Aequat. Math. 89 (2015), 1389–1400

0001-9054/15/051389-12
published online November 23, 2014
DOI 10.1007/s00010-014-0319-x

Aequationes Mathematicae

Relative convexity and its applications

Constantin P. Niculescu and Ionel Rovenţa

Abstract. We discuss a rather general condition under which the inequality of Jensen works for certain convex combinations of points not all in the domain of convexity of the function under attention. Based on this fact, an extension of the Hardy–Littlewood–Pólya theorem of majorization is proved and a new insight is given into the problem of risk aversion in mathematical finance.

Mathematics Subject Classification. Primary 26B25, Secondary 26D15.

Keywords. Convex function, Supporting hyperplane, Positive measure, Doubly stochastic matrix.

1. Introduction

The important role played by the classical inequality of Jensen in mathematics, probability theory, economics, statistical physics, information theory etc. is well known. See the books of Niculescu and Persson [12], Pečarić et al. [18] and Simon [19]. The aim of this paper is to discuss a rather general condition under which the inequality of Jensen works in a framework that includes a large variety of nonconvex functions and to provide on this basis applications to majorization theory and mathematical finance.

The possibility to extend the inequality of Jensen outside the framework of convex functions was first noticed twenty years ago by Dragomirescu and Ivan [4]. Later, Pearce and Pečarić [17] and Czinder and Páles [3] have considered the special case of mixed convexity (assuming the symmetry of the graph with respect to the inflection point). For related results, see the papers of Florea and Niculescu [5], Niculescu and Spiridon [14], and Mihai and Niculescu [11].

The inequality of Jensen characterizes the behavior of a continuous convex function with respect to a mass distribution on its domain. More precisely, if f is a continuous convex function on a compact convex subset K of $\mathbb{R}^N$ and μ is a Borel probability measure on K having the barycenter

$$b_\mu = \int_K x d\mu(x),$$

then the value of f at b_μ does not exceed the mean value of f over K, that is,

$$f(b_\mu) \leq \int_K f(x) d\mu(x).$$

A moment's reflection reveals that the essence of this inequality is the fact that b_μ is a point of convexity of f relative to its domain K. The precise meaning of the notion of point of convexity is given in Definition 1.1 below, which is stated in the framework of real-valued continuous functions f defined on a compact convex subset K of $\mathbb{R}^N$.

Definition 1.1. A point $a \in K$ is a *point of convexity* of the function f relative to the convex subset V of K if $a \in V$ and

$$f(a) \leq \sum_{k=1}^{n} \lambda_k f(x_k), \tag{J}$$

for every family of points $x_1, \ldots, x_n$ in V and every family of positive weights $\lambda_1, \ldots, \lambda_n$ with $\sum_{k=1}^{n} \lambda_k = 1$ and $\sum_{k=1}^{n} \lambda_k x_k = a$.

The point a is a *point of concavity* if it is a point of convexity for $-f$ (equivalently, if the inequality (J) works in the reversed way).

In what follows, the set V that appears in Definition 1.1 will be referred to as a *neighborhood of convexity* of a. Here, the term of *neighborhood* has an extended meaning and is not necessarily ascribed to the topology of $\mathbb{R}^N$. In order to avoid the trivial case where $V = \{a\}$, we will always assume that V is an infinite set; for example, this happens when a belongs to the relative interior of V (the interior within the affine hull of V).

For the function $f(x, y) = x^2 - y^2$, the origin is a point of convexity relative to the Ox axis, and a point of concavity relative to the Oy axis. With respect to the plane topology, both axes have empty interior.

If a function $f : K \to \mathbb{R}$ is convex, then every point of K is a point of convexity relative to the whole domain K (and this fact characterizes the property of convexity of f).

Definition 1.1 is motivated mainly by the existence of nonconvex functions that admit points of convexity relative to the whole domain (or at least to a neighborhood of convexity where the function is not convex). An illustration is offered by the nonconvex function $g(x) = \left|x^2 - 1\right|$; all points in $(-\infty, -1] \cup [1, \infty)$ are points of convexity relative to the entire real set $\mathbb{R}$.

Every point of local minimum of a continuous function $f : [0, 1] \to \mathbb{R}$ is a point of convexity. Thus, every nowhere differentiable continuous function $f : [0, 1] \to \mathbb{R}$ admits points of convexity despite the fact that it is not convex on any nondegenerate interval.

The idea of point of convexity is not entirely new. In an equivalent form, it is present in the paper of Dragomirescu and Ivan [4]. The technique of convex minorants, described by Steele [20] at pp. 96–99, is also close to the concept of point of convexity.

A different concept of *punctual convexity* is discussed in the recent paper of Florea and Păltănea [6].

2. The existence of points of convexity

The following lemma indicates a simple geometric condition under which a point is a point of convexity relative to the whole domain.

Lemma 2.1. *Assume that f is a real-valued continuous function defined on a compact convex subset K of $\mathbb{R}^N$. If f admits a supporting hyperplane at a point a, then a is a point of convexity of f relative to K.*

In other words, every point at which the subdifferential is nonempty is a point of convexity.

Proof. Indeed, the existence of a supporting hyperplane at a is equivalent to the existence of an affine function $h(x) = \langle x, v\rangle + c$ (for suitable $v \in \mathbb{R}^N$ and $c \in \mathbb{R}$) such that

$$f(a) = h(a) \text{ and } f(x) \geq h(x) \quad \text{for all } x \in K.$$

If μ is a Borel probability measure, its barycenter is given by the formula

$$b_\mu = \int_K x d\mu(x),$$

so that if $b_\mu = a$, then

$$f(a) = h(a) = h\left(\int_K x d\mu(x)\right) = \int_K h(x) d\mu(x) \leq \int_K f(x) d\mu(x).$$

□

Remark 2.2. Another sufficient condition for convexity at a point, formulated in terms of secant line slopes, can be found in the papers of Niculescu and Stephan [15,16]. However, as shown by the case of polynomials of fourth degree, that condition does not overcome the result of Lemma 2.1.

As is well known, the usual property of convexity assures the existence of a supporting hyperplane at each interior point. See [12], Theorem 3.7.1, p. 128. This explains why Jensen's inequality works nicely in the context of continuous convex functions.

In the case of differentiable functions, the supporting hyperplane is unique and coincides with the tangent hyperplane. For such functions, Lemma 2.1 asserts that every point where the tangent hyperplane lies below/above the graph is a point of convexity/concavity.

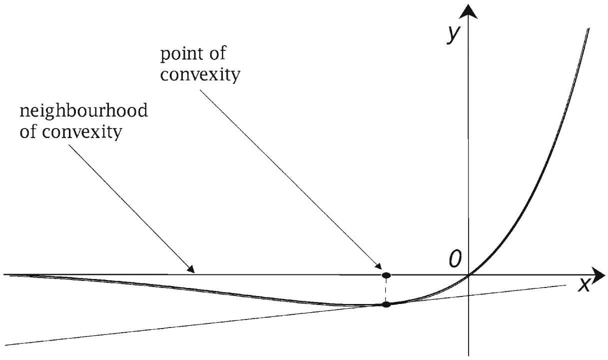

Figure 1. A point of convexity of the function xe^x relative to the whole real axis

Example. In the one real variable case, the existence of points of convexity of a nonconvex differentiable function (such as xe^x, x^2e^{-x}, $\log^2 x$, $\frac{\log x}{x}$ etc.) can be easily proved by looking at the position of the tangent line with respect to the graph.

For example, the function xe^x is concave on $(-\infty, -2]$ and convex on $[-2, \infty)$ (and attains a global minimum at $x = -1$). See Fig. 1.

Based on Lemma 2.1, one can easily show that every point $x \geq -1$ is a point of convexity of f relative to the whole real axis. Therefore

$$\sum_{k=1}^{n} \lambda_k x_k e^{x_k} \geq \left(\sum_{k=1}^{n} \lambda_k x_k \right) e^{\sum_{k=1}^{n} \lambda_k x_k},$$

whenever $\sum_{k=1}^{n} \lambda_k x_k \geq -1$.

In the special case where $\sum_{k=1}^{n} \lambda_k x_k \geq 0$, this inequality can be deduced from Chebyshev's inequality and the convexity of the exponential function. Borwein and Girgensohn [2] proved that

$$\sum_{k=1}^{n} x_k e^{x_k} \geq \frac{\max\{2, e(1-1/n)\}}{n} \sum_{k=1}^{n} x_k^2,$$

for every family of real numbers $x_1, x_2, \ldots, x_n$ such that $\sum_{k=1}^{n} x_k \geq 0$. The extension of their result to the weighted case (subject to the condition $\sum_{k=1}^{n} \lambda_k x_k \geq 0$) is an open problem.

Example. The two real variables function

$$f(x, y) = e^{-x^2-y^2}, \quad (x, y) \in \mathbb{R}^2,$$

exhibits the phenomenon of relative concavity. Indeed, its graph is the rotation graph of the function $z = e^{-x^2}$ around the Oz axis and this makes it possible to apply Lemma 2.1 by means of calculus of one real variable. See Fig. 2.

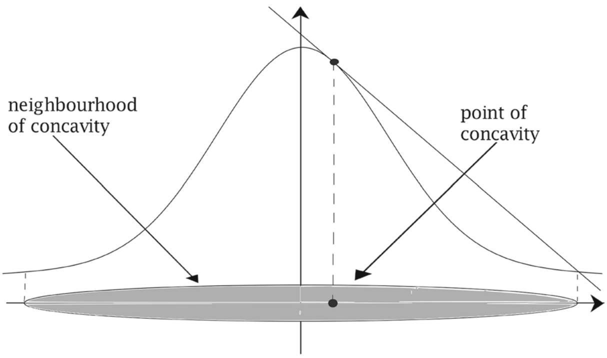

FIGURE 2. A point of concavity of the function $e^{-x^2-y^2}$ and a neighborhood of concavity of it

The convexity properties of the function f can be described in a more convenient way by viewing it as a function of complex variable, via the formula $f(w) = e^{-|w|^2}$.

The function f is strictly concave on the compact disc $\overline{D}_{1/\sqrt{2}}\,(0)$ and attains a global maximum at the origin. The tangent plane at the graph of f at any point $w_0 = (x_0, y_0)$ with $\|w_0\| \leq 1/2$, is above the graph over a neighborhood of w_0 including the closed disc $\overline{D}_{r^*}\,(0)$, where $r^* = 1.183\,802\ldots$ is the solution of the equation $e^{-1/4}(\frac{3}{2} - x) = e^{-x^2}$. As a consequence,

$$\sum_{k=1}^{n} \lambda_k e^{-|w_k|^2} \leq e^{-M^2}$$

for all points $w_1, \ldots, w_n \in \overline{D}_{r^*}\,(0)$ and all $\lambda_1, \ldots, \lambda_n > 0$ such that $\sum_{k=1}^{n} \lambda_k = 1$ and $\left|\sum_{k=1}^{n} \lambda_k w_k\right| = M \leq 1/2$. Notice that Jensen's inequality yields this conclusion only when $w_1, \ldots, w_n \in \overline{D}_{1/\sqrt{2}}\,(0)$.

The real variable case also has nontrivial implications in the case of matrix functions. The function $F(X) = \text{trace}(f(X))$ is convex/concave on the linear space $\text{Sym}(n, \mathbb{R})$, of all self-adjoint (that is, symmetric) matrices in $\text{M}_n(\mathbb{R})$, whenever $f : \mathbb{R} \to \mathbb{R}$ is convex/concave. See the paper of Lieb and Pedersen [8] for details. Thus, in the case of $f(x) = xe^x$, the function F is concave on the convex set $\text{Sym}_{sp\subset(-\infty,-2]}(n, \mathbb{R})$, of all symmetric matrices in $\text{M}_n(\mathbb{R})$ whose spectrum is included in $(-\infty, -2]$; the function F is convex on the set $\text{Sym}_{sp\subset[-2,\infty)}(n, \mathbb{R})$, of all symmetric matrices in $\text{M}_n(\mathbb{R})$ whose spectrum is included in $(-2, \infty]$.

The following result is a direct consequence of functional calculus with self-adjoint matrices. If $\lambda_1, \ldots, \lambda_n$ are positive numbers such that $\sum_{k=1}^{n} \lambda_k = 1$

and $A_1, \ldots, A_n$ are matrices in $\mathrm{Sym}_{sp\subset(-\infty,-2]}(n,\mathbb{R}) \cup \mathrm{Sym}_{sp\subset[-2,\infty)}(n,\mathbb{R})$ such that $\sum_{k=1}^{n} \lambda_k A_k \geq -I_n$, then

$$\sum_{k=1}^{n} \lambda_k \operatorname{trace}\left(A_k e^{A_k}\right) \geq \operatorname{trace}\left[\left(\sum_{k=1}^{n} \lambda_k A_k\right) e^{\sum_{k=1}^{n} \lambda_k A_k}\right].$$

3. The extension of Hardy–Littlewood–Pólya theorem of majorization

The notion of point of convexity leads to a very large generalization of the Hardy–Littlewood–Pólya theorem of majorization. Given a vector $\mathbf{x} = (x_1, \ldots, x_N)$ in $\mathbb{R}^N$, let $\mathbf{x}^{\downarrow}$ be the vector with the same entries as $\mathbf{x}$ but rearranged in decreasing order,

$$x_1^{\downarrow} \geq \cdots \geq x_N^{\downarrow}.$$

The vector $\mathbf{x}$ is *majorized* by $\mathbf{y}$ (abbreviated, $\mathbf{x} \prec \mathbf{y}$) if

$$\sum_{i=1}^{k} x_i^{\downarrow} \leq \sum_{i=1}^{k} y_i^{\downarrow} \quad \text{for } k = 1, \ldots, N-1$$

and

$$\sum_{i=1}^{N} x_i^{\downarrow} = \sum_{i=1}^{N} y_i^{\downarrow}.$$

The concept of majorization admits an order-free characterization based on the notion of doubly stochastic matrix. Recall that a matrix $A \in \mathrm{M}_n(\mathbb{R})$ is *doubly stochastic* if it has nonnegative entries and each row and each column sums to unity.

Theorem 3.1. (Hardy, Littlewood and Pólya [7]). *Let* $\mathbf{x}$ *and* $\mathbf{y}$ *be two vectors in* $\mathbb{R}^N$*, whose entries belong to an interval* I*. Then the following statements are equivalent:*

- (a) $\mathbf{x} \prec \mathbf{y}$;
- (b) *there is a doubly stochastic matrix* $A = (a_{ij})_{1\leq i,j\leq N}$ *such that* $\mathbf{x} = A\mathbf{y}$;
- (c) *the inequality* $\sum_{i=1}^{N} f(x_i) \leq \sum_{i=1}^{N} f(y_i)$ *holds for every continuous convex function* $f : I \to \mathbb{R}$.

An alternative characterization of the relation of majorization is given by the Schur–Horn theorem: $\mathbf{x} \prec \mathbf{y}$ in $\mathbb{R}^N$ if and only if the components of $\mathbf{x}$ and $\mathbf{y}$ are respectively the diagonal elements and the eigenvalues of a self-adjoint matrix. The details can be found in the book of Marshall, Olkin and Arnold [10], pp. 300–302.

The notion of majorization is generalized by weighted majorization, which refers to probability measures rather than vectors. This is done by identifying any vector $\mathbf{x} = (x_1, \ldots, x_N)$ in $\mathbb{R}^N$ with the probability measure $\frac{1}{N}\sum_{i=1}^{N} \delta_{x_i}$, where δ_{x_i} denotes the Dirac measure concentrated at x_i.

We define the relation of majorization

$$\sum_{i=1}^{m} \lambda_i \delta_{\mathbf{x}_i} \prec \sum_{j=1}^{n} \mu_j \delta_{\mathbf{y}_j}, \tag{2}$$

between two positive discrete measures supported at points in $\mathbb{R}^N$, by requiring the existence of an $m \times n$-dimensional matrix $A = (a_{ij})_{i,j}$ such that

$$a_{ij} \geq 0, \qquad \text{for all } i, j \tag{3}$$

$$\sum_{j=1}^{n} a_{ij} = 1, \qquad i = 1, \ldots, m \tag{4}$$

$$\mu_j = \sum_{i=1}^{m} a_{ij}\lambda_i, \quad j = 1, \ldots, n \tag{5}$$

and

$$\mathbf{x}_i = \sum_{j=1}^{n} a_{ij}\mathbf{y}_j, \quad i = 1, ..., m. \tag{6}$$

The matrices verifying the conditions 3 and 4 are called *stochastic on rows*. When $m = n$ and all weights λ_i and μ_j are equal to each other, the condition (5) assures the *stochasticity on columns*, so in that case we deal with doubly stochastic matrices.

We are now in a position to state the following generalization of the Hardy–Littlewood–Pólya theorem of majorization:

Theorem 3.2. *Suppose that f is a real-valued function defined on a compact convex subset K of $\mathbb{R}^N$ and $\sum_{i=1}^{m} \lambda_i \delta_{\mathbf{x}_i}$ and $\sum_{j=1}^{n} \mu_j \delta_{\mathbf{y}_j}$ are two positive discrete measures concentrated at points in K. If $\mathbf{x}_1, \ldots, \mathbf{x}_m$ are points of convexity of f relative to K and*

$$\sum_{i=1}^{m} \lambda_i \delta_{\mathbf{x}_i} \prec \sum_{j=1}^{n} \mu_j \delta_{\mathbf{y}_j},$$

then

$$\sum_{i=1}^{m} \lambda_i f(\mathbf{x}_i) \leq \sum_{j=1}^{n} \mu_j f(\mathbf{y}_j). \tag{7}$$

Proof. By our hypothesis, there exists an $m \times n$-dimensional matrix $A = (a_{ij})_{i,j}$ that is stochastic on rows and verifies the conditions (5) and (6). The last condition shows that each point $\mathbf{x}_i$ is the barycenter of the probability measure $\sum_{j=1}^n a_{ij}\delta_{\mathbf{y}_j}$. By Jensen's inequality, we infer that

$$f(\mathbf{x}_i) \leq \sum_{j=1}^n a_{ij} f(\mathbf{y}_j).$$

Multiplying each side by λ_i and then summing up over i from 1 to m, we conclude that

$$\sum_{i=1}^m \lambda_i f(\mathbf{x}_i) \leq \sum_{i=1}^m \left(\lambda_i \sum_{j=1}^n a_{ij} f(\mathbf{y}_j)\right) = \sum_{j=1}^n \left(\sum_{i=1}^m a_{ij}\lambda_i\right) f(\mathbf{y}_j) = \sum_{j=1}^n \mu_j f(\mathbf{y}_j),$$

and the proof of (7) is done. □

Example. The well known Gauss–Lucas theorem on the distribution of the critical points of a polynomial asserts that the roots $(\mu_k)_{k=1}^{n-1}$ of the derivative P' of any complex polynomial $P \in \mathbb{C}[z]$ of degree $n \geq 2$ lie in the smallest convex polygon containing the roots $(\lambda_j)_{j=1}^n$ of the polynomial P. This led Malamud [9] to the interesting remark that the two families of roots are actually related by the relation of majorization. Based on this remark, he was able to prove the following conjecture raised by de Bruijn and Springer in 1947: for any convex function $f : \mathbb{C} \to \mathbb{R}$ and any polynomial P of degree $n \geq 2$,

$$\frac{1}{n-1}\sum_{k=1}^{n-1} f(\mu_k) \leq \frac{1}{n}\sum_{j=1}^n f(\lambda_j),$$

where $(\lambda_j)_{j=1}^n$ and $(\mu_k)_{k=1}^{n-1}$ are respectively the roots of P and P'.

Theorem 3.2 allows us to relax the condition of convexity by asking only that all the roots μ_k of P' be points of convexity for f. According to a remark above concerning the function $e^{-|w|^2}$, this implies that

$$\frac{1}{n-1}\sum_{k=1}^{n-1} e^{-|\mu_k|^2} \geq \frac{1}{n}\sum_{j=1}^n e^{-|\lambda_j|^2},$$

whenever the roots $\mu_1, \ldots, \mu_{n-1}$ belong to $\overline{D}_{1/2}(0)$ and $\lambda_1, \ldots, \lambda_n$ belong to $\overline{D}_{1.18}(0)$. An example of a polynomial verifying these conditions is $P(z) = 4z^3 - 3z$.

Example. A second application of Theorem 3.2 refers to the function $f(x) = \log^2 x$. This function is convex on the interval $(0, e]$ and concave on $[e, \infty)$. The Hardy–Littlewood–Pólya theorem of majorization easily yields the implication

$$(x_1, \ldots, x_n) \prec (y_1, \ldots, y_n) \Rightarrow \sum_{i=1}^n \log^2 x_i \leq \sum_{i=1}^n \log^2 y_i \tag{8}$$

whenever $x_1, \ldots, x_n$ and $y_1, \ldots, y_n$ belong to $(0, e]$. According to Lemma 2.1, all points in $(0, 2]$ are points of convexity of f relative to $(0, a^*]$, where

$$a^* = 5.495\,869\,874\ldots$$

is the solution of the equation $\log^2 x - \log^2 2 = (\log 2)(x - 2)$. By Theorem 3.2, the implication (8) still works when $x_1, \ldots, x_n \in (0, 2]$ and $y_1, \ldots, y_n \in (0, a^*]$. Recently, Bîrsan, Neff and Lankeit [1] noticed still another case where an inequality of the form (8) holds true. Precisely, they proved that for every two triplets x_1, x_2, x_3 and y_1, y_2, y_3 of positive numbers which satisfy the conditions

$$x_1 + x_2 + x_3 \leq y_1 + y_2 + y_3, \quad x_1x_2 + x_2x_3 + x_3x_1 \leq y_1y_2 + y_2y_3 + y_3y_1$$

and $x_1x_2x_3 = y_1y_2y_3$, we have

$$\sum_{i=1}^{3} \log^2 x_i \leq \sum_{i=1}^{3} \log^2 y_i.$$

This suggests a new concept of majorization for n-tuples of positive elements, based on elementary symmetric functions. As it is beyond the scope of this paper, we will not go into the details.

Theorem 3.2 provides the following extension of Popoviciu's inequality:

Theorem 3.3. *Suppose that f is a real-valued function defined on an interval I. If a, b, c belong to I and $\frac{a+b}{2}$, $\frac{a+c}{2}$ and $\frac{b+c}{2}$ are points of convexity of f relative to the entire interval I, then*

$$\frac{f(a) + f(b) + f(c)}{3} + f\left(\frac{a+b+c}{3}\right) \geq \frac{2}{3}\left[f\left(\frac{a+b}{2}\right) + f\left(\frac{a+c}{2}\right) + f\left(\frac{b+c}{2}\right)\right]. \quad (9)$$

Proof. Without loss of generality we may assume that $a \geq b \geq c$. Then

$$\frac{a+b}{2} \geq \frac{a+c}{2} \geq \frac{b+c}{2} \quad \text{and} \quad a \geq \frac{a+b+c}{3} \geq c.$$

We attach to the points a, b, c two sextic families of points:

$$x_1 = x_2 = \frac{a+b}{2},\ x_3 = x_4 = \frac{a+c}{2},\ x_5 = x_6 = \frac{b+c}{2}$$
$$y_1 = a,\ y_2 = y_3 = y_4 = \frac{a+b+c}{3},\ y_5 = b,\ y_6 = c$$

if $a \geq (a + b + c)/3 \geq b \geq c$, and

$$x_1 = x_2 = \frac{a+b}{2}, \quad x_3 = x_4 = \frac{a+c}{2}, \quad x_5 = x_6 = \frac{b+c}{2}$$
$$y_1 = a, \quad y_2 = b, \quad y_3 = y_4 = y_5 = \frac{a+b+c}{3}, \quad y_6 = c$$

if $a \geq b \geq (a+b+c)/3 \geq c$. In both cases $\frac{1}{6}\sum_{i=1}^{6}\delta_{x_i} \prec \frac{1}{6}\sum_{i=1}^{6}\delta_{y_i}$, and thus the inequality (9) follows from Theorem 3.2. $\square$

Popoviciu noticed that under the presence of continuity, the inequality (9) works for all triplets $a, b, c \in I$ if and only if the function f is convex. See [12], p. 12. Theorem 3.3 allows this inequality to work for *certain* triplets a, b, c even when f is not convex. For example, this is the case for the function $\log^2 x$, and all points $a, b, c \in (0, a^*]$ such that $\frac{a+b}{2}, \frac{a+c}{2}, \frac{b+c}{2} \in (0, 2]$.

Remark 3.4. The theory of points of convexity and our generalization of the Hardy–Littlewood–Pólya theorem stated in Theorem 3.2 extend verbatim to the context of spaces with global nonpositive curvature. See [13] for the theory of convex functions on such spaces.

4. An application to mathematical finance

In the context of probability theory, Jensen's inequality is generally stated in the following form: if X is a random variable and f is a continuous convex function on an open interval containing the range of X, then

$$f(E(X)) \leq E(f(X)),$$

provided that both expectations $E(X)$ and $E(f(X))$ exist and are finite.

A nice illustration of this inequality in mathematical finance refers to the so called *risk aversion*, the reluctance of someone who wants to invest his life savings into a stock that may have high expected returns (but also involves a chance of losing value), preferring to put his or her money into a bank account with a low but guaranteed interest rate. Indeed, if the utility function f is concave, then

$$f(E(X)) \geq E(f(X)).$$

Using the technique of pushing-forward measures (i.e., of image measures), we will show that this inequality still works when f is continuous and $E(X)$ is a point of concavity of f relative to its whole domain. This follows from the following technical result.

Theorem 4.1. *Suppose that f is a real-valued continuous function defined on an open interval I and X is a random variable associated with a probability space (Ω, Σ, μ) such that*

(i) *the range of X is included in the interval I;*
(ii) *the expectations $E(X)$ and $E(f(X))$ exist and are finite;*
(iii) *$E(X)$ is a point of convexity of f relative to I.*

Then

$$f(E(X)) \leq \int_{\Omega} f(X(\omega))d\mu(\omega).$$

Proof. Since $X : \Omega \to I$ is a μ-integrable map, the push-forward measure ν, given by the formula $\nu(A) = \mu(X^{-1}(A))$, is a Borel probability measure on I with barycenter $b_\nu = \int_\Omega X(\omega)d\mu(\omega) = E(X)$. We have to prove that

$$f(b_\nu) \leq \int_I f(x)d\nu(x).$$

When ν is a discrete measure, this follows from the fact that b_ν is a point of convexity. If the range of X is included in a compact subinterval K of I, then the support of ν is included in K and we have to use the following approximation argument proved in [12], Lemma 4.1.10, p. 183: every Borel probability measure ν on a compact convex set K is the pointwise limit of a net of discrete Borel probability measures ν_α on K, each having the same barycenter as ν.

In the general case, we approximate X by the sequence of bounded random variables $X_n = \sup \{\inf \{X, n\}, -n\}$. □

5. Concluding remarks

In this paper we introduced the concept of convexity at a point relative to a convex subset of the domain. This fact made Jensen's inequality available to a large variety of nonconvex functions and shed new light on the Hardy–Littlewood–Pólya theorem of majorization. In turn, the probabilistic form of Jensen's inequality (as stated in Theorem 4.1) is put in a more general perspective of the problem of risk aversion.

Most likely the notion of convexity at a point could have a practical purpose in optimization theory, information theory, the design of communication systems etc.

Acknowledgements

The first author was supported by a grant of the Romanian National Authority for Scientific Research, CNCS—UEFISCDI, project number PN-II-ID-PCE-2011-3-0257. The second author was supported by the strategic grant POSDRU/159/1.5/S/133255, Project ID 133255 (2014), co-financed by the European Social Fund within the Sectorial Operational Program Human Resources Development 2007–2013.

References

[1] Bîrsan, M., Neff, P., Lankeit, J.: Sum of squared logarithms—an inequality relating positive definite matrices and their matrix logarithm, J. Inequal. Appl., paper no. 168. (2013)

[2] Borwein, J., Girgensohn, R.: A class of exponential inequalities. Math. Inequal. Appl. **6**(3), 397–411 (2003)

[3] Czinder, P., Páles, Z.: An extension of the Hermite-Hadamard inequality and an application for Gini and Stolarsky means. J. Inequal. Pure Appl. Math. **5**(2), Article no. 42. (2004)

[4] Dragomirescu, M., Ivan, C.: Convexity-like inequalities for averages in a convex set. Aequ. Math. **45**, 179–194 (1993)

[5] Florea, A., Niculescu, C.P.: A Hermite-Hadamard inequality for convex-concave symmetric functions, Bull. Math. Soc. Sci. Math. Roumanie **50**(98), no. 2, pp. 149–156 (2007)

[6] Florea, A., Păltănea, E.: On a class of punctual convex functions. Math. Inequal. Appl. **17**(1), 389–399 (2014)

[7] Hardy, G.H., Littlewood, J.E., Pólya, G.: Inequalities. Cambridge Mathematical Library, 2nd Edition, 1952, Reprinted (1988)

[8] Lieb, E.H., Pedersen, G.K.: Convex multivariable trace functions. Rev. Math. Phys. **14**, 631–648 (2002)

[9] Malamud, S.M.: Inverse spectral problem for normal matrices and the Gauss-Lucas theorem, Trans. Amer. Math. Soc. **357**, pp. 4043–4064 (2005) Posted on arXiv:math/0304158 at July 6, (2003)

[10] Marshall, A.W., Olkin, I., Arnold, B.: Inequalities: theory of majorization and its application, 2nd Edition. Springer, New York (2011)

[11] Mihai, M., Niculescu, C.P.: A simple proof of the Jensen type inequality of Fink and Jodeit. Mediterr. J. Math. doi:10.1007/s00009-014-0480-4

[12] Niculescu, C.P., Persson, L.-E.: Convex Functions and their Applications. A Contemporary Approach, CMS Books in Mathematics vol. 23. Springer, New York (2006)

[13] Niculescu, C.P., Rovenţa, I.: An approach of majorization in spaces with a curved geometry. J. Math. Anal. Appl. **411**(1), 119–128 (2014)

[14] Niculescu, C.P., Spiridon, C.: New Jensen-type inequalities. J. Math. Anal. Appl. **401**(1), 343–348 (2013)

[15] Niculescu, C.P., Stephan, H.: A generalization of Lagrange's algebraic identity and connections with Jensen's inequality. Weierstraß—Institut für Angewandte Analysis und Stochastik, Preprint no. 1756/2012.

[16] Niculescu, C.P., Stephan, H.: Lagrange's Barycentric Identity From An Analytic Viewpoint, Bull. Math. Soc. Sci. Math. Roumanie **56**(104), no. 4, pp. 487–496 (2013)

[17] Pearce, C.E.M., Pečarić, J.: On convexity-like inequalities. Revue Roumaine Math. Pure Appl. Math. **42**(1-2), 133–138 (1997)

[18] Pečarić, J., Proschan, F., Tong, Y.L.: Convex functions, partial orderings, and statistical applications. Academic Press, New York (1992)

[19] Simon, B.: Convexity. An analytic viewpoint. Cambridge University Press, Cambridge (2011)

[20] Steele, J.M.: The Cauchy-Schwarz master class: an introduction to the art of mathematical inequalities. Cambridge University Press, Cambridge (2004)

C. P. Niculescu and I. Rovenţa
Department of Mathematics
University of Craiova
200585 Craiova
Romania
e-mail: cpniculescu@gmail.com
e-mail: ionelroventa@yahoo.com

Received: July 16, 2014
Revised: September 23, 2014