

Chapter 2

Approaches to Criteria Combination

Abstract Two main approaches to consider the importance of the criteria in a multicriteria decision may be employed. The preferences according to each of them may be combined into a global preference by a weighted average where the importance of the criteria enters separately as weights. Otherwise the interaction between the evaluations by the multiple criteria must be taken into account when combining them and the importance of sets of criteria taken together must be computed. A simple procedure to determine the importance in the first case can be based on pairwise comparison of the criteria. A procedure to compute the importance of the sets of criteria to apply the Choquet integral in the second case may be based on pairwise comparison of preferences between distributions of probability on the space of criteria.

Keywords Weighted average • Analytic hierarchy process • Pairwise comparison • Capacity • Choquet integral

2.1 Uncertainty About Preferences

The combination of multiple criteria to decide on alternatives of action is a practical problem that every person could face at any moment. Moreover, at different times, even if the problem is formulated in the same way, the solution chosen by the same decision maker may be different. Preference depends on the criteria that are taken into account, on the importance assigned to each of them, on the relations assumed between them and, aside from all these and other features of the problem that can be objectively modeled, on subjective disturbances, which change quickly and cannot be accurately determined.

Inaccuracy affects even the simplest criteria, based on registering the presence or absence of traits considered relevant to describe each element of the set of alternatives, as in the double-entry tables of Bourdieu (1992).

Bourdieu suggests using double-entry tables to comparatively identify objects such as different types of educational institutions, martial arts or newspapers. One row in the table is assigned to each institution. In an inductive initial stage, a new column is opened whenever a property needed to characterize one of the individuals is found. This results in posing the issue of the presence or absence of such property for all the others. In a final stage, repetitions possibly introduced are eliminated.

By proceeding in this way, a functional or structural feature is associated with each column in such a way to retain all the features—and only those—that allow for discriminating more or less rigorously the different institutions. In these instances, preference will be related to the presence of desirable features or to the absence of undesirable ones.

In this framework, a numerical representation for the preference according to each criterion may be built by assigning the value 1 to the presence of a desirable attribute, -1 to the presence of an undesirable one and 0 to absence of the attribute, whatever its type.

When we use it for comparison purposes, this evaluation will be inaccurate because the presence or absence of even the simplest attributes may be subject to discussion. It will vary, for instance, the value that the presence or absence of a feature may have to make the object useful for the evaluator.

2.1.1 General Criteria Features

The decision becomes a bit more complex when the comparison, rather than being based on the ability to satisfy certain conditions, is based on the usefulness for a particular purpose to have different amounts of a certain attribute, where utility grows with the amount possessed. Thus, applying a criterion consists in evaluating such ability or utility.

Keeney (1992) suggests that for a candidate to effectively become a criterion an analysis of its properties should be conducted. In this examination, the criterion candidate must prove to be controllable, measurable, operational, isolable, concise, and understandable within the decision context. These and other properties must be judged with respect to the alternatives to be evaluated. These alternatives may have a high level of complexity, formed, for instance, by considering distinct results from the same experiment as more or less satisfactory.

2.1.2 Criteria Combination

Different relationships between the criteria determine different algorithms for the combination of the assessments according to them. The final results depend on the combination algorithm chosen as much as on the evaluations according to the multiple criteria.

The set of criteria must be both exhaustive, in the sense of enabling the decision maker to take into account all relevant aspects of the alternatives, and non-redundant, with each one adding some relevant aspect to discriminate between alternatives.

Roy (1996, 2005) adds to these two properties a third, called cohesiveness, as a necessary condition to have a coherent family of criteria. A family of criteria would present cohesion if a move that is not for the better in the evaluation according to any particular criteria would never lead to a move for a better general evaluation.

These and other properties are required of the set of criteria, and all have to do with the general fact that the model must represent as approximately as possible the reality, no model being able to ever completely cover all features of the decision problem. A gamma of different approaches to reduce the reality to a multi-criteria model is developed, for instance, in Greco and Ehrgott (2005) and in Ehrgott et al. (2010).

However, besides that, the form of combining the criteria depends on the goal that the decision maker has in mind. This adds to the difficulty of adequately making clear the composition rules. Modeling should not only enhance the ability to conduct the evaluation to produce an outcome that achieves the requirements of the evaluator but also the ability to explain how the values declared by the evaluator lead to the final outcome. The composition rules should allow for relating as clearly as possible the final preferences to previously exposed motives.

2.2 Weighted Averages

The classic criteria composition form, developed precisely in Keeney and Raiffa (1976), is conducted by assigning weights to the criteria and obtaining final scores as weighted averages of the measurements of preferences according to each of them. These final scores are sometimes called expected utilities.

From the point of view of making the rules clear, this form of composition has the advantage of simplicity. The model is built by defining the criteria to be taken into account and, through the weights, the importance assigned to each of them.

The idea behind this form of composition is that the decision maker starts by choosing one objective from among multiple options. Here, choosing an objective means preferring one among the multiple criteria. This choice may be not univocal but randomized, i.e., it can be given by a probability distribution of preference among the criteria.

The concept of a probability distribution will be formulated more clearly in the Appendix. This term is used here in the context of a lottery where each possible prize has a different chance of being won. In the present case, the criteria will be thought of as the prizes and the weights as their chances. This corresponds to the decision maker running a possibly biased roulette game to pick the preferred criterion.

Another way to look at this form of composition is by associating each criterion with a different evaluator and considering the averaging as a rule to satisfy the

group of evaluators, with the weights corresponding to the importance assigned to each different evaluator.

A strategy to reach a distribution of weights for the criteria is to ask the decision maker to compare the criteria pairwise and, afterwards, extract from the results of these comparisons a probability for the choice of each isolated criteria.

Even establishing preferences between the elements of a pair of criteria is not free of error. However, limiting the object of each evaluation to a pair and limiting the set of outcomes of the comparison to a small set of possible results (indifference of preference for one or another, or a little more in cases where a few different degrees of preference are employed) presents considerably less difficulty than evaluating each criterion directly against some fixed pattern.

2.2.1 The Analytic Hierarchy Process

Saaty (1980) developed an elegant, though laborious, method to find the weights for the desired criteria as part of a methodology named Analytic Hierarchy Process (AHP). It involves the pairwise comparison of the criteria using a scale of values for this comparison, with a criterion being at most 9 times more important than any other.

When performing this pairwise comparison, one must keep in mind that the effect of the weights depends on the different scales on which the evaluations according to the two criteria will be measured. Thus, the comparison between the weights implicitly involves a comparison between these scales. This inner scale adjustment may be avoided only if the application of all the criteria is conceived in such a way as to involve the same scale.

To tackle this problem, Saaty proposes to start the modeling by prioritizing criteria conceived in an abstract form instead of derived from the analysis of the observed attributes of the alternatives. He first defines the criteria and compares their importance. Arranging the goals, attributes, issues, and stakeholders in a hierarchy, AHP provides an overall view of the complex relationships inherent to the situation and helps the decision maker assess whether the issues in each level are of the same order of magnitude so that he can compare homogeneous elements accurately (Saaty 1990). By this approach, only when the criteria are applied to compare the alternatives by the values of its attributes do the scales on which these attributes are effectively measured appear.

A different way to address this problem of scaling is to replace each vector of attribute measurements that appear naturally with the probabilities of the different alternatives presenting the best value for such measurements. When establishing, in the next step, the priorities for the use of each of these vectors of probabilities, we have, at the same time, values measured on the same scale and conceptual criteria built on a concrete basis to compare the alternatives. The prioritization of the criteria thus defined can be made on a sounder basis than if we start from abstract concepts.

2.2.2 AHP Tools

The most noticeable feature of AHP is the form employed to address the inconsistencies arising from the pairwise comparisons. The relative preferences are registered in a square matrix M , where the ij -th entry, m_{ij} , measures the ratio between the preference for the j -th and the i -th criterion. Thus, the m_{ij} are positive numbers with

$$m_{ij} = 1/m_{ji}.$$

A square matrix with these properties is called a positive reciprocal matrix. This matrix of preference ratios is consistent if and only if, not only

$$m_{ij} * m_{ji} = m_{ii} = 1$$

for every pair (i, j) but also, for every triple (j_1, j_2, j_3) ,

$$m_{j_1 j_2} * m_{j_2 j_3} = m_{j_1 j_3}.$$

Obviously, given a row or a column of a reciprocal matrix, consistency determines the rest of the matrix. However, when informing the preference ratios for each pair of criteria on the scale from 1 to 9, the difficulty in evaluating abstract criteria leads to inaccuracies in such a way that it is not expected that these reciprocal matrices will be consistent in practice.

The Analytic Hierarchy Process (AHP) is designed to allow for inconsistencies due to the fact that, in making judgments, people are more likely to be cardinally inconsistent than cardinally consistent (Saaty 2003). The decision makers are not able to estimate precisely measurement values even with a tangible scale, and the difficulty is worse when they address abstract concepts.

If a reciprocal matrix is consistent, all its rows and columns are proportional to each other. This means that they span a linear space of dimension 1. In other words, the rank of consistent reciprocal matrices is equal to 1.

By the rank of a matrix we mean the dimension of the space generated by their columns (or their rows), i.e., the maximal number of linear independent columns (or rows).

In addition to the concept of rank, the concepts of the trace of a square matrix and their eigenvalues and eigenvectors play an important role in the weighting of the criteria in AHP.

The eigenvalues of a matrix M are the real or complex numbers λ such that, for some vector v ,

$$Mv = \lambda v.$$

The eigenvectors are those vectors v such that, for some eigenvalue λ ,

$$Mv = \lambda v.$$

The trace of a matrix is the sum of its eigenvalues. It is also equal to the sum of the diagonal elements. For reciprocal matrices, since all diagonal elements are equal to 1, the trace is equal to the number of rows or columns. Because the rank of any consistent reciprocal matrix is 1, its non-null eigenvectors are all in the same direction and, consequently, the non-null eigenvalue is that number of rows or columns. In particular, it is a real number.

In contrast, if the matrix is inconsistent, it has negative eigenvalues; thus, its highest eigenvalue is larger than the number of rows and columns. A detailed proof of this result may be found in Saaty (1990).

If the matrix of pairwise preferences between criteria is consistent, the vector of weights is the normalized eigenvector of the matrix. Saaty proposes then, to deal with inconsistency, to take as the vector of weights the unitary eigenvector associated with the highest eigenvalue, employing the value of this highest eigenvalue to decide if the level of inconsistency in the matrix is sufficiently small.

For high levels of inconsistency, the pairwise comparison of the criteria must be revised.

Saaty employs a measure of consistency called the Consistency Index, which is based on the deviation of the highest eigenvalue to m , the number of criteria:

$$CI = (\lambda_{\max} - m)/(m - 1).$$

This index may also be seen as the negative average of the other eigenvalues of the inconsistent matrix.

After knowing the Consistency Index, the next question is how to use it. Saaty proposed to use this index by comparing it with an appropriate threshold. To determine such an appropriate threshold, one can employ the Random Consistency Index, RI, an index obtained by examining reciprocal matrices randomly generated by Vargas (1982) using the scale $1/9, 1/8 \dots 1, \dots 8, 9$. The random consistency index, computed as the average of a sample of 500 matrices, for a number of criteria varying from 3 to 10, is shown in Table 2.1.

Then, Saaty employs what is called the Consistency Ratio, which is a comparison between the observed Consistency Index and the Random Consistency Index, or, formally,

$$CR = CI/RI.$$

Table 2.1 Random consistency index

n	3	4	5	6	7	8	9	10
RI	0.58	0.90	1.12	1.24	1.32	1.41	1.45	1.49

If the value of the Consistency Ratio is smaller or equal to 10 %, the inconsistency is acceptable. If the Consistency Ratio is greater than 10 %, the subjective judgment must be revised.

2.2.3 Example of AHP Application

This subsection deals with the problem of choice of a car model. Suppose the space of alternatives formed of 20 models and six criteria based on a satisfactory answer for the presence or absence of seven attributes: beauty, comfort, gas consumption, power, acquisition price, reliability and safety.

Table 2.2 presents the ratios a given decision maker consider to more adequately reflect the preference between the seven criteria corresponding to the presence or absence of each of the seven attributes.

The highest eigenvalue for this positive reciprocal matrix is 7.335. So, its consistency index is $(7.335 - 7)/(7 - 1) = 0.056$. The consistency ratio is $0.056/1.32 = 0.042 < 0.1$. So, the inconsistency is acceptable and a unitary eigenvector corresponding to this eigenvalue will be used as the vector of weights. Given by this unitary eigenvector, the weights are those in Table 2.3.

Table 2.4 describes the evaluation of 20 models according to the seven criteria directly based on these attributes.

Table 2.2 Criteria pairwise evaluation

	Beauty	Comfort	Consumption	Power	Price	Reliability	Safety
Beauty	1	1/3	1/5	1	1/3	1/9	1/9
Comfort	3	1	1/3	3	3	1/5	1/5
Consumption	5	3	1	3	5	1/3	1/3
Power	1	1/3	1/3	1	1/3	1/9	1/9
Price	3	1/3	1/5	3	1	1/7	1/7
Reliability	9	5	3	9	7	1	1
Safety	9	5	3	9	7	1	1

Table 2.3 Criteria weights

Beauty	Comfort	Consumption	Power	Price	Reliability	Safety
0.03	0.08	0.15	0.03	0.05	0.33	0.33

Table 2.4 Cars evaluations

	Beauty	Comfort	Consumption	Power	Price	Reliability	Safety
Car1	1	1	1	0	1	1	0
Car2	1	1	1	0	1	0	1
Car3	1	1	1	0	0	1	1
Car4	1	1	0	1	0	1	1
Car5	1	1	0	1	0	1	0
Car6	1	1	0	1	0	0	1
Car7	1	1	0	0	1	1	1
Car8	1	1	0	0	1	1	0
Car9	1	1	0	0	0	1	1
Car10	1	0	1	1	0	1	1
Car11	1	0	1	0	1	0	1
Car12	1	0	0	1	1	1	1
Car13	1	0	0	1	0	1	1
Car14	1	0	0	0	1	1	1
Car15	0	1	1	1	1	1	0
Car16	0	1	1	1	0	1	1
Car17	0	1	1	0	1	1	1
Car18	0	1	0	1	0	1	1
Car19	0	0	1	1	0	1	1
Car20	0	0	0	1	1	1	1

Comparing the weighted averages generated multiplying the value 0 or 1 assigned to the model according to each criterion by the weight of the criterion and adding the products, Car 17, with a score of 0.94, would be chosen.

2.3 Capacities

The classical approach to the composition of multiple criteria, described in the preceding section, employs weighted averages of the evaluations according to the multiple criteria. This form of composition is justified if the decision can be thought of as a two-stage structure: first, one of the criteria is chosen, with the chance of being chosen depending on the importance that the decision maker wants to give it; then, the chosen criterion is applied alone. In that case, the probability of an alternative being chosen is determined by the Total Probability Theorem, provided in the Appendix, as the weighted average of its probabilities of being chosen

according to each criterion with weights given by the probabilities of the criteria being chosen in the first stage.

In this case, to build the composition algorithm, possible correlations between the events corresponding to being preferred by the different criteria need only be taken into account in determining the distribution of weights among the criteria. That is, this correlation must be taken into account in the initial stage of weighting the criteria: the probability of each one being chosen must reduce the likelihood of choosing all others positively correlated with it.

However, the problem cannot always be formulated in this manner and in general the determination of weights for the weighted average is inefficient by not taking into account these correlations. A more general form of composition that draws attention to the need to consider the possible presence of correlations is to replace the weighted average of a probability distribution by the Choquet integral with respect to a capacity (Choquet 1953).

To use this new form of composition of preferences the criteria must be comparable, i.e., the preference measurement according to the various criteria must employ the same scale, or scales between which a precise relationship is known. This problem of comparability is eliminated if the preferences are given as probabilities of being the best alternative, the scale, then, being always that of the probability of being the best.

2.3.1 Choquet Integral

To make expected utility models more flexible, additive subjective probabilities are replaced by non-additive probabilities, or capacities.

Capacities may be used to model different types of behavior. Most decision makers, for example, overestimate small and underestimate large probabilities. Furthermore, most decision makers prefer decisions where more criteria are combined rather than decisions based on less available information. These behaviors cannot be expressed through an additive model.

A (normalized) capacity on the finite set of criteria S is a set function $\mu: 2^S \rightarrow [0, 1]$ satisfying the three properties:

- (1) $\mu(\emptyset) = 0$ (a set function satisfying this property is also called a cooperative game),
- (2) $\mu(S) = 1$ (normality),
- (3) $\forall A, B \in 2^N, [A \subseteq B \Rightarrow \mu(A) \leq \mu(B)]$ (monotonicity).

Thus, a capacity is a monotonic (normalized) cooperative game.

The capacity μ on S is said to be additive if

$$\mu(A \cup B) = \mu(A) + \mu(B)$$

for all disjoint subsets A and B of S .

Capacities generalize probabilities in the sense that an additive capacity is a probability.

The Choquet integral of $x = (x_1, \dots, x_m)$, an \mathbb{R}^{m+} valued function, with respect to the capacity μ on $S = \{1, \dots, m\}$ is defined as:

$$C_\mu(x) = \sum_{j=1}^m (x_{\tau(j)} - x_{\tau(j-1)}) \mu(\{\tau(j), \dots, \tau(m)\}),$$

for τ , a permutation on S such that

$$x_{\tau(1)} \leq x_{\tau(2)} \leq \dots \leq x_{\tau(m-1)} \leq x_{\tau(m)} \text{ and } x_{\tau(0)} = 0.$$

Let $x: S \rightarrow \mathbb{R}^+$ and μ a capacity. The Choquet integral of x with respect to μ satisfies

$$C_\mu(x) = \sum_{i=1}^m x(\tau(i)) [\mu(A\tau(i)) - \mu(A\tau(i+1))]$$

for $A\tau(i) = \{\tau(i), \dots, \tau(m)\}$ for every i from 1 to m , and $A\tau(m+1) = \emptyset$.

A fundamental property of the Choquet integral is that

$$C_\mu(1_A) = \mu(A), \forall A \subseteq S,$$

for 1_A , the indicator of A , the function x defined by

$$x(i) = 1 \text{ if } i \in A \text{ and } x(i) = 0 \text{ otherwise.}$$

The expected value of a function x with domain S with respect to a probability P in the finite space S is the weighted average

$$\sum_{i \in S} x(i) P(i).$$

Thus, the Choquet integral with respect to a capacity extends the expected value with respect to a probability.

But this definition makes sense only if $x_{\tau(i)}$ and $x_{\tau(j)}$, for the different possible values of i and j are commensurable. Commensurability of the measures of preference according to different criteria means that they make us able to compare the results of the evaluations according to the different criteria. This property holds for the case of evaluations according to the criteria in S given in terms of probabilities of being the best.

To compute a capacity μ , the modeler needs to define the 2^n coefficients corresponding to the capacities of the 2^n subsets of S . Modeling the capacity by means of its Moebius transform may simplify this task.

For μ a capacity on S , the Moebius transform of μ is the function $v: P(S) \rightarrow R$ defined by

$$v(A) = \sum_{B \subseteq A} (-1)^{|A-B|} \mu(B), \forall A \in 2^S$$

The Moebius transform determines the capacity by:

$$\mu(A) = \sum_{B \subseteq A} v(B).$$

The determination of the capacity may employ the Penrose-Banzhaf or Shapley interaction indices (Grabisch and Roubens 1999) for limited levels of iteration.

Given a capacity μ on S , the Penrose-Banzhaf joint index for any subset $A \subseteq S$ is given by (Penrose 1946; Banzhaf 1965)

$$\text{Banzhaf}(A) = 2^{-\#(S \setminus A)} \sum_{K \subseteq S \setminus A} \sum_{L \subseteq A} (-1)^{\#(A-L)} \mu(K \cup L),$$

for $\#$ the cardinality function, i.e., the function that associates to each set the number of elements in it.

Analogously the Shapley joint index is defined by

$$S\mu(A) = \sum_{(K \subseteq S \setminus A)} [(\#(S \setminus A \setminus K))! (\#(K))! / (\#(S \setminus A) + 1)!] \sum_{L \subseteq A} (-1)^{\#(A \setminus L)} \mu(K \cup L)$$

For an isolated criterion i , $S_\mu(\{i\})$ is called the Shapley value (Shapley 1953).

The capacity μ is said to be k -additive, for a positive integer k , if its Moebius transform v satisfies (Grabisch 1997):

- (1) $\forall T \in 2^S, v(T) = 0$ if $\#(T) > k$,
- (2) $\exists B \in 2^S$ such that $\#(B) = k$ and $v(B) \neq 0$.

By assuming 2-additivity, the complexity of the problem of determining the capacity is reduced. The capacity can then be determined employing only the coefficients $\mu(\{i\})$ and $\mu(\{i, j\})$ for i and $j \in S$.

Necessary and sufficient conditions for 2-additivity are:

- (1) $\sum_{\{i, j\} \subseteq S} \mu(\{i, j\}) - (m - 2) \sum_{i \in S} \mu(\{i\}) = 1$ (normality),
- (2) $\mu(\{i\}) \geq 0, \forall i \in S$ (nonnegativity) and
- (3) $\forall A \subseteq S$ with $\#(A) \geq 2, \forall k \in A, \sum_{i \in A \setminus \{k\}} (\mu(\{i, k\}) - \mu(\{i\})) \geq (\#(A) - 2) \mu(\{k\})$ (monotonicity).

For a 2-additive capacity μ , the Shapley value of an isolated criterion i is given by

$$S_\mu(\{i\}) = \sum_{K \subseteq S \setminus \{i\}} [(\#(S \setminus K) - 1)! (\#(K))! / \#(S)!] \mu(K \cup \{i\}) - \mu(K)$$

or

$$S_\mu(\{i\}) = \mu(\{i\}) + 1/2 \sum_{j \in S \setminus \{i\}} I_{\mu_{ij}}$$

for

$$I_{\mu_{ij}} = \mu(\{i, j\}) - \mu(\{i\}) - \mu(\{j\}).$$

$I_{\mu_{ij}}$ represents an interaction between i and j , in the sense that

$I_{\mu_{ij}} = 0$ corresponds to independence between i and j ;

$I_{\mu_{ij}} > 0$ means some complementarity between i and j , i.e., for the decision maker, both criteria have to be satisfactory in order to get a satisfactory alternative; and

$I_{\mu_{ij}} < 0$ means some substitutability or redundancy between i and j , i.e., for the decision maker, the satisfaction of one of the two criteria is sufficient to have a satisfactory alternative.

With this notation, for any $x = (x_1, \dots, x_m)$, the Choquet integral of x with respect to the 2-additive capacity μ is given by:

$$C\mu(x_1, \dots, x_m) = \sum_{i=1}^m S_\mu(i) x_i - 1/2 \sum_{i,j \in S} I_{\mu_{ij}} |x_i - x_j|.$$

2.3.2 Example of Application of the Choquet Integral

An example of application of the concept of capacity and the Choquet integral may be constructed by revising the car models choice problem presented to show how to use the AHP approach.

Essentially, a car has no value if it does not move. Thus a model evaluated as unsatisfactory from the point of view of power suffers from a basic limitation. In that sense, the value of the presence of any of the other attributes depends on the presence of power. To take this into account, a capacity might be employed to improve that study. It would derive from the weights there employed, which assumed additivity, the capacity of any set of criteria that includes power, but would assign a null capacity to any set that does not include power. This would result in a final score of zero for those cars which were assigned a value of zero with respect to power.

Table 2.5 Combined Car Scores

	Capacity score	Rank	Additive score	Rank
Car1	0	16	0.64	15
Car2	0	16	0.64	15
Car3	0	16	0.92	2.5
Car4	0.8	4	0.8	7
Car5	0.47	10.5	0.47	19.5
Car6	0.47	10.5	0.47	19.5
Car7	0	16	0.82	6
Car8	0	16	0.49	18
Car9	0	16	0.77	9
Car10	0.87	2	0.87	4
Car11	0	16	0.56	17
Car12	0.77	5.5	0.77	9
Car13	0.72	8	0.72	13
Car14	0	16	0.74	11.5
Car15	0.64	9	0.64	15
Car16	0.92	1	0.92	2.5
Car17	0	16	0.94	1
Car18	0.77	5.5	0.77	9
Car19	0.84	3	0.84	5
Car20	0.74	7	0.74	11.5

After this change, Car17 would no longer be chosen on the basis of the evaluations in Table 2.4. It would be replaced by Car16. The score of Car16 would be $0 * (0.08) + 1 * (1 - 0.08) = 0.92$.

Table 2.5 presents the final scores for the composition employing this capacity.

2.3.3 Pairwise Comparisons in the Space of Sets of Criteria

The strategy of pairwise comparisons of sets of criteria may be employed to determine the capacity. This strategy becomes more feasible if it is not needed to quantify the preference, but only to tell whether there is indifference between the elements of the pair, and, otherwise, which one is preferred.

This simplification in the comparison can be applied at the cost of complicating the objects of comparison, the decision maker being called to compare not pairs of criteria, but pairs of distributions of weights on the set of criteria. Under these

conditions and certain conditions of rationality that constitute the basis of the Expected Utility Theory of von Neumann and Morgenstern (1944), the relation of preference between pairs uniquely determines the distribution of weights desired.

The work of von Neumann and Morgenstern extends that of Daniel Bernoulli, in the 18th century for the utility of money, and has further extensions designed to deal with much complex sets than the finite set of criteria that is the object of analysis here.

To clarify the hypotheses of von Neumann and Morgenstern is first necessary to formulate precisely the concepts. The necessary concepts are made clear in the next section. Immediately after, is presented the simple version of the result of von Neumann and Morgenstern (1944) here employed.

2.3.4 Binary Relations

A binary relation \succsim on a set C is any subset of the Cartesian product $C \times C$, i.e., any set of ordered pairs of elements of C . To denote that an ordered pair (c_1, c_2) of elements of C belongs to the binary relation \succsim , we write $c_1 \succsim c_2$ and say that c_1 precedes c_2 .

A binary relation \succsim is complete on C if and only if, for all c_1 and c_2 of C , at least one of the ordered pairs (c_1, c_2) and (c_2, c_1) belongs to \succsim , that means, $c_1 \succsim c_2$ or $c_2 \succsim c_1$.

A binary relation \succsim is anti-symmetric if and only if

$c_1 \succsim c_2$ and $c_2 \succsim c_1$ only if $c_1 = c_2$.

A binary relation \succsim on C is transitive if and only if,

for all c_1, c_2 and c_3 of C , if $c_1 \succsim c_2$ and $c_2 \succsim c_3$, then $c_1 \succsim c_3$.

A binary relation \succsim is an order relation on C if it is anti-symmetric, transitive and complete on C .

A binary relation \succsim is a preference relation on C if it is transitive and complete on C . So, order relations are preference relations, but these need not be anti-symmetric. For those criteria for which $c_1 \succsim c_2$ and $c_2 \succsim c_1$, it will be said that according to \succsim the decision maker is indifferent between c_1 and c_2 and it will be used the notation $c_1 \sim c_2$.

A distribution of weights on a set C is any positive function u with domain C , i.e., any subset of the Cartesian product $C \times \mathbb{R}^+$ such that for every element x of C there is a unique positive number y for which $(x, y) \in u$. Instead of $(x, y) \in u$ is usually employed the notation $y = u(x)$ or is said that y is the weight of x , the preference value of x or the utility of x . To simplify the arguments, is usually assumed that u is a probability.

The von Neumann and Morgenstern theory involves extending the preference relations on C to preference relations on the set $D(C)$ of distributions of weights on C . Let us denote by \succsim^* the extension of \succsim to $D(C)$. For the theorem to hold, these

relations must present, besides transitivity and completeness, other properties of continuity, monotonicity, substitutability and decomposability.

Continuity means that for any weight distributions \mathbf{p} , \mathbf{q} and \mathbf{r} on C , with $\mathbf{p} \succ \mathbf{q}$ and $\mathbf{q} \succ \mathbf{r}$, there is a real $\alpha \in [0, 1]$ such that $\mathbf{q}(c) = \alpha \mathbf{p}(c) + (1 - \alpha)\mathbf{r}(c)$ for any $c \in C$.

Monotonicity means that if \mathbf{p} and \mathbf{q} are weight distributions concentrated on $\{c_1, c_2\}$ (that means, such that $\mathbf{p}(c_1) + \mathbf{p}(c_2) = \mathbf{q}(c_1) + \mathbf{q}(c_2) = 1$), for a pair of criteria (c_1, c_2) , if $c_1 \succ c_2$ and $\mathbf{p}(c_1) > \mathbf{q}(c_1)$, then $\mathbf{p} \succ \mathbf{q}$.

Substitutability means that if $\mathbf{p}(c_1) = \mathbf{q}(c_2)$ and \mathbf{p} and \mathbf{q} assign the same value for any other criterion c , then $c_1 \sim c_2$ implies $\mathbf{p} \sim \mathbf{q}$.

Decomposability employs the definition of \mathbf{p}_ω : for any distribution ω on $D(C)$, \mathbf{p}_ω denotes the distribution on C determined by

$$\mathbf{p}_\omega(c) = \sum_{\mathbf{p} \in D(C)} \omega(\mathbf{p}) \mathbf{p}(c).$$

Decomposability holds if and only if

$$\omega_1 = \omega_2 \text{ is equivalent to } \mathbf{p}_{\omega_1} = \mathbf{p}_{\omega_2}.$$

These properties are not as natural as they may seem to be. But to determine weights for preference criteria, what is going to be useful from Expected Utility Theory is the representation theorem asserting that for any set of criteria C and any preference relation \succ on C for which there exists an extension \succ to $D(C)$ with some properties, \succ identifies a unique distribution of weights u on C such that

$$u(c_1) \geq u(c_2) \text{ if and only if } c_1 \succ c_2$$

and a unique distribution of weights u on $D(C)$ such that, for ω_1 concentrated in c_1 and ω_2 concentrated in c_2 ,

$$u(\omega_1) \geq u(\omega_2) \text{ if and only if } c_1 \succ c_2.$$

This u is defined on $D(C)$ by

$$u(\omega) = \sum_{j=1}^m \omega(c_j) u(c_j) \text{ for any } \omega \in D(C).$$

If there is an outcome c_0 such that

$$u(c_0) = \sum_j \omega(c_j) u(c_j),$$

this outcome c_0 may be seen as the certainty equivalent of ω , in the sense that a distribution of weights with the unique outcome c_0 has the same expected utility of ω .

When the outcomes in C have numerical values, besides computing the expected utility we can compute the expected outcome

$$\sum \omega(c)c.$$

Sometimes we can compute also the utility of the expected outcome

$$u\left(\sum \omega(c)c\right).$$

A concave utility means risk aversion and a convex utility means risk proclivity, in the sense that the utility of the distribution on $D(C)$ that gives probability 1 to the expected outcome is, respectively, greater and smaller than its expected utility. Thus concavity represents a utility-decreasing evaluation of pure risk-bearing and convexity the contrary.

To understand the concept, suppose there are two lotteries, one that pays the expected value with certainty and another that pays the different values with their different probabilities. The utility of the first lottery is larger than the utility of the second for a risk-averse evaluation. On the other hand, giving risk a positive value would lead to a convex utility. Finally, neutrality with respect to risk would make indifferent the choice between the certain outcome and the same outcome in the average, so that

$$u\left(\sum \omega(x)x\right) = \sum \omega(x)u(x).$$

2.3.5 Example of Capacity Determination

Von Neumann and Morgenstern representation theorem provides the basis for the design of complex tools to derive the capacity of each set. Accepting the above listed conditions, instead of directly assigning a value to the set, its capacity may be derived from preferences between distributions. The evaluator will find easier to compare simple distributions involving the set than choosing a numeric value for the capacity of that set. The key idea consists of asking the decision maker appropriate questions about extreme distributions involving the set, to determine if its capacity is closer to one of two extreme values than to the other.

The procedure starts by determining the capacities of unitary sets $\{c\}$. For each such set, the evaluator answers a question about preference between two distributions: one of them assigns the value 1 to $\{c\}$; the other, a free choice between the distributions assigning the value 1 to any other unitary set in C (and consequently 0 to $\{c\}$).

If the evaluator prefers the first distribution, we conclude that the decision maker assigns to $\{c\}$ a value closer to 1 than to 0, that means a value larger than $\frac{1}{2}$; if the other is preferred, we conclude that the decision maker assigns to $\{c\}$ a value smaller than $\frac{1}{2}$; in the case of indifference, the quest ends, with the value $\frac{1}{2}$ assigned to $\{c\}$.

Suppose the answer to this first question is a preference for the distribution that assigns the value 1 to $\{c\}$. Then we proceed by asking the preference between the

distribution assigning the value 1 to $\{c\}$ and another assigning the value $\frac{1}{2}$ to $\{c\}$ and freely assigning the value $\frac{1}{2}$ to other element of C . If the evaluator prefers the first of these distributions, assigns to $\{c\}$ a value closer to 1 than to $\frac{1}{2}$, that means a value larger than $\frac{3}{4}$. If the other distribution is preferred, then that value is smaller than $\frac{3}{4}$. Indifference means that $\frac{3}{4}$ is the value assigned and this will be the capacity of $\{c\}$.

If the Von Neumann and Morgenstern assumptions are satisfied, indifference will appear after a while. After the capacities of the unitary sets are obtained, capacities for the sets of size two will be determined.

Suppose, for instance, the capacities of $\{c_1\}$ and $\{c_2\}$ are found to be 0.3 and 0.2. Then the capacity of $\{c_1, c_2\}$ is between 0.3 and 1. To determine this capacity, the decision maker will be asked about preference between a distribution assigning the value 1 to that set and another assigning the value 0.7 to a freely chosen subset of its complement. From preference for the value 0.7 for the complement follows that the capacity of $\{c_1, c_2\}$ for the decision maker is closer to 0.3 than to 1, what means that it is between 0.3 and 0.65. From preference for the other distribution follows a capacity from 0.65 to 1. From indifference, follow the final value of 0.65.

So, if, for instance, the interval from 0.3 to 0.65 follows from the answer obtained, the next choice will be between a distribution assigning to the set a value of 0.3 and another assigning the value 0.65. If the preference is for that assigning 0.65, we get a restriction to the interval between 0.475 and 0.65. If the preference is for the other, the capacity is between 0.3 and 0.475. If the answer is indifference, the value is 0.475. And so on.

After a logical sequence of such questions, we would eventually find a capacity representing the preference of the decision maker.

References

- Banzhaf, J. F. (1965). Weighted voting doesn't work: A mathematical analysis. *Rutgers Law Review*, 19(2), 317–343.
- Bourdieu, P. (1992). The Practice of Reflexive Sociology(Chap. III, pp. 217–240). In in P. Bourdieu, & L. Wacquant (Eds.), *Invitation to a reflexive sociology*. Chicago: University Chicago Press.
- Choquet, G. (1953). Theory of capacities. *Annales de l'Institut Fourier*, 5, 131–295.
- Ehrgott, M., Figueira, J. R., & Greco, S. (2010). *Trends in multicriteria decision analysis*. Dordrecht: Springer.
- Grabisch, M. (1997). k-order additive discrete fuzzy measures and their representation. *Fuzzy Sets and Systems*, 92, 167–189.
- Grabisch, M., & Roubens, M. (1999). An axiomatic approach to the concept of interaction among players in cooperative games. *International Journal of Game Theory*, 28, 547–565.
- Greco, J., & Ehrgott, S. (2005). *Multicriteria Decision Analysis: state of the art survey*, 3-24. Boston: Springer.
- Keeney, R. L. (1992). *Value-focused thinking: A path to creative decision making*. Cambridge, MA: Harvard University Press.
- Keeney, R. L., & Raiffa, H. (1976). *Decisions with multiple objectives' preferences and value tradeoffs*. New York: Wiley.

- Penrose, L. (1946). The Elementary statistics of majority voting. *Journal of the Royal Statistical Society*, 109(1), 53–57.
- Roy, B. (1996). *Multicriteria methodology for decision aiding*. Dordrecht: Kluwer.
- Roy, B. (2005). Paradigms and challenges. In J. Greco & S. Ehrgott (Eds.), *Multicriteria decision analysis: State of the art survey* (pp. 03–24). Boston: Springer.
- Saaty, T. L. (1980). *The analytic hierarchy process: Planning, priority Setting, resource allocation*. New York: McGraw-Hill. ISBN 0-07-054371-2.
- Saaty, T. L. (1990). How to make a decision: The analytic hierarchy process. *European Journal of Operational Research*, 48(1), 9–26.
- Saaty, T. L. (2003). Decision-making with the AHP: Why is the principal eigenvector necessary. *European Journal of Operational Research*, 145(1), 85–91.
- Shapley, L. S. (1953). A value for n -person games. In H. W. Kuhn, & A. W. Tucker, (Eds.), *Contributions to the theory of games* (Vol. II, pp. 307–317).
- Vargas, L. G. (1982). Reciprocal Matrices with Random Coefficients. *Mathematical Modelling*, 3, 69–81.
- Von Neumann, J., & Morgenstern, O. (1944). *Theory of games and economic behavior*. Princeton, NJ. Princeton University Press.

Probabilistic Composition of Preferences, Theory and
Applications

Parracho Sant'Anna, A.

2015, VII, 141 p., Hardcover

ISBN: 978-3-319-11276-3