Mathematical
Inequalities
& Applications
Volume 18, Number 3 (2015), 1013–1023 doi:10.7153/mia-18-77

A GENERALIZATION OF MERCER'S RESULT ON CONVEX FUNCTIONS, II

MAREK NIEZGODA

(*Communicated by I. Perić*)

Abstract. In this paper, we apply Jensen's inequality and two kinds of majorization for matrices in order to generalize some recent results on Mercer's inequality for convex functions.

1. Introduction and summary

A. McD. Mercer [9, Theorem 1.2] proved the following variant of Jensen's inequality.

THEOREM A. [9, Theorem 1.2] *Let f be a real convex function on an interval $[a_1,a_2]$, $a_1 < a_2$, such that*

$$a_1 \leqslant x_i \leqslant a_2 \quad \text{for } i = 1,\dots,n. \tag{1}$$

Then

$$f\left(a_1 + a_2 - \sum_{i=1}^{n} w_i x_i\right) \leqslant f(a_1) + f(a_2) - \sum_{i=1}^{n} w_i f(x_i), \tag{2}$$

where $\sum\limits_{i=1}^{n} w_i = 1$ *with* $w_i > 0$.

An m-tuple $\mathbf{b} = (b_1,\dots,b_m) \in \mathbb{R}^m$ is said to be *majorized* by m-tuple $\mathbf{a} = (a_1,\dots,a_m) \in \mathbb{R}^m$, written as $\mathbf{b} \prec \mathbf{a}$, if

$$\sum_{j=1}^{k} b_{[j]} \leqslant \sum_{j=1}^{k} a_{[j]} \quad \text{for } k = 1,\dots,m, \text{ and } \quad \sum_{j=1}^{m} b_j = \sum_{j=1}^{m} a_j, \tag{3}$$

where $a_{[1]} \geqslant \cdots \geqslant a_{[m]}$ and $b_{[1]} \geqslant \cdots \geqslant b_{[m]}$ are the entries of $\mathbf{a}$ and $\mathbf{b}$, respectively, in nonincreasing order [6, p. 8].

A majorization generalization of Mercer's inequality (2) has been shown in [10, Theorem 2.1].

Mathematics subject classification (2010): 26A51, 26D15, 06F20.
Keywords and phrases: Convex function, Jensen's inequality, Mercer's inequality, majorization, doubly stochastic matrix, column stochastic matrix.

THEOREM B. [10, Theorem 2.1] *Let $f : I \to \mathbb{R}$ be a continuous convex function on an interval $I \subset \mathbb{R}$. Suppose $\mathbf{a} = (a_1,\ldots,a_m)$ with $a_j \in I$, and $\mathbf{X} = (x_{ij})$ is a real $n \times m$ matrix such that $x_{ij} \in I$ for all i, j.*

If $\mathbf{a}$ majorizes each row of $\mathbf{X}$, that is

$$\mathbf{x}_{i\cdot} = (x_{i1},\ldots,x_{im}) \prec (a_1,\ldots,a_m) = \mathbf{a} \quad \text{for each } i = 1,\ldots,n, \tag{4}$$

then we have the inequality

$$f\left(\sum_{j=1}^{m} a_j - \sum_{j=1}^{m-1}\sum_{i=1}^{n} w_i x_{ij}\right) \leqslant \sum_{j=1}^{m} f(a_j) - \sum_{j=1}^{m-1}\sum_{i=1}^{n} w_i f(x_{ij}), \tag{5}$$

where $\sum\limits_{i=1}^{n} w_i = 1$ with $w_i \geqslant 0$.

Pavić [13] has recently proved the following result (with $\gamma = -1$).

THEOREM C. [13, Theorem 2.5] *Let $a_j, x_i \in \mathbb{R}$ be points so that*

$$\{a_1,\ldots,a_m\} \cap \operatorname{conv}\{x_1,\ldots,x_n\} = \emptyset \quad \text{or } \{endpoint(s)\}.$$

Let $\alpha_j, w_j \in [0,1]$ be coefficients of sums $\sum\limits_{j=1}^{m} \alpha_j - 1 = \sum\limits_{j=1}^{n} w_j = 1$.

If

$$\frac{1}{\sum\limits_{j=1}^{m} \alpha_j} \sum_{j=1}^{m} \alpha_j a_j \in \operatorname{conv}\{x_1,\ldots,x_n\},$$

then every convex function $f : \operatorname{conv}\{a_1,\ldots,a_n\} \to \mathbb{R}$ verifies the inequality

$$f\left(\sum_{j=1}^{m} \alpha_j a_j - \sum_{i=1}^{n} w_i x_i\right) \leqslant \sum_{j=1}^{m} \alpha_j f(a_j) - \sum_{i=1}^{n} w_i f(x_i). \tag{6}$$

Some generalizations, extensions and refinements of Mercer's inequality and Theorem A can be found in [2, 4, 7, 8, 12].

In the present paper our aim is to demonstrate further results related to Theorems A, B and C by using a majorization method.

Here is the summary of the paper. Our approach is based on two types of majorization for matrices. In Section 2 we utilize some majorization relations for rows of matrices. We prove Theorem 2.1, which extends Theorem B. In addition, Corollary 2.2 is a generalization of Theorem A. In Section 3 we pay attention on consequences of majorization relations for columns of matrices. In Theorem 3.1 and Corollary 3.2 we provide inequalities with row sums of some involved column stochastic matrix inducing the column majorization. This leads to results similar as in Theorem C and in [4, Theorem 5.1].

2. Majorization and Mercer type inequalities

An $m\times m$ real matrix $\mathbf{S}=(s_{ij})$ is said to be *doubly stochastic* if $s_{ij}\geqslant 0$ for $i,j=1,2,\dots,m$, and all row and column sums of $\mathbf{S}$ are equal to 1, i.e., $\sum_{j=1}^{m} s_{ij}=1$ for $i=1,2,\dots,m$, and $\sum_{i=1}^{m} s_{ij}=1$ for $j=1,2,\dots,m$.

It is worth noting that an $m\times m$ matrix $\mathbf{S}$ is doubly stochastic iff $\mathbf{S}\geqslant 0$ (entrywise) and $\mathbf{eS}=\mathbf{e}=\mathbf{eS}^T$, where $\mathbf{e}=(1,\dots,1)$ is the vector of ones of dimension m.

For given two real row m-tuples $\mathbf{a}=(a_1,\dots,a_m)$ and $\mathbf{b}=(b_1,\dots,b_m)$ it is well-known that

$$\mathbf{b}\prec\mathbf{a}\quad\text{iff}\quad \mathbf{b}=\mathbf{aS} \tag{7}$$

for some doubly stochastic $m\times m$ matrix $\mathbf{S}=(s_{ij})$ (see [6, Theorem B.2., p. 33]).

Majorization and convexity are related as follows. If $f:I\to\mathbb{R}$ is a continuous convex function on an interval $I\subset\mathbb{R}$ then for $\mathbf{a}=(a_1,\dots,a_m)$ and $\mathbf{b}=(b_1,\dots,b_m)$ with $a_j,b_j\in I$ for $j=1,\dots,m$,

$$\mathbf{b}\prec\mathbf{a}\quad\text{implies}\quad \sum_{j=1}^{m} f(b_j)\leqslant\sum_{j=1}^{m} f(a_j) \tag{8}$$

(see [5, p. 75], [6, p. 156]).

The ith row and jth column of a matrix $\mathbf{Z}=(z_{ij})$ are denoted by $\mathbf{z}_{i\cdot}$ and $\mathbf{z}_{\cdot j}$, respectively.

Given a function $f:I\subset\mathbb{R}\to\mathbb{R}$ and an $n\times 1$ column matrix $\mathbf{z}=(z_1,\dots,z_n)^T$ such that $z_i\in I$ for $i=1,\dots,n$, we define $f(\mathbf{z})$ as $n\times 1$ column matrix $(f(z_1),\dots,f(z_n))^T$.

In addition, we equipp the space $\mathbb{R}^n$ with the standard inner product $\langle\cdot,\cdot\rangle$ given by

$$\langle\mathbf{x},\mathbf{y}\rangle=\sum_{i=1}^{n} x_iy_i\quad\text{for } \mathbf{x}=(x_1,\dots,x_n)^T,\mathbf{y}=(y_1,\dots,y_n)^T\in\mathbb{R}^n.$$

THEOREM 2.1. *Let $f:I\to\mathbb{R}$ be a continuous convex function on an interval $I\subset\mathbb{R}$. Suppose $\mathbf{X}=(x_{ij})$ and $\mathbf{A}=(a_{ij})$ are real $n\times m$ matrices such that $x_{ij}\in I$ and $a_{ij}\in I$ for $i=1,\dots,n,\ j=1,\dots,m$.*

If for each $i=1,\dots,n$, the ith row of $\mathbf{X}$ is majorized by the ith row of $\mathbf{A}$, that is

$$\mathbf{x}_{i\cdot}=(x_{i1},\dots,x_{im})\prec(a_{i1},\dots,a_{im})=\mathbf{a}_{i\cdot}\quad\textit{for each } i=1,\dots,n, \tag{9}$$

then we have the inequality

$$f\left(\sum_{j=1}^{m}\sum_{i=1}^{n} w_ia_{ij}-\sum_{j=1}^{m-1}\sum_{i=1}^{n} w_ix_{ij}\right)\leqslant\sum_{j=1}^{m}\sum_{i=1}^{n} w_if(a_{ij})-\sum_{j=1}^{m-1}\sum_{i=1}^{n} w_if(x_{ij}), \tag{10}$$

where $\sum_{i=1}^{n} w_i=1$ with $w_i\geqslant 0$.

Proof. We begin by showing that the left-hand side of inequality (10) is well-defined.

By a routine algebra, we have

$$\sum_{j=1}^{m}\sum_{i=1}^{n} w_i a_{ij} - \sum_{j=1}^{m-1}\sum_{i=1}^{n} w_i x_{ij} = \sum_{j=1}^{m}\langle \mathbf{w}, \mathbf{a}_{\cdot j}\rangle - \sum_{j=1}^{m-1}\langle \mathbf{w}, \mathbf{x}_{\cdot j}\rangle$$
$$= \langle \mathbf{w}, \sum_{j=1}^{m}\mathbf{a}_{\cdot j} - \sum_{j=1}^{m-1}\mathbf{x}_{\cdot j}\rangle. \tag{11}$$

In light of (3) and (9) we see that

$$\sum_{j=1}^{m} a_{ij} = \sum_{j=1}^{m} x_{ij} \quad \text{for each } i = 1,\ldots,n,$$

and therefore

$$\sum_{j=1}^{m} a_{ij} - \sum_{j=1}^{m-1} x_{ij} = x_{im} \quad \text{for each } i = 1,\ldots,n,$$

which means

$$\sum_{j=1}^{m}\mathbf{a}_{\cdot j} - \sum_{j=1}^{m-1}\mathbf{x}_{\cdot j} = \mathbf{x}_{\cdot m}. \tag{12}$$

Because $\mathbf{x}_{\cdot m} \in I^n$ and $\mathbf{w}$ is a probability vector in $\mathbb{R}^n$, we infer that $\langle \mathbf{w}, \mathbf{x}_{\cdot m}\rangle \in I$. Hence, by (12), we get $\sum\limits_{j=1}^{m}\sum\limits_{i=1}^{n} w_i a_{ij} - \sum\limits_{j=1}^{m-1}\sum\limits_{i=1}^{n} w_i x_{ij} \in I$, as was to be proven.

From (11) and by the Jensen's inequality for convex function f, we can write

$$f\left(\sum_{j=1}^{m}\sum_{i=1}^{n} w_i a_{ij} - \sum_{j=1}^{m-1}\sum_{i=1}^{n} w_i x_{ij}\right)$$
$$= f\left(\langle \mathbf{w}, \sum_{j=1}^{m}\mathbf{a}_{\cdot j} - \sum_{j=1}^{m-1}\mathbf{x}_{\cdot j}\rangle\right) \leqslant \langle \mathbf{w}, f\left(\sum_{j=1}^{m}\mathbf{a}_{\cdot j} - \sum_{j=1}^{m-1}\mathbf{x}_{\cdot j}\right)\rangle. \tag{13}$$

It now follows from (12) and (13) that

$$f\left(\sum_{j=1}^{m}\sum_{i=1}^{n} w_i a_{ij} - \sum_{j=1}^{m-1}\sum_{i=1}^{n} w_i x_{ij}\right) \leqslant \langle \mathbf{w}, f(\mathbf{x}_{\cdot m})\rangle. \tag{14}$$

On account of (8) and (9) we obtain

$$f(x_{i1}) + \ldots + f(x_{im}) \leqslant f(a_{i1}) + \ldots + f(a_{im}) \quad \text{for } i = 1,\ldots,n.$$

From this we deduce that

$$f(\mathbf{x}_{\cdot 1}) + \ldots + f(\mathbf{x}_{\cdot m}) \leqslant f(\mathbf{a}_{\cdot 1}) + \ldots + f(\mathbf{a}_{\cdot m}),$$

where here $\leqslant$ is the componentwise ordering on $\mathbb{R}^n$ induced by the cone $\mathbb{R}^n_+$. Therefore

$$f(\mathbf{x}_{\cdot m}) \leqslant f(\mathbf{a}_{\cdot 1}) + \ldots + f(\mathbf{a}_{\cdot m}) - f(\mathbf{x}_{\cdot 1}) - \ldots - f(\mathbf{x}_{\cdot m-1}).$$

In addition, $\mathbf{w} \in \mathbb{R}^n_+$, so we get

$$\langle \mathbf{w}, f(\mathbf{x}_{\cdot m}) \rangle \leqslant \langle \mathbf{w}, \sum_{j=1}^{m} f(\mathbf{a}_{\cdot j}) - \sum_{j=1}^{m-1} f(\mathbf{x}_{\cdot j}) \rangle,$$

which implies

$$\begin{aligned} \langle \mathbf{w}, f(\mathbf{x}_{\cdot m}) \rangle &\leqslant \sum_{j=1}^{m} \langle \mathbf{w}, f(\mathbf{a}_{\cdot j}) \rangle - \sum_{j=1}^{m-1} \langle \mathbf{w}, f(\mathbf{x}_{\cdot j}) \rangle \\ &= \sum_{j=1}^{m} \sum_{i=1}^{n} w_i f(a_{ij}) - \sum_{j=1}^{m-1} \sum_{i=1}^{n} w_i f(x_{ij}). \end{aligned} \tag{15}$$

Finally, by combining (14) and (15), we establish (10) completing the proof. □

In order to establish an extension of Mercer's inequality (see Theorem A), we employ Theorem 2.1 for $m = 2$. Here we relax the condition (1) from $a \leqslant x_i \leqslant b$ to $a_i \leqslant x_i \leqslant b_i$ for $i = 1, \ldots, n$.

COROLLARY 2.2. *Let $f : I \to \mathbb{R}$ be a continuous convex function on an interval $I \subset \mathbb{R}$. Let $a_i, b_i, x_i \in I$ satisfy $a_i \leqslant x_i \leqslant b_i$ for $i = 1, \ldots, n$.*

Then we have the inequality

$$f\left(\sum_{i=1}^{n} w_i a_i + \sum_{i=1}^{n} w_i b_i - \sum_{i=1}^{n} w_i x_i \right) \leqslant \sum_{i=1}^{n} w_i f(a_i) + \sum_{i=1}^{n} w_i f(b_i) - \sum_{i=1}^{n} w_i f(x_i),$$

where $\sum_{i=1}^{n} w_i = 1$ with $w_i \geqslant 0$.

Proof. We introduce $n \times 2$ matrices $\mathbf{A} = (a_{ij})$ and $\mathbf{X} = (x_{ij})$ by putting $a_{i1} = a_i$, $a_{i2} = b_i$, $x_{i1} = x_i$, $x_{i2} = y_i = a_i + b_i - x_i$ for $i = 1, \ldots, n$.

It is not hard to check that the ith row of $\mathbf{X}$ is majorized by the ith row of $\mathbf{A}$, i.e.,

$$\mathbf{x}_{i\cdot} = (x_{i1}, x_{i2}) \prec (a_{i1}, a_{i2}) = \mathbf{a}_{i\cdot} \quad \text{for each } i = 1, \ldots, n.$$

Now, the required result is due to Theorem 2.1. □

REMARK 2.3. In Theorem 2.1, if all rows of the matrix $\mathbf{A}$ are identical, i.e.,

$$\mathbf{a}_{i\cdot} = (a_{i1}, \ldots, a_{im}) = (a_1, \ldots, a_m) = \mathbf{a} \quad \text{for each } i = 1, \ldots, n$$

(cf. (4)), then inequality (10) becomes inequality (5).

In fact, since $a_{ij} = a_j$ for $i = 1, \ldots, n$, from Theorem 2.1 we get

$$f\left(\sum_{j=1}^{m} a_j - \sum_{j=1}^{m-1}\sum_{i=1}^{n} w_i x_{ij}\right) = f\left(\sum_{j=1}^{m}\sum_{i=1}^{n} w_i a_{ij} - \sum_{j=1}^{m-1}\sum_{i=1}^{n} w_i x_{ij}\right)$$

$$\leqslant \sum_{j=1}^{m}\sum_{i=1}^{n} w_i f(a_{ij}) - \sum_{j=1}^{m-1}\sum_{i=1}^{n} w_i f(x_{ij}) = \sum_{j=1}^{m} f(a_j) - \sum_{j=1}^{m-1}\sum_{i=1}^{n} w_i f(x_{ij}),$$

which gives (5).

Therefore Theorem B is a special version of Theorem 2.1.

To give motivation for our further studies, notice that the majorization assumption (9) in Theorem 2.1, i.e.,

$$\mathbf{x}_{i\cdot} \prec \mathbf{a}_{i\cdot} \quad \text{for } i = 1, \ldots, n \tag{16}$$

is equivalent to

$$\mathbf{x}_{i\cdot} = \mathbf{a}_{i\cdot}\mathbf{S}_i \quad \text{for } i = 1, \ldots, n \tag{17}$$

with some doubly stochastic $m \times m$ matrices $\mathbf{S}_i$ (see (7)).

If in addition all matrices $\mathbf{S}_i$, $i = 1, \ldots, n$, are identical:

$$\mathbf{S}_1 = \ldots = \mathbf{S}_n = \mathbf{S} = (s_{ij}) \tag{18}$$

then (17) implies

$$\mathbf{x}_{i\cdot} = \mathbf{a}_{i\cdot}\mathbf{S} \quad \text{for } i = 1, \ldots, n,$$

which can be rewritten in the form

$$x_{ij} = \sum_{l=1}^{m} a_{il} s_{lj} \quad \text{for } i = 1, \ldots, n, \tag{19}$$

where $\mathbf{X} = (x_{ij}) = (\mathbf{x}_{\cdot 1}, \ldots, \mathbf{x}_{\cdot m})$ and $\mathbf{A} = (a_{ij}) = (\mathbf{a}_{\cdot 1}, \ldots, \mathbf{a}_{\cdot m})$, and $\mathbf{x}_{\cdot j}$ and $\mathbf{a}_{\cdot j}$ denote the jth columns of $\mathbf{X}$ and $\mathbf{A}$, respectively. Hence

$$\mathbf{x}_{\cdot j} = \begin{pmatrix} x_{1j} \\ \vdots \\ x_{nj} \end{pmatrix} = \begin{pmatrix} \sum\limits_{l=1}^{m} a_{1l} s_{lj} \\ \vdots \\ \sum\limits_{l=1}^{m} a_{nl} s_{lj} \end{pmatrix} = \sum_{l=1}^{m} \begin{pmatrix} a_{1l} \\ \vdots \\ a_{nl} \end{pmatrix} s_{lj} = \sum_{l=1}^{m} \mathbf{a}_{\cdot l} s_{lj}.$$

Since $\mathbf{S}$ is doubly stochastic, we obtain

$$\mathbf{x}_{\cdot j} \in \operatorname{conv}\{\mathbf{a}_{\cdot 1}, \ldots, \mathbf{a}_{\cdot m}\} \quad \text{for } j = 1, \ldots, m. \tag{20}$$

Thus we see that (20) is a consequence of (16) with assumption (18).

In the next section, we focus on the special majorization (20) in place of (16). This approach has the advantage that column stochastic matrices can be used instead of doubly stochastic ones.

3. Vector majorization and its applications

We now extend our framework by admitting m- and k-tuples with vector components in place of scalar ones and by making use of column stochastic matrices.

An $m \times k$ real matrix $\mathbf{S} = (s_{ij})$ is said to be *column stochastic* if $s_{ij} \geqslant 0$ for $i = 1,2,\ldots,m$, $j = 1,2,\ldots,k$, and all column sums of $\mathbf{S}$ are equal to 1, i.e., $\sum_{i=1}^{m} s_{ij} = 1$ for $j = 1,2,\ldots,k$.

It is readily seen that an $m \times k$ matrix $\mathbf{S}$ is column stochastic iff $\mathbf{S} \geqslant 0$ (entrywise) and $\mathbf{eS} = \mathbf{e}$, where $\mathbf{e} = (1,\ldots,1)$ is the vector of ones of an appropriate dimension.

We say that a vector k-tuple $\mathbf{Y} = (\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_k)$ with $\mathbf{y}_j \in \mathbb{R}^n$ for $j = 1,\ldots,k$, is *pre-majorized* (resp. *majorized*) by a vector m-tuple $\mathbf{X} = (\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_m)$ with $\mathbf{x}_j \in \mathbb{R}^n$ for $j = 1,\ldots,m$, written as $\mathbf{Y} \prec_p \mathbf{X}$ (resp. $\mathbf{Y} \prec \mathbf{X}$), if there exists an $m \times k$ column stochastic matrix (resp. doubly stochastic matrix) $\mathbf{S} = (s_{ij})$ so that

$$(\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_k) = (\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_m)\mathbf{S} \tag{21}$$

(see [11], cf. [3, 14]). Here and hereafter the notation (21) means

$$\mathbf{y}_j = s_{1j}\mathbf{x}_1 + s_{2j}\mathbf{x}_2 + \ldots + s_{mj}\mathbf{x}_m \quad \text{for } j = 1,2,\ldots,k. \tag{22}$$

Observe that (22) is equivalent to

$$\mathbf{y}_j \in \operatorname{conv}\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_m\} \quad \text{for } j = 1,2,\ldots,k$$

(cf. [14, Proposition 3.3]).

If $(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_m)$ is identified with $n \times m$ matrix $\mathbf{X}$, and $(\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_k)$ is identified with $n \times k$ matrix $\mathbf{Y}$, then (21) can be interpreted as

$$\mathbf{Y} = \mathbf{XS},$$

which is equivalent to

$$\mathbf{Y}^T = \mathbf{S}^T \mathbf{X}^T \tag{23}$$

with row stochastic $k \times m$ matrix $\mathbf{S}^T$.

In this situation, one sees that pre-majorization (21) means the *weak matrix majorization* (23) for matrices $\mathbf{X}^T$ and $\mathbf{Y}^T$ in terminology of [14] (see also [6, pp. 616–617]).

It is not hard to verify that if $f : I \to \mathbb{R}$ is convex on an interval $I \subset \mathbb{R}$ then the extended function f on $I^n \subset \mathbb{R}^n$ is $\leqslant$-convex on I^n (see (24)), where $\leqslant$ is the *componentwise ordering* on $\mathbb{R}^n$ given by

$$(x_1,\ldots,x_m)^T \leqslant (y_1,\ldots,y_m)^T \quad \text{iff} \quad x_i \leqslant y_i \quad \text{for } i = 1,\ldots,n.$$

That is, the convexity of f on I implies

$$f(\alpha(x_1,\ldots,x_m)^T + \beta(y_1,\ldots,y_m)^T) \leqslant \alpha f((x_1,\ldots,x_m)^T) + \beta f((y_1,\ldots,y_m)^T) \tag{24}$$

for $\alpha, \beta \geqslant 0$, $\alpha + \beta = 1$, $x_i, y_i \in I$.

THEOREM 3.1. *Let $f: I \to \mathbb{R}$ be a continuous convex function on an interval $I \subset \mathbb{R}$. Suppose $\mathbf{X} = (x_{ij})$ is a real $n \times k$ matrix such that $x_{ij} \in I$ for $i = 1, \ldots, n$, $j = 1, \ldots, k$, and $\mathbf{A} = (a_{ij})$ is a real $n \times m$ matrix such that $a_{ij} \in I$ for $i = 1, \ldots, n$, $j = 1, \ldots, m$.*

Assume the k-tuple $(\mathbf{x}_{\cdot 1}, \ldots, \mathbf{x}_{\cdot k})$ of columns of $\mathbf{X}$ is pre-majorized by the m-tuple $(\mathbf{a}_{\cdot 1}, \ldots, \mathbf{a}_{\cdot m})$ of columns of $\mathbf{A}$, i.e.,

$$(\mathbf{x}_{\cdot 1}, \ldots, \mathbf{x}_{\cdot k}) \prec_p (\mathbf{a}_{\cdot 1}, \ldots, \mathbf{a}_{\cdot m}) \tag{25}$$

in the sense

$$(\mathbf{x}_{\cdot 1}, \ldots, \mathbf{x}_{\cdot k}) = (\mathbf{a}_{\cdot 1}, \ldots, \mathbf{a}_{\cdot m})\mathbf{S} \tag{26}$$

for some column stochastic $m \times k$ matrix $\mathbf{S} = (s_{ij})$.

Then we have the inequality

$$f\left(\sum_{j=1}^{m}\sum_{i=1}^{n} S_j w_i a_{ij} - \sum_{j=1}^{k-1}\sum_{i=1}^{n} w_i x_{ij}\right) \leqslant \sum_{j=1}^{m}\sum_{i=1}^{n} S_j w_i f(a_{ij}) - \sum_{j=1}^{k-1}\sum_{i=1}^{n} w_i f(x_{ij}), \tag{27}$$

where $\sum\limits_{i=1}^{n} w_i = 1$ with $w_i \geqslant 0$, and $S_j = \sum\limits_{l=1}^{k} s_{jl}$ is the jth row sum of $\mathbf{S}$, $j = 1, \ldots, m$.

Proof. First of all, we show that the left-hand side of inequality (27) is well-defined. It follows that

$$\begin{aligned}\sum_{j=1}^{m}\sum_{i=1}^{n} S_j w_i a_{ij} - \sum_{j=1}^{k-1}\sum_{i=1}^{n} w_i x_{ij} &= \sum_{j=1}^{m} S_j \langle \mathbf{w}, \mathbf{a}_{\cdot j}\rangle - \sum_{j=1}^{k-1} \langle \mathbf{w}, \mathbf{x}_{\cdot j}\rangle \\ &= \langle \mathbf{w}, \sum_{j=1}^{m} S_j \mathbf{a}_{\cdot j} - \sum_{j=1}^{k-1} \mathbf{x}_{\cdot j}\rangle. \end{aligned} \tag{28}$$

It is sufficient to show that for each $i = 1, \ldots, n$,

$$\sum_{j=1}^{m} S_j a_{ij} - \sum_{j=1}^{k-1} x_{ij} \in I. \tag{29}$$

From (25) and (26) we get

$$\mathbf{x}_{\cdot j} = \sum_{l=1}^{m} s_{lj}\, \mathbf{a}_{\cdot l} \quad \text{for each } j = 1, \ldots, k. \tag{30}$$

Hence

$$\sum_{j=1}^{k} \mathbf{x}_{\cdot j} = \sum_{j=1}^{k}\sum_{l=1}^{m} s_{lj}\, \mathbf{a}_{\cdot l} = \sum_{l=1}^{m}\left(\sum_{j=1}^{k} s_{lj}\right)\mathbf{a}_{\cdot l} = \sum_{l=1}^{m} S_l \mathbf{a}_{\cdot l},$$

which yields

$$\mathbf{x}_{\cdot k} = \sum_{j=1}^{m} S_j \mathbf{a}_{\cdot j} - \sum_{j=1}^{k-1} \mathbf{x}_{\cdot j}. \tag{31}$$

But $x_{ij} \in I$ for all i, j, so (31) gives (29), as claimed.

On the other hand, from (28) and by the convexity of f on I (see (24)), we get

$$f\left(\sum_{j=1}^{m}\sum_{i=1}^{n} S_j w_i a_{ij} - \sum_{j=1}^{k-1}\sum_{i=1}^{n} w_i x_{ij}\right) = f\left(\langle \mathbf{w}, \sum_{j=1}^{m} S_j \mathbf{a}_{\cdot j} - \sum_{j=1}^{k-1} \mathbf{x}_{\cdot j}\rangle\right) \leqslant \langle \mathbf{w}, f\left(\sum_{j=1}^{m} S_j \mathbf{a}_{\cdot j} - \sum_{j=1}^{k-1} \mathbf{x}_{\cdot j}\right)\rangle. \quad (32)$$

According to (31) we have

$$f\left(\sum_{j=1}^{m} S_j \mathbf{a}_{\cdot j} - \sum_{j=1}^{k-1} \mathbf{x}_{\cdot j}\right) = f(\mathbf{x}_{\cdot k})$$

and further

$$\langle \mathbf{w}, f\left(\sum_{j=1}^{m} S_j \mathbf{a}_{\cdot j} - \sum_{j=1}^{k-1} \mathbf{x}_{\cdot j}\right)\rangle = \langle \mathbf{w}, f(\mathbf{x}_{\cdot k})\rangle. \quad (33)$$

Now, from (32) and (33) we find that

$$f\left(\sum_{j=1}^{m}\sum_{i=1}^{n} S_j w_i a_{ij} - \sum_{j=1}^{k-1}\sum_{i=1}^{n} w_i x_{ij}\right) \leqslant \langle \mathbf{w}, f(\mathbf{x}_{\cdot k})\rangle. \quad (34)$$

We shall prove that

$$f(\mathbf{x}_{\cdot 1}) + \ldots + f(\mathbf{x}_{\cdot k}) \leqslant S_1 f(\mathbf{a}_{\cdot 1}) + \ldots + S_m f(\mathbf{a}_{\cdot m}), \quad (35)$$

where here $\leqslant$ is the componentwise ordering on $\mathbb{R}^n$ induced by the cone $\mathbb{R}^n_+$ (cf. [14, Theorem 3.9]).

In fact, from (30) and by the column stochasticity of $\mathbf{S}$ and the convexity of f (w.r.t. the componentwise $\leqslant$ ordering on $\mathbb{R}^n$ (see (24))), we obtain

$$\sum_{j=1}^{k} f(\mathbf{x}_{\cdot j}) = \sum_{j=1}^{k} f\left(\sum_{l=1}^{m} s_{lj}\mathbf{a}_{\cdot l}\right) \leqslant \sum_{j=1}^{k}\sum_{l=1}^{m} s_{lj} f(\mathbf{a}_{\cdot l})$$
$$= \sum_{l=1}^{m}\left(\sum_{j=1}^{k} s_{lj}\right) f(\mathbf{a}_{\cdot l}) = \sum_{l=1}^{m} S_l f(\mathbf{a}_{\cdot l}).$$

Thus (35) is proven.

In consequence,

$$f(\mathbf{x}_{\cdot k}) \leqslant \sum_{j=1}^{m} S_j f(\mathbf{a}_{\cdot j}) - \sum_{j=1}^{k-1} f(\mathbf{x}_{\cdot j}).$$

In addition, $\mathbf{w} \in \mathbb{R}^n_+$, so we get

$$\langle \mathbf{w}, f(\mathbf{x}_{\cdot k})\rangle \leqslant \langle \mathbf{w}, \sum_{j=1}^{m} S_j f(\mathbf{a}_{\cdot j}) - \sum_{j=1}^{k-1} f(\mathbf{x}_{\cdot j})\rangle,$$

which forces

$$\langle \mathbf{w}, f(\mathbf{x}_{\cdot k})\rangle \leqslant \sum_{j=1}^{m} S_j \langle \mathbf{w}, f(\mathbf{a}_{\cdot j})\rangle - \sum_{j=1}^{k-1} \langle \mathbf{w}, f(\mathbf{x}_{\cdot j})\rangle$$
$$= \sum_{j=1}^{m}\sum_{i=1}^{n} S_j w_i f(a_{ij}) - \sum_{j=1}^{k-1}\sum_{i=1}^{n} w_i f(x_{ij}). \tag{36}$$

It is now sufficient to combine (34) and (36) in order to see that (27) is valid. This completes the proof. □

COROLLARY 3.2. *Under the assumptions of Theorem 3.1, if all rows of the matrix* **A** *are identical, i.e.,*

$$\mathbf{a}_{i\cdot} = (a_{i1}, \ldots, a_{im}) = (a_1, \ldots, a_m) \quad \textit{for each } i = 1, \ldots, n, \tag{37}$$

then we have the inequality

$$f\left(\sum_{j=1}^{m} S_j a_j - \sum_{j=1}^{k-1}\sum_{i=1}^{n} w_i x_{ij}\right) \leqslant \sum_{j=1}^{m} S_j f(a_j) - \sum_{j=1}^{k-1}\sum_{i=1}^{n} w_i f(x_{ij}), \tag{38}$$

where $\sum_{i=1}^{n} w_i = 1$ *with* $w_i \geqslant 0$*, and* $S_j = \sum_{l=1}^{k} s_{jl}$ *is the* j*th row sum of* $\mathbf{S}$*,* $j = 1, \ldots, m$*.*

Proof. By (37) one has $a_{ij} = a_j$ for $i = 1, \ldots, n$. Therefore Theorem 3.1 gives

$$f\left(\sum_{j=1}^{m} S_j a_j - \sum_{j=1}^{k-1}\sum_{i=1}^{n} w_i x_{ij}\right) = f\left(\sum_{j=1}^{m}\sum_{i=1}^{n} S_j w_i a_{ij} - \sum_{j=1}^{k-1}\sum_{i=1}^{n} w_i x_{ij}\right)$$
$$\leqslant \sum_{j=1}^{m}\sum_{i=1}^{n} S_j w_i f(a_{ij}) - \sum_{j=1}^{k-1}\sum_{i=1}^{n} w_i f(x_{ij}) = \sum_{j=1}^{m} S_j f(a_j) - \sum_{j=1}^{k-1}\sum_{i=1}^{n} w_i f(x_{ij}).$$

This implies (38), as desired. □

The case $k = 2$ of Corollary 3.2 leads to a result similar to that of Pavić [13] with $\gamma = -1$ (cf. Theorem C). See also [12, Theorem 16, (34)].

We conclude with the observation that the variant of Theorem 3.1 with doubly stochastic matrix **S** is a special case of Theorem 2.1.

REFERENCES

[1] W. W. BRECKNER AND T. TRIF, *Convex Functions and Related Functional Equations: Selected Topics*, Cluj University Press, Cluj, 2008.

[2] W. S. CHEUNG, A. MATKOVIĆ AND J. PEČARIĆ, *A variant of Jensen's inequality and generalized means*, J. Ineq. Pure & Appl. Math. **7**, (1) (2006), Article 10.

[3] G. DAHL, *Matrix majorization*, Linear Algebra Appl., **288**, (1999), 53–73.

[4] A. GUESSAB, *Direct and converse results for generalized multivariate Jensen-type inequalities*, J. Nonlinear Convex Anal. **13**, (4) (2012), 777–797.
[5] G. M. HARDY, J. E. LITTLEWOOD AND G. PÓLYA, *Inequalities*, 2nd Ed., Cambridge University Press, Cambridge, 1952.
[6] A. W. MARSHALL, I. OLKIN AND B. C. ARNOLD, *Inequalities: Theory of Majorization and Its Applications*, Second Edition, Springer, New York, 2011.
[7] A. MATKOVIĆ AND J. PEČARIĆ, *A variant of Jensen's inequality for convex functions of several variables*, J. Math. Ineq. **1**, (1) (2007), 45–51.
[8] A. MATKOVIĆ, J. PEČARIĆ AND I. PERIĆ, *Refinements of Jensen's inequality of Mercer's type for operator convex functions*, Math. Ineq. Appl. **11**, (1) (2008), 113–126.
[9] A. MCD. MERCER, *A variant of Jensen's inequality*, J. Ineq. Pure & Appl. Math. **4**, (4) (2003), Article 73.
[10] M. NIEZGODA, *A generalization of Mercer's result on convex functions*, Nonlinear Anal.: Theory Methods Appl., **71**, (2009), 2771–2779.
[11] M. NIEZGODA, *Vector majorization and Schur-concavity of some sums generated by the Jensen and Jensen-Mercer functionals*, Math. Inequal. Appl., **18**, 2 (2015), 769–786.
[12] Z. PAVIĆ, *The applications of functional variants of Jensen's inequality*, J. Function Spaces Appl., Article ID 194830 (2013), 5 pages.
[13] Z. PAVIĆ, *Generalizations of Jensen-Mercer's inequality*, J. Pure Appl. Math. Adv. Appl., **11**, (1) (2014), 19–36.
[14] F. D. PERIA, P. G. MASSEY AND L. E. SILVESTRE, *Weak matrix majorization*, Linear Algebra Appl. **403**, (2005), 343–368.

(Received August 2, 2014)

Marek Niezgoda
Department of Applied Mathematics and Computer Science
University of Life Sciences in Lublin
Akademicka 15, 20-950 Lublin, Poland
e-mail: marek.niezgoda@up.lublin.pl, bniezgoda@wp.pl