ISSN - 0103-2577

A Bayesian analysis of the generalized least-squares procedure to functional relationship

JOSEMAR RODRIGUES

Nº 83

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

São Carlos (SP)
1991

A Bayesian analysis of the generalized least-squares procedure to functional relationships

by
Josemar Rodrigues
ICMSC - USP, C.P. - 668
13.560 - São Carlos - SP.

Summary

In this paper we present an exact Bayesian justification of the least-squares estimates for the functional relationships situation (Sprent, 1966) considering the Jeffreys prior. It is shown that the profile posterior and the marginal posterior do not give conflicting information about the parameter of interest if an orthogonal reparametrization (Cox and Reid, 1987) is adopted for the nuisance parameter.

Key Words: Orthogonality; profile posterior; marginal posterior; generalized least-squares.

Introduction

Sprent (1966), an important paper, applied the "Generalized Least-Squares Principle" to estimate the slope of a regression line relating response variable Y_k to a scalar covariate x_k.

The estimation is based on the observations $(Y_1, X_1), \cdots, (Y_n, X_n)$, where

$$\begin{aligned} Y_k &= y_k + e_k \\ X_k &= x_k + u_k \\ Y_k &= \beta x_k \quad , k = 1,\cdots,n. \end{aligned} \tag{1}$$

In that paper, e_k and u_k were independent measurement error variables with mean zero and known variances σ_{ee} and σ_{uu}, respectively.

The justification of the generalizad least-squares procedure under normality assumptions was given by Lindley in his discussion of Sprent's paper (1966), from the Bayesian view point. In this paper, adopting the Jeffreys prior, we show the generalized least-squares estimator maximizes exactly the marginal posterior of β given the data (X,Y), that is, we do not need to ignore a certain skewness in the marginal posterior due the term $\sigma_{ee} + \beta^2\sigma_{uu}$ in expression (4) of Lindley's discussion, (Sprent,1966). Also, we show the profile posterior is equal to the marginal posterior if an orthogonal reparametrization is adopted for the nuisance paramenter $x = (x_1,\cdots,x_n)$. Let us follows Lindley's discussion (Sprent, 1966) considering the linear functional model (1) with

$$e_k \sim N(o,\sigma_{ee}) \quad and \quad u_k \sim N(o,\sigma_{uu}), \tag{2}$$

where σ_{uu} and σ_{ee} are known variances. In this paper we consider the problem of inference about the parameter of interest

$\varphi = \beta$ in the presence of the nuisance parameter $\phi = x = (x_1,\ldots,x_n)$.

2 - An exact Bayesian justification of the generalized least-Squares Procedure

Considering the model (1) and (2) are appropriate the log of the likelihood function for (φ,ϕ) is given by

$$\ell(\varphi,\phi) \propto -\frac{1}{2}\left\{\frac{\sum_{k=1}^{n}(Y_k - \varphi x_k)^2}{\sigma_{ee}} + \frac{\sum_{k=1}^{n}(X_k - x_k)^2}{\sigma_{uu}}\right\}. \tag{3}$$

The Fisher information matrix for φ and ϕ is

$$I(\varphi,\phi)=\begin{bmatrix} \frac{\sum_{k=1}^{n} x_k^2}{\sigma_{ee}} & \frac{\varphi x_1}{\sigma_{ee}} & \cdots\cdots & \frac{\varphi x_n}{\sigma_{ee}} \\ \frac{\varphi x_1}{\sigma_{ee}} & \frac{\varphi^2}{\sigma_{ee}}+\frac{1}{\sigma_{uu}} & 0\ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\varphi x_n}{\sigma_{ee}} & 0\ldots\ldots\ 0 & & \frac{\varphi^2}{\sigma_{ee}}+\frac{1}{\sigma_{uu}} \end{bmatrix} \tag{4}$$

Let us assume a constant prior for φ. Applying the Jeffrey's rule (Box and Tiao,1973) we obtain the following prior for ϕ given φ:

$$\Pi(\phi/\varphi)\ \alpha\ \mid I_\phi(\varphi)\mid^{\frac{1}{2}} = \left(\frac{\varphi^2}{\sigma_{ee}}+\frac{1}{\sigma_{uu}}\right)^{\frac{n}{2}}, \tag{5}$$

where $I_\phi(\varphi)$ is a submatrix of (4) w.r.t. ϕ. So, the joint prior for (φ,ϕ) is

$$\Pi(\varphi,\phi)\ =\ \Pi(\varphi)\Pi(\phi/\varphi)\alpha\left(\frac{\varphi^2}{\sigma_{ee}}+\frac{1}{\sigma_{uu}}\right)^{\frac{n}{2}}. \tag{6}$$

For a Bayesian the inferences are typically based on the marginal posterior

$$\Pi(\varphi/(X,Y))\alpha\left(\frac{\varphi^2}{\sigma_{ee}}+\frac{1}{\sigma_{uu}}\right)^{\frac{n}{2}}.\int e^{\ell(\varphi,\phi)}d\phi. \tag{7}$$

For fixed φ, expanding $l(\varphi,\phi)$ in a second order Taylor's series about the maximizing value $\hat{\phi}_\varphi$ of $\ell(\varphi,\phi)$, we have the Laplace approximation

$$\Pi(\varphi/(X,Y))\alpha\left(\frac{\varphi^2}{\sigma_{ee}}+\frac{1}{\sigma_{uu}}\right)^{\frac{n}{2}}.e^{l_p(\varphi)}\mid J(\varphi,\hat{\phi}_\varphi)\mid^{\frac{-1}{2}} \tag{8}$$

where

$$J(\varphi,\phi)=-\frac{\partial^2\ell(\varphi,\phi)}{\partial\phi^2}\quad and\quad \ell_p(\varphi)\qquad is$$

the likelihood profile of φ, that is,

$$\ell_p(\varphi) = \ell(\varphi, \hat{\phi}_\varphi).$$

For a general discussion of the Laplace method in Bayesian Analysis, see Tierney and Kadane (1986). From (3) we have that

$$\hat{\phi}_\varphi = (\hat{x}_1, \ldots, \hat{x}_k),$$

$$\hat{x}_k = \frac{\frac{\varphi Y_k}{\sigma_{ee}} + \frac{X k}{\sigma_{uu}}}{\frac{\varphi^2}{\sigma_{ee}} + \frac{1}{\sigma_{uu}}}, \quad k = 1, \ldots, n \tag{9}$$

$$| J(\varphi, \hat{\phi}_\varphi) | = \left(\frac{\varphi^2}{\sigma_{ee}} + \frac{1}{\sigma_{uu}} \right)^n,$$

$$\ell_p(\varphi) \alpha - \frac{1}{2(\sigma_{uu}\varphi^2 + \sigma_{ee})} \{S_{YY} - 2\varphi S_{XY} + S_{XX}\varphi^2\},$$

where

$$S_{XX} = \frac{\sum_{k=1}^n X_k^2}{n}, \quad S_{YY} = \frac{\sum_{k=1}^n Y_k^2}{n} \quad and$$

$$S_{XY} = \frac{\sum_{k=1}^n X_k Y_k}{n}$$

The expression of $\ell_p(\varphi)$ in (9), as in Lindley's discussion of the Sprent's paper (1986), can be conveniently be written, so that,

$$\Pi(\varphi/(X,Y)) \; \alpha \; e^{-\frac{1}{2}(S_{XX} - \lambda\sigma_{uu})\frac{(\varphi - \hat{\varphi}_p)^2}{\sigma_{uu}\varphi^2 + \sigma_{ee}}}, \tag{10}$$

where

$$\hat{\varphi}_p = \frac{S_{XY}}{S_{XX} - \lambda\sigma_{uu}} \quad and \quad \lambda \quad is \quad the$$

smallest root of the determinantal equation

$$| B - \lambda \textstyle\sum | = 0,$$

$$B = \begin{bmatrix} S_{YY} & S_{XY} \\ S_{XY} & S_{XX} \end{bmatrix} \; and \; \sum = diag\{\sigma_{ee}, \sigma_{uu}\}.$$

The estimative $\hat{\varphi}_p$ is the generalized least-square estimate of Sprent (1966). It is interesting to observe that $\Pi(\varphi/(X,Y))$ is exactly the profile $l_p(\varphi)$ and we do not have to ignore anything in (10) as in Lindley's discussion (1966). An alternative method of removing the nuisance parameter ϕ is by maximization, that is, the profile posterior given by

$$\Pi_p(\varphi) \;=\; \Pi(\varphi, \hat{\phi}_\varphi) l_p(\varphi) \; \alpha \; \left(\frac{\varphi^2}{\sigma_{ee}} + \frac{1}{\sigma_{uu}}\right)^{\frac{n}{2}} e^{-\frac{S_{XX} - \lambda\sigma_{uu}}{2(\sigma_{uu}\varphi^2 + \sigma_{ee})}\left(\varphi - \hat{\varphi}_p\right)^2} \quad (11)$$

The profile posterior (11), as in Lindley's discussion (1966), do not justify exactly the generalized-least estimate unless we ignore our prior distribution $\left(\frac{\varphi^2}{\sigma_{ee}} + \frac{1}{\sigma_{uu}}\right)^{\frac{n}{2}}$ which does not make sense from the Bayesian wiowpoint. Since this prior can not be ignored we have from (10) and (11) confliting inferences about φ. To avoid this problem an orthogonal reparametrization (Cox and Reid, 1987) is adopted for the nuisance parameter ϕ.

3 - The orthogonal reparametrization of the parameter ϕ.

As discussed by Box and Reid (1987), the orthogonal reparametrization is a transformation which orthogonalizes the Fisher matrix (4). Therefore, we transform the origin parameters (φ, ϕ) in model (3) to the orthogonal parameters (φ, λ).

The orthogonal parameter λ is obtained by solving the differential equations,

$$\left(\frac{\varphi^2}{\sigma_{ee}} + \frac{1}{\sigma_{uu}}\right)\frac{dx_k}{d\varphi} = -\frac{\varphi x_k}{\sigma_{ee}} \qquad k = 1, \ldots, n \tag{12}$$

These differential equations give a one-to-one transformations:

$$x_k = \frac{\lambda_k}{a_\varphi}, \quad k = 1, \ldots, n,$$

where

$$a_\varphi^2 = \frac{\varphi^2}{\sigma_{ee}} + \frac{1}{\sigma_{uu}}. \tag{13}$$

Then the new likelihood function in terms of φ and $\lambda = (\lambda_1, \ldots, \lambda_n)$ is

$$\ell(\varphi, \lambda) \propto -\frac{1}{2}\left\{\frac{1}{\sigma_{ee}}\sum (Y_k - \varphi\frac{\lambda_k}{a_\varphi})^2 + \frac{1}{\sigma_{uu}}\sum(X_k - \frac{\lambda_k}{a_\varphi})^2\right\} \tag{14}$$

For fixed φ, let $\hat{\lambda}_\varphi = (\hat{\lambda}_{\varphi\,1}, \ldots, \hat{\lambda}_{\varphi\,n})$ the value maximizing $\ell(\varphi, \lambda)$. Then we have that

$$\hat{\lambda}_{\varphi\,k} = a_\varphi\,\hat{x}_k, \quad k = 1, \ldots, n. \tag{15}$$

and the Fisher matrix w.r.t. (φ, λ) is given by

$$I(\varphi,\lambda) = \begin{bmatrix} \frac{\sum_{i=1}^{n} X_i^2}{\varphi^2\sigma_{uu} + \sigma_{ee}} & 0 & \cdots\cdots & 0 \\ 0 & 1 & \cdots\cdots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ 0\cdots\cdots\cdots & \vdots\cdots & \cdots\cdots & 1 \end{bmatrix} \qquad (16)$$

Let us denote $\ell_p^{\perp}(\varphi), \Pi_p^{\perp}(\varphi)$ and $\Pi^{\perp}(\varphi/(X,Y))$, the profile, the profile posterior and the marginal posterior in the new parametrization respectively. From (15) we have that

$$\ell_p^{\perp}(\varphi) = \ell_p(\varphi) \; \alpha \; exp\left\{\frac{(S_{XX} - \lambda\,\sigma_{uu})}{\sigma_{ee}\varphi^2 + \sigma_{uu}}(\varphi - \hat{\varphi}_p)^2\right\}. \qquad (17)$$

Since the Jeffrey's prior is constant in this new situation and

$J(\varphi,\lambda) = -\frac{\partial^2\ell(\varphi,\lambda)}{\partial\lambda^2}$ is the identity matrix, then it follows from (17) and the Laplace approximation that

$$\Pi_p^{\perp}(\varphi) = \Pi^{\perp}(\varphi/(X,Y)) = \ell_p^{\perp} = \ell_p(\varphi) \; \alpha \; exp\left\{\frac{(S_{XX} - \lambda\sigma_{uu})}{\varphi^2\sigma_{ee} + \sigma_{uu}}(\varphi - \hat{\varphi}_p)\right\} (18)$$

In this orthogonalized situation we have an exactly justification of the generalized least-squares estimative and no conflicting inferences from the Bayesian viewpoint.

Also, we show that the marginal posterior is invariant under the orthogonal transformation, that is

$$\Pi(\varphi/(X,Y)) = \Pi^{\perp}(\varphi/(X,Y)). \qquad (19)$$

References

[Box. G.E.P. and Tiao, G.C.(1973).] - Bayesian inference in Statistical Analysis, Wesley.

[Cox, D. R. and Reid, N. (1987)] - Parameter orthogonality and approximate conditional inference, J. R. Statist. Soc., B, 49, 1-39.

[Sprent , P. (1986)] - A generalized least-squares approach to linear functional relationships. J.R. Statist. Soc. B., 28, 278-297.

[Tierney, L. and Kadane, J.B. (1986)] - Accurate approximations for posterior moments and marginal densities, J. of the Amer. Statist. Assoc., 81,82-86.

NOTAS DO ICMSC - USP

Nº 82/90 - LIZANA PEÑA, M. - Exponential dichotomy for singularity perturbed linear functional differential equations with small delays

Nº 81/90 - BERGAMASCO, A.P. - Perturbations of globally hipoelliptic operators

Nº 80/90 - MARAR, W.L. - On the image of $\mathcal{A}$-simple map-germs from $\mathbb{R}^2$ to $\mathbb{R}^3$

Nº 79/90 - RAPOPORT- CAMPODONICO, D.L. - On the stochastic processes associated to the conformal Riemann-Cartan-Weyl structures of the theory of gravitation

Nº 78/90 - CUMINATO, J.A. - On the convergence of a perturbed collocation method for a class of Cauchy integral equations

Nº 77/90 - TERENCE, G. - Aureoles and integral closure of modules

Nº 76/90 - ACHCAR, J.A. - Use of predective densities in quality control with accelerated life tests assuming a power rule model and an exponenetial distribution

Nº 75/90 - RODRIGUES, J.; ACHCAR, J.A.; LOUZADA NETO, F. - A Bayesian analysis of the accelerated life tests via the orthogonal parameters

Nº74/90 - MARAR, W.L. - The Euler characteristic of the disentanglement of the image of a corank 1 map germ

Nº 73/90 - ACHCAR, J.A.; LOUZADA NETO, F. - Accelerated life tests with an exponential distribution: A Bayesian approach with the generalized eyring model and type II censored data