

PROBABILISTIC CHARACTERIZATION OF STRONG CONVEXITY

Teresa Rajba, Szymon Wąsowicz

Abstract. Strong convexity is considered for real functions defined on a real interval. Probabilistic characterization is given and its geometrical sense is explained. Using this characterization some inequalities of Jensen-type are obtained.

Keywords: convexity, strong convexity, Jensen's inequality, Jensen gap of a function, distribution of a random variable, variance.

Mathematics Subject Classification: Primary: 26A51. Secondary: 26D15, 60E05, 60E15.

1. INTRODUCTION

Recall that the function $\varphi : D \rightarrow \mathbb{R}$ defined on a convex subset of a linear space is convex, if the inequality

$$\varphi(tx + (1-t)y) \leq t\varphi(x) + (1-t)\varphi(y)$$

holds for all $x, y \in D$ and for all $t \in [0, 1]$.

In this paper we consider convex functions fulfilling some stronger condition (cf. [3, 6]).

Definition 1.1. Let $D \subset \mathbb{R}^n$ be a convex set and let $c > 0$. We say that the function $\varphi : D \rightarrow \mathbb{R}$ is *strongly convex with modulus c* , if

$$\varphi(tx + (1-t)y) \leq t\varphi(x) + (1-t)\varphi(y) - ct(1-t)\|x - y\|^2 \quad (1.1)$$

for all $x, y \in D$ and for all $t \in [0, 1]$.

Obviously every strongly convex function is convex. Observe also that, for instance, affine functions are not strongly convex, because they fulfil (1.1) only with $x = y$.

Strong convexity has a nice characterization ([3, p. 73, Proposition 1.1.2]).

Proposition 1.2. *Let $D \subset \mathbb{R}^n$ be a convex set. The function $\varphi : D \rightarrow \mathbb{R}$ is strongly convex with modulus c if and only if the function $\varphi - c\|\cdot\|^2$ is convex.*

To prove this result it is enough to use the formula $\|x\| = \sqrt{\langle x, x \rangle}$, $x \in \mathbb{R}^n$. Thus strong convexity can be considered also for functions defined on convex subsets of inner product spaces with exactly the same characterization. Going a step further, we could replace in Definition 1.1 the Euclidean space \mathbb{R}^n with any normed space \mathcal{X} . In this setting it is worth mentioning that if the statement of Proposition 1.2 holds, then \mathcal{X} is necessarily the inner product space. It was recently proved in [5].

The goal of this paper is to give some probabilistic interpretations of strong convexity. First let us rephrase standard convexity in the language of random variables. Given a random variable X , by $\mathbb{E}[X]$ we denote its expectation and by $\mathbb{D}^2[X]$ its variance. We will always assume that all random variables are real-valued and non-degenerate and their expectations do exist. One of the most familiar and elementary inequalities in the probability theory reads as follows:

$$\mathbb{E}[f(X)] \geq f(\mathbb{E}[X]), \quad (1.2)$$

where f is convex over the convex hull of the range of the random variable X (see [2]). Conversely, if (1.2) holds, then f is a convex function.

2. RESULTS

Let $\mathcal{I} \subset \mathbb{R}$ be an interval. We have the following probabilistic characterization of strong convexity.

Theorem 2.1. *The function $\varphi : \mathcal{I} \rightarrow \mathbb{R}$ is strongly convex with modulus c if and only if*

$$\varphi(\mathbb{E}[X]) \leq \mathbb{E}[\varphi(X)] - c\mathbb{D}^2[X] \quad (2.1)$$

for any integrable random variable taking values in \mathcal{I} .

Proof. By Proposition 1.2, φ is strongly convex with modulus c if and only if $g(x) = \varphi(x) - cx^2$ is convex, which, by (1.2), is equivalent to

$$\varphi(\mathbb{E}[X]) - c(\mathbb{E}[X])^2 \leq \mathbb{E}[\varphi(X)] - c\mathbb{E}[X^2],$$

This inequality can be rewritten as

$$\varphi(\mathbb{E}[X]) \leq \mathbb{E}[\varphi(X)] - c(\mathbb{E}[X^2] - (\mathbb{E}[X])^2).$$

Because $\mathbb{D}^2[X] = \mathbb{E}[X^2] - (\mathbb{E}[X])^2$, the proof is complete. \square

Now let us turn attention to some particular cases of the “if part” of Theorem 2.1.

For the arbitrary $t \in (0, 1)$ and $x_1, x_2 \in \mathcal{I}$ consider the random variable X such that $P(X = x_1) = t$, $P(X = x_2) = 1 - t$. Then $\mathbb{E}[X] = \bar{x} = tx_1 + (1 - t)x_2$ and $\mathbb{D}^2[X] = t(x_1 - \bar{x})^2 + (1 - t)(x_2 - \bar{x})^2$. Hence we obtain some inequality, which, in fact, is equivalent to the inequality (1.1) defining strong convexity.

Corollary 2.2. *The function $\varphi : \mathcal{I} \rightarrow \mathbb{R}$ is strongly convex with modulus c if and only if*

$$\varphi(tx_1 + (1-t)x_2) \leq t\varphi(x_1) + (1-t)\varphi(x_2) - c(t(x_1 - \bar{x})^2 + (1-t)(x_2 - \bar{x})^2)$$

for any $x_1, x_2 \in \mathcal{I}$ and $t \in (0, 1)$.

The next result we state concerns the Jensen–type inequality for strongly convex functions.

Corollary 2.3. *If the function $\varphi : \mathcal{I} \rightarrow \mathbb{R}$ is strongly convex with modulus c , then*

$$\varphi\left(\sum_{i=1}^n t_i x_i\right) \leq \sum_{i=1}^n t_i \varphi(x_i) - c \sum_{i=1}^n t_i (x_i - \bar{x})^2$$

for any $x_1, \dots, x_n \in \mathcal{I}$ and $t_1, \dots, t_n > 0$ summing up to 1.

Proof. Let X be a random variable such that $P(X = x_i) = t_i$, $i = 1, \dots, n$. Then

$$\mathbb{E}[X] = \bar{x} = \sum_{i=1}^n t_i x_i, \quad \mathbb{D}^2[X] = \sum_{i=1}^n t_i (x_i - \bar{x})^2.$$

Now it is enough to use Theorem 2.1. □

By the similar reasoning we arrive at the integral Jensen–type inequality for strongly convex functions.

Corollary 2.4. *Let (Ω, Σ, μ) be a probability measure space, $h : \Omega \rightarrow \mathcal{I}$ be a Lebesgue integrable function and let $\varphi : \mathcal{I} \rightarrow \mathbb{R}$ be a strongly convex function with modulus c . Then*

$$\varphi\left(\int_{\Omega} h \, d\mu\right) \leq \int_{\Omega} (\varphi \circ h) \, d\mu - c \int_{\Omega} (h - m)^2 \, d\mu,$$

where $m = \int_{\Omega} h \, d\mu$.

Proof. By Proposition 1.2 the function $g(x) = \varphi(x) - cx^2$ is convex. It is enough to apply to g the integral Jensen inequality

$$g\left(\int_{\Omega} h \, d\mu\right) \leq \int_{\Omega} (g \circ h) \, d\mu$$

and observe that

$$\int_{\Omega} h^2 \, d\mu - \left(\int_{\Omega} h \, d\mu\right)^2 = \int_{\Omega} (h - m)^2 \, d\mu.$$

□

The above two results were recently proved in [4] by using the support technique.

3. GEOMETRICAL INTERPRETATIONS

Fix $c > 0$ and for arbitrary $a, b \in \mathbb{R}$ consider the function $h(x) = cx^2 + ax + b$. Take $x_1, x_2 \in \mathcal{I}$ and the random variable X such that $P(X = x_1) = t$, $P(X = x_2) = 1 - t$, where $0 < t < 1$. On the Figure 1 we can see that the expectation $\mathbb{E}[h(X)]$ lies on the chord joining points $(x_1, h(x_1))$ and $(x_2, h(x_2))$. Moreover, the quantity

$$\mathbb{E}[h(X)] - h(\mathbb{E}[X]) = c\mathbb{D}^2[X]$$

is independent on a and b (because, geometrically speaking, we could translate the picture to another place).

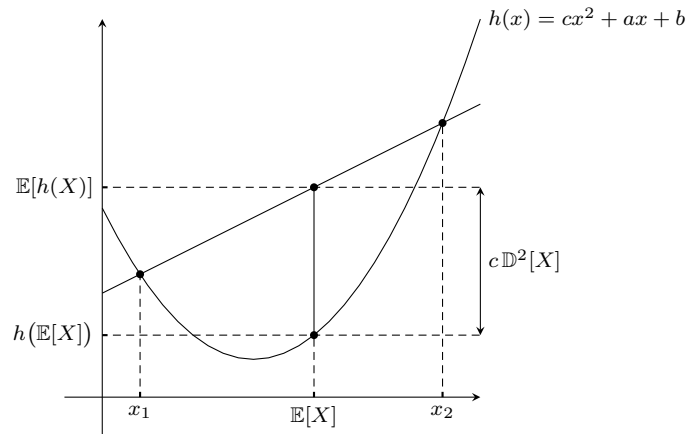


Fig. 1. The geometrical interpretation of variance

Now take $x_1, x_2 \in \mathcal{I}$ with $x_1 < x_2$ and fix $x_0 \in (x_1, x_2)$. For $t \in (0, 1)$ such that $x_0 = tx_1 + (1 - t)x_2$ and for the random variable X constructed as above we have $\mathbb{E}[X] = x_0$. Let the function $\varphi : \mathcal{I} \rightarrow \mathbb{R}$ be strongly convex with modulus c . We choose the constants a, b such that for $h(x) = cx^2 + ax + b$ there is $h(x_1) = \varphi(x_1)$, $h(x_2) = \varphi(x_2)$. Using the interpretation given on Figure 1 we can easily see that $c\mathbb{D}^2[X] = \mathbb{E}[h(X)] - h(\mathbb{E}[X])$. Using the inequality (2.1) we arrive at

$$\mathbb{E}[\varphi(X)] - \varphi(\mathbb{E}[X]) \geq c\mathbb{D}^2[X] = \mathbb{E}[h(X)] - h(\mathbb{E}[X]). \quad (3.1)$$

By the construction (see also Figure 2) we have $\mathbb{E}[h(X)] = \mathbb{E}[\varphi(X)]$. Hence $\varphi(\mathbb{E}[X]) \leq h(\mathbb{E}[X])$, which means that $\varphi(x_0) \leq h(x_0)$. The geometrical interpretation of this inequality is shown on Figure 2: the graph of a strongly convex function (with modulus c) between any $x_1, x_2 \in \mathcal{I}$ lies below the graph of its quadratic interpolant $h(x) = cx^2 + ax + b$ at the points x_1, x_2 . This also shows the connections between strong convexity and generalized convexity in the sense of Beckenbach (cf. [1]): any strongly convex function with modulus c is convex with respect to a two-parameter family of quadratic functions $\{x \mapsto cx^2 + ax + b : a, b \in \mathbb{R}\}$. This is proved and explained in detail in the paper [4].

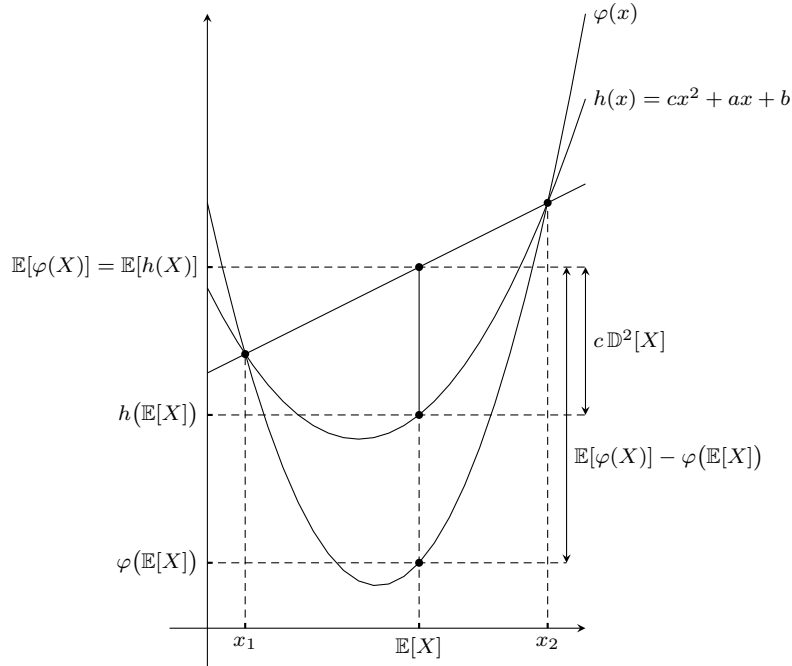


Fig. 2. The geometrical interpretation of strong convexity

Observe now that the inequality (3.1), as a consequence of Theorem 2.1, holds in fact for any integrable random variable taking values in \mathcal{I} . Its left hand side equals to the so-called *Jensen gap* of φ (which is strongly convex with modulus c), while the right hand side is the Jensen gap of an arbitrary quadratic function of the form $h(x) = cx^2 + ax + b$ (this gap is independent on a and b). Thus inequality (3.1) means that the Jensen gap of any strongly convex function with modulus c is greater or equal to the Jensen gap of any quadratic polynomial with leading coefficient c . Figure 3 illustrates it for a random variable with discrete distribution and Figure 4 — for a random variable with continuous distribution.

Notice that on Figure 4 the locations of the points $\mathbb{E}[X]$ and $\varphi(\mathbb{E}[X])$ are determined by the density functions of the appropriate random variables (which are drawn as dotted lines).

It is interesting that the converse is also true: if the Jensen gap of some function φ is (for any random variable X) not less than the Jensen gap of any quadratic polynomial with leading coefficient c , then φ is necessarily strongly convex with modulus c . It easily follows by Theorem 2.1.

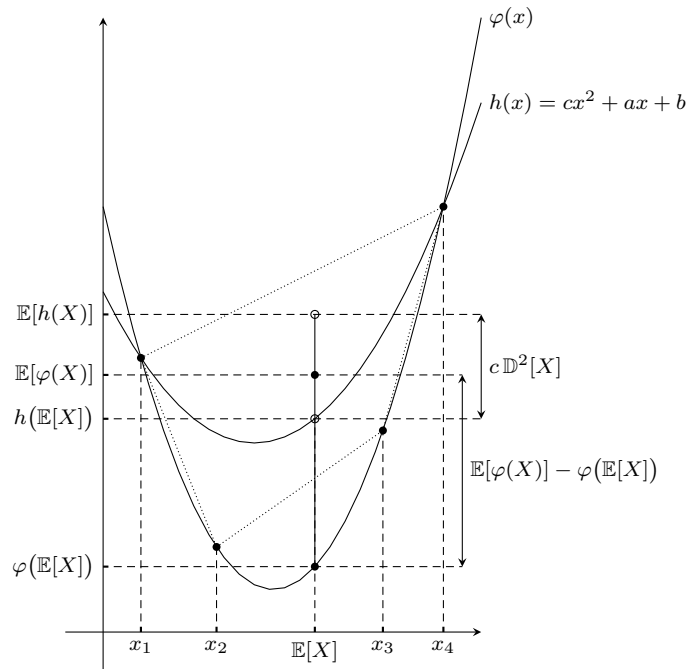


Fig. 3. The inequality between Jensen gaps: discrete distribution

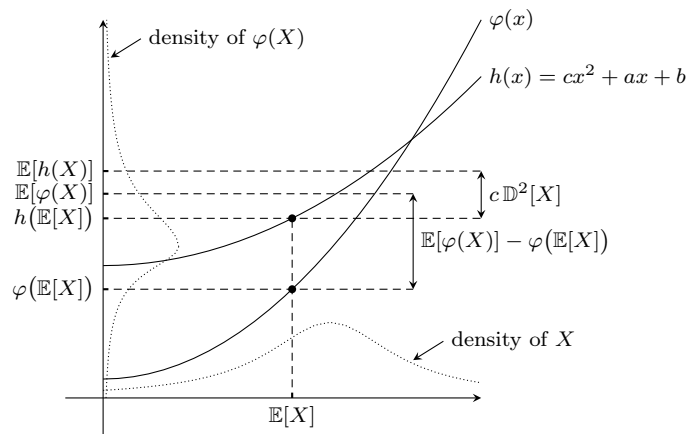


Fig. 4. The inequality between Jensen gaps: continuous distribution

Acknowledgments

The research of the second author was supported by the University of Bielsko-Biała internal grant no. 08/II/GW/2009

REFERENCES

- [1] E.F. Beckenbach, *Generalized convex functions*, Bull. Amer. Math. Soc. **43** (1937), 363–371.
- [2] P. Billingsley, *Probability and Measure*, John Wiley & Sons, New York, 1995.
- [3] J.B. Hiriart-Urruty, C. Lemaréchal, *Fundamentals of convex analysis*, Springer-Verlag, Berlin, Heidelberg 2001.
- [4] N. Merentes, K. Nikodem, *Remarks on strongly convex functions*, Aequationes Math. **80** (2010), 193–199.
- [5] K. Nikodem, Zs. Páles, *Characterizations of inner product spaces by strongly convex functions*, Banach J. Math. Anal. **5** (2011) 1, 83–87.
- [6] A.W. Roberts, D.E. Varberg, *Convex Functions*, Academic Press, New York-London, 1973.

Teresa Rajba
trajba@ath.bielsko.pl

University of Bielsko-Biała
Department of Mathematics and Computer Science
Willowa 2, 43–309 Bielsko-Biała, Poland

Szymon Wąsowicz
swasowicz@ath.eu

University of Bielsko-Biała
Department of Mathematics and Computer Science
Willowa 2, 43–309 Bielsko-Biała, Poland

Received: April 23, 2010.

Accepted: June 6, 2010.