Auto Regressive Process (1) with Change Point: Bayesian Approach

Dr. Mayuri Pandya, Paras Sheth
Department of Statistics, Maharaja Krishnakumarnsinhji Bhavnagar University, Bhavnagar, Gujarat, India.

Abstract: Here we consider first order autoregressive process with changing autoregressive coefficient at some point of time m. This is called change point inference problem. For Bayes estimation of m and autoregressive coefficient we used MHRW (Metropolis Hasting Random Walk) algorithm and Gibbs sampling. The effects of prior information on the Bayes estimates are also studied.

Keywords: Auto Regressive Model (1), Bayes Estimator, Change Point, Gibbs Sampling and MHRW Algorithm.

I. INTRODUCTION

Mayuri Pandya (2013) had studied the Bayesian analysis of the autoregressive model $X_t = \beta_1 X_{t-1} + \epsilon_t$, $t=1,2,...,m$ and $X_t = \beta_2 X_{t-1} + \epsilon_t$, $t=m+1,...,n$ where $0 < \beta_1 < 1$, $\beta_2 < 1$, and $\epsilon_t$ was independent random variable with an exponential distribution with mean $\theta_t$ and is reflected in the sequence after $\epsilon_m$ is changed in mean $\theta_2$. M. Pandya, K. Bhatt, H. Pandya, C. Thakar (2012) had studied the Bayes estimators of $m, \beta_1$ and $\beta_2$ under Asymmetric loss functions namely Linex loss & General Entropy loss functions of changing auto regression process with normal error. Tsurumi (1987) and Zacks (1983) are useful references on structural changes.

II. PROPOSED AR (1) MODEL:

Let our AR(1) model be given by,

$$X_i = \begin{cases} \beta_1 X_{i-1} + \epsilon_i, & i = 1,2,...,m. \\ \beta_2 X_{i-1} + \epsilon_i, & i = m+1,...,n. \end{cases}$$

where, $\beta_1$ and $\beta_2$ are unknown autocorrelation coefficients, $x_t$ is the $i^{th}$ observation of the dependent variable, the error terms $\epsilon_i$ are independent random variables and follow a N(0, $\sigma_i^2$) for $i=1,2,...,m$ and a N(0, $\sigma_f^2$) for $i=m+1,...,n$ and $\sigma_1^2$ and $\sigma_f^2$ both are known. $m$ is the unknown change point and $x_0$ is the initial quantity.

III. BAYES ESTIMATION

The Bayes procedure is based on a posterior density, say, $g(\beta_1, \beta_2, m \mid Z)$, which is proportional to the product of the likelihood function $L(\beta_1, \beta_2, m \mid Z)$, with a joint prior density, say, $g(\beta_1, \beta_2, m)$ representing uncertainty on the parameters values.

The likelihood function of $\beta_1$, $\beta_2$ and $m$, given the sample information $Z_t = (x_{t-1}, x_t)$, $t = 1, 2,..., m$, $m+1,..., n$, is,

$$L(\beta_1, \beta_2, m \mid Z) = K_1 \cdot \exp \left( -\frac{1}{2} \beta_1^2 \frac{S_{m1}}{\sigma_1^2} + \beta_1 \left( \frac{S_{m2}}{\sigma_f^2} - \frac{A_{1m}}{2\sigma_1^2} \right) - \frac{1}{2} \beta_2^2 \frac{S_{n1} - S_{m1}}{\sigma_2^2} + \beta_2 \left( \frac{S_{n2} - S_{m2}}{\sigma_2^2} \right) \right)$$

Where,

$$S_{k1} = \sum_{i=1}^{k} x_{i-1}^2 \quad S_{k2} = \sum_{i=1}^{k} x_i x_{i-1}$$

$$A_{1m} = \sum_{i=1}^{m} x_i^2 \quad A_{2m} = \sum_{i=m+1}^{n} x_i^2$$

$$k_1 = (2\pi)^{-\frac{n}{2}}$$


www.ijres.org 2017 | PP. 27-31
3.1 Using Informative (Normal) Priors On $\beta_1, \beta_2$

In this section, we derive posterior density of change point $m$, $\beta_1$ and $\beta_2$ of the model explained in equation (1) under informative priors.

We consider the AR(1) model shown in equation (1) with unknown $\sigma^{-2}$. We suppose uniform prior of change point same as Broemeling (1987), we also suppose that $m, \beta_1$ and $\beta_2$ are independent.

$$g(m) = \frac{1}{n-1}$$

We have normal prior density on $\beta_1$ and $\beta_2$ as,

$$g(\beta_1) = \frac{1}{\sqrt{2\pi a_1}} e^{-\frac{1}{2} \left( \frac{\beta_1}{a_1} \right)^2}$$

$$g(\beta_2) = \frac{1}{\sqrt{2\pi a_2}} e^{-\frac{1}{2} \left( \frac{\beta_2}{a_2} \right)^2}$$

Hence, joint prior p.d.f. of $\beta_1$, $\beta_2$ and $m$, say $g(\beta_1, \beta_2, m)$ is

$$g(\beta_1, \beta_2, m) = \frac{1}{2\pi a_1 a_2 (n-1)} e^{-\frac{1}{2} \left( \frac{\beta_1}{a_1} \right)^2} e^{-\frac{1}{2} \left( \frac{\beta_2}{a_2} \right)^2}$$

Using Likelihood function (2) with the joint prior density (4), the joint posterior density of $\beta_1, \beta_2, m$ say $g(\beta_1, \beta_2, m|Z)$ is

$$g(\beta_1, \beta_2, m|Z) = \frac{K_1}{h_1(Z)} [L(\beta_1, \beta_2, m|Z) \cdot g(\beta_1, \beta_2, m)]$$

$$= K_2 \frac{h_1(Z)}{h(Z)} e^{-\frac{1}{2} \beta_1^2 A_1 + \beta_1 B_1} e^{-\frac{1}{2} \beta_2^2 A_2 + \beta_2 B_2} e^{-\left( \frac{4A_1 \bar{m} + 4A_2 \bar{m}}{2A_1} \right)}$$

$$= \sigma_1^{-m} \sigma_2^{-(n-m)}$$

Where,

$$h_1(Z) = \sum_{m=1}^{n-1} e^{-\frac{1}{2} \beta_1^2 A_1 + \beta_1 B_1} e^{-\frac{1}{2} \beta_2^2 A_2 + \beta_2 B_2}$$

$$h(Z) = \sum_{m=1}^{n-1} e^{-\frac{1}{2} \beta_1^2 A_1 + \beta_1 B_1} e^{-\frac{1}{2} \beta_2^2 A_2 + \beta_2 B_2}$$

$$T_1(m) = k_m G_1 m G_2 m$$

$$G_1 m = \int_{-\infty}^{\infty} \exp \left[ -\frac{1}{2} \beta_1^2 A_1 + \beta_1 B_1 \right] d\beta_1 = \frac{e^{\frac{1}{2} A_1 \sqrt{2\pi}}}{\sqrt{A_1}}$$

$$G_2 m = \int_{-\infty}^{\infty} \exp \left[ -\frac{1}{2} \beta_2^2 A_2 + \beta_2 B_2 \right] d\beta_2 = \frac{e^{\frac{1}{2} A_2 \sqrt{2\pi}}}{\sqrt{A_2}}$$

$$k_m = e^{\left( \frac{4A_1 \bar{m} + 4A_2 \bar{m}}{2A_1} \right)}$$

Marginal posterior density of change point $m, \beta_1$ and $\beta_2$ are.

$$g_1(m|x) = \frac{T_1(m)}{\sum_{m=1}^{n-1} T_1(m)}$$

$$g_1(\beta_1|X) = \frac{k_3}{h_1(X)} \sum_{m=1}^{n-1} k_m e^{-\frac{1}{2} \beta_1^2 A_1 + \beta_1 B_1}$$

$$g_2(\beta_2|X) = \frac{k_3}{h_1(X)} \sum_{m=1}^{n-1} k_m e^{-\frac{1}{2} \beta_2^2 A_2 + \beta_2 B_2}$$

$G_1 m, G_2 m$ and $k_m$ are define as shows in equation (9), (10) and (11) respectively.

Now, the Bayes estimator of any function of parameter $\alpha$, say $g(\alpha)$ under the squarred loss function is,

$$E_{\alpha}(g(\alpha|Z)) = \int_{\alpha} g(\alpha|Z) \cdot d\alpha$$

Where, $g(\alpha|Z)$ is marginal posterior density of $\alpha$. It is complicate to compute equation (*) analytically in this case. Therefore, we use MCMC methods to find the Bayes estimator of $\beta_1, \beta_2$ and $m$. 

www.ijres.org 28 | Page
Gibbs Sampling algorithm:

Given a posterior distribution \( g(\alpha | Z) \) for unknown parameters \( \alpha = (\alpha_1, \ldots, \alpha_k) \) defined, at least up to proportionality, by multiplying the likelihood function with the corresponding prior distribution, we can easily identify the full conditional distribution \( g(\alpha_i | Z, \alpha_{-i}) \), up to proportionality, by regarding \( g(\alpha | Z) \) as a function of \( \alpha_1, \ldots, \alpha_{-i}, \alpha_i \), to be fixed.

To implement the Gibbs sampling procedure, we re-write (13) as full conditional of \( \beta_1 \), by fixing all other parameters i.e. \( \beta_2 \) and \( m \). Hence full conditional density of \( \beta_1 \) given \( \beta_2 \) and \( m \) is as follows,

\[
g(\beta_1 | \beta_2, m, Z) \propto N\left( \frac{\beta_1}{\beta_2}, \left( \frac{1}{\sqrt{\beta_2}} \right)^2 \right)
\]

where \( A_1 \) and \( B_1 \) shows in equation (6).

We re-write (14) as full conditional density of \( \beta_2 \) by fixing all other parameters \( \beta_1 \) and \( m \), we get the full conditional density of \( \beta_2 \) given \( \beta_1, \sigma^2 \) and \( m \) as follows,

\[
g(\beta_2 | \beta_1, m, Z) \propto N\left( \frac{\beta_2}{\beta_1}, \left( \frac{1}{\sqrt{\beta_1}} \right)^2 \right)
\]

where \( A_2 \) and \( B_2 \) shows in equation (6).

In order to estimate the parameter \( \beta_1 \), and \( \beta_2 \) we use Gibbs sampling to generate sample from the full conditional density of \( \beta_1 \) and \( \beta_2 \) given respectively in (15) and (16). We use following algorithm:

Algorithm:

Initialize \( \beta_1 = \beta_{10}, \beta_2 = \beta_{20} \) and \( m = m_0 \) then,

Step-1: Generate \( \beta_1 \sim N\left( \frac{A_1}{B_1}, \left( \frac{1}{\sqrt{B_1}} \right)^2 \right) \), using Gibbs Sampling.

Step-2: Generate \( \beta_2 \sim N\left( \frac{A_2}{B_2}, \left( \frac{1}{\sqrt{B_2}} \right)^2 \right) \), using Gibbs Sampling.

Step-3: Repeat the above steps.

MCMC techniques:

Since the posterior distribution of change point (12) has no closed form, we propose to use MCMC techniques to generate the samples from the posterior distribution. To implement the MCMC Techniques, we re-write (12) as target function of \( m \), by fixing all other parameters i.e. \( \beta_1 \) and \( \beta_2 \). Hence target function of \( m \) given \( \beta_1 \) and \( \beta_2 \) is as follows,

\[
g(m | \beta_1, \beta_2, Z) \propto k_m e^{-\frac{1}{2} \beta_1^2 \alpha_1 + \beta_1 b_1} e^{-\frac{1}{2} \beta_2^2 \alpha_2 + \beta_2 b_2}
\]

where \( A_1, B_1, A_2, B_2 \) and \( k_m \) shows in equation (6) and (11) respectively.

IV. NUMERICAL STUDY

Application To Generated Data

Let us consider AR(1) model as

\[
X_t = \begin{cases} 
0.1X_{t-1} + \epsilon_t, & i = 1, 2, \ldots, 10 \\
0.3X_{t-1} + \epsilon_t, & i = 11, 12, \ldots, 20
\end{cases}
\]

(18)

Where, the error terms \( \epsilon_t \) are independent random variables and follow a \( N(0, 1) \) for \( i=1,2,\ldots,10 \). and a \( N(0,4) \) for \( i=11, \ldots,20 \) and \( \sigma^2_1 \) and \( \sigma^2_2 \) known. \( m \) is the unknown change point and \( x_0 = 0.1 \) is the initial quantity. We have generated 20 random observations from proposed AR(1) model given in (18). The first ten observations are from normal distribution with \( \sigma^2 = 1 \) and next 10 are from normal distribution with \( \sigma^2 = 4 \). \( \beta_1 \) and \( \beta_2 \) themselves were random observations from normal distributions with prior means \( \mu_1 = 0.1, \mu_2 = 0.3 \) and variances \( \sigma_1 = 0.1 \) and \( \sigma_2 = 0.1 \). These observations are given in table-1.

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \epsilon_t )</td>
<td>0.167</td>
<td>-0.208</td>
<td>0.399</td>
<td>-0.259</td>
<td>-0.784</td>
<td>-1.058</td>
<td>0.819</td>
<td>0.404</td>
<td>1.215</td>
<td>1.537</td>
</tr>
<tr>
<td>( \beta_1 )</td>
<td>0.157</td>
<td>-0.221</td>
<td>0.420</td>
<td>-0.299</td>
<td>-0.758</td>
<td>-0.979</td>
<td>0.925</td>
<td>0.322</td>
<td>1.175</td>
<td>1.416</td>
</tr>
<tr>
<td>( \beta_2 )</td>
<td>11</td>
<td>12</td>
<td>13</td>
<td>14</td>
<td>15</td>
<td>16</td>
<td>17</td>
<td>18</td>
<td>19</td>
<td>20</td>
</tr>
</tbody>
</table>

To generate a random sample from (9.3) using the RWM-H algorithm, the selected proposal is uniform (2, 19) same as prior, which is symmetric around 10 with small steps. Since the target function is bounded. The initial distribution is chosen as uniform (1, 19). Further we truncate the initial distribution and we get integer value of the Bayes estimate of change point \( m \) is 10 when Selected Proposal is U(1, 19) and Initial
Distribution is $U(3, 14)$. The results are shown in Table 2 for data given in Table1 when given value of $\beta_1=0.1$, $\beta_2=0.3$, $\sigma_1^2=1$ and $\sigma_2^2=16$.

Table 2: Bayes Estimates of Change point (m) using RWM-H algorithm under SEL

<table>
<thead>
<tr>
<th>Bounded proposal</th>
<th>Selected proposal</th>
<th>Initial distribution</th>
<th>Bayes Estimate of Change Point (m)</th>
<th>Integer value of Bayes estimate of Change Point (m)</th>
</tr>
</thead>
<tbody>
<tr>
<td>BD(2,19)</td>
<td>U(1,19)</td>
<td>U(1,19)</td>
<td>8.4</td>
<td>8</td>
</tr>
<tr>
<td>BD(2,19)</td>
<td>U(2,19)</td>
<td>U(2,19)</td>
<td>8.6</td>
<td>9</td>
</tr>
<tr>
<td>BD(3,19)</td>
<td>U(1,19)</td>
<td>U(1,19)</td>
<td>10.3</td>
<td>10</td>
</tr>
<tr>
<td>BD(3,19)</td>
<td>U(1,19)</td>
<td>U(3,14)</td>
<td>10.2</td>
<td>10</td>
</tr>
</tbody>
</table>

We also compute the Bayes estimators of m using RWM-H algorithm for different prior consideration for data given in Table 1. The results are shown in Table 3.

Table 3: Bayes Estimates of Change point (m) using RWM-H algorithm under SEL for different prior consideration.

<table>
<thead>
<tr>
<th>Sr. No.</th>
<th>$a_1^2$</th>
<th>$a_2^2$</th>
<th>Bayes Estimate of change point (m) (Posterior Mean)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.0100</td>
<td>0.01</td>
<td>10</td>
</tr>
<tr>
<td>2</td>
<td>0.0400</td>
<td>0.04</td>
<td>10</td>
</tr>
<tr>
<td>3</td>
<td>0.0490</td>
<td>0.04</td>
<td>10</td>
</tr>
<tr>
<td>4</td>
<td>0.0550</td>
<td>0.09</td>
<td>10</td>
</tr>
<tr>
<td>5</td>
<td>0.0600</td>
<td>0.25</td>
<td>10</td>
</tr>
<tr>
<td>6</td>
<td>0.0625</td>
<td>0.49</td>
<td>10</td>
</tr>
<tr>
<td>7</td>
<td>0.0900</td>
<td>0.64</td>
<td>10</td>
</tr>
<tr>
<td>8</td>
<td>0.4900</td>
<td>0.81</td>
<td>10</td>
</tr>
<tr>
<td>9</td>
<td>0.8100</td>
<td>1.00</td>
<td>10</td>
</tr>
<tr>
<td>10</td>
<td>1.0000</td>
<td>4.00</td>
<td>10</td>
</tr>
</tbody>
</table>

Now we compute the Bayes estimators of $\beta_1$ (when given value of $\beta_2=0.3$, $m=10$, $\sigma_1^2=1$ and $\sigma_2^2=16$) and $\beta_2$ (when given value of $\beta_1=0.1$, $m=10$, $\sigma_1^2=1$ and $\sigma_2^2=16$) using Gibbs sampling MCMC algorithm for different prior consideration for data given in Table 1. The results are shown in Table 4.

Table 4: Bayes Estimates of $\beta_1$ and $\beta_2$ using Gibbs Sampling MCMC algorithm under SEL for different prior consideration.

<table>
<thead>
<tr>
<th>Sr. No.</th>
<th>$a_1^2$</th>
<th>$a_2^2$</th>
<th>Bayes Estimates of $\beta_1$</th>
<th>S.D. of Bayes Estimates of $\beta_1$</th>
<th>Bayes Estimates of $\beta_2$</th>
<th>S.D. of Bayes Estimates of $\beta_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.0100</td>
<td>0.01</td>
<td>0.025</td>
<td>0.048</td>
<td>0.255</td>
<td>0.048</td>
</tr>
<tr>
<td>2</td>
<td>0.0400</td>
<td>0.04</td>
<td>0.090</td>
<td>0.048</td>
<td>0.305</td>
<td>0.048</td>
</tr>
<tr>
<td>3</td>
<td>0.0490</td>
<td>0.04</td>
<td>0.107</td>
<td>0.048</td>
<td>0.305</td>
<td>0.048</td>
</tr>
<tr>
<td>4</td>
<td>0.0550</td>
<td>0.09</td>
<td>0.118</td>
<td>0.048</td>
<td>0.344</td>
<td>0.048</td>
</tr>
<tr>
<td>5</td>
<td>0.0600</td>
<td>0.25</td>
<td>0.126</td>
<td>0.048</td>
<td>0.367</td>
<td>0.048</td>
</tr>
<tr>
<td>6</td>
<td>0.0625</td>
<td>0.49</td>
<td>0.130</td>
<td>0.048</td>
<td>0.374</td>
<td>0.048</td>
</tr>
<tr>
<td>7</td>
<td>0.0900</td>
<td>0.64</td>
<td>0.172</td>
<td>0.048</td>
<td>0.376</td>
<td>0.048</td>
</tr>
<tr>
<td>8</td>
<td>0.4900</td>
<td>0.81</td>
<td>0.415</td>
<td>0.048</td>
<td>0.377</td>
<td>0.048</td>
</tr>
<tr>
<td>9</td>
<td>0.8100</td>
<td>1.00</td>
<td>0.475</td>
<td>0.048</td>
<td>0.378</td>
<td>0.048</td>
</tr>
<tr>
<td>10</td>
<td>1.0000</td>
<td>4.00</td>
<td>0.496</td>
<td>0.048</td>
<td>0.381</td>
<td>0.048</td>
</tr>
</tbody>
</table>

Figure 1 graph the full conditional of $\beta_1$ when a sample of size 10000 is generated from (9.1), Gibbs Sampling with MCMC algorithm has been run. ($\beta_2=0.3$, $m=10$, $\sigma_1^2=1$ & $\sigma_2^2=16$)
Figure 2 graph the full conditional of $\beta_2$ when a sample of size 10000 is generated from (9.2), Gibbs Sampling with MCMC algorithm has been run. ($\beta_1=0.1$, $m=10$, $\sigma_1^2=1$ & $\sigma_2^2=16$)

![Figure 2: Full Conditional of $\beta_2$](image)

REFERENCES


