Songklanakarin J. Sci. Technol.
43 (2), 372-383, Mar. - Apr. 2021

http://www.sjst.psu.ac.th

Original Article

A new non-probabilistic divergence measure of fuzzy matrix and its applications

Omdutt Sharma[1], Pratiksha Tiwari[2*], and Priti Gupta[3]

[1] *PDM University, Bahadurgarh, Haryana, 124507 India*

[2] *Delhi Institute of Advanced Studies, Rohini, Delhi, 110085, India*

[3] *Maharshi Dayanand University, Rohtak, Haryana 124001, India*

Received: 17 December 2018; Revised: 6 November 2019; Accepted: 4 February 2020

Abstract

Uncertainties and fuzziness are basic phenomena in human thinking and in many real-world objectives. In the existing literature of information theory various divergence measures are available for studying such phenomena in accordance to their ethos. In general, some are probabilistic and some are non-probabilistic by nature. In the present communication, an attempt is made to introduce non-probabilistic divergence measures for fuzzy matrices that are exponential in nature. In the present study, we prove their validity and study their properties. The different applications of proposed non-probabilistic measure are discussed in the amphitheater of decision making and feature selection.

Keywords: probabilistic measure, non-probabilistic measure, fuzzy set, fuzzy matrix, divergence measure, fuzzy divergence measure, feature selection, decision making

1. Introduction

Statistical and mathematical procedures are significant in modern scientific theories and technology. In many real life problems, the available information may be vague, ambiguous and uncertain. Shannon (1948) used "entropy" to measure the degree of randomness in a probability distribution. The concept of entropy has been widely used in various areas, e.g., communication theory, statistical mechanics, finance, pattern recognition and neural networks, etc. Because of the limitations of Shannon measure in certain situations, Renyi (1961) took the first step and generalized the Shannon measure. After Renyi, many generalized measures were developed for different situations. Kullback & Leibler (1951) initiated the measure of discrimination between two probability distributions, one being the standard and the other an observed distribution. A measure $D(A:B)$ of divergence or cross entropy was found to be significant in mathematical, physical and biological sciences. This measure is probabilistic in character and is characterized as the discrepancy of the probability distribution P from another probability distribution Q. Literature on the development of divergence measures has expanded significantly in the last two decades of the 20th century. Taneja (2001, 2005), Sharma & Mittal (1975), Besseville (2010), Sharma *et al.* (2017, 2018) and Sharma & Gupta (2017, 2019) contributed to the development of generalized information and divergence measures.

The concept of fuzziness introduced by Zadeh (1965) revolutionized research and development in the area. De Luca & Termini (1971) defined the measure of fuzzy entropy corresponding to Shannon (1948) measure of entropy. Bhandari and Pal (1993) defined measure of fuzzy entropy corresponding to Renyi (1961) entropy and measure of fuzzy directed divergence corresponding to Kullback & Leibler (1951) divergence measure. Fuzzy-divergence measures have been studied by Kapur (1997), Gupta *et al.* (2014), and Gupta & Kumari (2014), and Bhatia & Singh (2013) presented a survey of fuzzy information and divergence measures.

*Corresponding author
Email address: parth12003@yahoo.co.in

In this report the first section is introduction, and the second section is preliminaries in which we discuss the related concepts of our works. In third section we propose our divergence measures and show validity and also characterize their properties. In fourth section applications of the proposed measured are discussed, along with a case study example of a decision making problem. Finally, conclusions are drawn.

2. Preliminaries

In this section we discuss some related terms used in this paper.

2.1 Information measure

Shannon (1948) defined the measure of information

$$H(P) = \sum_{i=1}^{n} p_i \, log \, p_i \,, \qquad P \in S$$

where $S = \{P = (p_1, p_2, p_3, \dots \dots \dots, p_n); p_i \geq 0, \sum_{i=1}^{n} p_i = 1;\ n \geq 2\}$ is the complete finite discrete probability distribution.

2.2 Divergence measure

The relative entropy or divergence is a measure of the distance between two probability distributions. In statistics, it arises as the expected logarithm of the likelihood ratio. The relative entropy $D(P:Q)$ is the measure of inefficiency of assuming that the distribution is Q when the true distribution is P. For example, if we knew the true distribution of the random variable, then we could construct a code with average description length $H(P)$. If, instead, we used the code for a distribution Q, we would need $H(P) + D(P:Q)$ bits on the average to describe the random variable. The relative entropy or Kullback & Leibler (1951) distance between two probability distributions is defined as

$$D(P:Q) = \sum_{i=1}^{n} p_i \log \frac{p_i}{q_i}$$

A correct measure of divergence must satisfy the following postulates:

$$D(P:Q) \geq 0$$

$$D(P:Q) = 0 \; iff \; P = Q$$

$D(P:Q)$ is a convex function of both p_i and q_i where $p_i \in P$ and $q_i \in Q$.

$D(P:Q) = D(Q:P)$ that is the measure is symmetric.

Thus properties (2.2.1) to (2.2.4) are essential to define a new divergence measure. If in addition triangle inequality is also satisfied by $D(P:Q)$ then it is called a distance measure.

2.3 Fuzzy sets

In mathematics, fuzzy sets are sets whose elements have degree of membership. Fuzzy sets were introduced by Zadeh (1965) as an extension of the classical notion of a set. In classical set theory, the membership of elements in a set is either 0 or 1 i.e. an element either belongs (1) or does not belong to the set (0). By contrast, in a fuzzy set the membership of elements in a set lies in the interval $[0, 1]$. Thus we say that it is a class of objects with a continuum of grades of membership i.e. a fuzzy set is defined by its membership function which allots to each object a grade of membership ranging between zero and one. Since the indicator functions of classical sets are a particular case of the membership functions of fuzzy sets, only taking values 0 or 1, we say that fuzzy sets generalize classical sets. In fuzzy set theory, classical bivalent sets are usually called crisp sets.

Definition 1. A fuzzy set is a pair (U, m) where U is a universal set and $m: U \to [0, 1]$ for each $z \in U$, the value $m(z)$ is called the grade of membership of z in (U, m). For a finite set $U = \{z_1, z_2, \dots \dots, z_n\}$, the fuzzy set (U, m) is often denoted by $\left\{\frac{m(z_1)}{(z_1)}, \dots \dots, \frac{m(z_n)}{(z_n)}\right\}$. Let $z \in U$. Then z is called not included in the fuzzy set (U, m) if $m(z) = 0$, z is called a fuzzy number if $0 < m(z) < 1$. The set $\{z \in U \mid m(z) > 0\}$ is called the support of (U, m) and the set $\{z \in U \mid m(z) = 1)$ is called its kernel or core. The function m is called the membership function of the fuzzy set (U, m).

2.3.1 Standard operator and operation of fuzzy sets

With the min-max system proposed by Zadeh (1965), fuzzy set operators are defined component-wise as:

Complement: The complement of a set 'S' is denoted as S^c. Membership degree can be calculated as following:

$$m_{S^c}(z) = 1 - m_S(z); \quad \forall\, z \in Z.$$

Union: Membership value of member x in the union takes the maximum value of membership between S and T:

$$m_{S \cup T}(Z) = \max[m_S(z),\ m_T(z)]\,; \ \forall\, z \in Z.$$

Two alternative choices of membership function for the union ($S \cup T$) are:

$$m_{S \cup T}(Z) = m_S(z) + m_T(z) - m_S(z) * m_T(z); \ \forall\, z \in Z.$$

$$m_{S \cup T}(Z) = \min\{1, m_S(z) + m_T(z)\}\,; \ \forall\, z \in Z.$$

Intersection: Intersection of fuzzy sets S and T takes minimum value of membership function between S and T:

$$m_{S \cap T}(Z) = \min[m_S(z), m_T(z)]; \ \forall z \in Z.$$

Two alternative choices of membership function for the intersection $S \cap T$ are:

$$m_{S \cap T}(Z) = \{m_S(z) * m_T(z)\}; \ \forall z \in Z.$$

$$m_{S\cap T}(Z) = \max\{0, m_S(z) + m_T(z) - 1\};\ \forall z \epsilon Z.$$

The complement, union and intersection are applicable even if the membership function is restricted to 0 or 1.

2.4 Fuzzy divergence measure

As we have already discussed, the measure of divergence or cross entropy is the difference between two sets. Thus, taking two fuzzy sets Bhandari & Pal (1993) initiated the first fuzzy directed divergence measure corresponding to Kullback & Leibler's (1951) divergence, defined as:

$$D(A:B) = \sum_{i=1}^{n}[\mu_A(x_i)\, log\frac{\mu_A(x_i)}{\mu_B(x_i)} + (1-\mu_A(x_i))log\frac{1-\mu_A(x_i)}{1-\mu_B(x_i)}]$$

Later various studies related to fuzzy directed divergence have been published. Jain & Chhabra (2016) defined a new exponential measure of fuzzy directed divergence based on Bajaj & Hooda (2010) as:

$$G_{exp}(A,B) = \sum_{i=1}^{n}[\mu_A(x_i) - \mu_B(x_i)]\left[e^{\frac{\mu_A(x_i)}{\mu_B(x_i)}} - e^{\frac{1-\mu_A(x_i)}{1-\mu_B(x_i)}}\right]$$

Definition 2. Let a universal set be X and $F(X)$ be the set of all fuzzy subsets. A mapping $D : F(X) \times F(X) \to R$ is called a divergence between fuzzy subsets if and only if the following axioms hold:

$D(A:B) \geq 0$, i.e. divergence measure is non negative.

$D(A:B) = 0$ when A=B, i.e. divergence measure is equal for two equal fuzzy set.

$D(A:B) = D(B:A)$, i.e. divergence measure is symmetric in nature.

$D(A:B)$ divergence measure is convex in A and B.

A measure that satisfies above axioms is a valid divergence measure.

2.5 Fuzzy matrices

Definition 3. Fuzzy matrix**:** A fuzzy matrix S of order $m \times n$ is defined as $S = [< s_{ij}, s_{ij\mu} >]_{m\times n}$ where $s_{ij\mu}$ is the membership value of the element s_{ij} in S. For our convinence, we write S as $S = [s_{ij}]_{m\times n}$.

Definition 4. Boolean fuzzy matrix**:** A fuzzy matrix $S = [s_{ij}]_{m\times n}$ is said to be a Boolean fuzzy matrix or crisp matrix of order m×n if all the elements of S are either 0 or 1.

Definition 5. Most fuzzy matrix**:** A fuzzy matrix $S = [s_{ij}]_{m\times n}$ is said to be the most fuzzy matrix of order m×n if all the elements of S are 0.5.

Definition 6. Two fuzzy matrices are equal if they have same order and their corresponding elements are equal.

Definition 7. Let $S = [s_{ij\mu}] \in [F(M)]_{m\times n}$, where $s_{ij\mu}$ is the membership value of element s_{ij}. If m≠n, then S is called a fuzzy rectangular matrix.

Definition 8. Let $S = [s_{ij\mu}] \in [F(M)]_{m\times n}$, where $s_{ij\mu}$ is the membership value of element s_{ij}. If m = n, then S is called a fuzzy square matrix.

Definition 9. Let $S = [s_{ij\mu}] \in [F(M)]_{m\times n}$, where $s_{ij\mu}$ is the membership value of element s_{ij}. If m = 1, then S is called a fuzzy row matrix.

Definition 10. Let $S = [s_{ij\mu}] \in [F(M)]_{m\times n}$, where $s_{ij\mu}$ is the membership value of element s_{ij}. If n = 1, then S is called a fuzzy column matrix.

Definition 11. Let $S = [s_{ij\mu}] \in [F(M)]_{m\times n}$, where $s_{ij\mu}$ is the membership value of element s_{ij}. If n = m, and $s_{ij\mu} = 0$ for all i ≠ j then S is called a fuzzy diagonal matrix.

Definition 12. Let $S = [s_{ij\mu}] \in [F(M)]_{m\times n}$, where $s_{ij\mu}$ is the membership value of element s_{ij}. If n = m, and $s_{ij\mu} = 0$ for all i ≠ j and $s_{ij\mu} = \omega \in [0,1]\ \forall\ i = j$ then S is called a fuzzy scalar matrix.

Definition 13. Let $S = [s_{ij\mu}] \in [F(M)]_{m\times n}$, where $s_{ij\mu}$ is the membership value of element s_{ij}. If m = n, and $s_{ij\mu} = 0$ for all i > j then S is called a fuzzy upper triangular matrix.

Definition 14. Let $S = [s_{ij\mu}] \in [F(M)]_{m\times n}$, where $s_{ij\mu}$ is the membership value of element s_{ij}. If m = n, and $s_{ij\mu} = 0$ for all I < j then S is called a fuzzy lower triangular matrix.

A fuzzy matrix is said to be triangular if it is either fuzzy lower or fuzzy upper triangular matrix.

Definition 15. Let $S = [s_{ij\mu}] \in [F(M)]_{m\times n}$, where $s_{ij\mu}$ is membership value of element s_{ij}. Then the elements $s_{11}, s_{22}, \ldots, s_{mm}$ are called the diagonal elements and the line along which they lie is called the principal diagonal of the fuzzy matrix.

Definition 16. Let $S = [s_{ij\mu}]$, $T = [t_{ij\mu}] \in [F(M)]_{m\times n}$. Then union of S, T is defined by $S_{m\times n} \cup T_{m\times n} = R_{m\times n} = [r_{ij\mu}]_{m\times n}$, where $r_{ij\mu} = s_{ij\mu} + t_{ij\mu} - s_{ij\mu}t_{ij\mu}$ for all i and j.

Example 2.5.1.

Let $S = \begin{bmatrix}0.7 & 0.5 & 0.2\\ 0.3 & 0.2 & 0.1\\ 0.3 & 0.7 & 0.6\end{bmatrix}$ and $T = \begin{bmatrix}0.2 & 0.3 & 0.1\\ 0.5 & 0.2 & 0.6\\ 0.8 & 0.7 & 0.2\end{bmatrix}$; then

$$S_{3\times3} \cup T_{3\times3} = R_{3\times3} = \begin{bmatrix} 0.76 & 0.65 & 0.28 \\ 0.65 & 0.36 & 0.64 \\ 0.86 & 0.91 & 0.68 \end{bmatrix}$$

Definition 17. Let $S = [s_{ij\mu}],\ T = [t_{ij\mu}] \in [\mathbb{F}(\mathrm{M})]_{m\times n}$. Then maximum operation of S, T is defined by $Max(S_{m\times n}, T_{m\times n}) = R_{m\times n} = [r_{ij\mu}]_{m\times n}$, where $r_{ij\mu} = \max(p_{ij\mu}, q_{ij\mu})$ for all i and j.

Example 2.5.2. Hence for the matrix S and T given in example (2.5.1) we have,

$$Max(S_{3\times3}, T_{3\times3}) = R_{3\times3} = \begin{bmatrix} 0.7 & 0.5 & 0.2 \\ 0.5 & 0.2 & 0.6 \\ 0.8 & 0.7 & 0.6 \end{bmatrix}$$

Definition 18. Let $S = [s_{ij\mu}],\ T = [t_{ij\mu}] \in [\mathbb{F}(\mathrm{M})]_{m\times n}$. Then intersection of S, T is defined by $S_{m\times n} \cap T_{m\times n} = R_{m\times n} = [r_{ij\mu}]_{m\times n}$, where $r_{ij\mu} = s_{ij\mu} * t_{ij\mu} = s_{ij\mu}t_{ij\mu}$ for all i and j.

Example 2.5.6. Hence for the matrix S and T given in example (2.5.1) we have,

$$S_{3\times3} \cap T_{3\times3} = R_{3\times3} = \begin{bmatrix} 0.14 & 0.15 & 0.02 \\ 0.15 & 0.04 & 0.06 \\ 0.24 & 0.49 & 0.12 \end{bmatrix}$$

Definition 19. Let $S = [s_{ij\mu}],\ T = [t_{ij\mu}] \in [\mathbb{F}(\mathrm{M})]_{m\times n}$. Then minimum operation of S, T is defined by $Min(S_{m\times n}, T_{m\times n}) = R_{m\times n} = [r_{ij\mu}]_{m\times n}$, where $r_{ij\mu} = \min(p_{ij\mu}, q_{ij\mu})$ for all i and j.

Example 2.5.3. Hence for the matrix S and T given in example (2.5.1) we have,

$$Min(S_{3\times3}, T_{3\times3}) = R_{3\times3} = \begin{bmatrix} 0.2 & 0.3 & 0.1 \\ 0.3 & 0.2 & 0.1 \\ 0.3 & 0.7 & 0.2 \end{bmatrix}$$

Some experts used Max operation as Union and Addition of two fuzzy matrices and Min operation as Intersection and Subtraction of fuzzy matrices.

Definition 20. Let $S = [s_{ij\mu}],\ T = [t_{ij\mu}] \in [\mathbb{F}(\mathrm{M})]_{m\times n}$. Then Max-min operation of S,T is defined by $Max - min(S_{m\times n}, T_{m\times n}) = R_{m\times n} = [r_{ij\mu}]_{m\times n}$, where $r_{ij\mu} = \max\{\min[(s_{ij\mu}, t_{ji\mu})\ for\ j = 1\ to\ n]\}\ for\ i = 1\ to\ m.$

Example 2.5.4. Hence for the matrix S and T given in example (2.5.1) we have,

$$Max - min(S_{3\times3}, T_{3\times3}) = R_{3\times3} = \begin{bmatrix} 0.5 & 0.3 & 0.5 \\ 0.2 & 0.3 & 0.2 \\ 0.6 & 0.6 & 0.6 \end{bmatrix}$$

Definition 21. Let $S = [s_{ij\mu}],\ T = [t_{ij\mu}] \in [\mathbb{F}(\mathrm{M})]_{m\times n}$. Then Min-max operation of S,T is defined by $Min - max(S_{m\times n}, T_{m\times n}) = R_{m\times n} = [r_{ij\mu}]_{m\times n}$, where $r_{ij\mu} = \min\{\max[(s_{ij\mu}, t_{ji\mu})\ for\ j = 1\ to\ n]\}\ for\ i = 1\ to\ m.$

Example 2.5.5. Hence for the matrix S and T given in example (2.5.1) we have,

$$Min - max(S_{3\times3}, T_{3\times3}) = R_{3\times3} = \begin{bmatrix} 0.5 & 0.5 & 0.2 \\ 0.3 & 0.2 & 0.2 \\ 0.3 & 0.3 & 0.3 \end{bmatrix}$$

Some experts adopt Max-min and Min-max operation as a product of two fuzzy soft matrices according to their requirements

Definition 22. Let $S = [s_{ij\mu}],\ R = [r_{ij\mu}] \in [\mathbb{F}(\mathrm{M})]_{m\times n}$, then R is complement of S denoted by $S^c = R = [r_{ij\mu}]$, where $r_{ij\mu} = 1 - s_{ij\mu}$ for all i and j.

Example 2.5.7. Hence for the matrix S given in example (2.5.1) we have,

$$S^c = R = \begin{bmatrix} 0.3 & 0.5 & 0.8 \\ 0.7 & 0.8 & 0.9 \\ 0.7 & 0.3 & 0.4 \end{bmatrix}$$

Definition 23. Let $S = [s_{ij\mu}] \in [\mathbb{F}(\mathrm{M})]_{m\times n},\ R = [r_{ji\mu}] \in [\mathbb{F}(\mathrm{M})]_{n\times m}$, then R is transpose of S denoted by $S^T = R = [r_{ji\mu}]$, where $r_{ji\mu} = s_{ij\mu}$ for all i and j.

Example 2.5.8. Hence for the matrix S in example (2.5.1) we have,

$$S^T = R = \begin{bmatrix} 0.7 & 0.3 & 0.3 \\ 0.5 & 0.2 & 0.7 \\ 0.2 & 0.1 & 0.6 \end{bmatrix}$$

3. Proposed Divergence measure

In this section non-probabilistic divergence measures of fuzzy matrices are proposed and their axioms based on fuzzy divergence measure are also discussed.

Here we first define a non-probabilistic divergence measure of fuzzy matrices:

Definition 24. Let F_M be the set of all fuzzy matrices having m rows and n columns and X & $Y \in F_M$. Then a mapping $J :\ F_M \times F_M \rightarrow R$ is called non-probabilistic divergence measure of fuzzy matrices if and only if

a. $J(X{:}Y) \geq 0$, i.e. divergence measure is non-negative
b. $J(X{:}Y) = 0$ when X=Y i.e. $x_{ij} = y_{ij}$, i. e., divergence measure is zero when fuzzy matrix are equal
c. $J(X{:}Y) = J(Y{:}X)$ i.e. divergence measure is symmetric in nature.
d. $J(X{:}Y)$ divergence measure is convex in X and Y.

A measure is a non-probabilistic divergence measure of fuzzy matrices if it satisfies axioms (a) to (d).

Here an exponential divergence measure of fuzzy matrices is proposed as

$$J(X{:}Y) = \sum_{i=1}^{m}\sum_{j=1}^{n}\{1 - (x_{ij} - y_{ij})\}e^{(1-x_{ij})-(1-y_{ij})} + \left\{1 - \left((1 - x_{ij}) - (1 - y_{ij})\right)\right\}e^{x_{ij}-y_{ij}} - 2$$

$$J(X:Y)=\sum_{i=1}^{m}\sum_{j=1}^{n}[(1-x_{ij}+y_{ij})e^{(y_{ij}-x_{ij})}+(1-y_{ij}+x_{ij})e^{x_{ij}-y_{ij}}-2] \quad (3.1)$$

where X and Y are fuzzy matrices of order m × n and $x_{ij} \in X$ & $y_{ij} \in Y$

Now to show that the proposed measure is a valid measure, the following theorem show satisfaction of the required axioms:

Theorem 3.1. J(X:Y) is non-negative if $X\ and\ Y \in [F_M]_{m\times n}$.

Proof. It can be clearly shown in Figure (1) that the measure is non-negative for each a & b (where a $= x_{ij}X$ & $b = y_{ij} \in Y$).

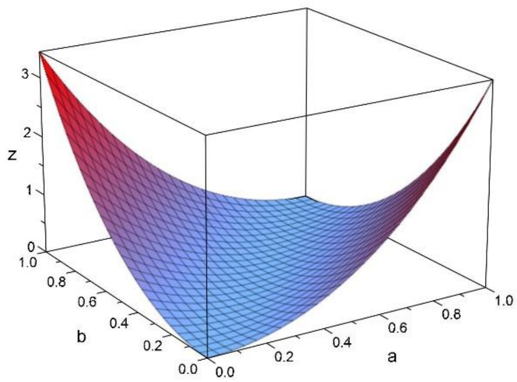

Figure 1. Divergence

Theorem 3.2. J(X:Y)=0 when X & Y are equal fuzzy matrices, or X=Y or $x_{ij} = y_{ij}$.

Proof.

$$J(X:Y)=\sum_{i=1}^{m}\sum_{j=1}^{n}[(1-x_{ij}+y_{ij})e^{(y_{ij}-x_{ij})}+(1-y_{ij}+x_{ij})e^{x_{ij}-y_{ij}}-2]$$

And if X=Y or $x_{ij} = y_{ij}$ then

$$J(X:Y)=\sum_{i=1}^{m}\sum_{j=1}^{n}\left[(1-x_{ij}+x_{ij})e^{(x_{ij}-x_{ij})}+(1-x_{ij}+x_{ij})e^{x_{ij}-x_{ij}}-2\right]=0$$

Hence the claim is proven.

Theorem 3.3. J(X:Y) is symmetric, i.e. $J(X:Y) = J(Y:X)$.

Proof. To prove that J(X:Y) is symmetric we show that

$$J(X:Y) - J(Y:X) = 0$$

$$J(X:Y)=\sum_{i=1}^{m}\sum_{j=1}^{n}\left[(1-x_{ij}+y_{ij})e^{(y_{ij}-x_{ij})}+(1-y_{ij}+x_{ij})e^{x_{ij}-y_{ij}}-2\right]$$

$$J(Y:X)=\sum_{i=1}^{m}\sum_{j=1}^{n}[(1-y_{ij}+x_{ij})e^{(x_{ij}-y_{ij})}+(1-x_{ij}+y_{ij})e^{y_{ij}-x_{ij}}-2]$$

$$J(X:Y)-J(Y:X)=\sum_{i=1}^{m}\sum_{j=1}^{n}\left[(1-x_{ij}+y_{ij}-1+x_{ij}-y_{ij})e^{(y_{ij}-x_{ij})}+(1-y_{ij}+x_{ij}-x_{ij}+y_{ij}-1)e^{x_{ij}-y_{ij}}-2+2\right]=0$$

Hence the claim is proven.

Theorem 3.4. J(X:Y) is convex in X and Y.

Proof. First we show J(X:Y) convex in X and similarly we prove for Y.

$$J(X:Y)=\sum_{i=1}^{m}\sum_{j=1}^{n}\left[(1-x_{ij}+y_{ij})e^{(y_{ij}-x_{ij})}+(1-y_{ij}+x_{ij})e^{x_{ij}-y_{ij}}-2\right]$$

$$\frac{\partial J(X:Y)}{\partial x_{ij}}=\sum_{i=1}^{m}\sum_{j=1}^{n}\left[-(1-x_{ij}+y_{ij})e^{(y_{ij}-x_{ij})}-e^{(y_{ij}-x_{ij})}+(1-y_{ij}+x_{ij})e^{x_{ij}-y_{ij}}+e^{x_{ij}-y_{ij}}\right]$$

$$\frac{\partial J(X:Y)}{\partial x_{ij}}=\sum_{i=1}^{m}\sum_{j=1}^{n}\left[(x_{ij}-y_{ij}-2)e^{(y_{ij}-x_{ij})}+(2-y_{ij}+x_{ij})e^{x_{ij}-y_{ij}}\right]$$

$$\frac{\partial J(X:Y)}{\partial x_{ij}}=0\ when\ \ X=Y\ or\ x_{ij}=y_{ij}$$

$$\frac{\partial^2 J(X:Y)}{\partial {x_{ij}}^2}=\sum_{i=1}^{m}\sum_{j=1}^{n}\left[e^{(y_{ij}-x_{ij})}-(x_{ij}-y_{ij}-2)e^{(y_{ij}-x_{ij})}+(2-y_{ij}+x_{ij})e^{x_{ij}-y_{ij}}+e^{x_{ij}-y_{ij}}\right]$$

$$\frac{\partial^2 J(X:Y)}{\partial {x_{ij}}^2}=\sum_{i=1}^{m}\sum_{j=1}^{n}\left[(3-x_{ij}+y_{ij})e^{(y_{ij}-x_{ij})}+(3-y_{ij}+x_{ij})e^{x_{ij}-y_{ij}}\right]$$

$$\frac{\partial^2 J(X:Y)}{\partial {x_{ij}}^2}=6mn>0\ when\ X=Y\ or\ x_{ij}=y_{ij}$$

This shows that J(X:Y) is a convex function of x_{ij}.

Similarly we can show that J(X:Y) is a convex function of y_{ij}.

Since the proposed divergence measure satisfied all the four required axioms, we can say that this measure is a valid divergence measure of fuzzy matrices.

Now we prove some properties of the proposed measure in the form of theorems.

Theorem 3.5. Let $X\ and\ Y \in [F_M]_{m\times n}$ then the following properties are satisfied by J(X:Y).

a. $J(X:Y) = J(X^c:Y^c)$
b. $J[\max(X,Y):\min(X,Y)] = J(X:Y)$
c. $J(X:\max(X,Y)) = J(Y:\min(X,Y))$
d. $J(X:\min(X,Y)) = J(Y:\max(X,Y)$

Proof. For this purpose we divide the elements of each pair of fuzzy matrices having equal order into two sets as given below:

$$\left.\begin{array}{l} S_1 = \{x_{ij}\ or\ y_{ij}\ ;\ x_{ij} \in X\ or\ y_{ij} \in Y;\ x_{ij}\ \geq\ y_{ij}\} \\ S_2 = \{x_{ij}\ or\ y_{ij}\ ;\ x_{ij} \in X\ or\ y_{ij} \in Y;\ x_{ij}\ <\ y_{ij}\} \end{array}\right\}$$

(a) We have

$$J(X:Y) = \sum_{i=1}^{m}\sum_{j=1}^{n}\Big[\left(1 - x_{ij} + y_{ij}\right)e^{(y_{ij}-x_{ij})} + \left(1 - y_{ij} + x_{ij}\right)e^{x_{ij}-y_{ij}} - 2\Big]$$

$$J(X^c:Y^c) = \sum_{i=1}^{m}\sum_{j=1}^{n}\begin{bmatrix} \Big(1-(1-x_{ij}) + (1-y_{ij})\Big)e^{\left((1-y_{ij})-(1-x_{ij})\right)} \\ +\left(1-(1-y_{ij}) + (1-x_{ij})\right)e^{(1-x_{ij})-(1-y_{ij})} - 2 \end{bmatrix}$$

$$J(X^c:Y^c) \sum_{i=1}^{m}\sum_{j=1}^{n}\Big[\left(1 - x_{ij} + y_{ij}\right)e^{(y_{ij}-x_{ij})} + \left(1 - y_{ij} + x_{ij}\right)e^{x_{ij}-y_{ij}} - 2\Big]$$

$$J(X^c:Y^c) = J(X:Y)$$

Hence the claim is proven.

(b) We have

$$J[\max(X,Y):\min(X,Y)]$$

$$= \sum_{i=1}^{m}\sum_{j=1}^{n}\begin{bmatrix} \left(1-\max(x_{ij},y_{ij}) + \min(x_{ij},y_{ij})\right)e^{\left(\min(x_{ij},y_{ij})-\max(x_{ij},y_{ij})\right)} \\ +\left(1-\min(x_{ij},y_{ij}) + \max(x_{ij},y_{ij})\right)e^{\left(\max(x_{ij},y_{ij})-\min(x_{ij},y_{ij})\right)} - 2 \end{bmatrix}$$

$$= \sum_{x_{ij},y_{ij}\in S_1}\left[\left(1 - x_{ij} + y_{ij}\right)e^{(y_{ij}-x_{ij})} + \left(1 - y_{ij} + x_{ij}\right)e^{x_{ij}-y_{ij}} - 2\right]$$

$$+ \sum_{x_{ij},y_{ij}\in S_2}\Big[\left(1 - x_{ij} + y_{ij}\right)e^{(y_{ij}-x_{ij})} + \left(1 - y_{ij} + x_{ij}\right)e^{x_{ij}-y_{ij}} - 2\Big]$$

$$= \{\sum_{x_{ij},y_{ij}\in S_1}\Big[\left(1 - x_{ij} + y_{ij}\right)e^{(y_{ij}-x_{ij})}\Big] + \sum_{x_{ij},y_{ij}\in S_2}\Big[\left(1 - x_{ij} + y_{ij}\right)e^{(y_{ij}-x_{ij})}\Big]\}$$

$$+\{\sum_{x_{ij},y_{ij}\in S_1}\left[\left(1 - y_{ij} + x_{ij}\right)e^{x_{ij}-y_{ij}} - 2\right] + \sum_{x_{ij},y_{ij}\in S_2}\Big[\left(1 - y_{ij} + x_{ij}\right)e^{x_{ij}-y_{ij}} - 2\Big]\}$$

$$= \sum_{i=1}^{m}\sum_{j=1}^{n}\Big[\left(1 - x_{ij} + y_{ij}\right)e^{(y_{ij}-x_{ij})}\Big] + \sum_{i=1}^{m}\sum_{j=1}^{n}\Big[\left(1 - y_{ij} + x_{ij}\right)e^{x_{ij}-y_{ij}} - 2\Big]$$

$$= \sum_{i=1}^{m}\sum_{j=1}^{n}\Big[\left(1 - x_{ij} + y_{ij}\right)e^{(y_{ij}-x_{ij})} + \left(1 - y_{ij} + x_{ij}\right)e^{x_{ij}-y_{ij}} - 2\Big] = J(X:Y)$$

Hence the proof is complete.

(c) $J(X:\max(X,Y)) = J(Y:\min(X,Y))$

Now first taking left hand side

$$J(X:\max(X,Y)) = \sum_{i=1}^{m}\sum_{j=1}^{n}\begin{bmatrix} \left(1-x_{ij} + \max(x_{ij},y_{ij})\right)e^{\left(\max(x_{ij},y_{ij})-x_{ij}\right)} \\ +\left(1-\max(x_{ij},y_{ij}) + x_{ij}\right)e^{\left(x_{ij}-\max(x_{ij},y_{ij})\right)} - 2 \end{bmatrix}$$

$$= \sum_{x_{ij},y_{ij}\in S_1}\Big[\left(1 - x_{ij} + x_{ij}\right)e^{(x_{ij}-x_{ij})} + \left(1 - x_{ij} + x_{ij}\right)e^{x_{ij}-x_{ij}} - 2\Big]$$

$$+ \sum_{x_{ij},y_{ij}\in S_2}\Big[\left(1 - x_{ij} + y_{ij}\right)e^{(y_{ij}-x_{ij})} + \left(1 - y_{ij} + x_{ij}\right)e^{x_{ij}-y_{ij}} - 2\Big]$$

$$= \sum_{x_{ij},y_{ij}\in S_2}\Big[\left(1 - x_{ij} + y_{ij}\right)e^{(y_{ij}-x_{ij})} + \left(1 - y_{ij} + x_{ij}\right)e^{x_{ij}-y_{ij}} - 2\Big]$$

Now taking right hand side

$$J(Y:\min(X,Y)) = \sum_{i=1}^{m}\sum_{j=1}^{n}\begin{bmatrix} \left(1-y_{ij} + \min(x_{ij},y_{ij})\right)e^{\left(\min(x_{ij},y_{ij})-y_{ij}\right)} \\ +\left(1-\min(x_{ij},y_{ij}) + y_{ij}\right)e^{\left(y_{ij}-\min(x_{ij},y_{ij})\right)} - 2 \end{bmatrix}$$

$$= \sum_{x_{ij},y_{ij}\in S_1}\Big[\left(1 - y_{ij} + y_{ij}\right)e^{(y_{ij}-y_{ij})} + \left(1 - y_{ij} + y_{ij}\right)e^{y_{ij}-y_{ij}} - 2\Big]$$

$$+\sum_{x_{ij},y_{ij}\in S_2}\Big[\big(1-x_{ij}+y_{ij}\big)e^{(y_{ij}-x_{ij})} + \big(1-y_{ij}+x_{ij}\big)e^{x_{ij}-y_{ij}}-2\Big]$$

$$=\sum_{x_{ij},y_{ij}\in S_2}\Big[\big(1-x_{ij}+y_{ij}\big)e^{(y_{ij}-x_{ij})} + \big(1-y_{ij}+x_{ij}\big)e^{x_{ij}-y_{ij}}-2\Big]$$

Thus L.H.S = R.H.S.

Hence the claim is proven.

(d) $J(X{:}\min(X,Y)) = J(Y{:}\max(X,Y))$

Now first taking left hand side

$$J(X{:}\min(X,Y)) = \sum_{i=1}^{m}\sum_{j=1}^{n}\begin{bmatrix}\big(1-x_{ij}+\min(x_{ij},y_{ij})\big)e^{(\min(x_{ij},y_{ij})-x_{ij})}\\ +\big(1-\min(x_{ij},y_{ij})+x_{ij}\big)e^{(x_{ij}-\min(x_{ij},y_{ij}))}-2\end{bmatrix}$$

$$=\sum_{x_{ij},y_{ij}\in S_1}\Big[\big(1-x_{ij}+y_{ij}\big)e^{(y_{ij}-x_{ij})} + \big(1-y_{ij}+x_{ij}\big)e^{x_{ij}-y_{ij}}-2\Big]$$

$$+\sum_{x_{ij},y_{ij}\in S_2}\Big[\big(1-x_{ij}+x_{ij}\big)e^{(x_{ij}-x_{ij})} + \big(1-x_{ij}+x_{ij}\big)e^{x_{ij}-x_{ij}}-2\Big]$$

$$=\sum_{x_{ij},y_{ij}\in S_1}\Big[\big(1-x_{ij}+y_{ij}\big)e^{(y_{ij}-x_{ij})} + \big(1-y_{ij}+x_{ij}\big)e^{x_{ij}-y_{ij}}-2\Big]$$

Now taking right hand side

$$J(Y{:}\max(X,Y)) = \sum_{i=1}^{m}\sum_{j=1}^{n}\begin{bmatrix}\big(1-y_{ij}+\max(x_{ij},y_{ij})\big)e^{(\max(x_{ij},y_{ij})-y_{ij})}\\ +\big(1-\max(x_{ij},y_{ij})+y_{ij}\big)e^{(y_{ij}-\max(x_{ij},y_{ij}))}-2\end{bmatrix}$$

$$=\sum_{x_{ij},y_{ij}\in S_1}\Big[\big(1-x_{ij}+y_{ij}\big)e^{(y_{ij}-x_{ij})} + \big(1-y_{ij}+x_{ij}\big)e^{x_{ij}-y_{ij}}-2\Big]$$

$$+\sum_{x_{ij},y_{ij}\in S_2}\Big[\big(1-y_{ij}+y_{ij}\big)e^{(y_{ij}-y_{ij})} + \big(1-y_{ij}+y_{ij}\big)e^{y_{ij}-y_{ij}}-2\Big]$$

$$=\sum_{x_{ij},y_{ij}\in S_1}\Big[\big(1-x_{ij}+y_{ij}\big)e^{(y_{ij}-x_{ij})} + \big(1-y_{ij}+x_{ij}\big)e^{x_{ij}-y_{ij}}-2\Big]$$

Thus L.H.S = R.H.S.

Hence the claim is proven.

Corollary 3.1. If $X\ and\ Y \in [F_M]_{m\times n}$ then we have $J(X{:}\max(X,Y)) + J(X{:}\min(X,Y)) = J(X{:}Y)$

Proof. By using parts c and d of theorem (3.5) we can easily verify the claim.

Corollary 3.2. If $X\ and\ Y \in [F_M]_{m\times n}$ then we have $J(Y{:}\max(X,Y)) + J(Y{:}\min(X,Y)) = J(X{:}Y)$

Proof. Again by using parts c and d of theorem (3.5) the claim is verified.

Corollary 3.3. If $X\ and\ Y \in [F_M]_{m\times n}$ then we have $J(\min(X,Y):X) = J(\max(X,Y){:}Y) \le J(X{:}Y)$

Proof. By symmetry of the divergence measure and part d of theorem (3.5) we have

$$J(X{:}\min(X,Y)) = J(\min(X,Y):X) = J(\max(X,Y){:}Y) = J(Y{:}\max(X,Y))$$

Now using corollary (3.1) we can prove the result.

Corollary 3.4. If $X\ and\ Y \in [F_M]_{m\times n}$ then we have $J(\max(X,Y):X) = J(\min(X,Y){:}Y) \le J(X{:}Y)$

Proof. By symmetry of divergence measure and part c of theorem (3.5) we have

$$J(X{:}\max(X,Y)) = J(\max(X,Y):X) = J(\min(X,Y){:}Y) = J(Y{:}\min(X,Y))$$

Now using corollary (3.2) we can prove the result.

Theorem 3.6. Let $X\ and\ Y \in [F_M]_{m\times n}$ then the following properties are satisfied by $J(X{:}Y)$.

a. $J(X{:}Y^c) = J(X^c{:}Y)$
b. $J(X{:}X^c) = 2mn(e-1)$ when $x_{ij} = 0\ or\ 1$ for all $i\ and\ j$.
c. $J(X{:}X^c) = 0$ when X is a standard fuzzy matrix or $x_{ij} = 0.5$ for all $i\ and\ j$.

Proof. We have

$$J(X{:}Y) = \sum_{i=1}^{m}\sum_{j=1}^{n}\Big[\big(1-x_{ij}+y_{ij}\big)e^{(y_{ij}-x_{ij})} + \big(1-y_{ij}+x_{ij}\big)e^{x_{ij}-y_{ij}}-2\Big]$$

(a) First taking left hand side

$$J(X{:}Y^c) = \sum_{i=1}^{m}\sum_{j=1}^{n}\Big[\big(1-x_{ij}+(1-y_{ij})\big)e^{((1-y_{ij})-x_{ij})} + \big(1-(1-y_{ij})+x_{ij}\big)e^{x_{ij}-(1-y_{ij})}-2\Big]$$

$$J(X:Y^c) = \sum_{i=1}^{m}\sum_{j=1}^{n}\Big[(2 - x_{ij} - y_{ij})e^{(1-y_{ij}-x_{ij})} + (y_{ij} + x_{ij})e^{x_{ij}+y_{ij}-1} - 2\Big]$$

Now taking right hand side

$$J(X^c:Y) = \sum_{i=1}^{m}\sum_{j=1}^{n}\Big[\big(1 - (1 - x_{ij}) + y_{ij}\big)e^{(y_{ij}-(1-x_{ij}))} + \big(1 - y_{ij} + (1 - x_{ij})\big)e^{(1-x_{ij})-y_{ij})} - 2\Big]$$

$$J(X:Y^c) = \sum_{i=1}^{m}\sum_{j=1}^{n}\Big[(y_{ij} + x_{ij})e^{x_{ij}+y_{ij}-1} + (2 - x_{ij} - y_{ij})e^{(1-y_{ij}-x_{ij})} - 2\Big]$$

Thus L.H.S. = R.H.S.

Hence the claim is proven.

(b) We have

$$J(X:X^c) = \sum_{i=1}^{m}\sum_{j=1}^{n}\Big[\big(1 - x_{ij} + (1 - x_{ij})\big)e^{((1-x_{ij})-x_{ij})} + \big(1 - (1 - x_{ij}) + x_{ij}\big)e^{x_{ij}-(1-x_{ij})} - 2\Big]$$

$$J(X:X^c) = \sum_{i=1}^{m}\sum_{j=1}^{n}\Big[\big(2(1 - x_{ij})\big)e^{((1-2x_{ij})} + (2x_{ij})e^{2x_{ij}-1} - 2\Big]$$

When $x_{ij} = 0 \; or \; 1$ for all $i \; and \; j$.

$$J(X:X^c) = \sum_{i=1}^{m}\sum_{j=1}^{n}[2(e - 1)] = 2mn(e - 1)$$

Hence the claim is proven.

(c) Proceeding like (b) we have when $x_{ij} = 0.5$ for all $i \; and \; j$.

$$J(X:X^c) = \sum_{i=1}^{m}\sum_{j=1}^{n}\Big[\big(2(1 - x_{ij})\big)e^{((1-2x_{ij})} + (2x_{ij})e^{2x_{ij}-1} - 2\Big]$$

$$J(X:X^c) = \sum_{i=1}^{m}\sum_{j=1}^{n}\left[\left(2\left(\frac{1}{2}\right)\right)(1) + \left(2\left(\frac{1}{2}\right)\right)(1) - 2\right] = \sum_{i=1}^{m}\sum_{j=1}^{n}[1 + 1 - 2] = 0$$

$$J(X:X^c) = 0$$

Theorem 3.7. Let $X, \; Y \; and \; Z \in [F_M]_{m\times n}$ then

a) $J(X:Z) + J(Y:Z) - J(\max(X,Y):Z) = J(\min(X,Y):Z)$
b) $J(X:Z) + J(Y:Z) - J(\min(X,Y):Z) = J(\max(X,Y):Z)$

Proof. For this purpose we divide the elements of each pair of fuzzy matrices having equal order into two sets as given below:

$$\left.\begin{matrix} S_1 = \{x_{ij} \; or \; y_{ij} \; ; \; x_{ij} \in X \; or \; y_{ij} \in Y; \; x_{ij} \; \geq \; y_{ij}\} \\ S_2 = \{x_{ij} \; or \; y_{ij} \; ; \; x_{ij} \in X \; or \; y_{ij} \in Y; \; x_{ij} \; < \; y_{ij}\} \end{matrix}\right\}$$

(a) We have

$$J(X:Z) + J(Y:Z) - J(\max(X,Y):Z)$$
$$= \left\{\sum_{i=1}^{m}\sum_{j=1}^{n}\big[(1 - x_{ij} + z_{ij})e^{(z_{ij}-x_{ij})} + (1 - z_{ij} + x_{ij})e^{x_{ij}-z_{ij}} - 2\big] + \sum_{i=1}^{m}\sum_{j=1}^{n}\big[(1 - y_{ij} + z_{ij})e^{(z_{ij}-y_{ij})} + (1 - z_{ij} + y_{ij})e^{y_{ij}-z_{ij}} - 2\big] - \sum_{i=1}^{m}\sum_{j=1}^{n}\left[\begin{matrix}(1 - \max(x_{ij}, y_{ij}) + z_{ij})e^{(z_{ij}-\max(x_{ij},y_{ij}))} \\ +(1 - z_{ij} + \max(x_{ij}, y_{ij}))e^{(\max(x_{ij},y_{ij})-z_{ij})} - 2\end{matrix}\right]\right\}$$

$$= \sum_{x_{ij},y_{ij},z_{ij}\in S_1}\Big[(1 - x_{ij} + z_{ij})e^{(z_{ij}-x_{ij})} + (1 - z_{ij} + x_{ij})e^{x_{ij}-z_{ij}} - 2\Big] + \Big[(1 - y_{ij} + z_{ij})e^{(z_{ij}-y_{ij})} + (1 - z_{ij} + y_{ij})e^{y_{ij}-z_{ij}} - 2\Big] - \Big[(1 - x_{ij} + z_{ij})e^{(z_{ij}-x_{ij})} + (1 - z_{ij} + x_{ij})e^{x_{ij}-z_{ij}} - 2\Big]$$
$$+ \sum_{x_{ij},y_{ij},z_{ij}\in S_2}\Big[(1 - x_{ij} + z_{ij})e^{(z_{ij}-x_{ij})} + (1 - z_{ij} + x_{ij})e^{x_{ij}-z_{ij}} - 2\Big] + \Big[(1 - z_{ij} + y_{ij})e^{(y_{ij}-z_{ij})} + (1 - y_{ij} + z_{ij})e^{z_{ij}-y_{ij}} - 2\Big] - \Big[(1 - z_{ij} + y_{ij})e^{(y_{ij}-z_{ij})} + (1 - y_{ij} + z_{ij})e^{z_{ij}-y_{ij}} - 2\Big]$$

$$= \sum_{x_{ij},y_{ij},z_{ij}\in S_1}\Big[(1 - y_{ij} + z_{ij})e^{(z_{ij}-y_{ij})} + (1 - z_{ij} + y_{ij})e^{y_{ij}-z_{ij}} - 2\Big] + \sum_{x_{ij},y_{ij},z_{ij}\in S_2}\Big[(1 - x_{ij} + z_{ij})e^{(z_{ij}-x_{ij})} + (1 - z_{ij} + x_{ij})e^{x_{ij}-z_{ij}} - 2\Big]$$

$$= \sum_{i=1}^{m}\sum_{j=1}^{n}\left[\begin{matrix}(1 - \min(x_{ij}, y_{ij}) + z_{ij})e^{(z_{ij}-\min(x_{ij},y_{ij}))} \\ +(1 - z_{ij} + \min(x_{ij}, y_{ij}))e^{(\min(x_{ij},y_{ij})-z_{ij})} - 2\end{matrix}\right] = J(Min(X,Y):Z)$$

Hence the claim is proven.

(b) In part (a) it is proved that $J(X:Z) + J(Y:Z) - J(\max(X,Y):Z) = J(\min(X,Y):Z)$ thus by using this it is easy to show that $J(X{:}Z) + J(Y{:}Z) - J(\min(X,Y):Z) = J(\max(X,Y):Z)$.

Corollary 3.5. If $X,\ Y\ and\ Z \in [F_M]_{m\times n}$ then

a. $J(\max(X,Y):Z) \le J(X{:}Z) + J(Y{:}Z)$
b. $J(\min(X,Y):Z) \le J(X{:}Z) + J(Y{:}Z)$

Proof.
(a) Since $J(min(X,Y){:}Z) \ge 0$ so by using part 'a' of theorem (3.7) we prove it.
(b) Since $J(max(X,Y){:}Z) \ge 0$ so by using part 'b' of theorem (3.7) we prove it.

Theorem 3.8. Let $X,\ Y\ and\ Z \in [F_M]_{m\times n}$ then

a. $J(\max(X,Y):Z) + J(\min(X,Y):Z) = J(X{:}Z) + J(Y{:}Z)$
b. $J(X{:}\max(Y,Z)) + J(X{:}\min(Y,Z)) = J(X{:}Y) + J(X{:}Z)$

Proof. For this purpose we divide the elements of each pair of fuzzy matrices having equal order into two sets as given below:

$$\left.\begin{array}{l} S_1 = \{x_{ij}\ or\ y_{ij}\ ;\ x_{ij} \in X\ or\ y_{ij} \in Y;\ x_{ij} \ \ge\ y_{ij}\} \\ S_2 = \{x_{ij}\ or\ y_{ij}\ ;\ x_{ij} \in X\ or\ y_{ij} \in Y;\ x_{ij} \ <\ y_{ij}\} \end{array}\right\}$$

(a) By using part 'a' of theorem (3.7) we know that

$$J(X{:}Z) + J(Y{:}Z) - J(\max(X,Y):Z) = J(\min(X,Y):Z)$$

Thus

$$J(\max(X,Y):Z) + J(\min(X,Y):Z) = J(X{:}Z) + J(Y{:}Z)$$

(b) We have

$$J(X{:}Y) + J(X{:}Z) - J(X{:}\max(Y,Z)$$
$$= \left\{ \sum_{i=1}^{m}\sum_{j=1}^{n}\left[(1 - x_{ij} + y_{ij})e^{(y_{ij}-x_{ij})} + (1 - y_{ij} + x_{ij})e^{x_{ij}-y_{ij}} - 2\right] + \sum_{i=1}^{m}\sum_{j=1}^{n}\left[(1 - x_{ij} + z_{ij})e^{(z_{ij}-x_{ij})} + (1 - z_{ij} + x_{ij})e^{x_{ij}-z_{ij}} - 2\right] - \sum_{i=1}^{m}\sum_{j=1}^{n}\left[\begin{array}{c}(1 - x_{ij} + \max(z_{ij}, y_{ij}))e^{(\max(z_{ij},y_{ij})-x_{ij})} \\ +(1 - \max(z_{ij}, y_{ij}) + x_{ij})e^{(x_{ij}-\max(z_{ij},y_{ij}))} - 2\end{array}\right]\right\}$$

$$= \sum_{x_{ij},y_{ij},z_{ij}\in S_1}\Big[(1 - x_{ij} + y_{ij})e^{(y_{ij}-x_{ij})} + (1 - y_{ij} + x_{ij})e^{x_{ij}-y_{ij}} - 2\Big] + \Big[(1 - x_{ij} + z_{ij})e^{(z_{ij}-x_{ij})} + (1 - z_{ij} + x_{ij})e^{x_{ij}-z_{ij}} - 2\Big] - \Big[(1 - x_{ij} + z_{ij})e^{(z_{ij}-x_{ij})} + (1 - z_{ij} + x_{ij})e^{x_{ij}-z_{ij}} - 2\Big]$$

$$+ \sum_{x_{ij},y_{ij},z_{ij}\in S_2}\Big[(1 - x_{ij} + y_{ij})e^{(y_{ij}-x_{ij})} + (1 - y_{ij} + x_{ij})e^{x_{ij}-y_{ij}} - 2\Big] + \Big[(1 - x_{ij} + z_{ij})e^{(z_{ij}-x_{ij})} + (1 - z_{ij} + x_{ij})e^{x_{ij}-z_{ij}} - 2\Big] - \Big[(1 - x_{ij} + y_{ij})e^{(y_{ij}-x_{ij})} + (1 - y_{ij} + x_{ij})e^{z_{ij}-x_{ij}} - 2\Big]$$

$$= \sum_{x_{ij},y_{ij},z_{ij}\in S_1}\Big[(1 - x_{ij} + y_{ij})e^{(y_{ij}-x_{ij})} + (1 - y_{ij} + x_{ij})e^{x_{ij}-y_{ij}} - 2\Big] + \sum_{x_{ij},y_{ij},z_{ij}\in S_2}\Big[(1 - x_{ij} + z_{ij})e^{(z_{ij}-x_{ij})} + (1 - z_{ij} + x_{ij})e^{x_{ij}-z_{ij}} - 2\Big]$$

$$= \sum_{i=1}^{m}\sum_{j=1}^{n}\left[\begin{array}{c}(1 - x_{ij} + \min(z_{ij}, y_{ij}) +)e^{(\min(z_{ij},y_{ij})-x_{ij})} \\ +(1 - \min(z_{ij}, y_{ij}) + x_{ij})e^{(x_{ij}-\min(z_{ij},y_{ij}))} - 2\end{array}\right] = J(X{:}min(Z,Y))$$

$$J(X{:}Y) + J(X{:}Z) = J(X{:}\max(Y,Z)) + J(X{:}\min(Z,Y))$$

Thus properties **(**1**)** and **(**2**)** are satisfied.

4. Application in Decision Making and Feature Selection Problems

In this section, we introduce a method to solve decision making and feature selection problems by using the proposed non-probabilistic exponential measure of fuzzy matrices.

Now, let us consider a decision making problem involving a set of options $O = \{o_1, o_2, \ldots\ldots, o_n\}$ to be considered on the basis of certain criteria $C = \{c_1, c_2, \ldots\ldots, c_n\}$. For decision making and feature selection, the characteristic sets for each option are determined by assigning appropriate membership values, and an ideal solution O to the problem has maximum membership values in each criterion. We calculate the divergence for each case and select the option with minimum divergence.

Now to exhibit the applicability of the proposed measure we consider an example:
Suppose a person wants to buy a car and he needs to select the brand out of six brands $\{X_1, X_2, \ldots\ldots, X_6\}$ having feature conditions $\{Y_1, Y_2, \ldots\ldots, Y_7\}$. Now, for evaluating the six brands, the companies form six fuzzy sets as given below:

$$X_1 = \{(Y_1, 0.4), (Y_2, 0.7), (Y_3, 0.5), (Y_4, 0.9), (Y_5, 0.4), (Y_6, 0.6), (Y_7, 0.6)\}$$
$$X_2 = \{(Y_1, 0.7), (Y_2, 0.9), (Y_3, 0.6), (Y_4, 0.7), (Y_5, 0.6), (Y_6, 0.6), (Y_7, 0.8)\}$$
$$X_3 = \{(Y_1, 0.9), (Y_2, 0.6), (Y_3, 0.4), (Y_4, 0.5), (Y_5, 0.7), (Y_6, 0.5), (Y_7, 0.3)\}$$
$$X_4 = \{(Y_1, 0.5), (Y_2, 0.5), (Y_3, 0.6), (Y_4, 0.3), (Y_5, 0.6), (Y_6, 0.8), (Y_7, 0.7)\}$$
$$X_5 = \{(Y_1, 0.6), (Y_2, 0.5), (Y_3, 0.7), (Y_4, 0.6), (Y_5, 0.7), (Y_6, 0.5), (Y_7, 0.5)\}$$

$X_6 = \{(Y_1, 0.4), (Y_2, 0.3), (Y_3, 0.2), (Y_4, 0.5),$
$(Y_5, 0.5), (Y_6, 0.4), (Y_7, 0.3)\}$

Now these six sets are represented as a fuzzy matrix $'P'$ having order (6×7), where rows represent brands and columns represent features:

$$P = \begin{bmatrix} 0.4 & 0.7 & 0.5 & 0.9 & 0.4 & 0.6 & 0.6 \\ 0.7 & 0.9 & 0.6 & 0.7 & 0.6 & 0.6 & 0.8 \\ 0.9 & 0.6 & 0.4 & 0.5 & 0.7 & 0.5 & 0.3 \\ 0.5 & 0.5 & 0.6 & 0.3 & 0.6 & 0.8 & 0.7 \\ 0.6 & 0.5 & 0.7 & 0.6 & 0.7 & 0.5 & 0.5 \\ 0.4 & 0.3 & 0.2 & 0.5 & 0.5 & 0.4 & 0.3 \end{bmatrix}$$

Now we take the optimal solution from the above six sets which is given by:

$X = \{(Y_1, 0.9), (Y_2, 0.9), (Y_3, 0.7), (Y_4, 0.9), (Y_5, 0.7),$
$(Y_6, 0.8), (Y_7, 0.8)\}$

The set X is represented as a fuzzy row matrix B.
$B = [\,0.9 \quad 0.9 \quad 0.7 \quad 0.9 \quad 0.7 \quad 0.8 \quad 0.8]$

Now partition the above fuzzy matrix P into six row matrices $\{B_1, B_2, \dots\dots, B_6\}$. We find the divergence between these six row matrices and matrix B by using the proposed divergence measure

$$J(X{:}Y) = \sum_{i=1}^{m}\sum_{j=1}^{n}[(1 - x_{ij} + y_{ij})e^{(y_{ij}-x_{ij})} + (1 - y_{ij} + x_{ij})e^{x_{ij}-y_{ij}} - 2]$$

$$J(B_1{:}B) = \sum_{i=1}^{1}\sum_{j=1}^{7}[(1 - 0.4 + 0.9)e^{(0.9-0.4)} + (1 - 0.9 + 0.4)e^{0.4-0.9} - 2] = 1.53240$$

Similarly we find divergences of B from the other options as

$J(B_2{:}B) = 0.42209,$ $J(B_3{:}B) = 2.08724,$
$J(B_4{:}B) = 2.20652,$ $J(B_5{:}B) = 1.58428,$
$J(B_6{:}B) = 4.56607.$

We find that the optimum solution of B is 0.42209, which is defined with option B_2. Thus it can be easily said that the person prefers the car of brand X_2 with rank order of lesser preferences X_1, X_5, X_3, X_4, X_6.

As we know, in the above matrix P the columns represent the features of brands, thus for feature selection we first find the order of significance of these features.

Now we take the optimal solution according to the features from the matrix P in the form of a set:

$Y = \{(X_1, 0.9), (X_2, 0.9), (X_3, 0.7), (X_4, 0.8),$
$(X_5, 0.7), (X_6, 0.5)\}$

The set Y is represented as a fuzzy column matrix F.
$F = [\,0.9 \quad 0.9 \quad 0.9 \quad 0.8 \quad 0.7 \quad 0.5]^T$

Now partition the above fuzzy matrix P into seven column matrices $\{F_1, F_2, \dots\dots, F_7\}$. We find the divergence between these seven column matrices and matrix B by using the proposed divergence measure

$$J(X{:}Y) = \sum_{i=1}^{m}\sum_{j=1}^{n}[(1 - x_{ij} + y_{ij})e^{(y_{ij}-x_{ij})} + (1 - y_{ij} + x_{ij})e^{x_{ij}-y_{ij}} - 2]$$

$$J(F_1{:}F) = \sum_{i=1}^{6}\sum_{j=1}^{1}[(1 - 0.4 + 0.9)e^{(0.9-0.4)} + (1 - 0.9 + 0.4)e^{0.4-0.9} - 2] = 1.23048$$

Similarly we find divergences of B from the other options

$J(F_2{:}F) = 0.90878,$ $J(F_3{:}F) = 1.93453,$
$J(F_4{:}F) = 1.41779,$ $J(F_5{:}F) = 1.29107,$
$J(F_6{:}F) = 1.18823,$ $J(F_7{:}F) = 1.70972$

The optimum solution of F is 0.90878, which is achieved with option F_2. Thus it can be easily said that the feature F_2 is more significant than the other features, whose rank order is $F_6, F_1, F_5, F_4, F_7, F_3$.

Now we remove those features that do not alter the preference order of brands or maintain the optimality of brands. Since F_3 is the least significant we first remove it and check for optimality.

$J(B_1{:}B) = 1.41174,$ $J(B_2{:}B) = 0.39205,$
$J(B_3{:}B) = 1.81386,$ $J(B_4{:}B) = 2.17648,$
$J(B_5{:}B) = 1.58428,$ $J(B_6{:}B) = 3.78973.$

The optimality of brands is maintained so remove F_3.
When we remove F_5 then

$J(B_1{:}B) = 1.25901,$ $J(B_2{:}B) = 0.39205,$
$J(B_3{:}B) = 2.08724,$ $J(B_4{:}B) = 2.17648,$
$J(B_5{:}B) = 1.58428,$ $J(B_6{:}B) = 4.44541.$

We can remove F_5 because optimality is maintained after removal of it.

Similarly when we remove F_6, F_1, F_4, F_7, F_2 individually then the optimality of brands is not maintained, so we cannot remove these features.
When we remove F_3 & F_5 both then

$J(B_1{:}B) = 1.13835,$ $J(B_2{:}B) = 0.36201,$
$J(B_3{:}B) = 1.81386,$ $J(B_4{:}B) = 2.14644,$
$J(B_5{:}B) = 1.58428,$ $J(B_6{:}B) = 3.66907.$
Since optimality is maintained we can eliminate these features.
When we remove F_3 & F_6 both then

$J(B_1{:}B) = 1.29107,$ $J(B_2{:}B) = 0.27138,$
$J(B_3{:}B) = 1.54047,$ $J(B_4{:}B) = 2.17648,$
$J(B_5{:}B) = 1.31091,$ $J(B_6{:}B) = 3.29899.$

Since optimality is maintained so we can eliminate these features. When we remove F_3, F_5 & F_6 then

$J(B_1:B) = 1.01768,\quad J(B_2:B) = 0.24134,$
$J(B_3:B) = 1.54047,\quad J(B_4:B) = 2.14644,$
$J(B_5:B) = 1.31091,\quad J(B_6:B) = 3.17834.$

Since optimality is maintained we can eliminate these features. Thus we can eliminate the following features $\{\{F_3, F_5, F_6\}, \{F_3, F_5\}, \{F_3, F_6\}, \{F_3\}, \{F_5\}\}$.

4.1 Application in decision making

.Here a case study is considered to demonstrate application of proposed measure. Prof. S. C. Malik (Head of Department of Statistics, M.D.U. Rohtak, India) wants to add a software related subject in his curriculum. With suggestions from the faculty of his department, he considers different software subjects with various parameters, such as Job Efficiency, Latest software, Useful for Statisticians, Low cost to Purchase, Easy to Learn, and Curriculum Related. Fuzzy values for the defined parameters are shown in Table 1.

According to Prof. Malik the optimal fuzzy value of the standard software with respect to parameters is this:

$$S = \{0.8, 0.6, 0.9, 0.8, 0.5, 0.9\}$$

The divergences of given software and standard software are shown in Table 2, which is obtained by using the proposed measure.

Since in Table 2 the divergence between R-software and optimal value of standard software is less than for the others, it is concluded that Prof. Malik preferred R-software for the curriculum.

5. Conclusions

In this paper a non-probabilistic divergence measure of fuzzy matrix was introduced. It was also shown that the proposed measure is a valid measure. Some properties of the proposed measure were also discussed. Eventually this measure was applied to a decision making problem and to a feature selection problem. In the decision making problem, we select the best alternatives preferred over the other alternatives. In the feature selection problem, we eliminate those features that are irrelevant and after removal our original decision is maintained. Lastly a case study example was also discussed to demonstrate a real-world application of the proposed measure.

References

Bajaj, R. K., & Hooda, D. S. (2010). Generalized measures of fuzzy directed-divergence, total ambiguity and information improvement. J*ournal of Applied Mathematics, Statistics and Informatics*, *6*(2).

Basseville, M. B. (2010). *Divergence measures for statistical data processing* (Research Report PI-1961), pp.23. Retrieved from https://hal.inria.fr/inria-00542337/document

Bhandari, D., & Pal, N. R. (1993). Some new information measures for fuzzy sets*. Information Sciences*, *67*(3), 209-228.

Bhatia, P. K., & Singh, S. (2013). On some divergence measures between fuzzy sets and aggregation operations. *AMO-Advanced Modeling and Optimization*, *15*(2), 235-248.

da Costa, G. A. T. F., & Taneja, I. J. (2011). Generalized symmetric divergence measures and metric spaces. Retrieved from https://arxiv.org/abs/1105.2707

De Luca, A., & Termini, S. (1972). A definition of a nonprobabilistic entropy in the setting of fuzzy sets theory. *Information and Control*, *20*(4), 301-312.

Priti Gupta, H D Arora and Pratiksha Tiwari. Article: A Measure of Divergence between Fuzzy Sets with Advancements in Information Theory. *IJCA Proceedings on International Conference on Advances in Computer Engineering and Applications ICACEA*(3), 6-10.

Gupta P., & Kumari S. (2014). On bounds for weighted fuzzy mean difference-divergence measures. *International Journal of Scientific Research, 3*(6), 19-29.

Jain, K. C., & Chhabra, P. (2016). A new exponential directed divergence information measure. *Journal Applied Mathematics and Informatics*, *34*(3-4), 295-308.

Kapur, J. N. (1997). *Measures of fuzzy information*. New Delhi, India: Mathematical Sciences Trust Society.

Kullback, S., & Leibler, R. A. (1951). On information and sufficiency. *The Annals of Mathematical Statistics, 22*(1), 79-86.

Rényi, A. (1961). On measures of entropy and information. *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability, Volume 1: Contributions to the Theory of Statistics*. Oakland, CA: The Regents of the University of California.

Table 1. Fuzzy relation between software and criteria parameters

S/P	Job efficiency	Latest software	Useful for statisticians	Low cost to purchase	Easy to learn	Curriculum related
SPSS	0.9	0.8	0.9	0.1	0.8	1.0
C++	0.2	0.3	0.4	1.0	0.2	0.1
R	0.9	0.8	0.8	1.0	0.9	0.9
MATLAB	0.6	0.6	0.6	0.1	0.5	0.7
C	0.3	0.2	0.3	1.0	0.7	0.2
TORA	0.4	0.7	0.7	0.3	0.7	0.8

Table 2. Software divergences from standard software

Software	SPSS	C++	R	MATLAB	C	TORA
Divergence value	2.026496	4.674548	0.792166	2.08708	4.483641	1.568513

Shannon, C. E. (1948). The mathematical theory of communication. *Bell System Technical Journal, 27*, 423-467.

Sharma, B. D., & Mittal, D. P. (1975). New non-additive measures of entropy for discrete probability distributions. *Journal of Mathematical Sciences (Calcutta), 10*, 28-40.

Sharma, O., Rani, A., & Gupta, P. (2017). Some similarity and distance measures on fuzzy rough sets & it's applications. *International Journal of Engineering, Science and Mathematics, 6*(5), 85-105.

Sharma, O., Gupta, P., & Sheoran, A. (2018). Dimensionality reduction using fuzzy soft set theory. *International Journal of Statistics and Reliability Engineering*, *4*(2), 154-158.

Sharma, O., Rani, A., & Gupta, P. (2017). Some similarity and distance measures on fuzzy rough sets & it's applications. *International Journal of Engineering, Science and Mathematics, 6*(5), 85-105.

Sharma, O., Tiwari, P., & Gupta, P. (2018). Fuzzy soft matrices entropy: Application in data-reduction. *International Journal of Fuzzy System Applications (IJFSA)*, *7*(3), 56-75.

Sharma, O., & Gupta, P. (2019). Probabilistic entropy measures derived by using conic-section equation and their application in dimension reduction. *Journal of Statistics and Management Systems*, 1-19.

Taneja, I. J. (2001). Generalized information measures and their applications. Retrieved from http://www. mtmufsc.br/~taneja/book/book.html.

Zadeh, L. A. (1965). Fuzzy sets. *Information and Control, 8*(3), 338-353.