Applied Mathematics and Computation 243 (2014) 528–536

Contents lists available at ScienceDirect

Applied Mathematics and Computation

journal homepage: www.elsevier.com/locate/amc

Comparison of the normalized Jensen functionals of two convex functions with applications

Jamal Rooin *, Hossein Dehghan, Akram Alikhani

Department of Mathematics, Institute for Advanced Studies in Basic Sciences (IASBS), Gava Zang, Zanjan 45137-66731, Iran

ARTICLE INFO

Keywords:
Convex function
Jensen's inequality
Kullback–Leibler divergence
Shannon's entropy
Inner product space

ABSTRACT

In this paper, we compare normalized Jensen functionals of two given convex functions defined on an interval of the real line. As applications, we give some valuable upper and lower bounds for the AGM inequality, which lead to some comparison results regarding Kullback–Leibler divergence and Shannon's entropy. Some applications in inner product spaces are also included.

1. Introduction

Throughout this paper, we denote by I an arbitrary nondegenerate interval of the real line $\mathbb{R}$, I° the interior of I, $C^2(I)$ the set of all twice continuously differentiable real-valued functions on I and $n \geqslant 2$ an integer number. Consistent with [5], we denote by $\mathcal{P}_n$ the set of all nonnegative n-tuples $\mathbf{p} = (p_1, \dots, p_n)$ with the property that $\sum_{i=1}^n p_i = 1$. Consider the *normalized Jensen functional*

$$\mathcal{J}_n(f, \mathbf{x}, \mathbf{p}) = \sum_{i=1}^n p_i f(x_i) - f\left(\sum_{i=1}^n p_i x_i\right) \geqslant 0,$$

where $f : C \to \mathbb{R}$ is a convex function on a convex set C in a real linear space, $\mathbf{x} = (x_1, \dots, x_n) \in C^n$ and $\mathbf{p} \in \mathcal{P}_n$.

Recently, Dragomir established the following theorem which compares two different normalized Jensen functionals.

Theorem A [5, Theorem 1]. *Given $\mathbf{p}, \mathbf{q} \in \mathcal{P}_n$, $q_i > 0$ for each $i \in \{1, \dots, n\}$, we have*

$$(0 \leqslant) \min_{1 \leqslant i \leqslant n} \left\{\frac{p_i}{q_i}\right\} \mathcal{J}_n(f, \mathbf{x}, \mathbf{q}) \leqslant \mathcal{J}_n(f, \mathbf{x}, \mathbf{p}) \leqslant \max_{1 \leqslant i \leqslant n} \left\{\frac{p_i}{q_i}\right\} \mathcal{J}_n(f, \mathbf{x}, \mathbf{q}) \qquad (1.1)$$

for any convex function $f : C \to \mathbb{R}$ and $\mathbf{x} \in C^n$.

The following two natural questions arise:

(**I**) *Given $\mathbf{x}, \mathbf{y} \in C^n$, what are (preferably the best) possible constants $\lambda, \mu \geqslant 0$, depending only on $\mathbf{x}$ and $\mathbf{y}$, such that*

$$(0 \leqslant) \lambda \mathcal{J}_n(f, \mathbf{y}, \mathbf{p}) \leqslant \mathcal{J}_n(f, \mathbf{x}, \mathbf{p}) \leqslant \mu \mathcal{J}_n(f, \mathbf{y}, \mathbf{p}) \qquad (1.2)$$

for any $\mathbf{p} \in \mathcal{P}_n$ and convex function $f : C \to \mathbb{R}$.

* Corresponding author.
E-mail addresses: rooin@iasbs.ac.ir (J. Rooin), h_dehghan@iasbs.ac.ir (H. Dehghan), alikhani@iasbs.ac.ir (A. Alikhani).

http://dx.doi.org/10.1016/j.amc.2014.05.131
0096-3003/

(**II**) *Given two convex functions $f, g : C \to \mathbb{R}$, what are (preferably the best) possible constants $\lambda, \mu \geqslant 0$, depending only on f and g, such that*

$$(0 \leqslant)\ \lambda \mathcal{J}_n(g, \mathbf{x}, \mathbf{p}) \leqslant \mathcal{J}_n(f, \mathbf{x}, \mathbf{p}) \leqslant \mu \mathcal{J}_n(g, \mathbf{x}, \mathbf{p}) \tag{1.3}$$

for any $\mathbf{p} \in \mathcal{P}_n$ and $\mathbf{x} \in C^n$.

Regarding to the question (I) we have the following result which is an immediate consequence of Theorem 1.4.1 of [10] due to Niculescu.

Theorem B. *If $c, d \in [a, b]$, then*

$$p_1 f(c) + p_2 f(d) - f(p_1 c + p_2 d) \leqslant p_1 f(a) + p_2 f(b) - f(p_1 a + p_2 b) \tag{1.4}$$

for any convex function $f : [a, b] \to \mathbb{R}$ and $(p_1, p_2) \in \mathcal{P}_2$.

This yields the left hand inequality in (1.2) in the special case of $n = 2,\ C = [a, b] \subseteq \mathbb{R},\ \mathbf{x} = (a, b)$ and $\mathbf{y} = (c, d)$, with $\lambda = 1$.

In 2002, Dragomir and Scarmozzino [7] without exposing the question (II), considered implicitly the convex functions $\ln((1-t)/t))$ and $-\ln t$ on $(0, 1/2]$ and gave a refinement and a converse to the Ky Fan inequality (3.1).

In this paper, motivating by Theorem A and generalizing the idea of [7], we give a positive answer to the existence of the best possible constants $\lambda, \mu \geqslant 0$ in the question (II) in the important case of $C = I \subseteq \mathbb{R}$ and $f, g \in C^2$. This yields some interesting new inequalities in Information Theory and inner product spaces.

2. Main results

We may state the main theorem of this paper as follows which answers positively the question (II).

Theorem 2.1. *Let $f, g : I \to \mathbb{R}$ be two continuous and convex functions on I belonging to $C^2(I^\circ)$. Now, if $g'' > 0$ on I° and $\sup_{t \in I^\circ}\left\{\frac{f''(t)}{g''(t)}\right\} < \infty$, then for any $\mathbf{p} \in \mathcal{P}_n$ and $\mathbf{x} \in I^n$,*

$$(0 \leqslant)\ \inf_{t \in I^\circ}\left\{\frac{f''(t)}{g''(t)}\right\}\ \mathcal{J}_n(g, \mathbf{x}, \mathbf{p}) \leqslant \mathcal{J}_n(f, \mathbf{x}, \mathbf{p}) \leqslant \sup_{t \in I^\circ}\left\{\frac{f''(t)}{g''(t)}\right\}\ \mathcal{J}_n(g, \mathbf{x}, \mathbf{p}). \tag{2.1}$$

Moreover, the constants in (2.1) are best possible.

If the infimum (supremum) is not taken on I° and $p_i > 0$ $(i = 1, \ldots, n)$, then equality holds in the left (right) hand of (2.1) if and only if $x_1 = \cdots = x_n$.

Proof. Let

$$\lambda := \inf_{t \in I^\circ}\left\{\frac{f''(t)}{g''(t)}\right\} \quad \text{and} \quad \mu := \sup_{t \in I^\circ}\left\{\frac{f''(t)}{g''(t)}\right\}.$$

Setting $h_\lambda(t) := f(t) - \lambda g(t)$ for all $t \in I$, we have

$$h''_\lambda(t) = f''(t) - \lambda g''(t) = g''(t)\left(\frac{f''(t)}{g''(t)} - \lambda\right) \geqslant 0 \quad (t \in I^\circ).$$

Therefore, h_λ is convex on I, and so

$$f\left(\sum_{i=1}^n p_i x_i\right) - \lambda g\left(\sum_{i=1}^n p_i x_i\right) = h_\lambda\left(\sum_{i=1}^n p_i x_i\right) \leqslant \sum_{i=1}^n p_i h_\lambda(x_i) = \sum_{i=1}^n p_i f(x_i) - \lambda \sum_{i=1}^n p_i g(x_i),$$

which implies that

$$\lambda\ \mathcal{J}_n(g, \mathbf{x}, \mathbf{p}) \leqslant \mathcal{J}_n(f, \mathbf{x}, \mathbf{p}).$$

Similarly, setting $h_\mu(t) := f(t) - \mu g(t)$ for all $t \in I$, we have

$$h''_\mu(t) = f''(t) - \mu g''(t) = g''(t)\left(\frac{f''(t)}{g''(t)} - \mu\right) \leqslant 0 \quad (t \in I^\circ).$$

Hence, concavity of h_μ on I implies that

$$f\left(\sum_{i=1}^n p_i x_i\right) - \mu g\left(\sum_{i=1}^n p_i x_i\right) = h_\mu\left(\sum_{i=1}^n p_i x_i\right) \geqslant \sum_{i=1}^n p_i h_\mu(x_i) = \sum_{i=1}^n p_i f(x_i) - \mu \sum_{i=1}^n p_i g(x_i),$$

that is,

$$\mathcal{J}_n(f, \mathbf{x}, \mathbf{p}) \leqslant \mu\ \mathcal{J}_n(g, \mathbf{x}, \mathbf{p}).$$