Filomat 36:6 (2022), 1937–1942
https://doi.org/10.2298/FIL2206937N

Published by Faculty of Sciences and Mathematics, University of Niš, Serbia
Available at: http://www.pmf.ni.ac.rs/filomat

On Slater Like Inequality for Vectors Transformed by a Doubly Stochastic Matrix with Control Function

Marek Niezgoda[a]

[a] *Institute of Mathematics, Pedagogical University of Cracow, Podchorążych 2, 30-084 Kraków, Poland*

Abstract. In this work, we deal with a Slater type inequality designed for a symmetric convex function and for a collection of vectors transformed by a doubly stochastic matrix. In doing so, we use an additional convex control function. In the case when the composition of the control function and of the underlying convex function is Schur-concave, such an approach leads to a refinement of the standard Slater inequality. Special cases are also considered.

1. Background and preliminaries

In [10] Slater proved the following result.

Theorem A (Slater [10, Theorem 1].) *Suppose that $f : I \to \mathbb{R}$ is a convex and increasing function on interval $I = (a, b)$. Then for $x_i \in I$ and $p_i \geq 0$, $i = 1, \ldots, m$, with $\sum_{i=1}^m p_i > 0$ and $\sum_{i=1}^m p_i f'_+(x_i) > 0$, it holds that*

$$\frac{\sum_{i=1}^m p_i f(x_i)}{\sum_{i=1}^m p_i} \leq f\left(\frac{\sum_{i=1}^m p_i f'_+(x_i) x_i}{\sum_{i=1}^m p_i f'_+(x_i)}\right). \tag{1}$$

A multidimensional generalization of Slater's inequality is due to Pečarić [8].

Theorem B (Pečarić [8, Theorem].) *Let $f : I \to \mathbb{R}$ be a convex function on an open set $I \subset \mathbb{R}^n$, and let $x_i \in I$ and $p_i \geq 0$, $i = 1, \ldots, m$, with $P_m = \sum_{i=1}^m p_i > 0$. If $A \in I$ exists such that*

$$\langle A, \sum_{i=1}^m p_i f'_+(x_i)\rangle \geq \sum_{i=1}^m p_i \langle x_i, f'_+(x_i)\rangle, \tag{2}$$

where $f'_+(x) = (f'_{1+}(x), \ldots, f'_{n+}(x))$ and $f'_{1+}, \ldots, f'_{n+}$ are right partial derivatives of f, then

$$\frac{1}{P_m} \sum_{i=1}^m p_i f(x_i) \leq f(A). \tag{3}$$

2020 *Mathematics Subject Classification*. Primary 26B25, 26D10 ; Secondary 26D15, 52A41
Keywords. convex function, gradient inequality, Slater inequality, majorization, doubly stochastic matrix, Schur-concave function
Received: 27 June 2019; Accepted: 28 January 2020
Communicated by Dragan S. Djordjević
Email address: bniezgoda@wp.pl (Marek Niezgoda)

See [1, 4, 5] for further results on Slater type inequalities.

In the present paper, we demonstrate a further result of Slater type in the framework of majorization theory (see Section 2).

To do so, in the rest of this section we collect some basic notation, definitions and facts.

As usual, the n-dimensional Euclidean space $\mathbb{R}^n$ is equipped with the following inner product

$$\langle \mathbf{x}, \mathbf{y} \rangle = \sum_{i=1}^{n} x_i y_i \quad \text{for } \mathbf{x} = (x_1, \ldots, x_n)^T \text{ and } \mathbf{y} = (y_1, \ldots, y_n)^T.$$

Hereafter the symbol $(\cdot)^T$ stands for the transpose of a matrix.

The space $\mathbb{R}^n$ is also endowed with the entrywise order $\leq$ given by

$$\mathbf{y} \leq \mathbf{x} \quad \text{iff} \quad y_i \leq x_i \quad \text{for each } i = 1, \ldots, n, \tag{4}$$

where $\mathbf{x} = (x_1, \ldots, x_n)^T$ and $\mathbf{y} = (y_1, \ldots, y_n)^T$.

A function $\Phi : I \to \mathbb{R}$ defined on a convex set $I \subset \mathbb{R}^n$ is said to be *convex* on I, if for any points $\mathbf{x}_i \in I$ and scalars $t_i \geq 0$, $i = 1, \ldots, m$, with $\sum\limits_{i=1}^{m} t_i = 1$, the following *Jensen's inequality* holds:

$$\Phi\left(\sum_{i=1}^{m} t_i \mathbf{x}_i\right) \leq \sum_{i=1}^{m} t_i \Phi(\mathbf{x}_i).$$

If a convex function $\Phi : I \to \mathbb{R}$ is differentiable then the following *gradient inequality* holds:

$$\langle \nabla\Phi(\mathbf{y}), \mathbf{x} - \mathbf{y} \rangle \leq \Phi(\mathbf{x}) - \Phi(\mathbf{y}) \quad \text{for } \mathbf{x}, \mathbf{y} \in I.$$

Hereafter the symbol ∇ denotes the gradient, and $\langle \cdot, \cdot \rangle$ is the inner product on $\mathbb{R}^n$.

Let $\mathbf{x} = (x_1, \ldots, x_n)^T \in \mathbb{R}^n$ and $\mathbf{y} = (y_1, \ldots, y_n)^T \in \mathbb{R}^n$. We say that $\mathbf{x}$ *majorizes* $\mathbf{y}$, and write $\mathbf{y} \prec \mathbf{x}$, if

$$\sum_{i=1}^{l} y_{[i]} \leq \sum_{i=1}^{l} x_{[i]} \quad \text{for } l = 1, \ldots, n, \quad \text{and} \quad \sum_{i=1}^{n} y_i = \sum_{i=1}^{n} x_i,$$

where $x_{[1]} \geq \cdots \geq x_{[n]}$ and $y_{[1]} \geq \cdots \geq y_{[n]}$ are the entries of $\mathbf{x}$ and $\mathbf{y}$, respectively, stated in decreasing order [6, p. 8].

Throughout the symbol conv stands for "the convex hull of", and $\mathbb{P}_n$ denotes the group of $n \times n$ permutation matrices.

It is a result of Rado [9] that

$$\mathbf{y} \prec \mathbf{x} \quad \text{if and only if} \quad \mathbf{y} \in \operatorname{conv} \mathbb{P}_n \mathbf{x} \tag{5}$$

for $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ (see also [6, p. 10]).

An $n \times n$ real matrix $\mathbf{S} = (s_{ij})$ is said to be *doubly stochastic* provided that $s_{ij} \geq 0$ for $i, j = 1, \ldots, n$, and all column sums and row sums of $\mathbf{S}$ are equal to 1, i.e., $\sum\limits_{i=1}^{n} s_{ij} = 1$ for $j = 1, \ldots, n$, and $\sum\limits_{j=1}^{n} s_{ij} = 1$ for $i = 1, \ldots, n$.

By Ω_n we denote the set of all $n \times n$ doubly stochastic matrices.

It is known by Birkhoff's Theorem (see [6, Theorem A.2.]) that

$$\Omega_n = \operatorname{conv} \mathbb{P}_n.$$

Therefore for $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, (5) can be restated as

$$\mathbf{y} \prec \mathbf{x} \quad \text{if and only if} \quad \mathbf{y} = \mathbf{S}\mathbf{x} \tag{6}$$

for some doubly stochastic $n \times n$ matrix $\mathbf{S}$ [6, p. 33].

A function $\Phi : J^n \to \mathbb{R}$ with an interval $J \subset \mathbb{R}$ is called *Schur-convex* (resp. *Schur-concave*) on J^n if for $\mathbf{x}, \mathbf{y} \in J^n$,

$$\mathbf{y} \prec \mathbf{x} \text{ implies } \Phi(\mathbf{y}) \leq \text{(resp. } \geq \text{)} \ \Phi(\mathbf{x})$$

(see [6, p. 79-154]).

Theorem C (Hardy-Littlewood-Pólya [2] and Karamata [3].) *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 $\mathbf{y} = (y_1, y_2, \ldots, y_n)^T \in J^n$,

$$\mathbf{y} \prec \mathbf{x} \quad \textit{implies} \quad \sum_{i=1}^{n} f(y_i) \leq \sum_{i=1}^{n} f(x_i). \tag{7}$$

Throughout the symbols J and J_0 represent any intervals in $\mathbb{R}$.

A function $\Phi : J^n \to \mathbb{R}$ is said to be $\mathbb{P}_n$-*invariant* (also called, *symmetric*), if

$$\Phi(\mathbf{p}\mathbf{x}) = \Phi(\mathbf{x}) \quad \text{for } \mathbf{p} \in \mathbb{P}_n \text{ and } \mathbf{x} \in J^n.$$

A function $F : J_0^n \to \mathbb{R}^n$ is said to be $\mathbb{P}_n$-*equivariant* if

$$F(\mathbf{p}\mathbf{x}) = \mathbf{p}F(\mathbf{x}) \quad \text{for } \mathbf{p} \in \mathbb{P}_n \text{ and } \mathbf{x} \in J_0^n.$$

In the sequel, by $\leq$ we denote the componentwise order on $\mathbb{R}^l$ with any $l \in \mathbb{N}$ (see (4)).

A function $F : J_0^n \to \mathbb{R}^l$ is said to be *convex* on J_0^n if

$$F\left(\sum_{i=1}^{k} t_i \mathbf{x}_i\right) \leq \sum_{i=1}^{k} t_i F(\mathbf{x}_i)$$

for any $k \in \mathbb{N}$, $\mathbf{x}_i \in J_0^n$, $t_i \geq 0$, $i = 1, \ldots, k$, $\sum_{i=1}^{k} t_i = 1$.

The next result provides a majorization gradient inequality for differentiable symmetric convex functions.

Theorem D ([7, Theorem 2.3].) *Let $\Phi : J^n \to \mathbb{R}$ be a differentiable $\mathbb{P}_n$-invariant convex function, and $F : J_0^n \to J^n$ be a $\mathbb{P}_n$-equivariant convex function. Let $\mathbf{x}, \mathbf{y} \in J_0^n$ and $\mathbf{a}, \mathbf{b} \in J^n$ with $\nabla\Phi(\mathbf{b}) \geq 0$.*

If

$$\mathbf{y} = \mathbf{S}\mathbf{x} \quad \textit{and} \quad \mathbf{a} = \mathbf{S}^T \mathbf{b} \tag{8}$$

for some $n \times n$ doubly stochastic matrix $\mathbf{S}$, then

$$\langle \nabla\Phi(\mathbf{b}), F(\mathbf{y}) - \mathbf{b} \rangle \leq \Phi(F(\mathbf{x})) - \Phi(\mathbf{a}). \tag{9}$$

In Section 2, by using Theorem **D**, we derive a majorization Slater like inequality. Next, we discuss some special cases.

2. A Slater like inequality for transformed vectors

As previously, the symbols J and J_0 stand for any intervals in $\mathbb{R}$.

In the forthcoming theorem we present a majorization extension of Slater inequality (3) for a differentiable symmetric convex function Φ and for a collection of vectors $\mathbf{a}_1, \ldots, \mathbf{a}_m$ transformed by a doubly stochastic matrix.

Theorem 2.1. *Let $\Phi : J^n \to \mathbb{R}$ be a differentiable $\mathbb{P}_n$-invariant convex function, and $F : J_0^n \to J^n$ be a $\mathbb{P}_n$-equivariant convex function. Let $\mathbf{y} \in J_0^n$, $\mathbf{a}_i \in J^n$ and $p_i \geq 0$, $i = 1, \dots, m$, with $P_m = \sum_{i=1}^m p_i > 0$.*

If there exist $\mathbf{x} \in J_0^n$ and $\mathbf{b}_i \in J^n$ with $\nabla\Phi(\mathbf{b}_i) \geq 0$, $i = 1, \dots, m$, satisfying

$$\mathbf{y} = \mathbf{S}\mathbf{x} \quad \textit{and} \quad \mathbf{a}_i = \mathbf{S}^T \mathbf{b}_i \tag{10}$$

for some $n \times n$ doubly stochastic matrix $\mathbf{S}$, and

$$\sum_{i=1}^m p_i \langle \nabla\Phi(\mathbf{b}_i), \mathbf{b}_i \rangle \leq \sum_{i=1}^m p_i \langle \nabla\Phi(\mathbf{b}_i), F(\mathbf{y}) \rangle, \tag{11}$$

then the following Slater type inequality with control function F holds

$$\frac{1}{P_m} \sum_{i=1}^m p_i \Phi(\mathbf{a}_i) \leq \Phi(F(\mathbf{x})). \tag{12}$$

If in addition the composite function $\Phi \circ F$ is Schur-concave on J_0^n, then

$$\frac{1}{P_m} \sum_{i=1}^m p_i \Phi(\mathbf{a}_i) \leq \Phi(F(\mathbf{x})) \leq \Phi(F(\mathbf{y})). \tag{13}$$

Proof. By virtue of [7, Theorem 2.3] (see (8)-(9) in Theorem **D**), we get

$$\langle \nabla\Phi(\mathbf{b}_i), F(\mathbf{y}) - \mathbf{b}_i \rangle \leq \Phi(F(\mathbf{x})) - \Phi(\mathbf{a}_i) \quad \text{for each } i = 1, \dots, m.$$

Hence we obtain

$$\frac{1}{P_m} \sum_{i=1}^m p_i \langle \nabla\Phi(\mathbf{b}_i), F(\mathbf{y}) - \mathbf{b}_i \rangle \leq \Phi(F(\mathbf{x})) - \frac{1}{P_m} \sum_{i=1}^m p_i \Phi(\mathbf{a}_i). \tag{14}$$

On account of (11) we have

$$0 \leq \frac{1}{P_m} \sum_{i=1}^m p_i \langle \nabla\Phi(\mathbf{b}_i), F(\mathbf{y}) - \mathbf{b}_i \rangle. \tag{15}$$

By combining (14) and (15) we get the desired assertion (12).

Now, assume that the composite function $\Phi \circ F$ is Schur-concave. Then (10) implies $\mathbf{y} \prec \mathbf{x}$ (see (6)), so $\Phi(F(\mathbf{x})) \leq \Phi(F(\mathbf{y}))$. This alone with (12) gives (13), as wanted. ∎

Remark 2.2. In Theorem 2.1, the usage of the control function F has the advantage that the composition $\Phi \circ F$ can be Schur-concave for some F, which leads to the double inequality (13) (see Theorems 2.5-2.6).

Remark 2.3. In Theorem 2.1, for given points $\mathbf{b}_i$ and $A = F(\mathbf{y})$, it follows by Theorem **B** from condition (11) (which corresponds to (2)) that

$$\frac{1}{P_m} \sum_{i=1}^m p_i \Phi(\mathbf{b}_i) \leq \Phi(F(\mathbf{y})). \tag{16}$$

If moreover the matrix $\mathbf{S}$ is such that $\mathbf{a}_i = \mathbf{b}_i$ (and the composite function $\Phi \circ F$ is Schur-concave on J_0^n), then (12) can be viewed as a refinement of (16) (see (13)).

Remark 2.4. It is worth emphasizing that if in Theorem 2.1 $F = \mathrm{id}$ (the identity map) with $J_0 = J$, then the assumption $\nabla\Phi(\mathbf{b}) \geq 0$ can be dropped. In fact, in this case the proof of Theorem **D** does not utilize this assumption, because some needed inequalities become equalities and hold trivially.

Theorem 2.5. *Let $\varphi : J \to \mathbb{R}$ be a differentiable convex function, and $f : J_0 \to J$ be a convex function. Let $\mathbf{y} = (y_1, \ldots, y_n)^T \in J_0^n$, $\mathbf{a}_i = (a_{1i}, \ldots, a_{ni})^T \in J^n$ and $p_i \geq 0$, $i = 1, \ldots, m$, with $P_m = \sum_{i=1}^m p_i > 0$.*

If there exist $\mathbf{x} = (x_1, \ldots, x_n)^T \in J_0^n$ and $\mathbf{b}_i = (b_{1i}, \ldots, b_{ni})^T \in J^n$ with $\varphi'(b_{ji}) \geq 0$ for $i = 1, \ldots, m$, such that

$$\mathbf{y} = \mathbf{S}\mathbf{x} \quad and \quad \mathbf{a}_i = \mathbf{S}^T\mathbf{b}_i \tag{17}$$

for some $n \times n$ doubly stochastic matrix $\mathbf{S}$, and

$$\sum_{i=1}^m \sum_{j=1}^n p_i\varphi'(b_{ji})b_{ji} \leq \sum_{i=1}^m \sum_{j=1}^n p_i\varphi'(b_{ji})f(y_j), \tag{18}$$

then

$$\frac{1}{P_m}\sum_{i=1}^m \sum_{j=1}^n p_i\varphi(a_{ji}) \leq \sum_{j=1}^n \varphi(f(x_j)). \tag{19}$$

If additionally the composite function $\varphi \circ f$ is concave on J_0, then

$$\frac{1}{P_m}\sum_{i=1}^m \sum_{j=1}^n p_i\varphi(a_{ji}) \leq \sum_{j=1}^n \varphi(f(x_j)) \leq \sum_{j=1}^n \varphi(f(y_j)). \tag{20}$$

Proof. Let $\Phi : J^n \to \mathbb{R}$ and $F : J_0^n \to J^n$ be functions of n-variables defined by

$$\Phi(\mathbf{c}) = \sum_{j=1}^n \varphi(c_j) \quad \text{for } \mathbf{c} = (c_1, \ldots, c_n)^T \in J^n ,$$

and

$$F(\mathbf{z}) = (f(z_1), \ldots, f(z_n))^T \quad \text{for } \mathbf{z} = (z_1, \ldots, z_n)^T \in J_0^n.$$

It is not hard to check that $\Phi : J^n \to \mathbb{R}$ is a differentiable $\mathbb{P}_n$-invariant convex function with

$$\nabla\Phi(\mathbf{c}) = (\varphi'(c_1), \ldots, \varphi'(c_n))^T \quad \text{for } \mathbf{c} = (c_1, \ldots, c_n)^T \in J^n.$$

Also, $F : J_0^n \to J^n$ is a $\mathbb{P}_n$-equivariant function convex with respect to the componentwise order $\leq$ on $\mathbb{R}^n$. Thus we obtain

$$\sum_{i=1}^m p_i\langle\nabla\Phi(\mathbf{b}_i), \mathbf{b}_i\rangle = \sum_{i=1}^m \sum_{j=1}^n p_i\varphi'(b_{ji})b_{ji}$$

and

$$\sum_{i=1}^m p_i\langle\nabla\Phi(\mathbf{b}_i), F(\mathbf{y})\rangle = \sum_{i=1}^m \sum_{j=1}^n p_i\varphi'(b_{ji})f(y_j).$$

Hence assumption (18) takes the form

$$\sum_{i=1}^m p_i\langle\nabla\Phi(\mathbf{b}_i), \mathbf{b}_i\rangle \leq \sum_{i=1}^m p_i\langle\nabla\Phi(\mathbf{b}_i), F(\mathbf{y})\rangle.$$

Therefore it now follows from Theorem 2.1 that

$$\frac{1}{P_m}\sum_{i=1}^{m} p_i\Phi(\mathbf{a}_i) \le \Phi(F(\mathbf{x})).$$

In other words,

$$\frac{1}{P_m}\sum_{i=1}^{m}\sum_{j=1}^{n} p_i\varphi(a_{ji}) \le \sum_{j=1}^{n}\varphi(f(x_j)). \tag{21}$$

This proves (19).

To see (20), suppose that $\varphi \circ f$ is concave on J_0. Then the function $(\Phi \circ F)(\mathbf{z}) = \sum\limits_{j=1}^{n}\varphi(f(z_j))$ for $\mathbf{z} = (z_1, \dots, z_n)^T \in J_0^n$ is Schur-concave on J_0^n by Theorem **C**.

So, since $\mathbf{y} \prec \mathbf{x}$ by (17), we conclude that

$$\sum_{j=1}^{n}\varphi(f(x_j)) \le \sum_{j=1}^{n}\varphi(f(y_j)),$$

which together with (21) leads to (20). ■

In the next result for a convex function φ we demonstrate a construction of a convex function f in order that the composition $\varphi \circ f$ be concave, as required in the last part of Theorem 2.5.

Theorem 2.6. *Let $\varphi : J \to \mathbb{R}$ be a differentiable convex function on $J = (\alpha, \beta)$. Assume that φ is strictly decreasing on $J_1 = (\alpha, \gamma)$ and φ is increasing on $J_2 = [\gamma, \beta)$.*

Let ψ be the restriction to J_1 of φ, and $\psi^{-1} : \psi(J_1) \to J_1$ be the inverse function of ψ. Let $g : J_0 \to \psi(J_1)$ be a concave function and let $f = \psi^{-1} \circ g : J_0 \to J_1$.

Let $\mathbf{y} = (y_1, \dots, y_n)^T \in J_0^n$, $\mathbf{a}_i = (a_{1i}, \dots, a_{ni})^T \in J^n$ and $p_i \ge 0$, $i = 1, \dots, m$, with $P_m = \sum\limits_{i=1}^{m} p_i > 0$.

Suppose that there exist $\mathbf{x} = (x_1, \dots, x_n)^T \in J_0^n$ and $\mathbf{b}_i = (b_{1i}, \dots, b_{ni})^T \in J^n$ with $\varphi'(b_{ji}) \ge 0$ for $i = 1, \dots, m$, such that $\mathbf{y} = \mathbf{S}\mathbf{x}$ and $\mathbf{a}_i = \mathbf{S}^T\mathbf{b}_i$ for some $n \times n$ doubly stochastic matrix $\mathbf{S}$.

If condition (18) is satisfied then the double inequality (20) holds valid.

Proof. The composite function $f = \psi^{-1} \circ g : J_0 \to J_1 \subset J$ is well-defined. Also, f is convex on J_0, because g is concave and ψ^{-1} is decreasing and convex.

Furthermore, it is clear that

$$\varphi \circ f = \psi \circ f = \psi \circ \psi^{-1} \circ g = g.$$

For this reason the composite function $\varphi \circ f : J_0 \to \mathbb{R}$ is concave on J_0, because g is so.

Finally, all assumptions of Theorem 2.5 are satisfied. In consequence, we infer that (20) holds. ■

References

[1] S. S. Dragomir, Some Slater's type inequalities for convex functions defined on linear spaces and applications, Abstract and Applied Analysis 2012 (2012) art. ID 168405 pp. 16, doi: 10.1155/2012/168405.
[2] G. M. Hardy, J. E. Littlewood, G. Pólya, Inequalities, (2nd edition), Cambridge University Press, Cambridge, 1952.
[3] J. Karamata, Sur une inégalité rélative aux fonctions convexes, Publications Mathématiques de l'Université de Belgrade 1 (1932) 145-148.
[4] M. A. Khan, J. E. Pečarić, Improvement and reversion of Slater's inequality and related results, Journal of Inequalities and Applications 2010 (2010) art. ID 646034 pp. 14.
[5] M. A. Khan, J. E. Pečarić, On Slater's integral inequality, Journal of Mathematical Inequalities 5 (2) (2011) 231–241.
[6] A. W. Marshall, I. Olkin, B. C. Arnold, Inequalities: Theory of Majorization and Its Applications, (2nd edition), Springer, New York 2011.
[7] M. Niezgoda, A majorization gradient inequality for symmetric convex functions, Journal of Convex Analysis 28 (2021) 1-10.
[8] J. E. Pečarić, A multidimensional generalization of Slater's inequality, Journal of Approximation Theory 44 (1985) 292–294.
[9] R. Rado, An inequality, Journal of the London Mathematical Society 27 (1952) 1–6.
[10] M. S. Slater, A companion inequality to Jensen's inequality, Journal of Approximation Theory 32 (2) (1981) 160–166.