Reaching social consensus family budgets: The Spanish case

J.M. Cascón[a], T. González-Arteaga[b], R. de Andrés Calle[c,*]

[a]*Department of Economics and Economic History,
Institute on Fundamental Physics and Mathematics University of Salamanca,
E37007 Salamanca, Spain*
[b]*BORDA and PRESAD Research Groups and
Multidisciplinary Institute of Enterprise (IME), University of Valladolid,
E47011 Valladolid, Spain*
[c]*BORDA Research Unit, PRESAD Research Group and
Multidisciplinary Institute of Enterprise (IME), University of Salamanca,
E37007 Salamanca, Spain*

Abstract

The study of family budgets has been traditionally used to analyse consumers' behaviour and estimate cost-of-living since the end of 19th century. Generally speaking, the computation of the budgets has been based on two different methodologies, the prescriptive and the descriptive method. Both present several drawbacks like the comparison among different areas, family types and over time.

This paper proposes a new methodology for reaching family budgets, namely *social consensus family budgets*, to overcome such problems and examine the main features of the novel approach. The suggested method uses the minimization of the differences with respect to the consumer's preferences to obtain a solution that summarizes single behaviour into a social preference. This approach is especially conceived for preferences on possibly related-expenditure groups. In addition, several algorithms are introduced to compute the social family budgets. Finally, the contribution includes the Spanish case as an example of reaching some social consensus family budgets in order to show the operational character and intuitive interpretation of the proposal approach.

---

[*]Corresponding author

Email addresses: `casbar@usal.es` (J.M. Cascón),
`teresa.gonzalez.arteaga@uva.es` (T. González-Arteaga), `rocioac@usal.es` (R. de Andrés Calle)

---

1. Introduction

Taking a decision means select among several alternatives. When decision depends only on the opinion of one person, its study is related to *Rational Decision Making Theory*. If decision depends on a group of people, the process adds a layer of complexity and its study is included in *Group Decision Making Theory*. A group decision making problem involves, basically, a set of individuals and set of available alternatives. Under these simple assumptions, each individual expresses her/his preferences on alternatives and the problem is solved when individual's preferences are mixed in a group's preference[1]. Generally speaking, the achievement of an overarching solution has been considered like an aggregation problem from Borda [4], Kendall [5] (in a voting context) to Saari and Merlin [6], Meskanen and Nurmi [7], Klamler [8, 9] and de Andrés Calle, García-Lapresta and González-Pachón [10] (using distance-based aggregation rules).

Nowadays, obtaining of a group's preference entails not only aggregating but also finding the best alternative or the solution that more consensus conveys among decision-makers. Due to this fact, there has been an increasing interest in developing approaches to solve this new paradigm in this kind of problems like *goal programming* techniques [11], operational procedures [12], cluster models [13], p-indicators [14], multi-objective programming theories [15], control strategies [16] and conditional probabilities [17], among others.

This paper is focused on the contribution of González-Arteaga, Alcantud and de Andrés Calle [18] which proposes a distance-based methodology to obtain group consensus solutions, namely *Mahalanobis social consensus solutions*. This particular methodology is especially designed for correlated alternatives and obtains group solutions minimizing the dissensus among individual's preferences[2]. However, it is needed to stress that the main weakness of the approach presented in [18] is its computational complexity even for a small set of alternatives[3]. In

---

[1]Preferences on alternatives can be expressed by means of different types of preference relations such as ordinal [1], linguistic [2], multiplicative [3], and so on.

[2]The paper is focused on preferences on alternatives express by means of complete preorders.

[3]In group decision making context, persons' maximum efficiency on one-dimensional absolute judgement coincides with the ability to differentiate among four and eight alternatives (see [19]).

particular, the formulated optimization problem can be tackled by a force-brute attack, but this strategy must be rejected because the cardinality of the set of all complete preorders on the alternatives set grows dramatically in terms of the number of alternatives.

Taking into account the aforementioned limitations, the objectives of this research are to improve and to extend the proposal of [18], by means of designing a strategy to compute efficiently Mahalanobis consensus solutions, particularly when the number of alternatives is not small. To achieve these goals, the central thesis of this paper is to writte the optimization problem presented in [18] as a set of *mixed-integer* quadratic programming problems and solve them, assuming that they are NP-hard decision problems [20]. Then, this primary aim is divided into three main research contributions:

- Firstly, prior to transform the original problem into mixed-integer quadratic problems, the *canonical codification* included in [18] to codify ordinal information must be adapted due to the feasible set wich is determined by equality linear constraints using continuous and binary variables. Therefore, the concept of α-*index* is introduced as well as the characterization of the canonical codified complete preorder based on the α-index.

- Secondly, the feasible set is setted by means of a novel recursive algorithm, the *Recursive α-index algorithm* (RAI). As has already been evoked, one of the main obstacles in the resolution of this type of optimization problems is the setting of the feasible set for a "not small" set of alternatives. To be precise, the number of preorders on a set of k alternatives is equal to $\sum_{j=0}^{k} j! \cdot S(k, j)$, being $S(k, j)$ the Stirling number of the second kind[4]. To our knowledge, there is not a procedure to determine feasible set on set of preorders in the specialized literature[5].

- Finally, an algorithm to efficiently compute Mahalanobis consensus solutions, the *Mahalanobis consensus solution algorithm* (MCSA), is defined.

In closing and with the goal of putting in practice the aforementioned researches, this study aims to contribute to this growing area of research by proposing a new methodology to compute family's budgets and then analyse consumers'

---

[4]see [21] and The On-Line Encyclopedia of Integer Sequences® (OEIS®) Wiki, http://oeis.org/A000670

[5]In [22] there is another procedure to compute the number of preorders associated to a set of alternatives, but not to identified them.

behaviour. In these times of a new global economy, consumers' preferences are changing all the time (modifications in living standards, tendency, technology, et cetera) and the establishment of consumers' behaviour patterns is a key research area. The study of consumers' preferences and their behaviour is one of the greatest challenges for companies because product marketing is largely dependent on them (see [23], [24] and [25], ammong others) and for Governments because the efficacy and efficiency of public policies (regulatory, welfare, etc.) is also dependent on them (see [26]). Therefore, this paper sets out to investigate the usefulness of the proposed methodologies to determine consumers' preferences and the ranking of those expenditure groups that best agrees with family preferences. That is, the ranking that minimizes the disagreement among families' ranking expenditure groups considering the relation among expenditure groups. Specifically, the focus of the work is on the Spanish Economy. Based on Spanish data on household budgets from 2016, the ranking of the expenditure groups for the Spanish consumers, i.e., the Spanish consumers' behaviour by the year 2016, is obtained by means of the proposed methodologies. Among other research consequences, getting Spanish consumers' behaviour patterns will allow companies and also the Spanish Government to improve their marketing and public policies, respectively.

This paper is organized as follows. Section 2 introduces some starting points like notation and basic definitions. In Section 3, the main research contributions of this paper to compute efficiently Mahalanobis consensus solutions are introduced. Section 4 includes the real case of study on Spanish Economy to obtain family budgets. This section analyses the results obtained and provides discussions on them. Finally, some conclusions and further research work are presented in Section 5.

2. Starting points

Prior to commencing the study, this section reviews some previous definitions as well as the approach proposed by González-Arteaga, Alcantud and de Andrés Calle [18] to obtain social consensus solutions under the possibility of alternatives correlated.

2.1. Notation

Let us denote by $\mathbf{N} = \{1, 2, ..., n\}$, $n > 1$ a society of individuals and by $\mathbf{X} = \{x_1, ..., x_k\}$, a finite set of alternatives, options or issues. It is assumed that

4

the cardinality of $\mathbf{X}$ is at least 2, $|\mathbf{X}| \geqslant 2$. For simplicity of notation, we write s instead of issue x_s.

Without loss of generality it is required that experts rank alternatives by means of complete preorders[6]. Centring on reflecting real situations, the representation of agents' opinions allows ties among alternatives[7]. $\mathbf{W}(\mathbf{X})$ denotes the set of all complete preorders on $\mathbf{X}$.

Let $\mathcal{R} \in \mathbf{W}(\mathbf{X})$ be a complete preorder on $\mathbf{X}$, then $x_i \succ_{\mathcal{R}} x_j$ means x_i is strictly preferred to x_j, $x_i \sim_{\mathcal{R}} x_j$ means x_i and x_j are equally preferred and $x_i \succcurlyeq_{\mathcal{R}} x_j$ means alternative x_i is at least as good as x_j. Any permutation π of the alternatives $\{x_1, \ldots, x_k\}$ determines another preorder ${}^{\pi}\mathcal{R}$ given by:

$$x_i \, {}^{\pi}\mathcal{R} \, x_j \;\Leftrightarrow\; x_{\pi^{-1}(i)} \, \mathcal{R} \, x_{\pi^{-1}(j)}, \text{ for every } i, j \in \{1, \ldots, k\}.$$

Let $\mathcal{P} = (\mathcal{R}_1, ..., \mathcal{R}_n) \in \mathbf{W}(\mathbf{X}) \times \ldots \times \mathbf{W}(\mathbf{X}) = \mathbf{W}(\mathbf{X})^n$ be the *profile* of $\mathbf{N}$ (the society) on $\mathbf{X}$ (the set of alternatives). The $\mathcal{R}_i \in \mathcal{P}$ element stands for the i-th individual's preferences on the k alternatives for each $i = 1, ..., n$.

Dealing with ordinal information like preorders, necessarily involves to establish how it is represented. The first serious discussion and analysis of transforming or codifying ordinal information into numerical values emerged with Borda's work [4]. Subsequently, several procedures have been suggested to that purpose as [27], [28] and [29], among others.

The choice of a strong codification procedure is an essential issue to accomplish any methodology on ordinal information. Therefore, the next subsection provides the codification method proposed in [18], the *canonical codification* and includes a slightly modified characterization of it.

2.2. Codification of preferences

A *codified complete preorder* of $\mathcal{R} \in \mathbf{W}(\mathbf{X})$ is a real vector $\mathbf{m}_{\mathcal{R}} = (m_1, \ldots, m_k)$, being m_j the codification value corresponding to alternative x_j and satisfying $m_i \geq m_j \Leftrightarrow x_i \succcurlyeq_{\mathcal{R}} x_j$. Note that any $\mathbf{m} \in \mathbb{R}^n$ produces a unique preorder $\mathcal{R}$, however every preorder $\mathcal{R}$ can be associated with infinitely many vector $\mathbf{m} \in \mathbb{R}^n$.

---

[6]Complete preorders or weak orders, i.e., a complete and transitive binary relation on $\mathbf{X}$.

[7]Throughout this paper, calligraphic letters are used to denote preorders and its corresponding profiles, capital letters to represent matrices, capital bold typeface letters for sets, bold lowercase letters to denote vectors, and finally scalars are represented by lowercase letters.

Analogously, a *codified profile* of $\mathcal{P} = (\mathcal{R}_1, \ldots, \mathcal{R}_n)$ is a $n \times k$ real matrix:

$$M_{\mathcal{P}} = (\mathbf{m}_{\mathcal{R}_1}, \ldots, \mathbf{m}_{\mathcal{R}_n}) \in \mathbb{M}_{n \times k}$$

where m_{ij} is the codification value of agent i over the alternative x_j. We denote $\mathbb{M}_{n \times k}$ the set of all $n \times k$ real matrices.

Definition 1. *Given* $\mathcal{R} \in \mathbf{W}(\mathbf{X})$ *be a complete preorder on* X, *its* canonical codified complete preorder *is defined by the vector:*

$$\mathbf{c}_{\mathcal{R}} = (c_1, ..., c_k) \in (\{1, \ldots, k\})^k$$

where c_j *is the number of alternatives that are graded at most as good as* x_j, $c_j = |\{q : x_j \succcurlyeq_{\mathcal{R}} x_q\}|$.

The set of all canonical codified complete preorders associated to $\mathbf{W}(\mathbf{X})$ is denoted by $\mathcal{F} = \mathcal{F}(\mathbf{W}(\mathbf{X}))$.

A *canonical codified profile* associated with $\mathcal{P} = (\mathcal{R}_1, ..., \mathcal{R}_N) \in \mathbf{W}(\mathbf{X})^n$ is an $n \times k$ real matrix, namely $M_{\mathcal{P}} = (\mathbf{c}_{\mathcal{R}_1}, ..., \mathbf{c}_{\mathcal{R}_n}) \in \mathbb{M}_{n \times k}$, where the row i, denoted by $\mathbf{c}_{\mathcal{R}_i}$, represents the canonical codified complete preorder associated with $\mathcal{R}_i$.

This particular codification was characterised in [18, Proposition 1]. This contribution includes such proposition although slightly modified. The same proof is still valid and may be omitted.

Proposition 1. *A vector* $\mathbf{c} = (c_1, \ldots, c_k) \in (\{1, \ldots, k\})^k$ *is the canonical codified complete preorder* $\mathbf{c}_{\mathcal{R}}$ *associated with* $\mathcal{R} \in \mathbf{W}(\mathbf{X})$ *(i.e.* $\mathbf{c}_{\mathcal{R}} \in \mathcal{F}$*) if and only if the increasingly ordered vector* $\uparrow\mathbf{c} = (c_{(1)}, \ldots, c_{(k)})$ *satisfies:*

$$c_{(k)} = k. \tag{1}$$
$$c_{(j)} = c_{(j+1)} - t_{j+1}d_{j+1}, \quad j \in k-1, \ldots, 1. \tag{2}$$

where t_{j+1} *is the number of values equal to* $c_{(j+1)}$ *and* $d_{j+1} = \begin{cases} 0 & \text{if } c_{(j+1)} = c_{(j)}, \\ 1 & \text{otherwise.} \end{cases}$

For abbreviation, a preorder $\mathcal{R}$ can be identified with its canonical codification $\mathbf{c}_{\mathcal{R}}$, i.e., $\mathcal{R} \equiv \mathbf{c}_{\mathcal{R}}$.

Once the canonical codified procedure has been reviewed, the methodology used in this paper to find the best option or solution from the set of alternatives is now presented. The following subsection reminds the proposal of [18] to obtain global solutions based on the Mahalanobis distance. These solutions are the rankings of issues or alternatives that best concurs with individual preferences, that is, the ranking that maximizes the agreement among agents.

2.3. The Mahalanobis consensus solution for ordinal information

The *Mahalanobis social consensus solution* provides a complete preorder $\widehat{\mathcal{R}}$ that is the *best agreement* taking into consideration the Mahalanobis distance. Although the Mahalanobis consensus solution can be computed for any codification of preorders, our presentation is restricted to the canonical codification introduced at Subsection 2.2.

First, the *Mahalanobis consensus distance function* (MCDF) is defined. It is the objective function that must be optimized to get the consensus solution.

Definition 2. *Let* $\Sigma \in \mathbb{M}_{k \times k}$ *be a definite positive matrix. Given a profile* $\mathcal{P} = (\mathcal{R}_1, \ldots, \mathcal{R}_n) \in \mathbf{W}(\mathbf{X})^n$ *of complete preorders and its canonical codified profile,* $M_{\mathcal{P}} = (\mathbf{c}_{\mathcal{R}_1}, \ldots, \mathbf{c}_{\mathcal{R}_n}) \in \mathbb{M}_{n \times k}$. *The* Mahalanobis consensus distance function *(MCDF) is a mapping*

$$\mathcal{C}_{\Sigma, \mathcal{P}} : \mathbf{W}(\mathbf{X}) \longrightarrow [0, \infty)$$

that assigns to any complete preorder $\mathcal{R} \in \mathbf{W}(\mathbf{X})$ *with canonical codification* $\mathbf{c}_{\mathcal{R}}$ *the real number:*

$$\mathcal{C}_{\Sigma, \mathcal{P}}(\mathcal{R}) = \sum_{i=1}^{n} d_{\Sigma}(\mathbf{c}_{\mathcal{R}_i}, \mathbf{c}_{\mathcal{R}}) = \sum_{i=1}^{n} (\mathbf{c}_{\mathcal{R}_i} - \mathbf{c}_{\mathcal{R}}) \Sigma^{-1} (\mathbf{c}_{\mathcal{R}_i} - \mathbf{c}_{\mathcal{R}})^t \qquad (3)$$

Remark 1. *If* $C_{\mathcal{P}, \mathcal{R}} \in \mathbb{M}_{n \times k}$ *denotes a matriz whose i-th row is* $\mathbf{c}_{\mathcal{R}_i} - \mathbf{c}_{\mathcal{R}}$, *we can rewrite MCDF as:*

$$\mathcal{C}_{\Sigma, \mathcal{P}}(\mathcal{R}) = \operatorname{tr}(C_{\mathcal{P}, \mathcal{R}} \, \Sigma^{-1} \, C_{\mathcal{P}, \mathcal{R}}^t), \qquad (4)$$

where $\operatorname{tr}(\cdot)$ *is the trace operator (* $\operatorname{tr}(A) = \sum_{i=1}^{n} a_{ii}$ *).*

Now, the *Mahalanobis consensus solution* can be introduced as the preorder that minimizes the Mahalanobis distance to a given profile. Formally:

Definition 3. *The* Mahalanobis consensus solution $\widehat{\mathcal{R}} \in \mathbf{W}(\mathbf{X})$ *of a given profile* $\mathcal{P} = (\mathcal{R}_1, \ldots, \mathcal{R}_n) \in \mathbf{W}(\mathbf{X})^n$ *of complete preorders, is a complete preorder that solves:*

$$\min_{\mathcal{R} \in \mathbf{W}(\mathbf{X})} \mathcal{C}_{\Sigma, \mathcal{P}}(\mathcal{R}) = \min_{\mathbf{c} \in \mathcal{F}} \sum_{i=1}^{n} d_{\Sigma}(\mathbf{c}_{\mathcal{R}_i}, \mathbf{c}).$$

That is, $\widehat{\mathcal{R}} = \arg \min_{\mathcal{R} \in \mathbf{W}(\mathbf{X})} \mathcal{C}_{\Sigma, \mathcal{P}}(\mathcal{R})$, *being* $\mathbf{c}_{\widehat{\mathcal{R}}}$ *its canonical codification.*

7

It should be noted that the existence of Mahalanobis consensus solution is guaranteed because the feasible set is finite. However, in general its uniqueness can not be assured.

The folowing theorem, proved at [18], establishes an equivalence between rankings computed by the minimization of the MCDF and rankings closest to the mean vector defined by the component-wise averages. This strategy simplify the computation of Mahalanobis consensus solution.

Theorem 1. *Let $\Sigma \in \mathbb{M}_{k \times k}$ be a positive definite matrix and $\mathcal{P} = (\mathcal{R}_1, \ldots, \mathcal{R}_n)$ a profile of complete peorders, where $M_{\mathcal{P}} = (\mathbf{c}_{\mathcal{R}_1}, \ldots, \mathbf{c}_{\mathcal{R}_n}) \in \mathbb{M}_{n \times k}$ denotes its canonical codified profile.*

Then, the following are equivalent statements:

1. $\widehat{\mathcal{R}} \equiv \widehat{\mathbf{c}}$ *is the Mahalanobis consensus solution associate to pair $(\Sigma, \mathcal{P})$:*

$$\widehat{\mathbf{c}} = \arg\min_{\mathbf{c} \in \mathcal{F}} \sum_{i=1}^{n} d_{\Sigma}(\mathbf{c}_{\mathcal{R}_i}, \mathbf{c}). \tag{5}$$

2. $\widehat{\mathcal{R}} \equiv \widehat{\mathbf{c}}$ *is given by:*

$$\widehat{\mathbf{c}} = \arg\min_{\mathbf{c} \in \mathcal{F}} d_{\Sigma}(\overline{\boldsymbol{m}}_{\mathcal{P}}, \mathbf{c}). \tag{6}$$

where $\overline{\boldsymbol{m}}_{\mathcal{P}}$ is the average of the canonical codified profile:

$$\overline{\boldsymbol{m}}_{\mathcal{P}} = (\overline{m}_1, \ldots, \overline{m}_k), \qquad \overline{m}_j = \frac{1}{n} \sum_{i=1}^{n} (\mathbf{c}_{\mathcal{R}_i})_j.$$

the properties of Mahalanobis social consensus solution (anonymity, unanimity, weak neutrality, consistency, compatibility, reciprocity and non-dictatorship) are analyzed and diccussed at [18] .

3. Practical computation of Mahalanobis consensus solution

This section is focused in designing a strategy to compute efficiently Mahalanobis consensus solutions, particularly when the number of alternatives is not small. Note that problem of Eq. (6) can be tackled by a force-brute attack, but this strategy must be rejected even for small k, because the cardinality of set $\mathbf{W}(\mathbf{X})$ grows dramatically in terms of k.

More specifically, the number of preorders $\phi(\cdot)$ of a set $\mathbf{X}$ of cardinality k, is given by:

$$\phi(k) = \sum_{j=0}^{k} j! \cdot S(k,j),$$

where $S(k,j) = \frac{1}{j!} \sum_{l=0}^{j} (-1)^{j-l} \cdot \binom{j}{l} \cdot l^k$ is the Stirling number of the second kind (see [21]).

For that aforementioned purpose, the optimization problem presented in Definition 3 is written as a set of *mixed-integer* constrained quadratic programming problems. Therefore, the feasible set is determined by equality linear constraints, and it is then required the use of continuous and binary auxiliary variables.

Prior to transform the original problem into quadratic problems, the canonical codification included in Subsection 2.2 requires to be adapted. Consequently, some relations and definitions are introduced bellow.

3.1. Canonical codification adaptation

Some elaborations are required in order to relate a preorder $\mathcal{R}$, and its corresponding canonical codification $\mathbf{c}_\mathcal{R}$ via basic algebra. The concept of $\boldsymbol{\alpha}$-index is now introduced:

Definition 4. *Let $\mathcal{R} \in \mathbf{W}(\mathbf{X})$ be a complete preorder on $\mathbf{X}$, and $\mathbf{c}_\mathcal{R} = (c_1, \ldots, c_k)$ its associate canonical codification. Then, the $\boldsymbol{\alpha}_\mathcal{R}$-index of $\mathcal{R}$, is a vector*

$$\boldsymbol{\alpha}_\mathcal{R} = \boldsymbol{\alpha}(\mathbf{c}_\mathcal{R}) = (\alpha_1, \ldots, \alpha_k) \in (\{0, \ldots, k\})^k$$

where the i-th component is the number of occurrences of "i" in its canonical codifications, $\mathbf{c}_\mathcal{R}$:

$$\alpha_i = |\{q \,:\, i = c_q, \, q = 1, \ldots, k\}|$$

Note that two different preorders could have the same $\boldsymbol{\alpha}$-index. The relation between them is established below as well as the characterization of the canonical codified complete preorder based on the $\boldsymbol{\alpha}$-index.

This codification is further illustrated in Example 1 in order to improve its understanding.

Proposition 2. *A vector $\mathbf{c} = (c_1, \ldots, c_k) \in (\{1, \ldots, k\})^k$ is the canonical codified complete preorder $\mathbf{c}_\mathcal{R}$ associated with $\mathcal{R} \in \mathbf{W}(\mathbf{X})$ if and only if the $\boldsymbol{\alpha}_\mathcal{R}$-index of $\mathbf{c}$ verifies:*

$$1 \leq \alpha_k \leq k. \tag{7}$$

$$0 \leq \alpha_j \leq j, \text{ for } j = 1, \ldots, k-1. \tag{8}$$

$$\text{If } \alpha_j = \beta > 0 \Rightarrow \left\{ \begin{array}{l} \alpha_{j-1} = \ldots = \alpha_{j-\beta+1} = 0, \\ \alpha_{j-\beta} \neq 0. \end{array} \right. \tag{9}$$

Proof 2. *Since each vector $\mathbf{c} = (c_1, \ldots, c_k) \in (\{1, \ldots, k\})^k$ induces a unique complete preorder on $\mathbf{X}$, we only have to prove necessity. Conditions $1 \leq \alpha_k$ and $0 \leq \alpha_j$ ($j = 1, \ldots, k-1$) are consequence of Definitions 1 and 4. We now examine:*

$$\alpha_j \leq j, \text{ for } j = 1, \ldots, k.$$

Arguing by contradiction, assume that there exits $\alpha_l = l+1$. That means that $l+1$ alternatives are graded with l in $\mathbf{c}$: $c_{(1)} = c_{(2)} = \ldots = c_{(l+1)} = l$. Thus, there are $l+1$ alternatives that are equally preferred, and in particular any of them, e.g. issue x_i, is least at good at $l+1$ options, which is a contradiction with the value of $c_{(i)} = l < l+1$ (see Definition 1).

To conclude we have to prove Eq. (9). Let $\alpha_j = \beta > 0$, this implies that there are β indifferent alternatives graded with j in $\mathbf{c}$. Therefore, there exits a chain:

$$x_{j_1} \sim x_{j_2} \sim \ldots \sim x_{j_\beta} \succ x_{i_1} \succeq x_{i_2} \succeq \ldots \succeq x_{i_\gamma}$$

that yields

- *$\beta + \gamma = j$, because the alternatives x_{j_l}, $l \in \{1, \ldots, \beta\}$ are graded with j in $\mathbf{c}$.*

- *The alternative x_{i_1}, is at least as good as $j - \beta$ options, that implies $c_{i_1} = j - \beta$. Thus, $\alpha_{j-\beta} > 0$.*

- *In view of there are not elements that are at least as good as l alternatives for $j - 1 \leq l \leq j - \beta + 1$, we deduce $\alpha_{j-1} = \ldots = \alpha_{j-\beta+1} = 0$.*

$\square$

Remark 2. *A simple consequence of Definition 4 is that the $\boldsymbol{\alpha}_{\mathcal{R}}$-index of any canonical codified complete preorder $\mathbf{c}_{\mathcal{R}}$ satisfies :*

$$\sum_{i=1}^{k} \alpha_i = k.$$

Now and in order to improve understanding of the notation, definitions and results, the following illustrative example is introduced.

Example 1. Let $\mathbf{X} = \{x_1, x_2, x_3\}$ be a set of alternatives ($k = 3$). All the possible preorders, $\mathcal{R}$, for these alternatives are summarized in the first column of Table 1. The second column includes the canonical codification of the aforementioned preorders, $\mathbf{c}_{\mathcal{R}}$, according to Definition 1. Their corresponding ordered vectors, $\uparrow\mathbf{c}_{\mathcal{R}}$, are showed in the third column of the table. Their associated $\boldsymbol{\alpha}_{\mathcal{R}}$-indices are summarized in the forth column (Definition 4).

$\mathcal{R}$	$\mathbf{c}_{\mathcal{R}}$	$\uparrow\mathbf{c}_{\mathcal{R}}$	$\boldsymbol{\alpha}_{\mathcal{R}}$
$x_1 \sim x_2 \sim x_3$	$(3,3,3)$	$(3,3,3)$	$(0,0,3)$
$x_1 \sim x_2 \succ x_3$	$(3,3,1)$		
$x_1 \sim x_3 \succ x_2$	$(3,1,3)$	$(1,3,3)$	$(1,0,2)$
$x_2 \sim x_3 \succ x_1$	$(1,3,3)$		
$x_1 \succ x_2 \sim x_3$	$(3,2,2)$		
$x_2 \succ x_1 \sim x_3$	$(2,3,2)$	$(2,2,3)$	$(0,2,1)$
$x_3 \succ x_1 \sim x_2$	$(2,2,3)$		
$x_1 \succ x_2 \succ x_3$	$(3,2,1)$		
$x_1 \succ x_3 \succ x_2$	$(3,1,2)$		
$x_2 \succ x_1 \succ x_3$	$(2,3,1)$	$(1,2,3)$	$(1,1,1)$
$x_2 \succ x_3 \succ x_1$	$(1,3,2)$		
$x_3 \succ x_1 \succ x_2$	$(2,1,3)$		
$x_3 \succ x_2 \succ x_1$	$(1,2,3)$		

Table 1: Preorders, canonical codification, ordered vectors and $\boldsymbol{\alpha}_{\mathcal{R}}$-indices for $k = 3$.

The $\boldsymbol{\alpha}_{\mathcal{R}}$-index (or increasingly ordered vector $\uparrow\mathbf{c}_{\mathcal{R}}$) provides a partition of $\mathbf{W}(\mathbf{X})$. This will be fundamental in Subsection 3.3. First, an equivalence relation in $\mathbf{W}(\mathbf{X})$ is established.

Definition 5. *Given $\mathcal{R}$, $\mathcal{R}' \in \mathbf{W}(\mathbf{X})$, we state that both complete preorders are related (or belong to the same class), $\mathcal{R} \equiv \mathcal{R}'$ if and only if there exists a permutation π such that $\mathcal{R} = {}^\pi\mathcal{R}'$.*

It is easy to prove that this binary relation is in fact an equivalence relation. In fact, each equivalence class can be identified by its $\alpha_\mathcal{R}$-index. The quotient set is denoted by $\mathbf{W}(\mathbf{X})/\sim$ (set of equivalence classes).

Proposition 3. *Let $\mathcal{R}$, $\mathcal{R}'$ be two complete preorder, and $\mathbf{c}_\mathcal{R}$, $\mathbf{c}_{\mathcal{R}'}$ their corresponding canonical codifications. The following statements are equivalent:*

1. $\mathcal{R} \equiv \mathcal{R}'$.
2. $\uparrow\mathbf{c}_\mathcal{R} = \uparrow\mathbf{c}_{\mathcal{R}'}$
3. $\alpha_\mathcal{R} = \alpha_{\mathcal{R}'}$

Proof 3. *We proceed by steps.*

$\boxed{1 \Rightarrow 2}$ *Since $\mathcal{R}$ and $\mathcal{R}'$ belong to same class, there exits a permutation π such that:*

$$\mathcal{R}' = {}^\pi\mathcal{R} \quad (x_i \, {}^\pi\mathcal{R} \, x_j \Leftrightarrow x_{\pi^{-1}(i)} \, \mathcal{R} \, x_{\pi^{-1}(j)}),$$

thus $(\mathbf{c}_\mathcal{R})_i = (\mathbf{c}_{\mathcal{R}'})_{\pi(i)}$, and them both preorders have the same ordered vector, i.e., $\uparrow\mathbf{c}_\mathcal{R} = \uparrow\mathbf{c}_{\mathcal{R}'}$.

$\boxed{2 \Rightarrow 1}$ *Let π and π' be two permutations that map $\mathbf{c}_\mathcal{R}$ and $\mathbf{c}_{\mathcal{R}'}$ onto $\uparrow\mathbf{c}_\mathcal{R} = \uparrow\mathbf{c}_{\mathcal{R}'}$. Then, $\sigma = (\pi')^{-1} \circ \pi$ satisfies $\mathcal{R}' = {}^\sigma\mathcal{R}$ and therefore $\mathcal{R} \equiv \mathcal{R}'$.*

$\boxed{2 \Leftrightarrow 3}$ *For a given $\mathcal{R} \in \mathbf{W}(\mathbf{X})$, the relation between $\uparrow\mathbf{c}_\mathcal{R}$ and $\alpha_\mathcal{R}$ is one-to-one. Definition 4 implies that exits a bijection between the set of ordered vector and the set of α-indices. A simple consequence is that if two preorders, $\mathcal{R}$ and $\mathcal{R}'$ have the same ordered vector, then they have also the same α-index, and reciprocally.*

$\square$

The main consequence of the previous result is that allows to identify the quotient set, $\mathbf{W}(\mathbf{X})/\sim$, with the set of all feasible $\alpha_\mathcal{R}$-indices, that from now on is denoted by $\mathcal{A}$:

$$
\begin{array}{ccc}
\mathbf{W}(\mathbf{X})/\sim: & \xrightarrow{\ \sim\ } & \mathcal{A} \\
[\mathcal{R}] & \longrightarrow & \alpha_\mathcal{R}
\end{array}
$$

This bijection will be an important key to achieve the goal of this paper. In other words, $\alpha_{\mathcal{R}}$-indices induce a partition of the set of canonical codified complete preorder $\mathcal{F}$

$$\mathcal{F} = \bigsqcup_{\alpha \in \mathcal{A}} \mathcal{F}_\alpha, \qquad \mathcal{F}_\alpha = \{\mathbf{c} \in \mathcal{F} \,:\, \alpha(\mathbf{c}) = \alpha\}, \tag{10}$$

where $\mathcal{F}_\alpha$ groups all canonical codified complete preorder with the same α-index.

The following results establishes the cardinality of the quotient set, $\mathbf{W}(\mathbf{X})/\sim$ (equivalently the cardinality of $\mathcal{A}$).

Proposition 4. *The cardinality of the* $\mathbf{W}(\mathbf{X})/\sim$ *is* 2^{k-1}.

Proof 4. *From Proposition 3, we only have to count the number of ordered vectors* $\mathbf{c} = \uparrow \mathbf{c}$ *satisfying Eqs. (1)-(2). We proceed by steps to fill the coefficient c_j, $j = k, \ldots, 1$, and so build all possible $\mathbf{c}$. In the first step the value of c_k is fixed to k. Now, filled c_{j+1}, in the j-th step we only have two choices that depends of the value $d_{j+1} = 0, 1$. Since we have $k - 1$ effective steps and two choice in each step, we then conclude that the cardinality of $\mathbf{W}(\mathbf{X})/\sim$ is 2^{k-1}.* □

3.2. The establishment of the feasible set

One of the main obstacles in the resolution of the optimization problems presented in this contribution is the setting of the feasible set for a "not small" set of alternatives. A recursive algorithm, Algorithm 1, is proposed to compute explicitly $\mathcal{A}$ for a given k. This novel algorithm benefits from Proposition 2 and it is called *Recursive α-index algorithm* (RAI).

The algorithm receives as unique argument the number of alternatives, k. In line 2, algorithm distinguishes between the trivial case ($k = 1$, where $\mathcal{A} = (1)$), and the generic one ($k > 1$). Lines 4-7 build $\mathcal{A}$, when $k > 1$. In line 4, the unique α-index with $\alpha_k = k$ is generated (this index corresponds to k indifferent alternatives). Then, the for-loop in lines 5-7, builds α-indices with $\alpha_k \in \{k - 1, \ldots, 1\}$. Line 6 implements the main step of the algorithm. The α-indices that have their k-*th* component equals to $k - i$ are generated. According to Eq. (9), these indices take value $\alpha_j = 0$ for $j = i + 1, \ldots, k - 1$, and their first i-*th* components correspond to the α-indices associated to a set of cardinality i. Note that the algorithm is called recursively.

Algorithm 1 : **RAI** (Recursive α-index algorithm)

1: **procedure** $\mathcal{A} \leftarrow$ GETALPHASET(k)
2:    **if** $k = 1$ **then return** $\mathcal{A} \leftarrow \{(1)\}$
3:    **else**
4:       $\mathcal{A}_k \leftarrow (0, \ldots, 0, k) \boxed{\in \mathbb{M}_{1 \times k}}$
5:       **for** $i \leftarrow 1$ to $k - 1$ **do**
6:       $\mathcal{A}_{k-i} \leftarrow \left(\text{getAlphaSet}(i) \left| \begin{array}{ccc} & \vdots & \\ \ldots & 0 & \ldots \\ & \vdots & \end{array} \right| \begin{array}{c} k - i \\ \ldots \\ k - i \end{array} \right) \boxed{\in \mathbb{M}_{2^{i-1} \times (i + (k-i-1)+1)}}$
7:       **end for**
8:       **return** $\mathcal{A} \leftarrow (\mathcal{A}_k; \mathcal{A}_{k-1}; \ldots; \mathcal{A}_1) \boxed{\in \mathbb{M}_{2^{k-1} \times k}}$
9:    **end if**

3.3. Algorithm to compute the Mahalanobis consensus solution

Our procedure depends on two fundamental keys:

1. The partition of canonical codified complete peorders, $\mathcal{F}(\equiv \mathbf{W}(\mathbf{X}))$ as disjoint union, Eq. (10).
2. The canonical codified complete preorders in $\mathcal{F}_\alpha$ (preorders with the same α-index) can be characterized via linear constrains.

Taking into account these issues, the optimization problem defined in Theorem 1 Eq. (6) can be rewrited as:

$$\min_{\mathbf{c} \in \mathcal{F}} d_\Sigma(\overline{\mathbf{m}}_\mathcal{P}, \mathbf{c}) \quad \equiv \quad \min_{\alpha \in \mathcal{A}} \left(\min_{\mathbf{c} \in \mathcal{F}_\alpha} d_\Sigma(\overline{\mathbf{m}}_\mathcal{P}, \mathbf{c}) \right)$$

where $\mathcal{F}_\alpha$ denotes the set of canonical codified complete preorders which has $\alpha \in \mathcal{A}$ as α-index (Definition 4).

Focusing on the optimization problem on $\mathcal{F}_\alpha$ that can be formulated in *standard* way. The issue is to write $\mathbf{c} \in \mathcal{F}_\alpha$ using basic algebra.

Let $\{b_{ij}\}_{1 \leq i,j \leq k}$ be $k \times k$ binary variables ($b_{ij} = 0, 1$). The following conditions characterize $\mathbf{c} \in \mathcal{F}_\alpha$:

$$c_i = \sum_{j=1}^{k} j b_{ij}, \qquad\qquad \forall i = 1, \ldots, k, \qquad\qquad (11)$$

$$\sum_{j=1}^{k} b_{ij} = 1 \qquad\qquad \forall i = 1, \ldots, k, \qquad\qquad (12)$$

$$\sum_{i=1}^{k} b_{ij} = \alpha_i \qquad\qquad \forall j = 1, \ldots, k, \qquad\qquad (13)$$

$$b_{ij} \in \{0, 1\} \qquad\qquad \forall i, j = 1, \ldots, k, \qquad\qquad (14)$$

$$c_i \in \mathbb{R} \qquad\qquad \forall i = 1, \ldots, k. \qquad\qquad (15)$$

Note that, on one hand Eqs. (11), (12) and (14) ensure that the components $\{c_i\}_{i=1}^{k}$ of the codified preorder $\mathbf{c}$ are integer and belong to $\{1, \ldots, k\}$ ($c_i = j \Leftrightarrow b_{ij} = 1$). On the other hand, Eq. (13) provides the number of occurrences i in $\mathbf{c}$ (i.e., the $\boldsymbol{\alpha}$-index of $\mathbf{c}$).

In addition, several remarks should be highlighted:

- Since $\boldsymbol{\alpha} \in \mathcal{A}$ is a *feasible* $\boldsymbol{\alpha}$-index, any $\mathbf{c} = (c_1, \ldots, c_k)$, satisfying Eqs. from (11) to (15), is a canonical codified complete preorder.

- Each canonical codified complete preorder $\mathbf{c} \in \mathcal{F}$ is solution of Eqs. from (11) to (15). It is enough to choose $b_{ij} = \begin{cases} 1 & \text{if } c_i = j, \\ 0 & \text{otherwise.} \end{cases}$

- Equations (11), (12), (13), (14) and (15) involve $3k$ equality linear constrains and $k(k+1)$ variables, of which k^2 are binary and k are continuous and unbounded.

15

Collecting these results, the optimization problem on $\mathcal{F}_\alpha$ can be represented in matrix form as:

$$\min \quad (\overline{\mathbf{m}}_\mathcal{P} - \mathbf{c})\, \Sigma^{-1}\, (\overline{\mathbf{m}}_\mathcal{P} - \mathbf{c})^t$$

$$\text{s.a.}$$

$$\begin{pmatrix} Id & B_1 & \dots & B_k \\ 0 & C_1 & \dots & C_k \\ 0 & Id & \dots & Id \end{pmatrix} \begin{pmatrix} \mathbf{c}^t \\ \mathbf{b}_1^t \\ \vdots \\ \mathbf{b}_k^t \end{pmatrix} = \begin{pmatrix} \mathbf{0}^t \\ \mathbf{1}^t \\ \boldsymbol{\alpha}^t \end{pmatrix} \tag{16}$$

where $Id, \{B_i\}_{i=1}^k, \{C_i\}_{i=1}^k \in \mathbb{M}_{k\times k}$ and $\mathbf{0}, \mathbf{1}, \{\mathbf{b}_i\}_{i=1}^k \in \mathbb{R}^k$ are given by:

$$B_i = \begin{pmatrix} 0 & \dots & \dots & 0 \\ \dots & \dots & \dots & \dots \\ -1 & -2 & \dots & -k \\ \dots & \dots & \dots & \dots \\ 0 & \dots & \dots & 0 \end{pmatrix} \qquad C_i = \begin{pmatrix} 0 & \dots & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 1 & 1 & \dots & 1 \\ \dots & \dots & \dots & \dots \\ 0 & \dots & \dots & 0 \end{pmatrix}$$

being gray row, the i-*th* row;

$$Id = \begin{pmatrix} 1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & 1 \end{pmatrix}$$

$$\mathbf{0} = (0, \dots, 0) \qquad \mathbf{1} = (1, \dots, 1) \qquad \mathbf{b}_i = (b_{i1}, \dots, b_{ik}).$$

To finish this section, Algorithm 2 to compute Mahalanobis consensus solutions, the *Mahalanobis consensus solution algorithm* (MCSA), is presented.

This algorithm requires as input parameters: the Σ matrix associated to the Mahalanobis consensus measure, and the codified profile $M_\mathcal{P}$. Initially, some preliminar calculations are done: the average of the canonical codified profile $\overline{\mathbf{m}}_\mathcal{P}$ (line 2), the cardinaltiy of the alternatives set k (line 3), and the set of $\boldsymbol{\alpha}$-indices $\mathcal{A}$ (line 4). Then, the consensus solution value and its associated Mahalanobis distance are initialized. The for-loop of lines 7 to 13 solves the optimization problem on the set $\mathcal{F}_\alpha$ with $\boldsymbol{\alpha} \in \mathcal{A}$ ($|\mathcal{A}| = 2^{k-1}$). In each step, the quadratic optimization problem of Eq. (16) is assembled (line 8), and solved (line 9). If the former

---
Algorithm 2 : MCSA

1: **procedure** $\quad \widehat{\mathbf{c}} \leftarrow$ COMPUTEMSCM(Σ, $M_{\mathcal{P}}$)
2: $\quad \overline{\mathbf{m}}_{\mathcal{P}} \leftarrow$ getAverage($M_{\mathcal{P}}$)
3: $\quad k \leftarrow$ getNoAlternatives($M_{\mathcal{P}}$)
4: $\quad \mathcal{A} \leftarrow$ getAlphaSet(k)
5: $\quad$ fmin $\leftarrow$ Inf
6: $\quad \widehat{\mathbf{c}} \leftarrow \mathbf{0}$
7: $\quad$ **for** $i \leftarrow 1$ to 2^{k-1} **do**
8: $\quad\quad$ OP = assemble(Σ, $\overline{\mathbf{m}}$, $\boldsymbol{\alpha}_i$)
9: $\quad\quad \{\mathbf{c}$, fobj$\} \leftarrow$ solve(OP)
10: $\quad\quad$ **if** fobj $<$ fmin **then**
11: $\quad\quad\quad \widehat{\mathbf{c}} \leftarrow \mathbf{c}$
12: $\quad\quad$ **end if**
13: $\quad$ **end for**
14: $\quad$ **return** $\widehat{\mathbf{c}}$
---

consensus solution provides a better agree (the Mahalanobis distance is the smallest one) then, the consensus solution is updated. This loop can be parallelized. Finally, the best consensus solution is returned.

About the solution of the quadratic programing problem, note that Algorithm 2 (MCSA) has been implemented in Matlab [30]. The optimization problem was addressed with Gurobi software [31], using a mixed-integer quadratic programming solver (MIQP). It is based on branch-and-bound algorithm, and combines: pre-solving, cutting planes, heuristics and branching.

By way of illustration of the procedure proposed, the next example is presented.

Example 2. Let $\boldsymbol{\alpha} = (0, 2, 1)$ be an $\boldsymbol{\alpha}$-index for $k = 3$. The set $\mathcal{F}_{\boldsymbol{\alpha}}$ is characterized by the following constrains:

$$
\begin{aligned}
c_1 &= b_{11} + 2b_{12} + 3b_{13} \\
c_2 &= b_{21} + 2b_{22} + 3b_{23} \\
c_3 &= b_{31} + 2b_{32} + 3b_{33} \\
1 &= b_{11} + b_{12} + b_{13} \\
1 &= b_{21} + b_{22} + b_{23} \\
1 &= b_{31} + b_{32} + b_{33} \\
0 &= b_{11} + b_{21} + b_{31} \\
2 &= b_{12} + b_{22} + b_{32} \\
1 &= b_{13} + b_{23} + b_{33}
\end{aligned}
$$

$$
\begin{aligned}
b_{i,j} &= \{0,1\} & i,j &= 1,2,3 \\
c_i &\in \mathbb{R} & i &= 1,2,3
\end{aligned}
$$

Expressing the constrains in matrix structure:

$$
\begin{pmatrix}
1 & 0 & 0 & -1 & -2 & -3 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 & 0 & 1 & -2 & -3 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & -2 & -3 \\
0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 \\
0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1
\end{pmatrix}
\begin{pmatrix}
c_1 \\ c_2 \\ c_3 \\ b_{11} \\ b_{12} \\ b_{13} \\ b_{21} \\ b_{22} \\ b_{23} \\ b_{31} \\ b_{32} \\ b_{33}
\end{pmatrix}
=
\begin{pmatrix}
0 \\ 0 \\ 0 \\ 1 \\ 1 \\ 1 \\ 0 \\ 2 \\ 1
\end{pmatrix}
$$

A straightforward computation reveals that the solutions (codified complete preorders) are given by:

$$
\{(2,2,3),(2,3,2),(2,2,3)\}
$$

that correspond with the third row-group of Table 1.

For the purpose of showing and justifying the computation efficiency of the RAI-Algorithm and MSCA-Algorithm, some arguments are now introduced.

3.4. About the complexity of the RAI and MSCA Algorithm

The computation of Mahalanobis consensus solutions can be tackled by a force-brute attack but this strategy involves some technical difficulties:

a) First off it is necessary to emphasize that the use of the force-brute computation requires an exhaustive inspection of all preorders to define the feasible set $\mathcal{F}$. Due to the fact that the number of preorder grows dramatically in terms of the number of alternatives, the establishment of the feasible set in a reasonable time frame for a not small number of alternatives is a daunting task. Table 2 shows that the number of preorders to be defined for $k = 20$ is 2677687796244384203115.

b) Secondly and to the best of our knowledge, there is not a parameterization of the set of preorders in the specialized literature. Therefore, there not exists an algorithm to explicit generate the feasible set $\mathcal{F}$.

c) Lastly, it could be possible to produce the feasible set $\mathcal{F}$ selecting the set of preorders from the k-tuples set although this process is high demanding.

In the light of the above, the RAI and MSCA Algorithm and force-brute attack strategy are compared. To this end and to measure the computational cost in term of the number of alternatives, a complexity rate is proposed. This ratio is defined like the quotient of the computational cost to obtain the Mahalanobis consensus solution between two consecutive values of k. Thereby, the complexity rate for the force-brute attack strategy is defined by

$$C_{\text{BF}}(k) = \frac{|\mathcal{F}_k|}{|\mathcal{F}_{k-1}|},$$

where $|\mathcal{F}_k|$ and $|\mathcal{F}_{k-1}|$ is de cardinality of the feasible set for k and $k-1$ alternatives, respectively.

For the MSCA-Algorithm, the complexity rate is defined by

$$C_{\text{MSCA}}(k) = \frac{|\mathcal{A}_k| \cdot t_k}{|\mathcal{A}_{k-1}| \cdot t_{k-1}},$$

where $|\mathcal{A}_k|$ and $|\mathcal{A}_{k-1}|$ denote the cardinality of the set of all feasible $\alpha_{\mathcal{R}}$-indices for k and $k-1$ alternatives, respectively and t_k denotes the average time to solve the mixed-integer quadratic programming problem [8].

---

[8]To solve the mixed-integer quadratic programming problem, the associated matrix Σ and the average canonical profile $\overline{\mathbf{m}}_{\mathcal{P}}$ were randomly generated. Moreover, the α-index was fixed

The results of the comparative study are shown in Table 2 and Figure 1. Note that the complexity rate for MCSA-Algorithm remains between $[2, 4]$ for $k \leq 20$. However the rate for the brute-force strategy grows linearly. That reveals that our proposal is efficient and competitive.

The reported computation time in Table 2 corresponds to an average over 100 problems on a Dell Precision T7500. Data management and analysis was performed using a equipped with two Intel Xeon X5550 v2 processors - 4 cores, each working at 2.66 GHz - and 24 GB RAM. Matlab and Gurobi optimizer were combined to solve the problems. In addition, since RAI-Algorithm allows to generate the $\mathcal{A}$ set, the computation of Mahalanobis consensus solution can be trivially parallelized. Therefore the computations time on Table 2 can be reduced using the current generation of calculation platforms.

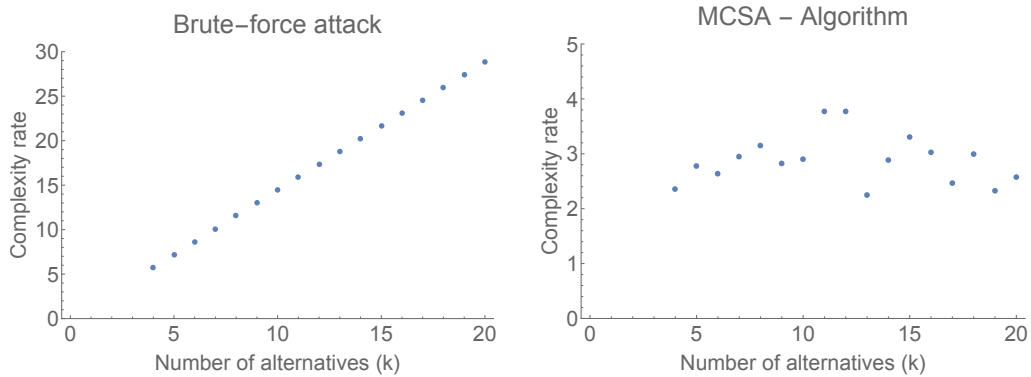Figure 1: Complexity rate for a brute-force attack strategy and for the MSCA-Algorithm versus the number of alternatives (k).

In the coming section the novel procedures and algorithms exposed in this work are put in practice in a real case.

___

to $(1, \ldots, 1)$ due to the fact that the associated feasible set is the biggest and heuristically, the mixed-integer quadratic programming problem seems to be the hardest one.

	Brute-force attack strategy		MSCA-Algorithm		
k	$\|\mathcal{F}_k\|$	$C_{\mathrm{BF}}(k)$	$\|\mathcal{A}_k\|$	t_k	$C_{\mathrm{MSCA}}(k)$
3	13	-	4	0.0044	-
4	75	5.7	8	0.0052	2.4
5	541	7.2	16	0.0072	2.8
6	4683	8.6	32	0.0095	2.6
7	47293	10.1	64	0.014	2.9
8	545835	11.5	128	0.022	3.1
9	7087261	13.0	256	0.031	2.8
10	102247563	14.4	512	0.035	2.9
11	1622632573	15.9	1024	0.055	3.7
12	28091567595	17.3	2048	0.16	3.7
13	526858348381	18.7	4096	0.18	2.3
14	10641342970443	20.2	8192	0.26	2.8
15	230283190977853	21.6	16384	0.43	3.3
16	5315654681981355	23.1	32768	0.65	3.0
17	130370767029135901	24.5	65536	0.80	2.5
18	3385534663256845323	26.0	131072	1.2	3.0
19	92801587319328411133	27.4	262144	1.4	2.3
20	2677687796244384203115	28.8	524288	1.8	2.6

Table 2: Main features of the two proposed Mahalanobis consensus solution strategies: brute-force attack and MSCA-Algorithm .

4. A real case: The Spanish social consensus family budget

The experimental work presented here provides one of the first investigations into how to compute the family budgets that more consensus conveys among families and then, to establish household consumption behaviour[9]. Particularly, the focus of the work is on Spanish Economy and social consensus family budgets are computed based on Spanish data on 2016 household budgets. As mentioned in Introduction, the knowledge of the Spanish household consumption behaviour allows companies and Governments to improve their marketing and public policies. It is hoped that this real study will contribute to a deeper understanding of the methodologies proposed in the previous Sections.

The remaining part of the section proceeds as follows. Firstly, a a brief overview of the framework and the traditional approach used in Spain to determine the family budgets is introduced. Secondly, the data provided by the Spanish

---

[9]In this particular case, families play the role of agents or consumers.

National Statistics Institute for the households consumption are manipulated to obtain families' preferences. Subsequently, the canonical codification of the families preferences are computed as well as averages group expenditure. Thereupon, different social consensus family budgets are calculated for different Σ matrices to promote the debate. Finally, the aforementioned results are analysed and discussed.

4.1. The Spanish framework and approach

The Spanish *National Statistics Institute* performs the *Household Budget Survey* (HBS) since 1958. This survey is one of the oldest in Spain and its main goal is to obtain information on the nature and destination of consumption cost, as well as on various characteristics relating to the conditions of household life.

The survey was last reformed in 2006, it is nowadays annual, includes about 24000 dwellings in its sample and provides essential information on the estimates on households consumption expenditure and on updating the *Consumer Price Index* (CPI) weightings.

The HBS classifies expenses using the COICOP (national adaptation of the international classification used by *Eurostat* for budget surveys) and it structures them as follows:

- Group 1: Food and non-alcoholic beverages

- Group 2: Alcoholic beverages, tobacco and narcotics

- Group 3: Clothing and footwear

- Group 4: Housing, water, electricity, gas and other fuels

- Group 5: Furniture, household equipment and ordinary expenses for the maintenance of the dwelling

- Group 6: Health

- Group 7: Transport

- Group 8: Communication

- Group 9: Leisure, performances and culture

- Group 10: Education

- Group 11: Restaurants, coffee and hotels

- Group 12: Miscellaneous goods and services

The National Statistics Institute to estimate the characteristics of households uses the following estimator [32]:

$$\widehat{X}_A = \sum_{h \in A} \frac{P_h \cdot T}{\sum_{g \in h} R_g \cdot \sum_{i \in g}[c_{hi} \cdot \sum_{j \in i} p_{hij}]} \cdot \left[\sum_{g \in h} R_g \cdot \sum_{i \in g}[c_{hi} \cdot \sum_{j \in i} x_{hij}] \right]$$

being:

- $\widehat{X}_A$: Estimate of the total annual household expenditure on a good or service X in geographical area A.

- P_h: Population projection of stratum h, referring to half of the surveying year.

- T: Temporary elevation factor. This factor depends on the reference period of the good or service X.

- c_{hi}: Update coefficient. It is a value that depends on the selection and represents its growth from the moment of sample selection until the year of the data in the survey. For the year 2016 this coefficient is equal to 1.

- p_{hij}: Population formed by the household members in household j of the sample, section i, stratum h.

- x_{hij}: Value of the expenditure on the good or service X in household j of the sample, section i, stratum h.

- R_g: Non-response correction factor in group g. It is obtained as the quotient between the total number of households of the theoretical sample of this group and the household of the effective sample.

The previous $\widehat{X}_A$ estimator can also be expressed in the next simplified version by *calibration* approach:

$$\hat{X}_A = \sum_{k \in A} w_k x_k$$

where k ranges over all of the effective sample units in geographical area A.

To obtained this final estimator, it is necessary to find a new weight d_k enables $\hat{X}_A$ to verify:

i. The estimate based on the sample of a specific feature must match the value of this feature on the population (this value is obtained through an external source).

ii. Fixed a distance function[10], the distance between d_k and w_k should be minimal.

In recent years, there has been an increasing interest in analysing the aforementioned traditional approach and the main challenges faced by many researchers are (see [33] and [34]):

- Selection of the appropriate equivalence scale to adapt the household budgets for different sizes and types.

- Consideration of interrelationships among expenditure groups.

- Establishment of index/weights for geographical differences in prices.

- Updating budgets over time.

Although there are several approaches in the specialized literature to overcome these drawbacks, there is no consensus on it. Therefore, in the pages that follow, this research shows how to implement the new methodology and algorithms raised above that improving such drawbacks.

4.2. Our proposal

Following the group decision making problem framework presented in Section 2, the set of agents considered is composed of the households included in the sample selected by the Spanish National Statistics Institute in 2016: 22130 household, then $\mathbf{N} = \{1, \ldots, 22130\}$. Additionally, the set of alternatives is formed by the twelve expenditure groups[11] , $\mathbf{X} = \{x_1, x_2, \ldots, x_{12}\}$, $k = 12$.

---

[10]A truncated linear distance function is traditionally applied and the computations are performed by the CALMAR software (CALibration of MARgins). This specific software is used by many National Statistics Offices on the world like UK, Ireland, France, and so on.

[11]The structure of the expenditure groups as well as the sample distribution are preserved in time.

By means of the HBS, the data of the households consumption expenditure is collected and classified according the expenditure groups[12]. For the purpose of obtaining the households preferences on the expenditure groups, the consumption for each household is aggregated for each expenditure group. After aggregating, the results are ordered to grade the groups, obtained a complete preorder $\mathcal{R}_i \in \mathbf{W(X)}$ on $\mathbf{X}$ for each household $i = \{1, \ldots, n = 22\,130\}$. These complete preorders generate a particular profile:

$$\mathcal{P} = (\mathcal{R}_1, \ldots, \mathcal{R}_n) \in \mathbf{W(X)} \times \ldots \times \mathbf{W(X)} = \mathbf{W(X)}^{n=22130}$$

where $\mathcal{R}_i$ represents the preferences of the household i on the twelve expenditure groups for each $i = 1, \ldots, 22\,130$.

Applying Definition 1 to each complete preorder the codified profile for $\mathcal{P}$, $\mathcal{M}_{\mathcal{P}} \in \mathbb{M}_{22\,130 \times 12}$ is obtained.

By way of illustration of this first procedure, Table 3 shows the expenditure of three families from the sample in the groups of goods. Taking into account the data from Table 3, the household preferences are:

- Household 1:

  $\mathcal{R}_1 : x_2 \prec x_9 \prec x_5 \prec x_6 \prec x_8 \prec x_7 \prec x_{10} \prec x_3 \prec x_{12} \prec x_4 \prec x_{11} \prec x_1$

- Household 2:

  $\mathcal{R}_2 : x_3 \sim x_6 \sim x_7 \sim x_{10} \prec x_5 \prec x_8 \prec x_2 \prec x_9 \prec x_{12} \prec x_1 \prec x_4 \prec x_{11}$

- Household 3:

  $\mathcal{R}_3 : x_6 \sim x_{10} \prec x_9 \prec x_5 \prec x_2 \prec x_8 \prec x_{12} \prec x_1 \prec x_7 \prec x_3 \prec x_4 \prec x_{11}$

---

[12]All data available through the website: http://www.ine.es/dyngs/INEbase/en/operacion.htm?c=Estadistica_C&cid=1254736176806&menu=resultados&idp=1254735976608

	Expenditure group*					
Household	x_1	x_2	x_3	x_4	x_5	x_6
1	11.66	0.08	2.31	7.07	0.81	0.93
2	3.54	2.32	0.00	5.36	0.24	0.00
3	2.48	0.12	5.11	9.75	0.12	0.00

	Expenditure group*					
Household	x_7	x_8	x_9	x_{10}	x_{11}	x_{12}
1	1.61	1.54	0.65	1.75	7.82	5.14
2	0.00	0.59	2.71	0.00	7.40	2.84
3	3.04	1.18	0.11	0.00	15.14	1.92

*: data in thousands of euro.

Table 3: Expenditure for group from original data.

Once the household preferences have been established, the corresponding pre-orders, $\mathcal{R}_1, \mathcal{R}_2$ and $\mathcal{R}_3$ are codified following Definition 1. Table 4 shows these codified complete preorders:

	$\mathbf{c}_{\mathcal{R}_i}$					
Household	c_1	c_2	c_3	c_4	c_5	c_6
1	12	1	8	10	3	4
2	10	7	4	11	5	4
3	8	5	10	11	4	2

	$\mathbf{c}_{\mathcal{R}_i}$					
Household	c_7	c_8	c_9	c_{10}	c_{11}	c_{12}
1	6	5	2	7	11	9
2	4	6	8	4	12	9
3	9	6	3	2	12	7

Table 4: Codified complete preorders $\mathcal{R}_i$ for $i = 1, 2, 3$

Returning to the general issue, to obtain the social consensus family budgets by means of the optimization problem presented in Eq. (6), it is necessary:

1. To compute the average of the canonical codified profile:

$$\overline{m}_{\mathcal{P}} = (\overline{m}_1, \ldots, \overline{m}_{12})$$

   The elements of $\overline{m}_{\mathcal{P}}$ vector are showed in Table 5.

2. To fix a Σ matrix. This matrix should convey the specific characteristics of the problem at hand. The role of the Σ matrix allows to take into account the variances of the expenditure groups and the covariances among them. In this contribution, several different Σ matrices were considered to promote the discussion and to capture the interrelations among the twelve expenditure groups:

   - $\Sigma_1 = Id$ matrix. The simplest case. The Σ matrix is the identity matrix, then all expenditure groups are equally treated.

   - Σ_2 matrix. This Σ matrix accounts the case where the expenditure groups are considered differently by means of a diagonal matrix. Specifically, Σ_2 includes like diagonal elements the variances of the expenditure groups from the original data. Table 6 includes this particular matrix.

   - Σ_3 matrix. A natural choice could be to consider the Σ matrix like the empirical variance-covariance matrix computed directly from the original data (without codification). This matrix involves not only that all expenditure groups are not equally treated but they are also correlated. Figure 2 presents a graphical interpretation of such relations[13]. Table 7 shows Σ_3 and its corresponding correlation matrix is in Table 8.

   - Σ_4 matrix. Similar to Σ_2 matrix, Σ_4 is a diagonal matrix which includes like diagonal elements the variances of the expenditure groups but from the codified profile as shows Table 9.

   - Σ_5 matrix. Since the objective function in the optimization problem uses the Mahalanobis distance among canonical codified preorders, it is reasonable to propose like Σ matrix the statistical variance-covarince matrix computed from the codified profile. This matrix Σ_5 and its corresponding correlation matrix are in Table 10 and Table 11, respectively. Figure 3 shows a graphical interpretation of the respective

correlation matrix[13].

- Σ_6 matrix. This special Σ matrix has been designed to capture the correlations among the group of expenditures from original data and to adapt them to the scale of the codified profile. To this aim, Σ_6 is defined by

$$\Sigma_6 = D \cdot R \cdot D$$

where R is the correlation matrix computed from the twelve expenditure groups (Table 8) and D the diagonal matrix which diagonal elements are the square roots of the variances of the codified profile (Table 10). The Σ_6 matrix is showed in Table 12.

3. To establish the feasible set. As mentioned in Subsection 2.3, the feasible set in this problem is finite, although for $k = 12$, the number of complete preorders to consider is $\phi(12) = 28\,091\,567\,595 \approx 3e + 10$ (see [35]). To manage such a number of preorders, Algorithm 1 is put in practice, to compute the α-indices set $\mathcal{A}_{12}$ ($|\mathcal{A}_{12}| = 2^{11} = 2048$).

4. To solve the quadratic optimization problems. For each $\alpha \in \mathcal{A}_{12}$ we solve:

$$\min_{\mathbf{c} \in \mathcal{F}_\alpha} d_{\Sigma_m}(\overline{\mathbf{m}}_{\mathcal{P}}, \mathbf{c})$$

or in matrix form:

$$\min \quad (\overline{\mathbf{m}}_{\mathcal{P}} - \mathbf{c})\, \Sigma_m^{-1}\, (\overline{\mathbf{m}}_{\mathcal{P}} - \mathbf{c})^t$$

s.a.

$$\begin{pmatrix} Id & B_1 & \ldots & B_{12} \\ 0 & C_1 & \ldots & C_{12} \\ 0 & Id & \ldots & Id \end{pmatrix} \begin{pmatrix} \mathbf{c}^t \\ \mathbf{b}_1^t \\ \vdots \\ \mathbf{b}_{12}^t \end{pmatrix} = \begin{pmatrix} \mathbf{0}^t \\ \mathbf{1}^t \\ \boldsymbol{\alpha}^t \end{pmatrix}$$

where $Id, \{B_i\}_{i=1}^{12}, \{C_i\}_{i=1}^{12} \in \mathbb{M}_{12 \times 12}$ and $\mathbf{0}, \mathbf{1}, \{\mathbf{b}_i\}_{i=1}^{12} \in \mathbb{R}^{12}$.
Finally, among 2048 solutions associated to each $\mathcal{F}_\alpha$, we select the codified preorder $\mathbf{c}$ that produces the best agreement.

---

[13]Ellipses in Figures 2 and 3 symbolize the correlation between two expenditure groups. Ellipses slant upward (resp., downward) show a positive (resp., negative) correlation. Moreover, the tighter the ellipse the stronger correlation.

Algorithm 2 reveals that the social consensus family budgets for the six Σ matrices are given by:

Σ_1: $x_4 \succ x_1 \succ x_{12} \succ x_7 \succ x_9 \succ x_{11} \succ x_3 \sim x_5 \succ x_2 \sim x_6 \sim x_8 \sim x_{10}$.

Σ_2: $x_4 \succ x_1 \succ x_7 \succ x_{11} \succ x_{12} \succ x_9 \succ x_3 \sim x_8 \succ x_2 \sim x_5 \sim x_6 \sim x_{10}$.

Σ_3: $x_4 \succ x_1 \succ x_7 \succ x_{11} \succ x_{12} \succ x_9 \succ x_3 \sim x_8 \succ x_2 \sim x_5 \sim x_6 \sim x_{10}$.

Σ_4: $x_4 \succ x_1 \succ x_7 \succ x_{11} \succ x_{12} \succ x_9 \succ x_3 \sim x_5 \sim x_8 \succ x_2 \sim x_6 \sim x_{10}$.

Σ_5: $x_4 \succ x_7 \succ x_1 \succ x_{11} \succ x_{12} \succ x_9 \succ x_8 \succ x_3 \succ x_5 \sim x_2 \sim x_6 \sim x_{10}$.

Σ_6: $x_4 \succ x_1 \succ x_7 \succ x_{12} \succ x_{11} \succ x_9 \succ x_8 \sim x_5 \sim x_3 \succ x_2 \sim x_6 \sim x_{10}$.

Moreover the group position for different Σ as well as the values of the objective function are displayed in Table 13 and showed in Figures 4 and 5.

Note that the calculation time employed to obtain the solutions aforementioned on a Dell Precision T7500 (equipped with two Intel Xeon X5550 v2 processors – 4 cores, each working at 2.66 GHz – and 24 GB RAM) is on the order of 80 seconds. A brute-force procedure should calculate the distance between preorders in a time of approximately $3.5e + 8$ preorders per second to be competitive with the procedure implemented by Algorithms 1 and 2.

4.3. Discussion

The present case of study was designed to analyse consumer behaviour from a non-standard point of view. The results of this study indicate that different social consensus family budgets are obtained if the expenditure groups are not equally treated as well as if the interrelations among them are taken into account.

In reviewing the literature, no data was found on the inclusion of such relationships in the traditional computation on the family budgets. This fact may disturb the results of the analyses of the consumer behaviour obtained and also the computation of the consumer price index (CPI).

Closer inspection of the outcomes obtained in Subsection 4.2 and Figures 4 and 5, reveals that x_4 the expenditure group 4 (Housing, water, electricity, gas and other fuels), appears on the top of the social consensus family budget for each Σ matrix.

The solutions obtained for Σ_2 and Σ_3 coincide among them although Σ_3 takes into account the interrelation among the groups. This fact could be due to the

magnitude of the original data, that is, the large figures included in the Σ matrix skew the preorders included in the codified profile and then the results. This problem is solved by including the special matrix, Σ_6. This particular matrix has been built to adapt the correlations among the original data from expenditure groups to the scale of the codified profile.

For Σ_4 and Σ_5 the solutions obtained are similar but not equal. To be precise, for Σ_5 the social solution presents less expenditure groups indifferent among them, for instance: $x_8 \succ x_3 \succ x_5$ for Σ_5 against $x_8 \sim x_3 \sim x_5$ for Σ_3, due to the incorporation of the interrelation among them even if their correlations are relatively small (see Tables 10, 11 and Figures 3, 4 and 5).

About the change of expenditure group position respect to Σ_i, note that groups x_7 (transport) and x_{12} (services) are the more affected. Whereas the solution for Σ_1 orders x_7 in the 2nd position, the choice Σ_5 sets this group at 4th position. A similar situation is presented for the alternative x_{12} (see Figure 5). It reveals that the correlations among expenditure groups affects the final ranking of alternatives.

Together these results provide important insights into the relevance of considering the relationship among expenditure groups and the data scale in order to compute the consumers' behaviour (goods ranking) that more consensus conveys to the society. Hence, it could conceivably be hypothesised that the Σ_6 matrix provides the best solution:

$$x_4 \succ x_1 \succ x_7 \succ x_{12} \succ x_{11} \succ x_9 \succ x_8 \sim x_5 \sim x_3 \succ x_2 \sim x_6 \sim x_{10}$$

5. Conclusions and futher research work

The present study makes several noteworthy contributions to find the best solution in group decision making problems. Focusing on the problem presented by González-Arteaga, Alcantud and de Andrés Calle in [18], the key strengths of this study is provided a strategy to compute efficiently Mahalanobis consensus solutions when the number of alternatives is not small. Particularly, this study writes the original optimization problem like a set of mixed-integer quadratic programming problems where the feasible set is determined by equality linear constraints by continuous and binary variables. Regarding this general goal, three mains contributions are proposed: a new codification procedure based on the new concept of α-index and its characterization; a recursive algorithm, the recursive

α-index algorithm (RAI), to establish the feasible set and an algorithm to compute Mahalanobis consensus solutions, the Mahalanobis consensus solution algorithm (MCSA).

Finally and in order to show the several practical applications of this research, a new approach is proposed to analyse consumers' behaviour. The empirical findings in this study provide a new understanding of computing family budgets, the consensus family budgets. Based on Spanish data on household budgets from 2016, the rankings of the expenditure groups that best agrees with family preferences are obtained under different assumptions.

Regarding future research, this contribution has thrown up many questions in need further investigation among which the following are noteworthy:

i. The computation of Mahalanobis consensus solutions requires to have a reference matrix suitable to the problem at hand. Therefore, it is essential to develop procedures to determine such a reference matrix. In this sense, it could be interesting to design methodologies for generating this matrix by means of an endogenous way.

ii. The α-codification proposed in this paper could also be used to generated the feasible set $\mathcal{F}$.

iii. Future studies related to Social Choice Theory and Preference Behaviour Theory could be inspired by this research particularly, decision making problems that have a not small set of alternatives.

Acknowledgements

References

[1] N. M. Fraser, Ordinal preference representations, Theory and Decision 36 (1) (1994) 45–67.

[2] Z. Xu, Linguistic Preference Relations, Springer Berlin Heidelberg, Berlin, Heidelberg, 2012, pp. 87–175.

[3] F. Chiclana, F. Herrera, E. Herrera-Viedma, Integrating multiplicative preference relations in a multipurpose decision-making model based on fuzzy preference relations, Fuzzy Sets and Systems 122 (2) (2001) 277 – 291.

[4] J. Borda, Mémoire sur les Elections au Scrutin, Histoire de l'Academie Royale des Sciences, Paris, 1781.

[5] M. Kendall, Rank Correlation Methods, Hafner Publishing Company, New York, 1962.

[6] D. Saari, V. Merlin, A geometric examination of Kemeny's rule, Social Choice and Welfare 17 (2000) 403–438.

[7] T. Meskanen, H. Nurmi, Mathematics and Democracy, Springer Berlin Heidelberg, 2006, Ch. Distance from consensus: A theme and variations, pp. 117–132.

[8] C. Klamler, The Dodgson ranking and its relation to Kemeny's method and Slater's rule, Social Choice and Welfare 23 (1) (2004) 91–102.

[9] C. Klamler, A distance measure for choice functions, Social Choice and Welfare 30 (2008) 419–425.

[10] R. de Andrés Calle, J. García-Lapresta, J. González-Pachón, Performance appraisal based on distance function methods, European Journal of Operational Research 207 (3) (2010) 1599 – 1607.

[11] J. González-Pachón, C. Romero, Inferring consensus weights from pairwise comparison matrices without suitable properties, Annals of Operations Research 154 (2004) 123 – 132.

[12] J. González-Pachón, C. Romero, The design of socially optimal decisions in a consensus scenario, Omega 39 (2) (2011) 179 – 185.

[13] S. France, W. Batchelder, Unsupervised consensus analysis for on-line review and questionnaire data, Information Sciences 283 (2014) 241 – 257.

[14] F. Franceschini, D. Maisano, Checking the consistency of the solution in ordinal semi-democratic decision-making problems, Omega 57 (2015) 188 – 195.

[15] Z. Gong, X. Xu, H. Zhang, U. A. Ozturk, E. Herrera-Viedma, C. Xu, The consensus models with interval preference opinions and their economic interpretation, Omega 55 (2015) 81 – 90.

[16] H. Zhang, R. Yang, H. Yan, F. Yang, H_∞ consensus of event-based multi-agent systems with switching topology, Information Sciences In Press.

[17] Y. Akiyama, J.N., M. Darrah, M. Rahem, L. Wang, A method for measuring consensus within groups: An index of disagreement via conditional probability, Information Sciences 345 (2016) 116 – 128.

[18] T. González-Arteaga, R. de Andrés Calle, J. Alcantud, A new consensus ranking approach for correlated ordinal information based on mahalanobis distance, Information Sciences 372 (Supplement C) (2016) 546 – 564.

[19] G. A. Miller, The magical number seven, plus or minus two: some limits on our capacity for processing information, Psychological Review 63 (1956) 81–97.

[20] D. Li, X. Sun, Nonlinear Integer Programming, Springer US, 2006.

[21] I. J. Good, The number of orderings of n candidates when ties are permitted, Fibonacci Quart. 13 (1975) 11 – 18.

[22] R. W. Bailey, The number of weak orderings of a finite set, Social Choice and Welfare 15 (4) (1998) 559–562.

[23] M. Solomon, R. Russell-Bennett, J. Previte, Consumer Behaviour, Pearson Higher Education, 2012.

[24] A. Biswas, M. Roy, Green products: an exploratory study on the consumer behaviour in emerging economies of the East, Journal of Cleaner Production 87 (2015) 463 – 468.

[25] M. M. Wei, F. Zhang, Recent research developments of strategic consumer behavior in operations management, Computers & Operations Research 93 (2018) 166 – 176.

[26] T. S. Teo, Y. Yu, Online buying behavior: a transaction cost economics perspective, Omega 33 (5) (2005) 451 – 465.

[27] D. Black, Partial justification of the Borda count, Public Choice 28 (1976) 1–16.

[28] W. Cook, L. Seiford, On the Borda–Kendall consensus method for priority ranking problems, Management Science 28 (1982) 621–637.

[29] J. L. García-Lapresta, D. Pérez-Román, Measuring consensus in weak orders, in: E. Herrera-Viedma, J. L. García-Lapresta, J. Kacprzyk, M. Fedrizzi, H. Nurmi, S. Zadrozny (Eds.), Consensual Processes, Vol. 267 of Studies in Fuzziness and Soft Computing, Springer Berlin Heidelberg, 2011, pp. 213–234.

[30] Matlab r2017a, The MathWorks, Inc., Natick, Massachusetts, USA, http://www.mathworks.com.

[31] Gurobi optimizer, v7, Houston, USA. Free Academic License, http://www.gurobi.com.

[32] Household budget survey. Methodology, INE, http://www.ine.es/dyngs/INEbase.

[33] H. Watts, Special panel suggests changes in bls family budget program, Monthly Labor Review 103 (12) (1980) 3–10.

[34] D. Johnson, T. Garner, Developing poverty thresholds using expenditure data, in: Proceedings of the Government and Social Statistics Section, 1997, pp. 28–37.

[35] N. Sloane, The on-line encyclopedia of integer sequences, http://oeis.org/A000670 (1964).

	Group					
	x_1	x_2	x_3	x_4	x_5	x_6
Average Expenditure*	4.47	0.55	1.5	9.03	1.23	1.05
$\overline{m}_i$	10.08	4.22	6.06	11.61	5.71	4.78

	Group					
	x_7	x_8	x_9	x_{10}	x_{11}	x_{12}
Average Expenditure*	3.31	0.82	1.69	0.41	2.68	2.15
$\overline{m}_i$	7.51	5.7	6.39	3.14	7.43	7.9

*: data in thousands of euro.

Table 5: Expenditure averages for original data and codified data.

	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	x_{10}	x_{11}	x_{12}
x_1	**9.24**	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
x_2	0.00	**0.92**	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
x_3	0.00	0.00	**4.85**	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
x_4	0.00	0.00	0.00	**24.99**	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
x_5	0.00	0.00	0.00	0.00	**4.61**	0.00	0.00	0.00	0.00	0.00	0.00	0.00
x_6	0.00	0.00	0.00	0.00	0.00	**8.35**	0.00	0.00	0.00	0.00	0.00	0.00
x_7	0.00	0.00	0.00	0.00	0.00	0.00	**30.47**	0.00	0.00	0.00	0.00	0.00
x_8	0.00	0.00	0.00	0.00	0.00	0.00	0.00	**0.30**	0.00	0.00	0.00	0.00
x_9	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	**6.48**	0.00	0.00	0.00
x_{10}	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	**2.14**	0.00	0.00
x_{11}	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	**14.12**	0.00
x_{12}	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	**7.70**

Table 6: Σ_2: Variance matrix of the expenditure groups from original data.

	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	x_{10}	x_{11}	x_{12}
x_1	9.24	0.48	1.71	3.54	1.23	0.83	3.01	0.44	1.92	0.65	2.33	1.83
x_2	0.48	0.92	0.17	0.21	0.09	0.08	0.46	0.08	0.23	0.02	0.53	0.18
x_3	1.71	0.17	4.85	2.20	1.04	0.63	2.70	0.31	1.69	0.55	2.65	1.58
x_4	3.54	0.21	2.20	24.99	3.31	1.42	5.16	0.67	3.23	1.84	4.99	3.99
x_5	1.23	0.09	1.04	3.31	4.60	0.53	1.70	0.22	1.22	0.60	1.68	1.29
x_6	0.83	0.08	0.63	1.42	0.53	8.35	1.09	0.14	0.74	0.11	0.75	0.80
x_7	3.01	0.46	2.70	5.16	1.70	1.09	30.47	0.79	3.10	1.38	5.88	2.94
x_8	0.44	0.08	0.31	0.67	0.22	0.14	0.79	0.30	0.40	0.16	0.66	0.35
x_9	1.92	0.23	1.69	3.23	1.22	0.74	3.10	0.40	6.48	0.75	3.41	1.61
x_{10}	0.65	0.02	0.55	1.84	0.60	0.11	1.38	0.16	0.75	2.14	1.29	0.73
x_{11}	2.33	0.53	2.65	4.99	1.68	0.75	5.88	0.66	3.41	1.29	14.12	2.48
x_{12}	1.83	0.18	1.58	3.99	1.29	0.80	2.94	0.35	1.61	0.73	2.48	7.70

*: data in thousands of euro.

Table 7: Σ_3: Variance-Covariance matrix of the expenditure groups from original data.

	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	x_{10}	x_{11}	x_{12}
x_1	1.00	0.17	0.26	0.23	0.19	0.09	0.18	0.26	0.25	0.15	0.20	0.22
x_2	0.17	1.00	0.08	0.05	0.05	0.03	0.09	0.16	0.09	0.02	0.15	0.07
x_3	0.26	0.08	1.00	0.20	0.22	0.10	0.22	0.26	0.30	0.17	0.32	0.26
x_4	0.23	0.05	0.20	1.00	0.31	0.10	0.19	0.24	0.25	0.25	0.27	0.29
x_5	0.19	0.05	0.22	0.31	1.00	0.09	0.14	0.19	0.22	0.19	0.21	0.22
x_6	0.09	0.03	0.10	0.10	0.09	1.00	0.07	0.09	0.10	0.03	0.07	0.10
x_7	0.18	0.09	0.22	0.19	0.14	0.07	1.00	0.26	0.22	0.17	0.28	0.19
x_8	0.26	0.16	0.26	0.24	0.19	0.09	0.26	1.00	0.29	0.21	0.32	0.23
x_9	0.25	0.09	0.30	0.25	0.22	0.10	0.22	0.29	1.00	0.20	0.36	0.23
x_{10}	0.15	0.02	0.17	0.25	0.19	0.03	0.17	0.21	0.20	1.00	0.24	0.18
x_{11}	0.20	0.15	0.32	0.27	0.21	0.07	0.28	0.32	0.36	0.24	1.00	0.24
x_{12}	0.22	0.07	0.26	0.29	0.22	0.10	0.19	0.23	0.23	0.18	0.24	1.00

Table 8: Correlation matrix computed from original data.

	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	x_{10}	x_{11}	x_{12}
x_1	**2.44**	0.00	0.00	0.00	0.00	0.00	0.00	0.000	0.00	0.00	0.00	0.00
x_2	0.00	**6.18**	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
x_3	0.00	0.00	**7.07**	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
x_4	0.00	0.00	0.00	**0.58**	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
x_5	0.00	0.00	0.00	0.00	**6.10**	0.00	0.00	0.00	0.00	0.00	0.00	0.00
x_6	0.00	0.00	0.00	0.00	0.00	**7.34**	0.00	0.00	0.00	0.00	0.00	0.00
x_7	0.00	0.00	0.00	0.00	0.00	0.00	**8.55**	0.00	0.00	0.00	0.00	0.00
x_8	0.00	0.00	0.00	0.00	0.00	0.00	0.00	**3.88**	0.00	0.00	0.00	0.00
x_9	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	**6.19**	0.00	0.00	0.00
x_{10}	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	**4.88**	0.00	0.00
x_{11}	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	**7.30**	0.00
x_{12}	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	**3.40**

Table 9: Σ_4: Variance matrix of the codified profile.

	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	x_{10}	x_{11}	x_{12}
x_1	2.44	0.49	-0.34	0.03	0.14	0.08	-0.97	0.50	-0.426	-0.09	-1.15	0.16
x_2	0.49	6.18	-0.95	0.12	0.13	-0.13	-0.86	0.93	-0.81	-0.03	-0.74	0.00
x_3	-0.34	-0.95	7.07	-0.13	-0.39	-0.61	-1.03	-0.64	-0.61	-0.17	-0.66	-0.34
x_4	0.03	0.12	-0.13	0.58	0.11	0.03	-0.49	0.27	-0.07	0.13	-0.23	0.13
x_5	0.14	0.13	-0.39	0.11	6.10	0.36	-1.65	0.35	-0.59	0.36	-1.52	0.39
x_6	0.08	-0.13	-0.61	0.03	0.36	7.34	-1.49	0.37	-0.92	-0.01	-1.69	0.22
x_7	-0.97	-0.86	-1.04	-0.49	-1.65	-1.49	8.55	-0.95	-0.82	-0.48	0.17	-1.07
x_8	0.50	0.93	-0.64	0.27	0.35	0.37	-0.95	3.88	-0.58	0.61	-1.27	0.67
x_9	-0.42	-0.81	-0.61	-0.07	-0.59	-0.92	-0.82	-0.58	6.19	-0.03	-0.12	-0.66
x_{10}	-0.09	-0.03	-0.17	0.13	0.36	-0.01	-0.48	0.61	-0.03	4.88	-0.32	0.15
x_{11}	-1.15	-0.74	-0.66	-0.23	-1.52	-1.69	0.16	-1.27	-0.12	-0.32	7.30	-1.14
x_{12}	0.16	0.00	-0.35	0.13	0.39	0.23	-1.07	0.67	-0.65	0.15	-1.14	3.39

Table 10: Σ_5: Variance-Covariance matrix computed of the codified profile.

	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	x_{10}	x_{11}	x_{12}
x_1	1.00	0.13	-0.08	0.03	0.04	0.02	-0.21	0.16	-0.11	-0.03	-0.27	0.06
x_2	0.13	1.00	-0.14	0.06	0.02	-0.02	-0.12	0.19	-0.13	-0.01	-0.11	0.00
x_3	-0.08	-0.14	1.00	-0.07	-0.06	-0.08	-0.13	-0.12	-0.09	-0.03	-0.09	-0.07
x_4	0.03	0.06	-0.07	1.00	0.06	0.02	-0.22	0.18	-0.04	0.08	-0.12	0.10
x_5	0.04	0.02	-0.06	0.06	1.00	0.05	-0.23	0.07	-0.10	0.07	-0.23	0.09
x_6	0.02	-0.02	-0.08	0.02	0.05	1.00	-0.19	0.07	-0.14	-0.00	-0.23	0.05
x_7	-0.21	-0.12	-0.13	-0.22	-0.23	-0.19	1.00	-0.16	-0.11	-0.07	0.02	-0.20
x_8	0.16	0.19	-0.12	0.18	0.07	0.07	-0.16	1.00	-0.12	0.14	-0.24	0.19
x_9	-0.11	-0.13	-0.09	-0.04	-0.10	-0.14	-0.11	-0.12	1.00	-0.01	-0.02	-0.14
x_{10}	-0.03	-0.01	-0.03	0.08	0.07	-0.00	-0.07	0.14	-0.01	1.00	-0.05	0.04
x_{11}	-0.27	-0.11	-0.09	-0.12	-0.23	-0.23	0.02	-0.24	-0.02	-0.05	1.00	-0.23
x_{12}	0.06	0.00	-0.07	0.10	0.09	0.05	-0.20	0.19	-0.14	0.04	-0.23	1.00

Table 11: Correlation matrix from the codified profile $\mathcal{M}_{\mathcal{P}}$.

	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	x_{10}	x_{11}	x_{12}
x_1	2.45	0.65	1.07	0.28	0.73	0.40	0.82	0.82	0.97	0.51	0.86	0.63
x_2	0.65	6.19	0.54	0.09	0.28	0.21	0.64	0.80	0.59	0.11	1.00	0.32
x_3	1.07	0.54	7.07	0.41	1.45	0.72	1.73	1.37	2.00	1.00	2.31	1.27
x_4	0.28	0.09	0.41	0.58	0.58	0.20	0.42	0.37	0.48	0.42	0.55	0.40
x_5	0.73	0.28	1.45	0.58	6.10	0.57	1.04	0.94	1.38	1.04	1.40	0.99
x_6	0.40	0.21	0.72	0.20	0.57	7.34	0.54	0.49	0.69	0.17	0.51	0.50
x_7	0.82	0.64	1.73	0.42	1.04	0.54	8.55	1.51	1.61	1.11	2.24	1.04
x_8	0.82	0.80	1.37	0.37	0.94	0.49	1.51	3.88	1.41	0.91	1.70	0.85
x_9	0.97	0.59	2.00	0.48	1.38	0.69	1.61	1.41	6.19	1.12	2.40	1.05
x_{10}	0.51	0.11	1.00	0.42	1.04	0.17	1.11	0.91	1.12	4.88	1.40	0.74
x_{11}	0.86	1.00	2.31	0.55	1.40	0.51	2.24	1.70	2.40	1.40	7.30	1.19
x_{12}	0.63	0.32	1.27	0.40	0.99	0.50	1.04	0.85	1.05	0.74	1.19	3.40

Table 12: Σ_6. This matrix contains the covariances computed from the gadget and the variances computed from the codified profile
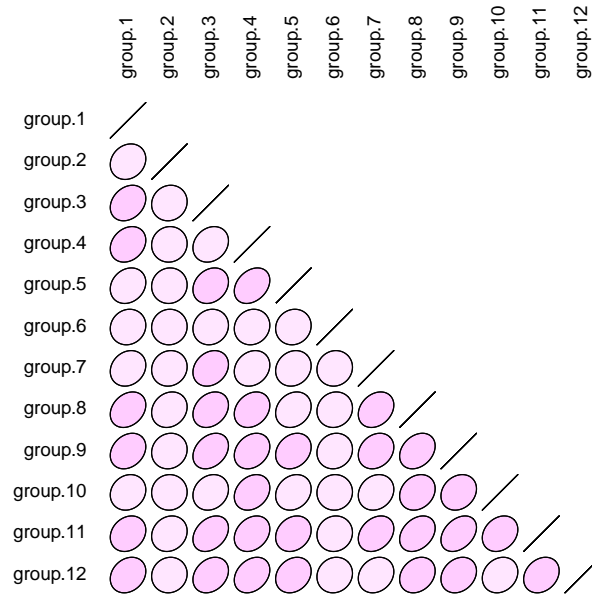
Figure 2: Graphical interpretation of the correlation matrix (Table 8) of the Spanish expenditure groups from original data.

	O. F. value	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	x_{10}	x_{11}	x_{12}
								Group					
Σ_1	12.64	11	4	6	12	6	4	9	4	7	4	8	10
Σ_2	1.9	11	4	6	12	4	4	10	6	7	4	9	8
Σ_3	2.04	11	4	6	12	4	4	10	6	7	4	9	8
Σ_4	2.43	11	3	6	12	6	3	10	6	7	3	9	8
Σ_5	3.5	10	3	5	12	4	3	11	6	7	3	9	8
Σ_6	2.41	11	3	6	12	6	3	10	6	7	3	8	9

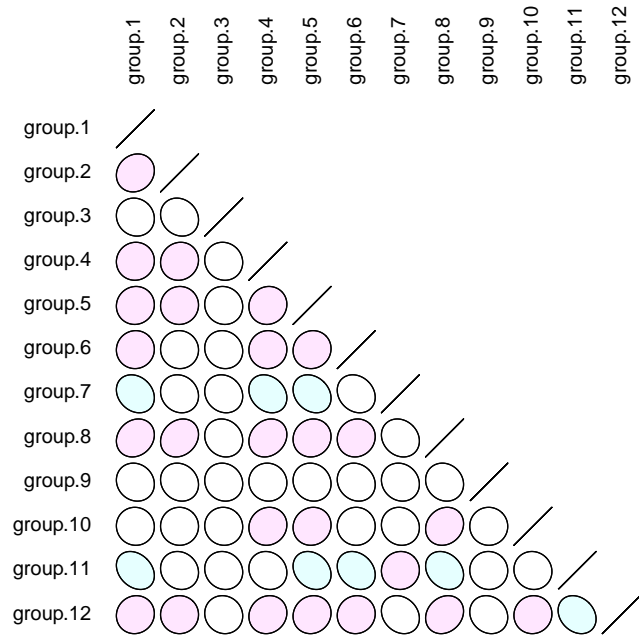Table 13: Group position for different Σ matrices.

Figure 3: Graphical interpretation of the correlation matrix (Table 11) of the Spanish expenditure groups from codified profile.

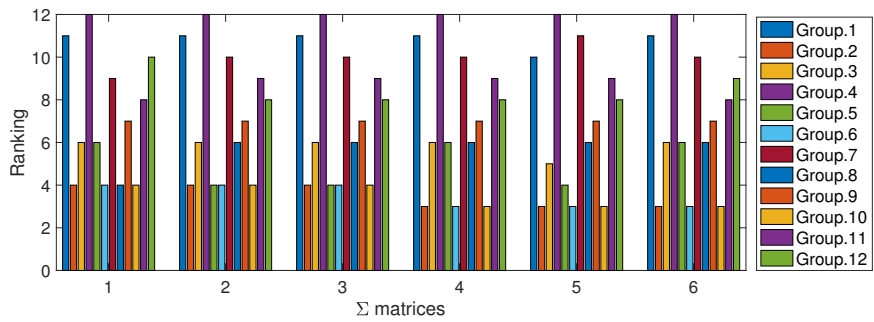Figure 4: Group position/ranking for $\{\Sigma_i\}_{i=1}^{6}$ matrices. Results are clustered according the matrix Σ_i used to compute the Mahalanobis distance.
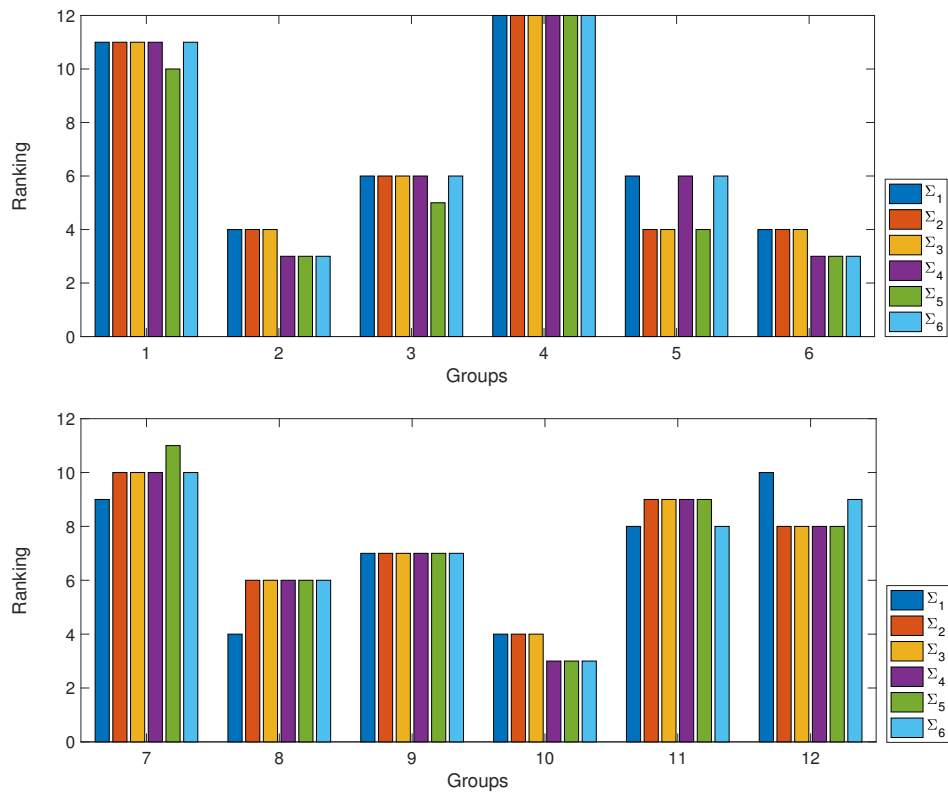
Figure 5: Group position/ranking for $\{\Sigma_i\}_{i=1}^{6}$ matrices. Results are clustered by expenditure group.