Hindawi Publishing Corporation
Abstract and Applied Analysis
Volume 2012, Article ID 168405, 16 pages
doi:10.1155/2012/168405

Research Article

Some Slater's Type Inequalities for Convex Functions Defined on Linear Spaces and Applications

S. S. Dragomir[1, 2]

[1] *Mathematics, School of Engineering and Science, Victoria University, P.O. Box 14428, Melbourne City, MC 8001, Australia*
[2] *School of Computational and Applied Mathematics, University of The Witwatersrand, Private Bag 3, Wits, Johannesburg 2050, South Africa*

Correspondence should be addressed to S. S. Dragomir, sever.dragomir@vu.edu.au

Received 28 September 2011; Accepted 18 December 2011

Academic Editor: Elena Braverman

Some inequalities of the Slater type for convex functions defined on general linear spaces are given. Applications for norm inequalities and f-divergence measures are also provided.

1. Introduction

Suppose that I is an interval of real numbers with interior $\mathring{I}$, and $f : I \to \mathbb{R}$ is a convex function on I. Then f is continuous on $\mathring{I}$ and has finite left and right derivatives at each point of $\mathring{I}$. Moreover, if $x, y \in \mathring{I}$ and $x < y$, then $f'_-(x) \le f'_+(x) \le f'_-(y) \le f'_+(y)$ which shows that both f'_- and f'_+ are nondecreasing functions on $\mathring{I}$. It is also known that a convex function must be differentiable except for at most countably many points.

For a convex function $f : I \to \mathbb{R}$, the subdifferential of f denoted by ∂f is the set of all functions $\varphi : I \to [-\infty, \infty]$ such that $\varphi(\mathring{I}) \subset \mathbb{R}$ and

$$f(x) \ge f(a) + (x - a)\varphi(a), \quad \text{for any } x, a \in I. \tag{1.1}$$

It is also well known that if f is convex on I, then ∂f is nonempty, $f'_-, f'_+ \in \partial f$ and if $\varphi \in \partial f$, then

$$f'_-(x) \le \varphi(x) \le f'_+(x), \quad \text{for any } x \in \mathring{I}. \tag{1.2}$$

In particular, φ is a nondecreasing function.

If f is differentiable and convex on $\mathring{I}$, then $\partial f = \{f'\}$.

The following result is well known in the literature as *the Slater inequality.*

Theorem 1.1 (Slater, 1981, [1]). *If $f : I \to \mathbb{R}$ is a nonincreasing (nondecreasing) convex function, $x_i \in I$, $p_i \ge 0$ with $P_n := \sum_{i=1}^n p_i > 0$ and $\sum_{i=1}^n p_i\varphi(x_i) \ne 0$, where $\varphi \in \partial f$, then*

$$\frac{1}{P_n}\sum_{i=1}^n p_i f(x_i) \le f\left(\frac{\sum_{i=1}^n p_i x_i \varphi(x_i)}{\sum_{i=1}^n p_i \varphi(x_i)}\right). \tag{1.3}$$

As pointed out in [2] (see also [3, p. 64] and [4, p. 208]), the monotonicity assumption for the derivative φ can be replaced with the condition

$$\frac{\sum_{i=1}^n p_i x_i \varphi(x_i)}{\sum_{i=1}^n p_i \varphi(x_i)} \in I, \tag{1.4}$$

which is more general and can hold for suitable points in I and for not necessarily monotonic functions.

For recent works on Slater's inequality, see [5–7].

The main aim of the present paper is to extend Slater's inequality for convex functions defined on general linear spaces. A reverse of the Slater's inequality is also obtained. Natural applications for norm inequalities and f-divergence measures are provided as well.

2. Slater's Inequality for Functions Defined on Linear Spaces

Assume that $f : X \to \mathbb{R}$ is a *convex function* on the real linear space X. Since for any vectors $x, y \in X$ the function $g_{x,y} : \mathbb{R} \to \mathbb{R}$, $g_{x,y}(t) := f(x + ty)$ is convex, it follows that the following limits exist

$$\nabla_{+(-)} f(x)(y) := \lim_{t \to 0+(-)} \frac{f(x+ty) - f(x)}{t}, \tag{2.1}$$

and they are called the *right (left) Gâteaux derivatives* of the function f in the point x over the direction y.

It is obvious that for any $t > 0 > s$ we have

$$\begin{aligned}\frac{f(x+ty)-f(x)}{t} &\geq \nabla_+ f(x)(y) = \inf_{t>0}\left[\frac{f(x+ty)-f(x)}{t}\right] \\ &\geq \sup_{s<0}\left[\frac{f(x+sy)-f(x)}{s}\right] = \nabla_- f(x)(y) \geq \frac{f(x+sy)-f(x)}{s},\end{aligned} \tag{2.2}$$

for any $x, y \in X$ and, in particular,

$$\nabla_- f(u)(u-v) \geq f(u) - f(v) \geq \nabla_+ f(v)(u-v), \tag{2.3}$$

for any $u, v \in X$. We call this *the gradient inequality* for the convex function f. It will be used frequently in the sequel in order to obtain various results related to Slater's inequality.

The following properties are also of importance:

$$\nabla_+ f(x)(-y) = -\nabla_- f(x)(y), \tag{2.4}$$

$$\nabla_{+(-)} f(x)(\alpha y) = \alpha \nabla_{+(-)} f(x)(y), \tag{2.5}$$

for any $x, y \in X$ and $\alpha \geq 0$.

The right Gâteaux derivative is *subadditive* while the left one is *superadditive*, that is,

$$\begin{aligned}\nabla_+ f(x)(y+z) &\leq \nabla_+ f(x)(y) + \nabla_+ f(x)(z), \\ \nabla_- f(x)(y+z) &\geq \nabla_- f(x)(y) + \nabla_- f(x)(z),\end{aligned} \tag{2.6}$$

for any $x, y, z \in X$.

Some natural examples can be provided by the use of normed spaces.

Assume that $(X, \|\cdot\|)$ is a real normed linear space. The function $f : X \to \mathbb{R}$, $f(x) := (1/2)\|x\|^2$ is a convex function which generates *the superior* and *the inferior semi-inner products*

$$\langle y, x \rangle_{s(i)} := \lim_{t \to 0+(-)} \frac{\|x+ty\|^2 - \|x\|^2}{2t}. \tag{2.7}$$

For a comprehensive study of the properties of these mappings in the Geometry of Banach Spaces, see the monograph [8].

For the convex function $f_p : X \to \mathbb{R}$, $f_p(x) := \|x\|^p$ with $p > 1$, we have

$$\nabla_{+(-)} f_p(x)(y) = \begin{cases} p\|x\|^{p-2}\langle y, x\rangle_{s(i)} & \text{if } x \neq 0, \\ 0 & \text{if } x = 0, \end{cases} \tag{2.8}$$

for any $y \in X$.

If $p = 1$, then we have

$$\nabla_{+(-)} f_1(x)(y) = \begin{cases} \|x\|^{-1}\langle y, x\rangle_{s(i)} & \text{if } x \neq 0, \\ +(-)\|y\| & \text{if } x = 0, \end{cases} \tag{2.9}$$

for any $y \in X$.

For a given convex function $f : X \to \mathbb{R}$ and a given n-tuple of vectors $\mathbf{x} = (x_1, \dots, x_n) \in X^n$, we consider the sets

$$\begin{aligned} \mathrm{Sla}_{+(-)}(f, \mathbf{x}) &:= \{v \in X \mid \nabla_{+(-)} f(x_i)(v - x_i) \geq 0 \ \forall i \in \{1, \dots, n\}\}, \\ \mathrm{Sla}_{+(-)}(f, \mathbf{x}, \mathbf{p}) &:= \left\{v \in X \mid \sum_{i=1}^{n} p_i \nabla_{+(-)} f(x_i)(v - x_i) \geq 0\right\}, \end{aligned} \tag{2.10}$$

where $\mathbf{p} = (p_1, \dots, p_n) \in \mathbb{P}^n$ is a given probability distribution, that is, $p_i \geq 0$ for $i \in \{1, \dots, n\}$ and $\sum_{i=1}^{n} p_i = 1$.

The following properties of these sets hold.

Lemma 2.1. *For a given convex function $f : X \to \mathbb{R}$, a given n-tuple of vectors $\mathbf{x} = (x_1, \dots, x_n) \in X^n$, and a given probability distribution $\mathbf{p} = (p_1, \dots, p_n) \in \mathbb{P}^n$, one has*

(i) $\mathrm{Sla}_{-}(f, \mathbf{x}) \subset \mathrm{Sla}_{+}(f, \mathbf{x})$ *and* $\mathrm{Sla}_{-}(f, \mathbf{x}, \mathbf{p}) \subset \mathrm{Sla}_{+}(f, \mathbf{x}, \mathbf{p})$;

(ii) $\mathrm{Sla}_{-}(f, \mathbf{x}) \subset \mathrm{Sla}_{-}(f, \mathbf{x}, \mathbf{p})$ *and* $\mathrm{Sla}_{+}(f, \mathbf{x}) \subset \mathrm{Sla}_{+}(f, \mathbf{x}, \mathbf{p})$ *for all* $\mathbf{p} = (p_1, \dots, p_n) \in \mathbb{P}^n$;

(iii) *the sets* $\mathrm{Sla}_{-}(f, \mathbf{x})$ *and* $\mathrm{Sla}_{-}(f, \mathbf{x}, \mathbf{p})$ *are convex.*

Proof. The properties (i) and (ii) follow from the definition and the fact that $\nabla_{+} f(x)(y) \geq \nabla_{-} f(x)(y)$ for any x, y.

(iii) Let us only prove that $\mathrm{Sla}_{-}(f, \mathbf{x})$ is convex.

If we assume that $y_1, y_2 \in \mathrm{Sla}_{-}(f, \mathbf{x})$ and $\alpha, \beta \in [0, 1]$ with $\alpha + \beta = 1$, then by the superadditivity and positive homogeneity of the Gâteaux derivative $\nabla_{-} f(\cdot)(\cdot)$ in the second variable we have

$$\begin{aligned} \nabla_{-} f(x_i)(\alpha y_1 + \beta y_2 - x_i) &= \nabla_{-} f(x_i)[\alpha(y_1 - x_i) + \beta(y_2 - x_i)] \\ &\geq \alpha \nabla_{-} f(x_i)(y_1 - x_i) + \beta \nabla_{-} f(x_i)(y_2 - x_i) \geq 0, \end{aligned} \tag{2.11}$$

for all $i \in \{1, \dots, n\}$, which shows that $\alpha y_1 + \beta y_2 \in \mathrm{Sla}_{-}(f, \mathbf{x})$

The proof for the convexity of $\mathrm{Sla}_{-}(f, \mathbf{x}, \mathbf{p})$ is similar and the details are omitted. □

For the convex function $f_p : X \to \mathbb{R}$, $f_p(x) := \|x\|^p$ with $p \geq 1$, defined on the normed linear space $(X, \|\cdot\|)$ and for the n-tuple of vectors $\mathbf{x} = (x_1, \dots, x_n) \in X^n \setminus \{(0, \dots, 0)\}$ we have, by the well-known property of the semi-inner products,

$$\langle y + \alpha x, x\rangle_{s(i)} = \langle y, x\rangle_{s(i)} + \alpha\|x\|^2, \quad \text{for any } x, y \in X, \ \alpha \in \mathbb{R}, \tag{2.12}$$

that

$$\mathrm{Sla}_{+(-)}\left(\|\cdot\|^{p},\mathbf{x}\right)=\mathrm{Sla}_{+(-)}\left(\|\cdot\|,\mathbf{x}\right):=\left\{v\in X\mid\left\langle v,x_j\right\rangle_{s(i)}\geq\left\|x_j\right\|^{2}\ \forall j\in\{1,\dots,n\}\right\} \tag{2.13}$$

which, as can be seen, does not depend on p. We observe, by the continuity of the semi-inner products in the first variable, that $\mathrm{Sla}_{+(-)}(\|\cdot\|,\mathbf{x})$ is closed in $(X,\|\cdot\|)$. Also, we should remark that if $v\in\mathrm{Sla}_{+(-)}(\|\cdot\|,\mathbf{x})$, then for any $\gamma\geq 1$ we also have that $\gamma v\in\mathrm{Sla}_{+(-)}(\|\cdot\|,\mathbf{x})$.

The larger classes, which are dependent on the probability distribution $\mathbf{p}\in\mathbb{P}^n$, are described by

$$\mathrm{Sla}_{+(-)}\left(\|\cdot\|^{p},\mathbf{x},\mathbf{p}\right):=\left\{v\in X\mid\sum_{j=1}^{n}p_j\left\|x_j\right\|^{p-2}\left\langle v,x_j\right\rangle_{s(i)}\geq\sum_{j=1}^{n}p_j\left\|x_j\right\|^{p}\right\}. \tag{2.14}$$

If the normed space is smooth, that is, the norm is Gâteaux differentiable in any nonzero point, then the superior and inferior semi-inner products coincide with the Lumer-Giles semi-inner product $[\cdot,\cdot]$ that generates the norm and is linear in the first variable (see for instance [8]). In this situation,

$$\begin{gathered}\mathrm{Sla}\left(\|\cdot\|,\mathbf{x}\right)=\left\{v\in X\mid\left[v,x_j\right]\geq\left\|x_j\right\|^{2}\ \forall j\in\{1,\dots,n\}\right\},\\ \mathrm{Sla}\left(\|\cdot\|^{p},\mathbf{x},\mathbf{p}\right)=\left\{v\in X\mid\sum_{j=1}^{n}p_j\left\|x_j\right\|^{p-2}\left[v,x_j\right]\geq\sum_{j=1}^{n}p_j\left\|x_j\right\|^{p}\right\}.\end{gathered} \tag{2.15}$$

If $(X,\langle\cdot,\cdot\rangle)$ is an inner product space, then $\mathrm{Sla}(\|\cdot\|^{p},\mathbf{x},\mathbf{p})$ can be described by

$$\mathrm{Sla}\left(\|\cdot\|^{p},\mathbf{x},\mathbf{p}\right)=\left\{v\in X\mid\left\langle v,\sum_{j=1}^{n}p_j\left\|x_j\right\|^{p-2}x_j\right\rangle\geq\sum_{j=1}^{n}p_j\left\|x_j\right\|^{p}\right\}, \tag{2.16}$$

and if the family $\{x_j\}_{j=1,\dots,n}$ is orthogonal, then obviously, by the Pythagoras theorem, we have that the sum $\sum_{j=1}^{n}x_j$ belongs to $\mathrm{Sla}(\|\cdot\|,\mathbf{x})$ and therefore to $\mathrm{Sla}(\|\cdot\|^{p},\mathbf{x},\mathbf{p})$ for any $p\geq 1$ and any probability distribution $\mathbf{p}=(p_1,\dots,p_n)\in\mathbb{P}^n$.

We can state now the following results that provide a generalization of Slater's inequality as well as a counterpart for it.

Theorem 2.2. *Let $f:X\to\mathbb{R}$ be a convex function on the real linear space X, $\mathbf{x}=(x_1,\dots,x_n)\in X^n$ an n-tuple of vectors, and $\mathbf{p}=(p_1,\dots,p_n)\in\mathbb{P}^n$ a probability distribution. Then for any $v\in\mathrm{Sla}_{+}(f,\mathbf{x},\mathbf{p})$, one has the inequalities*

$$\nabla_{-}f(v)(v)-\sum_{i=1}^{n}p_i\nabla_{-}f(v)(x_i)\geq f(v)-\sum_{i=1}^{n}p_if(x_i)\geq 0. \tag{2.17}$$

Proof. If we write the gradient inequality for $v \in \mathrm{Sla}_+(f, \mathbf{x}, \mathbf{p})$ and x_i, then we have that

$$\nabla_- f(v)(v - x_i) \geq f(v) - f(x_i) \geq \nabla_+ f(x_i)(v - x_i), \tag{2.18}$$

for any $i \in \{1, \ldots, n\}$.

By multiplying (2.18) with $p_i \geq 0$ and summing over i from 1 to n, we get

$$\sum_{i=1}^{n} p_i \nabla_- f(v)(v - x_i) \geq f(v) - \sum_{i=1}^{n} p_i f(x_i) \geq \sum_{i=1}^{n} p_i \nabla_+ f(x_i)(v - x_i). \tag{2.19}$$

Now, since $v \in \mathrm{Sla}_+(f, \mathbf{x}, \mathbf{p})$, then the right hand side of (2.19) is nonnegative, which proves the second inequality in (2.17).

By the superadditivity of the Gâteaux derivative $\nabla_- f(\cdot)(\cdot)$ in the second variable, we have

$$\nabla_- f(v)(v) - \nabla_- f(v)(x_i) \geq \nabla_- f(v)(v - x_i), \tag{2.20}$$

which, by multiplying with $p_i \geq 0$ and summing over i from 1 to n, produces the inequality

$$\nabla_- f(v)(v) - \sum_{i=1}^{n} p_i \nabla_- f(v)(x_i) \geq \sum_{i=1}^{n} p_i \nabla_- f(v)(v - x_i). \tag{2.21}$$

Utilising (2.19) and (2.21), we deduce the desired result (2.17). □

Remark 2.3. The above result has the following form for normed linear spaces. Let $(X, \|\cdot\|)$ be a normed linear space, $\mathbf{x} = (x_1, \ldots, x_n) \in X^n$ an n-tuple of vectors from X, and $p = (p_1, \ldots, p_n) \in \mathbb{P}^n$ a probability distribution. Then for any vector $v \in X$ with the property

$$\sum_{j=1}^{n} p_j \|x_j\|^{p-2} \langle v, x_j \rangle_s \geq \sum_{j=1}^{n} p_j \|x_j\|^p, \quad p \geq 1, \tag{2.22}$$

we have the inequalities

$$p\left[\|v\|^p - \sum_{j=1}^{n} p_j \|x_j\|^{p-2} \langle v, x_j \rangle_i\right] \geq \|v\|^p - \sum_{j=1}^{n} p_j \|x_j\|^p \geq 0. \tag{2.23}$$

Rearranging the first inequality in (2.23), we also have that

$$(p-1)\|v\|^p + \sum_{j=1}^{n} p_j \|x_j\|^p \geq p \sum_{j=1}^{n} p_j \|x_j\|^{p-2} \langle v, x_j \rangle_i. \tag{2.24}$$

If the space is smooth, then the condition (2.22) becomes

$$\sum_{j=1}^{n} p_j \|x_j\|^{p-2}[v, x_j] \geq \sum_{j=1}^{n} p_j \|x_j\|^{p}, \quad p \geq 1, \tag{2.25}$$

implying the inequality

$$p\left[\|v\|^p - \sum_{j=1}^{n} p_j \|x_j\|^{p-2}[v, x_j]\right] \geq \|v\|^p - \sum_{j=1}^{n} p_j \|x_j\|^{p} \geq 0. \tag{2.26}$$

Notice also that the first inequality in (2.26) is equivalent with

$$(p-1)\|v\|^p + \sum_{j=1}^{n} p_j \|x_j\|^{p} \geq p\sum_{j=1}^{n} p_j \|x_j\|^{p-2}[v, x_j] \left(\geq p\sum_{j=1}^{n} p_j \|x_j\|^{p} \geq 0 \right). \tag{2.27}$$

Corollary 2.4. *Let $f : X \to \mathbb{R}$ be a convex function on the real linear space X, $x = (x_1, \ldots, x_n) \in X^n$ an n-tuple of vectors, and $p = (p_1, \ldots, p_n) \in \mathbb{P}^n$ a probability distribution. If*

$$\sum_{i=1}^{n} p_i \nabla_+ f(x_i)(x_i) \geq (<)0, \tag{2.28}$$

and there exists a vector $s \in X$ with

$$\sum_{i=1}^{n} p_i \nabla_{+(-)} f(x_i)(s) \geq (\leq)1, \tag{2.29}$$

then

$$\begin{aligned}\nabla_- f\left(\sum_{j=1}^{n} p_j \nabla_+ f(x_j)(x_j)s\right)\left(\sum_{j=1}^{n} p_j \nabla_+ f(x_j)(x_j)s\right) - \sum_{i=1}^{n} p_i \nabla_- f\left(\sum_{j=1}^{n} p_j \nabla_+ f(x_j)(x_j)s\right)(x_i) \\ \geq f\left(\sum_{j=1}^{n} p_j \nabla_+ f(x_j)(x_j)s\right) - \sum_{i=1}^{n} p_i f(x_i) \geq 0.\end{aligned} \tag{2.30}$$

Proof. Assume that $\sum_{i=1}^{n} p_i \nabla_+ f(x_i)(x_i) \geq 0$ and $\sum_{i=1}^{n} p_i \nabla_+ f(x_i)(s) \geq 1$ and define $v := \sum_{j=1}^{n} p_j \nabla_+ f(x_j)(x_j)s$. We claim that $v \in \mathrm{Sla}_+(f, \mathbf{x}, \mathbf{p})$.

By the subadditivity and positive homogeneity of the mapping $\nabla_+ f(\cdot)(\cdot)$ in the second variable, we have

$$\begin{aligned}
\sum_{i=1}^{n} p_i \nabla_+ f(x_i)(v - x_i) &\geq \sum_{i=1}^{n} p_i \nabla_+ f(x_i)(v) - \sum_{i=1}^{n} p_i \nabla_+ f(x_i)(x_i) \\
&= \sum_{i=1}^{n} p_i \nabla_+ f(x_i)\left(\sum_{j=1}^{n} p_j \nabla_+ f(x_j)(x_j) s \right) - \sum_{i=1}^{n} p_i \nabla_+ f(x_i)(x_i) \\
&= \sum_{j=1}^{n} p_j \nabla_+ f(x_j)(x_j) \sum_{i=1}^{n} p_i \nabla_+ f(x_i)(s) - \sum_{i=1}^{n} p_i \nabla_+ f(x_i)(x_i) \\
&= \sum_{j=1}^{n} p_j \nabla_+ f(x_j)(x_j) \left[\sum_{i=1}^{n} p_i \nabla_+ f(x_i)(s) - 1 \right] \geq 0,
\end{aligned} \tag{2.31}$$

as claimed. Applying Theorem 2.2 for this v, we get the desired result.

If $\sum_{i=1}^{n} p_i \nabla_+ f(x_i)(x_i) < 0$ and $\sum_{i=1}^{n} p_i \nabla_- f(x_i)(s) \leq 1$, then for

$$w := \sum_{j=1}^{n} p_j \nabla_+ f(x_j)(x_j) s, \tag{2.32}$$

we also have that

$$\begin{aligned}
\sum_{i=1}^{n} p_i \nabla_+ f(x_i)(w - x_i) &\geq \sum_{i=1}^{n} p_i \nabla_+ f(x_i)\left(\sum_{j=1}^{n} p_j \nabla_+ f(x_j)(x_j) s \right) - \sum_{i=1}^{n} p_i \nabla_+ f(x_i)(x_i) \\
&= \sum_{i=1}^{n} p_i \nabla_+ f(x_i)\left(\left(-\sum_{j=1}^{n} p_j \nabla_+ f(x_j)(x_j) \right)(-s) \right) - \sum_{i=1}^{n} p_i \nabla_+ f(x_i)(x_i) \\
&= \left(-\sum_{j=1}^{n} p_j \nabla_+ f(x_j)(x_j) \right) \sum_{i=1}^{n} p_i \nabla_+ f(x_i)(-s) - \sum_{i=1}^{n} p_i \nabla_+ f(x_i)(x_i) \\
&= \left(-\sum_{j=1}^{n} p_j \nabla_+ f(x_j)(x_j) \right) \left(1 + \sum_{i=1}^{n} p_i \nabla_+ f(x_i)(-s) \right) \\
&= \left(-\sum_{j=1}^{n} p_j \nabla_+ f(x_j)(x_j) \right) \left(1 - \sum_{i=1}^{n} p_i \nabla_- f(x_i)(s) \right) \geq 0,
\end{aligned} \tag{2.33}$$

where, for the last equality, we have used the property (2.4). Therefore, $w \in \mathrm{Sla}_+(f, \mathbf{x}, \mathbf{p})$ and by Theorem 2.2 we get the desired result. □

It is natural to consider the case of normed spaces.

Remark 2.5. Let $(X, \|\cdot\|)$ be a normed linear space, $\mathbf{x} = (x_1, \ldots, x_n) \in X^n$ an n-tuple of vectors from X, and $\mathbf{p} = (p_1, \ldots, p_n) \in \mathbb{P}^n$ a probability distribution. Then for any vector $s \in X$ with the property that

$$p\sum_{i=1}^{n} p_i \|x_i\|^{p-2} \langle s, x_i \rangle_s \geq 1, \tag{2.34}$$

we have the inequalities

$$\begin{aligned} p^p \|s\|^{p-1} & \left(\sum_{j=1}^{n} p_j \|x_j\|^p \right)^{p-1} \left(p\|s\| \sum_{j=1}^{n} p_j \|x_j\|^p - \sum_{j=1}^{n} p_j \langle x_j, s \rangle_i \right) \\ & \geq p^p \|s\|^p \left(\sum_{j=1}^{n} p_j \|x_j\|^p \right)^p - \sum_{j=1}^{n} p_j \|x_j\|^p \geq 0. \end{aligned} \tag{2.35}$$

The case of smooth spaces can be easily derived from the above; however, the details are left to the interested reader.

3. The Case of Finite Dimensional Linear Spaces

Consider now the finite dimensional linear space $X = \mathbb{R}^m$ and assume that C is an open convex subset of $\mathbb{R}^m$. Assume also that the function $f : C \to \mathbb{R}$ is differentiable and convex on C. Obviously, if $x = (x^1, \ldots, x^m) \in C$, then for any $y = (y^1, \ldots, y^m) \in \mathbb{R}^m$ we have

$$\nabla f(x)(y) = \sum_{k=1}^{m} \frac{\partial f(x^1, \ldots, x^m)}{\partial x^k} \cdot y^k. \tag{3.1}$$

For the convex function $f : C \to \mathbb{R}$ and a given n-tuple of vectors $\mathbf{x} = (x_1, \ldots, x_n) \in C^n$ with $x_i = (x_i^1, \ldots, x_i^m)$ with $i \in \{1, \ldots, n\}$, we consider the sets

$$\begin{aligned} \mathrm{Sla}(f, \mathbf{x}, C) &:= \left\{ v \in C \mid \sum_{k=1}^{m} \frac{\partial f(x_i^1, \ldots, x_i^m)}{\partial x_i^k} \cdot v^k \right. \\ & \qquad \left. \geq \sum_{k=1}^{m} \frac{\partial f(x_i^1, \ldots, x_i^m)}{\partial x_i^k} \cdot x_i^k \ \forall i \in \{1, \ldots, n\} \right\}, \\ \mathrm{Sla}(f, \mathbf{x}, \mathbf{p}, C) &:= \left\{ v \in C \mid \sum_{i=1}^{n} \sum_{k=1}^{m} p_i \frac{\partial f(x_i^1, \ldots, x_i^m)}{\partial x_i^k} \cdot v^k \right. \\ & \qquad \left. \geq \sum_{i=1}^{n} \sum_{k=1}^{m} p_i \frac{\partial f(x_i^1, \ldots, x_i^m)}{\partial x_i^k} \cdot x_i^k \right\}, \end{aligned} \tag{3.2}$$

where $\mathbf{p} = (p_1, \ldots, p_n) \in \mathbb{P}^n$ is a given probability distribution.

As in the previous section the sets, $\mathrm{Sla}(f, \mathbf{x}, C)$ and $\mathrm{Sla}(f, \mathbf{x}, \mathbf{p}, C)$ are convex and closed subsets of $\mathrm{clo}(C)$, the closure of C. Also $\{x_1, \ldots, x_n\} \subset \mathrm{Sla}(f, \mathbf{x}, C) \subset \mathrm{Sla}(f, \mathbf{x}, \mathbf{p}, C)$ for any $\mathbf{p} = (p_1, \ldots, p_n) \in \mathbb{P}^n$ is a probability distribution.

Proposition 3.1. *Let $f : C \to \mathbb{R}$ be a convex function on the open convex set C in the finite dimensional linear space $\mathbb{R}^m$, $(x_1, \ldots, x_n) \in C^n$ an n-tuple of vectors and $(p_1, \ldots, p_n) \in \mathbb{P}^n$ a probability distribution. Then for any $v = (v^1, \ldots, v^m) \in \mathrm{Sla}(f, \mathbf{x}, \mathbf{p}, C)$, one has the inequalities*

$$\begin{aligned} &\sum_{k=1}^{m} \frac{\partial f(v^1, \ldots, v^m)}{\partial x^k} \cdot v^k - \sum_{i=1}^{n} \sum_{k=1}^{m} p_i \frac{\partial f(x_i^1, \ldots, x_i^m)}{\partial x_i^k} \cdot v^k \\ &\quad \geq f\left(v^1, \ldots, v^n\right) - \sum_{i=1}^{n} p_i f\left(x_i^1, \ldots, x_i^m\right) \geq 0. \end{aligned} \tag{3.3}$$

The unidimensional case, that is, $m = 1$ is of interest for applications. We will state this case with the general assumption that $f : I \to \mathbb{R}$ is a convex function on an *open* interval I. For a given n-tuple of vectors $\mathbf{x} = (x_1, \ldots, x_n) \in I^n$, we have

$$\begin{aligned} \mathrm{Sla}_{+(-)}(f, \mathbf{x}, I) &:= \left\{ v \in I \mid f'_{+(-)}(x_i) \cdot (v - x_i) \geq 0 \ \forall i \in \{1, \ldots, n\} \right\}, \\ \mathrm{Sla}_{+(-)}(f, \mathbf{x}, \mathbf{p}, \mathbf{I}) &:= \left\{ v \in I \mid \sum_{i=1}^{n} p_i f'_{+(-)}(x_i) \cdot (v - x_i) \geq 0 \right\}, \end{aligned} \tag{3.4}$$

where $(p_1, \ldots, p_n) \in \mathbb{P}^n$ is a probability distribution. These sets inherit the general properties pointed out in Lemma 2.1. Moreover, if we make the assumption that $\sum_{i=1}^{n} p_i f'_+(x_i) \neq 0$, then for $\sum_{i=1}^{n} p_i f'_+(x_i) > 0$ we have

$$\mathrm{Sla}_{+}(f, \mathbf{x}, \mathbf{p}, \mathbf{I}) = \left\{ v \in I \mid v \geq \frac{\sum_{i=1}^{n} p_i f'_+(x_i) x_i}{\sum_{i=1}^{n} p_i f'_+(x_i)} \right\}, \tag{3.5}$$

while for $\sum_{i=1}^{n} p_i f'_+(x_i) < 0$ we have

$$\mathrm{Sla}_{+}(f, \mathbf{x}, \mathbf{p}, \mathbf{I}) = \left\{ v \in I \mid v \leq \frac{\sum_{i=1}^{n} p_i f'_+(x_i) x_i}{\sum_{i=1}^{n} p_i f'_+(x_i)} \right\}. \tag{3.6}$$

Also, if we assume that $f'_+(x_i) \geq 0$ for all $i \in \{1, \ldots, n\}$ and $\sum_{i=1}^{n} p_i f'_+(x_i) > 0$, then

$$v_s := \frac{\sum_{i=1}^{n} p_i f'_+(x_i) x_i}{\sum_{i=1}^{n} p_i f'_+(x_i)} \in I, \tag{3.7}$$

due to the fact that $x_i \in I$ and I is a convex set.

Proposition 3.2. *Let $f : I \to \mathbb{R}$ be a convex function on an open interval I. For a given n-tuple of vectors $\mathbf{x} = (x_1, \ldots, x_n) \in I^n$ and $(p_1, \ldots, p_n) \in \mathbb{P}^n$ a probability distribution, one has*

$$f'_-(v)\left(v - \sum_{i=1}^n p_i x_i\right) \geq f(v) - \sum_{i=1}^n p_i f(x_i) \geq 0, \tag{3.8}$$

for any $v \in \mathrm{Sla}_+(f, \mathbf{x}, \mathbf{p}, \mathbf{I})$.

In particular, if one assumes that $\sum_{i=1}^n p_i f'_+(x_i) \neq 0$ and

$$\frac{\sum_{i=1}^n p_i f'_+(x_i) x_i}{\sum_{i=1}^n p_i f'_+(x_i)} \in I, \tag{3.9}$$

then

$$\begin{aligned} &f'_-\left(\frac{\sum_{i=1}^n p_i f'_+(x_i) x_i}{\sum_{i=1}^n p_i f'_+(x_i)}\right)\left[\frac{\sum_{i=1}^n p_i f'_+(x_i) x_i}{\sum_{i=1}^n p_i f'_+(x_i)} - \sum_{i=1}^n p_i x_i\right] \\ &\quad \geq f\left(\frac{\sum_{i=1}^n p_i f'_+(x_i) x_i}{\sum_{i=1}^n p_i f'_+(x_i)}\right) - \sum_{i=1}^n p_i f(x_i) \geq 0. \end{aligned} \tag{3.10}$$

Moreover, if $f'_+(x_i) \geq 0$ for all $i \in \{1, \ldots, n\}$ and $\sum_{i=1}^n p_i f'_+(x_i) > 0$, then (3.10) *holds true as well.*

Remark 3.3. We remark that the first inequality in (3.10) *provides a reverse inequality for the classical result due to Slater.*

4. Some Applications for f-Divergences

Given a convex function $f : [0, \infty) \to \mathbb{R}$, the *$f$-divergence functional*

$$I_f(\mathbf{p}, \mathbf{q}) := \sum_{i=1}^n q_i f\left(\frac{p_i}{q_i}\right), \tag{4.1}$$

where $\mathbf{p} = (p_1, \ldots, p_n)$, $\mathbf{q} = (q_1, \ldots, q_n)$ are positive sequences, was introduced by Csiszár in [9], as a generalized measure of information, a "distance function" on the set of probability distributions $\mathbb{P}^n$. As in [9], we interpret undefined expressions by

$$\begin{gathered} f(0) = \lim_{t \to 0+} f(t), \qquad 0f\left(\frac{0}{0}\right) = 0, \\ 0f\left(\frac{a}{0}\right) = \lim_{q \to 0+} q f\left(\frac{a}{q}\right) = a \lim_{t \to \infty} \frac{f(t)}{t}, \qquad a > 0. \end{gathered} \tag{4.2}$$

The following results were essentially given by Csiszár and Körner [10]:

(i) *if* f *is convex, then* $I_f(\mathbf{p},\mathbf{q})$ *is jointly convex in* $\mathbf{p}$ *and* $\mathbf{q}$;

(ii) *for every* $\mathbf{p},\mathbf{q}\in\mathbb{R}_+^n$, *we have*

$$I_f(\mathbf{p},\mathbf{q}) \geq \sum_{j=1}^{n} q_j f\left(\frac{\sum_{j=1}^{n} p_j}{\sum_{j=1}^{n} q_j}\right). \tag{4.3}$$

If f *is strictly convex, equality holds in* (4.3) *if and only if*

$$\frac{p_1}{q_1} = \frac{p_2}{q_2} = \cdots = \frac{p_n}{q_n}. \tag{4.4}$$

If f *is normalized, that is,* $f(1)=0$, *then for every* $\mathbf{p},\mathbf{q}\in\mathbb{R}_+^n$ *with* $\sum_{i=1}^{n} p_i = \sum_{i=1}^{n} q_i$, *we have the inequality*

$$I_f(\mathbf{p},\mathbf{q}) \geq 0. \tag{4.5}$$

In particular, if $\mathbf{p},\mathbf{q}\in\mathbb{P}^n$, then (4.5) holds. This is the well-known positivity property of the f-divergence.

It is obvious that the above definition of $I_f(\mathbf{p},\mathbf{q})$ can be extended to any function $f:[0,\infty)\to\mathbb{R}$; however, the positivity condition will not generally hold for normalized functions and $\mathbf{p},\mathbf{q}\in\mathbb{R}_+^n$ *with* $\sum_{i=1}^{n} p_i = \sum_{i=1}^{n} q_i$.

For a normalized convex function $f:[0,\infty)\to\mathbb{R}$ and two probability distributions $\mathbf{p},\mathbf{q}\in\mathbb{P}^n$, we define the set

$$\mathrm{Sla}_+(f,\mathbf{p},\mathbf{q}) := \left\{ v\in[0,\infty) \mid \sum_{i=1}^{n} q_i f'_+\left(\frac{p_i}{q_i}\right)\cdot\left(v-\frac{p_i}{q_i}\right) \geq 0 \right\}. \tag{4.6}$$

Now, observe that

$$\sum_{i=1}^{n} q_i f'_+\left(\frac{p_i}{q_i}\right)\cdot\left(v-\frac{p_i}{q_i}\right) \geq 0, \tag{4.7}$$

is equivalent with

$$v\sum_{i=1}^{n} q_i f'_+\left(\frac{p_i}{q_i}\right) \geq \sum_{i=1}^{n} p_i f'_+\left(\frac{p_i}{q_i}\right). \tag{4.8}$$

If $\sum_{i=1}^{n} q_i f'_+(p_i/q_i) > 0$, then (4.8) is equivalent with

$$v \geq \frac{\sum_{i=1}^{n} p_i f'_+(p_i/q_i)}{\sum_{i=1}^{n} q_i f'_+(p_i/q_i)}, \tag{4.9}$$

therefore in this case

$$\mathrm{Sla}_{+}(f,\mathbf{p},\mathbf{q}) = \begin{cases} [0,\infty) & \text{if } \sum_{i=1}^{n} p_i f'_{+}(p_i/q_i) < 0, \\ \left[\dfrac{\sum_{i=1}^{n} p_i f'_{+}(p_i/q_i)}{\sum_{i=1}^{n} q_i f'_{+}(p_i/q_i)}, \infty\right) & \text{if } \sum_{i=1}^{n} p_i f'_{+}(p_i/q_i) \geq 0. \end{cases} \tag{4.10}$$

If $\sum_{i=1}^{n} q_i f'_{+}(p_i/q_i) < 0$, then (4.8) is equivalent with

$$v \leq \frac{\sum_{i=1}^{n} p_i f'_{+}(p_i/q_i)}{\sum_{i=1}^{n} q_i f'_{+}(p_i/q_i)}, \tag{4.11}$$

therefore

$$\mathrm{Sla}_{+}(f,\mathbf{p},\mathbf{q}) = \begin{cases} \left[0, \dfrac{\sum_{i=1}^{n} p_i f'_{+}(p_i/q_i)}{\sum_{i=1}^{n} q_i f'_{+}(p_i/q_i)}\right] & \text{if } \sum_{i=1}^{n} p_i f'_{+}\left(\dfrac{p_i}{q_i}\right) \leq 0, \\ \emptyset & \text{if } \sum_{i=1}^{n} p_i f'_{+}\left(\dfrac{p_i}{q_i}\right) > 0. \end{cases} \tag{4.12}$$

Utilising the extended f-divergences notation, we can state the following result.

Theorem 4.1. *Let $f : [0,\infty) \to \mathbb{R}$ be a normalized convex function and $\mathbf{p},\mathbf{q} \in \mathbb{P}^n$ two probability distributions. If $v \in \mathrm{Sla}_{+}(f,\mathbf{p},\mathbf{q})$, then one has*

$$f'_{-}(v)(v-1) \geq f(v) - I_f(\mathbf{p},\mathbf{q}) \geq 0. \tag{4.13}$$

In particular, if one assumes that $I_{f'_{+}}(\mathbf{p},\mathbf{q}) \neq 0$ and

$$\frac{I_{f'_{+}(\cdot)(\cdot)}(\mathbf{p},\mathbf{q})}{I_{f'_{+}}(\mathbf{p},\mathbf{q})} \in [0,\infty), \tag{4.14}$$

then

$$f'_{-}\left(\frac{I_{f'_{+}(\cdot)(\cdot)}(\mathbf{p},\mathbf{q})}{I_{f'_{+}}(\mathbf{p},\mathbf{q})}\right)\left[\frac{I_{f'_{+}(\cdot)(\cdot)}(\mathbf{p},\mathbf{q})}{I_{f'_{+}}(\mathbf{p},\mathbf{q})} - 1\right] \geq f\left(\frac{I_{f'_{+}(\cdot)(\cdot)}(\mathbf{p},\mathbf{q})}{I_{f'_{+}}(\mathbf{p},\mathbf{q})}\right) - I_f(\mathbf{p},\mathbf{q}) \geq 0. \tag{4.15}$$

Moreover, if $f'_{+}(p_i/q_i) \geq 0$ for all $i \in \{1,\ldots,n\}$ and $I_{f'_{+}}(\mathbf{p},\mathbf{q}) > 0$, then (4.15) *holds true as well.*

The proof follows immediately from Proposition 3.2 and the details are omitted.

The K. Pearson χ^2-*divergence* is obtained for the convex function $f(t) = (1-t)^2$, $t \in \mathbb{R}$ and given by

$$\chi^2(\mathbf{p},\mathbf{q}) := \sum_{j=1}^{n} q_j \left(\frac{p_j}{q_j} - 1\right)^2 = \sum_{j=1}^{n} \frac{(p_j - q_j)^2}{q_j} = \sum_{j=1}^{n} \frac{p_i^2}{q_i} - 1. \tag{4.16}$$

The *Kullback-Leibler divergence* can be obtained for the convex function $f : (0,\infty) \to \mathbb{R}$, $f(t) = t\ln t$ and is defined by

$$\mathrm{KL}(\mathbf{p},\mathbf{q}) := \sum_{j=1}^{n} q_j \cdot \frac{p_j}{q_j} \ln\left(\frac{p_j}{q_j}\right) = \sum_{j=1}^{n} p_j \ln\left(\frac{p_j}{q_j}\right). \tag{4.17}$$

If we consider the convex function $f : (0,\infty) \to \mathbb{R}$, $f(t) = -\ln t$, then we observe that

$$I_f(\mathbf{p},\mathbf{q}) := \sum_{i=1}^{n} q_i f\left(\frac{p_i}{q_i}\right) = -\sum_{i=1}^{n} q_i \ln\left(\frac{p_i}{q_i}\right) = \sum_{i=1}^{n} q_i \ln\left(\frac{q_i}{p_i}\right) = \mathrm{KL}(\mathbf{q},\mathbf{p}). \tag{4.18}$$

For the function $f(t) = -\ln t$, we will obviously have that

$$\begin{aligned}
\mathrm{Sla}(-\ln,\mathbf{p},\mathbf{q}) &:= \left\{ v \in [0,\infty) \mid -\sum_{i=1}^{n} q_i \left(\frac{p_i}{q_i}\right)^{-1} \cdot \left(v - \frac{p_i}{q_i}\right) \geq 0 \right\} \\
&= \left\{ v \in [0,\infty) \mid v \sum_{i=1}^{n} \frac{q_i^2}{p_i} - 1 \leq 0 \right\} \\
&= \left[0, \frac{1}{\chi^2(\mathbf{q},\mathbf{p}) + 1}\right].
\end{aligned} \tag{4.19}$$

Utilising the first part of Theorem 4.1, we can state the following.

Proposition 4.2. *Let $\mathbf{p},\mathbf{q} \in \mathbb{P}^n$ be two probability distributions. If $v \in [0, (1/(\chi^2(\mathbf{q},\mathbf{p}) + 1))]$, then one has*

$$\frac{1-v}{v} \geq -\ln(v) - \mathrm{KL}(\mathbf{q},\mathbf{p}) \geq 0. \tag{4.20}$$

In particular, for $v = 1/(\chi^2(\mathbf{q},\mathbf{p}) + 1)$, one gets

$$\chi^2(\mathbf{q},\mathbf{p}) \geq \ln\left[\chi^2(\mathbf{q},\mathbf{p}) + 1\right] - \mathrm{KL}(\mathbf{q},\mathbf{p}) \geq 0. \tag{4.21}$$

If we consider now the function $f : (0,\infty) \to \mathbb{R}$, $f(t) = t\ln t$, then $f'(t) = \ln t + 1$ and

$$\begin{aligned}
\operatorname{Sla}((\cdot)\ln(\cdot),\mathbf{p},\mathbf{q}) &:= \left\{ v \in [0,\infty) \mid \sum_{i=1}^{n} q_i \left(\ln\left(\frac{p_i}{q_i}\right) + 1 \right) \cdot \left(v - \frac{p_i}{q_i} \right) \geq 0 \right\} \\
&= \left\{ v \in [0,\infty) \mid v \sum_{i=1}^{n} q_i \left(\ln\left(\frac{p_i}{q_i}\right) + 1 \right) - \sum_{i=1}^{n} p_i \cdot \left(\ln\left(\frac{p_i}{q_i}\right) + 1 \right) \geq 0 \right\} \\
&= \{ v \in [0,\infty) \mid v(1 - \mathrm{KL}(\mathbf{q},\mathbf{p})) \geq 1 + \mathrm{KL}(\mathbf{p},\mathbf{q}) \}.
\end{aligned} \tag{4.22}$$

We observe that if $\mathbf{p},\mathbf{q} \in \mathbb{P}^n$ are two probability distributions such that $0 < \mathrm{KL}(\mathbf{q},\mathbf{p}) < 1$, then

$$\operatorname{Sla}((\cdot)\ln(\cdot),\mathbf{p},\mathbf{q}) = \left[\frac{1 + \mathrm{KL}(\mathbf{p},\mathbf{q})}{1 - \mathrm{KL}(\mathbf{q},\mathbf{p})}, \infty \right). \tag{4.23}$$

If $\mathrm{KL}(\mathbf{q},\mathbf{p}) \geq 1$, then $\operatorname{Sla}((\cdot)\ln(\cdot),\mathbf{p},\mathbf{q}) = \emptyset$.

By the use of Theorem 4.1, we can state now the following.

Proposition 4.3. *Let $\mathbf{p},\mathbf{q} \in \mathbb{P}^n$ be two probability distributions such that $0 < \mathrm{KL}(\mathbf{q},\mathbf{p}) < 1$. If $v \in [(1 + \mathrm{KL}(\mathbf{p},\mathbf{q}))/(1 - \mathrm{KL}(\mathbf{q},\mathbf{p})), \infty)$, then one has*

$$(\ln v + 1)(v - 1) \geq v\ln(v) - \mathrm{KL}(\mathbf{p},\mathbf{q}) \geq 0. \tag{4.24}$$

In particular, for $v = (1 + \mathrm{KL}(\mathbf{p},\mathbf{q}))/(1 - \mathrm{KL}(\mathbf{q},\mathbf{p}))$, one gets

$$\begin{aligned}
&\left(\ln\left[\frac{1 + \mathrm{KL}(\mathbf{p},\mathbf{q})}{1 - \mathrm{KL}(\mathbf{q},\mathbf{p})} \right] + 1 \right) \left(\frac{1 + \mathrm{KL}(\mathbf{p},\mathbf{q})}{1 - \mathrm{KL}(\mathbf{q},\mathbf{p})} - 1 \right) \\
&\quad \geq \frac{1 + \mathrm{KL}(\mathbf{p},\mathbf{q})}{1 - \mathrm{KL}(\mathbf{q},\mathbf{p})} \ln\left[\frac{1 + \mathrm{KL}(\mathbf{p},\mathbf{q})}{1 - \mathrm{KL}(\mathbf{q},\mathbf{p})} \right] - \mathrm{KL}(\mathbf{p},\mathbf{q}) \geq 0.
\end{aligned} \tag{4.25}$$

Similar results can be obtained for other divergence measures of interest such as the *Jeffreys divergence and Hellinger discrimination*. However, the details are left to the interested reader.

Acknowledgment

The author would like to thank the anonymous referees for their valuable comments that have been implemented in the final version of the paper.

References

[1] M. S. Slater, “A companion inequality to Jensen’s inequality,” *Journal of Approximation Theory*, vol. 32, no. 2, pp. 160–166, 1981.

[2] J. E. Pečarić, “A multidimensional generalization of Slater’s inequality,” *Journal of Approximation Theory*, vol. 44, no. 3, pp. 292–294, 1985.

[3] J. E. Pečarić, F. Proschan, and Y. L. Tong, *Convex Functions, Partial Orderings, and Statistical Applications*, vol. 187 of *Mathematics in Science and Engineering*, Academic Press, New York, NY, USA, 1992.

[4] S. S. Dragomir, *Discrete Inequalities of the Cauchy-Bunyakovsky-Schwarz Type*, Nova Science Publishers, Hauppauge, NY, USA, 2004.

[5] M. A. Khan and J. E. Pečarić, "Improvement and reversion of Slater's inequality and related results," *Journal of Inequalities and Applications*, vol. 2010, Article ID 646034, 14 pages, 2010.

[6] S. S. Dragomir, "Some Slater's type inequalities for convex functions defined on linear spaces and applications," *RGMIA Research Report Collection*, vol. 12, article 8, 2009.

[7] M. Matić and J. Pečarić, "Some companion inequalities to Jensen's inequality," *Mathematical Inequalities & Applications*, vol. 3, no. 3, pp. 355–368, 2000.

[8] S. S. Dragomir, *Semi-Inner Products and Applications*, Nova Science Publishers, Hauppauge, NY, USA, 2004.

[9] I. Csiszár, "Information-type measures of difference of probability distributions and indirect observations," *Studia Scientiarum Mathematicarum Hungarica*, vol. 2, pp. 299–318, 1967.

[10] I. Csiszár and J. Körner, *Information Theory: Coding Theorems for Discrete Memoryless Systems*, Probability and Mathematical Statistics, Academic Press, New York, NY, USA, 1981.