Controlling robustness in ordinal regression

Salvatore Greco
Faculty of Economics, University of Catania, Corso Italia 55, 95129 Catania, Italy, e-mail: salgreco@unict.it

Yannis Siskos
University of Piraeus, 80, Karaoli & Dimitriou Street, GR-18534 Piraeus, Greece e-mail: ysiskos@unipi.gr

Roman Slowinski
Institute of Computing Science, Poznan University of Technology, 60-965 Poznan,
and Systems Research Institute, Polish Academy of Sciences, 01-447 Warsaw, Poland, e-mail: roman.slowinski@cs.put.poznan.pl

Within Multiple Criteria Decision Aiding (MCDA) ordinal regression methods builds a preference model expressed either as a value function or as an outranking relation on the basis of some preference expressed by the Decision Maker (DM). The first method using this approach was UTA in which one looks for a piecewise additive value function representing DM preference information in a ranking decision problem. The same methodology has been used in the UTADIS method for decision sorting problems. Ordinal regression has also widely used for outranking methods, both for ranking and sorting decision problems. Recently ordinal regression has been reconsidered under the light of the following observation: very often there exists not only one instance of a preference model of a given class (e.g. additive value function or outranking model), but there exists a plurality of instances that represents as well the preference information given by the DM. However, each one of those instances of the a preference model expresses different preferences when applied to pairs of alternatives different from those ones considered in the preference information given by the DM. On the basis of this observation Robust Ordinal Regression (ROR) has been proposed. ROR consider the whole set of instance of a given preference model representing the preference information given by the DM and expresses the recommendations of the MCDA procedure in terms of possible and necessary preferences. For example, if the considered decision model is an additive value function, ROR considers the whole set $\boldsymbol{U}$ of additive value U functions compatible with the preference information given by the DM and for any pair of alternatives a and b, we say that

- a is necessarily preferred to b, if $U(a)\geq U(b)$ for all $U\in\boldsymbol{U}$,
- a is necessarily preferred to b, if $U(a)\geq U(b)$ for at least one $U\in\boldsymbol{U}$.

Consideration of the whole set of instances of a given model compatible with the preference information of the DM permits to avoid the arbitrariness of choosing only one compatible instance. ROR has been firstly proposed in the UTA^{GMS} method, after generalized in the GRIP method. UTA^{GMS} and GRIP considered a ranking decision problem dealt with a preference model expressed in terms of an additive value function. ROR has been applied also to decision sorting problem with the UTADIS^{GMS} method. Moreover, ROR has been considered also for outranking methods (ELECTRE^{GKMS}, PROMETHEE^{GKS}) and for non additive value functions expressed in terms of Choquet integral. It has also been successfully applied to group decision problems (UTA^{GMS}-GROUP, UTADIS^{GMS}-GROUP, ELECTRE^{GKMS}-GROUP and so on).

Let us observe that even if the idea of considering the whole set of instances of a given preference model compatible with the preference information expressed by the DM is appealing, there is always the risk that the results that one could obtain in terms of necessary and possible preferences are too vague because the set of compatible instances of the considered preference model is in some form "too large". In order to control this aspect of ROR methodology we propose to measure the set of compatible instances of the considered preference model. In fact, each compatible instance of preference model is defined by a set of decision parameters (for example, in case of additive value functions, the preference parameters are the value assigned by the marginal value function to the evaluations of the representative alternatives considered by the DM for her preference information). Thus, we propose to measure the set of compatible instances of the considered preference model as the hypervolume of the set of corresponding preference parameters. This hypervolume can be approximated using a Montecarlo approach, i.e. randomly generating a set of vectors of preferential parameters. For the sake of simplicity we propose this methodology taking into account UTA^{GMS} method.

The set of constraints COMP to be satisfied by a compatible additive value functions are

$$\begin{cases} U(a)\geq U(b)+\varepsilon \text{ for all } a\succ b,\ U(a)=U(b) \text{ for all } a\sim b \\ u_j(x_j^{r+1})\geq u_j(x_j^r) \text{ where } x_j^r\in X_j=\{g_j(a),\ a\in A\},\text{ such that } x_j^1<x_j^1<\ldots<x_j^{card(X_j)} \text{ for all } g_j\in G, \\ u_j(x_j^1)=0, \text{ for all } g_j\in G,\ u_1(x_1^{card(X_1)})+\ldots+u_m(x_m^{card(X_m)})=1, \\ \text{where } \varepsilon \text{ is a small positive value.} \end{cases}$$

To compute an approximation of the volume of the set of parameters corresponding to the set of compatible additive value functions one can proceed as follows. First we compute the volume of the whole set of preferential parameters when there is no preference information. It can be done in this way. We generate a large number of random vectors

$$\alpha=\left[\alpha_j^r, x_j^r\in X_j-\left\{x_j^{card(X_j)}\right\}, g_j\in G\right]$$

such that

$$\alpha_j^r\geq 0, \text{for all } x_j^r\in X_j-\left\{x_j^{card(X_j)}\right\}, g_j\in G$$

and

$$\sum_{x_j^r \in X_j - \left\{x_j^{card(X_j)}\right\}, g_j \in G} \alpha_j^r = 1 .$$

After we solve the following linear programming problem for each one of those vectors:

$$\text{Max} \sum_{x_j^r \in X_j - \left\{x_j^{card(X_j)}\right\}, g_j \in G} \alpha_j^r \varepsilon_j^r$$

$$\begin{cases} u_j(x_j^{r+1}) \geq u_j(x_j^r) + \varepsilon_j^r \quad \text{where } x_j^r \in X_j - \{ x_j^{card(X_j)} \} \text{ where } X_j = \{g_j(a),\ a \in A\}, \text{ such that } x_j^1 < x_j^1 < \ldots < x_j^{card(X_j)} \text{ for all } g_j \in G \\ u_j(x_j^1) = 0, \text{ for all } g_j \in G,\ u_1(x_1^{card(X_1)}) + \ldots + u_m(x_m^{card(X_m)}) = 1 \\ \varepsilon_j^r \geq 0 \text{ for all where } x_j^r \in X_j - \{ x_j^{card(X_j)} \} \text{ and } g_j \in G. \end{cases}$$

Each one of the solutions we obtain from the previous linear programming problems gives the preferential parameters of an additive value function. Let us denote by TOT the total number of the additive value functions we generated. For each one of those solutions we can verify if it satisfies the constraints from COMP. Let us denote by COMPTOT the total number of the additive value functions satisfying COMP. The ratio between COMPTOT and TOT gives an approximated measure of the hypervolume of the set of compatible value functions considering as unitary the hypervolume of the set of all additive value functions.

With the same methodology we can compute a credibility of the weak preference of alternative a^* over alternative b^*. Let us generate a large number of random vectors α with

$$\alpha = \left[\alpha_j^r, x_j^r \in X_j - \left\{x_j^{card(X_j)}\right\}, g_j \in G; \alpha_{(a,b)}, a \succ b\right],$$

such that

$$\alpha_j^r \geq 0, \text{for all } x_j^r \in X_j - \left\{x_j^{card(X_j)}\right\}, g_j \in G,\ \ \alpha_{(a,b)} \geq 0, \text{for all } a, b \text{ such that } a \succ b, \text{ and}$$

$$\sum_{x_j^r \in X_j - \left\{x_j^{card(X_j)}\right\}, g_j \in G} \alpha_j^r + \sum_{a,b: a \succ b} \alpha_{(a,b)} = 1 .$$

After we solve the following linear programming problem for each one of those vectors:

$$\text{Max} \sum_{x_j^r \in X_j - \left\{x_j^{card(X_j)}\right\}, g_j \in G} \alpha_j^r \varepsilon_j^r$$

$$\begin{cases} U(a) \geq U(b) + \varepsilon_{(a,b)} \text{ for all } a \succ b,\ U(a) = U(b) \text{ for all } a \sim b, \\ u_j(x_j^{r+1}) \geq u_j(x_j^r) + \varepsilon_j^r \quad \text{where } x_j^r \in X_j - \{ x_j^{card(X_j)} \} \text{ where } X_j = \{g_j(a),\ a \in A\}, \text{ such that } x_j^1 < x_j^2 < \ldots < x_j^{card(X_j)} \text{ for all } g_j \in G, \\ u_j(x_j^1) = 0, \text{ for all } g_j \in G,\ u_1(x_1^{card(X_1)}) + \ldots + u_m(x_m^{card(X_m)}) = 1, \\ \varepsilon_j^r \geq 0 \text{ for all where } x_j^r \in X_j - \{ x_j^{card(X_j)} \} \text{ and } g_j \in G. \end{cases}$$

Each one of the solutions we obtain from the previous linear programming problems gives the preferential parameters of a compatible additive value function. Let us denote by COMP the set of those value function and by COMPTOT the total number of the additive value functions we generated. For each one of those solutions we can verify if it satisfies the constraint $U(a^*) \geq U(b^*)$. Let us denote by COMP(a^*,b^*) the set of those compatible value functions and let us denote by COMPTOT(a^*,b^*). The ratio between COMPTOT(a^*,b^*) and COMPTOT gives an approximated measure of the hypervolume of COMP(a^*,b^*) considering as unitary the volume of COMP and can be interpreted as the probability that random picking a compatible value function alternative a^* is at least as good as alternative b^*.