MATH 523: ITERATIVE PROPORTIONAL FITTING

A Poisson log-linear model is often used to model data in a multi-way contingency table. For a three-way table, defined by three factors A, B and C, it is assumed that $Y_{jkl} \sim Poisson(\mu_{jkl})$ where the saturated model can be written

$$\log \mu_{jkl} = \beta_0 + \beta_j^{(A)} + \beta_k^{(B)} + \beta_l^{(C)} + \beta_{jk}^{(AB)} + \beta_{jl}^{(AC)} + \beta_{kl}^{(BC)} + \beta_{jkl}^{(ABC)}$$

and all other log-linear models are nested inside the saturated model. If the factors take J, K and L levels, then there are a maximum of JKL parameters. The log likelihood is

$$l_n(\beta) = \sum_j \sum_k \sum_l (y_{jkl} \log \mu_{jkl} - \mu_{jkl}) + \text{const.}$$

Constraints for identifiability must be introduced: for example, sum-to-zero constraints include

$$\text{Main effects:} \qquad \sum_{j=1}^{J} \beta_j^{(A)} = 0$$

$$\text{Interactions:} \qquad \sum_{k=1}^{K} \beta_{kl}^{(AB)} = 0 \qquad \text{for each } l,$$

$$\sum_{k=1}^{K} \sum_{l=1}^{L} \beta_{jkl}^{(ABC)} = 0 \qquad \text{for each } j$$

etc. To find the ML estimates, note first that for **any** of the parameters β that appear in the form of μ_{jkl}, for the Poisson model, we have

$$\frac{\partial \mu_{jkl}}{\partial \beta} = \mu_{jkl}.$$

Therefore, solving the score equation for β_0 reduces to

$$\frac{\partial l_n(\beta)}{\partial \beta_0} = \sum_j \sum_k \sum_l (y_{jkl} - \mu_{jkl}) = 0 \tag{1}$$

But note that $\mu_{jkl} = \exp\{\beta_0 + (\beta_j^{(A)} + \cdots + \beta_{jkl}^{(ABC)})\} = \exp\{\beta_0\}\omega_{jkl}$, say, so equation (1)

$$\exp\{\widehat{\beta}_0\} = \frac{\sum_j \sum_k \sum_l y_{jkl}}{\sum_j \sum_k \sum_l \omega_{jkl}} = \frac{n}{\omega_{...}}$$

say, where n is the total count, whatever the ω values are. Therefore the log likelihood using this reparameterization can be written, up to an additive constant, as

$$l_n(\omega) \;=\; \sum_j \sum_k \sum_l \left(y_{jkl} \log \frac{n\omega_{jkl}}{\omega_{...}} \right) - n = \sum_j \sum_k \sum_l \left(y_{jkl} \log \frac{\omega_{jkl}}{\omega_{...}} \right) + n \log n - n$$

This is a complicated function of the ω parameters, so differentiating and solving to compute the ML estimates is not possible analytically; the exception to this is the case of the **saturated** model.

For a non-saturated model, the Poisson log likelihood is not amenable to direct maximization, but a "block-based" method can be used. Consider the model written in traditional notation as

$$A + B + C + A.B + A.C + B.C.$$

In this model, we have that

$$\log \mu_{jkl} = \beta_0 + \beta_j^{(A)} + \beta_k^{(B)} + \beta_l^{(C)} + \beta_{jk}^{(AB)} + \beta_{jl}^{(AC)} + \beta_{kl}^{(BC)}$$

and the identifiability constraints are

$$\sum_j \beta_j^{(A)} = 0 \qquad \sum_k \beta_{jk}^{(AB)} = 0 \qquad \sum_l \beta_{jl}^{(AC)} = 0$$

and similarly for the other factors.

Step 1: We write

$$\mu_{jkl} = \omega_{jk}\alpha_{jkl}$$

where

$$\omega_{jk} = \exp\{\beta_0 + \beta_j^{(A)} + \beta_k^{(B)} + \beta_{jk}^{(AB)}\} \qquad \alpha_{jkl} = \exp\{\beta_l^{(C)} + \beta_{jl}^{(AC)} + \beta_{kl}^{(BC)}\}$$

and consider the score equation for a single parameter, $\beta_{jk}^{(AB)}$. We have for each j and k

$$\frac{\partial l_n(\beta)}{\partial \beta_{jk}^{(AB)}} = \sum_l (y_{jkl} - \mu_{jkl}) = y_{jk.} - \mu_{jk.} \tag{2}$$

say.

Step 2: We presume that the α_{jkl} terms are **known**. We have $\mu_{jk.} = \omega_{jk}\alpha_{jk.}$, so therefore solving (2) equal to zero yields $\widehat{\omega}_{jk} = y_{jk.}/\alpha_{jk.}$. By similar calculations, we can also conclude that

$$\frac{\partial l_n(\beta)}{\partial \beta_0} = \sum_j \sum_k \left\{ \sum_l (y_{jkl} - \mu_{jkl}) \right\} = 0 \quad \frac{\partial l_n(\beta)}{\partial \beta_j^{(A)}} = \sum_k \left\{ \sum_l (y_{jkl} - \mu_{jkl}) \right\} = 0 \quad \frac{\partial l_n(\beta)}{\partial \beta_k^{(B)}} = \sum_j \left\{ \sum_l (y_{jkl} - \mu_{jkl}) \right\} = 0.$$

These remaining score equations can be solved analytically for known α_{jkl}:

$$\begin{aligned}
\widehat{\beta}_0 &= \frac{1}{JK} \sum_j \sum_k \log \widehat{\omega}_{jk} = \frac{1}{JK} \sum_j \sum_k \log \left\{ \frac{y_{jk.}}{\alpha_{jk.}} \right\} \\
\widehat{\beta}_j^{(A)} &= \frac{1}{K} \sum_k \log \widehat{\omega}_{jk} - \widehat{\beta}_0 \\
\widehat{\beta}_k^{(B)} &= \frac{1}{J} \sum_j \log \widehat{\omega}_{jk} - \widehat{\beta}_0 \\
\widehat{\beta}_{jk}^{(AB)} &= \log \widehat{\omega}_{jk} - \widehat{\beta}_j^{(A)} - \widehat{\beta}_k^{(B)} - \widehat{\beta}_0
\end{aligned}$$

These parameters are thus computed according to these formulae, and satisfy the identifiability constraints.

Step 3: To proceed, we consider now holding parameters

- $\beta_0, \beta_j^{(A)}, \beta_l^{(C)}, \beta_{jl}^{(AC)}$
- $\beta_0, \beta_k^{(B)}, \beta_l^{(C)}, \beta_{kl}^{(BC)}$

constant in turn, and then compute the parameter updates are recursively until convergence. Note that the parameter blocks here are not disjoint.

It can be shown that this algorithm leads to a non-decreasing sequence of log likelihood values. The algorithm is termed ***Iterative Proportional Fitting.***