Proof the Positive Definiteness for the Jaccard Index Matrix Based on Fuzzy Sets

Linge Sun¹, Yafei Song² and Lei Lei¹

¹College of Information and Navigation, Air Force Engineering University, Xi’an 710077, China
²College of Air and Missile Defense, Air Force Engineering University, Xi’an 710051, China

afeu_sunlg@163.com, yafei_song@163.com, afeu_leilei@163.com

Abstract. This paper addresses the proof process of the positive definiteness property for the Jaccard index matrix, which is applied to define a metric distance measure between belief functions. Our proof is implemented in the framework of fuzzy sets. Based on the decomposition theorem of Jaccard index matrix proposed by M. Bouchard et al, and the transformation between crisp sets and fuzzy sets, the positive definiteness of the Jaccard index matrix can be proved in a concise way.

1. Introduction
Evidence theory was developed by Dempster and Shafer [1,2]. So it was also called as Dempster-Shafer evidence theory, shorten as D-S theory or evidence theory. Eevidence theory has inherent advantages in expressing uncertain and unknown situations. Since it can model uncertainty in a more intuitionistic way, it has been successfully applied in lots of areas such as information fusion, uncertainty reasoning and decision making [3-8]. In evidence theory, the measurement of the difference between two basic probability assignments (BPAs) is important for information processing. Many researchers has proposed different similarity or distance measures for BPAs in evidence theory.

To give a comprehensive analysis on existing distance measures between BPAs or belief functions, Jousselme and Maupin reviewed these main measures in [9],[10]. Some main properties of these distance measures were discussed. They divided these distance measures into five classes, named as, Composite, Inner product, Minkowski, Fidelity and Information-based. It was noted that the most populated kind of distance measure is the Minkowski, denoted as $p$-distance or $L_p$. In the Minkowski distance measures, Euclidean distance ($L_2$) is a kind of distance measure widely used. Based on Euclidean distance, many distance measures between BPAs have been proposed. As a widely used Euclidean distance, the Jaccard distance written as $d_{(2)}^{J}$ has attracted much attention. Jaccard distance was defined by Jousselme and Bosse in [11]. Due to its good properties, Jaccard distance has been applied in many areas.

$d_{(2)}^{J}$ is defined by introducing a weighting matrix denoted as Jac. It was pointed that the positive definiteness property of the weighting matrix can guarantee the metric properties of $d_{(2)}^{J}$ [10]. Although numerical experiments have shown that $d_{(2)}^{J}$ is metric [9],[10], there is few formal proof of this property, which is necessary to confirm the validity of $d_{(2)}^{J}$ as a full metric distance between BPAs.
Such a formal proof did not appear until the positive definiteness of $\text{Jac}$ was proved by Bouchard et al. in [12]. The proof in [12] depends on the matrix decomposition, where the matrix was decomposed into the infinite sum of several matrices with positive semidefinite property. The proposed process of proof provides reasonable results and a constructive proof structure. However, it is somewhat complicated. It is necessary to develop a simple proof process. So we propose another proof process for the positive definiteness property of the $\text{Jac}$ based on the decomposition theorem proposed in [12]. Our proof is carried out based on fuzzy sets, by regarding the subsets of a reference frame $X$ as fuzzy sets.

The remaining part of this paper will be structured as following. Section 2 proposed a brief introduction on fuzzy set theory, definitions and properties of positive definite matrices, and the decomposition theorem of Jaccard index matrix. In Section 3, the transformation between crisp sets and fuzzy sets are discussed Section 4 presents the method for proving the positive definiteness property.

2. Some Preliminaries

Since the background material about evidence theory can be found in many works, for simplicity, it will not appear in this paper. We mainly recall definitions about fuzzy set theory, definitions and properties related to positive definite matrices, and the decomposition theorem of the Jaccard index matrix.

**Definition 1.** Let $\Omega = \{x_1, x_2, \ldots, x_n\}$ be the discourse universe, then a fuzzy set $F$ in $\Omega$ is defined as [13]:

$$F = \{\langle x, \mu_x(x) \rangle | x \in \Omega \}$$  \hspace{1cm} (1)

where $\mu_x(x): \Omega \rightarrow [0,1]$ is the degree of membership degree. Conventionally, we take $\nu_x(x) = 1 - \mu_x(x)$ as the non-membership degree of $x$ to $A$. The non-membership value $\nu_x(x)$ is also a value between zero and one.

Since all the results related to positive definite matrices presented as following are well known, only definitions and theorems about positive definite matrices will be put forward, without detailed proofs.

**Definition 2.** Let $A$ be a real symmetric matrix with $n$ dimension. It is positive semidefinite if $x^T Ax \geq 0$ for every column vector $x \neq 0$ with $n$ dimension. Additionally, if $x^T Ax = 0 \Rightarrow x = 0$, $A$ is positive definite, also called as strictly positive definite.

**Theorem 1.** For a real symmetric matrix $A$, its eigenvalues are all real numbers.

**Theorem 2.** If and only if the eigenvalues are all nonnegative numbers, then $A$ is positive semidefinite. If and only if its eigenvalues are all positive numbers, then $A$ is positive definite.

**Theorem 3.** For two matrices $A$ and $B$, if they are both positive semidefinite, their sum $A + B$ is also positive semidefinite.

**Theorem 4.** Let $x$ be a column vector, $A = xx^T$ is a positive semidefinite matrix.

In [11], the weighting matrix $\text{Jac}$ was used to measure the relation between focal elements of different BPAs. The use of $\text{Jac}$ is helpful for the definition of a metric distance measure between BPAs. The elements of matrix $\text{Jac}$ are the Jaccard indices of any two subsets of $\Omega$, expressed as:

$$\text{Jac}(P, Q) = \frac{|P \cap Q|}{|P \cup Q|} \text{ for all } P, Q \in \mathcal{P}(\Omega)$$  \hspace{1cm} (2)

where $\mathcal{P}(\Omega)$ is the non-empty power set of $\Omega$ with $|\Omega| = n$. Then, based on the Jaccard matrix and Euclidean distance, the distance between two BPAs $m_1$ and $m_2$ can be defined as:

$$d_{\text{Jac}}(m_1, m_2) = \sqrt{\text{Jac}(m_1 - m_2)^T (m_1 - m_2)}$$  \hspace{1cm} (3)
Definition 3. Let $R$ and $S$ be two matrices, their Hadamard product denoted as $R \odot S$ is a matrix $T$ with elements $T(i,j) = R(i,j) \cdot S(i,j)$.

In this paper, we use the terminology $R^k$ to denote the $k$ power Hadamard product matrices $R$, i.e., $R^k(i,j) = (R(i,j))^k$. This is different to the classical definition of power production for matrix.

Theorem 5 [14]. For two positive semidefinite matrices $R$ and $S$ with identical dimension, their Hadamard product $T = R \odot S$ is also a matrix with positive semidefinite.

It has been claimed in [12] that the matrix \textbf{Jac} can be decomposed as the infinite sum of several matrices by the following theorem.

Theorem 6. Suppose that $V$ and $W$ are two matrices. Their columns and rows are indexed by the subsets of $\Omega$. We define $V(P,Q) = |P \cap Q|$, and $W(P,Q) = n - |P \cup Q|$, where $|.|$ denotes the cardinality function. Let $J_i = \frac{V}{n^k} \left( \frac{W}{n^k} \right)^i$. The sum of $r$ matrices is expressed as $\text{Jac} = \sum_{i=1}^{r} J_i$. Then the finite sum $\text{Jac}$ converges to $\text{Jac}$, i.e., $\lim_{r \to \infty} \text{Jac}_r = \text{Jac}$, or $\text{Jac} = \sum_{i=1}^{\infty} J_i$.

This theorem has been proved in [12], where the Chebyshev distance between $\text{Jac}$ and $\text{Jac}_r$ is also computed to provide an illustration.

3. Transform Crisp Sets to Fuzzy Sets

Considering the relations between fuzzy set and crisp set, we can express a crisp set in the form of fuzzy set. Therefore, any subsets of the discernment frame (or the universe of discourse) $X = \{x_1, x_2, \ldots, x_n\}$, can be regarded as fuzzy sets following Definition 4.

Definition 4. Let $X = \{x_1, x_2, \ldots, x_n\}$ be the frame of discernment. For $A \subseteq X$ and $A \neq \emptyset$, it can be expressed as a fuzzy set: $A = \{\langle x_i, \mu_A(x_i) \rangle \}$, where $\mu_A(x_i) = 1$ for $x_i \in A$ and $\mu_A(x_i) = 0$ for $x_i \not\in A$, $i = 1, 2, \ldots, n$.

For example, in the universe $X = \{x_1, x_2, x_3, x_4\}$, the subset $A = \{x_1, x_2\}$, $B = \{x_1, x_3, x_4\}$ can, respectively, be transformed to fuzzy sets as:

$\tilde{A} = \{\langle x_1, 1 \rangle, \langle x_2, 1 \rangle, \langle x_3, 0 \rangle, \langle x_4, 0 \rangle\}$, $\tilde{B} = \{\langle x_1, 1 \rangle, \langle x_2, 0 \rangle, \langle x_3, 1 \rangle, \langle x_4, 0 \rangle\}$.

Theorem 7. Let $A$ and $B$ be two non-empty subsets of $X = \{x_1, x_2, \ldots, x_n\}$. $\tilde{A}$ and $\tilde{B}$ are two IF sets generated from $A$ and $B$, respectively. Then we can get the following relations:

(R1). $|\tilde{A}| = \sum_{i=1}^{n} \mu_{\tilde{A}}(x_i)$,

(R2). $|A \cap B| = \sum_{i=1}^{n} \mu_A(x_i) \cdot \mu_B(x_i)$,

(R3). $|A \cup B| = n - \sum_{i=1}^{n} \nu_A(x_i) \cdot \nu_B(x_i)$, where $\nu_A(x_i) = 1 - \mu_A(x_i)$, $\nu_B(x_i) = 1 - \mu_B(x_i)$.

Proof (R1). $|\tilde{A}| = \sum_{i=1}^{n} \mu_{\tilde{A}}(x_i)$ is obviously indicated by the construction of $\tilde{A}$.

(R2). Setting $C = A \cap B$, we have: $\tilde{C} = \{\langle x, \mu_C(x) \rangle \}$ with $\mu_C(x) = 1$ for $x_i \in C$, $\mu_C(x) = 0$ for $x_i \not\in C$.

For $x_i \in A \cap B$, we have $x_i \in A$ and $x_i \in B$. Hence, $\mu_A(x_i) = \mu_B(x_i) = 1$. $x_i \not\in A \cap B$ indicates three cases: (1) $x_i \not\in A$, $x_i \not\in B$; (2) $x_i \in A$, $x_i \not\in B$; and (3) $x_i \not\in A$, $x_i \in B$. Therefore, $\mu_A(x_i) \cdot \mu_B(x_i) = 0$ holds for $x_i \not\in A \cap B$.

So we have $\mu_{\tilde{A}}(x_i) = \mu_A(x_i) \cdot \mu_B(x_i)$. Hence, $|A \cap B| = \sum_{i=1}^{n} \mu_A(x_i) \cdot \mu_B(x_i)$.

(R3). Considering $|A \cup B| = |A| + |B| - |A \cap B|$, we can get:
\[ |A \cup B| = \sum_{i=1}^{n} \mu_i(x) + \sum_{i=1}^{k} \mu_{ij}(x) - \sum_{i=1}^{n} \mu_i(x) \cdot \mu_{ij}(x) \]
\[ = n - \left(n - \sum_{i=1}^{n} \mu_i(x) - \sum_{i=1}^{k} \mu_{ij}(x) + \sum_{i=1}^{n} \mu_i(x) \cdot \mu_{ij}(x) \right) \]
\[ = n - \sum_{i=1}^{n} (1 - \mu_i(x) - \mu_{ij}(x) + \mu_i(x) \cdot \mu_{ij}(x)) \]
\[ = n - \sum_{i=1}^{n} (1 - \mu_i(x))(1 - \mu_{ij}(x)) \]

Given \( v_i(x) = 1 - \mu_i(x) \), \( v_{ij}(x) = 1 - \mu_{ij}(x) \), we have \( |A \cup B| = n - \sum_{i=1}^{n} v_i(x) \cdot v_{ij}(x) \).

4. Prove the Positive Definiteness Property of Jac Based on Fuzzy Sets

**Theorem 8.** The Jaccard index matrix developed by \( N \) nonempty different subsets of \( X = \{x_1, x_2, \cdots, x_n\} \) is a square, symmetric and non-singular matrix.

**Proof.** Let \( A_1, A_2, \cdots, A_N \) denote \( N \) arbitrary nonempty different subsets of \( X \). For clarity, we define:

\[ S(A_i, A_j) = \frac{|A_i \cap A_j|}{|A_i \cup A_j|}, \quad i, j = 1, 2, \cdots, N. \]

The Jaccard index matrix can be written as:

\[ \text{Jac} = \begin{bmatrix}
S(A_1, A_1) & S(A_1, A_2) & \cdots & S(A_1, A_N) \\
S(A_2, A_1) & S(A_2, A_2) & \cdots & S(A_2, A_N) \\
\vdots & \vdots & \ddots & \vdots \\
S(A_N, A_1) & S(A_N, A_2) & \cdots & S(A_N, A_N)
\end{bmatrix}.
\]

Since \( S(A_i, A_i) = 1 \) and \( S(A_i, A_j) = S(A_j, A_i) \) for \( i, j = 1, 2, \cdots, N \), \( \text{Jac} \) is a square and symmetric matrix, with 1 as its diagonal elements.

\( S(A_i, A_i) = 1 \) indicates \( |A_i \cap A_i| = |A_i \cup A_i| \), so we have \( A_i = A_i \). Thus, \( S(A_i, A_i) = 1 \Leftrightarrow A_i = A_i \).

Furthermore, \( |A_i \cap A_j| < |A_i \cup A_j| \) for \( A_i \neq A_j \) suggests that \( S(A_i, A_j) < 1 \) for \( A_i \neq A_j \).

The matrix \( \text{Jac} \) can be written in the form of column vector: \( \text{Jac} = [y_1, y_2, \cdots, y_N]^T \),

where \( y_i = [S(A_i, A_1), S(A_i, A_2), \cdots, S(A_i, A_N)]^T = [y_{i1}, y_{i2}, \cdots, y_{in}]^T \) with \( y_{ij} = 1, \ 0 \leq y_{ij} < 1 \) for \( i \neq j \).

Let us suppose that \( \text{Jac} \) is a singular matrix, which indicates that there exist linearly dependent column vectors. We can assume that \( y_i \) and \( y_k \) are linearly dependent. So we have: \( y_j = t \cdot y_k \).

Hence, \( y_p = t \cdot y_k \) for all \( p = 1, 2, \cdots, N \).

Given \( p = j \) and \( p = k \), we have:

\( y_{ij} = t \cdot y_{jk} = t \cdot y_{ij} = t \)

Finally, we can get two contradictory equations as:

\( t = 1 / y_{ij} > 1 \) and \( t = y_{jk} < 1 \).

So the assumption that the matrix \( \text{Jac} \) is singular cannot stand up. Thus \( \text{Jac} \) is non-singular.

Finally, it turns out that \( \text{Jac} \) is a square, symmetric and non-singular matrix.

**Theorem 9.** Jaccard matrix generated by \( N \) nonempty different subsets of discourse universe \( X = \{x_1, x_2, \cdots, x_n\} \), is positive definite.

**Proof.** Let \( A_1, A_2, \cdots, A_N \) denote \( N \) arbitrary subsets of \( X \). \( \tilde{A}_j = \{\{x, \mu_{ij}(x)\}\} \) \((i = 1, 2, \cdots, n, \ j = 1, 2, \cdots, N) \) is the fuzzy sets generated by \( A_j \).
Since $|A_1 \cap A_2| = \sum_{x} \mu_1(x) \cdot \mu_2(x)$ and $n - |A_1 \cup A_2| = n - \left( \sum_{x} v_1(x) \cdot v_2(x) \right) = \sum_{x} v_1(x) \cdot v_2(x)$, the matrix $V$ and $W$ can be written as:

$$V = \begin{bmatrix}
\sum_{x} \mu_2(x) \cdot \mu_2(x) & \sum_{x} \mu_1(x) \cdot \mu_2(x) & \cdots & \sum_{x} \mu_n(x) \cdot \mu_2(x) \\
\sum_{x} \mu_2(x) \cdot \mu_1(x) & \sum_{x} \mu_1(x) \cdot \mu_1(x) & \cdots & \sum_{x} \mu_n(x) \cdot \mu_1(x) \\
\vdots & \vdots & \ddots & \vdots \\
\sum_{x} \mu_2(x) \cdot \mu_n(x) & \sum_{x} \mu_1(x) \cdot \mu_n(x) & \cdots & \sum_{x} \mu_n(x) \cdot \mu_n(x)
\end{bmatrix}$$

$$W = \begin{bmatrix}
\sum_{x} v_2(x) \cdot v_2(x) & \sum_{x} v_1(x) \cdot v_2(x) & \cdots & \sum_{x} v_n(x) \cdot v_2(x) \\
\sum_{x} v_2(x) \cdot v_1(x) & \sum_{x} v_1(x) \cdot v_1(x) & \cdots & \sum_{x} v_n(x) \cdot v_1(x) \\
\vdots & \vdots & \ddots & \vdots \\
\sum_{x} v_2(x) \cdot v_n(x) & \sum_{x} v_1(x) \cdot v_n(x) & \cdots & \sum_{x} v_n(x) \cdot v_n(x)
\end{bmatrix}$$

where

$$S_2(x) = \begin{bmatrix}
\mu_2(x) \cdot \mu_2(x) & \mu_1(x) \cdot \mu_2(x) & \cdots & \mu_n(x) \cdot \mu_2(x) \\
\mu_2(x) \cdot \mu_1(x) & \mu_1(x) \cdot \mu_1(x) & \cdots & \mu_n(x) \cdot \mu_1(x) \\
\vdots & \vdots & \ddots & \vdots \\
\mu_2(x) \cdot \mu_n(x) & \mu_1(x) \cdot \mu_n(x) & \cdots & \mu_n(x) \cdot \mu_n(x)
\end{bmatrix},$$

$$S_1(x) = \begin{bmatrix}
\nu_2(x) \cdot \nu_2(x) & \nu_1(x) \cdot \nu_2(x) & \cdots & \nu_n(x) \cdot \nu_2(x) \\
\nu_2(x) \cdot \nu_1(x) & \nu_1(x) \cdot \nu_1(x) & \cdots & \nu_n(x) \cdot \nu_1(x) \\
\vdots & \vdots & \ddots & \vdots \\
\nu_2(x) \cdot \nu_n(x) & \nu_1(x) \cdot \nu_n(x) & \cdots & \nu_n(x) \cdot \nu_n(x)
\end{bmatrix}.$$ 

Let $\mu = [\mu_1(x), \mu_2(x), \ldots, \mu_n(x)]^T$, $v = [v_1(x), v_2(x), \ldots, v_n(x)]^T$, then we can get: $S_\mu(x) = \mu \mu^T$, $S_v(x) = v v^T$.

By Theorem 4, $S_\mu(x)$ and $S_v(x)$ are positive semidefinite. Taking Theorem 3 into account, we can conclude that both $V$ and $W$ are positive semidefinite matrices.

Form $J_k = V \cdot \left( \frac{W}{n} \right) ^T = V \cdot \left( \frac{W}{n} \right) \cdot \left( \frac{W}{n} \right) \cdots \left( \frac{W}{n} \right)$ and Theorem 5, we can easily obtain that $J_k$ is positive semidefinite. Hence, $\text{Jac} = \sum_{k=1}^{n} J_k$ is positive semidefinite.

Since $\text{Jac}$ is symmetric, its eigenvalues denoted by $\lambda_1, \lambda_2, \ldots, \lambda_n$ are all real numbers. Positive semidefinite $\text{Jac}$ indicates that its eigenvalues are non-negative. From Theorem 8 we get $\det \text{Jac} \neq 0$.

Then we have $\prod_{k=1}^{n} \lambda_k = \det \text{Jac} \neq 0$. Hence, $\lambda_1, \lambda_2, \ldots, \lambda_n$ are all strictly positive.

Given Theorem 2, we can finally conclude that the matrix $\text{Jac}$ is (strictly) positive definite.

5. Conclusion

In this paper, we address the proof for the positive definiteness of the Jaccard index matrix, which is applied to define a metric distance measurement between two belief functions. Our proof is implemented in the framework of fuzzy sets. Based on the decomposition theorem of Jaccard index
matrix proposed by M. Bouchard et al, and the transformation between crisp sets and fuzzy sets, the positive definiteness property of the Jaccard matrix can be proved in a concise way.

It has been pointed that Jaccard index matrix is helpful for defining effective distance or similarity measures for BPAs in evidence theory. Since no existing distance measures are suitable for all applications related to evidence theory, it is necessary to seek for reasonable distance or similarity measures. In the definition of these measures, we must consider the mathematical properties and practical demands simultaneously. The good property of Jaccard matrix may provide guidance for us in the definition of new distance/similarity measures.

References
[1] A.P. Dempster, “Upper and lower probabilities induced by a multivalued mapping”, Annals of Mathematical Statistics, vol. 38, pp. 325-339, 1967.
[2] G. Shafer, “A Mathematical Theory of Evidence”, Princeton Univ. Press, Princeton, NJ, 1976.
[3] J. Wang, J. Zhu, Y. Song, “A Self-Adaptive Combination Method in Evidence Theory Based on the Power Pignistic Probability Distance”, Symmetry, vol. 12, no. 4, pp. 1-13, 2020.
[4] K. Wen, Y. Song, C. Wu, T. Li, “A Novel Measure of Uncertainty in the Dempster-Shafer Theory”, IEEE Access, Vol. 8, no.3, pp. 51550-51559, 2020.
[5] Q. Fu, Y. Song, C. Fan, L. Lei, X. Wang, “Evidential model for intuitionistic fuzzy multi-attribute group decision making”, Soft Computing, vol. 24, no. 10, pp. 7615–7635, 2020.
[6] P. Ni, J. Gao, Y Song, W. Quan, Q. Xing, “A New Method for Dynamic Multi-Objective Optimization Based on Segment and Cloud Prediction”, Symmetry, vol. 12, no. 3, pp. 1-13, 2020.
[7] W. Wu, Y. Song, W. Zhao, “Evaluating Evidence Reliability on the Basis of Intuitionistic Fuzzy Sets”, Information, vol. 9, no. 11, pp. 1-12, 2018.
[8] X. Li, Y. Song, W. Quan, “Evaluating evidence reliability based on intuitionistic fuzzy MCDM model”, Journal of Intelligent & Fuzzy Systems, vol.31, no. 6, pp. 1167–1182, 2016.
[9] A.-L. Jousselme, P. Maupin, “On some properties of distances in evidence theory”, in: Workshop on the Theory of Belief Functions, Brest, France, 2010.
[10] A.-L. Jousselme, P. Maupin, “Distances in evidence theory: Comprehensive survey and generalizations”, International Journal of Approximate Reasoning, vol. 53, pp. 118-145, 2012.
[11] A.-L. Jousselme, D. Grenier, E. Bossé, “A new distance between two bodies of evidence”, Information Fusion, vol. 2, pp. 91-101, 2001.
[12] M. Bouchard, A.-L. Jousselme, P.-E. Doré, “A proof for the positive definiteness of the Jaccard index matrix”, International Journal of Approximate Reasoning, vol. 54, pp. 615-626, 2013.
[13] L.A. Zadeh, “Fuzzy sets”, Information and Control, vol. 8, pp. 338-353, 1965.
[14] J.C. Gower, “A general coefficient of similarity and some of its properties”, Biometrics, vol. 27, pp. 857-871, 1971.