DIMACS Technical Report 97-04
February 1997

On a Global Optimization Problem in the Study of Information Discrepancy [1]

by

Weiwu FANG
DIMACS Center, Rutgers University
IAM-CAS and APORC[2]

[1]This work has been partially supported by Chinese National Science Foundation under Grant 79370082
[2]IAM-CAS: Institute of Applied Mathematics, Chinese Academy of Sciences
APORC : Asian-Pacific Operations Research Center

DIMACS is a partnership of Rutgers University, Princeton University, AT&T Research, Bellcore, and Bell Laboratories.

DIMACS is an NSF Science and Technology Center, funded under contract STC–91–19999; and also receives support from the New Jersey Commission on Science and Technology.

ABSTRACT

A nonlinear function has been introduced for indexing the disagreement degree of a group of judgment matrices (in FW, 1994). It has many good properties and may be applied in decision making and information process. In this paper, we will discuss a global optimization problem concerned with the global maximum of this function which is constrained on some sets of matrices. Because the size of matrix groups in the problem is arbitrary and the number of local maximum solutions increases exponentially, numerical methods are not suitable and formalized results are desired for the problem. By an approach somewhat similar to the branch and bounded method, we have obtained some formulae on global maximums, a sufficient and necessary condition of the function's taking the maximums, and its some maximum solution sets.

Key Words: discrepancy, entropy, global maximization

1 Introduction

Based on an axiom set, a nonlinear function FDOD has been introduced for indexing disagreement degree of a group of expert judgment matrices in the papers (Fang Weiwu, 1994, hereinafter abbreviated FW). This function has many good properties, such as non-negativity, symmetry, monotonicity, invariance for average, extensive agreement, proportional principle, linear homogeneity, uniform continuity, boundedness and maximality. In the case of vector set, FDOD function is convex function w.r.t. each vector. This function can be applied in decision making and information process.

For example, Kullback-Leibler entropy is one of the most important entropies for measuring information discrepancy between two distributions (Kullback, 1978; Kapur and Kesavan, 1992; Ullah, 1996). In FW (1996) we have compared this function with Kulback-Leibler entropy, as is given in the following table.

Basic properties	Shannon entropy	K-L measure	FDOD function
Basic meaning of concept	Uncertainty, disorder, ... on a distribution	discrepancy of two distributions	discrepancy of a group of distributions
Data amount (number of distributions)	1	2	$s \geq 2$
Non-negativity	Yes	Yes	Yes
Identity		Yes	Yes
Symmetry	Yes	No	Yes
Boundedness	Yes	No	Yes
Uniform continuity	Yes	No	Yes
Upper bound	$\ln t$	∞	$s \cdot \ln t$
Convexity or concavity w.r.t $\mathbf{x} \in \mathbf{\Gamma_t}$	concavity	convexity	convexity
Linear homogeneity		Yes	Yes
Monotonicity			Yes
Invariance for averages			Yes
Limitation for data except for $\sum_{i=1}^{t} x_i = 1$	No	Yes (if $q_i = 0$)	No

The FDOD function not only satisfies the basic properties of K-L measure , but also has other good ones, such as the continuity, monotonicity, boundedness, linear homogeneity, invariance for averages, and so on. For example, let

$$\Gamma_t := \{(x_1, x_2, \cdots, x_t) \mid \sum_{k=1}^{t} x_k = 1 \ \text{ and } \ x_k \geq 0\}, \quad (t = 2, 3, \cdots).$$

and given s distributions $\mathbf{p_i} \in \Gamma_t$ $(i = 1, \cdots, s)$, $s = t$, and

$$\begin{aligned}
\mathbf{p_1} &= (1, 0, \cdots, 0), \\
\mathbf{p_2} &= (0, 1, \cdots, 0), \\
&\vdots \\
\mathbf{p_s} &= (0, 0, \cdots, 1),
\end{aligned}$$

For these distributions, each outcome has complete certainty and all outcomes are completely different from each other, from the meaningfulness of entropy, their discrepancy should have a maximum value. In this case, the K-L method simply can not be used to measure their discrepancies even for any two distributions, but the FDOD function just has the maximum,

$$B(\mathbf{p_1}, \mathbf{p_2}, \cdots, \mathbf{p_s}) = s \cdot \ln s$$

from Theorem 3 in FW (1994). The result is a little surprising because the maximum of Shannon entropy is, as is well known, $\ln s$ for any $\mathbf{p_i}$.

The question naturally arises: how about FDOD function's maximums in the cases of $s \neq t$? This function will be in closer connection with the entropy family if there indeed exists the results similar to $s \cdot \ln s$ for other cases. In addition, the study on maximums of a measure function is always one of the most important topics from measurement theory's point of view. This is why we want to discuss this problem in the paper.

Suppose set $I = \{1, 2, \cdots, t\}$, $J = \{1, 2, \cdots, r\}$ and $K = \{1, 2, \cdots, s\}$ ($s, t \geq 2$ and $r \geq 1$). Let $G(t, s, r)$ or $U := \{U_k | k = 1, \cdots, s\}$ denote a group of matrices, where

$$U_k = [u_{ikj}] := \begin{bmatrix} u_{1k1}, & u_{1k2}, & \cdots, & u_{1kr} \\ \vdots & \vdots & \cdots & \vdots \\ u_{tk1}, & u_{tk2} & \cdots, & u_{tkr} \end{bmatrix}, \quad (k = 1, \cdots, s),$$

$$u_{ikj} \geq 0 \quad \text{and} \quad \sum_{i=1}^{t} \sum_{j=1}^{r} u_{ikj} = 1. \tag{1}$$

We also let $\mathcal{U}(t, s, r) := \{G(t, s, r)\}$, $\mathcal{U}$ is the set of all matrix groups. Further, we let U_k^j denote the j-th column vector $(u_{1kj}, u_{2kj}, \cdots, u_{tkj})^T$ of U_k, and will assume that no U_k^j is a zero "vector".

In FW (1994), the function FDOD, which satisfies seven axioms, for measuring the degree of disagreement of expert judgments is found as the following:

$$B(U_1, U_2, \cdots, U_s) = \sum_{j=1}^{r} BJ_j(U_1^j, U_2^j, \cdots, U_s^j) \tag{2}$$

where

$$\begin{aligned}
BJ_j(U_1^j, U_2^j, \cdots, U_s^j) &= \sum_{k=1}^{s} \sum_{i=1}^{t} u_{ikj} \cdot \ln \frac{u_{ikj} \cdot \sum_{k=1}^{s} \sum_{i=1}^{t} u_{ikj}}{\sum_{k=1}^{s} u_{ikj} \cdot \sum_{i=1}^{t} u_{ikj}} \\
&= \sum_{k=1}^{s} \sum_{i=1}^{t} u_{ikj} \ln \frac{u_{ikj} \cdot \sum_{k=1}^{s} I_{kj}}{K_{ij} \cdot s \cdot I_{kj}},
\end{aligned} \tag{3}$$

$$K_{ij} = \sum_{k=1}^{s} u_{ikj}/s, \; I_{kj} = \sum_{i=1}^{t} u_{ikj}/t, \tag{4}$$

and $0 \cdot \ln(0/0) = 0$ and $0 \cdot \ln 0 = 0$ are defined.

In the formulae mentioned above, $B(t,s,r)$ or $B(U_1, U_2, \cdots, U_s)$ denotes a measure of information discrepancy among the matrices of a $G(t,s,r)$, $B(t,s,r)$ is defined on s $r \times t$ matrices, and B is a function from (t,s,r)-tuples u_{ikj} of non-negative real numbers to a real number. $BJ_j(U_1^j, U_2^j, \cdots, U_s^j)$ denotes a measure of discrepancy on column j, and BJ_j is defined on s column vectors of the matrices.

In this paper, we will investigate the global maximum and its solutions of this nonlinear function defined on the set $\mathcal{U}(t,s,r)$, i.e.

$$\max_{U \in \mathcal{U}} \quad B(U_1, U_2, \cdots, U_s)$$

$$\text{s.t.} \quad u_{ikj} \geq 0 \;\text{ and }\; \sum_{i=1}^{t} \sum_{j=1}^{r} u_{ikj} = 1 \qquad k = 1, 2, \cdots, s.$$

Because the size of matrix groups is arbitrary (t, s,and r are arbitrary) and formalized results are desired, the numerical methods are not suitable for such a problem. So this is a difficult optimization problem in which the number of local maximum solutions will increase exponentially when the size of matrix groups increases.

In Section 2, we will define a function and prove three inequalities. In Section 3, we will define some special matrix groups and get their FDOD values. Both Section 2 and 3 are provided as a convenience in proofs of the following sections. In Section 4, we obtain a sufficient and necessary condition of the function's taking the global maximum $s \cdot \ln t$ by an approach somewhat similar to the branch and bound method. In Section 5, we find out some maximum solution sets for the matrix groups with symmetric structure or $s \leq t$. Finally, for $s > t$ and non-symmetric cases we will illustrate that there is at least some sort of matrix groups, whose FDOD values are close to $s \cdot \ln t$ in Section 6.

2 A function and three inequalities

In this section, we define a continuous function and prove three inequalities as a convenience in the proofs of the following sections. We define

$$f(a,b) := (a+b)\ln(a+b) - a\ln a - b\ln b = a\ln\frac{a+b}{a} + b\ln\frac{a+b}{b},$$

where $(a,b) \in R_+^2$ and $R_+^2 = \{(x,y) \mid x \in R \text{ and } x > 0, y \in R \text{ and } y > 0\}$.

$f(a,b)$ has the following properties:

1. $f(a,b) > 0.$ (5)
2. $f(a,b)$ is a symmetric function.
3. $\lim_{a\to 0} f(a,b) = \lim_{b\to 0} f(a,b) = \lim_{\substack{a\to 0\\ b\to 0}} f(a,b) = 0.$ (6)
4. $\frac{\partial f}{\partial a} = \ln(1+\frac{b}{a}) > 0,\ \frac{\partial f}{\partial b} = \ln(1+\frac{a}{b}) > 0.$
5. $\frac{\partial^2 f}{\partial a^2} = -\frac{b}{a(a+b)} < 0,\ \frac{\partial^2 f}{\partial b^2} = -\frac{a}{b(a+b)} < 0,\ \frac{\partial^2 f}{\partial b \partial a} = \frac{1}{a+b}$, and the Hessian matrix of second-order partial derivatives of $f(a,b)$ is seminegative definite, so $f(a,b)$ is a concavity function.

We use this function to prove three inequalities.

Lemma 1. *Suppose $x > b > 0, y > a > 0$,*

1. if $y - a \leq x - b$, then

$$x^x y^y (a+b)^{(a+b)} > (y+b)^{(y+b)} (x-b)^{(x-b)} a^a b^b; \tag{7}$$

2. if $y - a \geq x - b$, then

$$x^x y^y (a+b)^{(a+b)} > (y-a)^{(y-a)} (x+a)^{(x+a)} a^a b^b. \tag{8}$$

Proof. Given points $O(b;a)$ and $X(x;y) \in R^2_+$ and the function $f(a,b)$ defined on R^2_+, from $x > b > 0$ and $y > a > 0$, we can know that the feasible region of $f(x,y)$ is AOB in Figure 1. Let line OP ($y = x + a - b$) divide AOB into two parts (the line OP is a common region) and $\angle AOP = \angle BOP = 45°$.

1. Further, we assume that the points $O(b;a)$, $X(x;y)$, $X_1(x;y+e), and\ X_2(x-e;y)$ are on the region POA, and $X_1(x;y+e) and\ X_2(x-e;y)$ on OP (see Fig.1). If X is on OP, we have $y = x + a - b$, so $X(x;y) \in POA$ implies $y - a \leq x - b$.

Using the function f, we define a function

$$ff(b,a,x,y) := f(b,a) - f(x,y) + f(x-b, y+b) \tag{9}$$

on POA.

From (9), $\frac{\partial ff}{\partial x} = \ln\frac{x}{x-b} > 0$ and $\frac{\partial ff}{\partial y} = \ln\frac{y}{b+y} < 0$, we can obtain

$$ff(b,a,x_2,y_2) = ff(b,a,x-e,y) \leq ff(b,a,x,y) \text{ and} \tag{10}$$

$$ff(b,a,x_1,y_1) = ff(b,a,x,y+e) \leq ff(b,a,x,y). \tag{11}$$

Suppose $(u;v)$ is a point on the line OP and $(u;v) \neq (b;a)$, i.e., $v = u + a - b$. Let $u = b + r\cos 45°$ and $v = a + r\sin 45°$, according to (9), one gets

$$\begin{array}{l} ff(b,a,u,v) = f(b,a) - f(b + r\cos 45°, a + r\sin 45°) \\ +f(b + r\cos 45° - b, a + b + r\sin 45°) \end{array}.$$

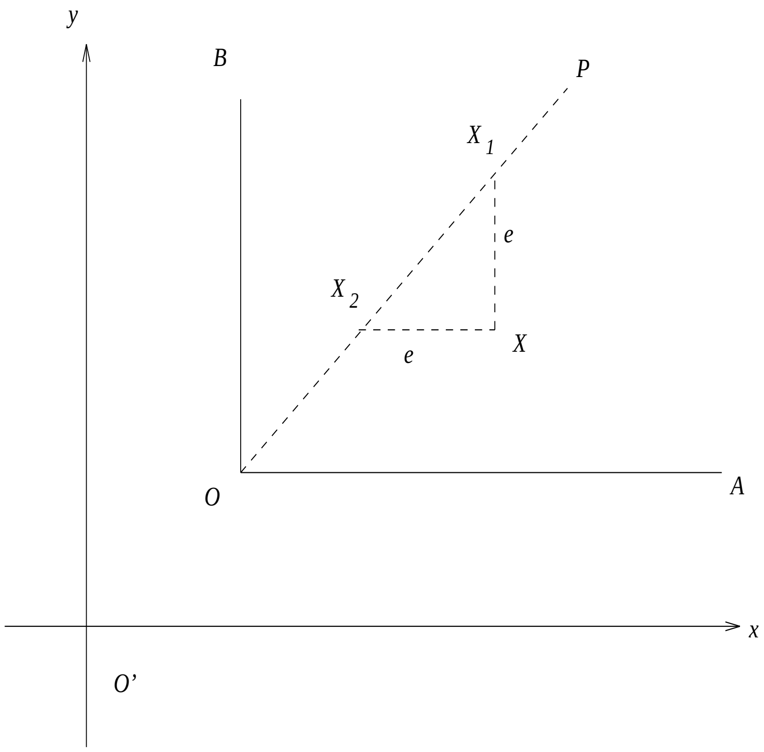

Figure 1: The domains of functions $f(x, y)$ and $ff(a, b, x, y)$

Let $\cos 45^\circ = \sin 45^\circ = c$

$$\begin{aligned} \frac{\partial ff}{\partial r}\Big|_{op} &= -c\ln cr + c\ln(a+cr) + c\ln(b+cr) - c\ln(a+b+cr) \\ &= c\ln\frac{(a+cr)(b+cr)}{(a+b+cr)\cdot cr} = c\ln\frac{c^2r^2+bcr+acr+ab}{c^2r^2+bcr+acr} > 0 \end{aligned} . \tag{12}$$

It implies that ff is an increasing function on the line OP. According to (11) (or (10)), (12) and (6)

$$ff(b,a,x,y) \geq ff(b,a,x,y+e) > \lim_{r\to 0} ff(b,a,u,v) = 0,$$

i.e.,

$$\begin{aligned} ff(b,a,x,y) &= f(b,a) - f(x,y) + f(x-b, y+b) \\ &= (a+b)\ln(a+b) - a\ln a - b\ln b - (x+y)\ln(x+y) + x\ln x + y\ln y \\ &+(x+y)\ln(x+y) - (x-b)\ln(x-b) - (y+b)\ln(y+b) \\ &= \ln\frac{(a+b)^{(a+b)}x^x y^y}{a^a b^b (x-b)^{(x-b)}(y+b)^{(y+b)}} > 0. \end{aligned}$$

So

$$\frac{(a+b)^{(a+b)}x^x y^y}{a^a b^b (x-b)^{(x-b)}(y+b)^{(y+b)}} \ > \ 1,$$

i.e., (7) holds.

2. On the region POB, $y - a \geq x - b$ holds and we define

$$gg(b,a,x,y) := f(b,a) - f(x,y) + f(x+a, y-a).$$

Using the last function and a proof similar to that of 1 of this lemma, Inequality (8) can be obtained.

Remark: It is easy to see that some conditions mentioned in Lemma 1 can be relaxed, i.e., if one use $x > b > 0$, $y \geq a \geq 0$ instead of $x > b > 0$, $y > a > 0$ and define $0 \cdot \ln 0 = 0$, then Inequality (7) also holds; similarly, if one use $x \geq b \geq 0, y > a > 0$ instead of $x > b > 0, y > a > 0$, then Inequality (7) also holds.

Lemma 2. *If* $x > b > 0$, *then*

$$x^x > (x-b)^{(x-b)} b^b.$$

Proof. The proof is similar to Lemma 1. We also can use Lemma 1 to prove it as follows:

We have known that $x > b > 0$, let $y = a \geq 0$, then $y - a \leq x - b$, so

$$x^x y^y (a+b)^{(a+b)} > (y+b)^{(y+b)}(x-b)^{(x-b)} a^a b^b$$

holds from Lemma 1, i.e.,

$$x^x > (x-b)^{(x-b)} b^b.$$

3 Some special matrix groups and their FDOD values

In this section, we define five sorts of disagreement matrix groups and discuss their FDOD values.

Definition 1. *First pseudo-typical disagreement group* PT1(t, s, r)

A $G(t, s, r)$ is called as a *first pseudo-typical disagreement group* denoted by *PT1(t, s, r)* if there exists only one non-zero entry in each column of each matrix.

An example of $PT1(3, 2, 2)$ is

$$\begin{pmatrix} 1/3 & 0 \\ 0 & 2/3 \\ 0 & 0 \end{pmatrix} \quad \begin{pmatrix} 2/3 & 0 \\ 0 & 0 \\ 0 & 1/3 \end{pmatrix}.$$

Remark: In a $PT1(t, s, r)$, for each $j \in J$ together there are s non-zero entries of all matrices, hereinafter "for each $j \in J$" actually means "for all j-th columns of all matrices".

Definition 2. *Second pseudo-typical disagreement group* PT2(t, s, r)

A $PT1(t, s, r)$ is called as a *second pseudo-typical disagreement* denoted by $PT2(t, s, r)$ if for each $j \in J$ there do not exist such two rows i and m ($i, m \in I$ and $m \neq i$), that all $u_{ikj} = 0$ ($k = 1, 2, \cdots, s$) but there are more than one non-zero entries among the set $\{u_{mkj} |\ k = 1, 2, \cdots, s\}$.

The example of Definition 1 is not a $PT2(3, 2, 2)$. An example of $PT2(3, 2, 2)$ is

$$\begin{pmatrix} 1/3 & 0 \\ 0 & 2/3 \\ 0 & 0 \end{pmatrix} \quad \begin{pmatrix} 0 & 0 \\ 2/3 & 0 \\ 0 & 1/3 \end{pmatrix}.$$

Definition 3. *Third pseudo-typical disagreement group $PT3(t, s, r)$*

A $PT2(t, s, r)$ is called as a *third pseudo-typical disagreement* denoted by $PT3(t, s, r)$ if the entries of matrices are distributed in such a manner, that all non-zero entries are exactly located in the same s rows for $s \leq t$ or in t rows for $s \geq t$. The example of Definition 2 is not a $PT3(3, 2, 2)$. An example of $PT3(3, 2, 2)$ is

$$\begin{pmatrix} 1/3 & 0 \\ 0 & 2/3 \\ 0 & 0 \end{pmatrix} \quad \begin{pmatrix} 0 & 1/3 \\ 2/3 & 0 \\ 0 & 0 \end{pmatrix}.$$

Remark.: If $s \geq t$, then a $PT2(t, s, r)$ is also a $PT3(t, s, r)$. In a $PT3(t, s, r)$, suppose all rows with non-zero entries are denoted by $i_1, i_2, \cdots, i_s$, and $s \in I$, then $K_{ij} \neq 0$, for $i = i_1, i_2, \cdots, i_s$ and $j = 1, 2, \cdots, r$.

Definition 4. *Fourth pseudo-typical disagreement group $PT4(t, s, r)$*

A $PT3(t, s, r)$ is called as a *fourth pseudo-typical disagreement group* denoted by PT4(t, s, r) if the entries of matrices are distributed in such a manner, that the total number of non-zero entries, which are in the same row and the same column, of these matrices is $[s/t]$ or $[s/t]+1$ (assume $s = m \cdot t + n$, $[s/t]$ is the greatest integer not greater than s/t, i.e., $[s/t] = m$).

An example of $PT4(3,4,3)$ is

$$\begin{pmatrix} 1/2 & 0 & 0 \\ 0 & 1/4 & 0 \\ 0 & 0 & 1/4 \end{pmatrix} \begin{pmatrix} 0 & 1/3 & 1/3 \\ 1/3 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1/6 \\ 2/3 & 1/6 & 0 \end{pmatrix} \begin{pmatrix} 3/5 & 1/5 & 0 \\ 0 & 0 & 1/5 \\ 0 & 0 & 0 \end{pmatrix}.$$

Remark.: The non-zero entries of different matrices may appear in the same place (the same row and same column). For any $PT4(t,s,r)$, it is easy to understand that for each $j \in J$ there exists n places with $m+1$ non-zero entries, and there exists $t-n$ places with m non-zero entries. For all j's, the product $n \cdot r$ gives the total number of places with $m+1$ non-zero entries, we call these places as the first places of a matrix group; and $r \cdot t - r \cdot n$ gives the total number of places with m, we call them as the second places.

Definition 5. *Typical disagreement group $T(t,s,r)$*

A $PT4(t,s,r)$ is called a *typical disagreement group* denoted by *T(t, s, r)* if the value of each non-zero entry equals to $1/r$.

An example of $T(3,4,3)$ is

$$\begin{pmatrix} 1/3 & 0 & 0 \\ 0 & 1/3 & 0 \\ 0 & 0 & 1/3 \end{pmatrix} \begin{pmatrix} 0 & 1/3 & 1/3 \\ 1/3 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1/3 \\ 1/3 & 1/3 & 0 \end{pmatrix} \begin{pmatrix} 1/3 & 1/3 & 0 \\ 0 & 0 & 1/3 \\ 0 & 0 & 0 \end{pmatrix}.$$

We denote the FDOD value of any group of pseudo-typical disagreement matrices by $PTB(t,s,r)$, and, further, denote the FDOD values of $PT1(t,s,r)$, $PT2(t,s,r)$, $PT3(t,s,r)$, or $PT4(t,s,r)$ by corresponding each $PTB_1(t,s,r)$, $PTB_2(t,s,r)$, $PTB_3(t,s,r)$, or $PTB_4(t,s,r)$.

Theorem 1. *For any group of pseudo-typical disagreement matrices (PT1, PT2, PT3 or PT4),*

$$PTB(r,s,t) = s\sum_{i=1}^{t}\sum_{j=1}^{r} K_{ij} \ln \frac{\sum_{i=1}^{t} K_{ij}}{K_{ij}}, \tag{13}$$

and $PTB(r,s,t)$ is a concave function with respect to $K_{i1}, K_{i2}, \cdots, K_{ir}$.

Proof. For any group of pseudo-typical disagreement matrices (PT1, PT2, PT3 or PT4), there is only one non-zero entry for each column of each matrix, thus we have $u_{ikj} = tI_{kj}$ from (4), and according to (1) and (4), also have $K_{ij} = \sum_{k=1}^{s} \frac{u_{ikj}}{s}$ and $\sum_{k=1}^{s} I_{kj} = \frac{s}{t}\sum_{i=1}^{t} K_{ij}$. Substituting them into the FDOD function (2) in Introduction, it follows that

$$\begin{aligned} PTB(r,s,t) &= \sum_{i=1}^{t}\sum_{j=1}^{r}\sum_{k=1}^{s} u_{ikj} \ln \frac{u_{ikj}\sum_{k=1}^{s} I_{kj}}{s \cdot K_{ij} \cdot I_{kj}} = \sum_{i=1}^{t}\sum_{j=1}^{r}\sum_{k=1}^{s} u_{ikj} \ln \frac{t\sum_{k=1}^{s} I_{kj}}{s \cdot K_{ij}} \\ &= \sum_{i=1}^{t}\sum_{j=1}^{r}\sum_{k=1}^{s} u_{ikj} \ln \frac{\sum_{i=1}^{t} K_{ij}}{K_{ij}} = s\sum_{i=1}^{t}\sum_{j=1}^{r} (\ln \frac{\sum_{i=1}^{t} K_{ij}}{K_{ij}}) \cdot (\sum_{k=1}^{s} \frac{u_{ikj}}{s}) \\ &= s\sum_{i=1}^{t}\sum_{j=1}^{r} K_{ij} \ln \frac{\sum_{i=1}^{t} K_{ij}}{K_{ij}}. \end{aligned}$$

With respect to $K_{i1}, K_{i2}, \cdots, K_{ir}$,

$$\frac{\partial PTB}{\partial K_{ij}} = \ln \frac{\sum_{i=1}^{t} K_{ij}}{K_{ij}} > 0, \quad \frac{\partial^2 PTB}{\partial K_{ij}^2} = \frac{1}{\sum_{i=1}^{t} K_{ij}} - \frac{1}{K_{ij}} < 0, \ \textit{and} \ \frac{\partial^2 PTB}{\partial K_{ij} \partial K_{im}} = 0$$

for $j, m = 1, 2, \cdots, r$, $j \neq m$. So the Hessian matrix of second order partial derivatives of PTB is negative definite, i.e., $PTB(t, s, r)$ is a concave function with respect to $K_{i1}, K_{i2}, \cdots, K_{ir}$.

4 A sufficient and necessary condition on global maximums

In this section, at first we introduce three lemmas concerned with the entropy and the upper bound of FDOD function, and define a sort of special operations, then we use an approach somewhat similar to the branch and bound method to obtain a sufficient and necessary condition of FDOD function's taking the global maximum $s \cdot \ln t$; more detail, we will prove that any group of matrices can be transformed into some special matrix group by a series of special operations, and the FDOD value of any new matrix group obtained by each operation is greater or not less than that of the previous group. Thus, we can investigate the maximization problem on a much smaller subset of set $\mathcal{U}(t, s, r)$.

Lemma 3 (cf. Theorem 2 of FW,1994). *Given t, s and r, an upper bound of the FDOD function of $G(r, s, t)$ is $s \cdot \ln t$.*

Let $\Gamma_t := \{(x_1, x_2, \cdots, x_t) \mid \sum_{k=1}^{t} x_k = 1 \text{ and } x_k \geq 0\}$, $(t = 2, 3, \cdots)$ and

$$H_t(x_1, x_2, \cdots, x_t) := -\sum_{k=1}^{t} x_k \ln x_k$$

where $(x_1, x_2, \cdots, x_t) \in \Gamma_t$, then the following result is well known.

Lemma 4 *(Hartley's entropy, see Aczél and Daróczy book's, 1975)*

$$H_t(x_1, x_2, \cdots, x_t) \leq H_t(\frac{1}{t}, \frac{1}{t}, \cdots, \frac{1}{t}) = \ln t$$

$$\textit{for all} \quad (x_1, x_2, \cdots, x_t) \in \Gamma_t, \textit{and}$$

with equality iff $x_i = \frac{1}{t}$, $(i = 1, \cdots, t)$.

Lemma 5 *If there exists only h non-zero entries in $(x_1, x_2, \cdots, x_t) \in \Gamma_t$, then*

$$H_n(x_1, x_2, \cdots, x_t) \leq \ln h \quad \textit{and}$$

with equality iff $x_i = 1/h$ for all $x_i \neq 0$.

Proof. From Lemma 4 and the N-symmetry and the Expansibility of entropy (Aczél, 1970), where the Expansibility means

$$H_t(x_1, x_2, \cdots, x_{t-1}, 0) = H_{t-1}(x_1, x_2, \cdots, x_{t-1}),$$

the result is immediate.

We also need to define a sort of operations:

Definition 6. *A merging operation MO in column j of a matrix*

Suppose there are two entries $u_{akj} \geq 0$ and $u_{bkj} \geq 0$ in column j of the k-th matrix,

1. if we replace u_{akj} by the sum of u_{akj} and u_{bkj} and replace u_{bkj} by 0, we call it as *a merging operation* $MO(b \to a)$;

2. if we replace u_{bkj} by the sum of u_{akj} and u_{bkj} and replace u_{akj} by 0, we call it as *a merging operation* $MO(a \to b)$.

Each merging operation will produce a new group $\{U'_k\}$ and the symbols corresponding the new group will be denoted by B', BJ'_j, u'_{ikj}, K'_{ij} and I'_{kj}, etc. In the following Lemma 6 and 7 , we will discuss the change of FDOD function value after each operation.

Lemma 6.

Suppose $a, b \in I$, and $a \neq b$, for each merging operation MO$(a \to b)$, or MO$(b \to a)$, it follows that

$$\begin{aligned} B' - B &= BJ'_j - BJ_j \\ &= \sum_{k=1}^{s}(u'_{akj} \ln \frac{u'_{akj}}{K'_{aj}} - u_{akj} \ln \frac{u_{akj}}{K_{aj}}) + \sum_{k=1}^{s}(u'_{bkj} \ln \frac{u'_{bkj}}{K'_{bj}} - u_{bkj} \ln \frac{u_{bkj}}{K_{bj}}). \end{aligned} \tag{14}$$

Proof. From (2), (3) and (4), we know that $I_{kj}, \sum I_{kj}$, and $BJ_m (m \neq j)$ will be not changed for each merging operation , i.e., $I'_{kj} = I_{kj}, \sum I'_{kj} = \sum I_{kj}$, and $BJ'_m = BJ_m (m \neq j)$. Thus,

$$\begin{aligned} B' - B &= BJ'_j - BJ_j \\ &= \sum_{k=1}^{s}\sum_{i=1}^{t} u'_{ikj} \ln \frac{u'_{ikj} \sum_{k=1}^{s} I_{kj}}{s \cdot K'_{ij} I_{kj}} - \sum_{k=1}^{s}\sum_{i=1}^{t} u_{ikj} \ln \frac{u_{ikj} \sum_{k=1}^{s} I_{kj}}{s \cdot K_{ij} I_{kj}} \\ &= \sum_{k=1}^{s}\sum_{i=1}^{t}(u'_{ikj} \ln \frac{u'_{ikj}}{K'_{ij}} - u_{ikj} \ln \frac{u_{ikj}}{K_{ij}}) + \sum_{k=1}^{s}[(\ln \frac{\sum_{k=1}^{s} I_{kj}}{s \cdot I_{kj}})(\sum_{i=1}^{t} u'_{ikj} - \sum_{i=1}^{t} u_{ikj})] \\ &= \sum_{k=1}^{s}\sum_{i=1}^{t}(u'_{ikj} \ln \frac{u'_{ikj}}{K'_{ij}} - u_{ikj} \ln \frac{u_{ikj}}{K_{ij}}) + 0. \end{aligned} \tag{15}$$

For $i \neq a$ and $i \neq b$, $u'_{ikj} = u_{ikj}$ and $K'_{ij} = K_{ij}$, we have

$$u'_{ikj} \ln \frac{u'_{ikj}}{K'_{ij}} = u_{ikj} \ln \frac{u_{ikj}}{K_{ij}}. \tag{16}$$

So (15) reduces to

$$BJ'_j - BJ_j = \sum_{k=1}^{s}(u'_{akj} \ln \frac{u'_{akj}}{K'_{aj}} - u_{akj} \ln \frac{u_{akj}}{K_{aj}}) + \sum_{k=1}^{s}(u'_{bkj} \ln \frac{u'_{bkj}}{K'_{bj}} - u_{bkj} \ln \frac{u_{bkj}}{K_{bj}}).$$

Lemma 7.

Suppose $a, b \in I$, $a \neq b$, u_{akj} and u_{bkj} are two entries of the $k-th$ matrix, and u_{bmj} is an entry of the $m-th$ matrix,

1. If $K_{aj} > u_{akj}/s > 0$ and $K_{bj} > u_{bkj}/s > 0$, then there is at least one operation between $\mathrm{MO}(a \to b)$ and $\mathrm{MO}(b \to a)$, so that the new group of matrices produced by the operation has the property $BJ'_j - BJ_j > 0$;

2. If $K_{aj} = u_{akj}/s > 0$ and $K_{bj} > u_{bkj}/s > 0$, do operation $\mathrm{MO}(b \to a)$; if $K_{aj} > u_{akj}/s > 0$ and $K_{bj} = u_{bkj}/s > 0$, do operation $\mathrm{MO}(a \to b)$. In both cases $BJ'_j - BJ_j > 0$;

3. If $K_{aj} = u_{akj}/s > 0$ and $K_{bj} = u_{bkj}/s > 0$, either $\mathrm{MO}(a \to b)$ or $\mathrm{MO}(b \to a)$ will have $BJ'_j - BJ_j = 0$;

4. If $K_{aj} = 0$, $K_{bj} > u_{bkj}/s > 0$, and $K_{bj} > u_{bmj}/s > 0$, do $\mathrm{MO}(b \to a)$ and no matter which one of u_{bkj} and u_{bmj} is merged into the a-th row of matrices, then in both cases $BJ'_j - BJ_j > 0$;

5. If $K_{aj} = 0$ and $K_{bj} = u_{bkj}/s > 0$, $\mathrm{MO}(b \to a)$ will produce a new group of matrices with $BJ'_j - BJ_j = 0$.

Proof. For $i \neq a$ and $i \neq b$, $u'_{ikj} = u_{ikj}$ and $K'_{ij} = K_{ij}$ hold after a merging operation $\mathrm{MO}(a \to b)$ or $\mathrm{MO}(b \to a)$ is finished.

1.(a). At first, assume

$$K_{aj} - u_{akj}/s \leq K_{bj} - u_{bkj}/s. \tag{17}$$

Doing $\mathrm{MO}(b \to a)$, one gets

$$u'_{akj} = u_{akj} + u_{bkj}, \quad u'_{bkj} = 0, \quad K'_{aj} = K_{aj} + u_{bkj}/s, \quad \text{and } K'_{bj} = K_{bj} - u_{bkj}/s \tag{18}$$

For clarity of presentation, we change the subscript variable k of summations in the formula (14) into the variable l, and we have known that $u'_{alj} = u_{alj}$ and $u'_{blj} = u_{blj}$ also hold for $l \neq k$, thus we can transfer (14) into

$$\begin{aligned} BJ'_j - BJ_j \;&= \Big(\sum_{\substack{l=1\\ l\neq k}}^{s} u_{alj} \ln(K'_{aj})^{-1} + u'_{akj} \ln\frac{u'_{akj}}{K'_{aj}} - \sum_{\substack{l=1\\ l\neq k}}^{s} u_{alj} \ln(K_{aj})^{-1} \\ &\quad - u_{akj} \ln\frac{u_{akj}}{K_{aj}}\Big) + \Big(\sum_{\substack{l=1\\ l\neq k}}^{s} u_{blj} \ln(K'_{bj})^{-1} + u'_{bkj} \ln\frac{u'_{bkj}}{K'_{bj}} - \sum_{\substack{l=1\\ l\neq k}}^{s} u_{blj} \cdot \\ &\quad \ln(K_{bj})^{-1} - u_{bkj} \ln\frac{u_{bkj}}{K_{bj}}\Big). \end{aligned}$$

Substituting (18) into the last formula, one gets

$$
\begin{aligned}
BJ'_j \ -BJ_j =&\\
&\Big(\sum_{\substack{l=1\\ l\neq k}}^{s} u_{alj}\Big)\ln(K'_{aj})^{-1} - \Big(\sum_{\substack{l=1\\ l\neq k}}^{s} u_{alj}\Big)\ln(K_{aj})^{-1} + \Big(\sum_{\substack{l=1\\ l\neq k}}^{s} u_{blj}\Big)\ln(K'_{bj})^{-1}\\
&-\Big(\sum_{\substack{l=1\\ l\neq k}}^{s} u_{blj}\Big)\ln(K_{bj})^{-1} + (u_{akj}+u_{bkj})\ln(u_{akj}+u_{bkj}) + (u_{akj}+u_{bkj})\ln(K'_{aj})^{-1}\\
&-u_{akj}\ln u_{akj} - u_{akj}\ln(K_{aj})^{-1} - u_{bkj}\ln u_{bkj} - u_{bkj}\ln(K_{bj})^{-1}\\
=&\Big(\sum_{\substack{l=1\\ l\neq k}}^{s} u_{alj}\Big)\ln\frac{K_{aj}}{K'_{aj}} + \Big(\sum_{\substack{l=1\\ l\neq k}}^{s} u_{blj}\Big)\ln\frac{K_{bj}}{K'_{bj}} + u_{akj}\ln\frac{K_{aj}}{K'_{aj}}\\
&+u_{bkj}\ln\frac{K_{bj}}{K'_{bj}} - u_{bkj}\ln(K'_{bj})^{-1} + u_{bkj}\ln(K'_{aj})^{-1} + (u_{akj}+u_{bkj})\ln(u_{akj}+u_{bkj})\\
&-u_{akj}\ln u_{akj} - u_{bkj}\ln u_{bkj}\\
=&(\sum_{l=1}^{s} u_{alj})\ln\frac{K_{aj}}{K'_{aj}} + (\sum_{l=1}^{s} u_{blj})\ln\frac{K_{bj}}{K'_{bj}} + u_{bkj}\ln(\frac{K'_{bj}}{K'_{aj}})\\
&+(u_{akj}+u_{bkj})\ln(u_{akj}+u_{bkj}) - u_{akj}\ln u_{akj} - u_{bkj}\ln u_{bkj}\\
=&s(K_{aj}\ln\frac{K_{aj}}{K'_{aj}} + K_{bj}\ln\frac{K_{bj}}{K'_{bj}}) + u_{bkj}\ln\frac{K'_{bj}}{K'_{aj}} + (u_{akj}+u_{bkj})\ln(u_{akj}+u_{bkj})\\
&-u_{akj}\ln u_{akj} - u_{bkj}\ln u_{bkj}\\
=&\ln\frac{(K_{aj})^{s\cdot K_{aj}}\cdot(K_{bj})^{s\cdot K_{bj}}\cdot(K'_{bj})^{u_{bkj}}\cdot(u_{akj}+u_{bkj})^{(u_{akj}+u_{bkj})}}{(K'_{aj})^{s\cdot K_{aj}}\cdot(K'_{bj})^{s\cdot K_{bj}}\cdot(K'_{aj})^{u_{bkj}}\cdot(u_{akj})^{u_{akj}}\cdot(u_{bkj})^{u_{bkj}}}.
\end{aligned}
\tag{19}
$$

Let $K_{aj} = y,\ K_{bj} = x, u_{bkj} = s\cdot B,\ u_{akj} = s\cdot A$ and substitute them into (19); according to (18) and (17), we have $y - A \leq x - B, K'_{aj} = y + B$ and $K'_{bj} = x - B,\ y > A > 0$ and $x > B > 0$; thus, using 1 of Lemma 1 and (19), one gets

$$BJ'_j - BJ_j = s\ln\frac{x^x\cdot y^y\cdot(A+B)^{(A+B)}}{(y+B)^{(y+B)}(x-B)^{(x-B)}\cdot A^A\cdot B^B} \ > \ 0$$

(b). If $K_{aj} - u_{akj}/s \geq K_{bj} - u_{bkj}/s$, doing operation MO($a \to b$), one gets

$$u'_{bkj} = u_{akj} + u_{bkj},\quad u'_{akj} = 0,\quad K'_{aj} = K_{aj} - u_{akj}/s,\quad \text{and}\quad K'_{bj} = K_{bj} + u_{akj}/s. \tag{20}$$

Using a similar proof to that of 1.(a) of this lemma and according to 2 of Lemma 1, we have

$$BJ'_j - BJ_j = s\ln\frac{x^x\cdot y^y\cdot(A+B)^{(A+B)}}{(y-A)^{(y-A)}(x+A)^{(x+A)}\cdot A^A\cdot B^B} \ > \ 0$$

where $y = K_{aj},\ x = K_{bj},\ s\cdot B = u_{bkj},\ s\cdot A = u_{akj}$, and $y - A \geq x - B$ (from (20)).

2. (a). For the case that $K_{aj} = u_{akj}/s > 0$ and $K_{bj} > u_{bkj}/s > 0$, do operation MO($b \to a$), (18) holds, and $K_{aj} - u_{akj}/s = 0 < K_{bj} - u_{bkj}/s$. Using the same proof as that of 1.(a) of this lemma, we also have (19). So one can obtain $BJ'_j - BJ_j \ > \ 0$.

(b). For $K_{aj} > u_{akj}/s > 0$ and $K_{bj} = u_{bkj}/s > 0$, do MO($a \to b$), (20) holds, and $K_{aj} - u_{akj}/s > K_{bj} - u_{bkj}/s = 0$. So from 1.(b) of this lemma, $BJ'_j - BJ_j \ > \ 0$ holds.

3. If $K_{aj} = u_{akj}/s > 0$ and $K_{bj} = u_{bkj}/s > 0$, i.e., $K_{aj} - u_{akj}/s = 0 = K_{bj} - u_{bkj}/s$, do $\mathrm{MO}(a \to b)$ or $\mathrm{MO}(b \to a)$, and assume one does $\mathrm{MO}(b \to a)$, then

$$K'_{aj} = (u_{akj} + u_{bkj})/s, \quad u'_{bkj} = 0, \quad u'_{akj} = u_{akj} + u_{bkj}, \text{ and } K'_{bj} = 0.$$

Using a proof similar to that of 1.(a) of this lemma, we have

$$BJ'_j - BJ_j = s \ln \frac{x^x \cdot y^y \cdot (A+B)^{(A+B)}}{(y+B)^{(y+B)} \cdot A^A \cdot B^B}$$

where $y = K_{aj} = u_{akj}/s = A$, $x = K_{bj} = u_{bkj}/s = B$. So we have $BJ'_j - BJ_j = 0$.

It is the same for doing $\mathrm{MO}(a \to b)$.

4. For $K_{aj} = 0$, $K_{bj} > u_{bkj}/s > 0$, and $K_{bj} > u_{bmj}/s > 0$, assume that u_{bkj} is merged into the a-th row of the k-th matrix, then

$$u'_{akj} = u_{bkj}, \quad u'_{bkj} = 0, \quad K'_{aj} = u_{bkj}/s, \quad \text{and} \quad K'_{bj} = K_{bj} - u_{bkj}/s.$$

Using a proof similar to that of 1.(a) of this lemma, one gets

$$BJ'_j - BJ_j = s \ln \frac{x^x}{(x-B)^{(x-B)} B^B}$$

where $x = K_{bj}$, and $s \cdot B = u_{bkj}$. According to 1 of Lemma 2, we have $BJ'_j - BJ_j > 0$. It is the same for merging u_{bmj} into the a-th row of the m-th matrix.

5. For this case, $u'_{akj} = u_{bkj}$, $u'_{bkj} = 0$, $K'_{aj} = u_{bkj}/s = K_{bj}$, $K'_{bj} = 0$, and $u_{akj} = 0$, so from (14), one gets

$$BJ'_j - BJ_j = u'_{akj} \ln \frac{u'_{akj}}{K'_{aj}} - u_{bkj} \ln \frac{u_{bkj}}{K_{bj}} = 0.$$

We will prove that any matrix group can always be transformed into some pseudo-typical disagreement group, and FDOD value of new group is not less than that of the original group in the following Theorem 2 and Theorem 3.

Theorem 2. *Given $t, s,$ and r, the global maximum among $\{B(t,s,r)\}$ equals to that among $\{PTB_1(t,s,r)\}$; further, if $s \geq t$, the global maximums can be found only among $\{PTB_1(t,s,r)\}$.*

Proof. For any two non-zero entries u_{akj} and u_{bkj} of any column of any matrix in any group of matrices, there exists only three possibilities of relationship among u_{akj}, u_{bkj}, K_{aj} and K_{bj} : (a). $K_{aj} > \frac{u_{akj}}{s} > 0$ and $K_{bj} > \frac{u_{bkj}}{s} > 0$; (b). $K_{aj} = \frac{u_{akj}}{s} > 0$ and $K_{bj} > \frac{u_{bkj}}{s} > 0$, or $K_{aj} > \frac{u_{akj}}{s} > 0$ and $K_{bj} = \frac{u_{bkj}}{s} > 0$; (c). $K_{aj} = \frac{u_{akj}}{s} > 0$ and $K_{bj} = \frac{u_{bkj}}{s} > 0$. According to 1, 2, and 3 of Lemma 7, we can always merge these two non-zero entries into one, and the $B(t,s,r)$ of the new matrices obtained by the merging operation is always not less than that of the original. We can continuously do it until there exists only one non-zero entry in each column. This is a $\mathrm{PT1}(t,s,r)$ and its $PTB_1(t,s,r)$ is not less than that of the original one.

In the case of $s \geq t$, assume that there is a $G(t,s,r) \notin PT1(t,s,r)$, which has the global maximum, thus there must be at least two non-zero entries in some column j of some matrix

k of the $G(t,s,r)$ due to Definition 1, and we can denote the non-zero entries by u_{akj} and u_{bkj}. The possible relationship among u_{akj}, u_{bkj}, K_{aj} and K_{bj} is (a) or (b) or (c) as above. If (a) or (b) holds, we can always merge these two non-zero entries into one , and the $B(t,s,r)$ of the new matrices obtained by the merging operation is greater than that of the original according to 1 or 2 of Lemma 7. If (c) holds, at first, we can do $MO(b \to a)$ (or $MO(a \to b)$), and obtain a new matrix group, $B(t,s,r)$ of which equals to that of the original from 3 of Lemma 7. In this new group, $K_{bj} = 0$, so there must be at least a row i ($i \notin a$ and $i \notin b$) with two non-zero entries u_{ilj}, u_{imj} due to $s \geq t$, which means $K_{ij} > u_{ilj}/s > 0$ and $K_{ij} > u_{imj}/s > 0$; then, we can move u_{ilj} (or u_{imj}) into the row b , and get another new group, $B(t,s,r)$ of which is greater than that of the former group according to 4 of Lemma 7. All of them indicate that $B(t,s,r)$ of the original $G(t,s,r)$ is not the maximum of the problem. In other words, the global maximum can be found only among $PTB_1(t,s,r)$ when $s \geq t$.

Theorem 3. *Given $t, s,$ and r,*

1. the maximum value among $\{B(t,s,r)\}$ equals to that among $\{PTB_3(t,s,r)\}$;

2. for each j of any $PT3(t,s,r)$, the greatest number of non-zero elements in set $\{K_{ij} | i = 1, 2, \cdots, t\}$ is at most

$$t \text{ (} if\ s \geq t \text{)} \quad or \quad s \text{ (} if\ s \leq t \text{).}$$

Proof.

1. For any $PT1(t,s,r)$, if there are $K_{aj} = 0$, $u_{bkj} > 0$, $u_{bmj} > 0$, and $a \neq b$ for some j, then we can obtained a new PT1(t,s,r) with $K'_{aj} \neq 0$ and $K'_{bj} \neq 0$ by moving u_{bkj} or u_{bmj} into the a-th row, and its FDOD value is greater than that of the original one according to 4 of Lemma 7 . We can continuously do it until a PT2(t,s,r) is produced, and its $PTB_2(t,s,r)$ of this PT2(t,s,r) is greater than that of the original $PT1(t,s,r)$. From the Remark of Definition 3, a $PT2(t,s,r)$ is also a $PT3(t,s,r)$ when $s \geq t$. For a PT2(t,s,r) with $s < t$, one can also get a $PT3(t,s,r)$ by moving non-zero entries to some place with zero entries, and its FDOD value of the new group equals to that of the original PT2(t,s,r) according to 5 of Lemma 7. We can continuously do it until a $PT2(t,s,r)$ is transformed into a $PT3(t,s,r)$. All of these imply that FDOD maximum can be found in $\{PTB_3(t,s,r)\}$.

2. Given a $PT3(t,s,r)$, for each $j \in J$ there are only s non-zero entries because there is only one non-zero entry for each column of each matrix. So it is easy to see that if $s < t$, we can at most get s non-zero K_{ij}'s for each j; moreover, if $s \geq t$, the number of non- zero K_{ij}'s is at most t for each j because there are only t rows.

Lemma 8

Given a continuous function $f(x_{11}, \cdots, x_{1j}, x_{21}, \cdots, x_{2j}, \cdots, x_{t1}, \cdots, x_{tj}) = \sum_{i=1}^{t} f_i(x_{i1}, x_{i2}, \cdots, x_{ij})$ defined on $\mathbf{R}^{t \times j}$, if a point $(x_{11}^0, \cdots, x_{1j}^0, x_{21}^0, \cdots, x_{2j}^0, \cdots, x_{t1}^0, \cdots, x_{tj}^0) \in \mathbf{R}^{t \times j}$ is a global maximum solution of the following maximization problem:

$$\max f(x_{11}, \cdots, x_{1j}, x_{21}, \cdots, x_{2j}, \cdots, x_{t1}, \cdots, x_{tj})$$

s.t. $\sum_{i=1}^{t} a_{i1} \cdot x_{i1} + a_{i2} \cdot x_{i2} + \cdots + a_{ij} \cdot x_{ij} = b$, then the point $(x_{i1}^0, x_{i2}^0, \cdots, x_{ij}^0)$ is also a global maximum solution of $f_i(x_{i1}, x_{i2}, \cdots, x_{ij})$ with respect to $x_{i1}, x_{i2}, \cdots, x_{ij}$ on some hyperplane $a_{i1} \cdot x_{i1} + a_{i2} \cdot x_{i2} + \cdots + a_{ij} \cdot x_{ij} = b_i$.

Proof. It is trivial to prove it by reductio ad absurdum. Suppose $(x_{11}^0, \cdots, x_{21}^0, \cdots, x_{t1}^0, \cdots, x_{tj}^0)$ is a global maximum solution and $a_{i1} \cdot x_{i1}^0 + a_{i2} \cdot x_{i2}^0 + \cdots + a_{ij} \cdot x_{ij}^0 = b_i$. If $(x_{i1}^0, x_{i2}^0, \cdots, x_{ij}^0)$ is not a global maximum point of $f_i(x_{i1}, x_{i2}, \cdots, x_{ij})$ on the hyperplane $a_{i1} \cdot x_{i1} + a_{i2} \cdot x_{i2} + \cdots + a_{ij} \cdot x_{ij} = b_i$, then there is a point $(x'_{i1}, x'_{i2}, \cdots, x'_{ij})$, so that $a_{i1} \cdot x'_{i1} + a_{i2} \cdot x'_{i2} + \cdots + a_{ij} \cdot x'_{ij} = b_i$ and $f_i(x_{i1}^0, x_{i2}^0, \cdots, x_{ij}^0) < f_i(x'_{i1}, x'_{i2}, \cdots, x'_{ij})$, and it will make a contradiction to the assumption.

Finally, according to all results before, we can get the formulae and sufficient and necessary condition on global maximums as follows.

Theorem 4. *Given a $PT3(t,s,r)$, a sufficient and necessary condition of its $PTB(t,s,r)$'s taking maximum $s \cdot \ln h$ is*

$$\sum_{j=1}^{r} K_{ij} = \frac{1}{h} \ for\ i = i_1, \cdots, i_h, \quad \sum_{j=1}^{r} K_{ij} = 0 \ for\ i \in I\ and\ i \notin i_1, \cdots, i_h, \tag{21}$$

$$and\ \ K_{i_1 j} = K_{i_2 j} = \cdots = K_{i_h j} \quad (j = 1, \cdots, r) \tag{22}$$

where $h = s$ when $s \le t$ or $h = t$ when $s \ge t$, and $(i_1, \cdots, i_h,)$ is an index subset of I.

Proof. Using (1) and (4), we have

$$\sum_{i=1}^{t} \sum_{j=1}^{r} K_{ij} = 1. \tag{23}$$

From Theorem 3, we have known that the maximum of FDOD function can be found among $\{PTB_3(t,s,r)\}$. We use Lagrange's method to find PTB$(t,s,r)'$s maximum. Using formulae (13) and (23), one gets the Lagrangian function

$$L \equiv -s \cdot \sum_{i=1}^{t} \sum_{j=1}^{r} K_{ij} \ln \frac{\sum_{i=1}^{1} K_{ij}}{K_{ij}} - \lambda(\sum_{i=1}^{t} \sum_{j=1}^{r} K_{ij} - 1).$$

Differentiating with respect to $K_{i1}, K_{i2}, \cdots, K_{ir}$, we get

$$\ln \sum_{i=1}^{t} K_{ij} - \ln K_{ij} - \lambda/s = 0, \quad (j = 1, 2, \cdots, r), \quad i.e.,$$

$$\ln((\sum_{i=1}^{t} K_{ij})/K_{ij}) = \lambda/s, \qquad (j = 1, 2, \cdots, r). \tag{24}$$

From (24), we have

$$\sum_{i=1}^{t} K_{ij} = e^{\lambda/s} K_{ij} \qquad (j = 1, 2, \cdots, r). \tag{25}$$

Using (25) and (23), one gets

$$\begin{aligned} e^{\lambda/s} \textstyle\sum_{j=1}^{r} K_{ij} &= \textstyle\sum_{j=1}^{r} \sum_{i=1}^{t} K_{ij} = 1 \qquad \text{and} \\ \textstyle\sum_{j=1}^{r} K_{ij} &= e^{-\lambda/s}. \end{aligned} \tag{26}$$

Substituting (26) into (25), we have

$$K_{ij} = (\sum_{i=1}^{t} K_{ij})(\sum_{j=1}^{r} K_{ij}) \qquad (j = 1, 2, \cdots, r). \tag{27}$$

Due to $\sum_{i=1}^{t} K_{ij} \neq 0$ (cf. Introduction), (27) also implies that $\sum_{j=1}^{r} K_{ij} \neq 0$ iff $K_{ij} \neq 0$. Because PTB(t, s, r) is a concavity function w.r.t. $K_{i1}, K_{i2}, \cdots, K_{ir}$ (see Theorem 1) and the constraint condition is linear, (27) holds iff PTB(t, s, r) takes the maximum w.r.t. $K_{i1}, K_{i2}, \cdots, K_{ir}$. For the original maximization problem, (27) is a necessary condition from Lemma 8, thus we can substitute (27) into (13) and get

$$PTB(t, s, r) = s \cdot \sum_{j=1}^{r} \sum_{i=1}^{t} K_{ij} \ln \frac{\sum_{i=1}^{t} K_{ij}}{(\sum_{i=1}^{t} K_{ij})(\sum_{j=1}^{r} K_{ij})} = -s \cdot \sum_{i=1}^{t} (\sum_{j=1}^{r} K_{ij})(\ln \sum_{j=1}^{r} K_{ij}).$$

Due to $\sum_{i=1}^{t} \sum_{j=1}^{r} K_{ij} = 1$ and Lemma 5, finally, one gets the global maximum

$$PTB^{*}(t, s, r) = s \cdot (\sum_{i=1}^{h} \frac{1}{h})(\ln h) = s \cdot \ln h \tag{28}$$

iff

$$\sum_{j=1}^{r} K_{ij} = \frac{1}{h} \quad \text{for} \ \ i = i_1, \cdots, i_h \tag{21}$$

where $(i_1, \cdots, i_h,)$ is an index subset of I, and h is such a number as great as possible, that $\sum_{j=1}^{r} K_{ij}$ is exactly not equal to zero for $i \in (i_1, i_2, \cdots, i_h)$. From (27) we have already known that $\sum_{j=1}^{r} K_{ij} \neq 0$ iff $K_{ij} \neq 0$, so h is also such a number as great as possible, that K_{ij} is exactly not equal to zero for $i \in (i_1, i_2, \cdots, i_h)$.

Substituting (21) into (27), it follows that

$$K_{ij} = (\sum_{i=i_1}^{i_h} K_{ij}) \cdot 1/h \qquad (i = i_1, i_2, \cdots, i_h) \ and \ \ (j = 1, \cdots, r). \tag{29}$$

(29) implies

$$K_{i_1 j} = K_{i_2 j} = \cdots = K_{i_h j} \ \ (j = 1, \cdots, r). \tag{22}$$

Thus, (21) and (22) are a sufficient and necessary condition, so that $PTB(t, s, r)$ takes maximum value $s \cdot \ln h$.

For a $PT3(t, s, r)$, we have that $h = t$ when $s \geq t$, and $h = s$ when $s \leq t$ according to Theorem 3.

Remark: Maximum $s \ln t$ here exactly equals to the upper bound in Lemma 3.

5 The cases with symmetry structure or $s \leq t$

In Section 4, we have obtained one sufficient and necessary condition of $PTB(t, s, r)$'s taking the global maximum, here we indicate that there indeed exist such groups in the matrix group sets, that they satisfy conditions (21) and (22).

Given t, s and r, we denote s by $s = m \cdot t + n$ ($m = [s/t]$ is the greatest integer not greater than s/t).

Definition 7. A $G(t, s, r)$ is called as a matrix group with symmetric structure if $s \equiv 0 \pmod t$, or $n \neq 0$ but $n \cdot r \equiv 0 \pmod t$ or $n \cdot r \equiv 0 \pmod s$.

Theorem 5. *For any $t, s,$ and r, the maximum of $\{B(t, s, r)\}$ is indeed*

1. *$s \cdot \ln\ t$ if $s \geq t$ and $\{G(t, s, r)\}$ is the matrix group set with symmetry structure; or*
2. *$s \cdot \ln s$ if $s \leq t$.*

Proof.

1. For $s \geq t$ and a $PT4(t, s, r)$, $s = mt + n$, and $h = t$ from Theorem 3.

A. Assume $s \equiv 0$ (mod t), i.e., $s = m \cdot t$,
further, assume

(a). the total of non-zero entries, which are in the same row and in the same column, of the $TP4(t, s, r)$ is m (it is easily placed for a $TP4(t, s, r)$ or a $T(t, s, r)$) and

(b). every non-zero entry u_{ikj} has the same value $1/r$,

then

$$K_{ij} = \frac{m}{s \cdot r}, \quad \sum_{j=1}^{r} K_{ij} = \frac{m}{s} = \frac{1}{t}, \quad \text{and} \quad \sum_{j=1}^{r}\sum_{i=1}^{t} K_{ij} = 1.$$

Thus, the $PTB_4(t, s, r)$ of this TP4(t, s, r) has the maximum $s \cdot \ln t$ from Theorem 4.

B. Assume $n \neq 0$, $r \geq 1$, and $n \cdot r = 0 \pmod s$,
further, assume

(a). $\dfrac{(m+1) \cdot n \cdot r}{s}$ entries of each matrix are distributed among $n \cdot r$ first places, (cf. the remark of Definition 4) and their values are $a_{m+1} = \dfrac{s}{r \cdot (m+1) \cdot t}$, and

(b). $r - \dfrac{(m+1) \cdot n \cdot r}{s}$ entries of each matrix are distributed among $r \cdot t - r \cdot n$ second places and their values are $a_m = \dfrac{s}{r \cdot m \cdot t}$
(these entries can always be placed because each matrix has r entries,
$r = \frac{r \cdot s}{s} = r \cdot (m \cdot t + n)/s = (r \cdot m \cdot t + n \cdot r)/s = (r \cdot m \cdot t + n \cdot r \cdot (m+1) - (n \cdot r) \cdot m)/s =$
$((n \cdot r) \cdot (m+1) + (r \cdot t - r \cdot n)m)/s > n \cdot r \cdot (m+1)/s$, and
$n \cdot r \cdot (m+1)/s$ is an integer),

then for this $PT4(t, s, r)$, we get

$K_{ij} = \dfrac{m+1}{s} \cdot a_{m+1} = \dfrac{m+1}{s} \cdot \dfrac{s}{r \cdot (m+1) \cdot t} = \dfrac{1}{r \cdot t}$ for first places,
$K_{ij} = \frac{m}{s} \cdot a_m = \frac{m}{s} \cdot \frac{s}{r \cdot m \cdot t} = \frac{1}{r \cdot t}$ for second places, $\sum_{j=1}^{t} K_{ij} = \frac{1}{t}$, and
$\dfrac{(m+1) \cdot n \cdot r}{s} \cdot a_{m+1} + (r - \dfrac{(m+1) \cdot n \cdot r}{s}) \cdot a_m = 1$ (the last formula means that the sum of all entries of each matrix equals to 1).

This $PT4(t, s, r)$ satisfies the conditions (22) and (23), it has the maximum $s \cdot \ln t$ from Theorem 4.

C. Assume $n \neq 0,\ r \geq 1,$ and $n \cdot r = 0(\text{mod t})$,
further, assume

(a). the matrices of this $PT4(t, s, r)$ can be divided into two groups, the first group with n matrices and the second one with $m \cdot t$;
(b). all $n \cdot r$ non-zero entries of the first group are respectively distributed in $n \cdot r$ first places, and the values of all entries are equal to $1/r$; and
(c). $\frac{m \cdot n \cdot r}{m \cdot t}$ entries of each matrix of the second group are distributed in the same places as those of the first group, it can be done because $m \cdot n \cdot r + n \cdot r = (m+1)n \cdot r$, and the entries' values are $a_{m+1} = \frac{s-t}{r \cdot m \cdot t}$. The remaining $(r - \frac{m \cdot n \cdot r}{m \cdot t})$ entries of each matrix are distributed in the second places, , and the entries' values are $a_m = \frac{s}{r \cdot m \cdot t}$.

then one gets

$K_{ij} = \frac{1}{s} \cdot (m \cdot a_{m+1} + \frac{1}{r}) = \frac{1}{r \cdot t}$ for the first places,
$K_{ij} = \frac{1}{s} \cdot m \cdot a_m = \frac{1}{r \cdot t}$ for the second places, $\sum_{j=1}^{r} K_{ij} = \frac{1}{t}$, and
$\frac{m \cdot n \cdot r}{m \cdot t} \cdot a_{m+1} + (r - \frac{m \cdot n \cdot r}{m \cdot t}) \cdot a_m = 1.$

Its FDOD value of the $PT4(t, s, r)$ is $s \cdot \ln\, t$ from Theorem 4.

2. For $s \leq t$ and a $PT3(t, s, r)$, $h = s$ from Theorem 3. If all u_{ikj} have the same value $1/r$ when $u_{ikj} \neq 0$, then all $K_{ij} = \frac{1}{r \cdot s}$ and $\sum_{j=1}^{r} K_{ij} = \frac{1}{s}$, and the result $s \cdot \ln s$ can be obtained from Theorem 4.

6 The cases with non-symmetric structure and $s > t$

In the case of non-symmetric structure and $s > t$, instead of general analysis we compare the FDOD value of typical disagreement cases with $s \cdot \ln t$, the results illustrate that the FDOD value of this sort of matrix groups are close to $s \cdot \ln t$.

The FDOD value of a typical disagreement group $T(t, s, r)$ (see Definition 5) is denoted by $TB(t, s, r)$. We define $TM(t, s, r) = TB(t, s, r)/(s \cdot \ln t)$

Lemma 9 (see Section 6 in FW). $TB(t, s, r) = (t-n)m \ln(s/m) + n(m+1) \ln(s/(m+1))$ *where* $n = s(mod\ \ t)$.

From Lemma 9, we have that $TB(t, s, r)$ equals to the upper value $s \cdot \ln t$ when $s \geq t$ and $s = m * t$ (i.e., $n = 0$), and that $TB(t, s, r)$ equals to the maximum value $s \cdot \ln s$ when $s < t$.

Theorem 6. $1 \geq TM(t, s, r) > 0.9182$

Proof. From Lemma 3 and Lemma 9, one gets $1 \geq TM(t, s, r)$. We prove $TM(t, s, r) > 0.918$ as the following:

1. We assume $s = m \cdot t + n$, and substituting it into $TM(t, s, r)$. So

$$TM(t, s, r) = \frac{(t-n) \cdot m \cdot \ln((mt+n)/m) + (m+1) \cdot n \cdot \ln((mt+n)/(m+1))}{(mt+n) \ln t}. \tag{30}$$

Differentiating (30) with respect to m, one gets

$$\begin{aligned}\frac{\partial TM(t,s,r)}{\partial m} &= \frac{m^2(t-n)(\frac{t}{m}-\frac{mt+n}{m^2})+(1+m)^2\cdot n(\frac{t}{1+m}-\frac{n+mt}{(1+m)^2})}{(mt+n)(mt+n)\ln t}\\ &+\left((t-n)\ln\frac{mt+n}{m}+n\ln\frac{mt+n}{1+m}\right)\frac{1}{(mt+n)\ln t}\\ &-\frac{t(m(t-n)\ln\frac{mt+n}{m}+(1+m)n\ln\frac{mt+n}{1+m})}{(mt+n)^2\ln t}.\end{aligned}$$

Reducing the last formulae, we have obtained

$$\frac{\partial TM(t,s,r)}{\partial m}=(t-n)\cdot n\ln\frac{1+m}{m}/((mt+n)^2\ln t)>0$$

due to $t>n$. It implies that $TM(t,s,r)$ is an increasing function w.r.t. variable m when t and n are arbitrary constants and $t>n$, i.e., $TM(t,mt+n,r)>TM(t,t+n,r)$ when $m>1$.

2. we further assume $m=1$, and from (30) we have

$$TM_1(t,s,r)=\frac{(t-n)\ln(n+t)+2n\ln\frac{t+n}{2}}{(t+n)\ln t}, \tag{31}$$

$$\frac{\partial TM_1(t,s,r)}{\partial n}=\frac{t+n+2t\cdot\ln\frac{n+t}{2}-2t\ln(t+n)}{(t+n)^2\cdot\ln t}=\frac{n+t-2t\cdot\ln 2}{(n+t)^2\cdot\ln t}, \tag{32}$$

and

$$\frac{\partial^2 TM_1(r,s,t)}{\partial n^2}\Big|_{n=t(\ln 4-1)}=-\frac{\ln 4-4\ln 2}{t^2(\ln 4)^3\ln t}>0. \tag{33}$$

According to (32) and (33), when

$$n=2\cdot t\ln 2-t=t(\ln 4-1), \tag{34}$$

(31) will take minimum value w.r.t. variable n. Substituting (34) into (31), one gets

$$\begin{aligned}TM_1(t,s,r) &= \frac{(\ln(t\cdot\ln 4))(t-t(-1+\ln 4))+2t\ln(\frac{t\ln 4}{2})\cdot(-1+\ln 4)}{t(\ln t)\cdot(\ln 4)}\\ &= \frac{\ln(t\ln 4)\cdot\ln 4+2\ln 2(\ln 2-\ln 4)}{(\ln t)\cdot(\ln 4)}\\ &= 1+\frac{\ln(\ln 4)+1-2\ln 2}{\ln t}>1-\frac{0.05967}{\ln t}.\end{aligned} \tag{35}$$

Suppose $t_2>t_1$. Using (35), we have

$$TM_1(t_2,s,r)-TM_1(t_1,s,r)=(\ln(\ln 4)+1-2\ln 2)(\frac{\ln t_1-\ln t_2}{(\ln t_2)\cdot(\ln t_1)})>0$$

due to $\ln t_1-\ln t_2<0$ and $\ln(\ln 4)+1-2\ln 2\simeq -0.5967<0$. This means $TM_1(t,s,r)$ is also an increasing function w.r.t variable t and $n=t(\ln 4-1)$.

3. Thus, in the case that $t=2$ and $n=1$, $TM(2,3,r)\simeq 0.9183$ is the lower bound of all $TM(t,s,r)$. In fact, $TM(3,4,r)=0.9464$, $TM(3,5,r)=0.9602$, $TM(4,5,r)=0.9610$, and $TM(4,7,r)=0.9751,\cdots$.

Acknowledgments

This work has been partially supported by Chinese National Science Foundation under Grant 79370082 and DIMACS Center, and the author thanks F.Roberts, Prof. H.K. Kim. Prof. F. Roush, Prof. G. Torrance, Prof. G. Steiner, and Prof. T. Daniel for their help.

References

References

[1] J.Aczél and Z.Daróczy, On Measures of Information and Their Characterization, Academic Press, New York, London, 1975.

[2] Weiwu Fang, The Disagreement Degree of Multi-person judgments in additive structure, Mathematical Social Sciences, vol. 28, no. 2 (1994) 85-111.

[3] Weiwu Fang, Some properties of FDOD function and kullback measure, CSIAM'96 VII:9, 1996

[4] J.N.Kapur and H.K.Kesavan, Entropy Optimization Principles with Applications. Academic Press, San Diego, 1992.

[5] P.M. Pardalos and J.B. Rosen, Constrained Global Optimization: Algorithms and Applications, Springer-Verlag, Lecture Notes in Computer Sciences 268 (1987).

[6] R. Rockafellar, Convex Analysis, Princeton University Press, Princeton,1970.

[7] A.Ullah, Entropy, divergence and distance measures with econometric applications, J.stati. planning and inference (49), 137-162,1996.