J. Austral. Math. Soc. (*Series A*) **58** (1995), 404–410

TWO GRADIENT PROPERTIES OF EXPLICITLY CONVEX FUNCTIONS

XINMIN YANG

(Received 20 July 1992; revised 4 October 1992)

Communicated by B. Mond

Abstract

In this paper, two gradient properties of explicit convex functions are given.

1991 *Mathematics subject classification* (*Amer. Math. Soc.*): 26B25, 52A41.
Keywords and phrases: explicitly convex function, differentiable, mean-value theorem.

1. Introduction

Convexity plays a central role in mathematical economics, engineering, management science and optimization theory, and the subject is one currently being discussed in the mathematical programming literature, [1, 3, 4].

The definition of explicitly convex functions appears in [5].

DEFINITION 1.1. Let C be an open convex set of R^n, and $f : C \to R$. The function f is said to be an *explicitly convex function* on C if, for every pair of points $x \in C$, $y \in C$, $f(x) \neq f(y)$, we have

$$f\big(\lambda x + (1-\lambda) y\big) < \lambda f(x) + (1-\lambda) f(y), \quad \forall \lambda \in (0, 1).$$

EXAMPLE 1.1. Consider

$$f(x) = x, \quad x \in R.$$

Then f is a convex function on R, but f is not an explicitly convex function on R.

This work was supported by the National Science Foundation of China.

Obviously, a strictly convex function is an explicitly convex function. But the converse is not true.

EXAMPLE 1.2. This example illustrates that an explicitly convex function need not be either a convex function or a strictly convex function.

$$f(x) = \begin{cases} 1 & \text{if} \quad x = 0, \\ 0 & \text{if} \quad x \neq 0. \end{cases}$$

Then f is an explicitly convex function on R, but f is not a convex function, nor a strictly convex function on R: for $x_1 = 1$ and $x_2 = -1$, $f(x_1) = f(x_2) = 0$, but

$$f\left[(1/2)\,x_1 + (1/2)\,x_2\right] = f(0) = 1 > (1/2)\,f(x_1) + (1/2)\,f(x_2).$$

The following proposition shows that a local minimum of an explicitly convex function over a convex set is also a global minimum.

PROPOSITION 1.1. *Let C be a non-empty convex set in R^n and $f : C \to R$. Consider the problem of minimizing $f(x)$ subject to $x \in C$. Suppose that $\bar{x} \in C$ is a local optimal solution to the problem and f is an explicitly convex function. Then $\bar{x}$ is a global optimal solution.*

PROOF. Since $\bar{x}$ is a local optimal solution, then there exists an ϵ-neighborhood $N_\epsilon(\bar{x})$ around x such that

$$(*) \qquad f(x) \geq f(\bar{x}) \quad \text{for each} \quad x \in C \cap N_\epsilon(\bar{x}).$$

By contradiction, suppose that $\bar{x}$ is not a global optimal solution so that $f(\hat{x}) < f(\bar{x})$ for some $\hat{x} \in C$. By the explicit convexity of f, the following is true for each $\alpha \in (0, 1)$

$$f\left(\alpha\hat{x} + (1-\alpha)\,\bar{x}\right) < \alpha f(\hat{x}) + (1-\alpha)\,f(\bar{x}) < f(\bar{x}).$$

But for $\alpha > 0$ and sufficiently small, $\alpha\hat{x} + (1-\alpha)\,\bar{x} \in C \cap N_\epsilon(\bar{x})$. Hence the above inequality contradicts (*); this completes the proof.

From Examples 1.1 and 1.2, and Proposition 1.1, it is clear that explicitly convex functions are a very useful class and that research of such functions is worthwhile from a mathematical point of view.

In this note, based on Karamardian and Schaible's idea in [2], we give two gradient properties of explicitly convex functions.

https://doi.org/10.1017/S1446788700038386 Published online by Cambridge University Press

2. Main results

LEMMA 2.1. *Let C be a non-empty open convex set of R^n, and let $f : C \to R$ be an explicitly convex function. If f is lower semi-continuous, then f is a convex function on C.*

PROOF. Let $x, y \in C$. If $f(x) \neq f(y)$, then by the definition of an explicitly convex function, we must have $f[\lambda x + (1-\lambda) y] < \lambda f(x) + (1-\lambda) f(y)$ for each $\lambda \in (0,1)$. Now suppose that $f(x) = f(y)$. To show that f is a convex function, we need to show that $f[\lambda x + (1-\lambda) y] \leq f(x)$ for each $\lambda \in (0,1)$. By contradiction, suppose that $f[\alpha x + (1-\alpha) y] > f(x)$ for some $\alpha \in (0,1)$. Denote $\alpha x + (1-\alpha) y$ by z. Since f is lower semi-continuous and explicitly convex, there exists a $\beta \in (0,1)$ such that

$$\text{(A)} \qquad f(z) > f[\beta x + (1-\beta) z] > f(x) = f(y).$$

Note that z can be represented as a convex combination of $u = \beta x + (1-\beta) z$ and y. Hence by the explicit convexity of f, and since $f(u) > f(y)$, $f(z) < f(u)$, contradicting (A). This completes the proof.

THEOREM 2.2. *Let C be an open convex set of R^n, and let $f : C \to R$ be a differentiable function. Then f is an explicitly convex function if and only if, for every pair of points $x \in C$, $y \in C$, $f(x) \neq f(y)$, we have*

$$f(y) > f(x) + (y-x)^T \nabla f(x).$$

PROOF. Suppose that f is an explicitly convex function on C. By Definition 1.1, for every pair of points $x \in C$, $y \in C$, $f(x) \neq f(y)$, we have

$$f(\lambda x + (1-\lambda) y) < \lambda f(x) + (1-\lambda) f(y), \quad \forall \lambda \in (0,1).$$

This yields

$$\left\{f[x + \lambda (y-x)] - f(x)\right\}/\lambda < f(y) - f(x), \quad \forall \lambda \in (0,1).$$

From Lemma 2.1 above and Lemma 3.1.5 in [1], we get

$$(y-x)^T \nabla f(x) = \inf_{\lambda \geq 0}\left\{f[x + \lambda (y-x)] - f(x)\right\} \Big/ \lambda < f(y) - f(x),$$

that is, $f(y) > f(x) + (y-x)^T \nabla f(x)$.

Conversely, suppose that for every pair of points $x \in C$, $y \in C$, $f(x) \neq f(y)$, we have

$$\text{(B)} \qquad f(y) > f(x) + (y-x)^T \nabla f(x).$$

https://doi.org/10.1017/S1446788700038386 Published online by Cambridge University Press

Now let $z_\alpha = \alpha x + (1-\alpha) y$, $\forall \alpha \in (0, 1)$. Without loss of generality, we assume $f(x) < f(y)$. We now show that $f(z_\alpha) \neq f(y)$.

Assume to the contrary that

$$f(z_{\alpha_0}) = f(y), \tag{1}$$

for some $\alpha_0 \in (0, 1)$. Now from (1), we will show that $f\left[\lambda y + (1-\lambda) z_{\alpha_0}\right] = f(y)$, for any $\lambda \in (0, 1)$.

Indeed, if there exists $\bar{\lambda} \in (0, 1)$, such that $f\left[\bar{\lambda} y + \left(1-\bar{\lambda}\right) z_{\alpha_0}\right] \neq f(y)$, then:

(i) If $f\left[\bar{\lambda} y + \left(1-\bar{\lambda}\right) z_{\alpha_0}\right] > f(y)$, let

$$g(\lambda) = f\left[\lambda y + (1-\lambda) z_{\alpha_0}\right], \quad \forall \lambda \in [0, 1].$$

Then g attains a maximum on $(0, 1)$. Assume that g attains its maximum at $\lambda_0 \in (0, 1)$. So $\left(y - z_{\alpha_0}\right)^T \nabla f\left[\lambda_0 y + (1-\lambda_0) z_{\alpha_0}\right] = 0$, yielding

$$\left\{\left[\lambda_0 y + (1-\lambda_0) z_{\alpha_0}\right] - z_{\alpha_0}\right\}^T \nabla f\left[\lambda_0 y + (1-\lambda_0) z_{\alpha_0}\right] = 0. \tag{2}$$

From (1) and (2), obtain $f(z_{\alpha_0}) > f\left[\lambda_0 y + (1-\lambda_0) z_{\alpha_0}\right]$, which contradicts g attaining a maximum at λ_0.

(ii) If $f\left[\bar{\lambda} y + \left(1-\bar{\lambda}\right) z_{\alpha_0}\right] < f(y)$, then since $f(z_{\alpha_0}) = f(y)$ and $f(x) < f(y)$, we see that $f(x) < f(z_{\alpha_0})$ and $f\left[\bar{\lambda} y + \left(1-\bar{\lambda}\right) z_{\alpha_0}\right] < f(z_{\alpha_0})$. Hence the function

$$g(\lambda) = f\left[\lambda x + (1-\lambda)\left(\bar{\lambda} y + \left(1-\bar{\lambda}\right) z_{\alpha_0}\right)\right]$$

attains a maximum on $(0, 1)$. Suppose this maximum occurs at λ_0. Then

$$g'(\lambda_0) = \left\{x - \left[\bar{\lambda} y + \left(1-\bar{\lambda}\right) z_{\alpha_0}\right]\right\}^T \nabla f\left\{\lambda_0 x + (1-\lambda_0)\left[\bar{\lambda} y + \left(1-\bar{\lambda}\right) z_{\alpha_0}\right]\right\} = 0.$$

Now, $f(x) < f\left\{\lambda_0 x + (1-\lambda_0)\left[\bar{\lambda} y + \left(1-\bar{\lambda}\right) z_{\alpha_0}\right]\right\}$ implies that

$$\begin{aligned} f(x) > {} & f\left\{\lambda_0 x + (1-\lambda_0)\left[\bar{\lambda} y + \left(1-\bar{\lambda}\right) z_{\alpha_0}\right]\right\} \\ & + \left\{x - \lambda_0 x - (1-\lambda_0)\left[\bar{\lambda} y + \left(1-\bar{\lambda}\right) z_{\alpha_0}\right]\right\}^T \\ & \nabla f\left\{\lambda_0 x + (1-\lambda_0)\left[\bar{\lambda} y + \left(1-\bar{\lambda}\right) z_{\alpha_0}\right]\right\} \\ = {} & f\left\{\lambda_0 x + (1-\lambda_0)\left[\bar{\lambda} y + \left(1-\bar{\lambda}\right) z_{\alpha_0}\right]\right\} \\ & + (1-\lambda_0)\left\{x - \left[\bar{\lambda} y + \left(1-\bar{\lambda}\right) z_{\alpha_0}\right]\right\}^T \\ & \nabla f\left\{\lambda_0 x + (1-\lambda_0)\left[\bar{\lambda} y + \left(1-\bar{\lambda}\right) z_{\alpha_0}\right]\right\} \\ = {} & f\left\{\lambda_0 x + (1-\lambda_0)\left[\bar{\lambda} y + \left(1-\bar{\lambda}\right) z_{\alpha_0}\right]\right\}, \quad \text{a contradiction.} \end{aligned}$$

Combining (i) and (ii), we have

$$f\left[\lambda y + (1-\lambda)\, z_{\alpha_0}\right] = f(y), \quad \forall \lambda \in [0, 1]. \tag{3}$$

Let $h(\lambda) = f\left[\lambda y + (1-\lambda)\, z_{\alpha_0}\right]$. From (3) we get

$$0 = h'(1) = \left(y - z_{\alpha_0}\right)^T \nabla f(y). \tag{4}$$

By the hypothesis of the theorem and (4), we obtain $f(x) > f(y)$, a contradiction. Therefore

$$f(z_\alpha) \neq f(y), \quad \forall \alpha \in (0, 1). \tag{5}$$

If $f(z_\alpha) = f(x)$ for some $\alpha \in (0, 1)$, then $f(z_\alpha) < \alpha f(x) + (1-\alpha) f(y)$ is trivial. If $f(z_\alpha) \neq f(x)$ for some $\alpha \in (0, 1)$, then by Hypothesis (B) we have

$$f(x) > f(z_\alpha) + (x - z_\alpha)^T \nabla f(z_\alpha), \tag{6}$$

$$f(y) > f(z_\alpha) + (y - z_\alpha)^T \nabla f(z_\alpha), \tag{7}$$

in view of (5). Multiplying (6) by α, and (7) by $(1-\alpha)$, and then adding, yields

$$f(z_\alpha) < \alpha f(x) + (1-\alpha) f(y).$$

This completes the proof of Theorem 2.2.

THEOREM 2.3. *Let f be differentiable on an open convex subset C of R^n. Then f is explicitly convex on C if and only if for every pair of points $x \in C$, $y \in C$, $f(x) \neq f(y)$, we have*

$$(y - x)^T \left[\nabla f(y) - \nabla f(x)\right] > 0.$$

PROOF. Suppose that f is an explicitly convex function on C. Let $x, y \in C$, $f(x) \neq f(y)$. From Theorem 2.2 we have

$$f(y) > f(x) + (y - x)^T \nabla f(x), \tag{8}$$

$$f(x) > f(y) + (x - y)^T \nabla f(y). \tag{9}$$

Adding these we obtain $(y - x)^T \left[\nabla f(y) - \nabla f(x)\right] > 0$.

Conversely, suppose that for every pair of points $x, y \in C$, $f(x) \neq f(y)$, we have

$$(y - x)^T \left[\nabla f(y) - \nabla f(x)\right] > 0. \tag{C}$$

From the mean-value theorem we obtain

$$f(y) - f(x) = (y - x)^T \nabla f(\bar{x}), \tag{10}$$

where

(11) $$\bar{x} = \lambda x + (1-\lambda)\, y$$

for some $0 < \lambda < 1$.

(I) If $f(x) \neq f(\bar{x})$, then from (C) we obtain $(\bar{x} - x)^T \left[\nabla f(\bar{x}) - \nabla f(x)\right] > 0$. This yields

(12) $$(y - x)^T\, \nabla f(\bar{x}) > (y - x)^T\, \nabla f(x),$$

in view of (11). Now from (10) and (12), we have

$$f(y) > f(x) + (y - x)^T\, \nabla f(x).$$

(II) If $f(x) = f(\bar{x})$, we want to show that $f(x) = f(\bar{x}) \neq f(u)$, where $u = \alpha x + (1-\alpha)\,\bar{x}$, for some $0 < \alpha < 1$.

Assume to the contrary that

(13) $$f(x) = f(\bar{x}) = f(u),$$

where $u = \alpha x + (1-\alpha)\,\bar{x}$, for any $0 < \alpha < 1$.

$$\text{Let} \qquad \phi(\alpha) = f\left[x + \alpha\,(\bar{x} - x)\right], \quad \forall \alpha \in [0, 1].$$

Then (13) implies that

$$\phi(\alpha) = \text{const} = f(\bar{x}), \quad \forall \alpha \in [0, 1].$$

This yields

$$0 = \phi'(1) = (\bar{x} - x)^T\, \nabla f(\bar{x}) = (1 - \lambda)\,(y - x)^T\, \nabla f(\bar{x}).$$

Hence, $(y - x)^T\, \nabla f(\bar{x}) = 0$, which together with (10) contradicts $f(x) \neq f(y)$. Thus

(14) $$f(x) = f(\bar{x}) \neq f(u)$$

where

(15) $$u = \alpha x + (1-\alpha)\,\bar{x},$$

for some $0 < \alpha < 1$.

Now from (B) and (15) we have

$$(\bar{x} - u)^T \left[\nabla f(\bar{x}) - \nabla f(u)\right] > 0, \quad (x - u)^T \left[\nabla f(x) - \nabla f(u)\right] > 0.$$

This yields

(16) $$(\bar{x} - x)^T \left[\nabla f(\bar{x}) - \nabla f(u)\right] > 0,$$

(17) $$(x - \bar{x})^T \left[\nabla f(x) - \nabla f(u)\right] > 0,$$

https://doi.org/10.1017/S1446788700038386 Published online by Cambridge University Press

in view of (15). Now (16) and (17) together imply

$$(\bar{x} - x)^T \left[\nabla f(\bar{x}) - \nabla f(x)\right] > 0. \tag{18}$$

Multiplying (18) by $1/(1-\lambda)$ and noting (11), we obtain

$$(y - x)^T \left[\nabla f(\bar{x}) - \nabla f(x)\right] > 0. \tag{19}$$

That is,

$$(y - x)^T \nabla f(\bar{x}) > (y - x)^T \nabla f(x). \tag{20}$$

Combining (10) and (20) we have $f(y) > f(x) + (y-x)^T \nabla f(x)$. Given (I) and (II), this implies that for every pair of points $x, y \in C$, $f(x) \neq f(y)$, we have

$$f(y) > f(x) + (y-x)^T \nabla f(x).$$

From Theorem 2.2 we conclude that f is an explicitly convex function on C.

Acknowledgement

The author wishes to thank the referee for several valuable suggestions which improved the presentation of this paper.

References

[1] M. S. Bazaraa and C. M. Shetty, *Nonlinear programming: Theory and algorithms* (Wiley, New York, 1979).
[2] S. Karamardian and S. Schaible, 'Seven kinds of monotone maps', *J. Optim. Theory Appl.* **66** (1990), 37–46.
[3] A. W. Roberts and D. E. Varberg, *Convex functions* (New York, 1973).
[4] R. T. Rockafellar, *Convex analysis* (Princeton University Press, Princeton, 1970).
[5] S. Xue and S. Shen, 'Eleven kinds of convexity', *Chinese Journal Oper. Res.* **8** (1989), 72–75.

Department of Mathematics
Chongqing Normal University
Chongqing 630047
China