Mathematical Inequalities & Applications
Volume 22, Number 1 (2019), 297–306
doi:10.7153/mia-2019-22-23

EQUALITY CASES OF INEQUALITIES INVOLVING GENERALIZED CSISZÁR AND TSALLIS TYPE f–DIVERGENCES

MAREK NIEZGODA

(*Communicated by S. Varošanec*)

Abstract. In this note, we study the problem of equality case of two inequalities involving generalized Csiszár f-divergences and generalized Tsallis f-divergences, respectively, with a convex function f. To this end we use generalized inverses of matrices and inverse-positive matrices.

1. Introduction

Throughout the paper, $\mathbb{R}_+ = [0,\infty)$ and $\mathbb{R}_{++} = (0,\infty)$. Elements of $\mathbb{R}^n$ will be referred to as row n-vectors.

For a convex function $f : [0,\infty) \to \mathbb{R}$ and two nonnegative n-tuples $\mathbf{p} = (p_1,\ldots,p_n)$ and $\mathbf{q} = (q_1,\ldots,q_n)$, the *Csiszár f-divergence* is defined by

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

Here $0f\left(\frac{0}{0}\right) = 0$ and $0f\left(\frac{c}{0}\right) = c \lim\limits_{t\to\infty} \frac{f(t)}{t}$, $c > 0$ (see [1, 2, 3]).

The Csiszár-Körner inequality states that

$$\sum_{i=1}^{n} p_i f\left(\frac{\sum_{i=1}^{n} q_i}{\sum_{i=1}^{n} p_i}\right) \leqslant C_f(\mathbf{p},\mathbf{q}) \tag{2}$$

(see [2, 11]). See [3, 4, 8] for other inequalities for f-divergence.

An extension of definition (1) is given as follows.

Let $f : \mathbb{R}_+ \to \mathbb{R}$ be a convex function on $\mathbb{R}_+$, and $\mathbf{p} = (p_1,\ldots,p_n) \in \mathbb{R}^n_+$, $\mathbf{q} = (q_1,\ldots,q_n) \in \mathbb{R}^n_+$, and $\mathbf{r} = (r_1,\ldots,r_n) \in \mathbb{R}^n_+$. Then the *generalized Csiszár f-divergence* is defined by

$$C_f(\mathbf{p},\mathbf{q};\mathbf{r}) = \sum_{i=1}^{n} r_i p_i f\left(\frac{q_i}{p_i}\right) \tag{3}$$

(see [9]).

An $n\times m$ real matrix $R = (r_{ij})$ is said to be *positive* (entrywise), written as $R \geqslant 0$, if $r_{ij} \geqslant 0$ for all $i = 1,\ldots,n$ and $j = 1,\ldots,m$. The symbol R^T is used to denote the transpose of a matrix R.

In [9] the authors proved the following result.

Mathematics subject classification (2010): 94A17, 26D15, 15B48.

Keywords and phrases: Convex function, Csiszár f-divergence, Tsallis type f-divergence, positive (entrywise) matrix, generalized inverse of matrix, inverse-positive (entrywise) matrix.

© ELEMENT, Zagreb
Paper MIA-22-23

THEOREM A. [9] *Let $f:\mathbb{R}_+\to\mathbb{R}$ be a convex function on $\mathbb{R}_+$. Let $\mathbf{p}=(p_1,\dots,p_n)\in\mathbb{R}^n_{++}$, $\mathbf{q}=(q_1,\dots,q_n)\in\mathbb{R}^n_+$ and $\mathbf{d}=(d_1,\dots,d_m)\in\mathbb{R}^m_+$.*

Let R be an $n\times m$ positive (entrywise) matrix. Denote

$$\widetilde{\mathbf{p}}=\mathbf{p}R\,,\quad \widetilde{\mathbf{q}}=\mathbf{q}R \quad and \quad \mathbf{c}=\mathbf{d}R^T. \tag{4}$$

Assume $\widetilde{\mathbf{p}}\in\mathbb{R}^m_{++}$.

Then

$$C_f\left(\widetilde{\mathbf{p}},\widetilde{\mathbf{q}};\mathbf{d}\right)\leqslant C_f\left(\mathbf{p},\mathbf{q};\mathbf{c}\right). \tag{5}$$

Let $f:I\times[0,\infty)\to\mathbb{R}$ be a two variables function on $I\times[0,\infty)$, where I is an interval in $\mathbb{R}$. We denote $f_u(t)=f(t,u)$ for $t\in I$ and $u\geqslant 0$. Then $\{f_u:u\in[0,\infty)\}$ is a family of real functions on I. We use the notation

$$g_u(t)=g(t,u)=\frac{f(t,u)-f(t,0)}{u} \quad \text{for } t\in I \text{ and } u>0, \tag{6}$$

$$g_0(t)=g(t,0)=\lim_{u\to 0^+}\frac{f(t,u)-f(t,0)}{u} \quad \text{for } t\in I \tag{7}$$

(see [7, p. 854]).

For instance, in the standard case $f_u(t)=t^u$ for $t>0$, $u\geqslant 0$, one obtains $g_u(t)=\frac{t^u-1}{u}=\ln_u t$ for $u>0$, and $g_0(t)=\lim\limits_{u\to 0^+}\frac{t^u-1}{u}=\ln t$.

In what follows we also deal with the Tsallis type divergence (entropy):

$$T_{f_u}\left(\mathbf{p},\mathbf{q};\mathbf{r}\right)=C_{g_u}\left(\mathbf{p},\mathbf{q};\mathbf{r}\right) \quad \text{for } u\in(0,\infty), \tag{8}$$

with $\mathbf{p}=(p_1,\dots,p_n)\in\mathbb{R}^n_+$, $\mathbf{q}=(q_1,\dots,q_n)\in\mathbb{R}^n_+$, $\mathbf{r}=(r_1,\dots,r_n)\in\mathbb{R}^n_+$ (see [9]).

THEOREM B. [9] *Let $f_u:\mathbb{R}_+\to\mathbb{R}$ be a convex function on $\mathbb{R}_+$ for some $u>0$. Let $f_0:\mathbb{R}_+\to\mathbb{R}$ be a constant function. Let $\mathbf{p}=(p_1,\dots,p_n)\in\mathbb{R}^n_{++}$, $\mathbf{q}=(q_1,\dots,q_n)\in\mathbb{R}^n_+$ and $\mathbf{d}=(d_1,\dots,d_m)\in\mathbb{R}^m_+$.*

Let R be an $n\times m$ positive (entrywise) matrix. Denote

$$\widetilde{\mathbf{p}}=\mathbf{p}R\,,\quad \widetilde{\mathbf{q}}=\mathbf{q}R \quad and \quad \mathbf{c}=\mathbf{d}R^T. \tag{9}$$

Assume $\widetilde{\mathbf{p}}\in\mathbb{R}^m_{++}$.

Then

$$T_{f_u}\left(\widetilde{\mathbf{p}},\widetilde{\mathbf{q}};\mathbf{d}\right)\leqslant T_{f_u}\left(\mathbf{p},\mathbf{q};\mathbf{c}\right). \tag{10}$$

The purpose of the present note is to discuss the equality cases of the inequalities (5) and (10). To do so, we utilize generalized inverses of matrices and inverse-positive matrices. In particular, we also use both left and right inverses of matrices.

2. Results for Csiszár type f-divergence

Let S be a given $m\times n$ real matrix.

We say that $n\times m$ real matrix S^- is a *generalized inverse* of S if $SS^-S=S$ (see [13]). When $m=n$ and S is invertible, then S^- is unique and $S^-=S^{-1}$.

We say that $n \times m$ real matrix S^+ is a *Moore-Penrose inverse* of S if $SS^+S = S$, $S^+SS^+ = S^+$, $(SS^+)^T = SS^+$ and $(S^+S)^T = S^+S$ (see [13]). The Moore-Penrose inverse of S is unique.

Remind that elements of $\mathbb{R}^n$ and of $\mathbb{R}^m$ are row vectors. Therefore, in the sequel, we also treat S as a linear map $S : (\mathbb{R}^n)^T \to (\mathbb{R}^m)^T$ via $\mathbf{v}^T \to S \cdot \mathbf{v}^T$ for $\mathbf{v} \in \mathbb{R}^n$, where $(\mathbb{R}^k)^T = \{\mathbf{a}^T : \mathbf{a} \in \mathbb{R}^k\}$. Analogously, S^T can be viewed as a linear map $S^T : (\mathbb{R}^m)^T \to (\mathbb{R}^n)^T$ via $\mathbf{u}^T \to S^T \cdot \mathbf{u}^T$ for $\mathbf{u} \in \mathbb{R}^m$.

The symbols $\operatorname{ran} S$ and $\operatorname{ran} S^T$ stand for the *ranges* (i.e., column spaces) of S and S^T, respectively. That is, $\operatorname{ran} S = \{S\mathbf{v}^T : \mathbf{v} \in \mathbb{R}^n\}$ and $\operatorname{ran} S^T = \{S^T\mathbf{u}^T : \mathbf{u} \in \mathbb{R}^m\}$.

An equality case of inequality (5) in Theorem A is described in the following.

THEOREM 1. *Let $f : \mathbb{R}_+ \to \mathbb{R}$ be a convex function on $\mathbb{R}_+$. Let $\mathbf{p} = (p_1, \ldots, p_n) \in \mathbb{R}^n_{++}$, $\mathbf{q} = (q_1, \ldots, q_n) \in \mathbb{R}^n_+$ and $\mathbf{d} = (d_1, \ldots, d_m) \in \mathbb{R}^m_+$.*

Let S and R be positive (entrywise) real matrices of sizes $m \times n$ and $n \times m$, respectively, such that $R = S^-$. Denote

$$\widetilde{\mathbf{p}} = \mathbf{p}R\,, \quad \widetilde{\mathbf{q}} = \mathbf{q}R \quad \textit{and} \quad \mathbf{c} = \mathbf{d}R^T. \tag{11}$$

Assume $\widetilde{\mathbf{p}} \in \mathbb{R}^m_{++}$.

If

$$\mathbf{p}^T \in \operatorname{ran} S^T\,, \quad \mathbf{q}^T \in \operatorname{ran} S^T \quad \textit{and} \quad \mathbf{d}^T \in \operatorname{ran} S, \tag{12}$$

then

$$C_f(\widetilde{\mathbf{p}}, \widetilde{\mathbf{q}}; \mathbf{d}) = C_f(\mathbf{p}, \mathbf{q}; \mathbf{c})\,. \tag{13}$$

Proof. In light of Theorem A in Section 1 we obtain

$$C_f(\widetilde{\mathbf{p}}, \widetilde{\mathbf{q}}; \mathbf{d}) \leqslant C_f(\mathbf{p}, \mathbf{q}; \mathbf{c})\,. \tag{14}$$

On the other hand, it follows from (12) that

$$\mathbf{p}^T = S^T\mathbf{u}^T\,, \quad \mathbf{q}^T = S^T\mathbf{w}^T\,, \quad \mathbf{d}^T = S\mathbf{v}^T$$

for some $\mathbf{u}, \mathbf{w} \in \mathbb{R}^m$ and $\mathbf{v} \in \mathbb{R}^n$. Thus we get $\mathbf{p} = \mathbf{u}S$, $\mathbf{q} = \mathbf{w}S$ and $\mathbf{d} = \mathbf{v}S^T$. So, (11) implies that $\widetilde{\mathbf{p}} = \mathbf{u}SR$ and $\widetilde{\mathbf{q}} = \mathbf{w}SR$. Furthermore, by $SS^-S = S$ and $R = S^-$, we have

$$\widetilde{\mathbf{p}}S = \mathbf{u}SRS = \mathbf{u}S = \mathbf{p}$$

and analogously,

$$\widetilde{\mathbf{q}}S = \mathbf{w}SRS = \mathbf{w}S = \mathbf{q}.$$

Now, recall that $\mathbf{d} = \mathbf{v}S^T$. Moreover, (11) gives $\mathbf{c} = \mathbf{d}S^{-T}$. Hence $\mathbf{c} = \mathbf{v}S^TS^{-T}$ and therefore $\mathbf{c}S^T = \mathbf{v}S^TS^{-T}S^T = \mathbf{v}(SS^-S)^T = \mathbf{v}S^T = \mathbf{d}$.

In summary, we have

$$\mathbf{p} = \widetilde{\mathbf{p}}S\,, \quad \mathbf{q} = \widetilde{\mathbf{q}}S \quad \text{and} \quad \mathbf{d} = \mathbf{c}S^T. \tag{15}$$

Simultaneously the matrix S is positive (entrywise). So, by (15) and Theorem A applied to S, $\widetilde{\mathbf{p}}$, $\widetilde{\mathbf{q}}$ and $\mathbf{c}$, we derive the inequality

$$C_f(\mathbf{p},\mathbf{q};\mathbf{c}) \leqslant C_f(\widetilde{\mathbf{p}},\widetilde{\mathbf{q}};\mathbf{d}). \tag{16}$$

By combining (14) and (16) we get (13), as required. □

If $\operatorname{rank} S = n$ then there exists an $n \times m$ real matrix S_l^{-1}, called a *left inverse* of S, such that $S_l^{-1} S = I_n$, where I_n denotes the $n \times n$ identity matrix [13].

If $\operatorname{rank} S = m$ then there exists an $n \times m$ real matrix S_r^{-1}, called a *right inverse* of S, such that $S S_r^{-1} = I_m$, where I_m denotes the $m \times m$ identity matrix [13].

It is not hard to verify that left- and right-inverses S_l^{-1} and S_r^{-1} are generalized inverses of S.

For one-sided inverses of S condition (12) simplifies.

COROLLARY 1. *Let $f : \mathbb{R}_+ \to \mathbb{R}$ be a convex function on $\mathbb{R}_+$. Let $\mathbf{p} = (p_1,\ldots,p_n) \in \mathbb{R}^n_{++}$, $\mathbf{q} = (q_1,\ldots,q_n) \in \mathbb{R}^n_+$ and $\mathbf{d} = (d_1,\ldots,d_m) \in \mathbb{R}^m_+$.*

Let S and R be positive (entrywise) real matrices of sizes $m \times n$ and $n \times m$, respectively. Denote

$$\widetilde{\mathbf{p}} = \mathbf{p}R\,, \quad \widetilde{\mathbf{q}} = \mathbf{q}R \quad \text{and} \quad \mathbf{c} = \mathbf{d}R^T. \tag{17}$$

Assume $\widetilde{\mathbf{p}} \in \mathbb{R}^m_{++}$.

(i) *If $\operatorname{rank} S = n$ and $R = S_l^{-1}$ and*

$$\mathbf{d}^T \in \operatorname{ran} S, \tag{18}$$

then

$$C_f(\widetilde{\mathbf{p}},\widetilde{\mathbf{q}};\mathbf{d}) = C_f(\mathbf{p},\mathbf{q};\mathbf{c}). \tag{19}$$

(ii) *If $\operatorname{rank} S = m$ and $R = S_r^{-1}$ and*

$$\mathbf{p}^T \in \operatorname{ran} S^T \quad \text{and} \quad \mathbf{q}^T \in \operatorname{ran} S^T, \tag{20}$$

then

$$C_f(\widetilde{\mathbf{p}},\widetilde{\mathbf{q}};\mathbf{d}) = C_f(\mathbf{p},\mathbf{q};\mathbf{c}). \tag{21}$$

Proof. **(i)**. Since $\operatorname{rank} S = n$, we have $\operatorname{rank} S^T = n$. For this reason, $S^T : (\mathbb{R}^m)^T \to (\mathbb{R}^n)^T$, where $(\mathbb{R}^k)^T = \{\mathbf{a}^T : \mathbf{a} \in \mathbb{R}^k\}$, and $\operatorname{ran} S^T = (\mathbb{R}^n)^T$. But $\mathbf{p} \in \mathbb{R}^n_+$ and $\mathbf{q} \in \mathbb{R}^n_+$, so $\mathbf{p}^T \in \operatorname{ran} S^T$ and $\mathbf{q}^T \in \operatorname{ran} S^T$. In addition, (18) holds. Thus condition (12) is met, as wanted.

Now, in order to get (19), it is sufficient to apply Theorem 1.

(ii). Since $\operatorname{rank} S = m$ and $S : (\mathbb{R}^n)^T \to (\mathbb{R}^m)^T$, therefore $\operatorname{ran} S = (\mathbb{R}^m)^T$. But $\mathbf{d} \in \mathbb{R}^m_+$, so $\mathbf{d}^T \in \operatorname{ran} S$. Moreover, (20) is fulfilled. Thus condition (12) is satisfied.

Now, by making use of Theorem 1 we get (21). □

An $n \times n$ invertible real matrix R is said to be *inverse-positive* if the matrix R^{-1} is positive [12].

A consequence of Corollary 1 is the following result for an invertible matrix R.

COROLLARY 2. *Let* $f:\mathbb{R}_+\to\mathbb{R}$ *be a convex function on* $\mathbb{R}_+$. *Let* $\mathbf{p}=(p_1,\ldots,p_n)\in\mathbb{R}^n_{++}$, $\mathbf{q}=(q_1,\ldots,q_n)\in\mathbb{R}^n_+$ *and* $\mathbf{d}=(d_1,\ldots,d_n)\in\mathbb{R}^n_+$.

Let R *be an* $n\times n$ *real matrix such that*

(i) *R is invertible,*

(ii) *R is positive (entrywise),*

(iii) *R is inverse-positive (entrywise).*

Denote

$$\widetilde{\mathbf{p}}=\mathbf{p}R\,,\quad \widetilde{\mathbf{q}}=\mathbf{q}R\quad and\quad \mathbf{c}=\mathbf{d}R^T. \tag{22}$$

Assume $\widetilde{\mathbf{p}}\in\mathbb{R}^n_{++}$.

Then

$$C_f(\widetilde{\mathbf{p}},\widetilde{\mathbf{q}};\mathbf{d})=C_f(\mathbf{p},\mathbf{q};\mathbf{c})\,. \tag{23}$$

Proof. By putting $S=R^{-1}$ and employing (22) we see that S is invertible and $R=S^{-1}=S_l^{-1}$, $\operatorname{rank}S=n$ and $\operatorname{ran}S=(\mathbb{R}^n)^T$. Therefore (18) holds true.

It is now sufficient to apply Corollary 1, item (i). □

REMARK 1. The class of the $n\times n$ matrices R satisfying conditions **(i)**, **(ii)** and **(iii)** in Corollary 2 is not empty, since the $n\times n$ identity matrix I_n is so.

In the next example we show that there are matrices R satisfying the conditions **(i)** and **(ii)** but not **(iii)**. Thus, in general, the inequality (5) need not be an equality.

EXAMPLE 1. Take $n=2$ and

$$R=\begin{pmatrix}5&2\\2&1\end{pmatrix}. \tag{24}$$

Then R is positive (entrywise), and $\det R=1\neq 0$, so R is invertible. Additionally, R is *not* inverse-positive (entrywise), since

$$R^{-1}=\begin{pmatrix}1&-2\\-2&5\end{pmatrix}.$$

We end this section by quoting some definitions and results from Peris' paper [12] (with only minor modifications).

For a matrix R, we say that the splitting $R=B-A$ is *positive* if $A\geqslant 0$ and $B\geqslant 0$ (positive entrywise) (see [12, p. 47]).

A positive splitting $R=B-A$ of a square matrix R is said to be a B*-splitting* if B is nonsingular and

(a) for all $\mathbf{x}\in\mathbb{R}^n$, $B\mathbf{x}^T\geqslant 0$ implies $A\mathbf{x}^T\geqslant 0$,

(b) for all $\mathbf{x}\in\mathbb{R}^n$,

$$\begin{pmatrix}R\\B\end{pmatrix}\mathbf{x}^T\geqslant 0 \text{ implies } \mathbf{x}\geqslant 0$$

(see [12, p. 52]).

A criterion for the inverse-positivity of a matrix is incorporated in the following result of Peris [12].

THEOREM C. ([12, Theorems 1 and 5]) *For a square nonsingular matrix R, the following conditions are equivalent:*

(a) *R is inverse-positive.*

(b) *For all positive splittings of R,*

$$R = B - A, \quad B \geqslant 0, \quad A \geqslant 0,$$

there exist a vector $\mathbf{v} \geqslant 0$ with $\mathbf{v} \neq 0$ and scalar $\mu \in [0,1)$ such that $A\mathbf{v}^T = \mu B \mathbf{v}^T$.

(c) *R allows a B-splitting $R = B - A$ such that $\mu < 1$.*

By making use of Theorem C and Corollary 2 we obtain the following.

COROLLARY 3. *Under the assumptions of Corollary 2 with condition* **(iii)** *replaced by condition* **(b)** *or* **(c)** *in Theorem C, then equality (23) is satisfied.*

Proof. Combine Corollary 2 and Theorem C. □

3. Results for Tsallis type f_u-divergence

Throughout this section $f : I \times [0,\infty) \to \mathbb{R}$ is a two variables function on $I \times [0,\infty)$, where I is an interval in $\mathbb{R}$. We also use the notation $f_u(t) = f(t,u)$ for $t \in I$ and $u \geqslant 0$. In addition, the functions $g_u(t) = g(t,u)$ for $t \in I$ and $u > 0$, and $g_0(t) = g(t,0)$ for $t \in I$ are defined by (6)-(7).

It is easily seen that the convexity of f_u implies the convexity of g_u (see (6)).

An equality case of inequality (10) in Theorem B for the Tsallis type f_u-divergence is given below.

THEOREM 2. *Let $f_u : \mathbb{R}_+ \to \mathbb{R}$ be a convex function on $\mathbb{R}_+$ for some $u > 0$. Let $f_0 : \mathbb{R}_+ \to \mathbb{R}$ be a constant function. Let $\mathbf{p} = (p_1, \ldots, p_n) \in \mathbb{R}^n_{++}$, $\mathbf{q} = (q_1, \ldots, q_n) \in \mathbb{R}^n_+$ and $\mathbf{d} = (d_1, \ldots, d_m) \in \mathbb{R}^m_+$.*

Let S and R be positive (entrywise) real matrices of sizes $m \times n$ and $n \times m$, respectively, such that $R = S^-$. Denote

$$\widetilde{\mathbf{p}} = \mathbf{p}R\,, \quad \widetilde{\mathbf{q}} = \mathbf{q}R \quad \text{and} \quad \mathbf{c} = \mathbf{d}R^T. \tag{25}$$

Assume $\widetilde{\mathbf{p}} \in \mathbb{R}^m_{++}$.

If

$$\mathbf{p}^T \in \operatorname{ran} S^T\,, \quad \mathbf{q}^T \in \operatorname{ran} S^T \quad \text{and} \quad \mathbf{d}^T \in \operatorname{ran} S\,,$$

then

$$T_{f_u}(\widetilde{\mathbf{p}}, \widetilde{\mathbf{q}}; \mathbf{d}) = T_{f_u}(\mathbf{p}, \mathbf{q}; \mathbf{c})\,. \tag{26}$$

Proof. Apply Theorem 1 for the convex function g_u (see (6) and (8)). □

COROLLARY 4. *Let $f_u : \mathbb{R}_+ \to \mathbb{R}$ be a convex function on $\mathbb{R}_+$ for some $u > 0$. Let $f_0 : \mathbb{R}_+ \to \mathbb{R}$ be a constant function. Let $\mathbf{p} = (p_1, \ldots, p_n) \in \mathbb{R}^n_{++}$, $\mathbf{q} = (q_1, \ldots, q_n) \in \mathbb{R}^n_+$ and $\mathbf{d} = (d_1, \ldots, d_m) \in \mathbb{R}^m_+$.*

Let S and R be positive (entrywise) real matrices of sizes $m \times n$ and $n \times m$, respectively. Denote

$$\widetilde{\mathbf{p}} = \mathbf{p}R\,, \quad \widetilde{\mathbf{q}} = \mathbf{q}R \quad \text{and} \quad \mathbf{c} = \mathbf{d}R^T. \tag{27}$$

Assume $\widetilde{\mathbf{p}} \in \mathbb{R}^m_{++}$.

(i) *If $\operatorname{rank} S = n$ and $R = S_l^{-1}$ and*

$$\mathbf{d}^T \in \operatorname{ran} S, \tag{28}$$

then

$$T_{f_u}(\widetilde{\mathbf{p}}, \widetilde{\mathbf{q}}; \mathbf{d}) = T_{f_u}(\mathbf{p}, \mathbf{q}; \mathbf{c}). \tag{29}$$

(ii) *If $\operatorname{rank} S = m$ and $R = S_r^{-1}$ and*

$$\mathbf{p}^T \in \operatorname{ran} S^T \quad \text{and} \quad \mathbf{q}^T \in \operatorname{ran} S^T, \tag{30}$$

then

$$T_{f_u}(\widetilde{\mathbf{p}}, \widetilde{\mathbf{q}}; \mathbf{d}) = T_{f_u}(\mathbf{p}, \mathbf{q}; \mathbf{c}). \tag{31}$$

Proof. Apply Corollary 1 for the convex function g_u and use (8). □

COROLLARY 5. *Let $f_u : \mathbb{R}_+ \to \mathbb{R}$ be a convex function on $\mathbb{R}_+$ for some $u > 0$. Let $f_0 : \mathbb{R}_+ \to \mathbb{R}$ be a constant function. Let $\mathbf{p} = (p_1, \ldots, p_n) \in \mathbb{R}^n_{++}$, $\mathbf{q} = (q_1, \ldots, q_n) \in \mathbb{R}^n_+$ and $\mathbf{d} = (d_1, \ldots, d_n) \in \mathbb{R}^n_+$.*

Let R be an $n \times n$ real matrix such that

(i) *R is invertible,*

(ii) *R is positive (entrywise),*

(iii) *R is inverse-positive (entrywise).*

Denote

$$\widetilde{\mathbf{p}} = \mathbf{p}R\,, \quad \widetilde{\mathbf{q}} = \mathbf{q}R \quad \text{and} \quad \mathbf{c} = \mathbf{d}R^T. \tag{32}$$

Assume $\widetilde{\mathbf{p}} \in \mathbb{R}^n_{++}$.

Then

$$T_{f_u}(\widetilde{\mathbf{p}}, \widetilde{\mathbf{q}}; \mathbf{d}) = T_{f_u}(\mathbf{p}, \mathbf{q}; \mathbf{c}). \tag{33}$$

Proof. It is enough to employ Corollary 2 for the convex function g_u, and next use (8). □

COROLLARY 6. *Under the assumptions of Corollary 5 with condition* **(iii)** *replaced by condition* **(b)** *or* **(c)** *in Theorem C, then equality (33) is satisfied.*

Proof. Combine Theorem C in Section 2 and Corollary 5. □

4. Examples

In this section we present some examples illustrating the results of the previous sections.

Let f be a convex function on $I=(0,\infty)$. Remind that the *generalized Csiszár f-divergence* of $\mathbf{p}$ and $\mathbf{q}$ with respect to $\mathbf{r}$ is

$$C_f(\mathbf{p},\mathbf{q};\mathbf{r})=\sum_{i=1}^{n} r_i p_i f\left(\frac{q_i}{p_i}\right), \tag{34}$$

where

$$\mathbf{p}=(p_1,\ldots,p_n), \quad \mathbf{q}=(q_1,\ldots,q_n) \quad \text{and} \quad \mathbf{r}=(r_1,\ldots,r_n)$$

with positive numbers p_i and q_i, and nonnegative r_i for $i=1,\ldots,n$.

We now give definitions of some relative entropies corresponding to the functions $\log t$, $t^u \log t$, $\ln_u t=\frac{t^u-1}{u}$ and $\frac{[1-s+st^u]^{1/u}-1}{s}$, as follows

$$S(\mathbf{p},\mathbf{q})=\sum_{i=1}^{n} p_i \log\frac{q_i}{p_i}, \tag{35}$$

$$S_u(\mathbf{p},\mathbf{q})=\sum_{i=1}^{n} p_i \left(\frac{q_i}{p_i}\right)^u \log\left(\frac{q_i}{p_i}\right), \tag{36}$$

$$T_u(\mathbf{p},\mathbf{q})=\sum_{i=1}^{n} p_i \ln_u\left(\frac{q_i}{p_i}\right), \tag{37}$$

where $u\in(0,1]$, and

$$T_{s,u}(\mathbf{p},\mathbf{q})=\sum_{i=1}^{n} p_i \frac{\left[1-s+s\left(\frac{q_i}{p_i}\right)^u\right]^{1/u}-1}{s}, \tag{38}$$

where $s\in(0,1]$ and $u\in[-1,1]$, $u\neq 0$ (see [5, 6, 10, 14]).

In the sequel we use the following notation and assumptions.

Let $\mathbf{p}=(p_1,\ldots,p_n)\in\mathbb{R}^n_{++}$, $\mathbf{q}=(q_1,\ldots,q_n)\in\mathbb{R}^n_{++}$ and $\mathbf{d}=(d_1,\ldots,d_m)\in\mathbb{R}^m_{+}$. Let R be an $n\times m$ positive (entrywise) matrix. We denote $\widetilde{\mathbf{p}}=\mathbf{p}R$, $\widetilde{\mathbf{q}}=\mathbf{q}R$ and $\mathbf{c}=\mathbf{d}R^T$ with $\widetilde{\mathbf{p}}=(\widetilde{p}_1,\ldots,\widetilde{p}_m)$, $\widetilde{\mathbf{q}}=(\widetilde{q}_1,\ldots,\widetilde{q}_m)$ and $\mathbf{c}=(c_1,\ldots,c_n)$.

EXAMPLE 2. We are now ready to illustrate Theorem A and Theorem 1 in the context of the convex function $f_1(t)=-\log t$, $t>0$ (see [5, 10]).

By Theorem A, we get the inequality

$$\sum_{i=1}^{m} d_i\widetilde{p}_i \log\left(\frac{\widetilde{p}_i}{\widetilde{q}_i}\right)=C_{-\log}(\widetilde{\mathbf{p}},\widetilde{\mathbf{q}};\mathbf{d})\leqslant C_{-\log}(\mathbf{p},\mathbf{q};\mathbf{c})=\sum_{i=1}^{n} c_i p_i \log\left(\frac{p_i}{q_i}\right). \tag{39}$$

According to Theorem 1, equality holds in inequality (39) whenever $R=S^-$ for some positive (entrywise) real matrix S of size $m\times n$ such that

$$\mathbf{p}^T\in\operatorname{ran} S^T, \quad \mathbf{q}^T\in\operatorname{ran} S^T \quad \text{and} \quad \mathbf{d}^T\in\operatorname{ran} S. \tag{40}$$

EXAMPLE 3. We now verify our previous results for the convex function $f_2(t) = -\ln_u t = -\frac{t^u-1}{u}$, $t > 0$ (see [10, 14]).

By virtue of Theorem A, we find that

$$-\sum_{i=1}^{m} d_i\widetilde{p}_i \ln_u\left(\frac{\widetilde{q}_i}{\widetilde{p}_i}\right) = C_{-\ln_u}(\widetilde{\mathbf{p}},\widetilde{\mathbf{q}};\mathbf{d}) \leqslant C_{-\ln_u}(\mathbf{p},\mathbf{q};\mathbf{c}) = -\sum_{i=1}^{n} c_i p_i \ln_u\left(\frac{q_i}{p_i}\right). \quad (41)$$

On account of Theorem 1, equality is met in inequality (41) provided that $R = S^-$ for some positive (entrywise) real matrix S of size $m\times n$ such that (40) holds.

EXAMPLE 4. In this example we show some applications for the convex function $f_3(t) = t\log t$, $t > 0$ (see [5, 10]).

Thanks to Theorem A, we establish the inequality

$$\sum_{i=1}^{m} d_i\widetilde{q}_i \log\left(\frac{\widetilde{q}_i}{\widetilde{p}_i}\right) = C_{f_3}(\widetilde{\mathbf{p}},\widetilde{\mathbf{q}};\mathbf{d}) \leqslant C_{f_3}(\mathbf{p},\mathbf{q};\mathbf{c}) = \sum_{i=1}^{n} c_i q_i \log\left(\frac{q_i}{p_i}\right). \quad (42)$$

In light of Theorem 1, equality is satisfied in inequality (42) if $R = S^-$ for some positive (entrywise) real matrix S of size $m\times n$ such that (40) is fulfilled.

EXAMPLE 5. We now deal with the parametric Tsallis relative entropy $T_{s,u}(\mathbf{p},\mathbf{q})$ generated by the concave function

$$f_4(t) = \frac{(1-s+st^u)^{1/u}-1}{s}, \quad t>0$$

(see [6, 10]).

It follows from Theorem A that

$$\sum_{i=1}^{m} d_i\widetilde{p}_i \frac{\left[1-s+s\left(\frac{\widetilde{q}_i}{\widetilde{p}_i}\right)^u\right]^{1/u}-1}{s} = C_{f_4}(\widetilde{\mathbf{p}},\widetilde{\mathbf{q}};\mathbf{d})$$

$$\geqslant C_{f_4}(\mathbf{p},\mathbf{q};\mathbf{c}) = \sum_{i=1}^{n} c_i p_i \frac{\left[1-s+s\left(\frac{q_i}{p_i}\right)^u\right]^{1/u}-1}{s}. \quad (43)$$

By making use Theorem 1, we conclude that equality appears in inequality (43) if $R = S^-$ for some positive (entrywise) real matrix S of size $m\times n$ such that (40) is met.

Acknowledgement. The author is grateful to the anonymous referee for careful reading the paper and giving useful suggestions that improved the previous version of the manuscript.

REFERENCES

[1] I. CSISZÁR, *Information-type measures of differences of probability distributions and indirect observations*, Studia Sci. Math. Hung. **2** (1967), 299–318.

[2] I. CSISZÁR AND J. KÖRNER, *Information Theory: Coding Theorems for Discrete Memory-less Systems*, Academic Press, New York, 1981.

[3] S. S. DRAGOMIR, *Upper and lower bounds for Csiszár f-divergence in terms of the Kullback-Leibler distance and applications, Inequalities for the Csiszár f-divergence in Information Theory*, (Ed. S. S. Dragomir), 2000, `http://rgmia.vu.edu.au/monographs/csiszar.htm`.

[4] S. FURUICHI AND F.-C. MITROI, *Mathematical inequalities for some divergences*, Physica A, **391** (2012), 388–400.

[5] T. FURUTA, *Parametric extensions of Shannon inequality and its reverse one in Hilbert space operators*, Linear Algebra Appl. **381** (2004), 219–235.

[6] H. ISA, M. ITO, E. KAMEI, H. TOHYAMA AND M. WATANABE, *Generalizations of operator Shannon inequality based on Tsallis and Rényi relative entropies*, Linear Algebra Appl. **439** (10) (2013), 3148–3155.

[7] P. A. KLUZA AND M. NIEZGODA, *Inequalities for relative operator entropies*, Electron. J. Linear Algebra **27** (2014), 851–864.

[8] P. KLUZA AND M. NIEZGODA, *Generalizations of Crooks and Lin's results on Jeffreys–Csiszár and Jensen–Csiszár f-divergences*, Physica A, **463** (2016), 383–393.

[9] P. KLUZA AND M. NIEZGODA, *On Csiszár and Tsallis type f-divergences induced by superquadratic and convex functions*, Math. Inequal. Appl. **21** (2018), 455–467.

[10] M. NIEZGODA, *On Shannon like inequalities for f-connections of positive linear maps and positive operators*, Linear Algebra Appl. **481** (2015), 186–201.

[11] M. NIEZGODA, *Vector joint majorization and generalization of Csiszár–Körner's inequality for f-divergence*, Discrete Applied Math. **198** (2016), 195–205.

[12] J. E. PERIS, *A new characterization of inverse-positive matrices*, Linear Algebra Appl. **154–156** (1991), 45–58.

[13] C. R. RAO AND S. K. MITRA, *Generalized Inverse of Matrices and its Applications*, John Wiley and Sons, New York, 1971.

[14] K. YANAGI, K. KURIYAMA AND S. FURUICHI, *Generalized Shannon inequalities based on Tsallis relative operator entropy*, Linear Algebra Appl. **394** (2005), 109–118.

(Received November 29, 2017)

Marek Niezgoda
Department of Applied Mathematics and Computer Science
University of Life Sciences in Lublin
Akademicka 13, 20-950 Lublin, Poland
e-mail: bniezgoda@wp.pl; marek.niezgoda@up.lublin.pl