93.4.4. *Reduced Rank Regression Asymptotics in Multivariate Regression* — Solution, proposed by Peter C.B. Phillips.

(a) The log-likelihood function is

$$L(\alpha,\beta,\Omega) = -(nT/2)\ln(2\pi) - (T/2)\ln|\Omega| - (\tfrac{1}{2})\mathrm{tr}[\Omega^{-1}(Y' - \Pi X')(Y - X\Pi')].$$

Concentrating out $\Omega$, the first-order conditions give

$$-(T/2)\mathrm{tr}(\Omega^{-1}d\Omega) + (T/2)\mathrm{tr}(\Omega^{-1}d\Omega\Omega^{-1}U'U/T) = 0,$$

and hence

$$\Omega = U'U/T, \text{ with } U = Y - X\Pi'. \quad (1)$$

The concentrated log-likelihood is (up to a constant)

$$L^a(\alpha,\beta) = -(T/2)\ln|(Y' - \alpha\beta'X')(Y - X\beta\alpha')/T|.$$

First-order conditions with respect to $\alpha$ give

$$(\tfrac{1}{2})\mathrm{tr}[\Omega^{-1}(Y' - \alpha\beta'X')X\beta d\alpha'] = 0,$$

and hence

$$\alpha = Y'X\beta(\beta'X'X\beta)^{-1}. \tag{2}$$

Concentrating out $\alpha$, we get the second-level concentrated log-likelihood (again, up to a constant)

$$L^b(\beta) = -(T/2)\ln|T^{-1}\{Y'Y - Y'X\beta(\beta'X'X\beta)^{-1}\beta'X'Y\}|. \tag{3}$$

To simplify (3) we write

$$\begin{aligned}
&|Y'Y - Y'X\beta(\beta'X'X\beta)^{-1}\beta'X'Y| \\
&\quad = |Y'Y||I - (Y'Y)^{-1/2}Y'\beta(\beta'X'X\beta)^{-1}\beta'X'Y(Y'Y)^{-1/2}| \\
&\quad = |Y'Y||I - DD'|, \text{ say for } D = (Y'Y)^{-1/2}Y'X\beta(\beta'X'X\beta)^{-1/2} \\
&\quad = |Y'Y||I - D'D| \\
&\quad = |Y'Y||I - (\beta'X'X\beta)^{-1/2}\beta'X'Y(Y'Y)^{-1}Y'X\beta(\beta'X'X\beta)^{-1/2}| \\
&\quad = |Y'Y||\beta'X'X\beta|^{-1}|\beta'X'X\beta - \beta'X'Y(Y'Y)^{-1}Y'X\beta|.
\end{aligned}$$

This argument holds for matrix $\beta$ as well as vector $\beta$. Because $Y'Y$ is not dependent on $\beta$, we have

$$\hat\beta = \operatorname{argmax} L^b(\beta) = \operatorname{argmin}|\beta'X'Q_YX\beta|/|\beta'X'X\beta|, \tag{4}$$

as required, where $Q_Y = I - Y(Y'Y)^{-1}Y'$.

Substituting (4) in (2), we obtain

$$\hat\alpha = Y'X\hat\beta(\hat\beta'X'X\hat\beta)^{-1},$$

and from (1) we have

$$\hat\Omega = T^{-1}(Y' - \hat\alpha\hat\beta'X')(Y - X\hat\beta\hat\alpha').$$

We standardize the moment matrices in (4) by $T^{-1}$, use the fact that $\beta'(T^{-1}X'X)\beta = \beta'\beta = 1$, and then

$$\hat\beta = \operatorname{argmin} \beta'(X'Q_YX/T)\beta.$$

The solution of this minimization problem is the eigenvector

$$(T^{-1}X'Q_YX - \lambda_T I)\hat\beta = 0 \tag{5}$$

corresponding to

$$\lambda_T = \lambda_{\min}(T^{-1}X'Q_YX).$$

Using the fact that

$$T^{-1}X'Q_YX = I - (T^{-1}X'Y)(T^{-1}Y'X)^{-1}(T^{-1}Y'X),$$

we can rewrite (5) as

$$(M_T - \mu_T I)\hat{\beta} = 0, \tag{6}$$

where

$$M_T = (T^{-1}X'Y)(T^{-1}Y'Y)^{-1}(T^{-1}Y'X), \tag{7}$$

and

$$\mu_T = 1 - \lambda_T = \lambda_{\max}(M_T). \tag{8}$$

(b) From (6) and (8) we have

$$\hat{\beta} = \operatorname{argmax} h'M_Th, \text{ subject to } h'h = 1.$$

Let $S_T(h) = h'M_Th$ and consider its limit as $T \to \infty$. We have

$$T^{-1}Y'X = \alpha\beta' + T^{-1}U'X \to_{\text{a.s.}} \alpha\beta'$$

$$T^{-1}Y'Y = T^{-1}U'U + \alpha\beta'(T^{-1}X'U) + (T^{-1}U'X)\beta\alpha' + \alpha\beta'(T^{-1}X'X)\beta\alpha' \to_{\text{a.s.}} \Omega + \alpha\alpha'.$$

Hence,

$$S_T(h) \to_{\text{a.s.}} h'\beta\alpha'(\Omega + \alpha\alpha')^{-1}\alpha\beta'h = S(h)$$

uniformly in $h$. Observe that the quadratic form $S(h)$ is maximized over $h'h = 1$ for $h = \beta$ where

$$S(\beta) + \alpha'(\Omega + \alpha\alpha')^{-1}\alpha.$$

Moreover, with the sign condition on $\beta$ (i.e., that its first nonzero element is positive), the solution vector $h = \beta$ to maximizing the form $S(h)$ is unique in the sphere $h'h = 1$. Because

$$S_T(\hat{\beta}) = \max_h S_T(h) \leq S_T(\beta)$$

and $S_T(\beta) \to_{\text{a.s.}} S(\beta) = \max_h S(h)$, it follows by the uniqueness of the solution to the latter problem that $\hat{\beta} \to_{\text{a.s.}} \beta$, as required.

(c) To find the limit distribution of $\hat{\beta}$ from (6), we show how $\hat{\beta}$ depends on the elements of $M_T$, at least to the first order. We proceed by taking differentials of (6) and the normalization

$$\hat{\beta}'\hat{\beta} = 1. \tag{9}$$

We have

$$dM_T\hat{\beta} + (M_T - \mu_T I)d\hat{\beta} - d\mu_T\hat{\beta} = 0, \tag{10}$$

$$\hat{\beta}'d\hat{\beta} = 0. \tag{11}$$

As our concern is with the limit distribution of $\hat{\beta}$, these differentials of (6) and (9) can be evaluated in a neighborhood of the probability limits of the component matrices. Note that

$$T^{-1}Y'X \to_p \alpha\beta', \qquad T^{-1}Y'Y \to_p \Omega + \alpha\beta'\beta\alpha' = \Omega + \alpha\alpha'$$

so that

$$M_T \to_p \beta\alpha'(\Omega + \alpha\alpha')^{-1}\alpha\beta' = M, \text{ say}$$

and

$$\mu_T \to_p \alpha'(\Omega + \alpha\alpha')^{-1}\alpha = \lambda_{\max}(M) = \mu, \text{ say,}$$

when evaluating (10) and (11) at these limits, we have

$$dM_T\beta + (M - \mu I)d\hat{\beta} - d\mu_T\beta = 0$$
$$\beta' d\hat{\beta} = 0.$$

By combining these equations and noting that $Md\hat{\beta} = 0$ and (because $\mu_T = \hat{\beta}'M_T\hat{\beta}$)

$$d\mu_T = \hat{\beta}'dM_T\hat{\beta} + 2\hat{\beta}M_T d\hat{\beta} = \hat{\beta}'dM_T\hat{\beta},$$

we have (again, evaluating at the probability limits)

$$\mu d\hat{\beta} = dM_T\beta - d\mu_T\beta = (I - \beta\beta')dM_T\beta. \tag{12}$$

This equation shows how $d\hat{\beta}$ is determined uniquely by the elements of $dM$. Observe that (12) automatically satisfies the condition $\beta' d\hat{\beta} = 0$ that arises from the normalizations $\beta'\beta = 1$, $\hat{\beta}'\hat{\beta} = 1$.

To use (12) to find the limit distribution of $\sqrt{T}(\hat{\beta} - \beta)$, we simply interpret the differential $d\hat{\beta}$ as $d\hat{\beta} = \sqrt{T}(\hat{\beta} - \beta)$ and $dM_T$ as $dM_T = \sqrt{T}(M_T - M)$. To find the limit distribution of $\hat{\beta}$, we must therefore first consider the moment matrix $M_T$. Now

$$M_T = M_{xy}M_{yy}^{-1}M_{yx},$$

where $M_{xy} = T^{-1}X'Y$ and $M_{yy} = T^{-1}Y'Y$. Thus,

$$dM_T = dM_{xy}M_{yy}^{-1}M_{yx} + M_{xy}M_{yy}^{-1}dM_{yx} - M_{xy}M_{yy}^{-1}dM_{yy}M_{yy}^{-1}M_{yx}, \tag{13}$$

where

$$dM_{xy} = \sqrt{T}[T^{-1}X'Y - \beta\alpha'] = T^{-1/2}X'U \to_d N(0, I \otimes \Omega) \tag{14}$$

and

$$\begin{aligned} dM_{yy} &= \sqrt{T}[T^{-1}Y'Y - (\Omega + \alpha\alpha')] \qquad (15) \\ &= \sqrt{T}[(T^{-1}U'U - \Omega) + \alpha\beta'(T^{-1}X'U) + (T^{-1}U'X)\beta\alpha']. \end{aligned}$$

When evaluating (13) at probability limits, we have

$$dM_T = dM_{xy}(\Omega + \alpha\alpha')^{-1}\alpha\beta' + \beta\alpha'(\Omega + \alpha\alpha')dM_{yx} - \beta\alpha'(\Omega + \alpha\alpha')^{-1}dM_{yy}(\Omega + \alpha\alpha')^{-1}\alpha\beta',$$

so that

$$(I - \beta\beta')dM_T\beta = (I - \beta\beta')dM_{xy}(\Omega + \alpha\alpha')^{-1}\alpha.$$

Hence, from (12) and (14) we deduce that

$$\begin{aligned}\mu d\hat{\beta} &= (I - \beta\beta')dM_{xy}(\Omega + \alpha\alpha)^{-1}\alpha \\ &\to_d N(0,(I - \beta\beta') \otimes \alpha'(\Omega + \alpha\alpha')^{-1}\Omega(\Omega + \alpha\alpha')^{-1}\alpha) \\ &= N(0,(I - \beta\beta')\alpha'(\Omega + \alpha\alpha')^{-1}\Omega(\Omega + \alpha\alpha')^{-1}\alpha).\end{aligned}$$

Because $\mu = \alpha'(\Omega + \alpha\alpha')^{-1}\alpha$, we have

$$\begin{aligned}d\hat{\beta} &= \sqrt{T}(\hat{\beta} - \beta) \\ &\to_d N(0,(I - \beta\beta')\alpha'(\Omega + \alpha\alpha')^{-1}\Omega(\Omega + \alpha\alpha')^{-1}\alpha/(\alpha'(\Omega + \alpha\alpha')^{-1}\alpha)^2).\end{aligned} \quad \textbf{(16)}$$

To simplify the variance matrix, we use the following formula:

$$(\Omega + \alpha\alpha')^{-1} = \Omega^{-1} - \Omega^{-1}\alpha\alpha'\Omega^{-1}/(1 + \alpha'\Omega^{-1}\alpha).$$

Then

$$\alpha'(\Omega + \alpha\alpha')^{-1} = \alpha'\Omega^{-1}/(1 + \alpha'\Omega^{-1}\alpha),$$

and consequently

$$\alpha'(\Omega + \alpha\alpha')^{-1}\Omega(\Omega + \alpha\alpha')^{-1}\alpha = \alpha'\Omega^{-1}\alpha/(1 + \alpha'\Omega^{-1}\alpha)^2$$

and

$$\mu = \alpha'(\Omega + \alpha\alpha')^{-1}\alpha = \alpha'\Omega^{-1}\alpha/(1 + \alpha'\Omega^{-1}\alpha).$$

It follows from these last formulae that (16) can be written more simply as

$$\sqrt{T}(\hat{\beta} - \beta) \to_d N(0,(I - \beta\beta')/\alpha'\Omega^{-1}\alpha). \quad \textbf{(17)}$$

Observe that the support of this singular normal distribution in the limit is $\mathbb{R}(I - \beta\beta') = \mathbb{R}(\beta)^{\perp}$.

Next consider

$$\hat{\alpha} = Y'X\hat{\beta}(\hat{\beta}'X'X\hat{\beta})^{-1} = \alpha + U'X\hat{\beta}(\hat{\beta}'X'X\hat{\beta})^{-1},$$

and then

$$d\hat{\alpha} = \sqrt{T}(\hat{\alpha} - \alpha) = T^{-1/2}U'X\hat{\beta} = dM_{yx}\beta \to_d N(0,\Omega)$$

in view of (14). The joint limit distribution is obtained from the expression

$$\begin{bmatrix} d\hat{\alpha} \\ d\hat{\beta} \end{bmatrix} = \begin{bmatrix} dM_{yx}\beta \\ (I - \beta\beta')dM_{xy}(\Omega + \alpha\alpha')^{-1}\alpha/\mu \end{bmatrix}$$

$$= \begin{bmatrix} dM_{yx}\beta \\ (I - \beta\beta')dM_{yx}\Omega^{-1}\alpha/(\alpha'\Omega^{-1}\alpha) \end{bmatrix}.$$

Observe that

$$dM_{yx}\beta = \operatorname{vec}(dM_{yx}\beta) = (I \otimes \beta')\operatorname{vec}(dM_{yx}) = (I \otimes \beta')K_{nm} \operatorname{vec}(dM_{yx}),$$

where $K_{nm}$ is the commutator matrix for which vec $A = K_{nm}$ vec $A'$, where $A$ is an $n \times m$ matrix. Hence, we may write

$$\begin{bmatrix} d\hat{\alpha} \\ d\hat{\beta} \end{bmatrix} = \begin{bmatrix} K_{nm}(\beta' \otimes I) \\ (I - \beta\beta') \otimes \alpha'\Omega^{-1}/\alpha'\Omega^{-1}\alpha \end{bmatrix} \operatorname{vec}(dM_{xy})$$

$$\to_d N\left(0, \begin{bmatrix} K_{nm}(I \otimes \Omega)K_{nm} & 0 \\ 0 & (I - \beta\beta')/\alpha'\Omega^{-1}\alpha \end{bmatrix}\right)$$

$$\equiv N\left(0, \begin{bmatrix} \Omega \otimes I & 0 \\ 0 & (I - \beta\beta')/\alpha'\Omega^{-1}\alpha \end{bmatrix}\right).$$

It follows that $\sqrt{T}(\hat{\alpha} - \alpha)$ and $\sqrt{T}(\hat{\beta} - \beta)$ are independent in the limit.

This solution *derives* the limit distribution by first principles. Note how it allows for the singularity of the limit distribution of $\hat{\beta}$. A direct approach using the information matrix is also possible after controlling for the singularity in the matrix limit.