



Estimate the location matrix of a multivariate semiparametric regression model when the random error follows a matrix–variate generalized hyperbolic distribution

Sarmad A. Salih^a, Emad H. Aboudi^{b,*}

^aNineveh Agriculture Directorate, Mosul, Iraq.

^b College of Administration and Economics, University of Baghdad, Iraq.

(Communicated by Madjid Eshaghi Gordji)

Abstract

The matrix-variate generalized hyperbolic distribution is heavy-tailed mixed continuous skewed probability distribution. This distribution has multi applications in the field of economics, risk management, especially in stock modeling.

This paper includes the estimate of the location matrix θ for the multivariate partial linear regression model, which is one of the multivariate semiparametric regression models when the random error follows a matrix-variate generalized hyperbolic distribution in the Bayesian technique depending on non-informative and informative prior information, estimating the location matrix under balanced and unbalanced loss function and the shape parameters (λ, ψ, ν) , skewness matrix (δ) , the scale matrix (Σ) are known. In addition, estimation the smoothing parameter by a proposed method depending on the rule of thumb, the proposed kernel function depending on the mixed Gaussian kernel. the researchers concluded when non-informative and informative prior information is available that the posterior probability distribution for the location matrix θ is a matrix-variate generalized hyperbolic distribution, through the experimental side, it was found that the proposed kernel function is overriding than the Gaussian kernel function in estimate the location matrix and under informative prior information.

Keywords: Multivariate Partial Linear Regression Model, Matrix-variate generalized hyperbolic distribution, Kernel functions, Smoothing Parameter, Bayesian technique.

2020 MSC: mandatory

*Corresponding author

Email addresses: hsarmad012@gmail.com (Sarmad A. Salih), dr.emad.aboudi@gmail.com (Emad H. Aboudi)

1. Introduction

Multivariate regression models are one of the statistical models that have great importance in different areas of life and especially in economic fields. Among these models are multivariate semiparametric regression model, and as their name suggests, they are a mixture of multivariate parametric model and multivariate nonparametric model, the parametric part is the regression function that is supposed to be linear in the observations of its explanatory variables, while the nonparametric part is an unknown smoothing function and is a nonlinear function, the multivariate semiparametric regression model provide an intermediate solution between the parametric and nonparametric model.

Therefore, many researchers have been interested in estimating multivariate semiparametric regression models in which the random error term follows a matrix normal distribution, but there are cases in which the random error follows the heavy tail distribution or extremes (heavier than the matrix normal distribution), in the case above, it is more appropriate to pay attention to alternative probability distributions than the matrix normal distribution, meaning that the mixed distributions are more fits as the matrix-variate generalized hyperbolic distribution . [2] [6] [15]

The first to use the multivariate parametric regression model when the random error follows a matrix-variate generalized hyperbolic distribution are the two researchers (**Thabane & Haq 2004**) and they studied the estimation of the model parameters in the Bayesian technique when the priors distributions belong to the conjugate families as the prior distribution of the location matrix represented the matrix normal distribution and the prior distribution of the scale matrix is a matrix generalized inverse Gaussian distribution as well as finding the Bayesian prediction [15]. Followed by the researcher (**Deschamps, 2012**) as he studied two models of GARCH when the white noise limit is twisted and heavy tails, the first model is GHT-GARCH that represents the threshold for GARCH and the second model is ODLV-GARCH and the variables of both models are twisted variables In addition to the two models overlap with the T-GARCH threshold as a constraint, and by comparing the two models, the researcher concluded that the GHT-GARCH model is better than the ODLV-GARCH model through the Bayesian method and the Bayes factor criterion for information and based on the MCMC algorithm and for five sets of real data [5]. (**Choi, et al. 2009**) test a statistical hypothesis in the Bayesian method of the normal multiple partial linear regression model and assumed that the parametric part of the model is a linear multidimensional function while the nonparametric part is an infinite series of trigonometric functions and deduced upon increasing the sample size that the bayes factor criterion is under the null hypothesis H_0 of the linear function is consistent, that is, it approaches infinity while it approaches zero under the alternative hypothesis H_1 of the partial linear function [4].

The second section deals with the description of the multivariate partial regression model when the random error follows a matrix-variate generalized hyperbolic distribution. In the third section, proposal of the kernel function and the smoothing parameter. The fourth section included finding the posterior probability distribution of the location matrix based on non-informative and informative prior information is available. The fifth section includes Bayesian estimate under unbalanced, balanced loss function, The sixth section includes an experimental side using the language of Matlab, while in the last section show the most important conclusions and future works.

2. Description the Model

The multivariate partial regression model is described according to the following equation: [13] [16]

$$Y_{im} = X_i' \beta_m + g_m(T_i) + \epsilon_{im} \quad i = 1, 2, \dots, n \quad , \quad m = 1, 2, \dots, k \quad (2.1)$$

Since $X'_i\beta_m$ represents the parametric part of the model which β_m is estimated by the parametric methods, such as the method of maximum likelihood, ordinary least squares, moments, or Bayes method ..., and $g_m(T_i)$ represents the nonparametric part of the model, which is an unknown smoothing function that is estimated by the nonparametric methods, such as the kernel smoother, spline smoother and the k-nearest neighbor smoother, it is possible to write the model defined in equation (2.1) in the form of matrices as follows: [1]

$$Y = X\beta + W\gamma + \epsilon \tag{2.2}$$

$$Y = \begin{bmatrix} y_{11} & \cdots & y_{1k} \\ \vdots & \ddots & \vdots \\ y_{n1} & \cdots & y_{nk} \end{bmatrix}_{n \times k}, \quad X = \begin{bmatrix} 1 & x_{11} & \cdots & x_{1p} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n1} & \cdots & x_{np} \end{bmatrix}_{n \times p+1}, \quad \beta = \begin{bmatrix} \beta_{01} & \cdots & \beta_{0k} \\ \vdots & \ddots & \vdots \\ \beta_{p1} & \cdots & \beta_{pk} \end{bmatrix}_{p+1 \times k}$$

$$W = \begin{bmatrix} ker_h(t_1 - T_{11}) & \cdots & ker_h(t_s - T_{1s}) \\ \vdots & \ddots & \vdots \\ ker_h(t_1 - T_{n1}) & \cdots & ker_h(t_s - T_{ns}) \end{bmatrix}_{n \times s}, \quad \gamma = \begin{bmatrix} \gamma_{11} & \cdots & \gamma_{1k} \\ \vdots & \ddots & \vdots \\ \gamma_{s1} & \cdots & \gamma_{sk} \end{bmatrix}_{s \times k}$$

$$\epsilon = \begin{bmatrix} \epsilon_{11} & \cdots & \epsilon_{1k} \\ \vdots & \ddots & \vdots \\ \epsilon_{n1} & \cdots & \epsilon_{nk} \end{bmatrix}_{n \times k}$$

As \mathbf{Y} is a matrix of response variables with dimension $(n \times k)$, \mathbf{n} represents the number of observations, \mathbf{k} represents the number of response variables, \mathbf{X} is a non-random matrix that represents observations of parametric explanatory variables with dimension $(n \times p + 1)$. The p represents the number of parametric explanatory variables, and the β represents the parameters matrix of the parametric part with dimension $(p + 1 \times k)$. As for \mathbf{W} , it is the design matrix which indicates the kernel weighted with dimension $(n \times s)$. The s represents the number of nonparametric explanatory variables and $ker_h(.)$ * represents the kernel function, γ the matrix of parameters of the nonparametric part (Added parameters) with dimension $(s \times k)$, and ϵ is the matrix of random errors with dimension $(n \times k)$. The model can be rewritten in equation (2.2) as follows: [1]

$$Y_{n \times k} = C_{n \times (p+s+1)}\theta_{(p+s+1) \times k} + \epsilon_{n \times k} \tag{2.3}$$

As:

$$C = [X \quad W] \quad , \quad \theta = [\beta \quad \gamma]^T$$

It is assumed that the matrix of errors (ϵ) follows a matrix-variate generalized hyperbolic distribution, using the concept of mixed distributions can be found the probability density function of the matrix(ϵ) from the matrix normal variance-mean mixture distribution with the generalized inverse Gaussian distribution as follows: [6][15]

$$\epsilon | Z \sim MN_{n,k}(\delta Z, Z\Sigma, I_n) \quad , \quad Z \sim GIG(\lambda, \psi, \nu)$$

Whereas, the probability density function of the matrix (ϵ) conditional by random (Z) ($\epsilon | Z$) takes the following formula:

* Kernel function and is as follows ($ker_h(.) = \frac{1}{h} ker(\frac{\cdot}{h})$), and that this function is a real, symmetric and continuous function and that h represents the smoothing parameter and is positive value, they will be mentioned later.

$$f(\epsilon | Z) = \frac{1}{(2\pi)^{\frac{nk}{2}} |\Sigma|^{\frac{n}{2}}} e^{-\frac{1}{2Z} \text{tr}(\epsilon - \delta Z)^T (\epsilon - \delta Z) \Sigma^{-1}} \tag{2.4}$$

The probability density function of the random variable (Z) is as follows: [11]

$$P(Z) = \frac{\left(\frac{\lambda}{\psi}\right)^{\frac{v}{2}}}{2 K_v(\sqrt{\lambda\psi})} Z^{v-1} \exp\left[-\frac{1}{2} \left(\left(\frac{\psi}{Z}\right) + \lambda Z\right)\right], \quad Z > 0 \tag{2.5}$$

According to the concept of mixed distributions, the probability distribution of the matrix of unconditional random errors by (Z) is as follows:

$$f(\epsilon) = \int_0^\infty f(\epsilon | Z) P(Z) dZ$$

$$f(\epsilon) = \frac{\left(\frac{\lambda}{\psi}\right)^{\frac{nk}{4}} e^{\text{tr}(\epsilon)^T \delta \Sigma^{-1}} K_{\frac{2v-nk}{2}} \left(\sqrt{\lambda\psi \left(1 + \frac{\text{tr}(\epsilon^T \epsilon \Sigma^{-1})}{\psi}\right) \left(1 + \frac{\text{tr}(\delta^T \delta \Sigma^{-1})}{\lambda}\right)}\right)}{(2\pi)^{\frac{nk}{2}} |\Sigma|^{\frac{n}{2}} K_v(\sqrt{\lambda\psi})} * \left(1 + \frac{\text{tr}(\epsilon^T \epsilon \Sigma^{-1})}{\psi}\right)^{\frac{2v-nk}{4}} \left(1 + \frac{\text{tr}(\delta^T \delta \Sigma^{-1})}{\lambda}\right)^{\frac{nk-2v}{4}} \tag{2.6}$$

As:

λ, ψ, v : shape parameters.

$K_v(\cdot)$: the modified Bessel function of the third kind of order v which takes the following form: [9]

$$K_v(x) = 0.5 \int_0^\infty t^{v-1} \exp(-0.5 x (t + t^{-1})) dt \quad x > 0 \tag{2.7}$$

δ : skewness matrix of degree $(n \times k)$.

Equation (2.6) represents a matrix-variate generalized hyperbolic distribution for the matrix (ϵ) which is described as follows:

$$\epsilon \sim MGH_{n,k}(0, \Sigma, I_n, \lambda, \psi, \nu, \delta) \leftrightarrow \text{vec}(\epsilon) \sim MGH_{nk}(\text{vec}(0), \Sigma \otimes I_n, \lambda, \psi, \nu, \text{vec}(\delta))$$

Since the matrix (Y) in Equation (2.3) is a linear combination in terms of the matrix (ϵ) that follows a matrix-variate generalized hyperbolic distribution, the probability distribution of (Y) follows a matrix-variate generalized hyperbolic distribution as follows:

The probability density function of the matrix of response variables (Y) conditional by (Z) (Y | Z) that follows a matrix normal variance-mean mixture distribution is as follows:

$$f(Y | Z) = \frac{1}{(2\pi)^{\frac{nk}{2}} |\Sigma|^{\frac{n}{2}}} e^{-\frac{1}{2Z} \text{tr}(Y - C\theta - \delta Z)^T (Y - C\theta - \delta Z) \Sigma^{-1}} \tag{2.8}$$

The probability distribution of (Y) unconditional by (Z), depending on the concept of mixed distri-

butions is as follows:

$$\begin{aligned}
 f(Y) = & \frac{\left(\frac{\lambda}{\psi}\right)^{\frac{nk}{4}} e^{tr(Y-C\theta)^T \delta \Sigma^{-1}}}{(2\pi)^{\frac{nk}{2}} |\Sigma|^{\frac{n}{2}} K_v(\sqrt{\lambda\psi})} K^{\frac{2v-nk}{2}} \\
 & \left(\sqrt{\lambda\psi \left(1 + \frac{tr(Y-C\theta)^T (Y-C\theta) \Sigma^{-1}}{\psi}\right)} \left(1 + \frac{tr \delta^T \delta \Sigma^{-1}}{\lambda}\right) \right) \\
 & * \left(1 + \frac{tr(Y-C\theta)^T (Y-C\theta) \Sigma^{-1}}{\psi}\right)^{\frac{2v-nk}{4}} \left(1 + \frac{tr \delta^T \delta \Sigma^{-1}}{\lambda}\right)^{\frac{nk-2v}{4}} \tag{2.9}
 \end{aligned}$$

As:

θ : The location matrix with a dimension $(p + s + 1 \times k)$.

Σ : The scale matrix with a dimension $(k \times k)$.

This distribution can be expressed descriptively as follows:

$$Y \sim MGH_{(n,k)}(C\theta, \Sigma, I_n, \lambda, \psi, v, \delta) \leftrightarrow vec(Y) \sim MGH_{nk}(vec(C\theta), \Sigma \otimes I_n, \lambda, \psi, v, vec(\delta))$$

3. Kernel functions and smoothing parameter

Using the kernel functions can estimation the regression functions, the spectral functions and the probability density functions, this function is a real, symmetric, continuous and definite function, and its integral is equal to one. the kernel function has other names, including (weight, shape, and window function) and the following table shows some types of kernel functions: [10]

Table 1: Some of the kernel functions

| Kernel | Ker(x) | |
|--------------------|------------------------------|-------------------|
| Epanchnikov | $(3/4)(1 - x^2)$ | $I(x \leq 1)$ |
| Quartic | $(15/16)(1 - x^2)^2$ | $I(x \leq 1)$ |
| Triweight | $(35/32)(1 - x^2)^3$ | $I(x \leq 1)$ |
| Gauss | $(2\pi)^{-0.5} \exp(-x^2/2)$ | $I(x < \infty)$ |
| Uniform | 0.5 | $I(x \leq 1)$ |

Depending on the mixed Gaussian kernel function, the researchers proposal a new kernel function as follows:

We assume the mixed Gaussian kernel function is described as follows:

$$(x | Z) \sim N(0, Z)$$

That (Z) is a random variable which follows the generalized inverse Gaussian distribution defined in equation (2.5) and uses the mixed distributions. The proposed kernel function is as follows:

$$ker(x) = \int_0^\infty ker(x | Z) P(Z) dZ$$

$$ker(x) = \frac{\left(\frac{\lambda}{\psi}\right)^{0.25}}{\sqrt{2\pi} K_v(\sqrt{\lambda\psi})} K_{\frac{2v-1}{2}} \left(\sqrt{\lambda\psi \left(1 + \frac{x^2}{\psi}\right)} \right) \left(1 + \frac{x^2}{\psi}\right)^{\frac{2v-1}{4}} \tag{3.1}$$

Equation (3.1) represents the proposed kernel function, and can be called the symmetric generalized hyperbolic kernel function and is described as follows:

$$ker(x) \sim GH(0, 1, \lambda, \psi, v, 0)$$

For example, the kernel function used in the estimation is the Gaussian Kernel function, and based on the rule of thumb. the smoothing parameter at the second kernel degree is as follows: [7][14]

$$h_{thumb} = 1.06 \hat{\sigma} n^{-\frac{1}{5}} \tag{3.2}$$

As h_{thumb} is a non-random, and positive parameter, the smoothing parameter is usually chosen according to the researcher’s experience or iteration methods to obtain the best bandwidth parameter. The smoothing parameter greatly affects the bias and variance. Moreover, the researchers suggested counting on the rule of normal distribution and a Gaussian kernel function and assuming the normal probability density function is a mixed function, so the suggested smoothing parameter is as follows:

$$h_{sug.} = 1.06 * \hat{\sigma} * n^{-\frac{1}{5}} * \left(\frac{K_v(\sqrt{\lambda\psi}) \left(\frac{\lambda}{\psi}\right)^{\frac{-5}{4}}}{K_{\frac{2v-5}{2}}(\sqrt{\lambda\psi})} \right)^{1/5} \tag{3.3}$$

As $\hat{\sigma}$ represents the standard deviation of the sample, it is possible to use any other kernel function according to the following rule: [14]

$$h^* = C_K * h_{sug.} \tag{3.4}$$

As:

$$C_K = \left(\frac{2\sqrt{\pi} R(K)}{\mu_2^2(K)} \right)^{\frac{1}{5}} \tag{3.5}$$

Whereas $h_{sug.}$ represents the proposed smoothing parameter depending on the Gaussian kernel function, so any derived rule based on the Gaussian kernel function can be inverted by relying on other kernel functions by multiplying it by a multiplier. Therefore, the proposed smoothing parameter and based on the multiplier of the proposed kernel function is as follows:

$$h^* = \left(\frac{K_v(\sqrt{\lambda\psi}) K_{\frac{2v-1}{2}}(\sqrt{\lambda\psi})}{\left(\frac{\lambda}{\psi}\right)^{\frac{-5}{4}} (K_{v+1}(\sqrt{\lambda\psi}))^2} \right)^{\frac{1}{5}} * 1.06 * \hat{\sigma} * n^{-\frac{1}{5}} * \left(\frac{K_v(\sqrt{\lambda\psi}) \left(\frac{\lambda}{\psi}\right)^{\frac{-5}{4}}}{K_{\frac{2v-5}{2}}(\sqrt{\lambda\psi})} \right)^{1/5} \tag{3.6}$$

4. Posterior probability distribution of the location matrix

In this topic, the location matrix (θ) is defined in equation (2.3) is estimated based on non-informative and informative prior information under the assumption that the shape parameters, the scale matrix (Σ) are known.

4.1. Non-informative prior information

The prior distribution of location matrix (θ) is found from Fisher's information by taking the natural logarithm of the two sides of the probability density function of (Y) conditional by (Z) $(Y|Z)$ and knowing in equation (2.8) and taking the second partial derivative relative to (θ) , the prior distribution of $(\theta|Z)$ is as follows: [12]

$$I_Y^{(\theta)} = \left[\left[-E \frac{\partial^2 \ln f(Y | \theta, \Sigma)}{\partial \theta^T \partial \theta} \right] \right] \tag{4.1}$$

$$= \frac{K_{v-1}(\sqrt{\lambda\psi})}{K_v(\sqrt{\lambda\psi})} \left(\frac{\lambda}{\psi}\right)^{0.5} (C^T C) \otimes \Sigma^{-1} \tag{4.2}$$

We notice from equation (4.2) that the prior distribution of matrix (θ) is a constant distribution, meaning that:

$$P(\theta) \propto \text{Constant Matrix} \tag{4.3}$$

By merge the probability density function of (Y) (conditional by (Z) (defined in equation (8) with the prior probability distribution defined in equation (4.3), we obtain the kernel of the posterior probability distribution for the location matrix (θ) conditional by the random variable (Z) (as follows:

$$P(\theta | Y, \Sigma, Z) \propto P(\theta) f(Y | \theta, \Sigma, Z) \propto \exp\left(-\frac{1}{2Z} \text{tr} (Y - C\theta - \delta Z)^T (Y - C\theta - \delta Z) \Sigma^{-1}\right) \tag{4.4}$$

By adding and subtracting the amount $C\hat{\theta}_{m|Z}$ to the exponential function in the equation (4.4) and that $\hat{\theta}_{m|Z}$ represents the conditional maximum likelihood estimator, which was found by deriving normal logarithm of equation (2.8) a partial derivation of θ as follows:

$$\hat{\theta}_{m|Z} = (C^T C)^{-1} C^T Y - (C^T C)^{-1} C^T \delta Z \tag{4.5}$$

And procedure some mathematical simplifications, we get the following:

$$P(\theta | Y, \Sigma, Z) \propto \exp\left(-\frac{1}{2Z} \text{tr} (\theta - \hat{\theta}_{m|Z})^T C^T C (\theta - \hat{\theta}_{m|Z}) \Sigma^{-1}\right) \tag{4.6}$$

Equation (4.6) represents the kernel of the matrix normal variance-mean mixture distribution. Therefore, the posterior distribution of the $(\theta | Y, \Sigma, Z)$ is as follows:

$$P(\theta | Y, \Sigma, Z) = \frac{|C^T C|^{\frac{k}{2}} e^{-\frac{1}{2Z} \text{tr} (\theta - \hat{\theta}^* + (C^T C)^{-1} C^T \delta Z)^T C^T C (\theta - \hat{\theta}^* + (C^T C)^{-1} C^T \delta Z) \Sigma^{-1}}}{(2\pi Z)^{\frac{(p+s+1)k}{2}} |\Sigma|^{\frac{(p+s+1)}{2}}} \tag{4.7}$$

Express equation (4.7) descriptively as follows:

$$(\theta | Z) \sim MN_{(p+s+1), k} \left((C^T C)^{-1} C^T Y + (C^T C)^{-1} C^T (-\delta) Z, Z\Sigma, (C^T C)^{-1} \right)$$

Accordingly, the posterior probability distribution of location matrix (θ) unconditional of Z and according to mixed distributions is as follows:

$$P(\theta | Y, \Sigma) = \int_0^\infty P(\theta | Y, \Sigma, Z) P(Z) dZ$$

$$\begin{aligned}
 P(\theta | Y, \Sigma) &= \frac{\left(\frac{\lambda}{\psi}\right)^{\frac{(p+s+1)k}{4}} |C^T C|^{\frac{k}{2}} e^{tr(\theta - \hat{\theta}^*)^T C^T (-\delta) \Sigma^{-1}}}{(2\pi)^{\frac{(p+s+1)k}{2}} |\Sigma|^{\frac{(p+s+1)k}{2}} K_v(\sqrt{\lambda\psi})} \\
 &* K_{\frac{2v-(p+s+1)k}{2}} \left(\sqrt{\lambda\psi \left(1 + \frac{tr(\theta - \hat{\theta}^*)^T C^T C (\theta - \hat{\theta}^*) \Sigma^{-1}}{\psi}\right)} \left(1 + \frac{tr \delta^T C (C^T C)^{-1} C^T \delta \Sigma^{-1}}{\lambda}\right) \right) \\
 &* \left(1 + \frac{tr(\theta - \hat{\theta}^*)^T C^T C (\theta - \hat{\theta}^*) \Sigma^{-1}}{\psi}\right)^{\frac{2v-(p+s+1)k}{4}} * \left(1 + \frac{tr \delta^T C (C^T C)^{-1} C^T \delta \Sigma^{-1}}{\lambda}\right)^{\frac{-2v+(p+s+1)k}{4}} \quad (4.8)
 \end{aligned}$$

As:

$$\hat{\theta}^* = (C^T C)^{-1} C^T Y$$

Equation (4.8) represents a matrix-variate generalized hyperbolic distribution and is described as follows:

$$\theta \sim MGH_{(p+s+1), k}(\hat{\theta}^*, \Sigma, (C^T C)^{-1}, \lambda, \psi, v, (C^T C)^{-1} C^T (-\delta))$$

Or:

$$vec(\theta) \sim MGH_{(p+s+1)k}(vec(\hat{\theta}^*), \Sigma \otimes (C^T C)^{-1}, \lambda, \psi, v, vec((C^T C)^{-1} C^T (-\delta)))$$

4.2. Informative prior information (Conjugate family)

We know that the prior distribution of $(\theta | \Sigma, Z)$ has the following form:

$$P(\theta | \Sigma, Z) \propto \exp\left(-\frac{1}{2Z} tr(\theta - \theta_0)^T U_0^{-1} (\theta - \theta_0) \Sigma^{-1}\right) \quad (4.9)$$

Express equation (4.9) descriptively as follows:

$$(\theta | Z) \sim MN_{(p+s+1), k}(\theta_{0(p+s+1) \times k}, Z \Sigma_{k \times k}, U_{0(p+s+1) \times (p+s+1)})$$

The posterior probability distribution of the matrix $(\theta | Y, \Sigma, Z)$ is the distribution resulting from merging the equation (2.8) with the equation (4.9), as follows:

$$\begin{aligned}
 P(\theta | Y, \Sigma, Z) &\propto P(\theta | \Sigma, Z) f(Y | \theta, \Sigma, Z) \\
 &\propto \exp\left(-\frac{1}{2Z} tr(\theta - \theta_0)^T U_0^{-1} (\theta - \theta_0) \Sigma^{-1}\right) \\
 &* \exp\left(-\frac{1}{2Z} tr(Y - C\theta - \delta Z)^T (Y - C\theta - \delta Z) \Sigma^{-1}\right) \quad (4.10)
 \end{aligned}$$

By adding and subtracting the amount $C\hat{\theta}_{m|Z}$ to the second exponential function in the equation (4.10) and that $\hat{\theta}_{m|Z}$ represents the conditional maximum likelihood estimator, which previously defined in equation (4.5) And by making some mathematical simplifications, we get the following:

$$\begin{aligned}
 P(\theta | Y, \Sigma, Z) &\propto \exp\left(-\frac{1}{2Z} tr(\theta - \theta_0)^T U_0^{-1} (\theta - \theta_0) \Sigma^{-1}\right) \\
 &* \exp\left(-\frac{1}{2Z} tr(\theta - \hat{\theta}_{m|Z})^T C^T C (\theta - \hat{\theta}_{m|Z}) \Sigma^{-1}\right) \quad (4.11)
 \end{aligned}$$

We transform the matrices in the equation (4.11) into the vector transformation (vector operator) formula as follows:

$$P(\theta | Y, \Sigma, Z) \propto \exp \left[-\frac{1}{2Z} \left\{ (vec \theta - vec \theta_0)^T (U_0 \otimes \Sigma)^{-1} (vec \theta - vec \theta_0) + (vec \theta - vec \hat{\theta}_{m|Z})^T \left((C^T C)^{-1} \otimes \Sigma \right)^{-1} (vec \theta - vec \hat{\theta}_{m|Z}) \right\} \right] \quad (4.12)$$

We resemble the quadratic form in the equation (4.12) with the following quadratic form: [3]

$$(\underline{X} - \underline{a})^T A (\underline{X} - \underline{a}) + (\underline{X} - \underline{b})^T B (\underline{X} - \underline{b}) = (\underline{X} - \underline{d})^T (A + B) (\underline{X} - \underline{d}) + (\underline{a} - \underline{b})^T A (A + B)^{-1} B (\underline{a} - \underline{b}) \quad (4.13)$$

As $(\underline{X}, \underline{a}, \underline{b}, \underline{d})$ are vectors with dimension $(k^2 \times 1)$ and (A, B) are matrices with dimension $(k^2 \times k^2)$ and that:

$$\underline{d} = (A + B)^{-1} (A \underline{a} + B \underline{b})$$

Returning to equation (4.13), then:

$$\begin{aligned} \underline{X} &= vec \theta, \quad \underline{a} = vec \theta_0, \quad \underline{b} = vec \hat{\theta}_{m|Z}, \quad A = (U_0 \otimes \Sigma)^{-1}, \quad B = \left((C^T C)^{-1} \otimes \Sigma \right)^{-1} \\ \underline{d} &= \left((U_0 \otimes \Sigma)^{-1} + \left((C^T C)^{-1} \otimes \Sigma \right)^{-1} \right)^{-1} \left[(U_0 \otimes \Sigma)^{-1} vec \theta_0 + \left((C^T C)^{-1} \otimes \Sigma \right)^{-1} vec \hat{\theta}_{m|Z} \right] \end{aligned}$$

We assume that:

$$\begin{aligned} Q_{11} &= (U_0 \otimes \Sigma)^{-1} + \left((C^T C)^{-1} \otimes \Sigma \right)^{-1} \\ \underline{d} &= (Q_{11})^{-1} \left[(U_0 \otimes \Sigma)^{-1} vec \theta_0 + \left((C^T C)^{-1} \otimes \Sigma \right)^{-1} vec \hat{\theta}_{m|Z} \right] = vec \theta^{**} \end{aligned} \quad (4.14)$$

Therefore, the kernel of the posterior probability distribution of (θ) conditional by the variable (Z) is as follows:

$$P(\theta | Y, \Sigma, Z) \propto e^{-\frac{1}{2Z} (vec \theta - vec \theta^{**})^T \left[(U_0 \otimes \Sigma)^{-1} + \left((C^T C)^{-1} \otimes \Sigma \right)^{-1} \right] (vec \theta - vec \theta^{**})} \quad (4.15)$$

Return equation (4.15) to the matrix formula depending on the properties of the vector and the Kronecker product, we get the following:

$$P(\theta | Y, \Sigma, Z) \propto \exp \left[-\frac{1}{2Z} tr \left((\theta - \theta^{**})^T [U_0^{-1} + (C^T C)] (\theta - \theta^{**}) \Sigma^{-1} \right) \right] \quad (4.16)$$

As:

$$\theta^{**} = [U_0^{-1} + (C^T C)]^{-1} [U_0^{-1} \theta_0 + (C^T C) \hat{\theta}_{m|Z}] \quad (4.17)$$

Equation (4.16) represents the kernel of the matrix normal variance-mean mixture distribution. Therefore, the posterior distribution of the $(\theta | Y, \Sigma, Z)$ is as follows:

$$P(\theta | Y, \Sigma, Z) = \frac{|U_0^{-1} + (C^T C)|^{\frac{k}{2}}}{(2\pi Z)^{\frac{(p+s+1)k}{2}} |\Sigma|^{\frac{(p+s+1)}{2}}} e^{-\frac{1}{2Z} tr \left((\theta - \theta^{**})^T [U_0^{-1} + (C^T C)] (\theta - \theta^{**}) \Sigma^{-1} \right)} \quad (4.18)$$

Express equation (4.18) descriptively as follows:

$$(\theta | Y, \Sigma, Z) \sim MN_{(p+s+1), k} \left(\theta^{**}_{(p+s+1) \times k}, Z \Sigma_{k \times k}, [U_0^{-1} + (C^T C)]^{-1}_{(p+s+1) \times (p+s+1)} \right)$$

Depending on mixed distributions, the posterior probability distribution of location matrix (θ) unconditional of Z is as follows:

$$\begin{aligned}
 P(\theta | Y, \Sigma) &= \int_0^\infty P(\theta | Y, \Sigma, Z) P(Z) dZ \\
 P(\theta | Y, \Sigma) &= \frac{\left(\frac{\lambda}{\psi}\right)^{\frac{(p+s+1)k}{4}} |U_0^{-1} + C^T C|^{\frac{k}{2}} e^{tr (\theta - Q_{22})^T C^{*T} (-\delta) \Sigma^{-1}}}{(2\pi)^{\frac{(p+s+1)k}{2}} |\Sigma|^{\frac{(p+s+1)}{2}}} \frac{K_v(\sqrt{\lambda\psi})}{K_v(\sqrt{\lambda\psi})} \\
 &\quad * K_{\frac{2v-(p+s+1)k}{2}} \left(\sqrt{\lambda\psi \left(1 + \frac{tr Q_{22}^* \Sigma^{-1}}{\psi}\right) \left(1 + \frac{tr \delta^T C^* C^{*T} \delta \Sigma^{-1}}{\lambda}\right)} \right) \\
 &\quad * \left(1 + \frac{tr Q_{22}^* \Sigma^{-1}}{\psi}\right)^{\frac{2v-(p+s+1)k}{4}} \left(1 + \frac{tr \delta^T C^* C^{*T} \delta \Sigma^{-1}}{\lambda}\right)^{\frac{-2v+(p+s+1)k}{4}} \tag{4.19}
 \end{aligned}$$

As:

$$\begin{aligned}
 Q_{22}^* &= (\theta - Q_{22})^T [U_0^{-1} + (C^T C)] (\theta - Q_{22}) \\
 Q_{22} &= [U_0^{-1} + (C^T C)]^{-1} [U_0^{-1} \theta_0 + (C^T C) \hat{\theta}^*] \\
 C^{*T} &= [U_0^{-1} + (C^T C)]^{-1} C^T
 \end{aligned}$$

Equation (4.19) represents a matrix-variate generalized hyperbolic distribution and is described as follows:

$$\theta \sim MGH_{(p+s+1), k} \left(Q_{22}, \Sigma, (U_0^{-1} + (C^T C))^{-1}, \lambda, \psi, v, C^{*T}(-\delta) \right)$$

Or:

$$vec(\theta) \sim MGH_{(p+s+1)k} \left(vec(Q_{22}), \Sigma \otimes (U_0^{-1} + (C^T C))^{-1}, \lambda, \psi, v, vec(C^{*T}(-\delta)) \right)$$

5. Loss functions of location matrix (θ)

In this section, the location matrix (θ) is defined in equation (2.3) is estimated based on entropy loss function and entropy balanced loss function.

5.1. Entropy loss function

The formula for this function is as follows:[8]

$$l_E(\hat{\theta}, \theta) = \left(\frac{\theta}{\hat{\theta}}\right) - \ln\left(\frac{\theta}{\hat{\theta}}\right) - 1 \tag{5.1}$$

$$l_E(\hat{\theta}, \theta) = \sum_{i=1}^{p+s+1} \sum_{j=1}^k \left[\frac{\theta_{ij}}{\hat{\theta}_{ij}} - \log\left(\frac{\theta_{ij}}{\hat{\theta}_{ij}}\right) - 1 \right] \tag{5.2}$$

To find the Bayesian estimator, we make the mathematical expectation of the entropy loss function relative to (θ) as low as possible, which means the risk of the estimator $(\hat{\theta})$, meaning that:

$$\frac{\partial R_E(\hat{\theta}, \theta)}{\partial \hat{\theta}_{ij}} = \begin{bmatrix} \frac{\partial R_E(\hat{\theta}, \theta)}{\partial \hat{\theta}_{11}} & \frac{\partial R_E(\hat{\theta}, \theta)}{\partial \hat{\theta}_{12}} & \dots & \frac{\partial R_E(\hat{\theta}, \theta)}{\partial \hat{\theta}_{1k}} \\ \frac{\partial R_E(\hat{\theta}, \theta)}{\partial \hat{\theta}_{21}} & \frac{\partial R_E(\hat{\theta}, \theta)}{\partial \hat{\theta}_{22}} & \dots & \frac{\partial R_E(\hat{\theta}, \theta)}{\partial \hat{\theta}_{2k}} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial R_E(\hat{\theta}, \theta)}{\partial \hat{\theta}_{(p+s+1)1}} & \frac{\partial R_E(\hat{\theta}, \theta)}{\partial \hat{\theta}_{(p+s+1)2}} & \dots & \frac{\partial R_E(\hat{\theta}, \theta)}{\partial \hat{\theta}_{(p+s+1)k}} \end{bmatrix}$$

In order to find the elements of the matrix, we derive the entropy risk function relative to $(\hat{\theta}_{ij})$ as follows:

$$\begin{aligned} \frac{\partial R_E(\hat{\theta}, \theta)}{\partial \hat{\theta}_{ij}} &= \int_{\theta_{ij}} \frac{\partial}{\partial \hat{\theta}_{ij}} l_E(\hat{\theta}, \theta) P(\theta_{ij} | Y) d\theta_{ij} \\ \frac{\partial R_E(\hat{\theta}, \theta)}{\partial \hat{\theta}_{ij}} &= \int_{\theta_{ij}} \left[-\frac{\theta_{ij}}{\hat{\theta}_{ij}^2} - \frac{\hat{\theta}_{ij}}{\theta_{ij}} \left(-\frac{\theta_{ij}}{\hat{\theta}_{ij}^2} \right) \right] P(\theta_{ij} | Y) d\theta_{ij} \end{aligned}$$

To find the Bayesian estimator, we set $\frac{\partial R_E(\hat{\theta}, \theta)}{\partial \hat{\theta}_{ij}} = 0$ as follows:

$$\begin{aligned} \int_{\theta_{ij}} \left[-\frac{\theta_{ij}}{\hat{\theta}_{ij}^2} + \frac{\hat{\theta}_{ij}}{\theta_{ij}} \right] P(\theta_{ij} | Y) d\theta_{ij} &= 0 \\ \frac{1}{\hat{\theta}_{ij}^2} [\hat{\theta}_{ij} - E_{\theta_{ij}}(\theta_{ij} | Y)] &= 0 \\ \hat{\theta}_{E(ij)} &= E_{\theta_{ij}}(\theta_{ij} | Y) \end{aligned} \tag{5.3}$$

Equation (5.3) represents the Bayesian estimator under the entropy loss function of the element (ij) from the location matrix (θ) and represents the mean posterior distribution for the element. Therefore, the Bayesian estimator for the location matrix represents the mean posterior probability distribution of the matrix.

$$\hat{\theta}_E = E_{\theta}(\theta | Y) \tag{5.4}$$

5.2. Entropy balanced loss function

The formula for the entropy balanced loss function is as follows:[8]

$$l_{Eb}(\hat{\theta}, \theta) = w \sum_{i=1}^{p+s+1} \sum_{j=1}^k \left[\frac{\hat{\theta}_{m(ij)}}{\hat{\theta}_{ij}} - \log \left(\frac{\hat{\theta}_{m(ij)}}{\hat{\theta}_{ij}} \right) - 1 \right] + (1-w) \sum_{i=1}^{p+s+1} \sum_{j=1}^k \left[\frac{\theta_{ij}}{\hat{\theta}_{ij}} - \log \left(\frac{\theta_{ij}}{\hat{\theta}_{ij}} \right) - 1 \right] \tag{5.5}$$

As:

$\hat{\theta}_{m(ij)}$: Represents the maximum likelihood estimator of θ .

To find the Bayesian estimator, we make the mathematical expectation of the entropy balanced loss function relative to (θ) as low as possible, which means the risk of the estimator $(\hat{\theta})$,

meaning that:

$$\frac{\partial R_{Eb}(\hat{\theta}, \theta)}{\partial \hat{\theta}_{ij}} = \begin{bmatrix} \frac{\partial R_{Eb}(\hat{\theta}, \theta)}{\partial \hat{\theta}_{11}} & \frac{\partial R_{Eb}(\hat{\theta}, \theta)}{\partial \hat{\theta}_{12}} & \dots & \frac{\partial R_{Eb}(\hat{\theta}, \theta)}{\partial \hat{\theta}_{1k}} \\ \frac{\partial R_{Eb}(\hat{\theta}, \theta)}{\partial \hat{\theta}_{21}} & \frac{\partial R_{Eb}(\hat{\theta}, \theta)}{\partial \hat{\theta}_{22}} & \dots & \frac{\partial R_{Eb}(\hat{\theta}, \theta)}{\partial \hat{\theta}_{2k}} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial R_{Eb}(\hat{\theta}, \theta)}{\partial \hat{\theta}_{(p+s+1)1}} & \frac{\partial R_{Eb}(\hat{\theta}, \theta)}{\partial \hat{\theta}_{(p+s+1)2}} & \dots & \frac{\partial R_{Eb}(\hat{\theta}, \theta)}{\partial \hat{\theta}_{(p+s+1)k}} \end{bmatrix}$$

In order to find the elements of the matrix, we derive the entropy balanced risk function relative to $(\hat{\theta}_{ij})$ as follows:

$$\begin{aligned} \frac{\partial R_{Eb}(\hat{\theta}, \theta)}{\partial \hat{\theta}_{ij}} &= \int_{\theta_{ij}} \frac{\partial}{\partial \hat{\theta}_{ij}} l_{Eb}(\hat{\theta}, \theta) P(\theta_{ij} | Y) d\theta_{ij} \\ \frac{\partial R_{Eb}(\hat{\theta}, \theta)}{\partial \hat{\theta}_{ij}} &= \left[w \left(-\frac{\hat{\theta}_{m(ij)}}{\hat{\theta}_{ij}^2} - \frac{\hat{\theta}_{ij}}{\hat{\theta}_{m(ij)}} \left(-\frac{\hat{\theta}_{m(ij)}}{\hat{\theta}_{ij}^2} \right) \right) \right] \\ &\quad + (1-w) \int_{\theta_{ij}} \left[-\frac{\theta_{ij}}{\hat{\theta}_{ij}^2} - \frac{\hat{\theta}_{ij}}{\theta_{ij}} \left(-\frac{\theta_{ij}}{\hat{\theta}_{ij}^2} \right) \right] P(\theta_{ij} | Y) d\theta_{ij} \end{aligned} \tag{5.6}$$

To find the Bayes estimator, we set $\frac{\partial R_{Eb}(\hat{\theta}, \theta)}{\partial \hat{\theta}_{ij}} = 0$ as follows:

$$\begin{aligned} \frac{1}{\hat{\theta}_{ij}^2} \left[w \left(\hat{\theta}_{ij} - \hat{\theta}_{m(ij)} \right) + (1-w) \left(\hat{\theta}_{ij} - E_{\theta_{ij}}(\theta_{ij} | Y) \right) \right] &= 0 \\ w \left(\hat{\theta}_{ij} \right) + (1-w) \left(\hat{\theta}_{ij} \right) &= w \left(\hat{\theta}_{m(ij)} \right) + (1-w) E_{\theta_{ij}}(\theta_{ij} | Y) \\ \hat{\theta}_{Eb(ij)} &= w \left(\hat{\theta}_{m(ij)} \right) + (1-w) E_{\theta_{ij}}(\theta_{ij} | Y) \end{aligned} \tag{5.7}$$

Equation (5.7) represents the Bayesian estimator under the entropy balanced loss function of the element (ij) from the matrix (θ) and represents the mean posterior distribution for the element. Therefore, the Bayesian estimator for the location matrix represents the mean posterior probability distribution of the matrix.

$$\hat{\theta}_{Eb} = w \left(\hat{\theta}_m \right) + (1-w) E_{\theta}(\theta | Y) \quad , \quad 0 \leq w \leq 1 \tag{5.8}$$

6. Experimental side

This section discusses the simulate of the mechanism reached in the theoretical sections to data generated from a multivariate partial linear regression model with random error following the matrix-variate generalized hyperbolic distribution.

6.1. data generation

It is resorted to generate data through mixed distributions, as the matrix normal variance-mean mixture distribution and the generalized inverse Gaussian distribution previously mentioned were used, because it is difficult to generate data from the multivariate partial regression model with the random error following the matrix-variate generalized hyperbolic distribution.

Whereas, random observations were generated from the multivariate standard normal (Z), and since $Z = (\epsilon | Z - \delta Z) (Z \Sigma)^{-0.5}$ and ϵ represents the matrix of random errors of the model from which the observations are to be generated and through the concept of mixed distributions as follow:

$$\epsilon | Z = Z * (Z \Sigma)^{0.5} + \delta Z \tag{6.1}$$

$$\epsilon = \int_Z \epsilon | Z P(Z) dZ$$

$$\epsilon = Z * \Sigma^{0.5} * \frac{K_{\frac{2v+1}{2}}(\sqrt{\lambda\psi})}{K_v(\sqrt{\lambda\psi}) \left(\frac{\lambda}{\psi}\right)^{\frac{1}{4}}} + \delta * \frac{K_{v+1}(\sqrt{\lambda\psi})}{K_v(\sqrt{\lambda\psi}) \left(\frac{\lambda}{\psi}\right)^{\frac{1}{2}}} \tag{6.2}$$

Equation (6.2) represents the matrix of random errors, which follows a matrix-variate generalized hyperbolic distribution.

The below algorithm shows the method for generating data from a matrix-variate generalized hyperbolic distribution:

Step1: Assume we the number of observations ($n = 100$) and the number of response variables ($k = 2$).

Step2: Generate random numbers from the multivariate standard normal distribution with (n) observations, let the multivariate standard normal random number matrix be (Z).

Step3: We put $(\epsilon | Z = Z * (Z \Sigma)^{0.5} + \delta Z)$, (Z) represents **Step2**.

Step4: Find the generated observations (ϵ) which represent the random error observations generated from the matrix-variate generalized hyperbolic distribution taking into account the assumed values of the shape parameters (λ, ψ, ν), the scale matrix, and skewness matrix (δ) defined in table 2.

Step5: The purpose of generating data from the multivariate partial linear regression model, we generate the data of the two explanatory variables ($p=2$), ($s = 2$) for the parametric and nonparametric part (X_1, X_2) and (t_1, t_2) respectively, the parametric variables through the following equation:

$$X_j = 2 \overline{X_j} u_j$$

Since u_j represents the standard uniform distribution, $\overline{X_j}$ represents the arithmetic mean and they are usually assumed values, and the nonparametric part (W) represents the kernel weights one time we take it to the Gaussian kernel function and another we take it to the proposed kernel function which was previously defined in equation (3.1) and depending on the rule of normal distribution to choose the smoothing parameter for both functions, the nonparametric variables (t_1, t_2) is a standard normal variables.

Step6: Randomly assumed values are given for the location matrix θ , the scale matrix S , and for shape parameters they are given random values depending on the state of the studied distribution $\lambda, \psi, \nu > 0$ and as in table 2 below.

Table 2: Approved default values for all parameters

| | λ | ψ | η | $\theta_{(p+s+1) \times k}$ | $\Sigma_{k \times k}$ | δ |
|---|-----------|--------|--------|--|--|--|
| 1 | 3 | 0.5 | 2.5 | $\begin{bmatrix} 2 & 1.4 & 5 & 3 & 4.1 \\ 1.5 & 3 & 2.5 & 1 & 3.4 \end{bmatrix}^T$ | $\begin{bmatrix} 1.5 & 2 \\ 2 & 3 \end{bmatrix}$ | $\begin{bmatrix} 0.6498 \\ 0.9347 \end{bmatrix}^T$ |
| 2 | 5 | | 3.5 | | | |
| 3 | 7 | | | | | |

Step7: After substituting **Step4** and data of the two explanatory variables for the parametric part, the kernel weights matrix for the nonparametric part defined in **Step5**, and the assumed values of the parameters defined in **Step6**, we obtain (2.6) models depending on the combination between the assumed values of the response matrix Y .

6.2. Estimation of the location matrix

The location matrix θ is estimated in a Bayesian technique based on non-informative, informative prior information and under the entropy, entropy balanced loss function. the comparison between the estimators was made using the **MSE** and depending on all the combinations between the default values shown in table 2 by using a program Matlab-R2016a.

Table 3: MSE for the estimator of location matrix θ under entropy loss function

| Models (λ, ψ, v) | Gaussian kernel function | | Proposed kernel function | | Non-informative prior 0.0431 |
|----------------------------------|--------------------------|-------------------|--------------------------|-----------------------|---------------------------------|
| | Non-informative prior | Informative prior | Non-informative prior | Non-informative prior | |
| (3, 0.5, 2.5) | 0.0751 | 0.0661 | 0.0431 | 0.0431 | 0.1173 |
| (3, 0.5, 3.5) | 0.1029 | 0.0910 | 0.1173 | 0.1173 | 0.0173 |
| (5, 0.5, 2.5) | 0.0472 | 0.0413 | 0.0173 | 0.0173 | 0.0405 |
| (5, 0.5, 3.5) | 0.0633 | 0.0556 | 0.0405 | 0.0405 | 0.0101 |
| (7, 0.5, 2.5) | 0.0351 | 0.0306 | 0.0101 | 0.0101 | 0.0206 |
| (7, 0.5, 3.5) | 0.0463 | 0.0405 | 0.0206 | 0.0206 | Non-informative prior |
| | | | | | 0.0431 |
| | | | | | 0.1173 |
| | | | | | 0.0173 |
| | | | | | 0.0405 |
| | | | | | 0.0101 |
| | | | | | 0.0206 |

We notice from table 3 that the best estimator for θ it was at the proposed kernel function, informative prior information, and of the model $(7, 0.5, 2.5)$, this estimate is as follows:

$$\hat{\theta}_E = \begin{bmatrix} 1.8617 & 1.3913 & 5.1040 & 3.1080 & 4.0650 \\ 1.5501 & 2.8621 & 2.6503 & 1.1532 & 3.2851 \end{bmatrix}^T$$

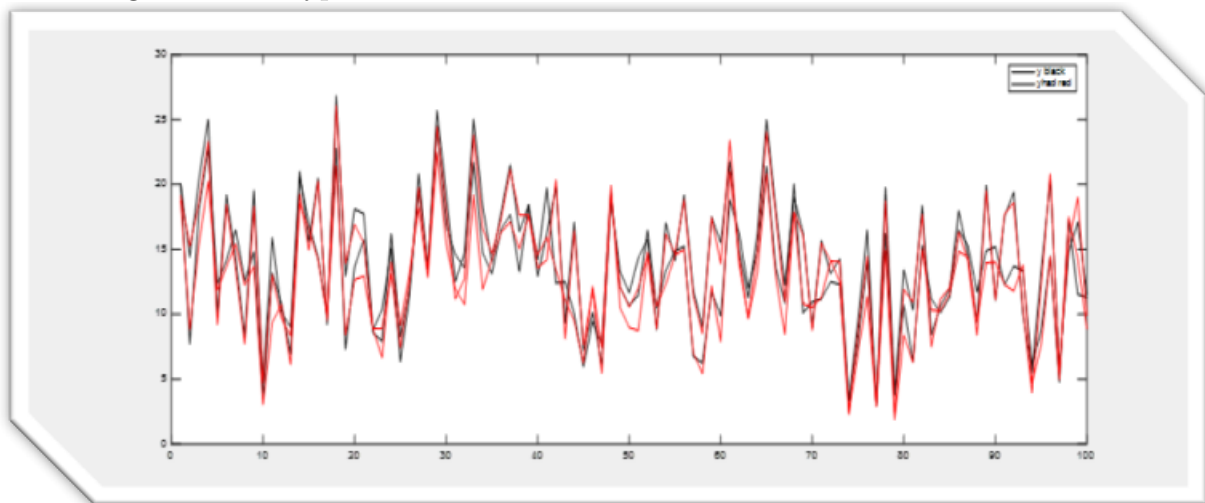
Table 4: MSE for the estimator of location matrix θ under entropy balanced loss function

| Models (λ, ψ, ν) | Gaussian kernel function | | | | Proposed kernel function | | | | Rank |
|------------------------------------|--------------------------|-------|-------------------|--------|--------------------------|-------|-------------------|--------|------|
| | Non-informative prior | | Informative prior | | Non-informative prior | | Informative prior | | |
| | W=0.1 | W=0.9 | W=0.1 | W=0.9 | W=0.1 | W=0.9 | W=0.1 | W=0.9 | |
| (3, 0.5, 2.5) | 0.0751 | | 0.0668 | 0.0739 | 0.0739 | | 0.0399 | 0.0426 | 5 |
| (3, 0.5, 3.5) | 0.1029 | | 0.0920 | 0.1015 | 0.1015 | | 0.1034 | 0.1154 | 6 |
| (5, 0.5, 2.5) | 0.0472 | | 0.0417 | 0.0464 | 0.0464 | | 0.0167 | 0.0171 | 2 |
| (5, 0.5, 3.5) | 0.0633 | | 0.0562 | 0.0623 | 0.0623 | | 0.0372 | 0.0400 | 4 |
| (7, 0.5, 2.5) | 0.0351 | | 0.0309 | 0.0344 | 0.0344 | | 0.0098 | 0.0100 | 1 |
| (7, 0.5, 3.5) | 0.0463 | | 0.0409 | 0.0455 | 0.0455 | | 0.0197 | 0.0204 | 3 |

We notice from table 4 that the best estimator for θ it was at the proposed kernel function, informative prior information, and of the model $(7, 0.5, 2.5)$, this estimate is as follows:

$$\hat{\theta}_{Eb} = \begin{bmatrix} 1.8730 & 1.3867 & 5.1065 & 3.1033 & 4.0595 \\ 1.5515 & 2.8622 & 2.6498 & 1.1511 & 3.2852 \end{bmatrix}^T$$

The following figure show the matrix of generated and estimated response variables that follows the matrix-variate generalized hyperbolic distribution based on the best estimator of θ .



7. Conclusions and future works

In this paper, a multivariate partial linear regression model is used when the error follows a matrix-variate generalized hyperbolic distribution as an alternative to the model in which the error follows a matrix normal distribution to find the Bayesian estimate of the location matrix. The posterior probability distribution of the location matrix (θ) in the case of availability of non-informative and informative prior information is a matrix-variate generalized hyperbolic distribution. The superiority of the proposed kernel function over the Gaussian kernel function for all loss functions at estimate of the location matrix. When non-informative prior information is available, the estimation for the location matrix under the entropy loss function are equal to the estimation under the entropy balanced loss function. Steadfastly the parameters ($\psi=0.5, \nu=2.5$), as the value of the parameter (λ) increases, we get the smaller value of the MSE criterion for the $\hat{\theta}$ and for all loss functions, and in the case of availability of non-informative and informative prior information. The researchers recommend conducting an application side to implement what was reached in the research, and estimate the scale matrix when the location matrix is known, depending on the kernel functions and bandwidth parameter defined in Sections (3) and (4) respectively and under different loss functions.

References

- [1] A. S. AL-Mouel and A. J. Mohaisen , *Study on Bayes Semiparametric Regression* , J. Adv. Appl. Math., 2 (4) (2017) 197-207.
- [2] O. Barndorff-Nielsen, *distribution and distributions on hyperbolae* , Scand.J. Statist., 5(1978) 151-157.
- [3] G. P. Box and G. C. Tiao , *Bayesian Inference in Statistical Analysis* , Addison Wesley publishing company, Inc. London, U.K., 1973.
- [4] T. Choi, J. Lee and A. Roy , *A note on the bayes factor in a semiparametric regression model* , J. Multivar. Anal., 100(6)(2009) 1316-1327.
- [5] P. J. Deshmays , *Bayesian estimation of generalized hyperbolic skewed student GARCH models* , Comput. Stat. Data Anal., Elsevier, 56(11) (2012) 3035-3054.
- [6] M. P.B. Gallagher and P.D. McNicholas , *Three Skewed Matrix Variate Distributions* , Stat. Probab. Lett., Elsevier, 145 (2019) 103-109.
- [7] W. Hardle, *Smoothing Technique with Implementation in S*, Spring-Verlage, New York., 1991.
- [8] M. J. Jozani, E. Marchand and A. Parzian, *Bayesian and robust Bayesian analysis under a general class of balanced loss function* , Stat. Pap., 53 (2012) 51-60, <https://doi.org/10.1007/s00362-010-0307-8>.
- [9] A. E. Koudou and C. Ley , *Characterizations of GIG Laws: A survey Complemented with Two New Result*, Proba. Surv., 11 (2014) 161-176.
- [10] N. Langrene and X. Warin, *Fast and Stable Multivariate Kernel Density Estimation by Fast Sum Updating*, J. Comput. Graphical Stat., 28 (2019) 596-608.
- [11] A. J. Lemonte and G. M. Cordeiro , *The exponentiated generalized inverse Gaussian distribution*, J. Stat. Probab. Lett., 81(4) (2011) 506-517, <https://doi.org/10.1016/j.spl.2010.12.016>.
- [12] S. J. Press , *Subjective and Objective Bayesian Statistics Principles, Models and Application*, John Wiley & Sons, Canada. , 2003.
- [13] M. Przystalski , *Estimation of the Covariance Matrix in Multivariate Partially Linear Models*, J. Multivar. Anal., 123 (2014) 380 -385.
- [14] B. W. Silverman, *Density Estimation for Statistics and Data Analysis*, Chapman and Hall, London, 1986.
- [15] L. Thabane and M. S. Haq , *On the Matrix-variate Generalized Hyperbolic Distribution and its Bayesian Applications* , J. Theor. . Appl. Stat. Sci. , 38(6) (2004) 511-526.
- [16] J. You, Y. Zhou, G. Che , *Statistical Inference for Multivariate Partially Linear Regression Models*, Can. J. Stat., 41(1)(2013) 1-22.