Computational Statistics & Data Analysis 18 (1994) 331–340
North-Holland

Generalized canonical analysis based on optimizing matrix correlations and a relation with IDIOSCAL *

Henk A.L. Kiers
University of Groningen, Groningen, The Netherlands

Robert Cléroux
Université de Montréal, Montréal, Que, Canada

Jos M.F. Ten Berge
University of Groningen, Groningen, The Netherlands

Received December 1991
Revised April 1993

Abstract: Carroll's method for generalized canonical analysis of two or more sets of variables is shown to optimize the sum of squared inner-product matrix correlations between a consensus matrix and matrices with canonical variates for each set of variables. In addition, the method that analogously optimizes the sum of squared RV matrix correlations (proposed by Escoufier, 1973) between a consensus matrix and matrices with canonical variates, can be interpreted as an application of Carroll and Chang's IDIOSCAL. A simple algorithm is developed for this and other applications of IDIOSCAL where the similarity matrices are positive semi-definite.

Keywords: Matrix correlation; Generalized canonical analysis; IDIOSCAL

1. Introduction

Canonical correlation analysis is a method for finding "common" dimensions in two sets of variables measured on the same observation units. In Canonical correlation analysis, linear combinations (canonical variates) of the variables in

* This research has been made possible by a fellowship from the Royal Netherlands Academy of Arts and Sciences to the first author, and by a grant from the Natural Sciences and Engineering Research Council of Canada to the second author.
Correspondence to: H.A.L. Kiers, Department of Psychology, Grote Kruisstr. 2/1, 9712 TS Groningen, The Netherlands.

SSDI 0167-9473(93)E0022-V

both sets are constructed such that the canonical variates in one set have optimal correlations with corresponding canonical variates in the other. Carroll's (1968) generalized canonical analysis (GCA) is an extension to the analysis of two or more sets of variables where we maximize the sum of squared correlations of canonical variates with a "consensus" variable to be found. Explicitly, let X_k denote the $n \times m_k$ data matrix for set k, $k = 1, \dots, K$, and $\mathbf{a}_k$ a vector of weights to form a canonical variate $X_k \mathbf{a}_k$ for set k. Then GCA finds a consensus variable $\mathbf{z}$ such that

$$f_1(\boldsymbol{a}_1, \dots, \boldsymbol{a}_K, z) = \sum_{k=1}^{K} w_k r^2(z, X_k \boldsymbol{a}_k), \tag{1}$$

is maximized over $\mathbf{a}_1, \dots, \mathbf{a}_K$ and $\mathbf{z}$, where $r(\cdot,\cdot)$ denotes the correlation between the variables between parentheses, and w_k denotes a fixed (nonnegative) weight for set k. Carroll has shown that the solution for this problem is to take $\mathbf{z}$ equal to the first eigenvector of

$$Q = \sum_{k=1}^{K} w_k X_k (X_k' X_k)^{-1} X_k'. \tag{2}$$

The vectors $\mathbf{a}_1, \dots, \mathbf{a}_K$ can be obtained by regression as $\mathbf{a}_k = (X_k' X_k)^{-1} X_k' \mathbf{z}$. If more than one dimension is required, Carroll suggests taking a second dimension such that it is the best dimension orthogonal to the first one; the third is the best one orthogonal to the first two dimensions, etc. The complete r-dimensional solution is then given by a matrix Z containing the first r eigenvectors of Q. The solution for the $\mathbf{a}_k$ columns, now collected in $(m_k \times r)$ matrices A_k, is given by $A_k = (X_k' X_k)^{-1} X_k' Z$.

In the way Carroll proposed GCA, it is conceived of as a method in which dimensions are found successively and afterwards collected in one 'consensus matrix' Z. Although this does not imply that the complete matrix Z is optimal in any particular way, it is readily verified that GCA maximizes

$$\sum_{l=1}^{r} \sum_{k=1}^{K} w_k r^2(z_l, X_k \boldsymbol{a}_{kl}), \tag{3}$$

over arbitrary vectors $\mathbf{a}_{kl}$, $k = 1, \dots, K$, $l = 1, \dots, r$, and a matrix Z with columns $\mathbf{z}_1, \dots, \mathbf{z}_r$, subject to $Z'Z = I_r$ (e.g., Tenenhaus and Young, 1985, p. 100). Thus, GCA optimizes (3) both simultaneously and successively.

Lazraq et al. (1992) proposed a different way of handling more than one dimension simultaneously in (variants of) GCA by optimizing the sum of squared matrix correlations between $X_k A_k$ and Z, $k = 1, \dots, K$, instead of the sum of dimension-wise squared correlations. A number of different matrix correlation coefficients have been proposed (see Ramsay et al. 1984, for a review), and Lazraq et al. (1992) have developed variants of GCA based on two such matrix correlations coefficients. The first is the one implicitly proposed by Lingoes and Schönemann (1974), which they call RLS. This coefficient measures agreement between two matrices after they have been rotated to optimal agreement. If X and Y are two matrices of the same row order, then the RLS matrix correlation between X and Y is defined as

$$RLS(X, Y) = \frac{\operatorname{tr}\left[(X'YY'X)^{1/2}\right]}{(\operatorname{tr}(X'X)\operatorname{tr}(Y'Y))^{1/2}}. \tag{4}$$

Lazraq et al. proposed to maximize

$$f_2(A_1, \ldots, A_K, Z) = \sum_{k=1}^{K} w_k RLS^2(Z, X_k A_k), \tag{5}$$

over arbitrary $A_1, \ldots, A_K$ and columnwise orthonormal Z, as a method for the canonical analysis of more than two sets employing more than one dimension simultaneously. They demonstrate that maximizing the sum of squared matrix correlations in (5) is equivalent to Carroll's original GCA method, thus showing again that Carroll's method can be seen as a method in which the dimensions are optimized simultaneously, albeit based on a different criterion. However, this is not the only way GCA can be considered as maximizing a sum of squared matrix correlations between Z and $X_k A_k$, $k = 1, \ldots, K$. The first purpose of the present paper is to show that GCA also maximizes the sum of squared correlations computed according to a more popular matrix correlation coefficient, namely the "inner product" matrix correlation coefficient defined by Ramsay et al. (1984).

The main contribution of Lazraq et al.'s (1992) paper is in showing that Carroll's GCA is just *one* member of a family of methods for generalized canonical analysis based on matrix correlation coefficients. Employing other matrix correlation coefficients might lead to other potentially interesting methods for generalized canonical analysis. Lazraq et al. (1992) discussed one such method, based on the RV matrix correlation coefficient proposed by Escoufier (1973). The RV -coefficient expresses the correlation between X and Y as

$$RV(X, Y) = \frac{\operatorname{tr}(X'YY'X)}{\left[\operatorname{tr}(X'X)^2 \operatorname{tr}(Y'Y)^2\right]^{1/2}}. \tag{6}$$

Like the RLS-coefficient, the RV-coefficient measures agreement between two matrices up to a rotation. An advantage of the RV-coefficient is that it is easier to compute because it does not involve computation of a square root of a matrix. The RV-based variant of GCA proposed by Lazraq et al. optimizes

$$f_3(A_1, \ldots, A_K, Z) = \sum_{k=1}^{K} w_k RV^2(Z, X_k A_k), \tag{7}$$

in analogy to the RLS-based criterion. In the optimization of the function f_3, the matrix Z is again constrained to be columnwise orthonormal. The resulting method differs from Carroll's GCA. In fact, it seems that the RV-based method for GCA will tend to differentiate the contribution of different sets of variables (of which the scores are collected in the matrices $X_1, \ldots, X_K$) such that sets that contribute much to the ordinary matrix correlation will do even more so in the RV-based GCA. This can be understood most easily by considering the one-dimensional case in which $\mathrm{RV}(\mathbf{z}, X_k\mathbf{a}_k)$ is simply the *squared* correlation between $\mathbf{z}$ and $X_k\mathbf{a}_k$. Maximizing a sum of squared RV-coefficients then comes down to maximizing a sum of *fourth powers* of correlations, which will result in more extreme correlations than maximizing a sum of squared correlations would. In the case of more than one dimension a similar reasoning can be used. Thus, RV-based GCA finds a kind of implicit weighting of the sets, that might well obviate the use of a priori known weights w_k. The second purpose of the present paper is to propose an algorithm for the RV-based variant of GCA. In order to derive this algorithm, we first show that the RV-based variant of GCA is equivalent to IDIOSCAL applied to a particular set of (weighted) projector matrices. Next, a simple algorithm for this and similar applications of IDIOSCAL is proposed.

2. GCA as maximizing a sum of squared inner product matrix correlations

According to Ramsay et al. (1984), the most common matrix correlation coefficient is the 'inner product' matrix correlation coefficient, defined as

$$R(X, Y) = \frac{\mathrm{tr}(X'Y)}{(\mathrm{tr}(X'X)\ \mathrm{tr}(Y'Y))^{1/2}}. \tag{8}$$

It will now be shown that maximizing

$$g(A_1, \ldots, A_K, Z) = \sum_{k=1}^{K} w_k R^2(Z, X_k A_k), \tag{9}$$

over $A_1, \ldots, A_K$, subject to $Z'Z = I_r$ is equivalent to Carroll's GCA. Explicitly, maximizing (9) over A_k, $k = 1, \ldots, K$, disregarding Z, reduces to maximizing

$$g_k(A_k) = \frac{[\mathrm{tr}(Z'X_k A_k)]^2}{\mathrm{tr}(A_k' X_k' X_k A_k)} = \frac{\left[\mathrm{tr}\left(Z'X_k(X_k'X_k)^{-1/2}(X_k'X_k)^{1/2}A_k\right)\right]^2}{\mathrm{tr}(A_k' X_k' X_k A_k)}. \tag{10}$$

From the Cauchy–Schwartz inequality we have

$$\begin{aligned} g_k(A_k) &\leqslant \frac{\mathrm{tr}\left(Z'X_k(X_k'X_k)^{-1}X_k'Z\right)\ \mathrm{tr}(A_k'X_k'X_kA_k)}{\mathrm{tr}(A_k'X_k'X_kA_k)} \\ &= \mathrm{tr}\left(Z'X_k(X_k'X_k)^{-1}X_k'Z\right) \\ &= \mathrm{tr}(Z'P_kZ), \end{aligned} \tag{11}$$

where P_k is the projector matrix $P_k \equiv X_k(X_k'X_k)^{-1}X_k'$. The inequality in (11) gives an upper bound for $g_k(A_k)$, which is attained by taking $A_k = (X_k'X_k)^{-1}X_k'Z$. This shows that for every Z, the A_k that maximizes g_k can be expressed in terms of Z as $A_k = (X_k'X_k)^{-1}X_k'Z$. Note that this expression for

A_k is identical to the one obtained from GCA. Substituting this expression for A_k, $k = 1, \ldots, K$ in (9) we can rewrite g as

$$g(*, Z) = \sum_{k=1}^{K} w_k \operatorname{tr}(Z'P_k Z)$$
$$= \operatorname{tr} Z'\left(w_k \sum_{k=1}^{K} P_k\right)Z. \tag{12}$$

The original problem of maximizing (9) over A_k and Z has been reduced to maximizing (12) over Z only. The maximum of (12) over Z, subject to $Z'Z = I_r$, is obtained by taking Z equal to the matrix containing the first r eigenvectors of $Q = (\Sigma_{k=1}^{K} w_k P_k)$, which shows that the solution for Z also is equal to the one obtained from GCA. This establishes the equivalence of GCA and maximizing the (weighted) sum of squared inner-product matrix correlations between Z and A_k, $k = 1, \ldots, K$. Lazraq et al.'s (1992) result that GCA optimizes the sum of squared RLS matrix correlations between Z and A_k, $k = 1, \ldots, K$, follows at once. Specifically, the RLS-coefficient is the inner-product matrix correlation between two matrices after (optimal) rotation of one (say X_k) of the matrices to the other (Z) (see Lazraq et al., 1992). Because GCA maximizes $R(Z, X_k A_k)$ over arbitrary transformations A_k of X_k, it subsumes the optimal rotation of X_k to Z; hence, after optimizing A_k, $k = 1, \ldots, K$, $RLS(Z, X_k A_k) = R(Z, X_k A_k)$. This implies that GCA maximizes both $\Sigma_k w_k R^2(Z, X_k A_k)$ and $\Sigma_k w_k RLS^2(Z, X_k A_k)$.

3. RV-based GCA as IDIOSCAL applied to projector matrices

Lazraq et al. (1992) discussed a variant of GCA based on the (matrix correlation) coefficient RV, which will now be shown to be an application of IDIOSCAL (Carroll and Chang, 1970, 1972) to the projector matrices $P_1, \ldots, P_K$. Lazraq et al. demonstrated that the maximization problem in their RV-based variant of GCA can be solved as follows. Maximizing (7) over A_k, $k = 1, \ldots, K$, disregarding Z, reduces to maximizing

$$h_k(A_k) = \frac{[\operatorname{tr}(Z'X_k A_k A_k' X_k' Z)]^2}{\operatorname{tr}(X_k A_k A_k' X_k')^2}$$
$$= \frac{[\operatorname{tr}(ZZ'X_k A_k A_k' X_k')]^2}{\operatorname{tr}(X_k A_k A_k' X_k')^2}. \tag{13}$$

Using $P_k X_k = X_k$ and applying the Cauchy–Schwartz inequality to the numerator of (13), we have

$$h_k(A_k) \leqslant \frac{\operatorname{tr}(P_k ZZ'P_k)^2 \operatorname{tr}(X_k A_k A_k' X_k')^2}{\operatorname{tr}(X_k A_k A_k' X_k')^2} = \operatorname{tr}(Z'P_k Z)^2. \tag{14}$$

The upper bound (14) for h_k is attained for $A_k = (X_k'X_k)^{-1}X_k'Z$, which shows that for every Z the maximum over A_k is obtained in this way. Substituting this expression for A_k in the original function (7), and thereby eliminating A_k from the optimization problem, we end up with the problem of maximizing

$$f_3(*, Z) = \sum_{k=1}^{K} w_k \operatorname{tr}(Z'P_kZ)^2, \tag{15}$$

over Z subject to $Z'Z = I_r$. It will now be shown that maximizing (15) is equivalent to applying IDIOSCAL to the matrices $w_k^{1/2}P_k$, $k = 1,\ldots,K$.

IDIOSCAL is a method for multidimensional scaling of a number of $(n \times n)$ similarity matrices, $S_1,\ldots,S_K$. IDIOSCAL minimizes the function

$$\sigma_1(Z, C_1,\ldots,C_K) = \sum_{k=1}^{K} \| S_k - ZC_kZ' \|^2, \tag{16}$$

over arbitrary matrices Z $(n \times r)$ and C_k $(r \times r)$, $k = 1,\ldots,K$, where C_k is required to be positive semi-definite (p.s.d.). Without loss of generality, it can be assumed that Z is columnwise orthonormal. Then σ_1 can be written as

$$\sigma_1(Z, C_1,\ldots,C_K) = \sum_{k=1}^{K} \| S_k - ZZ'S_kZZ' \|^2 + \sum_{k=1}^{K} \| C_k - Z'S_kZ \|^2, \tag{17}$$

which is minimized over C_k by taking $C_k = Z'S_kZ$, for $k = 1,\ldots,K$. Clearly, if S_k is p.s.d., the obtained matrix C_k is p.s.d., and hence a valid IDIOSCAL solution. In the present paper, we only consider cases in which S_k is p.s.d., $k = 1,\ldots,K$. Substituting the above expression for $C_1,\ldots,C_K$ in σ_1, simplifies the problem to that of minimizing

$$\begin{aligned}\sigma_1(Z, *) &= \sum_{k=1}^{K} \| S_k - ZZ'S_kZZ' \|^2 \\ &= \sum_{k=1}^{K} \| S_k \|^2 - \sum_{k=1}^{K} \operatorname{tr}(Z'S_kZ)^2. \end{aligned} \tag{18}$$

It follows from (18) that minimizing σ_1 subject to $Z'Z = I_r$, and hence IDIOSCAL, is equivalent to maximizing

$$k(Z) = \sum_{k=1}^{K} \operatorname{tr}(Z'S_kZ)^2, \tag{19}$$

over Z, subject to $Z'Z = I_r$.

Above, it has been shown that the RV-based variant of GCA maximizes $\Sigma_{k=1}^{K} w_k \operatorname{tr}(Z'P_kZ)^2$, which can be rewritten as $\Sigma_{k=1}^{K} \operatorname{tr}[Z'(w_k^{1/2}P_k)Z]^2$. It follows from (19) that this method is equivalent to applying IDIOSCAL to the "similarity" matrices $S_k = w_k^{1/2}P_k$. Note that the condition that S_k be p.s.d. is satisfied for these weighted projector matrices.

An important application of GCA is multiple correspondence analysis (MCA; e.g., Tenenhaus and Young, 1985; also, see Gifi, 1990). MCA can be considered as an application of GCA to the indicator matrices for the nominal variables (each of which are considered a set of variables). Recently, Marchetti (1988) has proposed some alternatives for MCA, one of which is based on applying IDIOSCAL to the (centered) projectors derived from the indicator matrices. It follows directly from the relation derived above that Marchetti's method can be seen as the RV-based variant of GCA applied to the indicator matrices of the nominal variables.

4. A simple algorithm for IDIOSCAL applied to positive semi-definite matrices

Algorithms for approximating the IDIOSCAL solution have been suggested by Carroll and Chang (1970), and De Leeuw and Pruzansky (1978). The first monotonically convergent algorithm for IDIOSCAL was given by Kroonenberg and De Leeuw (1980, p. 78) who suggested using their TUCKALS2 algorithm to find the IDIOSCAL solution. The TUCKALS2 algorithm, however, minimizes the function

$$\sigma_2(X, Y, C_1, \ldots, C_K) = \sum_{k=1}^{K} \| S_k - XC_kY' \|^2, \tag{20}$$

which is equal to the IDIOSCAL function only if X is constrained to be equal to Y. However, it is not guaranteed that the optimizing X and Y matrices will be equal (Kiers, 1989, p. 516). For this reason, Kiers (1989) proposed a different algorithm that does minimize the IDIOSCAL function and converges monotonically. This algorithm is based on updating one column of Z at a time, and cycling through all columns during all iterations. However, in case TUCKALS2 is applied to p.s.d. matrices, the TUCKALS2 algorithm does converge to a solution with $X = Y$, as has been proved recently by Ten Berge et al. (in press). In the present paper, we follow an almost equivalent approach, where no distinction between X and Y is made at all. Specifically, Z is started with initial values $Z^{(0)}$ (e.g., a random columnwise orthonormal matrix). Next, $Z^{(0)}$ is updated by $Z^{(1)}$, $Z^{(1)}$ by $Z^{(2)}$, etc. In general, a current matrix Z is updated by a matrix

$$Z^{\mathrm{u}} = UV', \tag{21a}$$

where U and V are taken from the singular value decomposition

$$\sum_{k=1}^{K} S_k ZZ'S_k Z = UDV', \tag{21b}$$

with $U'U = V'V = I$ and D diagonal and nonnegative. This procedure is repeated until convergence. It will now be shown that this procedure monotonically increases the function $k(Z) = \sum_{k=1}^{K} \mathrm{tr}(Z'S_kZ)^2$, which is the function maximized in IDIOSCAL, subject to $Z'Z = I_r$. That is, it will be shown that $k(Z^{\mathrm{u}}) \geqslant k(Z)$.

To prove that $k(Z^{\mathrm{u}}) \geqslant k(Z)$, we develop the following inequalities. From

$$\| S_k^{1/2} ZZ' S_k^{1/2} - S_k^{1/2} Z^{\mathrm{u}} Z^{\mathrm{u}\prime} S_k^{1/2} \|^2 \geqslant 0, \tag{22}$$

it follows that

$$k(Z)+k(Z^{u}) = \sum_{k=1}^{K} \operatorname{tr}(Z'S_kZ)^2 + \sum_{k=1}^{K} \operatorname{tr}(Z^{u\prime}S_kZ^{u})^2 \geqslant 2\sum_{k=1}^{K} \operatorname{tr}(Z^{u\prime}S_kZ)(Z'S_kZ^{u}), \tag{23}$$

and from

$$\| Z'S_kZ - Z^{u\prime}S_kZ \|^2, \tag{24}$$

it follows that

$$\sum_{k=1}^{K} \operatorname{tr}(Z^{u\prime}S_kZ)(Z'S_kZ^{u}) \geqslant 2\sum_{k=1}^{K} \operatorname{tr}(Z^{u\prime}S_kZ)(Z'S_kZ) - k(Z). \tag{25}$$

Finally, the particular choice (21) of the update Z^{u} guarantees that

$$\sum_{k=1}^{K} \operatorname{tr} Z^{u\prime}(S_kZZ'S_kZ) \geqslant \sum_{k=1}^{K} \operatorname{tr} Z'(S_kZZ'S_kZ) = k(Z), \tag{26}$$

(Cliff, 1966). Rewriting (23) as

$$k(Z^{u}) \geqslant 2 \operatorname{tr} \sum_{k=1}^{K} (Z^{u\prime}S_kZ)(Z'S_kZ^{u}) - k(Z)$$

and combining this with (25), we find

$$\begin{aligned} k(Z^{u}) &\geqslant 2\sum_{k=1}^{K} \operatorname{tr}(Z^{u\prime}S_kZ)(Z'S_kZ^{u}) - k(Z) \\ &\geqslant 4\sum_{k=1}^{K} \operatorname{tr}(Z^{u\prime}S_kZ)(Z'S_kZ) - 3k(Z). \end{aligned} \tag{28}$$

Finally, from the combination of (28) with (26) it follows at once that $k(Z^{u}) \geqslant 4k(Z) - 3k(Z) = k(Z)$. So updating Z according to (21) increases the value of k, and because the function k is bounded from above, the function value will converge to a stable value.

Instead of updating Z according to (21), a slightly simpler procedure is to update Z as

$$Z^{v} = GS\left(\sum_{k=1}^{K} S_kZZ'S_kZ \right), \tag{29}$$

where GS denotes the Gram–Schmidt orthonormalized version of the matrix between parentheses. The matrices Z^{v} in (29) and Z^{u} in (21b) are columnwise orthonormal and span the same column spaces. Hence, they are the same up to an orthonormal transformation, that is, $Z^{v} = Z^{u}T$ for a certain orthonormal matrix T. It is readily verified that $\mathrm{k}(Z^{v}) = \mathrm{k}(Z^{u})$, and thus the update Z^{v} is just

Table 1
Average performance of old and new IDIOSCAL algorithms on 40 random data sets

Type of data	r	Computation time (in seconds)		Number of iterations		Comp. time per iter.	
		Old	New	Old	New	Old	New
$3\times10\times10$	3	37.8	14.8	8.9	27.5	4.3	0.5
$3\times20\times20$	3	402.5	43.1	12.1	34.5	33.3	1.3
$6\times10\times10$	3	58.0	23.1	11.8	30.3	4.9	0.7
$3\times10\times10$	6	108.5	24.3	20.6	37.3	5.3	0.7

as good as update Z^u but typically easier to calculate. This procedure has been proposed in a different context by Gifi (1990, p.99).

To get an impression of the performance of the new algorithm, we programmed it in PCMATLAB and compared it to the algorithm by Kiers (1989), which, for this occasion was also programmed in PCMATLAB. We first generated ten data sets consisting of three p.s.d. matrices, and analyzed these in three dimensions. For comparative purposes, we also analyzed ten $3 \times 20 \times 20$ data sets and ten $6 \times 10 \times 10$ data sets, all in three dimensions, and we analyzed ten $3 \times 10 \times 10$ data sets in six dimensions. All analyses were started by taking Z equal to the first r eigenvectors of $\Sigma_{k=1}^{K} S_k$. The algorithm was considered to have converged when the value of σ_2 decreased by less than 0.0001%. In all analyses the two algorithms obtained the same function value (up to at most 0.0001%). Average computation times (using a 80386/80387 processor), numbers of iterations, and computation times per iteration are reported in Table 1. It can be seen that the new algorithm was considerably faster in all conditions than Kiers' (1989) algorithm, even though the latter algorithm consistently used fewer iterations. As can be seen in the last columns of Table 1, this is caused by the fact that the iterations in Kiers' (1989) algorithm are considerably more expensive (i.e., at least 8 times as expensive) as those in the new algorithm.

5. Discussion

In the present paper it has been demonstrated that Carroll's GCA is just one possibility of analyzing a set of matrices by optimizing a well-known matrix correlation coefficient, and that alternatives, employing different matrix correlation coefficients, are of potential interest as well. Obviously, still other coefficients than RV or RLS might be used for that purpose.

It has been found in the present paper that the new algorithm for IDIOSCAL is considerably faster than the old one. One might contend that this difference is due to the use of an interpreter based language like PCMATLAB, which favors procedures with a minimum number of internal loops (as in our new algorithm). However, it can be verified that, for one single iteration, Kiers' (1989) algorithm not only computes a similar matrix product as (29) but, in addition, requires a full eigendecomposition of an $(n - r + 1) \times (n - r + 1)$ matrix, which, for moderate or large n and relatively small r can be expected to dominate the computational process. The fact that the old algorithm uses fewer iterations (2 to 3 times less) than the new algorithm will not counterbalance this computational drawback if n is at least of moderate size. Apart from this difference between the algorithms, the new algorithm has a considerable advantage in terms of ease of programming.

References

Carroll, J.D., Generalization of canonical correlation analysis to three or more sets of variables, *Proceedings of the 76th Convention of the American Psychological Association* 3 (1968), 227–228.

Carroll, J.D. and J.J. Chang, Analysis of individual differences in multidimensional scaling via an n-way generalization of "Eckart–Young" decomposition, *Psychometrika*, **35** (1970), 283–319.

Carroll, J.D. and J.J. Chang, *IDIOSCAL: A generalization of INDSCAL allowing IDIOsyncratic reference systems as well as an analytic approximation to INDSCAL*, Paper presented at the Spring Meeting of the Psychometric Society, Princeton, NJ, March 30-31, 1972.

Cliff, N., Orthogonal rotation to congruence, *Psychometrika*, **31** (1966) 33–42.

De Leeuw, J. and S. Pruzansky, A new computational method to fit the weighted Euclidean distance model, *Psychometrika*, **43** (1978) 479–490.

Escoufier, Y., Le traitement des variables vectorielles, *Biometrics*, **29** (1973) 751–760.

Gifi, A., *Nonlinear multivariate analysis*, (Wiley, New York, 1990).

Kiers, H.A.L., An alternating least squares algorithm for fitting the two- and three-way DEDICOM model and the IDIOSCAL model, *Psychometrika*, **54** (1989) 515–521.

Kroonenberg, P.M. and De Leeuw, J. (1980) Principal component analysis of three-mode data by means of alternating least squares algorithms, *Psychometrika*, **45**, 69–97.

Lazraq, A., R. Cléroux and H.A.L. Kiers, Mesures de liaison vectorielle et généralisation de l'analyse canonique (Measures of vectorial correlation and generalized canonical analysis), *Revue de Statistique Appliquée*, **39** (1992) 23–35.

Lingoes, J.C. and P.H. Schönemann, Alternative measures of fit for the Schönemann–Carroll matrix fitting algorithm, *Psychometrika*, **39** (1974) 423–427.

Marchetti, G.M., Three-way analysis of two-mode matrices of qualitative data, Research report no. 10. (Dept. of Statistics, University of Florence, 1988).

Ramsay, J.O., J.M.F. Ten Berge and G.P.H. Styan, Matrix correlation, *Psychometrika*, **49** (1984) 403–423.

Ten Berge, J.M.F., P.A. Bekker, and H.A.L. Kiers, Some clarifications of the TUCKALS2 algorithm applied to the IDIOSCAL problem, to appear to: *Psychometrika.*

Tenenhaus, M., and F.W. Young, An analysis and synthesis of multiple correspondence analysis, optimal scaling, dual scaling, homogeneity analysis and other methods for quantifying categorical multivariate data, *Psychometrika*, **50** (1985) 91–119.