Introduction to Bayesian Econometrics Course

Norges Bank

May, 2007

Overheads for Lecture on

Bayesian State Space Modeling

Gary Koop, University of Strathclyde

1 Introduction

- State space models are a popular way of modeling time series.
- Useful for trend/cycle/irregular/seasonal decompositions
- Also many models in macroeconomics/finance have state space structure (e.g. time varying VARs, stochastic volatility)
- Falls into familiar framework: Gibbs sampling/hierarchical priors

2 A General State Space Model

We will work with the state space model with measurement equation:

$$y_t = X_t \beta + Z_t \alpha_t + \varepsilon_t,$$

and state equation:

$$\alpha_{t+1} = T_t \alpha_t + u_t.$$

- y_t is scalar dependent variable and X_t and Z_t are $1 \times k$ and $1 \times p$ vectors, respectively, containing explanatory variables and/or known constants.
- ullet Extension to multivariate y is straightforward
- α_t is $p \times 1$ vector containing p states.

- ε_t is i.i.d. $N\left(0,h^{-1}\right)$
- u_t is $p \times 1$ vector which is i.i.d. $N\left(0, H^{-1}\right)$
- ullet ε_t and u_s are independent of one another for all s and t.
- T_t is a $p \times p$ matrix of known constants.
- Note: if T_t contains unknown parameters can be handled easily.
- Lots of interesting models structural time series models fall in this framework
- State equation can be interpreted as a hierarchical prior

2.1 Bayesian Computation in the State Space Model

 Conditional on knowing states, model is the Normal regression model:

$$y_t^* = X_t \beta + \varepsilon_t,$$

where $y_t^* = y_t - Z_t \alpha_t$.

- All our results for Normal linear regression model can be used (conditional on states).
- But, conditional on other parameters, can draw states.
 Thus, Gibbs sampler can be used. Formally, Gibbs sampler involves

$$p(\beta|y,\alpha_1,..,\alpha_T,h),p(h|y,\alpha_1,..,\alpha_T,\beta),$$

$$p(\alpha_1,..,\alpha_T|y,\beta,h,H)$$
 and $p(H|y,\alpha_1,..,\alpha_T)$.

2.2 A Prior

 We will derive Gibbs sampler using following prior (although anything possible):

$$p(\beta) = f_N(\beta|\underline{\beta},\underline{V}),$$

$$p(h) = f_G(h|\underline{s}^{-2},\underline{\nu}),$$

$$p(H) = f_W(H|\underline{\nu}_H, \underline{H}).$$

- Assume $\alpha_0 = 0$ (there are various ways of treating initial condition, but we will not discuss)
- State equation is a hierarchical prior:

 $p\left(\alpha_1,..,\alpha_T|H\right) = p\left(\alpha_1|H\right)p\left(\alpha_2|\alpha_1,H\right)..p\left(\alpha_T|\alpha_{T-1},H\right),$ where

$$p\left(\alpha_{t+1}|\alpha_t,H\right) = f_N\left(\alpha_{t+1}|T_t\alpha_t,H\right)$$

 $\quad \text{and} \quad$

$$p(\alpha_1|H) = f_N(\alpha_1|0,H).$$

2.3 Gibbs Sampler

- Gibbs sampler involves drawing from posterior conditionals
- Posterior conditionals for β and h based on usual regression results (with y replaced by y^* , see page 197)

$$\beta|y,h,\alpha_1,..,\alpha_T \sim N(\overline{\beta},\overline{V})$$
.

and:

$$h|y,\beta,\alpha_1,..,\alpha_T \sim G(\overline{s}^{-2},\overline{\nu})$$

- What about H? Conditional on $\alpha_1, ..., \alpha_T$, state equations are like SUR model (with no explanatory variables).
- Thus (using SUR results from page 140):

$$H|y, \alpha_1, ..., \alpha_T \sim W\left(\overline{\nu}_H, \overline{H}\right)$$

where $\overline{
u}_H$ and \overline{H} are given on page 197

- To complete Gibbs sampler, need $p(\alpha_1, ..., \alpha_T | y, \beta, h, H)$ and a means of drawing from it.
- This is multivariate Normal distribution, but hard draw from it since T—dimensional (elements can be highly correlated with one another).
- But we can draw on standard state space algorithms (e.g. Carter and Kohn, 1994, and DeJong and Shephard, 1995, Durbin and Koopman, 2002)
- Textbook described DeJong and Shephard's algorithm. I won't repeat details here.
- But note these papers are not explicitly Bayesian, but it turns out that there is an equivalence between standard non-Bayesian methods.

- ullet E.g. Kalman filter produces posterior means of $lpha_t$ that are required for our Gibbs sampler
- Since writing my textbook, it has come to my attention that the following is the most efficient algorithm for drawing states:
- Durbin, J. and Koopman, S., 2002, A simple and efficient simulation smoother for state space time series analysis, *Biometrika*.
- Thus, a Gibbs sampler can be set up which draws on results for Normal linear regression model, SUR model and a standard algorithm for drawing states in state space models.

2.4 Empirical Illustration: A TVP-AR Model

- A simple example taken from Koop and Potter (2001, Econometrics Journal).
- Economic history data set percentage change in UK industrial production from 1701-1992.
- AR(p) model with time varying parameters:

$$y_t = \alpha_{0t} + \alpha_{1t} y_{t-1} + ... + \alpha_{pt} y_{t-p} + \varepsilon_t,$$
 where for $i=0,..,p$

$$\alpha_{it+1} = \alpha_{i,t} + u_{it}.$$

- ε_t is i.i.d. $N\left(0,h^{-1}\right)$ and u_{it} is i.i.d. $N\left(0,\lambda_i h^{-1}\right)$ where ε_t , u_{is} and u_{jr} are independent of one another
- Special case of our state space model if we exclude X_t and define:

$$lpha_t = \left(egin{array}{c} lpha_{0t} \ lpha_{1t} \ \ \ \ \ \ lpha_{pt} \end{array}
ight),$$

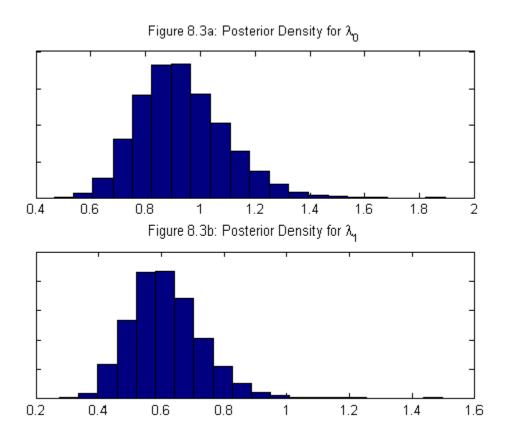
$$u_t = \left(egin{array}{c} u_{0t} \\ u_{1t} \\ \vdots \\ u_{pt} \end{array}
ight),$$

$$Z_t = \left(\begin{array}{ccccc} 1 & y_{t-1} & \dots & y_{t-p} \end{array}\right)$$

and set $T_t = I_{p+1}$ and

$$H^{-1} = h^{-1} \begin{bmatrix} \lambda_0 & 0 & 0 & . & 0 \\ 0 & \lambda_1 & . & . & . \\ . & 0 & . & . & . \\ . & . & . & . & 0 \\ 0 & . & . & 0 & \lambda_p \end{bmatrix}.$$

- Posterior results obtained by running Gibbs sampler.
- As an example of empirical results.



3 Extensions

- Note that blocking nature of Gibbs sampler means we can handle many extensions by just adding one more block (i.e. one more conditional posterior distribution) to the Gibbs sampler
- For instance, adding *Student-t* errors to the state space model can be done by combining Gibbs sampler Normal state space model with Gibbs sampler we talked about before for regression model with Student-t errors.
- Many other nonlinear/non-normal extensions of state space models
- Of particular importance: stochastic volatility model:

$$y_t = \exp\left(rac{lpha_t}{2}
ight) arepsilon_t,$$

where ε_t is i.i.d. N(0,1)

$$\alpha_{t+1} = \mu + \phi \alpha_t + u_t$$
 where u_t is i.id. $N\left(\mathbf{0}, \sigma_u^2\right)$.

• Rewrite (nonlinear) measurement equation as:

$$\log\left(y_t^2\right) = \alpha_t + \log\left(\varepsilon_t^2\right)$$

 \bullet The only thing which stops direct use of our previous Gibbs sampler is that $\log\left(\varepsilon_t^2\right)$ is not Normal

- ullet But as noted in Shephard (1993, Biometrika), Carter and Kohn (1997, JRSS,B) and Kim, Shephard and Chib (ReStud, 1998), $\log(\varepsilon_t^2)$ can be approximated extremely well by a mixture of Normal distributions.
- Hence, add block to Gibbs sampler drawing on results from mixture of Normals literature (remember Student-t errors is an example of mixture, but more general mixtures possible).