MORE ON TIME SERIES REGRESSION

1. Rational forecast regression

- It is indeed standard practice in models involving overlapping multi-step forecasts like the $y_{t+k} = \alpha + \beta F_t + \varepsilon_{t+k}$ model we discussed last time, to base inference on the approximation
$$\mathrm{Var}(\hat{\beta}_{OLS}) = (X'X)^{-1}X'\Omega X(X'X)^{-1}\,,$$
which is exact for the model with exogenous X and $E[\varepsilon\varepsilon'] = \Omega$.
- For the k-step forecast model, because the $X's$ are not exogenous, it is not true that $\mathrm{Var}(\varepsilon \mid X) = \Omega$, even though unconditionally $E[\varepsilon\varepsilon'] = \Omega$.
- Nonetheless under some regularity conditions, satisfied for example when the y and F vectors are jointly normal, the standard variance matrix formula becomes accurate asymptotically. Proving this is not terribly hard, but we don't have time for it.
- Note that this kind of regression equation is the building block for the expectational theory of the term structure of interest rates, among others.

2. Likelihood-based inference for the forecast regression model

- One can't do likelihood-based inference here without taking a stand on how the F's depend on the y's.
- Simple case: $k = 2$, all the dependence is taken care of with one lag:
$$\begin{aligned} y_t &= \alpha_{10} + \alpha_{11}y_{t-1} + \alpha_{12}F_{t-1} + \eta_{1t} \\ F_t &= \alpha 20 + \alpha_{21}y_{t-1} + \alpha_{22}F_{t-1} + \eta_{2t}\,, \end{aligned}$$
where we are assuming that $E_t\eta_{t+1} = 0$ (and that the unsubscripted η is the two η_j's stacked).
- Letting $X_t = [y_t\ F_t]'$, this is in the form $X_t = c + AX_{t-1} + \eta_t$. By substituting this equation into itself we arrive at
$$X_{t+2} = c + Ac + A^2X_t + \eta_t + A\eta_{t-1}\,,$$
from which we can read off $E_t[X_{t+2}]$
- If we want to test, then, that $E_t[y_{t+2}] = F_t$, we actually are testing that
$$\begin{aligned} \alpha_{11}^2 + \alpha_{12}\alpha_{21} &= 0 \\ \alpha_{11}\alpha_{12} + \alpha_{22}^2 &= 1 \\ \alpha_{11}\alpha_{10} + \alpha_{12}\alpha_{20} &= 0\,. \end{aligned}$$
So we should explore the likelihood to see how much probability is near the region where these constraints are satisfied.

3. Finding the unconditional distribution to use for IC's

- In a pure AR model with i.i.d. ε_t, assuming there is a stationary unconditional joint distribution for $\{y_t, \dots, y_{t-\ell+1}\}$, it is a matter of some slightly messy algebra to derive the covariance matrix of initial conditions from the ρ_s's and σ^2.
- The AR model is
$$y_t = \alpha + \sum_{s=1}^{\ell} \rho_s y_{t-s} + \varepsilon_t \,.$$
- Assuming Ey_t is constant across t, $Ey_t = \alpha/(1 - \sum \rho_s)$.
- Assuming $\mathrm{Cov}(y_t, y_{t-s}) = R_s = R_{-s}$, we can use the AR equation to derive
$$R_s = \sum_{v=1}^{\ell} \rho_s R_{|s-v-1|} + \delta_s \sigma^2 \;\; s = 0, \dots, \ell \,.$$
Using these equations for $s = 0, \dots, \ell - 1$, we have ℓ equations in the ℓ unknowns $R_0, \dots, R_s$, from which, under normality, we can construct the covariance matrix of $y_o, \dots, y_{-\ell}$ and hence, under normality, the joint distribution of the initial conditions.
- Note that this whole argument obviously depends on a) the sum of the ρ_s's not being 1 (for the mean calculation) and b) on the existence of a solution for the R_s's in which they can be used to populate a covariance matrix $[R_{|i-j|}]$ that turns out to be positive definite. In particular, we must have $R_0 > 0$ in the solution. This is not automatic. There are ρ_s vectors that imply that no stationary distribution for the y's exists.
- Using this approach to forming a distribution for initial conditions is only possible if nonstationary behavior for y is ruled out a priori. The conditional likelihood does not take a special form in such nonstationary cases, and they may represent economic behavior we don't want to rule out.