Adv. Nonlinear Anal. 2020; 9: 168–175 **DE GRUYTER**

Research Article

Marek Niezgoda*

Nonlinear Sherman-type inequalities

https://doi.org/10.1515/anona-2018-0098
Received April 23, 2018; revised June 28, 2018; accepted July 9, 2018

Abstract: An important class of Schur-convex functions is generated by convex functions via the well-known Hardy–Littlewood–Pólya–Karamata inequality. Sherman's inequality is a natural generalization of the HLPK inequality. It can be viewed as a comparison of two special inner product expressions induced by a convex function of one variable. In the present note, we extend the Sherman inequality from the (bilinear) inner product to a (nonlinear) map of two vectorial variables satisfying the Leon–Proschan condition. Some applications are shown for directional derivatives and gradients of Schur-convex functions.

Keywords: Majorization, convex function, convex-concave map, HLPK inequality, Sherman inequality, directional derivative, gradient

MSC 2010: 26B25, 26D15, 26D10

1 Introduction

We say that an n-tuple $\mathbf{y} = (y_1, \dots, y_n)^T \in \mathbb{R}^n$ is *majorized* by an n-tuple $\mathbf{x} = (x_1, \dots, x_n)^T \in \mathbb{R}^n$, and write $\mathbf{y} \prec \mathbf{x}$, if

$$\sum_{i=1}^{l} y_{[i]} \le \sum_{i=1}^{l} x_{[i]} \quad \text{for } l = 1, \dots, n, \qquad \text{and} \qquad \sum_{i=1}^{n} y_i = \sum_{i=1}^{n} x_i.$$

Here $x_{[1]} \ge \dots \ge x_{[n]}$ and $y_{[1]} \ge \dots \ge y_{[n]}$ are the entries of $\mathbf{x}$ and $\mathbf{y}$, respectively, arranged in decreasing order [13, p. 8].

It is known that for $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$,

$$\mathbf{y} \prec \mathbf{x} \quad \text{if and only if} \quad \mathbf{y} \in \operatorname{conv} \mathbb{P}_n \mathbf{x}, \tag{1.1}$$

where the symbol conv means "the convex hull of", and $\mathbb{P}_n$ denotes the group of $n \times n$ permutation matrices (see [5, p. 16], [6, p. 12] and [7]).

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

An $n \times n$ real matrix $\mathbf{S} = (s_{ij})$ is said to be *doubly stochastic* if it is column stochastic and row stochastic [13, pp. 29–30]. The set of all $n \times n$ doubly stochastic matrices is denoted by $\mathbb{D}_n$.

A doubly stochastic matrix is a convex combination of some permutation matrices, and vice versa [13, Theorem A.2.]. That is, $\mathbb{D}_n = \operatorname{conv} \mathbb{P}_n$. Therefore, (1.1) takes the following form: for $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$,

$$\mathbf{y} \prec \mathbf{x} \quad \text{if and only if} \quad \mathbf{y} = \mathbf{S}\mathbf{x}$$

for some doubly stochastic $n \times n$ matrix $\mathbf{S}$ (see [13, p. 33]).

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

Open Access. © 2019 Walter de Gruyter GmbH, Berlin/Boston.
This work is licensed under the Creative Commons Attribution 4.0 Public License.

A function $F : J^n \to \mathbb{R}$ with an interval $J \subset \mathbb{R}$ is said to be *Schur-convex* on J^n if for $\mathbf{x}, \mathbf{y} \in J^n$,

$$\mathbf{y} \prec \mathbf{x} \quad \text{implies} \quad F(\mathbf{y}) \le F(\mathbf{x}).$$

See [13, pp. 79–154] for applications of Schur-convex functions.

Some important examples of Schur-convex functions are included in the following theorem.

Theorem A ([8, 11]). *Let $f : J \to \mathbb{R}$ be a real convex function defined on an interval $J \subset \mathbb{R}$. Then, for $\mathbf{x} = (x_1, x_2, \ldots, x_n)^T \in J^n$ and $\mathbf{y} = (y_1, y_2, \ldots, y_n)^T \in J^n$,*

$$\mathbf{y} \prec \mathbf{x} \quad \textit{implies} \quad \sum_{i=1}^n f(y_i) \le \sum_{i=1}^n f(x_i).$$

Throughout, the symbol $(\cdot)^T$ denotes the operation of taking the transpose of a matrix. So, $\mathbf{S}$ is a column stochastic matrix if and only if $\mathbf{S}^T$ is a row stochastic matrix.

A generalization of Theorem A is the following result (see [17], cf. also [4]).

Theorem B ([17]). *Let f be a real convex function defined on an interval $J \subset \mathbb{R}$. Let $\mathbf{a} = (a_1, \ldots, a_n)^T \in \mathbb{R}_+^n$, $\mathbf{b} = (b_1, \ldots, b_m)^T \in \mathbb{R}_+^m$, $\mathbf{x} = (x_1, \ldots, x_n)^T \in J^n$ and $\mathbf{y} = (y_1, \ldots, y_m)^T \in J^m$.*

If

$$\mathbf{y} = \mathbf{S}\mathbf{x} \quad \textit{and} \quad \mathbf{a} = \mathbf{S}^T\mathbf{b} \tag{1.2}$$

for some $n \times m$ row stochastic matrix $\mathbf{S} = (s_{ij})$, then

$$\sum_{j=1}^m b_j f(y_j) \le \sum_{i=1}^n a_i f(x_i). \tag{1.3}$$

If f is concave, then inequality (1.3) *is reversed.*

Statements (1.2) and (1.3) are referred to as *Sherman's condition* and *Sherman's inequality*, respectively. Consult [1–4, 9, 10, 14–16] for generalizations and applications of Theorem B.

Observe that, when $m = n$, inequality (1.3) can be rewritten as

$$\langle \mathbf{b}, f(\mathbf{y})\rangle \le \langle \mathbf{a}, f(\mathbf{x})\rangle, \tag{1.4}$$

where $\langle \cdot, \cdot \rangle$ is the standard inner product on $\mathbb{R}^n$, and $f(\mathbf{x}) = (f(x_1), \ldots, f(x_n))^T$ and $f(\mathbf{y}) = (f(y_1), \ldots, f(y_n))^T$. Further, (1.4) can be restated as

$$\Psi(\mathbf{b}, f(\mathbf{y})) \le \Psi(\mathbf{a}, f(\mathbf{x})), \tag{1.5}$$

where Ψ is the inner product map on $\mathbb{R}^n$, i.e.,

$$\Psi(\mathbf{c}, \mathbf{z}) = \langle \mathbf{c}, \mathbf{z}\rangle \quad \text{for } \mathbf{c}, \mathbf{z} \in \mathbb{R}^n.$$

In the next section, we study inequalities of the form (1.5) for an arbitrary (nonlinear) map Ψ of two variables in $\mathbb{R}^n$.

2 Sherman-type inequality for nonlinear maps

In [12], Leon and Proschan gave some interesting inequalities for the Hadamard product map $\Psi(\mathbf{x}, \mathbf{y}) = \mathbf{x} \circ \mathbf{y}$, where $\mathbf{x} \circ \mathbf{y} = (x_1y_1, \ldots, x_ny_n)^T$ for $\mathbf{x} = (x_1, \ldots, x_n)^T \in \mathbb{R}^n$ and $\mathbf{y} = (y_1, \ldots, y_n)^T \in \mathbb{R}^n$. They applied a finite reflection group G acting on $\mathbb{R}^n$ with the property that for each $g \in G$ there exist $h, k \in G$ such that

$$\mathbf{x} \circ g\mathbf{y} = k((h\mathbf{x}) \circ \mathbf{y}) \quad \text{for } \mathbf{x}, \mathbf{y} \in \mathbb{R}^n. \tag{2.1}$$

Example 2.1. Let $\Psi : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$ be defined by $\Psi(\mathbf{x}, \mathbf{y}) = \mathbf{x} \circ \mathbf{y}$, the Hadamarad product on $\mathbb{R}^n$.

If $G = \mathbb{P}_n$ (the permutation group acting on $\mathbb{R}^n$), then

$$\mathbf{x} \circ p\mathbf{y} = p(p^{-1}\mathbf{x} \circ \mathbf{y}) \quad \text{for } \mathbf{x}, \mathbf{y} \in \mathbb{R}^n \text{ and } p \in \mathbb{P}_n.$$

So, (2.1) is met with $g = p$, $h = p^{-1} = p^T$ and $k = p$.

If $G = \mathbb{C}_n$ (the sign changes group acting on $\mathbb{R}^n$), then

$$\mathbf{x} \circ c\mathbf{y} = (c\mathbf{x}) \circ \mathbf{y} \quad \text{for } \mathbf{x}, \mathbf{y} \in \mathbb{R}^n \text{ and } c \in \mathbb{C}_n.$$

Therefore, (2.1) holds with $g = c$, $h = c$ and $k = \mathrm{id}$.

In what follows, we adapt this idea for any map Ψ of two vectorial variables and for the group $G = \mathbb{P}_n$ of $n \times n$ permutation matrices acting on $\mathbb{R}^n$.

We say that a map $\Psi : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^l$ admits the *Leon–Proschan property* for the permutation group $\mathbb{P}_n$ if for each $g \in \mathbb{P}_n$ there exist $h \in \mathbb{P}_n$ and $k \in \mathbb{P}_l$ such that

$$\Psi(\mathbf{x}, g\mathbf{y}) = k\Psi(h\mathbf{x}, \mathbf{y}) \quad \text{for } \mathbf{x}, \mathbf{y} \in \mathbb{R}^n. \tag{2.2}$$

With $k = \mathrm{id}$, statement (2.2) is called the *simplified Leon–Proschan property*.

Example 2.2. Take $\Psi : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^l$ to be given by $\Psi(\mathbf{x}, \mathbf{y}) = \Phi(\mathbf{x} + \mathbf{y})$, where $\Phi : \mathbb{R}^n \to \mathbb{R}^l$ is a permutation-invariant function.

We have

$$\Psi(\mathbf{x}, p\mathbf{y}) = \Psi(p^{-1}\mathbf{x}, \mathbf{y}) \quad \text{for } \mathbf{x}, \mathbf{y} \in \mathbb{R}^n \text{ and } p \in \mathbb{P}_n.$$

Therefore, (2.2) holds with $g = p$, $h = p^{-1} = p^T$ and $k = \mathrm{id}$.

More generally, let $\Psi : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^l$ be a permutation-invariant function in the sense

$$\Psi(p\mathbf{x}, p\mathbf{y}) = \Psi(\mathbf{x}, \mathbf{y}) \quad \text{for all } \mathbf{x}, \mathbf{y} \in \mathbb{R}^n \text{ and } p \in \mathbb{P}_n.$$

Then it follows that

$$\Psi(\mathbf{x}, p\mathbf{y}) = \Psi(p^{-1}\mathbf{x}, \mathbf{y}) \quad \text{for all } \mathbf{x}, \mathbf{y} \in \mathbb{R}^n \text{ and } p \in \mathbb{P}_n,$$

which is the simplified Leon–Proschan property with $g = p$ and $h = p^{-1} = p^T$.

Throughout, $\leq$ stands for the componentwise order on $\mathbb{R}^l$ with $l \in \mathbb{N}$.

For a given function $f : \mathbb{R} \to \mathbb{R}$, we extend f to $\mathbb{R}^n$ by

$$f((x_1, \ldots, x_n)^T) = (f(x_1), \ldots, f(x_n))^T \quad \text{for } x_1, \ldots, x_n \in \mathbb{R} \tag{2.3}$$

In the sequel, for a map $\Psi : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^l$, we consider the set

$$\mathcal{A}_\Psi = \{\mathbf{x} \in \mathbb{R}^n : \text{ for } \mathbf{y}, \mathbf{z} \in \mathbb{R}^n \text{ inequality } \mathbf{y} \leq \mathbf{z} \text{ implies } \Psi(\mathbf{x}, \mathbf{y}) \leq \Psi(\mathbf{x}, \mathbf{z})\}.$$

In other words,

$$\mathcal{A}_\Psi = \{\mathbf{x} \in \mathbb{R}^n : \text{ the one-variable map } \Psi(\mathbf{x}, \cdot\,) \text{ is nondecreasing on } \mathbb{R}^n\}. \tag{2.4}$$

For example, if Ψ is the inner product map, then

$$\mathcal{A}_\Psi = \{\mathbf{x} \in \mathbb{R}^n : \text{ for } \mathbf{y}, \mathbf{z} \in \mathbb{R}^n \text{ inequality } \mathbf{y} \leq \mathbf{z} \text{ implies } \langle \mathbf{x}, \mathbf{y}\rangle \leq \langle \mathbf{x}, \mathbf{z}\rangle\} = \mathbb{R}^n_+.$$

Theorem 2.3. *Let $\Psi : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^l$ be a map. Let $f : \mathbb{R} \to \mathbb{R}$ be a convex function. Assume the following conditions:*

(i) *The map Ψ admits the simplified Leon–Proschan property, that is, for each $g \in \mathbb{P}_n$ there exists $h \in \mathbb{P}_n$ such that*

$$\Psi(\mathbf{x}, g\mathbf{y}) = \Psi(h\mathbf{x}, \mathbf{y}) \quad \text{for } \mathbf{x}, \mathbf{y} \in \mathbb{R}^n.$$

(ii) *For each $\mathbf{x} \in \mathbb{R}^n$ the one-variable map $\Psi(\mathbf{x}, \cdot\,)$ is convex (with respect to $\leq$) on $\mathbb{R}^n$, i.e., for $\mathbf{y}_1, \ldots, \mathbf{y}_m \in \mathbb{R}^n$, $t_1, \ldots, t_m \geq 0$, $\sum_{i=1}^m t_i = 1$,*

$$\Psi\Big(\mathbf{x}, \sum_{i=1}^m t_i\mathbf{y}_i\Big) \leq \sum_{i=1}^m t_i\Psi(\mathbf{x}, \mathbf{y}_i).$$

(iii) *For each* $\mathbf{y} \in \mathbb{R}^n$ *the one-variable map* $\Psi(\cdot, \mathbf{y})$ *is concave (with respect to* $\leq$*) on* $\mathbb{R}^n$*, i.e., for* $\mathbf{x}_1, \dots, \mathbf{x}_m \in \mathbb{R}^n$*,* $t_1, \dots, t_m \geq 0$*,* $\sum_{i=1}^m t_i = 1$*,*

$$\Psi\Big(\sum_{i=1}^m t_i \mathbf{x}_i, \mathbf{y}\Big) \geq \sum_{i=1}^m t_i \Psi(\mathbf{x}_i, \mathbf{y}).$$

Fix any $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ *and* $\mathbf{a}, \mathbf{b} \in \mathbb{R}^n$ *with* $\mathbf{b} \in \mathcal{A}_\Psi$*. If*

$$\mathbf{y} = \sum_{i=1}^m t_i g_i \mathbf{x} \quad \textit{and} \quad \mathbf{a} = \sum_{i=1}^m t_i h_i \mathbf{b} \tag{2.5}$$

for some $g_i \in \mathbb{P}_n$ *and* $h_i \in \mathbb{P}_n$ *such that* $\Psi(h_i \mathbf{x}, \mathbf{y}) = \Psi(\mathbf{x}, g_i \mathbf{y})$*,* $i = 1, \dots, m$*, then the following Sherman-type inequality holds:*

$$\Psi(\mathbf{b}, f(\mathbf{y})) \leq \Psi(\mathbf{a}, f(\mathbf{x})). \tag{2.6}$$

Proof. For any $\mathbf{z} \in \mathbb{R}^n$ we get

$$\Psi\Big(\mathbf{b}, \sum_{i=1}^m t_i g_i \mathbf{z}\Big) \leq \sum_{i=1}^m t_i \Psi(\mathbf{b}, g_i \mathbf{z}) = \sum_{i=1}^m t_i \Psi(h_i \mathbf{b}, \mathbf{z}) \leq \Psi\Big(\sum_{i=1}^m t_i h_i \mathbf{b}, \mathbf{z}\Big) = \Psi(\mathbf{a}, \mathbf{z}). \tag{2.7}$$

In fact, the first inequality is due to (ii). The first equality follows from (i). The second inequality is a consequence of (iii). And the last equality is valid by (2.5).

By setting $\mathbf{z} = f(\mathbf{x})$, from (2.7) we have

$$\Psi\Big(\mathbf{b}, \sum_{i=1}^m t_i g_i f(\mathbf{x})\Big) \leq \Psi(\mathbf{a}, f(\mathbf{x})). \tag{2.8}$$

Because the extension (2.3) of f is convex on $\mathbb{R}^n$ (with respect to $\leq$), we find that

$$f\Big(\sum_{i=1}^m t_i g_i \mathbf{x}\Big) \leq \sum_{i=1}^m t_i f(g_i \mathbf{x}).$$

Hence, by the monotonicity of $\Psi(\mathbf{b}, \cdot)$ on $\mathbb{R}^n$ (with respect to $\leq$) (see (2.4)), we obtain

$$\Psi\Big(\mathbf{b}, f\Big(\sum_{i=1}^m t_i g_i \mathbf{x}\Big)\Big) \leq \Psi\Big(\mathbf{b}, \sum_{i=1}^m t_i f(g_i \mathbf{x})\Big).$$

It is not hard to check that

$$f(g_i \mathbf{x}) = g_i f(\mathbf{x}), \quad i = 1, \dots, m.$$

Therefore, the last inequality becomes

$$\Psi\Big(\mathbf{b}, f\Big(\sum_{i=1}^m t_i g_i \mathbf{x}\Big)\Big) \leq \Psi\Big(\mathbf{b}, \sum_{i=1}^m t_i g_i f(\mathbf{x})\Big). \tag{2.9}$$

Finally, by combining (2.5), (2.8) and (2.9) we derive a Sherman-type inequality as follows:

$$\Psi(\mathbf{b}, f(\mathbf{y})) = \Psi\Big(\mathbf{b}, f\Big(\sum_{i=1}^m t_i g_i \mathbf{x}\Big)\Big) \leq \Psi\Big(\mathbf{b}, \sum_{i=1}^m t_i g_i f(\mathbf{x})\Big) \leq \Psi(\mathbf{a}, f(\mathbf{x})).$$

This completes the proof. □

By $\mathbb{M}_n$ we denote the space of all $n \times n$ real matrices. Clearly, $\mathbb{P}_n \subset \mathbb{D}_n \subset \mathbb{M}_n$.

Corollary 2.4. *Let* $f : \mathbb{R} \to \mathbb{R}$ *be a convex function and let* $\Psi : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^l$ *be a map satisfying assumptions* (ii) *and* (iii) *of Theorem 2.3. Additionally, let* $\theta : \mathbb{M}_n \to \mathbb{M}_n$ *be a linear map such that*

(i') *for each* $g \in \mathbb{P}_n$*,*

$$\Psi(\mathbf{x}, g\mathbf{y}) = \Psi(\theta(g)\mathbf{x}, \mathbf{y}) \quad \textit{for } \mathbf{x}, \mathbf{y} \in \mathbb{R}^n. \tag{2.10}$$

Fix any $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ *and* $\mathbf{a}, \mathbf{b} \in \mathbb{R}^n$ *with* $\mathbf{b} \in \mathcal{A}_\Psi$*. If*

$$\mathbf{y} = \mathbf{S}\mathbf{x} \quad \textit{and} \quad \mathbf{a} = \theta(\mathbf{S})\mathbf{b} \tag{2.11}$$

for some $\mathbf{S} \in \mathbb{D}_n$*, then inequality* (2.6) *holds.*

Proof. Since $\mathbf{S} \in \operatorname{conv} \mathbb{P}_n$, we obtain that $\mathbf{S} = \sum_{i=1}^m t_i g_i$ for some $g_1, \ldots, g_m \in \mathbb{P}_n$ and $t_1, \ldots, t_m \geq 0$ with $t_1 + \ldots + t_m = 1$. Then (2.11) implies (2.5). So, it is enough to apply Theorem 2.3. □

Corollary 2.5. *Let $f : \mathbb{R} \to \mathbb{R}$ be a convex function and let $\Psi : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^l$ be a map satisfying assumptions* (ii) *and* (iii) *of Theorem 2.3. Additionally, assume that*
(i') *a map Ψ is permutation-invariant in the sense that for each $g \in \mathbb{P}_n$*

$$\Psi(g\mathbf{x}, g\mathbf{y}) = \Psi(\mathbf{x}, \mathbf{y}) \quad \text{for } \mathbf{x}, \mathbf{y} \in \mathbb{R}^n.$$

Fix any $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ and $\mathbf{a}, \mathbf{b} \in \mathbb{R}^n$ with $\mathbf{b} \in \mathcal{A}_\Psi$. If

$$\mathbf{y} = \mathbf{S}\mathbf{x} \quad \text{and} \quad \mathbf{a} = \mathbf{S}^T\mathbf{b} \tag{2.12}$$

for some $\mathbf{S} \in \mathbb{D}_n$, then inequality (2.6) *holds.*

Proof. It follows from (i') that

$$\Psi(\mathbf{x}, g\mathbf{y}) = \Psi(g^{-1}\mathbf{x}, \mathbf{y}) \quad \text{for } \mathbf{x}, \mathbf{y} \in \mathbb{R}^n \text{ and } g \in \mathbb{P}_n.$$

However, for $g \in \mathbb{P}_n$ one has $g^{-1} = g^T$. So, (2.10) is met with $\theta(g) = g^{-1} = g^T$ for $g \in \mathbb{P}_n$. Therefore, the usage of Corollary 2.4 with $\theta = (\cdot)^T$ leads us to (2.6) via (2.11) and (2.12), as desired. □

3 Sherman-type inequalities induced by directional derivative of a Schur-convex function

We remind that for a function $\psi : \mathbb{R}^n \to \mathbb{R}$ the directional derivative $\nabla_{\mathbf{y}}\psi(\mathbf{x})$ of ψ at the point $\mathbf{x}$ in the direction $\mathbf{y}$ is given by

$$\nabla_{\mathbf{y}}\psi(\mathbf{x}) = \lim_{t\to 0} \frac{\psi(\mathbf{x} + t\mathbf{y}) - \psi(\mathbf{x})}{t}$$

(provided the limit there exists).

It is readily seen that if ψ is permutation-invariant, i.e., $\psi(p\mathbf{x}) = \psi(\mathbf{x})$ for $\mathbf{x} \in \mathbb{R}^n$ and $p \in \mathbb{P}_n$, then

$$\nabla_{p\mathbf{y}}\psi(p\mathbf{x}) = \nabla_{\mathbf{y}}\psi(\mathbf{x}) \quad \text{for } \mathbf{x}, \mathbf{y} \in \mathbb{R}^n \text{ and } p \in \mathbb{P}_n. \tag{3.1}$$

Thus the directional derivative of ψ is a permutation-invariant map.

By taking $\Psi(\mathbf{x}, \mathbf{y}) = \nabla_{\mathbf{y}}\psi(\mathbf{x})$ with $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, we find that

$$\mathcal{A}_\Psi = \{\mathbf{x} \in \mathbb{R}^n : \text{ the one-variable map } \Psi(\mathbf{x}, \cdot) = \nabla_{\cdot}\psi(\mathbf{x}) \text{ is nondecreasing on } \mathbb{R}^n\}.$$

Theorem 3.1. *Let $\psi : \mathbb{R}^n \to \mathbb{R}$ be a Schur-convex function. Assume that for any $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ there exists the directional derivative $\nabla_{\mathbf{y}}\psi(\mathbf{x})$ of ψ at the point $\mathbf{x}$ in the direction $\mathbf{y}$. Let $f : \mathbb{R} \to \mathbb{R}$ be convex. Assume that assumptions* (ii) *and* (iii) *of Theorem 2.3 are satisfied for $\Psi(\mathbf{x}, \mathbf{y}) = \nabla_{\mathbf{y}}\psi(\mathbf{x})$ with $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$.*

Fix any $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ and $\mathbf{a}, \mathbf{b} \in \mathbb{R}^n$ with $\mathbf{b} \in \mathcal{A}_\Psi$. If

$$\mathbf{y} = \mathbf{S}\mathbf{x} \quad \text{and} \quad \mathbf{a} = \mathbf{S}^T\mathbf{b} \tag{3.2}$$

for some $\mathbf{S} \in \mathbb{D}_n$, then the following Sherman-type inequality holds:

$$\nabla_{f(\mathbf{y})}\psi(\mathbf{b}) \leq \nabla_{f(\mathbf{x})}\psi(\mathbf{a}). \tag{3.3}$$

Proof. Since ψ is Schur-convex, it is permutation-invariant. By virtue of (3.1), the directional derivative of ψ is permutation-invariant. That is, Corollary 2.5 (i') is fulfilled with $\Psi(\mathbf{x}, \mathbf{y}) = \nabla_{\mathbf{y}}\psi(\mathbf{x})$ for $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$. Simultaneously, (3.2) gives (2.12) which implies (2.6). Thus we get (3.3), completing the proof. □

We now consider Theorem 3.1 in the context of ψ with Gâteaux differentiability. That is, we assume that the directional derivative $\nabla_{\mathbf{y}}\psi(\mathbf{x})$, viewed as a function of a direction $\mathbf{y}$, is linear and continuous on $\mathbb{R}^n$. Then there exists the gradient $\nabla\psi(\mathbf{x})$ of ψ at the point $\mathbf{x}$ such that

$$\nabla_{\mathbf{y}}\psi(\mathbf{x}) = \langle \nabla\psi(\mathbf{x}), \mathbf{y} \rangle \quad \text{for } \mathbf{x}, \mathbf{y} \in \mathbb{R}^n,$$

where $\langle \cdot, \cdot \rangle$ is the standard inner product on $\mathbb{R}^n$.

For the map $\Psi(\mathbf{x}, \mathbf{y}) = \langle \nabla\psi(\mathbf{x}), \mathbf{y} \rangle$ with $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, we find that

$$\begin{aligned}\mathcal{A}_\Psi &= \{\mathbf{x} \in \mathbb{R}^n : \text{ the one-variable map } \langle \nabla\psi(\mathbf{x}), \cdot \rangle \text{ is nonnegative on } \mathbb{R}^n\} \\ &= \{\mathbf{x} \in \mathbb{R}^n : \nabla\psi(\mathbf{x}) \in \mathbb{R}^n_+\}.\end{aligned}$$

Thus the condition $\mathbf{b} \in \mathcal{A}_\Psi$ means that $\nabla\psi(\mathbf{b}) \in \mathbb{R}^n_+$.

Corollary 3.2. *Let $\psi : \mathbb{R}^n \to \mathbb{R}$ be a Gâteaux differentiable Schur-convex function. Let $f : \mathbb{R} \to \mathbb{R}$ be convex. Assume that the assumption* (iii) *of Theorem 2.3 is satisfied for $\Psi(\mathbf{x}, \mathbf{y}) = \langle \nabla\psi(\mathbf{x}), \mathbf{y} \rangle$ with $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$.*

Fix any $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ and $\mathbf{a}, \mathbf{b} \in \mathbb{R}^n$ with $\nabla\psi(\mathbf{b}) \in \mathbb{R}^n_+$. If

$$\mathbf{y} = \mathbf{S}\mathbf{x} \quad \text{and} \quad \mathbf{a} = \mathbf{S}^T\mathbf{b}$$

for some $\mathbf{S} \in \mathbb{D}_n$, then the following Sherman-type inequality holds:

$$\langle \nabla\psi(\mathbf{b}), f(\mathbf{y}) \rangle \le \langle \nabla\psi(\mathbf{a}), f(\mathbf{x}) \rangle. \tag{3.4}$$

Proof. Condition (ii) of Theorem 2.3 holds by the linearity of $\Psi(\mathbf{x}, \mathbf{y}) = \langle \nabla\psi(\mathbf{x}), \mathbf{y} \rangle$ with respect to $\mathbf{y}$. Therefore, it is sufficient to apply Theorem 3.1. □

By putting

$$\psi(\mathbf{x}) = \|\mathbf{x}\|^2 = \langle \mathbf{x}, \mathbf{x} \rangle \quad \text{for } \mathbf{x} \in \mathbb{R}^n, \tag{3.5}$$

which is a Schur-convex function by virtue of its convexity and permutation-invariance, we have

$$\nabla\psi(\mathbf{x}) = 2\mathbf{x} \quad \text{for } \mathbf{x} \in \mathbb{R}^n. \tag{3.6}$$

In this case, condition $\nabla\psi(\mathbf{b}) \in \mathbb{R}^n_+$ means that $\mathbf{b} \in \mathbb{R}^n_+$.

It is interesting that under (3.5) and (3.6) inequality (3.4) in Corollary 3.2 holds in the form

$$\langle \mathbf{b}, f(\mathbf{y}) \rangle \le \langle \mathbf{a}, f(\mathbf{x}) \rangle,$$

which is the classical Sherman's inequality (see Theorem B). Thus Corollary 3.2 is a generalization of Theorem B (whenever $m = n$). Moreover, by setting $\mathbf{a} = (1, \dots, 1) \in \mathbb{R}^n$ with a doubly stochastic $\mathbf{S}$, we can get Theorem A.

It is not hard to verify for the map $\Psi(\mathbf{x}, \mathbf{y}) = \nabla_\mathbf{x}\psi(\mathbf{y})$ with $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ that

$$\mathcal{A}_\Psi = \{\mathbf{x} \in \mathbb{R}^n : \text{ the one-variable map } \Psi(\mathbf{x}, \cdot) = \nabla_\mathbf{x}\psi(\cdot) \text{ is nondecreasing on } \mathbb{R}^n\}.$$

Theorem 3.3. *Let $\psi : \mathbb{R}^n \to \mathbb{R}$ be a Schur-convex function. Assume that for any $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ there exists the directional derivative $\nabla_\mathbf{x}\psi(\mathbf{y})$ of ψ at the point $\mathbf{x}$ in the direction $\mathbf{y}$. Let $f : \mathbb{R} \to \mathbb{R}$ be convex. Assume that assumptions* (ii) *and* (iii) *of Theorem 2.3 are satisfied for $\Psi(\mathbf{x}, \mathbf{y}) = \nabla_\mathbf{x}\psi(\mathbf{y})$ with $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$.*

Fix any $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ and $\mathbf{a}, \mathbf{b} \in \mathbb{R}^n$ with $\mathbf{b} \in \mathcal{A}_\Psi$. If

$$\mathbf{y} = \mathbf{S}\mathbf{x} \quad \text{and} \quad \mathbf{a} = \mathbf{S}^T\mathbf{b} \tag{3.7}$$

for some $\mathbf{S} \in \mathbb{D}_n$, then the following Sherman-type inequality holds:

$$\nabla_\mathbf{b}\psi(f(\mathbf{y})) \le \nabla_\mathbf{a}\psi(f(\mathbf{x})). \tag{3.8}$$

Proof. Because ψ is Schur-convex, its directional derivative is permutation-invariant (see (3.1)). For this reason, Corollary 2.5 (i') is satisfied for the map $\Psi(\mathbf{x}, \mathbf{y}) = \nabla_\mathbf{x}\psi(\mathbf{y})$, $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$. Furthermore, (3.7) gives (2.12), which implies (2.6). Finally, we get (3.8), as claimed. □

For the map $\Psi(\mathbf{x}, \mathbf{y}) = \langle \nabla\psi(\mathbf{y}), \mathbf{x} \rangle$ with $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ we have

$$\mathcal{A}_\Psi = \{\mathbf{x} \in \mathbb{R}^n : \text{ the one-variable map } \Psi(\mathbf{x}, \cdot) = \langle \nabla\psi(\cdot), \mathbf{x} \rangle \text{ is nondecreasing on } \mathbb{R}^n\}.$$

Condition $\mathbf{b} \in \mathcal{A}_\Psi$ means that the one-variable map $\Psi(\mathbf{b}, \cdot) = \langle \nabla\psi(\cdot), \mathbf{b} \rangle$ is nondecreasing on $\mathbb{R}^n$, that is, for $\mathbf{y}, \mathbf{z} \in \mathbb{R}^n$,

$$\mathbf{y} \le \mathbf{z} \quad \text{implies} \quad \langle \nabla\psi(\mathbf{y}), \mathbf{b} \rangle \le \langle \nabla\psi(\mathbf{z}), \mathbf{b} \rangle. \tag{3.9}$$

Corollary 3.4. *Let* $\psi : \mathbb{R}^n \to \mathbb{R}$ *be a Gâteaux differentiable Schur-convex function. Let* $f : \mathbb{R} \to \mathbb{R}$ *be convex. Assume that assumptions* (ii) *and* (iii) *of Theorem 2.3 are satisfied for* $\Psi(\mathbf{x}, \mathbf{y}) = \langle \nabla\psi(\mathbf{y}), \mathbf{x} \rangle$ *with* $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$.

Fix any $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ *and* $\mathbf{a}, \mathbf{b} \in \mathbb{R}^n$ *with* $\mathbf{b} \in \mathcal{A}_\Psi$. *If*

$$\mathbf{y} = \mathbf{S}\mathbf{x} \quad \text{and} \quad \mathbf{a} = \mathbf{S}^T\mathbf{b}$$

for some $\mathbf{S} \in \mathbb{D}_n$, *then the following Sherman-type inequality holds:*

$$\langle \nabla\psi(f(\mathbf{y})), \mathbf{b} \rangle \le \langle \nabla\psi(f(\mathbf{x})), \mathbf{a} \rangle. \tag{3.10}$$

Proof. It is sufficient to apply Theorem 3.3. □

To illustrate the last result, choose

$$\psi(\mathbf{x}) = \exp\langle \mathbf{x}, \mathbf{e} \rangle \quad \text{for } \mathbf{x} \in \mathbb{R}^n, \tag{3.11}$$

where $\mathbf{e} = (1, \dots, 1) \in \mathbb{R}^n$. This is a Schur-convex function.

It is readily seen that

$$\nabla\psi(\mathbf{x}) = (\exp\langle \mathbf{x}, \mathbf{e} \rangle)\mathbf{e} \quad \text{for } \mathbf{x} \in \mathbb{R}^n. \tag{3.12}$$

Here

$$\Psi(\mathbf{x}, \mathbf{y}) = \langle \nabla\psi(\mathbf{y}), \mathbf{x} \rangle = \big\langle (\exp\langle \mathbf{y}, \mathbf{e} \rangle)\mathbf{e}, \mathbf{x} \big\rangle = \langle \mathbf{e}, \mathbf{x} \rangle \exp\langle \mathbf{y}, \mathbf{e} \rangle.$$

Evidently, this map is convex with respect to $\mathbf{y}$ and concave with respect to $\mathbf{x}$, which proves the validity of conditions (ii) and (iii) in Theorem 2.3.

In this case, condition $\mathbf{b} \in \mathcal{A}_\Psi$ means that (3.9) holds in the form

$$\mathbf{y} \le \mathbf{z} \quad \text{implies} \quad \langle \mathbf{e}, \mathbf{b} \rangle \exp\langle \mathbf{y}, \mathbf{e} \rangle \le \langle \mathbf{e}, \mathbf{b} \rangle \exp\langle \mathbf{z}, \mathbf{e} \rangle,$$

which is true whenever $\langle \mathbf{e}, \mathbf{b} \rangle \ge 0$.

It follows from (3.11) and (3.12) that inequality (3.10) in Corollary 3.4 holds in the form

$$\langle \mathbf{e}, \mathbf{b} \rangle \exp\langle f(\mathbf{y}), \mathbf{e} \rangle \le \langle \mathbf{e}, \mathbf{a} \rangle \exp\langle f(\mathbf{x}), \mathbf{e} \rangle.$$

References

[1] M. Adil Khan, S. Ivelić Bradanović and J. Pečarić, On Sherman's type inequalities for n-convex function with applications, *Konuralp J. Math.* **4** (2016), no. 2, 255–270.

[2] M. Adil Khan, J. Khan and J. Pečarić, Generalizations of Sherman's inequality by Montgomery identity and Green function, *Electron. J. Math. Anal. Appl.* **5** (2017), no. 1, 1–16.

[3] R. P. Agarwal, S. Ivelić Bradanović and J. Pečarić, Generalizations of Sherman's inequality by Lidstone's interpolating polynomial, *J. Inequal. Appl.* **2016** (2016), DOI 10.1186/s13660-015-0935-6.

[4] A.-M. Burtea, Two examples of weighted majorization, *An. Univ. Craiova Ser. Mat. Inform.* **37** (2010), no. 2, 92–99.

[5] M. L. Eaton, On group induced orderings, monotone functions, and convolution theorems, Inequalities in statistics and probability (Lincoln 1982), IMS Lecture Notes Monogr. Ser. 5, Institute of Mathematical Statistics, Hayward (1984), 13–25.

[6] M. L. Eaton, Group induced orderings with some applications in statistics, *CWI Newslett.* (1987), no. 16, 3–31.

[7] M. L. Eaton and M. D. Perlman, Reflection groups, generalized Schur functions, and the geometry of majorization, *Ann. Probab.* **5** (1977), no. 6, 829–860.

[8] G. H. Hardy, J. E. Littlewood and G. Pólya, *Inequalities*, 2nd ed., Cambridge, at the University Press, 1952.

[9] S. Ivelić Bradanović, N. Latif and J. Pečarić, Generalizations of Sherman's theorem by Taylor's formula, *J. Inequal. Spec. Funct.* **8** (2017), no. 2, 18–30.

[10] S. Ivelić Bradanović and J. Pečarić, Generalizations of Sherman's inequality, *Period. Math. Hungar.* **74** (2017), no. 2, 197–219.

[11] J. Karamata, Sur une inégalité rélative aux fonctions convexes, *Publ. Math. Univ. Belgrade* **1** (1932), 145–148.

[12] R. V. León and F. Proschan, An inequality for convex functions involving G-majorization, *J. Math. Anal. Appl.* **69** (1979), no. 2, 603–606.

[13] A. W. Marshall, I. Olkin and B. C. Arnold, *Inequalities: Theory of Majorization and its Applications*, 2nd ed., Springer Ser. Statist., Springer, New York, 2011.
[14] M. Niezgoda, Remarks on Sherman like inequalities for (α, β)-convex functions, *Math. Inequal. Appl.* **17** (2014), no. 4, 1579–1590.
[15] M. Niezgoda, Vector joint majorization and generalization of Csiszár–Körner's inequality for f-divergence, *Discrete Appl. Math.* **198** (2016), 195–205.
[16] M. Niezgoda, On Sherman method to deriving inequalities for some classes of functions related to convexity, preprint (2018).
[17] S. Sherman, On a theorem of Hardy, Littlewood, Polya, and Blackwell, *Proc. Natl. Acad. Sci. USA* **37** (1951), 826–831.