Mathematica Aeterna, Vol. 4, 2014, no. 6, 651 - 657

Some Characterizations of Strongly Convex Functions in Inner Product Spaces

Teodoro Lara

Departamento de Física y Matemáticas. Universidad de los Andes. Venezuela.
tlara@ula.ve

Nelson Merentes

Escuela de Matemáticas. Universidad Central de Venezuela. Caracas. Venezuela
nmerucv@gmail.com

Edgar Rosales
Departamento de Física y Matemáticas. Universidad de los Andes. Venezuela.
edgarr@ula.ve

Mayrelly Valera
Universidad Nacional Experimental Simón Rodríguez. Valera. Venezuela.
mayrellyvalera@yahoo.es

Abstract

In this research we show some characterizations in terms of inequalities, for strongly convex functions defined on inner product spaces. These results generalize the ones for functions defined on real intervals. They involve ideas of the first and second derivatives on inner product spaces.

Mathematics Subject Classification: 26A51, 39B62

Keywords: Strongly Convex Functions, Support of a Convex Function, Inner Product Space.

1 Introduction

Strongly convex functions have been introduced by Polyak ([6]), he used them for proving the convergence of a gradient type algorithm for minimizing a function. They play an important role in optimization theory and mathematical

[1]This research has been partially supported by Central Bank of Venezuela.

economics. Many properties and applications are available in the literature, for instance in [1, 2, 3, 4, 5, 7]. In this paper we show characterizations of strongly convex functions in terms of inequalities similar to the known for the real case but now on inner product spaces.

2 Preliminaries

In this section we recall some results for strongly convex functions with modulus c defined on real intervals. A function $f : (a, b) \to \mathbf{R}$ is called strongly convex with modulus $c > 0$ if

$$f(tx + (1-t)y) \leq tf(x) + (1-t)f(y) - ct(1-t)(x-y)^2,$$

with $x, y \in (a, b)$ and $t \in [0, 1]$. It is known that $f : (a, b) \to \mathbf{R}$ is strongly convex with modulus $c > 0$ if and only if at every point $x_0 \in (a, b)$ it has a support of the form

$$f(x) \geq f(x_0) + L(x - x_0) + c(x - x_0)^2;$$

where $L \in [f'_-(x_0), f'_+(x_0)]$ and $f'_\pm(x_0)$ are the right and left derivative respectively of f at x_0 ([7]). Actually, in the case f differentiable $L = f'(x_0)$.

For differentiable f the following statements take place ([7])

1. f is strongly convex with modulus c if and only if f' is strongly increasing, that is,
$$(f'(x) - f'(y))(x - y) \geq 2c(x - y)^2.$$
2. For twice differentiable f, f is strongly convex with modulus c if and only if $f''(x) \geq 2c$.

3 Strongly Convex Functions on Inner Product Spaces

Here we state and prove our main results, those are similar to the case on real intervals but now on inner product spaces. First we recall some facts about derivatives on normed spaces; let X and Y be normed spaces, $g : U \subseteq X \to Y$ a function and U an open subset. Then g is said to be differentiable at $x_0 \in U$ if there exists a linear transformation $S : X \to Y$ such that, for $h \in X$ small enough,

$$g(x_0 + h) = g(x_0) + S(h) + ||h||\epsilon(x_0, h),$$

where $\epsilon(x_0, h) \in Y$ and goes to zero as $||h|| \to 0$. This linear transformation is called the derivative and is denoted by $g'(x_0)$. It is worth to notice that we can think of g' as a mapping $g' : U \to \mathcal{F}(X, Y)$, where $\mathcal{F}(X, Y)$ is the set of linear transformations from X to Y. If g' is continuous and differentiable at x then we may define $g''(x)$ as a linear transformation from X to $\mathcal{F}(X, Y)$,

$$g''(x) : X \to \mathcal{F}(X, Y),$$

so $g''(x)(h) \in \mathcal{F}(X, Y)$ and $[g''(x)(h)(k)]$ makes sense for $k \in X$. Again we may notice that the expression $[g''(x)(h)(k)]$ is linear in both variables h and k, so $g''(x)$ is bilinear transformation from $X \times X$ to Y and it is denoted as $g''(x)(h, k)$.

Through this section, let X be an inner product space, $\Omega \subseteq X$ a convex set and $f : \Omega \to \mathbf{R}$ a function.

Definition 1 *([1, 3, 4]). $f : \Omega \to \mathbf{R}$ is called strongly convex with modulus $c > 0$ if*

$$f(tx + (1 - t)y) \leq tf(x) + (1 - t)f(y) - ct(1 - t) \parallel x - y \parallel^2$$

for any $x, y \in \Omega$, $t \in [0, 1]$.

Theorem 2 *Let $f : \Omega \to \mathbf{R}$ be a differentiable function. Then f is strongly convex with modulus $c > 0$ if and only if*

$$f(x) \geq f(x_0) + f'(x_0)(x - x_0) + c \parallel x - x_0 \parallel^2,$$

for any $x, x_0 \in \Omega$.

Proof. If f is strongly convex with modulus $c > 0$ then

$$f(tx + (1 - t)x_0) \leq tf(x) + (1 - t)f(x_0) - ct(1 - t) \parallel x - x_0 \parallel^2,$$

so

$$tf(x) + (1 - t)f(x_0) \geq f(tx + (1 - t)x_0) + ct(1 - t) \parallel x - x_0 \parallel^2,$$

but then

$$tf(x) + f(x_0) - tf(x_0) \geq f(tx + x_0 - tx_0) + ct(1 - t) \parallel x - x_0 \parallel^2,$$

that is,

$$t[f(x) - f(x_0)] \geq f(tx + x_0 - tx_0) - f(x_0) + ct(1 - t) \parallel x - x_0 \parallel^2, \quad (1)$$

dividing both side of (1) by t we get

$$f(x) - f(x_0) \geq \frac{f(t(x - x_0) + x_0) - f(x_0)}{t} + c(1 - t) \parallel x - x_0 \parallel^2$$

and letting $t \to 0^+$,

$$f(x) - f(x_0) \geq f'(x_0)(x - x_0) + c \parallel x - x_0 \parallel^2 .$$

For the converse let $x_1, x_2 \in \Omega$, $t \in [0,1]$ and $x_0 = tx_1 + (1-t)x_2$. It is clear that

$$f(x_0) = f(x_0) + f'(x_0)[t(x_1 - x_0) + (1-t)(x_2 - x_0)],$$

or better

$$f(x_0) = f(x_0) + tf'(x_0)(x_1 - x_0) + (1-t)f'(x_0)(x_2 - x_0),$$

therefore

$$f(x_0) = t[f(x_0) + f'(x_0)(x_1 - x_0)] + (1-t)[f(x_0) + f'(x_0)(x_2 - x_0)].$$

By hypothesis,

$$f(x) \geq f(x_0) + f'(x_0)(x - x_0) + c \parallel x - x_0 \parallel^2, \tag{2}$$

now we change x by x_1 in (2) and multiply by t the resulting expression to get

$$tf(x_1) \geq t[f(x_0) + f'(x_0)(x_1 - x_0) + c \parallel x_1 - x_0 \parallel^2],$$

that is,

$$tf(x_0) + tf'(x_0)(x_1 - x_0) + ct \parallel x_1 - x_0 \parallel^2 \leq tf(x_1). \tag{3}$$

In the same way we change x by x_2 in (2) and multiply the expression that comes out by $(1-t)$,

$$(1-t)f(x_2) \geq (1-t)[f(x_0) + f'(x_0)(x_2 - x_0) + c \parallel x_2 - x_0 \parallel^2],$$

or

$$(1-t)f(x_0) + (1-t)f'(x_0)(x_2 - x_0) + c(1-t) \parallel x_2 - x_0 \parallel^2 \leq (1-t)f(x_2). \tag{4}$$

Now by adding up (3) and (4),

$$f(x_0) \leq tf(x_1) + (1-t)f(x_2) - ct \parallel x_1 - x_0 \parallel^2 - c(1-t) \parallel x_2 - x_0 \parallel^2 . \tag{5}$$

But $x_1 - x_0 = x_1 - (tx_1 + (1-t)x_2) = (1-t)(x_1 - x_2)$ and $x_2 - (tx_1 + (1-t)x_2) = t(x_2 - x_1)$. Thus, using the fact that Ω is an inner product space we have

$$\parallel x_1 - x_0 \parallel = (1-t) \parallel x_1 - x_2 \parallel \text{ and } \parallel x_2 - x_0 \parallel = t \parallel x_1 - x_2 \parallel .$$

Hence, (5) can be written as

$$f(tx_1 + (1-t)x_2) \leq tf(x_1) + (1-t)f(x_2) - ct(1-t)^2 \parallel x_1 - x_2 \parallel^2 - c(1-t)t^2 \parallel x_1 - x_2 \parallel^2$$

or better

$$f(tx_1 + (1-t)x_2) \leq tf(x_1) + (1-t)f(x_2) - ct(1-t) \parallel x_1 - x_2 \parallel^2 .$$

Therefore f is strongly convex with modulus $c > 0$.

The following result follows ideas from [2] and [7].

Theorem 3 *Let $f : \Omega \to \mathbf{R}$ be differentiable. Then f is strongly convex with modulus $c > 0$ if and only if*

$$[f'(x) - f'(y)](x - y) \geq 2c \parallel x - y \parallel^2,$$

for any $x, y \in \Omega$.

Proof.

If f is strongly convex with modulus $c > 0$, then by foregoing Theorem 2

$$f(x) \geq f(x_0) + f'(x_0)(x - x_0) + c \parallel x - x_0 \parallel^2,$$

for any $x, x_0 \in \Omega$. Then for $x_0 = y$

$$f(x) \geq f(y) + f'(y)(x - y) + c \parallel x - y \parallel^2 . \tag{6}$$

Similarly, switching x and y

$$f(y) \geq f(x) + f'(x)(y - x) + c \parallel y - x \parallel^2, \tag{7}$$

and adding up (6) and (7)

$$f(x) + f(y) \geq f(y) + f(x) + f'(y)(x - y) + f'(x)(y - x) + 2c \parallel x - y \parallel^2 .$$

Thus,

$$f'(x)(x - y) - f'(y)(x - y) \geq 2c \parallel x - y \parallel^2,$$

that is

$$[f'(x) - f'(y)](x - y) \geq 2c \parallel x - y \parallel^2 .$$

Conversely, let $x_0, x_1 \in \Omega$, and consider the univariate function $\varphi(t) = f(x_t)$, where $x_t = x_0 + t(x_1 - x_0)$, for $t \in [0, 1]$; φ is well defined because $x_t \in \Omega$, for any $t \in [0, 1]$ and is differentiable because f is, even more, $\varphi'(t) = f'(x_t)(x_1 - x_0)$, but then for $0 \leq t' < t \leq 1$,

$$\varphi'(t) - \varphi'(t') = (f'(x_t) - f'(x_{t'}))(x_1 - x_0).$$

Because $x_t - x_{t'} = (t - t')(x_1 - x_0)$,

$$\begin{aligned} (f'(x_t) - f'(x_{t'}))(x_1 - x_0) &= (f'(x_t) - f'(x_{t'})) \left(\frac{x_t - x_{t'}}{t - t'} \right) \\ &= \frac{1}{t - t'}(f'(x_t) - f'(x_{t'}))(x_t - x_{t'}), \end{aligned}$$

thus

$$\varphi'(t) - \varphi'(t') = \frac{1}{t - t'}(f'(x_t) - f'(x_{t'}))(x_t - x_{t'}).$$

By hypothesis, the right hand side is greater than or equal to $\frac{1}{t-t'}[2c \parallel x_t - x_{t'} \parallel^2]$. Now if we set $t' = 0$,

$$\varphi'(t) - \varphi'(0) \geq \frac{1}{t} 2c \parallel x_t - x_0 \parallel^2 = 2tc \parallel x_1 - x_0 \parallel^2 .$$

Therefore we can deduce

$$\varphi(1) - \varphi(0) - \varphi'(0) = \int_0^1 [\varphi'(t) - \varphi'(0)]dt \geq c \parallel x_1 - x_0 \parallel^2,$$

or

$$f(x_1) - f(x_0) - f'(x_0)(x_1 - x_0) \geq c \parallel x_1 - x_0 \parallel^2,$$

hence

$$f(x_1) \geq f(x_0) + f'(x_0)(x_1 - x_0) + c \parallel x_1 - x_0 \parallel^2,$$

conclusion follows from Theorem 2.

Theorem 4 *If $f''(x_0)$ exists for any $x_0 \in \Omega$, then f is strongly convex with modulus $c > 0$ if and only if for all $x_0, x \in \Omega$*

$$f''(x_0)(x, x) \geq 2c \parallel x \parallel^2 .$$

Before going over the the proof we recall that $f : \Omega \to \mathbf{R}$ and, as done at the beginning of this section, $f''(x_0)$ can be thought as a bilinear transformation from $\Omega \times \Omega$ to $\mathbf{R}$.

Proof. For any $x, y \in \Omega$, f can be written as

$$f(x) = f(y) + f'(y)(x - y) + \frac{1}{2} f''(y + s(x - y))(x - y, x - y)$$

for some $s \in (0, 1)$ ([7]), but then

$$f''(y + s(x - y))(x - y, x - y) = 2[f(x) - f(y) - f'(y)(x - y)].$$

Now, by the hypothesis and Theorem 2, the right hand side is greater than or equal to $2c \parallel x - y \parallel^2$. In other words,

$$f''(y + s(x - y))(x - y, x - y) \geq 2c \parallel x - y \parallel^2$$

or

$$f''(y + sh)(h, h) \geq 2c \parallel h \parallel^2, \text{ where } h = x - y.$$

Conversely, given x_0, $x \in \Omega$ again we may write f as

$$f(x) = f(x_0) + f'(x_0)(x - x_0) + \frac{1}{2} f''(x_0 + s(x - x_0))(x - x_0, x - x_0),$$

for some $s \in (0, 1)$. By hypothesis

$$\frac{1}{2} f''(x_0 + s(x - x_0))(x - x_0, x - x_0) \geq c \parallel x - x_0 \parallel^2$$

that is

$$f(x) \geq f(x_0) + f'(x_0)(x - x_0) + c \parallel x - x_0 \parallel^2,$$

and by Theorem 2, f is strongly convex.

References

[1] A. Azócar, J. Giménez, K. Nikodem and J.L. Sánchez, On strongly mid-convex functions, Opuscula Math., 31/1 (2011), 15-26.

[2] J. B. Hiriart-Urruty and C. Lemaréchal, Fundamentals of Convex Analysis, Springer-Verlag, Berlin-Heidelberg, 2001.

[3] J. Makó, K. Nikodem and Zs. Páles, On strong (α, F)- convexity, Math. Inequal. Appl., 15 (2012), 289–299.

[4] N. Merentes, K. Nikodem, Remarks on Strongly Convex Functions, Aeqationes. Math. 80 (2010), 193-199.

[5] K. Nikodem and Zs. Páles, Characterizations of Inner Product Spaces by Strongly Convex Functions, Banach J. Math. Anal. **1**(2011), 83-87.

[6] B. T. Polyak, Existence Theorems and Convergence of Minimizing Sequences in Extremum Problems with Restrictions, Soviet Math. Dokl. 7 (1966), 7275.

[7] A. W. Roberts and D. E. Varberg, Convex Functions, Academic Pres. NY. 1973.

Received: June, 2014