INDEPENDENT COMPONENT AND FACTOR ANALYSIS

JAN DE LEEUW

ABSTRACT. We describe two ways to compute independent component analyses. Derivations do not depend on using a conceptualization in terms of random variables but are defined directly in term of finite matrices. The first technique is related to the methods used in polynomial component analysis, the second to array decomposition.

1. Fixed Rank Approximation

The problem we study in this paper is to approximate an $n \times m$ data matrix Y by a product XA', where X is $n \times r$ and A is $m \times r$. In *regression analysis* the matrix X is known, and we merely have to compute A. In *component* and *factor analysis* both X and A are unknown and must be computed. Using terminology familiar from psychometrics we call X the *components* and A the *loadings*.

Suppose $Y = XA'$. Take S to be any non-singular matrix of order r, with S^{-T} the transpose of its inverse. Define $\tilde{X} = XS$ and

Date: Monday 3rd May, 2010 — 16h 39min — Typeset in LUCIDA BRIGHT.
Key words and phrases. Independent Component Analysis, Principal Component Analysis, Factor Analysis.

$\tilde{A} = AS^{-T}$. Then $\tilde{X}\tilde{A}' = XA' = Y$. Thus components and loadings are not uniquely determined, and we generally impose some additional identification conditions. The most useful and familiar one is $X'X = I$, although this is still not enough to uniquely define the decomposition. The previous indeterminacy relation still applies if we choose S to be square orthonormal.

Two cases must be distinguished. If $r < \min(m, n)$ the approximation is called *component analysis* (CA), if $r > m$ it is called *factor analysis* (FA). In the case of factor analysis we need additional identification conditions. The usual one, in addition to $X'X = I$, is

$$\underset{m\times r}{A} = \left[\underset{b\times(r-m)}{L} \quad | \quad \underset{m\times m}{D}\right], \tag{1}$$

with D diagonal.

There is an huge amount of literature in psychometrics, much of it rather confusing, about the differences between component and factor analysis. Some of it is philosophical, and concerns the use and testability of "models" or "theories" in multivariate analysis. Some of it is algebraic, and discusses the remaining indeterminacies of the decomposition $Y = XA'$ under the constraints of factor analysis.

In this paper we are discuss the matrix approximation problem $Y \approx XA'$ under various constraints on X and/ or A, say $X \in \mathcal{X} \subseteq \mathbb{R}^{n\times r}$ and $A \in \mathcal{A} \subseteq \mathbb{R}^{m\times r}$. The approximation problem also covers (multivariate) regression, but we shall concentrate on the case in which both X and A are partially or completely unknown.

In order to compute the optimal approximation we use a *loss map* $F : \mathbb{R}^{n\times m} \Rightarrow \mathcal{R}$ and a *norm-like function* δ on $\mathcal{R}$. Here δ is norm-like if it is continuous, non-negative, and vanishes only at the origin. The loss map F maps the data, and the approximation, into the *loss space* $\mathcal{R}$, where we actually measure the quality of the approximation. Approximating the data then means minimizing the *loss* *function*

$$\sigma(X,A) \triangleq \delta(F(Y) - F(XA')). \tag{2}$$

An important special case uses the identity map $F(Y) = Y$. Another familiar example of a loss map is $F(Y) = Y'Y$. If X is constrained by $X'X = I$, then this means we must minimize $\delta(Y'Y - AA')$, and X has been eliminated from the approximation problem. The metric then can be defined by least squares, multinormal maximum likelihood, and so on. More complicated loss maps use the empirical distribution function or empirical characteristic function for F.

But no matter what loss map and what norm we choose, in all cases we have to answer the question if the best approximation exists, if it is unique, and how it can be computed.

2. Classical PCA/FA

Let us first look at case in which F is the identity. Many choices for the metric δ on $\mathbb{R}^{n\times m}$ have been proposed, but we shall limit ourselves to the least squares loss function

$$\sigma(X,A) = \mathbf{tr}\ (Y - XA')'W(Y - XA'), \tag{3}$$

with W is positive definite diagonal matrix of *weights*. Much more general weight structures are discussed in De Leeuw [1984].

We now minimize $\sigma(X,A)$ over all $m \times r$ matrices A and over the $n \times r$ matrices X that satisfy $X'WX = I$. Note that we do not assume $r \leq m$, but we do assume $r \leq n$. If $X'WX = I$ then

$$\min_A \sigma(X,A) = \sigma(X, Y'WX) = \mathbf{tr}\ Y'WY - X'WYY'WX$$

and thus, if $r < m$,

$$\min_{X'WX=I} \min_A \sigma(X,A) = \sum_{s=r+1}^{m} \lambda_s^2(W^{\frac{1}{2}}Y).$$

Here $\lambda_1() \geq \cdots \geq \lambda_m()$ are the ordered singular values of its matrix argument. If $r \geq \mathbf{rank}(Y)$, for instance because $r \geq m$, then $\min_{X'WX=I} \min_A \sigma(X,A) = 0$. It is shown in appendix A that

$$\min_{X'WX=I} \sigma(X,A) = \mathbf{tr}\ Y'WY + \mathbf{tr}\ A'A - 2\sum_{s=1}^{r} \lambda_s(W^{\frac{1}{2}}YA),$$

Thus minimizing the loss function can be done by maximizing $\sum_{s=1}^{r} \lambda_s(W^{\frac{1}{2}}YA) = \sum_{s=1}^{r} \sqrt{\lambda_s(A'Y'WYA)}$ over A satisfying $\mathbf{tr}\ A'A = 1$. If $r \geq m$ we must find A of unit Frobenius norm that maximizes the spectral norm of $W^{\frac{1}{2}}YA$.

In FA, using equation (1), the loss function is

$$\sigma(X,L,D) = \mathbf{tr}\ (Y - ZL' - UD)'W(Y - ZL' - UD).$$

The constraints are $Z'Z = I$, $U'U = I$, $Z'U = 0$ and D is diagonal. Alternating least squares methods to compute the best approximation are in De Leeuw [2004]. It is straightforward to adapt these methods to simple modifications of the technique that result, for example, from dropping the constraint $Z'Z = I$ or from requiring D to be block-diagonal.

3. Independence

Orthogonality of the components may not be a strong enough requirement to isolate independent effects. There are many situations in multivariate analysis in which successive components are functions of a single first component [Guttman, 1950; Kendall, 1971; Hill and Gauch, 1980; Diaconis et al., 2008]. This leads to polynomial components, also known as "horseshoes" [Van Rijckevorsel, 1987] or the "Effet Guttman" [Flament and Milland, 2005].

Consequently, it may be of interest to require "independence" instead of "uncorrelatedness". This is problematic, because independence is not an algebraic but a probabilistic concept, and it cannot readily be translated into a finite number of linear algebra constraints on the components. We can approximate the notion of independence, however, by requiring the X satisfies a set of *independence conditions* of the form

$$\sum_{i=1}^{n} w_i \prod_{s=1}^{p} x_{is}^{r_s} = \prod_{s=1}^{p} \sum_{i=1}^{n} w_i x_{is}^{r_s}.$$

Note that each vector $r = (r_1, \cdots, r_p)$ of non-negative integers defines an independence condition. Thus there is an infinite number these conditions, and from that infinite number we have to select a finite number to use as constraints.

Consider the integer vectors $\mathcal{R}_\rho$ of length p, whose elements add up to ρ. The independence constraints in $\mathcal{R}_0$ and $\mathcal{R}_1$ are trivially statisfied. The independence constraints in $\mathcal{R}_2$ are equivalent to the orthogonality constraint $X'WX = I$, and thus if we only impose $\mathcal{R}_2$ the constraints define linear CA (or FA, if the appropriate restrictions are imposed on A).

The $\mathcal{R}_3$ constraints are

$$\sum_{i=1}^{n} w_i x_{is} x_{it} x_{iu} = 0$$

for all triples $s < t < u$ and

$$\sum_{i=1}^{n} w_i x_{is}^2 x_{it} = 0$$

for all pairs $s \neq t$.

For the $\mathcal{R}_4$ constraints we have

$$\sum_{i=1}^{n} w_i x_{is} x_{it} x_{iu} x_{iv} = 0$$

for all quadruples $s < t < u < v$, and

$$\sum_{i=1}^{n} w_i x_{is} x_{it}^3 = 0$$

for all pairs $s \neq t$, and

$$\sum_{i=1}^{n} w_i x_{is} x_{it} x_{iu}^2 = 0$$

for all triples (s, t, u) with $s < t$ and $u \neq s, t$. We also have

$$\sum_{i=1}^{n} w_i x_{is}^2 x_{it}^2 = 1$$

for all pairs $s < t$.

4. Direct Least Squares

Minimize the loss function (3), with X constrained by a set of independence conditions, and optionally A constrained by linear constraints such as (1).

For an algorithm we could explore ALS with the columns of X as blocks. Partition X and A as

$$X = \begin{bmatrix} x & \mid & \overline{X} \end{bmatrix},$$
$$A = \begin{bmatrix} a & \mid & \overline{A} \end{bmatrix}.$$

Let $Z \triangleq Y - \overline{X}\overline{A}'$. Then

$$\sigma(X, A) = \mathbf{tr}\ (Z - xa')'W(Z - xa') =$$
$$= \mathbf{tr}\ Z'WZ - 2a'Z'Wx + x'Wx.a'a$$

If there are no restrictions on a, then finding the optimal x means maximizing the quadratic form $x'WZZ'Wx$ under the independence conditions, which include $x'Wx = 1$. If, as in FA, some elements of a are restricted to be equal to zero, we use a slightly modified quadrtaic form.

The situations is relatively simple if we require $\sum_{i=1}^n w_i x_i \overline{x}_{is} = 0$ for $s = 2, \cdots, p$, and $\sum_{i=1}^n w_i x_i \overline{x}_{is} \overline{x}_{it} = 0$ for $s, t = 2, \cdots, p$. These are simple orthogonality conditions, which can easily be extended to higher orders. Conditions such as $\sum_{i=1}^n w_i x_i^r \overline{x}_{is} = 0$ for $s = 2, \cdots, p$ are more complicated, however. If $r = 2$ they can be written as $x'V_s x = 0$, but if $r > 2$ we move into the more complicated realm of multilinear algebra.

Thus it is computationally convenient in an ALS approach to limit ourselves to independence constraints of the form

$$\sum_{i=1}^{n} w_i \prod_{s=1}^{p} x_{is}^{r_s} = 0,$$

with all r_s either zero or one. But these relatively simple constraints may not prevent horseshoes from occurring. We need some additional theory to deal with, for example, some constraints of the form

$$\sum_{i=1}^{n} w_i x_{is}^{r_s} x_{it}^{r_t} = \sum_{i=1}^{n} w_i x_{is}^{r_s} \sum_{i=1}^{n} w_i x_{it}^{r_t}.$$

for all $s \neq t$ and for selected $r_s, r_t \geq 1$.

5. Array Decomposition

Define

$$\psi(\eta) \stackrel{\Delta}{=} \log \sum_{i=1}^{n} w_i \exp\{\eta' A x_i\}.$$

A power series expansion gives

$$\psi(\eta) = \frac{1}{2}\sum_{i=1}^{n} w_i(\eta' A x_i)^2 + \frac{1}{6}\sum_{i=1}^{n} w_i(\eta' A x_i)^3 + \\ + \frac{1}{24}\left\{\sum_{i=1}^{n} w_i(\eta' A x_i)^4 - 3\{\sum_{i=1}^{n} w_i(\eta' A x_i)^2\}^2\right\} + \cdots$$

Define the symmetric k-dimensional array, of dimension $\underbrace{p \times \cdots \times p}_{k \text{ times}}$,

$$M^{(k)} \stackrel{\Delta}{=} \sum_{i=1}^{n} w_i\{\underbrace{x_i \otimes \cdots \otimes x_i}_{k \text{ times}}\}$$

Thus $M^{(1)} = 0$ and $M^{(2)} = X'WX$. We also use $M^{(k)}$ for the corresponding symmetric multilinear forms.

$$\psi(\eta) = \frac{1}{2}M^{(2)}(A'\eta, A'\eta) + \frac{1}{6}M^{(3)}(A'\eta, A'\eta, A'\eta) + \\ + \frac{1}{24}\left\{M^{(4)}(A'\eta, A'\eta, A'\eta, A\eta) - 3M^{(2)}(A'\eta, A'\eta)^2\right\} + \cdots$$

If the independence constraints are true, then $M^{(2)} = I$ and $M^{(3)}$ only has non-zero elements, the skewness γ_s, along the body diagonal. Thus

$$M^{(2)}(A'\eta, A'\eta) = \eta' AA'\eta,$$

and

$$M^{(3)}(A'\eta, A'\eta, A'\eta) = \sum_{s=1}^{p} \gamma_s (a_s'\eta)^3.$$

Element (s_1, s_2, s_3, s_4) of $M^{(4)}$ is equal to 1 if the four indices are made up of two pairs of equal indices. It is equal to the kurtosis κ_s if all four indices are equal. Thus

$$M^{(4)}(A'\eta, A'\eta, A'\eta, A\eta) = \sum_{s=1}^{p} (\kappa_s - 3)(a_s'\eta)^4 + 3\{\sum_{s=1}^{p} (a_s'\eta)^2\}^2$$

Thus, if the independence constraints are true,

$$\psi(\eta) = \frac{1}{2}\sum_{s=1}^{p}(a_s\eta)^2 + \frac{1}{6}\sum_{s=1}^{p}\gamma_s(a_s\eta)^3 + \frac{1}{24}\sum_{s=1}^{p}(\kappa_s - 3)(a_s\eta)^4 + \cdots$$

If we differentiate we find

$$\begin{aligned}
\mathcal{D}^{(1)}\psi(0) &= 0,\\
\mathcal{D}^{(2)}\psi(0) &= \sum_{s=1}^{p}\{a_s \otimes a_s\},\\
\mathcal{D}^{(3)}\psi(0) &= \sum_{s=1}^{p}\gamma_s\{a_s \otimes a_s \otimes a_s\},\\
\mathcal{D}^{(4)}\psi(0) &= \sum_{s=1}^{p}(\kappa_s - 3)\{a_s \otimes a_s \otimes a_s \otimes a_s\}.
\end{aligned}$$

Thus we can find A by solving one or several of the equation arrays

$$\begin{aligned}
\sum_{i=1}^{n} w_i\{y_i \otimes y_i\} &= \sum_{s=1}^{p}\{a_s \otimes a_s\},\\
\sum_{i=1}^{n} w_i\{y_i \otimes y_i \otimes y_i\} &= \sum_{s=1}^{p}\gamma_s\{a_s \otimes a_s \otimes a_s\},
\end{aligned}$$

and

$$\sum_{i=1}^{n} w_i\{y_i\otimes y_i\otimes y_i\otimes y_i\}-3\left\{\sum_{i=1}^{n} w_i\{y_i\otimes y_i\}\right\}\bigotimes\left\{\sum_{i=1}^{n} w_i\{y_i\otimes y_i\}\right\}= \\ =\sum_{s=1}^{p}(\kappa_s-3)\{a_s\otimes a_s\otimes a_s\otimes a_s\}.$$

More generally all multivariate cumulants of Y under independence have this symmetric array decomposition (INDSCAL-PARAFAC) structure. Note that these array decomposition methods do not give a solution for X, only for A. Also note that there is no restriction that $p \leq m$, only that $p \leq n$.

There are many algorithms available to (approximately) solve these equations. One simple way to proceed is to use the method proposed by De Leeuw and Pruzansky [1978], but the ICA literature [Hyvärinen et al., 2001, Chapter 11] also discusses methods such as FOBI and JADE.

6. Algorithms

6.1. **Using Second and Third Order Information.**

6.2. **Simultaneous Diagonalization of the Third Order Array.**

6.3. **Coordinate Descent.**

$$\tilde{a}_{is}=a_{is}+\delta^{i\ell}\delta^{s\tau}\theta_{is}$$

$$\tilde{a}_{is}\tilde{a}_{js}\tilde{a}_{ks}=(a_{is}+\delta^{i\ell}\delta^{s\tau}\theta)(a_{js}+\delta^{j\ell}\delta^{s\tau}\theta)(a_{ks}+\delta^{k\ell}\delta^{s\tau}\theta)$$

$$\tilde{a}_{is}\tilde{a}_{js}\tilde{a}_{ks}=a_{is}a_{js}a_{ks}+\theta\delta^{s\tau}(a_{is}a_{js}\delta^{k\ell}+a_{is}a_{ks}\delta^{j\ell}+a_{js}a_{ks}\delta^{i\ell})+ \\ +\theta^2\delta^{s\tau}(a_{is}\delta^{j\ell}\delta^{k\ell}+a_{js}\delta^{i\ell}\delta^{k\ell}+a_{ks}\delta^{i\ell}\delta^{j\ell})+\theta^3\delta^{s\tau}\delta^{i\ell}\delta^{j\ell}\delta^{k\ell}.$$

$$\sum_{s=1}^{p}\gamma_s\tilde{a}_{is}\tilde{a}_{js}\tilde{a}_{ks}=\sum_{s=1}^{p}\gamma_s\tilde{a}_{is}\tilde{a}_{js}\tilde{a}_{ks}+\theta\gamma_\tau(a_{i\tau}a_{j\tau}\delta^{k\ell}+a_{i\tau}a_{k\tau}\delta^{j\ell}+a_{j\tau}a_{k\tau}\delta^{i\ell})+$$
$$+\theta^2\gamma_\tau(a_{i\tau}\delta^{j\ell}\delta^{k\ell}+a_{j\tau}\delta^{i\ell}\delta^{k\ell}+a_{k\tau}\delta^{i\ell}\delta^{j\ell})+\theta^3\gamma_\tau\delta^{i\ell}\delta^{j\ell}\delta^{k\ell}$$

$$\tilde{r}_{ijk}=\sum_{s=1}^{p}\gamma_s\tilde{a}_{is}\tilde{a}_{js}\tilde{a}_{ks}-c_{ijk}$$

$$\tilde{r}_{ijk}=r_{ijk}+\theta\gamma_\tau(a_{i\tau}a_{j\tau}\delta^{k\ell}+a_{i\tau}a_{k\tau}\delta^{j\ell}+a_{j\tau}a_{k\tau}\delta^{i\ell})+$$
$$+\theta^2\gamma_\tau(a_{i\tau}\delta^{j\ell}\delta^{k\ell}+a_{j\tau}\delta^{i\ell}\delta^{k\ell}+a_{k\tau}\delta^{i\ell}\delta^{j\ell})+\theta^3\gamma_\tau\delta^{i\ell}\delta^{j\ell}\delta^{k\ell}$$

$$\tilde{\sigma}=\sum_{i=1}^{m}\sum_{j=1}^{m}\sum_{k=1}^{m}\tilde{r}_{ijk}^2$$

Obviously $\tilde{\sigma}$ is a polynomial of degree 6 in θ. Choose θ to be the real root which minimize $\tilde{\sigma}$, and cycle through the $m\times p$ coordinates.

Appendix A. Augmented Procrustus

Suppose X is an $n \times m$ matrix of rank r. Consider the problem of maximizing $\mathbf{tr}\ U'X$ over the $n \times m$ matrices U satisfying $U'U = I$. This is known as the *Procrustus* problem, and it is usually studied for the case $n \geq m = r$. We want to generalize to $n \geq m \geq r$. For this, we use the singular value decomposition

$$X = \begin{bmatrix} \underset{n\times r}{K_1} & \underset{n\times(n-r)}{K_0} \end{bmatrix} \begin{bmatrix} \underset{r\times r}{\Lambda} & \underset{r\times(m-r)}{0} \\ \underset{(n-r)\times r}{0} & \underset{(n-r)\times(m-r)}{0} \end{bmatrix} \begin{bmatrix} \underset{r\times m}{L_1'} \\ \underset{(m-r)\times m}{L_0'} \end{bmatrix}.$$

Theorem A.1. *The maximum of* $\mathbf{tr}\ U'X$ *over* $n \times m$ *matrices* U *satisfying* $U'U = I$ *is* $\mathbf{tr}\ \Lambda$*, and it is attained for any* U *of the form* $U = K_1L_1' + K_0VL_0'$*, where* V *is any* $(n-r)\times(m-r)$ *matrix satisfying* $V'V = I$*.*

Proof. Using a symmetric matrix of Lagrange multipliers leads to the stationary equations $X = UM$, which implies $X'X = M^2$ or $M = \pm(X'X)^{1/2}$. It also implies that at a solution of the stationary equations $\mathbf{tr}\ U'X = \pm\mathbf{tr}\ \Lambda$. The negative sign corresponds with the minimum, the positive sign with the maximum.

Now

$$M = \begin{bmatrix} \underset{m\times r}{L_1} & \underset{m\times(m-r)}{L_0} \end{bmatrix} \begin{bmatrix} \underset{r\times r}{\Lambda} & \underset{r\times(m-r)}{0} \\ \underset{(m-r)\times r}{0} & \underset{(m-r)\times(m-r)}{0} \end{bmatrix} \begin{bmatrix} \underset{r\times m}{L_1'} \\ \underset{(m-r)\times m}{L_0'} \end{bmatrix}.$$

If we write U in the form

$$U = \begin{bmatrix} \underset{n\times r}{K_1} & \underset{n\times(n-r)}{K_0} \end{bmatrix} \begin{bmatrix} \underset{r\times m}{U_1} \\ \underset{(n-r)\times m}{U_0} \end{bmatrix}$$

then $X = UM$ can be simplified to

$$U_1L_1 = I,$$

$$U_0L_1 = 0,$$

with in addition, of course, $U_1'U_1 + U_0'U_0 = I$. It follows that $U_1 = L_1'$ and

$$\underset{(n-r)\times m}{U_0} = \underset{(n-r)\times(m-r)}{V} \underset{(m-r)\times m}{L_0'},$$

with $V'V = I$. Thus $U = K_1L_1' + K_0VL_0'$. □

Appendix B. Decomposing a symmetric matrix

Suppose C is positive semi-definite or order n and rank $r \leq n$. Suppose $p \geq r$. Describe all $n \times p$ matrices X such that $C = XX'$. The eigenvalue-decomposition of C is

$$C = \begin{bmatrix} \underset{n\times r}{K} & \underset{n\times(n-r)}{\overline{K}} \end{bmatrix} \begin{bmatrix} \underset{r\times r}{\Lambda^2} & \underset{r\times(n-r)}{0} \\ \underset{(n-r)\times r}{0} & \underset{(n-r)\times(n-r)}{0} \end{bmatrix} \begin{bmatrix} \underset{r\times n}{K'} \\ \underset{r\times(n-r)}{\overline{K}'} \end{bmatrix}$$

Now write X in the form

$$X = \begin{bmatrix} \underset{n\times r}{K} & \underset{n\times(n-r)}{\overline{K}} \end{bmatrix} \begin{bmatrix} \underset{r\times p}{P} \\ \underset{(n-r)\times p}{Q} \end{bmatrix}.$$

Then $C = XX'$ becomes

$$\begin{bmatrix} \underset{r\times r}{\Lambda^2} & \underset{r\times(n-r)}{0} \\ \underset{(n-r)\times r}{0} & \underset{(n-r)\times(n-r)}{0} \end{bmatrix} = \begin{bmatrix} \underset{r\times r}{PP'} & \underset{r\times(n-r)}{PQ'} \\ \underset{(n-r)\times r}{QP'} & \underset{(n-r)\times(n-r)}{QQ'} \end{bmatrix}$$

Thus $Q = 0$ and $PP' = \Lambda^2$. It follows that $X = K\Lambda L'$, where L is $p \times r$ and satisfies $L'L = I$.

References

J. De Leeuw. Fixed-rank Approximation with Singular Weight Matrices. *Computational Statistics Quarterly*, 1:3–12, 1984.

J. De Leeuw. Least Squares Optimal Scaling of Partially Observed Linear Systems. In K. van Montfort, J. Oud, and A. Satorra, editors, *Recent Developments on Structural Equation Models*, chapter 7. Kluwer Academic Publishers, Dordrecht, Netherlands, 2004.

J. De Leeuw and S. Pruzansky. A New Computational Method to fit the Weighted Euclidean Distance Model. *Psychometrika*, 43: 479–490, 1978.

P. Diaconis, S. Goel, and S. Holmes. Horseshoes in Multidimensional Scaling and Kernel Methods. *Annals of Applied Statistics*, 2:777–807, 2008.

C. Flament and L. Milland. Un Effet Guttman en ACP. *Mathématiques & Sciences Humaines*, 43:25–49, 2005.

L. Guttman. The Principal Components of Scale Analysis. In S.A. Stouffer and Others, editors, *Measurement and Prediction*. Princeton University Press, Princeton, 1950.

M.O. Hill and H.G. Gauch. Detrended Correspondence Analysis: An Improved Ordination Technique. *Vegetatio*, 42:47–58, 1980.

A. Hyvärinen, J. Karhunen, and E. Oja. *Independent Component Analysis*. John Wiley & Sons Inc., New York, N.Y., 2001.

D.G. Kendall. Seriation from Abundance Matrices. In F.R. Hodson, D.G. Kendall, and P. Tăutu, editors, *Mathematics in the Archeological and Historical Sciences*, pages 215–252, Edinburgh, 1971. Edinburgh University Press.

J.L.A. Van Rijckevorsel. *The Application of Fuzzy Coding and Horseshoes in Multiple Correspondence Analysis*. PhD thesis, University of Leiden, The Netherlands, 1987. Also published in 1987 by DSWO-Press, Leiden, The Netherlands.

Department of Statistics, University of California, Los Angeles, CA 90095-1554

E-mail address, Jan de Leeuw: deleeuw@stat.ucla.edu

URL, Jan de Leeuw: http://gifi.stat.ucla.edu