TAKEHOME EXAM

The exam is due by 9AM on Thursday, January 18, preferably in electronic form, by email to me (`sims.@princeton.edu`). Exam answers on paper are also acceptable, and should be handed in either directly to me or under the door of my office (209 JRRB). Unlike on problem sets, you are not to discuss this exam with any other students or collaborate with anyone in answering it, except of course after 9AM Thursday.

(1) In class we discussed the idea of using the unconditional joint distribution of the initial conditions in a VAR as part of the likelihood function, and also the idea of allowing for trend-like non-stationarity by using the unconditional joint distribution only of the components of the system corresponding to roots λ satisfying $|\lambda| < 1$ and $|1 - \lambda| > 1/T$, where T is sample size.

(a) Show that if we use the full joint unconditional distribution, assuming stationarity, the likelihood function is continuous and zero at any point in the parameter space where any root (including complex roots, of course) has absolute value 1.

(b) The idea of not imposing stationarity on components of the system corresponding to roots within $1/T$ of 1 is to allow simple linear trend-like non-stationarity, while ruling out unrealistically precise, nonlinear, forecasts from initial conditions. But why single out roots such that $|1 - \lambda_i| < 1/T$? Might it be better to avoid imposing stationarity on all components corresponding to roots close to any point on the unit circle, i.e. satisfying $1 - |\lambda_i| < 1/T$? And where does the $1/T$ come from? Could we just as well choose an arbitrary small number rather than $1/T$? Explain your answers.

(2) Here is a simple dynamic factor model:

$$y_t = c + Bf_t + u_t \tag{1}$$

$$u_t = \Omega u_{t-1} + \varepsilon_t \tag{2}$$

$$f_t = Af_{t-1} + \nu_t \tag{3}$$

$$B \text{ is } 4 \times 1\,, \quad \Omega \text{ is diagonal,} \tag{4}$$

$$\varepsilon_t \sim N(0, \Sigma) \text{ with } \Sigma \text{ diagonal, } \nu_t \sim N(0, \tau^2) \tag{5}$$

$$\nu_t, \varepsilon_t \text{ i.i.d. across time } t \text{ and independent of each other.} \tag{6}$$

(a) Explain how to use the Kalman filter to evaluate the likelihood function for this model. Include a discussion of how to handle initial conditions.

Date: January 17, 2018.

(b) One approach to estimating the posterior for this model is to use a random walk Metropolis MCMC iteration, working directly with prior and likelihood via the Kalman filter. Another approach is to use Gibbs sampling by including a step where a sequence of values for f_t is drawn, conditional on other parameters, by a backwards iteration similar to (but not the same as) Kalman smoothing. Explain both approaches in detail and discuss what might be advantages or disadvantages of each.

(3) In a structural VAR with constant coefficients $A(L)$ but time varying innovation covariance matrix Σ_t, we argued in class that identification through heteroskedasticty was possible. That argument assumed, though, that each period's distinct Σ_t could be estimated consistently, which of course implicitly assumes a limiting process in which each period becomes longer as sample size increases. Suppose instead time is divided into finite blocks (even blocks of fixed length, to keep the argument simpler), but we consider the number of blocks increasing with sample size, while block length remains fixed. Are there conditions under which we still could still consistently estimate $A(L)$ in such a model, even though the individual Σ_t matrices are not consistently estimated?