MID-TERM EXAM

(1) Consider the following systems, in all of which ε_t and ν_t are the innovations in y and x:

(i)

$$y_t = -.5y_{t-1} + 1.5x_{t-1} + \varepsilon_t \tag{A}$$

$$x_t = -1.5y_{t-1} + 2.5x_{t-1} + \nu_t\,. \tag{B}$$

(ii)

$$y_t = x_{t-1} + \varepsilon_t \tag{C}$$

$$x_t = y_{t-1} + \nu_t\,. \tag{D}$$

(iii)

$$y_t = -3.5y_{t-1} + 3.8x_{t-1} + \varepsilon_t \tag{E}$$

$$x_t = -4.2y_{t-1} + 4.5x_{t-1} + \nu_t\,. \tag{F}$$

Which, if any, of the systems are consistent with joint stationarity of y and x? For those, if any, that imply non-stationarity, which make y and x cointegrated? Which imply that impulse responses will converge to non-zero constants at distant horizons? Which imply persistent oscillations in the impulse responses?

These are all first-order AR systems, and the roots of the right-hand-side coefficient matrices are:

(i) (1.0,1.0)
(ii) (1.0, -1.0)
(iii) (.7,.3)

These can be found by solving the characteristic equations, which are quadratic. So e.g. for (i) we form $(-.5-\lambda)\cdot(2.5-\lambda)+1.5^2$, which reduces to $\lambda^2-2\lambda+1$. From these roots alone we can conclude that only (iii) is consistent with joint stationarity, since it is the only one with all roots having absolute value less than one. Since the two non-stationary systems have no roots less than one in absolute value, neither one implies any cointegration. System (iii), with stable roots, will imply that impulse responses converge to zero, not a non-zero constant, at long horizons. System (ii) has a negative root on the unit circle, so at least some of its impulse responses must oscillate indefinitely, not converge to a constant. System 1 has two unit roots. If they are not "repeated",

Date: November 6, 2008.

they imply that impulse responses converge to a non-zero constant. But if they are repeated, they imply a linear trend component in at least some impulse responses. They are indeed repeated. Proving this by hand directly is not easy. A clever quick argument, not expected under exam time pressure, notes that the Jordan decomposition of the matrix, if the roots are not repeated, will have the form $P\Lambda P^{-1}$, where Λ is diagonal with the two unit roots on the diagonal. But that means Λ is the identity, which in turn means that the system matrix has to be $PIP^{-1} = I$. But the actual system matrix for (i) above is not the identity. So the two unit roots repeat. In case you are interested (this was not expected on the exam), the first 4 powers of the system matrix for (i) are

$$\begin{bmatrix} -0.5 & 1.5 \\ -1.5 & 2.5 \end{bmatrix}, \quad \begin{bmatrix} -2 & 3 \\ -3 & 4 \end{bmatrix}, \quad \begin{bmatrix} -3.5 & 4.5 \\ -4.5 & 5.5 \end{bmatrix}, \quad \begin{bmatrix} -5 & 6 \\ -6 & 7 \end{bmatrix},$$

from which the linear trending behavior of the impulse responses is clear.

(2) Consider a standard linear regression model, with $y_t = X_t\beta + \varepsilon_t$, ε_t i.i.d. $N(0, \sigma^2)$, X_t strictly exogenous. The sample is $t = 1, \ldots, T$. We can treat β as the state in a Kalman filter (with a somewhat degenerate plant equation) and the regression equation as the observation equation. Display the plant equation that accomplishes this. Prove that as we progress throught the sample from $t = 1$ to $t = T$, the sequence of Kalman filtered estimates $\hat{\beta}_t$ will form a Martingale process — i.e. will satisfy $E_t\hat{\beta}_{t+s} = \hat{\beta}_t$ for all t and all $s \geq 0$. Will the same be true of the smoothed estimates? Why or why not? The plant equation is just $s_t = s_{t-1}$, where $s_t = \beta$ is the state. At each date the Kalman filter delivers as its filtered estimate of the state $\hat{\beta}_t = E[s_t \mid \{y_s, s \leq t\}] = E_t\beta$. Since this is a sequence of conditional expectations of the same random variable β, conditioning on an increasing sequence of information sets, the law of iterated expectations tells us immediately that it is a martingale process. This of course is true under the probability measure defined jointly by the plant and observation equations and the prior distribution for the state. Since β is unchanging, the smoothed estimates of the state are just a constant, equal to $\hat{\beta}_T$ at all dates in the sample. This is of course a trivial martingale.

(3) Consider the system

$$x_t = .2x_{t-1} + .7z_{t-1} + \varepsilon_t \tag{G}$$
$$y_t = -.1x_{t-1} + .9y_{t-1} - .3z_{t-1} + \nu_t \tag{H}$$
$$z_t = -x_{t-1} + .8z_{t-1} + \xi_t\,. \tag{I}$$

Is x Granger causally prior (GCP) to y in this system? Does it make any difference to your answer whether the variables are stationary or not? Does it make any difference to your answer whether $\varepsilon_t, \nu_t, \xi_t$ are the innovations in x, y, and z or not?

If we rearrange the system matrix by interchanging the first and second equations and interchanging the positions of y and x, the system matrix is converted to be

$$\begin{bmatrix} .9 & -.1 & -.3 \\ 0 & .2 & .7 \\ 0 & -1 & .8 \end{bmatrix}$$

This is block triangular, with x in the lower block and y in the upper block, so y is GCP to x. The definition of GCP in a VAR depends on the VAR (by definition) having one-step-ahead forecast errors (innovations) as its error terms, since Granger causality in its original form is about forecasting relationships. So long as the residuals are innovations, it doesn't matter whether the system is stationary or not, and the connection between GCP orderings and strict exogeneity holds. If the residuals are not one-step-ahead forecast errors, then the equations given do not define the properties of the (x_t, y_t, z_t) stochastic process, so the equations cannot be connected to GCP assertions.

(4) X and Y are two independent random variables. Each is distributed as $\text{Gamma}(1, \alpha)$, i.e. with p.d.f. $\alpha e^{-\alpha x}$ for X, for example. Suppose we have a prior pdf of the form $e^{-\alpha}$ on α (i.e. a $\text{Gamma}(1, 1)$) and we observe $U = X + Y$. What is the conditional distribution for X given our observation on U? (We condition only on U, not on the unknown parameter α, of course.) What if instead of $\text{Gamma}(1, \alpha)$ as the distribution for X and Y we had $\text{Gamma}(2, \alpha)$? (I.e., a p.d.f for X of the form $\alpha^2 x e^{-\alpha x}$)

The joint pdf of X, Y, α in the first version of the problem is $\alpha^2 e^{-\alpha(x+y+1)}\, dx\, dy\, d\alpha$. Transforming variables from (x, y, α) to u, x, α we have

$$\frac{\partial(u, x, \alpha)}{\partial(x, y, \alpha} = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}\,.$$

The absolute value of the determinant of this matrix is identically 1, so there is no Jacobian term when we transform the pdf to become

$$\alpha^2 e^{-\alpha u}\, du\, dx\, d\alpha$$

This does not depend on x at all, except through the fact that, because x and y are both positive with probability one, the range of possible X values once U is observed is just $(0, u)$. There is no need to integrate out the α component of the distribution explicitly, because the resulting pdf for $X \mid U$ is still uniform on $0, u$. If we change the distribution of X and Y to $\mathrm{Gamma}(2, \alpha)$, the joint distribution becomes $\alpha^4 xye^{-\alpha(x+y+1)}$, which transforms to $\alpha^4 x(u - x)e^{-\alpha u}$. Here again, with u fixed the pdf factors into a piece that depends on x, $x(u - x)$, and a piece that depends only on α and u. Therefore without doing any integration we know that $X \mid U$ is distributed as $\mathrm{Beta}(2, 2)$, scaled to the support $(0, u)$, i.e. with a pdf whose kernel is $x(u - x)$ on $(0, u)$.