A NEW DIMENSION IN THE FACTORIAL TECHNIQUES: THE RESPONSE SURFACE

Giuseppe Giordano

Dipartimento di Scienze Economiche e Statistiche, Università di Salerno
ggiordan@unisa.it

Abstract

In this paper we propose to bring together some facets of the Response Surface Methodology used in Design of Experiments and the graphical displays arising from Multidimensional Data Analysis in order to enhance the interpretation of the traditional plots of the Factorial Techniques. By exploiting the peculiar interpretation of a response surface we describe a data-fitting based approach in order to analyse the relationships between an outer quantitative response variable and a set of principal axes. Moreover, we to use this interpretative-aid tool to represent by a surface response the information about the quality of the representation derived by a multidimensional data analysis.

Keywords: Contour Plot, Design of Experiment, Exploratory Data Analysis, Preference Map, Principal Component Analysis.

1. INTRODUCTION

In many application fields the problem of reducing a multidimensional construct is addressed through the use of multidimensional data analysis methods. The use of techniques such as Principal Component Analysis (Jolliffe, 2002), or Correspondence Analysis (Greenacre and Hastie, 1987), is nowadays very common for analysing survey data and made even easier by standard procedures in the most diffuse statistical packages.

In *soft*-modelling analysis, when multivariate data are structured as *response* and *predictor* variables, several multivariate techniques - such as *Reduced Rank Regression* (Anderson 1951), *Partial Least Squares* techniques (Wold, 1982, 1985), etc. - are better suited at both explorative and synthesis aims. Their use is widespread in the fields of chemometrics, ecology, psychometrics, etc.

All these different techniques share the powerful capability of a geometrical interpretation and a subsequent graphical representation. The construction of a reduced factorial subspace is a common device, very useful in assessing the relationships among statistical units and variables.

The different ways to plot the *modes* (rows and columns) of a data matrix can be covered under the framework of Multidimensional Scaling (MDS) which comprises a wide set of multidimensional data techniques applicable to a variety of research designs allowing for the analysis of any kind of distance or similarity matrix. A typical application area is *Marketing* where, for example, the problem of product positioning or customers' preference can be addressed by means of *perceptual maps*. These plots are able to visually display the perceptions of consumers or potential customers with respect to products, brands, competitors and so on.

There are various statistical methods that can be used to convert raw data collected in a survey into a perceptual map (Carroll, 1972). The *Unfolding* technique (Green and Carmone, 1969) and the *Conjoint Analysis* (a regression based estimation technique, Green and Srinivasan, 1990) are the most used. A combined approach is proposed in the *Factorial Conjont Analysis* (Lauro, Giordano, Verde, 1998).

The peculiarity of a Conjoint Analysis study lies in the collecting data phase where Design of Experiment theory is used to select the different stimuli to be evaluated. Usually this is done according to a fractional factorial plan. At a first stage, a screening design is used to estimate few main effects, therefore to allow for interactions and quadratic effect a Response Surface method is better suited, either to find improved process settings, or to make the product insensitive against non-controllable (outer) factors.

In the framework of DOE the use of the Response Surface Methodology (*RSM*, Box and Wilson, 1951) allows to analyse the relationships between the response variable and a set of input factors[1]. The analysis is carried out through successive steps of experimentation, modelling, data analysis and optimization. The aim is to obtain an accurate approximation of the response surface and to identify an optimum design region.

Several statistical packages allows to carry out the RSM (such as, fore example, the *ADX* menu system in *SAS/QC*). The typical graphical output is the three-dimensional representation of the surface and the *Contour Plot*. A contour plot is a graphical technique for representing a three-dimensional surface by plotting constant slices, called contours, on a two-dimensional format. That is, given a value for the response y, lines are drawn for connecting the (x_1, x_2) coordinates where that y value occurs. These lines are the iso-response values.

---

[1] *The use of the term "factor" can be misleading because of the different meaning in the frameworks of Design of Experiment and Factorial Analysis, we preserve the DOE terminology and will use the term "component" for the correspondent meaning in factorial techniques.*

The independent variables are usually restricted to a regular grid. The actual techniques for determining the correct iso-response values are rather complex and are almost always computer generated. Detailed routines can be found in *Matlab*, *SAS*, *S-plus*, *Statgraphics* and in many others general purpose graphics and mathematics programs.

The contour plot is generated for two factors. Typically, this would be the two most important factors as determined by previous analyses. It can also be generated a matrix of all pairwise contour plots for a number of important factors (similar to the scatter plot matrix for scatter plots).

In this paper we propose to consider the philosophy of the RSM as a further graphical resource in analysing and interpreting the results of a multidimensional data analysis.

Once the principal plan is drawn by means of the most meaningful components, we suggest to combine in the same plot a (three-dimensional) response surface, or a (bidimensional) contour plot, together with any two components considered as independent variables.

The proposed methodology apply both to symmetrical (unstructured data) and non-symmetrical data analysis, it will be discussed in section 2, while a Matlab procedure is given in section 2.1.

The dependent variables to consider in building the surface can be of various kind. By means of a case study (section 3), we distinguish an *internal* (3.1) and an *external* (3.2) analysis. In the first case, the surface represents information about the quality of the analysis results. For this kind of analysis we may wish to represent information such as a function of the *relative contribution* (Lebart et al., 1998). It is also possible to represent the response of a continuous variable used in the factorial decomposition. If we are interested in such a variable at illustrative purpose, we can study its behaviour according to its own surface shape.

In the external analysis we are interested to consider new information not previously involved in the determination of the factorial components. These external variables could help us to better understand the underlying phenomenon to be analysed. This approach is in the view of a *Principal Component Regression* (Kendall, 1957). Section 4 concludes.

2. THE RESPONSE SURFACE FACTORIAL MAP (RSFM)

The response surface methodology introduced by Box and Wilson (1951) is a collection of statistical and mathematical methods useful for modelling and analyzing engineering problems. In this technique, the main objective is to optimize the response surface that is influenced by various process parameters. Response Surface Methodology also quantifies the relationship between the controllable input parameters and the obtained response surfaces.

Usually the design procedure of response surface methodology is sequential as follows:

i) designing of a series of experiments for reliable measurement of the response of interest;

ii) developing a mathematical model of the second order response surface with the best fittings;

iii) finding the optimal set of experimental parameters that produce a maximum or minimum value of response;

iv) representing the main and interactive effects of process parameters through two and three dimensional plots.

The *Response Surface Plot* is a graphical device used in the procedure to represent and visually analyse how the experimental factors interact with the response variable. If all variables are assumed to be measurable, the response surface can be expressed as follows.

$$y = f\left(x_1, x_2, \ldots x_p\right) + \varepsilon \tag{1}$$

The goal is to optimize the response variable *y*. It is assumed that the independent variables are continuous and controllable by experiments with negligible errors. It is required to find a suitable approximation for the true functional relationship between independent variables and the response surface. Usually a second-order model is utilized in response surface methodology.

$$y = \beta_0 + \sum_{i=1}^{p} \beta_i x_i + \sum_{i=m+1}^{p} \beta_{ii} x_i^2 + \sum_{i=1}^{p-1} \sum_{j=i+1}^{p} \beta_{ij} x_i x_j + \varepsilon \tag{2}$$

where ε is a random error. In general the equation (2) can be stated in matrix form:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon} \tag{3}$$

where **y** is defined to be a vector of measured values, **X** to be a matrix of independent variables. The vectors β and ε consist of coefficients and errors, respectively. The solution of Eq. (3) can be obtained by the ordinary least square method.

Dealing with a number of independent factors *p* greater than *2*, it needs different plots for each pairs of them, it is also possible to represent a surface plot in a triangular reference system. The experimental facet of the methodology lead to an in-depth analyse of the surfaces in order to find critical points (minimum, maximum, saddle point) of the response curves. We refer to the work of **Myers** and Montgomery (2002) for further details on the RSM.

Now we are interested in apply the philosophy of the RSM to the *factorial* techniques defined in Multidimensional Data Analysis framework. We aims at showing how this kind of graphical tool may enrich the interpretation of a factorial technique results. The basic idea consists in an overlapping representation of a traditional factorial plan with a surface response derived from an idiosyncratic "*response variable*".

In this paper we main refer to the Principal Component Analysis as the leading multidimensional data analysis technique, however it should be straightforward to extend the proposed method to other multivariate techniques.

Let us consider the eigen-equation found in PCA carried out on the data matrix **Z** holding p standardised variables:

$$\frac{1}{n}\mathbf{Z}'\mathbf{Z}u_\alpha = u_\alpha \qquad \alpha = 1,\ldots,rank(\mathbf{Z}) \tag{4}$$

where **Z** is the $(n \times p)$ individual per variable data matrix, Λ the $(1/n)(\mathbf{Z'Z})$ is the correlation matrix, is the diagonal matrix holding the eigenvalues λ_α arranged in non-increasing order, the u_α are the α ortho-normalised eigenvectors associated to the correspondent α largest eigenvalues.

The factorial coordinates for variables and individuals are given, respectively, by:

$$\begin{cases} \varphi_\alpha = \sqrt{\lambda_\alpha}\, u_\alpha \\ \psi_\alpha = \mathbf{Z}u_\alpha \end{cases} \tag{5}$$

In many real-data applications, two-dimensional representations are enough for detecting patterns in the data. Some Authors (see Jolliffe, 2002, pp.80) stresses that the range of structures revealed by a PCA plot may be limited due to the orthogonality constraint imposed to the components and, for example, a linear relation between two PC's is impossible. However, if the p-dimensional data cloud lies close to a two-dimensional subspace, we can have a realistic visualisation of the most important aspect of the data structure.

The analysis of the factorial plan allows to visualise the correlation structure among the original variables and to focus on groups of individuals that share similar values for the same variables. In many application fields (for example, in Marketing, in a Customer Relationship Management strategy) we are also interested to the statistical units individually. In this case it needs to consider the relative contribution of the unit *i* to the variability along the α-th axis:

$$RC_{\alpha}\left(i\right)=\frac{1}{n}\frac{\psi_{\alpha i}^{2}}{\lambda_{\alpha}}=\frac{\psi_{\alpha i}^{2}}{\sum_{i=1}^{n}\psi_{\alpha i}^{2}} \tag{6}$$

where the $\psi_{\alpha i}$, according to the (5), are the coordinates of the *i*-th statistical units onto the α-th principal component, λ_{α} is the α-th eigenvalue which is a measure of the total inertia along the α-th principal component, it also represents the squared distance of the units *i* from the origin. Thus, the last term in (6) is expressed as the ratio between the squared distance of the *i*-th unit from the origin and the total inertia along the α-th axis. In the (6) we have considered the usual case of statistical units with equal mass (1/*n*), in such case the relative contribution is just proportional to the squared coordinate. This is not even true for all techniques, for example in Correspondence Analysis, where we deal with the rows and the columns of a contingency table, each "*unit*" is the modality of a categorical variable which is affected by a mass equal to its own marginal value (Lebart et al, 1998). As a general rule, one needs pay attention to statistical units with poor contribution in order to provide right interpretation.

2.1 BUILDING THE RESPONSE SURFACE FACTORIAL MAP

Once obtained the coordinates by eq. (5), we need to compute a suitable grid defined on the first factorial plan and fitting a response surface in the same way as in the common *RSM*. We describe a procedure that exploits suitable functions defined in the *Matlab* software.

Let **Z** be the input matrix holding the original data, we first obtain the PC-scores:

```
[ coeff, score] = princomp(Z)
F1 = score(:,1);
F2 = score(:,2);
```

where `COEFF` is a $p \times p$ matrix whose column containing coefficients (*factor loadings*) for a principal axis. The columns are in order of decreasing component variance. `SCORE` ($n \times p$) are the principal component scores, that is the coordinates of the individuals as defined in eq. (5).

Once the scores `F1` and `F2 are` obtained, we set the grid density parameter "`dens`". This is an user-defined parameter that fixes the number of horizontal and vertical lines of the grid. Depending on the number of observations and by the intrinsic process variability, values in the range from *20* to *40* should be explored. A proper value should produce a smooth surface while avoiding overfitting:

```
dens = 30;
thick1 = (max(F1) - min(F1)) / dens;
thick2 = (max(F2) - min(F2)) / dens;
t1 = min(F1) - tick1 : tick1 : max(F1) + tick1;
t2 = min(F2) - tick1 : tick2 : max(F2) + tick2;
```

the values `t1` and `t2` determine the equally spaced grid thick-marks along the two dimensions.

The next commands are built-in Matlab functions to generate the interpolated surface. The default fitting method used in the "`griddata`" function is based on the *Delaunay* data triangulation (Barber, et al. 1996). It produces a linear interpolation of the surface at the uniformly spaced points specified by (`Z1I,Z2I`) to produce YI.

```
[ Z1I,Z2I] = meshgrid(t1,t2);
YI = griddata(F1,F2,Y,Z1I,Z2I);
```

The graphical representation can be produced as a two-dimensional plot which overlaps the response variable Contour Plot and the factorial coordinates:

```
figure;
contour(Z1I,Z2I,YI), hold
plot3(F1,F2,Y,’ .k’ ), hold off
```

Moreover, a three-dimensional surface can be produced too:

```
figure;
surfc(X1,X2,YI), hold
plot3(F1,F2,Y,’ .k’ ), hold off
```

We denote this kind of graphical representations as *Response Surface Factorial Map* (*RSFM*). According to the different kind of response variables, we distinguish between an *Internal Analysis* and an *External Analysis*. In the first case, the response variable is used for better understand the results of the factorial technique and can be chosen among the initial set of variables or by using a quality index, for example the relative contribution as defined in eq. (6).

In the second case, an outer information is introduced in the analysis. The response variable should be a new variable not used in defining the factorial subspace but that can be of interest to better understand some pattern in the data. This is the situation for some economical and demographic variables in social surveys or for cost-benefit functions in product positioning problem as arising in preference data analysis and in Marketing surveys. For example the values of a cost function could reveal a technical frontier in the space of the optimal product positioning defined by a PrefMap (Carroll, 1962) or in a Factorial Conjoint Analysis experiment (Lauro et al., 1998).

In the next sections two illustrations of both Internal and External analysis will be provided.

3. ILLUSTRATIVE EXAMPLES OF RSFM IN INTERNAL AND EXTERNAL ANALYSIS.

In this section we show different way to produce the RSFM in presence of a response variable induced by the original data set or by an outer response variable.

The case study is a quite famous one, it was reported in Jackson (1991) and in SAS/STAT manuals. The data (here denoted as "*Linnerud Data*") consists of three physiological variable (predictors) and three fitness exercise responses measured on 20 middle-aged men in a health fitness club. This data set is interesting because it is a typical non-symmetrical data structure. Thus, we will use it both in Internal and External Analysis illustrations.

In Internal analysis the "*dependent*" variable plays an interpretative role, it is useful to better understand the quality of the representation in a subspace, or to visualise some correlation pattern between one or more original variables and the derived components. Let us examine the situation in the following case study.

3.1 THE INTERNAL ANALYSIS: QUALITY OF REPRESENTATION.

We performed a PCA on the *Linnerud* physiological variables (*Weight, Waist, Pulse*) and considered the exercises responses (*Chins, Situps, Jumps*) as supplementary variables. The results provided by the software (SPAD) are reported in the tables 1-3.

Tab. 1: Eigenvalues and explained variance.

	eigenvalue	*%*	*cum %*
1	2.1041	70.14	70.14
2	0.7662	25.54	95.68
3	0.1296	4.32	100.00

Tab. 2: Factor loadings of active and supplementary variables.

Active Variables	PC-loadings 1	PC-loadings 2
Weight	0.93	0.25
Waist	0.93	0.27
Pulse	-0.61	0.80
Supplementary Variables		
Chins	-0.46	-0.16
Situps	-0.57	-0.15
Jumps	-0.20	-0.10

Tab. 3: Scores and Relative contributions for the first two PCs.

Ind.	PC-Scores 1	PC-Scores 2	Rel. Contr. 1	Rel. Contr. 2	Total contribution
1	0.82	-0.58	0.63	0.32	0.95
2	0.85	-0.25	0.92	0.08	1.00
3	0.81	0.67	0.58	0.40	0.98
4	-0.88	0.53	0.64	0.23	0.87
5	0.80	-1.22	0.28	0.66	0.94
6	0.22	0.09	0.85	0.13	0.88
7	1.41	0.63	0.79	0.16	0.95
8	-0.83	0.23	0.93	0.07	1.00
9	-2.04	1.85	0.49	0.40	0.89
10	-1.14	-0.54	0.80	0.18	0.98
11	-0.18	-1.04	0.03	0.97	1.00
12	-0.59	-0.91	0.28	0.69	0.97
13	-1.42	0.59	0.80	0.14	0.94
14	4.37	1.06	0.94	0.06	1.00
15	1.11	-1.08	0.50	0.47	0.97
16	0.60	1.20	0.19	0.75	0.94
17	0.38	-0.15	0.41	0.06	0.47
18	-1.03	-1.12	0.46	0.54	1.00
19	-0.97	-0.77	0.61	0.38	0.99
20	-2.29	0.82	0.83	0.11	0.94

The correlation structure is a quite simple one, let us notice the collinearity between the two predictors variables *Weigth* and *Waist*, and betweeen the two responses *Situps* and *Chins*. These relations are also well evidentiate in the correlation circle. The first two components explain the 95.68% of the total variability.

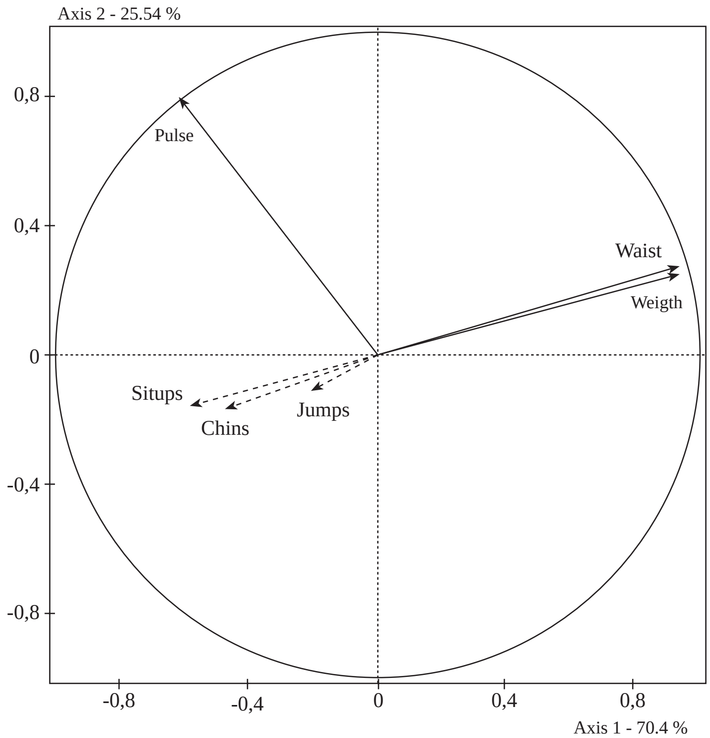

Fig. 1: The correlation circle.

By looking at the 20 individuals, we may note the good quality of representation for all but one athlete (see Tab. 3). The unit 17 shows a value of 0.47, which is somewhat low with respect to the whole group. In the Principal Component scatter plot (Fig. 2) this unit is located near the origin of the coordinate system.

To illustrate how the Internal analysis works, we analyse the global quality representation onto the principal plan. The (relative) Total Contribution (Tab. 2) is the response variable and according the Matlab procedure described in section 2.1, the Contour Plot (Fig. 3) and the Response Surface (Fig. 4), are produced.

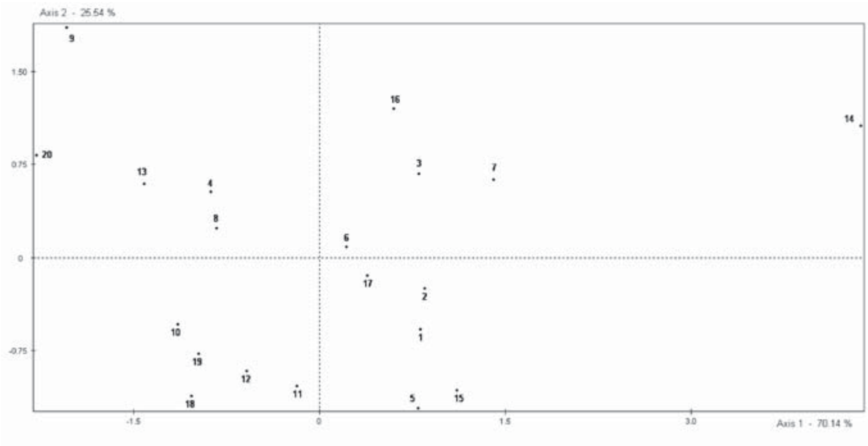

Fig. 2: The PCA scatter plot of the 20 athletes.

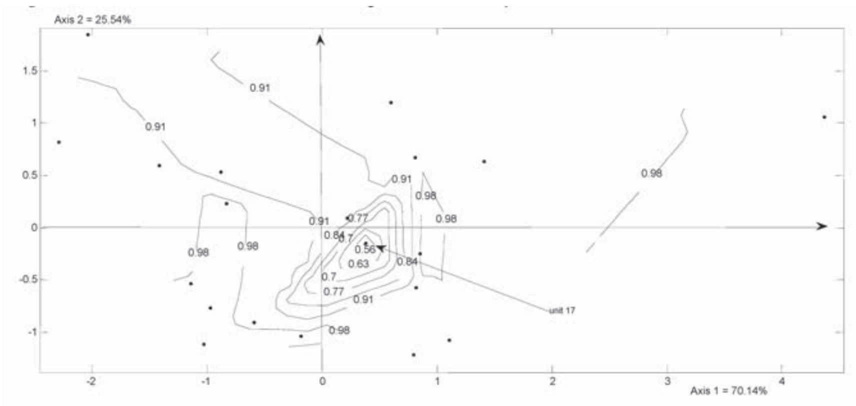

Fig. 3: The Contour Plot: the isocurves represent values of the total relative contributions.

In the Contour Plot in Fig. 3, the iso-curves represent equal values of Total Contribution. Values which tend to 1 represent zone where the statistical units are well represented.

In this case study, it is clear the effect in the zone of the unit 17 (highlighted by the arrow in Fig. 3), its poor quality cause a “hole” in the response surface.

In the three dimensional plot provided by the response surface map in Fig. 4, the “hole effect” is much more evident. The third dimension represents the quality of the representation and the unit 17, with a value of 0.47, lies in a deeper region.

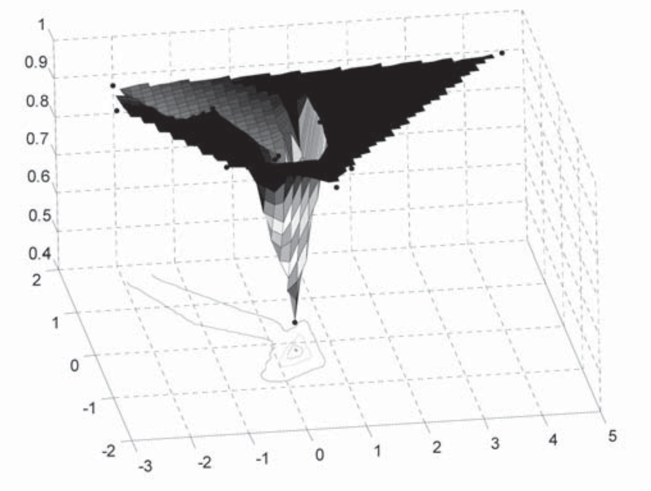

Fig. 4 – The Response Surface Factorial Map for the Quality of the representation.

3.2 EXTERNAL ANALYSIS: THE RESPONSE VARIABLE ANALYSIS

In the second example we perform an External Analysis by considering the response variable "*Situps*". The term *External* is referred to the circumstance that the response variable do not contribute to the definition of the principal components. In the previous example, the response variable was handled as a supplementary one: just the correlations with the two first components were computed and used as coordinates in the correlation circle (Fig. 1). It is undoubted the inverse relation with the active variables *Waist* and *Weigth*. Now we try to produce a more in-depth analysis of this relation by means of the RSFM.

At this purpose the Contour plot (Fig. 5) and the Response Surface (Fig. 6) are produced with the procedure described in section 2.1.

The analysis of the Contour plot allows to highlight the hillside effect of the response variable. Although it is inversely correlated with the first principal component (this is the result highlighted by the correlation circle), we may appreciate a more complex correlation structure. In particular, it is clear a twofold structure: a predominant one, in the direction – we may compare it with the correlation circle – of the first component, and a secondary one who appear in the direction of the second principal component. This effect is better illustrated in the three-dimensional RSFM in Fig. 6.

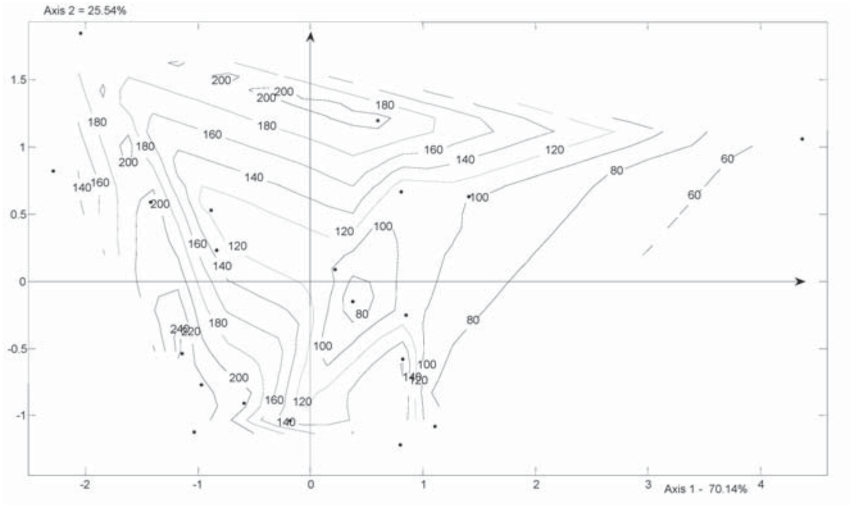

Fig. 5: The Contour Plot: the isocurves represent values of the response variable *Situps.*

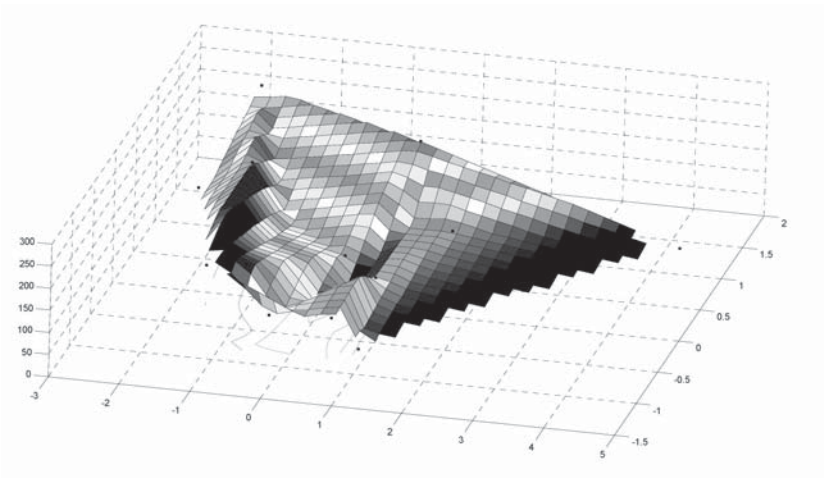

Fig. 6: The Response Surface Factorial Map for the External Analysis: The *Situps* surface.

4. CONCLUDING REMARKS

In this work we have proposed an exploratory strategy useful in multidimensional data analysis to visualise some interesting pattern by means of the Response Surface Methodology.

Our interest in using this kind of visualization is twofold: *i)* as interpretative tool, by using a response variable to investigate about the quality of the representation in a factorial map; *ii)* as exploratory tool, allowing to produce a surface onto the factorial plan which allows a more in-depth visualisation of the relationship between the principal components and other variables. When we deal with outer variables which do not contribute to the analysis, we define the procedure as “External Analysis”, whereas the response is an inner variable we define our procedure as “Internal Analysis”.

In the given examples we have just dealt with Principal Component Analysis, it needs further research to extent this approach to both symmetric and non-symmetrical factorial techniques (Lauro and D’Ambra, 1984). In the same way, some validation technique and inferential analytic tool should be provided in order to extent this procedure to sample survey data.

REFERENCES

ANDERSON T.W. (1951), Estimating linear restrictions on regression coefficients for multivariate normal distributions, *Ann. Math. Statist.*, **22**, 327-351. [Correction, Ann. Statist., 8, (1980), p. 1400.]

BARBER C. B., D.P. DOBKIN, and H.T. HUHDANPAA (1996), The Quickhull Algorithm for Convex Hulls, ACM Transactions on Mathematical Software, Vol. **22**, No. 4, Dec. 1996, p. 469-483. http://www.qhull.org

BOX G.E.P, DRAPER, N.R. (1987), *Empirical model building and response surfaces,* NewYork: JohnWiley & Sons.

BOX G.E.P., HUNTER W.G., HUNTER J.S. (1978), *Statistics for experimenters: An introduction to design,data analysis, and model building,* New York: John Wiley & Sons.

BOX G.E.P., WILSON K.B. (1951), On the Experimental Attainment of Optimum Conditions, *Journal of the Royal Statistical Society B,* **13**, 1-45.

CARR D.B. (1998), *Multivariate graphics*, in Encyclopedia of Biostatistics, P. Armitage and T. Colton eds., 2864-2886, Chichester: Wiley & Sons.

CARROLL J.D. (1972), Individual Differences and Multidimensional Scaling, In: R.N. Shepard, A.K. Romney, and S.B. Nerlove Eds., *Multidimensional Scaling: Theory and Applications in the Behavioral Sciences*, **1**. New York: Seminar Press.

CORNELL J.A. (1990), *How to App/y Response Surface Methodology*, Revised Edition, ASQC: Milwaukee, W1.

GREEN P.E., CARMONE F.J. (1969), Multidimensional Scaling: An Introduction and Comparison of Nonmetric Unfolding Techniques, *Journal of Marketing Research*, **6**, No. 3, 330-341.

GREEN P.E., SRINIVASAN V. (1990), Conjoint analysis in Marketing: New Development with Implications for Research and Practice, *Journal of Marketing,* 4-19.

GREENACRE M.J., HASTIE T.J. (1987), The geometric interpretation of correspondence analysis, *Journal of the American Statistical Association*, ***82***, *437-447*.

JACKSON J.E., (1992) *A User's Guide to Principal Component*, New York: John Wiley & Sons.

JOBSON J.D., (1992) *Applied Multivariate Data Analysis*, Vol. II: Categorical and Multivariate Methods, New York: Springer & Verlag, Inc.

JOLLIFFE I.T. (2002), *Principal Component Analysis*, 2nd Edition, New York: Springer & Verlag, Inc.

KENDALL M.G. (1957), *A Course in Multivariate Analysis*, London: Griffin.

KHURI A.I., CORNELL J.A. (1996), *Response surfaces: designs and analyses,* New York: Marcel Dekker, Inc.

LAURO N.C., D'AMBRA L. (1984). L'analyse non symetrique des correspondances. In: E. Diday et al. (eds.), *Data Analysis and Informatics*, III, North-Holland, Amsterdam.

LAURO N.C., GIORDANO G., VERDE R. (1998), A Multidimensional approach to Conjoint Analysis, *Applied Sthocastic Model and Data Analysis,* **14**, pp. 265-274.

LEBART L., MORINEAU A., PIRON M. (1997), *Statistique exploratoire multidimensionelle,* 2e Edition, Paris: Dunod.

MYERS, R. H., MONTGOMERY**,** D. C. (2002), *Response Surface Methodology: Process and Product Optimization Using Designed Experiments*, 2nd Edition, New York: John Wiley & Sons.

SAS Institute Inc. (1999), *SAS/STATÆ User's Guide*, Version 8, Cary, NC: SAS Institute Inc.

WOLD H. (1982), Soft modeling: the basic design and some extensions, In: J.-K. Joreskog and H. Wold, eds., *Systems Under Indirect Observation*, **2**, 1–53, North Holland, Amsterdam.

WOLD H. (1985), Partial least squares. In S. Kotz and N.L. Johnson, editors, *Encyclopedia of the Statistical Sciences*, **6**, 581–591, New York: John Wiley & Sons.

UNA NUOVA DIMENSIONE NELLE TECNICHE FATTORIALI: LA SUPERFICIE DI RISPOSTA

Riassunto

Nel presente lavoro si propone un'integrazione della Metodologia delle Superfici di Risposta, tipicamente applicata nell'ambito del Disegno degli Esperimenti, insieme alle rappresentazioni grafiche dei metodi fattoriali di Analisi Multidimensionale dei Dati. L'obiettivo è di migliorare e facilitare l'interpretazione dei risultati di tali tecniche tradizionalmente espressi con grafici cartesiani a due dimensioni. Grazie alla tecnica di rappresentazione delle superfici di risposta, si propone uno strumento descrittivo in grado di analizzare l'adattamento di una funzione dei dati definiti da una variabile quantitativa rispetto al sottospazio fattoriale. Il modello adattato può riguardare sia una variabile di risposta esterna *che non ha contribuito alla determinazione degli stessi assi fattoriali, sia una variabile* interna *al sistema per la quale si ha interesse ad approfondire l'analisi rispetto alle singole osservazioni. Nello stesso modo, è possibile rappresentare funzioni dei dati che descrivono la qualità della rappresentazione nel sottospazio al fine di una corretta interpretazione delle prossimità tra i punti-individuo.*