UNBOUNDED LIKELIHOODS IN ARRAY MULTINORMAL GROWTH CURVE MODELS

JAN DE LEEUW

ABSTRACT. Meet the abstract. This is the abstract.

1. INTRODUCTION

The negative log-likelihood for a matrix normal model with data $Y \in \mathbb{R}^{n\times m}$ is

$$\sigma(V,W,X,B,Z) = \\ m\log\mathbf{det}(V) + n\log\mathbf{det}(W) + \mathbf{tr}\, V^{-1}(Y - XBZ')W^{-1}(Y - XBZ')'.$$

Suppose

- $\mathcal{V}$ is a set of positive definite matrices or order n,
- $\mathcal{W}$ is a set of positive definite matrices or order m,
- $\mathcal{X}$ is a set of $n \times p$ matrices,
- $\mathcal{Z}$ is a set of $m \times q$ matrices,
- $\mathcal{B}$ is a set of $p \times q$ matrices.

Define

$$\sigma(\mathcal{V},\mathcal{W},\mathcal{X},\mathcal{B},\mathcal{Z}) = \inf_{V\in\mathcal{V}}\inf_{W\in\mathcal{W}}\inf_{X\in\mathcal{X}}\inf_{Z\in\mathcal{Z}}\inf_{B\in\mathcal{B}}\sigma(V,W,X,B,Z)$$

Theorem 1.1. *If $\mathcal{V}$ and $\mathcal{W}$ contain the set of all diagonal matrices, and if*

$$\mathcal{M} = \begin{bmatrix}\mu'\\ \mu'\\ \vdots\\ \mu'\end{bmatrix}$$

Date: Saturday 8th August, 2009 — 0h 0min — Typeset in LUCIDA BRIGHT.

then $\sigma(Y \mid \mathcal{V}, \mathcal{W}, \mathcal{M}) = -\infty$*, i.e. the maximum likelihood estimates of* $(\mathcal{V}, \mathcal{W}, \mathcal{M})$ *do not exist.*

Proof. Take $W = I$ and μ equal to the first row of Y. V is diagonal, with all elements equal to one, except the first which is equal to ϵ. Then

$$m \log \mathbf{det}(V) + n \log \mathbf{det}(W) + \mathbf{tr}\ V^{-1}(Y - M)W^{-1}(Y - M)' = \\ = m \log \epsilon + \mathbf{tr}(Y - M)'(Y - M)$$

which goes to $-\infty$ if $\epsilon \to 0$. □

Obviously this result remains true if $\mathcal{M}$ is defined any way that makes it possible to fit one row or one column of Y exactly. This implies the usual Anderson-Rubin result for fixed score ML factor analysis. The result is not true, at least not necessarily true, if $\mathcal{V}$ is a set of correlation matrices, or if $M = XBZ'$ with X and Z fixed.

Alternatively we can use residual maximum likelihood. If K is $n \times (n-1)$ and orthonormal, with columns that add up to zero, then $K'M = 0$, and thus we can compute REML estimates by minimizing

$$m \log \mathbf{det}(K'VK) + (n-1) \log \mathbf{det}(W) + \mathbf{tr}\ (K'VK)^{-1}K'YW^{-1}Y'K.$$

For $M = XBZ'$ with X and Z known, we can compute the likelihood of $X_\perp' Y Z_\perp$, using bases for the orthogonal complements of X and Z. This keep us in the domain of the matrix normal. In particular, if $V = \sum \theta_s A_s$ and $W = \sum \xi_s B_s$ have linear structure, then in REML they still have linear structure.

NB: If we apply this REML approach to fixed score factor analysis, we must define the likelihood of $Y - \Pi_r(Y)$, where P_r defines the best rank p approximation to Y. This can be done, at least in principle, by using the Jacobian of the singular value decomposition (see http://preprints.stat.ucla.edu/547/lowrank.pdf). This provides an alternative to the Anderson-Rubin non-central Wishart method, and to MacDonald's maximum likelihood ratio method. That, however, is another project.

2. General

Suppose $\mathcal{P}_n$ is the cone of all positive definite matrices of order n, and $\overline{\mathcal{P}}_n$ is its closure, i.e. the set of all positive semi-definite matrices. For $V \in \mathcal{P}_n$ we also write $V \succ 0$ and for $V \in \overline{\mathcal{P}}_n$ we write $V \succeq 0$.

Suppose $S \succeq 0$. Define $f_S : \mathcal{P}_n \Rightarrow \mathbb{R}$ by

$$f_S(V) = \log \mathbf{det}(V) + \mathbf{tr}\, V^{-1}S,$$

and

$$f_S(\mathcal{V}) = \inf_{V \in \mathcal{V}} f_S(V).$$

Theorem 2.1. *$f_S(\mathcal{P}_n) > -\infty$ if and only if $S \succ 0$.*

Proof. If $S \succ 0$ then let $W = S^{-\frac{1}{2}} V S^{-\frac{1}{2}}$. Now

$$f_S(\mathcal{P}_n) = \log \mathbf{det}(S) + \inf_{W \in \mathcal{P}_n} \left\{ \log \mathbf{det}(W) + \mathbf{tr}\, W^{-1} \right\}.$$

If $\lambda_i > 0$ are the eigenvalues of W, then

$$\log \mathbf{det}(W) + \mathbf{tr}\, W^{-1} = \sum_{i=1}^{n} \left\{ \log \lambda_i + \frac{1}{\lambda_i} \right\},$$

which attains a minimum equal to n if $\lambda_i = 1$ for all i, which means that $W = I$ and $V = S$.

Suppose $S \succeq 0$ and $\mathbf{rank}(S) = r < n$. Let $K_\perp$ be an $n \times (n-r)$ matrix with an orthonormal basis for the null space of S. Define $V(\epsilon) = S + \epsilon K_\perp K_\perp'$. Then $\mathbf{tr}\, V^{-1}(\epsilon) S = r$ and $\log \mathbf{det}(V(\epsilon)) = \sum_{i=1}^{r} \log \sigma_i + (n-r) \log \epsilon$, where the σ_i are the non-zero eigenvalues of S. Thus $\lim_{\epsilon \to 0} f_S(V(\epsilon)) = -\infty$. □

Corollary 2.2. *If $S \succ 0$ then $f_S(\mathcal{V}) \geq \log \mathbf{det}(S) + n$.*

Proof. $f_S(\mathcal{V}) \geq f_S(\mathcal{P}_n)$. The proof of Theorem 2.1 shows $f_S(\mathcal{P}_n) = \log \mathbf{det}(S) + n$. □

Thus if $S \succ 0$ the loss function is bounded below. Thus in the sequel we suppose $S \succeq 0$ is singular of rank r. Then $S = K\Lambda K'$, with K an $n \times r$ orthonormal, $K'K_\perp = 0$, and Λ diagonal with $\Lambda \succ 0$.

Theorem 2.3. *If there is a* $\delta > 0$ *such that*

$$\inf_{V \in \mathcal{V}} \min_s \lambda_s(K_\perp' V K_\perp) \geq \delta$$

then $f_S(\mathcal{V}) > -\infty$.

Proof. Define $A = K'VK$, $B = K'VK_\perp$, and $C = K_\perp' V K_\perp$. Then

$$f_S(V) = \log \mathbf{det} \begin{bmatrix} A & B \\ B' & C \end{bmatrix} + \mathbf{tr} \begin{bmatrix} A & B \\ B' & C \end{bmatrix}^{-1} \begin{bmatrix} \Lambda & 0 \\ 0 & 0 \end{bmatrix},$$

and, using the expression for inverses and determinants of partitioned matrices,

$$f_S(V) = \log \mathbf{det}(A - B'C^{-1}B) + \log \mathbf{det}(C) + \mathbf{tr}\ (A - B'C^{-1}B)^{-1}\Lambda.$$

Because $\Lambda \succ 0$ we have from Theorem 2.2

$$f_S(V) \geq \log \mathbf{det}(\Lambda) + \log \mathbf{det}(C) + r,$$

and thus, from the condition in the theorem,

$$f_S(V) \geq \log \mathbf{det}(\Lambda) + (n - r)\log \delta + r,$$

which implies $f_S(\mathcal{V}) > -\infty$. □

Theorem 2.4. *If the closure of* V *contains at least one singular matrix* V_0 *and there is a* $\delta > 0$ *such that*

$$\sup_{V \in \mathcal{V}} \max_s \lambda_s(K'V^{-1}K) \leq \delta$$

then $f_S(\mathcal{V}) = -\infty$.

Proof. For the second result we use

$$f_S(V) \leq \log \mathbf{det}(V) + \delta \mathbf{tr}\,(\Lambda),$$

and thus $V_k \rightarrow V_0$ implies $f_S(V_k) \rightarrow -\infty$. □

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