Journal of the Nigerian Association of Mathematical Physics
Volume **14 (May, 2009), pp 283 – 288**

A derivative and integral characterization of real-valued convex functions of single variable through the geometric chord property

[1]P. E. Ezimadu and [2]J. N. Igabari
Department of Mathematics and Computer Science,
Delta State University, Abraka, Nigeria.

Abstract

Although convex functions have been characterized using the derivative, integral and monotonicity of the derivative, a characterization which involves a combination of these concepts has not been achieved. This is the centre of this work. In particular we show that for functions enjoying the geometric chord property, this characterization gives equivalence for the definitions of convexity.

Keywords
Real-valued Function, Non-decreasing Function, Convex Function.

1.0 Introduction

A search through calculus text reveals that there are essentially three approaches used in defining convex functions: *A function f defined over an open interval I is said to be convex if*

i. f' *exists and the graph of* f *lies on or above the tangent at* $(x, f(x))$ *for each* $x \in I$, [10];

ii. $f'(x_1) \le f'(x_2) \quad \forall\, x_1 < x_2$, [7];

iii. $f''(x) \ge 0$, [3].

However, the geometric chord property requires neither the existence of *f´* nor *f˝* on *I* and as such will apply to a wider class of functions including non-smooth functions. We shall see how this is equivalent to those definitions more commonly used.

Perhaps it is a surprise that this simple geometric chord property forces *f* to be continuous, and both the right- and left-hand derivatives to exist and be non-decreasing. Moreover, for functions enjoying this chord property, the characterization in Sction 7 (Theorem 7.1) which is not mentioned in calculus texts can be used to easily show the equivalence, under the appropriate existence hypotheses, of the definitions of convexity.

2.0 Convexity

In this work we consider the function $f: I \to \mathbb{R}$, where I is an open interval of the real line $\mathbb{R}$. f is said to be convex if for each pair of points $x_1, x_2 \in I$, $x_1 < x < x_2$, the graph of f lies on or below the chord joining the points$(x_1, f(x_1))$ and $(x_2, (f(x_2))$.

We recall that for any two points x_1, x_2 the gradient m is given by

$$m = \frac{f(x_2) - f(x_1)}{x_2 - x_1} \ , \qquad (2.1)$$

[1]**Corresponding author:**
[1]**Telephone: 08028831349**

so that for any $x \in [x_1, x_2]$ the equation of the chord joining x_1 and x_2 is

$$y = m(x - x_1) + f(x_1)\,, \tag{2.2}$$

and by convexity

$$f(x) \leq m(x - x_1\,) + f(x_1) \tag{2.3}$$

$$\Longrightarrow f(x)(x_2 - x_1) \leq [f(x_2) - f(x_1)](x - x_1) + f(x_1)(x_2 - x_1)\,. \tag{2.4}$$

We observe that $x = \lambda x_1 + (1-\lambda)x_2,\ \lambda \in [0,1]$, so that (4.4) becomes

$$f(\lambda x_1 + (1-\lambda)x_2) \leq \lambda f(x_1) + (1-\lambda) f(x_2) \tag{2.5}$$

Thus we have shown that:

***Lemma* 2.1** *For any real-valued function f defined on a closed and bounded interval I*

i) *f is convex if and only if*

$$f(\lambda x_1 + (1-\lambda)x_2) \leq \lambda f(x_1) + (1-\lambda) f(x_2) \quad \forall\, x_1, x_2 \in I, \qquad \lambda \in [0,1].$$

ii) *f is strictly convex if, and only if,*

$$f(\lambda x_1 + (1-\lambda)x_2) \leq \lambda f(x_1) + (1-\lambda) f(x_2) \quad \forall\, x_1, x_2 \in I,\ \ x_1 \neq x_2,\ \ \lambda \in [0,1].$$

An important generalization of this Lemma is Jensen's Inequality.

3.0 Jensen's inequality

Lemma 2.1 is easily extended to convex combinations of more than two points: *If f is convex, $x_i \in I$ and $\lambda_i \geq 0,\ i = 1 \dots n$, with $\sum_{i=1}^{n} \lambda_i = 1$, then*

$$f\left(\sum_{i=1}^{n} \lambda_i x_i\right) \leq \sum_{i=1}^{n} \lambda_i f(x_i)$$

This is known as Jensen's Inequality.

***Theorem* 3.1**

$f: I \to \mathbb{R}$ is convex if and only if f satisfies Jensen's Inequality

A proof of this result is found in [1]. A very important application of the above result is that the geometric mean of n positive numbers does not exceed their arithmetic mean.

4.0 The geometric chord property of convex functions

From [2, 6, 8, 9, 11], we have the following results which will be very useful in the proof of the result in Section 7.

***Lemma* 4.1**

$f: I \to \mathbb{R}$ is convex if and only if for any $x_1 < x_2 < x_3$, we have

$$\frac{f(x_2) - f(x_1)}{x_2 - x_1} \leq \frac{f(x_3) - f(x_1)}{x_3 - x_1} \leq \frac{f(x_3) - f(x_2)}{x_3 - x_2}. \tag{4.1}$$

f is strictly convex if and only if the inequalities are strict.

Proof

Choose any $x_1 < x_2 < x_3$ and define $\lambda = \frac{x_2 - x_1}{x_3 - x_1}$, then $1 - \lambda = \frac{x_3 - x_2}{x_3 - x_1}$ and $x_2 = \lambda x_3 + (1 - \lambda) x_1$. Thus

$$f(x_2) \leq \lambda f(x_3) + (1-\lambda) f(x_1)$$

is equivalent to

$$(1-\lambda)\big(f(x_2) - f(x_1)\big) \leq \lambda\big(f(x_3) - f(x_2)\big)$$

Substituting for λ and multiplying by $x_3 - x_1$ yields the equivalent statement

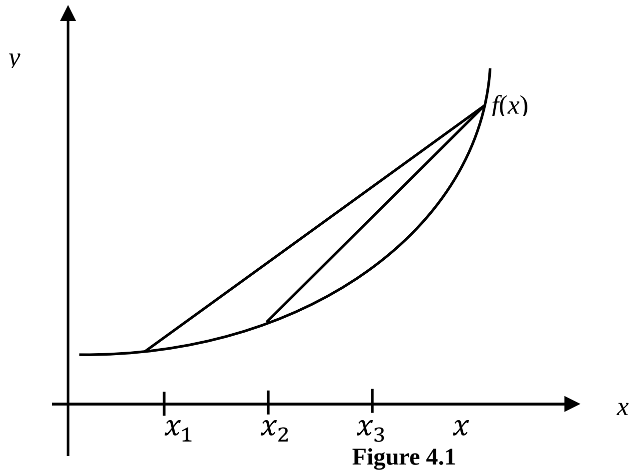

Figure 4.1

$$(x_3 - x_2)\big(f(x_2) - f(x_1)\big) \le (x_2 - x_1)\big(f(x_3) - f(x_2)\big) \qquad (4.2)$$

Adding $(x_2 - x_1)\big(f(x_2) - f(x_1)\big)$ to each side of (4.2) and simplifying we obtain

$$(x_3 - x_1)\big(f(x_2) - f(x_1)\big) \le (x_2 - x_1)\big(f(x_3) - f(x_1)\big) \qquad (4.4)$$

which is equivalent to

$$\frac{f(x_2)-f(x_1)}{x_2-x_1} \le \frac{f(x_3)-f(x_1)}{x_3-x_1} \qquad (4.5)$$

Similarly adding $(x_3 - x_2)\big(f(x_3) - f(x_2)\big)$ to each side of (4.1) we have that

$$(x_3 - x_2)\big(f(x_3) - f(x_1)\big) \le (x_3 - x_1)\big(f(x_3) - f(x_1)\big)$$

and therefore the equivalent statement

$$\frac{f(x_3)-f(x_1)}{x_3-x_1} \le \frac{f(x_3)-f(x_2)}{x_3-x_2}.$$

The equivalence of each statement also holds if each inequality is strict. The lemma above (Lemma 4.1) is known as the geometric chord property.

***Lemma* 4.2**

Suppose $f: I \to \mathbb{R}$ *is convex. Then* $x \in int\ (I)$ *implies* $f'_-(x)$ *and* $f'_+(x)$ *exist and* $f'_-(x) \le f'_+(x)$.

Proof

Note that Lemma 4.1 implies that $\frac{f(x_3)-f(x_1)}{x_3-x_1}$ is nondecreasing in x_1 and x_3 for $x_1 \ne x_3$. Therefore for all $x_1 < x_2 < x_3$, we have

$$f'_+(x_2) \equiv \lim_{x_3 \to x_2} \frac{f(x_3)-f(x_2)}{x_3-x_2} \ge \frac{f(x_2)-f(x_1)}{x_2-x_1} \qquad (\quad 4.6$$

and

$$f'_+(x_2) \ge \lim_{x_1 \to x_2} \frac{f(x_2)-f(x_1)}{x_2-x_1} \equiv f'_-(x_2). \qquad 4.7$$

5.0 Boundedness

***Theorem* 5.1**

A convex function $f: I \to \mathbb{R}$ *is bounded over a closed subinterval of* I.

Proof

The convexity of f over $[x_1, x_2]$ implies

$$f(x) \le \max\{f(x_1),\ f(x_2)\} = M\ \forall\, x \in [x_1, x_2] \qquad 5.1 \qquad (12)$$

Thus f is bounded above.

Let $\bar{x} = \frac{1}{2}(x_1 + x_2)$ so that $x = \bar{x} + \lambda$. By convexity

$$f(\bar{x}) \le \frac{1}{2} f(\bar{x} - \lambda) + \frac{1}{2} f(\bar{x} + \lambda)$$

since $\bar{x} \in\ [\bar{x} - \lambda,\ \bar{x} + \lambda]$

$$\Rightarrow f(x) \ge 2f(\bar{x}) - f(\bar{x} - \lambda) \qquad 5.2 \qquad (13)$$

But $f \le M\ \forall\, x$, so that (4.6) becomes

$$f(x) \ge f(\bar{x}) - M = m$$

Hence f is bounded.

6.0 Continuity

***Theorem* 6.1**

If $f: I \to \mathbb{R}$ is convex over the open interval I, then f is continuous on I.

Proof

Let M and m be respectively upper and lower bounds for f over $[\alpha, \beta] \subset I$. For $\varepsilon > 0$ with $[\alpha - \varepsilon, \beta + \varepsilon] \subset I$ and arbitrary $x_1, x_2 \in [\alpha, \beta]$, let $\bar{x} = x_2 + \varepsilon \frac{x_2 - x_1}{|x_2 - x_1|}$

Note that for both $x_1 > x_2$ and $x_2 > x_1$, $\bar{x} \in [\alpha - \varepsilon, \beta + \varepsilon]$. If we let $y = \frac{x_2 - x_1}{\varepsilon + |x_2 - x_1|}$, we have that $x = (1 - y)x_1 + yx_2$

$$\Rightarrow f(x_2) \le (1 - y)f(x_1) + yf(\bar{x}) = f(x_1) + y[f(\bar{x}) - f(x_1)]$$

$$\Rightarrow f(x_2) - f(x_1) \le y(M - m) < c|x_2 - x_1| \qquad 6.1$$

where $c = (M - m)/\varepsilon$. Also

$$f(x_1) - f(x_2) < c|x_1 - x_2| \qquad 6.2 \qquad (15)$$

From (6.1) and (6.2)

$$|f(x_2) - f(x_1)| < c|x_2 - x_1|$$

Set $\delta = \frac{\varepsilon}{c}$. If $|x_2 - x_1| < \delta$ then $|f(x_2) - f(x_1)| < \varepsilon$. Thus f is continuous. Thus we have shown that f is locally Lipschitz continuous on I and as such is absolutely continuous over I. This is a stronger condition.

7.0 Derivative and integral characterization of convex function using the geometric chord property

In [5] convex functions of several variables were characterized using monotone mapping and the Hessian matrix. Using similar approach a characterization of these functions through the derivative and the integral was achieved in [4]. This is very insightful in the next result. [4] and [5] suggest that since convex functions of several variables can be thus characterized, a version of these results are also possible for real-valued convex functions of single variable as shown in Theorem 7.1 below.

***Theorem* 7.1**

Let $f: I \to \mathbb{R}$ be differentiable over the open interval $I \subseteq \mathbb{R}$ then the following statements are equivalent.

(*i*) $f(x_2) \le \lambda f(x_1) + (1 - \lambda) f(x_3),\ \ \lambda \in [0,1],\ \ x_2 = \lambda x_1 + (1 - \lambda)x_3,\ x_1, x_2 x_3 \in I$

(*ii*) $f'(x_1) \le f'(x_2),\ \ \forall\, x_1 < x_2\ \in I$

(*iii*) $f(x) - f(x_0) = \int_{x_0}^{x} f'(t)dt \quad x, x_0 \in I$

(*iv*) $f(x) \ge f'(x_0)(x - x_0) + f(x_0) \quad x, x_0 \in I$

Proof

We show that $(i) \Rightarrow (ii) \Rightarrow (iii) \Rightarrow (iv) \Rightarrow (i)$ and we are done

$(i) \Rightarrow (ii)$

Let $x_1, x_2, x_3 \in I$ with $x_1 < x_2 < x_3$. By the geometric chord property

$$\frac{f(x_2)-f(x_1)}{x_2-x_1} \le \frac{f(x_3)-f(x_1)}{x_3-x_1} \le \frac{f(x_3)-f(x_2)}{x_3-x_2} \qquad (7.1)$$

Considering the first inequality; as $x \to x_1$ from the right the slopes decrease. This implies that the Newton quotients $\frac{f(x)-f(x_1)}{x-x_1}$ used to compute $f_+^{'}(x_1)$ are increasing. Also the second inequality reveals that the ratios representing the slopes increase as $x \longrightarrow x_3$ from the left.

Now at x_2, these inequalities show that for any $x_1 < x_2$ the ratios $\frac{f(x_2)-f(x_1)}{x_2-x_1}$ are bounded above by $\frac{f(x_3)-f(x_1)}{x_3-x_1}$ which does not depend on x_2. Thus the Newton quotients used to compute the left-hand derivative at x_2 are increasing and bounded above implying that $f_-^{'}(x_2)$ exists. By similar argument $f_+^{'}(x_2)$ also exists. Furthermore, it follows that

$$f_+^{'}(x_1) \le \frac{f(x_2)-f(x_1)}{x_2-x_1} \le \frac{f(x_3)-f(x_2)}{x_3-x_1} \le f_-^{'}(x_2), \qquad (7.2)$$

so that

$$f_-^{'}(x_1) \le f_+^{'}(x_1) \le f_-^{'}(x_3) \le f_+^{'}(x_3) \qquad (7.3)$$

Thus $f_-^{'}$ and $f_+^{'}$ are non-decreasing

$(ii) \Rightarrow (iii)$

Since a function which is non-decreasing on an interval is integrable on that interval, it follows that $f_-^{'}$ and $f_+^{'}$ are Riemann integrable. Suppose $x_0 < x \in I$ (the argument for $x < x_0$ is similar and omitted). For any partition

$x_0 < x_1 < \cdots < x_n = x$, by (7.2) and (7.3)

$$f_-^{'}(x_{k-1}) \le f_+^{'}(x_{k-1}) \le \frac{f(x_k)-f(x_{k-1})}{x_k-x_{k-1}} \le f_-^{'}(x_k) \le f_+^{'}(x_k) \qquad (7.4)$$

Since

$$\sum_{k=1}^{n}[f(x_k) - f(x_{k-1})] = \sum_{k=1}^{n} \frac{f(x_k)-f(x_{k-1})}{x_k-x_{k-1}}(x_k - x_{k-1}) = f(x) - f(x_0) \qquad (7.5)$$

$$\Longrightarrow \int_{x_0}^{x} f_-^{'}(t)dt = \int_{x_0}^{x} f_+^{'}(t)dt = f(x) - f(x_0) \qquad (7.6)$$

$(iii) \Rightarrow (iv)$: From (7.4), (7.5) and (7.6), we observe that

$$f'(x_0) \le \frac{f(x)-f(x_0)}{x-x_0}, \text{ if } x_0 < x \text{ and } f'(x_0) \ge \frac{f(x)-f(x_0)}{x-x_0}, \text{ if } x_0 > x$$

In either case

$f(x) - f(x_0) \ge f^{'}(x_0)(x - x_0)$.

$(iv) \Rightarrow (i)$

The existence of $f'(x)$ implies that of $f_-'(x)$ and $f_+'(x)$. Now let $x_2 = \lambda x_1 + (1-\lambda)x_3$, $\lambda \in [0,1]$. By (iv)

$$f(x_3) - f(x_2) \ge f_+'(x_2)(x_3 - x_2)$$

and

$$f(x_2) - f(x_1) \le f_-'(x_2)(x_2 - x_1)$$

But $x_3 - x_2 = \lambda(x_3 - x_1)$ and $x_2 - x_1 = (1-\lambda)(x_3 - x_1)$. Therefore

$$\lambda f_+'(x_2)(x_3 - x_2) \le f(x_3) - f(x_2)$$

$$(1-\lambda)f_-'(x_2)(x_3 - x_1) \ge f(x_2) - f(x_1)$$

Now multiplying the first relation by $(1-\lambda)$ and the second by λ and subtracting we have

$$0 \le (1-\lambda)\lambda[f_+'(x_2) - f_-'(x_2)](x_3 - x_1) \le (1-\lambda)f(x_3) + \lambda f(x_1) - f(x_2)$$

which, from the definition of x_2, implies

$$f(\lambda x_1 + (1-\lambda)x_3) \leq \lambda f(x_1) + (1-\lambda) f(x_3).$$

***Remark* 7.1**

This characterization shows that while a convex function need not be necessarily differentiable, it has both right- and left-hand derivatives at every point in the interval I and is thus almost everywhere differentiable. These derivatives are non-decreasing functions, showing that convex functions are anti-derivatives of non-decreasing functions. Furthermore, it is also clear that the graph of a convex function $f: I \longrightarrow \mathbb{R}$ lies on or above the tangent line drawn at each point $x \in I$, which under the assumption of differentiability means that

$$f(x) \geq f^{'}(x_0)(x - x_0) + f(x_0).$$

8.0 Conclusion

We observe that without a pre-information on the convexity of a function, a presentation with any of the four statements in the result above implies a presentation with all the other statements. It also shows that if a given definition cannot be used in a given scheme we can resort to another with a little refinement.

Despite the fact that these properties already exist in optimization materials, this characterization, in particular through the geometric chord property, has not been achieved. This gives a wider definition of convexity. Thus under appropriate hypothesis, this characterization gives equivalence of the definitions of convexity.

References

[1] Boas, R. P. Jr (1981), *A Primer of Real Functions*, Carus Mathematical Monograph 13, Mathematical Association of America.

[2] Boyd, S. and Vandenberg L. (2004), *Convex Optimization*, Cambridge University Press.

[3] Bronson, R. and Naadimuthu, G. (1997), *Operations Research*, Schaum's Outlines, Tata McGraw-Hill, New Delhi.

[4] Ezimadu, P. E. and Igabari, J. N. (2008), *On a Differential and Integral Characterization of Real-valued Convex Functions of Several Variables*, Journal of the Nigerian Association of Mathematical Physics, 13: 83-86.

[5] Ezimadu, P. E. and Okonta, P. N. (2009), *On a Characterization of Convex Functions*, International Journal of Numerical Mathematics, 4: 98-109.

[6] Fenchel, W. (1953), *Convex Cones, Sets, and Functions*. Lecture notes, Princeton University,Department of Mathematics. From notes taken by D. W. Blackett, Spring 1951.

[7] Larson, R. E., Hostetler, R. P. and Edwards, B. H.(1994), *Calculus*, D. C. Heath and Co., Lexington, M.

[8] Phelps, R. R. (1993), *Convex Functions, Monotone Operators and Differentiability*, Second. Edition, Number 1364 in Lecture Notes in Mathematics, Springer–Verlag, Berlin.

[9] Royden, H. L. (1988), *Real Analysis*, Third Edition, Macmillan, New York.

[10] Stewart, J (1995), *Calculus*, Third Edition Books/Cole, Pacific Grove, CA.

[11] Wilson, C. (2008), *Concave Functions of a Single Variable*, Mathematics for Economists Vol. 31.