Econometrics III: Solution to Exam 2020

Solution to problem 1

(a) Companion form:

$$\mathbf{Y}_t = \mathbf{c}^* + \mathbf{A}\mathbf{Y}_{t-1} + \mathbf{U}_t = \begin{pmatrix} y_{1t} \\ y_{2t} \\ y_{1,t-1} \\ y_{2,t-1} \end{pmatrix} \begin{pmatrix} 0.1 \\ 0 \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} \alpha_1 & \beta_1 & \alpha_2 & \beta_2 \\ \gamma_1 & \delta_1 & \gamma_2 & \delta_2 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} y_{1,t-1} \\ y_{2,t-1} \\ y_{1,t-2} \\ y_{2,t-2} \end{pmatrix} + \begin{pmatrix} u_{1t} \\ u_{2t} \\ 0 \\ 0 \end{pmatrix} \quad (1)$$

The process is stable if all eigenvalues of coefficient matrix $\mathbf{A}$ are strictly smaller than one in modulus. Alternatively, the process is stable if the roots of the reverse characteristic polynomial are strictly larger than one in modulus.

(b) The process is weakly stationary because it is stable (as shown in part (a)), it started in the infinite past (no influence of initial values), and $\mathbf{u}_t$ is white noise, i.e. it has finite, time-invariant second moments.
The process is also strictly stationary, since the errors are Gaussian white noise. In this case, the joint density only depends on the first and second moment, which are time-invariant. Therefore, the entire distribution is time-invariant.

(c) (i)

$$\mathbf{y}_t = \begin{pmatrix} 0.1 \\ 0 \end{pmatrix} + \begin{pmatrix} \alpha_1 & \beta_1 \\ 0 & \delta_1 \end{pmatrix} \mathbf{y}_{t-1} + \begin{pmatrix} \alpha_2 & \beta_2 \\ 0 & \delta_2 \end{pmatrix} \mathbf{y}_{t-2} + \mathbf{u}_t \quad (2)$$

(ii)

$$\begin{aligned}
\mu &= (\mathrm{I}_2 - \mathbf{A}_1 - \mathbf{A}_2)^{-1}\mathbf{c} \\
&= \left[\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} - \begin{pmatrix} \alpha_1 & \beta_1 \\ 0 & \delta_1 \end{pmatrix} - \begin{pmatrix} \alpha_2 & \beta_2 \\ 0 & \delta_2 \end{pmatrix} \right]^{-1} \begin{pmatrix} 0.1 \\ 0 \end{pmatrix} \\
&= \begin{pmatrix} 1 - \alpha_1 - \alpha_2 & -\beta_1 - \beta_2 \\ 0 & 1 - \delta_1 - \delta_2 \end{pmatrix}^{-1} \begin{pmatrix} 0.1 \\ 0 \end{pmatrix} \\
&= \frac{1}{(1 - \alpha_1 - \alpha_2)(1 - \delta_1 - \delta_2)} \begin{pmatrix} 1 - \delta_1 - \delta_2 & \beta_1 + \beta_2 \\ 0 & 1 - \alpha_1 - \alpha_2 \end{pmatrix} \begin{pmatrix} 0.1 \\ 0 \end{pmatrix} \\
&= \begin{pmatrix} \frac{1}{1-\alpha_1-\alpha_2} & \frac{\beta_1+\beta_2}{(1-\alpha_1-\alpha_2)(1-\delta_1-\delta_2)} \\ 0 & \frac{1}{1-\delta_1-\delta_2} \end{pmatrix} \begin{pmatrix} 0.1 \\ 0 \end{pmatrix} \\
&= \begin{pmatrix} \frac{0.1}{1-\alpha_1-\alpha_2} \\ 0 \end{pmatrix}
\end{aligned}$$

(d) Computation of autocovariances of $VAR(p)$ processes; Computation of forecast errors/forecast error variances; Impulse response analysis

(e) Information criteria are used to balance the tradeoff between estimation uncertainty (the more parameters, the higher) and goodness of fit (the more parameters, the better). They typically contain an estimate of the error covariance matrix (in logs) as goodness of fit measure and a penalty term that increases with the number of parameters in the model.

Solution to problem 2

(a) The condition is that matrix Π has rank 1.

(b) Cointegration means that there exists a linear combination of $I(1)$ variables that is $I(0)$. If this is the case, the information can be used for forecasting and structural analysis. Simply making the time series stationary by taking first differences would eliminate this information.

(c) (i)

$$\Delta\mathbf{Y} = \begin{pmatrix}\Pi & \Gamma\end{pmatrix}\begin{pmatrix}\mathbf{Y}_{-1} \\ \Delta\mathbf{X}\end{pmatrix} + \mathbf{U} \tag{3}$$

(ii) $\text{vec}(\mathbf{U}) \sim N(\mathbf{0}_{(2T\times 1)}, \mathrm{I}_{2T})$. (Generally, $\Sigma_{\text{vec}(\mathbf{U})} = \mathrm{I}_T \otimes \Sigma_{\mathbf{u}}$.)

(d) Procedure: Replace Γ in the likelihood by its least squares estimator in the model

$$\Delta\mathbf{Y} - \Pi\mathbf{Y}_{-1} = \Gamma\Delta\mathbf{X} + \mathbf{U}, \tag{4}$$

which equals

$$\widehat{\Gamma} = (\Delta\mathbf{Y} - \Pi\mathbf{Y}_{-1})\Delta\mathbf{X}'(\Delta\mathbf{X}\Delta\mathbf{X}')^{-1} \tag{5}$$

Re-arranging yields the concentrated likelihood function

$$\ell^c = -\frac{KT}{2}\ln(2\pi) - \frac{1}{2}\text{tr}\left[(\Delta\mathbf{Y}M_{\Delta\mathbf{X}} - \Pi\mathbf{Y}_{-1}M_{\Delta\mathbf{X}})'(\Delta\mathbf{Y}M_{\Delta\mathbf{X}} - \Pi\mathbf{Y}_{-1}M_{\Delta\mathbf{X}})\right] \tag{6}$$

with

$$M_{\Delta\mathbf{X}} := \mathrm{I}_T - \Delta\mathbf{X}'(\Delta\mathbf{X}\Delta\mathbf{X}')^{-1}\Delta\mathbf{X} \tag{7}$$

Solution to problem 3

(a) For projection matrices, the rank is equal to the trace. We have

$$\text{rank}(M_G) = \text{tr}(M_G) = \text{tr}(I_{NT}) - \text{tr}(G(G'G)^{-1}G') = NT - \text{tr}(\underbrace{G'G(G'G)^{-1}}_{=\mathrm{I}_N}) = NT - N \tag{8}$$

where we have used the circularity property of the trace ($\text{tr}(ABC) = \text{tr}(BCA) = \text{tr}(CAB)$).

(b) We get

$$(P_G - P_0)y = P_G y - P_0 y = \begin{pmatrix}\mathbb{1}_T \cdot \frac{1}{T}\sum_{t=1}^{T} y_{1t} \\ \vdots \\ \mathbb{1}_T \cdot \frac{1}{T}\sum_{t=1}^{T} y_{Nt}\end{pmatrix} - \begin{pmatrix}\frac{1}{NT}\sum_t\sum_i y_{it} \\ \vdots \\ \frac{1}{NT}\sum_t\sum_i y_{it}\end{pmatrix} \tag{9}$$

So via this transformation, the vector containing the overall means over time and individuals is subtracted from the vector of individual/group means.

(c) Potential advantage: More efficient estimation if individual effects are random. Needs additional assumption that μ and e are independently distributed.
Potential problems/disadvantages: If individual effects are not independent of e, GLS is inconsistent. Also, GLS contains unknown variance components that need to be estimated.