318

Nonlinear Redundancy Analysis

Research Report 88-1

Eeke van der Burg
Jan de Leeuw

department of EDUCATION

Division of Educational Measurement and Data Analysis

University of Twente

378

Nonlinear Redundancy Analysis

Eeke van der Burg

Jan de Leeuw[1]

1 The address of Jan de Leeuw is: University of California, Department of Psychology and Mathematics, 405 Hilgard Ave, Los Angeles CA 90024, USA

Abstract

A nonlinear version of redundancy analysis is introduced. The technique is called REDUNDALS. It is implemented within the computer program for canonical correlation analysis called CANALS (Van der Burg & De Leeuw, 1983). The REDUNDALS algorithm is of an alternating least squares (ALS) type. The technique is defined as minimization of a squared distance between criterion variables and weighted predictor variables. With the help of optimal scaling the variables are transformed nonlinearly (cf. Young, 1981). An application of redundancy analysis is provided.

Key words: redundancy analysis, canonical correlation analysis, optimal scaling, nonlinear transformation.

Nonlinear Redundancy Analysis

Introduction

In many situations data are available from different sources. Suppose the data are of the form: objects x variables, and let us suppose the data from one source correspond with a *subset* of variables. In case two (sub)sets of variables are available a possible technique to relate the sets to each other is *canonical correlation analysis* (CCA). This technique is described in many multivariate analysis textbooks (e.g. Tatsuoka, 1971, chap. 6;, Gnanadesikan, 1977, chap. 3.3). In CCA the two sets of variables are treated symmetrically. But a symmetric treatment is not always natural. It also happens that it is clear from the data which variables are *predictors* and which ones are *criteria*. In such cases *redundancy analysis* (RA) is a possible technique.

The name redundancy analysis originates from Van den Wollenberg (1977). Although he was the first one to name the technique, it actuallydates back from an earlier period. De Leeuw (1986) discusses the history of RA. We briefly summarize it. Horst (1955), Rao (1962), Stewart & Love (1968) and Glahn (1969) all propose the *Redundancy Index*. Rao (1964) and Robert & Escoufier (1976) discuss techniques for decomposing this Redundancy Index to uncorrelated components. Fortier (1966) proposes 'simultaneous linear predictions' which is equivalent with RA (cf. Ten Berge, 1985). Izenman (1975) and Davies & Tso (1982) also treat RA, but under the name Reduced Rank Regression. So far the discussion of De Leeuw (1986). Johansson (1981) proposes several forms of RA, which vary with orthogonality constraints, and DeSarbo (1981) disscusses a technique which is a mixture between CCA and RA. Van de Geer (1984) places various types of RA in a larger framework of *k* sets CCA. Israëls (1986) treats RA with various normalizations and rotations. Meulman (1986, chap. 5.2.1) discusses a version of RA which can be shown to be a generalization of Van den Wollenberg's RA. However Meulman uses a completely different approach, formulating RA in terms of distances between objects or individuals. We come will back to this later.

A nonlinear version of RA has been proposed by Israëls (1984). His technique makes it possible to incorporate qualitative variables by the use of 'dummies'. Also Meulman (1986, chap. 5.2.1) discusses a nonlinear version of RA, dealing with variables on an ordinal measurement level. In this paper another version of nonlinear RA is proposed. A larger choice of measurement levels is possible for each variable than in case of Israëls (1984).

As the algorithm for nonlinear redundancy analysis shows many correspondences with the algorithm for nonlinear CCA proposed by van der Burg & De Leeuw (1983), the computer program for nonlinear RA, called REDUNDALS, is embedded in the canonical correlation analysis program, called CANALS.

Redundancy analysis

Suppose the data consist of observations for n objects on m variables, and assume that the m variables can be divided into m_1 criterion variables and m_2 predictors. In addition assume that each variable is standardized, i.e. it has zero mean and unit variance. Collect the criterion variables in the matrix $\mathbf{H}_1$ of dimensions $(n \times m_1)$ and the predictors in $\mathbf{H}_2$ $(n \times m_2)$. The Redundancy Index of Stewart & Love (1968) is obtained by a *multivariate multiple regression* of $\mathbf{h}_i$, the columns of $\mathbf{H}_1$, $(i=1,\ldots,m_1)$ on $\mathbf{H}_2$. Thus

(1) minimize $\Sigma_{i=1}^{m_1}(\mathbf{h}_i - \mathbf{H}_2\mathbf{b}_i)'(\mathbf{h}_i - \mathbf{H}_2\mathbf{b}_i)/nm_1$

over $\mathbf{b}_1,\ldots,\mathbf{b}_{m_1}$.

where the vector $\mathbf{b}_i$ (m_2 elements) consists of regression weights. The squared distance or *loss* is divided by a factor nm_1 for the sake of comparing the various techniques. The matrix formulation of (1) is:

(2) minimize $\mathrm{tr}(\mathbf{H}_1 - \mathbf{H}_2\mathbf{B})'(\mathbf{H}_1 - \mathbf{H}_2\mathbf{B})/nm_1$ over $\mathbf{B}$.

This expression is minimized by

(3) $\mathbf{B} = (\mathbf{H}_2'\mathbf{H}_2)^{-1}\mathbf{H}_2'\mathbf{H}_1$.

provided that $\mathbf{H}_2'\mathbf{H}_2$ is of full rank. Substitution of (3) in (2) gives the minimum:

$$(4)\quad \mathrm{tr}(\mathbf{H}_1'\mathbf{H}_1 - \mathbf{H}_1'\mathbf{H}_2(\mathbf{H}_2'\mathbf{H}_2)^{-1}\mathbf{H}_2'\mathbf{H}_1)/nm_1$$

Denoting $\mathbf{R}_{11}$ for $\mathbf{H}_1'\mathbf{H}_1/n$ and $\mathbf{R}_{12}$ and $\mathbf{R}_{22}$ for $\mathbf{H}_1'\mathbf{H}_2/n$ and $\mathbf{H}_2'\mathbf{H}_2/n$ respectively, expression (4) is equivalent to

$$(5)\quad 1 - \mathrm{tr}(\mathbf{R}_{12}\mathbf{R}_{22}^{-1}\mathbf{R}_{21})/m_1.$$

The expression $\mathrm{tr}(\mathbf{R}_{12}\mathbf{R}_{22}^{-1}\mathbf{R}_{21})/m_1$ is equal to the Redundancy Index of Stewart & Love (1968). Thus minimizing (1) corresponds to computing the Redundancy Index.

However this is not the same as performing a redundancy analysis in the sense of Van den Wollenberg (1977). He searches for (normalized) weights that optimize the explained variance between criterion variables and weighted predictors. These weight vectors $\mathbf{v}$ (m_2 elements) are eigenvectors of the matrix $\mathbf{R}_{22}^{-1}\mathbf{R}_{21}\mathbf{R}_{12}$. Denote the corresponding eigenvalues by μ. Then

$$(6)\quad \mathbf{R}_{22}^{-1}\mathbf{R}_{21}\mathbf{R}_{12}\mathbf{v} = \mu\mathbf{v} \text{ with } \mathbf{v}'\mathbf{R}_{22}\mathbf{v} = 1.$$

When all $\mathbf{v}$'s are solved, the sum of eigenvalues equals the Redundancy Index (cf. Israëls, 1984). In fact we can see Van den Wollenberg's analysis as a specialization of our minimization problem (2), namely the case in which there are rank restrictions on matrix $\mathbf{B}$, i.e. $\mathbf{B}=\mathbf{VW}'$ with $\mathbf{V}$ ($m_2 \times r$), $\mathbf{W}$

(m_1 x r), $1 \leq r \leq \min(m_1, m_2)$, and normalization constraints on **V**, i.e. $\mathbf{V}'\mathbf{R}_{22}\mathbf{V}=\mathbf{I}$. Expression (2) is rewritten in terms of **V** and **W** as follows

(7) minimize $\text{tr}(\mathbf{H}_1 - \mathbf{H}_2\mathbf{V}\mathbf{W}')'(\mathbf{H}_1 - \mathbf{H}_2\mathbf{V}\mathbf{W}')/nm_1$ over **V** and **W**

subject to the condition that $\mathbf{V}'\mathbf{R}_{22}\mathbf{V}=\mathbf{I}$.

Some computational work shows that the columns of **V** correspond to the vectors **v** discussed above. Note that Van den Wollenberg has the choice of r, i.e. how many eigenvectors **v** will be computed. In our case automatically all weights **B** are solved for, as this is implicit to the way (2) is formulated. Although (7) is more restrictive than (2), we can argue that formulation (7) is the more general one, as (7) can be solved for $r=m_1$ (assuming that $m_1 \leq m_2$), and for lower values of r.

Expression (7) also shows the relation between reduced rank regression and redundancy analysis, as reduced rank regression corresponds to (7) with small r (c.f. De Leeuw, Mooijaart & Van der Leeden, 1985). To recognize other forms of RA it is necessary to formulate expression (7) in a different way. Define matrix **X** (n x r) as $\mathbf{H}_2\mathbf{V}$. Then we get

(8) minimize $\{\text{tr}(\mathbf{X}-\mathbf{H}_2\mathbf{V})'(\mathbf{X}-\mathbf{H}_2\mathbf{V}) + \text{tr}(\mathbf{H}_1-\mathbf{X}\mathbf{W}')'(\mathbf{H}_1-\mathbf{X}\mathbf{W}')\}/nm_1$

over **X**, **V** and **W**, subject to the conditions that

$$\mathbf{X} = \mathbf{H}_2\mathbf{V} \text{ and } \mathbf{R}_{XX} = \mathbf{I}.$$

Matrix $\mathbf{R}_{XX}$ is equal to $\mathbf{X}'\mathbf{X}/n$. Meulman (1986, chap. 5.2.1) discusses the minimization of the loss as formulated in (8), subject to the condition that only $\mathbf{R}_{XX}=\mathbf{I}$. Thus $\mathbf{X}$ does not have to be in the column space of $\mathbf{H}_2$. De Leeuw & Bijleveld (1987) deal with the same loss function, but they use the condition $\mathbf{R}_{XX}=\alpha^2\mathbf{I}$, where α is a parameter. They show that different values of α correspond to various multivariate techniques, e.g. $\alpha=0$ boils down to principal component analysis (PCA), and $\alpha\rightarrow\infty$ corresponds to reduced rank regression.

Optimal scaling

In many ways *nonlinear transformations* can be implemented in redundancy analysis. To do so Israëls (1984) employed dummies for variables measured on a nominal measurement level. Meulman (1986, chap. 5.2.1) uses monotone regression in her version of nonlinear RA. Monotone regression is a form of *optimal scaling* (cf. Young, 1981). This means that the transformations (scaling parameters) minimize the loss, and at the same time *measurement restrictions* are maintained. We also use optimal scaling. The nonlinear transformations treated in this article are *nominal* and *ordinal* (a definition will follow). In addition, of course, *linear* or *numerical* transformations are dealt with. 'Dummy transformations', as employed by Israëls (1984), are not discussed, however they can always be obtained by simply coding variables as dummies, and, in addition, by treating these dummies numerically. Another way to obtain these 'dummy transformations' is by using *copies* of a variable within the corresponding set, and by treating these copies as nominal. This gives a multiple nominal (or dummy) transformation (cf. Gifi, 1981, chap. 5.2.7). Using copies instead of dummies has the advantage that one may choose both the dimensionality of the transformation and the measurement level of each copy separately. More information about copies can be found in De Leeuw (1984) and Van der Burg & De Leeuw (1987).

The nominal, ordinal and numerical transformations employed in this article agree with the transformations used by Van der Burg & De Leeuw (1983) in their version of nonlinear CCA (CANALS). Together these three transformations form the optimal scaling. Our definition of optimal scaling corresponds to the definition of Young (1981). We mentioned already that optimal scaling refers to the fact that variables are optimally scaled in the sense of the model. This means that the data matrices $\mathbf{H}_1$ and $\mathbf{H}_2$ are replaced by parameter matrices $\mathbf{Q}_1$ ($n \times m_1$) and $\mathbf{Q}_2$ ($n \times m_2$) such that they optimize the model, i.e. minimize the original loss, but at the same time satisfy the measurement restrictions. The original loss was formulated in (2). If the parameter matrix $\mathbf{Q}_1$ is subsituted for $\mathbf{H}_1$ and $\mathbf{Q}_2$ for $\mathbf{H}_2$, this expression can be rewritten as follows. Denote the set of possible transformations for the *i*th variable, i.e. *i*th column of

$[\mathbf{H}_1, \mathbf{H}_2]$, by C_i and use the notation $\mathbf{q}_i$ for the ith column of $[\mathbf{Q}_1, \mathbf{Q}_2]$. Nonlinear redundancy analysis or REDUNDALS is

(9) minimize $\mathrm{tr}(\mathbf{Q}_1 - \mathbf{Q}_2\mathbf{B})'(\mathbf{Q}_1 - \mathbf{Q}_2\mathbf{B})/nm_1$

over $\mathbf{Q}_1$, $\mathbf{Q}_2$ and $\mathbf{B}$, subject to the condition that

$\mathbf{q}_i \in C_i$ $(i=1,\ldots,m)$.

The sets of possible transformations are determined by *tie* and *normalization restrictions* for nominal variables, and, in addition, by *monotone* constraints for ordinal variables or by *linear* constraints for numerical variables (cf. De Leeuw, 1977). Tie restrictions imply that ties in the data correspond to ties in the transformation. Normalization restrictions result in standardized transformations (i.e. zero mean and unit variance). The monotone transformations discussed here correspond to the secondary approach of Kruskal & Shephard (1974). Finally linear transformations are equal to the variables itself, as standardization of the columns of the data matrix was supposed. A more extensive discussion of optimal scaling restrictions can be found in Young, De Leeuw & Takane (1976) and Young (1981).

REDUNDALS algorithm

The algorithm for nonlinear redundancy analysis follows easily from (9). Using an alternating least squares method results in solving the parameters in the following order (δ is a very small number)

```
        a  initialize Q1, Q2
 |---> b  compute B
 |      c  compute the columns of Q1
 |      d  compute the columns of Q2
 |no
 |<--- e  (loss_previous - loss_present) ≤ δ
 |yes
 |---> f  end
```

If one set of parameters is updated the remaining ones are kept at a constant level. As nonlinear RA can be viewed as a special case of nonlinear CCA the solutions for the various parameters can be found in Van der Burg & De Leeuw (1983). These authors formulate nonlinear CCA as follows. Define **A** (m_1 x p) and **B** (m_2 x p) as the weight matrices for the first set and the second set respectively (this new definition of **B** does not interfere with the earlier one). The p corresponds to the number of dimensions or the number of canonical variates. Then nonlinear CCA is, according to Van der Burg & De Leeuw (1983),

(10) minimize $\mathrm{tr}(\mathbf{Q}_1\mathbf{A} - \mathbf{Q}_2\mathbf{B})'(\mathbf{Q}_1\mathbf{A} - \mathbf{Q}_2\mathbf{B})/np$

over $\mathbf{Q}_1$, $\mathbf{Q}_2$, $\mathbf{A}$ and $\mathbf{B}$, subject to the conditions that

$\mathbf{A}'\mathbf{Q}_1{}'\mathbf{Q}_1\mathbf{A} = n\mathbf{I}$ or $\mathbf{B}'\mathbf{Q}_2{}'\mathbf{Q}_2\mathbf{B} = n\mathbf{I}$ and

$\mathbf{q}_i \in C_i$ $(i=1,\ldots\ldots,m)$.

This technique is called *CANALS*. In comparing this definition with the definition of nonlinear RA in (9) we see that, from a CCA point of view,

a the number of dimensions p is fixed to m_1,

b the weight matrix $\mathbf{A}$ is equal to the identity matrix,

c no normalizations are used.

The CANALS algorithm is based on one normalization, either of $\mathbf{A}$ or of $\mathbf{B}$. In addition the CANALS algorithm uses transfer of normalization in the iterative process (i.e. weight matrices $\mathbf{A}$ and $\mathbf{B}$ are rescaled such that the matrix to be updated is *not* normalized). The sequence of solving the parameters in the CANALS program is the following.

1 initialize $\mathbf{Q}_1$, $\mathbf{Q}_2$, $\mathbf{A}$

2 rescale $\mathbf{A}$ such that $\mathbf{A}'\mathbf{Q}_1{}'\mathbf{Q}_1\mathbf{A} = n\mathbf{I}$

——> 3 compute $\mathbf{B}$

4 compute $\mathbf{Q}_2$

5 rescale $\mathbf{A}$ and $\mathbf{B}$

6 compute $\mathbf{A}$

7 compute $\mathbf{Q}_1$

8 rescale $\mathbf{A}$ and $\mathbf{B}$ such that $\mathbf{A}'\mathbf{Q}_1{}'\mathbf{Q}_1\mathbf{A} = n\mathbf{I}$

no

<—— 9 $(\text{loss}_{\text{previous}} - \text{loss}_{\text{present}}) \leq \delta$

yes

——> 10 rescale $\mathbf{A}$ and $\mathbf{B}$ such that both $\mathbf{A}$ and $\mathbf{B}$ normalized

11 end

Again the remaining parameters are supposed to be at a constant level when one set of parameters is updated. As the REDUNDALS solutions are similar to the CANALS solutions (as long as $\mathbf{I}$ is substituted for $\mathbf{A}$ and p is taken as m_1), we see that REDUNDALS corresponds to steps 1,3,4,7,9, and 11 of CANALS. Therefore it is easy to combine the two algorithms. The REDUNDALS program is simply embedded in the CANALS program by employing only the equivalent steps and by skipping the other ones. A difference between the CANALS and the REDUNDALS program is the fact that in case of REDUNDALS matrix $\mathbf{A}$ is initialized on $\mathbf{I}$ and in case of CANALS $\mathbf{A}$ starts with random values. In addition, the final solution for REDUNDALS is not rotated while the CANALS solution is (for CCA rotated weights give a similar loss).

As long as the variables are treated numerically the REDUNDALS program will iterate to a global minimum. However if nominal or ordinal variables are dealt with, a local minimum may occur. We do not know how serious this problem is. Compare Van der Burg *et al.* (1986) for a discussion of convergence in the case of nonlinear CCA with k sets of variables.

Application

For illustration of REDUNDALS an example is taken which was also used to demonstrate the CANALS technique (Van der Burg & De Leeuw, 1983). A detailed description of this example can be found in the latter article. The data are from a Parliamentory Survey carried out in 1972. Among other things, the Dutch members of parliament (MPs) gave their opinion on seven issues, and their preference votes for the political parties of which only the four larger parties interest us. The opinions were measured on a nine–point scale of which the lowest and the highest category were described (Table 1). The preference votes were recorded as a table of rank orders. As there were 14 parties, the preference votes take values 2 (highest preference) to 15 (lowest preference) (Table 1).

INSERT TABLE 1 ABOUT HERE

We wondered whether party preferences could predict the opinions on the issues. The idea behind it is that MPs have their, traditional, party sympathies and take over the official party viewpoints. In Dutch politics a rather strong party discipline exists, so that many parties carry a clear image.

Results are shown in Table 2. In the first place the multiple correlation coefficients are rather high. Thus the preference votes are good predictors for the opinions on the issues. In Table 2 also the correlations between the preferences and the issues (both transformed monotonically) are given. We do not interprete the weights as they do not give a clear idea of the relations between the variables due to multicollinearity (cf. Gnanadesikan, 1977, p. 22). Table 2 shows that the preference votes for PvdA and VVD are more strongly correlated with the issues than the preference votes for ARP and KVP are. The highest correlations are between PvdA and LAW, and VVD and INC. This means that the amount of sympathy for the socialists (PvdA) goes together with ideas about law & order: more sympathy corresponds to 'too strong action', and antipathy to 'stronger action'. Great sympathy of the MPs for the VVD agrees with 'income differences should remain', and antipathy with 'income differences much less'. Law & order is a hot topic for the PvdA (but also TAX and DEF) and income differences for the VVD. Both parties take up clear positions on these issues(cf. Van der Burg & De Leeuw, 1983). All the other opinions are also correlated strongly with the preference votes for these two parties, and nearly always in different directions (except for ABO). The socialistic party (PvdA) and the conservative party (VVD) are antagonists in Dutch politics, as in most countries. Apparently the profilation of those parties is clear to many members of parliament.

The subject abortion needs some extra explanation. As the VVD originally is a liberal party, there exists still some liberal ideas. Especially with regard to abortion several VVD—members kept liberal thoughts (cf. Van der Burg & De Leeuw, 1983), so that socialistic MPs and those conservative MPs agree on this subject. One has to know the historical background to understand such behaviour. Apparently sympathy (or antipathy) for the VVD comes from both people against abortion and people pro abortion, as the correlation between ABO and the VVD preference is not very high.

The christian democratic parties (KVP and ARP) appear to be less clear (or extreme) than the socialists or conservatives are. Sympathy for the KVP includes a strong position against free abortion, but other issues hardly correlate with a KVP preference. A similar thing holds for the ARP. This agrees with the fact that the christian democrates form a middle party (they combined after 1972). Sometimes they co—operate with left, sometimes with right. Both socialists and conservatives need christian democrats to have a majority in the parliament. Therefore it is clear that MPs from both left and right have sympathy for the KVP or

ARP, while having completely different ideas on the issues (except for ABO).

The difference between seven separate (nonlinear) multiple correlations and a REDUNDALS analysis lays mainly in the fact that only one transformation is obtained for all the analyses together. This is a great advantage as analysis results have to be interpreted for each set of tranformations separately. This problem is avoided by using one set of transformations. As all variables had a natural ordering in the categories, the data were treated at an ordinal measurement level. The monotone transformations are given in Fig. 1. The original scores (horizontal) are plotted against the transformed values, the so-called category quantifications (vertical). The most striking transformation is the one belonging to the KVP preference. We see that the lowest score is separated from the rest. This means that one has either a very large sympathy (i.e. one is a member of the KVP party) or not. The nots are not distiguished from each other. Therefore it is clear that the KVP preference hardly correlates with the opinions on the issues. The ties accentuate once more the middle position of the KVP.

INSERT FIGURE 1 ABOUT HERE

The transformation of the ARP preference also shows ties, but only for categories regarding great antipathy. Thus strong or very strong aversion to the ARP does not matter in this analysis. The issue INC shows several ties in the low categories. This means that the differentiation is in the fact of how much income differences should decrease. The amount of decrease corresponds to the PvdA or VVD preference. The remaining transformations look normal, i.e. they do not contain many ties, nor very big jumps.

Discussion

The nonlinear redunduncy analysis presented in this article corresponds to a multivariate multiple regression with optimal scaling. The technique maximizes the Redundancy Index of Stewart & Love (1968). The algorithm is realized in the computer program REDUNDALS which can handle several types of nonlinear transformations. This is a great advantage over other redundancy analysis programs. In addition only one transformation for each variable is obtained for all the multiple regressions together, which makes interpretation more simple than in case of separate analyses.

As the REDUNDALS program is implemented within a program for nonlinear canonical correlation analysis (CANALS) the approach to missing data is the same. This means that missing observations are quantified such that the model is fitted optimally. Only one quantification for each missing value is computed instead of as many as there are predictors. Thus even in case of (incomplete) data with a numerical measurement level, one may prefer the REDUNDALS program over multivariate multiple regression with listwise or pairwise deletion.

The REDUNDALS program is written in Fortran. As CANALS and REDUNDALS are combined, the program can only be obtained together with the CANALS program. Another computer program which can also perform multivariate multiple regression together with nonlinear transformations is TRANSREG (Kuhfeld, Young & Kent, 1987). This program is implemented as a SAS procedure.

A disadvantage of the REDUNDALS technique is that no space reduction is obtained. Van den Wollenberg (1977) can choose how many components must be obtained. In fact he solves the generalized eigenvalue problem ($R_{21}R_{12}$, R_{22}), which gives directions in the predictor space that explain the larger proportion of variance of the criterion variables. Of course, this generalized eigenvalue problem can always be solved after a REDUNDALS analysis. Then the transformed variables must be used to compute the correlation matrices.

The difference between CANALS and REDUNDALS results is that CANALS finds direction(s) in both sets of variables (subspaces), that correlate maximally, independent of how much variance is explained, while REDUNDALS explains as much variance as possible in every criterion direction. This means that results are hardly comparable unless one of the canonical variates correlates strongly with one or more criterion variables. However, they should never contradict each other.

In case of the Dutch Parliamentory data the REDUNDALS results are mostly comparable with the numerical CANALS analysis (cf. Van der Burg & De Leeuw, 1983). The first dimension of this analysis is dominated by all the issues (except for ABO, and DEV) and the VVD and PvdA preference, and the second dimension by ABO and the KVP and ARP preference. Of course the transformations differ, but we have seen from Fig. 1 that no large deviations exist from linearity (except for the KVP transformation). The first CANALS dimension corresponds to the left-right contrast and the second dimension to the con-pro-abortion contrast. Indeed, REDUNDALS shows a similar pattern although not in two dimensions. Even subtle results (e.g. the KVP and ARP preferences are tended towards the VVD preference on the issues LAW and DEF (Table 2)), are recovered in the CANALS results.

References

Davies, P.T. and Tso, M.K. (1982). Procedures for reduced-rank regression. *Applied Statistics, 31*, 244–255.

De Leeuw, J. (1977) Normalized cone regression. University of Leiden: Department of Data Theory. Mimeo.

De Leeuw, J. (1986). On the history of redundancy analysis. University of Leiden: Department of Data Theory.

De Leeuw, J. (1984). The Gifi-system of Nonlinear Multivariate Analysis. In E. Diday, M. Jambu, L. Lebart, J. Pagès & R. Thomassone (Eds.), *Data Analysis and Informatics III*, 415–424. Amsterdam: North Holland.

De Leeuw, J. and Bijleveld, C. (1987). Fitting reduced rank regression models by alternating least squares. University of Leiden: Department of Data Theory. RR-87-05.

De Leeuw, J., Mooijaart, A., and Van der Leeden, R. (1985). Fixed factor score models with linear restrictions. University of Leiden, Department of Data Theory, RR-85-06.

DeSarbo, W.S. (1981). Canonical/redundancy factorial analysis. *Psychometrika, 46,* 307–329.

Fortier, J.J. (1966). Simultaneous linear prediction. *Psychometrika*,31, 369–381.

Gifi, A. (1981). *Nonlinear Multivariate Anlysis.* University of Leiden, Department of Data Theory. In press, Leiden: DSWO Press.

Glahn, H.R. (1969). Some relationships derived fron canonical correlation theory. *Econometrika, 37,* 252–256

Gnanadesikan, R. (1977). *Methods for Statistical Data Analysis of Multivariate Observations*. New York: Wiley.

Horst, P. (1955). A technique for the development of a multiple absolute prediction battery. *Psychological Monographs, 69, no 5,* 1–22.

Israëls, A.Z. (1984). Redundancy analysis for qualitative variables. *Psychometrika, 31,* 369–381.

Israëls, A.Z. (1986). Interpretation of redundancy analysis: rotated vs. unrotated solutions. *Applied Stochastic Models and Data Analysis, 2,* 121–130.

Izenman, A.J. (1975). Reduced-rank regression for the multivariate linear model. *Journal of Multivariate Analysis, 5,* 248–264

Johansson, J.K. (1981). An extension of Wollenberg's redundancy analysis. *Psychometrika, 46,* 93–104.

Kruskal, J.D. and Shephard, R.N. (1974). A nonmetric variety of linear factor analysis. *Psychometrika, 39,* 123–157.

Kuhfeld, W.F , Young, F.W. and Kent, D.P. (1987). New developments in psychometric amd market research procedures. *In SUGI Proceedings*. Cary, NC: SAS Institute.

Meulman, J. (1986). *A Distance Approach to Nonlinear Multivariate Analysis*. Leiden: DSWO Press.

Rao, C.R. (1962). *Use of discriminant and allied functions in multivariate analysis. Sankya A, 24,* 149–154.

Rao, C.R. (1964). The use and interpretation of principal components analysis in applied research. *Sankya A, 26,* 329–358.

Robert, P. & Escoufier, Y. (1976). A unifying tool for linear multivariate statistical methods: the RV–coefficient. *Applied Statistics, 25*, 257–265

Stewart, D. and Love, W. (1968). A general canonical correlation index. *Psychological Bulletin, 70,* 160–163.

Tatsuoka, M.M. (1971). *Multivariate Analysis: Techniques for Educational and Psychological Research*. New York: Wiley.

Ten Berge, J.M.F. (1985). On the relationship between Fortier's simultaneous linear predition and Van den Wollenberg's redundancy analysis. *Psychometrika, 50,* 121–122.

Van de Geer, J.P. (1984). Linear relations between *k* sets of variables. *Psychometrika, 49*, 79–94.

Van den Wollenberg, A.L. (1977). Redundancy analysis: an alternative for canonical correlation analysis. *Psychometrika, 42,* 207–219.

Van der Burg , E. and De Leeuw, J. (1983). Non–linear canonical correlation. *British Journal of Mathematical and Statistical Psychlogy, 36*, 54–80.

Van der Burg , E. and De Leeuw, J. (1987). Nonlinear canonical correlation analysis with *k* sets of variables. University of Twente: Department of Education, RR–87–8. Submitted for publication.

Van der Burg , E., De Leeuw, J. and Verdegaal, R. (1986). Homogeneity analysis with *k* sets of variables: an alternating least squares method with optimal scaling features. University of Twente: Department of Education, RR–86–5. To appear in *Psychometrika*

Young, F. (1981). Qualitative Analysis of Qualitative Data. *Psychometrika, 46*, 247–388.

Young, F, De Leeuw, J. & Takane, Y. (1976). Regression with Qualitative and Quantitative variables: an alternating least sqiares method with optimal scaling features. *Psychometrika, 41*, 505–529.

<u>Table 1.</u> Dutch Parliament. The issues and party preferences and the meaning of the lowest and the highest category.

DEV: development aid
(1) the government should spend more money on aid to developing countries
(9) the government should spend less money on aid to developing countries

ABO: abortion
(1) the government should prohibit abortion completely
(9) a woman has the right to decide for herself about abortion

LAW: law and order
(1) the government takes too strong action against public disturbances
(9) the government should take stronger action against public disturbances

INC: income differences
(1) income differences should remain as they are
(9) income differences should become much less

PAR: participation
(1) only management should decide important matters in industry
(9) workers must also have participation in decisions important for industry

TAX: taxation
(1) taxes should be decreased for general welfare
(9) taxes should be decreased so that people can decide for themselves how to spend their money

DEF: defence
(1) the government should insist on shrinking Western armies
(9) the government should insist on maintaining strong Western armies

PvdA: socialists
(2) highest preference, (15) lowest preference

ARP: christian democrats (protestants)
(2) highest preference, (15) lowest preference

KVP: christian democrates (katholics)
(2) highest preference, (15) lowest preference

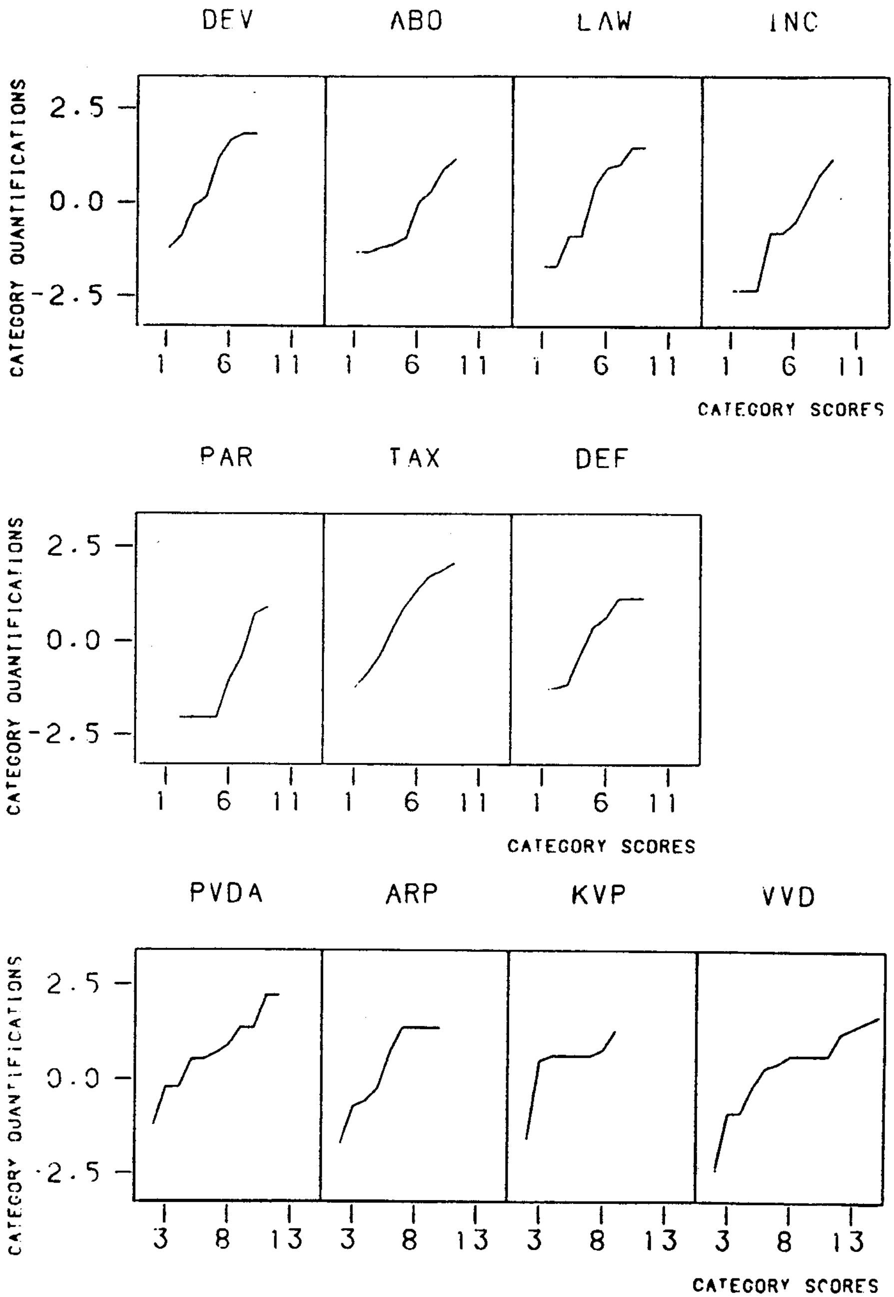
VVD: conservatives
(2) highest preference, (15) lowest preference

Table 2. Dutch Parliament. Multiple correlations (MC) and correlations between issues (columns) and preference votes (rows).

	DEV	ABO	LAW	INC	PAR	TAX	DEF
MC	.662	.782	.786	.804	.767	.805	.765
PvdA	.591	−.430	.719	−.624	−.632	.716	.705
ARP	.000	.542	−.336	−.022	−.152	.031	−.294
KVP	−.001	.637	−.274	.051	−.100	.033	−.182
VVD	−.537	−.212	−.418	.722	.667	−.654	−.450

Figure Caption

Figure 1. Monotone transformations of the variables. Original scores (horizontal) against category quantifications (vertical).


DEV
ABO
LAW
INC
CATEGORY QUANTIFICATIONS
2.5
0.0
-2.5
1
6
11
CATEGORY SCORES
PAR
TAX
DEF
CATEGORY QUANTIFICATIONS
2.5
0.0
-2.5
1
6
11
CATEGORY SCORES
PVDA
ARP
KVP
VVD
CATEGORY QUANTIFICATIONS
2.5
0.0
-2.5
3
8
13
CATEGORY SCORES