COWLES COMMISSION DISCUSSION PAPERS: STATISTICS, NO. 302

ALTERNATIVE PROOF OF RUBIN'S RESULT (3.23) REGARDING THE LIKELIHOOD FUNCTION FOR AN INCOMPLETE LINEAR SYSTEM †

by Tjalling C. Koopmans

The alternative proof is based on replacing Rubin's restrictions

(1) $$A_{u_{II} x} = \begin{pmatrix} 0_{u_{II} y_I} & I_{u_{II} y_{II}} & -\Pi_{u_{II} z} \end{pmatrix}$$

on the "unknown" part of the system by the restrictions

(2) $$A_{u_{II} x} M_{xx} A'_{ux} = \begin{pmatrix} 0_{u_{II} u_I} & I_{u_{II} u_{II}} \end{pmatrix}.$$

Since any matrix

(3) $$A_{ux} = \begin{pmatrix} A_{u_I x} \\ \\ A_{u_{II} x} \end{pmatrix}$$

can be made to satisfy condition (2) by a transformation

(4) $$\begin{cases} A^*_{u^* x} = \Upsilon_{u^* u} \; A_{ux} , \qquad \Upsilon_{u^* u} = \begin{pmatrix} I_{u^*_I u_I} & 0_{u^*_I u_{II}} \\ \\ \Upsilon_{u^*_{II} u_I} & \Upsilon_{u^*_{II} u_{II}} \end{pmatrix} \\ \\ \Sigma^*_{u^* u^*} = \Upsilon_{u^* u} \; \Sigma_{uu} \; \Upsilon'_{u^* u} \end{cases}$$

which leaves $A_{u_I x}$ unaffected and preserves the form of the likelihood function*

(5) $$\log L = \log \left| A_{uy} \right| - \frac{1}{2} \log \left| \Sigma_{uu} \right| - \frac{1}{2} \operatorname{tr} \Sigma^{-1}_{uu} A_{ux} M_{xx} A'_{ux}$$

the addition of (2) to whatever a priori restrictions are imposed on $A_{u_I x}$ does not constitute a further restriction of the form of the probability distribution of the observations.

---

* omitting irrelevant additive constants

† This paper has been delayed through temporary loss of the stencil. It was prepared in the Autumn of 1947.

Following Rubin, we first maximize (5) with respect to A_{uu} and insert the result to obtain* (footnote on page 1)

(6) $$\log \hat{L} = \log \left| A_{uy} \right| - \frac{1}{2} \log \left| A_{ux} \; M_{xx} \; A'_{ux} \right| .$$

Next we maximize (5) with respect to $A_{u_{II}x}$ for a given value of $A_{u_I x}$.

Without the restrictions (2) on $A_{u_{II}x}$ the first order conditions for a maximum would be, if $I_{u_{II}u} \equiv (O_{u_{II}u_I} \quad I_{u_{II}u_{II}})$,

(7) $$\frac{\partial \log \hat{L}}{\partial A_{u_{II}x}} = I_{u_{II}u} \cdot \frac{\partial \log \hat{L}}{\partial A_{ux}} = I_{u_{II}u} \left[A'^{-1}_{uy} I_{yx} - (A_{ux} M_{xx} A'_{ux})^{-1} A_{ux} M_{xx} \right] = 0$$

However, since (2) does not further restrict the distribution function, any point where (6) is stationary with respect to variations of $A_{u_{II}x}$ subject to (2) must also be a point where (6) is stationary with respect to any variations of $A_{u_{II}x}$, and hence satisfy** (7). Using (2) in (7) we have

(8) $$I_{u_{II}u} \; A'^{-1}_{uy} \; I_{yx} - A_{u_{II}x} \; M_{xx} = 0 ,$$

or

(9) $$A_{u_{II}x} = I_{u_{II}u} \; A'^{-1}_{uy} \; I_{yx} \; M^{-1}_{xx} .$$

Without evaluating the maximizing value $A_{u_{II}x}$ (which occurs also in the right hand member of (9)) we eliminate it from (9) and (6) as follows. Writing

(10) $$W_{yy} \equiv (I_{yx} \; M^{-1}_{xx} \; I_{xy})^{-1} ,$$

** For a full justification of this statement see the last para. of this note.

we have

$$(11) \qquad A_{\underline{uy}} = \begin{pmatrix} A_{\underline{u}_I\underline{y}} \\ A_{\underline{u}_{II}\underline{y}} \end{pmatrix} = \begin{pmatrix} A_{\underline{u}_I\underline{y}} \\ I_{\underline{u}_{II}\underline{u}} \cdot A'^{-1}_{\underline{uy}} \cdot W^{-1}_{\underline{yy}} \end{pmatrix},$$

and hence

$$(12) \qquad A_{\underline{uy}} \; W_{\underline{yy}} \; A'_{\underline{uy}} = \begin{pmatrix} A_{\underline{u}_I\underline{y}} \cdot W_{\underline{yy}} \cdot A'_{\underline{uy}} \\ I_{\underline{u}_{II}\underline{u}} \end{pmatrix}.$$

From (6), (2) and (12), it follows that the likelihood function after maximization with respect to $A_{\underline{u}_{II}\underline{x}}$ can be written as

$$\log \hat{L} = \frac{1}{2} \log \left| A_{\underline{uy}} \; W_{\underline{yy}} \; A'_{\underline{uy}} \right| - \frac{1}{2} \log \left| W_{\underline{yy}} \right| - \frac{1}{2} \log \left| A_{\underline{ux}} \; M_{\underline{xx}} \; A'_{\underline{ux}} \right| =$$

$$= \frac{1}{2} \log \left| A_{\underline{u}_I\underline{y}} \; W_{\underline{yy}} \; A'_{\underline{u}_I\underline{y}} \right| - \frac{1}{2} \log \left| W_{\underline{yy}} \right| - \frac{1}{2} \log \left| A_{\underline{u}_I\underline{x}} \; M_{\underline{xx}} \; A'_{\underline{u}_I\underline{x}} \right|,$$

which is Rubin's equation (8.83).

This proof avoids the use of "perpendiculars" of the type $M_{\underline{y}_{II},\underline{y}_{II}\cdot\underline{u}_I\cdot\underline{z}}$ (except that $W_{\underline{y},\underline{y}} = M_{\underline{y},\underline{y}\cdot\underline{z}}$).

However, to make the statement following (7) above quite rigorous, it is necessary to specify $\Upsilon_{\underline{u}^*_{II}\underline{u}}$ in (4) as a continuous function of $A_{\underline{ux}}$ which assumes the value $I_{\underline{u}^*_{II}\underline{u}}$ whenever $A_{\underline{ux}}$ satisfies the restrictions (2). One possible choice is given by

$$A^{\dagger}_{\underline{u}_{II}\underline{x}} \equiv - A_{\underline{u}_{II}\underline{x}} \; M_{\underline{xx}} \; A'_{\underline{u}_I\underline{x}} \; (A_{\underline{u}_I\underline{x}} \; M_{\underline{xx}} \; A'_{\underline{u}_I\underline{x}})^{-1} A_{\underline{u}_I\underline{x}} + A_{\underline{u}_{II}\underline{x}}$$

$$A^{*}_{\underline{u}_{II}\underline{x}} \equiv (A^{\dagger}_{\underline{u}_{II}\underline{x}} \; M_{\underline{xx}} \; A'^{\dagger}_{\underline{u}_{II}\underline{x}})^{-\frac{1}{2}} A^{\dagger}_{\underline{u}_{II}\underline{x}},$$

taking the positive definite square root.