Ordinal Sorting in the Presence of Interacting Points of View: TOMASO

Patrick Meyer
Complex Enterprise Systems Institute
Faculty of Law, Economy and Finance
University of Luxembourg
162a, avenue de la Faïencerie
L-1511 Luxembourg (G.D. Luxembourg)
patrick.meyer@internet.lu

Marc Roubens
MATHRO
Faculté Polytechnique of Mons
9, rue de Houdain
B-7OOO Mons (Belgium)
m.roubens@ulg.ac.be

1 Introduction

This paper presents an ordered sorting procedure based on the Choquet integral as a discriminant function. It uses information provided by the Decision Maker (DM) in terms of a set of prototypes (alternatives well-known to the DM). The capacities of the Choquet integral are assessed through the solving of a linear program or a quadratic program. An interpretation of the results is provided by means of importance and interaction indexes of the points of view.

We analyze a sorting procedure for ordinal data in a very general case, where the points of view can have interactions. Its name, TOMASO stands for **T**ool for **O**rdinal **M**ulti**A**ttribute **S**orting and **O**rdering. The first version of this method has been described in [7] and [9]. Later, in [6] the authors present further evolutions to the first ideas, and describe a software which is directly inspired from the sorting procedure.

Three important features differentiate this procedure from other multiple criteria sorting methods. First of all, the possibility to treat purely ordinal data. Secondly, the use of a Choquet integral [1] as a discriminant function. And finally, the way the capacities ("weights") of the aggregator are learnt from a reference set of alternatives called prototypes. These three key features allow to treat a quite large set of problems. In particular, the learning feature of the method is interesting as it allows to ask the Decision Maker (DM) a minimal set of technical details. In order to allow a more effective and objective analysis of the problem, we think that it is useful to have a permanent interaction with the DM. But this questioning should mainly be restricted to his expertise domain and not to technical parameters of the method. The use of the prototypes fits to this philosophy.

The method works in two steps. First of all, the ordinal data is transformed into partial net scores, where each alternative is compared to all the other ones for each point of view. Then, the Choquet integral is used to aggregate these partial net scores. As already mentioned earlier, the capacities of the aggregator are learnt from the reference set of prototypes. Here, two options appear: either the prototypes don't violate the axioms ([11]) for the use of a Choquet integral as a discriminant function, or the structure of the prototypes does not allow its use as an aggregator. In the first case, the capacities are learnt by solving a linear constraints satisfaction problem. This procedure is briefly recalled in section 3.1. In the second case, the capacities are learnt by trying to be as close as possible to the original sorting imposed by the prototypes. This part is described in section 3.2.

This paper is organized as follows. First of all, general concepts are introduced in section 2. Then, in section 3.1 we recall the first ideas of TOMASO already published in [6]. In section 3.2 we present how to work in case the classical way fails. Finally, in 4 we draw some conclusions, and discuss further improvements.

2 Preliminary considerations

Let A be a set of q potential alternatives which are to be assigned to disjoint ordered classes. Let $F = \{g_1, \ldots, g_n\}$ be a set of points of view. For each index of point of view $j \in \mathcal{J} = \{1, \ldots, n\}$, the alternatives are evaluated according to a s_j-point ordinal performance scale represented by a totally ordered set

$$X_j := \{g_1^j \prec_j \ldots \prec_j g_{s_j}^j\}.$$

Therefore, an alternative $x \in A$ can be identified with its corresponding profile

$$(x_1, \ldots, x_n) \in \prod_{j=1}^{n} X_j =: X,$$

where for any $j \in \mathcal{J}$, x_j is the partial evaluation of x on point of view j.

Let us consider a partition of $X := \Pi_{j=1}^{n} X_j$ into m nonempty increasingly ordered classes $\{Cl_t\}_{t=1}^{m}$. This means that for any $r, s \in \{1, \ldots, m\}$, with $r > s$, the elements of Cl_r are considered as better than the elements of Cl_s. The sorting problem we are dealing with consists in partitioning assigning the alternatives of A to the classes $\{Cl_t\}_{t=1}^{m}$.

In Roubens [9] it is justified how an n-place Choquet integral as a discriminant function and normalised scores as criteria function can be used to solve this problem. Hereafter we present the sorting procedure derived from this particular case.

3 The Tomaso method

The Tomaso method (**T**echnique for **O**rdinal **M**ulti**a**ttribute **S**orting and **O**rdering) is mainly based on two techniques (which can lead to the same results under certain conditions). The original method has first been described in [9]. In the following Subsection, we present its basics. In Subsection 3.2 we show how it is possible to deal with a larger set of problems.

3.1 The classical way

The different stages of the original Tomaso are listed below:

1. Modification of the criteria evaluations into normalised scores;
2. Use of a Choquet integral as a discriminant function;
3. Assessment of fuzzy measures by questionning the DM and by solving a linear constraint satisfaction problem;
4. Calculation of the borders of the classes and assignment of the alternatives to the classes;
5. Analysis of the results (interaction, importance, leave one out, visualisation).

In this Section we roughly present these different elements.

First of all, concerning the scales on the points of view, two natural approaches can be considered: either the score of each alternative is built on the basis of all the alternatives in A or this score is constructed in a context-free manner, that is, independently of the other alternatives. The decision maker must be aware that the final results may significantly differ according to the considered approach. Therefore, a prior analysis of the problem is recommended to choose the scores appropriately.

In the first approach, one possible way to build the scores is to consider comparisons of the alternatives on each of the points of view. We consider $S_j(x)$, the jth partial net score of alternative $x \in A$ along point of view $j \in \mathcal{J}$, as the number of times that x is preferred to any other alternative of A minus the number of times that any other alternative of A is preferred to x for point of view j. We furthermore normalize these scores so that they range in the unit interval, i.e.,

$$S_j^N(x) := \frac{S_j(x) + (q-1)}{2(q-1)} \in [0,1] \qquad (j \in \mathcal{J},$$

where $q = |A|$. Clearly, this normalized score is not a utility, and should not be considered as such. Indeed, observing an extreme value (close to 0 or 1) means that x is rather "atypical" compared to the other alternatives along point of view j. Thus, the resulting evaluations strongly depend on the alternatives which have been chosen to build A.

Consider now the second approach, that is, where the score of each alternative does not depend on the other alternatives in A. In this case, we suggest the decision maker provides the score functions as utility functions. Alternatively, we can approximate these utility functions by the following linear formula:

$$S_j^N(x) := \frac{\mathrm{ord}_j(x) - 1}{s_j - 1} \in [0,1] \qquad (j \in \mathcal{J}),$$

where $\mathrm{ord}_j : A \to \{1, \ldots, s_j\}$ is a mapping defined by $\mathrm{ord}_j(x) = r$ if and only if $x_j = g_r^j$. In this latter case, S_j^N does not necessarily represent a real utility and probably does not correspond to the utility the decision maker has in mind. We therefore continue to call it a score.

We now come to the crucial part of the aggregation of the normalised partial net scores of a given alternative x by means of a Choquet integral [1]. The advantage of this aggregator is mainly that it allows to deal with interacting (depending) points of view. According to the general definition of the Choquet integral, we have:

$$\mathcal{C}_v(S^N(x)) := \sum_{j=1}^{n} S_{(j)}^N(x)[v(A_{(j)}) - v(A_{(j+1)})]$$

where v is a fuzzy measure on $\mathcal{J}$; that is a monotone set function $v : 2^{\mathcal{J}} \to [0,1]$ fulfilling $v(\emptyset) = 0$ and $v(\mathcal{J}) = 1$. The parentheses used for indexes stand for a permutation on $\mathcal{J}$ such that

$$S_{(1)}^N(x) \leq \ldots \leq S_{(n)}^N(x),$$

and for any $j \in \mathcal{J}$, $A_{(j)}$ represents the subset $\{(j), \ldots, (n)\}$. The characterisation of the Choquet integral by Marichal ([4], [5]) clearly justifies the way the partial scores are aggregated.

The next step of this method is to assess the fuzzy measures in order to classify the alternatives of A. One can easily understand that it is impossible to ask the DM to give values for the $2^n - 2$ free parameters of the fuzzy measure v. Practically, the assessment of the fuzzy measures is done by asking the DM to provide a set of prototypes $P \subseteq A$ and their assignments to the given classes; that is a partition of P into prototypic classes $\{P_t\}_{t=1}^m$ where $P_t := P \cap Cl_t$ for $t \in \{1, \ldots, m\}$. The values of the fuzzy measure are then derived from this information as described hereafter.

We would like the Choquet integral to strictly separate the classes Cl_t. Therefore, the following necessary condition is imposed

$$\mathcal{C}_v(S^N(x)) - \mathcal{C}_v(S^N(x')) \geq \varepsilon \qquad (1)$$

for each ordered pair $(x, x') \in P_t \times P_{t-1}$ and each $t \in \{2, \ldots, m\}$, where ε is a given strictly positive thershold.

Due to the increasing monotonicity of the Choquet integral, the number of separation constraints 1 can be reduced significantly. Thus, it is enough to consider *border elements* of the classes. To formalise this concept, we first define a dominance relation D (partial order) on X by

$$xDy \iff x_j \succeq_j y_j, \text{ for all } j \in \mathcal{J}.$$

As *upper border* of the prototypic class P_t we use the set of non-dominated alternatives of P_t defined by

$$ND_t := \{x \in P_t \text{ s.t. } \nexists x' \in P_t \setminus \{x\} : x'Dx\}.$$

Similarly, the *lower border* of the prototypic class is given by the set of non-dominating alternatives of P_t which is defined by

$$Nd_t := \{x \in P_t \text{ s.t. } \nexists x' \in P_t \setminus \{x\} : xDx'\}.$$

The separation conditions restricted to the prototypes of the subsets $ND_t \cup Nd_t$, $t \in \{1, \ldots, m\}$ put together with the monotonicity constraints on the fuzzy measure, form a linear program [7] whose unknowns are the capacities $v(S)$, $S \subset \mathcal{J}$ and where ε is a non-negative variable to be maximised in order to deliver well separated classes.

We use the principle of parsimony for the resolution of this problem. If there exists a k-additive fuzzy measure v^* ([3]), k being kept as low as possible, then we determine the boundaries of the classes as follows:

- lower boundary of Cl_t: $z(t) := \min_{x \in Nd_t} \mathcal{C}_{v^*}(S^N(x))$;
- upper boundary of Cl_t: $Z(t) := \max_{x \in ND_t} \mathcal{C}_{v^*}(S^N(x))$.

At this point, any alternative $x \in A$ can be classified in the following way:

- x is assigned to class Cl_t if $z_t \leq \mathcal{C}_{v^*}(S^N(x)) \leq Z_t$;
- x is assigned to class $Cl_t \cup Cl_{t-1}$ if $Z_{t-1} < \mathcal{C}_{v^*}(S^N(x)) < z_t$.

A final step of the classical TOMASO method concerns the evaluation of the results and the interpretation of the behavior of the Choquet integral. The meaning of the values $v(T)$ is not clear to the DM. They don't immediatly indicate the global importance of the points of view, nor their degree of interaction. It is possible to derive some indexes from the fuzzy measure which are helpful to interpret its behavior. Among them, the TOMASO method proposes to have a closer look at the importance indexes [10] and the interaction indexes [8].

3.2 An alternate way

It may happen that the linear program described in Subsection 3.1 has no solution. This occurs when the prototypic elements violate the axioms that are imposed to produce a discriminant function of Choquet type ([5] [11]), in particular the triple cancellation axiom.

In such a case, and in order to present a solution to the DM, we suggest to find a fuzzy measure by solving the following quadratic program

$$\min \sum_{x \in \cup_{t \in \{1,\ldots,m\}} \{ND_t \cup ND_t\}} [\mathcal{C}_v(S^N(x)) - y(x)]^2,$$

where the unkowns are

- the capacities $v(S)$ which determine the fuzzy measure;
- some global evaluations $y(x)$ for each $x \in \cup_{t \in \{1,\ldots,m\}} \{ND_t \cup Nd_t\}$.

The capacities $v(S)$ are constrained by the monotonicity conditions (as previously shown in Section 3.1). The global evaluations $y(x)$ must verify the classification imposed by the DM. In other words, for every ordered pair $(x, x') \in Nd_t \times ND_{t-1}$, $t \in \{2, \ldots, m\}$ the condition $y(x) - y(x') \geq \varepsilon'$, $\varepsilon' > 0$ must be satisfied.

Intuitively, for a given alternative $x \in A$, its Choquet integral $\mathcal{C}_v(S^N(x))$ should be as close as possible to the global evaluation $y(x)$, without being constrained by monotonicity conditions which might violate the triple cancellation axiom for example. On the other hand, the evaluation $y(x)$ is constrained by the conditions derived from the original classification given by the DM on the prototypes.

Unlike the method described in Section 3.1, in this case, ε' plays the role of a parameter, which needs to be fixed by the DM. As previously, we use the principle of parsimony when searching for a solution (keep k as low as possible; at worst k equals the number of points of view). A correct choice of ε' remains one of the main challenges of our future research. It is clear that ε' has to be chosen in $]0, 1/n[$.

As in the classical method, the next step is to determine the structure of the classes. We determine an assignment for every alternative of X in terms of intervals of contiguous classes on the basis of the information provided by the Choquet integrals related to the prototypes of $P \subseteq A$.

First of all, let us suppose that $S^N(x^-) := (0, \ldots, 0)$ is classified to the worst class, Cl_1 and that $S^N(x^+) := (1, \ldots, 1)$ is classified to the best class, Cl_m.

To each assignment $I(x)$ correspond a lower class label $\underline{l}(x)$ and an upper class label $\bar{l}(x)$, $\underline{l}, \bar{l} \in \mathcal{J}$. We say that the alternative $x \in X$ is *precisely assigned* to $Cl_{l(x)}$ if for the assignment $I(x)$ we have $\underline{l}(x) = \bar{l}(x) =: l(x)$. Else, the alternative x is said to be *ambiguously assigned* to the interval of labels $I(x) = [\underline{l}(x), \bar{l}(x)]$. The *degree of the assignment* corresponds to the number of contiguous classes contained in $I(x)$, $d(x) = \bar{l}(x) - \underline{l}(x) + 1$.

The assignments are done according to the procedure described hereafter. Starting from the prototypes $x \in P$, their Choquet integrals $\mathcal{C}_v(S^N(x))$ and their original classification label $Cl(x)$ (according to the DM's choice), we define for every $u \in [0, 1]$,

$$m(u) = \max_{x \in P : \mathcal{C}_v(S^N(x)) \leq u} Cl(x), \text{and}$$

$$M(u) = \min_{x \in P : \mathcal{C}_v(S^N(x)) \geq u} Cl(x).$$

m (resp. M) is a right (resp. left) continuous stepwise function of argument u with values belonging to the discrete finite set $\mathcal{J}$.

We now define for each $u \in [0, 1]$ an interval of contiguous classes $I(u) = [\underline{l}(u), \bar{l}(u)]$ where

$$\underline{l}(u) = \min\{m(u), M(u)\}$$

$$\bar{l}(u) = \max\{m(u), M(u)\}.$$

Obviously $\underline{l}(u) \leq \bar{l}(u)$ and due to monotonicity of m and M we have: $\underline{l}(u) \leq \underline{l}(v), \bar{l}(u) \leq \bar{l}(v), \forall u, v \in [0, 1]$ with $u \leq v$.

The interval $[0, 1]$ is partitioned into (closed, semi-open or open) intervals $I_s, s = 1, \ldots, S$, and each of those intervals of $[0, 1]$ receives an assignment of the type $[\underline{l}(s), \bar{l}(s)]$ (or semi-open or open) in such a way that: if $u, v \in [0, 1], u \leq v$ and if u is assigned to $I_r := [\underline{l}(r), \bar{l}(r)]$ and v is assigned to $I_{r'} := [\underline{l}(r'), \bar{l}(r')]$ then $\underline{l}(r) \leq \underline{l}(r')$ and $\bar{l}(r) \leq \bar{l}(r')$.

Moreover if $u = \mathcal{C}_v(S^N(x)), x \in P$ then $\underline{l}(u) \leq Cl(x) \leq \bar{l}(u)$. This means that each prototype is *correctly classified*, possibly with ambiguity if $d(x) \geq 1$.

The assignment of a prototype a to the intervals of classes leads now to two scenarios:

- a is assigned to a single class ($d(a) = 1$) which corresponds to the original class decided by the DM
- a is assigned to an interval of classes and the original class decided by the DM belongs to this interval.

The quality of a model (classifier) depends on different ratios. A good model has the following *natural* properties:

- a simple model according to parsimony (low k);
- a high number of precise assignments of the elements of P;
- a low number of ambiguous assignments of the elements of P (and the lower the degree of the assignment, the better the model)

For a given ε', the DM has to select a model (k) which seems the best compromise to him in terms of the previously described assignments. The simplest additive model ($k = 1$) can in certain situations be this ideal compromise between simplicity and quality. But in more complex problems, k has to be increased in order to obtain a satisfying number of precisely assigned prototypes.

3.3 Behavioral analysis of aggregation

Now that we have a sorting model for assigning alternatives to classes (based on the linear program or the quadratic program), an important question arises: How can we interpret the behavior of the Choquet integral or that of its associated fuzzy measure? Of course the meaning of the values $v(T)$ is not always clear for the DM. These values do not give immediately the global importance of the points of view, nor the degree of interaction among them.

In fact, from a given fuzzy measure, it is possible to derive some indexes or parameters that will enable us to interpret the behavior of the fuzzy measure. These indexes constitute a kind of *id card* of the fuzzy measure. The TOMASO method presently allows to analyse both the importance of points of view (Shapley indexes [10]), and their interactions ([8]).

3.4 Interpretation of the behaviour of the fuzzy measure

In this Section we briefly show the main advantage to use a Choquet integral rather than the weighted sum as a discriminant function. We therefore take the simple case of two points of view, which can be represented in a plane. Figure 1 presents 5 possible ranges of values for the weights v and the corresponding structures of the limits of the classes. One can see that the main difference between the classical weighted sum and the Choquet integral is the greater flexibility of the borders of the classes. The Choquet integral creates piecewise linear borders, which allows to build more precise classes. The different possibilities are summarised by the following list:

- I: $v(1) + v(2) < v(12)$: synergy
- II: $v(1) + v(2) > v(12)$: redundancy
- III: $v(1) + v(2) = v(12) = 1$: additivity
- IV: $v(1) = v(2) = 0$: limit case; maximal synergy
- V: $v(1) = v(2) = 1$: limit case; maximal redundancy

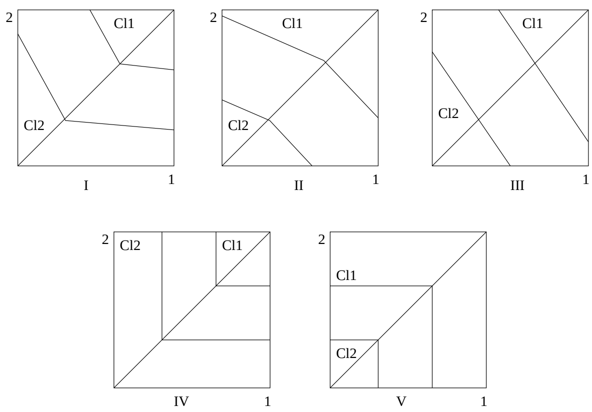

Figure 1: Interpretation of the discriminant functions

In [2] the authors give an interpretation to the first two cases. In case of synergy, although the importance of a single criterion for the decision is rather low, the importance of the pair is large. The criteria are said to be *complementary*. In case of redundance, or negative synergy, the union of criteria does not bring much, and the importance of the pair might be roughly the same as the importance of a single criterion.

The limit case (IV) occurs for maximal synergy. In that case, the Choquet integral corresponds to the aggregation by the min function. Maximal redundancy occurs for case (V), where the Choquet integral is the max function.

In case the number of points of view is larger than two, it becomes quite hard to represent the problem. Nevertheless, the previous short example helps to understand how the borders of the classes are built in such more general examples.

3.5 The software TOMASO

In this short part of the paper we briefly present the key-characteristics of the software TOMASO . It can be downloaded on http://patrickmeyer.tripod.com. It is an implementation of the algorithms which were presented previously. Its name stands for "Tool for Ordinal MultiAttribute Sorting and Ordering". It is written in Visual Basic and uses two external solvers: a free linear program solver (lp_solve 3.0, ftp://ftp.ics.ele.tue.nl/pub/lp_solve/, released under the LGPL license), and a non free quadratic program solver (bpmpd, free trial version at http://www.sztaki.hu/ meszaros/bpmpd/).

It is still under development and many improvements are added on a regular basis. The general steps of the software are outlined hereafter:

- Loading of the ordinal data;
- Choice of a scoring method according to the problem's specificities and calculation of the normalised partial net scores;
- Definition of the prototypes by the DM;
- Search for a fuzzy measure (either by the linear program, or the quadratic program)
- Analysis of the results (classes, Shapley indexes, interaction indexes, accuracies, ...)

A detailed description of the software can be obtained from the author.

4 Concluding remarks

We have presented a procedure for ordinal sorting in the presence of interacting points of view. It has already been applied to real life cases (in particular to a noise annoyance problem) quite successfully. Future work will concern the simplification of the software in order to make it even more user-friendly. Furthermore, the automatic determination of ε' will also be one of our main concerns. The implementation of other indexes (veto, favour, ...) is also planned.

References

[1] G. Choquet, Theory of capacities, *Annales de l'Institut Fourier*, 5, (1953) 131-295.

[2] M. Grabisch, M. Roubens, Application of the Choquet Integral in Multicriteria Decision Making, In: M. Grabisch, T. Murofushi, M. Sugeno (eds.): *Fuzzy Measures and Integrals*, Physica Verlag, Heidelberg, (2000) 348-374.

[3] M. Grabisch, k-order additive discrete fuzzy measure and their representation, *Fuzzy Sets and Systems*, 92, (1997) 167-189.

[4] J.-L. Marichal, *Aggregation operators for multicriteria decision aid*, Ph.D. thesis, Institute of Mathematics, University of Liège, Liège, Belgium, (1998).

[5] J-L. Marichal, An axiomatic approach of the discrete Choquet integral as a tool to aggregate interacting criteria, *IEEE Transactions on Fuzzy Systems*, 8, (2000) 800-807.

[6] J.-L. Marichal, P. Meyer and M. Roubens, Sorting multiattribute alternatives: The TOMASO method, *International Journal of Computers & Operations Research*, 2003, in press.

[7] J-L. Marichal, M. Roubens, On a sorting procedure in the presence of qualitative interacting points of view. In: J. Chojean, J. Leski (eds.):*Fuzzy Sets and their Applications*. Silesian University Press, Gliwice, (2001) 217-230.

[8] T. Murofushi, S. Soneda, Techniques for reading fuzzy measures (III): interaction index, *9th Fuzzy Sytem Symposium*, Sapporo, Japan, (1993) 693-696. In Japanese.

[9] M. Roubens, Ordinal multiattribute sorting and ordering in the presence of interacting points of view. In: D. Bouyssou, E. Jacquet-Lagrèze, P. Perny, R. Slowinsky, D. Vanderpooten and P. Vincke (eds.): *Aiding Decisions with Multiple Criteria: Essays in Honour of Bernard Roy*. Kluwer Academic Publishers, Dordrecht, (2001) 229-246.

[10] L.S. Shapley, A value for n-person games. In: H.W. Kuhn, A.W. Tucker (eds.): *Contributions to the Theory of Games*, Vol. II, Annals of Mathematics Studies, 28, Princeton University Press, Princeton, NJ, (1953) 307-317.

[11] P. Wakker, *Additive Representations of Preferences: A new Foundation of Decision Analysis*, Kluwer Academic Publishers, Dordrecht, Boston, London, (1989).