MATRIX-VARIATE NORMAL FIXED FACTOR ANALYSIS

JAN DE LEEUW

ABSTRACT. Meet the abstract. This is the abstract.

1. INTRODUCTION

Suppose $\underline{Y}$ is an $n \times m$ matrix with matrix-variate normal distribution. We suppose the means have *reduced-rank* structure, i.e. there exists and $n \times p$ matrix A and and $m \times p$ matrix B such that

$$\mathbf{E}(\underline{Y}) = AB', \tag{1a}$$

From the definition of the matrix-variate normal we suppose the dispersions have *direct product structure*, i.e. there exist positive definite matrices Θ and Ω of orders n and m such that

$$\mathbf{C}(\underline{y}_{ij}, \underline{y}_{k\ell}) = \theta_{ik}\omega_{j\ell}. \tag{1b}$$

People who are so inclined often write $\mathbf{C}(\mathbf{vec}(Y)) = \Theta \otimes \Omega$, but we shall refrain from using this notation.

The deviance, i.e. minus two times the log-likelihood, is

$$\mathcal{D}(A, B, \Theta, \Omega) = m \log |\Theta| + n \log |\Omega| + \mathbf{tr}\ \{\Theta^{-1}(Y - AB')\Omega^{-1}(Y - AB')'\}, \tag{2}$$

except for irrelevant constants.

Now set $\alpha_i = y_{i1}$ and $\beta_j = 1$. All other β_j are zero. Also set $\omega_1^2 = \epsilon^2$, while all other ω_j and all θ_i are equal to one. For this solution we have $\mathcal{D} = n \log \epsilon^2 + \sum_{i=1}^{n} \sum_{j=2}^{m} y_{ij}^2$, and thus $\mathcal{D} \rightarrow -\infty$ if $\epsilon^2 \rightarrow 0$. The deviance is unbounded below, and maximum likelihood estimates do not exist.

Date: August 30, 2010 — 12h 42min — Typeset in LUCIDA BRIGHT.
2000 *Mathematics Subject Classification.* 00A00.
Key words and phrases. Binomials, Normals, LaTeX.

If $(y_{ij} - \alpha_i\beta_j)^2 > 0$ for all i, j we have

$$\mathcal{D}_1(\alpha, \beta) = \min_{\theta,\omega} \mathcal{D}(\alpha, \beta, \theta, \omega) > -\infty \tag{3}$$

and the minimum is attained at θ and ω that are unique, up to a scale constant. This is the Sinkhorn theorem, familiar from iterative proportional fitting theory.

Also define

$$\mathcal{D}_0(\alpha, \beta, \sigma) = \sum_{i=1}^{n}\sum_{j=1}^{m} \log \sigma_{ij}^2 + \sum_{i=1}^{n}\sum_{j=1}^{m} \left(\frac{y_{ij} - \alpha_i\beta_j}{\sigma_{ij}}\right)^2. \tag{4}$$

If $(y_{ij} - \alpha_i\beta_j)^2 > 0$ for all i, j we have

$$\mathcal{D}_0(\alpha, \beta) = \min_{\sigma} \mathcal{D}(\alpha, \beta, \sigma) = \sum_{i=1}^{n}\sum_{j=1}^{m} \log(y_{ij} - \alpha_i\beta_j)^2 + nm \tag{5}$$

Clearly $\mathcal{D}_0(\alpha, \beta) \leq \mathcal{D}_1(\alpha, \beta)$, so

$$\mathcal{D}_{10}(\alpha, \beta) = \mathcal{D}_1(\alpha, \beta) - \mathcal{D}_0(\alpha, \beta) \geq 0. \tag{6}$$

Our proposal is to estimate α and β by minimizing $\mathcal{D}_{10}(\alpha, \beta)$, i.e. by minimizing the non-negative function

$$\Delta(\alpha, \beta, \theta, \omega) = m\sum_{i=1}^{n} \log \theta_i^2 + n\sum_{j=1}^{m} \log \omega_j^2 + \\ + \sum_{i=1}^{n}\sum_{j=1}^{m} \left\{\left(\frac{y_{ij} - \alpha_i\beta_j}{\theta_i\omega_j}\right)^2 - \log(y_{ij} - \alpha_i\beta_j)^2 - 1\right\}. \tag{7}$$

over $\alpha, \beta, \theta, \omega$. One interesting problem is to find out how well this estimate behaves compared with the OLS estimate, i.e. the first left and right singular vectors.

In the special case that $\Theta = I$, i.e. the rows are i.i.d., we find that

$$\mathcal{D}_1(\alpha, \beta) = \min_{\Omega} \mathcal{D}(\alpha, \beta, I, \Omega) = n\sum_{j=1}^{m} \log \frac{1}{n}\sum_{i=1}^{n} (y_{ij} - \alpha_i\beta_j)^2 + nm \tag{8}$$

and thus

$$\mathcal{D}_{10}(\alpha, \beta) = n\sum_{j=1}^{m} \log \frac{1}{n}\sum_{i=1}^{n} (y_{ij} - \alpha_i\beta_j)^2 - \sum_{i=1}^{n}\sum_{j=1}^{m} \log(y_{ij} - \alpha_i\beta_j)^2. \tag{9}$$

Compare this McDonald's maximum likelihood ratio method for fixed effect factor analysis. For this we go back to the more general

$$(10)\quad \mathcal{D}(\alpha,\beta,\Theta,\Omega)=m\log|\Theta|+n\log|\Omega|+\mathbf{tr}\,\Theta^{-1}(Y-\alpha\beta')\Omega^{-1}(Y-\alpha\beta')',$$

which becomes for $\Theta = I$

$$(11)\quad \mathcal{D}(\alpha,\beta,I,\Omega)=m+n\log|\Omega|+\mathbf{tr}\,(Y-\alpha\beta')\Omega^{-1}(Y-\alpha\beta')'.$$

Now

(12)

$$\mathcal{D}_M(\alpha,\beta)=\min_{\Omega}\mathcal{D}(\alpha,\beta,I,\Omega)=m+n\log|(Y-\alpha\beta')'(Y-\alpha\beta')|+nm,$$

and thus

$$(13)\quad \mathcal{D}_1(\alpha,\beta)-\mathcal{D}_M(\alpha,\beta)=\log|\Gamma(Y-\alpha\beta')|,$$

where $\Gamma(Y-\alpha\beta')$ is the correlation matrix of the residuals $Y-\alpha\beta'$.

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`