GIFI GOES LOGISTIC

JAN DE LEEUW

1. Data

2. Loss Function

Minimize

$$\mathcal{L}(\theta)=-\sum_{i=1}^{n}\sum_{j=1}^{m}\sum_{\ell=1}^{k_j} y_{ijl}\log\frac{\exp(\phi_{ij\ell}(\theta))}{\sum_{\nu=1}^{k_j}\exp(\phi_{ij\nu}(\theta))}$$

where $\phi_{ij\ell}$ can be any function with parameters θ.

Clearly $\mathcal{L}(\theta)\geq 0$ and $\mathcal{L}(\theta)=0$ if and only if $\pi_{ij\ell}(\theta)=y_{ij\ell}$ for all i,j,ℓ. Now suppose ϕ is homogeneous, in the sense that $\phi(\lambda\theta)=\lambda\phi(\theta)$ for all $\lambda\geq 0$. Then

$$\lim_{\lambda\rightarrow\infty}\pi_{ij\ell}(\lambda\theta)=\begin{cases}1 & \text{if } \pi_{ij\ell}(\theta)=\max_\nu \pi_{ij\nu}(\theta),\\ 0 & \text{otherwise.}\end{cases}$$

3. Majorization

$$\mathcal{L}(\theta)\leq\mathcal{L}(\tilde\theta)+\sum_{i=1}^{n}\sum_{j=1}^{m}\sum_{\ell=1}^{k_j} g_{ij\ell}(\lambda(\tilde\theta))(\lambda_{ij\ell}(\theta)-\tilde\lambda_{ij\ell}(\theta))+$$

$$+\frac{1}{2}\sup_{0\leq\xi\leq 1}\sum_{i=1}^{n}\sum_{j=1}^{m}\sum_{\ell=1}^{k_j}\sum_{i'=1}^{n}\sum_{j'=1}^{m}\sum_{\ell'=1}^{k_j} h_{ij\ell i'j'\ell'}(\xi\lambda+(1-\xi)\tilde\lambda)(\lambda_{ij\ell}-\tilde\lambda_{ij\ell})(\lambda_{i'j'\ell'}-\tilde\lambda_{i'j'\ell'})$$

$$g_{ij\ell}(\lambda)=-(y_{ij\ell}-\pi_{ijl}(\lambda))$$

Date: March 2, 2005.

2000 *Mathematics Subject Classification.* 62H25.

Key words and phrases. Multivariate Analysis, Correspondence Analysis.

$$h_{ij\ell i'j'\ell'}(\lambda) = \delta^{ii'}\delta^{jj'}\{\pi_{ij\ell}(\lambda)\delta^{\ell\ell'} - \pi_{ij\ell}(\lambda)\pi_{ij\ell'}(\lambda)\}$$

Define the matrices $H_{ij}(\lambda)$ with elements $\pi_{ij\ell}(\lambda)\delta^{\ell\ell'} - \pi_{ij\ell}(\lambda)\pi_{ij\ell'}(\lambda)$. Then

$$H_{ij}(\lambda) \leq \frac{1}{2}I$$

Proof. We show that the largest eigenvalue of $H_{ij}(\lambda)$ is less than or equal to $\frac{1}{2}$. Since the largest eigenvalue is smaller than any norm, it is smaller than

$$\|H_{ij}\|_1 = \max_{\nu=1}^{k_j} \sum_{\ell=1}^{k_j} |h_{ij}(\lambda)|_{\ell\nu} = 2\max_{\nu=1}^{k_j} \pi_{ij\nu}(1-\pi_{ij\nu}) \leq \frac{1}{2}.$$

□

Thus it follows that

$$\mathcal{L}(\lambda) \leq \mathcal{L}(\tilde{\lambda}) + \sum_{i=1}^{n}\sum_{j=1}^{m}\sum_{\ell=1}^{k_j} g_{ij\ell}(\tilde{\lambda})(\lambda_{ij\ell} - \tilde{\lambda}_{ij\ell}) + \frac{1}{4}\sum_{i=1}^{n}\sum_{j=1}^{m}\sum_{\ell=1}^{k_j}(\lambda_{ij\ell} - \tilde{\lambda}_{ij\ell})^2$$

Thus it suffices to minimize

$$\sum_{i=1}^{n}\sum_{j=1}^{m}\sum_{\ell=1}^{k_j}[\lambda_{ij\ell} - \tilde{z}_{ij\ell}]^2$$

with

$$\tilde{z}_{ij\ell} = \tilde{\lambda}_{ij\ell} + 2g_{ij\ell}(\tilde{\lambda}) = \tilde{\lambda}_{ij\ell} - 2(y_{ij\ell} - \pi_{ijl}(\tilde{\lambda})).$$

Thus we can solve the logistic optimization problem by using iterative least squares. If we know how to fit $\lambda_{ij\ell}$ to a matrix $z_{ij\ell}$ by least squares, then we can also fit it logistically.

4. Geometry

5. Inequalities

6. Rank orders

7. Special Cases

8. Random scores

In particular this is true for choice models of the form $\lambda_{ij\ell} = \beta' x_{i\ell}$.

And suppose $\lambda_{ij\ell} = -\frac{1}{2}\|a_i - b_{jl}\|^2$.

And suppose $\lambda_{ij\ell} = a_i' b_{jl} + c_{jl}$.

And suppose $\updownarrow_{jl}$ are parallel straight lines, and $\lambda_{ij\ell} = -\frac{1}{2} \min_{y \in \updownarrow_{jl}} \|x_i - y\|^2$.

And suppose $\updownarrow_{jl}$ are balls, and $\lambda_{ij\ell} = -\frac{1}{2} \min_{y \in \updownarrow_{jl}} \|x_i - y\|^2$.

Nonmetric Unfolding

$$-\sum_{i=1}^{n}\sum_{j=1}^{m} \log \frac{\exp(-\|x_i - y_{\pi_i(j)}\|)}{\sum_{\ell=j}^{m} \exp(-\|x_i - y_{\pi_i(\ell)}\|)}$$

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`