Article

AHP-Like Matrices and Structures—Absolute and Relative Preferences

David Koloseni ¹, Tove Helldin ² and Vicenç Torra ^{2,3,*}

¹ Department of Mathematics, University of Dar es Salaam, Dar es Salaam 35065, Tanzania; dkdavykol@gmail.com

² School of Informatics, University of Skövde, 54128 Skövde, Sweden; tove.helldin@his.se

³ Hamilton Institute, Maynooth University, W23 A5Y6 Maynooth, Ireland

* Correspondence: vtorra@his.se

Received: 12 April 2020; Accepted: 10 May 2020; Published: 18 May 2020

Abstract: Aggregation functions are extensively used in decision making processes to combine available information. Arithmetic mean and weighted mean are some of the most used ones. In order to use a weighted mean, we need to define its weights. The Analytical Hierarchy Process (AHP) is a well known technique used to obtain weights based on interviews with experts. From the interviews we define a matrix of pairwise comparisons of the importance of the weights. We call these AHP-like matrices absolute preferences of weights. We propose another type of matrix that we call a relative preference matrix. We define this matrix with the same goal—to find the weights for weighted aggregators. We discuss how it can be used for eliciting the weights for the weighted mean and define a similar approach for the Choquet integral.

Keywords: aggregation functions; weight selection; fuzzy measures; AHP (Analytical Hierarchy Process)

1. Introduction

Saaty [1,2] introduced the Analytical Hierarchy Process (AHP). It is widely used to define weights for aggregation functions. The key idea is to ask an expert about the relative importance of a criteria with respect to another criteria. By means of comparing all pairs of criteria, we obtain a matrix with all pairwise comparisons. The matrix is later used to extract the weights of each criteria. This approach was defined for extracting the weights for the weighted mean, but it can be applied in the same way for other aggregation functions that require weights.

At present there exists a large number of aggregation functions [3–6]. Aggregation functions are used to combine and fuse information. Different aggregation functions exist because there are different types of data (e.g., numerical, categorical, partitions, sequences, and dendrograms) as well as different requirements on the properties of the aggregation functions (e.g., robust to outliers, conjunctive, weighted), and so forth.

For numerical data, the arithmetic mean and the weighted mean are the most well known aggregation functions. The quasi-arithmetic mean, the ordered weighted operator (OWA) [7], the weighted OWA (WOWA) [8], and the fuzzy integrals [9,10] are other examples of functions for numerical data. The Choquet [11] and the Sugeno [12] integrals are examples of fuzzy integrals. Weighted minimum, weighted maximum, and the Sugeno integral are aggregation functions that can be used not only for numerical data but also for data in ordinal scales. This is so because they are defined in terms of minimum and maximum of some terms. These functions have been used in a broad range of applications including decision making [5,13] and data integration [14].

The arithmetic mean and the Choquet integral assume that the information sources that supply the data (e.g., criteria, experts) are independent. In contrast, fuzzy integrals permit to consider interactions between the information sources. The parameters of these fuzzy integrals will be used not only to represent the importance of the sources but also to express their interactions.

In the weighted mean, the importance of each information source is represented by a weight. These weights define a weighting vector with one weight for each source. It is well known that all weights are positive and add up to one. The OWA operator also requires a weighting vector. Similarly, the weights are positive and add up to one. Nevertheless, the meaning or interpretation of the weights are different. In the OWA operator, weights are used to represent the relative importance of small and large values in the aggregation process. This permits to define functions that have a conjunctive behaviour (that give importance to smallest values) and functions that are disjunctive (that give importance to the largest values), and, thus, we can model compensation (a large value compensates a small value). The WOWA operator uses two types of weights, one to define the importance of the sources (weighted mean like) and the other to define the relative importance of small/large values (OWA-like).

Fuzzy integrals use fuzzy measures to represent the importance of the sources. Fuzzy measures are set functions and, in this case, they are defined on the set of information sources. Therefore, we define measures for any subset of information sources. In this way, the measure of a subset of sources correspond to its importance. Fuzzy measures are not forced to be additive, this means that we may have that the measure of $A \cup B$ (for disjoint A and B) is not the measure of A plus the measure of B . This is used to model positive and negative interactions of criteria.

The Analytical Hierarchical Process (AHP) deals with the problem of defining the weights. This is not an easy problem in real applications. Outcomes heavily depend on weights, and these weights are not easy to determine. In AHP, the definition of weights is based on questioning an expert about the relative importance of criteria. That is, we ask for a_{ij} , how many times is information source or criteria c_i more important than criteria c_j . This approach has been used for the weighted mean but can be applied as well to the other aggregation functions that need weights.

In a recent paper [15] we proposed to distinguish between absolute and relative preferences for AHP-like matrices. Let us use a for the absolute preferences and $\tilde{a}$ for the relative ones. Absolute preferences refer to classical AHP ones where each a_{ij} compares the importance of c_i and c_j . In the relative preference, $\tilde{a}_{ij}$ is defined assuming that we are already considering criteria c_i and then evaluate the importance of c_j knowing that c_i is taken. The importance of c_j is relative to this fact. Observe that we may have two criteria that are equally important, so, $a_{ij} = 1$ but if they are correlated we may use $\tilde{a}_{ij} = 0$. In contrast, if they are not correlated $\tilde{a}_{ij} = 1$. In this paper we discuss this new type of matrices and how to define the weights from them.

Our motivation for introducing relative preferences is on using correlation coefficients and dissimilarity matrices as a basis for defining weights and fuzzy measures. More concretely, for two highly correlated criteria, the presence of one makes the other one unnecessary in the aggregation process. The same applies when the two criteria are very similar. We will see that we can build a relative preference matrix from the correlation coefficients. The same applies for dissimilarity matrices. Correlation coefficients take values in $[-1, 1]$ and dissimilarities can be expressed in the range $[0, 1]$. These values are then processed to define the relative preference matrix and later to obtain weighting vectors or fuzzy measures.

This paper follows the following structure—Section 2 reviews some preliminaries needed in the rest of the paper. Section 3 discusses absolute and relative preferences for AHP-like matrices. Section 4 discusses the problem of matrix elicitation. We finish the paper with some conclusions and lines for future work.

2. Preliminaries

In this section we review some concepts that we need later in this work. We begin with reviewing some aggregation functions, and then give an outline of the AHP method. See for example, references [3,6] for additional details.

2.1. Some Aggregation Functions

Let us consider a set of n information sources (e.g., experts, criteria, sensors, etc.) that supply n numerical values $b = (b_1, b_2, \dots, b_n)$. Given these data, the arithmetic mean is defined by $\sum b_i / n$.

We define a weighted vector of dimension n as $w = (w_1, \dots, w_n)$ with $w_i \geq 0$ and $\sum w_i = 1$. Given data b and weights w we define the weighted mean by $\sum_i w_i b_i$. In this case we interpret w_i as the weight or importance of the i th information source. When $w_i = 1/n$ (all the sources have the same importance), the weighted mean reduces to the arithmetic mean.

The Choquet integral is a type of fuzzy measure that can be seen as a generalization of the weighted mean. It expresses the importance of the information sources by means of a fuzzy measure. Let $X = \{x_1, \dots, x_n\}$ be the set of n information sources and let f be a function that assigns to each information source the value that it supplies. Using b , as above, for the values supplied by the sources this corresponds to $f(x_i) = b_i$.

A fuzzy measure is a monotonic set function. In our context, we need a set function on X . The definition is as follows:

Definition 1. A fuzzy measure (capacity or non-additive measure) μ on a set X is a set function $\mu : \wp(X) \rightarrow [0, 1]$ satisfying the following axioms:

- (i) $\mu(\emptyset) = 0, \mu(X) = 1$ (boundary conditions)
- (ii) $A \subseteq B$ implies $\mu(A) \leq \mu(B)$ (monotonicity)

Additive measures are a particular family of fuzzy measures that satisfy $\mu(A \cup B) = \mu(A) + \mu(B)$ when $A \cap B = \emptyset$. Naturally, probabilities and the Lebesgue measure are examples of additive measures. Sugeno λ -measures [12], are another family of measures. In contrast to the ones mentioned here, they are not additive. Sugeno λ -measures belong to a larger family of measures that are the $\perp$ -decomposable fuzzy measures. Below follows a definition of Sugeno λ -measures.

Definition 2. Let μ be a fuzzy measure; then, μ is a Sugeno λ -measure if for some fixed $\lambda > -1$ it holds that

$$\mu(A \cup B) = \mu(A) + \mu(B) + \lambda \mu(A) \mu(B) \quad (1)$$

for all $A \cap B = \emptyset$.

An interesting property of Sugeno λ -measures is that we can determine the measure from non-additive weights on the sources. Details are given in Section 5.3.1 of Reference [6].

Fuzzy integrals permit us to aggregate data taking into account fuzzy measures. In this case, the measures play the role of weights. Their relevance is that fuzzy measures allow us to take into account interactions between criteria, which is not the case when we are constrained by additivity. For example, we may consider that the importance of two criteria c_1 and c_2 taken together is less than the addition of their importance (i.e., $\mu(\{c_1, c_2\}) \leq \mu(\{c_1\}) + \mu(\{c_2\})$). This means that there is a negative interaction among the criteria or, from a decision making perspective, that the information contained in one is redundant (similar or correlated) to the one contained in the other. In contrast, we may need to consider the case that the two criteria are complementary expressing $\mu(\{c_1, c_2\}) \geq \mu(\{c_1\}) + \mu(\{c_2\})$.

A scenario like this appears in a multicriteria decision making problem that include *price*, *security*, and *comfort* as criteria for buying a car. Then, we may need to consider that the importance of the

criteria $\{price, security\}$ is not necessarily the addition of the importance of the individual criteria. Such measure will imply that a car with a very good price and a very good comfort will have much larger relevance than any one only satisfying only one of the criteria.

Fuzzy integrals permit to aggregate data with respect to a given fuzzy measure. Formally, we will integrate a function f with respect to the fuzzy integral. We give below the definition of the Choquet integral, one of the fuzzy integrals defined and studied so far.

Definition 3. Let μ be a fuzzy measure on X ; then, the Choquet integral of a function $f : X \rightarrow \mathbb{R}^+$ with respect to the fuzzy measure μ is defined by

$$(C) \int f d\mu = \sum_{i=1}^N [f(x_{s(i)}) - f(x_{s(i-1)})] \mu(A_{s(i)}), \quad (2)$$

where $f(x_{s(i)})$ indicates that the indices have been permuted so that $0 \leq f(x_{s(1)}) \leq \dots \leq f(x_{s(N)}) \leq 1$, and where $f(x_{s(0)}) = 0$ and $A_{s(i)} = \{x_{s(i)}, \dots, x_{s(N)}\}$.

The Choquet integral can be seen as a generalization of the weighted mean. Formally, the Choquet integral of a function with respect to an additive measure corresponds to the Lebesgue integral. This means, that the integral of a function will be the expectation if the measure is a probability. Therefore, in discrete domains, the Choquet integral reduces to the weighted mean when the measure is additive and add up to one.

2.2. The AHP Method

The Analytical Hierarchy Process (AHP) has been extensively used to extract weights for weighted means. That is, to assign weights to each of the criteria.

Let us denote the criteria by $c_1, \dots, c_n$. Then, the AHP process starts with interviewing experts and asking them to evaluate each pair of criteria. For each pair c_i, c_j the expert is asked how many times c_i is more relevant than c_j . In this way, we define a matrix $\{a_{ij}\}$. From these values $\{a_{ij}\}$ we extract the weights for each criteria. Observe the following: if w_i and w_j are weights for criteria c_i and c_j , then the following equation holds $a_{ij} = w_i / w_j$. Then, if experts are consistent we have that $a_{ij} = 1/a_{ji}$ and also that $a_{ij} = a_{ik}a_{kj}$. In practice, these conditions do not hold, leading to an inconsistent matrix.

Alternative approaches exist in the literature to find weights w_i for each criteria c_i that approximate the matrix when it is inconsistent. Following References [16,17], we classify these methods in two groups: (i) the eigenvalue approach and (ii) the methods minimizing the distance between the user-defined matrix and the nearest consistent matrix. For example, Crawford and Williams [18] proposed an approach that minimizes the difference that leads to an expression for the weights that is a geometric mean of the values in the matrix. This way to derive the weights is known as the logarithmic least square method and also as the geometric mean.

3. On the Definition of the AHP Matrices

As we have explained in the previous section, AHP bases its definition on comparing pairs of criteria. For each pair of criteria c_i, c_j the expert is asked to what degree c_i is more important than c_j . Then, the value given by the expert, say a_{ij} , is put into the matrix.

The process implicitly assumes that there is an order of the criteria with respect to their importance. This order is used to define the weights. In short, if c_i is said to be more important than c_j , then the weight of c_i is presumed to be larger than the weight of c_j .

It is also relevant here to underline that whichever criteria we take first, the matrix we obtain will be the same, up to inconsistency errors introduced by the expert. That is, whatever order of criteria we use to define the matrix, we will get approximately the same matrices.

We call this type of matrix an *absolute preference* matrix. We think that this situation applies well when the criteria are independent.

We introduced in Reference [15] an alternative type of AHP-like matrices. We call them *relative preference* matrices. This is to model an alternative situation. We have a set of criteria to evaluate a set of alternatives, but they are not independent. Because they are not independent, our preference for a new criteria will be dependent of what we have already considered.

Let us consider the case that c_1, c_2, c_3 are three highly correlated criteria, and that c_4, c_5 are also highly correlated but that any criterion of the first set of criteria is not correlated with any of the other belonging to the second set. Then, depending on which criterion is selected first, the relevance of another criterion can be different. This implies that preference ratios will also be different.

For example, if we start with c_2 , the relevance to add c_2 will be zero and what is relevant is to take c_4 or c_5 , instead. If we consider c_1 the relevance of c_4 is also high. Similarly, if we start with c_4 then the relevance of c_1 is also high. More particularly, we may expect that these two relevances are the same. In other words, we may expect that the matrix is symmetric. In addition, we may also expect that the diagonal is zero. Note that if we have already c_i , the importance or relevance of including c_i (again) is zero.

More specifically, this process defines the relative preference matrix $A = \{a_{ij}\}$ and the elements a_{ij} of the matrix correspond to answers to the question: *If we take attribute c_i , to which degree would you also include c_j ?* The degree value is taken in the $[0, 1]$ interval.

Then, the relative preference matrix is a matrix in which the elements are between 0 and 1 (i.e., $a_{ij} \in [0, 1]$), where we expect symmetry ($a_{ij} = a_{ji}$) and in which the diagonal is zero ($a_{ii} = 0$). A matrix satisfying these properties will be called consistent. This is established in the following definition.

Definition 4. Let $c_1, \dots, c_n$ be a set of criteria. Then, a matrix $A = \{a_{ij}\}$ where $a_{ij} \in [0, 1]$ refers to the importance of criteria c_j taken criteria c_i is a consistent relative preference matrix if

- $a_{ii} = 0$ for all $i \in \{0, \dots, n\}$,
- $a_{ij} = a_{ji}$ for all $i, j \in \{0, \dots, n\}$.

In the same way that absolute preferences in AHP are usually not consistent when elicited from experts, we do not claim relative preference matrices to be consistent in general. Because of that we may consider algorithms to approximate inconsistent matrices by a consistent one. We also consider the need of defining the weights from the matrices.

We illustrate this definition with an example formalizing the case of the five criteria discussed above.

Example 1. Let us consider the criteria $C = \{c_1, c_2, c_3, c_4, c_5\}$. Then, let the importances of the criteria be defined according to the following matrix:

$$A = \begin{pmatrix} 0 & 0.1 & 0.1 & 0.9 & 0.9 \\ 0.1 & 0 & 0.1 & 0.9 & 0.9 \\ 0.1 & 0.1 & 0 & 0.9 & 0.9 \\ 0.9 & 0.9 & 0.9 & 0 & 0.1 \\ 0.9 & 0.9 & 0.9 & 0.1 & 0 \end{pmatrix}$$

This matrix is a consistent relative preference matrix.

3.1. Correlation Matrices and Dissimilarity Matrices

We have discussed above that when establishing the importance of a criteria, we may give a low importance when the new criteria is highly correlated with another criteria already considered. The following result shows that we may define consistent relative preference matrices based on correlation coefficient matrices.

Proposition 1. Let $C = \{c_1, \dots, c_n\}$ be a set of criteria. Let $X_1, \dots, X_n$ be random variables that model the values that these criteria take. Let $\text{cor}(X_i, X_j)$ be the correlation coefficients and let $P = \{\rho_{ij}\}$ be the matrix defined with these correlation coefficients.

Then, the matrix defined by

$$A = J - \text{abs}(P)$$

is a consistent relative preference matrix.

Note. Here J is the $n \times n$ matrix with ones in all positions (i.e., $J = \mathbf{1}\mathbf{1}^T$).

Proof. As $\rho_{ii} = \text{cor}(X_i, X_i) = 1$, we have that $a_{ii} = 0$. As $\text{cor}(X_i, X_j) = \text{cor}(X_j, X_i)$ then $a_{ij} = a_{ji}$. $\square$

We have a similar result if we consider dissimilarity matrices for pairs of criteria, where these dissimilarities are in the $[0,1]$ interval. The matrices need to be defined so that the dissimilarity between a criteria and itself is 0, and, naturally, the dissimilarity is symmetric. In this case, we can just define:

$$A = J - D.$$

We consider correlation and dissimilarity matrices as a natural way to construct relative preference matrices. Weights for AHP are usually extracted from (absolute) preference matrices elicited from experts. Nevertheless, this is often not the only information available on the criteria. For example, we may have data on the different ways to evaluate (i.e., criteria) the alternatives, and we can use this information to compute the correlation between criteria. This information can then be used to aggregate the information in AHP. Similarly, experts may be able to provide information about the similarity or dissimilarity of the criteria. This information can also be of relevance in the aggregation process. Our approach permits to define a relative preference matrix from both correlation matrices and dissimilarity matrices. Then, we can use our approach to extract the weights. So, in short, our approach follows the following steps: (1) obtain data and determine correlation coefficients; (2) build the relative preference matrix; and (3) determine weights from the relative preference matrix (see Section 3.3).

3.2. On Relative Preference Hierarchical Structures

In this section we consider a generalization of the relative preference matrices. When we define the relative preference matrices we consider pairs of elements c_i and c_j . The matrix is defined by pairs a_{ij} and a_{ji} where the former corresponds to the degree when already having c_i we add c_j , and the latter to the degree when already having c_j we add c_i . The generalization consists of having sets of criteria $C_i \subseteq C$ and adding a criteria c_j .

We can then consider two consistency conditions, that roughly correspond to the ones in Definition 4. The first implies that adding a criteria c_j to the set C_i has degree 0 if c_j is already in C_i . This corresponds to the first condition in the definition above (i.e., $a_{ii} = 0$). The other one is that when a set and a criteria result into a set C_t , the degree of this composition is independent of the path in which the elements in C_t are added. That is, if we consider the criteria c_1 , then c_2 , and finally c_3 (i.e., the degree $(\emptyset, c_1) + (\{c_1\}, c_2) + (\{c_1, c_2\}, c_3)$) should be equal to considering first criteria c_2 , then c_3 and finally c_1 (i.e., the degree $(\emptyset, c_2) + (\{c_2\}, c_3) + (\{c_2, c_3\}, c_1)$).

We formalize the definition below. Our definition needs the concept of a chain.

Definition 5. Let X be a set. Given $A, B \subset X$ with $A \neq B$, we say that A is covered by B and write $A \prec B$ or $B \succ A$ when for all C such that $A \subseteq C \subseteq B$ with $C \neq B$, then $C = A$. That is, $A \prec B$ if there are no elements of $\wp(X)$ between A and B .

Let $\mathcal{C} = (c_0, c_1, \dots, c_n)$, with $c_i \subseteq X$ for $i = 1, \dots, n$ we say that $\mathcal{C}$ is a chain of $\wp(X)$ for $A \subseteq X$ if it satisfies that $\emptyset = c_0 \prec \dots \prec c_{n-2} \prec c_{n-1} \prec c_n = A$.

Observe that for a given chain $\mathcal{C} = (c_0, c_1, \dots, c_n)$, the set difference between two consecutive sets c_i and c_{i+1} is just an element. I.e., $c_{i+1} \setminus c_i$ is an element of X , the one we are adding to build the chain. We use this property in the following definition.

Definition 6. Let $X = \{c_1, \dots, c_n\}$ be a set of criteria. Then, a relative preference hierarchical structure is a set of chains $\mathcal{C} = \{\mathcal{C}^i\}_i$ on $\wp(X)$ and a function $d : \wp(X) \times X \rightarrow [0, 1]$ if

- $d : (A, c) = 0$ for all $c \in A$,
- Let $A \subseteq X$, then for any pair of chains $\mathcal{C}^1 = (c_0^1, c_1^1, \dots, c_n^1 = A)$ and $\mathcal{C}^2 = (c_0^2, c_1^2, \dots, c_n^2 = A)$ with $\mathcal{C}^1, \mathcal{C}^2 \in \mathcal{C}$ we have

$$\sum_{i=0}^{n-1} d(c_i^1, c_{i+1}^1 \setminus c_i^1) = \sum_{i=0}^{n-1} d(c_i^2, c_{i+1}^2 \setminus c_i^2).$$

We will represent a relative preference hierarchical structure by $(\mathcal{C}, d)$.

Note that the second condition of this definition does not require that d is defined for all possible chains, but only for the ones in $\mathcal{C}$.

Observe that when chains are restricted to at most two elements, a relative preference hierarchical structure $(\mathcal{C}, d)$ can be seen as equivalent as a relative preference matrix. The only difference is that the matrix is required to be defined by all pairs of criteria and the hierarchical structure is not.

3.3. Weight Determination

Given a relative preference matrix, we consider how to infer a set of weights. That is, given $\{a_{ij}\}$ obtain weights $w(i)$ for all criteria $c_i \in \mathcal{C}$. To do so we propose an algorithm based on the one in Reference [15]. The algorithm follows. $0 < \alpha \leq 1$ is a parameter of the method.

- **Step 1.** $k = 1; s = \emptyset$
- **Step 2.** $c(k) = i; s = s \cup \{i\}$ // Select criteria i at random
- **Step 3.** $w(i) = \alpha$; // Assign a high weight α to the i th criterion
- **Step 4.** $k = k + 1$
- **Step 5. while** there are criteria connected to s not yet in c **loop**
 - **Step 5.1.** $c(k) =$ select a criteria connected to $c(k-1)$ not yet in s according to a probability distribution built from $a_{c(k-1),j}$.
 - **Step 5.2.** $s = s \cup \{c(k)\}$
 - **Step 5.3.** $w(c(k)) = \alpha \cdot \min_{j \in c(1), \dots, c(k-1)} a_{j,c(k)}$;
 - **Step 5.4.** $k = k + 1$
- **Step 6. end while**
- **Step 7. while** $s \neq \mathcal{C}$ **loop** // there are criteria to select
 - **Step 7.1** $c(k) = i; s = s \cup \{i\}; w(c(k)) = 0; k = k + 1$ // add missing criteria
- **Step 8. end while**

This algorithm can be observed from a graph perspective. The matrix can be seen as a graph (see Figure 1) that connects some pairs of criteria. Randomness corresponds to selecting a path that traverses nodes in the graph (at most once). All criteria not connected will have a weight of zero (they are seen as redundant to the ones existing).

We first select a criteria (i th criteria). Then, the first loop of the algorithm (Steps 5–6) deals with selecting the remaining criteria with non-zero weights. The second loop (Steps 7–8) will assign the weights of all other criteria to zero.

In the first loop we have that we select in Step 5.1 a new criteria among the ones that are not yet selected (not in $c(\cdot)$). We select the criteria at random with a probability based on its relevance

or eligibility. That is, if the last criterion considered is $c(k-1) = l$ and if $S^k = C \setminus c(\cdot)$ are the criteria pending to be selected at iteration k , we consider a probability distribution where for each criterion $c_j \in S^k$ its probability of being selected is a function of $a_{l,j}$. For example, we can just use $p(j) = a_{l,j} / \sum_{r \in S^k} a_{l,r}$.

Figure 1. Graph representing matrix A in Example 1. Each criteria is a node in the graph. Edges represent the non-zero values in the matrix. Thick edges represent a value of 0.9 in the matrix and thin edges a value 0.1. As the matrix is symmetric we represent it as an undirected graph.

Step 5.3 assigns a weight to the new selected criterion $c(k)$. The higher the relevance of the new criterion with respect to all the ones already selected ($\min_j a_{j,c(k)}$), the higher the weight.

Using this algorithm, and assuming that the iteration selects c_2 , c_4 , c_1 , c_5 , and c_3 in this order, we will get their weights equal to α , 0.9α , 0.1α , 0.1α , and 0.1α . That is, $w = (0.1\alpha, \alpha, 0.1\alpha, 0.9\alpha, 0.1\alpha)$.

When weights are extracted to be used for the weighted mean, the weights obtained from the algorithm need to be normalized. Therefore, we will have $w'_i = w_i / \sum_j w_j$.

We give another example detailing all the steps.

Example 2. We use again the matrix introduced above in Example 1. We will proceed according to the following steps. We use $\alpha = 0.8$.

- $k = 1; s = \emptyset$
- $i = 3; c(1) = 3; s = \{3\}$
- $\alpha = 0.8; w(3) = 0.8$
- $k = 2$
- $c(2) = 1$
- $s = \{2, 3\}$
- $w(c(2)) = w(1) = 0.8 * 0.1$
- $k = 3$
- $c(3) = 2$
- $s = \{1, 2, 3\}$
- $w(c(3)) = w(2) = 0.8 * \min(0.1, 0.1) = 0.8 * 0.1$
- $k = 4$
- $c(4) = 5$
- $s = \{1, 2, 3, 5\}$
- $w(c(4)) = w(5) = 0.8 * \min(0.9, 0.9, 0.9) = 0.8 * 0.9$
- $k = 5$

- $c(5) = 4$
- $s = \{1, 2, 3, 4, 5\}$
- $w(c(5)) = w(4) = 0.8 * \min(0.9, 0.9, 0.9, 0.1) = 0.8 * 0.1$

When the aggregation operator is the Choquet integral, and the measure is of the family of Sugeno λ -measures, we can use the approach proposed in Reference [19], see also References [6] (Section 5.3.1) and [20].

We can prove some properties of the algorithm above.

Proposition 2. *Let $w(i)$ be the weights obtained from the algorithm above. Then, the following holds.*

- *The weights $w(i) = \alpha$ are the largest importance for any criterion. When the values in the matrix are at most one, all the other weights will be at most α . This follows from Step 5.3 above.*
- *When all criteria are independent, the relative preference matrix is defined with $a_{ij} = 1$ for $i \neq j$. This implies that $w(i) = \alpha$ for all i .*
- *Using different criteria as starting points will result into different sets of weights. Differences can also arise because of the random selection of a new criteria in Step 5.1.*

The first two properties were given in Reference [15]. We can illustrate the last property considering three criteria c_1 , c_2 and c_3 where the first two are rather redundant (similar) ($a_{12} = a_{21} = 0.1$) and the third one is complementary ($a_{13} = a_{23} = a_{31} = a_{32} = 1$). Then, considering the order c_1, c_2, c_3 we get $(\alpha, 0.1\alpha, \alpha)$ and considering the order c_2, c_1, c_3 we get $(0.1\alpha, \alpha, \alpha)$. These results are also obtained with other orderings. An average of these weights will result into $(0.55\alpha, 0.55\alpha, \alpha)$ which will average the two similar criteria.

We formulate this idea in the following hypothesis.

Note 1. *The average of the weights obtained from different paths will result into weights that average redundant criteria.*

When the weights are used to build a Sugeno λ -measure, the following holds which makes the definition consistent with the interpretation of fuzzy measures.

Proposition 3. *(See Reference [15] for a proof) Let $a_{ij} = 1$ for all $i \neq j$ and $a_{ii} = 0$; then, the following can be proven:*

- *When $\alpha = 1/n$ with n the number of criteria, we have that $\lambda = 0$ and the measure is additive. This produces the following measure: $\mu(\{c_i\}) = 1/n$ for all i .*
- *When $\alpha > 1/n$, then $\lambda < 0$, and the measure is a plausibility.*
- *When $\alpha < 1/n$, then $\lambda > 0$, and the measure is a belief.*

The first condition shows that when all criteria are independent and $\alpha = 1/n$, we will obtain an additive fuzzy measure, which means that the fuzzy integral will reduce to the arithmetic mean. This makes the approach consistent, as equally important independent criteria are usually aggregated using the arithmetic mean.

3.4. Measure Determination

When a hierarchical structure is considered we can proceed in a way similar to what has been defined for a matrix.

- **Step 1.** $w(A) = 0$ for all $A \subseteq C$
- **Step 2.** Apply the algorithm for a matrix to obtain weights $w(i)$.
- **Step 3.** For each chain C^i in C do

- **Step 3.1** For each set but last $c_j \in C^i$ do
 - * **Step 3.1.1** $e = c_{j+1} \setminus c_j$
 - * **Step 3.1.2** if $w(c_{j+1} \cup \{e\}) = 0$ then $w(c_{j+1} \cup \{e\}) = d(c_j, e)$
- **Step 3.2** End for
- **Step 4.** End for
- **Step 5.** $\mu'(A) = \sum_{B \subseteq A} w(A)$ for all $A \subseteq C$
- **Step 6.** $\mu(A) = \mu'(A) / \mu'(C)$ for all $A \subseteq C$

This definition will construct a fuzzy measure that is normalized (i.e., $\mu(C) = 1$) and monotonic (as $w(A)$ is positive for all A). Observe that $w(A)$ can be seen as the Möbius transform of the measure. This means that (as all $w(A)$ are positive) the measure built will be a belief measure.

In Step 3.1.2 we only assign a value to $w(A)$ when this is not already assigned. Note that if the relative preference hierarchical structure is consistent, any assignment will just assign the same value. We may consider non consistent structures by means of averaging values $d(c_j, e)$ (in line with Note 1).

Given a relative preference structure it is easy to define a relative preference hierarchical structure.

Definition 7. Given a relative preference structure defined by a matrix $A = \{a_{ij}\}$ where a_{ij} corresponds to the pair of criteria c_i and c_j , we define a relative preference hierarchical structure as follows. The set of chains is defined in terms of all possible increasing sets. Then, we use the following function

$$d(A, c_j) = \min_{c_i \in A} a_{ij}$$

For this definition, Step 3.1.2 means to consider a set c_{j+1} and add an element e into the set. For illustration consider $c_{j+1} = \{c_5, c_4, c_3\}$ and $e = c_2$. We illustrate the definition of $w(c_{j+1} \cup \{e\})$ in Figure 2. In the figure we represent the edges between c_2 and the set $\{c_5, c_4, c_3\}$. According to the definition of d above, we have that $d(\{c_5, c_4, c_3\}, c_2) = \min_{c_i \in A} a_{i2} = 0.1$.

Figure 2. Representation of Step 3.1.2 in the Algorithm for Measure determination. We use $c_{j+1} = \{c_5, c_4, c_3\}$ and we add the element $e = c_2$. Using $d(A, c_j) = \min_{c_i \in A} a_{ij}$ we have that $d(\{c_5, c_4, c_3\}, c_2) = 0.1$.

4. Knowledge Elicitation from Experts

Extraction of knowledge from experts is not easy. The same applies to the case of weights and correlations. For this purpose, different approaches have been explored. Some of them are based on Bayesian theory [21], AHP [22], Conjoint Analysis [23], Discrete Choice experiments [24] and Best-Worst Scaling [25]. Some methods based on interviews to experts pose the problem that

information elicited from experts can be inconsistent. This is the case of preferences elicited for AHP. Because of that weight derivation methods for AHP need to be resistant to inconsistencies in the data (see for example, References [17,26,27]). Most methods suffer from the problem with how to best model the results from experts that vary in their subjective opinions, as well as that the context in which the questions are posed can highly influence the answers collected. In addition, in many cases, and this is the case for decision-making processes in the medical domain, the results from the knowledge elicitation process need to be presented to the experts in a coherent and transparent way. The use of model specific and complex mathematical expressions is often unsuitable for domain experts.

As stated by Pecchia et al. [22], the AHP approach towards decision-making enables a hierarchical process where a consistent framework can be built step-by-step, structuring a complex problem into smaller, less complicated ones that a decision maker can more easily solve and understand. The knowledge to be elicited is often extracted through interviews or questionnaires where the experts are asked to compare the relative importance of a set of variables. The questions posed are often phrased in the form of—“according to your experience, how important do you consider the element i compared to the element j ?”, where the participants are requested to answer using a 5-point Likert scale (see for instance Reference [22]), thus hiding the more complex model assumptions from the experts, making it easier for them to express their knowledge and review their results.

5. Conclusions and Future Work

This work studies relative preference matrices and relative preference hierarchical structures as a way to extract weights for the weighted mean and fuzzy measures for the fuzzy integrals.

We have first defined these matrices and hierarchical structures and then proposed algorithms to extract the weights and the measures from them. A few properties have been studied.

This work was motivated by the problem of building fuzzy measures from correlation matrices. We have shown how to build a relative preference matrix from a correlation matrix, which permits us later to obtain the weights with our algorithm.

As future work we consider problems similar to the ones of AHP matrices as for example, non consistent matrices, non complete matrices, matrices with other types of data other than numeric, and so forth.

Author Contributions: All authors have contributed equally. All authors have read and agreed to the published version of the manuscript.

Funding: D.K. visit to U. of Skövde was funded by the Swedish International Development Cooperation Agency. V.T. is partially funded by the Swedish Research Council.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

References

1. Saaty, T.L. *The Analytic Hierarchy Process*; McGraw-Hill: New York, NY, USA, 1980.
2. Saaty, R.W. The analytic hierarchy process—What it is and how it is used. *Math. Model.* **1987**, *9*, 161–176. [[CrossRef](#)]
3. Beliakov, G.; Pradera, A.; Calvo, T. *Aggregation Functions: A Guide for Practitioners*; Springer: Berlin/Heidelberg, Germany, 2008.
4. Beliakov, G.; Bustince, H.; Calvo, T. *A Practical Guide to Averaging Functions*; Springer: Berlin/Heidelberg, Germany, 2016.
5. Grabisch, M. *Set Functions, Games and Capacities in Decision Making*; Springer: Berlin/Heidelberg, Germany, 2016.
6. Torra, Y.V. Narukawa, *Modeling Decisions: Information Fusion and Aggregation Operators*; Springer: Berlin/Heidelberg, Germany, 2007.

7. Yager, R.R. On ordered weighted averaging aggregation operators in multi-criteria decision making. *IEEE Trans. SMC* **1988**, *18*, 183–190.
8. Torra, V. The weighted OWA operator. *Int. J. Intell. Syst.* **1997**, *12*, 153–166. [[CrossRef](#)]
9. Pap, E. (Ed.) *Handbook of Measure Theory*; North-Holland: Amsterdam, The Netherlands, 2002.
10. Torra, V.; Narukawa, Y.; Sugeno, M. (Eds.) *Non-Additive Measures: Theory and Applications*; Springer: Berlin/Heidelberg, Germany, 2013.
11. Choquet, G. Theory of capacities. *Ann. Inst. Fourier* **1953**, *5*, 131–295. [[CrossRef](#)]
12. Sugeno, M. Fuzzy measures and fuzzy integrals (in Japanese). *Trans. Soc. Instrum. Control. Eng.* **1972**, *8*, 2.
13. Beliakov, G. Learning weights in the generalized OWA operators. *Fuzzy Optim. Decis. Mak.* **2005**, *4*, 119–130. [[CrossRef](#)]
14. Abril, D.; Navarro-Arribas, G.; Torra, V. Choquet Integral for Record Linkage. *Ann. Oper. Res.* **2012**, *195*, 97–110. [[CrossRef](#)]
15. Koloseni, D.; Helldin, H.; Torra, V. Absolute and relative preferences in AHP-like matrices. *Proc. FLINS* **2018**, *2018*, 260–267.
16. Golany, B.; Kress, M. A multicriteria evaluation of the methods for obtaining weights from ratio-scale matrices. *Eur. J. Oper. Res.* **1993**, *69*, 210–220. [[CrossRef](#)]
17. Ishizaka, A.; Lusti, M. How to derive priorities in AHP: a comparative study. *CEJOR* **2006**, *14*, 387–400. [[CrossRef](#)]
18. Crawford, G.; Williams, C. A note on the analysis of subjective judgment matrices. *J. Math. Psychol.* **1985**, *29*, 387–405. [[CrossRef](#)]
19. Leszczyski, K.; Penczek, P.; Grochulski, W. Sugeno's fuzzy measure and fuzzy clustering. *Fuzzy Sets Syst.* **1985**, *15*, 147–158. [[CrossRef](#)]
20. Tahani, H.; Keller, J.M. Information fusion in computer vision using the fuzzy integral. *IEEE Trans. Syst. Man Cybern.* **1990**, *20*, 733–741. [[CrossRef](#)]
21. O'Hagan, A.; Buck, C.E.; Daneshkhah, A.; Eiser, J.R.; Garthwaite, P.H.; Jenkinson, D.J.; Oakley, J.E.; Rakow, T. *Uncertain Judgements: Eliciting Experts' Probabilities*; John Wiley & Sons: Hoboken, NJ, USA, 2006.
22. Pecchia, L.; Martin, J.L.; Ragozzino, A.; Vanzanella, C.; Scognamiglio, A.; Mirarchi, L.; Morgan, S.P. User needs elicitation via analytic hierarchy process (AHP). A case study on a Computed Tomography (CT) scanner. *BMC Med Inform. Decis. Mak.* **2013**, *13*, 2. [[CrossRef](#)] [[PubMed](#)]
23. Louviere, J.J.; Flynn, T.N.; Carson, R.T. Discrete choice experiments are not conjoint analysis. *J. Choice Model.* **2010**, *3*, 57–72. [[CrossRef](#)]
24. de Bekker-Grob, E.W.; Ryan, M.; Gerard, K. Discrete choice experiments in health economics: A review of the literature. *Health Econ.* **2012**, *21*, 145–172. [[CrossRef](#)] [[PubMed](#)]
25. Mühlbacher, A.C.; Kaczynski, A.; Zweifel, P.; Johnson, F.R. Experimental measurement of preferences in health and healthcare using best-worst scaling: An overview. *Health Econ. Rev.* **2016**, *6*, 2. [[CrossRef](#)] [[PubMed](#)]
26. Torra, V. *On the Derivation of Weights Using the Geometric Mean Approach for Set-Valued Matrices*; IOS Press: Amsterdam, The Netherlands, 2014; pp 193–201.
27. Torra, V. Derivation of priorities and weights for set-valued matrices using the geometric approach. *Appl. Artif. Intell.* **2015**, *29*, 500–513. [[CrossRef](#)]