JOURNAL OF APPROXIMATION THEORY 32, 160–166 (1981)

A Companion Inequality to Jensen's Inequality

MORTON L. SLATER

Department of Mathematics, Texas Christian University, Fort Worth, Texas 76129

Communicated by R. Bojanic

Received February 10, 1980

1. INTRODUCTION

In this paper we consider a convex function $f: (a, b) \to R$, where $-\infty \leqslant a < b \leqslant \infty$. We recall that f is *convex* if for all $x_1, x_2 \in (a, b)$ and $\alpha_1, \alpha_2 \geqslant 0$ such that $\alpha_1 + \alpha_2 = 1$,

$$f(\alpha_1 x_1 + \alpha_2 x_2) \leqslant \alpha_1 f(x_1) + \alpha_2 f(x_2).$$

f is *strictly convex* if, in addition, whenever $\alpha_1 x_1 + \alpha_2 x_2$ is strictly between x_1 and x_2,

$$f(\alpha_1 x_1 + \alpha_2 x_2) < \alpha_1 f(x_1) + \alpha_2 f(x_2).$$

If f is convex, it is continuous; its left and right derivatives, f'_- and f'_+, exist, are finite, and are non-decreasing; $f'_- \leqslant f'_+$, and except for at most countably many $x \in (a, b)$, $f'_-(x) = f'(x) = f'_+(x)$.

Jensen's inequality can be stated as follows:

Suppose that f is convex on (a, b). Then, for $x_1,..., x_n$ in (a, b) and $p_1,..., p_n \geqslant 0$, $p_1 + \cdots + p_n > 0$,

$$f\left(\frac{p_1 x_1 + \cdots + p_n x_n}{p_1 + \cdots + p_n}\right) \leqslant \frac{p_1 f(x_1) + \cdots + p_n f(x_n)}{p_1 + \cdots + p_n}. \tag{1}$$

The inequality in the title of the paper states that, under the additional assumption of monotonicity of f, there is a specific point in (a, b) at which the value of f is greater than or equal to the right-hand side of (1).

THEOREM 1. *Suppose that f is convex and increasing on (a, b). Then for $x_1,..., x_n \in (a, b)$, $p_1,..., p_n \geqslant 0$, $p_1 + \cdots + p_n > 0$, and $p_1 f'_+(x_1) + \cdots + p_n f'_+(x_n) > 0$, we have*

$$\frac{p_1 f(x_1) + \cdots + p_n f(x_n)}{p_1 + \cdots + p_n} \leqslant f\left(\frac{p_1 f'_+(x_1) x_1 + \cdots + p_n f'_+(x_n) x_n}{p_1 f'_+(x_1) + \cdots + p_n f'_+(x_n)}\right). \tag{2}$$

0021-9045/81/060160–07$02.00/0

A more general version of Jensen's inequality can be stated as follows:

Suppose that f is convex on (a, b), that $(E, \mathscr{E}, \mu)$ is a probability measure space, and that $X\colon (E, \mathscr{E}) \to (a, b)$ is measurable. If X and $f \circ X$ are in $L(\mu)$, then

$$f\left(\int_E X\, d\mu\right) \leqslant \int_E (f \circ X)\, d\mu. \tag{3}$$

The corresponding generalization of Theorem 1 is

THEOREM 2. *Suppose that f is convex and increasing on (a, b). Suppose also that $X\colon (E, \mathscr{E}) \to (a, b)$ is measurable, that $f \circ X$, $f'_+ \circ X$, and $(f'_+ \circ X)\, X$ are in $L(\mu)$ and that $\int_E (f'_+ \circ X)\, d\mu > 0$. Then*

$$\int_E (f \circ X)\, d\mu \leqslant f\left(\frac{\int_E (f'_+ \circ X)\, X\, d\mu}{\int_E (f'_+ \circ X)\, d\mu}\right). \tag{4}$$

If, in addition we assume that f is strictly convex, then equality holds in (4) *if and only if X is constant μ a.e.*

Both theorems remain valid if at any occurrence of $f'_+(x)$ we write instead any value in the interval $[f'_-(x), f'_+(x)]$. Note merely that inequalities (5) and (6) below, on which the proofs depend, continue to hold with such replacement. ("f'_+" was chosen for ease in stating the theorems.) Furthermore, both theorems are also true if f is a convex and decreasing function; the only changes needed in the hypotheses are to require $p_1 f'_+(x_1) + \cdots + p_n f'_+(x_n) < 0$ in Theorem 1, and $\int_E f'_+ \circ X\, d\mu < 0$ in Theorem 2. For if $f\colon (a, b) \to R$ is a convex and decreasing function, then $\tilde{f}\colon (-b, -a) \to R$, defined by $\tilde{f}(x) = f(-x)$, is a convex and increasing function, and Theorems 1 and 2 applied to $\tilde{f}$ yield the same theorems with f, just as before.

In Section 2, inequalities (2) and (4) will be established and the case of equality will be discussed. Section 3 contains several applications. Section 4 contains concluding remarks.

2. The Proofs of the Companion Inequalities

The proofs of (2) and (4) are based on the inequality

$$f(y) \geqslant f(x) + (y - x)\, f'_+(x), \tag{5}$$

valid for any convex function on (a, b) and arbitrary $x, y \in (a, b)$. The proof that (in the case of strict convexity) equality holds in (4) only if X is constant μ a.e. depends on the inequality

$$f(y) > f(x) + (y-x) f'_+(x), \tag{6}$$

valid for any strictly convex function on (a, b) and arbitrary distinct $x, y \in (a, b)$.

To simplify typesetting, set

$$A = \frac{p_1 f'_+(x_1) x_1 + \cdots + p_n f'_+(x_n) x_n}{p_1 f'_+(x_1) + \cdots + p_n f'_+(x_n)}.$$

To prove Theorem 1, observe that $A \in (a, b)$ since A is a convex combination of $x_1, \ldots, x_n$, and so by (5),

$$f(A) \geqslant f(x_k) + (A - x_k) f'_+(x_k), \qquad k = 1, \ldots, n.$$

Multiply the kth inequality by p_k and add the inequalities thus obtained; we find

$$(p_1 + \cdots + p_n) f(A) \geqslant \sum_{k=1}^{n} p_k f(x_k) + A \sum_{k=1}^{n} p_k f'_+(x_k) - \sum_{k=1}^{n} p_k f'_+(x_k) x_k,$$

and Theorem 1 is proved since

$$A \sum_{k=1}^{n} p_k f'_+(x_k) - \sum_{k=1}^{n} p_k f'_+(x_k) x_k = 0.$$

We begin the proof of Theorem 2 by establishing inequality (4). The proof is similar to that of (2). Set

$$A = \frac{\int_E (f'_+ \circ X) X \, d\mu}{\int_E (f'_+ \circ X) \, d\mu}.$$

$A \in (a, b)$, and by (5), for every $t \in E$,

$$f(A) \geqslant (f'_+ \circ X)(t) + (A - X(t))(f'_+ \circ X)(t).$$

Integrate this inequality with respect to μ and observe that $A \int_E (f'_+ \circ X) \, d\mu - \int_E (f'_+ \circ X) X \, d\mu = 0$; (4) is immediate.

To complete the proof of Theorem 2, we consider the case of equality in (4). If X is constant μ a.e., then (4) is an obvious equality. We now show that if we require f to be *strictly* convex, then for equality to hold in (4), it is also necessary that X be constant μ a.e.

Since f is increasing and convexity is strict on (a, b), have $f'_+ > 0$, and $Z: E \to R$ as defined by the equation

$$\int_E f \circ X \, d\mu = (f \circ X)(t) + (f'_+ \circ X)(t)(Z(t) - X(t)), \tag{7}$$

is measurable on E.

From (3) and (4), and the continuity and strict monotonicity of f, we see that there is a unique $x_0 \in (a, b)$ such that $f(x_0) = \int_E f \circ X \, d\mu$. We rewrite (7) as

$$Z(t) - x_0 = \frac{f(x_0) - ((f \circ X)(t) + (x_0 - X(t))(f'_+ \circ X)(t))}{(f'_+ \circ X)(t)}. \tag{8}$$

From (6), and the fact that $f'_+ > 0$, we see that for all $t \in E$, $X(t) \neq x_0$ if and only if $Z(t) > x_0$, while $X(t) = x_0$ if and only if $Z(t) = x_0$.

Integrate both sides of (7) with respect to μ and divide by $\int_E f'_+ \circ X \, d\mu$. We find

$$A = \frac{\int_E (f'_+ \circ X) \, Z \, d\mu}{\int_E (f'_+ \circ X) \, d\mu}, \tag{9}$$

where A is as defined above. If X is not constant μ a.e., then $E_1 = \{t \in E: X(t) \neq x_0\} = \{t \in E: Z(t) > x_0\}$ has positive measure. Since $E \backslash E_1 = \{t \in E: Z(t) = x_0\}$, we see from (9) that $A > x_0$, and so $\int_E f \circ X \, d\mu = f(x_0) < f(A)$; thus the inequality in (4) is strict and the proof of Theorem 2 is complete.

The following example illustrates one of the possibilities for equality in (4) (more specifically, in (2)) when convexity is not strict. (A sketch will make the construction clear.)

Suppose $f_i(x) = m_i x + b_i$, $i = 1, 2$ and $x \in R$, where $0 < m_1 < m_2$. Define $f = \operatorname{Sup}(f_1, f_2)$ and let $\bar{x}$, $\bar{y}$ be defined by $f_1(\bar{x}) = f_2(\bar{x}) = \bar{y}$. Suppose x_1 and x_2 are fixed points such that $x_1 < \bar{x} < x_2$. $\bar{y}$ is in the open interval $(f(x_1), f(x_2))$ and so there exist $\alpha_1, \alpha_2 > 0$ such that $\alpha_1 + \alpha_2 = 1$ and $\alpha_1 f(x_1) + \alpha_2 f(x_2) = \bar{y}$. An easy calculation shows that

$$\bar{x} = \frac{\alpha_1 f'(x_1) \, x_1 + \alpha_2 f'(x_2) \, x_2}{\alpha_1 f'(x_1) + \alpha_2 f'(x_2)}.$$

($f'(x_1) = m_1$ and $f'(x_2) = m_2$.) Thus

$$\alpha_1 f(x_1) + \alpha_2 f(x_2) = \bar{y} = f(\bar{x}) = f\left(\frac{\alpha_1 f'(x_1) \, x_1 + \alpha_2 f'(x_2) \, x_2}{\alpha_1 f'(x_1) + \alpha_2 f'(x_2)}\right).$$

3. Applications

In this section we give two applications, one of inequality (4) (which yields a familiar result), and one of inequality (2) (which might be new). Since all integrals are over E, we will drop "E" from the integral sign.

1°. Define $f: R \to R$ by $f(x) = 0$ for $x < 0$ and $f(x) = x^p$, $p > 1$, for $x \geqslant 0$. If $X: E \to R$ is non-negative, bounded, and measurable, and not a null function, then from (4)

$$\int X^p \, d\mu \leqslant \left(\frac{\int X^p \, d\mu}{\int X^{p-1} \, d\mu} \right)^p ; \tag{10}$$

after rearranging we obtain

$$\left(\int X^{p-1} \, d\mu \right)^{1/(p-1)} \leqslant \left(\int X^p \, d\mu \right)^{1/p} . \tag{11}$$

By familiar approximation techniques of real analysis, notably the Lebesgue monotone convergence theorem, (11) can easily be shown to hold for all non-negative extended real valued measurable X. With $q > 0$, replace X by X^q and raise both sides of (11) to the $1/q$. If we now set $q = t - s$ and $p = t/(t - s)$, where $0 < s < t$, we recover the well known result

$$\left(\int X^s \, d\mu \right)^{1/s} \leqslant \left(\int X^t \, d\mu \right)^{1/t} , \qquad \text{if} \quad 0 < s < t. \tag{12}$$

It is also well known that (12) holds if $s < t < 0$; this too can be obtained from (11) by first replacing X by $1/X$, and then proceeding in a similar fashion.

2°. Suppose $g: \{z: |z| < R\} \to C$ is analytic. Let $M(r) = \operatorname{Max} \{|g(z)|: |z| = r\}$ for $0 < r < R$. From the maximum modulus theorem and the Hadamard three circles theorem we are able to conclude that $\log M(r)$ is a convex increasing function of $\log r$. That is, the composite function $\log \circ M \circ \exp$ is convex and increasing on $(0, R)$. Since log and exp have strictly positive derivatives, we conclude that the one-sided derivatives M'_+ and M'_- exist and satisfy $M'_- \leqslant M'_+$ and that they are equal except for at most countably many $r \in (0, R)$.

Take any $r_1, r_2 \in (0, R)$ and $\alpha_1, \alpha_2 \geqslant 0$ such that $\alpha_1 + \alpha_2 = 1$. Hadamard's inequality states

$$M(r_1^{\alpha_1} r_2^{\alpha_2}) \leqslant (M(r_1))^{\alpha_1} (M(r_2))^{\alpha_2} ; \tag{13}$$

the companion inequality, which will be proved below, states

$$(M(r_1))^{\alpha_1}(M(r_2))^{\alpha_2} \leqslant M(r_1^{\beta_1} r_2^{\beta_2}), \tag{14}$$

where

$$\beta_i = \frac{\alpha_i M'_+(r_i)\, r_i/M(r_i)}{\alpha_1 M'_+(r_1)\, r_1/M(r_1) + \alpha_2 M'_+(r_2)\, r_2/M(r_2)}, \qquad i = 1, 2.$$

We begin the proof by setting $u_1 = \log r_1$, $u_2 = \log r_2$, and with $f\colon(-\infty, \log R) \to R$ defined by $f(u) = \log M(e^u)$, we apply inequality (2):

$$\begin{aligned} &\alpha_1 \log M(e^{u_1}) + \alpha_2 \log M(e^{u_2}) \qquad (16)\\ &\quad\leqslant \log M\left(\exp\left\{\frac{\alpha_1[M'_+(e^{u_1})/M(e^{u_1})]\, e^{u_1}u_1 + \alpha_2[M'_+(e^{u_2})/M(e^{u_2})]\, e^{u_2}u_2}{\alpha_1[M'_+(e^{u_1})/M(e^{u_1})]\, e^{u_1} + \alpha_2[M'_+(e^{u_2})/M(e^{u_2})]\, e^{u_2}}\right\}\right). \end{aligned}$$

If we now apply exp to both sides of (16) and replace u_1 and u_2 by $\log r_1$ and $\log r_2$ respectively, we find

$$\begin{aligned} &(M(r_1))^{\alpha_1}(M(r_2))^{\alpha_2}\\ &\quad\leqslant M\left(\exp\left\{\frac{\alpha_1[M'_+(r_1)/M(r_1)]\, r_1 \log r_1 + \alpha_2[M'_+(r_2)/M(r_2)]\, r_2 \log r_2}{\alpha_1[M'_+(r_1)/M(r_1)]\, r_1 + \alpha_2[M'_+(r_2)/M(r_2)]\, r_2}\right\}\right)\\ &\quad= M(\exp(\beta_1 \log r_1 + \beta_2 \log r_2))\\ &\quad= M(r_1^{\beta_1} r_2^{\beta_2}), \text{ and the proof is complete.} \end{aligned}$$

4. Concluding Remarks

Hardy, Littlewood, and Polyà's "Inequalities," is still an excellent source; all our "well known" results are to be found there.

For a discussion of the differentiability of $M(r)$, Otto Blumenthal, "Über ganze transzendente Funktionen," Jahresbericht d. Deutschen Mathem.-Vereinigung, XVI, Heft 2, 1905, seems to be the best we could find. $M'_+(r) > M'_-(r)$ *can* occur. We remark that in (14), any M'_+ may be replaced by M'_- (or any value inbetween) and the inequality remains valid. (In this connection, compare the remarks following the statement of Theorem 2.)

Besides the elementary convex monotone functions (e.g., $f(x) = e^x$, $f(x) = x(\log_+ x)^k$, $k \geqslant 1$) which often provide useful companion inequalities, there are other functions from analytic function theory whose companion inequalities may be worth investigating. For example, the Nevanlinna characteristic $T(r, f)$ of a meromorphic f (see Einar Hille, "Analytic Function Theory," Vol. II) is known to be an increasing convex function of $\log r$. Its differentiability properties were recently investigated by D. W. Townsend (Abstracts A. M. S. Vol. 1, No. 1, January 1980, Abstract 773-30-12).

Printed in Belgium