NON-EXISTENCE OF NON-METRIC COMMON FACTOR ANALYSIS SOLUTIONS

JAN DE LEEUW

ABSTRACT. We show that in most cases of practical interest non-metric common factor analysis solutions do not exist.

1. INTRODUCTION

In non-metric common factor analysis (NMCFA), as implemented for example in the `FACTALS` program [Takane et al., 1979], the loss function

$$\sigma(X,A,D) = \|R(X) - AA' - D\| \tag{1}$$

is minimized over X, A, and D. In the NMCFA minimization X varies over the set $\mathscr{X} \subseteq \mathbb{R}^{n\times m}$ of the cone of column-wise monotone transformations of the data matrix Y. We also suppose matrices in $\mathscr{X}$ are standardized, i.e. all columns add up to zero and have unit sum-of-squares. The matrix $R(X) = X'X$ is the correlation matrix corresponding with $X \in \mathscr{X}$. Matrix $A \in \mathscr{A} = \mathbb{R}^{m\times p}$ is the matrix of common factor loadings, while $D \in \mathscr{D}$, the set of all non-negative diagonal matrices of order m, are the unique variances. The norm used in defining loss function (1) is usually a least squares norm, but for our results below any norm will do.

In NMCFA problems, as in similar non-metric problems using the approach of Kruskal [1964], we distinguish the primary and secondary approach to ties (see also De Leeuw [1977]). In the primary approach tied data can be untied. Thus if the data are $z_1 < z_2 = z_3 < z_4$ then we require $x_1 \leq x_2 \leq x_4$ and $x_1 \leq x_3 \leq x_4$, but the order of x_2 and x_3 is undecided. In the secondary approach we require $x_1 \leq x_2 = x_3 \leq x_4$, so ties are maintained. Clearly $\mathscr{X}$ is larger for the primary approach than for the secondary approach.

Date: Sunday 5th April, 2009 — 20h 36min — Typeset in TIMES ROMAN.

2000 *Mathematics Subject Classification.* 00A00.

Key words and phrases. Binomials, Normals, LaTeX.

Any feasible triple (X,A,D) where a local minimum of the loss function is attained is called an *NMCFA solution*. If $\sigma(X,A,D)=0$ then (X,A,D) is called a *perfect NMFCA solution*.

2. Result

Lemma 2.1. *Suppose (X,A,D) is an NMCFA solution. Suppose there exists an $n\times m$ matrix Z such that $Z'X=0$ and $Z'Z=I$ and such that*

$$X(\varepsilon)=\frac{1}{\sqrt{1+\varepsilon^2}}(X+\varepsilon Z)$$

is feasible for some $\varepsilon>0$. Then (X,A,D) is a perfect NMCFA solution.

Proof. The correlation matrix of $X(\varepsilon)$ is

$$R(X(\varepsilon))=\frac{1}{1+\varepsilon^2}R(X)+\frac{\varepsilon^2}{1+\varepsilon^2}I.$$

Define $A(\varepsilon)=\frac{1}{\sqrt{1+\varepsilon^2}}A$ and $D(\varepsilon)=\frac{1}{1+\varepsilon^2}D+\frac{\varepsilon^2}{1+\varepsilon^2}I$. Then

$$\sigma(X(\varepsilon),A(\varepsilon),D(\varepsilon))=\frac{1}{1+\varepsilon^2}\sigma(X,A,D),$$

and unless (X,A,D) is perfect $\sigma(X(\varepsilon),A(\varepsilon),D(\varepsilon))<\sigma(X,A,D)$. □

Remark 1. The feasibility condition on Z is equivalent to $z_{ij}\geq z_{kj}$ for all i,j,k such that $y_{ij}>y_{kj}$ and $x_{ij}=x_{kj}$.

Theorem 2.2. *Suppose $R(X)=AA'+D$ and there exists Z with $Z'Z=I$ and $X'Z=0$ such that $X(\varepsilon)=X+\varepsilon ZA'$ is feasible for some $\varepsilon>0$. If*

$$A(\varepsilon)=\sqrt{1+\varepsilon^2}\{(1+\varepsilon^2)\mathbf{diag}(AA')+D\}^{-\frac{1}{2}}A,$$
$$D(\varepsilon)=\{(1+\varepsilon^2)\mathbf{diag}(AA')+D\}^{-1}D,$$

then

$$R(X(\varepsilon))=A(\varepsilon)A(\varepsilon)'+D(\varepsilon).$$

Theorem 2.3. *Let f be a mapping of $\{1,2,\cdots,n\}$ into $\{1,2,\cdots,r\}$. If $y_{ij}\leq y_{kj}$ for all $j=1,\cdots,m$ whenever $f(i)<f(k)$ then there is an $X\in\mathscr{X}$ such that $\mathbf{rank}(R(X))=1$.*

Proof. Simply set $x_{ij}=f(i)$. □

Remark 2. The theorem applies if one observation is dominant (higher on all variables), or if one observation is dominated (lower on all variables).

3. Discussion

Note that the theorem depends critically on the fact that the unique variances are available to fit the diagonal elements of the correlation matrix. In non-metric principal component analysis [De Leeuw, 2006] the sum of the largest p eigenvalues of $R(X)$ is maximized. The eigenvalues of $R(X(\varepsilon))$ are $\frac{\lambda_s+\varepsilon^2}{1+\varepsilon^2}$, where the λ_s are the eigenvalues of $R(X)$. It follows that the sum of the p largest ones is a decreasing function of ε^2, which is maximized for $\varepsilon^2 = 0$.

References

J. De Leeuw. Correctness of Kruskal's Algorithms for Monotone Regression with Ties. *Psychometrika*, 42:141–144, 1977.

J. De Leeuw. Nonlinear Principal Component Analysis and Related Techniques. In M. Greenacre and J. Blasius, editors, *Multiple Correspondence Analysis and Related Methods*. Chapman and Hall, 2006.

J.B. Kruskal. Nonmetric Multidimensional Scaling: a Numerical Method. *Psychometrika*, 29:115–129, 1964.

Y. Takane, F.W. Young, and J. De Leeuw. Nonmetric Common Factor Analysis: an Alternating Least Squares Method with Optimal Scaling Features. *Behaviormetrika*, 6:45–56, 1979.

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

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