THE BAYESIAN MODELLING OF INFLATION RATE IN ROMANIA

Mihaela Simionescu

Institute for Economic Forecasting of the Romanian Academy
An introduction in the Bayesian inference from econometrics made by Zellner (1996);
The rapid progress of the computational techniques made possible the application on a larger scale of the specific methods of the Bayesian Econometrics
Most of the researchers use R and Matlab programs for Bayesian estimations
The regression models used in forecasting do not fulfil the necessity of development and frequent update. The classical regression model does not succeed in achieving these objectives. Therefore, the Bayesian regression is a very good solution to extend the low volume data series. It is actually an adaptation of famous Bayes theorem, taking into account two types of information:

1. A prior information;
2. Experimental data.
Bayes’ theorem applied in Economics

If A and B are two events, the first one being unknown and the second one being known, in Bayesian approach, B is associated to known data and A to model coefficients. The following notations are used:
y - set of data
y* - set of unobserved data
Mi - set of models, where i=1,2,...,m
θi - parameters that Mi depend on
\( p(\theta^i / Mi, y) \) - posterior density
p(Mi/y) - posterior probability of the model (the model is based on this model)
p(y*/y) - predictive density that the forecast is based on
The conditional probability of A, when B is known, represents the probability that A takes place, when B has already taken place: \( pr(A/B) = pr(A,B)/pr(B) \).

According to Bayes’ theorem: \( pr(A/B) = \frac{pr(B|A) \cdot pr(A)}{pr(B)} \).

We consider only one regression model that depends on parameters \( \theta_i \). according to Bayesian theory, \( pr(B/A) = \frac{pr(A|B) \cdot pr(B)}{pr(A)} \), but if A is replaced by y and B with \( \theta \), then: \( pr(\theta /y) = \frac{pr(y|\theta) \cdot pr(\theta)}{pr(y)} \) (y data being known, what can we know about \( \theta \)). There is a controversy between econometricians, more of them considering that \( \theta \) is not a random variable. \( p(y) \) not depending on \( \theta \), it can be ignored and the following approximation could be done for the posterior density being computed after the knowledge of the data: \( pr(\theta /y) \approx pr(y|\theta) \cdot pr(\theta) \).
$p(y|\theta)$ - likelihood function

$pr(\theta)$ - posterior density, that does not depend on data

Marginal density is based on integration: 
\[ (y^*|y) = \int p(y^*, \theta^*|y) d\theta \]

with 
\[ p(y^*|y) = \int p(y^*, y, \theta)p(\theta|y)d\theta \]
The linear regression model in the Bayesian approach. The estimation algorithm Gibbs Sampling

Gibbs sampling is a numerical method used for estimation and it is applied in 3 possible cases:

Estimation of prior distribution of $A$ under the hypothesis of known variance of errors;

Estimation of prior distribution of variance under the hypothesis of known $A$ matrix;

Estimation when both parameters are unknown.
Let us consider the following regression model (AR(2) model) for $Y_t$ (monthly index of consumer prices for Romania), used in computing the inflation rate, in the period 1991: January – 2013: April:

$$Y_t = \alpha + A_1 Y_{t-1} + A_2 Y_{t-2} + u_t, u_t \sim N(0, \sigma^2) \quad (17)$$

The RHS variables are: 1, $Y_{t-1}, Y_{t-2}$

$A = \{\alpha, A_1, A_2\}$ - vector of coefficients

Objective: the approximation of the marginal distribution for coefficients $(\alpha, A_1, A_2)$ and variance $(\sigma^2)$
Step 1: The priors and initial values are set. A normal distribution is set for the coefficients. This implies that the prior averages are specified for each coefficient ($A_0$)- a $3x1$ vector and the prior variance ($\Sigma_0$ - a $3x3$ matrix).

\[
p(A) \rightarrow N\left(\begin{pmatrix} \alpha^0 \\ A_0^0 \\ A_1^0 \\ A_2^0 \end{pmatrix}, \begin{pmatrix} \Sigma_\alpha & 0 & 0 \\ 0 & \Sigma_{A_1} & 0 \\ 0 & 0 & \Sigma_{A_2} \end{pmatrix}\right)
\]  \hspace{1cm} (18)

An inverse Gamma distribution with prior for $\sigma^2$ and the prior degrees of freedom $T_0$ and the prior scale matrix as $\theta_0$. We will work with inverse Gamma distribution.

\[
p(\sigma^2) \rightarrow \Gamma^{-1}\left(\frac{T_0}{2}, \frac{\theta_0}{2}\right)
\]  \hspace{1cm} (19)

The OLS estimator for $\sigma^2$ is set as starting value. The large number of Gibbs iterations will determine an insignificant influence of the starting value on the results for linear regressions.
Step 2: We sample from the conditional posterior repartition of $A$, having a starting value for $\sigma^2$ \textit{(normal distribution)}.

$$H(A \mid \sigma^2, Y_t) \rightarrow N(M^*, V^*)$$

$$M^* = (\Sigma_0^{-1} + \frac{1}{\sigma^2} X_t^T X_t)^{-1} (\Sigma_0^{-1} A_0 + \frac{1}{\sigma^2} X_t^T Y_t)$$

$$V^* = (\Sigma_0^{-1} + \frac{1}{\sigma^2} X_t^T X_t)^{-1}$$

(20)

The following algorithm is used to compute the draw for $A$.

Algorithm a: Let $z$ be a $k \times 1$ vector that is sampled from a normal distribution of average $m$ and variance $v$. Let $z_0$ be the first $k \times 1$ numbers from the standard normal distribution, numbers that could be transformed in order to have the mean $m$ and the variance $v$: $z = m + z_0 \ast v^{0.5}$. In our case, we have the following relationship:

$$A^1 = M^* + [\tilde{A} + (V^*)^{0.5}]$$

The draw for $A^1$ is computed using the previous formula, where $A^1 \textit{ is the first Gibbs iteration.}$

$\tilde{A}$ - a vector from the standard normal distribution
Step 3: The variance $\sigma^2$ is drawn from the conditional posterior distribution, $A^1$ being given, the repartition being an inverse Gamma one:

$$H(\sigma^2 \mid A, Y_t) \rightarrow \Gamma^{-1}\left(\frac{T_0}{2}, \frac{\theta_0}{2}\right)$$

(21)

$$T = T_0 + T$$

(22)

$$\theta_1 = \theta_0 + (Y_t - A^1X_t)^T(Y_t - A^1X_t)$$

The following algorithm is used in order to sample a scalar denoted by $z$ from the Inverse Gamma distribution ($T/2$ degrees of freedom and scale parameter $D/2$).

Algorithm b. We generate $T$ numbers from a standard normal distribution ($z^0$). Then, $z$ is computed as: $z = \frac{D}{(z^0)^Tz^0}$, being drawn from an Inverse Gamma repartition.

Step 4: The second and the third steps are repeated in order to get $A^1, \ldots, A^P$. The last $M$ values of $A$ are used to form the empirical repartition of the parameters (an approximate of the marginal posterior distribution). The first iterations (P-M iteration) that are not taken into consideration are considered burn-in iteration.
The prior mean for the coefficients is: \( \begin{pmatrix} \alpha^0 \\ A_1^0 \\ A_2^0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \), while the prior variance is represented by the identity matrix: \( \begin{pmatrix} \Sigma_\alpha \\ 0 \\ 0 \\ 0 \\ \Sigma_{A_1} \\ 0 \\ 0 \\ 0 \\ \Sigma_{A_2} \end{pmatrix} \). 10 000 replications were saved for this application, the total number of iterations being 50 000. Each coefficient has a posterior mean and the standard deviation (table no. 1).
Table no. 1 The coefficients, the posterior means and the standard deviations

<table>
<thead>
<tr>
<th>Coefficient</th>
<th>Posterior mean</th>
<th>Standard deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\alpha$</td>
<td>-0.0391</td>
<td>0.099883</td>
</tr>
<tr>
<td>$A_1$</td>
<td>0.0063</td>
<td>0.010486</td>
</tr>
<tr>
<td>$A_2$</td>
<td>0.0611</td>
<td>0.09542</td>
</tr>
</tbody>
</table>

Source: own computations

The posterior means are rather low, while the standard deviations are less than 0.1, fact that suggests the parameters’ stability, the model being better than the classical autoregressive one.
The Bayesian approach, which knew a major development in the last 20 years, finds its applicability in many domains of economics, being the support for taking decisions in various conditions.

The Bayesian Econometrics has many different applications, the Bayesian linear regression model being another perspective of modelling the variables’ dependences. The inclusion of prior information determines better estimations for the parameters, the situation being also reflected by the low values of the coefficients compared to the models from classical econometrics.

The future research directions should take into account the selection of the best prior values of the parameters based on the results provided by classical econometrics.
REFERENCES


THANK YOU!