Parsimonious preference model for robust ordinal regression

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

Vincent Mousseau
LGI, Ecole Centrale Paris, Grande Voie des Vignes, 92 295 Châtenay-Malabry, France, e-mail: vincent.mousseau@ecp.fr

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

We consider the ordinal regression approach to determining a complete set of preference models of a given class (e.g., additive value functions, like in UTA^{GMS}, or outranking models, like in $ELECTRE^{GKMS}$ or $PROMETHEE^{GKS}$), compatible with preference information provided by the Decision Maker (DM). An original component of this approach consists in ensuring not only the compatibility of the model with the preference information, but also keeping the model as parsimonious as possible. In this case, "parsimonious" means simple and intuitive. To present the main idea of this approach, we shall consider here a typical preference model of Multiple Attribute Utility Theory: an additive value function. Thus, we consider a set alternatives A={a,b,…}, card(A)=n, evaluated by means of a consistent family of criteria $G=\{g_1, \dots, g_m\}$, where criterion g_j:A→**R**, such that for all a,b∈A, $g_j(a)\geq g_j(b)$ means that a is at least as good as b with respect to $g_j\in G$. The preference model is the following value function: $U(a)= u_1(g_1(a))+\ldots+u_m(g_m(a))$, for all a∈A, where marginal value function u_j: **R**→**R** is nondecreasing, j=1,…,m, such that for all a,b∈A, U(a)≥U(b) means that a is comprehensively at least as good as b. We shall take into account UTA and UTA^{GMS} methods. UTA looks for additive value functions with piecewise-linear marginal value functions, while UTA^{GMS} considers all compatible additive value functions with marginal value functions which are generically nondecreasing. In technical terms, this means that in UTA of the characteristic points of a marginal value function u_j are denoted by $u_j(x_j^k)$, k=0,1,…,h_j, where h_j is equal to the admitted number of linear pieces, i.e. the number of equal intervals of values of criterion $g_j\in G$. For the intermediate values a linear interpolation is considered. In case of UTA^{GMS}, all the values $u_j(g_j(a))$, for all a∈A and $g_j\in G$, are considered as characteristic points, thus a linear interpolation is not necessary. In this way, UTA^{GMS} is considering the most general form of the additive value function. As to the capacity of representation, the general additive model is able to represent the most complex preferences that an additive model is able to express, while the capacity of representation by the piecewise-linear model is smaller. On the other hand, the general additive model is more complex and less intuitive than the piecewise-linear model of UTA, and it is not always necessary for preference representation. This is why we want to propose the ordinal regression approach which seeks for a compatible value function which is as simple as possible and nevertheless capable to represent DM's preferences. This is obtained by solving two linear programming problems that permit to find the "most linear" value function able to represent the considered preferences, and the value function with the minimal number of linear pieces, respectively. The two problems are formulated as follows.

The first linear programming problem computes a compatible additive value function whose marginal value functions minimally deviate from linear functions:

Min δ

$U(a)\geq U(b)+\varepsilon$ for all $a\succ b$

$U(a)=U(b)$ for all $a\sim b$

$u_j(x_j^{r+1})\geq u_j(x_j^r)$ where $x_j^r\in X_j=\{g_j(a), a\in A\}$, such that $x_j^1 < x_j^2 <\ldots< x_j^{card(X_j)}$ for all $g_j\in G$

$u_j(x_j^1)=0$, for all $g_j\in G$,

$u_1(x_1^{card(X_1)})+\ldots+u_m(x_m^{card(X_m)})=1$

$$\frac{u_j\left(x_j^{r+2}\right)-u_j\left(x_j^{r+1}\right)}{x_j^{r+2}-x_j^{r+1}}-\frac{u_j\left(x_j^{r+1}\right)-u_j\left(x_j^{r}\right)}{x_j^{r+1}-x_j^{r}}\leq\delta,\ \text{ for all } g_j\in G \text{ and } r=1,\ldots,card(X_j)-2$$

$$\frac{u_j\left(x_j^{r+1}\right)-u_j\left(x_j^{r}\right)}{x_j^{r+1}-x_j^{r}}-\frac{u_j\left(x_j^{r+2}\right)-u_j\left(x_j^{r+1}\right)}{x_j^{r+2}-x_j^{r+1}}\leq\delta,\ \text{ for all } g_j\in G \text{ and } r=1,\ldots,card(X_j)-2$$

where ε is a small positive value.

The second linear programming problem computes a compatible additive value function whose marginal value functions are composed of a minimal number of linear pieces:

$$\text{Min} \sum_{g_j\in G,\ r\in card\left(X_j\right)} \delta_{rj}$$

$U(a)\geq U(b)+\varepsilon$ for all $a\succ b$

$U(a)=U(b)$ for all $a\sim b$

$u_j(x_j^{r+1})\geq u_j(x_j^{r})$ where $x_j^{r}\in X_j=\{g_j(a),\ a\in A\}$, such that $x_j^1<x_j^2<\ldots<x_j^{card\left(X_j\right)}$ for all $g_j\in G$

$u_j(x_j^1)=0$, for all $g_j\in G$,

$u_1(x_1^{card(X_1)})+\ldots+u_m(x_m^{card(X_m)})=1$

$$\frac{u_j\left(x_j^{r+2}\right)-u_j\left(x_j^{r+1}\right)}{x_j^{r+2}-x_j^{r+1}}-\frac{u_j\left(x_j^{r+1}\right)-u_j\left(x_j^{r}\right)}{x_j^{r+1}-x_j^{r}}\leq M_j\delta_{rj},\ \text{ for all } g_j\in G \text{ and } r=1,\ldots,card(X_j)-2$$

$$\frac{u_j\left(x_j^{r+1}\right)-u_j\left(x_j^{r}\right)}{x_j^{r+1}-x_j^{r}}-\frac{u_j\left(x_j^{r+2}\right)-u_j\left(x_j^{r+1}\right)}{x_j^{r+2}-x_j^{r+1}}\leq M_j\ \delta_{rj},\ \text{ for all } g_j\in G \text{ and } r=1,\ldots,card(X_j)-2$$

$\delta_{rj}\in\{0,1\}$ for all $g_j\in G$ and $r=1,\ldots,card(X_j)-2$

where ε is a small positive value and M_j is a "big value" such that $M_j>\max_{r=1,\ldots,card(X_j)-1}\dfrac{1}{x_j^{r+1}-x_j^{r}}$, for all $g_j\in G$.

The two above linear programming problems can be used conjointly. For example, one could compute the minimal possible deviation from linearity solving the first problem and after adding a constraint that limits the deviation from linearity to the maximal value already found, one could look for a value function with a minimal number of linear pieces solving the second linear programming problem. It is also possible to invert the order of using the two problems. Finally it is possible to solve the two linear programming problems and after knowing the minimal deviation from linearity and the minimal number of linear pieces, one could compute the possible and the necessary preference within the set of compatible value functions that deviate from those minimal limits within a prefixed range of tolerance.

Observe also that the two above linear programming problems have a solution in case it is possible to represent the DM's preference with an additive value function. If this is not the case, some more complex value function can be considered, as it is the case in the UTA^{GMS}–INT method. Also in this case, one can consider the above approach in order to search for a parsimonious model, i.e. a model with marginal value functions and marginal synergy functions "as much linear as possible". The same approach can be applied in ordinal regression of outranking methods such as ELECTRE and PROMETHEE.