

Bayesian Methods for Macroeconometrics

Bayesian Analysis of a Linear Regression Model

Frank Schorfheide

Department of Economics

University of Pennsylvania

Introduction to Bayesian Statistics

- Ingredients of Bayesian Analysis:
 - Likelihood function $\mathcal{L}(\theta|Y^T) = p(Y^T|\theta)$
 - Prior density $p(\theta)$
 - Marginal data density $p(Y^T) = \int p(Y^T|\theta)p(\theta)d\theta$

- Bayes Theorem:

$$p(\theta|Y^T) = \frac{\mathcal{L}(\theta|Y^T)p(\theta)}{p(Y^T)} \quad (1)$$

Linear Regression

- Consider linear regression model:

$$y_t = x_t' \theta + u_t, \quad u_t \sim iid \mathcal{N}(0, 1), \quad (2)$$

or

$$Y = X\theta + U.$$

Assume θ is $k \times 1$.

- Notice: we treat the variance of the errors as known. The generalization to unknown variance is straightforward but tedious.
- Likelihood function:

$$\mathcal{L}(\theta|Y, X) = (2\pi)^{-T/2} \exp \left\{ -\frac{1}{2} (Y - X\theta)' (Y - X\theta) \right\}. \quad (3)$$

A Convenient Prior

- Prior:

$$\theta \sim \mathcal{N}\left(0_{k \times 1}, \tau^2 \mathcal{I}_{k \times k}\right), \quad p(\theta) = (2\pi\tau^2)^{-k/2} \exp\left\{-\frac{1}{2\tau^2}\theta'\theta\right\} \quad (4)$$

- Large τ means diffuse prior.
- Small τ means tight prior.

Deriving the Posterior (I)

- Bayes Theorem:

$$p(\theta|Y, X) \propto p(\theta)\mathcal{L}(\theta|Y, X). \quad (5)$$

- Right-hand-side is given by

$$\begin{aligned} & p(\theta)\mathcal{L}(\theta|Y, X) \\ & \propto (2\pi)^{-\frac{T+k}{2}} \tau^{-k} \exp \left\{ -\frac{1}{2} [Y'Y - \theta'X'Y - Y'X\theta + \theta'X'X\theta + \tau^{-2}\theta'\theta] \right\}. \quad (6) \end{aligned}$$

- Rewrite exponential term

$$\begin{aligned} & Y'Y - \theta'X'Y - Y'X\theta + \theta'X'X\theta + \tau^{-2}\theta'\theta \\ & = Y'Y - \theta'X'Y - Y'X\theta + \theta'(X'X + \tau^{-2}\mathcal{I})\theta \quad (7) \\ & = \left(\theta - (X'X + \tau^{-2}\mathcal{I})^{-1}X'Y \right)' \left(X'X + \tau^{-2}\mathcal{I} \right) \left(\theta - (X'X + \tau^{-2}\mathcal{I})^{-1}X'Y \right) \\ & \quad + Y'Y - Y'X(X'X + \tau^{-2}\mathcal{I})^{-1}X'Y. \end{aligned}$$

Deriving the Posterior (II)

- Exponential term is a quadratic function of θ .
- Deduce: posterior distribution of θ must be a multivariate normal distribution

$$\theta|Y, X \sim \mathcal{N}(\tilde{\theta}_T, \tilde{V}_T) \quad (8)$$

with

$$\tilde{\theta}_T = (X'X + \tau^{-2}\mathcal{I})^{-1}X'Y$$

$$\tilde{V}_T = (X'X + \tau^{-2}\mathcal{I})^{-1}.$$

- $\tau \longrightarrow \infty$:

$$\theta|Y, X \overset{approx}{\sim} \mathcal{N}\left(\hat{\theta}_{mle}, (X'X)^{-1}\right). \quad (9)$$

- $\tau \longrightarrow 0$:

$$\theta|Y, X \overset{approx}{\sim} \text{Pointmass at } 0 \quad (10)$$

Point Estimation (I)

- Interpret point estimation as decision problem.
- Consider quadratic loss:

$$L(\theta, \delta) = (\theta - \delta)^2 \quad (11)$$

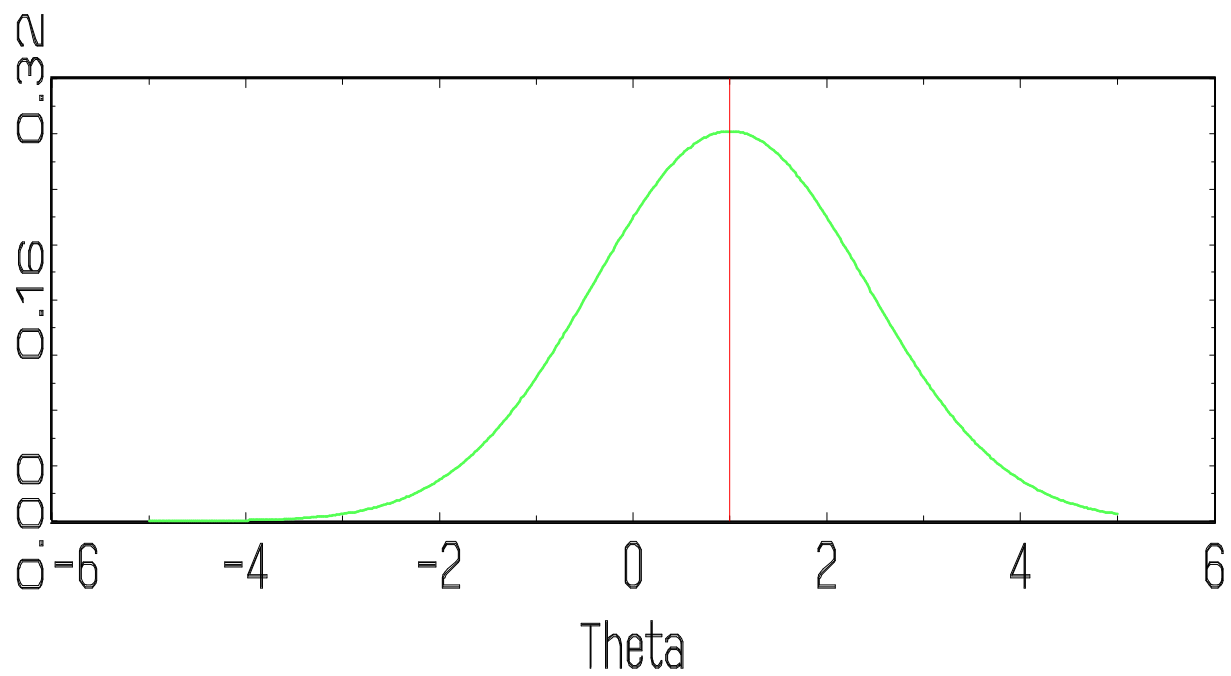
- Optimal decision rule is obtained by minimizing

$$\min_{\delta \in \mathcal{D}} \mathbf{E}[(\theta - \delta)^2 | Y, X] \quad (12)$$

- Solution: $\delta = \mathbf{E}[\theta | Y, X]$, i.e., posterior mean.

(Figure)

Posterior Density



Point Estimation (II)

- Consistency: Suppose data are generated from the model $y_t = x_t'\theta_0 + u_t$. Asymptotically the Bayes estimator converges to the “true” parameter θ_0 .
- Consider

$$\begin{aligned}\tilde{\theta}_T &= (X'X + \tau^{-2}\mathcal{I})^{-1}X'Y && (13) \\ &= \theta_0 + \left(\frac{1}{T}X'X + \frac{1}{\tau^2 T}\mathcal{I}\right)^{-1} \left(\frac{1}{T}X'U\right) \\ &\xrightarrow{p} \theta_0\end{aligned}$$

- Disagreement between two Bayesians who have different priors will asymptotically vanish.

Confidence Sets

- Bayesian: $C_Y \subseteq \Theta$ is $1 - \alpha$ credible if

$$IP_Y\{\underbrace{\theta}_{r.v.} \in C_Y\} \geq 1 - \alpha \quad (14)$$

- Bayesian: a highest posterior density region (HPD) is of the form

$$C_Y = \{\theta : p(\theta|Y) \geq k_\alpha\} \quad (15)$$

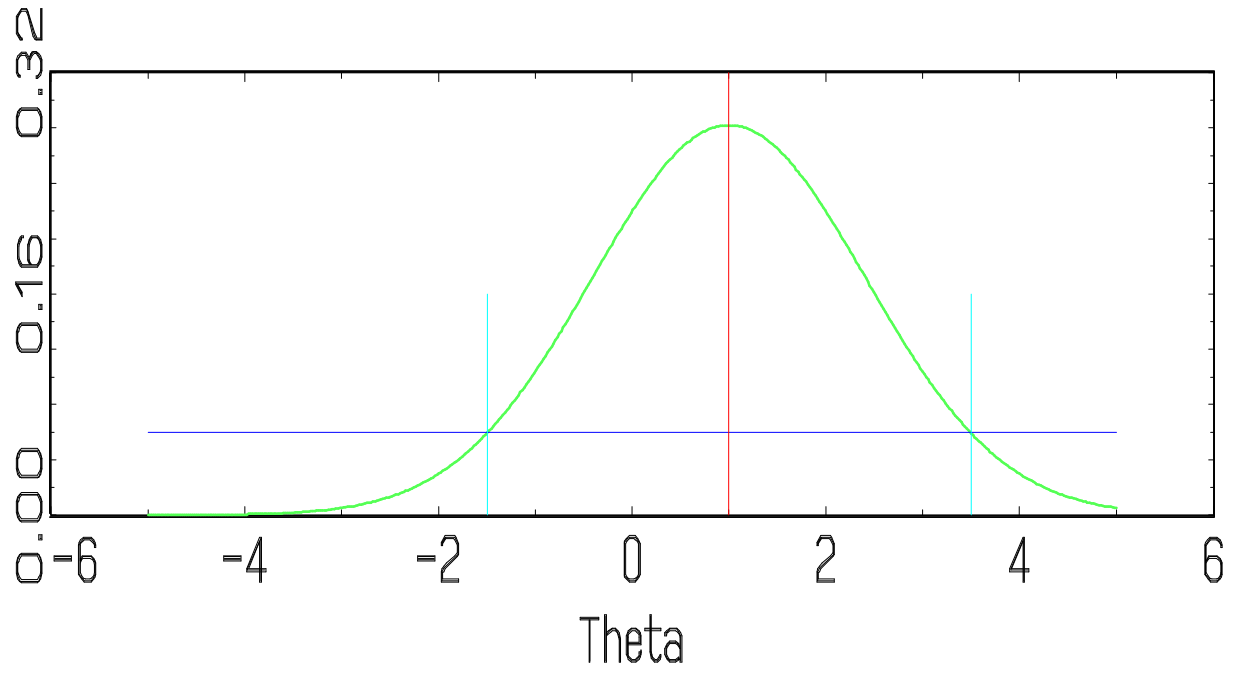
where k_α is the largest bound such that

$$IP_Y\{\theta \in C_Y\} \geq 1 - \alpha$$

HPD regions have the smallest volume among all α credible regions of the parameter space Θ .

(Figure)

Posterior Density



Testing (I)

- $H_0 : \theta \in \Theta_0$ versus $H_1 : \theta \in \Theta_1$.
- Decision space is 0 (“reject”) and 1 (“accept”).
- Loss function

$$L(\theta, \delta) = \begin{cases} 0 & \delta = \{ \theta \in \Theta_0 \} & \text{correct decision} \\ a_0 & \delta = 0, \theta \in \Theta_0 & \text{Type 1 error} \\ a_1 & \delta = 1, \theta \in \Theta_1 & \text{Type 2 error} \end{cases} \quad (16)$$

Note that the parameters a_1 and a_2 are part of the econometricians preferences.

- Optimal decision:

$$\delta(Y^T) = \begin{cases} 1 & IP_Y\{\theta \in \Theta_0\} \geq \frac{a_1}{a_0+a_1} \\ 0 & \text{otherwise} \end{cases} \quad (17)$$

Testing (II)

- Posterior odds:

$$\frac{\mathbb{P}_Y\{\theta \in \Theta_0\}}{\mathbb{P}_Y\{\theta \in \Theta_1\}}$$

- Often, hypotheses are evaluated according to Bayes factors:

$$B(Y^T) = \frac{\text{Posterior Odds}}{\text{Prior Odds}} = \frac{\mathbb{P}_Y\{\theta \in \Theta_0\}/\mathbb{P}_Y\{\theta \in \Theta_1\}}{\mathbb{P}\{\theta \in \Theta_0\}/\mathbb{P}\{\theta \in \Theta_1\}} \quad (18)$$

- How do we test $\theta = 0$? We'll come back to that later...