PLISKA
STUDIA MATHEMATICA
BULGARICA
ПЛИСКА
БЪЛГАРСКИ
МАТЕМАТИЧЕСКИ
СТУДИИ

For further information on
Pliska Studia Mathematica Bulgarica
visit the website of the journal http://www.math.bas.bg/~pliska/
or contact: Editorial Office
Pliska Studia Mathematica Bulgarica
Institute of Mathematics and Informatics
Bulgarian Academy of Sciences
Telephone: (+359-2)9792818, FAX:(+359-2)971-36-49
e-mail: pliska@math.bas.bg

Pliska Stud. Math. Bulgar. **17** (2005), 71–84

PLISKA
STUDIA MATHEMATICA
BULGARICA

ANALYZING CONJOINT ANALYSIS DATA BY A RANDOM COEFFICIENT REGRESSION MODEL

Roberto Furlan, Roberto Corradetti

Since late 1960s conjoint analysis has been applied in estimating consumer preferences in marketing research [8]. This article discusses how to model the data coming from a full or a fractional factorial design within a unique regression model, as an alternative to the estimation done by n independent multiple linear regression models, one for each subject. The advantage of the method presented here resides in the possibility of computing correct standard errors for the conjoint analysis utility values based on a particular group of subjects [7]. The model assumes that the utility values within subjects could be correlated.

1. Introduction

Conjoint analysis (CA) is a multivariate statistical analysis technique based on the study of the joint effects on consumers of the elements that compose a product or service.

It was first developed in the early 1970s by Green and Rao [5] which got advantage by Sidney Addelman's efforts spent in developing methods for determining fractional factorial designs [1] [2]. Since then, CA has received more and more academic and private sector attention. As a result, after Green and Rao's initial presentation, many articles have been written about CA, and several computer programs implementing CA have been put on the market. One of the most

2000 *Mathematics Subject Classification:* 62J12, 62K15, 91B42, 62H99

Key words: conjoint analysis, random coefficient regression model, full factorial design, fractional factorial design, design matrix

important statistical problems concerning CA studies resides in the CA design. The choice of the design depends strongly on the characteristics of the study, such as the number of variables involved, the potential presence of interaction effects among the variables, the ability and the motivation of the target respondents to the CA survey, the statistical ability of the researcher, and the availability of software.

One of the most common ways to prepare a CA study is by using a full factorial design or a fractional factorial design [6] [8]. In this CA design, each subject participating in the study evaluates all possible profiles or a particular subset of them. The analysis is done at the subject level by fitting a different multiple linear regression model for each subject [12]. Eventually, the researcher can combine the n vectors of utilities coming from different subjects in order to get a vector of utilities for the whole sample or for a particular group of subjects. Analogously, the researcher can combine the covariance matrices (necessary for the computation of the standard errors and of the confidence intervals for the utility values) of different subjects in order to summarize the information for a particular group of subjects. The problem with this approach consists in the possibility of incorrect computation of the standard error for each CA utility value based on a particular group of subjects. As a matter of fact, this approach assumes that the utility values are not correlated within subjects. This is a strong assumption that usually does not hold true. More correct standard errors for the utility values based on a particular group of subjects can be computed by assuming that the utility values within subjects could be correlated. In this article the authors solve this problem by using a random coefficient regression model to analyze the data from a full or a fractional factorial design within a unique regression model.

The necessity to adopt a model able to deal with correlated data within subject is shown in this article by an application using data from a CA survey conducted in Italy in Spring 2002.

2. The Traditional Setup

First, let

$$i = 1, \ldots, n \tag{1}$$

denote the set of n subjects evaluating a particular set of p profiles, each of which is composed by f factors. In a full or fractional factorial design, the multiple linear regression model for the i^{th} subject is

$$y_i = X\beta_i + \varepsilon_i, \tag{2}$$

where $y_i = (y_{i1}, , y_{ip})'$ is a p-dimensional vector of responses for the i^{th} subject, whose generic j^{th} element contains the evaluation of the profile j; X is a design matrix of the type (p x T), whose creation is explained below in the section 'The Design Matrix X;' β_i is a T-dimensional vector of parameters for the i^{th} subject; and ε_i is a p-dimensional vector, whose j^{th} element represents the random error for the subject i associated with the profile j.

Note that in a full factorial design, p is the number of possible profiles, while in a fractional factorial design, p represents the number of profiles presented for that particular fraction of the factorial design [4][9][12].

In this setup, the assumptions are: (a) $\varepsilon_i \sim normal(0, \sigma_i^2 I_p)$; (b) $Rank(X) = T$; and (c) $T < p$, where T, the number of parameters in the mean model, is assumed to be fixed. Under these assumptions, by the linear models theory [3][10][11], it is known that the least squares estimators of the β_i's for the model (2) are

$$\widehat{\beta}_i = (X'X)^{-1}X'y_i \tag{3}$$

and that each $\widehat{\beta}_i$ is distributed as a $normal(\beta_i, \sigma_i^2(X'X)^{-1})$. The unknown quantity σ_i^2 can be estimated by the unbiased estimator:

$$\widehat{\sigma}_i^2 = \frac{1}{p-T}(y_i'y_i - \widehat{\beta}_i{}'X'y_i). \tag{4}$$

In this model each vector of responses $\widehat{y}_i$ is distributed as a $normal(X\beta_i, \sigma_i^2 I_p)$.

Usually, in order to summarize the CA results, the analysis of the data includes the estimation of the vector of utilities $\widehat{\beta}$ for a particular group of n subjects, obtained as a linear combination of the n least squares estimators $\widehat{\beta}_i$'s:

$$\widehat{\beta} = \frac{1}{n}\sum_{i=1}^{n} \widehat{\beta}_i. \tag{5}$$

Note that $\widehat{\beta}$ is the best linear unbiased estimator (BLUE) of β. In addition, the analysis usually includes the estimation of the covariance matrix for $\widehat{\beta}$:

$$Cov(\widehat{\beta}) = \frac{1}{n^2}\sum_{i=1}^{n}(\sigma_i^2(X'X)^{-1}), \tag{6}$$

which is done by plugging-in the estimated values for the σ_i^2's.

As stated in the introduction, the limitation of this approach consists in the frequent incorrect computation of the standard errors for the CA utilities based on a group of n subjects, due to the fact that the β_i's are assumed to be uncorrelated within subjects, which in fact is not often the case.

3. The Design Matrix X

The design matrix X contains the description of the CA profiles, each of which comes from a particular combination of factors' levels. Here, the authors develop Green and Srinivasan's idea [6] about the creation of pseudo-attributes, which are the values that represent the columns of the design matrix X in this approach. The nature of the factors, either qualitative or quantitative, involved in the CA study has to be kept in consideration.

A qualitative (or discrete) factor with g levels is represented in the matrix X by $(g-1)$ columns. After assuming that the levels within the factor have been ordered in an arbitrary way, a value is assigned to each position X_{jc} (j^{th} row, c^{th} column of X), according to the level observed for the profile j with respect to each column. In particular, X_{jc} assumes '-1' if the level g is observed for the profile j; '1' if the level c is observed for the profile j; and '0' if the level c is not observed for the profile j. Note that other alternative ways of defining X_{jc} are possible, but the one used here is convenient for its particular interpretation, for when each level of an attribute appears with equal frequency in the CA design, the sum of the utilities of each level is zero, so that the model is centered on zero.

A quantitative factor with g levels, is represented in the matrix X by $(g-1)$ columns. The generic value X_{jc} is obtained by

$$X_{jc} = \left(l_{j0} - g^{-1}\sum_{e=1}^{g} l_e\right)^c, c = 1, \ldots, (g-1) \tag{7}$$

where l_{jo} is the level observed of the factor for the profile j, and l_e is the e^{th} level of the factor, $e = 1, \ldots, g$.

In a CA study characterized by *df* qualitative factors and by pf quantitative factors, the design matrix X is of the type $(p \times T)$, where T, the number of columns of X, is obtained by

$$T = 1 + \sum_{m=1}^{df}(l_m - 1) + \sum_{m=1}^{pf}(g_m - 1) \tag{8}$$

where lm is the number of levels of the m^{th} qualitative factor $(m = 1, \ldots, df)$, g_m is the number of levels of the m^{th} quantitative factor $(m = 1, \ldots, pf)$, and $df + pf = f$ is the total number of factors included in the study.

Note that the first column of the design matrix X is a p-dimensional vector of 1's, which has been included for estimating the model intercept μ.

4. The Random Coefficient Regression Model

Instead of estimating the regression coefficients separately for each subject by fitting the model (2) n times, it is possible to analyze all the data available within a unique model. In order to do this, the authors consider the version of the random coefficient regression model (RCR) presented by Gumpertz and Pantula [7], which is a special case of the general mixed linear models. The RCR model is expressed by

$$y_i = X\beta_i + \varepsilon_i \tag{9}$$

where $y_i = (y_{i1}, \ldots, y_{ip})'$ is a p-dimensional vector of responses for the i^{th} subject, whose generic j^{th} element contains the evaluation of the profile j; X is a design matrix of the type $(p \times T)$; β_i is a T-dimensional vector of parameters for the i^{th} subject; and ε_i is a p-dimensional vector, whose j^{th} element represents the random error for the subject i associated with the profile j.

In the present setup the model is fitted not subject by subject, but for all the n subjects at the same time. Here, the n subjects are assumed to be randomly selected from a larger population. Therefore, the n regression coefficient vectors β_i's can be seen as random drawings from a T-dimensional population of parameters, from which the name RCR model derives. The assumptions for the RCR model are: (a) $\beta_i \sim^{iid} normal(\beta, \Sigma)$, where Σ is a nonsingular $(T \times T)$ covariance matrix; (b) $\varepsilon_i \sim^{iid} normal(0, \sigma_i^2 I_p)$; (c) β_i and $\varepsilon_{i'}$ are independent random variables for all i and i'; (d) $Rank(X) = T$; and (e) $T < n$ and $T < p$, where T, the number of parameters in the mean model, is assumed to be fixed. Under these assumptions, the least squares estimators of the β_i's for the RCR model are the same as in (3). By the linear models theory [3] [10] [11], the distributions of $\widehat{\beta}_i$, $\widehat{y}_i$, and $\widehat{\beta}$ can be easily derived:

$$\widehat{\beta}_i \sim^{iid} normal(\beta, \Sigma + \sigma_i^2 (X'X)^{-1}) \tag{10}$$

$$\widehat{y}_i \sim^{iid} normal(X\beta, X\Sigma X' + \sigma_i^2 I_p) \tag{11}$$

$$\widehat{\beta} \sim^{iid} normal\left(\beta, \frac{\Sigma}{n} + \frac{1}{n^2}\sum_{i=1}^{n}(\sigma_i^2 (X'X)^{-1})\right) \tag{12}$$

Note that in these distributions β, Σ, and σ_i^2 are unknown quantities. Gumpertz and Pantula [7] show that for the RCR model, the estimator $\widehat{\beta}$ in (5) is the BLUE of β. Note that an unbiased estimator of σ_i^2 is still provided by (4). Also, from their results, it is possible to derive an unbiased estimator of Σ:

$$\widehat{\Sigma} = \frac{1}{n-1}\sum_{i=1}^{n}(\widehat{\beta}_i - \widehat{\beta})(\widehat{\beta}_i - \widehat{\beta})' - \frac{1}{n}\sum_{i=1}^{n}(\widehat{\sigma}_i^2 (X'X)^{-1}) \tag{13}$$

The covariance matrix for $\widehat{\beta}$ in the RCR model provided in (12) can also be written as a function of the covariance matrix for $\widehat{\beta}$ in the traditional model provided in (6):

$$Cov(\widehat{\beta}) = \frac{\Sigma}{n} + Cov(\widehat{\beta})_{\text{traditional setup}} \tag{14}$$

This expression indicates that the covariance matrix $Cov(\widehat{\beta})$ computed by the RCR model is always different than that computed by the traditional model, except for the case when Σ is a matrix of zeros. In this particular case, verified if and only if the utility values within subjects are uncorrelated, the two models provide the same results.

The elements of $Cov(\widehat{\beta})$ allow the computation of the standard errors for the CA utility parameters based on the n subjects. Consider the notation $\widehat{\beta}_{km}$ for the element in $\widehat{\beta}$ which refers to the level m of the factor k. Analogously, consider the notation β_{km} for the element in β which refers to the level m of the factor k. Also, consider the notation $Cov(km, k'm')$ for referring to the element in $Cov(\widehat{\beta})$ which corresponds to the covariance between $\widehat{\beta}_{km}$ and $\widehat{\beta}_{k'm'}$, where m is a level of the factor k, and m' is a level of the factor k'. Note that by this notation, $Cov(km, km)$ is the variance for $\widehat{\beta}_{km}$. Also, note that the first element of the diagonal of $Cov(\widehat{\beta})$ represents the variance for $\widehat{\mu}$, which is the estimated value for the intercept μ.

In order to get the final CA utilities for the factors levels and the standard errors associated with each parameter, the following procedure is used. For a qualitative factor k with g levels, there are $(g-1)$ regression coefficients β_{km}, where $m = 1, \ldots, (g-1)$. The estimation of the parameter associated with the last level of the factor k is obtained by adding together the first $(g-1)$ parameters and by inverting the sign of the result:

$$\widehat{\beta}_{km} = \begin{cases} \widehat{\beta}_{km}, & \text{for m=1, \ldots, g-1;} \\ -1\left(\sum\limits_{m=1}^{g-1} \widehat{\beta}_{km}\right), & \text{for m=g.} \end{cases} \tag{15}$$

The standard error associated with the first $(g-1)$ parameters of the factor k are directly obtained from the diagonal of $Cov(\widehat{\beta})$, while the standard error associated with the last parameter of the factor k can be derived as follows:

$$SE(\widehat{\beta}_{km}) = \begin{cases} Cov(km, km)^{1/2}\text{,for m=1, \ldots, g-1;} \\ \left(\sum\limits_{m=1}^{g-1} Cov(km, km) + 2\sum\limits_{m=1}^{g-1}\sum\limits_{m'<m} Cov(km, km')\right)^{1/2}\text{,for m=g.} \end{cases} \tag{16}$$

For a quantitative factor k with g levels there are $(g-1)$ regression coefficients β_{km}, where $m = 1, \ldots, (g-1)$. The utility value for any level of the factor can be estimated, if it belongs to the range of levels of the same factor included in the CA study. By assuming that the levels included in the CA design for the factor k range between l_1 and l_2, then the utility value for the level $l_0 \in [l_1, l_2]$ is obtained by

$$\widehat{\beta}_{kl_0} = \sum_{m=1}^{c} [\Delta_0^m \widehat{\beta}_{km}] \tag{17}$$

where

$$\Delta_0 = \left(l_0 - g^{-1} \sum_{e=1}^{g} l_e \right) \tag{18}$$

is the difference between the level l_0 of the factor k and the average of all the values included in the CA study for that factor. In (18), l_e represents the e^{th} level included in the study for the factor k, with $e = 1, \dots, g$. The standard error for $\widehat{\beta}_{kl_0}$ is provided by

$$SE(\widehat{\beta}_{kl_0}) =$$

$$\left(\sum_{m=1}^{c} \Delta_0^{2m} Cov(km, km) + 2\Delta_0^{m} \Delta_0^{m'} \sum_{m=1}^{g-1} \sum_{m'<m}^{g-1} Cov(km, km') \right)^{1/2} . \tag{19}$$

Finally, the final CA utilities associated with the factors levels are obtained by adding $1/f$ of $\widehat{\mu}$ to each parameter $\widehat{\beta}_{km}$ and $\widehat{\beta}_{kl_0}$.

5. An Application

The authors report some empirical results from real world data to illustrate the application of the RCR model to a resolution III fractional factorial design, and to point out the difference of the results obtained by this new approach and the ones obtained by a traditional model. The CA survey analyzed here was conducted in Italy, in Spring 2002. The object of the survey was a non-digital camcorder. Three brands (JVC, Sony, and Panasonic), three formats (Video 8 mm, Compact VHS, and Digital 8), two levels for the zoom (10x optical - 40x digital and 20x optical - 80x digital), two types of viewfinder (B/W and color), three levels for the LCD screen (absent, 2.5-inch color, and 3.5-inch color), and four levels for the price after taxes (EUR 500.00, EUR 650.00, EUR 800.00, and EUR 950.00) have been selected for the study. The design chosen for the CA study was a main-effects-only fractional factorial design with $p = 16$ profiles, each to be evaluated by all the subjects.

Respondents were chosen among the free-lancers of an Italian marketing research company. The sample size was fixed in 80 subjects; however, at the end of the fieldwork 87 respondents took part in the study. None of them had an earlier knowledge of any CA methodology. A personal briefing has been provided to each respondent before the beginning of the interview. A card containing all the variables/levels involved with a short description has been presented during the briefing and left on the operative table throughout the interview. The subjects have been instructed to consider only the factors included in the CA; the same conditions for the omitted attributes had to be considered. The subjects' task was to rate each of the 16 camcorder alternatives on a scale 1:100. Even thought they have been instructed in order to face correctly the CA tasks, a supervisor was always on place to provide technical help. The fieldwork required 16 days during Spring 2002. Each day no more than six interviews were completed.

All attributes considered are qualitative, except the price that is quantitative. From the full set of profiles, the 16 profiles shown in Tabl.1 have been selected. Hence, according to (8) the design matrix X has $T = 12$ columns, as shown in Tabl.2. In this table, the authors use the notation X_{km} to indicate the column in the design matrix X, which represents the m^{th} column created for the factor k. Note that X_{00} indicates the column created for the intercept μ.

Profile #	Brand	Format	Zoom	Viewfinder	LCD	Price
1	Samsung	Video 8	20x-80x	B/W	absent	650.00
2	JVC	VHS-C	10x-40x	color	absent	650.00
3	Sony	Video 8	10x-40x	B/W	absent	800.00
4	Samsung	VHS-C	10x-40x	B/W	2.5"	950.00
5	Sony	VHS-C	20x-80x	B/W	3.5"	500.00
6	JVC	Video 8	20x-80x	B/W	absent	950.00
7	Sony	Digital 8	10x-40x	color	absent	950.00
8	Sony	Video 8	20x-80x	color	2.5"	650.00
9	Samsung	Digital 8	20x-80x	color	absent	500.00
10	JVC	Video 8	10x-40x	color	2.5"	500.00
11	Samsung	Video 8	10x-40x	color	3.5"	800.00
12	JVC	Digital 8	20x-80x	B/W	2.5"	800.00
13	JVC	Video 8	20x-80x	color	3.5"	950.00
14	JVC	VHS-C	20x-80x	color	absent	800.00
15	JVC	Video 8	10x-40x	B/W	absent	500.00
16	JVC	Digital 8	10x-40x	B/W	3.5"	650.00

Table 1: Profiles selected for the study

By considering the RCR model (9) and by applying (3), the authors got the least squares estimators $\widehat{\beta}_i$'s (for $i = 1, \ldots, 87$). By applying (4), the authors got unbiased estimators of the σ_i^2's, while by applying (13), the authors got the unbiased estimator of Σ, that is $\widehat{\Sigma}$:

Profile#	X_{00}	X_{11}	X_{12}	X_{21}	X_{22}	X_{31}	X_{41}	X_{51}	X_{52}	X_{61}	X_{62}	X_{63}
1	1	-1	-1	1	0	-1	1	1	0	-75	$(-75)^2$	$(-75)^3$
2	1	1	0	0	1	1	-1	1	0	-75	$(-75)^2$	$(-75)^3$
3	1	0	1	1	0	1	1	1	0	75	$(75)^2$	$(75)^3$
4	1	-1	-1	0	1	1	1	0	1	225	$(225)^2$	$(225)^3$
5	1	0	1	0	1	-1	1	-1	-1	-225	$(-225)^2$	$(-225)^3$
6	1	1	0	1	0	-1	1	1	0	225	$(225)^2$	$(225)^3$
7	1	0	1	-1	-1	1	-1	1	0	225	$(225)^2$	$(225)^3$
8	1	0	1	1	0	-1	-1	0	1	-75	$(-75)^2$	$(-75)^3$
9	1	-1	-1	-1	-1	-1	-1	1	0	-225	$(-225)^2$	$(-225)^3$
10	1	1	0	1	0	1	-1	0	1	-225	$(-225)^2$	$(-225)^3$
11	1	-1	-1	1	0	1	-1	-1	-1	75	$(75)^2$	$(75)^3$
12	1	1	0	-1	-1	-1	1	0	1	75	$(75)^2$	$(75)^3$
13	1	1	0	1	0	-1	-1	-1	-1	225	$(225)^2$	$(225)^3$
14	1	1	0	0	1	-1	-1	1	0	75	$(75)^2$	$(75)^3$
15	1	1	0	1	0	1	1	1	0	-225	$(-225)^2$	$(-225)^3$
16	1	1	0	-1	-1	1	1	-1	-1	-75	$(-75)^2$	$(-75)^3$

Table 2: The Design Matrix X

$$
\begin{gathered}
\widehat{\Sigma} = \left(\begin{array}{rrrrrrrrrr}
20.9435 & -1.2325 & -2.5638 & -1.1728 & 0.3460 & 1.8252 & 5.6801 & 1.7130 & -0.8012 & \ldots \\
 & 9.9551 & -3.2650 & 0.2231 & -1.4348 & -1.8327 & -1.1614 & -3.6853 & 1.7933 & \ldots \\
 & & 8.8213 & 0.6652 & 0.3042 & 1.1387 & -0.0578 & 2.2703 & -0.2594 & \ldots \\
 & & & 8.5861 & -6.7384 & -1.0997 & -0.5871 & -1.0729 & 1.1328 & \ldots \\
 & & & & 16.0006 & 1.6188 & -0.5052 & 0.6788 & -2.0393 & \ldots \\
 & & & & & 1.4920 & 1.1229 & 0.8907 & -0.4397 & \ldots \\
 & & & & & & 8.0790 & 0.4642 & 0.1508 & \ldots \\
 & & & & & & & 11.8721 & -4.8954 & \ldots \\
 & & & & & & & & 1.8626 & \ldots \\
 & & & & & & & & & \ldots
\end{array}\right. \\
\left.\begin{array}{rrrr}
\ldots & -6.6\cdot10^{-3} & -1.1\cdot10^{-5} & 2.2\cdot10^{-7} \\
\ldots & -1.6\cdot10^{-2} & 3.8\cdot10^{-6} & 2.1\cdot10^{-8} \\
\ldots & 2.1\cdot10^{-2} & -3.4\cdot10^{-6} & -7.5\cdot10^{-8} \\
\ldots & -8.9\cdot10^{-3} & -8.0\cdot10^{-6} & 1.4\cdot10^{-7} \\
\ldots & 1.3\cdot10^{-2} & -2.3\cdot10^{-6} & -3.9\cdot10^{-7} \\
\ldots & -5.2\cdot10^{-3} & -6.5\cdot10^{-6} & 1.0\cdot10^{-7} \\
\ldots & -1.2\cdot10^{-2} & -1.5\cdot10^{-6} & 4.0\cdot10^{-8} \\
\ldots & -7.5\cdot10^{-3} & 1.6\cdot10^{-5} & 8.8\cdot10^{-8} \\
\ldots & -1.8\cdot10^{-3} & -6.6\cdot10^{-6} & 3.2\cdot10^{-8} \\
 & 6.1\cdot10^{-4} & 1.4\cdot10^{-7} & -1.9\cdot10^{-9} \\
 & & 1.7\cdot10^{-10} & -2.7\cdot10^{-12} \\
 & & & 3.1\cdot10^{-15}
\end{array}\right)
\end{gathered}
\tag{20}
$$

Note that for the sake of clarity, only the upper-right triangle of the symmetric matrix $\widehat{\Sigma}$ has been displayed in (20). The same visualization will be used again in this section. Therefore, according to (12) $\widehat{\beta}$ is normally distributed with mean β, whose estimate is obtained by (5):

$$\widehat{\beta} = (83.49 \quad -1.46 \quad 1.88 \quad -.71 \quad -.23 \quad -.46 \quad -2.28 \quad -3.33 \quad 1.54 \quad -2.02 \cdot 10^{-2} \quad 1.45 \cdot 10^{-7})' \tag{21}$$

and with covariance matrix $Cov(\widehat{\beta})$, whose estimate is obtained by (14):

$$Cov(\widehat{\beta}) = \left(\begin{array}{cccccccccc}
0.2732 & -0.0192 & -0.0270 & -0.0185 & 0.0065 & 0.0210 & 0.0653 & 0.0147 & -0.0067 & \ldots \\
 & 0.1344 & -0.0475 & 0.0026 & -0.0165 & -0.0211 & -0.0133 & -0.0424 & 0.0206 & \ldots \\
 & & 0.1288 & 0.0076 & 0.0035 & 0.0131 & -0.0007 & 0.0261 & -0.0030 & \ldots \\
 & & & 0.1186 & -0.0874 & -0.0126 & -0.0067 & -0.0123 & 0.0130 & \ldots \\
 & & & & 0.2113 & 0.0186 & -0.0058 & 0.0078 & -0.0234 & \ldots \\
 & & & & & 0.0284 & 0.0129 & 0.0102 & -0.0051 & \ldots \\
 & & & & & & 0.1041 & 0.0053 & 0.0017 & \ldots \\
 & & & & & & & 0.1564 & -0.0662 & \ldots \\
 & & & & & & & & 0.0488 & \ldots \\
 & & & & & & & & & \ldots
\end{array} \right.$$

$$\left. \begin{array}{cccc}
\ldots & -7.6 \cdot 10^{-5} & -7.4 \cdot 10^{-7} & 2.5 \cdot 10^{-9} \\
\ldots & -1.8 \cdot 10^{-4} & 4.3 \cdot 10^{-8} & 2.5 \cdot 10^{-10} \\
\ldots & 2.4 \cdot 10^{-4} & -3.9 \cdot 10^{-8} & -8.6 \cdot 10^{-10} \\
\ldots & -1.0 \cdot 10^{-4} & -9.2 \cdot 10^{-8} & 1.6 \cdot 10^{-9} \\
\ldots & 1.5 \cdot 10^{-4} & -2.7 \cdot 10^{-8} & -4.5 \cdot 10^{-9} \\
\ldots & -5.9 \cdot 10^{-5} & -7.5 \cdot 10^{-8} & 1.2 \cdot 10^{-9} \\
\ldots & -1.4 \cdot 10^{-4} & -1.7 \cdot 10^{-8} & 4.6 \cdot 10^{-10} \\
\ldots & -8.6 \cdot 10^{-5} & 1.9 \cdot 10^{-7} & 1.0 \cdot 10^{-9} \\
\ldots & -2.1 \cdot 10^{-5} & -7.6 \cdot 10^{-8} & 3.7 \cdot 10^{-10} \\
 & 1.2 \cdot 10^{-5} & 1.6 \cdot 10^{-9} & -1.2 \cdot 10^{-10} \\
 & & 2.4 \cdot 10^{-11} & -3.1 \cdot 10^{-14} \\
 & & & 2.2 \cdot 10^{-15}
\end{array} \right) \tag{22}$$

From (20) it is evident that $\widehat{\Sigma}$ is not a matrix of zeros. It implies that the β_i's are correlated within subjects, and thus the standard errors should be computed by the RCR model and not by the traditional model. As a matter of fact, the RCR model provides the matrix $Cov(\widehat{\beta})$ displayed in (22), which is substantially different from the same matrix computed by the expression (6) relative to the traditional model:

$$Cov(\widehat{\beta}) = \left(\begin{array}{rrrrrrrrrl} 0.0325 & -0.0050 & 0.0025 & -0.0050 & 0.0025 & 0 & 0 & -0.0050 & 0.0025 & \ldots \\ & 0.0199 & -0.0100 & 0 & 0 & 0 & 0 & 0 & 0 & \ldots \\ & & 0.0274 & 0 & 0 & 0 & 0 & 0 & 0 & \ldots \\ & & & 0.0199 & -0.0100 & 0 & 0 & 0 & 0 & \ldots \\ & & & & 0.0274 & 0 & 0 & 0 & 0 & \ldots \\ & & & & & 0.0112 & 0 & 0 & 0 & \ldots \\ & & & & & & 0.0112 & 0 & 0 & \ldots \\ & & & & & & & 0.0199 & -0.0100 & \ldots \\ & & & & & & & & 0.0274 & \ldots \\ & & & & & & & & & \ldots \\ & & & & & & & & & \end{array} \right.$$

$$\left. \begin{array}{lrrr} \ldots & 0 & -6.23 \cdot 10^{-7} & 0 \\ \ldots & 0 & 0 & 0 \\ \ldots & 0 & 0 & 0 \\ \ldots & 0 & 0 & 0 \\ \ldots & 0 & 0 & 0 \\ \ldots & 0 & 0 & 0 \\ \ldots & 0 & 0 & 0 \\ \ldots & 0 & 0 & 0 \\ \ldots & 0 & 0 & 0 \\ & 5.06 \cdot 10^{-6} & 0 & -1.01 \cdot 10^{-10} \\ & & 2.22 \cdot 10^{-11} & 0 \\ & & & 2.19 \cdot 10^{-15} \end{array} \right) \tag{23}$$

Hence, by (15) one can easily get the parameter value associated with the last level of each qualitative factor, while by (17) one can estimate the parameter value for any level $l_0 \in [500, 950]$ of the quantitative factor price. In addition, (16) and (19) provide the way to get the standard error from the values in (22) for any parameter $\widehat{\beta}_{km}$ and $\widehat{\beta}_{kl_0}$.

6. Conclusions

By the approach developed in this article it is possible to get an estimate of the covariance matrix for the vectors of coefficients $\widehat{\beta}$, based on the assumption that the correlation of the utility values within subjects can be nonzero. As a consequence, it is possible to obtain more correct estimates of the standard errors of the CA utility values based on the whole sample or on a particular group of subjects. In addition, the estimators provided here are unbiased. Therefore, in this article the authors provided a useful tool that should be adopted whenever the researcher decides to use a full or a fractional factorial design for a CA study.

A further research could investigate the application of a RCR model in the case when the design matrix X is not the same for all the subjects, the situation when the subjects receive different sets of profiles. In this generalization, the estimator of β is the weighted least squares estimator of the β_i's, where the weights are the inverse covariance matrices of β_i's.

REFERENCES

[1] Addelman, S. Orthogonal Main-Effect Plans for Asymmetrical Factorial Experiments *Technometrics* **4(1)** (1962), 21–46.

[2] Addelman, S. Symmetrical and Asymmetrical Fractional Factorial Plans *Technometrics* **4(1)** (1962), 47–58.

[3] Christensen, R. Plane Answers to Complex Questions: The theory of Linear Models. New York: Springer-Verlag, 1996 .

[4] Green, P. E., D. J. Carroll, F. J. Carmone Some New Types of Fractional Factorial Designs for Marketing Experiments. In: Research in Marketing (ed. J. N. Sheth), Greenwich, CT: JAI Press (1978).

[5] Green, P. E., V. R. Rao Conjoint Measurement for Quantifying Judgemental Data *Journal of Marketing Research* **8** (1971), 355–63.

[6] Green, P. E., S. V. Srinivasan Conjoint Analysis in Consumer Research: Issue and Outlook *Journal of Consumer Research* **5** (1978), 103–123.

[7] Gumpertz, M. L., S. G. Pantula A Simple Approach to Inferences in Random Coefficient Models *The American Statistician* **43(4)** (1989), 203–210.

[8] Gustafsson, A., A. Hermann, F. Huber Conjoint Analysis as an Instrument of Market Research Practice. In: Conjoint Measurement: Methods and Applications (eds A. Gustafsson, A. Hermann, and F. Huber), Berlin: Springer (2001), 5–46.

[9] Hahn, G. J., S. S. Shapiro A Catalog and Computer Program for the Design and Analysis of Orthogonal Symmetric and Asymmetric Fractional Factorial Experiments *General Electric Research and Development Center* **66-C-165** May, Schenectady, NY, 1966.

[10] Rawlings, J. O., S. G. Pantula, D. A. Dickey Applied Regression Analysis: A Research Tool. New York: Springer-Verlag, 1998.

[11] Rencher, A. C. Linear Models in Statistics. New York: John Wiley & Sons, Inc. 2000.

[12] SPSS Conjoint 8.0. Chicago, 1997.

Roberto Furlan
GfK Martin Hamblin,
London, UK
e-mail: `roberto.furlan@libero.it`

Roberto Corradetti
University of Torino
Department of Statistics
and Applied Mathematics
'Diego de Castro',
Torino, Italy
e-mail: `roberto.corradetti@unito.it`