Discussiones Mathematicae
Probability and Statistics 26 (2006) 19–45

LEAST SQUARES ESTIMATOR CONSISTENCY: A GEOMETRIC APPROACH

João Tiago Mexia

Universidade Nova de Lisboa, Departamento de Matemática
da Faculdade de Ciências e Tecnologia
Quinta da Torre, 2825–114 Monte de Caparica, Portugal

e-mail: parcr@fct.unl.pt

and

João Lita da Silva

Universidade Nova de Lisboa, Departamento de Matemática
da Faculdade de Ciências e Tecnologia
Quinta da Torre, 2825–114 Monte da Caparica, Portugal

e-mail: jfls@fct.unl.pt

Abstract

Consistency of LSE estimator in linear models is studied assuming that the error vector has radial symmetry. Generalized polar coordinates and algebraic assumptions on the design matrix are considered in the results that are established.

Keywords: linear models, least squares estimator, consistency, radial symmetry, generalized polar coordinates.

1. Introduction

Consistency of the least square estimator (LSE) in linear models has been lately derived by several authors from distinct approaches (see for example [3], [4], [5], [15], [16], [17], [18], [19] and [20]). We will assume the random error sequence $e_1, e_2, \dots$ to have radial symmetry in the study of this problem.

It is worthwhile to point out that no assumption of error independence or of identical distribution for the $e_1, e_2, \ldots$ will be made. As a matter of fact radial symmetry ensures, as we shall see, the independence of the new random variables that we get when we use generalized polar coordinates. In connection with the use of this coordinates we will obtain the distributions of the relevant random variables. These results will be useful in establishing consistency for the LSE.

We now state the following definition.

Definition 1. A random vector $(X_1, \ldots, X_n)$ has *radial symmetry* if it has joint density

$$f_{X_1,\ldots,X_n}(x_1, \ldots, x_n) = g(r), \qquad r = \sqrt{x_1^2 + \ldots + x_n^2}\,,$$

which depends only on the distance to the origin through some non-negative function g.

Given the linear model

$$\mathbf{y}_n = \mathbf{X}_n \boldsymbol{\beta} + \mathbf{e}_n\,, \tag{1.1}$$

where the random vector $\mathbf{e}_n := (e_1, \ldots, e_n)$ has radial symmetry, we will study, as mentioned above, the consistency of LSE estimator of the vector of (unknown) parameters

$$\boldsymbol{\beta} := (\beta_1, \ldots, \beta_\kappa)\,.$$

2. Notations and preliminaries

Let us now recall some relevant notations and results. The *spectral radius* of $\mathbf{A} \in \mathcal{M}_\kappa(\mathbb{R})$ is defined by $\rho(\mathbf{A}) := \sup\big\{\, |\lambda| :\ \lambda \in \mathrm{Spec}(\mathbf{A})\big\}$ where $\mathrm{Spec}(\mathbf{A})$ is the *spectrum* of $\mathbf{A}$ and the transpose matrix of $\mathbf{A}$ will be $\mathbf{A}^T$. When there is no ambiguity, we will write

$$\rho_n := \rho\Big((\mathbf{X}_n^T \mathbf{X}_n)^{-1}\Big)\,.$$

Moreover, F_X (respectively $F_{X_1,\dots,X_n}$) will be the distribution function of a given random variable X (respectively joint distribution function of the random vector $(X_1, \dots, X_n)$), f_X (respectively $f_{X_1,\dots,X_n}$) its probability density function (respectively joint probability density function), the symbol $\sim$ will be used to indicate distributed as, $\approx$ will mean asymptotically (or approximately) equal to, Ω_n will be the range space of $\mathbf{X}_n$ and $\mathbf{P}_{\Omega_n}\mathbf{e}_n$ (respectively $\mathbf{P}_{\Omega_n^\perp}\mathbf{e}_n$) the orthogonal projection of $\mathbf{e}_n$ on Ω_n (respectively on $\Omega_n^\perp$).

Our main purpose will be the study of the convergence of LSE using a geometrical approach, assuming the error vector to have radial symmetry. Thus it will be quite natural to avail ourselves of generalized polar coordinates. The transformation from cartesian to the new coordinates corresponds to the mapping,

$$\mathbb{R}^n \ni (e_1, \dots, e_n) \mapsto (r, \theta_1, \dots, \theta_{n-1}) \in]0, +\infty[\times]0, \pi[\times \dots \times]0, \pi[\times]0, 2\pi[$$

defined by

$$\begin{cases} e_1 = r\cos\theta_1 \\ e_2 = r\sin\theta_1\cos\theta_2 \\ \qquad\vdots \\ e_{n-1} = r\sin\theta_1\sin\theta_2\dots\sin\theta_{n-2}\cos\theta_{n-1} \\ e_n = r\sin\theta_1\sin\theta_2\dots\sin\theta_{n-1} \end{cases} .$$

This mapping has the jacobian

$$J = r^{n-1}(\sin\theta_1)^{n-2}(\sin\theta_2)^{n-3}\dots\sin\theta_{n-2} \geqslant 0 .$$

We now have a new pair of random variables $(R_n, \mathbf{\Theta}_{n-1})$ with

$$R_n := \sqrt{e_1^2 + \dots + e_n^2}$$

and $\mathbf{\Theta}_{n-1} := (\Theta_1, \dots, \Theta_{n-1})$ the vector of central angles. The joint density of $(R_n, \mathbf{\Theta}_{n-1})$ will be given by

$$
\begin{aligned}
&f_{R_n,\boldsymbol{\Theta}_{n-1}}(r,\theta_1,\ldots,\theta_{n-1}) = \\
&\qquad = g(r)r^{n-1}(\sin\theta_1)^{n-2}(\sin\theta_2)^{n-3}\ldots\sin\theta_{n-2}\,.
\end{aligned}
\tag{2.1}
$$

Moreover, integrating this joint density on $]0,\pi[\times\ldots\times]0,\pi[\times]0,2\pi[$ in order to $\theta_1,\ldots,\theta_{n-1}$ leads to the probability density function of R_n

$$
f_{R_n}(r) = \frac{n\pi^{n/2}}{\Gamma\left(\frac{n}{2}+1\right)}\,g(r)r^{n-1}, \quad r>0\,. \tag{2.2}
$$

Likewise, if we integrate the joint density on $]0,+\infty[$ in order to r then the joint probability density function of $\boldsymbol{\Theta}_{n-1}$ will be

$$
f_{\boldsymbol{\Theta}_{n-1}}(\theta_1,\ldots,\theta_{n-1}) = \frac{\Gamma\left(\frac{n}{2}+1\right)}{n\,\pi^{n/2}}\,(\sin\theta_1)^{n-2}(\sin\theta_2)^{n-3}\ldots\sin\theta_{n-2} \tag{2.3}
$$

which does not depend on the real function g. We extract now an important result.

Proposition 2.1. *The random variables $R_n,\Theta_1,\ldots,\Theta_{n-1}$ are (mutually) independent.*

Proof. From (2.3) it follows that the densities of the angles $\Theta_1,\ldots,\Theta_{n-1}$ are,

$$
f_{\Theta_1}(x) = \frac{(\sin x)^{n-2}}{\displaystyle\int_0^\pi (\sin t)^{n-2}\,dt}, \quad 0<x<\pi\,,
$$

$$
f_{\Theta_2}(x) = \frac{(\sin x)^{n-3}}{\displaystyle\int_0^\pi (\sin t)^{n-3}\,dt}, \quad 0<x<\pi\,,
$$

$$
\vdots
$$

$$
f_{\Theta_{n-2}}(x) = \frac{\sin x}{2}, \quad 0<x<\pi\,,
$$

$$
f_{\Theta_{n-1}}(x) = \frac{1}{2\pi}, \quad 0<x<2\pi
$$

and thus

$$f_{\boldsymbol{\Theta}_{n-1}}(x_1, \ldots, x_{n-1}) = f_{\Theta_1}(x_1) \ldots f_{\Theta_{n-1}}(x_{n-1}) .$$

Hence

$$f_{R_n,\boldsymbol{\Theta}_{n-1}}(r, x_1, \ldots, x_{n-1}) = f_{R_n}(r) \cdot f_{\boldsymbol{\Theta}_{n-1}}(x_1, \ldots, x_{n-1}) =$$
$$= f_{R_n}(r) f_{\Theta_1}(x_1) \ldots f_{\Theta_{n-1}}(x_{n-1}) ,$$

which proves the (mutual) independence of $R_n, \Theta_1, \ldots, \Theta_{n-1}$. ∎

The factorization,

$$\|\mathbf{P}_{\Omega_n} \mathbf{e}_n\|^2 = \frac{\|\mathbf{P}_{\Omega_n} \mathbf{e}_n\|^2}{\|\mathbf{e}_n\|^2} \, \|\mathbf{e}_n\|^2 = Z_n \, \|\mathbf{e}_n\|^2 = Z_n \, R_n^2 , \tag{2.4}$$

where

$$Z_n := \frac{\|\mathbf{P}_{\Omega_n} \mathbf{e}_n\|^2}{\|\mathbf{e}_n\|^2} \tag{2.5}$$

gives us a first result about this random variable.

Lemma 2.1. *The random variable Z_n is bounded.*

Proof. By Pitagoras formulae (see [13]) we have

$$0 \leqslant Z_n := \frac{\left\|\mathbf{P}_{\Omega_n} \mathbf{e}_n\right\|^2}{\left\|\mathbf{e}_n\right\|^2} = \frac{\left\|\mathbf{P}_{\Omega_n} \mathbf{e}_n\right\|^2}{\left\|\mathbf{P}_{\Omega_n} \mathbf{e}_n\right\|^2 + \left\|\mathbf{P}_{\Omega_n^{\perp}} \mathbf{e}_n\right\|^2} \leqslant 1 ,$$

which establishes the thesis. ∎

When there is not multicollinearity, the LSE $\widetilde{\boldsymbol{\beta}}$ of $\boldsymbol{\beta}$ is given by:

$$\widetilde{\boldsymbol{\beta}} = (\mathbf{X}_n^T \mathbf{X}_n)^{-1} \mathbf{X}_n^T \, \mathbf{y}_n = \boldsymbol{\beta} + (\mathbf{X}_n^T \mathbf{X}_n)^{-1} \mathbf{X}_n^T \, \mathbf{e}_n .$$

We now present an upper bound for the LSE error

Lemma 2.2. *In the linear model* (1.1) *we have*

$$\left\|\widetilde{\boldsymbol{\beta}} - \boldsymbol{\beta}\right\|^2 \leqslant \rho_n \, R_n^2 \, Z_n \quad \text{a.s.} \tag{2.6}$$

Proof. We have, see [16],

$$\left\|\widetilde{\boldsymbol{\beta}} - \boldsymbol{\beta}\right\|^2 \leqslant \rho_n \, \|\mathbf{P}_{\Omega_n} \mathbf{e}_n\|^2 \quad \text{a.s.}$$

so the thesis follows from factorization (2.4) . ■

An fundamental result for the last section are annunciated above.

Proposition 2.2. *There exist an orthonormal basis* $\{\mathbf{w}_1(n), \ldots, \mathbf{w}_n(n)\}$ *of* $\mathbb{R}^n$ *such that*

$$Z_n = \big\langle \mathbf{w}_1(n), (\cos\Theta_1, \sin\Theta_1\cos\Theta_2, \ldots, \sin\Theta_1 \ldots \sin\Theta_{n-1})\big\rangle^2 +$$

$$+ \ldots + \big\langle \mathbf{w}_\kappa(n), (\cos\Theta_1, \sin\Theta_1\cos\Theta_2, \ldots, \sin\Theta_1 \ldots \sin\Theta_{n-1})\big\rangle^2, \; n \geqslant \kappa \,,$$

where $\langle \, \cdot \, , \, \cdot \, \rangle$ *is the usual inner product defined on the vector space* $\mathbb{R}^n$*.*

Proof. Since the design matrix $\mathbf{X}_n := \big[x_{ij}\big]_{\substack{i=1,2,\ldots,n \\ j=1,2,\ldots,\kappa}}$ has rank κ let us consider a basis of Ω_n given by

$$\mathbf{x}_1 := (x_{11}, x_{21}, \ldots, x_{n1}) \,,$$
$$\vdots$$
$$\mathbf{x}_\kappa := (x_{1\kappa}, x_{2\kappa}, \ldots, x_{n\kappa}) \,.$$

Using the Gram-Schmidt orthogonalization we can construct a orthonormal basis $\{\mathbf{w}_1(n), \ldots, \mathbf{w}_\kappa(n)\}$ of Ω_n and it is well known that the orthogonal complement $\Omega_n^\perp$ of Ω_n will admit a orthonormal basis $\{\mathbf{w}_{\kappa+1}(n), \ldots, \mathbf{w}_n(n)\}$. Supposing

$$\mathbf{w}_1(n) := \big(w_{11}(n), w_{21}(n), \ldots, w_{n1}(n)\big) \,,$$
$$\vdots$$
$$\mathbf{w}_n(n) := \big(w_{1\kappa}(n), w_{2\kappa}(n), \ldots, w_{n\kappa}(n)\big) \,.$$

We can take the matrix

$$\mathbf{W}_n(n) = \begin{bmatrix} w_{11}(n) & w_{12}(n) & \dots & w_{1n}(n) \\ w_{21}(n) & w_{22}(n) & \dots & w_{2n}(n) \\ \vdots & \vdots & \ddots & \vdots \\ w_{n1}(n) & w_{n2}(n) & \dots & w_{nn}(n) \end{bmatrix}$$

and the error vector $\mathbf{e}_n$ can be expressed on the basis $\{\mathbf{w}_1(n), \dots, \mathbf{w}_n(n)\}$ by $\mathbf{e}'_n = \left(\mathbf{W}_n(n)\right)^T \mathbf{e}_n$. Therefore

$$\|\mathbf{P}_{\Omega_n}\mathbf{e}_n\|^2 = \langle \mathbf{w}_1(n), \mathbf{e}_n\rangle^2 + \ldots + \langle \mathbf{w}_\kappa(n), \mathbf{e}_n\rangle^2$$

and the conclusion follows from the generalized polar coordinates. ∎

Remark 1. Let us observe that from Proposition 2.1 we can conclude the independence of R_n and Z_n since Z_n only depends of $\Theta_1, \dots, \Theta_{n-1}$ by the last Proposition 2.2.

Let us consider, on a probability space $(\Sigma, \mathcal{F}, \mathbb{P})$, a sequence $\{X_n\}$ of random variables and a random variable X. Given $p > 0$ we write:

1. $X_n \xrightarrow{\text{a.s.}} X$ if X_n *converges almost surely* to X;

2. $X_n \xrightarrow{\mathbb{P}} X$ if X_n *converges in probability* to X;

3. $X_n \xrightarrow{\mathscr{L}_p} X$ if X_n *converges in mean of order* p to X.

Let X_n and X be random variables with distribution functions F_n and F, respectively. If

$$\lim_{n\to+\infty} F_n(x) = F(x)$$

for every continuity point x of F, then X_n is said to *converge in distribution* or *in law* to X, and we write $X_n \xrightarrow{\text{d}} X$.[1]

[1] Clearly, convergence in distribution is a property of the distribution functions of the random variables and not of the random variables themselves. Recall that the random variables X_n may be defined on entirely different probability spaces. Moreover, given a distribution function F there always exists, on some probability space, a random variable X for which $F(x) = \mathbb{P}(X \leqslant x)$ (see [1]).

Let $(X_1, \dots, X_n)$ be a random vector induced by n observations $x_1, \dots, x_n$, which has joint distribution function depending, among others, on a vector of (unknown) parameters $\boldsymbol{\lambda}$ belonging to a parameter space $\boldsymbol{\Lambda} \subseteq \mathbb{R}^\kappa$. The estimator $\mathbf{t}_n := \mathbf{t}_n(X_1, \dots, X_n)$ will be called *strongly consistent* for $\boldsymbol{\lambda}$ if $\mathbf{t}_n \xrightarrow{\text{a.s.}} \boldsymbol{\lambda}$ for each fixed $\boldsymbol{\lambda} \in \boldsymbol{\Lambda}$. Given $s > 0$ the estimator $\mathbf{t}_n := \mathbf{t}_n(X_1, \dots, X_n)$ will be called *consistent in mean of order* s for $\boldsymbol{\lambda}$, if $\mathbf{t}_n \xrightarrow{\mathscr{L}_s} \boldsymbol{\lambda}$ for each fixed $\boldsymbol{\lambda} \in \boldsymbol{\Lambda}$. Convergence in mean of order 2 will be called *consistent in mean square.*

3. The distribution of Z_n

On the previous section, we showed that if $\mathbf{e}_n$ had radial symmetry we could replace it by the pair $(R_n, \boldsymbol{\Theta}_{n-1})$ with joint density given by (2.1). Now Z_n only depends on $\boldsymbol{\Theta}_{n-1}$ so that its density will not depend on the real function g. From Proposition 2.2 we get,

$$\begin{aligned} Z_n = &\big\langle \mathbf{w}_1(n), (\cos\Theta_1, \sin\Theta_1\cos\Theta_2, \dots, \sin\Theta_1 \dots \sin\Theta_{n-1})\big\rangle^2 + \\ &+ \dots + \big\langle \mathbf{w}_\kappa(n), (\cos\Theta_1, \sin\Theta_1\cos\Theta_2, \dots, \sin\Theta_1 \dots \sin\Theta_{n-1})\big\rangle^2 \end{aligned}$$

for some orthonormal basis $\big\{\mathbf{w}_1(n), \dots, \mathbf{w}_\kappa(n)\big\}$ of Ω_n. Thus, to obtain the probability density function of Z_n it can be assumed that $g(r)$ is whatever non-negative function. Choosing

$$g(r) = \frac{1}{(2\pi)^{n/2}}\, e^{-\frac{r^2}{2}} ,$$

we get for $\mathbf{e}_n$ the joint density

$$f_{e_1,\dots,e_n}(x_1, \dots, x_n) = \frac{1}{(2\pi)^{n/2}}\, e^{-\frac{x_1^2+\dots+x_n^2}{2}} ,$$

which is the standard multinormal distribution, i.e. $\mathbf{e}_n \sim N(\mathbf{0}, \mathbf{I})$. Hence (see [22]) the components $e_1, \dots, e_n$ are independent having each of them (univariate) standard normal distribution. By Cochran theorem (see [8]) the random variables $\big\|\mathbf{P}_{\Omega_n}\mathbf{e}_n\big\|^2$, $\big\|\mathbf{P}_{\Omega_n^\perp}\mathbf{e}_n\big\|^2$ are independent and

$$\big\|\mathbf{P}_{\Omega_n}\mathbf{e}_n\big\|^2 \sim \chi^2(\kappa), \qquad \big\|\mathbf{P}_{\Omega_n^\perp}\mathbf{e}_n\big\|^2 \sim \chi^2(n-\kappa) .$$

We now establish

Proposition 3.1. *Let $X_1, \dots, X_m$ be independent random variables with densities $f_{X_i}(x_i)$ $(i = 1, \dots, m)$, and $Y_1, \dots, Y_m$ random variables given by*

$$Y_j := X_1 + \ldots + X_j, \qquad j = 1, \dots, m\,.$$

Then $(Y_1, \dots, Y_m)$ has joint probability density function given by

$$f_{Y_1,\dots,Y_m}(y_1, \dots, y_m) = f_{X_1}(y_1)\, f_{X_2}(y_2 - y_1) \dots f_{X_m}(y_m - y_{m-1})\,.$$

Proof. The system of m linear equations in m unknowns $x_1, \dots, x_m$

$$\begin{cases} x_1 = y_1 \\ x_1 + x_2 = y_2 \\ \qquad \vdots \\ x_1 + x_2 + \ldots + x_m = y_m \end{cases}$$

has an unique solution given by,

$$\begin{cases} x_1 = y_1 \\ x_2 = y_2 - y_1 \\ \qquad \vdots \\ x_m = y_m - y_{m-1} \end{cases} .$$

Hence, the joint probability density function of $(Y_1, \dots, Y_m)$ is expressed by

$$\begin{aligned} f_{Y_1,\dots,Y_m}(y_1, \dots, y_m) =& \frac{f_{X_1,\dots,X_m}(x_1, \dots, x_m)}{|J(x_1, \dots, x_m)|} \\ =& f_{X_1}(y_1) \cdot f_{X_2}(y_2 - y_1) \dots f_{X_m}(y_m - y_{m-1}) \end{aligned}$$

since the transformation

$$\begin{cases} \varphi_1(x_1,\dots,x_m) = x_1 \\ \varphi_2(x_1,\dots,x_m) = x_1 + x_2 \\ \qquad\vdots \\ \varphi_n(x_1,\dots,x_m) = x_1 + x_2 + \dots + x_m \end{cases}$$

has jacobian

$$J(x_1,\dots,x_m) = \det\left(\begin{bmatrix} \dfrac{\partial\varphi_1}{\partial x_1} & \cdots & \dfrac{\partial\varphi_1}{\partial x_m} \\ \vdots & \ddots & \vdots \\ \dfrac{\partial\varphi_m}{\partial x_1} & \cdots & V\frac{\partial\varphi_m}{\partial x_m} \end{bmatrix}\right) = \det\left(\begin{bmatrix} 1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 1 & \dots & 1 \end{bmatrix}\right) = 1.$$

■

Setting $X_1 := \left\|\mathbf{P}_{\Omega_n}\mathbf{e}_n\right\|^2$, $X_2 := \left\|\mathbf{P}_{\Omega_n^\perp}\mathbf{e}_n\right\|^2$, $Y_1 := \left\|\mathbf{P}_{\Omega_n}\mathbf{e}_n\right\|^2$ and $Y_2 := \left\|\mathbf{e}_n\right\|^2$ the pair of random variables (Y_1, Y_2) has, according to Proposition 3.1, the joint density

$$f_{Y_1,Y_2}(x,y) = f_{X_1}(x)\, f_{X_2}(y-x)\ .$$

Hence, the density of Z_n will be

$$\begin{aligned} f_{Z_n}(z) &= \int_{-\infty}^{\infty} |\,t\,|\ f_{Y_1,Y_2}(zt,t)\,dt \\ &= \int_{-\infty}^{\infty} |\,t\,|\ f_{X_1}(zt)\, f_{X_2}(t-zt)\,dt \\ &= \frac{1}{\Gamma\left(\dfrac{\kappa}{2}\right)\,\Gamma\left(\dfrac{n-\kappa}{2}\right)} z^{\frac{\kappa}{2}-1}(1-z)^{\frac{n-\kappa}{2}-1} \int_0^\infty \frac{1}{2^{\frac{n}{2}}}\, t^{\frac{n}{2}-1} e^{-\frac{t}{2}}\,dt \\ &= \frac{\Gamma\left(\dfrac{n}{2}\right)}{\Gamma\left(\dfrac{\kappa}{2}\right)\,\Gamma\left(\dfrac{n-\kappa}{2}\right)} z^{\frac{\kappa}{2}-1}(1-z)^{\frac{n-\kappa}{2}-1}\,, \end{aligned}$$

if $0 < z < 1$ and $f_{Z_n}(z) = 0$ otherwise. Therefore, the random variable Z_n has distribution beta with parameters $\left(\frac{\kappa}{2}, \frac{n-\kappa}{2}\right)$.

We could have obtained this last result through a different approach in which Proposition 3.1 is not used. The random variable W_n defined by

$$W_n := \frac{\dfrac{\|\mathbf{P}_{\Omega_n}\mathbf{e}_n\|^2}{\kappa}}{\dfrac{\left\|\mathbf{P}_{\Omega_n^\perp}\mathbf{e}_n\right\|^2}{n-\kappa}}$$

has F distribution with κ and $(n-\kappa)$ degrees of freedom i.e.

$$f_{W_n}(w) =$$

$$= \frac{\Gamma\left(\frac{n}{2}\right)}{\Gamma\left(\frac{\kappa}{2}\right)\Gamma\left(\frac{n-\kappa}{2}\right)} \left(\frac{\kappa}{n-\kappa}\right)^{\frac{\kappa}{2}} w^{\frac{\kappa}{2}-1} \left(1+\frac{\kappa w}{n-\kappa}\right)^{-\frac{n}{2}}, \quad w > 0\,.$$

Therefore, the probability density function of the random variable $V_n := \frac{\kappa}{n-\kappa} W_n$ will be

$$f_{V_n}(v) = \frac{\Gamma\left(\frac{n}{2}\right)}{\Gamma\left(\frac{\kappa}{2}\right)\Gamma\left(\frac{n-\kappa}{2}\right)} v^{\frac{\kappa}{2}-1} (1+v)^{-\frac{n}{2}}, \quad v > 0\,.$$

Applying the transformation $u = \frac{v}{v+1}$ we get the integral identity

$$\int_t^{\infty} v^{a-1} (1+v)^{-a-b}\, dv = \int_{\frac{t}{t+1}}^{1} u^{a-1} (1-u)^{b-1}\, du, \quad t > 0\,.$$

Thus, taking $a = \frac{\kappa}{2}$ and $b = \frac{n-\kappa}{2}$ we get

$$F_{V_n}(t) = F_{U_n}\left(\frac{t}{t+1}\right), \quad t > 0\,, \tag{3.1}$$

where U_n has the beta distribution with parameters $\left(\frac{\kappa}{2}, \frac{n-\kappa}{2}\right)$.
Moreover, since

$$Z_n \leqslant V_n \tag{3.2}$$

we will have

$$\mathbb{P}(Z_n > t) \leqslant \mathbb{P}\left(V_n > t\right) = \mathbb{P}\left(U_n > \frac{t}{t+1}\right), \qquad t > 0 \tag{3.3}$$

and we can use the distribution of U_n to obtain upper bounds for $\mathbb{P}(Z_n > t)$.

4. Estimator consistency

The main purpose of this section is to establish the consistency of LSE. Nevertheless, we will start with some preparatory results for the LSE strong consistency.

Lemma 4.1. *For any $0 \leqslant \alpha < 1$ there exists $n_0 \in \mathbb{N}$ such that*

$$\mathbb{P}\left(Z_n > \frac{\varepsilon}{n^\alpha}\right) < \frac{\Gamma\left(\frac{n}{2}\right)}{\Gamma\left(\frac{\kappa}{2}\right)\Gamma\left(\frac{n-\kappa}{2}\right)}\left(\frac{\varepsilon}{n^\alpha}\right)^{\frac{\kappa}{2}-1}\left(1-\frac{\varepsilon}{n^\alpha}\right)^{\frac{n-\kappa}{2}} \qquad (\varepsilon > 0)$$

for all $n \geqslant n_0$.

Proof. If $\kappa \geqslant 2$ and $n > \kappa + 2$ then Z_n has a unique mode (see [10]) given by

$$z_0 = \frac{\kappa - 2}{n - 4}$$

(if $\kappa = 1$ and $n > 3$ then $f_{Z_n}(z)$ is monotonically decreasing on $]0,1[$). For all $0 < \alpha < 1$ we have

$$\lim_{n\to+\infty} \frac{\dfrac{\kappa-2}{n-4}}{\dfrac{\varepsilon}{n^\alpha}} = 0, \qquad \varepsilon > 0\,,$$

which implies the existence of an order $n_0 \in \mathbb{N}$ such that

$$z_0 = \frac{\kappa-2}{n-4} < \frac{\varepsilon}{n^\alpha}, \qquad \forall n \geqslant n_0, \quad \varepsilon > 0\;.$$

Since $f_{Z_n}(z)$ is monotonically decreasing on $[z_0, 1[$ we can write

$$\mathbb{P}\left(Z_n > \frac{\varepsilon}{n^\alpha}\right) < \left(1 - \frac{\varepsilon}{n^\alpha}\right) f_{Z_n}\left(\frac{\varepsilon}{n^\alpha}\right), \quad \forall n \geqslant n_0, \quad \varepsilon > 0\,,$$

completing the Lemma proof. ■

Lemma 4.2. *For any* $0 < \alpha < 1$ *we have*

$$\lim_{n\to+\infty} n^\xi \left(1 - \frac{\varepsilon}{n^\alpha}\right)^{\frac{n-\kappa}{2}} = 0, \qquad \xi \in \mathbb{R}, \quad \varepsilon > 0\,.$$

Proof. Choosing $\xi \in \mathbb{R}$ we have for all $\varepsilon > 0$

$$\begin{aligned} n^\xi \left(1 - \frac{\varepsilon}{n^\alpha}\right)^{\frac{n-\kappa}{2}} &= e^{\log\left(n^\xi \left(1-\frac{\varepsilon}{n^\alpha}\right)^{\frac{n}{2}}\right)} \left(1 - \frac{\varepsilon}{n^\alpha}\right)^{-\frac{\kappa}{2}} \\ &= e^{\xi \log n + \frac{n^{1-\alpha}}{2} \log\left(1-\frac{\varepsilon}{n^\alpha}\right)^{n^\alpha}} \left(1 - \frac{\varepsilon}{n^\alpha}\right)^{-\frac{\kappa}{2}} \\ &= e^{n^{1-\alpha}\left(\frac{\xi \log n}{n^{1-\alpha}} + \frac{1}{2} \log\left(1-\frac{\varepsilon}{n^\alpha}\right)^{n^\alpha}\right)} \left(1 - \frac{\varepsilon}{n^\alpha}\right)^{-\frac{\kappa}{2}}\;. \end{aligned}$$

Since $0 < \alpha < 1$ it follows

$$\lim_{n\to+\infty} \log\left(1-\frac{\varepsilon}{n^\alpha}\right)^{n^\alpha} = -\varepsilon \qquad \text{and} \qquad \lim_{n\to+\infty} \frac{\xi \log n}{n^{1-\alpha}} = 0$$

so that,

$$\lim_{n\to+\infty} n^\xi \left(1-\frac{\varepsilon}{n^\alpha}\right)^{\frac{n-\kappa}{2}} = 0\,.$$

■

Nextly, we present two important results with stringent and direct influence on the LSE strong consistency.

Proposition 4.1. *For any $0 < \alpha < 1$ we have*

$$\mathbb{P}\left(\limsup_{n\to+\infty} \left\{Z_n > \frac{1}{m\,n^\alpha}\right\}\right) = 0, \qquad m = 1, 2, \ldots$$

Proof. Since

$$\lim_{n\to+\infty} \frac{\Gamma\left(\frac{n}{2}\right)}{\Gamma\left(\frac{n-\kappa}{2}\right) n^{\frac{\kappa}{2}}} = 2^{-\frac{\kappa}{2}}$$

(see [14]) we obtain from Lemma 4.2, with $\xi = \alpha\left(1-\frac{\kappa}{2}\right) + s$ $(s > 1)$ and $\varepsilon = \frac{1}{m}$ $(m = 1, 2, \ldots)$

$$\lim_{n\to+\infty} \frac{\dfrac{\Gamma\left(\frac{n}{2}\right)}{\Gamma\left(\frac{\kappa}{2}\right)\Gamma\left(\frac{n-\kappa}{2}\right)} \left(\dfrac{1}{m\,n^\alpha}\right)^{\frac{\kappa}{2}-1} \left(1-\dfrac{1}{m\,n^\alpha}\right)^{\frac{n-\kappa}{2}}}{\dfrac{1}{n^s}} =$$

$$= \frac{1}{m^{\frac{\kappa}{2}-1}\,\Gamma\left(\frac{\kappa}{2}\right) 2^{\frac{\kappa}{2}}} \lim_{n\to+\infty} n^{\alpha+\frac{\kappa}{2}(1-\alpha)+s} \left(1-\frac{1}{m\,n^\alpha}\right)^{\frac{n-\kappa}{2}} = 0\,.$$

Thus, Lemma 4.1 ensures

$$\lim_{n\to+\infty}\frac{\mathbb{P}\left(Z_n>\dfrac{1}{m\,n^{\alpha}}\right)}{\dfrac{1}{n^s}}\leqslant$$

$$\leqslant \lim_{n\to+\infty}\frac{\dfrac{\Gamma\left(\frac{n}{2}\right)}{\Gamma\left(\frac{\kappa}{2}\right)\Gamma\left(\frac{n-\kappa}{2}\right)}\left(\dfrac{1}{m\,n^{\alpha}}\right)^{\frac{\kappa}{2}-1}\left(1-\dfrac{1}{m\,n^{\alpha}}\right)^{\frac{n-\kappa}{2}}}{\dfrac{1}{n^s}}=0$$

and consequently,

$$\sum_{n=1}^{+\infty}\mathbb{P}\left(Z_n>\frac{1}{m\,n^{\alpha}}\right)<+\infty,\qquad m=1,2,\dots$$

provided that series $\sum_{n=1}^{+\infty}\frac{1}{n^s}$ $(s>1)$ converges. The thesis now follows from the First Borel-Cantelli Lemma. ■

Remark 2. If we take the alternative path described at the end of the last section it is clear that, for any $0<\alpha<1$, there exits an order $n_0\in\mathbb{N}$ such that,

$$\mathbb{P}\left(U_n>\frac{\varepsilon}{n^{\alpha}+\varepsilon}\right)<\frac{\Gamma\left(\frac{n}{2}\right)}{\Gamma\left(\frac{\kappa}{2}\right)\Gamma\left(\frac{n-\kappa}{2}\right)}\left(\frac{\varepsilon}{n^{\alpha}}\right)^{\frac{\kappa}{2}-1}\left(1-\frac{\varepsilon}{n^{\alpha}+\varepsilon}\right)^{\frac{n-\kappa}{2}}\quad(\varepsilon>0)$$

for all $n\geqslant n_0$. Moreover,

$$\lim_{n\to+\infty}n^{\xi}\left(1-\frac{\varepsilon}{n^{\alpha}+\varepsilon}\right)^{\frac{n-\kappa}{2}}=0,\qquad \xi\in\mathbb{R},\quad \varepsilon>0$$

and the thesis of Proposition 4.1 follows from (3.2).

Proposition 4.2. *For any* $0 < \alpha < 1$ *we have,*

$$n^{\alpha} Z_n \xrightarrow[n \to +\infty]{\text{a.s.}} 0 \, .$$

Proof. According to Proposition 4.1 we get,

$$\mathbb{P}\left(\limsup_{n \to +\infty} \left\{ |n^{\alpha} Z_n| > \frac{1}{m} \right\} \right) = \mathbb{P}\left(\limsup_{n \to +\infty} \left\{ Z_n > \frac{1}{m\, n^{\alpha}} \right\} \right) = 0$$

for all $m = 1, 2, \ldots$ so that (see [6])

$$\mathbb{P}\left(\lim_{n \to +\infty} n^{\alpha} Z_n = 0 \right) = 1 \, .$$

■

Let us present now the first results on strong consistency of LSE.

Theorem 4.1. *If* $\mathbf{e}_n$ *has radial symmetry and for some* $0 < \alpha < 1$,

$$\exists K > 0 \colon \quad \limsup_{n \to +\infty} \left(n^{-\alpha} \rho_n \, R_n^2 \right) \leqslant K \quad a.s., \tag{4.1}$$

then $\widetilde{\boldsymbol{\beta}}$ *is strongly consistent.*

Proof. The thesis follows from estimate (2.6) and Proposition 4.2. ■

Corollary 4.1. *If* $\mathbf{e}_n$ *has radial symmetry,* $R_n \xrightarrow{\text{a.s.}} R_\infty$ *and* $\rho_n = O(n^{\alpha})$ *for some* $0 < \alpha < 1$ *then* $\widetilde{\boldsymbol{\beta}}$ *is strongly consistent.*

Proof. According to the Corollary's assumptions, the term $n^{-\alpha} \rho_n \, R_n^2$ is bounded almost surely and the thesis follows from Theorem 4.1 ■

The consistency in mean square of LSE can be obtained from the independence of R_n and Z_n (see Remark 1). We now consider assumptions on $\mathbb{E}\left(R_n^2\right)$. Firstly we get

Theorem 4.2. *If* $\mathbf{e}_n$ *has radial symmetry,* $\mathbb{E}\left(R_n^2\right) = O\left(n^{\alpha}\right)$ *and* $\rho_n = o\left(n^{1-\alpha}\right)$, *then* $\widetilde{\boldsymbol{\beta}}$ *is mean square consistent.*

Proof. According to estimate (2.6) we have

$$\mathbb{E}\left(\left\|\widetilde{\boldsymbol{\beta}} - \boldsymbol{\beta}\right\|^2\right) \leqslant \rho_n \, \mathbb{E}\left(R_n^2\right) \, \mathbb{E}\left(Z_n\right) \leqslant C \, \rho_n \, n^{\alpha} \, \frac{\kappa}{n} = \kappa \, C \cdot \frac{\rho_n}{n^{1-\alpha}} = o(1)$$

and so the thesis follows. ■

More generally, the independence of the random variables R_n and Z_n leads to s order mean consistency.

Theorem 4.3. *If* $\mathbf{e}_n$ *has radial symmetry,* $\mathbb{E}\left(R_n^s\right) = O\left(n^{\alpha}\right)$ *and* $\rho_n = o\left(n^{1-\frac{2\alpha}{s}}\right)$, *then* $\widetilde{\boldsymbol{\beta}}$ *is* s *order mean consistent.*

Proof. Given $s > 0$ the estimate (2.6) yields

$$\mathbb{E}\left(\left\|\widetilde{\boldsymbol{\beta}} - \boldsymbol{\beta}\right\|^s\right) \leqslant {\rho_n}^{\frac{s}{2}} \cdot \mathbb{E}\left(R_n^s\right) \, \mathbb{E}\left({Z_n}^{\frac{s}{2}}\right)$$

$$= {\rho_n}^{\frac{s}{2}} \cdot \mathbb{E}\left(R_n^s\right) \frac{\Gamma\left(\frac{n}{2}\right) \Gamma\left(\frac{s+\kappa}{2}\right)}{\Gamma\left(\frac{\kappa}{2}\right) \Gamma\left(\frac{n+s}{2}\right)} = o(1)\,,$$

since $\Gamma\left(\frac{n+s}{2}\right) \approx \left(\frac{n}{2}\right)^{s/2} \Gamma\left(\frac{n}{2}\right)$ for n large enough. ■

Remark 3. The LSE consistency in mean of order s still remains valid if $\mathbf{e}_n$ has radial symmetry and

$$\rho_n = o\left(\frac{n}{\left[\mathbb{E}\left(R_n^s\right)\right]^{2/s}}\right).$$

5. Extension to case $\alpha = 1$

On Section 3 we saw that, if the random error sequence had radial symmetry then the random variable Z_n will had distribution beta with parameters $\left(\frac{\kappa}{2}, \frac{n-\kappa}{2}\right)$. Therefore,

$$f_{nZ_n}(z) = \frac{\Gamma\left(\frac{n}{2}\right)}{\Gamma\left(\frac{\kappa}{2}\right)\Gamma\left(\frac{n-\kappa}{2}\right) n^{\frac{\kappa}{2}}} z^{\frac{\kappa}{2}-1}\left(1-\frac{z}{n}\right)^{\frac{n-\kappa}{2}-1}, \qquad 0 < z < n$$

and $f_{nZ_n}(z) = 0$ otherwise. From Proposition 2.2 we know that the random variable nZ_n can be expressed by

$$n\, Z_n = \left(S_{(n-1)1}(n)\right)^2 + \ldots + \left(S_{(n-1)\kappa}(n)\right)^2,$$

where for each $i = 1, \ldots, \kappa$,

$$\begin{aligned} S_{(n-1)i}(n) :=& \sqrt{n}\cos\Theta_1 w_{1i}(n) + \sqrt{n}\sin\Theta_1\cos\Theta_2 w_{2i}(n) + \\ &+ \ldots + \sqrt{n}\sin\Theta_1\ldots\sin\Theta_{n-2}\cos\Theta_{n-1}w_{(n-1)i}(n) + \\ &+ \sqrt{n}\sin\Theta_1\ldots\sin\Theta_{n-2}\sin\Theta_{n-1}w_{ni}(n)\ . \end{aligned}$$

Let us consider the triangular array $T_{mi}(n)$, $n \geqslant 2$, $1 \leqslant m \leqslant n-1$ of random variables defined by

$$\left\{\ T_{1i}(2) := \sqrt{2}\left[\cos\Theta_1 w_{1i}(2) + \sin\Theta_1 w_{2i}(2)\right]\right.$$

$$\begin{cases} T_{1i}(3) := \sqrt{3}\,\cos\Theta_1 w_{1i}(3) \\ T_{2i}(3) := \sqrt{3}\,\sin\Theta_1\left[\cos\Theta_2 w_{2i}(3) + \sin\Theta_2 w_{3i}(3)\right] \end{cases}$$

$$\vdots$$

$$
\begin{cases}
\vdots \\
T_{1i}(n) := \sqrt{n}\,\cos\Theta_1 w_{1i}(n) \\
T_{2i}(n) := \sqrt{n}\,\sin\Theta_1\cos\Theta_2 w_{2i}(n) \\
\vdots \\
T_{(n-2)i}(n) := \sqrt{n}\,\sin\Theta_1\ldots\sin\Theta_{n-3}\cos\Theta_{n-2}w_{(n-2)i}(n) \\
T_{(n-1)i}(n) := \sqrt{n}\,\sin\Theta_1\ldots\sin\Theta_{n-2}\big[\cos\Theta_{n-1}w_{(n-1)i}(n)+\sin\Theta_{n-1}w_{ni}(n)\big]
\end{cases}
$$

for each $i = 1, \ldots, \kappa$. Thus

$$
\begin{aligned}
S_{1i}(2) &= T_{1i}(2) \\
S_{2i}(3) &= T_{1i}(3) + T_{2i}(3) \\
&\vdots \\
S_{(n-1)i}(n) &= T_{1i}(n) + T_{2i}(n) + \ldots + T_{(n-1)i}(n)
\end{aligned}
$$

for each $i = 1, \ldots, \kappa$ and since

$$
\begin{aligned}
&\mathbb{E}\big[T_{1i}(2)\big] = 0 \quad \text{a.s.} \\
&\mathbb{E}\big[T_{2i}(3) \mid T_{1i}(3)\big] = \sqrt{3}\sin\Theta_1\big[w_{2i}(3)\,\mathbb{E}(\cos\Theta_2) + w_{3i}(3)\,\mathbb{E}(\sin\Theta_2)\big] = 0 \text{ a.s.} \\
&\qquad\vdots \\
&\mathbb{E}\big[T_{(n-1)i}(n) \mid T_{1i}(n), \ldots, T_{(n-2)i}(n)\big] = \\
&= \sqrt{n}\sin\Theta_1\ldots\sin\Theta_{n-2}\big[w_{(n-1)i}(n)\mathbb{E}(\cos\Theta_{n-1}) + w_{ni}(n)\mathbb{E}(\sin\Theta_{n-1})\big] = 0 \text{ a.s.},
\end{aligned}
$$

we can apply the Extended Kolmogorov Inequality and the Extended Bienaymé Equality (see [12]) to each term of the sequence $S_{(n-1)i}(n)$, $n \geqslant 2$, which give us[2]

$$\mathbb{P}\left\{\max_{1\leqslant j\leqslant 1}|S_{ji}(2)| \geqslant \varepsilon\right\} \leqslant \frac{1}{\varepsilon^2}\cdot\mathbb{E}\left[\Big(S_{1i}(2)\Big)^2\right] \leqslant \frac{\mathbb{E}\big(2\,Z_2\big)}{\varepsilon^2} = \frac{\kappa}{\varepsilon^2}$$

$$\mathbb{P}\left\{\max_{1\leqslant j\leqslant 2}|S_{ji}(3)| \geqslant \varepsilon\right\} \leqslant \frac{1}{\varepsilon^2}\cdot\mathbb{E}\left[\Big(S_{2i}(3)\Big)^2\right] \leqslant \frac{\mathbb{E}\big(3\,Z_3\big)}{\varepsilon^2} = \frac{\kappa}{\varepsilon^2}$$

$$\vdots$$

$$\mathbb{P}\left\{\max_{1\leqslant j\leqslant n-1}|S_{ji}(n)| \geqslant \varepsilon\right\} \leqslant \frac{1}{\varepsilon^2}\cdot\mathbb{E}\left[\Big(S_{(n-1)i}(n)\Big)^2\right] \leqslant \frac{\mathbb{E}\big(n\,Z_n\big)}{\varepsilon^2} = \frac{\kappa}{\varepsilon^2}$$

provided that $\mathbb{E}\big(nZ_n\big) = \kappa$, $\forall n \in \mathbb{N}$. Hence

$$\mathbb{P}\left\{\max_{1\leqslant j\leqslant n-1}|S_{ji}(n)| < \varepsilon\right\} \geqslant 1 - \frac{\kappa}{\varepsilon^2}, \quad \forall n \in \mathbb{N}$$

and choosing the subsequence (η_n) which give us

$$\lim_{n\to+\infty}\left(\max_{1\leqslant j\leqslant \eta_n-1}|S_{ji}(\eta_n)|\right) = \limsup_{n\to+\infty}\left(\max_{1\leqslant j\leqslant n-1}|S_{ji}(n)|\right),$$

it follows

$$\mathbb{P}\left\{\max_{1\leqslant j\leqslant \eta_n-1}|S_{ji}(\eta_n)| < \varepsilon\right\} \geqslant 1 - \frac{\kappa}{\varepsilon^2}, \quad \forall n \in \mathbb{N}\,.$$

[2]Observe that $\mathbb{E}S_{(n-1)i}(n) = 0$, $\forall n \geqslant 2$.

Thus, for n sufficiently large we get

$$\mathbb{P}\left\{\sup_{n\leqslant m}\left(\max_{1\leqslant j\leqslant n-1}|S_{ji}(m)|\right)<\varepsilon\right\}\geqslant 1-\frac{\kappa}{\varepsilon^2}$$

that is, the random sequence $\left\{S_{(n-1)i}(n)\right\}$, $i=1,\ldots,\kappa$, is bounded in probability. Therefore, almost surely

$$\limsup_{n\to+\infty}\, nZ_n$$

exists and is finite. On the other side, it is easy to check that the sequence of functions $f_{nZ_n}(z)$ converges pointwise on $\mathbb{R}$ as $n\to+\infty$ i.e.

$$f_{nZ_n}(z)\xrightarrow[n\to+\infty]{} f(z)=\begin{cases}\dfrac{1}{2^{\frac{\kappa}{2}}\,\Gamma\left(\dfrac{\kappa}{2}\right)}\,z^{\frac{\kappa}{2}-1}e^{-\frac{z}{2}} & \text{se}\quad z>0\\[2ex] 0 & \text{se}\quad z\leqslant 0\end{cases}.$$

Consequently

$$nZ_n\xrightarrow[n\to+\infty]{\mathrm{d}}\chi^2(\kappa) \tag{5.1}$$

that is, nZ_n converges in law to a random variable which has the χ^2 distribution with κ degrees of freedom, which implies $\limsup\limits_{n\to+\infty} nZ_n>0$ almost surely.

Natural improvements of the results of the last section follows now immediately

Theorem 5.1. *Let $p\geqslant 0$. If $\mathbf{e}_n$ has radial symmetry, $\rho_n=o\left(n^{-p}\right)$ and*

$$\exists K>0\colon\quad \limsup_{n\to+\infty}\frac{R_n^2}{n^{p+1}}\leqslant K\quad \text{a.s.}\,, \tag{5.2}$$

then $\widetilde{\boldsymbol{\beta}}$ is strongly consistent.

Proof. The condition (5.2) leads to

$$\rho_n\,R_n^2\,Z_n=\rho_n\,n^p\cdot\frac{R_n^2}{n^{p+1}}\cdot nZ_n\leqslant K\,C\,\rho_n\,n^p\quad\text{a.s.}$$

and the conclusion follows from the condition on spectral radius. ∎

Corollary 5.1. *Let $p \geqslant 0$. If* $\mathbf{e}_n$ *has radial symmetry,* $\dfrac{R_n}{n^p} \xrightarrow{\text{a.s.}} R_\infty$ *and* $\rho_n = o\left(n^{1-2p}\right)$, *then* $\widetilde{\boldsymbol{\beta}}$ *is strongly consistent.*

Proof. We have $\dfrac{R_n^2}{n^{2p}} \xrightarrow{\text{a.s.}} R_\infty^2$ (see [6]) and the thesis follows from Theorem 5.1. ■

Remark 4. If, see [22], we assume the radial symmetry of $\mathbf{e}_n$ with differentiable function g or continuous marginal densities and also the independency of $e_1, e_2, \dots,$ then each error has normal distribution with zero mean and equal variance.

6. EXAMPLES

To complete this geometrical approach to the consistency of LSE we present a few examples of application.

1. Multivariate normal distribution
Suppose

$$g(r) = \frac{1}{(2\pi\,\sigma^2)^{n/2}}\, e^{-\frac{r^2}{2\,\sigma^2}}, \qquad \sigma > 0\,,$$

then each component of the random vector $\mathbf{e}_n$ has distribution $N(0, \sigma^2)$. Hence, the sequence $e_1, e_2, \dots$ is i.i.d. (see [22]) and according to classical Kolmogorov's Strong Law of Large Numbers we get

$$\frac{R_n^2}{n} = \frac{e_1^2 + \ldots + e_n^2}{n} \xrightarrow[n\to+\infty]{\text{a.s.}} \sigma^2\,,$$

since $\mathbb{E}\left(e_i^2\right) = \sigma^2$ for all i. According to Corollary 5.1 the LSE is strongly consistent if $\rho_n = o(1)$. Moreover, R_n^2 has χ^2 distribution with n degrees of freedom so that $\mathbb{E}\left(R_n^2\right) = n$ and the LSE mean square consistency holds for $\rho_n = o(1)$.

2. Kotz type distribution
Let $a, b > 0$ and $n + 2q > 2$. If

$$g(r) = \frac{a\,\Gamma\left(\frac{n}{2}\right)}{b^{-\frac{2q+n-2}{2a}}\,\pi^{\frac{n}{2}}\,\Gamma\left(\frac{2q+n-2}{2a}\right)}\, r^{2q-2} e^{-b\,r^{2a}}\,,$$

then the random vector $\mathbf{e}_n$ has Kotz type distribution[3] (see [23]). For example, taking $q = a - \frac{n}{2} + 1$ we obtain

$$f_{R_n^2}(r) = \frac{1}{2\sqrt{r}}\, f_{R_n}\left(\sqrt{r}\right) = ab\, r^{a-1} e^{-b\,r^a}, \quad r > 0\,,$$

that is, R_n^2 has Weibull-Gnedenko distribution with parameters (a, b). Since

$$\mathbb{E}\big(R_n^2\big) = b^{-\frac{1}{a}}\,\Gamma\left(1 + \frac{1}{a}\right)\,,$$

the LSE mean square consistency holds if $\rho_n = o(n)$. The strong consistency of LSE holds if $\rho_n = o(n)$ since

$$\sup_{m>n} \mathbb{E}\left|R_m^2 - R_n^2\right| = \sup_{m>n} \Big(\mathbb{E}\big(R_m^2\big) - \mathbb{E}\big(R_n^2\big)\Big) = 0\,,$$

implies $R_n^2 \xrightarrow{\mathscr{L}_1} R_\infty$ for some $R_\infty \in \mathscr{L}_1$. Therefore, $R_n^2 \xrightarrow{\text{a.s.}} R_\infty$ since R_n^2 is an increasing sequence and $R_n^2 \xrightarrow{\mathbb{P}} R_\infty$.

3. Multivariate uniform distribution
Let $a > 0$ and consider

$$g(r) = \frac{\Gamma\left(\frac{n}{2} + 1\right)}{\left(a\sqrt{\pi}\right)^n}, \qquad r < a\,,$$

[3] Note that if $q = a = 1$ and $b = \frac{1}{2\sigma^2}$ $(\sigma > 0)$ then we recover the multivariate normal distribution of last example.

which corresponds to the situation where the random variable $\mathbf{e}_n$ is distributed uniformly on an open ball centered in origin with radius a. It is easy to check that R_n has density

$$f_{R_n}(r) = \frac{n}{a^n}\, r^{n-1}, \qquad 0 < r < a$$

and distribution

$$F_{R_n}(r) = \begin{cases} 0 & \text{when} \quad r \leqslant 0 \\ \left(\dfrac{r}{a}\right)^n & \text{when} \quad 0 < r < a \\ 1 & \text{when} \quad r \geqslant a \end{cases} .$$

Thus

$$F_{R_n}(r) \xrightarrow[n \to +\infty]{} F(r) = \begin{cases} 1 & \text{when} \quad r \geqslant a \\ 0 & \text{when} \quad r < a \end{cases} ,$$

and so $R_n \xrightarrow{\mathbb{P}} a$ which implies $R_n^2 \xrightarrow{\text{a.s.}} a^2$ since R_n^2 is an increasing sequence. From Corollary 5.1 the strong consistency of LSE is ensured if $\rho_n = o(n)$. On the other hand, the random variable R_n^2 has density,

$$f_{R_n^2}(r) = \frac{n}{2\,a^n}\, r^{\frac{n}{2}-1}, \qquad 0 < r < a^2 ,$$

which implies $\mathbb{E}\left(R_n^2\right) = \dfrac{n\,a^2}{n+2}$ and LSE mean square consistency holds if $\rho_n = o(n)$.

4. Multivariate *t* distribution

Let $q \in \mathbb{N}$ and consider

$$g(r) = \frac{\Gamma\left(\dfrac{n+q}{2}\right) n^{\frac{n}{2}}}{\Gamma\left(\dfrac{q}{2}\right) (q\pi)^{\frac{n}{2}}} \left(1 + \frac{n}{q}\, r^2\right)^{-\frac{n+q}{2}} ,$$

which corresponds to the situation in which $\mathbf{e}_n := (e_1, \dots, e_n)$ has multivariate t distribution with q degrees of freedom, precision matrix

$$\mathbf{T} = \mathrm{diag}\big(n, \dots, n\big)$$

and null skewness. The density of R_n^2 will be

$$f_{R_n^2}(r) = \frac{1}{2\sqrt{r}}\, f_{R_n}\left(\sqrt{r}\right) = \frac{n\,\pi^{n/2}}{2\,\Gamma\left(\dfrac{n}{2}+1\right)}\, g\left(\sqrt{r}\right)\, r^{\frac{n}{2}-1}, \qquad r > 0\,,$$

so that

$$f_{R_n^2}(r) = \frac{\Gamma\left(\dfrac{n+q}{2}\right)}{\Gamma\left(\dfrac{n}{2}\right)\,\Gamma\left(\dfrac{q}{2}\right)} \left(\frac{n}{q}\right)^{\frac{n}{2}} r^{\frac{n}{2}-1} \left(1 + \frac{n\,r}{q}\right)^{-\frac{n+q}{2}}, \qquad r > 0\,.$$

Therefore, R_n^2 has F distribution with degrees of freedom n and q, so

$$\mathbb{E}\big(R_n^2\big) = \frac{q}{q-2} \qquad (q > 2)\,.$$

Hence, LSE mean square consistency is guaranteed if $\rho_n = o(n)$. The strong consistency of LSE holds if $\rho_n = o(n)$: as a matter of fact

$$\sup_{m>n} \mathbb{E}\left|R_m^2 - R_n^2\right| = \sup_{m>n} \Big(\mathbb{E}\big(R_m^2\big) - \mathbb{E}\big(R_n^2\big)\Big) = 0$$

implies $R_n^2 \xrightarrow{\mathscr{L}_1} R_\infty$ for some $R_\infty \in \mathscr{L}_1$. Thus, $R_n^2 \xrightarrow{\text{a.s.}} R_\infty$ since R_n^2 is an increasing sequence and $R_n^2 \xrightarrow{\mathbb{P}} R_\infty$.

References

[1] P. Billingsley, *Probability and Measure*, (third edition), John Wiley & Sons 1995.

[2] S. Cambanis, S. Huang and G. Simons, *On the theory of elliptically contoured distributions*, Journal of Multivariate Analysis **11** (1981), 368–385.

[3] X. Chen, *Some results on consistency of LS estimates*, Chin. Sci. Bull. **39** (22) (1994), 1872–1876.

[4] X. Chen, *Consistency of LS estimates of multiple regression under a lower order moment condition*, Sci. Chin. **38** (12) (1995), 1420–1431.

[5] X. Chen, *A note on the consistency of LS estimates in linear models*, Chin. Ann. Math. Ser. B, **22** (4) (2001), 471–474.

[6] Y.S. Chow and H. Teicher, *Probability Theory: Independence, Interchangeability, Martingales* (third edition), Springer 1997.

[7] K.L. Chung, *A Course in Probability Theory* (third edition), Academic Press 2001.

[8] D. Dacunha-Castelle et M. Duflo, *Probabilités et Statistiques: Problèmes à Temps Fixe*, Masson 1982.

[9] K. Fang, S. Kotz and K. Ng, *Symmetric Multivariate and Related Distributions*, Monographs on Statistics and Applied Probability 36, Chapman & Hall 1990.

[10] V. Koroliouk, N. Portenko, A. Skorokhod and A. Tourbine, *Aide-Mémoire de Théorie des Probabilités et de Statistique Mathématique*. Mir. (1983).

[11] M. Loève, *Probability Theory I* (fourth edition), Springer 1977.

[12] M. Loève, *Probability Theory II* (fourth edition), Springer 1978.

[13] L.T. Magalhães, *Álgebra Linear como Introdução á Matemática Aplicada*, Texto Editora 1992.

[14] B.M. Makarov, M.G. Goluzina, A.A. Lodkin, and A.N. Podkorytov, *Selected Problems in Real Analysis*, American Mathematical Society 1992.

[15] J.T. Mexia, and P. Corte Real, *Extension of Kolmogorov's strong law to multiple regression*, Revista de Estatística, (2^{o} quadrimestre de 2001), **24** (2001), 277–278.

[16] J.T. Mexia, P. Corte Real, M.L. Esquível, e J. Lita da Silva, *Convergência do estimador dos mínimos quadrados em modelos lineares*, Estatística Jubilar. Actas do XII Congresso da Sociedade Portuguesa de Estatística, Edições SPE, (2005), 455–466.

[17] J.T. Mexia e J. Lita da Silva, *A consistęncia do estimador dos mínimos quadrados em domínios de atracção maximais*, (to appear) 2005.

[18] J.T. Mexia and J. Lita da Silva, *Least squares estimator consistency: on error stability*, (to appear) 2005.

[19] J. Mingzhong, *Some new results of the strong consistency of multiple regression coefficients*, Proceedings of the Second Asian Mathematical Conference 1995 (Tangmanee, S. & Schulz, E. eds.), World Scientific (1995), 514–519.

[20] J. Mingzhong and X. Chen, *Strong consistency of least squares estimate in multiple regression when the error variance is infinite*, Stat. Sin. **9** (1) (1999), 289–296.

[21] W. Pestman, *Mathematical Statistics*, Walter de Gruyter Berlin 1998.

[22] C.R. Rao, *Linear Statistical Inference and Its Applications*, (second edition), John Wiley & Sons (1973).

[23] R. Schmidt, *Tail dependence for elliptically contoured distributions*, Mathematical Methods of Operations Research **55** (2002), 301–327.

[24] D. Williams, *Probability with Martingales*, Cambridge University Press, Cambridge 1991.

Received 9 September 2005