Kragujevac Journal of Mathematics
Volume 37(2) (2013), Pages 269–273.

ON A CLASS OF SELF-IMPROVING INEQUALITIES

SLAVKO SIMIĆ

Abstract. We establish new lower and upper bounds for Jensen's discrete inequality. Applying those results we improve some classical inequalities and obtain new and more precise bounds for Shannon's entropy.

1. Introduction

In this article we shall consider a class of inequalities with remarkable property that they can be improved by themselves. To make this idea clear we give an example.

Example 1.1. Take for instance the well-known arithmetic-geometric inequality, written in the form

$$\frac{\left(\frac{\sum_1^n x_i}{n}\right)^n}{\prod_1^n x_i} \ge 1,$$

where $0 < a = x_1 \le x_2 \le \cdots \le x_n = b$. This inequality can be improved by itself in the following way:

$$\begin{aligned}\Big(\frac{1}{n}\sum_1^n x_i\Big)^n &= \Big(\frac{1}{n}\Big(\sum\nolimits_2^{n-1} x_i + \frac{a+b}{2} + \frac{a+b}{2}\Big)\Big)^n \\ &\ge \Big(\frac{a+b}{2}\Big)^2 \prod_2^{n-1} x_i \\ &= \frac{(a+b)^2}{4ab} \prod_1^n x_i.\end{aligned}$$

Key words and phrases. Jensen's inequality, Global bounds, Entropy.
2010 *Mathematics Subject Classification.* Primary: 05C69. Secondary: 05C20.
Received: February 1, 2013.
Revised: October 21, 2013.

Hence,

$$\frac{\left(\frac{\sum_1^n x_i}{n}\right)^n}{\prod_1^n x_i} \geq 1 + \frac{1}{4}\Big(\sqrt{\frac{b}{a}} - \sqrt{\frac{a}{b}}\Big)^2,$$

which is a considerable improvement of the target inequality, especially in the case $b \gg a$.

Moreover, taking $x_1 = x_2 = \cdots = x_{n-2} = a,\ x_{n-1} = a(1-\epsilon),\ x_n = a(1+\epsilon)$; $a > 0,\ 0 < \epsilon < 1$; it can be seen that the constant $1/4$ is the best possible.

We shall consider now a class of well-known inequalities involving convex functions.

For a positive weight sequence $\mathbf{p} = \{p_i\}_1^n, \sum_1^n p_i = 1$ and a sequence $\mathbf{x} = \{x_i\}_1^n,\ a = x_1 \leq x_2 \leq \cdots \leq x_n = b$, the classical Jensen's inequality states that if f is convex on $I := [a, b]$, then

$$0 \leq \sum_1^n p_i f(x_i) - f\Big(\sum_1^n p_i x_i\Big), \tag{1.1}$$

with the equality sign only if all members of x are equal to a or b or if f is linear on I (cf [1], p. 70).

Jensen's inequality is one of the most known and extensively used inequalities in various fields of Mathematics. Some important inequalities are just particular cases of this inequality such as the weighted $A-G-H$ inequality, the Cauchy's inequality, the Ky Fan and Hölder inequalities, etc.

2. Results

One can see that the lower bound zero in (1.1) is of global nature; it depends only on f and I but does not depend on sequences $\mathbf{p}$ and $\mathbf{x}$.

This bound can be improved by no other means than the inequality (1.1) itself to the following

Theorem 2.1. *If f is convex on I, then*

$$\begin{aligned} &\max_{1\leq\mu<\nu\leq n}\Big[(p_\mu + p_\nu)\Big(\frac{p_\mu f(x_\mu) + p_\nu f(x_\nu)}{p_\mu + p_\nu} - f\Big(\frac{p_\mu x_\mu + p_\nu x_\nu}{p_\mu + p_\nu}\Big)\Big)\Big] \\ &\leq \sum_1^n p_i f(x_i) - f\Big(\sum_1^n p_i x_i\Big). \end{aligned} \tag{2.1}$$

For fixed μ, ν, equality sign holds for $n = 2$ or $x_i = \frac{p_\mu x_\mu + p_\nu x_\nu}{p_\mu + p_\nu}, i \neq \mu, \nu$.

Proof. Choose arbitrary $x_r, x_s \in \mathbf{x}, 1 \leq r < s \leq n$ with corresponding weights $p_r, p_s \in \mathbf{p}$. Note that, if $x_r, x_s \in I$, then also $\frac{p_r x_r + p_s x_s}{p_r + p_s} \in I$.

By (1.1), we get

$$f\Big(\sum_1^n p_i x_i\Big) = f\Big(\sum_{i\neq r,s} p_i x_i + (p_r+p_s)\Big(\frac{p_r x_r + p_s x_s}{p_r+p_s}\Big)\Big) \le \sum_{i\neq r,s} p_i f(x_i) + (p_r+p_s) f\Big(\frac{p_r x_r + p_s x_s}{p_r+p_s}\Big).$$

Hence

$$\sum_1^n p_i f(x_i) - f\Big(\sum_1^n p_i x_i\Big) \ge p_r f(x_r) + p_s f(x_s) - (p_r+p_s) f\Big(\frac{p_r x_r + p_s x_s}{p_r+p_s}\Big).$$

Since $x_r, x_s \in \mathbf{x}$ are arbitrary, the desired result follows.

It is obvious that the equality sign holds in (2.1) for $n=2$. The same is valid for $n>2$ and $x_i = \frac{p_\mu x_\mu + p_\nu x_\nu}{p_\mu + p_\nu}$, $i \neq \mu, \nu$. Indeed, in this case we get

$$\begin{aligned}\sum_1^n p_i f(x_i) - f\Big(\sum_1^n p_i x_i\Big) &= p_\mu f(x_\mu) + p_\nu f(x_\nu) + f\Big(\frac{p_\mu x_\mu + p_\nu x_\nu}{p_\mu+p_\nu}\Big)\times \\ &\times \sum_{i\neq\mu,\nu} p_i - f\Big(p_\mu x_\mu + p_\nu x_\nu + \Big(\frac{p_\mu x_\mu + p_\nu x_\nu}{p_\mu+p_\nu}\Big)\sum_{i\neq\mu,\nu} p_i\Big) \\ &= p_\mu f(x_\mu) + p_\nu f(x_\nu) + f\Big(\frac{p_\mu x_\mu + p_\nu x_\nu}{p_\mu+p_\nu}\Big)(1-p_\mu-p_\nu) \\ &- f\Big(p_\mu x_\mu + p_\nu x_\nu + \Big(\frac{p_\mu x_\mu + p_\nu x_\nu}{p_\mu+p_\nu}\Big)(1-p_\mu-p_\nu)\Big) \\ &= p_\mu f(x_\mu) + p_\nu f(x_\nu) - (p_\mu+p_\nu) f\Big(\frac{p_\mu x_\mu + p_\nu x_\nu}{p_\mu+p_\nu}\Big).\end{aligned}$$

In the case of uniform weights we obtain

$$\text{(2.2)} \qquad \frac{1}{n}\Big(f(a)+f(b)-2f\Big(\frac{a+b}{2}\Big)\Big) \le \frac{1}{n}\sum_1^n f(x_i) - f\Big(\frac{1}{n}\sum_1^n x_i\Big).$$

□

An interesting fact is that the expression $T_f(a,b) := f(a)+f(b)-2f(\frac{a+b}{2})$ represents also a global *upper* bound for Jensen's functional (cf [4, Theorem 1]) i.e. for any $\mathbf{p}$ and $\mathbf{x} \in [a,b]$, the inequality

$$\text{(2.3)} \qquad \sum_1^n p_i f(x_i) - f\Big(\sum_1^n p_i x_i\Big) \le T_f(a,b).$$

holds for any f which is convex over $[a,b]$.

Hence, merging the assertions (2.2) and (2.3) into one we obtain the following important conclusion.

Corollary 2.1. *For any sequence* $\mathbf{x} = \{x_i\}_1^n,\ a = x_1 \le x_2 \le \cdots \le x_n = b$ *and any* f *which is convex on* $[a, b]$*, we have*

$$\frac{1}{n}\Big[f(a) + f(b) - 2f\Big(\frac{a+b}{2}\Big)\Big] \le \frac{1}{n}\sum_1^n f(x_i) - f\Big(\frac{1}{n}\sum_1^n x_i\Big)$$

$$\le f(a) + f(b) - 2f\Big(\frac{a+b}{2}\Big). \tag{2.4}$$

The left-hand side of (8) is saturated for $n = 2$ or

$$x_1 = a, x_2 = x_3 = \ldots = x_{n-1} = \frac{a+b}{2}, x_n = b.$$

3. Applications

The above results can be useful in different parts of Analysis, Probability Theory, etc. As an illustration we give the following examples.

Example 3.1. For a sequence $\mathbf{x}$ of positive numbers, defined as above, denote by

$$A_n(\mathbf{x}) := \frac{1}{n}\sum_1^n x_i; \quad G_n(x) := \Big(\prod_1^n x_i\Big)^{\frac{1}{n}}$$

its arithmetic and geometric mean, respectively.

It is well known that

$$0 \le A_n(\mathbf{x}) - G_n(\mathbf{x}).$$

We can improve this classical inequality to the following one.

Theorem 3.1. *We have*

$$\frac{1}{n}(\sqrt{b} - \sqrt{a})^2 \le A_n(\mathbf{x}) - G_n(\mathbf{x}) \le (\sqrt{b} - \sqrt{a})^2,$$

where $a := \min_{1\le i\le n} x_i; \quad b := \max_{1\le i\le n} x_i$.

Proof. Applying Corollary 2.1 with $f(t) = e^t$, we get

$$\frac{1}{n}(e^a + e^b - 2e^{\frac{a+b}{2}}) \le \frac{1}{n}\sum_1^n e^{t_i} - e^{(\sum_1^n t_i)/n} \le e^a + e^b - 2e^{\frac{a+b}{2}}.$$

Changing variables $t_i = \log x_i,\ i = 1, 2, \cdots, n$, we obtain the desired result. □

Example 3.2. In the next example we shall give new bounds for Shannon's entropy $H(X)$ [2, 3], which is of utmost importance in Information Theory. Those bounds will be expressed as a combination of some classical means and are more precise than already existing ones.

Definition 3.1. If the probability distribution F is given by $P(X = i) = p_i,\ p_i > 0$, $i = 1, 2, \cdots, r;\ \sum_1^r p_i = 1$, then $H(X) := \sum_1^r p_i \log \frac{1}{p_i}$.

Theorem 3.2. *Let* $a := \min(p_i) < \max(p_i) := b,\ i = 1, 2, \cdots, r.$ *We have the following estimation*

$$2A(a,b)\log\Big(\frac{S(a,b)}{A(a,b)}\Big) \le \log r - H(X) \le 2\log\Big(\frac{A(a,b)}{G(a,b)}\Big), \tag{3.1}$$

where $A(a,b) := \frac{a+b}{2}$; $G(a,b) := \sqrt{ab}$; $S(a,b) := a^{\frac{a}{a+b}} b^{\frac{b}{a+b}}$, *are the arithmetic, geometric and Gini means, respectively.*

Proof. Indeed, applying inequalities (2.4) and (3.1) with $f(x) = -\log x,\ x_i = 1/p_i$, $i = 1, 2, \cdots, r$, we get

$$a\log a + b\log b + (a+b)\log\Big(\frac{2}{a+b}\Big) \le \log\Big(\sum_1^r p_i(1/p_i)\Big) - \sum_1^r p_i \log(1/p_i)$$
$$\le -\log(1/a) - \log(1/b) + 2\log\Big(\frac{1}{2}(1/a + 1/b)\Big),$$

which is equivalent to (3.1). □

Remark 3.1. It is interesting to compare (3.1) with [5], where the following result is stated

$$0 < \log r - H(X) \le \frac{(b-a)^2}{4ab}. \tag{3.2}$$

Since $\log(1+x) < x,\ x > 0$, putting $x = (b-a)^2/4ab$, it follows that the estimation (3.1) is much better than (3.2).

References

[1] G. H. Hardy, J. E. Littlewood, G. Polya, *Inequalities*, Cambridge University Press, Cambridge, 1978.

[2] I. Csiszar, J. Korner, *Information Theory*, Academic Press, New York, 1981.

[3] R. J. McEliece, *The Theory of Information and Coding*, Addison Wesley Publishing Company, Reading, 1977.

[4] S. Simic, *On a global upper bound for Jensen's inequality*, J. Math. Anal. and Appl., **313/1** (2008), 414-419 .

[5] S. S. Dragomir, C. J. Goh, *Some bounds on entropy measures in Information Theory*, Appl. Math. Letters, **10/3** (1997), 23-28.

[6] J. A. Adell, A. Lekuona, Y. Yu, *Sharp bounds on the entropy of the Poisson law and related quantities*, Information Theory, **56/5** (2010), 2299-2306.

[7] A. Sayyareh, *A new upper bound for Kullback-Leibler divergence*, Appl. Math. Sci., **5/65-68** (2011), 3303-3317.

Mathematical Institute SANU,
Kneza Mihaila 36,
11000 Belgrade, Serbia
E-mail address: ssimic@turing.mi.sanu.ac.rs