Chapter 9
Robust Ordinal Regression

Salvatore Greco, Roman Słowiński, José Rui Figueira, and Vincent Mousseau

Abstract Within disaggregation–aggregation approach, *ordinal regression* aims at inducing parameters of a preference model, for example, parameters of a value function, which represent some holistic preference comparisons of alternatives given by the Decision Maker (DM). Usually, from among many sets of parameters of a preference model representing the preference information given by the DM, only one specific set is selected and used to work out a recommendation. For example, while there exist many value functions representing the holistic preference information given by the DM, only one value function is typically used to recommend the best choice, sorting, or ranking of alternatives. Since the selection of one from among many sets of parameters of the preference model compatible with the preference information given by the DM is rather arbitrary, *robust ordinal regression* proposes taking into account all the sets of parameters compatible with the preference information, in order to give a recommendation in terms of necessary and possible consequences of applying all the compatible preference models on the considered set of alternatives. In this chapter, we present the basic principle of robust ordinal

S. Greco (✉)
Faculty of Economics, University of Catania, Corso Italia 55, 95 129 Catania, Italy
e-mail: salgreco@unict.it

R. Słowiński
Institute of Computing Science, Poznań University of Technology, Piotrowo 2,
60-965 Poznań, Systems Research Institute, Polish Academy of Sciences,
00-441 Warsaw, Poland
e-mail: roman.slowinski@cs.put.poznan.pl

J.R. Figueira
CEG-IST Center for Management Studies, Instituto Superior Técnico, Technical University
of Lisbon and TagusPark, Av. Cavaco Silva, 2780-990 Porto Salvo, Portugal
LAMSADE Université Paris-Dauphine, France
e-mail: figueira@ist.utl.pt

V. Mousseau
Laboratoire de Génie Industriel, Ecole Centrale Paris, Grande Voie des Vignes,
92 295 Châtenay-Malabry Cedex, France
e-mail: vincent.mousseau@ecp.fr

M. Ehrgott et al. (eds.), *Trends in Multiple Criteria Decision Analysis*,
International Series in Operations Research and Management Science 142,
DOI 10.1007/978-1-4419-5904-1_9,

regression, and the main multiple criteria decision methods to which it has been applied. In particular, *UTA*^{GMS} and *GRIP* methods are described, dealing with choice and ranking problems, then *UTADIS*^{GMS}, dealing with sorting (ordinal classification) problems. Next, we present robust ordinal regression applied to Choquet integral for choice, sorting, and ranking problems, with the aim of representing interactions between criteria. This is followed by a characterization of robust ordinal regression applied to outranking methods and to multiple criteria group decisions. Finally, we describe an interactive multiobjective optimization methodology based on robust ordinal regression, and an evolutionary multiobjective optimization method, called *NEMO*, which is also using the principle of robust ordinal regression.

Keywords Robust ordinal regression · Multiple criteria · Choice, sorting and ranking · Additive value functions · Choquet integral · Outranking methods · Multiple criteria group decisions · Interactive multiobjective optimization · Evolutionary multiobjective optimization

9.1 Introduction

In Multiple Criteria Decision Analysis (MCDA) (for a recent state of the art see [14]), an alternative a, belonging to a finite set of alternatives $A = \{a, b, \ldots\}$ ($|A| = m$), is evaluated on the basis of a family of n criteria $F = \{g_1, g_2, \ldots, g_i, \ldots, g_n\}$, with $g_i : A \rightarrow \mathbb{R}$. For example, in a decision problem regarding a recruitment of new employees, the alternatives are the candidates and the criteria can be a certain number of characteristics useful to give a comprehensive evaluation of the candidates, such as educational degree, professional experience, age, and interview assessment. From here on, we will use the term criterion g_i, or criterion i interchangeably ($i = 1, 2, \ldots, n$). For the sake of simplicity, but without loss of generality, we suppose that the evaluations on criteria are increasing with respect to preference, i.e., the more the better, defining a marginal weak preference relation as follows:

$$\text{“}a \text{ is at least as good as } b\text{” with respect to criterion } i \Leftrightarrow g_i(a) \geq g_i(b).$$

The purpose of Multi-Attribute Utility Theory (MAUT) [13, 42] is to represent the preferences of the Decision Maker (DM) on a set of alternatives, A, by an overall value function $U(g_1(\cdot), \ldots, g_n(\cdot)) : \mathbb{R}^n \rightarrow \mathbb{R}$, such that:

- a is indifferent to b $\Leftrightarrow$ $U(\underline{g}(a)) = U(\underline{g}(b))$;
- a is preferred to b $\Leftrightarrow$ $U(\underline{g}(a)) > U(\underline{g}(b))$,

where for simplicity of notation, we used $U(\underline{g}(a))$, instead of $U((g_1(a), \ldots, g_n(a)))$.

The principal value function aggregation model is the multiple attribute additive utility [42]:

$$U(\underline{g}(a)) = u_1(g_1(a)) + u_2(g_2(a)) + \cdots + u_n(g_n(a)) \quad \text{with} \quad a \in A,$$

where u_i are nondecreasing marginal value functions, $i = 1, 2, \ldots, n$.

Even if multiple attribute additive utility is the most well-known aggregation model, some critics have been advanced to it because it does not permit to represent interactions between the considered criteria. For example, in evaluating a car one can consider criteria such as maximum speed, acceleration, and price. In this case, very often there is a negative interaction (redundancy) between maximum speed and acceleration of cars: in fact, a car with a high maximum speed has, usually, also a good acceleration and thus, even if these two criteria can be very important for a person who likes sport cars, their comprehensive importance is smaller than the importance of the two criteria considered separately. In the same decision problem, very often there is a positive interaction (synergy) between maximum speed and price of cars: in fact, a car with a high maximum speed has, usually, also a high price, and thus a car with a high maximum speed and not so high price is very much appreciated. So, the comprehensive importance of these two criteria is greater than the importance of the two criteria considered separately. To handle the interactions between criteria one can consider *nonadditive integrals*, such as Choquet integral [11] and Sugeno integral [61] (for a comprehensive survey on the use of nonadditive integrals in MCDA, see [21, 25]).

Another interesting decision model permitting representation of interactions between criteria is the Dominance-based Rough Set Approach (DRSA) [29, 59]. In DRSA, the DM's preference model is a set of decision rules, i.e., easily understandable "*if..., then...*" statements, such as "if the maximum speed is at least 200 km/h and the price is not greater than $50,000, then the car is attractive." In general, we shall call the decision models, which, differently from multiple attribute additive utility, permit to represent the interaction between criteria *nonadditive decision models*.

Each decision model requires the specification of some parameters. For example, using MAUT, the parameters are related to the formulation of the marginal value functions $u_i(g_i(a))$, $i = 1, 2, \ldots, n$, while using nonadditive integrals, the parameters are related to so-called fuzzy measures, which permit to model the importance not only of each criterion $g_i \in F$, but also of any subset of criteria $R \subseteq F$. Within MCDA, many methods have been proposed to determine the parameters characterizing the considered decision model in a direct way, i.e., asking them directly to the DM, or in an indirect way, i.e., inducing the values of such parameters from some holistic preference comparisons of alternatives given by the DM. In general, this is a difficult task for several reasons. For example, it is acknowledged that the DM's preference information is often incomplete because the DM is not fully aware of the multiple criteria approach adopted, or because the preference structure is not well defined in DM's mind [43, 62].

Recently, MCDA methods based on indirect preference information and on the *disaggregation approach* [40] are considered more interesting, because they require a relatively smaller cognitive effort from the DM than methods based on direct preference information. In these methods, the DM provides some holistic preference comparisons on a set of reference alternatives A^R, and from this information the parameters of a decision model are induced using a methodology called *ordinal regression*. Then, a consistent decision model is taken into consideration to evaluate the alternatives from set A (*aggregation approach*.) Typically, ordinal regression has been applied to MAUT models, such that in these cases we speak of *additive ordinal regression*. For example, additive ordinal regression is applied by the well-known method called *UTA* [39]. The principle of ordinal regression has also been applied to some nonadditive decision models. In this case, we speak of *nonadditive ordinal regression* exemplified by some *UTA*-like methods involving the Choquet integral [1, 47], and by the DRSA methodology [29, 59].

Usually, from among many sets of parameters of a preference model representing the preference information given by the DM, only one specific set is selected and used to work out a recommendation. For example, while there exist many value functions representing the holistic preference information given by the DM, only one value function is typically used to recommend the best choice, sorting, or ranking of alternatives. Since the selection of one from among many sets of parameters compatible with the preference information given by the DM is rather arbitrary, *robust ordinal regression* proposes taking into account all the sets of parameters compatible with the preference information, in order to give a recommendation in terms of necessary and possible consequences of applying all the compatible preference models on the considered set of alternatives.

The first method of robust ordinal regression is a recent generalization of the *UTA* method, called UTA^{GMS} [34]. The UTA^{GMS} is a multiple criteria method, which, instead of considering only one additive value function *compatible* with the preference information provided by the DM, as *UTA* does, takes into consideration the whole set of compatible additive value functions.

In particular, the UTA^{GMS} method requires from the DM a set of pairwise comparisons on a set of reference alternatives $A^R \subseteq A$ as preference information.

Then, using linear programming, one obtains two relations in set A: the *necessary* weak preference relation, which holds for any two alternatives $a, b \in A$ if and only if all compatible value functions give to a a value greater than the value provided to b, and the *possible* weak preference relation, which holds for this pair if and only if at least one compatible value function gives to a a value greater than the value given to b.

More recently, an extension of UTA^{GMS} has been proposed: the *GRIP* method [18]. The *GRIP* method builds a set of additive value functions, taking into account not only a set of pairwise comparisons of reference alternatives, but also the intensities of preference among reference alternatives.

This kind of preference information is often required in other well-known MCDA methods such as *MACBETH* [6] and *AHP* [54, 55].

Both UTA^{GMS} and *GRIP* apply the robust ordinal regression to the MAUT models, so we can say that these methods apply the *additive robust ordinal regression.*

Finally, *nonadditive robust ordinal regression* has been proposed, applying the basic ideas of robust ordinal regression to a value function expressed as Choquet integral in order to represent positive and negative interactions between criteria. More precisely, the disaggregation–aggregation approach used in this context has been inspired by UTA^{GMS} and *GRIP* methods, but in addition to preference information required by these methods, it includes some preference information on the sign and intensity of interaction between couples of criteria.

The chapter is organized as follows. Section 9.2 is devoted to a presentation of a general scheme of the constructive learning interactive procedure. It provides a brief reminder on learning of one compatible additive piecewise-linear value function for multiple criteria ranking problems using the *UTA* method. In Section 9.3, the *GRIP* method is presented, which is presently the most general of all *UTA*-like methods. Section 9.4, makes a comparison of *GRIP* to its main competitors in the field of MCDA. First *GRIP* is compared to *AHP* method, which requires pairwise comparisons of alternatives and criteria, and yields a priority ranking of solutions. Then *GRIP* is compared to *MACBETH* method, which also takes into account a preference order of alternatives and intensity of preference for pairs of alternatives. The preference information used in *GRIP* does not need, however, to be complete: the DM is asked to provide comparisons of only those ordered pairs of selected alternatives on particular criteria for which his/her judgment is sufficiently certain. This is an important advantage comparing to methods which, instead, require comparison of all possible pairs of alternatives on all the considered criteria. Section 9.5 presents robust ordinal regression applied to sorting problems. Section 9.6 presents the concept of "most representative" value function. Section 9.7 deals with nonadditive robust ordinal regression considering an application of robust ordinal regression methodology to a decision model formulated in terms of Choquet integral. Section 9.8 describes an interactive multiobjective optimization method based on robust ordinal regression. Section 9.9 presents *NEMO*, being an evolutionary multiobjective optimization method based on robust ordinal regression. Section 9.10 shows how robust ordinal regression can deal with outranking methods. Section 9.11 deals with robust ordinal regression applied to multiple criteria group decisions. Section 9.12 presents a didactic example relative to an interactive application of the robust ordinal regression to a multiple objective optimization problem. In Section 9.13, some conclusions and further research directions are provided.

9.2 Ordinal Regression for Multiple Criteria Ranking Problems

The preference information may be either direct or indirect, depending upon whether it specifies directly values of some parameters used in the preference model (e.g., trade-off weights, aspiration levels, discrimination thresholds, etc.) or, whether it specifies some examples of holistic judgments from which compatible values of the preference model parameters are induced. Eliciting direct preference information from the DM can be counterproductive in real-world decision-making situations because of a high cognitive effort required. Consequently, asking directly the DM to provide values for the parameters seems to make the DM uncomfortable. Eliciting indirect preference is less demanding of the cognitive effort. Indirect preference information is mainly used in the ordinal regression paradigm. According to this paradigm, a holistic preference information on a subset of some reference or training alternatives is known first and then a preference model compatible with the information is built and applied to the whole set of alternatives in order to rank them.

The ordinal regression paradigm is concordant with the posterior rationality postulated by March in [46]. It has been known for at least 50 years in the field of multidimensional analysis. It is also concordant with the induction principle used in machine learning. This paradigm has been applied within the two main MCDA approaches mentioned above: those using a value function as preference model [39, 51, 58, 60], and those using an outranking relation as preference model [44, 49, 50]. This paradigm has also been used since mid-1990s in MCDA methods involving a new, third family of preference models – a set of dominance decision rules induced from rough approximations of holistic preference relations [28, 29, 31, 59].

Recently, the ordinal regression paradigm has been revisited with the aim of considering the whole set of value functions compatible with the preference information provided by the DM, instead of a single compatible value function used, for example, in *UTA*-like methods [39, 58]. This extension has been implemented in a method called UTA^{GMS} [34], further generalized in another method called *GRIP* [18]. UTA^{GMS} and *GRIP* are not revealing to the DM only one compatible value function, but they are using the whole set of compatible (general, not piecewise-linear only) additive value functions to set up a necessary weak preference relation and a possible weak preference relation in the whole set of considered alternatives.

9.2.1 Concepts: Definitions and Notation

We are considering an MCDA problem where a finite set of alternatives $A = \{x, \ldots, y, \ldots, w, \ldots, z\}$ ($|A| = m$), is evaluated on a family $F = \{g_1, g_2, \ldots, g_n\}$ of n criteria. Let $I = \{1, 2, \ldots, n\}$ denote the set of criteria indices. We assume, without loss of generality, that the greater $g_i(x)$, the better alternative x on criterion g_i, for all $i \in I$, $x \in A$. A DM is willing to rank the alternatives of A from the best to the worst, according to his/her preferences. The ranking can be complete or partial, depending on the preference information provided by the DM and on the way of exploiting this information. The family of criteria F is supposed to satisfy consistency conditions, i.e., completeness (all relevant criteria are considered), monotonicity (the better the evaluation of an alternative on the considered criteria, the more it is preferable to another), and nonredundancy (no superfluous criteria are considered) [53].

Such a decision-making problem statement is called *multiple criteria ranking problem*. It is known that the only information coming out from the formulation of this problem is the dominance ranking. Let us recall that in the dominance ranking, alternative $x \in A$ is preferred to alternative $y \in A$, $x \succ y$, if and only if $g_i(x) \geq g_i(y)$ for all $i \in I$, with at least one strict inequality. Moreover, x is indifferent to y, $x \sim y$, if and only if $g_i(x) = g_i(y)$ for all $i \in I$. Hence, for any two alternatives $x, y \in A$, one of the four situations may arise in the dominance ranking: $x \succ y$, $y \succ x$, $x \sim y$ and $x?y$, where the last one means that x and y are incomparable. Usually, the dominance ranking is very poor, i.e., the most frequent situation is $x?y$.

In order to enrich the dominance ranking, the DM has to provide preference information, which is used to construct an aggregation model making the alternatives more comparable. Such an aggregation model is called preference model. It induces a preference structure on set A, whose proper exploitation permits to work out a ranking proposed to the DM.

In what follows, the evaluation of each alternative $x \in A$ on each criterion $g_i \in F$ will be denoted either by $g_i(x)$ or x_i. Let G_i denote the value set (scale) of criterion g_i, $i \in I$. Consequently,

$$G = \prod_{i \in I} G_i$$

represents the evaluation space, and $x \in G$ denotes a profile of an alternative in such a space. We consider a weak preference relation $\succsim$ on A which means, for each pair of vectors, $x, y \in G$,

$$x \succsim y \Leftrightarrow \text{“}x \text{ is at least as good as } y\text{”}.$$

This weak preference relation can be decomposed into its asymmetric and symmetric parts, as follows,

(1) $x \succ y \equiv [x \succsim y \text{ and not } y \succsim x] \Leftrightarrow$ “x is preferred to y”, and
(2) $x \sim y \equiv [x \succsim y \text{ and } y \succsim x] \Leftrightarrow$ “x is indifferent to y”.

From a pragmatic point of view, it is reasonable to assume that $G_i \subseteq \mathbb{R}$, for $i = 1, \ldots, n$. More specifically, we will assume that the evaluation scale on each criterion g_i is bounded, such that $G_i = [\alpha_i, \beta_i]$, where α_i, β_i, $\alpha_i < \beta_i$ are the worst and the best (finite) evaluations, respectively. Thus, $g_i : A \rightarrow G_i$, $i \in I$. Therefore, each alternative $x \in A$ is associated with an evaluation vector denoted by $\underline{g}(x) = (x_1, x_2, \ldots, x_n) \in G$.

9.2.2 The UTA Method

In this section, we recall the principle of the ordinal regression *via* linear programming, as proposed in the original *UTA* method, see [39].

9.2.2.1 Preference Information

The preference information is given in the form of a complete preorder on a subset of reference alternatives $A^R \subseteq A$ (where $|A^R| = p$), called *reference preorder*. The reference alternatives are usually those contained in set A for which the DM is able to express holistic preferences. Let $A^R = \{a, b, c, \ldots\}$ be the set of reference alternatives.

9.2.2.2 Additive Model

The additive value function is defined on A such that for each $\underline{g}(x) \in G$,

$$U(\underline{g}(x)) = \sum_{i \in I} u_i(g_i(x_i)), \tag{9.1}$$

where, u_i are nondecreasing marginal value functions, $u_i : G_i \rightarrow \mathbb{R}, i \in I$. For the sake of simplicity, we shall write (1) as follows,

$$U(x) = \sum_{i \in I} u_i(x_i) \quad \text{or} \quad U(x) = \sum_{i=1}^{n} u_i(x_i). \tag{9.2}$$

In the *UTA* method, the marginal value functions u_i are assumed to be piecewise-linear functions. The ranges $[\alpha_i, \beta_i]$ are divided into $\gamma_i \geq 1$ equal sub-intervals $\left[x_i^0, x_i^1\right], \left[x_i^1, x_i^2\right], \ldots, \left[x_i^{\gamma_i - 1}, x_i^{\gamma_i}\right]$, where $x_i^j = \alpha_i + \frac{j}{\gamma_i}(\beta_i - \alpha_i)$, $j = 0, \ldots, \gamma_i$, and $i \in I$. The marginal value of an alternative $x \in A$ is obtained by linear interpolation,

$$u_i(x) = u_i\left(x_i^j\right) + \frac{x_i - x_i^j}{x_i^{j+1} - x_i^j}\left(u_i\left(x_i^{j+1}\right) - u_i\left(x_i^j\right)\right), \; x_i \in \left[x_i^j, x_i^{j+1}\right]. \tag{9.3}$$

The piecewise-linear additive model is completely defined by the marginal values at the breakpoints, i.e., $u_i\left(x_i^0\right) = u_i(\alpha_i)$, $u_i\left(x_i^1\right)$, $u_i\left(x_i^2\right)$, $\cdots$, $u_i\left(x_i^{\gamma_i}\right) = u_i(\beta_i)$.

In what follows, the principle of the *UTA* method is described as it was recently presented in [58]. Therefore, a value function $U(x) = \sum_{i=1}^{n} u_i(x_i)$ is compatible if it satisfies the following set of constraints.

$$\begin{array}{l} \left.\begin{array}{l} U(a) > U(b) \Leftrightarrow \; a \succ b \\ U(a) = U(b) \Leftrightarrow \; a \sim b \end{array}\right\} \quad \forall \; a, b \in A^R \\ u_i\left(x_i^{j+1}\right) - u_i\left(x_i^j\right) \geq 0, \; i = 1, ..., n, \; j = 0, ..., \gamma_i - 1 \\ u_i(\alpha_i) = 0, \; i = 1, ..., n \\ \sum_{i=1}^{n} u_i(\beta_i) = 1 \end{array} \tag{9.4}$$

9.2.2.3 Checking for Compatible Value Functions Through Linear Programming

To verify if a compatible value function $U(x) = \sum_{i=1}^{n} u_i(x_i)$ restoring the reference preorder $\succsim$ on A^R exists, one can solve the following linear programming problem, where $u_i\left(x_i^j\right), i = 1, \ldots, n, \ j = 1, \ldots, \gamma_i$, are unknown, and $\sigma^+(a), \sigma^-(a)$ $(a \in A^R)$ are auxiliary variables:

$$
\begin{array}{l}
\text{Min } Z = \sum_{a \in A^R} \left(\sigma^+(a) + \sigma^-(a)\right) \\
\text{s.t.} \\
\left.\begin{array}{l}
U(a) + \sigma^+(a) - \sigma^-(a) \geq \\
\quad U(b) + \sigma^+(b) - \sigma^-(b) + \varepsilon \ \Leftrightarrow \ a \succ b \\
U(a) + \sigma^+(a) - \sigma^-(a) = \\
\quad U(b) + \sigma^+(b) - \sigma^-(b) \ \Leftrightarrow \ a \sim b
\end{array}\right\} \forall a, b \in A^R \\
u_i\left(x_i^{j+1}\right) - u_i\left(x_i^j\right) \geq 0, \ i = 1, ..., n, \ j = 0, ..., \gamma_i - 1 \\
u_i(\alpha_i) = 0, \ i = 1, ..., n \\
\sum_{i=1}^{n} u_i(\beta_i) = 1 \\
\sigma^+(a), \ \sigma^-(a) \geq 0, \ \forall a \in A^R,
\end{array}
\tag{9.5}
$$

where ε is an arbitrarily small positive value so that $U(a) + \sigma^+(a) - \sigma^-(a) > U(b) + \sigma^+(b) - \sigma^-(b)$ in case of $a \succ b$.

If the optimal value of the objective function of program (9.5) is equal to zero ($Z^* = 0$), then there exists at least one value function $U(x) = \sum_{i=1}^{n} u_i(x_i)$ satisfying (9.4), i.e., compatible with the reference preorder on A^R. In other words, this means that the corresponding polyhedron (9.4) of feasible solutions for $u_i\left(x_i^j\right), \ i = 1, ..., n, \ j = 1, ..., \gamma_i$, is not empty.

Let us remark that the transition from the preorder $\succsim$ to the marginal value function exploits the ordinal character of the criterion scale G_i. Notice, however, that the scale of the marginal value function is a conjoint interval scale. More precisely, for the considered additive value function $\sum_{i=1}^{n} u_i(x_i)$, the admissible transformations on the marginal value functions $u_i(x_i)$ have the form $u_i^*(x_i) = k \times u_i(x_i) + h_i$, $h_i \in \mathbb{R}, \ i = 1, \ldots, n, k > 0$, such that for all $[x_1, ..., x_n], [y_1, ..., y_n] \in \prod_{i=1}^{n} G_i$

$$\sum_{i=1}^{n} u_i(x_i) \geq \sum_{i=1}^{n} u_i(y_i) \Leftrightarrow \sum_{i=1}^{n} u_i^*(x_i) \geq \sum_{i=1}^{n} u_i^*(y_i).$$

An alternative way of representing the same preference model is:

$$U(x) = \sum_{i \in I} w_i \hat{u}_i(x), \tag{9.6}$$

where $\hat{u}(\alpha_i) = 0,\ \hat{u}(\beta_i) = 1,\ w_i \geq 0\ \ \forall i \in I$, and $\sum_{i \in I} w_i = 1$. Note that the correspondence between (9.6) and (2) is such that $w_i = u_i(\beta_i),\ \forall i \in I$. Due to the cardinal character of the marginal value function scale, the parameters w_i can be interpreted as trade-off weights among marginal value functions $\hat{u}_i(x)$. We will use, however, the preference model (2) with normalization constraints bounding $U(x)$ to the interval [0, 1].

When the optimal value of the objective function of the program (9.5) is greater than zero ($Z^* > 0$), then there is no value function $U(x) = \sum_{i \in I} u_i\ (x_i)$ compatible with the reference preorder on A^R. In such a case, three possible moves can be considered:

- Increasing the number of linear pieces γ_i for one or several marginal value function u_i could make it possible to find an additive value function compatible with the reference preorder on A^R.
- Revising the reference preorder on A^R could lead to find an additive value function compatible with the new preorder.
- Searching over the relaxed domain $Z \leq Z^* + \eta$ could lead to an additive value function giving a preorder on A^R sufficiently close to the reference preorder (in the sense of Kendall's τ).

9.3 Robust Ordinal Regression for Multiple Criteria Ranking Problems

Recently, two new methods, *UTA*^{GMS} [34] and *GRIP* [18], have generalized the ordinal regression approach of the *UTA* method in several aspects:

- Taking into account all additive value functions (1) compatible with the preference information, while *UTA* is using only one such function.
- Considering marginal value functions of (1) as general nondecreasing functions, and not piecewise-linear, as in *UTA*.
- Asking the DM for a ranking of reference alternatives, which is not necessarily complete (just pairwise comparisons).
- Taking into account additional preference information about intensity of preference, expressed both comprehensively and with respect to a single criterion.
- Avoiding the use of the exogenous, and not neutral for the result, parameter ε in the modeling of strict preference between alternatives.

UTA^{GMS} and *GRIP* produce two rankings on the set of alternatives A, such that for any pair of alternatives $a, b \in A$:

- In the *necessary* ranking, a is ranked at least as good as b if and only if, $U(a) \geq U(b)$ for *all* value functions compatible with the preference information.
- In the *possible* ranking, a is ranked at least as good as b if and only if, $U(a) \geq U(b)$ for *at least one* value function compatible with the preference information.

The necessary ranking can be considered as *robust* with respect to the preference information. Such robustness of the necessary ranking refers to the fact that any pair of alternatives compares in the same way whatever the additive value function compatible with the preference information. Indeed, when no preference information is given, the necessary ranking boils down to the dominance relation, and the possible ranking is a complete relation. It allows for taking into account the incomparability between alternatives. Every new pairwise comparison of reference alternatives, for which the dominance relation does not hold, is enriching the necessary ranking and it is impoverishing the possible ranking, so that they converge with the growth of the preference information.

Moreover, such an approach gives space for interactivity with the DM. Presentation of the necessary ranking, resulting from a preference information provided by the DM, is a good support for generating reactions from part of the DM. Namely, (s)he could wish to enrich the ranking or to contradict a part of it. Such a reaction can be integrated in the preference information considered in the next iteration.

The idea of considering the whole set of compatible value functions was originally introduced in UTA^{GMS}. *GRIP* (*G*eneralized *R*egression with *I*ntensities of *P*reference) can be seen as an extension of UTA^{GMS} permitting to take into account additional preference information in the form of comparisons of intensities of preference between some pairs of reference alternatives. For alternatives $x, y, w, z \in A$, these comparisons are expressed in two possible ways (not exclusive): (i) comprehensively, on all criteria, like "x is preferred to y at least as much as w is preferred to z"; and, (ii) partially, on any criterion, like "x is preferred to y at least as much as w is preferred to z, on criterion $g_i \in F$". Although UTA^{GMS} was historically the first method among the two, as *GRIP* incorporates and extends UTA^{GMS}, in the following we shall present only *GRIP*.

9.3.1 The Preference Information Provided by the Decision Maker

The DM is expected to provide the following preference information:

- A partial preorder $\succsim$ on A^R whose meaning is: for $x, y \in A^R$

  $$x \succsim y \Leftrightarrow x \text{ is at least as good as } y.$$

  Moreover, $\succ$ (preference) is the asymmetric part of $\succsim$ and $\sim$ (indifference) is the symmetric part given by $\succsim \cap \succsim^{-1}$. ($\succsim^{-1}$ is the inverse of $\succsim$, i.e., for all $x, y \in A^R$, $x \succsim^{-1} y \Leftrightarrow y \succsim x$).
- A partial preorder $\succsim^*$ on $A^R \times A^R$, whose meaning is: for $x, y, w, z \in A^R$,

  $$(x, y) \succsim^* (w, z) \Leftrightarrow x \text{ is preferred to } y \text{ at least as much as } w \text{ is preferred to } z.$$

Also in this case, $\succ^*$ is the asymmetric part of $\succsim^*$ and $\sim^*$ is the symmetric part given by $\succsim^* \cap \succsim^{*^{-1}}$ ($\succsim^{*^{-1}}$ is the inverse of $\succsim^*$, i.e., for all $x, y, w, z \in A^R$, $(x, y) \succsim^{*^{-1}} (w, z) \Leftrightarrow (w, z) \succsim^* (x, y)$).

- A partial preorder $\succsim_i^*$ on $A^R \times A^R$, whose meaning is: for $x, y, w, z \in A^R$, $(x, y) \succsim_i^* (w, z) \Leftrightarrow x$ is preferred to y at least as much as w is preferred to z on criterion g_i, $i \in I$.

In the following, we also consider the weak preference relation $\succsim_i$ being a complete preorder whose meaning is: for all $x, y \in A$,

$$x \succsim_i y \quad \Leftrightarrow \quad x \text{ is at least as good as } y \text{ on criterion } g_i, \quad i \in I.$$

Weak preference relations $\succsim_i$, $i \in I$, are not provided by the DM, but it is obtained directly from the evaluation of alternatives x and y on criterion g_i, i.e., $x \succsim_i y \Leftrightarrow g_i(x) \geq g_i(y)$.

9.3.2 Possible and Necessary Rankings

While the preference information provided by the DM is rather similar to that of *UTA*, the output of *GRIP* is quite different. In *GRIP*, the preference information has the form of a partial preorder in a set of reference alternatives $A^R \subseteq A$ (i.e., a set of pairwise comparisons of reference alternatives), augmented by information about intensities of preferences.

A value function is called *compatible* if it is able to restore the partial preorder $\succsim$ on A^R, as well as the given relation of intensity of preference among ordered pairs of reference alternatives. Each compatible value function induces, moreover, a complete preorder on the whole set A. In particular, for any two alternatives $x, y \in A$, a compatible value function orders x and y in one of the following ways: $x \succ y$, $y \succ x$, $x \sim y$. With respect to $x, y \in A$, it is thus reasonable to ask the following two questions:

- Are x and y ordered in the same way by *all* compatible value functions?
- Is there *at least one* compatible value function ordering x at least as good as y (or y at least as good as x)?

Having answers to these questions for all pairs of alternatives $(x, y) \in A \times A$, one gets a *necessary weak preference relation* $\succsim^N$ (partial preorder), whose semantics is $U(x) \geq U(y)$ for all compatible value functions, and a *possible weak preference relation* $\succsim^P$ in A (strongly complete and negatively transitive relation), whose semantics is $U(x) \geq U(y)$ for at least one compatible value function.

Let us remark that preference relations $\succsim^N$ and $\succsim^P$ are meaningful only if there exists at least one compatible value function. Observe also that in this case, for any $x, y \in A^R$,

$$x \succsim y \Rightarrow x \succsim^N y$$

and

$$x \succ y \Rightarrow \text{ not}\left(y \succsim^P x\right).$$

In fact, if $x \succsim y$, then for any compatible value function, $U(x) \geq U(y)$ and, therefore, $x \succsim^N y$. Moreover, if $x \succ y$, then for any compatible value function, $U(x) > U(y)$ and, consequently, there is no compatible value function such that $U(y) \geq U(x)$, which means not($y \succsim^P x$).

9.3.3 Linear Programming Constraints

In this section, we present a set of constraints that interprets the preference information in terms of conditions on the compatible value functions.

The value function $U : A \to [0, 1]$ should satisfy the following constraints corresponding to DM's preference information,

(a) $U(w) > U(z)$ if $w \succ z$
(b) $U(w) = U(z)$ if $w \sim z$
(c) $U(w) - U(z) > U(x) - U(y)$ if $(w, z) \succ^* (x, y)$
(d) $U(w) - U(z) = U(x) - U(y)$ if $(w, z) \sim^* (x, y)$
(e) $u_i(w) \geq u_i(z)$ if $w \succsim_i z, i \in I$
(f) $u_i(w) - u_i(z) > u_i(x) - u_i(y)$ if $(w, z) \succ_i^* (x, y), i \in I$
(g) $u_i(w) - u_i(z) = u_i(x) - u_i(y)$ if $(w, z) \sim_i^* (x, y), i \in I$

Let us remark that within *UTA*-like methods, constraint (a) is written as $U(w) \geq U(z) + \varepsilon$, where $\varepsilon > 0$ is a threshold exogenously introduced. Analogously, constraints (c) and (f) should be written as,

$$U(w) - U(z) \geq U(x) - U(y) + \varepsilon$$

and

$$u_i(w) - u_i(z) \geq u_i(x) - u_i(y) + \varepsilon.$$

However, we would like to avoid the use of any exogenous parameter and, therefore, instead of setting an arbitrary value of ε, we consider it as an auxiliary variable, and we test the feasibility of constraints (a), (c), and (f) (see Section 9.3.4). This permits to take into account all possible value functions, even those which satisfy the constraints for having a very small threshold ε. This is safer also from the viewpoint of "objectivity" of the selected methodology. In fact, the value of ε is not meaningful in itself and it is useful only because it permits to discriminate preference from indifference.

Moreover, the following normalization constraints should also be taken into account:

(h) $u_i\left(x_i^*\right) = 0$, where x_i^* is such that $x_i^* = \min\{g_i(x) : x \in A\}$
(i) $\sum_{i \in I} u_i\left(y_i^*\right) = 1$, where y_i^* is such that $y_i^* = \max\{g_i(x) : x \in A\}$

If the constraints from (a) to (i) are fulfilled, then the partial preorders $\succsim$ and $\succsim^*$ on A^R and $A^R \times A^R$ can be extended on A and $A \times A$, respectively.

9.3.4 Computational Issues

In order to conclude the truth or falsity of binary relations $\succsim^N$, $\succsim^P$, $\succsim^{*N}$, $\succsim^{*P}$, $\succsim_i^{*N}$ and $\succsim_i^{*P}$, we have to take into account that, for all $x, y, w, z \in A$ and $i \in I$:

(1) $x \succsim^N y \Leftrightarrow \inf\Big\{U(x) - U(y)\Big\} \geq 0$
(2) $x \succsim^P y \Leftrightarrow \inf\Big\{U(y) - U(x)\Big\} \leq 0$
(3) $(x, y) \succsim^{*N} (w, z) \Leftrightarrow \inf\Big\{\Big(U(x) - U(y)\Big) - \Big(U(w) - U(z)\Big)\Big\} \geq 0$
(4) $(x, y) \succsim^{*P} (w, z) \Leftrightarrow \inf\Big\{\Big(U(w) - U(z)\Big) - \Big(U(x) - U(y)\Big)\Big\} \leq 0$
(5) $(x, y) \succsim_i^{*N} (w, z) \Leftrightarrow \inf\Big\{\Big(u_i(x_i) - u_i(y_i)\Big) - \Big(u_i(w_i) - u_i(z_i)\Big)\Big\} \geq 0$
(6) $(x, y) \succsim_i^{*P} (w, z) \Leftrightarrow \inf\Big\{\Big(u_i(w_i) - u_i(z_i)\Big) - \Big(u_i(x_i) - u_i(y_i)\Big)\Big\} \leq 0$

with the infimum computed on the set of value functions satisfying constraints from (a) to (i). Let us remark, however, that the linear programming is not able to handle strict inequalities such as the above (a), (c), and (f). Moreover, linear programming permits to compute the minimum or the maximum of an objective function and not an infimum. Nevertheless, reformulating properly the above properties (1) to (6), a result presented in [47] permits to use linear programming for testing the truth of binary relations, $\succsim^N$, $\succsim^P$, $\succsim^{*N}$, $\succsim^{*P}$, $\succsim_i^{*N}$ and $\succsim_i^{*P}$.

In order to use such a result, constraints (a), (c) and (f) have to be reformulated as follows:

(a') $U(x) \geq U(y) + \varepsilon$ if $x \succ y$
(c') $U(x) - U(y) \geq U(w) - U(z) + \varepsilon$ if $(x, y) \succ^* (w, z)$
(f') $u_i(x) - u_i(y) \geq u_i(w) - u_i(z) + \varepsilon$ if $(x, y) \succ_i^* (w, z)$

with $\varepsilon > 0$.

Then, properties $(1) - (6)$ have to be reformulated such that the search of the infimum is replaced by computing the maximum value of ε on the set of value functions satisfying constraints from (a) to (i), with constraints (a), (c) and (f) transformed to (a'), (c') and (f'), plus constraints specific for each point:

$(1')$ $x \succsim^P y \Leftrightarrow \varepsilon^* > 0$,
where $\varepsilon^* = \max\ \varepsilon$, subject to the constraints (a'), (b), (c'), (d), (e), (f'), plus the constraint $U(x) \geq U(y)$
$(2')$ $x \succsim^N y \Leftrightarrow \varepsilon^* \leq 0$,
where $\varepsilon^* = \max\ \varepsilon$, subject to the constraints (a'), (b), (c'), (d), (e), (f'), plus the constraint $U(y) \geq U(x) + \varepsilon$
$(3')$ $(x, y) \succsim^{*P} (w, z) \Leftrightarrow \varepsilon^* > 0$,
where $\varepsilon^* = \max\ \varepsilon$, subject to the constraints (a'), (b), (c'), (d), (e), (f'), plus the constraint $\Big(U(x) - U(y)\Big) - \Big(U(w) - U(z)\Big) \geq 0$

(4′) $(x, y) \succsim^{*N} (w, z) \Leftrightarrow \varepsilon^* \leq 0$,
where $\varepsilon^* = \max \ \varepsilon$, subject to the constraints (a'), (b), (c'), (d), (e), (f'), plus the constraint $\Big(U(w) - U(z)\Big) - \Big(U(x) - U(y)\Big) \geq \varepsilon$

(5′) $(x, y) \succsim_i^{*P} (w, z) \Leftrightarrow \varepsilon^* > 0$,
where $\varepsilon^* = \max \ \varepsilon$, subject to the constraints (a'), (b), (c'), (d), (e), (f'), plus the constraint $\Big(u_i(x_i) - u_i(y_i)\Big) - \Big(u_i(w_i) - u_i(z_i)\Big) \geq 0$

(6′) $(x, y) \succsim_i^{*N} (w, z) \Leftrightarrow \varepsilon^* \leq 0$,
where $\varepsilon^* = \max \ \varepsilon$, subject to the constraints (a'), (b), (c'), (d), (e), (f'), plus the constraint $\Big(u_i(w_i) - u_i(z_i)\Big) - \Big(u_i(x_i) - u_i(y_i)\Big) \geq \varepsilon$.

9.4 Comparison of *GRIP* with other MCDA Methods

9.4.1 Comparison of GRIP with the AHP

In *AHP* (*A*nalytical *H*ierarchy *P*rocess) [54, 55], criteria should be pairwise compared with respect to their importance. Alternatives are also pairwise compared on particular criteria with respect to intensity of preference. The following nine point scale is used:

1 – Equal importance-preference
3 – Moderate importance-preference
5 – Strong importance-preference
7 – Very strong or demonstrated importance-preference
9 – Extreme importance-preference

2, 4, 6, and 8 are intermediate values between the two adjacent judgements. The ratio of importance of criterion g_i over criterion g_j is the inverse of the ratio of importance of g_j over g_i. Analogously, the intensity of preference of alternative x over alternative y is the inverse of the intensity of preference of y over x. The above scale is a ratio scale. Therefore, the difference of importance is read as the ratio of weights w_i and w_j, corresponding to criteria g_i and g_j, and the intensity of preference is read as the ratio of the attractiveness of x and the attractiveness of y, with respect to the considered criterion g_i. In terms of value functions, the intensity of preference can be interpreted as the ratio $\frac{u_i(g_i(x))}{u_i(g_i(y))}$. Thus, the problem is how to obtain values of w_i and w_j from ratio $\frac{w_i}{w_j}$, and values of $u_i(g_i(x))$ and $u_i(g_i(y))$ from ratio $\frac{u_i(g_i(x))}{u_i(g_i(y))}$.

In *AHP* it is proposed that these values are supplied by the principal eigenvectors of the matrices composed of the ratios $\frac{w_i}{w_j}$ and $\frac{u_i(g_i(x))}{u_i(g_i(y))}$. The marginal value functions $u_i(g_i(x))$ are then aggregated by means of a weighted-sum using the weights w_i.

Comparing *AHP* with *GRIP*, we can say that with respect to single criteria the type of questions addressed to the DM is the same: express intensity of preference in qualitative-ordinal terms (equal, moderate, strong, very strong, extreme). However, differently from *GRIP*, this intensity of preference is translated in *AHP* into quantitative terms (the scale from 1 to 9) in a quite arbitrary way. In *GRIP*, instead, the marginal value functions are just a numerical representation of the original qualitative-ordinal information, and no intermediate transformation in quantitative terms is exogenously imposed.

Other differences between *AHP* and *GRIP* are related to the following aspects.

1. In *GRIP*, the value functions $u_i(g_i(x))$ depend mainly on comprehensive preferences involving jointly all the criteria, while this is not the case in *AHP*.
2. In *AHP*, the weights w_i of criteria g_i are calculated on the basis of pairwise comparisons of criteria with respect to their importance; in *GRIP*, this is not the case, because the value functions $u_i(g_i(x))$ are expressed on the same scale and thus they can be summed up without any further weighting.
3. In *AHP*, all unordered pairs of alternatives must be compared from the viewpoint of the intensity of preference with respect to each particular criterion. Therefore, if m is the number of alternatives, and n the number of criteria, then the DM has to answer $n \times \frac{m \times (m-1)}{2}$ questions. Moreover, the DM has to answer questions relative to $\frac{n \times (n-1)}{2}$ pairwise comparisons of considered criteria with respect to their importance. This is not the case in *GRIP*, which accepts partial information about preferences in terms of pairwise comparison of some reference alternatives. Finally, in *GRIP* there is no question about comparison of relative importance of criteria.

As far as point 2 is concerned, observe that the weights w_i used in *AHP* represent trade-offs between evaluations on different criteria. For this reason it is doubtful that if they could be inferred from answers to questions concerning comparison of importance. Therefore, *AHP* has a problem with meaningfulness of its output with respect to its input, and this is not the case of *GRIP*.

9.4.2 Comparison of GRIP with MACBETH

MACBETH (*M*easuring *A*ttractiveness by a *C*ategorical *B*ased *E*valuation *T*ec*H*nique) is a method for MCDA [5,6], which builds a value function from qualitative judgements obtained from DMs about differences of values quantifying the relative attractiveness of alternatives or criteria.

When using *MACBETH*, the DM is asked to provide the following preference information about every two alternatives from set A:

- First, through an (ordinal) judgement on their relative attractiveness.
- Second, (if the two alternatives are not considered to be equally attractive), through a qualitative judgement about the difference of attractiveness between these two alternatives.

Seven semantic categories of difference of attractiveness are considered in *MACBETH*: null, very weak, weak, moderate, strong, very strong, extreme.

The main idea of *MACBETH* is to build an interval scale from the preference information provided by the DM. It is, however, necessary that the above categories correspond to disjoint intervals (represented in terms of the real numbers). The bounds for such intervals should not be arbitrarily fixed a priori, but they should be calculated simultaneously with the numerical values of all particular alternatives from A, so as to ensure compatibility between these values [5]. Linear programming models are used for these calculations. In case of inconsistent judgments, *MACBETH* provides the DM with information permitting to eliminate such inconsistency.

When comparing *MACBETH* with *GRIP* the following aspects should be considered:

- Both deal with qualitative judgements.
- Both need a set of comparisons of alternatives or pairs of alternatives to work out a numerical representation of preferences, however, *MACBETH* depends on the definition of two characteristic levels on the original scale, "neutral" and "good," to obtain the numerical representation of preferences.
- *GRIP* adopts the disaggregation–aggregation approach and, therefore, it considers also comprehensive preferences relative to comparisons involving jointly all the criteria, which is not the case of *MACBETH*.
- *GRIP* is, however, more general than *MACBETH* since it can take into account the same kind of qualitative judgments as *MACBETH* (the difference of attractiveness between pairs of alternatives) and the intensity of preferences of the type "x is preferred to y at least as much as z is preferred to w".

As for the last item, it should be noticed that the intensity of preference considered in *MACBETH* and the intensity coming from comparisons of the type "x is preferred to y at least as strongly as w is preferred to z" (i.e., the quaternary relation $\succsim^*$) are substantially the same. In fact, the intensities of preference are equivalence classes of the preorder generated by $\succsim^*$. This means that all the pairs (x, y) and (w, z), such that x is preferred to y with the same intensity as w is preferred to z, belong to the same semantic category of difference of attractiveness considered in *MACBETH*. To be more precise, the structure of intensity of preference considered in *MACBETH* is a particular case of the structure of intensity of preference represented by $\succsim^*$ in *GRIP*. Still more precisely, *GRIP* has the same structure of intensity as *MACBETH* when $\succsim^*$ is a complete preorder. When this does not occur, *MACBETH* cannot be used while *GRIP* can naturally deal with this situation.

Comparison of *GRIP* and *MACBETH* could be summarized in the following points:

1. *GRIP* is using preference information relative to: (a) comprehensive preference on a subset of reference alternatives with respect to all criteria, (b) partial intensity of preference on some single criteria, and (c) comprehensive intensity of preference with respect to all criteria, while *MACBETH* requires preference information on all pairs of alternatives with respect to each one of the considered criteria.

2. Information about partial intensity of preference is of the same nature in *GRIP* and *MACBETH* (equivalence classes of relation $\succsim_i^*$ correspond to qualitative judgements of *MACBETH*), but in *GRIP* it may not be complete.
3. *GRIP* is a "disaggregation–aggregation" approach while *MACBETH* makes use of the "aggregation" approach and, therefore, it needs weights to aggregate evaluations on the criteria.
4. *GRIP* works with all compatible value functions, while *MACBETH* builds a single interval scale for each criterion, even if many such scales would be compatible with preference information.

9.5 Robust Ordinal Regression for Multiple Criteria Sorting Problems

Robust ordinal regression has been proposed also for sorting problems [32, 35, 45]. In the following, we present the new *UTADIS*^{GMS} method [32, 35]. *UTADIS*^{GMS} considers an additive value function

$$U(a) = \sum_{i=1}^{n} u_i(g_i(a))$$

as a preference model ($a \in A$). Let us remember that sorting procedures consider a set of p predefined preference ordered classes $C_1, C_2, \ldots, C_p$, where $C_{h+1} \gg C_h$ ($\gg$ a complete order on the set of classes), $h = 1, \ldots, p-1$. The aim of a sorting procedure is to assign each alternative to one class or to a set of contiguous classes. The robust ordinal regression uses a value function U to decide the assignments in such a way that if $U(a) > U(b)$, then a is assigned to a class not worse than b.

We suppose the DM provides preference information in form of possibly imprecise assignment examples on a reference set A^*, i.e., for all $a^* \in A^*$ the DM defines a desired assignment $a^* \rightarrow [C_{L^{DM}(a^*)}, C_{R^{DM}(a^*)}]$, where $[C_{L^{DM}(a^*)}, C_{R^{DM}(a^*)}]$ is an interval of contiguous classes $C_{L^{DM}(a^*)}$, $C_{L^{DM}(a^*)+1}$, ..., $C_{R^{DM}(a^*)}$. Each such alternative is called a reference alternative. $A^* \subseteq A$ is called the set of reference alternatives. An assignment example is said to be precise if $L^{DM}(a^*) = R^{DM}(a^*)$, and imprecise, otherwise.

Given a value function U, a set of assignment examples is said to be *consistent with U* iff

$$\forall a^*, b^* \in A^*, \quad U(a^*) \geq U(b^*) \Rightarrow R^{DM}(a^*) \geq L^{DM}(b^*) \tag{9.7}$$

which is equivalent to

$$\forall a^*, b^* \in A^*, \quad L^{DM}(a^*) > R^{DM}(b^*) \Rightarrow U(a^*) > U(b^*) \tag{9.8}$$

On the basis of (9.8), we can state that, formally, a general additive compatible value function is an additive value function $U(a) = \sum_{i=1}^{n} u_i(a)$ satisfying the following set of constraints:

$$\left.\begin{array}{l} U(a^*) > U(b^*) \;\Leftrightarrow\; L^{DM}(a^*) > R^{DM}(b^*) \quad \forall a^*, b^* \in A^* \\ u_i(g_i(a_{\tau_i(j)})) - u_i(g_i(a_{\tau_i(j-1)})) \geq 0,\; i = 1,...,n,\; j = 2,...,m \\ u_i(g_i(a_{\tau_i(1)})) \geq 0, u_i(g_i(a_{\tau_i(m)})) \leq u_i(\beta_i),\; i = 1,...,n, \\ u_i(\alpha_i) = 0,\; i = 1,...,n \\ \sum_{i=1}^{n} u_i(\beta_i) = 1, \end{array}\right\} (E^{A^*})$$

where α_i and β_i are, respectively, the worst and the best evaluations on each criterion g_i, and τ_i is the permutation on the set of indices of alternatives from A^* that reorders them according to the increasing evaluation on criterion g_i, i.e.,

$$g_i(a_{\tau_i(1)}) \leq g_i(a_{\tau_i(2)}) \leq \ldots \leq g_i(a_{\tau_i(m-1)}) \leq g_i(a_{\tau_i(m)}).$$

Let us observe that the set of constraints (E^{A^*}) is equivalent to

$$\left.\begin{array}{l} U(a^*) \geq U(b^*) + \varepsilon \Leftrightarrow L^{DM}(a^*) > R^{DM}(b^*) \quad \forall a^*, b^* \in A^* \\ u_i(g_i(a_{\tau_i(j)})) - u_i(g_i(a_{\tau_i(j-1)})) \geq 0,\; i = 1,...,n,\; j = 2,...,m \\ u_i(g_i(a_{\tau_i(1)})) \geq 0, u_i(g_i(a_{\tau_i(m)})) \leq u_i(\beta_i),\; i = 1,...,n, \\ u_i(\alpha_i) = 0,\; i = 1,...,n \\ \sum_{i=1}^{n} u_i(\beta_i) = 1, \end{array}\right\} (E^{A^*})'$$

with $\varepsilon > 0$. Thus, to verify that the set of all compatible value functions $\mathcal{U}_{\mathcal{A}^*}$ is not empty, it is sufficient to verify that $\varepsilon^* > 0$, where $\varepsilon^* = \max\ \varepsilon$, subject to set of constraints $(E^{A^*})'$.

Taking into account a single value function $U \in \mathcal{U}_{\mathcal{A}^*}$ and its associated assignment examples relative to the reference set A^*, an alternative $a \in A$ can be assigned to an interval of classes $[C_{L^U(a)}, C_{R^U(a)}]$, in the following way:

$$L^U(a) = Max\left(\{1\} \cup \left\{L^{DM}(a^*) : U(a^*) \leq U(a),\; a^* \in A^R\right\}\right), \tag{9.9}$$

$$R^U(a) = Min\left(\{p\} \cup \left\{R^{DM}(a^*) : U(a^*) \geq U(a),\; a^* \in A^R\right\}\right). \tag{9.10}$$

For each nonreference alternative $a \in A \setminus A^*$ the indices satisfy the following condition:

$$L^U(a) \leq R^U(a). \tag{9.11}$$

In order to take into account the whole set of value functions one can proceed as follows. Given a set A^* of assignment examples and a corresponding set $\mathcal{U}_{\mathcal{A}^*}$ of compatible value functions, for each $a \in A$, we define the possible assignment $C_P(a)$ as the set of indices of classes C_h for which there exist at least one value function $U \in \mathcal{U}_{\mathcal{A}^*}$ assigning a to C_h, and the necessary assignment $C_N(a)$ as set of indices of classes C_h for which all value functions $U \in \mathcal{U}$ assign a to C_h, that is:

$$C_P(a) = \left\{h \in H : \exists U \in \mathcal{U}_{\mathcal{A}^*} \text{ for which } h \in \left[L^U(a), R^U(a)\right]\right\} \tag{9.12}$$

$$C_N(a) = \left\{h \in H : \forall U \in \mathcal{U}_{\mathcal{A}^*} \text{ it holds } h \in \left[L^U(a), R^U(a)\right]\right\} \tag{9.13}$$

To compute the possible and necessary assignments $C_P(a)$ and $C_N(a)$, we can consider the following indices:

- minimum possible class:

$$L_P^{\mathcal{U}}(a) = Max\left(\{1\} \cup \left\{L^{DM}(a^*) : \forall U \in \mathcal{U}_{\mathcal{A}^*}, U(a^*) \leq U(a),\ a^* \in A^*\right\}\right) \tag{9.14}$$

- minimum necessary class:

$$L_N^{\mathcal{U}}(a) = Max\left(\{1\} \cup \left\{L^{DM}(a^*) : \exists U \in \mathcal{U}_{\mathcal{A}^*} \text{ for which } U(a^*) \leq U(a),\ a^* \in A^*\right\}\right) \tag{9.15}$$

- maximum necessary class:

$$R_N^{\mathcal{U}}(a) = Min\left(\{p\} \cup \left\{R^{DM}(a^*) : \exists U \in \mathcal{U}_{\mathcal{A}^*} \text{ for which } U(a) \leq U(a^*),\ a^* \in A^*\right\}\right) \tag{9.16}$$

- maximum possible class:

$$R_P^{\mathcal{U}}(a) = Min\left(\{p\} \cup \left\{R^{DM}(a^*) : \forall U \in \mathcal{U}_{\mathcal{A}^*}, U(a) \leq U(a^*),\ a^* \in A^*\right\}\right) \tag{9.17}$$

Using indices $L_P^{\mathcal{U}}(a)$, $L_N^{\mathcal{U}}(a)$, $R_N^{\mathcal{U}}(a)$ and $R_P^{\mathcal{U}}(a)$, the possible and necessary assignments $C_P(a)$ and $C_N(a)$ can be expressed as follows:

$$C_P(a) = \left[L_P^{\mathcal{U}}(a), R_P^{\mathcal{U}}(a)\right]$$

and, if $L_N^{\mathcal{U}}(a) \leq R_N^{\mathcal{U}}(a)$, then

$$C_N(a) = \left[L_N^{\mathcal{U}}(a), R_N^{\mathcal{U}}(a)\right]$$

while, if $L_N^{\mathcal{U}}(a) > R_N^{\mathcal{U}}(a)$, then

$$C_N(a) = \emptyset.$$

As in the methods *UTA*^{GMS} and *GRIP*, on the basis of all compatible value functions $\mathcal{U}_{\mathcal{A}^*}$, we can define two binary relations on the set of alternatives A:

- *Necessary* weak preference relation $\succsim^N$, in case $U(a) \geq U(b)$ for all compatible value functions
- *Possible* weak preference relation $\succsim^P$, in case $U(a) \geq U(b)$ for at least one compatible value function

Using necessary weak preference relation $\succsim^N$ and possible weak preference relation $\succsim^P$ we can redefine indices $L_P^{\mathcal{U}}(a)$, $L_N^{\mathcal{U}}(a)$, $R_N^{\mathcal{U}}(a)$ and $R_P^{\mathcal{U}}(a)$ as follows:

- minimum possible class:

$$L_P^{\mathcal{U}}(a) = Max\left(\{1\} \cup \left\{L^{DM}(a^*) : a \succsim^N a^*,\ a^* \in A^*\right\}\right), \tag{9.18}$$

- minimum necessary class:

$$L_N^{\mathcal{U}}(a) = Max\left(\{1\} \cup \left\{L^{DM}(a^*) : a \succsim^P a^*,\ a^* \in A^*\right\}\right), \qquad (9.19)$$

- maximum necessary class:

$$R_N^{\mathcal{U}}(a) = Min\left(\{p\} \cup \left\{R^{DM}(a^*) : a^* \succsim^P a,\ a^* \in A^*\right\}\right), \qquad (9.20)$$

- maximum possible class:

$$R_P^{\mathcal{U}}(a) = Min\left(\{p\} \cup \left\{R^{DM}(a^*) : a^* \succsim^N a,\ a^* \in A^*\right\}\right). \qquad (9.21)$$

Thus, using necessary weak preference relation $\succsim^N$ and possible weak preference relation $\succsim^P$, it is possible to deal quite simply with the sorting problem.

Therefore, on the basis of the above observations, the following example-based sorting procedure can be proposed:

1. Ask the DM for an exemplary sorting.
2. Verify that the set of compatible value functions $\mathcal{U}_{A^*}$ is not empty.
3. Calculate the necessary and the possible weak preference relations $a \succsim^N a^*$, $a \succsim^P a^*$, $a^* \succsim^N a$ and $a^* \succsim^P a$, with $a^* \in A^*$ and $a \in A$.
4. Calculate for each $a \in A$ the indices $L_P^{\mathcal{U}}(a)$, $L_N^{\mathcal{U}}(a)$, $R_N^{\mathcal{U}}(a)$ and $R_P^{\mathcal{U}}(a)$ using (9.18), (9.19), (9.20) and (9.21).
5. Assign to each $a \in A$ its possible assignment $C_P(a) = \left[L_P^{\mathcal{U}}(a), R_P^{\mathcal{U}}(a)\right]$.
6. Assign to each $a \in A$ its necessary assignment, which is $C_N(a) = \left[L_N^{\mathcal{U}}(a), R_N^{\mathcal{U}}(a)\right]$ in case $L_N^{\mathcal{U}}(a) \leq R_N^{\mathcal{U}}(a)$, and $C_N(a) = \emptyset$ otherwise.

In [16], one can find a proposal how to handle within *UTADIS*^{GMS} an additional preference information about intensity of preference.

9.6 The Most Representative Value Function

The robust ordinal regression builds a set of additive value functions compatible with preference information provided by the DM and results in two rankings, necessary and possible. Such rankings answer to robustness concerns, since they are in general "more robust" than a ranking made by an arbitrarily chosen compatible value function. However, in practice, for some decision-making situations, a score is needed to assign to the different alternatives, and despite the interest of the rankings provided, some users would like to see, and they indeed need to know, the "most representative" value function among all the compatible ones. This allows assigning a score to each alternative.

Recently, a methodology to identify the "most representative" value function in *GRIP*, without loosing the advantage of taking into account all compatible value functions, has been proposed in [17]. The idea is to select among compatible value functions that one which better highlights the necessary ranking maximizing the difference of values between alternatives for which there is a preference in the necessary ranking. As secondary objective, one can consider minimizing the difference of values between alternatives for which there is no preference in the necessary ranking. This comprehensive "most representative" value function can be determined via the following procedure:

1. Determine the necessary and the possible rankings in the considered set of alternatives.
2. For all pairs of alternatives (a, b), such that a is necessarily preferred to b, add the following constraints to the linear programming constraints of *GRIP*: $U(a) \geq U(b) + \varepsilon$.
3. Maximize the objective function ε.
4. Add the constraint $\varepsilon = \varepsilon^*$, with $\varepsilon^* = \max \varepsilon$ from the previous point, to the linear programming constraints of robust ordinal regression.
5. For all pairs of alternatives (a, b), such that neither a is necessarily preferred to b nor b is necessarily preferred to a, add the following constraints to the linear programming constraints of *GRIP* and to the constraints considered in above point 4): $U(a) - U(b) \leq \delta$ and $U(b) - U(a) \leq \delta$.
6. Minimize the objective function δ.

This procedure maximizes the minimal difference between values of alternatives for which the necessary preference holds. If there is more than one such value function, the above procedure selects the most representative compatible value function giving the greatest minimal difference between values of alternatives for which the necessary preference holds, and the smallest maximal difference between values of alternatives for which the possible preference holds.

Notice that the concept of the "most representative" value function thus defined is still based on the necessary and possible preference relations, which remain crucial for *GRIP*, and, in a sense, it gives the most faithful representation of this necessary and possible preference relations.

In [27] the concept of the "most representative" value function has been extended to robust ordinal regression applied to sorting problems within *UTADIS* ^{GMS}.

The idea is to select among all compatible value functions that one which better highlights the possible sorting considered as the most stable part of the robust sorting obtained by *UTADIS*^{GMS}. In consequence, the selected value function is that one which maximizes the difference of values between alternatives for which the intervals of possible sorting are disjoint. As secondary objective, to tie-breaking, one can wish to maximize the minimal difference between values of alternatives a and b such that for any compatible value function U a is assigned to a class not worse than the class of b and for at least one compatible value function a is assigned to a class which is better than the class of b. In case there is still more than one such value function, the "most representative" function minimizes the maximal difference between values of alternatives a and b being in the same class for all compatible value functions U or such that the order of classes is not univocal in the sense that for some compatible value functions U a is assigned to a class better than b and for other compatible value function b is assigned to a class better than a.

The following three-stage procedure for determining the most representative value function can be proposed:

1. Determine the possible sorting $C_P(a)$ and the necessary sorting $C_N(a)$ for each considered alternative $a \in A$.
2. For all pairs of alternatives (a, b), such that $L_P^{\mathcal{U}}(a) > R_P^{\mathcal{U}}(b)$, add the following constraint to the linear programming constraints of *UTADIS*^{GMS}, E^{A^R}:

$$U(a) \geq U(b) + \varepsilon.$$

3. Maximize the objective function ε subject to the set of linear constraints from point 2.
4. Add the constraint $\varepsilon = \varepsilon^*$, with $\varepsilon^* = \max \varepsilon$ from the previous point, to the linear programming constraints of *UTADIS*^{GMS}, E^{A^R}.
5. For all pairs of alternatives (a, b), such that for any compatible value function U a is assigned to a class not worse than the class of b and for at least one compatible value function a is assigned to a class which is better than the class of b, add the following constraint to the linear programming constraints from point 4:

$$U(a) \geq U(b) + \gamma.$$

6. Maximize the objective function γ subject to the set of linear constraints from point 5.
7. Add the constraint $\gamma = \gamma^*$, with $\gamma^* = \max \gamma$ from the previous point, to the linear programming constraints from point 5.
8. For all pairs of alternatives (a, b), such that they are in the same class for all compatible value functions U, or such that the order of classes is not univocal, add the following constraints to the linear programming constraints from point 7:

$$U(a) - U(b) \leq \delta \text{ and } U(b) - U(a) \leq \delta.$$

9. Minimize the objective function δ subject to the set of linear constraints from point 8.

Notice that the concept of the "most representative" value function thus defined is based on the possible assignments and supplies the most faithful representation of the recommendation given by *UTADIS*^{GMS}. Therefore, it can play a significant role in supporting the DM to understand the results of the robust sorting. Moreover, the most representative value function U^R chosen according to the above principles, can be used along with the assignment examples supplied at the beginning by the DM to drive an autonomous example-based sorting procedure. In such a way the most representative assignment for each alternative $a \in A$ can be determined.

9.7 Nonadditive Robust Ordinal Regression

To take into account interactions between criteria, robust ordinal regression has been applied to Choquet integral [2].

Let 2^G be the power set of G (i.e., the set of all the subsets of the set of criteria G); a fuzzy measure on G is defined as a set function $\mu : 2^G \rightarrow [0, 1]$ which satisfies the following properties:

(1a) $\mu(\emptyset) = 0$ and $\mu(G) = 1$ (boundary conditions)
(2a) $\forall\, T \subseteq R \subseteq G,\ \ \mu(T) \leq \mu(R)$ (monotonicity condition)

In the framework of multiple criteria decision problems, a fuzzy measure $\mu(R)$ is related to the importance weight given by the DM to every subset of criteria R that can be evaluated by the Shapley value [56], defined later in this section.

Let $x \in A$ and μ be a fuzzy measure on G, then the *Choquet integral* [11] is defined by:

$$C_\mu(x) = \sum_{i=1}^{n} \left[\left(g_{(i)}(x) \right) - \left(g_{(i-1)}\,(x) \right) \right] \mu\,(A_i), \tag{9.22}$$

where $_{(\cdot)}$ stands for a permutation of the indices of evaluations of criteria such that:

$$g_{(1)}\,(x) \leq g_{(2)}\,(x) \leq g_{(3)}\,(x) \leq ... \leq g_{(n)}\,(x),$$

with $A_i = \{(i), \ldots, (n)\}$, $i = 1, \ldots, n$, and $g_{(0)} = 0$.

One of the main drawbacks of the Choquet integral is the necessity to elicit and give an adequate interpretation of $2^{|G|} - 2$ parameters. In order to reduce the number of parameters to be computed and to eliminate a too strict description of the interactions among criteria, which is not realistic in many applications, one can consider the concept of fuzzy k-additive measure [22].

Given a partial preorder $\succeq$ on A^R, a set of fuzzy measures $\boldsymbol{\mu}$ is called compatible if the Choquet integral, calculated with respect to it, restores the DM's ranking on A^R, i.e.,

$$a \succeq b \;\;\Leftrightarrow\;\; C_{\boldsymbol{\mu}}(a) \geq C_{\boldsymbol{\mu}}(b) \;\;\; \forall a, b \in A^R.$$

The procedure proposed is composed of three successive phases:

(I) Elicitation of preference information on a reference set $A^R \subseteq A$ of alternatives
(II) Evaluation of all the *compatible* fuzzy measures to establish the preference relations $a \succeq^P b$ and $a \succeq^N b$ for every ordered pair of alternatives $(a, b) \in A \times A$
(III) Exploitation of the results obtained to detect possible DM's inconsistencies or to revise the preference model obtained

In the phase of elicitation of preference information, the DM is asked to provide the following preference information:

(a) A partial preorder $\succeq$ on A^R, i.e., for $a, b \in A^R$:

$$a \succeq b \;\Leftrightarrow\; a \text{ is at least as good as } b.$$

(b) A partial preorder $\succeq^*$ on $A^R \times A^R$, i.e., for $a, b, c, d \in A^R$

$$(a, b) \succeq^* (c, d) \Leftrightarrow a \text{ is preferred to } b \text{ at least as much as } c \text{ is preferred to } d.$$

(c) A partial preorder $\triangleright$ on G, for $i, j \in G$, whose definition is:

$$i \triangleright j \;\Leftrightarrow\; \text{criterion } i \text{ is more important than criterion } j.$$

(d) A partial preorder $\triangleright^*$ on $G \times G$, whose definition is: for $i, j, l, k \in G$ $(i, j) \triangleright^* (l, k) \;\Leftrightarrow$ the difference of importance between criteria i and j is at least as much as difference of importance between criteria l and k.

(e) A sign (positive or negative) of interaction of couples of criteria.

(f) A partial preorder $\triangleright_{\text{Int}}$ on $G \times G$, whose definition is: for $i, j, l, k \in G$,

$$(i, j) \triangleright_{\text{Int}} (l, k) \;\Leftrightarrow$$

intensity of interaction between criteria i and j is at least as strong as intensity of interaction between criteria l and k.

(g) A partial preorder $\triangleright^*_{\text{Int}}$ on G^4, whose definition is: for $i, j, l, k, r, s, t, w \in G$,

$$[(i, j), (l, k)] \triangleright^*_{\text{Int}} [(r, s), (t, w)] \;\Leftrightarrow$$

difference of intensity of interaction between criteria i and j, and intensity of interaction between criteria l and k is at least as strong as difference of intensity of interaction between criteria r and s, and intensity of interaction between criteria t and w. In this phase, the DM compares the intensity of interaction for pairs of criteria, both redundant or synergic.

The preference information of type (b), (d), (f) and (g) can be provided by the DM using a semantic scale in a similar way to the approaches of *MACBETH* [6], *AHP* [54] and *GRIP* [18]. More precisely, given an ordinal scale such as "null," "small," "medium," "large," and "extreme," the DM can give information of the type: "the preference of alternative a over alternative b is large" or "the difference of importance between criteria g_i and g_j is medium" or "the synergy between criteria g_i and g_j is small".

In Phase II, the set of all *compatible* fuzzy measures is determined as those fuzzy measures satisfying a system of linear constraints representing all the preference information given by the DM in Phase I, plus the monotonicity and boundary conditions of fuzzy measures.

In Phase III, the obtained preference model, i.e., the system of linear constraints determining the set of all compatible fuzzy measures, is used to determine the *necessary* preference relation $\succsim^N$ and the *possible* preference relation $\succsim^P$ on A, as follows:

$$x \succsim^N y \;\Leftrightarrow\; C_\mu(x) \geq C_\mu(y)$$

for all compatible sets of fuzzy measures $\boldsymbol{\mu}$, with $x, y \in A$, and

$$a \succsim^P b \;\Leftrightarrow\; C_\mu(x) \geq C_\mu(y)$$

for at least one compatible set of fuzzy measures $\boldsymbol{\mu}$, with $x, y \in A$.

In [3], nonadditive robust ordinal regression has been proposed to deal with sorting problems. In simple words, the methodology follows the principle of *UTADIS*^{GMS}, but considering the Choquet integral, instead of an additive value function. In [4], nonadditive robust ordinal regression has been extended in turn to deal with some generalizations of Choquet integral, such as bipolar Choquet integral [23, 24, 30] and the level dependent Choquet integral [26].

9.8 Robust Ordinal Regression in Interactive Multiobjective Optimization

Classical ordinal regression methods have been applied in Multiobjective Optimization (MOO) in [38] and in [57], where an additive value function interactively built using the *UTA* method is optimized within the feasible region. In the same spirit, robust ordinal regression has been applied to MOO problems in [15], as explained below. We assume that the Pareto optimal set of an MOO problem is generated prior to an interactive exploration of this set. Instead of the whole and exact Pareto optimal set of a MOO problem, one can also consider a proper representation of this set, or its approximation. In any case, an interactive exploration of this set should lead the DM to a conviction that either there is no satisfactory solution to the considered problem, or there is at least one such solution. We will focus our attention on the interactive exploration, and the proposed interactive procedure will be valid for any finite set of solutions to be explored. Let us denote this set by A. Note that such set A can be computed using evolutionary multiobjective optimization. For a recent state of the art of interactive and evolutionary approaches to MOO, see [8].

In the course of the interactive procedure, the preference information provided by the DM concerns a small subset of A, called reference or training sample, and denoted by A^R. The preference information is transformed by an ordinal regression method into a DM's preference model. We propose to use at this stage the *GRIP* method, thus the preference model is a set of general additive value functions compatible with the preference information. A compatible value function compares the solutions from the reference sample in the same way as the DM. The obtained preference model is then applied on the whole set A, which results in possible and necessary rankings of solutions. These rankings are used to select a new sample of reference solutions, which is presented to the DM, and the procedure cycles until a satisfactory solution is selected from the sample or the DM comes to conclusion that there is no satisfactory solution for the current problem setting.

The proposed interactive procedure is composed of the following steps:

- ***Step* 1.** Select a representative reference sample A^R of solutions from set A.
- ***Step* 2.** Present the sample A^R to the DM.
- ***Step* 3.** If the DM is satisfied with at least one solution from the sample, then this is the satisfactory solution and the procedure stops. The procedure also stops in this step if the DM concludes that there is no satisfactory solution for the current problem setting. Otherwise continue.
- ***Step* 4.** Ask the DM to provide information about his/her preferences on set A^R in the following terms:
  - Pairwise comparison of some solutions from A^R
  - Comparison of intensities of comprehensive preferences between some pairs of solutions from A^R
  - Comparison of intensities of preferences on single criteria between some pairs of solutions from A^R
- ***Step* 5.** Use the *GRIP* method to build a set of additive monotonically nondecreasing value functions compatible with the preference information obtained from the DM in *Step* 4.
- ***Step* 6.** Apply the set of compatible value functions built in *Step* 5 on the whole set A, and present the possible and necessary rankings (see Section 9.4.2) resulting from this application to the DM.
- ***Step* 7.** Taking into account the possible and necessary rankings, let the DM select a new reference sample of solutions $A^R \subseteq A$, and go to *Step* 2.

In *Step* 4, the information provided by the DM may lead to a set of constraints, which define an empty polyhedron of the compatible value functions. In this case, the DM gets information about which items of his/her preference information make the polyhedron empty, so as to enable revision in the next round. This point is explained in detail in [18, 34]. Moreover, information provided by the DM in *Step* 4 cannot be considered as irreversible. Indeed, the DM can retract to one of previous iterations and continue from this point. This feature is concordant with the spirit of a learning oriented conception of multiobjective interactive optimization, i.e., it confirms the idea that the interactive procedure permits the DM to learn about his/her preferences and about the "shape" of the Pareto optimal set (see [7]).

Notice that the proposed approach allows to elicit incrementally preference information from the DM. In *Step* 7, the "new" reference sample A^R is not necessarily different from the previously considered, however, the preference information elicited from the DM in the next iteration is richer than previously, due to the learning effect. This permits to build and refine progressively the preference model: in fact, each new item of information provided in *Step* 4 restricts the set of compatible value functions and defines the DM's preferences more and more precisely.

Let us also observe that information obtained from the DM in *Step* 4 and information given to the DM in *Step* 6 is composed of very simple and easy to understand statements: preference comparisons in *Step* 4, and possible and necessary rankings in *Step* 6 (i.e., a necessary ranking that holds for all compatible value functions, and a possible ranking that holds for at least one compatible value function; see Section 9.4.2). Thus, the nature of information exchanged with the DM during the interaction is purely ordinal. Indeed, monotonically increasing transformations of evaluation scales of considered criteria have no influence on the final result.

Finally, observe that a very important characteristic of our method from the point of view of learning is that the DM can observe the impact of information provided in *Step* 4 in terms of possible and necessary rankings of solutions from set A.

9.9 Robust Ordinal Regression in Evolutionary Interactive Multiobjective Optimization

Most of the research in evolutionary multiobjective optimization (EMO) attempts to approximate the complete Pareto optimal front by a set of well-distributed representatives of Pareto optimal solutions. The underlying reasoning is that in the absence of any preference information, all Pareto optimal solutions have to be considered equivalent.

On the other hand, in most practical applications, the DM is eventually interested in only a single solution. In order to come up with a single solution, it is necessary to involve the DM. This is the underlying idea of another multiobjective optimization paradigm: interactive multiobjective optimization (IMO). IMO deals with the identification of the most preferred solution by means of a systematic dialogue with the DM. Only recently, the scientific community has discovered the great potential of combining the two paradigms (for a recent survey, see [41]). From the point of view of EMO, involving the DM in an interactive manner will allow to focus the search on the area of the Pareto front which is most relevant to the DM. This, in turn, may allow to find more appropriate solutions faster. In particular, in the case of many objectives, EMO has difficulties, because the number of Pareto-optimal solutions becomes huge, and Pareto-optimality is not sufficiently discriminative to guide the search into better regions. Integrating user preferences promises to alleviate these problems, allowing to converge faster to the preferred region of the Pareto-optimal front.

Robust ordinal regression has been applied to EMO in a methodology called *NEMO* (*N*ecessary preference-based *E*volutionary *M*ultiobjective *O*ptimization) presented in [9, 10]. *NEMO* combines *NSGA*-II [12], a widely used EMO technique, with the IMO methodology based on robust ordinal regression presented in Section 9.4. The *NEMO* methodology takes into account the information about necessary preferences, given by the robust ordinal regression, in order to focus the search on the most promising parts of the Pareto optimal front. More specifically, robust ordinal regression based on information obtained through interaction with the DM determines the set of compatible value functions, and an EMO procedure searches for all nondominated solutions taking into account all compatible value functions in parallel.

We believe that the integration of robust ordinal regression into EMO is particularly promising for two reasons:

1. The preference information required by robust ordinal regression is very basic and easy to provide by the DM. All that the DM is asked for is to compare two nondominated solutions, and to reveal whether one is preferred over the other.
2. The resulting set of compatible value functions reveals implicitly an appropriate scaling of the criteria, an issue that is largely ignored by the EMO community so far.

A crucial step in *NSGA*-II, is the ranking of solutions (individuals) in a current population according to two criteria.

The primary criterion is the so-called dominance-based ranking. This criterion ranks individuals by iteratively determining the nondominated solutions in the population (nondominated front), assigning those individuals the next best rank, and removing them from the population. The result is a partial ordering, favoring individuals closer to the Pareto optimal front.

According to the secondary criterion, individuals which have the same dominance-rank (primary criterion) are sorted with respect to the crowding distance, which is defined as the sum of distances between a solution and its neighbors on either side in each dimension of the objective space. Individuals with a large crowding distance are preferred, as they are in a less crowded region of the objective space, which is concordant with the goal of preserving diversity in the population.

NEMO combines the robust ordinal regression with *NSGA*-II in three different variants:

- *NEMO*-0: a single compatible value function is used to rank solutions in a population. For example, one can consider the value function obtained by the *UTA* method.
- *NEMO*-I: the whole set of compatible value functions is considered and the dominance relation used in *NSGA*-II to rank solutions is replaced by the necessary preference relation of robust ordinal regression.
- *NEMO*-II: the whole set of compatible value functions is also considered, but differently from *NEMO*-I, the solutions in the population are ranked according to a score calculated as the max–min difference of values between a given solution and other solutions in the population, for the whole set of compatible value functions.

In *NEMO*-0, *NEMO*-I, and *NEMO*-II, the following types of value functions are considered:

- Linear value function, i.e.,

$$U(g(a)) = \lambda_1(g_1(a)) + \lambda_2(g_2(a)) + \ldots + \lambda_n(g_n(a)) \quad \text{with} \quad a \in A,$$

  $\lambda_1 \geq 0, \lambda_2 \geq 0, \ldots, \lambda_n \geq 0, \quad \lambda_1 + \lambda_2 + \ldots + \lambda_n = 1$
- Piecewise-linear value function, as in the *UTA* method (see Section 9.3)

- General additive value function, as in the *UTA*^{GMS} and *GRIP* methods (see Section 9.4)

Observe that the *NEMO*-I variant with a linear value function corresponds to the method proposed in [37].

In the following, we present *NEMO*-I with a general additive value function, which was the first variant proposed in [9]. The modifications of *NEMO*-I with respect to *NSGA*-II are the following:

1. *NEMO*-I replaces the dominance-based ranking procedure by the necessary ranking procedure. The necessary ranking procedure works analogously to the dominance-based ranking procedure, but taking into account the preference information by the DM through the necessary preference relations. More precisely, the procedure first puts in the best rank all solutions, which are not preferred by any other solution in the population, then removes them from the population and creates the second best rank composed of solutions, which are not preferred by any other solution in the reduced population, and so on.
2. *NEMO*-I replaces the crowding-distance by a distance calculated in the space of marginal values, taking into account the multidimensional scaling given by the "the most representative" value function among the whole set of compatible value functions (see Section 9.7). More precisely, the crowding distance is calculated according to the procedure used in *NSGA*-II with the only difference that in calculating the average side-length of the cuboid the distance is measured in terms of marginal values of the "most representative" value function.

Preferences are elicited by asking the DM to compare pairs of nondominated solutions, and specify a preference relation between them.

The overall *NEMO*-I algorithm is outlined in Algorithm 1. Although the general procedure is rather straightforward, there are several issues that need to be considered:

Algorithm 1: Basic *NEMO*-I

Generate initial solutions randomly
Elicit DM's preferences {Present to the DM a pair of nondominated solutions and ask for a preference comparison}
Determine *necessary* ranking {Replaces dominance ranking in *NSGA*-II}
Determine *secondary* ranking {Order solutions within the same rank, based on the crowding distance measured in terms of the "most representative value function"}
repeat
    Mating selection and offspring generation
    if Time to ask the DM **then**
        Elicit DM's preferences
    end if
    Determine *necessary* ranking
    Determine *secondary* ranking
    Environmental selection
until Stopping criterion met
Return all preferred solutions according to necessary ranking

1. How many pairs of solutions are shown to the DM, and when? In [9], one pairwise comparison of nondominated solutions was asked every k generations, i.e., every k generations, *NEMO*-I is stopped, and the user is asked to provide preference information about one given pair of individuals. Preliminary experiments show that $k = 20$ in 300 generation runs gives satisfactory results.
2. Which pairs of solutions should be presented to the DM for comparison? In [9], each pair of solutions was picked randomly from among the best solutions not related by the necessary preference relation, i.e., from solutions having the best rank. This avoids that the DM can specify inconsistent information, inverting the necessary preference relation (including dominance) between two solutions. To speed up convergence, it would be reasonable, however, to pick pairs of solutions having the best rank and being close with respect to the overall value but diversified on respective marginal values, for "the most representative" value function.

An important remark about the *NEMO* methodology regards its approximation power. In fact, *NSGA*-II can identify all nondominated solutions, even improper ones, i.e., nondominated points that allow unbounded trade-off between objective functions [20], in problems where the nondominated frontier has discontinuities or it is nonconvex. From this point of view, *NEMO* methodology maintains this good property. More precisely, considering linear value functions in *NEMO*-0 or *NEMO*-II, one cannot deal with improper solutions and discontinuous or nonconvex frontier, because there can be no linear value function giving the best value to some efficient solutions. *NEMO*-I can find, however, all nondominated points because it compares pairs of solutions and, therefore, there can be linear compatible value functions for which the considered nondominated solution, possibly improper, is preferred to other nondominated solutions, even in case of discontinuities of nonconvexity. Using a general additive value function in *NEMO*-0, *NEMO*-I, or *NEMO*-II, improper efficient points, discontinuous or nonconvex nondominated frontiers can be dealt without any difficulty. To explain this ability, remark that:

(a) The class of value functions, which can be expressed as additive value functions is very large, including, for instance, value functions of the form

$$U(g(a)) = u_1(g_1(a))^{\lambda_1} \times u_2(g_2(a))^{\lambda_2} \times \ldots \times u_n(g_n(a))^{\lambda_n}$$

with $a \in A, \lambda_1 \geq 0, \lambda_2 \geq 0, \ldots, \lambda_n \geq 0$, whose logarithm takes the form

$$U^*[g(a)] = log[U(g(a))] = \lambda_1 \times log[u_1(g_1(a))] + \lambda_2 \times log[u_2(g_2(a))] \\ + \ldots + \lambda_n \times log[u_n(g_n(a))].$$

(b) Marginal value functions $u_i(g_i(a)), i = 1, \ldots, n$, can be constant in some parts of their domains

Remark (a) explains why *NEMO*-I and *NEMO*-II are able to deal with discontinuous and nonconvex nondominated frontiers, while remark (b) explains why *NEMO*-I and *NEMO*-II are able to deal with improper points.

Piecewise linear value functions have a behavior, which is intermediate between the linear value functions and general additive value functions. One can say, in general, that the greater the number of linear pieces assumed for each marginal value function, the more similar the final results are to the case of general additive value functions. This means that increasing the number of linear pieces, one improves the capacity of dealing with improper solutions, discontinuities and nonconvexities. However, the more flexible the value function model, the more preference information, and thus more interactions with the DM, is required to focus the search on the most preferred region of the Pareto optimal front.

9.10 Robust Ordinal Regression for Outranking Methods

Outranking relation is a noncompensatory preference model used in the *ELECTRE* family of MCDA methods [52]. Its construction involves two concepts known as concordance and discordance. Outranking relation, usually denoted by S, is a binary relation on a set A of alternatives. For an ordered pair of alternatives $(a, b) \in A$, aSb means "a is at least as good as b." The assertion aSb is considered to be true if the coalition of criteria being in favor of this statement is "strong enough" comparing to the rest of criteria, and if among the criteria opposing to this statement, there is no one for which a is "significantly worse" than b. The first condition is called concordance test, and the second, non-discordance test.

Let us denote by k_i the weight assigned to criterion g_i, $i = 1, \ldots, n$; it represents a relative importance of criterion g_i within family F of n criteria. The indifference, preference and veto thresholds on criterion g_i are denoted by q_i, p_i and v_i, respectively. For consistency, $v_i > p_i > q_i \geq 0$, $i = 1, \ldots, n$. In all formulae that follow, we suppose, without loss of generality, that all these thresholds are constant, that preferences are increasing with evaluations on particular criteria, and that criteria are identified by their indices.

The concordance test involves calculation of concordance index $C(a, b)$. It represents the strength of the coalition of criteria being in favor of aSb. This coalition is composed of two subsets of criteria:

- Subset of criteria being clearly in favor of aSb, i.e., such that $g_i(a) \geq g_i(b) - q_i$.
- Subset of criteria that do not oppose to aSb, while being in an ambiguous position with respect to this assertion; these are those criteria for which a weak preference relation bQa holds; i.e., such that $g_i(b) - p_i \leq g_i(a) < g_i(b) - q_i$.

Consequently, the concordance index is defined as

$$C(a, b) = \frac{\sum_{i=1}^{n} \phi_i(a, b) \times k_i}{\sum_{i=1}^{n} k_i}, \tag{9.23}$$

where, for $i = 1, \ldots, n$,

$$\phi_i(a,b) = \begin{cases} 1, & \text{if } g_i(a) \geq g_i(b) - q_i, \\ \dfrac{g_i(a) - [g_i(b) - p_i]}{p_i - q_i}, & \text{if } g_i(b) - p_i \leq g_i(a) < g_i(b) - q_i, \\ 0, & \text{if } g_i(a) < g_i(b) - p_i. \end{cases} \tag{9.24}$$

$\phi_i(a,b)$ is a marginal concordance index, indicating to what extent criterion g_i contributes to the concordance index $C(a,b)$. As defined by (9.24), $\phi_i(a,b)$ is a piecewise linear function, nondecreasing with respect to $g_i(a) - g_i(b)$.

Remark that $C(a,b) \in [0,1]$, where $C(a,b) = 0$ if $g_i(a) \leq g_i(b) - p_i$, $i = 1, \ldots, n$ (b is strictly preferred to a on all criteria), and $C(a,b) = 1$ if $g_i(a) \geq g_i(b) - q_i$, $i = 1, \ldots, n$ (a outranks b on all criteria).

The result of the concordance test for a pair $(a,b) \in A$ is positive if $C(a,b) \geq \lambda$, where $\lambda \in [0.5, 1]$ is a cutting level, which has to be fixed by the DM.

Once the result of the concordance test has been positive, one can pass to the non-discordance test. Its result is positive for the pair $(a,b) \in A$ unless "a is significantly worse than b" on at least one criterion, i.e., if $g_i(b) - g_i(a) < v_i$ for $i = 1, \ldots, n$.

It follows from above that the outranking relation for a pair $(a,b) \in A$ is true, and denoted by aSb if both the concordance test and the non-discordance test are positive. On the other hand, the outranking relation for a pair $(a,b) \in A$ is false, and denoted by aS^cb, either if the concordance test or the non-discordance test is negative.

Knowing S or S^c for all ordered pairs $(a,b) \in A$, one can proceed to exploitation of the outranking relation in set A, which is specific for the choice, or sorting or ranking problem, as described in [19].

Experience indicates that elicitation of preference information necessary for construction of the outranking relation is not an easy task for a DM. In particular, the inter-criteria preference information concerning the weights of criteria and the veto thresholds are difficult to be expressed directly.

For this reason, some disaggregation–aggregation procedures have been proposed in the past to assist the elicitation of the weights of criteria and all the thresholds required to construct the outranking relation [48–50]. The most general proposal, however, has been presented in [33,36]. It permits to asses the whole set of outranking relations compatible with some exemplary pairwise comparisons of few real or fictitious reference alternatives, using a robust ordinal regression approach. Below, we briefly sketch this proposal.

We assume that the preference information provided by the DM is a set of pairwise comparisons of some reference alternatives. The set of reference alternatives is denoted by A^R, and it is usually, although not necessarily, a subset of set A. The comparison of a pair of alternatives $(a,b) \in A^R$ states the truth or falsity of the outranking relation, denoted by aSb or aS^cb, respectively. It is worth stressing that the DM does not need to provide all pairwise comparisons of reference alternatives, so this comparison can be confined to a small subset of pairs.

We also assume that the intra-criterion preference information concerning indifference and preference thresholds $p_i > q_i \geq 0$, $i = 1, \ldots, n$, is given. The last assumption is not unrealistic because these thresholds are relatively easy to provide by an analyst who is usually aware what is the precision of criteria, and how much difference is nonsignificant or relevant.

In order to simplify calculations of the ordinal regression, we assume that the weights of criteria sum up to one, i.e., $\sum_{i=1}^{n} k_i = 1$. Thus, (9.23) becomes

$$C(a,b) = \sum_{i=1}^{n} \phi_i(a,b) \times k_i = \sum_{i=1}^{n} \psi_i(a,b), \tag{9.25}$$

where the marginal concordance index $\psi_i(a,b) = \phi_i(a,b) \times k_i$ is a monotone nondecreasing function with respect to $g_i(a) - g_i(b)$, such that $\psi_i(a,b) \geq 0$ for all $(a,b) \in A^R \times A^R$, $i = 1, \ldots, n$, $\psi_i(a,b) = 0$ for all $g_i(b) - g_i(a) \geq p_i$, $i = 1, \ldots, n$, and $\sum_{i=1}^{n} \psi_i(a,b) = 1$ in case $g_i(a) - g_i(b) \geq -q_i$ for all $i = 1, \ldots, n$.

The ordinal regression constraints defining the set of concordance indices $C(a,b)$, cutting levels λ and veto thresholds v_i, $i = 1, \ldots, n$, compatible with the pairwise comparisons provided by the DM have the following form:

$$\left.\begin{array}{l}
C(a,b) = \sum_{i=1}^{n} \psi_i(a,b) \geq \lambda \text{ and } g_i(b) - g_i(a) \leq v_i - \varepsilon, i = 1, \ldots, n, \\
\quad \text{if } aSb, \text{ for } (a,b) \in A^R \times A^R, \\
C(a,b) = \sum_{i=1}^{n} \psi_i(a,b) \leq \lambda - \varepsilon + M_0(a,b) \text{ and } g_i(b) - g_i(a) \leq v_i - \delta M_i(a,b), \\
M_i(a,b) \in \{0,1\}, \sum_{i=0}^{n} M_i(a,b) \leq n, \ i = 1, \ldots, n, \\
\quad \text{if } aS^c b, \text{ for } (a,b) \in A^R \times A^R, \\
1 \geq \lambda \geq 0.5, \quad v_i \geq p_i, \ i = 1, \ldots, n, \\
\psi_i(a,b) \geq 0, \text{ for all } (a,b) \in A^R \times A^R, \ i = 1, \ldots, n, \\
\psi_i(a,b) = 0 \text{ if } g_i(b) - g_i(a) \geq p_i, \text{ for all } (a,b) \in A^R \times A^R, \ i = 1, \ldots, n, \\
\sum_{i=1}^{n} \psi_i(a,b) = 1 \text{ if } g_i(a) - g_i(b) \geq -q_i \text{ for all } (a,b) \in A^R \times A^R, \ i = 1, \ldots, n, \\
\psi_i(a,b) \geq \psi_i(c,d) \text{ if } g_i(a) - g_i(b) \geq g_i(c) - g_i(d), \\
\quad \text{for all } a,b,c,d \in A^R, \ i = 1, \ldots, n,
\end{array}\right\} E(A^R)$$

where ε is a small positive value and δ is a big positive value. Remark that $E(A^R)$ are constraints of a 0-1 mixed linear program.

Given a pair of alternatives $(x,y) \in A$, x *necessarily* outranks y, which is denoted by $xS^N y$, if and only if $d(x,y) \geq 0$, where

$$d(x,y) = Min \left\{ \sum_{i=1}^{n} \psi_i(x,y) - \lambda \right\},$$

subject to constraints $E(A^R)$, plus constraints $\psi_i(x,y) \geq 0$, $\psi_i(x,y) = 0$ if $g_i(y) - g_i(x) \geq p_i$, $\psi_i(a,b) \geq \psi_i(c,d)$ if $g_i(a) - g_i(b) \geq g_i(c) - g_i(d)$, for all $a,b,c,d \in A^R \cup \{x,y\}$, $i = 1,\ldots,n$, $\sum_{i=1}^{n} \psi_i(x,y) = 1$ if $g_i(x) - g_i(y) \geq -q_i$ for all $i = 1,\ldots,n$, and $g_i(y) - g_i(x) \leq v_i$, $i = 1,\ldots,n$.

$d(x,y) \geq 0$ means that for all compatible outranking models x outranks y. Obviously, for all $(x,y) \in A^R$, $xSy \Rightarrow xS^N y$.

Analogously, given a pair of alternatives $(x,y) \in A$, x *possibly* outranks y, which is denoted by $xS^P y$, if and only if $D(x,y) \geq 0$, where

$$D(x,y) = Max\left\{\sum_{i=1}^{n} \psi_i(x,y) - \lambda\right\},$$

subject to constraints $E(A^R)$, plus constraints $\psi_i(x,y) \geq 0$, $\psi_i(x,y) = 0$ if $g_i(y) - g_i(x) \geq p_i$, $\psi_i(a,b) \geq \psi_i(c,d)$ if $g_i(a) - g_i(b) \geq g_i(c) - g_i(d)$, for all $a,b,c,d \in A^R \cup \{x,y\}$, $i = 1,\ldots,n$, $\sum_{i=1}^{n} \psi_i(x,y) = 1$ if $g_i(x) - g_i(y) \geq -q_i$ for all $i = 1,\ldots,n$, and $g_i(y) - g_i(x) \leq v_i$, $i = 1,\ldots,n$.

$D(x,y) \geq 0$ means that for at least one compatible outranking model x outranks y.

Moreover, for any pair of alternatives $(x,y) \in A$:

$$xS^N y \Leftrightarrow not(xS^{cP} y) \quad \text{and} \quad xS^P y \Leftrightarrow not(xS^{cN} y)$$

so, only $xS^N y$ and $xS^P y$ are to be checked.

The necessary and the possible outranking relations are to be exploited as usual outranking relations in the context of choice, sorting, and ranking problems.

9.11 Robust Ordinal Regression for Multiple Criteria Group Decisions

The robust ordinal regression can be adapted to the case of group decisions [36]. In this case, several DMs cooperate in a decision problem to make a collective decision. DMs share the same "description" of the decision problem (the same set of alternatives, family of criteria and performance matrix). Each DM provides his/her own preference information, composed of pairwise comparisons of some reference alternatives. The collective preference model accounts for the preference expressed by each DM.

Let us denote the set of DMs by $\mathcal{D} = \{d_1,\ldots,d_p\}$.

In case of ranking and choice problems, for each DM $d_h \in \mathcal{D}' \subseteq \mathcal{D}$, we consider all compatible value functions. Four situations are interesting for a pair $(a,b) \in A$:

- $a \succsim_{\mathcal{D}'}^{N,N} b$: $a \succsim^N b$ for *all* $d_h \in \mathcal{D}'$
- $a \succsim_{\mathcal{D}'}^{N,P} b$: $a \succsim^N b$ for *at least one* $d_h \in \mathcal{D}'$

- $a \succsim_{\mathcal{D}'}^{P,N} b : a \succsim^{P} b$ for *all* $d_h \in \mathcal{D}'$
- $a \succsim_{\mathcal{D}'}^{P,P} b : a \succsim^{P} b$ for *at least one* $d_h \in \mathcal{D}'$

In case of sorting problems, for each DM $d_r \in \mathcal{D}' \subseteq \mathcal{D}$, we consider the set of all compatible value functions $\mathcal{U}_{A^R}^{d_r}$. Given a set A^R of assignment examples, for each $a \in A$ and for each DM $d_r \in \mathcal{D}'$, we define his/her possible and necessary assignments as

$$C_P^{d_r}(a) = \left\{h \in H : \exists U \in \mathcal{U}_{A^R}^{d_r} \text{ assigning } a \text{ to } C_h\right\}, \tag{9.26}$$

$$C_N^{d_r}(a) = \left\{h \in H : \forall U \in \mathcal{U}_{A^R}^{d_r} \text{ assigning } a \text{ to } C_h\right\}. \tag{9.27}$$

Moreover, for each subset of DMs $D' \subseteq \mathcal{D}$, we define the following assignments:

$$C_{P,P}^{D'}(a) = \bigcup_{d_r \in \mathcal{D}'} C_P^{d_r}(a), \tag{9.28}$$

$$C_{N,P}^{D'}(a) = \bigcup_{d_r \in \mathcal{D}'} C_N^{d_r}(a), \tag{9.29}$$

$$C_{P,N}^{D'}(a) = \bigcap_{d_r \in \mathcal{D}'} C_P^{d_r}(a), \tag{9.30}$$

$$C_{N,N}^{D'}(a) = \bigcap_{d_r \in \mathcal{D}'} C_N^{d_r}(a). \tag{9.31}$$

Possible and necessary assignments $C_P^{d_r}(a)$ and $C_N^{d_r}(a)$ are calculated for each decision maker $d_r \in \mathcal{D}$ using *UTADIS*^{GMS}, and then the four assignments $C_{P,P}^{D'}(a)$, $C_{N,P}^{D'}(a)$, $C_{P,N}^{D'}(a)$ and $C_{P,P}^{D'}(a)$ can be calculated for all subsets of decision makers $\mathcal{D}' \subseteq \mathcal{D}$.

In case of application of robust ordinal regression to outranking methods, for each DM $d_h \in D' \subseteq D$, we consider all compatible outranking models. Four situations are interesting for a pair $(x, y) \in A$:

- $x\ S_{D'}^{N,N}\ y : xS^N y$ for all $d_h \in D'$
- $x\ S_{D'}^{N,P}\ y : xS^N y$ for at least one $d_h \in D'$
- $x\ S_{D'}^{P,N}\ y : xS^P y$ for all $d_h \in D'$
- $x\ S_{D'}^{P,P}\ y : xS^P y$ for at least one $d_h \in D'$.

9.12 An Illustrative Example

In this section, we present a didactic example proposed in [15], illustrating how robust ordinal regression can support the DM to specify his/her preferences in a multiobjective optimization problem. In this didactic example, we shall imagine an interaction with a fictitious DM so as to exemplify and illustrate the type of interaction proposed in our methodology.

We consider an MOO problem involving five objectives that are to be maximized. Let us consider a subset A of the Pareto frontier of the MOO problem consisting of

Table 9.1 The set A of Pareto optimal solutions for the illustrative MOO problem

s_1	=	(14.5, 147, 4, 1014, 5.25)
s_2	=	(13.25, 199.125, 4, 1014, 4)
s_3	=	(15.75, 164.375, 16.5, 838.25, 5.25)
s_4	=	(12, 181.75, 16.5, 838.25, 4)
s_5	=	(12, 164.375, 54, 838.25, 4)
s_6	=	(13.25, 199.125, 29, 662.5, 5.25)
s_7	=	(13.25, 147, 41.5, 662.5, 5.25)
s_8	=	(17, 216.5, 16.5, 486.75, 1.5)
s_9	=	(17, 147, 41.5, 486.75, 5.25)
s_{10}	=	(15.75, 216.5, 41.5, 662.5, 1.5)
s_{11}	=	(15.75, 164.375, 41.5, 311, 6.5)
s_{12}	=	(13.25, 181.75, 41.5, 311, 4)
s_{13}	=	(12, 199.125, 41.5, 311, 2.75)
s_{14}	=	(17, 147, 16.5, 662.5, 5.25)
s_{15}	=	(15.75, 199.125, 16.5, 311, 6.5)
s_{16}	=	(13.25, 164.375, 54, 311, 4)
s_{17}	=	(17, 181.75, 16.5, 486.75, 5.25)
s_{18}	=	(14.5, 164.375, 41.5, 838.25, 4)
s_{19}	=	(15.75, 181.75, 41.5, 135.25, 5.25)
s_{20}	=	(15.75, 181.75, 41.5, 311, 2.75)

20 solutions (see Table 9.1). Note that this set A is to be computed using MOO or EMO algorithms (see [8]). Let us suppose that the reference sample A^R of solutions from set A is the following: $A^R = \{s_1, s_2, s_4, s_5, s_8, s_{10}\}$. For the sake of simplicity, we shall consider the set A^R constant across iterations (although the interaction scheme permits A^R to evolve during the process). For the same reason, we will suppose that the DM expresses preference information only in terms of pairwise comparisons of solutions in A^R (intensity of preference will not be expressed in the preference information).

The DM does not see any satisfactory solution in the reference sample A^R (s_1, s_2, s_4 and s_5 have too weak evaluations on the first criterion, while s_8 and s_{10} have the worst evaluation in A on the last criterion), and wishes to find a satisfactory solution in A. Obviously, solutions in A are not comparable unless preference information is expressed by the DM. In this perspective, he/she provides a first pairwise comparison: $s_1 \succ s_2$.

Considering the provided preference information, we can compute the necessary and possible rankings on set A. The DM decided to consider the necessary ranking only, as it has more readable graphical representation than the possible ranking at the stage of relatively poor preference information. The partial preorder of the necessary ranking is depicted in Fig. 9.1 and shows the comparisons that hold for all additive value functions compatible with the information provided by the DM (i.e., $s_1 \succ s_2$). It should be observed that the computed partial preorder contains the preference information provided by the DM (dashed arrow), but also additional comparisons that result from the initial information (continuous arrows); for instance, $s_3 \succ^N s_4$ holds as $U(s_3) > U(s_4)$ holds for all compatible value functions.

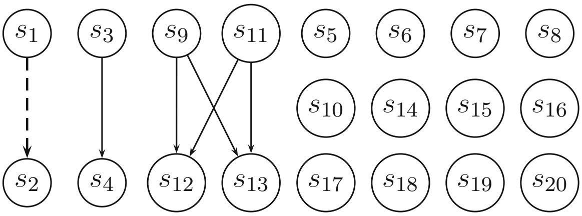
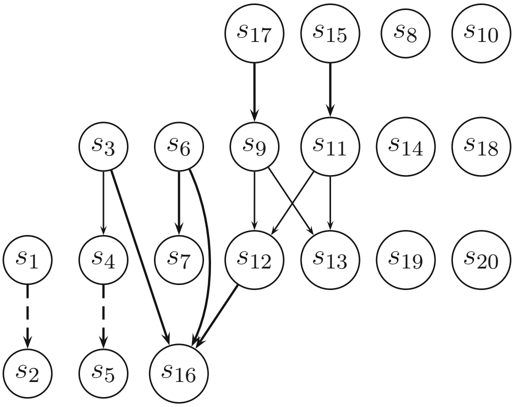


Fig. 9.1 Necessary partial ranking at the first iteration

Fig. 9.2 Necessary partial ranking at the second iteration



Analyzing this first result, the DM observes that the necessary ranking is still very poor, which makes it difficult to discriminate among the solutions in A. He/she reacts by stating that s_4 is preferred to s_5. Considering this new piece of preference information, the necessary ranking is computed again and shown in Fig. 9.2. At this second iteration, it should be observed that the resulting necessary ranking has been enriched as compared to the first iteration (bold arrows), narrowing the set of "best choices," i.e., solutions that are not preferred by any other solution in the necessary ranking: $\{s_1, s_3, s_6, s_8, s_{10}, s_{14}, s_{15}, s_{17}, s_{18}, s_{19}, s_{20}\}$.

The DM believes that this necessary ranking is still insufficiently decisive and adds a new pairwise comparison: s_8 is preferred to s_{10}. Once again, the necessary ranking is computed and shown in Fig. 9.3.

At this stage, the set of possible "best choices" has been narrowed down to a limited number of solutions, among which s_{14} and s_{17} are judged satisfactory by the DM. In fact, these two solutions have a very good performance on the first criterion without "dramatically" bad evaluation on the other criteria.

The current example stops at this step, but the DM could then decide to provide further preference information to enrich the necessary ranking. He/she could also compute new Pareto optimal solutions "close" to s_{14} and s_{17} to focus the search in this area. In this example, we have shown that the proposed interactive process

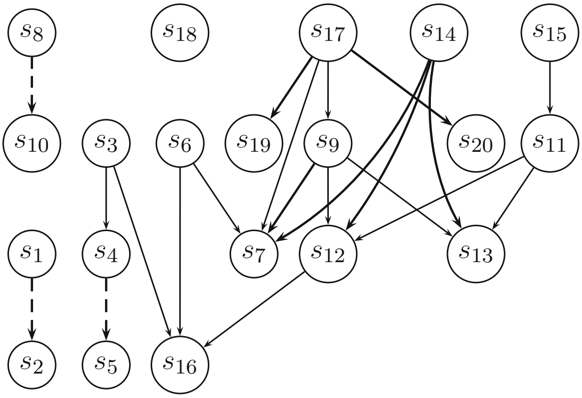


Fig. 9.3 Necessary partial ranking at the third iteration

supports the DM in choosing most satisfactory solutions, without imposing any strong cognitive effort, as the only information required is a holistic information.

9.13 Conclusions and Further Research Directions

In this chapter we presented the basic principle of robust ordinal regression, which is to take into account all the sets of parameters of a preference model compatible with the preference information given by the DM. We recalled the main multiple criteria decision methods to which it has been applied, in particular UTA^{GMS} and *GRIP* dealing with choice and ranking problems, and $UTADIS^{GMS}$ dealing with sorting (ordinal classification) problems. We presented also robust ordinal regression applied to Choquet integral for choice, ranking, and sorting problems, with the aim of representing interactions between criteria. Moreover, we described an interactive multiobjective optimization methodology based on robust ordinal regression, and an evolutionary multiobjective optimization methodology, called *NEMO*, which is also using the principle of robust ordinal regression. In order to show that robust ordinal regression is a general paradigm, independent of the type of preference model involved, we described the robust ordinal regression methodology for outranking methods, and for multiple criteria group decisions. Finally, we presented an exemplary application of robust ordinal regression methodology. Future research will be related to the development of a user friendly software and to specialization of robust ordinal regression methodology to specific real-life problems, such us environmental management, financial planning, and bankruptcy risk evaluation.

Acknowledgments The second author wishes to acknowledge financial support from the Polish Ministry of Science and Higher Education, grant N N519 314435. The third author acknowledges the FCT/CNRS 2009 Luso-French grant, between CEG-IST and LAMSADE.

References

1. S. Angilella, S. Greco, F. Lamantia, and B. Matarazzo. Assessing non-additive utility for multicriteria decision aid. *European Journal of Operational Research*, 158:734–744, 2004.
2. S. Angilella, S. Greco, and B. Matarazzo. Non-additive robust ordinal regression: a multiple criteria decision model based on the Choquet integral. *European Journal of Operational Research*, 201:277–288, 2010.
3. S. Angilella, S. Greco, and B. Matarazzo. Sorting decisions with interacting criteria. http://services.economia.unitn.it/AttiAMASES2008/Lavori/angilella.pdf. Presented at the A.M.A.S.E.S. conference, Trento, September 1–4, 2008.
4. S. Angilella, S. Greco, and B. Matarazzo. Non-additive robust ordinal regression with Choquet integral, bipolar and level dependent Choquet integrals. In J.P. Carvalho, D. Dubois, U. Kaymak, and J.M.C. Sousa, editors, *Proceedings of the Joint 2009 International Fuzzy Systems Association World Congress and 2009 European Society of Fuzzy Logic and Technology Conference, Lisbon, Portugal, July 20–24, 2009*. European Society for Fuzzy Logic and Technology, July 2009.
5. C.A. Bana e Costa, J.M. De Corte, and J.C. Vansnick. On the mathematical foundation of MACBETH. In J. Figueira, S. Greco, and M. Ehrgott, editors, *Multiple Criteria Decision Analysis: State of the Art Surveys*, pages 409–443. Springer, New York, 2005.
6. C.A. Bana e Costa and J.C. Vansnick. MACBETH: An interactive path towards the construction of cardinal value functions. *International Transactions in Operational Research*, 1(4):387–500, 1994.
7. V. Belton, J. Branke, P. Eskelinen, S. Greco, J. Molina, F. Ruiz, and R. Słowiński. Interactive multiobjective optimization from a learning perspective. In J. Branke, K. Deb, K. Miettinen, and R. Słowiński, editors, *Multiobjective Optimization: Interactive and Evolutionary Approaches*, volume 5252 of *Lecture Notes in Computer Science*, pages 405–434. Springer, Berlin, 2008.
8. J. Branke, K. Deb, K. Miettinen, and R. Słowiński, editors. *Multiobjective Optimization: Interactive and Evolutionary Approaches*, volume 5252 of *Letcure Notes in Computer Science*. Springer, Berlin, 2008.
9. J. Branke, S. Greco, R. Słowiński, and P. Zielniewicz. Interactive evolutionary multiobjective optimization using robust ordinal regression. In M. Ehrgott, C.M. Fonseca, X. Gandibleux, J.-K. Hao, and M. Sevaux, editors, *Proceedings of the 5th International Conference on Evolutionary Multi-Criterion Optimization, EMO'09. Nantes, April 07–10, 2009*, volume 5467 of *Lecture Notes in Computer Science*, pages 554–568. Springer, Berlin, 2009.
10. J. Branke, S. Greco, R. Słowiński, and P. Zielniewicz. NEMO-II: integrating evolutionary multiobjective optimization and decision making. Presented at the 23rd European Conference on Operational Research, Bonn, July 5–8, 2009.
11. G. Choquet. Theory of capacities. *Annales de l'Institut Fourier*, 5:131–295, 1953.
12. K. Deb, A. Pratap, A. Agarwal, and T. Meyarivan. A fast and elitist multiobjective genetic algorithm: NSGA-II. *IEEE Transactions on Evolutionary Computation*, 6:182–197, 2002.
13. J.S. Dyer. Multiattribute utility theory. In J. Figueira, S. Greco, and M. Ehrgott, editors, *Multiple Criteria Decision Analysis: State of the Art Surveys*, pages 265–295. Springer, New York, 2005.
14. J. Figueira, S. Greco, and M. Ehrgott. *Multiple Criteria Decision Analysis: State of the Art Surveys*. Springer, New York, 2005.
15. J. Figueira, S. Greco, V. Mousseau, and R. Słowiński. Interactive multiobjective optimization using a set of additive value functions. In J. Branke, K. Deb, K. Miettinen, and R. Słowiński, editors, *Multiobjective Optimization: Interactive and Evolutionary Approaches*, volume 5252 of *Lecture Notes in Computer Science*, pages 97–120. Springer, Berlin, 2008.
16. J. Figueira, S. Greco, V. Mousseau, and R. Słowiński. UTA^{GMS} and *GRIP* methodology for multiple criteria decision problems. presented at the 19th International Conference on Multiple Criteria Decision Making, Auckland, January 7–12, 2008.

17. J. Figueira, S. Greco, and R. Słowiński. Identifying the "most representative" value function among all compatible value functions in the grip method. Presented at the 68th Meeting of the European Working Group on Multiple Criteria Decision Aiding, Chania, October 2–3, 2008.
18. J. Figueira, S. Greco, and R. Słowiński. Building a set of additive value functions representing a reference preorder and intensities of preference: GRIP method. *European Journal of Operational Research*, 195:460 – 486, 2009.
19. J. Figueira, V. Mousseau, and B. Roy. ELECTRE methods. In J. Figueira, S. Greco, and M. Ehrgott, editors, *Multiple Criteria Decision Analysis: State of the Art Surveys*, chapter 4, pages 133–162. Springer, New York, 2005.
20. A.M. Geoffrion. Proper effciency and the theory of vector maximization. *Journal of Mathematical Analysis Application*, 22:618–630, 1968.
21. M. Grabisch. The application of fuzzy integrals in multicriteria decision making. *European Journal of Operational Research*, 89:445–456, 1996.
22. M. Grabisch. k-order additive discrete fuzzy measures and their representation. *Fuzzy Sets and Systems*, 92:167–189, 1997.
23. M. Grabisch and C. Labreuche. Bi-capacities–I: Definition, Möbius transform and interaction. *Fuzzy Sets and Systems*, 151:211–236, 2005.
24. M. Grabisch and C. Labreuche. Bi-capacities–II: The Choquet integral. *Fuzzy Sets and Systems*, 151:237–259, 2005.
25. M. Grabisch and C. Labreuche. Fuzzy measures and integrals in MCDA. In J. Figueira, S. Greco, and M. Ehrgott, editors, *Multiple Criteria Decision Analsysis: State of the Art Surveys*, pages 563–608. Springer, New York, 2005.
26. S. Greco, S. Giove, and B. Matarazzo. The Choquet integral with respect to a level dependent capacity. Submitted to *Fuzzy Sets and Systems*, 2009.
27. S. Greco, M. Kadziński, and R. Słowiński. The most represenative value function for multiple criteria sorting based on robust ordinal regression. Submitted to *Computers & Operations Research*, 2009.
28. S. Greco, B. Matarazzo, and R. Słowiński. The use of rough sets and fuzzy sets in MCDM. In T. Gal, T. Hanne, and T. T. Stewart, editors, *Multicriteria Decision Making: Advances in MCDM Models, Algorithms, Theory and Applications*, pages 14.1–14.59. Kluwer, Dordrecht, 1999.
29. S. Greco, B. Matarazzo, and R. Słowiński. Rough sets theory for multicriteria decision analysis. *European Journal of Operational Research*, 129:1–47, 2001.
30. S. Greco, B. Matarazzo, and R. Słowiński. Bipolar Sugeno and Choquet integrals. In B. De Baets, J. Fodor, and G. Pasi, editors, *Proceedins of the EUROFUSE Workshop on Informations Systems*, pages 191–196, Varenna, September 2002.
31. S. Greco, B. Matarazzo, and R. Słowiński. Decision rule approach. In J. Figueira, S. Greco, and M. Ehrgott, editors, *Multiple Criteria Decision Analsysis: State of the Art Surveys*, pages 507–561. Springer, New York, 2005.
32. S. Greco, V. Mousseau, and R. Słowiński. Multiple criteria sorting with a set of additive value functions. Presented at the 18th International Conference on Multiple Criteria Decision Making, Chania, June 19–23, 2006.
33. S. Greco, V. Mousseau, and R. Słowiński. The necessary and the possible in a perspective of robust decision aiding. Presented at the 66th Meeting of the European Working Group on Multiple Criteria Decision Aiding, Marrakech, October 18–20, 2007.
34. S. Greco, V. Mousseau, and R. Słowiński. Ordinal regression revisited: Multiple criteria ranking with a set of additive value functions. *European Journal of Operational Research*, 191:415–435, 2008.
35. S. Greco, V. Mousseau, and R. Słowiński. Multiple criteria sorting with a set of additive value functions. *European Journal of Operational Research*, in print, doi:10.1016/j.ejor.2010.05.021, 2010.
36. S. Greco, V. Mousseau, and R. Słowiński. The possible and the necessary for multiple criteria group decision. In F. Rossi and A. Tsoukias, editors, *Proceedings of the First International Conference on Algorithmic Decision Theory (ADT 2009)*, volume 5783 of *Lecture Notes in Artificial Intelligence*, pages 203–214. Springer Verlag, Berlin, 2009.

37. G. W. Greenwood, X.S. Hu, and J.G. D'Ambrosio. Fitness functions for multiple objective optimization problems: Combining preferences with pareto rankings. In R.K. Belew and Vose M.D., editors, *Foundations of Genetic Algorithms*, pages 437–455. Morgan Kaufmann, San Francisco, 1997.
38. E. Jacquet-Lagrèze, R. Meziani, and R. Słowiński. MOLP with an interactive assessment of a piecewise-linear utility function. *European Journal of Operational Research*, 31:350–357, 1987.
39. E. Jacquet-Lagrèze and Y. Siskos. Assessing a set of additive utility functions for multicriteria decision-making, the UTA method. *European Journal of Operational Research*, 10:151–164, 1982.
40. E. Jacquet-Lagrèze and Y. Siskos. Preference disaggregation: 20 years of MCDA experience. *European Journal of Operational Research*, 130:233–245, 2001.
41. A. Jaszkiewicz and J. Branke. Interactive multiobjective evolutionary algorithms. In J. Branke, K. Deb, K. Miettinen, and R. Słowiński, editors, *Multiobjective Optimization: Interactive and Evolutionary Approaches*, volume 5252 of *Lecture Notes in Computer Science*, pages 179–193. Springer, Berlin, 2008.
42. R.L. Keeney and H. Raiffa. *Decision with Multiple Objectives – Preferences and Value Trade-offs*. Wiley, New York, 1976.
43. S.H. Kim and B.S. Ahn. Interactive group decision making procedure under incomplete information. *European Journal of Operational Research*, 116:498–507, 1999.
44. L. Kiss, J.M. Martel, and R. Nadeau. ELECCALC - an interactive software for modelling the decision maker's preferences. *Decision Support Systems*, 12:757–777, 1994.
45. M. Koksalan and S. Bilgin Ozpeynirci. An interactive sorting method for additive utility functions. *Computers & Operations Research*, 36:2565–2572, 2009.
46. J. March. Bounded rationality, ambiguity and the engineering of choice. *Bell Journal of Economics*, 9:587–608, 1978.
47. J.L. Marichal and M. Roubens. Determination of weights of interacting criteria from a reference set. *European Journal of Operational Research*, 124:641–650, 2000.
48. V. Mousseau and L. Dias. Valued outranking relations in ELECTRE providing manageable disaggregation procedures. *European Journal of Operational Research*, 156:467–482, 2004.
49. V. Mousseau and R. Słowiński. Inferring an ELECTRE TRI model from assignment examples. *Journal of Global Optimization*, 12:157–174, 1998.
50. V. Mousseau, R. Słowiński, and P. Zielniewicz. A user-oriented implementation of the ELECTRE-TRI method integrating preference elicitation support. *Computers & Operations Research*, 27:757–777, 2000.
51. D. Pekelman and S.K. Sen. Mathematical programming models for the determination of attribute weights. *Management Science*, 20:1217–1229, 1974.
52. B. Roy. The outranking approach and the foundations of ELECTRE methods. *Theory and Decision*, 31:49–73, 1991.
53. B. Roy and D. Bouyssou. *Aide Multicritère à la Décision : Méthodes et Cas*. Economica, Paris, 1993.
54. T.L. Saaty. *The Analytic Hierarchy Process*. McGraw-Hill, New York, 1980.
55. T.L. Saaty. The analytic hierarchy and analytic network processes for the measurement of intangible criteria and for decision-making. In J. Figueira, S. Greco, and M. Ehrgott, editors, *Multiple Criteria Decision Analysis: State of the Art Surveys*, pages 345–407. Springer, New York, 2005.
56. L.S. Shapley. A value for n-person games. In H.W. Kuhn and A.W. Tucker, editors, *Contributions to the Theory of Games II*, pages 307–317. Princeton University Press, Princeton, 1953.
57. J. Siskos and D.K. Despotis. A DSS oriented method for multiobjective linear programming problems. *Decision Support Systems*, 5:47–55, 1989.
58. Y. Siskos, V. Grigoroudis, and N. Matsatsinis. UTA methods. In J. Figueira, S. Greco, and M. Ehrgott, editors, *Multiple Criteria Decision Analysis: State of the Art Surveys*, pages 297–343. Springer, New York, 2005.

59. R. Słowiński, S. Greco, and B. Matarazzo. Rough set based decision support. In E.K. Burke and G. Kendall, editors, *Search Methodologies: Introductory Tutorials in Optimization and Decision Support Techniques*, chapter 16, pages 475–527. Springer, New York, 2005.
60. V. Srinivasan and A.D. Shocker. Estimating the weights for multiple attributes in a composite criterion using pairwise judgments. *Psychometrika*, 38:473–493, 1973.
61. M. Sugeno. *Theory of fuzzy integrals and its applications*. Ph.D. thesis, Tokyo Institute of Technology, 1974.
62. M. Weber. A method of multiattribute decision making with incomplete information. *Management Science*, 31(11):1365–1371, 1985.