Division of Research
Graduate School of Business Administration
The University of Michigan

January 1985

A GENERAL MODEL AND SIMPLE ALGORITHM FOR REDUNDANCY ANALYSIS

Working Paper No. 412

Claes Fornell
and
Donald W. Barclay

A GENERAL MODEL AND SIMPLE ALGORITHM FOR REDUNDANCY ANALYSIS

Abstract

Stewart and Love proposed redundancy as an index for measuring the amount of shared variance between two sets of variables. Van den Wollenberg presented a method for maximizing redundancy. Johansson extended the approach to include derivation of optimal Y-variates, given the X-variates. This paper shows that redundancy maximization with Johansson's extension can be accomplished via a simple iterative algorithm based on Wold's Partial Least Squares.

1. Introduction

Following the publication of van den Wollenberg's [1977] paper on redundancy maximization, discussions and extensions of his approach have been developed by Johansson [1981], DeSarbo [1981], Muller [1981], Dawson-Saunders [1983], and Tyler [1982].

A limitation of van den Wollenberg's redundancy solution was pointed out by Johansson and later by Tyler. Specifically, van den Wollenberg's approach implies that the Y-variates are derived independently of the X-variates using the eigenvectors of $\tilde{R}_{yy}^{-1}\tilde{R}_{yx}\tilde{R}_{xy}$, i.e., the transformation of Y-variates is not related to the transformation of the X-variates. As a result, as opposed to canonical correlations, the correlations between X and Y variates are not optimal. Johansson [1981] extended van den Wollenberg's method to include the derivation of Y-variates which are maximally correlated with the X-variates constructed to maximize the redundancy of the y-variables. He also shows that these Y-variates have some desirable properties of orthogonality. Thus, Johansson's approach appears very attractive.

The purpose of this paper is to demonstrate that Johansson's extended version of redundancy analysis can be accomplished via a very simple iterative algorithm involving nothing more than a series of simple and multiple regressions. A useful feature of the algorithm lies in its simplicity. There is no need for the analyst to construct his own computer program; all that is necessary is a program for standard multiple regression such as MIDAS, SAS, or TROLL. As a result, Johansson's extended redundancy analysis is easily available to most applied researchers.

2. Redundancy Analysis

Stewart and Love [1968] developed a nonsymmetric index of redundancy which represents the mean variance in one set of variables predicted by a linear composite or variate of another set of variables. The redundancy index is:

$$RD_{y_\ell} = \gamma_\ell^2 \frac{1}{q_y} \underset{\sim}{\Lambda}'_{y\eta_\ell} \underset{\sim}{\Lambda}_{y\eta_\ell}$$

where RD_{y_ℓ} is the redundancy of the criteria given the ℓ^{th} canonical variate of the predictors, γ is the canonical correlation coefficient, q_y is the number of y-variables, $\underset{\sim}{\Lambda}_{y\eta_\ell}$ is the vector of loadings of the y-variables on their ℓ^{th} canonical variate η_ℓ.

It has also been shown [e.g., Johansson 1981] that the redundancy index can be viewed as the mean squared loadings of the variables of the criteria set on the canonical variate of the predictor set, ξ, i.e.,

$$RD_{y_\ell} = \frac{1}{q_y} \underset{\sim}{\Lambda}'_{y\xi_\ell} \underset{\sim}{\Lambda}_{y\xi_\ell}.$$

In canonical correlation analysis, <u>only</u> the γ_ℓ^2 portion of the redundancy index is maximized. Van den Wollenberg [1977] suggested that redundancy per se could be maximized. Maximization of redundancy results in two general characteristic equations:

$$(\underset{\sim}{R}_{xy}\underset{\sim}{R}_{yx} - \lambda_\alpha \underset{\sim}{R}_{xx})\underset{\sim}{\alpha} = 0$$

$$(\underset{\sim}{R}_{yx}\underset{\sim}{R}_{xy} - \lambda_\beta \underset{\sim}{R}_{yy})\underset{\sim}{\beta} = 0$$

where λ_α and λ_β are eigenvalues while $\underset{\sim}{\alpha}$ and $\underset{\sim}{\beta}$ are weight vectors.

Van den Wollenberg goes on to develop the case for extracting successive variates such that these are orthogonal to variates already constructed within the same variable set. It is not, in general, possible to have biorthogonal variates in redundancy analysis.

Johansson [1981] presents two solutions for deriving Y-variates given X-variates derived to maximize the redundancy in the y-variables. The first solution, a least squares approach, satisfies the orthogonality condition $\beta_{\ell}' R_{yy} \beta_m = 0$, $\ell \neq m$ but not the condition $\alpha_{\ell}' R_{xy} \beta_m = 0$, $\ell \neq m$ where α_{ℓ} defines the weights of the ℓ^{th} X variate and β_m defines the weights of the corresponding Y variate. The second approach, a restandardized solution, fulfills the opposite orthogonality condition, i.e., the solutions are complementary.

We first present the iterative algorithm for redundancy maximization. In the next section we show the equivalence of this approach to Johansson's solution. Following this, the numerical example used by both van den Wollenberg and Johansson is applied.

3. The Algorithm

The algorithm derives from Wold's [1966] extension of his fix-point method to nonlinear iterative Partial Least Squares (PLS) which has been shown to be a general model for principal components and canonical correlation analysis [e.g., Areskoug 1982]. As we will show here, Johansson's extension of redundancy analysis can also be obtained via this algorithm. Thus, even before the term "redundancy" was coined by Stewart and Love in 1968 and well before van den Wollenberg presented his analytical solution to redundancy maximization in 1977, a method existed that not only maximized redundancy but also incorporated the attractive features of Johansson's [1981] extension. And, as we will see, the approach involves nothing but traditional OLS regressions in an iterative fashion using a fixed point constraint.

The following series of OLS regressions is executed:

<u>Initialize</u>

Let $\eta = y_1 + \ldots + y_q$; $\xi = x_1 + \ldots + x_p$.

Loop

Normalize ξ, η to variance unity. Regress η on $x_1, \ldots, x_p$ jointly:

$$\eta = \alpha_1 x_1 + \ldots + \alpha_p x_p + \varepsilon_\eta .$$

Compute

$$\hat{\xi} = \sum_{i=1}^{p} \alpha_i x_i .$$

Regress $y_1, \ldots, y_q$ separately on $\hat{\xi}$:

$$\begin{aligned} y_1 &= \beta_1 \hat{\xi} + \varepsilon_1 \\ &\vdots \\ y_q &= \beta_q \hat{\xi} + \varepsilon_q . \end{aligned}$$

Compute

$$\hat{\eta} = \sum_{k=1}^{q} \beta_k y_k .$$

Test

If $\hat{\xi}$ not equal to ξ or $\hat{\eta}$ to η (within some chosen convergence criteria), loop again. Otherwise

Finish

Regress η on ξ for the parameter γ (the correlation between the Y-variate and the X-variate).

4. Equivalence to Johansson's Least Squares Solution

Following analysis similar to that outlined by Areskoug [1982] in the case of canonical correlation, we can show that the above algorithm leads to eigenvalue equations which could be solved for $\underset{\sim}{\alpha}$ and $\underset{\sim}{\beta}$. It also permits us to compare these eigenvalue equations with van den Wollenberg's and Johansson's to assess the equivalence of the approaches.

In PLS, the estimated latent variables (components) are defined as the linear forms:

$$\hat{\xi} = f_1 \underset{\sim}{x}\underset{\sim}{\alpha} \quad (1)$$

and $$\hat{\eta} = f_2 \underset{\sim}{y}\underset{\sim}{\beta} \quad (2)$$

where: $\underset{\sim}{x}$ and $\underset{\sim}{y}$ are matrices of observations from T cases and: f_1 and f_2 are scalar constants to give unit variance to $\hat{\xi}$ and $\hat{\eta}$.

$$\frac{1}{f_1^2} = T^{-1}\underset{\sim}{\alpha}'(x'x)\underset{\sim}{\alpha} = \underset{\sim}{\alpha}'\underset{\sim}{R}_{xx}\underset{\sim}{\alpha} \ ;$$

$$\frac{1}{f_2^2} = T^{-1}\underset{\sim}{\beta}'(y'y)\underset{\sim}{\beta} = \underset{\sim}{\beta}'\underset{\sim}{R}_{yy}\underset{\sim}{\beta} \ .$$

f is the scalar constant which transforms the eigenvector corresponding with the largest eigenvalue in classical eigenvalue equations to standardized weights.

The parameters of the redundancy model were estimated by PLS in the following iterative fashion.

<u>Start</u> Choose arbitrary weights for $\underset{\sim}{\beta}$ and from (2) let

$$\hat{\eta}^{(0)} = f_2^{(0)} \underset{\sim}{y}\underset{\sim}{\beta}^{(0)} \ . \quad (3)$$

<u>Step 1</u> Regress $\hat{\eta}^{(0)}$ on $\underset{\sim}{x}$ jointly to get $\underset{\sim}{\alpha}^{(0)}$

$$\underset{\sim}{\alpha}^{(0)} = (\underset{\sim}{x}'\underset{\sim}{x})^{-1}\underset{\sim}{x}'\hat{\eta}^{(0)}$$

and from (3)

$$\underset{\sim}{\alpha}^{(0)} = f_2^{(0)} \underset{\sim}{R}_{xx}^{-1}\underset{\sim}{R}_{xy}\underset{\sim}{\beta}^{(0)} \ . \quad (4)$$

Now from (1) let:

$$\hat{\xi}^{(0)} = f_1^{(0)} \underset{\sim}{x}\underset{\sim}{\alpha}^{(0)} \tag{5}$$

and regress $y_1, \ldots, y_q$ on $\hat{\xi}^{(0)}$ separately

$$\underset{\sim}{\beta}^{(0)} = r(\hat{\xi}^{(0)}, \underset{\sim}{y})$$

and from (5)

$$\underset{\sim}{\beta}^{(0)} = f_1^{(0)} \underset{\sim}{R}_{yx} \underset{\sim}{\alpha}^{(0)} \ . \tag{6}$$

Substitution of (6) into (4) gives $\underset{\sim}{\alpha}^{(1)}$ in terms of $\underset{\sim}{\alpha}^{(0)}$

$$\underset{\sim}{\alpha}^{(1)} = f_1^{(0)} f_2^{(0)} \underset{\sim}{R}_{xx}^{-1} \underset{\sim}{R}_{xy} \underset{\sim}{R}_{yx} \underset{\sim}{\alpha}^{(0)} \tag{7}$$

$$= f^{(0)} M_\alpha \underset{\sim}{\alpha}^{(0)} \tag{8}$$

where $f = f_1 f_2$ and $M_\alpha = \underset{\sim}{R}_{xx}^{-1} \underset{\sim}{R}_{xy} \underset{\sim}{R}_{yx}$.

Substitution of (4) into (6) gives $\underset{\sim}{\beta}^{(1)}$ in terms of $\underset{\sim}{\beta}^{(0)}$

$$\underset{\sim}{\beta}^{(1)} = f_1^{(0)} f_2^{(0)} \underset{\sim}{R}_{yx} \underset{\sim}{R}_{xx}^{-1} \underset{\sim}{R}_{xy} \underset{\sim}{\beta}^{(0)} \tag{9}$$

$$= f^{(0)} M_\beta \underset{\sim}{\beta}^{(0)} \tag{10}$$

where $M_\beta = \underset{\sim}{R}_{yx} \underset{\sim}{R}_{xx}^{-1} \underset{\sim}{R}_{xy}$.

<u>Step n</u>

$$\underset{\sim}{\alpha}^{(n)} = f^{(n-1)} M_\alpha \underset{\sim}{\alpha}^{(n-1)} \tag{11}$$

and

$$\underset{\sim}{\beta}^{(n)} = f^{(n-1)} M_\beta \underset{\sim}{\beta}^{(n-1)} \ . \tag{12}$$

Now if the iterative procedure converges we must have:

$$\lim_{n \to \infty} \underset{\sim}{\alpha}^{(n)} = \lim_{n \to \infty} \underset{\sim}{\alpha}^{(n-1)} = \underset{\sim}{\alpha} \tag{13}$$

and

$$\lim_{n\to\infty} \underset{\sim}{\beta}^{(n)} = \lim_{n\to\infty} \underset{\sim}{\beta}^{(n-1)} = \underset{\sim}{\beta} \; . \tag{14}$$

The iterative procedure $v^{(n)} = Av^{(n-1)}$ where v, scaled in some fashion known as the power method, converges to the eigenvector of A that corresponds to its largest eigenvalue [Morrison 1976, pp. 279-82].

Then the solution must satisfy the general eigenvalue equation:

$$(M_v - \lambda_v I)\underset{\sim}{v} = 0 \text{ where } v = \alpha, \beta \; . \tag{15}$$

To obtain nonzero solutions consider first $\underset{\sim}{\alpha}$:

From (15) and (7)

$$(\underset{\sim}{R}_{xx}^{-1}\underset{\sim}{R}_{xy}\underset{\sim}{R}_{yx} - \lambda_\alpha I)\underset{\sim}{\alpha} = 0$$

or equivalently

$$(\underset{\sim}{R}_{xy}\underset{\sim}{R}_{yx} - \lambda_\alpha \underset{\sim}{R}_{xx})\underset{\sim}{\alpha} = 0 \; . \tag{16}$$

Solving (16) for the largest root λ_α, we can then proceed to solve for $\underset{\sim}{\alpha}$.

Similarly for $\underset{\sim}{\beta}$, from (15) and (9) we have the eigenvalue equation

$$(\underset{\sim}{R}_{yx}\underset{\sim}{R}_{xx}^{-1}\underset{\sim}{R}_{xy} - \lambda_\beta I)\underset{\sim}{\beta} = 0 \tag{17}$$

which we solve for the largest eigenvalue λ_β and hence proceed to solve for $\underset{\sim}{\beta}$.

When comparing the above eigenvalue equations to those derived by van den Wollenberg we find that (16) is consistent with his general characteristic equation:

$$(\underset{\sim}{R}_{xy}\underset{\sim}{R}_{yx} - \lambda \underset{\sim}{R}_{xx})\underset{\sim}{\alpha} = 0$$

and hence the PLS solution and van den Wollenberg's redundancy solution give identical solutions for $\underset{\sim}{\alpha}$, i.e., the weights of the ξ variate, as expected.

Van den Wollenberg does not present a solution for β given the optimal ξ variate, however Johansson does for his least squares and restandardized solutions. Examining the least squares solution for the first η variate, Johansson arrives at his equation (6):

$$\beta' R_{yx} \alpha = \lambda \tag{18}$$

where λ indicates the correlation between the first variate pair.

Now, from (4) and (13)

$$\alpha = f_2 R_{xx}^{-1} R_{xy} \beta$$

and substituting into (18) we get

$$f_2 \beta' R_{yx} R_{xx}^{-1} R_{xy} \beta - \lambda I = 0 \tag{19}$$

post-multiplying by β and dividing by $f_2 \beta' \beta$ where $\beta' \beta$ is a scalar we arrive at

$$(R_{yx} R_{xx}^{-1} R_{xy} - \frac{\lambda}{f_2 \beta' \beta} I)\beta = 0 \quad \text{or} \quad (R_{yx} R_{xx}^{-1} R_{xy} - \lambda_\beta I)\beta = 0 \tag{20}$$

which is the same eigenvalue equation as (17) for the first η variate. Hence the PLS solution is equivalent to Johansson's least squares solution for the weights of both the ξ and η variates in the first redundancy pair.

The PLS results would be consistent with Johansson's for the second and higher order redundancy variates based on his argument regarding the general case for the j^{th} variate [Johansson 1981, p. 96]. Also similar equivalencie s can be found if the redundancy of the x-variables is examined given η variates.

5. Numerical Example

Van den Wollenberg's [1977] "artificial" data will be used to illustrate the equivalency of results using the PLS algorithm and results from Johansson's least squares extension. The input correlation matrix is presented in Table 1.

Using van den Wollenberg's data as analyzed by both van den Wollenberg and Johansson, we note that our results are completely consistent with

Johansson's (see Table 2). As expected, the correlation between ξ and η in the iterative algorithm and Johansson's extension is larger than in van den Wollenberg's case, while the opposite is true for the sum of the squared loadings of the y-variables on η. The net result is that the redundancy of the y-variables in all solutions is the same.

6. Conclusion

This paper has shown that a simple, easily implementable algorithm based on Wold's PLS, when applied to redundancy analysis, not only derives optimal X-variates to maximize redundancy with respect to the y-variables but also produces Y-variates which maximally correlate with the derived X-variates. In addition, this algorithm produces Y-variates with the desirable orthogonality properties of Johansson's least squares solution, i.e., $\underset{\sim}{\beta}'_{\ell}\underset{\sim}{R}_{yy}\underset{\sim}{\beta}_{m} = 0$, $\ell \neq m$.

TABLE 1. Van den Wollenberg's Correlation Matrix

	x_1	x_2	x_3	x_4	y_1	y_2	y_3
x_2	.800						
x_3	.140	.060					
x_4	.060	.140	.800				
y_1	-.003	.062	.422	.710			
y_2	.265	.203	.714	.440	.400		
y_3	.404	.709	-.142	.089	.200	.000	
y_4	.723	.461	-.012	-.037	.000	.200	.400

Table 2. Redundancy Analysis

	Wold's PLS	Van den Wollenberg's Redundancy Analysis	Johansson's Least Squares Extension
Correlation between ξ, η i.e. γ	.732	.689	.732
Redundancy of y-variables	.210	.210	a)
Weights of x-variables for ξ(i.e. X)			
x_1	.508	.508	a)
x_2	.413	.413	
x_3	-.266	-.266	
x_4	.606	.606	
Loadings of x-variables on ξ (i.e. X)			
x_1	.837	.837	a)
x_2	.888	.888	
x_3	.315	.315	
x_4	.482	.482	

	Wold's PLS	Van den Wollenberg's Redundancy Analysis	Johansson's Least Squares Extension
Weights of y-variables for η (i.e.Y)		b)	
y_1	.298	.294	.298
y_2	.257	.211	.257
y_3	.513	.598	.514
y_4	.468	.500	.467
Loadings of y-variables on η(i.e. Y)		b)	
y_1	.503	.498	c)
y_2	.470	.428	
y_3	.760	.856	
y_4	.725	.781	
Loadings of y-variables on ξ(i.e. X)		d)	
y_1	.343	.343	a)
y_2	.295	.295	
y_3	.589	.590	
y_4	.538	.538	

Table 2. (Cont'd)

Notes: a. Results, although not given, must be the same as in PLS and van den Wollenberg's Redundancy Analysis since Johansson's extensions do not change the redundancy or the ξ variate.

b. As van den Wollenberg does not present these results, nor does the REDANAL program [Thissen and van den Wollenberg 1975] produce them, these loadings and weights have been calculated from:

$$\lambda_{y_k} = \left(\frac{\left(\lambda_{y\xi_k}\right)^2}{\gamma^2}\right)^{1/2} \qquad k = 1, \ldots, q$$

and $\underset{\sim}{\beta} = \underset{\sim}{R}_{yy}^{-1}\underset{\sim}{\Lambda}_y$.

c. Since the weights of the y-variables for η in PLS and Johansson's extension are equivalent, so must be the loadings since

$$\underset{\sim}{\Lambda}_y = \underset{\sim}{R}_{yy}\underset{\sim}{\beta} .$$

d. It appears that there has been a transposition error in van den Wollenberg's Table 5. Thus the loadings presented here are incorrectly labeled in his Table 5 as the x-variable loadings on the Y-variates.

REFERENCES

Areskoug, B., The First Canonical Correlation: Theoretical PLS Analysis and Simulation Experiments. In K. G. Jöreskog and H. Wold (Eds.), Systems Under Indirect Observation--II. Amsterdam: North-Holland, 1982, pp. 95-117.

Dawson-Saunders, B. K., The Effect of Affine Transformation on Redundancy Analysis, Psychometrika, 1983, 48, 299-302.

DeSarbo, W. S., Canonical/Redundancy Factoring Analysis, Psychometrika, 1981, 46, 307-29.

Johansson, J. K., An Extension of Wollenberg's Redundancy Analysis, Psychometrika, 1981, 46, 93-103.

Morrison, D. F., Multivariate Statistical Methods, 2nd Edition. New York: McGraw-Hill, 1976.

Muller, K. E., Relationships Between Redundancy Analysis, Canonical Correlation, and Multivariate Regression, Psychometrika, 1981, 46, 139-42.

Stewart, D., and W. A. Love, A General Canonical Correlation Index, Psychological Bulletin, 1968, 70, 160-63.

Thissen, M., and A. L. van den Wollenberg, REDANAL: A Fortran G/H Program for Redundancy Analysis, Research Bulletin 26. Nijmegen, the Netherlands: University of Nijmegen, Department of Mathematical Psychology, 1975.

Tyler, D. E., On the Optimality of the Simultaneous Redundancy Transformations, Psychometrika, 1982, 47, 77-86.

van den Wollenberg, A. L., Redundancy Analysis--An Alternative for Canonical Correlation Analysis, Psychometrika, 1977, 42, 207-19.

Wold, H., Estimation of Principal Components and Related Models by Iterative Least Squares. In P. R. Krishnaiah (Ed.), Multivariate Analysis. New York: Academic Press, 1966, pp. 391-420.