J. Math. Anal. Appl. 479 (2019) 903–925

Contents lists available at ScienceDirect

Journal of Mathematical Analysis and Applications

www.elsevier.com/locate/jmaa

A new look at Popoviciu's concept of convexity for functions of two variables

Sorin G. Gal [a,*], Constantin P. Niculescu [b,1]

[a] *Department of Mathematics and Computer Science, University of Oradea, Street Universtatii No. 1, Oradea, 410087, Romania*
[b] *Department of Mathematics, University of Craiova, Craiova 200585, Romania*

ARTICLE INFO

Article history:
Received 25 February 2019
Available online 19 June 2019
Submitted by B.S. Thomson

Keywords:
Divided differences
Differentiable functions
Higher order convex functions
Higher order monotonicity

ABSTRACT

One proves that most results known for the usual convex functions (including Jensen's inequality, Jensen's characterization of convexity, the duality between monotone and convex functions, the existence of affine supports etc.), have analogs for Popoviciu's convex functions. The details are presented in the case of two variables, but the theory exposed here can be easily extended to higher dimensions.

1. Introduction

Many important results in real analysis of one variable are related to the so called divided differences. Given a function f defined on an interval I and a family $x_1, x_2, ..., x_{n+1}$ of distinct points of I, the divided differences of order $0, 1, \ldots, n$ are respectively defined by the formulas

$$[x_1; f] = f(x_1)$$
$$[x_1, x_2; f] = \frac{f(x_1) - f(x_2)}{x_1 - x_2}$$
$$...$$
$$[x_1, x_2, ..., x_{n+1}; f] = \frac{[x_1, x_2, ..., x_n; f] - [x_2, x_3, ..., x_{n+1}; f]}{x_1 - x_{n+1}}.$$

Notice that all these divided differences are invariant under the permutation of points $x_1, x_2, ..., x_{n+1}$.

* Corresponding author.
E-mail addresses: galso@uoradea.ro (S.G. Gal), cpniculescu@gmail.com (C.P. Niculescu).
[1] *Current address:* Academy of Romanian Scientists, Splaiul Independenţei No. 54, Bucharest 050085, Romania.

https://doi.org/10.1016/j.jmaa.2019.06.057
0022-247X/

A function f is respectively nonnegative, nondecreasing or convex when all divided differences of order 0, 1 or 2 are nonnegative, for all increasingly ordered system of points. Indeed, in the later case we have to remark that due to the property of invariance the inequality

$$[x_1, x_2, x_3; f] = \frac{\frac{f(x_1)-f(x_2)}{x_1-x_2} - \frac{f(x_2)-f(x_3)}{x_2-x_3}}{x_1 - x_3} \geq 0 \tag{1.1}$$

holds for all triplets x_1, x_2, x_3 of distinct points in I if and only if it holds for all triplets $x_1 < x_3 < x_2$. Since every intermediate point x_3 can be uniquely represented as a convex combination of the extremities x_1 and x_2, this translates into the usual condition of convexity,

$$f((1-\lambda)x_1 + \lambda x_2) \leq (1-\lambda)f(x_1) + \lambda f(x_2)$$

for all $x_1 < x_2$ in I and all $\lambda \in (0,1)$. When $f \in C^2(I)$, a repeated application of Lagrange's mean value theorem yields the existence of a point $\xi \in (\min_k x_k, \max_k x_k)$ such that

$$[x_1, x_2, x_3; f] = \frac{f''(\xi)}{2}. \tag{1.2}$$

This formula together with the fact that

$$\lim_{x_1 \to x} \left(\lim_{x_2 \to x_1} [x_1, x_2, x; f] \right) = f''(x)/2$$

show that a function $f \in C^2(I)$ is convex if and only if its second derivative is nonnegative.

The above considerations led H. Hopf [14] and T. Popoviciu [21] to initiate the study of higher order convexity on intervals, by calling a function f of one real variable *n-convex* (*respectively n-concave*) if $[x_1, x_2, \ldots, x_{n+1}; f] \geq 0$ (≤ 0) for any $n+1$ distinct points $x_1, x_2, \ldots, x_{n+1}$. There are plenty of examples of such functions. For example, a function of class C^n is n-convex if and only if its nth derivative is nonnegative.

T. Popoviciu went a step further by sketching a higher dimensional analog. His main motivation was interpolation theory and not the study of a specific counterpart to usual convexity. Probably this explains why neither Popoviciu nor his followers made no attempt to see at what extent his new concept of convexity parallels the classical one (referring to convexity along linear segments). It is the aim of the present paper to fill out the gap.

In order to emphasize the main idea rather than technical details, we will restrict ourselves to the two dimensional case but the results obtained can be extended to all dimensions.

In N dimensions, the analog of real intervals are the boxes with edges parallel to the coordinate axes. They can be equally described as products of 1-dimensional intervals or as order intervals in $\mathbb{R}^N$, associated to the coordinatewise ordering. In two dimensions, the boxes are usually called rectangles, so we will work mostly with functions defined on nondegenerate rectangles, supposed to be either open or closed.

However, the theory exposed in this paper works for any function whose domain D play the following property (verified by every open set): if $x \in D$, then there is a nondegenerate compact rectangle R such that $x \in R \subset D$.

In this framework two rather exotic concepts, of monotonicity and convexity, reflecting the geometry of rectangles, will be discussed.

If $f = f(x, y)$ is a function defined on a rectangle $I \times J$, and $x_1, x_2, \ldots, x_m$ are distinct points in I and $y_1, y_2, \ldots, y_n$ are distinct points in J, one defines the *divided double difference* of order two by the formula

$$\begin{bmatrix} x_1, & x_2, & \dots & x_m \\ y_1, & y_2, & \dots & y_n \end{bmatrix}; f \Bigg] = [x_1, x_2, \dots, x_m; [y_1, y_2, \dots, y_n; f(x, \cdot)]] \tag{1.3}$$

$$= [y_1, y_2, \dots, y_n; [x_1, x_2, \dots, x_m; f(\cdot, y)]].$$

It is worth noticing that this formula is invariant under the permutation of variables x_k (and also under the permutation of variables y_k). This property also works when the function f takes values in a Banach space.

Drawing a parallel to the one dimensional case, T. Popoviciu [21], p. 78, has called a function $f : I \times J \to \mathbb{R}$ *convex of order* (m, n) (*box-convex of order* (m, n) in our terminology) if all divided differences

$$\left[\begin{matrix} x_1, & x_2, & \dots & x_m \\ y_1, & y_2, & \dots & y_n \end{matrix}; f\right]$$

are nonnegative, for all distinct points $x_1, x_2, ..., x_m$ and $y_1, y_2, ..., y_n$. In this paper we are interested in two particular cases, when $m = n = 1$ and $m = n = 2$, the hypothesis of continuity being added. Thus, a continuous function $f : I \times J \to \mathbb{R}$ verifying the condition

$$\left[\begin{matrix} x_1, & x_2 \\ y_1, & y_2 \end{matrix}; f\right] = \frac{1}{(x_1 - x_2)(y_1 - y_2)} \cdot [f(x_1, y_1) - f(x_1, y_2) - f(x_2, y_1) + f(x_2, y_2)] \geq 0 \tag{1.4}$$

for all distinct points $x_1, x_2 \in I$ and $y_1, y_2 \in J$, will be referred to as *box-monotone* (or *box-convex of order* $(1, 1)$). Among the first who noticed the usefulness of this class of functions in mathematical analysis we cite here G. H. Hardy [12], [13] and W. H. Young [24]. A brief account on some recent applications makes the objective of Section 3.

Many interesting monotone function are *box-monotone nondecreasing* (respectively *box-monotone nonincreasing*), meaning that also their coordinate restrictions are nondecreasing (respective nonincreasing). Unlike the case of functions of a single variable, a box-monotone function can be neither box-monotone nondecreasing nor box-monotone nonincreasing. An example is provided by the function

$$f(x, y) = (2x - 1)(2y - 1), \quad (x, y) \in [0, 1] \times [0, 1].$$

The functions for which the inequalities (1.4) work in the opposite direction are the *alternating* ones.

As we shall show in Section 4, the box-monotone functions are related to the following box-analog of the class of convex functions of one real variable.

Definition 1. *(T. Popoviciu [21], p. 78)* A function $f : I \times J \to \mathbb{R}$ is called box-convex (or box-convex of order $(2, 2)$) if it is continuous and all its divided double differences

$$\left[\begin{matrix} x_1, & x_2, & x_3 \\ y_1, & y_2, & y_3 \end{matrix}; f\right]$$

associated to triplets of distinct elements $x_j \in I$ and $y_k \in J$ are nonnegative.

The related notions of *box-concave function* and *box-affine function* can be introduced in the standard way (the condition of continuity being included).

The following remark collects some simple examples and properties of box-convex functions.

Remark 1. *a*) The set $\mathcal{C}_{2d}(I \times J)$ of all box-convex functions defined on a rectangle $I \times J$ constitutes a convex cone in the space $C(I \times J)$ (of all continuous functions defined on $I \times J$). This cone contains all products $f(x)g(y)$ between two continuous functions $f : I \to \mathbb{R}$ and $g : J \to \mathbb{R}$ which are both convex or both concave.

Moreover, the cone $\mathcal{C}_{2d}(I \times J)$ is closed under pointwise limits (provided that the limits are continuous).

b) The restriction of every box-convex function $f \in \mathcal{C}_{2d}(I \times J)$ to a rectangle $K \times L$ included in $I \times J$, is also a box-convex function.

c) $\mathcal{C}_{2d}(\mathbb{R}^2)$ includes also all continuous functions of the form $h(x,y) = c_1 f(x) + c_2 g(y)$ with $c_1, c_2 \geq 0$; for these functions the inequality (1.3) becomes an equality, so that they are actually examples of box-affine functions. The product of an arbitrary continuous function of x and a function of y of the form $my + n$ (with $m, n \in \mathbb{R}$) is also a box-affine function. In particular, the polynomial function xy is box-affine. See Corollary 2 for the general form of box-affine functions of class C^4.

d) If (Ω, Σ, μ) is a probability space and $f : \Omega \times I \times J \to \mathbb{R}$, $f = f(\omega, x, y)$, is a function integrable with respect to ω and box-convex with respect to the pair of variables (x, y), then the integral

$$F(x,y) = \int_{\Omega} f(\omega, x, y)\, d\mu(\omega)$$

defines a box-convex function on $I \times J$.

According to Remark 1 *a*), the function e^{x+y} is box-convex on $\mathbb{R}^2$.

Remark 2. Of a special interest are the *fully* box-convex functions, that is, the functions that combines box-convexity with separate convexity. Some simple examples are the functions

$$x^m \exp(\lambda x + \mu y) \text{ and } y^n \exp(\lambda x + \mu y) \text{ on } \mathbb{R}_+^2 \quad (\text{for all } \lambda, \mu \in \mathbb{R}_+ \text{ and } m, n \in \mathbb{N})$$

and

$$\Pi(x,y) = x^p y^q \text{ on } \mathbb{R}_+^2 \quad (\text{for all } p, q \in [1, \infty)).$$

As noticed in [19], the cosine coefficients (either simple or double) of any fully box convex function defined on $[0, 2\pi] \times [0, 2\pi]$ are nonnegative.

A simple (but tedious) computation shows that the condition

$$\begin{bmatrix} x_1, & x_2, & x_3 \\ y_1, & y_2, & y_3 \end{bmatrix}; f \Bigg] = \begin{bmatrix} x_1, & x_3, & x_2 \\ y_1, & y_3, & y_2 \end{bmatrix}; f \Bigg] \geq 0 \tag{1.5}$$

is equivalent to the following one,

$$\begin{aligned} f(x_1,y_1)&(x_2-x_3)(y_2-y_3) + f(x_1,y_3)(x_2-x_3)(y_1-y_2) \\ &+ f(x_1,y_2)(x_2-x_3)(y_3-y_1) + f(x_3,y_1)(x_1-x_2)(y_2-y_3) \\ &+ f(x_3,y_3)(x_1-x_2)(y_1-y_2) + f(x_3,y_2)(x_1-x_2)(y_3-y_1) \\ &+ f(x_2,y_1)(x_3-x_1)(y_2-y_3) + f(x_2,y_3)(x_3-x_1)(y_1-y_2) \\ &+ f(x_2,y_2)(x_3-x_1)(y_3-y_1) \geq 0 \end{aligned} \tag{1.6}$$

whenever $x_1 < x_3 < x_2$ and $y_1 < y_3 < y_2$.

Noticing that

$$x_3 = (1-\lambda)x_1 + \lambda x_2 \text{ and } y_3 = (1-\mu)y_1 + \mu y_2$$

for suitable $\lambda, \mu \in (0,1)$ one arrives at the following formulation of the notion of box convexity: continuity plus the fulfillment of the inequality

$$\begin{aligned}(1-\lambda)(1-\mu)f(x_1,y_1) + (1-\lambda)\mu f(x_1,y_2) + \lambda(1-\mu)f(x_2,y_1) \\ + \lambda\mu f(x_2,y_2) + f((1-\lambda)x_1 + \lambda x_2, (1-\mu)y_1 + \mu y_2) \\ \geq (1-\lambda)f(x_1,(1-\mu)y_1 + \mu y_2) + (1-\mu)f((1-\lambda)x_1 + \lambda x_2, y_1) \\ + \mu f((1-\lambda)x_1 + \lambda x_2, y_2) + \lambda f(x_2,(1-\mu)y_1 + \mu y_2), \end{aligned} \tag{1.7}$$

whenever $x_1 < x_2$ are points in I, $y_1 < y_2$ are points in J and $\lambda, \mu \in (0,1)$. Under this form, the concept of box-convexity can be extended verbatim to the case of functions taking values in $\mathbb{R} \cup \{\infty\}$.

From a probabilistic point of view, the above inequality describes the barycentric behavior of a box-convex function f on a rectangle $[x_1, x_2] \times [y_1, y_2]$, in the presence of a discrete probability measure

$$m = (1-\lambda)(1-\mu)\delta_{(x_1,y_1)} + \lambda(1-\mu)\delta_{(x_2,y_1)} + \lambda\mu\delta_{(x_2,y_2)} + (1-\lambda)\mu\delta_{(x_1,y_2)}.$$

More precise, this relates the values taken by f at the four corners of the rectangle $[x_1, x_2] \times [y_1, y_2]$,

$$A = (x_1,y_1),\ B = (x_2,y_1),\ C = (x_2,y_2) \text{ and } D = (x_1,y_2),$$

with the value taken at the barycenter of m,

$$\begin{aligned}\operatorname{bar}(m) &= \left(\iint\limits_{[x_1,x_2]\times[y_1,y_2]} x\,dm(x,y), \iint\limits_{[x_1,x_2]\times[y_1,y_2]} y\,dm(x,y) \right) \\ &= \left((1-\lambda)x_1 + \lambda x_2, (1-\mu)y_1 + \mu y_2\right),\end{aligned}$$

and the values at the barycenters of the restriction of m at the sides AB, BC, CD and DA. See [6], Vol. II, p. 143, for the general concept of barycenter. In terms of barycenters, the inequality (1.7) can be rewritten as

$$\begin{aligned}\iint_{ABCD} f\,dm + f(\operatorname{bar}(m)) \geq m(AB)f\left(\operatorname{bar}\frac{m_{|AB}}{m(AB)}\right) + m(BC)f\left(\operatorname{bar}\frac{m_{|BC}}{m(BC)}\right) \\ + m(CD)f\left(\operatorname{bar}\frac{m_{|CD}}{m(CD)}\right) + m(DA)f\left(\operatorname{bar}\frac{m_{|DA}}{m(DA)}\right).\end{aligned}$$

The box-convex functions differ considerably from the usual convex functions. Indeed, if $f(x)$ is an arbitrary function defined on I and $g(y)$ is an arbitrary function defined on J, then the function

$$h(x,y) = f(x) + g(y)$$

verifies the inequality (1.7)! In particular, the fulfillment of this inequality is not able to assure the property of continuity at any point. That's why in order to avoid the presence of pathological functions, in this paper the notion of box-convex function includes continuity as a precondition.

This paper is organized as follows. In Section 2 we present several characterizations of box-convexity including analogs of Jensen's classical results. The differential test of box-convexity (Theorem 3) shows that the class of box-convex functions is strictly larger than the class of completely positive functions, while examples of nondifferentiable such functions are exhibited via Theorem 4 (which offers a characterization of box-convexity by the positivity of Schwarz's upper differential of second order. A brief review on box-monotonicity makes the objective of Section 3. The main result of the next section is Theorem 6, that asserts the duality of the notions of box-convexity and box-monotonicity with respect to the operations of double integration and mixed derivation. As a consequence, many new examples (such as the double logarithmic integral function, the double Clausen function etc.) can be added to the gallery of box-convex functions.

Section 5 focuses on the box-analog of the subdifferential inequality (see Theorem 9). As an immediate consequence we infer the fact that every box-convex function is the pointwise supremum of the family of its box-affine minorants. Also a consequence of this inequality is the integral form of Jensen's inequality for box-convex functions proved in Section 6.

The paper ends with a short list of open problems.

2. Characterizations of box-convexity

The following characterization of the property of box-convexity asserts that under the presence of continuity the basic inequality (1.7) is equivalent to its particular case where $\lambda = \mu = 1/2$.

Theorem 1. *A continuous function $f : I \times J \to \mathbb{R}$ is box-convex if and only if*

$$\frac{1}{2}\left[\frac{f(x_1,y_1)+f(x_1,y_2)+f(x_2,y_1)+f(x_2,y_2)}{4}+f(\frac{x_1+x_2}{2},\frac{y_1+y_2}{2})\right] \geq \frac{f(x_1,\frac{y_1+y_2}{2})+f(\frac{x_1+x_2}{2},y_1)+f(\frac{x_1+x_2}{2},y_2)+f(x_2,\frac{y_1+y_2}{2})}{4}$$

for all $x_1 < x_2$ in I and $y_1 < y_2$ in J.

Corollary 1. *A continuous function $f : I \times J \to \mathbb{R}$ is box-convex if and only if*

$$\frac{1}{2}\left[\frac{f(x-h,y-k)+f(x-h,y+k)+f(x+h,y-k)+f(x+h,y+k)}{4}+f(x,y)\right] \geq \frac{f(x-h,y)+f(x,y-k)+f(x,y+k)+f(x+h,y)}{4}$$

for all points $(x, y) \in I \times J$ and all $h, k > 0$ such that $x \pm h \in I$ and $y \pm k \in J$.

The nontrivial part of Theorem 1 is the sufficiency, which can be easily deduced from the following lemma.

Lemma 1 *(The connection of box-convexity with convexity on line segments). Given a function $f : I \times J \to \mathbb{R}$, one can attach to each subrectangle $[u_1, u_2] \times [v_1, v_2]$ two new functions, the horizontal function, defined by*

$$H_{v_1,v_2}(x) = f(x, v_1) - 2f(x,(v_1 + v_2)/2) + f(x, v_2) \quad \textit{for } x \in [u_1, u_2]$$

and the vertical function, by

$$V_{u_1,u_2}(y) = f(u_1, y) - 2f((u_1 + u_2)/2, y) + f(u_2, y) \quad \textit{for } y \in [v_1, v_2].$$

Then

$$\frac{1}{2}\left[\frac{f(u_1,v_1)+f(u_1,v_2)+f(u_2,v_1)+f(u_2,v_2)}{4}+f(\frac{u_1+u_2}{2},\frac{v_1+v_2}{2})\right]$$
$$-\frac{f(u_1,\frac{v_1+v_2}{2})+f(\frac{u_1+u_2}{2},v_1)+2f(\frac{u_1+u_2}{2},v_2)+f(u_2,\frac{v_1+v_2}{2})}{4}$$
$$=\frac{1}{8}\left[H_{v_1,v_2}(u_1)-2H_{v_1,v_2}((u_1+u_2)/2)+H_{v_1,v_2}(u_2)\right]$$
$$=\frac{1}{8}\left[V_{u_1,u_2}(v_1)-2V_{u_1,u_2}((v_1+v_2)/2)+V_{u_1,u_2}(v_2)\right].$$

As a consequence, under the presence of continuity, f is box-convex if and only if the horizontal and/or the vertical functions are convex.

The proof of Lemma 1 reduces to a straightforward computation and will be omitted.

Remark 3. A property of convexity related to the box-convexity is that of $2D$-convexity (inspired by a beautiful inequality due to T. Popoviciu [22]). According to [5], a real-valued function f defined on a convex subset U of $\mathbb{R}^2$ is called $2D$-convex if

$$\frac{1}{2}\left[\frac{f(x_1,y_1)+f(x_2,y_2)+f(x_3,y_3)}{3}+f\left(\frac{x_1+x_2+x_3}{3},\frac{y_1+y_2+y_3}{3}\right)\right]$$
$$\geq\frac{1}{3}\left[f\left(\frac{x_1+x_2}{2},\frac{y_1+y_2}{2}\right)+f\left(\frac{x_2+x_3}{2},\frac{y_2+y_3}{2}\right)+f\left(\frac{x_3+x_1}{2},\frac{y_3+y_1}{2}\right)\right]$$

for all triplets $(x_1,y_1),(x_2,y_2),f(x_3,y_3)$ of points in U. Every $2D$-convex continuous function is convex in the usual sense. See [5], Theorem 2. However, one can indicate simple examples (like $\max\{|x|,|y|\}$) of $2D$-convex functions that are not box-convex.

Lemma 1 offers an easy way to transfer the inequalities of convexity from one variable to two variables. Indeed, assuming for a moment that the box-convex function $f: I\times J\to\mathbb{R}$ is of class C^2, we infer that $\frac{\partial^2 f}{\partial x^2}(x,y)$ is convex in the variable y, which yields, according to Jensen's classical inequality, that

$$\frac{1}{N}\sum_{k=1}^{N}\frac{\partial^2 f}{\partial x^2}(x,v_k)-\frac{\partial^2 f}{\partial x^2}(x,\frac{1}{N}\sum_{k=1}^{N}v_k)\geq 0$$

whenever $x\in I$ and $v_1,...,v_N\in J$. In turn, this implies the convexity of all the functions

$$\frac{1}{N}\sum_{k=1}^{N}f(x,v_k)-f(x,\frac{1}{N}\sum_{k=1}^{N}v_k)$$

and a new appeal to Jensen's inequality allows us to conclude that

$$\frac{1}{MN}\sum_{j=1}^{M}\sum_{k=1}^{N}f(u_j,v_k)+f(\frac{1}{M}\sum_{j=1}^{M}u_j,\frac{1}{N}\sum_{k=1}^{N}v_k)\geq\frac{1}{M}\sum_{j=1}^{M}f(u_j,\frac{1}{N}\sum_{k=1}^{N}v_k)+\frac{1}{N}\sum_{k=1}^{N}f(\frac{1}{M}\sum_{j=1}^{M}u_j,v_k),\quad (2.1)$$

whenever $u_1,...,u_M\in I$ and $v_1,...,v_N\in J$. The inequality (2.1) represents the box-analog of Jensen's inequality for functions of two variables. The differentiability requirement on f can be replaced by continuity due to the following result:

Theorem 2. (S. Gal [10], Theorem 2.2.2, p. 116) *If a continuous function $f : [0,1] \times [0,1] \to \mathbb{R}$ is box-convex of order (r,s), then so are the bivariate Bernstein polynomials*

$$B_{n,m}(f)(x,y) = \sum_{i=0}^{n} \sum_{j=0}^{m} \binom{n}{i} \binom{m}{j} x^i (1-x)^{n-i} y^j (1-y)^{n-j} f\left(\frac{i}{n}, \frac{j}{m}\right)$$

associated to it. Moreover, by the well-known property of simultaneously uniform approximation of a univariate function and its derivatives by the univariate Bernstein polynomials and its derivatives, it follows that $B_{n,m}(f)$ and any partial derivative (of any order) of it, converge uniformly to f (and to its partial derivatives), correspondingly.

The following consequence of Lemma 1 offers a practical way to detect the box-convexity.

Theorem 3 *(The differential test of box-convexity). Suppose that $f : I \times J \to \mathbb{R}$ is a function continuous on the compact rectangle $I \times J$ and of class C^2 on the interior of $I \times J$. Then f is box-convex if and only if $\frac{\partial^2 f}{\partial x^2}$ is convex in the second variable (if and only if $\frac{\partial^2 f}{\partial y^2}$ is convex in the first variable).*

If f of class C^4 on the interior of $I \times J$, then f is box-convex if and only if $\frac{\partial^4 f}{\partial x^2 \partial y^2} \geq 0$.

The last part of Theorem 3 was known to T. Popoviciu [21]. The necessity of the condition $\frac{\partial^4 f}{\partial x^2 \partial y^2} \geq 0$ for the box-convexity of f was derived by him via the formula

$$\begin{bmatrix} x_1, & x_2, & x_3 \\ y_1, & y_2, & y_3 \end{bmatrix}; f = \frac{1}{4} \cdot \frac{\partial^4 f}{\partial x^2 \partial y^2}(\xi, \eta)$$

for a suitable $(\xi, \eta) \in]x_1, x_3[\times]y_1, y_3[$; here $x_1 < x_2 < x_3$ and $y_1 < y_2 < y_3$. The sufficiency part admits a similar argument. Indeed, at every interior point (x,y),

$$\lim_{h \to 0} \begin{bmatrix} x-h, & x, & x+h \\ y-h, & y, & y+h \end{bmatrix}; f = \frac{1}{4} \cdot \frac{\partial^4 f}{\partial x^2 \partial y^2}(x,y).$$

The differential test of box-convexity yields the general form of box-affine functions of class C^4:

Corollary 2. *A function $f \in C^4(I \times J)$ is box-affine if and only if $\frac{\partial^4 f}{\partial x^2 \partial y^2} = 0$. The general form of the box-affine functions is*

$$f(x,y) = u(x)\,(py+q) + v(y)\,(rx+s) + t$$

where $u \in C^2(I)$, $v \in C^2(J)$ and $p, q, r, s, t \in \mathbb{R}$.

The differential test of box-convexity is the source of numerous important examples. For example, if φ is a function of class C^4 defined on $(0,\infty)$, and $\varphi^{(4)} \geq 0$, then the function

$$f(x,y) = \varphi(\lambda x + \mu y + \gamma)$$

is box-convex on $(0,\infty) \times (0,\infty)$, whenever $\lambda, \mu, \gamma > 0$. In particular, this is the case of the functions

$$\Gamma(\lambda x + \mu y + \gamma) \text{ and } (\lambda x + \mu y + \gamma)\log(\lambda x + \mu y + \gamma).$$

With appropriate changes, this remark also works for functions defined on other intervals. For example, the function $\sin(x+y)$ is box-convex on $[0,\pi/2]\times[0,\pi/2]$ and the function

$$\frac{\log(x+y)}{x+y}$$

is box-convex on $[\frac{1}{2}e^{25/12},\infty)\times[\frac{1}{2}e^{25/12},\infty)$ (though not on $(0,\infty)\times(0,\infty)$).

Other interesting examples are provided by the class of completely monotone functions of two variables. Recall that an infinitely differentiable function $f:(0,\infty)\times(0,\infty)\to\mathbb{R}$ is completely monotone if it verifies the inequalities

$$(-1)^{i+j}\frac{\partial^{i+j}f}{\partial x^i\partial y^j}\geq 0 \quad \text{for all } i,j=0,1,2,\ldots .$$

Every such function is nonnegative, box-increasing and box-convex (as well as decreasing and convex in each variable). So is the case of the function

$$\frac{1}{1+x+y+\alpha xy},$$

where $\alpha\in[0,1]$ is a parameter. Notice that if $\varphi:(0,\infty)\to\mathbb{R}$ is a completely monotone functions of one variable (that is, $(-1)^k\varphi^{(k)}\geq 0$ for $k=0,1,2,\ldots$), then $f(x,y)=\varphi(x+y)$ is nonnegative, box-monotone and box-convex. The class of completely monotone functions of one variable includes the functions

$$e^{-x},\ e^{-1/x},\ \frac{1}{(1+x)^2},\ \frac{\log(x+1)}{x} \text{ and } \frac{e^x\Gamma(x+1)}{x^{x+1/2}},$$

and is closed under linear combinations with positive coefficients, products, derivation of even order etc. See Miller and Samko [15] for full details.

Also a consequence of the differential test of box-convexity is the case of perspective functions. The *perspective function* associated to a function $\varphi:(0,\infty)\to\mathbb{R}$ is defined by the formula

$$f(x,y)=y\varphi(x/y), \quad \text{for } x,y>0.$$

The function f is box-convex provided that φ is of class C^4 and $\left(x^2\varphi''(x)\right)''\geq 0$.

In particular, the function x^2/y (the perspective companion of x^2) is box-convex on $\mathbb{R}\times(0,\infty)$ while the Kullback–Leibler divergence function $x\log x-x\log y$ (the perspective companion of $x\log x$) is box-convex on $(0,\infty)\times(0,\infty)$.

The box-convex functions can be suitably characterized outside differentiability, by using the Schwarz derivative of second order for functions of two variables:

$$S\mathcal{D}^2f(x,y)=\lim_{\substack{h,k\to 0\\ hk\neq 0}}\frac{\begin{array}{c}f(x-h,y+k)-2f(x,y+k)+f(x+h,y+k)\\ -2f(x-h,y)+4f(x,y)-2f(x+h,y)\\ +f(x-h,y-k)-2f(x,y-k)+f(x+h,y-k)\end{array}}{h^2k^2}.$$

By replacing lim by lim sup and lim inf one obtains respectively the notions of upper and lower Schwarz derivative of second order (denoted $\overline{S\mathcal{D}}^2f(x,y)$ and $\underline{S\mathcal{D}}^2f(x,y)$).

It is not difficult to prove that if f is of class C^4, then

$$S\mathcal{D}^2 f(x,y) = \frac{\partial^4 f}{\partial x^2 \partial y^2}(x,y).$$

Theorem 4. *A continuous function $f : I \times J \to \mathbb{R}$ is box-convex if and only if $\overline{S\mathcal{D}}^2 f(x,y) \geq 0$ at all points (x,y) interior to $I \times J$.*

Before to detail the proof of Theorem 4 we will comment some of its consequences.

Corollary 3 *(The local character of box-convexity). Suppose that f is a continuous function defined on an open rectangle R such that every point x in R is interior to an open subrectangle S on which f is box-convex. Then f is box-convex on the entire rectangle R.*

As was noticed by C. Freiling and D. Rinne (see [2]) the nondifferentiable continuous function $f(x,y) = (x+y)\,|x+y|$ verifies the condition $S\mathcal{D}^2 f(x,y) = 0$ at all points of $\mathbb{R}^2$. According to Theorem 4, this function is box-convex. Other functions, like $|x+y|$ and $\max\{x,y\}$, are neither box convex nor box-concave.

The proof of Theorem 4 needs some preparation.

Lemma 2. *If $f : [c,d] \to \mathbb{R}$ is a function such that $f(c) - 2f((c+d)/2) + f(d) < 0$, then there exist points $\alpha < \beta$ in $[c,d]$ such that*

$$f(\alpha) - 2f((\alpha+\beta)/2) + f(\beta) < 0 \text{ and } \beta - \alpha = \frac{1}{2}(d-c).$$

Proof. By considering the auxiliary function $g : [c,(d+c)/2] \to \mathbb{R}$ defined by

$$g(x) = f(x) - 2f(x + (d-c)/4) + f(x + (d-c)/2),$$

it suffices to prove the existence of a point $\alpha \in [c,(d+c)/2]$ such that $g(\alpha) = f(\alpha) - 2f(\alpha + (d-c)/4) + f(\alpha + (d-c)/2) < 0$. Indeed, in this case the statement of Lemma 2 works for $\beta = \alpha + (d-c)/2)$.

If no such α would exist, it follows that

$$g(x) = f(x) - 2f(x + (d-c)/4) + f(x + (d-c)/2) \geq 0 \quad \text{for all } x \in [c,(d+c)/2].$$

For $x = c$ and $x = (d+c)/2$ this gives

$$f(c) - 2f((d+c)/2 - (d-c)/4) + f((d+c)/2) \geq 0,$$

and

$$f((d+c)/2) - f((d+c)/2 + (d-c)/4)) + f(d) \geq 0$$

whence, by adding term by term these two inequalities, we obtain

$$f(c) - 2f((d+c)/2) + f(d) - 2\left[f((d+c)/2 + (d-c)/4) - 2f((d+c)/2) + f((d+c)/2 - (d-c)/4)\right] \geq 0.$$

By our hypothesis, $f(c) - 2f((c+d)/2) + f(d) < 0$, which yields

$$f((d+c)/2 + (d-c)/4) - 2f((d+c)/2) + f((d+c)/2 - (d-c)/4) < 0$$

in contradiction with the fact that

$$g((d+c)/2-(d-c)/4) \geq 0.$$

The proof is done. □

Proof of Theorem 4. The necessity part follows from the fact that

$$\overline{S\mathcal{D}}^2 f(x,y) \geq \underline{S\mathcal{D}}^2 f(x,y) \geq 0$$

for every box-convex function.

For the sufficiency, suppose first that $\overline{S\mathcal{D}}^2 f > 0$ but f were not box-convex. This yields points $a_1^{(1)} < a_2^{(1)}$ in I and $b_1^{(1)} < b_2^{(1)}$ in J such that

$$\frac{1}{2}\left[\frac{f(a_1^{(1)},b_1^{(1)})+f(a_1^{(1)},b_2^{(1)})+f(a_2^{(1)},b_1^{(1)})+f(a_2^{(1)},b_2^{(1)})}{4}\right.$$
$$\left.+f(\frac{a_1^{(1)}+a_2^{(1)}}{2},\frac{b_1^{(1)}+b_2^{(1)}}{2})\right]$$
$$<\frac{f(a_1^{(1)},\frac{b_1^{(1)}+b_2^{(1)}}{2})+f(\frac{a_1^{(1)}+a_2^{(1)}}{2},b_1^{(1)})+f(\frac{a_1^{(1)}+a_2^{(1)}}{2},b_2^{(1)})+f(a_2^{(1)},\frac{b_1^{(1)}+b_2^{(1)}}{2})}{4}.$$

According to Lemma 2 (when applied to $H_{b_1^{(1)},b_2^{(1)}}(x)$) there exist points $a_1^{(2)} < a_2^{(2)}$ in $\left[a_1^{(1)},a_2^{(1)}\right]$ such that $a_2^{(2)}-a_1^{(2)}=\frac{a_2^{(1)}-a_1^{(1)}}{2}$ and

$$H_{b_1^{(1)},b_2^{(1)}}(a_1^{(2)})-2H((a_1^{(2)}+a_2^{(2)})/2)+H(a_2^{(2)})<0.$$

A similar argument yields the existence of points $b_1^{(2)} < b_2^{(2)}$ in $\left[b_1^{(1)},b_2^{(1)}\right]$ such that $b_2^{(2)}-b_1^{(2)}=\frac{b_2^{(1)}-b_1^{(1)}}{2}$ and

$$V_{b_1^{(2)},b_2^{(2)}}(a_1^{(2)})-2V((a_1^{(2)}+a_2^{(2)})/2)+V(a_2^{(2)})<0.$$

Continuing this way, we obtain the existence of a sequence of nested rectangles $[a_1^{(n)},a_2^{(n)}]\times[b_1^{(n)},b_2^{(n)}]$ (whose diameters converge to 0) such that

$$\frac{1}{2}\left[\frac{f(a_1^{(n)},b_1^{(n)})+f(a_1^{(n)},b_2^{(n)})+f(a_2^{(n)},b_1^{(n)})+f(a_2^{(n)},b_2^{(n)})}{4}\right.$$
$$\left.+f(\frac{a_1^{(n)}+a_2^{(n)}}{2},\frac{b_1^{(n)}+b_2^{(n)}}{2})\right]$$
$$<\frac{f(a_1^{(n)},\frac{b_1^{(n)}+b_2^{(n)}}{2})+f(\frac{a_1^{(n)}+a_2^{(n)}}{2},b_1^{(n)})+f(\frac{a_1^{(n)}+a_2^{(n)}}{2},b_2^{(n)})+f(a_2^{(n)},\frac{b_1^{(n)}+b_2^{(n)}}{2})}{4},$$

for all indices $n \geq 1$. See Lemma 1. According to a well known results due Cantor, there is a point (a,b) at which both sequences $(a_1^{(n)},b_1^{(n)})_n$ and $(a_2^{(n)},b_2^{(n)})_n$ converge. Taking into account that f is a continuous function, we infer from the last inequality that $\overline{S\mathcal{D}}^2 f(a,b) \leq 0$, a contradiction.

If $\overline{S\mathcal{D}}^2 f \geq 0$, we apply the above reasoning to the sequence of functions $f_n(x,y) = f(x,y) + x^2y^2/4n$. Since $\overline{S\mathcal{D}}^2 f_n \geq 1/n$ for every n, the functions f_n are box-convex and thus the function f verifies the inequality stated in Theorem 1 (being a pointwise limit of such functions). According to this theorem, f is box convex. □

3. Generalities on the box-monotone functions

The box-monotonicity of a real-valued function f defined on the rectangle $R = I \times J$ can be discusses in terms of box-increments.

The *box-increment* of f over a compact subrectangle $S = [a_1, a_2] \times [b_1, b_2]$ of R is defined by the formula

$$\Delta(f;S) = f(a_1,b_1) - f(a_1,b_2) - f(a_2,b_1) + f(a_2,b_2).$$

Accordingly, the function f is box-monotone if and only if

$$\Delta(f;S) \geq 0 \tag{3.1}$$

for all compact subrectangles S of R, equivalently,

$$\Delta(f;S) \leq \Delta(f;\tilde{S})$$

whenever $S \subset \tilde{S}$ are two compact rectangles.

When f is a product function of the form $f(x,y) = g(x)h(y)$, then the property of box-monotonicity means the *comonotonicity* of g and h, that is,

$$(g(x_2) - g(x_1))(h(y_2) - h(y_1)) \geq 0$$

whenever $x_1 \leq x_2$ in I and $y_1 \leq y_2$ in J. As a consequence, the concept of box-monotonicity differs from the usual one, associated to the coordinatewise ordering of $\mathbb{R}^2$.

The box-monotonicity can be characterized in a convenient manner via the Schwarz derivative of first order for functions of two variables:

$$\begin{aligned} S\mathcal{D}f(x,y) &= \lim_{h,k\to 0} \frac{\Delta(f;[x,x+h]\times[y,y+k])}{hk} \\ &= \lim_{h,k\to 0} \frac{f(x+h,y+k) - f(x+h,y) - f(x,y+k) + f(x,y)}{hk}; \end{aligned}$$

by replacing lim by lim sup and lim inf one obtains respectively the notions of upper and lower Schwarz derivative of first order (denoted $\overline{S\mathcal{D}}f(x,y)$ and $\underline{S\mathcal{D}}f(x,y)$).

Remark 4. The Schwarz derivative of first order $S\mathcal{D}f$ is also known as the hyperbolic derivative because in the case of functions f of class C^1 which admit a continuous mixed derivative $\frac{\partial^2 f}{\partial y \partial x}$, the other mixed derivative $\frac{\partial^2 f}{\partial x \partial y}$ also exists and

$$S\mathcal{D}f(x,y) = \frac{\partial^2 f}{\partial y \partial x} = \frac{\partial^2 f}{\partial x \partial y}$$

See J. M. Ash, J. Cohen, C. Freiling, D. Rinne [3], Proposition 2, and A. G. Aksoy and M. Martelli [1], Theorem 3.

Theorem 5. *A continuous function $f : I \times J \to \mathbb{R}$ is box-monotone if and only if $\overline{S\mathcal{D}}f(x, y) \geq 0$ at all points.*

The proof is similar to that of Theorem 4, so we omit the details.

Corollary 4 *(The differential criterion of box-monotonicity). If $f \in C^1(I \times J)$ and the mixed derivative $\frac{\partial^2 f}{\partial y \partial x}$ exists and is continuous, then:*

(i) f is box-monotone if and only if

$$\frac{\partial^2 f}{\partial x \partial y} \geq 0; \tag{3.2}$$

(ii) f is box-monotone nondecreasing if and only if

$$\frac{\partial f}{\partial x} \geq 0, \ \frac{\partial f}{\partial y} \geq 0 \text{ and } \frac{\partial^2 f}{\partial x \partial y} \geq 0.$$

An immediate consequence of Corollary 4 is the box-monotonicity of a series of functions like xy and $-\log(e^x + e^y)$ on $\mathbb{R}^2$ and $-\log(x + y)$ on $(0, \infty) \times (0, \infty)$. Two more exotic examples are the incomplete gamma function

$$\gamma(a, x) = \int_0^x e^{-t} t^{a-1} dt \quad \text{for } a > 0 \text{ and } x \geq 1,$$

and the elliptic integral of the first kind,

$$F(\varphi, k) = \int_0^\varphi \left(1 - k^2 \sin^2 \theta\right)^{-1/2} d\theta \quad \text{for } \varphi \in [0, \pi/2], \ k \in [0, 1).$$

As was noticed in [17], if $\varphi : \mathbb{R} \to \mathbb{R}$ and $\psi : [0, \infty) \to [0, \infty)$ are continuous convex functions, then the formulas

$$f(x, y) = -\varphi(x - y) \quad \text{and} \quad g(x, y) = \psi(x + y)$$

define box-monotone functions, respectively on $\mathbb{R} \times \mathbb{R}$ and $[0, \infty) \times [0, \infty)$.

Remark 5 *(From local to global).* Corollary 4 easily yields the following gluing principle for box-monotone functions: If f is a real-valued function of class C^2 defined on a compact rectangle R and the restriction of f to each compact subrectangle of a division of R is box-monotone, then f is box-monotone on the entire rectangle R.

Probability and statistics constitute a major source of box-monotone functions: the functions (known as copulas, with a prominent role in risk management) that couple multivariate distribution functions to their one-dimensional marginal distribution functions. Some relevant papers in this respect are P. Embrechts [7] and C. Genest and J. Nešlehová [11]. Also, a thorough introduction to the theory of copulas is offered by the book of R. B. Nelsen [16].

The most popular examples of copulas are the *Archimedean copulas*. They are constructed through a continuous, strictly decreasing and convex generator φ via the formula

$$A_\varphi(x, y) = \varphi^{-1}\left(\varphi(x) + \varphi(y)\right),$$

where φ^{-1} is the inverse of φ. Several concrete examples of one-parameter families of Archimedean copulas are presented in the book of R. B. Nelsen [16]. Notice that all Archimedean copulas are box-monotone *increasing* functions.

It is also worthwhile to mention that two important mathematical concepts, namely those of doubly stochastic measure and of Markov operator, are intimately related to copulas (see, e.g., E. T. Olsen, W. F. Darsow and B. Nguyen [20]).

Remark 6. Comparative statistics and combinatorial optimization provide yet another valuable source of box-monotone/box-alternating functions defined on Cartesian products $I \times I$, where $I = \mathbb{R}$ or $I = \mathbb{R}_+$. A function $L : I \times I \to \mathbb{R}$ is called *supermodular* if for each pair of vectors (x, y) and (x', y') in $I \times I$ we have the inequality

$$L(x,y) + L(x',y') \le L(\min\{x,x'\}, \min\{y,y'\}) + L(\max\{x,x'\}, \max\{y,y'\}).$$

The function L is called submodular if $-L$ is supermodular. Since supermodularity can be characterized by the condition

$$L(x,y+v) + L(x+u,y) \le L(x,y) + L(x+u,y+v)$$

for all $x, y \in I$ and $u, v \in [0,\infty)$, the supermodular/submodular functions are nothing but box-monotone/box-alternating functions defined on $I \times I$.

The function $L(x,y) = |x-y|^p$ is submodular on $\mathbb{R} \times \mathbb{R}$ for $p \in [1,\infty)$; the case $p = 1$ follows from the definition of submodularity, while for $p > 1$ we have to notice that $\frac{\partial^2 L}{\partial y \partial x} \le 0$.

Applications of submodular functions are available from many sources. For example, see F. Bach [4] and B. M. Topkis [23]. There is also a large literature devoted to the submodular set functions with important applications to economics, game theory, machine learning and computer vision. See the book of S. Fujishige [9] for a gentle introduction and the connection with submodular functions.

4. The connection between box-convexity and box-monotonicity

Many results concerning the functions of one real variable can be extended to the context of functions defined on rectangles, by replacing the usual derivative with the mixed derivative.

For example, the derivative of a differentiable convex function of one real variable is a nondecreasing function and the primitive of a nondecreasing function is a convex function. The analog of this relationship for functions defined on rectangles is as follows:

Theorem 6. (i) *If $f \in C(I \times J)$ is a box-monotone function and $(a,b) \in I \times J$, then the function*

$$F(x,y) = \int_a^x \int_b^y f(u,v)\, du dv$$

is box-convex.

(ii) *Conversely, if $f \in C^2(I \times J)$ is a box-convex function, then its mixed derivative $\frac{\partial^2 f}{\partial x \partial y}$ is box-monotone.*

Proof. (i) If f is of class C^2, then $\frac{\partial^2 f}{\partial x \partial y}(x,y) \ge 0$ according to Corollary 4. Therefore

$$\frac{\partial^4 F}{\partial x^2 \partial y^2}(x,y) = \frac{\partial^2}{\partial x \partial y}\left(\frac{\partial^2 F}{\partial x \partial y}(x,y)\right) = \frac{\partial^2 f}{\partial x \partial y}(x,y) \ge 0$$

and from Theorem 3 we infer that F is box-convex. The case where f is not differentiable reduces to the preceding one via Theorem 2.

(ii) If f is of class C^4, then $\frac{\partial^2}{\partial x \partial y}\left(\frac{\partial^2 f}{\partial x \partial y}(x,y)\right) = \frac{\partial^4 f}{\partial x^2 \partial y^2}(x,y) \geq 0$, which shows that the mixed derivative $\frac{\partial^2 f}{\partial x \partial y}$ is box-monotone.

The case where f is only of class C^2 is now a consequence of Theorem 2. □

Corollary 5 *(From local to global). If f is a real-valued function of class C^2 defined on a compact rectangle R and the restriction of f to each compact subrectangle of a division of R, is box-convex, then f is box-convex on the entire rectangle R.*

Two examples illustrating Theorem 6 are provided by the *double logarithmic integral function,*

$$\mathrm{Li}(x,y) = \int_0^x \int_0^y \frac{1}{\log(u+v)}\, dudv, \quad x, y > 0,$$

and the *double Clausen function,*

$$\mathrm{Cl}(x,y) = -\int_0^x \int_0^y \log\left(2\sin\frac{u+v}{2}\right) dudv, \quad x, y \in [0,\pi].$$

Notice that according to Theorem 3, the function $\frac{1}{\log(x+y)}$ is box-convex on $(1/4,\infty) \times (1/4,\infty)$ (but not on $(0,\infty) \times (0,\infty)$).

The equivalence between convexity and increasing monotonicity of the slope function (as known for functions defined on intervals) also admits a box-analog:

Theorem 7. *A continuous function $f : [a_1, a_2] \times [b_1, b_2] \to \mathbb{R}$ is box-convex if and only if for every pair of points $a \in [a_1, a_2]$ and $b \in [b_1, b_2]$ the associated function*

$$F(x,y) = \begin{bmatrix} a, & x \\ b, & y \end{bmatrix}; f$$

is box-monotone on each subrectangle $S = [u_1, u_2] \times [v_1, v_2]$ of $[a_1, a_2] \times [b_1, b_2]$.

Proof. Indeed,

$$\begin{aligned}
\left[\begin{matrix} u_1, & a, & u_2 \\ v_1, & b, & v_2 \end{matrix}; f\right] &= [u_1, a, u_2; [v_1, b, v_2; f((x,\cdot)]] \\
&= \left[u_1, a, u_2; \frac{[b, v_2; f(x,\cdot)] - [v_1, b; f(x,\cdot)]}{v_2 - v_1}\right] \\
&= \frac{1}{u_2 - u_1}\left(\left[a, u_2; \frac{[b, v_2; f] - [v_1, b; f]}{v_2 - v_1}\right] - \left[u_1, a; \frac{[b, v_2; f] - [v_1, b; f]}{v_2 - v_1}\right]\right) \\
&= \frac{1}{(u_2 - u_1)(v_2 - v_1)}\left(\left[\begin{matrix} a, & u_2 \\ b, & v_2 \end{matrix}; f\right] - \left[\begin{matrix} a, & u_2 \\ v_1, & b \end{matrix}; f\right]\right. \\
&\quad \left. - \left[\begin{matrix} u_1, & a \\ b, & v_2 \end{matrix}; f\right] + \left[\begin{matrix} u_1, & a \\ v_1, & b \end{matrix}; f\right]\right)
\end{aligned}$$

$$= \frac{1}{(u_2-u_1)(v_2-v_1)}\left(F(u_2,v_2)-F(u_2,v_1)-F(u_1,v_2)+F(u_1,v_1)\right)$$
$$= \frac{1}{(u_2-u_1)(v_2-v_1)}\cdot\Delta(F;S),$$

and the assertion of the theorem is now obvious. □

5. The analog of the subgradient inequality

We start this section with the following box-analog of the Leibniz-Newton formula:

Theorem 8. *Assume that $f: I\times J\to\mathbb{R}$ is a continuously differentiable function which admits a continuous second order partial derivative $\frac{\partial^2 f}{\partial x\partial y}$. Then for every pair of points (a,b) and (x,y) in $I\times J$ we have*

$$f(x,y) = -f(a,b)+f(x,b)+f(a,y)+\int_a^x\int_b^y \frac{\partial^2 f}{\partial u\partial v}(u,v)\,dvdu.$$

Proof. Indeed, taking into account Remark 4, we have

$$\begin{aligned}\int_a^x\int_b^y \frac{\partial^2 f}{\partial u\partial v}(u,v)\,dvdu &= \int_a^x\int_b^y \frac{\partial}{\partial v}\left(\frac{\partial f}{\partial u}\right)(u,v)\,dvdu\\ &= \int_a^x\left[\frac{\partial f}{\partial u}(u,y)-\frac{\partial f}{\partial u}(u,b)\right]du\\ &= f(x,y)-f(a,y)-f(x,b)+f(a,b). \quad \square\end{aligned}$$

For functions of class C^4, the box-analog of the Leibniz-Newton formula can be iterated by replacing $f(x,y)$ by $\frac{\partial^2 f}{\partial u\partial v}(x,y)$. As a consequence,

$$\begin{aligned}f(x,y) &= -f(a,b)+f(x,b)+f(a,y)+\int_a^x\int_b^y \frac{\partial^2 f}{\partial u\partial v}(u,v)\,dudv\\ &= -f(a,b)+f(x,b)+f(a,y)\\ &+\int_a^x\int_b^y\left[-\frac{\partial^2 f}{\partial u\partial v}(a,b)+\frac{\partial^2 f}{\partial u\partial v}(u,b)+\frac{\partial^2 f}{\partial u\partial v}(a,v)+\int_a^u\int_b^v \frac{\partial^4 f}{\partial s^2\partial t^2}(s,t)\,dsdt\right]dudv\\ &= -f(a,b)+f(x,b)+f(a,y)+(x-a)\left[\frac{\partial f}{\partial u}(a,y)-\frac{\partial f}{\partial u}(a,b)\right]\\ &+(y-b)\left[\frac{\partial f}{\partial v}(x,b)-\frac{\partial f}{\partial v}(a,b)\right]-\frac{\partial^2 f}{\partial u\partial v}(a,b)(x-a)(y-b)\\ &+\int_a^x\int_b^y\left[\int_a^u\int_b^v \frac{\partial^4 f}{\partial s^2\partial t^2}(s,t)\,dsdt\right]dudv.\end{aligned}$$

This yields the following box-analog of the subdifferential inequality:

Theorem 9. *A function $f \in C^2(I \times J)$ is box-convex if and only if*

$$f(x,y) \geq -f(a,b) + f(x,b) + f(a,y) + (x-a)\frac{\partial f}{\partial u}(a,y) + (y-b)\frac{\partial f}{\partial v}(x,b) - (x-a)\frac{\partial f}{\partial u}(a,b) - (y-b)\frac{\partial f}{\partial v}(a,b) - \frac{\partial^2 f}{\partial u \partial v}(a,b)(x-a)(y-b)$$

for all points (x,y) and (a,b) in $I \times J$.

Proof. The case of functions of class C^4 is clear. The case of functions of class C^2 can be reduced to the precedent one via an approximation argument. Indeed, according to Theorem 2, every box-convex function can be uniformly approximated by the bivariate tensor product Bernstein polynomials, $B_{n,m}$, which are box-convex polynomials, and any partial derivative (of any order) of f can be uniformly approximated by the corresponding partial derivatives of $B_{n,m}$. $\square$

Corollary 6. *If $f \in C^2(I \times J)$ is box-convex, then f is the pointwise supremum of the family of its box-affine minorants.*

For $f(x,y) = e^{x+y}$ $(x, y \in \mathbb{R})$ and $a = b = 0$, the inequality stated by Theorem 9 is equivalent to

$$e^x \geq 1 + x \quad \text{for all } x \in \mathbb{R},$$

while in the case $f(x,y) = \log(1 + x + y)$ $(x, y \geq 0)$ and $a = b = 0$, one obtains a consequence of this later inequality. In the same time, Theorem 9 is the source of many intriguing inequalities, a fact illustrated even by some simple functions such as $f(x,y) = \frac{1}{1+x+y+\alpha xy}$ $(x, y \geq 0)$, with $\alpha \in [0,1]$.

Remark 7. It is worth mentioning that Theorem 9 can also be deduced from Theorem 7. Indeed, according to this theorem the function

$$F(x,y) = \begin{bmatrix} a, & x \\ b, & y \end{bmatrix}; f \Bigg],$$

verifies the inequality

$$\left[\begin{matrix} a, & x \\ b, & y \end{matrix}; f\right] - \left[\begin{matrix} a, & x \\ v, & b \end{matrix}; f\right] - \left[\begin{matrix} u, & a \\ b, & y \end{matrix}; f\right] + \left[\begin{matrix} u, & a \\ v, & b \end{matrix}; f\right] \geq 0 \tag{5.1}$$

whenever u is an intermediate point between a and x and v is an intermediate point between b and y. Passing to the limit as $u \to a$ and $v \to b$ one obtains

$$\begin{aligned} &\left[\begin{matrix} a, & x \\ b, & y \end{matrix}; f\right] - \left[a, \quad x\,; \frac{\partial f}{\partial v}(\cdot, b)\right] - \left[b, \quad y\,; \frac{\partial f}{\partial u}(a, \cdot)\right] + \frac{\partial^2 f}{\partial u \partial v}(a,b) \\ &\qquad = \frac{f(x,y) - f(a,y) - f(x,b) + f(a,b)}{(y-b)(x-a)} - \frac{\frac{\partial f}{\partial v}(x,b) - \frac{\partial f}{\partial v}(a,b)}{x-a} \\ &\qquad\qquad - \frac{\frac{\partial f}{\partial u}(a,y) - \frac{\partial f}{\partial u}(a,b)}{y-b} + \frac{\partial^2 f}{\partial u \partial v}(a,b) = \frac{1}{(y-b)(x-a)} \cdot SD \geq 0, \end{aligned}$$

where

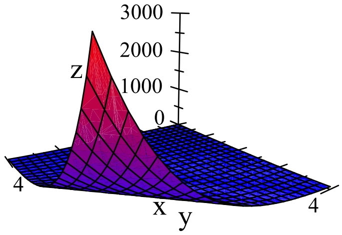

Fig. 1. The graph of e^{x+y}.

$$SD = f(x,y) - f(a,y) - f(x,b) + f(a,b) + (y-b)\left(\frac{\partial f}{\partial v}(x,b) - \frac{\partial f}{\partial v}(a,b)\right) + (x-a)\left(\frac{\partial f}{\partial u}(a,y) - \frac{\partial f}{\partial u}(a,b)\right) + (x-a)(y-b)\frac{\partial^2 f}{\partial u \partial v}(a,b).$$

This implies $SD \geq 0$ when $(x-a)(y-b) > 0$. The case $(x-a)(y-b) < 0$ can be settled in a similar manner.

Definition 2. Under the assumptions of Theorem 9, we call the function

$$Af(a,b)(x,y) = -f(a,b) + f(x,b) + f(a,y) + (x-a)\frac{\partial f}{\partial u}(a,y) + (y-b)\frac{\partial f}{\partial v}(x,b) - (x-a)\frac{\partial f}{\partial u}(a,b) - (y-b)\frac{\partial f}{\partial v}(a,b) - \frac{\partial^2 f}{\partial u \partial v}(a,b)(x-a)(y-b)$$

the box-affine part at (a,b) of the function f.

A box-convex function and its box-affine part are very tight at the point of contact. For example, if $f \in C^2(I \times J)$, then f and $Af(a,b)$ coincide along the cross $I \times \{b\} \cup \{a\} \times J$ together with all their partial derivatives of second order, except the mixed derivative $\frac{\partial^2}{\partial x \partial y}$ where the coincidence occurs only at the point (a,b). Indeed,

$$\begin{aligned} Af(a,b)(a,y) &= f(a,y), \quad & Af(a,b)(x,b) &= f(x,b), \\ \frac{\partial Af(a,b)}{\partial x}(x,b) &= \frac{\partial f}{\partial x}(x,b), \quad & \frac{\partial Af(a,b)}{\partial y}(a,y) &= \frac{\partial f}{\partial y}(a,y), \\ \frac{\partial^2 Af(a,b)}{\partial x^2}(x,b) &= \frac{\partial^2 f}{\partial x^2}(x,b), \quad & \frac{\partial^2 Af(a,b)}{\partial y^2}(a,y) &= \frac{\partial^2 f}{\partial y^2}(a,y), \end{aligned}$$

and

$$\frac{\partial^2 Af(a,b)}{\partial x \partial y}(x,y) = \frac{\partial^2 f}{\partial x \partial y}(a,y) + \frac{\partial^2 f}{\partial y \partial x}(x,b) - \frac{\partial^2 f}{\partial x \partial y}(a,b)$$

for all $x, y \in I \times J$.

If $f \in C^4(I \times J)$, then

$$\frac{\partial^4 Af(a,b)}{\partial x^2 \partial y^2}(x,y) = 0 \quad \text{on } I \times J.$$

Taking into account Corollary 2, this motivates the name of *box-affine part* for the function $Af(a,b)$.

Compare the graph of the box convex function e^{x+y} (shown in Fig. 1) to the graph of its box-affine part $e^x + e^y + xe^y + ye^x - 1 - x - y - xy$, at the origin (shown in Fig. 2).

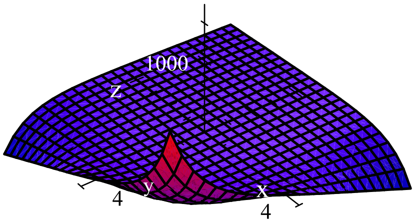

Fig. 2. The graph of the box affine part of the function e^{x+y} at the origin.

6. Jensen type integral inequalities

Since the box-convex functions are not necessarily convex, the Jensen integral inequality does not work in their context. For example

$$\int_{1/2}^{3/2}\int_{1/2}^{3/2} \frac{\log(x+y)}{x+y}dxdy \approx 0.326\,27 \text{ but } \frac{\log(1+1)}{1+1} \approx 0.346\,57.$$

However, a useful box-substitute of Jensen's inequality can be obtained by integrating the inequality stated in Theorem 9 with respect to any Borel probability measure that admits a box-barycenter.

Definition 3. The *box-barycenter* of a Borel probability measure μ defined on a compact rectangle $R = [a_1, a_2] \times [b_1, b_2]$ is the unique point $(a, b) \in R$ such that each of the following integrals

$$\iint_R (x-a)\,d\mu(x,y), \quad \iint_R (y-b)\,d\mu(x,y) \text{ and } \iint_R (x-a)(y-b)\,d\mu(x,y)$$

is equal to zero.

The fact that the first two double integrals appearing in Definition 3 vanish means that a box-barycenter (if exists) is also a barycenter in the usual sense.

Taking into account that every Borel probability measure supported by a compact convex set admits a barycenter in the usual sense, one can easily show that every product measure $\mu_1 \otimes \mu_2$ associated to Borel probability measures μ_1 and μ_2 (defined respectively on $[a_1, a_2]$ and $[b_1, b_2]$) admits as a box-barycenter the point of coordinates

$$a = \int_{a_1}^{a_2} x\,d\mu_1(x) \text{ and } b = \int_{b_1}^{b_2} y\,d\mu_2(y).$$

For $\lambda, \mu \in (0, 1)$, the discrete probability measure

$$m = (1-\lambda)(1-\mu)\delta_{(x_1,y_1)} + \lambda(1-\mu)\delta_{(x_2,y_1)} + \lambda\mu\delta_{(x_2,y_2)} + (1-\lambda)\mu\delta_{(x_1,y_2)} \tag{6.1}$$

defined on the rectangle $[x_1, x_2] \times [y_1, y_2]$ has the box-barycenter the point of coordinates

$$a = (1-\lambda)x_1 + \lambda x_2 \text{ and } b = (1-\mu)y_1 + \mu y_2.$$

However, not every Borel probability measure admits a box-barycenter. So is the case of the absolutely continuous probability measure $\frac{3}{2}(x^2 + y^2)dxdy$ on the rectangle $[0, 1] \times [0, 1]$.

By integrating both sides of the inequality stated in Theorem 9, one obtains the following Jensen type integral inequality:

Theorem 10. *If $f \in C(I \times J)$ is box-convex and μ is a Borel probability measure on a compact subrectangle $[a_1, a_2] \times [b_1, b_2]$ that admits the box-barycenter (a, b), then*

$$\int_{a_1}^{a_2}\int_{b_1}^{b_2} f(x,y)\,d\mu(x,y) \geq -f(a,b) + \int_{a_1}^{a_2}\int_{b_1}^{b_2} [f(x,b) + f(a,y)]\,d\mu(x,y).$$

For the measure defined by formula (6.1), Theorem 10 yields the Jensen type inequality (1.7) defining box-convexity.

Corollary 7. *Suppose that $(X, \mathcal{M}, \mu)$ and $(Y, \mathcal{N}, \nu)$ are two Borel probability spaces and $f : X \to [a_1, a_2]$ and $g : Y \to [b_1, b_2]$ are two Borel measurable functions. Then every box-convex function $F \in C^2([a_1, a_2]\times[b_1, b_2])$ verifies the inequality*

$$\int_X\int_Y F\left(f(x), g(y)\right) d\mu(x)d\nu(y) + F\left(\int_X f(x)d\mu(x), \int_Y g(y)d\nu(y)\right)$$
$$\geq \int_X F\left(f(x), \int_Y g(y)d\nu(y)\right) d\mu(x) + \int_Y F\left(\int_X f(x)d\mu(x), g(y)\right) d\nu(y).$$

Proof. We use the technique of pushing-forward measures applied to the product probability measure $\pi = \mu \otimes \nu$ and the π-integrable map $T : X \times Y \to [a_1, a_2] \times [b_1, b_2]$ defined by the formula

$$T(x, y) = (f(x), g(y)).$$

The push-forward measure $\lambda = T\#\pi$ is defined on $[a_1, a_2] \times [b_1, b_2]$ by the formula

$$\lambda(A \times B) = \pi(T^{-1}(A \times B))$$

and verifies

$$\iint_{[a_1,a_2]\times[b_1,b_2]} F(s,t)d\lambda(s,t) = \iint_{X\times Y} F(f(x), g(y))d\mu(x)d\nu(y)$$

for all continuous functions $F : [a_1, a_2] \times [b_1, b_2] \to \mathbb{R}$. See V. Bogachev [6], Vol. I, Section 3.6, for details. As a consequence, the barycenter of λ is the point of coordinates

$$a = \iint_{[a_1,a_2]\times[b_1,b_2]} s\,d\lambda(s,t) = \int_X f(x)d\mu(x)$$

and

$$b = \iint_{[a_1,a_2]\times[b_1,b_2]} t\,d\lambda(s,t) = \int_Y g(y)d\nu(y).$$

According to Theorem 9,

$$F(x,y) \geq -F(a,b) + F(x,b) + F(a,y) + (x-a)\frac{\partial F}{\partial u}(a,y) + (y-b)\frac{\partial F}{\partial v}(x,b)$$
$$- (x-a)\frac{\partial F}{\partial u}(a,b) - (y-b)\frac{\partial F}{\partial v}(a,b) - \frac{\partial^2 F}{\partial u \partial v}(a,b)(x-a)(y-b),$$

whence, by integrating it with respect to $\lambda = T\#\pi = (f\#\mu) \otimes (g\#\nu)$, we obtain the inequality stated by Corollary 7. □

Corollary 8. *Suppose that $p(x)dx$ and $q(y)dy$ are two absolutely continuous probability measures defined respectively on the intervals $[a_1, a_2]$ and $[b_1, b_2]$. Denote by a and b their barycenters, that is,*

$$a = \int_{a_1}^{a_2} xp(x)dx \quad \text{and} \quad b = \int_{b_1}^{b_2} yq(y)dy.$$

Then every box-convex function $F \in C([a_1, a_2] \times [b_1, b_2])$ verifies the inequality

$$\int_{a_1}^{a_2}\int_{b_1}^{b_2} F(x,y)p(x)q(y)dxdy + F(a,b) \geq \int_{a_1}^{a_2} F(x,b)p(x)dx + \int_{b_1}^{b_2} F(a,y)q(y)dy.$$

For $p(x)dx = (a_2 - a_1)^{-1}dx$, $q(y)dy = (b_2 - b_1)^{-1}dy$ the result of Corollary 8 becomes

$$\frac{1}{(a_2-a_1)(b_2-b_1)}\int_{a_1}^{a_2}\int_{b_1}^{b_2} f(x,y)dxdy + f(\frac{a_1+a_2}{2}, \frac{b_1+b_2}{2})$$
$$\geq \frac{1}{(a_2-a_1)}\int_{a_1}^{a_2} f(x, \frac{b_1+b_2}{2})dx + \frac{1}{(b_2-b_1)}\int_{b_1}^{b_2} f(\frac{a_1+a_2}{2}, y)dy.$$

For $f(x,y) = e^{x+y}$, the last inequality is equivalent to the following one,

$$L(e^{a_1}, e^{a_2})L(e^{b_1}, e^{b_2}) + e^{(a_1+a_2+b_1+b_2)/2} \geq e^{(b_1+b_2)/2}L(e^{a_1}, e^{a_2}) + e^{(a_1+a_2)/2}L(e^{b_1}, e^{b_2}),$$

where

$$L(u,v) = \begin{cases} \frac{\log v - \log u}{v-u} & \text{if } u, v > 0,\ u \neq v \\ u & \text{if } u = v > 0, \end{cases}$$

represents the logarithmic mean. Entering the geometric mean $G(u,v) = \sqrt{uv}$, one can restate the last inequality as

$$L(u_1, v_1)L(u_2, v_2) + G(u_1, v_1)\, G(u_2, v_2) \geq G(u_2, v_2)\, L(u_1, v_1) + G(u_1, v_1)\, L(u_2, v_2)$$

for all $u_1, v_1, u_2, v_2 > 0$. This shows that $L(u_1, v_1) - L(u_2, v_2)$ and $G(u_1, v_1) - G(u_2, v_2)$ have the same sign (when different from 0).

7. Open problems

The theory developed in this paper represents a first attempt towards the study of box-convexity in its own. The reader can easily generalize the results presented here to the case of N variables, by using the concepts of divided differences for functions of three or more variables.

However, many important questions remained open even in the case of two variables. Some of them are listed below.

Problem 1. Is there any box-analog of Hardy-Littlewood-Pólya theory of majorization? The proof of the box-analog of Jensen's inequality (2.1) (as detailed in Section 2) suggests an affirmative answer by replacing the use of doubly stochastic matrices with that of multi-stochastic tensors.

Problem 2. Describe the box-analogs of the classes of log-convex functions and strongly convex functions. See [18] for the classical theory of these functions.

Problem 3. As was noticed in Section 3, the box-monotonicity plays a role in a series of fields like probability and statistics, risk management and combinatorial optimization. What are the practical applications of the theory of box-convexity? Since the concepts of box-monotonicity and box-convexity are dual to each other (via derivation and integration) it seems reasonable to believe in the existence of such applications. At the moment the authors are aware only of some few results involving the presence of this theory in shape preserving approximation and the theory of double Fourier series. See [10] and respectively [2], [3] and [19] for details.

Problem 4. As is well known, the Cauchy problem for the homogeneous wave equation

$$\begin{aligned} \frac{\partial^2 u}{\partial t^2}(x,t) &= c^2 \frac{\partial^2 u}{\partial x^2}(x,t) \quad \text{for } t \geq 0,\ x \in \mathbb{R} \\ u(x,0) &= \Phi(x) \quad \text{for } x \in \mathbb{R} \\ \frac{\partial u}{\partial t}(x,0) &= \Psi(x) \quad \text{for } x \in \mathbb{R} \end{aligned}$$

with continuous initial data and a positive coefficient $c > 0$ has the solution

$$u(x,t) = \frac{1}{2}[\Phi(x+ct) + \Phi(x-ct)] + \frac{1}{2c}\int_{x-ct}^{x+ct} \Psi(s)ds.$$

See L. C. Evans [8], pp. 67-68. Then is a simple exercise to prove the following two assertions showing how the properties of monotonicity and convexity of the initial data transfer to the solution u:

(i) If Φ is continuous and convex of order 3 in the sense of Popoviciu and Ψ is continuous and nondecreasing on $\mathbb{R}$, then the solution u is box-monotone;

(ii) If Φ and Ψ are continuous and convex of order 4 in the sense of Popoviciu, then the solution of u is box-convex.

Find more general results *describing* when the solution of a boundary value problem is box-monotone and/or box-convex.

References

[1] A.G. Aksoy, M. Martelli, Mixed partial derivatives and Fubini's theorem, College Math. J. 33 (2002) 126–130.

[2] J.M. Ash, A new, harder proof that continuous functions with Schwarz derivative 0 are lines, in: W.O. Bray, C.V. Stanojevic (Eds.), Contemporary Aspects of Fourier Analysis, 1994, pp. 35–46.
[3] J.M. Ash, J. Cohen, C. Freiling, D. Rinne, Generalizations of the wave equation, Trans. Amer. Math. Soc. 338 (1) (1993) 57–75.
[4] F. Bach, Learning with submodular functions: a convex optimization perspective, Preprint, arXiv:1111.6453.
[5] M. Bencze, C.P. Niculescu, F. Popovici, Popoviciu's inequality for functions of several variables, J. Math. Anal. Appl. 365 (1) (2010) 399–409.
[6] V.I. Bogachev, Measure Theory, vol. 1, Springer Verlag, Berlin, 2007.
[7] P. Embrechts, Copulas: a personal view, J. Risk Insur. 76 (2009) 639–650.
[8] L.C. Evans, Partial Differential Equations, second edition, Graduate Studies in Mathematics, vol. 19, Amer. Math. Soc., 2010.
[9] S. Fujishige, Submodular Functions and Optimization, 2nd ed., Elsevier, 2005.
[10] S.G. Gal, Shape Preserving Approximation by Real and Complex Polynomials, Birkhäuser, Boston, 2008.
[11] C. Genest, J. Nešlehová, A primer on copulas for count data, Astin Bull. 37 (2007) 475–515.
[12] G.H. Hardy, On the convergence of certain multiple series, Proc. Lond. Math. Soc. 1 (1904) 124–128.
[13] G.H. Hardy, On double Fourier series and especially those which represent the double zeta-function with real and incommensurable parameters, Q. J. Math. 37 (1906) 53–79.
[14] E. Hopf, Über die Zusammenhänge zwischen gewissen höheren Differenzenquotienten reeller Funktionen einer reellen Variablen und deren Differenzierbarkeitseigenschaften, Dissertation, Univ. Berlin, 1926.
[15] K.S. Miller, S.G. Samko, Completely monotonic functions, Integral Transforms Spec. Funct. 12 (4) (2001) 389–402.
[16] R.B. Nelsen, An Introduction to Copulas, 2nd ed., Springer, 2006.
[17] C.P. Niculescu, The Abel-Steffensen inequality in higher dimensions, Carpathian J. Math. 35 (1) (2019) 69–78.
[18] C.P. Niculescu, L.-E. Persson, Convex Functions and Their Applications. A Contemporary Approach, 2nd ed., CMS Books in Mathematics, vol. 23, Springer-Verlag, New York, 2018.
[19] C.P. Niculescu, I. Rovenţa, Convex functions and Fourier coefficients, Positivity (2019), https://doi.org/10.1007/s11117-019-00670-8.
[20] E.T. Olsen, W.F. Darsow, B. Nguyen, Copulas and Markov operators, in: L. Ruschendorf, B. Schweizer, M.D. Taylor (Eds.), Proceedings of the Conference on Distributions with Fixed Marginals and Related Topics, in: IMS Lecture Notes and Monograph Series, vol. 28, 1996, pp. 244–259.
[21] T. Popoviciu, Sur quelques propriétés des fonctiones d'une ou de deux variables réelles, Thèse, Faculté des Sciences de Paris, 1933, see also Mathematica (Cluj) VIII (1934) 1–85.
[22] T. Popoviciu, Sur certaines inégalités qui caractérisent les fonctions convexes, An. Ştiinţ. Univ. Al. I. Cuza Iaşi. Secţ. Mat. 11 (1965) 155–164.
[23] D.M. Topkis, Supermodularity and Complementarity, Princeton University Press, Princeton, 1998.
[24] W.H. Young, On multiple integrals, Proc. R. Soc. Ser. A 93 (647) (1917) 28–41.