

Jensen convex functions and doubly stochastic matrices

MÁTYÁS BARCZY^{*,}, ZSOLT PÁLES^{**,}

* HUN-REN–SZTE Analysis and Applications Research Group, Bolyai Institute, University of Szeged, Aradi vértanúk tere 1, H–6720 Szeged, Hungary.

** Institute of Mathematics, University of Debrecen, Pf. 400, H–4002 Debrecen, Hungary.

E-mails: barczy@math.u-szeged.hu (M. Barczy), pales@science.unideb.hu (Zs. Páles).

◇ Corresponding author.

Abstract

Given an $n \times n$ doubly stochastic matrix P satisfying an appropriate condition of linear algebraic-type, and a function f defined on a nonempty interval, we show that the validity of a convexity-type functional inequality for f in terms P implies that f is Jensen convex. We also prove that if f is convex, then the functional inequality in question holds for all doubly stochastic matrices of any order. The particular case when the doubly stochastic matrix is a circulant one is also considered.

1 Introduction

The theory of convex functions is an old and important field of classical analysis, and it has a large number of applications in many branches of mathematics. For a recent book containing numerous aspects, generalizations and applications connected to convex functions, see, e.g., Niculescu and Persson [7]. In this paper, we establish some new connections between Jensen convex functions defined on a nonempty open interval and doubly stochastic matrices in terms of a functional inequality of convexity type, and we also specialize our results to the setting of doubly stochastic circulant matrices.

For a comprehensive treatment of doubly stochastic matrices and Schur convex functions, we refer to the monograph [6] by Marshall and Olkin. Concerning some recent inequalities for convex functions and doubly stochastic matrices, we can mention the paper [8] of Niezgoda (see also the subsequent ones that can be found in the literature). According to our knowledge, in all of these papers, given a convex function and a doubly stochastic matrix, Niezgoda derived some inequalities for the given convex function, but we could not find any result for the reversed directions which would be similar to the assertion (i) of our main Theorem 2.1.

2020 Mathematics Subject Classifications: 39B62, 26A51.

Key words and phrases: Jensen convex function, functional inequality, doubly stochastic matrix, circulant matrix, Schur convexity.

Our paper is organized as follows. In Section 2, as the main result of the paper, we show that, given an $n \times n$ doubly stochastic matrix P satisfying a certain condition of linear algebraic type, and a function f defined on a nonempty interval, if a convexity-type functional inequality holds for f in terms of P , then f has to be Jensen convex (see part (i) of our main Theorem 2.1). Furthermore, we also prove that if f is convex, then the functional inequality in question holds for all doubly stochastic matrices of any order (see part (ii) of Theorem 2.1). In Remark 2.3, we point out a connection between Theorem 2.1 and the notion of Schur convexity with respect to the doubly stochastic matrix with identical entries. Section 3 is devoted to derive some corollaries of Theorem 2.1 for doubly stochastic circulant matrices (see Proposition 3.1 and Corollary 3.3). We also study in detail the cases corresponding to 2×2 , 3×3 and 4×4 doubly stochastic circulant matrices whose entries are not identically equal to each other, respectively (see Corollaries 3.5, 3.6 and Proposition 3.7).

Throughout this paper, let $\mathbb{N}$, $\mathbb{Z}_+$, $\mathbb{R}$ and $\mathbb{R}_+$ denote the sets of positive integers, non-negative integers, real numbers and non-negative real numbers, respectively. For a finite subset A of $\mathbb{N}$, its cardinality is denoted by $|A|$. An interval $I \subseteq \mathbb{R}$ will be called nondegenerate if it contains at least two distinct points. Given a non-degenerate interval $I \subseteq \mathbb{R}$, the difference $I - I$ stands for the set $\{x - y : x, y \in I\}$. To avoid misunderstandings, in some case we write $(I - I)$ instead of $I - I$. The natural basis in $\mathbb{R}^n$ is denoted by $\{\mathbf{e}_1, \dots, \mathbf{e}_n\}$, where $n \in \mathbb{N}$.

Next, we recall the notion of t -convex functions (where $t \in [0, 1]$).

1.1 Definition. Let $I \subseteq \mathbb{R}$ be a nondegenerate interval and $t \in [0, 1]$. A function $f : I \rightarrow \mathbb{R}$ is called t -convex if

$$f(tx + (1 - t)y) \leq tf(x) + (1 - t)f(y), \quad x, y \in I.$$

If f is $\frac{1}{2}$ -convex, then it is called Jensen convex or midpoint convex as well. If f is t -convex for all $t \in (0, 1)$, then it is said to be convex. Note that every function $f : I \rightarrow \mathbb{R}$ is automatically 0-convex and 1-convex as well.

In the next result, we collect some known implications among the convexity notions introduced in Definition 1.1.

1.2 Theorem. Let $I \subseteq \mathbb{R}$ be a nonempty open interval and $f : I \rightarrow \mathbb{R}$. Then the following two assertions hold.

- (i) If f is t -convex for some $t \in (0, 1)$, then it is Jensen convex.
- (ii) If f is t -convex for some $t \in (0, 1)$ and f is bounded from above on some nonempty open subinterval of I , then f is convex and continuous.

For a proof of Theorem 1.2, see the proof of Theorem 1.1 in Barczy and Páles [1].

2 Jensen convexity and doubly stochastic matrices

As the main result of the paper, we establish a connection between Jensen convex functions and doubly stochastic matrices in terms of a functional inequality of convexity-type.

2.1 Theorem. *Let $I \subseteq \mathbb{R}$ be a nonempty open interval and $f : I \rightarrow \mathbb{R}$ be a function.*

- (i) *Let $n \in \mathbb{N}$ and let $P = (p_{i,j}) \in \mathbb{R}^{n \times n}$ be a doubly stochastic matrix and assume that there exist disjoint nonempty subsets $A, B \subseteq \{1, \dots, n\}$ such that $\frac{1}{|A|} \sum_{\alpha \in A} \mathbf{e}_\alpha - \frac{1}{|B|} \sum_{\beta \in B} \mathbf{e}_\beta$ belongs to the linear span of the column vectors of P . If the inequality*

$$(2.1) \quad f\left(\frac{x_1 + \dots + x_n}{n}\right) \leq \frac{1}{n} \sum_{i=1}^n f(p_{i,1}x_1 + \dots + p_{i,n}x_n)$$

holds for all $x_1, \dots, x_n \in I$, then the function f is Jensen convex on I .

- (ii) *If the function f is Jensen convex on I , then the inequality (2.1) holds for all $n \in \mathbb{N}$, for all doubly stochastic matrices $P = (p_{i,j}) \in \mathbb{R}^{n \times n}$ and for all $x_1, \dots, x_n \in I$.*

Proof. First, we prove part (i). Let $n \in \mathbb{N}$, let $P = (p_{i,j}) \in \mathbb{R}^{n \times n}$ be a doubly stochastic matrix, and let $A, B \subseteq \{1, \dots, n\}$ disjoint nonempty sets such that $\frac{1}{|A|} \sum_{\alpha \in A} \mathbf{e}_\alpha - \frac{1}{|B|} \sum_{\beta \in B} \mathbf{e}_\beta$ belongs to the linear span of the column vectors of P . Therefore, there exists $\mathbf{x}_{A,B} \in \mathbb{R}^n$ such that $\frac{1}{|A|} \sum_{\alpha \in A} \mathbf{e}_\alpha - \frac{1}{|B|} \sum_{\beta \in B} \mathbf{e}_\beta = P\mathbf{x}_{A,B}$. For the sake of simplicity, denote $a := |A|$ and $b := |B|$. Suppose that (2.1) holds for all $x_1, \dots, x_n \in I$.

Given $\mathbf{x} = (x_1, \dots, x_n) \in I^n$, introduce the vector $\mathbf{y} = (y_1, \dots, y_n)$ by

$$(2.2) \quad y_i := p_{i,1}x_1 + \dots + p_{i,n}x_n, \quad i \in \{1, \dots, n\}.$$

In other words, $\mathbf{y} = P\mathbf{x}$. Using that the rows of P are finite probability distributions, it follows that y_i is a convex combination of $x_1, \dots, x_n$ for each $i \in \{1, \dots, n\}$, and hence $y_i \in I$, $i \in \{1, \dots, n\}$, in other words, $\mathbf{y} \in I^n$.

Using also that the column-sums of P are equal to 1, we have that

$$y_1 + \dots + y_n = x_1 + \dots + x_n,$$

and hence (2.1) can be rewritten as

$$(2.3) \quad f\left(\frac{y_1 + \dots + y_n}{n}\right) \leq \frac{1}{n} \sum_{i=1}^n f(y_i)$$

for all $\mathbf{y} = (y_1, \dots, y_n) \in \{P\mathbf{x} \mid \mathbf{x} \in I^n\} \subseteq I^n$.

We are going to show that f is locally Jensen convex on I , i.e., for all $p \in I$, there exists $r > 0$ such that $(p - r, p + r) \subseteq I$ and f is Jensen convex on $(p - r, p + r)$. Let $p \in I$ be fixed.

For $u, v \in I$, define

$$(2.4) \quad y_i := \begin{cases} u & \text{if } i \in A, \\ v & \text{if } i \in B, \\ \frac{au + bv}{a + b} & \text{if } i \in \{1, \dots, n\} \setminus (A \cup B). \end{cases}$$

Using that the map $\mathbb{R}^n \ni \mathbf{x} \mapsto P\mathbf{x}$ is linear and the diagonal elements of I^n are fixed points of this map, we can obtain that

$$(2.5) \quad \begin{aligned} (y_1, \dots, y_n) &= \frac{au + bv}{a + b}(\mathbf{e}_1 + \dots + \mathbf{e}_n) + \frac{ab(u - v)}{a + b} \left(\frac{1}{a} \sum_{\alpha \in A} \mathbf{e}_\alpha - \frac{1}{b} \sum_{\beta \in B} \mathbf{e}_\beta \right) \\ &= \frac{au + bv}{a + b} P(\mathbf{e}_1 + \dots + \mathbf{e}_n) + \frac{ab(u - v)}{a + b} P\mathbf{x}_{A,B} \\ &= P \left(\frac{au + bv}{a + b} (\mathbf{e}_1 + \dots + \mathbf{e}_n) + \frac{ab(u - v)}{a + b} \mathbf{x}_{A,B} \right). \end{aligned}$$

Further, we have that

$$\lim_{(u,v) \rightarrow (p,p)} \left(\frac{au + bv}{a + b} (\mathbf{e}_1 + \dots + \mathbf{e}_n) + \frac{ab(u - v)}{a + b} \mathbf{x}_{A,B} \right) = p(\mathbf{e}_1 + \dots + \mathbf{e}_n) = (p, \dots, p) \in I^n.$$

Consequently, since I and I^n are open, there exists $r > 0$ such that $(p - r, p + r) \subseteq I$ and

$$\frac{au + bv}{a + b} (\mathbf{e}_1 + \dots + \mathbf{e}_n) + \frac{ab(u - v)}{a + b} \mathbf{x}_{A,B} \in I^n \quad \text{for all } u, v \in (p - r, p + r),$$

and hence, by (2.5),

$$(y_1, \dots, y_n) \in \{P\mathbf{x} \mid \mathbf{x} \in I^n\} \quad \text{for all } u, v \in (p - r, p + r),$$

where $(y_1, \dots, y_n)$ is defined in (2.4). We have that

$$\frac{y_1 + \dots + y_n}{n} = \frac{1}{n} \left(au + bv + (n - a - b) \frac{au + bv}{a + b} \right) = \frac{au + bv}{a + b}.$$

Using this identity, the inequality (2.3) takes the form

$$f \left(\frac{au + bv}{a + b} \right) \leq \frac{1}{n} \left(af(u) + bf(v) + (n - a - b) f \left(\frac{au + bv}{a + b} \right) \right),$$

which reduces to

$$(2.6) \quad f \left(\frac{au + bv}{a + b} \right) \leq \frac{af(u) + bf(v)}{a + b}$$

for all $u, v \in (p - r, p + r)$, that is, f is $\frac{a}{a+b}$ -convex on $(p - r, p + r)$, where $\frac{a}{a+b} \in (0, 1)$. According to part (i) of Theorem 1.2, it follows that f is Jensen convex on $(p - r, p + r)$.

All in all, we have proved that, for all $p \in I$, there exists $r > 0$ such that $(p - r, p + r) \subseteq I$ and f is Jensen convex on $(p - r, p + r)$, in other words, f is locally Jensen convex on I . This implies that f is Jensen convex on I , i.e.,

$$f\left(\frac{u+v}{2}\right) \leq \frac{f(u) + f(v)}{2}, \quad u, v \in I,$$

see, for example, Corollary 2 with the choices $n = 2$ and $t_1 = t_2 = 1$ in Gilányi and Páles [4] or Corollary 6 with the choice $t = \frac{1}{2}$ in Nikodem and Páles [9] (which, roughly speaking, states that Jensen convexity is a localizable convexity property).

Finally, we prove part (ii). Suppose that the function $f : I \rightarrow \mathbb{R}$ is Jensen convex. Then, for all $n \in \mathbb{N}$, for all doubly stochastic matrices $P = (p_{i,j}) \in \mathbb{R}^{n \times n}$ and for all $x_1, \dots, x_n \in I$, we have that

$$\sum_{i=1}^n (p_{i,1}x_1 + \dots + p_{i,n}x_n) = x_1 + \dots + x_n,$$

yielding that

$$\frac{1}{n} \sum_{k=1}^n x_k = \frac{1}{n} \sum_{i=1}^n (p_{i,1}x_1 + \dots + p_{i,n}x_n).$$

Now the Jensen convexity of f together with Lemma 5.3.1 in Kuczma [5] imply that the inequality (2.1) holds. $\square$

In the next remark, among others, we provide a sufficient condition under which the condition on P in part (i) of Theorem 2.1 holds, and we also formulate an open problem.

2.2 Remark. (i). If $P \in \mathbb{R}^{n \times n}$ is an invertible doubly stochastic matrix, then, for all disjoint nonempty subsets $A, B \subseteq \{1, \dots, n\}$, we have that $\frac{1}{|A|} \sum_{\alpha \in A} \mathbf{e}_\alpha - \frac{1}{|B|} \sum_{\beta \in B} \mathbf{e}_\beta$ belongs to the linear span of the column vectors of P , i.e., the condition in part (i) of Theorem 2.1 holds trivially. Note also that, for any disjoint nonempty subsets $A, B \subseteq \{1, \dots, n\}$, the vectors $\frac{1}{|A|} \sum_{\alpha \in A} \mathbf{e}_\alpha - \frac{1}{|B|} \sum_{\beta \in B} \mathbf{e}_\beta$ and $\mathbf{e}_1 + \dots + \mathbf{e}_n$ are linearly independent, and $\mathbf{e}_1 + \dots + \mathbf{e}_n$ belongs to the linear span of the column vectors of P (since the row-sums of P are equal to 1). Consequently, under the condition on P in part (i) of Theorem 2.1, the rank of P is at least 2. However, we do not know whether the assertion in part (i) of Theorem 2.1 holds for every doubly stochastic matrix having rank at least 2. We leave this question as an open problem.

(ii). If $n \in \mathbb{N}$, then, by the trinomial theorem, the number of ways how one can choose disjoint (possibly empty) subsets A and B of $\{1, \dots, n\}$ is

$$\sum_{a,b \in \mathbb{Z}_+ : a+b \leq n} \binom{n}{a} \binom{n-a}{b} = \sum_{a,b \in \mathbb{Z}_+ : a+b \leq n} \frac{n!}{a!b!(n-a-b)!} 1^a 1^b 1^{n-a-b} = (1+1+1)^n = 3^n.$$

On the other hand, the number of ways how one can choose a (possibly empty) subset (A or B) of $\{1, \dots, n\}$ is $\sum_{a=0}^n \binom{n}{a} = 2^n$. Therefore, by the inclusion–exclusion principle (which is sometimes

referred to as the sieve formula), the cardinality of the set of vectors $\frac{1}{|A|} \sum_{\alpha \in A} \mathbf{e}_\alpha - \frac{1}{|B|} \sum_{\beta \in B} \mathbf{e}_\beta$ that satisfy the assumption in part (i) of Theorem 2.1 is $3^n - 2 \cdot 2^n + 1$, which is of order 3^n .

(iii) If the assumptions of part (i) of Theorem 2.1 hold, f is bounded from above on some nonempty open subset of I and the inequality (2.1) is satisfied for all $x_1, \dots, x_n \in I$, then f is convex on I . This statement easily follows from part (i) of Theorem 2.1 combined with part (ii) of Theorem 1.2. $\square$

The next remark is devoted to give a reformulation of the inequality (2.1) in Theorem 2.1, and to point out a connection between Theorem 2.1 and the notion of Schur convexity with respect to the doubly stochastic matrix with identical entries.

2.3 Remark. Let $I \subseteq \mathbb{R}$ be a nonempty open interval, $n \in \mathbb{N}$, and $f : I \rightarrow \mathbb{R}$ be a function. Let us define $F : I^n \rightarrow \mathbb{R}$,

$$(2.7) \quad F(x_1, \dots, x_n) := f(x_1) + \dots + f(x_n), \quad (x_1, \dots, x_n) \in I^n.$$

Let S_n denote the doubly stochastic matrix with identical entries equal to $\frac{1}{n}$. Then, for a doubly stochastic matrix $P \in \mathbb{R}^{n \times n}$ and a vector $\mathbf{x} \in I^n$, the inequality (2.1) in Theorem 2.1 can be also written in the form

$$(2.8) \quad F(S_n \mathbf{x}) \leq F(P \mathbf{x}).$$

We claim that, for an arbitrary function $G : I^n \rightarrow \mathbb{R}$, the inequality

$$(2.9) \quad G(S_n \mathbf{x}) \leq G(P \mathbf{x})$$

is equivalent to

$$(2.10) \quad G(S_n \mathbf{y}) \leq G(\mathbf{y}), \quad \mathbf{y} \in P(I^n) = \{P \mathbf{x} \mid \mathbf{x} \in I^n\} \subseteq I^n.$$

First, suppose that (2.10) holds. Then, using that $S_n P = S_n$ (since P is doubly stochastic), with $\mathbf{y} := P \mathbf{x}$, where $\mathbf{x} \in I^n$, we get that

$$G(S_n \mathbf{x}) = G(S_n P \mathbf{x}) \leq G(P \mathbf{x}),$$

as desired. Now suppose that (2.9) holds. Then, using again that $S_n P = S_n$, we have that

$$G(S_n P \mathbf{x}) = G(S_n \mathbf{x}) \leq G(P \mathbf{x}), \quad \mathbf{x} \in I^n,$$

and hence (2.10) is valid for all $\mathbf{y} \in P(I^n)$, as desired.

Furthermore, if the condition on P in part (i) of Theorem 2.1 is satisfied and the inequality (2.8) holds for all $\mathbf{x} \in I^n$ (where F is defined by (2.7)), then f is Jensen convex on I (following from Theorem 2.1), and, hence, using also Lemma 5.3.1 in Kuczma [5], it implies that

$$(2.11) \quad \begin{aligned} F(S_n \mathbf{x}) &= F\left(\frac{x_1 + \dots + x_n}{n}, \dots, \frac{x_1 + \dots + x_n}{n}\right) = n f\left(\frac{x_1 + \dots + x_n}{n}\right) \\ &\leq n \cdot \frac{1}{n} \sum_{i=1}^n f(x_i) = \sum_{i=1}^n f(x_i) = F(\mathbf{x}), \quad \mathbf{x} \in I^n. \end{aligned}$$

According to Burai et al. [2, Definition 3], if the inequality (2.11) holds, then we say that F is S_n -Schur convex. $\square$

3 Jensen convexity and doubly stochastic circulant matrices

In this section, we derive some corollaries of Theorem 2.1 for doubly stochastic circulant-type matrices (also called left-circulant matrices). To simplify the formulation of our subsequent results, for each $n \in \mathbb{N}$, $i \in \{1, \dots, n\}$ and $j \in \{0, \dots, n-1\}$, let us define the truncated sum $i \oplus j$ of i and j by

$$(3.1) \quad i \oplus j := \begin{cases} i + j & \text{if } i + j \leq n, \\ i + j - n & \text{if } i + j > n. \end{cases}$$

3.1 Proposition. *Let $I \subseteq \mathbb{R}$ be a nonempty open interval and $f : I \rightarrow \mathbb{R}$ be a function.*

(i) *Let $n \in \mathbb{N}$ and denote $\omega_n := e^{\frac{2\pi i}{n}}$. Let $\lambda_1, \dots, \lambda_n \in [0, 1]$ such that $\lambda_1 + \dots + \lambda_n = 1$ and*

$$(3.2) \quad \sum_{j=1}^n \lambda_j \omega_n^{k(j-1)} \neq 0, \quad k \in \{1, \dots, n-1\}.$$

If the inequality

$$(3.3) \quad f\left(\frac{x_1 + \dots + x_n}{n}\right) \leq \frac{1}{n} \sum_{j=0}^{n-1} f(\lambda_{1 \oplus j} x_1 + \dots + \lambda_{n \oplus j} x_n)$$

holds for all $x_1, \dots, x_n \in I$, then the function f is Jensen convex on I , where $\oplus$ is defined in (3.1).

(ii) *If the function f is Jensen convex on I , then the inequality (3.3) holds for all $n \in \mathbb{N}$, for all $\lambda_1, \dots, \lambda_n \in [0, 1]$ with $\lambda_1 + \dots + \lambda_n = 1$ and $x_1, \dots, x_n \in I$.*

Proof. Let $n \in \mathbb{N}$ and $\lambda_1, \dots, \lambda_n \in [0, 1]$ be such that $\lambda_1 + \dots + \lambda_n = 1$. Let us define

$$(3.4) \quad P := (\lambda_{i \oplus j})_{i=1, \dots, n}^{j=0, \dots, n-1} = \begin{bmatrix} \lambda_1 & \lambda_2 & \lambda_3 & \cdots & \lambda_{n-1} & \lambda_n \\ \lambda_2 & \lambda_3 & \lambda_4 & \cdots & \lambda_n & \lambda_1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \lambda_n & \lambda_1 & \lambda_2 & \cdots & \lambda_{n-2} & \lambda_{n-1} \end{bmatrix} \in \mathbb{R}^{n \times n}.$$

Then P is a left circulant matrix with parameters $\lambda_1, \dots, \lambda_n$, see Carmona et al. [3, formula (5)]. By part (vii) of Lemma 1.1 and Lemma 1.2 in Carmona et al. [3], we get that P is invertible if and only if $\sum_{j=1}^n \lambda_j \omega_n^{k(j-1)} \neq 0$, $k \in \{0, \dots, n-1\}$. Since $\lambda_1 + \dots + \lambda_n = 1$, it yields

that P is invertible if and only if (3.2) holds. Further, P is doubly stochastic as well due to that $\lambda_1 + \dots + \lambda_n = 1$. Moreover, we have that

$$\begin{aligned} \sum_{i=1}^n f(\lambda_{i \oplus 0} x_1 + \dots + \lambda_{i \oplus (n-1)} x_n) &= \sum_{j=0}^{n-1} f(\lambda_{(j+1) \oplus 0} x_1 + \dots + \lambda_{(j+1) \oplus (n-1)} x_n) \\ &= \sum_{j=0}^{n-1} f(\lambda_{1 \oplus j} x_1 + \dots + \lambda_{n \oplus j} x_n). \end{aligned}$$

Therefore the statement follows from Theorem 2.1 taking into account part (i) of Remark 2.2 as well. $\square$

In the first part of the next remark, we point out the fact that if, in addition, f is bounded from above on some nonempty open subset of I , then in part (i) of Proposition 3.1 we can obtain that f is convex on I . The second part of next remark is about the case when the condition (3.2) does not hold.

3.2 Remark. (i). If the assumptions of part (i) of Proposition 3.1 hold, f is bounded from above on some nonempty open subset of I and the inequality (3.3) is satisfied for all $x_1, \dots, x_n \in I$, then f is convex on I . This statement easily follows from part (i) of Proposition 3.1 combined with part (ii) of Theorem 1.2.

(ii). Let $I \subseteq \mathbb{R}$ be a nonempty open interval, $f : I \rightarrow \mathbb{R}$ be a function, $n \in \mathbb{N} \setminus \{1\}$, and $\lambda_1, \dots, \lambda_n \in [0, 1]$ be such that $\lambda_1 + \dots + \lambda_n = 1$. Let us consider the matrix P given by (3.4). By the proof of Proposition 3.1, P is not invertible if and only if the condition (3.2) does not hold, i.e., the set

$$K := \left\{ k \in \{1, \dots, n-1\} : \sum_{j=1}^n \lambda_j \omega_n^{k(j-1)} = 0 \right\}$$

is nonempty. Since

$$\omega_n^{k(j-1)} = \exp\left(\frac{2k(j-1)\pi}{n} \mathbf{i}\right) = \cos\left(\frac{2k(j-1)\pi}{n}\right) + \mathbf{i} \sin\left(\frac{2k(j-1)\pi}{n}\right), \quad j \in \{1, \dots, n\},$$

we get that

$$(3.5) \quad \sum_{j=1}^n \lambda_j \cos\left(\frac{2k(j-1)\pi}{n}\right) = 0, \quad \sum_{j=1}^n \lambda_j \sin\left(\frac{2k(j-1)\pi}{n}\right) = 0, \quad k \in K.$$

Observe that, if n is even and $k = n/2$, then the second condition in (3.5) is trivial, while the first condition in (3.5) together with $\lambda_1 + \dots + \lambda_n = 1$ yields

$$(3.6) \quad \lambda_1 + \lambda_3 + \dots + \lambda_{n-1} = \lambda_2 + \lambda_4 + \dots + \lambda_n = \frac{1}{2}.$$

$\square$

Next, we present a corollary of Proposition 3.1.

3.3 Corollary. Let $I \subseteq \mathbb{R}$ be a nonempty open interval and $f : I \rightarrow \mathbb{R}$ be a function.

(i) Let $n \in \mathbb{N} \setminus \{1\}$. If the inequality

$$(3.7) \quad f\left(\frac{x_1 + \cdots + x_n}{n}\right) \leq \frac{1}{n} \sum_{i=1}^n f\left(\frac{x_1 + \cdots + x_n - x_i}{n-1}\right)$$

holds for all $x_1, \dots, x_n \in I$, then the function f is Jensen convex on I .

(ii) If the function f is Jensen convex on I , then the inequality (3.7) holds for all $n \in \mathbb{N} \setminus \{1\}$ and $x_1, \dots, x_n \in I$.

Proof. First, we prove part (i). Let $n \in \mathbb{N} \setminus \{1\}$. Let us apply part (i) of Proposition 3.1 with choices $\lambda_1 = \cdots = \lambda_{n-1} := \frac{1}{n-1}$ and $\lambda_n := 0$. Then, for each $k \in \{1, \dots, n-1\}$, we have that

$$\sum_{j=1}^n \lambda_j \omega_n^{k(j-1)} = \frac{1}{n-1} \sum_{j=1}^{n-1} e^{\frac{2\pi k(j-1)i}{n}} = \frac{1}{n-1} \frac{e^{\frac{2\pi k(n-1)i}{n}} - 1}{e^{\frac{2\pi ki}{n}} - 1} \neq 0,$$

yielding that condition (3.2) holds. Further, for all $x_1, \dots, x_n \in I$ and each $j \in \{0, \dots, n-1\}$, we have that

$$\lambda_{1 \oplus j} x_1 + \cdots + \lambda_{n \oplus j} x_n = \frac{x_1 + \cdots + x_n - x_{j_*}}{n-1},$$

where $j_* \in \{1, \dots, n\}$ is such that $j_* \oplus j = n$. Taking into account that $\{x_{j_*} : j \in \{0, \dots, n-1\}\} = \{x_1, \dots, x_n\}$, part (i) of Proposition 3.1 implies part (i).

The second assertion follows immediately from assertion (ii) of Proposition 3.1 with choices $\lambda_1 = \cdots = \lambda_{n-1} := \frac{1}{n-1}$ and $\lambda_n := 0$. $\square$

3.4 Remark. If the assumptions of Corollary 3.3 hold, f is bounded from above on some nonempty open subset of I , $n \in \mathbb{N} \setminus \{1\}$, and the inequality (3.7) is satisfied for all $x_1, \dots, x_n \in I$, then f is convex on I . Indeed, by part (i) of Corollary 3.3, f is Jensen convex on I , and, using part (ii) of Theorem 1.2, it implies that f is convex on I , as desired. $\square$

Next, we specialize Proposition 3.1 to the cases $n = 2$ and $n = 3$, respectively.

3.5 Corollary. Let $I \subseteq \mathbb{R}$ be a nonempty open interval and $f : I \rightarrow \mathbb{R}$ be a function. Let $\lambda_1, \lambda_2 \in [0, 1]$ be such that $\lambda_1 + \lambda_2 = 1$.

(i) The condition (3.2) in case of $n = 2$ is equivalent to $\lambda_1 \neq \lambda_2$, i.e., not both of λ_1 and λ_2 are equal to $\frac{1}{2}$.

(ii) If $\lambda_1 \neq \lambda_2$ and the inequality (3.3) with $n = 2$ holds for all $x_1, x_2 \in I$, i.e.,

$$f\left(\frac{x_1 + x_2}{2}\right) \leq \frac{1}{2} \left(f(\lambda_1 x_1 + \lambda_2 x_2) + f(\lambda_2 x_1 + \lambda_1 x_2) \right), \quad x_1, x_2 \in I,$$

then the function f is Jensen convex on I .

Proof. (i). Since $\omega_2 = e^{\pi i} = -1$, the condition (3.2) in case of $n = 2$ is equivalent to $\lambda_1 - \lambda_2 \neq 0$, i.e., $\lambda_1 \neq \lambda_2$. Since $\lambda_1 + \lambda_2 = 1$, we have that $\lambda_1 \neq \lambda_2$ is equivalent to the fact that not both of λ_1 and λ_2 are equal to $\frac{1}{2}$.

(ii). Assume that $\lambda_1 \neq \lambda_2$. Then, by part (i), the condition (3.2) in case of $n = 2$ holds, and hence part (i) of Proposition 3.1 yields the assertion. $\square$

3.6 Corollary. *Let $I \subseteq \mathbb{R}$ be a nonempty open interval and $f : I \rightarrow \mathbb{R}$ be a function. Let $\lambda_1, \lambda_2, \lambda_3 \in [0, 1]$ be such that $\lambda_1 + \lambda_2 + \lambda_3 = 1$.*

(i) *The condition (3.2) in case of $n = 3$ is equivalent that $\lambda_1, \lambda_2, \lambda_3$ are not all equal to each other (i.e., not all of them are equal to $\frac{1}{3}$).*

(ii) *If $\lambda_1, \lambda_2, \lambda_3$ are not all equal to each other and the inequality (3.3) with $n = 3$ holds for all $x_1, x_2, x_3 \in I$, then the function f is Jensen convex on I .*

Proof. (i). For each $k \in \{1, 2\}$, we have that

$$\omega_3^{k(j-1)} = \exp\left(\frac{2k(j-1)\pi i}{3}\right) = \begin{cases} 1 & \text{if } j = 1, \\ e^{\frac{2k\pi i}{3}} = \left(-\frac{1}{2} + \frac{\sqrt{3}}{2}i\right)^k & \text{if } j = 2, \\ e^{\frac{4k\pi i}{3}} = \left(-\frac{1}{2} - \frac{\sqrt{3}}{2}i\right)^k & \text{if } j = 3. \end{cases}$$

Consequently, the condition (3.2) in case of $n = 3$ is equivalent to

$$\lambda_1 + \left(-\frac{1}{2} + \frac{\sqrt{3}}{2}i\right)^k \lambda_2 + \left(-\frac{1}{2} - \frac{\sqrt{3}}{2}i\right)^k \lambda_3 \neq 0, \quad k \in \{1, 2\},$$

i.e.,

$$2\lambda_1 + (-1 + \sqrt{3}i)\lambda_2 - (1 + \sqrt{3}i)\lambda_3 \neq 0 \quad \text{and} \quad 2\lambda_1 - (1 + \sqrt{3}i)\lambda_2 + (-1 + \sqrt{3}i)\lambda_3 \neq 0.$$

Using that $\lambda_1 = 1 - \lambda_2 - \lambda_3$, these conditions are equivalent to

$$(2 - 3\lambda_2 - 3\lambda_3) + \sqrt{3}(\lambda_2 - \lambda_3)i \neq 0 \quad \text{and} \quad (2 - 3\lambda_2 - 3\lambda_3) + \sqrt{3}(-\lambda_2 + \lambda_3)i \neq 0.$$

Here the first condition (and the second one as well) is equivalent to

$$(2 - 3\lambda_2 - 3\lambda_3)^2 + (\lambda_2 - \lambda_3)^2 > 0.$$

Therefore, the condition (3.2) in case of $n = 3$ does not hold if and only if $\lambda_2 + \lambda_3 = \frac{2}{3}$ and $\lambda_2 = \lambda_3$, i.e., if and only if $\lambda_2 = \lambda_3 = \frac{1}{3}$. Hence, since $\lambda_1 = 1 - \lambda_2 - \lambda_3$, we have that the condition (3.2) in case of $n = 3$ does not hold if and only if $\lambda_1 = \lambda_2 = \lambda_3 = \frac{1}{3}$, which yields part (i).

We mention that one could give an alternative proof as well. By part (vii) of Lemma 1.1 and Lemma 1.2 in Carmona et al. [3], we get that P (introduced in (3.4)) is invertible if and

only if condition (3.2) holds. Then one could calculate directly the determinant of P in case of $n = 3$, and study when it can be zero. The determinant in question can be found, e.g., in Carmona et al. [3, page 787].

(ii). Assume that $\lambda_1, \lambda_2, \lambda_3$ are not equal to each other. Then, by part (i), the condition (3.2) in case of $n = 3$ holds, and hence part (i) of Proposition 3.1 yields the assertion. $\square$

In the next proposition, on the one hand, we make the condition (3.2) in case of $n = 4$ more explicit, and, on the other hand, we prove part (i) of Proposition 3.1 without this condition for a 4×4 doubly stochastic matrix whose entries are not identically equal to $\frac{1}{4}$.

3.7 Proposition. *Let $I \subseteq \mathbb{R}$ be a nonempty open interval and $f : I \rightarrow \mathbb{R}$ be a function. Let $\lambda_1, \dots, \lambda_4 \in [0, 1]$ be such that $\lambda_1 + \dots + \lambda_4 = 1$.*

(i) *The condition (3.2) in case of $n = 4$ is equivalent to*

$$(3.8) \quad (\lambda_1 - \lambda_3)^2 + (\lambda_2 - \lambda_4)^2 > 0 \quad \text{and} \quad \lambda_1 + \lambda_3 \neq \lambda_2 + \lambda_4.$$

(ii) *If $\lambda_i, i \in \{1, \dots, 4\}$, are not all equal to each other (i.e., not all of them are equal to $\frac{1}{4}$) and the inequality (3.3) with $n = 4$ holds for all $x_1, \dots, x_4 \in I$, then the function f is Jensen convex on I .*

Note that, in general, the two conditions in (3.8) are not dependent on each other, since, for example, in the case of $\lambda_1 = \frac{1}{4}, \lambda_2 = \frac{1}{6}, \lambda_3 = \frac{1}{4}$ and $\lambda_4 = \frac{1}{3}$, we have $(\lambda_1 - \lambda_3)^2 + (\lambda_2 - \lambda_4)^2 = \frac{1}{36}$, but $\lambda_1 + \lambda_3 = \frac{1}{2} = \lambda_2 + \lambda_4$. On the other hand, if $\lambda_1 = \frac{1}{3}, \lambda_2 = \frac{1}{6}, \lambda_3 = \frac{1}{3}$ and $\lambda_4 = \frac{1}{6}$, then $(\lambda_1 - \lambda_3)^2 + (\lambda_2 - \lambda_4)^2 = 0$, but $\lambda_1 + \lambda_3 = \frac{2}{3} \neq \frac{1}{3} = \lambda_2 + \lambda_4$.

Proof of Proposition 3.7. (i). For each $k \in \{1, 2, 3\}$, we have that

$$\omega_4^{k(j-1)} = \exp\left(\frac{2k(j-1)\pi}{4}i\right) = \begin{cases} 1 & \text{if } j = 1, \\ e^{\frac{k\pi}{2}i} = i^k & \text{if } j = 2, \\ e^{k\pi i} = (-1)^k & \text{if } j = 3, \\ e^{\frac{3k\pi}{2}i} = (-i)^k & \text{if } j = 4. \end{cases}$$

Consequently, the condition (3.2) in case of $n = 4$ is equivalent to

$$\lambda_1 + i^k \lambda_2 + (-1)^k \lambda_3 + (-i)^k \lambda_4 \neq 0, \quad k \in \{1, 2, 3\},$$

i.e.,

$$\lambda_1 + i\lambda_2 - \lambda_3 - i\lambda_4 \neq 0, \quad \lambda_1 - \lambda_2 + \lambda_3 - \lambda_4 \neq 0, \quad \lambda_1 - i\lambda_2 - \lambda_3 + i\lambda_4 \neq 0,$$

where $\lambda_1, \dots, \lambda_4 \in [0, 1]$ are such that $\lambda_1 + \dots + \lambda_4 = 1$. Here the first condition (and the third condition as well) is equivalent to $(\lambda_1 - \lambda_3)^2 + (\lambda_2 - \lambda_4)^2 > 0$, and the second condition is equivalent to $\lambda_1 + \lambda_3 \neq \lambda_2 + \lambda_4$. All in all, the condition (3.2) in case of $n = 4$ is equivalent to (3.8).

(ii). If the condition (3.2) with $n = 4$ holds, then the statement follows by part (i) of Proposition 3.1.

If the condition (3.2) with $n = 4$ does not hold, then we are going to show that part (i) of Theorem 2.1 can be applied. Assume that the condition (3.2) in case of $n = 4$ does not hold, i.e.,

$$\sum_{j=1}^4 \lambda_j \omega_4^{k(j-1)} = 0 \quad \text{for some } k \in \{1, 2, 3\}.$$

Then, by part (i), we have that

$$(\lambda_1 - \lambda_3)^2 + (\lambda_2 - \lambda_4)^2 = 0 \quad \text{or} \quad \lambda_1 + \lambda_3 = \lambda_2 + \lambda_4.$$

Case 1. If $(\lambda_1 - \lambda_3)^2 + (\lambda_2 - \lambda_4)^2 = 0$, then $\lambda_1 = \lambda_3$ and $\lambda_2 = \lambda_4$, yielding that the matrix P introduced in (3.4) takes the form

$$P = \begin{bmatrix} \lambda_1 & \lambda_2 & \lambda_1 & \lambda_2 \\ \lambda_2 & \lambda_1 & \lambda_2 & \lambda_1 \\ \lambda_1 & \lambda_2 & \lambda_1 & \lambda_2 \\ \lambda_2 & \lambda_1 & \lambda_2 & \lambda_1 \end{bmatrix},$$

where $\lambda_1, \lambda_2 \in \mathbb{R}_+$ are such that $\lambda_1 + \lambda_2 = \frac{1}{2}$. If $\lambda_1 = \lambda_2$, i.e., $\lambda_1 = \lambda_2 = \frac{1}{4}$ were true, then all the entries of P would be equal to $\frac{1}{4}$, but we excluded this case. Consequently, we have that $\lambda_1 \neq \lambda_2$. Then the linear span of the column vectors of P is

$$\left\{ \alpha \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_1 \\ \lambda_2 \end{pmatrix} + \beta \begin{pmatrix} \lambda_2 \\ \lambda_1 \\ \lambda_2 \\ \lambda_1 \end{pmatrix} : \alpha, \beta \in \mathbb{R} \right\} = \left\{ \begin{pmatrix} u \\ v \\ u \\ v \end{pmatrix} : u, v \in \mathbb{R} \right\} = \left\{ u(\mathbf{e}_1 + \mathbf{e}_3) + v(\mathbf{e}_2 + \mathbf{e}_4) : u, v \in \mathbb{R} \right\},$$

where, at the first equality, we used that, for any $u, v \in \mathbb{R}$, the system of linear equations $\alpha\lambda_1 + \beta\lambda_2 = u$ and $\alpha\lambda_2 + \beta\lambda_1 = v$ can be solved uniquely for $\alpha, \beta \in \mathbb{R}$ (since $\lambda_1 \neq \lambda_2$). Consequently, the condition on P in part (i) of Theorem 2.1 is satisfied by choosing $A := \{1, 3\}$ and $B := \{2, 4\}$, and hence we can apply part (i) of Theorem 2.1 in case of $\lambda_1 = \lambda_3$ and $\lambda_2 = \lambda_4$ with $\lambda_1 \neq \lambda_2$.

Case 2. If $\lambda_1 + \lambda_3 = \lambda_2 + \lambda_4$, then $\lambda_4 = \lambda_1 + \lambda_3 - \lambda_2$, and hence the matrix P introduced in (3.4) takes the form

$$P = \begin{bmatrix} \lambda_1 & \lambda_2 & \lambda_3 & \lambda_1 + \lambda_3 - \lambda_2 \\ \lambda_2 & \lambda_3 & \lambda_1 + \lambda_3 - \lambda_2 & \lambda_1 \\ \lambda_3 & \lambda_1 + \lambda_3 - \lambda_2 & \lambda_1 & \lambda_2 \\ \lambda_1 + \lambda_3 - \lambda_2 & \lambda_1 & \lambda_2 & \lambda_3 \end{bmatrix}.$$

We will verify that $\mathbf{e}_1 + \mathbf{e}_2 - \mathbf{e}_3 - \mathbf{e}_4$ is in the linear span of the column vectors of P . Since the fourth row (column) of P is nothing else but the sum of its first and third rows (columns) minus its second row (column) (and hence the rank of P is at most 3), it is enough to check that the system of linear equations

$$(3.9) \quad \begin{bmatrix} \lambda_1 & \lambda_2 & \lambda_3 \\ \lambda_2 & \lambda_3 & \lambda_1 + \lambda_3 - \lambda_2 \\ \lambda_3 & \lambda_1 + \lambda_3 - \lambda_2 & \lambda_1 \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \\ \gamma \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix}$$

can be solved for $\alpha, \beta, \gamma \in \mathbb{R}$. One can calculate that the determinant of the 3×3 (base) matrix of the system of linear equations (3.9) is

$$-(\lambda_1 + \lambda_3)((\lambda_1 - \lambda_2)^2 + (\lambda_2 - \lambda_3)^2),$$

which is 0 if and only if $\lambda_1 + \lambda_3 = 0$ or $\lambda_1 = \lambda_2 = \lambda_3$. If $\lambda_1 + \lambda_3$ were 0, then $\lambda_2 + \lambda_4$ would be 0 as well, and, using that $\lambda_i \in \mathbb{R}_+$, $i \in \{1, \dots, 4\}$, we would have that $\lambda_i = 0$, $i \in \{1, \dots, 4\}$, leading us to a contradiction (since $\lambda_1 + \lambda_2 + \lambda_3 + \lambda_4 = 1$). If $\lambda_1 = \lambda_2 = \lambda_3$ were true, then, since $\lambda_4 = \lambda_1 + \lambda_3 - \lambda_2$, we would have that $\lambda_4 = \lambda_1$, and, using that $\lambda_1 + \lambda_2 + \lambda_3 + \lambda_4 = 1$, it would imply that all the entries of P would be equal to $\frac{1}{4}$, but we excluded this case. All in all, the determinant of the (base) matrix of the system of linear equations (3.9) cannot be zero, and hence there exist a unique $\alpha, \beta, \gamma \in \mathbb{R}$ such that (3.9) holds. Consequently, the condition on P in part (i) of Theorem 2.1 is satisfied by choosing $A := \{1, 2\}$ and $B := \{3, 4\}$, and hence we can apply part (i) of Theorem 2.1 in case of $\lambda_1 + \lambda_3 = \lambda_2 + \lambda_4$ as well. $\square$

Declarations

Funding

The research of the second author was supported by the K-134191 NKFIH Grant.

Competing Interests

The authors have no relevant financial or non-financial interests to disclose.

Authors Contribution

Both authors contributed equally to the preparation of this manuscript.

Data availability

Not applicable.

References

- [1] M. Barczy and Zs. Páles. A convexity-type functional inequality with infinite convex combinations. *Ann. Univ. Sci. Budapest. Sect. Comput.*, 58:47–55, 2025.
- [2] P. Burai, J. Makó, and P. Szokol. Convexity generated by special circulant matrices. *Bul. Ştiinţ. Univ. Politeh. Timiş. Ser. Mat. Fiz.*, 64:47–53, 2019.
- [3] A. Carmona, A. M. Encinas, S. Gago, M. J. Jiménez, and M. Mitjana. The inverses of some circulant matrices. *Appl. Math. Comput.*, 270:785–793, 2015.
- [4] A. Gilányi and Zs. Páles. On Dinghas-type derivatives and convex functions of higher order. *Real Anal. Exchange*, 27(2):485–493, 2001/02.
- [5] M. Kuczma. *An Introduction to the Theory of Functional Equations and Inequalities*. Birkhäuser Verlag, Basel, 2nd edition, 2009. Cauchy’s equation and Jensen’s inequality, Edited and with a preface by Attila Gilányi.
- [6] A. W. Marshall, I. Olkin, and B. C. Arnold. *Inequalities: Theory of Majorization and its Applications*. Springer Series in Statistics. Springer, New York, second edition, 2011.
- [7] C. P. Niculescu and L.-E. Persson. *Convex Functions and their Applications—A Contemporary Approach*, volume 14 of *CMS/CAIMS Books in Mathematics*. Springer, Cham, third edition, 2025.
- [8] M. Niezgoda. Inequalities for convex functions and doubly stochastic matrices. *Math. Inequal. Appl.*, 16(1):221–232, 2013.
- [9] K. Nikodem and Zs. Páles. On t -convex functions. *Real Anal. Exchange*, 29(1):219–228, 2003/04.