89.3.4. *Estimation and Testing in Linear Models with Singular Covariance Matrices* – Solution, proposed by Peter C.B. Phillips. Our model is:

$$y_t = Z_t\beta + u_t; \qquad (t = 1, \ldots, T) \tag{1}$$

with

$$R'u_t = 0 \quad \text{a.s.}$$

where

$$\Sigma R = 0.$$

For convenience, let the columns of R be orthonormal (if not, simply use $R(R'R)^{-1/2}$ instead). Construct the orthogonal matrix

$$C = \begin{bmatrix} S & R \\ n-r & r \end{bmatrix}$$

and use this transform (1) as follows:

$$C'y_t = C'Z_t\beta + C'u_t. \tag{2}$$

Here

$$C'u_t \equiv N\left(0, \begin{pmatrix} \Sigma_s & 0 \\ 0 & 0 \end{pmatrix}\right)$$

with $\Sigma_s = S'\Sigma S$. Now

$$C'\Sigma C + \begin{pmatrix} 0 \\ I \end{pmatrix}(0 \quad I) = \begin{pmatrix} \Sigma_s & 0 \\ 0 & I \end{pmatrix}$$

and so

$$\Sigma + RR' = C\begin{pmatrix} \Sigma_s & 0 \\ 0 & I \end{pmatrix}C'$$

and

$$(\Sigma + RR')^{-1} = C\begin{pmatrix} \Sigma_s & 0 \\ 0 & I \end{pmatrix}C' = \Sigma^{-}.$$

Take the first $n - r$ rows of (2) and write (in an obvious notation)

$$y_{st} = Z_{st}\beta + u_{st}.$$

Here $u_{st} \equiv N(0,\Sigma_s)$ and

$$pdf(u_{st}) = (2\pi)^{(n-r)/2}|\Sigma_s|^{-1/2}\exp\left\{-\frac{1}{2}\, u'_{st}\Sigma_s^{-1}u_{st}\right\}.$$

The likelihood is

$$pdf(Y_s) = (2\pi)^{(n-r)T/2}|\Sigma_s|^{-T/2}\exp\left\{-\frac{1}{2}\,\Sigma_1^T(y_{st} - Z_{st}\beta)'\Sigma_s^{-1}(y_{st} - Z_{st}\beta)\right\}. \tag{3}$$

Now consider

$$\begin{aligned} u'_{st}\Sigma_s^{-1}u_{st} &= (u'_{st} \quad 0)\begin{pmatrix} \Sigma_s^{-1} & 0 \\ 0 & I \end{pmatrix}\begin{pmatrix} u_{st} \\ 0 \end{pmatrix} \\ &= (u'_{st} \quad 0)C'C\begin{pmatrix} \Sigma_s^{-1} & 0 \\ 0 & I \end{pmatrix}C'C\begin{pmatrix} u_{st} \\ 0 \end{pmatrix} \\ &= (u'_{st} \quad 0)C'(\Sigma + RR')^{-1}C\begin{pmatrix} u_{st} \\ 0 \end{pmatrix} \\ &= u'_t(\Sigma + RR')^{-1}u_t \end{aligned}$$

since $u_t'C = (u_{st}', 0)$. Note also that

$$|\Sigma_s^{-1}| = \begin{vmatrix} \Sigma_s^{-1} & 0 \\ 0 & I \end{vmatrix} = |(\Sigma + RR')^{-1}|$$

and thus the likelihood (3) may be written in the form

$$pdf(Y_s) = (2\pi)^{(n-r)T/2} |\Sigma + RR'|^{-T/2} \times \exp\left\{ -\frac{1}{2} \Sigma_1^T (y_t - Z_t\beta)'(\Sigma + RR')^{-1}(y_t - Z_t\beta) \right\}. \quad \textbf{(4)}$$

To find the estimating equations for the MLE's $(\hat{\beta}, \hat{\Sigma})$ we write the log likelihood as

$$\mathcal{L} = \text{const} - \left(\frac{T}{2}\right) \ln |\Sigma + RR'| - \left(\frac{1}{2}\right) \Sigma_1^T (y_t - Z_t\beta)'(\Sigma + RR')^{-1}(y_t - Z_t\beta)$$

and note that we must maximize subject to $\Sigma R = 0$.

Set up the Lagrangean

$$\mathcal{L}^* = -T/2 \ln |\Sigma + RR'| - T/2 \operatorname{tr}[(\Sigma + RR')^{-1}M] + \operatorname{tr}(\Lambda \Sigma R)$$

with

$$M = T^{-1} \Sigma_1^T (y_t - Z_t\beta)(y_t - Z_t\beta)'.$$

Take differentials, giving

$$-T/2 \operatorname{tr}((\Sigma + RR')^{-1} d\Sigma) + T/2 \operatorname{tr}[(\Sigma + RR')^{-1} d\Sigma (\Sigma + RR')^{-1} M] + \operatorname{tr}(d\Sigma R \Lambda) = 0 \quad \textbf{(5)}$$

and, of course, $\Sigma R = 0$. Then (5) gives us

$$(\hat{\Sigma} + RR')^{-1} - (\hat{\Sigma} + RR')^{-1} M (\hat{\Sigma} + RR')^{-1} - \frac{2}{T} R\Lambda = 0 \quad \textbf{(6)}$$

so that

$$\frac{2}{T} \Lambda = (R'R)^{-1} R'(\hat{\Sigma} + RR')^{-1} - (R'R)^{-1} R'(\hat{\Sigma} + RR')^{-1} M (\hat{\Sigma} + RR')^{-1}.$$

But $R'(\hat{\Sigma} + RR') = (R'R)R'$ so that $R'(\hat{\Sigma} + RR')^{-1} = (R'R)^{-1}R'$. Thus

$$\frac{2}{T} \Lambda = (R'R)^{-2} R' - (R'R)^{-2} R' M (\hat{\Sigma} + RR')^{-1}$$

and (6) gives us

$$\hat{\Sigma} + RR' = M + (\hat{\Sigma} + RR')R(R'R)^{-2}R'(\hat{\Sigma} + RR') - (\hat{\Sigma} + RR')R(R'R)^{-2}R'M = M + RR' - R(R'R)^{-1}R'M.$$

But $R'M = 0$ so that we have

$$\hat{\Sigma} = M = T^{-1}\Sigma_1^T(y_t - Z_t\hat{\beta})(y_t - Z_t\hat{\beta})'. \tag{7}$$

Concentrating the likelihood function, we now have

$$\mathcal{L}^{**} = \text{const} - (T/2)\ln|\hat{\Sigma} + RR'| - T/2\,\text{tr}[(\hat{\Sigma} + RR')^{-1}\hat{\Sigma}] = \text{const} - (T/2)\ln|\hat{\Sigma} + RR'| \tag{8}$$

since

$$\text{tr}[(\hat{\Sigma} + RR')^{-1}\hat{\Sigma}] = \text{tr}[(\hat{\Sigma} + RR')^{-1}((\hat{\Sigma} + RR') - RR')] = \text{tr}[I - (\hat{\Sigma} + RR')^{-1}RR'] = \text{tr}[I - R(R'R)^{-1}R'].$$

Differentiating (8), we have

$$-T/2\,\text{tr}[(M + RR')^{-1}\,dM] = 0$$

so

$$\text{tr}[(M + RR')^{-1}(T^{-1}\Sigma_1^T Z_t\,d\beta(y_t - Z_t\beta)')] = 0$$

and thus

$$\Sigma_1^T d\beta' Z_t'(M + RR')^{-1}(y_t - Z_t\beta) = 0$$

giving

$$\hat{\beta} = [\Sigma_1^T Z_t'(M + RR')^{-1}Z_t]^{-1}[\Sigma_1^T Z_t'(M + RR')^{-1}y_t].$$

Hence, the MLE $(\hat{\beta}, \hat{\Sigma})$ is

$$\hat{\beta} = (\Sigma_1^T Z_t'\hat{\Sigma}^- Z_t)^{-1}(\Sigma_1^T Z_t'\hat{\Sigma}^- y_t)$$

$$\hat{\Sigma} = T^{-1}\Sigma_1^T(y_t - Z_t\hat{\beta})(y_t - Z_t\hat{\beta})'.$$

Remark 1. An alternative approach is to work directly from (3), giving

$$\hat{\Sigma}_s = T^{-1}\Sigma_1^T(y_{st} - Z_{st}\beta)(y_{st} - Z_{st}\beta)'$$

and

$$\hat{\beta} = (\Sigma_1^T Z_{st}'\hat{\Sigma}_s^{-1}Z_{st})^{-1}(\Sigma_1^T Z_{st}'\hat{\Sigma}_s^{-1}y_{st}).$$

Then

$$\hat{\Sigma} = C\begin{bmatrix}\hat{\Sigma}_s & 0\\ 0 & I\end{bmatrix}C' - RR' = S\hat{\Sigma}_s S' = T^{-1}\Sigma_1^T(y_t - Z_t\hat{\beta})(y_t - Z_t\hat{\beta})'$$

since

$$y_t = C\begin{pmatrix} y_{st} \\ 0 \end{pmatrix} = Sy_{st}, \qquad Z_t = C\begin{pmatrix} Z_{st} \\ 0 \end{pmatrix} = SZ_{st}$$

and, correspondingly,

$$\begin{aligned}\hat{\beta} &= (\Sigma_1^T Z_t' S\hat{\Sigma}_s^{-1} S' Z_t)^{-1}(\Sigma_1^T Z_t' S\hat{\Sigma}_s^{-1} S' y_t) \\ &= (\Sigma_1^T Z_t' \hat{\Sigma}^- Z_t)^{-1}(\Sigma_1^T Z_t' \hat{\Sigma}^- y_t)\end{aligned}$$

since

$$(\hat{\Sigma} + RR')^{-1} = C\begin{pmatrix} \hat{\Sigma}_s^{-1} & 0 \\ 0 & I \end{pmatrix} C' = S\hat{\Sigma}_s^{-1} S' = \hat{\Sigma}^-.$$

Obviously, this is the shorter method. But both are of interest.

(b) The OLS estimates are:

$$\beta^* = (\Sigma_1^T Z_t' Z_t)^{-1}(\Sigma_1^T Z_t' y_t)$$

$$\Sigma^* = T^{-1}\Sigma_1^T (y_t - Z_t\beta^*)(y_t - Z_t\beta^*)'.$$

Assume:

(i) The elements of Z_t are bounded uniformly in t;

(ii) $T^{-1}\Sigma_1^T Z_t' Z_t \to K > 0$, $T^{-1}\Sigma_1^T Z_t' \Sigma Z_t \to V > 0$.

Then

$$\begin{aligned}\sqrt{T}(\beta^* - \beta) &= (T^{-1}\Sigma_1^T Z_t' Z_t)^{-1}(T^{-1}\Sigma_1^T Z_t' y_t) \\ &\Rightarrow N(0, K^{-1}VK^{-1}).\end{aligned}$$

Now

$$\begin{aligned}\Sigma^* &= T^{-1}\Sigma_1^T u_t u_t' + T^{-1}\Sigma_1^T Z_t(\beta - \beta^*)u_t' \\ &\quad + T^{-1}\Sigma_1^T u_t(\beta - \beta^*)' Z_t' + T^{-1}\Sigma_1^T Z_t(\beta - \beta^*)(\beta - \beta^*)' Z_t' \\ &\underset{p}{\to} \Sigma.\end{aligned}$$

Since $\beta - \beta^* = o_p(1)$, we have

$$\begin{aligned}\sqrt{T}(\Sigma^* - \Sigma) &\sim \sqrt{T}(T^{-1}\Sigma_1^T u_t u_t' - \Sigma) \\ &= T^{-1/2}\Sigma_1^T (u_t u_t' - \Sigma) \\ &\Rightarrow N(0, 2P_D(\Sigma \otimes \Sigma))\end{aligned}$$

by the multivariate extension of the Lindeberg Lévy theorem.

(c) Note that the hypothesis

$$H_0 : \beta'\Sigma\beta = 0$$

is equivalent to

$H_0' : \beta = Ra$ for some a

since R spans $N(\Omega)$. We can also write H_0' in the form

$H_0'' : P_R\beta = \beta$ or $H_0''' : Q_R\beta = 0$

where $Q_R = I - P_R = I - R(R'R)^{-1}R'$.

To test H_0''' we use

$$W = \beta^{*\prime}Q_R[Q_R(\Sigma_1^T Z_t'Z_t)^{-1}(\Sigma_1^T Z_t'\Sigma^* Z_t)(\Sigma_1^T Z_t'Z_t)^{-1}Q_R]^- Q_R\beta^*.$$

Set

$$V_\beta^* = (\Sigma_1^T Z_t'Z_t)^{-1}(\Sigma_1^T Z_t'\Sigma^* Z_t)(\Sigma_1^T Z_t'Z_t)^{-1}$$

and then

$$(Q_R V_\beta^* Q_R)^- = (Q_R V_\beta^* Q_R + RR')^{-1}$$

and

$$W \Rightarrow \chi_{n-r}^2$$

where $Q_R V_\beta Q_R = Q_R K^{-1} V K^{-1} Q_R$ and $\operatorname{rank}(Q_R V_\beta Q_R) = n - r$. (We assume Q_R reduces the rank of V_β by r and no more. This will be so if V_β has full rank $k = n$.)

Alternatively, take the matrix S whose columns span $R(R)^\perp$, so that, as before, $C = [S \;\; R]$ is orthogonal. Then

$$R(R) = R(S)^\perp = N(S')$$

and H_0' is equivalent to

$$H_0^{iv} : S'\beta = 0.$$

(In effect, H_0''' is then H_0^{iv} with $Q_R = SS'$.) The Wald test is just

$$\begin{aligned} W &= (S'\beta^*)'(S'V_\beta^* S)^{-1}S'\beta^* \\ &= \beta^{*\prime}S(S'V_\beta^* S)^{-1}S'\beta^* \\ &\Rightarrow \chi_s^2 \end{aligned}$$

where $s = n - r = k - r$ here.

Remark 2. This shows, incidentally, that another g inverse of $Q_R V_\beta^* Q_R$ is

$$(Q_R V_\beta^* Q_R)^- = S(S'V_\beta^* S)^{-1}S'$$

as

$$S(S'V_\beta^* S)^{-1}S' = SS'S(S'V_\beta^* S)^{-1}S'SS' = Q_R S(S'V_\beta^* S)^{-1}S'Q_R$$

and

$$Q_R V_\beta^* Q_R S(S'V_\beta^* S)^{-1} S' Q_R V_\beta^* Q_R = SS'V_\beta^* S(SV_\beta^* S)^{-1} S' V_\beta^* SS'$$
$$= SS'V_\beta^* SS' = Q_R V_\beta^* Q_R.$$

Remark 3. We could attempt to test H_0 directly by using $\beta^{*\prime}\Sigma^*\beta^*$ or, equivalently, $\Sigma^*\beta^*$. Write

$$\Sigma^*\beta^* = \Sigma^*(\beta^* - \beta).$$

This holds because, under the null H_0, β satisfies $\Sigma\beta = 0$ and thus $\beta \in R(R)$, so that $\beta'\Sigma^* = 0$ (or, by construction, $R'\Sigma^* = 0$). Now as $T \to \infty$ we have

$$\sqrt{T}\Sigma^*\beta^* = \Sigma^*\sqrt{T}(\beta^* - \beta) \sim \Sigma\sqrt{T}(\beta^* - \beta) = \Sigma\sqrt{T}\beta^*$$

and then the Wald test is based on

$$\beta^{*\prime}\Sigma^*(\Sigma^* V_\beta^* \Sigma^*)^{-}\Sigma^*\beta \sim \beta^{*\prime}\Sigma(\Sigma V_\beta^* \Sigma)^{-}\Sigma\beta^*. \tag{9}$$

Now note that $\Sigma = S\Sigma_s S'$ and

$$S(S'V_\beta^* S)^{-1}S' = S\Sigma_s S' S \Sigma_s^{-1}(S'V_\beta^* S)^{-1}\Sigma_s^{-1} S' S \Sigma_s S'$$
$$= S\Sigma_s S'(S\Sigma_s S' V_\beta^* S \Sigma_s S')^{-} S\Sigma_s S'$$
$$= \Sigma(\Sigma V_\beta^* \Sigma)^{-}\Sigma \tag{10}$$

the middle line following because

$$[S\Sigma_s S' V_\beta^* S \Sigma_s S'][S\Sigma_s^{-1}(S'V_\beta^* S)^{-1}\Sigma_s^{-1}S'][S\Sigma_s S' V_\beta^* S\Sigma_s S']$$
$$= S\Sigma_s S' V_\beta^* S \Sigma_s S'$$
$$= \Sigma V_\beta^* \Sigma. \tag{11}$$

We deduce from (9) and (10) that

$$W = \beta^{*\prime}\Sigma^*(\Sigma^* V_\beta^* \Sigma^*)^{-}\Sigma^*\beta^*$$
$$= \beta^{*\prime} S(S'V_\beta^* S)^{-1} S'\beta^*$$

so that this test is again equivalent to the others.

Remark 4. The generalized inverse of (11) can also be written as follows:

$$(\Sigma V_\beta^* \Sigma)^{-} = S\Sigma_s^{-1}(S'V_\beta^* S)^{-1}\Sigma_s^{-1}S'$$
$$= S\Sigma_s^{-1}S'S(S'V_\beta^* S)^{-1}S'S\Sigma_s^{-1}S'$$
$$= \Sigma^{-}\Sigma(\Sigma V_\beta^* \Sigma)^{-}\Sigma\Sigma^{-}.$$

EDITOR'S COMMENT

The following solution has been proposed by H. Peter Boswijk. It does not correspond to the problem as originally stated, since part 2 and part 3 of Boswijk's solution involve the ML estimators, whereas the problem was about least-squares estimators. However, it was felt that the reader might also be interested in this elegant solution in the ML case.