J. of Inequal. & Appl., 1997, Vol. 1, pp. 311–325
Reprints available directly from the publisher
Photocopying permitted by license only

Some Counterpart Inequalities for a Functional Associated with Jensen's Inequality

S.S. DRAGOMIR[a,*] and C.J. GOH[b]

[a] *Department of Mathematics, University of Transkei, Private Bag XI, Unitra, Umtata 5100, South Africa*
[b] *Department of Mathematics, University of Western Australia Nedlands, WA 6907, Australia*

(Received 1 November 1996)

We derive several inequalities for a functional connected with the well-known Jensen inequality on IR^n. Some applications for arithmetic and geometric means, and for the entropy mapping in information theory are also discussed.

Keywords: Jensen's inequality; entropy function; superadditivity; monotonicity; arithmetic and geometric means.

AMS Subject Classification: 26D15, 26D99.

1 INTRODUCTION

Let $f : C \to IR$ be a convex mapping defined on a convex set C in a linear space X. Define the functional:

$$\mathcal{F}(f, \mathbf{p}, I, \mathbf{x}) := \sum_{i\in I} p_i f(\mathbf{x}_i) - P_I \, f\left(\frac{1}{P_I}\sum_{i\in I} p_i \mathbf{x}_i\right) \tag{1.1}$$

*SSD acknowledges the financial support from the University of Western Australia during his visit when this work was completed.

where f is as above, $\mathbf{p} = (p_i)_{i \in IN}$ is a sequence of positive real numbers, $I \in \mathcal{P}_f(IN)$, i.e., I is a finite set of indices and $\mathbf{x} = (\mathbf{x}_i)_{i \in IN} \subset X$ is a sequence of vectors in X. It is easy to see that, with the above assumptions, one has:

(i) $\mathcal{F}(f, \mathbf{p}, I, \mathbf{x}) \geq 0$, i.e., *Jensen's discrete inequality*, and

(ii) $\mathcal{F}(\alpha f + \beta g, \mathbf{p}, I, \mathbf{x}) = \alpha \mathcal{F}(f, \mathbf{p}, I, \mathbf{x}) + \beta \mathcal{F}(g, \mathbf{p}, I, \mathbf{x}) \geq 0$ for all $\alpha, \beta \geq 0$ and f, g are convex mappings.

It is instructive to examine the properties of this functional with respect to the second and the third arguments. The following results (Theorem 1.1 and 1.3) were established in [1]:

THEOREM 1.1 *(Properties of $\mathcal{F}$ with respect to the second argument) Let $f : C \subseteq X \to IR$ be a convex mapping on the convex subset C of the linear space X and $\mathbf{x} = (\mathbf{x}_i)_{i \in IN} \subset C$. Then:*

(i) For all $\mathbf{p}, \mathbf{q} > \mathbf{0}$ one has the inequality:

$$\mathcal{F}(f, \mathbf{p} + \mathbf{q}, I, \mathbf{x}) \geq \mathcal{F}(f, \mathbf{p}, I, \mathbf{x}) + \mathcal{F}(f, \mathbf{q}, I, \mathbf{x}) \geq 0, \qquad (1.2)$$

where I is fixed in $\mathcal{P}_f(IN)$, i.e. the mapping $\mathcal{F}(f, \cdot, I, \mathbf{x})$ is superadditive.

(ii) For all $\mathbf{p} \geqq \mathbf{q} \geqq \mathbf{0}$ (i.e., each component of $\mathbf{p}$ is greater or equal to the corresponding component in $\mathbf{q}$, and each component of $\mathbf{q}$ is greater or equal to 0), one has the inequality;

$$\mathcal{F}(f, \mathbf{p}, I, \mathbf{x}) \geq \mathcal{F}(f, \mathbf{q}, I, \mathbf{x}) \geq 0 \qquad (1.3)$$

for a fixed $I \in \mathcal{P}_f(IN)$, i.e. $\mathcal{F}(f, \cdot, I, \mathbf{x})$ is monotonically non-decreasing in the second argument.

Consider the following subset of nonnegative sequences

$$\mathcal{P}_r(I) := \{\mathbf{p} = (p_i)_{i \in IN},\ p_i \geq 0,\ i \in IN \text{ and } \sum_{i \in I} p_i = 1\}$$

for a fixed I in $\mathcal{P}_f(IN)$. It is obvious that $\mathcal{P}_r(I)$ is a convex set as for $\mathbf{p}$, $\mathbf{q} \in \mathcal{P}_r(I)$ and $\alpha, \beta \geq 0$ and $\alpha + \beta = 1$ we have that $\alpha\mathbf{p} + \beta\mathbf{q} \in \mathcal{P}_r(I)$.

We shall first derive some convexity properties of the functional $\mathcal{F}$ with respect to the second argument. The following corollary to Theorem 1.1 holds:

COROLLARY 1.2 *The mapping $\mathcal{F}(f, \cdot, I, \mathbf{x})$ is concave on $\mathcal{P}_f(I)$ for every fixed I in $\mathcal{P}_f(IN) \setminus \emptyset$.*

Proof Let $\alpha, \beta \geq 0$ with $\alpha + \beta = 1$. By (1) we have that

$$\mathcal{F}(f, \alpha\mathbf{p} + \beta\mathbf{q}, I, \mathbf{x}) \geq \mathcal{F}(f, \alpha\mathbf{p}, I, \mathbf{x}) + \mathcal{F}(f, \beta\mathbf{q}, I, \mathbf{x}) \text{ for } \mathbf{p}, \mathbf{q} \in \mathcal{P}_r(I).$$

Since $\mathcal{F}(f, \cdot, I, \mathbf{x})$ is homogeneous, we have

$$\mathcal{F}(f, \alpha\mathbf{p}, I, \mathbf{x}) = \alpha\mathcal{F}(f, \mathbf{p}, I, \mathbf{x})$$

and

$$\mathcal{F}(f, \beta\mathbf{q}, I, \mathbf{x}) = \beta\mathcal{F}(f, \mathbf{q}, I, \mathbf{x}+$$

and the proof is complete. ∎

THEOREM 1.3 *(Properties of $\mathcal{F}$ with respect to the third argument) Let $f : C \subseteq X \to IR$ be a convex mapping on the convex subset C of the linear space X and $\mathbf{x} = (\mathbf{x}_i)_{i \in IN} \subset C$. Then:*
(i) For all $J, K \in \mathcal{P}_f(IN)$ with $J \cap K = \emptyset$, one has the inequality

$$\mathcal{F}(f, \mathbf{p}, J \cup K, \mathbf{x}) \geq \mathcal{F}(f, \mathbf{p}, J, \mathbf{x}) + \mathcal{F}(f, \mathbf{p}, K, \mathbf{x}) \geq 0, \tag{1.5}$$

where $\mathbf{p} > \mathbf{0}$ is fixed, i.e., the mapping $\mathcal{F}(f, \mathbf{p}, \cdot, \mathbf{x})$ is superadditive as an index set mapping.
(ii) For all $J \subseteq K$, $I, K \in \mathcal{P}_f(IN)$, one has the inequality

$$\mathcal{F}(f, \mathbf{p}, K, \mathbf{x}) \geq \mathcal{F}(f, \mathbf{p}, J, \mathbf{x}) \geq 0, \tag{1.6}$$

for a fixed $\mathbf{p} > \mathbf{0}$, i.e. $\mathcal{F}(f, \mathbf{p}, \cdot, \mathbf{x})$ is monotonically non-decreasing as an index set function.

In this paper, we shall derive several counterpart inequalities for the functional of (1.1) for the case where the mapping f is defined on an open convex subset C of the linear vector space IR^n and f is differentiable on C. In particular, we extend the results in Theorem 1.1 and Theorem 1.3 to come up with further counterpart inequalities with respect to the second and third arguments. This wil be taken up in the next two sections, where we will also discuss some applications of these results.

2 COUNTERPART INEQUALITIES WITH RESPECT TO THE SECOND ARGUMENT

In this section, we show that the inequalities of Theorem 1.1 can be further refined if the first argument of the functional $\mathcal{F}$ is a differentiable function.

THEOREM 2.1 *Let $f: C \subseteq X \to IR$ be a differentiable convex mapping on the open convex set C. Thus for all* $\mathbf{p}$, $\mathbf{q} > 0$ *one has the inequality:*

$$0 \leq \mathcal{F}(f, \mathbf{p}+\mathbf{q}, I, \mathbf{x}) - \mathcal{F}(f, \mathbf{p}, I, \mathbf{x}) - \mathcal{F}(f, \mathbf{q}, I, \mathbf{x})$$
$$\leq \frac{P_I Q_I}{P_I + Q_I}\left\langle \nabla f\left(\frac{1}{P_I}\sum_{i\in I} p_i \mathbf{x}_i\right) - \nabla f\left(\frac{1}{Q_I}\sum_{i\in I} q_i \mathbf{x}_i\right),\right.$$
$$\left.\frac{1}{P_I}\sum_{i\in I} p_i \mathbf{x}_i - \frac{1}{Q_I}\sum_{i\in I} q_i \mathbf{x}_i\right\rangle \tag{2.1}$$

where $I \in \mathcal{P}_f(IN)$ and $\mathbf{x} = (\mathbf{x}_i)_{i\in IN} \subset C$ are fixed. Equality holds in both inequalities if and only if $\mathbf{p} = \mathbf{q}$.

Proof Since the mapping $f: C \to IR$ is differentiable and convex in the open convex set C we have that:

$$f(\mathbf{x}) - f(\mathbf{y}) \geq \langle \nabla f(\mathbf{y}), \mathbf{x} - \mathbf{y}\rangle \tag{2.2}$$

where $\mathbf{x} = (x_1, \ldots, x_n)$, $\mathbf{y} = (y_1, \ldots, y_n) \in C$. Using the inequality (4) we can write for all $\mathbf{x}, \mathbf{y} \in C$ and $\alpha, \beta \geq 0$ with $\alpha + \beta > 0$, that:

$$f\left(\frac{\alpha\mathbf{x} + \beta\mathbf{y}}{\alpha+\beta}\right) - f(\mathbf{x}) \geq \left\langle \nabla f(\mathbf{x}), \frac{\alpha\mathbf{x}+\beta\mathbf{y}}{\alpha+\beta} - \mathbf{x}\right\rangle$$
$$= \frac{\beta}{\alpha+\beta}\langle \nabla f(\mathbf{x}), \mathbf{y} - \mathbf{x}\rangle \tag{2.3}$$

and

$$f\left(\frac{\alpha\mathbf{x} + \beta\mathbf{y}}{\alpha+\beta}\right) - f(\mathbf{y}) \geq \left\langle \nabla f(\mathbf{y}), \frac{\alpha\mathbf{x}+b\mathbf{y}}{\alpha+\beta} - \mathbf{y}\right\rangle$$
$$= -\frac{\alpha}{\alpha+\beta}\langle \nabla f(\mathbf{y}), \mathbf{y} - \mathbf{x}\rangle. \tag{2.4}$$

Now, if we multiply the inequality (2.3) by α and the inequality (2.4) by β and summing the obtained results, we obtain:

$$(\alpha+\beta) f\left(\frac{\alpha\mathbf{x} + \beta\mathbf{y}}{\alpha+\beta}\right) - \alpha f(\mathbf{x}) - \beta f(\mathbf{y})$$
$$\geq \frac{\alpha\beta}{\alpha+\beta}\left[\langle \nabla f(\mathbf{x}), \mathbf{y}-\mathbf{x}\rangle - \langle \nabla f(\mathbf{y}), \mathbf{y}-\mathbf{x}\rangle\right]$$
$$= \frac{\alpha\beta}{\alpha+\beta}\langle \nabla f(\mathbf{x}) - \nabla f(y), \mathbf{y}-\mathbf{x}\rangle$$

and thus we obtain

$$\begin{aligned} 0 &\le \alpha f(\mathbf{x}) + \beta f(\mathbf{y}) - (\alpha+\beta) f\left(\frac{\alpha\mathbf{x}+\beta\mathbf{y}}{\alpha+\beta}\right) \\ &\le \frac{\alpha\beta}{\alpha+\beta}\left\langle \nabla f(\mathbf{x}) - \nabla f(\mathbf{y}), \mathbf{x}-\mathbf{y}\right\rangle . \end{aligned} \tag{2.5}$$

Now, it is easier to see that:

$$\begin{aligned} 0 &\le \mathcal{F}(f, \mathbf{p}+\mathbf{q}, I, \mathbf{x}) - \mathcal{F}(f, \mathbf{p}, I, \mathbf{x}) - \mathcal{F}(f, \mathbf{q}, I, \mathbf{x}) \\ &= P_I f\left(\frac{1}{P_I}\sum_{i\in I} p_i\mathbf{x}_i\right) + Q_I f\left(\frac{1}{Q_I}\sum_{i\in I} q_i\mathbf{x}_i\right) \\ &\quad - (P_I+Q_I) f\left(\frac{P_I\cdot\frac{1}{P_I}\sum_{i\in I} p_i\mathbf{x}_i + Q_I\cdot\frac{1}{Q_I}\sum_{i\in I} q_i\mathbf{x}_i}{P_I+Q_I}\right) \\ &\le \frac{P_I Q_I}{P_I+Q_I}\left\langle \nabla f\left(\frac{1}{P_I}\sum_{i\in I} p_i\mathbf{x}_i\right) - f\left(\frac{1}{Q_I}\sum_{i\in I} q_i\mathbf{x}_i\right), \right. \\ &\quad \left. \frac{1}{P_I}\sum_{i\in I} p_i\mathbf{x}_i - \frac{1}{Q_I}\sum_{i\in I} q_i\mathbf{x}_i \right\rangle, \end{aligned}$$

where in the last inequality we have used (2.5) with the choices:

$$\alpha = P_I,\ \beta = Q_I,\ \mathbf{x} = \frac{1}{P_I}\sum_{k\in I} p_k\mathbf{x}_k,\ \mathbf{y} = \frac{1}{Q_I}\sum_{k\in I} q_k\mathbf{x}_k.$$

∎

The following corollary is immediately obvious.

COROLLARY 2.2 *Let f be as above. If $I \in \mathcal{P}_f(IN) \setminus \emptyset$ is fixed and $\mathbf{p}$, $\mathbf{q} \in \mathcal{P}_r(I)$, thus for all $t \in [0,1]$ we have the inequality:*

$$\begin{aligned} 0 &\le \mathcal{F}(f, t\mathbf{p}+(1-t)\mathbf{q}, I, \mathbf{x}) - t\mathcal{F}(f, \mathbf{p}, I, \mathbf{x}) - (1-t)\mathcal{F}(f, \mathbf{q}, I, \mathbf{x}) \\ &\le t(1-t)\left\langle \nabla f(\sum_{i\in I} p_i\mathbf{x}_i) - \nabla f(\sum_{i\in I} q_i\mathbf{x}_i)\ ,\ \sum_{i\in I} p_i\mathbf{x}_i - \sum_{i\in I} q_i\mathbf{x}_i \right\rangle. \end{aligned}$$

3 COUNTERPART INEQUALITIES WITH RESPECT TO THE THIRD ARGUMENT

Similarly, if the first argument of $\mathcal{F}$ is differentiable, then it is also possible to refine the result of Theorem 1.3 further.

THEOREM 3.1 *Let $f: C \subseteq IR^n \to IR$ be a differentiable convex mapping defined on the open convex set C. Then for all $J, K \in P_f(IN)$ with $J \cap K = \emptyset$, one has the inequality:*

$$0 \leq \mathcal{F}(f, \mathbf{p}, J \cup K, \mathbf{x}) - \mathcal{F}(f, \mathbf{p}, J, \mathbf{x}) - \mathcal{F}(f, \mathbf{p}, K, \mathbf{x})$$
$$\leq \frac{P_J P_K}{P_{J \cup K}} \left\langle \nabla f(\frac{1}{P_J} \sum_{i \in J} p_i \mathbf{x}_i) - \nabla f(\frac{1}{P_K} \sum_{k \in K} p_k \mathbf{x}_k) , \right.$$
$$\left. \frac{1}{P_J} \sum_{i \in J} p_i \mathbf{x}_i - \frac{1}{P_K} \sum_{k \in K} p_k \mathbf{x}_k \right\rangle \tag{3.1}$$

where $\mathbf{p} \geqq \mathbf{0}$ is fixed and $\mathbf{x} = (\mathbf{x}_i)_{i \in IN} \subset C$.

Proof It is easy to see that

$$0 \leq \mathcal{F}(f, \mathbf{p}, J \cup K, \mathbf{x}) - \mathcal{F}(f, \mathbf{p}, J, \mathbf{x}) - \mathcal{F}(f, \mathbf{p}, K, \mathbf{x})$$
$$= P_J f(\frac{1}{P_J} \sum_{i \in J} p_i \mathbf{x}_i) + P_K f(\frac{1}{P_K} \sum_{k \in K} p_k \mathbf{x}_k)$$
$$- (P_J + P_K) f \left[\frac{P_J \cdot \frac{1}{P_J} \sum_{i \in J} p_i \mathbf{x}_i + P_K \cdot \frac{1}{P_K} \sum_{k \in K} p_k \mathbf{x}_k}{P_J + P_K} \right]$$
$$\leq \frac{P_J P_K}{P_J + P_K} \left\langle \nabla f(\frac{1}{P_J} \sum_{i \in J} p_i \mathbf{x}_i) - \nabla f(\frac{1}{P_K} \sum_{k \in K} p_k \mathbf{x}_k), \right.$$
$$\left. \frac{1}{P_J} \sum_{i \in J} p_i \mathbf{x}_i - \frac{1}{P_K} \sum_{k \in K} p_k \mathbf{x}_k \right\rangle$$

and for the last inequality we used (2.4) with the choices:

$$\alpha = P_J, \beta = P_K$$

and

$$\mathbf{x} = \frac{1}{P_J} \sum_{i \in J} p_i \mathbf{x}_i \text{ and } \mathbf{y} = \frac{1}{P_K} \sum_{k \in K} p_k \mathbf{x}_k,$$

and the proof is complete. ∎

4 APPLICATIONS IN THE A-G MEAN INEQUALITY

We now present an application of the above results to the following well-known Arithmetic mean – Geometric mean inequality (A-G Mean inequality, for short):

$$A_n(\mathbf{p}, \mathbf{x}) \geq G_n(\mathbf{p}, \mathbf{x}), \tag{4.1}$$

where $\mathbf{x} = (x_1, x_2, \cdots, x_n) \geqq \mathbf{0}$, $\mathbf{p} = (p_1, p_2, \cdots, p_n) \geqq \mathbf{0}$ and

$$A_n(\mathbf{p}, \mathbf{x}) := \frac{1}{P_n} \sum_{i=1}^{n} p_i x_i$$

$$G_n(\mathbf{p}, \mathbf{x}) := \left(\prod_{i=1}^{n} x_i^{p_i} \right)^{\frac{1}{P_n}},$$

where $P_n = \sum_{i=1}^{n} p_i$. If $p_i \neq 0\ \forall i = 1, 2, \cdots, n$, it is well-known that the equality holds in (2.6) if and only if $x_1 = x_2 = \cdots = x_n$..

For $I \in \mathcal{P}_f(IN)$, let us consider

$$A_I(\mathbf{p}, \mathbf{x}) := \frac{1}{P_I} \sum_{i \in I} p_i x_i$$

$$G_I(\mathbf{p}, \mathbf{x}) := \left(\prod_{i \in I} x_i^{p_i} \right)^{\frac{1}{P_I}},$$

where $x_i, p_i \geq 0$, $\forall i \in I$. It is easy to see that for $f(\cdot) = -\ln(\cdot)$, we have:

$$\mathcal{F}(-\ln(\cdot), \mathbf{p}, I, \mathbf{x}) = P_I \ln \left[\frac{A_I(\mathbf{p}, \mathbf{x})}{G_I(\mathbf{p}, \mathbf{x})} \right] \geq 0,$$

which is a direct proof of (4.1) by Jensen inequality. A more sophisticated extension of the A-G inequality is as follows.

PROPOSITION 4.1 *With the above assumptions, we have, for* $\mathbf{p}, \mathbf{q} \geqq \mathbf{0}$*:*

$$1 \leq \left[\frac{A_I(\mathbf{p}+\mathbf{q}, \mathbf{x})}{G_I(\mathbf{p}+\mathbf{q}, \mathbf{x})} \right]^{P_I+Q_I} \left[\frac{G_I(\mathbf{p}, \mathbf{x})}{A_I(\mathbf{p}, \mathbf{x})} \right]^{P_I} \left[\frac{G_I(\mathbf{q}, \mathbf{x})}{A_I(\mathbf{q}, \mathbf{x})} \right]^{Q_I}$$
$$\leq \exp \left[\frac{P_I Q_I}{P_I + Q_I} \cdot \frac{(A_I(\mathbf{p}, \mathbf{x}) - A_I(\mathbf{q}, \mathbf{x}))^2}{A_I(\mathbf{p}, \mathbf{x}) A_I(\mathbf{q}, \mathbf{x})} \right],$$

where $P_I = \sum_{i \in I} p_i > 0$ *and* $Q_I = \sum_{i \in I} q_i > 0$.

Proof Applying Theorem 2.1 to the convex mapping $f(\cdot) = \ln(\cdot)$, we deduce that:

$$0 \le (P_I + Q_I)\ln\left[\frac{A_I(\mathbf{p}+\mathbf{q},\mathbf{x})}{G_I(\mathbf{p}+\mathbf{q},\mathbf{x})}\right] - P_I\ln\left[\frac{A_I(\mathbf{p},\mathbf{x})}{G_I(\mathbf{p},\mathbf{x})}\right] - Q_I\ln\left[\frac{A_I(\mathbf{q},\mathbf{x})}{G_I(\mathbf{q},\mathbf{x})}\right]$$
$$\le \frac{P_I Q_I}{P_I + Q_I}\left(-A_I^{-1}(\mathbf{p},\mathbf{x}) + A_I^{-1}(\mathbf{q},\mathbf{x})\right)\cdot(A_I(\mathbf{p},\mathbf{x}) - A_I(\mathbf{q},\mathbf{x}))$$
$$= \frac{P_I Q_I}{P_I + Q_I}\frac{(A_I(\mathbf{p},\mathbf{x}) - A_I(\mathbf{q},\mathbf{x}))^2}{A_I(\mathbf{p},\mathbf{x})A_I(\mathbf{q},\mathbf{x})}.$$

The conclusion of the Theorem is obtained by taking the exponential of each term in the above. ∎

Similarly, applying Theorem 3.1 to the convex mapping $-\ln(\cdot)$ yields the following result:

PROPOSITION 4.2 *With the above assumptions and for $J, K \in \mathcal{P}_f(IN)$, $J \cap K \ne \emptyset$, we have the following inequality:*

$$1 \le \left[\frac{A_{J\cup K}(\mathbf{p},\mathbf{x})}{G_{J\cup K}(\mathbf{p},\mathbf{x})}\right]^{P_{J\cup K}}\left[\frac{G_J(\mathbf{p},\mathbf{x})}{A_J(\mathbf{p},\mathbf{x})}\right]^{P_J}\left[\frac{G_J(\mathbf{p},\mathbf{x})}{A_J(\mathbf{p},\mathbf{x})}\right]^{P_K}$$
$$\le \exp\left[\frac{P_J P_K}{P_{J\cup K}}\cdot\frac{(A_J(\mathbf{p},\mathbf{x}) - A_K(\mathbf{p},\mathbf{x}))^2}{A_J(\mathbf{p},\mathbf{x})A_K(\mathbf{p},\mathbf{x})}\right],$$

where $P_J = \sum_{i\in J} p_i > 0$, $P_K = \sum_{i\in K} p_i > 0$ and $P_{J\cup K} = \sum_{i\in J\cup K} p_i$.

5 APPLICATIONS FOR THE EXPONENTIAL MAPPING

In the A-G mean inequality, we use the convex mapping $f(\cdot) = -\ln(\cdot)$. If we use the other well-known convex mapping $f(\cdot) = \exp(\cdot)$ instead, further inequalities on the arithmetic mean can be obtained directly from Theorem 2.1 and 3.1 as follows:

PROPOSITION 5.1
(i) For $\mathbf{p}, \mathbf{q} \geqq \mathbf{0}$,

$$0 \le (P_I + Q_I)\exp[A_I(\mathbf{p}+\mathbf{q},\mathbf{x})] - P_I\exp[A_I(\mathbf{p},\mathbf{x})] - Q_I\exp[A_I(\mathbf{q},\mathbf{x})]$$
$$\le \frac{P_I Q_I}{P_I + Q_I}(\exp[A_I(\mathbf{p},\mathbf{x})] - \exp[A_I(\mathbf{q},\mathbf{x})])(A_I(\mathbf{p},\mathbf{x}) - A_I(\mathbf{q},\mathbf{x})). \quad (5.1)$$

(ii) For $J, K \in \mathcal{P}_f(IN)$, $J \cap K \neq \emptyset$,

$$0 \leq (P_{J \cup K}) \exp[A_{J \cup K}(\mathbf{p}, \mathbf{x})] - P_J \exp[A_J(\mathbf{p}, \mathbf{x})] - P_K \exp[A_K(\mathbf{p}, \mathbf{x})]$$
$$\leq \frac{P_J Q_K}{P_{J \cup K}} (\exp[A_J(\mathbf{p}, \mathbf{x})] - \exp[A_K(\mathbf{p}, \mathbf{x})]) (A_J(\mathbf{p}, \mathbf{x}) - A_K(\mathbf{p}, \mathbf{x})) \tag{5.2}$$

The above inequalities (5.1) and (5.2) can be used further to derive similar inequalities for geometric means. One simply replaces all occurence of x_i by $\ln y_i$ in (5.1) and (5.2) to yield the following:

PROPOSITION 5.2
(i) For $\mathbf{p}, \mathbf{q} \geqq \mathbf{0}$, and $y_i \geq 0 \; \forall i \in I$,

$$0 \leq (P_I + Q_I) G_I(\mathbf{p} + \mathbf{q}, \mathbf{y}) - P_I G_I(\mathbf{p}, \mathbf{y}) - Q_I G_I(\mathbf{q}, \mathbf{y})$$
$$\leq \frac{P_I Q_I}{P_I + Q_I} (G_I(\mathbf{p}, \mathbf{y}) - G_I(\mathbf{q}, \mathbf{y}))(\ln G_I(\mathbf{p}, \mathbf{y}) - \ln G_I(\mathbf{q}, \mathbf{y}))\,. \tag{5.3}$$

(ii) For $J, K \in \mathcal{P}_f(IN)$, $J \cap K \neq \emptyset$,

$$0 \leq (P_{J \cup K}) G_{J \cup K}(\mathbf{p}, \mathbf{y}) - P_J G_J(\mathbf{p}, \mathbf{y}) - P_K G_K(\mathbf{p}, \mathbf{y})$$
$$\leq \frac{P_J Q_K}{P_{J \cup K}} (G_J(\mathbf{p}, \mathbf{y}) - G_K(\mathbf{p}, \mathbf{y}))(\ln G_J(\mathbf{p}, \mathbf{y}) - \ln G_K(\mathbf{p}, \mathbf{y}))\,. \tag{5.4}$$

Note that the first inequality of (5.3) suggests some kind of concavity property of the geometric mean function G_I with respect to the first variable $\mathbf{p}$, and the first inequality of (5.4) suggests some kind of concavity property of the geometric mean function G_I as an index set function in I.

6 FURTHER INEQUALITIES FOR THE GEOMETRIC MEAN

Consider the mapping $f : IR \to IR_+$ defined by

$$f(\cdot) := \ln(1 + \exp(\cdot)) \tag{6.1}$$
$$\text{with } f'(\cdot) = \frac{\exp(\cdot)}{1 + \exp(\cdot)}$$
$$\text{and } f''(\cdot) = \frac{\exp(\cdot)}{(1 + \exp(\cdot))^2} > 0.$$

Clearly f is convex in IR. Consider the following mapping:

$$\Gamma(\mathbf{p}, I, \mathbf{x}, \mathbf{y}) := \left[\frac{G_I(\mathbf{p}, \mathbf{x} + \mathbf{y})}{G_I(\mathbf{p}, \mathbf{x}) + G_I(\mathbf{p}, \mathbf{y})} \right]^{P_I} \tag{6.2}$$

where $\mathbf{x}, \mathbf{y} \geqq 0$.

THEOREM 6.1 *With the above assumptions, we have,*
(i) If $\mathbf{p}, \mathbf{q} \geqq \mathbf{0}$*, then*

$$\Gamma(\mathbf{p}+\mathbf{q}, I, \mathbf{x}, \mathbf{y}) \geq \Gamma(\mathbf{p}, I, \mathbf{x}, \mathbf{y})\Gamma(\mathbf{q}, I, \mathbf{x}, \mathbf{y}) \geq 1, \tag{6.3}$$

i.e., the mapping $\Gamma(\cdot, I, \mathbf{x}, \mathbf{y})$ *is supermultiplicative in the first argument.*
(ii) If $\mathbf{p} \geq \mathbf{q} \geq \mathbf{0}$*, then*

$$\Gamma(\mathbf{p}, I, \mathbf{x}, \mathbf{y}) \geq \Gamma(\mathbf{q}, I, \mathbf{x}, \mathbf{y}) \geq 1, \tag{6.4}$$

i.e., the mapping $\Gamma(\cdot, I, \mathbf{x}, \mathbf{y})$ *is monotone nondecreasing in the first argument.*

Proof Let the vector $\mathbf{z} = (z_i)_{i\in I}$ be such that $z_i = \ln \frac{x_i}{y_i}$ $\forall i \in I$. For convenience, let $\frac{\mathbf{x}}{\mathbf{y}}$ denotes the vector $\left(\frac{x_i}{y_i}\right)_{i\in I}$ and $\mathbf{z} = \ln \frac{\mathbf{x}}{\mathbf{y}}$. Then, using the convex mapping f as defined in (6.1), we have,

$$\begin{aligned}
0 &\leq \mathcal{F}(\ln(1+\exp(\cdot)), I, \mathbf{p}, \mathbf{z}) \\
&= \sum_{i\in I} p_i \ln\left(1+\exp\left(\ln\frac{x_i}{y_i}\right)\right) - P_I \ln\left(1+\exp\left(\frac{1}{P_I}\sum_{i\in I} p_i \ln\frac{x_i}{y_i}\right)\right) \\
&= \sum_{i\in I} p_i \ln\left(1+\frac{x_i}{y_i}\right) - P_I \ln\left(1+\exp\left[\ln\left(\prod_{i\in I}\left(\frac{x_i}{y_i}\right)^{p_i}\right)^{\frac{1}{P_I}}\right]\right) \\
&= \sum_{i\in I} p_i \ln(x_i+y_i) - \sum_{i\in I} p_i \ln y_i - P_I \ln\left(1+\frac{G_I(\mathbf{p},\mathbf{x})}{G_I(\mathbf{p},\mathbf{y})}\right) \\
&= \ln\prod_{i\in I}(x_i+y_i)^{p_i} - \sum_{i\in I} p_i \ln y_i - P_I \ln(G_I(\mathbf{p},\mathbf{x}) \\
&\quad + G_I(\mathbf{p},\mathbf{y})) + \ln(G_I(\mathbf{p},\mathbf{y}))^{P_I} \\
&= \ln\left[\frac{G_I(\mathbf{p},\mathbf{x}+\mathbf{y})}{G_I(\mathbf{p},\mathbf{x})+G_I(\mathbf{p},\mathbf{y})}\right]^{P_I} \\
&= \ln\Gamma(\mathbf{p}, I, \mathbf{x}, \mathbf{y}).
\end{aligned}$$

Using the first inequality of Theorem 3.1, we have

$$\begin{aligned}
\ln\Gamma(\mathbf{p}+\mathbf{q}, I, \mathbf{x}, \mathbf{y}) &= \mathcal{F}(\ln(1+\exp(\cdot)), I, \mathbf{p}+\mathbf{q}, \mathbf{z}) \\
&\geq \mathcal{F}(\ln(1+\exp(\cdot)), I, \mathbf{p}, \mathbf{z}) + \mathcal{F}(\ln(1+\exp(\cdot)), I, \mathbf{q}, \mathbf{z}) \\
&= \ln[\Gamma(\mathbf{p}, I, \mathbf{x}, \mathbf{y})\Gamma(\mathbf{q}, I, \mathbf{x}, \mathbf{y})],
\end{aligned}$$

from which (6.3) follows. Similarly, (6.4) follows from the direct application of the first inequality of Theorem 3.1. ∎

The following corollary follows from Theorem 6.1 and the fact that $\Gamma(t\mathbf{p}, I, \mathbf{x}, \mathbf{y}) = (\Gamma(\mathbf{p}, I, \mathbf{x}, \mathbf{y}))^t$.

COROLLARY 6.2 *With the above assumptions, we have:*

$$\Gamma(t\mathbf{p} + (1-t)\mathbf{q}, I, \mathbf{x}, \mathbf{y}) \geq [\Gamma(\mathbf{p}, I, \mathbf{x}, \mathbf{y})]^t [\Gamma(\mathbf{q}, I, \mathbf{x}, \mathbf{y})]^{1-t}$$

for all $t \in [0, 1]$, *i.e., the mapping* $\Gamma(\cdot, I, \mathbf{x}, \mathbf{y})$ *is log-concave in the first variable.*

THEOREM 6.3
(i) If $J, K \in \mathcal{P}_f(IN)$ *with* $J \cap K = \emptyset$, *then*

$$\Gamma(\mathbf{p}, J \cup K, \mathbf{x}, \mathbf{y}) \geq \Gamma(\mathbf{p}, J, \mathbf{x}, \mathbf{y})\Gamma(\mathbf{p}, K, \mathbf{x}, \mathbf{y}) \geq 1 \tag{6.5}$$

i.e., the mapping $\Gamma(\mathbf{p}, \cdot, \mathbf{x}, \mathbf{y})$ *is supermultiplicative as an index set mapping.*
(ii) If $J \subseteq K$, $j \neq \emptyset$, *then*

$$1 \leq \Gamma(\mathbf{p}, J, \mathbf{x}, \mathbf{y}) \leq \Gamma(\mathbf{p}, K, \mathbf{x}, \mathbf{y}) \tag{6.6}$$

(iii) We have,

$$\Gamma(\mathbf{p}, I, \mathbf{x}, \mathbf{y}) = \sup_{\substack{J \subseteq I \\ J \neq \emptyset}} \Gamma(\mathbf{p}, J, \mathbf{x}, \mathbf{y}) \geq 1. \tag{6.7}$$

(iv) We have,

$$\Gamma(\mathbf{p}, I, \mathbf{x}, \mathbf{y}) \geq \max_{i,j \in I} \left\{ \frac{\left[(x_i + y_i)^{p_i}(x_j + y_j)^{p_j}\right]^{\frac{1}{p_i + p_j}}}{\left[x_i^{p_i} x_j^{p_j}\right]^{\frac{1}{p_i + p_j}} \left[y_i^{p_i} y_j^{p_j}\right]^{\frac{1}{p_i + p_j}}} \right\}^{p_i + p_j} \geq 1. \tag{6.8}$$

Proof The proofs of the above follow directly from Theorem 3.1, the details of which are omitted. ∎

THEOREM 6.4 *Let* $\mathbf{p}, \mathbf{q} \geqq \mathbf{0}$, *then we have:*

$$1 \le \frac{\Gamma(\mathbf{p}+\mathbf{q}, I, \mathbf{x}, \mathbf{y})}{\Gamma(\mathbf{p}, I, \mathbf{x}, \mathbf{y})\Gamma(\mathbf{q}, I, \mathbf{x}, \mathbf{y})} \le \left\{\frac{G_I(\mathbf{p}, \mathbf{x})\ G_I(\mathbf{q}, \mathbf{y})}{G_I(\mathbf{p}, \mathbf{y})\ G_I(\mathbf{q}, \mathbf{x})}\right\}^{\theta} \tag{6.9}$$

where

$$\theta = \frac{P_I Q_I}{P_I + Q_I}\left[\frac{G_I(\mathbf{p}, \mathbf{x})G_I(\mathbf{q}, \mathbf{y}) - G_I(\mathbf{q}, \mathbf{x})G_I(\mathbf{p}, \mathbf{y})}{(G_I(\mathbf{p}, \mathbf{x}) + G_I(\mathbf{p}, \mathbf{y}))(G_I(\mathbf{q}, \mathbf{x}) + G_I(\mathbf{q}, \mathbf{y}))}\right]$$

Proof The first inequality is just (6.3). To derive the second inequality, we use the convex mapping $f(\cdot) = \ln(1 + \exp(\cdot))$ in the second inequality of Theorem 2.1 to get

$$\begin{aligned} 0 &\le \mathcal{F}(f, \mathbf{p}+\mathbf{q}, I, \mathbf{z}) - \mathcal{F}(f, \mathbf{p}, I, \mathbf{z}) - \mathcal{F}(f, \mathbf{q}, I, \mathbf{z}) \\ &\le \frac{P_I Q_I}{P_I + Q_I}\left[f'(A_I(\mathbf{p}, \mathbf{z})) - f'(A_I(\mathbf{q}, \mathbf{z}))\right]\left[A_I(\mathbf{p}, \mathbf{z}) - A_I(\mathbf{q}, \mathbf{z})\right] \end{aligned} \tag{6.10}$$

Using the fact that

$$f'(\cdot) = \frac{\exp(\cdot)}{1 + \exp(\cdot)}$$

we have,

$$\begin{aligned} &f'(A_I(\mathbf{p}, \mathbf{z})) - f'(A_I(\mathbf{q}, \mathbf{z})) \\ &\quad = \frac{\exp(A_I(\mathbf{p}, \mathbf{z}))}{1 + \exp(A_I(\mathbf{p}, \mathbf{z}))} - \frac{\exp(A_I(\mathbf{q}, \mathbf{z}))}{1 + \exp(A_I(\mathbf{q}, \mathbf{z}))} \\ &\quad = \frac{G_I\left(\mathbf{p}, \frac{\mathbf{x}}{\mathbf{y}}\right)}{1 + G_I\left(\mathbf{p}, \frac{\mathbf{x}}{\mathbf{y}}\right)} - \frac{G_I\left(\mathbf{q}, \frac{\mathbf{x}}{\mathbf{y}}\right)}{1 + G_I\left(\mathbf{q}, \frac{\mathbf{x}}{\mathbf{y}}\right)} \\ &\quad = \frac{G_I(\mathbf{p}, \mathbf{x})}{G_I(\mathbf{p}, \mathbf{x}) + G_I(\mathbf{p}, \mathbf{y})} - \frac{G_I(\mathbf{q}, \mathbf{x})}{G_I(\mathbf{q}, \mathbf{x}) + G_I(\mathbf{q}, \mathbf{y})} \end{aligned}$$

Thus, by (6.10) we deduce that

$$\begin{aligned} 0 &\le \ln\Gamma(\mathbf{p}+\mathbf{q}, I, \mathbf{x}, \mathbf{y}) - \ln\Gamma(\mathbf{p}, I, \mathbf{x}, \mathbf{y}) - \ln\Gamma(\mathbf{q}, I, \mathbf{x}, \mathbf{y}) \\ &\le \frac{P_I Q_I}{P_I + Q_I}\left[\frac{G_I(\mathbf{p}, \mathbf{x})G_I(\mathbf{q}, \mathbf{y}) - G_I(\mathbf{q}, \mathbf{x})G_I(\mathbf{p}, \mathbf{y})}{(G_I(\mathbf{p}, \mathbf{x}) + G_I(\mathbf{p}, \mathbf{y}))(G_I(\mathbf{q}, \mathbf{x}) + G_I(\mathbf{q}, \mathbf{y}))}\right] \\ &\quad \times \left[\ln\frac{G_I(\mathbf{p}, \mathbf{x})}{G_I(\mathbf{p}, \mathbf{y})} - \ln\frac{G_I(\mathbf{q}, \mathbf{x})}{G_I(\mathbf{q}, \mathbf{y})}\right] \end{aligned}$$

which is equivalent to

$$0 \leq \ln \left[\frac{\Gamma(\mathbf{p}+\mathbf{q}, I, \mathbf{x}, \mathbf{y})}{\Gamma(\mathbf{p}, I, \mathbf{x}, \mathbf{y})\Gamma(\mathbf{q}, I, \mathbf{x}, \mathbf{y})} \right]$$
$$\leq \ln \left\{ \frac{G_I(\mathbf{p}, \mathbf{x})\, G_I(\mathbf{q}, \mathbf{y})}{G_I(\mathbf{p}, \mathbf{y})\, G_I(\mathbf{q}, \mathbf{x})} \right\}^{\mu}$$

where

$$\mu = \frac{P_I Q_I}{P_I + Q_I} \left[\frac{G_I(\mathbf{p}, \mathbf{x}) G_I(\mathbf{q}, \mathbf{y}) - G_I(\mathbf{q}, \mathbf{x}) G_I(\mathbf{p}, \mathbf{y})}{(G_I(\mathbf{p}, \mathbf{x}) + G_I(\mathbf{p}, \mathbf{y}))(G_I(\mathbf{q}, \mathbf{x}) + G_I(\mathbf{q}, \mathbf{y}))} \right].$$

The conclusion follows from taking the exponential of each term of the above. ■

Similarly, application of the second inequality of Theorem 3.2 leads to the following result, which we shall merely state without proof.

THEOREM 6.5 *Let $K, J \in \mathcal{P}_f(IN)$ with $K \cap J = \emptyset$ and $K, J \neq \emptyset$, then,*

$$1 \leq \frac{\Gamma(\mathbf{p}, J \cup K, \mathbf{x}, \mathbf{y})}{\Gamma(\mathbf{p}, J, \mathbf{x}, \mathbf{y})\Gamma(\beta, K, \mathbf{x}, \mathbf{y})}$$
$$\leq \left\{ \frac{G_J(\mathbf{p}, \mathbf{x})\, G_K(\mathbf{p}, \mathbf{y})}{G_J(\mathbf{p}, \mathbf{y})\, G_K(\mathbf{p}, \mathbf{x})} \right\}^{\lambda} \tag{6.11}$$

where

$$\lambda = \frac{P_J P_K}{P_{J \cup K}} \left[\frac{G_J(\mathbf{p}, \mathbf{x}) G_K(\mathbf{p}, \mathbf{y}) - G_K(\mathbf{p}, \mathbf{x}) G_J(\mathbf{p}, \mathbf{y})}{(G_J(\mathbf{p}, \mathbf{x}) + G_J(\mathbf{p}, \mathbf{y}))(G_K(\mathbf{p}, \mathbf{x}) + G_K(\mathbf{p}, \mathbf{y}))} \right].$$

7 APPLICATIONS IN INFORMATION THEORY

Another application of Theorem 3.1 can be found in Information Theory. Suppose X is a discrete random variable having range $R = \{x_i, i \in \mathcal{I}\}$ and having a probability distribution $\{0 < p_i = Pr(X = x_i),\ i \in \mathcal{I}\}$. Let $\mathbf{p}$ be the probability vector corresponding to the probability distribution of X. The b-Entropy of the random variable X is defined by [2]:

$$H_b(X) = H_b(\mathbf{p}) := \sum_{i \in \mathcal{I}} p_i \log_b \frac{1}{p_i}. \tag{7.1}$$

Several inequalities for the entropy function can be established merely by applying the Jensen ineqaulity, the following is one of them:

$$0 \le H_b(X) \le \log_b |\mathcal{I}|.$$

Now let's say we are interested in the entropy of two sub-probability vectors of $\mathbf{p}$ (upon appropriate normalization) and we wish to relate this to the entropy of the original probability vector as given in (7.1). For some index subset $J, K \subseteq \mathcal{I}$, $J \cap K = \emptyset$, $J \cup K = \mathcal{I}$, $J, K \neq \emptyset$, we define the new random variables X_J and X_K having range in $R_J := \{x_i,\ i \in J\}$ and $S_K := \{x_i, i \in K\}$ and respective probability distributions

$$\{p_i^J := \frac{p_i}{P_J} > 0,\ i \in J\},\ \text{and}\ \{p_i^K := \frac{p_i}{P_K} > 0,\ i \in K\},$$

where $P_J := \sum_{j\in J} p_j > 0$ and $P_K := \sum_{j\in K} p_j > 0$. Let $\mathbf{p}_J = \{p_i^J,\ j \in J\}$ and $\mathbf{p}_K = \{p_i^K,\ j \in K\}$ be the probability vectors corresponding to the probability distribution of X_J and X_K respectively. The entropies of the two sub-probability vectors are defined in the usual manner:

$$H_b(X_J) := \sum_{i\in J} p_i^J \log_b \frac{1}{p_i^J}$$

$$H_b(X_K) := \sum_{i\in K} p_i^K \log_b \frac{1}{p_i^K}.$$

THEOREM 7.1 *With the above assumptions, we have*

$$0 \le \log_b |\mathcal{I}| - H_b(X) - P_J[\log_b |J| - H_b(X_J)] - P_K[\log_b |K| - H_b(X_K)]$$
$$\le \frac{P_J P_K}{\ln b}\left(\gamma - \frac{1}{\gamma}\right)^2,$$

where

$$\gamma = \sqrt{\frac{|J| P_K}{|K| P_J}}.$$

Consequently, if $\gamma = 1$ *or*

$$\frac{|J|}{P_J} = \frac{|K|}{P_K}$$

then

$$\log_b |\mathcal{I}| - H_b(X) = P_J[\log_b |J| - H_b(X_J)] + P_K[\log_b |K| - H_b(X_K)].$$

Proof In Theorem 1.3, let $f(\cdot) = -\log_b(\cdot)$ which is convex with $\nabla f(\cdot) = -\frac{1}{\ln b\,(\cdot)}$; and let $\mathbf{x}_i = \frac{1}{p_i}$. Then,

$$
\begin{aligned}
&\mathcal{F}(f, \mathbf{p}, \mathcal{I}, \mathbf{x}) - \mathcal{F}(f, \mathbf{p}, J, \mathbf{x}) - \mathcal{F}(f, \mathbf{p}, K, \mathbf{x}) \\
&\quad = -\sum_{i\in\mathcal{I}} p_i \log_b p_i + \log_b\left(\sum_{i\in J\cup K} \frac{p_i}{p_i}\right) + P_J \sum_{i\in J} \frac{p_i}{P_J}\log_b p_i \\
&\quad - P_J \log_b\left(\frac{1}{P_J}\sum_{i\in J}\frac{p_i}{p_i}\right) + P_K \sum_{i\in K}\frac{p_i}{P_K}\log_b p_i - P_K \log_b\left(\frac{1}{P_K}\sum_{i\in K}\frac{p_i}{p_i}\right) \\
&\quad = \log_b |J\cup K| - H_b(X) - P_J[\log_b |J| - H_b(X_J)] \\
&\qquad - P_K[\log_b |K| - H_b(X_K)],
\end{aligned} \tag{7.2}
$$

and

$$
\begin{aligned}
&\frac{P_J P_K}{\ln b\; P_{\mathcal{I}}}\left\langle \nabla f\Big(\frac{1}{P_J}\sum_{i\in J} p_i \mathbf{x}_i\Big) - \nabla f\Big(\frac{1}{P_K}\sum_{k\in K} p_k \mathbf{x}_k\Big)\,,\right. \\
&\qquad \left. \frac{1}{P_J}\sum_{i\in J} p_i \mathbf{x}_i - \frac{1}{P_K}\sum_{k\in K} p_k \mathbf{x}_k \right\rangle \\
&\quad = \frac{P_J P_K}{\ln b}\left(-\frac{P_J}{|J|} + \frac{P_K}{|K|}\right) \times \left(\frac{|J|}{P_J} - \frac{|K|}{P_K}\right) \\
&\quad = \frac{P_J P_K}{\ln b}\left(\frac{|J|P_K}{|K|P_J} + \frac{|K|P_J}{|J|P_K} - 2\right) \\
&\quad = \frac{P_J P_K}{\ln b}\left(\gamma - \frac{1}{\gamma}\right)^2
\end{aligned} \tag{7.3}
$$

The conclusion thus follows from replacing the terms in Theorem 3.1 by (7.2) and (7.3). ∎

References

[1] S.S. Dragomir, J. Pečarić and L.E. Persson, Properties of some functionals related to Jensen's inequality, *Acta Math. Hungarica*, **60** (1996), 129–143.

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