A Bayesian Analysis of Structural Change in Linear Models

ARNULFO P. SUPE

Abstract

The problem of estimating break points in linear models which undergo structural change is analyzed using the Bayesian approach. Posterior distributions of the break points as well as the other parameters of the model are derived. A prior distribution based on past data is used.

1. Introduction

Structural change has, in the past, often been ignored in model building. The oil price "shock" of 1972-73 and the Ninoy Aquino assassination in 1983 are two major tests that many economic models, the Gross National Product (GNP) and the Consumer Price Index (CPI) among others, failed to pass. The widespread model problems associated with these events created an important stimulus to reevaluate and improve these models. It is thus the role of statistical analysis to detect the presence of structural change and to find ways to assimilate it in its models

In a regression framework, *Structural Change* may simply be defined as a change in one or more of the parameters of the model. Although coefficients in statistical models are usually assumed to be constant, it is often recognized that in applied work, some relationships change over time, especially after some sudden unforeseen events like war, revolutions, coup d'etats or major calamities.

Page (1955) was the first to study structural change in simple sequences of independent random variables based on cumulative sums or CUSUMS. Chow (1960) developed an F-test for a known break point while Quandt (1960) developed a test based on the likelihood function when the break point is unknown. Broemeling and Tsurumi (1987) discus-

ARNULFO P SUPE is an Associate Professor of Mathematics and Statistics at the College of Sciences and Mathematics. He finished his PhD in Statistics at the University of the Philippines, Diliman, Quezon City.

ses structural change in linear models using a Normal-Gamma prior distribution.

2. Methodology

Bayesian inferential procedures will be employed for the most part of the analysis. Although the Bayesian methodology faced a lot of criticisms in its earlier growth and especially during the 1970's, most of these criticisms have already been addressed with and the method has gained worldwide acceptance since then. Let

$\Theta = (\Theta_1, \Theta_2, \dots, \Theta_r)$ be the parameter of interest,

$h(\theta)$ the prior density associated with Θ, and

$f(x|\theta)$ the density from which the sample was taken.

Bayes' Theorem states that the ***posterior density*** of Θ given the sample information, denoted by $\pi(\theta|x)$, is for a continuous Θ,

$$\pi(\theta|x) = \frac{h(\theta) f(x|\theta)}{m(x)}, \tag{2.1}$$

where

$$m(x) = \int \cdots \int f(x|\cdot) h(\cdot) d\theta .$$

Since $m(x)$ does not involve θ, we may rewrite (2.1) as

$$\pi(\theta|x) \propto h(\theta) f(x|\theta)$$

where the symbol "$\propto$" means "is proportional to". This simplification is used in the illustration in Section 4.

3. The Model and the Prior

Consider the following model.

$$Y_i = \begin{cases} \mathbf{X}_i\beta_1 + \varepsilon_i; & i = 1, 2, \Lambda, \nu \\ \mathbf{X}_i\beta_2 + \varepsilon_i; & i = \nu+1, 2, \Lambda, n \end{cases} \tag{3.1}$$

where ν, $1 \le \nu \le n$, is the unknown break point,

$\varepsilon_i \sim N(0,\sigma^2)$; $i = 1, 2, \ldots, n$

$\mathbf{X}_i = (X_{1i}, X_{2i}, \ldots, X_{ri})$ is a $1 \times r$ vector of explanatory variables, $i = 1, \ldots, n$;

$\beta_j = (\beta_{1j}, \beta_{2j}, \ldots, \beta_{rj})'$ is the $r \times 1$ vector of parameters, $j = 1, 2$.

If *Structural Change* is present in Model (3.1), then $1 \le \nu \le n-1$. However, if no structural change occurs, $\nu = n$. Therefore, testing for the presence of structural change is like testing the hypothesis that

$\mathbf{H}_0 :\ 1 \le \nu \le n-1$ against

$\mathbf{H}_1 :\ \nu = n$.

The decision on whether to accept or reject the hypothesis will be based on the posterior probabilities of ν.

Assume that past data $(\mathbf{X}_1^p, Y_1^p), (\mathbf{X}_2^p, Y_1^p), \ldots, (\mathbf{X}_m^p, Y_m^p)$ exists. To find the prior for (β, δ), we fit the past data into the model

$$\mathbf{Y}^p = \mathbf{X}^p\beta + \varepsilon$$

where

$$\mathbf{Y}^p = \begin{bmatrix} \mathbf{Y}_1^p \\ Y_2^p \\ \vdots \\ Y_m^p \end{bmatrix}, \quad \mathbf{X}^p = \begin{bmatrix} \mathbf{X}_1^p \\ \mathbf{X}_2^p \\ \vdots \\ \mathbf{X}_m^p \end{bmatrix}, \tag{3.2}$$

and the other parameters are as defined in (3.1).

Using a noninformative prior for β_2 and ν, the joint prior of (β,δ,ν) is given by

$$\rho(\beta,\delta,\nu) \propto \delta^{\frac{m}{2}} \exp\left\{-\frac{\delta}{2}\left[(\mathbf{Y}^p - \mathbf{X}^p\beta_1)'(\mathbf{Y}^p - \mathbf{X}^p\beta_1)\right]\right\} \tag{3.3}$$

4. Posterior Analysis

Let $\delta = 1/\sigma^2$, and let

$$\mathbf{Y}_1 = \begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_v \end{bmatrix},\ \mathbf{Y}_2 = \begin{bmatrix} Y_{v+1} \\ Y_{v+2} \\ \vdots \\ Y_n \end{bmatrix},\ \mathbf{Z}_1 = \begin{bmatrix} \mathbf{X}_1 \\ \mathbf{X}_2 \\ \vdots \\ \mathbf{X}_v \end{bmatrix},\ \mathbf{Z}_2 = \begin{bmatrix} \mathbf{X}_{v+1} \\ \mathbf{X}_{v+2} \\ \vdots \\ \mathbf{X}_n \end{bmatrix}, \quad (4.1)$$

$$\mathbf{Y} = \begin{bmatrix} \mathbf{Y}_1 \\ \mathbf{Y}_2 \end{bmatrix},\ \mathbf{Z} = \begin{bmatrix} \mathbf{Z}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{Z}_2 \end{bmatrix},\ \beta = \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix}.$$

The likelihood function is

$$L(\beta,\delta,\nu \mid (\mathbf{Z},\mathbf{Y})) \propto \delta^{\frac{n}{2}} \exp\left\{\frac{-\delta}{2}\left[(\mathbf{Y}-\mathbf{Z}\beta)'(\mathbf{Y}-\mathbf{Z}\beta)\right]\right\}. \quad (4.2)$$

Combining (3.3) and (4.2) with Bayes' Theorem, we obtain

$$\pi(\beta,\delta,\nu|(\mathbf{Z},\mathbf{Y})) \propto$$

$$\propto \delta^{\frac{n+m}{2}} \exp\left\{-\frac{\delta}{2}\left[(\mathbf{Y}_1-\mathbf{Z}_1\beta_1)'(\mathbf{Y}_1-\mathbf{Z}_1\beta_1)+(\mathbf{Y}_2-\mathbf{Z}_2\beta_2)'(\mathbf{Y}_2-\mathbf{Z}_2\beta_2)\right]\right\} \times$$

$$\times \exp\left\{-\frac{\delta}{2}\left[(\mathbf{Y}^p - \mathbf{X}^p\beta_1)'(\mathbf{Y}^p - \mathbf{X}^p\beta_1)\right]\right\}$$

After some algebraic manipulations, the last equation can be rewritten as

$$\pi(\beta,\delta,\nu|(\mathbf{Z},\mathbf{Y})) \propto \delta^{\frac{n+m}{2}} \exp\left\{-\frac{\delta}{2}\left[(\beta-\beta^*)'\mathbf{Z}^{p'}\mathbf{Z}^p(\beta-\beta^*) + g(\nu)\right]\right\}, \quad (4.3)$$

where

$$\mathbf{Z}^{p'}\mathbf{Z}^p = \begin{bmatrix} \mathbf{Z}_1'\mathbf{Z}_1 + \mathbf{X}^{p'}\mathbf{X}^p & \mathbf{0} \\ \mathbf{0} & \mathbf{Z}_2'\mathbf{Z}_2 \end{bmatrix},\ \beta^* = \begin{bmatrix} \beta_1^* \\ \beta_2^* \end{bmatrix},$$

$$\beta_1^* = (\mathbf{Z}_1'\mathbf{Z}_1 + \mathbf{X}^{p'}\mathbf{X}^p)^{-1}(\mathbf{Z}_1'\mathbf{Y}_1 + \mathbf{X}^{p'}\mathbf{Y}^p),$$

$$\beta_2^* = (\mathbf{Z}_2'\mathbf{Z}_2)^{-1}(\mathbf{Z}_2'\mathbf{Y}_2), \text{ and}$$

$$g(\nu) = \mathbf{Y}_1'\mathbf{Y}_1 + \mathbf{Y}_2'\mathbf{Y}_2 - (\mathbf{Y}_2'\mathbf{Z}_2)(\mathbf{Z}_2'\mathbf{Z}_2)^{-1}(\mathbf{Z}_2'\mathbf{Y}_2) +$$

$$+ Y_1^{p'} Y_1^p - (Y_1' Y_1 + Y_1^{p'} Y_1^p)(Z_1' Z_1 + X^{p'} X^p)^{-1}(Z_1' Y_1 + Y_1^{p'} Y_1^p)$$

To find the marginal posterior density of ν, we integrate out β and δ from (4.3). To integrate out β, apply Aitken's integral

$$\int \cdots \int \exp\left\{-\frac{1}{2}(x' Ax)dx\right\} \propto |A|^{-\frac{1}{2}} \tag{4.4}$$

where $\mathbf{A}$ is positive definite. Integrating out β from (4.3), yields

$$\pi(\delta,\nu|(\mathbf{Z},\mathbf{Y})) \propto \delta^{\frac{n+m-2r}{2}} \exp\left\{-\frac{\delta}{2}[g(\nu)]\right\} \cdot \left|\mathbf{Z}^{p'} \mathbf{Z}^p\right|^{-\frac{1}{2}}$$

$$\propto \delta^{\frac{n+m-2r}{2}} \exp\left\{-\frac{\delta}{2}[g(\nu)]\right\} \cdot \left|\mathbf{Z}_1' \mathbf{Z}_1 + \mathbf{X}^{p'} \mathbf{X}^p\right|^{-\frac{1}{2}} \left|\mathbf{Z}_2' \mathbf{Z}_2\right|^{-\frac{1}{2}} \tag{4.5}$$

To integrate out δ, we use the property of the Gamma distribution

$$\int x^{k-1} \exp\{-\lambda k\} dx = \lambda^{-k} \tag{4.6}$$

Integrating out δ from (4.5), we therefore have

$$\pi_1(\nu|(\mathbf{Z},\mathbf{Y})) \propto \left|\mathbf{Z}_1' \mathbf{Z}_1 + \mathbf{X}^{p'} \mathbf{X}^p\right|^{-\frac{1}{2}} \left|\mathbf{Z}_2' \mathbf{Z}_2\right|^{-\frac{1}{2}} g(\nu)]^{-(\frac{n+m-2r}{2})} \tag{4.7}$$

Relation (4.7) is the posterior density of ν and we can choose as our point estimate of the break point the value of ν which attains the highest posterior density. However, a plot of the whole density will usually give a much clearer picture.

5. Estimation of Parameters

To estimate the parameters of the model, we also derive their posterior densities. If we are now sure that structural change is present, model (3.1) can be rewritten as

$$Y_i = \begin{cases} \mathbf{X}_i\beta_1 + \varepsilon_i & ; \quad i = 1,2,☹ \nu \\ \mathbf{X}_i\beta_2 + \varepsilon_i & ; \quad i = \nu+1,2,☹ n \end{cases}$$

where $1 \le \nu \le n-1$. Note that n is now excluded as a value of ν. From (4.3), the posterior density of (β,δ), for a fixed ν, is

$$\pi(\beta,\delta|\nu,(\mathbf{Z},\mathbf{Y})) \propto \delta^{\frac{n-m}{2}} \exp\left\{-\frac{\delta}{2}[(\beta-\beta^*)'\mathbf{Z}^{p'}\mathbf{Z}^p(\beta-\beta^*) + g(\nu)]\right\}$$

(5.1)

Integrating out δ from (5.1) using (4.6), yields

$$\pi_2(\beta|\nu,(\mathbf{Z},\mathbf{Y})) \propto g(\nu) + (\beta-\beta^*)'(\mathbf{Z}^{p'}\mathbf{Z}^p)(\beta-\beta^*)]^{-(\frac{n+m+2}{2})}$$

which can be rewritten as

$$\pi_2(\beta|\nu,(\mathbf{Z},\mathbf{Y})) \propto$$

$$\propto \left[1+\frac{\frac{(n+m-2r+2)(\beta-\beta^*)'(\mathbf{Z}^{p'}\mathbf{Z}^p)(\beta-\beta^*)}{g(\nu)}}{(n+m-2r+2)}\right]^{-(\frac{n+m-2}{2})} \tag{5.2}$$

We can distinguish (5.2) as the kernel of a multivariate t-distribution with

degrees of freedom $(n+m-2r+2)$,
mean vector β^*, and
precision matrix $\dfrac{(n+m-2r+2)(\mathbf{Z}^{p'}\mathbf{Z}^p)}{g(\nu)}$.

Summing now for all values of ν, the marginal conditional posterior density of β is therefore given by

$$\pi_3(\beta|(\mathbf{Z},\mathbf{Y}) = \sum_{\nu=1}^{n-1}[\pi_2(\beta|\nu,(\mathbf{Z},\mathbf{Y})) \cdot \pi_1(\nu|(\mathbf{Z},\mathbf{Y}))], \tag{5.3}$$

where $\pi_2(\beta|\nu,(\mathbf{Z},\mathbf{Y}))$ is defined in (5.2) and $\pi_1(\nu|(\mathbf{Z},\mathbf{Y}))$ is defined in (4.7).

Therefore, the marginal posterior distribution of β, is a mixture of multivariate t-distributions where the mixing probabilities are the marginal posterior probabilities of ν. To find the marginal density of δ, first fix the value of ν in (4.3). Integrating out β from (4.3) using (4.4), yields

$$\pi(\delta|\nu,(\mathbf{Z},\mathbf{Y})) \propto \delta^{\frac{n+m-2r}{2}} \exp\left\{-\frac{\delta}{2}[\mathbf{g}(\nu)]\right\} \cdot \left|[\mathbf{Z}^{p'}\mathbf{Z}^{p}]\right|^{-\frac{1}{2}}, \text{ or}$$

$$\pi_4(\delta|\nu,(\mathbf{Z},\mathbf{Y})) \propto \delta^{\frac{n+m-2r}{2}} \exp\left\{-\frac{\delta}{2}[\mathbf{g}(\nu)]\right\} \quad (5.4)$$

where $\left|[\mathbf{Z}^{p'}\mathbf{Z}^{p}]\right|^{-\frac{1}{2}}$ is absorbed into the constant of proportionality. We can recognize (5.4) as the kernel of a Gamma distribution with parameters $(n+m-2r)/2$ and $g(\nu)$. Summing now for all values of ν, we have

$$\pi_5(\delta|(\mathbf{Z},\mathbf{Y}) = \sum_{\nu=1}^{n-1}[\pi_4(\delta|\nu,(\mathbf{Z},\mathbf{Y})) \cdot \pi_1(\nu(\mathbf{Z},\mathbf{Y}))] \quad (5.5)$$

where $\pi_4(\delta|\nu,(\mathbf{Z},\mathbf{Y}))$ is defined in (5.4), and $\pi_1(\nu|(\mathbf{Z},\mathbf{Y}))$ is defined in (4.7).

The marginal posterior distribution of δ is a mixture of Gamma distributions where the mixing probabilities are the posterior probabilities of ν.

References

[1] Berger, J. O. (1985) : *Statistical Decision Theory and Bayesian Analysis*, pp. 118-138

[2] Broemeling and Tsurumi (1987) : *Econometrics and Structural Change*, pp. 65-86

[3] Chow (1960) : *Tests of the Equality of two sets of coefficients in two Linear Regresssions*, **Econometrica** 28, pp. 581-605

[4] Page , E. S. (1955) : *A Test for a change in a parameter occurring at an unknown time point*, **Biometrika** 42, pp. 523-527

[5] Quandt, R. E. (1960) : *Test of the Hypothesis that a linear regression system obeys two separate regimes*, JASA 55, pp 324-331