Journal of the Nigerian Association of Mathematical Physics
Volume 13 (November, 2008), 83 - 86

On a differential and integral characterization of real-valued convex functions of several variables

P. E. Ezimadu and J. N. Igabari
Department of Mathematics and Computer Science
Delta State University Abraka

Abstract

Convex functions play important roles in the study of optimization. These functions have many important properties which can be used to develop suitable optimality conditions and computational schemes for optimization problems. With a growing need for deeper understanding of these properties comes the need for a wider definition of convex functions. In this work we present a characterization of real valued convex functions of several variables through the derivative and integral. This characterization provides an equivalence for the definitions of convexity.

Keywords: Convex Functions, Convex Sets, Monotone Mapping

1.0 Introduction

The subject of convexity is often treated quite extensively in optimization texts. However the ever-increasing need for its application calls for suitable and easy means of recognizing convex functions. It is possible to give a quite strong (and simple) characterization of convexity by a combination of the first order condition (stated below in Theorem 2.5); a monotone mapping; and the integral, thereby revealing an equivalent definition of convexity.

2.0 Basic Concepts

***Definition* 2.1**

A set $S \subseteq \mathbb{R}^n$ *is convex if* $\forall x_1, x_2 \in S$

$$x = x_1 + (1-\alpha)x_2 \in S \;\; \forall \alpha \in [0,1] \tag{2.1}$$

It follows from this that S can have no re-entrant corners. A more general definition of convex set which readily follows is that:

$\forall\, x_i \in S,\; i = 1, \dots, n,$

$$x = \sum_{i=0}^{n} \alpha_i x_i \in S, \tag{2.2}$$

where $\sum_{i=0}^{n} \;\alpha_i = 1,\; \alpha_i \geq 0$. The vector x in (2.1) or (2.2) is referred to as a convex combination of the points $x_0, x_1, \dots$

***Definition* 2.2**

Let $S \subset \mathbb{R}^n$ *be a nonempty convex set. A function* $f: S \to \mathbb{R}$ *is said to be convex on* S *if for any* $x_1, x_2 \in S$ *and all* $\alpha \in [0,1]$ *we have*

$$f(\alpha x_1 + (1-\alpha)x_2) \leq \alpha f(x_1) + (1-\alpha) f(x_2) \tag{2.3}$$

The right hand side of (2.3) is the chord joining $(x_1, f(x_1))$ to $(x_2, f(x_2))$ on the graph of f, and the inequality expresses the fact that the graph of a convex function always lie below (or along) the chord.

Now for any $x_1, x_2 \in S$, if $f \in C^1(S)$, then

$$f(x_2) = f(x_1) + \int_{x_1}^{x_2} \nabla f(x) dx. \tag{2.4}$$

Thus

$$f(x_2) = f(x_1) + \nabla f(x)^T (x_2 - x_1), \; x \in (x_1, x_2) \tag{2.5}$$

Similarly, for $x_1, x_2 \in S$, we have

$$f(x_2) = f(x_1) + \nabla f(x_1 + \alpha(x_2 - x_1))^T (x_2 - x_1), \quad \alpha \in (0,1) \tag{2.6}$$

or

$$f(x_2) = f(x_1) + \nabla f(x_1)(x_2 - x_1) + o(\| x_2 - x_1 \|) \tag{2.7}$$

Now we will introduce the concept of monotone mapping, which is very useful in the characterization of convex functions

***Definition* 2.4**

Let $F: S \subset \mathbb{R}^n \to \mathbb{R}^n$, *then*

(1) F *is monotone on* S *if for any* $x_1, x_2 \in S$, $\left(F(x_1) - F(x_2)\right)^T (x_1 - x_2) \geq 0$.

(2) F is strictly monotone on S if for any $x_1, x_2 \in S$ $x_1 \neq x_2$, $\left(F(x_1) - F(x_2)\right)^T (x_1 - x_2) > 0$.

***Theorem* 2.5 (First order characterizations of convex functions)**

Let $S \subset \mathbb{R}^n$ *be a nonempty open convex set and* $f: S \to \mathbb{R}$ *be a differentiable function, then*

(a) *f is convex if, and only if, for any* $x_1, x_2 \in S$,

$$f(x_2) \geq f(x_1) + \nabla f(x_2)^T (x_2 - x_1)$$

(b) *f is strictly convex on* S *if, and only if, for any* $x, y \in S$ *with* $x_1 \neq x_2$

$$f(x_2) > f(x_1) + \nabla f(x_1)^T (x_2 - x_1)$$

(c) *f is strongly convex* (*or uniformly convex*) *on S if, and only if, for any* $x_1, x_2 \in S$

$$f(x_2) \geq f(x_1) + \nabla f(x_1)^T (x_2 - x_1) + \frac{1}{2} m \| x_2 - x_1 \|^2,$$

where $m > 0$ *is a constant.*

3.0 Characterization of convex functions through the derivative and the integral

***Theorem* 3.1**

Suppose $f: S \subseteq \mathbb{R}^n \to \mathbb{R}$, $f \in C^1(S)$ and $S \neq \emptyset$ is open and convex, then the following statements are equivalent

(*i*) *f* is convex.

(*ii*) ∇f is monotone.

(*iii*) $f(x_2) - f(x_1) = \int_{x_1}^{x_2} \nabla f(x) dx$, $x \in S$

(iv) $f(x_2) \geq f(x_1) + \nabla f(x)^T (x_2 - x_1)$.

Proof

(*i*)$\Longrightarrow$(*ii*): Since *f* is convex $f(\alpha x_1 + (1-\alpha) x_2) \leq \alpha f(x_1) + (1-\alpha) f(x_2)$

$$\Longrightarrow \frac{f(\alpha x_1 + (1-\alpha) x_2) - f(x_2)}{\alpha} \leq f(x_1) - f(x_2) \tag{3.1}$$

and

$$\frac{f(\alpha x_2 + (1-\alpha) x_1) - f(x_1)}{\alpha} \leq f(x_2) - f(x_1) \tag{3.2}$$

From (3.1) and (3.2), we have that

$$\frac{f(\alpha x_1+(1-\alpha)x_2)-f(x_2)}{\alpha}+\frac{f(\alpha x_2+(1-\alpha)x_1)-f(x_1)}{\alpha}\leq 0$$

As $\alpha\to 0$, we have that $\nabla f(x_2)^T(x_1-x_2)+\ \nabla f(x_1)^T(x_2-x_1)\leq 0$

$$\left(\nabla f(x_2)-\nabla f(x_1)\right)^T(x_2-x_1)\geq 0. \qquad (3.3$$

(*ii*)⟹(*iii*): Since ∇f is monotone, for any $x_1,x_2,x\in S$ with $x=x_1+(1-\alpha)x_2$, $\alpha\in(0,1)$, we have

$$\left(\nabla f(x_2)-\nabla f(x)\right)^T(x_2-x)\geq\left(\nabla f(x)-\nabla f(x_1)\right)^T(x-x_1)\geq 0.$$

$$\Rightarrow\nabla f(x_2)^T(x_2-x)+\ \nabla f(x_1)^T(x-x_1)\geq\nabla f(x)^T(x_2-x)+\ \nabla f(x)^T(x-x_2)$$

$$\Rightarrow\ f(x_2)-f(x)+o(\|x_2-x\|)+f(x)-f(x_1)+o(\|x-x_1\|)\geq\nabla f(x)^T(x_2-x_1)$$

By the mean value theorem

$$f(x_2)-f(x_1)\geq\int_{x_1}^{x_2}\nabla f(x)dx=f(x_2)-f(x_1)$$

$$\Rightarrow f(x_2)-f(x_1)=\int_{x_1}^{x_2}\nabla f(x)dx \qquad (3.4)$$

(Thus we observe that the value of integral of a gradient of a convex function between any two points is the same as the difference in the value of the function between the two points).

(iii)⟹(iv): From (iii), we have that

$$f(x_2)-f(x_1)=\left(\nabla f(x)-\nabla f(x_1)\right)^T(x_2-x_1)+\nabla f(x_1)^T(x_2-x_1) \qquad (3.5)$$

But we observe that

$$\left(\nabla f(x)-\nabla f(x_1)\right)^T(x_1+\alpha(x_2-x_1)-x_2)=\left(\nabla f(x)-\nabla f(x_1)\right)^T(t(x_2-x_1))$$

$$\Rightarrow\left(\nabla f(x)-\nabla f(x_1)\right)^T(x_2-x_1)=\frac{1}{\alpha}\left(\nabla f(x)-\nabla f(x_1)\right)^T(x-x_1)\geq 0 \qquad (3.6)$$

Thus from (3.5), (3.6) implies that

$$f(x_2)-f(x_1)\geq\nabla f(x)^T(x_2-x_1) \qquad (3.7)$$

(iv)⟹(i): Assume that (iv) holds, and consider the convex combination

$$x=\alpha x_1+(1-\alpha)x_2,\ \alpha\in(0,1).$$

Then

$$f(x_1)\geq f(x)+\nabla f(x)^T(x_1-x)$$

$$\Rightarrow f(x_2)\geq f(x)+\nabla f(x)^T(x_2-x)$$

$$\alpha f(x_1)+(1-\alpha)f(x_2)\geq f(x)+\nabla f(x)^T(\alpha x_1+(1-\alpha)x_2-x)$$

$$=f(\alpha x_1+(1-\alpha)x_2.$$

***Remark* 3.2**

Thus we can see from (*iv*) that the graph of *f* must lie above (along) the linearization of *f* about x_1 and hence this linearization acts as a supporting hyperplane for the convex function. A demonstration of the equivalence of (*iv)* is (*ii*) which illustrates that the slope of a convex function is non-decreasing along any line. In fact this result (for the directional derivative) can also be proved to hold for non-differentiable convex functions. The integral of this slope between any two points along any line is the same as the difference in the value of the function between the two points along this line. This is revealed in (*iii*). Thus under appropriate hypothesis this characterization gives an equivalence for the definitions of convexity.

4.0 Contribution

A presentation with any of (*ii)* to (*iv*) without a pre-information on the convexity of a function implies not only that the function is convex but also a presentation with all of (*i*) to (*iv*). Although these properties (the first order condition; the integral; and monotonicity) are in optimization texts, this combination which gives a wider definition of a convexity has not been achieved. Thus this work places us at a better horizon for recognizing convex functions which are widely used in optimization.

5.0 Conclusion

A version of this characterization for functions on $\mathbb{R}$ can be proved. For functions enjoying the Geometric Chord property this version can be proved to show that a convex function need not be necessarily differentiable. This result can also be extended to incorporate the Hessian matrix of a function on $\mathbb{R}^n$.

References

[1] Andersen, E.D. and Ye, Y. (1998): *A computational study of the homogeneous algorithm for large-scale convex optimization.* Computational Optimization and Applications, 10:243-269, 1998.

[2] Barvinok, A.(2002): *A Course in Convexity,* Volume 54 of Graduate Studies in Mathematics. American Mathematical Society.

[3] Ben-Tal, A. and Nemirovski, A. (1998): Robust convex optimization. Mathematics of Operations Research, 23(4):769-805

[4] Borwein, J. M. and Lewis, A. S.(2000): *Convex Analysis and Nonlinear Optimization.* Springer.

[5] Boyd, S. and Vandenber L.(2004): *Convex Optimization.* Cambridge University Press.

[6] Ekeland, I. and Temam, R.(1999): *Convex Analysis and Variational Inequalities.* Classics in Applied Mathematics. Society for Industrial and Applied Mathematics.

[7] Florenzano, M. and Le Van, C.(2001): *Finite Dimensional Convexity and Optimization.* Number 13 in Studies in Economic Theory. Springer, 2001.

[8] Forsgren, A; Gill, P. E. and Wright, M. H. (2002): *Interior methods for nonlinear optimization.* SIAM Review, 44(4):525-597.

[9] Hiriart-Urruty, J.-B. and Lemarechal, C.(2001) *Fundamentals of Convex analysis* Springer.

[10] Wenyu S. and Yaxiang Y.(2006) *Optimization Theory and Method*. Springer Science+Business Media, LLC.

[11] Luo, Z.-Q; Sturm, J. F. and Zhang, S. (2000): *Conic convex programming and self-dual embedding.* Optimization Methods and Software, 14:169-218.