Christiano
FINC 520, Spring 2009
Homework 3, due Thursday, April 23.

1. In class, we discussed the p^{th} order VAR:

$$y_t = c + \phi_1 y_{t-1} + \phi_2 y_{t-2} + ... + \phi_p y_{t-p} + \varepsilon_t,$$

where ε_t is a white noise with variance-covariance matrix, Ω, and orthogonal with y_{t-s}, $s > 0$. Find a definition of ξ_t, so that the VAR(p) can be written in the following VAR(1) form:

$$\xi_t = F\xi_{t-1} + v_t.$$

We assume that if λ satisfies

$$\left| \lambda^p I_n - \phi_1 \lambda^{p-1} - \phi_2 \lambda^{p-2} - ... - \phi_p \right| = 0, \tag{1}$$

then λ is less than unity in absolute value. In the above expression, $|A|$ denotes the determinant of the matrix, A.

2. Write the VAR(p) in transposed form:

$$y_t' = x_t' \Pi + \varepsilon_t',$$

where

$$\underset{(np+1)\times 1}{x_t} = \begin{pmatrix} 1 \\ y_{t-1} \\ \vdots \\ y_{t-p} \end{pmatrix}, \quad \underset{n\times(np+1)}{\Pi'} = \left[\begin{array}{cccc} c & \phi_1 & \cdots & \phi_p \end{array} \right].$$

Suppose we have data over the period $t = 1, ..., T$ and write:

$$Y = \begin{pmatrix} y_1' \\ y_2' \\ \vdots \\ y_T' \end{pmatrix}, \; X = \begin{pmatrix} x_1' \\ x_2' \\ \vdots \\ x_T' \end{pmatrix}, \; \varepsilon = \begin{pmatrix} \varepsilon_1' \\ \varepsilon_2' \\ \vdots \\ \varepsilon_T' \end{pmatrix},$$

so that the VAR can be written as follows:

$$\underset{T\times n}{Y} = \underset{T\times(np+1)}{X}\underset{(np+1)\times n}{\Pi} + \underset{T\times n}{\varepsilon}.$$

Premultiply this by X' and then take expectations:

$$E\left[X'Y\right] = E\left[X'X\right]\Pi + E\left[X'\varepsilon\right].$$

Note that

$$EX'\varepsilon = E\sum_{t=1}^{T} x_t\varepsilon_t' = 0,$$

because of the assumed orthogonality properties of ε_t. Then,

$$\Pi = \left\{E\left[X'X\right]\right\}^{-1} E\left[X'Y\right].$$

Ergodicity suggests using the following estimator for Π:

$$\hat{\Pi} = \left(X'X\right)^{-1} X'Y.$$

(a) Compute $\hat{\Pi}$ using the data in the MATLAB m file that has been provided, with $p = 4$. Let

$$y_t = \begin{pmatrix} R_t \\ \log \frac{GDP_t}{GDP_{t-1}} \\ \pi_t \end{pmatrix},$$

where R_t denotes the 3 month Tbill rate, π_t denotes the quarterly inflation rate and GDP_t denotes Gross Domestic Product in quarter t.

(b) Construct F, the matrix you constructed for question 1. Note how your way of handling the constant term causes there to be one eigenvalue equal to unity this matrix. Compute $\mu = Ey_t = \left[I - \phi_1 - \phi_2 - \phi_3 - \phi_4\right]^{-1} c$. Compare μ with:

$$P\left[y_{t+500}|y_t, y_{t-1}, ...\right],$$

Why should μ and this projection be so similar?

(c) Is the VAR(p) model you constructed covariance stationary and ergodic? Explain. (Hint: for this, you will want to know the roots of the polynomial in (1) and these can be found with Proposition 10.1 in the book.)

(d) Construct a 'predicted time series' on the long rate for the $T =$ 201 quarters, 1955Q4-2005Q4. Graph it together with the actual 5 year rate, R_t^l. To do this, you'll have to do the calculations discussed in class. However, the geometric sum formula described there will not work if you do it with the version of F in question 1, because that has a unit root in it. You can simply add $I + F + ... + F^{19}$. How does the expectations hypothesis do? Do the same graph, only this time do it with

$$y_t = \begin{pmatrix} R_t \\ R_t^l \\ \log \frac{GDP_t}{GDP_{t-1}} \\ \pi_t \end{pmatrix}.$$

Do the results make any difference with this adjustment?

3. In question 2, I expect you to find that the predicted long rate falls too rapidly in the early 1980s. This could be because (i) the linear prediction of the short rate based on the VAR does not correspond with people's actual expectations, or (ii) because the term structure hypothesis itself is not good. The possibility, (ii) is explored in the work of Piazzesi and Schneider (see, e.g., 'Equilibrium Yield Curves', NBER working paper 12609) and the papers they cite. However, this takes us too far afield of the subject matter of this course. Possibility (i) can be pursued in two ways. One could obtain survey data on forecasts of the short term interest rate and use those in place of $P[R_{t+j}|\Omega_t]$ in the formula for the long rate. This possibility would be fun to pursue, but would also take us too far off topic. Another way to attack (i) is to study the forecasts of the short rate that go into the calculation of the long rate in question (2). The idea is to see if they are 'any good'.

To investigate the quality of the forecast, you should graph the actual short rate data. At each two year interval you should graph the forecast of the short term rate that went into computing the predicted long term rate. Thus, suppose t is a quarter when you compute the forecast. On the graph of the actual short term rate, you should also graph $P[R_{t+j}|\xi_t]$, for $j = 0, .., 19$. Because the forecasted and actual interest rates appear in the same graph you can see if there is a systematic problem with the forecasts of the short rate. For example, is there a tendency to under-predict the short rate in the early 1980s? Your graph will look a little messy, because the forecasts will appear like long hairs sprouting at two year intervals and wiggling along for 5 years. The hairs associated with different forecast dates will become intertwined and be a little hard to distinguish. Hopefully, it will still be possible to see something.

4. The forecasting procedure not only provides formulas for $P[R_{t+j}|\xi_t]$, $j > 0$, but it also provides estimates of the uncertainty of the forecast. To see this, note that

$$R_{t+j} - P[R_{t+j}|\xi_t] = Dv_{t+j} + DFv_{t+j-1} + ... + DF^{j-1}v_{t+1},$$

where D is a row vector defined in class. Then the variance of the forecast error is:

$$v(j) \equiv var_t(R_{t+j} - P[R_{t+j}|\xi_t]) = DVD' + DFVF'D' + ... + DF^{j-1}V(F')^{j-1}D',$$

where V is the variance covariance matrix of v_t. This object can be constructed from the definition of v_t and an estimate of V, $\hat{V}$, where

$$\hat{V} = \frac{\hat{\varepsilon}'\hat{\varepsilon}}{T}.$$

Here, $\hat{\varepsilon} \equiv Y - X\hat{\Pi}$. Consider the three dates, $t_1 = 1980Q1$, $t_2 = 1982Q1$ and $t_3 = 1985Q1$. Produce three graphs, each containing the actual data as well as three other lines. Each of the three graphs corresponds to one of the dates, t_i, $i = 1, 2, 3$. Overlayed on the graph with the actual data, graph $P[R_{t_i+j}|\xi_t]$, for $j = 0, ..., 19$, as well as

$$P[R_{t_i+j}|\xi_t] \pm 2\sqrt{v(j)}, \; j = 0, ..., 19.$$

The graphs should be lined up so that $P[R_{t_i+j}|\xi_t]$ coincides with R_{t_i+j}, as in question 3. Do the uncertainty bands about the forecasts seem 'plausible' in light of the degree of fluctuations in the data? Do the uncertainty bands continue to expand, or does their width converge?

5. Consider the following time series representation for x_t :

$$x_t = \rho_x x_{t-1} + \varepsilon_t, \;\; \rho_x = 0.9, \;\; \sigma_\varepsilon = 0.01.$$

In class we have worked out that the variance of x_t is:

$$E\left(x_t\right)^2 = \frac{\sigma_x^2}{1-\rho_x^2},$$

and the τ^{th} lag covariance is

$$C\left(\tau\right) = E x_t x_{t-\tau} = \rho_x^{|\tau|} E x_t^2, \;\; \tau = 0, \pm 1, \pm 2, ... \;.$$

According to the discussion in class, the spectral density of x_t is:

$$S_x\left(e^{-i\omega}\right) = \frac{\sigma_x^2}{\left(1-\rho_x e^{-i\omega}\right)\left(1-\rho_x e^{i\omega}\right)}.$$

Applying the inverse Fourier transform to this yields the covariances:

$$C\left(\tau\right) = \frac{1}{2\pi}\int_{-\pi}^{\pi} S_x\left(e^{-i\omega}\right) e^{i\omega\tau} d\omega.$$

(a) Consider two parameterizations of the x_t process. One has the parameter values given above, and the other has $\rho_x = 0.0$ and σ_ε set so that the variance of the two x_t representations is the same. On the same figure, graph the spectral densities associated with the two representations of x_t for $\omega \in (0, \pi)$. Also, generate 100 artificial time series observations from each representation. Note how the one with the higher value of ρ_x generates a pattern of x_t 's that wanders around more lazily than does the representation with the smaller value of ρ_x. That 'lazy' pattern in the high-ρ_x representation shows up as relatively more power in the lower frequencies of its spectrum.

(b) Approximate the integral for $C\left(\tau\right)$ by the sum of the area of N rectangles, based on the Riemann sum interpretation of the integral. That is, if $\omega_j = 2\pi j/N$, with $j = -(N/2)+1, ..., N/2$,

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} S_x\left(e^{-i\omega}\right) e^{i\omega\tau} d\omega \approx \frac{1}{2\pi}\sum_{j=-(N/2)+1}^{N/2} S_x\left(e^{-i\omega_j}\right) e^{i\omega_j\tau}\left(\omega_j - \omega_{j-1}\right).$$

Is it correct to work with the lower value of $-(N/2)+1$ for j, or with the lower value of $-(N/2)$?

Evaluate $C(0)$, $C(2)$, $C(4)$ using the exact formulas above and the Riemann approximation for various values of N. What is the smallest values of N for which the Riemann approximation is good?