ON THE BAYESIAN ESTIMATORS IN MULTIDIMENSIONAL NONLINEAR REGRESSION MODELS

By Ung Ngoc Quang
Department of Mathematics – Informatics, University of Natural Sciences, Viet Nam National University – Ho Chi Minh City.
(Received on October 29th, 2001)

SUMMARY:

In this paper the Bayesian estimators in multidimensional regression models is presented by using the functional analysis method.

1. Introduction:

It is wellknow, that linear and nonlinear regression models are used in a cosiderable part of the application of mathematical statistics Nevertheless whereas the statistical theory of parameter estimation in linear regression models is almost comletely developed, in the nonlinear case many problems are unsolved. Of course, a similarly complete theory can be hardly expected. Therefore, investigations should be concentrated on somewhat specialized moddel types, in which the least squared method occurs as the most important estimation method (see [1] – [4]).

On the otherd hand, in [5] – [11], we are investigated Bayesian estimators in nonlinear regression models with the compact parameter space by the functional analysis technique.

Continuously, in this paper, we will present some results on the Bayesian estimators in multidimensional nonlinear regression models, in which the functional space $L^1(\mu)$ playes an essential role.

2. On the Bayesian estimators for the location parameter:

First of all, we give some notation (see [9], [11]).

$M(n \times q)$, $M(p \times r)$: Spaces of all $n \times q$ –matrices and $p \times r$ –matrices.

$\beta(n \times q)$, $\beta(p \times r)$: Borel σ –algebras in $M(n \times q)$ and $M(p \times r)$

Let us consider the following multidimensional nonlinear regression models

$$X = \varphi(\theta) + \varepsilon$$

Where,

X is a observed random variable, taking the values in $M(n \times q)$ and $EX = \varphi(\theta)$.

ε is a random error variable, taking the values in $M(n \times q)$ and $E\varepsilon = 0$.

Ⓗ is a compact subset of $M(p \times r)$.

θ is a unknown location parameter, $\theta \in$ Ⓗ

φ is a known fuction $\varphi :$ Ⓗ $\longrightarrow M(n \times q)$

It is wellknown, that for a random variable X, there exists a conditional regular distribution $P^{X|\theta}$. We denote $P^{X|\theta}$ by Q_θ, $\forall\theta \in$ Ⓗ . Assume that μ is a σ-finite measure in the measurable space $(M(n \times q), \beta(n \times q))$ and $Q_\theta \ll \mu$, $\forall\theta \in$ Ⓗ

Then, there exists a function $f_\theta(x)$ such that

$$f_\theta(x) = \frac{Q_\theta(dx)}{\mu(dx)}$$

Suppose that there exists C'> 0 such that

$$|f_\theta(x)| \leq C' \ (\text{mod}\mu), \ \forall\theta \in Ⓗ$$

Definition 2.1: A Borel function $h : (M(n \times q), \beta(n \times q)) \rightarrow (M(p \times r), \beta(p \times r))$ is called an estimator of the location parameter $\theta \in Ⓗ$.

We define $L^1(\mu) = L^1(\mu, M(n\ q), M(p\ r))$ to be the colletion of all Borel measurable functions h on M(n x q) for which $\int_{M(nxq)} \| h(x) \| \mu(dx) < +\infty$

Clearly, $L^1(\mu)$ is a Banach space with the norm $\| h \|_1 = \int_{M(nxq)} \| h(x) \| \mu(dx)$

The composed function defined by $L(h(.), .) : M(n \times q) \rightarrow \overline{R}^+$ is called a loss function (see [9]).

Definition 2.2: A functional $\Psi : L^1(\mu) \rightarrow \overline{R}^+$ is said to be a Bayesin risk function with a priori distribution τ if

$$\Psi(h) = \int_{M(nxq)} L(h(x),\theta).f_\theta(x)\mu(dx)\tau(d\theta)$$

An estimator $\hat{h} \in L^1(\mu)$ is said to be a Bayesian estimator with a priori distribution τ if

$$\Psi(\hat{h}) = \inf_{h \in L^1(\mu)} \Psi(h)$$

Theorem 2.1: Let $K \subset L^1(\mu)$ be a class of all estimators of the parameter $\theta \in Ⓗ \subset M(pxr)$ satisfying the following conditions:

(i) $h(M(nxq)) \subset Ⓗ \ (\text{mod}\mu), \forall h \in K$.

(ii) $\forall\varepsilon > 0, \exists$ finite partition $\{E_i\}_{i=1}^m \subset M(nxq)$ and points $x_i \in E_i$, i = 1,2, ..., m such that

(a) $C'' : \| h(x_i) \| \leq C''$, $\forall h \in K, \forall i = 1, 2,..., m$.

(b) $\int_{E_i} \| h(x) - h(x_1) \| \mu(dx) < \frac{\varepsilon}{m}, \forall h \in K, \forall i = 1,2,...,m$.

(iii) There exists C > 0 such that

$$| L(y',\theta) - L(y'',\theta) | \leq C \| y' - y'' \|, \forall y', y'' \in M(pxr), \forall\theta \in Ⓗ$$

Then K is a relatively compact subset of the space $L^1(\mu)$ and in the class $\overline{K}$ there exists a Bayesian estimator.

Proof: Let us consider the function $\phi : L^1(\mu) \to (M(pxr))^m$ defined by

$$\phi(h) = (h(x_1), h(x_2), ..., h(x_m))$$

Arguing similarly as in the proof of the theorem 2.1 in [6], we can show that there exists functions $h_j, j = 1, ..., s$ such that

$$K \subset \bigcup_{j=1}^{s} B(h_j, 4\varepsilon)$$

Then K is a totally bounded subset and it follows that K is a relatively compact subset of $L^1(\mu)$.

Next, we will prove that $h(M(nxq)) \subset (H)(mod\mu), \forall h \in \overline{K}$.

Indeed, for any $h \in \overline{K}$, there exists a sequence $(h_m) \subset K$ such that

$$\| h_m - h \|_1 \to 0, \text{ as } m \to \infty$$

Therefore, there exists a subsequence $(h_{m_k}) \subset (h_m)$ such that

$$\| h_{m_k}(x) - h(x) \| \to 0(mod\mu), \text{ as } k \to +\infty$$

Arguing similarly as in the proof of the theorem 2.2 in [6], we see that

$$h(M(nxq)) \subset (H)(mod\mu), \forall h \in \overline{K}.$$

Finally, consider the Bayesian risk function

$$\Psi(h) = \int_{M(nxq)} L(h(x),\theta) f_\theta(x)\mu(dx)\tau(d\theta)$$

It can be shown that Ψ is a continuous function on $L^1(\mu)$. It follows that there exists a Bayesian estimator $\hat{h} \in \overline{K}$ and the proof of the theorem is completed.

Let us consider the 1-dimentional nonlinear model. Then, we have following theorem.

Theorem 2.2: Let K be a class of all estimators of the location parameter $\theta \in (H)$ satisfying the condition of the Theorem 2.1 . Then the Bayesian estimators $\hat{h} \in \overline{K}$ can be arbitrarily closely approximated by a polinomial function. (See a similar argument in [7]).

3. On the Bayesian estimator for the variance parameter:

In this section we will investigate the Bayesian estimators for the variance parameter in the (q,r) – models. First, let R^{nq} be the nq-dimensional Euclidian space. Let us consider the following mapping $T : M(nxq) \to R^{nq}$, defined by

$$T(A) = \vec{A} = (a_{11}, a_{21}, ..., a_{n1}, a_{12}, a_{22}, ... , a_{n2}, ..., a_{1q}, a_{2q}, ... , a_{nq})$$

Where $A = (a_{ij}) \in M(nxq)$, $\vec{A}$ is a nq-dimensional vector.

Obviously, T is a linear isometry of M(nxq) onto R^{nq}.

The covariance matrix $Cov(\varepsilon)$ of the nq-dimensional random error vector $\vec{\varepsilon}$ is called the variance component of the random error matrix ε and denoting by $Var\varepsilon$. For each s, let

$M^{\geq}(sxs)$ denote the space of all non-negative definite sxs-matrices. We will assume that $Var = \psi(\sigma^2)$, where

$$\psi : M^{\geq}(sxs) \rightarrow M(nqxnq)$$

is a known nonlinear function and σ^2 is unknown parameter, $\sigma^2 \in M^{\geq}(sxs)$.
This unknown parameter is said to be the variance parameter.

Definition 3.1: A Borel function $h : (M(nxq), \beta(nxq)) \rightarrow (M(sxs), \beta(sxs))$ is called an estimator of the variance parameter $\sigma^2 \in M^{\geq}(sxs)$.

We define $L^1(\mu) = L^1(\mu, M(nxq), M(sxs))$ to be the collection of all Borel functions h on M(nxq) for which $\int_{M(nxq)} \| h(x) \| \mu(dx) < +\infty$

Definition 3.2: A functional $\Psi : L^1(\mu) \rightarrow \overline{R}^+$ is said to be a Bayesian risk function with a priori distribution ν if

$$\Psi(h) = \int_{M(sxs)} \int_{M(nxq)} L(h(x), \sigma^2).f_{\sigma^2}(x)\mu(dx)\upsilon(d\sigma^2)$$

An estimator $\hat{h} \in L^1(\mu)$ is said to be a Bayesian estimator of the variance parameter $\sigma^2 \in M^{>}(sxs)$ with a priori distribution ν if

$$\Psi(\hat{h}) = \inf_{h \in L^1(\mu)} \Psi(h)$$

Theorem 3.1: Let $K \subset L^1(\mu)$ be a class of estimotors of the variance parameter $\sigma^2 \in M^{\geq}(sxs)$ satisfying the following conditions

(i) $h(M(nxq)) \subset M^{\geq}(sxs)(mod\mu), \forall h \in K$

(ii) For any $\varepsilon > 0$, there exists a finite partition $\{E_i\}_{i=1}^m \subset M(nxq)$ and points $x_i \in E_i$, $i=1,2,...,m$ such that

(a) $\exists C" : \| h(x_i) \| \leq C', \forall h \in K, \forall i = \overline{1,m}$.

(b) $\int_{E_i} \| h(x) - h(x_i) \| \mu(dx) < \frac{\varepsilon}{m}, \forall h \in K, \forall i = \overline{1,m}$.

(iii) There exists $C > 0$ such that

$$| L(y', \sigma^2) - L(y", \sigma^2) | \leq C \| y' - y" \|, \forall y', y" \in M(sxs), \forall \sigma^2 \in M^{\geq}(sxs)$$

Then K is a relatively compact subset of $L^1(\mu)$ and in the class $\overline{K}$ there exists a Bayesian estimator.

Proof: By a similar argument of the theorem 2.1, it can be seen that K is a relatively compact subset of $L^1(\mu)$.

Next, it can be shown that

$$h(M(nxq)) \subset M^{\geq}(sxs)(mod\mu), \forall h \in K.$$

Indeed, take any $h \in K$. Then there exists a sequence $(h_m) \subset K$ such that

$$\| h_m - h \|_1 \rightarrow 0, \text{ as } m \rightarrow \infty$$

It follows that, there exists a subsequence $(h_{m_k}) \subset (h_m)$ such that

$$\| h_{m_k}(x) - h(x) \| \rightarrow 0 \ (mod\mu), \text{ as } k \rightarrow +\infty$$

By a similar argument of the theorem 3.2 in [9], it can be shown that $h(x) \in M^{\geq}(sxs)(mod\mu)$, which implies that $h(M(nxq)) \subset M^{\geq}(sxs)(mod\mu), \forall h \in \overline{K}$.

Clearly, the Bayesian function risk function ψ is a continuous function on the compact set $\overline{K}$. Therefore, in $\overline{K}$ there exists a Bayesian estimator and theorem is proved.

Theorem 3.2: Let K be a class of all estimators of the variance parameter $\sigma^2 \in \overline{R}^+ = [0,+\infty)$ satisfying the condition of the theorem 3.1 . Then the Bayesian estimator $\hat{h} \in \overline{K}$ can be arbitrarily closely approximated by a polinomial function (See a similar argument in [8]).

4. On the Bayesian estimator for the compound parameter:

In this section we simultaneously consider the location parameter $\theta \in Ⓗ \subset M(pxr)$ and the variance parameter $\sigma^2 \in M^{\geq}(sxs) \subset M(sxs)$. First, consider the product-space M(pxr) x M(sxs). Clearly, it is a finite-dimensional normed linear space with the norm

$$\| y \| = \| y' \|_{M(pxr)} + \| y'' \|_{M(sxs)}$$

where, y = (y' , y"). Denote by $\beta(pxr) \times \beta(sxs)$ the smallest σ -algebra in M(pxr) x M(sxs) which contain every measurable rectangle. Next, consider the subset $Ⓗ \times M^{\geq}(sxs) \subset M(pxr) \times M(sxs)$. Denote by $\beta(Ⓗ) \times \beta^{\geq}(sxs)$ the σ -algebra in $Ⓗ \times M^{>}(sxs)$. Reccal that τ and ν are priori distributions of the location parameter θ and the variance parameter σ^2.

An element $\lambda = (\theta, \sigma^2)$, where $\theta \in Ⓗ$ and $\sigma^2 \in M^{>}(sxs)$ is called a compound parameter. Denote by $\eta = \tau \times \nu$ the product of the measures τ and ν.

Then η is a probability measure on the measurable space ($Ⓗ \times M^{>}(sxs)$, $\beta(Ⓗ) \times \beta^{\geq}(sxs)$ and η is called a priori distribution of the compound parameter $\lambda = (\theta, \sigma^2)$.

Definition 4.1: A Borel function

$$h : (M(nxq), \beta(nxq)) \rightarrow (M(pxr) \times M(sxs), \beta(pxr) \times \beta(sxs))$$

is called an estimator of the compound parameter $\lambda = (\theta, \sigma^2)$.

We difine $L^1(\mu) = L^1(\mu, M(nxq), M(pxr) \times M(sxs))$ to be the collections of all Borel measurable functions h on M(nxq) for which

$$\int_{M(nxq)} \| h(x) \| \mu(dx) < +\infty$$

Clearly, $L^1(\mu)$ is a Banach with the norm $\| h \|_1 = \int_{M(nxq)} \| h(x) \| \mu(dx)$

Definition 4.2: A functional $\Psi : L^1(\mu) \rightarrow R^+$, defined by

$$\Psi(h) = \int_{Ⓗ \times M^{>}(sxs)} \int_{M(nxq)} L(h(x), \lambda).f_\lambda(x) \mu(dx) \eta(d\lambda)$$

is called a Bayesian estimator with a priori distribution η.

An estimator $\hat{h} \in L^1(\mu)$ is said to be a Bayesian estimator with a priori distribution η if

$$\Psi(\hat{h}) = \inf_{h \in L^1(\mu)} \Psi(h)$$

Theorem 4.1: Let $K \subset L^1(\mu)$ be a class of all estimators of the compound parameter $\lambda = (\theta, \sigma^2) \in Ⓗ \times M^{>}(sxs)$ satisfying the following conditions

(i) $h(M(nxq)) \subset Ⓗ \times M^{>}(sxs)$ (modμ), $\forall h \in K$.

(ii) $\forall \varepsilon > 0$, $\exists$ finite partition $\{E_i\}_{i=1}^{m} \subset M(nxq)$

and points $x_i \in E_i$, i = 1, 2, ... , m such that

(a) $\exists C^* : \| h(x_i) \| \leq C^*, \forall h \in K, \forall i = \overline{1,m}$.

(b) $\int_{E_i} \| h(x) - h(x_i) \| \mu(dx) < \frac{\varepsilon}{m}, \forall h \in K, \forall i = \overline{1,m}$

(iii) There exists $C > 0$ such that

$$| L(y',\lambda) - L(y'',\lambda) | \leq C \| y' - y'' \|, \forall y',y'' \in M(pxr) \times M(sxs), \forall \lambda \in Ⓗ \times M^{>}(sxs)$$

Then K is a relatively compact subset of $L^1(\mu)$ and in the class $\overline{K}$ there exists a Bayesian estimator.

Proof: Clearly, K is a relatively compact subset of $L^1(\mu)$.

Next, it can be seen that $h(M(nxq)) \subset Ⓗ \times M^{>}(sxs) \pmod{\mu}, \forall h \in \overline{K}$.

In fact, take any $h \in \overline{K}$. Then, there exists a sequence $(h_m) \subset K$ such that

$$\| h_m - h \|_1 \rightarrow 0 , \text{ as } m \rightarrow +\infty$$

This implies that, there exists a subsequence $(h_{m_k}) \subset (h_m)$ such that

$$\| h_{m_k}(x) - h(x) \| \rightarrow 0 \pmod{\mu} , \text{ as } k \rightarrow +\infty$$

It can be seen that $h_{m_k} = (h'_{m_k}, h''_{m_k})$, $h = (h', h'')$,

where $h'_{m_k}, h' \in L^1(\mu, M(pxr))$ and $h''_{m_k}, h'' \in L^1(\mu, M(sxs))$

It follows that $h(M(nxq)) \subset Ⓗ \times M^{>}(sxs) \pmod{\mu}, \forall h \in \overline{K}$ as to be shown.

Finally, the Bayesian risk function Ψ is a continuous function on compact subset $\overline{K}$ and theorem is proved.

Theorem 4.2: Let K be a class of all estimator of the compound parameter $\lambda = (\theta, \sigma^2)$ satisfying the conditions of the theorem 4.1 . Then the Bayesian estimator $\hat{h} \in \overline{K}$ can be arbitrarily closely approximated by a polinomial function.

Proof: By a similar argument of the theorem 3.1 in [10], it can be shown that the Bayesian estimator $\hat{h}$ can be arbitrarily closely approximated by a polinomial function.

VỀ ƯỚC LƯỢNG BAYES TRONG MÔ HÌNH HỒI QUI PHI TUYẾN NHIỀU CHIỀU

Ung Ngọc Quang
Khoa Toán-Tin Trường Đại Học Khoa Học Tự Nhiên, ĐHQG-HCM.

TÓM TẮT: Bài báo khảo sát ước lượng Bayes của tham ẩn định vị, tham ẩn phương sai và tham ẩn hỗn hợp trong mô hình hồi qui phi tuyến nhiều chiều bằng kỹ thuật giải tích hàm.

REFERENCE

[1]. Nguyễn Bác Văn, Strong consistency of least estimates in poligonal regression with random explanatory variables, Acta Math.Vietnamica, Vol.19, No.19, 1994, 111 –137.

[2]. Nguyễn Bác Văn, Strong consistency of least estimates in poligonal regression with random explanatory variables, Acta Math.Vietnamica, Vol.23, No.2, 1998.

[3]. Bunker H., Paramater estimation in nonlinear regression models, Math.Operationsforsc.Statistik, Vol.8, No.1, 1977, 23 – 40.

[4]. Clarke G.P.F., Moments of the least squares estimators in a nonlinear regression model, J.R.Statist.Soc.B, Vol.42, No.2, 1980, 227 –237.

[5]. Ung Ngoc Quang, On the existence of Bayesian estimates in statistical models with compact parameter space, Journal of Mathematics, Vol.18, No.1 , 1990, 1 – 8 (in Vietnamese).

[6]. Ung Ngoc Quang, On the existence of Bayesian estimates in nonlinear statistical models with compact parameter space, Acta Math.Vietnamica, Vol.19, No.2, 1994, 149 – 160.

[7]. Ung Ngoc Quang, An approach to the Bayesian estimates in nonlinear statistical models, Jounal of Computer science and Cybernetics, Vol.10, No.4, 1994, 35 – 40 (in Vietnamese).

[8]. Ung Ngoc Quang, On the Bayesian estimators of the variance component in statistical models, Jounal of Computer science and Cybernetics, Vol.11, No.4, 1995, 53 – 63 (in Vietnamese).

[9]. Ung Ngoc Quang, On the existence of Bayesian estimates in multidimensional nonlinear statistical models with compact parameter space, Vietnam Journal of Mathematics, Vol.23, No. 2, 1995, 229 – 240.

[10]. Ung Ngoc Quang, On the Bayesian estimators of the compound paramater in multidimensional nonlinear regression models, Journal of Computer science and Cybernetics, Vol.14, No.2 , 1998, 19 – 29 (in Vietnamese).

[11]. Ung Ngoc Quang, On the Bayesian estimators in the $L^1(\mu, M(pxr))$-space. The 2nd scientific congress, University of Natural Sciences, Vietnam National University- Ho Chi Minh City, May, 2000.