Internat. J. Math. & Math. Sci.
Vol. 22, No. 3 (1999) 543–546
S 0161-1712⟨99⟩22543-4

INEQUALITIES VIA CONVEX FUNCTIONS

I. A. ABOU-TAIR and W. T. SULAIMAN

(Received 25 May 1998)

ABSTRACT. A general inequality is proved using the definition of convex functions. Many major inequalities are deduced as applications.

Keywords and phrases. Convex functions, inequalities.

1991 Mathematics Subject Classification. 26A51, 26D15.

1. Introduction. Kapur and Kumer (1986) have used the principle of dynamical programming to prove major inequalities due to Shannon, Renyi, and Hölder. See [1]. In this note, we prove a general inequality using convex functions. As a result, the inequalities of Shannon, Renyi, Hölder, and others are all deduced.

Let I be an interval in $\mathbb{R}$, $f : I \to \mathbb{R}$ is said to be convex if and only if, for all x, $y \in I$, all λ, $0 \le \lambda \le 1$,

$$f[\lambda x + (1-\lambda)y] \le \lambda f(x) + (1-\lambda)y. \tag{1}$$

Here, we give the following new definitions:

(a) Let f and g be two functions and let I be an interval in $\mathbb{R}$ for which $f \circ g$ is defined, then f is said to be g-convex if and only if, for all x, $y \in I$, all λ, $0 \le \lambda \le 1$,

$$f[\lambda g(x) + (1-\lambda)g(y)] \le \lambda f \circ g(x) + (1-\lambda) f \circ g(y). \tag{2}$$

(b) If the inequality is reversed, then f is said to be g-concave.

If $g(x) = x$, the two definitions of g-convex and convex functions become identical.

THEOREM 1.1. *Let f be g-convex, then*
(i) *if g is linear, then $f \circ g$ is convex, and*
(ii) *if f is increasing and g is convex, then $f \circ g$ is convex.*

PROOF.
(i)

$$\begin{aligned} f \circ g[\lambda x + (1-\lambda)y] &= f[\lambda g(x) + (1-\lambda)g(y)] \\ &\le \lambda f \circ g(x) + (1-\lambda) f \circ g(y). \end{aligned} \tag{3}$$

(ii)

$$\begin{aligned} f \circ g[\lambda x + (1-\lambda)y] &\le f[\lambda g(x) + (1-\lambda)g(y)] \\ &\le \lambda f \circ g(x) + (1-\lambda) f \circ g(y). \end{aligned} \tag{4}$$

□

Lemma 1.1. *Let f be g-convex and let $\sum_{i=1}^{n} t_i = T_n = 1$, $t_i \ge 0$, $i = 1,2,\dots,n$, then*

$$f\left(\sum_{i=1}^{n} t_i g(x_i)\right) \le \sum_{i=1}^{n} t_i f \circ g(x_i). \tag{5}$$

Proof.

$$\begin{aligned} f\left(\sum_{i=1}^{n} t_i g(x_i)\right) &= f\left(T_{n-1}\sum_{i=1}^{n-1}\frac{t_i}{T_{n-1}}g(x_i) + t_n g(x_n)\right) \\ &\le T_{n-1} f\left(\sum_{i=1}^{n-1}\frac{t_i}{T_{n-1}}g(x_i)\right) + t_n f\circ g(x_n) \\ &= T_{n-2} f\left(\frac{T_{n-2}}{T_{n-1}}\sum_{i=1}^{n-2}\frac{t_i}{T_{n-2}}g(x_i) + \frac{t_{n-1}}{T_{n-1}}g(x_{n-1})\right) + t_n f\circ g(x_n) \\ &\le T_{n-2} f\left(\sum_{i=1}^{n-2}\frac{t_i}{T_{n-2}}g(x_i)\right) + t_{n-1} f\circ g(x_{n-1}) + t_n f\circ g(x_n) \\ &\vdots \\ &\le \sum_{i=1}^{n} t_i f\circ g(x_i). \end{aligned} \tag{6}$$

□

Lemma 1.2. *For any function g, the exponential function $f(x) = e^x$ is g-convex.*

Proof. Define

$$F(x) = \lambda e^{g(x)} + (1-\lambda)e^{g(y)} - e^{\lambda g(x)+(1-\lambda)g(y)}. \tag{7}$$

Let

$$G(t) = (1-\lambda) + \lambda t - t^{\lambda}, \quad t > 0. \tag{8}$$

It follows that

$$G'(t) = \lambda(1-t^{\lambda-1}), \qquad G''(t) = \lambda(1-\lambda)t^{\lambda-2}. \tag{9}$$

Thus, $G'(t) = 0$ when $t = 1$ and $G''(1) = \lambda(1-\lambda) > 0$. Hence, G has its minimum value 0 at $t = 1$ and this implies $G(t) \ge 0$, $t > 0$. The result follows by putting $F(x) = e^{g(y)}G(e^{g(x)-g(y)})$. □

Corollary 1.3. *The function $f(x) = \ln(x)$ is concave for if $h(x) = e^x$, then, by Lemma 1.2, h is f-convex. Hence,*

$$e^{\lambda(\ln x)+(1-\lambda)\ln y} \le \lambda e^{\ln x} + (1-\lambda)e^{\ln y} = \lambda x + (1-\lambda y). \tag{10}$$

It follows that

$$\lambda \ln x + (1-\lambda)\ln y \le \ln[\lambda x + (1-\lambda)y]. \tag{11}$$

2. Main inequality

Theorem 2.1.

$$\sum_{j=1}^{n}\prod_{i=1}^{m}(p_{ij})^{q_i/\sum_{i=1}^{m} q_i} \le \frac{\sum_{i=1}^{m}\sum_{j=1}^{n} p_{ij}q_i}{\sum_{i=1}^{m} q_i}. \tag{12}$$

Proof. If $f(x) = e^x$ and $g(x) = \ln x$, then f is g-convex. By Lemma 1.2, we have

$$\begin{aligned}\prod_{i=1}^{m}(p_{ij})^{q_i/\sum_{i=1}^{m}q_i} &= e^{\ln\left(\prod_{i=1}^{m}(p_{ij})^{q_i/\sum_{i=1}^{m}q_i}\right)} \\ &= e^{\sum_{i=1}^{m}\ln(p_{ij})^{q_i/\sum_{i=1}^{m}q_i}} = e^{\sum_{i=1}^{m}(q_i/\sum_{i=1}^{m}q_i)\ln p_{ij}} \\ &\le \sum_{i=1}^{m}\left(\frac{q_i}{\sum_{i=1}^{m}q_i}\right)e^{\ln p_{ij}} = \frac{\sum_{i=1}^{m}q_ip_{ij}}{\sum_{i=1}^{m}q_i}.\end{aligned} \tag{13}$$

Therefore,

$$\sum_{j=1}^{n}\prod_{i=1}^{m}(p_{ij})^{q_i/\sum_{i=1}^{m}q_i} \le \frac{\sum_{j=1}^{n}\sum_{i=1}^{m}p_{ij}q_i}{\sum_{i=1}^{m}q_i} = \frac{\sum_{i=1}^{m}\sum_{j=1}^{n}p_{ij}q_i}{\sum_{i=1}^{m}q_i}. \tag{14}$$

□

3. Applications

Theorem 3.1 (Shannon's inequality). *Given* $\sum_{i=1}^{m}a_i = a$, $\sum_{i=1}^{m}b_i = b$, *then*

$$a\ln\left(\frac{a}{b}\right) \le \sum_{i=1}^{m}a_i\ln\left(\frac{a_i}{b_i}\right), \quad a_i, b_i \ge 0. \tag{15}$$

Proof. Applying Theorem 2.1 by putting

$$p_{ij} = \frac{b_i}{a_i}, \quad j = 1, \qquad q_i = a_i, \qquad \sum_{i=1}^{m}a_i = a, \qquad \sum_{i=1}^{m}b_i = b, \tag{16}$$

we have

$$\prod_{i=1}^{m}\left(\frac{b_i}{a_i}\right)^{a_i/\sum_{i=1}^{m}a_i} \le \frac{\sum_{i=1}^{m}b_i}{\sum_{i=1}^{m}a_i}. \tag{17}$$

That is

$$\prod_{i=1}^{m}\left(\frac{b_i}{a_i}\right)^{a_i/a} \le \frac{b}{a}. \tag{18}$$

It follows that

$$\frac{a}{b} \le \prod_{i=1}^{m}\left(\frac{a_i}{b_i}\right)^{a_i/a}. \tag{19}$$

Hence, we get

$$a\ln\left(\frac{a}{b}\right) \le \sum_{i=1}^{m}a_i\ln\left(\frac{a_i}{b_i}\right). \tag{20}$$

□

Theorem 3.2 (Renyi's inequality). *Given* $\sum_{i=1}^{m}a_i = a$, $\sum_{i=1}^{m}b_i = b$, *then, for* $\alpha > 0$, $\alpha \ne 1$,

$$\frac{1}{\alpha-1}(a^{\alpha}b^{1-\alpha}-a) \le \sum_{i=1}^{m}\frac{1}{\alpha-1}(a_i^{\alpha}b_i^{1-\alpha}-a_i), \quad a_i, b_i \ge 0. \tag{21}$$

Proof. Applying Theorem 2.1 with $i = 2$, $p_{1j} = c_j$, $p_{2j} = d_j$, $q_1 = \lambda$, $q_2 = 1-\lambda$, $0 < \lambda < 1$, we have

$$\sum_{j=1}^{m}c_j^{\lambda}d_j^{1-\lambda} \le \sum_{j=1}^{m}(\lambda c_j + (1-\lambda)d_j). \tag{22}$$

On putting $c_j = (a_j/\sum_{j=1}^m a_j)$ and $d_j = (b_j/\sum_{j=1}^m b_j)$, inequality (22) implies

$$\sum_{j=1}^m a_j^\lambda b_j^{1-\lambda} \le \left(\sum_{j=1}^m a_j\right)^\lambda \left(\sum_{j=1}^m b_j\right)^{1-\lambda}, \tag{23}$$

and this gives

$$\frac{a^\lambda b^{1-\lambda}}{\lambda - 1} \le \frac{1}{\lambda - 1}\sum_{j=1}^m a_j^\lambda b_j^{1-\lambda}. \tag{24}$$

Thus, for the case $0 < \alpha < 1$, the theorem follows from inequality (24) by setting $\lambda = \alpha$. Now, inequality (23) implies

$$\left(\sum_{j=1}^m a_j^\lambda b_j^{1-\lambda}\right)^{1/\lambda} \left(\sum_{j=1}^m b_j\right)^{1-1/\lambda} \le \sum_{j=1}^m a_j. \tag{25}$$

Let $a_j^\lambda b_j^{1-\lambda} = e_j$, $\lambda = 1/\alpha$, then inequality (25) gives

$$\frac{1}{\alpha - 1}\left(\sum_{j=1}^m e_j\right)^\alpha \left(\sum_{j=1}^m b_j\right)^{1-\alpha} \le \frac{1}{\alpha - 1}\sum_{j=1}^m e_j^\alpha b_j^{1-\alpha}. \tag{26}$$

This completes the proof of the theorem. □

Theorem 3.3 (Generalization of Hölder's inequality)**.**

$$\sum_{j=1}^n \prod_{i=1}^m (p_{ij})^{q_i} \le \prod_{i=1}^m \left(\sum_{j=1}^n p_{ij}\right)^{q_i}, \qquad \sum_{i=1}^m q_i = 1. \tag{27}$$

Proof. Applying Theorem 2.1 with $p_{ij}/\sum_{j=1}^n p_{ij}$ instead of p_{ij}, we get

$$\sum_{j=1}^n \prod_{i=1}^m \left(\frac{p_{ij}}{\sum_{j=1}^n p_{ij}}\right)^{q_i} \le \sum_{i=1}^m \left(\sum_{j=1}^n \left(\frac{p_{ij}}{\sum_{j=1}^n p_{ij}}\right)\right) q_i = \sum_{i=1}^m q_i = 1, \tag{28}$$

which implies

$$\sum_{j=1}^n \prod_{i=1}^m (p_{ij})^{q_i} \le \prod_{i=1}^m \left(\sum_{j=1}^n p_{ij}\right)^{q_i}. \tag{29}$$

□

Theorem 3.4 (Arithmetic-Geometric-Mean inequality)**.**

$$\left(\prod_{i=1}^m x_i\right)^{1/m} \le \frac{1}{m}\sum_{i=1}^m x_i. \tag{30}$$

Proof. Applying Theorem 2.1, with $j = 1$, $p_{ij} = x_i$, $q_i = 1$. □

References

[1] J. N. Kapur, V. Kumar, and U. Kumar, *A measure of mutual divergence among a number of probabilty distributions*, Internat. J. Math. Math. Sci. **10** (1987), no. 3, 597–607. MR 89d:94030. Zbl 641.94006.

Abou-Tair: Zarka Private University, Zarka, Jordan

Sulaiman: P.O. Box 120054, Doha, Qatar