Bayesian Estimation of Panel Models

C. Berka/S. Humer/M. Moser

15 June 2011

Agenda

1. Bayesian Theory
2. Pooled Model
3. Individual Effects
4. Random Coefficients

What is the Bayesian Approach?

- How information in the data modifies a reseacher's beliefs about the parameters θ
- Additional information is formalised into $p(\theta)$
- Combine prior via Bayes' theorem with the data

	Frequentist	Bayesian
θ	true value	random
Output	$\hat{\theta}$	$p(\theta \mid y)$
Beliefs	*ad hoc*	$p(\theta)$
Probability	frequency	degree of belief

Table: Frequentist vs. Bayesian Approach

Tasks/Problems

- Choosing a prior
  - informative vs. ignorance priors
  - results from previous studies
- Finding the posterior
  - mathematically demanding
  - advanced computational methods
- Convincing others
  - sensitivity analysis

Bayesian Econometrics

- Basic Probability Theory:

$$p(A, B) = p(A|B)p(B) = p(B|A)p(A)$$

- Rearranging gives **Bayes Rule**:

$$p(B|A) = \frac{p(A|B)p(B)}{p(A)}$$

- In Econometrics we search parameters (θ) given the data (y):

$$\underbrace{p(\theta|y)}_{posterior} \propto \underbrace{p(y|\theta)}_{likelihood} \underbrace{p(\theta)}_{prior}$$

Posterior = Likelihood × Prior

$p(\theta|y)$ — **the posterior**

- random distribution of θ

→ take uncertainty into account

$p(y|\theta)$ — **the likelihood**

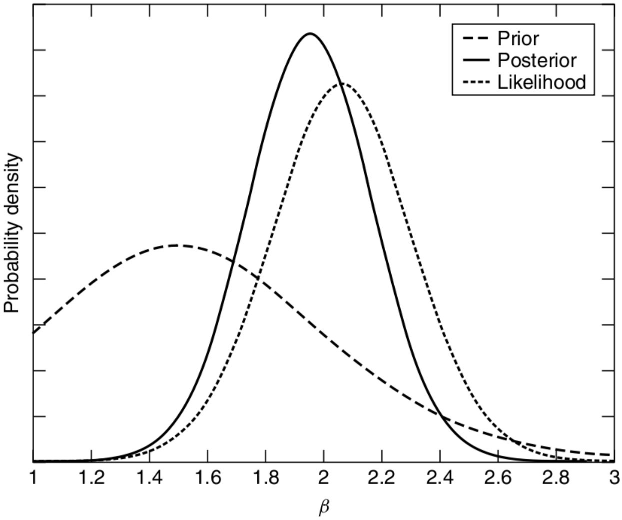
- best guess on θ out of y

$p(\theta)$ — **the prior**

- information before seeing the data
  - e.g. constant returns to scale in production function

Precision $\frac{1}{\sigma^2}$ of the prior and the likelihood are the relative weights which lead to the posterior distribution of θ.

Illustration of Prior, Likelihood & Posterior



The Pooled Model

Same linear relationship holds for every individual

$$y_i = X_i\beta + \epsilon_i$$

$i = 1, \dots, N$ β: k coefficients + intercept

Assumptions:

1. $\epsilon_i \sim N(0_T, h^{-1} I_T)$: multivariate Normal distribution ($\sigma^2 = \frac{1}{h}$)
2. ϵ_i and ϵ_j are independent for $i \neq j$
3. All elements of X_i are fixed or independent of ϵ_i with a probability function $p(X_i|\lambda)$, where $\beta, h \notin \lambda$

Pooled Model — Likelihood & Prior

Assumptions lead to a likelihood function of the form:

$$p(y|\beta,h)=\prod_{i=1}^{N}\frac{h^{\frac{T}{2}}}{2\pi^{\frac{T}{2}}}\left\{exp\left[-\frac{h}{2}(y_i-X_i\beta)'(y_i-X_i\beta)\right]\right\}$$

$$=\frac{1}{2\pi^{\frac{NT}{2}}}\left\{h^{\frac{k}{2}}exp\left[-\frac{h}{2}(\beta-\hat{\beta})'X_i'X_i(\beta-\hat{\beta})\right]\right\}\left\{h^{\frac{\nu}{2}}exp\left[-\frac{h\nu}{2s^{-2}}\right]\right\}$$

$\Rightarrow$ suggests a Normal–Gamma prior

$$\beta \sim N(\underline{\beta}, h^{-1}\underline{V}) \qquad h \sim G(\underline{s}^{-2}, \underline{\nu})$$

Pooled Model — Posterior

Analytical

Results in $\beta, h|y \sim NG(\overline{\beta}, \overline{V}, \overline{s}^{-2}, \overline{\nu})$

$$\begin{aligned}
\overline{V} &= (\underline{V}^{-1} + X_i' X_i)^{-1} \\
\overline{\beta} &= \overline{V}(\underline{V}^{-1}\underline{\beta} + X_i' X_i \hat{\beta}) \\
\overline{\nu} &= \underline{\nu} + NT
\end{aligned}$$

Numerical

- Rejection/ Importance Sampling
- Markov Chain Monte Carlo (MCMC)
  - Metropolis–Hastings algorithm
  - Gibbs sampling

The Individual Effects Model

Assuming that the intercept varies across inviduals, we can write

$$y_{it} = \alpha_i + \beta X_{it} + \epsilon_{it}$$

where α_i accounts for individual effects The likelihood function is based on the regression equation

$$y_i = \alpha_i \iota_T + \tilde{X}_i \tilde{\beta} + \epsilon_i$$

and is thus

$$p(y|\alpha, \tilde{\beta}, h) = \prod_{i=1}^{N} \frac{h^{\frac{T}{2}}}{2\pi^{\frac{T}{2}}} \left\{ exp\left[-\frac{h}{2}(y_i - \alpha_i - \tilde{X}_i\tilde{\beta})'(y_i - \alpha_i - \tilde{X}_i\tilde{\beta}) \right] \right\}$$

Priors in the Individual Effects Modell

Non-hierarchical Prior [FE]

- Prior Assumption for the coefficients and their variance

e.g. Normal-Gamma prior

$$\beta^* \sim N(\underline{\beta^*}, \underline{V})$$

$$h \sim G(\underline{s}^{-2}, \underline{\nu})$$

Priors in the Individual Effects Modell

Hierarchical Prior [RE]

- Impose a prior over the prior for the individual effects α_i

Say that for $i = 1, \ldots.., N$

$$\alpha_i \sim N(\mu_\alpha, V_\alpha)$$

we could treat μ_α and V_α as unknown paramaters which require their own prior

$$\mu_\alpha \sim N(\underline{\mu}_\alpha, \underline{\sigma}^2_\alpha)$$

$$V_\alpha^{-1} \sim G(\underline{V}_\alpha^{-1}, \underline{\nu}_\alpha)$$

The Random Coefficients Model

Intercept and slopes vary across inviduals

$$y_i = X_i \beta_i + \epsilon_i$$

$$i = 1, \dots, N \qquad \beta_i\text{: } k \text{ coefficients} + \text{intercept}$$

- $Nk + 1$ parameters to estimate: difficult unless $T \to \infty$

$\Rightarrow$ hierarchical prior for β_i: common distribution

Random Coefficients — Likelihood & Prior

Likelihood

$$p(y|\beta, h) = \prod_{i=1}^{N} \frac{h^{\frac{T}{2}}}{2\pi^{\frac{T}{2}}} \left\{ exp\left[-\frac{h}{2}(y_i - X_i\beta_i)'(y_i - X_i\beta_i) \right] \right\}$$

Hierarchical Prior

$\beta_i \ldots$ independent draws from a normal distribution

$$\beta_i \sim N(\mu_\beta, V_\beta)$$

$$\mu_\beta \sim N(\underline{\mu}_\beta, \underline{\Sigma}_\beta) \qquad V_\beta^{-1} \sim W(\underline{\nu}_\beta, \underline{V}_\beta^{-1})$$

$$h \sim G(\underline{s}^{-2}, \underline{\nu})$$

References

J. Albert.
Bayesian Computation with R.
Springer Science, 2009.

S. Chib.
Marginal likelihood from the gibbs sampler.
Journal of the American Statistical Association, 90(432):1313–1321, 1995.

P. Kennedy.
A Guide to Econometrics.
The MIT Press, 2003.

G. Koop.
Bayesian Econometrics.
John Wiley & Sons, 2003.