

# On improved estimation in a conditionally Gaussian regression

Pchelintsev Evgeny\*

May 26, 2011

## Abstract

The paper considers the problem of estimating a  $p \geq 2$  dimensional mean vector of a multivariate conditionally normal distribution under quadratic loss. The problem of this type arises when estimating the parameters in a continuous time regression model with a non-Gaussian Ornstein–Uhlenbeck process driven by the mixture of a Brownian motion and a compound Poisson process. We propose a modification of the James–Stein procedure of the form  $\theta^*(Y) = (1 - c/\|Y\|)Y$ , where  $Y$  is an observation and  $c > 0$  is a special constant. This estimate allows one to derive an explicit upper bound for the quadratic risk and has a significantly smaller risk than the usual maximum likelihood estimator for the dimensions  $p \geq 2$ . This procedure is applied to the problem of parametric estimation in a continuous time conditionally Gaussian regression model and to that of estimating the mean vector of a multivariate normal distribution when the covariance matrix is unknown and depends on some nuisance parameters.

*Keywords:* Conditionally Gaussian regression model; Improved estimation; James–Stein procedure; Non-Gaussian Ornstein–Uhlenbeck process.

*AMS 1991 subject classifications:* Primary: 62C20; Secondary: 62C15

---

\*Department of Mathematics and Mechanics, Tomsk State University, Lenin str. 36, 634050 Tomsk, Russia, and Laboratoire de Mathématiques Raphaël Salem, UMR 6085 CNRS, Université de Rouen, Avenue de l'Université BP.12, 76800 Saint Etienne du Rouvray Cedex, France, e-mail: evgen-pch@yandex.ru

# 1 Introduction

In 1961, James and Stein [7] considering the problem of estimating the mean vector  $\theta$  of a  $p$ -dimensional normal distributed random vector  $Y$  with a covariance matrix  $I_p$  introduced an estimator

$$\hat{\theta}_{JS} = \left(1 - \frac{p-2}{\|Y\|^2}\right) Y \quad (1)$$

which outperforms the maximum likelihood estimate (MLE)

$$\hat{\theta}_{ML} = Y \quad (2)$$

for dimension  $p \geq 3$ , under the common quadratic risk

$$R(\theta, \hat{\theta}) = \mathbf{E}_\theta \|\theta - \hat{\theta}\|^2, \quad (3)$$

in the sense that for all parameter values  $\theta$

$$R(\theta, \hat{\theta}_{JS}) < R(\theta, \hat{\theta}_{ML}).$$

This unexpected result draw a great interest of mathematical statisticians and stimulated a number of authors to contribute to the theory of improved estimation by extending the problem of James and Stein in different directions to more general models with unknown covariance matrix and considering other types of estimates (see [1, 2, 6, 9] for more details and other references). A considerable effort has been directed towards the problems of improved estimation in non-Gaussian models with the spherically symmetric distributions (see [3, 5]) and in the non-parametric regression models [4].

Now the James–Stein estimator and other improved shrinkage estimators are widely used in econometrics and the problems associated with the signal processing.

In this paper we will consider the problem of estimating the mean in a conditionally Gaussian distribution. Suppose that the observation  $Y$  is a  $p$ -dimensional random vector which obeys the equation

$$Y = \theta + \xi, \quad (4)$$

where  $\theta$  is a constant vector parameter,  $\xi$  is a conditionally Gaussian random vector with a zero mean and the covariance matrix  $\mathcal{D}(\mathcal{G})$ , i.e.  $Law(\xi|\mathcal{G}) = \mathcal{N}_p(0, \mathcal{D}(\mathcal{G}))$ , where  $\mathcal{G}$  is some fixed  $\sigma$ -algebra.

We propose to consider a shrinkage estimator of the form

$$\theta^* = \left(1 - \frac{c}{\|Y\|}\right) Y, \quad (5)$$

where  $c$  is a positive constant which will be specified below. It will be shown that such an estimator allows one to obtain an explicit upper bound for the quadratic risk in case of the regression model (4) with a conditionally Gaussian noise. Theorem 2.1 in Section 2 claims that the estimator (5) outperforms the maximum likelihood estimate  $\hat{\theta}_{ML}$  uniformly in  $\theta$  from any compact set  $\Theta \subset \mathbb{R}^p$  for any dimension  $p$  starting from two. In Section 3, we apply the estimator (5) to solve the problem of improved parametric estimation in the regression model in continuous time with a non-Gaussian noise.

The rest of the paper is organized as follows. In Section 2, we impose some conditions on the random covariance matrix  $\mathcal{D}(\mathcal{G})$  and derive to upper bound for the difference of risks

$$\Delta(\theta) := R(\theta^*, \theta) - R(\hat{\theta}_{ML}, \theta)$$

corresponding to  $\theta^*$  and  $\hat{\theta}_{ML}$  respectively. In Section 4, the estimate (5) is used for the parameter estimation in a discrete time regression with a Gaussian noise depending on some nuisance parameters. Appendix contains some technical results.

## 2 The upper bound for the estimate risk

In this section we will derive an upper bound for the risk of estimate (5) under some conditions on the random covariance matrix  $\mathcal{D}(\mathcal{G})$ .

Assume that

(C<sub>1</sub>) There exists a positive constant  $\lambda_*$ , such that the minimal eigenvalue of matrix  $\mathcal{D}(\mathcal{G})$  satisfies the inequality

$$\lambda_{min}(\mathcal{D}(\mathcal{G})) \geq \lambda_* \quad \text{a.s.}$$

(C<sub>2</sub>) The maximal eigenvalue of the matrix  $\mathcal{D}(\mathcal{G})$  is bounded on some compact set  $\Theta \subset \mathbb{R}^p$  from above, i.e.

$$\sup_{\theta \in \Theta} \mathbf{E}_\theta \lambda_{max}(\mathcal{D}(\mathcal{G})) \leq a^*,$$

where  $a^*$  is some known positive constant.

Let denote the difference of the risks of estimate (5) and that of (2) as

$$\Delta(\theta) := R(\theta^*, \theta) - R(\hat{\theta}_{ML}, \theta).$$

We will need also the following constant

$$\gamma_p = \frac{\sum_{j=0}^{p-2} 2^{\frac{j-1}{2}} (-1)^{p-j} \mu^{p-1-j} \Gamma\left(\frac{j+1}{2}\right) - (-\mu)^p I(\mu)}{2^{p/2-1} \Gamma\left(\frac{p}{2}\right) d},$$

where  $\mu = d/\sqrt{a^*}$ ,

$$I(a) = \int_0^\infty \frac{\exp(-r^2/2)}{a+r} dr \quad \text{and} \quad d = \sup\{\|\theta\| : \theta \in \Theta\}.$$

**Theorem 2.1.** *Let the noise  $\xi$  in (4) have a conditionally Gaussian distribution  $\mathcal{N}_p(0, \mathcal{D}(\mathcal{G}))$  and its covariance matrix  $\mathcal{D}(\mathcal{G})$  satisfy conditions  $(\mathbf{C}_1)$ ,  $(\mathbf{C}_2)$  with some compact set  $\Theta \subset \mathbb{R}^p$ . Then the estimator (5) with  $c = (p-1)\lambda_*\gamma_p$  dominates the MLE  $\hat{\theta}_{ML}$  for any  $p \geq 2$ , i.e.*

$$\sup_{\theta \in \Theta} \Delta(\theta) \leq -[(p-1)\lambda_*\gamma_p]^2.$$

**Proof.** First we will establish the lower bound for the random variable  $\|Y\|^{-1}$ .

**Lemma 2.2.** *Under the conditions of Theorem 2.1*

$$\inf_{\theta \in \Theta} \mathbf{E}_\theta \frac{1}{\|Y\|} \geq \gamma_p.$$

The proof of lemma is given in the Appendix.

In order to obtain the upper bound for  $\Delta(\theta)$  we will adjust the argument in the proof of Stein's lemma [9] to the model (4) with a random covariance matrix.

We consider the risks of MLE and of (5)

$$R(\hat{\theta}_{ML}, \theta) = \mathbf{E}_\theta \|\hat{\theta}_{ML} - \theta\|^2 = \mathbf{E}_\theta (\mathbf{E} \|\hat{\theta}_{ML} - \theta\|^2 | \mathcal{G}) = \mathbf{E}_\theta \text{tr} \mathcal{D}(\mathcal{G});$$

$$R(\theta^*, \theta) = R(\hat{\theta}_{ML}, \theta) + \mathbf{E}_\theta [\mathbf{E} ((g(Y) - 1)^2 \|Y\|^2 | \mathcal{G})]$$

$$+ 2 \sum_{j=1}^p \mathbf{E}_\theta [\mathbf{E} ((g(Y) - 1) Y_j (Y_j - \theta_j) | \mathcal{G})],$$

where  $g(Y) = 1 - c/\|Y\|$ .

Denoting  $f(Y) = (g(Y) - 1)Y_j$  and applying the conditional density of distribution of a vector  $Y$  with respect to  $\sigma$ -algebra  $\mathcal{G}$

$$p_Y(x|\mathcal{G}) = \frac{1}{(2\pi)^{p/2} \sqrt{\det \mathcal{D}(\mathcal{G})}} \exp\left(-\frac{(x - \theta)' \mathcal{D}^{-1}(\mathcal{G})(x - \theta)}{2}\right),$$

one gets

$$I_j := \mathbf{E}(f(Y)(Y_j - \theta_j)|\mathcal{G}) = \int_{\mathbb{R}^p} f(x)(x - \theta_j)p_Y(x|\mathcal{G})dx, \quad j = \overline{1, p}.$$

Making the change of variable  $u = \mathcal{D}^{-1/2}(\mathcal{G})(x - \theta)$  and assuming  $\tilde{f}(u) = f(\mathcal{D}^{1/2}(\mathcal{G})u + \theta)$ , one finds that

$$I_j = \frac{1}{(2\pi)^{p/2}} \sum_{l=1}^p \langle \mathcal{D}^{1/2}(\mathcal{G}) \rangle_{jl} \int_{\mathbb{R}^p} \tilde{f}(u)u_l \exp\left(-\frac{\|u\|^2}{2}\right) du, \quad j = \overline{1, p},$$

where  $\langle A \rangle_{ij}$  denotes the  $(i, j)$ -th element of matrix  $A$ . These quantities can be written as

$$I_j = \sum_{l=1}^p \sum_{k=1}^p \mathbf{E}(\langle \mathcal{D}^{1/2}(\mathcal{G}) \rangle_{jl} \langle \mathcal{D}^{1/2}(\mathcal{G}) \rangle_{kl} \frac{\partial f}{\partial u_k}(u)|_{u=Y}|\mathcal{G}), \quad j = \overline{1, p}.$$

Thus, the risk for an estimator (5) takes the form

$$R(\theta^*, \theta) = R(\hat{\theta}_{ML}, \theta) + \mathbf{E}_\theta((g(Y) - 1)^2 \|Y\|^2) + 2\mathbf{E}_\theta \left( \sum_{j=1}^p \sum_{l=1}^p \sum_{k=1}^p \langle \mathcal{D}^{1/2}(\mathcal{G}) \rangle_{jl} \langle \mathcal{D}^{1/2}(\mathcal{G}) \rangle_{kl} \frac{\partial}{\partial u_k} [(g(u) - 1)u_j]|_{u=Y} \right).$$

Therefore one has

$$R(\theta^*, \theta) = R(\hat{\theta}_{ML}, \theta) + \mathbf{E}_\theta W(Y),$$

where

$$W(z) = c^2 + 2c \frac{z' \mathcal{D}(\mathcal{G})z}{\|z\|^3} - 2tr \mathcal{D}(\mathcal{G})c \frac{1}{\|z\|}.$$

This implies that

$$\Delta(\theta) = \mathbf{E}_\theta W(Y).$$

Since  $z'Az \leq \lambda_{max}(A)\|z\|^2$ , one comes to the inequality

$$\Delta(\theta) \leq c^2 - 2c\mathbf{E}_\theta \frac{\text{tr}\mathcal{D}(\mathcal{G}) - \lambda_{max}(\mathcal{D}(\mathcal{G}))}{\|Y\|}.$$

From here it follows that

$$\Delta(\theta) \leq c^2 - 2c \sum_{i=2}^p \mathbf{E}_\theta \frac{\lambda_i(\mathcal{D}(\mathcal{G}))}{\|Y\|}.$$

Taking into account the condition  $(\mathbf{C}_1)$  and the Lemma 2.2, one has

$$\Delta(\theta) \leq c^2 - 2(p-1)\lambda_*\gamma_p c =: \phi(c).$$

Minimizing the function  $\phi(c)$  with respect to  $c$ , we come to the desired result

$$\Delta(\theta) \leq -[(p-1)\lambda_*\gamma_p]^2.$$

Hence Theorem 2.1. □

**Corollary 2.3.** *Let in (4) the noise  $\xi \sim \mathcal{N}_p(0, D)$  with the positive definite non random covariance matrix  $D > 0$  and  $\lambda_{min}(D) \geq \lambda_* > 0$ . Then the estimator (5) with  $c = (p-1)\lambda_*\gamma_p$  dominates the MLE for any  $p \geq 2$  and compact set  $\Theta \subset \mathbb{R}^p$ , i.e.*

$$\sup_{\theta \in \Theta} \Delta(\theta) \leq -[(p-1)\lambda_*\gamma_p]^2.$$

**Remark 2.1.** *Note that if  $D = \sigma^2 I_p$  then*

$$\sup_{\theta \in \Theta} \Delta(\theta) \leq -[(p-1)\sigma^2\gamma_p]^2.$$

**Corollary 2.4.** *If  $\xi \sim \mathcal{N}_p(0, I_p)$  and  $\theta = 0$  in model (4) then the risk of estimate (5) is given by the formula*

$$R(0, \theta^*) = p - \left[ \frac{(p-1)\Gamma((p-1)/2)}{\sqrt{2}\Gamma(p/2)} \right]^2 =: r_p. \quad (6)$$

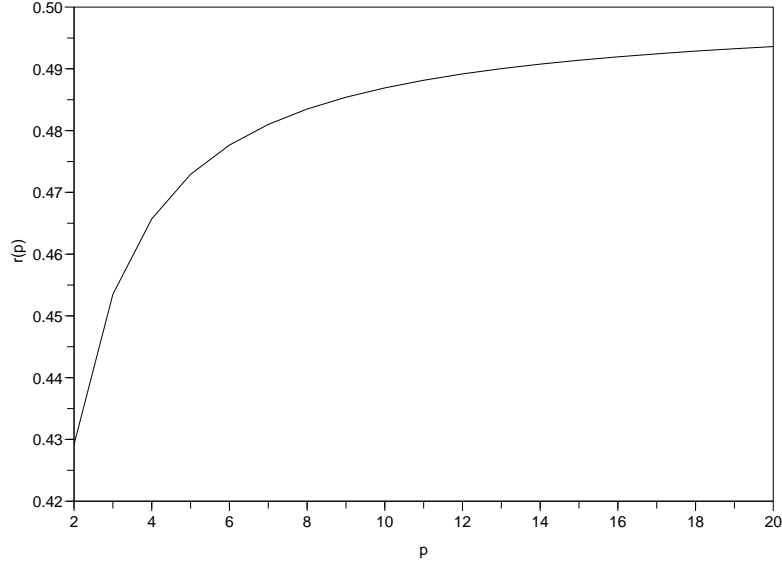


Figure 1: Risk of  $\theta^*$  at  $\theta = 0$ .

By applying the Stirling's formula for the Gamma function

$$\Gamma(x) = \sqrt{2\pi}x^{x-1/2} \exp(-x) (1 + o(1))$$

one can check that  $r_p \rightarrow 0.5$  as  $p \rightarrow \infty$ . The behavior of the risk (6) for small values of  $p$  is shown in Fig.1. It will be observed that in this case the risk of the James–Stein estimate remains constant for all  $p \geq 3$ , i.e.

$$R(0, \hat{\theta}_{JS}) = 2$$

and the risk of the MLE  $\hat{\theta}_{ML}$  is equal to  $p$  and tends to infinity as  $p \rightarrow \infty$ .

### 3 Improved estimation in a non-Gaussian Ornstein–Uhlenbeck–Levy regression model

In this section we apply the proposed estimate (5) to a non-Gaussian continuous time regression model. Let observations  $(y_t)_{0 \leq t \leq n}$  obey the equation

$$dy_t = \sum_{j=1}^p \theta_j \phi_j(t) dt + d\xi_t, \quad 0 \leq t \leq n. \quad (7)$$

Here a vector  $\theta = (\theta_1, \dots, \theta_p)'$  of unknown parameters from some compact set  $\Theta \subset \mathbb{R}^p$ . Assume that  $(\phi_j(t))_{j \geq 1}$  is a one-periodic  $\mathbb{R}_+ \rightarrow \mathbb{R}$  functions, bounded and orthonormal in  $L_2[0, 1]$ . The noise  $(\xi_t)_{t \geq 0}$  in (7) is a non-Gaussian Ornstein–Uhlenbeck process given by stochastic differential equation

$$d\xi_t = a\xi_t dt + du_t, \quad (8)$$

where  $a \leq 0$  is unknown parameter,  $(u_t)_{t \geq 0}$  is a Levy process satisfying the equation

$$u_t = \varrho_1 w_t + \varrho_2 z_t. \quad (9)$$

Here  $\varrho_1, \varrho_2$  are unknown constants,  $(w_t)_{t \geq 0}$  is a standard Brownian motion,  $(z_t)_{t \geq 0}$  is a compound Poisson process defined as

$$z_t = \sum_{j=1}^{N_t} Y_j, \quad (10)$$

where  $(N_t)_{t \geq 0}$  is a Poisson process with unknown intensity  $\lambda > 0$  and  $(Y_j)_{j \geq 1}$  is a sequence of i.i.d. Gaussian random variables with parameters  $(0, 1)$ .

The problem is to estimate the unknown vector parameter  $\theta$  on the basis of observations  $(y_t)_{0 \leq t \leq n}$ .

Let  $\mathcal{G} = \sigma\{N_t, t \geq 0\}$  denote the  $\sigma$ -algebra generated by the Poisson process.

It will be noted that the model (7) is conditionally Gaussian given the  $\sigma$ -algebra  $\mathcal{G}$ . Therefore one can use estimate (5) to obtain an improved estimate of the unknown vector parameter  $\theta$ . To this end we have to reduce the initial continuous time regression model (7) to a discrete time model of the form (4).

The quality of an estimator  $\tilde{\theta}$  will be measured by the quadratic risk

$$R(\theta, \tilde{\theta}) = \mathbf{E}_\theta \|\theta - \tilde{\theta}\|^2.$$



A commonly used estimator of an unknown vector  $\theta$  in model (7) is the least squares estimate (LSE)  $\hat{\theta} = (\hat{\theta}_1, \dots, \hat{\theta}_p)'$  with

$$\hat{\theta}_j = \frac{1}{n} \int_0^n \phi_j(t) dy_t, \quad j = \overline{1, p}.$$

From here taking into account (7), one has

$$\hat{\theta} = \theta + \varepsilon_n \zeta(n), \tag{11}$$

where  $\varepsilon_n = n^{-1/2}$  and  $\zeta(n)$  is the random vector with coordinates

$$\zeta_j(n) = \frac{1}{\sqrt{n}} \int_0^n \phi_j(t) d\xi_t.$$

Note that the vector  $\zeta(n)$  has a conditionally Gaussian distribution with a zero mean and conditional covariance matrix  $V_n(\mathcal{G}) = \text{cov}(\zeta(n), \zeta(n)' | \mathcal{G})$  with the elements

$$v_{ij}(n) = \mathbf{E}(\zeta_i(n) \zeta_j(n) | \mathcal{G}).$$

Thus the initial problem of estimating parameter  $\theta$  in (7) can be reduced to the that of estimating parameter  $\theta$  in conditionally Gaussian regression model (11).

**Theorem 3.1.** *Let the regression model be given by the equations (7)–(10),  $\varrho_1 > 0$ . Then, for any  $n \geq 2$  and  $p \geq 1$ , the estimator of  $\theta$*

$$\theta^* = \left( 1 - \frac{\varrho_1^2(p-1)\gamma_p}{n \|\hat{\theta}\|} \right) \hat{\theta},$$

*dominates the LSE  $\hat{\theta}$ :*

$$\sup_{\theta \in \Theta} \Delta(\theta) \leq - \left[ \frac{\varrho_1^2(p-1)\gamma_p}{n} \right]^2.$$

To proved this theorem one can apply Theorem 2.1 and it suffices to check conditions  $(\mathbf{C}_1)$ ,  $(\mathbf{C}_2)$  on the matrix  $V_n(\mathcal{G})$ . The proof of conditions  $(\mathbf{C}_1)$  and  $(\mathbf{C}_2)$  is given in the Appendix.

## 4 Improved estimation in an autoregression

In this section we consider the problem of improved estimating the unknown mean of a multivariate normal distribution when the dispersion matrix is unknown and depends on some nuisance parameters.

Let in (4) the noise  $\xi = (\xi_1, \dots, \xi_d)'$ , be described by a Gaussian autoregression process

$$\xi_k = a\xi_{k-1} + \varepsilon_k, \quad k = \overline{1, p}, \quad (12)$$

where  $|a| < 1$ ,  $\mathbf{E}\xi_0 = 0$  and  $\varepsilon_1, \dots, \varepsilon_p$  are independent Gaussian  $(0,1)$  random variables. Assume that the parameter  $a$  in (12) is unknown and belongs to interval  $[-\alpha, \alpha]$ , where  $0 < \alpha < 1$  is known number.

It is easy to check that the covariance of the noise  $\xi$  has the form

$$\mathcal{D}(a) = \frac{1}{1-a^2} \begin{pmatrix} 1 & a & \dots & a^{p-1} \\ a & 1 & \dots & a^{p-2} \\ & & \ddots & \\ a^{p-1} & a^{p-2} & \dots & 1 \end{pmatrix}$$

**Proposition 4.1.** *Let  $\xi$  in (4) be specified by (12) with  $a \in [-\alpha, \alpha]$ . Then for any  $p > 1/(1-\alpha)^2$  the MLE is dominated by the estimator*

$$\theta^* = \left( 1 - \left( p - \frac{1}{(1-\alpha)^2} \right) \frac{\gamma_p}{\|Y\|} \right) Y$$

and

$$\sup_{\theta \in \Theta} \Delta(\theta) \leq - \left( p - \frac{1}{(1-\alpha)^2} \right)^2 \gamma_p^2.$$

**Proof.** One has that  $\text{tr}\mathcal{D}(a) = p/(1-a^2)$ . Now we find the estimation of the maximal eigenvalue of matrix  $\mathcal{D}(a)$ . By definition

$$\lambda_{\max}(\mathcal{D}(a)) = \sup_{\|z\|=1} z' \mathcal{D}(a) z$$

one has

$$\begin{aligned} z' \mathcal{D}(a) z &= \sum_{i=1}^p \sum_{j=1}^p \langle \mathcal{D}(a) \rangle_{ij} z_i z_j = \frac{1}{1-a^2} \left( 1 + 2 \sum_{i=1}^{p-1} \sum_{j=1}^{p-i} a^j z_i z_{j+i} \right) \\ &= \frac{1}{1-a^2} \left( 1 + 2 \sum_{j=1}^{p-1} a^j \sum_{i=1}^{p-j} z_j z_{i+j} \right). \end{aligned}$$

By applying the Cauchy–Bunyakovskii inequality we obtain that

$$\lambda_{max}(\mathcal{D}(a)) \leq \frac{1}{1-\alpha^2} \left( 1 + 2 \sum_{j=1}^{\infty} \alpha^j \right) = \frac{1}{(1-\alpha)^2}.$$

Thus,

$$tr\mathcal{D}(a) - \lambda_{max}(\mathcal{D}(a)) \geq p - \frac{1}{(1-\alpha)^2}.$$

Hence, taking into account the Theorem 2.1 we come to assertion of Proposition.

□

## 5 Conclusions

In this paper we propose a new type improved estimation procedure. The main difference from the well-known James–Stein estimate is that in the dominator in the corrected term we take the first power of the observation norm  $\|Y\|$ . This allow us to improve estimation with respect to MLE begining with any dimension  $p \geq 2$ . Moreover, we apply this procedure to the estimation problem for the non-Gaussian Ornstein–Uhlenbeck–Levy regression model.

## 6 Appendix

6.1. Proof of the Lemma 2.2.

**Proof.** From (4) one has

$$J = \mathbf{E}_{\theta} \frac{1}{\|Y\|} = \mathbf{E}_{\theta} \frac{1}{\|\theta + \xi\|} \geq \mathbf{E}_{\theta} \frac{1}{d + \|\xi\|}.$$

Using a repeated conditional expectation and since the random vector  $\xi$  is distributed conditionally normal with a zero mean, then

$$J \geq \mathbf{E}_{\theta} \frac{1}{(2\pi)^{p/2} \sqrt{\det\mathcal{D}(\mathcal{G})}} \int_{\mathbb{R}^p} \frac{\exp(-x'\mathcal{D}(\mathcal{G})^{-1}x/2)}{d + \|x\|} dx.$$

Making the change of variable  $u = \mathcal{D}(\mathcal{G})^{-1/2}x$  and applying the estimation  $u'\mathcal{D}(\mathcal{G})u \leq \lambda_{max}(\mathcal{D}(\mathcal{G}))\|u\|^2$  we find

$$J \geq \frac{1}{(2\pi)^{p/2}} \int_{\mathbb{R}^p} \frac{\exp(-\|u\|^2/2)}{d + \sqrt{\lambda_{max}(\mathcal{D}(\mathcal{G}))}\|u\|} du.$$

Further making the spherical changes of the variables yields

$$J \geq \frac{1}{2^{p/2-1}\Gamma(p/2)} \mathbf{E}_\theta \int_0^\infty \frac{r^{p-1} \exp(-r^2/2)}{d + \sqrt{\lambda_{max}(\mathcal{D}(\mathcal{G}))}r} dr.$$

From here by Jensen and Cauchy–Bunyakovskii inequalities and by the condition  $(\mathbf{C}_2)$  we obtain

$$J \geq \frac{\mu}{2^{p/2-1}\Gamma(p/2)d} \int_0^\infty \frac{r^{p-1} \exp(-r^2/2)}{\mu + r} dr = \gamma_p.$$

This leads to the assertion of Lemma 2.2. □

6.2. The proof of conditions  $(\mathbf{C}_1)$  and  $(\mathbf{C}_2)$  on the matrix  $V_n(\mathcal{G})$ .  
The elements of matrix  $V_n(\mathcal{G})$  can be written as [8]

$$\begin{aligned} v_{ij}(n) &= \frac{\varrho_1^2}{n} \int_0^n \phi_i(t) \phi_j(t) dt \\ &+ \frac{\varrho_1^2}{2n} \int_0^n (\phi_i(t) \varepsilon_{\phi_j}(t) + \phi_j(t) \varepsilon_{\phi_i}(t)) dt + \frac{\varrho_2^2}{n} \sum_{l \geq 1} \phi_i(T_l) \phi_j(T_l) \chi_{(T_l \leq n)} \\ &+ \frac{\varrho_2^2}{n} \sum_{l \geq 1} \int_0^n (\phi_i(t) L_{\phi_j}(t - T_l, T_l) + \phi_j(t) L_{\phi_i}(t - T_l, T_l)) \chi_{(T_l \leq t)} dt, \end{aligned} \quad (13)$$

where

$$\begin{aligned} \varepsilon_g(t) &= a \int_0^t \exp(a(t-s)) g(s) (1 + \exp(2as)) ds, \\ L_g(x, y) &= a \exp(ax) \left( g(y) + a \int_0^x \exp(as) g(s+y) ds \right) \end{aligned}$$

and  $(T_l)_{l \geq 1}$  are the jump times of the Poisson process  $(N_t)_{t \geq 0}$ , i.e.

$$T_l = \inf\{t \geq 0 : N_t = l\}.$$

**Lemma 6.1.** *Let  $(\xi_t)_{t \geq 0}$  be defined by (8) with  $a \leq 0$ . Then a matrix  $V_n(\mathcal{G}) = (v_{ij}(\phi))_{1 \leq i, j \leq p}$  with elements defined by (13), satisfy the following inequality a.s.*

$$\inf_{n \geq 1} \inf_{\|z\|=1} z' V_n(\mathcal{G}) z \geq \varrho_1^2.$$

**Proof.** Notice that by (13) one can the matrix  $V_n(\mathcal{G})$  present as

$$V_n(\mathcal{G}) = \varrho_1^2 I_p + F_n + B_n(\mathcal{G}),$$

where  $F_n$  is non random matrix with elements

$$f_{ij}(n) = \frac{\varrho_1^2}{2n} \int_0^n (\phi_i(t)\varepsilon_{\phi_j}(t) + \phi_j(t)\varepsilon_{\phi_i}(t)) dt$$

and  $B_n(\mathcal{G})$  is a random matrix with elements

$$b_{ij}(n) = \varrho_2^2 \sum_{l \geq 1} [\phi_i(T_l)\phi_j(T_l)\chi_{(T_l \leq n)} + \int_0^n (\phi_i(t)L_{\phi_j}(t - T_l, T_l) + \phi_j(t)L_{\phi_i}(t - T_l, T_l)) \chi_{(T_l \leq t)} dt].$$

This implies that

$$z'V_n(\mathcal{G})z = \varrho_1^2 z'z + z'F_n z + z'B_n(\mathcal{G})z \geq \varrho_1^2 z'z,$$

therefore

$$\inf_{\|z\|=1} z'V_n(\mathcal{G})z \geq \varrho_1^2$$

and we come to the assertion of Lemma 6.1. □

**Lemma 6.2.** *Let  $(\xi_t)_{t \geq 0}$  be defined by (8) with  $a \leq 0$ . Then a maximal eigenvalue of the matrix  $V_n(\mathcal{G}) = (v_{ij}(n))_{1 \leq i, j \leq p}$  with elements defined by (13), satisfy the following inequality*

$$\sup_{n \geq 1} \sup_{\theta \in \Theta} \mathbf{E}_\theta \lambda_{\max}(V_n(\mathcal{G})) \leq Mp\varrho^*,$$

where  $\varrho^* = \varrho_1^2 + \lambda\varrho_2^2$  and  $M > 0$ .

**Proof.** One has

$$\mathbf{E}_\theta \lambda_{\max}(V_n(\mathcal{G})) \leq \mathbf{E}_\theta \text{tr}(V_n(\mathcal{G})) = \sum_{j=1}^p \mathbf{E}_\theta \zeta_j^2(n) = \frac{\varrho^*}{n} \sum_{j=1}^p \tau_j(n),$$

where

$$\begin{aligned} \tau_j(n) &= \int_0^n \phi_j^2(t) dt \\ &\quad + a \int_0^n \phi_j(t) \int_0^t \exp(a(t-u)) \phi_j(u) (1 + \exp(2au)) du dt. \end{aligned}$$

Since the  $(\phi_j)_{1 \leq j \leq p}$  is a one-periodic orthonormal functions therefore the first integral is equal to  $n$  and in view of the inequality  $\max |\phi_j(t)| \leq K$  for any  $a \leq 0$

$$\begin{aligned} a \int_0^n \phi_j(t) \int_0^t \exp(a(t-u)) \phi_j(u) (1 + \exp(2au)) du dt \\ \leq 2K^2 |a| \int_0^n \int_0^t \exp(a(t-u)) du dt \leq 2K^2 n. \end{aligned}$$

From here denoting  $M = 1 + 2K^2$  we obtain that

$$\mathbf{E}_\theta \lambda_{\max}(V_n(\mathcal{G})) \leq Mp\varrho^*.$$

Hence Lemma 6.2. □

Thus the matrix  $V_n(\mathcal{G})$  is positive definite and satisfies for any compact set  $\Theta \subset \mathbb{R}^p$ , the conditions  $(\mathbf{C}_1)$  and  $(\mathbf{C}_2)$  with  $\lambda_* = \varrho_1^2$  and  $a^* = Mp\varrho_*$ .

## References

- [1] J.O. Berger, L.R. Haff, A class of minimax estimators of a normal mean vector for arbitrary quadratic loss and unknown covariance matrix, *Statist. Decisions* 1 (1983) 105-129.
- [2] B. Efron, C. Morris, Families of minimax estimators of the mean of a multivariate normal distribution. *Ann. Statist.* 4 (1976) 11-21.
- [3] D. Fourdrinier, *Statistique inférentielle*, Dunod, 2002, p. 336.
- [4] D. Fourdrinier, S. Pergamenschikov, Improved selection model method for the regression with dependent noise, *Ann. of the Inst. of Statist. Math.*, 59 (3) (2007) 435-464.

- [5] D. Fourdrinier, W.E. Strawderman, E. William, A unified and generalized set of shrinkage bounds on minimax Stein estimates, *J. Multivariate Anal.* 99 (2008) 2221-2233.
- [6] L.J. Gleser, Minimax estimators of a normal mean vector for arbitrary quadratic loss and unknown covariance matrix, *Ann. Statist.* 14 (1986) 1625-1633.
- [7] W. James, C. Stein, Estimation with quadratic loss, in: *Proceedings of the Fourth Berkeley Symposium on Mathematics Statistics and Probability*, Vol. 1, University of California Press, Berkeley, 1961, pp. 361-380.
- [8] V. Konev, S. Pergamenchikov, Efficient robust nonparametric estimation in a semimartingale regression model, <http://hal.archives-ouvertes.fr/hal-00526915/fr/> (2010).
- [9] C. Stein, Estimation of the mean of a multivariate normal distribution, *Ann. Statist.* 9(6) (1981) 1135-1151.