COWLES COMMISSION DISCUSSION PAPER, STATISTICS, 317

Observations on the Computational Procedure for Maximum-Likelihood Estimates

by Herman Rubin
July 27, 1948

In his thesis, the author found that

$$(1)\quad -\frac{\partial \log L}{\partial A_{yx}} = S^{-1}AM - R^{-1}AW,$$

$$(2)\quad -\frac{\partial^2 \log L}{\partial (\mathrm{vec}_p A_{yx})^2} = S^{-1}\otimes(M - MA'S^{-1}AM) - R^{-1}\otimes(W - WA'R^{-1}AW)$$
$$- (MA'S^{-1})\otimes(S^{-1}AM) + (WA'R^{-1})\otimes(R^{-1}AW).$$

If A_{yx} strung out as a vector is $\bar{a}_q \Phi$, where $\bar{a}_q$ is the vector of essential parameters, we obtain

$$(3)\quad -\frac{\partial \log L}{\partial \bar{a}_q} = \mathrm{vec}_p\,(S^{-1}AM-R^{-1}AW)\,\Phi'$$

$$(4)\quad -\frac{\partial^2 \log L}{\partial \bar{a}_q^{\,2}} = \Phi\Big\{S^{-1}\otimes(M - MA'S^{-1}AM) - R^{-1}\otimes(W - WA'R^{-1}AW)$$
$$- (MA'S^{-1})\otimes(S^{-1}AM) + (WAR^{-1})\otimes(R^{-1}AW)\Big\}\Phi'$$
$$= L_{qq}$$

In this

$$(5)\quad S = AMA',$$

$$(6)\quad R = AWA'.$$

This reduces to the usual formulas for a complete system, as

$$(7)\quad R^{-1}A_{yx}W_{xx} = (A_{yx}W_{xx}A'_{yx})^{-1}A_{yx}W_{xx}$$
$$= A'^{-1}_{yy}W^{-1}_{yy}A^{-1}_{yy}A_{yy}W_{yy}I_{yx}$$
$$= A'^{-1}_{yy}I_{yx},$$

and

$$(8)\quad W_{xx} A'_{yx} R^{-1} A_{yx} W_{xx} = I'_{yx} W_{yy} A'_{yy} A'^{-1}_{yy} W^{-1}_{yy} A^{-1}_{yy} A_{yy} W_{yy} I_{yx}$$

$$= I'_{yx} W_{yy} I_{yx} = W_{xx}.$$

However, the ph method gives a positive "second derivative" matrix, which insures an increase for small h. The ph-method uses $S^{-1} \otimes M$ instead of (2) above.

Suppose, however, we use $S^{-1} \otimes (M - W)$. This is also positive. Furthermore, the roots of $\left| L_{qq} - \lambda \Phi \left(S^{-1} \otimes (M - W) \right) \Phi' \right| = 0$ are asymptotically 1, which means a p_1 method should converge rapidly for large samples.