Rad HAZU
Volume 515 (2013), 11-32

BREGMAN AND BURBEA-RAO DIVERGENCE FOR MATRICES

M. ADIL KHAN[1], M. NIEZGODA[2], AND J. PEČARIĆ[1,3]

Abstract. In this paper, the Bregman and Burbea-Rao divergences for matrices are investigated. Two mean-value theorems for the divergences induced by C^2-functions are derived. As application, certain Cauchy type means of the entries of the matrices are constructed. By utilizing three classes of parametrized convex functions, the exponential convexity of the divergences, thought as a function of the parameter, is proved. The monotonicity of the corresponding means of Cauchy type is shown. Power means are also considered.

1. Introduction and summary

For a real convex function ϕ defined on an interval $I \subset \mathbb{R}$, the *Bregman-divergence* $\mathfrak{B}$ and *Burbea-Rao divergence* $\mathfrak{J}$ between vectors $\mathbf{x} = (x_1, \dots, x_n)$ and $\mathbf{y} = (y_1, \dots, y_n)$, where $x_i, y_i \in I$, $(i = 1, 2, ..., n)$, are

$$\mathfrak{B}_{n,\phi}(\mathbf{x}, \mathbf{y}) = \sum_{i=1}^{n} \left[\phi(x_i) - \phi(y_i) - \phi'(y_i)(x_i - y_i)\right], \text{ if } \phi \text{ is differentiable,} \tag{1}$$

$$\mathfrak{J}_{n,\phi}(\mathbf{x}, \mathbf{y}) = \sum_{i=1}^{n} \left\{\frac{1}{2}[\phi(x_i) + \phi(y_i)] - \phi\left(\frac{x_i + y_i}{2}\right)\right\} \tag{2}$$

(cf. [4, 9]).

Assume that I is an interval in $\mathbb{R}$ with interior I° and $\phi : I \to \mathbb{R}$ is a convex function on I. It is well known (see e.g. [5]) that then ϕ is

---

2000 *Mathematics Subject Classification.* Primary 26A51, 26A24, 26A48 .

Key words and phrases. Convex function, Bregman-divergence, Burbea-Rao divergence, exponential convexity, Cauchy means.

Acknowledgement: This research work of the first and third authors is funded by Higher Education Commission Pakistan. The research of the third author was supported by the Croatian Ministry of Science, Education and Sports under the Research Grants 117-1170889-0888. Some results of the paper have been presented by the first author in the International Conference Mathematical Inequalities and Applications 2010, Lahore, March, 2010.

continuous on I°, and, in addition, ϕ has finite left and right derivatives at each point of I°. Moreover, if $x, y \in I^\circ$ and $x < y$, then $D^-\phi(x) \le D^+\phi(x) \le D^-\phi(y) \le D^+\phi(y)$. Therefore both $D^-\phi$ and $D^+\phi$ are nondecreasing functions on I°. Also, a convex function must be differentiable except for at most countably many points [7, pp. 271–272].

For a convex function $\phi : I \to \mathbb{R}$, the *subdifferential* of ϕ, denoted by $\partial\phi$, is the set of all functions $\varphi : I \to [-\infty, \infty]$ such that $\varphi(I^\circ) \subset \mathbb{R}$ and

$$\phi(x) \ge \phi(a) + (x-a)\varphi(a) \quad \text{for any } x, a \in I. \tag{3}$$

The convexity of ϕ on I ensures that $D^-\phi, D^+\phi \in \partial\phi$, which shows that $\partial\phi$ is nonempty, and

$$D^-\phi(x) \le \varphi(x) \le D^+\phi(x) \quad \text{for any } x \in I^\circ \text{ and } \varphi \in \partial\phi. \tag{4}$$

In particular, φ is nondecreasing function.

If ϕ is differentiable convex on I°, then $\partial\phi = \{\phi'\}$.

The following theorem has been proved in [5].

Theorem 1 ([5]). *Let $\phi : I \subseteq \mathbb{R} \to \mathbb{R}$ be a convex function on the interval I, $x_i, y_i \in I^\circ$ and $p_i \ge 0$ $(i = 1, ..., n)$.*

If $\varphi \in \partial\phi$, then we have the inequality

$$\sum_{i=1}^{n} p_i \left[\phi(x_i) - \phi(y_i) - \varphi(y_i)(x_i - y_i)\right] \ge 0. \tag{5}$$

If ϕ is strictly convex on I and $p_i > 0$ $(i = 1, ..., n)$, then the equality holds in (5) if and only if $x_i = y_i$ $(i = 1, ..., n)$.

In fact, (5) is an extension of the fact that for convex function ϕ, $\mathfrak{B}_{n,\phi}(x, y) \ge 0$.

In this paper, we study the Bregman and Burbea-Rao divergences for matrices. In Section 2 we investigate properties of the Bregman-divergence. We begin with an extension of Theorem 1 from n-vectors to $n \times m$ matrices (see Theorem 2). This allows to derive two mean-value theorems for the divergences induced by C^2-functions (see Theorems 3 and 4). As application, we construct certain Cauchy type means of the entries of the matrices (see Corollary 1 and Remark 1). By utilising three classes of parametrized convex functions, we prove the exponential convexity of the divergences, thought as a function of the parameter (see Theorems 5, 7, 9). In particular, we present a Gram type inequality for the divergences. We also show that the corresponding means of Cauchy type are monotone in each variable (see Theorems 6, 8, 10).

In Section 3 we present some corresponding results for the Burbea-Rao divergence.

Finally, Section 4 is devoted to power means.

2. Bregman-divergence for matrices

We will denote by $[a_{ij}]$ $n \times m$ matrix with entries $a_{ij} \in I$ $(i = 1,2,...,n, j = 1,2,...,m)$.

Definition 1. Let $\phi : I \subseteq \mathbb{R} \to \mathbb{R}$ be a convex function on the interval I and $\varphi \in \partial\phi$. The *Bregman divergence* of two matrices $X = [x_{ij}]$ and $Y = [y_{ij}]$ with weight $W = [w_{ij}]$, where $x_{ij}, y_{ij} \in I$ and $w_{ij} \geq 0$ $(i = 1,2,...,n,\ j = 1,2,...,m)$, is defined by

$$B_{n,m,\phi,\varphi}(X,Y;W) = \sum_{i=1}^{n}\sum_{j=1}^{m} w_{ij}\left[\phi(x_{ij}) - \phi(y_{ij}) - \varphi(y_{ij})(x_{ij} - y_{ij})\right]. \tag{6}$$

If $\phi : I \subseteq \mathbb{R} \to \mathbb{R}$ is a differentiable function (not necessarily a convex function (see e.g. Theorems 3-4)), then we can rewrite (6) in the form

$$B_{n,m,\phi}(X,Y;W) = \sum_{i=1}^{n}\sum_{j=1}^{m} w_{ij}\left[\phi(x_{ij}) - \phi(y_{ij}) - \phi'(y_{ij})(x_{ij} - y_{ij})\right]. \tag{7}$$

In particular, if $W = [w_{ij}] = [v_i u_j]$ and $y_{ij} = \sum_{j=1}^{m} u_j x_{ij}$ for $i = 1,2,...,n, j = 1,2,...,m$ with $\sum_{j=1}^{m} u_j = 1$, then from (6) we obtain *Burbea-Rao divergence* of the matrix X with weight W as follows:

$$J_{n,m,\phi}(X,W) = \sum_{i=1}^{n} v_i\left[\sum_{j=1}^{m} u_j\phi(x_{ij}) - \phi\left(\sum_{j=1}^{m} u_j x_{ij}\right)\right]. \tag{8}$$

Moreover, if $v_i = 1$, $u_j = \frac{1}{2}$, $m = 2$, $x_{i1} = x_i$ and $x_{i2} = y_i$, then from (8) we get Burbea-Rao divergence (2).

Theorem 2. *Let $\phi : I \subseteq \mathbb{R} \to \mathbb{R}$ be a convex function on the interval I, and $\varphi \in \partial\phi$.*

Then for Bregman–divergence $B_{n,m,\phi,\varphi}(X,Y;W)$ of two matrices $X = [x_{ij}]$, $Y = [y_{ij}]$ with weight $W = [w_{ij}]$, where $x_{ij}, y_{ij} \in I$ and $w_{ij} \geq 0$ $(i = 1,2,...,n,\ j = 1,2,..,m)$, the following inequality holds.

$$B_{n,m,\phi,\varphi}(X,Y;W) \geq 0. \tag{9}$$

If ϕ is strictly convex on I and $w_{ij} > 0$ $(i = 1,...,n,\ j = 1,2,...,m)$, then the equality holds in (9) if and only if $X = Y$.

Proof. If we apply (3) for the choice $x = x_{ij}$, $a = y_{ij}$ $(i = 1,2,...,n,\ j = 1,2,...,m)$, we may write

$$\phi(x_{ij}) - \phi(y_{ij}) - \varphi(y_{ij})(x_{ij} - y_{ij}) \geq 0 \tag{10}$$

for any $i = 1, 2, ..., n$ and $j = 1, 2, ..., m$.

By multiplying (10) by $w_{ij} \geq 0$ and summing over j from 1 to m and then summing over i from 1 to n, we obtain (9).

The case of equality for strictly convex functions follows by the fact that we have equality for such a function in (10) if and only if $X = Y$. $\square$

Theorem 3. *Let $\phi \in C^2(I)$, where I is a closed interval in $\mathbb{R}$, and let X, Y and W be matrices as in Theorem 2 with $X \neq Y$ and $w_{ij} > 0$ $(i = 1, 2, ..., n,\ j = 1, 2, ..., m)$.*

Then there exists $\xi \in I$ such that

$$B_{n,m,\phi}(X, Y; W) = \frac{\phi''(\xi)}{2} \sum_{i=1}^{n} \sum_{j=1}^{m} w_{ij} \left(x_{ij} - y_{ij}\right)^2. \tag{11}$$

Proof. Since ϕ'' is continuous on I, so $m \leq \phi''(x) \leq M$ for $x \in I$, where $m = \min_{x \in I} \phi''(x)$ and $M = \max_{x \in I} \phi''(x)$. Consider the functions ϕ_1 and ϕ_2 defined on I as

$$\phi_1(x) = \frac{Mx^2}{2} - \phi(x) \quad \text{and} \quad \phi_2(x) = \phi(x) - \frac{mx^2}{2} \quad \text{for } x \in I.$$

It is easily seen that

$$\phi_1''(x) = M - \phi''(x) \geq 0 \quad \text{and} \quad \phi_2''(x) = \phi''(x) - m \geq 0 \quad \text{for } x \in I.$$

So ϕ_1 and ϕ_2 are convex.

Now by applying ϕ_1 for ϕ in Theorem 2, we have

$$\sum_{i=1}^{n} \sum_{j=1}^{m} w_{ij} \left[\frac{Mx_{ij}^2}{2} - \phi(x_{ij}) - \frac{My_{ij}^2}{2} + \phi(y_{ij}) - \left(My_{ij} - \phi'(y_{ij})\right)\left(x_{ij} - y_{ij}\right) \right] \geq 0.$$

Hence we get

$$B_{n,m,\phi}(X, Y; W) \leq \frac{1}{2} M \sum_{i=1}^{n} \sum_{j=1}^{m} w_{ij} \left(x_{ij} - y_{ij}\right)^2. \tag{12}$$

Similarly, by applying ϕ_2 for ϕ in Theorem 2, we get

$$B_{n,m,\phi}(X, Y; W) \geq \frac{1}{2} m \sum_{i=1}^{n} \sum_{j=1}^{m} w_{ij} (x_{ij} - y_{ij})^2. \tag{13}$$

But $\sum_{i=1}^{n} \sum_{j=1}^{m} w_{ij}(x_{ij} - y_{ij})^2 > 0$ as $X \neq Y$ and $w_{ij} > 0$ $(i = 1, 2, ..., n,\ j = 1, 2, ..., m)$. So, by combining (12) and (13), we obtain

$$m \leq \frac{2B_{n,m,\phi}(X, Y; W)}{\sum_{i=1}^{n} \sum_{j=1}^{m} w_{ij}(x_{ij} - y_{ij})^2} \leq M.$$

Now by using the fact that for $m \leq \rho \leq M$ there exists $\xi \in I$ such that $\phi''(\xi) = \rho$, we get (11). □

Theorem 4. *Let $\phi, \psi \in C^2(I)$, where I is a closed interval in $\mathbb{R}$, and let X, Y and W be matrices as in Theorem 2 with $X \neq Y$ and $w_{ij} > 0$ $(i = 1, 2, ..., n,\ j = 1, 2, ..., m)$.*

Then there exists $\xi \in I$ such that

$$\frac{\phi''(\xi)}{\psi''(\xi)} = \frac{B_{n,m,\phi}(X, Y; W)}{B_{n,m,\psi}(X, Y; W)} \tag{14}$$

provided that the denominators are nonzero.

Proof. Let the function $k \in C^2(I)$ be defined by

$$k = c_1\phi - c_2\psi,$$

where

$$c_1 = B_{n,m,\psi}(X, Y; W) \quad \text{and} \quad c_2 = B_{n,m,\phi}(X, Y; W). \tag{15}$$

It is not hard to check that $B_{n,m,k}(X, Y; W) = 0$.

In consequence, by using Theorem 3 for the function k, we find that

$$0 = \left(\frac{c_1\phi''(\xi)}{2} - \frac{c_2\psi''(\xi)}{2}\right) \sum_{i=1}^{n} \sum_{j=1}^{m} w_{ij}(x_{ij} - y_{ij})^2. \tag{16}$$

Since $\sum_{i=1}^{n} \sum_{j=1}^{m} w_{ij}(x_{ij} - y_{ij})^2 > 0$, equality (16) gives us

$$\frac{\phi''(\xi)}{\psi''(\xi)} = \frac{c_2}{c_1},$$

which together with (15) proves (14). □

Corollary 1. *Let X, Y and W be matrices as in Theorem 2 with $X \neq Y$ and $w_{ij} > 0$ $(i = 1, 2, ..., n,\ j = 1, 2, ..., m)$, where I is a positive closed interval.*

Then for $-\infty < v \neq 0, 1 \neq u < \infty$, $u \neq v$, there exists $\xi \in I$ such that

$$\xi^{u-v} = \frac{v(v-1)\sum_{i=1}^{n}\sum_{j=1}^{m} w_{ij}\left[x_{ij}^{u} - u x_{ij} y_{ij}^{u-1} + y_{ij}^{u}(u-1)\right]}{u(u-1)\sum_{i=1}^{n}\sum_{j=1}^{m} w_{ij}\left[x_{ij}^{v} - v x_{ij} y_{ij}^{v-1} + y_{ij}^{v}(v-1)\right]}. \tag{17}$$

Proof. By setting $\phi(x) = x^u$ and $\psi(x) = x^v$, $x \in I$, in Theorem 4, we get (17). □

Remark 1. Note that we can consider the interval $I = [m, M]$, where $m = \min\{\min_{i,j} x_{ij}, \min_{i,j} y_{ij}\}$ and $M = \max\{\max_{i,j} x_{ij}, \max_{i,j} y_{ij}\}$.

Since the function $\xi \to \xi^{u-v}$ with $u \neq v$ is invertible, then from (17) it follows that

$$m \leq \left\{ \frac{v(v-1)\sum_{i=1}^{n}\sum_{j=1}^{m} w_{ij}\left[x_{ij}^{u} - u x_{ij} y_{ij}^{u-1} + y_{ij}^{u}(u-1)\right]}{u(u-1)\sum_{i=1}^{n}\sum_{j=1}^{m} w_{ij}\left[x_{ij}^{v} - v x_{ij} y_{ij}^{v-1} + y_{ij}^{v}(v-1)\right]} \right\}^{\frac{1}{u-v}} \leq M. \tag{18}$$

Therefore the expression in the middle of (18) is a mean of x_{ij} and y_{ij}.

In fact, similar result can also be given for (14). Namely, suppose that the function $\frac{\phi''}{\psi''}$ has inverse function. Then from (14) we have

$$\xi = \left(\frac{\phi''}{\psi''}\right)^{-1} \left(\frac{B_{n,m,\phi}(X,Y;W)}{B_{n,m,\psi}(X,Y;W)}\right). \tag{19}$$

So, the expression on the right-hand side of (19) is also a mean of the entries of X and Y.

In the sequel, we need the following lemmas.

Lemma 1 ([11]). *Let us define the function*

$$\eta_t(x) = \begin{cases} \frac{x^t}{t(t-1)}, & t \neq 1; \\ x\log x, & t = 1 \end{cases} \qquad \text{for } x > 0, \tag{20}$$

Then $\eta_t''(x) = x^{t-2}$ for $x, t > 0$, that is η_t is convex on $(0, +\infty)$ for every $t \in (0, \infty)$.

Lemma 2 ([3]). *Let us define the function*

$$\varphi_t(x) = \begin{cases} \frac{x^t}{t(t-1)}\ , & t \neq 0, 1; \\ -\log x\ , & t = 0; \\ x\log x\ , & t = 1 \end{cases} \qquad \text{for } x > 0. \tag{21}$$

Then $\varphi_t''(x) = x^{t-2}$ for $x > 0$, $t \in \mathbb{R}$, that is φ_t is convex on $(0, +\infty)$ for every $t \in \mathbb{R}$.

Lemma 3 ([3]). *Let us define the function*

$$\phi_t(x) = \begin{cases} \frac{1}{t^2} e^{tx}\ , & t \neq 0; \\ \frac{1}{2}x^2, & t = 0 \end{cases} \qquad \text{for } x \in \mathbb{R}. \tag{22}$$

Then $\phi_t''(x) = e^{tx}$ for $x, t \in \mathbb{R}$, that is ϕ_t is a convex on $\mathbb{R}$. for every $t \in \mathbb{R}$.

By using (20) in (6) for $x_{ij}, y_{ij}, t > 0$ $(i = 1, 2, ..., n,\ j = 1, 2, ..., m)$, we get

$$B_{n,m,t}(X, Y; W) = \begin{cases} \frac{1}{t(t-1)} \sum_{i=1}^{n} \sum_{j=1}^{m} w_{ij} [x_{ij}^{t} - t x_{ij} y_{ij}^{t-1} + (t-1)\, y_{ij}^{t}] \ , \\ \hfill t \neq 1; \\ \\ \sum_{i=1}^{n} \sum_{j=1}^{m} w_{ij} \left[x_{ij}(\log x_{ij} - \log y_{ij} - 1) + y_{ij}\right] \ , \\ \hfill t = 1. \end{cases} \tag{23}$$

By applying (21) in (6) for $x_{ij}, y_{ij} > 0$ and $t \in \mathbb{R}$ $(i = 1, 2, ..., n,\ j = 1, 2, ..., m)$, we obtain

$$\tilde{B}_{n,m,t}(X, Y; W) = \begin{cases} B_{n,m,t}(X, Y; W) \ , & t \neq 0; \\ \sum_{i=1}^{n} \sum_{j=1}^{m} w_{ij} \left[\frac{x_{ij} - y_{ij}}{y_{ij}} + \log y_{ij} - \log x_{ij}\right] \ , & t = 0. \end{cases} \tag{24}$$

Analogously, by utilizing (6) and (22) for $x_{ij}, y_{ij}, t \in \mathbb{R}$ $(i = 1, 2, ..., n,\ j = 1, 2, ..., m)$, we have

$$\overline{B}_{n,m,t}(X, Y; W) = \begin{cases} \frac{1}{t^2} \sum_{i=1}^{n} \sum_{j=1}^{m} w_{ij} \left[e^{t x_{ij}} - e^{t y_{ij}} - t e^{t y_{ij}} (x_{ij} - y_{ij})\right] \ , \\ \hfill t \neq 0; \\ \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{m} w_{ij} \left[x_{ij}^2 - y_{ij}^2 - 2 y_{ij}(x_{ij} - y_{ij})\right] \ , \\ \hfill t = 0. \end{cases} \tag{25}$$

Lemma 4 ([10, p. 2]). *If ϕ is convex on an interval $I \subseteq \mathbb{R}$, then*

$$(s_3 - s_2)\phi(s_1) + (s_1 - s_3)\phi(s_2) + (s_2 - s_1)\phi(s_3) \geq 0 \tag{26}$$

holds for every $s_1, s_2, s_3 \in I$ such that $s_1 < s_2 < s_3$.

In what follows, the notion of exponential convexity plays an important role (see [1] and references therein).

Definition 2 ([1]). A function $\phi : I \to \mathbb{R}$ is said to be *exponentially convex* if it is continuous and

$$\sum_{k,l=1}^{n} a_k a_l \phi(x_k + x_l) \geq 0$$

for all $n \in \mathbb{N}$, $a_k \in \mathbb{R}$ and $x_k \in I$, $k = 1, 2, ..., n$, such that $x_k + x_l \in I$, $k, l = 1, 2, ..., n$.

Proposition 1 ([1]). *Let $\phi : I \to \mathbb{R}$. Then the following statements are equivalent.*

(i): *ϕ is exponentially convex.*

(ii): *ϕ is continuous and*

$$\sum_{k,l=1}^{n} a_k a_l \phi\left(\frac{x_k + x_l}{2}\right) \geq 0$$

for all $n \in \mathbb{N}$, $a_k \in \mathbb{R}$, $x_k \in I$, $k = 1, 2, ..., n$.

Corollary 2 ([1]). *If ϕ is an exponentially convex function, then*

$$\det\left[\phi\left(\frac{x_k + x_l}{2}\right)\right]_{k,l=1}^{n} \geq 0$$

for all $n \in \mathbb{N}$, $x_k \in I$, $k = 1, 2, ..., n$.

Corollary 3 ([1]). *If $\phi : I \to (0, \infty)$ is an exponentially convex function, then ϕ is a log-convex function that is*

$$\phi(\lambda x + (1 - \lambda)y) \leq \phi^{\lambda}(x)\phi^{1-\lambda}(y), \ \textit{for all } x, y \in I, \ \lambda \in [0, 1].$$

In the remaining part of Section 2, we use in turn the three parametrized classes of convex functions defined in Lemmas 1-3, respectively, to define and study related Bregman-divergences and Cauchy type means.

We are now in a position to establish the exponential and logarithmic convexity and related properties of the Bregman-divergence $t \to B_{n,m,t}(X, Y; W)$, $t > 0$, introduced in (23) and connected with Lemma 1.

Theorem 5. *Let X, Y and W be matrices as in Theorem 2 with $x_{ij}, y_{ij} \geq 0$ $(i = 1, 2, ..., n, \ j = 1, 2, ..., m)$. Denote*

$$\Gamma_t = B_{n,m,t}(X, Y; W), \qquad t > 0.$$

Then

(a): *for all $n \in \mathbb{N}$, $p_k \in \mathbb{R}^+$, $k = 1, 2, ..., n$, the matrix $\left[\Gamma_{\frac{p_k+p_l}{2}}\right]_{k,l=1}^{n}$ is positive semi-definite. In particular,*

$$\det\left[\Gamma_{\frac{p_k+p_l}{2}}\right]_{k,l=1}^{n} \geq 0, \tag{27}$$

(b): *the function $t \to \Gamma_t$ is exponentially convex,*
(c): *if $\Gamma_t > 0$, then the function $t \to \Gamma_t$ is log-convex, i.e.,*

$$(\Gamma_s)^{t-r} \leq (\Gamma_r)^{t-s}(\Gamma_t)^{s-r} \quad \textit{for } 0 < r < s < t < \infty. \tag{28}$$

Proof. (a). As in [8], let us consider the function defined by

$$\mu(x) = \sum_{k,l=1}^{n} a_k a_l \eta_{p_{kl}}(x) \quad \text{for } x > 0,$$

where $a_k \in \mathbb{R}$ for all $k = 1, 2, ..., n$, and $p_{kl} = \frac{p_k + p_l}{2} > 0$, $k, l = 1, 2, ..., n$. By Lemma 1, it is easily seen that

$$\mu''(x) = \sum_{k,l=1}^{n} a_k a_l x^{p_{kl}-2} = \left(\sum_{k=1}^{n} a_k x^{\frac{p_k - 2}{2}} \right)^2 \geq 0 \quad \text{for } x > 0.$$

Therefore $\mu(\cdot)$ is convex on $[0, +\infty)$. By using (9) we obtain

$$B_{n,m,\mu}(X, Y; W) \geq 0.$$

Hence

$$\sum_{k,l=1}^{n} a_k a_l \Gamma_{p_{kl}} \geq 0,$$

so the matrix $\left[\Gamma_{\frac{p_k + p_l}{2}} \right]_{k,l=1}^{n}$ is positive semi-definite.

(b). Since $\lim_{t \to 1} \Gamma_t = \Gamma_1$, so the function $t \to \Gamma_t$ is continuous for all $t > 0$. By using Proposition 1 and the proved positive semi-definity of the matrix $\left[\Gamma_{\frac{p_k + p_l}{2}} \right]_{k,l=1}^{n}$, we obtain exponential convexity of $t \to \Gamma_t$.

(c). Let $\Gamma_t > 0$. Then, by Corollary 3, we have that Γ_t is log-convex, i.e., $t \to \log \Gamma_t$ is convex. By Lemma 4, we get

$$(t-s) \log \Gamma_r + (r-t) \log \Gamma_s + (s-r) \log \Gamma_t \geq 0 \quad \text{for } 0 < r < s < t < \infty,$$

which is equivalent to (28). □

By the inequality (18) we can give the following definition.

Let X, Y and W be matrices as in Theorem 2 with $X \neq Y$ and $x_{ij}, y_{ij}, w_{ij} > 0$ $(i = 1, 2, ..., n,\ j = 1, 2, ..., m)$. We define

$$M_{u,v} = \left(\frac{\Gamma_u}{\Gamma_v} \right)^{\frac{1}{u-v}} \quad \text{for } 0 < u \neq v < \infty. \tag{29}$$

Remind that $M_{u,v}$ are means of x_{ij} and y_{ij} (see Remark 1). Moreover, we can extend these means in other cases. So by limit we find that

$$M_{u,u} = \exp\left(\frac{\sum_{i=1}^{n} \sum_{j=1}^{m} w_{ij} \left[x_{ij}^u \log x_{ij} - y_{ij}^u \log y_{ij} - y_{ij}^{u-1}(x_{ij} - A \right]}{\sum_{i=1}^{n} \sum_{j=1}^{m} w_{ij} \left[x_{ij}^u - u x_{ij} y^{u-1} + y_{ij}^u (u-1) \right]} - \frac{2u-1}{u(u-1)} \right),\ u \neq 1,$$

$$M_{1,1} = \exp\left(\frac{\sum_{i=1}^{n} \sum_{j=1}^{m} w_{ij} \left[x_{ij} \log^2 x_{ij} - y_{ij} \log^2 y_{ij} - B \right]}{2 \sum_{i=1}^{n} \sum_{j=1}^{m} w_{ij} \left[x_{ij} \log x_{ij} - y_{ij} \log y_{ij} - C \right]} - 1 \right).$$

where $A = y_{ij})(u \log y_{ij} + 1), B = (x_{ij} - y_{ij})(\log y_{ij} + 2) \log y_{ij}$ and $C = (x_{ij} - y_{ij})(\log y_{ij} + 1)$.

Theorem 6. *Let $t, s, u, v \in \mathbb{R}^+$ such that $t \leq u, s \leq v$. Then the following inequality is valid.*

$$M_{t,s} \leq M_{u,v}. \tag{30}$$

Proof. Since by Theorem 5, Γ_t is log-convex, we get

$$\frac{\log \Gamma_s - \log \Gamma_t}{s - t} \leq \frac{\log \Gamma_v - \log \Gamma_u}{v - u}$$

(see [10, p. 2]). Hence we get (30) for $s \neq t$ and $u \neq v$.
For $s = t$ and/or $u = v$, we have the limiting case. □

We now study properties of the Bregman-divergence $t \to \tilde{B}_{n,m,t}(X, Y; W)$, $t \in \mathbb{R}$, introduced in (24) and related to Lemma 2. We also investigate the corresponding Cauchy type means $\tilde{M}_{u,v}$.

Theorem 7. *Let X, Y and W be matrices as in Theorem 2 with $x_{ij}, y_{ij} > 0$ $(i = 1, 2, ..., n,\ j = 1, 2, ..., m)$. Denote*

$$\tilde{\Gamma}_t = \tilde{B}_{n,m,t}(X, Y; W), \qquad t \in \mathbb{R}.$$

Then

(a): *for all $n \in \mathbb{N}$, $p_k \in \mathbb{R}$, $k = 1, 2, ..., n$, the matrix $\left[\tilde{\Gamma}_{\frac{p_k+p_l}{2}}\right]_{k,l=1}^{n}$ is positive semi-definite. In particular,*

$$\det \left[\tilde{\Gamma}_{\frac{p_k+p_l}{2}}\right]_{k,l=1}^{n} \geq 0, \tag{31}$$

(b): *the function $t \to \tilde{\Gamma}_t$ is exponentially convex,*
(c): *if $\tilde{\Gamma}_t > 0$, then the function $t \to \tilde{\Gamma}_t$ is log-convex, i.e.,*

$$(\tilde{\Gamma}_s)^{t-r} \leq (\tilde{\Gamma}_r)^{t-s}(\tilde{\Gamma}_t)^{s-r} \quad \text{for } -\infty < r < s < t < \infty. \tag{32}$$

Proof. The proof is similar to the proof of Theorem 5. □

Let X, Y and W be matrices as in Theorem 2 with $X \neq Y$ and $w_{ij}, x_{ij}, y_{ij} > 0$ $(i = 1, 2, ..., n,\ j = 1, 2, ..., m)$. Analogously to (29) we define

$$\tilde{M}_{u,v} = \left(\frac{\tilde{\Gamma}_u}{\tilde{\Gamma}_v}\right)^{\frac{1}{u-v}} \qquad \text{for } -\infty < u \neq v < \infty. \tag{33}$$

Also, by limit we have

$$\tilde{M}_{u,u} = \exp\left(\frac{\sum_{i=1}^{n}\sum_{j=1}^{m} w_{ij}\left[x_{ij}^{u}\log x_{ij} - y_{ij}^{u}\log y_{ij} - D)\right]}{\sum_{i=1}^{n}\sum_{j=1}^{m} w_{ij}\left[x_{ij}^{u} - ux_{ij}y^{u-1} + y_{ij}^{u}(u-1)\right]} - \frac{2u-1}{u(u-1)}\right),\ u \neq 0, 1,$$

$$\tilde{M}_{0,0} = \exp\left(\frac{\sum_{i=1}^{n}\sum_{j=1}^{m} w_{ij}\left[\log^2 x_{ij} - \log^2 y_{ij} - E\right]}{2\sum_{i=1}^{n}\sum_{j=1}^{m} w_{ij}\left[\log x_{ij} - \log y_{ij} - y_{ij}^{-1}x_{ij} + 1\right]} + 1\right),$$

$$\tilde{M}_{1,1} = \exp\left(\frac{\sum_{i=1}^{n}\sum_{j=1}^{m} w_{ij}\left[x_{ij}\log^2 x_{ij} - y_{ij}\log^2 y_{ij} - F\right]}{2\sum_{i=1}^{n}\sum_{j=1}^{m} w_{ij}\left[x_{ij}\log x_{ij} - y_{ij}\log y_{ij} - G\right]} - 1\right),$$

where $D = y_{ij}^{u-1}(x_{ij} - y_{ij})(u\log y_{ij} + 1, E = 2y_{ij}^{-1}\log y_{ij}(x_{ij} - y_{ij}), F = (x_{ij} - y_{ij})(\log y_{ij} + 2)\log y_{ij}, G = (x_{ij} - y_{ij})(\log y_{ij} + 1)$.

Theorem 8. *Let $t, s, u, v \in \mathbb{R}$ such that $t \leq u, s \leq v$. Then the following inequality is valid.*

$$\tilde{M}_{t,s} \leq \tilde{M}_{u,v}. \tag{34}$$

Proof. The proof is similar to the proof of Theorem 6. □

Finally, we deal with the Bregman-divergence $t \to \overline{B}_{n,m,t}(X, Y; W)$, $t \in \mathbb{R}$, defined in (25) (see also Lemma 3). We also consider the related Cauchy type means $\overline{M}_{u,v}$.

Theorem 9. *Let X, Y and W be matrices as in Theorem 2. Denote*

$$\overline{\Gamma}_t = \overline{B}_{n,m,t}(X, Y; W), \qquad t \in \mathbb{R}.$$

Then

- **(a):** *for all $n \in \mathbb{N}$, $p_k \in \mathbb{R}$, $k = 1, 2, ..., n$, the matrix $\left[\overline{\Gamma}_{\frac{p_k+p_l}{2}}\right]_{k,l=1}^{n}$ is positive semi-definite. In particular,*

$$\det\left[\overline{\Gamma}_{\frac{p_k+p_l}{2}}\right]_{k,l=1}^{n} \geq 0, \tag{35}$$

- **(b):** *the function $t \to \overline{\Gamma}_t$ is exponentially convex,*
- **(c):** *if $\overline{\Gamma}_t > 0$, then the function $t \to \overline{\Gamma}_t$ is log convex, i.e.,*

$$(\overline{\Gamma}_s)^{t-r} \leq (\overline{\Gamma}_r)^{t-s}(\overline{\Gamma}_t)^{s-r} \quad \textit{for } -\infty < r < s < t < \infty. \tag{36}$$

Proof. The proof is similar to the proof of Theorem 5. □

Let X, Y and W be matrices as in Theorem 2 with $X \neq Y$ and $w_{ij} > 0$ $(i = 1, 2, ..., n,\ j = 1, 2, ..., m)$. We define Cauchy type mean $\overline{M}_{u,v}$ of x_{ij} and y_{ij} as follows:

$$\overline{M}_{u,v} = \left(\frac{\overline{\Gamma}_u}{\overline{\Gamma}_v}\right)^{\frac{1}{u-v}} \quad \text{for } -\infty < u \neq v < \infty, \tag{37}$$

$$\overline{M}_{u,u} = \exp\left(\frac{\sum_{i=1}^{n}\sum_{j=1}^{m} w_{ij}[x_{ij}e^{ux_{ij}} - y_{ij}e^{uy_{ij}} - H]}{\sum_{i=1}^{n}\sum_{j=1}^{m} w_{ij}[e^{ux_{ij}} - e^{uy_{ij}} - ue^{uy_{ij}}(x_{ij} - y_{ij})]} - \frac{2}{u}\right), u \neq 0,$$

$$\overline{M}_{0,0} = \exp\left(\frac{\sum_{i=1}^{n}\sum_{j=1}^{m} w_{ij}[x_{ij}^3 - y_{ij}^3 - 3y_{ij}^2(x_{ij} - y_{ij})]}{3\sum_{i=1}^{n}\sum_{j=1}^{m} w_{ij}[x_{ij}^2 - y_{ij}^2 - 2y_{ij}(x_{ij} - y_{ij})]}\right),$$

where $H = e^{uy_{ij}}(uy_{ij} + 1)(x_{ij} - y_{ij})$.

Theorem 10. *Let $t, s, u, v \in \mathbb{R}$ such that $t \leq u, s \leq v$, then the following inequality is valid.*

$$\overline{M}_{t,s} \leq \overline{M}_{u,v}. \tag{38}$$

Proof. The proof is similar to the proof of Theorem 6. □

3. Burbea-Rao divergence for matrices

Remind that the Burbea-Rao divergence $J_{n,m,\phi}(X, W)$ is a special case of the Bregman–divergence $B_{n,m,\phi,\varphi}(X, Y; W)$ with the setting $W = [w_{ij}] = [v_i u_j]$ and $y_{ij} = \sum_{j=1}^{m} u_j x_{ij}$ for $i = 1, 2, ..., n, j = 1, 2, ..., m$ and $\sum_{j=1}^{m} u_j = 1$ (see (8)).

By making use of the results in the previous section, one can easily derive the forthcoming theorems. Their detailed proofs are omitted.

Theorem 11. *Let $\phi : I \subseteq \mathbb{R} \to \mathbb{R}$ be a convex function on the interval I, and let $X = [x_{ij}]$, $W = [u_i v_j]$ be matrices such that $x_{ij} \in I$, $u_i v_j \geq 0$ $(i = 1, 2, ..., n,\ j = 1, 2, ..., m)$ and $\sum_{j=1}^{m} u_j = 1$.*

Then

$$J_{n,m,\phi}(X, W) \geq 0. \tag{39}$$

Moreover, if ϕ is strictly convex on I and $u_i v_j > 0$ for $i = 1, ..., n$, $j = 1, 2, ..., m$, then we have strict inequality in (39).

Proof. The proof follows from Theorem 2. □

Theorem 12. *Let $\phi \in C^2(I)$, where I is closed interval in $\mathbb{R}$, and let $X = [x_{ij}]$, $W = [u_i v_j]$ be matrices such that $x_{ij} \in I$, $u_i v_j > 0$ $(i = 1, 2, ..., n,\ j = 1, 2, ..., m)$ and $\sum_{j=1}^{m} u_j = 1$.*

Then there exists $\xi \in I$ such that

$$J_{n,m,\phi}(X,W) = \frac{\phi''(\xi)}{2} \sum_{i=1}^{n} v_i \left[\sum_{j=1}^{m} u_j x_{ij}^2 - \left(\sum_{j=1}^{m} u_j x_{ij} \right)^2 \right]. \tag{40}$$

Proof. Apply Theorem 3. □

Theorem 13. *Let $\phi, \psi \in C^2(I)$, where I is closed interval in $\mathbb{R}$, and let X, W be matrices as in Theorem 12.*

Then there exists $\xi \in I$ such that

$$\frac{\phi''(\xi)}{\psi''(\xi)} = \frac{J_{n,m,\phi}(X,W)}{J_{n,m,\psi}(X,W)}, \tag{41}$$

provided that the denominators are nonzero.

Proof. See Theorem 4. □

Corollary 4. *Let X, W be matrices as in Theorem 12, where I is a positive closed interval. Then for $-\infty < u \neq 0, 1 \neq v < \infty$, $u \neq v$, there exists $\xi \in I$ such that*

$$\xi^{u-v} = \frac{v(v-1)\sum_{i=1}^{n} v_i \left[\sum_{j=1}^{m} u_j x_{ij}^u - \left(\sum_{j=1}^{m} u_j x_{ij} \right)^u \right]}{u(u-1)\sum_{i=1}^{n} v_i \left[\sum_{j=1}^{m} u_j x_{ij}^v - \left(\sum_{j=1}^{m} u_j x_{ij} \right)^v \right]}. \tag{42}$$

Proof. It is sufficient to set $\phi(x) = x^u$ and $\psi(x) = x^v$, $x \in I$, and to use Theorem 13. □

Remark 2. Since the function $\xi \to \xi^{u-v}$ with $u \neq v$ is invertible, then (42) implies

$$a \leq \left\{ \frac{v(v-1)\sum_{i=1}^{n} v_i \left[\sum_{j=1}^{m} u_j x_{ij}^u - \left(\sum_{j=1}^{m} u_j x_{ij} \right)^u \right]}{u(u-1)\sum_{i=1}^{n} v_i \left[\sum_{j=1}^{m} u_j x_{ij}^v - \left(\sum_{j=1}^{m} u_j x_{ij} \right)^v \right]} \right\}^{\frac{1}{u-v}} \leq b, \tag{43}$$

where $a = \min_{i,j}\{x_{ij}\}$ and $\max_{i,j}\{x_{ij}\} = b$. Thus the expression in the middle of (43) is a mean of x_{ij}.

More generally, if the function $\frac{\phi''}{\psi''}$ in (41) has inverse function, then we deduce that

$$\xi = \left(\frac{\phi''}{\psi''} \right)^{-1} \left(\frac{J_{n,m,\phi}(X,W)}{J_{n,m,\psi}(X,W)} \right). \tag{44}$$

In consequence, the expression on the right hand side of (44) is also a mean of x_{ij}.

Combining (20) with (8) for $x_{ij}, t > 0$ $(i = 1, 2, ..., n,\ j = 1, 2, ..., m)$ gives

$$J_{n,m,t}(X,W) = \begin{cases} \frac{1}{t(t-1)} \sum_{i=1}^{n} v_i \left[\sum_{j=1}^{m} u_j x_{ij}^t - \left(\sum_{j=1}^{m} u_j x_{ij} \right)^t \right], t \neq 1, \\ \\ \sum_{i=1}^{n} v_i \left[\sum_{j=1}^{m} u_j x_{ij} \log x_{ij} - \sum_{j=1}^{m} u_j x_{ij} \log(\sum_{j=1}^{m} u_j x_{ij}) \right], \\ \qquad t = 1. \end{cases} \tag{45}$$

Employing (21) in (8) for $x_{ij} > 0$ and $t \in \mathbb{R}$ $(i = 1, 2, ..., n,\ j = 1, 2, ..., m)$ yields

$$\tilde{J}_{n,m,t}(X,W) = \begin{cases} J_{n,m,t}(X,W)\ , & t \neq 0, \\ \sum_{i=1}^{n} v_i \left[\log(\sum_{j=1}^{m} u_j x_{ij}) - \sum_{j=1}^{m} u_j \log x_{ij} \right] , & t = 0. \end{cases} \tag{46}$$

Likewise, using (22) in (8) for $x_{ij}, t \in \mathbb{R}$ $(i = 1, 2, ..., n,\ j = 1, 2, ..., m)$, leads to

$$\overline{J}_{n,m,t}(X,W) = \begin{cases} \frac{1}{t^2} \sum_{i=1}^{n} v_i \left[\sum_{j=1}^{m} u_j e^{t x_{ij}} - e^{t \sum_{j=1}^{m} u_j x_{ij}} \right] , & t \neq 0, \\ \\ \frac{1}{2} \sum_{i=1}^{n} v_i \left[\sum_{j=1}^{m} u_j x_{ij}^2 - \left(\sum_{j=1}^{m} u_j x_{ij} \right)^2 \right] , & t = 0. \end{cases} \tag{47}$$

Analogously as in Section 2, we now present properties of the Burbea-Rao divergences (45)-(47) and of corresponding Cauchy type means.

Theorem 14. *Let X, W be matrices as in Theorem 12 with $x_{ij} > 0$ $(i = 1, 2, ..., n,\ j = 1, 2, ..., m)$. Denote*

$$\Lambda_t = J_{n,m,t}(X,W), \qquad t > 0.$$

Then

(a): *for all $n \in \mathbb{N}$, $p_k \in \mathbb{R}^+$, $k = 1, 2, ..., n$, the matrix $\left[\Lambda_{\frac{p_k+p_l}{2}} \right]_{k,l=1}^{n}$ is positive semi-definite. In particular,*

$$\det \left[\Lambda_{\frac{p_k+p_l}{2}} \right]_{k,l=1}^{n} \geq 0, \tag{48}$$

(b): *the function $t \to \Lambda_t$ is exponentially convex,*
(c): *the function $t \to \Lambda_t$ is log-convex, i.e.,*

$$(\Lambda_s)^{t-r} \leq (\Lambda_r)^{t-s} (\Lambda_t)^{s-r} \quad \text{for } 0 < r < s < t < \infty. \tag{49}$$

Proof. The proof is similar to the proof of Theorem 5, but using (39) instead of (9). □

Let X, W be matrices as in Theorem 12 with $x_{ij} > 0$ $(i = 1, 2, ..., n,\ j = 1, 2, ..., m)$. We define Cauchy type mean $L_{u,v}$ of x_{ij} as:

$$L_{u,v} = \left(\frac{\Lambda_u}{\Lambda_v}\right)^{\frac{1}{u-v}} \qquad \text{for } 0 < u \neq v < \infty, \tag{50}$$

and by limit we have

$$L_{u,u} = \exp\left(\frac{\sum_{i=1}^n v_i\left[\sum_{j=1}^m u_j x_{ij}^u \log x_{ij} - \left(\sum_{j=1}^m u_j x_{ij}\right)^u \log\left(\sum_{j=1}^m u_j x_{ij}\right)\right]}{\sum_{i=1}^n v_i\left[\sum_{j=1}^m u_j x_{ij}^u - \left(\sum_{j=1}^m u_j x_{ij}\right)^u\right]} - \frac{2u-1}{u(u-1)}\right),\ u \neq 1,$$

$$L_{1,1} = \exp\left(\frac{\sum_{i=1}^n v_i\left[\sum_{j=1}^m u_j x_{ij} \log^2 x_{ij} - \left(\sum_{j=1}^m u_j x_{ij}\right)\log^2\left(\sum_{j=1}^m u_j x_{ij}\right)\right]}{2\sum_{i=1}^n v_i\left[\sum_{j=1}^m u_j x_{ij}\log x_{ij} - \left(\sum_{j=1}^m u_j x_{ij}\right)\log\left(\sum_{j=1}^m u_j x_{ij}\right)\right]} - 1\right).$$

Theorem 15. *Let $t, s, u, v \in \mathbb{R}^+$ such that $t \le u, s \le v$, then the following inequality is valid.*

$$L_{t,s} \le L_{u,v}. \tag{51}$$

Proof. Use Theorem 6. □

Theorem 16. *Let X, W be matrices as in Theorem 12 with $x_{ij} > 0$ $(i = 1, 2, ..., n,\ j = 1, 2, ..., m)$. Denote*

$$\tilde{\Lambda}_t = \tilde{J}_{n,m,t}(X, W), \qquad t \in \mathbb{R}.$$

Then

- **(a):** *for all $n \in \mathbb{N}$, $p_k \in \mathbb{R}$, $k = 1, 2, ..., n$, the matrix $\left[\tilde{\Lambda}_{\frac{p_k+p_l}{2}}\right]_{k,l=1}^n$ is positive semi-definite. In particular,*

$$\det\left[\tilde{\Lambda}_{\frac{p_k+p_l}{2}}\right]_{k,l=1}^n \ge 0, \tag{52}$$

- **(b):** *the function $t \to \tilde{\Lambda}_t$ is exponentially convex,*
- **(c):** *the function $t \to \tilde{\Lambda}_t$ is log convex, i.e.,*

$$(\tilde{\Lambda}_s)^{t-r} \le (\tilde{\Lambda}_r)^{t-s}(\tilde{\Lambda}_t)^{s-r} \quad \text{for } -\infty < r < s < t < \infty. \tag{53}$$

Proof. The proof is similar to the proof of Theorem 5, but using (39) instead of (9). □

Let X, W be matrices as in Theorem 12 with $x_{ij} > 0$ $(i = 1, 2, ..., n,\ j = 1, 2, ..., m)$. We define Cauchy type mean $\tilde{L}_{u,v}$ of x_{ij} as:

$$\tilde{L}_{u,v} = \left(\frac{\tilde{\Lambda}_u}{\tilde{\Lambda}_v}\right)^{\frac{1}{u-v}} \quad \text{for } -\infty < u \neq v < \infty, \tag{54}$$

and, by limit,

$$\tilde{L}_{u,u} = \exp\left(\frac{\sum_{i=1}^n v_i \left[\sum_{j=1}^m u_j x_{ij}^u \log x_{ij}\ -A_1\right]}{\sum_{i=1}^n v_i \left[\sum_{j=1}^m u_j x_{ij}^u\ -\left(\sum_{j=1}^m u_j x_{ij}\right)^u\right]} - \frac{2u-1}{u(u-1)}\right),\ u \neq 0, 1,$$

$$\tilde{L}_{0,0} = \exp\left(\frac{\sum_{i=1}^n v_i \left[\sum_{j=1}^m u_j \log^2 x_{ij}\ -\log^2\left(\sum_{j=1}^m u_j x_{ij}\right)\right]}{2\sum_{i=1}^n v_i \left[\sum_{j=1}^m u_j \log x_{ij}\ -\log\left(\sum_{j=1}^m u_j x_{ij}\right)\right]} + 1\right),$$

$$\tilde{L}_{1,1} = \exp\left(\frac{\sum_{i=1}^n v_i \left[\sum_{j=1}^m u_j x_{ij} \log^2 x_{ij}\ -B_1\right]}{2\sum_{i=1}^n v_i \left[\sum_{j=1}^m u_j x_{ij} \log x_{ij}\ -B_2\right]} - 1\right),$$

where $A_1 = \left(\sum_{j=1}^m u_j x_{ij}\right)^u \log\left(\sum_{j=1}^m u_j x_{ij}\right), \left(\sum_{j=1}^m u_j x_{ij}\right) \log^2\left(\sum_{j=1}^m u_j x_{ij}\right),$ $B_2 = \left(\sum_{j=1}^m u_j x_{ij}\right) \log\left(\sum_{j=1}^m u_j x_{ij}\right)$.

Theorem 17. *Let $t, s, u, v \in \mathbb{R}$ such that $t \leq u, s \leq v$. Then the following inequality is valid.*

$$\tilde{L}_{t,s} \leq \tilde{L}_{u,v}. \tag{55}$$

Proof. The proof is similar to the proof of Theorem 6. □

Theorem 18. *Let X, W be matrices as in Theorem 12. Denote*

$$\overline{\Lambda}_t = \overline{J}_{n,m,t}(X, W), \qquad t \in \mathbb{R}.$$

Then

(a): *for all $n \in \mathbb{N}$, $p_k \in \mathbb{R}$, $k = 1, 2, ..., n$, the matrix $\left[\overline{\Lambda}_{\frac{p_k+p_l}{2}}\right]_{k,l=1}^n$ is positive semi-definite. In particular,*

$$\det\left[\overline{\Lambda}_{\frac{p_k+p_l}{2}}\right]_{k,l=1}^n \geq 0, \tag{56}$$

(b): *the function $t \to \overline{\Lambda}_t$ is exponentially convex,*

(c): *the function $t \to \overline{\Lambda}_t$ is log-convex, i.e.,*

$$(\overline{\Lambda}_s)^{t-r} \leq (\overline{\Lambda}_r)^{t-s}(\overline{\Lambda}_t)^{s-r} \quad \text{for } -\infty < r < s < t < \infty. \tag{57}$$

Proof. The proof is similar to the proof of Theorem 5, but using (39) instead of (9). □

Let X and W be matrices as in Theorem 12. We define Cauchy type mean $\overline{L}_{u,v}$ of x_{ij} as:

$$\overline{L}_{u,v} = \left(\frac{\overline{\Lambda}_u}{\overline{\Lambda}_v}\right)^{\frac{1}{u-v}} \qquad \text{for } -\infty < u \neq v < \infty, \tag{58}$$

$$\overline{L}_{u,u} = \exp\left(\frac{\sum_{i=1}^n v_i \left[\sum_{j=1}^m u_j x_{ij} e^{ux_{ij}} - \sum_{j=1}^m u_j x_{ij} e^{u\sum_{j=1}^m ux_{ij}}\right]}{\sum_{i=1}^n v_i \left[\sum_{j=1}^m u_j e^{ux_{ij}} - e^{\sum_{j=1}^m u_j x_{ij}}\right]} - \frac{2}{u}\right), \; u \neq 0,$$

$$\overline{L}_{0,0} = \exp\left(\frac{\sum_{i=1}^n v_i \left[\sum_{j=1}^m u_j x_{ij}^3 - \left(\sum_{j=1}^m u_j x_{ij}\right)^3\right]}{3\sum_{i=1}^n v_i \left[\sum_{j=1}^m u_j x_{ij}^2 - \left(\sum_{j=1}^m u_j x_{ij}\right)^2\right]}\right).$$

Theorem 19. *Let $t, s, u, v \in \mathbb{R}$ such that $t \leq u, s \leq v$, then the following inequality is valid.*

$$\overline{L}_{t,s} \leq \overline{L}_{u,v}. \tag{59}$$

Proof. The proof is similar to the proof of Theorem 6. □

We conclude Section 3 with the notion of $\mathbf{R}_\phi^h$ divergence.

Definition 3. Let $\phi : I \subseteq \mathbb{R} \to \mathbb{R}$ be a convex function on interval I and h be any increasing convex function on $\mathbb{R}$, and $X = [x_{ij}]$ and $W = [v_i u_j]$ be two matrices with $x_{ij} \in I$, $v_i, u_j \geq 0$ $(i = 1, 2, ..., n,\; j = 1, 2, ..., m)$ and $\sum_{j=1}^m u_j = 1$.

Then the $\mathbf{R}_\phi^h$ *divergence* of X with weight W is defined by

$$\mathbf{R}_\phi^h(X; W) = \sum_{j=1}^m u_j h\left(\sum_{i=1}^n v_i \phi(x_{ij})\right) - h\left(\sum_{i=1}^n v_i \phi\Big(\sum_{j=1}^m u_j x_{ij}\Big)\right). \tag{60}$$

Theorem 20. *Let ϕ, h, X and W be as stated in Definition 3. Then for $\mathbf{R}_\phi^h$ divergence of the matrix X with weight W, defined as in (60), the following inequality holds.*

$$\mathbf{R}_\phi^h(X; W) \geq 0. \tag{61}$$

Proof. Since h is convex, so we can write

$$h\left(\sum_{j=1}^m u_j \sum_{i=1}^n v_i \phi(x_{ij})\right) \leq \sum_{j=1}^m u_j h\left(\sum_{i=1}^n v_i \phi(x_{ij})\right).$$

Hence, by (60), we find that

$$\mathbf{R}_{\phi}^{h}(X;W) \geq h\left(\sum_{j=1}^{m} u_j \sum_{i=1}^{n} v_i \phi(x_{ij})\right) - h\left(\sum_{i=1}^{n} v_i \phi\Big(\sum_{j=1}^{m} u_j x_{ij}\Big)\right). \quad (62)$$

On the other hand, the convexity of ϕ gives

$$\phi\Big(\sum_{j=1}^{m} u_j x_{ij}\Big) \leq \sum_{j=1}^{m} u_j \phi(x_{ij}) \quad \text{for all } i = 1, 2, ..., n.$$

Consequently,

$$\sum_{i=1}^{n} v_i \phi\Big(\sum_{j=1}^{m} u_j x_{ij}\Big) \leq \sum_{i=1}^{n} v_i \sum_{j=1}^{m} u_j \phi(x_{ij}).$$

As h is increasing, we obtain

$$h\left(\sum_{j=1}^{m} u_j \sum_{i=1}^{n} v_i \phi(x_{ij})\right) - h\left(\sum_{i=1}^{n} v_i \phi\Big(\sum_{j=1}^{m} u_j x_{ij}\Big)\right) \geq 0. \quad (63)$$

Now by using (63) in (62), we get $\mathbf{R}_{\phi}^{h}(X;W) \geq 0$, as required. □

4. Power mean and Burbea-Rao divergence

The *power mean* of order p ($p \in \mathbb{R}$) of the positive m-tuple $x_i = (x_{i1}, ..., x_{im}) \in \mathbb{R}_{+}^{m}$, for all $i = 1, 2, ..., n$, with weights $u = (u_1, ..., u_m)$, where $u_j > 0$ for $j = 1, 2, ..., m$, is defined by

$$M_p(x_i) = \begin{cases} \left(\sum_{j=1}^{m} u_j x_{ij}^{p}\right)^{1/p}, & p \neq 0; \\ \prod_{j=1}^{m} x_{ij}^{u_j}, & p = 0. \end{cases}$$

Corollary 5. *Let X, W be matrices as in Theorem 12 with $u_j \geq 0$ $(j = 1, 2, ..., m)$ and with a positive closed interval I. Then for $r, s, l \in \mathbb{R}$ such that $r \neq s \neq l \neq r$, $s, r, l \neq 0$ and $x_{ij} \in I^{1/s} = \{\xi^{1/s} : \xi \in I\}$ $(i = 1, 2, ..., n,\ j = 1, 2, ..., m)$ there exists $\eta \in I^{1/s}$ satisfying*

$$\frac{\sum_{i=1}^{n} v_i \left[M_r^r(x_i) - M_s^r(x_i)\right]}{\sum_{i=1}^{n} v_i \left[M_l^l(x_i) - M_s^l(x_i)\right]} = \frac{r(r-s)}{l(l-s)} \eta^{r-l}, \quad (64)$$

provided that the denominator on the left-hand side of (64) is non-zero.

Proof. By setting $\phi(x) = x^{\frac{r}{s}}$ and $\psi(x) = x^{\frac{l}{s}}$, $x \in I$, in (41) (see Theorem 13), and then replacing $x_{ij}^{1/s}$ and $\xi^{1/s}$ by x_{ij} and η, respectively, we get (64). □

From (64), we can get the following

$$\inf_{\xi\in I}\xi^{1/s}\leq\left(\frac{l(l-s)}{r(r-s)}\frac{\sum_{i=1}^{n}v_i\left[M_r^r(x_i)-M_s^r(x_i)\right]}{\sum_{i=1}^{n}v_i\left[M_l^l(x_i)-M_s^l(x_i)\right]}\right)^{\frac{1}{r-l}}\leq\sup_{\xi\in I}\xi^{1/s},\quad(65)$$

where $r,l,s\in\mathbb{R}$, $r\neq l\neq s$, $s,r,l\neq 0$.

So from (65) we can define a new mean $\Upsilon_{r,l}^s$ as follows:

$$\Upsilon_{r,l}^s=\left(\frac{l(l-s)}{r(r-s)}\frac{\sum_{i=1}^{n}v_i[M_r^r(x_i)-M_s^r(x_i)]}{\sum_{i=1}^{n}v_i\left[M_l^l(x_i)-M_s^l(x_i)\right]}\right)^{\frac{1}{r-l}},\qquad l\neq r\neq s,l,r\neq 0;$$

and in the limiting cases we get the following forms

$$\Upsilon_{r,0}^s=\Upsilon_{0,r}^s=\left(\frac{s\sum_{i=1}^{n}v_i[M_r^r(x_i)-M_s^r(x_i)]}{r(r-s)\sum_{i=1}^{n}v_i[\log M_s(x_i)-\log M_0(x_i)]}\right)^{\frac{1}{r}},\qquad r\neq s,r,s\neq 0;$$

$$\Upsilon_{s,l}^s=\Upsilon_{l,s}^s=\left(\frac{l(l-s)}{s}\frac{\sum_{i=1}^{n}v_i[\sum_{j=1}^{m}u_jx_{ij}^s\log x_{ij}-M_s^s(x_i)\log M_s(x_i)}{\sum_{i=1}^{n}v_i[M_l^l(x_i)-M_s^l(x_i)]}\right)^{\frac{1}{s-l}},\quad l\neq s,l,s\neq 0;$$

$$\Upsilon_{s,0}^s=\Upsilon_{0,s}^s=\left(\frac{\sum_{i=1}^{n}v_i[\sum_{j=1}^{m}u_jx_{ij}^s\log x_{ij}-M_s^s(x_i)\log M_s(x_i)]}{\sum_{i=1}^{n}v_i[\log M_s(x_i)-\log M_0(x_i)]}\right)^{\frac{1}{s}},\qquad s\neq 0;$$

$$\Upsilon_{r,l}^0=\left(\frac{l^2\sum_{i=1}^{n}v_i[M_r^r(x_i)-M_0^r(x_i)]}{r^2\sum_{i=1}^{n}v_i[M_l^l(x_i)-M_0^l(x_i)]}\right)^{\frac{1}{r-l}},\qquad l,r\neq 0;$$

$$\Upsilon_{r,0}^0=\Upsilon_{0,r}^0=\left(\frac{2\sum_{i=1}^{n}v_i[M_r^r(x_i)-M_0^r(x_i)]}{r^2\sum_{i=1}^{n}v_i[M_2^2(\log x_i,\mu)-M_1^2(\log x_i)]}\right)^{\frac{1}{r}},\qquad r\neq 0.$$

$$\Upsilon_{t,t}^s=\exp\left(\frac{\sum_{i=1}^{n}v_i[\sum_{j=1}^{m}u_jx_{ij}^t\log x_{ij}-M_s^t(x_i)\log M_s(x_i)]}{\sum_{i=1}^{n}v_i[M_t^t(x_i)-M_s^t(x_i)]}-\frac{2t-s}{t(t-s)}\right),\qquad t\neq s;$$

$$\Upsilon_{t,t}^0=\exp\left(\frac{\sum_{i=1}^{n}v_i[\sum_{j=1}^{m}u_jx_{ij}^t\log x_{ij}-M_0^t(x_i)\log M_0(x_i)}{\sum_{i=1}^{n}v_i[M_t^t(x_i)-M_0^t(x_i)]}-\frac{2}{t}\right),\qquad t\neq 0;$$

$$\Upsilon_{0,0}^0=\exp\left(\frac{\sum_{i=1}^{n}v_i[\sum_{j=1}^{m}u_j\log^3 x_{ij})-\log^3 M_0(x_i)}{3\sum_{i=1}^{n}v_i[\sum_{j=1}^{m}u_j\log^2 x_{ij}-\log^2 M_0(x_i)}\right),$$

$$\Upsilon_{s,s}^s=\exp\left(\frac{\sum_{i=1}^{n}v_i[\sum_{j=1}^{m}u_jx_{ij}^s\log^2 x_{ij})-M_s^s(x_i)\log^2 M_s(x_i)}{2\sum_{i=1}^{n}v_i[\sum_{j=1}^{m}u_jx_{ij}^s\log x_{ij}-M_s^s(x_i)\log M_s(x_i)]}-\frac{1}{s}\right),\qquad s\neq 0;$$

$$\Upsilon_{0,0}^s=\exp\left(\frac{\sum_{i=1}^{n}v_i[\sum_{j=1}^{m}u_j\log^2 x_{ij}-\log^2 M_s(x_i)]}{2\sum_{i=1}^{n}v_i[\sum_{j=1}^{m}u_j\log x_{ij}-\log M_s(x_i)]}+\frac{1}{s}\right),\qquad s\neq 0.$$

Theorem 21. *Let* $t,r,u,v\in\mathbb{R}$ *such that* $t\leq v,r\leq u$. *Then we have*

$$\Upsilon_{t,r}^s\leq\Upsilon_{v,u}^s\quad\textit{for } s\in\mathbb{R}.\qquad(66)$$

Proof. For $s > 0$, by making use of (46) with p/s and x_{ij} instead of t and $x_{ij}^{1/s}$, we get

$$\tilde{\Lambda}_{p/s} = \begin{cases} \frac{s^2}{p(p-s)} \sum_{i=1}^n v_i[M_p^p(x_i) - M_s^p(x_i)], & p \neq 0, s, \\ s \sum_{i=1}^n v_i[\log M_s(x_i) - \log M_0(x_i)], & p = 0, \\ s \sum_{i=1}^n v_i[\sum_{j=1}^m u_j x_{ij}^s \log x_{ij} - M_s^s(x_i) \log M_s(x_i)], & p = s. \end{cases} \tag{67}$$

Since $t/s \leq v/s$, $r/s \leq u/s$, $t \neq r, v \neq u$, by virtue of Theorem 17 we can write

$$\Big(\frac{\tilde{\Lambda}_{t/s}}{\tilde{\Lambda}_{r/s}}\Big)^{\frac{1}{t-r}} \leq \Big(\frac{\tilde{\Lambda}_{v/s}}{\tilde{\Lambda}_{u/s}}\Big)^{\frac{1}{v-u}}. \tag{68}$$

By combining (67), (68) and the definition of the mean $\Upsilon^s_{\cdot,\cdot}$, we get (66).

For $s < 0$, the proof of (66) is similar as above.

For $s = 0$, we can derive our result by taking limit as $s \to 0$ in (66). Also in this case, we can consider $\overline{\Lambda}_t$ defined as in Theorem 18. By taking x_{ij} in place of $\log x_{ij}$ in (47), and by using Theorem 19, we conclude that

$$\Upsilon^0_{t,r} \leq \Upsilon^0_{v,u}.$$

This completes the proof. □

Remark 3. Let us note that the above results are equivalent to related results for vectors. Namely, observe that for given two matrices $X = [x_{ij}]$ and $Y = [y_{ij}]$ for $i = 1, 2, ..., n, j = 1, 2, ..., m$, as in Definition 1, if we construct the vectors

$$V_x = (x_{11}, x_{12}, ..., x_{1m}, x_{21}, x_{22}, ..., x_{2m}, ..., x_{n1}, x_{n2}, ..., x_{nm})$$

$$V_y = (y_{11}, y_{12}, ..., y_{1m}, y_{21}, y_{22}, ..., y_{2m}, ..., y_{n1}, y_{n2}, ..., y_{nm}),$$

we can deduce the above results by using results for vectors.

References

[1] M. Anwar, J. Jakšetić, J. Pečarić and Atiq ur Rehman, "Exponential convexity, positive semi–definite matrices and fundamental inequalities". J. Math. Inequal., to appear.

[2] M. Anwar and J. E. Pečarić, "New means of Cauchy's type". J. Inequal. and Appl., Vol. 2008, Article ID 163202.

[3] M. Anwar and J. E. Pečarić, "On logarithmic convexity for differences of power means and related results". Math. Inequal. and Appl., 12 (1) (2009), pp. 81–90.

[4] J. Burbea and C. R. Rao, "On Convexity of Some Divergence Measure Based on Entropy Functions". IEEE Transactions on Information Theory, Vol. IT-28, No. 3, May 1982.

[5] S. S. Dragomir, "Some majorisation type discrete inequalities for convex functions". Math. Inequal. and Appl., Vol. 7, No. 2 (2004), 207-216.
[6] M. E. Havrda and F. Charvát, "Quantification method of classification processes: Concept of structural $\alpha-$entropy," Kybernetica, Vol. 3, pp. 30-35, 1967.
[7] E. Hewitt and K. Stromberg, Real and Abstract Analysis, Springer-Verlag, 1965.
[8] J. Jakšetić and J. E. Pečarić, "Means of linear operators for n-convex functions". Math. Inequal. and Appl., to appear.
[9] M. D. C. Pardo and I. Vajda, "On Asymptotic Properties Of Information-Theoretic Divergences." IEEE Transactions on Information Theory, Vol. 49, No. 7, July 2003.
[10] J. E. Pečarić, F. Proschan and Y. L. Tong: Convex Functions, Partial Orderings and Statistcal Applications, Academic Press, New York, 1992.
[11] J. Pečarić and Atiq ur Rehman, "On logarithmic convexity for power sums and related results II". J. Inequal. and Appl., Vol. 2008, Article ID 305623.

1 -Abdus Salam School of Mathematical Sciences, GC University, Lahore, Pakistan
E-mail address: adilbandai@yahoo.com

2 -Department of Applied Mathematics and Computer Science, University of Life Sciences in Lublin, Poland
E-mail address: bniezgoda@wp.pl

3 -University of Zagreb, Faculty of Textile Technology Zagreb, Croatia
E-mail address: pecaric@mahazu.hazu.hr