UNIVERSIDADE DE SÃO PAULO

Bayesian estimation of the slope of a linear functional regression model

RODRIGUES, J. ; BOLFARINE, H.

Nº 72

NOTAS

Instituto de Ciências Matemáticas de São Carlos

ISSN 0103-2577

Bayesian estimation of the slope of a linear functional regression model

RODRIGUES, J. ; BOLFARINE, H.

Nº 72

N O T A S D O I C M S C

São Carlos (SP)
1990

BAYESIAN ESTIMATION OF THE SLOPE OF A LINEAR FUNCTIONAL REGRESSION MODEL

Josemar Rodrigues and Heleno Bolfarine
Universidade de São Paulo-Departamento de Estatistica
Caixa Postal 20570-Agencia Iguatemi
São Paulo-SP-Brasil

Summary

In this paper, we discuss the problem of obtaining posterior distributions for the slope parameter of a simple functional regression model. By making use of an orthogonal transformation, we show that the generalized least squares estimator proposed by Sprent (1966) is exactly the mode of the posterior distribution.

key words: orthogonal reparametrization; generalized least squares estimator; posterior mode

1. Introduction

The classical functional regression model is defined by the equations

$$
\begin{aligned}
Y_k &= y_k + e_k, \\
X_k &= x_k + u_k \qquad (1)\\
y_k &= \beta x_k,
\end{aligned}
$$

where e_k and u_k are normal independent measurement error variables with mean zero and variances σ_{ee} and σ_{uu}, respectively, that is, $e_k \sim N(0, \sigma_{ee})$ and $u_k \sim N(0, \sigma_{uu})$, $k = 1, \ldots, n$. The inference problem is to estimate β considering $\mathbf{x} = (x_1, \ldots, x_n)$ as nuisance parameters based on the data (X_k, Y_k), $k = 1, \ldots, n$. Let $\mathbf{X} = (X_1, \ldots, X_n)$ and $\mathbf{Y} = (Y_1, \ldots, Y_n)$. Extensive bibliographies on functional simple regression models are given in Kendall and Stuart (1961) and in Fuller (1987).

The main idea behind the equations (1) is that (y_k, x_k), $k = 1, \ldots, n$ are not observed directly in order to estimate the parameter β. An interesting situation is the case where x_k is the amount of nitrogen in the soil and y_k is the yield of a certain cereal. Values for X_k in this case can only be determined by laboratory analysis and are only estimates of the true x_k values.

Bayesian estimation of β is considered mainly by Lindley (1966), Lindley and El-Sayad (1968) and Zellner (1971). Lindley (1966) in his discussion of the paper by Sprent (1966), tried to give a Bayesian justification for Sprent's generalized least squares estimator. But in order to do it, an important factor in the posterior density had to be ignored. In this paper, we show that, by considering first an orthogonal transformation (Cox and Reid, 1987), the resulting marginal posterior density of β is free from the factor that Lindley (1966) ignored. Furthermore, the posterior mode is exactly Sprent's estimator. This gives an exact Bayesian justification for the estimator proposed by Sprent (1966) and is the subject of Section 2. In Section 3, we extend the results in Section 2 to the case where the variances are equal and unknown.

2. An Orthogonal Transformation for Known Error Variances

Considering that model (1) is appropriate, the likelihood function for $(\beta, \mathbf{x})$ is given by

$$(2) \qquad L(\beta,\mathbf{x}) \propto exp\{-\frac{1}{2}[\sum_{k=1}^{n}\frac{(Y_k-\beta x_k)^2}{\sigma_{ee}}+\sum_{k=1}^{n}\frac{(X_k-x_k)^2}{\sigma_{uu}}]\}.$$

The Fisher information matrix for $(\beta, \mathbf{x})$ is

$$(3) \qquad I(\beta,\mathbf{x}) = \begin{pmatrix} \sum_{i=1}^{n}\frac{x_k^2}{\sigma_{ee}} & \beta\frac{x_1}{\sigma_{ee}} & \cdots & \beta\frac{x_n}{\sigma_{ee}} \\ \beta\frac{x_1}{\sigma_{ee}} & \frac{\beta^2}{\sigma_{ee}}+\frac{1}{\sigma_{uu}} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ \beta\frac{x_n}{\sigma_{ee}} & 0 & \cdots & \frac{\beta^2}{\sigma_{ee}}+\frac{1}{\sigma_{uu}} \end{pmatrix}.$$

As pointed out by Cox and Reid (1987), the orthogonal transformation is a reparametrization which diagonalize the Fisher information matrix (3). Therefore, we transform the original parameter $(\beta, \mathbf{x})$ in model (2) to the orthogonal parametrization (β, ϕ). The orthogonal parameter $\phi = (\phi_1, \ldots, \phi_k)$ is obtained by solving the differential equations

$$(4) \qquad (\frac{\beta^2}{\sigma_{ee}}+\frac{1}{\sigma_{uu}})\frac{\partial x_k}{\partial \beta} = -\frac{\beta x_k}{\sigma_{ee}},$$

$k = 1, \ldots, n$. These differential equations give the following one to one transformations

$$(5) \qquad x_k = \frac{\phi_k}{\sqrt{\frac{\beta^2}{\sigma_{ee}}+\frac{1}{\sigma_{uu}}}},$$

$k = 1, \ldots, n$. Then, the new likelihood function in terms of β and ϕ is

$$(6) \quad L(\beta,\phi) \propto exp\{-\frac{1}{2}[\frac{1}{\sigma_{ee}}\sum_{k=1}^{n}(Y_k-\frac{\beta\phi_k}{\sqrt{\frac{\beta^2}{\sigma_{ee}}+\frac{1}{\sigma_{uu}}}})^2+\frac{1}{\sigma_{uu}}\sum_{k=1}^{n}(X_k-\frac{\phi_k}{\sqrt{\frac{\beta^2}{\sigma_{ee}}+\frac{1}{\sigma_{uu}}}})^2]\}.$$

The Fisher information matrix corresponding to (6) is

$$(7) \qquad I(\beta,\phi) = \begin{pmatrix} \frac{\sum_{k=1}^{n}x_k^2}{\beta^2\sigma_{uu}+\sigma_{ee}} & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}.$$

Based on (7), we suggest a noninformative prior for (β, ϕ) given by

$$(8) \qquad \pi(\beta,\phi) \propto constant.$$

Combining (6) and (8) and integrating out ϕ, we arrive, after some algebraic manipulations, at the marginal posterior density of β, which is given by

$$\pi(\beta|\mathbf{X},\mathbf{Y}) \propto exp\{\frac{-1}{2(\beta^2\sigma_{uu}+\sigma_{ee})}[S_{yy}-2\beta S_{xy}+\beta^2 S_{xx}]\},$$

or, equivalently,

(9) $$\pi(\beta|\mathbf{X},\mathbf{Y}) \propto exp\{-\frac{1}{2}\frac{(S_{xx}-\lambda\sigma_{uu})}{\beta^2\sigma_{uu}+\sigma_{ee}}(\beta-\hat{\beta})^2\},$$

where

$$S_{xx}=\frac{1}{n}\sum_{k=1}^{n}X_k^2,\ S_{yy}=\frac{1}{n}\sum_{k=1}^{n}Y_k^2,$$

$$S_{xy}=\frac{1}{n}\sum_{k=1}^{n}X_kY_k,\ \hat{\beta}=\frac{S_{xy}}{S_{xx}-\lambda\sigma_{uu}},$$

λ is the smallest root of the determinantal equation

$$|\mathbf{B}-\lambda\mathbf{\Sigma}|=0,$$

(10) $$\mathbf{B}=\begin{pmatrix} S_{xx} & S_{xy} \\ S_{xy} & S_{xx}\end{pmatrix} \text{ and } \mathbf{\Sigma}=\begin{pmatrix} \sigma_{ee} & 0 \\ 0 & \sigma_{uu}\end{pmatrix}.$$

Note that the mode $\hat{\beta}$ of the posterior density (9) is the generalized least squares estimator derived by Sprent (1966). The factor which had to be neglected in Lindley's (1966) posterior density does not appear in the posterior density (9). In this way, by making use of the orthogonal reparametrization (5), the posterior density (9) provides an exact Bayesian justification for Sprent's estimator.

3. An Orthogonal Transformatiom for Unknown Error Variances

In this Section, we consider that the ratio of the error variances σ_{ee}/σ_{uu} is known. Therefore, we can, without loss of generality, take $\sigma_{ee}=\sigma_{uu}=\sigma^2$, where σ^2 is unknown. An easy calculation shows that $(\beta,\mathbf{x})$ is orthogonal to σ^2. By making use of the differential equation (4), we obtain the one to one orthogonal transformation given by

$$x_k=\frac{\phi_k}{\sqrt{1+\beta^2}},$$

$k=1,\ldots,n$. Since the orthogonal parameters are by construction weakly dependent on β, it seems reasonable to consider the following prior

(11) $$\pi(\beta,\phi,\sigma^2)\propto\frac{1}{\sigma^2}.$$

The likelihood function in the new reparametrization is given by

$$(12) \quad L(\beta,\phi,\sigma^2) \propto (\frac{1}{\sigma^2})^{\frac{n+1}{2}} exp\{-\frac{1}{2\sigma^2}[\sum_{k=1}^{n}(Y_k - \frac{\phi_k}{\sqrt{1+\beta^2}})^2 + \sum_{k=1}^{n}(Y_k - \frac{\beta\phi_k}{\sqrt{1+\beta^2}})^2]\}.$$

Integrating out ϕ and σ^2, from the joint posterior density of (β, ϕ, σ^2) which follows by combining (11) and (12), we obtain the following marginal posterior density

$$\pi(\beta|\mathbf{X},\mathbf{Y}) \propto \{\frac{(S_{xx}-\lambda)}{\beta^2+1}(\beta-\hat{\beta})^2\}^{-\frac{n}{2}},$$

where λ is the smallest root of the determinantal equation

$$|\mathbf{B} - \lambda\mathbf{I}| = 0,$$

$\mathbf{B}$ is as given in (10), $\mathbf{I}$ is the 2x2 identity matrix and

$$\hat{\beta} = \frac{S_{xy}}{S_{xx}-\lambda}.$$

References

COX, D.R. & REID, N. (1987). Parameter Orthogonality and Approximate Conditional Inference (with discussion), JRSS, B, 49, 1-49.

FULLER, W.A. (1987). *Measurement Error Models.* Wiley

KENDALL, M.G. & STUART, A. (1961). *The Advanced Theory of Statistics*, 2, Griffin.

LINDLEY, D.V. (1966). Discussion on "A Generalized Least Squares Approach to Linear Functional Relationships", JRSS, B, 28, 279-297.

LINDLEY, D.V. & EL SAYAD, G. (1968). The Bayesian Estimation of a Linear Functional Relationship, JRSS, B, 30, 190-202.

SPRENT, P. (1966). A Generalized Least Squares Approach to Linear Functional Relationships, JRSS, B, 28, 279-297.

ZELLNER, A. (1971). *An Introduction to Bayesian Inference in Econometrics*, Wiley.

NOTAS DO ICMSC - USP