Prelim Examination

Friday August 11, 2017. Time limit: 150 minutes

Instructions:

(i) The total number of points is 80, the number of points for each problem is given below.

(ii) The exam is closed book and closed notes.

(iii) To receive full credit for your answers you have to explain your calculations. **You may state additional assumptions.**

Problem 1: Model Combination (18 Points)

An econometrician is estimating the time series regression model

$$y_t = x_t'\beta + u_t, \quad t = 1, \ldots, T, \tag{1}$$

where $x_t' = [1, t]$. The econometrician is concerned that u_t may be serially correlated and considers the following possibilities:

$$M_1 \;:\; u_t|(Y_{1:t-1}, X_{1:t-1}) \sim N(0,1) \tag{2}$$

$$M_2 \;:\; u_t = 0.95u_{t-1} + \epsilon_t, \quad \epsilon_t|(Y_{1:t-1}, X_{1:t-1}) \sim N\big(0, 1/(1-0.95^2)\big). \tag{3}$$

Note that $\big(0, 1/(1-0.95^2)\big) \approx 10$.

(i) (5 Points) Derive the (conditional) likelihood function for model M_2.

(ii) (2 Points) Derive the posterior distribution of $p(\beta|Y)$ for model M_2 under the improper prior $p(\beta) \propto 1$.

(iii) (6 Points) At what rates do the posterior variances of $\beta_1|Y$ and $\beta_2|Y$ shrink to zero as $T \longrightarrow \infty$?

(iv) (5 Points) The econometrician proposes using the estimator

$$\hat{\beta}_3 = \frac{1}{2}\big(\hat{\beta}_1 + \hat{\beta}_2\big),$$

where $\hat{\beta}_1$ and $\hat{\beta}_2$ are posterior mean estimators derived from specifications M_1 and M_2.

Discuss in detail the validity of the claim that under a quadratic loss function $\hat{\beta}_3$ is approximately the Bayes estimator for a prior that assigns prior probabilities of 0.5 to M_1 and M_2, respectively.

Problem 2: A Model with Time-Varying Intercept (37 Points)

Consider the following autoregressive model with time-varying intercept:

$$y_t = \phi_t + 0.95 y_{t-1} + u_t, \quad u_t \sim N(0,1), \quad t = 1, \ldots, T. \tag{4}$$

We will rewrite the problem slightly. Let $\phi^{(T)}$ be the $T \times 1$ vector of unknown intercepts and $Z^{(T)} = \mathcal{I}_T$ be the $T \times T$ identity matrix. Moreover, $Y^{(T)}$, $Y_{-1}^{(T)}$, and $U^{(T)}$ are the $T \times 1$ vectors with entries y_t, y_{t-1}, and u_t, respectively. Then we can express our model as

$$Y^{(T)} = Z^{(T)} \phi^{(T)} + 0.95 Y_{-1}^{(T)} + U^{(T)}. \tag{5}$$

If no ambiguity arises, we/you can drop the (T) superscripts.

(i) (2 Points) Derive the likelihood function for the model given by (4), utilizing the matrix notation in (5).

(ii) (4 Points) Derive the maximum likelihood estimator $\hat{\phi}_{mle}^{(T)}$. What is its frequentist sampling distribution? Can we use it to consistently estimate ϕ_t, $t = 1, \ldots, T$. What would happen if we replace the autocorrelation of 0.95 by an unknown parameter ρ that has to be estimated as well?

(iii) (4 Points) Now consider the following prior distribution, denoted by $p(\phi^{(T)}|\lambda)$, where λ is a hyperparameter. Define the $T \times T$ matrix R as

$$R = \begin{bmatrix} 1 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 & \cdots & 0 & 0 & 0 \\ 0 & -1 & 1 & 0 & \cdots & 0 & 0 & 0 \\ \vdots & & & & & & & \\ 0 & 0 & 0 & 0 & \cdots & 0 & -1 & 1 \end{bmatrix}$$

and let

$$R\phi^{(T)} \sim N\big(0_{T \times 1}, V(\lambda)\big), \quad \text{where} \quad V(\lambda) = \begin{bmatrix} 1 & 0_{1 \times (T-1)} \\ 0_{(T-1) \times 1} & \lambda I_{T-1} \end{bmatrix}.$$

What is the implied prior for $\phi^{(T)}$? What is the interpretation of this prior? What happens as $\lambda \longrightarrow 0$.

(iv) (6 Points) Derive the posterior distribution $p(\phi^{(T)}|Y, \lambda)$.

(v) (6 Points) For $T = 2$, compare the posterior mean $\tilde{\phi}(\lambda)$ to the maximum likelihood estimator. What are the limits of $\tilde{\phi}(\lambda)$ as $\lambda \longrightarrow 0$ and $\lambda \longrightarrow \infty$. Provide an interpretation of $\tilde{\phi}(0)$.

(vi) (7 Points) How is this model related to the time-varying coefficient models studied in class, its state-space representation, the Kalman filter, the Kalman smoother, and the Gibbs sampling approach to posterior inference in time-varying coefficient models?

(vii) (8 Points) Now calculate the log marginal data density

$$\ln p(Y|\lambda) = \ln \int p(Y|\phi^{(T)})p(\phi^{(T)}|\lambda)d\phi^{(T)}$$

as a function of λ. Decompose the log MDD into a goodness-of-fit term and a term that penalize dimensionality / model complexity. For $T = 2$ provide explicit formulas for these terms and show what happens to these terms as $\lambda \longrightarrow 0$ or $\lambda \longrightarrow \infty$.

Problem 3: VAR Analysis [25 Points]

Consider bivariate VAR of the form

$$y_t = \Phi_1 y_{t-1} + u_t = \Phi_1 y_{t-1} + \Phi_\epsilon \epsilon_t, \quad \epsilon_t \sim iid\mathcal{N}(0, \mathcal{I}). \tag{6}$$

where y_t is composed of log wages w_t and log hours h_t. The vector of structural shocks is composed of a productivity shock $\epsilon_{a,t}$ and a labor supply shock $\epsilon_{b,t}$.

(i) (1 Point) What condition does Φ_1 have to satisfy so that y_t is stationary?

(ii) (4 Points) Suppose y_t is stationary, derive the autocovariances of order zero and one, denoted by $\Gamma_{yy}(0)$ and $\Gamma_{yy}(1)$.

(iii) (2 Points) Derive the impulse response function of y_{t+h}, $h = 0, 1, \ldots$ with respect to the vector of structural shocks ϵ_t.

(iv) (4 Points) Describe the identification problem in the context of this VAR.

(v) (4 Points) Suppose we are willing to make the following assumptions:

- The inverse labor demand function $w_t = \varphi^D_{t-1}(h_t, \epsilon_{a,t})$ is contemporaneously affected by the technology shock, but not the labor supply shock.
- The inverse labor supply function $w_t = \varphi^S_{t-1}(h_t, \epsilon_{a,t}, \epsilon_{b,t})$ is affected both by the technology shock $\epsilon_{a,t}$ and the labor supply shock $\epsilon_{b,t}$.

Since labor supply and demand also depend on predetermined variables the functions φ are indexed by $t - 1$.

Shocks $\epsilon_{b,t}$ cause shifts along the labor demand schedule (a picture might help). Moreover, suppose that

$$\frac{\partial w_t}{\partial \epsilon_{b,t}} = (\alpha - 1)\frac{\partial h_t}{\partial \epsilon_{b,t}}. \tag{7}$$

Conditional on α, is it possible to uniquely identify the elements of Φ_ϵ? If yes, show how you can solve for Φ_ϵ based on α and the reduced-form VAR parameters.

(vi) (5 Points) Assume that our prior for α is $\mathcal{N}(0.66, 0.05^2)$ and our beliefs about α are independent of our beliefs about the reduced form VAR parameters. How you implement posterior inference with respect to the impulse response functions? Discuss in detail.

(vii) (5 Points) Alternatively, consider the following weaker sign restriction: a supply shock $\epsilon_{b,t}$ moves wages and hours in opposite directions upon impact, whereas a demand shock $\epsilon_{a,t}$ moves wages and hours in the same direction. What restrictions does this assumption impose on Φ_ϵ? Does it lead to point identification of the impulse responses?