$$\geq -\log(A)$$
$$\geq -\frac{1}{P_n}\sum_{j=1}^{n} p_j \log(A_j) - \frac{1}{P_n}\sum_{j=1}^{n} p_j \int_0^\infty (sI+A_j)^{-1}(A - A_j)(sI+A_j)^{-1} ds.$$

(2) We have the following inequalities for the operator convex function $f(x) = x\log(x)$:

$$-\frac{1}{P_n}\sum_{j=1}^{n} p_j \int_0^\infty \frac{1}{s+1}(sI+A_j)^{-1}(A_j - I)\left(\frac{1}{P_n}\sum_{i=1}^{n} p_i A_i - A_j\right) ds$$
$$\geq \frac{1}{P_n}\sum_{j=1}^{n} p_j A_j \log(A_j) - \left(\frac{1}{P_n}\sum_{i=1}^{n} p_i A_i\right)\log\left(\frac{1}{P_n}\sum_{i=1}^{n} p_i A_i\right)$$
$$\geq \frac{1}{P_n}\sum_{j=1}^{n} p_j \int_0^\infty \frac{1}{s+1}\left(sI + \frac{1}{P_n}\sum_{i=1}^{n} p_i A_i\right)^{-1}\left(\frac{1}{P_n}\sum_{i=1}^{n} p_i A_i - I\right)\left(A_j - \frac{1}{P_n}\sum_{i=1}^{n} p_i A_i\right) ds.$$

and

$$\frac{1}{P_n}\sum_{j=1}^{n} p_j A_j \log(A_j) - \frac{1}{P_n}\sum_{j=1}^{n} p_j \int_0^\infty \frac{1}{s+1}(sI+A)^{-1}(A - I)(A_j - A) ds$$
$$\geq A\log(A)$$
$$\geq \frac{1}{P_n}\sum_{j=1}^{n} p_j A_j \log(A_j) + \frac{1}{P_n}\sum_{j=1}^{n} p_j \int_0^\infty \frac{1}{s+1}(sI+A_j)^{-1}(A_j - I)(A - A_j) ds.$$

References

[1] J.S. Aujla, Matrix convexity of functions of two variables, Linear Algebra Appl. 194 (1993) 149–160.

[2] J.S. Aujla, On an operator inequality, Linear Algebra Appl. 310 No. 1-3 (2000) 43–47.

[3] E.F. Beckenbach, Convex functions, Bull. Amer. Math. Soc. 54 (1948) 439–460.

[4] T. Damm and D. Hinrichsen, Newton's method for concave operators with resolvent positive derivatives in ordered Banach spaces, Linear Algebra Appl. 363 (2003) 43–64, Special issue on nonnegative matrices, M-matrices and their generalizations (Oberwolfach, 2000).

[5] S.S. Dragomir, An improvement of Jensen's inequality, Bull. Math. Soc. Sci. Math. Roumanie 34 (82) No. 4 (1990) 291–296.

[6] S.S. Dragomir, Some refinements of Ky Fan's inequality, J. Math. Anal. Appl. 163 (2) (1992) 317–321.

[7] S.S. Dragomir, Some refinements of Jensen's inequality, J. Math. Anal. Appl. 163 (2) (1992) 518–522.

[8] S.S. Dragomir, A further improvement of Jensen's inequality, Tamkang J. Math. 25 (1) (1994) 29–36.

[9] S.S. Dragomir, A new improvement of Jensen's inequality, Indian J. Pure and Appl. Math. 26 (10) (1995) 959–968.

[10] S.S. Dragomir, An inequality improving the first Hermite-Hadamard inequality for convex functions defined on linear spaces and applications for semi-inner products, J. Inequal. Pure Appl. Math. 3 No. 2 (2002) Article 31.

[11] S.S. Dragomir, An inequality improving the second Hermite-Hadamard inequality for convex functions defined on linear spaces and applications for semi-inner products, J. Inequal. Pure Appl. Math. 3 No. 3 (2002) Article 35.
[12] S.S. Dragomir, Semi-inner Products and Applications, Nova Scienve Publishers Inc., New York, 2004.
[13] S.S. Dragomir, Inequalities in terms of the Gâteaux derivatives for convex functions on linear spaces with applications, Bull. Aust. Math. Soc. 83 No. 3 (2011) 500–517.
[14] S.S. Dragomir, A refinement of Jensen's inequality with applications for f-divergence measures, Taiwanese J. Maths., In Press.
[15] E. Eisenberg, A gradient inequality for a class of nondifferentiable functions, Operations Res. 14 (1966), 157–163.
[16] F. Hansen, An operator inequality, Math. Ann. 246 No. 3 (1979/80) 249–250.
[17] F. Hansen, Jensen's operator inequality for functions of two variables, Proc. Amer. Math. Soc. 125 No. 7 (1997) 2093–2102.
[18] F. Hansen, Operator convex functions of several variables, Publ. Res. Inst. Math. Sci. 33 No. 3 (1997) 443–463.
[19] F. Hansen, and G.K. Pedersen, Jensen's inequality for operators and Löwner's theorem, Math. Ann. 258 No. 3 (1981/82) 229–241.
[20] F. Hansen, and G.K. Pedersen, Jensen's operator inequality, Bull. London Math. Soc. 35 No. 4 (2003) 553–564.
[21] D.S. Mitrinović and I.B. Lacković, Hermite and convexity, Aequationes Math. 28 (1985) 229–232.
[22] H.-C. Lai, and J.-C. Liu, Duality for nondifferentiable minimax programming in complex spaces, Nonlinear Anal. 71 No. 12 (2009) e224–e233.
[23] J.E. Pečarić, Josip E., A multidimensional generalization of Slater's inequality, J. Approx. Theory 44 No. 3 (1985) 292–294.
[24] J.E. Pečarić and S.S. Dragomir, A refinement of Jensen inequality and applications, Studia Univ. Babeş-Bolyai, Mathematica 24 (1) (1989) 15–19.
[25] J.E. Pečarić and S.S. Dragomir, A generalisation of Hadamard's inequality for isotonic linear functional, Radovi Mat. (Sarajevo) 7 (1991) 103–107.
[26] J.E. Pečarić, F. Proschan and Y.L. Tong, Convex Functions, Partial Orderings, and Statistical Applications, Academic Press Inc., San Diego, 1992.
[27] G.K. Pedersen, Operator differentiable functions, Publ. Res. Inst. Math. Sci. 36 No. 1 (2000) 139–157.
[28] R.T. Rockafellar, R. Tyrrell, *Convex analysis*, Princeton Landmarks in Mathematics, Reprint of the 1970 original, Princeton Paperbacks, Princeton University Press, Princeton, NJ, 1997.
[29] M.L. Slater, A companion inequality to Jensen's inequality, J. Approx. Theory 32 No. 2 (1981) 160–166.
[30] J.S. Vandergraft, Newton's method for convex operators in partially ordered spaces, SIAM J. Numer. Anal. 4 (1967), 406–432.