HAL
open science

On the use of Weighted Regression in Conjoint Analysis

Salwa Benammou, Besma Souissi, Gilbert Saporta

To cite this version:

Salwa Benammou, Besma Souissi, Gilbert Saporta. On the use of Weighted Regression in Conjoint Analysis. COMPSTAT'2010, 19th International Conference on Computational Statistics, Aug 2010, Paris, France. pp.777-784. hal-01125764

HAL Id: hal-01125764

https://hal.science/hal-01125764v1

Submitted on 22 Mar 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

On the use of Weighted Regression in Conjoint Analysis

Salwa Benammou[1], Besma Souissi[2], and Gilbert Saporta[3]

[1] Faculté de Droit et des Sciences Economiques et Politiques, Sousse, Tunisie, *saloua.benammou@fdseps.rnu.tn*
[2] *Institut Supérieur de Gestion, Sousse, Tunisie, besma.swissi@yahoo.fr*
[3] *Chaire de statistique appliquée & CEDRIC, CNAM*
292 rue Saint Martin, Paris, France, gilbert.saporta@cnam.fr

Abstract. Conjoint analysis seeks to explain an ordered categorical ordinal variable according to several variables using a multiple regression scheme. A common problem encountered, there, is the presence of missing values in classification-ranks. In this paper, we are interested in the cases where consumers provide a ranking of some products instead of rating these products (i.e. explained variable presents missing values).In order to deal with this problem, we propose a weighted regression scheme. We empirically show (in several cases of weighting) that, if the number of missing values is not too large, the data remain useful, and our results are close to those of the complete order. A simulation study confirms these findings.

Keywords: conjoint analysis, missing values, weighted regression

1 Introduction

The conjoint analysis is a data analysis method. It links an explained categorical ordered variable to several explanatory variables either or unordered categories. It allows analyzing consumers' preferences for products defined by combinations of attributes, according to these last ones.
Initially developed by psychometrics, the conjoint analysis has been introduced in the marketing research field at the beginning of 1970s (Green and Rao1971). Its use knew a considerable development at the end of 1970s and in the 1980s (Wittink and Cattin on 1989). The conjoint analysis is a complete methodology composed of three phases. The first one, based on experimental design, consists in collecting observations, generally, by direct interviews where each interviewee evaluates a set of real or hypothetical products.
The second phase corresponds to data processing and the parameters' estimation. The conjoint analysis decomposes then the preferences according to a model of additive utility wich is specific for every interviewee. The last phase is dedicated to the simulation of market shares (Benammou et al. 2007).
We are interested here in the phase of treatment, and thus in the estimation of the parameters. When a consumer has to classify by order of preference a set of products, conjoint analysis is a particular case of the ordinary linear

model.
Generally, the estimation is done by ordinary least square method. It supposes that the consumer gives the same weight to all products and that the "distance" between two products of successive ranks is the same for all the ranks of classification. This hypothesis seems plausible in the case of the total order.
But, in actual fact, the cognitive capacity of the consumer decreases when the number of scenarios increases and the consumer tends to classify only the most favorite products. It is easier to imagine two or three products and to classify them rather than 10 or 12 products. At the end, we obtain missing ranks in the classification. Benammou et al (2003) show upon an example a good stability of the results when the missing values don't exceed half of the classified scenarios.
In some cases the distance between ranks of classification can be different. The difference between the second classified product and the first one is much smaller than the one between the last product and the next to last before. Thus we propose the use of weighted least square method to model this behavior.
Furthermore, if we suppose that the consumer gives intuitively more (respectively less) importance to the most (respectively least) preferred products, it would then be logical to give a higher weight (respectively weak) to the first classified products (respectively the last ones).
We propose the use of decreasing weight functions; what give weak weights for the last classified ranks. These functions seem to better describe the behavior of the consumer. The use of fast decreasing weight functions can give an alternative solution to the problem of partial classification of products, especially when the number of non classified product is important.

2 Reminder on weighted regression

Let's consider q products described by p qualitative variables X_1, X_2, ..., X_p in m_1, m_2, ...,m_p categories respectively. Generally, the associated linear model is given by (1)

$$\mathbf{y} = \mathbf{X}\beta + \boldsymbol{e} \tag{1}$$

Where β is the vector of parameters to estimate, $\mathbf{y}$ the vector of classification ranks, $\mathbf{X}$ the experimental matrix, (q,$\sum_{i=1}^{p} m_i$) the size of $\mathbf{X}$, and $\boldsymbol{e}$ the random errors vector associated such as $Var(e) = \sigma^2 \Sigma$.
Here Σ is the errors variance covariance matrix which equals the identity in the case of a classic linear model. The value of the Generalized Least Squares (GLS) estimator is then given by the relation (2)

$$\hat{\beta} = (\mathbf{X}'\Sigma^{-1}\mathbf{X})^{-1}(\mathbf{X}'\Sigma\mathbf{y}) \tag{2}$$

For the weighted least squares (see for example Weisberg , 1985) every coordinate of errors vector e is correlated to all the others; but the variances can not be the same and the matrix Σ is then in the form (3).

$$\Sigma = \begin{bmatrix} W_1 & 0 & \cdots & 0 \\ 0 & W_2 & \cdots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & 0 & W_q \end{bmatrix} \tag{3}$$

Where $W_i > 0$ and $\sum_{i=1}^{q} W_i = 1$.

Let us suppose that a diagonal matrix $\mathbf{C}$ such as Σ exists. The matrix $\mathbf{C}$ is then the "square root" of the inverse of the matrix Σ and we have $Var(\mathbf{C}\boldsymbol{e}) = \sigma^2\mathbf{I}$. By an appropriate transformation of variables, we can recover the common linear model.
And so if we suppose $\mathbf{y_w} = \mathbf{Cy}$, $\mathbf{X_w} = \mathbf{CX}$ and $\mathbf{e_w} = \mathbf{C}\boldsymbol{e}$,
the model becomes $\mathbf{y_w} = \mathbf{X_w}\beta + \mathbf{e_w}$ and we recover the same shape as described by the relation (1). This model verifies all the hypotheses required by the classical linear model and we can use the ordinary least squares method to estimate its parameters (relation (4)).

$$\hat{\beta} = (\mathbf{X'_w X_w})^{-1}(\mathbf{X'_w y_w}) \tag{4}$$

3 Exemple

3.1 The data

To be able to compare our results with those of the literature, we take back the data of the example treated by Benammou et al (2003), relatives to 263 consumers classifying scenarios of mobile phone subscriptions.

TABLE 1. *List and labels of variables*

Label	*Code*	*Categori*		
Device price (in euro)	*Device price*	0	100	
Subscription fee (in euro)	*Sub fee*	0	30	
Peak hours definition	*Peak hours def*	p_1	p_2	
Duration of subscription (in month)	*Duration*	6	24	
Monthly subscription price (in euro)	*Monthly sub*	0	4,5	9
Price/minute peak hours (in euro)	*Peak hours*	0,5	0,7	0,9
Price/minute off − peak hours (in euro)	*Off − peak hours*	0,008	0,1	0,15

The scenarios emanate of seven variables, where four variables are in two categories and three in three categories and the number of parameters linearly independent to be considered equals 10.The design used is a D-optimal one with 12 scenarios. The obtained products who presented to 263 consumers whom answered by a total ranking. We give variables description in table 1.

3.2 The results and their interpretations

3.2.1 Studies of R^2

Goodness of fit is measured by multiple correlation coefficient R^2. This coefficient is an indicator of the quality of adjustment of the model. When this coefficient is low it denotes an inadequacy of the model, or an incoherence of interviewee's responses. Generally the second hypothesis is retained since the model seems realistic. We should eliminate interviewee for whom R^2 is less than a critical value fixed by user. (Benammou et al 2003). In this example, Benammou et al (2003) showed that R^2 is close to 1 for almost all individuals, in the total order case. The authors proposed three simple procedures for estimating missing values. The first one consists in attributing to all the non classified products the rank of the last classified product increased by 1. In the second one, all the non classified products receive the average of missing ranks. For the third one, all the non classified products receive the maximum rank. The results given by the three procedures being equivalent, we give here those of the first one only.

The R^2 of the weighted linear model in the case of total order (see Fig. 1.) shows that they are better than those of the classic linear model. The adjustment quality improves with the decreasing speed of weight function.

For example in the case of the functions f(x)=e^{-2x} and f(x)=$(\frac{1}{2})^x$ the R^2 is very close or equal to 1 for the majority of the individuals.

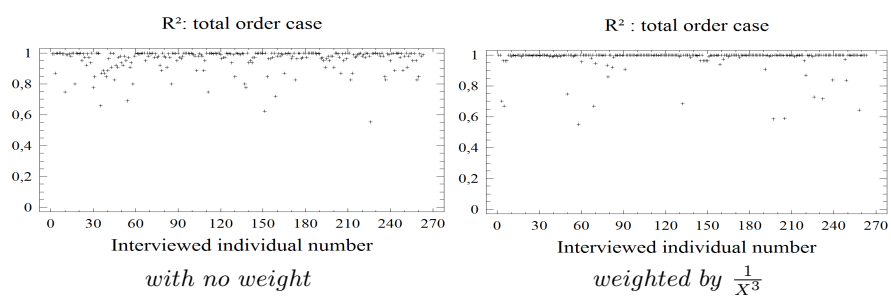
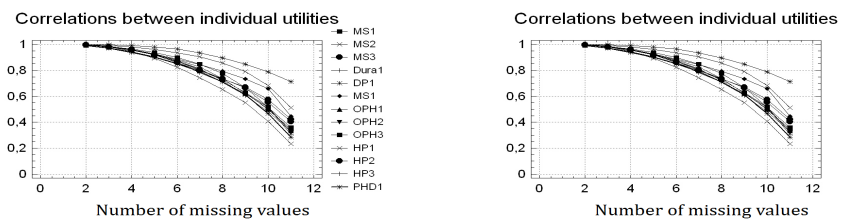


Fig. 1. R^2 *in the case of total order with and with no weight functions*

3.2.2 Study of the individual utilities

Benammou et al (2003) showed that the individual utilities remain stable when the number of missing values does not exceed six -that is half of the classified scenarios. The use of the weight functions improves considerably this result (see Fig. 2.). We notice that in the case of the weight functions of the form f(x)= $\frac{1}{(X)^n}$ and when the number of missing values equals 8 the correlation between individual utilities in the case of total order and the partial orders is $\geq 0,8$. It increases slightly with the decreasing speed of the weight function and stabilizes when n exceeds 4. For the same functions when $n > 2$ the results remain useful with 9 missing values (the correlation between individual utilities reaches 0.9 for some factor levels).



with no weight *weighted by* $\frac{1}{X^3}$

Fig. 2. *Correlation between individual utilities in case of total order and various partial orders with and without weight functions*

3.2.3 Study of the importance's of the utilities

Benammou et al (2003) showed that the importance of the factor utilities remain stable when the number of missing values does not exceed six. The use of the weight functions improves slightly this result (see fig. 3.). The importance of factors remains stable when the number of missing values is not very important (about 7). This stability degrades when this number increases. This is due to a strong correlation between these importances (and thoose of the total order case).

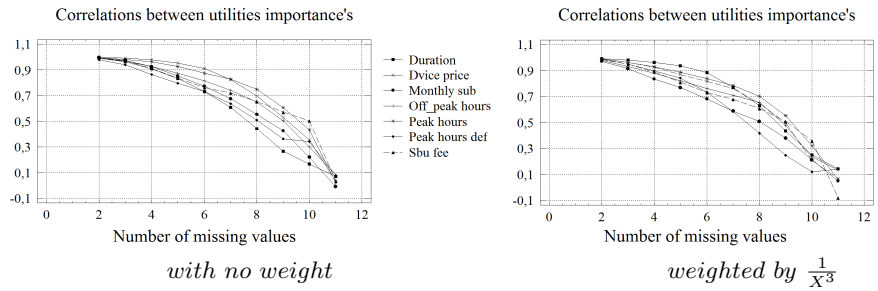


Fig. 3. *Correlation between importance's utilities in case of total order and various partial orders with and without weight functions*

4 Simulation

To generalize our results and to be able to compare them with the existing literature, we conduct a simulation study analogous to that used by Benammou et al (2003), where the authors have simulated ranks of classifications in a systematic way. To have a realistic model, we make choice of an utilities system and simulate ranks that are compatible and coherent adding noises ε with a given standard deviation to X β.

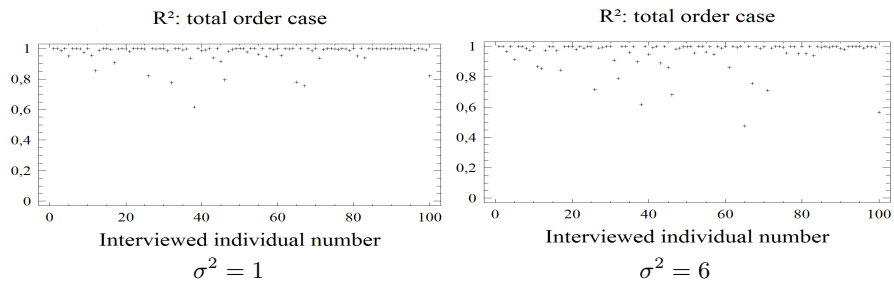
4.1 R^2 study

The simulation of the classification rank is based on the coefficients of a real data model.This is done in order to guarantee the coherence of the simulated data with a multiple correlation coefficient close to 1. We give in Fig.4. the values of R^2 for various values of σ^2 for the weight function $\frac{1}{X^3}$ which seems to give the best adjustment.

We point out that R^2 values are close to 0,99 for the majority of the individuals. Other functions such as $f(x) = \frac{1}{2^x}$, $f(x) = \exp(-2x)$ or $f(x) = \exp(\frac{1}{x})$ give comparable results.

4.2 Analysis of the individual utilities

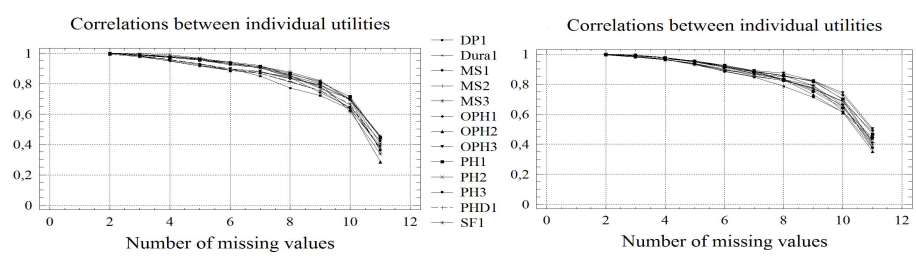
The individual utilities remain stable even for an important number of missing values (about 8). This stability decreases when the number of missing values increases (see Fig. 5.). The results improve slightly when σ^2 increases.



$\sigma^2 = 1$ $\sigma^2 = 6$

Fig. 4. *R^2 in the case of total order for different values of σ^2 weighted by $\frac{1}{X^3}$*

The use of weight functions yields superior results to those obtained by Benammou et al (2003). As an example, we give, in Fig. 5. the correlations between individual utilities for various values of σ^2 in the case of the weight functions $\frac{1}{X^3}$.

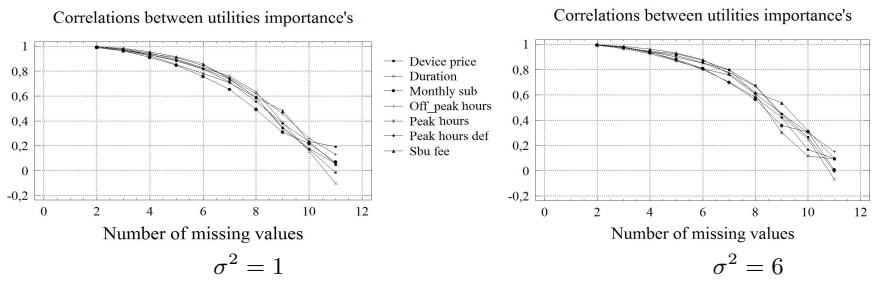


$\sigma^2 = 1$ $\sigma^2 = 6$

Fig. 5. *Correlation between individual utilities in case of total order and various partial orders for different values of σ^2 weighted by $\frac{1}{X^3}$*

4.3 Analysis of the importance's utilities

Importance's utilities remain stable for a large number of missing values (about 8). This stability decreases when the number of missing values increases (see Fig. 6.). The results improve slightly when σ^2 increase. The use of the weight functions improves the results obtained by Benammou et al (2003). As an example, we give (Fig. 6.) the correlations between the importances of the utilities for various values of σ^2 weighted by $\frac{1}{X^3}$



$\sigma^2 = 1$ $\sigma^2 = 6$

Fig. 6. *Correlation between importance's utilities in case of total order and various partial orders for different values of* σ^2 *weighted by* $\frac{1}{X^3}$.

5 Conclusion

The case of partial ordering in conjoint analysis, is very frequent especially when the number of proposed product exceeds ten. In this paper, we propose the use of the weighted least squares to estimate the parameters. Our experimentations showed a good stability of the results under the three quarter ranked scenarios.We confirm our findings by simulation. It should be remarqued that the results are better when the weight functions decreasing speed is faster.

References

Benammou, S., Harbi, S. and Saporta, G.(2003): Sur l'utilisation de l'analyse conjointe en cas de réponses incompltes ou de non réponses. *Revue de Statistique Appliquée, 51, 31-55.*

Benammou, S. , Saporta, G and Souissi, B.(2007): Une procédure de réduction du nombre de paires en analyse conjointe. *Journal de la Société Française de Statistique, 148, (4).57- 76.*

Cattin, P. and Wittink, D.R. (1989): Commercial Use of Conjoint analysis: An Update. *Journal of Marketing, 53, .91-96.*

Green, P.E. and Srinivasan, V. (1990): Conjoint analysis in marketing: new developments with implications for research and practice. *Journal of Marketing, 3-19.*

Green, P.E. and Rao, V.R.(1971): Conjoint Measurement for Quantifying Judgment Data. *Journal of Consumers Research, 5 September, 103-123.*

Weisberg, S. (1985): *Applied linear regression.* John Wiley and Sons, inc, second edition, New York.