A novel unsupervised capacity identification approach to deal with redundant criteria in multicriteria decision making problems

Guilherme Dean Pelegrina\textsuperscript{1} and Leonardo Tomazeli Duarte\textsuperscript{1}

\textsuperscript{1}Affiliation not available

February 14, 2024

Abstract

The use of the Choquet integral in multicriteria decision making problems has gained attention in the last two decades. Despite of its usefulness, there is the issue of how to define the Choquet integral parameters, called capacity coefficients, specially the ones associated with coalitions of criteria. A possible approach to address this issue is based on unsupervised learning, which aims to define such parameters with the goal of mitigating undesirable effects provided by intercriteria relations. However, current unsupervised approaches present some drawbacks, as there is no guarantee that the parameters are equally prioritized in the learning procedure. In this paper, we propose a novel unsupervised capacity identification approach which ensures a fair learning for all parameters. Moreover, in comparison with the existing methods, our proposal is less complex in terms of optimization, as it is based on a linear formulation. Experimental results in both synthetic and real datasets attest the applicability of our proposal.
A novel unsupervised capacity identification approach to deal with redundant criteria in multicriteria decision making problems

Guilherme Dean Pelegrina, Leonardo Tomazeli Duarte, Senior Member, IEEE.

Abstract—The use of the Choquet integral in multicriteria decision making problems has gained attention in the last two decades. Despite of its usefulness, there is the issue of how to define the Choquet integral parameters, called capacity coefficients, specially the ones associated with coalitions of criteria. A possible approach to address this issue is based on unsupervised learning, which aims to define such parameters with the goal of mitigating undesirable effects provided by intercriteria relations. However, current unsupervised approaches present some drawbacks, as there is no guarantee that the parameters are equally prioritized in the learning procedure. In this paper, we propose a novel unsupervised capacity identification approach which ensures a fair learning for all parameters. Moreover, in comparison with the existing methods, our proposal is less complex in terms of optimization, as it is based on a linear formulation. Experimental results in both synthetic and real datasets attest the applicability of our proposal.

Index Terms—multicriteria decision making, Choquet integral, 2-additive capacity, unsupervised capacity identification.

I. INTRODUCTION

Ranking alternatives is one of the problems addressed in multiple-criteria decision making (MCDM) [1]. A usual ranking approach in MCDM relies on data aggregation, which seeks to obtain global values for each alternative, by taking into account a set of decision criteria through an aggregation operator. In this approach, which stems from the multiattribute value theory (MAVT), non-redundancy among criteria is an expected property [2] and a cornerstone underlying the use of linear aggregation operators. However, intercriteria relations are often observed in practical scenarios [3], [4], which has been motivating the development of novel aggregation strategies that can, for instance, deal with redundant criteria so to avoid double counting of the same latent factor driving two different criteria [5]–[9].

There are several ways to deal with the issues related to the presence of intercriteria relations. For instance, in the case of linear aggregation approaches based on the weighted arithmetic mean (WAM), it is possible to mitigate redundancy among criteria through a procedure in which the criteria weights provided by the decision maker are subsequently adjusted, taking into account similarity measures between criteria (e.g., correlation and/or dependence measures) [10], [11]. Such an approach has the advantage of allowing the use of linear aggregation operators. On the other hand, adapting the weights in such a way may be uncomfortable for the decision maker as the weights subjectively defined by he/she according to his/her preference may be modified at the end of the process. In order to overcome this issue, nonlinear aggregation functions that can model interaction between criteria arise as an alternative approach in MCDM. This is the case of capacity-based aggregation functions, such as the Choquet integral [12] and the multilinear model [13].

Both the Choquet integral and the multilinear model have been used in several MCDM problems [14]–[16]. They are defined by means of a capacity, i.e., a normalized fuzzy measure. Aggregating the criteria information by using these functions allows the decision maker to model interaction between criteria and, therefore, overcome biased effects provided by redundancy in the decision data [17], [18]. However, the drawback in such approaches lies in the parameters identification, mainly because the number of capacity coefficients exponentially increases with the number of criteria.

To address the problem of capacity identification, several approaches have been proposed in the literature, which can be divided into supervised and unsupervised ones. In supervised approaches [16], [19]–[25], one identifies the capacity based on information provided by the decision maker, such as his/her preferences among criteria and/or preferences among a set of alternatives. Theses information are generally used into optimization problems in order to retrieve the capacity coefficients that satisfies the decision maker’s preferences. Some works deal with this task by assuming a 2-additive capacity, which drastically reduces the number of parameters to be identified [26], [27].

In unsupervised capacity identification, the adjustment of the capacity parameters relies exclusively on the decision matrix, i.e., without information provided by the decision-maker [28], [29]. The idea here is to retrieve a capacity that aligns with an expected goal, such as the mitigation of the redundancy among criteria. Such a goal must be achieved by exploiting statistics related to the decision matrix, with the main challenge being the task of associating those statistics with the capacity parameters governing the interaction between the decision criteria. In this work, we propose a
new formulation to the unsupervised capacity identification problem. Differently from previous approaches, such as [29], our proposal provides a more balanced strategy in terms of the association between cross-statistics among the criteria with the interactions indices obtained from the capacity adjustment step. In practical terms, such a feature avoids possible bias effects toward a criterion or a subset of them.

The paper is organized as follows. Section II provides a background comprising elementary definitions and an overview of existing unsupervised capacity identification approaches. Our proposal is introduced in Section III and assessed through a set of numerical experiments in Section IV. Finally, in Section V, our concluding remarks are presented.

II. BACKGROUND

A. Fuzzy measures in MCDM

Fuzzy measures have been successfully used in MCDM problems [30]. One of the most used capacity-based aggregation function is the Choquet integral [12]. The Choquet integral is a piecewise linear function whose parameters, the capacity coefficients \( \mu(A) \), are defined in all possible coalitions of criteria \( A \subseteq M \), where \( M = \{1, \ldots, m\} \) represents the set of \( m \) criteria. Three axioms must be satisfied to \( \mu \) be a capacity: if \( A \subseteq B \subseteq M \), \( \mu(A) \leq \mu(B) \leq \mu(M) \) (monotonicity), \( \mu(\emptyset) = 0 \) and \( \mu(M) = 1 \) (normalization).

Interestingly, the capacity \( \mu \) used in the definition of the Choquet integral has an association with the well-known solution concept in cooperative game theory called Shapley value [31]. Indeed, from the capacity coefficients, one may calculate the Shapley value \( \phi_j \) for each criterion \( j \) as follows:

\[
\phi_j = \sum_{A \subseteq M \setminus j} \frac{(m - |A| - 1)! |A|!}{m!} \left[ \mu(A \cup \{j\}) - \mu(A) \right],
\]

where \(|A|\) represents the cardinality of subset \( A \), \( \phi_j \geq 0 \) and \( \sum_{j=1}^{m} \phi_j = 1 \). One may interpret the Shapley values as the marginal contribution of features in the Choquet integral aggregation. Besides the Shapley values, one may also calculate the Shapley interaction index \( I_{j,j'} \) between pairs of features \( j, j' \) [32]. It is defined by

\[
I_{j,j'} = \sum_{A \subseteq M \setminus \{j,j'\}} \frac{(m - |A| - 2)! |A|!}{(m - 1)!} \left[ \mu(A \cup \{j,j'\}) - \mu(A \cup \{j\}) - \mu(A \cup \{j'\}) + \mu(A) \right],
\]

where \( I_{j,j'} \in [-1,1] \) and can be viewed as the interaction degree of coalition of criteria \( j, j' \). The interpretation depends on the sign of \( I_{j,j'} \). If \( I_{j,j'} < 0 \), we model a negative interaction (or redundant effect) between criteria \( j, j' \). If \( I_{j,j'} > 0 \), we model a positive interaction (or complementary effect) between criteria \( j, j' \). In the case where \( I_{j,j'} = 0 \), there is no interaction between criteria \( j, j' \). It is worth mentioning that (2) can be extended for all coalition of criteria [32]. However, for the scope of this paper, it is sufficient to take into account interactions only between pairs of criteria.

An issue in the Choquet integral is that the number of capacity coefficients exponentially increases with the number of criteria. Therefore, it is usual to consider a restricted version of the Choquet integral which is based on a 2-additive capacity, i.e., a capacity such that interaction exists only between pairs of criteria [33]. In this scenario and based on the Shapley values and interaction indices, the Choquet integral can be defined as follows:

\[
CI(x) = \sum_j x_j \left( \phi_j - \frac{1}{2} \sum_{j' > 0} (x_j \lor x_{j'}) |I_{j,j'}| \right) + \sum_{I_{j,j'} > 0} (x_j \land x_{j'}) I_{j,j'},
\]

where \( \lor \) and \( \land \) indicates the maximum and the minimum operator, respectively, and \( x = [x_1, \ldots, x_m] \) represents an alternative characterized by the criterion evaluations \( x_1, \ldots, x_m \). By using this formulation, the axioms of a capacity lead to the following conditions: \( I(\emptyset) = 0, \sum_{j=1}^{m} \phi_j = 1 \) and \( \sum_{j,j' \neq j} I_{j,j'} = 0 \).

B. Approaches for unsupervised capacity identification

There are several works that addressed the problem of unsupervised capacity identification [34], [35]. Among these works, a possible approach is based on principal component analysis (PCA) [28], a classical tool in unsupervised machine learning. While the solution proposed in [28] is based only on second-order statistics, which is interesting for matter of estimation, it cannot distinguish different types of relations between criteria, such as redundancy and synergy, from data. In order to overcome this limitation, the idea proposed in [29], which considered a simpler approach based on similarity measures and on the 2-additive capacity model. This approach will be further described in the sequence.

Let us define \( \psi_{j,j'} \) as a similarity measure between criteria \( j, j' \in M \), such as the Pearson's correlation coefficient [36]. The idea in [29] consists in associating such similarity measures to the interaction indices \( I_{j,j'} \). Note that, if one assumes the Pearson's correlation coefficient\(^1\) as the similarity measure, both \( \psi_{j,j'} \) and \( I_{j,j'} \) are in the range \([-1,1]\). However, as mentioned in Section II-A, negative (resp. positive) interaction indices model redundancies (resp. synergies) between criteria. Moreover, the greater the redundancy between criteria \( j, j' \) (i.e., a similarity measure close to 1), the lower the interaction index \( I_{j,j'} \) should be (i.e., close to -1) in order to model such an intercriteria relation. Therefore, for each pair of criteria \( j \) and \( j' \), the relation between the interaction index and the similarity measure is given by \( I_{j,j'} \approx -\psi_{j,j'} \).

As in this approach one assumed a 2-additive capacity, one also needs to estimate the Shapley power indices \( \phi_j \). In the absence of information provided by the decision maker about criteria preferences, [29] assigned the same marginal contribution for all criteria (principle of maximum entropy [37]). Mathematically, one defines \( \phi_j = 1/m \forall j = 1, \ldots, m \).

All Shapley power indices being defined, it only remains to adjust the interaction indices \( I_{j,j'} \) in such a way that they are as

\(^1\)Surely, other similarities measure could be used with the same characteristic.
close as possible from the negative of the associated similarity measure $\psi_{j,j'}$. However, as one needs to satisfy the axioms of a capacity, one may not assign $I_{j,j'} = -\psi_{j,j'}$. Indeed, the monotonicity condition reduces the $I_{i,i'}$ feasible region. One needs to find interaction indices $I_{j,j'} = -\psi_{j,j'}$ that is, an approximation of the similarity measures obtained from the decision matrix. On the other hand, for the normalization axioms, both are automatically satisfied as $\phi_j = 1/m \forall j = 1,\ldots,m$. The interaction indices are then obtained in [29] by solving the following optimization problem:

$$
\begin{align*}
\min_{\hat{\Psi}} & \quad \left\| \hat{\Psi} - \Psi \right\|_F \\
\text{s.t.} & \quad \frac{1}{m} - \frac{1}{2} \sum_{j\neq j'} \hat{\psi}_{j,j'} \geq 0, \ \forall j \in M \\
& \quad \hat{\psi}_{j,j'} - \hat{\psi}_{j',j} = 0, \ \forall j,j' \in M,
\end{align*}
$$

(4)

where $\hat{\Psi}$ is the positive semidefinite matrix (ensured by the constraint $\hat{\Psi} \geq 0$) which contains the approximated values for the interaction indices, and the matrix $\Psi$, given by

$$
\Psi = \begin{bmatrix} \psi_{1,1} & \psi_{1,2} & \cdots & \psi_{1,m} \\ \psi_{2,1} & \psi_{2,2} & \cdots & \psi_{2,m} \\ \vdots & \vdots & \ddots & \vdots \\ \psi_{m,1} & \psi_{m,2} & \cdots & \psi_{m,m} \end{bmatrix},
$$

(5)

contains the similarity measures obtained from the decision data. Finally, $\left\| \hat{\Psi} - \Psi \right\|_F$ is the Frobenius norm [38] between $\Psi$ and $\Psi$, i.e.,

$$
\left\| \hat{\Psi} - \Psi \right\|_F = \sqrt{\sum_{j=1}^{m} \sum_{j'=1}^{m} (\hat{\psi}_{j,j'} - \psi_{j,j'})^2}.
$$

(7)

As highlighted in [39], although the optimization problem proposed in [29] can adjust the interaction indices, there are some optimization issues that could be improved in order to deal with the unsupervised capacity identification by means of a quadratic problem. For instance, instead of minimizing the Frobenius norm between two matrices, one can directly minimize the difference between $I_{j,j'}$ and $-\psi_{j,j'}$. Besides reducing the number of variables from $m^2$ to $m(m-1)/2$, one also eliminates the second and the third constraints in (4). Another improvement can be achieved by eliminating the absolute values in the first set of constraints. One may consider all possible signals combinations of $I_{j,j'}$. The price to be paid is that one increases the number of such set of constraints from $m$ to $m(2^m)$. Given these modifications, one deals with the following quadratic optimization problem (see [39] for further details and [40] for an application):

$$
\begin{align*}
\min_{I} & \quad \sum_{j,j'} (I_{j,j'} + \psi_{j,j'})^2 \\
\text{s.t.} & \quad \phi_j - \frac{1}{2} \sum_{j \neq j'} \psi_{j,j'} \geq 0, \ \forall j \in M \\
& \quad \sum_j \phi_j = 1,
\end{align*}
$$

(8)

where $I = [\phi_1, \phi_2, \ldots, \phi_m, I_{1,2}, I_{1,3}, \ldots, I_{m-1,m}]$. Note that, in (8), we did not redefine the values of $\phi_j$, $j = 1,\ldots,m$. One may use any approach or available information to define them. Clearly, if one assumes $\phi_j = 1/m \forall j = 1,\ldots,m$, the interaction indices adjusted in (8) will be the same as solving (4).

III. PROPOSED APPROACH

As mentioned in Section I, there are drawbacks in the works discussed in the previous section. Moreover, the solution of the optimization problems (8) and (4) may lead to disparate ratios between the adjusted interaction indices and the target values. In other words, one may achieve a solution in which a $I_{j,j'}$ is much more close to $-\psi_{j,j'}$ than another $I_{j'',j'''}$ is to $-\psi_{j'',j'''}$. In this situation, as our aim is to adjust interaction indices in order to overcome biased effects provided by redundant criteria, one may introduce another source of bias in the aggregation procedure.

With the purpose of avoiding this weakness, we propose in this paper a novel approach whose idea is to ensure that the ratio between interaction indices and similarity measures are the same for all pairs of criteria. Let us define such a ratio by $t$. In order to ensure this ratio for all pairs of criteria, we introduce the following set of constraints:

$$
I_{j,j'} - \text{sign}(\psi_{j,j'}) t \psi_{j,j'} = 0, \ \forall j,j' \in M,
$$

(9)

where $\text{sign}(z)$ is the sign function (i.e., $\text{sign}(z) = 1$ if $z \geq 0$ and $0$ otherwise). In order to minimize the difference between interaction indices and similarity measures, our goal is to maximize $t$. As $t$ is a ratio, we consider that it lies in the range $[0,1]$. These bounds also ensure that $-\psi_{j,j'} \leq I_{j,j'} \leq 0$, if $\psi_{j,j'} \geq 0$, and $0 \leq I_{j,j'} \leq -\psi_{j,j'}$, if $\psi_{j,j'} \leq 0$, for all $j,j' \in M$.

Besides these novel constraints that ensure the same ratio between interaction indices and similarity measures for all pairs of criteria, we also adjusted the constraints that satisfy the monotonicity condition. Instead of considering all possible signals combinations of $I_{j,j'}$, we assume that $|I_{i,i'}| = \text{sign}(\psi_{i,i'}) |I_{i,i'}|$. Therefore, if $\psi_{i,i'} \geq 0$, $|I_{i,i'}| = I_{i,i'}$, as $I_{i,i'} \geq 0$. On the other hand, if $\psi_{i,i'} \geq 0$ and, therefore, $I_{i,i'} \leq 0$, $|I_{i,i'}| = -I_{i,i'}$. This reduces the number of such set of constraints from $m(2^m)$ to $m$ (as in the optimization problem (4), but without the use of absolute values).

Based on the aforementioned modifications, the optimization problem that we deal in our proposal is given by:

$$
\begin{align*}
\max_{t,I} & \quad t \\
\text{s.t.} & \quad \phi_j - \frac{1}{2} \sum_{j \neq j'} \text{sign}(\psi_{j,j'}) I_{j,j'} \geq 0, \ \forall j \in M \\
& \quad I_{j,j'} - \text{sign}(\psi_{j,j'}) t \psi_{j,j'} = 0, \ \forall j,j' \in M \\
& \quad \sum_j \phi_j = 1 \\
& \quad 0 \leq t \leq 1.
\end{align*}
$$

(10)

Note that we did not redefine the values of $\phi_j$, $j = 1,\ldots,m$. Indeed, we propose an approach that (i) can take into account the information provided by the decision maker with respect to the relative importance of each criterion (which is represented by the Shapley power indices) and (ii) softens the biased effect.
introduced by redundancy in the decision data by modeling intercriteria relations (expressed by means of the interaction indices). Therefore, our proposal mix both objective and subjective elicitation mechanisms to adjust the aggregation function parameters.

In summary, the benefits of our approach in comparison with the ones discussed in Section II-B are the following ones:

- We ensure that the ratio between interaction indices and similarity measures are the same for all pairs of criteria.
- We eliminates the use of absolute values or the excessive number of signals combinations when dealing with the set of constraints associated with the monotonicity condition.
- Our optimization problem is linear and, therefore, can be easily solved by linear programming solvers [41].

The negative points in our proposal is that we introduce \( m(m - 1)/2 \) constraints and the bounds for the ration \( t \). However, as we end up in a linear optimization problem, these novel elements are irrelevant in terms of computational effort.

### IV. NUMERICAL EXPERIMENTS

This section evaluates the application of our approach in both synthetic and real datasets. We compare our proposal with the weighted arithmetic mean (WAM) and the improved version of Duarte’s model, expressed by the optimization problem (8). As a similarity measure, we adopted the Pearson’s correlation coefficients. The numerical experiments and obtained results are presented in the sequel.

#### A. Experiments on synthetic dataset

Our first experiment evaluates the application of our approach in a synthetic scenario with redundant criteria. For this purpose, we randomly generated (according to a uniform distribution in the range [0, 1]) \( n = 200 \) samples described by \( m = 3 \) criteria. The scatter plots of pairs of criteria are presented in Figure 1. Note that there are intercriteria relations within this dataset. Indeed, the Pearson’s correlation coefficients are \( \rho_{1,2} = 0.6657, \rho_{1,3} = 0.1098 \) and \( \rho_{2,3} = -0.2370 \).

![Scatter plots (synthetic dataset)](image_url)

**Fig. 1: Scatter plots (synthetic dataset).**

We assumed no knowledge of further information from the decision maker with respect to the criteria weights. Therefore, we considered the same weights and the same Shapley power indices for all criteria, i.e., \( 1/3 \). The obtained interaction indices and ratios \( t = -I_{j,j'}/\rho_{j,j'} \) for all \( j, j' \in M \), for both Duarte’s model and our proposal are presented in Table I. Based on these parameters and on the criteria weights/power indices, the first 10 alternatives in the obtained ranking for each aggregation function are described in Table II. One may note from Table II some similarities between the rankings obtained by both Duarte’s model and our proposal. However, the ranking given by the weighted arithmetic mean is very different from those based on capacity coefficients. Indeed, as both Duarte’s model and our proposal can model criteria interactions and deal with redundant information, these methods could soften the similarity between criteria 1 and 2 in the aggregation function (see the interaction index between criteria 1 and 2 presented in Table I). Therefore, the ranking provided by these models favored alternatives with a very good evaluation in criteria 3 and a very good evaluation in at least one of the remaining criteria 1 and 2 (the redundant ones). See, for instance, alternatives #84, #42 and #88, which achieved a better position when using Duarte’s model or our proposal in comparison with the WAM. On the other hand, in the ranking provided by the WAM, the first positions are composed by alternatives with very good evaluations only in criteria 1 and 2. As the WAM does not take into account redundancy among criteria, this aggregator double-counted the same latent factor, which introduced bias towards alternatives with very good evaluations in only two criteria.

### TABLE I: Interaction indices and ratios (synthetic dataset).

<table>
<thead>
<tr>
<th>Interaction index (ratio)</th>
<th>Duarte’s model</th>
<th>Our proposal</th>
</tr>
</thead>
<tbody>
<tr>
<td>( I_{1,2} )</td>
<td>(-0.5477)</td>
<td>(-0.4917)</td>
</tr>
<tr>
<td>((-I_{1,2}/\rho_{1,2}))</td>
<td>( \approx 82.27% )</td>
<td>( \approx 73.86% )</td>
</tr>
<tr>
<td>( I_{1,3} )</td>
<td>(-0.1098)</td>
<td>(-0.0811)</td>
</tr>
<tr>
<td>((-I_{1,3}/\rho_{1,3}))</td>
<td>( 100%)</td>
<td>( 73.86%)</td>
</tr>
<tr>
<td>( I_{2,3} )</td>
<td>(-0.1790)</td>
<td>(-0.1790)</td>
</tr>
<tr>
<td>((-I_{2,3}/\rho_{2,3}))</td>
<td>( \approx 50.21% )</td>
<td>( \approx 73.86% )</td>
</tr>
</tbody>
</table>

### TABLE II: Ranking of alternatives (synthetic dataset).

<table>
<thead>
<tr>
<th>Alt. #</th>
<th>Criteria evaluations</th>
<th>Ranking position</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>( x_1 ) ( x_2 ) ( x_3 )</td>
<td>WAM</td>
</tr>
<tr>
<td>181</td>
<td>0.916 0.852 0.831</td>
<td>1</td>
</tr>
<tr>
<td>84</td>
<td>0.934 0.600 0.872</td>
<td>7</td>
</tr>
<tr>
<td>42</td>
<td>0.934 0.522 0.933</td>
<td>11</td>
</tr>
<tr>
<td>23</td>
<td>0.862 0.810 0.818</td>
<td>3</td>
</tr>
<tr>
<td>72</td>
<td>0.949 0.939 0.660</td>
<td>2</td>
</tr>
<tr>
<td>88</td>
<td>0.622 0.834 0.874</td>
<td>16</td>
</tr>
<tr>
<td>183</td>
<td>0.856 0.695 0.845</td>
<td>9</td>
</tr>
<tr>
<td>77</td>
<td>0.894 0.939 0.652</td>
<td>4</td>
</tr>
<tr>
<td>66</td>
<td>0.907 0.706 0.681</td>
<td>18</td>
</tr>
<tr>
<td>192</td>
<td>0.872 0.744 0.719</td>
<td>15</td>
</tr>
</tbody>
</table>

By comparing both Duarte’s model and our proposal, one may remark that the ratios between the adjusted interaction indices and the similarity measure are very different in Duarte’s model. On the other hand, in our proposal, all ratios are the same (\( t \approx 73.86\% \)). This is one of the benefits of our proposal. As we ensure the same ratio, we avoid another source of bias provided by disparate adjustments in the interaction indices. We may note this benefit when comparing the rank position of alternatives #84 and #42. Duarte’s method modeled more redundancy between criteria 1 and 2 (\( I_{1,2} = -0.5477 \)) in comparison with our proposal (\( I_{1,2} = -0.4917 \)). This helped alternative #42 to achieve the second position, as it has a
very good evaluation in criteria 3 and criteria 1 (it neglected criterion 2 as it modeled a strong redundancy with criterion 1). In our proposal, by balancing the adjusted interaction indices with respect to the similarity measures, we modeled a lower degree of redundancy between criteria 1 and 2 but we increased the synergy between criteria 2 and 3 \((I_{2,3} = 1750)\). This helped alternative \#84 to achieve the second position, as good evaluations in both criteria 2 and 3 are of importance in this aggregator.

B. Experiments on real dataset

We also evaluated the proposed unsupervised capacity identification approach in a real dataset. We considered the problem of ranking \(n = 171\) countries based on \(m = 4\) criteria: forest area (c1, in % of land area), CO2 emission (c2, in metric tons per capita), life expectancy at birth (c3, in years) and Gross National Income (GNI) per capita (c4, in US$). All these data were collected from The World Bank at https://data.worldbank.org/. In order to transform all criteria into the same scale, we normalize the data into the range \([0, 1]\) (1 is the better evaluation and 0 is the worst). Note that, in CO2 emission, lower the value, better the evaluation. In the remaining criteria, the higher the better.

Figure 2 presents the scatter plots of pairs of criteria. Note the positive correlation between life expectancy and GNI per capita (\(\rho_{c1,c4} = 0.6733\)) and the negative correlations between CO2 emission and both life expectancy (\(\rho_{c2,c3} = -0.5999\)) and GNI per capita (\(\rho_{c2,c4} = -0.6520\)). The remaining correlations are not significant (\(\rho_{c1,c2} = 0.1064, \rho_{c1,c3} = -0.0152\) and \(\rho_{c1,c4} = -0.0867\)). The obtained interaction indices by applying Duarte’s model and our proposal are presented in Table III. In both models (as well as in the WAM), we assumed the same value (i.e., 1/4) for all Shapley power indices (and WAM weights). As in the synthetic dataset, we also note disparate ratios when using the Duarte’s model. Our approach can ensure the maximum equal ratio for all interaction indices.

![Scatter plots (real dataset).](image)

A comparison in the derived rankings of alternatives (for the first 10 alternatives) is shown in Table IV. Remark that the ranking obtained by the WAM is very different from the capacity-based methods. As it does not model intercriteria relations, the ranking is biased by the presence of similar criteria. One may see, for instance, that the second position in its ranking (alternative \#74) is composed by an alternative with very good evaluations in redundant criteria (life expectancy at birth GNI per capita). This is also the case for the first position, whose evaluations for are \([c1, c2, c3, c4] = [0.3536, 0.4338, 0.9267, 1.0000]\). This alternative (alternative \#92) is far from the first positions in the rankings obtained by the Duarte’s model (20th position) and our proposal (21st position). As the capacity-based approaches model intercriteria relations, the redundant effect between life expectancy at birth and GNI per capita is softened and alternatives with good evaluations in CO2 emission (which has synergies with both life expectancy at birth and GNI per capita) and forest area (the independent criterion) are favored.

![Scatter plots (real dataset).](image)

### TABLE III: Interaction indices and ratios (real dataset).

<table>
<thead>
<tr>
<th>Interaction index (ratio)</th>
<th>Duarte’s model</th>
<th>Our proposal</th>
</tr>
</thead>
<tbody>
<tr>
<td>(I_{1,2}) ((-I_{1,2}/\rho_{1,2}))</td>
<td>(\approx 0) ((\approx 0%))</td>
<td>(-0.0377) ((\approx 35.40%))</td>
</tr>
<tr>
<td>(I_{1,3}) ((-I_{1,3}/\rho_{1,3}))</td>
<td>(\approx 0) ((0%))</td>
<td>0.0054 ((\approx 35.40%))</td>
</tr>
<tr>
<td>(I_{1,4}) ((-I_{1,4}/\rho_{1,4}))</td>
<td>(\approx 0) ((0%))</td>
<td>0.0807 ((\approx 35.40%))</td>
</tr>
<tr>
<td>(I_{2,3}) ((-I_{2,3}/\rho_{2,3}))</td>
<td>(\approx 0.735) ((\approx 44.65%))</td>
<td>0.1982 ((\approx 35.40%))</td>
</tr>
<tr>
<td>(I_{2,4}) ((-I_{2,4}/\rho_{2,4}))</td>
<td>(\approx 0.380) ((\approx 38.34%))</td>
<td>0.2309 ((\approx 35.40%))</td>
</tr>
<tr>
<td>(I_{3,4}) ((-I_{3,4}/\rho_{3,4}))</td>
<td>(\approx 0.250) ((\approx 37.13%))</td>
<td>(-0.2384) ((\approx 35.40%))</td>
</tr>
</tbody>
</table>

### TABLE IV: Ranking of alternatives (real dataset).

<table>
<thead>
<tr>
<th>Alt. #</th>
<th>Criteria evaluations</th>
<th>Ranking position</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(x_1)</td>
<td>(x_2)</td>
</tr>
<tr>
<td>113</td>
<td>0.386</td>
<td>0.721</td>
</tr>
<tr>
<td>69</td>
<td>0.005</td>
<td>0.822</td>
</tr>
<tr>
<td>82</td>
<td>0.661</td>
<td>0.501</td>
</tr>
<tr>
<td>139</td>
<td>0.631</td>
<td>0.731</td>
</tr>
<tr>
<td>137</td>
<td>0.222</td>
<td>0.651</td>
</tr>
<tr>
<td>58</td>
<td>0.336</td>
<td>0.671</td>
</tr>
<tr>
<td>143</td>
<td>0.382</td>
<td>0.806</td>
</tr>
<tr>
<td>76</td>
<td>0.332</td>
<td>0.786</td>
</tr>
<tr>
<td>74</td>
<td>0.116</td>
<td>0.693</td>
</tr>
<tr>
<td>51</td>
<td>0.586</td>
<td>0.758</td>
</tr>
</tbody>
</table>

Another remark in Table IV is the slightly difference between the rankings of Duarte’s model and our proposal. This is due to the disparities in the ratios between interaction indices and correlation coefficients. Therefore, we attest that disparate ratios can introduce another source of bias which may interfere in the ranking of alternatives.

V. Conclusions

In this work, we proposed a novel unsupervised approach to adjust the parameters (capacities) that are used in the Choquet integral to model the relevance of a set of criteria and the coalitions of them. Our approach extends previous works [29],...
in two respects. Firstly, we propose a strategy to mitigate bias towards the influence of different pairs of criteria into the aggregation process. As a second point, we provided a simpler optimization problem which is based on a linear formulation, which simplifies the implementation, paving the way for more complex scenarios, as, for example, the case with many criteria.

Experiments with both real and synthetic data showed that our method was able, as in [29], [39], to take into account redundancy between criteria in an unsupervised fashion, while promoting a more balanced solution in terms of the influence between different pairs of criteria. It is worth mentioning that our approach can be extended to other nonlinear aggregators founded on set functions, such as the multilinear models.

REFERENCES


